An agent-based model for modal shift in public transport

Thibaut Barbet[a], Amine Nacer-Weill[a], Changtao Yang[a], Juste Raimbault[b,*]

[a]Ecole des Ponts ParisTech, Champs-sur-Marne, France
[b]CASA, University College London, London, United Kingdom

Abstract

Modal shift in public transport as a consequence of a disruption on a line has in some cases unforeseen consequences such as an increase in congestion in the rest of the network. How information is provided to users and their behavior plays a central role in such configurations. We introduce here a simple and stylised agent-based model aimed at understanding the impact of behavioural parameters on modal shift. The model is applied on a case study based on a stated preference survey for a segment of Paris suburban train network. We systematically explore the parameter space and show non-trivial patterns of congestion for some values of discrete choice parameters linked to perceived wait time and congestion. We also apply a genetic optimisation algorithm to the model to search for optimal compromises between congestion in different modes.

1. Introduction

Disruptions in public transport networks are not rare events, often worsened by an increased complexity of these networks and their management (Dekker et al., 2018). Network resilience is then tightly linked to patterns of modal shift (Stamos et al., 2015), and a better understanding of these is crucial both from a theoretical and operational viewpoint. In that context, users have to make decisions in a limited time and under partial information (Lyons, 2006). The study of modal shift under disruption is therefore improved when taking into account users behavior and a detailed representation of users cognition (Brisbois, 2010). More particularly, the role of information provided in real time can in some cases become crucial, as rerouting may in fact increase the congestion in other parts of the network and ultimately increase the average travel time (Chatterjee et al., 2002; Chorus et al., 2006).

To study mechanisms linking information given to users with modal shift, and more generally the evolution of multi-modal network flows under disruptions, agent-based modeling has been highlighted as a relevant approach (Leng and Corman, 2020b). For example, Leng and Corman (2020a) show that the issue time of information has a significant impact on the total congestion. Agent-based models are used in similar studies of modal share, such as by Baindur and Viegas (2011) for freight. Ambra et al. (2019) include the reaction to disruptions in a multimodal agent-

---

* Corresponding author.
  E-mail address: j.raimbault@ucl.ac.uk

based model. Raney et al. (2003) describe an application of the MATSim model, which is a data-driven agent-based and activity-based transport model, to a large sample of Switzerland transport network. The MATSim model can be applied at large scales in a reproducible manner, such as in the case of Ile-de-France illustrated by Hörl and Balac (2020).

Large transport agent-based models such as MATSim however require an extensive parametrisation on real data, are difficult to systematically validate given their runtime and large parameter space, and despite their high modularity can be tuned only to some extent regarding a precise description of users behavior in public transport and their interactions with a network disruption. This paper therefore proposes to introduce a simple and stylised agent-based model to understand the role of information and user behavior in modal shift under disruptions in public transport. With a reduced computational complexity but also parameter space to explore, under limited requirements for data, such a model can be used to systematically explore if some qualitative stylised facts are robustly found under different scenario, and possibly extended into a decision-making tool. Our contribution is threefold: (i) we contextualise the construction of the model with a reduced stated preferences survey, applied on a case study in the Parisian public transport network on a specific segment often subject to disruption; (ii) we introduce the simple agent-based model as an open-source tool, which can be extended or modified to test concurrent hypotheses on user behavior; (iii) we proceed to a systematic exploration of the model parameter space, unveiling some non-linear patterns and a counter-productive reaction of users to the disruption in terms of congestion for some scenarios.

The rest of this paper is organised as follows: we first describe an exploratory stated preference survey; we then describe the agent-based model and results of numerical experiments. We finally discuss potential applications and developments.

2. Stated preferences survey

We first proceeded to a small size stated preferences survey (Kroes and Sheldon, 1988), in order to have a qualitative overview of processes needed in the model and to have a case study for model application. Martin et al. (2016) have indeed shown that there exists a high heterogeneity in user reaction to disruptions. We choose to study the line A of the *RER* suburban train in Paris, which has the highest load in the region and has a non-negligible frequency of disruption. We focus on a subpart of the network, in order to sample users which have a higher chance of realising a given origin-destination pattern, and for which several modal alternatives exist. Therefore, the segment *Etoile-La Défense* was chosen, as it features alternative trips with the *Metro Line 1* and several bus lines.

