Systematic Framework for Quantitative Assessment of Indoor Air Quality Under Future Climate Scenarios; 2100s Projection of a Belgian Case Study

Mohsen Pourkiaei[1*], Ramin Rahif[2], Claudia Falzone[1], Essam Elnagar[3], Sébastien Doutreloup[4], Justin Martin[1], Xavier Fettweis[4], Vincent Lemort[3], Shady Attia[2], Anne-Claude Romain[1]

[1] Sensing of Atmospheres and Monitoring (SAM) Lab, UR Spheres, Department of Environmental Science and Management, Faculty of Sciences, University of Liège, Arlon, Belgium.
[2] Sustainable Building Design Lab, Dept. UEE, Faculty of Applied Sciences, University of Liège, Belgium
[3] Thermodynamics Laboratory, Aerospace and Mechanical Engineering Department, Faculty of Applied Sciences, Université de Liège, Belgium.
[4] Lab. Of Climatology and Topoclimatology, SPHERES research unit, Department of Geography, Faculty of Sciences, University of Liège, Liège, Belgium.
[*]*Corresponding email: sm.pourkiaei@uliege.be*

S1. IAQ Measurements of the Case Study

Considering the selected IAQ indicators O_3, CO, NO, NO_2, $PM_{2.5}$, PM_{10}, VOCs, air temperature (T), Relative Humidity (RH), and air pressure (P), monitoring devices based on LCSs were fabricated. The calibration procedures were carried out both indoors and outdoors with the support of reference analyzers. Further information on the calibration procedures can be found in our previous study [1]. It should be mentioned that VOCs were not considered during the calibration process due to the lack of outdoor data. Afterward, the IAQ measurement campaign of the case study house (located in Arlon) was held from the 20th of June to the 31st of August 2021, both indoors and outdoors, concurrently. No recalibration of monitoring devices was considered during the campaign (over 73 days).

Also, a comprehensive questionnaire (hourly checklist) was designed to log the occupancy pattern (3 adult inhabitants), indoor activities incidence (sleeping, cooking, cleaning, showering, smoking), natural ventilation behavior (opening of windows), and 2 exhaust fan operations. Royal Meteorological Institute of Belgium afforded the weather data.

S2. IAQ Models

In this section, detailed information concerning the IAQ model descriptions, their general principles, their involved elements and parameters, their strengths and drawbacks, as well as their validation and calibration procedures are presented.

S2.1. Mass Balance Models

Mass balance models are formulated to enable the estimation of pollutant concentrations and the influence of sources, sinks, and IAQ control strategies on contaminant levels. These models are rooted in the conservation of mass principle [2]. They can be defined over either a single zone (compartment/room) or multi-zone ordonnances. In a single-zone model, the entire building is characterized as a shoe box model. On the other hand, multi-zone models characterize a building through interconnected spaces. The general mass balance equation for a well-mixed single zone is expressed as follows [3,4]:

$$V_i \frac{dC_i}{dt} = C_a Pt_a Q_{a,i} + C_h Q_{h,i} + \sum_{i=1, j\neq N}^{N} C_j Q_{j,i} - C_i Q_{i,a} - C_i Q_{i,h} - \sum_{i=1, j\neq N}^{N} C_i Q_{i,j} + E_i - S_i \quad \text{(S1)}$$

in which V_i denotes the zone volume, C_i denotes the indoor pollutant concentration in zone i, C_a denotes the outdoor concentration, Pt_a is the penetration factor for outdoor pollutants entering the indoors, $Q_{a,i}$ is the airflow from the outdoors into zone i, C_h is the concentration in the HVAC system, $Q_{h,i}$ is the airflow from the HVAC system into zone i, C_j is the concentration in zone j, $Q_{j,i}$ is the airflow from zone j into zone i, $Q_{i,a}$ is the airflow from zone i to the outdoors, $Q_{i,h}$ is the airflow from zone i into the HVAC system, $Q_{i,j}$ is the airflow from zone i into zone j, E_i denotes the emission source term, S_i denotes the pollutant removal term (e.g., air cleaners and sinks, etc.). It is important to recognize that the emission source and sink terms, might involve extra differential equations that define their characteristics.

A key assumption in various mass balance models is the assumption of well-mixed zones. This assumption is held under the following circumstances [3,5]:

- When time scales are several minutes or longer.
- When concentrations near significant sources are not of concern.
- When there are no local flow disruptions near the target zone.

The intrazonal airflows (airflows among zones themselves, and between zones and the ambient) can be influenced by the mechanical and natural ventilation. Some IAQ models (such as CONTAM) are capable of calculating various airflows, contingent on adequate data regarding door and window openings, temperature differentials, HVAC characteristics, and similar factors. In contrast, other models may necessitate the parameterization of input airflow as data.

Mass balance models are formulated to estimate the average indoor pollutant concentrations. In numerous scenarios, this average concentration serves as the focal point. Nonetheless, there are instances where the average concentration is not the primary focus. For instance, if the

concern is the exposure of an individual utilizing a potent emitting product the mean indoor concentration falls short. In such cases, mass balance models may prove insufficient [6,7].

S2.2. CFD Models

Certain scenarios in IAQ modeling require the prediction of punctual (local) concentrations rather than average zonal concentrations. In these cases, the mass balance models fail to achieve targeted objectives. In such situations, objectives can be better addressed by CFD models. CFD models have two significant differences with mass balance models. First, CFD models estimate air velocity and contaminant concentration at discrete points within a zone. Secondly, CFD models tackle a set of partial differential equations as opposed to the ordinary differential equations targeted by mass balance models [8,9]. As a consequence, CFD models are computationally more expensive compared to mass balance models.

CFD models prove particularly valuable for examining airflows and distribution within zones [10]. Additionally, emission source and sink models play a crucial role in estimating pollutant concentrations using CFD models. If the characterization of the emission source and sink is insufficient, concentration estimation will be inaccurate, despite accurate airflow calculations [11,12]. The partial differential equations governing fluid flow and concentration are rooted in the principles of mass, momentum, and contaminant concentration conservation. The variables of interest include velocity components, concentration, and certain turbulence parameters that consider turbulent flow effects [7]. These fundamental equations are typically expressed in the following manner [13]:

Conservation of mass:

$$\frac{\partial \rho}{\partial t}+\frac{\partial \rho u}{\partial x}+\frac{\partial \rho v}{\partial y}+\frac{\partial \rho w}{\partial z}=0 \tag{S2}$$

Conservation of momentum (x-direction):

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)=-\frac{\partial p}{\partial x}+\frac{\partial \sigma_{xx}}{\partial x}+\frac{\partial \sigma_{xy}}{\partial y}+\frac{\partial \sigma_{xz}}{\partial z}+f_{xx} \tag{S3}$$

where u, v, and w are horizontal, lateral and vertical velocities, respectively, ρ is density, p is pressure, f_x is velocity body force terms, σ_{xx} is normal viscous stress, σ_{xy}, σ_{xz} are tangential (shear) viscous stress terms. The CFD concentration formulation is presented in section S2.4.2. There are three fundamental CFD approaches typically employed for simulating flow and contaminants transport within zones: Finite Difference Method (FDM), Finite Volume Method (FVM), and the Finite Element Method (FEM) [13]. CFD solvers can run structured, unstructured, or hybrid meshing with 2D (triangle, quadrilateral) or 3D (tetrahedral,

hexahedron, etc.) forms with varying densities. Over the past decades, various other methods have emerged in scholarly literature, including the Boundary Element Method (BEM), the Lagrangian Particle Transport technique (LPT), the Particle-in-Cell method (PIC), and the Meshless method (MM) [14-18].

S2.3. Statistical Models

Projection of IAQ through mechanistic methods is based on comprehension of the fundamental mechanisms governing the displacement and transport of indoor air contaminants. As mechanistic models demand intricate inputs, securing sufficient information for their operation becomes challenging. This becomes even more critical in scenarios such as the study of building stocks, and real-world conditions where occupants interact with indoor settings. In such instances, statistical (also called numerical or mathematical) models based on mathematical approaches and Artificial intelligence (AI) offer an alternative pathway for IAQ prediction.

Machine learning and statistical models have gained substantial traction in outdoor settings for estimating atmospheric pollutant concentrations [19,21] and in indoor settings for predicting thermal comfort and building energy efficiency [22,24]. Although many statistical models have been applied to predict IAQ, research regarding the depth and scope of their applications is relatively new [25-30].

Statistical models offer the capacity to estimate IAQ through the utilization of questionnaires and/or measurements. Important statistical models employed for IAQ predictions include Artificial Neural Networks (ANNs), Regression models, and Decision Trees [31].

As a popular method in this category, ANNs operate on a network of interconnected nodes or neurons [32,33]. They employ an intuitive learning and prediction process making it particularly effective for solving non-linear problems. ANNs employ multicomplex combinations of weights and functions to transform input variables into predicted outputs, eliminating the need for predefined assumptions regarding variable relationships (black box modeling) [30,31].

Alternatively, Regression models such as Multiple Linear Regression (MLR), Kernel regression, and Partial Least Squares (PLS), approximate the relationships between variables. Among them, MLR is the main and extensively adopted model to assess the linear links between an output (dependent variable) and various inputs (explanatory variables). The MLR model basis can be formulated as [31,34]:

$$y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k + \varepsilon \quad \text{(S4)}$$

in which for k observations; y denotes the output, x_i denotes inputs, b_0 denotes the y-intercept (constant term), b_k represents the regression (slope) coefficients for each input, and ε denotes the stochastic error (residuals).

In addition to MLR, alternative regression models such as the Least Absolute Shrinkage and Selection Operator (LASSO) regression and stepwise regression might offer more advanced exploration and selection of input variables [35].

On the other hand, decision trees employ a tree-like structure to model decisions and their potential outcomes for data classification or regression. It serves for both classification (classification tree) and prediction (regression tree). An ensemble of regression trees (aggregation of multiple regression trees) is called Random Forest Regression (RFR). While a solitary regression tree may struggle with complex problems and lack of robustness, an RFR stands as the most frequently employed decision tree-oriented model [36].

Overall, mechanistic models may project a sense of reliability while statistical models offer a significant utility of black box (or gray box) understanding. In situations where the specific mechanisms or their dynamic variations lack well-established foundations and extensive datasets are available, statistical models are more favorable.

S2.4. Indoor Air Chemistry Models

The field of Indoor Air Chemistry (IAC) aims to comprehend the elements influencing exposure by investigating the chemical processes that take place in the air, aerosol particles, and surface reservoirs within indoor settings. Within this field, primary chemical sources include emissions originating indoors, or infiltrating from the outdoors. Contrarily, secondary sources are related to those by reactive chemistry happening indoors. These sources may be involved with permanent (sustained; i.e. from building materials), periodic (episodic; i.e. human activities), or transient characteristics. There are five major domains of indoor environments influencing indoor chemistry as follows [37]:

1. Indoor airflow features. Especially those linked to the building structure and interior characteristics, including AERs and mixing time (inter/intra-zonal).
2. Extremely high Surface-Area-to-Volume ratio (SA/V) of indoor spaces. This ratio is approximately around 3 m^{-1} when considering only macroscopic surface areas. This value serves as a conservative estimate, acknowledging that at the microscopic scale, surfaces may possess porosity or roughness. Additionally, building materials,

furnishings, and paintings may feature low viscosity or high porosity, facilitating molecular diffusion into them. These surface reservoirs are pivotal in influencing nonreactive partitioning processes and reactive chemistry.

3. Indoor photon fluxes. Particularly in the ultraviolet light range. These fluxes are notably lower compared to outdoor conditions. The intensity and spectral composition of light indoors are highly influenced by factors such as the efficiency of sunlight transmission through the glass, cleanliness of the glass, time of day, type and number of windows, distance from the window, ambient cloudiness, and the types of indoor lighting.
4. Indoor T and RH. In contrast to outdoor environments, indoor T and RH are frequently controlled, preventing the wet deposition process (less for kitchen and bathroom while using). Even in the absence of HVACs, indoor T and RH variations are relatively lower than those experienced outdoors.
5. The human presence indoors. Occupants' activities, including cooking and cleaning, play a crucial role, but humans also cause direct effects through emissions and multiphase chemistry, via their clothing and skin. These impacts can be significant in densely populated environments.

Indoor settings are characterized by low light levels, relatively low concentrations of gas-phase oxidants, and limited durations for the reactive processing of gaseous and particulate components, owing to air exchange. Nonetheless, significant gas-surface partitioning and reactive multiphase chemistry take place within the extensive surface reservoirs present in all indoor environments. These interactions are crucial in shaping the composition of indoor surfaces as well as the surrounding gases and aerosol particles, thereby influencing human chemical exposure.

