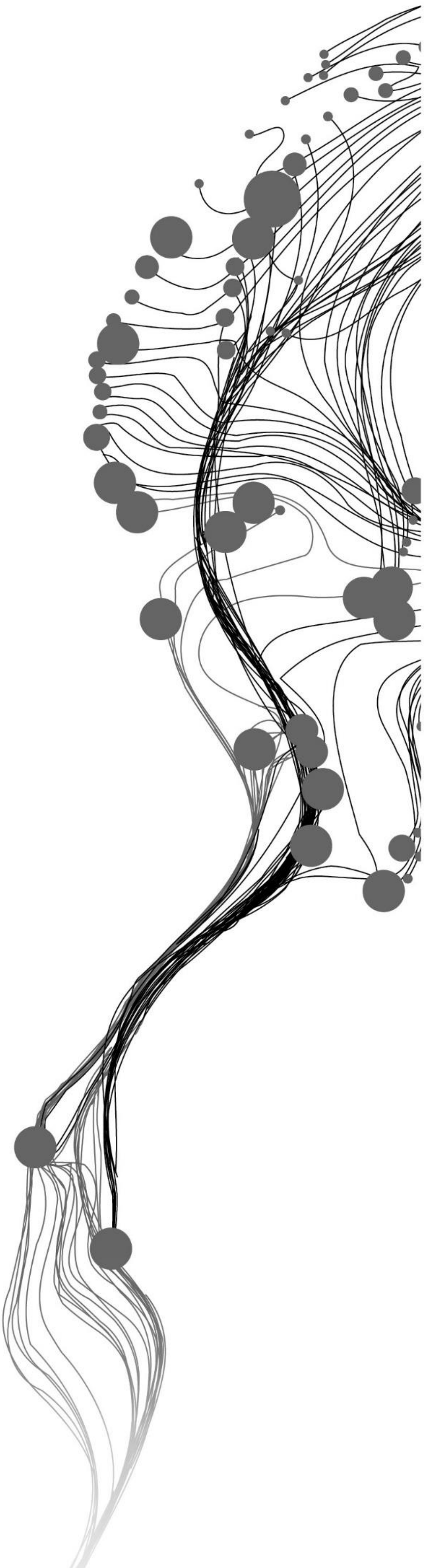
People Displacement in a Conflict Zone – A Case Study of Mosul Battle Mosul, Iraq

SARRA FEKIH
June, 2020

SUPERVISORS:
Dr. Ir. P.W.M. Augustijn
Dr. S. Amer



People Displacement in a Conflict Zone – A Case Study of Mosul Battle, Iraq

SARRA FEKIH
Enschede, The Netherlands, June, 2020

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Dr. ir. P.W.M. Augustijn
Dr. S. Amer

THESIS ASSESSMENT BOARD:
Prof.dr. M.J. Kraak (Chair)
Prof.dr. A. Voinov, (External Examiner, Department of Water Management, Faculty of Engineering Technology (FET))

ABSTRACT

The battle of Mosul was a major military liberation intervention that lasted nine months. Military operations by airstrike caused the loss of thousands of lives and generated, between 2016 and 2017, 83 % of the total worldwide population displacement. This thesis uses the intervention as a case study to develop an ABM simulation that predicts the flow of IDPs and Returnees.

Analysis driven from empirical data identified patterns of displacement. Import factors in displacement are risk perception of the household, and coping appraisal as defined in Protection Motivation Theory. Households need to decide that they are at risk before they leave their homes. Coping appraisal has to do with the location the household will move to. A literature review based on Protection Motivation Theory identified the number of deaths as a factor for perceived risk. It was observed that the number of deaths in Mosul was not the only factor that influences risk perception. Also, deaths in the households social network influence risk perception. On the other hand, the income level of a household and their social network (place to relocate to) were identified as having an impact on their coping strategy. A conceptual model was then designed using the ODD protocol.

After the implementation of the model, a sensitivity analysis was performed to identify the parameters that impacts the model output. The finding identified the number of deaths in social network of the households as a parameter that has significant impact on the perceived risk.

Model validation was performed using Pattern Oriented Modelling. The model output has predicted a lower percentage of IDPs than observed in the empirical data, this can be explained by using only the number of death troll as indicateur for risk perception. When comparing the origin of the IDPs, a high number of IDPs are generated in the west of Mosul. This may be due to the fact that more deaths occur in the west, and agents perceive risk based on the number of deaths in their sub-area.

When evaluating the location IDPs move to, the model evacuates a lower percentage of displaced households to the Provinces, compared to moving them to camps in the direct vicinity of Mosul. This can be explained by the fact that Netlogo takes into account the visualized features. The returnee sub-model has generated returnees that came back to Mosul in a later phase compared to reality. This indicates that deaths in the Mosul area are not the only factor influencing IDP behavior. By explicitly modeling the settlements around Modul, and giving them their own death rates, households from these areas can already return while the fighting in the city center of Mosul still continues.

ACKNOWLEDGEMENTS

As beautifully formulated by Umberto Ecco a thesis models a kind of self-actualization , a belief in the integrity of one's own voice.

I am indebted to the golden team : My supervisors !! for giving me the opportunity to work on a challenging, intellectually stimulating and exciting Msc research topic.

Thanks to Dr. E. W Augustijn for her support along this thesis process and this strange timing, she has been always available, giving feedbacks, reformulating , my questions (even those who hardly make sense sometimes) and stimulates my critical think on the Model.

Thanks to Dr. S. Amer who enriched our meetings by emphasising on the academic side of the research.

In a special way I would like to thank my friend Marki, who was always available in advising me on how to debugg the code .

No amount of thanks and gratitude would be enough to express the support that my parents gave me . The virtues that I claim today are a result of your investment in my education.

Thanks to my friend Crystele who advised me in pursuing this master degree at ITC , your encounter has been the golden buddha story on its own. Thanks to Marie Laure for la crème de Marron! and our "Apero-Zoom" . Thanks to Natsha for all your Post Cards and to my dearest friend Asma for your support.

My gratitude extend to the people that I met at De Tor and within the ukulele community at Enschede, who opened for me a window to the Dutch culture , especially to Jaap my tango partner , and Peter & Leanne for the Jazz on Fridays

Furthermore , I would like to thanks my friends at ITC , Sophia , Camilla , Myriam , Rachel, Alejandra and every UPM students form 2018 -2020 batch, for all the coffee breaks, sleepless nights and discussions that we had.

Special thanks to Hester to the kindness , love, and warmness of the heart , for supporting all my moods these last two months.

I would like to dedicate this work to the beautiful soul of Oscar , I hope you are happy where ever you are !

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Wars and armed conflicts are threatening for populations; they generates human and material losses and are major causes for people displacement. The term “humanitarian intervention” was commonly used in the 90s to describe the use of military force by states or international coalition in response to genocide or ethnic cleansing. (Hoag, 2008) . Although the intervenors’ purpose is to rescue and save people lives from persecutors, the intervention itself poses a moral question as it also causes human causalities and suffering, and the threshold of justifying it is blurry (ICISS, 2001). It is crucial to understand the elements that characterize a military intervention process. In such a situation, three elements are simultaneously making a cost-benefit decision to achieve a goal. Firstly, military intelligence is making decisions on which area to liberate and how to limit the movement of the rival?.Secondly, civilians trapped in the conflict are making decisions based on questions such as: When is it safe to leave? How to leave? And where to go? Thirdly, humanitarian organizations decide how to evacuate civilians and where to relocate them. The movement of civilians within space, in addition to the dynamic change of the battlefield, are essential to consider.

The term ”refugee” is defined as a migrant who crossed international frontiers because of well-funded threat (UNHCR, 2018). The term Internally Displaced People (IDP) is defined as people who are forcibly and involuntary displaced from their homes but remain within the boundaries of their own country (UNHCR, 2018). According to the UNHCR Global trend report, in 2020 one person is displaced every two seconds around the world (UNHCR, 2018). In 2018 the number of forcibly displaced people has reached 70.8 Million (UNHCR, 2018).

The Mosul battle is one of the recent armed conflicts that generated 83% of the total worldwide population displacement in 2016. , The city of Mosul fell under ISIS control on June 2014 shortly after the city’s economy collapsed. A rigid set of religious and civil codes were installed. In 2016 a military preparation to retake the City started, a coalition was made of Iraqi military units, Popular Militia Units and Kurdish Peshmerga troops. On October 17th of the same year, the Military intervention started in East Mosul and on January 24th, 2017, the eastern part of the city was declared liberated. On February 19 2017 a second military operation was launched on the western part of the city which ended on July 2017.

The most common reason for displacement around the world is armed conflict; the military intervention in Mosul is no exception. Other factors are drought, flooding, hunger, earthquakes or economic circumstances.

In extreme cases, climate change effects and risk of conflict coincide which expose populations to what the humanitarian sector defines as “double vulnerability” (Peters, Mayhew, Slim, Van Aalst, & Arrighi, 2019).

1.1. Problem Statement

Displacement flows are volatile and hard to anticipate. Due to their irregular movement, conflict induced displacement are even more difficult to predict. Understanding displacement flows is important to determine the number, capacity and location of refugee camps, etc.

Predictability of displacement flow depends on the quality of data, complexity of key drivers and the timeframe considered OECD/EASO (2018). Some approaches rely on quantitative data others on qualitative information; however, the goal remains to monitor information as close as possible to the event, in order to produce warnings of increasing displacement. The Displacement Tracking Matrix System developed by IOM is an example of such a monitoring system. Their project aims to estimates the flow of displacement. Preparedness in crisis situations is often linked to the correct prediction of migration flows.

The primary method of modelling displacement is via a statistical approach Sokolowski & Banks (2014). However statistical model’s output are static. In fact, when individuals have made a decision to move from their homes, other factors might change during the journey, in this case a statistical approach will not be suitable to model the reality. Therefore a dynamic modelling approach is more relevant and useful as tools to predict displacement movement and agent-based models are suitable tools.

Modeling human behavior through an Agent-Based Model (ABM) helps to simulate people’s actions and interactions. In some cases ABMs consist of rule-based agents that interact dynamically in space and time in the intention to simulate a real-life situation. ABM has a wide field of applications and is commonly used in evacuation reproduction (Heppenstall et al., 2012). Literature review agrees that models are useful to observe changes in the real world as well as to generate patterns we observe in complex systems.

Literature review has shown that migration studies have used modelling of forced displacement on different scales (Sokolowski, Banks, & Hayes, 2014 and Hébert, Perez, & Kim, 2018), however as far as this research is going, there is no record of modelling forced internal displacement during a conflict. This Msc thesis will contribute to resolving this problem and aim at predicting the flow of IDPs in a humanitarian and public health setting using the Mosul battle as case study.

1.2. Research Objective

The main objective of this study is to develop an agent-based model that can simulate the displacement of people before, during and shortly after a military intervention, with the Mosul Battle as a case study .

1.3. Research questions

In order to achieve the main objective, the research is split into three sub-objectives, and the following research questions are raised

Sub-Objective 1: Develop an ABM to simulate human behaviour based on risk perception, before, during and after a military intervention.

- What are the factors that influence the decision of a person to leave or stay before and during a military operation?
- Is there an existing (agent-based) model or theory using risk perception that can be used to develop the simulation model?
- How can the behavior of IDPs be simulated in an agent-based model?

Sub-Objective 2: Relocate returnees after the battle

- Which factors influence the return of IDPs to their home location?
- How can this inverse flow of refugees be modelled?

Sub-Objective 3: How can the model be validated

- How can Pattern-Oriented Modelling be used to validate the model?
- Which patterns of movement are present in the displacement data for the Mosul area?
- Can these patterns be re-generated using the simulation model?

1.4. Conceptual Framework

A model is a simple and schematic representation of reality. The agent-based simulation that will be developed during this research is a representation of a military intervention process. The model is therefore divided into two sub-models , the first one being risk perception model, , it represents the behavior of IDPs in the pre-conflict and conflict phase , the second one being relocation model and aim to represent the behavior of returnees in post conflict phase, The following section will discuss the conceptual framework.

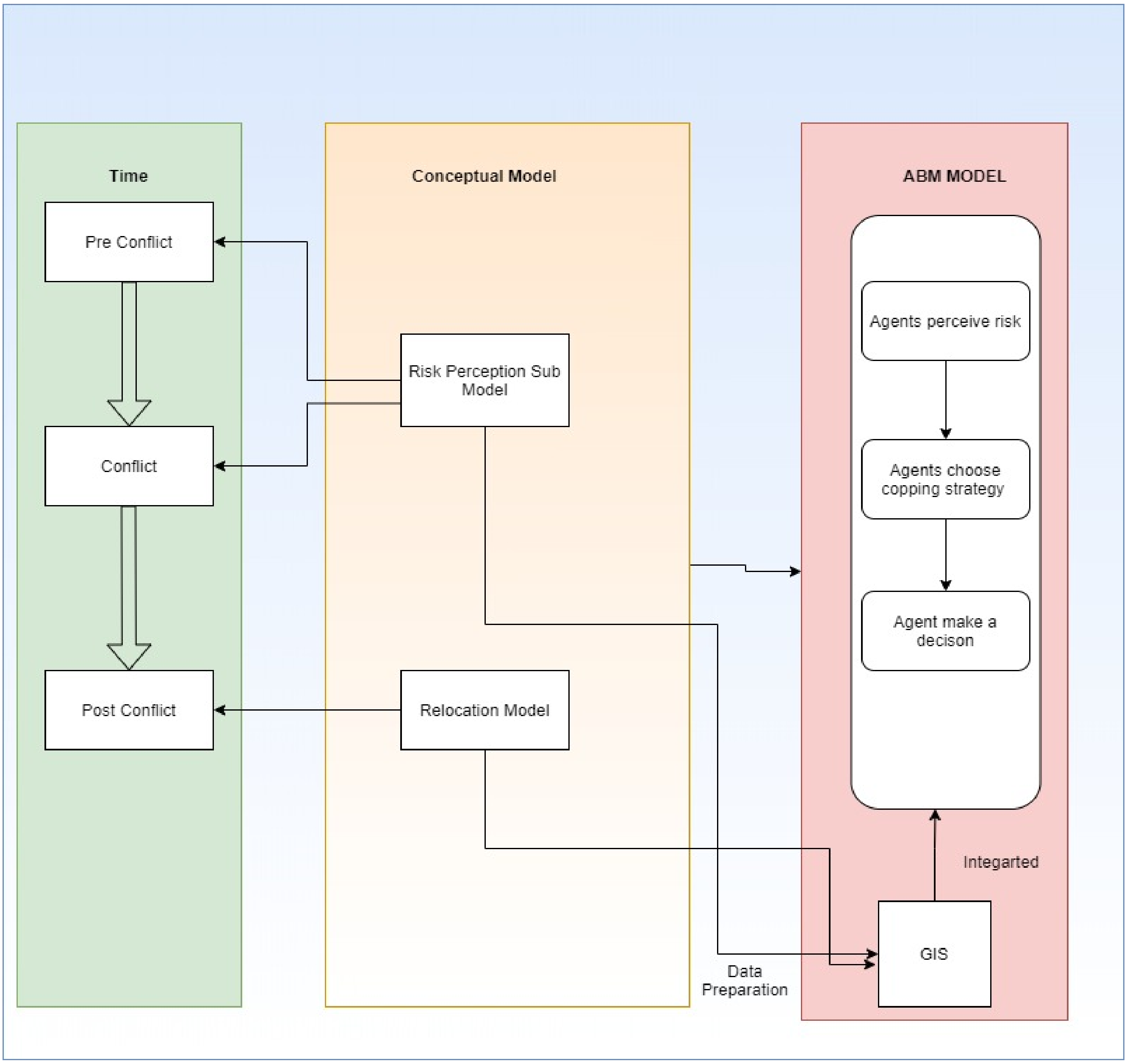


Figure 1.1 Conceptual Framework

a. Phases of the model and the corresponding flow of IDPs.

- Pre conflict phase: Simulation of IDPs' behavior on ABM, refugees will be traveling to join their friends or family, or move to refugee camps or stay in Mosul.
- Conflict phase: Refugees will be able to leave only after the area they live in is liberated.
- Post-conflict phase: Simulation of the inverse flow of IDPs, who will be returning to their homes once the war is over, and their houses are safe.

The military intervention happens during the conflict phase and can be simulated by ABM, where the agents are set up to observe their environment and make a decision based on their own judgment.

b. The decision making of IDPs will follow a set of factors that will impact the process.

According to (Hébert, Perez, & Kim, 2018) there are two important elements in the decision making process of refugees:

a. **Tolerance and the Decision to Leave**

The tolerance is an individual characteristic, and the decision to leave is based on risk perception.

b. **Destination Choice**

The destination choice can vary based on personal characteristics or on the phase of the conflict.

The decision making process will be modelled trough three distinguished phases of the conflict:

- Pre conflict phase : IDPs will be aware that a conflict is going to happen if their risk perception is high then, refugees who have a social network and financial capital will be traveling to their host relatives , Refugees who have financial capital but do not have social network will move to a refugee camps, refugees who do not have either of this options or are indifferent to the conflict will stay.
- Conflict phase: If refugees' risk perception of violence is high then refugees will decide to leave when the area is liberated.
- Post-conflict phase: Refugees will return to their homes when the fight is over, however depending on the degree of destruction of the house or the safety of the house refugees will decide either return to their homes or stay at the refugee camp.

1.5. Model components

1.5.1.1. Model overview

This study focusses on modelling decision making among IDPs during and after a military intervention. Therefore an agent-based model (ABM) that incorporates psychological aspects of agents will be developed. The methodology is sequential with respect to the research sub-objectives. The model will be developed as different sub-models. These sub-models are described below.

1.5.1.2. Evacuation sub-model:

This sub-model will first, decide on when agents become IDPs (timing of the evacuation). Second the type of shelter that agents will move to (family or camp).

Building on Abdulkareem et al., (2018) model, the first step is to look at literature review from psychology, economic and migration studies to investigate the factors that impact agents' decision-making when there is a threat. The second step is to develop a conceptual model, following the Overview, Design, Details (ODD) standard protocol for describing ABMs developed by Grimm et al. (2010). However, being a sub-model, not all the sections of ODD will be used. Therefore the focus will be on the overview part, which is divided into three-elements, 1) purpose, 2) entities, state variables and scales, and 3) process overview and scheduling.

The third step is to computerize the sub-model, via NetLogo using the network-based approach from Hébert et al. (2018) model of Syrian Migration Pathways. The evacuation sub-model will, therefore, add a decision-making process to agents in the form of an equation to evacuation decision, following different factors as variables.

Validation will be applied by comparing the outcome number of refugees to the actual number of refugees and Mosul's displacement graph (see figure 3.3) will also be compared to the outcome graph after running the model. This type of validation is also called Pattern-Oriented Modelling (POM). In order to apply this type of modelling, the empirical data about migration during the Mosul intervention has to be analyzed to determine patterns that the developed model should be able to reproduce.

1.5.1.3. Return of IDPs sub-model

The relocation sub-model will, based upon the literature, include factors that influence IDPs returning home. In addition, risk perception of IDPs will be considered.

On the other hand, the core of the model is composed of two main parts, the following section describes both elements :

1.5.1.4. Risk perception and Social networks :

Agents in an ABM can have social contacts and base their risk perception on the risk perception of their relatives and friends. An example of this is Abdulkareem et al. (2018)' s model who included social networks as a factor in risk perception. Social network can be inside Mosul such as the case of neighbors in Abdulkareem's modul, and outside Mosul, in the case of hosting families.

1.6. Research Methodology

The research methodology consists of the following steps illustrated in the following flowchart :

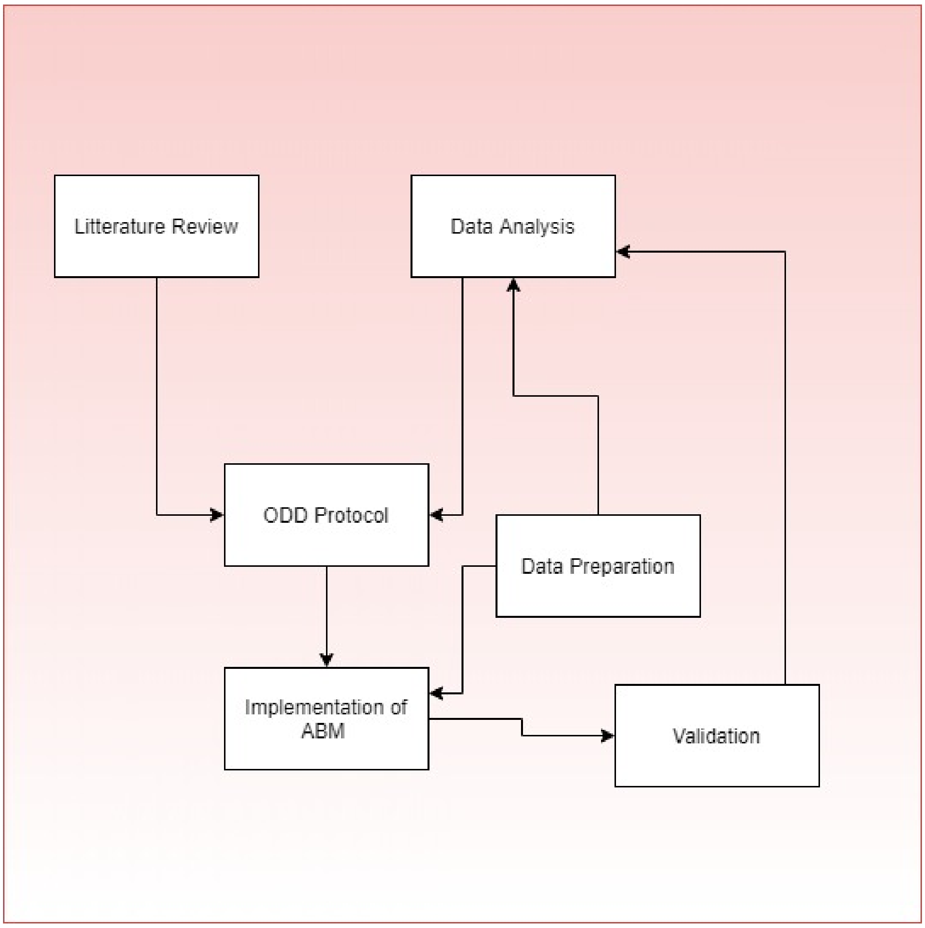


Figure 1.2 Flowchart Research Method

A. Literature review
    a. Identify other models simulating human migration (forced and not forced migration) , that are useful for the model development
    b. Find methods of risk assessment that are useful to integrate into the model
    c. Determine which phases can be identified in migration processes that should be integrated into the model or used in the analysis of the empirical data

    d. Which factors influence people displacement choice ?
B. Data Preparation : data retrieved from IOM website is preprocessed for analysis
C. Data Analysis to find important characteristics about the actual movement of refugees during the Mosul liberation
    a. Which phases are visible in the migration process
    b. Which destinations do displaced people select to migrate to during the different phases in the liberation process
D. Conceptual Modeling
    a. Which agents are needed and what should be their behavior
    b. Which environments are needed and how will they be represented.
E. Implementation and testing of the model
F. Comparison between the simulated patterns and the patterns determined in item B

2. LITERATURE REVIEW

The following chapter aim to give an overview of the theories that this study is researching. Covering multiple disciplinary. The chapter is divided into two main parts, literatures derived form social science and literature derived from computer science.

2.1. Conflict induced displacemnt

Forced migration is characterized by the movement of people from their homes to other destinations based on a life-threatening event (UNHCR, 2018). Armed conflicts are a common cause of the displacement of people. Conflicts usually generate waves of displaced people, and depending on the conflict intensity, displacement can be for a short or a longer period. This section discusses forced migration caused by conflict and the choice of destination of displaced populations.

2.1.1. Displacement during conflict

According to the International Committee of the Red Cross (ICRC), violence causes people to flee their homes; conflict duration might affect the period of displacement (ICRC, 2019). When a conflict happens in densely populated areas such as cities, a large number of people can become displaced. Waves of displacement in the countryside are generally smaller, because they are less densely populated (ICRC, 2019). From an origin perspective, Lozano-Gracia et al. (2010) in their study on conflict-induced migration, suggest that factors such as violence, absence of institutions in the case of a civil war, and dissatisfaction with basic needs are the pushing reasons for displacement. In the Democratic Republic of Congo, for instance, people flee when armed groups move into their village, civilians return as soon as the village becomes safe again. ICRC (2019) describes it as a short-term wave of displacement. Protracted displacement is generated by a cumulative effect of hostilities, that last longer in time and can generate multiple waves of displacement (ICRC, 2016). Iraq is an interesting example of short term and long-term displacement, in fact, since the 1990's the northern part of Iraq has experienced multiple waves of displacement. Anticipating the fighting to come, citizens moved back and forth, spending the day in their business and leaving at night to their place of displacement. Yet, the displacement during the 1990's was short term, as citizens went back to their homes when the situation was cleared. The 2003 invasion also generated short term displacement as the fighting had less direct impact on civilians (Chatelard, 2008). Early 2014, a slow wave of displacement occurred, but a more significant movement of displacement took place in the second half of 2014 at the time when ISIS seized major parts of northern Iraq. Later waves were observed during the military intervention (IOM, 2018).

2.1.2. The threshold approach

Crawley and Hagen-Zanker (2019) describe migrants' decision making as a process influenced by a complex interaction of macro and micro level factors that involves political, economic, social resources that migrants or refugees can assemble. Van der Velde & van Naerssen (2011) presented a theoretical framework to analyze the migrants decision-making process called the threshold approach. The authors suggest that there are a number of "thresholds" that need to be exceeded before migration will occur. The first one is the psychological barrier, for someone to migrate to another place, the idea of migration as a valid option should first take place in the individual mind. The second threshold is that the individual expects that the destination will bring improved wellbeing. Although the threshold approach studies voluntary migration, the indifference threshold can be reached in a violent conflict where chances of livelihood are becoming low. Furthermore, van der Velde et al. (2011) argue that two more thresholds need to be crossed for the migration to happen: a locational threshold and a trajectory threshold. The locational threshold is related to where to go as a destination. The trajectory threshold is regarding the journey to take, such as the distance to reach the destination, the cost of the journey and how dangerous it can get. Factors such as financial capital can explain the trajectory thresholds, where higher means can offer a safer journey. Social capital or social network explains the locational threshold, in fact, migrants with friends and family abroad will seek shelter with them. (Mallett & Hagen-Zanker, 2018).

2.1.3. Choice of destination

Migration research often tries to answer the question where people go and why people choose a specific location. Understanding the patterns of destination and the factors that trigger the choice of location in conflict-induced displacement is relatively limited (Lischer, 2007a). Factors that influence whether people flee or remain are further discussed in section (2.1.2). Ibáñez and Moya (2006) suggest that support from relatives and friends is the main reason for the selection of destination. Steele (2009) argues that a collectively victimized population has three options regarding destinations. The first is moving to a stronghold rival group to seek protection from the enemy. The second is clustering with a similar ethnic or religious group. The third is moving to a broader community and seek anonymity. Mekdjian (2018) argues that cities are the first refuge for the migrating population. From a destination perspective, Lozano-Gracia et al. (2010) suggest that more populated regions are attractive as well as areas that have a sufficient level of basic needs. Similarly, cities with a larger population, more extensive social networks, and closer to the place of origin attract more displaced people.

Repellent Factors	Attractive Factors
Violence	Populated region
Absence of institution	Social Network
No Basic needs provided	Sufficient level of basic needs
	Close to place of origin

Figure 2.1 Factors for choice of destination based on Lozano Garcia et al. (2010)

2.2. Protection Motivation Theory

Human behavior is not easy to predict. It is the response of each individual to internal and external stimuli. Behavioral studies (Ajzen, 2005) identify three human behavior laws. First, behavior follows the status quo meaning that people prefer that things remain the same. Second, behavior is a function of the person and his/her environment. Third, there is a tradeoff for every decision made. Protection Motivation Theory (PMT), a theory that emerged in the seventies, studying the motivation behind people taking action by awaking fear. Rogers (1983) indicates that threat appraisal and coping appraisal are two stages of PMT. The threat appraisal process considers the degree of harm from unsound behavior as well as the probability of experiencing damage. The coping appraisal process examines the effectiveness of favored actions in preventing potential damage as well as having the confidence of achieving it. This research applies the PMT framework to research human behavior in a conflict induced displacement and this section will look at factors that trigger threat and copping appraisal.

