A Generalizable Theory-Driven Agent-Based Framework to Study Conflict-Induced Forced Migration

Zakaria Mehrab[1], **Logan Stundal**[1], **Srinivasan Venkatramanan**[1], **Samarth Swarup**[1], **Bryan Lewis**[1], **Henning Mortveit**[1], **Christopher Barrett**[1], **Abhishek Pandey**[2], **Chad Wells**[2], **Alison Galvani**[2], **Burton Singer**[3], **Seyed Moghadas**[4], **David Leblang**[1], **Rita Colwell**[5], **Madhav Marathe**[1]

[1]University of Virginia
[2]Yale School of Public Health
[3]University of Florida
[4]York University
[5]University of Maryland

Abstract

Large-scale population displacements arising from conflict-induced forced migration generate uncertainty and introduce several policy challenges. Addressing these concerns requires an interdisciplinary approach that integrates knowledge from both computational modeling and social sciences. We propose a generalized computational agent-based modeling framework grounded by Theory of Planned Behavior to model conflict-induced migration outflows within Ukraine during the start of that conflict in 2022. Existing migration modeling frameworks that attempt to address policy implications primarily focus on destination while leaving absent a generalized computational framework grounded by social theory focused on the conflict-induced region. We propose an agent-based framework utilizing a spatiotemporal gravity model and a Bi-threshold model over a Graph Dynamical System to update migration status of agents in conflict-induced regions at fine temporal and spatial granularity. This approach significantly outperforms previous work when examining the case of Russian invasion in Ukraine. Policy implications of the proposed framework are demonstrated by modeling the migration behavior of Ukrainian civilians attempting to flee from regions encircled by Russian forces. We also showcase the generalizability of the model by simulating a past conflict in Burundi, an alternative conflict setting. Results demonstrate the utility of the framework for assessing conflict-induced migration in varied settings as well as identifying vulnerable civilian populations.

Introduction

The large-scale forced migration caused by the 2022 Russian invasion of Ukraine led to the largest refugee flows in Europe since World War II. Around 8.3 million Ukrainians took refuge in different parts of Europe while another 5.4 million remained in Ukraine as internally displaced persons (IDP), as of May 2023. This population displacement has caused almost 10.2 million Ukrainians to need humanitarian assistance (UNOCHA 2023). Understanding the dynamics of migration is essential for policymakers to plan resource distribution and other logistical decisions to support these people in need of assistance. Reactive approaches without adequate information have failed in the past to meet the requirements arising from refugee surges (Suhrke et al. 2000). Among different dimensions of migration, information on the time, location of migration, and demographics of the migrants are of interest to policymakers (Clemens et al. 2018). Due to the sudden nature of forced migration, the availability of such information is not trivial (Adhikari 2013). Thus, a computational framework to explain these dynamics of migration would immensely help inform policymaking.

Despite several works in computational modeling that have tried to understand these dynamics from the perspective of destination countries (Suleimenova et al. 2017; Davis, Bhattachan et al. 2018; Asgary, Solis et al. 2016), we observe few works attempting to study these questions from the perspective of the origin country. And, if we take a step back, we realize that prior to understanding the choice of destination, we need to understand which individuals among the population in the affected region undergo forced migration. Moreover, these studies are country-specific (Asgary, Solis et al. 2016; Davis, Bhattachan et al. 2018; Searle and van Vuuren 2021; Smith 2014). To the best of our knowledge, there is no existing work that attempts to explain conflict-induced migration outflows from a social-theoretic standpoint. A proper approach to this problem requires interdisciplinary collaboration between social scientists and computational scientists (Frydenlund and De Kock 2020).

In order to bridge the gap between social science and computational science in the context of conflict-induced forced migration, we must ponder carefully on the choice of the computational technique. One option is to choose traditional models of migration like the Gravity Model (Zipf 1946) or the Radiation Model (Simini et al. 2012). However, these are more suitable for voluntary migration, which is carried out over a longer period and less sudden, contrary to forced migration. Also, the functional forms of such models do not allow for straightforwardly accommodating conflict events. Another alternative is to use a Machine Learning (ML) approach, which has shown promise in studying migration destinations (Robinson and Dilkina 2018). However, ML techniques are inadequate in modeling the different decision processes of heterogeneous groups of people (fairness), and using such a technique makes it difficult to understand how different components interact with each other (explainability). Also, social theory comprises a set of rules that are difficult to model by these techniques.

