applied sciences

MDPI

Article

Modeling Indoor Particulate Matter and Small Ion Concentration Relationship—A Comparison of a Balance Equation Approach and Data Driven Approach

Miloš Davidović [1], Milena Davidović [2], Rastko Jovanović [1], Predrag Kolarž [3], Milena Jovašević-Stojanović [1] and Zoran Ristovski [4,*]

[1] VINČA Institute of Nuclear Sciences, National Institute of the Republic of Serbia, University of Belgrade, 11351 Belgrade, Serbia; davidovic@vin.bg.ac.rs (M.D.); virrast@vin.bg.ac.rs (R.J.); mjovst@vin.bg.ac.rs (M.J.-S.)

[2] Faculty of Civil Engineering, University of Belgrade, 11120 Belgrade, Serbia; milena@grf.bg.ac.rs

[3] Institute of Physics, University of Belgrade, 11080 Belgrade, Serbia; kolarz@ipb.ac.rs

[4] School of Earth and Atmospheric Sciences, Queensland University of Technology, 4000 Brisbane, Australia

* Correspondence: z.ristovski@qut.edu.au

Received: 17 July 2020; Accepted: 24 August 2020; Published: 27 August 2020

Featured Application: An ANN model successfully helped in harmonizing inputs from several instruments of different grade (low cost radon and lab grade particulate matter monitors) and enabled predictions of small ions concentration of comparable quality to the lab grade Gerdien type instrument.

Abstract: In this work we explore the relationship between particulate matter (PM) and small ion (SI) concentration in a typical indoor elementary school environment. A range of important air quality parameters (radon, PM, SI, temperature, humidity) were measured in two elementary schools located in urban background and suburban area in Belgrade city, Serbia. We focus on an interplay between concentrations of radon, small ions (SI) and particulate matter (PM) and for this purpose, we utilize two approaches. The first approach is based on a balance equation which is used to derive approximate relation between concentration of small ions and particulate matter. The form of the obtained relation suggests physics based linear regression modelling. The second approach is more data driven and utilizes machine learning techniques, and in this approach, we develop a more complex statistical model. This paper attempts to put together these two methods into a practical statistical modelling approach that would be more useful than either approach alone. The artificial neural network model enabled prediction of small ion concentration based on radon and particulate matter measurements. Models achieved median absolute error of about 40 ions/cm^3 and explained variance of about 0.7. This could potentially enable more simple measurement campaigns, where a smaller number of parameters would be measured, but still allowing for similar insights.

Keywords: indoor air quality; small ions; radon; particulate matter; linear regression; artificial neural networks

1. Introduction

Our health and wellbeing is a complex and multifaceted phenomenon, but clean air can be with certainty regarded as one of its most critical components. Not surprisingly, it has been shown that air pollution is the single largest environmental health risk in Europe [1]. While level of concentration of air pollutants can widely vary even locally, the reactions people may have in response to exposure even to the same level of air pollutants concentration can additionally vary due to the breathing volume (e.g., because of different age and levels of physical fitness and activity) and duration of exposure (e.g., large amount of time spent indoors, as a commuter, etc.). Furthermore, some age groups have behavioral patterns that may affect their exposure in a negative manner, such as elderly or young children, and result in various negative health effects including asthma, allergies, and other [2,3].

Monitoring of outdoor ambient air quality is usually done via networks at national and local level. However, despite having high quality of instrumentation these kinds of networks are usually very sparse, may not monitor all parameters of interest, and give little insight into personal exposure, for which indoor air quality may be of greater importance. This makes additional measuring campaigns necessary, especially for indoor places where sensitive age groups may spend considerable amount of time. The level and composition of air pollutant differs indoors and outdoors, and some air pollutants can be more prominent outdoors (e.g., gaseous pollutants sulfur dioxide or ozone), while others are usually more dominant in the indoor environment (e.g., formaldehyde, carbon monoxide or nitrogen oxides and radon) [4,5]. Even more so than the gas phase pollutants, PM concentration along with its size distribution and chemical composition is also a significant problem. It is well documented that serious health effects may result from long-term exposure to an elevated concentration of particulate matter [6]. Short term air pollution levels were found to correlate with reduced lung function, and remained visible up to 24 h after exposure [7]. In many European cities [6], particulate matter concentration is two to three times higher outdoors than that recommended by the World Health Organization (WHO). Outdoor air also contributes to indoor air quality, since it can diffuse easily into indoor environment. Air quality indoors can be further worsened by non-satisfactory ventilation quality.

In this paper we focus on a smaller subset of indoor air pollution phenomena with a focus on an interplay between primary pollutants radon and particulate matter (PM) [8] and small ions (SI) concentrations. Reasoning and arguments found in scientific literature behind the question why small air ions may have an impact on human health are the following. Small ions can be considered as natural air cleaners and sterilizers, and also biologically active constituents of the environmental air. Additionally, process of ion to aerosol attachment is leading to aggregation of ultrafine particles (UFP) in environmental air and thus reduces their number concentration (at the account of mass gain) and deposition on electrostatic surfaces. Recent scientific research shows that the health hazard possibly increases with the decrease of diameter of the inhaled particles. Peters et al. [9] demonstrated that the number concentration of nanoparticles is more strongly associated with health effects than the mass concentration. Other health impacts of air ions include psychological effects that have been reported in many studies and summarized in Perez et al. [10] and Pino and Ragione [11]. Most of the studies claimed that ionization was significantly associated with lower depression ratings. It is important to note that most of research refers to high levels of ion concentration exposure. Studies of background ion concentration relative to the ion-free state studies were not found but health benefits of rich ion background could be expected. Jiang et al. [12] claim that the reports where the presence of negative ions is credited for increasing psychological health are without reliable evidence in therapeutic practice. The studies showing that negative ions could help people with symptoms of allergies to dust, mold spores, and other allergens need additional confirmations. However, it is encouraging that there are no known negative effects of negative ions, so positive effect should be studied further. On the other hand, adverse health effects of radon [13] and particulate matter [14] are long well known. Since all three quantities (radon, SI and PM concentration) are linked via a balance equation, even though the health effects are only firmly established for radon and PM [8], it is useful to consider all three simultaneously. Understanding the way in which various air quality variables interact can be beneficial for developing predictive models, and also enable obtaining more knowledge about air quality based only on several key predictor measurements. Children spend a large part of their time at school microenvironment. In the last decade several large studies conducted within the framework of European projects BREATHE and SINPHONIE addressed the topics of level and chemical composition of particulate matter fractions including ultrafine particulate matter and gaseous air pollutants, also addressing differences during teaching and non-teaching hours and periods when there are no occupants in schools [15,16]. In the framework of numerous studies at national and international level radon concentration was measured in school classrooms, usually by utilizing passive samplers.

The objectives of this paper are the following. The first objective was to bring additional insight into indoor air quality by measuring a number of important air quality variables including quantities that are somewhat more rarely measured with higher temporal resolution namely SI concentration and radon concentration in two elementary schools' indoor environments. In addition, we have measured PM concentration and size with high temporal and size resolution. Descriptive statistics of the measured parameters is presented and discussed. We then proceed to studying association between SI, radon and PM concentrations, based on a balance equation. Parameters of the balance equation are estimated from the data in two schools. We investigate the hypothesis that small ion concentration can be predicted based on radon and particulate matter measurements predictors, by using artificial neural network model. If successful, this kind of modelling effort could enable obtaining more knowledge about air quality based only on several key predictor measurements. This could potentially enable more simple measurement campaigns, where a smaller number of parameters would be measured, but still allowing for similar insights. The outline of this paper is as follows. First, we provide a brief explanation of the physical processes involved in creation of small ions, and then utilize balance equation to describe it quantitatively. We then discuss the link between small ion concentration, volumetric production rate and particulate matter concentration. Following this discussion, we describe the method that was used in the indoor measuring campaign, in which a number of relevant parameters appearing in the balance equation are measured either directly or via an important proxy. We study an interplay between concentrations of radon, SI and PM. For this purpose, we utilize two approaches. The first approach is based on a balance equation which is used to derive approximate relation between concentration of small ions and particulate matter. The form of the obtained relation is transformed via Taylor expansion to enable meaningful linear regression modelling with several predictors. The second approach is data driven and utilizes machine learning techniques, namely shallow feed forward neural networks, and in this approach, we develop a more complex statistical model, but utilize predictors that were used in physics based linear regression modelling. Performance and trade-offs of the two approaches are then discussed.

