

21 Oct 2026 · Project 26.07.ALW · Stage 04 Live test (watch only) · Study: Backtest versus live evidence

# Turn back or run on? Watching a live gold grid basket by basket

Huyen X. Vu · Quantex Research, AlgoTrade Pte. Ltd., Singapore · ORCID 0000-0002-2209-842X

## Abstract

A grid model adds larger orders as price moves against it and closes them together as one basket. Most baskets turn back; a few run on and carry most of the risk. We ask whether a basket's size, speed and price state tell us how much risk is still ahead of it. We rebuilt every closed order of a gold (XAUUSD) grid model from 2 January to 9 October 2026 into baskets with one-minute prices; in the current setting (12 August to 9 October) this gives 2,091 cases from 788 baskets. Baskets whose largest order had reached 0.7 lots or more ran on (opened an order of at least twice that size) in 49% of cases, against 29% at 0.15 to 0.25 lots, a gap of 19 points (95% range 6 to 31). Speed and price-state measures are secondary questions. The tests were set in advance and are checked on the live account, with a public report every five minutes; the final read is in early February 2027.

**Keywords:** grid trading; averaging down; gold; XAUUSD; tail risk; live test; pre-registered tests

**JEL:** G11; G17; G32

## 1. Introduction

A grid model adds orders as price moves against it, and each order is bigger than the last. Most baskets turn back and close in profit. A few run on and carry most of the risk. This study asks what a basket's size, speed and price state tell us about the risk still ahead of it, and tests the answer on a live account, in public, before anything is changed in the model.

## Where this project stands

Project 26.07.ALW in the five Quantex stages

| Stage          | What it means here                                                                                      | Status                        |
|----------------|---------------------------------------------------------------------------------------------------------|-------------------------------|
| 01 Review      | Question and data                                                                                       | Done                          |
| 02 Backtest    | Past baskets, 2 Jan–9 Oct 2026, rebuilt from closed orders and one-minute prices                        | Done                          |
| 03 Fine-tune   | Not used: the model itself is not changed in this study                                                 | Skipped                       |
| 04 Live test   | Live watch of the running account, with tests set in advance. Watch only: the account trades as before. | <b>Now</b> , from 21 Oct 2026 |
| 05 Handed over | Only if a finding is confirmed and a later study shows it is worth using                                | Not started                   |

## Research questions

- Size.** Once the largest order in a basket has reached x lots, how often does it run on, and how much loss is still ahead of it?
- Speed.** Does the time a basket takes to grow from its first order to x lots tell us more?
- Price state.** At that moment, do simple measures (RSI on 15-minute bars, where price sits in the day's range, the size of the last 15 minutes' move) tell us more?

All questions are answered separately for buy and sell baskets. This study measures association, not cause, and does not test whether acting on it would make money (see Scope).

## 2. Data and method

### The model we watch

Altair, a gold (XAUUSD) grid model shown on the AlgoTrade live board. It runs on an MT4 demo account (balance about \$27 million) and reports every minute through AlgoTrade Connect. When price moves against it, it adds orders on the same side, each one larger than the one before, and closes them together. We call one such group a **basket**. As with all our models, we do not publish its rules.

### Definitions

- **Case:** a basket at the moment its largest order first reaches  $x$  lots ( $x = 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.7, 1.0, 1.5$ ). One basket gives one case at each level it reaches, so a deep basket gives several cases.  $x$  is the size of the largest single order, not the total open lots. In the current setting the total open lots rise with  $x$  in a fixed way, so  $x$  also ranks baskets by their total size. We use  $x$  as a measure of how deep the grid has gone, not of the account's full exposure.
- **Run on:** after that moment, the basket opens an order of at least  $2x$  lots. The rate at level  $x = \text{run-on cases} \div \text{all cases at level } x$ .
- **Worst floating loss after  $x$ :** the lowest open profit/loss of the basket between that moment and its close, from one-minute high and low prices. A **tail event** is a worst floating loss beyond 0.05% of the balance (about \$13,500) or beyond 0.2% (about \$54,000).
- **Speed:** minutes from the first order to the first order of  $x$  lots, split into fastest, middle and slowest thirds at each level.
- **95% range:** from resampling whole baskets (and, as a check, whole days) 2,000 times, so that cases of one basket, or of one market day, are not treated as independent.

