Root Cause Analysis of Anomalies in MultiVariate Time Series through Granger Causal Discovery

Xiao Han, Saima Absar, Lu Zhang, Shuhan Yuan

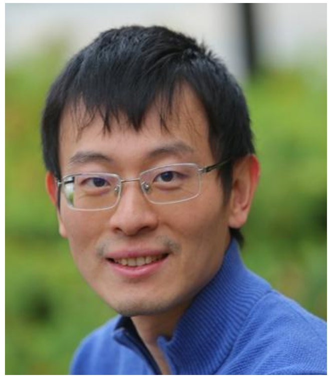
ICLR 2025 Oral

Scores: 8, 8, 8, 8, 8

Presented by Xin-Shuang Zhang

Authors

Xiao Han
Ph.D. Utah State University
Applied Scientist, Twitch



Shuhan Yuan
Assistant Professor
Utah State University

Supervisor

Saima Absar
Ph.D. University of Arkansas
Data Scientist,
Chevron Phillips Chemical Company

Lu Zhang
Associate Professor
University of Arkansas

Contribution

- **Learn the Granger causality among time series**
- **Explicitly model the distributions of exogenous variables** under normal conditions
- **Identify the root causes of anomalies** by highlighting exogenous variables that significantly deviate from their normal states
- **sota**

Granger Causality

- Granger causality is commonly used for modeling causal relationships in multivariate time series.
- Key assumption: If the prediction of the future value Y can be improved by knowing past elements of X, then X "Granger causes" Y.

Let a stationary time-series as $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_t, \ldots, \mathbf{x}_T)$, where $\mathbf{x}_t \in \mathbb{R}^d$ is a d-dimensional vector (e.g., d-dimensional time series data from d sensors) at a specific time t. Suppose that the true data generation mechanism is defined in the form of

$$\boxed{x_t^{(j)}} := \boxed{f^{(j)}(\mathbf{x}_{\leq t-1}^{(1)}, \cdots, \mathbf{x}_{\leq t-1}^{(d)})} + \boxed{u_t^{(j)}}, \text{ for } 1 \leq j \leq d, \quad (1)$$

where $\mathbf{x}_{\leq t-1}^{(j)} = [\cdots, x_{t-2}^{(j)}, x_{t-1}^{(j)}]$ denotes the past of series j; $u_t^{(j)} \in \mathbf{u}^{(j)}$ indicates exogenous variable for time series j at time step t; $f^{(j)}(\cdot)$ is a function for time series j that captures how the past values impact the future values of $\mathbf{x}^{(j)}$. The time series i Granger causes

Granger Causal Discovery

How to model the Granger causality? $\mathbf{x}_t := f(\mathbf{x}_{\leq t-1}) + \mathbf{u}_t$

- the causal relationships
- the distributions of exogenous variables

Motivation: Encoder-Decoder structure

- **Abductive** reasoning: Infer the most likely exogenous variables (causes) that could have generated the observed time series data:

$$\mathbf{u}_t := \mathbf{x}_t - \tilde{f}(\mathbf{x}_{\leq t-1})$$

- **Deductive** reasoning: Recursively resolve the previous time step, say x_{t-1},with their previous time step, i.e., x_{t-2} until the first time step:

$$\mathbf{x}_t = f(\mathbf{u}_{\leq t-1}) + \mathbf{u}_t$$

Encoder-Decorder Structure

we define a window with length K as $\mathbf{W}_t = (\mathbf{x}_{t-K+1}, ..., \mathbf{x}_t)$ and convert a time series $\mathbf{X}$ to a sequence of sliding windows $\mathcal{W} = (\mathbf{W}_K, \mathbf{W}_{K+1}, ..., \mathbf{W}_T)$.

Given a time series window, we first **parameterize** the Granger causality in time series:

$$\mathbf{x}_t := f(\mathbf{x}_{\leq t-1}) + \mathbf{u}_t \quad \Rightarrow \quad \mathbf{x}_t = \sum_{k=1}^{K} \boxed{\omega_{\theta_k}(\mathbf{x}_{t-k})} \mathbf{x}_{t-k} + \mathbf{u}_t$$

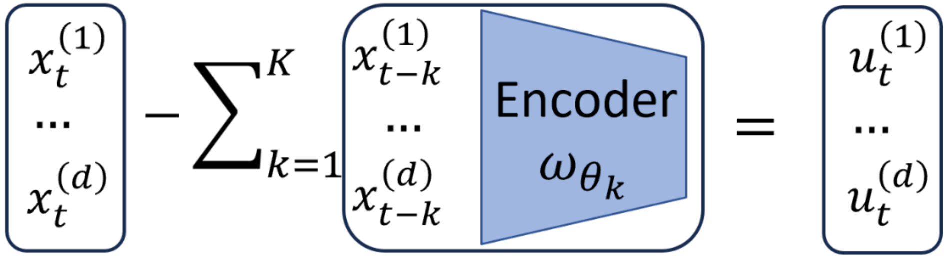
where $\omega_{\theta_k}(\mathbf{x}_{t-k})$ indicates the k-th neural network to predict the Granger causal relationship between $\mathbf{x}_{t-k}$ and $\mathbf{x}_t$. The output of $\omega_{\theta_k}(\mathbf{x}_{t-k})$ can be reshaped as a $d \times d$ coefficient matrix

