Simulating the movement of forcibly displaced Syrians using agent-based modelling

Christa de Kock

Thesis presented in partial fulfilment of the requirements for the degree of
Master of (Industrial) Engineering
in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr Brian John van Vuuren

March 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 1, 2018

Abstract

Over the course of the past decade, numerous calamities within the international community have led to phrases such as 'refugee' and 'undocumented migrant' becoming commonplace in public discourse. In particular, conflict-induced forced displacement and the various challenges it creates have received notable attention. The challenge posed by the management of sudden migration of large groups of people lies in the ability to accurately portray and predict the scale and dynamics of such movement. This is further complicated in light of the fact that associated data pertaining to the migration are largely incomplete or untrustworthy.

Presently, there exists a significant lack of data required to perform strategic, long-term planning with respect to both current and future crisis situations. One of the specific challenges faced by researchers and humanitarian aid organisations in addressing forced displacement is the ability to effectively predict the movement type and proposed destinations of those who are forcibly displaced. The provision of a reasonably accurate estimation of the number of forcibly displaced people is a potentially critical input when planning the logistics and management of structures which exist to support those fleeing violence and persecution.

In this thesis, an agent-based simulation model with the ability to model the outbreak of conflict and the associated movement patterns, based on the decision-making of those forcibly displaced, with a reasonable level of accuracy is proposed. This model aims to address the lack of complete data by estimating, as accurately as presently possible, the number of refugees, undocumented migrants and internally displaced persons fleeing conflict-affected Syria, whilst also proposing the anticipated destinations of these people, based on their personal characteristics and associated anticipated decision-making. The agent-based model is formulated and operates within the AnyLogic Simulation Software environment and allows for an animated output visualising the state of conflict over an area and the movement of agents.

Collaboration with various subject matter experts throughout the development of this model allowed for significant insight and knowledge to be gained into aspects pertaining to the coalition of research in the fields of forced migration, computer simulation and human decision-making, which do not necessarily appear presently in literature. The approaches taken to model people and their decision-making are therefore endorsed by the knowledge gained from this collaboration of research and expert opinion. To the best of the knowledge of the consulted subject matter experts, no such model which considers such a vast array of factors and implications pertaining to refugee modelling in the presence of conflict presently exists and, as such, the research has sparked significant interest and excitement in the international research community. To this effect, the author has been invited to serve as an expert panel member at a panel discussion to be held in Thessaloniki, Greece at the 2018 *International Association for the Study of Forced Migration* conference. Furthermore, numerous requests for collaboration and research visits at several prestigious refugee and conflict modelling centres worldwide have been received.

The agent-based simulation model presented in this thesis is subjected to a number of traditional verification and validation techniques which include the calibration of parameters pertaining to the modelling of conflict, the replication of visualised recorded data, a thorough face validation and a parameter variation analysis which, ultimately, facilitates the implementation of a graphical user interface. The model is therefore deemed to possess the ability to model any required scenario when equipped with the correct parameter values, owing to its flexibility and usability. Furthermore, the animation output allows for easy interpretation and deduction of the model development, particularly by parties who are not necessarily scientifically trained.

A model of this nature naturally provides numerous avenues for future work, both as improvements to or extension on the existing model. Such follow-up work may ultimately allow for the model to assist in the planning and logistic strategies pertaining to facilities and resources required to accommodate incoming refugees, internally displaced migrants and undocumented migrants in different areas, as well as predicting the population fluctuations in affected areas during times of conflict, natural disaster, or other refugee-causing events.

Uittreksel

Met die verloop van die laaste dekade het 'n verskeidenheid rampe in die internasionale gemeenskap daartoe gelei dat woorde soos "vlugteling" en "ongedokumenteerde emigrant" hul in die volksmond gevestig het. Verplasings weens geweld en die uitdagings daaraan verbonde het besonder baie aandag ontvang. Wat die bestuur van die skielike migrasie van groot groepe mense aangaan, lê die uitdaging daarin om die omvang en dinamiek van die beweging akkuraat voor te stel en te voorspel. Hierdie word verder bemoeilik deur dat inligting wat verband hou met migrasie grootliks onvolledig of onbetroubaar is.

Daar is tans 'n beduidende tekort aan die data wat benodig word vir strategiese langtermyn beplanning oor die huidige en toekomstige krisis situasies. Een van die spesifieke uitdagings wat navorsers en humanitêre hulporganisasies in die gesig staar met die hantering van gedwonge verskuiwing, is die vermoë om die bewegingstipe en voorgestelde bestemmings van die verplaasdes, effektief te voorspel. Toegang tot 'n redelik akkurate beraming van die aantal mense wat weens geweld ontwortel is, kan 'n kritiese inset wees met die beplanning van logistiek en die bestuur van strukture wat aan die vlugtelinge ondersteuning bied.

Hierdie proefskrif stel 'n agent-gebaseerde simulasiemodel voor met die vermoë om die uitbreek van konflik en die gepaardgaande bewegingspatrone gebaseer op die besluitneming van diegene wat met geweld verplaas word, met 'n redelike vlak van akkuraatheid te voorspel. Die model beoog om onvolledige data aan te spreek deur die aantal vlugtelinge, ongedokumenteerde migrante en plaaslik ontworteldes wat van die konflik in Sirië vlug, so akkuraat as moontlik te beraam. Terselfdertyd word verwagte bestemmings van die vlugtelinge voorgestel, gebaseer op hul persoonlike eienskappe en gepaardgaande verwagte besluitneming. Die agent-gebaseerde model is geformuleer in die AnyLogic Simulasie Sagteware omgewing en voorsien 'n geanimeerde uitset wat die toestand van konflik visualiseer oor 'n gebied asook die beweging van agente.

Samewerking met verskeie vakkundiges gedurende die ontwikkeling van die model het tot die verkryging van noemenswaardige insigte en kennis gelei in die navorsingskoalisie tussen die velde van geforseerde migrasie, rekenaar simulasie en die menslike besluitnemingsproses. Hierdie is nie noodwendig in huidige literatuur vervat nie. Die benadering gevolg vir die modellering van mense en hul besluitneming word gevolglik ondersteun deur die kennis wat uit hierdie navorsingskoalisie verkry is, tesame met ander relevante vakkennis. Sover die kennis van die gekonsulteerde vakkundiges strek, bestaan daar tot op hede nie 'n model wat so 'n groot verskeidenheid faktore en implikasies rakende die modellering van vlugtelinge in konflik geteisterde omgewings in ag neem nie. Gevolglik het die navorsing tot baie belangstelling en opgewondenheid in die internasionale navorsingsgemeenskap gelei. Om hierdie rede is die skrywer uitgenooi om as 'n deskundige paneellid te dien tydens 'n paneelbespreking wat in Thessaloniki, Griekeland, by die 2018 *International Association for the Study of Forced Migration* konferensie gehou sal word. Daarbenewens is talle versoeke vir samewerking en navorsingsbesoeke by verskeie gesogte vlugtelinge- en konflikmodelleringsentrums wêreldwyd ontvang.

Die agent-gebaseerde simulasiemodel wat in hierdie proefskrif aangebied word, is onderworpe aan ‘n aantal tradisionele verifikasie- en valideringstegnieke, insluitend die kalibrasie van parameters rakende die modellering van konflik, die replisering van opgetekende visuele data, ‘n deeglike gesigsvalidering en ‘n parameter variasie-analise wat uiteindelik die implementering van ‘n grafiese gebruikerskoppelvlak fasiliteer. Die model word dus geag die vermoë te h om enige vereiste scenario te modelleer wanneer dit toegerus is met die korrekte parameterwaardes, op grond van die model se buigsaamheid en bruikbaarheid. Verder bied die animasie-uitset 'n gebruikersvriendelike interpretasie en analise van die modelontwikkeling, veral vir diegene wat nie noodwendig die wetenskaplike agtergrond het nie.

‘n Model van hierdie aard bied natuurlik talle weë vir toekomstige werk, beide as verbeterings aan of uitbreiding op die bestaande model. Sodanige opvolgwerk kan uiteindelik daartoe lei dat die model gebruik kan word om te fasiliteer in die beplanning en logistieke strategie rakende fasiliteite en hulpbronne wat benodig word om inkomende vlugtelinge, plaaslik ontworteldes en ongedokumenteerde migrante in verskillende gebiede te akkommodeer, asook om sodoende die bevolkingsfluktuasies in geaffekteerde gebiede te voorspel tydens tye van konflik, natuurrampe of ander vlugtelinge-veroorsakende gebeurtenisse.

Acknowledgements

The author wishes to acknowledge the following people and institutions for their various contributions towards the completion of this work:

- God, My Father, for His way is perfect;
- Dr Brian van Vuuren, for taking me in and giving me the freedom to grow academically, you are an extraordinary supervisor and, above all, a dear friend;
- Dr Damla Aksel, Dr Erika Frydenlund, Dr Derek Groen, Dr Carlos Lemos, Wiji Shomary, Dr Christopher Smith and Prof Theodor Stewart, for sharing your expert knowledge and insight with me;
- Harry Crossley Foundation, for the financial support;
- My friends at the Department of Industrial Engineering and, in particular, the *Stellenbosch Unit for Operations Research and Engineering* (SUnORE), for all the "tea times" we had together, the many distractions and for just accepting me as is; and
- My family, this truly is unconditional love — home is where my bunch of crazies are.

Table of Contents

List of Acronyms

ABM: Agent-based modelling

AHP: Analytic hierarchy process

ELECTRE: Elimination and choice expressing reality

EVAMIX: Evaluation of mixed data

GDELT: Global Database of Events, Language, and Tone

GIS: Geographical Information System

GPS: Global Positioning System

GUI: Graphical User Interface

IDP: Internally displaced person

MADM: Multi-attribute decision-making

MAUT: Multiple attribute utility theory

MAVT: Multi-attributive value theory

MCDM: Multi-criteria decision-making

MODM: Multi-objective decision-making

NAIADE: Novel approach to imprecise assessment and decision environments

PROMETHEE: Preference ranking organization method for enrichment evaluation

SMART: Simple multi-attribute rating technique

TOPSIS: Technique for order preference by similarity to ideal solution

UNHCR: United Nations High Commissioner for Refugees

UNOCHA: United Nations Office for the Coordination of Humanitarian Affairs

UTA: Utility theory additive

List of Figures

List of Tables

CHAPTER 1

Introduction

Contents

1.1 Background

In 2015, one in every 113 people worldwide was forcibly displaced from their place of residence, whether as an asylum-seeker, a refugee or a person displaced within the borders of their own country. The *United Nations High Commissioner for Refugees* (UNHCR) estimated the total number of forcibly displaced people that year to be 65.3 million, 40.8 million of whom were people displaced within their country of origin, with 21.3 million as refugees fleeing across international borders [129].

The *Global Opportunity Report 2017* [47] identifies several topics pertinent to global unstable regions as some of the most important opportunities to pursue in 2017 with regard to social impact. The gross impact of persecution and violent conflicts are evident worldwide as millions of people are continually forced to flee their homes and seek refuge elsewhere. An index measuring the global instability worldwide, based on the level of safety and security, the extent of domestic or international conflict, as well as the degree of militarisation, is illustrated in Figure 1.1.

It is estimated that 46% of people in poverty will be living in unstable and conflict-affected areas by 2030. The scale, complexity, duration and reoccurring nature of crises faced in recent times requires a systematic approach in order to manage or curb such realities [47].

The conflict in Syria, for example, has caused mass population displacement, the ramifications of which have extended to neighbouring countries, Europe and beyond [2]. Political and economic deterioration may be a consequence of such a mass refugee surge, as is the case for some of Syria's neighbouring countries [1]. The implications of such humanitarian crises effect civilians, governments, international humanitarian organisations, as well as global governance [41]. Aiyar *et al.* note the economic challenges faced by Europe accompanying the influx of refugees, whilst Richmond [102] discusses forced migration from a sociological perspective, introducing determinant factors other than politics which influence forced migration.

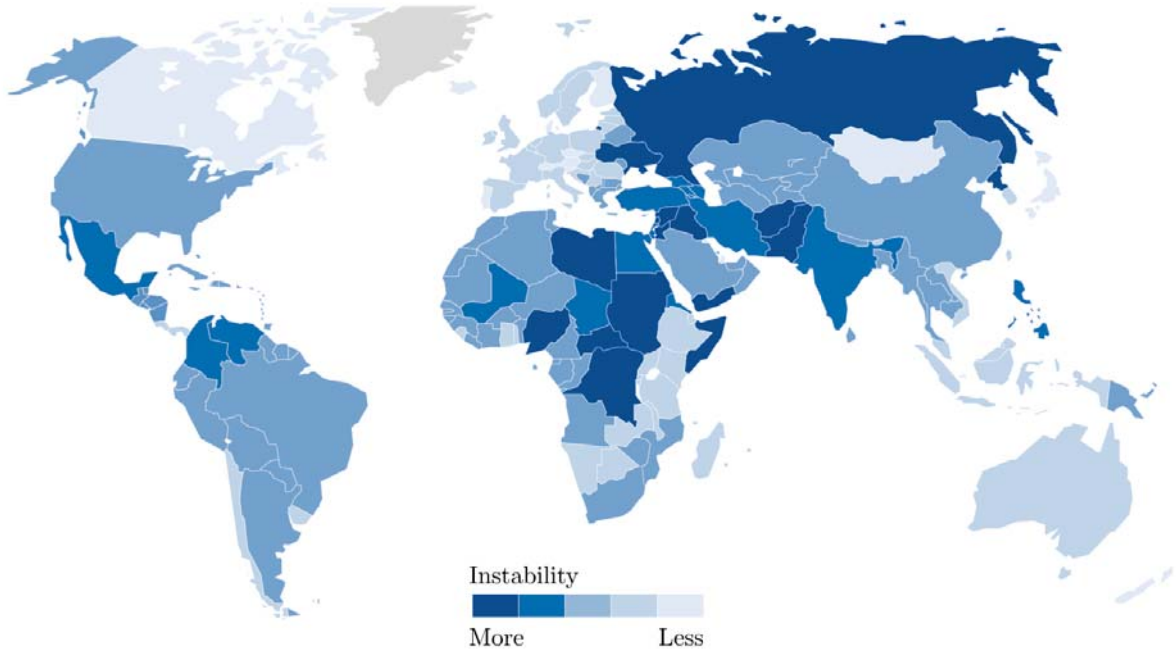


FIGURE 1.1: *A map indicating the global instability in 2015 [47, 60].*

The potential impact which may be made by humanitarian intervention efforts in these crises causes for many challenges and problem-solving opportunities to arise. Researchers such as Alhanaee and Csala [5] and Greenwood [49] have attempted to identify the motives of migrants in order to better understand the spatial phenomenon of migration with respect to economic and social aspects. It is imperative for humanitarian support organisations and policy makers to understand the motive behind a person's migration movements in an attempt to plan for necessary resources and logistics to facilitate their arrival [56]. The lack of adequate and complete data presents a serious problem to humanitarian aid, especially with regards to forced displacement. Efforts in improving the reliability, quality and scope of data concerning forcibly displaced people are necessary to address the gaps in data currently available to assist in long-term development planning during crises [106].

Edwards [37] discusses the potential of using computational models in order to predict the spatial patterns of forcibly displaced individuals, emphasising the assistance such models may be able to provide humanitarian aid organisations and policy makers. Groen [51] affirms the importance of utilising such models in order to capture the movements of refugees on a global scale. The principal of modelling is a tool utilised to study the behaviour of large complex systems and, in particular, its dynamic behaviour, when the complexity does not allow for analytical evaluation [112]. One of the most powerful tools available in comprehending the behaviour of complex systems and processes is simulation modelling [108].

Simulation modelling is the computer-based imitation of a real-world system [118]. Furthermore, agent-based simulation is the modelling of a collection of autonomous decision-making entities, where the behaviour of each agent depends on a basic set of rules which guide decision-making [18]. Agent-based modelling (ABM) allows for the consequences of individual decisions to be modelled, taking into account the complexities of social systems such as behaviours, motivations and relationships between agents [7]. The ability of agent-based models to simulate the interactions among individuals makes it well-suited for the purpose of modelling sociological and psychological behaviour, as well as human interaction with one another, as well as with the environment [63].

ABM is capable of facilitating a synthetic environment which allows for an understanding of the collective behaviour of forcibly displaced persons through computational experimentation [7]. With knowledge pertaining to the migration behaviour of people, an agent-based model potentially could be developed and implemented by researchers to predict the number of people displaced, identifying likely destinations for refugees, as well as the population size of refugees per destination. Furthermore, such a model should serve in identifying appropriate locations for aid distribution points, as well as predict the anticipated demand [37].

Simulation models of this nature which presently exist in literature include, amongst others, a model developed by Lemos *et al.* [77] which simulates social conflict and civil violence in order to capture the characteristics associated with its spread, a model developed by Crooks and Wise [31] which aids humanitarian relief after the occurrence of a natural disaster, a model simulating the autonomous decision-making of environmentally induced migrants, developed by Smith *et al.* [116] and a model which explicitly incorporates the social network between people migrating between Ecuador and the United States, developed by Rehm [101].

Anderson *et al.* [7] utilised ABM in simulating the effect of changes to humanitarian assistance policies with respect to the health and safety of refugee communities. The concept of simulating policies allows a user to evaluate the potential impact of decisions and test various strategies. Collins and Frydenlund [29] further proposed an agent-based model to simulate strategic group formulation of refugees when fleeing. This model investigates the evacuation time of refugees, assuming that refugees tend to form groups when travelling over long distances. Another agent-based model, developed by Orfano [93], simulates economic empirical evidence and long-term effects of forced migration.

Johnson *et al.* [63] calibrated an agent-based model for the use of peace-keeping within a refugee camp scenario. Owing to quantitative data not being available, the calibration was performed by relying on experimental designs and plausible considerations made with the help of subject matter experts. The spread of disease at refugee camps, another pertinent issue faced by humanitarian agencies at refugee camps, is modelled by Hailegiorgis and Crooks [55], taking the social behaviour of people and their movements into account. Another simulation model focussing on the displacement of Syrians within the city of Aleppo was developed by Sokolowski *et al.* [117] as a method of investigating the decision-making of citizens during forced migration.

Klabunde and Willekens [67] reviewed the use of agent-based models in modelling the decision-making of an agent during migration, concluding that ABM is still in its infancy when considering migration. Although a number of migration models exists, they differ in scale, complexity and documentations owing to the influence of different disciplines and limited knowledge. It was further found that the decision-making processes of agents are often modelled rudimentarily and that the criteria determining decisions should not only include behavioural rules (as in ABM), but also rates and probabilities (as in microsimulation). A further notable challenge in this field is the validation of agent-based models, owing to the lack of empirical data, and effectively modelling the manner in which migration decisions are influenced by a human's life course.

Awareness surrounding crises and, in particular, conflict-induced forced displacement, has increase notably within the international community over the last few years. The challenge, however, is to accurately portray the true scale and dynamics of the issue, as not all available data are credible or complete. There exists, currently, significant gaps in the data necessary to perform long-term development planning in crisis situations. The aggregate number of 65.3 million forcibly displaced people is only an estimate and data concerning those displaced within specific countries of origin are even less reliable [106].

More often than not, research does not to focus on those internally displaced or the undocumented migrants, which augments the gap in available data [4, 32]. The unknown number of forcibly displaced people creates an obstacle to humanitarian support organisations, especially in times where a crisis is critical in nature [55].

The greatest overall challenge faced by researchers and humanitarian support organisations addressing forced displacement is predicting the proposed destinations of people. Attempts have been made to address this issue, although the obstacle to predict such random movement remains in light of the fact that migration is a highly structured process dependent upon patterns, historical context and the manner in which an individual's decision-making process develops [36, 37]. The ability to predict the movement of forcibly displaced people with some measure of accuracy is critical to organisations in enabling the planning of logistics and procurement of resources which aim to support those fleeing violence and persecution [29]. Groen [51] endorses the use of simulation modelling to account for the shortfalls of incomplete empirical information in the monitoring infrastructure and predictions of refugee movements.

1.2 Problem description

In light of the aforementioned shortcomings in the modelling of conflict outbreaks and associated movement patterns of forcibly displaced people, an agent-based model is proposed which has the ability to model conflict and the associated decision-making of those forcibly displaced with a reasonable level of accuracy in an attempt to address the lack of complete data by estimating the accurate number of refugees, undocumented migrants and internally displaced persons (IDPs) fleeing conflict-affected Syria, whilst also proposing the anticipated destinations of these people, based on their personal characteristics and associated anticipated decision-making.

1.3 Thesis scope and objectives

In order to facilitate the simulation of refugee movement and consequent capture of inference data, a robust simulation framework is required. In an attempt to provide a platform for such framework, the following objectives are pursued in this thesis:

I To *conduct* a comprehensive study of the literature pertaining to:

  (a) an introduction to forced migration and a brief history on forcibly displaced people,
  (b) the displacement factors and types of movement of forcibly displaced people,
  (c) an overview of the current state of global conflict-induced forced migration,
  (d) the situation faced by those forcibly displaced in Syria,
  (e) computer simulation modelling techniques and, in particular, agent-based modelling,
  (f) a brief overview on the field of decision-making,
  (g) multi-criteria decision-making methods, and
  (h) the modelling of human decision-making.

II To *develop* an agent-based simulation model aimed at investigating the movement of forcibly displaced Syrians. This model should be capable of:

(a) reading input data pertaining to the initiation of conflict,
(b) allowing the user to manually initiate conflict,
(c) replicating the spread and depletion of conflict in a realistic manner, informed by historical observation,
(d) modelling an aggregated population of agents,
(e) determining an agent's ability to withstand conflict, and
(f) modelling the decision-making of agents under the influence of conflict.

III To *implement* the simulation model within a suitable framework so to assist model users in:

(a) estimating the total number of forcibly displaced persons more accurately,
(b) estimating the number of refugees, undocumented migrants and IDPs, which compose the forcibly displaced persons group,
(c) predicting the population fluctuations within countries affected by the crisis, and
(d) presenting an animation output of the simulation (*i.e.* the conflict present and the movement of forcibly displaced people) in a user-friendly manner.

IV To *verify* and *validate* the simulation model by means of a parameter calibration and variation in order to assess its performance, as well as illustrate its flexibility in simulating user-specified instances of conflict outbreak.

V To *recommend* possible future improvements or additions which may be included in the model, as well as sensible follow-up work which may stem from this study.

1.4 Research methodology

The methodology followed in this thesis in order to design and develop and agent-based simulation model for investigating the movement patterns of Syrian refugees, as described in §1.3, is as follows:

I *Consult* and *review* existing literature on forced migration and, in particular, the environment which Syrians reside, so as to formulate a more comprehensive understanding in context of the situation, as well as existing literature in the field of decision-making and, in particular, the modelling of human decision-making.

II *Develop* a competence in the AnyLogic Simulation Software Suite for the purpose of developing an agent-based model.

III Iteratively *construct* the agent-based model which employs the aspects of decision-making and predicts the movement of those forcibly displaced within Syria.

IV *Verify* the performance and execution of the model with respect to each aspect modelled.

V *Validate* the model by means of a face validation with the assistance of subject matter experts.

VI *Calibrate* the model by means of a parameter variation.

VII *Illustrate* the flexibility, usability and applicability of the model by performing a parameter establishment analysis to equip the model in replicating previously encountered conflict situations.

1.5 Thesis organisation

Apart from this introductory chapter, this thesis contains six additional chapters. Chapter 2 provides the reader with an understanding of the phenomenon of forced migration, the various factors which result in forced displacement and the typology of forced migration with regards to movement types. A brief history of the refugees is given, where after the current state of conflict-induced forced migration is discussed, before focussing on those forcibly displaced in Syria.

In Chapter 3, information pertaining to the different aspects and considerations of computer simulation modelling is provided. The chapter introduces the different types of simulation modelling, discusses levels of abstraction, various simulation modelling approaches, generic steps followed in completing a simulation study, and the verification and validation of simulation models. This is followed by a discussion on agent-based modelling, the advantages and disadvantages of this approach, and the components of an agent-based model. The chapter closes with a review on the application of agent-based modelling with regards to forced displacement, considering the modelling of different movement types, existing migration models focussing on forced displacement and the determinants of forced migration.

The aim of Chapter 4 is to provide the reader with an overview of the field of decision-making, taking into account existing prescriptive and descriptive theories. Multi-criteria decision-making methods are discussed, following a discussion on the modelling of human decision-making. Following this, the modelling of the decision-making of forcibly displaced is elaborated upon, considering both the choice of a movement type and of a proposed destination of a modelled migrant.

Chapter 5 describes the agent-based model developed and considers the various components modelled. The chapter opens with background to the model and the AnyLogic Simulation Software Suite utilised, before discussing the various assumptions and limitations made throughout the modelling process. The three elements modelled — the modelling of conflict, the modelling of people and the modelling of a person's decision-making — are described in detail, before an explanation of the graphical user interface created to control these elements.

In Chapter 6, the agent-based model developed is verified. The chapter opens with an explanation of the concept of verification, where after each of the three components modelled are individually verified. The verification process is performed by means of a number of cases being tested and its output analysed in order to determine if the model is accurately developed.

The model validation and analysis is performed in Chapter 7. The concepts pertaining to this analysis is briefly discussed before a calibration on the parameters pertaining to the modelling of conflict is performed. This is followed by a validation of the conflict as modelled which includes the comparison of real-world visualised data and the simulated output. The face validation is discussed thereafter, followed by the parameter establishment analysis as performed on the agent-based model. In the closing of this chapter the inclusion of a decision support and analysis tool is briefly discussed.

Finally, Chapter 8 provides a short summary of the thesis and an overview of the contributions made. Suggestions with regards to future work which may improve or build on the work conducted in this study, or serve as future projects which may stem from this study are then proposed.

CHAPTER 2

Forced migration

Contents

This chapter discusses the phenomenon of forced migration by examining various causes of forced displacement and illustrating a typology of forcibly displaced persons. Furthermore, a brief overview of historic refugee movement is given, followed by an overview of current global forced displacement, with specific focus on Syria.

2.1 Factors leading to forced displacement

Disasters, both natural and man-made, are increasingly disrupting the living conditions of individuals on a global scale. In most cases these incidents lead to the displacement of individuals residing in these areas [95]. The *International Federation of Red Cross and Red Crescent Societies* [59] defines a disaster as a sudden, calamitous events which disrupts the activities of a society or community and cause human, material, economic, or environmental losses which exceed the recovery capacity of the affected community or society using exclusively its own resources [95].

Disastrous events can occur as a result of numerous factors. Green and McGinnis [48] performed a study on the classification of disasters based on such factors. Three main classes were identified to describe the highest order range of disaster events, namely *natural disaster*, *human systems failure* and *conflict-based disaster* [48].

Natural disasters include any event which results from natural forces in which human intervention is not the primary cause of the forces [48]. Such naturally occurring phenomena, caused either by a rapid or slow onset events, can be geophysical (earthquakes, tsunamis, volcanic activity), hydrological (avalanches, floods), climatological (drought, wildfires, extreme temperatures), meteorological (cyclones, storms, wave surges) or biological (disease epidemics, plagues) [59].

Conflict-based disasters result from internal conflict within a country or external conflict directed at it. More specifically war, terrorism, genocide, internal state terrorism, as well as political, social and economic instability are examples of such disasters [48]. These disasters are intended to kill or persecute individuals based on their religion, race, nationality or political opinion.

Human systems failure encompasses disasters such as industrial accidents, transportation accidents (derailment of trains or collapsing bridges) or death due to poor control of hazardous materials [48, 59]. The primary differentiation between human systems failure and conflict-based disasters is intent [48]. For example, a nuclear testing plant which caused the contamination of land and long-term death due to radiation may have been unintentional, whilst the bombing of Hiroshima and Nagasaki in 1945 was a deliberate action intended to kill civilians and destroy land [48].

Apart from disasters, people may also lose their homes to developments. Development-induced displacement occurs when people are forced to leave their homes due to development projects. These projects may include natural resource extraction, urban renewal, industrial parks, infrastructure development or conservation programs [22]. Displacement typically progresses slowly as it usually only occurs after a development project commences [24].

2.2 A typology of forced migration

Migrants are people who leave or flee their homes, relocating either within or across international borders, to seek better or safer surroundings [94]. In the literature, migrants are classified into two distinct groups — those who voluntarily leave their homes and those who are forcibly displaced [9, 24, 94, 102]. *Voluntary migration* often occurs due to economic reasons, where a person moves from one place to another in pursuit of better living conditions [36]. *Forced migration* takes place when people flee as a result of a legitimate fear of being persecuted, or in reaction to crisis situations such as war, famine or other disasters [102].

Castles [22] reports that forced migration has grown in volume and political significance since the Cold War and has thus become a crucial dimension of globalisation. A major difference between voluntary and forced migrants is the fact that, while voluntary migrants may seek to escape uncomfortable circumstances, forced migrants could face imprisonment, deprivation of basic rights, physical injury or even death [83].

A diagram illustrating a typology of migration, with emphasis on forced migration, is shown in Figure 2.1. As indicated in the figure, individuals who are forced to relocate, either move within the borders of the country (*i.e.* internally displaced persons) or across its borders (*i.e.* international migrants).

Internally displaced persons (IDPs) are forced to leave their place of residence, but do not cross an international border. This may occur as a result of, or in an effort to avoid armed conflict or other disasters [9, 129]. For this reason, as indicated in Figure 2.1, IDPs relocate due to natural disasters, human failure systems, conflict-based disasters or development. The *Global Report on Internal Displacement* [24] states that, by the end of 2015, there were twice as many IDPs as there were refugees on a global scale. This phenomenon was caused by the inclination people to be less likely to completely abandon their country of residence, based on their implicit belief that returning to their home country would then not be possible. Furthermore, immigration is a financially costly exercise [93]. The report also suggests that countries with economic and political instability are more likely to have a greater number of IDPs resulting from a disaster, especially among people of low-income and those subject to socio-economic discrimination [24].

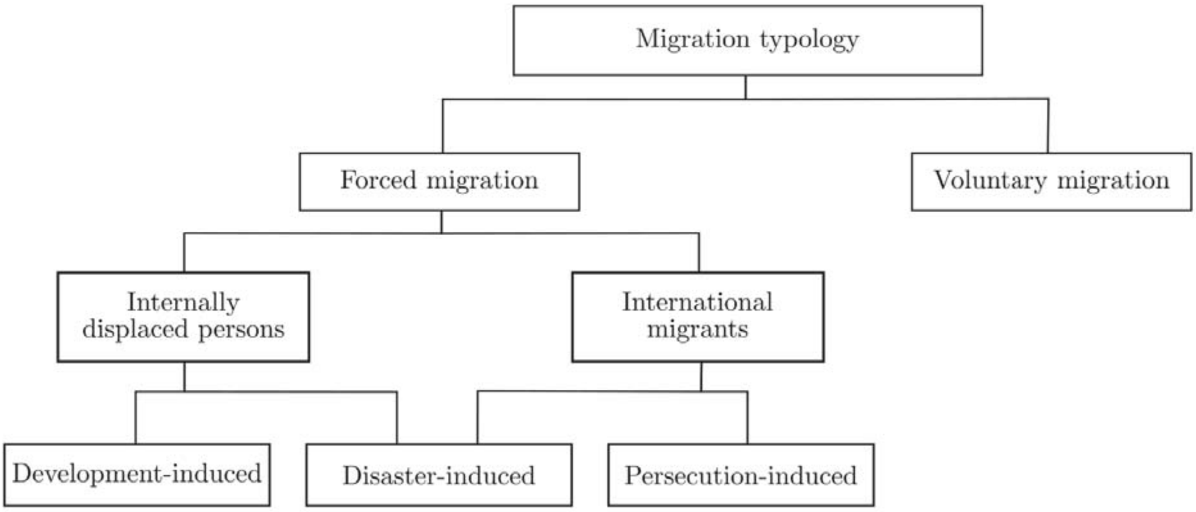


FIGURE 2.1: *A migration typology expanding forced migration.*

International migrants are migrants who flee their country of origin owing to natural disasters, human systems failure, conflict-based disasters or persecution. Those fleeing persecution tend to settle indefinitely in a new host country, whereas those who leave their country for other reasons tend to view the migration as temporary [93]. Individuals who relocate as a result of persecution are explicitly known, under international law, as refugees [127]. Refugees are formally defined by the 1951 United Nations Convention as "*individuals with a well-founded fear of being prosecuted, based on their race, religion, nationality, membership of a particular social group or political opinion in their country of origin*" [127]. These individuals are guaranteed, under international law, the right to request asylum in a country.

Asylum-seekers are individuals who sought international protection in a foreign country and are waiting to find out if their claim for refugee status has been granted [94, 97]. Such applications are reviewed by the relevant authorities in order to determine whether or not an individual can be classified as a refugee. Unsuccessful applicants may be deported back to their country of origin. The possibility of deportation leads to uncertainty surrounding this process on behalf of the migrant. For this reason, many migrants choose not to attempt seeking asylum, resulting in many undocumented migrants presently living in foreign countries [36].

The options available to a person fleeing persecution is illustrated by Figure 2.2. As may be seen, the person concerned can choose to apply for asylum and will then be classified as an asylum-seeker until the results of their application are known. When a persecution-induced migrant chooses not to seek asylum, they become an undocumented migrant, thereby living illegally in a foreign country [97]. Asylum-seekers, on the other hand, can either be approved as refugees or as unsuccessful applicants, the latter of which will lead to deportation.

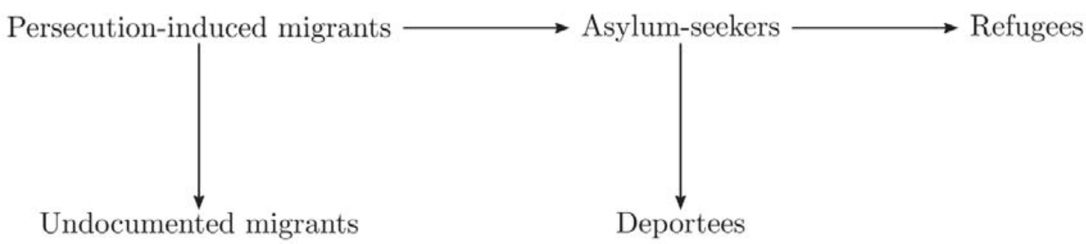


FIGURE 2.2: *The different statuses a persecution-induced migrant may take on.*

Those who seek asylum often have a high level of education and many are forced to relocate in spite of their jobs or studies. There are, however, many persecution-induced migrants (especially those displaced as a result of civil war) who are less educated. This suggests that less educated individuals typically choose to remain undocumented migrants, as opposed to risking deportation [14].

The *International Office for Migration* estimated in 2010 that worldwide, roughly 20–30 million migrants were undocumented, which comprises of about 10–15% of the global population of migrants [97]. This office refers to undocumented migrants as both those who arrive in a country without the necessary documentation, as well as those who violate tourist visas by applying for work.

Some literature, such as Dragostinova [36] and Sarzin [106], informally define refugees as people who attempt to flee their country due to armed conflict or persecution. Castello [99] agrees, extending the argument that the United Nations Convention's definition of refugees is too narrow, as those fleeing indiscriminate violence will struggle to find recognition as refugees. Richmond [102] also suggests that the formal definition of refugees should be revisited to include all those in peril from natural and unnatural disasters. The UNHCR, for example, recognises several groups of people as deserving of protection, including those forcibly displaced due to natural or man-made disasters, IDPs, stateless persons and those fleeing generalised violence [129]. International law, however, has not yet agreed upon a broader definition for refugees [102].

2.3 A brief history of forcibly displaced people

The concept of people being forced to leave their home country is not only a current phenomenon, but has occurred continuously throughout history.

In 1685, Louis XIV of France issued an edict, known as *The Edict of Fontainebleau*. This edict placed a ban on protestant schools and pastors and led to Huguenot churches being destroyed. The Huguenots therefore began to risk persecution for practising their faith freely [26]. *The Edict of Fontainebleau* is one of the first recognised displacements of people across nations and it is estimated that approximately 200 000 individuals fled France in the two decades that followed its institution [26].

Some time later, in the late 18^{th} century, an estimated 5–7 million Ottoman Muslim citizens moved to Anatolia (presently known as Turkey) in order to escape religious persecution [26]. This gave rise to the Ottoman Empire, now known as the Turkish Empire. Almost a century later, in 1881, the next mass exodus occurred after the assassination of Russian Tsar, Alexander II. It is approximated that 2 million Jews fled, primarily to the United Kingdom, the United States and Europe owing to the Russian anti-Jewish sentiment [26].

Before 1914, immigration control did not exist, allowing migrants and refugees to move freely between countries. World War I and the Russian revolution of 1917 put an end to this freedom of movement by advancing border controls, quotas and the like. This resulted in the first ever refugee crisis in Europe. Between 1914 and 1922, around 5 million people became refugees as a result of the aforementioned war [21]. As devastating as this refugee crisis appeared at the time, it was simply the foreshadow of what was to come. During World War II, more than 40 million refugees existed in Europe alone, uprooted and mostly without housing [21, 26]. This unprecedented disaster expedited the formation of international law and organisations which were tasked with the responsibility of managing the refugees, such as the *Intergovernmental Committee on Refugees* (1938), the *United Nations Relief and Rehabilitation*

Administration (1943), the *International Refugee Organisation* (1946) and the *Universal Declaration of Human Rights* (1948). Furthermore, the Geneva convention took place in 1949, where a series of treaties set out international law in terms of humanitarian policy during armed conflict [21, 26]. In 1950, the *United Nations High Commissioner for Refugees* was established. A year later, 147 governments signed the international *Convention on the Status of Refugees* [15].

Shortly after World War II, the 1948 Palestinian exodus resulted in almost 80% of the Arab population in Palestine fleeing to Israel. This was initiated by an Arab village that was attacked by a Zionist military group. The United Nations set up an agency to assist the 5 million refugees requiring assistance [26]. Then, in 1979, the Soviet Union occupied Afghanistan and an estimated 5 million people fled to Pakistan. Since 1990, the number of refugees in Afghanistan has not fallen below 2 million per annum [26].

During the 1990s, there was a further refugee crisis as a result of a number of incidents. In 1992, Yugoslavia was at war after the dissolution of the Soviet Union, resulting in 700 000 refugees fleeing to Serbia. Furthermore, the Bosnian war which took place between 1192–1995 resulted in a mass displacement of 2.7 million people (which, at the time, constituted half of the Bosnian population) [26]. African countries were also not excluded from this refugee crises. Rwanda, for example, suffered a mass genocide in 1994 which resulted in more than 2 million people seeking refuge in neighbouring countries [26].

In 2003 war broke out in the Darfur region of Sudan, causing the internal displacement of more than 2.5 million people [26]. Iraq experienced various humanitarian issues from the 1980s when at war with Iran. Furthermore, in 2003 the United States invaded Iraq and the number of refugees increased dramatically. It has been estimated that approximately 4.7 million Iraqi's were displaced, more than 2 million of which sought refuge across the border in Jordan, Lebanon and Syria [26].

Less than a decade later, civil war broke out in Syria as rebel opposition fought against the Syrian government. At present, after four years of war, the conflict in Syria exists between the Syrian government, the rebels, the Islamic State of Iraq and Syria and the Kurdish. As of 2016, there were an estimated 8 million people internally displaced within Syria, whilst an additional 4 million sought refuge in neighbouring countries [2].

2.4 Current global conflict-induced forced migration

The UNHCR gathers data from reports documented by their offices (represented in more than 120 countries), as well as information from governments and partner organisations, in order to track global data on refugees, asylum-seekers and IDPs [128]. At the time of their establishment in 1951, there were an approximate 1.5 million refugees worldwide [86]. This number increased to 2.4 million in 1975 before climbing further to 10.5 million in 1985, and 14.9 million in 1990. Numbers peaked after the end of the Cold War at 18.2 million refugees in 1993. By 2000, the global refugee population had declined, once again, to 12.1 million [22, 24]. At the end of 2015, however, there existed an estimated 65.3 million forcibly displaced people worldwide, 21.3 million of which are classified as refugees, 3.2 million as asylum seekers and 40.8 million as IDPs. These data are shown geographically in Figure 2.3. In comparison to the Earth's 7.39 billion population, these numbers indicate that, on a global scale, 1 in every 113 people is now either an asylum-seeker, an IDP or a refugee [128, 129].

The primary reason for the sudden increase in forcibly displaced people which occurred over the last decade is threefold. Firstly, conflict situations which initially lead to great number of

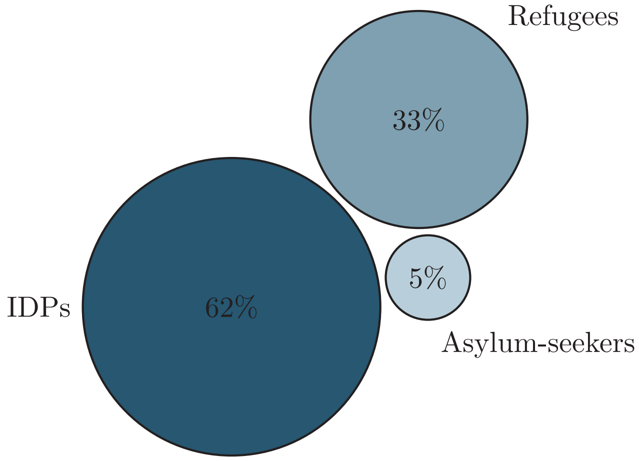


FIGURE 2.3: *The percentage of refugees, asylum-seekers and IDPs composing all forcibly displaced people worldwide, derived from the UNHCR [129].*

refugees are enduring for longer periods of time than previously witnessed (as seen in Somalia or Afghanistan where conflicts are now in their third and fourth decades, respectively). Secondly, the occurrence of new or reignited conflict situations are more frequent (*e.g.* in the last five years conflict escalated in South Sudan, Yemen Burundi and Ukraine [129]). Finally, feasible solutions on how to manage the crisis are not being proposed at a rate which is proportional to the global increase in refugee numbers. At the end of 2005, the UNHCR estimated that an average of six people are displaced every minute. Ten years later, the rate has increased to 24 people per minute being forcibly displaced [128, 129].

The UNHCR further estimates that only three countries are responsible for producing more than half of the world's refugees. These are Syria, Afghanistan and Somalia. Conflict in Colombia, Syria and Iraq caused the highest number of IDPs having approximately 6.9 million, 6.6 million and 4.4 million people displaced within each nation's borders, respectively [128]. The number of forcibly displaced people per nation worldwide as of the end of 2015 is shown in Figure 2.4. In the figure, a greater number of displaced people is illustrated by a larger and darker icon.

The increasing number of refugees arriving in Europe via the Mediterranean sea captured the attention of many over the past few years, however, this accounts for less than two percent of forcibly displaced people worldwide. Most of these individuals (more than 85%) relocated as refugees in low and middle-income countries which are in close proximity to the conflict. Major refugee hosting countries typically neighbour the country of origin. Countries surrounding Syria (Turkey, Lebanon and Jordan) account for 27% of the refugees globally, while the neighbours of Afghanistan, Somalia and South Sudan together account for another 27% [106]. Turkey is currently the country which hosts the largest number of refugees worldwide with 2.5 million refugees living within its borders. Other noteworthy refugee-hosting countries include Pakistan (1.6 million), Lebanon (1.1 million), Islamic Republic of Iran (979 400), Ethiopia (736 100) and Jordan (664 100). Lebanon, however, hosts the most refugees as a fraction of its total population, with 183 refugees per 1 000 inhabitants [128, 129].

It is estimated that women, youth and children typically account for at least two-thirds of forcibly displaced people worldwide over the last decade [106]. The UNHCR was able to gather data, which showed children below 18 years of age constituting about 51% of the refugee population in 2015. Furthermore, many of these children had been separated from their parents or were travelling alone. In 2015, there were 98 000 asylum requests from unaccompanied or separated children, primarily from Afghanistan, Syrian, Eritrea and Somalia [128, 129].

FIGURE 2.4: *Populations of forcibly displaced people worldwide as adapted from the UNHCR [129].*

2.5 Forcibly displaced Syrian

Since 2011, an estimated 5 million people have fled Syria to seek refuge in Lebanon, Turkey and beyond. In conjunction, even more Syrians are internally displaced within the borders of the country. By the end of 2015, an estimated total of 11.7 million Syrians (more than half of the Syrian population) had been forcibly displaced as a result of the civil war [64, 129]. The instability in certain areas of Syria forced families and individuals to abandon their homes to find safekeeping elsewhere. Syrian refugees typically choose one of four feasible options to find safety. These are internal displacement, encampment (*i.e.* refugee camps), self-settlement (in urban areas) or the challenging journey towards attempting to settle in European countries [15].

Figure 2.5 illustrates the displacement of Syrians on a national level, indicating the number of IDPs who settled in each region. At the end of 2015, at least 6.5 million people were internally displaced within Syria, although poor data collection owing to difficulties in monitoring the internally displaced Syrians may have led to significant under-reporting. Furthermore, certain areas in Syria also fall outside the reach of humanitarian agencies [24]. The *United Nations Office for the Coordination of Humanitarian Affairs* [131] approximates that 4.5 million Syrians are living in these areas. A further challenge which may lead to misleading statistical results is the fluctuation of Syrian IDPs, since they may be displaced more than once or choose to leave the country to seek refuge in neighbouring countries or beyond (thus no longer being classified as internally displaced) [24].

Of Syria's five immediate neighbouring countries, Israel remains inaccessible to refugees. In conjunction, Iraq, Jordan and Lebanon did not sign the Refugee Convention of 1951 and, whilst providing Syrians with protection, regards them as guests rather than refugees. Turkey is the

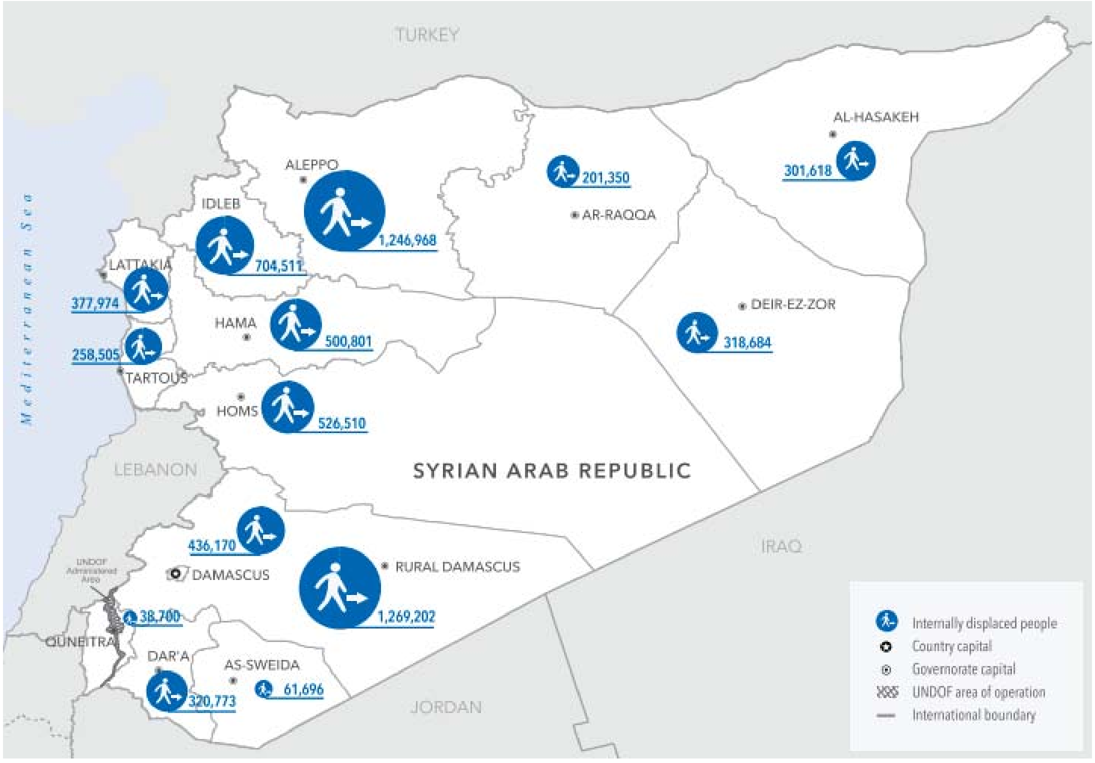


FIGURE 2.5: *A map of the Syrian IDPs per region at the end of 2015 [24].*

only signatory to the 1951 Refugee Convention, however, its interpretation of the *Convention on the Status of Refugees* only applies to Europeans seeking asylum [40]. Turkey and Jordan have set up large refugee camps for those most vulnerable, whilst Lebanon refused to allow international humanitarian aid to set up such camps [100]. Figure 2.6 shows the number of Syrians that were registered as refugees in neighbouring countries at the end of 2015.

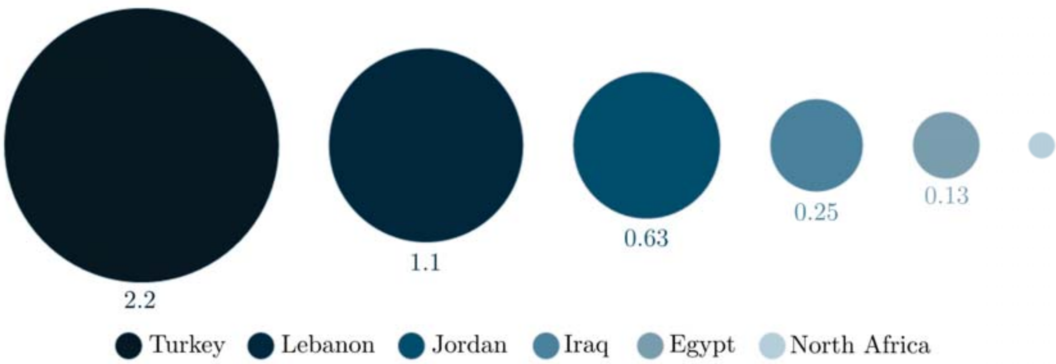


FIGURE 2.6: *The estimated number of Syrian refugees (in millions) in neighbouring countries at the end of 2015, as derived from data gathered by the UNHCR [129].*

Forcibly displaced Syrians have the option of seeking assistance in refugee camps, however, the prospects of living in these camps are dire [15]. Refugee camps are typically densely populated, chaotic settlements owing to overcrowding and the scarcity of resources [17]. These camps are mostly situated in arid regions where diseases are easily spread due to poor sanitation and housing conditions [55, 15]. Education is typically of poor quality and the economic activity within camps are restricted as refugees are not allowed to work. Research has shown that 80%

of individuals who enter refugee camps usually stay for at least five years before they are able to return to society. Considering these factors, it is understandable why only 9% of Syrians choose encampment [15]. The *Refugee Studies Centre* [100] corroborates that Syrians who are forcibly displaced prefer self-settlement to encampment.

It is further estimated that more than 60% of refugees fleeing across the Syrian border settle in urban areas where they have family or established social networks. Individuals who choose to settle in other countries without applying for asylum are not included in the United Nation's statistics, as it is assumed that such people are not in need of financial support [40]. Initially, neighbouring countries welcomed large numbers of Syrians, but, as the war intensified, these countries began restricting the influx from Syria, with some borders closing altogether [24]. Most official border crossings from Syria to Jordan and Turkey are strictly controlled, allow few admissions and migrants who want to cross over the border to Lebanon require a visa. Similarly, a visa is required for entering Turkey by sea, or by air entering into Turkey, Lebanon, Jordan and Eqypt. Apart from migration restrictions, lawful employment and poor housing in neighbouring countries may lead to urban destitution and many have to consider other alternatives [137].

Syrians hoping to seek asylum in Europe are required to physically migrate to Europe in order to do so since European Union member states closed their embassies in Syria at the start of the war. Furthermore, embassies in neighbouring countries are reluctant to process visa and asylum applications [137]. In the absence of safe, orderly and reliable pathways to European countries, migrants are often compelled to undertake perilous and circuitous journeys on land or by sea. Only slightly more than 10% of fleeing Syrians have attempted to migrate to Europe [36, 64]. The aggregated flow of Syrian refugees in January 2016 is depicted in Figure 2.7. This figure does not take exact routes followed by refugees into account [105].