2. Materials and Methods

2.1. Form of the Balance Equation Suitable for Statistical Modelling

The small air ion concentration ($n_{\pm}$) is determined by the following balance equation:

$$\frac{dn_{\pm}}{dt} = q - \alpha n_{\pm} n_{\mp} - n_{\pm} \beta Z \tag{1}$$

where q is the volumetric production rate, Z is the aerosol number concentration, α coefficient accounts for the losses of ion-to-ion recombination and β represents an effective ion-aerosol attachment coefficient, which is the integral over the size distribution of aerosol particles. The balance equation can include additional terms. If electrostatic deposition (occurring mainly in indoor air) is included in a model, there is an additional right-hand side term $-\delta^{\pm} n^{\pm}$, where $\delta^{\pm}$ is an electrostatic deposition rate coefficient of the air ions. Additional details about physics behind changes in SI concentration and relevant terms in balance equations are given in Appendix A.

While Equation (1) seems very intuitive, it is worth noting that it was discovered rather painstakingly, and some terms were added to increase its scope of validity. Namely, in the first half of last century Schweidler [17] showed that the quadratic law of recombination previously held valid (quadratic law is Equation (1) without the last (linear) term on the right-hand side. Since $n_{+} \approx n_{-}$, equality being only approximate due to the different mobility of positive and negative small ions, the middle term in Equation (1) is a quadratic term) is not valid in ordinary air, but only for clean air (Equation (1) when $Z \approx 0$). As another interesting historical note ion-aerosol attachment coefficient was previously referred to as "diminution coefficient of small ions in the presence of nuclei and large ions" [18].

If we neglect the quadratic term present in (1) we can obtain the following (under the assumption of a quasi-steady state):

$$\frac{dn^-}{dt} \approx 0 \approx q - n^-(\beta Z + \delta^-) \tag{2}$$

(since our campaign was situated indoors, the electrostatic deposition rate coefficient δ^- was also included). After expanding βZ term we obtain:

$$n^- \approx \frac{q}{\sum \beta_i Z_i + \delta^-} \tag{3}$$

or more conveniently

$$\sum \beta_i Z_i + \delta^- \approx \frac{q}{n^-} \tag{4}$$

The form of the above equation suggests linear regression is a justified modelling approach if we want to model the interdependence between the concentration of larger aerosol particles of various diameter and small ion concentration. The physical meaning of the coefficients in linear regression (β_i) are an ion-aerosol attachment coefficient and the intercept term corresponds to the electrostatic deposition rate coefficient of the air ions (δ^-). Note, however, that in a non-laboratory type of campaign, one cannot precisely control the aerosol distribution and there may be a significant correlation between individual channels corresponding to different particle sizes, which makes calculation (and interpretation) of the regression coefficients as attachment coefficients largely approximate.

Taking one more look into (4) brings up an important issue, that is taking a quotient of two noisy variables on the right-hand side of the Equation (4), that may also have values close to zero. The right hand side of (4) therefore may produce a quotient not suitable for further statistical modelling. Instead of using Equation (4) directly, we will do the following. By Taylor expanding expression (3) around some value of volumetric production rate q_0 and some value of particle concentrations Z_{i0} we obtain

$$\begin{aligned} n^- &\approx \frac{q}{\sum \beta_i Z_i + \delta^-} = f(q, Z_i) \approx \\ &\approx f(q_0, Z_{i0}) + \frac{\partial f(q,Z_i)}{\partial q}\Big|_{q_0,Z_{i0}}(q - q_0) + \frac{\partial f(q,Z_i)}{\partial Z_i}\Big|_{q_0,Z_{i0}}(Z_i - Z_{i0}) \end{aligned} \tag{5}$$

The derivatives in (5) are given by

$$\begin{aligned} \frac{\partial f(q,Z_i)}{\partial q} &= \frac{1}{\sum \beta_i Z_i + \delta^-}\Big|_{q_0,Z_{i0}} = \frac{1}{\sum \beta_i Z_{i0} + \delta^-} = \frac{n_0}{q_0} \\ \frac{\partial f(q,Z_i)}{\partial Z_i} &= \frac{-q}{(\sum \beta_i Z_i + \delta^-)^2}\beta_i\Big|_{q_0,Z_{i0}} = \frac{-q_0}{(\sum \beta_i Z_{i0} + \delta^-)^2}\beta_i = \frac{-{n_0}^2}{q_0}\beta_i \end{aligned}$$

And finally, we arrive at:

$$n^- \approx n_0 + \frac{n_0}{q_0}(q - q_0) - \sum \frac{{n_0}^2}{q_0}\beta_i(Z_i - Z_{i0}) \tag{6}$$

This equation is more suitable for developing linear regression model, since it doesn't include quotients of the noisy occasionally close-to-zero random variables as does Equation (4). If a linear regression model is developed based on (6) it would allow us to model and predict concentration of small ions based on the knowledge of rate of volumetric production rate (in this paper we use radon concentration as a proxy for volumetric production rate) and knowledge of particulate matter concentration. Additionally, it will allow us to estimate attachment coefficients.

2.2. Description of the Statistical Modeling Methodology

The first modelling approach that we will utilize is linear regression, with several input predictors (radon and aggregated particle channels) and one target variable (small ion concentration). General form of linear regression equation with several input predictors and one target is given by

$$y_i = a_0 + a_1 \cdot x_{i1} + \ldots + a_p \cdot x_{ip} + z_i \tag{7}$$

where p is number of predictors, i is number of points we use for fitting model parameters $a_0, a_1, \ldots a_p$, y_i is sample of target, $x_{i1}, x_{i2} \ldots x_{ip}$ are samples of predictors and z_i represents noise in i-th sample. Model parameters are determined by ordinary least squares i.e., by minimizing $\sum {z_i}^2$(summed over all samples). Input was transformed by subtracting median, since it is a more robust statistical measure than mean. Thus, the linear model built around Equation (6) can be considered as Taylor expansion around median. Note that particular linear scaling of input has no effect on R^2 score of the linear model, but serves to aid interpretation of coefficients in Equation (6).

Second approach we utilize for statistical modeling is a simple feed-forward artificial neural network (ANN) with one hidden layer. Such shallow feed-forward ANNs were previously used successfully in a number of contexts, and recently for calibration of low-cost sensors [19]. Input was scaled using standard normalization scaling (transforming it to zero mean and unit standard deviation). The hidden layer uses rectified linear unit (ReLU) activation function, which is known to have certain benefits such as non-vanishing gradient compared to commonly used, sigmoid transfer function. Implementation of the network was done via software library scikit-learn [20]. Network optimization procedure is the default 'adam' solver used in multi-layer perceptron (MLP)-regressor in scikit-learn, and further details can be found in [21]. Class MLPregressor implements a multi-layer perceptron (MLP) with no activation function in the output layer. It uses the squared error as the loss function, and the output is a set of continuous values. Model selection criterion was based on R2 score, and two requirements: to have as high as possible R2 score, and also a balanced result on training and test set. These criterions were examined via parametric sweep of number of input PCA components and neurons in the hidden layer. Since it is not generally possible to prescribe physical meaning to model parameters in the ANN models we will only compare the predictive power of the models.