Due to these issues, we approach this problem using the agent-based modeling (ABM) technique, which implicitly makes several AI methodologies and topical areas applicable in our context. ABM is well-suited for studying the problems where decisions and interactions of individuals result in an aggregated phenomenon (Asgary, Solis et al. 2016; Bijak, Higham et al. 2020). Thus, ABM can be used to observe the dynamics of emergent social phenomena in the advent of shock events (**Social Simulation**). Generally, in an ABM framework, each agent carries out some action depending on their own characteristics and their interaction with the environment or other agents following a set of simple rules (**Explainable AI**). ABM makes it flexible to incorporate various data as environments or agents. It also allows for different rules for different types of agents, ensuring fairness and equity (**Fairness**). Moreover, the rules of the agents can be incorporated from an appropriate social theory (**Social Science Guided AI**), making ABM the choice of a computational model that can be incorporated with social theory most seamlessly. Overall, in our context, ABM is like a planning algorithm in the form of a set of rules that help the agents reason about their actions toward the specific goal of migration (**Goal Directed Behavior**).

This study attempts to understand the following research questions. First, how to formally and computationally formulate a social theory within an ABM framework in the context of conflict-induced forced migration. Second, how does the model perform compared to existing models, and how generalizable it is for other countries. Third, how this model can be used in other policy-relevant use cases. In order to answer the first question, we employ the *Theory of Planned Behavior* (Ajzen 1991) to design a hierarchical agent-based framework that uses a spatiotemporal gravity model to capture the interaction between agents and events and a bi-threshold model atop a graph dynamical system (GDS) to capture the interaction between agents. To answer our second question, we compare the performance of our model with the model proposed by (Pandey et al. 2023) in capturing different dimensions of forced migration in Ukraine. We discover that our model outperforms the existing model in capturing the daily trend of refugees as well as generating other summary statistics. We also show that our model can easily be transferred to study other similar forced migration events, using the Burundi conflict as an example. Finally, we showcase how the model can be used to study *entrapped population* by imposing "entrapment" conditions for certain regions to limit migration for a period of time. After relaxing this restraint, we compute subsequent refugee surges to demonstrate the model's capability to describe previously trapped populations fleeing during windows of opportunity.

Related Work

Voluntary migration: Early models for *voluntary migration* include the Gravity Model (Zipf 1946) and Radiation Model (Simini et al. 2012). The Radiation model has recently been extended to incorporate pull factors other than population (Alis et al. 2021). Used primarily for *voluntary migration*, these models assume that the outflow of migration from the country of origin is already known and uses that as input to understand the destinations of intending migrants. Recently, AI communities have grown interested in using AI techniques to study migration. Yang et al. (2018) used regression techniques to distinguish migrants from locals in Shanghai using their call log records. Robinson and Dilkina (2018) showed that ML models can outperform traditional models in predicting migration destinations at county and country levels. However, they mostly focus on destinations and the data required to train such models are infeasible to obtain during forced migration.

Forced migration: Prior forced migration research primarily focuses on natural disaster-induced migration. Hassani-Mahmooei and Parris (2012) studied impending migration in Bangladesh due to climate change and concluded that in 40 years, there will be large population shift towards the eastern regions. Davis, Bhattachan et al. (2018) proposed a universal model for studying migration flows due to sea level rise and, using Bangladesh as a case study, estimated resulting end-of-century displacement and resource needs for the displaced. Temporal scale of these studies is quite long-term compared to *forced migration* as a result of forced events.

Several works have also tried to model conflict-induced *forced migration*. Searle and van Vuuren (2021) proposed a framework for modeling forced migration as a result of conflict events and used push and pull theory to determine the destination of migrants from Syria under that framework. Suleimenova et al. (2017) developed FLEE, a generalized simulation model that can be used to estimate the destination of migrants from a conflict-induced region. However, their work assumes that the people who want to migrate are already known and their model has to use this information as input. The work by Pandey et al. (2023) is perhaps the closest to us. They proposed an exponential decay model where they considered locations as agents and estimated the number of migrants from each location considering surrounding events. However, their method is neither backed by any social theory nor does it take interaction between agents into account, something we address in our study.

Methods

Underlying Social Theory

According to the *Theory of Planned Behavior*, the decision in response to an event is built upon three phases: a) *Attitude*, b) *Perceived behavior control (PBC)* and c) *Subjective norm*. **Attitude** can be correlated to the impact associated with an event observed by an individual (Kniveton, Smith, and Black 2012). **PBC** accounts for the fact that the same observed impact can be perceived differently across individuals, depending on their demographic attributes or experiences. **Subjective Norm** is correlated with the social acceptability of a particular action.

