24th EURO Working Group on Transportation Meeting, EWGT 2021, 8-10 September 2021, Aveiro, Portugal

Scientific workflow systems as platforms to integrate modular urban transportation models

Juste Raimbault[a,b,c,*], Mike Batty[a]

[a] CASA, University College London, London, United Kingdom
[b] UPS CNRS 3611 Complex Systems Institute, Paris, France
[c] UMR CNRS 8504 Géographie-cités, Paris, France

Abstract

Large scale urban transportation models such as four-step models require the integration of heterogenous data and the coupling of sub-models which can already be consequent in terms of complexity. Therefore, such integrated models are difficult to transfer, reproduce, and validate. We propose a modular and reproducible approach based on scientific workflow systems to build and validate such models. We illustrate it by coupling different open-source components within workflows to construct a four-step transportation model applied to all functional urban areas in the UK, and discuss its application to health indicators within public transport in the context of the COVID-19 crisis.

Keywords: Urban transport models; Scientific workflow systems; Modularity; Reproducibility

Urban transportation models such as four-step models, and more generally land-use transport interaction models (Wegener and Fürst, 2004), require the integration of heterogenous data and the coupling of various submodules with possibly high levels of complexity. This raises issues on the one hand for their implementation, transferability and reproducibility, and on the other hand for their validation which requires large scale numerical experiments to validate the submodules and the whole models (Lee Jr, 1973; Batty, 2014). This work proposes to tackle both issues by leveraging modularity and transparency for the construction of large urban models in a modular way, using scientific workflow systems (Barker and Van Hemert, 2007) to couple the different components of models and to launch numerical experiments for their validation.

More particularly, we demonstrate this approach by building a modular four-step multimodal transportation model using only open-source projects. We couple together the MATSim model (MATSim Community (Horni et al., 2016)) to simulate the transportation system, the SPENSER model (University of Leeds, `https://github.com/nismod/microsimulation`) for the generation of synthetic population, the QUANT model (University College London (Milton and Roumpani, 2019)) to estimate spatial interactions, and the spatialdata library (OpenMOLE Community (Raimbault et al., 2020)) for data preparation. The model parts are embedded as docker containers into the DAFNI facility (`https://dafni.ac.uk/`), which workflow system is used to couple them and build the integrated model. DAFNI

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

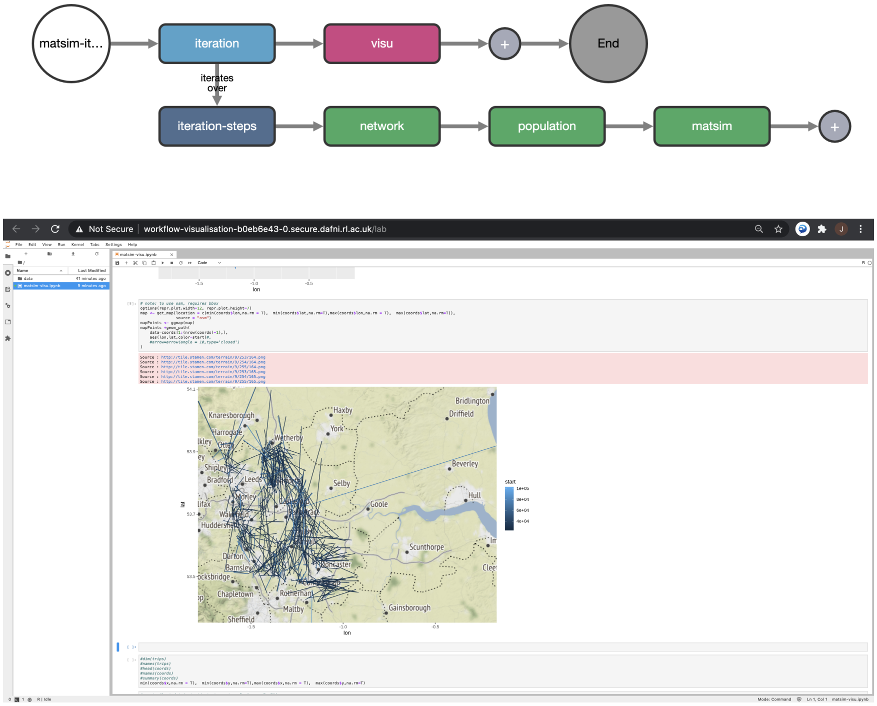

Fig. 1. Construction of the transport model using the DAFNI workflow system. *(Top)* DAFNI workflow including model steps and a computational experiment (Monte Carlo simulations); *(Bottom)* Example of visualisation of model output (MATSim generated trips) within the DAFNI platform.

provides a scientific workflow system for model integration and coupling, direct access to relevant open datasets, visualisation functionalities, and access to a High Performance Computing infrastructure. We show in Fig. 1 the workflow constructed with the interactive workflow builder within the platform, and an example of visualisation of model outputs.

