

二氧化氮浓度时空预测：一种区间二型直觉模糊神经网络方法

赵亮¹, 李梦威¹, 郑玉卿¹, 崔贝贝¹, 朱献超²

(1. 河南工业大学电气工程学院, 河南 郑州 450001;

2. 河南工业大学人工智能与大数据学院, 河南 郑州 450001)

摘要: 空气中二氧化氮浓度的高低对环境保护和公共健康具有重要影响。目前二氧化氮浓度预测方法在表征时空关联性方面存在不足。鉴于此, 提出了新的使用区间二型直觉模糊神经网络时空预测二氧化氮浓度的方法。首先, 阐述了该区间二型直觉模糊神经网络框架, 引入可变系数加权其隶属部分和非隶属部分的输出, 并采用随机向量泛函链接神经网络作为规则后件; 然后, 为确定网络结构和参数, 采用分层聚类算法得到模糊规则库, 并通过最小二乘法优化网络后件的输出权值; 最后, 使用2018年1月至3月采集的北京市二氧化氮浓度真实数据进行数值验证。实验结果表明, 与现有方法相比, 该方法在短期和长期时空预测方面均取得了较高的预测精度和效率。

关键词: 二氧化氮浓度时空预测; 区间二型直觉模糊神经网络; 结构辨识; 参数优化; 最小二乘法

中图分类号: TP399

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202425

Spatiotemporal prediction of nitrogen dioxide concentration: an interval type-2 intuitionistic fuzzy neural network approach

ZHAO Liang¹, LI Mengwei¹, ZHENG Yuqing¹, CUI Beibei¹, ZHU Xianchao²

1. College of Electrical Engineering, Henan University of Technology, Zhengzhou 450001, China

2. School of Artificial Intelligence and Big Data, Henan University of Technology, Zhengzhou 450001, China

Abstract: The concentration of nitrogen dioxide (NO₂) in the air significantly affects environmental protection and public health. Current methods for NO₂ concentration prediction lack sufficient characterization of spatiotemporal correlations. Therefore, this study proposes a novel approach using interval type-2 intuitionistic fuzzy neural networks (IT2IFNNs) for spatiotemporal prediction of NO₂ concentrations. Firstly, the framework of IT2IFNNs is elucidated, incorporating variable coefficient weighting for its membership and non-membership outputs, and employing a random vector functional-link neural network (RVFLNN) as the rule consequent. Secondly, a hierarchical clustering algorithm is employed to determine the fuzzy rule base and optimize the output weight values of the network consequents using least squares estimation. Finally, numerical validation is conducted using real NO₂ concentration data collected in Beijing from January to March 2018. Experimental results demonstrate that compared to existing methods, the proposed approach achieves superior prediction accuracy and efficiency in both short-term and long-term spatiotemporal prediction tasks.

Key words: nitrogen dioxide concentration spatiotemporal prediction, interval type-2 intuitionistic fuzzy neural networks, structural identification, parameter optimization, least squares method

收稿日期: 2024-04-07; 修回日期: 2024-05-23

通信作者: 赵亮, zhaoliang270@gmail.com

基金项目: 国家自然科学基金项目 (No.61473114); 河南省科技厅科技攻关项目 (No.242102211040); 河南省教育厅科学技术重点研究项目 (No.24B520006, No.24A520014)

Foundation Items: The National Natural Science Foundation of China (No.61473114), The Science and Technology Research Project of Henan Provincial Science and Technology Department (No.242102211040), The Key Scientific and Technological Research Project of Henan Provincial Department of Education (No.24B520006, No.24A520014)

0 引言

随着世界各国对空气污染问题的日益关注, 空气质量监测与污染治理已逐渐成为工业界与学术界的热点研究领域。根据国际标准化组织(ISO)的相关定义, 户外空气污染物主要包括可吸入颗粒物、二氧化硫、氮氧化物及一氧化碳等。这些污染物的构成会受到诸如气象状况、时间(包括一天中的不同时段及一周中的具体日期)、工业活动和交通流量等多重因素的显著影响, 从而产生大幅度的波动。在这些污染物中, 二氧化氮对动植物, 尤其是人类健康构成了严重威胁。因此快速且准确地预测二氧化氮的浓度, 对空气质量的有效监测与污染治理具有举足轻重的意义。

近几十年来, 二氧化氮浓度的预测问题已然成为学术研究的热点, 引发了众多学者的浓厚兴趣。伴随人工智能技术的崛起, 包括模糊理论和神经网络在内的多种智能技术被广泛应用于解决这类预测难题。Tao等^[1]利用2014至2020年中国1 609个空气质量监测站的数据, 构建了一个集成机器学习模型, 以实现二氧化氮浓度的准确预测。Muthukrishnan等^[2]基于2021年和2022年的气象观测数据, 建立了自回归模型, 用于预测印度泰米尔纳德邦纳盖科伊尔的日平均二氧化氮浓度。此外, 文献[3-6]也从不同角度对二氧化氮浓度的预测进行了深入探讨。然而, 这些研究存在预测结果难以解释的局限。使用模糊神经网络预测二氧化氮浓度是解决以上问题的有效途径之一。值得注意的是, 模糊神经网络作为一种已经发展了几十年的预测工具, 一型模糊神经网络的预测精度不如二型模糊神经网络。正如文献[7]所言: “在很多应用中, 我们都发现区间二型模糊系统的性能比一型要好, 但其原因尚无定论。”自Zadeh引入二型模糊集合概念以来, 基于二型模糊神经网络的预测方法在众多领域备受推崇并得到了广泛应用。Abiyev等^[8]创造性地将二型模糊系统与神经网络融合, 构建了二型模糊神经网络, 并将其应用于胎儿健康状况的检测。Zhao^[9]采用区间二型模糊神经网络对短期交通流进行预测, 并通过与其他预测方法的对比实验, 充分证实了其预测策略的准确性。Bai等^[10]提出了一种全新的数据-知识驱动的区间二型模糊神经网络, 该方法利用关联规则挖掘技术自动探索和利用具有可解释性的规则知识, 旨在提升模型的可解释性与泛化能