2.3. Description of the Experimental Setup

A wide range of relevant air quality parameters were measured in indoor environments of two elementary schools located far apart 20 km, namely: an elementary School "Aleksa Šantić" in Belgrade suburb Kaluđerica, and in an elementary school "20. oktobar" in a residential background of the New Belgrade municipality, referred to as School 1 and School 2 in further text, respectively. The measurement campaigns in both schools, that have natural ventilation, were conducted during March 2017. Measuring spots in both schools were in classrooms on the ground floor, occupied on workdays. Instruments were arranged rather densely, and we can assume that the sampled air was well-mixed. Area surrounding location of School 1 and 2 is depicted in Figure 1. School 1 is in the near vicinity of a major local road, and also in the neighborhood where there is a significant amount of domestic heating sources. School 1 itself has a coal-based heating system. School 2 on the other hand is located further apart from the major roads, and is situated in a block of buildings connected to the district heating system.

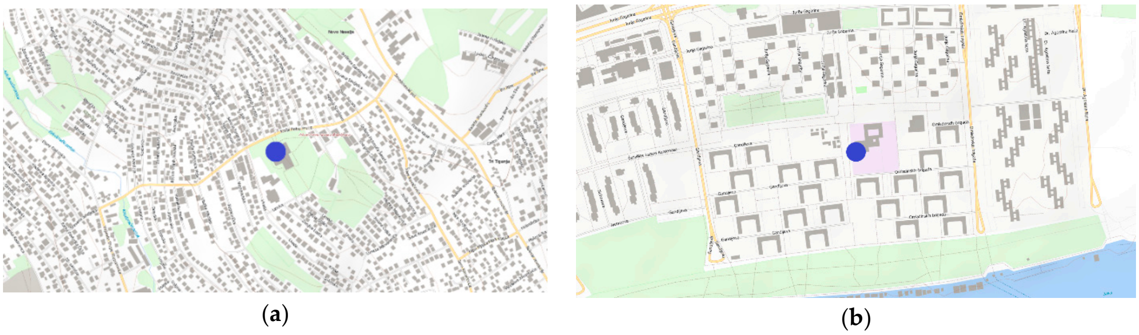


Figure 1. Larger area surrounding location of (**a**) School 1 (WGS84 20.556337261, 44.764855909) and (**b**) School 2 (WGS84 20.395445055, 44.799510267). Approximate location of indoor school space is marked as blue dot.

The measurement instrumentation setup included the following. SI concentration was measured using a Gerdien-type air ion detector [22]. PM concentrations for diameter of particles going from 10 nm to 420 nm in 13 size channels were detected using TSI NanoScan SMPS Model 3910, and PM concentrations for diameter of particles going from 0.3 µm to 10 µm in 16 channels were recorded using TSI Optical particle sizer 3330. In addition, the radon concentration level was measured hourly using a Radon Scout, along with the local temperature, pressure and humidity. The Radon Scout consists of a measurement chamber with high voltage collection and Si detector, and samples air by diffusion. Remaining instruments sample air actively, either via a pump (particle sizers) or a fan (air ion detector). Accuracy, traceability and manufacturer info of the used instruments if further described in Appendix B and Table A1.

The collected data describes all relevant processes quantified by a balance equation: 2-minute SI concentration measurements describe steady state; radon concentration gives insight into the rate of volumetric ion pair generation and 1-minute PM measurements give insight into main loss mechanism for SI.

3. Results

Based on the descriptive statistics given in Table 1, some remarks can be made. In School 1, both mean and median SI concentration were more than twice as large as in School 2. On the other hand, radon median concentration was similar in two schools, although the mean was higher in School 2. Looking at the particle concentration, we can observe that while 0.3–10 µm diameter particles were similar in concentration in both schools, this is not the case for 10–420 nm particles. This could explain lower SI concentration in School 2, since increased concentration of nanometer sized particles could explain main loss mechanism. While it is desirable to have sensitive indoor environments properly ventilated, this was not the case in the two schools where the experiments were conducted, which have natural ventilation. Due to this, radon and particulate matter concentration exhibited strong diurnal variations in both schools. Temperature had somewhat extreme values for indoor space (min and max values in Table 1), however, the standard deviation was less than 2 °C. Relative humidity had similar values of mean and median (suggesting no significant outliers) at both schools of about 30%, and standard deviation of about 6%. Particulate matter sizers flag data points that are not accurate due to some problem (e.g., low level of working fluid, or out of range inlet flow) and these data points were not taken into consideration. Table 1 also serves as a reminder on what is the scope of the validity of the predictive models and associated model quality estimates that will be developed and discussed in further text.

Table 1. Summary statistics (minimum, maximum, mean, median and standard deviation) of air quality variables for the campaign conducted in Schools 1 (S1) and 2 (S2).

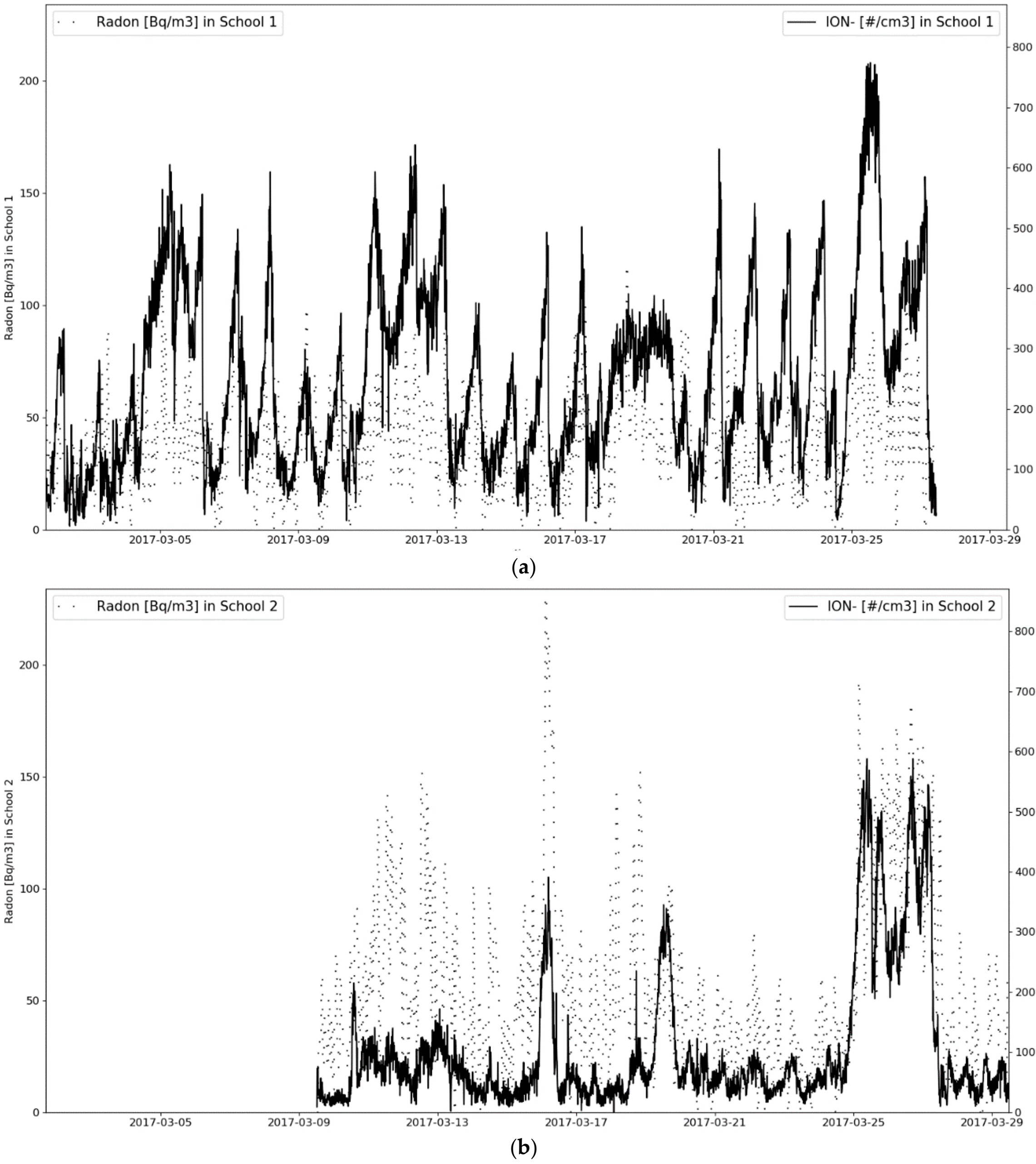
	Negative Small Ions [#/cm^3]		**Radon [Bq/m^3]**		**Particle conc. 10–420 nm [#/cm^3]**		**Particle conc. 0.3–10 um [#/cm^3]**		**Pressure (atm)**		**t [°C]**		**RH [%]**	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
Min	0.0	0.0	0.0	0.0	0.0	776.5	53.2	30.4	0.97	0.98	19.60	20.12	18.97	18.87
Max	871.0	643.0	118.0	234.0	95,023.6	116,127.2	541.9	572.6	1.00	1.01	28.81	32.62	46.14	49.18
Mean	239.9	104.8	36.8	56.5	2512.7	15,198.2	161.1	151.1	0.98	1.00	24.24	26.64	29.77	31.20
Med.	212.0	63.0	39.0	41.0	1577.5	3724.6	142.7	132.8	0.98	1.00	24.27	26.78	29.40	30.81
Std. dev	143.8	113.9	22.2	41.6	4283.2	31,759.4	82.3	81.3	0.01	0.01	1.80	2.00	6.23	5.37

