# Analyzing Manufacturing Data
# with Time-Series Foundation Models

**Sangwoo Mo**
POSTECH · Industrial & Management Engineering

TODAY

01 Manufacturing time series

02 Time-series foundation models

03 Making TSFMs interpretable

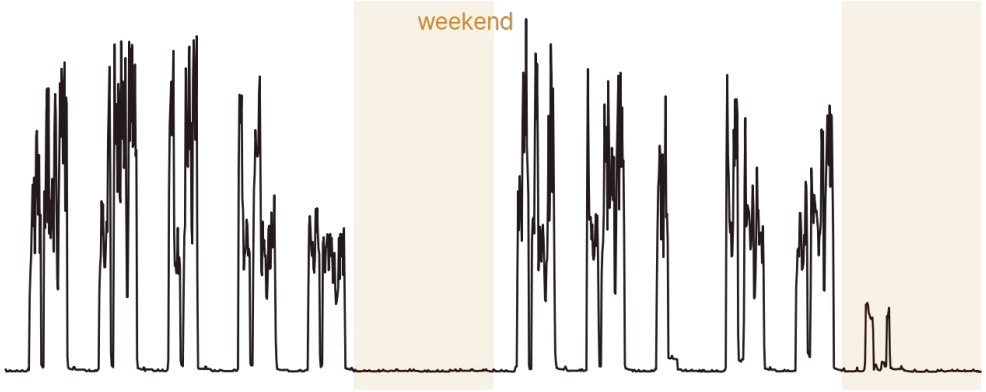
OUR WORK

# 1 Manufacturing time series

MANUFACTURING TIME SERIES

# Manufacturing runs on time

A modern line is instrumented end to end: nearly every signal, from sensor to shipment, is a time series. And the process never sits still: tools wear, materials drift, demand shifts.

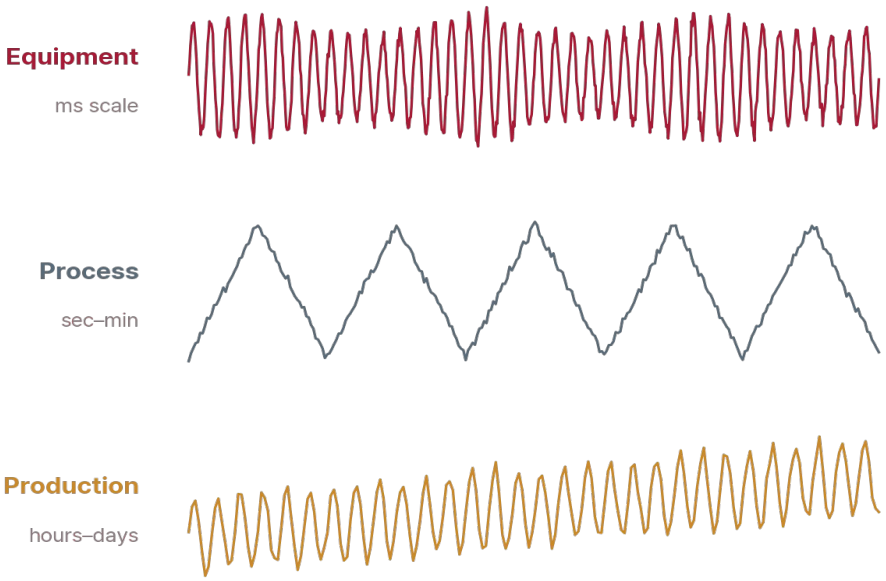



**A steel plant's power meter**
two weeks, 15-min readings

DAEWOO Steel, Gwangyang (UCI, 2018).

