Building and validating modular urban transportation models using scientific workflow systems

J. Raimbault[1,2,3,*] and M. Batty[1]

[1] Center for Advanced Spatial Analysis, University College London
[2] UPS CNRS 3611 Complex Systems Institute Paris Ile-de-France
[3] UMR CNRS 8504 Géographie-cités

[*] j.raimbault@ucl.ac.uk

Headline summary (50-100 words)

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.

Abstract (100-500 words)

Urban transportation models such as four-step models, and more generally land-use transport interaction models, 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. 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 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) to simulate the transportation system, the SPENSER model (University of Leeds) for the generation of synthetic population, the QUANT model (University College London) to estimate spatial interactions, and the spatialdata library (OpenMOLE Community) for data preparation. The model is integrated into the DAFNI facility (https://dafni.ac.uk/) which 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.

The model is run on all functional urban areas in the UK. We show first results of numerical experiments comparing the use of the spatial interaction model with a null model to generate transport demand. We also study the role of stochasticity on model outputs, and show that spatial configuration has a significant influence.

To illustrate the reproducibility of our approach, we sketch the construction of the model with the OpenMOLE workflow engine, 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.

We finally discuss ongoing developments on 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.