International Cartographic Association
Association Cartographique Internationale

Creating Geospatial Data Visualization and Exploration Tools for Simulated and Modeled Datasets

Stuart McGorman [a, *], Jon Corbett [a], Leandro Biondo [a]

[a] *University of British Columbia Okanagan, Department of Community, Global, and Cultural Studies, Stuart McGorman – stumcg@student.ubc.ca, Dr. Jon Corbett – jon.corbett@ubc.ca, Leandro Biondo – lmbiondo@student.ubc.ca*

* Corresponding author

Keywords: geovisualization, simulated spatial data, transportation models, data exploration, synthetic data

Abstract:

The CliMR (Climate Mobility Research Network) project at the University of British Columbia is a multifaceted research project studying mobility, transportation patterns, and emissions in the Vancouver and Okanagan Regions of the province of British Columbia, Canada. Researchers are using survey data to create an agent-based model that simulates the activities and travel patterns of the entire population of each study area. Other researchers are modelling lifecycle emissions of different vehicle types and emissions in various transportation scenarios (car, public transport, walk, etc.) to quantify the environmental impact of the activities simulated using the agent model.

This abstract discusses the Geoportal portion of the CliMR project which aims to create intentionally designed, dataset specific, geospatial data visualization and exploration tools which bring together and present the output data from the various other research teams. We have created cartographic data visualization and exploration tools which are both accessible and engaging to the public as well beneficial to city officials, researchers, and other project stakeholders with varying degrees of technical proficiency. This presentation will examine some of the unique challenges encountered working to visualize modelled datasets that are produced by other researchers in a multi-team project setting.

This project has been implemented as a React web application served by an Express backend using the Nest JS framework. We are storing data in a PostGIS database and have experimented with using Geoserver for increased database performance (Chandra et al., 2022). Geoserver can handle very large data transfers and includes built in geospatial functionality such as coordinate conversions and raster-vector transformations. These features do not have a significant impact on our data visualizations however they will be useful for future project objectives related to making raw data directly available. Frontend data visualizations have been built using the Vis.gl suite of WebGL based geospatial data visualization tools. Specifically, we are using Deck GL to create complex 3D data visualizations on top of Mapbox basemaps implemented using React Map GL. We are also using VisX to create custom non-spatial data visualizations.

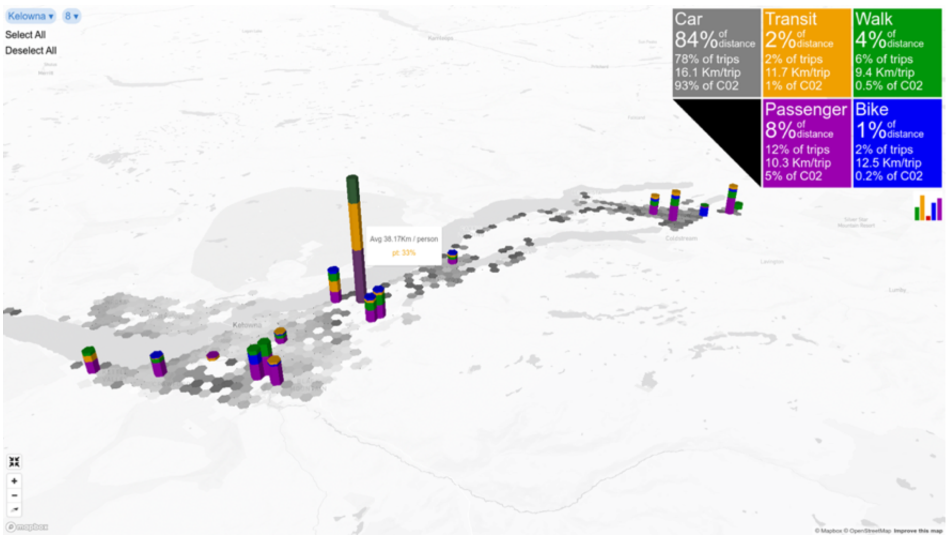

Figure 1: An interactive visualization showing the number of people traveling from various locations in the Okanagan study area, aggregated by mode of travel

We have experimented with creating various interactive visualizations which use extruded 3d shapes on a map to represent spatial data in a way that is engaging, visually appealing, and conveys interesting trends in our data. We have

https://doi.org/10.5194/ica-abs-10-192-2025

then animated these extruded shapes to display temporal changes in data. We have also used the concept of multiple coordinated views (Roberts, 2008) to create additional non-spatial data visualizations such as graphs and line charts which dynamically update and coordinate with the spatial visualizations. These additional views act as both an alternative view on the data, and a way to interact with the map.