Although gas-phase chemistry typically doesn't play a dominant role in determining the fate of most VOCs, it still involves substantial radical cycling and organic nitrate formation [37,38]. Additionally, the formation of Secondary Organic Aerosol (SOA) can occur through gas-phase oxidation of various precursors, including monoterpenes, unsaturated compounds from skin and cooking oils, and cigarette smoke [37]. While indoor SOA is not usually the primary component of indoor aerosols, its significance increases under specific conditions, such as elevated O_3 levels and low AERs. Episodic events with high precursor concentrations, like using a terpene-based cleaner or cigarette smoking, can also contribute to the formation of ultrafine particles [37,39,40]. Further detailed information about gas-phase and multi-phase chemistry (gas-phase autoxidation mechanisms, gas-phase and condensed-phase

photochemistry, multiphase thermodynamic partitioning, aerosol partitioning, equilibrium partitioning models, surface chemistry, and chemical reactions) are available in two valuable recent references of indoor chemistry [37,41].

Various techniques can model indoor chemical processes. Central to many models for indoor chemical processes, is the utilization of mass or concentration balances. These balances can be applied in both single- or multizonal well-mixed models to forecast the dynamic changes in gas- and particle-phase species over time. The 2 main approaches for indoor chemistry models are Box models and CFD models. The box model is the most commonly used for indoor chemistry studies since, in essence, a modeler should select between modeling chemical complexity with a box model or spatial complexity with other methods. Recently, CFD has been applied to simulate certain simple chemical situations that present spatial variation [42]. Neither model can fully represent the complexities of indoor chemistry, while each seeks to address a unique gap in knowledge. The user must define the models' parameters accurately to acquire the most efficient understanding of the indoor processes [43].

S.2.4.1. Box Models

To model the fate of a generic molecule F, the generation and removal reactions R1 and R2 are assumed as follows, respectively:

$$G+H \rightarrow y_F F+\ldots \quad \text{(R1)}$$

$$F+J \rightarrow \ldots \quad \text{(R2)}$$

The reaction between molecules "G" and "H" yields molecule "F", with the molar yield of y_F. Additionally, "F" can be removed from the indoor air by reaction with another molecule "J". With respect to R1 and R2, the concentration balance is as [42,44]:

$$\frac{dC_F}{d_t} = P_F + \lambda C_{F,out} + E_F - r_F C_F - \lambda C_F - v_{d,F}\frac{A}{V}C_F \quad \text{(S5)}$$

In which C_i is the concentration (ppb) of species i, t is time (h), P_F (ppb.h^{-1}) is the formation rate of "F" due to gas-phase reactions (in R1), r_F (h^{-1}) is the loss rate of "F" (in R2), $C_{F,out}$ (ppb) is the outdoor concentration of "F"; λ (h^{-1}) is the AER, E_F (ppb.h^{-1}) is the emission rate of "F", $v_{d,F}$ (m/h) is the deposition velocity of "F" to indoor surfaces; and A (m^2) is the surface area indoors. The gas-phase chemistry formation and loss rate of "F" in Equation (S5) are defined as follows:

$$P_F = y_F k_{G\text{-}H} C_G C_H \quad \text{(S6)}$$

$$r_F = k_{F\text{-}J} C_J \quad \text{(S7)}$$

In which $k_{G\text{-}H}$ and $k_{F\text{-}J}$ ($ppb^{-1}.h^{-1}$) are the biomolecular reaction rate constant between "G & H", and "F & J". Additional subset models of indoor SOA, inorganic aerosol, and surface chemistry are available in the reference of [42]. The developed box model INCHEM-Py v1.2, supports complex chemical mechanisms (> 6000 species, > 19000 reactions) [43].

S2.4.2. CFD Models

The fundamental equations of the indoor chemistry CFD model are typically expressed in the following manner [13]:

Species concentration:

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} + w\frac{\partial C}{\partial z} = \frac{\partial}{\partial x}\left(D_{xx}\frac{\partial C}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_{yy}\frac{\partial C}{\partial y}\right) + \frac{\partial}{\partial z}\left(D_{zz}\frac{\partial C}{\partial z}\right) + S \quad \text{(S8)}$$

where S denotes source/sink terms, and D_{xx}, D_{yy}, and D_{zz} are the species concentration diffusion coefficients.

S2.5. IAQ Model Validation & Calibration

Answering the question "How accurate is the model prediction?" depends on the model objectives and the type of data available for model inputs. The degree of agreement between model outputs and actual measurements is primarily influenced by the quality of the emission source and sink models. In situations where comprehension of the emission sources, sinks, and indoor-outdoor AER is robust, the disagreement between estimated and observed pollutant concentrations is rooted in measurement errors. Accordingly, for scenarios that require the assessment of a specific source using adequate source and sink models, predicted concentrations in a range of ±100% of measured values are expected [3].

Error in indoor-outdoor AERs are less crucial compared to those present in the emission source and sink models. Notably, predictions made by the IAQ model are minimally affected by interzonal airflows [3,45].

IAQ Models Validation

Concerning the mass-balance model validation, Sparks and colleagues proposed a range of quantitative criteria for evaluating the overall agreement between model outputs and experimental data [46]. Many of those criteria are from the ASTM D5157. The initial criteria

involve calculating the absolute value of the average fractional residual between the predicted concentration and the measured or observed concentration. Additional quantitative measures for gauging the alignment between model predictions and measurements are *r* (correlation), M (regression slope), *b* (regression intercept), NMSE (normalized mean square error), FB (fractional bias), and FS (fractional bias based on variance). A model may meet one or more criteria and still be inadequate, or conversely, a model may not satisfy one or more criteria and remain adequate for a given task [3,46,47].

For the validation of CFD models, in addition to the application of MAE, there are two valuable references in which descriptive information about the model validation can be found [48,49], however simple comparison between CFD and experimental results is very common in the literature [50].

On the other hand, the evaluation of statistical models' accuracy typically involves assessing the measured and predicted outputs via various performance metrics and approaches, including but not limited to Cross-validation (rotation estimation), Accuracy & Predictive Ability (PA), Root Mean Square Error (RMSE), Pearson correlation coefficient (*r*), coefficient of determination (R^2), Mean Absolute Percent Error (MAPE), Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Mean Absolute Error (MAE), and Normalized Mean Absolute Error (NMAE) [31].

Regarding the validation of indoor chemistry models, due to the enormous heterogeneity of indoor spaces, they have to be tested against measurements in a wide range of environments. This validation step is required to ensure that model predictions are quantitatively accurate and transferable [37,51].

IAQ Models Calibration

Regarding the IAQ model calibration (particularly mass balance and CFD models), this process is simply defined as the adjustment of statistical or physical modeling elements to improve the agreement of results with experimental data [52,53]. Accordingly, one typical calibration approach is the Bayesian calibration using the Markov Chain Monte Carlo (MCMC) technique. Bayesian calibration involves an iterative procedure wherein uncertainty distributions related to the IAQ model parameters are revised in a manner that is consistent with the observed data [54,55].

Concerning the calibration of statistical models based on machine learning, a typical model needs various constraints, weights, or learning rates to effectively assist various data patterns. These attributes are termed hyperparameters, and their calibration is essential to enable the

model to proficiently address the machine learning task [56]. The process of hyperparameter optimization identifies a combination of hyperparameters that results in an optimal model, minimizing a predetermined loss function on provided independent data. Cross-validation is frequently employed to determine this generalization performance. Other popular approaches are Grid search, Random search, Bayesian optimization, Gradient-based optimization, Evolutionary optimization (e.g., Genetic Algorithm (GA), etc.), Population Based Training (PBT), Early stopping-based, Radial Basis Function (RBF), Spectral methods, and Adaptive Moment Estimation (ADAM) optimization [57-59].

S3. IAQ Model; Design in CONTAM

To provide the baseline IAQ model of the case study, an IAQ "poly-contaminant" design of the test house was developed in CONTAM. CONTAM has been widely used to study various IAQ problems [60,61]. CONTAM enables indoor air multizone modeling, in which the building design is represented as a network of zones defined by the airflow paths over various zones (i.e., doors, windows, leakages, cracks, HVAC, etc.). The network nodes describe the zones concerning variant hydrostatic pressure, continuous temperature, and pollutant concentration. CONTAM is capable of the natural ventilation processes by applying the wind pressures acting on the exterior of the building, and buoyancy effects induced by temperature differences between zones, including the outdoors. In the following, correlated studies with CONTAM application in IAQ model design are presented, briefly.

CONTAM has been dynamically linked with energy analysis software such as EnergyPlus and TRNSYS [62,63]. Temenos and colleagues investigated the IAQ of Greek apartments, using the CONTAM model [64]. Their study showed that the variation of the baseline levels of the CONTAM inputs affects the output results and the modeled health effects. Silva and colleagues, evaluated different scenarios of natural ventilation operation and the IAQ at a classroom in Portugal, using CONTAM. They performed the validation with the aid of a 24-hour experimental dataset [65]. Fine & Touchie, investigated the ventilation system retrofits of high-rise residential buildings in Canada using a CONTAM model [66]. Yang and colleagues suggested integrating the IAQ model into healthy building design by developing a simulation toolbox, named i-IAQ, via MATLAB®. For the implementation of the airflow module, they principally followed the simulation setup of CONTAM. They carried out an experimental validation period of one week [67]. Na and colleagues, optimized the IAQ and acoustic levels in old schools' classrooms with air purifiers and heat recovery ventilation systems, in South Korea [68]. Alonso and colleagues, presented a methodology for the improvement of demand-controlled ventilation using

measurements of IAQ parameters with Low-Cost Sensors (LCSs), correlation analysis, and co-simulation EnergyPlus/CONTAM, in Norway [69]. Sung and colleagues investigated the building retrofit, ventilation, and filtration measures for IAQ in a school in South Korea [70]. Their CONTAM-based models were calibrated with the measured airflow and contaminant transport variables via ASTM D5157.

Regarding our case study, the floor plan properties were introduced to the software concerning the test house characteristics (area: 100m^2, volume: 320m^3, ceiling height: 3.2m). The envelope effective leakage area is defined at a pressure of 4 Pa, exponent of 0.65, and discharge coefficient of 1.

Two exhaust fans with on-off operating modes, one in the bathroom and one in the kitchen, were considered in the design by flow rates of 24 L.s^{-1}. In total, 8 zones with 34 airflow paths (doors, windows, cracks and leakages, and exhausts) were implemented in the model based on the case study geometry. The simulations ran with the use of Outdoor Contaminant files (.CTM), Outdoor Weather files (.WTH), and Continuous Value Files (.CVF) for continuous indoor temperature, indoor emission rates, natural ventilation activity, and occupancy pattern. Figure S1. illustrates the street view of the test house and the corresponding CONTAM sketchpad of the first floor.

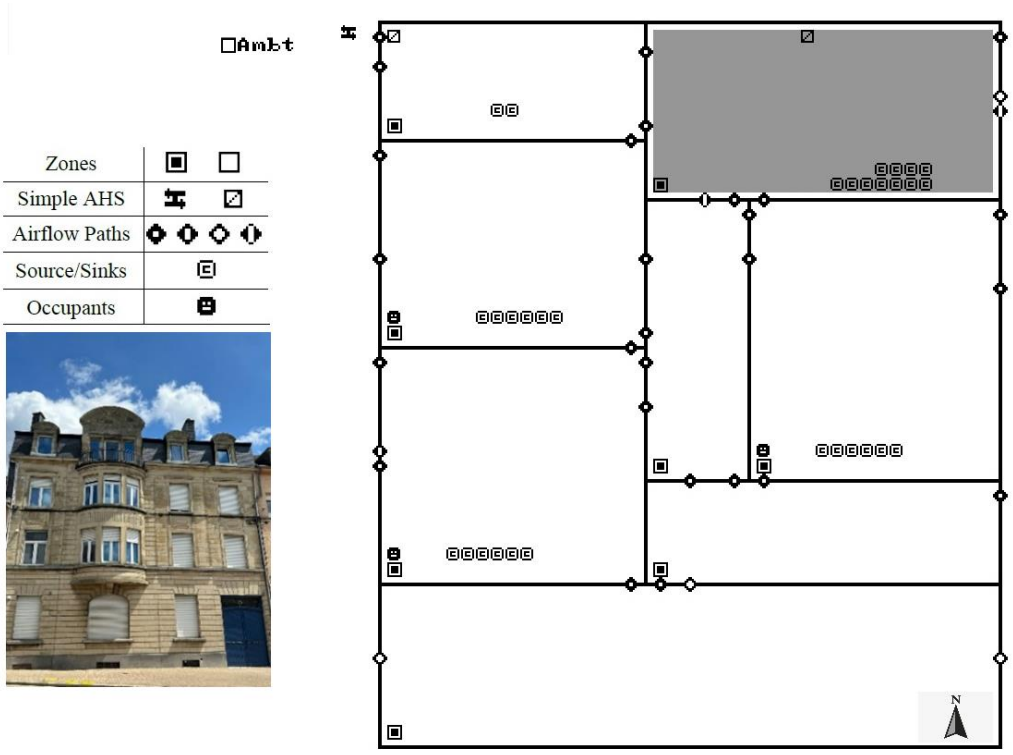

Figure S1. The test house in Arlon, and the corresponding CONTAM sketchpad of the first floor.

