Scale-Invariant Conditional VAE for Coarse-Grained Economic Time-Series Forecasting

Jianping Zhu[1] , **Lei Wang**[2] , **Yang Chen**[3] , **Bo Jin**[1*] and **Xiaopeng Wei**[1*]
[1]Dalian University of Technology, Dalian, China
[2]The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, Guangdong, China
[3]Zhejiang University, Hangzhou, China
zhujp@mail.dlut.edu.cn, leiwang3@hkust-gz.edu.cn, ychen2014@zju.edu.cn, jinbo@dlut.edu.cn, xpwei@dlut.edu.cn

Abstract

Coarse-grained time series (CGTS) are critical for business and macroeconomic analysis. However, CGTS are typically updated infrequently and contain few observations, so model-centric training on raw data is prone to overfitting and degraded forecast accuracy. To address this, we propose SI-CVAE, a scale-invariant conditional generative model built around variable-length subsequences, and adopt a data-centric "train on synthetic, test on real" paradigm to enhance downstream forecasting. Specifically, we first introduce a variable-length subsequence clustering-and-matching algorithm to capture cross-scale recurring patterns within a series. We then design a frequency-domain conditional VAE with a frequency-domain linear decoder to enable controllable, arbitrary-length sequence synthesis while enforcing scale-invariance constraints. Finally, we develop a time-domain reconstruction strategy with subsequence-conditioned fusion to ensure temporal continuity and spectral consistency across concatenated segments. We conduct extensive experiments on six ship-sales datasets and four U.S. macroeconomic indicators. Results show that replacing or augmenting the original training set with SI-CVAE–generated data yields consistent accuracy gains across multiple forecasting baselines, and that SI-CVAE attains higher synthetic-data quality than state-of-the-art generators on standard metrics.

1 Introduction

Coarse-grained time series (CGTS)—time series observed at sparse and infrequent intervals—often provide insufficient supervision for effective model-centric training and reliable forecasting [Roque *et al.*, 2024]. Yet in many market-facing and operational scenarios, coarse release is the norm: due to privacy regulations, statistical-release protocols, and reporting cycles, economic and business signals are frequently published at low temporal resolution with only a handful of observations [Xie *et al.*, 2025]. For example, ship-sales records

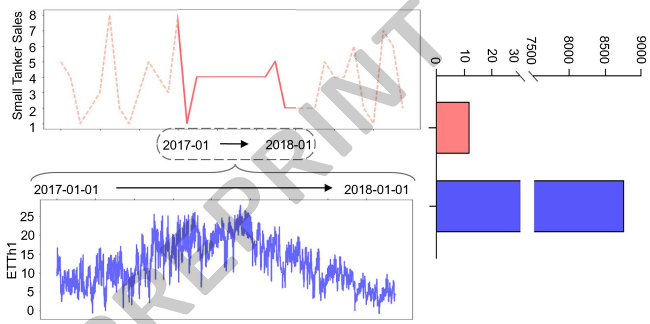


Figure 1: Comparison of the data volume generated by coarse-grained time series (Small Tanker Sales) and fine-grained time series (ETTh1) over one year.

are typically updated monthly, yielding merely 12 observations per year. As illustrated in Figure 1, compared with an hourly wind-power dataset [Zhou *et al.*, 2021], this corresponds to a 730× reduction in annual observations. Such sparsity obscures the evolution of volatility and regime shifts, causing forecasters trained directly on CGTS to overfit and generalize poorly [Zhu *et al.*, 2022]. These constraints make CGTS forecasting both challenging and practically important.

A natural remedy is a data-centric perspective: improving generalization by augmenting training with high-quality synthetic sequences [Yin *et al.*, 2024]. However, realistic CGTS generation is nontrivial under limited samples and measurement noise. First, scarce observations hinder learning diverse local fluctuation patterns, degrading both realism and diversity. Second, temporal continuity induces strong dependence on recent history, requiring generated segments to remain coherent with the prevailing volatility regime. Third, many economic signals arise from superposed multi-scale periodic components, calling for models that capture long-range periodic organization and scale-invariant regularities. While GANs [Goodfellow *et al.*, 2014] (*e.g.*, TimeGAN [Yoon *et al.*, 2019]), VAEs [Kingma and Welling, 2013] (*e.g.*, TimeVAE [Desai *et al.*, 2021], Ko-VAE [Naiman *et al.*, 2023]), and diffusion models [Sohl-Dickstein *et al.*, 2015; Brophy *et al.*, 2023; Gonen *et al.*, 2025] have been explored for time-series generation, many existing designs still emphasize local continuity and underrepresent global periodicity and scale-invariant struc-

[*]Corresponding authors.

ture—capabilities that are crucial for faithful CGTS synthesis and downstream forecasting.

To address these issues, we propose SI-CVAE, a scale-invariant conditional generative model over variable-length subsequences. We decompose CGTS generation into three stages: (i) we discover scale-invariant patterns via variable-length subsequence clustering and segment the series accordingly; (ii) conditioned on each subsequence's cluster label, we introduce a Fre-CVAE—a conditional VAE equipped with a frequency-domain linear-mapping decoder—to generate variable-length subsequences while preserving periodic structure; and (iii) we assemble generated subsequences through a conditional subsequence fusion mechanism, enforcing continuity across boundaries and coherent regime transitions.

