

# ĐÁNH GIÁ ĐỘ TIN CẬY CỦA THIẾT BỊ GIÁM SÁT CHẤT LƯỢNG MÔI TRƯỜNG KHÔNG KHÍ TRONG CÔNG TÁC QUAN TRẮC MÔI TRƯỜNG KHÔNG KHÍ XUNG QUANH

Bùi Thị Hiếu<sup>a,\*</sup>, Nguyễn Thành Trung<sup>a</sup>

<sup>a</sup>*Khoa Kỹ thuật Môi trường, Trường Đại học Xây dựng,  
55 đường Giải Phóng, quận Hai Bà Trưng, Hà Nội, Việt Nam*

**Lịch sử bài viết:**

*Nhận ngày 30/3/2018, Sửa xong 11/5/2018, Chấp nhận đăng 30/5/2018*

---

## Tóm tắt

Vấn đề ô nhiễm môi trường không khí ngày càng trở nên nghiêm trọng kéo theo yêu cầu chú trọng các chương trình quan trắc môi trường không khí. Do đó đề tài B2014-03-12 đã nghiên cứu và chế tạo thiết bị quan trắc môi trường không khí. Để đánh giá khả năng ứng dụng của thiết bị này, bài báo đánh giá độ chính xác của thiết bị đề tài bằng cách so sánh bộ số liệu các thông số về chất ô nhiễm môi trường không khí cung cấp bởi thiết bị đề tài và thiết bị kiểm chứng sử dụng các công cụ thông số thống kê và các bài kiểm định xác suất thống kê. Kết quả thu được chứng tỏ khả năng quan trắc các chất ô nhiễm bao gồm PM<sub>10</sub>, SO<sub>2</sub>, CO, NO<sub>2</sub> của thiết bị giám sát chất lượng môi trường không khí này.

*Từ khóa:* ô nhiễm môi trường không khí; nồng độ các chất độc hại; thiết bị quan trắc môi trường không khí; thống số thống kê; kiểm định xác suất thống kê.

EVALUATION THE ABILITY OF AIR QUALITY MONITORING EQUIPMENT FOR AIR QUALITY MONITORING

## Abstract

Serious air pollution requires scrupulous attention to air pollution substance monitoring programs. Therefore, the project B2014-03-12 carried out research on manufacturing air pollution monitoring equipment to record continuous toxic substances in the air environment. However, the accuracy of this equipment needs to be evaluated and verified. Therefore, this study evaluated the accuracy of this equipment by assessing the agreement between the test and reference equipment by statistical approaches. The obtained results demonstrate the ability of the air pollution monitoring equipment to measure continuously the toxic substances in the air environment including PM<sub>10</sub>, SO<sub>2</sub>, CO, NO<sub>2</sub>.

*Keywords:* air pollution; concentration of air toxic substances; air monitoring equipment; statistical approaches; statistic test.

[https://doi.org/10.31814/stce.nuce2018-12\(4\)-12](https://doi.org/10.31814/stce.nuce2018-12(4)-12) © 2018 Trường Đại học Xây dựng (NUCE)

---

## 1. Đặt vấn đề

Vấn đề ô nhiễm môi trường không khí do lượng khí thải từ các hoạt động giao thông, công nghiệp, sinh hoạt ngày càng trở nên nghiêm trọng và ảnh hưởng đến sức khỏe của dân cư đô thị, đặc biệt là đối với một số nhóm người dễ bị tổn thương như người già, trẻ em, người mắc bệnh tật ... Vì vậy, các

---

\*Tác giả chính. Địa chỉ e-mail: [buihieudhxd@gmail.com](mailto:buihieudhxd@gmail.com) (Hiếu, B. T.)

chương trình quan trắc xác định nồng độ các chất độc hại trong môi trường không khí cần được đầu tư và chú trọng. Hoạt động quan trắc môi trường không khí tại Việt Nam được thực hiện chủ yếu bằng phương pháp thủ công truyền thống với tần suất quan trắc 3-6 lần một năm. Tuy nhiên, quá trình lấy mẫu tại hiện trường và phân tích mẫu tại phòng thí nghiệm đòi hỏi phải đầu tư nhiều thời gian, nhân lực, vật tư, hóa chất và thiết bị kèm theo. Thêm vào đó, một trong những hạn chế của phương pháp quan trắc truyền thống là thời gian bảo quản mẫu ngắn. Do đó, số lượng mẫu đại diện trong ngày theo mục đích quan trắc khó có thể thu thập được. Vì vậy, số liệu không phản ánh thời gian thực đo do thông số môi trường biến đổi theo thời gian gây khó khăn cho việc thực hiện các biện pháp quản lý môi trường không khí.

