Experimental Setup and Machine Learning-Based Prediction Model for Electro Cyclone Filter Efficiency: Filtering of Ship Particulate Matter Emission

Aleksandr Šabanovič , Jonas Matijošius , Dragan Marinković * , Aleksandras Chlebnikovas , Donatas Gurauskis , Johannes H. Gutheil , Artūras Kilikevičius

Posted Date: 12 December 2024

doi: 10.20944/preprints202412.1023.v1

Keywords: ship emissions; air quality; particulate matter; emission control policies; health risks of PM; hybrid filtration systems; electrostatic precipitators (ESPs); machine learning prediction models

Article

Experimental Setup and Machine Learning-Based Prediction Model for Electro Cyclone Filter Efficiency: Filtering of Ship Particulate Matter Emission

Aleksandr Šabanovič [1], Jonas Matijošius [2], Dragan Marinkovič [2,3,*], Aleksandras Chlebnikovas [2], Donatas Gurauskis [2], Johannes H. Gutheil [4] and Artūras Kilikevičius [2]

[1] Department of Mechanical and Material Engineering, Faculty of Mechanics, Vilnius Gediminas Technical University-VILNIUS TECH; Plytinės str. 25, LT-10105 Vilnius, Lithuania
[2] Mechanical Science Institute, Vilnius Gediminas Technical University-VILNIUS TECH; Plytinės str. 25, LT-10105 Vilnius, Lithuania
[3] Department of Structural Analysis, TU Berlin, 10623 Berlin, Germany
[4] Institute of Particle Process Engineering, Rheinland-Pfälzische Technische Universität (RPTU); Gottlieb-Daimler-Straße 44, D-67663, Kaiserslautern, Germany
* Correspondence: dragan.marinkovic@tu-berlin.de; Tel.: +491 763 0167755

Abstract: Ship emissions significantly impact air quality, particularly in coastal and port regions, contributing to elevated concentrations of $PM_{2.5}$, and PM_{10}, with varying effects observed across different locations. This study investigates the effectiveness of emission control policies, inland and port-specific contributions to air pollution, and the health risks posed by particulate matter (PM). A regression discontinuity model at Ningbo Port revealed that ship activities show moderate $PM_{2.5}$ and PM_{10} variations. In Busan Port, container ships accounted for the majority of emissions, with social costs from pollutants estimated at USD 31.55 million annually. Inland shipping near the Yangtze River demonstrated significant PM contributions, emphasizing regional impacts. Health risks from $PM_{2.5}$, a major global toxic pollutant, were highlighted, with links to respiratory, cardiovascular, and cognitive disorders. Advances in air purification technologies, including hybrid electrostatic filtration systems, showed promising efficiency in removing submicron particles and toxic gases, reducing energy costs. A Random Forest machine learning model developed to predict particulate concentrations post-cleaning demonstrated robust performance (MAE=0.49 P/cm^3, R^2=0.97). These findings underscore the critical need for stringent emission controls, innovative filtration systems, and comprehensive monitoring to mitigate the environmental and health impacts of ship emissions.

Keywords: ship emissions; air quality; particulate matter; emission control policies; health risks of PM; hybrid filtration systems; electrostatic precipitators (ESPs); machine learning prediction models

1. Introduction

1.1. The Impact of Ship Particulate Matter Emissions on Air Quality

The impact of international shipping on air pollution in Europe research findings indicate that ship emissions significantly affect air quality, especially in coastal regions [1]. The importance of regulating ship emissions to improve air quality and reduce health risks is highlighted [2]. Primary Particulate Matter (PPM) is directly emitted during fuel combustion [3]. The sub-components and characteristics of PPM is ash which contains inorganic residues from fuel combustion, black carbon which absorbs sunlight and contributes to climate warming, organic carbon which is formed from incomplete combustion of fuels . These particulates are fine $PM_{2.5}$ and have health and environmental impacts. [4]

Using a regression discontinuity (RD) model, the effectiveness of ship emission control area (ECA) policies on air quality at Ningbo Port were evaluated. The study analysed the impact of ship

 doi:10.20944/preprints202412.1023.v1

emissions on concentrations of PM2.5, and PM10. The results indicate that PM2.5 and PM10 levels vary with ship activities. [5]

Inland ship emissions significantly contribute to air pollution in China along the river Yangtze in the Nanjing region. The estimated emissions PM10, and PM2.5 from September 2018 to August 2019 were 3.8 and 3.3 kton, respectively, with the highest emissions occurring in the summer. [6]

Container ships were identified as the largest contributors to particulate emissions at Busan Port. The average annual emissions from 2015 to 2019 were PM2.5 (0.05%), and PM10 (0.05%). The total social costs of air pollutants emitted from ships at Busan Port in 2019 were estimated to be between USD 11.06 million and USD 100.64 million, depending on the valuation method used. The Base Case estimate was USD 31.55 million. [7]

International shipping significantly impacts air quality, particularly in coastal regions. Effective regulation of ship particulate emissions is crucial to reduce health risks and environmental damage [8,9]. Emission control policies show promise, but challenges remain, especially with inland emissions and economic costs at major ports [10,11].

1.2. The Impact of Ship Emissions of Particulate Matter on Human Health

Air pollution was the 4th leading risk factor for early deaths globally in 2019, contributing to 6.67 million premature deaths. Particulate matter (PM), especially fine particulate matter (PM2.5), is a major toxic exposure risk, causing more than two million hospital admissions and premature deaths annually. PM is a major risk factor for mortality, causing 4.2 million deaths and 103.1 million disability-adjusted life-years (DALYs) globally. The International Agency for Research on Cancer classifies PM as a group 1 carcinogen due to its toxicity. PM exposure is linked to adverse health effects such as pulmonary inflammation, bronchoconstriction, and chronic respiratory diseases like asthma and COPD. Epidemiological studies show that PM contributes significantly to global disease burden, particularly in developing countries like China and India. [12–16]

Air pollutants such as particulate matter can trigger inflammation in the brain, leading to cognitive decline. Increased production of reactive oxygen species (ROS) can damage brain cells. Pollutants may disrupt the blood-brain barrier, allowing harmful substances to enter the brain. Air pollution can impair blood flow to the brain, affecting cognitive function. Even low levels of PM can lead to increased mortality and morbidity, affecting both adults and children. Short-term exposure is linked to respiratory and cardiovascular diseases. The effects are more pronounced in vulnerable groups such as the elderly, children, and individuals with preexisting cardiovascular and respiratory conditions. PM consists of a complex mixture of solid and liquid particles, including combustion particles, secondary inorganics, and crustal-derived particles. Fine particles (PM2.5) and ultrafine particles (less than 0.1 μm) are primarily produced by fossil fuel combustion. These particles can penetrate deep into the lungs and enter the bloodstream, causing systemic health effects. [15–19]

Air pollution, especially PM2.5, is a major global health risk, causing millions of premature deaths and severe health issues like respiratory and cardiovascular diseases, and cognitive decline. The impact is particularly severe in developing countries and vulnerable populations. Reducing PM emissions is essential to improve global health.