### Data

Every closed Altair order from 2 Jan to 9 Oct 2026, rebuilt into baskets with one-minute gold prices. On 12 Aug 2026 the model's order-size settings changed, so we report the two settings separately and use only the current one, which runs live, for the live tests:

- **Old setting**, 2 Jan–12 Aug: 4,321 cases from 1,142 baskets.
- **Current setting**, 12 Aug–9 Oct (8 weeks): 2,091 cases from 788 baskets.

**Rebuild quality.** We could place every order of a basket in time for 99.7% of baskets from April to 9 October, but only 60–63% of baskets (counted as baskets, not orders) in January to March, and the baskets we missed then were larger than average. In the current setting, 6 baskets could not be rebuilt, 4 of them among the 59 that reached 0.7 lots. They are left out of the results: we report the rebuilt sample as it is and do not assume the missing baskets behaved like the rest. In the worst case for the size result (all their deep cases turned back, all their shallow cases ran on), the gap in Result 1 would be +14.5 points (95% range +2.7 to +26.3) instead of +19.2. In a replay of 25 Aug–9 Oct 2026 through the live code, speed class and outcome matched the backtest on all 1,664 cases found by both.

### Checks we ran

- Results split by model setting; fast and slow cut-offs set within each setting.
- Every 95% range resamples whole baskets; resampling whole days gives the same conclusions (current setting: size +19.2, range +8.2 to +30.7; speed +0.1, −7.2 to +7.1).

- Speed was one of 24 measures we looked at. The backtest is therefore exploratory; the live watch below is the test.
- Unfinished baskets: the order history runs to 21 Oct, twelve days past the price data. Among closed baskets the longest took 73 hours, and on 21 Oct the only open basket had opened that day, so no basket from the backtest window was left out for being still open.

### What we test live (set 21 Oct 2026, before any live basket reached 0.15 lots)

| Test                                                                                                                            | Unit                          | Backtest, current setting | Confirmed if                                                      |
|---------------------------------------------------------------------------------------------------------------------------------|-------------------------------|---------------------------|-------------------------------------------------------------------|
| <b>H1, main:</b> run-on rate at 0.7 lots and above is higher than at 0.15–0.25 lots                                             | cases, resampled by basket    | 49% vs 29% (+19)          | gap at least +10 points and 95% range above 0                     |
| <b>H2:</b> slow sell baskets run on more than fast ones                                                                         | cases, resampled by basket    | +7.7                      | gap at least +8 points and 95% range above 0                      |
| <b>H3:</b> no speed gap for buy, or for all baskets                                                                             | cases, resampled by basket    | −6.3 / +0.1               | 95% range inside −8 to +8 points                                  |
| <b>H4, exploratory:</b> share of tail events (0.05% and 0.2%) rises with size; price at the day's extreme at 1.0 lots and above | one case per basket per level | Figure 2                  | reported with a false-discovery-rate control (Benjamini–Hochberg) |

**Sample and timing.** At the backtest rate the account gives about 13 deep cases a week from 7 baskets, and about 40 sell cases a week in each speed group; because one basket gives several cases, the effective sample is about half the number of cases (design effect 1.6 to 2.1). The final read is at 15 weeks (early February 2027). By simulation from the backtest, that gives about 95% power for H1 if the true gap is as large as in the backtest, but only about 35% if it is 10 points; H2 has about 50% power for an 8-point gap. The live sample is sized to find a large effect, not to rule out a smaller one: a result that misses its threshold will be reported as not confirmed at this sample size, not as no effect. The two size groups in H1 are compared as conditional case rates, not as independent groups of baskets; their 95% range resamples whole baskets, so the levels one basket reaches stay together. The power figures are planning estimates: they assume the backtest rates and resample baskets within each group separately, so they are not a promise of confirmation. The  $\pm 8$ -point margin in H3 is a practical threshold we set on 21 Oct 2026, before any live data; meeting it would not show that the true effect is zero. 18 Nov 2026 is a description only, with no test and no stopping. All results, including negative ones, will be published.

