Methods and models for developing smart environmental monitoring systems using machine learning and statistical approaches

Andrii Biloshchytskyi[1,2,†], Ilyas Kazambayev[1,*,†], Alexandr Neftissov[1,3,†], Aliya Aubakirova[1,†] and Magzhan Amangeldi[1,†]

[1] *Astana IT University, 010000, Astana, Kazakhstan*
[2] *Kyiv National University of Construction and Architecture, 03680 Kyiv, Ukraine*
[3] *Academy of Physical Education and Mass Sports, Astana 010000, Kazakhstan*

Abstract

Technological progress in recent decades has had a strong negative impact on the environment. For this reason, methods and measures for nature conservation have become relevant and necessary to achieve sustainable development. This article proposes the concept of a system for measuring emissions into the air based on the existing proposed methods. First of all, a method was proposed using the calculation of proportional coefficients based on the method of least squares relative to the measurements of reference gas analyzers. Then, a method for determining emissions based on multivariate regression analysis was proposed. Within the framework of this method, the application of the machine learning model is also considered. As a result, the use of the developed methods made it possible to develop a structural scheme for further visualization of duplicate measurements. The system will be tested in further studies.

Keywords

monitoring system, optical measurement, electrochemical measurement, coal power plants, Kazakhstan, statistical analysis

1. Introduction

At present, the achievement of sustainable development in the world is necessary and is achieved only by balancing between the technological world and the environment. Moreover, the consequences of scientific discoveries and technical breakthroughs are the main reasons for the environmental degradation. Therefore, the problem is relevant not just in Kazakhstan but worldwide [1]. One of the main approaches is the requirement to create and implement new approaches to environmental management. Moreover, the effectiveness of decisions is achieved only on the basis of reliable and accurate measurement data [2].

In environmental measurement, the most effective devices are those based on the optical and electrochemical methods. Moreover, optical methods operate mainly in the infrared and ultraviolet ranges [3-6]. Moreover, measurement in infrared makes it possible to measure the concentration of the CO_2 pollutant with less error and greater accuracy using the Fourier transform, on the other hand, it is possible to make forecasts based on ARIMA-ARCH models that take into account time trends and data volatility [7]. It is worth noting that the implementation of this solution is an expensive and complex system, which is also expensive. An alternative is portable spectrometers using an optical resonator and a Kalman filter [8-10]. These devices exhibit an accuracy of 5% or less in the measurement of volatile organic compounds (VOCs), glyoxal, nitrogen dioxide and ozone, while providing the ability to analyze multiple components in a comprehensive manner. Sulphur dioxide

[1] *STIoT 2025: Workshop on Smart Technologies and IoT, November 19-20, 2025, Almaty, Kazakhstan*
* Corresponding author.
† These authors contributed equally.
ab@astanait.edu.kz (A. Biloshchytskyi); i.kazambayev@astanait.edu.kz (I. Kazambayev)
0000-0001-9548-1959 (A. Biloshchytskyi); 0000-0003-0850-7490 (I. Kazambayev); 0000-0003-4079-2025 (A. Neftissov); 0000-0002-0884-2910 (A. Aubakirova); 0000-0003-4223-3200 (M. Amangeldi)

(SO_2) is controlled using a Bouguer-Lambert-Beer law technique [11] that provides high accuracy in the ultraviolet range, but is limited by laboratory conditions and does not take into account the effects of humidity. In recent years, special attention has been paid to methods for calibrating low-cost sensors. The most common are relative humidity correction (RHC), linear regression, and machine learning algorithms [12-14]. At the same time, each of the approaches has significant limitations: RHC is ineffective in high humidity conditions, linear regression is characterized by low flexibility, and machine learning methods are subject to overtraining and have low interpretability. A promising area is the use of nonlinear autoregressive models with exogenous variables (NARX), which demonstrate high accuracy and adaptability to changing conditions [15, 16].

In [17], a comparative analysis of the optical and electrochemical methods was carried out, as a result of which it was determined that in most cases the measurements of devices have proportionality. However, the analysis shows that many factors influence the resulting samples from devices. For this reason, the purpose of this study is to develop methods for determining emissions not only on the basis of measurements of the reference device, but also on the basis of technological parameters.