Figure 2.7: *The flow of asylum seekers from Syria to European countries as of the beginning of 2016, adapted from Saarinen [105] and the UNHCR [129].*

In light of the fact that Syrians can enter Turkey without a visa, many refugees choose to travel through the country en route to Greece or Bulgaria. From here, they can enter the European Union as an ordinary traveller or illegal immigrant, depending on whether or not they have a visa. Individuals without a visa have the option of applying for asylum [40]. There also exist individuals who leave Syria via the Mediterranean sea to Greece, Cyprus, Malta or other European countries situated along coast. Most of these people enter Europe as illegal or irregular immigrants. The data available on Syrian refugees in Europe only accounts for those who formally sought asylum in a European country [40].

2.6 Chapter summary

In this literature review chapter, an overview of the concept of forced migration was provided. This phenomenon has taken place throughout history and remains topical today, particularly in the Syria, where war broke out in 2011.

In §2.1, the factors which cause forced displacement, such as natural disasters, human systems failure, conflict-based disaster and development were discussed. Furthermore, a typology of migrants, particularly forced migrants, was defined in §2.2 to distinguish between IDPs, undocumented migrants, asylum-seekers and refugees. A brief history of refugees was given in §2.3 to contextualise the phenomenon, where after, in §2.4 worldwide cases where forced displacement is currently taking place was discussed, with respect to refugees, asylum-seekers and IDPs. The state of Syrian refugees, in particular, was discussed in §2.5, highlighting the extensive global impact which the Syrian conflict has had.

CHAPTER 3

Computer simulation modelling

Contents

In this chapter, computer simulation modelling is addressed by first giving an overview of the discipline, as well as an elaboration on the different types and approaches of simulation modelling. The agent-based approach is also discussed in more detail, followed by a discussion on the use of agent-based modelling in simulating refugee movement.

3.1 Simulation modelling

In order to understand a system, the relationships between its various components, and to achieve the ability to predict its behaviour under operating changes, simulation is required [74]. Simulation is defined as the process of designing a model of a real system and conducting experiments with this model for the purpose of either understanding the system's behaviour or evaluating various strategies (within the limits imposed by a criterion or set of criteria) for the operation of the system [109]. Simulation modelling therefore provides a risk-free environment wherein a real-world problem can be addressed firstly by simplifying the system and then experimenting with the model under various conditions [19]. The real system is simplified in that all irrelevant details are omitted, whereafter the model is developed within a virtual setting. When developing a simulation model it is crucial to identify and define characteristics of the system when ascertaining a suitable modelling approach [19].

3.1.1 Types of simulation modelling

Simulation models can be characterised according to their time dependency, the nature of the input data and the continuity of the model. Law & Kelton [74] discuss these attributes and classify them according to three dimensions: *static vs. dynamic*, *deterministic vs. stochastic* and *continuous vs. discrete.*

A simulation model which is independent of time is known as a *static* model, as such a system only takes a specific time instance into account. In contrast to this, a dynamic simulation model exhibits the progress of a system over a given period of time.

Simulation models can also differ according to the predictability of its behaviour. In a *deterministic* simulation model, the input variables are known and, thus, no probabilistic components exist. It is, however, more likely that a system will involve some form of randomness which will then be constructed according to the framework of a *stochastic* simulation model. It is important to note that this randomness causes a chance variation in the output of a stochastic model, meaning it is most likely only an estimate of a system's true characteristics [74].

Finally, a simulation model can either be *discrete* or *continuous*. The former allows for state variables of the model only to be altered at certain time-steps. A *continuous* model, on the other hand, deals with continuous change in system variables with regard to time.

3.1.2 Level of abstraction

Simulation models may be categorised by the quantity of information associated with the model, known as its level of abstraction [20]. Figure 3.1 illustrates the different levels and corresponding systems which can be modelled. The amount of information decreases as the abstraction level increases. Selecting the right level of abstraction is critical to the success of the simulation model [19].

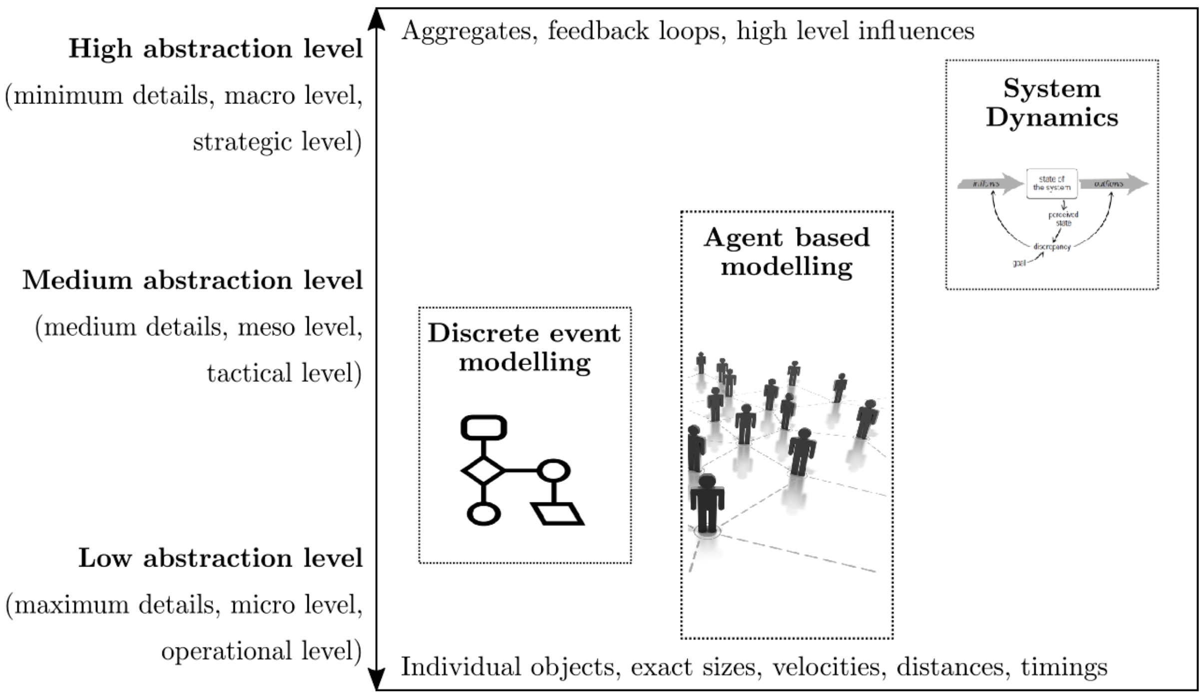


FIGURE 3.1: *The levels of abstraction relevant in modelling systems, adapted from [19].*

A high abstraction level refers to a model that has minimal detail, complementing the strategic level at which the model will be developed. The aim of such a model is to capture the essence of a global system, rather than focussing on micro details. Simulation models for social systems and economics, for example, will be highly abstract as they occur on a macro level. In systems wherein individuals form part of the whole, aggregates can be created instead of modelling each individual within a system. On the other hand, a simulation model of a tactical and operational nature would require a medium abstraction level and, thus, less detail than a model of a high abstraction level. Call centres, business processes and airports are examples of systems which would require a medium level of abstraction when modelled. Lastly, models of a low abstraction level require a large amount of detail pertaining to the system, hence encompassing more information than models of higher abstraction. These systems are modelled on a physical level, where individual entities are important to the simulation model, such as pedestrian movement, traffic, and control systems.

As indicated in Figure 3.1, the level of abstraction influences the simulation modelling approach or method employed. Three such approaches must be considered, where each is appropriate for a certain range of abstraction level. These methods are elaborated on in §3.1.3. According to Borshchev [19], discrete-event modelling operates on medium and medium-low abstraction levels, while system dynamics supports models with a high level of abstraction and, therefore, is more strategically focussed. Agent-based modelling covers the largest range of abstraction which includes medium and low levels and as such, this approach is ideal to model systems which require a great amount of detail.

3.1.3 Simulation modelling approaches

In modern computer simulation, three different approaches exist using which models can be developed for the purpose of simulating real-world systems namely discrete-event modelling, system dynamic modelling and agent-based modelling (ABM) [19]. These methods provide a general framework within which a user can build a model.

Discrete event modelling

Discrete event modelling supports a model developer with a transaction-flow world view of the system at hand. This means that the system is visualised as consisting of discrete units of traffic that move from point to point in the system while competing with each other for the use of scarce (capacity-constrained) resources [107]. Whenever entities compete for constrained resources, the existence of queues is likely. These form part of the events which take place within the system. The flow of entities through the system is visualised by means of a process flow diagram [19]. Entities can be clients, products, patients and tasks, amongst others, while resources might be doctors, workstations, staff and the like. Events are instantaneous occurrences which change the state of the system, these may include the arrival of an entity or the completion of a task [107]. Discrete-event models are stochastic as the service and arrival times of the system are drawn from a probability distribution. Due to its stochastic nature, the model requires adequate runtime or sufficient number of replications before the output thereof is of meaning to the model developer [19].

System dynamics

An endogenous view of a system is supported by a system dynamics approach to building a simulation model. This is a strategic approach which suggests a high level of abstraction. The system is modelled as a causally closed loop structure which determines its own behaviour [20]. Borshchev [19] explains this approach at the hand of an example: when considering a shop owner serving clients; as the number of people entering the shop increases, so the queue grows longer. As the queue grows, some clients may decide not to join the queue and rather leave the shop. Furthermore, some people might leave the queue due to the extended waiting time. As a result, the length of the queue impedes the growth of rate thereof. Causal loops, as such, need to be addressed in such a scenario, affording a suitable opportunity for the implementation of the system dynamics method.

Agent-based modelling

Of the three simulation modelling approaches discussed, the ABM approach is the most recent [20]. ABM allows a model developer to capture a large amount of detail within a model which is not typically achievable when employing other methods. The method suggests a bottom-up approach, as the development of the model commences with the agent and its behaviour, rather than initially considering the system as a whole . Examples of agents may include people, companies, animals or insects. Although the agent-based method of modelling allows for a greater amount of detail, it can also support a system with a higher level of abstraction. In such a case, the agents may be modelled to represent an aggregate of individuals, based on specific homogeneous characteristics, such as socio-demographics and the propensity to adapt their standard behaviour as a result of interaction with external stimuli [66]. Agents may interact with one another or be influenced by the environment. The system's overall behaviour is therefore determined by any number of concurrent individual behaviours [19].

Summary

When deciding on a modelling approach, the real-world system, as well as the purpose of the modelled system, should be taken into account. Both discrete-event modelling and system dynamics are methods that were created in the mid 1900's, while ABM is a fairly novel approach [19], offering a greater range in terms of the level of abstraction it supports. Discrete event modelling is of a medium to low abstraction, while the system dynamic approach is preferred for a more strategic modelling method. A primary difference between the agent-based and discrete-event approach is that the entities and resources of the latter are passive, whereas the agents of the ABM method are completely autonomous [20].

3.1.4 Generic steps within a simulation study

The process of designing and developing a simulation model is well documented in order to assist model developers in creating a thorough and reliable model [12]. Figure 3.2 illustrates the progression of the steps within a typical simulation model study. The framework presented is a consolidation of the research by Banks [12], Shannon [108], Law [72] and Stewart [118].

The generic steps of the framework are discussed in greater detail in the following text.

(1, 2) **Problem formulation and project planning**

The purpose of the study should first be clearly defined in terms of the project objectives [108]. A problem statement can be formulated either by the client or the model developer, but in both cases, the respective parties should both understand and agree upon the statement [12]. The project plan should include performance measures that will be used to evaluate the system, the scope of the model, the time frame agreed upon, the required resources, as well as the system configurations which will be modelled [72].

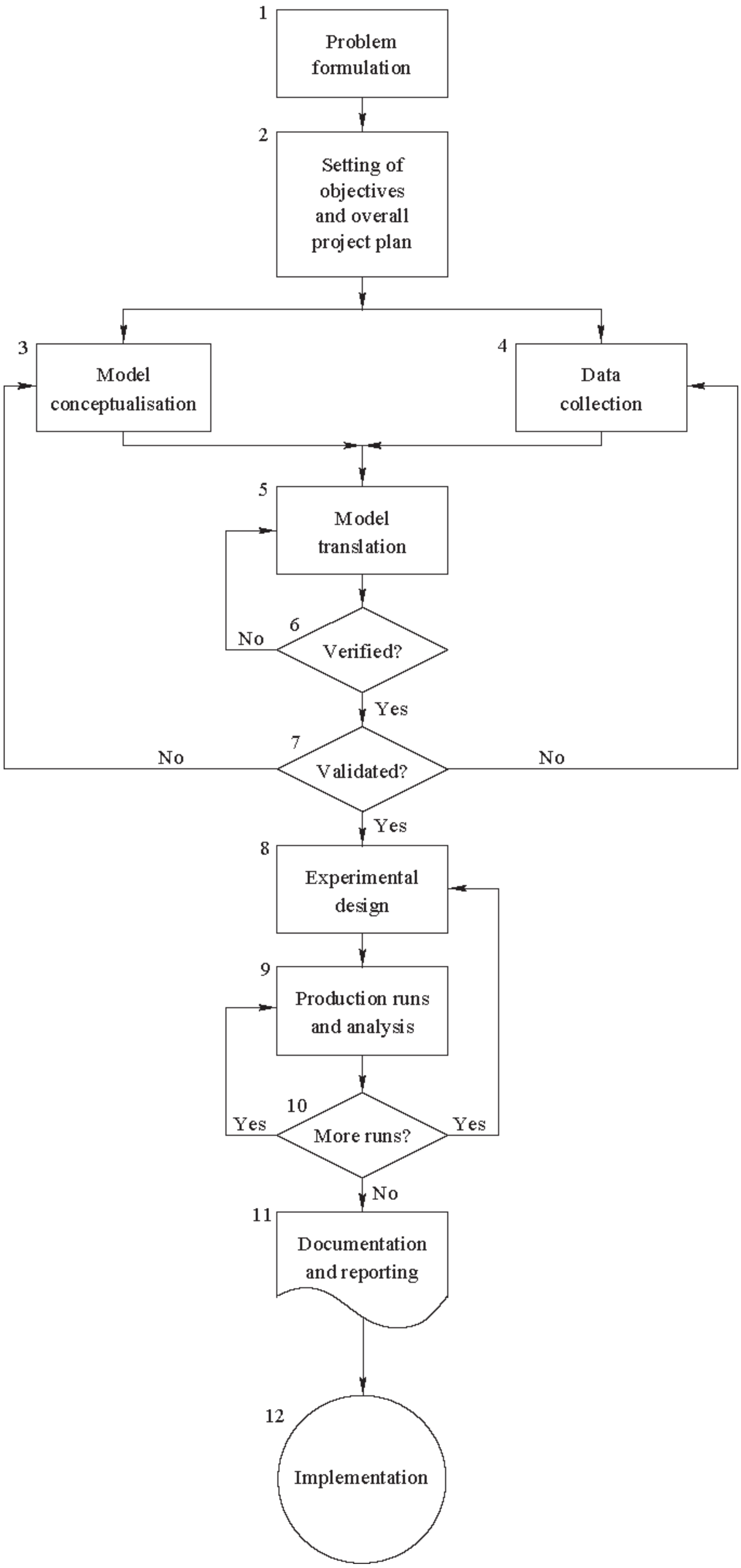


FIGURE 3.2: *The steps within a simulation model, as adapted from Banks [12].*

(3) **Conceptual modelling**

A conceptual model can be defined as a non-software description of the proposed simulation model. It should therefore comprise the objectives, assumptions, inputs, outputs, algorithms, data and clarification thereof [118]. This can be documented either graphically or by means of pseudo-code [108].

Part of designing the concept model involves determining whether simulation is indeed a suitable method in approaching the problem at hand, as well as choosing which of the modelling approaches discussed in §3.1.3 to employ [118]. Before finalising the concept model, the model developer should perform a structured walk-through for validation purposes [72].

(4) **Data collection**

Data collection and analysis occurs in conjunction with concept modelling, since contextual data is required in order to understand the system. The concept model should indicate which detailed data is necessary in terms of developing the simulation model [118]. Information regarding the system layout and its various operating processes should be collected, along with specific data regarding the model parameters and accompanying probability distributions [72]. The model developer should further ensure that the data is adequate in terms of quality, quantity and variety in order to perform reasonable analyses [12].

(5) **Model translation**

The next phase of the simulation study requires the conversion of the concept model into an operational model [12]. This is accomplished by programming the model in a commercial simulation software package, such as AnyLogic, Simio or VenSim [118, 72]. Although general-purpose program languages give the modeller better control when formatting the model into executable code, the use of specialised software packages may reduce the time spent on programming, as well as improve the efficiency and effectiveness of the model [108].

(6) **Model verification**

Verification applies to the operational model and confirms whether or not the model is performing properly (*i.e.* as expected and intended) [12]. The model developer should verify the model on a continual basis and it is advised to make use of an interactive run controller or a debugger to assist in the verification process [12]. An interactive run controller assesses the commands executed in the software and determines the success thereof, while a debugger aids in detecting errors within the model code.

(7) **Model validation**

The process of validation ensures that a model accurately represents the real-world system it depicts [12]. It allows the model developer to reach an agreeable level of confidence, suggesting that the model conjectures are legitimate and apply to the real-world system [108]. There are many methods of validating a model which are discussed later in this chapter.

(8, 9, 10) **Experimental design, production runs and analysis**

Tactical specifications such as run length, number of runs and starting conditions need to be determined for each scenario that is to be simulated [12, 72]. An experimental design must be executed in order to yield valuable information and test run specifications [108]. Once completed, production runs and the results of previous experimental designs may be used to modify input for future experiments. This is an iterative process where "what-if" analyses are often executed, results are evaluated and the next production run is adapted accordingly. The model developer may also decide to conduct additional runs of the initial experiment to check if the data set is independent, to assist in sensitivity analysis or the like [108, 72]. The key outcomes of this phase are to obtain sufficiently accurate results, search for the solution space and test the robustness of the solution [118].

(11) **Documentation and reporting**

At this point in the study all design, programming and execution of the model should be completed, as well as an analysis of the accompanying results [108]. The next step is to document the study by compiling a report which should include the conceptual model, a detailed description of the simulation model and the results obtained [72]. The document should convincingly present the outcome as clearly and concisely as possible [108]. Furthermore, the documentation should ideally include an animation of the simulation model and a proper discussion of the model verification process employed [72]. Documentation is essential for future reuse of the model with the added value of actuating the client's confidence in the simulation model [118].

(12) **Implementation**

Shannon [108] argues that a simulation study is only considered completely successful when the results are understood, accepted and used. Simulation studies may be implemented in various ways. Firstly, the simulation study's solution can be implemented in the real-world system and put into practice [118]. Secondly, the simulation study can be implemented by applying the model as a decision-support system. Finally, implementation can occur in the form of learning, such that the simulation allows for insight to be gained into the system's operation, which may aid in related future decision making [118].

3.1.5 Validation of simulation models

Validation is the process of ensuring that a model is sufficiently accurate in representing a real-world system [73, 118, 121]. A simulation model depicting a complex system can only be an approximation of the actual system, as an absolute model validation does not exist and the model is purely an abstraction and simplification of reality [72].

Stewart [118] suggests some difficulties that may be encountered during the validation process. A model can only be validated according to its objectives. Firstly, a model that is valid for one purpose may not be valid for another. Furthermore, the modelled system does not always exist in real-world, but it is essential that all models are validated, irrespective of whether or not it exists [72]. Another issue that may arise is the accuracy of real-world data, as probabilistic data is not definite. Validation should be considered as essential within the simulation model steps and it is the model developer's responsibility to ensure validation is performed.

A three-step approach to validation is discussed by Law *et al.* [73]. This include high face validity, testing model assumptions empirically and establishing whether simulation output data resembles the real system.

Face validation

A model should be developed in such a manner that people who are knowledgable in the system would regard the model, on a surface level, as being sensible and acceptable. In order to achieve face validation, the model developer should utilise current information such as the insight of experts, existing theories, the real-world system, general knowledge and intuition [121].

The model developer should approach people who are familiar with the system as their insights may be valuable to the model's validation. Further insight may be achieved by observing the real system (if it exists), although care should be taken when collecting data during observation, ensuring that it is truly representative of the system being modelled. It is also important for model developers to interact with the decision makers throughout the process of developing the model as to fully understand the thought process of these individuals.

Empirical testing

The second step in proving a model's validity, involves quantitatively testing of the assumptions made in the model building process. Observed data is often fitted to theoretical probability distributions, which are then used as input to the model. Goodness-of-fit tests, such as chi-square or Kolmogorov-Smirnov [71], can be used to test whether or not the fit is indeed adequate for a particular application to a model]. For further details on these and other goodness-of-fit tests, see Law [71]. Another useful tool is a sensitivity analysis, as it identifies parameters that have a greater influence on the output of a model. Through this, the model developer can identify certain parameters toward which the output of the model is sensitive and obtain a better estimate of these parameters. In such a manner the model will more accurately represent the real-world system, which would add to its validity.

Comparison of output data

The final step is to establish whether or not the model's output accurately depicts that of the concerned system. This is a conclusive step in validating a simulation model. The model developer should gather output data from the real-world system and compare it to the associated output of the model. If the comparison is favourable, the simulation model of the existing system is considered 'valid' [72]. The greater the commonality between the real-world and the simulated system, the greater the confidence in the simulated model. The model developer will never know if a system is completely valid, as its level of accuracy still remains an estimation, however, this step in the validation process will give confidence and credibility to the model.

3.2 Agent-based modelling

In its simplest form, ABM entails a system of autonomous and interacting agents [18]. It is a dynamic simulation technique where agents, embedded in and acting with the environment, possess the capacity to learn and adapt in response to changes in other agents and the environment [6]. The ABM technique incorporates a so-called 'bottom-up approach'. Within an agent-based simulation model the global characteristics of the system are never defined, while agent behaviour is specified on an individual level [20]. The global behaviour thus evolves as a result of collective individual decisions made by agents. Consequently, the model is able to explain macro-level phenomena developing from the micro-level interactions [53].

ABM allows for the modelling of social systems, as it is able to simulate the manner in which individuals or groups interact and adapt to changes in the environment or due to interaction with others [79]. A broad span of ABM applications exist and the discipline in which ABM is applied to varies. This can range from modelling consumer markets [44], to generating social networks [8] and modelling the diffusion of innovation [66], from coordinating supply chain and inventory control [61] to modelling organisational science [42] and more.

3.2.1 Advantages and disadvantages of agent-based simulation

Simulation in itself carries many advantages, as well as some notable disadvantages. The advantages and disadvantages of simulation in general are discussed briefly in this section, before referring specifically to those of ABM simulation.

The most integral benefit of simulation is its ability to investigate possible solutions and situations in a virtual reality without the risks associated with it in real life, such as committing resources or disrupting ongoing operations [19, 108]. Apart from the risks, the virtual time allows the model to run a simulation, which would have taken days or months in reality, in mere minutes [108]. The simulation can run much faster than real time; this enables the model developer to gather results of the system's performance much quicker than in reality [118, 108]. Furthermore, the simulation affords the model developer an animation of the system which is beneficial in conveying the operations or solution to the client [19, 108].

Despite the vast advantages of simulation, some drawbacks do exist. The process of developing an authoritative simulation model takes time and requires a specific level of skill and experience from the model developer. Furthermore, even though the simulated world is cost-free, specialised simulation software can be expensive [118, 73, 108]. In the same cases, although a model could be built accurately, the model developer must be able to interpret the results of the constructed model in order for it to have any significance [73]. Another challenge for the model developer is acquiring reliable data, however, this time consuming process should be handled with care for the model's credibility depends on the reliability of the model [108]. The credibility of a model is important and it requires proper validation and verification [6]. This is coupled with the limitation of reliable data for input to depict the real-world system accurately and the challenge of having sufficient output to compare with reality [108].

Advantages

The ABM method in particular also carries a number great advantages when chosen as a simulation modelling approach. According to Borschev and Filippov [20], the ABM technique is more general and powerful than the other existing methods, as it effectively captures real-world systems with their complexities and dynamics. Bonabeau [18] summarises the benefits of ABM in three aspects: the manner in which it captures emergent behaviour, ABM's ability to model the nature of a system and the flexibility of the tool.

Emergent phenomena display characteristics of individual behaviour that is non-linear, exhibiting memory or pat-dependence (amongst others) or when the interactions between agents are heterogeneous [18]. Such instances are challenging to model, unless in an agent-based environment, because ABM allows the model to imitate the behaviour as such.

Another favourable aspect is the fact that agent-based simulation models can be constructed in the absence of information on the system as a whole, by using the details on individual level. This aspect also makes it easier to maintain the model as changes and refinements to the system will occur locally and not on a global level [20].

Disadvantages

As with many modelling approaches, certain impeding factors apply. One such factor that is often mentioned in literature, is the difficulty of validating and verifying agent-based simulation models [6, 66]. Before commencing the ABM approach, it is the model developer's responsibility to ensure that adequate data are readily available for the purpose of validation.

In addition, Garifullin *et al.* [44] advise the model developer not to overreach in adding details to the model, as it would result in a model with too much sophisticated logic per agent. This in itself is not detrimental, but, as the detail increases, so the complexity thereof may cause difficulties due to computational intensity of ABM and a lack of adequate data.

3.2.2 Components of an agent-based model

Macal and North [79] speak of agent-based models as consisting of three main elements, namely a set of agents, their attributes and behaviours; a set of agent relationships and methods of interactions; and the environment in which agents are embedded. Others, such as Jennings [62], assert that the main elements include agents, global interactions and organisational relationships, as depicted in Figure 3.3.

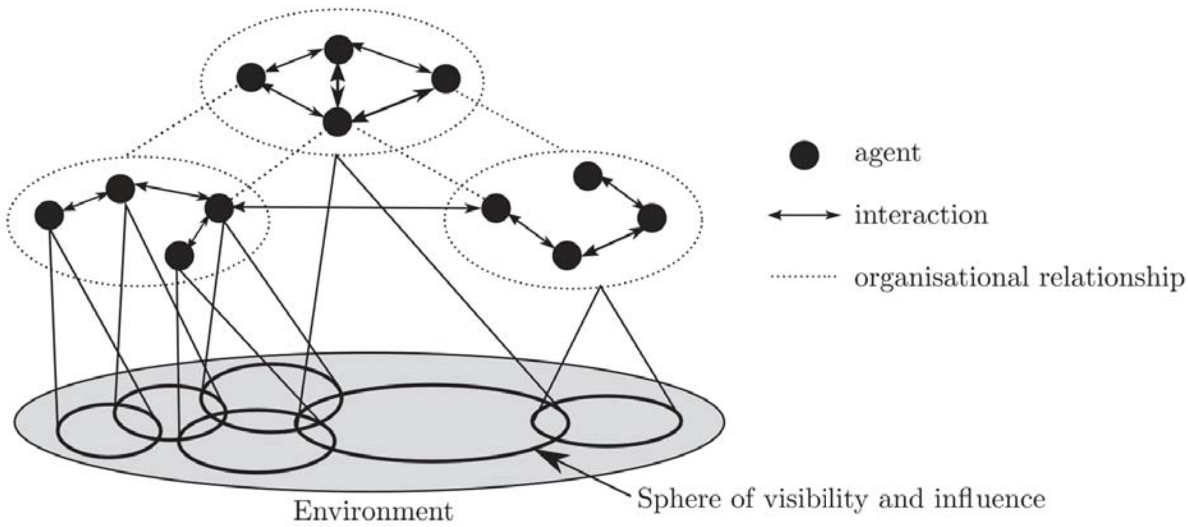


FIGURE 3.3: *The essential concepts of ABM as adapted from [62].*

All of these aforementioned components compose the structure of an agent-based model. The model developer should identify, plan and define these elements in order to construct the simulation model [79]. Agents, their interactions and the environment will now be elaborated on further.

Agents

Beyond the fundamental property of autonomy, the definition of an agent has not been formally agreed upon in literature. Some authors emphasise that an agent should mainly be defined by its autonomous properties [18], while others argue that an agent is adaptive and defined by its ability to learn from experience. Macal and North [80] defined an agent as a component that can learn from its environment and change its behaviour in response to its experiences.

The essential feature of an agent is its ability to act autonomously, making independent decisions as a self-directed individual and also its capability to exist independently in its environment [62, 79, 80]. Furthermore, agents possess the following characteristics:

- Agents are adaptable and uniquely identifiable as individuals [6, 62, 79, 80]. Each agent has a set of characteristics which determine their behaviour and decision-making abilities [80].

- Agents have specific design objectives [62] and are goal-orientated [62, 79, 80]. For this reason, agents can compare the outcome of certain behaviour with its objectives and alter future behaviour in order to achieve its goals [79].

- The state of an agent (*i.e.* its subset of attributes) changes over time and affects the behaviour of an agent. Many possible states will ensure a rich set of possible behaviours that an agent can possess [79].

- Agents are dynamic and social individuals which interact with the environment and other agents, therby impacting on the behaviour of these agents [79, 80]. Certain of rules conduct exist with regards to interactions with other agents in terms of communication, movement, competing of space and the like [79]. Agents are capable of recognising the characteristics of other agents [79].

- Agents are flexible problem-solvers in pursuit of achieving their design objectives [62]. They are capable of learning from previous experiences in order to adapt their behaviour as their circumstances change [80].

A typical agent and its interactions is illustrated in Figure 3.4. Agents have certain attributes and methods associated with them. Agent attributes can be of either static or dynamic nature [79]. The former refers to attributes which cannot be changed during the simulation, such as an agent's name, whereas dynamic attributes can change as the simulation progresses (*e.g.* the memory of an agent as it records past experiences) [79].

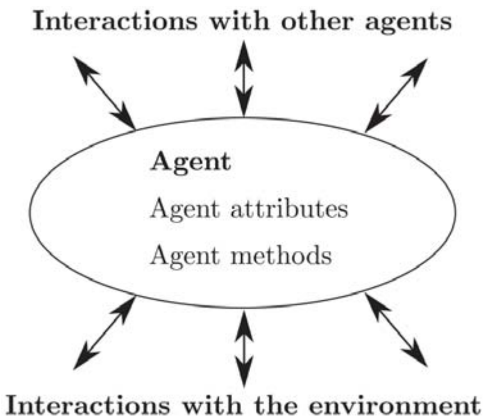


FIGURE 3.4: *A typical agent and its interactions as adapted from [79].*

Agent methods refer to the behaviour of the agent, these specifications can be in the form of simple rules or complex models (such as neural networks or genetic algorithms) [18, 79]. In order to model an agent's behaviours, a theory of agent behaviour for circumstances which may occur in the model is required. If such behavioural theory does not exist, a normative model, where the agent is required to optimise its design objective, can be utilised as a starting point [79]. Even though this is a simple model, the conclusions made from this model may assist in developing an elementary, more detailed (although sensible) heuristic behaviour model [79].

In conclusion, agents are independent decision-makers, this allows them to be dynamic entities, rather than mere passive components in the simulation model [80].

Agent relationships

An agent-based model should be able to imitate the interactions between agents, which are both repetitive and competitive [18]. Aspects to address when modelling these relationships are mainly to identify which agents could interact with one another and, secondly, to specify the dynamics of these interactions [79]. This needs to be taken into account when developing the simulation model.

Agent-based systems are decentralised and thus no central authority exists, but information rather spreads via interactions [79]. An agent would typically interact with its neighbouring agents and not the entire population (as is the case in real-world systems). Local information can be obtained by an agent through interacting with either its neighbours or its immediate environment [79]. The interaction with other agents may be crucial for an agent in achieving its objectives. Furthermore, agents interact with one another to manage relationships formed in the common environment [62].

As agents move throughout the simulation, their neighbours can change rapidly. In order to model agent relationships, a topology of agents need to be determined. Topology in ABM can be defined as the manner in which agents are connected (*i.e.* who transfers information to whom) [79]. The topology chosen to model the agent relationships can significantly change the outcome of the model in comparison to the predicted corporate behaviour [18]. Therefore, choosing the right topology is necessary in order to model the social influence of the system accurately [66]. Typical topologies include: spatial grids, networks of nodes, links and the like [79].

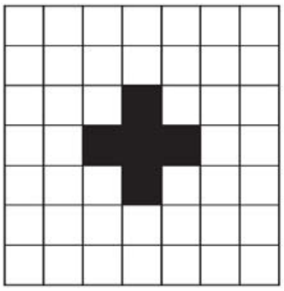
Figure 3.5 illustrates five different topologies used in ABM, as defined by Macal and North [79]. The cellular automata allows agents to move from one cell to the other on a grid, where no more than one agent can occupy a cell at a time. The von Neumann ‘5-neighbour’, as shown in Figure 3.5a, is an example of one of the many cellular automata neighbourhoods which exists. The Euclidean space model allows agents to move in two, three or higher dimensional spaces (Figure 3.5b). The geographical information system (GIS) topology shows a real geographic landscape which is divided in patches, as illustrated in Figure 3.5c, where an agent can move between these patches.

Figure 3.5d refers to a network topology, which is a more generally defined agent neighbourhood. Networks can be of either static or dynamic nature. The links between agents are set and cannot be altered in a static network. In a dynamic network, however, the links and nodes are endogenously decided on [79]. Lastly, in the ‘soup’ model depicted in Figure 3.5e, location is irrelevant and agents interact by means of a random selection [79].

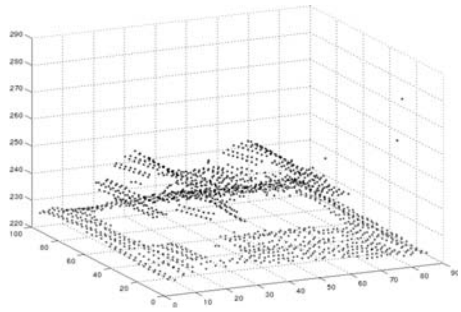
The environment

Agents within an agent-based model interact with and are enclosed in a dynamic environment [6]. This environment provides information to the agent which can vary from simply indicating an agent’s location relative to other agents, or by giving geographical information in the case of GIS [79]. The location of an agent is a dynamic attribute of the agent, used to track its movement.

An agent can learn from its environment and adapt as per necessity when changes occur [6]. The environment may also create certain spatial constraints for the agents. A transportation model is an example of such, as it would contain infrastructure and capacity constraints in the links and nodes which depict the road network [79].

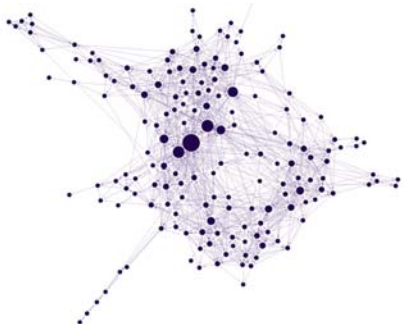


(a) *Cellular Automata*

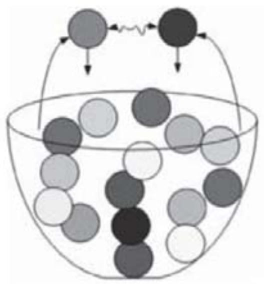


(b) *Euclidean 3D space*

(c) *GIS topology*



(d) *Network topology*



(e) *'Soup' model*

Figure 3.5: *The various topologies in ABM.*

3.3 Agent-based modelling with regards to forced displacement

Migration is a complex process which encompasses certain movement patterns, historical context and individual decision-making [36]. In the light of this, a likely challenge when investigating the movement of people via simulation is the availability of sufficiently high resolution, informative data. Hailegiorgis and Crooks [55] propose the integration of aggregated statistical data along with the review of relevant literature in order to accurately capture the complex process of migration within a modelling context.

By using ABM, agent behaviour shows potential to be explored in an attempt to better understand the various processes and consequences concerned with refugee movement. Such a model, once sufficiently developed and validated, could possibly then aid as a decision support tool for humanitarian relief [31, 51]. ABM is particularly well-suited for such applications for a number of reasons. Firstly, the spatial environment in which agents exist in ABM endorses the study of the population in terms of evacuation or movement which can be measured according to each individual agent's proximity to the disaster. Furthermore, the heterogeneity of agent control recognises that a person's decisions may differ according to certain personal attributes. For example, women with children might be more vulnerable in crisis situations than men and may act differently as a result. Finally, ABM allows for a certain degree of stochasticity to be incorporated in the model, with respect to both the environment and the agent, which assist in replicating realistic scenarios [31].

3.3.1 Types of movement modelling

Groen identifies three different manners in which human movement can be modelled, namely evacuation modelling, refugee modelling and migration modelling [51]. The difference of these movement modelling types lies in the variation of the temporal and spatial scale. Evacuation modelling occurs within seconds or hours, whilst migration intervals may comprise months or even years. Similarly, the spatial area considered in evacuation modelling typically varies within a few meters to a kilometre, while migration modelling could develop over an area of hundreds of kilometres. As may be seen in Figure 3.6, the refugee modelling domain falls between the evacuation and migration movement modelling types.

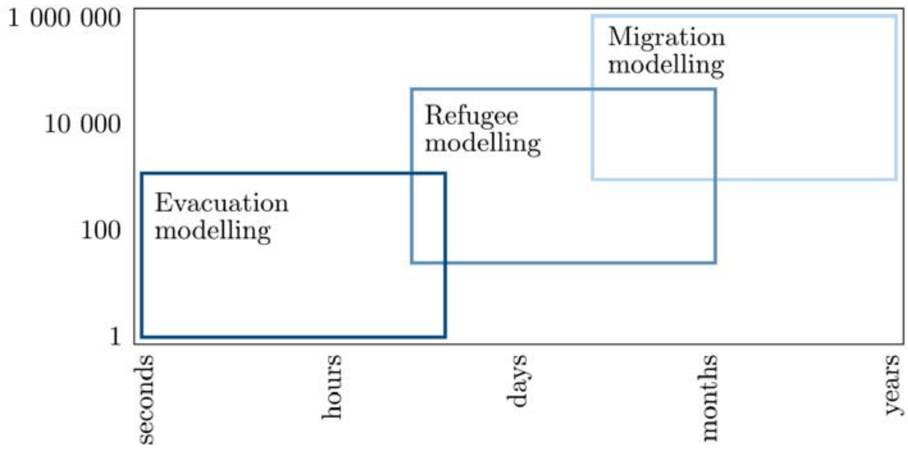


FIGURE 3.6: *The types of movement modelling as discussed by [51].*

Evacuation modelling has been used in literature to model pedestrian dynamics, crowd movement and urban evacuation [68, 141]. The spatial area in these cases refers to a city or a specific location within a city, and the evacuation time is typically simulated in seconds or minutes.

Migration modelling is widely considered in literature [14, 49, 96, 116], mostly as voluntary migration. Smith *et al.* [116] modelled a particular migration which occurred over a period of thirty years. In this model, migration was driven regionally and internationally due to climate change. Voluntary migration is driven by socio-economic factors, while refugees are typically forced to relocate for their own safety [102].

Forced migration modelling (*i.e.* refugee modelling) is unique in its spatial and temporal scale as it may span kilometres across the globe (as can be seen with Syrian refugees seeking safety in Scandinavian countries), while occurring within a matter of days or months [51]. Individuals forced to migrate can relocate within a country's border (*i.e.* IDPs) or flee across international borders (as discussed in §2.2).

3.3.2 Existing forced migration models using ABM

Agent-based models depicting refugee movement (although these studies are mostly published within the last decade) do exist in literature [7, 29, 31, 37, 51, 55, 116, 117].

In one such study, Collins and Frydenlund [29] investigated strategic group formations which exists among refugees when migrating. ABM methods were incorporated in this investigation, along with game theory methods. This model assumed that refugees, when travelling over long distances, attempt to form groups which provide them with security. Each individual was modelled according to an internal utility function which was based on the speed and size of the group, to determine whether or not the individual would find greater benefit in joining or leaving a specific group.

In another study, ABM was used, in conjunction with crowdsourced data and GIS data, to simulate the situation after a natural disaster in order to explore people's reactions to aid [31]. Geographic information pertaining to population density, existing transportation networks and aid centre locations was included as an input to the model. Crooks and Wise [31] demonstrated the use of data-rich ABM and further suggested the use of geographic explicit ABM in mapping future trajectories of similar events.

Anderson *et al.* [7] utilised ABM as a tool for simulating humanitarian assistance policy decisions executed by authorities who provide for the health and safety of refugees. The model was used to test various concepts and strategies by examining the response of refugee communities and the consequent affect on their health and well-being.

Agent-based models can also be used to simulate the spread of diseases, like cholera, within refugee camps, as shown by Hailegiorgis and Crooks [55]. The complex interaction of people and their environment were modelled in order to better understand the dynamics of cholera transmission within refugee camps. Factors influencing the spread of disease, such as a person's social behaviour and movements, surface elevation and rainfall were all taken in the model.

Sokolowski *et al.* [117] performed a study on population displacement within the Syrian city of Aleppo using ABM to characterise individuals, entities and the environment by means of integration between quantitative and qualitative data. The purpose of such a model is to provide a means to anticipate, measure and assess future displacements of a similar nature which might occur.

A simplified agent-based model has been developed by Groen [51] to explore the movement of displaced persons towards refugee camps, with particular focus on refugees during the Northern Mali conflict in 2012. A more complex approach was undertaken by Edwards [37] as counter-intuitive or 'aberrant' patterns of flight were considered in predicting spatial flow of conflict-induced migrants. According to Edwards, with the use of ABM and knowledge of the appropriate input variables, researchers ought to be able to estimate the extent of the resulting displacement, the likely destination choices of the displaced, refugee population sizes, as well as likely places where IDPs will settle.

3.3.3 Determinants of forced migration

A forced migration model should address several broad questions, such as 'Who migrates?', 'Why do the people migrate?', 'Where do these people migrant from and where do they migrate to?', 'When do they migrate?' and 'What are the consequences of the movement?' [49]. ABM provides model developers with the ability to simulate the autonomous decision-making of agents with regards to these determining factors.

Alhanaee and Csala [5] performed a study on the motives of Syrians when choosing to seek refuge in Lebanon. By means of a gravitational migration model and data from both social media and the UNHCR, the study concluded that a person's motives to flee change over time and, to some extent, as the person moves further from the source of danger. For example, when confronted with a life-threatening situation, a person might flee to the nearest and safest place without much consideration as to what the place would offer, other than safety. Once in a safer place, however, an individual's propensity to move might be influenced by other elements such as the shortage of resources or disruption of public services. Alhanaee and Csala [5] inferred that the primary determinants of forced migration were protection, economic status of the current location, health, education, family size and age.

In a study by Moore and Shellman [88], an individual's preference towards migration was determined by a certain threshold value. Assume there exist a probability, p, that a person is going to be forcibly removed, where $p \in [0, 1]$. As p increases from 0 to 1, there exists an inferred threshold value which, when exceeded, results in a person preferring migration over staying. Many factors exist that influence an individual's p-value. These include governmental conflict, culture, family ties, religion and the behaviour of politicians in the country. The study concluded that the primary determinant of forced migration was violent behaviour of both the government and dissidents. Individuals tend to move to the nearest location where violence can be avoided, typically choosing to relocate to places where other refugees have gone before [32, 90].

In attempting to predict displacement, or a person's inclination to migrate, the characteristics of the individuals, the environment and the relevant events which trigger displacement need to be considered. Parameters which model intricate considerations (linguistic homogeneity, wealth inequality and ethnic chasms), as well as more tangible influences (conflict intensity, spatial spread of the conflict and geographical features) should also be considered where possible [37]. In addition, Edwards [37] recommends the push-pull and the network analysis approaches, both of which form part of migration theory. The push-pull approach considers the initiation of migration, whereas network theory relates to the perpetuation thereof [54].

The push-pull approach explains the motivations of an individual to move. The 'push' factor provides a person with causal motivations to leave their place of residence, while the 'pull' factor of the chosen destination prompts the person to migrate [70]. The decision of an individual to move and their choice concerning where to go, is formed by a unique combination of circumstances. Smith *et al.* [116] acknowledges the push-pull effect as intervening obstacles which influence an individual's migration aim.

'Push' factors, with regards to forced migration, include conflict, insecurity and persecution, while 'pull' factors pertain to employment opportunities, established social networks, perceived economic and political conditions and education for children [9, 14, 33]. Apart from push-pull factors, Edwards [37] mentions that the decision-making of an individual will further be influenced by the social or psychological effects of trauma (caused by the conflict situation), as well as other individualistic behaviour. A broader range of considered factors will allow for a model to posses a more accurate predictive value.

In another study by Moore and Shellman [89], the process of becoming a refugee or an IDP was modelled as consisting of two stages. The fulfilment of these stages was determined by the person's perception of victimisation and socio-politico-economic opportunities. In the model, an individual must first decide whether or not to move from their place of residence. This is determined as a function of the person's perception of victimisation (in their place of residence). The second stage entails the person's decision on where to move. This decision is a function of the expectation of victimisation and of opportunities which might exist in other locations.

Another historical measure of motivation often used is Maslow's *hierarchy of needs* [84], which was developed as a theory of human motivation comprising of a five tiered model depicting hierarchical levels of motivation. Maslow argues that people are motivated to achieve certain needs, but that some needs take precedence over others. The hierarchy of needs is illustrated in Figure 3.7.

The most basic need identified is that of physical survival. The first and second tier, *physiological* and *safety* needs, respectively, form part of this basic need to survive. Psychological needs, tiers

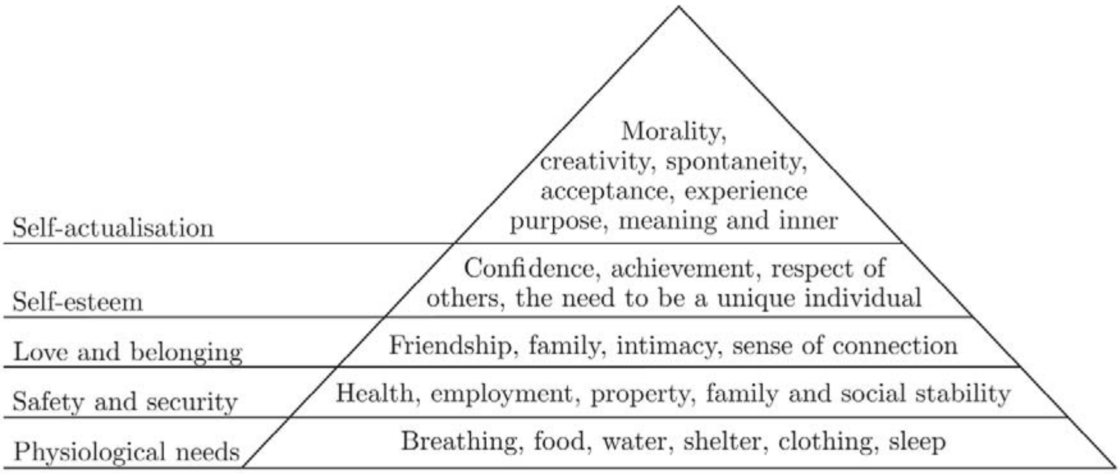


FIGURE 3.7: *The hierarchy of needs by Maslow [84].*

three and four, are *belongingness and love* needs, as well as *esteem* needs. The final tier represents *self-actualisation.* According to Maslow, all people strive towards self-actualisation, but this can only be achieved once all of the basic and psychological needs are met. This hierarchy of needs has been considered in literature as determinants of migration [5, 31].

3.4 Chapter summary

Computer simulation modelling, in particular ABM, was discussed within this chapter. More specifically, ABM was considered with regards to the modelling of refugee movement.

Simulation modelling was explained in §3.1, including aspects such as the different types of modelling, the various level of abstraction and different modelling approaches. The generic steps of a simulation study were detailed in §3.1.4, where after various methods used to validate simulation models were considered in §3.1.5. In §3.2, ABM, its advantages and disadvantages, as well as the various components that encompass an agent-based model were discussed. Along with the background of simulation modelling and ABM, the application of ABM in forced migration studies was reviewed in §3.3. The methods with which different kinds of movement are modelled were considered, as well as existing studies on forced migration modelling by means of ABM and the factors leading to forced migration.

CHAPTER 4

Decision-making

Contents

Influences and considerations in human decision-making are studied within this chapter. The categorisation of decision-making methods is discussed in §4.1, before the consideration of multi-criteria decision-making methods in §4.2 and performing a general study on human decision-making in §4.3. Finally, the application of human decision-making theory, with respect to modelling the decision-making of forcibly displaced people, is discussed in §4.4.

4.1 Decision-making as a field of research

Wang and Ruhe [138] define decision-making as a cognitive process where a preferred option is selected from a set of alternatives, based on a given set of criteria. The field of decision-making encompasses various decision theories, primarily categorised into *prescriptive* and *descriptive* theories [34, 35, 119, 138].

Prescriptive (or normative) decision-making theories provide guidance to decision-makers in order to make the most rational decision which would achieve some optimal benchmark outcome [119]. Prescriptive decision scientists are thus concerned with prescribing methods which are anticipated to result in optimal decisions [35]. The behaviour suggested by prescriptive decision-making differs from human behaviour, as it is rational and based on the strong theoretical foundation of normative theory, such as the Pareto optimal and efficiency [25, 34]. Prescriptive decision-making methods are typically used in the operations research and management sciences, where an algorithm offers some optimality. A classic example is the *Travelling Salesman Problem*, where a travelling salesman is required to visit a number of cities, without backtracking on the route and visiting each city exactly once, while minimising travel costs. The minimum cost, in this case, may be defined as the shortest route [13, 104].

Descriptive decision-making theories, on the other hand, imitate the manner in which people make decisions [119]. Such theories consider the actual behaviours and decisions made by humans in an attempt to describe this decision process [13]. Descriptive decision-making is described in this section by means of an example explored by Saad [104]. Consider the case where four cars, as given in Table 4.1, are scored based on four attributes, namely price, quality, safety and miles per gallon, on a scale from one to seven. One, in this case, is regarded as the worst and seven as the best. Each of these attributes have been allocated a weighting to indicate its importance. A cut-off value per attribute is also given. These cut-off values represent aspirational levels which suggest the minimal level at which an alternative is considered acceptable.

TABLE 4.1: *An example within descriptive decision-making as presented by Bell et al. [13].*

	Price $weight = 0.50$	Quality $weight = 0.25$	Safety $weight = 0.20$	Miles/gallon $weight = 0.05$
Car A	7	5	2	6
Car B	1	6	4	6
Car C	6	5	4	3
Car D	5	3	2	1
Cut-off	5	4	3	2

Given the selection of cars provided in the table, one car needs to be chosen. There are various ways in which to arrive at a winning alternative. For example, when implementing the *Lexicographic* rule, the alternative which scores the best on the most important attribute, which in this case is the price, will be chosen. Car A would thus be the best alternative according to the Lexicographic rule, as it received a score of seven (the highest) within the price category.

Another method that could be implemented is the *Satisficing* rule, which suggests that the first alternative to pass all of the minimum cut-off values should be chosen. From this table, Car B fails as an alternative, as it scores below cut-off in the price category. Similarly, Car A's score in the safety category is below the cut-off and is thus also eliminated as an option. Car C, however, succeeds the cut-off values in all categories and should, according to the Satisficing rule, be chosen as the best alternative. In this case, the descriptive decision rule used leads to completely different results. In other cases, however, different rules may still agree on alternatives.

There exists a significant gap between prescriptive and descriptive decision-making methods [34, 119]. Prescriptive modelling is concerned with optimality and to solve a problem by minimising or maximising some metric, whereas descriptive modelling simply describes human thought processes when making decisions [104]. The categorisation of decision-making into descriptive and prescriptive models, however, aids in understanding the field as a whole and various contributions toward explaining decision-making have been made by Churchman and Ackoff [27], Gigerenzer and Gaissmaier [46], Simon [113], and Wang and Ruhe [138].

Simon [113, 114] introduced a three-phase design process composed of the phases *Intelligence*, *Design* and *Choice*, depicted graphically in Figure 4.1. The first phase, namely intelligence, refers to the process in which the environment is explored and the need for improvement or change is identified. The design phase, following the realisation of such a need, investigates the problem and possible alternatives, thereby broadening the problem domain. In the final phase, the most appropriate alternative amongst the alternatives generated in the previous phase is chosen [34].

