

基于相关熵核学习演化模糊系统的混沌时间序列在线预测

胡磊¹ 许星晗² 刘建卫² 韩敏¹

摘要 演化模糊系统 (EFSs) 是在线学习领域中广泛应用的方法. 然而, 大部分 EFSs 往往隐含高斯噪声假设, 在重尾、偏态及强相关等非高斯扰动下易出现性能退化. 为此, 提出一种相关熵核学习演化模糊系统 (CKL-EFS). CKL-EFS 融合结构自组织与鲁棒递归学习两种机制, 以提升非高斯噪声环境下的在线建模能力. 在结构演化方面, 模型基于数据密度与激活度实现数据云的按需生成与效用驱动移除, 并通过在线归属更新维护数据云统计量, 从而抑制规则库无界膨胀并保持紧凑结构; 在参数更新方面, 采用核学习策略将输入映射至高维特征空间, 并引入改进相关熵准则对大误差样本自适应降权, 以增强对离群冲击与非高斯噪声的鲁棒性. 此外, 给出模型的计算复杂度分析. 在多个公开数据集上的实验结果表明, 所提 CKL-EFS 在预测精度与抗非高斯扰动稳定性方面均优于现有代表性方法.

关键词 演化模糊系统; 相关熵; 核学习策略; 混沌时间序列; 在线预测

引用格式 胡磊, 许星晗, 刘建卫, 韩敏. 基于相关熵核学习演化模糊系统的混沌时间序列在线预测. 自动化学报, 2026, 52(4): 724-737

DOI 10.16383/j.aas.c250343 **CSTR** 32138.14.j.aas.c250343

Online Prediction of Chaotic Time Series Based on Correntropy Kernel Learning Evolving Fuzzy Systems

HU Lei¹ XU Xing-Han² LIU Jian-Wei² HAN Min¹

Abstract Evolving fuzzy systems (EFSs) have been widely used in online learning. However, most existing EFSs implicitly rely on a Gaussian-noise assumption and may suffer performance degradation under non-Gaussian disturbances such as heavy-tailed, skewed, and strongly correlated noise. To address this issue, this paper proposes a correntropy kernel learning EFSs (CKL-EFS). CKL-EFS integrates structural self-organization and robust recursive learning mechanisms to enhance online modeling capability in non-Gaussian noise environments. For structural evolution, the model performs on-demand data-cloud generation and utility-driven pruning based on data density and activation, and maintains data-cloud statistics via online assignment updates, thereby preventing unbounded rule-base growth and preserving a compact structure; For parameter updating, a kernel learning strategy is employed to map inputs into a high-dimensional feature space, and an improved correntropy criterion is introduced to adaptively down-weight large-error samples, improving robustness against outlier perturbations and non-Gaussian noise. In addition, the computational complexity of the proposed model is analyzed. Experimental results on multiple public datasets demonstrate that CKL-EFS consistently outperforms representative existing methods in both predictive accuracy and stability against non-Gaussian disturbances.

Keywords evolving fuzzy systems; correntropy; kernel learning strategy; chaotic time series; online prediction

Citation Hu Lei, Xu Xing-Han, Liu Jian-Wei, Han Min. Online prediction of chaotic time series based on correntropy kernel learning evolving fuzzy systems. *Acta Automatica Sinica*, 2026, 52(4): 724-737

混沌时间序列广泛存在于天文^[1]、水文^[2]、电力^[3]和金融^[4]等领域中. 在当代社会, 基于数据的混沌

时间序列预测已成为计算机科学领域的研究重点^[5]. 然而, 现实中提取的混沌时间序列通常表现出较强的非线性和噪声, 这对传统的预测方法提出挑战^[6]. 随着信息技术和人工智能的不断发展^[7-8], 各种数据挖掘技术为混沌时间序列的分析提供新视角^[9-10]. 作为数据挖掘中的关键技术, 演化模糊系统 (evolving fuzzy systems, EFSs) 在处理复杂时间序列方面具有较强的竞争力^[11-12]. EFSs 的有效性已经在多个实际应用中得到验证^[13]. 相比其他预测方法, EFSs 推理和决策的过程类似于人类认知^[14], 具备在线更新、结构可调和结果可解释等优势^[15-16]. 现有工

收稿日期 2025-07-23 录用日期 2026-01-15
Manuscript received July 23, 2025; accepted January 15, 2026
国家自然科学基金 (62173063, 52309079) 资助
Supported by National Natural Science Foundation of China (62173063, 52309079)

本文责任编辑 王卓

Recommended by Associate Editor WANG Zhuo

1. 大连理工大学控制科学与工程学院 大连 116024 2. 大连理工大学建设工程学院 大连 116024

1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024 2. School of Infrastructure Engineering, Dalian University of Technology, Dalian 116024

作大致可分为三类. 第一类是典型 EFSs, 如基于经验数据分析的自主学习多模态系统 (autonomous learning multi-model systems, ALMMo)^[17], 能在数据漂移时自动生成或调整规则, 并通过递归最小二乘 (recursive least squares, RLS) 更新后件参数; Hu 等^[18] 通过多种群信息集成粒子群优化 (multi-population information integrated particle swarm optimization, MIIPSO) 同时调整前后件参数, 进一步提升模型自适应性能. 但这类方法大多仍以均方误差 (mean square error, MSE)^[19-20] 为主优化目标, 对强非高斯或脉冲噪声仍然敏感. 第二类工作将核学习或多层结构引入 EFSs, 例如多层集成演化模糊推理系统 (multilayer ensemble evolving fuzzy inference system, MEEFIS)^[21] 与基于核递归最小二乘和距离相关性的演化参与式学习 (evolving participatory learning with kernel recursive least square and distance correlation, ePL-KRLS-DISCO) 方法^[22], 显著增强了非线性表征能力, 但仍使用 MSE 或无鲁棒性误差度量, 在遇到离散化、结构性非高斯噪声时, 预测稳定性仍会下降^[23-24]. 第三类工作尝试把最大相关熵准则 (maximum correntropy criterion, MCC)^[25]、最大总相关熵 (maximum total correntropy, MTC)^[26] 等信息论准则^[27-28] 引入 EFSs, 如基于相关熵的演化模糊神经网络 (correntropy-based evolving fuzzy neural system, CEFNS)^[29] 和基于演化参与式学习的核一般损失算法 (kernel general loss algorithm based on evolving participatory learning, EPL-KGLA)^[30], 能在一定程度上抑制离群点影响^[31], 但多为单尺度或无梯度约束的相关熵形式, 且与在线核方法结合不紧密, 因而在强非线性和强非高斯场景下仍有改进空间. 现实世界中的混沌序列往往同时受到高斯噪声与非高斯噪声的干扰, 常表现为厚尾特性、间歇性离群点以及统计结构时变等复杂特征^[32-33]. 在此类环境下, 仅依赖二阶统计量的 MSE 难以兼顾精度与鲁棒性^[34]. 为此, 本文基于“EFSs + 核学习 + 相关熵”融合框架进行统一建模, 主要贡献包括:

1) 提出相关熵核学习演化模糊系统 (correntropy kernel learning EFSs, CKL-EFS). 该系统在 EFSs 的递归学习过程中引入相关熵驱动的鲁棒目标, 并与在线核递归更新机制紧密耦合, 从而在非平稳、非高斯扰动下实现稳定在线更新. 相比基于 MSE 的 EFSs, CKL-EFS 显著提升了对离群冲击与重尾噪声的抗扰能力.

2) 设计面向非高斯数据流的密度感知结构自

维护机制. 模型通过全局密度新颖性检测与局部覆盖门控识别新分布/稀疏区域并触发按需增生; 对于已覆盖样本, 采用在线归属更新 (样本归并到最相似数据云并更新其中心与统计量) 以自适应跟踪分布漂移; 同时引入效用驱动的规则移除策略, 自动剔除长期低贡献的数据云, 从而抑制规则库无界膨胀并保持紧凑可部署的结构复杂度.

3) 给出复杂度分析与多噪声场景下的实证验证. 本文分析模糊推理与核递推更新的时间复杂度上界, 并在 Lorenz-96 混沌系统、气象、水文及风电等多类数据集 (涵盖不同强度与类型的非高斯扰动) 上开展对比实验. 结果表明, CKL-EFS 在保证在线效率的同时取得更低预测误差, 且在非高斯性增强时呈现更小的性能退化幅度.

本文其他部分组织如下: 第 1 节给出相关预备知识; 第 2 节介绍所提 CKL-EFS 的结构与学习算法; 第 3 节分析计算复杂度; 第 4 节给出实验结果及讨论; 第 5 节为结论与展望.

1 相关理论

1.1 相关熵

线性系统可表示为 $y_n = w^* x_n^T$, 其中 w^* 是最优权重, $x_n \in \mathbf{R}^{P \times 1}$ 表示输入 (P 为输入变量维度), $y_n \in \mathbf{R}^{1 \times 1}$ 表示输出. 误差变量 (errors-in-variables, EIV) 系统^[34] 考虑了信号中的测量误差, 描述为:

$$\tilde{x}_n = x_n + u_n, \quad \tilde{y}_n = y_n + v_n \quad (1)$$

其中, \tilde{x}_n 和 \tilde{y}_n 分别表示被噪声污染的输入和输出观测值; u_n 和 v_n 表示输入和输出的噪声分量, 其协方差矩阵分别表示为 $\sigma_{\text{in}}^2 I_{P \times P}$ 和 $\sigma_{\text{out}}^2 I_{1 \times 1}$ (其中 $I_{P \times P}$ 是单位矩阵, σ_{in}^2 和 σ_{out}^2 分别表示输入噪声和输出噪声的方差), 如文献^[35] 所述. u_n 和 v_n 通常情况下是广义平稳的且彼此不相关^[36]. EIV 系统的目标是从 $\{\tilde{x}_n, \tilde{y}_n\}$ 中估计 w^* .

