CINT City Net-zero Tool

Silvio Carta
Centre for Climate Change Research
University of Hertfordshire
Hatfield, UK
s.carta@herts.ac.uk

Luigi Pintacuda
ARCH+ Research Group
University of Hertfordshire
Hatfield, UK
l.pintacuda@herts.ac.uk

Tommaso Turchi
Department of Computer Science
Universita' di Pisa
Pisa, Italy
tommaso.turchi@unipi.it

Foteini Papadopoulou
ARCH+ Research Group
University of Hertfordshire
Hatfield, UK
f.papadopoulou@herts.ac.uk

Marc McGurk
Project (EU) Ltd.
St Albans, UK
mmcgurk@project.eu.com

Alan Clark
Project (EU) Ltd.
St Albans, UK
aclark@project.eu.com

Candice Luper
St Albans City and District Council
St Albans, UK
candice.luper@stalbans.gov.uk

***Abstract*— We present CINT (City Net-zero Tool): an analytical and predictive model designed to quantitatively assess carbon data in urban areas. We developed a workflow to collect existing data from city councils on carbon footprint, consumption and production, and tested the interoperability between urban public data and GIS data. We implemented the model using Kernel Density Estimation (KDE) to infer the carbon emissions related to individual buildings based on a station-based dataset. We present initial testing on the integration of data from OpenStreetMap with a vector model and initial testing on modelling data into graph networks to generate enquiries and inference on carbon data and urban scenarios. This method shows how we can integrate more datasets into our base model (graph-based geo-referenced map) to infer unknown information (for example the estimated NO_2 emission per each building). (*Abstract*)**

***Keywords—SDG 11, SDG 12, Carbon Data, GIS Data, Predictive Modelling, Building Carbon Footprint** (key words)*

I. INTRODUCTION

Understanding the impact of urban communities on net zero targets and sustainable future agendas can be a daunting task. The behavior of residents is deeply related to the physical characteristics of our cities. But what if we could calculate and visualize in real-time how sustainable our cities and neighborhoods are today, how they have been in the past and –more importantly—how they could be in the future? CINT is an analytical and predictive tool designed to quantitatively assess carbon data in urban areas. Our tool can process existing urban situations, assessing the performance of buildings, streets and transport, previous urban configurations (establishing whether a particular area has improved/worsened in the past years), as well as future developments (to evaluate the impact of proposed housing, buildings, schools, pedestrianization of streets etc.). We present here the first stage of development of our CINT model that focuses on the assessment mode, discussing initial results of our case study. In this, the model returns an accurate analysis of the air quality of a given urban area. In future stages of development, we will be implementing an optimization mode, where neural networks are used to find the best carbon trade-offs to improve the existing situation, helping councils to achieve net zero targets. This study offers a novel contribution to the existing body of knowledge on environmental quality monitoring (air quality and carbon footprint) with urban carbon data. We present promising results that can help designers, planners and policymakers to monitor, predict and better communicate the urban dynamics and complexities that characterise cities to their residents.

II. OBJECTIVES

In this study we explored ways in which the urban data collected from St Albans can be modelled and visualized to yield relevant results. Some of the data used in this part are proxies, used to explore the potentialities of our model and approach, rather than providing accurate carbon data analytics of St Albans. In particular, we focused on: i) initial testing on how data can be manipulated and visualized; and ii) explore the inter-operability between initial data (from Council) and GIS data (using OpenStreetMap data). We decided to model the data using a graph structure, where we can combine data that we have (nodes) with incomplete or missing data (other nodes) that can be updated at a later stage or inferred or assigned a neutral/temporary value. By modelling St Albans as a graph, we can map urban typologies (parks, buildings, streets etc.) to nodes in the graph, assigning specific weights, analyzing configurations of equilibrium and evaluating how data can be passed through the graph.

