

DOI:10.19651/j.cnki.emt.2106016

基于 GPR-EN 时空混合模型的空气污染物浓度预测*

任静¹ 贾佳²

(1. 郑州科技学院 土木建筑工程学院 郑州 450064; 2. 郑州航空工业管理学院 管理工程学院 郑州 450064)

摘要: 针对现有的污染物浓度预测方法大多没有兼顾气象数据的时间和空间特征,导致预测精度大打折扣的问题,提出一种能够同时提取过程数据时间和空间特征的时空混合预测方法 GPR-EN。首先,利用弹性网算法(EN)全面分析样本点和目标点的空间关联性,并进行时空数据重构,为预测模型提供最优的变量输入。其次,借助于高斯回归模型(GPR)的强泛化能力,能够有效地处理时空数据的复杂非线性特性,更加全面地刻画了历史数据的时空依赖性。最后,在 AQI 数据集和空气 SO₂ 浓度数据集上进行仿真实验,实验结果表明所提方法对比方法的预测精度提高 22% 以上。

关键词: 空气污染物浓度预测;高斯过程回归;弹性网算法;时空特性分析

中图分类号: TP181 **文献标识码:** A **国家标准学科分类代码:** 120.30

Prediction of air pollutant concentration based on GPR-EN spatio-temporal hybrid model

Ren Jing¹ Jia Jia²

(1. College of Civil and Architectural Engineering, Zhengzhou University of Science and Technology, Zhengzhou 450064, China;

2. School of Management Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450064, China)

Abstract: Most of the current pollutant concentration prediction methods do not take into account the temporal and spatial characteristics of meteorological data, which results in a great loss of forecasting accuracy. To this end, this paper presents a spatio-temporal hybrid prediction method (GPR-EN), which can extract both temporal and spatial characteristics of process data and improve the prediction accuracy of pollutant concentration significantly. First, elastic net algorithm is used to comprehensively analyze the spatial correlation between the sample points and the target points, and reconstruct the spatiotemporal data, so as to provide the optimal variable input for the prediction model. Second, with the strong generalization ability of the Gaussian process regression model, the complex nonlinear characteristics of spatiotemporal data can be effectively handled, so as to obtain better prediction results. Finally, simulation experiments are carried out on AQI data set and air SO₂ concentration data set, and the experimental results show that the prediction accuracy of the proposed method is more than 22% higher than that of the contrast methods.

Keywords: prediction of air pollutant concentration; Gaussian process regression; elastic net algorithm; temporal and spatial characteristics analysis

0 引言

随着工业生产水平的提高、人口增加和能源消耗的增加,也带来了较为严重的环境污染物问题,给人们的身体健康和社会的经济发展带来了较大的隐患。特别是近 10 年来,我国多个城市都爆发出较为严重的空气质量问题,如 PM_{2.5}, SO₂ 等污染物浓度增高,雾霾天气增加等。因此,对空气污染物浓度的相关研究变得非常重要且必要,能够

为政府部门制定政策提供决策支持,并减少环境污染带来的危害,保护公众的健康。因此,空气污染物浓度的预测近年来也受到工业界和学术界的广泛关注,具有重要的理论和实践意义。

早期的污染物浓度预测方法主要有基于经验知识的方法^[1],基于物理模型的方法^[2]以及基于数据驱动的方法^[3]。其中,基于经验知识的方法主要依赖专家经验,容易受到主观因素的影响导致预测精度较低,而基于物理模型方法需

收稿日期:2021-03-17

* 基金项目:河南省软科学研究计划项目(192400410118)资助

要建立精确的数学模型,实际过程中不易实施,因此方法的可拓展性也较差。另一方面,随着数据采集和存储技术的发展,为获取大量包含丰富信息的气象数据提供了便利,进一步促进了基于数据驱动方法的空气污染物浓度预测研究^[4]。