MANUFACTURING TIME SERIES

# One factory, many time scales




## Equipment

vibration, motor current, temperature, pressure

## Process

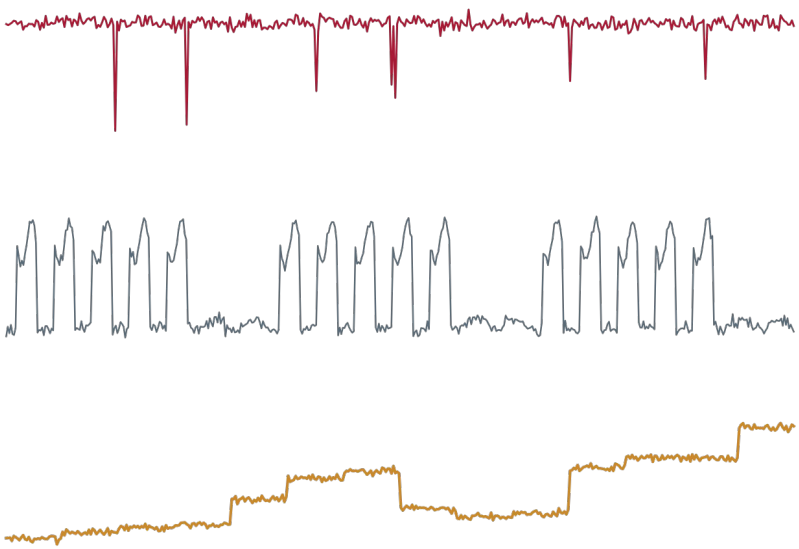
recipe & tool traces, in-line metrology

## Production

throughput, yield, energy, demand

MANUFACTURING TIME SERIES

# Three plant signals, three behaviors



## Regulation

a value the plant deliberately holds steady (here, acidity / pH). It should stay flat, so what matters is each sudden dip. (15 min)

## Rhythm

a value that rises and falls on a schedule (here, plant power use): it repeats with the workday, near zero on weekends. (15 min)

## Drift

a value with no fixed pattern (here, product color from lab tests). Its level slowly wanders between runs. (daily)

MANUFACTURING TIME SERIES

# From signal to decision

**Forecasting**
plan, stock, schedule

**Anomaly detection**
catch faults & excursions

**Diagnosis**
find the root cause

**Prognostics (RUL)**
time the maintenance

**Monitoring / control**
keep the process in spec

**Virtual metrology**
measure the unmeasured

MANUFACTURING TIME SERIES

# Today's practice: one model per process



MANUFACTURING TIME SERIES

# Today's practice: one model per process




↻ **and again, for every line, product, and changeover**

**4–5 years · KRW 6–10B** one bespoke plant model, on average

Averages across Korea's AI-Factory program projects (2025 program review).

# 2 Time-series foundation models

# Time-series foundation models (TSFM)

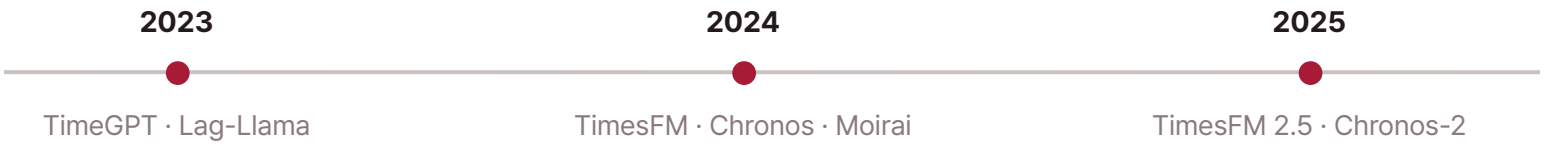
**Task-specific era**
ARIMA · Prophet · DeepAR
one model per series, hand-tuned
and fit to that series' own history

new process → **collect, refit, re-tune**

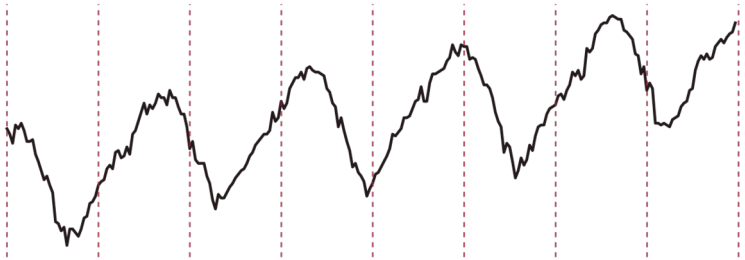
→

**Foundation era**
one transformer, pretrained once
on billions of points from many
domains, frequencies, and scales

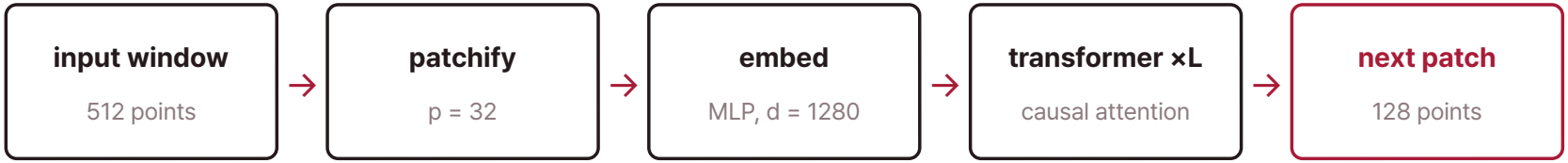
new process → **forecast zero-shot**



# Inside a TSFM: patch, attend, predict



**Patching.** The series is sliced into fixed windows (32 points each). Each patch is embedded as one token, like a word.