The main contributions of our work are summarized as follows:

- We formulate CGTS forecasting in market-facing settings as a data augmentation problem and propose SI-CVAE, a scale-invariant conditional VAE framework that supports arbitrary-length generation via variable-length subsequence modeling.
- We identify diverse fluctuation regimes through variable-length subsequence clustering, and integrate class-conditional generation with candidate selection conditioned on the immediately preceding subsequence, improving realism and regime consistency under coarse-grained supervision.
- We conduct extensive experiments on six real-world ship-sales datasets and four representative U.S. macroeconomic indicators. Results show that synthetic-data augmentation consistently improves baseline forecasters, and SI-CVAE achieves superior generation quality compared with competing methods.

2 Problem Definition

Coarse-grained generation for data augmentation. Coarse-grained economic time series (CGTS) are low-frequency sequences aggregated over extended reporting intervals (*e.g.*, weekly, monthly, or quarterly). Such coarse sampling yields scarce observations and thus weak supervision for learning robust temporal dynamics; forecasters trained directly on raw CGTS are prone to overfitting and poor out-of-sample generalization. We therefore adopt a data-centric formulation: learn a generator that expands a short observed CGTS into long, high-fidelity synthetic sequences, which serve as augmentation data for downstream forecasting. In this paper, we focus on the generative component.

Formally, given an observed CGTS $\mathcal{X} \in \mathbb{R}^{T_x}$, we aim to learn a generator f_{θ_G} that produces a substantially longer synthetic series $\hat{\mathcal{X}} \in \mathbb{R}^{\hat{T}_x}$ with $\hat{T}_x \gg T_x$:

$$\hat{\mathcal{X}} = f_{\theta_G}(\mathcal{X}). \tag{1}$$

The generated $\hat{\mathcal{X}}$ can then be used to train a downstream predictor f_{θ_P} for forecasting targets $\mathcal{Y}$:

$$\mathcal{Y} = f_{\theta_P}(\hat{\mathcal{X}}). \tag{2}$$

SI-CVAE generation via scale-invariant variable-length subsequences. SI-CVAE first applies an unsupervised, pretrained variable-length subsequence clustering-and-matching module $f_{\theta_{SC}}$ to segment $\mathcal{X}$ and assign each extracted subsequence to one of K categories:

$$\{\mathcal{S}^k\}_{k=1}^K = f_{\theta_{SC}}(\mathcal{X}), \tag{3}$$

where $\mathcal{S}^k = \{s_1^k, \dots, s_{m^k}^k\}$ denotes the set of extracted subsequences in category k. Each subsequence $s_j^k \in \mathbb{R}^{l_j^k}$ has variable length $l_j^k \in [l_{\min}, l_{\max}]$, where $l_{\min}$ and $l_{\max}$ control feasible segment durations. For each category k, SI-CVAE learns a corresponding Fre-CVAE to generate category-specific subsequences conditioned on context; a conditional subsequence fusion module then assembles generated segments into a coherent long-range synthetic series.

Learning objective. Let $c \in \mathbb{R}^{T_c}$ denote the immediately preceding subsequence used as conditional context during synthesis (from the observed prefix or previously generated segments). Let $\hat{S} = \{\hat{S}^k\}_{k=1}^K$ denote the collection of *generated* subsequence sets, where $\hat{S}^k = \{\hat{s}_1^k, \dots, \hat{s}_{\hat{m}^k}^k\}$. SI-CVAE aims to model the conditional distribution $p(\hat{\mathcal{X}}, \hat{S} \mid c)$. For tractable learning, we adopt the following factorization:

$$\log p(\hat{\mathcal{X}}, \hat{S} \mid c) = \sum_{k=1}^{K} \log p(\hat{S}^k \mid c) + \log p\Big(\hat{\mathcal{X}} \mid \{\hat{S}^k\}_{k=1}^K, c\Big), \tag{4}$$

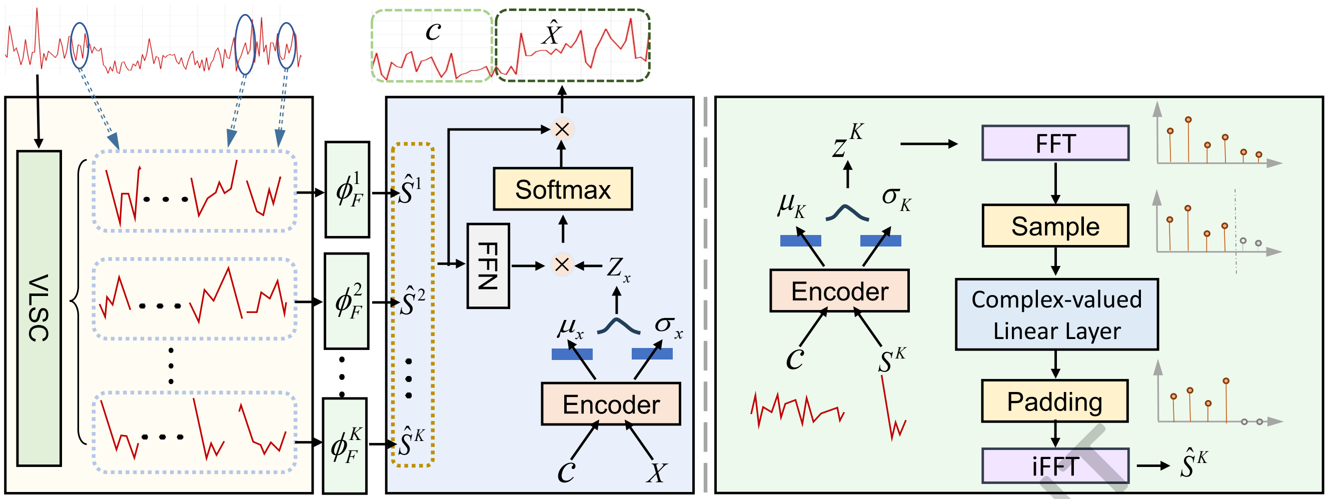
where $\log p(\hat{S}^k \mid c)$ is induced by the k-th Fre-CVAE, and $\log p\Big(\hat{\mathcal{X}} \mid \{\hat{S}^k\}_{k=1}^K, c\Big)$ corresponds to the conditional fusion module that composes the final synthetic sequence.