力。此外, Qiao^[11]提出的新的区间二型模糊神经网络被专门用于氮氧化物排放的预测。在经过基准模拟评估后, 其提出的模型在氮氧化物排放预测方面充分展现了其有效性和优越性。另外, Liu等^[12]提出了一种综合学习方法, 旨在优化区间二型模糊神经网络的结构和自由参数。该方法被成功应用于股票价格指数的预测, 显示了其实用性和准确性。与之类似, 赵亮等^[13]提出一种适用于多场景的数据驱动无线信号传输模型。该模型结合了深度学习的优势和区间二型模糊推理的方法, 能够有效地提高无线信号传输预测的准确性和适用性。Sun等^[14]开发了一种基于信息导向的模块化二型模糊神经网络, 其模型在高维复杂系统预测的可解释性可以与先进的方法相媲美。Liu等^[15]提出了一种用于预测和识别非线性系统的自组织区间二型模糊小波神经网络模型, 并将其应用到短期交通流量预测、麦基-格拉斯时间序列预测和上证指数开盘指数预测。与同类评估相比, 其具有更高的预测精度和速度。尽管区间二型模糊神经网络已经得到了广泛的研究和应用, 然而由于其前件和后件存在一定的局限性, 其应用也受到了制约。

1999年, Atanassov等^[16]考虑隶属函数的犹豫度, 首次阐述了直觉模糊集(intuitionistic fuzzy sets, IFS)及其基本性质, 这一理论随后在模糊决策等领域得到了广泛应用。在此基础上, 考虑隶属函数的不确定因素, 二型直觉模糊集^[17]应运而生, 其特点在于非隶属函数(non-membership function, NMF)和隶属函数(membership function, MF)的值分别被定义为单位区间内的一型模糊集。进一步地, 作为它的一种特殊情况, 区间二型直觉模糊集被重点研究并应用于模糊决策中。近年来, 直觉模糊推理与神经网络的结合逐渐成为模糊理论研究者关注的焦点。Li等^[18]将直觉模糊集与Hopfield神经网络结合, 进一步推导并验证了该结合方案的稳定性条件。Hájek等^[19]详细阐述了TSK型直觉模糊神经网络的模糊推理过程和学习方案, 并运用此方案为众多美国风险投资公司进行信用评分, 为潜在投资者提供了有价值的参考。

本文提出一种新的区间二型直觉模糊神经网络(interval type-2 intuitionistic fuzzy neural network, IT2IFNN), 旨在预测空气中的二氧化氮浓度。该方法的主要特色和贡献如下:

(1) 构建了新颖的区间二型直觉模糊神经网络

框架。在区间二型直觉模糊推理的核心环节中, 模糊降型被视为得到确定输出的关键步骤。通过引入可变系数, 对隶属部分和非隶属部分的输出进行加权处理, 有效规避了使用传统的 Karnik-Mendel (KM) 算法。更重要的是, 采用了随机向量泛函链接神经网络 (random vector functional-link neural network, RVFLNN) [20] 作为该网络的后件, 显著提升了系统对复杂非线性函数的逼近能力。

(2) 设计了全新的区间二型直觉模糊神经网络学习方案。本文遵循数据驱动的设计理念, 通过分层聚类算法对输入空间进行合理划分, 从而确定了模糊规则库的大小及部分前件初始参数。进一步采用广义逆矩阵理论和最小二乘法, 精确得到后件 RVFLNN 的输出权值。

(3) 实现了对二氧化氮浓度的时空预测。由于大气的开放特性, 二氧化氮的浓度会随着时间和空间的变化而波动。本文方法不仅充分考虑了时间因素, 更融合了空间状态, 从而使得预测结果更加准确和可靠, 能够满足对二氧化氮浓度的短期和长期预测需求。

本文结构如下: 第1节介绍了新颖的区间二型直觉模糊神经网络框架; 第2节介绍区间二型直觉模糊神经网络的学习方法; 第3节为实验分析; 第4节为总结。

1 新颖的区间二型直觉模糊神经网络框架

Atanassov 等 [16] 引入独立的 NMF, 从而提出了 IFS 的概念。此后, 他们在区间值模糊集合的启发下, 进一步推广了 IFS 的定义, 并对区间值直觉模糊集合 [21] 进行了深入研究。另外, 广义形式的二型模糊集合和系统被深入研究。在此基础上, 对直觉二型模糊推理的形式化进行了初步探讨。区间二型直觉模糊集合作为它的一种简化版本, 在理论和应用上都起着更为重要的作用。

1.1 区间二型直觉模糊逻辑系统

首先, 给出二型直觉模糊集合 (type-2 intuitionistic fuzzy set, T2IFS) 的定义。

定义1 X 为论域, $\text{Supp}_x^{\text{mf}}$ 为次 MF 和次 NMF 的支集, 对 $\forall x \in X$, $\forall u \in \text{Supp}_x^{\text{mf}}$, $\forall v \in \text{Supp}_x^{\text{nmf}}$, T2IFS $\tilde{A}^I$ 表示为:

$$\tilde{A}^I = \int_{x \in X} \int_{u \in \text{Supp}_x^{\text{mf}}} \int_{v \in \text{Supp}_x^{\text{nmf}}} \frac{\langle \mu_{\tilde{A}^I}(x, u), \nu_{\tilde{A}^I}(x, v) \rangle}{(x, u, v)} (1)$$

式中, $\text{Supp}_x^{\text{mf}} \subseteq [0, 1]$, $\text{Supp}_x^{\text{nmf}} \subseteq [0, 1]$, $\mu_{\tilde{A}^I}(x, u) \in [0, 1]$, $\nu_{\tilde{A}^I}(x, v) \in [0, 1]$ 且 $v \cdot \mu_{\tilde{A}^I}(x, u) + u \cdot \nu_{\tilde{A}^I}(x, v) \in [0, 1]$ 。由式 (1) 可以看出, 它比二型模糊集合具有更高的计算复杂度。因此, 为了研究和应用方便, 定义2给出了它的简化版本。