程阳洋等^[5]利用量子基因链编码双向神经网络用于旋转机械剩余使用寿命预测,仿真结果表明所提方法提高了预测精度和方法的实用性。王茜等^[6]开发了集合数值预报系统并成功应用于预测上海的污染物浓度。司志娟等^[7]将灰色 GM(1, 1)模型与神经网络相结合,提出改进的灰色神经网络模型来预测天津市 2001~2008 年的空气质量指数 (air quality index, AQI) 值。程蓉等^[8]利用优化后的随机森林方法作为预测模型,以南京市玄武湖地区近 5 年的历史空气质量数据作为训练集,并成功预测该地区下一日的 AQI。虽然上述的方法都取得了较好的预测结果,但是这些方法而忽略了空间上的相关性。因此,有学者提出了时空模型并成功应用于污染物浓度的预测,如时空 Kriging 模型^[9], 时空自回归滑动平均值^[10], 时空神经网络^[11-13]。

由上述的文献调研可知,气象数据不但具有时间关联性还具有空间关联性,两者都会对污染物的预测精度产生较大影响,而目前的研究中很少有同时考虑时间和空间关联性对空气污染物预测模型性能的影响。此外,现有的模型一般将整个空间变量作为输入,从而导致模型复杂且变量出现冗余,不利于模型预测精度的提升。基于此,本文提出一种基于高斯过程回归 (Gaussian process regression, GPR) 和弹性网算法 (elastic net, EN) 的时空混合模型 (GPR-EN)。本文的主要创新点如下:

1) 利用 EN 算法对多地点序列数据进行空间关系的分析,并筛选出关联性最前的空间变量,为预测模型提供最优的变量输入。

2) GPR 模型的强泛化性能能够很好地提取气象数据中的非线性空间依赖性,从而可以提高模型的预测性能。

3) 所提的混合预测模型是两种方法的互补与增强,为空气污染物浓度预测提供了新的思路,并在真实的气象数据中验证了 GPR-EN 模型的有效性。

1 高斯过程回归

1.1 先验概率分布

GPR 用于直接定义潜在函数的先验概率分布,它可以通过其平均函数和协方差函数完全确定,记为 $f(x) = GP(m(x), C(x, x'))$ 。其中,均值和协方差可通过式(1)计算得到。

$$\begin{cases} m(x) = E[f(x)] \\ C(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \end{cases} \quad (1)$$

其中, x 和 x' 表示任意随机变量。为了简单起见,一

般设均值函数 $m(x') = 0$ 。

令 $x_i = [T_i, \dots, T_{i+l-1}]^T$ 表示 l 维输入样本集,对应的输出 $y_i = T_{i+l}$ 为包含噪声 $\varepsilon \sim N(0, \sigma_n^2)$ 的实际观测值。设 x^* 为待预测的样本, y^* 为对应的预测值。根据高斯分布的性质,可得 y 和 y^* 的联合分布如式(2)所示。

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} C(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n & C(\mathbf{X}, x^*) \\ C(x^*, \mathbf{X}) & C(x^*, x^*) \end{bmatrix}\right) \quad (2)$$

其中, $y = [y_1, y_2, \dots, y_n]$, $\mathbf{X} = [x_1, x_2, \dots, x_n]$, $C(\mathbf{X}, \mathbf{X})$ 为 $n \times n$ 阶的半正定矩阵,其具体形式如式(3)所示。

$$C(\mathbf{X}, \mathbf{X}) = \begin{bmatrix} C(x_1, x_1) & C(x_1, x_2) & \dots & C(x_1, x_n) \\ C(x_2, x_1) & C(x_2, x_2) & \dots & C(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ C(x_n, x_1) & C(x_n, x_2) & \dots & C(x_n, x_n) \end{bmatrix} \quad (3)$$

其中, $C(x^*, \mathbf{X})$ 与 $C(\mathbf{X}, x^*)$ 互为转置,表示测试数据与训练集之间的相关程度。 $C(x^*, x^*)$ 表示测试集自身之间的相关性。通过推导可以得到后验分布 $P(y^* | x^*, S) \in GP(\mu(x^*), \sigma^2(x^*))$, $\mu(x^*), \sigma^2(x^*)$ 由式(4)、(5)定义:

$$\mu(x^*) = C(x^*, \mathbf{X})[C(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n]^{-1} y \quad (4)$$

$$\sigma^2(x^*) = C(x^*, x^*) - C(x^*, \mathbf{X}) \times [C(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n]^{-1} C(\mathbf{X}, x^*) \quad (5)$$