作为 MTC^[26] 的改进方案, 最大总广义相关熵 (maximum total generalized correntropy, MTGC)^[34] 被设计用于 EIV 系统. MTGC 准则在 MTC 基础上引入形状参数 α 和尺度参数 β , 包含更多非二次特征, 能进一步增强对非高斯环境的适应性. 基于 MTGC, w^* 可通过最大化式 (2) 计算得到.

$$J_{\text{MTGC}}(w) = G_{\alpha, \beta}(e'_n) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})} \exp\left(-\frac{|e'_n|^\alpha}{\beta^\alpha}\right) \quad (2)$$

其中, $e'_n = e_n / \sqrt{\bar{w}_n^T \bar{w}_n}$, $e_n = \tilde{y}_n - w_n^T \tilde{x}_n$ 为误差, w_n 是权值, $\bar{w}_n = [\sqrt{\gamma} \quad -w_n^T]$ 为修正后的增广权值, γ 是

$\sigma_{\text{out}}^2/\sigma_{\text{in}}^2$ 表示输出与输入噪声方差的比值; $G_{\alpha, \beta}(\cdot)$ 表示由 α 与 β 控制的广义高斯核函数; $\Gamma(\cdot)$ 表示 Gamma 函数. 式 (2) 的偏导数可以表示为:

$$\hat{\nabla} J_{\text{MTGC}}(w) = \lambda g(w_n) t(w_n) \quad (3)$$

其中, $\lambda = \alpha^2/(2\beta^{\alpha+1}\Gamma(1/\alpha))$; $g(w_n) = \exp(-|e_n|^\alpha/(\beta\|\bar{w}_n\|^\alpha))$; $t(w_n) = (|e_n|^{\alpha-1}S(e_n)\|\bar{w}_n\|^\alpha\bar{x}_n + |e_n|^\alpha \times \|\bar{w}_n\|^{\alpha-2}w_n)/\|\bar{w}_n\|^{2\alpha}$, $S(\cdot)$ 的函数式为:

$$S(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \quad (4)$$

基于式 (3), 设置 μ 表示步长, 权值更新可表示为:

$$w_{n+1} = w_n + \mu \lambda g(w_n) t(w_n) \quad (5)$$

与传统相关熵仅在单一尺度上抑制大残差不同, MTGC 准则采用总相关熵与广义核联合建模策略. 该策略在保留正常样本分辨率的同时, 能对重尾脉冲型非高斯噪声产生强抑制作用, 从而在非高斯环境下表现出更优的鲁棒性. 相较于仅对中等离群点敏感的 MCC^[25] 和主要进行总量抑制的 MTC 准则^[26], MTGC 准则能在同一目标中抑制异常且不影响主流样本, 更适用于在线模型的鲁棒学习.

1.2 演化模糊系统

EFSs 的通用数学结构可表示为^[17, 37]:

$$\mathcal{R}_i: \text{如果 } \underbrace{(x_1 \text{ 属于 } \mathcal{C}_i) \text{ 且 } \cdots \text{ 且 } (x_P \text{ 属于 } \mathcal{C}_i)}_{\text{前件部分}} \quad (6)$$

$$\text{则 } \underbrace{y_i = a_{i,0} + a_{i,1}x_1 + \cdots + a_{i,P}x_P}_{\text{后件部分}}$$

其中, \mathcal{R}_i 表示第 i 条模糊规则; \mathcal{C}_i 表示第 i 个数据云; $\mathbf{x} = [x_1, x_2, \cdots, x_P]^T$ 表示系统输入; $a_{i,p}$ ($p = 0, 1, \cdots, P$) 表示后件部分的权值向量; y_i 表示 \mathcal{C}_i 的输出. 如式 (6) 所示, 前件部分生成模糊规则并进行样本聚类, 通过模糊规则聚类成的各个数据云表示不同的线性子系统; 后件部分计算各个数据云的局部输出, 即每个规则控制下线性子系统的输出, 并根据各个规则的激活程度加权到局部输出, 最后得到整个 EFSs 的输出.

2 相关熵核学习演化模糊系统

本节首先介绍 CKL-EFS 基本原理, 其次解释学习机制, 最后提供完整算法.

2.1 模型结构

CKL-EFS 的内部工作流程如图 1 所示. CKL-

EFS 主要包括两个学习机制: 结构演化机制和参数更新机制. 在结构演化机制中, 首先通过全局/局部数据密度对输入样本进行在线分析与自组织聚类, 形成由数据云驱动的模糊规则库. 规则库采用增量式演化策略: 每次输入样本时, 会将其与已有数据云进行匹配并计算其局部密度与激活度. 当样本落入现有数据云的覆盖范围内时, 系统将其归并到最相容的数据云, 并在线更新该数据云的中心与统计量, 以自适应跟踪分布漂移; 当样本表现出明显的新颖性 (例如其全局密度超出当前数据云密度范围且局部覆盖不足) 时, 系统按需生成新的数据云以刻画新的运行模式. 同时, 为避免规则库无界膨胀, 模型引入效用驱动的自维护策略: 对长期低激活、低贡献的数据云进行在线剔除. 上述“按需增生-在线归并更新-效用剪枝”的闭环机制使规则库保持紧凑, 从而控制模型的计算复杂度与存储开销. 在参数更新机制中, 将核学习策略融入基于 MTGC 准则的损失函数中, 设计新的权值递归更新方式, 并应用到存储在各个数据云中的样本子集. 该机制中, 核学习策略通过高斯核函数将样本点从低维空间映射至高维空间, 增强了高维表达能力, 促进了模型对复杂非线性样本的学习能力. 最后, 选择最相容数据云的局部输出作为整体输出, 其计算方法如下:

$$\mathcal{R}_i: \text{如果 } \underbrace{\mathbf{u}_k \text{ 属于 } \mathcal{C}_i}_{\text{前件部分}} \text{ 则 } \underbrace{\hat{d}_k = f_i(\mathbf{u}_k, \boldsymbol{\theta}(k))}_{\text{后件部分}} \quad (7)$$

其中, k 表示当前样本序号; \mathbf{u}_k 表示样本输入; $\boldsymbol{\theta}(k)$ 表示权值矩阵; \hat{d}_k 表示样本输出 d_k 对应的模型计算输出; i 表示与 \mathbf{u}_k 相距最近数据云的序号; $f_i(\cdot, \cdot)$ 为数据云 \mathcal{C}_i 的后件学习器内置函数.

2.2 结构演化机制

2.2.1 数据云的初始化

接收到第 1 个样本 (\mathbf{u}_1, d_1) 后, 启动初始化过程, 得到第 1 个数据云. 数据云的初始化涉及全局参数和局部参数的初始化. 全局参数包括接收到输入的均值和 L_2 范数平方, 表示为:

$$\boldsymbol{\mu} \leftarrow \mathbf{u}_k, \quad \mathbf{X} \leftarrow \|\mathbf{u}_k\|_2^2, \quad k \leftarrow 1 \quad (8)$$

其中, $\|\mathbf{u}_k\|_2$ 表示 \mathbf{u}_k 的 L_2 范数; $\boldsymbol{\mu}$ 和 \mathbf{X} 分别表示样本均值和 L_2 范数平方. 局部参数主要为数据云内样本的均值、 L_2 范数平方和样本数. 每个数据云由对应规则 \mathcal{R}_i 控制, 以调节其内部样本的增长. 第 L 个数据云的局部参数初始化可表示为:

$$\boldsymbol{\mu}_L \leftarrow \mathbf{u}_k, \quad \mathbf{X}_L \leftarrow \|\mathbf{u}_k\|_2^2, \quad \mathcal{C}_L \leftarrow \{\mathbf{u}_k\}, \quad N_L \leftarrow 1 \quad (9)$$

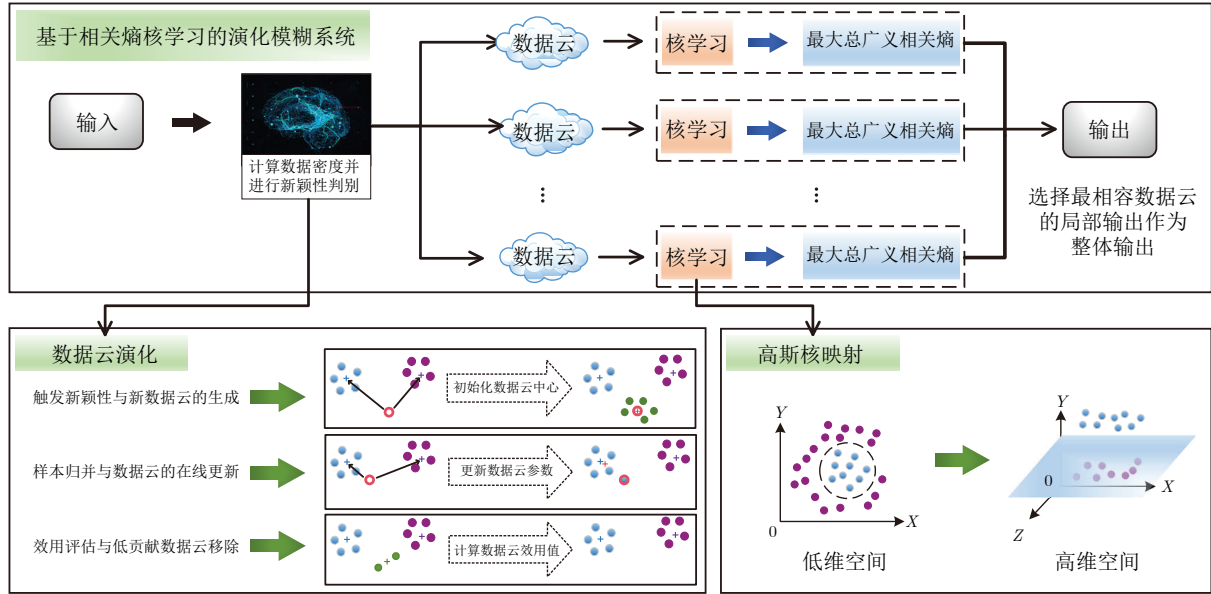


图1 CKL-EFS 结构及内部驱动原理

Fig.1 CKL-EFS architecture and internal driving principles

其中, \mathcal{C}_L 和 N_L 分别表示第 L 个数据云中样本及数量; μ_L 和 \mathbf{X}_L 分别表示局部均值和 L_2 范数平方. 当接收 \mathbf{u}_k ($k \leftarrow k+1$) 时, μ 和 \mathbf{X} 更新方法如下:

$$\mu \leftarrow \frac{(k-1)\mu + \mathbf{u}_k}{k}, \quad \mathbf{X} \leftarrow \frac{(k-1)\mathbf{X} + \|\mathbf{u}_k\|_2^2}{k} \quad (10)$$

μ_L 和 \mathbf{X}_L 可通过下式更新:

$$\mu_L \leftarrow \frac{N_L \mu_L + \mathbf{u}_k}{N_L + 1}, \quad \mathbf{X}_L \leftarrow \frac{N_L \mathbf{X}_L + \|\mathbf{u}_k\|_2^2}{N_L + 1} \quad (11)$$

2.2.2 数据云的生成

当接收到新输入 \mathbf{u}_k ($k = 2, 3, \dots, K$) 时, 通过计算数据密度 D [38-39] 来监控样本分布的变化. 如果满足条件 (12), 则创建 1 个新规则 \mathcal{R}_L ($L \leftarrow L+1$) 的数据云, 并初始化其中心为 $\mu_L \leftarrow \mathbf{u}_k$.