Within the framework of this study, methods of machine learning and artificial intelligence will also be used, which is also relevant in this field of research [18, 19]. The existing solutions have common disadvantage which lies in complexibility of the system and low accuracy due to water particle presence, which cannot be neglected fully. Moreover, the application of the existing methods of measurement requires additional setting coefficients to increase accuracy.

2. Materials and methods

The first known method is the method proposed in [18]. The main idea was the possibility of using a correction factor to take into account the presence of water particles in the exhaust gas. In turn, the operation of optical measuring devices is influenced by a certain dependence. First of all, areas with high humidity have a different refractive index. As a result, the intensity can be determined by the following expression [18]:

$$\mathbf{E}_{ji0}(f) = \frac{\mathbf{E}_{j0}}{k_i(f)}, \quad (1)$$

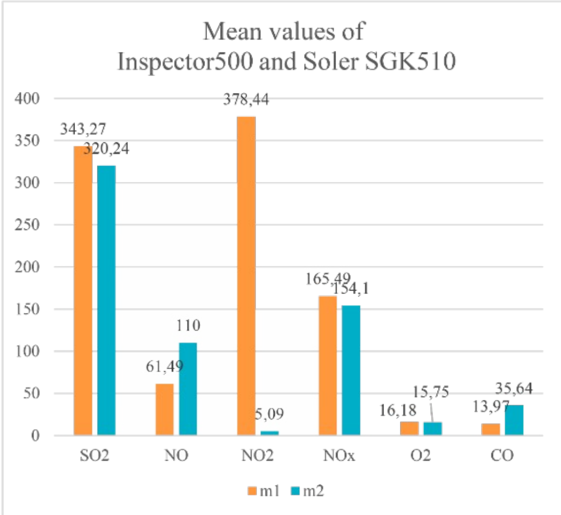
where i is the number of the area with a predominance of a certain component of the exhaust gas, $k_i(f)$ is the electrical constant of the area.

Omitting the successive mathematical transformations from [18], the intensity at the photodetector can be determined by the formula

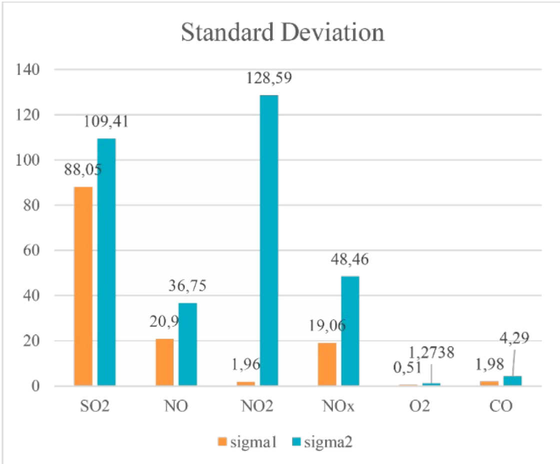
$$\sum_{j=1}^{n} \mathbf{E}_{jo0}(f) e^{i(\mathbf{u}\cdot\mathbf{r} - 2\pi jft)} = \mathbf{E}(\mathbf{r}, t) - \sum_{j=1}^{n} \sum_{i=1}^{p} \mathbf{E}_{ji0}(f) e^{i(\mathbf{u}\cdot\mathbf{r} - 2\pi jft)}. \quad (2)$$

The photomatrix perceives information as a scalar value of intensity. On the other hand, the nature of the electrochemical method is similar to the optical one. However, in this case, the difference in arc potentials is considered [18].

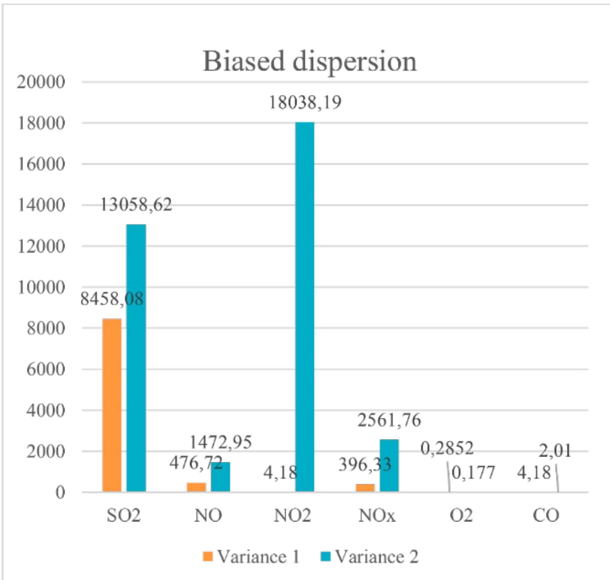
The results of the study are shown in Figure 1.



a)



b)



c)

Figure 1: Statistical indicators of Inspector500, Soler SGK510 measurements: a) mean values; b) standard deviation; c) biased dispersion.