1.3. Efficiency, Efficiency Factors and Lifetime of Different Technologies of Air Cleaning from Particulate Matter

Initial results of a holistic performance assessment showed that duct-type electrostatic precipitators (ESPs) used in heating, ventilation, and air conditioning (HVAC) systems with dielectric coatings (L1-IFD) showed enhanced removal efficiency of submicron particles compared to ESPs without dielectric coatings. The initial filtration efficiency for submicron particles was higher for ESPs with dielectric coatings, with overall efficiencies comparable to F7 filters but with 85% lower energy consumption. The pressure drop across ESPs increased with face velocity, and the resistance of L1-IFD was higher than that of ESPs without dielectric coatings due to the smaller gap between the collecting plates. Dust loading reduced the filtration efficiency of all ESPs, with L1-IFD showing a relatively large efficiency attenuation of 33.6%. Post-washing, the filtration efficiency of ESPs

 doi:10.20944/preprints202412.1023.v1

improved but did not fully recover to initial levels, indicating incomplete removal of dust on the surface. The maximum quality factor attenuation caused by dust loading was 55.3%, and the maximum attenuation after washing was 17.5%. The ozone emission rates (OERs) were significantly higher for ESPs without dielectric coatings compared to L1-IFD. The OERs for ESPs without dielectric coatings were approximately 2–7 times higher than for L1-IFD. [20]

Investigation of the performance and degradation of cylindrical corona electrodes in electrostatic precipitators (ESPs) showed that the collection efficiency of the ESPs is influenced by both the diameter and pitch of the cylindrical electrodes. Thinner electrodes result in higher collection efficiency due to stronger electric fields and larger ionization regions. Shorter pitch between electrodes also leads to higher collection efficiency because of the increased number of electrodes. The collection efficiency varies nonlinearly and inversely with the diameter of the electrodes, suggesting an optimal diameter for maximum efficiency. There is a trade-off between achieving high collection efficiency and maintaining the longevity of the electrodes. Stronger electric fields improve collection efficiency but accelerate electrode degradation due to oxidation. It is suggested that optimizing the pitch and diameter of the electrodes would balance collection efficiency and electrode life. [21]

High efficient particulate air filters (HEPA) and electrostatic precipitators (ESP) have limitations such as high energy consumption and low filtration efficiency, respectively. Electrostatic assisted air filtration systems combine HEPA and ESP to achieve high filtration efficiency and low energy consumption. The presence of the PMMA tank has a minimal effect on the V-I characteristics. When the applied voltage is 20 kV, the relative error of the total current value between the cases with and without the dielectric tank is only 3%. [22]

Jian Li et al. presented a novel air purification filter designed to efficiently remove fine particulate matter (PM2.5) from the air. The filter is based on a coaxial core–shell structure composed of CuO@NH2-MIL-53(Al) nanowire arrays grown on a copper mesh. This design leverages both local and external electric fields to enhance the removal efficiency of pollutants. The filter achieves a PM2.5 removal rate of 98.72% with an external electric field and 44.41% without it. The filter maintains stable air pollution removal efficiency after repeated filtration and cleaning cycles, demonstrating its reusability. The filter is constructed using a copper mesh for its excellent electrical conductivity and rigidity, with CuO nanowires providing a high surface area for pollutant capture. The NH2-MIL-53(Al) layer enhances the filter's ability to adsorb toxic gases. The filter operates through a combination of long-range electrostatic interactions (enhanced by an external electric field) and short-range electrostatic interactions (local electric field), which together improve the capture of PM2.5. [23]

Air temperature and humidity impacts the performance of electric precipitators used for air purification [24,25]. When the electric filter is powered by a constant voltage, the electric current (discharge current) increases as the air temperature goes up. When using pulsed voltage (a type of voltage that turns on and off rapidly), the discharge current stays almost the same regardless of the temperature. With constant voltage, the discharge current decreases as the air gets more humid. With pulsed voltage, the discharge current remains stable, even as humidity changes. Pulsed voltage generates a higher discharge current compared to constant voltage, making the filter more effective. Pulsed voltage makes the filter less affected by changes in temperature and humidity, ensuring consistent performance. [26]

The development and application of hybrid electrostatic filtration systems are designed to address the limitations of conventional filtration methods [27], particularly in removing submicron and nanoparticles from exhaust gases. High temperatures (>1000 K) in industrial processes limit the use of bag filters. Cyclones are ineffective for PM2.5 removal. Electrostatic precipitators have a "penetration window" where efficiency drops below 50% for particles in the 100 nm to 1 μm range. The solution for that is combining electrostatic devices (e.g., electrostatic precipitators, prechargers, agglomerators) with mechanical filters (e.g., fibrous filters, cyclones). Hybrid systems show significant improvements in collection efficiency for PM2.5 particles. The system achieved a PM2.5 collection efficiency of over 99.999%, with a significantly slower increase in pressure drop across the bag filter compared to conventional systems. Electrically energized filters and hybrid electrostatic

 doi:10.20944/preprints202412.1023.v1

filters reduce the pressure drop across the filter media. This is achieved by charging the particles, which enhances their deposition on the filter fibres and reduces the buildup of a dense dust cake. Lower pressure drop translates to reduced energy consumption for maintaining airflow through the filters, thereby lowering operational costs. Hybrid systems can handle high-temperature exhaust gases by using components like cyclones and electrostatic precipitators that are resistant to high temperatures, followed by cooling stages before mechanical filtration. Hybrid electrostatic filtration systems offer a highly efficient solution for exhaust gas cleaning, particularly for fine and ultrafine particles. [28]

Advancements in electrostatic precipitators (ESPs) and hybrid filtration systems enhance air purification efficiency and reduce energy consumption. ESPs with dielectric coatings improve submicron particle removal and lower ozone emissions. Hybrid systems combining ESPs with mechanical filters achieve high PM2.5 removal rates. Novel filters using CuO@NH2-MIL-53(Al) nanowire arrays effectively remove fine particulates and toxic gases. These innovations provide more effective and sustainable air purification solutions.