$$\text{如果 } D(\mathbf{u}_k) < \min_{i=[1, L]} (D(\mu_i))$$

$$\text{或者 } D(\mathbf{u}_k) > \max_{i=[1, L]} (D(\mu_i))$$

$$\text{且 } D_i(\mathbf{u}_k) \leq D_0$$

$$\text{则 新的数据云 } \mathcal{C}_L \text{ 建立, } L \leftarrow L+1 \quad (12)$$

其中, $D(\mathbf{u}_k) < \min_{i=[1, L]} (D(\mu_i))$ 和 $D(\mathbf{u}_k) > \max_{i=[1, L]} (D(\mu_i))$ 表示样本落在全局新区域/分布边界; $D_i(\mathbf{u}_k) \leq D_0$ 用于抑制在已有数据云附近因随机波动误增生新规则; D_0 为控制规则增长的门控阈值; $D(\mathbf{u}_k)$ 表示 \mathbf{u}_k 对整体样本序列的全局密度; $D(\mu_i)$ ($i = 1, 2, \dots, L$) 表示 \mathcal{C}_i 和整体样本序列间的全局密度; $D_i(\mathbf{u}_k)$ 表示 \mathbf{u}_k 和 \mathcal{C}_i 间的局部密

度. 具体公式如下 [17-18]:

$$\begin{cases} D(\mu_i) = e^{-\frac{\|\mu_i - \mu\|_2^2}{\mathbf{X} - \|\mu\|_2^2}} \\ D(\mathbf{u}_k) = e^{-\frac{\|\mathbf{u}_k - \mu\|_2^2}{\mathbf{X} - \|\mu\|_2^2}} \\ D_i(\mathbf{u}_k) = e^{-\frac{\|\mathbf{u}_k - \mu_i\|_2^2}{\mathbf{X}_i - \|\mu_i\|_2^2}} \end{cases} \quad (13)$$

2.2.3 样本归并与数据云的在线更新

如条件 (12) 所示, 当新样本未触发新数据云生成时, 系统将其归并到最相容的数据云中. 通过计算 \mathbf{u}_k 与各数据云中心之间的距离选择最近数据云, 并按式 (14) 在线更新该数据云的统计量.

$$\begin{aligned} \mathcal{C}_j &\leftarrow \mathcal{C}_j \cup \{\mathbf{u}_k\}, \quad \mu_j \leftarrow \frac{N_j \mu_j + \mathbf{u}_k}{N_j + 1}, \\ \mathbf{X}_j &\leftarrow \frac{N_j \mathbf{X}_j + \|\mathbf{u}_k\|_2^2}{N_j + 1}, \quad N_j \leftarrow N_j + 1 \end{aligned} \quad (14)$$

其中, $j = \operatorname{argmin}_{i=1, 2, \dots, L} \|\mu_i - \mathbf{u}_k\|_2^2$ 表示最接近当前样本的数据云序号.

2.2.4 数据云的移除

在数据云生成过程中, 可能出现两种情况: 1) 新数据云仅与少数样本规则一致 [18]; 2) 新数据云学习了样本中的异常数据 [20]. 冗余结构通常会降低计算效率 [39]. 为了提高模型的在线性能, 引入效用度量以移除冗余的数据云, 如式 (15) ~ (17) 所示.

$$\text{如果 } \eta_i^k < \eta_0 \text{ 则 第 } i \text{ 个数据云被移除} \quad (15)$$

其中, η_0 是修剪阈值; η_i^k 是第 i 个数据云 \mathcal{C}_i 的效用值, 即建立时刻起平均归一化激活度, 可由式 (16) 计算:

$$\eta_i^k = \frac{\sum_{r=I_i}^k \lambda_i^r}{k - I_i} \quad (16)$$

其中, I_i 表示 \mathcal{C}_i 建立时的迭代数; λ_i^r 表示 \mathcal{C}_i 的激活程度, 其计算方法为:

$$\lambda_i^r = \frac{D_i(\mathbf{u}_k)}{\sum_{l=1}^L D_l(\mathbf{u}_k)} \quad (17)$$

2.3 参数更新机制

该机制涉及参数更新和模型输出计算. 将核学习策略^[24]和 MTGC 准则^[34]结合, 用于迭代处理各数据云内样本子集和自适应更新权值向量. 这种结合具有内在互补性: 核学习可处理复杂非线性数据关系, 而 MTGC 准则能有效减轻噪声和异常值影响. 在同一递推框架下结合两者, 模型能在高维空间中准确拟合复杂模式, 并在非高斯环境下保持稳定, 避免单独使用核学习或相关熵损失函数时的过拟合或鲁棒性不足问题. 本研究的改进点在于不是简单把相关熵当作损失替换 MSE, 而是基于 MTGC 准则构建鲁棒目标, 并将其与核映射结合, 再推导递归闭式更新权值向量.

2.3.1 权值向量的更新

首先, 输入样本 \mathbf{u}_k 通过核映射 $\varphi(\cdot)$ 投影到高维特征空间, 得到新输入 $\varphi(\mathbf{u}_k)$. 基于 MTGC 准则, 建立目标函数如式 (18). 物理上, $J(\mathbf{w})$ 通过指数核 $\exp(-|e(k)|^\alpha / (\beta^\alpha \|\bar{\mathbf{w}}(k)\|^\alpha))$ 放大“误差接近 0 的样本区域”的贡献, 抑制“大幅度离群误差”的主导地位, 等价于在学习过程中自动降低离群点权重, 从而更适合非高斯噪声/重尾分布.

$$J(\mathbf{w}) = \max \sum_{k=1}^K \exp\left(-\frac{|e'(k)|^\alpha}{\beta^\alpha}\right) - \frac{\rho}{2} \|\mathbf{w}\|_2^2 \quad (18)$$

其中, ρ 为正则化因子; $e'(k) = e(k) / \sqrt{\mathbf{w}^T(k)\bar{\mathbf{w}}(k)}$ 为加权误差, $e(k) = d_k - \mathbf{w}^T \varphi(\mathbf{u}_k)$ 为预测误差. 对式 (18) 求关于 \mathbf{w} 的偏导并令其为零, 可计算权值向量如下:

$$\mathbf{w} = \left[\Phi \mathbf{G} \Phi^T + \left(\frac{\rho}{\alpha} \beta^\alpha \|\bar{\mathbf{w}}\|^{2\alpha} - \mathbf{R} \mathbf{T} \right) \mathbf{I} \right]^{-1} \Phi \mathbf{G} \mathbf{d} \quad (19)$$

在第 k 个样本中, 式 (19) 中的变量定义如下:

$$\begin{cases} \Phi(k) = [\varphi(\mathbf{u}_1), \varphi(\mathbf{u}_2), \dots, \varphi(\mathbf{u}_k)] \\ \mathbf{d}(k) = [d_1, d_2, \dots, d_k]^T \\ \mathbf{T} = [1, 1, \dots, 1]^T \in \mathbf{R}^{k \times 1} \\ \bar{\mathbf{w}}(k) = [\sqrt{\gamma} \quad -\mathbf{w}^T(k)] \\ \mathbf{G}(k) = \text{diag}\{g(1)\|\bar{\mathbf{w}}(1)\|^2, \dots, g(k)\|\bar{\mathbf{w}}(k)\|^2\} \\ \mathbf{R}(k) = (g(1)|e(1)|^2, \dots, g(k)|e(k)|^2) \mathbf{I} \\ g(k) = \exp\left(-\frac{|e(k)|^\alpha}{\beta^\alpha \|\bar{\mathbf{w}}(k)\|^\alpha}\right) |e(k)|^{\alpha-2} \|\bar{\mathbf{w}}(k)\|^{\alpha-2} \end{cases} \quad (20)$$

其中, $\text{diag}\{\cdot\}$ 表示对角矩阵. 为避免直接对式 (19) 中的大矩阵求逆, 引入矩阵求逆引理, 其一般形式为:

$$\begin{cases} (A + BCD)^{-1} = A^{-1} - A^{-1}BZ^{-1}DA^{-1} \\ Z = DA^{-1}B + C^{-1} \end{cases} \quad (21)$$

将式 (19) 的矩阵项分别对应到引理中, 即 $(\frac{\rho}{\alpha} \beta^\alpha \|\bar{\mathbf{w}}\|^{2\alpha} - \mathbf{R} \mathbf{T}) \mathbf{I} \rightarrow A$, $\Phi \rightarrow B$, $\mathbf{G} \rightarrow C$, $\Phi^T \rightarrow D$, 可将式 (19) 等价重写为:

$$\begin{cases} \mathbf{w}(k) = \Phi(k) [\Phi^T(k)\Phi(k) + \mathbf{P}(k)]^{-1} \mathbf{d}(k) \\ \mathbf{P}(k) = \left(\frac{\rho}{\alpha} \beta^\alpha \|\bar{\mathbf{w}}(k)\|^{2\alpha} - \mathbf{R}(k)\mathbf{T}(k) \right) \mathbf{G}^{-1}(k) \end{cases} \quad (22)$$

定义权值矩阵 $\boldsymbol{\theta}(k) = [\Phi^T(k)\Phi(k) + \mathbf{P}(k)]^{-1} \times \mathbf{d}(k)$, 结合式 (22), 权值向量可表示为:

$$\mathbf{w}(k) = \Phi(k)\boldsymbol{\theta}(k) \quad (23)$$

$\|\bar{\mathbf{w}}(k)\|^{2\alpha}$ 在 $\boldsymbol{\theta}(k)$ 中的计算如下:

$$\begin{aligned} \|\bar{\mathbf{w}}(k)\|^{2\alpha} &= (\mathbf{w}^T(k-1)\mathbf{w}(k-1) + \gamma)^\alpha = \\ &= (\boldsymbol{\theta}^T(k-1)\mathbf{K}(k-1)\boldsymbol{\theta}(k-1) + \gamma)^\alpha = \\ &= \|\bar{\boldsymbol{\theta}}(k)\|^{2\alpha} \end{aligned} \quad (24)$$

其中, $\bar{\boldsymbol{\theta}}(k) = [\sqrt{\gamma}, \sqrt{\boldsymbol{\theta}^T(k-1)\mathbf{K}(k-1)\boldsymbol{\theta}(k-1)}]$; $\mathbf{K}(k-1) = \Phi^T(k-1)\Phi(k-1)$ 是核矩阵. 将式 (24) 代入式 (23), $\boldsymbol{\theta}(k)$ 可重写为:

$$\begin{cases} \boldsymbol{\theta}(k) = [\Phi^T(k)\Phi(k) + \mathbf{P}^*(k)]^{-1} \mathbf{d}(k) \\ \mathbf{P}^*(k) = \left(\frac{\rho}{\alpha} \beta^\alpha \|\bar{\boldsymbol{\theta}}(k)\|^{2\alpha} - \mathbf{R}(k)\mathbf{T}(k) \right) \mathbf{G}^{-1}(k) \end{cases} \quad (25)$$