# Encoder vs decoder

## Encoder

masked · bidirectional

reconstructs masked spans;
sees context in both directions

USED FOR

**representation · imputation**
**classification · anomaly detection**

*e.g. Moirai*

## Decoder

causal · autoregressive

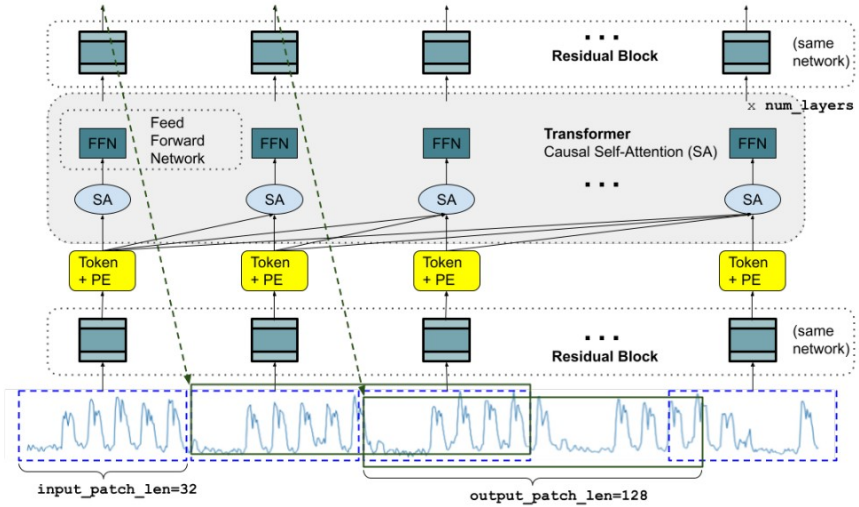
predicts the next patch;
generates future by rollout

USED FOR

**forecasting,**
**to any horizon**

*e.g. TimesFM · Chronos*

# TimesFM (Google): patch in, patch out



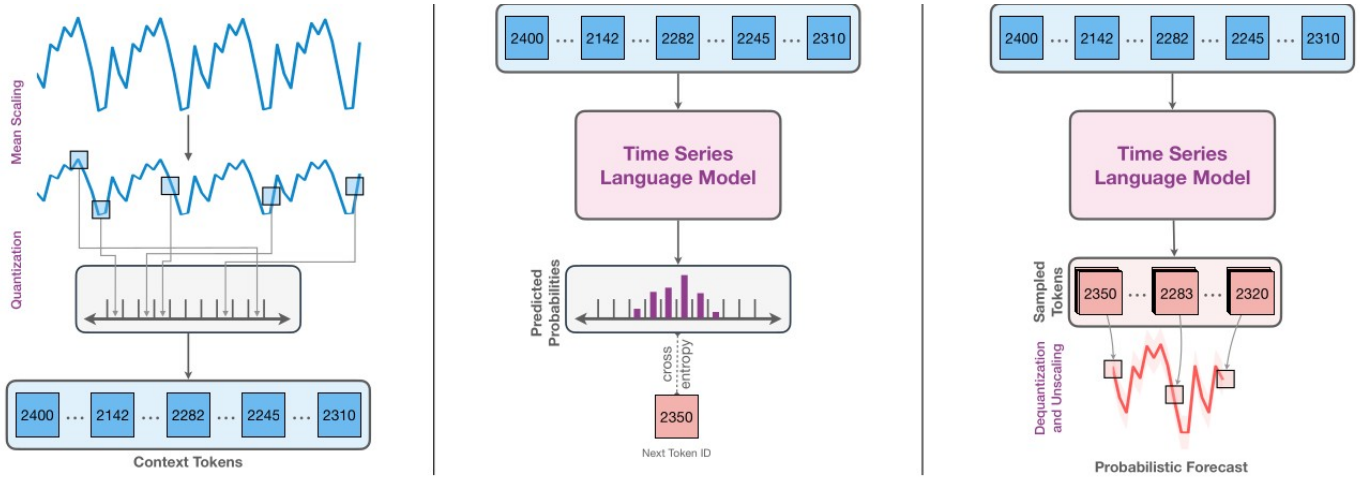

**decoder-only, causal**
patch 32 → 128

**MSE next-patch pretraining**
200M params
~100B points

**v2.5: context 512 → 16k**
point + optional quantile head

Das et al., ICML 2024. "A decoder-only foundation model for time-series forecasting."