In order to determine the detuning coefficients, the method of least squares was used [21]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x(t_i) - x_{m.c.}\right)^2}, \tag{3}$$

where $x(t_i)$ is the dimension calculated through the coefficient; $x_{m.c.}$ – actual measurement.

On the other hand, it is more efficient to perform calculations based on the relative error of the dimension:

$$RMSRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{\left(x(t_i) - x_{m.c.}\right)^2}{x_{m.c.}^{2}}}. \tag{4}$$

In turn, according to certain coefficients, it is possible to determine the dependence of the measured parameters of the working gas analyzer on the reference device. This approach allows you to accurately measure the concentration of contaminants in the exhaust gas.

3. Results

3.1. Development of a method for improving accuracy based on the calculation of proportional coefficients

The process of building a model for determining emissions and predicting them is carried out by measuring the system and transmitting it through the server to the computing server.

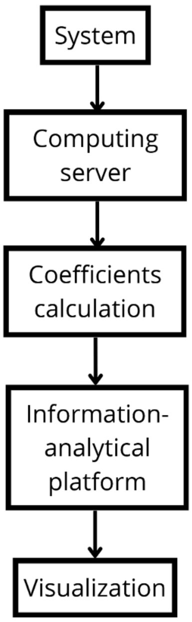


Figure 2: Structure diagram of the flow of information.

The process of building a model for determining emissions and predicting them is carried out by measuring the system and transmitting it through the server to the computing server. The result is transmitted to the information and analytical platform, which performs visualization.

In this study, in order to improve the accuracy of measurements, it is proposed to use machine learning methods to calibrate the initial data from the sensors. These methods have been successfully applied to the calibration of low-cost NO2 sensors [18] and the methane monitoring system [19]. In which trained machine learning models have contributed to a significant reduction in errors compared to traditional statistical methods. The model will be trained on environmental parameters in conjunction with raw data from the SGK510 and POLAR stationary sensors. Data from the Inspector 500 will be used as reference measurements and pointers for training.

The architecture of the methodology is provided in Figure 3.

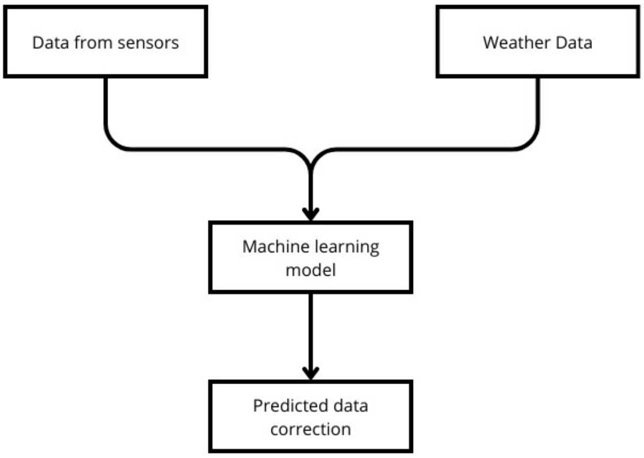


Figure 3: Data from SGK510/POLAR sensors and environmental parameters are processed by a model trained on Inspector 500 to obtain corrected values.

The use of this technique allows for accurate determination of emissions based on laboratory measurements. Despite this, it is necessary to ensure constant monitoring of the parameters on the basis of constant measurements.

3.2. Method of emission concentration determination with multivariate regression

In most cases, emissions are codependent on each other and can be used to build models that act as duplicate measurements of devices. At the same time, it is possible to use polynomial dependence for technological parameters.

The method can be implemented on the basis of existing monitoring systems, for example, SOLER SGK510 and will increase the accuracy of measurements of the concentration of pollutants when gas leaves the pipe. Moreover, OWEN HTP 10-2.3 can be used as a temperature sensor. And the PM2.5, PM10 particulate concentration meter – City Air Dust, SO2 measurement device – C-310A, CO and O2 measurement device – K-100, NO, NO2, NOx measurement device – P-205 [21, 22].