Bên cạnh phương pháp thủ công truyền thống, công tác quan trắc chất lượng môi trường không khí trong khu đô thị hiện nay còn được thực hiện bằng cách sử dụng các trạm quan trắc tự động liên tục. Tuy nhiên, mạng lưới các trạm quan trắc tự động tại nước ta có mật độ phân bố thưa, chưa có tính đại diện cho tất cả các khu vực chịu ảnh hưởng nặng nề do ô nhiễm môi trường không khí gây ra. Chi phí đầu tư ban đầu cũng như chi phí duy tu, vận hành của trạm quan trắc tự động hàng năm tương đối lớn. Do đó, một số trạm quan trắc tự động chất lượng môi trường không khí đã ngừng hoạt động do thiếu kinh phí duy trì, vận hành và hết thời gian khấu hao. Gần đây, đề tài B2014-03-12 đã nghiên cứu và chế tạo thiết bị quan trắc môi trường không khí với các thông số  $PM_{10}$ ,  $SO_2$ ,  $CO$ ,  $NO_2$  được cung cấp liên tục [1, 2]. Ưu điểm của thiết bị này so với trạm quan trắc tự động là có khả năng cung cấp được chuỗi số liệu liên tục trong ngày quan trắc với chi phí đầu tư, vận hành thấp hơn, thiết bị nhỏ gọn hơn so với hệ thống trạm quan trắc tự động. Tuy nhiên, độ chính xác của thiết bị này cần phải được đánh giá và kiểm định để đánh giá khả năng ứng dụng của thiết bị trong công tác quan trắc chất lượng môi trường không khí xung quanh.

Gần đây, trong nghiên cứu năm 2017 [3], tác giả Nguyễn Quốc Hoàn cũng thực hiện đánh giá độ chính xác của thiết bị quan trắc chất lượng môi trường không khí của đề tài B2014-03-12 bằng các phép so sánh sử dụng một số đại lượng thống kê của phương trình tuyến tính bình phương nhỏ nhất. Tuy nhiên, phương trình tuyến tính bình phương nhỏ nhất chỉ giả thiết thiết bị đề tài có sai số và mặc định thiết bị kiểm chứng (TBKC) không có sai số. Vì trong thực tế số liệu đo đạc của thiết bị kiểm chứng và thiết bị đề tài đều có sai số nên phương pháp hồi quy tuyến tính bình phương nhỏ nhất để đánh giá sự tương đồng giữa hai thiết bị là giải pháp để khắc phục nhược điểm của phương pháp hồi quy tuyến tính bình phương nhỏ nhất [4]. Do đó, trong nghiên cứu này, độ chính xác và độ tin cậy của thiết bị quan trắc chất lượng môi trường không khí này sẽ được đánh giá một cách có hệ thống bằng những phương pháp so sánh hai bộ số liệu sử dụng các bài kiểm định xác suất thống kê, các đại lượng thống kê bao gồm đại lượng thống kê của phương trình tuyến tính bình phương nhỏ nhất (phương pháp tuyến tính này xem xét sai số của cả thiết bị đề tài và thiết bị kiểm chứng).

## **2. Phương pháp đánh giá độ chính xác của thiết bị giám sát chất lượng môi trường không khí**

### *2.1. Phương pháp đánh giá độ chính xác của thiết bị giám sát chất lượng môi trường không khí sử dụng các công cụ thông số thống kê*

#### **a. Phương trình hồi quy tuyến tính**

Hồi quy tuyến tính [5] là một trong những phương pháp phổ biến nhất được sử dụng trong nghiên cứu so sánh [4]. Dựa vào phương pháp này, phương trình tuyến tính giữa số liệu đo đạc từ hai thiết bị sẽ được thiết lập. Nếu hai thiết bị tương quan với nhau, đường thẳng tuyến tính này sẽ không sai khác đáng kể so với đường thẳng 1:1 nghiêng 45 độ so với trục hoành. Đường thẳng 1:1 này được xác định bởi hai thông số: độ dốc bằng 1 và hệ số chặn bằng 0. Sự chênh lệch giữa đường thẳng tuyến tính và

đường thẳng 1:1 biểu thị sự thiếu tương quan giữa hai phương pháp. Do sai số luôn hiện diện trong các bài kiểm tra về tuyến tính, các thông số cung cấp bài kiểm tra cho mức độ tương quan của hai thiết bị. Trong các phương pháp hồi quy tuyến tính, phương pháp bình phương nhỏ nhất là một trong những phương pháp được sử dụng đầu tiên để thiết lập hệ số phương trình hồi quy tuyến tính [6]. Tuy nhiên, phương pháp hồi quy tuyến tính bình phương nhỏ nhất giả thiết rằng thiết bị kiểm chứng cho số liệu đo đạc không có sai số. Điều này rất khó xảy ra trong thực tế vì thiết bị nào trong thực tế cũng cho số liệu đo đạc với sai số nhất định [6]. Do đó, phương pháp hồi quy tuyến tính vuông góc bình phương nhỏ nhất được ứng dụng thay thế cho phương pháp hồi quy tuyến tính bình phương nhỏ nhất truyền thống. Phương pháp này thiết lập phương trình đường hồi quy tuyến tính bằng cách xác định hệ số độ dốc và hệ số chặn sao cho tổng bình phương từ điểm đo đạc đến đường thẳng tuyến tính theo phương vuông góc với đường thẳng tuyến tính là nhỏ nhất. Bằng cách này, phương pháp hồi quy tuyến tính vuông góc bình phương nhỏ nhất [7] có bao gồm sai số của cả thiết bị đề tài và thiết bị kiểm chứng. Đường hồi quy thể hiện sự tương đồng giữa kết quả đo của thiết bị đề tài với thiết bị kiểm chứng thông qua hệ số độ dốc và hệ số chặn. Bên cạnh hai hệ số này, hệ số tương quan cũng được sử dụng kết hợp để đánh giá độ tin cậy của thiết bị đề tài. Sự tương quan của hai thiết bị được diễn giải chỉ bằng hệ số tương quan là chưa đầy đủ và hợp lý vì giá trị hệ số tương quan cao không đồng nghĩa với việc hai thiết bị tương quan với nhau. Hai thiết bị cho kết quả đo đạc tương đồng với nhau khi các điểm trên đồ thị phân tán nằm xung quanh đường thẳng 1:1, tuy nhiên giá trị hệ số tương quan cao có thể đạt được khi các điểm phân tán nằm phân bố xung quanh bất kỳ đường thẳng nào. Do đó, trong nghiên cứu này, bên cạnh hệ số tương quan Pearson [8], tác giả sử dụng thêm hệ số tương quan phù hợp (Công thức (1)) để đánh giá thiết bị đề tài.