Before accommodating these constructs computationally, we want to paint a picture of how an agent in a conflict-induced region may, through these constructs, decide to migrate or not. First, the conflict events surrounding an agent may instill a sense of risk. Second, agents in the same households communicate among themselves to form an initial migration ***intention***. Finally, households in the same neighborhood communicate with each other to take the final ***decision***. The first step follows the *attitude* and *pbc* phase and *subjective norm* comes into play in the next two steps.

Input and Features

Holistically, the input space encompasses a set of conflict events $C = \{c_1, c_2, .., c_j, ..\}$, a set of person agents $A = \{a_1, a_2, .., a_i, ..\}$ and a set of household agents $H = \{h_1, h_2, ...h_k, ...\}$. Each event c_j is associated with the following features: its location L_j, time of the event T_j, and its severity S_j. We also assume that the following features are known about a person agent. These features are the location of an agent a_i at time t, x_i^t, and the risk perception, β_i; which is an indicator of their level of non-tolerance to the overall risk observed from events. We are also given the mapping function $\eta : A \rightarrow H$, which maps each person agent a_i to a household agent h_k. We assume that migration decision occurs at the household agent level and once a household agent migrates, all associated person agents migrate as well.

Problem Modeling and Framework

We model our problem by dividing it into two phases. The first phase is called the *Intention Phase* where an initial intention to migrate is formed based on interaction with events and limited interaction with agents within household and the last phase is where households communicate with each other to make the final decision of migration is called the *Decision Phase*. Each phase corresponds to one timestep. Agents update their states after every *decision phase*, indicating that the system evolves at every even timestep.

Intention Phase: The problem for the intention phase can be modeled as follows: At time t, given the remaining agents $A(t)$, the conflict events C, the agent-household mapping η; find the intention to migrate for each household $h_k^t \in H(t)$, where $H(t) = \{\eta(a_i) : a_i \in A(t)\}$, find the migration intention of the households at timestep $t + 1$. The intention phase is further subdivided into two sub-phases:

Risk-perception sub-phase: At time t, given the remaining agents $A(t)$ and the conflict events C, find the perceived risk of each agent $a_i^t \in A(t)$. To functionally formulate the perceived risk of an agent as a result of conflict events, we seek help from literature (Weidmann and Zürcher 2013). First, we calculate the perceived risk pr_i^t based on a modified version of their proposed *spatiotemporal gravity model* to take into account the severity of the event as well as an agent's own perception of the event.

$$pr_i^t = \begin{cases} \beta_i \sum\limits_{c_j \in C} \Delta_S(x_i^t, L_j)^\delta \Delta_T(t, T_j)^\tau S_j & \text{if } T_j <= t \\ \epsilon & \text{otherwise} \end{cases}$$

Here, δ_S and δ_T are functions to calculate the distance and time passed between the location of the agent and the event, respectively. The function also utilizes a spatial-decay parameter (δ) and a temporal-decay parameter (τ). Both of these parameters should be negative; ensuring that events less recent and further away should have less impact on the observed risk. Furthermore, since future events are not known beforehand, we do not consider them in the calculation of the observed risk. Since the same observed risk can be perceived differently by different agents according to the PBC construct of the *theory of planned behavior*, we scale the observed risk (summation term) by their non-tolerance (β_i) to calculate perceived risk. This ensures heterogeneity among various demographic groups.

Household-intention sub-phase: At time t, given the perceived risks pr_i^t for each agent $a_i^t \in A(t)$, calculate the migration intention for each household $h_k^t \in H(t)$ for time $t + 1$. The probability of migration intention of a household is calculated as follows:

$$Pr(h_k^t \text{ intends to migrate}) = \sum_{a_i \in \eta^{-1}(h_k^t)} \int(pr_i^t)/|\eta^{-1}(h_k^t)|$$

Here, $\int$ is an activation function to transform the perceived risk into a probabilistic value. In our case, we employ the general sigmoid function of the form $\int(x) = \frac{1}{1+Qe^{-rx}}$, where Q is the no-risk migration parameter and r is the growth-rate parameter. Afterward, the migration intention (MI) for each household h_k is calculated as follows.

$$MI_k^{t+1} \leftarrow X \sim \text{Bern}(Pr(h_k^t \text{ intends to migrate}))$$

Decision Phase: The problem for the decision phase can be modeled as follows: Given the network of households $G(V, E)$, the migration intention of the households at time $t+1$, the migration decisions of the households at time t; find the migration decisions of the households at time $t + 2$. In order to model the decision phase, we utilize Graphical Dynamical System (GDS), a popular framework that has been used to study bio-social systems (Adiga et al. 2019; Qiu et al. 2022; Alam et al. 2015). In its primary form, GDS is modeled with a graph G, a set K specifying the domain of the vertex states, a family of local state transition functions F, and an update scheme U specifying the sequence by which the local transition functions are processed. We define our update scheme to be synchronous, making our system a Synchronous dynamical system (SyDS).