III. METHODOLOGY

To build the prototype of our carbon measuring tool CINT, we firstly survey all available data for the area of study (St Albans, UK). The survey included data publicly available, data only accessible to councils and data unique to St Albans. We evaluated each dataset collected and selected the most viable for our study (Energy use of Council Buildings and street pollution data). We then create a model using a graph structure and included a number of features to evaluate missing datapoints through inference. We generated a Colab notebook (see Supplementary Materials) which includes the following steps:

1. Installing relevant dependencies (including geopandas pyrosm osmnx geopy ipywidgets plotly python-igraph swifter etc.);
2. Import data (tabular data from St Albans sets);

3. Data cleaning and format conversion. We converted easting/northing to latitude/longitude;
4. Plotting the first set. We plotted the geo-referenced values from the tabular data (dataset St Albans street pollution). Points are from NO2 stations and displayed using a heat-map based on their values per year (from 2016 to 2022) compared to the yearly average. The heatmap has a bottom slider that shows the change of value per year per each point, as shown in Figure 1.
5. Load graph data. In this stage we load St Albans data as a graph, where each building is a node and each node is connected through an edge, as shown in Figure 2.
6. Implementing Kernel Density Estimation (KDE). Once buildings are represented as a graph data structure, we can calculate and infer new information, as well as add additional data (write in the graph). To achieve this, we used the Kernel Density Estimation (KDE) method [1] [2] to calculate the NO2 emission of each building based on an approximation from the initial data points (from the stations). In our implementation, the KDE algorithm takes the id of the station (graph node), position (lat, lon) and a parameter called "bandwidth", that describes how "smooth" the approximate resulting curve is. The weighting of distances between points (say point x) in the graph can be expressed using the formulation of Conlen [3] as the function $\hat{f}(x)$:

$$\hat{f}(x) = \sum_{observations} K\left(\frac{x-observations}{bandwidth}\right) \quad (1)$$

where K is the Kernel function that affects directly the result of the estimation of the weighting on the observations taken from a population x. Our model finds the closest node for each building (centroid) for the entire urban area, and applies a label with the desired information, as displayed in Figure 3 and 4.

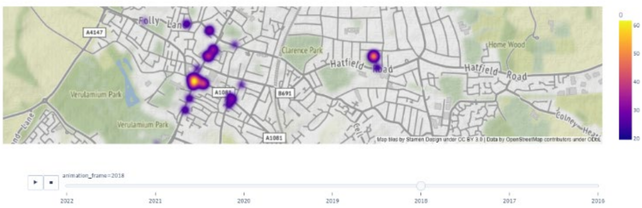
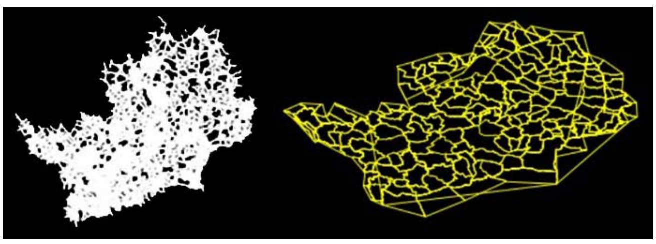


Fig. 1. Plot of tabular data on St Albans map.

Once the graph structure is set up, the model allows for correlations with other datasets. We used the Electricity and Fiscal Consumption data provided by St Albans and visualized them in the same map with the NO_2 values (Figure 5). We also tested the model with traffic data from the Local authority Hertfordshire, [4] and plotted the results in Figure 6,



where we can see side-by-side values of NO_2 for each building in the city and the intensity of traffic of each road.

Fig. 2. Representation of the graph structure underling the city map. Data on all existing roads and inter-sections are collected directly from OpenStreetMap (OSM) and plotted into a graph. Pyrosm uses a NetworkX API and renders a directed graph class that can store multiedges (MultiDiGraph).

Fig. 3. Example of a single data point (e.g. a building) where NO_2 values are estimated using KDE.