***K* neural networks** are used to predict the weights of past K time legs on deriving $\mathbf{x}_t$

Encoder

Then, given a time series window $\mathbf{W}_t$, we apply the encoder K times to derive the exogenous variables in a window, denoted as $\mathbf{U}_t = (\mathbf{u}_{t-K+1}, \ldots, \mathbf{u}_t)$.

$$\mathbf{u}_t = \mathbf{x}_t - \sum_{k=1}^{K} \omega_{\theta_k}(\mathbf{x}_{t-k})\mathbf{x}_{t-k}$$



Decoder

Proposition 1. *Consider a basic autoregressive model where $\omega_k = \omega_{\theta_k}(\mathbf{x}_{t-k})$ as a framework for analyzing Granger causality. The value at the current time step $\mathbf{x}_t$ can be derived by the exogenous variables from a previous window $[\mathbf{u}_{t-1}, ..., \mathbf{u}_{t-K}]$ and the observed time series from a previous window $[\mathbf{x}_{t-K-1}, ..., \mathbf{x}_{t-2K}]$ with the following equation:*

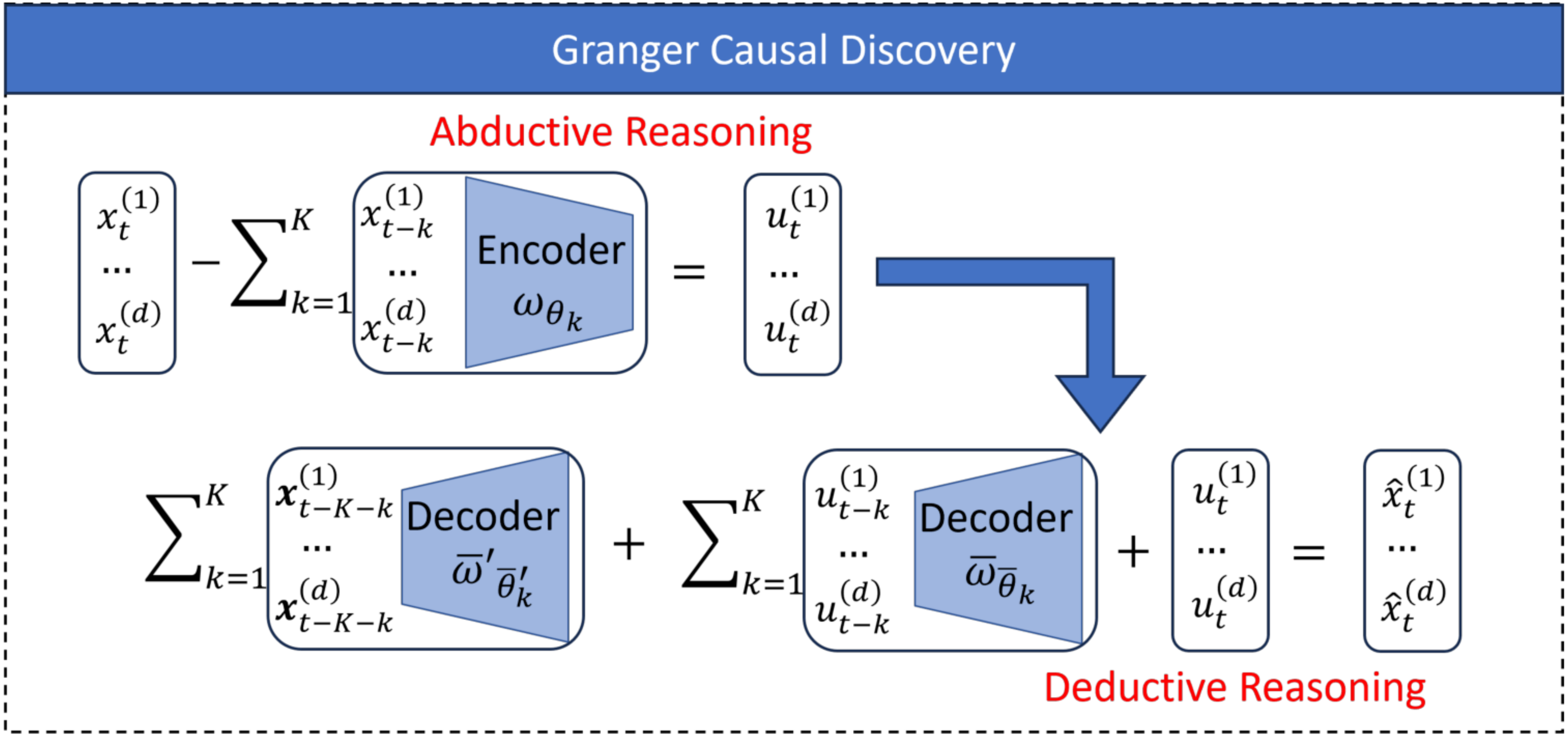
$$\boxed{\mathbf{x}_t} = \sum_{m=1}^{K} \alpha_{K-m} \boxed{\mathbf{u}_{t-(K-m)}} + \alpha_K \boxed{\mathbf{x}_{t-K}} + \sum_{m=2}^{K+1} \alpha_{K+1-m} \sum_{k=m}^{K} \omega_k \boxed{\mathbf{x}_{t-k-(K+1-m)}}, \tag{8}$$