Our SyDS is defined as a tuple $S = (G, F)$. Here $G(V, E)$ is the given household network where $V = H$ and $F = \{f_k : \forall h_k \in H\}$ is a set of local transition functions each vertex computes. Each vertex $v \in V$ has the following states at time $t + 1$ as described below.

$$x_v^{t+1} = (MI_v^{t+1}, M_v^t)$$

where $MI_v^{t+1} = \{0, 1\}$ is the migration intention calculated from the *intention phase* and $M_v^t = \{N, I, R\}$ denotes the migration status of household v at timestep t; N being the household has not migrated yet, I being the household has migrated as internally displaced and R being the household has migrated as refugee. Whether a household migrates or not next is decided by the following rules:

- Households already migrated will take no further action.
- Non-intending households will migrate if more than I_h neighbors are intending to migrate.
- Intending households will not migrate if less than I_l neighbors are intending to migrate.

B_i of agents				W_j of events	
Group	B_i (NF)	B_i (F)		Event Type	W_j
Elderly	0.15	0.20		Explosion	3
Child	0.50	0.70		Battle/Violence	2
Adult Male	0.02	0.05		Protest/Riots	1
Adult Female	0.20	0.70		Other	0

• NF = No Family • F = With Family

Table 1: Weights of different demographic groups and events

- The household would migrate based on their intention to migrate if none of the above holds.

We assume that if a household decides to migrate, with λ $(0 < \lambda < 1)$ probability they become a refugee and with $1 - \lambda$ probability they become internally displaced (IDP) (Pandey et al. 2023). Let, $\hat{f} : \mathbb{R} \rightarrow \{N, I, R\}$ be a function where $\hat{f}(\lambda)$ is R with probability λ and I with probability $1 - \lambda$ when $0 < \lambda < 1$. Otherwise, $\hat{f}(\lambda) = N$. Also, let $\sum_{u \in \mathcal{N}(v)} MI_u^{t+1} = \psi_v^{t+1}$ denote the number of neighbors of vertex v intending to migrate at time $t+1$. Given this, we can formally write the local transition function f_v^{t+2} for updating the migration decision state M_v^{t+2} of a vertex v at time $t + 2$ using a Bi-Threshold function (Kuhlman, Kumar et al. 2011) as follows.

$$f_v^{t+2}(x_v^{t+1}) = \begin{cases} M_v^t & \text{if } M_v^t = I \text{ or } M_v^t = R \\ \hat{f}(\lambda) & \text{if } M_v^t = N \text{ and } \psi_v^{t+1} \geq I_h \\ N & \text{if } M_v^t = N \text{ and } \psi_v^{t+1} < I_l \\ \hat{f}(\lambda MI_v^{t+1}) & otherwise \end{cases}$$

After this local transition function is computed, before the next intention phase at time $t + 2$, households with $M_v^{t+2} = N$ and corresponding agents are retained from $A(t)$ and $H(t)$ and the rest are removed. Moreover, by looking at the values $M_v^0, M_v^2, ...$, we can identify the migrating households at different timestamps.

Datasets

Agent Data: The synthetic population[1] developed by the Biocomplexity Institute (Mortveit et al. 2020) contains information about synthetic individuals (i.e. age, gender) and their households (i.e. location). Since we do not have real-time location information about person agents, we assume the location of each person agent to be the same as the location of their households across all timesteps. We also divide the person agents into four demographic groups based on their age and gender (elderly, child, adult male, adult female). We choose $\beta_i = b \times B_i$ for an agent where B_i is chosen based on the demographic group and family status of the agent based on Table 1 and b is a scaling parameter. These values were chosen based on previous literatures (Castelli 2018; Brunarska and Ivlevs 2022) and consultation with social scientists.

Conflict Data: We obtain the conflict information from Armed Conflict Location & Event Data Project (ACLED) (Raleigh et al. 2010). It contains records of conflicts and violent events happening across the globe and the data is updated daily. For each event, it has the location and time of the corresponding event, which we can feed directly into our proposed spatiotemporal gravity model. Ws calculate severity S_j as $S_j = w \times W_j \times I_j$. Here, I_j is the fatality or impact of event j which is recorded in ACLED as a field. W_j is a weight associated with the type of the event as defined in Table 1 and w is an associated scaling parameter. The weights are chosen from weights defined in the GDELT-CAMEO codebook (Carammia, Iacus, and Wilkin 2022).