The flow between these phases is complex. For example, further intelligence may be required in the design or choice phase. Furthermore, within any phase, a new decision process may exist containing the three-phase decision making process within the existing phase itself [34].

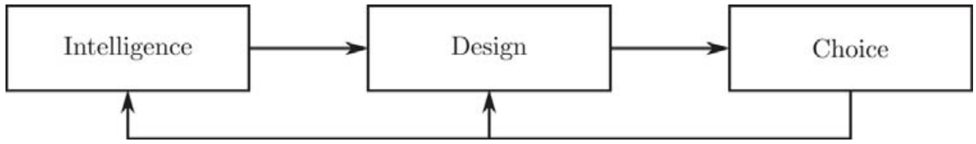


FIGURE 4.1: *The three phase decision making process proposed by Simon [114].*

4.2 Multi-criteria decision-making methods

The study of *multi-criteria decision-making* (MCDM), otherwise known as *multi-criteria decision aid*, relates to generic planning problems where the most attractive solution needs to be identified from a given set of alternatives. Each of these alternatives is characterised by scores given for a set of selected criteria. Furthermore, a set of interest groups' input is considered with respect to the selection criteria and associated relative importance. MCDM methods are used to define the attractiveness of numerous alternatives with the aim of identifying the most attractive solution [85]. As many conflicting priorities are considered in making the decision, there is no longer one optimal solution, but rather a set of satisfactory solutions [52, 136].

MCDM is not restricted to a specific field, but applies to all branches of operations research. Further, it can be prescriptive or descriptive in nature and may apply to many real-life problems [136]. Very few problems have a single criterion determining the choice of solution and decision makers are typically confronted with multiple objectives, attributes and criteria [140].

According to Guitouni and Martel [52], the MCDM methodology is a non-linear, recursive process which consists of four steps: (i) structuring of the decision-making situation, (ii) preference articulation and modelling, (iii) aggregating the alternative evaluations and (iv) making recommendations. Preference modelling (step ii) is an essential element of the MCDM methodology as the decision maker's preferences influence the resulting recommendations. This concept is discussed later in this chapter. The crux of a MCDM lies within the multi-criterion aggregation procedures (a combination of steps ii and iii), as considered in this section.

4.2.1 Classification of MCDM methods

There are various manners in which MCDM methods are classified in the literature, as discussed by Guitouni and Martel [52], Huang *et al.* [57], Hwang and Yoon [58], Malczewski [82], Massam [85], Mota *et al.* [91] and Vincke [136]. Only the classification methods most relevant to this study are considered in this literature review.

The first classification entails categorising MCDM methods as continuous or discrete, depending on the set of alternatives [30, 91, 92, 139]. Hwang and Yoon [58] refer to these categories as *Multi-Attributive Decision-Making* (MADM) and *Multi-Objective Decision-Making* (MODM) methods. Both methods involve a set of alternatives which are evaluated based on conflicting criteria. These criteria include attributes (the measurements of the system with regards to an objective) and objectives (the desired state of the system under certain conditions, derived from attributes) [38]. Table 4.2 lists criteria distinguishing MADM from MODM methods.

MADM methods include decision-making situations where a finite and feasible set of known alternatives exist and are specified in terms of attributes. These attributes are regarded as both decision variables and criteria. MODM methods, on the other hand, encompass a large number of feasible alternatives and the objectives (which also act as constraints) functionally relate to the decision variables. MODM methods therefore define the set of alternatives (from a larger set

TABLE 4.2: *The comparison of MODM and MADM approaches as presented by Malczewski [82].*

Distinguishing criteria	MODM	MADM
Criteria defined by	Objectives	Attributes
Objectives defined	Explicitly	Implictly
Attributes defined	Implicitly	Explicitly
Constraints defined	Explicitly	Implictly
Alternatives defined	Implicity	Explicitly
Number of alternatives	Infinite (large)	Finite (small)
Decision maker's control	Significant	Limited
Decision-modelling paradigm	Process-orientated	Outcome-oriented
Relevant to	Design/search	Evaluation/choice

of initial alternatives) from which a best solution is chosen, whereas MADM methods commence with a smaller, known set of alternatives [38, 85].

MCDM methods can also be classified according to the quality of available information in light of the complexities in modelling real-world problems which often involves imperfect knowledge and evaluations from humans. It is suggested that one should categorise information as *Crisp* (pertaining to precise data) or *Fuzzy* (pertaining to incomplete or vague data). MCDM methods can then be further divided into MODM/MADM methods when utilising crisp data, or into *Fuzzy MODM*/*Fuzzy MADM* regarding fuzzy knowledge [91].

Another classification method utilised by specialists is to classify MCDM methods according the following three categories: (i) *multiple attribute utility theory* (MAUT), (ii) *outranking methods* and (iii) *interactive methods* [52, 91, 103, 136].

MAUT is a unique synthesis criterion approach without incomparability which aggregates the different perspectives or point of views into a unique utility function. This can then be assessed to find the best solution [3, 136]. It therefore transforms the scores given at any level into utility functions, based on the preferences of an unitary decision maker. The MAUT approach facilitates rational decisions as the choice, with highest expected utility also being the preferred alternative [57].

Outranking methods accept incomparability among alternatives and develop a relation which represents the strongly established preferences of the decision maker in order to choose an alternative [136]. The overall attractiveness of each alternative is determined by weighting the scores given to specific dimensions relative to other alternatives. In opposition to MAUT, the calculated scores are utilised to steer a deliberative process amongst multiple stakeholders, rather than identifying a single solution [57].

Interactive methods utilise an collective local judgement approach with trial-error iterations [52]. This approach alternates between the computational steps and dialogue with the decision maker. Based on the reaction and preferences of the decision maker, the model is adapted in an attempt to build a solution [136].

4.2.2 Preference modelling

Many preference relations exist, but this section only considers those utilised in MCDM methods. The preference structures, as discussed by Colson and De Bruyn [30] and Guitouni and Martel [52] are therefore considered. Assuming a decision maker must compare two alternatives represented by a and b, four possible preferences can result in the form of the following binary relations:

Indifference situation (a **I** b): The decision maker is indifferent between alternative a and alternative b. There is thus no evidence that alternative a is preferred to alternative b, and vice versa.

Preference situation (a **P** b): The decision maker strictly and strongly prefers alternative a to alternative b, without any hesitation.

Weak Preference situation (a **Q** b): The decision maker strictly and weakly prefers alternative a to alternative b, although there is hesitance between the indifference and preference situations.

Incomparability situation (a **R** b): The decision maker is hesitant between a **P** b and b **P** a. These alternatives are seen as incomparable. This may be due to a lack of information.

These elementary preference relations may be combined with the use of logical operators to form performance structures (*e.g.* 'a $(\boldsymbol{P} \cup \mathbf{I})$ b' which combines the preference and indifference situations). An outranking relation, **S**, is obtained by combining **P**, **Q** and **I**, as $\mathbf{S} = \mathbf{P} \cup \mathbf{Q} \cup \mathbf{I}$ [30, 52]. In light of this, 'a **S** b' implies that alternative a is at least as good as alternative b.

4.2.3 MCDM problem formulation

Guitouni and Martel [52], Mota *et al.* [91], Ölçer and Odabaşi [92], and Zanakis *et al.* [139] employ the following MCDM problem formulation. Let $\mathbf{A} = \{a_1, \ldots, a_i, \ldots, a_n\}$ denote a set of alternatives when analysing a decision space. Let $\mathbf{C} = \{c_1, \ldots, c_j, \ldots, c_u\}$ denote the set of criteria or attributes pertinent to the decision at hand. A rating e_{ij} is given to the i^{th} alternative on the j^{th} criterion. The MCDM problem is therefore depicted by the following $n \times u$ *performance table* or *decision matrix*

$$\mathbf{E} = \begin{array}{c} \\ a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & \ldots & c_j & \ldots & c_u \end{array} \\ \begin{pmatrix} e_{11} & \ldots & e_{1j} & \ldots & e_{1u} \\ \vdots & & \vdots & & \vdots \\ e_{i1} & \ldots & e_{ij} & \ldots & e_{iu} \\ \vdots & & \vdots & & \vdots \\ e_{n1} & \ldots & e_{nj} & \ldots & e_{nu} \end{pmatrix} \end{array}.$$

The rating e_{ij} represents the score given when evaluating alternative a_i against criterion c_j. In addition, a weighting, or relative importance value, denoted by w_j, may be assigned to each criterion j, with the sum of the weightings summing to one.

4.2.4 Available multi-criterion aggregation procedures

A number of multi-criterion aggregation procedures exist and form part of the various MCDM methods. Guitouni and Martel [52] presented the main multi-criterion aggregation procedures in the following four MCDM categories: *elementary methods*, *single synthesizing criterion*, *outranking methods* and *mixed methods*.

Elementary methods are intended to simplify complex problems in order to select a preferred alternative. Due to the simplicity of the approach and analysis, these methods are suited for problems where a single decision maker is taken into account [78]. Weightings among criteria are not required, although multiple criteria may be present. The elementary MCDM methods include the *weighted sum*, *weighted product*, *lexicography*, *conjunctive*, *disjunctive* and *maximin* method. These methods may be of a descriptive decision-making nature. These methods are discussed in detail by Guitouni and Martel [52], Linkov *et al.* [78], MacCrimmon [81], Massam [85] and Steynberg [120].

Single synthesis criterion methods are considered as one of the most traditional approaches in MCDM. According to Akbulut [3], these methods aim to combine several considered criteria into a single comprehensive index. These methods also tolerate fuzzy theory, imperfect knowledge and consistency to a certain extent. A few prominent single synthesis criterion methods, as discussed by Akbulut [3], Guitouni and Martel [52], Steynberg [120], Velasquez and Hester [135], and Zanaksis *et al.* [139], include the *technique for order by similarity to ideal solution* (TOPSIS), *multi-attributive value theory* (MAVT), *utility theory additive* (UTA), MAUT, *simple multi-attribute rating technique* (SMART), *analytic hierarchy process* (AHP), *evaluation of mixed data* (EVAMIX), *fuzzy weighted sum* and *fuzzy maximin.*

Outranking methods follow a successive pairwise comparison of the alternatives under all criteria. They differ from single synthesising criterion methods in that they create a synthesising preference relation system rather than only considering each alternative in isolation [3]. The *elimination and choice expressing reality* (ELECTRE) method was the first to take an outranking synthesising approach and variations on this and the *preference ranking organization method for enrichment evaluation* (PROMETHEE) followed [52]. Further outranking methods include MELCHOR, ORESTE, REGIME and the *novel approach to imprecise assessment and decision environments* (NAIADE). Akbulut [3], Colson and De Bruyn [30], Guitouni and Martel [52], Velasquez and Hester [135], and Zanakis *et al.* [139] discuss these methods in detail.

The mixed methods, QUALIFLEX [3, 52], *Fuzzy conjunctive and disjunctive method* [78, 52] and the *Martel and Zaras method* [52] cannot be precisely categorised according to the three categories previously mentioned, however, these aggregation procedures are also often used in MCDM methods.

Traditional MCDM methods consider crisp sets of data (*i.e.* ordinal information) as opposed to uncertain or vague data, however, ambiguity and vagueness are frequently apparent in real-world decision-making problems. Fuzzy MCDM methods have therefore been developed to address the fuzzy data sets within these problems [52, 92].

Despite the great variety in MCDM methods, no one method can be considered appropriate to solve all decision-making situations [52]. It is important to apply a multi-criterion aggregation procedure which best suits the problem at hand. Guitouni and Martel [52], Mota et *al.* [91] and Velasquez and Hester [135] provide guidelines in selecting an appropriate MCDM method to suit a specific problem.

4.3 Modelling human decision-making

Human decision-making is defined as a cognitive and evolutionary process, based on an initial objective and a set of alternatives, from which the decision-maker choose the most appropriate action according to a set of criteria influenced by bias and individual preferences [25]. The best available alternative is not chosen if it is not the more appropriate alternative from the perspective of the decision-maker. Divekar *et al.* [35] found that the decision-maker's intuition with regards to risk routinely lead to decisions which deviate from rationality.

A central theme within the literature on human decision-making methods is the concept of *bounded rationality* (also known as limited rationality), meaning that any intended rational behaviour occurs within certain constraints [34]. Simon [111] first proposed this concept, stating that humans have physiological and psychological limitations which influence their rational choice and ability to arrive at an optimal solution. Linked to the concept of bounded rationality is the Satisficing model first proposed by Simon [111], where a decision-maker (acting as a satisficer) can only seek a satisfactory solution owing to cognitive limitations. The Satisficing model posits that the decision maker is satisfied by an alternative, even if it meets only some criteria, since decision makers are incapable of maximising in most situations [34, 113].

Conjunctive/Disjunctive models are decision-making methods which are also recognised as elementary MCDM methods. The Conjunctive model considers a group of potential solutions from a list of alternatives. The group of potential solutions includes all alternatives which exceed a certain threshold, while all alternatives which do not qualify are eliminated. Similar to the Satisficing model, the Conjunctive/Disjunctive models search for an adequate alternative, rather than an optimal solution. When minimising the evaluation function, a Conjunctive model will be used, whereas a Disjunctive model is used for a maximisation evaluation function [34].

Similar decision-making methods which include some form of rationality are the Additive and Additive Difference models regularly mentioned in the literature of descriptive decision-making methods [34]. These models provide good approximations of multi-attributive decision behaviour as they evaluate each multi-dimensional alternative independently. The Additive model considers and rates each alternative individually. These ratings are then compared holistically to determine the most appropriate solution. For the Additive Difference model, pair-wise comparisons are performed involving only two alternatives and one dimension at a time. Each of these comparisons are then multiplied by the weighting (or difference factor) to determine, for each attribute, the advantages and disadvantages per alternative. Finally, the performance values are summed for both alternatives in order to calculate the final subjective value for comparative purposes [34].

Cervantes *et al.* [25] followed a more multi-disciplinary approach in developing a comprehensive and coherent human decision-making model which encompasses an understanding in cognitive informatics, neuroscience and psychology. The primary underlying characteristics used to evaluate an alternative before making a decision are identified. These include the likelihood of success, reward, loss aversion, effort and delay. It is suggested that an individual will alternate between options in an attempt to maximise or minimise the value of these dimensions in order to identify the best alternative. The model also addresses issues which occur in modelling autonomous agents to exhibit human-like behaviour.

It is also suggested to use heuristics when considering human decision-making, instead of strict rigid rules of optimisation, owing to the complexity of a situation and an individual's rational inability [111]. Gigerenzer and Gaissmaier [46] define heuristics as a subset of strategies which choose to ignore information partially, with the aim of making quick, frugal and proper decisions. When using heuristics, there exists a trade-off of having some loss in accuracy for a faster and more economic solution. For the purpose of this study, a few one-reason decision-making heuristics are considered, namely *One-Clever-Cue*, *Take-the-Best*, and *Fast-and-Frugal* trees. One-reason decision-making is a class of heuristics which only consider one good reason when making judgements, therefore ignoring other cues [46].

Within nature, animal species tend to rely on a single 'clever' cue when searching for food, areas for nesting or mates. An example of such an One-Clever-Cue heuristic is the gaze heuristic. Take, for instance, an animal's pursuit of a prey or mate – the actual trajectory of the pursuing animal is not calculated in three-dimensional space, but rather a constant optical angle between themselves and the target is maintained. Another example where the One-Clever-Cue heuristic used is in geographical profiling [45, 46].

The Take-the-Best heuristic considers the manner in which one alternative is inferred to have a higher value on a criterion that another, based on binary cue values. Three rules form the basis of this heuristic: (1) The search rule, (2) the stopping rule and (3) the decision rule. The decision-making is simplified in that the search amongst cues is stopped as soon as the first cue which discriminates between alternatives is found. It is then inferred that the alternative with the positive cue value of one possesses the higher criterion value [46]. The Take-the-Best heuristic is similar to a Lexicographic model, as it evaluates alternatives considering only one attribute. Furthermore, it may be formally presented as a Fast-and-Frugal tree [65].

The final heuristic considered is Fast-and-Frugal trees. Typically depicting Bayes' rule as a tree, there would be a number of 2^m leaves, where m is the number of attributes or binary cues which may lead to a computationally unmanageable model. To put it into perspective, the Fast-and-Frugal tree will have only $m + 1$ leaves, which will allow the model to be more robust and tractable. The foundation rules of this heuristic are similar to that of the Take-the-Best heuristic, where cues are searched through in a predetermined order, stopping when such a cue leads to an exit. Examples where Fast-and-Frugal trees have been utilised include the modelling of physicians' thinking in the field of emergency medicine, as well as in the law when having to decide on bail [46].

4.4 Modelling decision-making of forcibly displaced people

The process of human decision-making is extremely complex as is influenced by a person's emotions, motivations and perception. For this reason, two people in similar circumstances may end up choosing completely different alternatives [25]. In this study of forcibly displaced people, the model aims to emulate the decision-making process of Syrians when confronted with conflict.

In an attempt to adequately model the decision-making of Syrians forcibly displaced, advice and insights were gathered from subject matter experts, such as Aksel [4] from Koç University, Frydenlund [43] from the Old Dominion University, Groen [50] from Brunell University, Lemos [75] from the University of Agder, Shomary [110] from Stockholm University, Smith [115] from the University of Sussex and Stewart [119] from the University of Cape Town. Their fields of expertise range from simulation and conflict modelling, to global studies, decision-making theories and social sciences.

A suggestion was made to simplify the modelling of the decision-making process of the people in order to grasp only the essential elements [75, 119]. There exists a strong appeal towards the simplification of decision analysis and support, especially considering the benefit of increased transparency and insight. The research performed by Katsikopoulos *et al.* [65] revealed that simple models tend to outperform complex models under the following two conditions:

> *Condition A*: When the data available are of low quality and not ample enough to estimate reliable model parameters, and

> *Condition B*: When there exists one alternative or attribute which dominates the others.

Condition A suggests that the decision at hand is difficult, which may be due to unpredictably regarding the problem. On the contrary, condition B implies that the decision is apparent and relatively straightforward. It is thus advised to use simple models rather than complex ones if either of these conditions can be considered to hold true [65]. In this study, which considers forced displacement, condition A holds true, as data are scarce and not typically of high quality (especially as most data are only estimates) [4, 43].

The decisions to be modelled are (1) choosing a movement type (*i.e.* whether to move as an IDP, refugee or undocumented migrant), and (2) choosing a proposed destination (dependent on the outcome of the first decision).

4.4.1 Choosing a movement type

Making a choice between the different movement types is mainly influenced by a person's characteristics and attributes [9]. Initially, a Fast-and-Frugal tree was considered for use in modelling this decision-making process, however, predetermining the order in which to present the alternatives based on the person's attributes is not feasible, since none of the attributes are predominant over the rest in determining the movement type of a person. As a result, the Additive model was considered and deemed fit for the purpose of determining a person's movement type, as it takes into account all attributes according to scores assigned by the modeller. This allows for a total score to be calculated for each alternative (movement type). This total score corresponds to the probability of that alternative being selected. Klabunde and Willekens [67] agree that the most apparent manner in which to evaluate such choices is to number the alternatives and select the option which achieved the highest valuation.

4.4.2 Choosing a proposed destination

The decision of where to move to depends on the movement type of a person. Factors such as safety, distance and population density are also taken into consideration. In the case of IDPs, an exact location within Syria needs to be identified as a new destination for the person, whereas the refugees and undocumented migrants must simply choose which neighbouring country to move to (not necessarily a specific location within that country).

Internally Displaced People

The decision-making method deemed fit in determining a new destination for an IDP is the Conjunctive model. A group of potential solutions is identified based on the criteria of safety and population density. The subject matter experts consulted, along with various sources from literature agree with this selected criteria [32, 43, 75, 90, 110, 115]. The weighted sum MCDM method is employed to identify the group of potential solutions by allocating weights to the two aforementioned criteria. Then, further utilising the Conjunctive model, the evaluation function aims to minimise the distance between the person and their new destination. The group of potential solutions is explored to find the alternative which minimises this distance. This allows the model to select an adequate, although not necessarily optimal, alternative. This feasibly replicates the decision-making of a person which, due to bounded rationality, also tends to not necessarily be optimal [34].

Refugees and undocumented migrants

Refugees and undocumented migrants must choose a destination country from the countries neighbouring Syria. To model this decision-making process, the weighted sum MCDM method is employed. Each alternative (neighbouring country) is scored, based on weighted factors such as the distance from the person to the border of that country, the ease with which a person will be able to enter that country (referred to as an openness score) and the population of Syrians presently within that country at a given time instance (referred to as popularity). The total score calculated per country will then correlate to a probability for a person to select that country as destination. Lemos [75] agreed with the method and also on the use of openness scores and further suggested the allocation of different openness scores per country towards refugees and undocumented migrants. In light of the fact that the openness scores toward these two groups of people would not typically differ significantly for the purpose of this model, it is excluded to avoid unnecessarily complicating the model.

4.5 Chapter summary

In this chapter, an introduction to the field of decision-making, as well as various modelling methods, was given. The different human decision-making situations captured within the simulation model were mentioned and the decision-making methods chosen to be employed within the simulation model were discussed.

The field of decision-making is discussed in §4.1. The difference between prescriptive and descriptive decision-making theories within this field was explained, with typical examples given in each case for illustrative purposes.

In §4.2, the study of MCDM methods was discussed. Some classification of these methods as employed in literature were discussed in §4.2.1, before considering the concept of preference modelling pertaining to MCDM and the problem formulation of MCDM methods in §4.2.2 and §4.2.3, respectively. In §4.2.4, various available multi-criteria aggregation procedures used in literature were mentioned.

The concept of human decision-making and the modelling thereof was discussed in §4.3. Various methods, such as the Satisficing model, Conjunctive/Disjunctive models, Additive and Additive Difference models, One-Clever-Cue heuristic, Take-the-Best heuristic and Fast-and-Frugal trees were mentioned and briefly discussed.

Finally, the chapter concluded in §4.4 where a discussion on the decisions captured within the simulation model developed in this thesis, as well as the decision-making methods employed to model these decisions, was given. This included insight and knowledge from various subject matter experts.

CHAPTER 5

An agent-based model depicting the movement of forcibly displaced people

Contents

An agent-based simulation model was designed and developed in order to simulate the movement of forcibly displaced people. This chapter discusses the development of the model within the AnyLogic software environment and the various limitations and assumptions considered in the modelling process are mentioned. Within this development process, three areas of modelling are identified — the modelling of conflict, the modelling of people and the modelling of a person's decision-making — and a graphical user interface is created to control these elements. Each of these components are considered in detail within this chapter.

5.1 Background to the model

The application of an agent-based simulation model in the social, political and economic sciences is highly complex and it is therefore important that such a model has a definite purpose and is developed with an appropriate level of detail [18].

The simulation model developed within this study is of a dynamic nature in that it simulates the conflict and associated population movement changes over a given period of time. This change over time is characteristic of a continuous simulation model, as changes occur continually rather than at specific time instances. The model is also stochastic in its behaviour since not all input variables are definitively known and, therefore, probabilistic estimates are widely utilised. With respect to the modelled level of detail, a medium level of abstraction is chosen, characteristic of a model which is more tactical than operational or strategic in nature. Agents, for example, do not represent an individual, but rather an aggregate of individuals with similar attributes. These agents are then capable of making decisions, in particular with regards to when and where to move in the presence of conflict. Furthermore, they can communicate with other agents and move within the modelled environment based on decision made in light of their profile, as well as their surroundings.

A map of Syria and its surrounding countries and ocean is employed within the simulation model, based on a scale of $1km : 1pixel$, established over the $1\,000 \times 600$ *pixel* model space. The comprehensiveness of the model scale, in that the entire country is modelled, allows for conflict initiation and spread to replicate reality while considering the whole of Syria. When confronted with conflict, agents decide whether to move or not, as well as where to move. The manner in which this decision-making process is captured is based on decision-making methods explained in §4.4 so as to incorporate the qualitative data which forms part of the decision-making process.

5.2 The AnyLogic Simulation Software Suite

The design and development of the agent-based model was conducted in the ANYLOGIC 7 Personal Learning Edition 7.3.6 [122] software suite. This multi-method simulation modelling tool enables a modeller to gain deeper insights into complex systems and processes across various industries [122]. ANYLOGIC includes a sophisticated suite of model development tools which, along with the JAVA modelling language, allows modellers to create complex graphical simulation models [134].

ANYLOGIC was found to be a suitable choice in facilitating the model development of the research at hand. Its agent-based approach and flexibility in terms of the level of abstraction employed supports the development of complex real-world systems, such as the movement modelling of people. The software is used to model not only the movement of people, but also the initiation, spread and depletion of conflict, along with the consequent decision-making processes of the people. ANYLOGIC supports the modelling of various aspects and also the interaction between them. Furthermore, the GIS animation and capabilities supported by ANYLOGIC will be useful for future work proposed to stem from this study.

Various ANYLOGIC components were employed in order to construct a model framework and code sequences which appropriately depict the conflict, people and their decision-making. These components are explained in Table 5.1.

5.3 Assumptions and limitations

There exist many challenges and important considerations when developing a model of social behaviour within conflict situations. These include the gathering of qualitative data, the quantification of that data and understanding the motives and constraints which affects the movement of forcibly displaced people. The primary assumptions made and limitations considered in the model are related to the geography, time, data and agent attributes within the simulation model.

5.3.1 Geographic

The developed model considers only neighbouring countries of Syria as plausible destinations for immigration by forcibly displaced people as, according to data from the UNHCR [129], more than 85% of Syrians who fled their country moved to a neighbouring country. These countries include Turkey, Iraq, Jordan, Lebanon and Greece, as seen in Figure 5.1 (Cyprus is representative of Greece). Other countries are not considered owing to modelling complexities.

TABLE 5.1: *The different components employed in the simulation environment [19, 134].*

Component	**Application in simulation**	ANYLOGIC **icon**
Object class	Represents a separate agent which contains its own internal structure governing its behaviour and decisions	
Parameter	Describes agent characteristics of an object class which only change as the behaviour of the agent changes	
Variable	Stores results of model characteristics during a simulation run which change over time	
Function	Executes a portion of code which may return the value of an argument and\or perform an action when called during a simulation run	
Event	Activates a scheduled action during the simulation triggered by a condition, specified rate or timeout	
EXCEL file object	Creates a link to a specified EXCEL file which allows for the reading of data from and writing of data to the spreadsheet	

FIGURE 5.1: *The map of Syria and its neighbouring countries utilised in the simulation model.*

Israel, although a neighbouring country of Syria, is not considered as a destination country for Syrian refugees owing to the ongoing war between Israel and Syria, deeming the border between these two countries as sealed [40].

The simulation model depicts the movement of agents within Syria, as well as the movement of agents from Syria to neighbouring countries. Once an agent has moved to a neighbouring country, the model no longer considers that agent's consequent decision-making processes and movement, since consideration is limited to the first country of entrance with respect to cross-border movement. Movement of displaced Syrians from one neighbouring country to another is therefore not considered in this model.

The geographic area of neighbouring countries utilised in the model, as illustrated in Figure 5.1, is not considered in detail. That is, the model only considers the area as a representation of the entire country. If an agent decides to move to a neighbouring country, the model tracks to which country it decides to move (in order to keep track of the associated fluctuating populations), but not the exact location within that country where the agent might choose to settle.

5.3.2 Time

The 'Arab Springs'[1] was initiated in Syria at the end of January 2011 [129, 87, 130]. The simulation therefore commences at the start of 2011, just before the occurrence of 'Arab Springs' and runs until the end of 2016. This aims to minimise the effect of *selection bias* which occurs when more than five years are considered. Selection bias can easily undermine the validity of research when restricting or narrowing the range of variation of data [28]. This is mitigated by initiating the simulation before the occurrence of the Syrian war and continuing the simulated period over five years.

Other time-based factors taken into account are the speed at which an agent moves and the waiting period which exists before an agent moves. To replicate the movement of forcibly displaced Syrians, who usually flee by foot, a fixed average movement speed of $4km/h$ is employed [105]. Agents do not always leave immediately and the waiting period correlates to the maturity of the conflict [87]. Aksel [4] affirmed that the duration of the war does have an effect on the likelihood of a person to move. Initially people might wait longer before deciding to move, whereas if the war has been ongoing for a few years, people would more easily move when they feel threatened as they have an indication of the effect of conflict based on what others experienced. During the beginning of 'Arab Springs', people were still reluctant to relocate, but as the conflict began to extent over years, increasing in its maturity, people, when confronted with conflict at a more mature state, are more inclined to relocate.

5.3.3 Data inputs and model definitions

The quantitative data employed by the model is available to the public and was gathered from various sources, such as the *Global Database of Events, Language, and Tone* (GDELT) [125], the *United Nations Department of Economics and Social Affairs Population Division* [126], *The World Bank* [11] and the *Central Intelligence Agency* [23]. Exact data do not always exist or are not always available — in such cases, sensible estimates, based on various sources and given arguments are made.

The developed model categorises agents fleeing conflict according to three types: (1) *Internally Displaced Persons* (IDPs), (2) *Asylum-seekers/refugees* (which will further be referred to as refugees) and (3) *Undocumented migrants.* IDPs are Syrians who were forcibly displaced, but choose to stay within the borders of Syria, whilst refugees are those who choose to cross the Syrian border legally, by either applying for asylum in the destination country, or being registered as refugees living in refugee camps. The category of undocumented migrants are those

[1] A series of anti-government protests, which sparked the initial uprising and armed rebellions and spread across the Middle East.

people who choose to cross the Syrian border without going through the legal and documented procedures.

Sarzin [106] proposes that analysts should integrate data sources in order to account for the different increasing and decreasing factors (as shown in Table 5.2) when considering the fluctuation of populations. The number of IDPs increases when people decide to flee their place of residence and relocate elsewhere within Syria. Similarly, if children are born while the mother is classified as an IDP, or if administrative errors are corrected, so the IDP data change. Administrative corrections are applicable to all of the categorised types of forced migrants in terms of the fluctuation in numbers.

Births are also an increasing factor, in the same way that deaths are a decreasing factor, to all the categorised types listed in Table 5.2. Decreasing factors require consideration in the modelling context. For example, when an IDP decides to cross an international border in order to seek asylum, the individual should not be counted as both an IDP and an asylum-seeker. Furthermore, when an IDP decides to move elsewhere within Syria, the individual should not be recorded for a second time as an IDP. In the case where an IDP decides to return to their original place of residence, the individual should then no longer be considered an IDP.

The number of refugees increases as individuals stochastically arrive or apply for asylum in a country. Similarly, the number of undocumented migrants will also increase owing to individuals impetuously arriving in a country, either temporarily, or with the aim of resettling in that country. Decreasing factors of the number of refugees and undocumented migrants include, amongst others, repatriation (individuals who return to their country of origin) and resettlement (individuals who decide to relocate to another country).

TABLE 5.2: *The increasing and decreasing factors of IDPs, refugees, as well as undocumented migrants, adapted from Sarzin [106].*

Type	Increasing factors	Decreasing factors
IDPs	New internal displacement Births Administrative corrections	Cross-border flight Return Settlement elsewhere in the country Deaths Administrative corrections
Refugees	New asylum applications or impetuous arrivals (separately identifying individuals who were previously IDPs) Births Administrative corrections	Positive political decisions (convention status, complementary protection status) Rejected asylum applications Otherwise closed asylum applications Repatriation Resettlement Deaths Administrative corrections
Undocumented migrants	Impetuous arrivals (separately identifying individuals who were previously IDPs) Resettlement arrivals Births Administrative corrections	Repatriation Resettlement Deaths Administrative corrections

This model considers internal displacement, new asylum applications and impetuous arrivals of migrants across the border. The aim is to eliminate the reliance on administrative data, thereby effectively removing 'administrative corrections' as an increasing or decreasing factor in population size. Furthermore, the model accounts for births and deaths, as well as IDPs fleeing across the border, but it does not consider the other decreasing factors of refugees and undocumented migrants, owing to the associated modelling complexities.

One noteworthy factor which is excluded from the model is the voluntary and impetuous repatriation of Syrians. As noted by Koepke [69] in a study on Afghans in the Islamic Republic of Iran nine years after the conflict in Afghanistan, the decision Afghans have to make in terms of whether or not to return to their country of origin is somewhat frightening, as most areas in Afghanistan are rural and offer very basic infrastructures, social services and employment opportunities after the conflict. Younger people also prefer to remain within Iran due to economic and educational opportunities. In the same manner, it would be daunting for Syrians to decide whether or not to return to Syria, especially while the conflict is still on-going. Younger Syrians might also prefer to remain in neighbouring countries, especially if economic and educational opportunities exist. Koepke also states reasons which limit repatriation, in particular with regards to Afghans in Iran, although they also apply to Syrians in neighbouring countries. These include the fact that many refugees may have no property to return to, along with few employment opportunities. Furthermore, there exists limited access to basic health care, education and other humanitarian support upon return.

5.3.4 Conflict

The model considers material conflict (*i.e.* physical conflict as opposed to verbal conflict, according to the GDELT [125] classification of conflict events) which occurred in Syria during the aforementioned timeline, as well as the intensity of these conflict occurrences. It does not, however, consider the other attributes of these conflict events, such as the type of actors involved (*e.g.* police forces, government or rebels) and the reason for occurrence, as the modelled outcome is not effected by these factors. Only conflict events initiated within Syria are captured by the model.

5.3.5 Agent attributes

Various attributes of an agent may influence when, why and to where an agent will move when confronted with a conflict-affected environment. Edwards [37] suggests considering attributes such as age, gender, wealth, ability to travel, ethnic identity and the like. This model considers, amongst other characteristics, an agent's gender, age, education and economic status. As mentioned, agents in this model constitute an aggregate of 10 000 individuals with similar characteristics. In light of this, aggregation attributes such as health status and social networks are not possible to model with a useful degree of accuracy. Furthermore, the health of a person should be considered in conjunction with other attributes (such as economic status and age). For example, an older individual with poor health and of low economic status might not be able to move at all, or only move within Syria; whereas a young working individual with a chronic health illness might choose to seek asylum in another country where proper medical facilities are available. Owing to the complexity of this attribute and the complexities surrounding the inclusion of social networks, these aspects are regarded as falling outside of the model scope. Lemos [75] affirms this decision, advocating that agent attributes be restricted to only essential inclusions.

5.4 Modelling of conflict

Modelling of conflict encompasses the initial occurrence of a conflict-related incident, the spread of this conflict, as well as its eventual depletion. In order to imitate a real-world situation in the simulation, historic data pertaining to conflict incidents which occurred in Syria during the specified timeline were acquired using the event data analysis service of the GDELT Project [125].

Arrays including `GroundState[][]`, `GroundStateInit[][]` and `GroundState_Time[][]`, are computationally overlayed across the physical model space, thereby mapping out 200×120 two-dimensional space to facilitate the modelling of conflict. Each of the cells within this space, as stored by the aforementioned arrays, represents a physical area of $5{\times}5\,km^2$ in the simulation model.

5.4.1 The initiation of conflict

The data gathered on the conflict events which occurred within Syria between 2011–2016 are used as input to model the conflict. For each occurrence, the processed datafile, named Conflict-Data, includes the date, *Global Positioning System* (GPS) coordinates and the total number of information sources citing this event. The total number of information sources citing an event is normalised so as to determine significance and intensity rating for each conflict event.

The array, `GroundState[][]`, represents the state of conflict within each cell as a quantitative value out of 100, where 0 refers to no conflict and 100 means fully-fledged war. The function, `EstablishGroundState`, allocates an initial value of zero to every cell within the `GroundState[][]` array, as it is assumed that no conflict is present at the outset of the simulation.

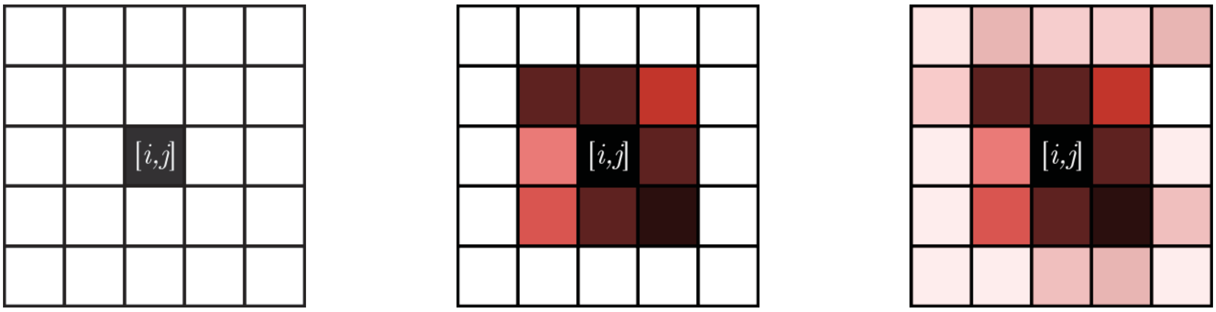
The event, `ConflictInitiate`, is set to be triggered once per simulation day (*i.e.* a day within the simulation runtime). With every trigger of the `ConflictInitiate` event, the current simulation date is compared to the list of recorded dates in the datafile. When the simulation date matches a date present in the datafile, the GPS coordinates and conflict intensity values are extracted from the datafile. The GPS coordinates are then converted to fit the positioning and scale of the model and the conflict is initiated at the specific cell located within the `GroundState[][]` array, with a value set to the intensity given in the datafile.

Before concluding the event, the function, `UpdateGroundState`, is called. Within this function, a search is conducted within the `GroundState[][]` array to identify cells in the array which contain a value greater than zero (*i.e.* the cells in which conflict exist). These cells are then coloured a scaled shade of red, according to the intensity of the conflict relative to the maximum conflict rating present within the simulation space, for observational purposes.

5.4.2 The spread of conflict

The spread of conflict is based on a combination of the concepts of *reaction-diffusion* (*i.e.* the uniform distribution of constituents within a space and the inverse counteract to elementary reactive steps) and *cellular automata* (as discussed in §3.2.2). If conflict occurs within an area, it is reasonable to assume that it will spread outwards and affect surrounding areas according to the reaction-diffusion principle. In quantifying conflict, its spread is modelled based on the principle of cellular automata, where the state of the surrounding cells affect the state of the current cell.

The spread of conflict, as modelled within the simulation, is illustrated in Figure 5.2. The intensity of conflict is shown by the darkness (or intensity) of colour, meaning a more intense conflict is indicated by a darker colour of the cell. Figure 5.2(a) indicates a cell where conflict originates with a $[x, y]$-position (where x represents the row and y the column) of $[i, j]$. The spread of conflict from the cell where conflict initially occurred to its immediate neighbours is depicted in Figure 5.2(b). In Figure 5.2(c), the spread of conflict to other cells is illustrated, as well as the depletion in intensity as the conflict spreads from the place of origin.



(a) *Conflict initiated*　(b) *Immediate neighbouring cells affected*　(c) *Further neighbouring cells affected*

FIGURE 5.2: *A graphical illustration of the spread of conflict.*

An event, `ConflictSpreading`, is triggered daily within the simulation runtime and incorporates the spread of conflict to neighbouring cells within the `GroundState[][]` array, according to a certain probability, `ProbabilityofInfection`.

The cells immediately surrounding the cell in which conflict is initiated (depicted in Figure 5.2(b) as $[i, j]$), will, with a probability of `ProbabilityofInfection`, be affected. The conflict intensity realised in each of these cells is calculated as a fraction of the conflict intensity which exists at the point of initiation. The `GroundState[][]` value of the immediate neighbouring cells will therefore be less than the `GroundState[i][j]` value (as indicated in Figure 5.2(b)) where conflict originated.

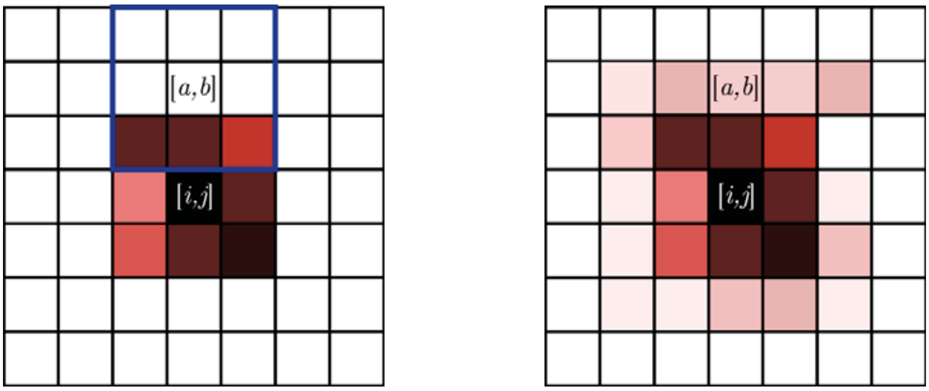
Furthermore, conflict continues to spread to cells beyond the immediate neighbouring cells of the origin point, as indicated in Figures 5.2(c) and 5.3. The `ConflictSpreading` event begins by searching through cells within the `GroundState[][]` array which fall within Syria, identifying cells with a zero conflict value. Assume, for example, that the identified cell has an $[x, y]$-position of $[a, b]$ (as shown in Figure 5.3(a)). The conflict `GroundState[][]` value of the immediate neighbouring cells of $[a, b]$ are considered in determining the average and maximum values (cell $[a, b]$ is excluded from these calculations). The average of these cells is calculated as

$$\left(\sum_{m=a-1}^{a+1} \sum_{\substack{n=b-1 \\ 1\{m=a \bigwedge n=b\} \neq 1}}^{b+1} GroundState[m, n] \right) \Big/ 8.$$

The notation sums the conflict intensity value of the cells surrounding cell $[a, b]$ and divides the value by the number of immediate neighbours. If this average conflict value exceeds the variable, specified as `SetNeighbourAverage`, such that

$$\left(\sum_{m=a-1}^{a+1} \sum_{\substack{n=b-1 \\ 1\{m=a \bigwedge n=b\} \neq 1}}^{b+1} GroundState[m, n] \right) \Big/ 8 > SetNeighbourAverage,$$

the GroundState[a][b] (*i.e.* the value of conflict at cell $[a, b]$) will, with a probability of ProbabilityofInfection, gain a conflict intensity value. That is, the indicator function $1\{m = a \bigwedge n = j\}$ takes on a value of 1 if the argument is true and 0 otherwise. The value allocated to GroundState[a][b] will then be a fraction of the maximum conflict value identified amongst the immediate neighbouring values, as seen in Figure 5.3(b). This ensures that, if a value is allocated to a cell where previously no conflict existed, the conflict intensity will be less than the maximum of its immediate neighbours.



(a) *The neighbourhood of the selected cell [a,b]*

(b) *Cell [a,b] gained a conflict intensity value*

FIGURE 5.3: *An illustration of the spread of conflict to other neighbouring cells.*

Finally, the ConflictSpreading event activates the UpdateGroundState function and cells with a GroundState[][] value greater than zero are identified and tinted red according to the intensity of the conflict relative to other existing conflict.

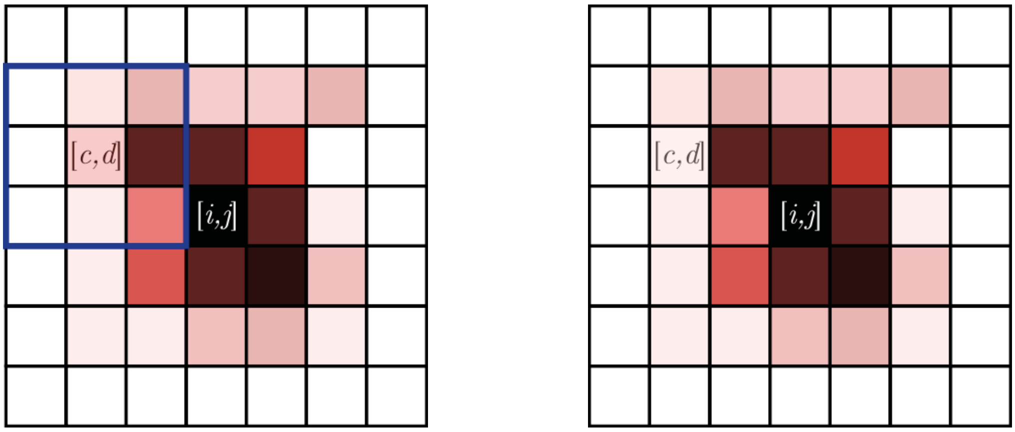
The method employed to model the spread of conflict allows for the spread to occur organically, in a reaction-diffusion-like manner, while utilising the rules of a typical cellular automata, as can be seen in Figures 5.2 and 5.3. The process detailed within this section occurs on a daily basis within the simulation runtime.

5.4.3 The depletion of conflict

The depletion of conflict is modelled in a similar fashion to its spread, as explained in §5.4.1, but effectively decreases overall conflict instead of advancing it. Figure 5.4 is used to illustrate conflict depletion. As explained in §5.4.2, the darker the shade of the cell, the higher its conflict intensity value. In Figure 5.4(a), a cell with an existing conflict intensity value and a certain average neighbouring conflict intensity value is identified. Figure 5.4(b) indicates the depletion of conflict within this identified cell.

The ConflictDepleting event executes once per simulated day and allows conflict zones located at the outer edges of the conflict area to be identified and decreased with a certain probability.

This event begins by first searching through the GroundState[][] array to identify cells within Syria which have a GroundState[][] value greater than zero (*i.e.* conflict exists within these cells). As an example, assume that a cell with a $[x, y]$-position of $[c, d]$, as seen in Figure 5.4(a), is identified as having a GroundState[][] value greater than zero.



(a) *The neighbourhood of the selected cell [c,d]*

(b) *The conflict intensity value of cell [c,d] decreased*

FIGURE 5.4: *An illustration of the conflict depletion.*

In a manner similar to the way in which conflict spreads, as explained in § 5.4.1, the `GroundState[][]` value of the immediate neighbouring cells are, once again, considered (see Figure 5.4(a)). This includes the average of neighbouring cells as well as the maximum `GroundState[][]` value amongst the neighbours. The average of these cells is determined by

$$\left(\sum_{r=c-1}^{c+1} \sum_{\substack{s=d-1 \\ 1\{r=c \bigwedge s=d\} \neq 1}}^{d+1} GroundState[r,s] \right) \Big/ 8.$$

If this average conflict value is less than the variable specified as `SetNeighbourAverage` such that

$$\left(\sum_{r=c-1}^{c+1} \sum_{\substack{s=d-1 \\ 1\{r=c \bigwedge s=d\} \neq 1}}^{d+1} GroundState[r,s] \right) \Big/ 8 < SetNeighbourAverage,$$

and the conflict has been present within the $[c, d]$ cell for more than a specified number of days, the `GroundState[c][d]` value will, with a probability of `ProbabilityofDepletion`, decrease. Once again, the indicator function $1\{r = c \bigwedge s = d\}$ takes on a value of 1 if the argument is true and 0 otherwise. The `GroundState[c][d]` value will decrease by a fraction of the maximum conflict value identified amongst the neighbouring cells (see Figure 5.4(b)), thereby ensuring that the conflict within that cell dissipates.

The `UpdateGroundState` function is then activated to appropriately tint the cells with a `GroundState[][]` value greater than zero with a shade of red related to the intensity of conflict.

5.5 Modelling of people

As mentioned previously in §5.3.5, each simulated agent represents 10 000 people who share similar demographics such as gender, age, economic status and level of education. An initial consideration was to allow each agent to represent a family, however, Aksel [4] suggested rather grouping individuals of similar ages and genders as, for example, men have different migrating patterns to women, with different age groups also exhibiting different behaviours.

The spread of agents over the simulated area of Syria mimics the actual population distribution as documented in the 'Syrian population by sex and governorate according to civil affairs records' of 2010 [133] and the natural fluctuation in population owing to births and deaths is also taken into account, as described in the following section.

The different stages and consequent decisions an agent may encounter in the model are illustrated in Figure 5.5 by means of a flow chart.

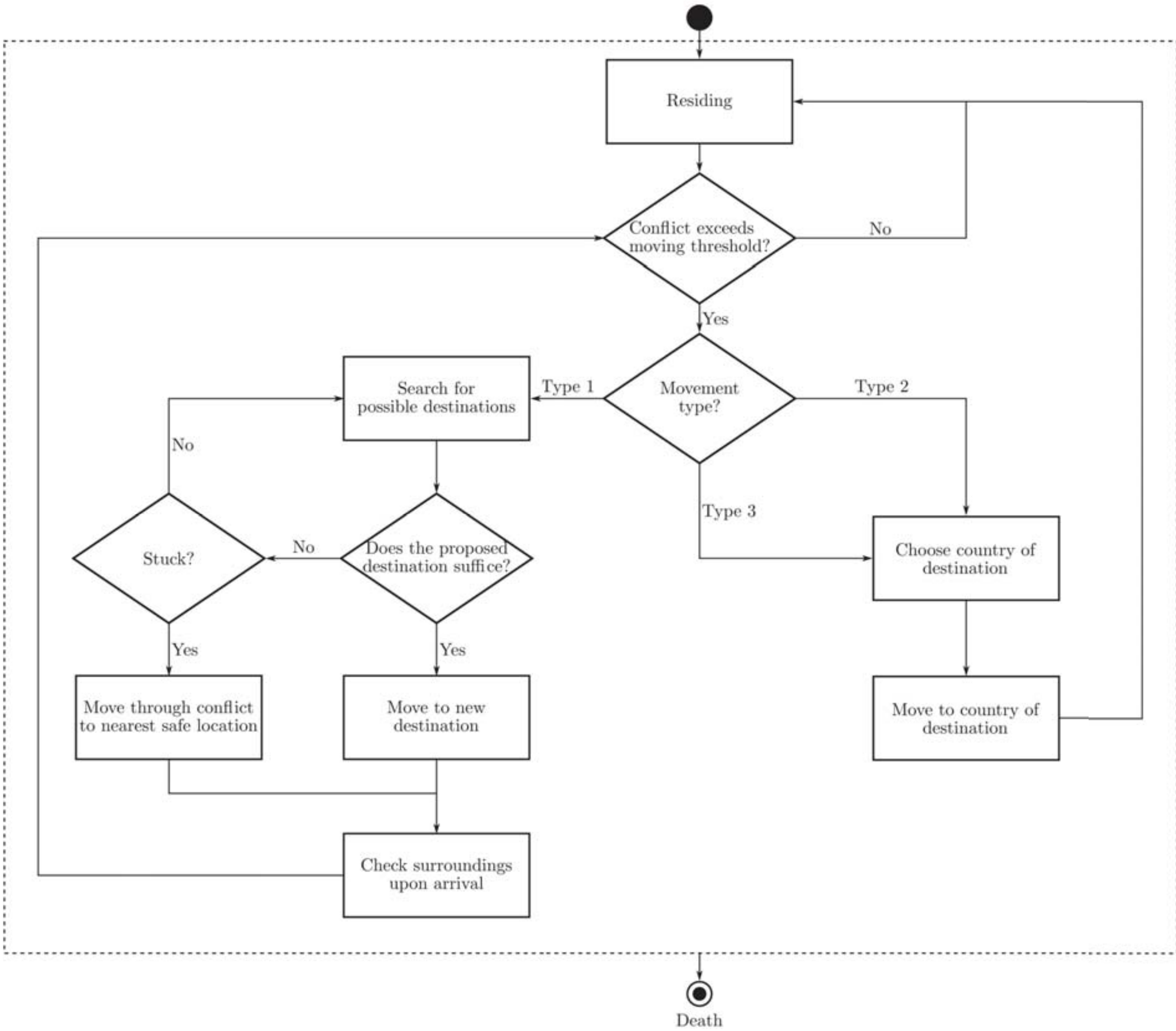


FIGURE 5.5: *A simplified flow chart depicting the stages an agent goes through.*

Initially, an agent exists in a state of residing, which is only disrupted when the conflict within that agent's immediate vicinity exceeds the agent's threshold to withstand violence. If this occurs, the agent decides whether to move as an IDP, a refugee or an undocumented migrant.