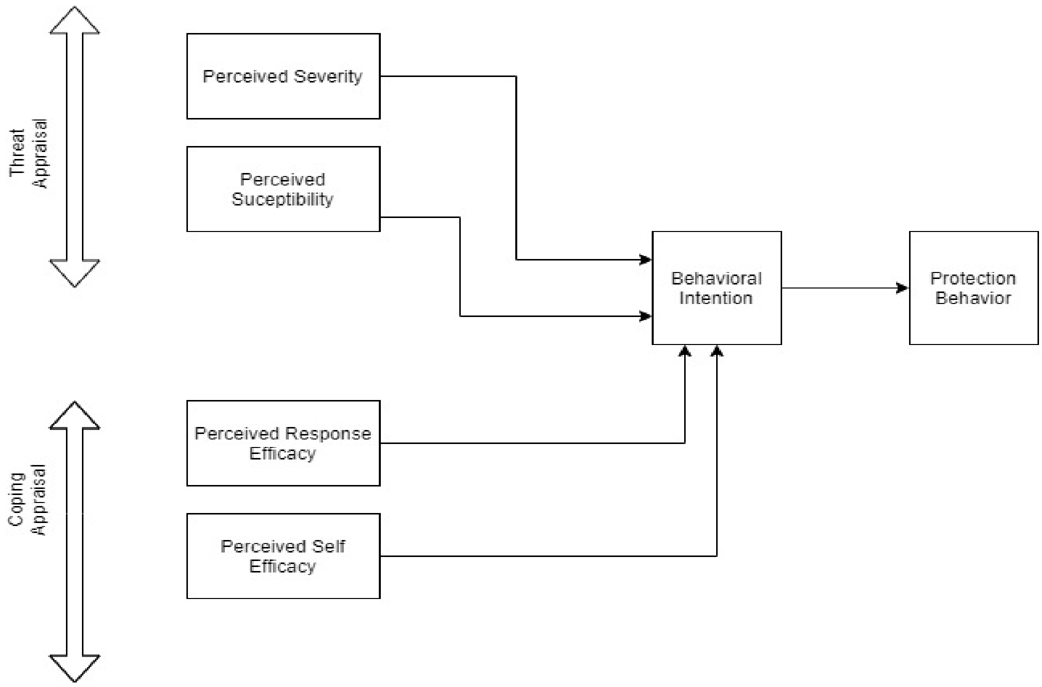


Figure 2.2 Protection Motivation Theory (adapted from a revised theory of protection motivation, Rogers 1983)

PTM has been applied in multiple scientific disciplines, in health studies for instance, threat appraisal is motivated by perceived severity, or negative consequences an individual associates with a health event or outcome. Consequences may be related to an anticipated event happening in the future or a current state (Miles, 2008.) along with perceived susceptibility which refers to a subjective assessment of the risk of developing a health problem. On the other hand, coping appraisal is motivated by perceived response efficacy which is an individual's belief that a certain action will be effective and by perceived self-efficacy an individual beliefs about its capabilities to produce effects.

2.2.1. Risk Perception

The threat perceived by a population regarding the proximity of the battle triggers more potential forced migrants than the intensity of the fighting. Davenport et al. (2003) argue that threat to personal integrity is the main reason that drives people to flee their homes. Furthermore, their research findings suggest that people will tend to go to places where they expect conditions to be better

Lischer (2007) identified armed conflict, genocide, war, etc. as the main determinant of forced migration. Melander and Öberg (2007) suggest that there are three different sources of threat. First, state violence such as repression and genocide. Second, dissident violence such as riots and guerrilla attacks. Finally, state-dissident interaction such as armed conflict. Melander and Öberg (2007) conclude that the geographical scope of fighting reaching urban centers increases the number of displaced people.

a) **Actual risk and Predicted risk**

Risk prediction can be optimistic or pessimistic. Literature suggests that risk perception is influenced by what information is available to the individual. For example, people will perceive a risk for a disease as high when a family member has been affected (Ferrer & Klein, 2015). The ability for humans to predict what will happen in their environment is an essential part of human cognition. Individuals are constantly anticipating what they will hear, see and feel. Evidence suggests that prediction is one of the primary functions of the brain and that people tend to underestimate the risk they are willing to take but overestimate risks in situations that they are not able to control (Bubic et al, 2010).

b) **Exposure to risk and risk perception**

In the 1960s, geographers started to relate risk perception and geography by carrying out surveys on the public's risk perception of flood and hazards (Burton & Kates, 1964). Soon after that, cognitive scientists carried surveys on people's risk perception regarding technological hazard (Slovic, 1996). In their study on the effect of proximity on risk perception, Maderthaner et al. (1978) argue that frequent contact with a potential threatening object can lower the risk perception. This confirms the hypothesis of mere-exposure that explains why it is easy for people to change their attitude and believes about remaining in a potentially risky location rather than moving to another place. On the other hand, Fischhoff et al. (2003) revealed that

American citizens felt a greater risk from terror when they live within 100 miles of the World Trade Center over those who live further away.

c) **Social Influence on Risk Perception**

Risks exposure leads most people into a complex process of decision making and communication to others of their own actions. Social influence may interfere with personal experience making coping strategies to risk an important factor within a group interaction. Abdulkareem et al. (2020) argue that selection of coping strategies by groups outperforms individual decisions in the case of reducing the incidence of a disease.

d) **Economic approach to risk**

When individuals are exposed to uncertainty, economists distinguish between two possible behaviors: risk aversion and risk-taking. Risk aversion is the behavior of humans who are willing to accept lower outcomes that are more secure, in contrast to risk-taking, which is characterized by individuals who are eager to take more risk to get a better outcome. Ceriani and Verme (2018) argue that when living under conflict, making a choice to leave or to stay is compared to a lottery. Staying means taking the risk of being harmed or killed. However, the probability of remaining in control of personal assets is higher. Leaving, on the other hand, implies a lower chance of being harmed or killed but, individuals will have lesser control over personal assets. In his research about factors that lead civilians to flee conflicts, a threat to personal integrity is the primary cause for people to flee their homes (Davenport et al. 2003). Adhikari (2013) argues that there is a misleading association between conflict and displacement; his study revealed that not only the role of violence but also the importance of economic infrastructure influences civilians' decision. For example, persons working in a factory that is still functioning in a conflict situation, are more likely to stay than people who have lost their agriculture land.

2.3. Models

Agent-Based Modelling (ABM) is a class of computational modelling to simulate the actions and interaction of agents within a system and better assess their effects on the system as a whole. In standard agent-based models, agents are programmed to obey a fixed set of conditions and rules (Bandini, 2009). ABM has a wide field of applications and is commonly used in, for example, evacuation modelling, and ecological modelling etc. (Heppenstall, Crooks, See, & Batty, 2012). This section focuses on human behavior modelling, agent interaction, decision making related to migration models, and examples of Agent-Based Modelling in migration studies.

2.3.1. Modelling Human Behavior with ABM

Modeling human behavior through an ABM helps to stimulate people's actions and interactions. ABM consists of rule-based agents that interact dynamically in space and time in the intention to simulate a real-life situation. Abdulkareem et al. (2018) propose to integrate the psychological aspect of decision making in a risky situation into a spatial ABM to allow the population of agents to perceive risk and share information through a social network which makes them undertake more effective choices. Agents interact and learn from each other; the behavior of some agents influences the decision making of other agents. Transmission of information within a social network is based on an individual's experience that is relayed in the form of anecdotal information. Nowak et al., (2017) argue that the more an experience is sensational, the more people will be informed. Abdulkareem et al. (2017) use social interaction in their model to inform the agents around a water collection point; each agent has a social interaction with seven other agents.

2.3.2. Decision Making in ABM migration Models

Decision making in ABM migration models is dominated by the study of economic migration. Klabunde and Willekens (2016) investigated agent-based models of human migration based on their decision-making. Their study identifies random utility theory as a prominent theory for explaining human behavior. Random utility theory assumes that an individual will evaluate all alternatives in a given choice set and will select the alternative that yields maximum utility. Klabunde and Willekens (2016) see the modelling of the decision process and social networks as critical in developing an ABM Model. Kniveton et al. (2011) developed a model to study environmental induced migration. According to their findings, an individual's attitude towards movement is influenced by the probability of migration of similar individuals (including age, gender, and wealth). Kniveton et al. (2012) also suggest that individuals can influence each other regarding the decision to migrate.

2.3.3. Modelling conflict-induced refugee behavior

The literature review conducted has not been able to identify specific research publications that discuss ABM modelling approaches for Internally Displaced People, which is an indication that forced migration studies using ABM are still in an early stage. Refugee models are dominated by statistical approaches (Pellegrini & Fotheringham, 2002; Fitzgerald & Arar, 2018) mostly based on socio-economic factors (e.g. violence and income level). Existing ABM models study the behavior of refugees on a larger geographic scale (Sokolowski et al., 2015; Herbet et al., 2018). One particular model explains refugee behavior and simulates migration pathways of Syrian refugees (Herbet et al., 2018). The model uses the death toll as an indicator of conflict escalation in different places and over time. Agents represent the human population, at a ratio of 1 agent per 1000 people. Decision making is based on the agent's perception of conflict escalation. The authors used the number of deaths to measure the severity of a conflict escalation, which in turn makes that agents either stay or leave. Once the decision of leaving is made, agents consider a variety of factors such as their wealth, destination capacity as well as means to travel. The choice of destination takes into account the wealth of each migrant. At the destination, agents decide whether to seek refuge based on capacity at the destination. For example, refugee camps being full or not. The model uses a time scale of 1 month for a period of 4 years. The model uses detailed series of indicators, such as ethnicity, religion, and gender. Indicators of destination attractiveness include infrastructure, airports, the efficiency of government, security, etc.

2.3.4. Models Validation

Model performance only has some relevance on the real system represented if the model is comparable to the described system. Model validation is the task of demonstrating that the model is a reasonable representation of the system, by reproducing similar behavior that satisfies the analysis objectives. Pattern oriented modelling (POM) is a strategy that provides a standardized framework for decoding the internal organization of agent-based complex systems (Grimm et al., 2005). POM was developed in ecology, a science that uses bottom-up modelling. The strategy follows a basic scientific approach by explaining the observed patterns (Grimm et al., 2005). POM identifies the trends observed at multiple scales that characterize a system. Emerging patterns from a model contain the right mechanism to address the system's problem (Grimm & Railsback, 2012). These patterns are threefold. They are used to determine the needs of the model in terms of scales, entities and variables. They test and select sub-models, such as adaptative behavior. Finally, they identify useful parameters for calibration. (Grimm & Railsback, 2012).

2.4. Summary of literature findings

The literature review has been a multidisciplinary approach giving the nature of this thesis subject. The finding from migration literature has proven that in fact, violence is the reason behind peoples displacement decision. The threshold approach theorized by Van der Velde et al. (2011) suggested that the choice of destination is part of the decision making process for a refugee. Attractive destinations, on the other hand, are influenced by factors such as social network, populated regions and the closeness to the place of origin. Our literature review has also looked at behavioral studies. The conclusion drawn was that risk perception was influenced by the available information and that social interaction was an important factor of coping strategy. The economic approach has proven that control over personal assets was a factor that people consider in an armed conflict situation. Computer science literature has shown one particular ABM model that models refugee behavior during a migration process. The model uses death toll as an indicator of conflict escalation in different places and over time. Finally literature has identified Pattern Oriented Modelling as an effective validation method in Agent Based Modelling.

3. DATA PRE-PROCESSING

The following chapter aim at telling the story of the battle from data perspective, the chapter is describing the data pre-processing step , data was retrieved from the IOM Displacement Tracking Matrix. The chapter is divided into five parts from data acquisition to data visualization.

Data pre-processing chapter include the process of data acquisition, data extraction, data cleaning and data visualization, the chapter includes also data description, Analysis of the outcome is discussed is in Chapter 6. Figure 1 illustrates the steps of data pre-processing. Data was collected from the International Organization for Migration (IOM) database. Python scripts and Qgis queries were performed in order to visualize Data. The code is provided in Appendix A.

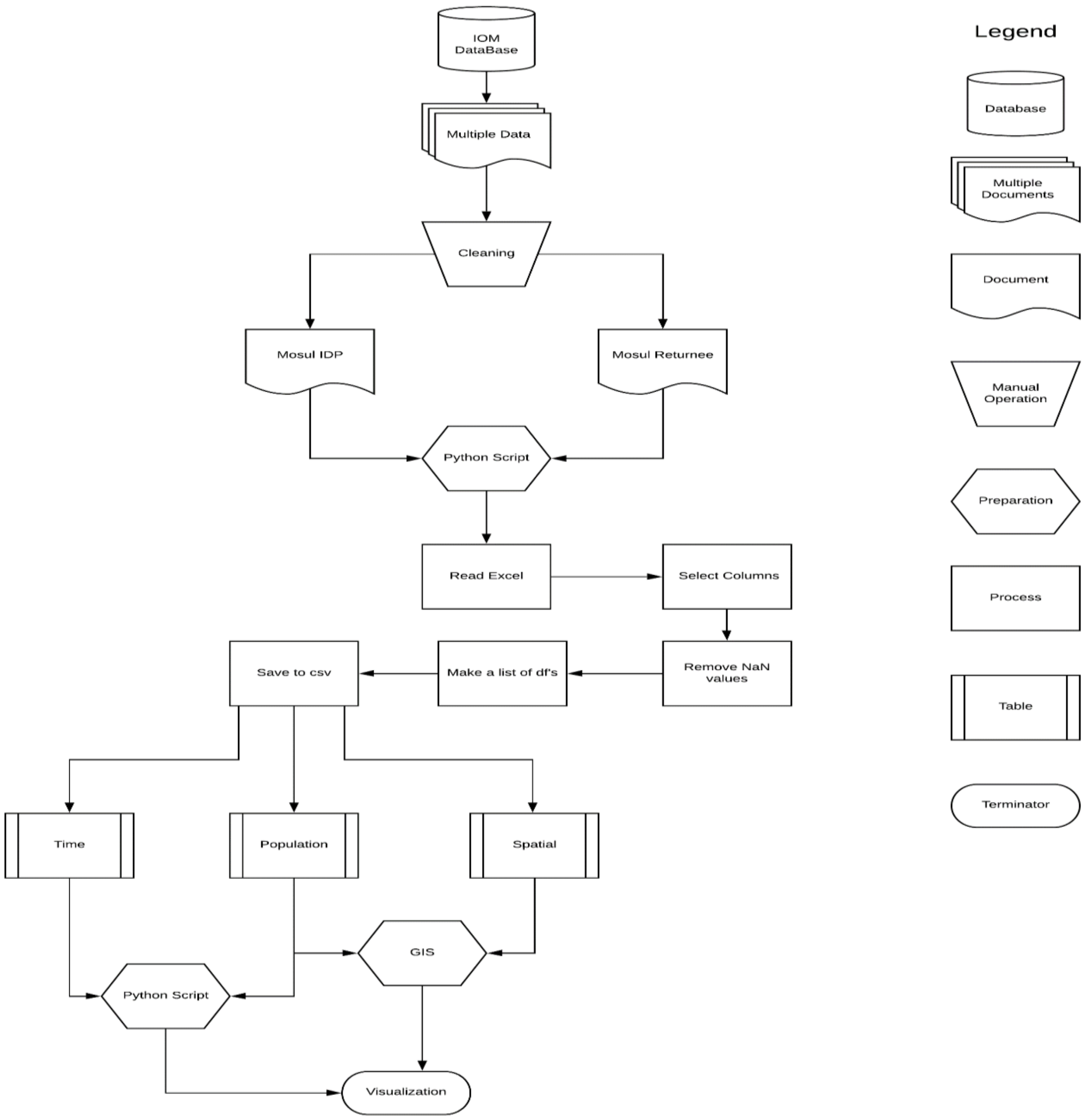


Figure 3.1 Flowchart Data Pre-processing

3.1. Data Acquisition:

Data was downloaded from the Displacement Tracking Matrix(DTM)- IOM Iraq Mission. DTM is an information management system that monitors and tracks population displacement during a crisis. DTM Framework has four main components. First, Mobility Tracking that aims at monitoring population displacement and capture mobility dynamics among IDPs and Returnees. Second, Location Assessment, which collects information on IDP and Returnees living at their identified location. Third, Safety Audit, which assesses site level-gender based violence and safety of the infrastructure. Fourth, Emergency Tracking activated during a complex crisis, it informs about the displacement of population through early emergency reports that are updated bi-weekly. The four different components use different time frames, geographic units and community of reference ("DTM-IOM-Iraq Mission,"). This thesis is based on the Mobility Tracking, and datasets are in the form of Master Lists, the DTM team uses Master Lists to produce monthly reports, figures, maps and dashboards to track better and monitor the current situation.

3.2. Data Description:

Master lists (ML) are in the form of Excel sheets, publicly available from IOM Iraq Mission Website, ML, collects information about the location and the number of IDP and Returnee at small geographic units. They are produced bi-weekly in the form of reports and dataset. The first round for IDP Master lists was launched in 2014 following the wave of displacement induced by the rise of ISIS till present. However, the Returnee Master List started tracking returnees from 2015 till today. Both Master Lists contains information about the number of households and individuals in each location (village level for rural areas and neighbourhood level for urban areas), The Master List takes into account the number of families displaced or returned to a location at the time of data collection. The ML also contains information on Governorate of Origin which are the provinces from where IDPs are initially from and areas from where Returnees are coming back.

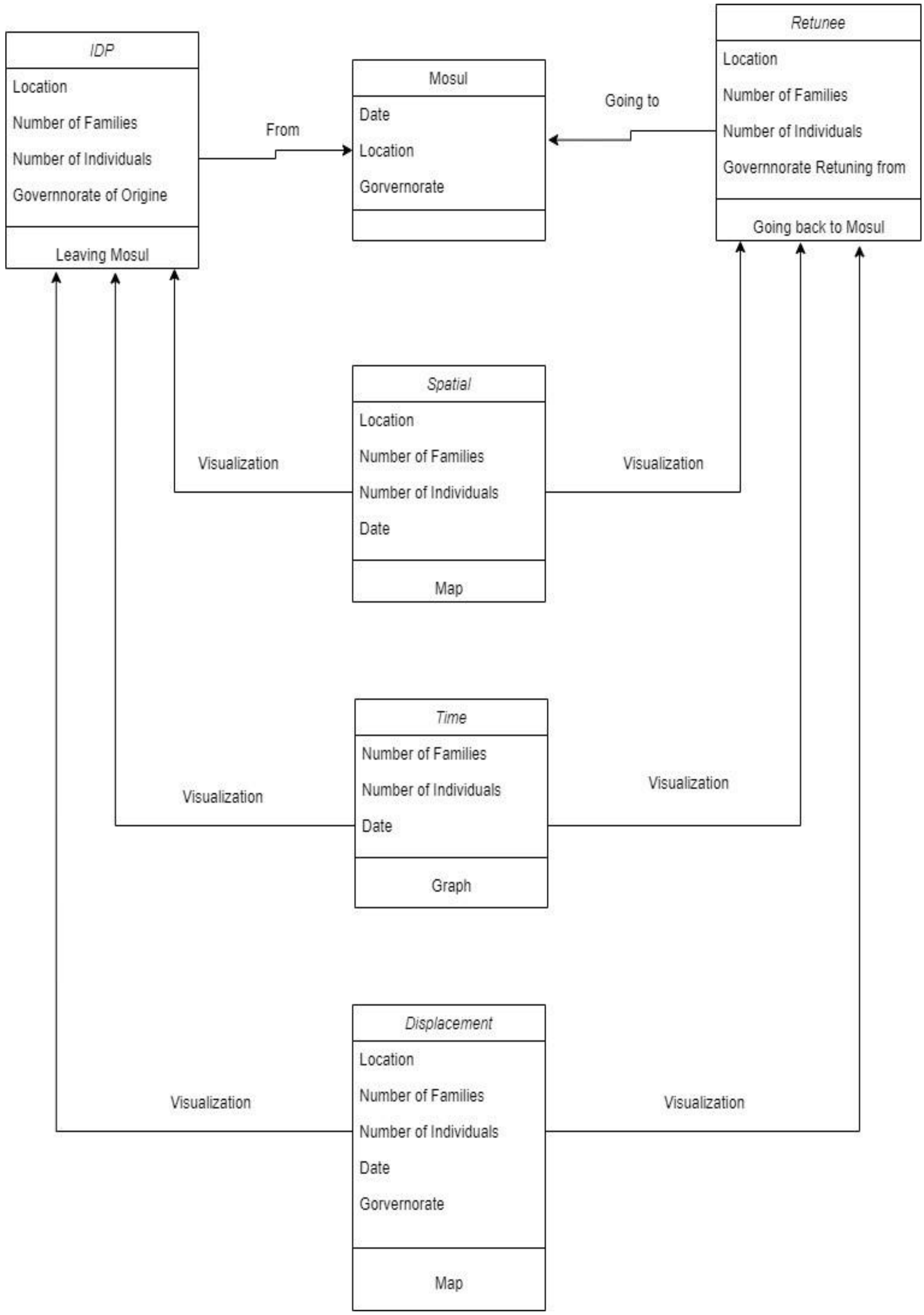


Figure 3.2 UML for Data Description

3.3. Data Cleaning

Master Lists for IDPs were uploaded in April 2014 till present. However, at the timing of the pre-processing data phase, this thesis is taking into account only lists released between April 2014 till December 2019. Same is applied for Lists of Returnees, which were, however, released a year later. The lists were not consistent in their formats due to the different versions produced during the last five years. The first step was to standardize the templates by giving headers the same typo (e.g., starting with a Capital Letter etc...). Second, Date – Month – Year, was added in each excel file to keep track of the time. Finally, excel files were stored in folders and sub-folders according to the month and year released on.

3.4. Data Extraction:

A python script see Appendix was written to iterate through folders and files to extract data related to Mosul and columns that where need for the visualization, therefore Libraries such as OS , Glob and Pandas where used. The data extraction resulted in three type of tables containing time information, spatial information , and displacement information. Each having the purpose of visualizing IDPs and Returnees.

3.5. Data Visualization

a) Time Visualization

A python script was written to plot three different graphs using Matplotlib library the first graph visualizes the number of IDPs and Returnees from April 2014 till december 2019 . The Second graph Visualizes the number of IDPs by shelter type, the graph shows camp, host families and unknown shelters for the same period of time (see figure 3. 5) . For both graphs population number was summed using a python script with Numpy libraby. The third graph visualizes the percentage of IDPs and Retunee during different displacement waves (see figure 3.6). Finally, although the graphs show the dynamic of Mosul's displacement, through time, the X axis takes into account only years, loosing key dates therefore for visualization purpose, a python script was written using Ploty library, developing interactive graphs, that keep track of the dates. The outcome is displayed in a story map.

b) Spatial Visualization:

Governorates of Origin for IDPs and Returnees were visualised on different maps eah map representing a year, and the number of IDPs and Returnees per Gouvernate. The number of population was calculated through a python script using Numpy Library and Sum Function. The maps were ploted using Qgis software.

c) Spatio-Temporal Visualization:

Location points provided within the dataset was plotted in a map at different Dates from April 2014 till December 2019, using a feature on a Qgis called Time Manager, that allowed to create an animation of points through time, giving an insight of the displacement dynamic for both IDP and Returnees. Location of camps where later on loaded , a screen picture of each map is added to this chapter.

4. METHODOLOGY

This section is divided into two main parts the first part (section 4.1) describes the Model following the ODD protocol structure. The second part (section 4.2) describes steps by step the implementation of the model.

4.1. ABM Description

The proposed model is based on a model developed by Hébert, et al. (2018) that simulates Syrian Refugee Pathway. However there are also clear differences:

Table 4.1 Comparison Hebert Model and the Mosul model

Hébert Model	**Mosul Model**
Models forced mobility to other countries	models internal displacement within Iraq
Includes country boundaries	Mosul Model do not include boundaries
One agent represents 1000 people	One agent represents one Household
Model simulates migrant	Model does not simulate the journey
Each agent has a health score that decreases through time, the loss of health score means the death of the agent.	Model does not take into account the health of agents
Destinations have an attractiveness score that include detailed factors (ethnicity, press freedom, life expectancy, quality of education)	Destination have no score
Agents status has two states: migrant or refugee Model only simulates agents leaving the area.	Agents status has three states : IDP, Stayers, and Returnee The model simulates not only the migration from the area, but also the return to the Mosul area.

To adjust for these shortcomings, the design of the adjusted model is described below, starting with the main composition of the model (1.1), the design concept (1.2), the details of the adjusted model are discussed in section 2.

4.1.1. Overview of the Model

Purpose:

The purpose of this model is to simulate phases of displacement through time during the Mosul Battle, in Iraq, from October 2016 to July 2017, the model also takes into account the few weeks before and after the military intervention. The existence of three Displacement phases will be examined:

- Pre conflict (early Leavers)
- During conflict (Later leavers)
- Post conflict (Returnees)

The model is constrained to Mosul city and its surroundings (spatial extent) from September 2016 to November 2017 (temporal extent). the prediction of the flow of IDPs in neighboring camps, will help estimate health and humanitarian need both in the short term and long term.

Agents:

The Hébert model contains one agent, representing 1000 people. The adjusted model will, on the other hand, contain one agent representing one household. We assume that the average size of a household is 7.7 based on survey's conducted by the United Nations in Iraq (United Nations, 2019). The attributes of this agent are shown in Table 2 below. The state variable of this agent is its status. This status can be Stayer, IDP or Returnee. All household agents are "stayers" at the beginning of the simulation. During the simulation, they perceive risk, if risk exceeds their risk tolerance ("risk threshold") they can change from "Stayer" to "IDP". When an agent becomes an IDP, this agent will have to choose a location outside Mosul to relocate to. All IDPs will continue to perceive risk, and when risk drops below the "risk threshold", the household agent can again change status to become a "Returnee".

Table 4.2 Agent attributes

Attribute	Description
ID	Unique identifier for each agent
Location	Longitude and Latitude of the Agent location at the neighbourhood or village level
Status	Stayer, IDP, or Returnee
Risk Threshold	The threshold above which the agents will change their status
Neighbour Network	Link with 7 other neighbouring agents
Wealth-weight	On a scale from 0 to 12 a weight for the income level. 0 being very high and 12 being very low

Social Network	a network outside the City of Mosul that represent host families.

The following properties distinguish IDPs and Returnees. For the Household agents to change their status, the agents have to be able to conduct the following behavioral elements:

- Perceive risk and compare this risk with their own Risk Threshold.
- Check if it is possible to relocate (restricted area)
- Find a location to relocate to
- Change their status

Table 4.3 Agent Behavior

Agents	Behavioral difference
IDPs	1) Leave Mosul when their risk perception is high 2) Move to camps or host families
Returnee	1) Return to Mosul when their risk perception is low 2) Move from host families or camps 3) Become IDPs if they remain in camps
Stayers	1) Stay in Mosul when factors are not met

In Hébert's model, agents decide upon creation on a destination based on their preferences with options available given their conditions. Factors such as their social network abroad, assuming that their contacts can provide shelter, influence the decision. Moreover, wealthier migrants will use better ways of transportation. If there is no destination assigned during the initialization of the model, the migrant chooses the nearest refugee camp. In the adjusted model, agents' destination is within Iraq, mostly in the northern provinces of the country where it is assumed that host families are located. Factors such as social network and income will also be applied in the destination choice. The adjusted model will not simulate the actual relocation, so no choice of transportation is needed. When factors are not met agents choose to relocate in the nearest camp.

Environments:

The environments required for this simulation include:

- Restricted area. These areas consists of a border that is dynamic and can change over time. The border can be closed or open. When the border is closed, agents will not be able to leave the area, or return back to the area, and will remain in their current position. The moment the status of the border will change is determined based on historical data of the liberation of the city. The shape of the border will be changes … times during the simulation.
- Camps outside and around the city of Mosul are represented by points. Camps have a maximum capacity indicating the maximum number of agents that can be accommodated in the camp. When a camp is full, no more IDPs will be admitted.
- Mosul is represented by polygons, determining the neighborhoods.
- Provinces of the northern part of Iraq and Baghdad are represented by points, they will be representing host families.

Scales

Table 4.4 Scale

Scale	**Description**
Temporal scale	Simulation will be run from September 1st, 2016 to November 26th 2017. Time scale will be weekly. (65 ticks).
Population scale	One point represents a households.
Spatial Extent	The study area covers Mosul city and the villages around it that are hosting camps. The map will be distorted so that and the distance to camps and cities will be reduced for visualization purposes.

Process overview and scheduling:

(1) **UML Class Diagram**

This diagram explains the relationships between the agents and the different environments. It also depicts the various attributes and behaviors of agents and the environment for each class and object.