3 Methodology

This section presents SI-CVAE, a scale-invariant conditional variational generation framework tailored to coarse-grained time series. As shown in Figure 2(a), SI-CVAE consists of three modules: (i) a *variable-length subsequence clustering* module that discovers scale-invariant local patterns and assigns subsequences to K categories; (ii) a *frequency-domain conditional variational autoencoder* (Fre-CVAE) that performs category-conditional generation while *aligning variable lengths* via explicit spectral length control and direct frequency-domain decoding; and (iii) a *conditional subsegment fusion* module that aggregates generated subsequences through attention, yielding a coherent long synthetic series.

3.1 Variable-Length Subsequence Clustering Module

Fixed-window subsequence clustering hinges on a predefined segment length, which can distort the empirical distribution when temporal scales vary. To robustly capture scale-invariant patterns, we introduce a variable-length clustering scheme that couples an adaptive greedy segmentation strategy with Dynamic Time Warping (DTW), enabling consistent matching across subsequences of different lengths.



(a) Overall Workflow

(b) Fre-CVAE

Figure 2: Overview of the proposed SI-CVAE architecture.

Adaptive Greedy Segmentation Strategy

Given an input series $\mathcal{X}$, for each starting time t we evaluate candidate lengths $l_t \in [l_{\min}, l_{\max}]$. For each candidate segment $\mathcal{X}_{t:t+l_t}$, we compute its DTW distance to each centroid $p \in \mathcal{P}$ and select the length that best matches the closest centroid:

$$l_t^* = \underset{l_t \in [l_{\min}, l_{\max}]}{\arg\min} \; [DTW(\mathcal{X}_{t:t+l_t}, p)], \qquad \forall p \in \mathcal{P}, \tag{5}$$

where $\mathcal{P}$ denotes the centroid set. After greedily segmenting the full series, we alternate between (i) assigning each subsequence to its nearest centroid under DTW and (ii) updating centroids with a DTW-consistent prototype update (details deferred to the appendix), until convergence or a fixed iteration budget.

Variable-Length Subsequence Similarity Measure

DTW is a canonical similarity measure for sequences with unequal lengths and temporal misalignment [Imamura and Nakamura, 2024]. For a subsequence $\mathcal{X}_{t:t+l_t}$ and centroid p, DTW optimizes over admissible alignment paths $A \in \mathcal{A}$:

$$DTW(\mathcal{X}_{t:t+l_t}, p) := \min_{A \in \mathcal{A}} \langle A, \Delta(\mathcal{X}_{t:t+l_t}, p) \rangle, \tag{6}$$

where $\Delta(\cdot, \cdot)$ is the pointwise distance matrix:

$$\Delta(\mathcal{X}_{t:t+l_t}, p) = [\delta(\mathcal{X}^i_{t:t+l_t}, p^j)]_{ij}, \tag{7}$$

with $\delta(\cdot, \cdot)$ denoting a pointwise distance. This DTW criterion allows subsequences spanning different temporal extents to be compared consistently, which is essential for adaptive segmentation and robust clustering.

3.2 Frequency-domain Conditional Variational Autoencoder

After clustering, we conditionally generate subsequences for each category via Fre-CVAE. The key difficulty is that subsequences across categories have variable lengths, whereas downstream fusion requires compatible representations. We address this by aligning lengths through frequency-domain operations and by introducing a direct frequency-domain linear decoder.

Domain Conversion/Inversion

Frequency-domain representations expose periodic structures that can be obscured in the time domain. We therefore map subsequences to the frequency domain, where low-pass filtering or masking preserves the dominant spectral characteristics while enabling flexible length adjustment in the reconstructed time-domain signal (e.g., stretching/compression). Concretely, we adopt the Fast Fourier Transform (FFT) [Brigham and Morrow, 1967] (implemented as rFFT for real-valued inputs), which maps N samples to $N/2+1$ complex coefficients, yielding a compact complex spectrum $\mathcal{T}$ for subsequent filtering and processing:

$$\begin{aligned} \mathcal{T}(f) &= \int_{-\infty}^{\infty} T(v) e^{-j2\pi f v}\, dv \\ &= \int_{-\infty}^{\infty} T(v)\cos(2\pi f v)\, dv \\ &\quad + j \int_{-\infty}^{\infty} T(v)\sin(2\pi f v)\, dv. \end{aligned} \tag{8}$$

where f is the frequency variable, v is the integral variable, and j is the imaginary unit. $\int_{-\infty}^{\infty} T(v)\cos(2\pi f v) dv$ represents the real part of $\mathcal{T}$, and $\int_{-\infty}^{\infty} T(v)\sin(2\pi f v) dv$ represents the imaginary part. After performing linear mapping or masking operations in the frequency domain, the signal is mapped back to the time domain via inverse Fourier Transform, restoring the time-domain signal $T(v)$, as follows:

$$T(v) = \int_{-\infty}^{\infty} \mathcal{T}(f) e^{j2\pi f v} df. \tag{9}$$

This frequency-domain conversion and inversion process effectively unifies the lengths of subsequences across different categories while preserving their main frequency characteristics.

