Visual Communication Tools for CO_2 Accumulation in Shared Spaces

J. Brazauskas

Fitzwilliam College

We have deployed 36 environmental sensors in a highly occupied lecture theatre to collect data on CO_2 and temperature distributions indoors. The sensor data was collected over the span of several months and was visualised using a set of novel data visualisation tools. It was found that by showing the building occupants the visualisation tools to inspect the data we could potentially affect their sitting positions in the lecture theatre.

Keywords: BIM-IoT fusion, smart buildings, Human-Building Interaction

Figure 1: Lecture Theatre 1 in the William Gates Building was selected as the main site for the sensor deployment project.

1 Introduction

In the domain of smart buildings, environmental sensor deployments have played a key role in making our inhabitable spaces smarter. Ranging from dozens to hundreds of sensors deployed in single buildings, the mass adoption of sensors in our living environments provides major benefits making the built environment safer, sustainable, and more comfortable for its inhabitants. The two key technologies enabling this are Building Information Modeling (BIM) systems and the Internet of Things (IoT) devices. For this reason, such smart building deployments are often defined as BIM-IoT fusion [8]. BIM provides detailed building information, including technical data and how the building changes during its lifecycle, which in return simplifies facility management. CAD building plans are generally considered to be a subset of the entire BIM package. Meanwhile, the IoT devices can range anywhere from discrete window-mounted reed switches that detect when windows are opened to full environmental monitoring systems with cameras [9]. Therefore, BIM-IoT fusion research raises a plethora of engineering challenges to be addressed as the scope of various projects differ in their used technology stacks and desired complexity [23]. Related literature shows a substantial focus on sustainability by decreasing the building's electricity consumption [11], other researchers aim to develop digital twin models with IoT devices being deeply embedded within virtual models [10], while some investigate the human-related aspects of digitising our physical environments [26].

Nevertheless, currently the majority of the smart environment research is focused on deploying sensor networks in buildings and creating the necessary infrastructure to support it [23]. Therefore, most of the current smart environments research solely focuses on sensor deployment and making physical spaces collect data. Yet, large scale sensor deployments are only one part of the puzzle as the data that these sensors transmit has to be visualised in human-usable format that is informative, intuitive, and easy to use. The traditional engineering-focused approach to smart buildings often disregards the more human-oriented research aspects of trying to create spaces where smart technology is able to impact humans in meaningful ways and leverage the utility of our inhabited environments beyond traditional smart building goals e.g. reduced electricity consumption [6]. Thus, this project aims to expand on the existing smart building deployment in the Williams Gates Building (WGB) by researching how we can communicate the data collected by the sensor infrastructure effectively and how this can influence human behaviour in smart environments. Rather than focusing on the entire building, the project only explores the impact of a dense sensor deployment in Lecture Theatre 1 (LT1) in WGB (shown in Figure 1). To do this, the existing sensor deployment in WGB was expanded to accommodate our research goals, then novel data communication tools were developed and tested in a questionnaire-style user study. Hence the remainder of this paper is structured based on these three general themes – sensor deployment, data visualisation toolkit creation, and user evaluation study design.

1.1 Smart Buildings

In smart buildings-related literature, Tang et al [23] presents an overview of recently published research on BIM-IoT fusion. Key criticisms included practical implementability of the described smart building testbeds, where IoT sensor deployments and their contextualisation were more conceptual rather than fully usable. Furthermore, even in cases where physical deployment did happen, the sensor networks were assessed under lab conditions rather than being subject to more rigorous real-

world conditions. This is directly in contrast with the sensor deployment presented in this paper, where a significant number of sensors were deployed in Lecture Theatre 1 alone to research the arising real-world challenges of tracking a set of environmental metrics in highly occupied spaces. Furthermore, the sensors were contextualised both within the building and LT1, where high granularity sensor readings could be extracted with sub-meter accuracy.

Our smart building infrastructure was built on several key tenets - sensing key environmental metrics, leveraging resilient network infrastructure, and deploying sensors at scale.

First, the target sensing metrics - temperature and CO_2. Choosing these two core metrics allowed us to achieve wide-scale environmental monitoring of the built environment. There was a particular interest in exhaled CO_2 as a proxy for possible coronavirus spread prevention, while temperature was essential to examine the effect of the human body heat radiation on the ambient lecture theatre temperature.

Second, resilient network infrastructure - all deployed sensors were connected via LoRaWAN network and their data was sent to the Adaptive City Platform (ACP) - our proprietary software that enables the BIM-IoT fusion research in the WGB. We deployed a set of gateways and ensured that a low power consumption network would enable our sensors to stay alive for several years.

Thirdly, the sensors were deployed at scale, so that the entire building would have a dense network of sensors (multiple sensors in every room) and even the smallest fluctuations in environmental readings could be detected. This was especially important in LT1 as it had the highest sensor density per square metre when compared with the rest of the WGB.

The combination of these three factors theoretically allowed for a complete in-building monitoring environment where we can both track the building occupancy levels as well as inform its occupants about the ambient temperature and CO_2 states of the building with minimal latency. However, these three core principles were largely set up as an engineering challenge to be solved, where the data communication or the human-computer interaction aspects of the project were largely lacking. In this case the infrastructure and the platform were not used to their full extent as the collected data did not have a full impact on the building's users due to the lack of efficient visual communication tools embedded within the platform.

This leads to the main research question of how such robust engineering infrastructure could be leveraged to better accommodate human needs. Specifically, how can we effectively use the collected data in data visualisations to impact our interactions with buildings?

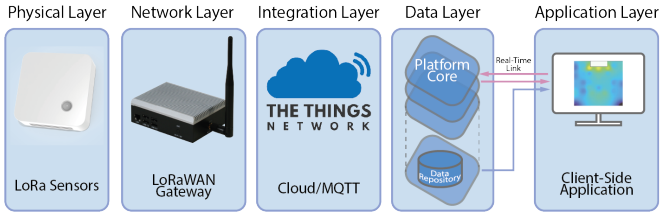


Figure 2: The Adaptive City Platform's architecture is comprised of 5 high-level layers. The core of the platform is the Data Layer ensuring low-latency, asynchronous sensor data streaming capabilities.

1.2 WGB as a smart building testbed

The described project builds on the pre-existing sensor deployment in the William Gates Building (also known as the Computer Laboratory) in Cambridge, UK. The building has become a testbed for large-scale sensor deployment since April 2020, with a proprietary BIM, sensor data-streaming and visualisation platform.

The following subsection describes the smart building deployment in the WGB in more detail, including the description of the installed infrastructure that allows for sensors to be deployed.

1.2.1 Adaptive City Platform

The backbone of the WGB smart building deployment is the Adaptive City Platform - a real-time building monitoring system capable of asynchronously handling spatiotemporal data from Building Information Modeling (BIM) and Internet of Things (IoT) sources with low latency and high throughput [7].

The platform's architecture consists of 5 layers, transitioning from the physical sensors layer to the client-side application layer, as seen in Figure 2. The data layer is the most important part, since it is there that the arriving sensor data is cleaned, filtered and recorded, as well as passed to the real-time data visualisations if necessary.

