# PAPER • OPEN ACCESS

# IoT network-based ANN for ventilation pattern prediction and actuation to optimize IAQ in educational spaces

To cite this article: L C Tagliabue et al 2019 IOP Conf. Ser.: Mater. Sci. Eng. 609 042042

View the article online for updates and enhancements.

**IOP** Publishing

# IoT network-based ANN for ventilation pattern prediction and actuation to optimize IAQ in educational spaces

# L C Tagliabue<sup>1,\*</sup>, F Re Cecconi<sup>2</sup>, S Rinadi<sup>3</sup> and A L C Ciribini<sup>1</sup>

<sup>1</sup>Department of Civil Environmental Architectural Engineering and Mathematics, University of Brescia, Via Branze 43, Brescia 25123, Italy

<sup>2</sup> Deapartment of Architecture, built environment and construction engineering, Politecnico di Milano, Via Ponzio 31, 20133 Milan, Italy

<sup>3</sup> Department of Information Engineering, University of Brescia, Via Branze 38, Brescia 25123, Italy

\* lavinia.tagliabue@unibs.it

Abstract. Nowadays, in a user centered design approach, one of the main parameters for assessing the well-being of building spaces is Indoor Air Quality (IAQ), which can assure a crucial level of comfort and optimal conditions to preserve users' productivity and cognitive performance. Research works in this direction mention that with 1000 ppm of CO<sub>2</sub> concentration, a reduction of the users' cognitive performance about 11-23% is reported and, for a concentration of 2500 ppm, the decrease reaches 44-94% compared to the performance at 600 ppm. Consequently, a correct buildings ventilation is crucial. The use of mechanical systems seems possibly to avoid the problem but indeed the existing buildings often have outdated and not flexible systems to face changing needs. Thereby, the ventilation rates are not related to people density and the static setup of HVAC systems might be an issue to maintain an acceptable level of CO<sub>2</sub> concentration. Moreover, in school buildings, mechanical ventilation is not diffusely adopted and insufficient rates of fresh air supplied to the classrooms are connected with inappropriate IAQ, occurrence of SBS symptoms among pupils. Current technology provides easy measurement of CO<sub>2</sub> through dedicated sensors networks. The present research uses the pilot educational building eLUX, located in the Smart Campus of the University of Brescia, to investigate the possibility to integrate IAQ data generated by IoT sensors to improve the estimation of occupancy rate in the educational spaces. The aim is to underline the relevance of the parameter to regulate properly the HVAC systems and to define opening/closing patterns for automated windows to enhance IAQ. The data collected during the monitoring phase are useful to train an Artificial Neural Network (ANN) that through an IoT communication protocol could actuate the ventilation rate control.

#### **1. Introduction**

The wellbeing into the indoor space is a fundamental accomplishment of the buildings where users spend their everyday life nonetheless the standard indoor conditions are not a satisfaction guarantee for a twofold reason a) the standard conditions, given by the regulations, are not always respected and b) the user is not standard. Studies in the office spaces show that the users experience dissatisfaction in 10% of that cases every day and 28% every week. The possibility to interact and customize their working spaces allows the users to feel most available to extend their comfort threshold beyond the standard levels and increase the potential of the adaptive comfort given by the building behavior. This individualization principle also enhance the productivity and smartness of the workers with an individual and social benefit. The paradigm of the Cognitive Building upsurges the concept of modulation of the indoor conditions providing an automatic response to the users' feedback or preferences learned, through an IoT infrastructure [1]. In the IoT-based smart city the variability of input and output can be managed through the building digital twins. Indoor conditions are critical also in the educational buildings and researches demonstrated that insufficient ventilation [2], increased CO<sub>2</sub> concentration and decrease of IAQ [3], implies weaker results in the exams pass rate [4]. The cognitive performance in the

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under licence by IOP Publishing Ltd 1

accomplishment of learning task and knowledge restitution suffers of increased value of  $CO_2$  concentration. Many researches conveyed that ventilation rates in school buildings are often substandard, and it is not exceptional to register 3000 ppm of  $CO_2$  concentration in the classrooms. Table 1 shows the air parameters affecting the learning performance in the educational spaces [5].