Conditional Variational Lower Bound

Within the Fre-CVAE framework, we achieve accurate generation and reconstruction of different category subsequences by maximizing the first likelihood term derived from the theoretical framework. Specifically, the Evidence Lower Bound (ELBO) for the Fre-CVAE corresponding to the k-th category can be factorized as follows:

$$\begin{aligned}\mathcal{L}^k_{Fre-CVAE} =&\mathbb{E}_{q_{\varphi_k}(z^k|S^k,c)}[\log p_{\theta_k}(S^k|z^k,c)] \\ &- KL[q_{\varphi_k}(z^k|S^k,c)||p_{\theta^p_k}(z^k|c)],\end{aligned} \tag{10}$$

where φ_k and θ_k parameterize the encoder and decoder, and θ^p_k parameterizes the conditional prior.

Given a subsequence $S^k \in \mathbb{R}^{l^k}$, we transform it into the frequency domain, apply an LPF (truncate/pad) to a fixed spectral size corresponding to a target length T_c, and invert back to obtain an aligned representation $\tilde{S}^k \in \mathbb{R}^{T_c}$:

$$\tilde{S}^k = iFFT\big(LPF(FFT(S^k))\big). \tag{11}$$

We then concatenate $\tilde{S}^k$ with condition c and predict posterior parameters:

$$\begin{aligned}&\mu_k = W^\mu_F \cdot [\tilde{S}^k, c] + b^\mu_F \\ &\sigma_k = \log(1+\exp(W^\sigma_F \cdot [\tilde{S}^k, c] + b^\sigma_F)) \\ &z^k \sim \mathcal{N}(\mu_k, \sigma^2_k),\end{aligned} \tag{12}$$

where $[\cdot,\cdot]$ denotes the concatenation operation, and W^μ_F, W^σ_F, b^μ_F, and b^σ_F denote trainable parameters. The first term in the formula represents the conditional reconstruction loss. Since the lengths of the input subsequence S^k and the reconstructed subsequence $\hat{S}^k$ are inconsistent, we impose a reconstruction constraint using Dynamic Time Warping (DTW). The DTW algorithm effectively aligns time series data, ensuring that the reconstructed sequence accurately reflects the characteristics of the original sequence:

$$\mathcal{L}^k_{rec} = DTW(S^k, \hat{S}^k). \tag{13}$$

To capture nonlinear temporal dependence beyond a fixed Gaussian prior, we model $p_{\theta^p_k}(z^k|c)$ with two GRU units:

$$\begin{aligned}&p_{\theta^p_k}(z^k_t|z^k_{<t},c) = \mathcal{N}(\mu^k_t, (\sigma^k_t)^2) \\ &\mu^k_t = \mathrm{GRU}^k_\mu(z^k_{<t}, c) \\ &\sigma^k_t = \log(1+\exp(\mathrm{GRU}^k_\sigma(z^k_{<t}, c))),\end{aligned} \tag{14}$$

where $\theta^p_k = \{\mathrm{GRU}^k_\mu, \mathrm{GRU}^k_\sigma\}$.

Frequency-Domain Linear Decoding

To generate a length-aligned subsequence from z^k, we introduce a direct frequency-domain linear decoder. We (i) apply FFT along the temporal dimension, (ii) retain salient spectral content via LPF and fix the spectral size to a hyperparameter η for stable generation, (iii) map complex coefficients with a complex-valued linear layer, (iv) zero-pad the spectrum to the target size $\lfloor T_x/2 \rfloor + 1$, and (v) apply iFFT to obtain $\hat{S}^k$ in the time domain. This design yields explicit, controllable length alignment while preserving dominant frequency characteristics.

3.3 Conditional Subsegment Fusion

Complex time series are often composed of multiple overlapping sub-patterns, and capturing and reconstructing these sub-patterns is crucial for generating high-quality time series. In the SI-CVAE framework, the conditional sub-segment fusion module integrates the subsequences generated by different categories to ensure the coherence and consistency of the final generated sequence. This process not only enhances the overall integrity of the generated sequence but also improves the model's robustness in handling diverse time series.

In the conditional sub-segment fusion module, we maximize the second likelihood term $\log p(X|\{S^k\}^K_{k=1}, c)$ in Equation 4 by effectively fusing the subsequences generated from different categories to produce the final sequence X. The variational lower bound (ELBO) for this process can be expressed as:

$$\begin{aligned}\mathcal{L}_F =&\mathbb{E}_{q_{\varphi_F}(z_x|X,c)}[\log p_{\theta_F}(X|z_x, c, \{S^k\}^K_{k=1})] \\ &- KL[q_{\varphi_F}(z_x|X,c)||p_{\theta^p_F}(z_x|c)],\end{aligned} \tag{15}$$

where φ_F represents the encoder of the fusion module, whose structure is consistent with Fre-CVAE and is mainly used to encode the input sequence X into the latent variable z_x. θ_F represents the sub-segment fusion decoder used to generate the final sequence, and θ^p_F is the conditional prior model used to learn the prior distribution under the condition. The first term in the formula represents the reconstruction loss, which we constrain using Mean Squared Error (MSE) to compare the original sequence X with the reconstructed sequence $\hat{X}$:

$$L^F_{rec} = -\left\|\hat{X} - X\right\|^2. \tag{16}$$

The second term in the formula represents the KL divergence constraint, for which we use two GRU units as the conditional prior model to learn the conditional prior distribution.