The platform's architecture is designed to accommodate large amounts of data arriving in real-time. The ACP stores the arriving data in the file system, as well as passes it to the front-end user interface with sub-second latencies. This allows for instant, real-time data visualisations that represent changes in the building's environmental data. The ACP's frontend is a web-based environment, where all of the information is displayed on a website. The data visualisations are written in JavaScript using the D3.js library, while the back-end of the platform uses a combination of Java, Python, and SQL programming languages.

1.3 LoRaWAN

The sensor infrastructure is built on the LoRaWAN network. LoRa (Long Range) Wide Area Network [2, 1] is a commonly used networking protocol in the domain of the Internet of Things and it is known for its low bandwidth

(typically 0.3-5.5 kb/s), low power consumption (under 20 milliwatts) and long operational range (up to several kilometres in urban environments). The deployed sensors connect to local LoRa gateways that act as routers between the LoRa network and some backhaul network, i.e. WiFi, Ethernet or 4G. The power consumption of LoRaWAN-enabled sensors largely depends on the network's strength in the area, hence the placement of LoRa want gateways is essential. The WGB already had two powerful LoRa gateways placed on the roof of the building, however, more gateways had to be installed inside the building, since LT1 was on the ground floor and the signal often had difficulties propagating through thick walls.

To connect the gateways and sensors to our local LoRa network, we used The Things Network (TTN) - an open source online LoRaWAN support service that enables us to easily manage the connected LoRaWAN devices [4]. TTN manages the sensor data traffic from the connected LoRa gateways and allows for the data to be easily accessible once the gateways and devices are registered online. This means that the end user does not have to worry about network routing or bandwidth throttling and can just access the live data using TTN's enabled online infrastructure. Finally, TTN is connected to the rest of the platform using MQTT messaging protocol [3] that allows for publish-subscribe data streams that are at the core of asynchronous ACP functionality.

1.3.1 Previously deployed sensors

Prior to the start of this project the building had a total of 453 environmental[1] sensors deployed in various rooms and corridors. The sensors were deployed over the three floors in the building. However, there were no sensors installed in the lecture theatres previously. The sensors that were deployed previously were a combination of off-the-shelf sensors available from Elsys - a Swedish sensor manufacturer specialising in environmental in-building monitoring sensors with LoRaWAN networking capabilities [12]. Identical Elsys sensors were used in LT1 sensor deployment as well.

1.3.2 Data visualisations for BIM-IoT fusion

Pre-existing data visualisations were using room-contained heatmaps overlaid over the building floor plans, provided by the BIM data. Heatmaps were constrained to every room so that the data from rooms with higher or lower temperatures (e.g. plant rooms) would not spillover to adjacent offices. This was especially useful, since this allowed us to work with single rooms at a time (e.g lecture theatres), without compromising the entire floor-wide heatmap for the building.

Heatmaps were comprised of a 2D grid of cells that were coloured based on the surrounding sensor values according to how far these sensors were from a selected cell on the heatmap. With a high resolution these cells are not visible and appear much more like pixels rather than individual cells. Figure 3 illustrates how heatmaps look with lower resolution and larger cell sizes.

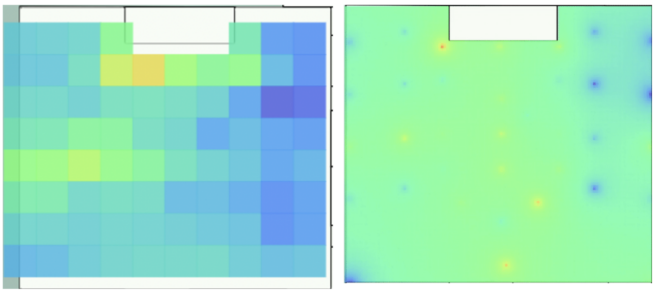


Figure 3: Two identical heatmaps are shown above with high and low resolutions.

1.3.3 Heatmaps and heterogeneous data

In larger spaces, sensors tend to have noticeable reading variations based on their placement locations. Sensors placed on window sills would be expected to have slightly higher temperature readings due to direct sunlight, compared to those tucked away in shadowy corners of the room. Similar issues apply for other environmental metrics, such as CO_2. For example, in LT1 CO_2 readings would be consistently higher next to where the lecturer was speaking, when compared to somewhere on the side of the theatre where no students usually sit. Without the heatmaps we would be left with a range of slightly different scatter plots unique to every sensor and no capability to see what the global state of readings look like at some specified time t. Heatmaps allow us to inspect such reading variations while looking at their spatial distribution at a room-level which in return provides a way to inspect the data with much finer granularity.

The previously designed visualisation system was real-time-focused only, meaning that all visualised data was spatiotemporal. Users looking at the visualisation would see a splash (or a blip) on screen indicating that one of the sensors had sent a message, however, there was no way for that event to be rewinded or backtracked. This raised substantial issues when it came to analysing the collected data. While looking at the heatmap one could tell that a lecture was taking place in real-time, however, at that point it would be impossible to compare it with the state of the heatmap during ambient hours. The lack of rewind functionality essentially made it unusable for any comparative analysis, which would allow to infer the effect people had on the building's environmental metrics over time.

2 Methods

The basis for this project was the installation of 36 new environmental sensors in LT1 in the WGB. This section will describe in detail how the deployment of these sen-

[1] The remainder of this paper refers to environmental sensors as those capable of sensing temperature, CO_2, and humidity

Figure 4: Numerous measurement were taken in LT1 to ensure that the CAD file matched the physical parameters of the site as closely as possible.

sors was executed, as well as what strategies were used to address the main research question of creating data visualisation tools that impact the users' interactions with the building.

2.1 Sensor Deployment

Before the sensors could be deployed, sufficient ground work had to be carried out to ensure that the space was suitable for the project. There were two large lecture theatres in the William Gates Building - LT1, comprising 264 fixed seats, and LT2, comprising 138 fixed seats. Both lecture theatres had seating arranged on a slope. Lecture Theatre 1 was selected to be more suitable due to its size and more common use that would allow for more data to be collected.

In a typical smart building deployment we could make use of the existing building's BIM files, including CAD drawings. However, while BIM data for floor-level floor plans were available, room-level CAD plans for the site were missing. Since no CAD or BIM files were available for LT1 at the sufficient data granularity, it had to be modelled from scratch. This included taking rigorous measurements of the site (Figure 4) and redrawing the room in detail in our CAD software of choice (Rhinoceros) to acquire precise chair, desk and staircase positions (shown in Figure 5).

2.1.1 Site Survey

Prior to the full sensor deployment, LT1 had to be surveyed to ensure that there was sufficient network connectivity from the LoRa gateways already installed on the roof of the WGB. The site was surveyed using Adeunis Field Test Devices [5], as shown in Figure 6. According to official documentation, LoRa modulation has a total of 6 spreading factors (SF) ranging from SF7 to SF12. These spreading factors can influence data rate, time-on-air, battery life, and receiver sensitivity [4]. Generally, lower spreading factors (e.g SF7) mean higher data transmission rates and shorter active times for the radio transceiver, that in return make battery life longer. However, this comes at the cost of shorter operating ranges. Since we did not care for long distances due to all of our sensors being deployed in a single building under

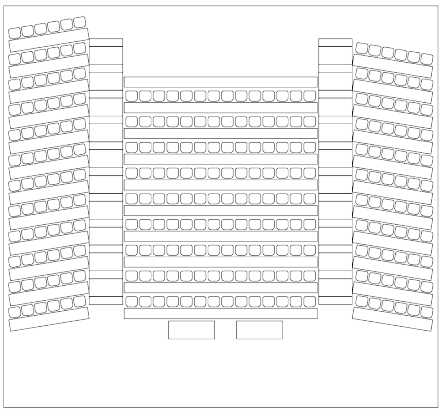


Figure 5: LT1's floorplan used for sensor deployment and data visualisations.