1.4. Modelling of ESPs

Investigation of the optimization of electrode structures in electrostatic precipitators (ESPs) to enhance dust removal efficiencies compared wire-plate, plate-hole, and hole-hole configurations through experiments and simulations Key findings revealed that the hole-hole structure achieves higher peak currents and better dust removal efficiency for particles larger than 1 μm compared to other configurations. Simulations showed that hole-hole structured electrodes increase electric field intensity at openings and reduce surface charge density, mitigating reverse corona phenomena. Additionally, perforations in collection plates help manage ion wind effects, improving the capture of fine particulate matter. The research concluded that the hole-hole electrode structure significantly enhances ESP efficiency, particularly for smaller particles, by optimizing electric and flow fields and reducing secondary dust re-entrainment. [29]

Investigation of Effect of High-Voltage Electrostatic Precipitator Dust Collection Plate Structure on Collection Efficiency explored how changes in the structure of dust collection plates affect the flow field and electric field distribution in electrostatic precipitators (ESPs). A multiphysics-coupled numerical model was used to analyse linear flat plates and folded plates with different pole configurations. Results showed that the folded plate design improves the near-plate electric field and flow field, reducing particle re-entrainment and enhancing dust removal efficiency. The folded plate structure decreases the flow velocity near the dust collection plate by approximately 20% compared to traditional linear flat plates, particularly at an inlet velocity of 0.5 m/s. This reduction in flow velocity increases particle residence time, improving dust collection efficiency. [30]

A new time-dependent model was developed to simulate continuous DC supply with impulse voltage supply, highlighting that impulse mode not only saves energy but also enhances precipitation efficiency, building on previous steady state models. The results indicated that during the "ON" state of impulse voltage, particles are effectively charged and precipitated, while during the "OFF" state, charged particles continue to precipitate due to residual charge. [31]

Accurate long-term modelling is essential for optimizing ESP performance, particularly in multi-zone configurations where different sections of the precipitator experience varying conditions and require tailored rapping cycles and energization strategies. [32]

Optimizing electrode structures in electrostatic precipitators (ESPs) significantly enhances dust removal efficiency [33]. The hole-hole configuration outperforms wire-plate and plate-hole setups, achieving higher peak currents and better efficiency for particles larger than 1 μm [34]. Perforated electrodes improve electric field intensity and reduce reverse corona effects, while folded plate designs enhance near-plate electric and flow fields, reducing particle re-entrainment [35]. Additionally, impulse voltage supply increases precipitation efficiency and energy savings [36]. Accurate long-term modelling is crucial for optimizing ESP performance, especially in multi-zone configurations with varying conditions [37].

 doi:10.20944/preprints202412.1023.v1

2. Materials and Methods

The data for our experimental prototype electro cyclone filter system efficiency prediction model was gathered from the conducted experiments. The gathered data was used for model training and validation implementing machine learning (ML) modelling the filtering efficiency of the particulate matter. The experimental setup featured a gas flow generation and management system, designed to deliver gas output with air flow rate levels ranging from 67 to 411 m^3/hour. Flow regulation was achieved using a frequency converter adjustable between 10–60 Hz, with a precision of 0.05–0.10 Hz [38]. The system's innovative architecture included a multi-channel helical cyclonic filter, which integrated three separate deposition zones within a single housing. To introduce controlled particulate matter, a Palas RGB 1000 mobile particulate generator was utilized, enabling precise aerosol feeding into the air duct just after the fan and before the cleaning system. Ship emissions primary particulate matter closely resembling test particles (density range 1.80-2.56 g/cm^3, refractive index range 0.45-1.52) were used for the experiments, and particle concentration and size were measured using a Palas Welas Digital 3000 light-scattering spectrometer, with a measurement range of 0.2 μm to 10 μm. A detailed grain composition of the tested particulate matter is presented in Table 1.

Table 1. Composition of the tested particulate matter..

PM diameter, μm	0.2	0.5	1	3	5	8	10
% under for full sample	0.8	1.2	5.2	19.8	29.5	39.3	44.8

The setup also included a low-voltage electrostatic air filter, featuring components such as a flow equalizer, aerodynamic test points, and particle concentration test points located before and after the cleaning process. The pipe diameter of the upstream (input) air flow is 140 mm. The downstream (output) pipe diameter is 250 mm.

The experimental results were validated with the experimental study of Šabanovič et.al [39] then data reformatted to a model training dataset excel file. The dataset sample of ten first rows is shown in Appendix A. Data in it is stated as follows: first column is the particle size that our measurement devices are able to count; second column states dosage speed of the particles, the bigger the number is stated the more particulate matter is being injected into the incoming air flow; third column states the input voltage that goes to power our electric air filter part; fourth column states the air flow into which the particulate matter is inserted; fifth column states the Particle concentration before air cleaning, this concentration comes in our filter with the incoming air flow; sixth column states the Particle concentration after air cleaning, this is the concentration which comes out of the filter after the cyclone and electric air filtering; seventh column states the manually calculated efficiency of the air filtering process – this directly corresponds with the difference of particle concentration before and after the cleaning process.

As mentioned above our goal was to develop a machine learning model to predict Particle Concentration After Cleaning based on these parameters. This is valuable because it allows to simulate and forecast filter performance under different conditions without running all possible experiments. we trained the model on data where the dosage device feed piston speeds which controls the feeding intensity of the particles which are being fed into the air flow before air cleaning (dosage speed) were 2, 4, and 16 mm/min and reserved data with a dosage speed of 8 mm/min as the test set. Dosage speeds of 2, 4, 8 and 16 mm/min correspond to 7.69, 15.25, 30.76 and 61.51 mg/s of particulate matter per unit time. However, to not make the simulation too complex we used the initial units of mm/min. This allowed us to validate the model's predictive accuracy on an intermediate dosage speed it had not seen in training.

There was selected a Random Forest Regressor as our model. This choice was based on several advantages such as model’s ability to capture complex relationships in the data, which is useful since filtering efficiency is not purely linear [40]. Random forests provide insights into which features are most influential in making predictions. Random forests handle noisy data well, which is important in experimental setups with some variability [41].

 doi:10.20944/preprints202412.1023.v1

The Python script was prepared for modelling. The script imports the dataset and extracts the necessary data such as features (particle size, dosage speed, voltage and air flow rate) and target (particle concentrations after air cleaning). The data set is split into training data (where dosage speeds are 2, 4, and 16 mm/min) and testing data (where dosage speed is 8 mm/min). The script initializes a Random Forest Regressor model. The model is trained on the training set features (X_train) to predict the target variable (Y_train). Predictions (Y_pred) are generated for the test set using the trained model. Performance metrics such as Mean Absolute Error (MAE) and R^2 are calculated.

After training, we used the test data (dosage speed of 8 mm/min) to evaluate the model's performance. This step provided insight into the model's ability to generalize to unseen data. We calculated two key metrics: MAE and R^2. The MAE is a measure of the average absolute error between actual and predicted values. It quantifies the model's accuracy by measuring how close the predictions are to the actual observations, regardless of whether the errors are positive or negative [42].

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|, \tag{1}$$

where y_i is the actual value (experiment data), $\hat{y}_i$ is the predicted value, n is the total number of predictions.

The R^2 score (coefficient of determination) measures how well the model explains the variance in the target variable. It is a metric commonly used for regression models and ranges from 0 to 1, where values closer to 1 indicate a better fit.

$$R^2 = 1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}, \tag{2}$$

where $\overline{y}$ is the mean of the actual values.