During the decoder phase, the subsequences $\{\hat{S}^k\}^K_{k=1}$ reconstructed from different categories are first encoded into feature vectors $\{h^k\}^K_{k=1}$ via a feedforward neural network (FFN):

$$h^k = FFN(\hat{S}^k). \tag{17}$$

These feature vectors $\{h^k\}^K_{k=1}$ work together with the latent variable z_x to calculate attention scores, thereby aggregating subsequences from different categories:

$$a_k = \frac{\exp(h^k \cdot z_x)}{\sum\limits_{j=1}^{K}\exp(h^j \cdot z_x)}, \tag{18}$$

where a_k represents the attention weight of the k-th subsequence, reflecting its contribution to generating the final sequence $\hat{X}$. The final generated sequence $\hat{X}$ can be represented as a weighted sum of all subsequences:

$$\hat{X} = \sum_{k=1}^{K} a_k \hat{S}^k. \tag{19}$$

Finally, the model loss function can be expressed as:

$$\mathcal{L} = \sum_{k=1}^{K} \mathcal{L}^k_{Fre-CVAE} + \mathcal{L}_F. \tag{20}$$

Base model	Train set	Ship-sales datasets						U.S. macroeconomic indicators			
		Small Tanker	VLCC	Suezmax Tanker	Aframax Tanker	Panamax Bulker	Handysize Bulker	CPI	Services PMI	Manufacturing PMI	PPI
LaST	$\mathcal{D}_T$	0.89	4.42	1.14	4.27	3.11	1.25	0.10	1.58	0.84	0.38
	$\mathcal{D}_G$	**0.63**	**3.43**	**0.85**	**3.32**	**1.68**	**1.03**	**0.05**	**1.30**	**0.49**	**0.27**
LSTM	$\mathcal{D}_T$	1.28	5.36	1.70	4.76	3.53	1.61	0.09	2.84	1.36	0.44
	$\mathcal{D}_G$	**0.69**	**3.91**	**0.86**	**3.84**	**1.92**	**1.11**	**0.04**	**1.41**	**0.78**	**0.23**
FITS	$\mathcal{D}_T$	0.75	3.44	1.06	3.48	1.39	1.11	0.08	2.15	1.01	0.50
	$\mathcal{D}_G$	**0.62**	**3.06**	**0.88**	**3.15**	**1.35**	**1.00**	**0.06**	**1.72**	**0.72**	**0.35**

Table 1: Performance comparison of different base models (including LaST, LSTM, and FITS) trained on synthetic data $\mathcal{D}_G$ generated by SI-CVAE versus original training data $\mathcal{D}_T$ across six ship sales datasets and four U.S. macroeconomic indicators.

Datasets	SI-CVAE	ImagenFew	KoVAE	FTS-Diffusion	GTGAN	TimeGAN	TimeVAE
Small Tanker	**0.63**	0.77	0.73	0.80	1.19	1.14	0.80
VLCC	**3.43**	3.95	4.01	4.30	5.58	5.31	4.12
Suezmax Tanker	**0.85**	0.94	0.93	0.98	1.54	1.52	0.97
Aframax Tanker	**3.32**	3.84	3.99	3.98	5.20	4.96	3.55
Panamax Bulker	**1.68**	2.91	2.64	2.39	2.40	4.23	2.26
Handysize Bulker	**1.03**	1.36	1.11	1.20	1.32	1.34	1.12
CPI	**0.05**	0.05	0.07	0.06	0.06	0.06	0.06
Services PMI	**1.30**	1.48	1.47	1.43	2.00	3.55	2.77
Manufacturing PMI	**0.49**	0.60	0.61	0.87	1.05	1.77	0.79
PPI	**0.27**	0.29	0.29	0.32	0.37	0.32	0.32

Table 2: Comparison of generation quality across six ship sales datasets and four U.S. macroeconomic indicators under the LaST baseline model.

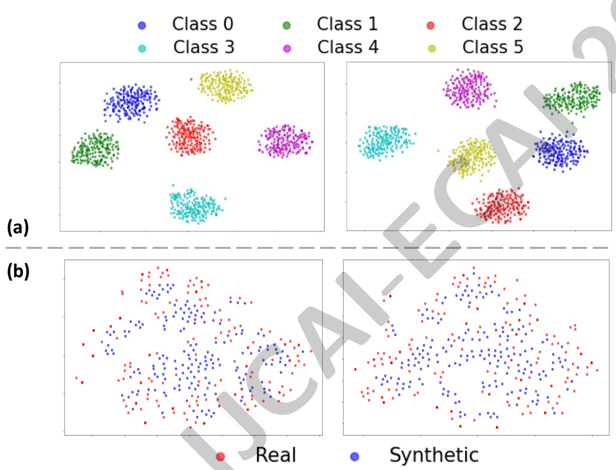


Figure 3: t-SNE-based representation visualization of SI-CVAE model performance on the Small Tanker Sales dataset. (a) Latent vector distribution for each clustering category. (b) Comparison of the distribution of synthetic and original data.

4 Experimental Settings

We evaluate SI-CVAE from three complementary perspectives: *(i) downstream utility*, *i.e.*, whether SI-CVAE-generated coarse-grained time series (CGTS) improve forecasting performance as data augmentation; *(ii) generative fidelity*, i.e., the realism and diversity of synthesized series compared with representative state-of-the-art generators; and *(iii) interpretability and robustness*, via visualization and ablations that isolate the contributions of key components. For fair comparison, all methods share identical *chronological* data splits and a unified training/evaluation pipeline.