Observed Data: Although not necessary by the framework, we require observed data to calibrate the model parameters. Obtaining ground truth about individual agents is realistically infeasible. Rather, we use border crossing data from Humanitarian Data Exchange (HUMDATA 2022) which collects border crossings from a conflict-induced country at a reasonable temporality, even daily in the best case. Note that, one can also use other summary statistics from surveys conducted by many humanitarian organizations. However, since this data has a good temporal resolution, calibrating against this data will make the model more robust in estimating other summary statistics automatically.

[1] Synthetic Population Data available at https://net.science/files/40e8d15e-d38b-48d4-aaff-79e85e1de87e/

Symbol	Parameter Name	Prior
δ	Spatial decay	$[-5, 0)$
τ	Temporal decay	$[-5, 0]$
Q	No-risk migration	$[0, 100]$
r	Growth rate	$(0, 5]$
b	Risk scale	$(0, 1]$
w	Event scale	$(0, 1]$
I_l	Low threshold	$\{x \in \mathbb{Z} : 1 \leq x \leq 5\}$
I_h	High threshold	$\{x \in \mathbb{Z} : 1 \leq x \leq 50\}$
λ	Refugee ratio	$(0, 1)$

Table 2: Parameters of the proposed framework

Calibration

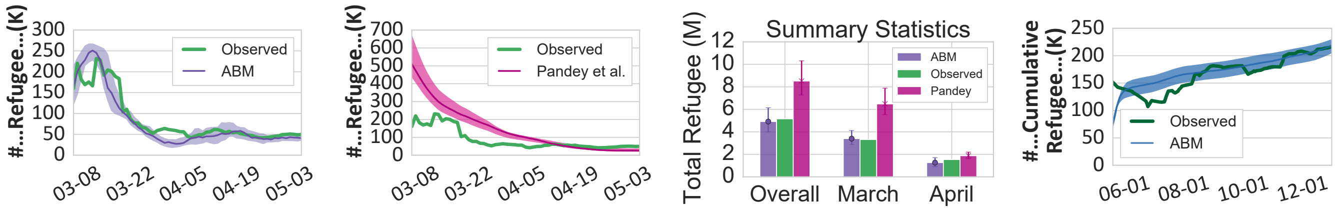
As per our proposed design, we have defined several parameters we have to calibrate. Table 2 shows these parameters and their priors. To calibrate the parameters, we employ the Bayesian Optimization strategy (Mockus 1989). Bayesian optimization strategy tries to best fit the objective function by creating a posterior distribution of functions. As the number of observations increases, it becomes more confident about which parameter space is worth exploring.

The objective function we want to optimize is as follows. Let us assume that $y(t)$ and $\hat{y}(t)$ are the refugee estimates by the model and reported border crossing in the observed data, respectively, at time t. Let $T_S = \{t_1, t_2, ...t_S\}$ be a sample of times which we will use for calibration. Thus, we want to minimize the mean squared error between corresponding model estimates and the observed border crossings along these time samples.

Results

Performance Study - Ukraine

We calibrate and apply our agent-based framework for the recent Russian invasion of Ukraine. Our simulation runs from February 24, 2022 upto May 15, 2022. The result for



(a) ABM (b) Pandey et al. (c) Summary Comparison (d) Burundi Study

Figure 1: Refugee Estimation

Method	MAPE	RMSE	PCC
ABM	$20.72\% \pm 3.22\%$	$2.6e4 \pm 1.3e3$	0.91 ± 0.01
Pandey	$74.95\% \pm 24.27\%$	$1.06e5 \pm 5.17e4$	0.87 ± 0.01

Table 3: Metric Comparison

R	Top 5 (IOM)	ABM		(Pandey et al.)	
		Top 5	Match	Top 5	Match
1	Kyiv City	**Kyiv City**	80%	**Kyiv City**	60%
	Kharkiv	**Donetsk**		Luhansk	
	Kyiv	**Kharkiv**		**Kyiv Region**	
	Dontesk	Luhansk		**Kharkiv**	
	Zaporizhzhia	**Kyiv**		Odesa	
2	Kyiv City	**Kyiv City**	80%	**Kyiv City**	80%
	Kharkiv	**Dontesk**		**Kharkiv**	
	Kyiv	Luhansk		**Dontesk**	
	Dontesk	**Kharkiv**		**Kyiv**	
	Chernihiv	**Chenihiv**		Luhansk	
3	Kharkiv	**Dontesk**	100%	**Kyiv City**	60%
	Kyiv City	**Kyiv City**		**Donetsk**	
	Kyiv	**Kharkiv**		**Kharkiv**	
	Dontesk	**Luhansk**		Dnipropetrovsk	
	Luhansk	**Kyiv**		Zaporizhzhia	