where ω_k indicates the parameter of Granger causality, and $\alpha_n = \sum_{i=1}^{n} \omega_n \alpha_{n-i}$, $1 \leq n \leq K$, is a recursive equation with $\alpha_0 = 1$.

$$\hat{\mathbf{x}}_t = \sum_{k=1}^{K} \boxed{\bar{\omega}_{\bar{\theta}_k}} (\mathbf{u}_{t-k}) \mathbf{u}_{t-k} + \sum_{k=1}^{K} \boxed{\bar{\omega}'_{\bar{\theta}'_k}} (\mathbf{x}_{t-K-k}) \mathbf{x}_{t-K-k} + \mathbf{u}_t$$

where $\hat{\mathbf{x}}_t$ indicates the reconstructed value at time step t, and $\mathbf{u}_{t-k}$ is computed by encoder

Encoder-Decorder



Independence Constraint

To enforce independence between the derived exogenous variables, we ensure that the distribution of $\mathbf{U}_t$ adheres to an isotropic standard Gaussian distribution Q.

Applying the KL divergence to quantify the distribution difference:

$$\begin{aligned} D_t^{KL}(P(\mathbf{U}_t)\|Q) &= \frac{1}{2}\left(\mathrm{tr}(\Sigma_Q^{-1}\Sigma_t) + (\mu_Q - \mu_t)^T\Sigma_Q^{-1}(\mu_Q - \mu_t) - d + \log\frac{\det\Sigma_Q}{\det\Sigma_t}\right) \\ &= \frac{1}{2}\left(\mathrm{tr}\{\Sigma_t\} + \mu_t^T\mu_t - d - \log\det\Sigma_t\right), \end{aligned} \tag{7}$$

where $\mu_Q = 0$ and $\Sigma_Q = I$ represent the mean and covariance matrix of the isotropic standard Gaussian distribution Q; μ_t and Σ_t are the mean and covariance matrix of $\mathbf{U}_t$.

Loss Function

The whole encoder-decoder structure can be defined as $\hat{\mathbf{x}}_t = AE_{\theta_k,\bar{\theta}_k,\bar{\theta}'_k}(\mathbf{x}_{<t})$. Given a time series with length T, the objective function to train the encoder neural netework ω_{θ_k} and decoder neural networks $\bar{\omega}_{\bar{\theta}_k}$, $\omega'_{\bar{\theta}'_k}$ is defined as:

$$\begin{aligned}\mathcal{L} = & \sum_{t=K+1}^{T} \left\{ \boxed{\|\hat{\mathbf{x}}_t - \mathbf{x}_t\|_2} + \beta \boxed{D_t^{KL}} + \lambda_{en} \boxed{R(\Omega_t)} + \lambda_{de} \boxed{R(\bar{\Omega}_t)} + \lambda_{de} \boxed{R(\bar{\Omega}'_t)} \right\} \\ & + \sum_{t=K+1}^{T-1} \left\{ \gamma_{en} S(\Omega_{t+1}, \Omega_t) + \gamma_{de} S(\bar{\Omega}_{t+1}, \bar{\Omega}_t) + \gamma_{de} S(\bar{\Omega}'_{t+1}, \bar{\Omega}'_t) \right\},\end{aligned} \tag{10}$$

