

## ARTIKEL TINJAUAN PUSTAKA

# POTENSI ALGORITMA BERBASIS NEURAL NETWORK DAN TURUNANNYA SEBAGAI PREDIKTOR KADAR PM<sub>2,5</sub> DAN PM<sub>10</sub>: SUATU TELAAH SISTEMATIS

## POTENTIAL OF NEURAL NETWORK-BASED ALGORITHM AND ITS DERIVATIVES AS PREDICTORS OF PM<sub>2.5</sub> AND PM<sub>10</sub> LEVELS: A SYSTEMATIC REVIEW

Robby Soetedjo, Leonardo Alexander\*, Nikolaus Tobian

Fakultas Kedokteran dan Ilmu Kesehatan Universitas Katolik Indonesia Atma Jaya, Jl. Pluit Raya No. 2, Jakarta 14440

\* Korespondensi: leoalexander137@gmail.com

### ABSTRACT

**Introduction:** Air pollution is a leading cause of Non-Communicable Diseases (NCDs) in Southeast Asia and was identified as the second largest global health issue in 2016. Effective predictive strategies are essential to mitigate its adverse impacts. This study systematically reviews the application of Artificial Neural Network (ANN)-based algorithms in forecasting air pollutant concentrations, particularly PM<sub>2.5</sub> and PM<sub>10</sub>.

**Methods:** This systematic review followed PRISMA guidelines. Systematic searches from 4 databases were conducted to search for studies that apply ANN models for air pollution prediction with meteorological and/or geographical data as inputs. Data extraction focused on model structure, predictive accuracy, and comparison with other artificial intelligence algorithms. ANN models were evaluated for their ability to handle complex, non-linear variable interactions, flexibility across datasets, and predictive performance relative to other methods.

**Results:** ANN-based algorithms consistently outperformed alternative models in predicting PM<sub>2.5</sub> and PM<sub>10</sub> levels. Their adaptive learning capacity and ability to integrate diverse inputs enhanced predictive accuracy. Several studies reported further improvement when ANN was combined with derivative methods.

**Conclusion:** ANN provides a reliable and accurate tool for forecasting air pollution, supporting evidence-based policies for prevention and environmental management. Its role in environmental science highlights innovation in predictive modeling and opportunities for integration with sustainable technologies.

**Key Words:** artificial intelligence, air pollution, prediction

### ABSTRAK

**Pendahuluan:** Polusi udara merupakan penyebab utama Penyakit Tidak Menular (PTM) di Asia Tenggara dan telah diidentifikasi sebagai masalah kesehatan global terbesar kedua pada tahun 2016. Strategi prediksi yang efektif sangat penting untuk mengurangi dampak buruknya. Studi ini secara sistematis meninjau penerapan algoritma berbasis *Artificial Neural Network* (ANN) dalam memprediksi konsentrasi polutan udara, khususnya PM<sub>2.5</sub> dan PM<sub>10</sub>.

**Metode:** Tinjauan sistematis ini mengikuti pedoman PRISMA. Pencarian sistematis dilakukan pada 4 basis data untuk menemukan studi yang menerapkan model ANN dalam prediksi polusi udara dengan menggunakan data meteorologi dan/atau geografis sebagai input. Ekstraksi data difokuskan pada struktur model, akurasi prediksi, serta perbandingan dengan algoritma kecerdasan buatan lainnya. Model ANN dievaluasi berdasarkan kemampuannya menangani interaksi variabel non-linear yang kompleks, fleksibilitas pada berbagai dataset, dan kinerja prediktif dibandingkan metode lain.

**Hasil:** Algoritma berbasis ANN secara konsisten menunjukkan performa lebih baik dibandingkan model alternatif dalam memprediksi kadar PM<sub>2.5</sub> dan PM<sub>10</sub>. Kemampuan ANN dalam pembelajaran adaptif serta integrasi berbagai input meningkatkan akurasi prediksi. Beberapa studi melaporkan adanya peningkatan lebih lanjut ketika ANN dikombinasikan dengan metode turunannya.

**Simpulan:** ANN merupakan alat yang andal dan akurat untuk memprediksi polusi udara serta mendukung kebijakan berbasis bukti dalam pencegahan dan pengelolaan lingkungan. Peran ANN dalam ilmu lingkungan menyoroti inovasi dalam pemodelan prediktif dan membuka peluang integrasi dengan teknologi berkelanjutan.

**Kata Kunci:** kecerdasan buatan, polusi udara, prediksi

## PENDAHULUAN

Polusi udara menjadi penyebab penyakit tidak menular (PTM) secara global ke-2 terbesar setelah merokok pada tahun 2016. Polusi udara juga merupakan penyebab terbesar PTM di Asia Tenggara. Penyakit paru yang tidak menular dan penyakit paru obstruktif kronik (PPOK), kanker paru, penyakit jantung iskemik, dan stroke adalah lima penyakit terbanyak yang termasuk ke dalam PTM. Riset tersebut juga menunjukkan bahwa 91% populasi manusia di dunia telah terpapar polusi udara yang dikarenakan oleh peningkatan jumlah industri, agrikultur, dan berbagai sektor pekerjaan manusia yang membutuhkan pembakaran bahan bakar untuk menghasilkan energi. Polusi udara telah menyebabkan kematian manusia dengan penyebab PTM sebesar 5,6 juta penduduk di dunia. Angka tersebut menunjukkan bahwa terjadi semakin banyak pencemaran akibat polusi udara. Peningkatan polusi udara akan menyebabkan tingginya prevalensi PTM dan peningkatan beban ekonomi pemerintah untuk menanggulangi penyakit tersebut.<sup>1</sup>

Peningkatan mortalitas dan morbiditas PTM karena polusi udara membuat polusi udara menjadi pusat perhatian dari beberapa negara, salah satunya adalah Cina. Beberapa studi yang dilakukan di Cina menemukan jika penyakit kardiovaskular, penyakit serebrovaskular, diabetes mellitus, serta penyakit ginjal dapat menjadi komplikasi dari PTM akibat polusi udara.<sup>2</sup>

Sebanyak 72% dari seluruh kasus kematian manusia di dunia disebabkan oleh PTM dan kebanyakan kematian akibat PTM

terjadi pada negara-negara dengan pendapatan sedang hingga rendah. Penurunan kualitas ekonomi dan pembangunan negara berkorelasi dengan penurunan kualitas sumber daya manusia akibat PTM. Oleh karena itu, terjadi penghambatan kemajuan negara berkembang yang memiliki penduduk dengan prevalensi PTM yang tinggi.<sup>3</sup>