Initial exploratory data analysis shows that there is a correspondence between radon concentration and small ion concentration, shown in Figure 2 for both schools. Note that radon measurements appear to be more noisy compared to the small ion measurements, despite having much larger sampling time of 1h. This could be attributed to sampling mechanism of Radon Scout, since it samples air via diffusion. It also appears that for some periods of time variation in radon and small ions is "matched", e.g., School 2, last week of March, where there is low concentration of radon and ions during workdays, and a sudden increase during weekend (25 and 26 March 2017). Similar, but less pronounced effect is seen in School 1 in the same week. This is probably due to different regimes of window opening and general use of indoor space during weekends. However, there are also periods where a "mismatch" between radon and SI concentrations seems to happen, probably due to loss mechanisms of attachment of small ions to particulate matter (see for example 9th of March in School 1, with a large spike in radon concentration not matched by ion concentration). In the following text we will see whether simple linear regression suggested by (6) could explain these and similar situations, and to what extent.

But before doing that, let us consider a correlation matrix of the wide array of quantities that we have measured simultaneously in two schools. Correlation plots are given in Figure 3, and include TSI Nanoscan channels (denoted in Figure 3 by channel size 11.5, 15.4, ... , 273.8 nm, concluding with total concentration), followed by TSI OPS channels (denoted in Figure 2 by 0.337, 0.419, ... , 9.015 μm, concluding with total concentration), followed by meteorological parameters, and finally small ions and radon. Notice that in both schools SI concentration negatively correlates with all particle channels (Pearson correlation coefficient is −0.36 for total particle concentration measured by TSI Nanoscan and −0.39 for total particle concentration measured by TSI OPS in School 1, and −0.09 and −0.37 in School 2), and that it positively correlates with radon (Pearson correlation coefficient is 0.33 and 0.59 in Schools 1 and Schools 2 respectively). This coincides with the conclusions that can be made from Equations (3), (4) and (6), and therefore it is justified to use these equations as a starting point for development of statistical models.

It is clear looking at TSI Nanoscan and OPS channels in the correlation matrix, that larger groups of channels correlate and thus it wouldn't be useful to consider all channels as independent predictors (signals) in the framework of regression modelling due to multicollinearity effects. We have, therefore, reduced the number of predictors for particulate matter we use in the modelling to the most significant ones.

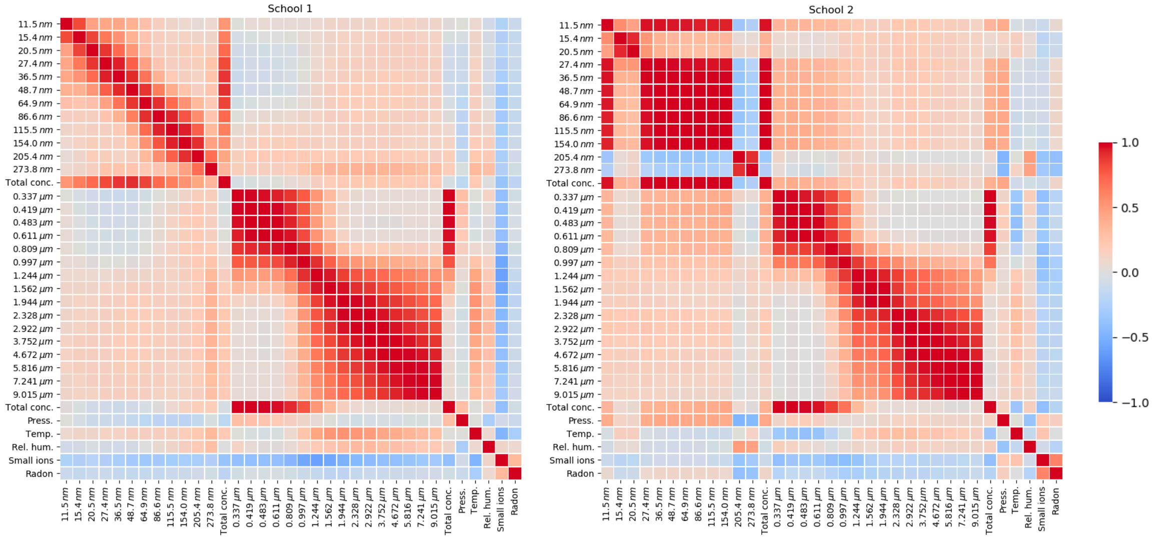
This can be done in a number of ways, and in statistical modelling this technique is known as feature selection [25]. However, at this stage, in order to keep physical interpretation of predictors possible, we have opted not to use principal component analysis and similar methods for feature selection, and aggregated particle channels into larger size bins, approximately corresponding to groups of correlated channels depicted in Figure 3. Aggregations that were used are shown in Table 2. Aggregations only include consecutive channels, and for that reason in School 2, 11.5 nm channel was excluded from Aggr2, despite the high correlation evident from Figure 3. Note that this kind of aggregation somewhat lacks in terms of desirable properties of predictors (such as low cross correlation) that could be achieved via use of principal component analysis (PCA) [25], however, it still preserves possibility for relatively simple interpretation of model parameters in linear regression. Later we will also explore the possible benefits of PCA for aggregation of particle channels in the context of predictive modeling.



(**a**)

(**b**)

Figure 2. Radon concentration (dotted) vs. small ion concentration (full line) in (**a**) School 1 and (**b**) School 2.



(**a**) (**b**)

Figure 3. Correlation matrix of quantities measured in (**a**) School 1 and (**b**) School 2. Shades of red are used for positive correlations, and shades of blue for negative correlations. (Plots were produced in Python 3.7.4 environment using libraries Seaborn 0.9.0 [23] for visualization and Pandas 0.25.2 [24] for data processing and calculations.)

Table 2. Aggregation of particle channels into larger bins based on correlation matrix in Figure 2.