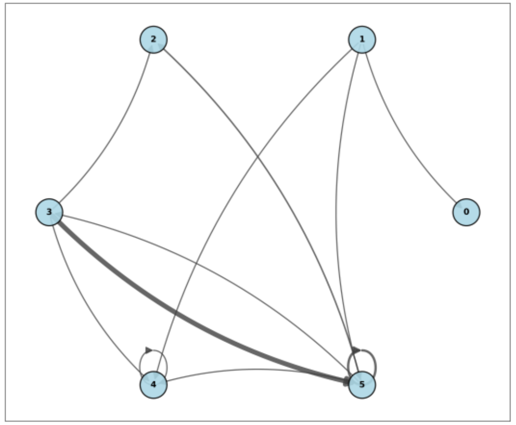
where D_t^{KL} indicates the independence constraint on $\mathbf{U}_t$ defined in Eq. 7; $\Omega_t := [\omega_{\theta_K}(\mathbf{x}_{t-K}) : \cdots : \omega_{\theta_1}(\mathbf{x}_{t-1})]$ indicates the concatenation of coefficient matrices over the past K time steps; similarly, we have $\bar{\Omega}_t := [\bar{\omega}_{\theta_K}(\mathbf{u}_{t-K}) : \cdots : \bar{\omega}_{\theta_1}(\mathbf{u}_{t-1})]$ and $\bar{\Omega}'_t := [\bar{\omega}'_{\theta'_K}(\mathbf{x}_{t-2K}) : \cdots : \bar{\omega}'_{\theta'_1}(\mathbf{x}_{t-K-1})]$; $R(\cdot)$ indicates the L1 and L2 norm penalty for sparsity of the coefficient matrices from the encoder and decoder; the $S(\cdot,\cdot)$ is a smoothness penalty, defined as $S(\Omega_{t+1}, \Omega_t) = \|\Omega_{t+1} - \Omega_t\|_2$; λ and γ are hyperparameters.

Causal Graph

Granger Causal Discovery. As the encoder-decoder is proposed to simulate the data generation process governed by Granger causality, we expect the function ω_{θ_k} can capture the causal relationships in time series. To further summarize the Granger causal relationships between variables as a summary causal graph, similar to (Marcinkevičs & Vogt, 2021), we aggregate the output from ω_{θ_k} into a summarized coefficient matrix as

$$S_{i,j} = \max_{1 \leq k \leq K} \{\text{median}_{K+1 \leq t \leq T}(\boxed{|(\omega_{\theta_k}(\mathbf{x}_{t-k}))_{i,j}|}), \text{ for } 1 \leq i, j \leq d,$$

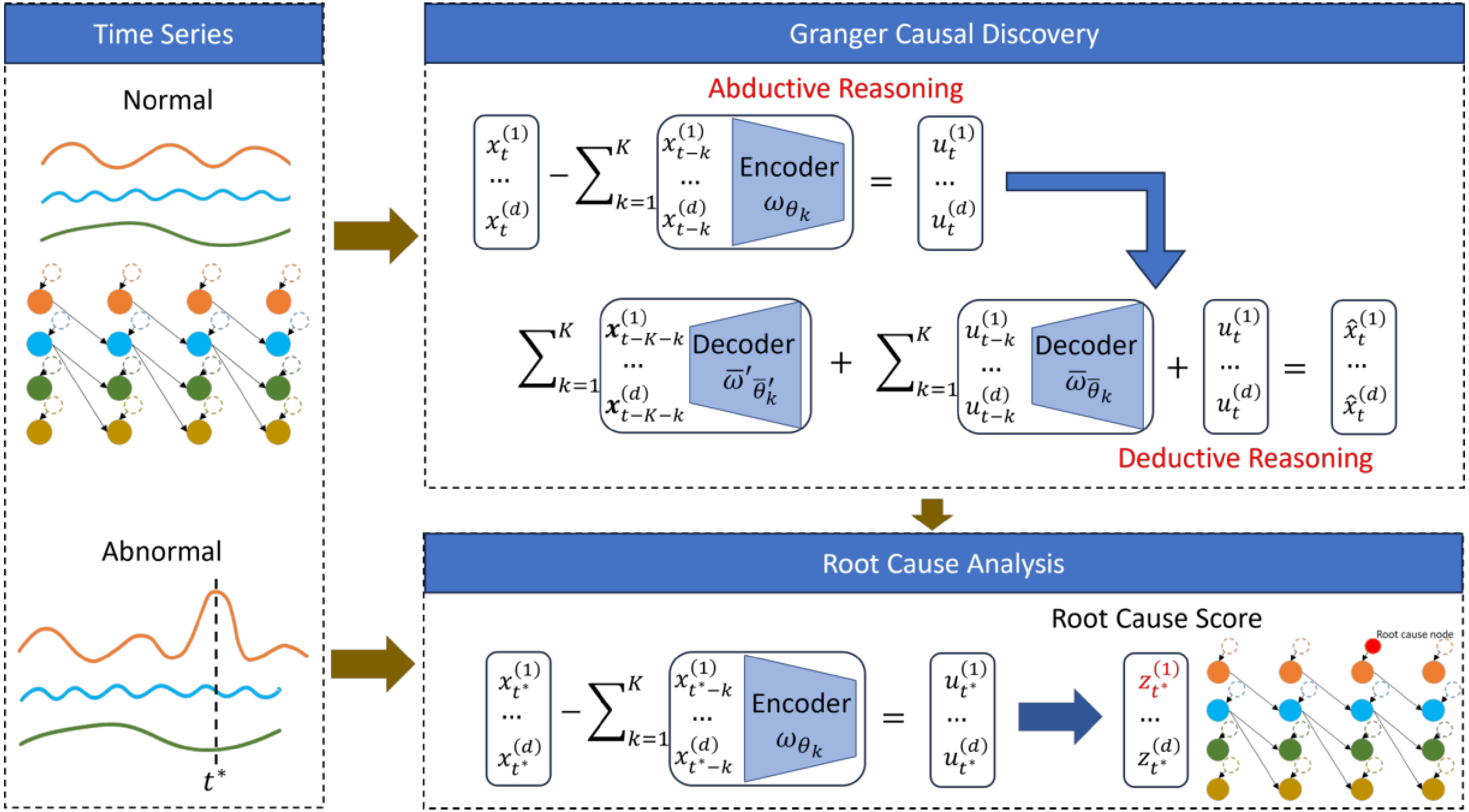
where $S_{i,j}$ indicates the strength of the Granger causal effect from $\mathbf{x}^{(i)}$ on $\mathbf{x}^{(j)}$. To further derive the adjacency matrix A, we set a threshold τ, if the value $S_{i,j} > \tau$, then $A_{i,j} = 1$. In experiments, the threshold is set based on the quantile of the coefficient matrix S.



