Predictive Models of Indoor Carbon Dioxide Concentration to Prevent Daily Decay of Productivity and Well-Being in Shared Offices

Hamed S. Alavi[1], Sailin Zhong[1], and Denis Lalanne[1]

University of Fribourg, Fribourg 1700, Switzerland
{hamed.alavi,sailin.zhong,denis.lalanne}@unifr.ch

Abstract. With the ultimate objective of enhancing health and well-being in shared office spaces, we have developed real-time predictive models of indoor CO2 level and then examined alternative user interaction scenarios with the prediction outcome. The specific aim is to persuade office workers to take timely preventive actions when a high level of carbon dioxide concentration is forecast. Through this novel approach that seeks to prevent the situations of discomfort, rather than rectifying them, we seek to reduce the daily exposure of office workers to poor air quality in shared spaces, which commonly harms their health and cognitive performance. Among the forecast models that we implemented, the best performing model provides 97.66% accuracy for predicting CO2 level 5 minutes ahead of time and 87.51% accuracy for predicting of 20 minutes. In the second step, we elicited the different parameters of the prediction that need to be communicated with the users (e.g. certainty, time lag, salience) and examined the space of possibilities for designing user interaction scenarios. We discuss these scenarios in this paper; user studies are being conducted (ongoing work) to evaluate the effectiveness of alternative interaction scenarios.

Keywords: Smart Environments, Human-Building Interaction, Well-being, Shared Workspace

1 Introduction

Indoor environmental conditions can have far-reaching implications for our health and well-being since we spend a substantial part of our time (90% in developed countries) inside buildings. More specifically, the quality of air indoors has been subject of numerous studies, which present evidence for direct links between relatively high levels of carbon dioxide concentration and problems such as fatigue, lethargy, headache, and cardiac arrhythmia as well as difficulty in retaining attention, concentration, and cognitive performance [3][4][10][11][12][14]. This is a commonly reoccurring situation in shared closed spaces. For example, a recent study of 100 schools in Switzerland has shown that in two-third of classrooms the level of CO2 falls in the range that can have a negative impact on the students learning capacity [17]. Also, a recent study revealed that in the UK poor environmental conditions in offices account for loss of productivity equivalent of 13B pounds every year [18].

Research in this area builds on the body of literature under the banner of Indoor Environmental Qualities (IEQ), where standards and quality norms have been developed to define the acceptable range of human comfort parameters and architectural design guidelines have been proposed to ensure those norms are met (e.g. [9]). However, over the past decade, as our buildings increasingly incorporate context-aware automation and new forms of interactivity, special attention has been payed to the role of humans and artificial intelligence embodied by buildings as well as the interaction between these two [16][19][2]. It has been widely acknowledged that the future of our experiences within buildings is bound with advances in artificial intelligence and particularly how intelligent components embodied by built environments can create a dialogue with users [1][5][15].

In our research, we pursue a similar perspective on the question of human comfort and well-being with a focus on the specific problem of high CO2 concentration in shared work spaces. We examine a novel approach, in which the objective is to predict and prevent the unhealthy situations rather than rectifying them after they occurred. Two reasons motivate this preventive approach: (1) recovering from discomfort exhausts both time and cognitive effort, (2) regular exposure to situations that are even mildly uncomfortable can harm health in long-term. We, therefore, needed to develop models that can accurately predict the level of carbon dioxide in office spaces, and to design interaction scenarios with the prediction results that can support the users take a preventive action. The predictive models and the comparison of their performance are described in Section 2, and four scenarios for interaction with the prediction results are presented in Section 3.

2 Prediction of CO2 Concentration Level

Human exhale is the main source of carbon dioxide indoors and its concentration in the air is highly impacted by the number of people in the room. Nevertheless, the evolution of CO2 level is also influenced by many other parameters such as room size, ventilation rate, relative humidity, as well as outdoor air quality. Since capturing all of these variables entails heavy instrumentation of the place and the people, the predication model that we seek to develop should function independently of their variations. More precisely, our objective in this step is to develop and compare algorithms that can predict at real-time whether or not the CO2 level in an office will exceed a threshold solely on the bases of previous measurements of CO2 in the same office.