4.1 Settings

Datasets and protocol. We consider two families of sparse, monthly economic CGTS: six ship-type sales series and four U.S. macroeconomic indicators. The ship-sales series (Small Tanker, VLCC, Suezmax Tanker, Aframax Tanker, Panamax Bulker, and Handysize Bulker) are collected from Clarksons Research [Clarksons Research, 2025]. The macroeconomic indicators include CPI, Services PMI, Manufacturing PMI, and PPI from the Wind database [Wind Information Co., Ltd., 2025]. Ship-sales span January 1995–December 2024 (360 time points), and macro series span January 2000–December 2024 (300 time points). For each series, we apply a chronological training/validation/test split with a 4:2:4 ratio to avoid temporal leakage.

Baselines. For generative modeling, we compare against representative methods from three paradigms: GAN-based (TimeGAN [Yoon *et al.*, 2019], GTGAN [Jeon *et al.*, 2022]), VAE-based (TimeVAE [Desai *et al.*, 2021], KoVAE [Naiman *et al.*, 2023]), and diffusion-based (FTS-Diffusion [Brophy *et al.*, 2023], ImagenFew [Gonen *et al.*, 2025]). To assess augmentation utility, we adopt three forecasting backbones as downstream predictors: LSTM [Hochreiter and Schmidhuber, 1997], LaST [Wang *et al.*, 2022], and FITS [Xu *et al.*, 2023].

Implementation. Unless otherwise specified, we optimize all models with Adam. For SI-CVAE, we set the learning rate to 1×10^{-5}, batch size to 8, and latent feature dimension to 64. For variable-length subsequence clustering, we set $l_{\min} = 4$, $l_{\max} = 8$, number of clusters $K = 6$, and run 50 iterations. During training, for each input sample X we randomly sample subsequences from each category to form $\{S^k\}_{k=1}^K$. At inference, we sample latent vectors from the conditional prior as decoder inputs; the generated subsequence length is drawn uniformly from $[l_{\min}, l_{\max}]$, and the previously synthesized subsequence serves as the condition c for the current generation step. All experiments are conducted on a single NVIDIA 3090 Ti GPU (24GB).

4.2 Downstream Prediction Analysis of the Synthetic Time Series

In coarse-grained economic time-series forecasting, scarce training data often cause forecasting baselines trained directly on the available set to overfit and underperform. To address this, we adopt the "training on synthetic, test on real" paradigm [Esteban *et al.*, 2017; Jordon *et al.*, 2018] and design two comparative studies to separately assess (i) the forecasting gains from using generated series and (ii) the quality of the generated series.

Table 1 reports results for three forecasting baselines—LSTM, LaST, and FITS—when trained on the original training set $\mathcal{D}_T$ versus on SI-CVAE–generated data $\mathcal{D}_G$. We evaluate with mean squared error (MSE) on a held-out real test set, fix the input window to 12 time steps, and consider prediction horizons $\{3, 6, 12\}$ (months). The table presents the mean MSE across horizons (full per-horizon results are in Appendix Tables C1 and C2). Models trained on $\mathcal{D}_G$ achieve substantially lower MSE than those trained on $\mathcal{D}_T$ with average improvements of 24.9% on the six ship-sales datasets and 36.5% on the four U.S. macroeconomic datasets.

Table 2 compares SI-CVAE against state-of-the-art time-series generators. We fix LaST as the downstream forecaster and hold its hyperparameters constant across conditions. For all generators, the synthesized sequence length is held fixed. Under data-scarce regimes, GAN-based generators exhibit higher training instability (*e.g.*, variance in discriminator–generator dynamics), yielding lower-quality synthetic series. Detailed results across different forecast horizons are reported in Appendix Tables C3. Relative to ImagenFew (the strongest baseline considered), SI-CVAE achieves average improvements of 20.2% on the six ship-sales datasets and 9.3% on the four U.S. macroeconomic datasets.

4.3 Representation Visualization

To gain a deeper understanding of the performance of the SI-CVAE model in the time series generation task, we designed two t-SNE-based representation visualization experiments on the Small Tanker Sales dataset. Figure 3(a) shows the latent vector distribution for each clustering category. From Figure 3(a), we can observe that the latent vectors for different categories form distinct clusters on the projection plane, with sample points tightly grouped, forming clear boundaries. This indicates that the model successfully captures the structural patterns within the time series and effectively categorizes them into corresponding clusters.

Figure 3(b) shows the distribution of synthetic and original data on the t-SNE projection plane to evaluate the realism and fitting ability of the generated data. From Figure 3(b), we observe that the distribution of synthetic data closely matches that of the original data, with significant overlap in their positions and shapes in the projection space. This consistency indicates that SI-CVAE effectively captures and reproduces the key features and patterns in the original data during time series generation.

4.4 Ablation Studies

To comprehensively evaluate the contribution and effectiveness of each module in the SI-CVAE framework, we conducted an extensive ablation study on six ship sales datasets across different ship types. The experimental results are shown in Table 3. Specifically, we first replaced the original variable-length subsequence clustering module with random partitioning of subsequences ($\text{SI-CVAE}_{\text{random}}$) to assess the role of this module in capturing inherent patterns within the time series. Next, we completely removed the variable-length subsequence clustering module, effectively setting the number of clusters K to 1, simplifying the model to a basic Fre-CVAE model ($\text{SI-CVAE}_{\text{Fre-CVAE}}$). In this configuration, the model loses its ability to capture diverse patterns, resulting in significant performance degradation. We then fixed the variable-length subsequence setting to a constant value by setting both $l_{\min}$ and $l_{\max}$ to 4, which reduces the model's flexibility in capturing variable-length patterns ($\text{SI-CVAE}_{l_{\min}=l_{\max}=4}$). Under this condition, the model loses flexibility in pattern capture, introducing more pattern bias. Furthermore, we conducted an ablation study on the conditional subsequence fusion module ($\text{SI-CVAE}_{\text{attention}-}$). By completely removing this module and using the mean of each category's reconstructed subsequence as the final generated sequence, i.e., $\hat{X} = \frac{1}{K}\sum_{k=1}^{K}\hat{S}^k$, we observed that the generated sequences lacked high-quality coherence and consistency, significantly impacting the downstream task performance. Finally, we removed the conditional prior model ($\text{SI-CVAE}_{p(z|c)-}$) and replaced it with a simple standard normal distribution $\mathcal{N}(0, 1)$. This operation substantially weakened the model's ability to capture complex prior information, leading to a significant decline in the diversity and accuracy of the generated sequences.