Berbagai pertemuan dan kebijakan telah dibuat oleh berbagai organisasi internasional seperti *United Nation General Assembly* dan *World Health Organization* (WHO) untuk mengontrol dan mencegah PTM. Kebijakan-kebijakan tersebut mengatur pola hidup manusia agar dapat hidup dengan sehat dan mencakup perubahan pola diet, pola kebiasaan manusia, serta aturan penggunaan alkohol. Namun, perencanaan dan kebijakan tersebut tidak mengatur polusi udara yang juga merupakan faktor risiko PTM sehingga timbul banyak penyakit PTM yang dikarenakan polusi udara.<sup>3</sup>

Berbagai teknik kalkulasi diciptakan manusia untuk memprediksi polusi udara, seperti *numerical forecasting models*, dan yang terkini adalah *machine learning models*. Penggunaan *machine learning models* juga mulai dikombinasikan dengan penghitungan *air quality index* (AQI) agar didapatkan konsentrasi polutan yang lebih optimal untuk menjaga kualitas udara. Kombinasi ini juga memiliki sistem yang dapat memberikan peringatan awal kepada manusia ketika terjadi penurunan kualitas udara karena polusi udara. Seiring dengan perkembangan teknologi, manusia menciptakan *artificial intelligence* (AI) dengan kapabilitas untuk mengukur dan

memprediksi konsentrasi polutan dalam udara. Dengan adanya AI yang dapat memprediksi, banyak peneliti mengharapkan jika alat tersebut dapat memprediksi lebih tepat dan dapat mengisyaratkan pemerintah untuk mengontrol polusi udara dalam kawasan tersebut.<sup>4</sup>

Kini model *machine learning* untuk memprediksi konsentrasi polutan telah banyak mengadopsi *artificial intelligence* (AI), dengan salah satu pendekatan unggulannya adalah *artificial neural network* (ANN). *Artificial Neural Network* merupakan algoritma berbasis jaringan saraf tiruan yang terdiri dari tiga jenis lapisan utama yaitu *input*, *hidden*, dan *output*, yang bekerja secara sinergis menyerupai cara kerja sistem saraf manusia. *Artificial Neural Network* mencakup berbagai jenis arsitektur, termasuk *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN), yang dirancang untuk menangani berbagai jenis data dan masalah kompleks. *Convolutional Neural Network* unggul dalam analisis data berbentuk *grid* seperti citra satelit, sedangkan RNN sangat efektif untuk menangani data sekuensial seperti tren waktu pada polusi udara. Kemampuan ANN untuk meniru pola pikir manusia ini memungkinkan algoritma ini mempelajari hubungan yang kompleks antara variabel *input* seperti data meteorologi dan geografis, sehingga dapat memprediksi konsentrasi polusi udara secara otomatis dengan tingkat akurasi yang lebih tinggi dibandingkan model statistik maupun model numerik tradisional.<sup>5</sup>

Tujuan studi ini adalah untuk menunjukkan efektivitas prediksi algoritma berbasis

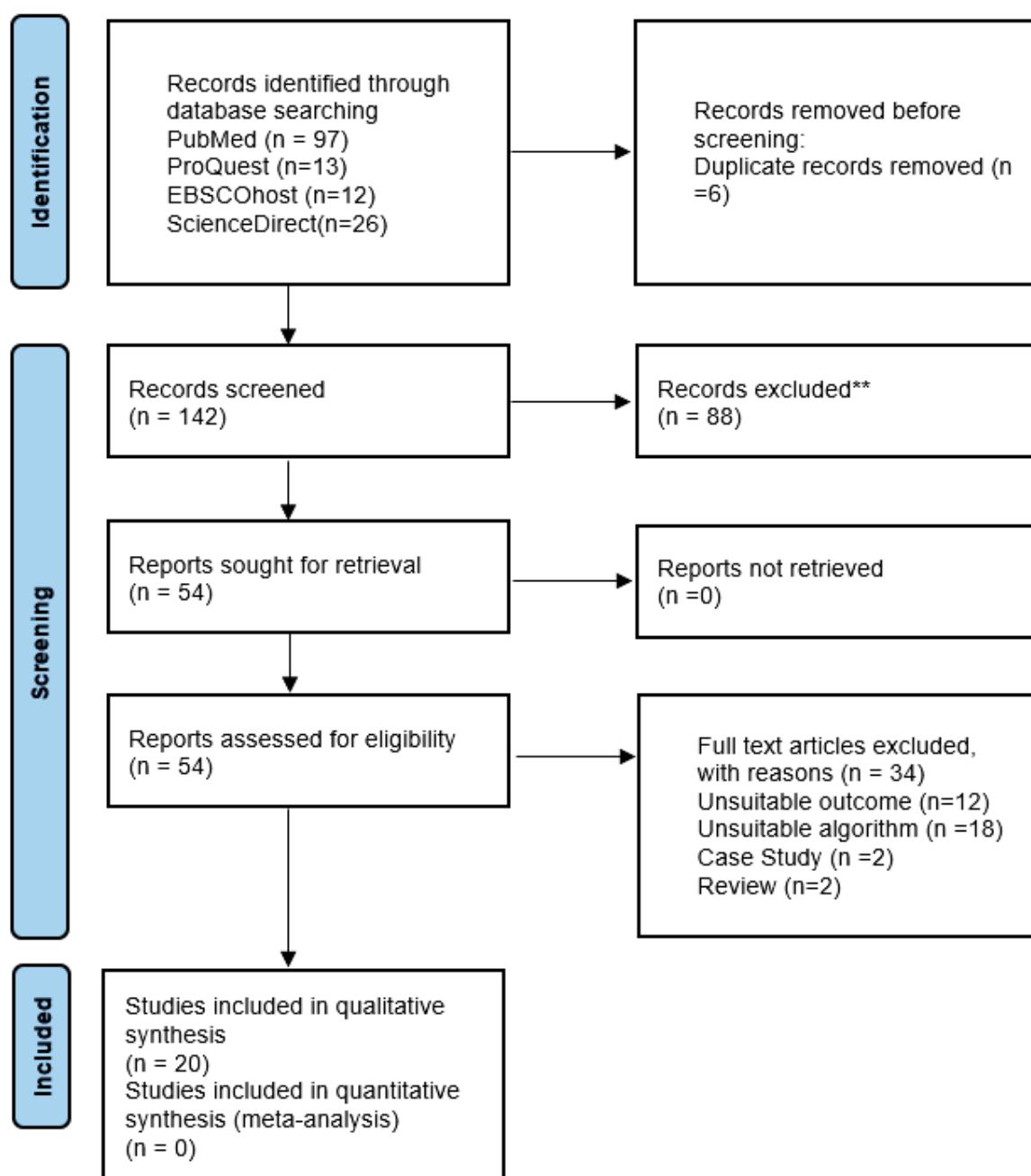
*Neural Network* terhadap konsentrasi polusi udara. Sama seperti *machine learning models*, ANN merupakan model komputasi yang terinspirasi dari otak manusia, yang belajar dari data dan dapat digunakan untuk membuat prediksi. Algoritma ini juga dapat dikombinasi dengan beberapa teknik lainnya yang berguna untuk meningkatkan performanya dalam memprediksi. peneliti memasukkan literatur yang membahas ANN secara tunggal maupun dikombinasikan dengan teknik lainnya dalam studi ini.