定义 $\mathbf{H}(k) = [\Phi^T(k)\Phi(k) + \mathbf{P}^*(k)]^{-1}$, 使用块矩阵形式计算矩阵 $\mathbf{H}(k)$ 的逆如下:

$$\mathbf{H}^{-1}(k) = \begin{bmatrix} \mathbf{H}^{-1}(k-1) & \Phi^T(k-1)\varphi(\mathbf{u}_k) \\ \varphi^T(\mathbf{u}_k)\Phi(k-1) & z(k) \end{bmatrix} \quad (26)$$

其中

$$\begin{cases} z(k) = \boldsymbol{\varphi}^T(\mathbf{u}_k)\boldsymbol{\varphi}(\mathbf{u}_k) + \\ \quad \left(\frac{\rho}{\alpha}\beta^\alpha\|\bar{\boldsymbol{\theta}}(k)\|^{2\alpha} - r(k)\right)g'^{-1}(k) \\ g'(k) = \exp\left(-\frac{|e(k)|^\alpha}{\beta^\alpha\|\bar{\boldsymbol{\theta}}(k)\|^\alpha}\right)|e(k)|^{\alpha-2}\|\bar{\boldsymbol{\theta}}(k)\|^\alpha \\ r(k) = \exp\left(-\frac{|e(k)|^\alpha}{\beta^\alpha\|\bar{\boldsymbol{\theta}}(k)\|^\alpha}\right)|e(k)|^\alpha\|\bar{\boldsymbol{\theta}}(k)\|^{\alpha-2} \end{cases} \quad (27)$$

令 $\mathbf{h}(k) = \boldsymbol{\Phi}^T(k-1)\boldsymbol{\varphi}(\mathbf{u}_k)$, $\mathbf{H}(k)$ 可表示为:

$$\mathbf{H}(k) = \nu^{-1}(k) \begin{bmatrix} \mathbf{H}(k-1)\nu(k) + \mathbf{m}(k)\mathbf{m}^T(k) & -\mathbf{m}(k) \\ -\mathbf{m}^T(k) & 1 \end{bmatrix} \quad (28)$$

其中, $\nu(k) = z(k) - \mathbf{m}^T(k)\mathbf{h}(k)$; $\mathbf{m}(k) = \mathbf{H}(k-1) \times \mathbf{h}(k)$.

结合式 (25) 和 (28), $\boldsymbol{\theta}(k)$ 可通过下式更新:

$$\begin{aligned} \boldsymbol{\theta}(k) &= \mathbf{H}(k)\mathbf{d}(k) = \\ & \begin{bmatrix} \mathbf{H}(k-1) + \mathbf{m}(k)\mathbf{m}^T(k)\nu^{-1}(k) & -\mathbf{m}(k)\nu^{-1}(k) \\ -\mathbf{m}^T(k)\nu^{-1}(k) & \nu^{-1}(k) \end{bmatrix} \times \\ & \begin{bmatrix} \mathbf{d}(k-1) \\ d_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}(k-1) - \mathbf{m}(k)\nu^{-1}(k)e(k) \\ \nu^{-1}(k)e(k) \end{bmatrix} \end{aligned} \quad (29)$$

2.3.2 模型输出的计算

基于式 (29), 可计算第 k 个样本的模型输出如下:

$$\begin{aligned} \hat{d}_k &= \mathbf{w}^T(k)\boldsymbol{\varphi}(\mathbf{u}_k) = [\boldsymbol{\Phi}(k)\boldsymbol{\theta}(k)]^T\boldsymbol{\varphi}(\mathbf{u}_k) = \\ & \boldsymbol{\theta}^T(k)\boldsymbol{\Phi}^T(k)\boldsymbol{\varphi}(\mathbf{u}_k) \end{aligned} \quad (30)$$

由于 $\boldsymbol{\Phi}^T(k)\boldsymbol{\varphi}(\mathbf{u}_k)$ 对应于核函数映射, 可将式 (30) 重写为:

$$\hat{d}_k = \sum_{j=1}^k \boldsymbol{\theta}(k)\kappa(\mathbf{u}_j, \mathbf{u}_k) \quad (31)$$

其中, $\kappa(\mathbf{u}_j, \mathbf{u}_k) = \boldsymbol{\varphi}^T(\mathbf{u}_j)\boldsymbol{\varphi}(\mathbf{u}_k)$ 表示核函数.

由于高斯核函数在计算过程中具有良好的数值稳定性^[20, 25], 在模型中使用高斯核函数, 其定义为:

$$\kappa(\mathbf{u}_j, \mathbf{u}_k) = \exp\left(-\frac{\|\mathbf{u}_j - \mathbf{u}_k\|_2^2}{2\sigma^2}\right) \quad (32)$$

其中, σ 表示核带宽. 因此, 模型通过使用式 (29) 迭代更新系数向量 $\boldsymbol{\theta}(k)$, 并基于高斯核映射计算预测值 \hat{d}_k , 从而确保非线性学习过程.

2.4 CKL-EFS 算法伪代码

CKL-EFS 的学习过程如算法 1 所示.

算法 1. CKL-EFS 算法伪代码

输入. 样本 $\{(\mathbf{u}_k, d_k)\}_{k=1}^K$, 门控阈值 D_0 , 修剪阈值 η_0 , 高斯核带宽 σ , 正则化因子 ρ , 形状参数 α , 尺度参数 β .

输出. 模型预测输出 \hat{d}_k .

- 1) **for** $k = 1, 2, \dots, K$ **do**
- 2) **if** $k = 1$ **then**
- 3) 使用式 (8) 和 (9) 初始化 $\boldsymbol{\mu}$ 、 \mathbf{X} 、 $\boldsymbol{\mu}_1$ 、 \mathcal{R}_1 、 \mathbf{X}_1 、 \mathcal{C}_1 ;
- 初始化: $\mathbf{H}(1) = 1$, $\boldsymbol{\theta}(1) = (\kappa(\mathbf{u}_1, \mathbf{u}_1))^{-1}d_1$, $e(1) = 0$;
- 4) **else**
- 5) **for** $i = 1, 2, \dots, L$ **do**
- 6) 使用式 (13) 计算 $D(\mathbf{u}_k)$ 、 $D(\boldsymbol{\mu}_i)$ 、 $D_i(\mathbf{u}_k)$;
- 7) **if** \mathbf{u}_k 满足条件 (12) **then**
- 8) 生成新规则即 $L \leftarrow L + 1$;
- 使用式 (9) 初始化 $\boldsymbol{\mu}_L$ 、 \mathcal{R}_L 、 \mathbf{X}_L 、 \mathcal{C}_L ;
- 初始化 $\bar{\mathbf{w}}_k$: $\bar{\mathbf{w}}(k) = [\sqrt{\gamma} \quad -\mathbf{w}^T(k)]$;
- 初始化 $g(k)$ 、 $\mathbf{R}(k)$.
- 9) **else**
- 10) 搜索距离样本最近数据云 j ,
- $j = \operatorname{argmin}_{i=1, 2, \dots, L} (\|\boldsymbol{\mu}_i - \mathbf{u}_k\|_2^2)$
- 使用式 (14) 更新 $\boldsymbol{\mu}_j$ 、 \mathcal{R}_j 、 \mathbf{X}_j ;
- 使用式 (27) 计算 $z(k)$ 、 $g'(k)$ 、 $r(k)$;
- 计算 $\nu(k)$ 、 $\mathbf{m}(k)$: $\nu(k) = z(k) - \mathbf{m}^T(k) \times \mathbf{h}(k)$, $\mathbf{m}(k) = \mathbf{H}(k-1)\mathbf{h}(k)$;
- 使用式 (28) 和 (29) 更新 $\mathbf{H}(k)$ 、 $\boldsymbol{\theta}(k)$.
- 11) **end if**
- 12) **end for**
- 13) **end if**
- 14) **for** $i = 1, 2, \dots, L$ **do**
- 15) **if** \mathbf{u}_k 满足条件 (15) **then**
- 16) 移除冗余规则 \mathcal{R}_i 和第 i 个数据云 \mathcal{C}_i ;
- 17) **end if**
- 18) **end for**
- 19) 使用式 (7), 式 (30) ~ (32) 计算最近数据云局部输出, 即模型最终输出 \hat{d}_k .
- 20) **end for**

3 计算复杂度分析

所提 CKL-EFS 的在线代价主要由两部分组成: 模糊推理和递归更新. 对每个新样本, 模型首先需要计算当前有效模糊规则对该样本的激活度并完成线性加权, 设当前激活规则数为 L , 则该过程的时间复杂度为 $O(L)$. 随后, 模型仅对与该样本相关的局部规则子集进行递归式参数更新, 设该子集大

小为 K' ($K' \leq K$), 则增量矩阵更新通常为 $O(K'^2)$, 在需要对该子块做完全重构时最高不超过 $O(K'^3)$, 与常见的 RLS 阶数相当. 因而, 在规则库进入稳定区间后, 单步在线计算的时间复杂度可写为 $O(L) + O(K'^3)$. 与传统仅增式方法推理复杂度最终增长到 $O(M_t)$ (其中 M_t 为时刻 t 前累积全部规则) 不同, CKL-EFS 的规则库在在线阶段同时执行增生、样本归并和低贡献剔除, 因此有效规则数 L 不会随时间线性累积, 而是稳定在一个受控范围内; 单次更新所涉及的局部规模 K' 也同步受到该上界约束. 这样, 上述复杂度表示的是当前活跃结构的开销, 而不是“运行时间越长就越慢”的累积开销. 为进一步降低最坏情况下的 $O(K'^3)$ 代价, 框架还可以采用低秩递归、仅更新高激活度规则或间隔式更新等工程策略.

4 实验结果与分析

本节在不同数据集中对所提 CKL-EFS 和对比模型进行了仿真实验. 为方便在线预测领域研究人员后续研究, 本文开放了 CKL-EFS 源代码, 相关代码可从网站 <https://gitee.com/hulei666888/ckl-efs-model-code> 获取.