Table 1	. Parameters	with standard	l and recomn	nended value	e with outcome	es on learning	performance.
---------	--------------	---------------	--------------	--------------	----------------	----------------	--------------

Parameters		Standard values	Recommended values	Outcomes	Standard	
	Indoor Air Temperature	20±2	20-25	+2-4% for each -1°C	°C	
Air	Ventilation rate	3	8	+7% from 5 to 15		
	CO <sub>2</sub> emissions	-	< 1000	+1-2.5%	ISO 7730	
	VOC emissions	-	<200	-		

#### 2. Case study

The analysis focuses on the eLUX - energy Laboratory as University eXpo [6], the demonstrator building at the Smart Campus of the University of Brescia [7]. The building, located in the Faculty of Engineering, develops the concept of cognitive building and uses the digital twin of the educational building to perform a data mapping of the data coming from the sensors installed in the classrooms. In the present study one of the two ICT laboratories in the underground floor (namely MLAB2) are adopted as case study. All the spaces, i.e. n. 2 laboratories in the underground floor; n. 2 lecture rooms in the ground floor; n. 1 auditorium in the first floor and n. 2 levels of atrium and the main eLUX lab in the underground floor, are sensorized. Data are gathered and visualized in the classroom to inform the users and mined to unveil failures and issues in the educational spaces. The Internet of Things (IoT) paradigm is used in the modeling of the data generated by the sensors and in the communication protocols used to exchange this data (based on RESTful Web Services). In such a way, it is possible a horizontal integration among the devices (sensors and actuators) of different domains (HVAC, EMS, Building Automation), and it easy to export data to machine learning applications. The ANN model is feed with data directly provided by the sensors through a dedicated IoT interface. The actuation commands (e.g. the opening of motorized windows) can be sent to the actuator on the field (e.g. motors) by means of the same IoT interface. The living lab is ongoing since October 2017.

## 3. Methodology

#### 3.1. ANN for $CO_2$ prediction

A common model used for modeling temporal data is the Recurrent Neural Network (RNN). RNNs are neural networks that are able to learn sequences that are not composed of independent, identically distributed observations. Rather, they are able to elicit the context of observations within sequences and accurately classify sequences that have strong temporal correlations [8]. In RNN, the information cycles through a loop. When it makes a decision, it takes into consideration the current input and also what it has learned from the inputs it received previously. An historical limitation of RNN is the poor performance when training models with more than 10–20 time steps [9], but this weakness can be overcome using Long Short-Term Memory (LSTM) layer. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. LSTM solves complex, artificial long-time-lag tasks that have never been solved by previous recurrent network algorithms [10]. The network used in this research is shown in Figure 1, where:

- Layer 1 (BR): is a recurrent layer that takes a vector and produces vector of size n. A recurrent layer is able to process a sequence of arbitrary length by recursively applying a transition function. The activation of the hidden states at timestamp t is computed as a function f of the current input symbol xt and the previous hidden states ht-1;
- Layer 2 (LSTM): is a layer that contains a recurrently connected memory cells and three multiplicative units—the input, output and forget gates—that provide continuous analogues of write, read and reset operations for the cells;

#### IAQVEC

IOP Conf. Series: Materials Science and Engineering 609 (2019) 042042 doi:10.1088/1757-899X/609/4/042042

- Layer 3 (LSTM): is another Long Short Term Memory layer;
- Layer 4 (SL): represents a net that takes a sequence of inputs and returns the last element of the sequence;
- Layer 5: is a fully connected net layer that computes w.x+b with output vector of size n.

Figure 1. Structure of the neural network adopted in the study.

#### 3.2. Dataset and ANN training

The initial dataset was made of 1078 records given by the sensors, 22 records each day representing a value of  $CO_2$  each half an hour from 8 am to 6.30 pm. The dataset was split into two parts, the first part containing the first 75% of the records was used to train the ANN, the second 25% was used as test.

#### 4. Data preparation

The analyses were focused on the winter (heated) period from 15/10/2017 to 22/12/2017. The whole dataset consists of 570 records of CO<sub>2</sub> measures and 17470 records of Temperature and Relative Humidity measures. Measures are registered only if a sensor is activated. Thus, data corresponding to the two sensors (n. 1 for CO<sub>2</sub> and n. 1 for T, RH) are be collected separately and at different time intervals. Consequently, some work on data must be done to make them comparable. CO<sub>2</sub>, T and RH data were interpolated to obtain a value of each parameters every 30 minutes from 8 am to 18.30 am. Figure 2, for example, shows the evolution of the measured CO<sub>2</sub> in the chosen period of time and compares it to the interpolated values.



