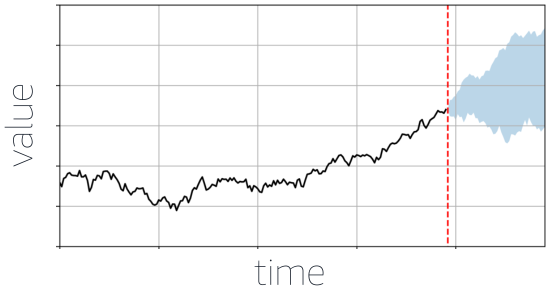
Efficiently Generating Correlated Sample Paths from Multi-step Time Series Foundation Models

Ethan Baron, Boris Oreshkin, Ruijun Ma, Hanyu Zhang, Kari Torkkola, Michael Mahoney, Andrew Wilson, Tatiana Konstantinova

BERT^{2}S Workshop at NeurIPS, December 7th, 2025

Demand Forecasting at Amazon Supply Chain Optimization Technology

Probabilistic demand forecasting



Buying systems

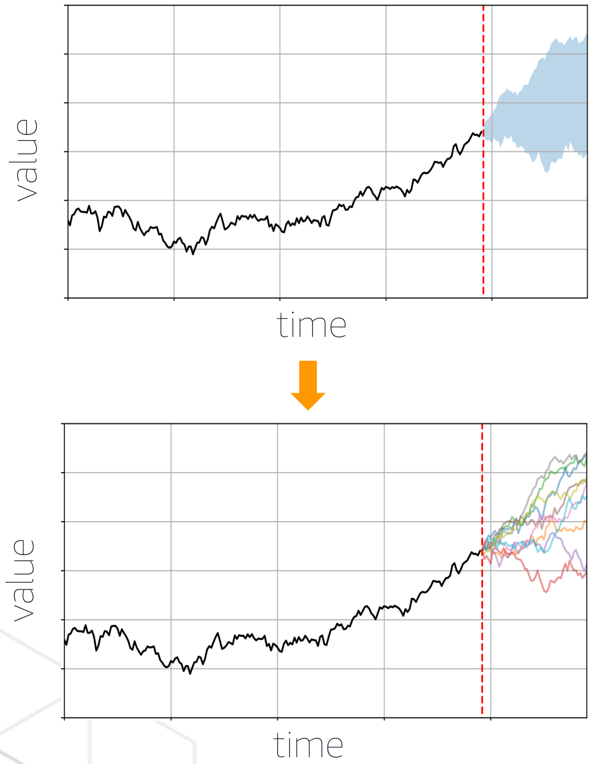
Product placements at fulfillment centers

Inventory planning

Labor planning

Complex systems require various forecast representations

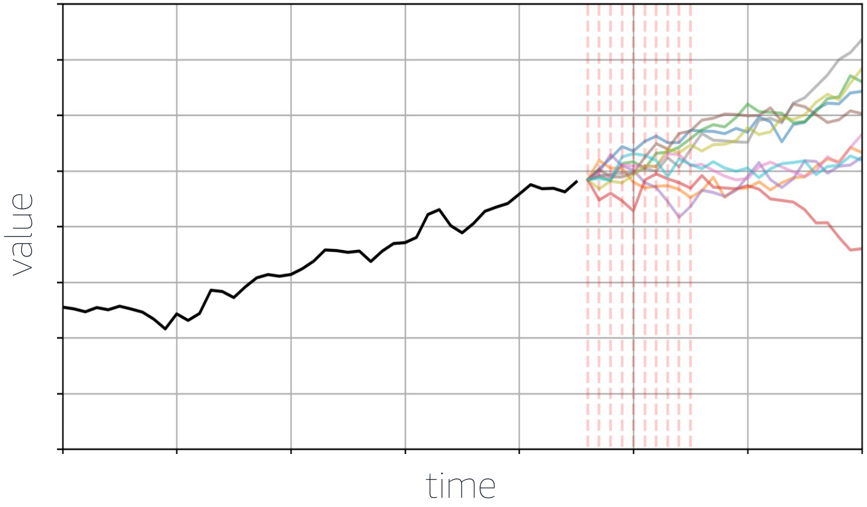
Forecasting Correlated Sample Paths



Use cases:

- Aggregated forecast for multiple horizons
- Conditional forecast for custom queries
- Inputs for simulations
- Inputs for RL training

Sample Path Generation: Naïve Approach



<u>Algorithm:</u>

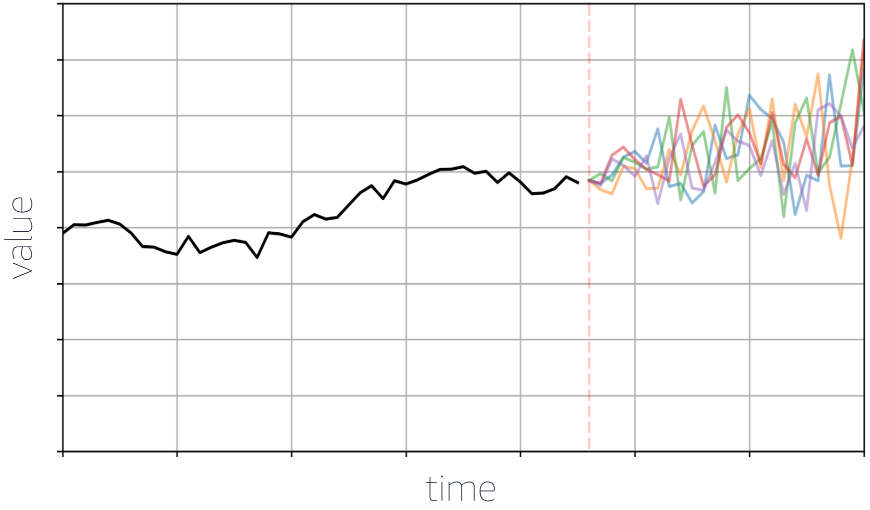
For each t in [0, H]:

1. Forecast distribution for 1 step
2. Sample from the distribution
3. Append to historical series
4. Pass as input to model for next step

<u>Concerns:</u>

1. Slow
2. Accumulation of error

Sample Path Generation: Fast Approach



Algorithm:

1. Generate a multi-horizonal probabilistic forecast
2. Sample N times from distribution at each horizon

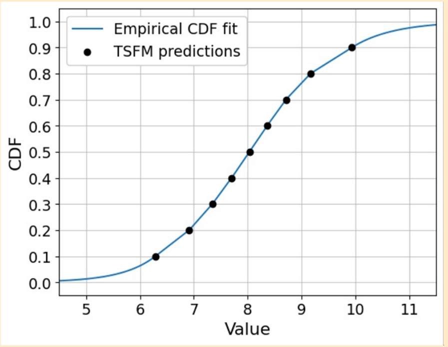
Concern:

1. Not a realistic correlation structure

amazon | science

Sample Path Generation: Copula Approach

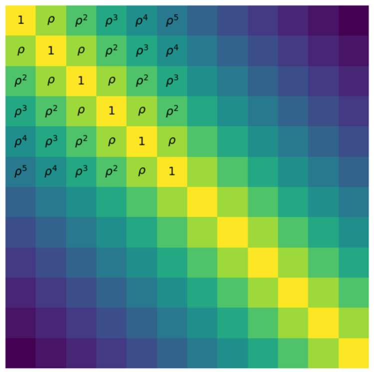
Joint distribution = Marginal distributions + Copula



Empirical CDF fit to quantile predictions

Gaussian with AR-style covariance

$$\Sigma_{ij} = \rho^{|i-j|}$$



Evaluating Sample Path Forecast

Marginal accuracy -- Continuous Ranked Probability Score:

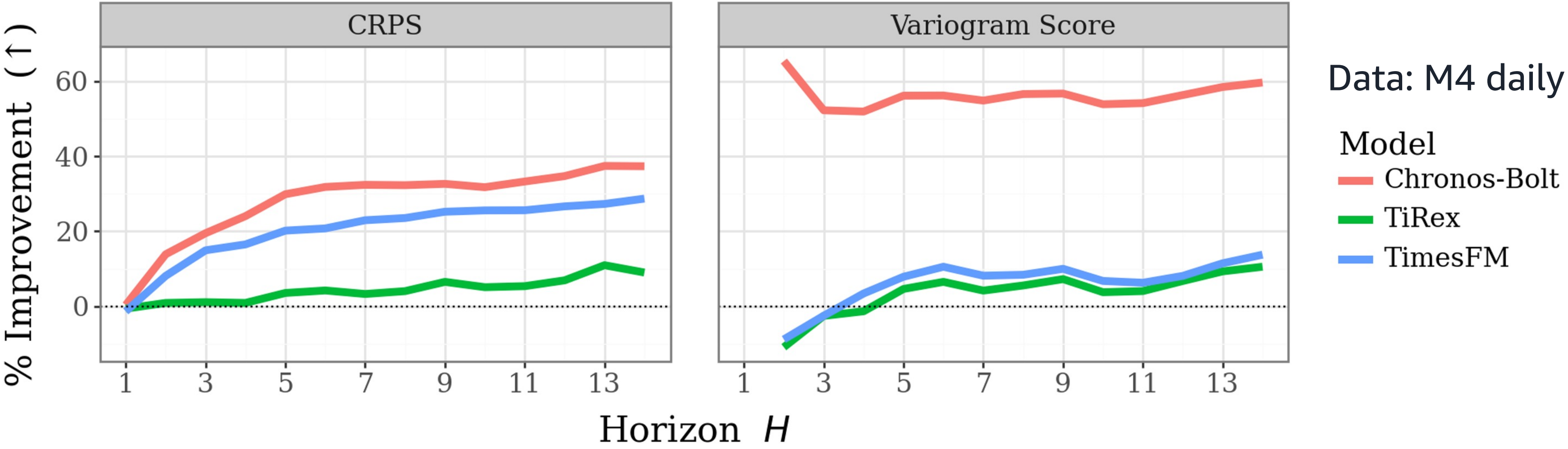
$$CRPS\ (p, \mathbf{x}) = \sum_{i=1}^{H} \mathbb{E}_{y\sim p}[|\mathbf{x}_i - \mathbf{y}_i|] - \frac{1}{2}\mathbb{E}_{y,z\sim p}[|\mathbf{y}_i - \mathbf{z}_i|]$$

Correlation structure of the joint predictive -- Variogram Score:

$$VS(p, \mathbf{x}) = \sum_{i=1}^{H}\sum_{j=1}^{H}\left(|\mathbf{x}_i - \mathbf{x}_j|^{0.5} - \mathbb{E}_{y\sim p}|\mathbf{y}_i - \mathbf{y}_j|^{0.5}\right)^2$$

The goal is to minimize both CRPS and VS.

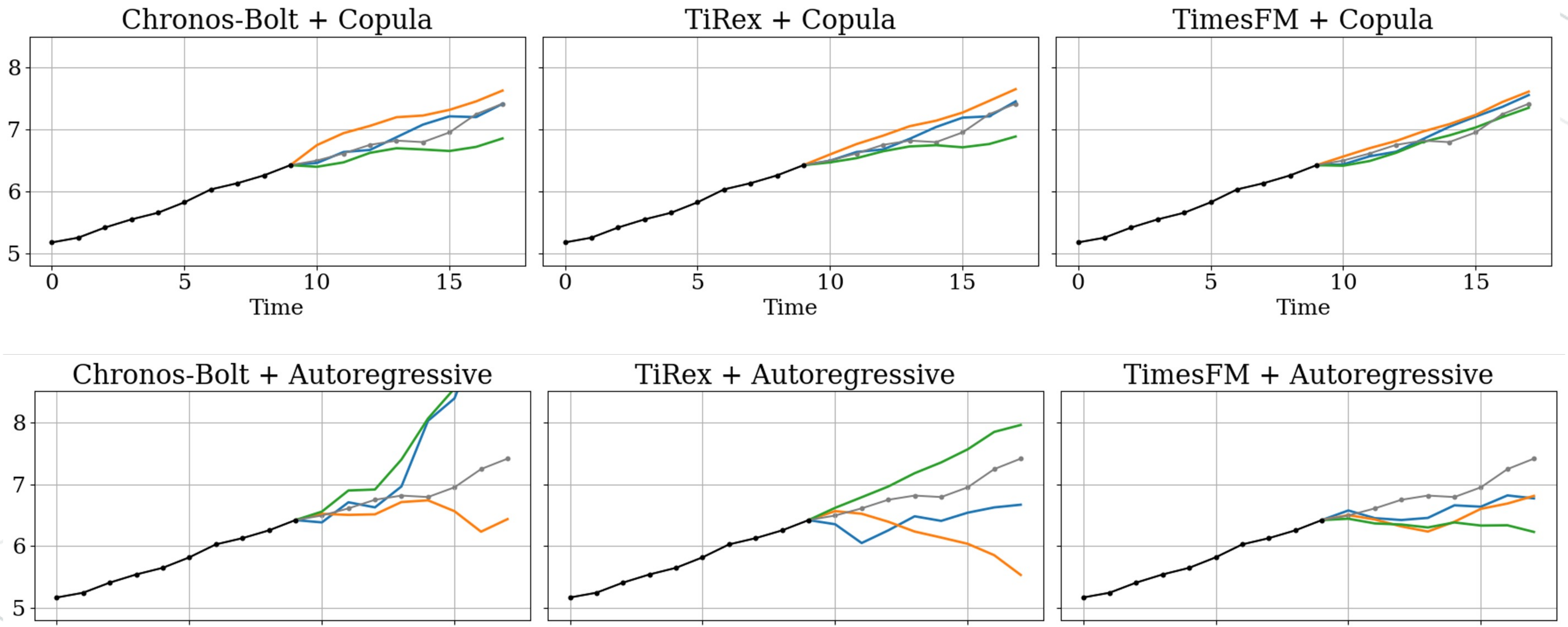
Copula vs Autoregressive Sample Path Generation