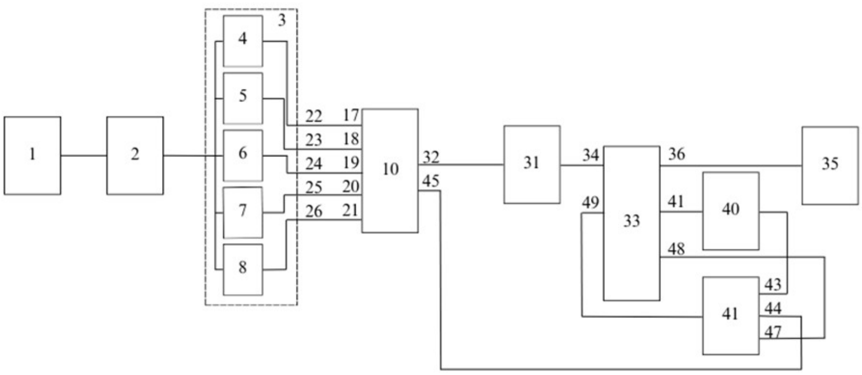


Figure 4: Structural diagram of the method for determining emission concentrations based on multivariate regression analysis.

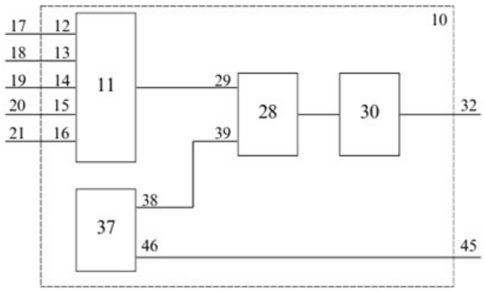


Figure 5: A structural diagram of the control unit is presented.

In the process of operation of a thermal power plant, namely the combustion of solid fuel, a multicomponent gas is released, for the purpose of measuring which samples are taken using the gas gathering unit 1. Then, in order to ensure that interference from cross-absorption of light is eliminated, the sample is dried in preconditioning unit 2. Moreover, the gas enters the measurement unit 3, in which the temperature is measured using the appropriate sensor 4, the concentration of particulate matter PM2.5, PM10 using unit 5, the SO2 content in block 6, the CO and O2 content in block 7, the content of NO, NO2, NOx block 8. The measurements are transmitted via the input unit 11 and stored in the control unit 10, namely in the first computing unit 28. In most cases, the values obtained are transmitted as particles per millimeter (ppm), whereby a conversion function is used to convert the reference oxygen concentration of 6% to the mg/m3 unit in the first computing unit 28. For further presentation of the data, the measurements are transmitted via the appropriate unit 31 to the second computing unit 33. In this case, statistical indicators and data storage are calculated. Then, the calculation unit 40 performs an analysis of variance and a Student's test to determine the similarity of the actual measurement sample and the reference sample stored in the data acquisition unit 37. In addition, the main parameters of the technological process are also used to compile a multi-component regression model using the calculation of the correlation coefficient:

$$r_{ij} = \frac{\sum_{l=1}^{n}(x_{il} - \frac{\sum_{l=1}^{n} x_{il}}{n})(x_{jl} - \frac{\sum_{l=1}^{n} x_{jl}}{n})}{\sqrt{\sum_{l=1}^{n}(x_{il} - \frac{\sum_{l=1}^{n} x_{il}}{n})^2 \sum_{l=1}^{n}(x_{il} - \frac{\sum_{l=1}^{n} x_{il}}{n})^2}}, \quad (5)$$

where x_{il} is the *i*-th variable from the *l*-order data collection unit in the data array.

Based on the obtained value from (1), a correlation matrix is constructed:

$$\begin{bmatrix} 1 & |r_{12}| & \cdots & |r_{1n}| \\ |r_{21}| & 1 & \cdots & |r_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |r_{n1}| & |r_{n2}| & \cdots & 1 \end{bmatrix}. \quad (6)$$

The obtained values from the matrix (2) are then compared in the comparison block 40, which selects the calculated data according to the main criterion, within which the most correlated samples are determined to build a multivariate dependence:

$$r_{ij} < r_s, \quad (7)$$

where r_s is the established value of the correlation coefficient.