GPR 在计算过程中,不仅可以得到预测值 y^* , 还可以得到与之对应的方差 $\sigma^2(x^*)$ 。而这是其他一般方法所不具备的。

1.2 超参数求取

不同的核函数都含有多个核参数,在 GPR 训练过程中一般使用极大似然法来估计超参数。理论上,选取合适的核函数以及对应的核函数超参数 Θ (Θ 表示核函数中多个参数构建的向量),则可以使得 GPR 能够逼近任意函数^[14],即为:

$$\tilde{C} = C(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n \quad (6)$$

并且可以得到似然函数如式(7)所示。

$$\mathcal{L}(\Theta) = -\frac{1}{2}(n \log(2\pi) + \log |\tilde{C}| + y^T \tilde{C}^{-1} y) \quad (7)$$

一般,对式(7)使用共轭梯度法来计算极值,而与极值对应的参数则为需要求取的最优超参数。其中对 Θ 中的每个参数 θ_i 的偏导如式(8)所示。

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = -\frac{1}{2} \left(\text{Tr}(\tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial \theta_i}) - y^T \tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial \theta_i} \tilde{C}^{-1} y - 1 \right) \quad (8)$$

2 时空数据重构

2.1 弹性网回归算法

将待预测观测点序列记为 $y = [y_1, y_2, \dots, y_n]^T$, 其

他样本观测点的序列记为 $\mathbf{X} = [x_1, x_2, \dots, x_p]$, 其中 $x_i = [x_{1i}, x_{2i}, \dots, x_{ni}]^T$ 。包括待预测观测点的所有样本观测点个数即为 $num(\mathbf{y}) + num(\mathbf{X}) = 1 + p = q$ 。弹性网算法的目标函数为:

$$L(\lambda_1, \lambda_2, \boldsymbol{\varphi}) = |\mathbf{y} - \mathbf{X}\boldsymbol{\varphi}|^2 + \lambda_1 \|\boldsymbol{\varphi}\|_1 + \lambda_2 \|\boldsymbol{\varphi}\|_2^2 \quad (9)$$

其中,

$$\begin{cases} \|\boldsymbol{\varphi}\|_1 = \sum_{i=1}^{n-1} |\varphi_i| \\ \|\boldsymbol{\varphi}\|_2^2 = \sum_{i=1}^{n-1} |\varphi_i|^2 \end{cases} \quad (10)$$

本文利用弹性网回归算法得到待预测观测点和其他样本观测点之间的回归系数 $\hat{\boldsymbol{\varphi}}$ 。

$$\hat{\boldsymbol{\varphi}} = \min\{L(\lambda_1, \lambda_2, \boldsymbol{\varphi})\} \quad (11)$$

若将上述过程中的 $\lambda_2/\lambda_1 + \lambda_2$ 记为 β , 则式(11)可以等价写成式(12):

$$\begin{cases} \hat{\boldsymbol{\varphi}} = \min\{|\mathbf{y} - \mathbf{X}\boldsymbol{\varphi}|^2\} \\ (1-\beta)\|\boldsymbol{\varphi}\|_2^2 + \beta\|\boldsymbol{\varphi}\|_1 \leq t, \forall t \end{cases} \quad (12)$$

在式(12)中, $(1-\beta)\|\boldsymbol{\varphi}\|_2^2 + \beta\|\boldsymbol{\varphi}\|_1$ 表示弹性网惩罚项, 该项为岭回归惩罚和套索 (least absolute shrinkage and selection operator, LASSO) 回归惩罚的凸组合。因为弹性网回归算法为两者的凸组合, 所以弹性网回归算法具有变量选择和良好的群组效应等优点。此外, 当 $\beta = 1$ 时, 弹性网回归算法退变为岭回归算法; 当 $\beta = 0$ 时, 弹性网回归算法退变为 LASSO 回归算法。