## METODE

Penulisan studi literatur ini mengacu kepada protokol PRISMA.<sup>6</sup> Pencarian literatur menggunakan sumber referensi dari *PubMed*, *EBSCOhost*, *ProQuest*, dan *ScienceDirect* tanpa ada batasan bahasa dan waktu. Kata kunci “*Air Pollution, Artificial Intelligence, dan Forecasting*” dengan sinonimnya digunakan untuk menyusun literatur ini (Tabel 1). Kriteria inklusi dalam studi literatur ini adalah sebagai berikut (1) literatur membahas kemampuan kecerdasan buatan berbasis *neural network* sebagai prediktor kadar *particulate matter* 2,5 (PM2,5) dan *particulate matter* 10 (PM10); (2) literatur menilai keakuratan prediksi algoritma ANN terhadap kadar PM2,5 dan PM10 (3) literatur merupakan penelitian eksperimental. Hasil dari pencarian ditinjau terlebih dahulu dengan melakukan skrining judul dan abstrak untuk menilai relevansi literatur. Studi yang relevan diperiksa kembali dengan membaca literatur secara lengkap dan dinilai berdasarkan kriteria inklusi. Hasil dari pencarian ditelaah oleh dua penulis (RS dan LA) dan bila

ada perbedaan pendapat dimediasi oleh satu penulis (NT). Semua perbedaan pendapat diselesaikan dengan diskusi. Semua pencari-

an diselesaikan pada tanggal 23 Desember 2020.



Gambar 1. Diagram Prisma Penelitian

## HASIL

Hasil pencarian literatur yang didapat berjumlah 157 literatur, yakni 97 literatur dari *PubMed*, 13 literatur dari *ProQuest*, 12 literatur dari *EBSCOhost*, dan 36 literatur dari

*ScienceDirect*. Berikutnya, 15 duplikat dieksklusi menggunakan *EndNote X9*. Kemudian, tahap skrining judul dan abstrak yang dilakukan peneliti mengeksklusi 88 literatur. Setelah membaca literatur secara

penuh, penulis mengeksklusi 34 literatur dari 54 literatur yang tersisa sehingga mendapatkan 20 literatur yang sesuai dengan kriteria inklusi dan dipakai dalam studi ini.

Penulis kemudian menelaah dan mengkaji setiap artikel pada literatur ini dengan mengkompilasi hasil dari setiap studi dalam sebuah tabel (Tabel 2).

**Tabel 1.** Hasil Pencarian Database

Databases	Key Words	Articles
PubMed	((("Air Pollution"[Mesh] OR ("Air Pollution"[Title/Abstract]) OR ("Air Pollutions"[Title/Abstract]) OR ("Air Quality"[Title/Abstract])) AND (("Artificial Intelligence"[Mesh]) OR ("Artificial Intelligence"[Title/Abstract]) OR ("Computational Intelligence"[Title/Abstract]) OR ("Artificial Intelligence"[Title/Abstract]) OR ("Machine Intelligence"[Title/Abstract]) OR ("Computer Reasoning"[Title/Abstract]) OR (AI[Title/Abstract]) OR ("Computer Vision Systems"[Title/Abstract]) OR ("Computer Vision System"[Title/Abstract]) OR ("Knowledge Acquisition"[Title/Abstract]) OR ("Knowledge Representation"[Title/Abstract]) OR ("Knowledge Representations"[Title/Abstract]))) AND (((("Forecasting"[Mesh]) OR (Forecasting[Title/Abstract]) OR (Projections[Title/Abstract] AND Predictions[Title/Abstract])) OR (Predictions[Title/Abstract] AND Projections[Title/Abstract])) OR (Future[Title/Abstract])) OR (Futurology[Title/Abstract]))	97
ProQuest	ab("air pollution" OR "Air Pollutions" OR "Air Quality") AND ab("Artificial Intelligence" OR "Computational Intelligence" OR "Machine Intelligence" OR "Computer Reasoning AI" OR "Computer Vision Systems" OR "Computer Vision System" OR "Knowledge Acquisition" OR "Knowledge Representation" OR "Knowledge Representations") AND ab(Forecasting OR (Projections AND Predictions) OR (Predictions AND Projections) OR Future OR Futurology)	13
ScienceDirect	("air pollution" OR "Air Pollutions" OR "Air Quality") AND ("Artificial Intelligence" OR "Computational Intelligence" OR "Machine Intelligence") AND (Forecasting OR Projection OR Prediction)	36
EBSCO	"AB ( "air pollution" OR "Air Pollutions" OR "Air Quality" ) AND AB ("Artificial Intelligence" OR "Computational Intelligence" OR "Machine Intelligence" OR "Computer Reasoning AI" OR "Computer Vision Systems" OR "Computer Vision System" OR "Knowledge Acquisition" OR "Knowledge Representation" OR "Knowledge Representations") AND AB (Forecasting OR (Projections AND Predictions) OR (Predictions AND Projections) OR Future OR Futurology)	12

**Tabel 2.** Karakteristik Studi

Penulis	Desain Studi	Tempat Studi	Algoritma yang Dipakai	Hasil
Bozdağ, et al., 2020	Eksperimental	Ankara, Turki	LASSO, SVR, RF, kNN, xGBoost, dan ANN	Dibandingkan dengan algoritma lainnya performa terbaik dicapai oleh ANN untuk memprediksi kadar PM <sub>10</sub> .
Dong, et al., 2020	Eksperimental	Tiongkok	RF dan BPNN	BPNN dibandingkan RF lebih sensitif terhadap data geografis dalam memprediksi konsentrasi PM <sub>2,5</sub> .
Dunea, et al., 2015	Eksperimental	Romania	FANN dan WFANN	Algoritma WFANN tidak meningkatkan performa dalam memprediksi konsentrasi PM <sub>2,5</sub> dan PM <sub>10</sub> .