The obtained values in the comparison unit 40 are sent to the second input 49 of the computational unit 33, which in turn formulates a multi-component formula:

$$y = \mathbf{A} \times \mathbf{X} + B;$$
$$\begin{cases} \mathbf{A} = \begin{bmatrix} a_{10} & a_{11} & \dots & a_{1n} \\ a_{20} & a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m0} & a_{m1} & \dots & a_{mn} \end{bmatrix}; \\ \mathbf{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{10} & x_{21} & \dots & x_{m1} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{bmatrix}, \end{cases}, \tag{8}$$

where A is the coefficient matrix, X is the variable matrix, and B is the displacement constant.

In turn, the computing unit 33 performs the recalculation, determining the measurement error relative to the reference value according to standard expressions:

$$\begin{cases} \delta = y - x_r; \\ \varepsilon = \dfrac{y - x_r}{x_r} \cdot 100\%, \end{cases} \tag{9}$$

where δ is the absolute error, ε is the relative error, and x_r is the reference value.

The compute unit 33 makes a decision and then sends the data to the visualization unit 35, which presents the data as trends and alerts when threshold values have been exceeded. Moreover, correlation analysis makes it possible to identify the main reason for the increase in the concentration of pollutants.

A method for determining emissions generated during the operation of a thermal power plant based on solid fuel, containing a gas collection unit, a preliminary treatment unit connected by the inlet to the output of the gas collection unit, the first measurement unit consisting of a temperature sensor, a PM2.5, PM10 particulate concentration meter, SO2 measurement devices, CO and O2 measurement devices, NO, NO2, NOx measurement devices connected by inputs through the input of the measuring unit to the output of the preconditioning unit, second measuring unit, control unit in the form of an input unit connected first, the second, third, fourth and fifth inputs through the corresponding inputs of the control unit to the first, second, third, fourth and fifth outputs of the first measurement unit, and the sixth – to the output of the second measurement unit, the first computing unit connected by the first input to the output of the unit inputs, an output unit connected by the input to the output of the first computing unit, a data transmission unit connected through the output of the control unit to the output unit, the second computing unit connected by the first input to the output of the data transmission unit, the visualization unit, connected by the input to the first output of the second computing unit, differing in that the data acquisition unit connected by the first output to the second input of the first computing unit, the calculation unit connected by the input to the second output of the second computing unit, the comparison unit, connected by the first input to the output of the calculation unit, the second to the second output of the data acquisition unit, the third to the third output of the second computing unit, and the output to the second input of the second computing unit.

3.3. Development of the concept of an intelligent measurement system

Taking into account the developed system and the proposed method for determining emissions based on technological parameters, the use of machine learning makes it possible to increase the accuracy of measurement of the concentration and content of various components in the exhaust gas. Therefore, this study proposes an upgrade of the existing system, consisting in the introduction of additional modules for calculating proportional coefficients and polynomial dependencies carried out in the cloud into the proposed scheme [17], using a machine learning model for further visualization. Figure X shows a structural diagram of information flows between modules.

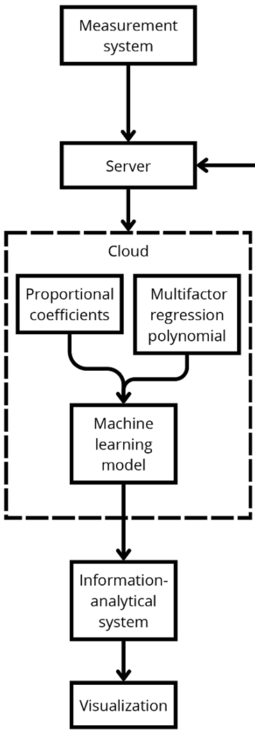


Figure 6: Structural model of a measurement system using smart technologies.