We surveyed a total of $N = 48$ users, among which a subsample of $N = 27$ were regular users for which the full questionnaire was given. Main results which can inform our model context are shown in Fig. 1. We first confirm that disruptions are indeed frequent on the line, as around 50% of users experience a disruption at least twice a month. We then find that under no information, user behavior is heterogeneous and a large part shifts to alternative itineraries. Finally, when users are informed of a 10 minutes disruption, on the contrary more than 75% decide to wait on place. Therefore, the model must include various modal shifts, but also take into account how perceived wait time will influence user choice.

This survey was designed in an exploratory manner and could not be continued on a larger scale for various practical reasons. Our sample size thus remains relatively small and this part of the study must be considered as qualitative. Significantly larger sample sizes are needed to obtain robust inferential results, for example to forecast travel demand (Horváth and Horváth, 2015) or to estimate discrete choice parameters (Bliemer and Rose, 2009). In our case, we consider the simulation model as a tool to go beyond such data limitations and still study stylised behavioural processes.

3. Model description

The agent-based model is based on two types of agents: users and RER trains. The simulated segment of the network is included for the main mode (RER train) and alternative modes (metro, bus, taxi, bike, walking). Main mode is simulated in full granularity, i.e. including train boarding, train capacity and train alighting. Other modes are simulated as queues with a given capacity. One simulation covers peak hours, which we take as a duration of 4 hours,
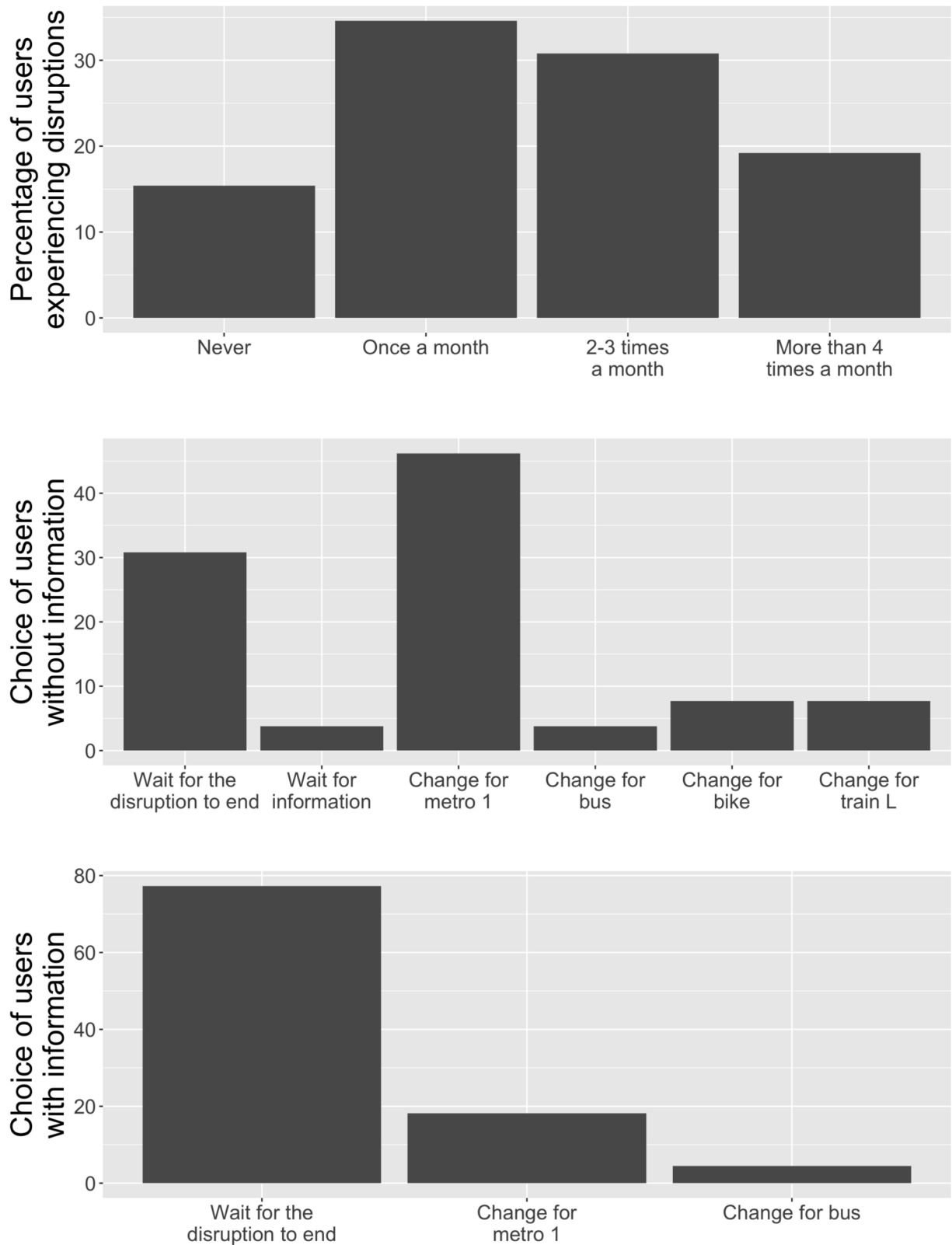
Fig. 1. Results of the stated preference survey. *(Top)* Number of times a month users experience a disruption; *(Middle)* Behavior of users when no information is given; *(Bottom)* Behavior of users when a 10 minutes traffic stop is announced.