Table S1. shows the specified air flow paths in the model. The model parameters of different flow path elements were extracted from the ASHRAE 2015 Handbook [71].

Table S1. Properties of airflow paths in the IAQ model.

Element	Model Summary	Formula	Model Parameter
Exterior wall leakage	One-way flow using power law	Leakage area per unit length	15 cm^2/m
Interior wall leakage	One-way flow using power law	Leakage area per unit length	20 cm^2/m
Windows Close	One-way flow using power law	Leakage area per item	2 cm^2
Doors Closed (old)	One-way flow using power law	Leakage area per item	150 cm^2
Windows Open	Two-way flow	One opening	Cross section area
Doors Open	Two-way flow	One opening	Cross section area

S3.1. Indoor Emission Rates

Regarding the indoor emission rates, two approaches were carried out. In the first approach, the average indoor emission rates were extracted from the literature and introduced into CONTAM to model a random 7-day period among the whole measurement period. In the second approach, calculated continuous emission rates were introduced to CONTAM for modeling a period of about 3 summer months (73 days). These 2 approaches are described in the following.

S3.1.1. Averaged Indoor Emission Rates from Literature

In this approach, we employed reported indoor emission rates available in the literature and introduced them to the software based on the questionnaire and activity pattern of the occupants for one week (18-24 July 2021). Our motive was to obtain acceptable approximates of typical indoor emission rates of the activities in our case study. For the VOCs, even though source strengths and activity patterns will vary from case to case, an approximate volumetric steady-state indoor emission rate of 40 $\mu g/m^3.h$ was considered [72,73]. For the $PM_{2.5}$, PM_{10}, NO_2, and CO, indoor emission rates were considered following the most common source (cooking activity) from comprehensive studies. Regarding the different values reported in the literature, the median values among them have been considered, due to the variety of cooking utilities and activities (boiling, frying, grilling, toasting, and microwaving), and food ingredients. Correspondingly, the estimated $PM_{2.5}$, PM_{10}, NO_2, and CO emission rates were determined to be 1.6, 4.1, 3.1, and 5.3 mg/min, respectively [74-79].

S3.1.2. Continuous Emission Rates by Mass-Balance Approach (Calibration)

To calculate higher resolution continuous emission rates, the method based on the widely used mass balance model was employed [80-83]:

$$C(t)=C_b+\frac{\lambda}{(\lambda+k)V}+\left(C(0)+C_b+\frac{G}{(\lambda+k)V}\right)e^{-(\lambda+k)t} \quad \text{(S9)}$$

In which, $C(t)$ is the concentration at time t, C_b is the background concentration, $C(0)$ is the initial concentration, G is the emission rate, V is the mixing volume, and $(\lambda+k)$ is the total decay rate due to ventilation, deposition, and coagulation (λ is the ventilation rate and k is the deposition rate). With the assumption of a well-mixed decay period and to obtain the emission rates over an emission period T, Equation (S9) could be simplified to Equation (S10) [84], as follows:

$$G=(\lambda+k)V\left(\frac{(C_p-C_b)-(C(0)-C_b)e^{-(\lambda+k)T}}{1-e^{-(\lambda+k)T}}\right) \quad \text{(S10)}$$

In which, C_p is the peak concentration at the time t_p. Nonetheless, the substantial fluctuations in daily indoor-generated pollutant source intensities contradict the presumption of a singular total decay rate, applicable to all days. The continuous total decay rate can be quantified as the negative slope in the logarithm of the indoor concentration as a function of time [53]:

$$(\lambda+k)=\frac{1}{t_2-t_1}ln\frac{C_{t1}}{C_{t2}} \quad \text{(S11)}$$

To calculate the emission rates of gases, and with the assumption of $t \ll \tau$ (τ is the residence time), the "peak estimation approach" presented by Ott and colleagues, was applied [82]:

$$G \cong \frac{VC_{\max}}{t} \quad \text{(S12)}$$

In which, $C_{\max}$ is the maximum concentration. Finally, calculated indoor emission rates were fed to the CONTAM by ".CVF" input files. Table S2. presents the Standard Deviation (SD), average, and maximum values of calculated indoor emission rates.

Table S2. Calculated indoor emission rates by mass-balance approach, SD, Average, and the Maximum

Emission Rates (mg/min)	CO	NO_2	$PM_{2.5}$	PM_{10}
Standard Deviation	6.87	2.11	15.28	31.81
Average	2.40	1.26	2.96	8.32
Maximum	121.80	13.92	278.82	565.90

Following the instructions provided by the systematic framework (see Fig. 2), this section is dedicated to presenting the outcomes of a total of 21 simulation cases. These cases are the combinations of 2 approaches of obtaining emission rates within 7 days, and a single approach of obtaining emission rates (more accurate; mass-balance method) within 73 days, for 7 indoor contaminants.

Figure S2 shows the hourly indoor, and outdoor measured concentrations and indoor concentrations results by CONTAM for the kitchen zone (most complex zone) of the test-house, during 168 hours (7 days, 18-24 July 2021). Figure S2a presents the CONTAM results while it is fed by average event emission rates from the literature, and Figure S2b presents the CONTAM results when it is fed by calibrated continuous emission rates derived by the mass-balance approach. Figure S3 presents the same parameters as Figure S2 but for 1752 hours (73 days, 20 June-31 August 2021) with calibrated emission rates. For better presentation purposes in Figure S3, the upper concentration range is kept limited for CO, $PM_{2.5}$, and PM_{10}. However maximum incidental records of CO: 150 mg/m^3, $PM_{2.5}$: 400, and PM_{10}: 800 µg/m^3 were logged during the measurement campaign.

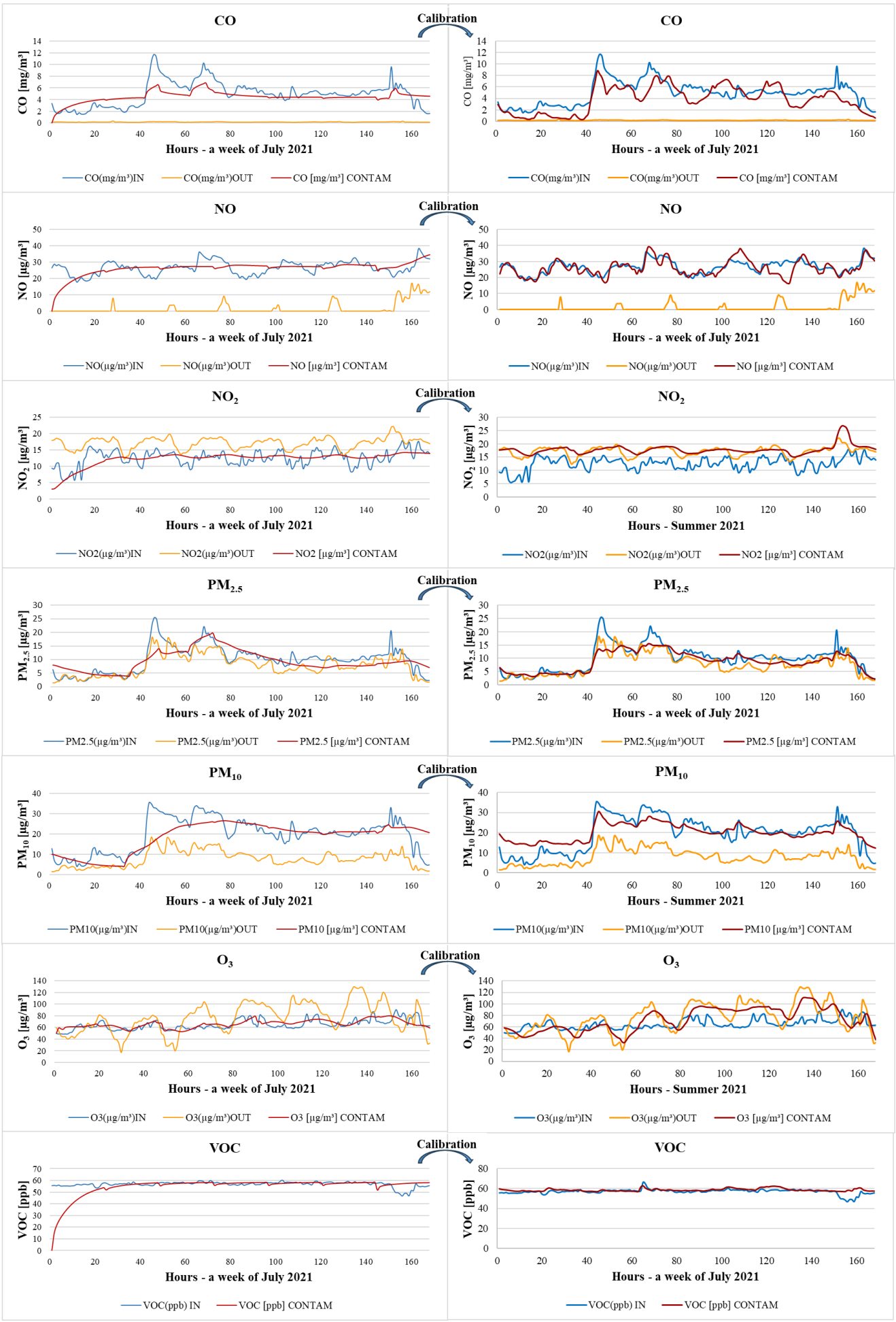

Figure S2. Hourly concentration of indoor and outdoor measurements and indoor results by CONTAM (168 hours = 7 days; 18-24 July 2021) – Case study house, a) Average emission rates b) Calibrated emission rates.

Figure S3. Hourly concentration of indoor and outdoor measurements and results by CONTAM (1752 hours = 73 days; 20 June - 31 August 2021) – Case study house, Calibrated emission rates. The green shaded area shows the equivalent time range of Figure S2.

The equivalent time range of Figure S2 is highlighted by a green shading in Figure S3. Also, the initial concentrations of pollutants were introduced to CONTAM based on the mean outdoor concentration of contaminants during the IAQ measurement campaign. As it is recognizable from Figure S3., CONTAM simulation results are in high harmony with the measured values over the 73 days. To statistically examine the model performance and the agreement between model results and indoor measurement (see Figure S3.), ASTM D5157-19 was employed (ASTM 2019). Data characteristics were sufficient to evaluate CONTAM estimates of different zone pollutant levels. In contrast to our previous investigations [85], we calculated the NMSE parameter based on the definition presented by an exclusive EPA (U.S. Environmental Protection Agency) chapter of the McGraw-Hill Indoor Air Quality Handbook [3]. The aforementioned definition is as follows:

$$NMSE = \frac{\overline{(C_p - C_o)^2}}{(\overline{C_p}.\overline{C_o})} \tag{S13}$$

in which C_o represents the measured or observed concentration and C_p represents the predicted or modeled concentration. The bar accent denotes averaged values.

Table S3. presents the results of D5157 evaluation criteria for the model output and observed (measured) datasets, both for average emission rates of the literature and calculated continuous emission rates. The results falling in the D5157 expected ranges are remarked in green cells. Also, corresponding scatter plots are presented in Figure S4.

Table S3. Results of D5157 evaluation criteria over the modeled and measured datasets.