The residuals are the differences between the actual values and the predicted values. They help to assess model accuracy and can be used to analyse patterns in errors.

$$\text{Residual} = y_i-\hat{y}_i. \tag{3}$$

In Random Forest Regression, the final prediction for each input is the average of predictions made by each individual tree in the forest. Formula for each prediction is:

$$\hat{y}=\frac{1}{T}\sum_{t=1}^{T}\hat{y}_t. \tag{4}$$

These formulas formed the backbone of our approach to assessing model accuracy and interpreting results.

3. Results

The scatter plot (Figure 1) shows the relationship between the actual and predicted particle concentrations after cleaning, with data points color-coded by dosage speed.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 December 2024 doi:10.20944/preprints202412.1023.v1

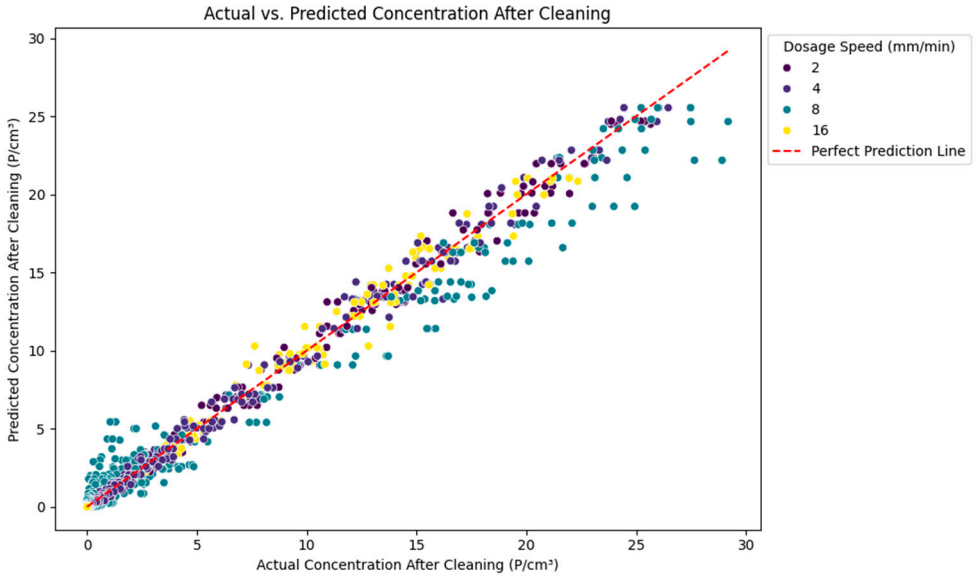


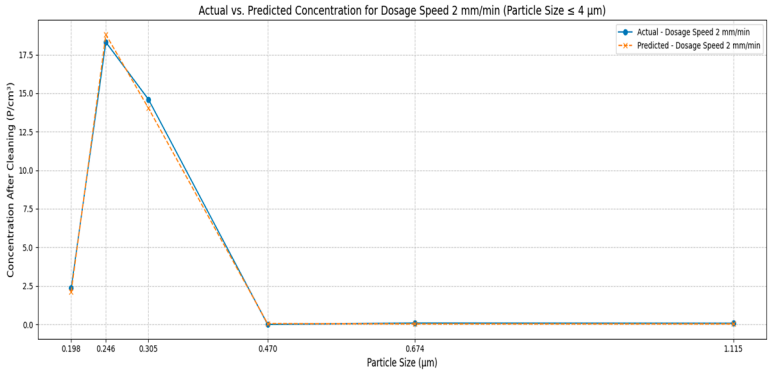
Figure 1. Actual vs. predicted concentration after cleaning.

Most points are tightly clustered around the red dashed line of perfect prediction ($y = x$). Each dosage speed (2, 4, 8, and 16 mm/min, or 7.69, 15.25, 30.76 and 61.51 mg/s of particulate matter per unit time) shows a similar clustering behaviour. For smaller dosage speeds (e.g., 2 mm/min and 4 mm/min), the predictions are almost indistinguishable from the actual values. At higher dosage speeds (e.g., 8 mm/min and 16 mm/min), there are slightly larger deviations for higher concentration values. As particle concentrations increase (> 20 P/cm^3), deviations from the perfect prediction line become more noticeable, particularly for the higher dosage speeds (yellow points).

Figure 2 shows a line plot of actual vs. predicted concentration by particle dosage speed. For all dosage speeds, the solid lines (actual) and dashed lines (predicted) closely overlap across the range of particle sizes.

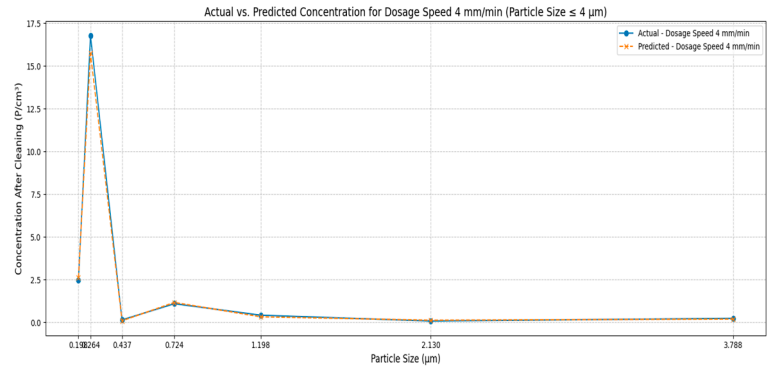
Concentration decreases steeply for particles smaller than 2 μm and stabilizes for larger particles (plots are not continued further). Larger particles are more efficiently removed by the cyclone section of the filter, leading to lower concentrations after cleaning. At higher speeds (8 mm/min and 16 mm/min), there is a slight separation between actual and predicted values for smaller particles.

Residuals by dosage speed plot (Figure 3) shows actual – predicted values are symmetrically distributed around zero, indicating no systematic bias in the model.