定义2 X 为论域, $\text{Supp}_x^{\text{mf}}$ 和 $\text{Supp}_x^{\text{nmf}}$ 为次 MF 和次 NMF 的支集, 对 $\forall x \in X$, $\forall u \in \text{Supp}_x^{\text{mf}}$, $\forall v \in \text{Supp}_x^{\text{nmf}}$, $\mu(x, u) \equiv 1$ 且 $\nu(x, v) \equiv 1$, T2IFS 简化为区间二型直觉模糊集合 (interval type-2 intuitionistic fuzzy set, IT2IFS), 即:

$$\tilde{A}^I = \int_{x \in X} \int_{u \in \text{Supp}_x^{\text{mf}}} \int_{v \in \text{Supp}_x^{\text{nmf}}} \frac{\langle 1, 1 \rangle}{(x, u, v)} (2)$$

式中, $0 \leq u + v \leq 1$ 。

当 $\text{Supp}_x^{\text{mf}}$ 和 $\text{Supp}_x^{\text{nmf}}$ 是单位区间的闭子区间时, 区间值直觉模糊集合实际上是 IT2IFS 的一个特例。本文对它们不作区分。

下面详细阐述了多输入单输出单值区间二型直觉模糊逻辑系统的推理过程。数据样本为 $\{\mathbf{X}^i, \mathbf{y}^i\}_{i=1}^N$, 其中 $\mathbf{X}^i \in \mathbf{R}^n$, $\mathbf{y}^i \in \mathbf{R}$, N 是样本数目。模糊规则表示如下。

R^j : 如果 x_1 是 $\tilde{A}_{1j}^I$ 且 x_n 是 $\tilde{A}_{nj}^I$, 那么 $y = \sum_{k=1}^n c_{kj} x_k + c_{0j}$, $j=1, \dots, M$ 。

式中, $\tilde{A}_{1j}^I, \dots, \tilde{A}_{nj}^I$ 表示规则的前件, c_{kj} 是规则后件系数, M 为模糊规则的数目。在上述模糊系统中, 模糊规则的前件部分为 IT2IFSs, 它们的 MF 和 NMF 表示为:

$$\bar{\mu}_{\tilde{A}_{kj}^I}(x_k) = \exp \left(- \left(\frac{x_k - m_{kj}}{\bar{\sigma}_{kj}} \right)^2 \right) (3a)$$

$$\underline{\mu}_{\tilde{A}_{kj}^I}(x_k) = \exp \left(- \left(\frac{x_k - m_{kj}}{\underline{\sigma}_{kj}} \right)^2 \right) (3b)$$

$$\bar{\nu}_{\tilde{A}_{kj}^I}(x_k) = s_{kj}(x_k) \left(1 - \exp \left(- \left(\frac{x_k - m_{kj}}{\bar{\sigma}_{kj}} \right)^2 \right) \right) (3c)$$

$$\underline{\nu}_{\tilde{A}_{kj}^I}(x_k) = s_{kj}(x_k) \left(1 - \exp \left(- \left(\frac{x_k - m_{kj}}{\underline{\sigma}_{kj}} \right)^2 \right) \right) (3d)$$

其中, m_{kj} 是高斯函数的中心参数, $[\underline{\sigma}_{kj}, \bar{\sigma}_{kj}]$ 是高斯函数的半径参数, $s_{kj}(\cdot) \in [0, 1]$ 是尺度函数, 可用 sigmoid 型函数、双曲正切函数等表示。将清晰的 X

输入模糊系统, 计算前件激活区间 $f_j(X)$ 和 $g_j(X)$ 的下界和上界为:

$$f_{-j}(X) = \min_{k=1}^n \left\{ \underline{\mu}_{\tilde{A}_{kj}}(x_k) \right\} \quad (4a)$$

$$\bar{f}_j(X) = \min_{k=1}^n \left\{ \bar{\mu}_{\tilde{A}_{kj}}(x_k) \right\} \quad (4b)$$

$$\underline{g}_j(X) = \max_{k=1}^n \left\{ \underline{\nu}_{\tilde{B}_{kj}}(x_k) \right\} \quad (4c)$$

$$\bar{g}_j(X) = \max_{k=1}^n \left\{ \bar{\nu}_{\tilde{B}_{kj}}(x_k) \right\} \quad (4d)$$

基于式(4), 进一步计算归一化的 $f_j^n(X)$ 、 $\bar{f}_j^n(X)$ 、 $\underline{g}_j^n(X)$ 和 $\bar{g}_j^n(X)$:

$$f_j^n(X) = \frac{f_{-j}(X)}{\sum_{j=1}^M f_{-j}(X)} \quad (5a)$$

$$\bar{f}_j^n(X) = \frac{\bar{f}_j(X)}{\sum_{j=1}^M \bar{f}_j(X)} \quad (5b)$$

$$\underline{g}_j^n(X) = \frac{\underline{g}_j(X)}{\sum_{j=1}^M \underline{g}_j(X)} \quad (5c)$$

$$\bar{g}_j^n(X) = \frac{\bar{g}_j(X)}{\sum_{j=1}^M \bar{g}_j(X)} \quad (5d)$$

最后, 清晰的输出 y 可计算为:

$$y = \sum_{j=1}^M \left(m f_j^n(X) + n \bar{f}_j^n(X) \right) \left(\sum_{k=1}^n (c^{\text{mf}})^j_k x_k + (c^{\text{mf}})^j_0 \right) + \left(p \underline{g}_j^n(X) + q \bar{g}_j^n(X) \right) \left(\sum_{k=1}^n (c^{\text{nmf}})^j_k x_k + (c^{\text{nmf}})^j_0 \right) \quad (6)$$

其中, $m+n+p+q=1$, $m, n, p, q \in \mathbf{R}$, $(c^{\text{mf}})^j_k$ 和 $(c^{\text{nmf}})^j_k$, $k=0, 1, \dots, n$ 是后件系数。