with one-minutes time steps. The model is therefore stopped at $t_f = 240$. Arrival of users for each mode are simulated using stationary Poisson random draws.

More precisely, the model simulates sequentially the following processes at each time step.

1. New users enters the network: for each mode, a fixed arrival rate λ_i is used for a random Poisson law draw giving the number of new users entering the segment on this mode. Users entering the main RER mode are set on the platform as waiting, while users on other modes are queued at the start of the segment.
2. Users waiting for a RER train on the platform evaluate a discrete choice utility difference (between waiting and shifting to an other mode) given by

$$\Delta U = \beta_c \cdot c + \beta_\tau \tau \tag{1}$$

   where c is perceived congestion given by the current number of user waiting (normalised by platform capacity) and τ is perceived time proportional to user current travel time (time since departure) and current waiting time between trains. β_c and β_τ are thus behavioural parameters capturing the influence of perceived time and congestion on user behaviour.
3. These users can then switch mode with the corresponding discrete choice probability probability given by

$$p = 1/(1 + \exp \Delta U) \tag{2}$$

   and with fixed nested probabilities for the choice of the other mode once a shift has been decided;
4. Next train possibly enter station (once the time interval I between trains has elapsed) and users board at a given speed and given train capacity. Train capacity C captures congestion upstream this segment, and is given as a maximal number of users which can board a train. In the meantime, users alight at the end of the segment if a train has arrived at the terminus station.
5. Train queue on the segment is simulated, i.e. each train is advanced at its maximal speed, if the next slot is not occupied by a preceding train.
6. Other modes are simulated as queues with a given user capacity: users enter the queue if it contains less users than its capacity; users in the queue are advanced according to the mode speed; arrived users are removed from the queue.

Model results are assessed with indicators giving the average travel time and average congestion over the peak hour and all users. Congestion indicators can be computed for each mode separately.

The model has several parameters which can be set with realistic values (see setup below), while others remain variables of interest which influence on model behavior will be studied in numerical experiments. In particular, varying parameters are the behavioural parameters β_c and β_τ, time interval between train I and train capacity C.

4. Results

4.1. Implementation and model setup

The model is implemented in NetLogo which is a platform and programming language specifically suited for such agent-based simulations (Tisue and Wilensky, 2004). We show a visualisation of model interface in Fig. 2. The opensource code of the model and results is available as a git repository at https://github.com/JusteRaimbault/ReportMasse. We systematically explored model parameter space by using the OpenMOLE model exploration software (Reuillon et al., 2013), which allows embedding any type of model, provides a seamless access to high performance computing infrastructures, and integrates state-of-the-art model validation and exploration methods.