The comprehensive experimental results clearly demonstrate that all components of the SI-CVAE model are indispensable, with each module playing a crucial role in enhancing the model's performance.

4.5 Hyperparameter Experiments

To assess the sensitivity of SI-CVAE to key design choices and to understand their impact on both *generative fidelity*

Models	Small Tanker	VLCC	Suezmax Tanker	Aframax Tanker	Panamax Bulker	Handysize Bulker
SI-CVAE$_{\text{random}}$	0.74	3.98	0.93	3.48	1.92	1.16
SI-CVAE$_{\text{Fre-CVAE}}$	0.78	5.33	1.04	4.39	2.68	1.37
SI-CVAE$_{l_{\min}=l_{\max}=4}$	0.68	3.59	0.91	3.42	1.87	1.06
SI-CVAE$_{\text{attention}-}$	0.81	4.87	1.12	3.86	2.06	1.22
SI-CVAE$_{p(z\|c)-}$	0.71	3.77	1.01	3.44	1.93	1.11
SI-CVAE	**0.62**	**3.43**	**0.84**	**3.21**	**1.67**	**1.02**

Table 3: Contribution of Different Components to SI-CVAE Performance on the six ship-sales datasets.

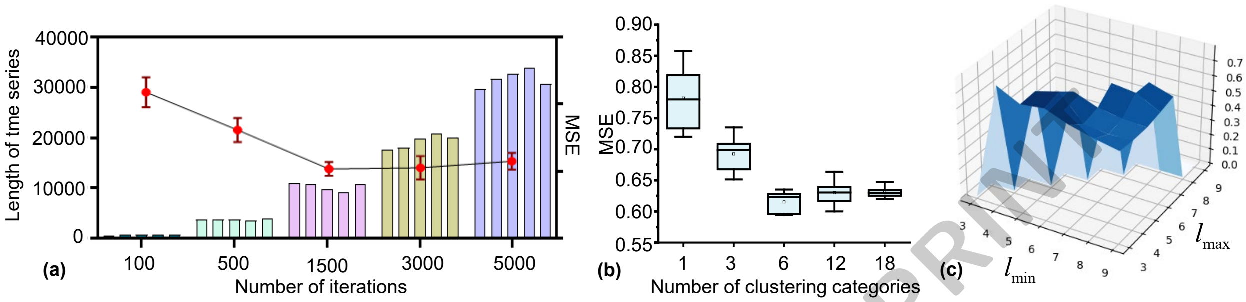


Figure 4: Impact of Hyperparameter Settings on Generated Sequence Quality and Model Performance. (a) The relationship between the number of iterations and both the length of generated time series (left axis) and the model's prediction performance (right axis, measured by MSE). (b) The effect of varying the number of clustering categories on the quality of generated sequences and prediction performance. (c) The impact of different minimum ($l_{\min}$) and maximum ($l_{\max}$) subsequence lengths on the model's mean squared error (MSE).

and *downstream utility*, we conduct controlled studies on the *Small Tanker Sales* series with input length 12 and prediction horizon 3. For each configuration, results are averaged over five random seeds. The main trends are summarized in Figure 4(a–c).

Generation iterations. We first vary the number of generation iterations in $\{100, 500, 1500, 3000, 5000\}$ to examine the trade-off between synthesis budget and forecasting gains. As shown in Figure 4(a), larger iteration budgets produce longer synthetic series, consistent with the fact that iterative synthesis controls the amount of generated data. In contrast, downstream forecasting performance exhibits a clear *inverted-U* pattern: it improves as iterations increase and peaks around 1500, but degrades slightly beyond that point. This suggests that overly large synthesis budgets may introduce redundant samples and accumulated noise, reducing the effective benefit of augmentation.

Number of clusters K. We then study the clustering granularity by sweeping $K \in \{1, 3, 6, 12, 18\}$. As reported in Figure 4(b), increasing K from small values yields substantial improvements, with the best performance attained around $K = 6$. When $K = 1$, SI-CVAE reduces to a single Fre-CVAE branch and thus cannot explicitly separate heterogeneous subsequence modes, leading to inferior generation and forecasting results. Further increasing K to 12 or 18 brings limited additional gains, indicating diminishing returns once the principal local patterns are sufficiently captured.

Variable-length range $[l_{\min}, l_{\max}]$. Finally, we evaluate the sensitivity to variable-length segmentation by testing $l_{\min}, l_{\max} \in \{3, 4, 5, 6, 7, 8, 9\}$ with the constraint $l_{\max} > l_{\min}$. As illustrated in Figure 4(c), the admissible length range materially affects performance: overly narrow ranges cause a notable drop, implying that fixed (or near-fixed) segmentation restricts the model's ability to represent multi-scale temporal structures in CGTS. Overall, these results provide empirical support for adopting variable-length subsequences, which offer a more expressive compositional basis for synthesizing realistic and diverse series.