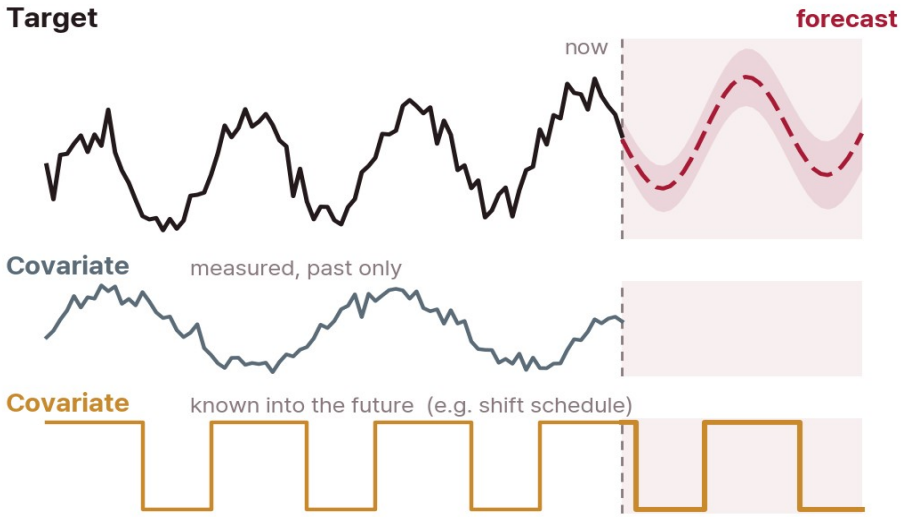
# Chronos (Amazon): a series as a language




scale → quantize (4096 bins) → LM → cross-entropy;  decode: sample token paths → dequantize → predictive distribution.

Ansari et al., TMLR 2024. "Chronos: Learning the Language of Time Series."

# Chronos-2 (Amazon): forecasting with covariates



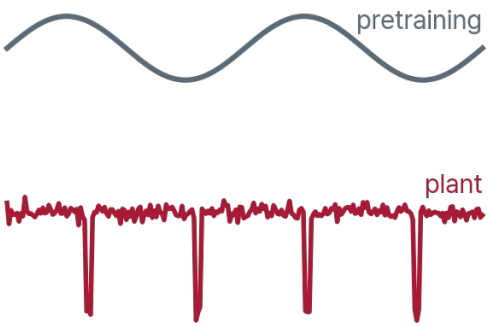

## Condition on covariates you already know

**The forecast reflects what you plan to do, not just past values.**

**Also new** multivariate · group attention · still zero-shot

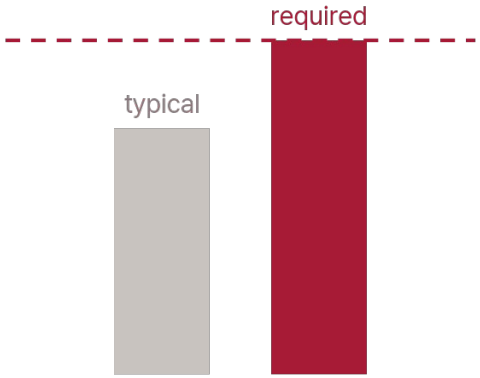
Ansari et al., 2025. Chronos-2, ~120M params.

