

Supplementary Methods and Reproduction Guide

S1 Scope and relationship to the main article

This supplement accompanies the English InventoryFusion manuscript by Jiexiao Zhang. It documents completed local experiments dated September 26, 2026. The workbook and reproduction archive retain weaker methods, ablations, and sensitivity results alongside the principal comparisons. The experiment is a forecast-output ensemble evaluation, not neural-network fine-tuning or neural-parameter merging.

S2 Data identities and selection

FreshRetailNet-50K is pinned to dataset revision 08c1fab7f9257bc73679d415d65d644165d351d4 and covers March 28 to July 2, 2024. The released normalized daily sale_amount is used without converting it into physical units. A positive stock_hour6_22_cnt flags a day with at least one recorded stockout interval. UCI Online Retail covers December 1, 2010 to December 9, 2011. Its source XLSX SHA256 is 43465a06f2ccf7c8b5bd2892bc7defb52f97487934fe93b16ae4c3936424676d. Filtering and daily aggregation retain positive, noncancelled UK sales. The all-false UCI stockout array is an interface placeholder, not evidence of availability.

Sampling uses only initial histories of 41 and 200 days. Eligible series have finite nonnegative histories and at least eight nonzero days. Four ranked zero-rate strata each contribute 10 series to each of three disjoint cohorts selected with SHA256 ordering and seeds 17, 29, and 43. The 120 series in each dataset jointly fit global and segment estimates; cohorts are not independently trained replications.

S3 Frozen forecasting and analysis protocols

The horizon is 14 days with a 7-day stride. FreshRetail tuning origins are 41, 48, and 55, followed by test origins 69, 76, and 83. UCI tuning origins are 248, 255, 262, 269, 276, and 283, followed by origins 297 through 360 at 7-day intervals. Origins are zero-based. Forecast tensors use expert, series, origin, horizon, and quantile axes, with quantiles 0.1, 0.5, and 0.9. Histories stop strictly before the forecast origin.

The seven experts are seasonal_naive_7, moving_average_28, croston_sba, calendar_ridge, ttm_r2_1_daily, chronos_2, and timesfm_2_5. Seven static combinations and one online combination yield 15 methods. The primary hierarchical weights use shrinkage strengths 8 and 4, temperature 7, and maximum expert weight 0.88. Online updates use rate 0.25 and process each fully matured 14-day window only once.

Bootstrap intervals resample 120 complete series trajectories with replacement, paired across methods, for 2,000 repetitions using seed 20260926. Intervals are 95% percentile intervals and are unadjusted for multiplicity. The method preserves overlapping windows within a series but does not independently resample store-level or calendar-level shocks.

S4 Inventory protocol and sensitivities

Inventory simulation has a 7-day review interval, a 7-day lead time, continuous quantities, lost sales, and common initial stock equal to seven times the preceding 28-day mean. Initial pipeline inventory is zero. Events are arrivals, review and ordering, sales fulfillment, and end-of-day holding. State carries between cycles. The simulation spans 21 days on FreshRetail and 70 days on UCI.

The main order-up-to target adds an empirical 90th-percentile cumulative residual correction to the summed median forecast. It is compared with the sum of daily P90 forecasts. The main correction shares tuning labels with weight fitting and is not a proved conformal guarantee. The strict sensitivity uses FreshRetail origin 41 for weights and 55

for calibration, and the first four UCI tuning origins for weights and 283 for calibration, without overlapping target dates.

Holding rates are 0.005, 0.02, and 0.05, with a unit lost-sales penalty of 1. These reprice the same trajectories rather than optimize separate policies. Terminal on-hand and pipeline inventory are retained without salvage or disposal costs. Purchasing, MOQ, budgets, capacity, transport, perishability, substitution, and latent stockout demand are outside the simulation.

S5 Supplementary file dictionary

Experimental_Results.xlsx contains 14 English worksheets. Overview links headline differences to the underlying result worksheets through formulas. Forecast metrics includes overall and sensitivity slices. Forecast bootstrap and Inventory bootstrap contain paired intervals. Inventory simulation reports cost, fill rate, holdings, lost sales, and terminal states. Sampling cohorts, Demand segments, Forecast origins, and Per SKU FRN/UCI provide disaggregated results. Separate calibration forecasts/inventory and Cumulative calibration retain the alternative calibration analysis. Parameters and sources records the frozen configuration and provenance.

Table_1.csv through Table_7.csv and Table_A1.csv match the displayed manuscript tables. They are rounded presentation tables, whereas the reproduction CSVs retain full numerical precision. The workbook is an inspection aid; Python scripts, not spreadsheet formulas, perform the bootstrap and inventory experiments.

The reproduction archive includes fixed data subsets, raw expert prediction caches, evaluated prediction arrays, data identities, protocols, environment information, and detailed analysis CSVs. Raw large download files, neural-network weights, virtual environments, dependency caches, credentials, and page-preview images are excluded.

S6 Reproduction levels

Level 1 recomputes analysis from the included fixed inputs and expert forecasts and requires no GPU or network. Level 2 regenerates the figures from existing analysis files. Level 3 downloads pinned models and source data and repeats inference in a new results directory. Exact commands and cache precautions are provided in REPRODUCE.md inside the archive.

Numerical reproducibility and page-layout reproducibility are separate. The English plots and manuscript use English labels and unchanged numerical data. The workbook builder requires the author's artifact-tool environment; full CSV outputs remain usable without it. Software versions, source hashes, and the original experiment's script hashes are retained so presentation changes can be distinguished from a new experiment.

S7 Source and license attribution

FreshRetailNet-50K: Wang, Y., Gu, J., Long, L., et al. (2025). FreshRetailNet-50K: A Stockout-Annotated Censored Demand Dataset for Latent Demand Recovery and Forecasting in Fresh Retail. arXiv:2505.16319. Dataset publisher: Dingdong-Inc. <https://huggingface.co/datasets/Dingdong-Inc/FreshRetailNet-50K>

UCI Online Retail: Chen, D. (2015). Online Retail [Dataset]. UCI Machine Learning Repository. <https://doi.org/10.24432/C5BW33>

Both datasets retain CC BY 4.0 licensing: <https://creativecommons.org/licenses/by/4.0/>. The subsets are filtered, aggregated where applicable, and deterministically sampled derivatives; no latent-demand imputation is performed. Project code retains its Apache-2.0 license. Third-party licenses and attributions are retained separately.