1.2 新颖的区间二型直觉模糊神经网络

在TSK模糊逻辑系统中, 模糊规则的后件部分表示为输入向量的线性函数, 这一设计虽然简洁, 但也在一定程度上限制了它对复杂非线性函数的逼近能力。为了弥补这一不足, 本文提出一种新颖的方法, 即采用RVFLNN代替传统的线性函数, 作为模糊规则的后件部分。网络结构示意图如图1所示。

图1中的网络由7层神经元节点构成, 各层分别执行不同的传输与计算功能。第1层为输入层,

主要作用是将输入向量传输至下一层, 不进行任何计算处理。接着, 第2层是MF/NMF层, 其主要任务是根据式(3)计算IT2 IFSs的MF或NMF值。第3层则是对前一层的输出进行 t -范数/ s -范数的运算, 如式(4)所示, 采用最小算子作为 t -范数的数学实现, 而最大算子则作为 s -范数的实现方式。第4层负责对模糊规则的激活区间进行归一化处理, 具体的计算过程如式(5)所示。然后, 在第5层, 每条规则的激活区间上界和下界的归一化值与后件部分(RVFLNN)的输出进行相乘运算。在第6层中, 通过对前一层MF/NMF部分的结果分别进行求和运算。最后, 如式(6)所示, 使用权值 m 、 n 、 p 、 q 分别对MF/NMF部分的上界或下界结果加权求和, 得到最终的输出结果。

2 区间二型直觉模糊神经网络的学习方法

第1节阐述了整个网络架构及信号传输过程, 然而, 合理确定网络的结构和参数仍是一个具有挑战性的问题。因此, 本文提出了一种有效的学习方案, 包括结构辨识和参数优化2个步骤。

2.1 结构辨识

迄今为止, 已探索并发展了多种神经网络结构识别方法, 如论域划分、聚类等。然而, 在采用论域划分法进行网络结构识别时, 随着输入向量维数的增加, 模糊规则的数量会呈现指数级增长, 这在很大程度上限制了它在实际应用中的可行性。近年来, 基于聚类的方法逐渐成为研究热点, 并在该领域得到了广泛应用。本文采用分层聚类方法识别该网络的结构, 实现对输入空间的合理划分, 该方法结合了减法聚类和模糊 c 均值聚类的优点。首先, 利用基于密度的减法聚类方法确定规则库的大小。通过调整其关键参数, 即影响范围因子, 可以指定更多更小的聚类, 从而构建更大规模的模糊规则库; 然后, 采用模糊 c 均值聚类方法初始化MFs/NMFs中心。这一步骤使得每个输入向量都能以一定的隶属度归属于多个聚类, 从而更全面地反映输入数据的内在结构; 最后, 根据隶属度值的大小, 将输入向量归属到隶属度最大的簇中。方差的上界和下界由下式确定:

$$\bar{\sigma}_{kj} = \text{cf}_{kj} \cdot \max \left\{ m_{kj} - \min_{i=1}^N X_k^i, \max_{i=1}^N X_k^i - m_{kj} \right\} \quad (7a)$$

$$\underline{\sigma}_{kj} = \frac{\bar{\sigma}_{kj}}{\text{uc}_{kj}} \quad (7b)$$