**Other analyses.** Checks of volatility, trading session or trend, and any early-warning model, are exploratory unless their tests are written down here before the outcomes are seen. An exploratory finding needs its own later test before we treat it as confirmed.

### How the live watch works

- The account sends its open orders every minute (open time, open price, size, side, profit). Our server does all the work; nothing runs on the trading terminal.
- The server rebuilds each basket, notes the exact time each lot level was first reached, classes it as fast, middle or slow, records the price measures and the basket's floating loss.
- Price comes from the open orders themselves (open price plus profit per lot). Measures are taken only when the minute report arrived within 3 minutes of the order. Floating loss is read once a minute, so live tail losses can be slightly smaller than one-minute highs and lows would show.
- Since 21 Oct 2026 a read-only recorder on the trading terminal also sends an account snapshot every 5 seconds. The strip at the top of the live report shows the latest one and checks for a new one every 10 seconds.

- Open baskets are shown and kept as unfinished cases in the final read. The p-values in the live tables treat cases as independent; the final read resamples whole baskets.

### 3. Results

#### What the backtest found

##### 1. A basket with a larger order is more likely to run on

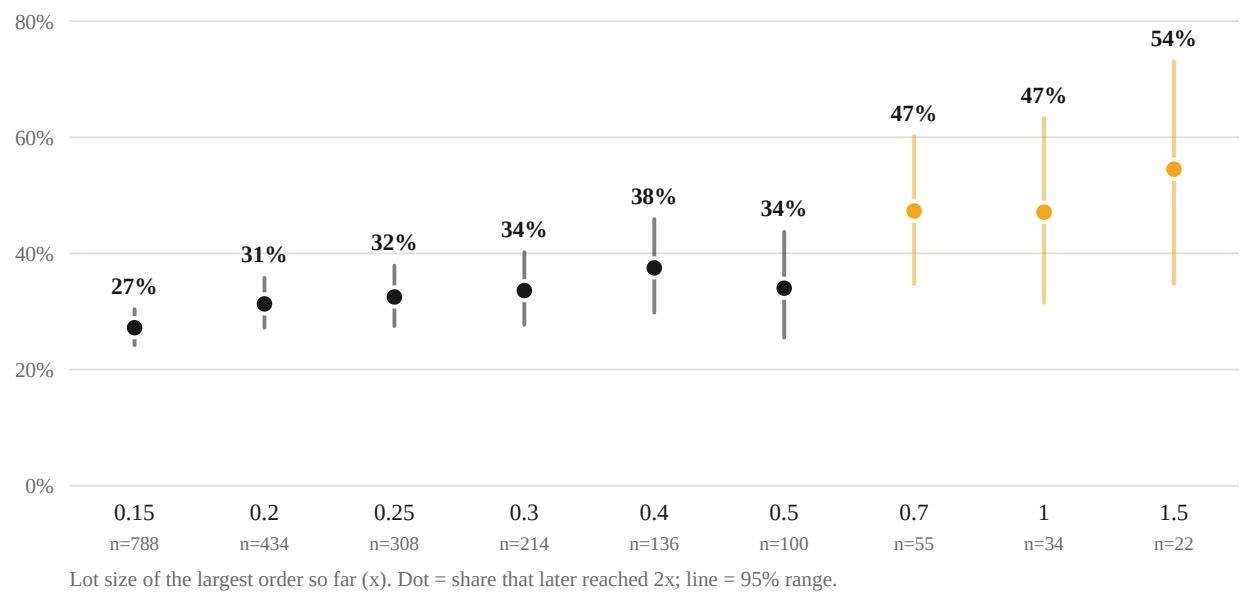


Figure 1. Current setting, 12 Aug–9 Oct 2026, buy and sell together. Of the cases at 0.15–0.25 lots, 29% ran on; at 0.7 lots and above, 49%.

The gap is +19 points (95% range +6 to +31 by baskets, +8 to +31 by days). This is a conditional rate: only baskets that have already grown reach the deep levels, so it describes the risk still ahead of a basket that has got there, not the cause of that risk. The deep end rests on 111 cases from 55 baskets.