Figure 6: LoRaWAN signal strength measurements were taken in LT1 to ensure the sensors could be deployed successfully.

centralised gateways, lower spreading factors were much more preferable. The Adeunis Field Test Devices are able to show the Received Signal Strength Indicator (RSSI), Signal to Noise Ratio (SNR), as well as the power consumption at each SF, which we were primarily interested in.

After the site was surveyed it was concluded that the existing LoRaWAN gateways set up on the roof of the building did not provide sufficient network coverage to support the 36 new sensors, hence new portable gateways had to be deployed.

2.1.2 Gateway Placement

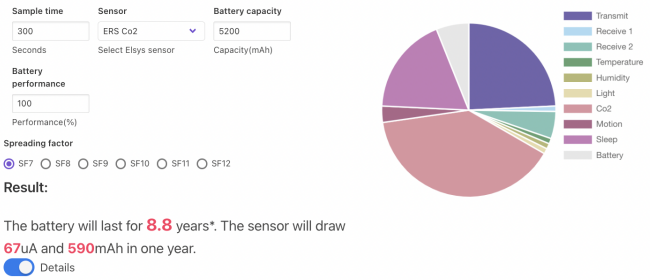
In total, two new Pycom LoPy-based gateways [19] were assembled, programmed and deployed in LT1 where the signal was measured to be the weakest (shown in Figure 7). Both gateways were assembled from scratch, flashed with proprietary software and connected to The Things Network.

2.1.3 Setting up the sensors

Prior to the deployment, our sensors had to be name-tagged (Figure 8 (left)), disassembled to insert batteries (Figure 8 (middle)), flashed with Elsys software, cal-

Figure 7: Left: LoPy LoRa gateways prior to assembling. Right: An assembled gateway deployed in LT1.

Figure 8: Sensor preparation process - name-tagging, battery placement, and calibration.



Name	Awake (mS)	Times (No/h)	Current (uA)	Current per hour (mA)	Current per year (mA)	Battery use (%)
Transmit	70	12	70000	0.0163	143	24
Receive 1	10	12	15000	0.0005	4	1
Receive 2	70	12	15000	0.0035	31	5
Temperature	20	12	15000	0.001	9	1
Humidity	20	12	15000	0.001	9	1
Light	20	12	15000	0.001	9	1
Co2	60	12	130000	0.026	228	39
Motion	1000	3600	2	0.002	18	3
Sleep	1000	3600	12	0.012	105	18
Battery	1000	3600	4	0.004	35	6

Figure 9: The energy consumption of a typical Elsys CO_2 sensor.

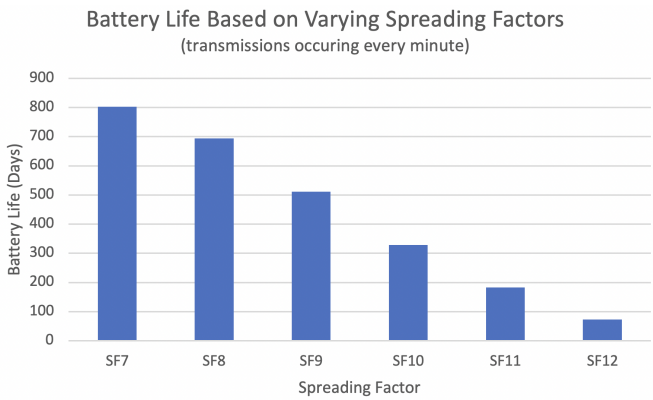


Figure 10: The impact on battery lifetime when using different LoRa Spreading Factors.

ibrated (Figure 8 (right)), connected to the Things Network and finally linked to the Adaptive City Platform.

The sensors were configured to send data every minute. Frequent sensor data transmissions were crucial as the existing data on CO_2 levels in the WGB had suggested that CO_2 tended to rise very rapidly (within minutes) in closed, densely occupied spaces and then plateaued shortly afterwards. This led to important design considerations since we had to maximise the amount of data we were receiving without sacrificing long-term battery life. After running battery lifetime calculations it was deduced that the installed batteries were large enough to support such frequent periodic readings for over a year. Battery-life calculations can be found in Figure 9.

As expected, the transfer of data consumes a substantial amount of the battery life, however, the integrated CO_2 actually consumes more power than anything else the device, as seen in Figure 9. Hence, optimising CO_2 sensors remains an important research topic in the future, especially taking into account that CO_2 sensors are key to human activity monitoring in smart building research. Finally, to illustrate how efficient LoRa network protocol is, we can have a look at varying periodic data transmission - had the periodic data been set to 5 minute intervals, the battery would last 8.8 years at SF7 (shown in Figure 10).

2.1.4 Sensor Calibration

After installing the batteries and turning the sensors on, the devices had to be calibrated by placing them outside for 15 minutes. This ensured that the sensors' baseline readings were identical. However, since not all of the 36 sensors were calibrated at once, there were still some inconsistencies between the baseline readings for the sensors. Hence in the future it is highly recommended to do the calibration procedure for all the sensors simultaneously. Especially this is the case when such sensors are all going to be in the same enclosed space and relatively close to one another.

2.1.5 Sensor Placement Options

There were numerous ways to deploy the 36 sensors in LT1, however, the main aim was to acquire the most even coverage of the site. Research papers suggest that infectious diseases spread in air over 1.5m in all directions [22]. Since using CO_2 as a proxy to measure the spread of the

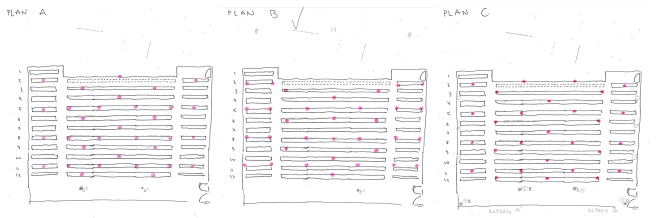


Figure 11: Several sensor placement options were analysed, until it was settled on the option B (middle) since it provided the most uniform sensor coverage in LT1.

Figure 12: Mounting sensor under the desks in LT1.

coronavirus was one of the early goals of the project, it was aimed to ensure that no two sensors were no further away from one another than the specified distance. Overall, several options were assessed before deciding on the final deployment locations for every sensor, as seen in Figure 11.

2.1.6 Sensor Deployment

The selected sensors' locations included a combination of under-desk mounted sensors (that had to be attached on the wooden panels behind the chairs (shown in Figure 13)), floor mounted sensors, and some placed on other available spots in LT1. The placement of these sensors turned out to be very important later on, as some of the deployed sensors were sending consistently higher temperature readings than other sensors due to unforeseen heat releasing devices in the close vicinity of the sensor, as explained in the Discussion Section.