一般地, 在进行空间关系处理时, 由于空间域内多个地点是随机选择, 无法保证空间域内所有的点都具有空间相关性。因此, 在进行时空重构的过程中, 本文利用弹性网回归算法进行回归系数计算时, 可以达到变量选择的目的, 防止受到与待预测点之间空间关系较弱或者不存在空间关系的观测点的影响, 进而减少重构后融合的时空数据所包含的时空信息的冗余, 提高了模型预测性能。

2.2 弹性网回归求解

弹性网算法的一般性解法是将式(9)转化为一般 LASSO 回归求解问题。

首先由数据 (\mathbf{y}, \mathbf{X}) 和 (λ_1, λ_2) , 然后构造一个自定义矩阵 $(\mathbf{y}^*, \mathbf{X}^*)$ 。其中:

$$\mathbf{X}_{(n+p) \times p}^* = (1 + \lambda_2)^{-\frac{1}{2}} \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda_2} \mathbf{I} \end{pmatrix} \quad (13)$$

$$\mathbf{y}_{(n+p)}^* = \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix} \quad (14)$$

然后, 记 $\mu = \lambda_1 / \sqrt{1 + \lambda_2}$, $\boldsymbol{\varphi}^* = \sqrt{1 + \lambda_2} \boldsymbol{\varphi}$, 则式(9)可表示为:

$$L(\mu, \boldsymbol{\varphi}^*) = L(\mu, \boldsymbol{\varphi}^*) = |\mathbf{y}^* - \mathbf{X}^* \boldsymbol{\varphi}^*|^2 + \mu |\boldsymbol{\varphi}^*|_1 \quad (15)$$

求解 $\hat{\boldsymbol{\varphi}}^*$ 记为:

$$\hat{\boldsymbol{\varphi}}^* = \min_{\boldsymbol{\varphi}^*} (L(\mu, \boldsymbol{\varphi}^*)) \quad (16)$$

其中, $\hat{\boldsymbol{\varphi}} = 1 / \sqrt{1 + \lambda_2} \hat{\boldsymbol{\varphi}}^*$ 。

以上通过使用增广矩阵将弹性网回归问题转化为 LASSO 回归问题。值得注意的是增广矩阵的扩充至 $n + p$, \mathbf{X}^* 的秩最大为 p 。这意味着弹性网回归算法具有可以最多选择 p 个变量的能力, 而在 LASSO 回归中, 当 $p > n$ 时, LASSO 回归算法由于凸优化问题的性质最多选择 n 个变量, 这对于变量选择方法是不利的。此时, 通过式(13)、(14)的矩阵变换证明弹性网算法可以克服 LASSO 回归的上述不足^[15]。

2.3 时空重构原理

时空序列预测过程, 模型的输入不仅需要包含时间信息, 空间信息也很重要。在时空预测建模过程中, 确定影响预测空间区域动态变化以及如何测量空间影响是非常关键的一步。一般使用空间权重矩阵用作空间相关性的度量, 空间权重矩阵可以反映了施加在待预测站点上的所有站点的影响程度。空间关系分析原理如图 1 所示。

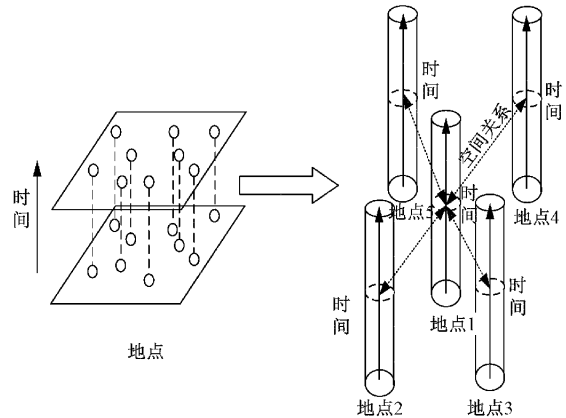


图 1 空间关系分析

在地统计学领域中, 一般认为空间距离越近, 则空间相关性越大。所以传统空间关系的计算一般是根据空间距离来判断。那么构建一般性空间权重矩阵 \mathbf{W} 的计算公式如式(17)所示。