##### 2. The loss still ahead grows faster than the size

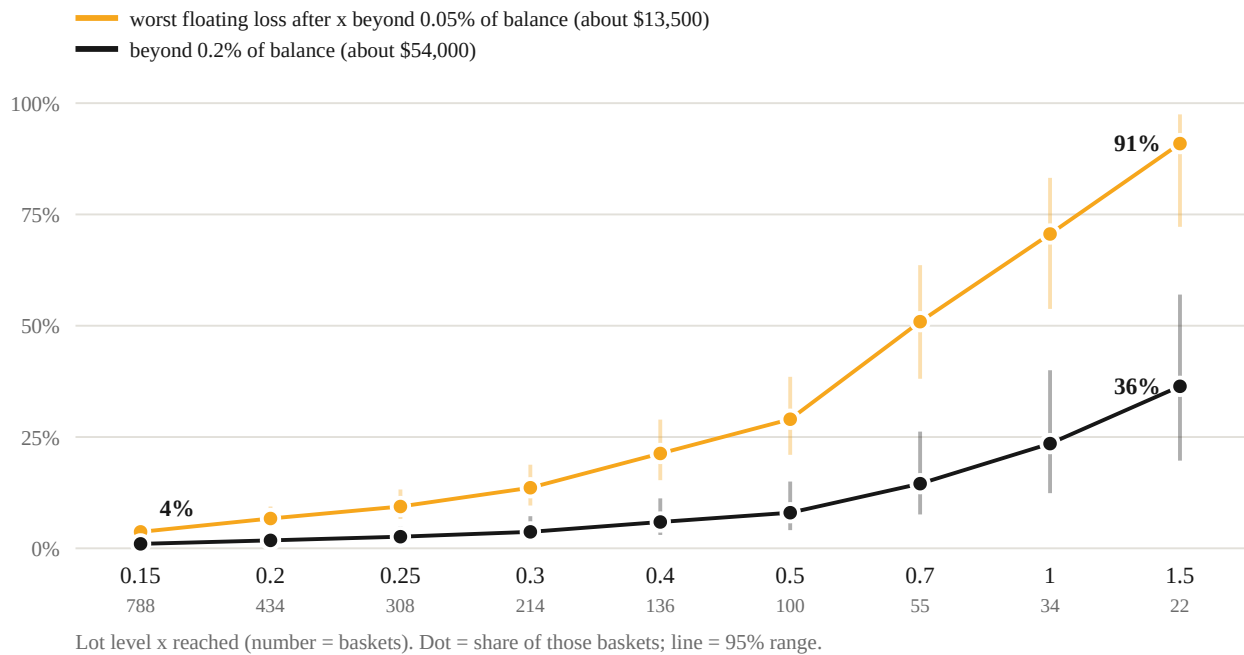
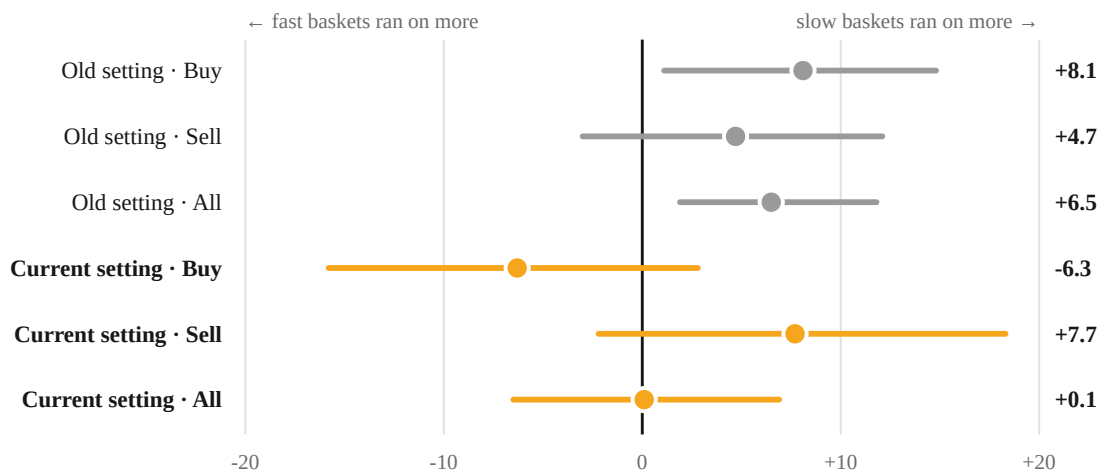


Figure 2. Current setting. From the moment a basket reached x lots, the share whose worst floating loss afterwards went past 0.05% or 0.2% of the balance. One case per basket at each level.