7 days (18-24 July 2021), CONTAM average emission rates								
Acceptable range	**Measure**	**CO**	**NO**	**NO_2**	**$PM_{2.5}$**	**PM_{10}**	**VOC**	**O_3**
$r \geq 0.9$	***r***	0.39	0.48	0.46	0.74	0.8	0.35	0.62
$0.75 \leq M \leq 1.25$	**M**	**1.90**	**0.79**	**0.76**	**0.96**	**0.9**	0.13	**0.81**
b≤0.25(observations)	**b**	3.9≥1.2	**5.32≤6.6**	**3.06≤3.1**	**0.79≤2.5**	**3.22≤4.9**	50.23≥14.22	**10.87≤16.2**
NMSE ≤ 0.25	**NMSE**	**0.09**	**0.01**	**0.02**	**0.11**	**0.09**	**0**	**0.01**
FB ≤ 0.25	**FB**	**-0.09**	**0.01**	**0.00**	**-0.04**	**0.04**	**-0.01**	**0.02**
FS ≤ 0.5	**FS**	-1.44	1.06	-0.84	-0.52	**-0.31**	0.62	0.62
73 days (20June-31August 2021), CONTAM calibrated emission rates								
$r \geq 0.9$	***r***	0.74	**0.92**	0.72	0.75	0.74	0.65	0.55
$0.75 \leq M \leq 1.25$	**M**	**0.82**	**0.99**	0.41	0.38	0.37	0.6	**1.04**
b≤0.25(observations)	**b**	**1.04≤1.14**	**0≤8.13**	12.68≥5.03	4.88≥1.93	6.2≥2.97	25.16≥14.22	**0.26≤14.97**
NMSE ≤ 0.25	**NMSE**	**0.11**	**0.01**	**0.04**	**0.22**	**0.13**	**0**	**0.04**
FB ≤ 0.25	**FB**	**0.2**	**–0.01**	**0.11**	**0.03**	**0.25**	**0.04**	**0.04**
FS ≤ 0.5	**FS**	**0.33**	**0.15**	–1.1	–1.18	–1.2	**–0.15**	1.12

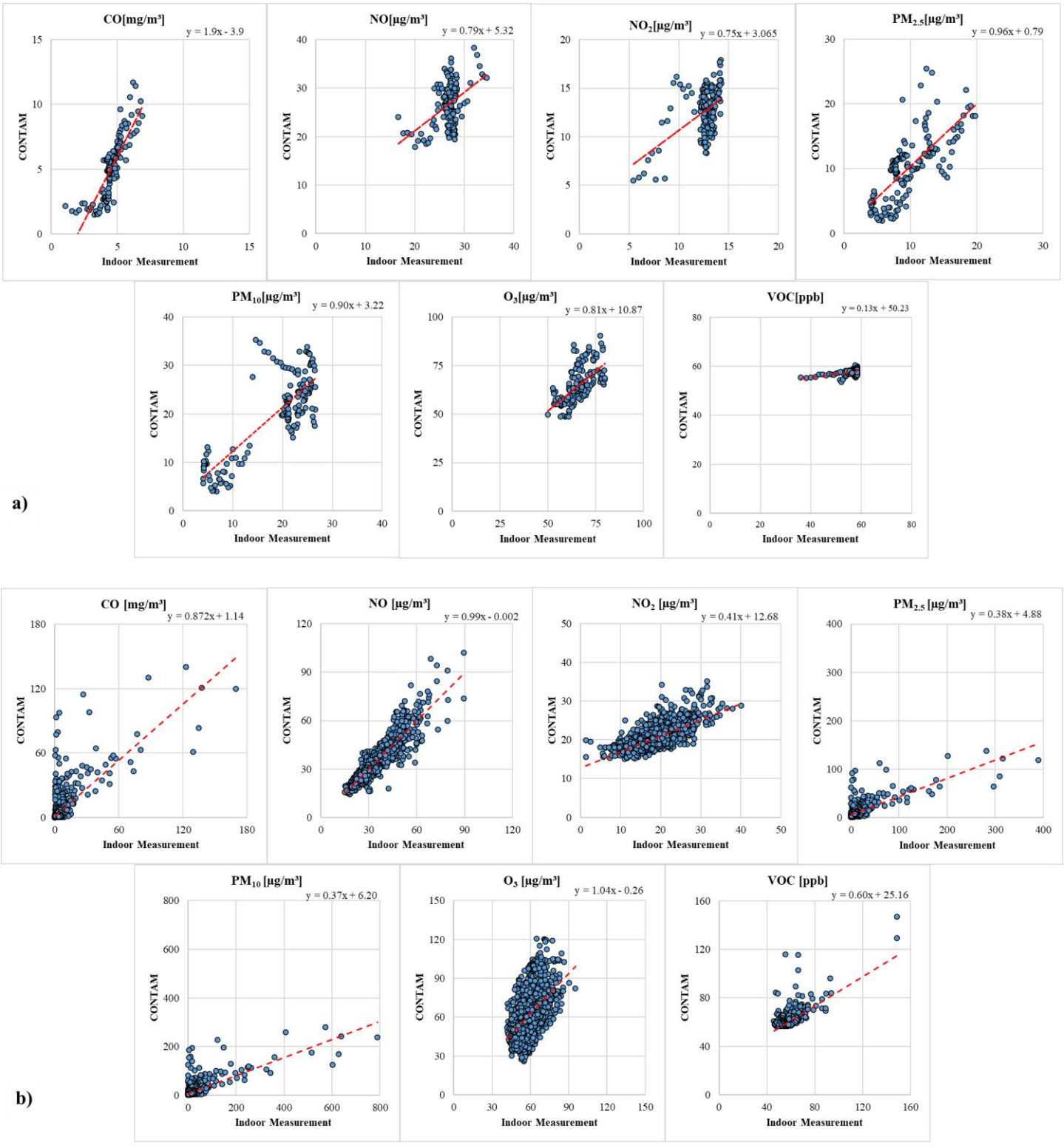

Figure S4. Scatter plots of the model concentration vs. measured values a) 7 days, 18-24 July 2021 (CONTAM averaged emission rates), b) 73 days, 20 June-31 August 2021 (CONTAM calibrated emission rates)

S4. Future Inputs for the IAQ Model

In this section, different steps for providing the future data input of the developed IAQ model (basis year 2021) are presented.

S4.1. Future Outdoor Weather (Meteorology Data)

The OCCuPANt project [86] partner has provided a historical and forthcoming weather database for dynamic building modeling in Belgium via the regional climate model "MAR" (Modèle Atmosphérique Régional) version 3.11.4 [87,88]. The database provides 13 weather variables including dry bulb temperature at 2m, RH at 2m, global horizontal radiation, diffuse solar radiation, direct normal radiation, wind speed at 10m, wind direction, dew point at 2m, atmospheric pressure, cloudiness, sky temperature, specific humidity at 2m, and precipitation. MAR is a 3D atmospheric model coupled to a one-dimensional (1D) transfer scheme between the surface, vegetation, and atmosphere. The spatial resolution of MAR is 5km atop an integration domain (120 x 90 grid cells) centered above Belgium as depicted in Figure S5 to derive hourly results. The central role of MAR is to downscale a global model or reanalysis to obtain weather data at a higher resolution of time (temporal) and space (spatial). This regional model simulates the past climate (1980-2020) and also provides various future forecasts and associated uncertainties for different scenarios based on SSP5-8.5.

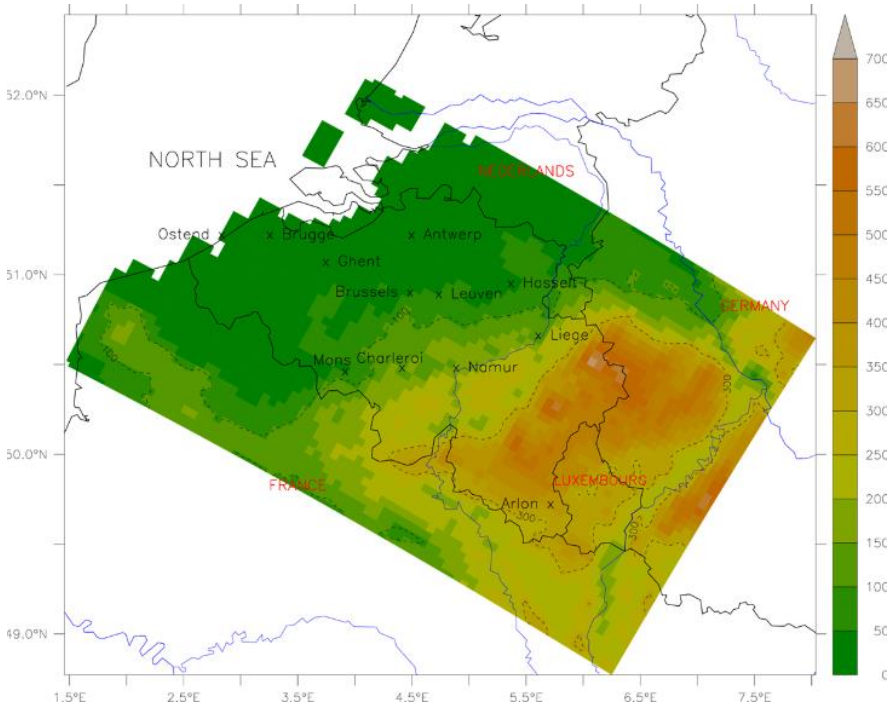

Figure S5. MAR Model topography.

S4.2. Future Outdoor Air Pollution

Applying representative CTM methods, such as CMAQ and WRF-Chem, are typical methods for urban air quality prediction. However, because of unreliable pollutant emission rates, complex underlying surface states, and inadequate theoretical groundwork, the calculated

results lack estimation accuracy. Although these approaches are helpful analysis of atmospheric dispersion, turbulent diffusion, wet and dry deposition, and decay; severe barriers to models' accuracy are still present [89-93]. The origin of error in a CTM is the unreliability of temporal variations of emission rates, though the sites of emission sources are normally detected. Mesoscale atmospheric interactions (e.g., convection, inversion) and the indeterminate wet deposition processes are significant origins of error as well. Also, the functionality of CTMs to calculate complex atmospheric photochemical reactions is partial, due to a couple of issues; such as unreliability in emission descriptions. Almost all research carried out by CTMs relies on hypothesizes or assumptions of main variables like emission rates, mixing heights, and cloudiness [89-95]. Also, other studies have been carried out to estimate the future impact of climate change on air pollution by land use models. However, these approaches are limited with high non-linearity and low accuracy efficiency [96].

S4.2.1. Artificial Intelligence

Deep Learning (DL) approach is a subdivision of Machine Learning (ML). Generally, it is a neural network with a more complex design structure. It is best to utilize multi-dimensional data efficiently due to its powerful learning ability, strong generalization, and flexible model structure. Due to its strong learning capability, powerful generalization, and adjustable architecture, it is efficient to be utilized for data with high dimensions. DL networks such as Convolutional Neural Networks, Recurrent Neural Networks (RNN), and Long Short-Term Memory recurrent networks (LSTM), as well as hybrid architectures, have been employed to estimate air pollution with high efficiency competently [19-21,97-99]. Recent studies carrying out comparisons between the accuracy performance of CTMs and AI techniques for outdoor air pollution prediction also show a better performance of ML and DL in terms of prediction accuracy [100].

S4.2.2. Deep Air-Quality Forecasting

Air quality data are high dimensional (with strong nonlinearities). In this regard, air quality prediction is an arduous task because of rapid weather variations, pollutant emission phenomena, and the presence of numerous influential elements. Additionally, the involved parameters in air quality are nonlinear and dynamic; including but not limited to wind speed and direction, solar radiation, air temperature, air humidity, as well as the pollutant concentrations themselves. All in all, air quality prediction in a complex and highly non-linear context is a challenging goal to be spatially and temporally precise. Since these elements are inherently interdependent, dealing with interdependencies and utilizing them for prediction

from multivariable time series data is not easy. To overcome these challenges, a hybrid deep learner algorithm consisting of multiplex 1D Convolutional Neural Networks and a Bi-directional Long Short-Term Memory recurrent network is developed based on the Deep Air Quality Forecasting Framework; DAQFF, in MATLAB software [101]. The CNN-LSTM deep network considers both spatial and temporal dependencies of air quality-related time series data and is explained more in detail in the following part [98, 99,101,102].

S4.2.2.1. Convolutional Neural Network (CNN)

A representative CNN consists of 3 layers (see Figure S6): convolutional, activation, and pooling layer. Dissimilar to the traditional convolutional network (classic 2D with application for images), multiples 1D filters convolved over all time steps of air quality time series data (1D-CNNs), are implemented. Also, the *ReLU* is set as the activation function. To learn regional pattern features 3 convolution layers are applied. After functioning 3 convolution layers, to alter the high-level expression to a feature vector, and employ a fully connected layer to decrease the final output vector dimension, a flattened layer is utilized. At this point, a concatenated layer delivers the final output. This enables receiving the regional pattern features of single station time series data (as the 1D filter is employed in each convolutional layer, the variation of regional pattern features, over time series can be apprehended), as well as integration of the probable spatial association features of multiple stations. Besides, the regional comprehension and weighted sharing features of the 1D convolutional network decrease the parameters for operating with multiplex time series data and lead to higher learning performance. Accordingly, with the aid of this approach, the learning takes place for more deep representation features of air quality data.