<table>
<tr><th colspan="6">(a)</th></tr>
<tr><th>Nanoscan ch. [nm]</th><th>11.5</th><th>15.4–20.5</th><th>27.4–64.9</th><th>86.6–154</th><th>205–273.8</th></tr>
<tr><td>S1</td><td colspan="3">Aggr1</td><td colspan="2">Aggr2</td></tr>
<tr><td>S2</td><td></td><td>Aggr1</td><td colspan="2">Aggr2</td><td>Aggr3</td></tr>
<tr><th colspan="6">(b)</th></tr>
<tr><th>OPS ch. [um]</th><th colspan="2">0.337–0.997</th><th colspan="3">1.244–9.015</th></tr>
<tr><td>S1</td><td colspan="2">Aggr3</td><td colspan="3">Aggr4</td></tr>
<tr><td>S2</td><td colspan="2">Aggr4</td><td colspan="3">Aggr5</td></tr>
</table>

As mentioned earlier we will test the predictive power of the two statistical modeling approaches, which will be inspired, in part, by Equation (6). Let us first consider linear regression model, with predictors being radon concentration and aggregated particle channels showed in Table 2. Since we are using radon concentration as a proxy for volumetric production rate we will also need a way of converting radon concentration to volumetric production rate. Since activity of 1 Bq/m^3 produces alpha particle of 5.49 MeV in cubic meter every second, and mean energy to create an ion pair in air is around 35.6 eV [26], one decay per second producing alpha particle will produce on average $A_0 \approx \frac{5.49\ \text{MeV}}{35.6\ \text{eV}\cdot 10^6 \text{cm}^3} = 0.15 \frac{\text{i.p}}{\text{cm}^3\text{s}}$. Using this approximate conversion constant, we can notice using data from Table 1, that in School 1 volumetric production rate has a median of around 6 i.p/cm^{3}s and maximum of 19 i.p/cm^{3}s, while in School 2 median is similar and maximum is about 37 i.p/cm^{3}s. Going back to linear regression model, which was implemented using software library [20], using whole data sets in Schools 1 and 2 we obtained coefficients given in Table 3. Explained variance of the model is 0.49 in School 1, and 0.52 in School 2. Note that this explained variance should not be interpreted as predictive power of the model, since it was calculated on the whole data set. We will further discuss the issue of data set splitting a bit later.

Looking at the results listed in Table 3, several conclusions can be made. In both schools, the linear model had a similar value of explained variance, despite School 2 having one additional predictor. The intercept (corresponding to n_0 in Equation (6)) of the linear model is larger in School 1, which is in accordance with relative value of descriptive statistics for ions for two schools listed in Table 1. Note that n_0 in Equation (6) is *not* the median of n^-, but rather a value corresponding to the medians of radon and particle concentrations. The radon term is very similar in both schools, indicating similar increase in ion concentration with radon concentration.

In the linear model for both schools, the sign of attachment coefficients is physically justified (positive), despite the fact that we have used ordinary least squares, and did not enforce the sign of predictors a priori. Attachment coefficients become larger for aggregations corresponding to larger particle diameters, in accordance to theoretical expectations [26]. However, the magnitude of the attachment coefficients is larger compared to theoretical expectation, and this effect could be due to non-controlled PM size distribution which is to be expected in non-laboratory conditions, and also differences in number of particles in different size channels that were aggregated.

So as an intermediate conclusion, we can note that the balance equation provided valuable guidance for choice of predictors, and also that due to understanding of underlying physics we could inspect and verify the sign of predictors. On the other hand, despite all these advantages, the explained variance seems low, which makes physical interpretation of model parameters rather approximate. Furthermore, the relatively low explained variance would negatively influence predictive ability of the model. Let us now examine possible predictive power of the statistical models developed around balance equation.

In the further text we will focus on School 1, since during this particular campaign we have collected more data, and furthermore data gaps in particulate predictors were smaller compared to School 2. Having sufficient data for training and testing is paramount for developing models and also for testing them in a meaningful way. Under optimal conditions, stages of model validation, selection, and predictive errors should be calculated using independent i.e., previously unseen data, however this is often not possible. For smaller datasets it can easily happen that data is incomplete, and there could be different ratios of incompleteness for the test and training sets. Best way for doing cross validation is when it resembles and mimics the way in which model is to be practically used. This brings issues of optimal training test split, both in terms of amount of data and also temporal position of the data. Similar issues are encountered in low cost sensor calibration [19,27]. Additionally, since we are dealing with time series analysis test data will always need to be temporally separated and have timestamp later in time compared to training data. Thus, to satisfy these requirements of complete data sets and optimal timestamps of training and test data we opted to use 50-50 training test split. One additional requirement for the predictive models of small ion concentration is that they must produce positive output. A way to ensure positive prediction of the statistical model, is to take logarithm of the concentration when training the model and then to exponentiate the prediction of the model [25], and this was applied here.

Figure 4 shows comparison of a linear model based on radon and particle aggregates for School 1 and a measurement of small ions. The score metric (R^2) on the training set is 0.44 and on test set, it is 0.49. While the model shows similar trend as the measurements it is also evident that it significantly under predicts small ions concentration. Gaps in the figure correspond to periods where data for particles was missing.

Table 3. Linear model based on Equation (6), with radon as a proxy for volumetric rate and aggregated particle channels.

Parameters of The Linear Model	**Intercept**	**Radon Term**	**Aggr1**	**Aggr2**		**Aggr3**	**Aggr4**
School 1 parameters	291.6	1.61	-5.08×10^{-3}	-8.24×10^{-3}		-5.08×10^{-1}	-4.88×10^{1}
School 1 attachment $\beta_i\left[\mathrm{cm}^3/\mathrm{s}\right]$			2.87×10^{-7}	4.65×10^{-7}		2.87×10^{-5}	2.75×10^{-3}
Parameters of The Linear Model	**Intercept**	**Radon Term**	**Aggr1**	**Aggr2**	**Aggr3**	**Aggr4**	**Aggr5**
School 2 parameters	111.8	1.52	-2.00×10^{-3}	-6.88×10^{-3}	-2.08×10^{-2}	-3.75×10^{-1}	-2.29×10^{0}
School 2 attachment $\beta_i\left[\mathrm{cm}^3/\mathrm{s}\right]$			1.05×10^{-6}	3.61×10^{-6}	1.09×10^{-5}	1.97×10^{-4}	1.20×10^{-3}

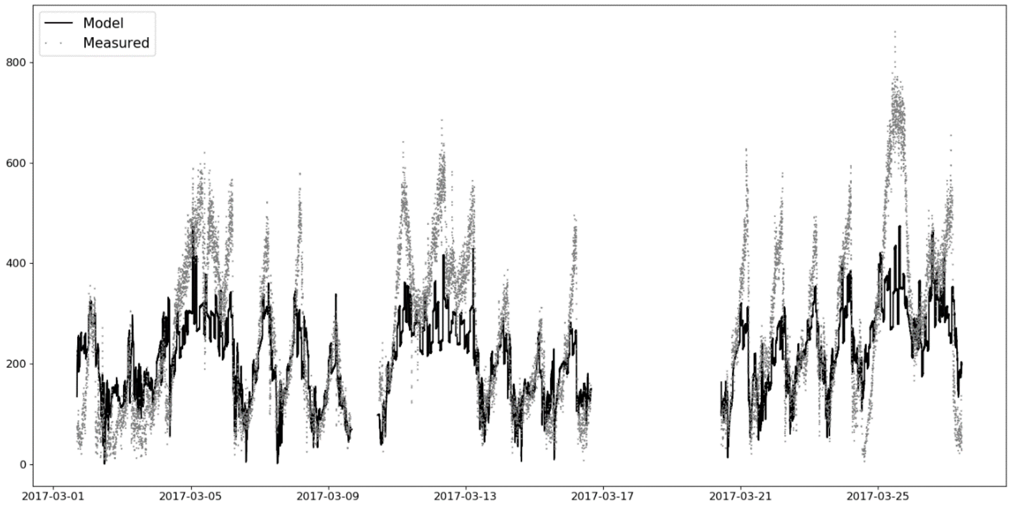


Figure 4. Comparison of a linear model based on radon and particle aggregates for School 1 (solid line) and a measurement of small ions (dotted line). Unit is [#/cm^3]. Training/test score 0.44/0.49.

