CATS: Enhancing Multivariate Time Series Forecasting by Constructing Auxiliary Time Series as Exogenous Variables

Jiecheng Lu [1] Xu Han [2] Yan Sun [1] Shihao Yang [1]

[1]Georgia Institute of Technology [2]Amazon Web Services

Introduction

For Multivariate Time Series Forecasting (MTSF), recent deep learning applications show that univariate models frequently outperform multivariate ones. To address the deficiency in multivariate models, we introduce a method to Construct Auxiliary Time Series (CATS) that functions like a 2D temporal-contextual attention mechanism, which generates Auxiliary Time Series (ATS) from Original Time Series (OTS) to effectively represent and incorporate inter-series relationships for forecasting. Key principles of ATS—continuity, sparsity, and variability—are identified and implemented through different modules. Even with a basic 2-layer MLP (2L) as the core predictor, CATS achieves state-of-the-art, significantly reducing complexity and parameters compared to previous multivariate models, marking it as an efficient and transferable MTSF solution.

Overall Architecture and Basic Intuition of CATS

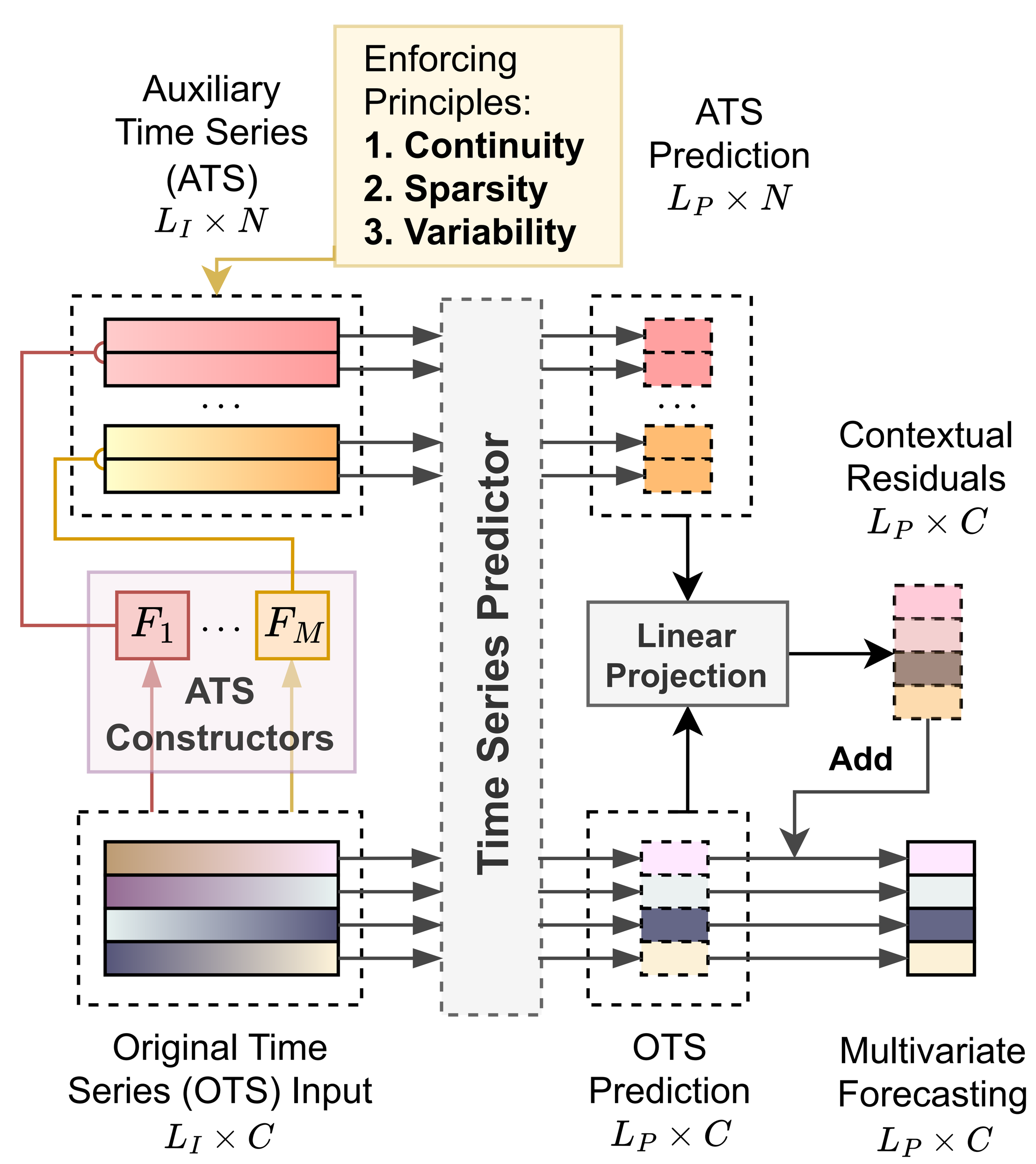

Figure 1. Overall Architecture of CATS

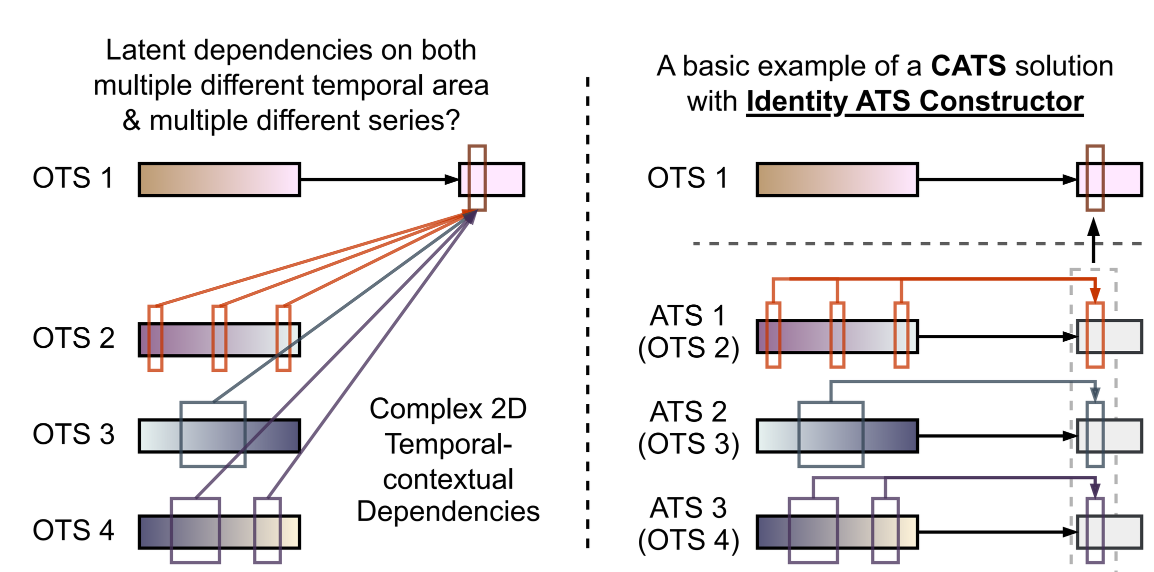

Figure 2. How CATS captures 2D Temporal-contextual Dependencies

How CATS Works: An Example on a Simple Shifting Problem