Root Cause Localization

After training on the normal time series, we expect that the exogenous variables can be approximated by the encoder. When deploying the model for root cause localization, we assume the time series is arrived in a streaming manner. When a new time step t^* is arrived, we first adopt the encoder to derive the exogenous variables $\mathbf{u}_{t^*}$ based on Eq. 6. Then, for each time series, $u_{t^*}^{(j)}$, we compute the z-score as the root cause score $z_{t^*}^{(j)} = \frac{u_{t^*}^{(j)} - \mu^{(j)}}{\sigma^{(j)}}$, where $\mu^{(j)}$ and $\sigma^{(j)}$ indicate the mean and standard deviation of the exogenous variable for the j-th time series in normal data. We then adopt streaming peaks-over-threshold (SPOT) (Siffer et al., 2017) to dynamically determine the threshold of labeling the potential root cause.

The overview of AERCA



Dataset

Dataset	Training	Test		
	# of Time Step	# of Sequences ($\lvert\mathcal{X}\rvert$)	Avg. Len. ($\mathbf{T}$)	Avg. # of Root Variables
Linear (4)	5,000	100	500	3.75
Nonlinear (6)	5,000	100	500	5.25
Lotka-Volterra (40)	40,000	100	2,000	30.75
Lorenz 96 (20)	200,000	100	2,000	15.75
SWaT (51)	49,500	20	51	13.35
MSDS (10)	29,268	4,255	21	3.05

Overall performance (mean±std.) of causal discovery

Model	Linear				Nonlinear			
	F1	AUC-PR	AUC-ROC	HD	F1	AUC-PR	AUC-ROC	HD
VAR	$0.969_{\pm0.019}$	$0.998_{\pm0.003}$	$0.999_{\pm0.001}$	$0.011_{\pm0.009}$	$0.473_{\pm0.164}$	$0.529_{\pm0.181}$	$0.676_{\pm0.140}$	$0.258_{\pm0.130}$
cMLP	$0.745_{\pm0.029}$	$0.595_{\pm0.038}$	$0.829_{\pm0.0.25}$	$0.229_{\pm0.033}$	$0.419_{\pm0.134}$	$0.327_{\pm0.079}$	$0.609_{\pm0.089}$	$0.340_{\pm0.217}$
cLSTM	$0.684_{\pm0.042}$	$0.522_{\pm0.048}$	$0.766_{\pm0.047}$	$0.312_{\pm0.062}$	$0.378_{\pm0.000}$	$0.233_{\pm0.000}$	$0.500_{\pm0.000}$	$0.767_{\pm0.000}$
TCDF	$0.943_{\pm0.070}$	$0.933_{\pm0.081}$	$0.950_{\pm0.061}$	$0.033_{\pm0.040}$	$0.473_{\pm0.107}$	$0.343_{\pm0.072}$	$0.655_{\pm0.087}$	$0.307_{\pm0.065}$
eSRU	$0.964_{\pm0.070}$	$0.958_{\pm0.082}$	$0.969_{\pm0.061}$	$0.021_{\pm0.041}$	$0.408_{\pm0.152}$	$0.332_{\pm0.071}$	$0.617_{\pm0.092}$	$0.267_{\pm0.069}$
PCMCI	$0.969_{\pm0.031}$	$0.981_{\pm0.040}$	$0.986_{\pm0.042}$	$0.025_{\pm0.038}$	$0.607_{\pm0.094}$	$0.456_{\pm0.172}$	$0.742_{\pm0.147}$	$0.273_{\pm0.175}$
PCMCI+	$1.000_{\pm0.000}$	$1.000_{\pm0.000}$	$1.000_{\pm0.000}$	$0.000_{\pm0.000}$	$0.505_{\pm0.141}$	$0.410_{\pm0.133}$	$0.669_{\pm0.134}$	$0.233_{\pm0.109}$
GVAR	$0.862_{\pm0.052}$	$0.981_{\pm0.040}$	$0.986_{\pm0.042}$	$0.131_{\pm0.066}$	$0.421_{\pm0.094}$	$0.562_{\pm0.145}$	$0.683_{\pm0.097}$	$0.487_{\pm0.103}$
CUTS	$0.810_{\pm0.076}$	$0.792_{\pm0.066}$	$0.844_{\pm0.050}$	$0.104_{\pm0.034}$	$0.357_{\pm0.040}$	$0.249_{\pm0.014}$	$0.536_{\pm0.032}$	$0.513_{\pm0.124}$
AERCA	$\mathbf{1.000_{\pm0.000}}$	$\mathbf{1.000_{\pm0.000}}$	$\mathbf{1.000_{\pm0.000}}$	$\mathbf{0.000_{\pm0.000}}$	$\mathbf{0.826_{\pm0.057}}$	$\mathbf{0.996_{\pm0.013}}$	$\mathbf{0.998_{\pm0.006}}$	$\mathbf{0.027_{\pm0.014}}$