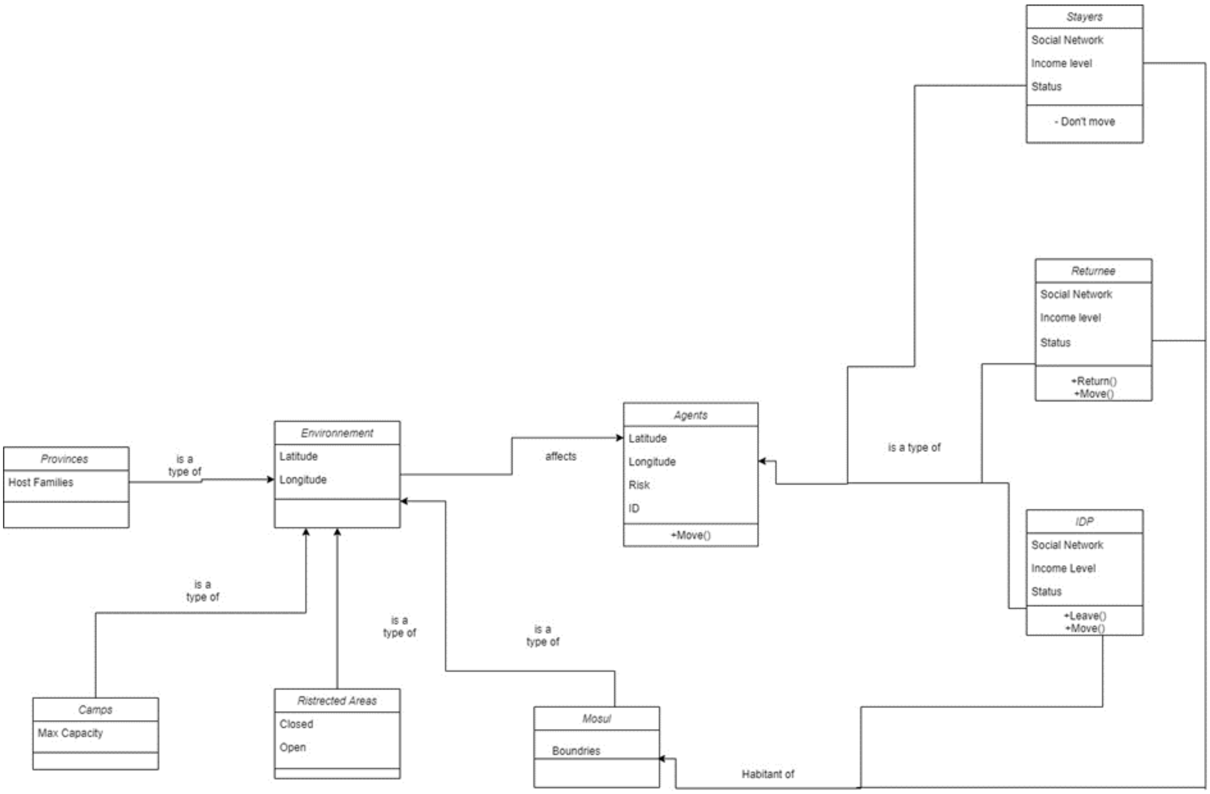


Figure 4.1 UML class diagram

(2) **UML Sequence Diagram**

This diagram explains the agents' behaviors, interactions with the environment. It is constructed in a chronological and a logical sequence of events for one simulation.

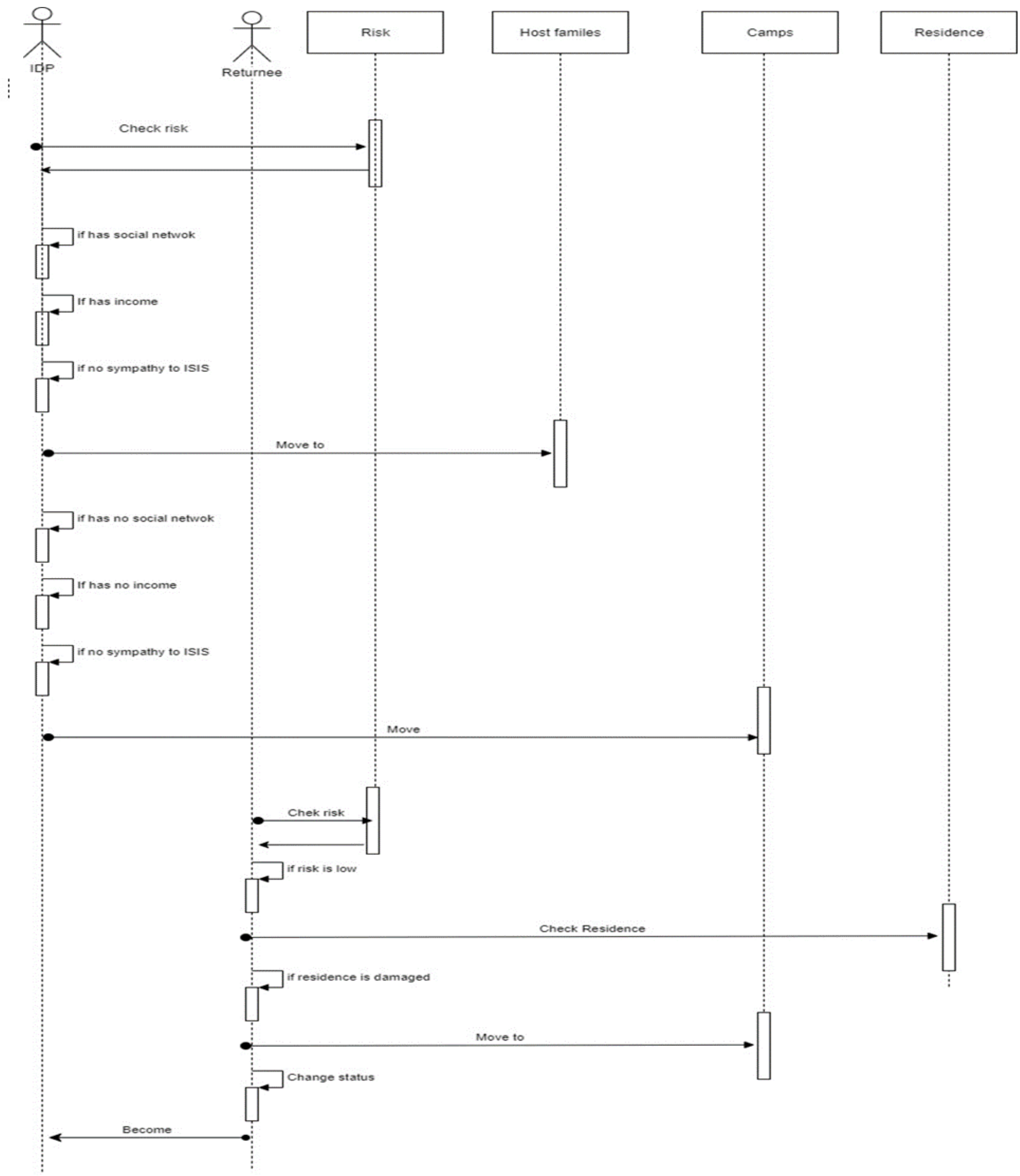


Figure 4.2 UML sequence diagram

4.1.2. Design Concepts

Basic principle

Protection motivation theory suggests that people protect themselves based on four factors: the perceived severity of a threat, the perceived probability of the occurrence, the efficacy of recommended behavior and the perceived self-efficacy. It has two principal components: Threat appraisal, it assesses how serious the situation is. On the other hand, the coping appraisal is the response to the situation.

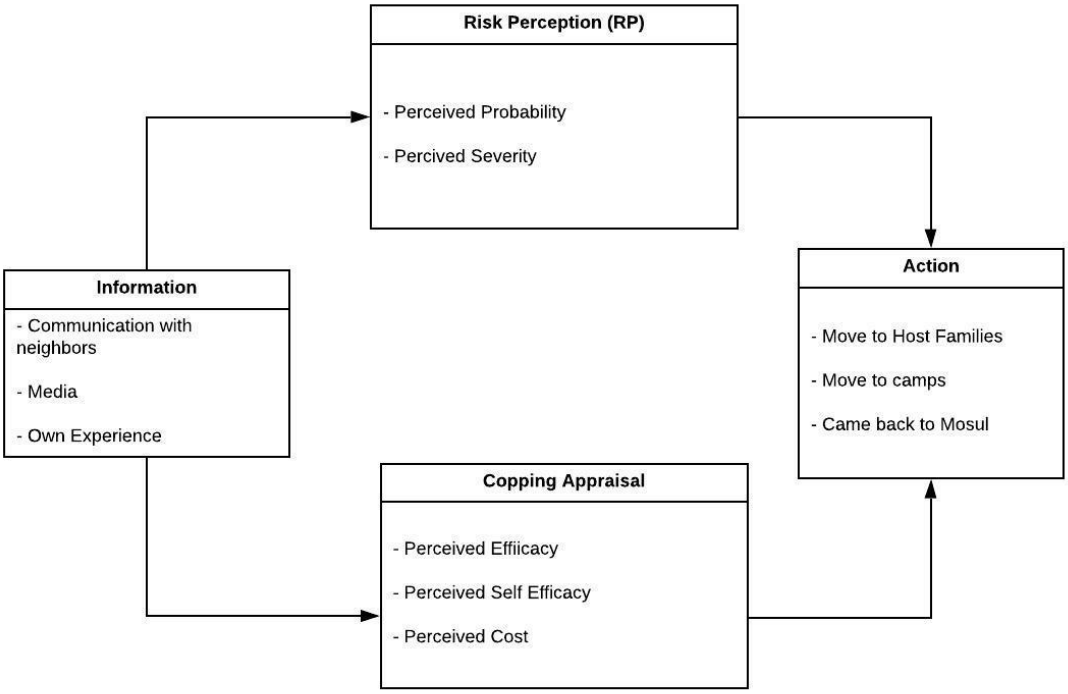


Figure 4.3 Cognitive process of Protection Motivation Theory

In a conflict zone situation the risk is high as a person is exposed to the risk of getting injured or being killed. Lafta, et al,. (2018) in their study on injury and death during the ISIS occupation of Mosul suggest that factors of risk were, injuries, kidnapping and death. Death and injuries were considerably higher during the military intervention than during ISIS occupation. In risk management, the risk is defined by Schneiderbauer & Ehrlich, (2004) as:

Equation 4.1 Risk Equation

$$R = H * E * V \qquad (1)$$

In wich R is the risk, H is hazard, E is exposure and V is vulnerability.

This general formula of risk perception is adjusted for the Mosul model in the following way:
Equation 4.2 Risk in Mosul equation

$$\boldsymbol{R_m = deaths_m * \ deaths_{social} * distance_m} \qquad (2)$$

In which m is Mosul and social is the social network of the agent. There is no mathematical relation between risk and risk perception at the time at the writing of this thesis, but the model assumes that risk perception always grows when risk grows.

Coping appraisal applied to the adjusted model :

Determines the action that the agent will take (when risk is higher than the risk threshold), in this case the spatial granularity of risk is distributed per neighberhood. Coping appraisal depends on the income level of the agent, restricted zones (freedom to leave the area) and the social network (family members and friends in non-occupied areas) and the condition of the house in Mosul (destroyed, or non-destroyed) that will be randomly assigned to the houses.

Fernandez, (2019) suggest that that if the household's income level exceed the migration cost, the agent will decide to migrate. The adjusted model will assume that when the destination is far, the migration cost will be higher.

Emergence :

The emergence is the different types of movement behaviour over the conflict phases. The key output of the model is the flow of IDPs and returnees through time. The following characteristics of temporal scale are expected.

IDPs during the pre-intervention phase will move to host families

IDPs during military intervention will move to camps

Returnee during post military intervention will return to residence if it is safe

Returnee during post military intervention will return to camps if their residence is not safe

IDPs will return to their residence if it is safe

IDPs will remain in camps if their residence is not safe

Adaptation:

(1) Agent Movement

Agents will move based on their risk perception and the quality of their condition (a higher income level, and a presence of social network will impact their decision to move). The quality of their condition during the pre-intervention phase have an influence on the agents. During a military intervention phase, the risk perception will have a higher influence on the agents behaviour. Agents will change status every time they decide to move

(2) Destination Choice

Destination choice will be influenced by agents' condition during the first phase and will be influenced by risk perception during the second and third phase. And will be adapted based on the capacity of the camps.

(3) Destination Selection

Destination selection will be applied to camps. Agents will decide to move to a camp if this camp is not full, otherwise agents will move to the next one. Euclidean distance will be calculated for the choice of the nearest camps. Camp's capacity will be also a factor that agents will take into account

Prediction:

Prediction is used to estimate the future risk in Mosul. Agents prediction of risk influence their behaviour. Based on actual data, the curve of deaths in the Mosul area will be loaded into the system. Data will be per neighbourhood, based on literature (Lafta, Al-Nuaimi, & Burnham, 2018). Agents will "predict" the risk for a moment into the future. The moment from the "current" tick is variable.

Sensing:

Agents are able to sense the risk. Their risk perception influences their behaviour. Hébert's model uses tolerance to conflict to asses people decisions. To determine the moment an agent will leave, a set of variables was associated with each agent. Hébert's model uses detailed population variables such as ethnicity, religion, wealth, gender, age and familial status. When these variables are combined, it determines the likelihood of an agent to leave. The model uses a simple comparison to simulate the decision rule: if the perceived severity of the conflict is higher than the tolerance, the population of agents decides to leave and change the status to IDP. The model uses death toll numbers to calculate the severity of the conflict. A cumulative stress factor was also used by Herbert's Model, contributing to the reduction of conflict's tolerance. The adjusted model will also use a set of variable such as income level, social network that will be combined and have a weight at each phase of the conflict. Using formula 2 above.

The element Hazard is already discussed under prediction. Exposure will be calculate, and will results in the number of deaths within the social network inside Mosul. Vulnerability will be calculate, and will result in the distance of agents from the deaths in Mosul, when deaths are closer to the neighborhood the agents will be more vulnerable.

Interaction:

In the adjusted model, what we refer to as Social network are, in fact, two distinct groups: neighbors and relatives. The first group is the neighbors inside Mosul, that will be assigned randomly to each agent. The second group are relatives outside Mosul, they are distributed through the top 5 provinces that attracted IDPs according to IOM data and will also be assigned randomly to each agent.

The interaction part concerns the first group, in this case agents will be able to inform an average of seven other agents about their own risk perception. Migration literature has shown that migration decision is not influenced only by individual factors but also by actions of other migrants (Fernandez, 2019).In the adjusted model, each agent at the initialization will be randomly assigned seven neighbor agents from the same neighborhood. Neighborhoods will take into account the income level.

Stochasticity:

(1) Agent initialization

Agents are randomly distributed among Mosul's neighbourhoods

(2) Choice of Destination

The decision making process included in the simulation is based on agents' risk perception and their reaction to it. The choice of destination is influenced by the quality of agent's condition.

(3) Risk perception

Risk perception is evaluated empirically with a threshold and values are set accordingly, risk perception will have a higher weight during military intervention phase.

Collectives:

The household agent is a collective as each agent represents a households of 8 individuals and all of these individuals will have the same behaviour. A second layer of collectives could be based on ethnic background.

4.1.3. Details:

The adjusted model is based on data provided by the IOM and UNHCR IraqThe terminology is based on IOM definition. The top 5 provinces attracting IDPs are : Ninewa, Dahuk, Erbil, Kirkuk, and Baghdad. To run the model the following parameters must be set: Households, Restricted Areas, Camps capacity, IDP, Returnees

Initialization:

Households agents

Agents will be assigned randomly a risk threshold , an income level, and a social network. Social network is divided in two distinct groups:

a group inside Mosul from the same neighbourhood, each agent will randomly be assigned of 7 neighbours

a second group of relatives outside Mosul , the number will be randomly assigned to each agent.

Table 4.5 Initialization

Object	**Details**
Households	1) Number of Mosul's Households (11000) 2) Income level of households 3) Have 7 Neighbours 4) Number of relatives outside Mosul

Number of households was retrieved from UNHCR report (1 223 374) the available data was per person, in our model the number was divided by 8 (the average number within a household) , the model approximatively represent 1% (11 000) of the total households in Mosul.

Spatial environment

Table 4.6 Spatial enviorment

Object	Details
Environment	1) Restricted Areas are open 2) Camps are empty 3) Number of death per neighbourhoods 4) Provinces

Spatial environment is an essential part of Agent Based Model as it shapes the environment in which agents interact (MacAl & North, 2010). In Mosul's case study the spatial environment are : The city , the surrounding camps and the surrounding provinces. As the geographical location of camps and provinces has no impact on the model behaviour , their location has been altered to make it more compact using Arcgis. The scale of camps and provinces were reduced, and extra step was done for the camps, the mean center was therefore identified and XY line was then drawn between camps and Mosul boundary , the camps where then dragged fallowing the line and where placed close to Mosul.

Table 4.7 input data

Object	Data
Camps Capacity	CampCapcity per camp (total of 65922 pp for the projected camps)
Provinces	Number of IDPs distributed through the top 5 provinces that attracted IDPs according to IOM data (total of 8,748,888 people)
Number of deaths	Retrieved from literature review, split in to two curves , east and west Mosul.
Percentage of income level among Mosul population	8% very high income level , 8% high income level , 13% above average, 15% average , 19% below average, 21% low, 13% very low retrieved from UNHCR
Number of neighbors	Each agent has 7 neighbours that is linked to
Hostfamilies	Up to 60 households have host family

- **Camp capacity:** the total camp capacity is in per person , the number was converted to number of households and was scaled down to 1 %

- **Number of deaths: was** retrieved from Lafta, et al,. (2018) , the numbers were extracted from a graph and written in an excel sheet, with 3 columns , time, number of deaths in east and number of deaths in west.
- **Host families :** a random number of households has host family

Submodel

The Model is split in to two sub models, (1) the flow of IDPs leaving Mosul and (2) the inverse flow of Returnee retuning to Mosul.

4.2. Model Implementation

The model described in the previous section was implemented in Netlogo software. This section will describe the steps that were used for the implementation of the model. GIS and CSV extensions were used; their purpose will be revealed in the following section.

Geodata was loaded using the GIS extension in the **setup procedure**. The Mosul boundary was set as the environment (see Figure 4.4). In this figure, camps are represented as green houses, provinces are represented as yellow stars, and buildings are shown in grey colour. Camps, provinces and buildings were loaded as global variables. The East side of Mosul is represented in Figure 4.4 by blue patches and the West Side of Mosul is represented by the green patches. The black patches are the environment of Netlogo

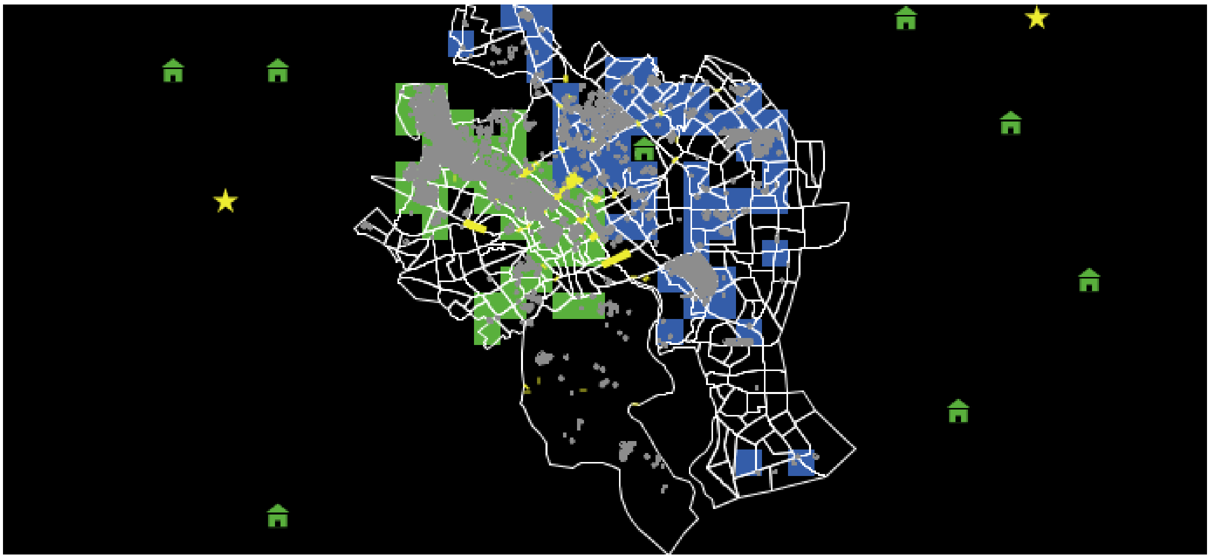


Figure 4.4 Mosul Environment

Households were created using the **create-agents procedure**. A random dwelling from the geodata is defined as their home location using the centroid of a building (See Figure 4.5). Each households is randomly assigned an income level ranging from very high income to very low. The percentage of households that fall into a given income range was set according to the income distribution classification of the UNHCR (UNHCR, 2019).

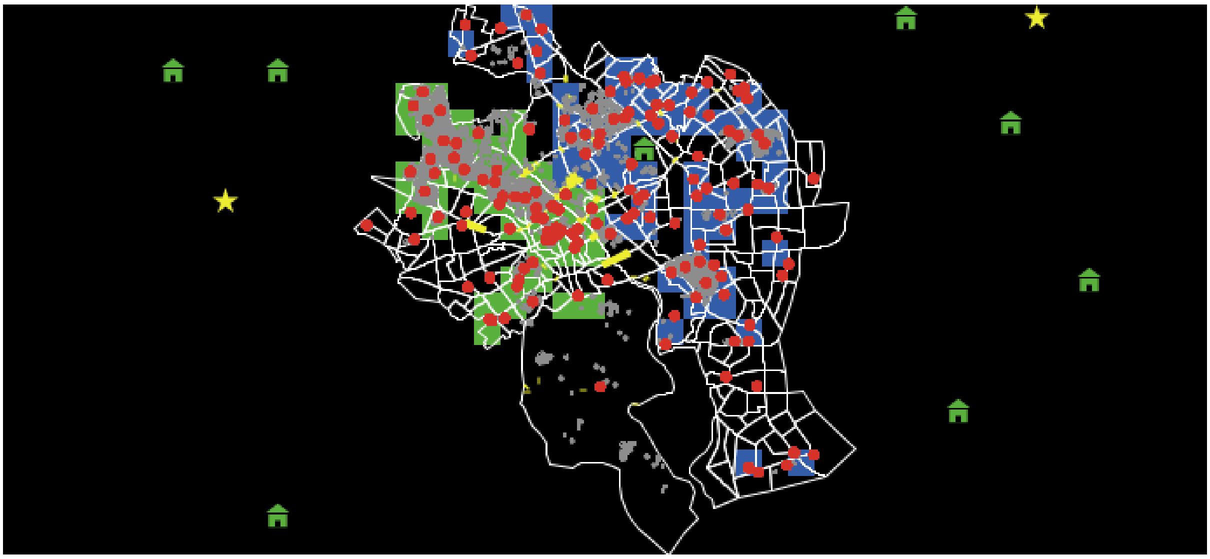


Figure 4.5 Households creation

Restricted areas were loaded under the **go procedure**. Four different restricted areas were loaded corresponding to the fighting area at a particular date. The model will load a new restricted area on tick 11 for the date of the starting of the military intervention in Mosul (21 st October 2016). On tick 13 for the date of November 03rd 2016 and tick 16 (November 24th 2016), for the development and the end of fight in the eastern part. Another restricted area is loaded on tick 23, for the development of the intervention on the west side (January 9th, 2017) and tick 51 (July 31st, 2017), once the city is fully liberated. Restricted areas were provided by WHO Iraq .

Social Network

To ensure that the households have links to other households and can exchange information, two types of social network were created. The first one is the neighbour network, where households create links with 7 neighbouring agents based on the minimum distance, and depending on which side of Mosul (east or west) they are located. This network was implemented using the **links** functionality of Netlogo (see Figure 4.6). A clustering was then performed based on income level. A random number of households was asked to create a "splotch" of patches around them, to attract other agents with the same attribute (income level). These households were moved to the splotch location.

The second type of social network is the host family network. For this network a random percentage of households (25 %) was set to have host families as an attribute. It is assumed that Host families are outside the Mosul area. If the households has an attribute with host family, the agent will move to one of the provinces.

Deaths in Mosul :

Deaths in Mosul for the period of time under study were retrieved from literature (Lafta et al. 2018) and were written in an excel file. The data describe the average number of deaths per month, the excel file was therefore adjusted to the number of ticks. The file was plotted in the Netlogo interface seen figure 4.6. The pattern in the number of deaths is different for the east and west side of Mosul. The death peak in East Mosul occurs earlier and is lower than that of West Mosul, which is a good indication of the timing and intensity of the liberation battle. In the model, agents have access to both the local death rate (east/west) and the total death rate for the complete Mosul area.

At a given tick, the actual number of deaths is distributed over the households, giving these households the status "death". This information is communicated via the social network of the agents.

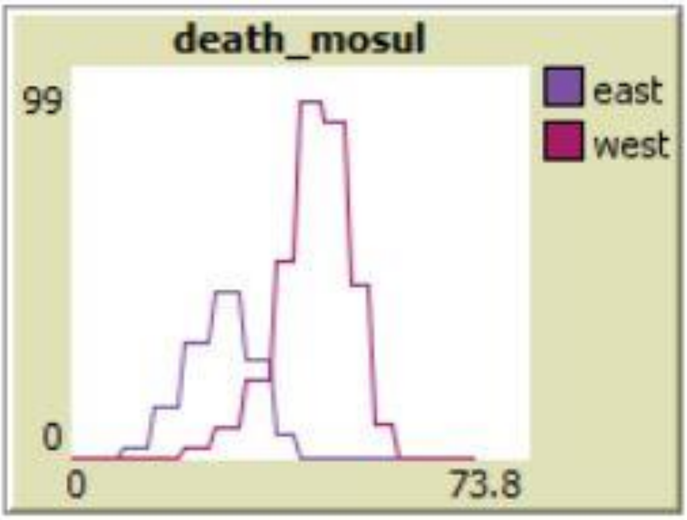


Figure 4.6 Deaths in Mosul

Risk perception

The households with deaths will communicate these deaths via their network to their neighbours. The risk level in Mosul is then calculated (Equation 4.1). Each household with an income level from average to very high is asked to compare its own risk perception threshold to the general risk and check its own location (inside or outside a restricted area). In case the household had a risk perception threshold lower than the general risk in their part of Mosul (East or West), and the restricted area is open, the household changes its status to "IDP" otherwise the household remains a "Stayer".

Choice of Destination

The choice of destination is described in Figure 4.7. Households with an income class high and very high that change their status to IDP, are asked to check their host family attribute. When this attribute is set to true, the IDP moves to the closest province. In case of a high and very high income level with no host family, the IDP could also choose the closest province. In the case the IDP has a host family but is of lower income level, the IDP can also choose the closest province.