Most of the sensors in LT1 were attached under desks. The sensors were mounted on the panels behind the seats, so that the PIR (Passive Infrared[2]) sensors on the Elsys devices would be pointing towards the students' legs, and thus would be additionally triggered whenever the students moved past them. The backs of the sensors were drilled into the panels with shallow screws so that they would remain in place, while still being easy to disassemble and leaving virtually unnoticeable holes after being taken off the panels. The other sensors were placed on the ground next to the walls with double sided tape securing them in place.

Finally, once the sensors were placed, a grid was overlaid over the CAD file with the official WGB coordinate scheme created for the Adaptive City Platform [24]. All

[2]Similar to those found in automatic light switches

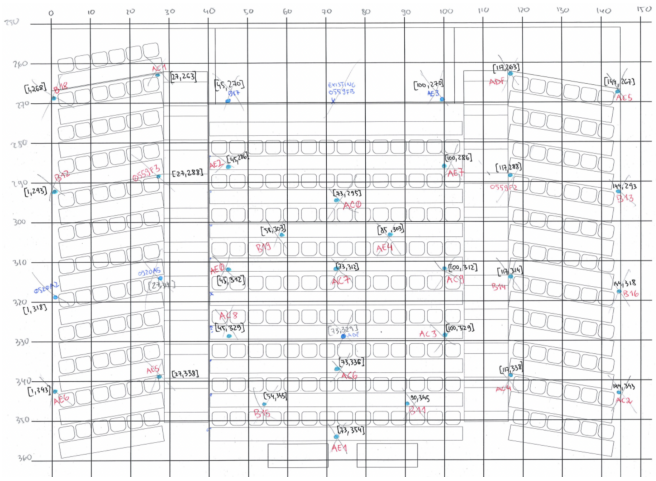


Figure 13: A grid was overlaid over the LT1's floorplan to acquire the sensors' coordinates in relation to the previously deployed sensors in the WGB.

of the sensors' coordinates were recorded on the grid and then transferred to the ACP's metadata database to prepare these sensors being connected to the platform.

2.1.7 Connecting to the ACP

All 36 sensors were first connected to The Things Network and then connected to the ACP servers running from the Computer lab.

2.2 Data Visualisation

The following subsection describes the need for effective data visualisations, as well as some major data visualisation challenges in BIM-IoT fusion research.

Data visualisation tools are essential for human-building interaction [6], as it is the primary interface that the building occupants use to interact with smart building systems. Effective data visualisations thus have the capacity to make such interactions better based on the information that the data visualisation is able to provide to its users.

Related literature [7] highlights three major challenges with smart building platforms and their visual data representation – efficient data contextualisation, capturing data flux and hierarchical data representation.

Contextualising spatiotemporal data: With BIM data providing contextual information and sensors providing real-world readings, one of the key challenges is effective contextualisation of sensor data that is immediately obvious and manages to convey underlying data patterns to the user [13]. While one option would be to simply have multiple time series plots for individual sensors and the entire rooms [14], ideally the data would be best displayed on the floor plan [18]. On the Adaptive City Platform, this is achieved by overlaying the heatmap over the floorplans.

Capturing high granularity data flux: In contrast to

traditional building sensor (e.g. thermostats), a large number of in-building environmental sensors spark the need for data management platforms that are able to cater for high-bandwidth and low-latency data streams [17]. While sensors can be configured to send data either periodically or after being triggered, capturing the high-granularity, high-volume data is a major challenge, especially in real-time building monitoring systems. While this is obvious for time series data, it becomes less obvious for some specific visualisations, e.g. heatmaps where they only convey spatiotemporal, present-moment-based information [13].

Hierarchical data representation layers: with smart building data arriving in relatively high frequency settings, e.g. once a minute, the data has to be able to be represented in numerous hierarchical layers [23]. Starting with a building-level layer, then increasing the granularity to allow for floor or room-level views of the collected data and finally culminating in highest granularity, single-sensor time series data [25]. This leads to a hierarchy of different visualisation where different levels represent different needs. For example, there could be a heatmap of room-level data, where the granularity can be further increased by clicking on a sensor for a time series scatter plot for a single sensor rather than the entire room.

The sensor deployment in the WGB and its extension in LT1 takes into account these three major challenges. However, we are mostly interested in expanding on these core challenges with the fourth one - exploratory past data analysis. The challenge here is to take a step back from the real-time-first data approach and to create a toolkit that allows for the data to be rewinded in time and contextualised on a high level. This in return allows the user to make more detailed adjustments on the kind of data they want to see and select the period for when that data was collected. Such an exploratory tool was missing from the platform and it is something that could allow for users to make more informed decisions about how they wish to interact with the building, i.e. selecting where they wish to sit in the lecture theatre based on the previously collected data.

3 Results

This section will discuss the data collected from the deployed sensors, as well the visualisations' impact on its users.

3.1 ROTAS

The new visualisation was titled Real-time Operated Time-Agnostic System or ROTAS, named after the Sator square [20] signifying its ability to be read backwards and forwards, just as time in the novel visualisation.

The visualisation is comprised of three parts. First, the heatmap showing the data at a specified time t, secondly a slider used to adjust the time, and finally a scatter plot showing how the sensor readings match those at the specified time on the slider. A screenshot of the final visualisation can be seen in Figure 14.

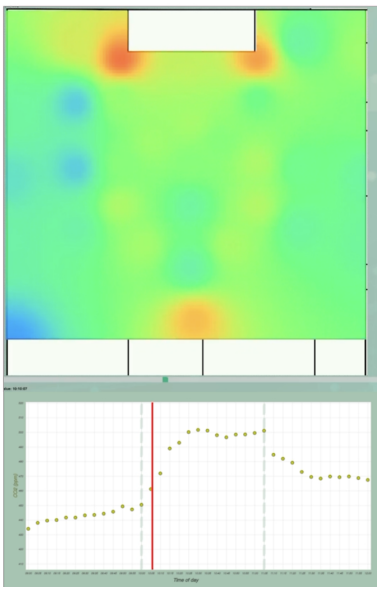


Figure 14: The final visualisation comprised of the heatmap, a slider, and a time-series plot underneath.

3.1.1 Rewinding Time

The main challenge for inspecting past historical data for heatmaps is the asynchronous nature of the platform with sensors transmitting their data packages in non-synchronised intervals. Even with the majority of sensors being configured to send their data in one minute intervals, some data transmissions may arrive in irregular intervals (e.g. after a sensor was triggered by someone walking past it). Hence, overall there is no orderly structure to the ways these packages arrive and are stored on the platform, which makes finding readings at a specific time for all sensors challenging task.

One upside is that all incoming sensor readings are saved in a chronological order, which means that for every sensor we already have a sorted list of messages that is relatively easy to search through using search algorithms like the binary search [21].