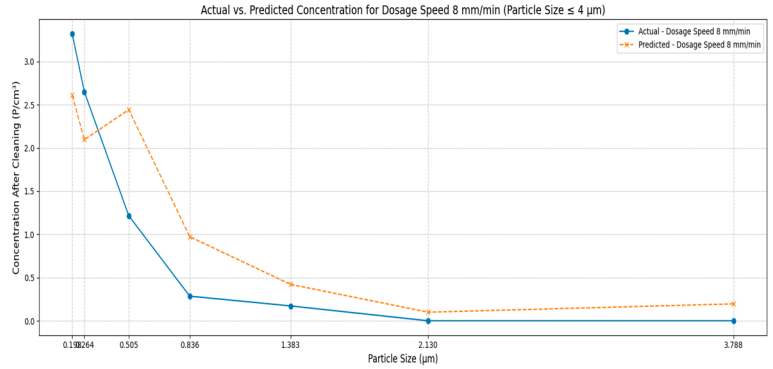


a)

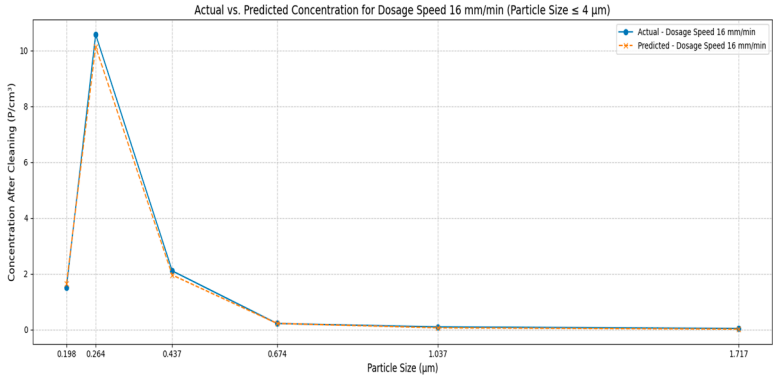
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 December 2024 doi:10.20944/preprints202412.1023.v1



b)



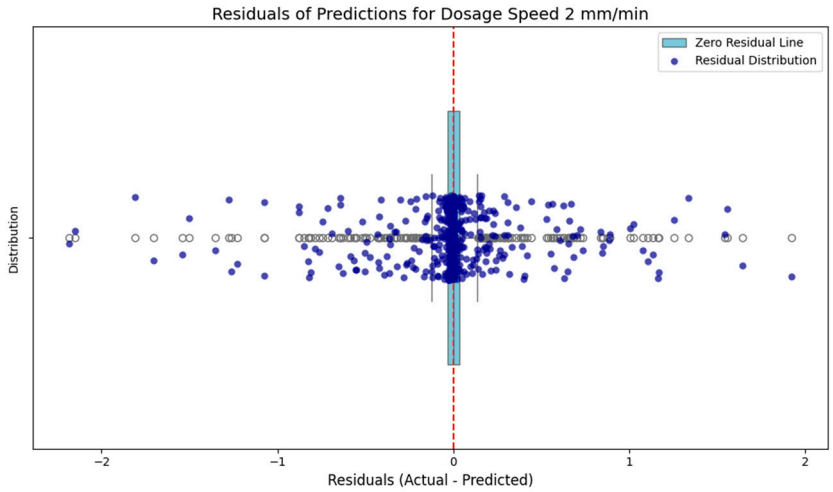
c)



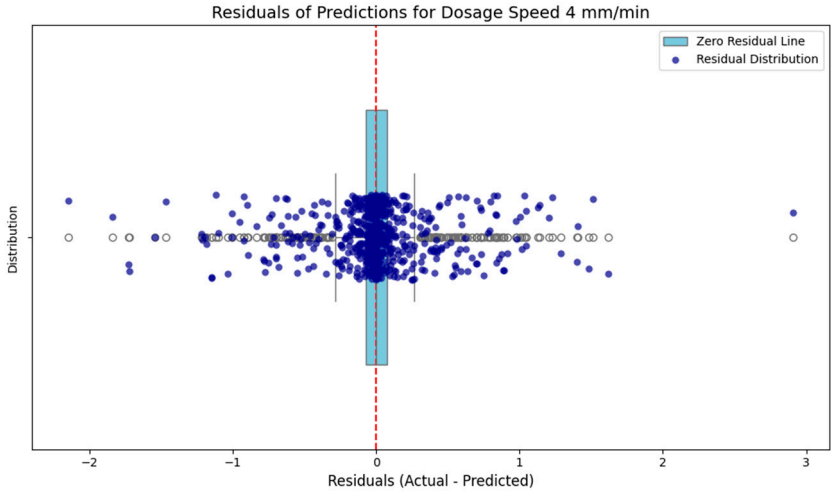
d)

Figure 2. Actual vs. predicted concentration after cleaning by particle dosage speed a) 2 mm/min; b) 4 mm/min; c) 8 mm/min; d) 16 mm/min.

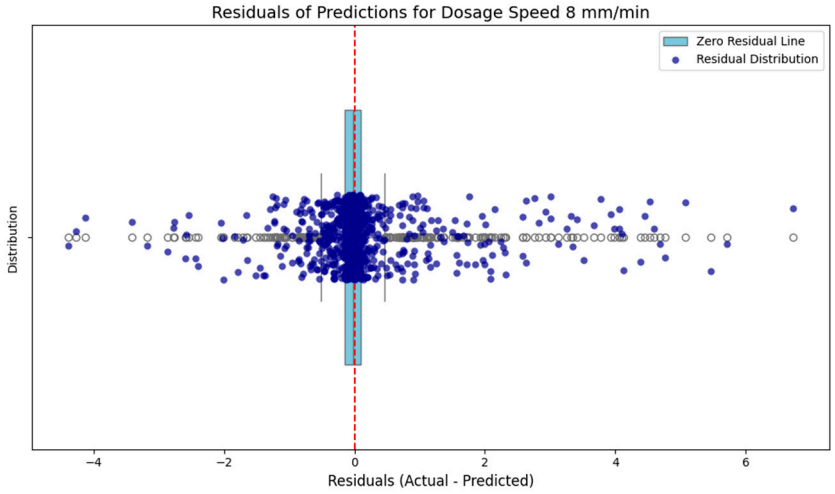
 doi:10.20944/preprints202412.1023.v1



a)

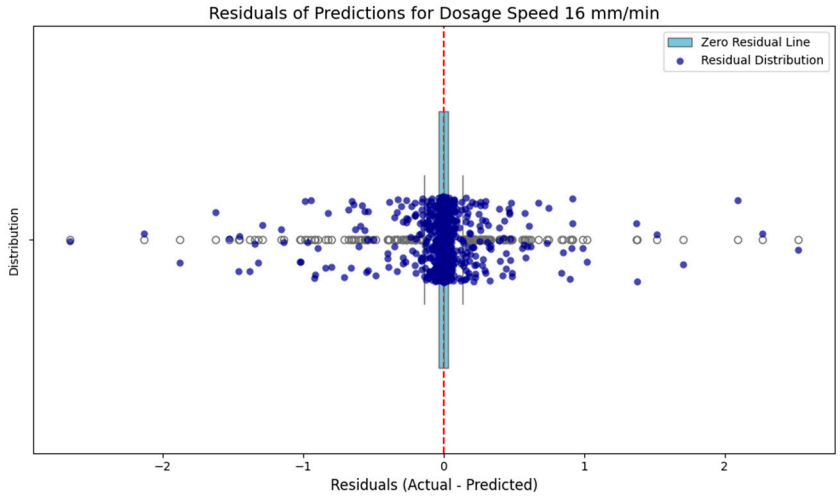


b)



c)

 doi:10.20944/preprints202412.1023.v1



d)

Figure 3. Residuals of predictions by dosage speed a) 2 mm/min; b) 4 mm/min; c) 8 mm/min; d) 16 mm/min.