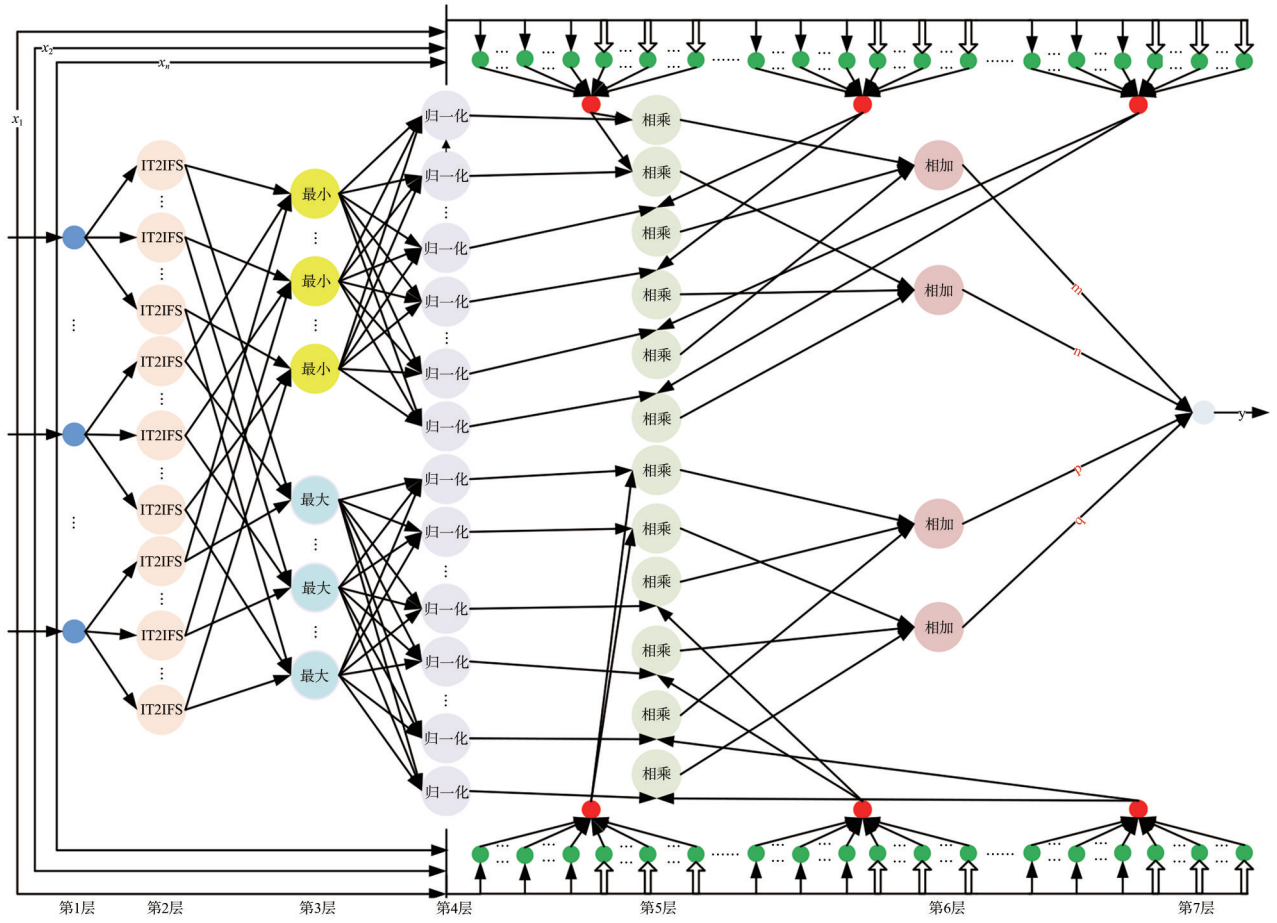


图1 网络结构示意图

其中, cf_{kj} 和 uc_{kj} 是待定系数, $uc_{kj} \geq 1$, $j=1, 2, \dots, M$, $k=1, 2, \dots, n$ 。

此外, NMF 的尺度函数表示如下:

$$s(x_k) = \begin{cases} 1, & x_k > m_{kj} + \delta_{kj} \\ \frac{4}{\delta_{kj}^2} \left(x_k - m_{kj} - \frac{\delta_{kj}}{2} \right)^2, & m_{kj} + \frac{\delta_{kj}}{2} < x_k \leq m_{kj} + \delta_{kj} \\ -\frac{4}{\delta_{kj}^2} \left(x_k - m_{kj} \right)^2 + 1, & m_{kj} - \frac{\delta_{kj}}{2} \leq x_k \leq m_{kj} + \frac{\delta_{kj}}{2} \\ \frac{4}{\delta_{kj}^2} \left(x_k - m_{kj} + \frac{\delta_{kj}}{2} \right)^2, & m_{kj} - \delta_{kj} \leq x_k < m_{kj} - \frac{\delta_{kj}}{2} \\ 1, & x_k < m_{kj} - \delta_{kj} \end{cases} \quad (8)$$

其中, $\delta_{kj} = \sigma_{kj}$ 。经过结构辨识, 确定了模糊规则库的大小和前件的初始参数。

2.2 参数优化

自从神经网络和模糊系统提出以来, 其参数学习问题便备受研究者的关注。一般而言, 参数学习

方法可分为基于梯度的学习和基于非梯度的学习两大类。在基于梯度的方法中, 随机梯度下降法及其变体、卡尔曼滤波及其变体、最小二乘法及其变体是常见的策略。然而, 随机梯度下降法易陷入局部极值, 这在全局优化前件自由参数时显得尤为不利。此外, 随着待优化参数的增多, 卡尔曼滤波方法需要计算更高维度的协方差矩阵, 这不仅会显著增加计算成本, 还可能降低其效率。另外, 基于非梯度的方法主要包括自然启发算法、遗传启发算法和物理启发算法, 如遗传算法及其变体、粒子群算法及其变体、引力搜索算法及其变体等。这些方法虽然具有一定的探索能力, 但通常收敛速度较慢, 且同样容易陷入局部极值。本文采用最小二乘法

对后件的自由参数进行优化。 $\{\mathbf{X}^i, \mathbf{y}^i\}_{i=1}^N$ 表示数据样本集, 激活力矩阵表示为 $F^n = \left(f_j^n(\mathbf{X}^i) \right)_{N \times M}$, $\bar{F}^n = \left(\bar{f}_j^n(\mathbf{X}^i) \right)_{N \times M}$, $\underline{G}^n = \left(\underline{g}_j^n(\mathbf{X}^i) \right)_{N \times M}$ 和 $\bar{G}^n = \left(\bar{g}_j^n(\mathbf{X}^i) \right)_{N \times M}$ 。

$(\bar{g}_j^n(\mathbf{X}^i))_{N \times M}$, $i=1, 2, \dots, N$, $j=1, 2, \dots, M$ 。每条规则的后件部分为RVFLNN, 记为 RN_{jj+M} 。每个

$$\mathbf{OM}_{jj+M} = \begin{bmatrix} x_1^1 & \cdots & x_k^1 & \cdots & x_n^1 & \text{EO}_1^{jj+M}(\mathbf{X}^1) & \cdots & \text{EO}_l^{jj+M}(\mathbf{X}^1) & \cdots & \text{EO}_{\text{EH}_{jj+M}}^{jj+M}(\mathbf{X}^1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^i & \cdots & x_k^i & \cdots & x_n^i & \text{EO}_1^{jj+M}(\mathbf{X}^i) & \cdots & \text{EO}_l^{jj+M}(\mathbf{X}^i) & \cdots & \text{EO}_{\text{EH}_{jj+M}}^{jj+M}(\mathbf{X}^i) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^N & \cdots & x_k^N & \cdots & x_n^N & \text{EO}_1^{jj+M}(\mathbf{X}^N) & \cdots & \text{EO}_l^{jj+M}(\mathbf{X}^N) & \cdots & \text{EO}_{\text{EH}_{jj+M}}^{jj+M}(\mathbf{X}^N) \end{bmatrix} \quad (9)$$

