

NGHIÊN CỨU ỨNG DỤNG MÔ HÌNH TRONG DỰ BÁO CHẤT LƯỢNG KHÔNG KHÍ

Mạc Duy Hưng¹, Nghiêm Trung Dũng², Hoàng Xuân Cơ³

¹*Khoa Xây dựng và Môi trường, trường Đại học Kỹ thuật Công nghiệp Thái Nguyên*

²*Viện Khoa học và Công nghệ Môi trường, trường Đại học Bách Khoa Hà Nội*

³*Khoa Môi trường, trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội*

TÓM TẮT

Ba loại mạng thần kinh nhân tạo (ANN – Artificial Neural Network), bao gồm mạng nơron truyền thuận (FFNN – Feed-forward Neural Network), mạng nơron hồi quy đa lớp (MLP – Multi Layer Perceptron) và mạng trạng thái phản hồi (ESN – Echo State Network) đã được ứng dụng để xây dựng thử nghiệm mô hình dự báo thống kê chất lượng không khí tại Việt Nam. Bước đầu cho thấy, kết quả dự báo của các mô hình ANN nghiên cứu là rất khả quan. Độ chính xác đạt được của các mô hình này trên ba đối tượng nghiên cứu thử nghiệm, gồm SO₂, NO₂ và PM₁₀ đều đạt trên 60%, đặc biệt đạt đến 83,8% (đối với ESN) khi dự báo nồng độ PM₁₀. Kết quả này khá phù hợp với các mô hình cùng loại đã công bố trên thế giới. Điều này cho thấy, việc ứng dụng công cụ ANN trong dự báo chất lượng không khí tại Việt Nam là một hướng tiếp cận rất triển vọng và khả thi.

Từ khóa: Dự báo, chất lượng không khí, ANN, FFNN, MLP, ESN, Việt Nam

ABSTRACT

Three types of artificial neural networks (ANN), including feed-forward neural network (FFNN), multi-layer perceptron (MLP) and echo state network (ESN), have been applied to develop the forecasting models of air quality in Vietnam. Preliminary results indicated that, the performances of these models are very promising. Accuracies of forecasting values of MLP and ESN models are higher than 60%, especially; accuracy of ESN model is up to 83.8% for the forecasting of PM₁₀ concentration. The results are in the same level with other published ones of the same model. This suggests that, the application of ANN tools to develop the forecasting models of air quality is a promising and feasible approach.

Keywords: Forecast, air quality, ANN, FFNN, MLP, ESN, Vietnam.

I. ĐẶT VẤN ĐỀ

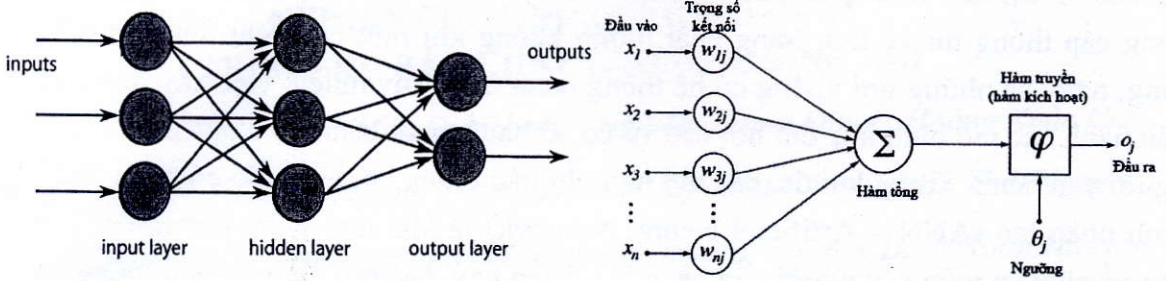
Trong hơn một thập kỷ trở lại đây, các mô hình dự báo chất lượng không khí đã và đang trở thành công cụ đắc lực trong công tác quản lý tổng hợp chất lượng không khí tại Việt Nam. Nó cung cấp thông tin một cách toàn diện và trực quan hơn về tình trạng cũng như diễn biến chất lượng không khí, nhờ đó, cơ quan hữu trách có thể đưa ra các giải pháp kịp thời để ứng phó với những ảnh hưởng tiêu cực do chất lượng không khí có thể gây ra cho người dân. Cho đến nay, các mô hình dự báo chất lượng không khí được ứng dụng tại Việt Nam chủ yếu là các loại mô hình