4.1 实验配置说明

在配置为 8 核 I7 CPU 和 32 GB 内存的台式计算机上, 使用 MATLAB R2023b 编程软件进行本节的仿真实验. 选取 CEFNS^[29]、PSO-ALMMo*^[40]、RMCEFS^[25]、MEEFIS^[21]、ePL-KRLS-DISCO^[22]、EPL-KGLA^[30] 和 MIIPSO-EFS^[18] 作为对比模型. 其中, CEFNS、RMCEFS 和 EPL-KGLA 都内置了相关熵算法, ePL-KRLS-DISCO 和 EPL-KGLA 使用递归核学习算法进行参数更新, PSO-ALMMo* 和 MIIPSO-EFS 结合了元启发式优化算法进行参数自适应优化. 运行时间被用于度量各模型的预测效率. 归一化均方根误差 (normalized root mean square error, NRMSE)、对称平均绝对百分比误差 (symmetric mean absolute percentage error, SMAPE)、平均绝对误差 (mean absolute error, MAE) 和决定系数 (coefficient of determination,

R^2) 被用于评估模型精度, 其计算方法如下:

$$\left\{ \begin{array}{l} \text{NRMSE} = \sqrt{\frac{\sum_{k=1}^K (\hat{d}_k - d_k)^2}{\sum_{k=1}^K (\bar{d}_k - d_k)^2}} \\ \text{SMAPE} = \frac{2}{K} \sum_{k=1}^K \frac{|d_k - \hat{d}_k|}{|d_k| + |\hat{d}_k|} \\ \text{MAE} = \frac{1}{K} \sum_{k=1}^K |d_k - \hat{d}_k| \\ R^2 = 1 - \frac{\sum_{k=1}^K (d_k - \hat{d}_k)^2}{\sum_{k=1}^K (d_k - \bar{d}_k)^2} \end{array} \right. \quad (33)$$

其中, d_k 为第 k 个真实值; \hat{d}_k 为第 k 个预测值; \bar{d}_k 是 d_k 的均值; K 是样本数. 表 1 为数据集参数设置.

为避免逐数据集手工调参, 本文对 CKL-EFS 的超参数采用训练集内验证的统一调优流程. 具体而言, 对每个数据集, 在训练集内部按时间顺序划分出一段验证集 (取训练集最后 10% 样本作为验证集), 并保持测试集完全独立仅用于最终评估, 以避免信息泄露. CKL-EFS 的超参数可分为三类: 1) 规则演化阈值 (门控阈值 D_0 、修剪阈值 η_0); 2) 核学习相关参数 (核带宽 σ 、正则化因子 ρ); 3) 鲁棒参数 (形状参数 α 、尺度参数 β). 1) 中参数用于控制规则库增长与修剪强度, 在所有数据集上固定为 $D_0 = e^{-\frac{1}{4}}$ 与 $\eta_0 = 0.1$, 从而保证结构复杂度的可控性与跨数据集一致性. 对于 2) 和 3) 中参数, 采用离散网格搜索, 在验证集上以 MAE 最小作为选择准则. 本文使用的候选集合为: $\sigma \in \{0.1, 1.0, 5.0, 10.0\}$, $\rho \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$, $\alpha \in \{1.4, 1.6, 1.8, 2.0\}$, $\beta \in \{0.5, 1.0, 2.0, 4.0\}$. 此外, 为进一步验证结论不依赖精细调参, 在新疆风电数据集上对关键超参数进行 $\pm 20\%$ 扰动, 结果显示性能波动较小且相对排序保持稳定.

4.2 Lorenz-96 混沌系统

Lorenz-96 混沌系统是一种描述混沌特性的数

表 1 数据集的参数设置
Table 1 Parameter settings of datasets

数据集	维度	样本数	训练集样本数	测试集样本数
Lorenz-96 混沌系统	40	2 400	1 800	600
伦敦气象数据集	15	4 000	3 000	1 000
利兹水文数据集	24	1 200	1 000	200
新疆风电数据集	15	3 000	2 500	500

学模型, 由 Edward Lorenz 于 1996 年提出, 是早期在大气对流和混沌理论方面工作的重要体现. Lorenz-96 由一组耦合的常微分方程组成, 通常表示为^[41]:

$$\frac{dv_i}{dt} = (v_{i+1} - v_{i-2})v_{i-1} - v_i + F \quad (34)$$

其中, v_i 表示系统状态, $i = 1, 2, \dots, D$, $v_{-1} = v_{i-1}$, $v_0 = v_i$, $v_1 = v_{i+1}$, $D \geq 4$ 表示维度; F 表示强迫系数. 本研究选择 $D = 40$, $F = 8$, 以诱导系统的混沌行为. 基于式 (34), 结合龙格-库塔方法生成了 2 400 个样本. 图 2 展示了不同模型在 1 步超前预测时的规则数量和运行时间方面的对比结果. 每个模型通过两根柱状条表示, 蓝色表示规则数, 橙色表示运行时间. MEEFIS 因多层结构生成最多的模糊规则, 但由于分布式计算模型, 运行时间最短. CKL-EFS 的最终规则数为 8, 和 PSO-ALMMo* 和 MIIPSO-EFS 在同一水平. 这说明在线规则库演化机制避免了无控制的规则爆炸, 从而保持了可实时部署的复杂度.

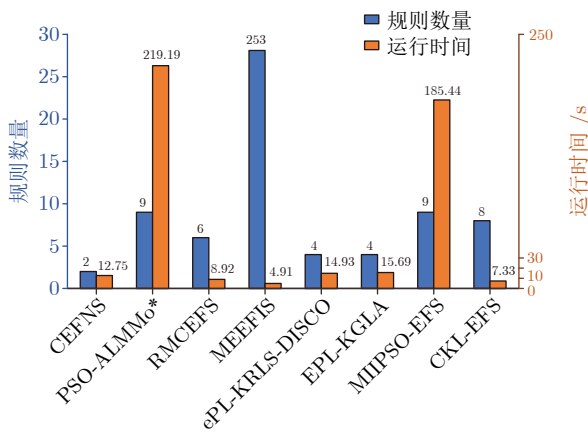


图 2 Lorenz-96 数据集: 不同模型的规则数量和运行时间
Fig.2 Lorenz-96 dataset: Number of rules and runtime for different models

为分析不同模型在非高斯噪声环境下的稳定性, 在预测目标序列中分别叠加两类典型非高斯扰动: 1) α -稳定噪声, 其稳定性指数设为 1.8、偏度为 1、尺度为 4、位置为 30, 用以刻画具有强正偏、重尾、冲击性离群点的脉冲干扰; 2) 粉红噪声, 其采样率为 1 800 Hz、持续 1 s, 并经一阶高通 Butterworth 滤波, 用以模拟带明显时间相关性的低频漂移型扰动. 需要指出的是, 本文的噪声参数用于构造标准化压力测试, 并不旨在复现某一具体物理噪声机理. 噪声序列在注入前均与目标序列对齐为相同长度, 并采用一致的幅值缩放规则以避免扰动完全主导信号. 所有对比模型均在相同噪声注入方案下评测, 以保证比较的公平性. 由表 2 可知, 对于大

多数基线模型, 一旦从无噪声切换到任一非高斯噪声环境, 其 MAE 均显著上升, 表明这类模型在重尾/相关型干扰出现后预测误差会迅速恶化, 并且这种退化在两类噪声下都会发生. 相比之下, CKL-EFS 的 MAE 仅由 0.359 提高到 0.372 和 0.397, 无论从误差绝对水平还是从误差增幅来看, 均保持最优. 也就是说, 随着噪声从无噪声到重尾脉冲型噪声, 再到具有相关漂移特征的粉红噪声, CKL-EFS 始终呈现出最小的误差放大趋势, 而其他模型往往出现明显的性能塌陷. 该结果说明: CKL-EFS 不仅在干净数据上具有可比精度, 在非高斯性扰动形式变化情况下也保持了最强的鲁棒性.

表 2 Lorenz-96 数据集: 在不同非高斯扰动下的预测性能 (度量标准: MAE)

Table 2 Lorenz-96 dataset: Prediction performance under various non-Gaussian disturbances (metric: MAE)

噪声类型	模型	值 ($\times 10^{-1}$)
理想观测 (无噪声)	CEFNS ^[29]	5.88
	PSO-ALMMo* ^[40]	5.01
	RMCEFS ^[25]	5.45
	MEEFIS ^[21]	3.01
	ePL-KRLS-DISCO ^[22]	3.15
	EPL-KGLA ^[30]	4.36
	MIIPSO-EFS ^[18]	3.75
	CKL-EFS	3.59
重尾脉冲型非高斯扰动 (加 α -稳定噪声)	CEFNS ^[29]	5.92
	PSO-ALMMo* ^[40]	7.48
	RMCEFS ^[25]	5.73
	MEEFIS ^[21]	6.61
	ePL-KRLS-DISCO ^[22]	6.47
	EPL-KGLA ^[30]	5.29
	MIIPSO-EFS ^[18]	6.41
	CKL-EFS	3.72
强相关低频非高斯扰动 (加粉红噪声)	CEFNS ^[29]	6.13
	PSO-ALMMo* ^[40]	7.98
	RMCEFS ^[25]	5.77
	MEEFIS ^[21]	6.85
	ePL-KRLS-DISCO ^[22]	7.32
	EPL-KGLA ^[30]	5.41
	MIIPSO-EFS ^[18]	7.13
	CKL-EFS	3.97

为进一步刻画非高斯性增强对预测稳定性的影响, 以 α -稳定噪声为载体, 将稳定性指数 α_0 从 2.0 逐步降低至 1.2. 需要指出的是, α_0 越小意味着噪声分布重尾特性越强、离群冲击越频繁, 因而该设

置可用于连续地增强非高斯扰动强度. 表 3 给出了不同 α_0 水平下各模型的 MAE 变化规律. 可以观察到: 1) 对于大多数基线模型, 随着 α_0 降低, MAE 呈单调上升趋势, 且在 $\alpha_0 \leq 1.6$ 时误差增长明显加速, 说明其对重尾离群点的敏感性较高; 2) 相比之下, CKL-EFS 的误差增长斜率最小, α_0 从 2.0 到 1.2 仅出现温和上升, 表明其在非高斯性持续增强时仍能保持稳定的在线建模能力. 这进一步验证了 CKL-EFS 在重尾脉冲型扰动下的鲁棒性优势: 不仅能抵抗单一强度设置下的误差恶化, 也能在扰动强度逐步提升的情况下维持最小的性能退化幅度.

表 3 Lorenz-96 数据集: 非高斯性增强时的性能变化规律 (以 α -稳定噪声为例, 度量标准: MAE ($\times 10^{-1}$))

Table 3 Lorenz-96 dataset: Performance trend as non-Gaussianity increases (via α -stable noise, metric: MAE ($\times 10^{-1}$))

模型	稳定性指数 α_0				
	2.0	1.8	1.6	1.4	1.2
CEFNS ^[29]	5.57	5.92	6.44	7.11	8.15
PSO-ALMMo* ^[40]	5.35	7.48	8.35	9.22	10.31
RMCEFS ^[25]	5.60	5.73	6.10	6.72	7.54
MEEFIS ^[21]	4.15	6.61	7.25	8.12	9.05
ePL-KRLS-DISCO ^[22]	4.02	6.47	7.05	7.88	8.96
EPL-KGLA ^[30]	4.95	5.29	5.70	6.22	6.95
MIIPSO-EFS ^[18]	4.40	6.41	7.20	8.05	9.28
CKL-EFS	3.62	3.72	3.85	4.10	4.45

4.3 伦敦气象数据集

伦敦气象数据集¹ 包含伦敦从 2020 年 1 月 1 日至 2020 年 12 月 31 日共 365 天的气象数据, 每 10 分钟记录 1 次. 该数据集包含 52 696 个数据样本, 涵盖 15 个维度. 尽管温度和风速作为气象指标具有重要意义, 但它们的数据样本通常包含非高斯噪声.