Several model parameters are set following real world values. Respective speed for each mode for example correspond to travel times for the segment in an uncongested setting. Mode capacities also have realistic values (for example for Metro 1, 700 users per train with 5 trains on the segment). We take a boarding speed for the RER of 1000 users per minute (MI09 trains have a total capacity of 2600 users and stop for not more than 1min30sec during peak hours). A train takes 4min for the segment. We assume a maximal waiting time in station of 2 minutes. We fix the
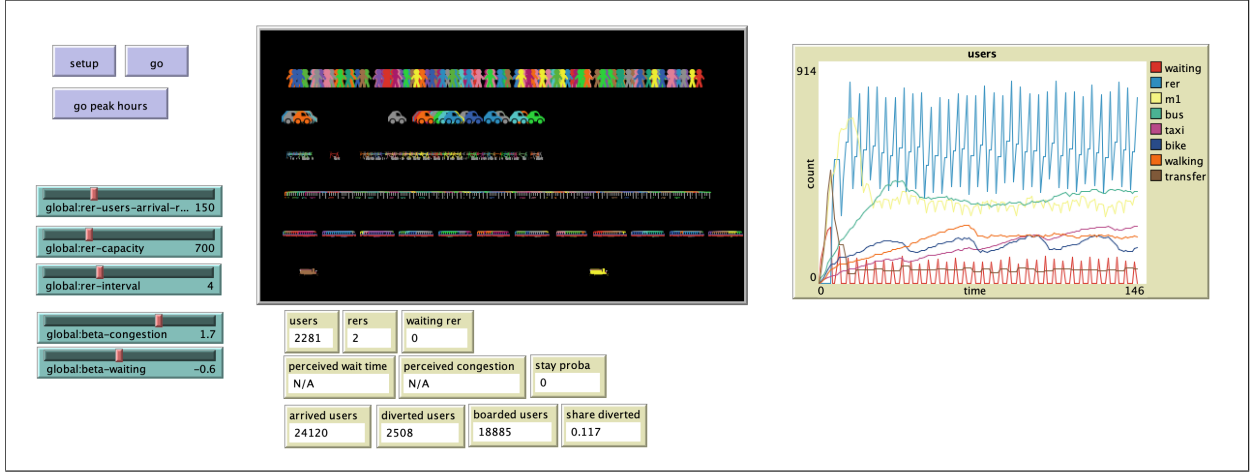
Fig. 2. Model interface. Main varying parameters are included as sliders, while mode queues are visualised in a spatialised manner and as a load plot in time. Various numerical indicators are provided.
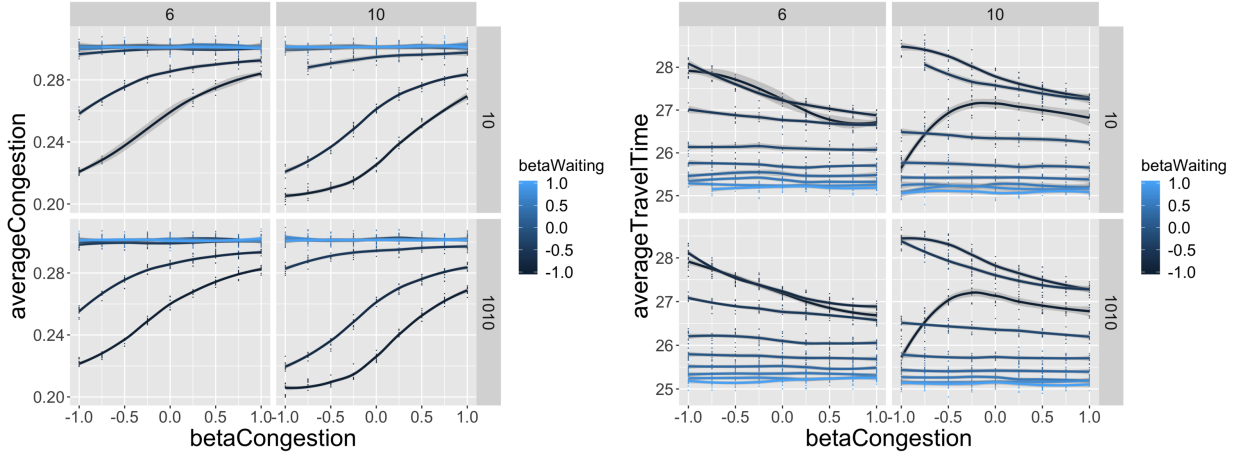

Fig. 3. Model exploration results. *(Left)* Average congestion as a function of β_c, for varying β_τ (colour), train interval (columns) and train capacity (rows); *(Right)* Average travel time for same parameter values.

user arrival rate $\lambda_{RER} = 100$ in experiments, to study borderline scenarios in terms of congestion. Finally, user transfer time between modes is taken as 5 minutes.