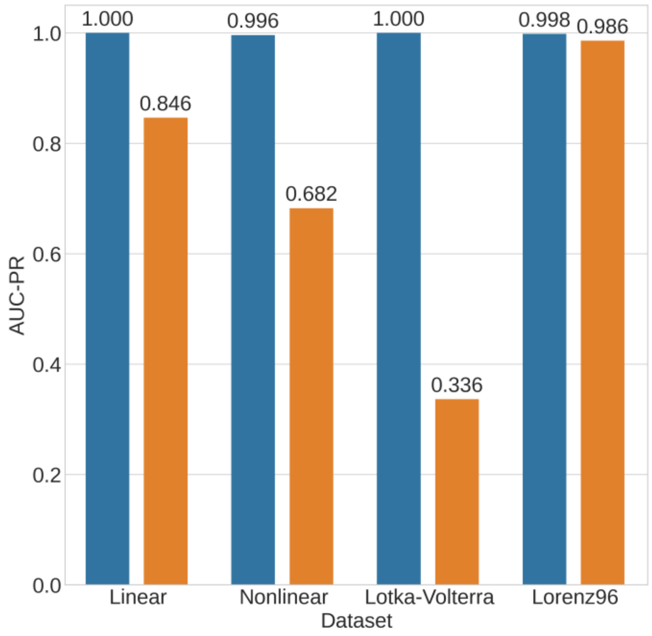
Model	Lotka-Volterra				Lorenz 96			
	F1	AUC-PR	AUC-ROC	HD	F1	AUC-PR	AUC-ROC	HD
VAR	$0.533_{\pm0.013}$	$\mathbf{1.000_{\pm0.000}}$	$\mathbf{1.000_{\pm0.000}}$	$0.044_{\pm0.003}$	$0.404_{\pm0.162}$	$0.562_{\pm0.376}$	$0.764_{\pm0.204}$	$0.360_{\pm0.121}$
cMLP	$0.511_{\pm0.011}$	$0.065_{\pm0.014}$	$0.508_{\pm0.007}$	$0.049_{\pm0.001}$	$0.472_{\pm0.058}$	$0.202_{\pm0.027}$	$0.569_{\pm0.038}$	$0.193_{\pm0.031}$
cLSTM	$0.356_{\pm0.176}$	$0.052_{\pm0.001}$	$0.500_{\pm0.000}$	$0.400_{\pm0.428}$	$0.453_{\pm0.048}$	$0.194_{\pm0.021}$	$0.572_{\pm0.031}$	$0.232_{\pm0.035}$
TCDF	$0.853_{\pm0.032}$	$0.749_{\pm0.050}$	$0.890_{\pm0.021}$	$0.019_{\pm0.002}$	$0.429_{\pm0.007}$	$0.290_{\pm0.006}$	$0.645_{\pm0.004}$	$0.260_{\pm0.011}$
eSRU	$0.422_{\pm0.039}$	$0.323_{\pm0.030}$	$0.634_{\pm0.016}$	$0.055_{\pm0.002}$	$0.195_{\pm0.024}$	$0.225_{\pm0.009}$	$0.539_{\pm0.009}$	$0.215_{\pm0.006}$
PCMCI	$0.465_{\pm0.025}$	$0.291_{\pm0.019}$	$0.906_{\pm0.017}$	$0.109_{\pm0.008}$	$0.368_{\pm0.004}$	$0.227_{\pm0.007}$	$0.680_{\pm0.013}$	$0.540_{\pm0.021}$
PCMCI+	$0.709_{\pm0.027}$	$0.651_{\pm0.121}$	$0.851_{\pm0.082}$	$0.024_{\pm0.005}$	$0.502_{\pm0.020}$	$0.329_{\pm0.022}$	$0.709_{\pm0.017}$	$0.163_{\pm0.009}$
GVAR	$0.787_{\pm0.011}$	$0.988_{\pm0.015}$	$0.999_{\pm0.002}$	$0.027_{\pm0.002}$	$0.568_{\pm0.330}$	$0.582_{\pm0.361}$	$0.776_{\pm0.194}$	$0.142_{\pm0.109}$
CUTS	$\mathbf{0.877_{0.031}}$	$0.791_{\pm0.047}$	$0.892_{\pm0.024}$	$\mathbf{0.011_{\pm0.002}}$	$0.341_{\pm0.003}$	$0.206_{\pm0.002}$	$0.621_{\pm0.004}$	$0.404_{\pm0.012}$
AERCA	$0.857_{\pm0.000}$	$\mathbf{1.000_{\pm0.000}}$	$\mathbf{1.000_{\pm0.000}}$	$0.026_{\pm0.000}$	$\mathbf{0.800_{\pm0.000}}$	$\mathbf{0.998_{\pm0.002}}$	$\mathbf{0.999_{\pm0.001}}$	$\mathbf{0.105_{\pm0.000}}$

