This space is reserved for the Procedia header, do not use it

# Simulating Refugee Movements: Where would you go?

Derek Groen[1,2]

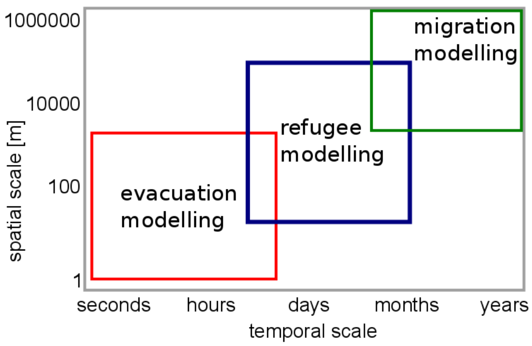
[1] Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UK
Derek.Groen@brunel.ac.uk
[2] Centre for Computational Science, University College London, London, UK.

**Abstract**
The challenge of understanding refugee movements is huge and affects countries worldwide on a daily basis. Yet, in terms of simulation, the challenge appears to have been largely ignored. I argue that we as researchers can, and should, harness our computational skills to better understand and predict refugee movements. I reflect on the computational challenges of modelling refugees, and present a simulation case study example focused on the Northern Mali Conflict in 2012. Compared to UNHCR data, the simulation predicts fewer refugees moving towards Mauritania, and more refugees moving towards Niger. This outcome aligns with UNHCR reports, which mention that unregistered refugees were known to reside outside of the official camps, though further investigations are required to rule out competing theories.

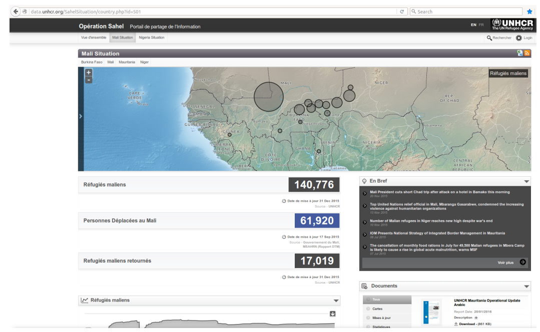
*Keywords:* modelling,simulation,refugees,migration,agent-based modelling

## 1 Introduction

People flee their homeland for very good reasons. Yet the macroscopic phenomenon of large influxes of refugees severely concerns both governments and members of the public. Some countries open their borders, and attempt to set up adequate facilities, whereas other countries close their borders, in an attempt to redirect the refugees. These decisions carry strong moral importance but the exact impact often remains unknown. What is clear, however, is that the area of impact exceeds that of single countries, and that a solution implemented by one country could result in major difficulties for another. Indeed, the challenge of understanding the movements of refugees is huge, global, and as a society we are confronted with it on a daily basis. Yet, in terms of simulation, the challenge appears to have been largely ignored. I argue that we as researchers can, and should, harness our computational skills to try and address it for three reasons: First, through the use of simulations, I believe we can estimate what would have been different if governments had chosen to apply different border and/or immigration policies, and help inform governments (as well as the general public) about the possible consquences of their decisions. Second, simulations could help predict where refugees are likely to go when a conflict erupts, allowing support organizations to make advance preparations. Third, most of




(a) Characteristic spatial and temporal scales of evacuation models, refugee models, and migration models.




(b) Screenshot of the UNHCR data portal at http://data.unhcr.org.

the empirical data currently available about refugees is incomplete; here simulations can help to find gaps in the monitoring infrastructure, and to make predictions of refugee populations in areas where little or no empirical information is available. Simulations have a particularly high added value potential in refugee settings, both because refugee data is ill-suited for determining causal inferences, and because running simulations does not carry the ethical burden that hinders empirical experiments.