# Powerful, but three catches in manufacturing




## Domain gap

plant signals are unlike the pretraining corpora




## Reliability bar

audits & safety demand high, explainable accuracy



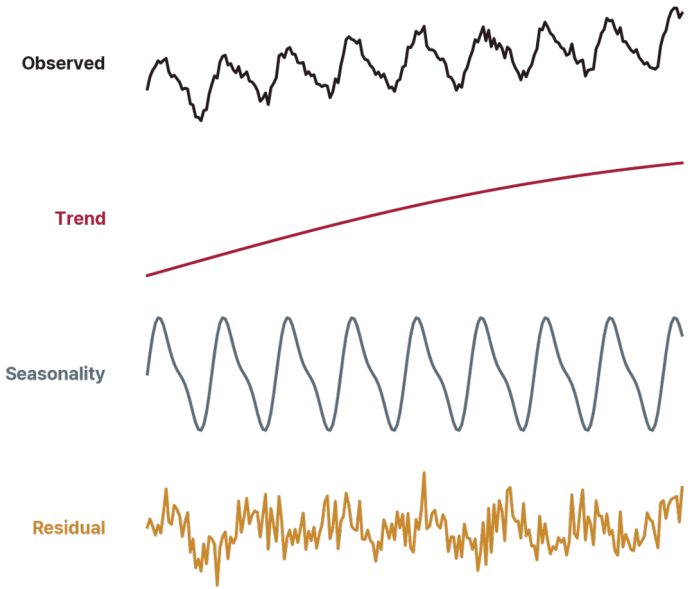

## Black box

values come out; reasons do not

# 3 Making TSFMs interpretable

OUR METHOD

# Lessons from statisticians




$$y = T + S + R$$

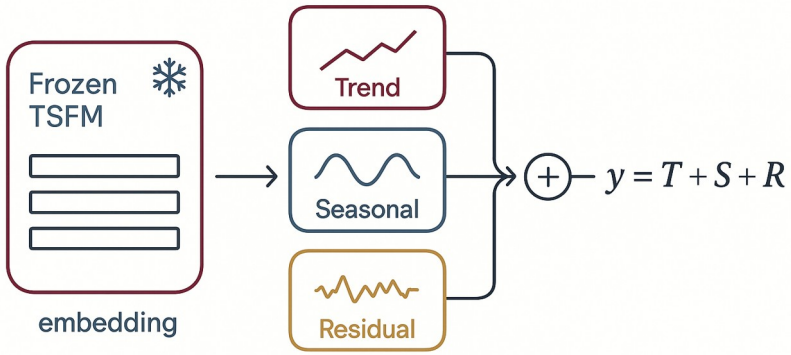
Classical decomposition
Holt–Winters 1960
STL 1990
X-13ARIMA-SEATS 1998

Prophet 2017

OUR METHOD

# Our idea: predict bases, not values

Decode a frozen TSFM into trend, seasonal, and residual, instead of forecasting raw values.



OUR METHOD

# The embedding already carries the structure

## CLASSICAL

**raw series**

just values, no structure yet

↓

compute the split from scratch,
for every new series

**Trend · Seasonal · Residual**

## OURS

**Frozen TSFM embedding**

already encodes **trend** · **seasonal** · **residual**

↓

**just read it out**

**Trend · Seasonal · Residual**

**The structure is already inside the embedding. We reuse it, not re-derive it.**

OUR METHOD

# In-context basis prediction

**In language:** a frozen LLM reads the examples in its prompt and answers, with no fine-tuning.




**In time series:** the frozen TSFM does the same, and what we read out are its bases.



OUR METHOD

# Three bases, one frozen embedding

**Trend**

$$\hat{T}(t) = (\delta_0 + \Gamma(t)^\top \delta) \cdot t + (\rho_0 + \Gamma(t)^\top \rho)$$

piecewise-linear · tuned change points, sparse penalty

**Seasonal**

$$\hat{S}_p(t) = g_p \sum_{j=1}^{k_p} \left[ a_j \cos \frac{2\pi j t}{p} + b_j \sin \frac{2\pi j t}{p} \right], \; p \in \{d, w, m, y\}$$

gated Fourier · orders 10/4/2/8, sigmoid gate

**Residual**

$$\hat{R}(t) = \mathrm{MLP}(\mathbf{e})$$

2-layer MLP on the frozen embedding

---

**Compose**

$$\hat{y} = \hat{T} + \sum_p \hat{S}_p + \hat{R}$$

the forecast

# Training: a staged curriculum

**1 · Synthetic warm-up**
learn the split where components are known

→

**2 · Mixed 9 : 1**
add real LOTSA batches; synthetic keeps the split honest

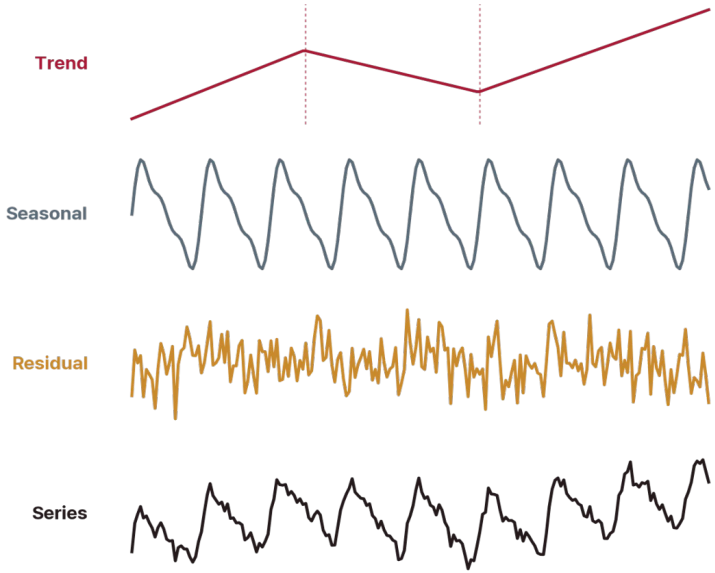
→

**3 · Residual-only**
freeze trend and seasonal, fit what remains

**Loss**

- prediction error
- orthogonality: decorrelate T, S, R
- sparsity: change points & gates