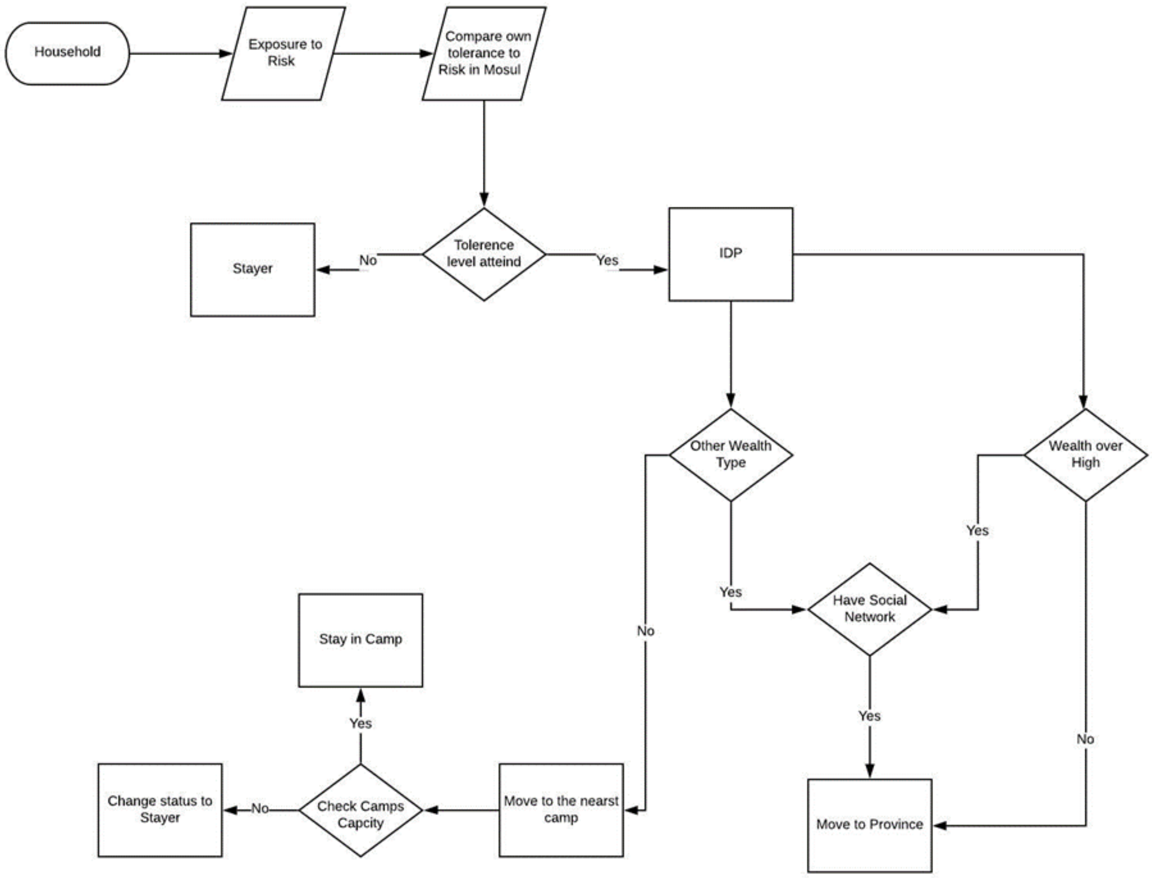


Figure 4.7 Flowchart of idp choice of destination

IDPs with no host family and lower income level, are asked to move to the nearest camp and check the camp capacity. In case the camp is full, IDPs are asked to check other camps. In case there is no camp available, the household remains a stayer.

Time was introduced as a factor in determining the destination for all incomes. When the time becomes shorter, IDPs choose camps over provinces. This helps in balancing the number of IDPs going to camps.

Reverse flow of IDPs

After becoming an IDP, this IDP can decide to return back to Mosul (become a returnee). IDPs are asked to compare the total deaths in Mosul in order to return. IDPs will calculate the trend in the number of deaths to determine if the area becomes safer. For this calculation, the IDP will use the death rates of the two preceding weeks. IDPs were asked to read the Excel file and compare the deaths of the two previous weeks; then, the risk perception condition was applied. If the number of deaths was decreasing and there were no further deaths in the social network, the IDP will check if their house is not damaged. When the house is not damaged, the IDP returns home and changes its status to returnee.

5. MODEL TESTING

In this chapter we will discuss the results of the verification of the model. It is split into four parts, the first part (section 5.1) will be the verification of the code. The second part (5.2) will be a sensitivity analysis of the model, the third part (5.3) will verify the robustness of the model, and the fourth part (section 5.4) will test its stability

5.1. Verification:

5.1.1. Number of links :

The first verification made on the model was the number of links that agents have in their social network. The results of the verification shown that households has a minimum of 7 links, a maximum of 24, and an average of 13 links.

5.1.2. Destination :

- **Provinces**

The results after verification shows that 0.54 % of IDPs were displaced to Provinces. Seventy eight percent (78%) of the displaced Households in Provinces have a very high income. Thirty percent (30%) of the Households who moved to provinces have a high income. Above average and average income levels score respectively 15 % and 1 %.

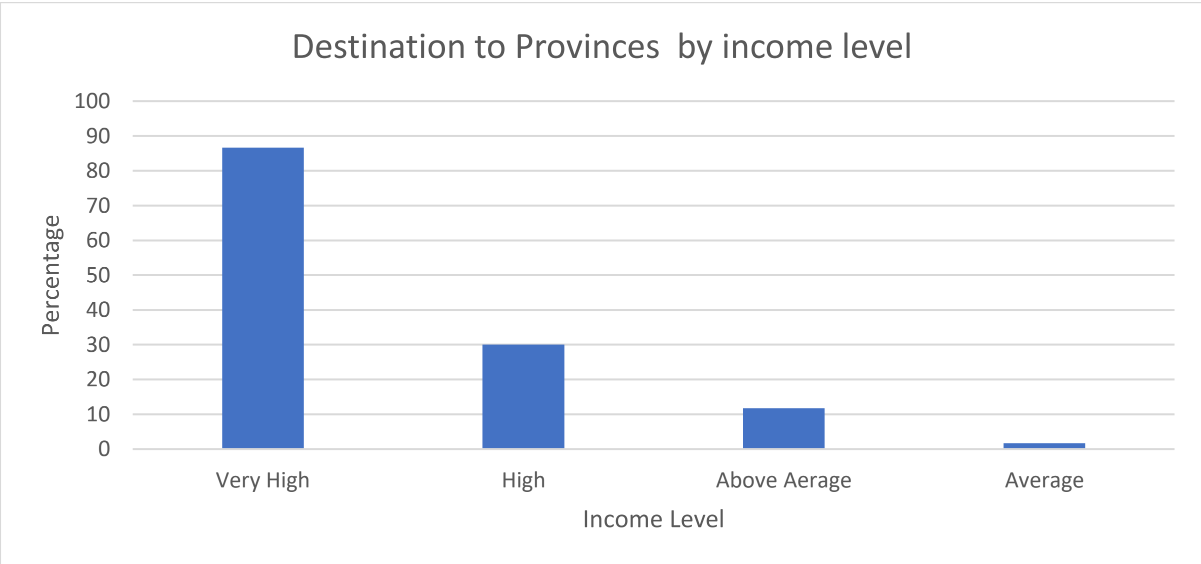


Figure 5.1 Destination to provinces by income level

The figure (5.2) on the provinces that hosted IDPs shows that only Ninewa and Erbil where selected as destination. Figure 5.2 shows that Ninewa has hosted the highest number of IDPs.

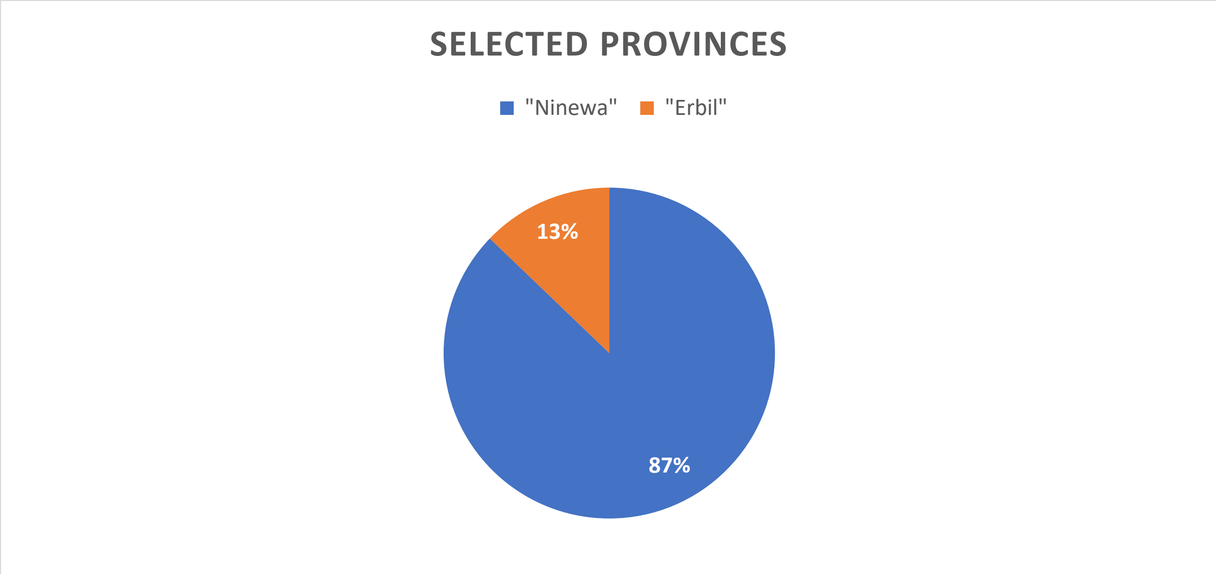


Figure 5.2 Selected Provinces

- Camps

The verification of camps shows that households with all income level have moved to camps. Figure 5.3 shows that High income level scored the highest percentages of income level within camps representing 60 % of IDPs, followed by Above Average and Average Income level representing 15 % and finally the high income level group has scored 10%.

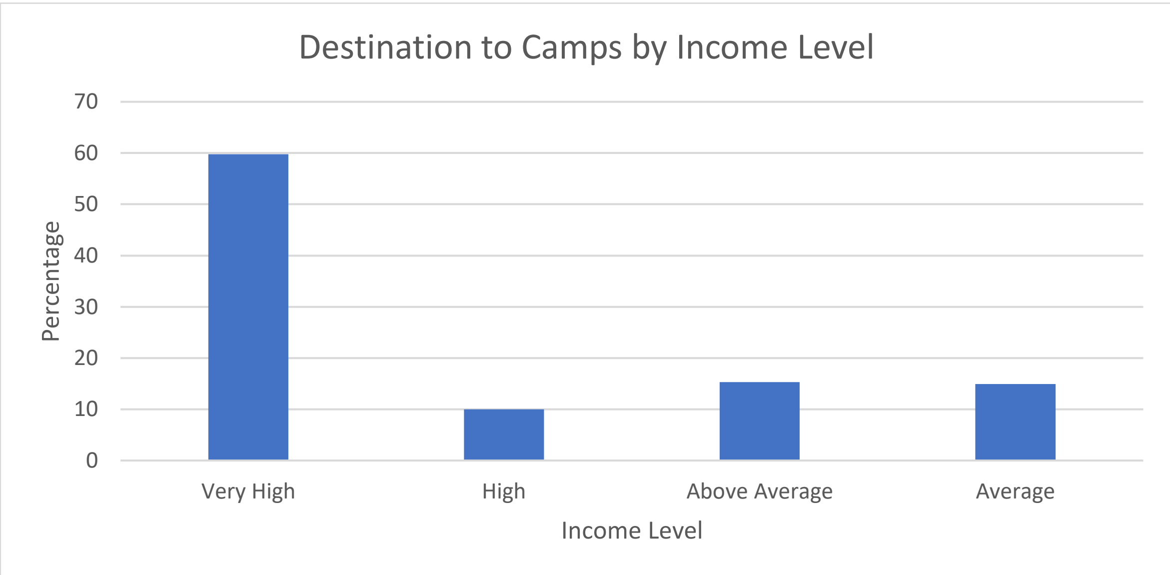


Figure 5.3 destination to Camps by income level

On the other hand the verification of hosting camps has shown that only 5 Camps seem to attract IDPs Figure 5.4 shows the percentage of IDPs per camps.

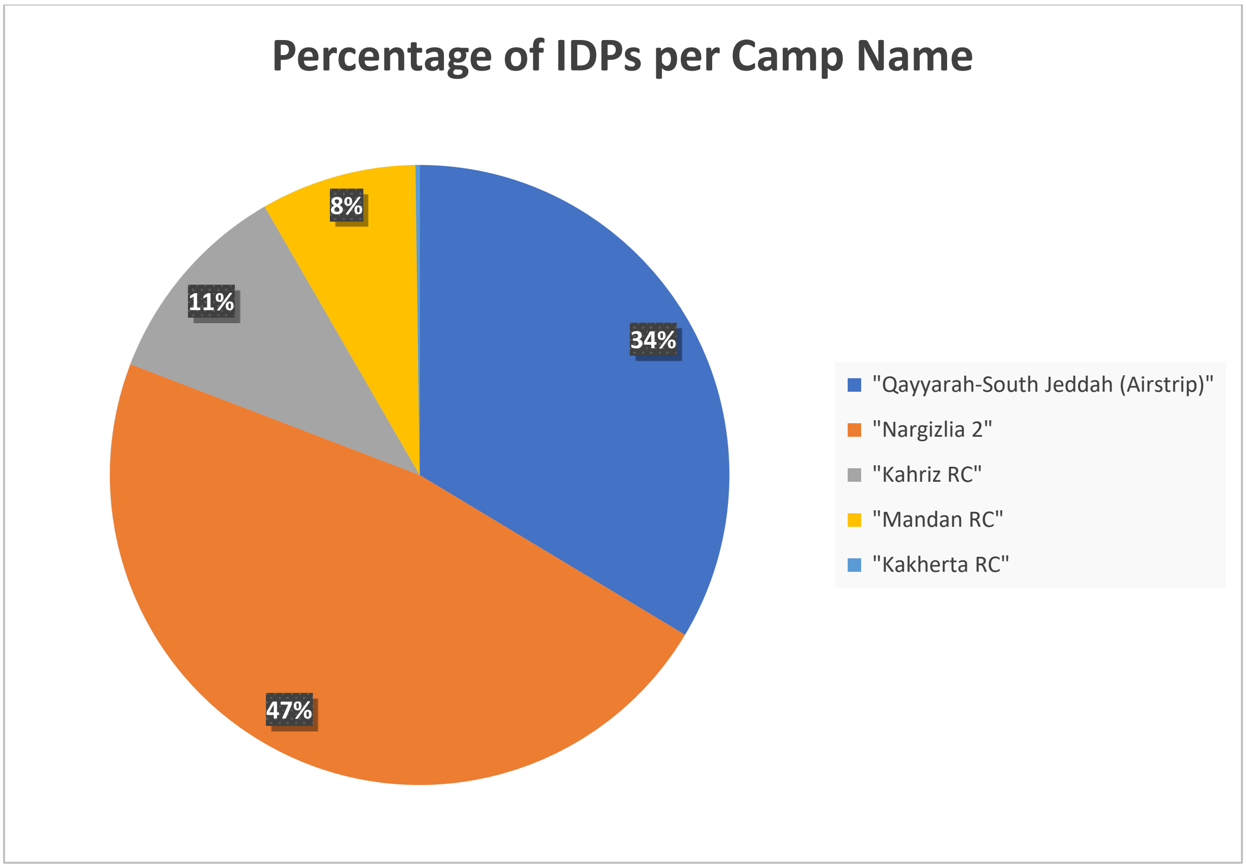


Figure 5.4 IDPs per Camp

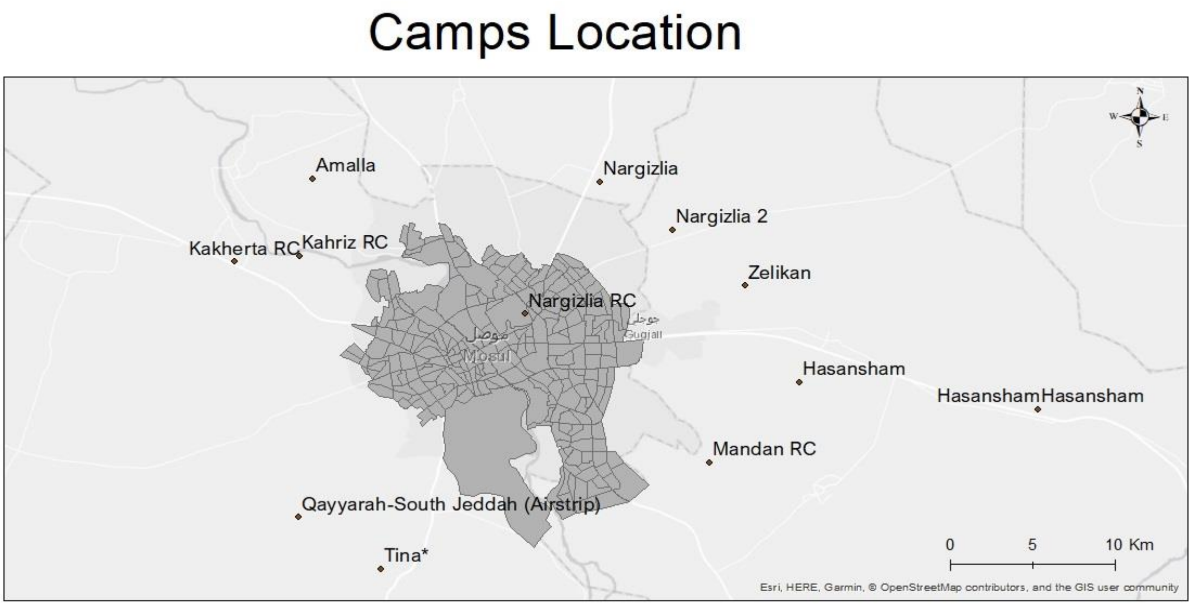


Figure 5.5 Location of Camps

Figures 5.5 shows that even though camps Amalla & Tina are close to Mosul, Inspected patches show that both camps ae not visualised, Suggesting that Netlogo takes into consideration only the visualized ones.

5.1.3. Perceived Risk for one household :

In order to verify that every single agent perceives risk, an output file was printed with risk perception for each household on each tick. Figure 5.6 shows the result for Household 10 after running the simulation.

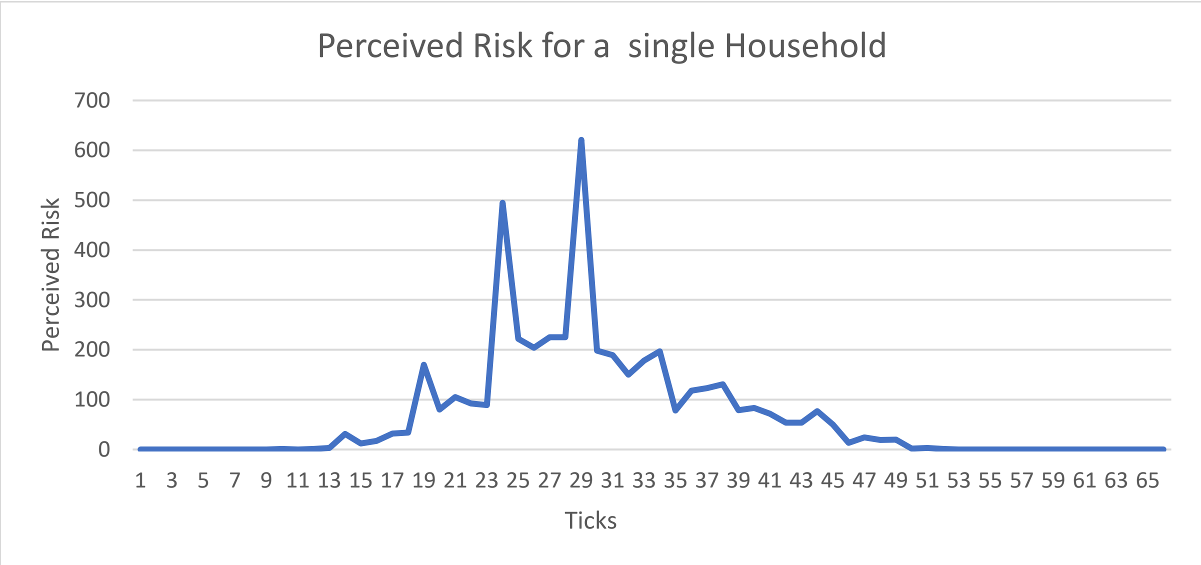


Figure 5.6 Risk perception for one household.

When plotting the number of deaths in both sides of Mosul with the risk perception graph, the peaks at tick 24 and tick 29 are triggered by the increase of the number of deaths in Mosul West, when both curves of Mosul East and Mosul west intersect, the risk perception scores it's highest number. However the high number of deaths in Mosul west didn't increase the risk perception. The peak at tick 14 is triggered by a change in restricted area. Some parts of Mosul were restricted until this tick. The liberation led to more household being able to leave the area (perceiving risk). The same applies for ticks 17 and tick 24. Restricted areas were loaded respectively at ticks 11, 13 , 16 and 23.

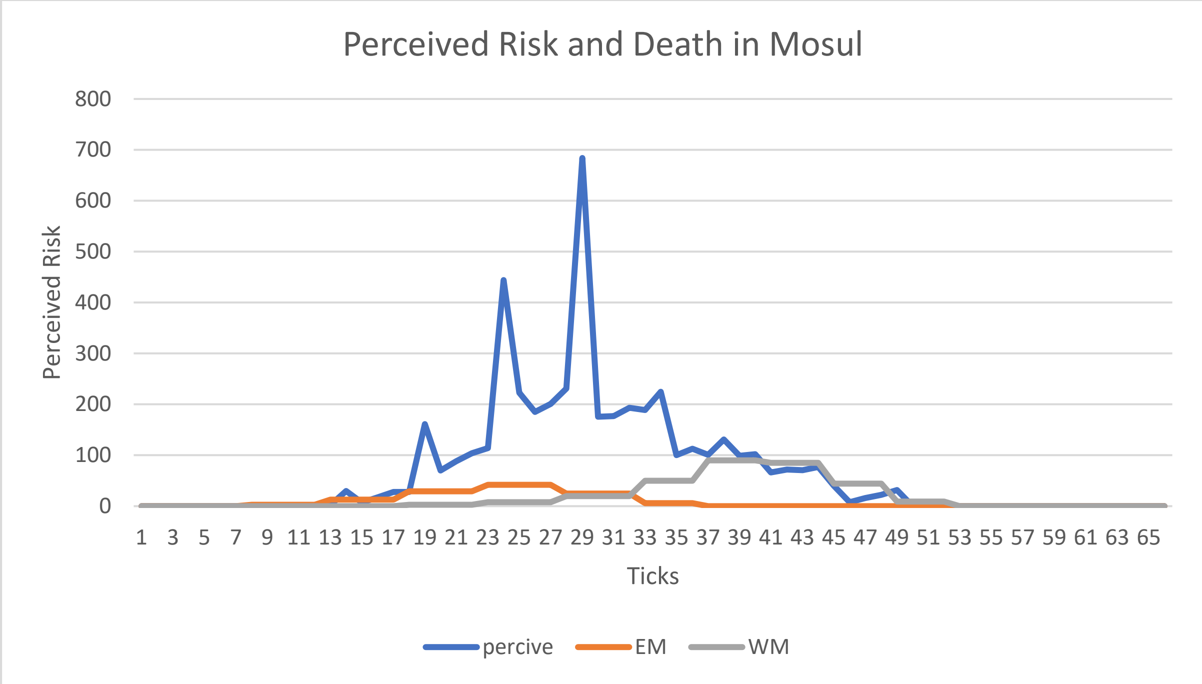


Figure 5.7 Impact of deaths in Mosul on perceived risk

Households 10 was selected to verify the behavior of a single household, verification show that the agent has a constant risk thresholds during the simulation. .In order to understand the relation between the perceived risk and the death in social network graph was plotted.

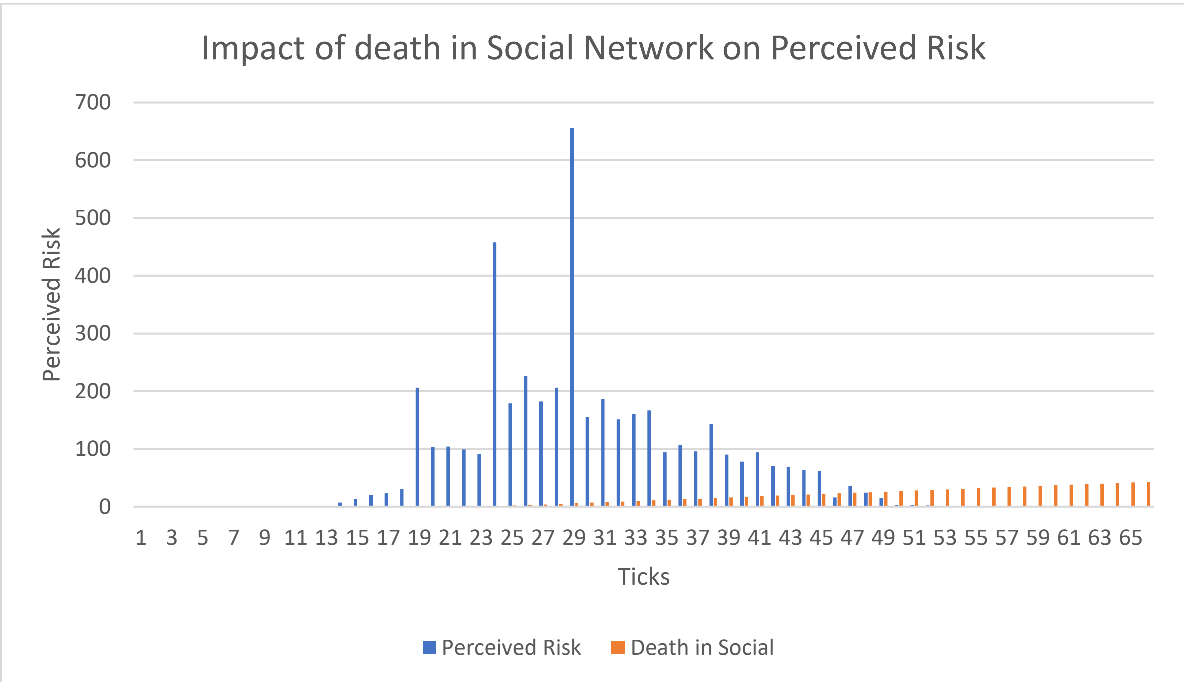


Figure 5.8 impact of death in social network on perceived risk

The results show that while risk perception was at its peak at tick 29, the deaths in the social network started in tick 24. Suggesting that the pick in perceived can be triggered by the number of deaths in the social network

5.1.4. Restricted area and movement :

Once the risk perception of a single household is verified, we proceed with the verification of the movement according to the restricted area. The movement of agents and the restricted areas are therefore printed into an output file. A graph was plotted to examine at which point of the restricted area status the households left Mosul. The result for household 10 (figure 5.8) shows that the agent changed its status to "IDP" at tick 32 and moved at tick 34 when the area was false (meaning open). The graph illustrate that once a restricted area is true it remains true for a number of ticks and the Households leave the city when the restricted area is open.

On the other hand in order to understand the impact that deaths in the social network have on the movement of agents, verification shows that Households 10 has left Mosul city 2 ticks after recording a death in his social network.

5.1.5. Patch damage verification

To verify how long a patch is damaged the patch coordinates were printed in an output file along with each tick and its status (damaged = 1 and not damaged = 0), a curve was then plotted in Excel to verify that once the patch was damaged, it remained damaged for the rest of the simulation. Figure 5.9 shows the example of Patch 28 8 that was damaged in tick 22 and that the patch remained damaged for the rest of the simulation.

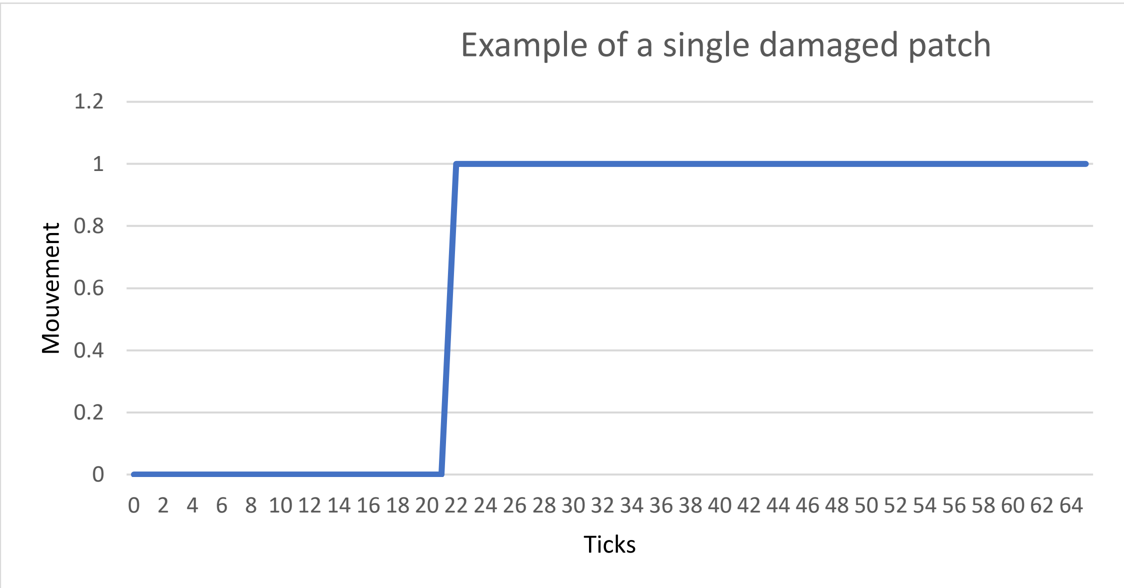


Figure 5.9 Damaged Houses

5.2. Sensitivity Analysis

Experiment 1: Perceived Risk

In order to test the impact that variables have on the risk equation, a sensitivity analysis using the Behavior Space tool was performed. A factor variable was introduced for the number of deaths as a weight and two experiments were conducted. The factor was applied for the deaths in the social network and the deaths on each side of Mosul. The perceived risk was selected to test the impact.