dự báo số trị, mô hình quang hóa và mô hình động lực học. Chúng có ưu điểm là cung cấp thông tin về tình trạng chất lượng không khí một cách chi tiết trên diện rộng, ngay cả những nơi không có hệ thống quan trắc. Tuy nhiên, việc xây dựng và vận hành các mô hình này đòi hỏi cao về cơ sở hạ tầng và trình độ chuyên môn của người vận hành. Trong khi đó, các mô hình dự báo thống kê, trong đó có mạng thần kinh nhân tạo (ANN – Artificial Neural Network) là khá đơn giản, không đòi hỏi cao về chuyên môn của người vận hành [1]. Hiện nay, các mô hình dự báo thống kê chất lượng không khí sử dụng mạng thần kinh nhân tạo đã được nghiên cứu và ứng dụng thành công tại nhiều quốc gia trên thế giới. Trong số đó, có thể kể đến như dự báo nồng độ bụi PM_{10} [2, 3], dự báo nồng độ ôzôn [4-7], hoặc một số chất ô nhiễm khác như SO_2 , NO_x , VOC... [8-14]. Điều đó cho thấy tiềm năng của chúng trong việc dự báo chất lượng môi trường là rất lớn. Bài báo này giới thiệu một số kết quả đạt được của mô hình dự báo thống kê chất lượng không khí đã được nghiên cứu áp dụng thành công tại Việt Nam.

II. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Cơ sở lý thuyết của mạng thần kinh nhân tạo (ANN)

Mạng thần kinh nhân tạo (ANN – Artificial Neural Network) được xây dựng trên cơ sở mô phỏng quá trình xử lý thông tin của một hệ thần kinh sinh học, bao gồm 03 hoặc nhiều lớp được cấu tạo từ các nơron. Trong đó, lớp đầu tiên được gọi là lớp đầu vào, lớp cuối cùng chứa kết quả được gọi là lớp đầu ra và (các) lớp còn lại được gọi là lớp ẩn (Hình 1). Mỗi nơron trong ANN nhận tín hiệu đầu vào và xử lý để đưa ra tín hiệu đầu ra, được kết nối với ít nhất một nơron khác, và mỗi kết nối này được thể hiện dưới dạng một số thực được gọi là trọng số liên kết - phản ánh mức độ quan trọng của liên kết. Tập hợp các trọng số này tạo thành một ma trận được gọi là ma trận trọng số liên kết $w_{i,j}$, nó mã hóa kiến thức (kinh nghiệm) tích lũy được của ANN từ tập dữ liệu đầu vào (Hình 2) [15]. Ưu điểm chính của một ANN là ở chỗ, nó có thể sử dụng một số thông tin ẩn trong bộ dữ liệu về mối liên hệ giữa các thông số đầu vào và đầu ra (các thông tin này không thể mô tả cụ thể bằng một hàm số toán học). Quá trình “nắm bắt” thông tin này được gọi là quá trình đào tạo ANN (quá trình học tập của một ANN). Về mặt nguyên tắc, học tập chính là quá trình hiệu chỉnh trọng số liên kết giữa các nơron.



Hình 1. Kiến trúc của một ANN cơ bản Hình 2. Mô hình toán một nơron nhân tạo

Quá trình xử lý thông tin của nơron thứ j (Hình 2) được mô tả vắn tắt như sau: Đầu tiên, nơron thứ j sẽ nhận các tín hiệu đầu vào x_i có trọng số liên kết w_{ij} tương ứng. Bộ tín hiệu này sẽ được tổng hợp thông qua hàm tổng (summing function) theo phương trình (1). Tiếp theo, tín hiệu đã được tổng hợp này sẽ được xử lý bởi hàm truyền (transfer function) tạo thành tín hiệu đầu ra (output), được mô tả theo phương trình (2).

$$\varphi(x) = \sum_{i=1}^n w_{ij}x_i - \theta_j \quad (1); \quad output = f\left(\sum_{i=1}^n w_{ij}x_i - \theta_j\right) \quad (2)$$