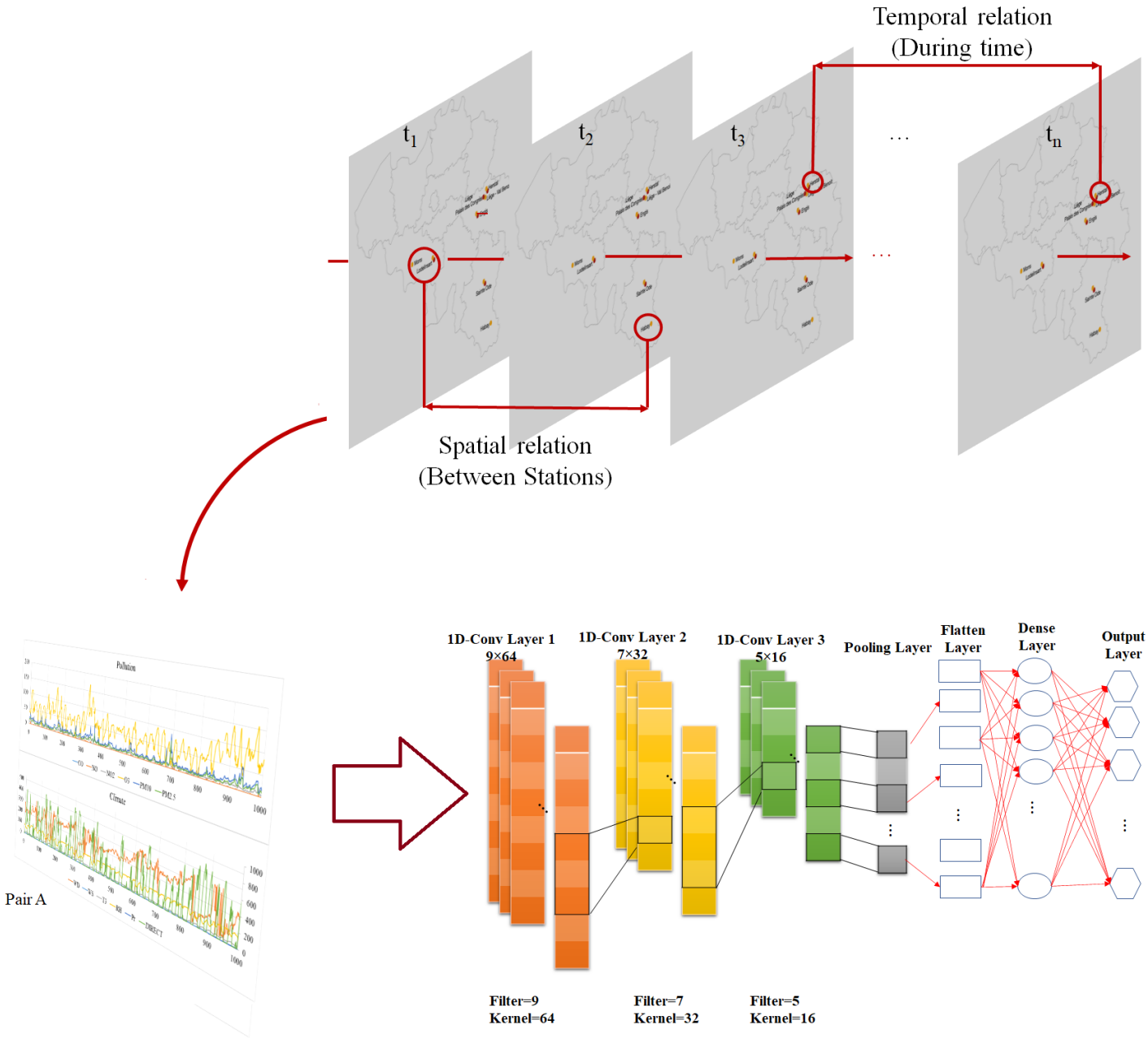

Figure S6. 1D-CNN configuration for the current study with 3 Coevolution, 1 Flatten, and 1 Dense layer.

S4.2.2.2. Bi-Directional Long Short-Term Memory Recurrent Network (Bi-LSTM)

A general LSTM segment consists of a cell with 3 gates (see Figure S7); input, output, and forget. The cell recalls values upon optional time intermissions, and the 3 gates control the input and output flow of data. Because of this specific memory cell architecture, the LSTM arrangement can consider long-term associations of time series data, and prevail over the disadvantages of typical recurrent networks (particularly the issue of gradient loss and burst). A chain of repeating cells forms the LSTM layer. Also, the *tanh* can be set for the activation function. With the aid of 2 independent hidden layers, a Bi-LSTM network can operate through

2 directions with time series data, at the same time. These data are concatenated and fed forward to the output layer. Simply put, Bi-LSTM networks are repeatedly functional with time series data in 2 directions.

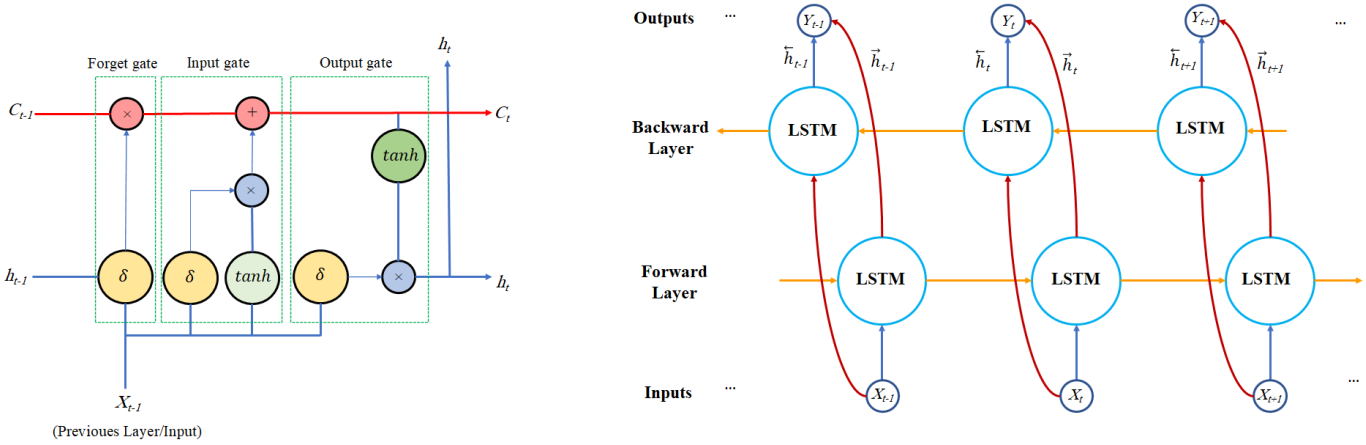

Figure S7. Bi-LSTM cell and network configuration.

S4.2.2.3. Data Preparation and Deep-Learner Network Setup

The hourly recorded air pollution data of the past 15 years (2006-2021) for the CO, NO_2, NO, $PM_{2.5}$, PM_{10}, and O_3, from 5 air quality stations in Belgium were collected [103]. VOCs were not considered during this study due to the lack of outdoor data. Accordingly, hourly weather data for the past 14 years (2008-2021) for the T, P, RH, wind speed (WS) and direction (WD), solar irradiance (IR), and precipitation (PR), were gathered [103]. The geographical (spatial) locations of the air quality stations (yellow) and weather stations (red) are illustrated in Figure S8. Information on the weather and air quality stations and their corresponding datasets are presented in Table S4. The primary data were randomly divided into calibration (80%), validation (10%), and test (10%) datasets. Also, the data of a whole month in the summer 2020 was kept blind for presenting the deep-learner network performance.

Figure S8. Locations of air quality (yellow) and weather stations (red) in Belgium with their corresponding data were employed for DL network training.

Table S4. Weather and air quality monitoring stations and their corresponding datasets.

Air Pollution stations	**Pairs**	**Weather stations** T, P, RH, WS, WD, IR, PR
Sainte-Ode (NO, NO_2, O_3, $PM_{2.5}$, PM_{10}) [07/04/2011– 18/02/2022] ***Habay (Arlon)*** (CO) [07/02/2008 – 18/02/2022]	Pair A	***Sainte-Ode (Arlon)*** [07/02/2008 – 18/02/2022]
Herstal (NO, NO_2, O_3, $PM_{2.5}$, PM_{10}) [03/01/2013 – 01/01/2020] ***Liege-Val Benoit*** (CO) [10/05/2011 – 01/01/2020]	Pair B	***Herstal*** [03/01/2013 – 01/01/2020]
Liege-Val Benoit (CO, NO, NO_2, O_3, $PM_{2.5}$, PM_{10}) [10/05/2011 – 01/01/2020]	Pair C	***Palais des congrès de Liège*** [10/05/2011 - 01/01/2020]
Engis (NO, NO_2, O_3, $PM_{2.5}$, PM_{10}) [11/02/2008 – 01/01/2020] ***Liege-Val Benoit*** (CO) [10/05/2011 – 01/01/2020]	Pair D	***Engis*** [11/02/2008 – 01/01/2020]
Lodelinsart (NO, NO_2, O_3, $PM_{2.5}$, PM_{10}) [06/02/08 - 01/01/2020]	Pair E	***Lodelinsart*** [06/02/2008 - 01/01/2020]

Also, to expedite the convergence of the DL network, and decrease the effect of outliers, a features' normalization step is performed for the raw data, to the range of [0,1] (by max-min function). To prevent the over-fitting issue, several methods were employed; such as a dropout policy with a probability of 0.3, which is utilized broadly among layers. Moreover, the early

stopping approach is applied for high-performance learning, in which a training procedure can be interrupted when the validation loss is reduced no more. Correspondingly, mean square error (MSE) was employed as the algorithm loss function representative. The hyper-parameters were initially regulated by the model performance over the validation dataset, and next, the Adam optimizer was applied. The designed CNN-BiLSTM employs singular hidden layers as default, consisting of 64 neurons. For spatial and temporal trend feature learning, 3 convolution layers, and Bi-LSTM structure with 128 hidden neurons, were applied respectively. The activation operator of the output layer is linear and is correspondingly employed for the final prediction. Missing features of experimental data are completed by the linear interpolation for single missing data points and the average value of the column in which they are placed for the remaining missing data points. Several statistical indexes, such as the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and the Pearson Correlation Coefficient (r) were applied to evaluate the performance of the proposed model. Table S5 represents the test error analysis of the CNN-BiLSTM model for the single-step prediction in paired weather and air quality stations (A-E). Convolutional networks are capable of both recursive or direct forecast tactics; where the network proceeds one-step estimation and outputs are being fed as inputs for following estimations, and where one model is established for each time-step to be estimated. Consecutively, Convolutional networks can be employed to estimate the whole output sequence, as a one-step estimation of the whole vector. This is a universal advantage of feed-forward ANN. Concerning the single-step prediction efficiency of the designed CNN-BiLSTM model, the prediction of contaminants over one blind month with untrained inputs was carried out. Figure S9 illustrates the hourly prediction performance of CNN-BiLSTM for 31 blind days of summer 2020 (no missing data completion) in the target city of Arlon for 5 contaminants (CO, NO, NO_2, O_3, $PM_{2.5}$, PM_{10}). Computational tasks were conducted on a PC server, with the AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10 GHz processor, and 16GB of memory.

Table S5. The Error analysis of CNN-BiLSTM model for the single-step prediction among paired weather and air quality stations (A-E).

CNN-BiLSTM	CO			NO			NO2			O_3			$PM_{2.5}$			PM_{10}		
	RMSE	*MAE*	*r*	*RMSE*	*MAE*	*r*	*RMSE*	*MAE*	*r*	*RMSE*	*MAE*	*r*	*RMSE*	*MAE*	*r*	*RMSE*	*MAE*	*r*
Pair A	0.86	0.05	0.57	17.43	0.42	0.64	80.82	1.94	0.67	443.7	12.9	0.87	148.4	4.65	0.63	251.7	5.94	0.61
Pair B	0.79	0.01	0.59	16.38	0.38	0.64	79.98	1.88	0.68	435.7	10.2	0.88	132.4	3.12	0.64	249.6	5.88	0.62
Pair C	0.93	0.06	0.54	19.3	0.56	0.63	82.9	2.18	0.64	466.4	14.9	0.81	152.6	3.19	0.6	263.7	6.17	0.61
Pair D	0.89	0.15	0.57	21.45	0.52	0.61	86.7	2.09	0.64	491.3	16.1	0.79	166.5	5.3	0.61	286.4	9.36	0.59
Pair E	0.96	0.17	0.55	23.1	0.62	0.6	92.06	2.18	0.62	501.6	15.7	0.76	184.4	7.04	0.59	311.5	12.1	0.57

Six future Typical Meteorological Year (TMY) weather files were selected among our database [87], to represent the temporal effects of climate change with 3 different SSP scenarios in Arlon. Additionally, for a better comparison with the past, the modeled average TMY of 2001-2020 period was taken into account, as well. The seven selected TMY weather files are as follows:

- Arlon TMY2001-2020_MAR
- Arlon TMY2041-2060_SSP2_MAR-BCC
- Arlon TMY2041-2060_SSP3_MAR-BCC
- Arlon TMY2041-2060_SSP5_MAR-BCC
- Arlon TMY2081-2100_SSP2_MAR-BCC
- Arlon TMY2081-2100_SSP3_MAR-BCC
- Arlon TMY2081-2100_SSP2_MAR-BCC

Figure S10, illustrates the box plots of the hourly predictions of outdoor air concentrations, derived from deep-learner model, for the city of Arlon, in “2000-2020”, “2050s”, and “2100s” : SSPs 2-4.5, 3-7.0, and 5-8.5.

Figure S9. Hourly prediction of Arlon's outdoor pollutant concentrations (CNN-BiLSTM) for 31 test days of summer 2020 (untrained input), Arlon.

