

神经网络在空气污染预报中的研究进展

权 钺, 漆 然, 邱 诚

成都工业学院材料与环境工程学院, 四川 成都

收稿日期: 2023年11月1日; 录用日期: 2023年12月1日; 发布日期: 2023年12月8日

摘 要

空气污染预报是一项复杂的系统工程, 是当今环境科学研究的热点, 国内外已有将神经网络法应用于大气污染预报的研究。人工神经网络(Artificial Neural Network, 即ANN), 是20世纪80年代以来人工智能领域兴起的研究热点。它从信息处理角度对人脑神经网络进行抽象, 建立某种简单模型, 按不同的连接方式组成不同的网络。在工程与学术界也常直接简称为神经网络或类神经网络。本论文旨在综述神经网络在空气污染预报中的研究进展, 探讨其在空气污染预报领域的未来展望。

关键词

神经网络, 空气污染预报, 研究进展, 未来展望

Research Progress of Neural Network in Air Pollution Prediction

Cheng Quan, Ran Qi, Cheng Qiu

Department of Material and Environmental Engineering, Chengdu Technological University, Chengdu Sichuan

Received: Nov. 1st, 2023; accepted: Dec. 1st, 2023; published: Dec. 8th, 2023

Abstract

Air Pollution Forecast is a complicated systems engineering, and it is a hotspot of environmental science research. The neural network method has been applied to air pollution forecast at home and abroad. Artificial Neural Network (ANN) is a research hotspot in the field of Artificial intelligence since 1980s. It abstracts the neural network of human brain from the angle of information processing, builds some simple model, and forms different networks according to different connection ways. In engineering and academic circles, it is often referred to as neural network or neural-like network. The purpose of this paper is to summarize the research progress of neural network in air pollution forecasting and discuss its future prospect in the field of air pollution forecasting.

Keywords

Neural Network, Air Pollution Prediction, Research Progress, Future Prospects

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

空气作为人类赖以生存的必需品，在生活、农业、工业等方面上有着至关重要的作用。空气污染，又称为大气污染，按照国际标准化组织(ISO)的定义，空气污染通常是指：由于人类活动或自然过程引起某些物质进入大气中，呈现出足够的浓度，达到足够的时间，并因此危害了人类的舒适、健康和福利或环境的现象。空气污染预报是根据气象条件(风、稳定性、降水及天气形势等)和污染源排放情况对某个区域未来的污染浓度及空间分布发出预报。对人类研究和评判空气质量起到关键作用。神经网络的学习过程就是调整权重参数的过程，确立好权重参数之后，神经网络就可以完成对输入样本的预测和分类任务。神经网络相比较与传统预测方法有着更便捷的预测方式和更包容的优点且能加强对环境污染的控制，以防止重大污染事件发生。特别是对于无法开展空气污染数值预报的城市，利用神经网络对空气污染的预测极其具有必要性。如果将空气污染的预报当成是一个映射过程，那么神经网络可以被看作是一个将空气污染因素(如气温、湿度、风速等)映射到污染浓度值的映射器。

