

Vehicle travel demand and built environment: A computational design tool using machine learning aided method

Manuscript: JCIT-D-21-01817
Review report

This manuscript presents the results of machine-learned predictive models to estimate the total traffic and temporal distribution of traffic from land use features. The author(s), assuming plural for the rest of the report, test the scalability and transferability of their approach in different study areas. A nice addition to the paper is that they then demonstrate the model's application by evaluating the impact of different interventions. Furthermore, they show how the model can be combined with a (meta)heuristic prescriptive model to achieve the most balanced traffic load, making the input ratios the decision variables.

I found the paper, overall, to be fairly well written and not too hard to read. My report only notes some issues to be addressed.

1. Despite my comment on being “fairly easy to read”, the paper requires general language editing.
2. The manuscript includes (too?) many figures. Indeed, it helps convey and visualise results, but I argue that the authors should be more concise and use figures where they truly add value. It becomes hard to read if you have to page back and forth sometimes two/three pages to just navigate between the text and the images.
3. Neural Networks are notoriously hard to interpret; they are black boxes. True, they yield high accuracy, but literature is abundant on when the predictions involve choices that affect *real people*, like medical diagnosis and travel interventions, then decision-makers would rather settle for predictive algorithms that are a) less accurate but b) more transparent.

The current literature review is mute on the advances that are made in *Explainable Artificial Intelligence* (XAI). I would argue that the authors consider the available literature. Consider, as but a few examples, Adler et al. (2018), Barredo Arrieta et al. (2020) and Gunning and Aha (2019).

4. Building on the previous point. At the start of Section 1.2, the authors identify the need of designers to “*determine the relationship between vehicle travel demand and urban functions...*” This should suggest that the need is an *explanatory* model and not simply a predictive model. This manuscript then later reports (p.12) that “[w]e have verified a link between our living environment and vehicle travel demand” but without the ability to add any value in terms of explainability.
5. There is a well-established body of knowledge on studying the patterns between land use and transport. What is the novelty of the manuscript. Simply applying a black-box neural network? Being machine-learned? The field has made much progress on developing behaviourally sensitive models. These neural networks are, in that sense, completely void of any behavioural sensitivity: you can only predict past patterns already observed (in the data).
6. Using taxis is a very specific mode. How generalisable is it to other modes and, more specifically, to other cities/regions outside your study areas?

7. Please explain concepts and acronyms when introducing them. Consider “*urban function ratio*”, “*entropy index*” and “*LMU strategy*”, all in Section 1.2. Check throughout the document. Another example is “*FFBS*” (p.2).
8. p.3, last paragraph: is there any way to consider the difference between value-adding and non-value-adding (empty) trips, or is the traffic order data simply vehicle activity? With the sizeable interest in automated traffic, this seems to become more and more of an issue: (too) many vehicles but very low occupancy rates.
9. General editing (but please check with journal guidelines): usually captions are *above* a table and *below* a figure.
10. Figure 2: I would think the more acceptable abbreviation for longitude would be *lon* and not *lng*. Please check.
11. Figure 4, and others: I, for one, has no idea where Haikou is. Could you please add a reference map so that the heat maps can be better understood and appreciated? Also, consider overlaying/superimposing the heat maps on a monochrome map. Since you are dealing with *sense of place*, this could add much value to readers.
12. Figure 6, for example, adds little value and could be replaced with a single sentence.
13. Table 2: are the results presented from a single instance? I would argue that there is randomness in the prediction workflow. For the sake of the repeatability of the ANN, it may be useful to reports the results over an ensemble of multiple runs. Then report the median over say 100 runs.
14. Generally, I want to thank the authors for using the median instead of the over-used but fragile mean. This may be a trivial issue, but I am still surprised as to how frequently researchers report *mean/average* results without cognisance of the shape of the distribution of the results.
15. p.14: Please expand and elaborate on the statement: “[t]his again illustrates the potential for misleading conclusions to be drawn from vaguely describing information in the urban environment”.
16. The results seem to not demonstrate particularly strong signals. Can it simply be random drift or just random densities?

References

- Adler, P., Falk, C., Friedler, S. A., Nix, T., Rybeck, G., Scheidegger, C., Smith, B., and Venkatasubramanian, S. (2018). Auditing black-box models for indirect influence. *Knowledge and Information Systems*, 54(1):95–122.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. (2020). Explainable Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58(October 2019):82–115.
- Gunning, D. and Aha, D. W. (2019). DARPA’s explainable artificial intelligence program. *AI Magazine*, 40(2):44–58.