Table 4: Comparison of IDP estimation

each day is obtained after two timesteps of the simulation, as outlined in the previous section. All pairs of households under the same S2 cell[2] are considered neighbors and we choose level 13 in our implementation (Mehrab et al. 2022). We select 10 data points as sample times for calibration, which is less than 60% of the entire observed data.

We compare our model[3] with the method proposed by *Pandey et al.* (Pandey et al. 2023) by accommodating the agent-event interaction model proposed in their work as the decision rules of our agents. We follow the parameter posteriors reported by the paper as best as possible except for two: the kernel dispersion parameter and the conflict time window; for which we did not find the final posterior of the parameters. We performed grid search to find the best combination of these two parameters.

Refugee Estimation: Figure 1 draws comparison between two methods. First, by visually comparing Figure 1a and Figure 1b, we can see that our ABM framework successfully captures the large early influx of migrants and the overall trend of the observed data. Conversely, the other method overestimates early migration and then decays as the simulation progresses. From Table 3 we can also see that our method has a higher Pearson Correlation Coefficient (PCC), lower Mean squared error (MSE) and lower Mean absolute percentage error (MAPE) with the observed data, compared to the other method. Finally, Figure 1c shows how closely the model estimates the number of refugees to the reported data, both overall and at the monthly level. While the other method performs comparably in estimating the total refugee for April, its overestimation in March is quite notable. This establishes that instead of just an agent-event interaction model, a framework incorporating both agent-event interaction and agent-agent interaction is likely able to model migration dynamics more accurately.

IDP Estimation - Ukraine: Similar to daily refugee estimates, our ABM can also generate daily IDP estimates. However, unlike border crossing data, IDP reporting is not available at daily temporality. Instead, we compare our IDP estimates against the IDP estimates provided by the International Organization of Migration (IOM) (IOM 2022). These data are collected through Random-digit-dial (RDD) approach over several rounds, conducted around every 2 weeks. The first three rounds overlap with the timeframe of our simulation. These surveys provide an estimate of the top 5 oblasts of IDP origins. We compare them with the top 5 oblasts of IDP origin estimated from our ABM and the method by Pandey et al.

The first round of survey was conducted in mid-March and we can see that our top 5 IDP origins match all but one region reported by IOM (80% match). On the other hand, the other method misses two of the top 5 reported origins of IDP (60% match). The match percentage of the two methods is comparable for the second round conducted on April 1. However, our model correctly identifies a new top 5 origin of IDP from the previous round, *Chernihiv*, which the other method failed to capture. For the round conducted on April 17, our model matches all top 5 oblasts with the observed data. Moreover, it successfully identifies *Luhansk* which was not reported as a top IDP origin in the previous two rounds.

Generalizability: Refugee Estimation - Burundi

Given agent and conflict data, the design of our model allows us to employ our ABM to examine migration dynam-

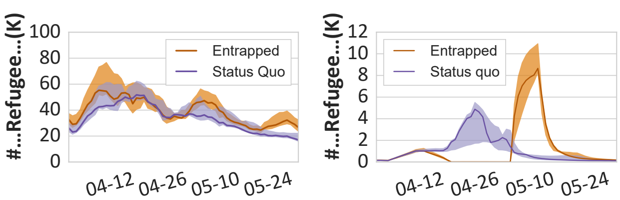
[2]https://s2geometry.io/

[3]Source Code available at https://github.com/dmehrab06/abm_generalized_migration

ics for any country. For this purpose, we tested our ABM for the 2015 Burundi conflict. We used the calibrated model of Ukraine and applied it to estimate the migrant outflow from Burundi. The result is shown in Figure 1d. The observed data is obtained from a previous study (Suleimenova et al. 2017). However, the data is associated with idiosyncrasies. This is evident from the figure that although supposed to be cumulative data, it shows a decline at some points from the previous point. Nevertheless, we find that our model generates very good estimates of the observed data, underscoring its generalizability across populations. An interesting future direction would be to study the significance and re-calibration aspects of these parameters in different contexts.