In order to find readings for all sensors at some specified time t, the binary search algorithm was used due its speed (runs in logarithmic time) and relative simplicity. Since the received sensor data is already sorted based on its time of arrival on the platform, we just have to issue a request to find the closest reading to a specified time t. All of the arriving data is recorded using the UNIX timestamp format, which makes it substantially easier to look for a specific data packet at a target time. Since the exact requested timestamp may not be available, we use a time window that may be good enough for our use, which in this case was set to 3 minutes. This was done

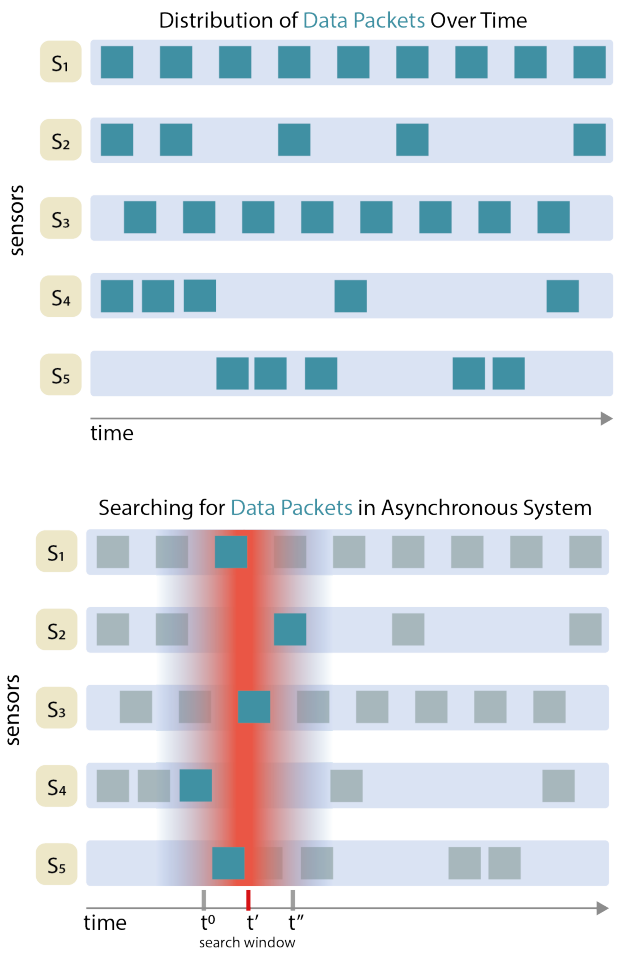


Figure 15: The binary search algorithm was used to find the required data packets from the ACP servers.

to account for any missing data that sensors may have due to network outages, packet losses or other various timeouts.

3.2 Usability Changes

The original visualisation had to be improved not only to include the time rewind functionality but also to implement important usability changes. In Human-Computer Interaction, *usability* can be defined as "a measure of how well a specific user in a specific context can use a product/design to achieve a defined goal effectively, efficiently and satisfactorily" [15]. Therefore, prior to user testing, these *contexts* had to be modified, which in our case was the colour schemes that were used for the original heatmap, as well as the way the data was being visualised.

3.2.1 Colour Schemes

The initial version of the heatmap visualisation on the Adaptive City Platform used a limited colour scheme (Plasma) that did not allow for high granularity visual

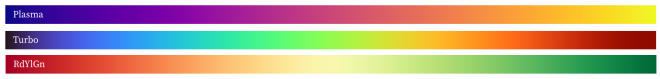


Figure 16: The three heatmap colour schemes shown (top to bottom) - Plasma, Turbo and Green-Red(RdYlGn).

feedback for the user. Thus, a new colour scheme with more transitional colours (Turbo) was selected as the primary colour scheme to reflect finer granularity sensor value changes on the heatmap with greater colour gradient. The used colour schemes can be seen in Figure 16.

3.2.2 Displaying the Rate of Change and Accumulation

The original visualisation was used to display the CO_2 accumulation and temperature readings in LT1, based on every sensor's readings. While this was a valid approach that was used both in the original and new heatmap visualisations, another alternative was proposed with the primary focus on CO_2 due to its dynamic changes when large groups of people are present in LT1. Instead of showing the CO_2 accumulation, a novel heatmap visualisation could show the CO_2's rate of change. This would be done by calculating the derivative values of two consecutive sensor readings[3]. The result would display if the CO_2 levels in the room were rising or falling, using a divergent green-red colour scheme. It was created aiming to analyse how dynamic CO_2 changes were in LT1 during the course of a lecture.

In total we had three distinct visualisations, as seen in Figure 16 - the original (CO_2 accumulation, Plasma), one with an updated colour scheme (CO_2 accumulation, Turbo), and one showing the CO_2 rate of change (Green-Red).

3.3 Heatmap Algorithm Changes

The previously created heatmap algorithm was limited in its visual output due to the way colour values were being calculated for each cell on the heatmap. In short, the previous heatmap algorithm did not allow for a wide range of values to be shown on screen since it was using a linear mapping between the sensors' readings and the cell colour values. This resulted in the entire room often being colored by single colour with differences between different regions in LT1 being very faint and only visible at the exact sensor placement location (seen on the left in Figure 16). Therefore, new ways to better visualise these differences between the readings on the heatmap were sought after.

A novel algorithm was proposed based on Inverse Distance Weighting interpolation [16]. The new algorithm

[3]Percentage rate calculations were also tested as an alternative to calculating the derivative, however, both produces almost identical results

would similarly calculate the heatmap's cell values based on the values of the sensors around it, however, it would amplify this value by raising the distance d from the sensor to a cell by a power of p.

$$z_p = \frac{\sum_{i=1}^{n} \left(\frac{z_i}{{d_i}^p} \right)}{\sum_{i=1}^{n} \left(\frac{1}{{d_i}^p} \right)} \tag{1}$$

Calculating the heatmap cell's value z. d is the distance from the cell z to the sensor i.

The new equation allows for a wider range of values to be visualised as well as clearer reading boundaries to be shown in the visualisation. This essentially acts as a visual trick to better illustrate the regions affected by different sensor readings.

As expected, the original heatmap calculation algorithm matched the novel one when the distance was raised by the power of 1, mimicking the original linear mapping between the sensor readings and the distance. However, as we keep raising the distance from the sensor to the power of p, the heatmap becomes fractioned into a Voronoi-cell-like structures, where each sensor is given its little region on the heatmap (Figure 17).

3.3.1 Room-scale time-series plots

Mirroring the problem of data heterogeneity described in the heatmap subsection, the 36 sensors deployed in the same lecture theatre all produced unique scatter plots due to various local environmental factors like HVAC vents being closer to some of the sensors (Figure 18). Combined with the lack of room-level time series plot pages on the platform[4], this created a challenge when we wanted for the novel visualisation tool to display a heatmap and a time series plot showing the time that the heatmap was rendered at.

However, due to the asynchronous nature of the platform, the collected data for these sensors could not be easily converted to a single time series plot as the timestamps for different readings would not match up for all 36 sensors. Figure 18 exemplifies these problems by showing different scatter plots vary in shape based on where the sensor is placed in LT1.

One approach to solve this problem was to subdivide the day into 5 minute segments and then run the binary search algorithm on each of those 5 minute segments checking what values each of the 36 sensors had at the requested time. These values were then averaged and a new time series plot was created.

The resulting output is a smoother time series scatter plot than for any of the individual sensors, since all the noise had disappeared in the averaging process (Figure 19).

[4]Previously scatter plots were only available for every sensor with no global view available for average readings from several sensors at a room level

3.4 Recording students' sitting locations in LT1

With the sensor infrastructure and the platform set up, the live sensor data started arriving to the servers. However, the heatmap patterns were seemingly different after every lecture in LT1 which meant that the students' sitting locations had an effect on the collected data.