At 0.15 lots, 4% of baskets later went past 0.05% of the balance; at 0.7 lots, 51%; at 1.5 lots, 91%. For the larger threshold (0.2%) the shares are 1%, 15% and 36%. The average of the worst 5% of these losses grows from 0.4% of the balance at 0.15 lots to 3.9% at 0.7 lots and 9.1% at 1.5 lots. Almost all baskets still closed in profit (99.5%). These are losses of single baskets on a very large demo balance; they are not the account's drawdown, which also depends on baskets open at the same time.

### 3. Speed: a signal in the old setting, none overall in the current one



Slowest third minus fastest third, in points. Line: 95% range, whole baskets resampled.

Figure 3. Run-on rate of the slowest third minus the fastest third. Thirds are set within each setting and level.

In the old setting, slow baskets ran on more often: +6.5 points (95% range +1.9 to +11.8). In the current setting the overall gap is +0.1 (-6.5 to +6.9). Sell baskets lean the same way as before (+7.7, -2.2 to +18.3) and buy baskets the

other way (-6.3, -15.8 to +2.8). Neither is clear. Speed may also stand in for volatility, session or trend; we test that on the final data.

#### 4. Price state and a combined model: weak

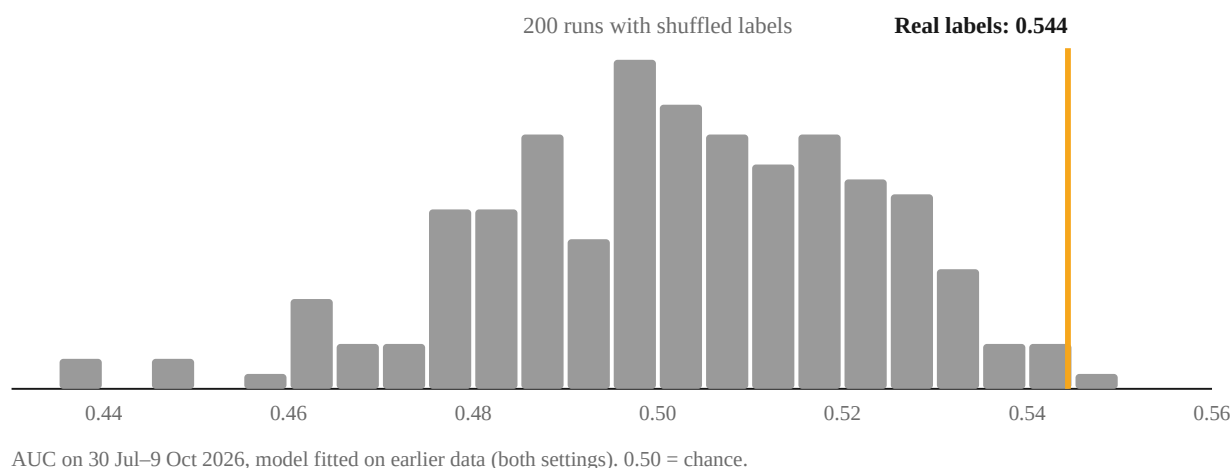


Figure 4. A model using 24 measures, fitted on earlier data and checked on 30 Jul–9 Oct 2026, against the same model trained on shuffled outcomes 200 times.

The model ranked cases only a little better than chance (AUC 0.544; one of 200 shuffled runs was as high). The shuffles moved single cases, not whole days, and the check period spans both settings, so even this is optimistic. At 1.0 lots and above, price at the day's extreme against the basket went with more run-ons in both settings, on small samples (current setting: 61% against 39%, 28 cases each).

#### Live report

Updated from the running account. How the watch works, the tests set in advance and the backtest are further down.

The live report on the web page updates every five minutes: <https://quantex.com.sg/research/altair-basket-watch/>. This PDF is a snapshot taken on 24 October 2026.