To my knowledge, few researchers have done computational studies on the movement of refugees. However, extensive research has been done both on evacuation, which focuses on smaller space and time scales, and migration modelling, which focuses on much larger time scales. Within this literature, refugee modelling is a unique application domain because its spatial scale can span up to half the globe (e.g., Syrian refugees are known to have reached North Scandinavia), while its temporal scale can be as short as a matter of days, or even hours (see Fig. 1(a)). Historically, Schmeidl et al. [6] argued that there are fundamental obstacles to developing quantitative prediction models for refugee movements. However, their fundamental concern on the lack of data availability arguably becomes less important now as the UNHCR has made refugee data publicly available, and refugees themselves increasingly have access to online platforms. On modelling refugees, Sokolowski et al. [8, 7] have proposed a detailed agent-based modelling (ABM) framework to model refugees and other factions in the Syrian city of Aleppo, while Latek et al. [5] calibrated an ABM tool to model decision-making and dynamic processes around refugees in Syria. Both models are primarily applied for sensitivity analysis, and feature a relatively large number of free parameters. Anderson et al. [1] proposed an ABM approach to model refugee communities. They focus on the interaction of refugees with other stakeholders; models which could serve to inform policy decisions for governments and NGOs. Lastly, Hailegiorgis et al. [4] modelled the spread of cholera within a refugee camp.

In this paper, I present a model on the 2012 Northern Mali Conflict and compare the predictions from the model with data provided by the UNHCR (see Fig. 1(b)).

## 2 Case study: Northern Mali Conflict (2012)

As a case study example, I examine the Northern Mali Conflict in 2012. This conflict erupted on January 16th 2012, when Touareg rebels began to conquer places in Northern Mali, starting a civil war. I focus particularly on the refugee movements that occur in this conflict after February 29th, as little refugee data is available prior to that day. Through simulation I aim to explore the patterns of refugee movements, and to compare the simulation results with the UNHCR data. For brevity, I abbreviate the neighbouring countries Mauretania (MAU), Burkina Faso (BF), and Niger (NI).

I present my network-based ABM model, including its key parameters and assumptions, in Fig. 1. The number of new refugees in my model is derived from the UNHCR-reported total refugee count (data.unhcr.org), using linear interpolation between data points. Locations are interconnected with paths and refugee agents are spawned in Kidal, and later in Timbuktu and Gao. The refugee agents move with a probability of *movechance* to different locations (movechance equals 1.0 in source locations). I assume that the average refugee stays in a camp for 1000 days, and I used UNHCR data from the first 60 days in my model design. Note that UNHCR registers refugees at the destination camps, while I use these numbers to populate the source points. As a results, simulated refugees depart for their journeys at least one day later than their real-life counterparts. I do not model any Internally Displaced People, as I do not possess systematic data of their exact locations. When flagged as moving, a refugee chooses its destination using a weighted probability function, using weights equal to 1 divided by the route length. I estimated travel distances between locations by using the shortest route planned for cars in Bing Maps (maps.bing.com). In doing so, I assume that refugees stick to major roads, and travel with (shared) vehicles. I assume all refugee movements take one day, as I do not possess data on travel times for refugees. To run the simulations I developed a Python toolkit, named Flee. I also use the Pandas library (pandas.pydata.org) in combination with matplotlib to analyze the data. To verify the model, I compare my results with UNHCR data, and calculate the sum of absolute differences in refugee counts for each camp as a proportion of the total number of refugees,

$$E(t) = \frac{\sum_{x \in S} (|n_{x,t} - m_{x,t}|)}{N}. \tag{1}$$

Here the number of refugees found in each camp $x$ of the set of all camps $S$ at time $t$ is given by $n_{x,t}$ based on the simulation predictions, and by $m_{x,t}$ based on the UNHCR data. The total number of refugees reported in the UNHCR data is given by $N$.

**Results:** I provide a comparison of the simulation results to the UNHCR data in Fig. 2. Here, organized refugee transfers from Fassala (MAU) cause a steep initial increase in refugee count in Mbera, both in the data and in the simulation. Decreases in refugee counts after Day 150 in the data are caused by new, more thorough, UNHCR registration requirements, which reduced refugee counts by as much as 68% after the "recount". Compared to our simulation predictions, the UNHCR data reports more refugees moving into Mauretania, and fewer of them moving into Niger (with the exception of Niamey). This may be caused by the fact that refugee tracking in Niger has not been as comprehensive as in Mauretania, or by the fact that the Niger border only permits refugees to cross it on foot (which is not captured in the model, due to lack of transportation data). The averaged difference $E(t)$ remains below 0.5 throughout the run (see bottom left panel in Fig. 2), and is severely affected by the new registration requirements after day 150. The total number of refugees in camps in the simulation (shown in the bottom right) is slightly less than reported in the data, as a fraction of the refugee agents remain in transit within Mali.