图 3 显示了两个预测变量的非高斯噪声强度. 温度样本的偏度为 0.5, 已经偏离了高斯噪声的对称性特征 (偏度 = 0), 表明噪声中含有非高斯成分. 风速的偏度和峰度更高, 偏离正态分布更明显, 说明数据中的非高斯噪声更强. 样本分位数与理论分位数之间的偏差进一步证实了这一观察结果.

图 4 为 CKL-EFS 的 1 步超前预测曲线. 预测温度时可以有效跟踪样本趋势, 误差波动较为稳定; 对风速预测时也能有效跟踪真实值, 且将误差控制在 -5 至 5 之间, 服从仍接近正态分布, 这表明 CKL-EFS 面对较强的非高斯噪声干扰时, 仍具有稳定的

¹ https://biendata.com/competition/kdd_2018

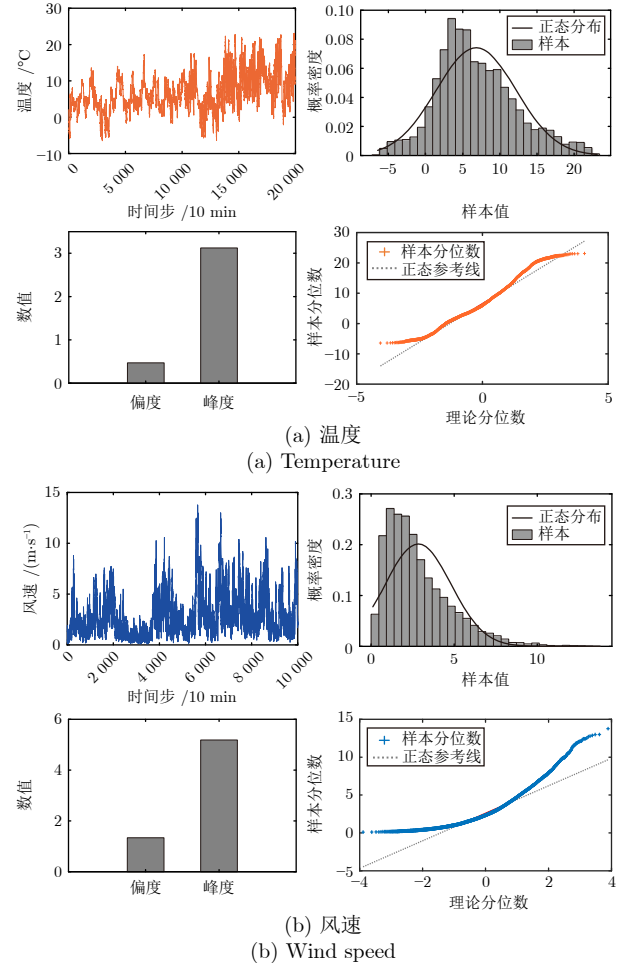


图 3 伦敦气象数据集: 不同预测因子的波动趋势和非高斯噪声分析

Fig. 3 London meteorological dataset: Fluctuation trend and non-Gaussian noise analysis of different predictors

性能. 整理预测指标如表 4 所示. 在温度预测任务中, CKL-EFS 的 NRMSE 为 0.054, SMAPE 为 0.087, 表明其具有较强的非线性建模能力. 在风速预测任务中, CEFNS、RMCEFS 和 EPL-KGLA 模型由于内置相关熵算法, 在非高斯噪声下表现出较强的适应性, 优于其他对比模型. 值得注意的是, CKL-EFS 达到了 0.213 的 NRMSE 和 0.074 的 SMAPE, 显著优于次优模型 EPL-KGLA, 精度分别提高了 23.93% 和 48.25%. 这说明, CKL-EFS 通过监控数据密度, 挖掘了丰富的模糊规则; 并且通过改进相关熵算法捕捉了误差的高阶矩信息, 在强非高斯噪声环境中适应能力更好.

4.4 利兹水文数据集

利兹水文数据集² 包含了利兹从 2017 年 11 月

² <https://www.kaggle.com/datasets/thomaswrightanderson/river-aire-discharge-time-series>

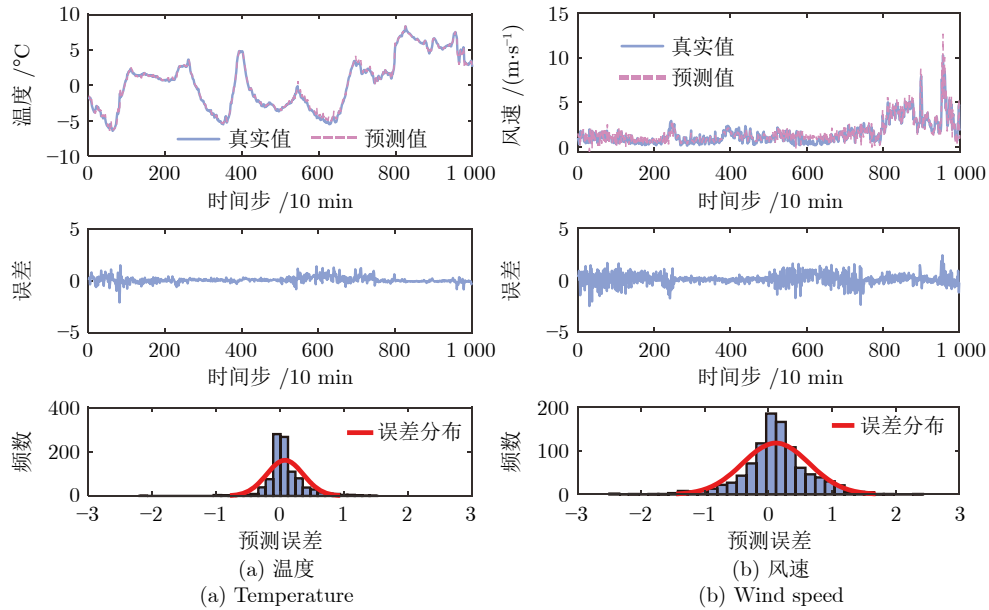


图 4 伦敦气象数据集: CKL-EFS 的 1 步超前预测结果

Fig.4 London meteorological dataset: 1-step-ahead prediction results of CKL-EFS

表 4 伦敦气象数据集: 不同模型的 1 步超前预测结果
Table 4 London meteorological dataset: 1-step-ahead prediction results by different models

性能指标	模型	温度 (°C)	风速 (m/s)
NRMSE ($\times 10^{-1}$)	CEFNS ^[20]	2.53	3.11
	PSO-ALMMo ^{*(40)}	1.71	8.03
	RMCEFS ^[25]	1.86	2.91
	MEEFIS ^[21]	0.75	3.47
	ePL-KRLS-DISCO ^[22]	1.68	4.35
	EPL-KGLA ^[30]	1.13	2.80
	MIIPSO-EFS ^[18]	1.39	4.52
	CKL-EFS	0.54	2.13
SMAPE ($\times 10^{-1}$)	CEFNS ^[20]	4.32	2.19
	PSO-ALMMo ^{*(40)}	5.03	5.13
	RMCEFS ^[25]	3.93	2.21
	MEEFIS ^[21]	1.84	3.02
	ePL-KRLS-DISCO ^[22]	2.28	2.87
	EPL-KGLA ^[30]	2.12	1.43
	MIIPSO-EFS ^[18]	2.75	2.52
	CKL-EFS	0.87	0.74
规则数量	CEFNS ^[20]	3	5
	PSO-ALMMo ^{*(40)}	9	9
	RMCEFS ^[25]	6	10
	MEEFIS ^[21]	379	329
	ePL-KRLS-DISCO ^[22]	2	4
	EPL-KGLA ^[30]	2	4
	MIIPSO-EFS ^[18]	12	8
	CKL-EFS	15	10
运行时间 (s)	CEFNS ^[20]	28.77	35.12
	PSO-ALMMo ^{*(40)}	527.69	172.41
	RMCEFS ^[25]	8.61	14.03
	MEEFIS ^[21]	11.26	6.59
	ePL-KRLS-DISCO ^[22]	50.73	64.32
	EPL-KGLA ^[30]	13.36	19.65
	MIIPSO-EFS ^[18]	363.21	81.05
	CKL-EFS	30.42	22.97

14 日到 2021 年 4 月 30 日共 1 265 个数据样本, 采样间隔为 1 天. 该数据集包括 24 个变量, 如空气湿度、空气温度、大气压力、土壤温度、风速、降水量、河流流量、河流水位等. 河流流量受到多种因素的影响, 如降雨量、融雪、地形、人类活动等, 这些因素的复杂性导致河流流量通常呈现出非高斯分布特征. 因此, 对河流流量数据进行分析 and 建模时, 需要考虑非高斯噪声的影响.

图 5(a) ~ 图 5(d) 中, 河流流量的波动显著, 说明数据非线性较强; 概率密度直方图偏离正态分布, Q-Q 图显示数据点偏离参考线. 偏度和峰度较高, 表明数据存在重尾和不对称性. 这些结论表明提取到的河流流量样本数据受到了非高斯噪声干扰. 图 5(e) 表明河流流量 (变量 12) 与其他变量间存在潜在的内在关联性, 且相互关系复杂, 十分依赖学习能力强的模型. 根据 1 步超前预测结果 (如表 5 和图 6), CKL-EFS 表现出明显的优势. 相比其他模型, CKL-EFS 的 MAE 最低, 为 2.19; R^2 高达 0.992, 展示了其强大的拟合能力和预测性能. 此外, CKL-EFS 的运行时间为 10.64 s, 说明其兼顾了预测效率和精度.

进一步整理多步超前预测结果如图 7 所示. 相比其他模型, CKL-EFS 在所有预测步长中均表现出了最低的误差值, 说明其在复杂非高斯噪声环境下进行多步预测时仍能保持较高的预测性能. 因此, CKL-EFS 不仅在预测精度和效率之间表现出良好的平衡, 而且在多步预测任务中性能稳定.