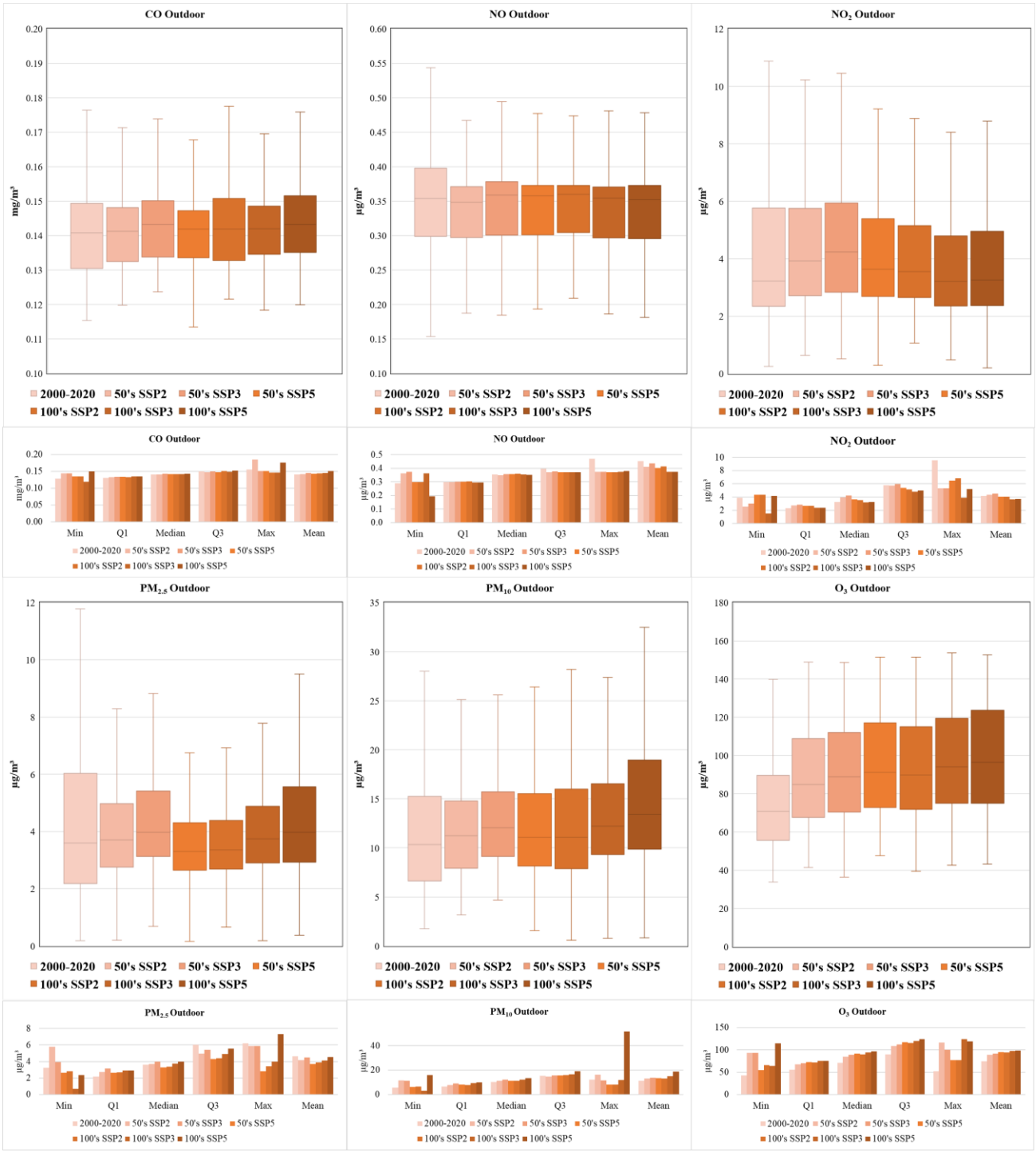

Figure S10. Outdoor pollutant concentrations of Arlon (CNN-BiLSTM) for TMYs: "2000-2020" and "2040-2060, 2080-2100: SSPs 2-4.5, 3-7.0, and 5-8.5", and corresponding statistics.

S4.3 Future Indoor Climate

For obtaining the future indoor climate; mainly the T and RH, we applied the I/O ratios recorded in the case study summer measurement campaign of 2021. These ratios are applied to the future outdoor T and RH of the selected TMYs to obtain the corresponding indoor values.

S4.4 Future Building Characteristics & Occupants Behavior

Owing to the focus of our study on indoor emission rates and ambient conditions in the context of climate change (weather and air pollution), the building characteristics and occupants' behavior patterns were assumed to be fixed. However, it is noteworthy to mention that there is a simple possibility to modify the air tightness and building air leakage information in CONTAM (reduction based on defined retrofit scenarios).

S5. Future IAQ state

After obtaining all the IAQ model inputs with their future values, we carried out the simulations to predict the future IAQ state of the case study house in CONTAM (basis year 2021). Figure S11, shows the CONTAM hourly indoor pollutant concentration estimates for CO, NO_2, NO, $PM_{2.5}$, PM_{10} and O_3, in "2000-2020", "2050s", and "2100s" : SSPs 2-4.5, 3-7.0, and 5-8.5. The estimated indoor pollutant concentrations by CONTAM, are derived from an internal integrated process of mass balance equations in which, different contributions of indoor and outdoor origins are taken into account.

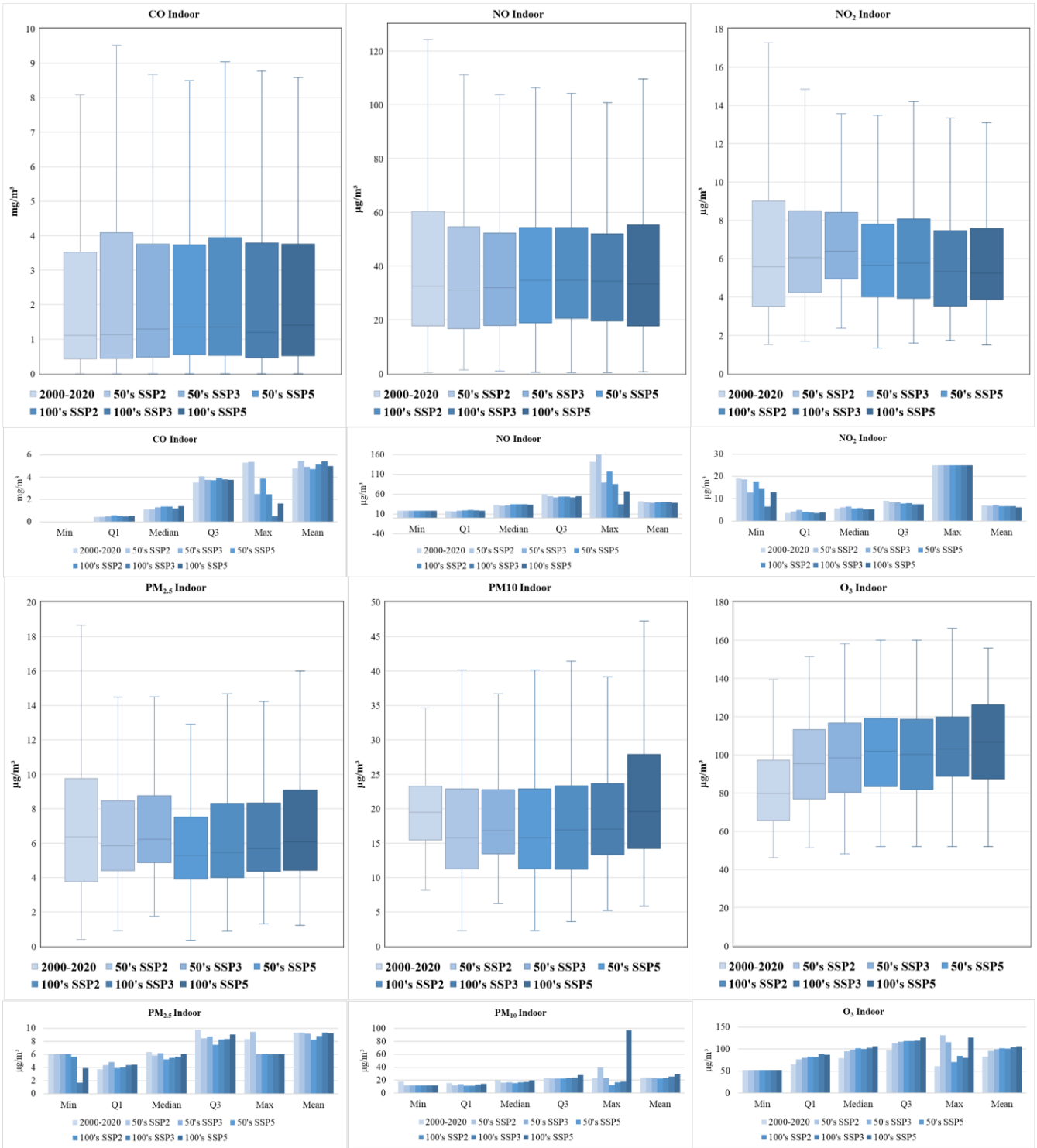

Figure S11. Indoor pollutant concentrations of Arlon test-house (CONTAM), for TMYs: "2000-2020" and "2040-2060, 2080-2100 : SSPs 2-4.5, 3-7.0, and 5-8.5", and corresponding statistics.

Considering long-term IAQ measurement campaigns with the help of low-cost sensors and establishing IAQ databases are crucial for better insight and future studies. The proposed hybrid deep-learner algorithm is capable of drawing out and learning the high-dimensional spatial and temporal features of air quality data time series in different regions. Our model performance is satisfactory since it can adopt regional pattern features by 1D-CNN, and long-term reliance features by Bi-LSTM. It is also approximated that mean indoor contaminant exposures are:

- Constant for CO (with substantial indoor sources, as indoor and outdoor emissions are not varied).
- Decreased for $PM_{2.5}$ and NO (those with periodic indoor sources that are naturally ventilated with decreased outdoor concentrations).
- Increased for NO_2 (slightly), PM_{10}, and O_3 (those with dominant outdoor sources which are naturally ventilated with increased outdoor concentrations).

References

[1] Pourkiaei, M., Falzone, C., & Romain, A. C. (2022). Exploring the Indoor Air Quality in the context of changing climate in residential buildings. Part A: developed measurement devices of low-cost sensors. In IAQ 2020: Indoor Environmental Quality Performance Approaches Transitioning from IAQ to IEQ. https://ashraem.confex.com/ashraem/iaq22/meetingapp.cgi/Paper/31952, https://hdl.handle.net/2268/292661

[2] Nazaroff, W. W., & Cass, G. R. (1989). Mathematical modeling of indoor aerosol dynamics. *Environmental Science & Technology*, *23*(2), 157-166, https://doi.org/10.1021/es00179a003

[3] Spengler, J. D., MDMS, J. M. S., & DCIH, J. F. M. S. (2001). *Indoor air quality handbook*. McGraw-Hill Education, https://www.accessengineeringlibrary.com/content/book/9780074455494

[4] Hussein, T., Kulmala, M. (2008). Indoor aerosol modeling: Basic principles and practical applications. *Water, Air, & Soil Pollution: Focus*, *8*, 23-34, https://doi.org/10.1007/s11267-007-9134-x

[5] Drescher, A. C., Lobascio, C., Gadgil, A. J., & Nazarofif, W. W. (1995). Mixing of a point-source indoor pollutant by forced convection. *Indoor Air*, *5*(3), 204-214, https://doi.org/10.1111/j.1600-0668.1995.t01-1-00007.x

[6] Norbeck, O. M. H., Sundsdal, O. M., Nambully, S. K., & Chaudhuri, A. (2022). CFD modeling of the transport of human respiratory droplets in an indoor environment. In *Proceedings of the 63rd International Conference of Scandinavian Simulation Society, SIMS 2022, Trondheim, Norway, September 20-21, 2022*. Linköping University Electronic Press, https://doi.org/10.3384/ecp192035

[7] Sedighi, A. A., Haghighat, F., Nasiri, F., Cao, S., & Ren, C. (2023). Approaches in CFD Modeling of Respiratory Droplet Dispersion–Issues and Challenges. *Sustainable Cities and Society*, 104696, https://doi.org/10.1016/j.scs.2023.104696

[8] Sajjadi, H., Tavakoli, B., Ahmadi, G., Dhaniyala, S., Harner, T., & Holsen, T. M. (2016). Computational fluid dynamics (CFD) simulation of a newly designed passive particle sampler. *Environmental Pollution*, *214*, 410-418, https://doi.org/10.1016/j.envpol.2016.04.020

[9] Xu, C., Xie, Y., Huang, S., Zhou, S., Zhang, W., Song, Y., ... & Tian, Z. (2023). A coupled analysis on human thermal comfort and the indoor non-uniform thermal environment through human exergy and CFD model. *Journal of Building Engineering*, *74*, 106845, https://doi.org/10.1016/j.jobe.2023.106845

[10] Cao, Z., An, Y., Wang, Y., Bai, Y., Zhao, T., & Zhai, C. (2023). Energy consumption of intermittent ventilation strategies of different air distribution modes for indoor pollutant removal. *Journal of Building Engineering*, 106242, https://doi.org/10.1016/j.jobe.2023.106242

[11] Liu, X., & Zhai, Z. (2008). Location identification for indoor instantaneous point contaminant source by probability-based inverse Computational Fluid Dynamics modeling. *Indoor air*, *18*(1), 2-11, https://doi.org/10.1111/j.1600-0668.2007.00499.x

[12] Zhang, T., & Chen, Q. (2007). Identification of contaminant sources in enclosed spaces by a single sensor. *Indoor air*, *17*(6), 439-449, https://doi.org/10.1111/j.1600-0668.2007.00489.x

[13] Pepper, D. W. (2009). *Modeling indoor air pollution*. Imperial College Press.