# 3 Discussion

I have presented a first case study on simulating the movements of refugees during the early stages of the North Mali conflict. The complex nature of civil wars and the lack of systematic reports makes it challenging to accurately model refugee movements. However, with the availability of UNHCR data and solid mapping platforms I have shown that it is possible to construct, and validate, basic simulations of refugee movements. The first results of my simulation allow for multiple interpretations, and give rise to a need for larger-scale investigations. Nevertheless, this model can be already used to gain some clarity on how the flow of refugees could have changed if Burkina Faso and Niger had opened their borders from Day 1 of the conflict, or if Mauretania has decided to shut them at a later point.

I this case study I used a simplistic ABM approach to explore the area of refugee modelling. However, sophisticated ABM frameworks certainly do exist [9, 2], and I do intend to adopt these as I move further towards more complicated and better informed refugee models. In particular, RePast HPC [2] is interesting, as it scales well to 10,000s of cores and allows users to use multiple ABM approaches concurrently. In future work I therefore aim to use RePast in conjunction with FabSim [3], to perform systematic large-scale investigations on this topic.

**Acknowledgments**

I am grateful to my students Rohit Prashar, Neaman Doushouki and Gaveen Basra, who worked in parallel with me on refugee modelling in different areas. In addition, I thank Dr. Moqi Groen-Xu, Dr. David Bell and Dr. Julie Eatock for their valuable feedback and constructive discussions regarding this work.

# References

[1] James Anderson, Alok Chaturvedi, and Mike Cibulskis. Simulation tools for developing policies for complex systems: Modeling the health and safety of refugee communities. *Health Care Management Science*, 10(4):331–339, 2007.

[2] Nicholson Collier and Michael North. Parallel agent-based simulation with repast for high performance computing. *SIMULATION*, 2012.

[3] D. Groen, A. Bhati, J. Suter, J. Hetherington, S. Zasada, and P. Coveney. FabSim: facilitating computational research through automation on large-scale and distributed e-infrastructures. *ArXiv:1512.02194*, December 2015.

[4] Atesmachew Hailegiorgis and Andrew T Crooks. Agent-based modeling for humanitarian issues: disease and refugee camps. In *The Computational Social Science Society of America Conference, Santa Fe, NM*, 2012.

[5] M. M. Latek, S. M. M. Rizi, and A. Geller. Verification through calibration: An approach and a case study of a model of conflict in syria. In *Simulation Conference (WSC), 2013 Winter*, pages 1649–1660, Dec 2013.

[6] Susanne Schmeidl and J Craig Jenkins. Issues in quantitative modelling in the early warning of refugee migration. *Refuge: Canada's Journal on Refugees*, 15(4), 1996.

[7] John A. Sokolowski and Catherine M. Banks. A methodology for environment and agent development to model population displacement. In *Proceedings of the 2014 Symposium on Agent Directed Simulation*, ADS '14, pages 3:1–3:11, San Diego, CA, USA, 2014. Society for Computer Simulation International.

[8] John A. Sokolowski, Catherine M. Banks, and Reginald L. Hayes. Modeling population displacement in the syrian city of aleppo. In *Proceedings of the 2014 Winter Simulation Conference*, WSC '14, pages 252–263, Piscataway, NJ, USA, 2014. IEEE Press.

[9] Weihao Yin, Pamela Murray-Tuite, Satish V. Ukkusuri, and Hugh Gladwin. An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transportation Research Part C: Emerging Technologies*, 42:44 – 59, 2014.