**Potensi Algoritma Berbasis Neural Network dan Turunannya Sebagai Prediktor Kadar PM<sub>2,5</sub> dan PM<sub>10</sub>:  
Suatu Telaah Sistematis**

<b>Penulis</b>	<b>Desain Studi</b>	<b>Tempat Studi</b>	<b>Algoritma yang Dipakai</b>	<b>Hasil</b>
Hur, <i>et al.</i> , 2016	Eksperimental	Seoul, Korea Selatan	<i>Neural Network Model</i>	Secara keseluruhan model <i>neural network</i> dapat dipercaya sebagai referensi yang reliabel untuk kebijakan pemerintah.
Kurt, <i>et al.</i> , 2008	Eksperimental	Istanbul, Turki	<i>Neural Network Model</i>	Model <i>neural network</i> digunakan secara online untuk memprediksi kadar polutan di udara didapatkan hasil yang memuaskan.
Li, <i>et al.</i> , 2017	Eksperimental	Beijing, Tiongkok	STDL, TDNN, ARMA, SVR, LSTM NN, dan LSTME	LSTME memiliki performa yang terbaik dibandingkan algoritma lainnya dalam mengaitkan data spasiotemporal, sehingga dapat menghasilkan prediksi yang lebih baik.
Ragosta, <i>et al.</i> , 2015	Eksperimental	Basilicata, Italia	ANN	Menggunakan parameter meteorologi membantu dalam memprediksi konsentrasi puncak dari PM <sub>10</sub> .
Chellali, <i>et al.</i> , 2016	Eksperimental	El Hamma, Algeria	<i>Artificial Neural Network (ANN)</i>	ANN lebih bagus memprediksi kualitas udara dengan data meteorologi model 3, yaitu kelembapan relatif.
Fernando, <i>et al.</i> , 2012	Eksperimental	Phoenix, Arizona	EnviINNet dan CMAQ	EnviINNet memiliki kemampuan yang lebih unggul dalam memprediksi PM <sub>10</sub> dibandingkan CMAQ.
Lu, <i>et al.</i> , 2004	Eksperimental	Mong Kok, Hong Kong	PCA/RBF dan simpel RBF	Sistem MLP (Multi-Layered Perceptron) berbasis PCA/RBF memberikan hasil prediksi polusi udara yang lebih baik dibandingkan metode simpel RBF.
Lu, <i>et al.</i> , 2005	Eksperimental	Hong Kong	SVM dan RBF	SVM memiliki kemampuan yang lebih akurat dalam memprediksi kadar polusi udara dan memberikan hasil yang lebih tergeneralisasi dibandingkan dengan RBF.
Pone, <i>et al.</i> , 2015	Eksperimental	Kallilo dan Vallila, Finlandia	ANN dan HS	Sistem prediksi dengan koreksi HS dalam mengukur kadar polusi udara memberikan performa yang lebih bagus dibandingkan dengan sistem prediksi dengan koreksi ANN.
Zhou, <i>et al.</i> , 2019	Eksperimental	Taipei, Taiwan	DM-LTSM dan SM-LTSM	DM-LTSM memiliki nilai RMSE yang lebih rendah sehingga memberikan prediksi kadar udara yang lebih akurat daripada SM-LTSM.

Penulis	Desain Studi	Tempat Studi	Algoritma yang Dipakai	Hasil
Zhang, et al., 2020	Eksperimental	Tiongkok	BiAGRU dan <i>machine learning</i> konvensional seperti LSTM	BiAGRU memberikan kemampuan yang lebih baik dibandingkan dengan <i>machine learning</i> konvensional dengan memprediksi kualitas polusi udara dengan bantuan sistem spatiotemporal.
Zhang, et al., 2017	Eksperimental	Sham Shui Po dan Tap Mun , Hong Kong	ELM, MLR, dan FFANN-BP	ELM memiliki keunggulan dalam beberapa metrik dibandingkan dengan MLR dan FFANN-BP dalam mengukur PM <sub>2,5</sub> .
Thomas dan Jacko, 2007	Eksperimental	Borman, India	Model Regresi dan MLP	Model regresi memberikan prediksi PM <sub>2,5</sub> yang sedikit lebih akurat dari berbagai metrik dibandingkan dengan MLP.
Carnevale, et al., 2011	Eksperimental	Milan, Italia	ANN, <i>Cokriging Technique</i>	ANN dapat dikombinasikan dengan <i>Cokriging Technique</i> untuk memprediksi kadar PM <sub>10</sub> secara <i>real time</i> .
Paschalidou, et al., 2010	Eksperimental	Cyprus, <i>Republic of Cyprus</i>	NN MLP, NN RBF	NN MLP memiliki nilai R <sub>2</sub> yang lebih mendekati 1 daripada NN RBF sehingga dapat lebih baik memprediksi kadar PM <sub>10</sub> .
Sun, et al., 2014	Eksperimental	Iowa. Amerika Serikat	NN RBF	NN RBF dapat memprediksi tingkat polusi udara lebih baik jika dicampurkan dengan analisis multivariat.
Voukantsis, et al., 2011	Eksperimental	Thessaloniki dan Helsinki	ANN-MLP	ANN-MLP dapat memprediksi kadar PM dan MLP akan meningkatkan akurasi dari prediksi.