Let us now examine if ANNs can improve the situation compared to the linear model. If we move away from a requirement that parameters of the statistical models have simple and precisely defined physical interpretation, we can bring several improvements to our modeling methodology. Choice of predictor variables can now become less stringent compared to the linear model that was based on a balance equation, so we can actually make statistically more justified aggregations of particle channels, e.g., by using PCA. Since for the linear model aggregation was based on correlation matrices derived from complete dataset, this could introduce so called "knowledge leak" from training to test data, and is thus best to avoid it, e.g., by using PCA. Furthermore, when using PCA it is more clear what amount of variance is left out of the model predictors when reducing the number of particulate matter related features. For example, first two PCA components explain 62% of the particle channels variance. The number of neurons in the hidden layer, and the optimal number of input PCA components was determined in optimization procedure. The most optimal training/test ratio was observed for ANN that has 4 neurons with ReLU activation function in the hidden layer and uses 2 PCA components and radon concentration, a total of 3 signals, as input. The optimal ANN model achieved median absolute error of about 40 ions/cm^3 and explained variance of about 0.70. Some additional details about network architecture, hyperparameters and tuning procedures are given in Appendix C.

Predictive power of the optimized ANN model is illustrated in Figure 5. It seems that the model now doesn't significantly under predict small ion concentration, and it improves on the linear model, despite having smaller number of input predictors. Furthermore, the score on test set is significantly improved. It also appears that some of the noise introduced by radon measurements is now less pronounced compared to the linear model output shown in Figure 4. However, it seems that while the overall trend in small ions concentration is well modelled, the shape of the peaks is not always preserved. Since final ANN has rather simple architecture and there are no discrepancies between training and test scores it can be concluded that the model is not overly complex to introduce overfit, and additionally, training and test datasets are sufficiently complete.

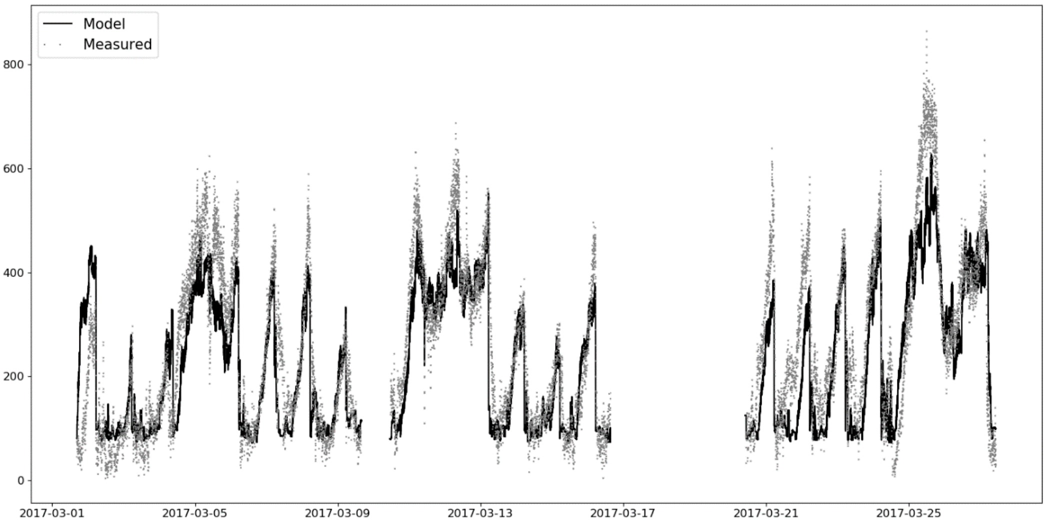


Figure 5. Comparison of an artificial neural network (ANN) model based on radon and 2 PCA particle components for School 1 (solid line) and a measurement of small ions (dotted line). Unit is [#/cm^3]. Training/test score 0.69/0.69.

4. Conclusions

In this work, we have studied indoor air quality related to important, but more rarely continuously measured parameters: small ion concentration and its association with radon and PM concentration in two elementary schools′ indoor environments. We have analyzed the association using two approaches: descriptive statistical analysis coupled with linear modelling and artificial neural network predictive modelling. The following conclusions can be made.

- Descriptive statistics showed that for similar median radon concentrations larger number of nanosized particles corresponds to smaller number of small ions. This observation is coherent with the balance equation.
- The linear model derived directly from balance equation allowed estimation of balance equation parameters. The parameters corresponding to the radon term were similar in both schools, indicating similar increase in small ion concentration with radon concentration in both schools. Regarding particulate matter parameters, it was observed that attachment coefficients become larger for particle aggregations corresponding to larger particle diameters, in accordance to theoretical expectations. However, these parameters were different in two schools, possibly due to different air pollution composition.
- The hypothesis that small ion concentration, which may have certain impact on human health and wellbeing, can be predicted based on radon and particulate matter measurements predictors was successfully tested.
- Explained variance for the linear predictive model was under 0.5, and for the artificial neural network (ANN) predictive model with similar predictors was around 0.7. ANN predictive model has achieved median absolute error of about 40 ions/cm^3 on test data.

These modelling efforts enable several future work directions and applications that may be of wider interest for indoor air quality monitoring. Since small ions can be an important part of air quality consideration, their concentration could be to a certain extent modelled based on several more easily obtained/measured predictors. We have showed that the ANN model successfully helped in harmonizing inputs from several instruments of different grade (low cost radon and lab grade particulate matter monitors) and enabled predictions of small ions concentration of comparable quality to the lab grade Gerdien type instrument. Furthermore, since the particulate matter concentration was one of the important predictors, and having in mind recent uptake of low-cost PM monitors, data driven solution for estimating small ions concentration based on these sensors as supporting predictors is an interesting future research topic.

Author Contributions: Conceptualization, M.D. (Miloš Davidović); data curation, M.D. (Miloš Davidović) and M.J.-S.; formal analysis, M.D. (Miloš Davidović) and M.D. (Milena Davidović); funding acquisition, M.D. (Miloš Davidović), M.D. (Milena Davidović) and M.J.-S.; investigation, M.D. (Miloš Davidović), P.K. and M.J.-S.; methodology, M.D. (Miloš Davidović), P.K., M.J.-S. and Z.R.; project administration, M.D. (Miloš Davidović); resources, P.K. and M.J.-S.; software, M.D. (Miloš Davidović); supervision, M.J.-S. and Z.R.; validation, M.D. (Miloš Davidović) and M.D. (Milena Davidović); visualization, M.D. (Miloš Davidović); writing—original draft, M.D. (Miloš Davidović), M.D. (Milena Davidović), R.J., P.K., M.J.-S. and Z.R.; writing—review and editing, M.D. (Miloš Davidović), M.D. (Milena Davidović), R.J., P.K., M.J.-S. and Z.R. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Acknowledgments: We would like to thank the participating pupils and staff of Elementary School "Aleksa Šantić" in Kaluđerica, Belgrade, and Elementary School "20. oktobar" New Belgrade, Belgrade, for helping with the smooth realization of the experiments. The authors are thankful to the COST Action INDAIRPOLLNET CA17136, supported by COST (European Cooperation in Science and Technology).

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

Appendix A

There are several main physical processes that are involved in changes in SI concentration. Firstly, SI are constantly created, in pairs, by ionizing radiation that exists in the environment. SI are continuously created when neutral air molecules are irradiated by cosmic rays or radiations from natural radioactive materials in soil and air. Neutral air molecules may be ionized into positive ions and free electrons that are attached to other hydrous or oxygen molecules in air, and in order to become stable these ions are adsorbed to neutral molecules forming cluster ions.

UV radiation is mainly responsible for ionization of molecules and atoms in high atmosphere, but it is exhausted at higher levels and doesn't arrive to lower troposphere [28]. Several natural sources of ionizing radiation are responsible for air-ion production in the lower troposphere, namely radioactive gases (particularly ^{222}Rn and its progenies), radioactive substances at ground level (e.g., natural α and β emitters in the air and soil) and cosmic rays. These three contributions are comparable, approximately 20% of the total surface ionization rate is due to ionization from cosmic rays, and remaining 80% arises from natural α and β emitters in the air and soil.