Case Study - Entrapment in Ukraine

Here, we consider the implication of imposing an "entrapment" condition on Raions behind the frontlines in Russian-controlled territory of Ukraine. Confronted with violence that would ordinarily compel civilians to flee, conditions on the ground can sever plausible escape routes leaving civilians trapped in active conflict zones. Faced with the prospect of crossing active battlefields to reach safety or sheltering in their home, civilians will rationally opt to remain stationary. During the last few weeks of our modeling period (beginning April 14) we qualitatively identified Raions within Ukraine where the advance and positioning of Russian military forces had effectively closed off westward escape routes for civilians within identified raions.



(a) Ukraine, total (b) Polohivskyi raion

Figure 2: Entrapment case study

Figure 2 compares status quo estimates with estimates obtained with entrapped Raions between late April and early May. We only focus on the simulation period from April 1 onwards for this case study because the prior outflow trend is similar. We also let the simulation continue until the end of May to see some observable differences. Figure 2a shows that around late May we would observe an obvious difference in the outflow of refugees from the scenario where the entrapment condition was enforced to the status quo condition. To investigate this further, we looked at Polohivskyi Raion (Figure 2b). This particular Raion had observed a large outflow in the status quo scenario during the entrapment period (April 14 - May 5). When we look at the scenario, we find that although people were not able to move during the entrapment period, as soon as the enforcement was relaxed, a large amount of migration happened. The surge of migration is almost twice as large as the surge observed in the status quo scenario. Due to the accumulation of fear during this entrapped period for individuals who wanted to migrate but could not, they fled at the first window of opportunity, causing this large surge of migration.

These case study results indicate the policy use of the model and indicate the sorts of valuable data the model can generate. Civilians in entrapped conflict zones in need of humanitarian aid is a largely unknown quantity during a crisis and therefore informed estimates of the sort produced here represent important data points for both policymakers and humanitarian aid agencies alike. These estimates can help to inform likely future refugee flows as escape routes from entrapped regions open when conflict conditions change. Such analysis can also help humanitarian organizations identify pockets of civilians most in need of aid and therefore can help those organizations to advocate for limited ceasefires in geographically explicit regions to provide civilian aid - a common humanitarian objective in armed conflicts.

Conclusion

We have developed a generalized ABM framework guided by social theory to model conflict-induced migration from a host country. Our model has shown promise in estimating migration at high temporal and spatial resolution for different conflict settings. In future, we can extend this framework to account for other conflict factors (e.g. infrastructure damage, perpetrator of conflict events) or deploy the model in new contexts such as natural disasters (e.g., point process events such as earthquakes or landslides), climate-induced migration (e.g., scenario analysis of migration in response to sea-level rise), or even economic modeling (e.g., migration in response to business failures). The model also has several public policy applications. For example, our entrapment analysis highlights regions most in need of humanitarian aid and therefore would help aid organizations to better identify optimal locations for resource staging in anticipation of future refugee flows as conflict conditions evolve. Additionally, incorporating return migration would strengthen the model. By considering returns, the model can run to longer time horizons allowing for better civilian displacement mapping. Finally, we only consider agent age and gender demographic characteristics. The availability of other demographic features (e.g. income) would allow further refinements to modeling conflict-induced migration.

Acknowledgements

This work was supported in part by NSF grant #2053013 Focused CoPe: Building Capacity for Adaptation in Rural Coastal Communities, DTRA contract HDTRA1-19-D-0007, NSF grant OAC-1916805 CINES: A Scalable Cyberinfrastructure for Sustained Innovation in Network Engineering and Science, NSF Expeditions in Computing grant CCF-1918656, University of Virginia Strategic Investment Fund award number SIF160, and the Strategic Investment Fund Award to the University of Virginia's Humanitarian Collaborative.

References

Adhikari, P. 2013. Sampling Hard to Reach Populations: A Survey on Forced Migration During Maoist Insurgency in Nepal. *Survey Practice*, 6(3).

Adiga, A.; et al. 2019. Graphical dynamical systems and their applications to bio-social systems. *International Journal of Advances in Engineering Sciences and Applied Mathematics*, 11: 153–171.

Ajzen, I. 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2).

Alam, M.; et al. 2015. Sensitivity analysis of an ENteric Immunity SImulator (ENISI)-based model of immune responses to Helicobacter pylori infection. *PLoS One*, 10(9).

Alis, C.; et al. 2021. Generalized radiation model for human migration. *Scientific reports*, 11.

Asgary, A.; Solis, A.; et al. 2016. An agent-based modeling and simulation tool for estimation of forced population displacement flows in Iraq. In *Proc. of DHSS*, 75–81.