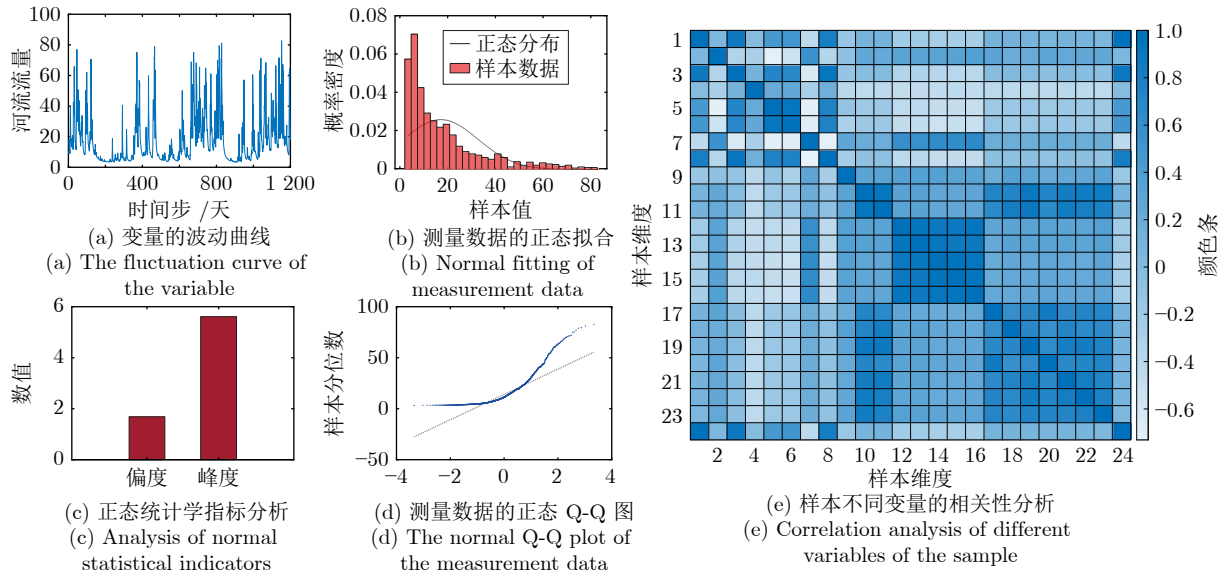


图 5 利兹水文数据集: 非高斯噪声强度分析和特征可视化分析

Fig. 5 Leeds hydrological dataset: Non-Gaussian noise intensity analysis and feature visualization analysis

表 5 利兹水文数据集: 不同模型的 1 步超前预测结果
Table 5 Leeds hydrological dataset: 1-step-ahead prediction results by different models

模型	MAE	规则数量	运行时间 (s)
CEFNS ^[29]	2.87	6	23.99
PSO-ALMMo ^{*[40]}	4.13	9	128.63
RMCEFS ^[25]	2.35	8	22.87
MEEFIS ^[21]	3.62	74	2.81
ePL-KRLS-DISCO ^[22]	4.04	2	2.58
EPL-KGLA ^[30]	2.72	2	13.87
MIIPSO-EFS ^[18]	3.87	9	84.35
CKL-EFS	2.19	11	10.64

4.5 新疆风电数据集

新疆风电数据集³ 涵盖新疆哈密地区某风电场从 2019 年 1 月 1 日至 2019 年 12 月 31 日的风电及相关气象信息, 数据每 15 min 采样 1 次. 该数据集包括 15 维特征, 如温度、气压、湿度、风速、风向、发电功率等. 分析图 8 发现, 风电功率数据的非高斯性并非普通重尾噪声, 而是结构性和离散化的, 远超出高斯噪声假设, 比河流流量的重尾单峰偏斜或温度/风速的轻重尾偏态更大. 因此, 传统以 MSE 为目标的递归估计方法在该场景下稳定性和鲁棒性不足, 必须引入对非高斯噪声友好的鲁棒准则以及自适应结构机制, 以保证预测可靠性.

整理 CKL-EFS 和其他内置相关熵算法的对比模型预测结果如表 6 所示. 在该数据集上, 所提

³ <https://www.xjwind.com/>

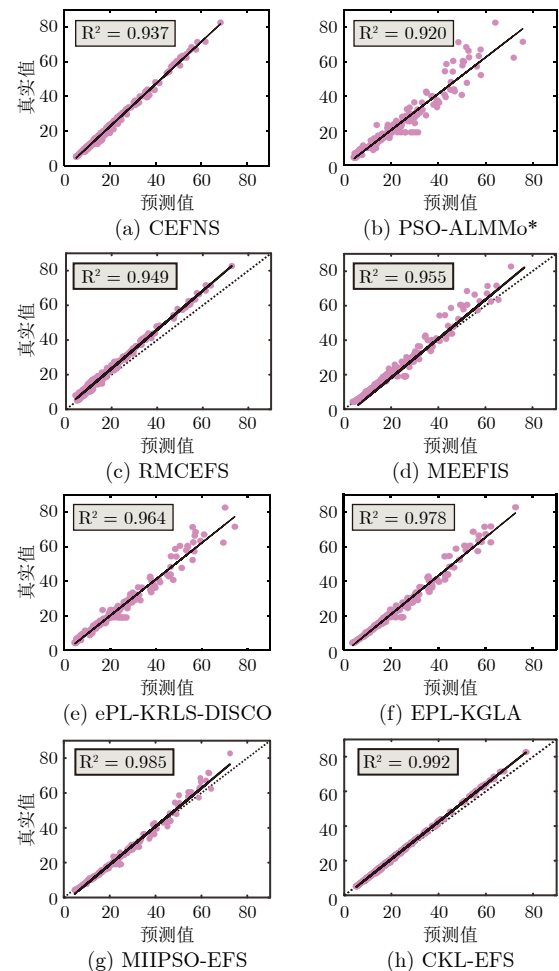


图 6 利兹水文数据集: 不同模型的 1 步超前预测散点图
Fig. 6 Leeds hydrological dataset: 1-step-ahead prediction scatter plot by different models

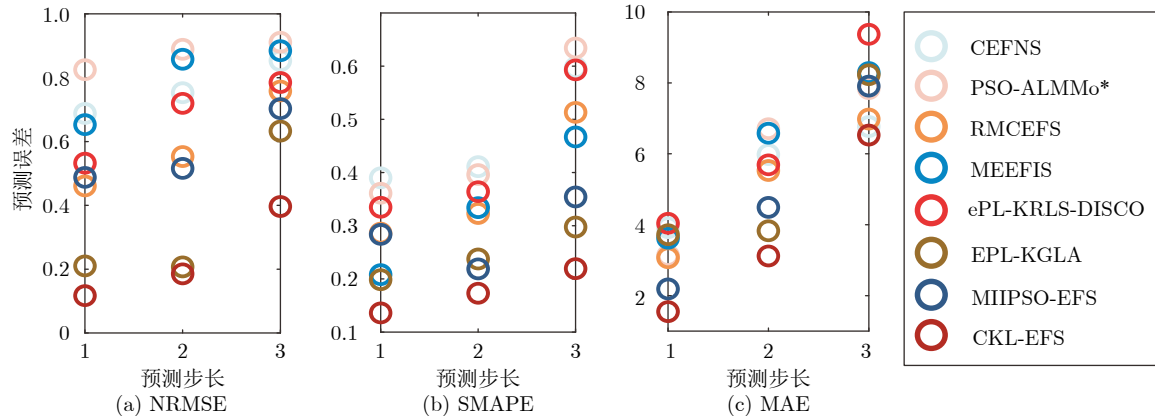


图7 利兹水文数据集: 多步超前预测的 NRMSE、SMAPE 和 MAE

Fig. 7 Leeds hydrological dataset: NRMSE, SMAPE and MAE of multi-step-ahead prediction

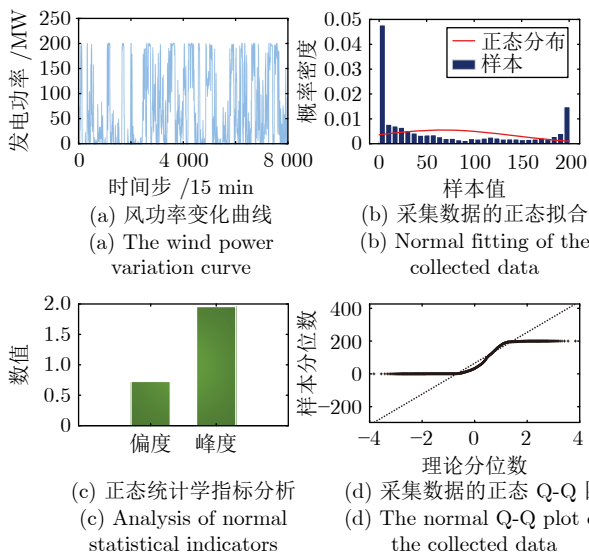


图8 新疆风电数据集: 风功率波动及非高斯噪声强度分析

Fig. 8 Xinjiang wind power dataset: Analysis of wind power fluctuations and non-Gaussian noise intensity

CKL-EFS 在保持适中规则数量 (10 条) 的同时, 将 1 步超前预测 SMAPE 降至最低的 0.211, 运行时间仅 18.52 s, 较规则数更多的 CEFNS (13 条)、RMCEFS (15 条) 获得更高精度和更低时间, 也优于规则数更少但精度不足的 EPL-KGLA (6 条). 这说明模糊规则库对结果具有显著影响: 规则过多会引入冗余、增加计算负担, 规则过少则难以充分刻画风功率的非线性特征, 而 CKL-EFS 通过自适应压缩与保留关键规则, 在规则数量与表达能力之间取得合理平衡, 从而实现紧凑结构下的高预测性能.

由表 7 分析可知, 在最优参数的基础上对部分超参数做 $\pm 20\%$ 扰动时, SMAPE 的增幅均小于 5%, 且模型仍优于 CEFNS、RMCEFS 与 EPL-

表6 新疆风电数据集: 不同模型的 1 步超前预测结果
Table 6 Xinjiang wind power dataset: 1-step-ahead prediction results by different models

模型	SMAPE ($\times 10^{-1}$)	规则数量	运行时间 (s)
CEFNS ^[29]	6.93	13	29.18
RMCEFS ^[25]	5.15	15	32.01
EPL-KGLA ^[30]	3.97	6	15.75
CKL-EFS	2.11	10	18.52

表7 新疆风电数据集: CKL-EFS 在不同超参数扰动下的 SMAPE 变化

Table 7 Xinjiang wind power dataset: SMAPE variation of CKL-EFS under different hyperparameter disturbances

设置	SMAPE ($\times 10^{-1}$)	SMAPE 差值	相对变化
最优参数	2.11	—	—
核带宽 $\sigma \uparrow 20\%$	2.17	0.06	+2.8%
正则化因子 $\gamma \uparrow 20\%$	2.14	0.03	+1.4%
形状参数 $\alpha \uparrow 20\%$	2.18	0.07	+3.3%
尺度参数 $\beta \uparrow 20\%$	2.16	0.05	+2.4%
核带宽 $\sigma \downarrow 20\%$	2.15	0.04	+1.9%
正则化因子 $\gamma \downarrow 20\%$	2.13	0.02	+0.9%
形状参数 $\alpha \downarrow 20\%$	2.18	0.07	+3.3%
尺度参数 $\beta \downarrow 20\%$	2.16	0.05	+2.4%

KGLA. 说明本文 CKL-EFS 的性能提升源于其自适应规则库和鲁棒递归更新机制本身, 并不依赖精细调参.