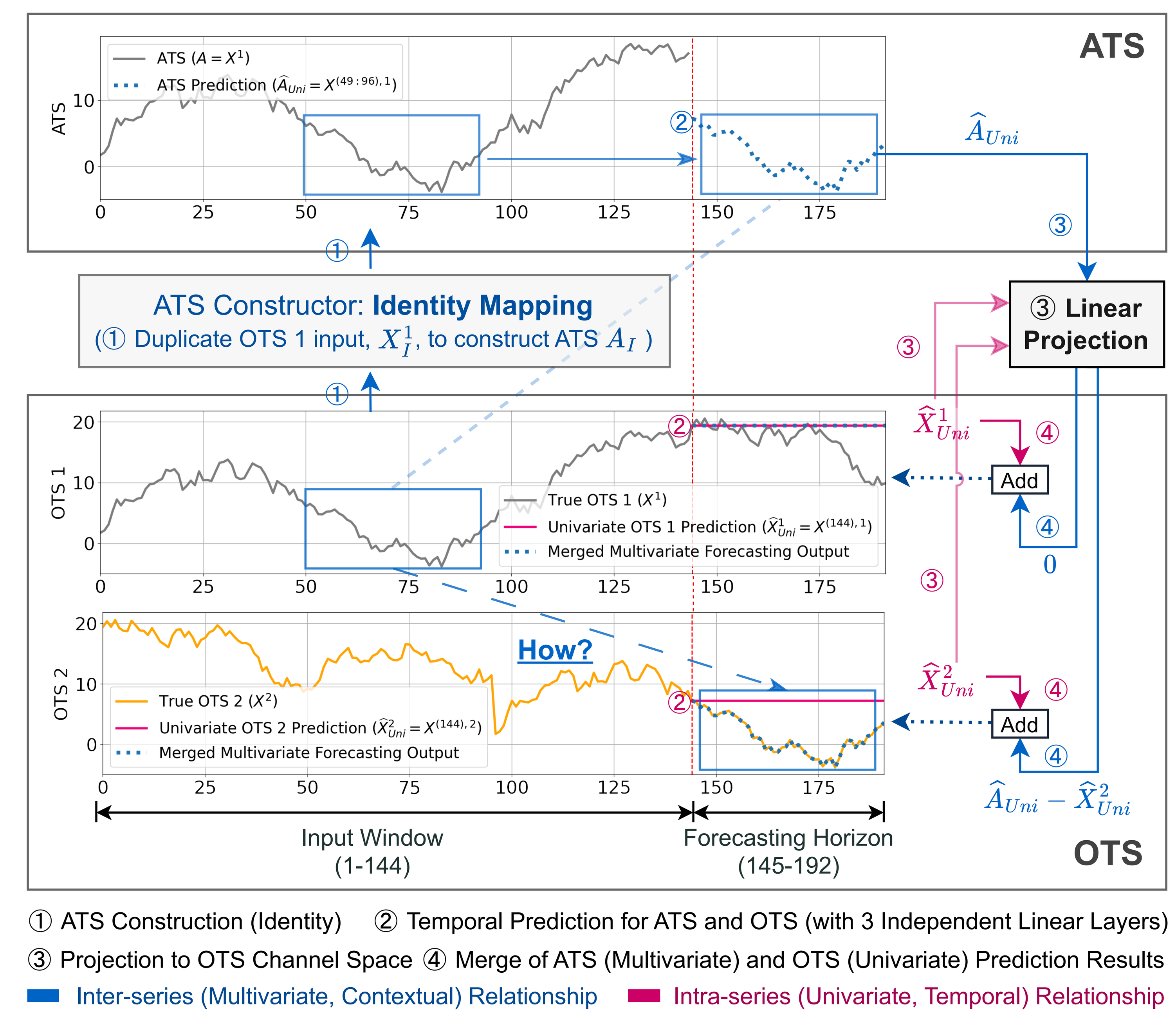

Figure 3. The functionality of CATS through a straightforward shifting problem. By simply duplicating the OTS to build the ATS (identity constructor) and using independent linear predictors, CATS transforms these basic univariate predictors to handle multivariate forecasting challenges. This approach is similar to a 2D temporal-contextual attention mechanism, where the ATS created from the original effectively captures inter-series relationships. Despite without complicated ATS constructors, CATS still efficiently captures multivariate series dependencies, demonstrating the adaptability fo the structure.