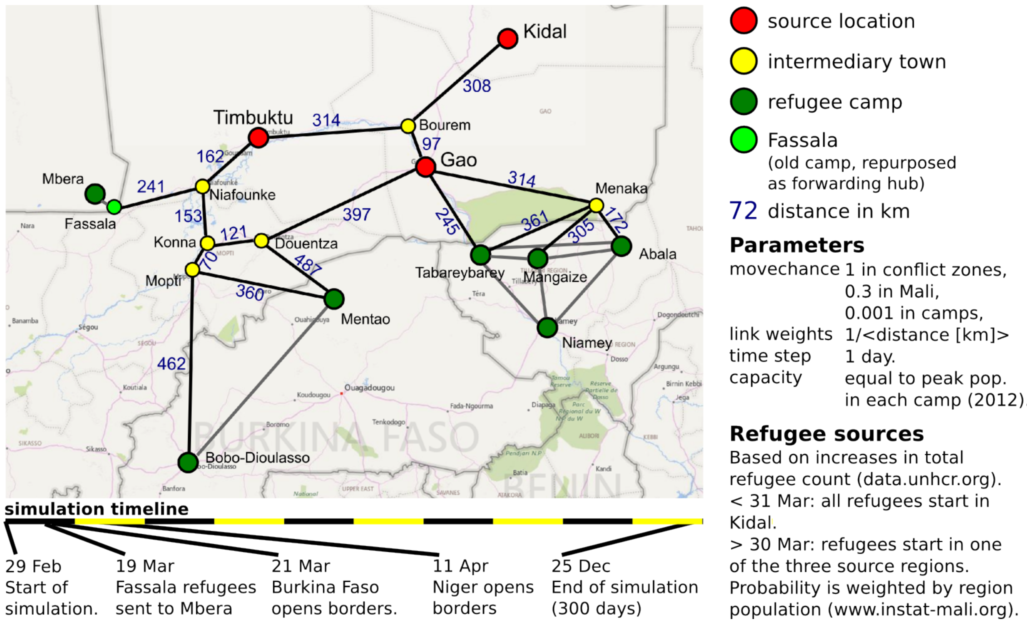



Figure 1: Overview of the simulation model.

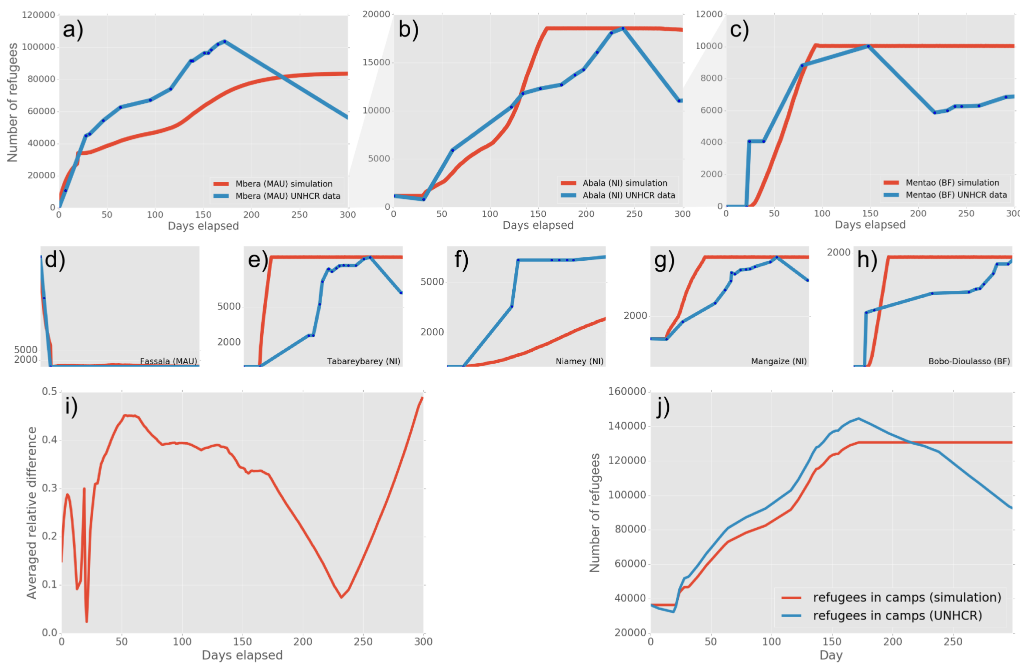



Figure 2: (a-c) comparison of simulation and UNHCR data for the three largest camps, (a) Mbera, (b) Abala and (c) Mentao (c). (d-h) Quick reference comparison for the other five camps. (i) Relative difference $E(t)$ averaged across locations, calculated using Eq. 1. (j) Number of refugees residing in camps, based on the simulation results (red) and the interpolated UNHCR data (blue).