In the first experiment, the weight of the number of deaths in the risk perception equation (Equation 4.1) when multiplied by 2 and ten runs were performed. In the second experiment, the impact of the weight was multiplied by 5, also performing ten simulation runs. The results of both experiments were compared to the values of the base runs (Table 5.1).

Table 5.1 Comparison between the perceived risk in experiments and the original model outcomes (base)

VARIABLE	BASE	FACTOR X 2	FACTOR X 5
DEATH IN SOCIAL NETWORK IN WEST	Mean: 22.6 Max: 140.1	Mean : 22.1 Max : 145.1	Mean : 22.2 Max : 144.9
DEATH IN WEST OF MOSUL	Mean: 23.0 Max: 146.6	Mean : 22.4 Max : 141.9	Mean : 22.2 Max : 146.6
DEATH IN EAST OF MOSUL	Mean : 47.5 Max : 598	Mean : 48.0 Max : 595.1	Mean : 48.3 Max 611.5
DEATH IN SOCIAL NETWORK EAST	Mean: 48.25 Max: 606	Mean : 48.18 Max: 595.3	Mean: 47.81 Max : 597.7

Results show the risk perception in east Mosul is higher than in west Mosul,. and that for the people in the east it mattered less if a death was in their social network , then it did for the people in the west. Because the people in the east have been exposed longer to acute danger and fighting – fighting stated in the east – so their risk perception was higher and lasted longer. This causes a significantly higher number of IDPs in east Mosul compared to west Mosul. However, the increase in the factor variable does not have an impact on the perceived risk and the number of generated IDPs.

Experiment 2 : Time Factor

In the second experiment , time as a factor for the selection of the destination was tested. Perceived risk was divided by a factor in order to balance the number of IDP to camp destination. Results shows that the division by the factor impact results in a low number of IDPs in provinces.

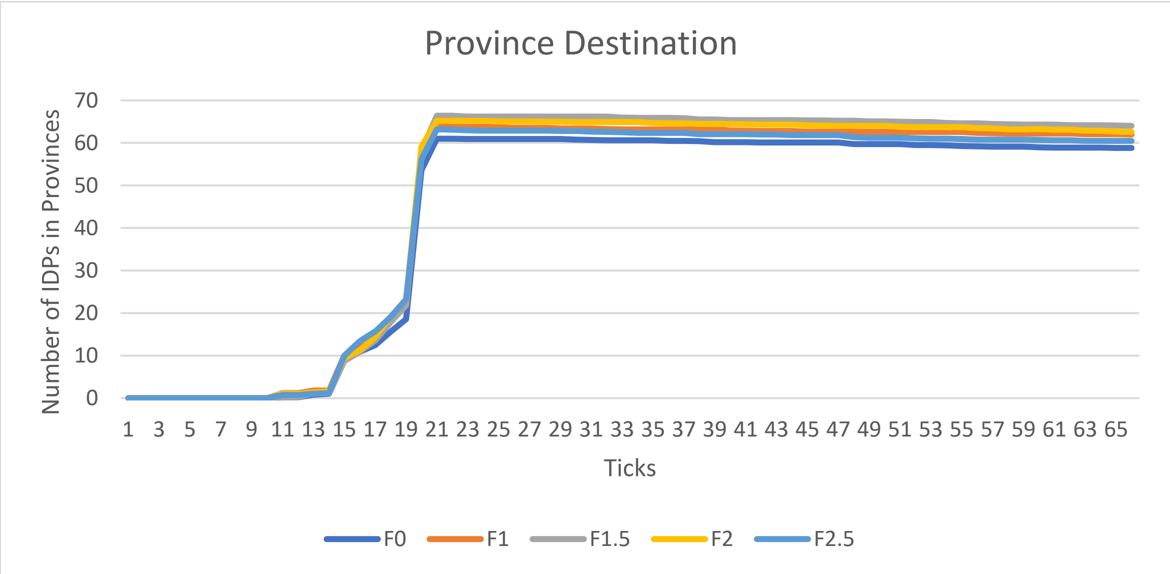


Figure 5.10 Generated number of IDPs in provinces

Table 5.2 generated number of IDPs to provinces

	F0	F1	F1.5	F2	F2.5
[max]	61	64	66.4	65.2	63.2
[mean]	43.6697	45.99545	47.55909	46.87424	45.20758

However the change in the variables does not impact the number of IDPs in camps.

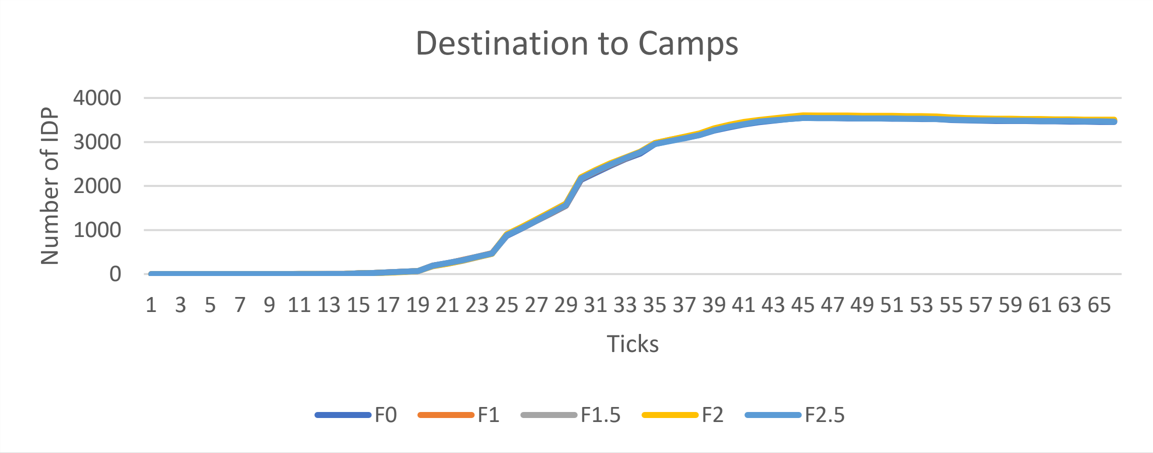


Figure 5.11 Generated number of IDPs in camps

Table 5.3 generated number of IDPs to Camps

	F0	F1	F1.5	F2	F2.5
[MAX]	3546.7	3573.2	3569.4	3605.5	3546.3
[MEAN]	1966.638	1983.256	1983.982	1997.291	1972.064

5.3. ABM Robustness

To test the robustness of the model, behavior space was used to run the model ten times. The total number of IDPs over time is plotted in Figure 5.12. The results of the experiment show that the results of the model do not vary a lot per model run. This is to be expected as the model is mostly deterministic. The shape of the curve is also almost the same for all runs.

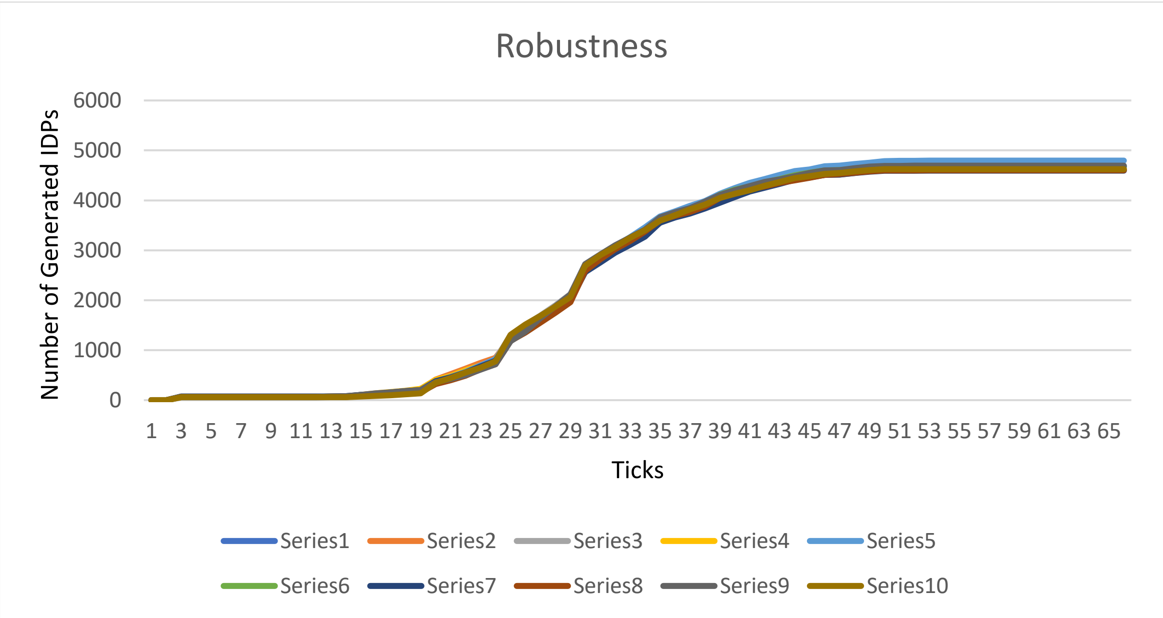


Figure 5.12 : model robustness

5.4. ABM Stability

To test for stability of the model, 100 runs were performed using behaviour space, and the accumulative average number of IDPs was plotted. The result is shown in Figure 5.13, showing that the model becomes stable after approximatively 67 iterations, at a value of 4630 IDPs.

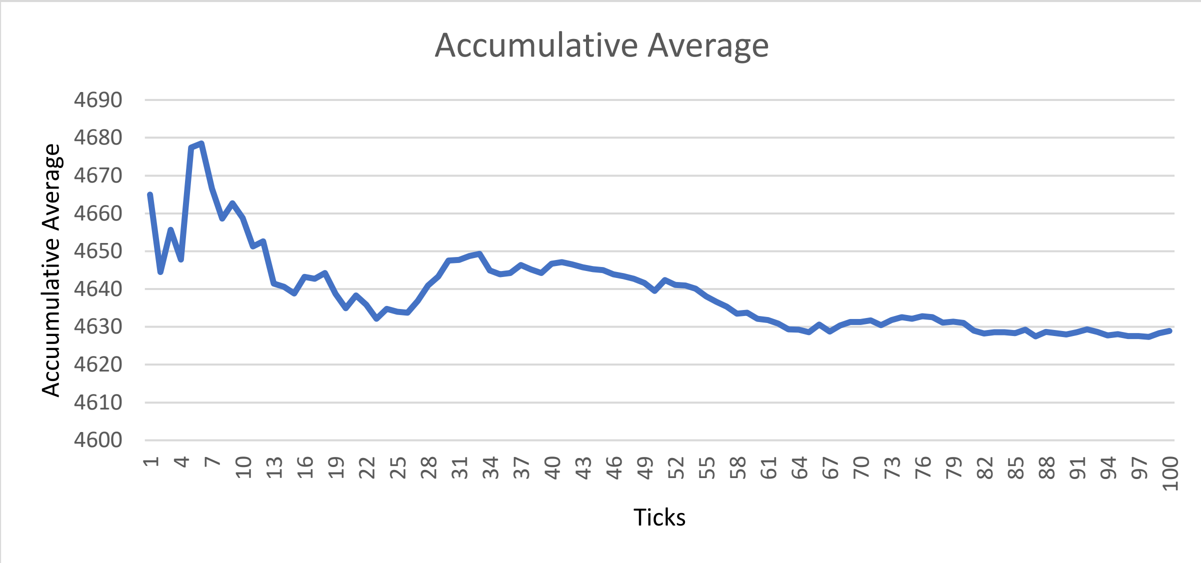


Figure 5.13 : accumulative average of IDPs

6. RESULTS

Validation differs from verification and calibration in the sense that it evaluates whether or not the model is a true representation of the real system. Almost 20 years ago, Edwards (2003) argued on the importance of computational tools and ABM in understanding the pattern of displacement. Pattern-Oriented Modelling (POM) looks at the emerging pattern and identifies generative mechanisms by examining patterns observed in the system.

This chapter will first discuss the results of the analysis of the empirical data (section 6.1) it will look at the calibration of two model parameters – the factor in the social network and the number of damaged buildings (section 6.2), and then describe the validation process, the comparison of the empirical data and the model outcomes (section 6.3) of the model using POM

6.1. Data Analysis

The next sections discuss the results of the empirical data analysis. This will be done by first looking at the analysis in time, followed by the analysis in space and time.

6.1.1. Analysis in Time :

Figure 6.1 illustrates Mosul's population displacement. The IDP group, represented by a blue line, shows the pattern of a bell curve with an increase of displacement starting from November 2015, reaching a peak in September 2017 when 113,424 households were displaced. Although a decrease in IDP households is noticed starting from December 2017, the curve shows a high number of displacement is still going on, raising the question on the profile of IDPs leaving after the battle was over. On the other hand, the Returnee line shows a concave curve, starting from September 2016. The line didn't yet reach its peak on December 2019, with a current number of 170.056 households returned. Both graphs intersect in December 2017. Between April 2014 and November 2015, there is a wave of the first IDP households, from December 2015 till November 2017, a second wave of IDP households leaves Mosul, and a last wave of IDPs that start's from March 2018 till December 2019.

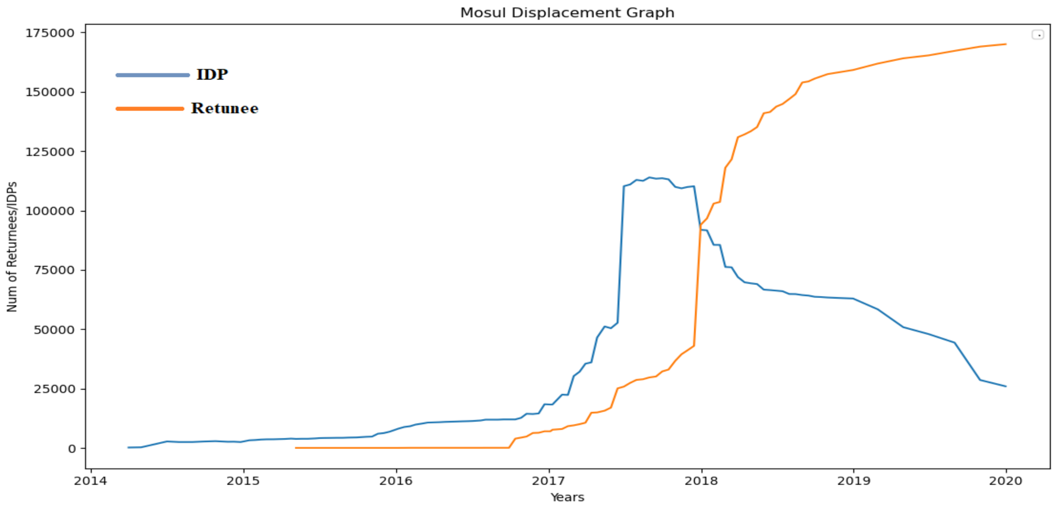


Figure 6.1 Mosul Displacement Graph

6.1.2. Space and Time analysis

Figure 6.2 reflects on the location of IDPs and Returnees through time. For IDPs, the movement is oriented from the center of Mosul to the south and the outskirts of Mosul. For returnees, the movement starts from the south of Mosul in 2017, where there is a large number of camps, and increases in the center of Mosul, north and the outskirts of the city between 2018 and 2019.

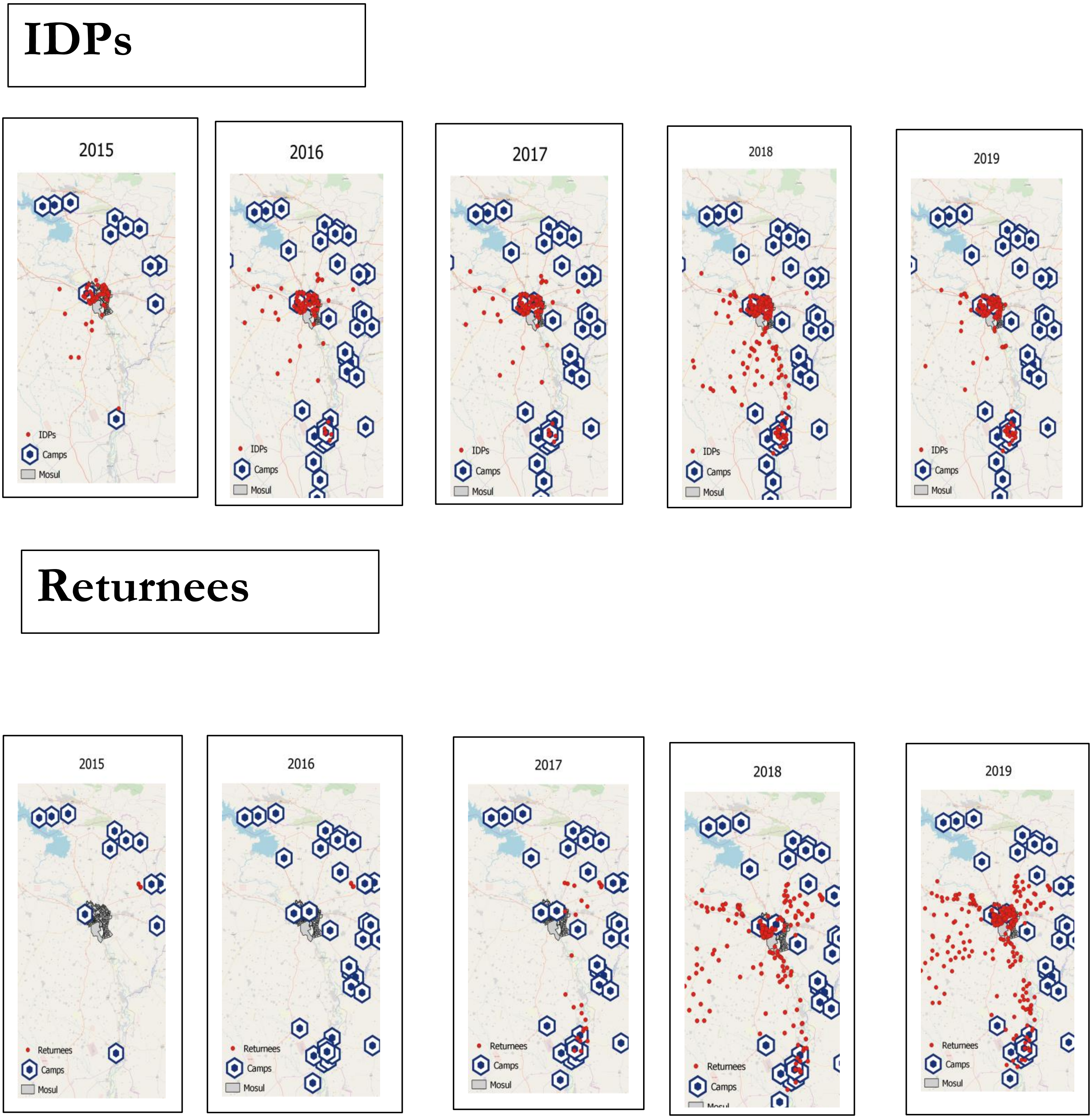


Figure 6.2 Movement in space and time

6.1.3. Shelter Types :

Empirical data analysis of IDPs by shelter type, illustrated in figure 6.3 shows a high percentage of IDPs moving to camps (50%), unknown shelter type (30%) and Host Families (12%). Compared to 75 % choosing camps and only 3 % choosing provinces after the ABM simulation

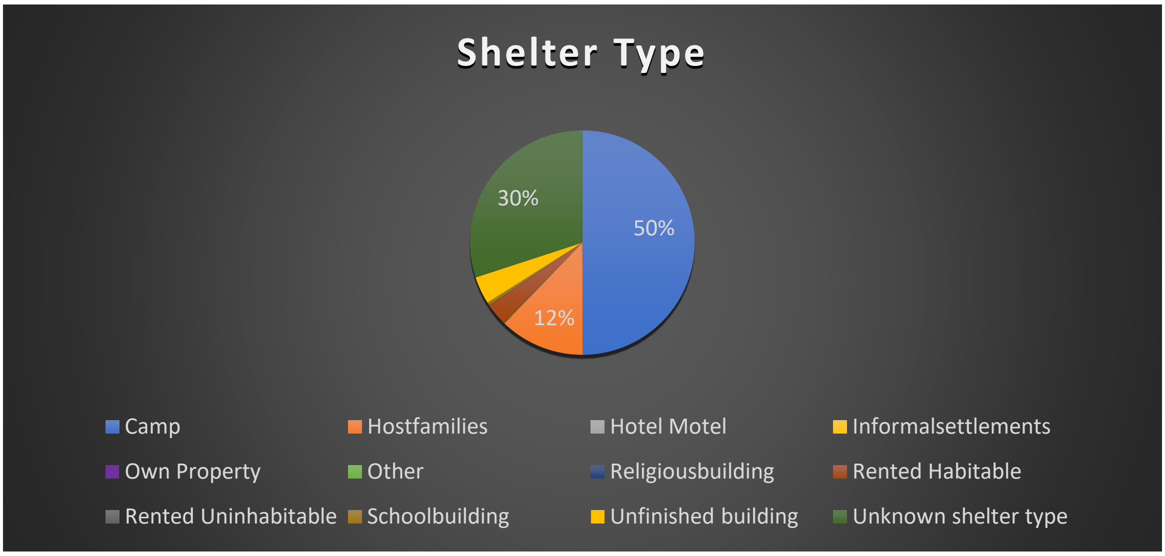


Figure 6.3 Type of Shelter

However, figure 6.4 illustrates the number of IDPs by shelter type over time. The graph shows that the first wave of IDPs had a higher number, 4388 households choosing host families starting from July 2014 and a later peak in March 2018 of 21,466 households. On the other hand, the camp lines show a wave of later leavers with an increase of household displacement in November 2016, reaching the highest peak in August 2017 of 78.014 displaced households. The figure shows large numbers of IDPs in camps after August 2017 till December 2019. This large number can be explained by the destruction of houses and infrastructure after the battle, causing IDPs to remain in nearby camps. IDPs that are going to unknown shelters leave Mosul after the end of the battle. Figure 6.4 shows an increasing number of IDPs after the end of the battle, this can be explained by a group of fighters leaving the city after the end of the battle to an unknown destination.

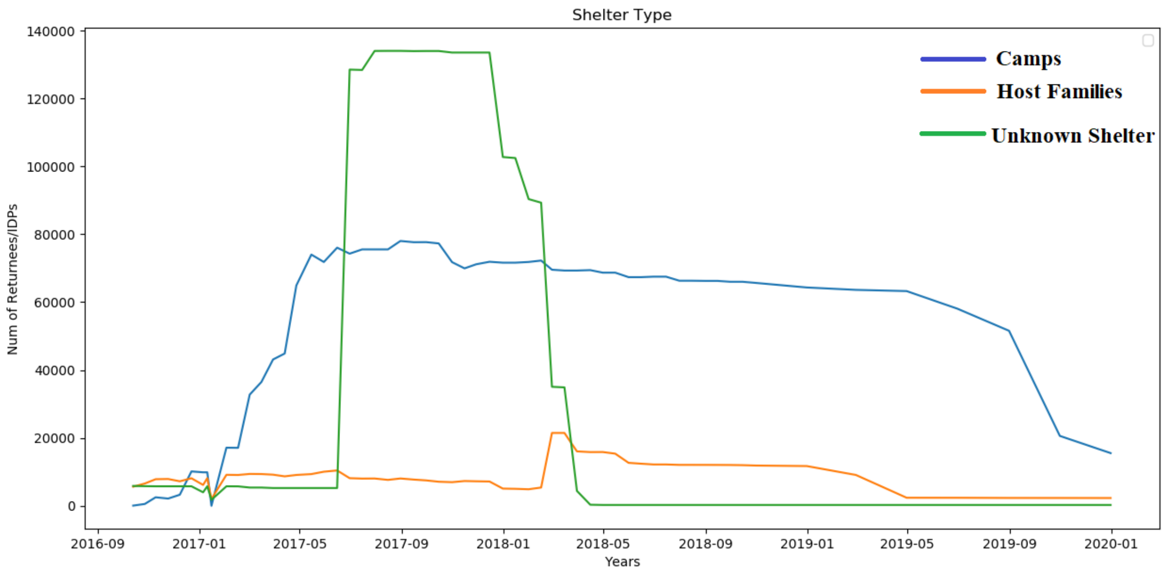


Figure 6.4 Shelters by Type

Empirical data has identified 5 top province destination the first one being Ninawa governorate, where Mosul is located, hosted the highest number of IDPs suggesting that IDPs were displaced inside the governorate, meaning that they were either in camps or host families.

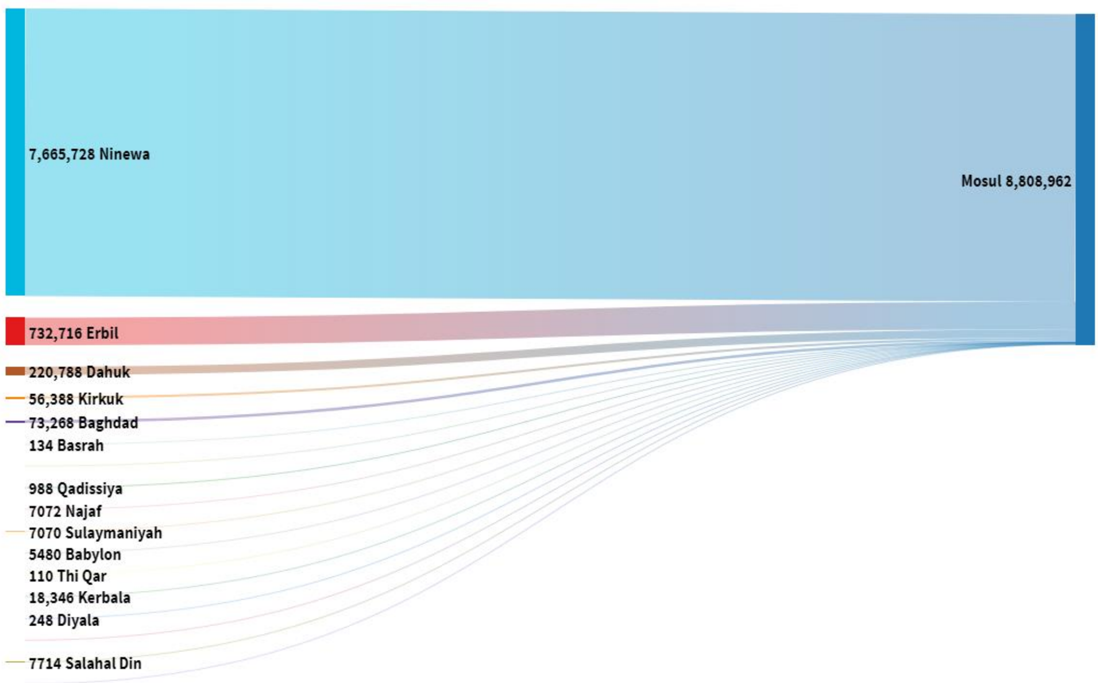


Figure 6.5 IDPs by province

Provinces such as Duhok, Erbil and Kirkuk, also recorded a high number of IDPs, showing that displacement happened mostly in the northern part of Iraq. An exception is Baghdad that is ranked the 5th hosting governorate according to data. This can be explained by the importance of the capital city that is also located in the same governorate.