4.2. Parameter space exploration

We run 10 stochastic replications of the model, for parameters β_c, β_τ, train capacity C, and train arrival interval I varying within a grid. This corresponds to $\simeq 24,000$ runs of the model. We show in Fig. 3 the variation of average travel time and congestion as a function of β_c. We find an interesting non-linear behavior of travel time (Fig. 3, Right panel) for low values of β_τ (dark blue, right column), where travel time is maximal around a neutral position of users to congestion: in that context of a low tolerance to waiting, either a low or high tolerance to congestion are better than being indifferent. Larger values of β_τ induce a decreasing travel time as a function of β_c (the more users are not tolerant to congestion, the lower the travel time), and similarly travel time is decreasing as a function of β_τ. Regarding the congestion indicator (Left panel of Fig. 3), it increases with both discrete choice parameters but differently in the various disruption scenarios.
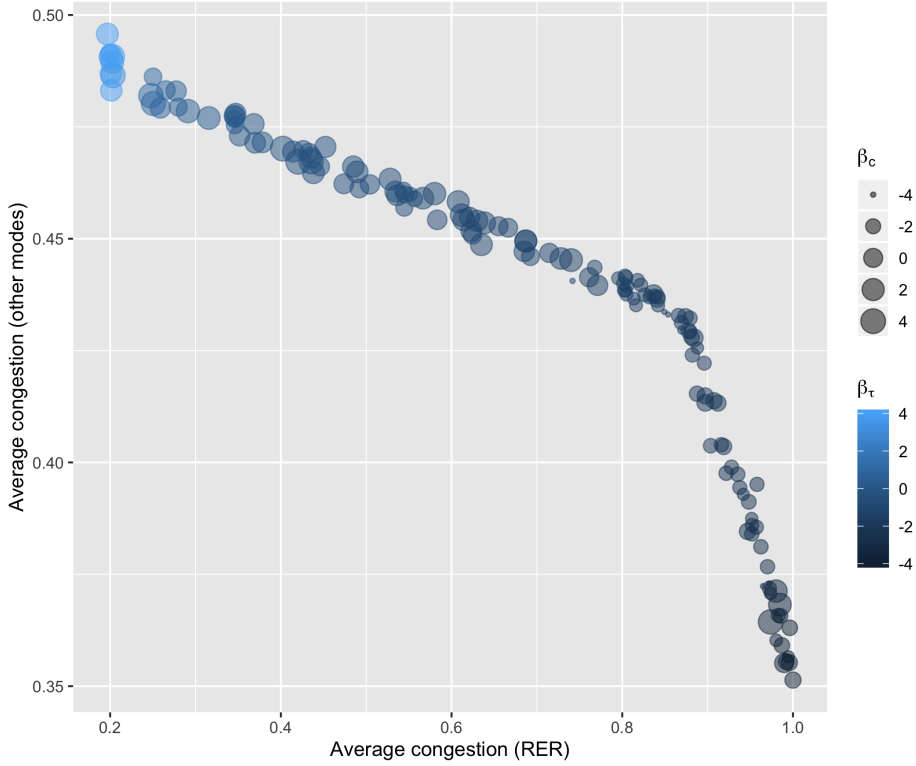
Fig. 4. Model optimisation. We plot the Pareto front of compromise points between the two optimisation objectives. Point colour level gives β_τ while point size gives β_c.

4.3. Model optimisation

We then propose to apply the model in an optimisation context. We use a genetic algorithm for multi-objective optimisation (more precisely the NSGA2 algorithm implemented into the OpenMOLE platform) to search for compromises between congestion in the train and congestion in alternative modes. We run the algorithm in a congested context (arrival rate $\lambda_{RER} = 100$, train capacity $C = 500$ and train interval $I = 5$), with free parameters β_c and β_τ. Optimisation objectives are the average congestion in the RER and the average congestion in other modes.

Simulation results obtained with an algorithm population of $\mu = 200$ after 2000 generations are shown in Fig. 4. We find that acting on different dimensions of user behavior can mitigate the different dimensions of congestion. More precisely, a Pareto front is obtained between congestion in the RER and congestion in the other modes. Congestion in the RER can be minimised, at the detriment of other modes. This occurs when congestion and perceived time parameters have both high values. On the contrary, a low congestion in other modes implies a high tolerance of users and a high congestion in the RER. The shape of the Pareto front is interesting, witnessing two linear regimes. A good compromise is thus found at the breakpoint. At this point, users are rather tolerant to perceived time (negative β_τ values) but can have various behaviours regarding congestion. This shows that mitigating congestion can be achieved by acting on user behaviour, and that a compromise for the multimodal network can be found.