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1q} \\ w_{21} & w_{22} & \dots & w_{2q} \\ \vdots & \vdots & \vdots & \vdots \\ w_{q1} & w_{q2} & \dots & w_{qq} \end{bmatrix} \quad (17)$$

其中, w_{ij} 表示第 i 个观测点和第 j 观测点的空间相关系数。 w_{ij} 的计算方法一般有两种, 第 1 种如式(18)所示当距离在阈值小于 δ 时可以认为两地之间具有空间相关关系, 观测地点之间互相影响。当超过这个阈值则表示没有空间相关关系。

$$w_{ij} = \begin{cases} 1, & d_{ij} < \delta \\ 0, & \text{其他} \end{cases} \quad (18)$$

第2种如式(19)所示,其中 ϵ 表示幂指数。从式(19)中可以看出随着两观测地点间距离的增加,第 i 个观测点和第 j 观测点之间的空间相关性呈现指数下降。通过相关研究发现,空间关系的强度随着距离增加,其减弱程度强于线性比例关系。两式中的 d_{ij} 均表示空间距离。

$$\omega_{ij} = \begin{cases} 1/d_{ij}^\epsilon, & d_{ij} < \delta \\ 0, & \text{其他} \end{cases} \quad (19)$$

通过空间权重矩阵构建的时空数据如式(20)所示。

$$\mathbf{Z}_{ST} = \mathbf{W}_{q \times q} \cdot [\mathbf{y} \mathbf{X}] \quad (20)$$

研究区域内的空间关系衡量不仅受距离因素影响,还会因选取观测变量空间位置的不同,受到地形、地势、气候等各种环境因素的影响。因此仅用空间距离来衡量空间关系具有很大的局限性。本文采用弹性网回归算法,通过样本观测点序列与待预测点序列数据之间内在关系作为彼此间空间关系的度量指标。该方法通过变量之间的关系作为空间关系衡量,综合了影响空间关系的各个因素,进而保证重构时空数据准确。同时,由于弹性网回归算法具有变量选择的能力,可以筛选研究区域内空间关系弱的观测点数据。其融合比例由 η 决定。因此,可得到如式(21)所示的时空重构数据。

$$\mathbf{Z}_{ST} = \eta \times \mathbf{X}\hat{\boldsymbol{\phi}} + (1 - \eta) \times \mathbf{y} \quad (21)$$

其中, η 表示时空融合系数,表示融合信息的时空比例。

3 仿真实验

3.1 AQI数据集实验结果

本文所使用的AQI时空数据集包含山东省2015年1月~2016年12月枣庄、青岛、菏泽、莱芜、泰安、潍坊、济宁、日照、济南、淄博、聊城、临沂、德州的AQI日值数据(该数据来源于中国气象网)。

为了定量地表示时空预测效果,本文选取均方根误差(root mean squared error, RMSE),平均对称绝对百分率误差(symmetric mean absolute percentage error, SMAPE),标准化均方根误差(normalized root mean square error, NRMSE)指标。

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n [\hat{y}_i - y_i]^2}{n-1}} \quad (22)$$

$$\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^n [\hat{y}_i - y_i]^2 / \sum_{i=1}^n [y_i - \bar{y}]^2}{n}} \quad (23)$$

$$\text{SMAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{1/2 | \hat{y}_i + y_i |} \quad (24)$$

此外,还选取80%的样本作为训练集,剩余20%样本作为测试集,并选取山东济南观测站点的数据作为测试站点。时空融合系数 η 设定为0.4。AQI数据实验结果如图2所示。从图2中可以定性地看出预测值曲线能够较好的拟合观测值曲线,并且误差曲线也保持在一定的误差范围之内,震荡的范围较小。说明所提的预测效果较好。