Về mặt bản chất, dự báo chất lượng không khí là quá trình ước tính giá trị nồng độ các chất ô nhiễm không khí dựa trên dữ liệu về nồng độ của chúng trong quá khứ và hiện tại. Điều này hoàn toàn giống với cơ chế xử lý thông tin dựa trên quá trình học tập hành vi thông qua những kinh nghiệm đã có của ANN. Chính vì vậy, bài này giới thiệu ứng dụng của ba dạng ANN, bao gồm mạng nơron truyền thuận (FFNN - Feed-Forward Neural Network), mạng nơron hồi quy đa lớp (MLP - Multi-Layer Perceptron) - là hai trong số những ANN được ứng dụng rộng rãi nhất trong lĩnh vực dự báo chất lượng không khí, và mạng trạng thái phản hồi (ESN - Echo State Network) để xây dựng mô hình dự báo chất lượng không khí tại Việt Nam.

2.2. Các bước tiến hành nghiên cứu

Việc ứng dụng mạng thần kinh nhân tạo để xây dựng thử nghiệm mô hình dự báo chất lượng không khí đã được thực hiện theo các bước sau: chuẩn bị dữ liệu, thiết kế mạng và xây dựng mô hình dự báo, đào tạo mạng và đánh giá độ tin cậy của mô hình.

Chuẩn bị dữ liệu: Dữ liệu sử dụng cho các mô hình nghiên cứu được trích xuất từ bộ dữ liệu quan trắc chất lượng không khí tại các trạm quan trắc chất lượng không khí ở Việt Nam. Bộ dữ liệu này được chia làm hai phần, một phần được dùng để đào tạo mô hình (training data), phần còn lại được dùng để đánh giá và kiểm tra độ tin cậy của mô hình đã xây dựng (testing data).



Thông thường, bộ mẫu dữ liệu đầu vào (input sample) cho mô hình được thiết kế bao gồm: thông số chất lượng không khí cần dự báo và các tham số có ảnh hưởng đến diễn biến nồng độ của nó trong không khí (như tham số khí tượng và nồng độ các chất tham gia trong phản ứng hóa học với/tạo ra chất ô nhiễm cần dự báo).

Thiết kế mạng và xây dựng mô hình dự báo: Kiến trúc ANN bao gồm số lớp ẩn, số nơron trong lớp ẩn, tốc độ học (learning rate) ... được thiết kế phù hợp với đối tượng nghiên cứu, dựa trên kinh nghiệm của người thiết kế mạng thông qua thực nghiệm hoặc có sự hỗ trợ của một số công cụ tối ưu hóa như thuật toán di truyền (GA – Genetic Algorithm).

Trong những thử nghiệm này, kiến trúc của ANN đã được xác định đối với một số đối tượng nghiên cứu tại Việt Nam như sau (Bảng 1):

Bảng 1. Kiến trúc ANNs đã được nghiên cứu tại Việt Nam [16, 17]

Loại mạng	Kiến trúc	Thông số cần dự báo	Biến đầu vào
FFNN	2 – 15 – 1	SO ₂ , PM ₁₀	SO _{2(t-n)} , t (giờ)
MLP	5 – 10 – 1	SO ₂ , NO ₂ , PM ₁₀	SO _{2(t-n)} ; NO _{2(t-n)} ; PM _{10(t-n)} WS; RH; T; Rain
ESN	5 – 50 – 1	SO ₂ , NO ₂ , PM ₁₀	SO _{2(t-n)} ; NO _{2(t-n)} ; PM _{10(t-n)} WS; RH; T; Rain

Đào tạo mạng: Các mô hình nghiên cứu đã được huấn luyện dựa trên kiến trúc mạng và bộ dữ liệu đào tạo (training data) thu được ở các bước trên. Về mặt bản chất, quá trình đào tạo là một quá trình thay đổi trọng số liên kết giữa các nơron để đưa ra kết quả phù hợp với kết quả mong muốn (target).