Bijak, J.; Higham, P. A.; et al. 2020. Modelling migration: Decisions, processes and outcomes. In *Proc. of WSC*, 2613–2624. IEEE.

Brunarska, Z.; and Ivlevs, A. 2022. Family Influences on Migration Intentions: The Role of Past Experience of Involuntary Immobility. *Sociology*, 00380385221136060.

Carammia, M.; Iacus, S. M.; and Wilkin, T. 2022. Forecasting asylum-related migration flows with machine learning and data at scale. *Scientific Reports*, 12(1): 1457.

Castelli, F. 2018. Drivers of migration: why do people move? *Journal of travel medicine*, 25(1): tay040.

Clemens, M.; Huang, C.; Graham, J.; and Gough, K. 2018. *Migration is what you make it: seven policy decisions that turned challenges into opportunities*. Center for Global Development Washington, DC.

Davis, K. F.; Bhattachan, A.; et al. 2018. A universal model for predicting human migration under climate change: examining future sea level rise in Bangladesh. *Environmental Research Letters*, 13(6).

Frydenlund, E.; and De Kock, C. 2020. Agent-based modeling within forced migration research: A review and critique. *Refugee Review*, 4(1): 53–68.

Hassani-Mahmooei, B.; and Parris, B. W. 2012. Climate change and internal migration patterns in Bangladesh: an agent-based model. *Environment and Development Economics*, 17.

HUMDATA. 2022. The Humanitarian Data Exchange . https://data.humdata.org/. [Online; accessed December 2, 2022].

IOM. 2022. Ukraine Internal Displacement Report-General Population Survey Round 2 (24 March - 1 April 2022). https://rb.gy/po93t. [Online; accessed May 30, 2023].

Kniveton, D. R.; Smith, C. D.; and Black, R. 2012. Emerging migration flows in a changing climate in dryland Africa. *Nature Climate Change*, 2(6): 444–447.

Kuhlman, C.; Kumar, V.; et al. 2011. A bi-threshold model of complex contagion and its application to the spread of smoking behavior. In *Proc. of SNA-KDD*.

Mehrab, Z.; et al. 2022. Data-driven real-time strategic placement of mobile vaccine distribution sites. In *Proc. of AAAI*, volume 36, 12573–12579.

Mockus, J. 1989. *The Bayesian approach to local optimization*. Springer.

Mortveit, H. S.; et al. 2020. Synthetic Populations and Interaction Networks for the U.S. Technical report, University of Virginia. NSSAC Technical Report: #2019-025.

Pandey, A.; et al. 2023. Disease burden among Ukrainians forcibly displaced by the 2022 Russian invasion. *PNAS*.

Qiu, Z.; et al. 2022. Understanding the coevolution of mask wearing and epidemics: A network perspective. *PNAS*.

Raleigh, C.; Linke, R.; Hegre, H.; and Karlsen, J. 2010. Introducing ACLED: An armed conflict location and event dataset. *Journal of peace research*, 47(5): 651–660.

Robinson, C.; and Dilkina, B. 2018. A machine learning approach to modeling human migration. In *Proc. of COMPASS*, 1–8.

Searle, C.; and van Vuuren, J. H. 2021. Modelling forced migration: A framework for conflict-induced forced migration modelling according to an agent-based approach. *Computers, Environment and Urban Systems*, 85: 101568.

Simini, F.; et al. 2012. A universal model for mobility and migration patterns. *Nature*, 484(7392): 96–100.

Smith, C. D. 2014. Modelling migration futures: development and testing of the Rainfalls Agent-Based Migration Model–Tanzania. *Climate and Development*, 6(1).

Suhrke, A.; et al. 2000. The Kosovo refugee crisis: An independent evaluation of UNHCR's emergency preparedness and response. Technical report, United Nations.

Suleimenova, D.; et al. 2017. A generalized simulation development approach for predicting refugee destinations. *Scientific reports*, 7(1): 13377.

UNOCHA. 2023. Ukraine Humanitarian Response 2023. https://reports.unocha.org/en/country/ukraine/. [Accessed May 29, 2023].

Weidmann, N. B.; and Zürcher, C. 2013. How wartime violence affects social cohesion: the Spatial–Temporal gravity model. *Civil Wars*, 15(1): 1–18.

Yang, Y.; et al. 2018. Urban dreams of migrants: A case study of migrant integration in Shanghai. In *Proc. of AAAI*.

Zipf, G. K. 1946. The P 1 P 2/D hypothesis: on the intercity movement of persons. *American sociological review*, 11(6): 677–686.