Hence, in order to compare how the distribution of temperature and CO_2 related to the students' sitting locations it was decided to take photos from the back of the lecture theatre during lectures and then manually place dots over the LT1's CAD floor plan. This was done over the span of the entire term, in total over a dozen lectures were photographed and students sitting location were marked.

3.4.1 Ethics

An ethics approval form was submitted and posters were put up on the doors of all LT1's entrances and exits. The students were only photographed from the back and it was made sure that the contents of their laptop screens were not visible.

3.5 User Study

In order to assess the effectiveness of the new ROTAS visualisations (and see how they compared to the original), an online user study was conducted. The study was designed to ask the participants a mix of general (would they use a similar visualisation tool) and specific (if they liked the colour schemes) questions regarding the visualisations. The participants had to rate the visualisations by rating the question statement using a 7-point Likert scale, 1 meaning they strongly disagreed with the statement, while 7 meaning they strongly agreed. The questions can be found below.

- I think that I would like to use this visualisation frequently
- I think that the visualisation was easy to understand
- The colours clearly show the highest/lowest CO_2 concentration in the lecture theatre
- This visualisation helps my understanding of how CO_2 moves around in the lecture theatre
- This visualisation helps my understanding of how CO_2 moves around in the lecture theatre

3.5.1 Running the User Study

In total, 9 participants were recruited for the study. During the study, the participants were introduced to the sensor deployment in LT1 and then shown a video[5] of CO_2 levels changing over the course of a lecture. Three visualisations were shown side by side with the time-series scatter plots shown under them, a red bar was moving

[5]The video can be found here: https://www.youtube.com/watch?v=riESn1mJTlw

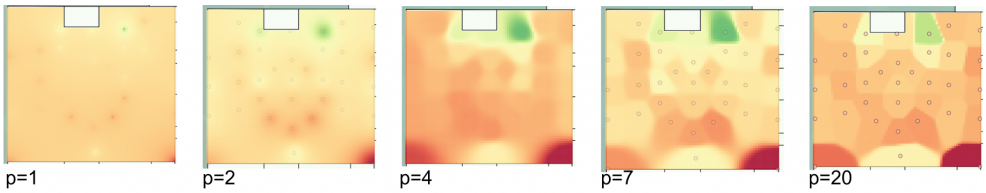


Figure 17: With raising distance *d* to the power *p* the extreme changes in sensor readings become more pronounced, despite time being similar for all heatmaps. In return the spatial distribution of heatmap cells start to resemble a Voronoi diagram. Circles represent the locations of individual sensors.

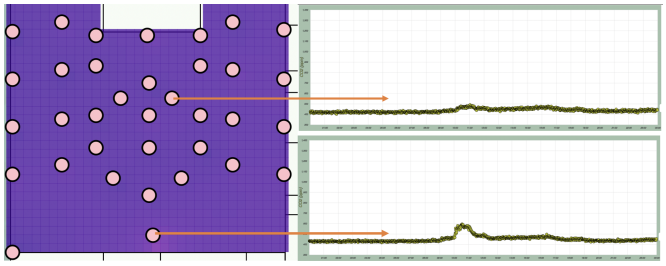


Figure 18: Sensor readings can differ greatly depending on where they were place in the lecture theatre.

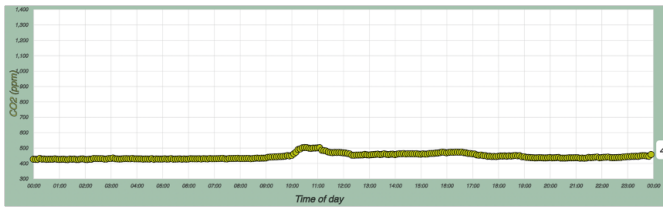


Figure 19: Averaged readings from all of the sensor in LT1. The graph is much smoother since any noise in the data was averaged out.

across the time series plot to indicate how the real-time clock matched the time axis, as seen in Figure 20. After watching the video, the participants were asked to answer the 5 questions regarding the visualisations.

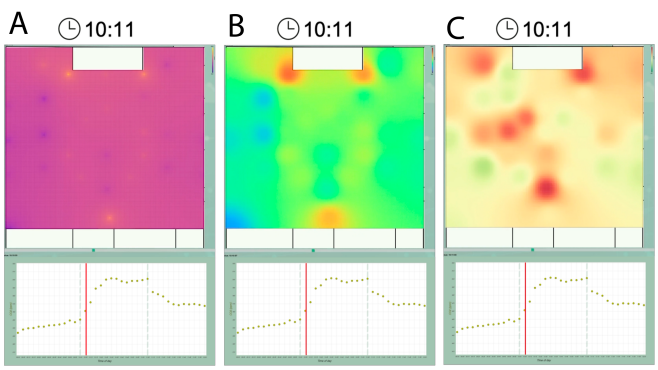


Figure 20: The three visualisations shown during the user study.

3.5.2 Study Results

Overall the results indicate that there is a clear preference for the new visualisations (Figure 21). Furthermore, all participants stated that they would like to use such a data visualisation frequently, and that it would impact their sitting preference in the lecture theatre. This directly ties in with the main research question showing how the collected data can be used to affect the building occupants. Therefore, it gives an indication that data visualisations like this can amplify the existing sensor deployment in the building to create deeper, more meaningful interactions that affect how building users interact with the building.

4 Discussion

The collected results generally indicate that all of the measured metrics in LT1 are highly dependent on the number of people present in the lecture theatre. This means that people attending lectures have a great deal of influence over temperature and CO_2 in the lecture theatre. In this section, we will look at how the measured metrics are affected by the student distribution in the lecture theatre.

4.1 Temperature data analysis

First, the average temperature can fluctuate greatly in the lecture theatre based on the ambient outside temperatures. This means that the HVAC system in the building is not always able to maintain a consistent temperature as on warmer days the ambient temperature in the lecture theatre is higher by several degrees, when compared to cooler days. This becomes especially apparent when analysing the collected heatmap data on a fixed value scale, as seen in Figure 24. This means that the starting state of the heatmaps were not identical on all days, which is important when doing a visual analysis of the data.

Overall, the temperature distribution showed that the lecture theatre heated up considerably over the course of lectures, however, there was also noticeable noise in the data. It was also found that temperature increased consistently throughout the lecture duration and decreased slowly after the lecture concluded. Furthermore, after examining several day's worth of collected data from lectures, it was noted that some of the sensors were return-