In this section, we present the examined predictive algorithms, describe how the training and test datasets were collected, and report the accuracy and computation time of each algorithm.

2.1 Predictive Models

Due to the nature of CO2 variation being highly sensitive to the momentary indoor environmental conditions, we limit our exploration to the forecasting models that can be trained using only the values recorded during a recent past interval (up to 60 minutes prior to the current moment). We refer to this interval as "Observation Buffer". Variants of short-term forecasting methods for time-series were applied, namely, Autoregressive (AR) and Autoregressive Integrated Moving Average (ARIMA). In addition, we examined Long Short-Term Memory (LSTM)—a recurrent neural network architecture—in which the problem was formulated as a multiple parallel input and multi-step output case. Precisely, the following models were implemented:

- **ARIMA**: `AutoARIMA` function available in the python library `pmdarima` was used to train the model from the first observation buffer to identify the optimal values of ARIMA parameters *(p, d, q)* [1]. For every observation buffer afterwards, a new ARIMA model was created with the saved parameters using the `ARIMA` function from the same python library.
- **ARIMA Log**: Model configuration remained the same as the ARIMA described above except that the log-transformed observation buffers were used to train the model.
- **ARIMA with update**: `AutoARIMA` function was used to train the model once for every 10 observation buffers. For the remainder of the observation buffers, the ARIMA model was updated using the `update` function of `AutoARIMA` when a observation was available.
- **ARIMA Log with update**: Model configuration remained the same as the "ARIMA with update" except that the log-transformed observation buffers were used to train the model.
- **ARIMA with new parameters**: A new ARIMA model by `AutoARIMA` was created for every observation buffer.
- **ARIMA Log with new parameters**: Model configuration remained the same as the "ARIMA with new parameters" except that the log-transformed observation buffers were used to train the model.
- **AR**: Autoregressive algorithm using function `AR` from python library `statsmodels` was applied on an observation buffer to build the model.
- **AR Log**: Model configuration remained the same as the AR except that the log-transformed observation buffers were used to train the model.
- **LSTM**: Data from the four sources were used as parallel input, and the output was a multi-step prediction. Basic Encoder-Decoder LSTM architecture with `adam` as the optimizer and mean squared logarithmic error as the loss function were applied.

The Python source code implementing the above algorithms are available in a git repository: [link anonymized]

[1] p: the number of autoregressive terms; d: the number of nonseasonal differences needed for stationarity; q: the number of lagged forecast errors in the prediction equation.

2.2 Data Collection

Data was collected using four devices deployed in a shared office, which recorded the CO2 concentration in the air every five seconds. CO2 concentration is measured as the number of CO2 molecules in one million molecules of air, and is expressed as part per million (ppm). For the measurement, we used state-of-the-art technology that operates based on an optical gas molecule detection method termed as *non-dispersive infrared spectroscopy*. The data collection was done during Autumn time in a Central European country. The data was divided into observation buffers for training the AR and ARIMA models, while LSTM was trained using data of an entire day.

2.3 Evaluation of Predictive Models

To assess the performance of the models, we consider two criteria: (1) the accuracy of prediction compared against the actual values, and (2) the training time, as the algorithms may be required to run several times during a short time interval.

Accuracy of a prediction is computed as the percentage of predicted values that fall within the confidence interval around the actual value (confidence interval is fixed to ±30ppm, based on the technical error range of the sensor). The overall accuracy of a model, as we report in the next section, is the average accuracy of 464 instances of prediction executed on one day of data (four devices, 12 times prediction per hour, 10 hours, and excluding the last 20 minutes).

Regarding the training time, the performance of AR and ARIMA were measured as they were run on a Mac Pro with 2.7 GHz 12-Core Intel Xeon E5 processor. LSTM was run on Google Colab GPU using Keras. For the case of LSTM, since the training needs to be done only once using the four parallel inputs, its overall performance cannot be compared to the other models based on the training time. Therefore, in our comparitive analysis, we consider only the accuracy of LSTM.