Fig. 4. Label for each data point in the graph. Each point has a label that shows the position of each point (lon/lat).

Fig. 5. Plot of the KDE NO_2 data (see key on the right) with Electricity and Fiscal Consumption data per year (key on the left).

Fig. 6. Example of side-by-side plotting of traffic and KDE estimated NO_2 data. Buildings are colored based on NO_2 KDE (red = high values, blue = low values) and roads based on traffic intensity (blue = low traffic, purple = high traffic)

IV. Results and Discussion

The model presented offers a structured approach to map different urban data on OpenStreetMap, identifying key points (e.g. buildings and roads) in a graph-based structure and estimating new values for those parts of the city for which data is not available. In our example, we used energy consumption data of publicly owned buildings in St Albans available from the Council to estimate the energy consumption of any other building (privately owned) in the city.

In our initial testing of the model, we used NO_2 data (nitrogen dioxide), as this was part of a detailed dataset available. NO_2 is part of the NOx family and can be used as an indicator of air quality in urban areas, as it provides a measure of pollutants resulting from fossil fuel combustion (e.g. in cars, along with CO_2 and CH_4). The NOx family includes nitrous oxide (N_2O), which is one of the main greenhouse gases (GHG) components. Our model is useful to estimate urban carbon data of precise points in cities based on multiple carbon datasets, where other data are not available. This method does not consider building carbon data per se, as such values can be more accurately and robustly calculated using established methods for embodied carbon (see for example, ASHRAE [5]).

A. Limitations

The model presented is in a first prototype stage and, although it yielded promising results, it has been limited so far to the available carbon urban data gathered in this study. We used a lightweight and effective architecture, but more robust implementation may require models that allow for a more reliable scale-up, like neo4J or NetworkX. Secondly, the results obtained are related to the data we were able to combine. We note that there are parts of the city with no relevant data, so the KDE and other calculations are coarse in some part. More accurate datapoints would yield more detailed results.

B. Next steps

In future developments of the CINT model, we will include more extensive datasets that include operational and embodied carbon values of buildings to compare the results from our model and evaluate its degree of accuracy. Secondly, we plan to externally validate the results of CINT using sensed data collected directly in situ. Options include the use of Remote Sensing of Real Driving Emission (RDE) using the Plume Chasing method [6] , Portable Emissions Measurement Systems (PEMS) [7] [8], and simple handheld devices, like portable gas analyzers. Thirdly, we will test and implement a Graph Neural Network (GNN), for example successfully used by Peng et al. [9] to optimize carbon trade-offs (production, absorption, off-setting etc.). In this iteration of CINT, we will include tailor the model to yield results that are directly more impactful for residents. For instance, we will include people's behaviors and daily routines as labels to pass in the graph (encoding social behaviors as we did in Carta et al. [10]), predict scenarios based on certain variables that can be provided by both the Council and residents.

V. Conclusion and Recommendations

With this first stage of development of CINT, we explored the carbon and pollution datasets publicly available, as well as those exclusively available to St Albans. We then selected three sets to use in our model. We geo-referenced each data point (air quality station) from the NO_2 dataset. We then modelled St Albans as a graph and integrated the map with the street pollution and Council-owned buildings energy consumption dataset. We used the KDE (Kernel Density Estimation) algorithm to ap-proximate each point in the map (buildings) to the closest node in the graph. Through KDE we estimated the value of NO_2 produced per each building (as a node in the graph). This method shows how we can integrate more datasets into our base model (graph-based geo-referenced map) to infer unknown information (for example the estimated NO_2 emission corresponding to each building in cities). CINT can be a very powerful tool for councils to i) assess current urban configurations against net-zero and sustainable agendas; ii) evaluate future plans and possible alternatives as proposed by developers and private parties; and iii) visually illustrate to residents the impact of changes in the city. Visualizations could also be rendered interactive (including features where users can change parameters and see how urban configurations change) as well as live, including live stream of sensed data that are inputted in the model that visualizes results via browser. This could help to monitor in almost real-time urban carbon emissions.