Air-ion pair generation near the ground varies mostly with the concentration of ^{222}Rn and its progenies. The half-life of ^{222}Rn is approximately 3.8 days and the decay product is an alpha particle with energy 5.49 MeV. The decay of ^{222}Rn generates a large number of nitrogen and oxygen molecular ions (order of magnitude ~10^5) per each α-particle. As a consequence, the near-ground ionization rate caused by background ionization, is about 10 ion pairs/cm^{3}s in continental areas [28]. Within microseconds of the ionization process, primary ions evolve through the process of hydration to form small cluster ions, also known as small air ions or nano-air ions. This class of air ions can survive much longer, up to 100 s, depending predominantly on air pollution and air density [22].

While small air ions technically are particulate matter belonging to ultrafine particulate fraction (albeit only a few nanometers in diameter), it is worth pointing out the following. Small air ions are electrically charged clusters consisting of several molecules in which ordinarily neutral atmospheric molecules/atoms have gained or lost electrons. While particulate matter can also be charged, it is composed of a much larger number of molecules and is thus up to several orders of magnitude larger in diameter compared to the small ion clusters. Since cluster ions readily attach to particles, it is know that their concentration decreases sharply within few tens of meters from the road [29], this making indoor ion sources most significant, which is important to keep in mind for indoor air monitoring.

SI are also continually being destroyed in a process of recombination, producing neutral molecular clusters. In addition to the process of recombination, SI can attach to PM. Because of this, a change in PM concentration directly results in a change in SI concentration. A significant portion of PM in the urban environment is a result of human activities, where smaller particles are typically associated with the process of combustion occurring in vehicles, industrial activities, biomass burning, and similar, while larger particles are typically due to construction and demolition activities, entrainment of outdoor dust and similar. The SI balance equation can be used to quantitatively describe the above-mentioned processes.

Appendix B

Table A1. Description of the instruments used in data gathering campaigns conducted in this study.

Instrument	Specification Based on Datasheets, Application Notes and Calibration Certificates	Type of Calibration
TSI NanoScan SMPS Model 3910	Relative standard deviations in total concentration 2.7% Sizing of the particles: standard deviations in median particle diameter 1.1% Discrepancy relative to certified size ranges of 20, 60, 80, 200 and 300 nm less than 8%.	NIST traceable using TSI calibration system, conducted in test atmosphere of polystyrene latex particles
TSI Optical particle sizer 3330	Counting efficiency at 0.5 µm (90–110%) Inlet flow: 0.95–1.05 L/min Sizing of 1 um particles: 90–100% *Allowable range is given in parenthesis, calibration certificate includes traceably measured single value.*	NIST traceable using TSI calibration system, conducted in test atmosphere of polystyrene latex particles
Radon Scout	Sampling type: Diffusion Sensitivity: 1.8 count per minute/kBq/m^3 (4 cph/pCi/L) Measurement range: 0. . . 2 MBq/m^3 Error: ±5% within the whole range or smaller	Factory calibration, instrument class certified by the US-EPA/NRSB
Gerdien-type air ion detector	Sensitivity of the current measurement is limited by AD converter resolution and amounts 1.6 fA. Using Equation (1) in [22], this value equals to 2 ions/cm^3. Measuring sensitivity is limited by noise induced by various sources (uncertainties of air-flow, calibration, temperature drift, gain error, etc.) and is experimentally obtained to be ±5 ions/cm^3.	Calibrated using Equation (1) in [22] and Keithley 261 small current generator (output signal ~10 fA). Flow tuning was done via hot-wire anemometer.

Particle sizers were manufactured by TSI Incorporated, 500 Cardigan Road, Shoreview, Minnesota 55126 USA. Radon Scout was manufactured by SARAD GmbH, Wiesbadener Straße 10, DE-01159 Dresden, Germany. Gerdien-type air ion detector was developed by Institute of Physics, University of Belgrade, Belgrade, Serbia. The divide between low-cost and lab grade instruments is not strict, and does not always refer to price of the sensor itself. The line is further blurred by the fact that low cost sensors may require extensive, i.e., costly (re)calibration efforts. While operating principles of both type of instruments can be very similar, lab grade instruments are typically characterized by better implementation of these principles, for example better sampling method, higher quality of the sensors that allow better temporal resolution etc. and consequently better accuracy. In this paper, all of the instruments used, except for the Radon Scout (due to instruments sampling method, time resolution, cost and lack of traceability) could be considered to be lab grade instruments.

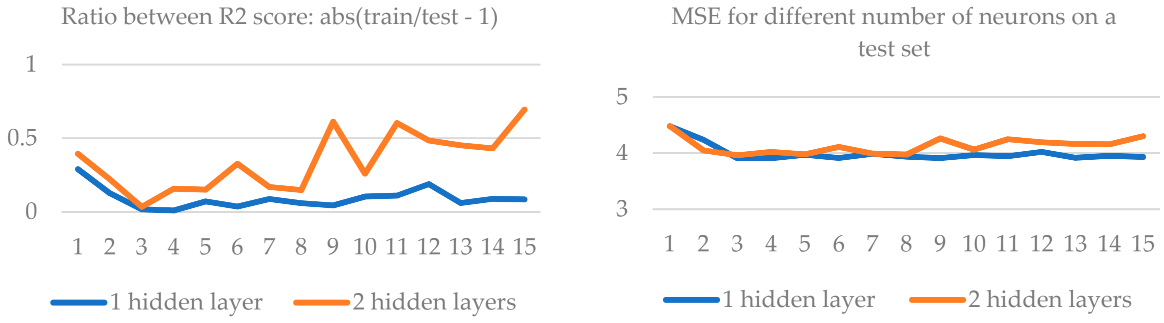
Appendix C

There are two classes of parameters of neural networks that can be tuned in order to improve the predictions. The first class consists of model parameters that are calculated during the process of neural network training. This is done by optimizing a loss function, which in our case was squared error (MSE). The second class of parameters are hyperparameters. These parameters define the structure of the model and need to be defined a priori. Hyperparameters used in this study are shown in Table A2. Parameter tuning can easily become an exhaustive task, and it is desirable to use simple models if they can achieve similar accuracy to the more complex ones, in order to keep solution space more manageable.

Table A2. Hyperparameters of the artificial neural network model used in this study.

Hyperparameters	Minimum	Maximum	Step Size
Hidden layers	1	2	1
Neurons per hidden layer	1	15	1
Neurons in input layers (Radon + PCA components)	2	5	1
Early stopping	Not used		
Activation function	'relu'		
Cost function	MSE		
Solver	'adam'		

Figure A1a shows what happens with the R2 score on training and test data sets when number of neurons is increased in the hidden layer. It can be observed that only a small network, with 3–4 neurons results in a balanced training/test score, and that further increase of number of neurons results in increased test set discrepancy, due to an overfit. This was the reason why no additional layers were introduced, since this would certainly lead to an overfit, i.e., an ANN model that behaves excellently on the training data, and very poorly on the test data. Furthermore, similar conclusion can be made by observing MSE observed on a test set, shown in Figure A1b. The network with two hidden layers had a greater MSE compared to the simpler one hidden layer network. The simplest models that have optimal statistics are ANN models with three and four hidden neurons (network architecture and activation function is shown in Figure A2), with more complex model being slightly better in terms of R2 score and MSE. MSE is susceptible to outliers, and a more robust statistical measure of model performance is a median absolute error, which turned out to be around 40 ions/cm^3 for optimal ANN models. The explained variance of around 0.70 was from the test set for the optimal ANN models.



(**a**) (**b**)

Figure A1. Statistics for different ANN models where several hyperparameters are changed (number of hidden layers and number of neurons per hidden layer). (**a**) R2 score training test ratio (**b**) MSE on a test set. All models have radon and 2 particulate matter related PCA components as inputs. Figure A1b uses log10 scale.