b. Hệ số tương quan phù hợp

Hệ số tương quan phù hợp [9] biểu thị sự liên quan chặt chẽ giữa số liệu đo đạc phân bố xung quanh đường thẳng 1:1 và được tính bằng công thức:

$$r_c = \frac{2S_{TR}}{S_T^2 + S_R^2 + (\bar{T}_i - \bar{R}_i)^2} \quad (1)$$

trong đó  $(\bar{R}_i)$  là nồng độ trung bình của thiết bị kiểm chứng;  $(\bar{T}_i)$  là nồng độ trung bình của thiết bị đề tài;  $S_{TR}$  là hiệp phương sai của hai biến thông số đo đạc từ thiết bị đề tài và thiết bị kiểm chứng;  $S_T^2$ ,  $S_R^2$  là phương sai của hai biến thông số đo đạc từ thiết bị đề tài và thiết bị kiểm chứng.

c. Độ sai phương chuẩn

Bên cạnh các hệ số của phương trình hồi quy tuyến tính, hệ số tương quan và hệ số tương quan phù hợp, để đánh giá độ tin cậy của thiết bị đề tài được chặt chẽ và kỹ lưỡng hơn, độ sai phương chuẩn cũng được kết hợp vào công cụ đánh giá. Công thức (2) xác định độ sai phương chuẩn:

$$NMSE = \overline{(R_i - T_i)^2} / [\bar{R}_i \bar{T}_i] \quad (2)$$

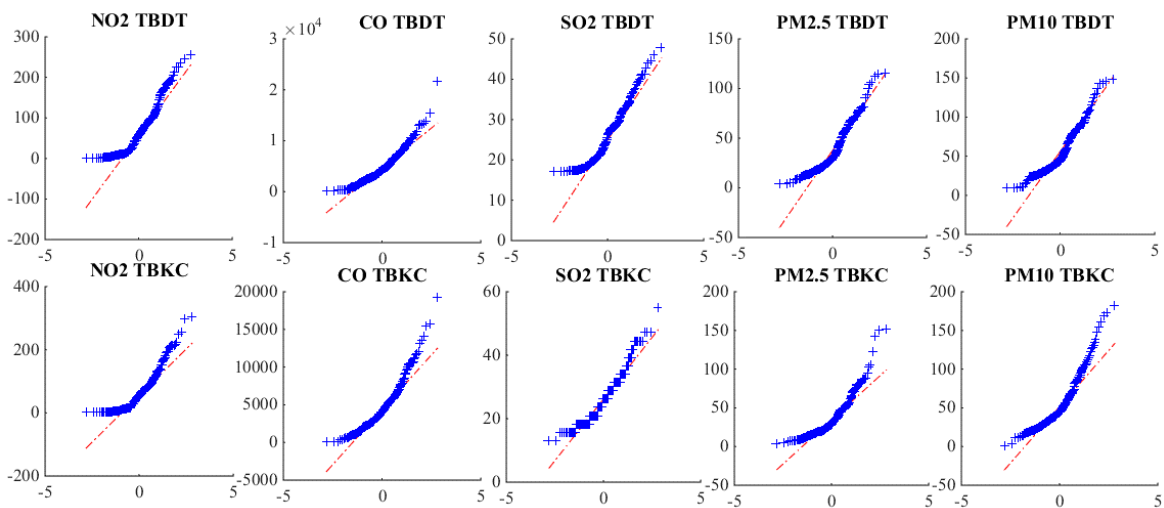
trong đó  $\bar{R}_i$  là nồng độ trung bình của thiết bị kiểm chứng;  $\bar{T}_i$  là nồng độ trung bình của thiết bị đề tài;  $R_i$  là nồng độ đo được tại thời điểm  $i$  bởi thiết bị kiểm chứng;  $T_i$  là nồng độ đo được tại thời điểm  $i$  bởi thiết bị đề tài.

Trong nghiên cứu này, tiêu chuẩn hướng dẫn về so sánh, đánh giá thống kê cho các cặp thông số của Hiệp Hội Thí nghiệm và Vật liệu Hoa Kỳ ASTM-D5157 [10] được sử dụng để đánh giá mối quan hệ giữa hai bộ số liệu từ hai thiết bị. Hai bộ số liệu tương quan chặt chẽ với nhau khi: Hệ số tương quan lớn hơn hoặc bằng 0,9. Trong bài báo này, tác giả sử dụng cả hệ số tương quan cả hệ số tương

quan phù hợp với giá trị giới hạn dưới của hai hệ số này là 0,9. Hệ số độ dốc nằm trong khoảng từ 0,75 đến 1,25. Hệ số chặn nhỏ hơn hoặc bằng 25 phần trăm giá trị đo trung bình từ thiết bị kiểm chứng. Độ sai phương chuẩn NMSE nhỏ hơn hoặc bằng 0,25.