其中, $\text{EO}_l^{jj+M}(\mathbf{X}^i) = f(\ln \mathbf{W}_l^{jj+M} \mathbf{X}^i + b_l^{jj+M})$ 。 $\text{EO}_l^{jj+M}(\mathbf{X}^i)$ 为第 l 个增强节点的输出, $f(\cdot)$ 为激活函数, $\ln \mathbf{W}_l^{jj+M}$ 为连接第 l 个增强节点的输入权重向量, b_l^{jj+M} 为其偏置, 它们是均匀分布在区间 $[-1, 1]$ 的随机数。其中, $\mathbf{OutW}^{jj+M} = [\text{outw}_1^{jj+M}, \dots, \text{outw}_l^{jj+M}, \dots, \text{outw}_{n+\text{EH}_{jj+M}}^{jj+M}]$ 为 RN_{jj+M} 的输出权重值向量。因此有线性方程如下:

$$\mathbf{Y} = \mathbf{WOM} \cdot \mathbf{OutW}^T \quad (10)$$

式中, $\mathbf{OutW} = [\mathbf{OutW}^1, \dots, \mathbf{OutW}^j, \dots, \mathbf{OutW}^{2M}]$, 以及 $\mathbf{WOM} = [\mathbf{WOM}_1, \dots, \mathbf{WOM}_j, \dots, \mathbf{WOM}_{2M}]$ 。 $\mathbf{WOM}_j = \left((m \bar{F}^n(i, j) + n \bar{F}^n(i, j)) \times \mathbf{OM}_j(i, 1:n+\text{EH}_j) \right)_{N \times (n+\text{EH}_j)}$,

以及 $\mathbf{WOM}_{j+M} = \left((p \bar{G}^n(i, j) + q \bar{G}^n(i, j)) \times \mathbf{OM}_{j+M}(i, 1:n+\text{EH}_{j+M}) \right)_{N \times (n+\text{EH}_{j+M})}$, 其中 m, n, p, q 是 $[0, 1]$

取值的随机数且它们的和为1。利用 Moore-Penrose 逆, 得到 $\mathbf{OutW} = (\mathbf{WOM}^+ \mathbf{Y})^T$ 。与反向传播算法相比, 最小二乘法具有更高的效率。

3 实验分析

本文以北京市为研究对象, 运用本文所构建的预测方法, 对北京市特定区域的二氧化氮浓度进行预测。

3.1 数据准备和预处理

本文使用的数据集来源于北京市环境监测中心, 收集时间为2018年1月1日至2018年3月31日。目前, 北京市设有35个空气质量监测站, 这些监测站分布在城市的各个角落, 以全面监测空气质量状况。值得注意的是, 二氧化氮浓度随空间分布呈现出复杂的变化规律, 因此其他监测站所收集的数据可能会对某一特定监测站的预测值产生影响。为了

RN_{jj+M} 的增强节点数为 EH_{jj+M} 。输出矩阵 $\mathbf{OM}_{jj+M}$ 为:

准确挖掘各监测站间的复杂关系, 获得更为可靠和合理的预测结果, 本文采用了时空联合预测的方法, 即利用邻近5个监测站的数据进行综合分析。具体而言, 通过综合考虑这些监测站的数据, 以捕捉二氧化氮浓度的时空变化规律, 从而提高预测的准确性。此外, 采集到的数据集由于监测仪器的故障或其他不可控因素, 存在部分数据缺失情况。在这种情况下, 可应用下面的表达式来估计缺失的数据:

$$\hat{C}_{\text{NO}_2}^{s_i}(s) = \frac{C_{\text{NO}_2}^{s_i}(t-\tau) + C_{\text{NO}_2}^{s_i}(t+\tau)}{2} \quad (11)$$

式中, $\hat{C}_{\text{NO}_2}^{s_i}(s)$ 为估计值, $s=t-\tau+1, \dots, t+\tau$ 。另外, 为了消除每日浓度的季节性变化对精确预报的影响, 采用了以下数据处理方法:

$$C_{\text{NO}_2}^{s_i}(h+24(r-1)) = C_{\text{NO}_2}^{s_i}(h+24(r-1)) - \frac{\sum_{r=1}^{d_n} C_{\text{NO}_2}^{s_i}(h+24(r-1))}{d_n} \quad (12)$$

式中, $h=1, 2, \dots, 24$ 。 d_n 表示天数。最后得到本实验使用的数据集。

3.2 短期预测

短期预测以其较高的预测精度, 在二氧化氮浓度预测问题中扮演至关重要的角色, 通常被称为一步预测。本文使用的二氧化氮浓度数据的采样周期是1小时, 故短期预测的时长为1小时。在此基础上, 短期预测模型可表达为:

$$C_{\text{NO}_2}^{s_i}(t+1) = f_{s_i} \left(C_{\text{NO}_2}^{s_i}(t-3), \dots, C_{\text{NO}_2}^{s_i}(t), \dots, C_{\text{NO}_2}^{s_i}(t-3), \dots, C_{\text{NO}_2}^{s_i}(t) \right) \quad (13)$$

式中, $C_{\text{NO}_2}^{s_i}(k)$ 表示第 s_i 个监测站在 k 时刻的二氧化氮浓度, $i=1, 2, 3, 4, 5$ 和 $k=t-3, t-2, t-1, t, t+1$ 。由式(13)可以看出, 考虑二氧化氮浓度的空间分布特性, 使用地理位置毗邻的东城东四 (s_1)、东城

天坛 (s_2)、西城官园 (s_3)、西城万寿西宫 (s_4) 和朝阳奥体中心 (s_5) 5 个监测站的二氧化氮浓度测量数据进行相应站点二氧化氮浓度的一步预测。整个预测过程分为 3 个步骤。

(1) 产生输入-输出数据集。首先, 使用式 (14) 将收集到的数据集归一化为 $[0, 1]$:

$$\tilde{C}_{\text{NO}_2}(t) = \frac{C_{\text{NO}_2}(t) - \min_{t=1}^{N_t} \{C_{\text{NO}_2}(t)\}}{\max_{t=1}^{N_t} \{C_{\text{NO}_2}(t)\} - \min_{t=1}^{N_t} \{C_{\text{NO}_2}(t)\}} \quad (14)$$