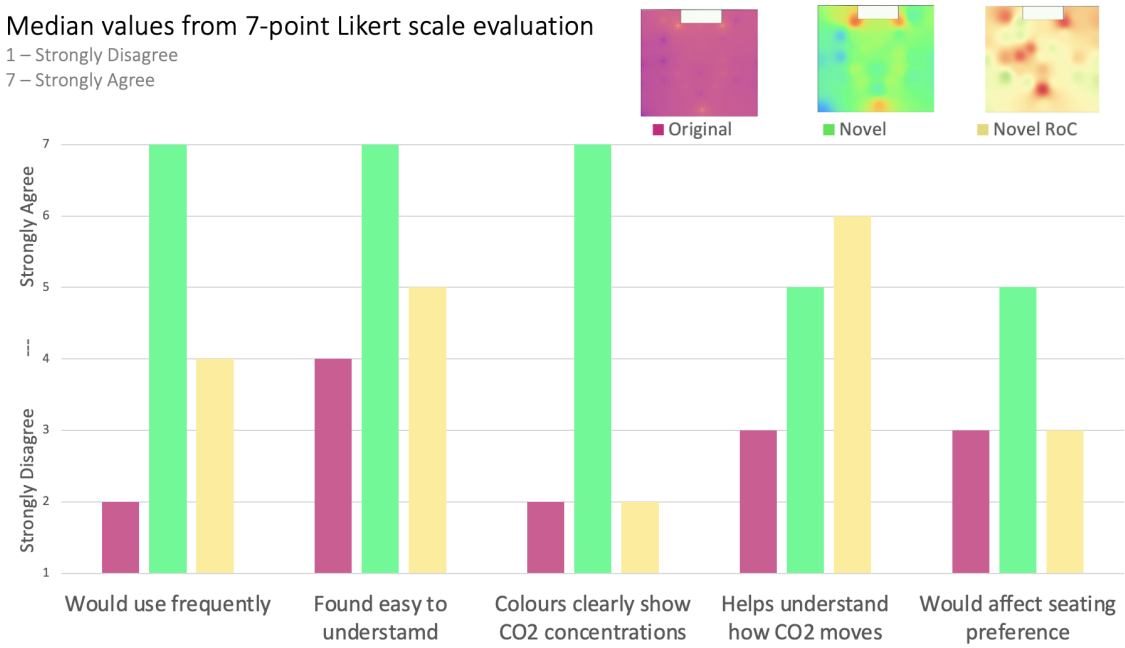


Figure 21: The results of the study.

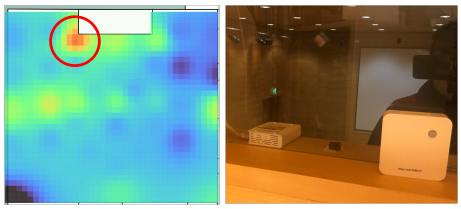


Figure 22: An example of unforseen factors influencing sensor readings. A router behind the glass warmed up the temperature sensor in return creating a bias in readings (circled in red).

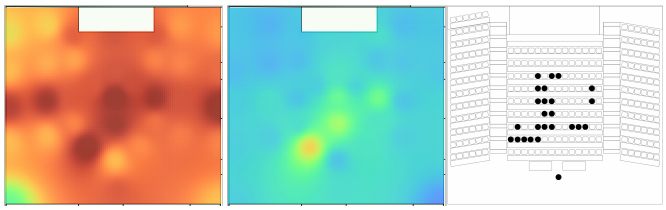


Figure 23: Temperature readings after a lecture. (Left - original data. Middle - midnight data subtracted to account for uneven baselines. Right - student distribution in LT1.

ing consistently higher readings than others, as seen in Figure 22.

One explanation for this was that sensor calibration was not done at the same time for all sensors, hence some could have a different baseline than others. Alternatively, some sensors could have been from a different manufacturing batch, and these have different factory settings for baseline temperature. Finally, it was also found that one of the sensors was accidentally placed next to a heat emitting router that was hidden behind a glass, thus consistently warming up the sensors adjacent to the glass (Figure 22). One way to get around the problem of uneven baselines was to subtract the temperature readings recorded at midnight from the data we were interested in. Results can be seen in Figure 23.

Midnight was selected as there would be no students in the lecture theatre and all sensors were assumed to fall into their natural baselines at that time. The results of this made the heatmaps more readable. For example, when comparing the student distribution in lecture theatre we can see a loose correlation between warmer regions and places where groups of students were sitting. Furthermore, by adjusting the slider that modifies the heatmap algorithm (raising sensor-cell distance to the power of p), these regions can be further enhanced to see which sensors had higher readings compared to the rest of the heatmap- something that was impossible using the original algorithm.

4.2 CO_2 data analysis

After analysing the heatmap data it was found that CO_2 generally tended to increase at the front of the lecture theatre where the lecturer was presenting as well as in the back of the lecture theatre where the sensors were not too far from the ceiling. Furthermore, it was found that CO_2 levels tended to rise very rapidly from the beginning of lecture, reaching its apex at around 15 minutes into the lecture and then staying there during the course of the lecture. While opening the doors was found to temporarily decrease CO_2 levels, as soon as the doors were

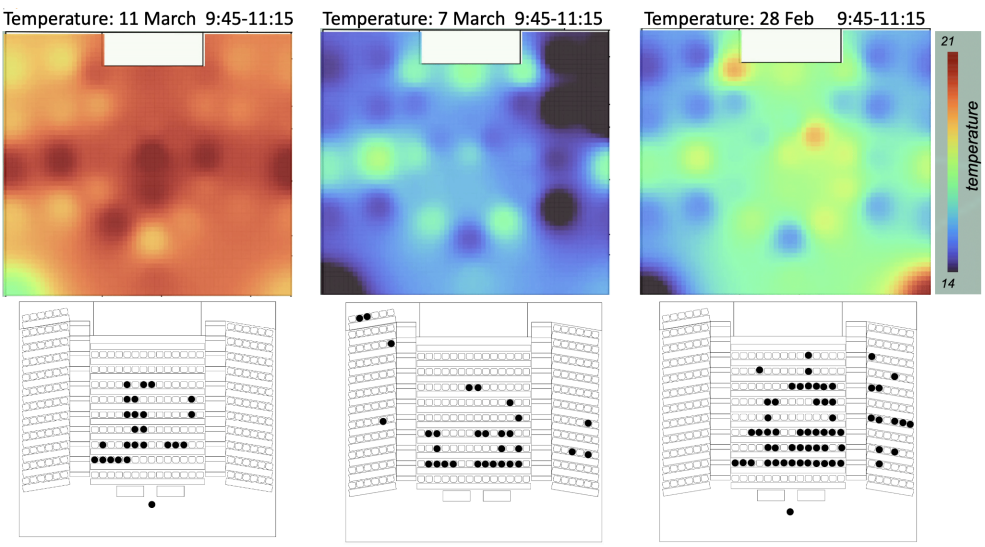


Figure 24: Temperature and student distribution in LT1.

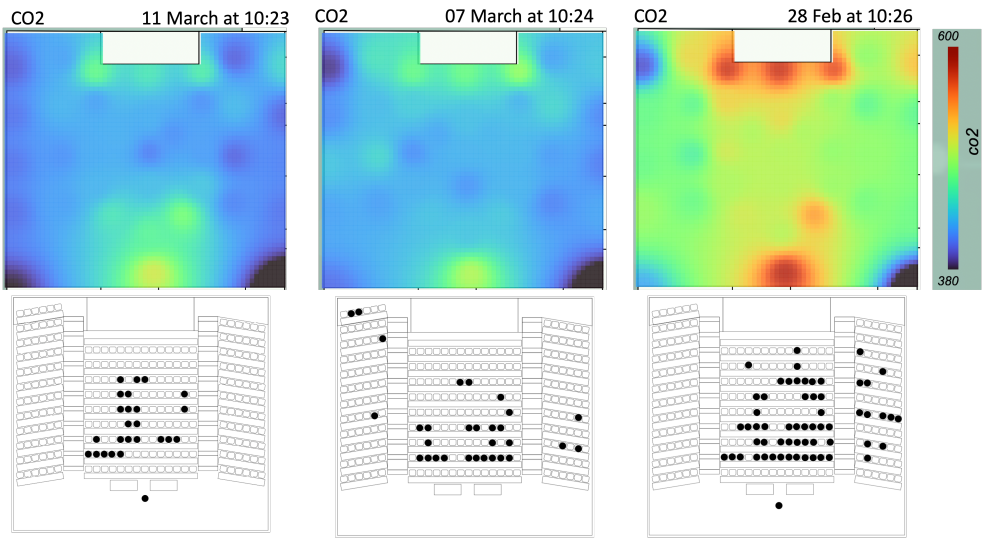


Figure 25: CO_2 and student distribution in LT1.

shut CO_2 levels jumped up again rapidly within minutes.