Table 1. Performance of the models with 20 minutes observation buffer and five minutes prediction interval (ARIMA/AR); 10 hours observations and 10 hours predictions (LSTM)

Model	Accuracy (%)	Training time (s)
ARIMA	97.62	73.7486
ARIMA Log	96.87	80.9518
ARIMA update	97.66	7.9522
ARIMA Log update	93.70	16.6990
ARIMA new param	97.63	73.7192
ARIMA Log new param	96.87	80.9293
AR	97.66	0.0029
AR Log	97.66	0.0026
LSTM	65.46	53.4 (per epoch)

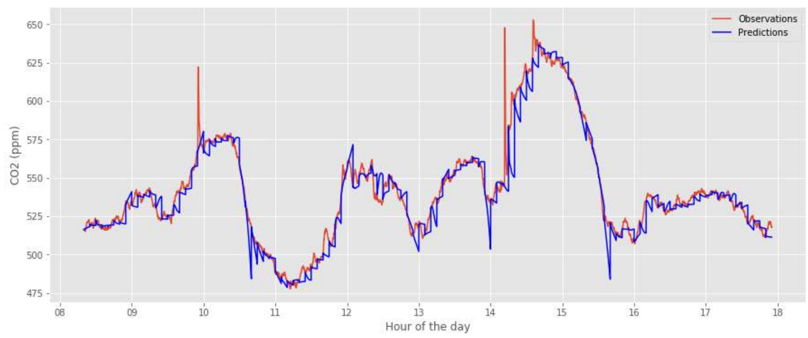

Fig. 1. CO2 prediction with 20 minutes observation buffer and five minutes prediction interval by AR model

2.4 Results

Combinations of observation buffer sizes of 10 and 20 minutes, and prediction interval of 5, 10, and 15 minutes were tested. In all of the performed tests, AR outperformed the other methods both in terms of accuracy and training time. The second highest performing candidate was "ARIMA with update".

Table 1 shows the performance of the models for an example setup where the observation buffer size is 20 minutes and the prediction interval is 5 minutes, and Figure 1 shows the prediction results versus observations during one day by AR model.

3 Interaction with Indoor Co2 Forecast

Once a high level of CO2 is predicted, it should be communicated to the user. This is an interaction design challenge that we begin to explore in this section.

3.1 Objective

As previously described, the ultimate objective here is to prevent the situations in which a high level of CO2 harms the occupants' health and productivity. This can be accomplished by either inputting the prediction results to the building ventilation system, or by informing the occupants and supporting them to take preventive action. In this work, we favor the latter, for two major reasons: (1) many buildings do not provide a standard digital interface to their ventilation system, and (2) by keeping the users informed and engaged, we may be able to increase their general awareness, leading to change of routine and behavior also in places where such sensing and predictive technology is not available. Therefore, we seek to design interaction scenarios that can motivate the users to carry out a timely preventive action (e.g. opening a window, taking a break, reorganizing the location of collaborative tasks). On the other hand, to ensure that the system will be accepted by the users in long-term, it must be designed to be unobtrusive, and be able to fade in the background of the user's attention.

3.2 Elements of Prediction

In this section, as we shift our focus from the technical description of predictive models to the human interactive experiences, we also change the perspective of our discussion and use a different nomenclature. From the user perspective, three concepts can be elicited from the CO2 forecast: the *Risk* itself including its nature and significance, the *Temporal Proximity* of the risk, and the *Certainty* that it will occur. Mapping these concepts with the output variables of the prediction models, one can see a link between Risk and the value of prediction, between Temporal Proximity and the prediction interval, as well as between Certainty and prediction accuracy. It is, however, worth noting that, depending on the interpretation of these concepts they can be overlapping or inter-related; it is not required that all of them be included in the presentation of a forecast; and they might be influenced by factors other than the outcome of the predictive model (subjective and contextual interpretation of their presentation). In the following, we first elaborate on possible interpretations of these three concepts and then explore how they may be integrated into interactive scenarios.

Risk. The Risk can be simply a trigger notifying that the CO2 level is predicted to exceed the standard threshold for healthy indoor air quality. Alternatively, it can be the precise value of the predicted CO2 concentration level, or the range within which the predicted value falls (e.g. comfortable, slightly uncomfortable, uncomfortable, intolerable).