# Synthetic data: components known by construction




**Trend** piecewise-linear basis
**Seasonal** Fourier, only physically valid harmonics
**Residual** Gaussian · t · gamma · Pareto · Weibull

| freq | fd (days/step) | 512-ctx span |
|---|---|---|
| 5-min | 1/288 | 1.78 d |
| 10-min | 1/144 | 3.56 d |
| 30-min | 1/48 | 10.67 d |
| hourly | 1/24 | 21.33 d |
| daily | 1.0 | 512 d |
| weekly | 7.0 | 3584 d |
| monthly | 30.44 | 15584 d |

Valid harmonics per granularity.

# Real data: LOTSA

27B observations across 9 domains; energy, transport and climate dominate, and most series are hourly or finer.

*Table 2.* Key statistics of LOTSA by domain.

| | Energy | Transport | Climate | CloudOps | Web | Sales | Nature | Econ/Fin | Healthcare |
|---|---|---|---|---|---|---|---|---|---|
| # Datasets | 30 | 23 | 6 | 3 | 3 | 6 | 5 | 23 | 6 |
| # Obs. | 16,358,600,896 | 4,900,453,419 | 4,188,011,890 | 1,518,268,292 | 428,082,373 | 197,984,339 | 28,547,647 | 24,919,596 | 1,594,281 |
| % | 59.17% | 17.73% | 15.15% | 5.49% | 1.55% | 0.72% | 0.09% | 0.10% | 0.01% |

*Table 3.* Key statistics of LOTSA by frequency.

| | Yearly | Quarterly | Monthly | Weekly | Daily | (Multi) Hourly | (Multi) Minute-level | (Multi) Second-level |
|---|---|---|---|---|---|---|---|---|
| # Datasets | 4 | 5 | 10 | 7 | 21 | 31 | 25 | 2 |
| # Obs. | 873,297 | 2,312,027 | 11,040,648 | 18,481,871 | 709,017,118 | 19,875,993,973 | 7,013,949,430 | 14,794,369 |
| % | 0.003% | 0.008% | 0.040% | 0.067% | 2.565% | 71.893% | 25.370% | 0.054% |

Evaluation: public benchmarks, 512 context → 96 horizon, MAE vs zero-shot backbone. · Woo et al., ICML 2024 (Moirai).