If, for example, the agent decides to stay within Syria as an IDP (Type 1), a new possible destination within Syria will be sought. If the newly discovered destination is sufficiently free of conflict, the agent will move there. If not, consider whether they are stuck (*i.e.* unable to find a sufficient destination while having searched for more than a certain time period). In this case, the agent will choose to move through conflict to the nearest location which will ensure safety, although this may lead to death. If the agent is not stuck, but failed to find a suitable destination, another search iteration will continue in an attempt to find a safe route to a conflict-free zone. Eventually, when an agent arrives at the identified destination, it will again check its immediate surroundings for the presence of conflict in case the conflict has progressed during the journey.

If, instead, the agent decides to move to a neighbouring country as either a refugee (Type 2) or an undocumented migrant (Type 3), the agent selects a neighbouring country as a suitable destination country and moves there. Upon arrival at the destination country, the agent will remain in a state of residing, as the model does not consider conflict within countries other than Syria itself and, in light of this, the agent will not be confronted with further conflict situations.

Within this model, the primary decisions to be made by each fleeing agent are (i) when to move and (ii) where to move to. These decision-making processes and the modelling thereof are thoroughly explained in §5.6. Before discussing the decision-making of an agent, however, it is necessary to contextualise agents within the model space, having certain combinations of attributes which attempt to replicate reality. This population of agents is considered within the simulated model space.

5.5.1 Agent attributes

Each agent in the simulation model is characterised by six parameters — gender, age, whether or not they received tertiary education, economic status, whether or not the agent has family living outside of Syria and anticipated age at death. These attributes and their associated values are contained within the `people` agent environment and are set according to the probabilities distribution indicated in Table 5.3.

TABLE 5.3: *Agent attributes*

Parameter	Determine value
Gender (Male)	$\text{P}(true) = 0.5$
Age	$\text{P}(0 < age < 15) = 0.364$ $\text{P}(15 \leq age < 65) = 0.602$ $\text{P}(age \geq 65) = 0.034$
Tertiary Education	$\text{If}(age > 18) \rightarrow \text{P}(true) = 0.33$
Economic Status	$\text{P}(LowEconomicStatus) = 0.119$ $\text{P}(MediumEconomicStatus) = 0.6$ $\text{P}(HighEconomicStatus) = 0.281$
International Family	$\text{P}(true) = \text{uniform}(0.05, 0.2)$
Anticipated Age at Death	$\text{If}(male) \rightarrow AgeAtDeath = \text{normal}(12, 64.7)$
	$\text{If}(female) \rightarrow AgeAtDeath = \text{normal}(12, 76.6)$

The "Syrian population by sex and governorate according to civil affairs records on 1/1/2010" [133] states that there are, in total, 101 males for every 100 females in Syria. It is therefore valid to assume that a person would be either male or female with an equal probability of 50%. The model incorporates a boolean parameter called `Male` to indicate an agent's gender. The parameter will, according to the chosen probability, be set to true.

The next agent attribute, stored as the `Age` parameter, specifies the age of the agent. In 2010, the percentage of Syrian Arab Republic's total population by broad age groups were as follows: 36.4% were 14 years of age or younger, 60.2% were between the ages of 15 and 64 and 3.5%

were older than 65 years of age [126]. Agents added to the population by means of a function, AgeDistribution, ensure a similar distribution of age as suggested by the data. The ageing of an agent is also taken into account in the model and is discussed in the following section.

Another important attribute taken into account is the economic status of an agent. It is estimated that, in 2016, 11.9% of the Syrian population fell below the poverty line. This measure is based on surveys of subgroups, with the results weighted by the number of people within each subgroup. In addition, the unemployment rate was 8.3% in Syria during 2010 [23]. Further information suggests that an estimated 28.1% of the Syrian population were using the internet in 2014 [126] and, for modelling purposes, these individuals are considered as those of high economic status. For the case of simplicity, the parameter EconomicStatus is divided into three categories which include low, medium and high economic status. The probability of a person being of low, medium or high economic status is implemented in the model based on the aforementioned data as 11.9%, 60% and 28.1%, respectively.

The level of education of agents in the model is necessary for consideration and, for modelling purposes, a boolean parameter, called TertiaryEducation, is employed to indicate whether or not an agent has received tertiary education. Although 86.4% of Syrians over the age of 15 are literate, it is estimated that only 33 people per 100 in the population will receive tertiary education [23]. It is therefore modelled that, if a person is older than 18 years of age, there exists a 33% probability of them having been educated at tertiary level.

Another agent attribute considered within this model is the whether or not an agent has family living within across the Syrian border. According to Fargues [39], an estimated twenty million individuals from Arab countries were not living in their country of origin by 2011. Syria, along with eight other countries, were identified as 'major senders' and it is further estimated that between 5-20% of nationals from these countries were living abroad. To simplify the inclusion of this aspect in the model, the boolean agent parameter InternationalFamily will have a uniformly distributed probability between 0.05 and 0.2 of being true.

The final attribute employed to characterise an agent is its approximated age at death. The World Bank [11] provides the life expectancy (in years) of Syrians at birth between 2010 and 2015. For the purpose of this study, the data from 2011 to 2015 are considered, while the data for 2016 are generated as a duplicate of 2015. These values are shown in Table 5.4, with the average being utilised in the model.

TABLE 5.4: *The life expectancy (in years) of Syrians at birth from 2011 to 2016.*

	2011	2012	2013	2014	2015	2016	Average
Male	66.6	65.3	64.4	64.0	63.9	63.9	64.7
Female	76.7	76.5	76.5	76.5	76.6	76.6	76.6
Both genders	71.5	70.8	70.3	70.1	70.1	70.1	70.5

Within the model, an agent's anticipated age at death, stored by parameter AgeAtDeath, is determined by generating a random number from a normal distribution with the mean as the average anticipated age at death over the six years investigated. The parameter value has a lower bound of 0 and an upper bound of 100, and the variance is determined at the discretion of the modeller.

5.5.2 Simulating agent population

In an attempt to accurately simulate the Syrian population, the agents are geographically distributed across Syria, matching the population data per administrative governorate. Furthermore, the population fluctuation (births and deaths), as well as the ageing of agents are taken into account.

Certain variables, such as `BirthRate` and `InitialPopulation`, are defined and utilised to control initialisation and growth of the population. According to the data from the *United Nations Office for the Coordination of Humanitarian Affairs* (UNOCHA) [133], the population size of Syria at the onset of 2011 was 23 720 000, while the average annual number of births between 2010 and 2015 was 24.4 per 1 000 persons in the population. The annual number of deaths during the same period was 5.4 per 1 000 persons [126].

To manage the geographic spread of agents, presentation elements within the AnyLogic environment are employed to underlay a background representing the shape of each governorate within Syria. These shape files, as can be seen in Figure 5.6, allows for interaction with the agents.

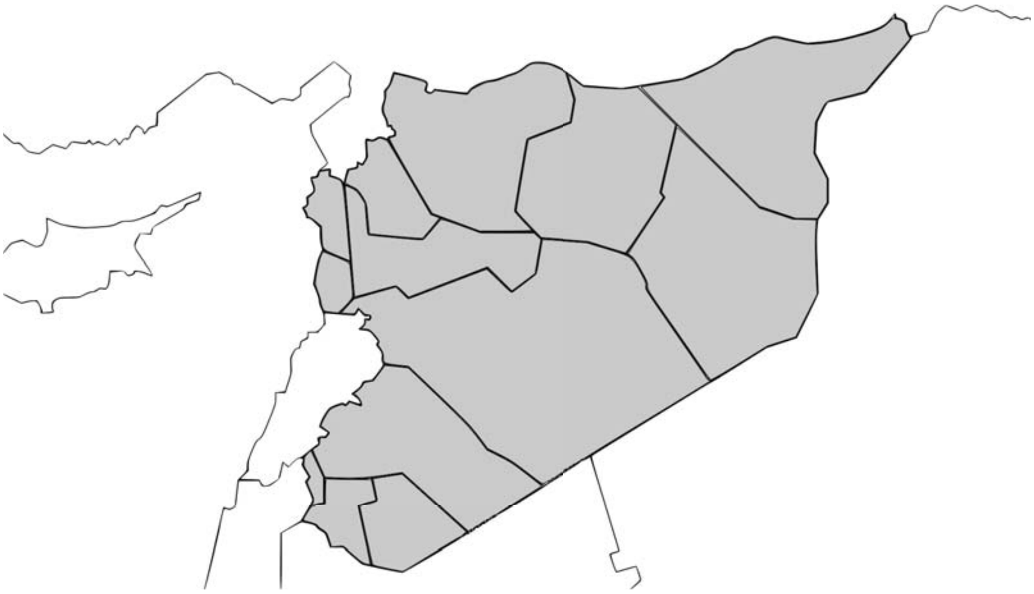


FIGURE 5.6: *The shape files representing the Syrian governorates.*

At the start of the simulation a function called `Populate` is called to populate the modelled area of Syria with agents. The percentage of agents populated per governorate correlates to the percentage of the total Syrian population per governorate as recorded in 2010 [133]. These percentages are stored as variables within the model as a set of variables shown in Figure 5.7. Within each governorate the agents are, however, placed randomly.

The `UpdatePopulation` event calls the `Births` function annually to account for the growth of the population. Within this function, the agent population is increased according to the specified `BirthRate`. The agents added to the population take on an age of 0, with the rest of the agent attributes allocated according to Table §5.3. Instead of implementing the death rate in this manner, the age at which each agent dies is determined as a local parameter, `AgeAtDeath`, of each agent. A local function called `Death` is called annually within each `people` agent's environment. This function assesses whether or not the agent has reached its anticipated death age and, if this is the case, the agent is removed from the simulation.

FIGURE 5.7: *The variables containing the percentages of the population per governorate.*

Data analyses conducted by Alhanaee and Csala [5] revealed that the number of deaths of children between the ages of 0 and 4 years of age increased during the winter months of December and January. A further finding was a declining adult male population between the ages of 18 and 59, owing to the number of males who die fighting in the war. These aspects are not investigated further or explicitly added to the model, as they are encompassed in the data pertaining to the life expectancy previously mentioned.

Another annual event which is included in the modelling of the agent population is the ageing of agents. The `Ageing` event simply increments the age of every agent annually. This is explicitly incorporated in the model since age is one of the parameters that influence a person's movement choices and, as this simulation takes place over a number of years, ageing is a factor which is necessary to account for.

5.5.3 The agent statechart

The behaviour exhibited by an agent in an agent-based model is specified within the statechart of that agent [20]. This statechart portrays the actions taken by the agent and shows the different states in which an agent can reside. Figure 5.8 shows the statechart constructed within the `people` agent class environment of the simulation model. The changing of states is either condition-triggered (?), arrival-triggered ( ) or timeout-triggered ( ).

This statechart correlates to the flow chart depicting the agent stages in Figure 5.5. At the point where an agent enters the system, attributes are assigned to the agent according to the description in §5.5.1 and the agent is placed within Syria according to the geographic distribution of people in Syria described in §5.5.2. The agent immediately enters the `Residing` state and remains there unless the conflict within the agent's immediate vicinity exceeds the agent's moving threshold. In this case, the agent's movement type will be determined as the agent enters `WaitingPeriod` state and after a certain number of days the agent will move to the `SearchMove` state. Within this state, the function associated with the movement type to which the agent belongs to will be activated, prompting the agent to search for a safe location to move to. As soon as the agent arrives at the identified destination, it will, if it is an IDP, enter the `CheckSurroundings` state where it will scout its immediate environment for any further conflict and if no threats exist, the agent will return to the `Residing` state. If there is indeed a threat, the agent will return to the `SearchMove` state and continue the search for an alternative safe destination. If the agent is either a refugee or an undocumented migrant it will bypass the `CheckSurroundings` state when arriving at its destination and directly revert to the `CheckSurroundings` state.

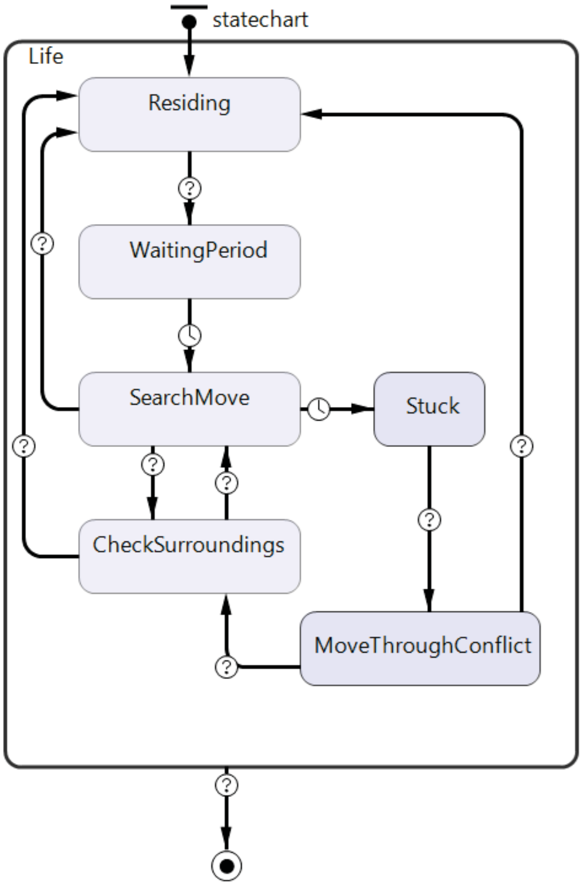


FIGURE 5.8: *The statechart of a person as modelled within the simulation.*

If an agent spends more than a specified time within the `SearchMove` state without successfully finding a safe destination, it is assumed that the agent is stuck within the conflict. The agent is then directed to the `Stuck` state where there is a 50% probability that it will die, or decide to move through the conflict towards a safer location. In the latter case, the agent will move to the `MoveThroughConflict` state where, based on the agent's movement type, a destination will be chosen. Once the agent finds and moves to the new destination, it moves to the `CheckSurroundings` state and, if it is indeed a safe zone, progress further to the `Residing` state. At any moment during the simulation an agent may die, either because it reached it's anticipated age at death, or due to being stuck amid conflict. If this occurs, the agent simply exits the system.

5.5.4 The moving threshold of an agent

Within the model, agents may choose to leave their current place of residence and migrate to another location. It is assumed that agents value their physical security and will relocate when they feel it is threatened. External push factors and other agent attributes, or a combination thereof, determine an agent's tolerance for conflict within its immediate environment, as discussed in §3.3.3. Within this model, the conflict is the external push factor, with an agent's attributes determining its ability to endure conflict before choosing to flee. Owing to this influence, two agents being exposed to similar conflict intensity may react differently. The influencing attributes accounted for in this model are gender, age, whether or not the agent has tertiary education, economic status and whether or not the agent has family living outside of Syria.

An agent's tolerance towards violence is defined by a variable, `MovingThreshold`, within the `people` agent class. Each agent's `MovingThreshold`, which is defined as a value between 0 and 1, is empirically calculated by a function called `DetMovingThreshold`. This function is recalled annually by the event `CheckMovingThreshold`. The annual update of the `MovingThreshold` variable considers the ageing of agents and consequences this may have on the value of the variable.

According to Aksel [4], men (especially those between the ages of 15 and 64 years old) typically choose to leave home when fleeing violence ahead of the women and children in order to seek a safe place for their family to relocate to. In light of this, male agents experience a slight decrease in the `MovingThreshold` value, whilst the value for women is increased. With respect to different age groups, data provided by the UNOCHA [132] state that children, adults and the elderly account for approximately 40%, 53.5% and 6.5%, of those in need of humanitarian assistance, respectively. People over 65 years of age are thus the least inclined to flee their place of residence, while children and adults relocate more easily. The process of relocating, especially emigration, is costly and it follows that agents with a low economic status are not expected to relocate as easily as those with a higher economic status [32]. People who have tertiary education are also more inclined to relocate in order to find new jobs in a better economy [110]. It is also assumed that, once members within an agent's social network relocate, that agent itself is then more likely to do similar, as the cost and risks involved in moving is decreased when more information pertaining to the experience or opportunity is available directly from someone within a person's social network [32]. Although the simulation model does not directly consider the social networks of people, it does take into account the family that an individual might have living across the Syrian border. If a person has international family, they are likely to be less reluctant towards moving.

By the end of 2015, the UNHCR estimated that more than half of the Syrian population have been forcibly displaced [129]. In light of this, it is assumed that the normalised value of the average conflict intensity experienced over the simulated time period should correspond with the moving threshold of more than half of the agent population. An initial `MovingThreshold` value is therefore assigned to each agent from of a triangular distribution with zero as the minimum, 1 as the maximum and the normalised value of the average conflict intensity as the mode. Within the `DetMovingThreshold` function, this initial value is then increased or decreased, according to the attributes of the agent, in order to determine the final `MovingThreshold` value of each agent.

The simulation daily initiates an event called `CheckForConflict`, which in turn activates the function called `CheckGroundState`. This function calculate the average `GroundState` value of the cell the agent is in, and the immediate neighbouring cells. If this average value exceeds the agent's `MovingThreshold` value, the agent switches from a residing state to a state of searching for a new proposed destination. This process depends on the `MovementType` of the agent, which will be explained in the next section.

5.6 Modelling the decision-making of an agent

In reality, the decision-making process with regards to migration is complex and, in order to keep the model tractable, it is necessary to conceptualise a simplified representation of this process. Three types of movement are considered in the model, as explained in §5.3. Movement type 1 pertains to those individuals who decide to leave their place of residence, but remain within Syria (referred to as IDPs). Movement type 2 refers to people who cross the Syrian border in order to seek refuge in refugee camps and/or apply for asylum in the country of destination (referred to as refugees). Movement type 3 considers those individuals who cross the Syrian border and relocate to another country without following documented procedures (referred to as undocumented migrants).

Depending on the movement type of an agent, it will select a proposed destination. Agents possess segmented knowledge about the current state of the modelling environment (conflict and population spread) depending on their location and may use this information to predict what their future might entail. This assists agents in accounting for their future well-being when making decisions [67].

5.6.1 Choosing a movement type

Alhanaee and Csala [5] analysed the motives of forcibly displaced people and concluded that safety and the proposition of a better life appear to be their strongest motivators. A person's attributes in totality, however, affect their decision when it comes to deciding whether to leave their country or relocate within its borders.

Shomary [110] explained that the people with little wealth, for example, will either stay within Syria or move to the refugee camps in neighbouring countries, as they cannot afford to pay smugglers to travel as undocumented migrants. Aksel [4] mentions, however, that those who choose to seek asylum are not necessarily of high economic status. Furthermore, individuals who choose to settle in other countries without applying for asylum are assumed to not have financial need [40, 137]. Shomary [110] concluded that the people who leave Syria as undocumented migrants are typically middle class, or of higher economic status, as there are fairly high costs involved. People who apply for asylum typically fall in a medium economic class and posses tertiary education, knowing that they will be able to look after themselves once granted asylum [4, 10, 98].

While some younger Syrians want to relocate and seek for job opportunities, having the necessary funds available to do so, others are reluctant to move. Shomary [110] spoke of her family who, like many others, chose to remain within the country, while only two of the younger members (both in their twenties) decided to move. The rest of the family, especially those over 40 years of age, are without education, have little work aspirations and therefore chose to stay within Syria. There are also numerous stories told within communities of how difficult it is for Syrian refugees to adapt in other countries (especially within Europe), which increases people's reluctance to leave Syria. Many Syrians who originally moved to other countries as refugees or undocumented migrants return to Syria owing to difficulty adapting. The return of refugees or undocumented migrants are, however, not taken into account within the simulation model.

Fargues and Fandrich [40] stated that people with relatives or friends across the border may use those connections to relocate outside of Syria. Those with no pre-existing networks would rather move to the nearest neighbouring country from which they will apply for asylum or stay in the refugee camps. The *Refugee Studies Centre* [100] estimated that more than 60% of people crossing the Syrian border settle with families or connections, choosing self-settlement to encampment, as discussed in §2.5. In general, Syrians who move across the Syrian border prefer self-settlement (being an undocumented migrant) to encampment (living in refugee camps).

Within the simulation model, an agent's V `MovementType` is determined as it enters the `Waiting-Period` state. As discussed in §4.4, an additive model is employed to determine the movement type of each agent. Various agent attributes such as age, economic status, whether or not they have tertiary education and whether or not they have family living outside of the Syrian border, influence this decision.

A probability matrix is constructed per attribute to indicate, for each class of that attribute, the probability that an agent will be of movement type 1, 2 or 3. Let $\mathbf{A} = \{a_1, a_2, a_3\}$ denote the set of alternatives where a_i represent movement type i. Let $\mathbf{C} = \{c_1, \ldots, c_j, \ldots, c_u\}$ denote the set of criteria constituting to each attribute. A probability p_{ij} is given to the i^{th} movement type at the j^{th} criterion. Let $\mathbf{E}^{(T)} = [p_{ij}]$, where $T \in \{1, \ldots, n\}$ represents the probability matrix for attribute T as

$$\mathbf{E}^{(T)} = \begin{matrix} \\ a_1 \\ a_2 \\ a_3 \end{matrix} \begin{matrix} \begin{matrix} c_1 & \cdots & c_j & \cdots & c_u \end{matrix} \\ \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1u} \\ p_{21} & \cdots & p_{2j} & \cdots & p_{2u} \\ p_{31} & \cdots & p_{3j} & \cdots & p_{3u} \end{pmatrix} \end{matrix}.$$

The values used to populate these probability tables represent quantitative values derived and inferred from qualitative data pertaining to the attributes of forced migrants within each movement type. For each movement type or alternative, a_i, the probability that an agent will move according to that movement type is given by

$$p(a_i(j)) = \frac{\sum_{T=1}^{n} p_{ij}(T)}{n}.$$

For each attribute probability matrix, $\mathbf{E}^{(T)}$, the criterion j corresponds to the specific criterion of the attribute which is associated with the specific agent. According to the probability associated with each movement type $p(a_i(j))$, the agent will be allocated a movement type. Within the `SearchMove` state, the agent will, depending on the movement type it has been allocated, call the function Ⓕ `MovementType1` or Ⓕ `MovementType23`.

5.6.2 Choosing a proposed destinations

The actions of an agent correlate directly to the movement type attributed to it. Distance and population density (or popularity) are important criteria, as individuals tend to move to the nearest location when fleeing conflict, typically choosing a destination where others have gone before [32, 90]. Shomary [110] agreed that people fleeing will take the population density of possible destinations into account when deciding where to move. When an agent is of movement type 1, the main considerations will be the safety of the proposed destination, the distance between the current location and the proposed destination, and the population density of the proposed destination. For movement types 2 and 3, the agent will consider the openness score of each neighbouring country, the distance from their current location to the neighbouring borders, and the number of Syrians who already relocated to the countries under consideration.

Modelling movement type 1

Harrison [56] mentions that IDPs generally move within the same area in which the conflict or disaster occurred (trying to minimise distance), although Alhanaee and Csala [5] state that these individuals prioritise safety. Both authors agree that the destinations IDPs move to highly correlate with areas of high population density [5, 56]. This corresponds with the study of Zipf [142] which states that the inter-community movement of people between two communities, P_1 and P_2, which are separated by a distance D will be directly proportional to the product of the population sizes of P_1 and P_2, and inversely proportional to the distance, D.

Within the `Main` workspace of the simulation, arrays are superimposed over the physical modelling space, each of granularity 100×60, in order to facilitate the modelling of the conflict, the population density, and the attractiveness of destinations within Syria, based on these two weighted criteria. Each of these cells represent a $100km^2$ area.

The array, `GS`, translates the array `GroundState` (which indicates the conflict intensity within Syria over a 200×120 space, as explained in 5.4) into a 100×60 array. Within the simulation, an event called `UpdateAttrZones` calls the `PopulateGS` function on a monthly cycle which is responsible for this translation. Another function, `UpdatePopDensity`, is simultaneously executed to count the number of agents within each of the simulated cells covering Syria and then populate the `PopDensity` array with these values. Furthermore, these population density values are normalised to a value between 0 and 100, and stored within another array called `PopDensityNorm`, for further calculation purposes. For visualisation purposes, the cells are tinted blue according to the population density of that cell. The greater the population density, the darker the hue of that cell.

The `UpdateAttrZones` event also calls a function called `UpdateCombinePopGS` which populates the `CombinePopGS` array so as to consider all attractive zones within Syria. Each cell, $[i, j]$ (the i^{th} column and j^{th} row), is allocated a value which considers the sum of the weighted criteria, conflict and population density. The greater the conflict intensity, the less attractive the cell, whereas greater population density increases attractiveness. People are drawn to places with a higher population density and therefore the inverse of the conflict intensity value is used when calculating the attractiveness value of a cell. The attractiveness value per cell is therefore calculated as

$$w_1(100 - GS[i][j]) + w_2(PopDensityNorm[i][j]),$$

where w_1 and w_2 represent the weight allocated to the conflict and population density, respectively. Both of these weights are user inputs. Furthermore, these attractiveness values are normalised and stored within the `CombinePopGSNorm` array as a value between 0 and 100. These normalised values are used to illustrate the attractiveness of each cell by applying a purple hue to the cell which corresponds to its level of attractiveness. The darker the hue of a cell, the more attractive it is as proposed destination for IDPs.

The `MovementType1` function within the `people` agent environment utilises the `CombinePopGSNorm` array to find good alternative destinations from which to propose a destination. The agent then searches through the list of alternatives, trying to minimise the distance between their current location and the destination before finally selecting a location as its destination.

Modelling movement types 2 and 3

When considering cross-border movement, this model only takes into account Syrians moving to neighbouring countries either as refugees or undocumented migrants. Section 5.6.1 described the modelled decision-making process which indicates how refugees (movement type 2) and undocumented migrants (movement type 3) differ according to their agent attributes. When it comes to modelling the decision-making process of these agents in choosing a specific country as new destination, however, the modelling thereof is similar for both movement types.

Similar to the decision-making process of IDPs, distance and population is taken into account when choosing a destination. The refugees and undocumented migrants consider the distance between their current location and the borders of the various neighbouring countries, as well as the popularity of each of these countries. Popularity refers to the number of Syrians who have already decided to move to that country since the start of the conflict. Another factor taken into account is the openness score of each neighbouring country (*i.e.* the ease with which a person would be able to enter the country). The objective is to maximise the openness score and popularity, whilst minimising the distance.

When an agent is classified as either of movement type 2 or 3, the F `MovementType23` function is utilised. User input variables, each representing the weighting factor of a certain criterion, are considered within this function. These variables are the weight factors which correspond to the distance V `MT23_factDist`, the openness score V `MT23_factOS`, and the popularity V `MT23_factPop` of the countries.

The F `MovementType23` function first considers the distance between the agent and each of the neighbouring countries' borders. Shape files are utilised to calculate this distance. Similar to the shape files that represent the different Syrian governorates as explained in §5.5.2, shape files representing the different neighbouring countries are also constructed, as illustrated in Figure 5.9. The inverse values of the various distances to each country are normalised over the maximum inverse distance, in order to equate all distances to a value between 0 and 100.

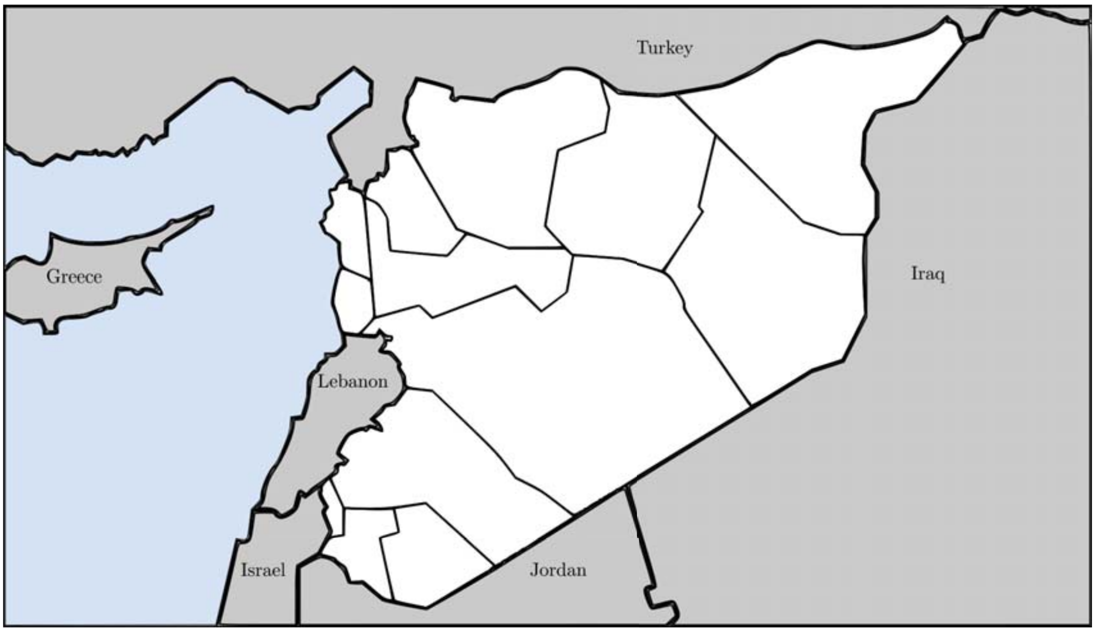


FIGURE 5.9: *The shape files representing the neighbouring countries to Syria.*

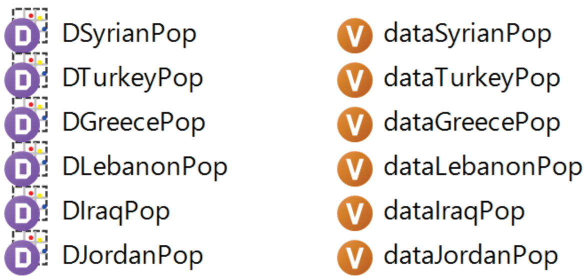


FIGURE 5.10: *The series datasets and corresponding variables which stores the population data per country.*

Next, the number of Syrians who have moved to the respective neighbouring countries is determined by means of a function `PopCounter` within the `people` agent environment which is called monthly during the simulation by an event `CheckPopCounter`. The function allows for the number of people located within each country to be counted, while also taking into consideration their `MovementType`.

These data are stored in the `Main` object class within a series of datasets which are accessed via a series of variables, depicted in Figure 5.10. The values representing the population of Syrians in the respective neighbouring countries are normalised over the maximum population in order to indicate the popularity of each of these countries as a value between 0 and 100.

Finally, the openness score of the neighbouring countries are considered. The openness scores of the neighbouring countries are stored within parameters, as shown in Figure 5.11. Each country has an initial openness score which can be specified by the user. Within the `Main` object class, an event called `AdjustOpenScores` calls a function called `OpennessScores` each month which allows for fluctuation of the openness scores.

FIGURE 5.11: *The openness scores stored as parameters within the* `Main` *object class.*

The Refugee Studies Centre [100] regards Turkey as being far more humane and pragmatic than the other countries neighbouring Syria, considering its approach towards the mass influx of Syrians. It is therefore assumed that Turkey initially possesses the highest openness score. When comparing the neighbouring countries of Syria, Turkey is the only signatory of the 1951 Refugee Convention, although other neighbours such as Iraq, Jordan and Lebanon regard Syrian refugees as temporary visitors and initially had an open-door policy [40]. Turkey and Jordan had set up large refugee camps for those most vulnerable and, although Lebanon refused to allow international humanitarian aid to set up such camps, Syrians were still granted access into Lebanon [100]. In light of this, Turkey and Jordan have initial openness scores which are greater than that of Lebanon.

Initially, neighbouring countries welcomed large numbers of Syrians, but, as the war intensified, these countries began restricting the influx of Syrians, with some borders closing altogether [24]. With this occurrence in 2013, Turkey became increasingly popular as it was still open to Syrians, as well as Iraq which experienced an influx owing to a number of camps that had been set up for Syrian refugees [87, 137]. Towards the end of 2014, Jordan's attitude towards the Syrians started changing, as the government realised their approach is unsustainable [1].

In January 2015, the Lebanese government introduced new regulations which required Syrians to apply for a visa before being granted entrance, leading to a decrease in Lebanon's openness. Furthermore, in March 2015, the Turkish government decided not to further accept any asylum applications, thereby leading to a decrease in its openness. Most official border crossings from Syria to Jordan and Turkey were strictly controlled. In January 2016, Turkey began requiring Syrians to obtain a visa before entering the country and Iraq started closing most of their borders which resulted in a decreased openness for both these countries [137]. Then, when Germany opened their borders to refugees in June 2015, it resulted in an increase of Syrians in Greece, as they travelled via this country towards Europe [10].

The `OpennessScores` function compares the time within the simulation to the timeline of openness adjustments as described above and when a correlation occurs, the openness score of the neighbouring country involved is adjusted in the model accordingly. The fluctuation is managed by employing a certain percentage increase or decrease, rather than simply adding or subtracting values. Figure 5.12 depicts a timeline showing when changes to neighbouring countries' openness scores occurred. The notation ↑ and ↓ indicate an increase and a decrease in openness score, respectively.

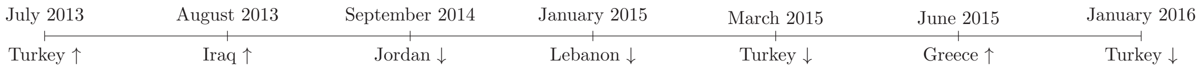


Figure 5.12: *A timeline illustrating the fluctuation of the openness scores of neighbouring countries.*

When an agent of movement types 2 or 3 select a country as proposed destination, the distance, popularity and openness score are taken into account within the `MovementType23` function. Let $\mathbf{N} = \{n_1, \ldots, n_i, \ldots, n_I\}$ denote the set of alternatives or neighbouring countries. Then, $Dist_i$ refers to the normalised inverse value of the distance between the agent and the border of country i, whilst $Popularity_i$ represents the popularity of country i and $OpennessScore_i$ refers to the openness score of country i. The overall score, S, of country i, can therefore be calculated as

$$S_{n_i} = w_1(Dist_i) + w_2(Popularity_i) + w_3(OpennessScore_i),$$

where w_1, w_2, and w_3 are the weight factors associated with each of the criteria. This overall score is stored as a variable within the `people` agent's modelling environment, shown in Figure 5.13. The probability of an agent selecting a specific country as proposed destination directly correlates to the overall score of that country. According to this probability, an agent then selects a country which results in the agent moving to a random point within that country.

Figure 5.13: *The overall scores calculated per neighbouring country.*

5.7 The graphical user interface (GUI)

A *graphical user interface* (GUI) acts as a platform for the simulation model user to engage with the functionality of the simulation system in an intuitive and informative manner. The GUI employed consists of a configuration screen and a primary screen which facilitate the workings of the user with regards to the simulation model. The following section describes the GUI of the simulation model as seen by its user, along with the underlying modelling structure which regulates and facilitates the various model components.

5.7.1 The configuration screen

The configuration screen, as shown in Figure 5.14, is constructed within the `Simulation:Main` tab of the model. Upon initiation of the simulation model, the configuration screen will appear. It is designed to prompt the user for certain user inputs required in order for the model to be executed. If the user chooses not to enter any inputs, the model will still execute by employing default values. A number of user inputs have been added to the configuration screen, as labelled in Figure 5.14, and are discussed below.

(A) **Data input**

The simulation model allows for both data and manual initiation with regards to the modelling of conflict. A user can choose whether to manually initiate conflict by means of clicking on a specific location during the simulation run, or to utilise existing data on the conflict as input. The data may include dates, GPS locations and conflict intensity values. Within the `Simulation:Main` tab, a variable called `Input` links to the `Main` object class parameter called `SetInput`. Upon simulation execution, this parameter determines whether to activate the manual conflict initiation environment or to initiate the `ConflictInitiate` event, which reads the data input from a EXCEL file linked to the EXCEL file object, `ConflictData`.

(B) **Criteria weights regarding refugees and undocumented migrants**

An agent, when classified as either a refugee or an undocumented migrant, is required to decide on a country as a proposed destination. This decision-making process, as explained in §5.6, considers three criteria (distance, openness score and popularity), which are weighted in order to model the decision. These variables are employed whenever an agent has to decide between neighbouring countries. The user's input values are linked to a set of variables within the `Simulation:Main` tab, which determines the initial values of a set of parameters within the `Main` object class environment. Each of these parameters represents the openness score of the associated neighbouring country.

(C) **Criteria weights regarding IDPs**

The "Attractive Zones" layer illustrates the attractiveness of areas within Syria when considering proposed destinations for IDPs. Attractiveness is calculated by taking the conflict and population density of the physical area into consideration. The user may input the weights associated with each of these criteria in the calculation. These input values link to the variables within the `Simulation:Main` tab which are utilised within the `Main` object class on a monthly basis within the simulation run in order to determine the attractive zones.

(D) **Initial openness scores of neighbouring countries**

When an agent classified as a refugee or undocumented migrant needs to select a neighbouring country as their new destination, one of the criteria taken into account is the openness score of the various countries. This is a score out of 100, where 0 reflects extreme reluctance of a country to accept migrants, while 100 indicates that the country is open and welcoming towards migrants. The openness scores fluctuate during the course of the simulation run as affecting incidents occur. The initial openness score of each neighbouring country at the start of the war may be controller as user input.

When the user finished the configuration, the model execution can be initiated by the user be means of a click on the "Run" button.

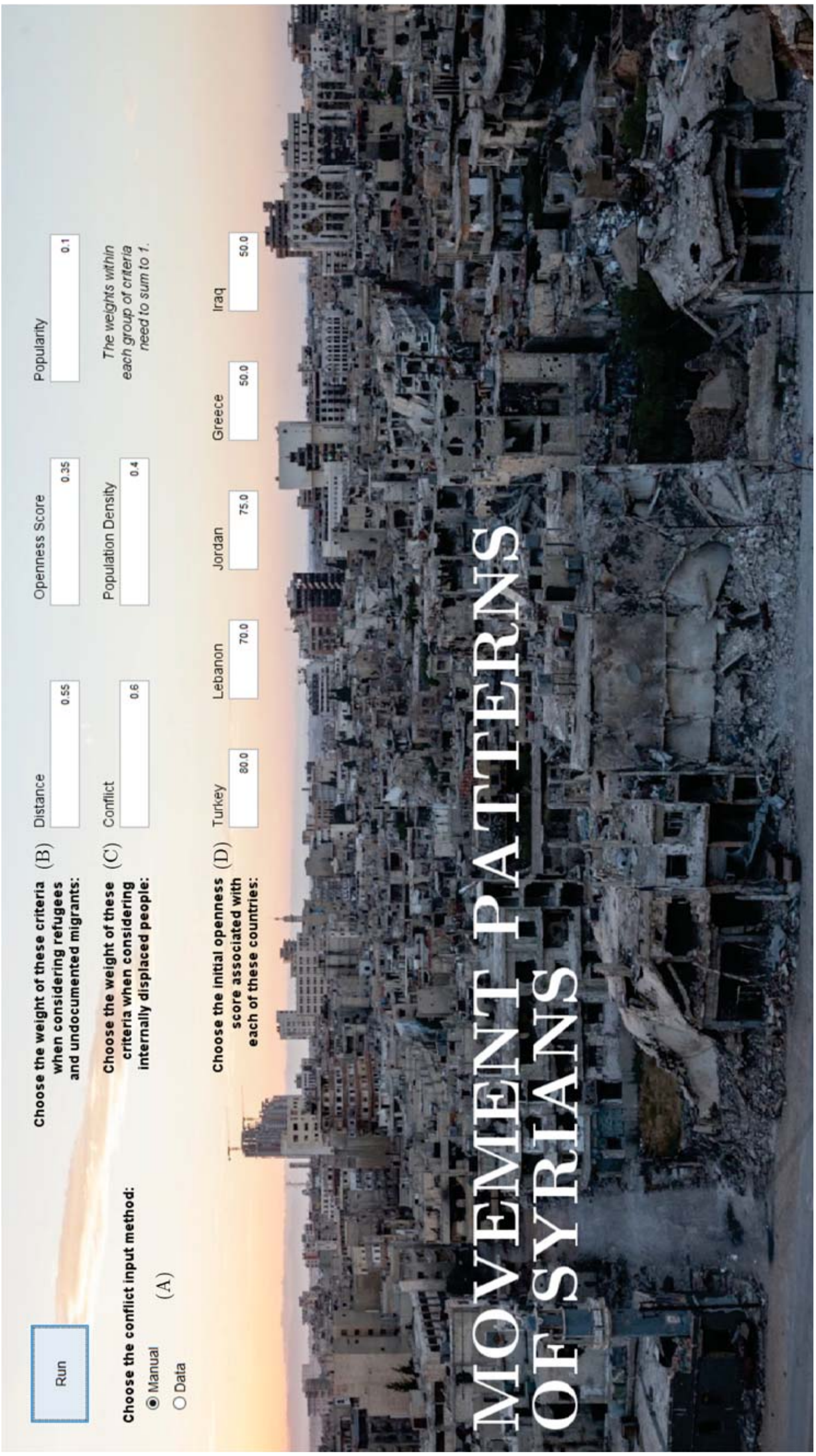


FIGURE 5.14: *A screenshot of the simulation model's configuration screen.*

5.7.2 The primary screen

The primary screen of the simulation model, as shown in Figure 5.15, has been developed in the `Main` object class and during the execution of the model, this screen will be visible to the user. The user-specified values on the primary screen may be altered during the simulation run, while the user-input settings on the configuration screen may only be set before the execution of the simulation run. The user inputs included in the primary screen, as labelled in Figure 5.15, are discussed below.

(E) **Visibility of layers**

The checkbox feature allows the user the functionality of choosing certain indicative animation layers to be shown or hidden. A conditional statement controls the dynamic visibility feature for each instance and, when the box is checked, the corresponding layer will be superimposed as a display over the primary modelling space. Three layers are constructed in the form of grids overlaying the simulation model space. The "Conflict Zones" layer illustrates conflict and its associated intensity within Syria at any given simulation time by means of a red hue. A greater conflict intensity within a cell will result in a dark hue. The "Population Density" layer shows the different population densities across Syria by means of a blue hue. Similarly, cells with a dark blue hue indicate high population densities within those cells. The "Attractive Zones" layer gives an indication of the attractiveness of an area with regards to locations proposed for IDPs and, as previously explained, the attractiveness is visually shown with a purple hue, with a darker hue pertaining to a more attractive area.

(F) **Graphs**

Another checkbox feature is employed to allow the user to view graphs depicting data from the simulation run, plotted against the simulation timeline. These graphs, shown in Figure, include the number of Syrians who moved to each of the neighbouring countries, the fluctuation of the Syrian population living in Syria, the total number of Syrian IDPs, the total number of Syrians who fled as refugees, the total number of Syrians who fled as undocumented migrants, the number of Syrian refugees per neighbouring country, as well as the number of Syrian undocumented migrants per neighbouring country. The data pertaining to these graphs are stored within data sets by means of an event, `UpdatePopulationData`, which retrieves the data per simulated year within the simulation run by means of the `PopData` function. Refer to Figure 5.16 for an example of a simulation run output illustrated on the graphs.

(G) **Conflict spread adjustment bar**

The probability of conflict spread is another input the user may alter before or during the simulation model execution. This value links to the `ProbabilityofInfection` parameter. The `ConflictSpreading` event, which occurs three times per simulated week, employs this variable in modelling the spread of conflict.

(H) **Conflict depletion adjustment bar**

Similar to the conflict spread adjustment bar, the conflict depletion adjustment bar is a user input which links to the `ProbabilityofDepletion` parameter. The `ConflictDepleting` event executes three times per simulated week, utilising this input variable.

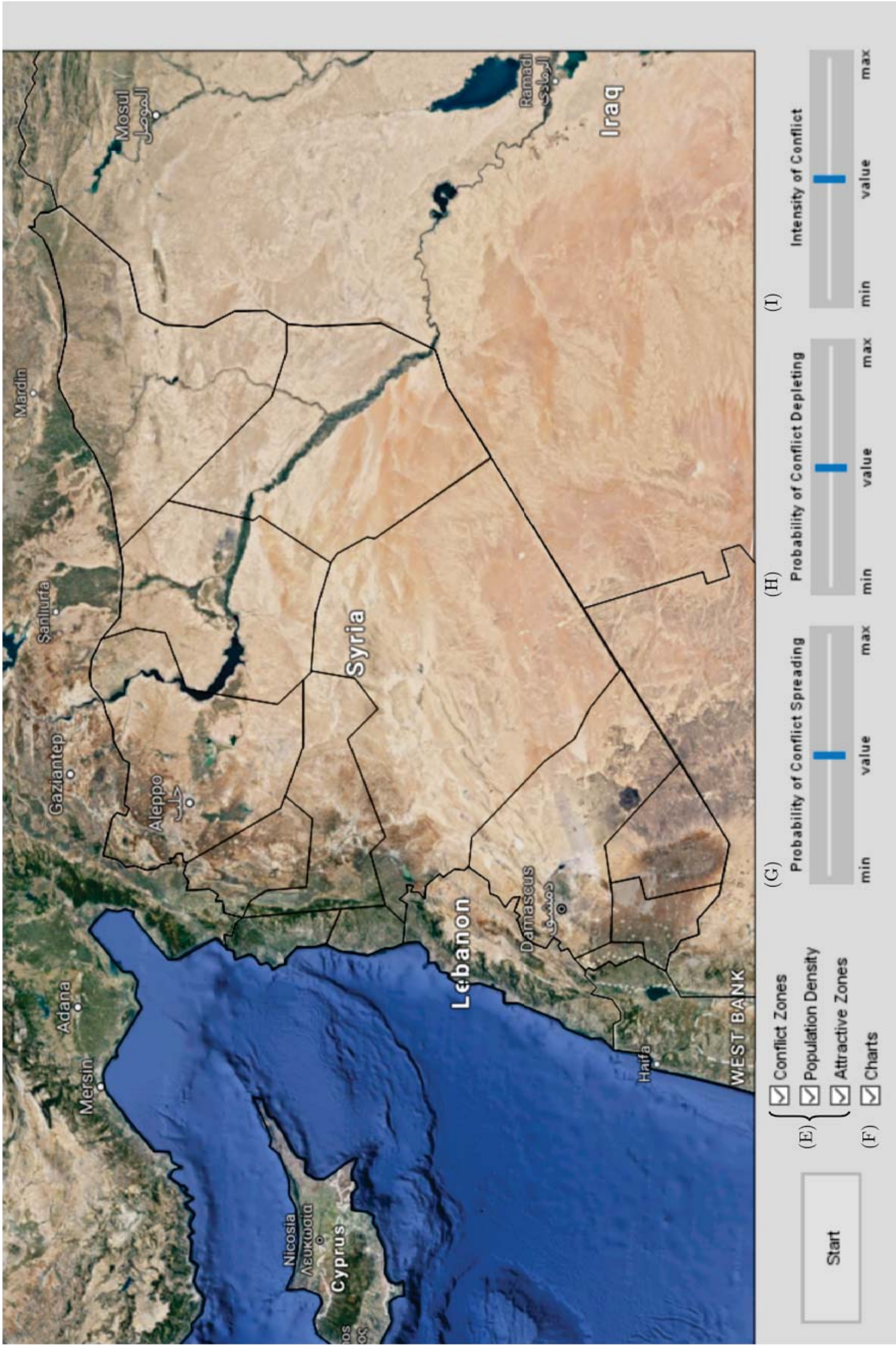
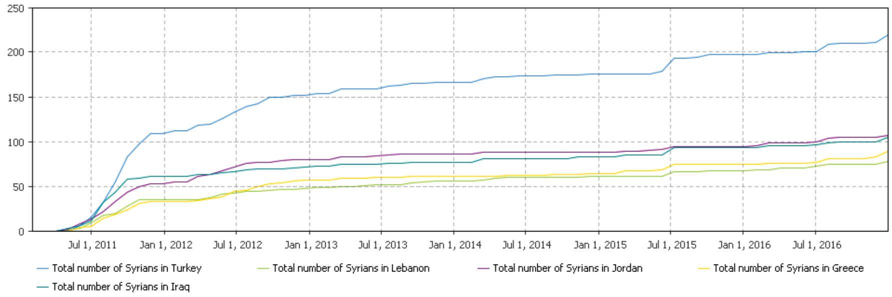
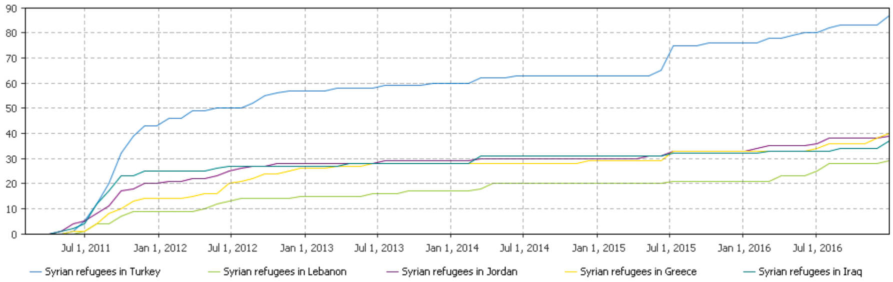


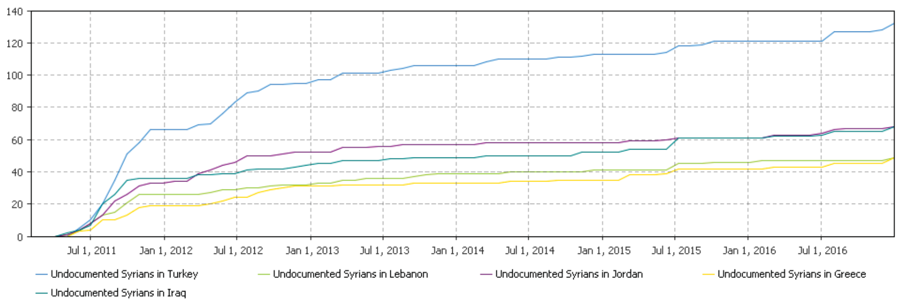
FIGURE 5.15: *A screenshot of the simulation model's primary screen.*



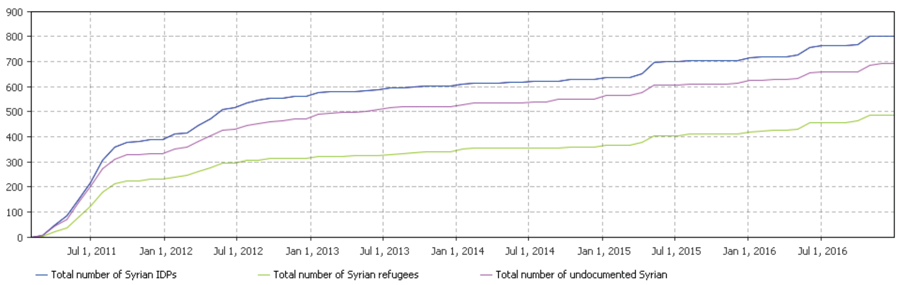
(a) *Total number of Syrians per neighbouring country*



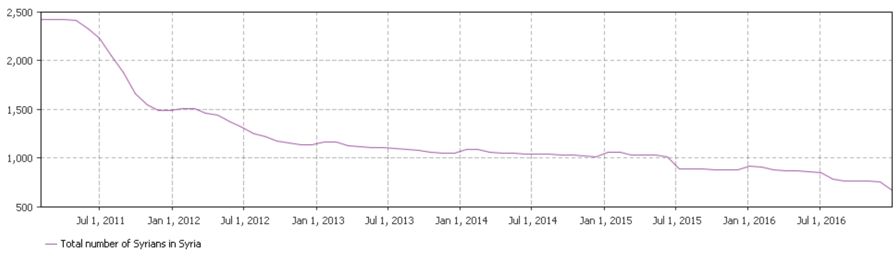
(b) *Number of Syrian refugees per neighbouring country*



(c) *Number of undocumented Syrians per neighbouring country*



(d) *The number of displaced Syrians per movement type*



(e) *Total number of Syrians living within Syria*

FIGURE 5.16: *A screenshot of the graphs displayed during the simulation run.*

(I) **Conflict intensity adjustment bar**

The conflict intensity adjustment bar is only available for user manipulation when conflict data is manually initiated and selected as such within the configuration screen. Conflict is manually initiated in that the user sets the intensity of the conflict by means of the adjustment bar and then clicks on a location within the simulation modelling area from where the conflict should initiate. The conflict intensity user input is stored within a variable called V `ConflictIntensity`.