Table 1
Measurements performed using the Portable device

Instrument	SO2 (mg/m3)	NO (mg/m3)	NO2 (mg/m3)	NOx (mg/m3)	O2 (%)	CO (mg/m3)
Portable	228.2	119.3	4.3	165.5	16.7	36.3
Portable	246.5	102.2	8.5	148.2	15.1	38.9
Portable	118.6	75.8	10.2	115.3	15.1	43.1
Portable	321.0	132.4	4.8	183.9	15.3	36.5
Portable	316.2	95.4	6.0	135.7	15.9	36.5
Portable	133.2	3.8	4.0	10.4	19.7	23.3
Portable	375.1	127.3	3.2	174.8	15.4	35.5
Portable	405.4	132.9	3.5	182.5	15.1	35.6
Portable	416.2	133.0	4.2	183.6	15.1	35.4
Portable	419.7	133.8	4.0	184.5	15.2	35.5
Portable	430.8	132.6	4.2	183.2	15.3	35.4
Portable	432	131.5	4.2	181,6	15.1	35.7

In this case, the information from the measuring system is transmitted via telecommunications servers to the cloud, where the coefficients and dependencies are calculated. The results obtained are transmitted through the server to the information and analytical platform, where visualization is carried out. Figure 7 shows the possible outcome of the Learning Machine Work Representation shown in Figure 6. The system were tested on the data presented in Table 1.

Table 1 presents measurements from portable devices. Nevertheless, in Table 2 the data presented for stationary devices.

Table 2
Measurements performed using the Stationary devices

Instrument	SO2 (mg/m3)	NO (mg/m3)	NO2 (mg/m3)	NOx (mg/m3)	O2 (%)	CO (mg/m3)
Stationary	218.3	87.0	535.3	139.3	16.3	16.7
Stationary	285.9	64.7	398.1	166.6	16.0	19.3
Stationary	171.5	41.9	258	115.9	17.7	13.0
Stationary	277.5	72.2	444.5	151.2	16.7	13.5
Stationary	336.8	66.9	411.9	176.2	15.9	14.1
Stationary	302.8	67.6	415.8	176.1	15.9	14.3
Stationary	387.1	67.4	415	177.8	15.9	12.9
Stationary	416.7	67.8	417	178.1	15.9	12.8
Stationary	419.4	67.5	415.6	174.1	16	11.9
Stationary	414.6	-0.4	-2.3	174.8	16.1	12.5
Stationary	442.2	68.2	419.7	179	15.9	13.1
Stationary	446.4	67.1	412.7	176.8	15.9	13.5

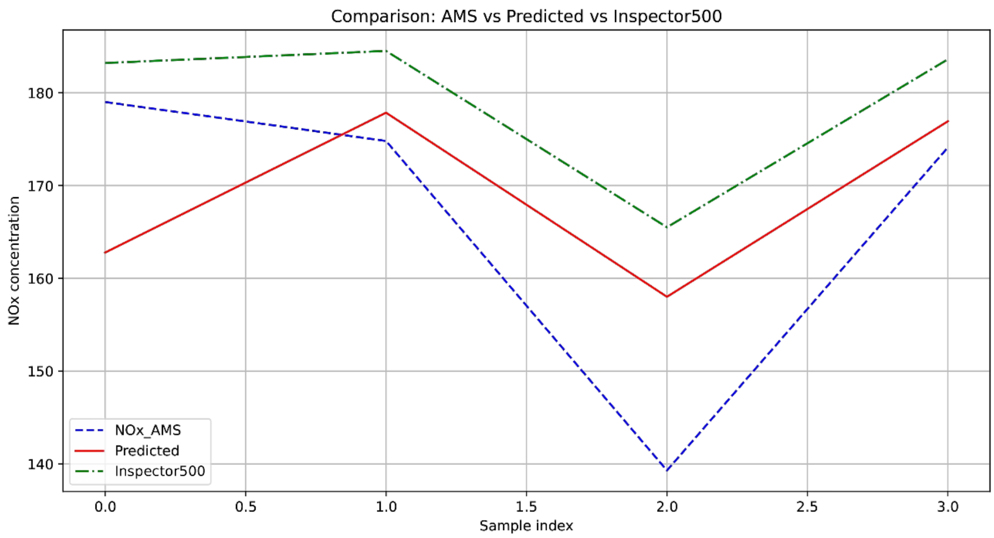


Figure 7: Output of the machine learning model.

The results of model testing are shown on Figure 1, it can be analyzed that the predicted trend matches the Inspector500 values, highlighting the potential of the approach. However, the results are to be considered as indicative, which will give a concept meaning for the further data collection and analysis.