Đánh giá độ tin cậy của mô hình: Mô hình dự báo ứng dụng ANN là một trong những dạng mô hình dự báo thống kê. Chính vì vậy, độ tin cậy của các mô hình này đã được kiểm chứng dựa trên độ sai lệch của kết quả dự báo với giá trị thực (measured value) của bộ dữ liệu kiểm tra (testing data), thông qua một số chỉ số thống kê, bao gồm: RMSE (Root Mean Square Error), nRMSE (normalized Root Mean Square Error), MAE (Mean Absolute Error), MAPA (Mean Absolute Percentage Accuracy) được chỉ ra trong các phương trình dưới đây:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (C_i^{pred} - C_i^{observ})^2}$$

$$nRMSE = \frac{RMSE}{\frac{1}{N} \sum_{i=1}^N C_i^{observ}} \cdot 100\%$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |O_i - P_i|$$

$$MAPA = \left(1 - \frac{MAE}{\frac{1}{N} \sum_{i=1}^N P_i} \right) * 100\%$$

III. KẾT QUẢ VÀ THẢO LUẬN

Trong nghiên cứu này, mô hình sử dụng ANN để dự báo chất lượng không khí, bao gồm FFNN, ANN-MLP và ESN, đã được nghiên cứu thử nghiệm tại Việt Nam trên một số chất ô nhiễm không khí điển hình. Trong đó, FFNN được nghiên cứu thử nghiệm để bù dữ liệu quan trắc nồng độ SO_2 và PM_{10} bị khuyết thiếu dựa trên cơ chế dự báo; ANN-MLP và ESN được ứng dụng thử nghiệm để dự báo nồng độ SO_2 , PM_{10} và NO_2 cực đại ngày với chu kỳ dự báo 07 ngày. Kết quả thu được là rất khả quan, hiệu suất của các mô hình nghiên cứu là tương đương với các mô hình cùng loại đã được thực hiện trên thế giới. Kết quả cụ thể được trình bày trong Bảng 2, 3 và 4.

Bảng 2. So sánh kết quả dự báo SO_2 của các mô hình ANN

Quốc gia	Mô hình	RMSE (ppb)	nRMSE (%)	MAE (ppb)	MAPA (%)
Việt Nam	FFNN [16]	5,8 - 7,1	-	4,4 - 5,7	-
	MLP [17]	5,2 - 8,5	27,8 - 57,6	4,3 - 7,1	45,5 - 80,7
	ESN [17]	4,7 - 9,0	23,4 - 38,3	3,9 - 7,7	65,7 - 81,1
Trung Quốc	SVM [18]	5,2 - 5,86	-	3,4 - 5,0	-
Ba Lan	SVM [11]	-	15,8 - 18,2	0,4 - 0,6	-
	MLP [11]	-	18,6 - 27,0	-	-
Italy	MLP [19]	6,6 - 12,5	-	4,3 - 9,1	-
	MLP [12]	0,3 - 6,9	-	0,3 - 4,9	-

Bảng 3. So sánh kết quả dự báo PM_{10} của các mô hình ANN

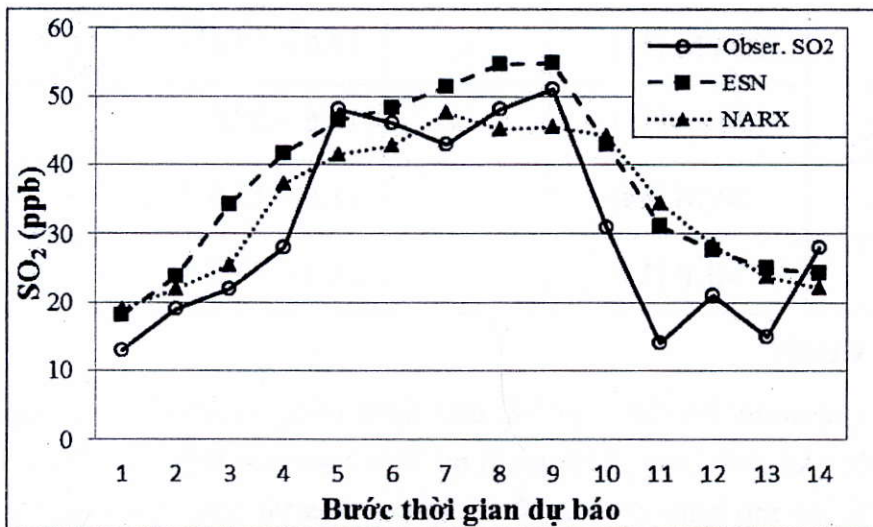
Quốc gia	Mô hình	RMSE ($\mu g/m^3$)	nRMSE (%)	MAE ($\mu g/m^3$)	MAPA (%)
Việt Nam	FFNN [16]	55,1	-	43,4	-
	MLP [17]	-	22,4	68,2	76,4
	ESN [17]	-	16,2	53,5	83,8
Trung Quốc	SVM [13]	-	-	17,7 - 19,0	-
	ANN-RBF [13]	-	-	46,2 - 70,7	-
Ba Lan	SVM [11]	-	13,1 - 21,5	3,2 - 3,4	-
Italy	FFNN [20]	-	-	8,59	-
Chile	FFNN [21]	-	16,0 - 21,0	-	-