5 结束语

针对非高斯噪声环境下的在线预测任务, 本文提出了 CKL-EFS. 在结构演化机制中, CKL-EFS 通过数据密度动态监测样本空间结构的变化, 生成

不同的数据云, 并且通过相应规则控制内部样本数量增长. 在参数更新机制中, 根据样本信息, 结合核学习策略和 MTGC 准则实现数据云内的参数更新. 在不同类型和强度的非高斯噪声数据集上的实验结果表明, 所提 CKL-EFS 的鲁棒性并非依赖于特定噪声分布, 而是能够在非高斯性逐步增强的情况下保持稳定的预测行为, 并且在不同领域中具有良好的泛化性能. 未来研究将从以下层面开展: 1) 将 CKL-EFS 与深度学习的高维表征学习模块深度融合, 强化模型的特征提取能力; 2) 引入滑动窗口策略进一步降低在线更新的计算成本.

参考文献

- Alley R B, Emanuel K A, Zhang F Q. Advances in weather prediction: Better weather and environmental forecasting will continue to improve well-being. *Science*, 2019, **363**(6425): 342–344
- Li D C, Han M, Wang J. Chaotic time series prediction based on a novel robust echo state network. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(5): 787–799
- Chen Y S, Yu S S, Lim C P, Shi P. Multi-objective estimation of optimal prediction intervals for wind power forecasting. *IEEE Transactions on Sustainable Energy*, 2024, **15**(2): 974–985
- Li K, Príncipe J C. Transfer learning in adaptive filters: The nearest instance centroid-estimation kernel least-mean-square algorithm. *IEEE Transactions on Signal Processing*, 2017, **65**(24): 6520–6535
- Han M, Zhang S H, Xu M L, Qiu T, Wang N. Multivariate chaotic time series online prediction based on improved kernel recursive least squares algorithm. *IEEE Transactions on Cybernetics*, 2018, **49**(4): 1160–1172
- Zanin M, Olivares F. Ordinal patterns-based methodologies for distinguishing chaos from noise in discrete time series. *Communications Physics*, 2021, **4**(1): Article No. 190
- Wang Y S, Xu X H, Hu L, Fan J C, Han M. A time series continuous missing values imputation method based on generative adversarial networks. *Knowledge-Based Systems*, 2024, **283**: Article No. 111215
- Dubayan M R, Ershadi-Nasab S, Zomorodi M, Plawiak P, Tadeusiewicz R, Roui M B. Automated classification of thyroid disease using deep learning with neuroevolution model training. *Engineering Applications of Artificial Intelligence*, 2025, **146**: Article No. 110209
- Luo T Y, Liu M, Shi P, Duan G R, Cao X B. A hybrid data preprocessing-based hierarchical attention BiLSTM network for remaining useful life prediction of spacecraft lithium-ion batteries. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(12): 18076–18089
- Xu X H, Ren W J. A hybrid model of stacked autoencoder and modified particle swarm optimization for multivariate chaotic time series forecasting. *Applied Soft Computing*, 2022, **116**: Article No. 108321
- Angelov P P, Filev D P. An approach to online identification of Takagi-Sugeno fuzzy models. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, **34**(1): 484–498
- Yu H, Lu J, Zhang G Q. Topology learning-based fuzzy random neural networks for streaming data regression. *IEEE Transactions on Fuzzy Systems*, 2022, **30**(2): 412–425
- Gu X W, Han J G, Shen Q, Angelov P P. Autonomous learning for fuzzy systems: A review. *Artificial Intelligence Review*, 2023, **56**(8): 7549–7595
- Pratama M, Lu J, Lughofer E, Zhang G Q, Er M J. An incremental learning of concept drifts using evolving type-2 recurrent fuzzy neural networks. *IEEE Transactions on Fuzzy Systems*, 2017, **25**(5): 1175–1192
- Ge D J, Zeng X J. A self-evolving fuzzy system which learns dynamic threshold parameter by itself. *IEEE Transactions on Fuzzy Systems*, 2019, **27**(8): 1625–1637
- Hu L, Xu X H, Ren W J, Han M. Hierarchical evolving fuzzy system: A method for multidimensional chaotic time series online prediction. *IEEE Transactions on Fuzzy Systems*, 2024, **32**(6): 3329–3341
- Angelov P P, Gu X W, Príncipe J C. Autonomous learning multimodel systems from data streams. *IEEE Transactions on Fuzzy Systems*, 2018, **26**(4): 2213–2224
- Hu L, Xu X H, Liu J W, Yan X H, Han M. MIIPSO-EFS: Learning system with self-optimized parameters for chaotic time series online prediction. *Knowledge-Based Systems*, 2025, **310**: Article No. 112878
- Pratama M, Lu J, Zhang G Q. Evolving type-2 fuzzy classifier. *IEEE Transactions on Fuzzy Systems*, 2016, **24**(3): 574–589
- Mei Z, Zhao T, Gu X W. A dynamic evolving fuzzy system for streaming data prediction. *IEEE Transactions on Fuzzy Systems*, 2024, **32**(8): 4324–4337
- Gu X W. Multilayer ensemble evolving fuzzy inference system. *IEEE Transactions on Fuzzy Systems*, 2021, **29**(8): 2425–2431
- Alves K S T R, de Aguiar E P. A novel rule-based evolving fuzzy system applied to the thermal modeling of power transformers. *Applied Soft Computing*, 2021, **112**: Article No. 107764
- Shi L, Zhao H Q, Zakharov Y. An improved variable kernel width for maximum correntropy criterion algorithm. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2020, **67**(7): 1339–1343
- Duarte M S, Barreto G A. Fully adaptive dictionary for online correntropy kernel learning using proximal methods. *Expert Systems With Applications*, 2021, **178**: Article No. 114976
- Rong H J, Yang Z X, Wong P K. Robust and noise-insensitive recursive maximum correntropy-based evolving fuzzy system. *IEEE Transactions on Fuzzy Systems*, 2020, **28**(9): 2277–2284
- Wang F, He Y C, Wang S Y, Chen B D. Maximum total correntropy adaptive filtering against heavy-tailed noises. *Signal Processing*, 2017, **141**: 84–95
- Shi L, Shen L, Chen B D. An efficient parameter optimization of maximum correntropy criterion. *IEEE Signal Processing Letters*, 2023, **30**: 538–542
- Wang J, Lu L, Shi L, Zhu G Y, Yang X M. Euclidean direction search algorithm based on maximum correntropy criterion. *IEEE Signal Processing Letters*, 2023, **30**: 1032–1036
- Bao R J, Rong H J, Angelov P P, Chen B D, Wong P K. Correntropy-based evolving fuzzy neural system. *IEEE Transactions on Fuzzy Systems*, 2018, **26**(3): 1324–1338
- Han M, Xia H J, Ren W J. Kernel general loss algorithm based on evolving participatory learning for online time series prediction. *Engineering Applications of Artificial Intelligence*, 2023, **124**: Article No. 106600
- Shen T Y, Ren W J, Han M. Quantized generalized maximum correntropy criterion based kernel recursive least squares for online time series prediction. *Engineering Applications of Artificial Intelligence*, 2020, **95**: Article No. 103797
- Wu Z Z, Shi J H, Zhang X, Ma W T, Chen B D. Kernel recursive maximum correntropy. *Signal Processing*, 2015, **117**: 11–16
- Lv S H, Zhao H Q, Zhou L J. Maximum mixture total correntropy adaptive filtering against impulsive noises. *Signal Processing*, 2021, **189**: Article No. 108236
- He J C, Wang G, Zhang X, Wang H W, Peng B. Maximum total generalized correntropy adaptive filtering for parameter es-

- timation. *Signal Processing*, 2023, **203**: Article No. 108787
- 35 Shen P C, Li C G. Minimum total error entropy method for parameter estimation. *IEEE Transactions on Signal Processing*, 2015, **63**(15): 4079–4090
- 36 Zhu Y Y, Zhao H Q, Zeng X P, Chen B D. Robust generalized maximum correntropy criterion algorithms for active noise control. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020, **28**: 1282–1292
- 37 Angelov P. Fuzzily connected multimodel systems evolving autonomously from data streams. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2011, **41**(4): 898–910
- 38 Angelov P P, Gu X W, Príncipe J C. A generalized methodology for data analysis. *IEEE Transactions on Cybernetics*, 2018, **48**(10): 2981–2993
- 39 Gu X W, Ni Q, Tang G L. A novel data-driven approach to autonomous fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 2022, **30**(6): 2073–2085
- 40 Gu X W, Shen Q, Angelov P P. Particle swarm optimized autonomous learning fuzzy system. *IEEE Transactions on Cybernetics*, 2021, **51**(11): 5352–5363
- 41 Wang Q X, Jiang L, Yan L S, He X C, Feng J C, Pan W, et al. Chaotic time series prediction based on physics-informed neural operator. *Chaos, Solitons & Fractals*, 2024, **186**: Article No. 115326



胡磊 大连理工大学控制科学与工程学院博士研究生. 主要研究方向为时间序列在线预测.

E-mail: hl666888@mail.dlut.edu.cn

(**HU Lei** Ph.D. candidate at the School of Control Science and Engineering, Dalian University of

Technology. His main research interest is online time series prediction.)



许星晗 大连理工大学建设工程学院博士研究生. 主要研究方向为水文时间序列分析及建模. 本文通信作者.

E-mail: xuxh2023@dlut.edu.cn

(**XU Xing-Han** Ph.D. candidate at the School of Infrastructure Engineering, Dalian University of Techno-

logy. His research interests include hydrological time series analysis and modeling. Corresponding author of this paper.)



刘建卫 大连理工大学建设工程学院副教授. 主要研究方向为水文系统机理分析及数据建模.

E-mail: jwliu@dlut.edu.cn

(**LIU Jian-Wei** Associate professor at the School of Infrastructure Engineering, Dalian University of

Technology. His research interests include hydrological system mechanism analysis and data modeling.)



韩敏 大连理工大学控制科学与工程学院教授. 主要研究方向为复杂系统建模及时间序列预测.

E-mail: minhan@dlut.edu.cn

(**HAN Min** Professor at the School of Control Science and Engineering, Dalian University of Technology.

Her research interests include complex system modeling and time series prediction.)