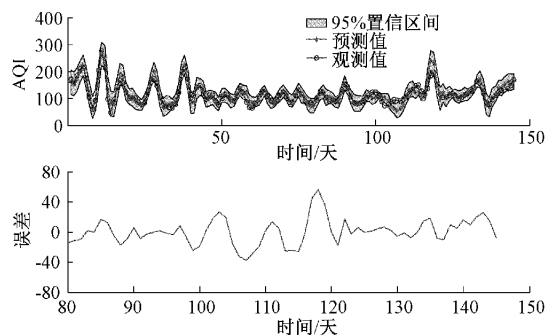


图2 AQI数据的GPR-EN预测结果

为了进一步定量刻画本文所提方法的有效性,实验选取自适应神经网络的模糊推理系统(adaptive network-based fuzzy inference system, ANFIS)、回声状态高斯过程(echo state gaussian process, ESGP)、回声状态网络(echo state network, ESN)、GPR方法进行对比实验。各个方法均是选取80%的样本作为训练集,剩余20%样本作为测试集,对山东济南市观测站点数据进行预测,得到对比结果如表1所示。

表1 各个方法预测结果比较

方法	RMSE	SMAPE	NRMSE
ANFIS	39.509 9	0.110 7	0.900 7
ESGP	30.600 9	0.101 0	0.959 9
ESN	36.455 9	0.106 4	0.901 4
GPR	15.389 5	0.045 7	0.374 7
GPR-EN	11.146 4	0.035 2	0.256 7

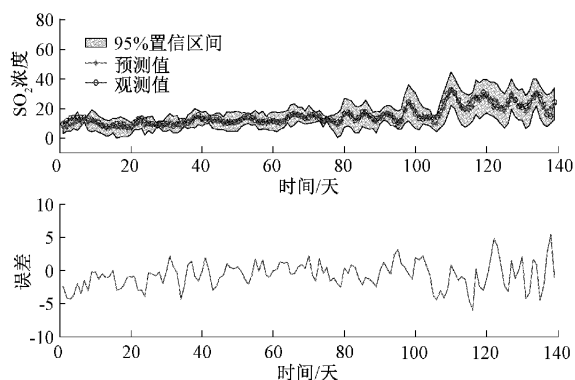
由表1的结果可知,ANFIS方法的3个指标的预测效果都比较差,类似的情况也出现在ESGP和ESN方法中。本文所提方法是以GPR方法为基础进行的改进,因此使用原始GPR方法作为对比算法,虽然GPR方法相比于前2种模型的预测精度有了一定的提升,但是还是难以取得令人满意的效果。本文提出的改进方法GPR-EN能够有效提取最相关特征变量,为模型提供最好的输入,而且可以同时利用气象数据的时空和时间信息,从而提高了模型的预测性能,取得最高的AQI预测精度。

3.2 空气SO₂浓度数据仿真

本文还选取2017年1月1日~2018年11月30日的辽宁省葫芦岛、朝阳、锦州、盘锦、营口、辽阳、铁岭、鞍山、沈阳、本溪、抚顺、丹东共12个城市的SO₂浓度数据(该数据来源于中国气象网)。

本文将数据集分为训练和测试两部分,两者比例为4:1。且高斯核函数中的时空融合系数 η 设置为0.6。针对本溪观测站点数据进行验证,得到预测曲线如图3所示。

为了进一步验证本文所提方法的有效性,实验选取ANFIS、ESGP、ESN、GPR方法进行对比实验。各个方法

图 3 STGPR 模型的 SO₂ 预测结果

均是选取 80% 的数据作为训练, 20% 的数据作为测试, 对于本溪观测站点数据进行预测, 得到对比结果如表 2 所示。

表 2 各个方法预测结果比较

方法	RMSE	SMAPE	NRMSE
ANFIS	7.169 8	0.128 3	0.900 7
ESGP	4.575 4	0.097 5	0.695 1
ESN	6.477 8	0.114 9	0.901 4
GPR	4.483 4	0.095 1	0.672 7
GPR-EN	2.988 2	0.075 4	0.451 4