2.2. Phương pháp đánh giá độ chính xác của thiết bị quan trắc chất lượng môi trường không khí sử dụng các bài kiểm định xác suất thống kê

Trong nghiên cứu này, bài kiểm định t-test [11] được sử dụng để kiểm tra sự tương đồng của bộ số liệu đo đạc từ thiết bị đề tài và bộ số liệu đo đạc từ thiết bị kiểm chứng thông qua đánh giá sự khác biệt giá trị trung bình của số liệu đo đạc từ hai thiết bị. Các bước ứng dụng t-test trong việc so sánh hai bộ số liệu được thể hiện ở Hình 1. Kiểm định t-test trung bình hai mẫu độc lập phải thỏa mãn điều kiện biến phụ thuộc phải có phân phối chuẩn và phương sai của biến phụ thuộc phải có tính đồng nhất (homogeneity of variance). Do đó trước khi làm bài kiểm định t-test, tác giả sẽ kiểm tra số liệu từ hai thiết bị có phân phối chuẩn hay không thông qua bài kiểm tra phân phối chuẩn Anderson–Darling và bài kiểm tra so sánh phương sai của hai bộ số liệu.



Hình 1. Đồ thị Q-Q của các số liệu thông số chất ô nhiễm môi trường tại điểm đo 1

*Bài kiểm định phân phối chuẩn của bộ số liệu Anderson-Darling test:* Giả thiết của bài kiểm định phân phối chuẩn Anderson-Darling test như sau:  $H_0$ : Bộ số liệu có phân phối chuẩn.  $H_1$ : Bộ số liệu không có phân phối chuẩn. Nếu giá trị  $p$  lớn hơn 0,05, giả thiết  $H_0$  được chấp nhận ở khoảng tin cậy 95%. Chúng ta có thể kết luận rằng không có đủ bằng chứng để kết luận bộ số liệu không có phân phối chuẩn. Bên cạnh bài kiểm định phân phối chuẩn bằng bài kiểm định Anderson-Darling test, bài báo sử dụng đồ thị xác suất chuẩn Q-Q để kiểm tra phân phối chuẩn của các số liệu quan trắc các chất ô nhiễm môi trường không khí. Biểu đồ xác suất chuẩn bao gồm trục hoành là giá trị quan sát, trục tung là giá trị mong đợi. Số liệu có phân phối chuẩn khi các điểm trị số quan sát và trị số mong đợi đều nằm gần trên đường thẳng. Trong trường hợp ngược lại, bộ số liệu không có phân phối chuẩn, trước khi thực hiện t-test, tác giả sẽ chuyển đổi bộ số liệu sang phân phối chuẩn sử dụng phương pháp chuyển đổi Box-Cox.

*Bài kiểm định so sánh phương sai của hai bộ số liệu:* Giả thiết kiểm tra so sánh phương sai của hai bộ số liệu như sau:  $H_0$ : Hai bộ số liệu có tỉ số phương sai bằng 1.  $H_1$ : Hai bộ số liệu có tỉ số phương sai khác 1. Nếu số liệu có phân phối chuẩn, F-test được sử dụng để so sánh phương sai của

hai bộ số liệu. Trong trường hợp số liệu không có phân phối chuẩn, Levene-test được sử dụng để so sánh phương sai của hai bộ số liệu. Nếu giá trị  $p$  lớn hơn 0,05 thì giả thiết  $H_0$  được chấp nhận và có thể kết luận không có sự sai khác đáng kể về phương sai của hai bộ số liệu.

*Bài kiểm định t-test:* Bước đầu tiên của bài kiểm định xác suất thống kê t-test là đưa ra giả thiết  $H_0$  và giả thiết thay thế. Trong nghiên cứu này, giả thiết  $H_0$  và giả thiết thay thế  $H_1$  lần lượt là:  $H_0$ : giá trị trung bình của số liệu đo đạc từ thiết bị đề tài và thiết bị kiểm chứng là như nhau.  $H_1$ : giá trị trung bình của số liệu đo đạc từ thiết bị đề tài và thiết bị kiểm chứng là khác nhau. Hai thông số cơ bản của bài kiểm định t-test là thông số giá trị  $t$  và giá trị  $p$ . Giả thiết  $H_0$  có thể bị loại bỏ ở khoảng tin cậy 95% nếu giá trị  $p$  nhỏ hơn 0,05 và kết luận rằng có sự khác nhau về ý nghĩa thống kê giữa giá trị trung bình giữa hai bộ số liệu. Hay nói một cách khác, hai bộ số liệu không tương đồng với nhau nếu giá trị  $p$  nhỏ hơn 0,05. Cần lưu ý rằng kiểm định trung bình t-test không phải là cách tốt nhất để dùng trong so sánh tương đồng bởi vì giá trị trung bình của hai bộ số liệu có thể không có sự sai khác đáng kể về mặt xác suất thống kê, nhưng hai bộ số liệu có thể có sự sai khi ta xem xét ở những khoảng xác định. Tuy nhiên, nghiên cứu này sử dụng bài kiểm định giá trị trung bình t-test để làm một trong những tiêu chí đánh giá sự tương quan giữa hai bộ số liệu để kết quả đánh giá độ tin cậy của thiết bị đề tài B2014-03-12 có độ tin cậy cao hơn.