Ghi chú: SVM (Support Vector Machine) - một công cụ trí tuệ nhân tạo tương đương với ANN; ANN-RBF (Radial Basis Function Network) - một dạng của ANN

Có thể thấy rõ rằng, mặc dù mô hình FFNN trong nghiên cứu này được xây dựng với mục đích bù dữ liệu quan trắc nồng độ SO_2 và PM_{10} bị khuyết thiếu theo cơ chế dự báo, với số điểm khuyết thiếu dữ liệu tương ứng với số giờ cần dự báo là

khá lớn (lên đến 360 giờ đối với SO_2 và trên 2000 giờ đối với PM_{10}). Thêm vào đó, dữ liệu khuyết thiếu không chỉ có nồng độ của SO_2 và PM_{10} mà còn cả những thông số khí tượng có ảnh hưởng đến diễn biến của chúng trong không khí. Chính vì vậy, mô hình FFNN đã được nghiên cứu với chỉ hai biến đầu vào được trình bày trong Bảng 1. Tuy nhiên, qua các chỉ tiêu thống kê được tính toán cho thấy, kết quả của mô hình FFNN đạt được là rất khả quan, tương đương với một số mô hình cùng loại đã được công bố (Bảng 2 và Bảng 3). Ngoài ra, hàm phân bố mật độ xác suất của bộ dữ liệu về nồng độ SO_2 và PM_{10} , sau khi được bù các giá trị khuyết thiếu bằng mô hình này, đã được kiểm tra và cho thấy hoàn toàn phù hợp với luật phân bố xác suất đã được chỉ ra trong nghiên cứu [22]. Hệ số tương quan giữa phân bố xác suất của bộ dữ liệu đã được bù bằng phương pháp dự báo và bộ dữ liệu nguồn của SO_2 và PM_{10} đều là 1,0. Còn với phân bố mật độ xác suất, hệ số tương quan của SO_2 và PM_{10} lượt là 0,98 và 0,97 [16].

Kết quả nghiên cứu đối với hai mô hình MLP và ESN khi dự báo chất lượng không khí tại Việt Nam với chu kỳ 07 ngày là rất khả qua. Độ chính xác trung bình (MAPA) của kết quả dự báo trong 07 ngày của mô hình MLP là từ 45,5% đến 80,7%, và của ESN là từ 65,7% đến 81,1% đối với SO_2 . Còn đối với PM_{10} , giá trị này của MLP và ESN lần lượt là 77,6% và 83,8%. Kết quả này là tương đương với hầu hết các mô hình cùng loại đã được công bố (Bảng 2 và Bảng 3). Diễn biến nồng độ dự báo cũng khá phù hợp với diễn biến thực tế quan sát được. Thêm nữa, theo [1] nhược điểm lớn nhất của các mô hình ANN thông thường nói chung là ít nhạy cảm với các trường hợp có đột biến lớn về nồng độ của chất cần dự báo (dự báo không tốt). Tuy nhiên, kết quả thử nghiệm trên một số khoảng thời gian có xảy ra biến động mạnh nồng độ của chất cần dự báo cho thấy, cả hai mô hình MLP và ESN đều thích ứng tốt độ chính xác trong trường hợp này đạt trên 70% (Hình 3) [17].



Hình 3. So sánh diễn biến nồng độ SO_2 được dự báo bằng mô hình MLP và ESN với thực tế trong trường hợp có biến động mạnh về nồng độ [17].