根据表 2 所示, STGPR 在时间信息的基础上融合空间信息, 因此本文所提方法 STGPR 可以得到较好的预测效果。STGPR 相对 ANFIS、ESGP、ESN 以及 GPR 等时间序列预测方法, 能够结合空间相关性信息, 提高预测精度。

4 结 论

针对气象数据存在时空特征, 本文提出基于弹性网回归算法对时空数据进行时空重构, 并结合高斯过程回归算法的概率输出的优点, 进而得到较好的时空预测结果。在 AQI 和 SO₂ 数据集上进行仿真实验, 实验结果表明本文所提的 GPR-EN 模型具有良好的时空预测效果。

虽然所提方法能够同时考虑污染物数据的时间和空间特性并取得较好的预测效果。但是, 随着研究区域的增加和数据量的增大, 如何在保障预测精度的同时提高模型的计算速度, 也是未来研究的一个重点和难点。

参考文献

- [1] 姚红岩, 施润和. 基于周边站点优化选取的随机森林 PM_{2.5} 小时浓度预测研究[J]. 环境科学学报, 2021, 41(4): 1565-1573.
- [2] SICARD P, CRIPPA P, DE A, et al. High spatial resolution WRF-Chem model over Asia: Physics and chemistry evaluation[J]. Atmospheric Environment, 2021, 244: 118004.
- [3] 张婷曼, 丁凰. 依赖不确定性关联规则的城市交通流

大数据挖掘[J]. 国外电子测量技术, 2020, 39(11): 39-45.

- [4] 赵佰秋, 孟立新, 于笑楠, 等. 多通道超声的双向单程伪距测量研究[J]. 电子测量与仪器学报, 2020, 34(9): 174-180.
- [5] 程阳洋, 李锋, 汤宝平, 等. 量子基因链编码双向神经网络用于旋转机械剩余使用寿命预测[J]. 仪器仪表学报, 2020, 41(7): 164-174.
- [6] 王茜, 黄蕊珠, 肖宇. 长三角区域空气质量数值预报系统及其在重大活动保障中的应用[J]. 中国环境监测, 2020, 36(2): 15-21.
- [7] 司志娟, 孙宝盛, 李小芳. 基于改进型灰色神经网络组合模型的空气质量预测[J]. 环境工程学报, 2013, 7(9): 3543-3547.
- [8] 程蓉, 钱雪忠. 基于神经随机森林的局部空气质量预测模型[J]. 计算机工程与设, 2020, 41(7): 1958-1966.
- [9] YANG Y, WU J, CHRISTAKOS G. Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model[J]. Ecological Indicators, 2015, 56: 125-133.
- [10] AGOUA X G, GIRARD R, KARINIOTAKIS G. Probabilistic models for spatio-temporal photovoltaic power forecasting [J]. IEEE Transactions on Sustainable Energy, 2018, 10(2): 780-789.
- [11] TASCIKARAOGLU A, SANANDAJI B M, POOLLA K, et al. Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform [J]. Applied Energy, 2016, 165: 735-747.
- [12] 黄伟建, 李丹阳, 黄远. 面向空气质量的时空混合预测模型[J]. 计算机应用, 2020, 40(11): 3385-3391.
- [13] 沈廷鳌, 管金发, 华卫星, 等. 基于 FFT 的极端频率信号相位差测量方法[J]. 电子测量技术, 2020, 43(24): 6-10.
- [14] ZHANG C, WEI H, ZHAO X, et al. A Gaussian process regression based hybrid approach for short-term wind speed prediction[J]. Energy Conversion & Management, 2016, 126: 1084-1092.
- [15] 李鑫, 孙晋, 肖甫. 基于弹性网稀疏表示的芯片参数成品率估算方法[J]. 电子学报, 2017, 45(12): 2917-292.

作者简介

任静, 硕士, 讲师, 主要研究方向为信号处理、污染物浓度预测。

E-mail: renjing850315@163.com

贾佳, 博士, 讲师, 主要研究方向为气象数据挖掘、时间序列预测。

E-mail: jiajia90001@163.com