### 3. Kết quả

#### 3.1. Kết quả của phương pháp so sánh sử dụng các công cụ thông số thống kê

Thiết bị giám sát chất lượng môi trường không khí sau khi được lắp đặt và kiểm nghiệm được sử dụng để đo đồng thời và liên tục các thông số ô nhiễm môi trường không khí xung quanh cùng với thiết bị kiểm chứng. Tại địa điểm quan trắc 1 nằm trên đường Nguyễn Văn Cừ, thiết bị đề tài được đặt cùng vị trí với trạm quan trắc môi trường không khí tự động tại số 556 Nguyễn Văn Cừ để đo đạc các thông số  $SO_2$ ,  $NO_2$ , CO và bụi  $PM_{10}$  từ 19/10/2016 đến 22/10/2016 và từ 11/11/2016 đến 14/11/2016. Tại điểm đo đạc số 1, các số liệu đo đạc tháng 10 và tháng 11 được kết hợp với nhau để sử dụng trong các phương pháp đánh giá độ chính xác của thiết bị. Tại địa điểm 2 (cổng Trường Đại học Xây dựng (ĐHXD)), thiết bị kiểm chứng được sử dụng là thiết bị IR-208 cung cấp các thông số  $SO_2$ ,  $NO_2$ , CO và bụi  $PM_{10}$ . Chuỗi số liệu đo đạc tại cổng Trường Đại học Xây dựng từ ngày 13/9/2016 đến ngày 17/9/2016. Bảng 1 và Bảng 2 tổng kết các thông số thống kê để làm cơ sở so sánh bộ số liệu đo đạc nồng độ các chất ô nhiễm môi trường không khí bao gồm các khí  $SO_2$ ,  $NO_2$ , CO và bụi hạt ở hai địa điểm đo đạc 1 và 2.

Bảng 1 cho thấy các đại lượng thống kê bao gồm hệ số độ dốc, hệ số chặn, hệ số tương quan, hệ số tương quan phù hợp, độ sai phương trung bình chuẩn đều có giá trị thỏa mãn điều kiện hai bộ số liệu tương quan chặt chẽ theo tiêu chuẩn ASTM-5157 của Hoa Kỳ. Nói một cách khác, thiết bị đề tài và thiết bị kiểm chứng cung cấp hai bộ số liệu quan trắc các thông số môi trường không khí xung quanh bao gồm các khí  $SO_2$ ,  $NO_2$ , CO và bụi hạt có mức tương đồng cao. Tuy nhiên, trước đây tại điểm đo 1, nghiên cứu [1] có kết luận rằng trong các thông số về các chất ô nhiễm xem xét, chỉ có thông số  $SO_2$  có hai đại lượng không đạt tiêu chuẩn về bộ số liệu tương đồng (đại lượng hệ số tương quan  $r = 0,82$ ,  $r = 0,8$  và hệ số độ dốc  $a = 0,6$ ,  $a = 0,64$  lần lượt cho tháng 10 và tháng 11). Có thể thấy một vài nguyên nhân dẫn đến sự khác biệt trong kết quả so sánh này: (1) Phương trình tuyến tính mà nghiên cứu [3] sử dụng là phương trình tuyến tính bình phương nhỏ nhất chỉ giả thiết đến sai số của thiết bị đề tài mà bỏ qua sai số của thiết bị kiểm chứng. (2) Nghiên cứu [3] phân tách bộ số liệu quan trắc thành hai tháng riêng biệt để đánh giá mức độ tương đồng của bộ số liệu dẫn đến số lượng mẫu số liệu quan trắc của nghiên cứu [3] bé hơn số lượng mẫu nghiên cứu này xem xét.

Bảng 1. So sánh, đánh giá sự đồng nhất giá trị đo giữa thiết bị đề tài với thiết bị kiểm chứng tại trạm quan trắc tự động tại điểm đo 1

Đại lượng	Hệ số độ dốc	Hệ số chặn	Hệ số tương quan	Hệ số tương quan phù hợp	Độ sai phương chuẩn
Tiêu chí đánh giá	0,75–1,25	$\leq 25\%TBKC$	$\geq 0,9$	$\geq 0,9$	$\leq 0,25$
NO <sub>2</sub>	0,94	3,32	0,96	0,96	0,08
CO	1,00	181,98	0,96	0,96	0,09
SO <sub>2</sub>	0,85	3,57	0,93	0,92	0,01
PM <sub>2.5</sub>	0,88	4,38	0,93	0,92	0,07
PM <sub>10</sub>	0,88	6,92	0,92	0,91	0,06

Bảng 2. So sánh, đánh giá sự đồng nhất giá trị đo giữa thiết bị đề tài với thiết bị kiểm chứng tại điểm đo 2