Residuals for 8 mm/min show slightly larger variability. Outliers are more frequent for higher speeds (e.g., 8 mm/min and 16 mm/min), particularly for positive residuals (model underpredicting). Across all speeds, residuals are small relative to the scale of the particle concentrations (±2P/cm^3).

Model performance metrics (Table 2) showed that predictions deviate by an average of less than 0.5 particles per cubic centimetre from actual values. The R^2 value of 0.97.

Table 2. Model performance metrics.

Dosage Speed (mm/min)	MAE	R^2
2	0.196	0.995
4	0.232	0.994
8	0.491	0.967
16	0.156	0.993

4. Discussion

Since most points are tightly clustered around the perfect prediction line (Figure 1) and the slight deviations are distributed symmetrically, it means that there's no systematic underprediction or overprediction bias. It also indicates that the model accurately predicts the particle concentration after cleaning.

For smaller dosage speeds (e.g., 2 mm/min and 4 mm/min), the predictions are almost indistinguishable from the actual values. At higher dosage speeds (e.g., 8 mm/min and 16 mm/min), there are slightly larger deviations for higher concentration values, likely due to the greater complexity of particle interactions at these conditions.

Limitations in capturing dynamic behaviours of particle accumulation in the cyclone or measurement noise in experimental data may reflect the deviations from the perfect prediction line becoming more noticeable as particle concentrations increase. The high accuracy across dosage speeds and concentration ranges validates the generalizability of the Random Forest model. Small deviations at higher concentrations highlight areas where model refinement could improve precision, such as accounting for dust resuspension effects in the cyclone.

The solid lines (actual) and dashed lines (predicted) aligning closely across all dosage speeds and particle sizes indicates that the model accurately captures the filtration efficiency of the cyclone for particles of varying sizes. Concentration drops sharply for particles under 2 μm and stabilizes for larger ones, as the cyclone filter more effectively removes larger particles, resulting in lower post-cleaning concentrations. At higher speeds (8 mm/min and 16 mm/min), slight separation between

 doi:10.20944/preprints202412.1023.v1

actual and predicted values for smaller particles suggests more pronounced particle accumulation or dust resuspension effects. The model effectively captures particle concentration patterns across all speeds, demonstrating robustness in representing the electro cyclone filter's behaviour. Deviations at higher speeds and smaller particle sizes suggest potential for improvement through fine-tuning, such as accounting for particle size and flow rate interactions.

The residuals are symmetrically distributed around zero across all speeds, indicating no systematic bias in the model. Residual spread is consistent, though slightly larger at 8 mm/min due to its intermediate nature and complex particle behaviours. Outliers, more frequent at higher speeds, suggest unmodeled phenomena like dust resuspension. Overall, the low residual range (±2 P/cm³) confirms the model's strong predictive performance, with outliers offering insights for further refinement.

The model's MAE of 0.49 P/cm³ highlights its accuracy, with deviations averaging less than 0.5 P/cm³ across a concentration range up to 30 P/cm³. Its R^2 of 0.97 confirms the model captures 97% of variance, effectively reflecting complex interactions among particle size, dosage speed, voltage, and airflow.

Particles >2 μm are effectively filtered, resulting in lower post-cleaning concentrations, consistent with the cyclone filter's expected performance. Higher speeds (e.g., 16 mm/min) show increased residuals and deviations, likely due to dynamic effects like resuspension or turbulence. Concentration decreases with increasing particle size across all speeds, demonstrating the system's efficiency in capturing larger particles.

Addressing outliers at higher speeds by integrating factors like dust resuspension dynamics or advanced flow modelling could boost accuracy. Reducing experimental noise through preprocessing, such as filtering extreme data points, can further enhance model reliability.

The Random Forest model performs exceptionally well in predicting particle concentrations after cleaning, with strong alignment between actual and predicted values across different dosage speeds and particle sizes. The insights gained from the residuals and deviations provide actionable areas for refinement, but the overall performance metrics (MAE=0,49, R^2=0.97) confirm the model's robustness and reliability for practical use.

5. Conclusions

The Random Forest model demonstrated exceptional performance in predicting particle concentrations after cleaning, with strong alignment between actual and predicted values across various dosage speeds and particle sizes. Most data points closely followed the perfect prediction line (y=xy = xy=x), indicating minimal bias and high accuracy. The model's overall performance metrics, including an MAE of 0.49 P/cm³ and an R^2 of 0.97, reflect its ability to capture 97% of the variance in particle concentrations effectively.

Predictions were nearly indistinguishable from actual values at lower speeds (e.g., 2 mm/min and 4 mm/min), while slightly larger deviations occurred at higher speeds (e.g., 8 mm/min and 16 mm/min), particularly for smaller particles. These deviations suggest potential refinement opportunities, such as accounting for dynamic particle accumulation or dust resuspension effects.

The residuals were symmetrically distributed around zero across all dosage speeds, with low variability (±2P/cm³\pm 2 P/cm³±2P/cm³), confirming consistent and unbiased predictions. Larger particles (>2μm>2 \mu m>2μm) were filtered more efficiently, leading to lower concentrations after cleaning, consistent with the expected performance of the cyclone filter.

While the model excels in representing particle concentration patterns, addressing outliers at higher speeds through advanced modelling (e.g., incorporating resuspension dynamics) could further enhance accuracy. Overall, the model's robustness and reliability make it a valuable tool for predicting and optimizing electro cyclone filter performance in practical applications.

Author Contributions: Conceptualization, A. S., A. C., J. M., A.K. and D.M.; methodology, A. S; A. C., J. M., A.K. and D.M.; software, A. S., A. C., D.G., J. M. and A.K.; validation, A. C., J. M., A.K. and D.M.; formal analysis, A. S., A. C., J. M., D.G., and A.K.; investigation, A. S., A. C., J. M., D.G., A.K. , J. G. and D.M.; data curation, A. S., A. C., J. M. and A.K.; writing—original draft preparation, A. S., A. C., D.G., J. G., J. M. and A.K.; writing—review

 doi:10.20944/preprints202412.1023.v1

and editing,, A. C., J. M., A.K. and D.M.; visualization, A. C., J. M., and A.K.; supervision, J. M. and D. M.; project administration, J. M., A. K., A. C. and D.M.; funding acquisition, D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data Availability after request.

Acknowledgments: This research was supported by the center of excellence project „Civil Engineering Research Centre“ (Grant No. S-A-UEI-23-5). This project has received financial support from the Research Council of Lithuania (LMTLT), agreement No [S-MIP-24-88].