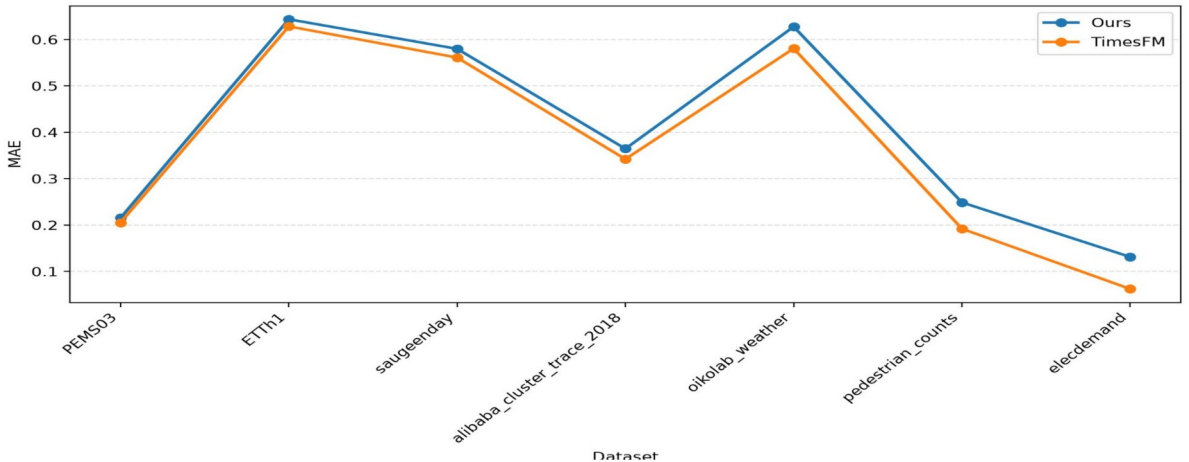
RESULTS · ACCURACY

# On par on accuracy, ahead on interpretation

**≈ 0.04**

**average MAE gap to TimesFM v2.5 zero-shot**

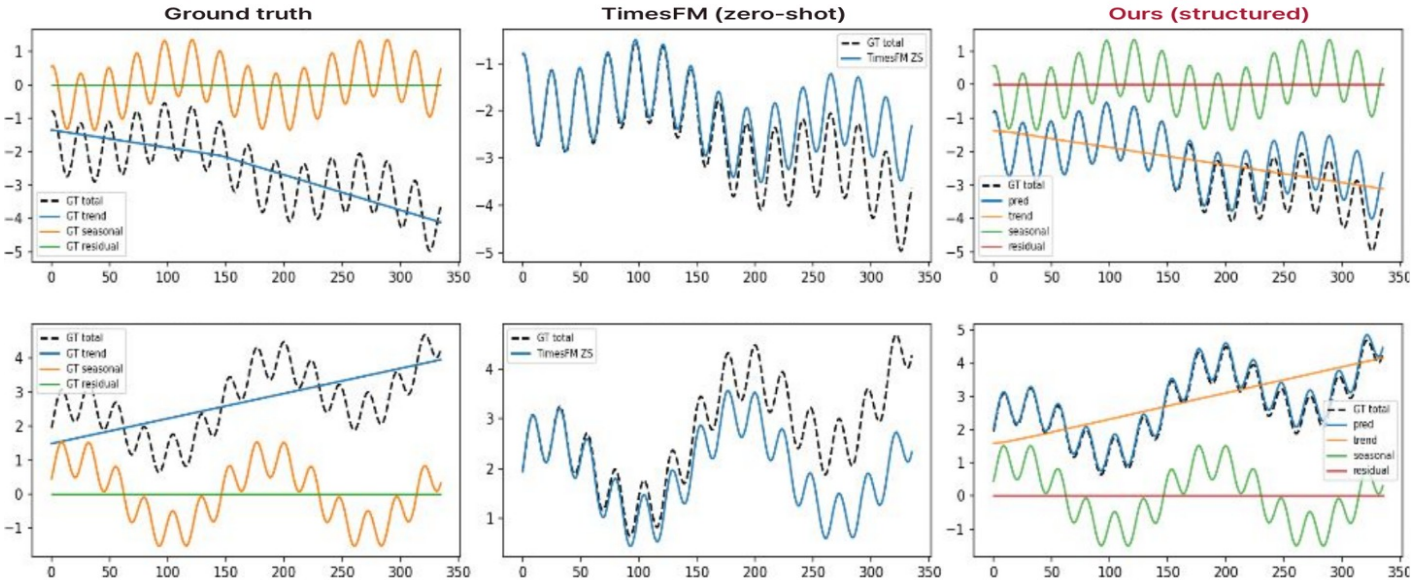
512 context → 96 horizon




Per-dataset MAE, 512 context → 96 horizon.

RESULTS · SYNTHETIC DATA

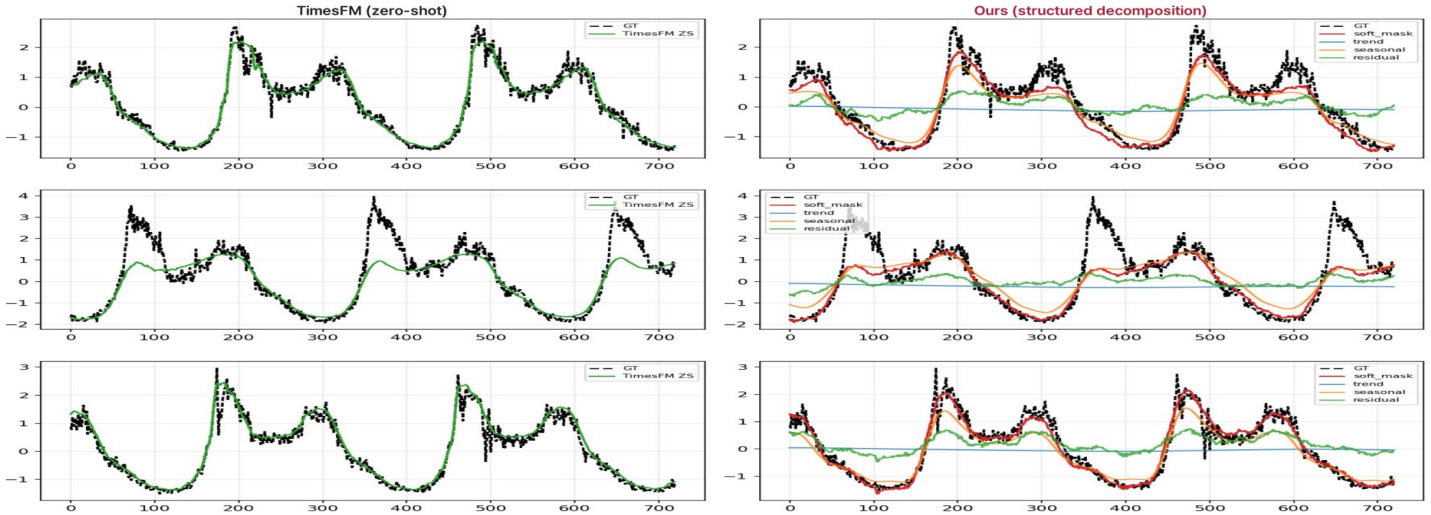
# On synthetic data, the split recovers ground truth




