



Workshop: Modelling migration and decisions

Eastleigh, Southampton, 21 January 2020

Workshop Report

General information

The Workshop on Modelling migration and decisions took place at the Lakeside Centre in Eastleigh, on 21 January 2020. The workshop was organised as a part of the ERC project on Bayesian Agent-Based Population Studies, and the organisational support has been kindly provided by the ESRC Centre for Population Change.

The aim of the workshop was to discuss, in an informal setting, some promising avenues of modelling migration and migrant decisions, with focus on four areas: migration data, decisions, model construction and analysis. For each area, we will first briefly share our preliminary findings from the project so far, which were subsequently followed by an unfettered exchange of ideas and exploring areas for potential collaboration. Five invited guests took part alongside five members of the project team, with the size of the group providing ideal environment for intense and very thorough discussions throughout the day.

Programme

The workshop was opened at 9:30 by Jakub Bijak (University of Southampton) with a brief introduction to the programme and to the project at large. Throughout the day, five topics were discussed in turn, starting from the broad tenets of agent-based modelling of migration and presentation of a model developed in the project so far, introduced by Martin Hinsch (Southampton), as well as an overview of different data sources on asylum migration, and the data assessment framework, presented on the example on the recent flows from Syria to Europe, which was introduced by Jakub Bijak.



Eastleigh Lakeside Centre

The discussions continued after lunch, with cognitive aspects of migrant decision-making and examples of psychological experiments to inform agent-based modelling introduced by Toby Prike (Southampton), and the choice of software languages for modelling and for documenting models presented by Oliver Reinhardt (University of Rostock). In the final part, Jakub Bijak outlined emulator-based statistical meta-models for analysing model results. The workshop concluded with a roundtable of exchanging ideas, finishing by 17:00.

Summary of the discussions

The discussions offered a wide range of insights into the practice, opportunities and caveats regarding the modelling of migration and decision processes; requirements as to the data and other sources of information for modelling; as well as practical and theoretical consideration of using different programming languages, modelling and analytical approaches.

The discussion started by observing that while different models can serve different purposes, they are useful to expand the imagination of modellers and users alike and to frame the conversation. The modelling process forces the assumptions, concepts and outcome measures to be made and operationalised explicitly.

Even though no canonical modelling approaches exist, a catalogue of models (and their individual sub-modules) can offer a toolbox for describing and explaining of mechanisms being modelled. The modellers need to be clear about the model epistemology and limitations, and it is best when a model serves to describe one, well-defined phenomenon. In this way, models can serve as 'theories of the middle range' (as per Robert Merton and Peter Hedström), and an answer to the question about the 'right' level of analysis may well be a pragmatic and empirical one.

As for the practicalities of modelling, it was stressed that domain-specific languages, despite being to some extent restricted by their construction, enable separating the model logic (formal description of the model and the underlying processes) from the logic of the programming language. Internal domain-specific languages, embedded as libraries in well-known general-purpose languages, seem to offer a sound compromise solution. The exploration of the model space can additionally involve tools of artificial intelligence, such as neural networks, especially when the more traditional methods (e.g. Gaussian process emulators) do not work very well, for example in the presence of tipping points or phase transitions between different model regimes.

In terms of cognitive input, the key challenge was confirmed to be the external validity of the experiments carried out, and their capacity for generalisation, especially with the cultural and socio-economic differences in mind. One suggested way around it (subject to ethical considerations) was to collect experimental data via social media targeting the groups of interest, and to compare the demographic profiles with the samples collected with traditional methods (lab, MTurk). Furthermore, necessary psychological input on the structures of decision-making into the modelling process can be offered by formal description (e.g. the belief-desire-intention model), augmented by formal models for memory, information exchange, etc. It was noted that for migration and similar problems, such as evacuations, modelling the decision processes for 'stayers' can be as important as for 'movers', and that information on perceived needs and expectations of both groups is key.

Besides any statistical analysis, the use of data in modelling can involve face validity tests of the individual model output trajectories, which would confirm the viability of individual-level assumptions. In the discussion it was noted that this approach would provide a *confirmation*, rather than *validation* of the model workings, and that the process of identifying data gaps and requirements could be iterative. At a more general level,



Participants (l-r): Oliver Reinhardt, Elzemies Kortlever, Katarzyna Jaško, André Grow, Eric Silverman, Jason Hilton, Toby Prike, Sarah Wise and Martin Hinsch. Fot: Jakub Bijak

having specific principles and guidelines for using different types of individual data sources in modelling endeavours would be helpful – especially that it would directly feed into the provenance description of the formal relationships within the model, in a modular fashion. In the discussion, the need for minimum reporting requirements for documentation was identified, and the provenance models were found to be complementary, rather than competing with narrative-based approaches, such as the ODD(+D) protocol.

Conclusions and ways forward

For modelling to succeed, open channels for conversation and collaboration across disciplines are crucial, despite constraints in terms of publication and conference ‘silos’. For informed modelling of complex processes such as migration, research needs to be carried out by interdisciplinary teams, with modelling and analytical experts, and diverse, yet complementary expertise of subject matter. Virtual meeting spaces, cross-disciplinary publication and dissemination outlets, journals, blogs, fora, as well as networks such as [CoMSES](#), can help reach the different communities with interest in modelling and its results. A more widespread methods training provision could also speed up the uptake of computational modelling.

The pragmatic considerations around how to do agent-based migration modelling are difficult, but we can help by identifying examples of existing good practice and being precise about the type of research questions such models can answer, as well as being mindful of the epistemological limitations of the various modelling approaches. A related issue of how to make any modelling exercises suitable and attractive for users and policy-makers additionally requires a careful managing of the expectations, to highlight the novelty and potential of the proposed modelling approaches, whilst making sure that what is offered remains realistic in terms of what can be actually delivered.

One important research challenge, when we envisage more work to concentrate in the coming years, is how to combine the different constituting elements of the modelling process together. Here again, having agreed guidelines and examples of good practice would be helpful, both for the research community, as well as for the users. In terms of the quality of input data and other information sources, there is a need to be explicit about what they can tell us, as well as about the quality aspects – in the latter case, explicit provenance modelling can help. At the moment, there is a still untapped potential with using digital trace data, for example from mobile phones or social media, to inform modelling.

Similar issues and challenges are related to experimental information about human cognition: one of the key challenges is how to map the data from the experimental (general) population to a specific one of interest, such as migrants or refugees. One promising possibility consists in ‘dual track’ experimentation on both populations at the same time, in order to try to estimate the biases involved. Finally, the analysis of model results does not at present rely on a standard toolkit of approaches, but the various methods of uncertainty quantification and emulation offer substantial promise. Here, more work needs to be carried out on comparing the results, applicability and trade-offs of using different meta-models for analysis.

The conclusion, shared by everyone, was that open discussions around good practice, exploring different approaches to migration modelling and decisions, matter a lot both for the practitioners as well as theorists and methodologists, especially in such a complex and uncertain area as migration.

More detailed information and agenda can be found on the workshop website:

<https://www.southampton.ac.uk/baps/news/events/2020/01/2020-01-21-workshop-migration-decisions.page?>