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The sample of the model training dataset (10 first rows) is shown below in Table A1:

Table A1. Sample of the model training dataset.

Particle size, X [μm]	Dosage Speed (mm/min)	Voltage (kV)	Air Flow Rate (m^3/s)	Particle concentration before air cleaning, dCnup [P/cm^3]	Particle concentration after air cleaning, dCndown [P/cm^3]	Efficiency, manually recalculated, %
0.198	2	2.5	255	4.307	2.345	45.55375
0.198	2	2.5	255	5.284	2.034	61.50643
0.198	2	2.5	255	5.087	1.413	72.22331
0.198	2	2.5	255	5.176	2.271	56.12442
0.198	2	2.5	255	5.484	2.04	62.80088
0.198	2	2.5	255	6.071	2.329	61.63729
0.198	2	2.5	349	5.897	1.88	68.11938
0.198	2	2.5	349	4.853	1.834	62.20894
0.198	2	2.5	349	4.933	2.016	59.13237
0.198	2	4	167	6.055	1.413	76.66391

References

1. Jaroń, A.; Borucka, A.; Deliś, P.; Sekrecka, A. An Assessment of the Possibility of Using Unmanned Aerial Vehicles to Identify and Map Air Pollution from Infrastructure Emissions. *Energies* **2024**, *17*, 577. https://doi.org/10.3390/en17030577.
2. Krzysiak, Z. Analiza Występowania Poważnych Awarii Przemysłowych w Polsce. *CHEMICAL REVIEW* **2021**, *1*, 56–59. https://doi.org/10.15199/62.2021.7.4.
3. Borucka, A.; Wiśniowski, P.; Mazurkiewicz, D.; Świderski, A. Laboratory Measurements of Vehicle Exhaust Emissions in Conditions Reproducing Real Traffic. *Measurement* **2021**, *174*, 108998. https://doi.org/10.1016/j.measurement.2021.108998.
4. Jonson, J.E.; Gauss, M.; Schulz, M.; Jalkanen, J.-P.; Fagerli, H. Effects of Global Ship Emissions on European Air Pollution Levels. *Atmos. Chem. Phys.* **2020**, *20*, 11399–11422. https://doi.org/10.5194/acp-20-11399-2020.
5. Lu, S.; Zhou, F. Impact of Ship Emission Control Area Policies on Port Air Quality—A Case Study of Ningbo Port, China. *Sustainability* **2024**, *16*, 3659. https://doi.org/10.3390/su16093659.
6. Zhang, X.; Van Der A, R.; Ding, J.; Zhang, X.; Yin, Y. The Impact of Inland Ship Emissions on Air Quality 2022.
7. Yoo, Y.; Moon, B.; Kim, T.-G. Estimation of Pollutant Emissions and Environmental Costs Caused by Ships at Port: A Case Study of Busan Port. *JMSE* **2022**, *10*, 648. https://doi.org/10.3390/jmse10050648.
8. Khan, D.; Burdzik, R. Measurement and Analysis of Transport Noise and Vibration: A Review of Techniques, Case Studies, and Future Directions. *Measurement* **2023**, *220*, 113354. https://doi.org/10.1016/j.measurement.2023.113354.

9. Khan, D.; Burdzik, R. A Review on Different Regulation for the Measurement of Transport Noise and Vibration. *J. meas. eng.* **2023**, *11*, 196–213. https://doi.org/10.21595/jme.2023.23279.
10. Kozłowski, E.; Borucka, A.; Oleszczuk, P.; Jałowiec, T. Evaluation of the Maintenance System Readiness Using the Semi-Markov Model Taking into Account Hidden Factors. *Eksploatacja i Niezawodność – Maintenance and Reliability* **2023**, *25*. https://doi.org/10.17531/ein/172857.
11. Žvirblis, T.; Hunicz, J.; Matijošius, J.; Rimkus, A.; Kilikevičius, A.; Gęca, M. Improving Diesel Engine Reliability Using an Optimal Prognostic Model to Predict Diesel Engine Emissions and Performance Using Pure Diesel and Hydrogenated Vegetable Oil. *Eksploatacja i Niezawodność – Maintenance and Reliability* **2023**, *25*. https://doi.org/10.17531/ein/174358.
12. Valavanidis, A. Airborne Particulate Matter (PM) as a Significant Air Pollution Risk Factor for Human Health. New Research on Adverse Health Effects of the Respiratory and Cardiovascular Systems. **2023**, *1*, 1–25.
13. Verykioy, M.; Sardi, C.; Skanavis, C. HEALTH EFFECTS OF THE INCREASE IN PARTICULATE MATTER OF PM2.5 AND PM10 DIAMETER IN THE AREA OF VOLOS. In; 2022; pp. 41–64 ISBN 978-960-08-0924-4.
14. Mahalingam, S.; Ramsundram, N.; Sathyamoorthy, G.L. Particulate Matter and Its Health Effects - A Review. **2020**, 2582–5208.
15. Lim, E.Y.; Kim, G.-D. Particulate Matter-Induced Emerging Health Effects Associated with Oxidative Stress and Inflammation. *Antioxidants* **2024**, *13*, 1256. https://doi.org/10.3390/antiox13101256.
16. Garcia, A.; Santa-Helena, E.; De Falco, A.; De Paula Ribeiro, J.; Gioda, A.; Gioda, C.R. Toxicological Effects of Fine Particulate Matter (PM2.5): Health Risks and Associated Systemic Injuries—Systematic Review. *Water Air Soil Pollut* **2023**, *234*, 346. https://doi.org/10.1007/s11270-023-06278-9.
17. Meo, S.A.; Shaikh, N.; Alotaibi, M.; AlWabel, A.A.; Alqumaidi, H. Effect of Air Pollutants Particulate Matter (PM2.5, PM10), Sulfur Dioxide (SO2) and Ozone (O3) on Cognitive Health. *Sci Rep* **2024**, *14*, 19616. https://doi.org/10.1038/s41598-024-70646-6.
18. Shin, D.C. Health Effects of Ambient Particulate Matter. *J Korean Med Assoc* **2007**, *50*, 175. https://doi.org/10.5124/jkma.2007.50.2.175.
19. Bae, S.; Hong, Y.-C. Health effects of particulate matter. *J Korean Med Assoc* **2018**, *61*, 749. https://doi.org/10.5124/jkma.2018.61.12.749.
20. Wang, P.; Liu, J.; Wang, C.; Zhang, Z.; Li, J. A Holistic Performance Assessment of Duct-Type Electrostatic Precipitators. *Journal of Cleaner Production* **2022**, *357*, 131997. https://doi.org/10.1016/j.jclepro.2022.131997.
21. Wen, T.-Y.; Su, J.-L. Corona Discharge Characteristics of Cylindrical Electrodes in a Two-Stage Electrostatic Precipitator. *Heliyon* **2020**, *6*, e03334. https://doi.org/10.1016/j.heliyon.2020.e03334.
22. Feng, Z.; Cao, S.-J.; Long, Z. Determining V-I Characteristics of Energy-Efficient Electrostatic Assisted Air Filtration System by Utilizing the Back-Corona Induced Current Model. *Energy and Built Environment* **2023**, *4*, 227–235. https://doi.org/10.1016/j.enbenv.2021.11.004.
23. Li, J.; Sun, Q.; Ping, Z.; Gao, Y.; Chen, P.; Huang, F. Electric Field-Driven Air Purification Filter for High Efficiency Removal of PM2.5 and SO2: Local Electric Field Induction and External Electric Field Enhancement. *Atmosphere* **2022**, *13*, 1260. https://doi.org/10.3390/atmos13081260.
24. Kaczmarzyk, P.; Warguła, Ł.; Janik, P.; Krawiec, P.; Bąk, D.; Klapsa, W. Problems of Operation of Positive Pressure Ventilators on the Basis of Surveys of Polish Officers of the State Fire Service. *Sci Rep* **2024**, *14*, 10795. https://doi.org/10.1038/s41598-024-61507-3.
25. Wieczorek, B.; Kaczmarzyk, P.; Warguła, Ł.; Giedrowicz, M.; Bąk, D.; Gierz, Ł.; Stambolov, G.; Kostov, B. Research on the Distribution of Axial Excitation of Positive Pressure Ventilators in the Aspect of Stability Safety of the Load-Bearing Frame. *Adv. Sci. Technol. Res. J.* **2024**, *18*, 142–154. https://doi.org/10.12913/22998624/174845.
26. Muzafarov, S.; Balitsky, V.; Babayev, A.; Kilichov, O.; Ruzikulov, Z.; Tashkhodzheva, G.; Rakhmataliev, M. Influence of Temperature and Humidity of the Air on Discharge Processes in Electric Filters. *E3S Web Conf.* **2024**, *563*, 01019. https://doi.org/10.1051/e3sconf/202456301019.
27. Buczaj, A. Bezpieczeństwo Podczas Przewożenia Chemicznych Towarów Niebezpiecznych. *CHEMICAL REVIEW* **2019**, *1*, 106–110. https://doi.org/10.15199/62.2019.8.13.
28. Jaworek, A.; Sobczyk, A.T.; Marchewicz, A.; Krupa, A.; Czech, T.; Śliwiński, Ł. Hybrid Electrostatic Filtration Systems for Exhaust Gas Cleaning. *J. Phys.: Conf. Ser.* **2024**, *2702*, 012009. https://doi.org/10.1088/1742-6596/2702/1/012009.
29. Li, H.; Mou, H.; Pan, P.; Zhou, X.; Wang, Z. Design of the Electrode Structure for the Dust Removal Efficiency Optimization of Electrostatic Precipitator. *Phys. Scr.* **2024**, *99*, 125013. https://doi.org/10.1088/1402-4896/ad8af7.
30. Ning, L.; Wang, N.; Fu, J.; Ma, Y.; Gu, S.; Cheng, M. Effect of High-Voltage Electrostatic Precipitator Dust Collection Plate Structure on Collection Efficiency. *ACS Omega* **2024**, *9*, 37396–37407. https://doi.org/10.1021/acsomega.4c06500.
31. Iváncsy, T.; Kiss, I.; Suda, J.; Berta, I. EFFICIENCY OF THE PRECIPITATION OF FINE PARTICLES INFLUENCED BY THE ESP SUPPLY MODE.