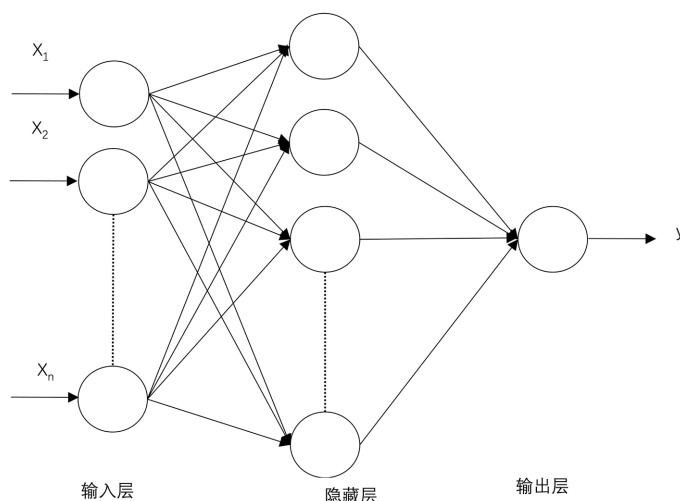


Figure 1. Typical neural network structure

图 1. 典型神经网络结构

2. 利用神经网络预报空气污染的意义

神经网络一般也称为人工神经，是科学家们从生物学和信息处理的角度出发，通过借助数学和物理的方法对人脑神经网络进行抽象后建立的简化模型。空气污染预报普遍采用空气污染指数的形式向公众发布空气质量的信息，其中常见的对空气污染指数有影响的作用气象因子有：降水，极大风速，极大风速的风向，平均风速，平均气温，平均水气压等等。空气污染与气象条件密切相关，所以在建立环境空

气污染预报模型时, 如果考虑各种气象因子与环境空气质量之间的关系, 应该有利于提高所建立模型的预报精度。

相比较于传统的空气污染预报方式, 因空气污染浓度的分布具有较强的非线性特性, 传统的计算方法会需要大量的人工观察, 耗费时间和人力, 并且结果误差偏大。而人工神经网络具有自学习, 自组织, 容错率大的特点, 利用神经网络进行抽象后建立的简化模型, 能更加准确的预报出空气污染的指数, 特别是对骤升骤降趋势也能得到准确度较高的预报结果, 并且神经网络应用于空气污染预报中, 计算简单, 所需数据要求相对较低, 尤其适用于无法开展数值预报的城市和地区。

3. 神经网络在空气污染预报中的研究进展

人工神经网络是 20 世纪 80 年代迅速兴起的非线性科学, 特别适用于对具有多因素性、不确定性、随机性、非线性等特点的对象进行研究, 而空气污染预报正是这样的一类问题。随着经济的高速发展、人口的快速增加以及人类对自然资源的过度开发和利用, 空气污染问题日益严重, 已经成为全球人类共同面临和关注的问题。开展空气污染预报可以更好地了解空气污染变化趋势, 为环境管理决策提供及时、准确、全面的空气质量信息, 以加大环境污染控制力度预防严重污染事件的发生。人工神经网络(Artificial Neural Network, ANN)简称神经网络, 是以人脑结构为参考模型, 由大量简单神经元广泛连接而成的复杂网络, 可用来模拟人类大脑神经的思维活动。

神经网络的研究始于 20 世纪 40 年代信息科学的开创, 经过 60 多年发展, 神经网络已经成为一种多学科、综合性的交叉技术研究领域, 吸引着国内外许多不同领域专家和学者对其进行研究和探索。近年来大气污染问题越来越受到普通民众的关注, 大气污染的预报成为一项很重要的任务。通过利用大数据分析方法来研究环境问题的学者越来越多。郑跃君等[1]提出在环境污染防治管理中应用大数据方法, 侧重在污染防治管理方式的创新方面提出了一些中肯意见。潘本锋[2]等对构建环境或大气监测平台较为感兴趣, 试图通过数据分析构建一个可行的监测平台, 这对高层的管理决策起来很重要的作用。其中最为出名的 BP 神经网络 1986 年由 Rumelhart 和 McClelland 提出。是一种基于误差逆向传播的多层前馈网络, 主要包含了三层结构: 输入层、输出层和隐含层(见图 1) [3] [4]。其主要思想是在输入层输入学习样本, 然后通过正向传播得到预期的输出, 如果误差过大再使用逆向传播, 调整 BP 网络的权值和阈值, 经过若干次的反复调整和学习训练, 不断逼近期望值, 当误差满足要求时停止学习训练, 记录网络权值和阈值。近些年大气环保产业一直作为国家大力倡导的产业, 环保产业的良性发展国家一直致力于改善大气环保产业发展的大环境, 在这样的大背景下有不少企业的营利能力依旧不强。国内学者高明、黄清煌[5]从提升大气环保产业竞争力的角度, 何欢浪、陈璐[6]从关联市场、环境政策强度的角度对环保产业的发展进行研究。但目前关于环保产业的研究大多基于宏观角度, 从要素投入视角的微观层面的具体量化研究尚不多见, 从而无法准确确定影响环保产业发展关键因素。RBF 是一种有效的神经网络预测模型。例如, 张兆同、余潜[7]通过 RBF 神经网络与灰色关联相结合构建神经网络模型对我国棉花价格进行预测; 海涛、王路[8]运用 RBF 神经网络建立预测模型对太阳光照度进行预测, 但目前尚未应用于环保产业发展预测中(见表 1)。RBF 神经网络为径向基函数(Radial Basis Function, RBF)是一种单隐层的前馈神经网络, 其原理是输入层将低维度的输入矢量映射至隐层(此过程为非线性的), 然后隐层与输出层通过权值连接(此过程为线性的), 通过网络不断训练确定权值, 最终构建 RBF 神经网络。 $PM_{2.5}$ 是指在空气动力学领域中直径不大于 2.5 微米的可吸入颗粒物, 是空气质量评价的主要指标之一。全面掌握 $PM_{2.5}$ 浓度的空间分布规律, 表征大气污染的空间过程和环境行为, 对于支撑大气污染监测预警与综合治理、保护人类健康与社会可持续发展, 具有重大的现实意义和指导价值。截至 2017 年底, 中国环境监测总站已建成超过 1400 个地面空气质量监测站点, 并对外发布包括 $PM_{2.5}$ 在内的每小时空气质量监测数据, 提供了高精度、高可靠的