5 Conclusion

This paper addresses the challenge of data scarcity in coarse-grained economic time series forecasting by proposing a Variable-Length Subsequence Scale-Invariant Conditional Generative Model (SI-CVAE) in conjunction with a data-centric approach. By incorporating subsequence scale invariance with frequency-domain mapping and decoding methods, SI-CVAE effectively tackles the challenges posed by data scarcity, noise interference, and the difficulty of capturing complex patterns in time series data. Extensive experimental results demonstrate that SI-CVAE significantly enhances the generalization ability of ship sales forecasting models on synthetic datasets, with data generation quality surpassing that of existing state-of-the-art methods. We also successfully addressed the data scarcity problem for this task through time series data generation strategies. In the future, we plan to explore the potential and value of SI-CVAE in more practical application scenarios.

Acknowledgments

This research was supported by the National Key R&D Program of China (No. 2023YFC3604205), the Key Research and Development Program of Liaoning Province (No. 2024JH2/102400073), the National Natural Science Foundation of China (No. 62172074), and the Program of Introducing Talents of Discipline to Universities (111 Project) (No. B20070).

References

[Brigham and Morrow, 1967] E Oran Brigham and RE Morrow. The fast fourier transform. *IEEE spectrum*, 4(12):63–70, 1967.

[Brophy *et al.*, 2023] Eoin Brophy, Zhengwei Wang, Qi She, and Tomás Ward. Generative adversarial networks in time series: A systematic literature review. *ACM Computing Surveys*, 55(10):1–31, 2023.

[Clarksons Research, 2025] Clarksons Research. Shipping Intelligence Network. Online database, https://sin.clarksons.net/, 2025. Data accessed in 2025.

[Desai *et al.*, 2021] Abhyuday Desai, Cynthia Freeman, Zuhui Wang, and Ian Beaver. Timevae: A variational auto-encoder for multivariate time series generation. *arXiv preprint arXiv:2111.08095*, 2021.

[Esteban *et al.*, 2017] Cristóbal Esteban, Stephanie L Hyland, and Gunnar Rätsch. Real-valued (medical) time series generation with recurrent conditional gans. *arXiv preprint arXiv:1706.02633*, 2017.

[Gonen *et al.*, 2025] Tal Gonen, Itai Pemper, Ilan Naiman, Nimrod Berman, and Omri Azencot. Time series generation under data scarcity: A unified generative modeling approach. *arXiv preprint arXiv:2505.20446*, 2025.

[Goodfellow *et al.*, 2014] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[Hochreiter and Schmidhuber, 1997] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[Imamura and Nakamura, 2024] Makoto Imamura and Takaaki Nakamura. Efficient discovery of time series motifs under both length differences and warping. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1188–1198, 2024.

[Jeon *et al.*, 2022] Jinsung Jeon, Jeonghak Kim, Haryong Song, Seunghyeon Cho, and Noseong Park. Gt-gan: General purpose time series synthesis with generative adversarial networks. *Advances in Neural Information Processing Systems*, 35:36999–37010, 2022.

[Jordon *et al.*, 2018] James Jordon, Jinsung Yoon, and Mihaela Van Der Schaar. Pate-gan: Generating synthetic data with differential privacy guarantees. In *International conference on learning representations*, 2018.

[Kingma and Welling, 2013] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[Naiman *et al.*, 2023] Ilan Naiman, N Benjamin Erichson, Pu Ren, Michael W Mahoney, and Omri Azencot. Generative modeling of regular and irregular time series data via koopman vaes. *arXiv preprint arXiv:2310.02619*, 2023.

[Roque *et al.*, 2024] Luis Roque, Carlos Soares, and Luís Torgo. Rhiots: A framework for evaluating hierarchical time series forecasting algorithms. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2491–2499, 2024.

[Sohl-Dickstein *et al.*, 2015] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pages 2256–2265. pmlr, 2015.

[Wang *et al.*, 2022] Zhiyuan Wang, Xovee Xu, Weifeng Zhang, Goce Trajcevski, Ting Zhong, and Fan Zhou. Learning latent seasonal-trend representations for time series forecasting. In *Advances in Neural Information Processing Systems*, 2022.

[Wind Information Co., Ltd., 2025] Wind Information Co., Ltd. Wind Financial Terminal / Wind Economic Database. Online database, https://www.wind.com.cn/, 2025. Data accessed in 2025.

[Xie *et al.*, 2025] Yongzheng Xie, Hongyu Zhang, and Muhammad Ali Babar. Multivariate time series anomaly detection by capturing coarse-grained intra-and inter-variate dependencies. In *Proceedings of the ACM on Web Conference 2025*, pages 697–705, 2025.

[Xu *et al.*, 2023] Zhijian Xu, Ailing Zeng, and Qiang Xu. Fits: Modeling time series with $10k$ parameters. *arXiv preprint arXiv:2307.03756*, 2023.

[Yin *et al.*, 2024] Mingjia Yin, Hao Wang, Wei Guo, Yong Liu, Suojuan Zhang, Sirui Zhao, Defu Lian, and Enhong Chen. Dataset regeneration for sequential recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3954–3965, 2024.

[Yoon *et al.*, 2019] Jinsung Yoon, Daniel Jarrett, and Mihaela Van der Schaar. Time-series generative adversarial networks. *Advances in neural information processing systems*, 32, 2019.

[Zhou *et al.*, 2021] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11106–11115, 2021.

[Zhu *et al.*, 2022] Tuanfei Zhu, Xinwang Liu, and En Zhu. Oversampling with reliably expanding minority class regions for imbalanced data learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):6167–6181, 2022.