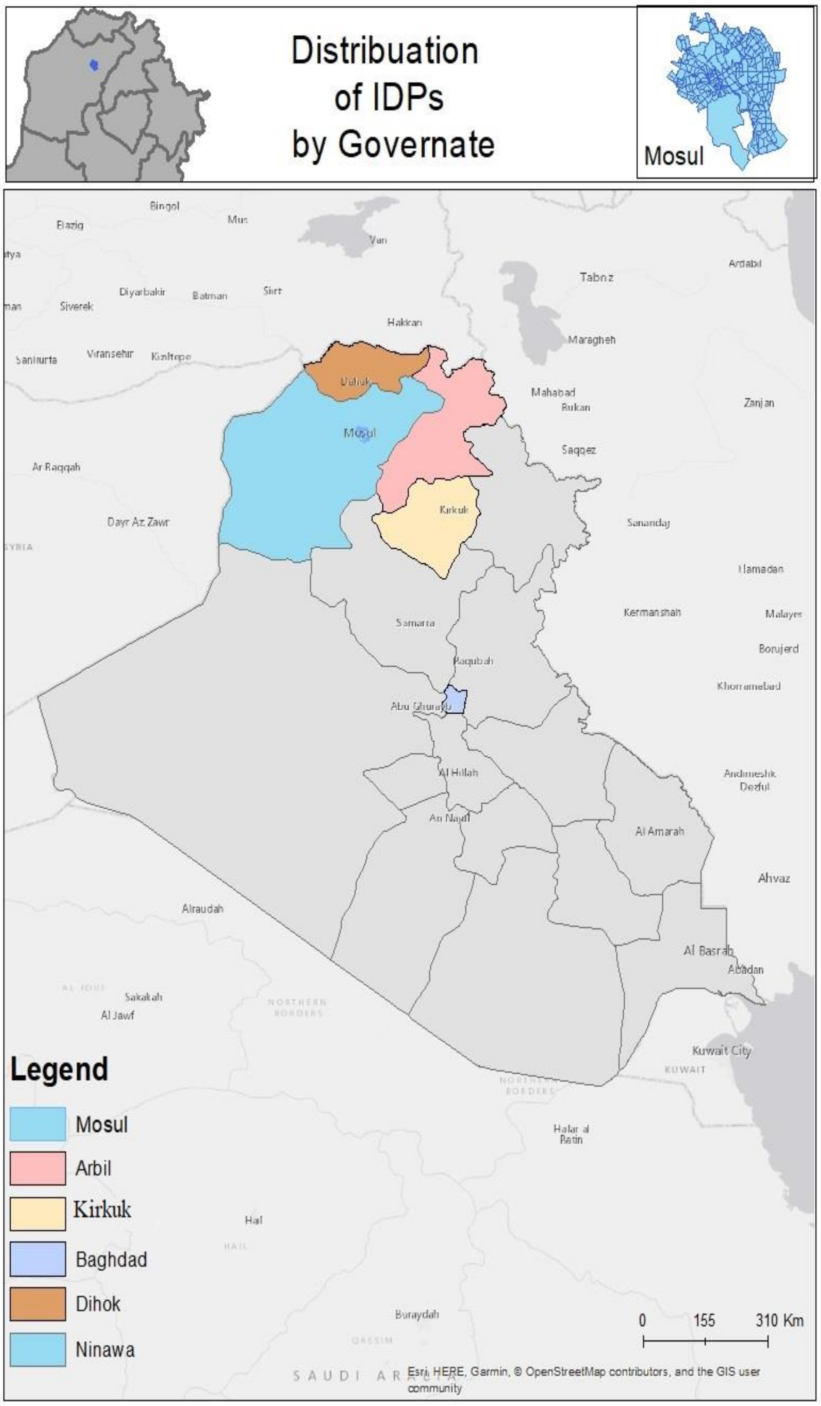


Figure 6.6 Spatial Distribution by province

Figure 6.5 shows only Ninewa and Erbil as the dominating Provinces that attracts IDPs. The explanation to this results lie in the limitation of Netlogo , in taking into account only the visualized features. Information retrieved from literature suggest that 20% of IDPs were generated from the east side and 80 % from the west side.

6.1.4. Returnees :

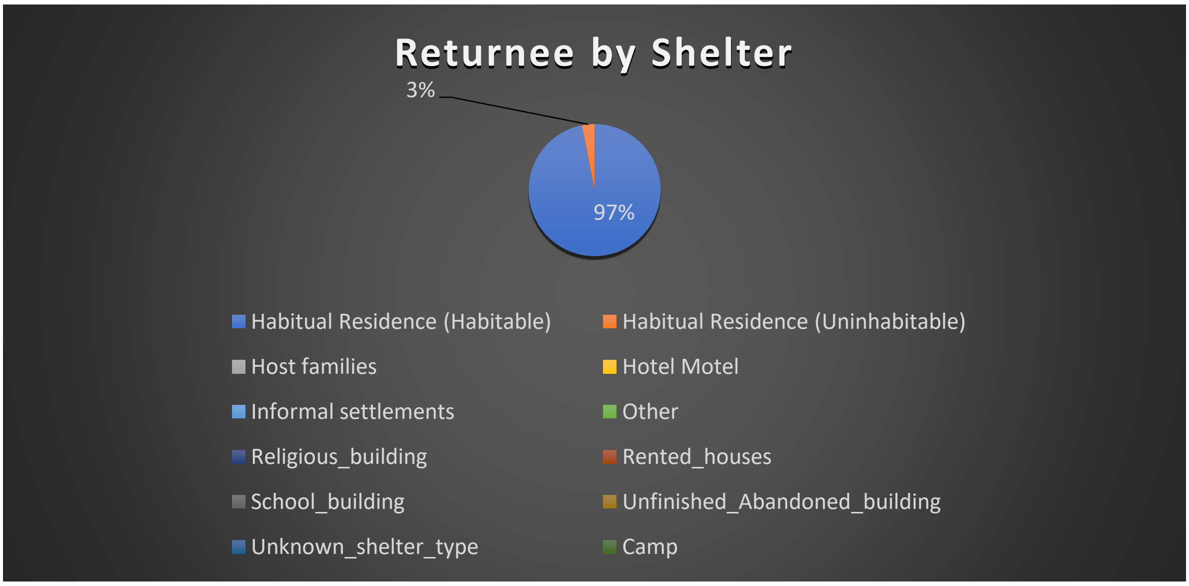


Figure 6.7 Returnee by shelter

Reports on returnees from IOM are not citing what the standards of Habitable Residence are. It can be either from safety perspective (cleared from explosives) or from an infrastructure perspective (having access to electricity and water). There is no data available that illustrates the percentage of returnees to camps, as the number is not recorded in the database, but the location is recorded.

6.2. Model Calibration

Two variables need to be calibrated before the final runs of the model area conducted. The first is a factor in the risk perception equation and the second are the number of damaged houses. The results of the analysis of the empirical data will be used to determine the correct values for these variables.

6.2.1. Factor in risk perception equation

A variable is used in the social network part of the risk perception equation. A series of runs was performed to test the impact of the increase of the number of deaths in the social network. Initially, the increase was set to 0.01 in equation 4.1The "tobl" factor was introduced to vary this synthetic value and 10 runs where performed with the values 0.01 , 0.5, 0.73 and 1. Figure 6.8 shows the impact of these values on the risk perception.

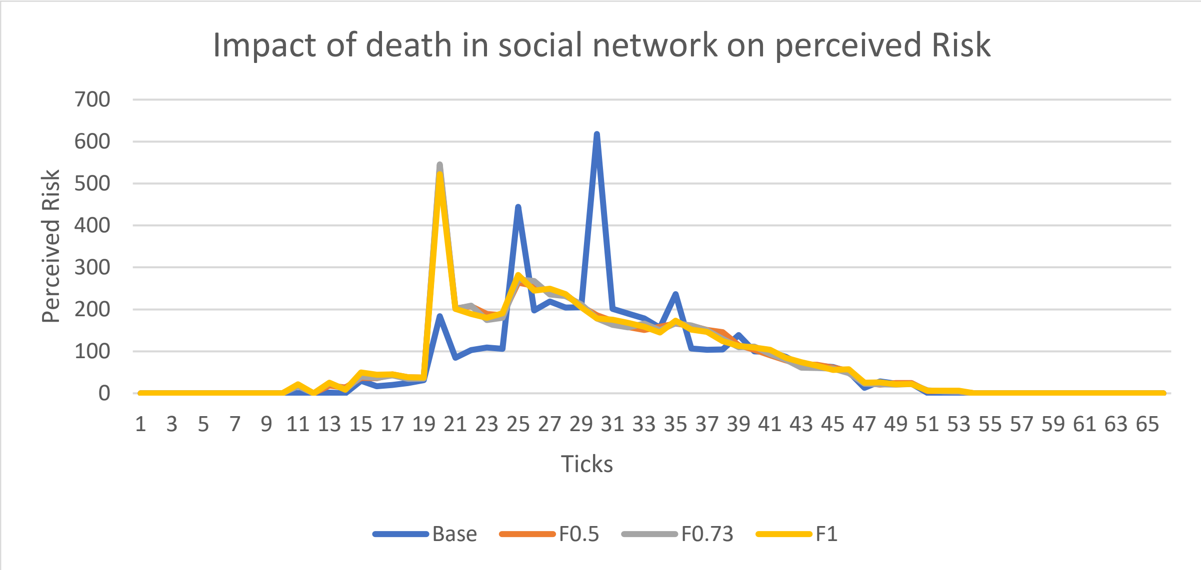


Figure 6.8 Sensitivity Analysis on Death in Social Network

While the base line for the value 0.01 shows four different peaks, the 0.5, 0.73 and 1 lines show an early major peak in tick 20 and a smoother curve. This suggests that the death in the social network impacts the risk perception. The death in the social network also increased the number of IDPs by 7,25 %. However, this is not the only change. The experiment has considerably impacted the Returnees curve, as it scored an increase of 26.18 % in the number of Returnees. The experiment has also revealed an impact on the behaviour of returnees as they moved back to Mosul in tick 30 compared to tick 55 in the base situation. When the patterns are compared to the values from the empirical data, we see that we should generate 113,424 IDPs by September 2017 and the number of IDPs should decline starting in December 2017. In addition, the number of returnees in September 2016 should be 170056.

<u>Risk perception and number of deaths :</u>

When comparing the perceived risk graph (figure 6.8) with the deaths in Mosul (figure 4.7), the results show that the perceived risk graph (base version) has three peaks. The first happens at tick 20 when the number of deaths in the east of Mosul start to increase. The second peak happens at tick 24 when the number of deaths in the east of Mosul reaches its peak and in parallel, the number of deaths in the west of Mosul start to increase. The third peak happens at tick 29 when both curves intersect. This suggests that in the model, the number of death contributes to the risk perception.

When the value of the factor is changed to 0.5, 0.73 and 1, these peaks disappear, leading to a more gradual risk perception. There is still a second peak around tick 24 (new restricted area loaded) but the later part of the curve only shows a decline.

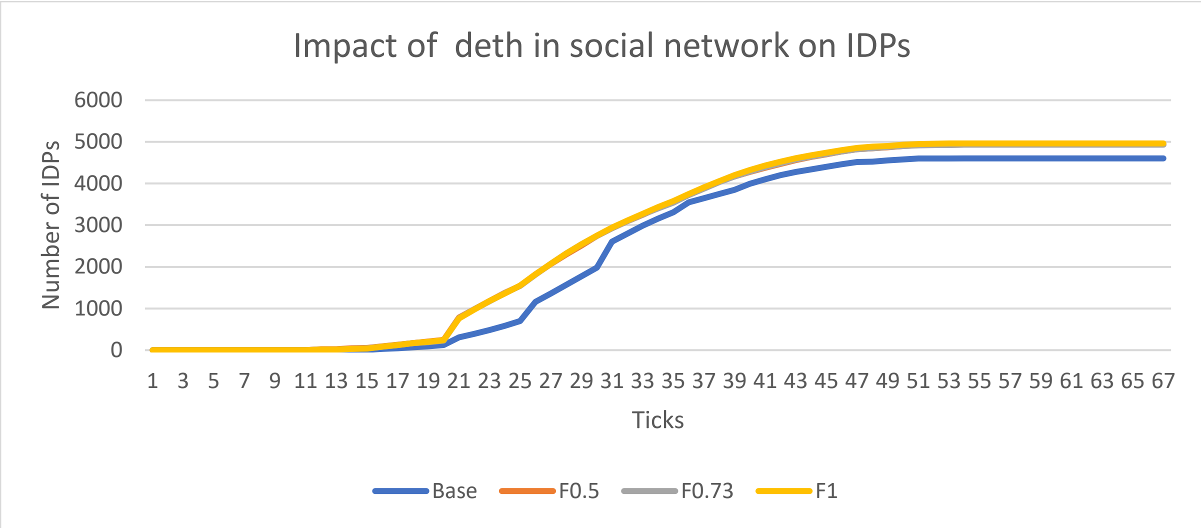


Figure 6.9 Impact of death in social network on generated IDPs

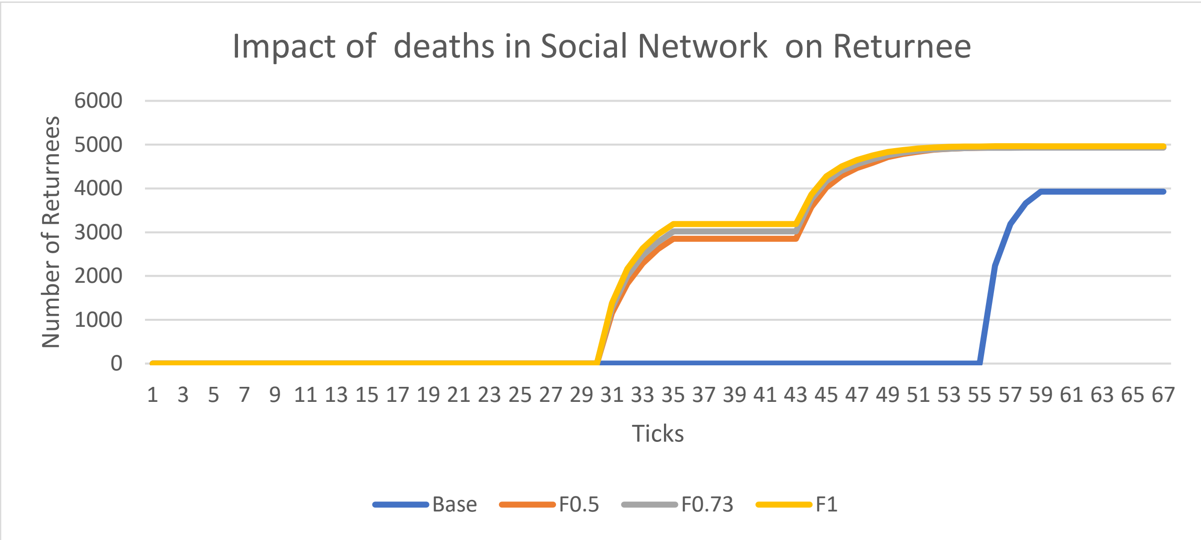


Figure 6.10 impact of death in social network on generated Returnees

The model will therefore use a value between 0.5 and 1 for the factor because it impact significantly the perceived risk , in our case 0.73 as it seem to not impact the number of generated IDPs but slightly impact the number of generated Returnees.

6.2.2. Calibration Damaged houses

The number of damaged buildings was initially set to 500. To test the impact of damaged houses on the generated number of Returnees, two experiments were performed using Behavior Space. The first 10 runs testing the impact of 250 houses damaged; the second 10 runs were testing the impact of 900 houses damaged out of 966 houses. No empirical data is available on the number of houses. The variable can only be calibrated based on the number and timing of the returnees.

The generated number of IDPs for the first 10 runs resulted in an average 3937.3 Returnees and 3920.6 Returnees for the remaining runs. Against 3933.5 in the initial simulation. This indicates that the number of damage houses has no impact on the number of Returnees. This is unexpected as agents are only able to return when their house remained undamaged.

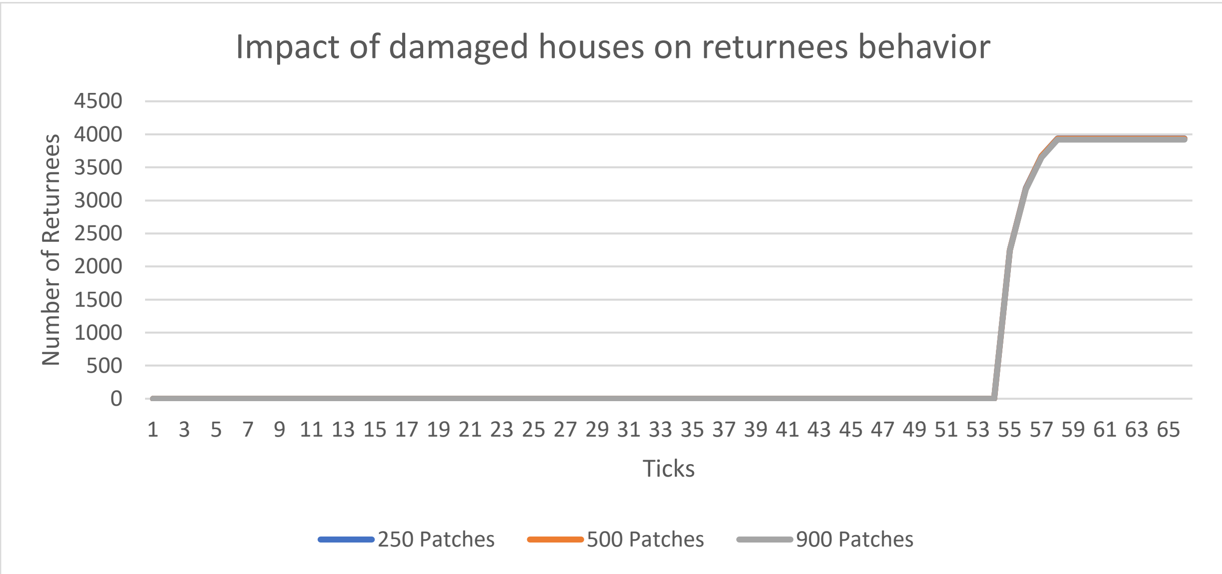


Figure 6.11 impact of damaged houses on returnees behavior

The model analysis shows that the number of damaged patches does not impact the number of returnees . The model calibration will therefore be set to 500 damaged houses. In the real situation Figure 6.6 shows that 97 % of IDPs returned to Habitable residency. The Model analysis cannot validate the result of damaged houses, this can be explained by the fact that the damage was randomly assigned and did not use real data instead.

6.3. Model Validation

In order to show the temporal dynamic of the model, five main graphs were plotted in the Netlogo interface. This section will explain each graph and it's results:

6.3.1 Total number of households

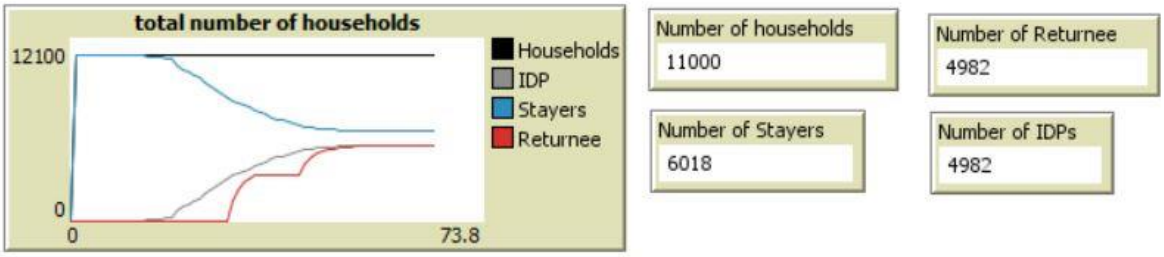


Figure 6.12 total number of households with their status

Figure 6.12 shows the three types of status of the households. While the IDP curve starts to increase around tick 15, the Stayer curve decreases at the same time and both curves become constant around tick 46. On the other hand, the returnee curve starts to increase around tick 53 and shortly after becomes constant at tick 57. The model has generated 45 % IDPs, 54 % stayers Returnees have the same percentage of IDPs

Section 6.1 has looked at displacement in Mosul from April 2014 to December 2019, the model simulation runs only for 65 weeks between September 2016 and November 2017. Figure 6.13 and 6.14 capture the concerned period of simulation.

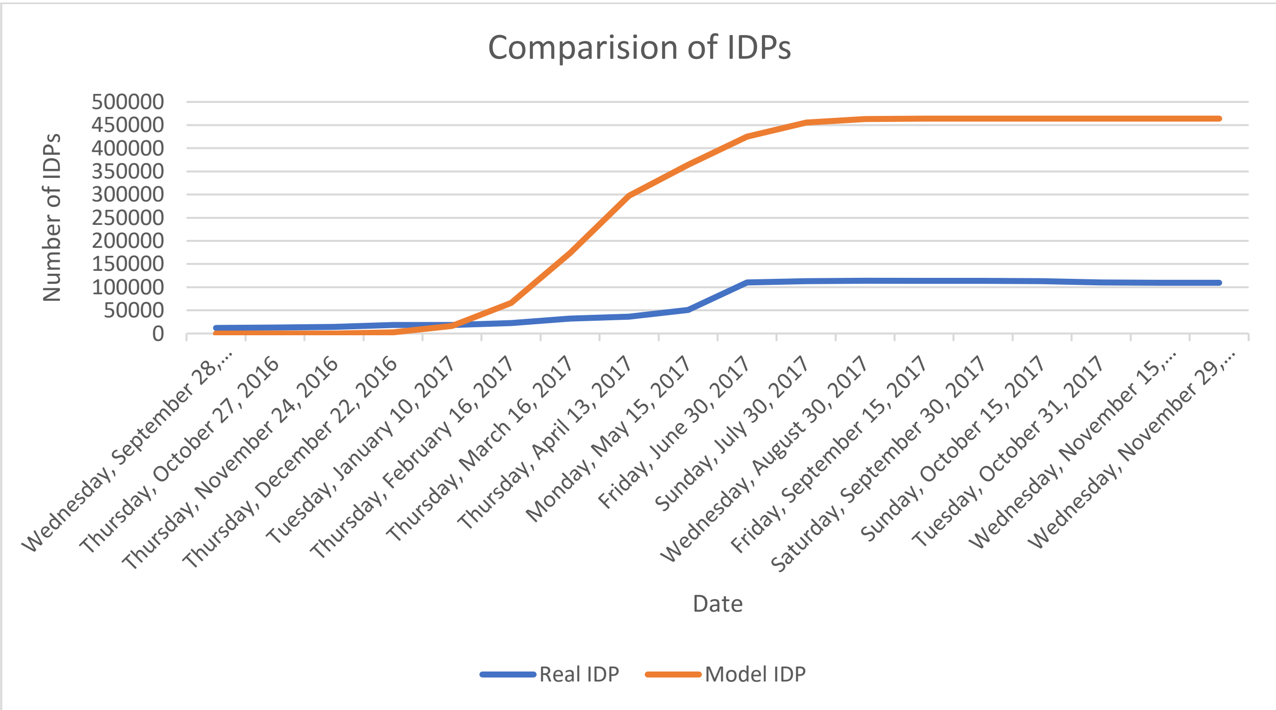


Figure 6.13 comparison between real number of IDP and generated number

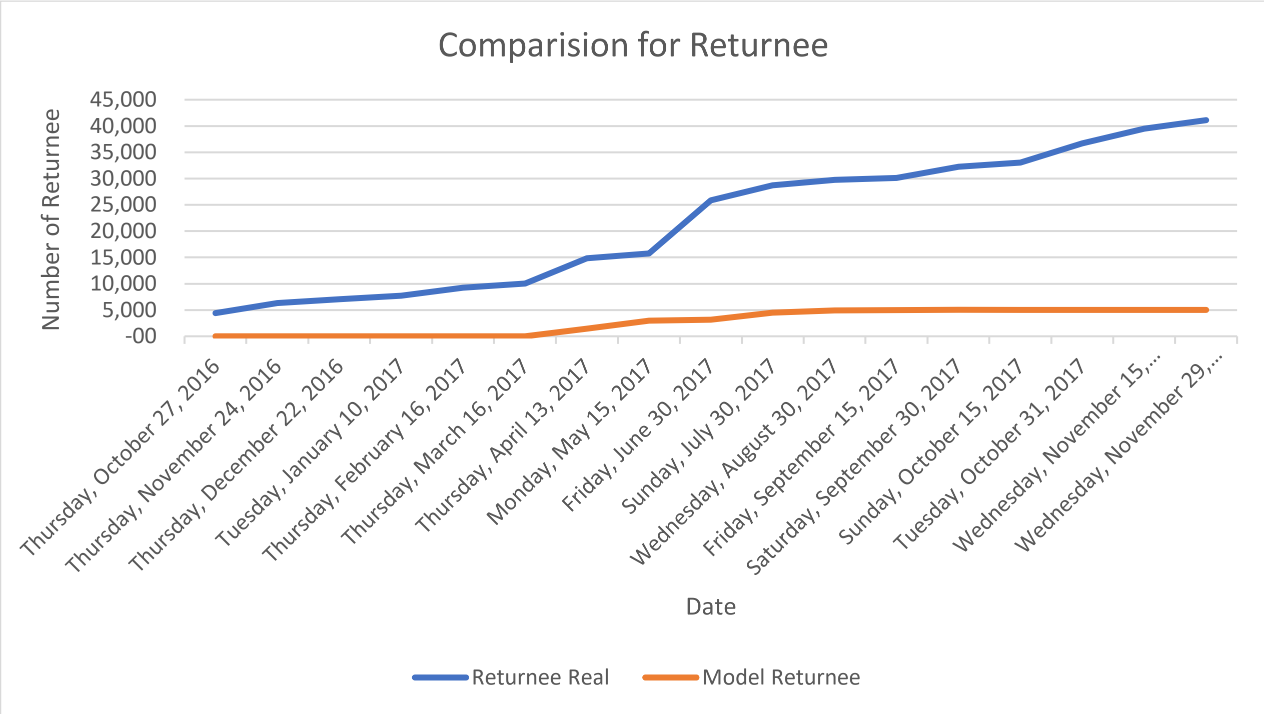


Figure 6.14 comparison between real number of Returnees and generated number

Both figures compares the real situation extracted from section 6.1 with the model output, they represent respectively the curves for IDPs and Returnees. For the simulated data, Figure 6.13 shows an increase in IDPs in December 2016, compared to March 2017 in the empirical data.

The generated IDP curve predict an increase in IDPs starting from December 2016, this can be explained by the increase number of deaths in east Mosul. On the other hand figure 6.14 compares the returnees, the model output predicts a return in March 2017, compared to empirical where the curve increases continuously, the results is explained by the fact that, the Mosul model does not take into account surrounding villages of Mosul where liberation happened in an early stage.

The analysis of the empirical data was not informative on the percentage of IDPs generated during the Mosul battle. However literature showed conflicting numbers. According UNOCHA (2018) the percentage of IDPs generated during the military intervention was 32% On the other hand, data retrieved from IOM (2017) suggest that the percentage was 84 %. The ABM model has generated 42% of IDPs, the comparison remains uncertain giving the fact that the scale of population is different, since the ABM model was simulating households, were available data was about individuals. At this point of the research, there was no available data on the percentage of Returnees during the battle.

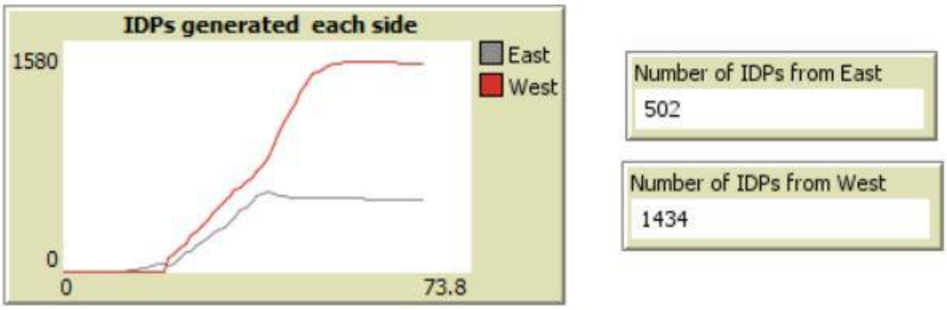


Figure 6.15 generated IDPs for each side of Mosul

Figure 6.15 shows an increasing number of IDPs from the east side of Mosul starting from tick 14. The curve reaches its first peak at tick 28 and it's second peak at tick 33. On the other hand, the west curve shows an increase at tick 18 and reaches its peak at tick 50, this result is due to the fact that the deaths in Mosul west happened later. The model has generated 2% of IDPs from the east side of Mosul and 18 % from the west side of Mosul.