实时监测结果。然而，由于地面监测站点空间分布不均、覆盖程度不高，现有研究难以对其监测数据进行有效地时空分析与深度挖掘[9]。大量研究表明，AOD 与 PM_{2.5} 浓度具有较强的相关性。研究 PM_{2.5} 浓度与基于遥感反演的 AOD 等相关因子之间的空间回归关系，能为获得整个研究区域的 PM_{2.5} 浓度分布提供有效解决方案。诸多学者建立了统计回归模型来表征 PM_{2.5} 和 AOD 等因子的内在联系，并通过将地面监测和遥感观测数据进行结合获得了与传统模拟方法相当或更高精度的 PM_{2.5} 浓度空间分布。例如，Liu 等[10]利用 AOD、季节因子、相对湿度等因子构建了广义线性模型对美国东部地区 PM_{2.5} 浓度进行空间估算，较大提升了 PM_{2.5} 浓度的反演精度。李啸天等[11]基于 AOD 数据、气象数据和植被覆盖数据，采用普通线性回归(Ordinary Linear Regression, OLR)模型进行了江苏省 PM_{2.5} 浓度的反演，其结果能较好地揭示 PM_{2.5} 浓度的时空分布规律。关于神经网络在大气污染预报中的研究例子还有很多，例如：梁春丽等人[12]基于 ARMA 模型，主成分分析模型和神经网络模型对黑龙江省空气质量数据进行了分析和预测。闫妍等人[13]运用 BP 神经网络对大气污染物浓度的实测值进行训练学习，利用建立的预测模型对污染物浓度进行预测和预测。马天成等[14]采用一种改进型 PSO 优化的模糊神经网络，将粒子群算法与模糊神经网络进行融合，发挥 PSO 算法全局寻优的特点，对某市 2013 年 PM_{2.5} 颗粒物浓度进行预测和验证，验证结果表明该算法具备良好的预测精度(见表 1)。诸多应用实例表明：神经网络应用于大气环境质量预测预警是比较理想的。在大气污染预测领域中运用神经网络技术，主要是利用神经网络的泛化能力，它是由所学习的知识和经验总结归纳后进行判断的能力。与传统的预测方法相比，人工神经网络在大气污染预测方面显示出了明显的优越性：(1) 由于人工神经网络具有较好的逼近效果，训练后的神经网络，在样本上输出期望值(在误差允许范围内)，在非样本点上，能够表现出网络的联想记忆功能；(2) 容错性较强，能够处理信息不全的预测问题；(3) 动态自适应能力强，能够适应外界新的学习样本，可以使网络知识不断更新；(4) 预测速度快；(5) 可以将新获取的数据信息应用到人工神经网络中，并及时建立新的预测模型，使预测的精度大幅度提高。总之，利用神经网络预测大气污染物浓度将是一种描述和刻画大气污染非线性现象的有利工具，并将取得比传统的统计方法更好的大气污染预测精度，以及与数值预测方法相比，更简单且易普及。神经网络应用于大气污染预测的研究方案与应用必将为大气污染预测开辟了一条新思路。