Data: M4 daily

Improvement for both Marginal and Joint distributions

Copula vs Autoregressive Sample Path Generation

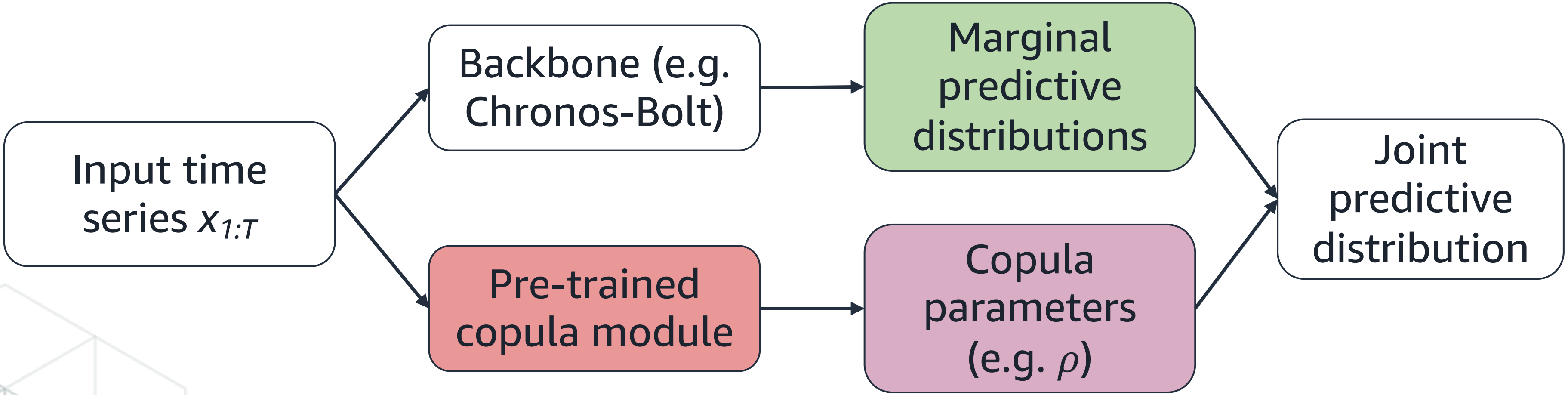


Can We Improve Further?