The primary screen appears with all the checkboxes unchecked and default values (as determined through model testing) set to the adjustment bars. The user can therefore choose whether or not to have any further influence, before clicking on the "Start" button for the simulation time to start. The user will then be able to adjust the various input values further during the simulation run.

5.8 Chapter summary

In this chapter, a discussion of the modelling approaches taken and assumptions made in the development of an agent-based model depicting the movement of forcibly displaced people is provided, along with a background to the AnyLogic simulation environment. The model comprised four major components namely the modelling of conflict, the modelling of people, the modelling of an agent's decision-making process, and the GUI implementation. The modelling of each of these divisions discussed are based on literature, subject matter expert opinions and sensible assumptions.

The modelling of conflict encompassed the initiation of conflict based on existing data or manual initiation, and the spread and depletion of conflict based on a combination of the concepts of reaction-diffusion and cellular automata. The modelling of the people considered the attributes allocated to agents, the births and deaths which allowed for fluctuation in the population size, the different states an agent may reside in and the moving threshold of each agent. The modelled decision-making processes of the agents included the allocation of an agent's movement type along with a proposed destination. The GUI platform allowing user-specified input were discussed along with the various elements which form part thereof.

CHAPTER 6

Verification of the agent-based model

Contents

The verification of the agent-based model developed in Chapter 5 is detailed within this chapter. The chapter opens with a short introduction to the concept of verification in §6.1, where after the manner in which conflict modelling in the simulation model was verified is explained through various tests in §6.2. These tests assess, amongst others, the manual conflict initiation function, the use of input data to initiate conflict and the subsequent spread or depletion of conflict present in the model. Following this, the correct functioning of the modelled population is addressed in §6.3 through a series of experiments in order to verify module performs as intended. This is followed by further verification pertaining to the decision-making of the agents in the model, detailed in §6.4. The chapter then closes with a chapter summary in §6.5

6.1 Model verification

Verification, as defined by Stewart [118], is the process of ensuring that the model is accurately developed and that it performs as intended. Verification occurs concurrently with the development of the model in that, as each phase or part of the model is added, it is subsequently verified, making the process iterative in nature. While verifying the model, the modeller should ensure that the model is correctly implemented in the computer software and that the input parameters and logical structure of the model are sensible [12].

The four fundamental building blocks of the simulation model, as described in §5.4–5.7, are verified individually before considering the model as a whole, when all of these building blocks are implemented together. These elements include the modelling of conflict, the population, the decision-making process and the GUI. A number of cases are tested within each of the modelling divisions for verification purposes.

6.2 Verification of the modelled conflict

The process followed in modelling the initiation, spread and depletion of conflict is discussed in §5.4. In order to verify the conflict, the following cases were tested in simulation experiments:

Case (i): Manual conflict initiation

Case (ii): A small set of input data points to test the accuracy of the GPS positioning

Case (iii): The effect of a varying probability of conflict spread

Case (iv): The effect of a varying probability of conflict depletion

Case (v): The effect of varying conflict intensity

These cases were implemented in the model and subsequently reflected upon. Figure 6.1 shows the results achieved for the different cases implemented. For each of these cases, the *Conflict Zones* layer checkbox within the primary screen was selected to show the presence of conflict.

In Case (i), the *Manual* conflict input method was chosen from the configuration page. During the execution of a simulation run, the *Probability of Conflict Spreading* was set to 1 and the *Probability of Conflict Depleting* was set to 0. The *Intensity of Conflict* was set to 100 and conflict was initiated at the city of Aleppo by means of a click on the map. The outbreak of conflict was indicated by a red area, as shown in Figure 6.1(a), which is in accordance with the manner in which the conflict's initiation is modelled.

A set of four data points, chosen by the model developer, were used as conflict input data to ensure that the latitude and longitude given as input are correctly transposed to the modelled Syrian map and that the associated conflict intensity is correctly read in. The dataset to be used as input for Case (ii) is shown in Table 6.1. The *Probability of Conflict Spreading* was set to 1 and the *Probability of Conflict Depleting* was set to 0 for consistency during the simulation runs. Three of the four locations imputed are in Syria, meaning the model should only show the conflict associated with those GPS coordinates, as this simulation model does not consider conflict that occurs outside of Syria. The conflict intensity was also fixed at 100 for consistency.

TABLE 6.1: *The dataset used for verification of the conflict data input.*

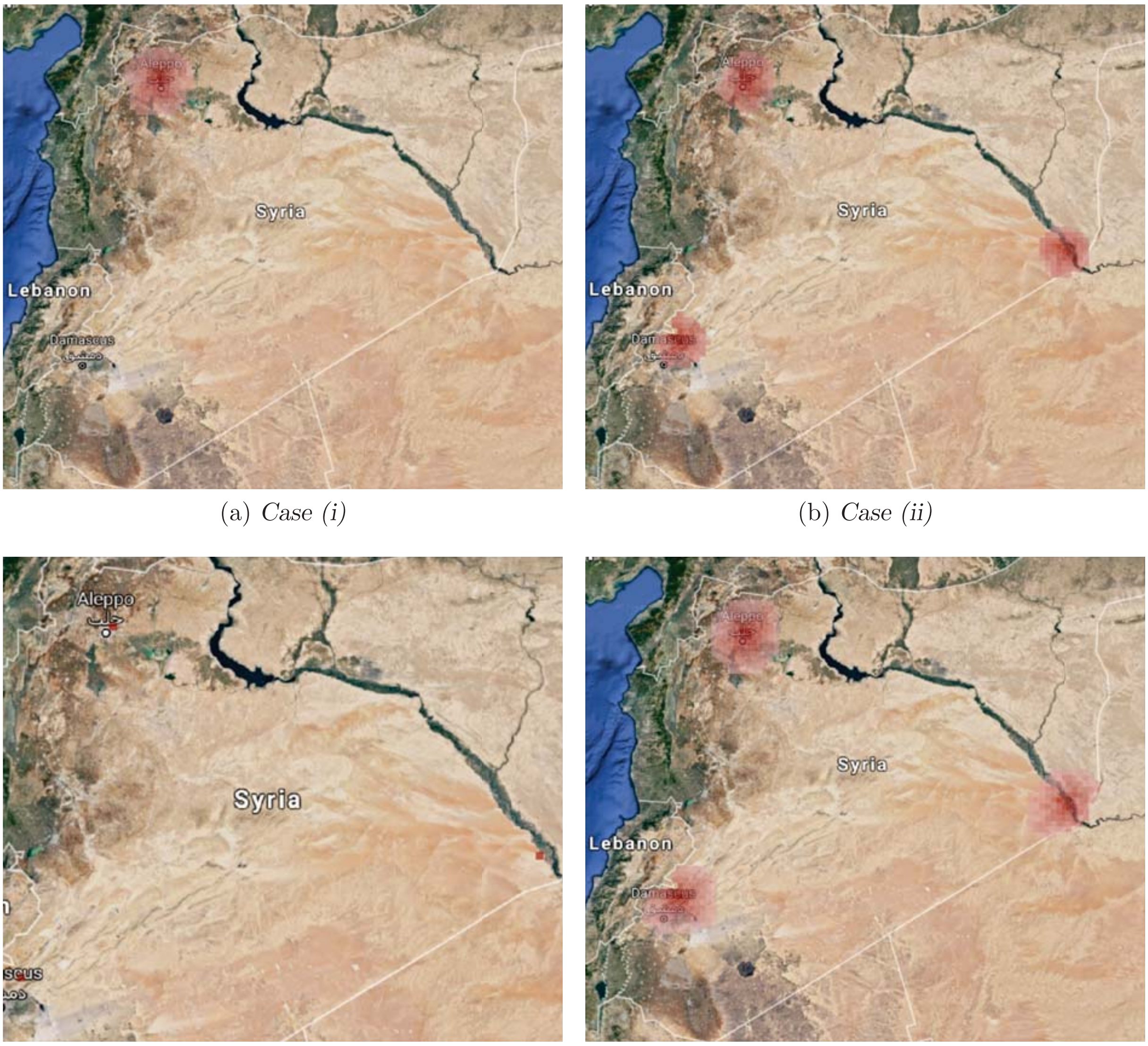
Date	Location	Latitude	Longitude	Conflict Intensity
03 January 2011	Aleppo, Syria	36.20	37.13	100
05 January 2011	Mosul, Iraq	36.36	43.16	100
07 January 2011	Damascus, Syria	33.51	36.28	100
09 January 2011	Abu Kamal, Syria	34.47	40.91	100

The *Data* conflict input method was chosen from the configuration page and the *Conflict Zones* layer checkbox was again selected to show the presence of conflict. Figure 6.1(b) illustrates the result from the simulation run at the end of January 2011, depicting the conflict within Syria. The output of the simulation run concurs with the input data given and the data input method is therefore verified.

In Case (iii), the *Probability of Conflict Spreading* was varied to ascertain the degree to which simulation model output accurately captures this phenomenon. The dataset specified in Table 6.1 was utilised as conflict input and the *Probability of Conflict Depleting* was set to 0 in order to determine exclusively the effect of the spreading. Two tests were ducted. In Test A, the

Probability of Conflict Spreading was set to 0 and, in Test B, the *Probability of Conflict Spreading* was set to 100. Figures 6.1(c) and 6.1(d) show the respective simulation model outputs for Test A and Test B at the end of January 2011.

In the output of Test A, only the point where the conflict was initiated is seen to be effected, while Test B correlates with Case (ii), where the dataset was simulated at a probability spread of 1, and a definite spread of conflict is visible. The output from these tests therefore correlates to the modelling of conflict spreading.



(a) *Case (i)*

(b) *Case (ii)*

(c) *Case (iii): Test A*

(d) *Case (iii): Test B*

FIGURE 6.1: *Cases (i)–(iii) in verifying the conflict modelling.*

Case (iv) tested the effect of the *Probability of Conflict Depletion* within a simulation run. For consistency, the *Probability of Conflict Spreading* was set to 1 and the dataset provided in Table 6.1 was implemented. Similar to Scenario (iii), two tests were executed — for Test A, the *Probability of Conflict Depletion* was set to 1 and, for Test B, the *Probability of Conflict Depletion* was set to 0. The outputs were analysed over a 12 month simulated period, in contrast to the previous tests, so as to accommodate for the effect of the depletion. The output of the simulation runs at the end of December 2011 are shown in Figures 6.2(a) and 6.2(b). Test A

(Figure 6.2(a)) shows no indication of the conflict depleting, whereas in Test B (Figure 6.2(b)) a notable decrease in conflict is seen.

Finally, the effect of variation of the *Intensity of Conflict* was tested in Case (v). For consistency the *Probability of Conflict Spreading* was set to 1 and the *Probability of Conflict Depleting* was set to 0. The data given as input to the conflict were adjusted to accommodate a varying conflict intensity, as shown in Table 6.2. Figure 6.2(c) indicates the simulation model output at the end of January 2011 for Case (v). The difference between the conflict intensity at 100, 50 and 10, as can be seen on the figure portraying the output, is notable and verifies the application of this input variable.

TABLE 6.2: *The dataset used for verification of the conflict intensity.*

Date	Location	Latitude	Longitude	Conflict Intensity
03 January 2011	Aleppo, Syria	36.20	37.13	100
07 January 2011	Damascus, Syria	33.51	36.28	10
09 January 2011	Abu Kamal, Syria	34.47	40.91	50

Through the implementation of Cases (i)–(v), the manner in which conflict is modelled within the simulation model is deemed to be successfully verified. A number of input measures which relate to the modelling of conflict were varied and their outputs compared to logical reasoning in the process of verifying the conflict modelling.

6.3 Verification of the modelled population

The modelling of people should reflect the Syrian population, irrespective of the presence of conflict. This includes agents possessing certain characteristics mimicking those of the actual population, the ageing of agents and the effect of births and deaths on the population size, the different stages an agent can be in and an agent's ability to withstand conflict based on their attributes. A number of cases are implemented in the model and the output from the simulation runs are explored. The following cases were considered:

Case (vi): Agent attribute statistics

Case (vii): The fluctuation of agent population

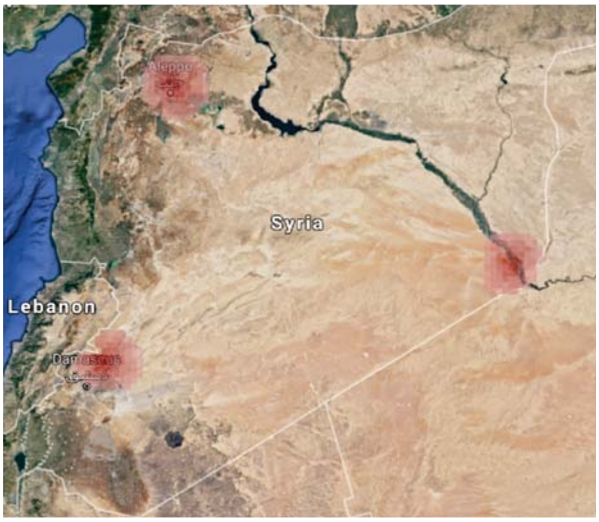
Case (viii): The ageing of a single agent

Case (ix): An agent changing states

Case (x): The moving threshold of agents

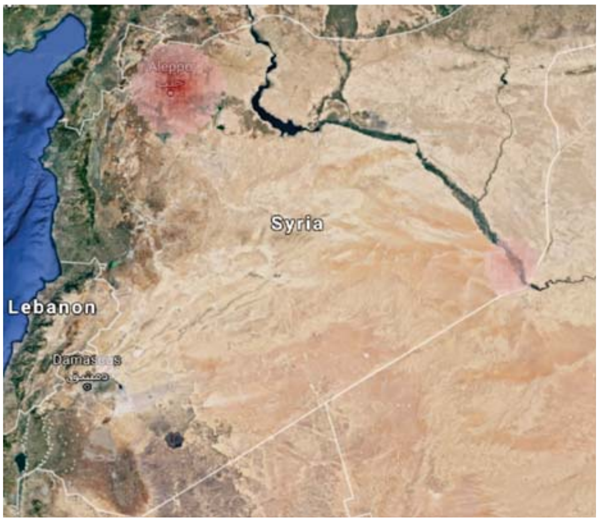
Case (xi): The geographic distribution of agents

In Case (vi), the agent attributes as modelled were investigated to verify their correlation to the agent attributes introduced in §5.5.1. A series of 30 simulation runs were executed and data pertaining to the current population was recorded in order to determine a mean distribution per attribute for the different criteria. The correlation is affirmed with the determined percentage error between the simulated and actual distributions falling below one percent, as can be seen in Table 6.3 (where AAD refers to the anticipated age at death). The only attribute not included in the table is the probability of a person having international family. The simulation is modelled



(a) *Case (iv): Test A*

(b) *Case (iv): Test B*



(c) *Case (v)*

FIGURE 6.2: *Cases (iv)–(v) in verifying the conflict modelling.*

such that every agent has a probability, p, where p is of a uniform distribution between 5% and 20%. The output indicates that, over the 30 runs executed, the average probability of an agent having family internationally was 12.57%, which aligns with the input data.

TABLE 6.3: *The percentage error between the actual and simulated agent attributes.*

Gender		Age			Tertiary Education		Economic Status			AAD	
Male	Female	A	B	C	True	False	Low	Medium	High	Male	Female
0.25%	0.25%	0.21%	0.19%	0.01%	0.35%	0.35%	0.11%	0.01%	0.13%	0.84%	0.86%

Case (vii) examined the fluctuation in population size over the 6-year simulated period, taking births and deaths into account. No conflict is initiated during the simulation experiment in order to disregard deaths which might occur as a result of conflict. As the simulation model accounts for the annual births at a single annual instance, the graph, as shown in Figure 6.3, indicates a sharp increase in population size yearly, followed by a slight decrease which accounts for the deaths. Over the 6-year simulated period the overall population increase was less than 4%, which indicates a small fluctuation in the total population size, verifying the implementation of births and deaths in the simulation model. Births are taken into account annually at a single instance and is modelled as such owing to it being less computationally intense than creating various events which would spread the occurrence of births over the year.

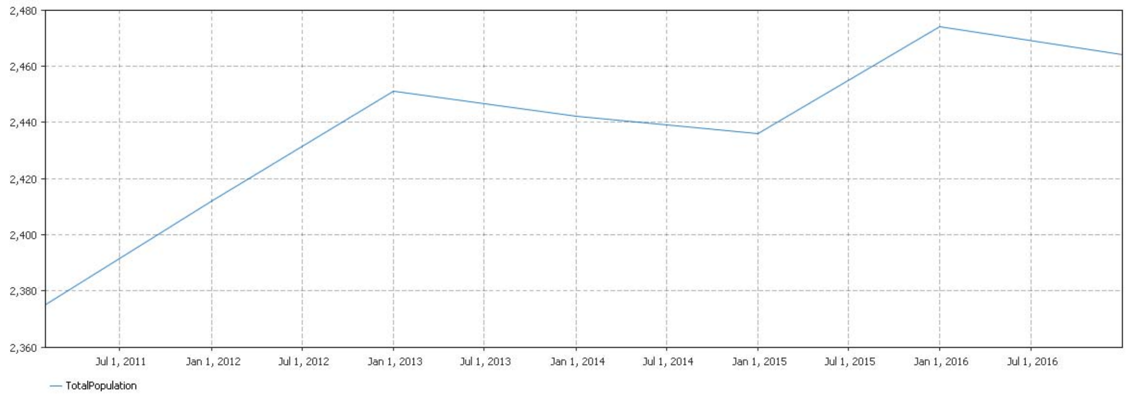


FIGURE 6.3: *The fluctuation in population size verified.*

Another modelled element which requires verification is the ageing of agents, as addressed by Case (viii). In order to test if the model incrementally increases the age of each agent annually, a simulation experiment was executed with only one agent in the population. A dynamic textbox indicating the age of the agent was implemented above the agent during the simulation to assist the modeller in verifying the increase in age as depicted in Figure 6.4. The agent had an initial age of 24 years which increased to an age of 29 years over the 6-year modelled period.

In order to test Case (ix), the simulation model was altered to change the colour of an agent when changing states. Each state was associated with a specific colour in order to ascertain whether or not an agent progressed to the correct state according to its modelled action. Agents in the `Residing` state are black ( ) and as soon as they move to the `SearchMove` state where they search for a safer destination, they turn to gold ( ). An agent that has chosen a new destination turns to light blue ( ) as it begins movement. When the agents arrive at their destination, whether within Syria or in another country, they turn green ( ) as they enter the `CheckSurroundings` state. If, however, an agent fails to find a new destination, thereby directing it to the `Stuck` state, it will turn pink ( ). If it were to escape and enter the `MoveThroughConflict` state, it will turn purple ( ). In the test simulation experiment executed, the visual animation of agents changing states confirmed to correspond to the imposed modelling rules. Figure 6.5 shows a screenshot of the simulation experiment which illustrates agents active in some of the states by means of the colour variation of their presentation element.

Case (x) investigates the `MovingThreshold` variable allocated per agent, which indicates an agent's ability to withstand conflict. A number of scenarios were considered where the average moving threshold of agents sharing similar attributes were compared to reality (as summarised in §5.6.1). Scenario A considers people above the age of 65, who do not have any family internationally. These people are expected to have a relatively high moving threshold. Scenario B refers to adults, between the ages of 20 and 40, who received tertiary education. According to research, these people should have a relatively low moving threshold. Scenario C investigates people of low economic status, who do not have tertiary education. These individuals should also have a relatively high moving threshold. A number of 30 simulation runs were executed in order to calculate the average `MovingThreshold` and the associated standard deviation of each scenario. Table 6.4 depicts the output values.

(a) *Case (viii) The test agent in January 2011.*

(b) *Case (viii) The test agent in January 2012.*

(c) *Case (viii) The test agent in January 2013.*

(d) *Case (viii) The test agent in January 2014.*

(e) *Case (viii) The test agent in January 2015.*

(f) *Case (viii) The test agent in January 2016.*

FIGURE 6.4: *The verification of an agent ageing.*

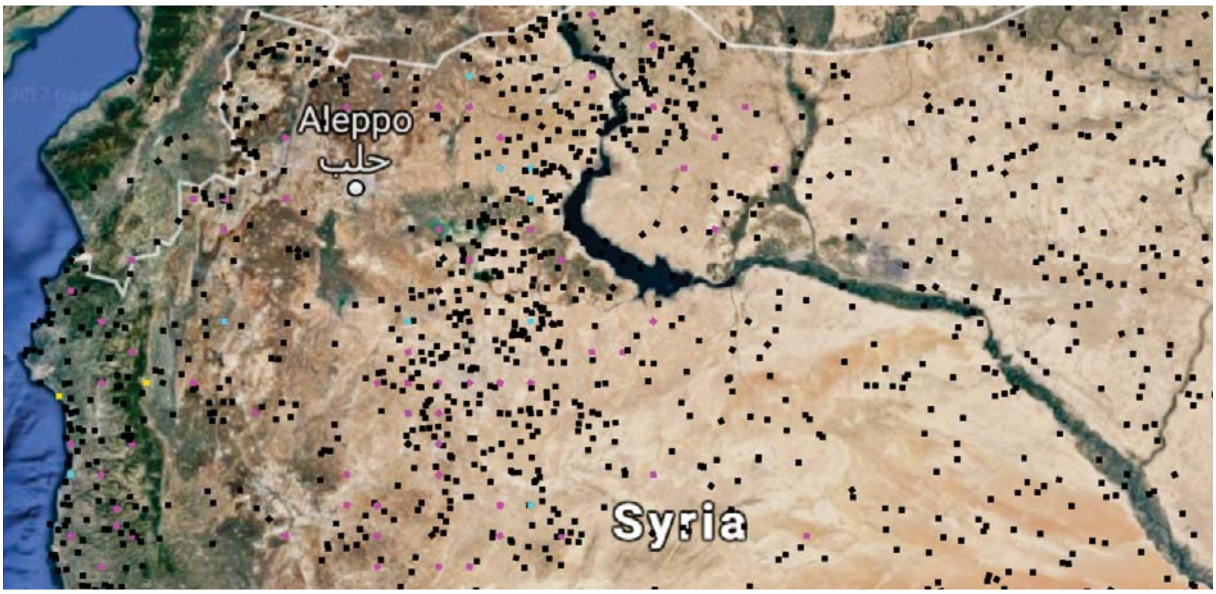


FIGURE 6.5: *An example of the agents' states and associated colours during a simulation experiment.*

The output showed that Scenarios A and C have equally high moving threshold values, while Scenario B's moving threshold is less than two-thirds of the value of Scenarios A and C. The output correlates with the assumptions made with respect to certain characteristics influencing a person's inclination to move and the small associated standard deviations prove that the MovingThreshold values do not deviate too greatly from this norm.

TABLE 6.4: *The output from Case (x) verifying the allocation of a moving threshold to agents.*

	Average	Standard Deviation
Scenario A	0.67206	0.029
Scenario B	0.42511	0.048
Scenario C	0.67143	0.014

The geographic distribution of people is considered in Case (xi). The modelling thereof, as explained in §5.5.2, ensures that the initial population of Syrians is dispersed across the country, mimicking reality. Figure 6.6 illustrates the spread of agents (marked as small black squares) over Syria. It can be seen that, as suggested in the data, the population density is greater in the North West and South West of the country, since these are the major cities in Syria.

FIGURE 6.6: *The geographical distribution of people in Syria.*

6.4 Verification of the modelling of decision-making

The modelling of the decision-making process of agents, as discussed in §5.6, includes an agent's decision in terms of which movement type to adopt and then, depending on this choice, the selection of a proposed destination. In order to verify this process and the manner in which it is modelled, the following cases were tested:

Case (xii): The agent movement type chosen based on attributes

Case (xiii): The IDPs' *Attractive Zones* and associated weighted criteria

Case (xiv): The effect of varying openness scores

Case (xv): A walk-through verification of an agent of each movement type

Case (xii) investigates the correlation between the combination of attributes of an agent and its associated movement type. The MovementType most commonly associated with a specified set of characteristics were tested and three types of people were considered. Type A refers to a person above the age of 65 who does not have tertiary education, has no international family and is of low economic status. Type B refers to a child of age 15 or younger, who is from a family of medium income and has no family living abroad. Type C refers to a person between the ages of 15 and 65 who has received tertiary education, is of a high income class and has international family.

According to research documented in §5.6.1, a person of Type A is likely to choose to move as an IDP (movement type 1), while a person of Type B is like to consider fleeing as a refugee (movement type 2) and a person of Type C is likely to move as an undocumented migrant (movement type 3). The simulation model was adapted for the testing of this case in order to calculate a person's choice of movement type at initiation of a simulation run. Three random agents, each prescribing to the characteristics of one of the types described above, were selected during each run and their movement types noted. After 30 simulation runs had been executed, the movement types most commonly associated with the three different types of people proved similar to what has been assumed in modelling this decision. Table 6.5 reflects the output of the simulation experiment. Cases A and C distinctly associates with movement types 1 and 3, respectively, and, although the mode of Case B associates with movement type 2, it can be seen that set of characteristics portrayed by Case B may associate with both movement types 1 and 2.

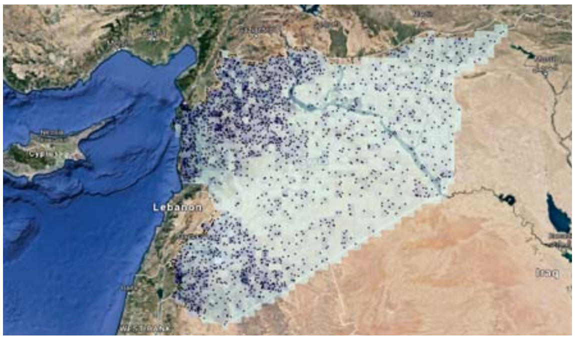
Table 6.5: *The verification of allocation of movement types.*

	Case A	Case B	Case C
Mode	1	2	3
Number of times Type 1 allocated	25	11	2
Number of times Type 2 allocated	4	13	9
Number of times Type 3 allocated	1	6	19

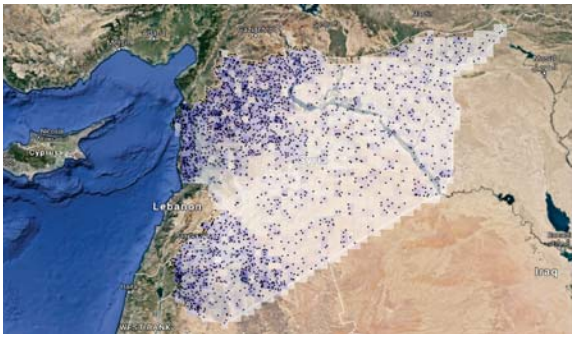
In order to verify the inclusion of so-called *Attractive Zones*, Case (xiii) employs existing data portraying the conflict in Syria during the simulated period in order to compare the changes in *Attractive Zones*, based on conflict zones and population density, as a simulation experiment progresses. A fixed seed and default parameters are employed to ensure consistency in the verification process.

At the initiation of a simulation experiment when no conflict has been initiated, according to the manner in which the *Attractive Zones* are calculated (as explained in §5.6.2), the attractiveness per cell should equal the normalised population density value per cell. Figure 6.7(a) depicts the *Population Density* layer and Figure 6.7(b) indicates the *Attractive Zones* at the initiation of the simulation run. As is apparent, the attractiveness of each cell relates directly to the population density thereof. Three scenarios were executed to indicate the influence of associating weights to the criteria, conflict and population density, in determining the *Attractive Zones*.

In Scenario I, the weights associated with the conflict and population density criteria were set to favour the areas with no conflict, having a 0.9:0.1 ratio, where conflict contribute 0.9 of the weight, with population density factor contributing the balance. The spread of conflict at the end of the simulated period, illustrated in Figure 6.8(a), is the same for all three scenarios owing to the fixed seed (configured for this case) and certain parameters kept constant. For this scenario, the *Population Density* is shown in Figure 6.8(b) and the *Attractive Zones* is shown in Figure 6.8(c). Figure 6.8(d) shows both the *Attractive Zones* and the *Conflict* superimposed over the simulated area within Syria.

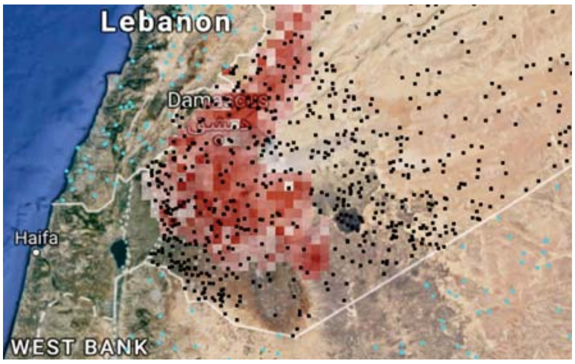


(a) *Case (xiii) Population Density* at the start of the simulation run

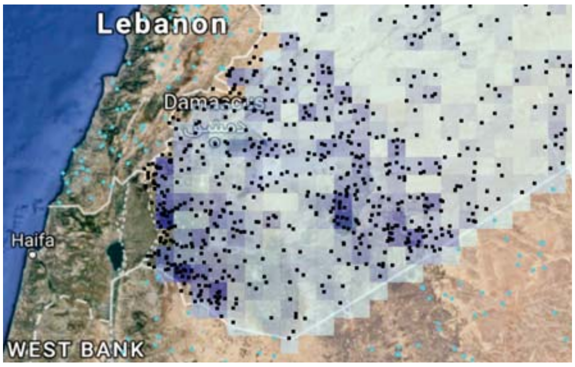


(b) *Case (xiii) Attractive Zones* at the start of the simulation run

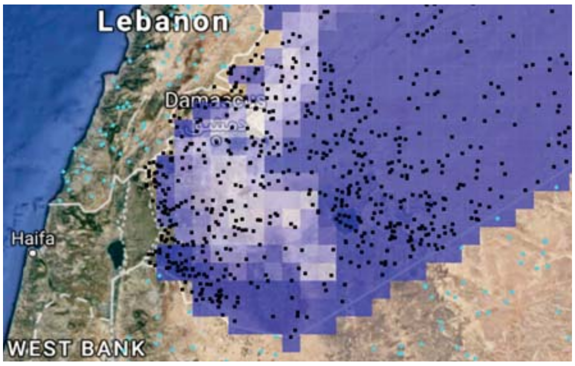
FIGURE 6.7: *The layers as shown at the start of the simulation.*



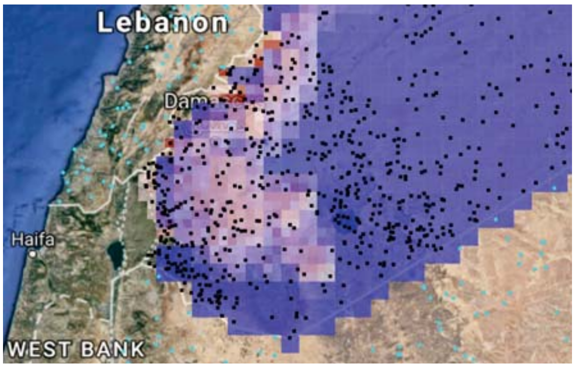
(a) *Case (xiii): Scenario I Conflict*



(b) *Case (xiii): Scenario I Population Density*



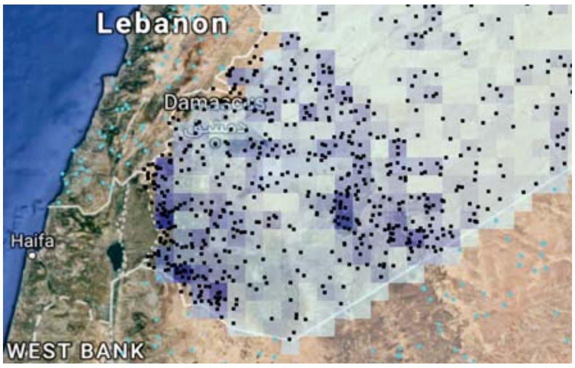
(c) *Case (xiii): Scenario I Attractive Zones*



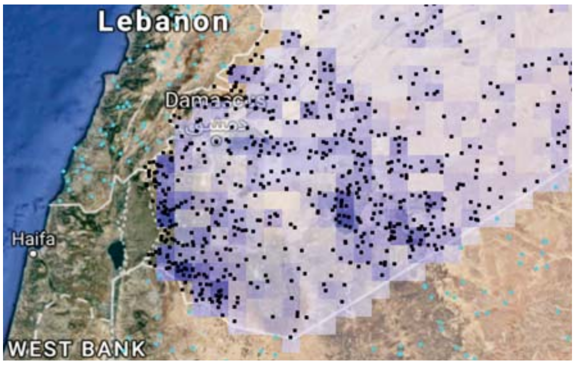
(d) *Case (xiii): Scenario I Attractive Zones* and *Conflict*

FIGURE 6.8: *Scenario I in verifying the Attractive Zones.*

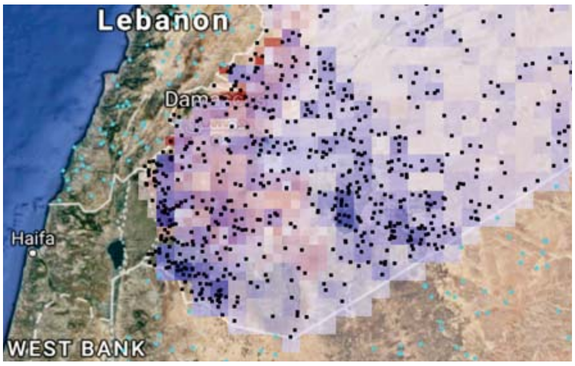
The areas which are more attractive to IDPs are identified by the dark purple hue, as shown in Figures 6.8(c) and 6.8(d), and it is apparent that the location of these areas is almost an exact inverse of areas where conflict is present. Scenario II considered a situation with altered criteria weights such that the conflict factor contributes 0.1 of the total weight, with the population density factor contributing to the balance. The *Population Density* as at the end of the simulation experiment is shown in Figure 6.9(a), with the *Attractive Zones* illustrated in Figure 6.9(b). Figure 6.9(c) shows the effect when the *Conflict* and *Attractive Zones* layers are superimposed over the simulated area.



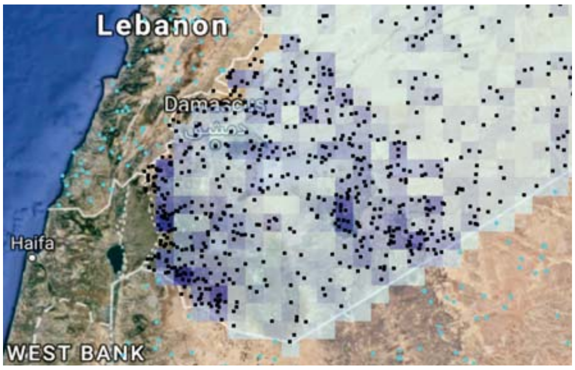
(a) *Case (xiii): Scenario II Population Density*



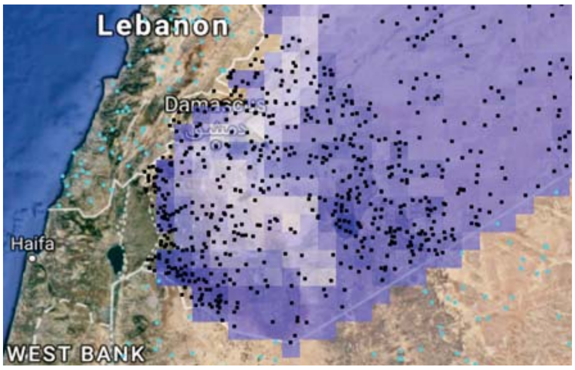
(b) *Case (xiii): Scenario II Attractive Zones*



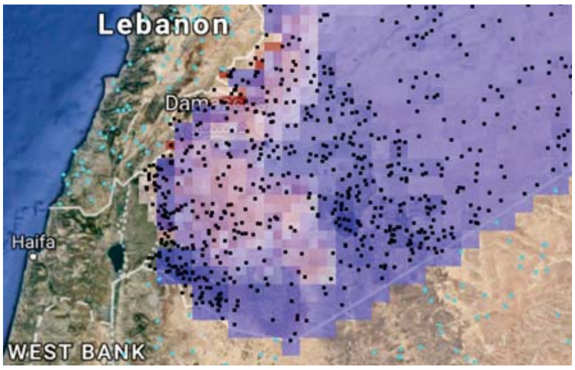
(c) *Case (xiii): Scenario II Attractive Zones and Conflict*



(d) *Case (xiii): Scenario III Population Density*



(e) *Case (xiii): Scenario III Attractive Zones*



(f) *Case (xiii): Scenario III Attractive Zones and Conflict*

FIGURE 6.9: *Scenarios II and III in verifying the Attractive Zones.*

This scenario naturally favours the population density in determining the attractiveness of an area. This is illustrated visibly as the *Attractive Zones* shown in Figure 6.9(b) are notably similar to the *Population Density* shown in Figure 6.9(a), indicating that the conflict impose a significant effect.

Finally, for Scenario III, the weights for the criteria influencing the *Attractive Zones* were adjusted to prescribe to the default model values such that conflict contributes 0.6 of the weighting and population density contributes the balance. Figures 6.9(d), 6.9(e) and 6.9(f) illustrate the *Population Density*, *Attractive Zones* and a combination of both these layers for this scenario, respectively.

In this case, it can be seen that, although the areas with less conflict are distinctly more attractive, the population density also influences the attractiveness of an area. By means of the three scenarios executed the process determining the *Attractive Zones* and the effect of the weighted criteria are successfully verified.

The effect of the scores portraying the openness of neighbouring countries towards refugees and undocumented migrants are considered in Case (xiv). In order to verify this, the initial openness score for one neighbouring country was set to 80, while all the other neighbouring countries were allocated an initial value of 10. The weights of the criteria determining which country a refugee or undocumented migrant will choose as destination are set to a distance, popularity and openness score ratio of 0.05:0.05:0.9 respectively, and kept constant for the purpose of consistency.

The fluctuation of openness scores during the simulated times were disabled in this case in order to focus on the effect of the openness score itself and the simulation runs were executed over the full 6-year period. Figures 6.10– 6.14 illustrate the output when varying these scores, allocating each neighbouring country the significantly higher openness score in a sequence. In each case, it is apparent that the neighbouring country with an allocated openness score of 80 has a significantly higher influx of Syrians over the simulated period than all the other neighbouring countries.

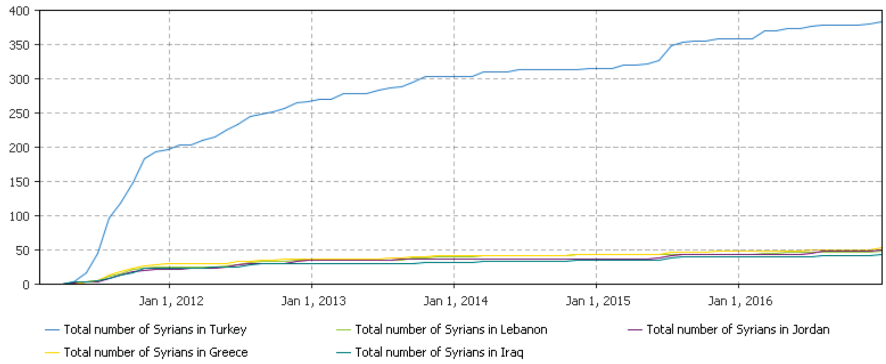


FIGURE 6.10: *Allocation of an Openness Score of 80 to Turkey.*

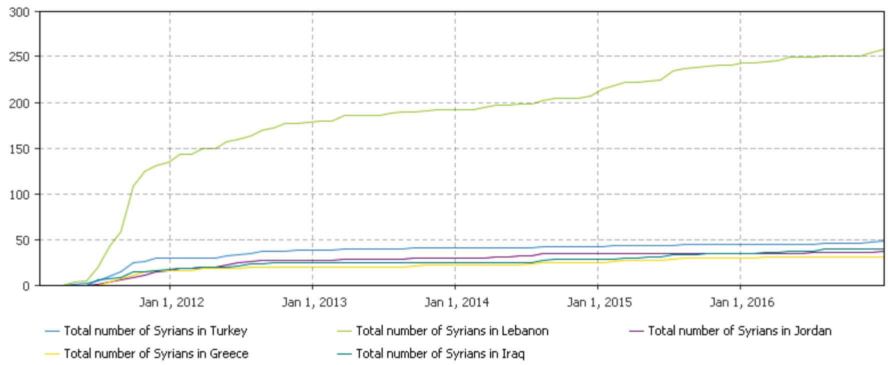


FIGURE 6.11: *Allocation of an Openness Score of 80 to Lebanon.*

Case (xv) investigated the process an agent goes through when observing one of each of the different modelled movement types. Scenarios considering IDPs, refugees and undocumented migrants where therefore investigated. The colour of the agent identified in each scenario was changed to white in order to make the agent more distinguishable for movement tracking. All parameters were set to default values and the conflict was initiated manually at the location

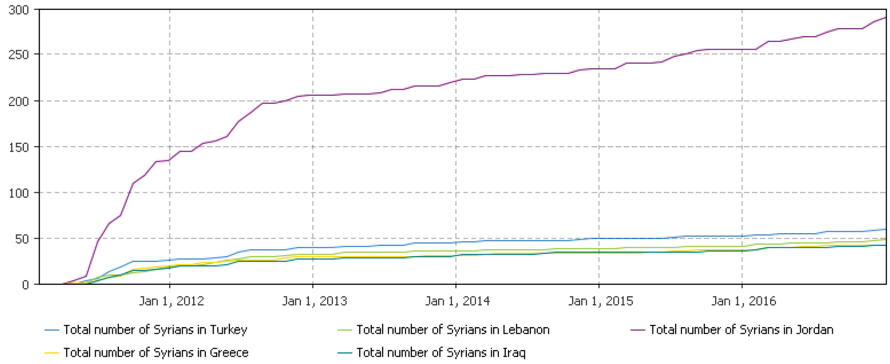


FIGURE 6.12: *Allocation of an Openness Score of 80 to Jordan.*

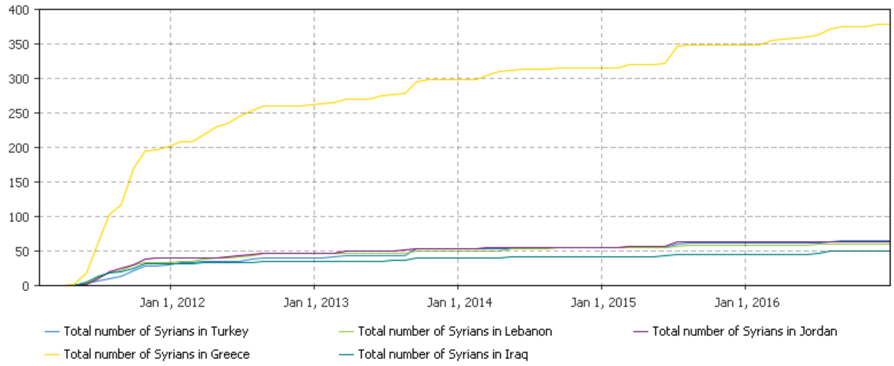


FIGURE 6.13: *Allocation of an Openness Score of 80 to Greece.*

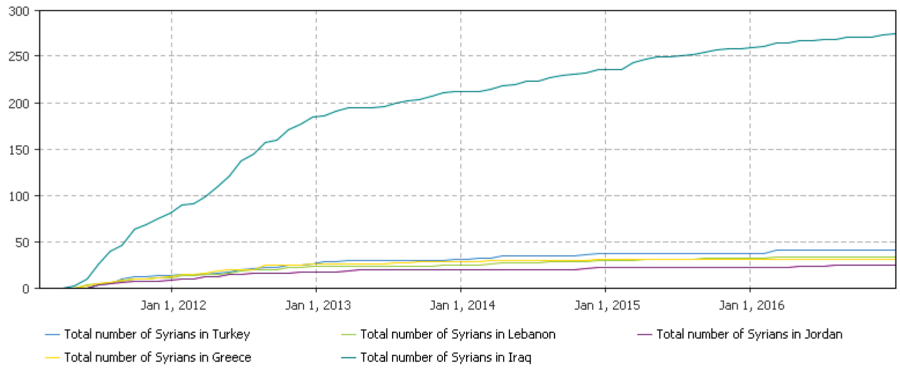
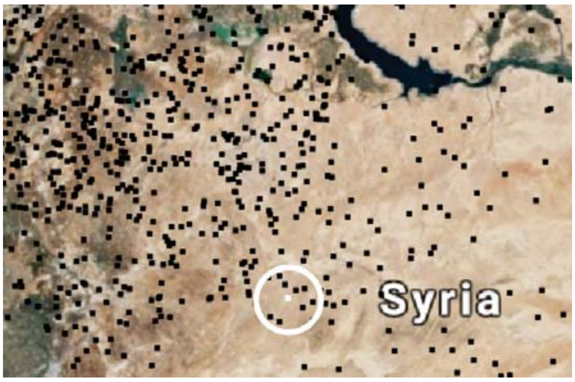


FIGURE 6.14: *Allocation of an Openness Score of 80 to Iraq.*

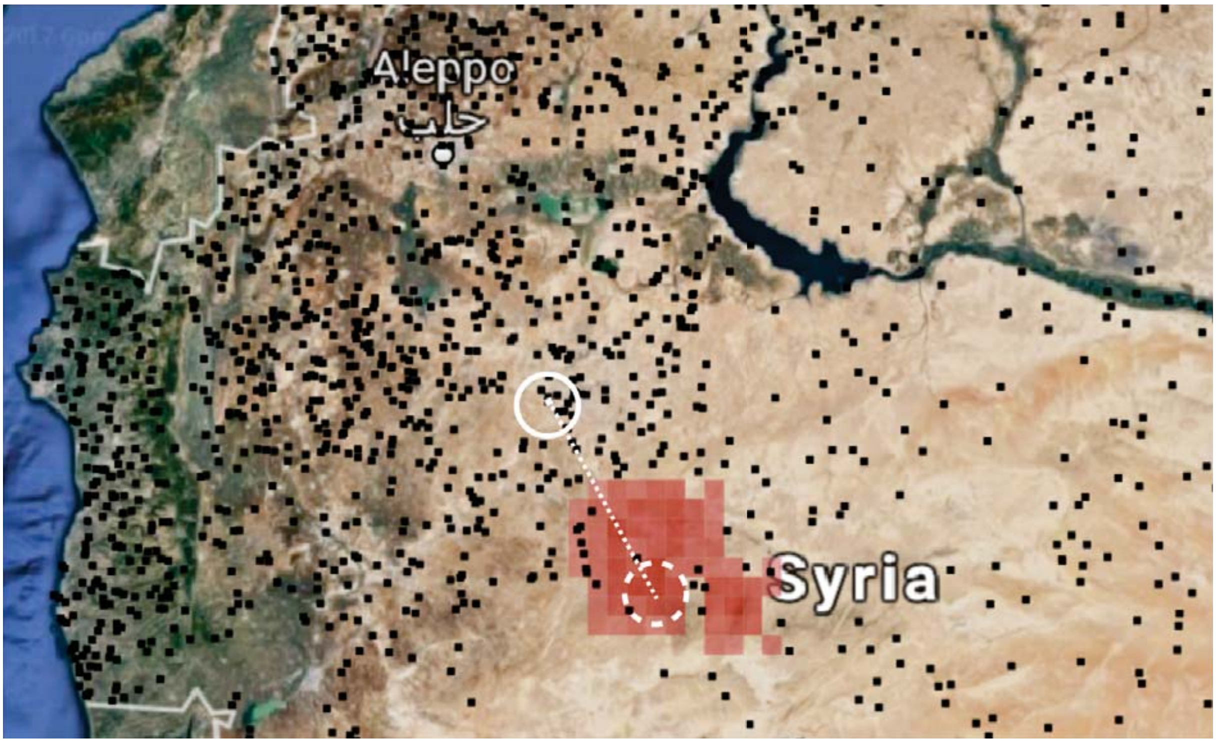
of the agent in order to investigate its reaction towards it. In each scenario, a random agent was chosen and its progress through its statechart was documented. The scenarios are depicted by Figures 6.15, 6.16 and 6.17, where a white circle has been placed over the agent under investigation for illustrative purposes.

Scenario IV depicts the movement of an agent who, when exposed to conflict exceeding its threshold, chooses to move to relocate within Syria. Figure 6.15 portrays the movement of the observed agent from the initiation of the simulation experiment to the point of relocation. The agent had a moving threshold of 0.49 (49%), which was exceeded by conflict, shown in Figure 6.15(b), with an intensity value of 80. The agent decided to move according to movement type 1 and identified a location North West of its original location. It then moved there, as shown in Figure 6.15(c). The agent then checked the surroundings upon arrival and deemed the area safe before progressing to the state of `Residing`.



(a) *Case (xv): Scenario IV experiment initiation*

(b) *Case (xv): Scenario IV exposed to conflict*



(c) *Case (xv): Scenario IV track movement*

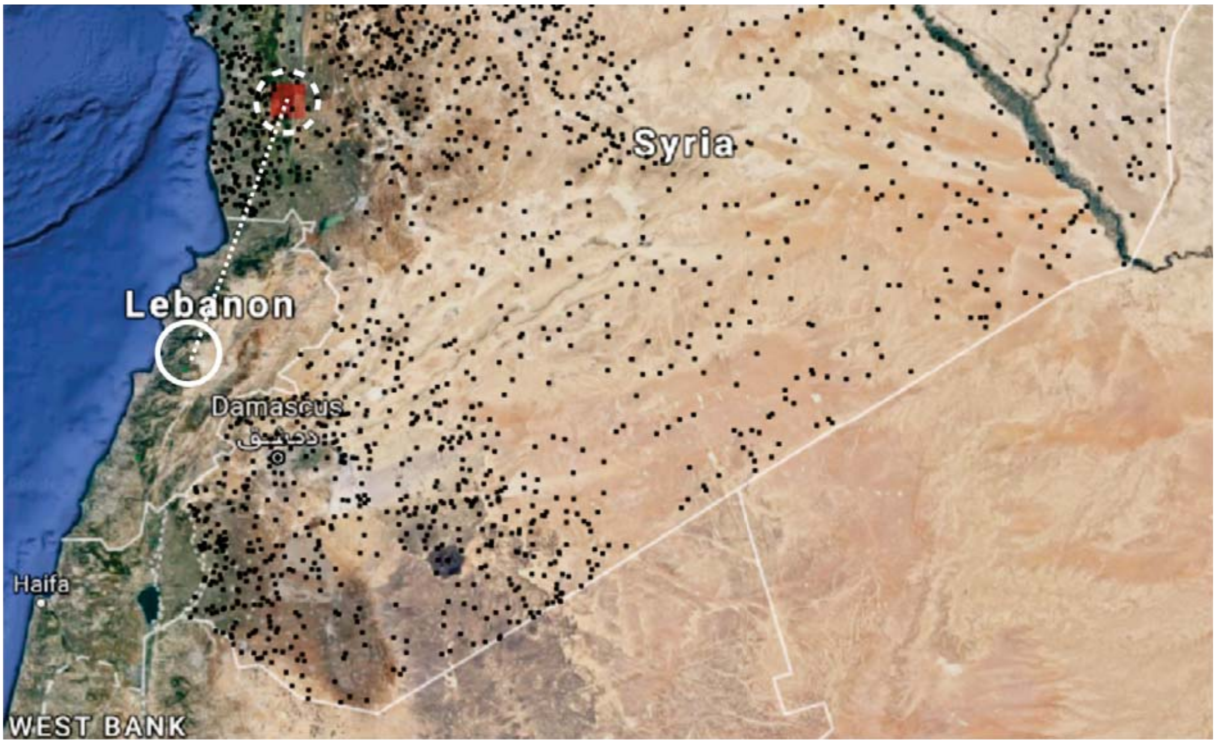
FIGURE 6.15: *The movement of Syrian IDPs.*

In Scenario V, the movement of a refugee is considered and illustrated in Figure 6.16. The agent had a moving threshold of 0.47 (47%) at the initiation of the simulation experiment which was exceeded by conflict of an intensity value of 62, as shown in Figure 6.16(b). The agent decided to move according to movement type 2 and chose Lebanon as the neighbouring country in which to seek refuge. Figure 6.16(c) illustrates the movement of the agent from its original location to Lebanon. Upon arrival, the agent immediately entered the `Residing` state, as conflict outside of Syria is not considered in the model.