6.3.2. Destination curve

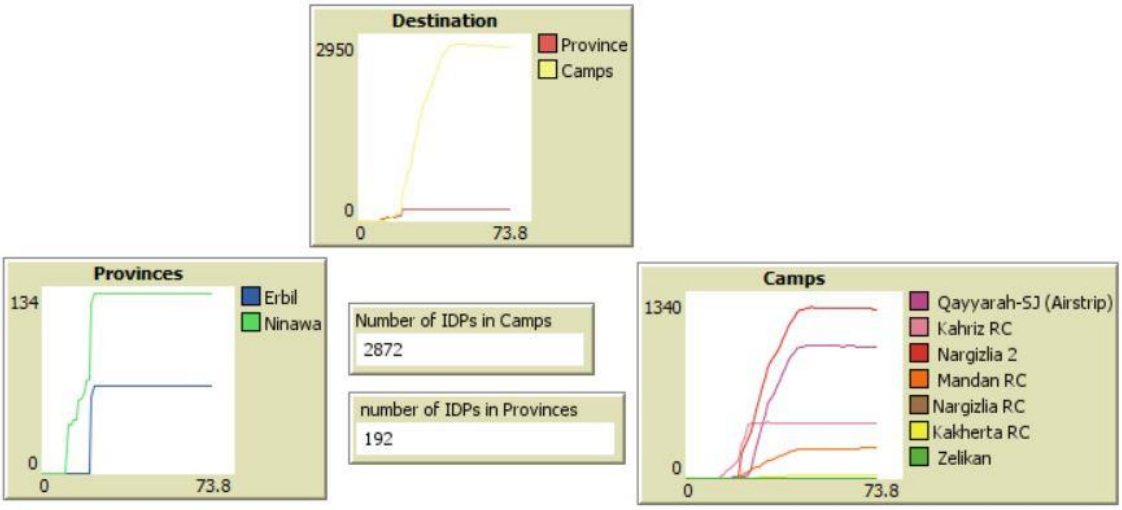


Figure 6.16 Destination graphs

The destination graph was split into two graphs, the province graph shows a higher number of IDPs going to Ninawa province and Erbil. The graphs also shows that IDP movement to these two provinces happened at the early stage of the simulation around tick 1. The model generated 3% of IDPs choosing Provinces as destination. The Second graph shows camps destination with the Nargizilia 2 camp scoring the highest number of IDPs. Unlike the provinces, the movement of IDPs to camps happened later in the simulation around tick 14 with Kahriz RC being the first chosen Camp. The model generated 75 % of IDPs choosing camps as their destination.

The model shows the camp “Kahriz RC” as being the first destination among the camps shelter group. The result is validated by checking the actual geographic locations of the refugee camps. However, movement to camps located in the south hasn’t emerged as a pattern in the generated model. As explained in the verification chapter, Netlogo takes into account only the visualized features. Camps located in the south where not visible on the Netlogo interface. However IDPs movement to provinces emerges instantly when the simulation is running. It is followed by the movement to camps see figure 6.16

Figure 6.16 compares the generated number of IDPs in camps to the real data, from the military intervention period. Model outputs predict an increase in the number of IDPs in Camps in January 2017. The Model predicts a higher number of IDPs compared to the real situation. Figure 6.17 compare the Number of IDPs in Host families, the model predicts an increase in January 2017, compared the empirical data where IDPs seem to have moved to host families before the intervention. This result can be explained by the fact that displacement in Iraq existed before the military intervention (see section 2.1.1)

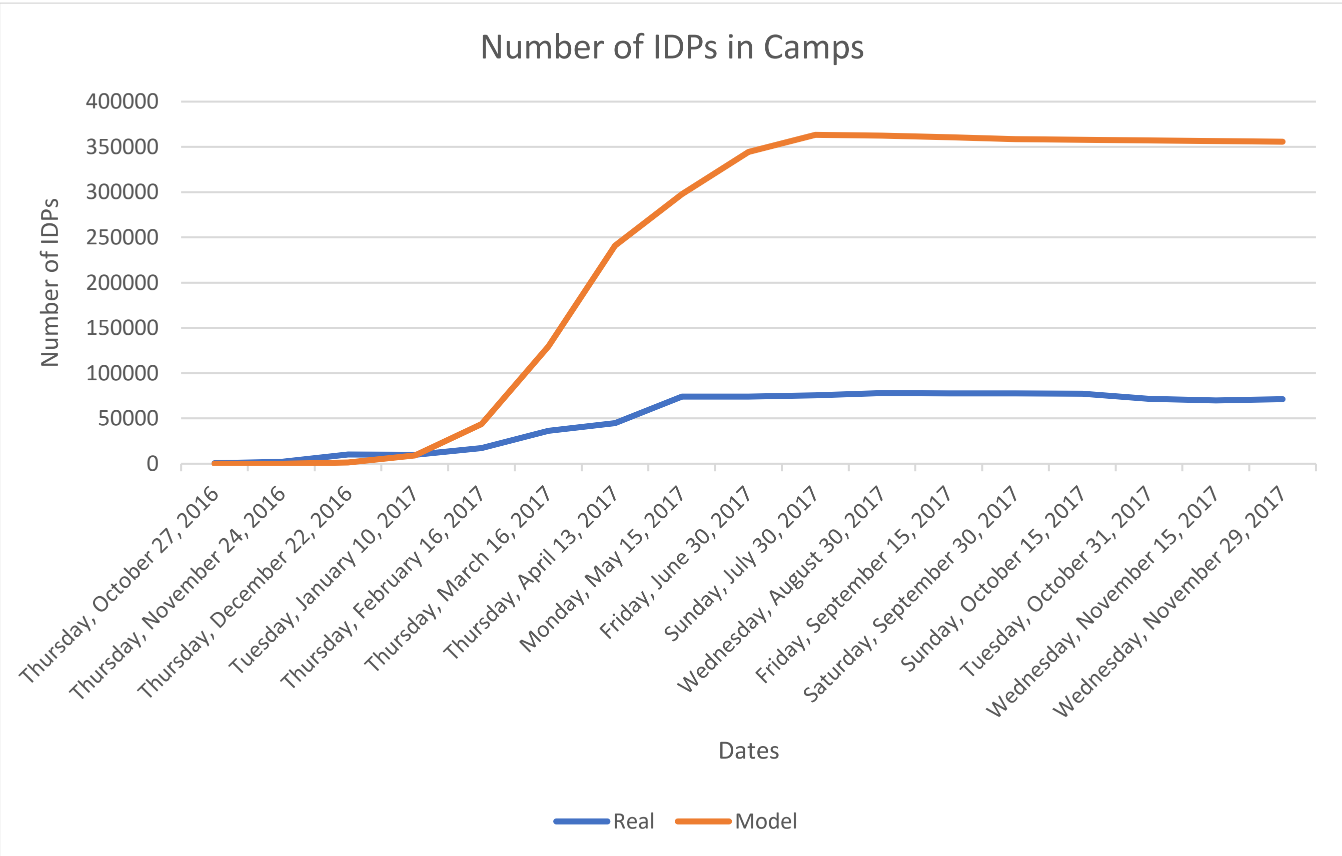


Figure 6.17 comparison between real number of IDPs in camps and generated number

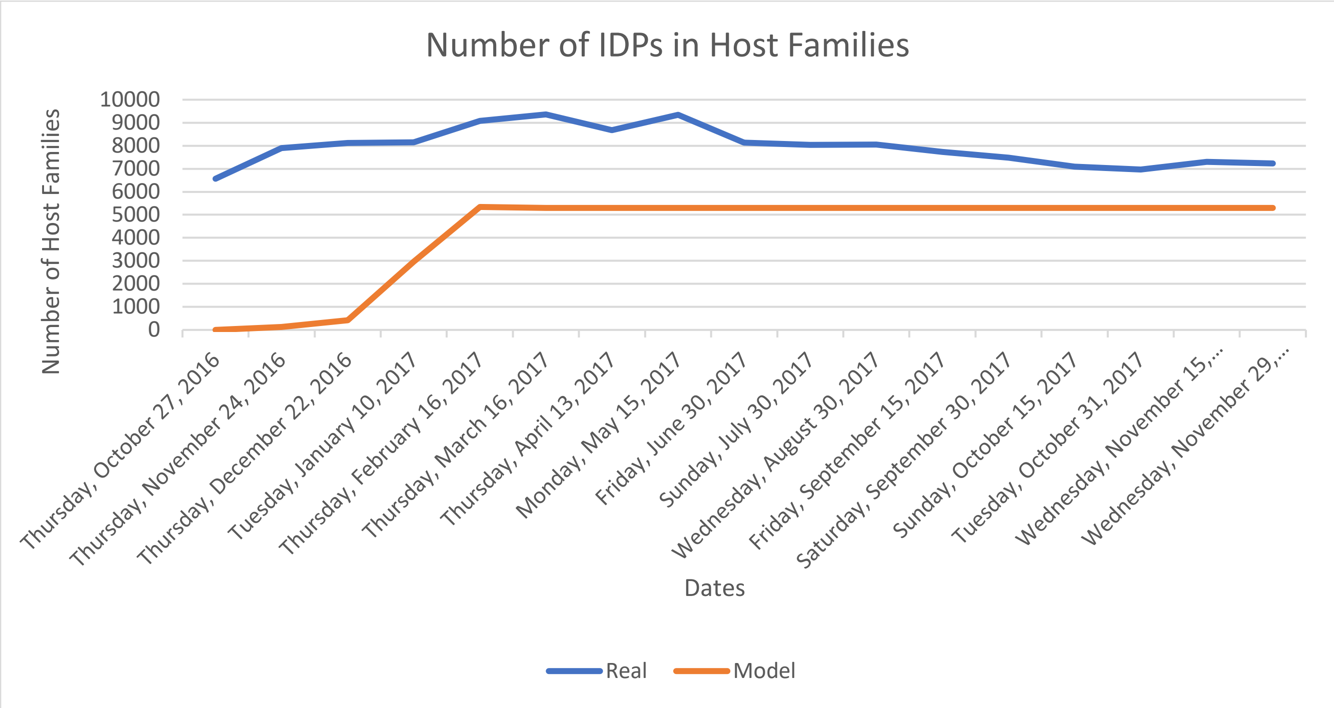


Figure 6.18 comparison between real number of Returnees in camps and generated number

7. CONCLUSION

In the course of this study a model was developed to predict the flow of IDPs and Returnees in the case of conflict induced displacement. The military intervention in Mosul was used as a case study.

7.1. Answer to research questions :

7.1.1. Human behaviour based on Risk perception

RQ1: What are the factors that influence the decision of a person to leave or stay before and during a military operation?

This question was answered via literature review. Van der Verld (2011) theorized these factors into three distinct threshold. The indifference threshold where the household has assessed the risk and decided to leave. The trajectory threshold where household asses their means and the safety of the journey. The locational threshold where households select their destination.

RQ2 : Is there an existing (agent-based) model or theory using risk perception that can be used to develop the simulation model?

The model of Hébert et al. (2018) was identified as a model that includes risk perception. The model used death toll to quantify the general risk of Syrian refugees. However, this model does not include restricted areas and assumes all refugees are free to leave the area.

RQ3: How can the behavior of IDPs be simulated in an agent-based model?

The simulation of IDPs behavior was conceptualised using Protection Motivation Theory, where death in social network and in the sub-area of Mosul (east – west) were identified as the risk perception factors and income level and social network were identified as factors that influenced the coping strategies.
The behavior was implemented in four consecutive steps. First households compare their own risk perception to the general risk in Mosul; second households assess their income level and their own location. Third households change their status to "idp" if the criteria are met, if not the household status remains "stayer". Finally households asses their income level and social network for the destination choice. IDPs move to Provinces in the situation where they have high income level or host families, the remaining IDPs assess the nearest camp, in case the camp reached its full capacity the IDPs move to the nearest other camp till it finds a place. IDPs that do not find a destination change their status to Stayer. Van der Verld (2011) theorized this four step into three distinct threshold. The first and third implemented steps

are identified as the indifference threshold where the household has assessed the risk and decided to leave. The second and fouth implemented step are identified as the locational threshold where households asses their means and select their destination. According to Van der Verld, the trajectory threshold , this last step was not implemented as our model does not take into account the journey that households are taking.

7.1.2. Inverse flow of displacement

RQ1: Which factors influence the return of IDPs to their home location?

The model used perceived risk and death in the social network as factors to determine the return of households. Two new elements were introduced: a delay in the risk perception and damaged houses. It is assumed that risk perception is not primarily determined by the perceived risk, but that the risk trend is equally important. In case the risk is low but it increases, this will not lead to a return to the home location. The opposite is also valid, a higher risk perception with a decreasing trend may lead to a return. In order to be able to return, the house of the IDP needs to be in good condition. Therefore, damage to houses is used as a limiting factor.

RQ2 : How can this inverse flow of refugees be modelled?

The inverse flow of IDPs (returnees) was implemented using three consecutive step. First, IDPs asses the risk in Mosul by comparing the number of deaths of the previous two weeks and asses the deaths in their social network. In the second step, IDP change their status to Returnee when they judge that the risk is low. Third, Returnees asses the condition of their residency, in case of no damage they move back to their home.

7.1.3. Model Validation

RQ1 How can Pattern-Oriented Modelling be used to validate the model?

The following patterns have been identified that help to compare the simulation output with the real data: the temporal patterns of IDPs and returnees, the choice of shelter location (camps versus host families). The choice of shelter differs over time and for different income levels. In addition to these patterns, results are evaluated for the eastern and western parts of Mosul. This helps to evaluate the impact of differences in deaths and time of liberation.

RQ2 Which patterns of movement are present in the population data for the Mosul area?

There are some important patterns present in the empirical data in both time and space. Both IDP and Returnee movement starts before the simulation time. This was unexpected, especially the fact that there was return behaviour before the end of the conflict. It is assumed that returnees return to the outskirts of Mosul when the liberation of the centre has not been completed. Different income levels seem to have different movement behaviours. An example of this is that high income households move to host families early in the conflict. Movement of population started in east then in west as the fighting happened first in east, the west on the other hand generated more IDPs.

RQ3 Can these patterns be re-generated using the simulation model?

There was considerable discrepancy between the real data and the simulated patterns. The number of IDPs and Returnees was difficult to match with the real data. Returnee movement in the simulation started later then in reality, the number of returnee was equal to the number of IDP , in reality it was not the case. The camps surrounding Mosul did also not fill up, only two camps were selected. The model generated IDPs only in two provinces Ninewa and Erbil, in reality IDP distribution was extended to more than two provinces

7.2. Research limitation

As all research methodologies have limitations, no exception is reserved for simulation modelling. This section focuses on the shortcomings of this study research and classifies them into two main sections. First data limitation will be discussed, followed by modelling limitation.

7.2.1. Data Limitations

Information on Returnee percentage among Mosul population was not found. In fact, available data is aggregated and takes into account surrounding villages of Mosul. The limitation of data concerns also the number of social contacts of the Mosul population which the model used as a coping strategy factor. A concrete number would have generated a truthful number of IDPs. Sensitivity analysis has demonstrated that the number of damaged home location did not impact the number of Returnees. Empirical data (see chapter 3) on building damage on the other hand, identified the condition of residence as an important factor for returnee behaviour. At this stage of the research, available data on damaged building is limited to the city centre. Available data covering the City would have generated a different number of returnees. Furthermore, the model is limited to two coping variables. in reality social factors such as ethnicity has played a role in the destination choice. Available data on ethnicity would have impacted the destination choice of IDPs.

7.2.2. Modelling Limitations

The results of returnee are also identified as a modelling limitation. The model took into account only the city of Mosul but the empirical data suggests that the military intervention started in the surrounding areas of Mosul. The pattern of returnees identified in the empirical data illustrates these limitations as people returned earlier than what the model have predicted, suggesting that returnees are also populations displaced in the early stage of the intervention, coming back once their area was liberated. This can easily be adjusted by creating more spatial areas (besides the east and west of Mosul) and loading death data specific to these areas.

The destination choice was designed based on factor that push IDPs to leave their location. The model seems to underestimate the provinces choice. Adjusting the model to take into account the destination attractiveness would have generated a larger number of IDPs in the provinces.

7.3. Model reproduction

The implemented model has succeeded in generating IDPs and Returnees regarding the limitations mentioned in the previous section. The model is able to predict the number of generated IDPs based on the geographical information of the fight. The model has predicted a higher number of generated IDPs in the western side. On the other hand, the destination to camps uses camps capacity and nearest distance, predicting the camps that are mostly concerned by the flow of IDPs, and the ones that are the primary destination choice.

LIST OF REFERENCES

Hébert, G. A. (2018). An Agent-Based Model to Identify Migration Pathways of Refugees: The Case of Syria. Advances in Geographic Information Science, 1–14. https://doi.org/10.1007/978-3-319-65993-0

Abdulkareem, S. A., Augustijn, E.-W., Filatova, T., Musial, K., & Mustafa, Y. T. (2020). Risk perception and behavioral change during epidemics: Comparing models of individual and collective learning. PLOS ONE, 15(1), e0226483. https://doi.org/10.1371/journal.pone.0226483

Abdulkareem, S. A., Augustijn, E.-W., Mustafa, Y. T., & Filatova, T. (2018). Intelligent judgements over health risks in a spatial agent-based model. International Journal of Health Geographics, 17(1), 1–19. https://doi.org/10.1186/s12942-018-0128-x

Abdulkareem, S., Augustin, E., Mustafa, Y. T., & Filatova, T. (2017). Integrating Spatial Intelligence for risk perception in an Agent Based Disease Model. GeoComputation, 1. Retrieved from https://pdfs.semanticscholar.org/a8ca/d9bbc7648c74622bc385c1c14c01e590e498.pdf

Adhikari, P. (2012). Conflict-Induced Displacement, Understanding the Causes of Flight. American Journal of Political Science, 57(1), 82–89. https://doi.org/10.1111/j.1540-5907.2012.00598.x

Ajezen, I. (2005). Symmetry, compatibility, and attitude- behavior correspondence. Multivariate Research Strategies. Researchgate, 3–19. Retrieved from https://www.researchgate.net/publication/264156930_Laws_of_human_behavior_Symmetry_compatibility_and_attitude-behavior_correspondence

BROWN, D., PAGE, S., RIOLO, R., & RAND, W. (2004). Agent-based and analytical modeling to evaluate the effectiveness of greenbelts. Environmental Modelling & Software, 19(12), 1097–1109. https://doi.org/10.1016/j.envsoft.2003.11.012

Bubic. (2010). Prediction, cognition and the brain. Frontiers in Human Neuroscience, 1. https://doi.org/10.3389/fnhum.2010.00025

Burton, & Kates. (1964). The perception of Natural Hazards in Resource Management. Natural Resources JournaL, 3, 412–441. Retrieved from https://www.jstor.org/stable/pdf/24879760.pdf?refreqid=excelsior%3A6340bfae94c914f22ba3b2acfab0f17d

Ceriani, L., & Verme, P. (2018). Risk Preferences and the Decision to Flee Conflict. Policy Research Working Papers, 1–41. https://doi.org/10.1596/1813-9450-8376

Chatelard, G. (2008). A quest for family protection: the fragmented social organisation of transnational Iraqi migration. HAL, 1–16. Retrieved from https://hal.archives-ouvertes.fr/hal-00338406/document

Crawley, H., & Hagen-Zanker, J. (2019). Deciding Where to go: Policies, People and Perceptions Shaping Destination Preferences. International Migration, 1–18. https://doi.org/10.1111/imig.12537

Davenport, C., Moore, W., & Poe, S. (2003). Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964-1989. International Interactions, 29(1), 27–55. https://doi.org/10.1080/03050620304597

Edwards, S. (2003). Computational Tools in Predicting and Assessing Forced Migration. Journal of Refugee Studies, 21(3), 347–359. https://doi.org/10.1093/jrs/fen024

Fernandez, L. E., Brown, D. G., Marans, R. W., & Nassauer, J. I. (2005). Characterizing Location Preferences in an Exurban Population: Implications for Agent-Based Modeling. Environment and Planning B: Planning and Design, 32(6), 799–820. https://doi.org/10.1068/b3071

Fernandez, R. J. (2019). A Theoretical Model of the Individual-Decision Making Processes for Refugees. ViXra, 1. Retrieved from https://www.semanticscholar.org/paper/A-Theoretical-Model-of-the-Individual-Decision-for-Fernandez/a34dd4093a3d82f50cd9b122c898bea7d6242f1c

Ferrer, R. A., & Klein, W. M. (2015). Risk perceptions and health behavior. Current Opinion in Psychology, 5, 85–89. https://doi.org/10.1016/j.copsyc.2015.03.012

Fischhoff, B., Gonzalez, R. M., Small, D. A., & Lerner, J. S. (2003). Judged Terror Risk and Proximity to the World Trade Center. The Journal of Risk and Uncertainty , 26, 1. Retrieved from https://link.springer.com/article/10.1023/A:1024163023174

FitzGerald, D. S., & Arar, R. (2018). The Sociology of Refugee Migration. Annual Review of Sociology, 44(1), 387–406. https://doi.org/10.1146/annurev-soc-073117-041204

Grimm, V. (2005). Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. Science, 310(5750), 987–991. https://doi.org/10.1126/science.1116681

Grimm, Volker, & Railsback, S. F. (2012). Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1586), 298–310. https://doi.org/10.1098/rstb.2011.0180

Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (Eds.). (2012). Agent-Based Models of Geographical Systems. Springer Link, 481–497. https://doi.org/10.1007/978-90-481-8927-4

Hilljegerdes, M., & Augustijn-Beckers, E.-W. (2018). Evaluating the effects of consecutive hurricane hits on evacuation patterns in Dominica. Semantic Scholar, 1–11. Retrieved from https://pdfs.semanticscholar.org/a84c/6dfddfb959b253b17fc76b48c171148b9bf7.pdf

Home Page - IRAQ DTM. (2014). Retrieved June 17, 2020, from http://iraqdtm.iom.int/

Ibáñez, A. M. (2006). The Impact of Intra-State Conflict on Economic Welfare and Consumption Smoothing: Empirical Evidence for the Displaced Population in Colombia. HiCN Working Papers, 1–20. Retrieved from https://ideas.repec.org/p/hic/wpaper/23.html

ICRC. (2016). Pprotracted conflict and Hhumanitarian action some recent icrc experiences development holds HUMANITARIAN ACTORS PRESENT. Retrieved from https://www.icrc.org/sites/default/files/document/file_list/protracted_conflict_and_humanitarian_action_icrc_report_lr_29.08.16.pdf

ICRC. (2019). Displacement in times of armed conflict: How international humanitarian law protects in war, and why it matters. Retrieved from https://www.icrc.org/en/publication/displacement-times-armed-conflict-how-international-humanitarian-law-protects-war-and

IOM. (2017). Mosul Crisis. Retrieved June 17, 2020, from https://utwente.maps.arcgis.com/home/signin.html?returnUrl=https%3A//utwente.maps.arcgis.com/apps/Cascade/index.html%3Fappid%3D4b952893b73146c9983e21f46c213de6

IOM. (2018). Reasons to Remain: Categorizing Protracted Displacement in Iraq. Retrieved from https://reliefweb.int/report/iraq/reasons-remain-categorizing-protracted-displacement-iraq-november

Klabunde, A., & Willekens, F. (2016). Decision-Making in Agent-Based Models of Migration: State of the Art and Challenges. European Journal of Population, 32(1), 73–97. https://doi.org/10.1007/s10680-015-9362-0

Kniveton, D. R., Smith, C. D., & Black, R. (2012). Emerging migration flows in a changing climate in dryland Africa. Nature Climate Change, 2(6), 444–447. https://doi.org/10.1038/nclimate1447

Kniveton, D., Smith, C., & Wood, S. (2011). Agent-based model simulations of future changes in migration flows for Burkina Faso. Global Environmental Change, 21, S34–S40. https://doi.org/10.1016/j.gloenvcha.2011.09.006

Lafta, R., Al-Nuaimi, M. A., & Burnham, G. (2018). Injury and death during the ISIS occupation of Mosul and its liberation: Results from a 40-cluster household survey. PLOS Medicine, 15(5), e1002567. https://doi.org/10.1371/journal.pmed.1002567

Lischer, S. K. (2007). Causes and Consequences of Conflict-Induced Displacement. Civil Wars, 9(2), 142–155. https://doi.org/10.1080/13698240701207302

Lozano-Gracia, N., Piras, G., Ibáñez, A. M., & Hewings, G. J. D. (2010). The Journey to Safety: Conflict-Driven Migration Flows in Colombia. International Regional Science Review, 33(2), 157–180. https://doi.org/10.1177/0160017609336998

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. Journal of Simulation, 4(3), 151–162. https://doi.org/10.1057/jos.2010.3

Maderthaner, R., Guttmann, G., Swaton, E., & Otway, H. J. (1978). Effect of distance upon risk perception. Journal of Applied Psychology, 63(3), 380–382. https://doi.org/10.1037/0021-9010.63.3.380

Mallett, R., & Hagen-Zanker, J. (2018). Forced migration trajectories: an analysis of journey- and decision-making among Eritrean and Syrian arrivals to Europe. Migration and Development, 7(3), 341–351. https://doi.org/10.1080/21632324.2018.1459244

Mekdjian, S. (2018). Urban artivism and migrations. Disrupting spatial and political segregation of migrants in European cities. Cities, 77, 39–48. https://doi.org/10.1016/j.cities.2017.05.008

Melander, E., & Öberg, M. (2007). The Threat of Violence and Forced Migration: Geographical Scope Trumps Intensity of Fighting. Civil Wars, 9(2), 156–173. https://doi.org/10.1080/13698240701207310

Melegaro, A., Del Fava, E., Poletti, P., Merler, S., Nyamukapa, C., Williams, J., … Manfredi, P. (2017). Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. PLOS ONE, 12(1), e0170459. https://doi.org/10.1371/journal.pone.0170459

Miles, A. (2003). Perceived Severity . Retrieved May 15, 2020, from https://cancercontrol.cancer.gov/brp/research/constructs/perceived_severity.html

Nowak, S. A., Matthews, L. J., & Parker, A. M. (2017). A General Agent-Based Model of Social Learning. Rand Health Q, 1. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5644776/

OECD/EASO. (2018). Can we anticipate future migration flows? Migration Policy Debates . Retrieved from https://www.oecd.org/els/mig/migration-policy-debate-16.pdf

Pellegrini, P. A., & Fotheringham, A. S. (2002). Modelling spatial choice: a review and synthesis in a migration context. Progress in Human Geography, 26(4), 487–510. https://doi.org/10.1191/0309132502ph382ra

Peters, K., Mayhew, L., Slim, H., Van Aalst, M., & Arrighi, J. (2019). Double vulnerability The humanitarian implications of intersecting climate and conflict risk. Retrieved January 2020, from https://www.odi.org/sites/odi.org.uk/files/resource-documents/12647.pdf

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. Journal of Deaf Studies and Deaf Education, 18(3), 427. https://doi.org/10.1093/deafed/ent031

Saldarriaga, J. F., & Hua, Y. (2019). A gravity model analysis of forced displacement in Colombia. Cities, 95, 102407. https://doi.org/10.1016/j.cities.2019.102407

Schneiderbauer, S., & Ehrlich, D. (2004). Risk, hazard and people’s vulnerability to natural hazards: A review of definitions, concepts and data. European Commission Joint Research Centre, 167–199. https://doi.org/10.1007/978-3-540-75162-5_7

Slovic, P. (1996). Perception of Risk From Radiation. Retrieved March 2020, from https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/22426/slovic_367.pdf?sequence=1

Sokolowski, J. A. (2014). A Methodology for Environment and Agent Development to Model Population Displacement people in a protracted situation conflict or persecution refugees spread throughout 70 individuals applied for asylum. ACM, 1–11. Retrieved from https://dl.acm.org/doi/pdf/10.5555/2665049.2665052?download=true

Sokolowski, John A., Banks, C. M., & Hayes, R. L. (2015). Modeling population displacement in the Syrian city of Aleppo. Proceedings of the Winter Simulation Conference 2014, 1–20. https://doi.org/10.1109/wsc.2014.7019893

Steele, A. (2009). Seeking Safety: Avoiding Displacement and Choosing Destinations in Civil Wars. Journal of Peace Research, 46(3), 419–429. https://doi.org/10.1177/0022343309102660

UNHCR. (2018). Global trends displacement, trends at a glance. Retrieved from https://www.unhcr.org/5d08d7ee7.pdf

UNHCR. (2019). Challenges Upon Return In West Mosul. Human Appeal, 53, 1–30. https://doi.org/10.1017/cbo9781107415324.004

United Nations. (2019). Household : Size and Composition 2018 . Retrieved February 2020, from https://population.un.org/Household/index.html#/countries/368

UNOCHA. (2018, December 17). Iraq: Humanitarian Needs Overview 2019 (November 2018) - Iraq. Retrieved May 30, 2020, from https://reliefweb.int/report/iraq/iraq-humanitarian-needs-overview-2019-november-2018

Van der Velde, M., & van Naerssen, T. (2011). People, borders, trajectories: an approach to cross-border mobility and immobility in and to the European Union. Area, 43(2), 218–224. https://doi.org/10.1111/j.1475-4762.2010.00974.x