Three Key Principles of ATS: Variability

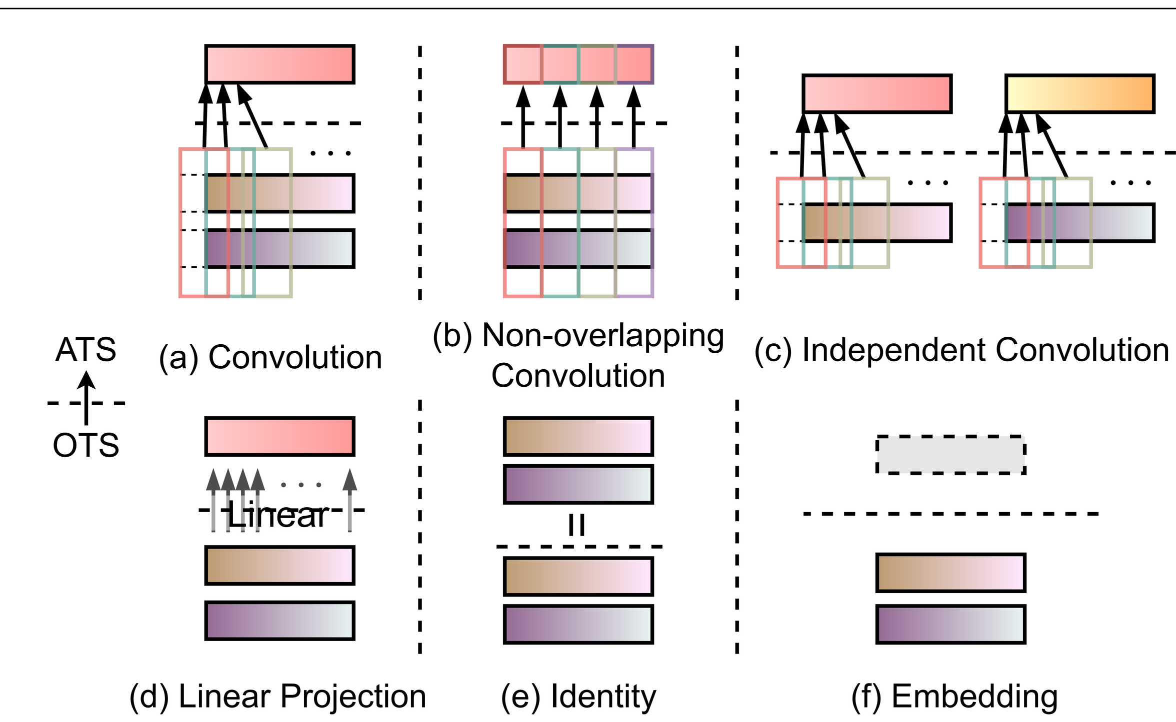

Figure 4. Variability represents the diversity of inter-series dependencies that ATS can capture, reflecting how effectively OTS information is combined in varied ATS from multiple perspectives. This approach helps to prevent the forecasting from being controlled by a specific type of ATS, allowing the model to adapt to diverse MTSF datasets with different characteristics of dependency. CATS achieves this through a variety of ATS constructors with differing structures.