Left: ground-truth components · center: TimesFM zero-shot forecast · right: our recovered trend / seasonal / residual.

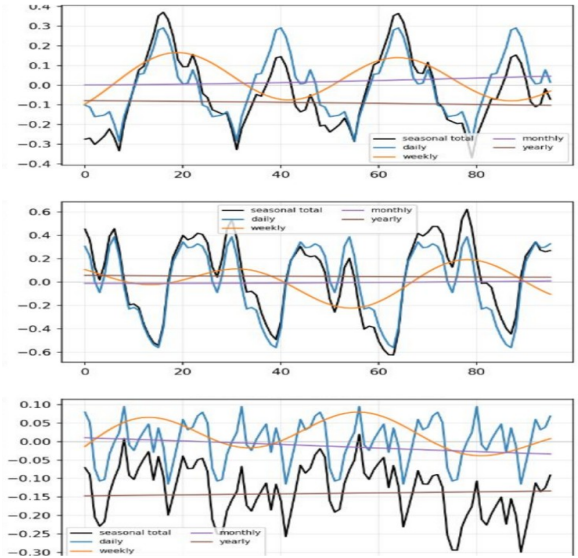
RESULTS · REAL DATA

# On real data, the same split holds




Transport / PEMS03, h = 720. Left: TimesFM zero-shot · right: our forecast with trend, seasonal, residual.

# Seasonality, split by cycle



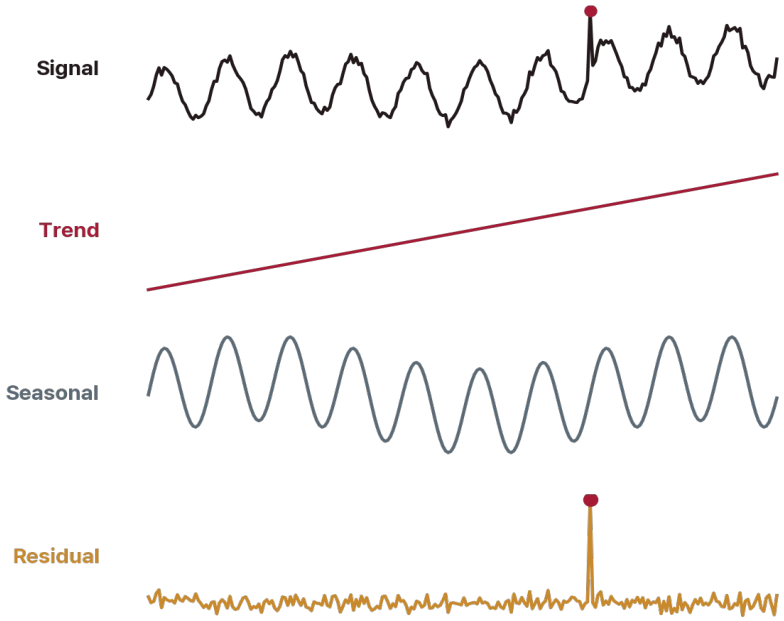

Within the seasonal component, **daily, weekly, monthly and yearly** cycles are separated and shown next to the total prediction.

Each cycle is the sum of the predicted Fourier harmonics with that period.

Black: seasonal total · colored: per-cycle parts.

BACK TO MANUFACTURING

# The three behaviors, decoded




**Trend**
the Drift from Part 1 → equipment wear → maintenance timing

**Seasonal**
the Rhythm of shifts and days → load & demand planning

**Residual**
a Regulation excursion → an anomaly alert

# Interpretable TSFMs

+ Decode a frozen TSFM into trend + seasonal + residual
+ Accuracy on par with the backbone, plus a decomposition it cannot give
+ Components map to maintenance, planning, and anomaly alerts