## DISKUSI

Algoritma ANN sudah banyak digunakan di berbagai negara sebagai alat prediksi kadar polusi udara yang akurat. Seperti pada penelitian yang dilakukan Bozdag, et al. yang menggunakan ANN untuk memprediksi kadar PM10 di daerah urban dan industri. Bozdag, et al. mendapatkan bahwa ANN dapat memprediksi kualitas udara di daerah tersebut secara akurat sesuai dengan tingkat urbanisasi dan industrialisasi di kota Ankara, Turki.<sup>7</sup> Berbagai studi yang dilakukan oleh Hur, et al., Kurt, et al., Ragosta, et al., Chellali, et al., dan

Voukantsis, et al. menyatakan bahwa penggunaan ANN dengan *input* parameter meteorologi dapat memberikan hasil yang memuaskan dalam memprediksi kadar polusi udara terutama kadar PM<sub>2,5</sub> dan PM<sub>10</sub>. Hasil prediksi yang didapatkan juga dapat menjadi referensi bagi pemerintah untuk menerapkan kebijakan yang lebih baik dalam mengatasi masalah polusi udara.<sup>8-12</sup>

Parameter *input* yang digunakan studi dapat memengaruhi keakuratan algoritma. Beberapa studi yang sudah dilaksanakan telah memakai berbagai jenis parameter *input*,

seperti data konsentrasi polusi tahun sebelum penelitian dilaksanakan, data meteorologi, dan data gabungan antara meteorologi dengan data lainnya, seperti kondisi geografis.<sup>7-10,13</sup> Studi-studi tersebut menyimpulkan ANN dapat memprediksi lebih baik dengan memasukkan data geografis ke dalam algoritma ANN.<sup>13</sup>

*Artificial Neural Network* adalah salah satu model algoritma berbasis *neural network* terbaik dalam memprediksi kadar PM2,5 dan kadar PM10. Perkembangan dan penyempurnaan ANN karena berjalannya perkembangan teknologi menghasilkan hasil prediksi ANN yang semakin akurat. Sebuah studi menunjukkan jika algoritma berbasis *neural network* memiliki keunggulan dalam keakuratan prediksi kadar PM10 dibandingkan dengan jenis algoritma lainnya, seperti *Random Forest* (RF), *Least Absolute Shrinkage and Selection Operator* (LASSO), *Support Vector Regression* (SVR), *k-nearest neighbor* (kNN), *eXtreme gradient boosting* (xGBoost), dan *community air quality modelling system* (CMAQ).<sup>7,14,15</sup>

*Neural network* terdiri dari beberapa tipe seperti *support vector machine* (SVM) dan *radial basis function* (RBF). Tiap tipe memiliki cara penghitungan algoritma yang berbeda. Lu, *et al.* membandingkan algoritma SVM dengan RBF untuk memprediksi kadar PM10 dan menemukan jika *neural network* tipe SVM dapat memprediksi lebih baik daripada RBF.<sup>14</sup> Beberapa studi perbandingan *neural network* dengan algoritma lainnya, seperti RF, *Spatio-temporal Deep Learning* (STDL), *autoregressive moving average* (ARMA), dan SVR yang dilakukan oleh Li, *et al.* dan Dong, *et al.*

mendapatkan hasil yang sama, yaitu keunggulan *neural network* dalam melakukan prediksi PM2,5.<sup>13-16</sup> Selain membandingkan ANN dengan algoritma kecerdasan buatan lainnya, terdapat satu studi yang membandingkan ANN dengan metode statistik yaitu *multiple linear regression* (MLR). Studi tersebut menjelaskan jika salah satu jenis model *neural network* yang Zhang, *et al.* gunakan yaitu, *extended machine learning* (ELM) memiliki derajat *error* lebih rendah dan dapat memproses data *input* dengan lebih cepat dibandingkan dengan MLR dan *feed forward artificial neural network-back propagation* (FFANN-BP).<sup>17</sup> Namun, tingkat efektivitas ANN dibandingkan dengan model regresi masih diperlukan uji lebih lanjut oleh karena Thomas dan Jacko mendapatkan hasil yang berbeda dengan Zhang, *et al.*, yang menyatakan model regresi memberikan hasil yang lebih akurat dibandingkan dengan model MLP (*multi-layered perceptron*) dalam memprediksi kadar PM2,5.<sup>18</sup>

Berdasarkan studi yang telah dipaparkan, maka dapat disimpulkan jika algoritma berbasis ANN ini mendapatkan hasil yang akurat dalam memprediksi kadar PM2,5 dan PM10.<sup>8-12</sup> Secara garis besar algoritma berbasis ANN ini dapat memberikan hasil yang lebih akurat dibandingkan dengan beberapa algoritma lainnya, tetapi terdapat beberapa studi yang mendapatkan bahwa kemampuan algoritma berbasis ANN kurang menghasilkan lebih baik dari algoritma seperti SVM dan RF.<sup>7,13,15-18</sup> Hal tersebut dimungkinkan karena penggunaan model *neural network* yang telah dikombinasi dengan komponen lain pada

beberapa studi sehingga terjadi peningkatan akurasi prediksi. Oleh karena itu, model ANN yang dipakai oleh sebuah studi dapat berpengaruh terhadap hasil prediksi PM<sub>2,5</sub> dan PM<sub>10</sub>.

Walaupun ANN sudah banyak digunakan di berbagai belahan dunia, algoritma ini masih belum sempurna dan perlu dikombinasikan dengan model lain untuk meningkatkan kinerjanya. Studi dari Sun, *et al.* dan Lu, *et al.* menyatakan bahwa penggunaan MLP yang diasosiasikan dengan model simple RBF dan *principal component analysis/radial basis function* (PCA/RBF) dapat membantu menambah akurasi ANN dalam memprediksi kadar PM<sub>10</sub> di udara. Kombinasi MLP dengan ANN dapat mencegah terjadinya *overfitting* dan kelebihan berat pada *hidden layer* yang mengganggu kinerja ANN.<sup>19-20</sup>

Walaupun beberapa penelitian mendukung ANN berjenis RBF, ada beberapa studi yang mengajukan penggunaan model lain yang memiliki kemampuan lebih unggul dari RBF. Paschalidou, *et al.* mengajukan model algoritma *neural network* lainnya, seperti NN-MLP yang memiliki kemampuan prediksi lebih baik daripada NN-RBF dan NN *principal component regression analysis* (PCRA) karena memiliki nilai regresi yang mendekati satu.<sup>21</sup>

Beberapa studi lainnya oleh Dong, *et al.* dan Dunea, *et al.* meneliti jenis model ANN lainnya yaitu *back-propagation neural network* (BPNN) dan *wavelet-feedforward artificial neural network* (WFANN). Model BPNN yang digunakan Dong, *et al.* membuktikan performa yang baik dalam memprediksi kadar PM<sub>2,5</sub>; sedangkan WFANN dari Dunea, *et al.* diduga

tidak meningkatkan performa dari ANN dalam memprediksi kadar PM<sub>10</sub>.<sup>13,22</sup>