Table 1. The results of neural networks in research progress

表 1. 神经网络在研究进展中的成果

研究者	项目	项目特点
权铖、漆然	项目研究进展总结	随着人工神经网络的理论发现与研究技术不断发展，其在空气污染预报领域中的使用越来越广泛，目前主要应用于空气污染浓度的短期预报，特别是目前对于空气污染数值难以作出准确数值的城市。
张兆同、余潜[7]等	RBF 神经网络	RBF 是一种有效的神经网络预测模型，具有结构简单、收敛速度快、能够逼近任意非线性函数的特点。研究者通过 RBF 神经网络与灰色关联相结合构建神经网络模型对我国棉花价格进行预测。海涛、王路运用 RBF 神经网络建立预测模型对太阳光照度进行预测，但目前尚未应用于环保产业发展预测中。
梁春丽[12]等	ARMA 模型神经网络	具有较好的预测精度，也是目前应用较为广泛的计量模型。研究者基于 ARMA 模型，主成分分析模型和神经网络模型对黑龙江省空气质量数据进行了分析和预测。
马天成[14]等	改进型 PSO 优化的神经网络	具有收敛速度快、参数少、算法简单、算法全局寻优的特点，研究者对某市 2013 年 PM _{2.5} 颗粒物浓度进行预测和验证，验证结果表明该算法具备良好的预测精度。

4. 神经网络预测与传统分析方法对比

神经网络预测和传统分析测试方法在环境空气质量测定方面各有优势和劣势(见表 2)。神经网络预测可能在测定时间短、经济成本较低、预测准确性高方面具有优势；而传统分析测试方法可能在准确性和可信度方面具有优势，但在测定时间较长、经济成本较高和操作难度较高方面存在一些限制。最终选择何种方法应根据具体需要、资源和实际情况进行综合考量。

Table 2. Neural networks vs. Traditional analysis methods

表 2. 神经网络与传统分析方法的对比

方法	测定时间	经济成本	准确性	操作难度
传统分析方法	时间周期长,至少需要一年的观测、需要进行样本采集和化学分析,步骤冗杂繁多	需要人力资源和仪器设备、人工费用与仪器使用成本较高	需要基于化学分析,准确性受到仪器和操作方式影响	需要掌握实验操作技巧、专业知识等技能、较为复杂
神经网络预测	时间短、需建立数据模型,例如 BP、MLP 模型	人力资源和数据采集成本较高、但预测成本低	对数据数量和精度要求不高,结果准确性和客观性较强[15]	操作较为简单,只需数据处理,整理结果等步骤

5. 神经网络在空气污染预报中的未来展望

随着神经网络理论和技术的不断发展,它在空气污染预报领域的应用也越来越广泛,以下几方面的研究将会得到更多关注。

(1) 研究网络输入因子选取的方法。建立空气污染预报的神经网络模型,关键是找出隐含于各气象要素和污染物浓度之间的规律,设计最佳网络结构。充分考虑影响空气污染物浓度的各种因素,采用合适的方法选取网络的输入因子,是确定网络结构必不可少的一个重要环节,这就要求研究人员在选取输入因子的方法上多下功夫,以确保网络结构更加合理。

(2) 探索适合于空气污染预报使用的神经网络模型和算法,进一步提高网络的泛化能力。目前神经网络在空气污染预报领域中的应用研究,尤其在国内外,大多还是集中在传统 BP 算法的尝试和应用方面,对于模型和算法的优化所做的工作较少。神经网络模型对于历史样本数据拟合的精度高并不能说明网络预报效果就一定很好,还需要考虑网络对于将来未知数据的预报效果。因此,获得更高的预报精度、更快的收敛速度、更小的误差、更高的泛化能力、实现长期预报,将是今后神经网络在空气污染预报领域应用的重要研究内容。