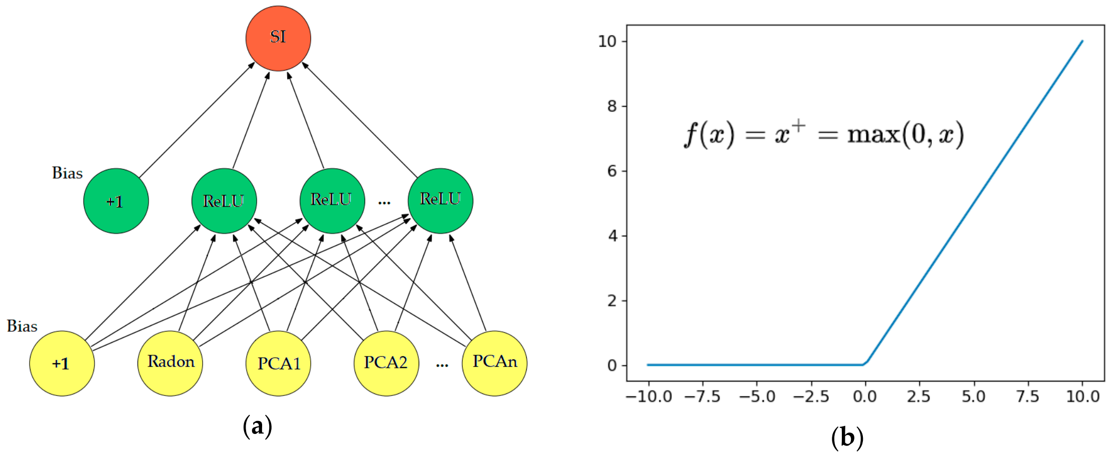


Figure A2. (**a**) ANN considered in this paper. Optimal model has two PCA components and 4 ReLU neurons. (**b**) Plot of a ReLU activation function.

References

1. EEA. *The European Environment—State and Outlook 2020, Knowledge for Transition to a Sustainable Europe*; European Environment Agency: Luxembourg, 2019.
2. Simoni, M.; Baldacci, S.; Maio, S.; Cerrai, S.; Sarno, G.; Viegi, G. Adverse effects of outdoor pollution in the elderly. *J. Thorac. Dis.* **2015**, *7*, 34.
3. Buka, I.; Koranteng, S.; Osornio-Vargas, A.R. The effects of air pollution on the health of children. *Paediatr. Child Health* **2006**, *11*, 513–516.
4. Höppe, P.; Martinac, I. Indoor climate and air quality. *Int. J. Biometeorol.* **1998**, *42*, 1–7. [CrossRef] [PubMed]
5. Pantelić, G.; Čeliković, I.; Živanović, M.; Vukanac, I.; Nikolić, J.K.; Cinelli, G.; Gruber, V. Qualitative overview of indoor radon surveys in Europe. *J. Environ. Radioact.* **2019**, *204*, 163–174. [CrossRef] [PubMed]
6. EEA. *Air Quality in Europe—2018*; European Environment Agency: Copenhagen, Denmark, 2018.
7. Moshammer, H.; Panholzer, J.; Ulbing, L.; Udvarhelyi, E.; Ebenbauer, B.; Peter, S. Acute effects of air pollution and noise from road traffic in a panel of young healthy adults. *Int. J. Environ. Res. Public Health* **2019**, *16*, 788. [CrossRef]
8. WHO. *WHO Guidelines for Indoor Air Quality: Selected Pollutants*; World Health Organization: Geneva, Switzerland, 2010.
9. Peters, A.; Wichmann, H.E.; Tuch, T.; Heinrich, J.; Heyder, J. Respiratory effects are associated with the number of ultrafine particles. *Am. J. Respir. Crit. Care Med.* **1997**, *155*, 1376–1383. [CrossRef]
10. Perez, V.; Alexander, D.D.; Bailey, W.H. Air ions and mood outcomes: A review and meta-analysis. *BMC Psychiatry* **2013**, *13*, 29. [CrossRef] [PubMed]
11. Pino, O.; La Ragione, F. There's something in the air: Empirical evidence for the effects of negative air ions (NAI) on psychophysiological state and performance. *Res. Psychol. Behav. Sci.* **2013**, *1*, 48–53.
12. Jiang, S.-Y.; Ma, A.; Ramachandran, S. Negative air ions and their effects on human health and air quality improvement. *Int. J. Mol. Sci.* **2018**, *19*, 2966. [CrossRef]
13. *Health Effects of Exposure to Radon: BEIR VI*; National Research Council: Washington, DC, USA, 1999.
14. Kim, K.-H.; Kabir, E.; Kabir, S. A review on the human health impact of airborne particulate matter. *Environ. Int.* **2015**, *74*, 136–143. [CrossRef]
15. Rivas, I.; Viana, M.; Moreno, T.; Pandolfi, M.; Amato, F.; Reche, C.; Bouso, L.; Àlvarez-Pedrerol, M.; Alastuey, A.; Sunyer, J. Child exposure to indoor and outdoor air pollutants in schools in Barcelona, Spain. *Environ. Int.* **2014**, *69*, 200–212. [CrossRef] [PubMed]
16. Baloch, R.M.; Maesano, C.N.; Christoffersen, J.; Banerjee, S.; Gabriel, M.; Csobod, É.; Fernandes, E.O.; Annesi-Maesano, I.; Szuppinger, P.; Prokai, R.; et al. Indoor air pollution, physical and comfort parameters related to schoolchildren's health: Data from the European SINPHONIE study. *Sci. Total. Environ.* **2020**, *739*, 139870. [CrossRef] [PubMed]

17. Schweidler, E. Ueber das Gleichgewicht zwischen ionenerzeugenden und ionenvernichtenden Vorgaengen in der Atmosphaere. *S. B. Akad. Wiss. Wien* **1918**, *128*, 947–955.
18. Donnelly, M.I. *A Study of the Nuclear Content of the Atmosphere and the Lifetime of Small Ions in New York City*; ETD Collection for Fordham University: New York City, NY, USA, 1950.
19. Topalović, D.B.; Davidović, M.D.; Jovanović, M.; Bartonova, A.; Ristovski, Z.; Jovašević-Stojanović, M. In search of an optimal in-field calibration method of low-cost gas sensors for ambient air pollutants: Comparison of linear, multilinear and artificial neural network approaches. *Atmos. Environ.* **2019**, *213*, 640–658. [CrossRef]
20. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
21. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
22. Kolarž, P.; Miljković, B.; Ćurguz, Z. Air-ion counter and mobility spectrometer. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Materials At.* **2012**, 219–222. [CrossRef]
23. Waskom, M.; Botvinnik, O.; O'Kane, D.; Hobson, P.; Ostblom, J.; Lukauskas, S.; Gemperline, D.C.; Augspurger, T.; Halchenko, Y.; Cole, J.B.; et al. mwaskom/seaborn: V0. 9.0 (July 2018). Available online: https://zenodo.org/record/1313201#.X0dZDSMRWUk (accessed on 6 July 2020).
24. McKinney, W. Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; pp. 51–56.
25. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.
26. University of Arizona, Hydrology & Atmospheric Sciences, Atmospheric Electricity ATMO/ECE 489/589 Spring Term, Lecture Notes. Available online: http://www.atmo.arizona.edu/students/courselinks/spring13/atmo589/ (accessed on 6 July 2020).
27. De Vito, S.; Di Francia, G.; Esposito, E.; Ferlito, S.; Formisano, F.; Massera, E. Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices. *Pattern Recognit. Lett.* **2020**. [CrossRef]
28. Iribarne, J.; Cho, H.-R. *Atmospheric Physics*; D. Reidel Publishing Company: Dordrecht, The Netherlands, 1980.
29. Jayaratne, E.; Ling, X.; Morawska, L. Observation of ions and particles near busy roads using a neutral cluster and air ion spectrometer (NAIS). *Atmos. Environ.* **2014**, *84*, 198–203. [CrossRef]