Our model can support i) designers in assessing the implications of different schemes in terms of carbon footprint and net zero; ii) companies that want to assess and improve the carbon footprint of their buildings; and, more importantly, iii) residents and local communities as they can actively engage with CINT if it offers some-thing with which they can resonate. Key aspects for residents' engagement may include carbon "quick wins" (small changes in daily routines that can have a com-pound effect in a larger net zero agenda), details about $C0_2$ equivalent and $C0_2$ per person, as well as visualizing future or alternative scenarios, including regression-based data trends (of pollution, urban congestion, demolished buildings, pedestrianized streets, new developments etc.).

Supplementary Materials

The datasets collected and used for this study as well as the Colab notebook with the implemented model in Python can be found at this link (anonymized): https://colab.research.google.com/drive/1TJk2jxpGsBTvLCw3jLJTW-MCGI5XaTur?usp=sharing

Acknowledgment

Phase I of CINT has been generously funded by Project EU and the University of Hertfordshire. The authors thank St Albans City and District Council for their help, feedback, and support and for supplying the data for this study. SC thanks Prof Jankovic for his feedback along the project.

References

[1] Węglarczyk, S.: Kernel density estimation and its application. *ITM Web of Conferences* 23, (00037) (2018). DOI: https://doi.org/10.1051/itmconf/20182300037

[2] Kamalov, F., 2020. Kernel density estimation based sampling for imbalanced class distribution. *Information Sciences*, 512, pp.1192-1201.

[3] Conlen, M.: Kernel Density Estimation (2018), https://mathisonian.github.io/kde/, last accessed 2022/10/20.

[4] Department of Transport: Local authority Hertfordshire - Road Traffic Statistics (2021), https://roadtraffic.dft.gov.uk/local-authorities/78, last accessed 2022/10/20.

[5] ASHRAE Task Force for Building Decarbonization: Embodied Carbon Codes and Policies Summary (2021), https://www.ashrae.org/file%20library/about/embodied-carbon-codes-and-policies-summary---final.pdf, last accessed 2022/10/20.

[6] Yang, Z., Tate, J.E., Rushton, C.E., Morganti, E., Shepherd, S.P.: Detecting candidate high NOx emitting light commercial vehicles using vehicle emission remote sensing. Science of The Total Environment 823, 153699 (2022). DOI: https://doi.org/10.1016/j.scitotenv.2022.153699

[7] O'Driscoll, R., ApSimon, H.M., Oxley, T., Molden, N., Stettler, M.E.J., Thiyagarajah, A.: A Portable Emissions Measurement System (PEMS) study of NOx and primary NO2 emissions from Euro 6 diesel passenger cars and comparison with COPERT emission factors. Atmospheric Environment 145, 81-91 (2016). DOI: 10.1016/j.atmosenv.2016.09.021

[8] Luján, J.M., Bermúdez, V., Dolz, V., Monsalve-Serrano, J.: An assessment of the real-world driving gaseous emissions from a euro 6 light-duty diesel vehicle using a portable emissions measurement system (PEMS). Atmospheric Environment 174, 112-121 (2018). DOI: 10.1016/j.atmosenv.2017.11.056

[9] Peng, H., Wang, H., Du, B., Bhuiyan, M.Z.A., Ma, H., Liu, J., Wang, L., Yang, Z., Du, L., Wang, S., Philip, S.Y.: Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. Information Sciences 521, 277-290 (2020). DOI: https://doi.org/10.1016/j.ins.2020.01.043

[10] Carta, S., Turchi, T., Spencer, N. & Vidal Calvet, M. 2022, Encoding social values of local communities in algorithmic-driven design methods. *Formal methods in architecture - Proceedings of the 6th International Symposium on formal methods in Architecture (6FMA)*. A Coruna 2022. Springer International Publishing AG, part of Springer Nature.