(3) 提高监测系统和监测网络的性能。神经网络预测的效果还取决于训练数据的代表性和数据量的大小,一般而言,训练数据量较大的网络,预测的精度也较高。另外,由于神经网络使用大量的参数,为避免出现过拟合现象,也需要大量观测数据。因此,对空气污染物浓度、气象条件、污染源源强等因子进行长期监测,以获得高质量和充足的观测数据,从而提高预测准确度,是进行空气污染预报一个非常重要的前期准备工作。

6. 结语

本文综述了神经网络在空气污染预报中的研究进展,通过对相关文献的梳理和分析,以及对不同神经网络的对比,我们发现神经网络在空气污染预报中的必要性,以及神经网络对空气预报工作所具有的时代意义。神经网络的特点包括用时短、经济成本低、准确性高和操作难度低。相比传统预测方法,神经网络只需建立模型,将空气中的测定量用作参数,再确认节点即可完成测试,提高了预报效率。然而,目前神经系统的研究仍存在节点选择不够精准、容错性较差或过度拟合等问题,且用于空气污染预

报的神经网络模型仍以 MLP 模型居多, 传统 BP 算法存在许多细微问题, 导致实际预报效果不理想。因此, 需要将多种预报方法联合应用, 发挥各自特点, 以准确预测空气污染物的未来变化情况。神经网络在空气污染预报中具有广阔前景。

项 目

2023 年四川省大学生创新创业训练计划项目, 编号 S202311116073。

参考文献

- [1] 郑跃君, 严翔. 环保大数据在环境污染防治管理中应用探究[J]. 资源节约与环保, 2017(10): 60, 62.
- [2] 潘本锋, 许人骥, 宫正宇, 等. 支撑京津冀区域大气污染联防联控的大气监测体系构建[J]. 中国环境监测, 2017, 33(5): 57-63.
- [3] 张同君. BP 神经网络在水利发电控制系统中的应用[J]. 河北科技师范学院学报, 2016, 30(1): 62-67.
- [4] 冯惠妍, 陈争光, 王淑云. BP 神经网络的土壤肥力评价研究[J]. 黑龙江八一农垦大学学报, 2014(1): 92-95.
- [5] 高明, 黄清煌. 基于产业链视角下我国大气污染治理产业分析[J]. 理论学刊, 2014(4): 49-54.
- [6] 何欢浪, 陈璐. 纵向关联市场、环境政策强度和中国环保产业的发展[J]. 商业研究, 2019(1): 71-77.
- [7] 张兆同, 余潜. 灰色关联分析与 RBF 神经网络在我国棉花价格预测中的应用研究[J]. 价格月刊, 2017(9): 31-36.
- [8] 海涛, 王路, 陈春华, 等. 基于 RBF 神经网络的太阳光辐照度预测[J]. 广西大学学报(自然科学版), 2016, 41(5): 1508-1513.
- [9] 蒲强, 邹滨, 翟亮, 等. 集成多源遥感数据的 PM_{2.5} 浓度空间分布制图[J]. 地球信息科学学报, 2016(12): 1717-1724.
- [10] Liu, Y., Sarnat, J.A., Kilaru, V., et al. (2005) Estimating Ground-Level PM_{2.5} in the Eastern United States Using Satellite Remote Sensing. *Environmental Science & Technology*, 9, 3269-3278. <https://doi.org/10.1021/es049352m>
- [11] 李啸天, 吴绍华, 徐于月, 等. 江苏省 PM_{2.5} 质量浓度的时空变化格局模拟[J]. 环境监测管理与技术, 2017, 29(2): 16-20.
- [12] 梁春丽, 周影. 黑龙江省空气质量数据的分析和预测[J]. 统计与咨询, 2014(6): 16-18.
- [13] 闫妍, 张云鹏, 李铠月, 等. 基于 BP 神经网络的西安环境空气质量的预测[J]. 电子设计工程, 2013, 21(21): 54-57.
- [14] 马天成, 刘大铭, 李雪洁, 等. 基于改进型 PSO 的模糊神经网络 PM_{2.5} 浓度预测[J]. 计算机工程与设计, 2014, 35(9): 3258-3262.
- [15] 赵慧宏. 兰州市 SO₂ 排放总量预测的 BP 人工神经网络模型[J]. 甘肃环境研究与监测, 2001, 14(3): 157-159.