Temporal Proximity. Temporal Proximity—defined as the time between the present moment and the time of prediction—is not a constant value. Once it is predicted that after a certain time interval the level of CO2 reaches a threshold, this interval reduces as time goes by. In addition, more data is fed to the model and thus the prediction is updated. An alarming situation forecast of 20 min, may not be valid after a minute to be shown as forecast of 19 minutes. This should be especially considered as we expect the users to take preventive actions (e.g. opening the window) which can change drastically the new predictions. How can we display the dynamic nature of Temporal Proximity while retaining the coherency of prediction presentation?

Certainty. The certainty of prediction may be defined in various ways and involve several factors, for example, the expected accuracy of predication, the previous prediction results, and the graveness of the situation (i.e. the difference between the predicted CO2 level and the standard threshold). How can we combine these parameters into a representation of Certainty that is most understandable for the users considering the common perception of certainty in everyday human interaction?

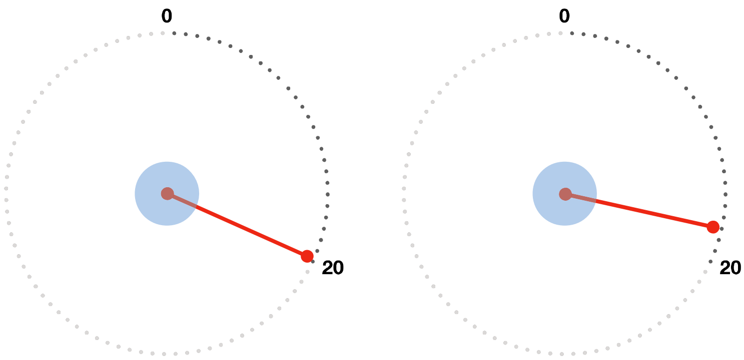

Fig. 2. Scenario I: the red line moves counterclockwise indicating the remaining time until when the CO2 level will be higher than the standard threshold.

3.3 Scenarios

In order to exemplify the ways in which Risk, Temporal Proximity, and Certainty can be translated to the constituents of interaction scenarios between users and a system that predicts the level of CO2 we propose four alternative visualizations. The ultimate objective is to prototype these scenarios and conduct user studies to compare their impact on the users' perception of environment and behaviour change. The implementation of the scenarios may manifest as dynamic visualization of forecast outcome on a personal interactive device placed on the user's desk. It could alternatively be displayed on a private screen (smartphone) or a public display in the shared office or meeting room.

Scenario I In the first scenario, only the Temporal Proximity and the most basic version of Risk are displayed. Once a high level of CO2 is predicted, a red line (resembling a clock arm) appears on a round screen and then moves counterclockwise with time. By examining such a minimal design, we aim to study whether the mere knowledge about when the situation of discomfort will occur can support the user to decide about its urgency. If the user takes a preventive action that leads to a drastic decrease in CO2 level the clock arm will disappear in a way that the user understands it is a reward for her action (e.g. moving clockwise and fading away).

On the round screen (resembling a clock), the default active part is shown as 20 minutes. The user can change this by turning the knob shown as blue circle inside the clock. The color of this knob can indicate the quality of air outdoors, depending on which opening the window may or may not be a good idea. The interaction with this knob is common across the four scenarios.

Scenario II Similar to the first scenario, a red clock arm indicates the Temporal Proximity. In scenario II, however, the Certainty of prediction is illustrated by either the length of the clock arm or the degree of its color opacity.