4. Discussion

The results of the comparative analysis showed in most cases the similarity of statistical indicators such as arithmetic mean expectation, standard deviation and biased variance (Figure 1). However, it is worth noting that in a previous study it was determined that proportionality can be established in most cases, however, for some variables this is not possible. Moreover, the samples were not large, which is also a strong drawback. Despite this, it can be noted that the availability of continuous collection of laboratory measurement data will allow the system to work with greater accuracy. As a result, the use of artificial intelligence is considered to ensure that the error is reduced.

On the other hand, in order to ensure the continuity of the data flow, the use of correlation multivariate analysis in order to determine emissions is recommended. The difference from the previous approach is the use of process variables rather than laboratory emission measurements. In this case, on the basis of existing knowledge, the dependencies of emissions on certain parameters make it possible to construct a polynomial expression, determining emissions into the atmosphere. Despite the limited amount of data for analysis, a machine learning approach could be applied to this task. It was considered to utilize synthetic data augmentation with controlled noise to enhance training set and balance regression. The primary goal for the model is to predict the value of the real NOx data based on AMS and compared to Inspector500 as a ground truth.

This concept can be implemented on various types of equipment. Duplication of indirect measurements allows you to determine the accuracy of measurements by comparison, which ensures reliable and continuous measurements. Moreover, the use of two methods within the framework of artificial intelligence also makes it possible to predict possible emissions.

In future studies, it is planned to use these methods for testing.

5. Conclusions

Comparative analysis made it possible to develop a method for calculating proportional coefficients based on measurements of reference gas analyzers. Moreover, in many cases, a linear dependence cannot be constructed. In such cases, it is necessary to use complex dependencies with several coefficients.

On the basis of the obtained technological parameters, it is possible to construct polynomial complex expressions with coefficients obtained by calculating correlations using multivariate regression analysis. The main indicator is a comparison based on a correlation matrix. The use of machine learning will ensure accuracy.

Based on the results of the proposed models, a measurement system concept was developed using cloud technology and machine learning to duplicate the measurement in order to ensure accuracy.

Acknowledgements

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number BR21882258 "Development of Intelligent Information and Communication Systems Complex for Environmental Emission Monitoring to Make Decisions on Carbon Neutrality".

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] United Nations Environment Programme. (2024). *Emissions Gap Report 2024: No more hot air … please! With a massive gap between rhetoric and reality, countries draft new climate commitments.* Retrieved August 20, 2024, from https://wedocs.unep.org/20.500.11822/46404.

[2] Xu, K., & Li, Z. (Eds.). (2024). *Gas sensors for monitoring environmental changes*. MDPI Books. https://mdpi-res.com/bookfiles/book/10138/Gas_Sensors_for_Monitoring_Environmental_Changes.pdf?v=1758157789.

[3] Lal, K., Jaywant, S. A., & Arif, K. M. (2023). Electrochemical and optical sensors for real-time detection of nitrate in water. *Sensors, 23*(16), 7099. https://doi.org/10.3390/s23167099.

[4] Mota, F. A. R., Passos, M. L. C., Santos, J. L. M., & Saraiva, M. L. M. F. S. (2024). Comparative analysis of electrochemical and optical sensors for detection of chronic wounds biomarkers: A review. *Biosensors & Bioelectronics, 241*, 116095. https://doi.org/10.1016/j.bios.2024.116095.

[5] Lin, J., Wang, Y., Chen, Q., et al. (2025). Near-Infrared mobile cloud OA-ICOS sensor system for atmospheric carbon dioxide monitoring. *Analytical Chemistry, 97*(4), 2345–2352. https://doi.org/10.1021/acs.analchem.4c05644.

[6] Ding, N., Xi, Y., Jiang, W., Li, H., Su, J., Yang, S., Lie, T. T., et al. (2025). State-of-the-art carbon metering: Continuous emission monitoring systems for industrial applications. *Heliyon, 11*(2), e42308. https://doi.org/10.1016/j.heliyon.2025.e42308.

[7] Dritsaki, M., & Dritsaki, C. (2020). Forecasting European Union CO_2 emissions using autoregressive integrated moving average-autoregressive conditional heteroscedasticity models. International Journal of Energy Economics and Policy, 10(3), 411–423. https://doi.org/10.32479/ijeep.9249.

[8] Womack, C. C., Min, K.-E., Washenfelder, R. A., et al. (2022). A lightweight broadband cavity-enhanced spectrometer for NO_2 measurement on uncrewed aerial vehicles. *Atmospheric Measurement Techniques, 15*(13), 4021–4039. DOI:10.5194/amt-2022-203.