Figure 2. Comparison between measured and interpolated data of CO<sub>2</sub> concentration in MLAB2.

# 5. Comfort and indoor air quality in winter time inside MLAB2

5.1. States of comfort conditions given by combined parameters

The comfort conditions are related to the values assumed by the main analysis parameters and thresholds to define the comfort level are introduced base on international and national standard as reported in **Table 2**. The three positions of the parameter referred to the two comfort thresholds entails n. 3 condition for each parameter (Table 2). The three states (below L1, between L1 and L2 and above L2) for the three measured parameters (T, RH and CO<sub>2</sub>) create n. 27 possible states of indoor conditions. They have been ordered and numerated starting from best-case scenario (number 1) to the worst-case scenario (number 27). The best-case scenario considers: a) both the indoor air temperature and the relative humidity in the comfort range between L1 and L2; b) CO<sub>2</sub> concentration lower than 600 ppm which means a healthy indoor air quality. The average condition (13) is defined when an issue is perceived however the condition is not strongly critical: i.e. the temperature is below 20°C (a little bit cold), the relative humidity is lower than 30% (dry air perceived) and CO<sub>2</sub> concentration is between 600 and 1000 ppm (between fresh air and lower limit of fresh air). The critical conditions are defined when temperature is higher than 22°C (too warm implies a reduction of optimal cognitive

performance), relative humidity above 45% (wet perception) and CO<sub>2</sub> concentration higher than 1000 ppm.

Table 2. Thresholds for indoor comfort conditions and IAQ: comfort range is between L1 and L2.

Parameter	Symbol	Ll	L2
Carbon dioxide concentration	$CO_2$	600	1000
Indoor air temperature	Т	20	22
Minimum Relative Humidity	RH min	35	30
Maximum Relative Humidity	RH max	40	45

#### 5.2. Indoor condition mapping

To ease the visualization of the comfort conditions and indoor air quality an hourly mean of the three measured parameters was computed and depicted in Figure 3. The horizontal axis represents the hours of the day, from 8 am to 6 pm, and the vertical axis represents the days of analyzed time span. Figure 3 shows how a)  $CO_2$  concentration, b) indoor air temperature and c) relative humidity changes during the days from 16/10/17 to 22/10/17.

#### *5.3. Prediction of the global comfort* Carefully analyzing

Figure **3**, it is possible to observe that the 27 conditions of comfort are not always present at all hours of the day. In particular, the maximum variability occurs at 12 o'clock and at 14 o'clock (Figure 3).



**Figure 3.** Data mapping for hours in the monitoring period: indoor air temperature (T), relative humidity (RH), CO<sub>2</sub> concentration (CO<sub>2</sub>) and states of the combined parameters (MLAB2 monitoring period 16/10/17 to 22/12/17).

Since the phenomenon described by the comfort condition values is not very complex, it was considered possible to construct a model to forecast the state of comfort using a traditional forecasting model such as the Markov model that is simpler and perform better compared to deep learning models in certain cases [11]. Simply stated, Markov model is a model that obeys Markov property. A stochastic process has the Markov property if the conditional probability distribution of future states of the process depends only upon the present state, not on the sequence of events that preceded it. Assuming that the sequence of measured hourly global comfort condition has the Markov property; it is possible to predict the global comfort in the next hour with a very high accuracy. For example,

given the global comfort in the first four hours of the day, from 8 am to 11 am, the predicted values of global comfort for 12 am have a Pearson's correlation coefficient  $R^2$  of more than 0.91 (Figure 4).



Figure 4. Comparison between predicted and measured global comfort.

#### 6. Results

The results of the training process are shown in **Figure 5**, on the training dataset the Pearson's correlation coefficient  $R^2$  between the measured values and the ANN forecasts is 0.93, the one computed for the test dataset is 0.88 and, eventually, the correlation between measured values and forecasts on the whole dataset is almost 0.92.



Figure 5. Structure Performance of the ANN after the training.

These very high correlations were obtained by the ANN despite the difficulty of the dataset,. If we plot the difference between the  $CO_2$  value at time  $t_{+1}$  and the value at time t on the y-axis of a graph where on the x-axis there are the time step, we see that there are big leaps between two successive measurements, some even up to 800 ppm (Figure 6a).



**Figure 6.** a) Differences between two successive CO<sub>2</sub> values; b) Comparison between the CO<sub>2</sub> concentration predicted by the trained neural network and the actual values during the test period.