1. What if the empirical autocorrelation is not the best choice for ρ?

2. What if we want a more flexible correlation structure?

Solution: Pre-trained Copula Module (PCM)



Algorithm for Gaussian Copula

1. Generate 9 quantiles {0.1, 0.2, …, 0.8, 0.9} for marginal distributions for horizons 1 … H
2. Fit empirical CDFs using incremental quantile function approach to be able to generate any quantile
3. Calculate ρ as AR(1) from $X_{1:T}$ and compose covariance matrix Σ

$$\rho = Corr(x_{1:T-1}, x_{2:T})$$

4. Calculate a lower triangular matrix (Cholesky decomposition) : $\Sigma = LL^T$

$$L_{ij} = \begin{cases} \rho^{i-j}, & j = 1 \\ \rho^{i-j}\sqrt{1-\rho^2}, & 2 \leq j \leq i-1 \\ 0, & j > i \end{cases}$$

5. Generate N independently Normally $\mathcal{N}(0,1)$ distributed vectors x of size H
6. Correlate them with L: $p = Lx$ to get a vector of forecast percentiles
7. Sample from forecast CDFs to get N sample paths

Pre-trained Copula Module

- Architectures: MLP, TCN, GRU
- Train on M1, M4, Tourism. Test on M3
- Train on Variogram Score for *H = 8* with Chronos-Bolt Tiny marginals
- Two possible copula parameterizations:

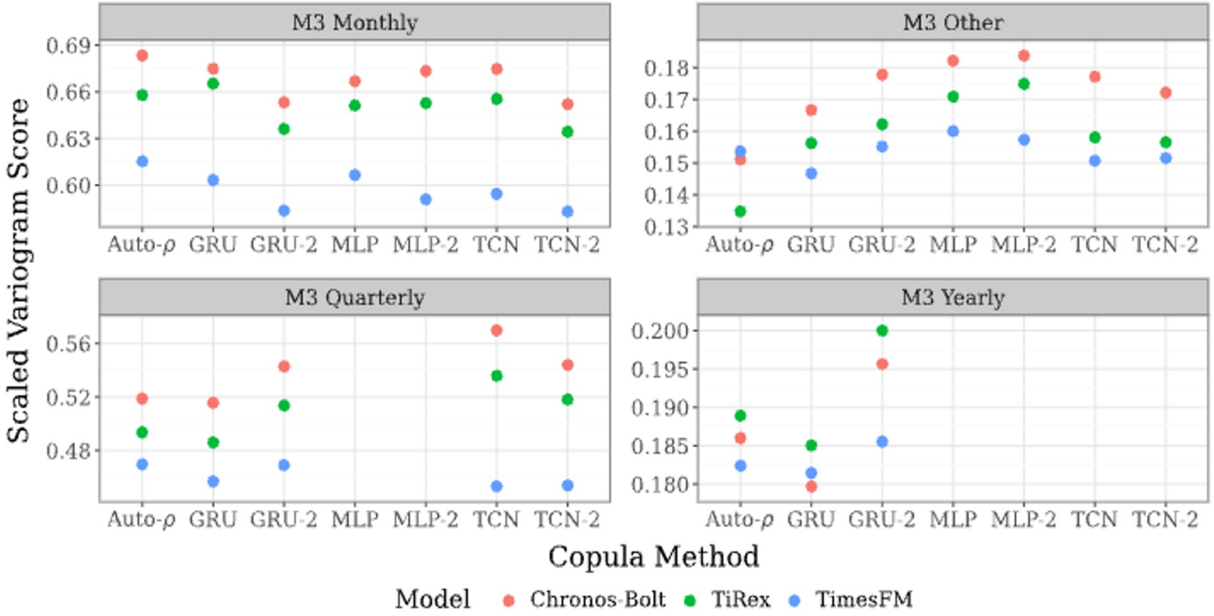
$$\Sigma_{ij} = \rho^{|i-j|}$$

ρ = **PCM**($x_{1:T}$)

$$\Sigma_{ij} = \beta\rho^{|i-j|} + (1-\beta)\delta_{ij}$$

ρ, β = **PCM**($x_{1:T}$)

Pre-trained Copula Module



- Can improve upon $\rho = AR(1)$
- Reusable across models
- GRU architecture worked the best

amazon | science

Summary

- We considered AR(1) Gaussian Copula as an efficient method for generating correlated sample path forecasts
- The method outputs realistic correlation structure while improving the accuracy comparing to stepping forecast
- Pre-trained copula module can potentially further improve the quality of forecast

What is next?

- PCM on larger dataset
- Increased parametrization of Σ
- Direct output of the correlated sample path by a model

Thank You.

amazon | science