Đại lượng	Hệ số độ dốc	Hệ số chặn	Hệ số tương quan	Hệ số tương quan phù hợp	Độ sai phương chuẩn
Tiêu chí đánh giá	0,75-1,25	$\leq 25\%TBKC$	$\geq 0,9$	$\geq 0,9$	$\leq 0,25$
NO <sub>2</sub>	0,98	1,01	0,93	0,93	0,06
CO	0,78	559,53	0,88	0,9	0,1
SO <sub>2</sub>	1,47	-17,37	0,72	0,75	0,01
PM <sub>10</sub>	0,99	1,12	0,9	0,9	0,03

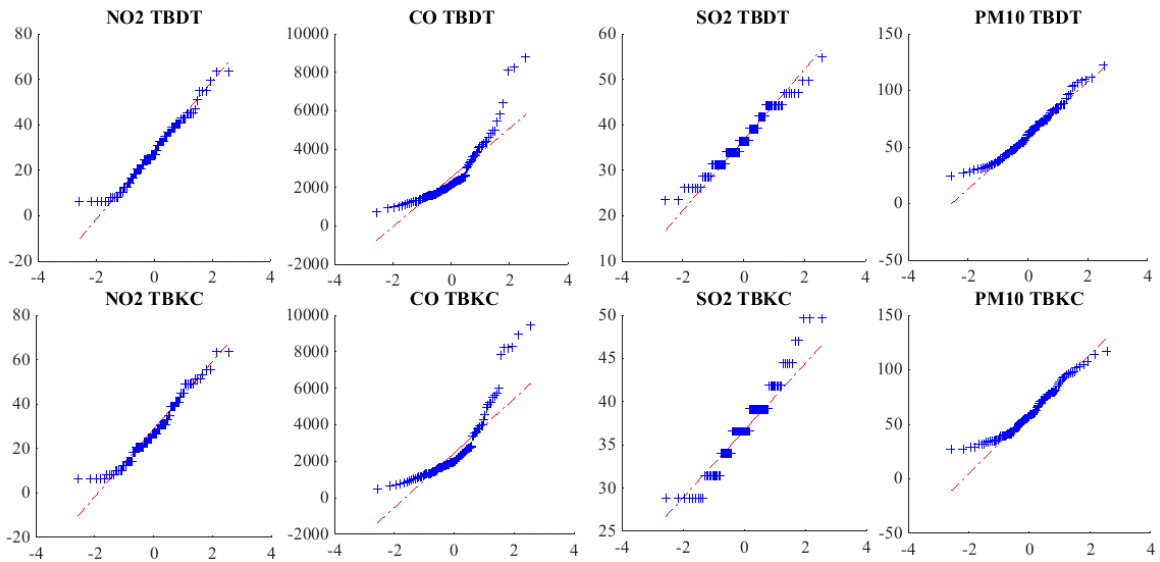
Bảng 2 cho thấy các đại lượng thống kê đối với các thông số NO<sub>2</sub>, CO và bụi PM<sub>10</sub> đều nằm trong giới hạn thỏa mãn hai bộ số liệu từ thiết bị đề tài và thiết bị kiểm chứng tương đồng nhau theo tiêu chuẩn ASTM-5257 của Hoa Kỳ. Tuy nhiên, đối với thông số SO<sub>2</sub>, hai đại lượng bao gồm hệ số độ dốc  $a = 1,47$  và hệ số tương quan  $r = 0,72$  là nằm ngoài khoảng thỏa mãn hai bộ số liệu tương đồng. Kết quả này tương tự như kết quả của nghiên cứu [3] (tại điểm đo 2, đối với thông số SO<sub>2</sub> đại lượng hệ số độ dốc  $a = 0,74$  và hệ số tương quan  $r = 0,64$ ). Tuy nhiên, để kết quả so sánh chính xác hơn thì chuỗi thời gian đo đạc các thông số ô nhiễm môi trường không khí cần được kéo dài hơn để cung cấp một bộ số liệu ổn định và nhiều thông tin hơn.

### 3.2. Kết quả của phương pháp so sánh sử dụng các bài kiểm định xác suất thống kê

Bảng 3 và Bảng 4 tổng kết các giá trị  $p$  của các bài kiểm định phân phối chuẩn của bộ số liệu quan trắc môi trường không khí xung quanh Anderson Darling test, bài kiểm định về so sánh sai phương bộ số liệu cung cấp từ thiết bị đề tài và thiết bị kiểm chứng F-test hoặc Levene-test và bài kiểm định t-test ở hai địa điểm đo đạc. Kết quả bài kiểm định về phân phối chuẩn của các số liệu quan trắc chất lượng môi trường không khí chỉ ra rằng tất cả các thông số quan trắc đều không theo phân phối chuẩn với giá trị  $p$  của tất cả các thông số quan trắc đều nhỏ hơn 0,05 tại điểm đo 1. Thêm vào đó, các đồ thị Q-Q của các thông số quan trắc đều không bám sát đường thẳng như Hình 1. Kết quả này góp phần khẳng định phân phối của các thông số quan trắc tại điểm đo 1 không phải là phân phối chuẩn. Do đó, trước khi thực hiện kiểm định t-test để kiểm tra sự sai khác giữa bộ số liệu từ thiết bị đề tài và thiết bị



kiểm chứng, các bộ số liệu được chuyển sang bộ số liệu phân phối chuẩn bằng phương pháp Box-Cox. Tương tự, hầu hết các số liệu đo đạc các thông số quan trắc môi trường của thiết bị đề tài và thiết bị kiểm chứng đều có giá trị  $p < 0,05$  (ngoại trừ số liệu  $\text{NO}_2$  của thiết bị đề tài). Do giá trị  $p < 0,05$  nên giả thiết rằng bộ số liệu có phân phối chuẩn bị loại bỏ. Bên cạnh giá trị  $p$ , ta có thể thấy các điểm đo đạc không bám sát đường thẳng trên đồ thị Q-Q như Hình 2. Do đó, để so sánh bộ số liệu từ thiết bị đề tài và thiết bị kiểm chứng sử dụng bài kiểm định t-test, tác giả chuyển đổi bộ số liệu quan trắc ban đầu sang bộ số liệu quan trắc có phân phối chuẩn sử dụng phương pháp Box-Cox.