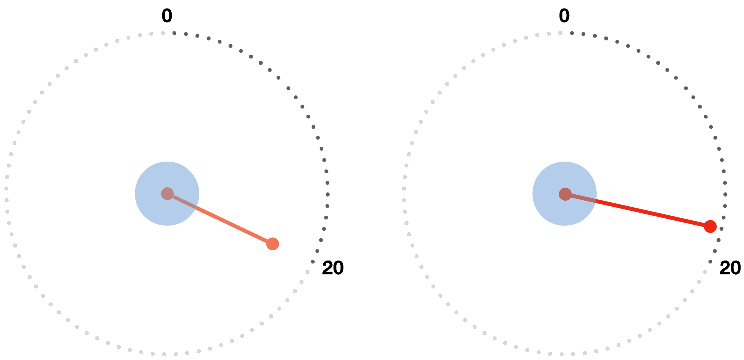

Fig. 3. Scenario II: the red line moves counterclockwise indicating the remaining time until the CO2 level will be higher than the standard threshold. The length of the red line shows the Certainty of prediction.

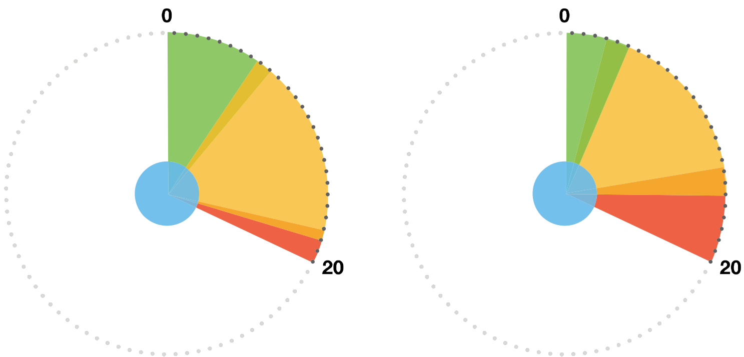

Fig. 4. Scenario III: the red, yellow, and green colors indicate if the predication of CO2 for that time is within the uncomfortable, borderline, or comfortable range, respectively.

Scenario III The Risk is presented in ranges and coded with a spectrum of colors between red to green. In this scenario, predication algorithm is run for every minute between 0 to what is set as active part of the clock (20 minutes by default). Certainty is not displayed in this scenario.

Scenario IV Certainty is added to the third scenario, illustrated by how far the areas are extended towards the border of the clock.

By presenting four imaginable scenarios, we aimed to demonstrate some of the ways in which the concepts of Risk, Temporal Proximity, and Certainty can be embedded into a representation of forecast results. Moreover, we will utilize these and similar scenarios to answer user experience design questions such as the followings: which design elements in these scenarios can impact the user engagement, trust, and long-term acceptability? How can the accuracy of

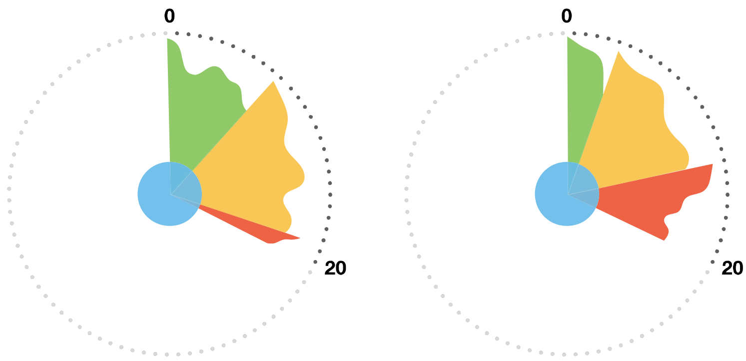

Fig. 5. Scenario IV: the colors indicate the range to which the predication of CO2 level corresponds, similar to the third scenario. The fluctuation corresponds to the change in the Certainty of prediction.

prediction translate to a contextualized representation of Certainty relevant to the user's daily communication experiences? How can we create the feeling of support and companionship for the users and avoid increasing distress?

4 Conclusion and Future Work

In this paper, we compared nine models of short-term indoor CO2 concentration prediction. The highest performing model, Autoregressive (AR), achieved 97.66% accuracy of prediction 5 minutes ahead of time and 87.51% of 20 minutes. We also briefly discussed the translation of forecast results from modeling space to user interaction space, and proposed four interaction scenarios that can dynamically display the outcome of prediction.

In the future development of this work, in addition to running user studies to investigate the questions mentioned in the last section, we will address some of the limitations in data collection and prediction. For example, we improve the predictive models by taking into account the noise in the data due to the inaccuracy of sensors [8]. We also aim to collect data from different office setups to cover larger variety of CO2 measurements with samples of extreme conditions.