5. Discussion

Our simulation results show how this model can be applied to explore complex congestion patterns in the stylised network following interactions between different components of user behavior. Our model can be applied to the simulation of disruption events, and possible scenarios to improve network resilience. It can also be applied to test interventions to mitigate congestion in targeted modes. Its advantage in comparison to more complicated models

relies on its simplicity, which allows systematic explorations of the parameter space and more advanced numerical experiments as we illustrated above.

Several developments can furthermore be considered. The behavioural model is rather simple, and as the implementation is generic, other functions or more complicated processes can be tested. The variables we used in the discrete choice utility are only proxies of perceived information, and how to account for the information effectively received by the users depending on the context remains an open question. Besides, the possible mitigation of congestion through interventions acting on the provided information may raise ethical issues - how to ensure an overall optimisation while remaining fair to all users. Finally, a more data-driven version of the model, including estimated discrete choice parameters, would be an interesting path to explore, assuming that a further survey and data collection is possible.

References

Ambra, T., Caris, A., Macharis, C., 2019. Should i stay or should i go? assessing intermodal and synchromodal resilience from a decentralized perspective. Sustainability 11, 1765.

Baindur, D., Viegas, J.M., 2011. An agent based model concept for assessing modal share in inter-regional freight transport markets. Journal of Transport Geography 19, 1093–1105.

Bliemer, M.C., Rose, J.M., 2009. Efficiency and sample size requirements for stated choice experiments. Technical Report.

Brisbois, X., 2010. Le processus de décision dans le choix modal: importance des déterminants individuels, symboliques et cognitifs. Ph.D. thesis. Université Pierre Mendès-France-Grenoble II.

Chatterjee, K., Hounsell, N., Firmin, P., Bonsall, P., 2002. Driver response to variable message sign information in london. Transportation research part C: Emerging technologies 10, 149–169.

Chorus, C.G., Molin, E.J., van Wee, B., 2006. Travel information as an instrument to change cardrivers travel choices: a literature review. European Journal of Transport and Infrastructure Research 6.

Dekker, M., van Lieshout, R., Ball, R., Bouman, P., Dekker, S., Dijkstra, H., Goverde, R., Huisman, D., Panja, D., Schaafsma, A., et al., 2018. A next step in disruption management: Combining operations research and complexity science. Technical Report.

Hörl, S., Balac, M., 2020. Reproducible scenarios for agent-based transport simulation: A case study for paris and île-de-france .

Horváth, B., Horváth, R., 2015. Estimation of sample size to forecast travel demand in urban public transport, in: 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE. pp. 300–303.

Kroes, E.P., Sheldon, R.J., 1988. Stated preference methods: an introduction. Journal of transport economics and policy , 11–25.

Leng, N., Corman, F., 2020a. How the issue time of information affects passengers in public transport disruptions: an agent-based simulation approach. Procedia Computer Science 170, 382–389.

Leng, N., Corman, F., 2020b. The role of information availability to passengers in public transport disruptions: An agent-based simulation approach. Transportation Research Part A: Policy and Practice 133, 214–236.

Lyons, G., 2006. The role of information in decision-making with regard to travel, in: IEE Proceedings-Intelligent Transport Systems, IET. pp. 199–212.

Martin, A., Adelé, S., Reutenauer, C., 2016. Stratégies du voyageur: analyse croisée d'entretiens semi-directifs, in: 13ème Journées internationales d'Analyse statistique des Données Textuelles (JADT 2016).

Raney, B., Cetin, N., Völlmy, A., Vrtic, M., Axhausen, K., Nagel, K., 2003. An agent-based microsimulation model of swiss travel: First results. Networks and Spatial Economics 3, 23–41.

Reuillon, R., Leclaire, M., Rey-Coyrehourcq, S., 2013. Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models. Future Generation Computer Systems 29, 1981–1990.

Stamos, I., Mitsakis, E., Salanova, J.M., Aifadopoulou, G., 2015. Impact assessment of extreme weather events on transport networks: A data-driven approach. Transportation research part D: transport and environment 34, 168–178.

Tisue, S., Wilensky, U., 2004. Netlogo: A simple environment for modeling complexity, in: International conference on complex systems, Boston, MA. pp. 16–21.