The model is run on the largest functional urban areas (following the definition of Florczyk et al. (2019)) in the UK. We show first results of numerical experiments studying the role of stochasticity on model outputs, for example in Fig. 2 for the statistical distribution of trip departure times (these are iteratively evolved by agents in the MATSim model) for a given urban area. We also show in Fig. 2 the distribution of car trip distances in the different urban areas.

Source code to prepare the model components, input data, and docker containers is available on an open-source git repository at `https://github.com/JusteRaimbault/UrbanDynamics`.

To illustrate the reproducibility of our approach, we test the construction of the model with the OpenMOLE workflow engine (Reuillon et al., 2013), which provides a scripted workflow engine and methods to calibrate and validate simulation models, and suggest advanced numerical experiments for the validation of the coupled model. For example, studying the role of spatial configuration on model outcomes (Raimbault et al., 2019) would be relevant to understand the influence of missing or imprecise data and sampling for the synthetic population.

Work in progress includes the application of this model to the development of health indicators within public transportation, and more particularly linking transportation and work-from-home policies with effective densities in public transport which provide potential exposure indicators in the context of the COVID-19 crisis.

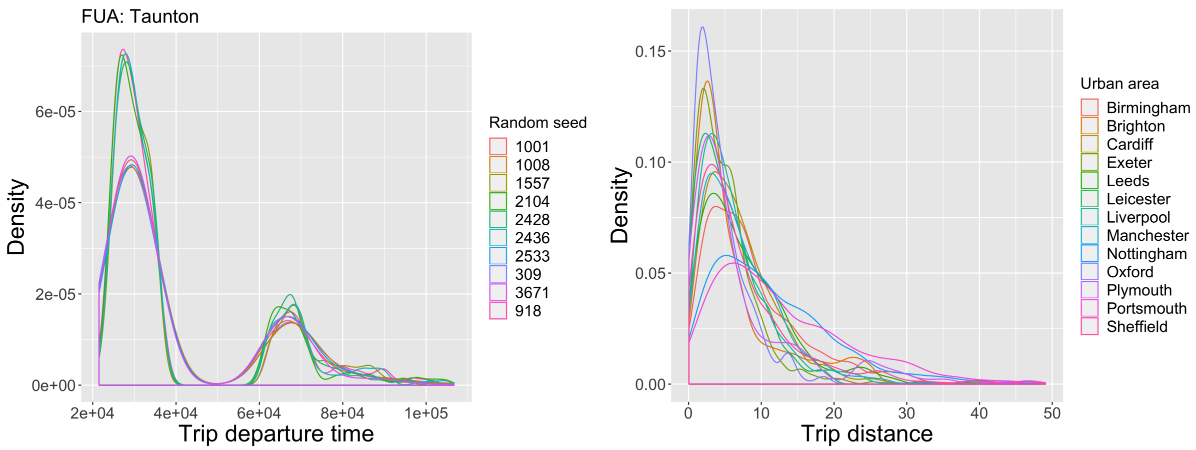

Fig. 2. Results of the simulation of the integrated model on the largest functional urban areas in the UK. *(Left)* Distribution of trip departure times, for several stochastic repetitions on the same urban area; *(Right)* Distribution of trip distances for all urban areas.

References

Barker, A., Van Hemert, J., 2007. Scientific workflow: a survey and research directions, in: International Conference on Parallel Processing and Applied Mathematics, Springer. pp. 746–753.

Batty, M., 2014. Can it happen again? planning support, lee's requiem and the rise of the smart cities movement. Environment and Planning B: Planning and Design 41, 388–391.

Florczyk, A.J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., et al., 2019. Ghsl data package 2019. Luxembourg, EUR 29788, 290498.

Horni, A., Nagel, K., Axhausen, K.W., 2016. The multi-agent transport simulation MATSim. Ubiquity Press.

Lee Jr, D.B., 1973. Requiem for large-scale models. Journal of the American Institute of planners 39, 163–178.

Milton, R., Roumpani, F., 2019. Accelerating urban modelling algorithms with artificial intelligence, in: Proceedings of the 5th International Conference on Geographical Information Systems Theory, Applications and Management, INSTICC. pp. 105–116.

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., Reuillon, R., 2019. Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models. Journal of Artificial Societies and Social Simulation 22.

Raimbault, J., Perret, J., Reuillon, R., 2020. A scala library for spatial sensitivity analysis. Proceedings of GISRUK 2020 .

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.

Wegener, M., Fürst, F., 2004. Land-use transport interaction: State of the art. Available at SSRN 1434678 .