The movement of an undocumented migrant portrayed by a random agent is followed in Scenario VI and depicted in Figure 6.17. The agent possessed a moving threshold of 0.15 (15%) which was exceeded by a conflict intensity of value 62, shown in Figure 6.17(b). After determining its movement type as of type 3, the agent selected Jordan as its proposed destination and moved there, as illustrated in Figure 6.17(c). The output of all three scenarios investigated in Case (xv) indicates that the behaviour of agents randomly chosen of each respective movement type correlates to the predicted behaviour as modelled.

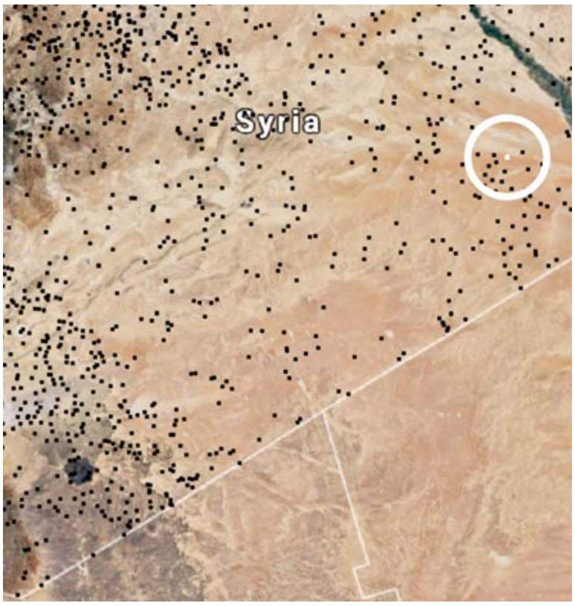
(a) *Case (xv): Scenario V experiment initiation*

(b) *Case (xv): Scenario V exposed to conflict*

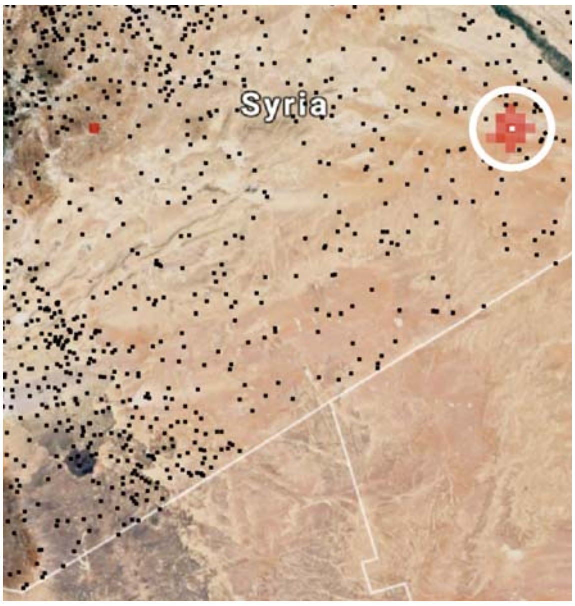


(c) *Case (xv): Scenario V track movement*

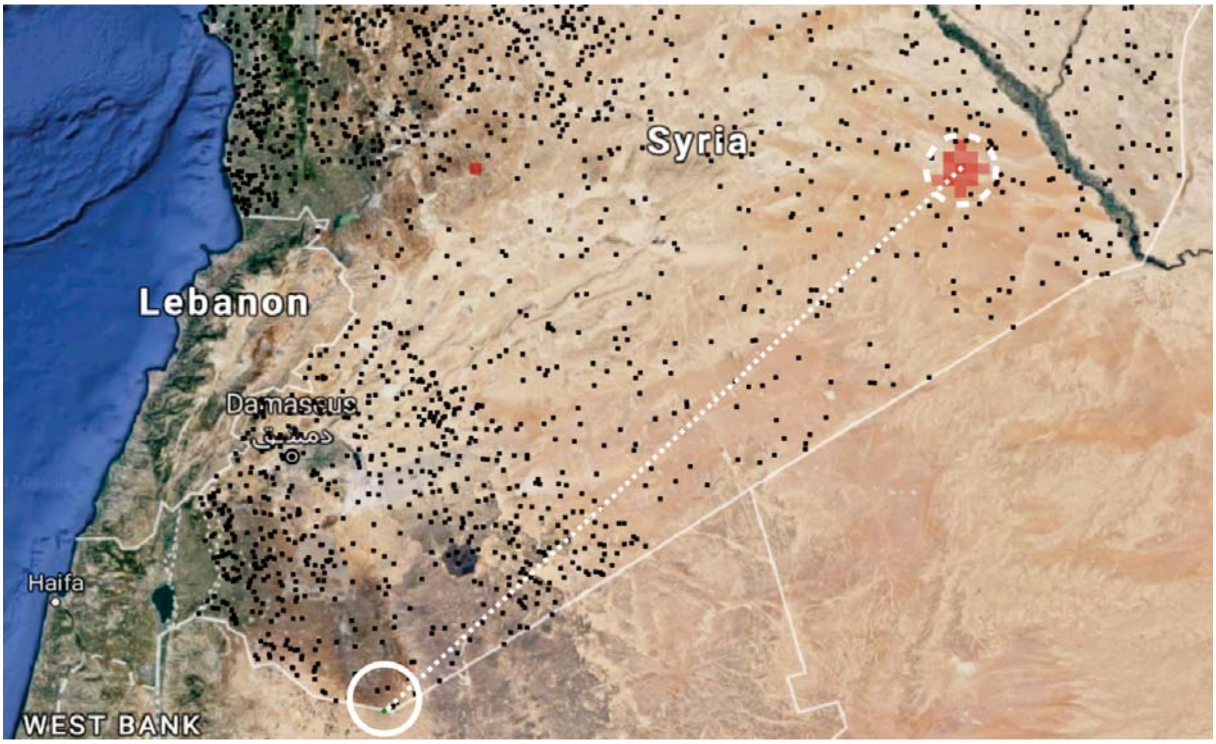
Figure 6.16: *The movement of a Syrian refugee.*



(a) *Case (xv): Scenario VI experiment initiation*



(b) *Case (xv): Scenario VI exposed to conflict*



(c) *Case (xv): Scenario VI track movement*

FIGURE 6.17: *The movement of an undocumented migrant from Syria.*

6.5 Chapter summary

The verification processes followed in order to ensure the correct operation of the different components of the model described in this thesis were discussed in this chapter. This included a verification of the modelled conflict, the population of agents residing in the model and the subsequent of decision-making of these agents when confronted with conflict. Illustrative and numeric outputs were gathered from the model during specific tests in order to assess the degree to which the aforementioned model components perform as intended. In all cases, the implementation was found to be sound and accurate and the response of the model appears to be in line with what was expected from its construction. In light of this, the model is deemed to have been appropriately verified.

CHAPTER 7

Model validation and analysis

Contents

This chapter comprises the model validation and output data analysis, as introduced in §7.1, of the agent-based model developed in this thesis. The model calibration and validation with respect to the manner in which conflict is modelled are performed in §7.2 and §7.3, respectively. This is followed by a face validation in §7.4, which includes collaboration with subject matter experts on various important components of the modelled. Existing data is then considered by means of a parameter establishment analysis in §7.5 where suitable model parameters are sought to recreate the documented scenarios. In light of the model's capability to recreate specific scenarios pertaining to conflict outbreak and associated fleeing of forcibly displaced people, the framework through which this means of parameter variation can be achieved by employing the model as a decision support and analysis tool is considered in §7.6.

7.1 Model validation and output data analysis

A significant component in the simulation model development process is determining the extent to which a simulation model accurately represents the real-world system under consideration. This phase consists primarily of the verification and validation of the simulation model. Verification of the agent-based model in this study was detailed previously in Chapter 6. *Validation* differs form verification in that it aims to determine whether or not the simulation model is a true representation of the actual system to the extent necessary in order to meet the model objectives [121].

In a similar vein, *output analysis* seeks to obtain an accurate estimated performance of the simulation model by determining the simulation model's true parameters and characteristics during model execution [118]. The output of this simulation model is dependent on the input values and initial configuration set by the user which, in turn, determines the values of parameters during and after a simulation execution. Unfortunately, detailed data pertaining to real-world elements which are modelled through the inclusion of specific input parameters are not widely available and, in light of this, the attribution of values to these parameters upon simulation start-up are left to the discretion of the user, with ‘base case’ or ‘best guess’ values included for convenience. This, in effect, aims to allow any desired experiment to be conducted through the selection of an appropriate set of input variables. The output of the simulation, using various input scenarios, may then be examined by experts in the field or compared to the output data which do exist as a form of validation [12]. Furthermore, AnyLogic lends itself to a visual output analysis which allows for easy assessment of the consequent model development, thereby contributing to this form validation.

7.2 Calibration of parameters used to model conflict

In an attempt to validate the conflict as modelled within the simulation, the population of agents was temporarily disregarded. The primary parameters associated with the modelling of conflict, `ProbabilityofInfection` and `ProbabilityofDepletion`, were then varied in an attempt to find values which best replicate the historical spread of conflict in Syria — historical data of which exists on record in the form of graphical mapping. The other parameters and variables were kept constant and a fixed seed random number generator was implemented in the model to ensure consistency between comparative experiments.

The `ProbabilityofInfection` and the `ProbabilityofDepletion` parameters were varied from a minimum value of 0.1 to a maximum value of 0.9, in intervals of 0.2, resulting in 25 simulation experiments. These parameter combinations are listed in Table 7.1. Appendix A shows the visual model outputs for each of these specifically configured experiments.

The general conflict distribution and spread is relatively similar in all of the experiments performed, although, as expected, the spread of conflict is greatest when the `ProbabilityofInfection` tends towards 1 and the `ProbabilityofDepletion` tends towards 0. The simulation model is therefore configured as having a `ProbabilityofInfection` and a `ProbabilityofDepletion` both of triangular distribution with 0 as minimum, 1 as maximum and 0.75 as mode. These distributions could then be subjected to a comparative validation, as described in the following section.

7.3 Validation of the modelling of conflict

Validation is commonly achieved by utilising the simulation model to replicate empirical real-world data, as the ability of the model to imitate existing data would constitute it valid. In this case, numerical data with respect to the spread and depletion of conflict in Syria do not exist, however, graphical data indicating the state of conflict during various time instances throughout the conflict’s lifetime in the area are available. AnyLogic facilitates the animation of the model output which was utilised for the validation of conflict in this instance. The inclusion of such a visual framework affords the ability to observe the simulated conflict as it spreads and depletes over time. The recorded visual data was then compared to this simulated output and, in doing so, the degree of visual replication contributes to the conviction of the model’s accuracy.

Visual data pertaining to the state of conflict in Syria during specific time instances were gathered from *The Carter Center* [123], GDELT [125] and *Palantir Technologies* [124]. Simulation experiments were conducted using a fixed seed for consistency and the `ProbabilityofInfec-`

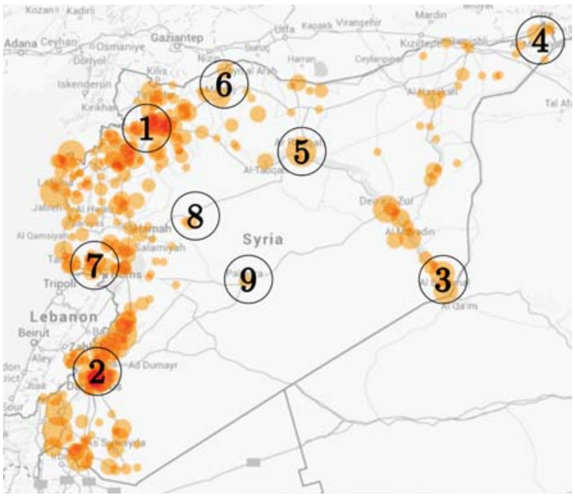
TABLE 7.1: *The list of experiments and corresponding parameter values in varying the spread of conflict with respect to the probability of spread and depletion.*

Experiment	Probability of Conflict Spreading	Probability of Conflict Depleting
1	0.1	0.1
2	0.3	0.1
3	0.5	0.1
4	0.7	0.1
5	0.9	0.1
6	0.1	0.3
7	0.3	0.3
8	0.5	0.3
9	0.7	0.3
10	0.9	0.3
11	0.1	0.5
12	0.3	0.5
13	0.5	0.5
14	0.7	0.5
15	0.9	0.5
16	0.1	0.7
17	0.3	0.7
18	0.5	0.7
19	0.7	0.7
20	0.9	0.7
21	0.1	0.9
22	0.3	0.9
23	0.5	0.9
24	0.7	0.9
25	0.9	0.9

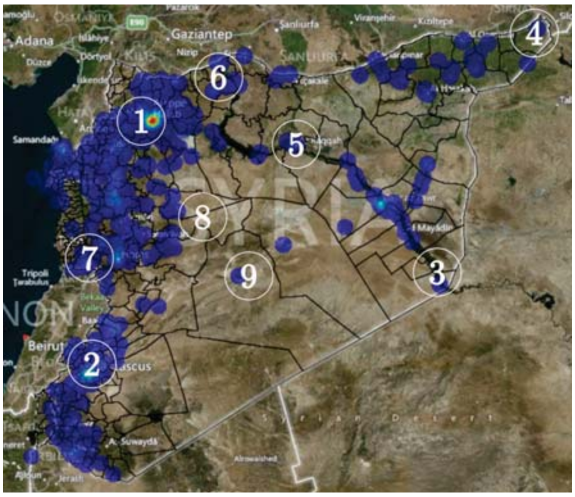
`tion` and `ProbabilityofDepletion` parameters were employed by means of a triangular distribution, as calibrated in §7.2, while all other parameters were held constant with no agents included in the model runs. The model output for three specific time instances are compared to the visual data available, as shown in Figures 7.1, 7.2 and 7.3. In each of these visual comparisons, numerical annotations are superimposed on the images for the purpose of analytical discussion.

Figure 7.1 shows a comparison between two sources who record the conflict situation in June 2014, namely GDELT and *Palantir Technologies*, with the conflict simulated for the same time period. The conflict intensity is shown in Figure 7.1(a) by means of a single colour intensity, whereas Figure 7.1(b) illustrates the intensity of conflict ranging over a colour spectrum from dark blue to red, where the latter suggests a high intensity.

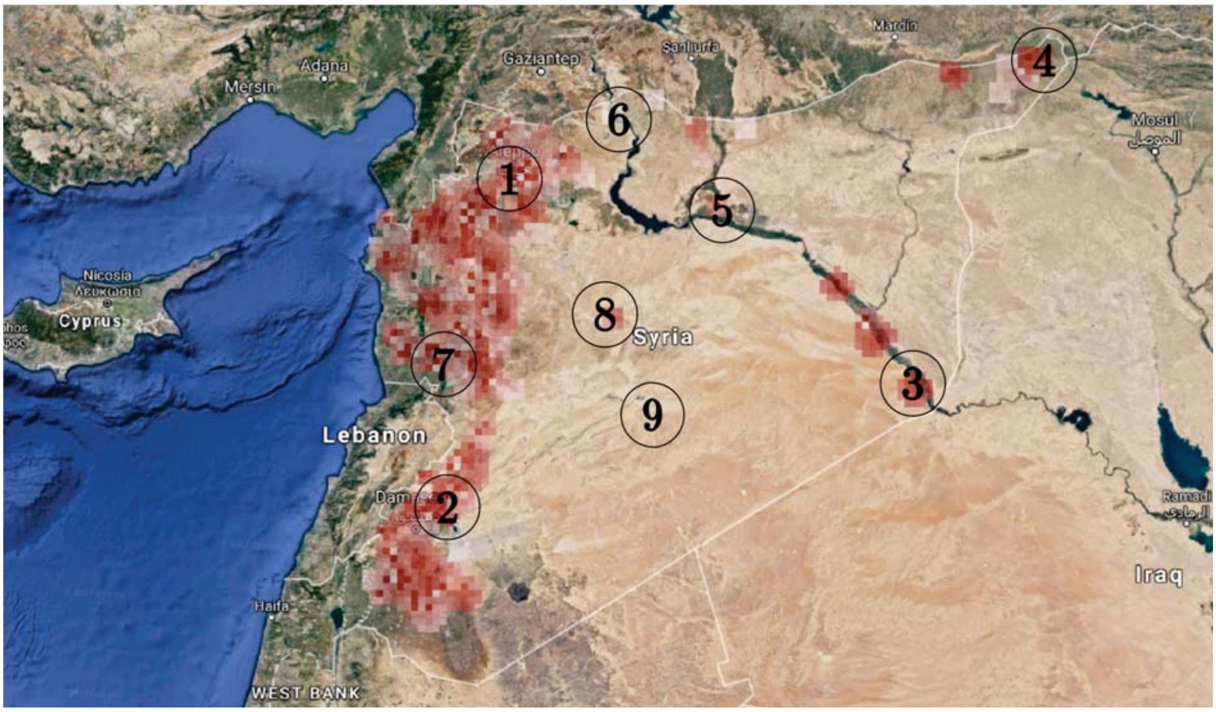
Points (1) and (2) indicate high conflict intensity on the visual data as recorded by GDELT and *Palantir Technologies*, as well as the simulated output. Furthermore, point (3) indicates conflict along the Euphrates river in both the simulated and recorded graphical representations of the conflict. The model further records conflict with a relatively high intensity at point (4), whereas both simulated visualisations, although agreeing with the presence of conflict at this point, present it as having less intensity. Although the simulated model considers the conflict at and immediately surrounding point (4), it does not notably consider the spread of this conflict



(a) *The spread of conflict as recorded in June 2014 by GDELT.*



(b) *The spread of conflict as recorded in June 2014 by Palantir Technologies.*



(c) *The spread of conflict as simulated in June 2014.*

Figure 7.1: *A comparison of the spread of conflict simulated and as recorded in June 2014 by GDELT and Palantir Technologies [124, 125].*

inland in the same manner that the records of Figures 7.1(a) and 7.1(b) do.

Along the Euphrates river, at points (5) and (6), correlation between the simulated and recorded outputs is evident, although the simulated version of conflict at point (6) is less intense and less far-reaching than that of the true data. The presence of conflict between points (1) and (7), spreading along the Syrian border and further inland, is similar in both the simulated output and the visualised data. Point (8) indicates the presence of conflict in central Syria with little to no spread. This is seen in both the visual data of GDELT and the simulated output, although *Palantir Technologies* did not account for this conflict in its record. Finally, point (9) shows the presence of conflict as recorded in Figures 7.1(a) and 7.1(b), although this area of conflict is not present in the simulated model output. In general, the spread of conflict at this time instance as recorded correlates well with the simulated output, especially alongside the Western border of the country at points (1), (2) and (7), along the Euphrates river between points (3) and (5), as well as along the Syrian border at point (4).

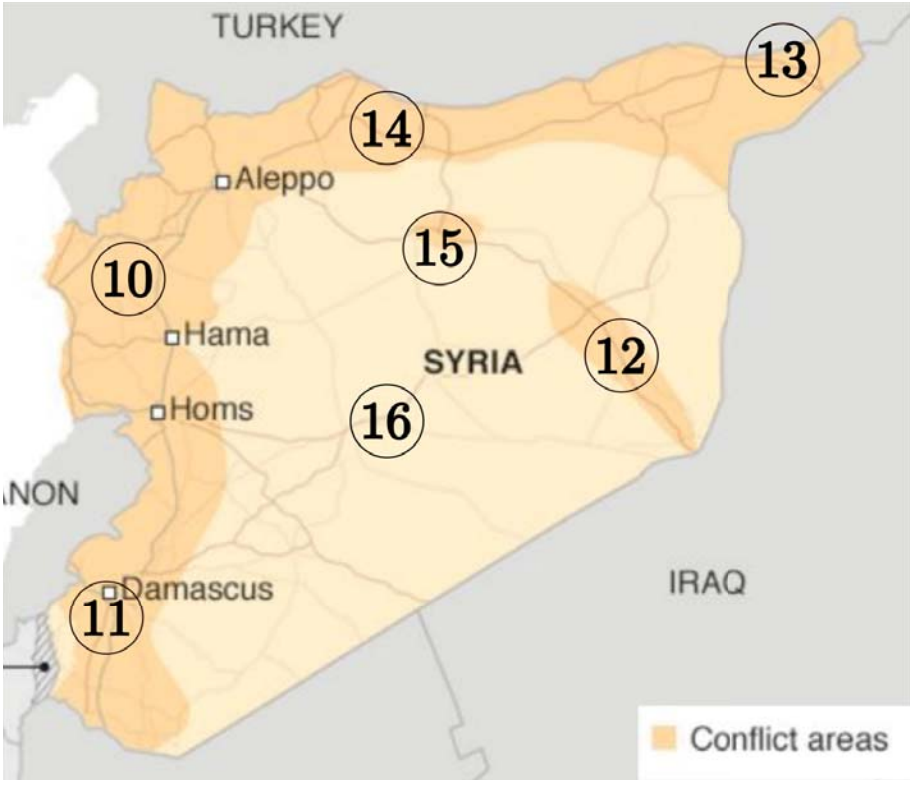
The conflict as simulated in July 2015 was then compared to the conflict as recorded for the same time period by the UNHCR in Figure 7.2. The visual record of the UNHCR shown in Figure 7.2(a) does not, however, indicate the intensity of conflict, but only its presence in certain regions.

Similarly to the previous time instance analysed, the majority of conflict lies along the Northern and North-Western borders of Syria, stretching between points (11), (10) and (14). The shape of the spread conflict is quite similar when comparing the recorded data with the simulated output. The wide-spread conflict stretching from point (14) towards point (13) is shown in the recorded data from the UNHCR, while the simulated output indicates a more sporadic spread in this region. This could, however, also be in light of the fact that the graphical data recorded by the UNHCR is notably aggregated, as well as excludes a consideration of the intensity of the conflict. A correlation in the conflict shown along the Euphrates river at point (12) is evident between both the simulated and recorded visualisations. Similarly, point (15) in Figure 7.2(a) indicates the presence of conflict at a specific location with little spread, which is accurately accounted for by the simulated output. At point (16), however, the simulated output, as shown in Figure 7.2(b), indicates the presence of conflict, where this is not the case in the recorded conflict from the UNHCR.

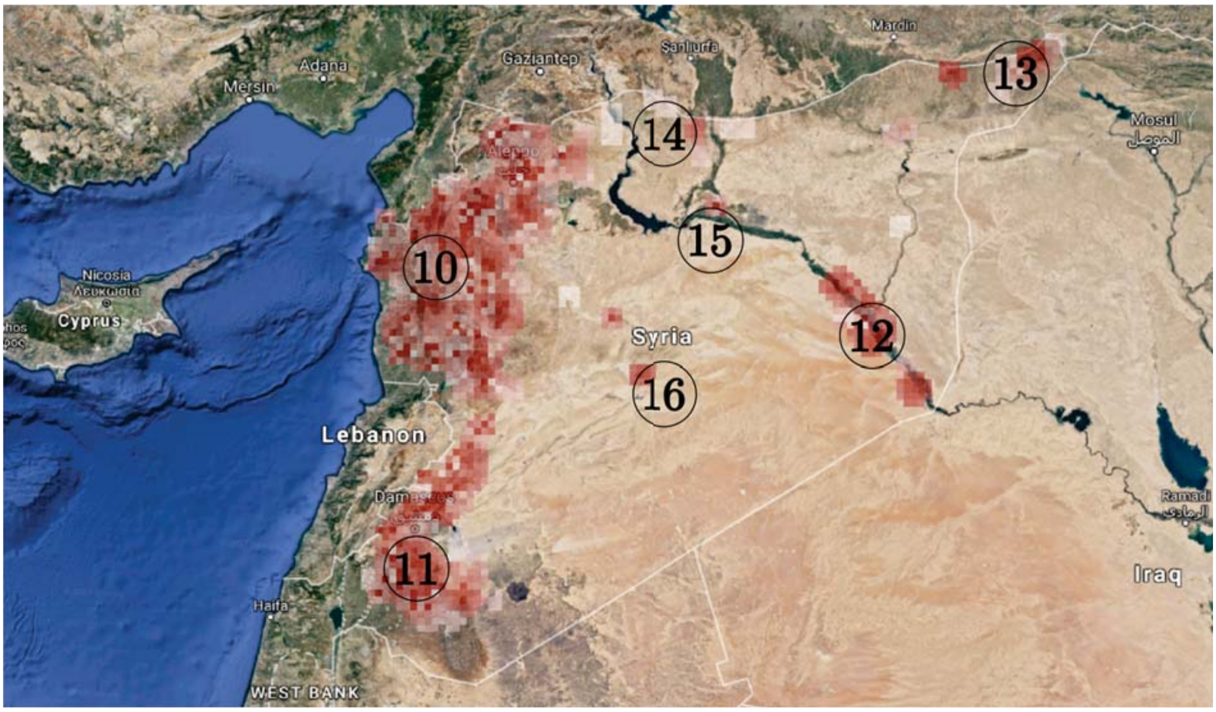
Finally, Figure 7.3 compares the conflict as simulated in December 2016 to the conflict recorded by *The Carter Center* at the same time instance. The illustration of *The Carter Center* shown in Figure 7.3(a) does not include the intensity of conflict, although it indicates the different parties involved in the conflict by means of colour. Red refers to the government, green refers to the opposition, black refers to ISIS, yellow refers to YPG and allies and blue refers to ceasefires.

When considering the conflict as recorded in Figure 7.3(a), it stretches from point (19), past points (17) and (18), towards point (23). The simulated output shows a similar spread of conflict between points (19) and (18), although it does not indicate such a wide spread of conflict between points (18) and (23). *The Carter Center* data furthermore indicates conflict present along the Euphrates river at point (21), as well as further North towards the Turkish border at points (22) and (20). The simulated output of the model accounts for the conflict along the Euphrates river at point (21), as well as conflict in the immediate vicinity of both points (20) and (22), although it does not necessarily indicate a similar spread of conflict between the two points. A further correlation between the simulated output and the visual data on record is the presence of conflict in central Syria at point (24). The simulated output mimics the scarcely dispersed conflict shown by the visual data recorded in Figure 7.3(a) in the triangular area mapped out by points (17), (23) and (24).

In general, the visualised data which exists on record for the three instances considered indicated high density of conflict occurring along the Northern and North-Western borders of Syria, in the vicinity of Aleppo and Damascus. Conflict is also present at the North-Eastern corner of the country, where it borders Turkey and Iraq. Furthermore, notable conflict exists alongside the Euphrates river from where it enters Syria at the Iraqi border. The simulated conflict indicates a reasonable replication of the actual spread of conflict as described, as well as visually correlates acceptably. This indicates that the manner in which conflict has been modelled in the agent-based model developed, bears meaningful similarities to the manner in which it exist, develops, festers and depletes in reality.

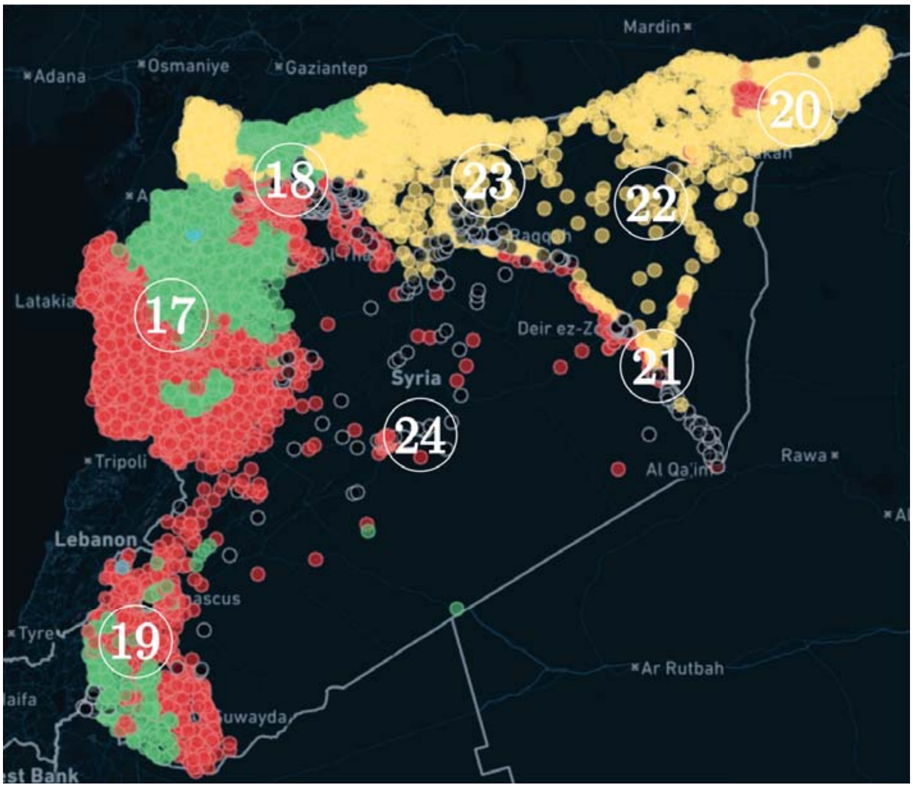


(a) *The spread of conflict as recorded in July 2015 by the UNHCR.*

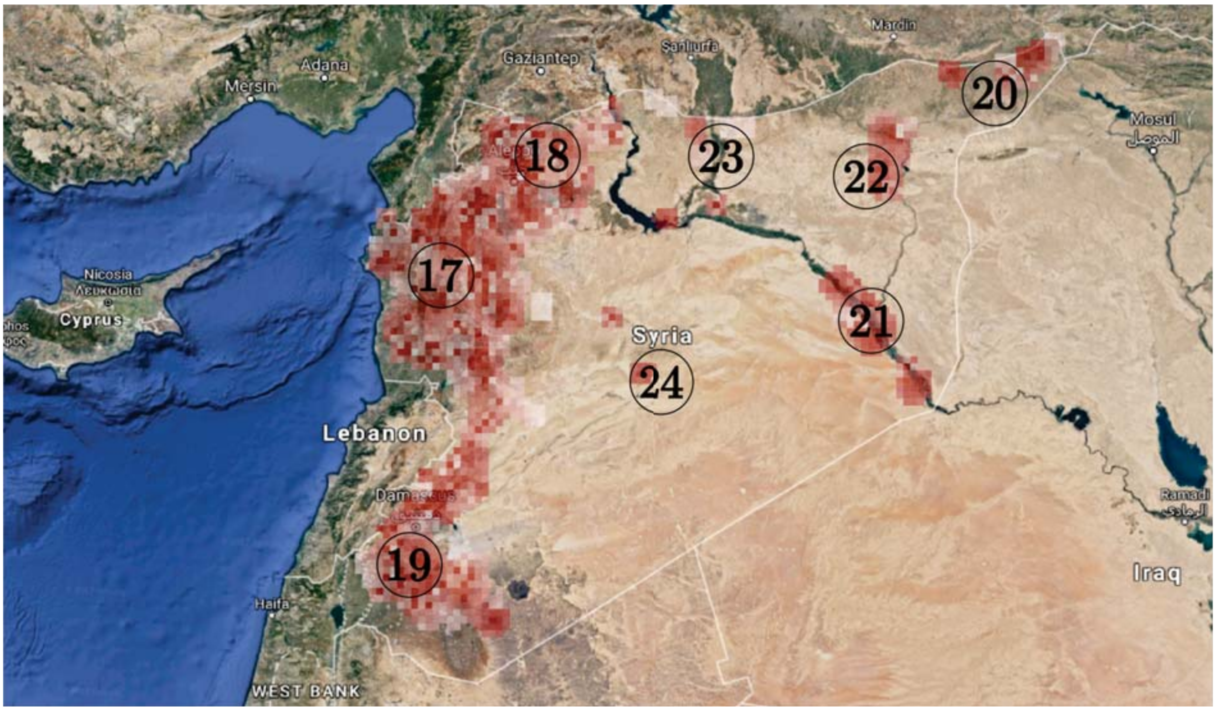


(b) *The spread of conflict as simulated in July 2015.*

FIGURE 7.2: *A comparison of the spread of conflict simulated and as recorded in July 2015 by the UNHCR [129].*



(a) *The spread of conflict as recorded in December 2016 by The Carter Center.*



(b) *The spread of conflict as simulated in December 2016.*

FIGURE 7.3: *A comparison of the spread of conflict simulated and as recorded in December 2016 by The Carter Center [123].*

7.4 Face validation

Face validation of a simulation model encompasses having subject matter experts review the model in order to test how reasonably the model replicates reality. A simulation model gains credibility when it is perceived by a subject matter expert as an appropriate representation of the intended real-world system. Law [72] proposed the following questions, amongst others, to be addressed during face validation:

(i) Do you agree with the assumptions being made in the model?

(ii) Are the processes employed in the model to recreate appropriate scenarios verified?

(iii) Is the animation output of the simulation model compelling and does it represent, to the best of your knowledge and expertise, a similar situation to that which is experienced in the real-life scenario which the model is attempting to recreate?

Subject matter experts, such as Aksel [4] from Koç University, Frydenlund [43] from the Old Dominion University, Groen [50] from Brunell University, Lemos [75] from the University of Agder, Shomary [110] from Stockholm University, Smith [115] from the University of Sussex and Stewart [119] from the University of Cape Town, were contacted throughout the progression of this study for inputs and expert recommendations in terms of the implementation of certain aspects of the model. As a final validation, the expertise of Lemos [75] was utilised to validated the modelling of conflict, while the expertise of Aksel [4] and Frydenlund [43] were utilised in validating the simulated Syrian people and their associated decision-making.

Lemos is a postdoctoral research fellow at the University of Agder in Norway, whose academic interest includes the social simulation of conflict using agent-based modelling He also recently published a book entitled "Agent-based modelling of social conflict" [76]on this topic. Frydenlund, a research assistant professor at the *Virginia Modeling, Analysis & Simulation Center* at the Old Dominion University in the United States of America, specialises in the modelling of migration, mobility and political dynamics, and is currently working on developing insightful simulation models focussing on protracted refugee situations. Frydenlund's insight and expertise were sought throughout the development of the simulation model described in this thesis. Aksel is a postdoctoral researcher at Koç University in Turkey and coordinator of the *Migration Research Center at Koç University*, with particular experience in the field of international migration and migration aspects related to Turkey. Aksel provided insight during the modelling of people and their decision-making, as referred to in Chapter 5.

Stewart is a professor at the University of Cape Town in South Africa and is an expert in the field of MCDM, having published numerous books on the topic. Groen is a lecturer of simulation and modelling at the Brunel University in England, with a research focus primarily on multi-scale modelling and optimisation. He only recently began considering the modelling of refugees by means of computer simulation modelling. Smith is a research fellow in geography at the University of Sussex in England, working on research related to migrants and, in particular, climate refugees. Shomary, a doctoral student in Early Childhood Education at Stockholm University in Sweden, is of Iraqi and Syrian descent and performs research pertaining to Syrian refugees in Sweden. Shomary was born in Iraq and raised in Syria before moving to Sweden as a political refugee nearly 25 years ago. In light of this, she has a first-hand understanding of the manner in which Syrians make decisions and the role which their personal characteristics play in this process.

A number of important components of the model were discussed with the subject matter experts. The face validation process was conducted at various stages of the model development to evaluate the model each of the assumptions made and to ascertain whether or not the model replicates reality. The different aspects of the model and the responses from the different subject matter experts are collected and summaries below.

Modelling assumptions

In order to surmount the challenges pertaining to the modelling of forced migration and the decision-making of people, various assumptions had to be made with regards to the simulation model, as described in §5.3. These assumptions are related to geography, time, data and agent attributes within the simulation model. With respect to geography, Groen advised to only consider neighbouring countries of Syria for the sake of computational expense. The model, however, does not consider movement between these neighbouring countries, or possible movement from these countries into Europe or back to Syria. Aksel explained the concept of transit countries, where people may plan to migrate to a certain country, but travel via another country to get to their final destination. Another factor suggested by Shomary to take into account is the repatriation of Syrians, where Syrians who relocated to another country choose to return to their place of origin. According to Shomary, this phenomenon is evident, even in some European countries. For the purpose of limiting model complexities, however, it was agreed that disregarding both the use of transit countries and the aspect of repatriation in the simulation model developed was a sensible assumption.

Concerning the influence of time, the only suggestion which came about was that of Aksel who suggested accounting for the maturity of conflict. This is understood to effect the manner in which a person reaction towards the conflict.

The data accounted for in this model pertains not only to information typically recorded by available data sources, but further considers the categorisation of movement types. Both Aksel and Groen advised that, aside from refugees and asylum-seekers, IDPs and undocumented migrants also need to be taken into account — the details for which there presently exist no adequate available data. In light of this, the model developed in this study has the potential to contribute to the research in this regard by providing the possibility of generate inference data for further investigation.

Aksel further recommended referring to studies completed on Afghan forced migrants in an attempt to better understand how the people react towards conflict situations. In terms of data with regards to conflict, a suggestion was made by Lemos to utilise the Social Conflict Analysis Database but, by this stage in the model's development, data from GDELT had already been utilised. Both data sources were discussed and it was found that GDELT is appropriate for the application at hand.

The final model assumptions are related to agent attributes. Initially it was proposed that each agent should represent a family or a neighbourhood. Aksel, however, suggested grouping the people based on age and gender instead, since males do not necessarily have the same migration patterns as females of the same age group. Lemos affirmed this decision and recommended only to account for the most essential agent characteristics to decrease model complexity, as well as minimise the number of required assumptions in modelling the Syrian population.

The initiation, spread and depletion of conflict

The modelling of conflict comprises the initial occurrence of a conflict-related incident, the spread of this conflict and its eventual depletion. The manner in which this is captured in the simulation model is described in §5.4. After reviewing this implementation of modelled conflict, Lemos affirmed it to be a sensible and acceptable implementation. Furthermore, he commented on the systematic approach utilised in the construction of the model, with regards to the conflict modelling, noting that it encompasses all fundamental issues without adding unnecessary complexity. Although the model does not account for the different parties involved in the conflict, the cellular automata approach in modelling the conflict was agreed to be more effective for the purpose of this model. Lemos and Frydenlund suggested that for an extension of the model with application in an environment other than Syria may be to include modelling conflict as an agents. Smith commented on the correlation between the conflict as simulated and the recorded visual data available, affirming the manner in which the simulated conflict replicates reality.

The agent population

The manner in which people are modelled in the simulation model presented in this study is described in §5.5. Attributes are assigned to an agent, based on research performed on the population in terms of the distribution of characteristics such as age, gender, tertiary education and the like. Smith agreed with the manner in which these deductions were made, but was concerned with the large number of people represented by a single agent, although Frydenlund agreed that, owing to restrictions placed on computational capability, the selected level of abstraction of the model is reasonable. It is, however, suggested by Frydenlund to increase the granularity of the model in future improvements with the use of supercomputers. Aksel and Frydenlund also recommended for future work to include an agent's social network, the inclusion of which could be captured with the use of a lower level of abstraction in the model.

The development of the agent population is modelled taking by into account births and deaths and the geographic dispersion of the agents across Syria, according to data from a census, in an attempt to replicate the population density within each governorate. Furthermore, each agent possesses a certain ability to withstand conflict in the form of a moving threshold which is determined based on literature, as well as the opinions of experts. Aksel stated that men, particularly those between the ages of 15 and 64, typically choose to leave their place of residence before the women and children, in order to find and prepare a place to which the rest of the family may relocate. It was also mentioned by Shomary that people with tertiary education are more inclined to relocate with the ambition of finding new job opportunities. Smith commented on the manner in which expert opinions are used to generate a general set of rules governing a person's propensity to move in the presence of conflict, confirming this as a suitable means of achieving reasonable estimations. Furthermore, Lemos agreed with the manner in which the interaction between agents and their environment, particularly with reference to the effects of conflict, was modelled.

The decision-making process of agents

The decision-making process of an agent in the simulation model, as described in §5.6, comprises an agent being required to take on a movement type and then selecting a suitable proposed destination. Initially, a consideration was made to employ MCDM methods in modelling these decisions, however, Stewart explained the implications of utilising prescriptive decision-making methods in an environment which requires descriptive methods, as explained in §4.1 and §4.4. It was agreed upon that employing a simplified modelling approach would more accurately reflect the reality of the decision-making process.

An agent's characteristics govern the movement type that is chooses. Lemos suggested utilising exploratory factor analysis in identifying underlying relationships between variables, although the time restrictions in this study did not allow for this. Shomary provided insights into the correlation between a person's attributes and their proposed movement type. A phenomenon commonly witnessed is that people pay smugglers for access across the Syrian border, travelling towards Europe. These individuals usually fall into a high income bracket, as this manner of travelling is expensive and may easily cost up to \$10 000 for a family. It is therefore reasonable to deduce that people of medium to low income would rather relocate as an IDP or a refugee in neighbouring countries.

Interestingly, some people of high income classes choose to seek asylum in European countries, only to leave that country once nationality is gained in order to relocate to countries along the gulf, such as the United Arab Emirates. Here they attempt to start businesses. These countries are chosen as they allow for living conditions and language which are similar to what the individuals were accustomed to in Syria. Furthermore, Shomary discussed the reluctance people who are of an advanced age, of a low income bracket or have without tertiary education have towards relocating.

The decision-making of an individual with respect to the selection of a proposed destination is modelled differently for IDPs than for refugees or undocumented migrants. An IDP considers destinations with little to no conflict, which have high population densities and are close to their place of origin. Aksel, Frydenlund, Lemos, Shomary and Smith explicitly affirmed the manner in which this is modelled. Refugees and undocumented migrants take into account the distance to the neighbouring countries, the popularity of these countries (*i.e.* the number of Syrians who typically choose to move there), as well as the 'openness' of these countries. Aksel commented that the use of popularity as a factor allows for an indirect consideration of social networks and affirms a simple the incorporation thereof. Lemos agreed with the implementation of openness scores, but suggested considering a country to have a different openness scores towards refugees than towards undocumented migrants. In Lesvos, for example, the attitude towards refugees differs from the attitude towards undocumented migrants, as the locals will have to compete for work opportunities against the latter group, while refugees only remain temporarily. Furthermore, Frydenlund and Aksel suggested further improvement of the model which may consider specific regions within countries as proposed destinations for relocation as opposed to the present implementation which considers countries in their entirety.

The discussions held with the various subject matter experts allowed for the incremental development the simulation model based on insight and expert knowledge, in conjunction with the literature, which was necessary in light of the lack of complete data. The subject matter experts affirmed the need for a model such as the one developed within this thesis and, to the best of the knowledge of this expert panel, no such model exists which considers such a vast collection of factors and implications pertaining to refugee modelling in the presence of conflict. Frydenlund further commented that this model allows for a good attempt at true estimations with regards to the number of forcibly displaced people per movement type and per country of destination. In combination, the insights, comments and affirmation received from the subject matter experts during the aforementioned discussion contributes largely to affirming the validity fo the constructed simulation model in light of the numerous assumptions and simplifications required for its construction.

7.5 Parameter variation and scenario replication

In light of the sporadic and largely incomplete data pertaining to the movement of forcibly displaced persons, a parameter variation analysis is performed in an attempt to fit the model to existing partial data and, by implication, illustrate the model's capability to be used to simulate specific scenarios. The model is not intended to be an absoloute, accurate representation of forced migration in Syria, but, rather, to serve as a generic tool which has the ability, when equipped with the correct input parameter values, to model any given scenario. In an attempt to illustrate this capability through a parameter variation analysis, two existing scenarios for which some data exist were chosen in an attempt to utlilise the model to replicate the outputs as recorded in the datasets.

The first set of experiments aims to replicate the ratio between IDPs, refugees and undocumented migrants who reside in or originate from Syria. In light of the fact that no definitive data are available on the number of undocumented migrants, a directed analysis is applied to the relationship between IDPs and refugees exclusively, allowing inference data pertaining to undocumented migrants to subsequently be generated. It is reported by the UNHCR [129] that the number of refugees at the end of 2015 accumulated to 5.2 million, while the number of IDPs within Syria at the same time was 6.5 million. It can therefore be assumed that the IDP:refugee ratio should be approximately 11:9 when considering only these two movement types.

The function which determines the movement type of an agent, F `DetMovementType`, utilises a probability matrix that, as discussed in §5.6.1, associates the various agent characteristics with a probability of being allocated a certain movement type. The probability matrices for the experiments performed for this scenario were therefore varied iteratively in an attempt to find a set of probabilities which results in a ratio between the number of IDPs and refugees which aligns with the existing estimation. The probability matrices associated with each of the experiments performed are given in Appendix B and a the results, with respect to the ratio of the number of IDPs and the number of refugees per experiment, are listed in Table 7.2.

Table 7.2: *The model output with respect to the various movement types.*

Experiment	Movement Type 1 Ratio	Movement Type 2 Ratio
1	0.7547	0.2452
2	0.7720	0.2280
3	0.7305	0.2695
4	0.7399	0.2601
5	0.7076	0.2924
6	0.6956	0.3044
7	0.6244	0.3756

In experiment 7, the best alternative probability matrix was determined which allows for a ratio which best matches the real-world approximation. The simulation model, as configured in experiment 7, modelled the number of undocumented migrants at the end of 2015 to account for 2.49 million people. Furthermore, the model indicated a number of 460 000 deaths resulting from people being trapped in conflict-effected areas. This correlates to the literature which approximates that, by the end of 2015, more than 400 000 Syrians had died in the civil war [64, 87].

The second series of experiments conducted is concerned with the weighted criteria and initial openness scores of neighbouring countries influencing an agent's decision with respect to selecting a proposed destination. An existing dataset, as reported by the UNHCR [129], depicting the number of refugees within each neighbouring country at the end of 2015, is shown in Table 7.3. Owing to the sporadic estimations of data with respect to undocumented migrants and the complexities associated with the consideration of specific destinations for the IDPs within Syria, the partial data regarding refugees in neighbouring countries are utilised in the form of a parameter variation experiment in an attempt to generate inference data pertaining to the other two aforementioned types of displaced persons.

Table 7.3: *The data recorded by the UNHCR [129] of the number of Syrian refugees within each of the neighbouring countries at the end of 2015.*

	Turkey	Lebanon	Jordan	Greece	Iraq
Number of people (in millions)	2.2	1.1	0.6	1.2	0.1
Ratio (as a percentage)	42.1	21.0	12.0	22.9	1.9

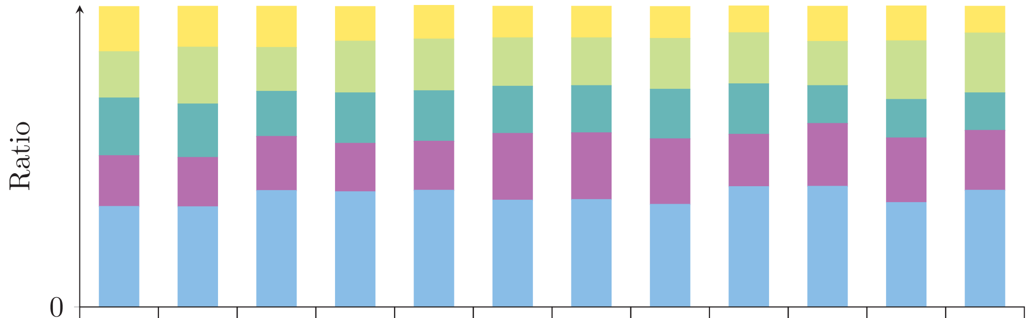
Simulation experiments were performed in an attempt to fit the simulated output to the existing data. A fixed seed was used to eliminate randomness and the experiments were set to pause at the end of December 2015 in order for comparisons to be made to the dataset, which is recorded as of the end of 2015. The parameters varied within these experiments included the weights of the criteria which influence refugees and undocumented migrants when selecting a proposed destination country, as well as the neighbouring countries' initial openness score towards these individuals. A list of the experiments performed and their associated set parameters are given in Table 7.4. The simulated output for these experiments with regards to the number of refugees per neighbouring country of Syria are noted in Table 7.5. Furthermore, Figure 7.4 illustrates the ratio of refugees and undocumented migrants as per neighbouring country, per experiment.

Table 7.4: *The list of experiments in the directed sensitivity analysis when considering choosing a proposed destination among neighbouring countries.*

Experiment	1	2	3	4	5	6	7	8	9	10	11	12
The weighted criteria												
Distance	0.55	0.55	0.55	0.45	0.40	0.40	0.45	0.45	0.45	0.45	0.45	0.45
Openness Score	0.35	0.35	0.35	0.45	0.40	0.40	0.45	0.45	0.45	0.45	0.45	0.45
Popularity	0.10	0.10	0.10	0.10	0.20	0.20	0.10	0.10	0.10	0.10	0.10	0.10
Initial openness scores of neighbouring countries												
Turkey	80	80	80	80	80	80	80	85	90	90	90	95
Lebanon	70	70	70	70	70	70	70	70	70	70	75	75
Jordan	75	60	45	45	45	40	40	40	35	30	30	30
Greece	50	50	50	50	50	50	50	55	60	60	65	65
Iraq	40	20	15	15	10	5	5	5	5	5	5	3

Table 7.5: *The model output (in millions) of the experiments performed considering choosing a proposed destination among neighbouring countries.*

Experiment	1	2	3	4	5	6	7	8	9	10	11	12
Turkey	1.6	1.69	1.70	1.76	1.76	1.65	1.66	1.61	2.03	1.88	1.71	1.74
Lebanon	0.81	0.83	0.79	0.74	0.75	1.03	1.03	1.03	0.88	0.98	1.06	0.89
Jordan	0.92	0.90	0.66	0.77	0.77	0.73	0.73	0.78	0.85	0.59	0.63	0.56
Greece	0.74	0.96	0.64	0.79	0.79	0.75	0.74	0.80	0.86	0.69	0.96	0.89
Iraq	0.72	0.69	0.60	0.53	0.52	0.49	0.49	0.50	0.45	0.55	0.57	0.40



	33	34	35	36	37	38	39	40	41	42	43	44
Turkey	33.4	33.3	38.7	38.3	38.8	35.5	35.7	34.1	40.0	40.1	34.7	38.8
Lebanon	16.9	16.4	18.0	16.1	16.3	22.2	22.2	21.8	17.4	20.9	21.5	19.9
Jordan	19.2	17.8	15.0	16.8	16.8	15.7	15.7	16.5	16.8	12.6	12.8	12.5
Greece	15.4	18.9	14.6	17.2	17.2	16.1	15.9	16.9	17.0	14.7	19.5	19.9
Iraq	15.0	13.6	13.7	11.5	11.3	10.5	10.5	10.6	8.9	11.7	11.6	8.9

Ratio of refugees and undocumented migrants per experiment

FIGURE 7.4: *The model output (ratios) of the experiments performed considering choosing a proposed destination among neighbouring countries.*

When comparing the experiment outputs to the ratios recorded by the UNHCR [129], as shown in Table 7.3, a noticeable difference is seen in the ratios associated with Jordan and Iraq. Each experiment was adjusted based on the results from previous experiments in an attempt to best replicate the existing data. Interestingly, the effect of adjusting the weight of criteria did not have as significant an effect as changing the initial openness scores, indicating that the latter mentioned parameters are more sensitive to change. The openness scores were therefore adjusted for each iterative experiment until the output was deemed sensible in its correlation to the existing data. Experiment 12 is regarded as the best-fit with respect to the ratio of people relocating to these neighbouring countries as refugees or undocumented migrants. A graphical output for experiment 12 with regards to the number of forcibly displaced people per movement type and per destination is given in Appendix C.