使用式 (13) 进一步处理数据, 得到输入-输出数据集, 即 $X = \left\{ \left(\tilde{C}_{\text{NO}_2}^{s_1}(t-3), \tilde{C}_{\text{NO}_2}^{s_1}(t-2), \tilde{C}_{\text{NO}_2}^{s_1}(t-1), \tilde{C}_{\text{NO}_2}^{s_1}(t-1), \tilde{C}_{\text{NO}_2}^{s_1}(t), \dots, \tilde{C}_{\text{NO}_2}^{s_5}(t-3), \tilde{C}_{\text{NO}_2}^{s_5}(t-2), \tilde{C}_{\text{NO}_2}^{s_5}(t-1), \tilde{C}_{\text{NO}_2}^{s_5}(t) \right) \right\}_{t=4}^{N_t-1}$, $Y = \left\{ \left(\tilde{C}_{\text{NO}_2}^{s_1}(t+1), \tilde{C}_{\text{NO}_2}^{s_2}(t+1), \dots, \tilde{C}_{\text{NO}_2}^{s_5}(t+1) \right) \right\}_{t=4}^{N_t-1}$ 。

(2) 通过 7 折交叉验证获取训练和测试数据集。对输入输出数据集 $\{X, Y\}$, 采用 7 折交叉验证的方法划分训练数据集和测试数据集。具体而言, 首先对 $\{X, Y\}$ 进行均匀采样, 将其平均分配为 7 个互不重叠的子集; 然后, 随机选取其中的 6 个子集作为训练数据集, 而将剩余的 1 个子集作为测试数据集。这一随机选择过程将不断重复进行, 直到每个子集都至少被用作一次测试数据集为止。

(3) 训练和验证预测模型。利用上述结构辨识和参数优化的学习算法设计预测方法, 对短期内的二氧化氮浓度进行预测。

短期预测实验结果如表 1 所示。

表 1 短期预测实验结果

预测方法	RMSE	
	训练数据集	测试数据集
最小二乘支持向量机 ^[22]	—	0.0305
极限学习机 ^[23]	—	0.0302
RVFLNN	0.0297	0.0302
本文方法	0.0287	0.0306

由表 1 可以看出, 由于训练过程通常是在理想条件下进行的, 所以训练结果的均方根误差 (root mean square error, RMSE) 普遍较好, 不能直接反映各方法在实际应用中的预测精度。但是, 训练过程中的表现仍然能反映出模型的学习能力和适应数据分布的能力。比较测试结果可以看出, 所有方法

的 RMSE 都较为接近, 这表明它们在应对新的二氧化氮浓度数据时, 能够准确地执行预测, 具有较高的泛化能力。这可能与充分考虑了二氧化氮浓度的空间分布特性, 并使用了多个监测站点的数据作为输入有关。值得注意的是, RVFLNN 在训练集上的 RMSE 为 0.0297, 但在测试集上的表现却与其他方法相近。而本文方法通过结构辨识和参数优化的学习算法, 在训练和测试结果上表现均衡。这为进一步探索和应用短期预测模型提供了有力的支持。

3.3 长期预测

与短期预测相比, 长期预测具有更多的随机干扰因素, 因此精度较低。本文的长期预测是指预测 24 小时后的二氧化氮浓度, 具体方法可以表示为:

$$C_{\text{NO}_2}^{s_i}(t+25) = f_{s_i} \left(C_{\text{NO}_2}^{s_1}(t-5), \dots, C_{\text{NO}_2}^{s_1}(t), \dots, C_{\text{NO}_2}^{s_5}(t-5), \dots, C_{\text{NO}_2}^{s_5}(t) \right) \quad (15)$$

对于长期预测, 不同站点间二氧化氮浓度空间分布特性更加复杂, 表现为不同站点的二氧化氮浓度测量数据的强耦合性, 从而导致式 (15) 中的预测函数 $f_{s_i}(\cdot)$ 相较短期预测构建更加困难。输入数据集 $X = \left\{ \left(\tilde{C}_{\text{NO}_2}^{s_1}(t-5), \dots, \tilde{C}_{\text{NO}_2}^{s_1}(t), \dots, \tilde{C}_{\text{NO}_2}^{s_5}(t-5), \dots, \tilde{C}_{\text{NO}_2}^{s_5}(t) \right) \right\}_{t=6}^{N_t-25}$, 输出数据集 $Y = \left\{ \left(\tilde{C}_{\text{NO}_2}^{s_1}(t+25), \dots, \tilde{C}_{\text{NO}_2}^{s_5}(t+25) \right) \right\}_{t=6}^{N_t-25}$; 然后, 使用 10 折交叉验证, 可以得到训练和测试数据集; 最后, 经过高效学习, 可以获得训练好的预测模型, 包括 3 条模糊规则和 6 个 RVFLNN, 每个 RVFLNN 有 10 个增强节点。长期预测实验结果如表 2 所示。

表 2 长期预测实验结果

预测方法	RMSE	
	训练数据集	测试数据集
最小二乘支持向量机 ^[22]	—	0.0970
极限学习机 ^[23]	—	0.1085
RVFLNN	0.1046	0.1085
本文方法	0.1019	0.1118

对比长期预测二氧化氮浓度的实验结果, 可以看出不同方法在处理这一复杂问题时所展现出的性能差异。首先, 长期预测由于存在更多的随机干扰因素, 因此整体精度普遍较低, 这与短期预测形成

了鲜明对比。由表2可以看出,最小二乘支持向量机在测试集上的RMSE为0.0970,表明该方法在一定程度上能捕捉到二氧化氮浓度的变化趋势。极限学习机、RVFLNN和本文方法通过10折交叉验证得到训练和测试结果,整体表现较为均衡。此外,本文方法结合了区间二型直觉模糊推理和RVFLNN,为长期预测提供了新的思路 and 实现途径。

4 结束语

作为大气污染防控领域的关键问题,二氧化氮浓度的精准预测一直备受关注。为了有效应对这一挑战,本文提出了一种新的区间二型直觉模糊神经网络预测方法,并将其创新性地应用于北京地区二氧化氮浓度短期与长期预测中。经过实验验证可以发现,相较于其他传统预测方法,本文提出的预测方法在性能上具有可比性,同时在效率上展现出了显著的优势。这一重要成果不仅为大气污染防控提供了有力的技术支持,也为相关领域的研究开辟了新的思路与方法。这一预测方法将在未来的实践中得到更广泛的应用,为改善空气质量、保护生态环境做出更加积极的贡献,推动大气污染防控工作不断向前发展。