Hình 2. Đồ thị Q-Q của các số liệu thông số chất ô nhiễm môi trường tại điểm đo 2

Bảng 3. Giá trị  $p$  của các bài kiểm định thống kê của bộ số liệu quan trắc chất lượng môi trường không khí từ hai thiết bị tại điểm đo 1

Đại lượng	Kiểm định Anderson–Darling test		Kiểm định so sánh phương sai	Kiểm định t-test
	Thiết bị đề tài	Thiết bị kiểm chứng		
$\text{NO}_2$	$< 0,005$	$< 0,005$	0,83	0,93
CO	$< 0,005$	$< 0,005$	0,82	0,76
$\text{SO}_2$	$< 0,005$	$< 0,005$	0,17	0,82
$\text{PM}_{2,5}$	$< 0,005$	$< 0,005$	0,58	0,64
$\text{PM}_{10}$	$< 0,005$	$< 0,005$	0,45	0,27

Trước khi thực hiện bài kiểm định t-test, phương sai của các bộ số liệu được so sánh với nhau sử dụng bài kiểm tra phương sai phi tham số Levene-test. Tại điểm đo 1, giá trị  $p$  của tất cả các thông số xem xét đều lớn hơn 0,05. Do đó, ở khoảng tin cậy 95%, giả thiết tỉ số giá trị phương sai của hai bộ số liệu bằng 1 được chấp nhận. Do đó, tất cả các bộ số liệu quan trắc môi trường tại điểm đo 1 có giá trị phương sai tương đồng nhau. Vì thế, chúng ta có thể áp dụng bài kiểm định t-test cho các bộ số liệu này. Từ kết quả tóm tắt tại Bảng 3, ta thấy giá trị  $p$  của tất cả các cặp thông số tại điểm đo tại điểm đo 1 đều lớn hơn 0,05. Từ đó chúng ta chấp nhận giả thiết giá trị trung bình của hai bộ số liệu quan trắc

Bảng 4. Giá trị  $p$  của các bài kiểm định thống kê của bộ số liệu quan trắc chất lượng môi trường không khí từ hai thiết bị tại điểm đo 2

Đại lượng	Kiểm định Anderson–Darling test		Kiểm định so sánh phương sai	Kiểm định t-test
	Thiết bị đề tài	Thiết bị kiểm chứng		
NO <sub>2</sub>	0,147	0,008	0,97	0,88
CO	< 0,005	< 0,005	0,25	0,31
SO <sub>2</sub>	0,008	< 0,005	0,002	-
PM <sub>10</sub>	0,045	< 0,005	0,98	0,8

môi trường là như nhau ở khoảng tin cậy 95%. Kết quả này là một trong những cơ sở để kết luận giá trị đo thiết bị đề tài và giá trị đo của thiết bị kiểm chứng có mức tương đồng cao.

Tại điểm đo 2, bài kiểm định so sánh phương sai phi tuyến tính Levene-test cho các bộ số liệu của các thông số NO<sub>2</sub>, CO và PM<sub>10</sub> đều cho kết quả giá trị  $p$  lớn hơn 0,05. Do đó, giả thiết tỉ số phương sai giữa hai cặp số liệu bằng 1 không thể bị loại bỏ. Từ đó, chúng ta có thể áp dụng bài kiểm định t-test cho bộ số liệu các thông số NO<sub>2</sub>, CO và PM<sub>10</sub>. Bài kiểm định t-test cho bộ số liệu các thông số NO<sub>2</sub>, CO và PM<sub>10</sub> cho kết quả giá trị  $p$  lớn hơn 0,05. Do đó, ở khoảng tin cậy 95% chúng ta có thể kết luận giá trị trung bình của các bộ số liệu các thông số NO<sub>2</sub>, CO và PM<sub>10</sub> không có sự sai khác đáng kể. Đây cũng là một trong những cơ sở để kết luận số liệu quan trắc các thông số NO<sub>2</sub>, CO và PM<sub>10</sub> từ thiết bị đề tài và thiết bị kiểm chứng có mối quan hệ chặt chẽ. Tuy nhiên, đối với thông số SO<sub>2</sub> tại điểm đo 2, bài kiểm định so sánh phương sai phi tham số Levene test cho giá trị  $p$  bằng 0,002 nhỏ hơn 0,05 nên tác giả không thực hiện bài kiểm định t-test cho bộ số liệu này.