The distribution of CO_2 at the front of the lecture theatre can be explained by most of the students' preference to sit at the front (shown in Figure 25), as well as the fact that lecturers presenting the lectures did not wear masks that would dissipate the CO_2 in space. The high concentration of CO_2 at the back was surprising since CO_2 is heavier than air and should not go up, especially given the steep slope in LT1 (Figure 1). However, since exhaled air from the human body is a mixture of gases that are warmer than the surrounding air, it generally tends to go up. The sensors located at the back of the theatre were also the highest-placed sensors due to the LT1's shape, hence the observed effect of CO_2 being highest at the back.

Finally, when compared with the student distributions, CO_2 had a smaller correlation with the seating positions than temperature, even when taking into account possible sensor miscalibration and subtracting their readings at midnight to form a better baseline.

Overall the results for CO_2 and temperature analysis show some correlation between the sitting location and the observed changes in environmental phenomena, however, more data has to collectd for definitive conclusions.

4.3 Future Research

The next step for the project could be to deploy the new visualisations to the Adaptive City Platform's production servers so that it could be used by the WGB researchers.

Further research questions could include deploying a similar amount of sensors to a flat lecture theatres to investigate the CO_2 distribution in space where no slopes

can affect it. Additionally, more user studies could be carried out to see how the visualisations were used on a day to day basis and how it affected the students sitting locations over longer periods of time.

5 Conclusion

This paper described a deployment of 36 environmental sensors in a single lecture theatre. Thorough description of the setting up of the sensors was used, as well as how this necessitated improvements to the pre-existing Adaptive City Platform. Novel visualisation tools were described, as well as their impact on the buildings occupants. It was found that the users preferred the new visualisation tools and expressed their interest in using them often. Finally, an analysis of the collected data was done when compared to students' distribution in the lecture theatre, possibly indicating how the environmental readings could be used to infer student sitting locations.

References

[1] Documentation: Lora. https://lora.readthedocs.io/en/latest/, 2022.

[2] Lora alliance. https://lora-alliance.org/, 2022.

[3] Mqtt messaging protocol. https://mqtt.org/, 2022.

[4] The things network. https://www.thethingsnetwork.org/, 2022.

[5] Adeunis. Adeunisftd: Network tester. https://www.adeunis.com/en/produit/ftd-network-tester/, 2022.

[6] Hamed S Alavi, Elizabeth F Churchill, Mikael Wiberg, Denis Lalanne, Peter Dalsgaard, Ava Fatah gen Schieck, and Yvonne Rogers. Introduction to human-building interaction (hbi) interfacing hci with architecture and urban design, 2019.

[7] Justas Brazauskas, Rohit Verma, Vadim Safronov, Matthew Danish, Ian Lewis, and Richard Mortier. Real-time data visualisation on the adaptive city platform. In *Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pages 210–211, 2021.

[8] Justas Brazauskas, Rohit Verma, Vadim Safronov, Matthew Danish, Jorge Merino, Xiang Xie, Ian Lewis, and Richard Mortier. Data management for building information modelling in a real-time adaptive city platform. *arXiv preprint arXiv:2103.04924*, 2021.

[9] Matthew Danish, Justas Brazauskas, Rob Bricheno, Ian Lewis, and Richard Mortier. Deepdish: multi-object tracking with an off-the-shelf raspberry pi. In *Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking*, pages 37–42, 2020.

[10] Bhargav Dave, Andrea Buda, Antti Nurminen, and Kary Främling. A framework for integrating bim and iot through open standards. *Automation in Construction*, 95:35–45, 2018.

[11] Giuseppe Desogus, Emanuela Quaquero, Giulia Rubiu, Gianluca Gatto, and Cristian Perra. Bim and iot sensors integration: A framework for consumption and indoor conditions data monitoring of existing buildings. *Sustainability*, 13(8), 2021.

[12] Elsys. Elsys ers-co2. https://www.elsys.se/en/ers-co2/, 2022.

[13] Matilda Kathryn Ferguson, Sudershan Boovaraghavan, and Yuvraj Agarwal. Vista: Spatial data representation for smart buildings. In *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pages 342–343, 2020.

[14] Gabriel Fierro and David E Culler. Xbos: An extensible building operating system. In *Proceedings of the 2nd acm international conference on embedded systems for energy-efficient built environments*, pages 119–120, 2015.

[15] Interaction Design Foundation. What is usability. https://www.interaction-design.org/literature/topics/usability, 2022.

[16] GISGeography. Inverse distance weighting (idw) interpolation, 2022.

[17] Mads Holten, Christian Aaskov Frausing Rasmussen, Christian Anker Hviid, and Jan Karlshø. Integrating building information modeling and sensor observations using semantic web. 2018.

[18] Aslak Johansen and Mikkel Baun Kjærgaard. Floorviz-a bricked dashboard: demo abstract. In *Proceedings of the 5th Conference on Systems for Built Environments*, pages 186–187, 2018.

[19] Pycom LoPy. Pycom LoPy4 development board, 2022.

[20] Walter Moeller. The mithraic origin and meanings of the rotas-sator square. In *The Mithraic Origin and Meanings of the Rotas-Sator Square*, pages 1–38. Brill, 1963.

[21] Robert Nowak. Generalized binary search. In *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 568–574. IEEE, 2008.

[22] Zhe Peng and Jose L Jimenez. Exhaled co2 as a covid-19 infection risk proxy for different indoor environments and activities. *Environmental Science & Technology Letters*, 8(5):392–397, 2021.

[23] Shu Tang, Dennis R Shelden, Charles M Eastman, Pardis Pishdad-Bozorgi, and Xinghua Gao. A review of building information modeling (bim) and the internet of things (iot) devices integration: Present status and future trends. *Automation in Construction*, 101:127–139, 2019.

[24] Rohit Verma, Justas Brazauskas, Vadim Safronov, Matthew Danish, Jorge Merino, Xiang Xie, Ian Lewis, and Richard Mortier. Sensert: A streaming architecture for smart building sensors. *arXiv preprint arXiv:2103.09169*, 2021.

[25] Thomas Weng, Anthony Nwokafor, and Yuvraj Agarwal. Buildingdepot 2.0: An integrated management system for building analysis and control. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8, 2013.

[26] Sailin Zhong, Denis Lalanne, and Hamed Alavi. The complexity of indoor air quality forecasting and the simplicity of interacting with it–a case study of 1007 office meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2021.