Bảng 4 trình bày kết quả của mô hình MLP và ESN đối với thông số NO₂. Đây là một thông số ô nhiễm không khí có diễn biến khá phức tạp do nó vừa là yếu tố có quan hệ mật thiết với nồng độ PM₁₀, vừa là một mắt xích quan trọng trong chu trình ôzôn tầng đối lưu [23]. Chính vì vậy, kết quả dự báo của các mô hình MLP và ESN được thực hiện trong nghiên cứu này thấp hơn so với SO₂ và PM₁₀. Điều này cho thấy, trong nghiên cứu này, việc chỉ mới tính đến ảnh hưởng của các tham số khí tượng có thể là chưa đủ và chưa phản ánh đầy đủ diễn biến thực tế của nồng độ NO₂. Mặc dù vậy, kết quả thu được vẫn cho thấy rằng, khả năng dự báo NO₂ của MLP và ESN là khá tốt. Độ chính xác trung bình (MAPA) của cả hai mô hình đều đạt trên 60% [17]. Mặc dù còn thấp hơn so với các mô hình đã được công bố trên thế giới (Bảng 4), tuy nhiên, kết quả này vẫn khá khả quan khi ứng dụng trong điều kiện tại Việt Nam. Thêm nữa, các mô hình thử nghiệm được xây dựng trên cấu trúc chuẩn, vì vậy, kết quả của các mô hình này có thể được cải thiện khi sử dụng các phương pháp tối ưu hóa cấu trúc mạng, sử dụng các bộ lọc dữ liệu...

Bảng 4. So sánh kết quả dự báo NO₂ bằng các mô hình ANN

Quốc gia	Mô hình	RMSE (ppb)	nRMSE (%)	MAE (ppb)	MAPA (%)
Việt Nam	MLP [17]	-	-	18,0	69,1
	ESN [17]	-	-	14,3	62,2
Trung Quốc	SVM [13]	-	-	6,1 – 7,0	-
	ANN-RBF [13]	-	-	9,4 – 9,6	-
Ba Lan	SVM [11]	-	13,0 – 17,4	0,9 – 1,8	-
	MLP [11]	-	18,6 – 27,0	-	-
Hungary	SVM [14]	-	11,0 – 12,6	-	-
	MLP [14]	-	11,5 – 13,2	-	-

IV. KẾT LUẬN

Kết quả bước đầu đạt được khi ứng dụng công cụ ANN để xây dựng mô hình dự báo thống kê chất lượng không khí tại Việt Nam cho thấy, công cụ này có lợi thế hơn so với các mô hình dự báo số trị, quang hóa và động lực học. Cụ thể là, thứ nhất, ANN có khả năng tự học và khái quát dữ liệu một cách mềm dẻo, chính vì vậy, không đòi hỏi quá cao về cơ sở hạ tầng và chuyên môn của người xây dựng và



vận hành, rất phù hợp với điều kiện thực tế tại Việt Nam. Thứ hai, kết quả thử nghiệm bước đầu đã cho thấy, độ chính xác của kết quả dự báo từ các mô hình ứng dụng ANN là khá cao, hoàn toàn đủ điều kiện để áp dụng vào thực tế. Cuối cùng, ngoài việc sử dụng trong công tác dự báo chất lượng không khí, công cụ ANN còn có thể được sử dụng để bù dữ liệu quan trắc chất lượng không khí bị khuyết thiếu. Điều này có ý nghĩa rất lớn trong công tác chuẩn hóa và khai thác dữ liệu chất lượng không khí. Chính vì vậy, có thể khẳng định công cụ ANN là một hướng tiếp cận đầy triển vọng và khả thi để xây dựng mô hình dự báo chất lượng không khí.