#### 4. Kết luận và kiến nghị

Tác giả đã sử dụng các công cụ thông số thống kê và sử dụng các bài kiểm định xác suất thống kê để đánh giá độ tin cậy của thiết bị [2] để đo các thông số về chất ô nhiễm môi trường không khí bao gồm SO<sub>2</sub>, NO<sub>2</sub>, CO và PM<sub>10</sub> và đưa ra các nhận xét:

- Tại điểm đo 1, các đại lượng thống kê của tất cả các bộ số liệu thông số chất ô nhiễm môi trường không khí cung cấp bởi thiết bị đề tài và thiết bị kiểm chứng đều nằm trong khoảng tiêu chuẩn hai bộ số liệu tương đồng với nhau theo tiêu chuẩn ASTM-5157. Thêm vào đó, giá trị  $p$  của bài kiểm định t-test của tất cả cặp số liệu của các thông số chất ô nhiễm đều lớn hơn 0,05. Do đó, các giá trị đo từ thiết bị đề tài có mối quan hệ chặt chẽ và tương quan cao với các giá trị đo của thiết bị kiểm chứng là trạm đo không khí tự động trên đường Nguyễn Văn Cừ.

- Tại điểm đo 2, các cặp số liệu đo đạc từ hai thiết bị của các thông số chất ô nhiễm NO<sub>2</sub>, CO và PM<sub>10</sub> đều có các đại lượng thống kê thỏa mãn điều kiện tương quan theo tiêu chuẩn ASTM-5157; bài kiểm định t-test của các cặp thông số này đều cho giá trị  $p$  lớn hơn 0,05. Tuy nhiên, cặp số liệu cung cấp bởi hai thiết bị cho thông số SO<sub>2</sub> cho kết quả hai đại lượng là hệ số tương quan, hệ số độ dốc nằm ngoài ngưỡng thỏa mãn điều kiện tương quan. Ngoài ra, giá trị  $p$  của bài kiểm định so sánh phương sai phi tham số Levene test cho thông số SO<sub>2</sub> nhỏ hơn 0,05; vì vậy, bài kiểm định t-test không thể sử dụng để đánh giá mối liên hệ của cặp số liệu SO<sub>2</sub>.

- Kết quả so sánh các cặp thông số các chất ô nhiễm môi trường không khí tại hai địa điểm đo đạc trên địa bàn thành phố Hà Nội cung cấp bởi thiết bị đề tài và thiết bị kiểm chứng cho thấy có thể sử dụng thiết bị giám sát chất lượng môi trường không khí [1, 2] nghiên cứu chế tạo để đo đạc và quan



trắc các thông số chất ô nhiễm môi trường  $\text{SO}_2$ ,  $\text{NO}_2$ ,  $\text{CO}$ ,  $\text{PM}_{10}$  và  $\text{PM}_{2.5}$  phục vụ cho công tác đánh giá hiện trạng, dự báo ô nhiễm môi trường không khí.

- Tuy nhiên, nếu thời gian đo đạc các thông số  $\text{SO}_2$ ,  $\text{NO}_2$ ,  $\text{CO}$ ,  $\text{PM}_{10}$  và  $\text{PM}_{2.5}$  dài hơn thì bộ liệu được ổn định và kết quả các đại lượng thống kê và các bài kiểm định thống kê sẽ có độ tin cậy cao hơn. Thêm nữa, độ bền của thiết bị của đề tài cũng sẽ được kiểm định bằng cách kéo dài thời gian đo đạc.

### Tài liệu tham khảo

- [1] Trung, N. T., Tới, P. V., Lan, Đ. T. P. (2016). Kết quả bước đầu trong nghiên cứu, thiết kế thiết bị quan trắc không khí khu vực đô thị. *Tạp chí Môi trường Đô thị Việt Nam*, 6(107):31–35.
- [2] Trung, N. T. (2014). *Nghiên cứu thiết kế, chế tạo hệ thống thiết bị giám sát chất lượng môi trường không khí ở khu vực dân cư đô thị, mã số B2014-03-12*. Bộ Giáo dục & Đào tạo.
- [3] Hoàn, N. Q. (2017). *Nghiên cứu ứng dụng thiết bị đo nhanh để quan trắc chất lượng không khí tại Hà Nội*. Luận văn Thạc sĩ, Trường Đại học Xây dựng, Hà Nội.
- [4] Yan, X., Su, X. G. (2009). *Linear regression analysis: Theory and computing*. World Scientific.
- [5] Robert, T. M. (2002). [Statistics for laboratory method comparison studies](#). *BioPharm Statistics*, 28–32.
- [6] Robert, F. M. (2000). [General deming regression for estimating systematic bias and its confidence interval in method-comparison studies](#). *Clinical Chemistry*, 46:100–104.
- [7] Leng, L., Zhang, T., Kleinman, L., Zhu, W. (2007). [Ordinary least square regression, orthogonal regression, geometric mean regression and their applications in aerosol science](#). *Journal of Physics*, 78: 1–5.
- [8] NCME.org (2014). *Correlation coefficient*. National Council on Measurement in Education.
- [9] Lin, L. K. (1989). [A concordance correlation coefficient to evaluate reproducibility](#). *Biometrics*, 45: 255–268.
- [10] ASTM D5157-97 (2014). *Standard guide for statistical evaluation of indoor air quality models*. ASTM International, West Conshohocken, PA.
- [11] John, A. R. (2006). *Mathematical statistics and data analysis*. Duxbury Advanced.