Pone, *et al.* menggunakan kombinasi ANN dan *hybrid system* (HS) pada penelitiannya dan menemukan hasil kombinasi tersebut akan menghasilkan pengukuran metrik yang lebih akurat dalam memprediksi kadar PM<sub>2,5</sub> dan PM<sub>10</sub>.<sup>23</sup> Hasil yang lebih efektif dan akurat juga didapatkan dalam studi Carnevalle, *et al.* yang menggabungkan ANN dan teknik *cokriging*.<sup>24</sup> Tidak hanya menggunakan teknik lain untuk dikombinasikan dengan ANN, beberapa studi seperti Li, *et al.* dan Zhou, *et al.* juga mengintegrasikan memori dalam *neural network*. Li, *et al.* menggunakan model *long short-term memory neural network extended* (LSTME) yang merupakan salah satu jenis ANN bertipe *long short-term memory neural network* (LTSM). *Artificial Neural Network* tipe ini memanfaatkan data *spatiotemporal* untuk memprediksi kadar PM, sedangkan Zhou, *et al.* menggunakan *Shallow Deep Multi-output LSTM* (DM-LSTM), sebuah pengembangan dari *Shallow Multi-output Long Short-Term Memory* (SM-LSTM), dengan integrasi berbagai teknik optimasi seperti *mini-batch gradient descent* (MBGD) *algorithm*, *dropout neuron algorithm*, dan *L2 regularization algorithm*. Teknik optimasi seperti MBGD berperan penting dalam mempercepat proses pembelajaran dengan membagi data menjadi kelompok-kelompok kecil (*mini-batches*), sehingga meningkatkan efisiensi komputasi dan mempercepat konvergensi model. Sementara itu, algoritma *dropout neuron* dan *L2 regularization* berfungsi untuk mengurangi risiko *overfitting*, meningkatkan

keandalan model, dan memastikan generalisasi yang lebih baik pada data baru dengan *dropout neuron* yang bekerja dengan mematikan secara acak sejumlah neuron selama pelatihan dan *L2 regularization* yang memberikan nilai penalti pada model agar bobot model tetap kecil dan terkontrol. Optimalisasi semacam ini menjadikan DM-LSTM lebih unggul dalam menangani tantangan yang dihadapi oleh SM-LSTM.<sup>16,25</sup> Namun, model LSTM masih belum dapat dikatakan sempurna karena terdapat sebuah penelitian terbaru oleh Zhang, *et al.* yang menemukan algoritma *bidirectional gated recurrent unit integrated with attention mechanism* (BiAGRU), sebuah modifikasi *recurrent neural network* (RNN) jenis *gated recurrent unit* (GRU), dengan nilai *error* yang lebih rendah daripada LSTM.<sup>26</sup> Dinamika kerja RNN, termasuk GRU, didasarkan pada kemampuan memproses data sekuensial secara lebih efisien dengan mempertahankan informasi penting dari input sebelumnya melalui mekanisme *gating*. Kombinasi dengan mekanisme perhatian (*attention mechanism*) dalam BiAGRU memungkinkan model untuk lebih fokus pada elemen-elemen data yang relevan, sehingga meningkatkan akurasi prediksi. *Artificial Neural Network* juga memiliki jaringan neuron seperti RNN dan memiliki kemampuan dalam menganalisis data sekuensial seperti deret waktu atau data polusi udara.<sup>26</sup>

Studi-studi di atas menunjukkan penggunaan algoritma turunan ANN yang beragam sehingga algoritma-algoritma tersebut dapat diaplikasikan untuk mengukur kadar polusi udara. Kombinasi penggunaan algoritma ter-

sebut menguntungkan disebabkan karena kelemahan dari algoritma berbasis *neural network* ini dapat diatasi, sehingga meningkatnya kapasitas prediksi algoritma-algoritma yang merupakan turunan dari ANN. Algoritma ini masih memiliki ruang untuk dikembangkan lebih lanjut agar dapat memprediksi kadar polusi udara lebih optimal.

Berdasarkan hasil yang diinklusi pada tinjauan pustaka ini, algoritma berbasis *neural network* dapat memprediksi kadar PM10 dan PM2,5 secara akurat.<sup>7-12</sup> Walaupun tingkat akurasi ANN dalam memprediksi masih menjadi kontradiksi dalam beberapa studi, tingkat akurasi ANN masih dapat ditingkatkan karena penggunaan input data yang masih belum maksimal dan belum tereksplorasi lebih lagi terutama data geografis yang merupakan salah satu faktor data yang sensitif untuk ANN.<sup>7,14-18</sup> Selain itu, ANN juga memiliki potensi untuk dipadukan dengan beberapa algoritma lainnya. Kombinasi ini dapat mencegah kekurangan dan meningkatkan potensi dari kelebihan ANN sendiri.<sup>16,19-21,23,26</sup>

Penelitian-penelitian yang telah dilakukan menggunakan beberapa jenis ANN yang berbeda, seperti Bozdag, *et al.* yang membandingkan ANN dengan algoritma lainnya, seperti LASSO, SVR, RF, dan beberapa algoritma lainnya menemukan ANN lebih unggul daripada algoritma lainnya. Hasil yang berbeda diperoleh Lu, *et al.* yang membandingkan ANN dengan jenis yang berbeda yaitu RBF dan dibandingkan dengan SVM. Hasil studi Lu, *et al.* menemukan jika SVM memang lebih unggul daripada ANN-RBF karena sistem pada SVM yang dapat meminimalkan

*error.*<sup>7,14</sup> Studi oleh Thomas dan Jacko juga menyatakan bahwa model regresi yang menggunakan sistem statistik diduga memiliki keunggulan yang sedikit lebih baik daripada ANN-MLP dalam mengukur kadar PM<sub>2,5</sub>.<sup>18</sup>

Penggunaan ANN dalam memprediksi kadar PM<sub>2,5</sub> dan PM<sub>10</sub> memerlukan suatu pelatihan dan simulasi berulang sehingga dibutuhkan waktu agar akurasi prediksi ANN semakin meningkat. *Artificial Neural Network* juga perlu dipadukan dengan algoritma lain untuk meningkatkan akurasi hasil prediksi dan mengurangi nilai *error*, serta menggunakan data yang lebih luas seperti geografis, meteorologi, dan spatiotemporal. Melalui penyempurnaan algoritma kombinasi ANN dan pelelangkapan data *input* diharapkan ANN dapat memprediksi kadar polusi udara, memperbaiki kualitas udara, dan mencegah terjadinya penumpukan polusi udara di masa mendatang.