By implementing an iterative approach to generate model outputs which correlate to recorded data, the model may be effectively calibrated for any given scenario. The model user could then decide to either implement these parameters in an attempt to infer future outcomes with regards to future forced migration under similar conditions, or adjust the parameters slightly in order to compare an expected simulated output to an actual simulated output, allowing for further insights to be gained through the use of so-called 'what-if' analyses.

7.6 A decision-support and analysis tool

A decision-support tool facilitates the selection and variation of parameters in a simulation model so as to allow for scenario analysis [134]. As illustrated in the previous section, the agent-based model developed in this thesis possesses the capability to be implemented as an investigative tool in an attempt to analyse various scenarios which may occur in the field of forced migration. The analysis of these different scenarios will then accommodate a better understanding of the behaviour and decision-making exhibited by people when confronted with conflict. The agent-based model developed in this thesis is equipped with a GUI, as detailed in §5.7, for the purpose of utilising the model as a decision support and analysis tool. The inclusion of this feature increases the usability and flexibility of the agent-based model.

The GUI includes a configuration screen designed to prompt the user for various selections and parameter input values, such as the criteria weights regarding refugees and undocumented migrants, the criteria weights regarding IDPs, as well as the initial openness scores of neighbouring countries. A brief explanation of each of the available choices and parameters is given below, as well as the default configuration of the model as presented to the user.

Choose the conflict input method

Input: Manual or Data.

Explanation: Allows for the initiation of conflict to occur as either induced manually by the user, or with the use of data input.

Default value: Manual.

Choose the weight of these criteria when considering refugees and undocumented migrants

Input: Distance, Openness Score and Popularity.

Explanation: Sets the weight of the corresponding criteria which influence agents of movement types 2 or 3 in their decision when choosing a neighbouring country as a proposed destination. The sum of these weights should be equal to 1.

Default value: Distance = 0.55, Openness Score = 0.35 and Popularity = 0.1.

Effect of increase: An agent would more strongly be influenced by the increased factor when having to choose a proposed destination.

Effect of decrease: An agent would be consider less significantly the decreased factor when having to choose a proposed destination.

Choose the weight of these criteria when considering internally displaced people

Input: Conflict and Population Density.

Explanation: Sets the weight of the corresponding criteria which influences an agent of movement type 1 in its decision when choosing a location within Syria as a proposed destination. The sum of these weights should be equal to 1.

Default value: Conflict = 0.6 and Population Density = 0.4.

Effect of increase: An agent would be more considerate towards the increased factor when having to choose a proposed destination.

Effect of decrease: An agent would be less considerate towards the decreased factor when having to choose a proposed destination.

Choose the initial openness score associated with each of these countries

Input: Turkey, Lebanon, Jordan, Greece and Iraq.

Explanation: Sets the initial willingness of the associated country to allow for the intake of refugees and undocumented migrants.

Default value: Turkey = 95, Lebanon = 75, Jordan = 30, Greece = 65 and Iraq = 3.

Effect of increase: An agent would be more prone to consider relocating to the associated country when having to choose a neighbouring country as proposed destination.

Effect of decrease: An agent would be less prone to consider relocating to the associated country when having to choose a neighbouring country as proposed destination.

While the configuration screen allows for input only at the initiation of a simulation experiment, the primary run-time screen allows for parameter variation during the course of the simulation run. The parameters which may be varied pertain exclusively to the modelling of conflict.

Probability of Conflict Spreading

Explanation: Set the probability that conflict will spread from conflict-infected areas to their immediate surrounding areas.

Default value: Triangular(0, 1, 0.75).

Effect of increase: The conflict would spread more easily from its source outwards.

Effect of decrease: The conflict would not spread as easily from its source outwards.

Probability of Conflict Depletion

Explanation: Set the probability that conflict will deplete from the edges of conflict-infected areas.

Default value: Triangular(0, 1, 0.75).

Effect of increase: The conflict would deplete more easily from the outskirts of the area where conflict is initiated.

Effect of decrease: The conflict would not deplete as easily from the outskirts of the area where conflict is initiated.

Intensity of Conflict

Explanation: Set the intensity at which conflict will be initiated when manually initiating conflict.

Default value: Triangular(20, 100, 75).

Effect of increase: The conflict initiated would have a higher level of intensity.

Effect of decrease: The conflict initiated would have a lower level of intensity.

7.7 Chapter summary

This chapter provided an introduction to the concepts of validation and output data analysis. Calibration of parameters constituting to the modelling of conflict was performed in § 7.2, followed by the validation of conflict modelled in §7.3 which considered the manner in which it spreads within the simulated area. The face validation as performed in consultation with international leaders in the field of refugee and conflict modelling was discussed in § 7.4. This was followed by a parameter variation analysis in §7.5, which illustrated the flexibility of the model, as well as its ability to simulate any desired situation as required by a model user. The implementation of the model as a decision-support tool and analysis was then discussed in §7.6. Each of these steps were discussed in detail within this chapter in an attempt to validate the model and increase the credibility of its outputs, in light of the restrictions which exist when attempting to validate the model according to traditional methods.

CHAPTER 8

Conclusion

Contents

This chapter includes a summary of the contents of this thesis where each of the project objectives is considered, followed by an overview of contributions made in §8.2. A number of suggestions is made with regards to possible future work stemming from this study.

8.1 Thesis summary

The introductory chapter of this thesis, Chapter 1, provided a brief description of global forced migration, the potential impact of humanitarian crises and the application of simulation models within this context. The problem description was given and, following this, the scope and objectives of the study were detailed, along with the proposed research methodology. The chapter then closed with a brief outline of the thesis organisation.

The aim of Chapter 2 was to introduce the reader to the concept of forced migration, in partial fulfilment of Thesis Objective I(a). This introduction discussed various causes of forced displacement and expanded on the typology of forced migration, explaining the various movement types of forcibly displaced people, in accordance with Thesis Objective I(b). This was followed by a brief overview of forced displacement throughout history, further fulfilling Thesis Objective I(a). Furthermore, the forced migration currently evident on a global scale was discussed, in fulfilment of Thesis Objective I(c). The chapter closed with a focus on the current forced migration taking place within Syria, thereby addressing Thesis Objective I(d).

Chapter 3 reviewed the different aspects related to computer simulation modelling, in partial fulfilment of Thesis Objective I(e). This discussion included the various simulation modelling types, levels of abstraction, various simulation modelling approaches, generic steps followed in completing a simulation study, as well as validation and verification procedures pertaining to simulation modelling. In further fulfilment of Thesis Objective I(e), §3.2 discussed agent-based simulation modelling, referring to the various advantages and disadvantages of this simulation method and the various components which compose such a model. Furthermore, specific examples of the application of ABM with regards to the modelling of forced migration were discussed in §3.3.

In Chapter 4, an introduction to the field of decision-making was given, in accordance with Thesis Objective I(f), where after the study of multi-criteria decision-making methods was elaborated upon, in fulfilment of Thesis Objective I(g). The classification of these methods was mentioned, along with the available multi-criterion aggregation procedures. In §4.3, the modelling of human decision-making was investigated, in fulfilment of Thesis Objective I(h). Various methods previously employed to model decision-making of humans were discussed and, following this, the modelling of decision-making pertaining to forcibly displaced people, in particular, was discussed in §4.4.

Chapter 5 provided a background to the model proposed in this thesis, as well as an overview of the AnyLogic Simulation Software Suite, describing its suitability for the development of this model. A number of limitations and assumptions made in this project were mentioned, after which a detailed description of the agent-based model developed, including its various elements was given. Thesis Objectives II(a), II(b) and II(c) were achieved in §5.4, with a detailed description of the manner in which conflict was modelled, including the initiation, spread and depletion of conflict. The manner in which the aggregated population of agents is modelled was then explained in §5.5.2, in fulfilment of Thesis Objective II(d), and a discussion on determining an agent's ability to withstand conflict was given in §5.5.4, in fulfilment of Thesis Objective II(e). Section 5.6 was dedicated to the discussion of the manner in which human decision-making was modelled, in accordance with Thesis Objective II(f), with the graphical user interface in which the model is embedded discussed in §5.7. This serves as the framework to assist model users in employing the model in their own prescribed scenarios, thereby partially fulfilling Thesis Objective III.

The verification of the agent-based model developed in this thesis was performed in Chapter 6 in partial fulfilment of Thesis Objectives III and IV. This included a short introduction to the concept of verification given in §6.1, followed by an explanation on the manner in which the conflict was modelled in this simulation model given in §6.2. The manner in which the people were modelled as a population of agents was verified in §6.3 by assessing certain cases. The chapter closed with the verification of the modelled decision-making of people in §6.4.

A structured model analysis was performed in Chapter 7, which included model validation, as well as parameter calibration and variation, in further fulfilment of Thesis Objectives III and IV. Section 7.1 provided an explanation of the concepts of model validation and output data analysis. Following this, the parameters pertaining to the modelling of conflict were calibrated in §7.2 and the conflict were validated in §7.3 by means of employing various experiments. The face validation was performed in §7.4, including the research opinions of subject matter experts. A parameter establishment analysis was performed in §7.5, where after the ability of implementing the model as a decision support tool was discussed in §7.6.

Thesis Objective V is achieved within this chapter where future improvements, as well as sensible follow-up work which may stem from this study, are discussed.

8.2 The contributions of this study

The primary contribution of this thesis is towards understanding the behaviour and actions of people when confronted with conflict and in light of this, developing a means of facilitating the modelling of the decision-making and movement of IDPs, refugees and undocumented migrants. The individual contributions of this project are presented in this section.

Contribution 8.1 *The formal, comprehensive coalition of research opinions and insights, in collaboration with subject matter experts, pertaining to the modelling of conflict-induced forced migration.*

In this thesis, significant effort was made towards formulating a comprehensive literature study pertaining to forced migration, computer simulation modelling and the field decision-making. Furthermore, collaboration with various subject matter experts such as Aksel [4] from Koç University, Frydenlund [43] from the Old Dominion University, Groen [50] from Brunell University, Lemos [75] from the University of Agder, Shomary [110] from Stockholm University, Smith [115] from the University of Sussex and Stewart [119] from the University of Cape Town, allowed for significant insight and knowledge to be gained in aspects pertaining to both social and engineering sciences relevant in this field which do not necessarily appear presently in literature. This facilitated a collated understanding of social concepts such as the human characteristics influencing decision-making of forcibly displace people, along with the practical understanding of ABM and the modelling of forced migration. The various opinions and insights were synthesised and formally documented to facilitate future development of such models to the most realistic degree of accuracy presently possible.

Contribution 8.2 *The design of a user-friendly, agent-based simulation model depicting the initiation, spread and depletion of conflict in Syria.*

An agent-based model which facilitates the modelling of conflict was developed in the AnyLogic Simulation Software Suite [122]. The initiation of conflict, as well as the spread and depletion of conflict were modelled according to the description given in §5.4. The graphical user interface accommodates the manual initiation of conflict, as well as the use of data input values, while taking into account the conflict intensity, probability of spread and probability of depletion. The model allows for a animated output visualising the state of conflict over an area, thereby making it accessible to a variety of potential users of varying degrees of technical ability.

Contribution 8.3 *A novel simulation framework encapsulating decision-making of agents pertaining to the classification of movement types and proposed destinations.*

The agent-based model mentioned above incorporates notable aspects in the decision-making process of forcibly displaced people. The model allows for the calculation of a person's ability to withstand conflict, based on personal characteristics, as described in §5.5.4, as well as the decision-making of a person in choosing a respective movement type, as explained in §4.4 and §5.6. The decision-making process modelled also accounts for the selection of a proposed destination based on the movement type of a person, its characteristics and other external factors. The graphical user interface accommodates user inputs with respect to the weights of the various decision-making criteria which influence the proposed destination selection process. The modelling of people and their decision-making is endorsed by research in social sciences, as well as the opinions and knowledge gathered from subject matter experts.

Contribution 8.4 *The design and implementation of a decision-support tool which provides information with regards to the influx of refugees and undocumented migrants into neighbouring countries.*

The graphical user interface developed in §5.7 may be utilised as a decision-support tool, as explained in §7.6. The interface accommodates user input with respect to the weights of decision-making criteria and the initial openness of neighbouring countries towards refugees and undocumented migrants. The decision-making pertaining to the selection of a proposed destination is described in detail in §5.6.2. The decision-support tool therefore facilitates a model output which provides graphical data pertaining to the number of refugees and undocumented migrants who flee to the respective neighbouring countries. This has the potential to assist governments and humanitarian support organisations in preparing for the influx of people requiring aid.

Contribution 8.5 *A novel contribution to the global research effort in this regard.*

During the development of the agent-based model presented in this thesis, regular contact was made with subject matter experts in order to gain insight and knowledge pertaining to the field of forced migration in order to accurately model this phenomenon. This knowledge, as documented, was implemented in modelling conflict and the decision-making of people in the presence of conflict by means of ABM. This study provided viable context to the state of computer simulation based research with regards to its application in the field of forced migration.

In light of the outcomes of this study, the author has been invited by Frydenlund [43] to partake as an expert member in a panel discussion on the theme of *Changing durable solutions* at the 17^{th} *International Association for the Study of Forced Migration* conference which will be hosted by the University of Macedonia in Thessaloniki, Greece in 2018. Furthermore, the *Emerging Scholars and Practitioners on Migration Issues Network* has invited the author to partake in a roundtable discussion occurring in 2018 at the *Canadian Association for Refugee and Forced Migration Studies* conference hosted in Ottawa, Canada. These invitations affirm the degree to which the author has contributed to the global state of research in this field, presenting exciting opportunities for further insights into improving or extending both the model developed in this thesis, as well as other existing models worldwide.

Contribution 8.6 *The recommendations of follow-up work which may stem from this study.*

The agent-based simulation model developed allows for various further improvements or extensions to be performed. Recommendations of specific future endeavours are made in the section to follow. The model was developed in a modular, structured fashion providing a good foundation with respect to the modelling of conflict and decision-making of individuals, for further improvements or extensions. These recommendations include novel areas in the communal research field of computer simulation and forced migration which may be investigated.

8.3 Possible future work

This study allows for various future work which involves improvements to the current model and extensions thereof. The simulation model presented is by no means complete and various improvements or adjustments with regards to simulation techniques, modelling approaches and assumptions made may increase the flexibility, level of accuracy and complexity of the model. This section therefore contains a number of suggestions with regards to possible future research, as model improvements or model extensions, which may be pursued as follow-up work to the contributions of this thesis.

8.3.1 Recommendations for current model improvements

Owing to computational expense and model complexity, the scope of the model was limited to the essential factors as discussed in §5.3, although the addition to or extension of certain aspects of the model may increase its performance. The following recommendations are therefore made with respect to improvements which could be made to the existing model.

Proposal 8.1 *Increase the model granularity.*

The model presented in this study aggregates 10 000 people to one agent in order to model the entire Syrian population. Owing to model complexities and computational expense, the model is limited either in scale or granularity. Frydenlund [43] proposes the use of super computers which would allow for the granularity of the simulation model to be increased. This would allow for the number of individuals represented by a single agent to be decreased to 1 000 or less. This will potentially also allow the inclusion of more social context in the model in terms of the agent characteristics, such as ethnicity and language. Furthermore, an agent's relationship to other modelled agents, as well as its existing social network could be considered in model execution.

ABM allows for the unique modelling of individual agents, their decision-making, interactions with one another and the environment, as well as the effect of such interaction on decision-making [67]. It therefore facilitates the explicit modelling of social networks resulting from social interactions and, as such, the ABM environment creates a platform for the existing model to, with an adjusted granularity, include social network modelling considerations in a model.

Proposal 8.2 *Consider the openness score of a country towards refugees to be different from its openness score towards undocumented migrants.*

The model described in this study assumes that a country's openness score towards refugees and undocumented migrants is the same, however, in reality, this might not always be the case. Lemos [75] affirms that the attitude of a country towards these two groups of people may differ. Incorporating this consideration may allow for an even better representation of reality in terms of the number of people who choose to relocate to these countries as either a refugee or an undocumented migrant.

Proposal 8.3 *Refine the model characteristics by means of an expert panel discussion.*

Although insight and knowledge has already been gained from various subject matter experts, these discussions occurred in isolation, where the modeller met with a single expert at a time. An open, expert panel discussion will allow for even more insight, as the opinions and knowledge of the various experts could be refined and synthesised, with new ideas and concepts possibly being considered. The panel discussions to which the modeller has been invited in 2018 could serve as an ideal platform to discuss the model in this study with various experts, aiming to refine the modelling aspects and achieve consensus surrounding simplifications and assumptions made in the model.

Proposal 8.4 *Consideration of specific regions within countries as proposed destinations for relocation.*

The simulation model described in this thesis allows agents to choose specific destinations when choosing to relocate. IDPs choose their destination based on the attractiveness of an area, considering the presence of conflict, the population density and distance. Refugees and undocumented migrants are modelled to choose a neighbouring country as proposed destination based on various criteria. When an agent choose to relocate across the border, however, the options of destinations are restricted to the neighbouring countries, each in their entirety only. Frydenlund [43] suggests investigating the option of modelling regions or cities within a neighbouring country as separate destination options, since the "pull" factors (*i.e.* the attraction of a place) of different regions or cities within a country may differ greatly. The true, accurate distance criteria should also have a greater impact on the decision-making of an agent in this case, as the distance to neighbouring countries in the simulation model presented in this thesis were simply calculated as the Euclidean distance between an agent and the nearest border of the country under consideration.

Proposal 8.5 *Consideration of the movement of agents between neighbouring countries as well as repatriation.*

In the simulation model, the consequent decision-making of an agent who relocates to a neighbouring country is not considered. In the real-world scenario, individuals who move across the Syrian border may further move between neighbouring countries, or move from one of these countries towards Europe, in which case the country of initial entry is known as a transit country. When modelling short-term effects of conflict, the modelling of transit countries might not have an effect to the model output, however, with the long-term modelling of forced migration, as is the case with modelling the forced migration in Syria, the modelling of transit countries may well have an influence. Another related factor is the repatriation of refugees and undocumented migrants, where the individual may decide to return to their country of origin. Including the movement of agents between various neighbouring countries, as well as the repatriation of agents, will allow for a greater level of accuracy with regards to the model outputs.

8.3.2 Recommendations for current model extensions

In conjunction ro improvements on the existing model, plausible extensions to the model as presented in this thesis could be applied in order to increase flexibility, versatility and complexity. The following recommendations are made in respect to extensions of the current model for future research.

Proposal 8.6 *Incorporate GIS capabilities.*

The model developed in this study utilises a static two-dimensional image as representation of the physical area modelled. The area modelled is thus restricted to this static image and geographic data pertaining to the area, such as terrain and roads, are not considered. Incorporating GIS capabilities in the model will allow for the model to expand the modelled area and make it dynamic in nature. The data associated to the physical area would then also be readily available without requiring it as input from the modeller or user. This would also allow the model to be easily adjusted to model various geographic areas worldwide.

Proposal 8.7 *Collaboration with humanitarian aid organisations.*

The model presented in this study is developed from academic knowledge and practical insights and opinions of experts, although the study did not have the opportunity of collaborating with a humanitarian organisation. Peres *et al.* [95] studied trends in humanitarian logistics and disaster relief research and concluded, however, that close collaboration between theory and practice is required, therefore suggesting that academia work together with humanitarian organisations in order to conduct case studies and empirical research. In this manner, knowledge and data on the subject will allow research to be more effective in assisting humanitarian aid. One of the subject matter experts involved in this study works in collaboration with UNOCHA and, in light of this, the author has been invited to spend time at one of the largest refugee camps in Greece for the purpose of extending the existing research.

Proposal 8.8 *Consideration of characteristics pertaining to conflict.*

The conflict as modelled in this study is not considered as having unique characteristics or behaviour rules, other than conflict intensity and the spread and depletion thereof. Frydenlund [43] suggests modelling conflict as entities so as to account for the different parties involved in conflict situations, such as militia groups, or the government army, as individuals may have different reactions to the various initiators of conflict. The characteristics of a person, such as religion and ethnicity, may further also influence the interaction of the person to the conflict, making this extension necessary in order to accurately track decision-making and associated movement of fleeing persons.

Proposal 8.9 *Adjust the model to fit other contexts.*

The model as developed in this study is specific to conflict-induced situations and the context of the existing situation in Syria. A further extension to the model could be to include accommodating for the modelling of forced migrations in other contexts. There exist other pertinent examples of conflict induced migration, such as in North Africa and Bangladesh, which would require a change in the simulation model to adapt for the different population characteristics and social aspects. Furthermore, forced migration caused by non-conflict situations, such as climate migration or natural disasters, could also be modelled with suitable extensions made to the model. The concept of climate refugees especially is becoming more apparent in the research, as climate change is expected to become a more prominent cause of forced migration [16]. This extension to the model could allow for a versatile, agent-based simulation model which is applicable in the general field of forced migration.

Proposal 8.10 *Utilise the model to perform techno-economic analyses.*

Within this study, the output of the simulation model is graphically recorded and provides the user with data pertaining to the number of people forcibly displaced with respect to the different movement types and the various destinations. An useful extension of this model could be to enable techno-economic analyses, based on the model output and utilising the data, to investigate resource allocations and the identification of locations to station refugee camps in an attempt to better cater for the anticipated number of forcibly displaced individuals who would require assistance.

References

[1] Achilli L, 2015, *Syrian refugees in Jordan: A reality check*, (Unpublished) Technical Report, Migration Policy Centre, European University Institute, Florence.

[2] Aiyar S, Barkbu B, Batini N, Berger H, Detragiache E, Dizioli A, Ebeke C, Lin H, Kaltani L, Sosa S, Spilimbergo A & Topalova P, 2016, *The Refugee Surge in Europe: Economic Challenges*, (Unpublished) Technical Report, International Monetary Fund, Washington (DC).

[3] Akbulut Y, 2014, *Autonomous resource allocation in clouds: A comprehensive analysis of single synthesizing criterion and outranking based multiple criteria decision analysis methods*, MSc thesis, University of Victoria, Victoria.

[4] Aksel DB, 2017, Coordinator at *Migration Research Center at Koç University*, [Personal Communication], Contactable at `dabayraktar@ku.edu.tr`.

[5] Alhanaee A & Csala D, 2016, *Motives to flee: Modeling Syrian refugee crisis*, (Unpublished) Technical Report, Masdar Institute of Science and Technology, Abu Dhabi.

[6] An L, 2011, *Modeling human decisions in coupled human and natural systems: Review of agent-based models*, Ecological Modelling, **229**, pp. 25–36.

[7] Anderson J, Chaturvedi A & Cibulskis M, 2007, *Simulation tools for developing policies for complex systems: Modeling the health and safety of refugee communities*, Health Care Management Science, **10(4)**, pp. 331–339.

[8] Arentze TA, Kowald M & Axhausen KW, 2013, *An agent-based random-utility-maximization model to generate social networks with transitivity in geographic space*, Social Networks, **35**, pp. 451–459.

[9] Arous R, 2013, *The predicament of Syrian refugees in navigating Cairo's urban spaces and the complexities of governance in turbulent times*, PhD dissertation, University of Stuttgart, Stuttgart.

[10] Baker A, 2017, *Promised Land*, TIME Magazine, **190(14)**, pp. 22–27.

[11] Bank TW, 2017, *World Development Indicators*, (Unpublished) Technical Report, The World Bank, Washington (DC).

[12] Banks J, 1998, *Handbook of Simulation*, John Wiley & Sons, New York (NY).

[13] Bell DE, Raiffa H & Tversky A (Eds), 1988, *Decision making: Descriptive, normative, and prescriptive interactions*, Cambridge University Press, Cambridge.

[14] Bell S, Alves S, de Oliviera ES & Zuin A, 2010, *Migration and Land Use Change in Europe: A Review*, Living Reviews in Landscape Research, **4(2)**.

[15] Betts A, 2016, *Our refugee system is failing. Here's how we can fix it*, [Online], [Cited March 2017], Available from `https://www.ted.com/talks/alexander_betts_our_refugee_system_is_failing_here_s_how_we_can_fix_it`.

[16] Biermann F & Boas I, 2010, *Preparing for a warmer world: Towards a global governance system to protect climate refugees*, Global Environmental Politics, **10(1)**, pp. 60–88.

[17] Bohnet H, 2012, *Large refugee populations, resource scarcity and conflict*, Proceedings of the 2^{nd} Annual General Conference of the European Political Science Association, Berlin, pp. 21–23.

[18] Bonabeau E, 2002, *Agent-based modeling: methods and techniques for simulating human systems*, Proceedings of the 3^{rd} Colloquium of the National Academy of Sciences, Irvine (CA), pp. 7280–7287.

[19] Borshchev A, 2013, *The Big Book of Simulation Modeling: Multimethod Modeling with AnyLogic 6*, AnyLogic North America.

[20] Borshchev A & Filippov A, 2004, *From system dynamics and discrete event to practical agent based modeling: Reasons, techniques, tools*, Proceedings of the 22^{nd} International Conference of the System Dynamics Society, Oxford, pp. 25–29.

[21] Bundy C, 2016, *Migrants, refugees, history and precedents*, Forced Migration Review, **51**, pp. 5–6.

[22] Castles S, 2003, *Towards a sociology of forced migration and social transformation*, Sociology, **77(1)**, pp. 13–34.

[23] Central Intelligence Agency, 2015, *The World Factbook*, [Online], [Cited July 2017], Available from `http://www.cia.gov/library/publications/the-worldfactbook`.

[24] Centre IDM, 2016, *Global report on internal displacement*, (Unpublished) Technical Report, Norwegian Refugee Council, Geneva.

[25] Cervantes JA, Ramos F & Wang Y, 2015, *A formal model inspired on human decision-making process*, Proceedings of the 2015 IEEE 14^{th} International Conference on Cognitive Informatics & Cognitive Computing, Beijing, pp. 375–383.

[26] Chalabi M, 2013, *What happened to history's refugees?*, [Online], [Cited July 2017], Available from `https://www.theguardian.com/news/datablog/interactive/2013/jul/25/what-happened-history-refugees`.

[27] Churchman CW, Ackoff RL & Smith NM, 1954, *An approximate measure of value*, Journal of the Operations Research Society of America, **2(2)**, pp. 172–187.

[28] Collier D & Mahoney J, 1996, *Selection bias in qualitative research*, World Politics, **49(1)**, pp. 56–91.

[29] Collins AJ & Frydenlund E, 2016, *Agent-based modeling and strategic group formation: A refugee case study*, Proceedings of the 2016 Winter Simulation Conference, pp. 1289–1300.

[30] Colson G & De Bruyn C, 1989, *Models and methods in multiple objectives decision making*, Mathematical and Computer Modelling, **12(10–11)**, pp. 1201–1211.

[31] Crooks AT & Wise S, 2013, *GIS and agent-based models for humanitarian assistance*, Computers, Environment and Urban Systems, **41**, pp. 100–111.

[32] Davenport C, Moore W & Poe S, 2003, *Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989*, International Interactions, **29(1)**, pp. 27–55.

[33] Davin E & Majidi N, 2009, *Study on cross border population movements between Afghanistan and Pakistan*, (Unpublished) Technical Report, United Nations High Commissioner for Refugees, Kabul.

[34] Dillon SM, 1998, *Descriptive decision making: Comparing theory with practice*, Proceedings of the 33^{rd} ORSNZ Conference, University of Auckland, Auckland.

[35] Divekar AA, Bangal S & Sumangala D, 2012, *The study of prescriptive and descriptive models of decision making*, International Journal of Advanced Research in Artificial Intelligence, **1(1)**, pp. 71–74.

[36] Dragostinova T, 2016, *Refugees or immigrants? The migration crisis in Europe in historical Perspective*, Origins, **9(4)**.

[37] Edwards S, 2008, *Computational tools in predicting and assessing forced migration*, Journal of Refugee Studies, **21(3)**, pp. 347–359.

[38] Eldrandaly K, Hadi A & Abdelaziz AN, 2009, *An expert system for choosing the suitable MCDM method for solving a spatial decision problem*, Proceedings of the 9^{th} lnternational Conference on Production Engineering, Design and Control, Alexandria.

[39] Fargues P, 2011, *Voice after exit: Revolution and migration in the Arab world*, (Unpublished) Technical Report, Migration Policy Institute, Washington (DC).

[40] Fargues P & Fandrich C, 2012, *The European Response to the Syrian Refugee Crisis - What Next?*, (Unpublished) Technical Report, European University Institute, San Domenico di Fiesole (FI).

[41] Ferris E, Kirişci K & Shaikh S, 2013, *Syrian crisis: Massive displacement, dire needs and a shortage of solutions*, (Unpublished) Technical Report, The Brookings Institution, Washington (DC).

[42] Fioretti G, 2013, *Agent-Based Simulation Models in Organization Science*, Organizational Research Methods, **16(2)**, pp. 227–242.

[43] Frydenlund E, 2017, Research Assistant Professor at *Virginia Modeling, Analysis & Simulation Center, Old Dominion University*, [Personal Communication], Contactable at `efrydenl@odu.edu`.

[44] Garifullin M, Borshchev A & Popkov T, 2007, *Using Anylogic and agent-based approach to model consumer market*, Proceedings of the 6^{th} EUROSIM Congress on Modelling and Simulation, Ljubljana, pp. 1–5.

[45] Gigerenzer G, 2007, *Gut feelings: The intelligence of the unconscious*, Penguin, New York (NY).

[46] Gigerenzer G & Gaissmaier W, 2011, *Heuristic decision making*, Annual Review of Psychology, **62**, pp. 451–482.

[47] Global Opportunity Network, 2017, *Global Opportunity Report 2017*, (Unpublished) Technical Report, DNV GLAS, Oslo.

[48] Green WG & Mcginnis SR, 2002, *Thoughts on the higher order taxonomy of disasters*, Notes on the Science of Extreme Situations, **(7)**.

[49] Greenwood MJ, 2005, *Modeling migration*, Encyclopedia of Social Measurement, **2**, pp. 725–734.

[50] Groen D, 2017, Lecturer in Simulation and Modelling at *Department of Computer Science, Brunel University London*, [Personal Communication], Contactable at `derek.groen@brunel.ac.uk`.

[51] Groen D, Bhati A, Suter J, Hetherington J, Zasada S & Coveney P, 2016, *FabSim: facilitating computational research through automation on large-scale and distributed e-infrastructures*, Computer Physics Communications, **207**, pp. 375–385.

[52] Guitouni A & Martel J.-M, 1998, *Tentative guidelines to help choosing an appropriate MCDA method*, European Journal of Operational Research, **109(2)**, pp. 501–521.

[53] Gulden T, Harrison JF & Crooks AT, 2011, *Modeling cities and displacement through an agent-based spatial interaction model*, Proceedings of the 2^{nd} Annual Conference of the Computational Social Science Society of the Americas CSSSA-2011, Santa Fe, (NM).

[54] Hagen-zanker J, 2008, *Why do people migrate? A review of the theoretical literature*, (Unpublished), Maastricht Graduate School of Governance, Maastricht.

[55] Hailegiorgis A & Crooks AT, 2008, *Agent-based modeling for humanitarian issues: Disease and refugee camps*, Proceedings of the Computational Social Science Society of America Conference, Santa Fe, (NM).

[56] Harrison E, 2016, *Modeling Syrian internally displaced person movements: A case study of conflict, travel, accessibility, and resource availability*, Student Writing, **13**, pp. 1–22.

[57] Huang IB, Keisler J & Linkov I, 2011, *Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends*, Science of the Total Environment, **409(19)**, pp. 3578–3594.

[58] Hwang CL & Yoon KP, 1981, *Multiple attribute decision making: Methods and applications*, Springer-Verlag, Berlin.

[59] IFRC, 2014, *Disaster and crisis management*, [Online], [Cited March 2017], Available from `https://www.ifrc.org/en/what-we-do/disaster-management/`.

[60] Institute for Economics and Peace, 2015, *Global Peace Index 2015: Quantifying Peace and its Benefits*, (Unpublished) Technical Report, Institute for Economics & Peace, Sydney.

[61] Ivanov D, 2016, *Operations and supply chain simulation with AnyLogic 7.2: Decision-oriented introductory notes for master students*, Berlin School of Economics & Law, Berlin.

[62] Jennings NR, 2000, *On agent-based software engineering*, Artificial Intelligence, **117(2)**, pp. 277–296.

[63] Johnson RT, Lampe TA & Seichter S, 2009, *Calibration of an agent-based simulation model depicting a refugee camp scenario*, Proceedings of the 2009 Winter Simulation Conference, pp. 1778–1786.

[64] Jonson B, Mouamar P, Huber C, Reid K & Koenig DC, 2017, *Syria refugee crisis: Facts you need to know*, World Vision, pp. 1778–1786.

[65] Katsikopoulos KV, Durbach IN & Stewart TJ, 2017, *When should we use simple decision models? A synthesis of various research strands*, Omega.

[66] Kiesling E, Günther M, Stummer C & Wakolbinger LM, 2012, *Agent-based simulation of innovation diffusion: A review*, Central European Journal of Operations Research, **20**, pp. 183–230.

[67] Klabunde A & Willekens F, 2016, *Decision-making in agent-based models of migration: State of the art and challenges*, European Journal of Population, **32**, pp. 73–97.

[68] Klüpfel H, Schreckenberg M & Meyer-König T, 2005, *Models for crowd movement and egress simulation*, Proceedings of the Traffic and Granular Flow 2003, Berlin, pp. 357–372.

[69] Koepke B, 2011, *The Situation of Afghans in the Islamic Republic of Iran Nine Years After the Overthrow of the Taliban Regime in Afghanistan*, Middle East Institute - Fondation pour la Recherche Stratégique, pp. 1–13.

[70] Kunz E, 1973, *The refugee in flight: Kinetic models and forms of dispacement*, International Migration Review, **7(2)**, pp. 125–146.

[71] Law AM, 2013, *A tutorial on how to select simulation input probability distributions*, Proceedings of the 2013 Winter Simulation Conference, Washington (DC), pp. 306–320.

[72] Law AM, 2008, *How to build valid and credible simulation models*, Proceedings of the 2008 Winter Simulation Conference, Fayetteville (AR), pp. 24–32.

[73] Law AM & Kelton WD, 1982, *Simulation Modeling & Analysis*, 1st Edition, McGraw-Hill, New York.

[74] Law AM & Kelton WD, 1991, *Simulation Modeling & Analysis*, 2nd Edition, McGraw-Hill, Singapore.

[75] Lemos C, 2017, Postdoctoral Research Fellow at *Department of Religion, Philosophy and History, University of Agder*, [Personal Communication], Contactable at `carlos.lemos@uia.no`.

[76] Lemos C, 2017, *Agent-based modeling of social conflict: From mechanisms to complex behavior*, Springer International Publishing, Cham.

[77] Lemos C, Coelho H & Lopes RJ, 2013, *Agent-based modeling of social conflict, civil violence and revolution: State-of-the-art-review and further prospects*, CEUR Workshop Proceedings, **1113**, pp. 124–138.

[78] Linkov I, Varghese A, Jamil S, Seager TP, Kiker G & Bridges T, "Multi-criteria decision analysis: A framework for structuring remedial decisions at contaminated sites", in: *Comparative risk assessment and environmental decision making*, Dodrecht: Kluwer Academic Publishers, 2004, pp. 15–54.

[79] Macal C & North M, 2010, *Tutorial on agent-based modelling and simulation*, Journal of Simulation, **4**, pp. 151–162.

[80] Macal CM & North MJ, 2008, *Agent-based modeling and simulation: ABMS Examples*, Proceedings of the 2008 Winter Simulation Conference, Miami, (FL), pp. 101–112.

[81] MacCrimmon KR, 1968, *Decisionmaking among multiple-attribute alternatives: a survey and consolidated approach*, (Unpublished) Technical Report, Santa Monica (CA).

[82] Malczewski J, 1999, *GIS and multicriteria decision analysis*, John Wiley & Sons, New York (NY).

[83] Martinez M, 2015, *Migrant vs refugee: What's the difference?*, [Online], [Cited September 2017], Available from `http://edition.cnn.com/2015/09/08/world/what-is-difference-migrants-refugees/`.

[84] Maslow AH, 1943, *A theory of human motivation*, Psychological Review, **50**, pp. 370–396.

[85] Massam BH, 1988, *Multi-criteria decision making (MCDM) techniques in planning*, Progress in planning, **30**, pp. 1–84.

[86] McMaster D, 2001, *Asylum seekers: Australia's response to refugees*, Melbourne University Press, Melbourne.

[87] Mercy Corps, 2017, *Quick facts: What you need to know about the Syria crisis*, Iraq, Jordan, Lebanon, Syria, Turkey.

[88] Moore WH & Shellman SM, 2004, *Fear of persecution: Forced migration*, The Journal of Conflict Resolution, **48(5)**, pp. 723–745.

[89] Moore WH & Shellman SM, 2006, *Refugee or Internally Displaced Person? To Where Should One Flee?*, Comparative Political Studies, **39(5)**, pp. 599–622.

[90] Moore WH & Shellman SM, 2007, *Whither will they go? A global study of refugees' destinations*, International Studies Quarterly, **51(4)**, pp. 811–834.

[91] Mota P, Campos AR & Neves-Silva R, 2013, *First look at MCDM: Choosing a decision method*, Advances in Smart Systems Research, **3(2)**, pp. 25–30.

[92] Ölçer AI & Odabaşi AY, 2005, *A new fuzzy multiple attributive group decision making methodology and its application to propulsion/manoeuvring system selection problem*, European Journal of Operational Research, **166(1)**, pp. 93–114.

[93] Orfano L, 2015, *Migrants and refugees, an agent-based simulation model for the economic empirical evidence and long-term effects*, MSc thesis, Università degli studi di Torino, Turin.

[94] Pauver B, Twigg J & Sagramola S, 2016, *Migrants, refugees, asylum seekers: Inclusion in disaster preparedness and response*, (Unpublished) Technical Report, EUR-OPA Council of Europe, Lisbon.

[95] Peres E, Brito Jr I, Leiras A & Yoshizaki H, 2012, *Humanitarian logistics and disaster relief research: Trends, applications, and future research directions*, Proceedings of the 4^{th} International Conference on Information Systems, Logistics and Supply Chain, Quebec.

[96] Petersen W, 1958, *A general typology of migration*, American Sociological Review, **23(3)**, pp. 256–266.

[97] Phillips J, 2011, *Asylum seekers and refugees: What are the facts?*, (Unpublished) Technical Report, Parliament of Australia, Melbourne.

[98] Pillimore J, 2011, *Refugees, acculturation strategies, stress and integration*, Journal of Social Policy, **40(3)**, pp. 575–593.

[99] Refugee Studies Centre, 2016, *Annual Report 2015 - 2016*, (Unpublished) Technical Report, University of Oxford, Oxford.

[100] Refugee Studies Centre, 2015, *The Syrian humanitarian disaster: Disparities in perceptions, aspirations and behaviour in Lebanon, Jordan and Turkey*, (Unpublished) Technical Report 3, University of Oxford, Oxford.

[101] Rehm M, 2012, *Migration and remittances. An agent-based model*, PhD dissertation, New School University, New York (NY).

[102] Richmond AH, 1993, *Reactive migration: Sociological perspectives on refugee movements*, Journal of Refugee Studies, **6(1)**, pp. 6–24.

[103] Roy B, 1985, *Méthodologie multicritère d'aide à la décision*, Economica, Paris.

[104] Saad G, 2016, *Normative, prescritive, and descriptive decision making*, [Online], [Cited October 2017], Available from `https://www.youtube.com/watch?v=%7B%5C_%7DjNISKlggms`.

[105] Saarinen V & Ojala J, 2017, *The flow towards Europe*, [Online], [Cited April 2017], Available from `https://www.lucify.com/the-flow-towards-europe/`.

[106] Sarzin Z, 2017, *Stocktaking of global forced displacement data*, (Unpublished) Technical Report, World Bank's Global Program on Forced Displacement.

[107] Schriber TJ & Brunner DT, 2013, *Inside discrete-event simulation software: How it works and why it matters*, Proceedings of the 2013 Winter Simulation Conference, Washington (DC), pp. 424–438.

[108] SHANNON R, 1998, *Introduction To the Art and Science of Simulation*, Proceedings of the 1998 Winter Simulation Conference, Washington (DC), pp. 7–14.

[109] SHANNON R, 1975, *Systems simulation: The art and science*, Prentice-Hall, Englewood Cliffs (NJ).

[110] SHOMARY WB, 2017, Doctoral Student at *Department of Child and Youth Studies, Stockholm University*, [Personal Communication], Contactable at wiji.bohme.shomary@buv.su.se.

[111] SIMON HA, 1955, *A behavioral model of rational choice*, The Quarterly Journal of Economics, **69(1)**, pp. 99–118.

[112] SIMON HA, 1990, *Prediction and prescription in system modeling*, Operations Research, **38(1)**, pp. 7–14.

[113] SIMON HA, 1956, *Rational choice and the structure of the environment*, Psychological Review, **63(2)**, pp. 129–138.

[114] SIMON HA, 1960, *The new science of management decision*, Prentice Hall, New Jersey.

[115] SMITH C, 2017, Research Fellow at *School of Global Studies, University of Sussex*, [Personal Communication], Contactable at c.d.smith@sussex.ac.uk.

[116] SMITH C, WOOD S & KNIVETON D, 2010, *Agent based modelling of migration decision-making*, Proceedings of the 8^{th} European Workshop on Multi-Agent Systems, Paris.

[117] SOKOLOWSKI JA, BANKS CM & HAYES RL, 2014, *Modeling population displacement in the Syrian city of Aleppo*, Proceedings of the 2014 Winter Simulation Conference, Savannah (GA), pp. 252–263.

[118] STEWART R, 2004, *Simulation: The practice of model development and use*, John Wiley & Sons, Chichester.

[119] STEWART TJ, 2017, Professor at *Department of Statistical Sciences, University of Cape Town*, [Personal Communication], Contactable at theodor.stewart@uct.ac.za.

[120] STEYNBERG R, 2016, *A framework for identifying the most likely successful underprivileged tertiary bursary applicants*, MEng thesis, Stellenbosch University, Stellenbosch.

[121] STURROCK DT, 2013, *Tutorial: Tips for successful practice of simulation*, Proceedings of the 2013 Winter Simulation Conference, Washington (DC).

[122] THE ANYLOGIC COMPANY, 2016, *Multimethod Simulation Software and Solutions*, [Online], [Cited October 2016], Available from http://www.anylogic.com.

[123] THE CARTER CENTER, 2017, *Tracking the front lines in Syria*, [Online], [Cited November 2017], Available from https://www.cartercenter.org/syria-conflict-map/.

[124] THE GDELT PROJECT, 2014, *Mapping the Syrian Conflict*, [Online], [Cited September 2017], Available from https://blog.gdeltproject.org/mapping-the-syrian-conflict/.

[125] THE GDELT PROJECT, 2016, *The GDELT Project: GDELT Analysis Service*, [Online], [Cited June 2017], Available from http://data.gdeltproject.org/analytics_user/20170627100301.24034.events.csv.

[126] UNITED NATIONS DEPARTMENT OF ECONOMICS AND SOCIAL AFFAIRS POPULATION DIVISION, 2017, *World Population Prospects: The 2017 Revision*, [Online], [Cited July 2017], Available from https://esa.un.org/unpd/wpp/DataQuery/.

[127] UNITED NATIONS HIGH COMMISSIONER FOR REFUGEES, 1961, *Convention an protocal relating to the status of refugees*, (Unpublished) Technical Report, UNHCR, Geneva.

[128] United Nations High Commissioner for Refugees, 2016, *Fact sheet: Global forced displacement*, (Unpublished) Technical Report, UNHCR, Geneva.

[129] United Nations High Commissioner for Refugees, 2015, *Global trends: Forced displacement in 2015*, (Unpublished) Technical Report, UNHCR, Geneva.

[130] United Nations High Commissioner for Refugees, 2014, *World at war: Global trends 2014 in review*, (Unpublished) Technical Report, UNHCR, Geneva.

[131] United Nations Office for the Coordination of Humanitarian Affairs, 2016, *Syrian Arab Republic - Humanitarian Needs Overview*, (Unpublished) Technical Report, UNOCHA.

[132] United Nations Office for the Coordination of Humanitarian Affairs, 2017, *Syrian Arab Republic - Humanitarian Needs Overview*, (Unpublished) Technical Report, UNOCHA.

[133] United Nations Office for the Coordination of Humanitarian Affairs, 2010, *Syrian Arab Republic: Population Statistics*, [Online], [Cited June 2017], Available from `https://data.humdata.org/dataset/syrian-arab-republic-other-0-0-0-0-0-0-0`.

[134] van Vuuren BJ, 2016, *An agent-based simulation model of Eldana saccharina Walker*, PhD dissertation, Stellenbosch University, Stellenbosch.

[135] Velasquez M & Hester PT, 2013, *An analysis of multi-criteria decision making methods*, International Journal of Operations Research, **10(2)**, pp. 56–66.

[136] Vincke P, 1992, *Multicriteria decision-aid*, John Wiley & Sons, Chichester.

[137] Vio E, 2016, *No way out: How Syrians are struggling to find an exit*, [Online], [Cited March 2017], Available from `https://www.irinnews.org/special-report/2016/03/10/no-way-out-how-syrians-are-struggling-find-exit`.

[138] Wang Y & Ruhe G, 2007, *The cognitive process of decision making*, International Journal of Cognitive Informatics and Natural Intelligence, **1(2)**, pp. 73–85.

[139] Zanakis SH, Solomon A, Wishart N & Dublish S, 1998, *Multi-attribute decision making: A simulation comparison of select methods*, European Journal of Operational Research, **107**, pp. 507–529.

[140] Zeleny M, 1984, *MCDM: Past decade and future trends: A source book of multiple criteria decision making*, Jai Pr, London.

[141] Zia K, Riener A, Farrahi K & Ferscha A, 2012, *A new opportunity to urban evacuation analysis: Very large scale simulations of social agent systems in repast HPC*, Proceedings of the ACM/IEEE/SCS 26^{th} Workshop on Principles of Advanced and Distributed Simulation, Zhangjiajie, pp. 233–242.

[142] Zipf GK, 1946, *The P1 P2 / D hypothesis: On the intercity movement of persons*, American Sociological Review, **11(6)**, pp. 677–686.

APPENDIX A

Model outputs illustrating conflict

Figures A.1–A.25 indicates the spread of conflict over Syria at the end of December 2016 as modelled having different probabilities of conflict spreading and depleting.