TÀI LIỆU THAM KHẢO

1. Yang Zhang, et al. (2012), *Real-time air quality forecasting, part I: History, techniques, and current status*, Atmospheric Environment **60**, pp. 632–655.
2. Hooyberghs Jef, et al. (2005), *A neural network forecast for daily average PM10 concentrations in Belgium*, Atmospheric Environment **39**(18), pp. 3279-3289.
3. Sabri Ghazi and Mohamed Tarek Khadir (2009), *Recurent Neural Network for Multi-Steps ahead prediction of PM₁₀ concentration*, J.Automation & System Engineering **3**(2), pp. 13-21.
4. Junsu Yi and Victor R. Prybutok (1996), *A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area*, Environmental Pollution **92**(3), pp. 349-357.
5. Wenjian Wang, et al. (2003), *Prediction of maximum daily ozone level using combined neural network and statistical characteristics*, Environment International **29**, pp. 555–562.
6. K. Elamparil and T. Chithambarathanu (2011), *A neural network model for the prediction of afternoon ozone level in a semi-urban tropical site, India*, International Journal of Engineering Science and Technology **3**(7), pp. 5546-5549.
7. K. P. Moustris, et al. (2012), *Application of Multiple Linear Regression Models and Artificial Neural Networks on the Surface Ozone Forecast in the Greater Athens Area, Greece*, Advances in Meteorology, pp. 1-8.
8. Enrico Pisoni, et al. (2009), *Forecasting peak air pollution levels using NARX models*, Engineering Applications of Artificial Intelligence **22**, pp. 593–602.
9. Eros Pasero and Luca Mesin (2010), *Artificial Neural Networks to Forecast Air Pollution*, in *Air Pollution*, Villanyi Vanda, Editor. InTech: Croatia. p. 221-240.
10. Ana Russo, Frank Raischel, and Pedro G. Lind (2013), *Air quality prediction using optimal neural networks with stochastic variables*, Atmospheric Environment **79**, pp. 822-830.
11. Stanislaw Osowski and Konrad Garanty (2007), *Forecasting of the daily meteorological pollution using wavelets and support vector machine*, Engineering Applications of Artificial Intelligence **20**, pp. 745–755.



12. U. Brunellia, et al. (2007), *Two-days ahead prediction of daily maximum concentrations of SO₂, O₃, PM₁₀, NO₂, CO in the urban area of Palermo, Italy*, Atmospheric Environment **41**, pp. 2967–2995.
13. Wei Zhen Lu and Wen Jian Wang (2005), *Potential assessment of the “support vector machine” method in forecasting ambient air pollutant trends*, Chemosphere **59**, pp. 693–701.
14. István Juhos, László Makra, and Balázs Tóth (2008), *Forecasting of traffic origin NO and NO₂ concentrations by Support Vector Machines and neural networks using Principal Component Analysis*, Simulation Modelling Practice and Theory **16**, pp. 1488–1502.
15. Daniel Svozil, Vladimir Kvasniska, and Jiri Pospichal (1997), *Introduction to multi-layer feed-forward neural networks* Chemometrics and Intelligent Laboratory Systems **39**(1), pp. 43-62
16. Mac Duy Hung, Nghiem Trung Dung, and Dinh Thu Hang (2015), *Application of Artificial Neural Network to fill in the missing monitoring data of air quality*, Journal of Science and Technology (VAST) **53**(3A), pp. 199-204.
17. Mac Duy Hung and Nghiem Trung Dung (2015), *Application of Echo State Network for the forecast of air quality*, Journal of Science and Technology (VAST) (đã được chấp nhận đăng).
18. Wenjian Wang, Changqian Men, and Weizhen Lu (2008), *Online prediction model based on support vector machine*, Neurocomputing **71**(4-6), pp. 550–558.
19. Giuseppe Nunnari, et al. (2004), *Modelling SO₂ concentration at a point with statistical approaches*, Environmental Modelling & Software **19**, pp. 887–905.
20. Giorgio Corani (2005), *Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning*, Ecological Modelling **185**(2-4), pp. 513–529.
21. Patricio Perez and Jorge Reyes (2006), *An integrated neural network model for PM₁₀ forecasting*, Atmospheric Environment **40**(16), pp. 2854-2851.
22. Chử Thị Hồng Nhung and NghiêM Trung DũNg (2012), *Xác định định luật phân bố xác suất của dữ liệu chất lượng không khí được quan trắc tại Hà Nội*, Tạp chí Khoa học và Công nghệ **50**(1), pp. 81-87.
23. Noel de Nevers (2000), *Air pollution control engineering*. 2 ed, United States of America: Waveland Press, Inc.