Algoritma ANN memiliki keunggulan dibandingkan beberapa metode lain yang pernah digunakan sebelumnya seperti metode regresi dan metode numerik. Algoritma ANN menggunakan metode *artificial intelligence* yang menirukan model seperti sistem saraf manusia. Metode ANN dapat memberikan korelasi antara kadar polusi udara dengan variabel meteorologi seperti suhu, kecepatan udara, dan kelembaban relatif secara akurat dibanding model lainnya. Lebih dari satu jenis data meteorologi dapat dimasukkan dalam kalkulasi ANN sehingga kadar polusi udara dapat dikorelasikan dengan berbagai variabel meteorologi. Hal ini membuat kita dapat mengetahui variabel meteorologi yang lebih berperan dalam meningkatkan kadar polusi

udara dan variabel meteorologi yang kurang berperan. Selain itu, keunggulan lainnya adalah kemampuan ANN dalam memproses data yang terbilang cepat dan praktis dibandingkan model lainnya.<sup>11</sup>

Walaupun ANN sudah dimanfaatkan dalam memprediksi kadar polusi udara, ANN masih dapat melakukan kesalahan dalam memprediksi kadar polusi udara yang ada.<sup>11</sup> Kesalahan tersebut seringkali terkait dengan tantangan dalam mengelola data besar dan risiko *overfitting*. Kondisi *overfitting* ini terjadi akibat ketidakseimbangan atau kompleksitas pola data yang tidak diimbangi dengan kapasitas model yang memadai. Dalam konteks ini, analisis akurasi (*accuracy*) dari algoritma ANN menjadi sangat penting. Aspek lain dari *confusion matrix*, seperti *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN), juga perlu diperhatikan karena nilai-nilai ini memberikan informasi tambahan tentang *precision*, *sensitivity*, dan *specificity model*, yang berperan penting dalam mendekripsi potensi *overfitting*.<sup>27</sup> *Artificial Neural Network* yang memanfaatkan data besar harus menjalani banyak latihan dan percobaan agar dapat mengurangi kejadian *local minima* (pengurangan fungsi ANN) atau dengan menggunakan teknik optimasi seperti *dropout neuron algorithm* atau *regularization* sehingga memberikan hasil yang lebih akurat. Banyak pelatihan yang dilakukan tentunya akan memakan waktu yang lebih lama. Setiap pelatihan ANN bersifat spesifik dikarenakan struktur model jaringan ANN yang berbeda-beda sehingga solusi dari suatu ANN mungkin tidak dapat diaplikasikan dalam jenis ANN

lainnya. Model struktur jaringan yang cocok untuk menyelesaikan beberapa masalah secara efektif juga menjadi sebuah kesulitan dalam meningkatkan kinerja ANN.<sup>14</sup>

Namun, beberapa kelemahan ANN sudah dapat ditangani, seperti *overfitting* dengan menggabungkan sistem MLP dengan PCA/RBF dengan PCA merupakan teknik yang berfungsi dalam mengurangi dimensi data dengan menurunkan jumlah fitur yang kurang signifikan (*noise* atau fitur korelasi tinggi) sehingga mengurangi *overfitting* dan meningkatkan akurasi karena hanya berfokus pada fitur-fitur yang signifikan, sedangkan RBF adalah jenis fungsi aktivasi yang dapat menerapkan sistem *local learning* yang membuat jarak antara *input* dan *centroid* tertentu yang membuatnya lebih fokus pada subset data lokal dan dapat menghasilkan fungsi aktivasi yang lebih terlokalisasi sehingga mengurangi dampak data *outlier* dan *noise*. *Radial Basis Function* juga memiliki fleksibilitas non-linear yang dapat menangkap hubungan non-linear yang kompleks dalam data dengan akurasi tinggi. Secara keseluruhan sistem PCA/RBF dinilai lebih efektif daripada sistem *simple RBF*. Sistem PCA/RBF ini juga terbukti memiliki *mean absolute error* (MAE) (fungsi *loss* untuk mengevaluasi kinerja model dengan menghitung rata-rata perbedaan absolut antara prediksi model dan nilai sebenarnya) yang lebih rendah dibandingkan dengan *simple RBF* sehingga kemungkinan *error* dapat diatasi.<sup>20</sup> Selain itu, solusi lain untuk mencegah *overfitting* ditemukan dalam studi Zhou, *et al.* yang menggunakan algoritma DM-LSTM dan SM-LTSM. Algoritma DM-

LSTM lebih unggul dalam memprediksi kadar polusi udara karena didukung dengan tiga algoritma yaitu *mini-batch gradient descent* (MBGD) *algorithm*, *dropout neuron algorithm*, dan *L2 regularization algorithm* yang dapat mengatasi instabilitas dan *overfitting* serta mempercepat pemrosesan data.<sup>25</sup>

Pengupayaan kecerdasan buatan sebagai suatu alat untuk memprediksi polusi udara telah diterapkan pada beberapa negara. Salah satu contoh penerapan dari kecerdasan buatan ini adalah penggunaan ANN sebagai sebuah sistem peringatan dini. Sistem peringatan dini ini dapat berfungsi sebagai rujukan masyarakat untuk beraktivitas berdasarkan kualitas udara terutama orang-orang yang perlu hidup dengan kualitas udara yang lebih baik seperti lansia, pengidap asma, dan anak-anak.<sup>4</sup> Penggunaan ANN juga telah diterapkan secara daring berbasis *website* pada kota Istanbul yaitu *AirPolTool* sehingga masyarakat dapat melihat hasil kualitas udara yang ada di lingkungannya. Inovasi ini dapat menjadi referensi dalam penentuan kebijakan pemerintah setempat ke depannya.<sup>9</sup>

Terdapat beberapa keterbatasan dalam tinjauan pustaka ini. Pertama, variasi algoritma ANN dalam studi inklusi, data *input* yang berbeda, serta lokasi studi yang berbeda sehingga terdapat beberapa faktor geografis dan meteorologi yang berbeda pula. Hal-hal tersebut mungkin memengaruhi hasil studi. Namun, dalam studi ini hanya diinklusikan 1 penggunaan algoritma *artificial intelligence* yakni ANN sehingga data menjadi lebih homogen dibandingkan jika terdapat 2 algoritma dalam kata kunci. Pengukuran keaku-

ratan algoritma juga telah distandarisasi dengan tiga parameter yaitu *mean squared error* (MSE), *mean absolute percentage error* (MAPE), dan *the squared correlation coefficient* (R<sup>2</sup>), sehingga dapat dilakukan perbandingan antar studi.