参考文献:

- [1] TAO C L, JIA M, WANG G Q, et al. Time-sensitive prediction of NO₂ concentration in China using an ensemble machine learning model from multi-source data[J]. *Journal of Environmental Sciences (China)*, 2024, 137: 30-40.
- [2] MUTHUKRISHNAN P, SHARMA R K. A short-term autoregressive model for the prediction of daily average NO₂ concentration in nagercoil, Tamil Nadu, India[J]. *Nature Environment and Pollution Technology*, 2024, 23(1): 525-535.
- [3] ZHANG C X, LIU C, LI B, et al. Spatiotemporal neural network for estimating surface NO₂ concentrations over north China and their human health impact[J]. *Environmental Pollution*, 2022, 307: Article No. 119510.
- [4] DOMHNAILL A Ó, BRODERICK B, O'MAHONY M. Integrated transportation and land use regression modelling for nitrogen dioxide mitigation[J]. *Transportation Research Part D: Transport and Environment*, 2023, 115: Article No. 103572.
- [5] VERMA A, RANGA V, VISHWAKARMA D K. BREATH-Net: a novel deep learning framework for NO₂ prediction using bi-directional encoder with transformer[J]. *Environmental Monitoring and Assessment*, 2024, 196(4): 340.
- [6] LIU C H, WU S S, DAI Z, et al. High-resolution daily spatiotemporal distribution and evaluation of ground-level nitrogen dioxide concentration in the Beijing-Tianjin-Hebei region based on TROPOMI data[J]. *Remote Sensing*, 2023, 15(15): 3878.
- [7] 伍冬睿, 曾志刚, 莫红, 等. 区间二型模糊集和模糊系统: 综述与展望[J]. *自动化学报*, 2020, 46(8): 1539-1556.
- [8] WU D R, ZENG Z G, MO H, et al. Interval type-2 fuzzy sets and systems: overview and outlook[J]. *Acta Automatica Sinica*, 2020, 46(8): 1539-1556.
- [9] ABIYEV R, IDOKO J B, ALTIPARMAK H, et al. Fetal health state detection using interval type-2 fuzzy neural networks[J]. *Diagnostics*, 2023, 13(10): 1690.
- [10] ZHAO L. Application of interval type-2 fuzzy neural networks to predict short-term traffic flow[J]. *International Journal of Computer Applications in Technology*, 2012, 43(1): 67-75.
- [11] BAI K Y, ZHANG W Y, WEN S P, et al. A data-knowledge-driven interval type-2 fuzzy neural network with interpretability and self-adaptive structure[J]. *Information Sciences*, 2024, 660: Article No. 120133.
- [12] QIAO J F, SUN Z J, MENG X. A comprehensively improved interval type-2 fuzzy neural network for NO_x emissions prediction in MSWI process[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(11): 11286-11297.
- [13] LIU C F, YEH C Y, LEE S J. Application of type-2 neuro-fuzzy modeling in stock price prediction[J]. *Applied Soft Computing*, 2012, 12(4): 1348-1358.
- [14] 赵亮, 谢志峰, 张坤鹏, 等. 无线网络信号传输建模: 一种区间二型模糊集成深度学习[J]. *智能科学与技术学报*, 2020, 2(4): 401-411.
- [15] ZHAO L, XIE Z F, ZHANG K P, et al. Modeling signal propagation in wireless network: an interval type-2 fuzzy ensemble deep learning approach[J]. *Journal of Intelligent Science and Technology*, 2020, 2(4): 401-411.
- [16] SUN C X, LIU Z, WU X L, et al. Information orientation-based modular type-2 fuzzy neural network[J]. *Information Sciences*, 2024, 672: Article No. 120716.
- [17] LIU X, ZHAO T Y, CAO J T, et al. Design and prediction of self-organizing interval type-2 fuzzy wavelet neural network[J]. *Information Sciences*, 2024, 661: Article No. 120157.
- [18] ATANASSOV K T. Intuitionistic fuzzy sets[M]//Intuitionistic Fuzzy Sets. Heidelberg: Physica-Verlag HD, 1999: 1-137.
- [19] 赵涛, 肖建. 二型直觉模糊集[J]. *控制理论与应用*, 2012, 29(9): 1215-1222.
- [20] ZHAO T, XIAO J. Type-2 intuitionistic fuzzy sets[J]. *Control Theory & Applications*, 2012, 29(9): 1215-1222.
- [21] LI L, YANG J, WU W. Intuitionistic fuzzy hopfield neural network and its stability[J]. *Neural Network World*, 2011, 21(5): 461-472.
- [22] HÁJEK P, OLEJ V. Intuitionistic fuzzy neural network: the case of credit scoring using text information[C]//Proceedings of the International Conference on Engineering Applications of Neural Networks. Cham: Springer, 2015: 337-346.
- [23] PAO Y H, PARK G H, SOBAJIC D J. Learning and generalization characteristics of the random vector functional-link net[J]. *Neurocomputing*, 1994, 6(2): 163-180.
- [24] ATANASSOV K, GARGOV G. Interval valued intuitionistic fuzzy sets[J]. *Fuzzy Sets and Systems*, 1989, 31(3): 343-349.
- [25] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers[J]. *Neural Processing Letters*, 1999, 9(3): 293-300.
- [26] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. *Neurocomputing*, 2006, 70: 489-501.

[作者简介]



赵亮（1978- ），男，博士，河南工业大学电气工程学院副教授，主要研究方向为深度学习、平行学习。



崔贝贝（1990- ），女，博士，河南工业大学电气工程学院讲师，主要研究方向为深度学习、图像处理。



李梦威（1998- ），男，河南工业大学电气工程学院硕士生，主要研究方向为深度学习。



朱献超（1990- ），男，博士，河南工业大学人工智能与大数据学院讲师，主要研究方向为强化学习、量子强化学习和机器学习。



郑玉卿（1987- ），男，博士，河南工业大学电气工程学院讲师，主要研究方向为数据驱动控制、智能控制、自适应控制。