[9] Liang, L. (2021). Calibrating low-cost sensors for ambient air monitoring: Techniques, trends, and challenges. Environmental Research, 200, 111163. https://doi.org/10.1016/j.envres.2021.111163.

[10] Guan, P. (2024). Carbon dioxide emissions prediction framework based on deep learning, combination model and optimization. In 2024 IEEE 2nd International Conference on Electrical, Automation and Computer Engineering (ICEACE) (pp. 1466–1470). IEEE. https://doi.org/10.1109/ICEACE63551.2024.10898916.

[11] Biloshchytskyi, A., Kazambayev, I., Toxanov, S., Kirichenko, L., Biloshchytska, S., & Abzhanova, D. (2025). Development of emission determination method concept. In 2025 IEEE Conference on Technologies for Sustainability (SusTech) (pp. 1–7). IEEE. https://doi.org/10.1109/SusTech63138.2025.11025756.

[12] Wang, F. (2021). A data calibration method for micro air quality detectors based on hybrid models. *Scientific Reports, 11*, 13843. https: 10.1038/s41598-021-00804-7.

[13] Bissengaliyeva, D., Amanzholov, T., Rakhimkul, A., & Yedilkhan, D. (2023). Determination of the reliability of air pollution measurement data based on vehicular emission recognized as concomitant in Astana. Scientific Journal of Astana IT University, 13(13), 42–51. https://doi.org/10.37943/13QOJG5081M.M.

[14] German Emissions Trading Authority (DEHSt). (2019). Application of continuous emissions measurement systems for the determination of CO_2 emissions: Experience and assessments (Report No. 34).

[15] Montrucchio, B., Giusto, E., Vakili, M. G., Quer, S., Ferrero, R., & Fornaro, C. (2020). A densely-deployed, high sampling rate, open-source air pollution monitoring WSN. IEEE Transactions on Vehicular Technology, 69(12), 15786–15799. https://doi.org/10.1109/TVT.2020.3035554.

[16] Zhou, H., Shan, Y., Han, Z., Meng, H., He, Y., & Xu, J. (2023). Design and application of CEMS-based measurement equipment for carbon dioxide emissions from stationary sources. In 2023 3rd International Signal Processing, Communications and Engineering Management Conference (ISPCEM) (pp. 201–205). IEEE. https://doi.org/10.1109/ISPCEM60569.2023.00043.

[17] Neftissov, A., Biloshchytskyi, A., Kazambayev, I., Kirichenko, L., Zhalmagambetova, U., & Biloshchytska, S. (2024). Analysis of the existing air emissions detection methods for stationary pollution sources monitoring. Applied Sciences, 14(23), 10934. https://doi.org/10.3390/app142310934.

[18] K. Kolesnikova, L. Naizabayeva, A. Myrzabayeva and R. Lisnevskyi, "Use the neural networks in prediction of environmental processes," 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, 2024, pp. 625-630, doi: 10.1109/SIST61555.2024.10629330.

[19] Koziel, S., Cheng, Q. S., Baltazar, J., Sudhakaran, M., & Mokhtari, S. (2024). High-performance machine-learning-based calibration of low-cost nitrogen dioxide sensor using environmental parameter differentials and global data scaling. Scientific Reports, 14(1), 26120. https://doi.org/10.1038/s41598-024-77214-y.

[20] Andrews, B., Schmidt, A., O'Brien, P., McFarlane, J., & Ingham, D. (2023). Application of machine learning for calibrating gas sensors for methane emissions monitoring. Sensors, 23(24), 9898. https://doi.org/10.3390/s23249898.

[21] A. Sarinova, R. Lisnevskyi, A. Biloshchytskyi and A. Akizhanova, "The Lossless Compression Algorithm of Hyperspectral Aerospace Images With Correlation and Bands Grouping," 2022 International Conference on Smart Information Systems and Technologies (SIST), Nur-Sultan, Kazakhstan, 2022, pp. 1-5, doi: 10.1109/SIST54437.2022.9945821.
[22] A. Neftissov, A. Sarinova, I. Kazambaev, L. Kirichenko, R. Lisnevskyi and L. Rzayeva, "Development of the Smart Office Concept," 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, 2023, pp. 382-386, doi: 10.1109/SIST58284.2023.10223512.