Overall performance (mean±std.) of root cause analysis

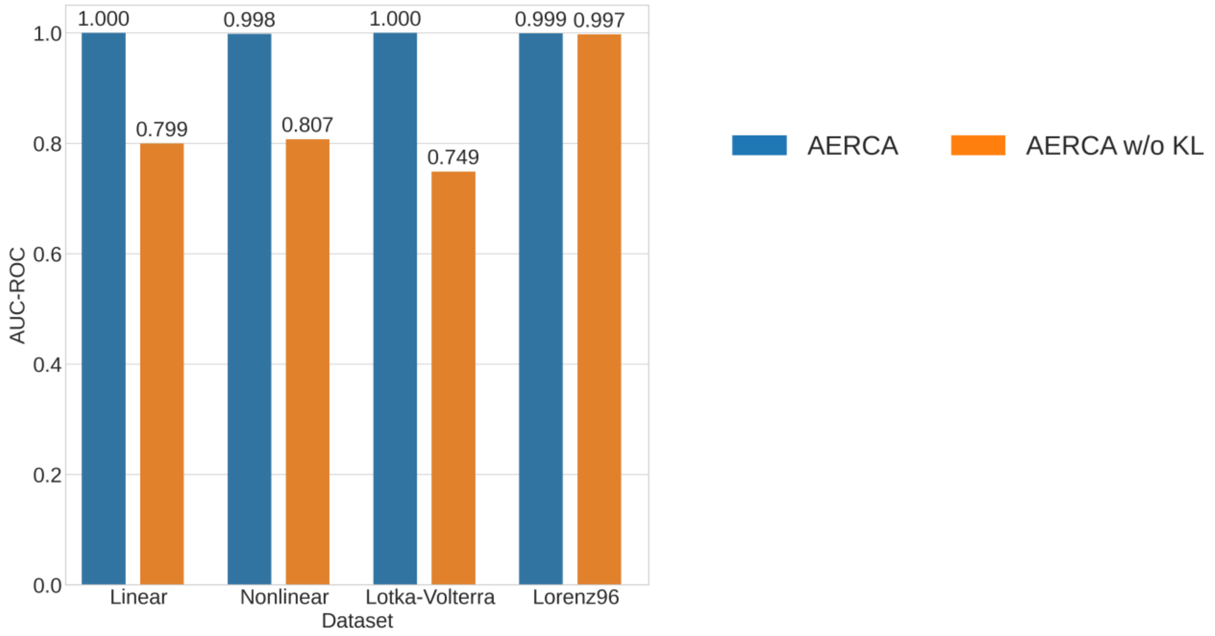
Dataset	Model	AC@1	AC@3	AC@5	AC@10	Avg@10
Linear	ϵ-Diagnosis	$0.900_{\pm 0.300}$	$0.850_{\pm 0.189}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.950_{\pm 0.043}$
	RCD	$0.500_{\pm 0.500}$	$0.817_{\pm 0.189}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.907_{\pm 0.076}$
	CIRCA	$0.600_{\pm 0.490}$	$0.800_{\pm 0.306}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.910_{\pm 0.106}$
	AERCA	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$
Nonlinear	ϵ-Diagnosis	$0.400_{\pm 0.490}$	$0.667_{\pm 0.325}$	$0.880_{\pm 0.165}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.837_{\pm 0.139}$
	RCD	$0.600_{\pm 0.490}$	$0.750_{\pm 0.344}$	$0.880_{\pm 0.165}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.878_{\pm 0.118}$
	CIRCA	$0.700_{\pm 0.458}$	$0.717_{\pm 0.395}$	$0.835_{\pm 0.295}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.863_{\pm 0.160}$
	AERCA	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$
Lotka-Volterra	ϵ-Diagnosis	$0.100_{\pm 0.300}$	$0.133_{\pm 0.163}$	$0.138_{\pm 0.149}$	$0.247_{\pm 0.188}$	$0.158_{\pm 0.131}$
	RCD	$0.100_{\pm 0.300}$	$0.133_{\pm 0.163}$	$0.138_{\pm 0.149}$	$0.247_{\pm 0.188}$	$0.158_{\pm 0.131}$
	CIRCA	$0.120_{\pm 0.325}$	$0.107_{\pm 0.169}$	$0.120_{\pm 0.150}$	$0.225_{\pm 0.230}$	$0.146_{\pm 0.163}$
	AERCA	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$
Lorenz96	ϵ-Diagnosis	$0.100_{\pm 0.300}$	$0.200_{\pm 0.221}$	$0.280_{\pm 0.312}$	$0.450_{\pm 0.330}$	$0.314_{\pm 0.225}$
	RCD	$0.200_{\pm 0.400}$	$0.333_{\pm 0.333}$	$0.400_{\pm 0.358}$	$0.556_{\pm 0.337}$	$0.421_{\pm 0.278}$
	CIRCA	$0.360_{\pm 0.480}$	$0.330_{\pm 0.244}$	$0.346_{\pm 0.249}$	$0.539_{\pm 0.263}$	$0.408_{\pm 0.220}$
	AERCA	$\mathbf{0.996}_{\pm \mathbf{0.009}}$	$\mathbf{0.996}_{\pm \mathbf{0.009}}$	$\mathbf{0.997}_{\pm \mathbf{0.008}}$	$\mathbf{0.996}_{\pm \mathbf{0.008}}$	$\mathbf{0.990}_{\pm \mathbf{0.011}}$