## SIMPULAN

Hasil studi menunjukkan bahwa model algoritma ANN akurat dalam memprediksi kadar PM<sub>2,5</sub> dan PM<sub>10</sub> yang merupakan salah satu parameter polusi udara. Berdasarkan studi yang diinklusi, ANN memiliki kemampuan prediksi yang superior dibandingkan dengan RF, LASSO, SVR, kNN, xGBoost, dan CMAQ dalam memprediksi PM<sub>10</sub> dan RF, STDL, ARMA, dan SVR dalam memprediksi PM<sub>2,5</sub>. Algoritma ANN dapat dikembangkan lebih lanjut melalui beberapa hal seperti menambahkan data input geografis untuk menambah performa prediksi. Walaupun ANN memiliki beberapa kekurangan, seperti fenomena *overfitting* dalam penggunaan tunggal dan *error* ketika mendapatkan hasil, kekurangan ini dapat diatasi dengan kombinasi ANN dan struktur model ANN lainnya. Kombinasi tersebut dapat mengatasi masalah-masalah seperti yang sudah dipaparkan oleh berbagai studi sebelumnya. Penggunaan ANN dapat mengukur kadar PM<sub>2,5</sub> dan PM<sub>10</sub> sebagai faktor penentu kualitas udara, serta dapat digunakan sebagai salah satu instrumen pertimbangan kebijakan negara dalam pencegahan polusi udara.

## DAFTAR PUSTAKA

1. Prüss-Ustün A, van Deventer E, Mudu P, Campbell Lendrum D, Vlckers C, Ivanov I, et al. Environmental risks and non-communicable diseases. *BMJ*. 2019;364:I265.
2. Chen S, Bloom DE. The macroeconomic burden of noncommunicable diseases associated with air pollution in China. *PLoS One*. 2019;14(4):e0215663.
3. Fuller R, Rahona E, Fischer S, Caravanos J, Webb D, Kass D, et al. Pollution and non-communicable disease: time to end the neglect. *Lancet Planet Health*. 2018;2(3):e96–e98.
4. Mo X, Zhang L, Li H, Qu Z. A novel air quality early-warning system based on artificial intelligence. *Int J Environ Res Public Health*. 2019;16(19):3505.
5. Asadollahfardi G, Mehdinejad M, Pam M, Parisa P, Asadollahfardi R, Farnad M. Predicting carbon monoxide concentrations in the air of Pardis City, Iran, using an artificial neural network. *Environ Qual Manag*. 2016;26(1):37–49.
6. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
7. Bozdağ A, Dokuz Y, Göçek ÖB. Spatial prediction of PM<sub>10</sub> concentration using machine learning algorithms in Ankara, Turkey. *Environ Pollut*. 2020;263(Pt A):114635.
8. Hur SK, Oh HR, Ho CH, Kim J, Song CK, Chang LS, et al. Evaluating the predictability of PM<sub>10</sub> grades in Seoul, Korea using a neural network model based on synoptic patterns. *Environ Pollut*. 2016;218:1324–33.
9. Kurt A, Gulbagci B, Karaca F, Alagha O. An online air pollution forecasting system using neural networks. *Environ Int*. 2008;34(5):592–8.
10. Ragosta M, D'Emilio M, Giorgio GA. Input strategy analysis for an air quality data modelling procedure at a local scale based on neural network. *Environ Monit Assess*. 2015;187(5):307.
11. Chellali MR, Abderrahim H, Hamou A, Nebatti A, Janovec J. Artificial neural network models for prediction of daily fine particulate matter concentrations in Algiers. *Environ Sci Pollut Res Int*. 2016;23(14):14008–17.
12. Voukantsis D, Karatzas K, Kukkonen J, Räsänen T, Karppinen A, Kolehmainen M. Intercomparison

- of air quality data using principal component analysis, and forecasting of PM10 and PM2.5 concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Sci Total Environ.* 2011;409(7):1266–76.
13. Dong L, Li S, Yang J, Shi W, Zhang L. Investigating the performance of satellite-based models in estimating the surface PM2.5 over China. *Chemosphere.* 2020;256:127051.
14. Lu WZ, Wang WJ. Potential assessment of the ‘support vector machine’ method in forecasting ambient air pollutant trends. *Chemosphere.* 2005;59(5):693–701.
15. Fernando HJS, Mammarella MC, Gradoni G, Fedele P, Di Marco R, Dimitrova R, et al. Forecasting PM10 in metropolitan areas: efficacy of neural networks. *Environ Pollut.* 2012;163:62–7.
16. Li X, Peng L, Yao X, Cui S, Hu Y, You C, et al. Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation. *Environ Pollut.* 2017;231(Pt 1):997–1004.
17. Zhang J, Ding W. Prediction of air pollutants concentration based on an extreme learning machine: the case of Hong Kong. *Int J Environ Res Public Health.* 2017;14(2):214.
18. Thomas S, Jacko RB. Model for forecasting expressway fine particulate matter and carbon monoxide concentration: application of regression and neural network models. *J Air Waste Manag Assoc.* 2007;57(4):480–8.
19. Sun G, Hoff SJ, Zelle BC, Nelson MA. Forecasting daily source air quality using multivariate statistical analysis and radial basis function networks. *J Air Waste Manag Assoc.* 2008;58(12):1571–8.
20. Lu WZ, Wang WJ, Wang XK, Yan SH, Lam JC. Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air, Hong Kong. *Environ Res.* 2004;96(1):79–87.
21. Paschalidou AK, Karakitsios S, Kleanthous S, Kassomenos PA. Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. *Environ Sci Pollut Res Int.* 2011;18(2):316–27.
22. Dunea D, Pohoata A, Iordache S. Using wavelet-feedforward neural networks to improve air pollution forecasting in urban environments. *Environ Monit Assess.* 2015;187(7):477.
23. de Mattos Neto PSG, Cavalcanti GDC, Madeiro F, Ferreira TAE. An approach to improve the performance of PM forecasters. *PLoS One.* 2015;10(9):e0138507.
24. Carnevale C, Finzi G, Pisoni E, Singh V, Volta M. An integrated air quality forecast system for a metropolitan area. *J Environ Monit.* 2011;13(12):3437–44.
25. Zhou Y, Chang F, Chang L, Kao IF, Wang Y. Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *J Clean Prod.* 2019;209:134–45.
26. Zhang K. Multi-step ahead forecasting of regional air quality using spatial-temporal deep neural networks: a case study of Huaihai Economic Zone. *J Clean Prod.* 2020;277:123231.
27. Kaur M, Singh D, Jabarulla MY, Chahar V, Kang J, Lee HN. Computational deep air quality prediction techniques: a systematic review. *Artif Intell Rev.* 2023;56:1–46.