32. Kiss, I.; Iváncsy, T.; Suda, J.M.; Berta, I. Some Aspects of "Long Term" Modelling of Electrostatic Precipitators.
33. Chung, J.H.; Kim, K.H.; Sohn, D.K.; Ko, H.S. Numerical Study on Staggered Needle-Plate Electrode Configuration for Improved Particulate Matter Collection in Electrostatic Precipitator. *Powder Technology* **2024**, *443*, 119865. https://doi.org/10.1016/j.powtec.2024.119865.
34. Czech, T.; Marchewicz, A.; Krupa, A.; Sobczyk, A.T.; Turzyński, T.; Gazda, M.; Jaworek, A. Properties of PM and Gaseous Products Generated by Low-Power Domestic Boiler after Wood Combustion and Removed by Electrostatic Precipitator. *Journal of Electrostatics* **2024**, *129*, 103931. https://doi.org/10.1016/j.elstat.2024.103931.
35. Li, J.; Duan, L.; Chen, J.; Li, D.; Bao, S.; Wang, Z.; Wang, J.; Liao, J. Research of the Effect of Different Corrugated Dust Collection Plates on Particle Removal in Electrostatic Precipitators. *Chemical Engineering Research and Design* **2023**, *197*, 323–333. https://doi.org/10.1016/j.cherd.2023.07.006.
36. Chen, B.; Guo, Y.; Li, H.; Zhou, W.; Liu, B. Discharge Characteristic of Barbed Electrodes in Electrostatic Precipitator. *Journal of Electrostatics* **2021**, *109*, 103528. https://doi.org/10.1016/j.elstat.2020.103528.
37. Duan, L.; Wang, J.; Huang, Q.; Xia, S. Experimental Investigation on the Performance of Hybrid Electrostatic-Fabric Precipitators with Different Structures. *Powder Technology* **2023**, *421*, 118404. https://doi.org/10.1016/j.powtec.2023.118404.
38. Wieczorek, A.N.; Konieczny, Ł.; Burdzik, R.; Wojnar, G.; Filipowicz, K.; Kuczaj, M. A Complex Vibration Analysis of a Drive System Equipped with an Innovative Prototype of a Flexible Torsion Clutch as an Element of Pre-Implementation Testing. *Sensors* **2022**, *22*, 2183. https://doi.org/10.3390/s22062183.
39. Šabanovič, A.; Matijošius, J.; Kilikevičius, A.; Chlebnikovas, A.; Steišūnas, S. Evaluation of Aerosol Filtration Efficiency in Low-Voltage Systems for Naval Applications. *8th Georgian-Polish International Scientific Conference "Transport Bridge Europe-Asia"* **2024**, 213–220.
40. Król, K. Random Forest Method to Identify Seepage in Flood Embankments. *ELECTROTECHNICAL REVIEW* **2022**, *1*, 193–196. https://doi.org/10.15199/48.2022.02.44.
41. Kozłowski, E.; Borucka, A.; Świderski, A.; Skoczyński, P. Classification Trees in the Assessment of the Road–Railway Accidents Mortality. *Energies* **2021**, *14*, 3462. https://doi.org/10.3390/en14123462.
42. Stewart, I.; Tall, D. *The Foundations of Mathematics*; 2. edition.; Oxford University Press: Oxford, 2015; ISBN 978-0-19-870643-4.