Three Key Principles of ATS: Channel Sparsity

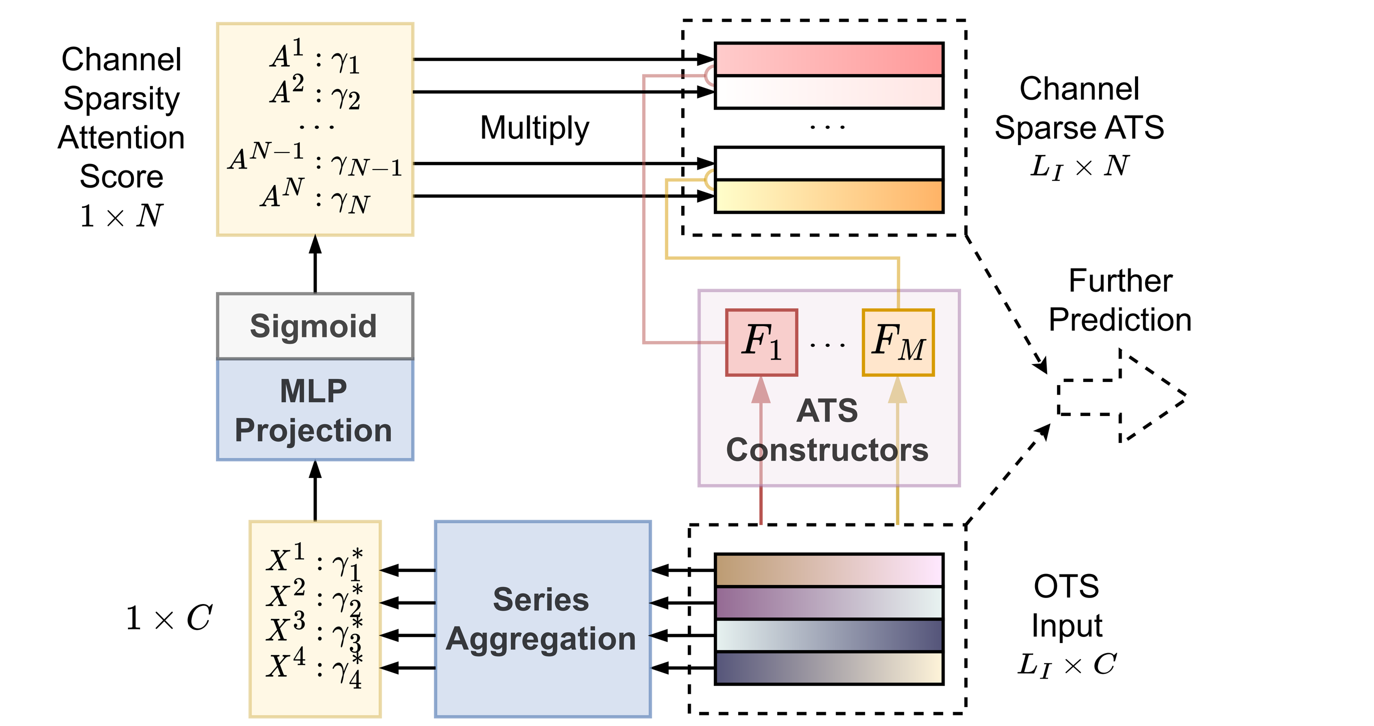

Figure 5. By activating only the most crucial ATS channels, model stability and performance would significantly improve. It also reduces the tuning cost for the hyperparameter of the number of ATS series. We refer to this as the principle of channel sparsity.

Three Key Principles of ATS: Continuity

Continuity is an important and inherent characteristic for most real-world time series datasets. We use a continuity loss term $\mathcal{L}_{\text{cont}}$ to enforce this principle of continuity in ATS:

$$\mathcal{L}_{\text{cont}} = \frac{\beta_{cont}}{L_I N}\sum_{t=2}^{L_I}\sum_{n=1}^{N}\left(\frac{A^{(t),n}-A^{(t-1),n}}{\sigma_n}\right)^2$$

where β_{cont} is a weighting factor that adjusts the influence of the continuity loss in the overall loss function. By applying continuity loss, ATS will smooth out temporal details and noise, focusing on overall trends that can be described by inter-series dependencies, leaving the remaining intra-series details to be better modeled by established methods on the OTS. This clear division of modeling focus using ATS and OTS enhances the reliability and reduces the complexity of forecasting.

Experiment Results

Table 1. Summary of MTSF results. Average MSE percentages (AvgMSE%) across all datasets and average rankings (AvgRank) of each model, with the count of first rankings (#Win), are included.

Models	CATS (2L)	ARM	PatchTST	DLinear	Autoformer	Informer	Repeat
AvgMSE%	**43.9%**	44.9%	55.4%	52.9%	79.4%	216.0%	100.0%
AvgRank	**1.25**	1.75	3.47	3.83	6.28	7.39	6.69
#Win	**28**	10	1	0	0	0	0

Conclusion

Our results reveal that CATS, even when paired with very simple predictor structures, significantly outperforms existing models in both strong and weak multivariate relationship scenarios. Furthermore, CATS offers a significant reduction in computational complexity and parameter size compared to other multivariate models, marking it as an easy-to-implement solution for MTSF.

jlu414@gatech.edu | icyxu@amazon.com | yansun@gatech.edu | shihao.yang@isye.gatech.edu