APPENDIX A

Full code for Data PreProcessing :

```
---------------------First----------------------------------------------
import os
from os import walk
import pandas as pd


path = r'C:\Users\Sarah\Desktop\IOMReturnee'
my_files = []
for (dirpath, dirnames, filenames) in walk(path):
    my_files.extend([os.path.join(dirpath, fname) for fname in filenames])


all_sheets = []
for file_name in my_files:

    #Display sheets names using pandas
    pd.set_option('display.width',300)
    mosul_file = file_name
    xl = pd.ExcelFile(mosul_file)
    mosul_df = xl.parse(0, header=[1], index_col=[0,1,2])

    #Read Excel and Select columns

    mosul_file = pd.read_excel(file_name, sheet_name = 0 ,
    index_clo=None, na_values= ['NA'], usecols = "A, E, G, H , L , M" )

    #Remove NaN values

    data_mosul_df = mosul_file.apply (pd.to_numeric, errors='coerce')
    data_mosul_df = mosul_file.dropna()

    #Make a list of df's
    all_sheets.append(data_mosul_df)


def save_frames(frames, output_path):
    frames.to_csv(output_path, mode='a', header=False)

if __name__ == '__main__':
   frames = pd.concat(all_sheets)
   save_frames(frames, r'C:\Users\Sarah\Desktop\DOReturnee\all.csv')
--------------------Second--------------------------------

# Setting directory and selecting file
rm(list=ls())
```

```
setwd("/Users/Sarah/Desktop") # Change directory
list.files()
getwd()
myData <- read.csv(file="2015.csv", sep=",", header=T, na.strings=c("NA","NaN", " ", ""), dec = ".") # change
name of file if necessary

# Taking first "table"
firsttable <- myData[, 1:4]

# Manually selecting second table
data2add1 <- myData[, 6:9]
# Renaming second table because rbind needs same colnames
colnames(data2add1) <- names(myData[, 1:4])

# Manually selecting third table
data2add2 <- myData[, 11:14]
# Renaming third table because rbind needs same colnames
colnames(data2add2) <- names(myData[, 1:4])

# Manually selecting third table
data2add3 <- myData[, 16:19]
# Renaming third table because rbind needs same colnames
colnames(data2add3) <- names(myData[, 1:4])

# Manually selecting third table
data2add4 <- myData[, 21:24]
# Renaming third table because rbind needs same colnames
colnames(data2add4) <- names(myData[, 1:4])

# Manually selecting third table
data2add5 <- myData[, 26:29]
# Renaming third table because rbind needs same colnames
colnames(data2add5) <- names(myData[, 1:4])

# Manually selecting third table
data2add6 <- myData[, 31:34]
# Renaming third table because rbind needs same colnames
colnames(data2add6) <- names(myData[, 1:4])

# Manually selecting third table
data2add7 <- myData[, 36:39]
# Renaming third table because rbind needs same colnames
colnames(data2add7) <- names(myData[, 1:4])

# Manually selecting third table
data2add8 <- myData[, 41:44]
# Renaming third table because rbind needs same colnames
colnames(data2add8) <- names(myData[, 1:4])
```

```
# Manually selecting third table
data2add9 <- myData[, 46:49]
# Renaming third table because rbind needs same colnames
colnames(data2add9) <- names(myData[, 1:4])

# Manually selecting third table
data2add10 <- myData[, 51:54]
# Renaming third table because rbind needs same colnames
colnames(data2add10) <- names(myData[, 1:4])

# Manually selecting third table
data2add11 <- myData[, 56:59]
# Renaming third table because rbind needs same colnames
colnames(data2add11) <- names(myData[, 1:4])

# Manually selecting third table
data2add12 <- myData[, 61:64]
# Renaming third table because rbind needs same colnames
colnames(data2add12) <- names(myData[, 1:4])

# Manually selecting third table
data2add13 <- myData[, 66:69]
# Renaming third table because rbind needs same colnames
colnames(data2add13) <- names(myData[, 1:4])

# Manually selecting third table
data2add14 <- myData[, 71:74]
# Renaming third table because rbind needs same colnames
colnames(data2add14) <- names(myData[, 1:4])

# Manually selecting third table
data2add15 <- myData[, 76:79]
# Renaming third table because rbind needs same colnames
colnames(data2add15) <- names(myData[, 1:4])

# Manually selecting third table
data2add16 <- myData[, 81:84]
# Renaming third table because rbind needs same colnames
colnames(data2add16) <- names(myData[, 1:4])

# Manually selecting third table
data2add17 <- myData[, 86:89]
# Renaming third table because rbind needs same colnames
colnames(data2add17) <- names(myData[, 1:4])

# Manually selecting third table
data2add18 <- myData[, 91:94]
```

```
# Renaming third table because rbind needs same colnames
colnames(data2add18) <- names(myData[, 1:4])

# Manually selecting third table
data2add19 <- myData[, 96:99]
# Renaming third table because rbind needs same colnames
colnames(data2add19) <- names(myData[, 1:4])

# Manually selecting third table
data2add20 <- myData[, 101:104]
# Renaming third table because rbind needs same colnames
colnames(data2add20) <- names(myData[, 1:4])

# Manually selecting third table
data2add21 <- myData[, 106:109]
# Renaming third table because rbind needs same colnames
colnames(data2add21) <- names(myData[, 1:4])

# Manually selecting third table
data2add22 <- myData[, 111:114]
# Renaming third table because rbind needs same colnames
colnames(data2add22) <- names(myData[, 1:4])

Year2015 <- rbind(firsttable, data2add1, data2add2, data2add3,data2add4, data2add5, data2add6, data2add7,
        data2add8, data2add9, data2add10, data2add11, data2add12, data2add13, data2add14, data2add15,
data2add16, data2add17, data2add18, data2add19, data2add20, data2add21, data2add22 )
# add other tables to list e.g. rbind(firsttable, data2add1, data2add2, data2add3)

# Saving file
write.csv(Year2015 ,".Year2015.csv", row.names = T)
```

```
---------------------Thrid----------------------------------
#drope duplicated values

df = pd.read_csv(r'C:\Users\Sarah\Desktop\tes\.Year2015.csv', usecols=['Place.ID', 'Location.Name',
'Latitude','Longitude' ]).drop_duplicates(keep='first').reset_index()
file_name = r'C:\Users\Sarah\Desktop\U2015.csv'
df.to_csv(file_name, index=False)
```

```
----------------------Fourth-------------------------------------
import pandas as pd
import glob as glob
import numpy as np


all_data = pd.DataFrame()
for f in glob.glob(r'C:\Users\Sarah\Desktop\IDPMosul\Data\*.xlsx'):
```

```
    df = pd.read_excel(f,index_col=None, na_values=['NA'])
    df['filename'] = f
    all_data = all_data.append(df,ignore_index=True)

# BUILD LIST OF DFs
df_list = [(pd.read_excel(f, index_col=None, na_values=['NA'])
            .assign(filename = f)
          ) for f in glob.glob(r'C:\Users\Sarah\Desktop\IDPMosul\Data\*.xlsx')]

# CONCATENATE ALL DFs
data = pd.concat(df_list, ignore_index=True)

# AGGREGATE DATA
result = data.groupby(["Date"])["Families", "Individuals"].agg([np.sum])

#SAVE FILE
file_name = r'C:\Users\Sarah\Desktop\IDPMosul\Output\Population\SumFamilies.csv'
result.to_csv(file_name, index=True)


--------------------Graph Plotting----------------------------------------
import pandas as pd
import matplotlib.pyplot as plt
from pandas.plotting import register_matplotlib_converters
register_matplotlib_converters()



df = pd.read_excel(r'C:\Users\Sarah\Desktop\IDPMosul\Output\Population\SumallIDP.xlsx')
df1 = pd.read_excel(r'C:\Users\Sarah\Desktop\ReturneeMosul\Output\Population\SumallReturnee.xlsx')

plt.plot(df['Date'], df['SumFamilies'])
plt.plot(df1['Date'], df1['SumFamilies'])

plt.xlabel('Years')
plt.ylabel('Num of Returnees/IDPs')
plt.title('Mosul Displacement Graph')
plt.legend()

#Show and Save
plt.show()
plt.savefig(r'C:\Users\Sarah\Desktop\Graph.jpeg')
---------------------Bar chart----------------------------------------
import pandas as pd
import matplotlib.pyplot as plt

df =pd.read_excel(r'C:\Users\Sarah\Desktop\wavetes.xlsx', index_col=0)
```

```
df.plot.bar()


plt.xlabel('Period')
plt.ylabel('Percentage')
plt.title('Mosul Displacement Graph')
plt.legend()
plt.show()
```

APPENDIX B

Full Code for ABM Implemenntation

```
extensions[gis csv]

globals[
  mosul-data
  bridges-data
  area1-data
  area2-data
  area3-data
  area4-data
  area5-data
  buildings-data
  camp-dataset
  population-dataset
  Province-dataset
  death_east
  death_west
  total_death
  shifted_death
  shifted_death_2
  ;death_total
 ; death_today
 ; death_before
  death_mosul
  month
  data
  variable
  ;risk_mosul
  ;host
  n_move_to_host
  n_move_to_camp
  counter
  factor
  damaging
  socialnetwork
  ristrected

  death-list
  prov_name_list
  s_factor

  ]


patches-own [
```

```
  ProvName
  neighbor
  MosulSide
  income-level
  death-rate
  Maxcap
  ;ristrected
  Open
  ;Gouvernorat
  CampName
  CampCurCap
  damaged
  ;host

]

turtles-own []

breed [camps camp]
breed [ households household ]
breed [provinces province]
breed [ nodes node ]
breed [idps idp]
breed [stayers stayer]
breed [returnees returnee]
;directed-link-breed [red-links red-link]


camps-own [capacity]
provinces-own [name pop]
households-own [ home-location home-node ;status-report
  wealth wealthweight householdID neigb
  risk-threshold stay-status deaths_ss dead risk-neighbor
  destinationprov hostfamilly destinationcamp  moved foundcamp postwar-status host]



to setup
  clear-drawing
  clear-all
  reset-ticks
  file-close-all
  ; zoom to study area
  resize-world 00 45 0 20
  set-patch-size 15

   ; upload mosul boundries
```

```
set mosul-data gis:load-dataset"data/MosulBoudrie.shp"
gis:set-world-envelope gis:envelope-of mosul-data
gis:apply-coverage mosul-data "Q_NAME_E" neighbor
;gis:apply-coverage mosul-data "BARIRER" Open
gis:set-drawing-color white
gis:draw mosul-data 1

set Province-dataset gis:load-dataset "data/provpoint.shp"
gis:apply-coverage Province-dataset "ADM1_EN" ProvName
;ask patches with [ProvName != 0]
;[set pcolor white]
ask patches  gis:intersecting Province-dataset
[sprout-provinces 1 [set name [ProvName] of patch-here
 set color yellow set size 1 set shape "star" ]]

  foreach gis:feature-list-of Province-dataset [
  y ->
  ask patches gis:intersecting y [
    set ProvName gis:property-value y "ADM1_EN"]]

 ;ask provinces[] ; test and debugging
gis:set-world-envelope-ds gis:envelope-of Province-dataset
print "Provinces loaded"
gis:set-world-envelope gis:envelope-of mosul-data

; upload bridges
set bridges-data gis:load-dataset "data/BridgesM.shp"
gis:set-drawing-color yellow
gis:draw bridges-data 4
print "Bridges loaded"

; upload buildings
set buildings-data gis:load-dataset "data/buildings.shp"
gis:apply-coverage buildings-data "SIDE" MosulSide
gis:set-drawing-color gray
gis:draw buildings-data 3
ask patches with [MosulSide != 0]
[set damaged 0]
print "Buildings loaded"

; upload restricted areas

set area1-data gis:load-dataset "RestrictedArea/Rest211016ISIS.shp"
;gis:apply-coverage area1-data "BARIRER" Open
print "21-10-16 loaded"

set area2-data gis:load-dataset "RestrictedArea/Rest031116.shp"
;gis:apply-coverage area2-data "BARIRER" Open
```

```
 print "03-11-16 loaded"

 set area3-data gis:load-dataset "RestrictedArea/Rest031116F.shp"
 ;gis:apply-coverage area3-data "BARIRER" Open
 print "03-11-16 loaded"

 set area4-data gis:load-dataset "RestrictedArea/Rest241116.shp"
 ;gis:apply-coverage area4-data "BARIRER" Open
 print "24-11-16 loaded"

 set area5-data gis:load-dataset "RestrictedArea/Rest090117.shp"
 ;gis:apply-coverage area5-data "BARIRER" Open
 print "09-01-17 loaded"

;opload 20oct camps
 set camp-dataset gis:load-dataset "data/20ctcamps.shp"
 gis:apply-coverage camp-dataset "MAX_CAPACI" Maxcap
 gis:apply-coverage camp-dataset "SITE_NAME" CampName
 ask patches  gis:intersecting camp-dataset
 [sprout-camps 1 [set capacity [Maxcap] of patch-here
    set color green set size 1 set shape "house" ]]

 foreach gis:feature-list-of camp-dataset [
   x ->
   ask patches gis:intersecting x [
   set Maxcap gis:property-value x "MAX_CAPACI"
   set Maxcap round Maxcap
   set CampName gis:property-value x "SITE_NAME"]
 ]

 ask camps[] ; test and debugging
 gis:set-world-envelope-ds gis:envelope-of camp-dataset
 print "Camps Loaded"

 gis:set-world-envelope gis:envelope-of mosul-data


 print count patches with [Maxcap > 0]

 ask patches with [MosulSide = "EAST" ][
   set pcolor green]
 ask patches with [MosulSide = "WEST"]  [
   set pcolor blue ]

 set death-list []

 ask patches [set income-level -1]
 set prov_name_list ["Al-Najaf" "Al-Basrah" "Al-Anbar" "Al-Muthanna" "Al-Qadissiya" "Al-Sulaymaniyah"
```

```
   "Babil" "Baghdad" "Diyala" "Duhok" "Erbil" "Kerbala" "Kirkuk" "Maysan" "Ninewa" "Salah Al-Din" "Thi
Qar" "Wassit"]
end
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;Restrected areas ;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to go

  ;print count patches with [MosulSide = "EAST"]
  ask patches with [MosulSide = "EAST"]
  [set death-rate death_east]
  ask patches with [MosulSide = "WEST"]
  [set death-rate death_west]

  ask patches with [Open = "Y"]
  [set ristrected ristrected = True]
  ask patches with [Open = "N"]
  [set ristrected ristrected = False]

   ;if (ticks = 45)[
   ; damage]
  ;damage
  restrictedarea
  set factor (ticks / 10) + 1
  set socialnetwork up-to-n-of 50 households
  ask socialnetwork [set host true]
  Risk-Perception
  ;set informeddeath up-to-n-of 50 households ;]
  ;ask informeddeath [set dead true]
  ;ask up-to-n-of total_death households [set dead true]
  deaths
  communicate
  ;clustering
  ;returneehome
  write-to-file
  outhouse-to-file
  outpatch-to-file
  file-final

  if (ticks = 65)
  [stop
  ]
  tick


end
```

```
to restrictedarea
;the number of weeks simulated is 45 weeks from october 1st 2016 to July 31st 2017 , it is assumed that one ticks
= one week
  if (ticks = 0)
     [gis:apply-coverage mosul-data "BARIRER" Open
      gis:set-drawing-color white gis:draw mosul-data 0

   ]
  if (ticks >= 11 and ticks < 12)
  [ gis:apply-coverage area1-data "BARIRER" Open
    gis:set-drawing-color red
    gis:draw area1-data 2]
  if (ticks >= 13 and ticks < 15 )
  [ gis:apply-coverage area2-data "BARIRER" Open
    gis:set-drawing-color red
    gis:draw area2-data 2
    gis:apply-coverage area3-data "BARIRER" Open
    gis:set-drawing-color yellow
    gis:draw area3-data 3
  ]
   if (ticks >= 16 and ticks < 22)
  [ gis:apply-coverage area4-data "BARIRER" Open
    gis:set-drawing-color cyan
    gis:draw area4-data 3
  ]

  if (ticks >= 23 and ticks < 50)
  [gis:apply-coverage area5-data "BARIRER" Open
   gis:set-drawing-color pink
   gis:draw area5-data 3
  ]

  if (ticks >= 51)
  [ gis:apply-coverage mosul-data "BARIRER" Open
    gis:set-drawing-color black
   gis:draw mosul-data 1
  ]
end
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;; Households ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to create-agents

  create-households 11000
  [
    set home-location gis:location-of ( gis:centroid-of
      (item random (length gis:feature-list-of buildings-data)
```

```
      gis:feature-list-of buildings-data))
      ;gis:apply-coverage buildings-data "SIDE" MosulSide
    setxy ( item 0 home-location ) ( item 1 home-location )

    set color red
    set shape "dot"
    set size 1



    ;set status-report ["Waiting..."]
    set home-node one-of nodes
    set risk-threshold random-float 1
    set stay-status "stayer"
    set moved 0
    set foundcamp 0



  ]
print count households
  ask n-of (0.25 *(count households)) Households [set host true]
  wealthness
  Neighbour-network
 ; Social-Network



end
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;VARIABLES;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
to wealthness
  ; set the wealth of the agent at random with those parameter
  ;  wealthness random 1 9999999999
  print "start wealth"

  ask n-of (0.13 *(count households)) Households [set wealth "very low"]
  ask n-of (0.21 *(count households)) Households [set wealth "low"]
  ask n-of (0.19 *(count households)) Households [set wealth "below average"]
  ask n-of (0.15 *(count households)) Households [set wealth "average"]
  ask n-of (0.13 *(count households)) Households [set wealth "above average"]
  ask n-of (0.08 *(count households)) Households [set wealth "high"]
  ask n-of (0.08 *(count households)) Households [set wealth "very high"]

  ask households [
    ifelse wealth = "very high" [set  wealthweight 0][
      ifelse wealth = "high" [set  wealthweight 2][
        ifelse wealth = "above average" [set  wealthweight 4][
          ifelse wealth = "average" [set  wealthweight 6][
            ifelse wealth = "below average" [set  wealthweight 8][
```

```
              ifelse wealth = "low" [set  wealthweight 10][
                if wealth = "very low" [set  wealthweight 12]]]]]]]

  print "end wealth"
end


to Neighbour-network

  ask households [
;;;;;;;;;;;;;;;;;;;; Second Version ;;;;;;;;;;;;;;;;;;

 create-links-with min-n-of 7 other turtles with  [MosulSide = [MosulSide] of myself] [distance myself]]
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
  ask one-of households with [count link-neighbors > 1]
  [
    ask link-neighbors [communicate]]
end

to communicate
  ;set informeddeath up-to-n-of 50 households [set dead true]
  ;ask up-to-n-of 50 households [set dead true]
 ; set dead true
  ask households [

  If any? link-neighbors with [dead = true]
    ;[ print " neighbor is dead"]
   [
      set deaths_ss deaths_ss + tolb
   ]

  ]
clustering

end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;; Clustering ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

to clustering
  ;ask n-of 2 households with [wealthweight = 8]
  ;[ print who
   ; ask patches in-radius 0.2
    ;  [ set pcolor blue
  ;set income-level 8] ]
  ;print patches with [income-level = 8]
  ;ask one-of households with [wealthweight = 8][
```

```
  ; if any? patches with [income-level = 8]
  ; [move-to one-of patches with [income-level = 8] print "moved"]
 ;print "other household" print who]

 let looplist [0 2 4 6 8 10 12]
 foreach looplist
 [ x ->
   ask n-of 2 households with [wealthweight = x]
   [
      ask patches in-radius 0.2
     [ set pcolor pink
       set income-level x]
       ask one-of households with [wealthweight = x][
         if any? patches with [income-level = x]
          [move-to one-of patches with [income-level = x and pcolor = pink] ]
   ]


 ]]

end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;Risk Perception;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
to deaths
 ;set dead true
 file-open "input/TesteCsv6.csv"
   ; ---------- Set values death_east and death_west----------

 if file-at-end? [stop]
 set death_mosul csv:from-row file-read-line
 set death_east item 1 death_mosul
 set death_west item 2 death_mosul
 set total_death item 3 death_mosul
 set shifted_death item 4 death_mosul
 set shifted_death_2 item 5 death_mosul
 ask up-to-n-of total_death households [set dead true]

end

to Risk-Perception
;; Mosul risk  defined as  R_m= 〖deaths〗_m*  〖deaths〗_social* 〖distance〗_m
ask households
 [
   if [MosulSide] of patch-here = "EAST" [
     set risk-neighbor  death_east  * deaths_ss * 1 ] ;* s_factor ]
```

```
    if [MosulSide] of patch-here = "WEST" [
      set risk-neighbor  death_west  * deaths_ss   * 1
    ]
  ]
  ;[ set risk-neighbor 0.01 + total_death  * deaths_ss * 1]
 set counter 0
;; change status from stayer to idp
  ask households with [stay-status = "stayer"] [

    if  risk-threshold < risk-neighbor and wealthweight <=  6 and ristrected = False [
      set stay-status "idp" set counter counter + 1] ;and ristrected = False

]
  ;; minimum distance to prince
 if (ticks < 10)[
  ask households with [stay-status = "idp"]
    [
     ; let targetpatch min-one-of (patches with [host = one-of [ 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18]])
[distance myself]
     ;     if targetpatch != nobody [
      ;    if moved = 0 [
       ;   ; move, set move status, update camp current capacity
       ;   move-to targetpatch
       ;   set moved 1
        ;  set destinationprov [host ] of targetpatch]

     ;;;;;;;;ProvName;;;;;;;;;;;;;
       let targetpatch min-one-of (patches with [member? ProvName prov_name_list]) [distance myself]

         if moved = 0 [
           ;move, set move status, update camp current capacity

          if  wealthweight <= 2 or host = true [
          move-to targetpatch
          set moved 1
            set destinationprov [ProvName] of targetpatch]]]
  ] ;];]

 if (ticks >= 10 and ticks < 20)[
    set n_move_to_host  round (counter / factor )
    ask n-of  n_move_to_host households with [stay-status =  "idp" ][

      ;let targetpatch min-one-of (patches with [host = one-of [ 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18]])
[distance myself]
       ;   if targetpatch != nobody [
        ;  if moved = 0 [
        ;  ; move, set move status, update camp current capacity
         ; move-to targetpatch
```

```
        ; set moved 1
        ; set destinationprov [host ] of targetpatch]

      ;;;;;;;;;ProvName;;;;;;;;;;;;;;
     let targetpatch min-one-of (patches with [member? ProvName prov_name_list]) [distance myself]
        if moved = 0 [
         if  wealthweight <= 2 or host = true[
          ;move, set move status, update camp current capacity
         move-to targetpatch
         set moved 1
           set destinationprov [ProvName] of targetpatch]]]
  ] ;]]

   set n_move_to_camp   counter - n_move_to_host
   if n_move_to_host >= counter
   [print counter print n_move_to_host ]
   if n_move_to_host < counter[
   if any? households with [stay-status =  "idp" and moved = 0][
   let movedcamps up-to-n-of n_move_to_camp households with [stay-status =  "idp" and moved = 0]; and
moved = 0]
   ask  movedcamps[
   ;; check camp

   let patchset sort-on [distance myself] (patches with [Maxcap > 0])

     foreach patchset [
       x ->  ; x is the current looping camp
       ; if x is "that" camp

       ; if current capacity smaller than max
       if ([CampCurCap] of x < [Maxcap] of x) and ( (ticks < 44 and [CampName] of x != "Nargizlia RC") or
(ticks > 44) )[ ;print [CampName] of  x ; if haven't moved yet
       ;print [CampName] of  x
         if foundcamp = 0 [
          ; print [CampName] of  x
         ; move, set move status, update camp current capacity
         move-to x
         set foundcamp 1
         set moved 1
         set destinationcamp [CampName] of x
         ask x [set CampCurCap CampCurCap + 1]
         ]

       ]
  ]]]]
  set damaging n-of 500 patches
  if (ticks >= 11 and ticks < 21)[
  ask damaging with [MosulSide = "EAST"]
```

```
 [
   set damaged 1
 ]]
 if (ticks >= 22 and ticks < 50)[
 ask damaging with [MosulSide = "WEST"]
 [
   set damaged 1
  ] ]
   let count-idp count households with [stay-status = "idp"]

 ask n-of round (0.5 * count-idp) households with [stay-status = "idp"][
   let homee patch ( item 0 home-location ) ( item 1 home-location )

   if total_death < shifted_death and  shifted_death < shifted_death_2  * deaths_ss * 1
   [  set postwar-status  "returnee"
      ;set stay-status "na"

     if [damaged] of homee  = 0  [ move-to homee ]
   ]
 ]

 ;ask households with [stay-status = "stayer"]
 ;[
 ;  if total_death  < shifted_death and   shifted_death < shifted_death_2
  ; [set postwar-status  "idp"]
; ]

end

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;Output file;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
to write-to-file

  if (file-exists? "output/attribute.csv")
 [
   carefully
   [file-delete "output/attribute.csv"]
   [print error-message]
 ]

file-open "output/attribute.csv"
 file-print (word "=============================================")
 file-print (word "============ simulation start ===============")
 file-type "ID " file-type ","
 file-type "Income " file-type ","
 file-type "Risk" file-type ","
 file-type " Status" file-type ","
```

```
  file-type "death social" file-type ","
  file-type "Neighbors" file-type ","
  file-write "Host Familly " file-type","
  file-type "Dead" file-type ","
  file-type "destination" file-type ","
  file-type "camps" file-type"\n"
  file-print (word)

  foreach sort households [aturtle -> ask aturtle [

    file-write who file-type ","
    file-write wealthweight file-type ","
    file-write risk-threshold file-type ","
    file-write stay-status file-type ","
    file-write deaths_ss file-type ","
    file-write count (link-neighbors) file-type ","
    file-write host file-type ","
    file-write dead file-type ","
    file-write destinationprov file-type ","
    file-write destinationcamp file-type"\n"]
  ]
  file-close
end
to outpatch-to-file
   if (file-exists? "output/tickspatch.csv")
  [
    carefully
    [file-delete "output/tickspatch.csv"]
    [print error-message]
  ]

 file-open "output/tickspatch.csv"
  file-print (word "==============================================")
  file-print (word "-------------- Tick Number ------------------")
  ;; use SORT so the turtles print their data in order by who number,
  ;; rather than in random order
  file-type "Tick Number" file-type ","
 ; file-type "ID " file-type ","
  file-type "xcor" file-type ","
  file-type "ycor" file-type ","
  file-type "damage" file-type "\n"
  foreach sort patches [ t ->
    ask t [
      file-write ticks file-type ","
      ;file-write who file-type ","
      file-write pxcor file-type ","
      file-write pycor file-type ","
      file-write damaged file-type "\n"
```

```
    ]
  ]
  file-print ""  ;; blank line
end
to outhouse-to-file
     if (file-exists? "output/tickshouse.csv")
  [
    carefully
    [file-delete "output/tickshouse.csv"]
    [print error-message]
  ]

 file-open "output/tickshouse.csv"
  file-print (word "==================================================")
  file-print (word "-------------- Tick Number ------------------")
  ;; use SORT so the turtles print their data in order by who number,
  ;; rather than in random order
  file-type "Tick Number" file-type ","
  file-type "ID" file-type ","
  file-type "xcor" file-type ","
  file-type "ycor" file-type ","
  file-type "Restricted" file-type ","
  file-type "percive" file-type ","
  file-type "Action" file-type "\n"
  foreach sort households [ t ->
    ask t [
      file-write ticks file-type ","
      file-write who file-type ","
      file-write pxcor file-type ","
      file-write pycor file-type ","
      file-write ristrected file-type ","
      file-write counter file-type ","
      file-write moved file-type "\n"

    ]
  ]
  file-print ""  ;; blank line
end
to file-final
    if (file-exists? "output/file.csv")
  [
    carefully
    [file-delete "output/file.csv"]
    [print error-message]
  ]

 file-open "output/file.csv"
```

```
  file-print (word "================================================")
  file-print (word "-------------- Tick Number ------------------")
  ;; use SORT so the turtles print their data in order by who number,
  ;; rather than in random order
  file-type "Tick Number" file-type ","
  file-type "ID" file-type ","
  file-type "Risk" file-type ","
  file-type " Status" file-type ","
  file-type "death social" file-type ","
  file-type "Restricted" file-type ","
  file-type "percive" file-type ","
  file-type "Action" file-type "\n"
  foreach sort households [ t ->
    ask t [
      file-write ticks file-type ","
      file-write who file-type ","
      file-write risk-threshold file-type ","
      file-write stay-status file-type ","
      file-write deaths_ss file-type ","
      file-write ristrected file-type ","
      file-write counter file-type ","
      file-write moved file-type "\n"

    ]
  ]
  file-print "" ;; blank line
end
```