SWaT	ϵ-Diagnosis	$0.075_{\pm 0.179}$	$0.125_{\pm 0.217}$	$0.125_{\pm 0.217}$	$0.375_{\pm 0.383}$	$0.180_{\pm 0.194}$
	RCD	$0.000_{\pm 0.000}$	$0.000_{\pm 0.000}$	$0.000_{\pm 0.000}$	$0.300_{\pm 0.458}$	$0.100_{\pm 0.161}$
	CIRCA	$0.000_{\pm 0.000}$	$0.000_{\pm 0.000}$	$0.000_{\pm 0.000}$	$0.300_{\pm 0.458}$	$0.100_{\pm 0.161}$
	AERCA	$\mathbf{0.220}_{\pm \mathbf{0.111}}$	$\mathbf{0.290}_{\pm \mathbf{0.088}}$	$\mathbf{0.330}_{\pm \mathbf{0.048}}$	$\mathbf{0.455}_{\pm \mathbf{0.044}}$	$\mathbf{0.342}_{\pm \mathbf{0.052}}$
MSDS	ϵ-Diagnosis	$0.004_{\pm 0.004}$	$0.266_{\pm 0.002}$	$0.452_{\pm 0.009}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.492_{\pm 0.001}$
	RCD	$0.412_{\pm 0.048}$	$0.573_{\pm 0.010}$	$\mathbf{0.984}_{\pm \mathbf{0.001}}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.821_{\pm 0.012}$
	CIRCA	$\mathbf{0.454}_{\pm \mathbf{0.238}}$	$0.860_{\pm 0.140}$	$0.917_{\pm 0.084}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$0.809_{\pm 0.035}$
	AERCA	$0.381_{\pm 0.408}$	$\mathbf{0.908}_{\pm \mathbf{0.062}}$	$0.974_{\pm 0.027}$	$\mathbf{1.000}_{\pm \mathbf{0.000}}$	$\mathbf{0.896}_{\pm \mathbf{0.037}}$

Ablation Study

To show the importance of the **independent constraint** for causal discovery, we conduct the ablation study to compare the performance of causal discovery when AERCA is trained with and without the independent constraint in the objective function.



(a) AUC-PR



(b) AUC-ROC

Improvement

1. 因果图：对自编码器的超参数比较敏感
2. 异常检测：用其他算法，比如随机森林
3. 知识如何体现：

（1）损失函数（例如体现周期性）

（2）KL 约束

（3）因果图不完备（已知某些变量之前强相关）

训练集: X_train形状 (45969, 6), y_train形状 (45969,)
测试集: X_test形状 (11493, 6), y_test形状 (11493,)
训练集异常样本比例: 49.82%
测试集异常样本比例: 50.72%
测试集准确率: 0.9685

分类报告:

	precision	recall	f1-score	support
0.0	0.97	0.96	0.97	5664
1.0	0.97	0.97	0.97	5829
accuracy			0.97	11493
macro avg	0.97	0.97	0.97	11493
weighted avg	0.97	0.97	0.97	11493

对建模而得的外生变量 u 的分布进行随机森林异常检测