[14] Sommerfeld, M., Cui, Y., & Schmalfuß, S. (2019). Potential and constraints for the application of CFD combined with Lagrangian particle tracking to dry powder inhalers. *European Journal of Pharmaceutical Sciences*, *128*, 299-324, https://doi.org/10.1016/j.ejps.2018.12.008

[15] McLoone, M., & Quinlan, N. J. (2020). Particle transport velocity correction for the finite volume particle method for multi-resolution particle distributions and exact geometric boundaries. *Engineering Analysis with Boundary Elements*, *114*, 114-126, https://doi.org/10.1016/j.enganabound.2020.02.003

[16] Shree, V., Marwaha, B. M., & Awasthi, P. (2019). Assessment of indoor air quality in buildings using CFD: A brief review. *International Journal of Mathematical, Engineering and Management Sciences*, *4*(5), 1154, https://dx.doi.org/10.33889/IJMEMS.2019.4.5-091

[17] Kim, S. H., Lee, J. H., & Braatz, R. D. (2020). Multi-phase particle-in-cell coupled with population balance equation (MP-PIC-PBE) method for multiscale computational fluid dynamics simulation. *Computers & Chemical Engineering*, *134*, 106686, https://doi.org/10.1016/j.compchemeng.2019.106686

[18] Ranganathan, P., Pandey, A. K., Sirohi, R., Hoang, A. T., & Kim, S. H. (2022). Recent advances in computational fluid dynamics (CFD) modelling of photobioreactors: Design and applications. *Bioresource technology*, *350*, 126920, https://doi.org/10.1016/j.biortech.2022.126920
[19] Cabaneros, S. M., Calautit, J. K., & Hughes, B. R. (2019). A review of artificial neural network models for ambient air pollution prediction. *Environmental Modelling & Software*, 119, 285-304, https://doi.org/10.1016/j.envsoft.2019.06.014
[20] Masood, A., & Ahmad, K. (2021). A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: Fundamentals, application and performance. *Journal of Cleaner Production*, *322*, 129072, https://doi.org/10.1016/j.jclepro.2021.129072
[21] Balogun, A. L., Tella, A., Baloo, L., & Adebisi, N. (2021). A review of the inter-correlation of climate change, air pollution and urban sustainability using novel machine learning algorithms and spatial information science. *Urban Climate*, *40*, 100989, https://doi.org/10.1016/j.uclim.2021.100989
[22] Fard, Z. Q., Zomorodian, Z. S., & Korsavi, S. S. (2022). Application of machine learning in thermal comfort studies: A review of methods, performance and challenges. *Energy and Buildings*, *256*, 111771, https://doi.org/10.1016/j.enbuild.2021.111771
[23] Zhang, W., Wu, Y., & Calautit, J. K. (2022). A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment. *Renewable and Sustainable Energy Reviews*, *167*, 112704, https://doi.org/10.1016/j.rser.2022.112704
[24] Lala, B., & Hagishima, A. (2022). A Review of Thermal Comfort in Primary Schools and Future Challenges in Machine Learning Based Prediction for Children. *Buildings*, *12*(11), 2007, https://doi.org/10.3390/buildings12112007
[25] Kotzias, D., Geiss, O., Tirendi, S., Barrero-Moreno, J., Reina, V., Gotti, A., ... & Sarigiannis, D. (2009). Exposure to multiple air contaminants in public buildings, schools and kindergartens—the European indoor air monitoring and exposure assessment (AIRMEX) study. *Fresenius Environmental Bulletin*, *18*(5a), 670-681.
[26] Symonds, P., Taylor, J., Chalabi, Z., Mavrogianni, A., Davies, M., Hamilton, I., ... & Macintyre, H. (2016). Development of an England-wide indoor overheating and air pollution model using artificial neural networks. *Journal of Building Performance Simulation*, *9*(6), 606-619, https://doi.org/10.1080/19401493.2016.1166265
[27] Sharma, P. K., Mondal, A., Jaiswal, S., Saha, M., Nandi, S., De, T., & Saha, S. (2021). IndoAirSense: A framework for indoor air quality estimation and forecasting. *Atmospheric Pollution Research*, *12*(1), 10-22, https://doi.org/10.1016/j.apr.2020.07.027
[28] Ma, N., Aviv, D., Guo, H., & Braham, W. W. (2021). Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality. *Renewable and Sustainable Energy Reviews*, *135*, 110436. https://doi.org/10.1016/j.rser.2020.110436
[29] Tien, P. W., Wei, S., Darkwa, J., Wood, C., & Calautit, J. K. (2022). Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality–a review. *Energy and AI*, 100198, https://doi.org/10.1016/j.egyai.2022.100198
[30] Dong, J., Goodman, N., & Rajagopalan, P. (2023). A Review of Artificial Neural Network Models Applied to Predict Indoor Air Quality in Schools. *International Journal of Environmental Research and Public Health*, *20*(15), 6441, https://doi.org/10.3390/ijerph20156441
[31] Wei, W., Ramalho, O., Malingre, L., Sivanantham, S., Little, J. C., & Mandin, C. (2019). Machine learning and statistical models for predicting indoor air quality. *Indoor Air*, *29*(5), 704-726, https://doi.org/10.1111/ina.12580
[32] Li, L., Fu, Y., Fung, J. C., Tse, K. T., & Lau, A. K. (2022). Development of a back-propagation neural network combined with an adaptive multi-objective particle swarm optimizer algorithm for predicting and optimizing indoor CO2 and PM2. 5 concentrations. *Journal of Building Engineering*, *54*, 104600, https://doi.org/10.1016/j.jobe.2022.104600

[33] Sun, X., Wu, H., & Wu, Y. (2021). Probability mass functions forecasting of occupants' sensation votes under the effects of temperature, illuminance, and sound level based on ANN. *Journal of Building Engineering*, *43*, 102882, https://doi.org/10.1016/j.jobe.2021.102882
[34] Elbayoumi, M., Ramli, N. A., & Yusof, N. F. F. M. (2015). Development and comparison of regression models and feedforward backpropagation neural network models to predict seasonal indoor PM2. 5–10 and PM2.5 concentrations in naturally ventilated schools. *Atmospheric Pollution Research*, *6*(6), 1013-1023, https://doi.org/10.1016/j.apr.2015.09.001
[35] Wei, W., Sivanantham, S., Malingre, L., Ramalho, O., & Mandin, C. (2020). Predicting the rate constants of semivolatile organic compounds with hydroxyl radicals and ozone in indoor air. *Environmental Pollution*, *266*, 115050, https://doi.org/10.1016/j.envpol.2020.115050
[36] Yuchi, W., Gombojav, E., Boldbaatar, B., Galsuren, J., Enkhmaa, S., Beejin, B., ... & Allen, R. W. (2019). Evaluation of random forest regression and multiple linear regression for predicting indoor fine particulate matter concentrations in a highly polluted city. *Environmental pollution*, *245*, 746-753, https://doi.org/10.1016/j.envpol.2018.11.034
[37] Abbatt, J. P., & Wang, C. (2020). The atmospheric chemistry of indoor environments. *Environmental Science: Processes & Impacts*, *22*(1), 25-48, https://doi.org/10.1039/C9EM00386J
[38] Carslaw, N. (2007). A new detailed chemical model for indoor air pollution. *Atmospheric Environment*, *41*(6), 1164-1179, https://doi.org/10.1016/j.atmosenv.2006.09.038
[39] Sarwar, G., Corsi, R., Allen, D., & Weschler, C. (2003). The significance of secondary organic aerosol formation and growth in buildings: experimental and computational evidence. *Atmospheric Environment*, *37*(9-10), 1365-1381, https://doi.org/10.1016/S1352-2310(02)01013-0
[40] Waring, M. S. (2014). Secondary organic aerosol in residences: predicting its fraction of fine particle mass and determinants of formation strength. *Indoor Air*, *24*(4), 376-389, https://doi.org/10.1111/ina.12092
[41] Zhang, Y., Hopke, P. K., & Mandin, C. (Eds.). (2022). *Handbook of indoor air quality*. Springer, https://doi.org/10.1007/978-981-16-7680-2
[42] Waring, M. S., & Shiraiwa, M. (2022). Indoor Chemistry Modeling of Gas-, Particle-, and Surface-Phase Processes. In *Handbook of Indoor Air Quality* (pp. 955-982). Singapore: Springer Nature Singapore, https://doi.org/10.1007/978-981-16-7680-2_36
[43] Shaw, D. R., Carter, T. J., Davies, H. L., Harding-Smith, E., Crocker, E. C., Beel, G., ... & Carslaw, N. (2023). INCHEM-Py v1. 2: A community box model for indoor air chemistry. *EGUsphere*, https://doi.org/10.5194/egusphere-2023-1328
[44] Nazaroff, W. W., & Cass, G. R. (1986). Mathematical modeling of chemically reactive pollutants in indoor air. *Environmental Science & Technology*, *20*(9), 924-934, https://doi.org/10.1021/es00151a012
[45] McGrath, J. A., Byrne, M. A., Ashmore, M. R., Terry, A. C., & Dimitroulopoulou, C. (2014). A simulation study of the changes in PM2. 5 concentrations due to interzonal airflow variations caused by internal door opening patterns. *Atmospheric Environment*, *87*, 183-188, https://doi.org/10.1016/j.atmosenv.2014.01.050
[46] Sparks, L. E., Tichenor, B. A., Chang, J., & Guo, Z. (1996). Gas-phase mass transfer model for predicting volatile organic compound (voc) emission rates from indoor pollutant sources. Indoor Air, 6(1), 31-40, https://doi.org/10.1111/j.1600-0668.1996.t01-3-00004.x
[47] ASTM D5157-19. 2019. Standard Guide for Statistical Evaluation of Indoor Air Quality Models 2019, https://www.astm.org/d5157-19.html
[48] Sørensen, D. N., & Nielsen, P. V. (2003). Guest editorial: CFD in indoor air. *Indoor air*, *13*(1), 1, https://doi.org/10.1111/j.1600-0668.2003.00199.x
[49] Li, Y., & Nielsen, P. V. (2011). CFD and ventilation research. *Indoor air*, *21*(6), 442-453, https://doi.org/10.1111/j.1600-0668.2011.00723.x
[50] Zhang, T., Li, X., Zhao, Q., & Rao, Y. (2019). Control of a novel synthetical index for the local indoor air quality by the artificial neural network and genetic algorithm. *Sustainable Cities and Society*, 51, 101714, https://doi.org/10.1016/j.scs.2019.101714