Creating useful data visualizations using simulated model data presents many challenges. One of the principal constraints is the size of datasets with which we are working. The agent model that provides us with most of our data simulates the movements of the entire populations of our study areas, which are around 700,000 people in the Okanagan and over 2,000,000 in the Vancouver region. The raw output of spatial data for a single day's simulation is on the order of 10s of Gigabytes. With this volume of data, we have found that it is simply not possible to load and process data spontaneously to generate visualizations, especially not in a tool running client side in a web browser. For this reason, we must do substantial preprocessing to the data. This must be done with the intent to preserve the ability for users to explore the dataset in depth, while still being computationally manageable. We need to strike a balance by making the data usable while also preserving its granularity. For example, the user may be interested in aggregating or filtering trips based on one of many attributes such as personal demographics, vehicle type, end activity, or geographic location. One solution to this has been to store multiple different versions of the data each aggregated on different attributes, as opposed to storing the entire raw data and preforming that aggregation at runtime. Although this technique does limit the potential to aggregate data on arbitrary combinations of multiple attributes.

Model data (or synthetic data) presents a unique visualization challenge because the data is not real and therefore should not be presented as if it were. Model data can accurately represent real world scenarios and predict meaningful trends, but individual data points should not be treated as meaningful on their own. We have had to think carefully about how to process and present model data to ensure that results we visualize are meaningful. For example, one of the first tools we created visualized each trip in our study area as a 3d arc on a map. While this visualization was successful in showing general trends regarding the locations where people travelled to and from, it also tended to direct users towards exploring individual trips which are not real or meaningful pieces of data on their own. We have used our own process of trial and error, as well as research into similar projects such as Todd et al., (2023) to ensure our visualizations accurately represent the meaning of our data.

There are additional challenges related to working within a multi-team project structure, where different research groups are working parallel to each other. This has meant that our data visualization team has had to work with unfinished versions of models which are under active development. Because of this, it is often difficult to test the effectiveness of our data visualizations - as in many cases we do not have any trends in the data to visualize. It can also be difficult to identify bugs or inaccuracies in our visualizations as unrealistic results may be due to our own mistakes or inaccuracies in the source data. One unintended outcome of this parallel work is that we have discovered that data visualization can act as an effective tool to validate a data model. For example, when creating an interactive map which visualized trips filtered by travel mode, we discovered that the model was simulating some agents commuting from Vernon to Kelowna (50km by highway) on foot. While this may not have shown up as an obvious inaccuracy of the model statistically, our visualization intuitively showed that the model was simulating trips which would not occur in the real world.

Another unexpected outcome related to working within a multiteam project is that as cartographers tasked with bringing the data together using geovisualization tools, we have become a central conduit that facilitates the interaction and collaboration between teams. We have found that in many cases the other teams do not have an adequate understanding of each other's projects, and our role has helped to create a more coherent vision of the project as a whole. Often in project meetings our team's requirements for usable data have created conversations between other teams, which in turn has helped them to align their diverse goals with one another.

References

Chandra, S., Raj, U., Sonkar, R., Yadav, U., Srivastava, P., & Gond, R. (2022). A Comparative Performance Analysis on GIS based Web Development using Geoserver versus without using Geoserver. *International Journal of Computer Science and Mobile Applications*, 1–6. https://doi.org/10.47760/ijcsma.2022.v10i04.001

Roberts, J. C. (2008). Coordinated Multiple Views for Exploratory GeoVisualization. In M. Dodge, M. McDerby, & M. Turner (Eds.), *Geographic Visualization* (1st ed., pp. 25–48). Wiley. https://doi.org/10.1002/9780470987643.ch3

Todd, J., Yano, K., & Hanaoka, K. (2023). A dashboard application to explore population distribution derived from GPS location data during the COVID-19 pandemic in Kyoto, Japan. *Abstracts of the ICA*, *6*, 1–2. https://doi.org/10.5194/ica-abs-6-258-2023

Abstracts of the International Cartographic Association, 10, 192, 2025
32nd International Cartographic Conference (ICC 2025), 17–22 August 2025, Vancouver, Canada.
https://doi.org/10.5194/ica-abs-10-192-2025