#### Interim reports

**Week 1, data to 23 Oct 2026, 20:56 UTC.** Three trading days (21 to 23 Oct): 419 closed baskets, 51 reached 0.15 lots, 124 level cases. This is an early read, not a result: the first look is on 18 Nov 2026 and the final read around early February 2027. The live tables above keep updating; the numbers below are frozen at the cut-off.

- **H1, depth (main test).** Share that ran on, live against the backtest: 0.15-0.25 lots 28.3% (26/92) against 29.4%; 0.3-0.5 lots 34.6% (9/26) against 34.9%; 0.7 lots and above 16.7% (1/6) against 48.6%. The two lower levels are close to the backtest. The deep level rests on 6 cases from about 3 baskets, so the live gap of -11.6 points (backtest +19.2) tells us nothing yet.
- **H2 and H3, speed (exploratory).** Sell: slow minus fast +3.1 points (6/26 against 1/5), p 0.44. Buy: +43.8 points (14/32 against 0/6), one-sided p 0.021, the opposite direction to the backtest (fast 34.7%, slow 28.4%). These p-values treat cases as independent, although cases from one basket are linked, and several tests are run, so they are not evidence. The final read resamples whole baskets.
- **H4, price state.** 3 cases at 1.0 lots and above; nothing to read.
- **How bad baskets got (descriptive).** The 3 closed baskets whose largest order reached 0.7 lots had a median worst floating loss of \$4,051 (backtest median \$14,137). With 3 baskets this does not show that live risk is lower. No

basket's worst floating loss after any level went past 0.2% of the balance. Largest order so far: 1.74 lots.

## Findings log

- **21 Oct 2026.** Backtest finished; tests set; live watch started.
- **24 Oct 2026.** Interim report, week 1 (see Interim reports). Too early for any reading.

## Result

Pending. First look on 18 Nov 2026, final read around early February 2027.

The final read will keep three levels apart. If only H1 is confirmed, it will show that, in this setting, baskets whose largest order has grown further run on more often. It will not show that speed adds anything, and it will not show that acting on either would improve the account's profit or drawdown.

## 4. Limitations

### Scope: what this study cannot tell

- It measures association at the moment a level is reached. An early-warning model that uses only what is known at 0.15 or 0.25 lots, and predicts a defined loss within a fixed time, is a separate analysis on the final data.
- It does not show that acting on a signal (an alert, no more orders, a loss limit) would improve the account. That needs its own study: a simulation that counts the profit given up by early action, then a shadow run next to the model, before any change to a live account.
- It does not measure account drawdown, margin or results on a smaller account.
- Backtest floating loss uses one-minute highs and lows; live floating loss is read once a minute from the account. Backtest and live tail rates are therefore not directly comparable.

### Limits of this study

- One model, one broker, a demo account, and ten months of history that included unusually large gold moves. The current setting has eight weeks of history.
- Price measures use another broker's one-minute prices; "run on" at 2x and the two loss thresholds are our choices.
- Baskets that open and close within one minute are not seen by the live watch; they rarely reach 0.15 lots.
- Past results, real or simulated, do not guarantee future results.

## Declarations

- **Author:** Huyen X. Vu, Quantex Research, AlgoTrade Pte. Ltd., Singapore (ORCID 0000-0002-2209-842X).
- **Data availability:** Live results are published in aggregate on this page and updated every five minutes. Order-level data can be shared with reviewers under a confidentiality agreement. Model rules are not published.
- **Use of AI tools:** Data processing, analysis code and drafting were assisted by AI tools (Claude, Anthropic). The author checked the results and is responsible for the content.
- **Conflict of interest:** The author is the founder of AlgoTrade Pte. Ltd., which develops and runs systematic trading models, including the model studied here.
- **Funding:** No external funding.
- **Peer review:** Not yet peer reviewed. Comments are welcome at [info@quantex.com.sg](mailto:info@quantex.com.sg).
- **How to cite:** Vu, H. X. (2026). Turn back or run on? Watching a live gold grid basket by basket. Quantex Research, AlgoTrade Pte. Ltd. <https://quantex.com.sg/research/altair-basket-watch/>

For education and information only. Not investment advice. See our Disclaimer.