[51] Weschler, C. J., & Carslaw, N. (2018). Indoor chemistry. *Environmental Science & Technology*, *52*(5), 2419-2428, https://doi.org/10.1021/acs.est.7b06387
[52] CFD Committee. (1998). Guide: Guide for the Verification and Validation of Computational Fluid Dynamics Simulations (AIAA G-077-1998 (2002)), https://doi.org/10.2514/4.472855
[53] Zhao, J., Birmili, W., Hussein, T., Wehner, B., & Wiedensohler, A. (2021). Particle number emission rates of aerosol sources in 40 German households and their contributions to ultrafine and fine particle exposure. Indoor Air, 31(3), 818-831, https://doi.org/10.1111/ina.12773
[54] Johnston, J. E., Sun, Q., & Gibson, J. M. (2014). Updating exposure models of indoor air pollution due to vapor intrusion: Bayesian calibration of the Johnson-Ettinger model. *Environmental science & technology*, *48*(4), 2130-2138, https://doi.org/10.1021/es4048413
[55] Muehleisen, R. T., & Bergerson, J. (2016). Bayesian calibration-what, why and how, https://docs.lib.purdue.edu/ihpbc/167/
[56] Claesen, M., & De Moor, B. (2015). Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*, https://doi.org/10.48550/arXiv.1502.02127
[57] Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295-316, https://doi.org/10.1016/j.neucom.2020.07.061
[58] Yu, T., & Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. *arXiv preprint arXiv:2003.05689*, https://doi.org/10.48550/arXiv.2003.05689
[59] Martínez-Comesaña, M., Eguía-Oller, P., Martínez-Torres, J., Febrero-Garrido, L., & Granada-Álvarez, E. (2022). Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost. *Sustainable Cities and Society*, *80*, 103723, https://doi.org/10.1016/j.scs.2022.103723
[60] Emmerich, S. J., & Persily, A. K. (1996). *Multizone modeling of three residential indoor air quality control options*. NIST, http://www.ibpsa.org/proceedings/BS1995/BS95_213_220.pdf
[61] Paralovo, S. L., Stranger, M., Lazarov, B., Spruyt, M., & Laverge, J. (2021). Adapted tracer gas test for passive measurement of total air change rates using alternative tracer substance. *Science and Technology for the Built Environment*, 1-15, https://doi.org/10.1080/23744731.2021.1997061
[62] Alonso, M. J., Dols, W. S., & Mathisen, H. M. (2022). Using Co-simulation between EnergyPlus and CONTAM to evaluate recirculation-based, demand-controlled ventilation strategies in an office building. *Building and Environment*, *211*, 108737, https://doi.org/10.1016/j.buildenv.2021.108737
[63] Tognon, G., Marigo, M., De Carli, M., & Zarrella, A. (2023). Mechanical, natural and hybrid ventilation systems in different building types: Energy and indoor air quality analysis. *Journal of Building Engineering*, 107060. https://doi.org/10.1016/j.jobe.2023.107060
[64] Temenos, N., Nikolopoulos, D., Petraki, E., & Yannakopoulos, P. H. (2015). Modelling of indoor air quality of Greek apartments using CONTAM (W) software. *Journal of Physical Chemistry & Biophysics*, *5*(6), 1. http://dx.doi.org/10.4172/2161-0398.1000190
[65] Silva, S., Monteiro, A., Russo, M. A., Valente, J., Alves, C., Nunes, T., ... & Miranda, A. I. (2017). Modelling indoor air quality: validation and sensitivity. *Air Quality, Atmosphere & Health*, *10*, 643-652. https://doi.org/10.1007/s11869-016-0458-4
[66] Fine, J. P., & Touchie, M. F. (2021). Evaluating ventilation system retrofits for high-rise residential buildings using a CONTAM model. *Building and Environment*, *205*, 108292. https://doi.org/10.1016/j.buildenv.2021.108292
[67] Yang, S., Mahecha, S. D., Moreno, S. A., & Licina, D. (2022). Integration of indoor air quality prediction into healthy building design. *Sustainability*, *14*(13), 7890. https://doi.org/10.3390/su14137890
[68] Na, H., Choi, H., Kim, H., Park, D., Lee, J., & Kim, T. (2023). Optimizing indoor air quality and noise levels in old school classrooms with air purifiers and HRV: A CONTAM simulation study. *Journal of Building Engineering*, *73*, 106645. https://doi.org/10.1016/j.jobe.2023.106645

[69] Alonso, M. J., Liu, P., Marman, S. F., Jørgensen, R. B., & Mathisen, H. M. (2023). Holistic methodology to reduce energy use and improve indoor air quality for demand-controlled ventilation. *Energy and Buildings*, *279*, 112692. https://doi.org/10.1016/j.enbuild.2022.112692
[70] Sung, H. J., Kim, S. H., & Kim, H. (2023). Analysis of Building Retrofit, Ventilation, and Filtration Measures for Indoor Air Quality in a Real School Context: A Case Study in Korea. *Buildings*, *13*(4), 1033, https://doi.org/10.3390/buildings13041033
[71] ASHRAE, A. H. 2015. HVAC Applications, GA. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlan, https://www.ashrae.org/about/news/2015/2015-ashrae-handbook-focuses-on-applications
[72] Holøs, S. B., Yang, A., Lind, M., Thunshelle, K., Schild, P., & Mysen, M. (2019). VOC emission rates in newly built and renovated buildings, and the influence of ventilation–a review and meta-analysis. International Journal of Ventilation, 18(3), 153-166. https://doi.org/10.1080/14733315.2018.1435026
[73] Fazli, T., & Stephens, B. (2018). Development of a nationally representative set of combined building energy and indoor air quality models for US residences. Building and Environment, 136, 198-212. https://doi.org/10.1016/j.buildenv.2018.03.047
[74] He, C., Morawska, L., Hitchins, J., & Gilbert, D. (2004). Contribution from indoor sources to particle number and mass concentrations in residential houses. Atmospheric environment, 38(21), 3405-3415. https://doi.org/10.1016/j.atmosenv.2004.03.027
[75] Dimitroulopoulou, C., Ashmore, M. R., Hill, M. T. R., Byrne, M. A., & Kinnersley, R. (2006). INDAIR: A probabilistic model of indoor air pollution in UK homes. Atmospheric Environment, 40(33), 6362-6379. https://doi.org/10.1016/j.atmosenv.2006.05.047
[76] Buonanno, G., Morawska, L., & Stabile, L. J. A. E. (2009). Particle emission factors during cooking activities. Atmospheric Environment, 43(20), 3235-3242. https://doi.org/10.1016/j.atmosenv.2009.03.044
[77] Kim, K. H., Pandey, S. K., Kabir, E., Susaya, J., & Brown, R. J. (2011). The modern paradox of unregulated cooking activities and indoor air quality. Journal of Hazardous materials, 195, 1-10. https://doi.org/10.1016/j.jhazmat.2011.08.037
[78] Zhou, S., Young, C. J., VandenBoer, T. C., Kowal, S. F., & Kahan, T. F. (2018). Time-resolved measurements of nitric oxide, nitrogen dioxide, and nitrous acid in an occupied New York home. Environmental science & technology, 52(15), 8355-8364, https://doi.org/10.1021/acs.est.8b01792
[79] Kang, K., Kim, H., Kim, D. D., Lee, Y. G., & Kim, T. (2019). Characteristics of cooking-generated PM10 and PM2. 5 in residential buildings with different cooking and ventilation types. Science of the total environment, 668, 56-66. https://doi.org/10.1016/j.scitotenv.2019.02.316
[80] Hussein, T., Korhonen, H., Herrmann, E., Hämeri, K., Lehtinen, K. E., & Kulmala, M. (2005). Emission rates due to indoor activities: indoor aerosol model development, evaluation, and applications. Aerosol Science and Technology, 39(11), 1111-1127, https://doi.org/10.1080/02786820500421513
[81] Olson, D. A., & Burke, J. M. (2006). Distributions of PM2. 5 source strengths for cooking from the Research Triangle Park particulate matter panel study. Environmental science & technology, 40(1), 163-169, https://doi.org/10.1021/es050359t
[82] Ott, W. R., Steinemann, A. C., & Wallace, L. A. (Eds.). (2006). Exposure analysis. CRC Press, https://doi.org/10.1201/9781420012637
[83] Dacunto, P. J., Cheng, K. C., Acevedo-Bolton, V., Jiang, R. T., Klepeis, N. E., Repace, J. L., ... & Hildemann, L. M. (2013). Real-time particle monitor calibration factors and PM 2.5 emission factors for multiple indoor sources. *Environmental Science: Processes & Impacts*, *15*(8), 1511-1519, https://doi.org/10.1039/C3EM00209H
[84] O'Leary, C., & Jones, B. (2017). A method to measure emission rates of $PM_{2.5}$s from cooking. In 38th Air Infiltration and Ventilation Centre Conference (Vol. 1). Nottingham UK.

[85] Pourkiaei, M., & Romain, A. C. (2022). Exploring the Indoor Air Quality in the context of changing climate in a naturally ventilated residential Building using CONTAM. In Indoor Air 2022. ISIAQ, INDOOR AIR 2022, open access link https://hdl.handle.net/2268/292660
[86] CCuPANt, 2020, https://www.occupant.uliege.be/cms/c_7968645/en/occupant
[87] Doutreloup, S., & Fettweis, X. (2021). Typical & Extreme Meteorological Year and Heat waves for Dynamic Building Simulations in Belgium based on MAR model Simulations. *Zenodo, Nov*, *9*., https://doi.org/10.5281/zenodo.5606983
[88] Doutreloup, S., Fettweis, X., Rahif, R., Elnagar, E. A., Pourkiaei, M. S., Amaripadath, D., & Attia, S. (2022). Historical and Future Weather Data for Dynamic Building Simulations in Belgium using the MAR model: Typical & Extreme Meteorological Year and Heatwaves. Earth System Science Data Discussions, 1-19, https://doi.org/10.5194/essd-14-3039-2022
[89] Vautard, R., Builtjes, P. H., Thunis, P., Cuvelier, C., Bedogni, M., Bessagnet, B., ... & Wind, P. (2007). Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the CityDelta project. Atmospheric environment, 41(1), 173-188, https://doi.org/10.1016/j.atmosenv.2006.07.039
[90] Stern, R., Builtjes, P., Schaap, M., Timmermans, R., Vautard, R., Hodzic, A., ... & Kerschbaumer, A. (2008). A model inter-comparison study focussing on episodes with elevated PM10 concentrations. Atmospheric Environment, 42(19), 4567-4588, https://doi.org/10.1016/j.atmosenv.2008.01.068
[91] Sharma, S., Sharma, P., & Khare, M. (2017). Photo-chemical transport modelling of tropospheric ozone: a review. Atmospheric Environment, 159, 34-54, https://doi.org/10.1016/j.atmosenv.2017.03.047
[92] Leelőssy, Á., Lagzi, I., Kovács, A., & Mészáros, R. (2018). A review of numerical models to predict the atmospheric dispersion of radionuclides. Journal of environmental radioactivity, 182, 20-33, https://doi.org/10.1016/j.jenvrad.2017.11.009
[93] Mirzaei, P. A. (2021). CFD modeling of micro and urban climates: Problems to be solved in the new decade. Sustainable Cities and Society, 69, 102839, https://doi.org/10.1016/j.scs.2021.102839
[94] Bernard, S. M., Samet, J. M., Grambsch, A., Ebi, K. L., & Romieu, I. (2001). The potential impacts of climate variability and change on air pollution-related health effects in the United States. Environmental health perspectives, 109(suppl 2), 199-209, https://doi.org/10.1289/ehp.109-1240667
[95] Li, X., Peng, L., Hu, Y., Shao, J., & Chi, T. (2016). Deep learning architecture for air quality predictions. Environmental Science and Pollution Research, 23(22), 22408-22417, https://doi.org/10.1007/s11356-016-7812-9
[96] Ma, J., Ding, Y., Cheng, J. C., Jiang, F., Tan, Y., Gan, V. J., & Wan, Z. (2020). Identification of high impact factors of air quality on a national scale using big data and machine learning techniques. Journal of Cleaner Production, 244, 118955, https://doi.org/10.1016/j.jclepro.2019.118955
[97] Feng, R., Zheng, H. J., Gao, H., Zhang, A. R., Huang, C., Zhang, J. X., ... & Fan, J. R. (2019). Recurrent Neural Network and random forest for analysis and accurate forecast of atmospheric pollutants: a case study in Hangzhou, China. Journal of cleaner production, 231, 1005-1015, https://doi.org/10.1016/j.jclepro.2019.05.319
[98] Li, S., Xie, G., Ren, J., Guo, L., Yang, Y., & Xu, X. (2020). Urban PM2. 5 concentration prediction via attention-based CNN–LSTM. *Applied Sciences*, *10*(6), 1953, https://doi.org/10.3390/app10061953
[99] Wang, H. W., Li, X. B., Wang, D., Zhao, J., & Peng, Z. R. (2020). Regional prediction of ground-level ozone using a hybrid sequence-to-sequence deep learning approach. Journal of Cleaner Production, 253, 119841, https://doi.org/10.1016/j.jclepro.2019.119841
[100] Fang, L., Jin, J., Segers, A., Liao, H., Li, K., Xu, B., ... & Lin, H. X. (2023). A gridded air quality forecast through fusing site-available machine learning predictions from RFSML v1. 0 and chemical transport model results from GEOS-Chem v13. 1.0 using the ensemble Kalman filter. *Geoscientific Model Development*, *16*(16), 4867-4882, https://doi.org/10.5194/gmd-16-4867-2023

[101] Du, S., Li, T., Yang, Y., & Horng, S. J. (2019). Deep air quality forecasting using hybrid deep learning framework. IEEE Transactions on Knowledge and Data Engineering, 33(6), 2412-2424, https://doi.org/10.1109/TKDE.2019.2954510
[102] Qin, D., Yu, J., Zou, G., Yong, R., Zhao, Q., & Zhang, B. (2019). A novel combined prediction scheme based on CNN and LSTM for urban PM 2.5 concentration. IEEE Access, 7, 20050-20059, https://doi.org/10.1109/ACCESS.2019.2897028
[103] ISSeP, 2021. Institut Scientifique De Service Public, Belgian Scientific Institute of Public Service https://cqaweb.issep.be