These differences are due, among other reasons, to the occasional use of the MLAB2 classroom and make the use of classical techniques to predict time series like parametric methods unsuitable on this dataset. Another view of the prediction capabilities of the trained ANN is given in Figure 6b. It shows a comparison between the  $CO_2$  concentration in ppm predicted by the trained neural network and the actual values during the test period, i.e. the period of time when the data in the test dataset were collected.

## IAQVEC

IOP Conf. Series: Materials Science and Engineering **609** (2019) 042042 doi:10.1088/1757-899X/609/4/042042

The mean square error (MSE) in the test period is around 75 ppm against an average  $CO_2$  measured of 712 ppm, thus the MSE is more or less 10.6% of the average  $CO_2$ .

# 7. Conclusions

Daily users are struggling with harsh comfort condition in the existing educational facilities and adaptive behaviour can reduce the commonly reported dissatisfaction that is related for the 35% to indoor air temperature and 63% to CO<sub>2</sub> concentration higher than 1000-1500 ppm (stuffy air perception). The possibility to calibrate the ventilation according to occupancy level and when comfort parameters such as temperature due to external factors (i.e. solar radiation) grows, could promote IAQ, users' learning performance and building energy saving. Using heat recovery systems coupled with HVAC implies a 70% thermal losses decrease due to higher ventilation rates required when 1000 ppm of CO<sub>2</sub> are exceeded and automated systems can preserve the IAQ and the learning and cognitive performance of the students. The main problem is to face the outdated systems in existing buildings while the IoT network development is not intrusive in the BACS implementation for real-time energy management.

# References

- [1] Bianchini D, De Antonellis V, Melchiori M, Bellagente P, Rinaldi S, Data management challenges for smart living, Lecture Notes of the Institute for Computer Sciences, *Social-Informatics and Telecommunications Engineering*, LNICST, 2018, pp. 131-137.
- [2] Bakó-Biró Zs., Clements-Croome DJ, Kochhar N, Awbi HB, Williams M J, Ventilation rates in schools and pupils' performance, *Building and Environment* 48 (2012) 215e223.
- [3] Szczurek A, Maciejewska M, Pietrucha T, CO<sub>2</sub> and volatile organic compounds as indicators of IAQ, *6th AIVC Conference*, *5th TightVent Conference*, 3rd venticool Conference, Madrid, Spain 23–24 September 2015.
- [4] Mendell M J, Eliseeva E A, Davies M M, Lobschied A, Do classroom ventilation rates in California elementary schools influence standardized test scores? Results from a prospective study, *Indoor Air*, 26, 546–557 (2016).
- [5] Myhrvold AN, Olsen E, Lauridsen O. Indoor environment in schools pupils health and performance in regard to CO<sub>2</sub> concentrations. *Proceedings of the 7th International Conference on Indoor Air Quality and Climate -Indoor Air* 1996 1996;4:369e74.
- [6] Tagliabue LC, Manfren M, Ciribini ALC, De Angelis E, Probabilistic behavioural modelling in building performance simulation – the Brescia eLUX lab, *Energy and Building*, 128 (2016) 119-131.
- [7] Re Cecconi F, Manfren M, Tagliabue LC, Cirbini ALC, De Angelis E, Probabilistic behavioural modeling in building performance simulation: A Monte Carlo approach, *Energy and Buildings*, 148, 128-141.
- [8] Graves A, Neural networks, Supervised Sequence Labelling with Recurrent Neural Networks, *Springer* (2012), pp. 15-35.
- [9] Hefron RG, Borghetti BJ, Christensen JC, SchubertKabban CM, Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation, *Pattern Recognition Letters*, Volume 94, 2017, Pages 96-104.
- [10] Nochreiter S., Schmidhuber J., 1997. Long Short-Term Memory. *Neural computation*, 9 (8), pp. 1735-1780.
- [11] Panzner M, & Cimiano P, 2016. Comparing Hidden Markov Models and Long Short Term Memory Neural Networks for Learning Action Representations. In P. M. Pardalos, et al., eds. *Machine Learning, Optimization, and Big Data: Second International Workshop*, MOD 2016, Volterra, Italy, August 26-29, 2016. Revised Selected Papers. Lecture Notes in Computer Science. no.10122 Cham: Springer International Publishing, pp. 94-105.