References

1. Ghaffarianhoseini, A., Berardi, U., AlWaer, H., Chang, S., Halawa, E., Ghaffarianhoseini, A., Clements-Croome, D. (2016). What is an intelligent building? Analysis of recent interpretations from an international perspective. Architectural Science Review, 59(5), 338-357.
2. Meurer, R. S., Fröhlich, A. A., Hübner, J. F. (2018, October). Ambient Intelligence for the Internet of Things Through Context-Awareness. In 2018 International Symposium on Rapid System Prototyping (RSP) (pp. 83-89). IEEE.

3. Erdmann, C. A., Steiner, K. C., Apte, M. G. (2002). Indoor carbon dioxide concentrations and sick building syndrome symptoms in the BASE study revisited: Analyses of the 100 building dataset.
4. Burge, P. S. (2004). Sick building syndrome. Occupational and environmental medicine, 61(2), 185-190.
5. Snow, S., Oakley, M., schraefel, M. (2019, June). Performance by design: supporting decisions around indoor air quality in offices. In Proceedings of the 2019 on Designing Interactive Systems Conference (pp. 99-111).
6. Satish, U., Mendell, M. J., Shekhar, K., Hotchi, T., Sullivan, D., Streufert, S., Fisk, W. J. (2012). Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance. Environmental health perspectives, 120(12), 1671-1677.
7. Siami-Namini, S., Tavakoli, N., Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1394-1401). IEEE.
8. Segù, M., Loquercio, A., Scaramuzza, D. (2019). A general framework for uncertainty estimation in deep learning. arXiv preprint arXiv:1907.06890.
9. Frontczak, M., Wargocki, P. (2011). Literature survey on how different factors influence human comfort in indoor environments. Building and environment, 46(4), 922-937.
10. Griffiths, M., Eftekhari, M. (2008). Control of CO2 in a naturally ventilated classroom. Energy and Buildings, 40(4), 556-560.
11. Fisk, W. J. (2010). CO2 monitoring for demand controlled ventilation in commercial buildings.
12. Apte, M. G., Erdmann, C. A. (2003). Indoor carbon dioxide concentrations, VOCs, environmental sensitivity associations with mucous membrane and lower respiratory sick building syndrome symptoms in the BASE study: Analyses of the 100 building dataset.
13. Fisk, W. J., Satish, U., Mendell, M. J., Hotchi, T., Sullivan, D. (2013). Is CO2 an indoor pollutant? Higher levels of CO2 may diminish decision making performance. ASHRAE JOURNAL, 55(LBNL-6148E).
14. Seppanen, O. A., Fisk, W. J., Mendell, M. J. (1999). Association of ventilation rates and CO2 concentrations with health andother responses in commercial and institutional buildings. Indoor air, 9(4), 226-252.
15. Alavi, H. S., Churchill, E. F., Wiberg, M., Lalanne, D., Dalsgaard, P., Fatah gen Schieck, A., Rogers, Y. (2019). Introduction to Human-Building Interaction (HBI) Interfacing HCI with Architecture and Urban Design.
16. Alavi, H. S., Verma, H., Papinutto, M., Lalanne, D. (2017, September). Comfort: A coordinate of user experience in interactive built environments. In IFIP conference on human-computer interaction (pp. 247-257). Springer, Cham.
17. Federal Office of Public Health, https://www.admin.ch/gov/fr/accueil/documentation /communiques.msg-id-74177.html. Last accessed 20 Jan 2020
18. Zaria Gorvett, https://www.bbc.com/worklife/article/ 20160617-the-never-ending-battle-over-the-best-office-temperature. Last accessed 20 Jan 2020
19. Hsu, Y. C., Cross, J., Dille, P., Tasota, M., Dias, B., Sargent, R., ... Nourbakhsh, I. (2019, March). Smell Pittsburgh: community-empowered mobile smell reporting system. In Proceedings of the 24th International Conference on Intelligent User Interfaces (pp. 65-79).