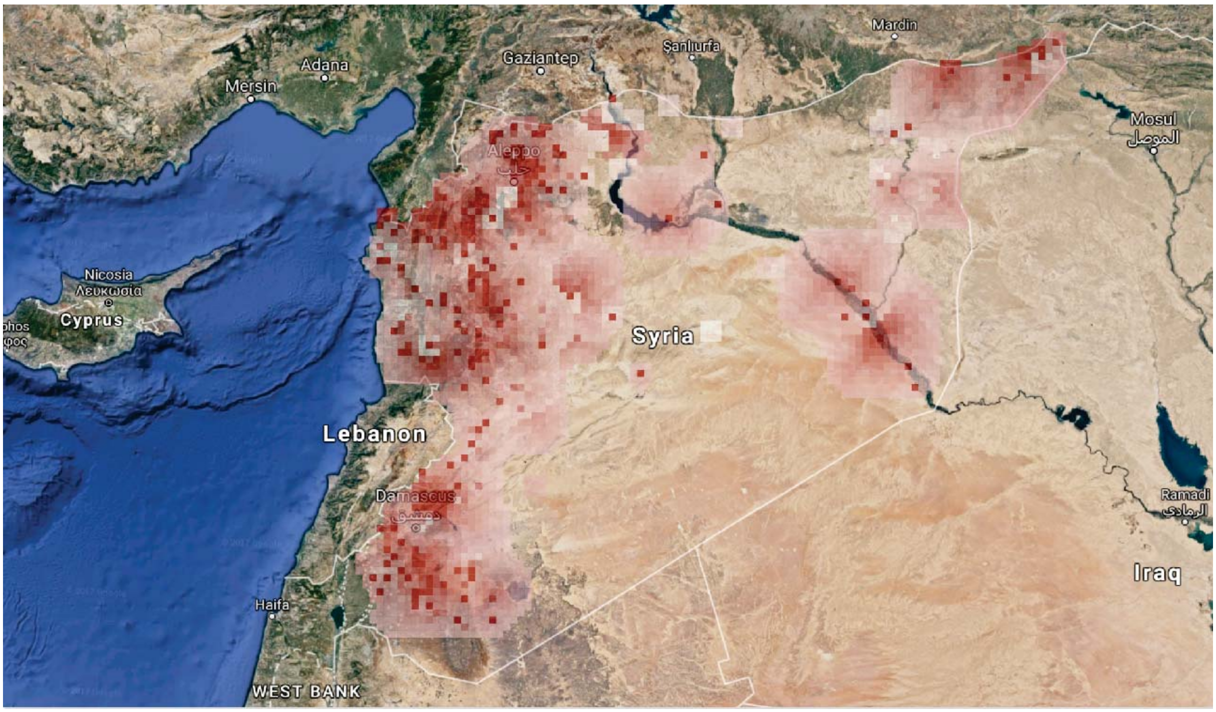


FIGURE A.1: *Experiment 1: The conflict modelled with 0.1 probability of spread and 0.1 probability of depletion.*

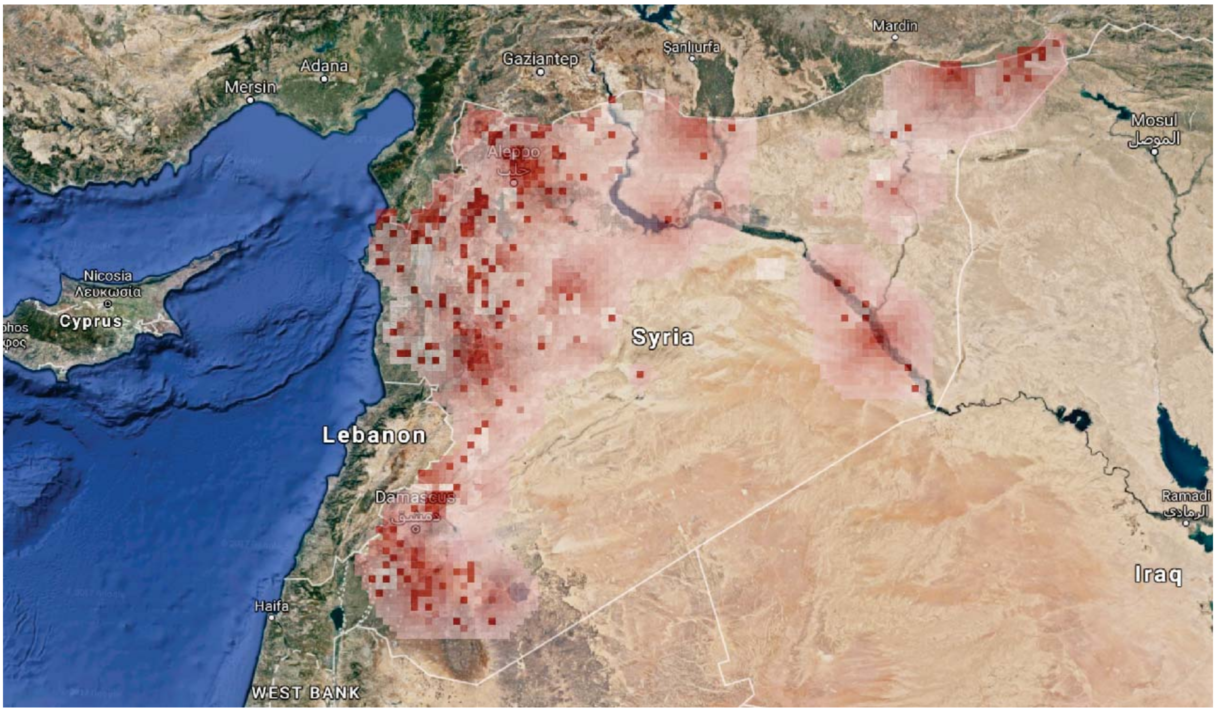


Figure A.2: *Experiment 2: The conflict modelled with 0.3 probability of spread and 0.1 probability of depletion.*

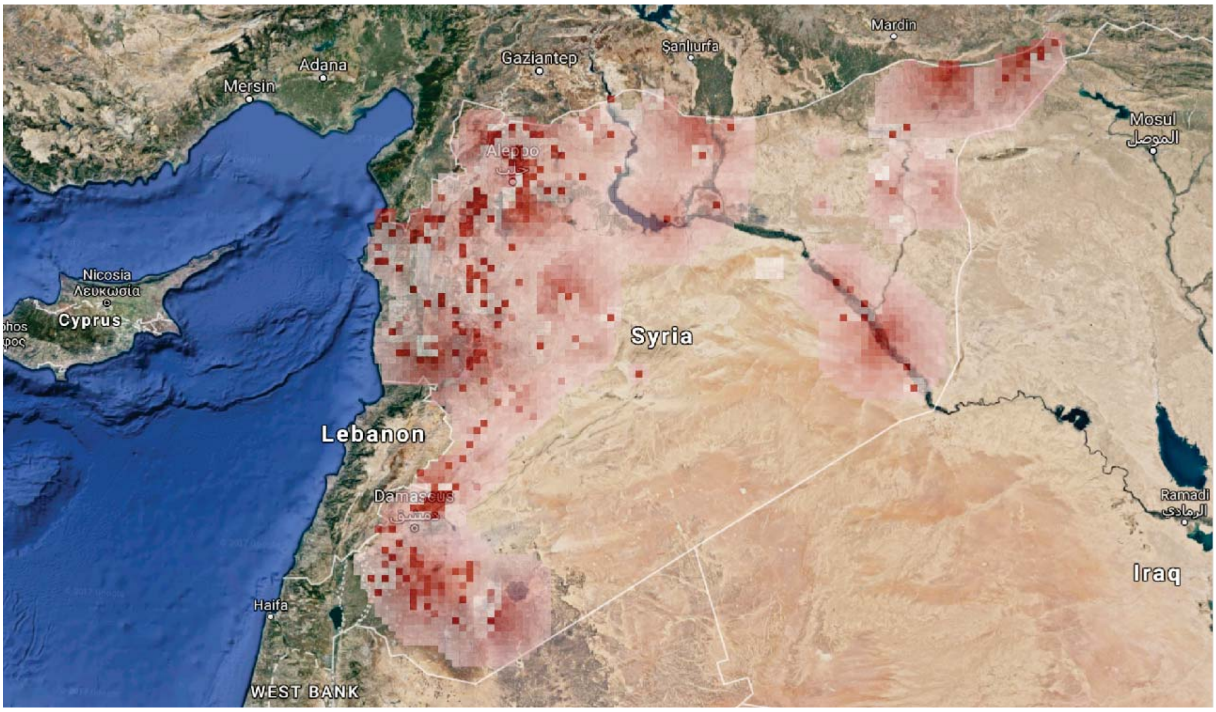


Figure A.3: *Experiment 3: The conflict modelled with 0.5 probability of spread and 0.1 probability of depletion.*

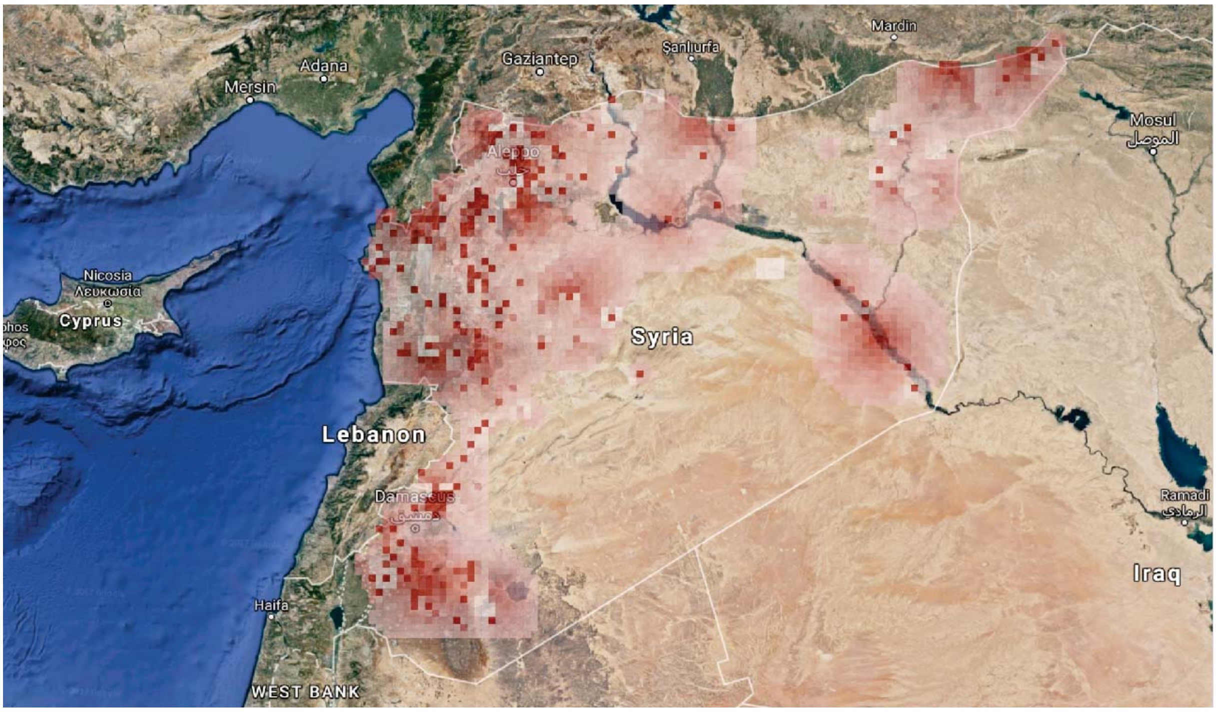


FIGURE A.4: *Experiment 4: The conflict modelled with 0.7 probability of spread and 0.1 probability of depletion.*

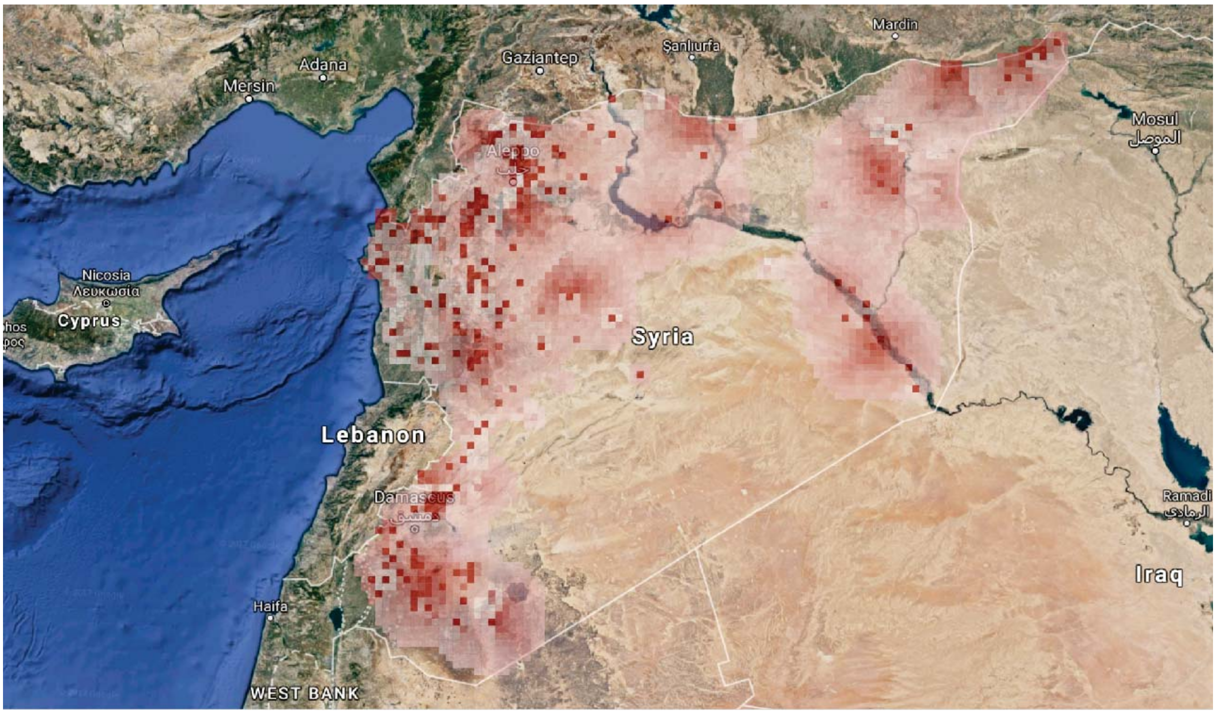


FIGURE A.5: *Experiment 5: The conflict modelled with 0.9 probability of spread and 0.1 probability of depletion.*

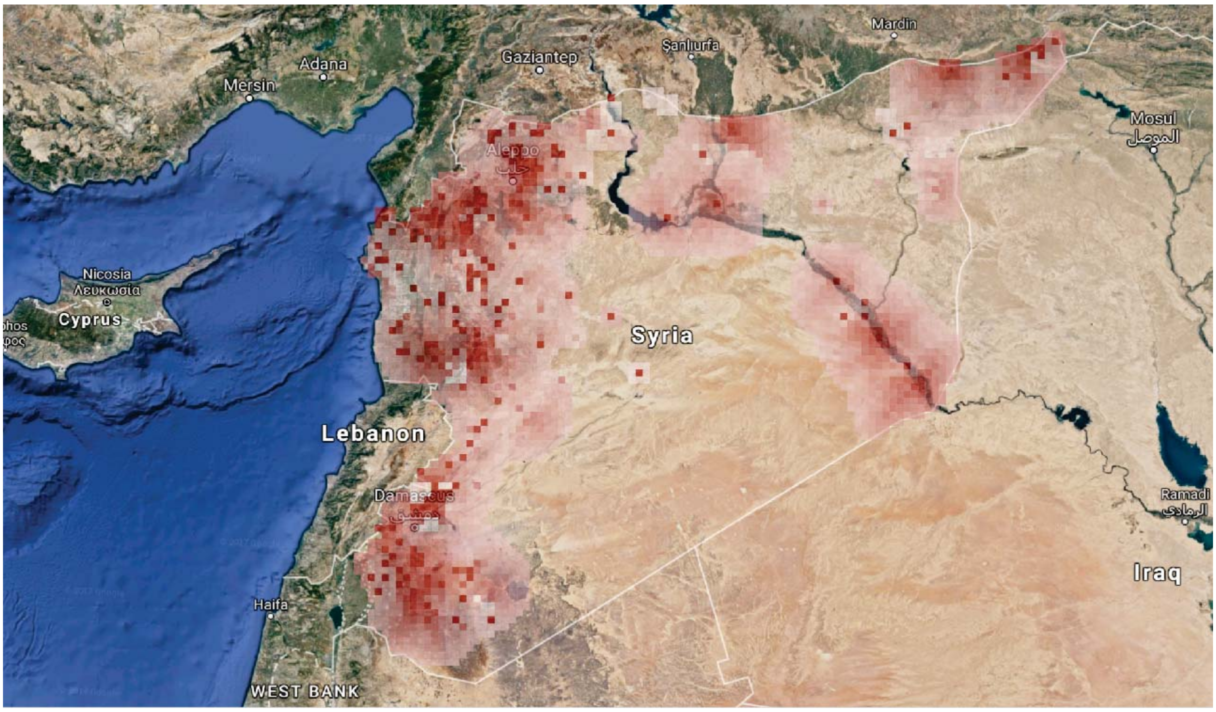


Figure A.6: *Experiment 6: The conflict modelled with 0.1 probability of spread and 0.3 probability of depletion.*

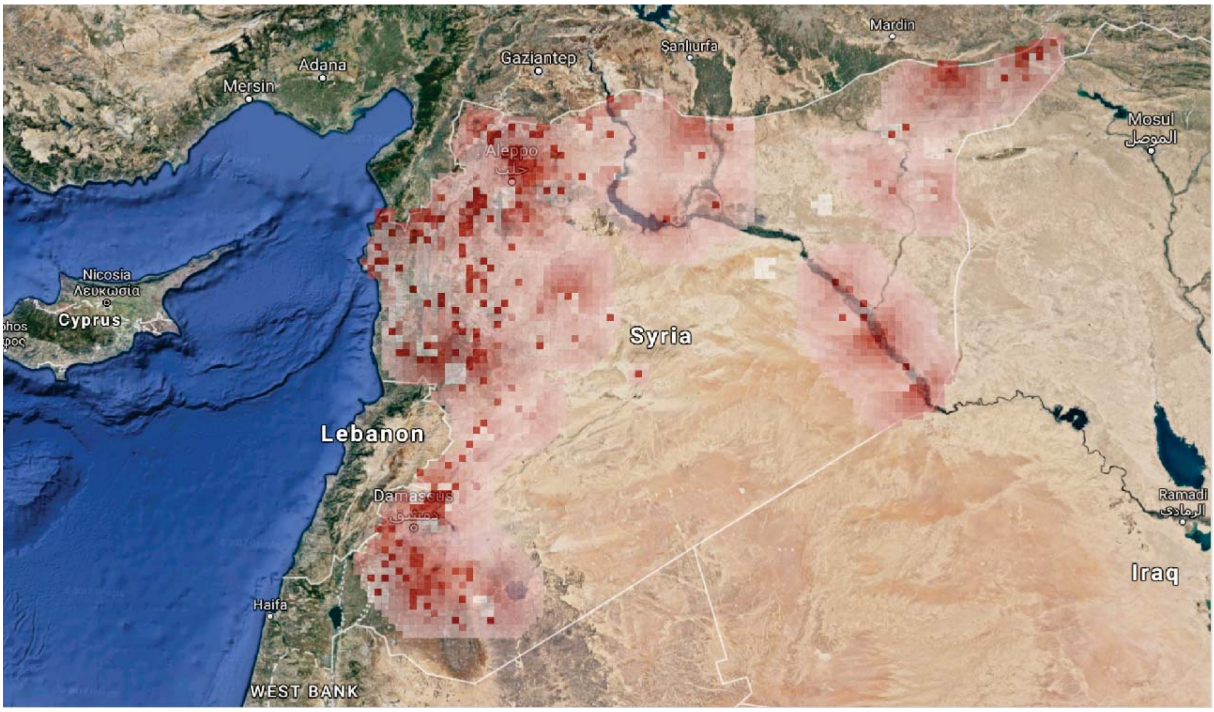


Figure A.7: *Experiment 7: The conflict modelled with 0.3 probability of spread and 0.3 probability of depletion.*

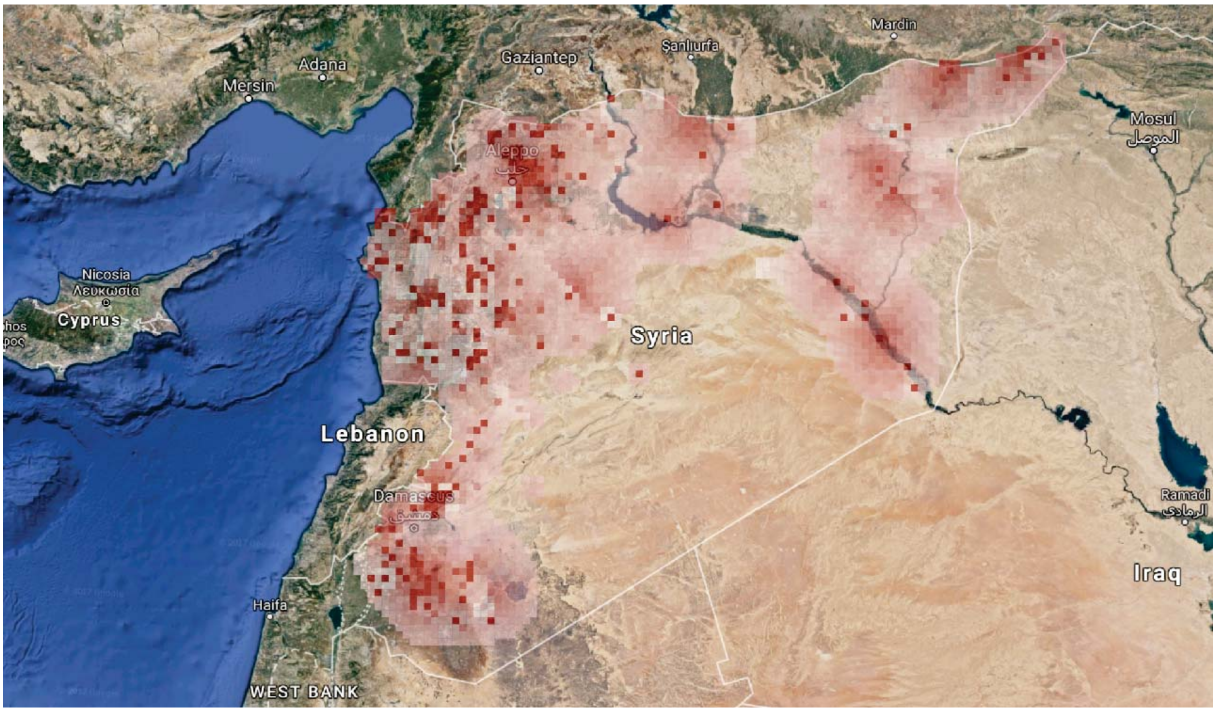


FIGURE A.8: *Experiment 8: The conflict modelled with 0.5 probability of spread and 0.3 probability of depletion.*

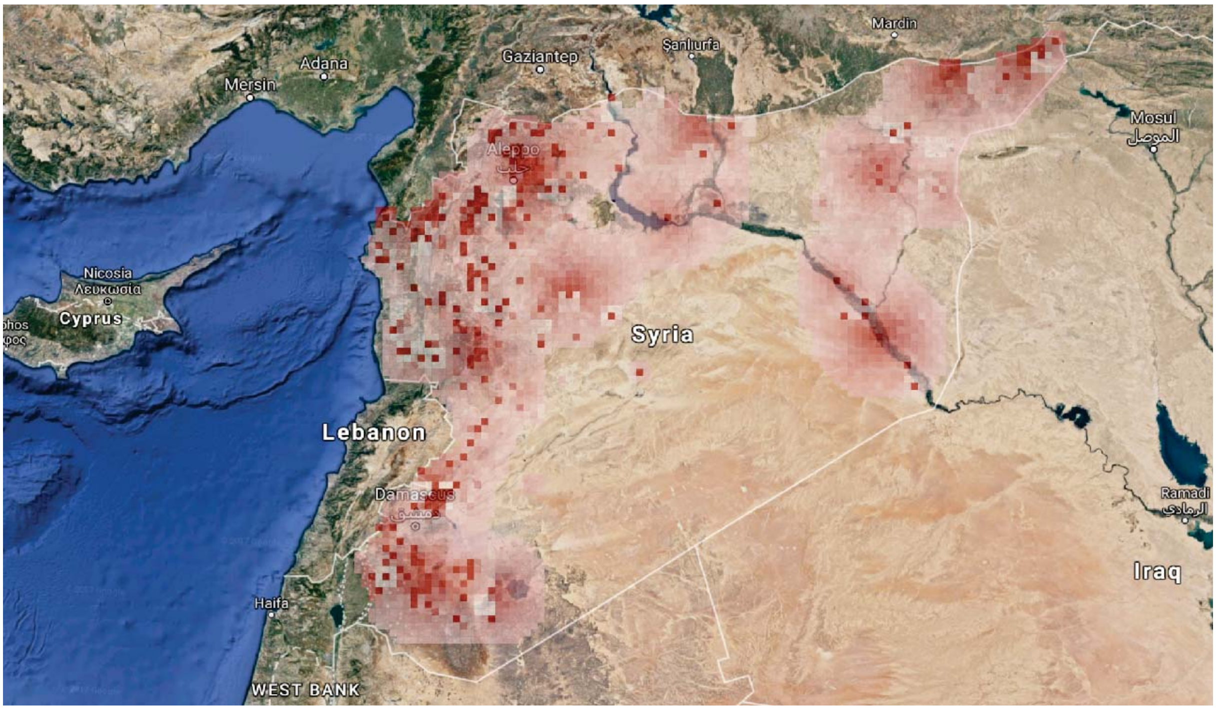


FIGURE A.9: *Experiment 9: The conflict modelled with 0.7 probability of spread and 0.3 probability of depletion.*

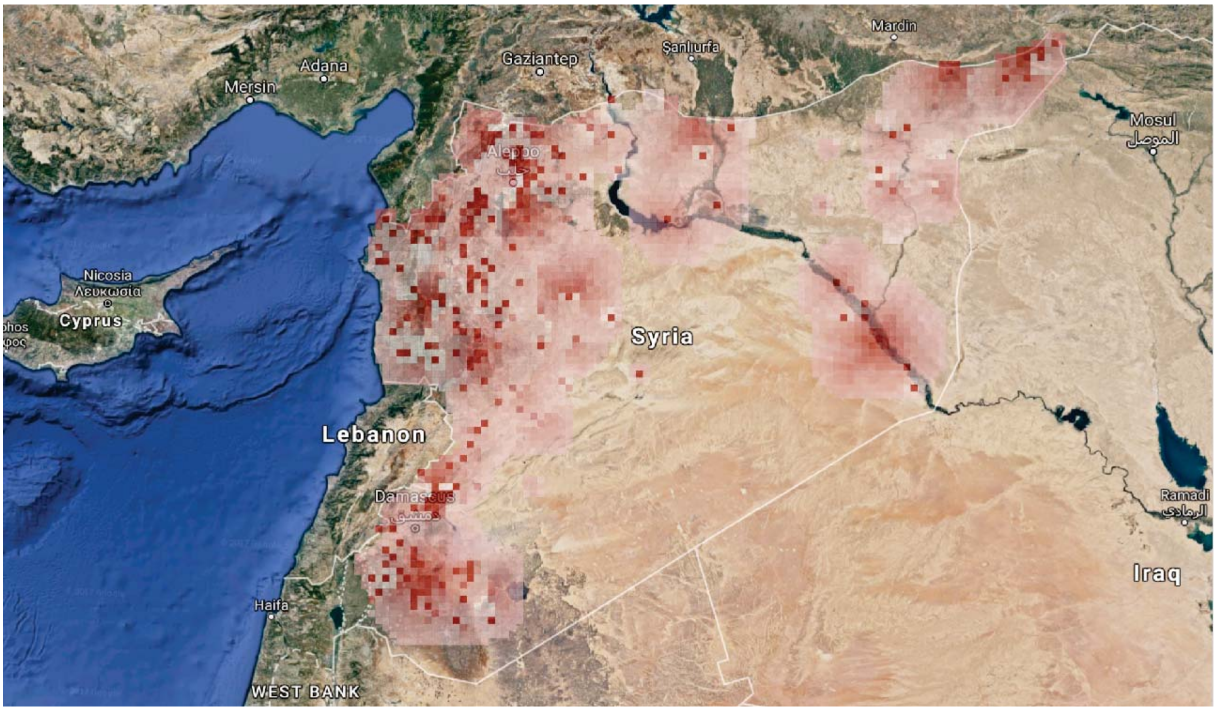


FIGURE A.10: *Experiment 10: The conflict modelled with 0.9 probability of spread and 0.3 probability of depletion.*

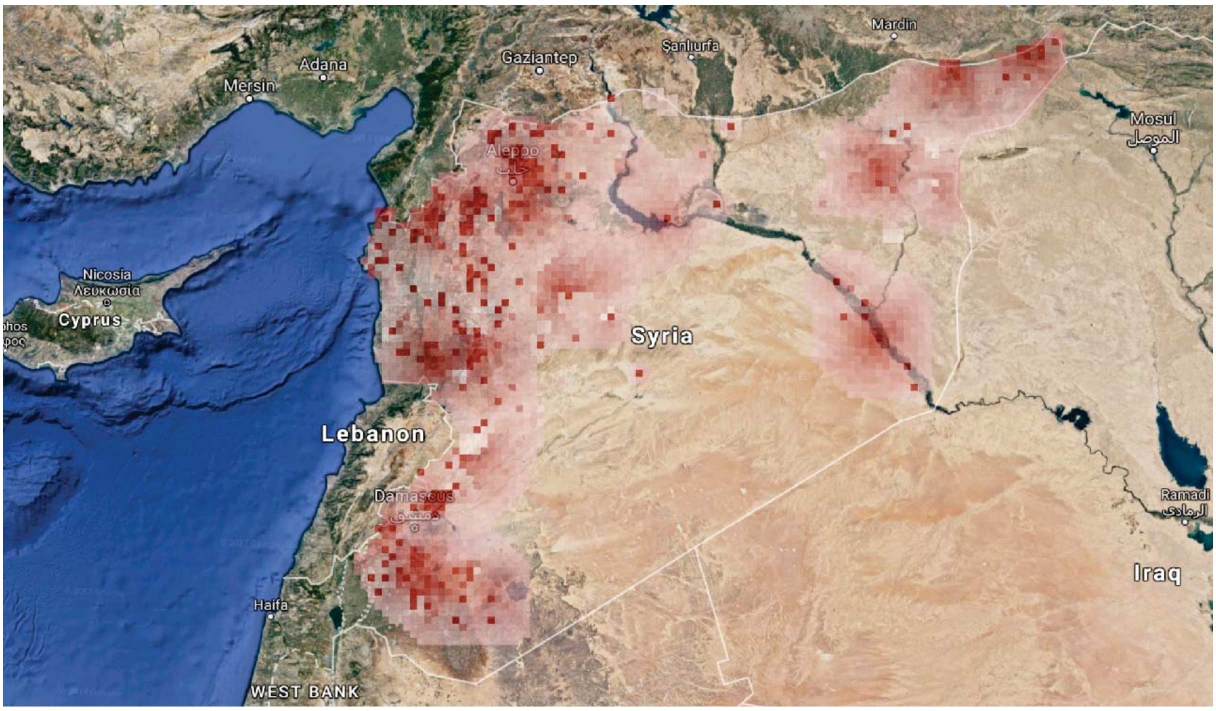


FIGURE A.11: *Experiment 11: The conflict modelled with 0.1 probability of spread and 0.5 probability of depletion.*

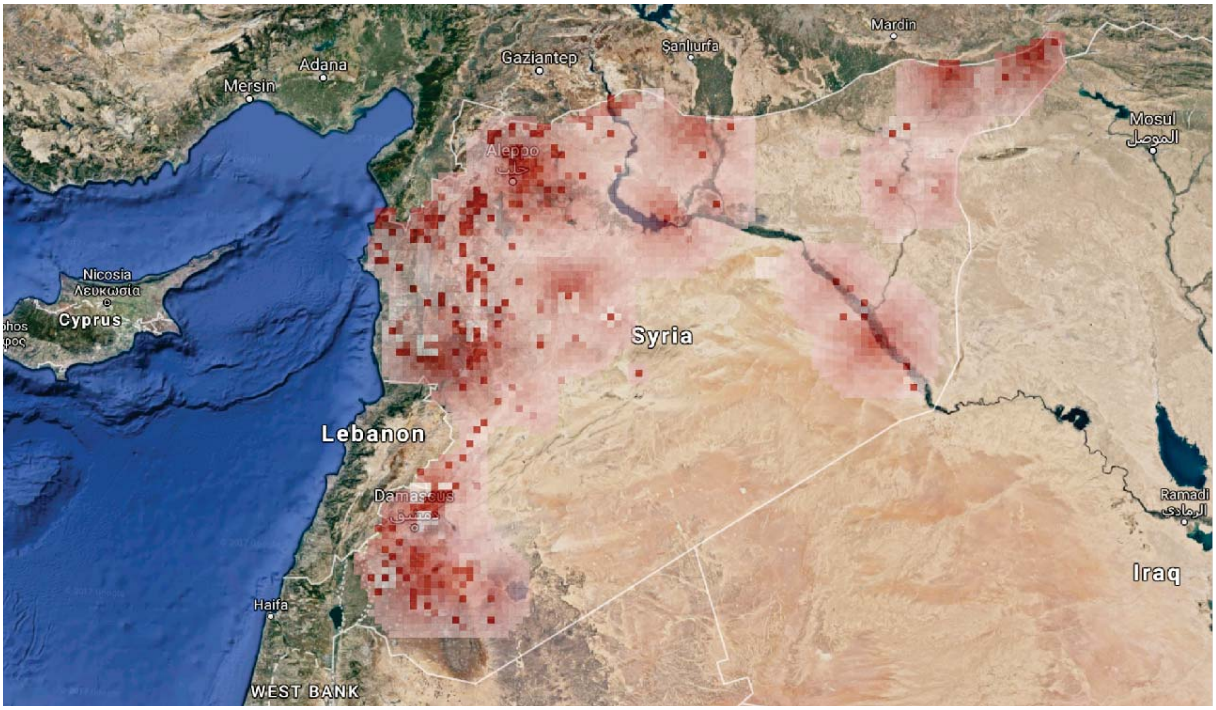


FIGURE A.12: *Experiment 12: The conflict modelled with 0.3 probability of spread and 0.5 probability of depletion.*

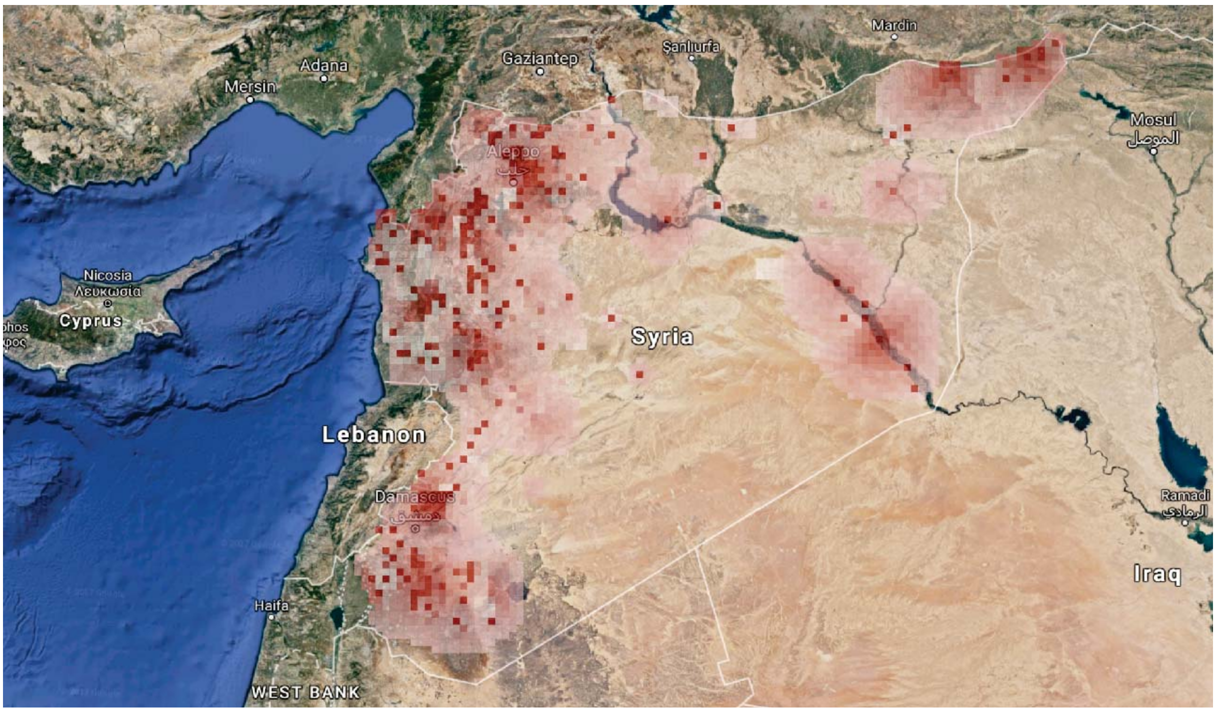


FIGURE A.13: *Experiment 13: The conflict modelled with 0.5 probability of spread and 0.5 probability of depletion.*

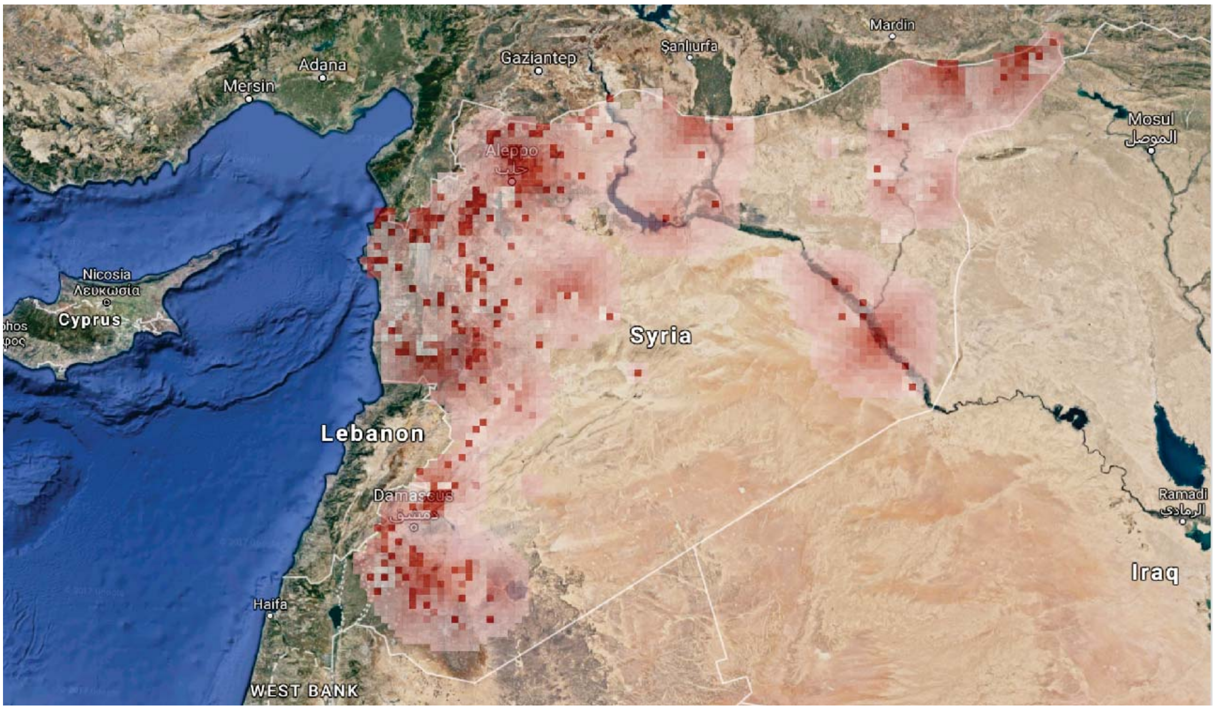


FIGURE A.14: *Experiment 14: The conflict modelled with 0.7 probability of spread and 0.5 probability of depletion.*

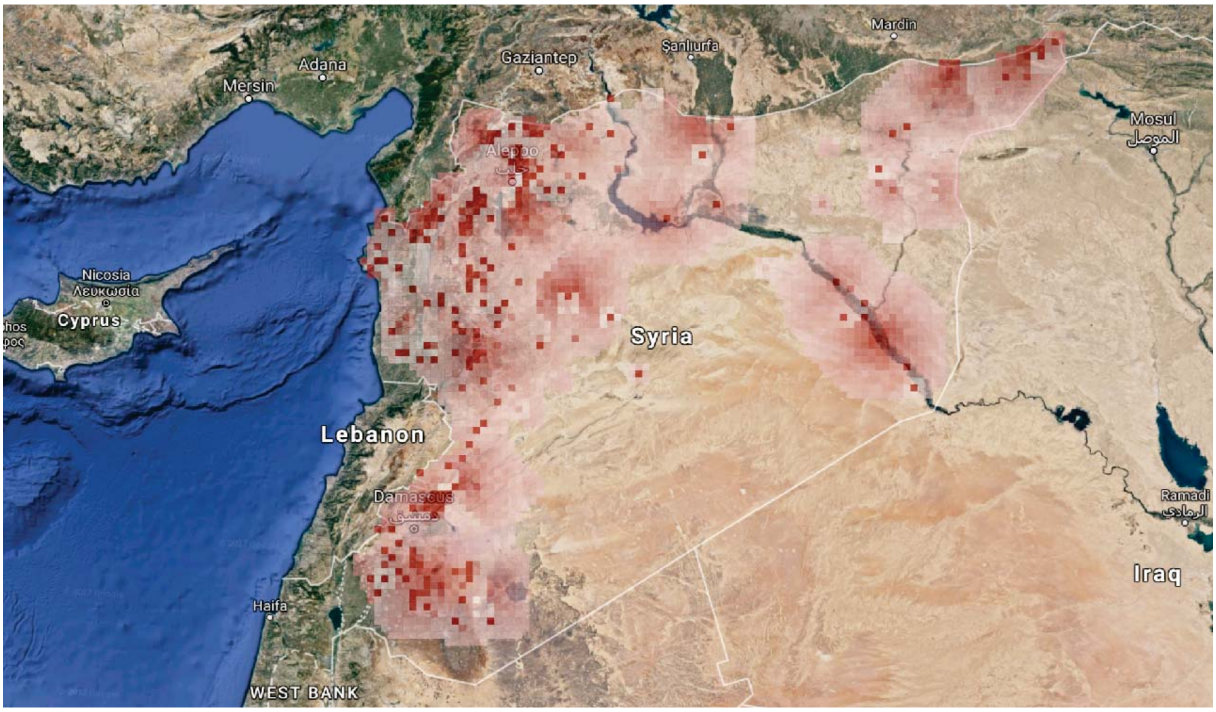


FIGURE A.15: *Experiment 15: The conflict modelled with 0.9 probability of spread and 0.5 probability of depletion.*

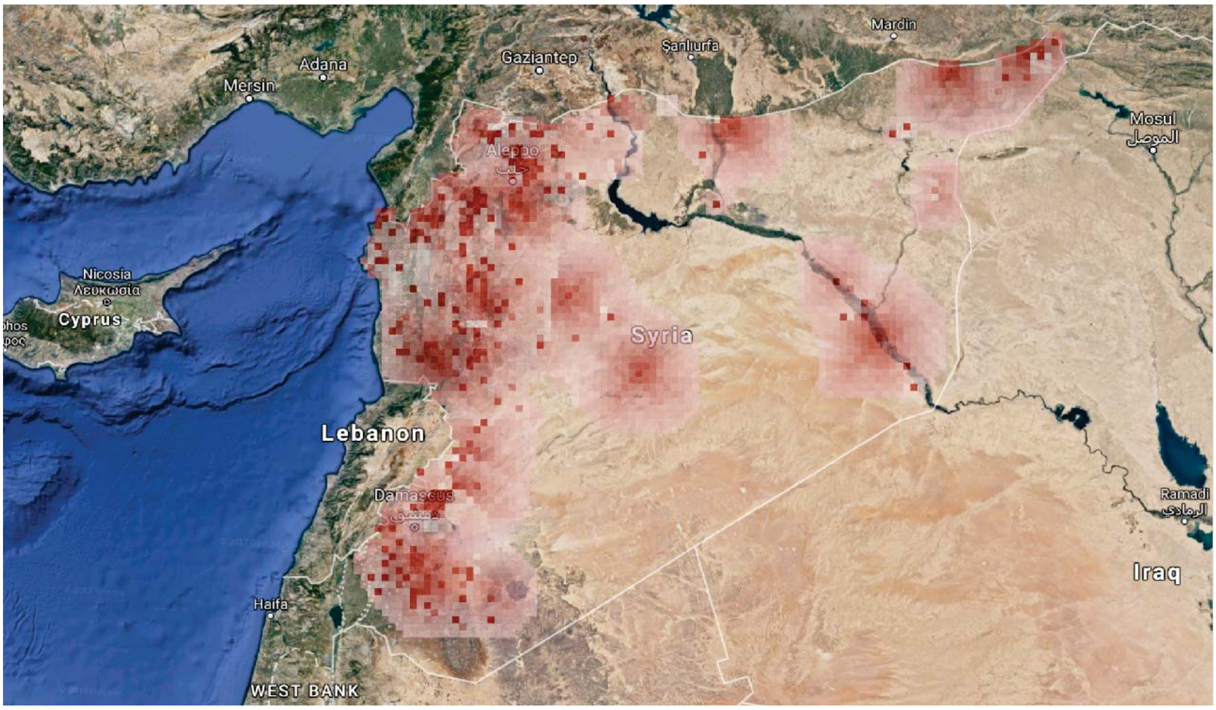


FIGURE A.16: *Experiment 16: The conflict modelled with 0.1 probability of spread and 0.7 probability of depletion.*

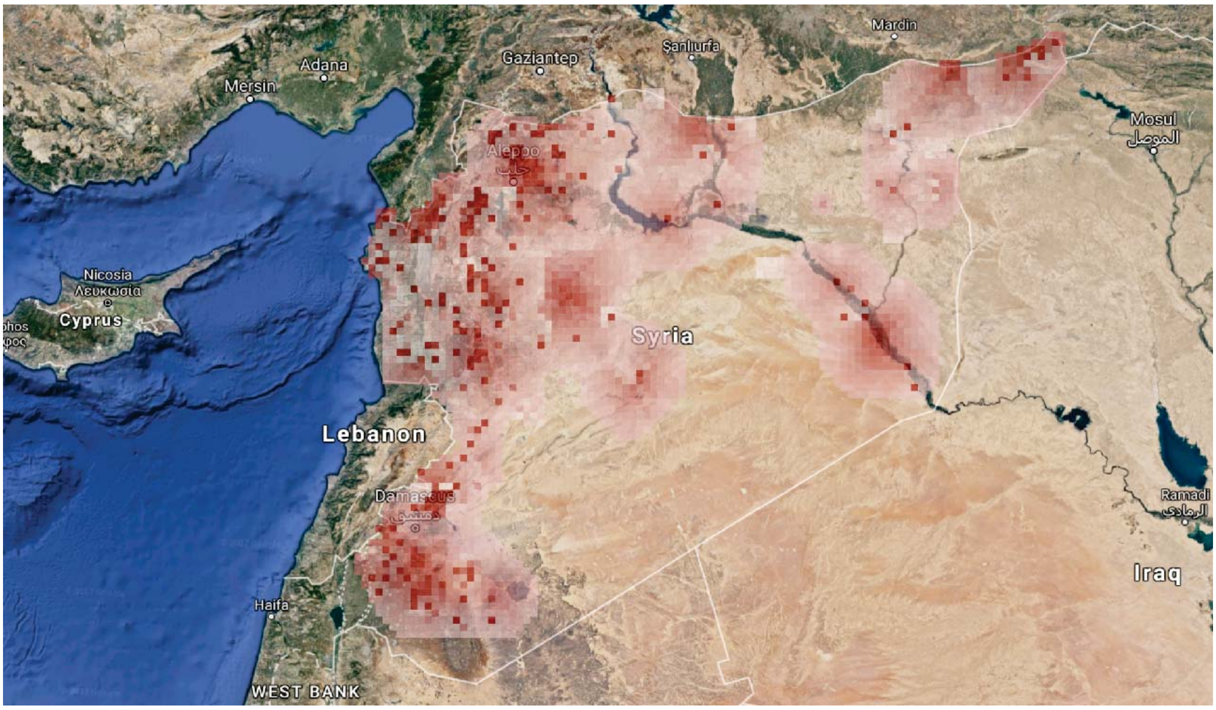


FIGURE A.17: *Experiment 17: The conflict modelled with 0.3 probability of spread and 0.7 probability of depletion.*

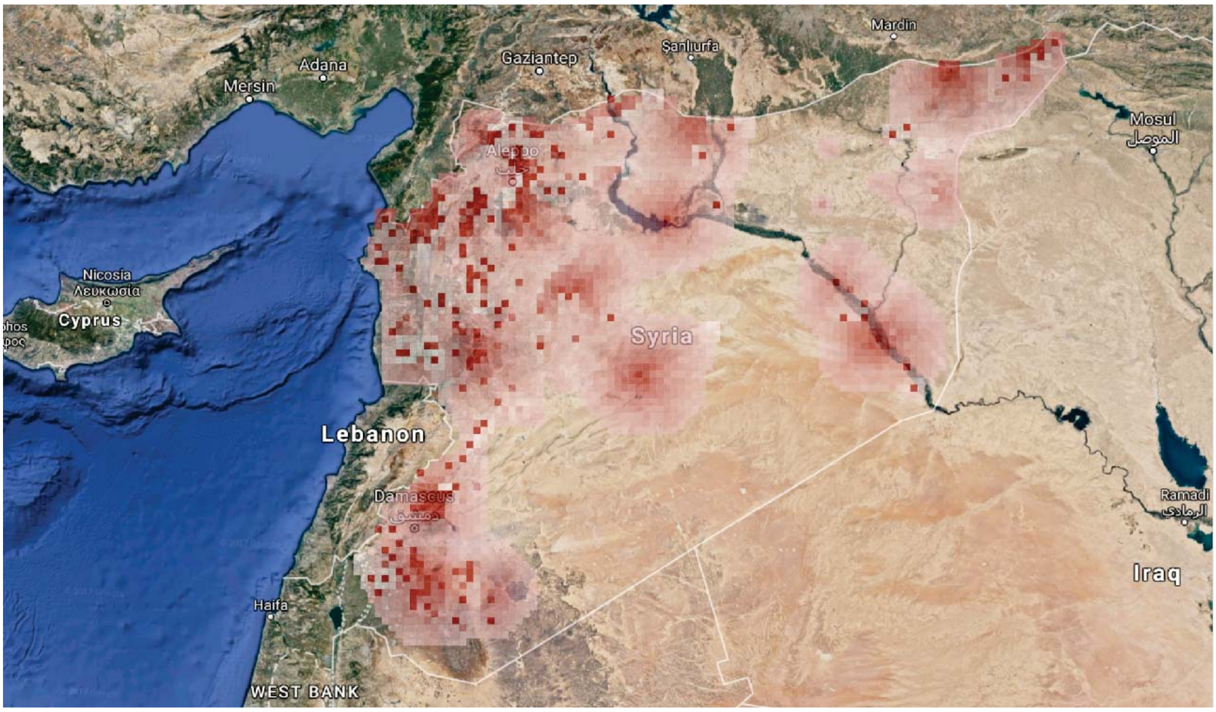


FIGURE A.18: *Experiment 18: The conflict modelled with 0.5 probability of spread and 0.7 probability of depletion.*

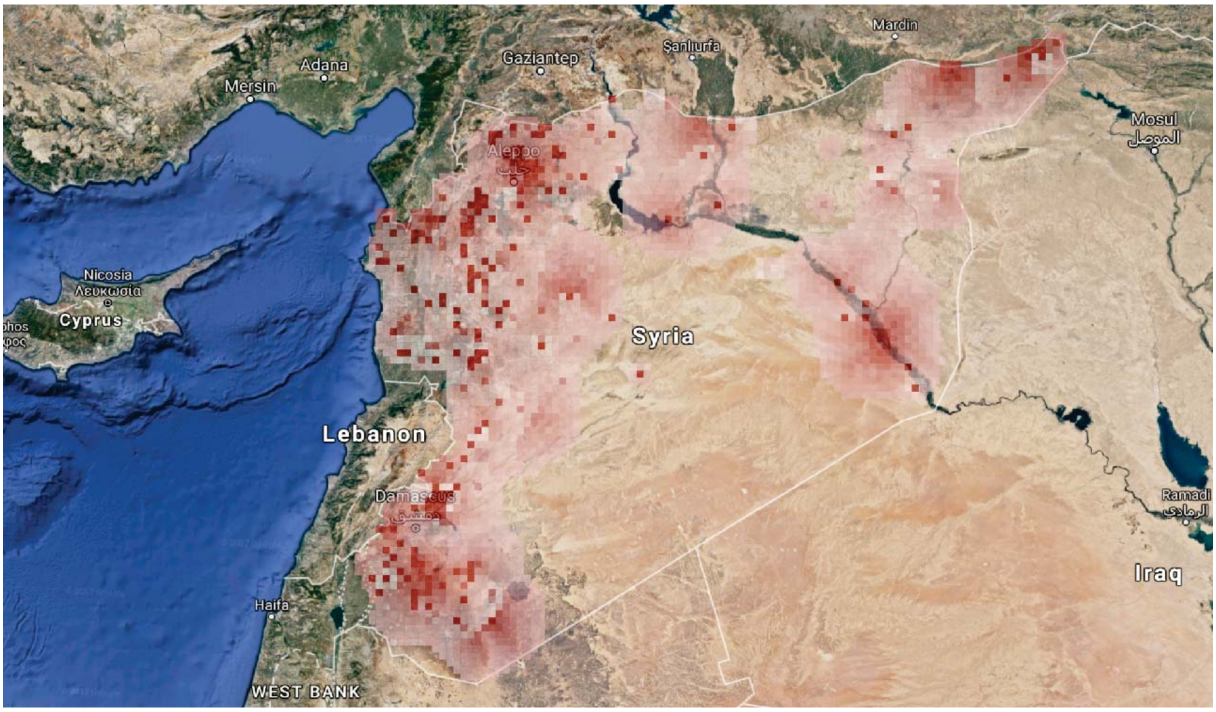


FIGURE A.19: *Experiment 19: The conflict modelled with 0.7 probability of spread and 0.7 probability of depletion.*

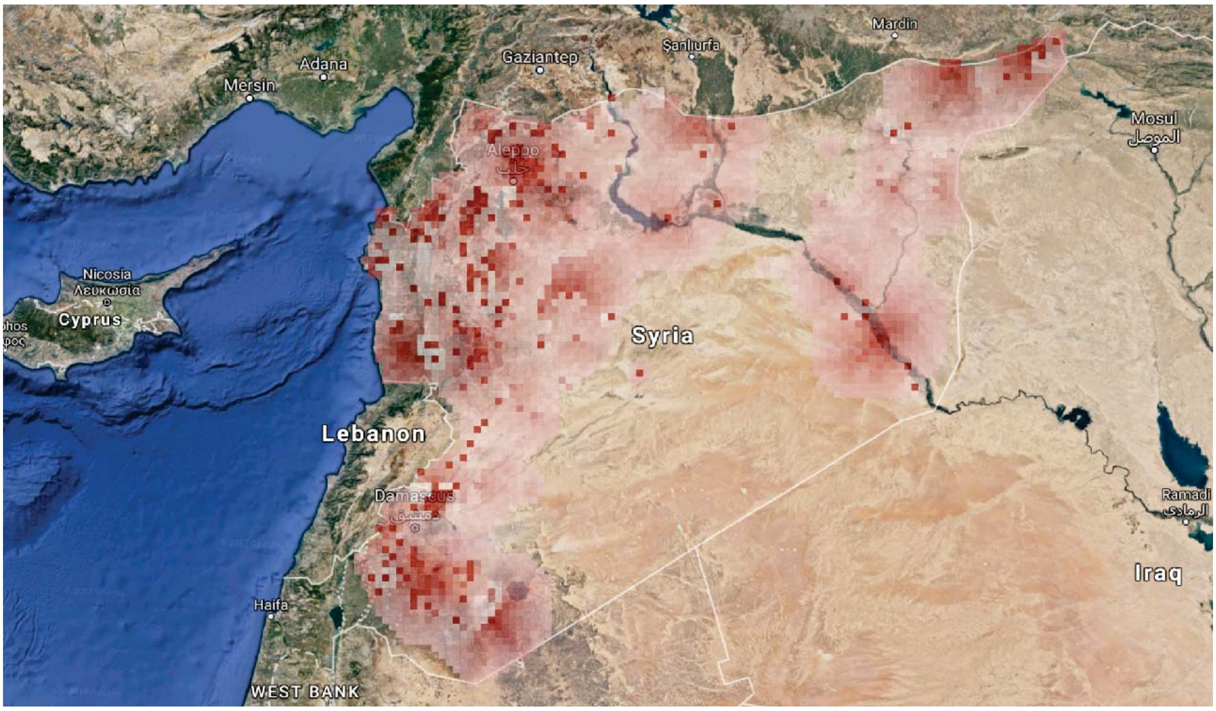


FIGURE A.20: *Experiment 20: The conflict modelled with 0.9 probability of spread and 0.7 probability of depletion.*

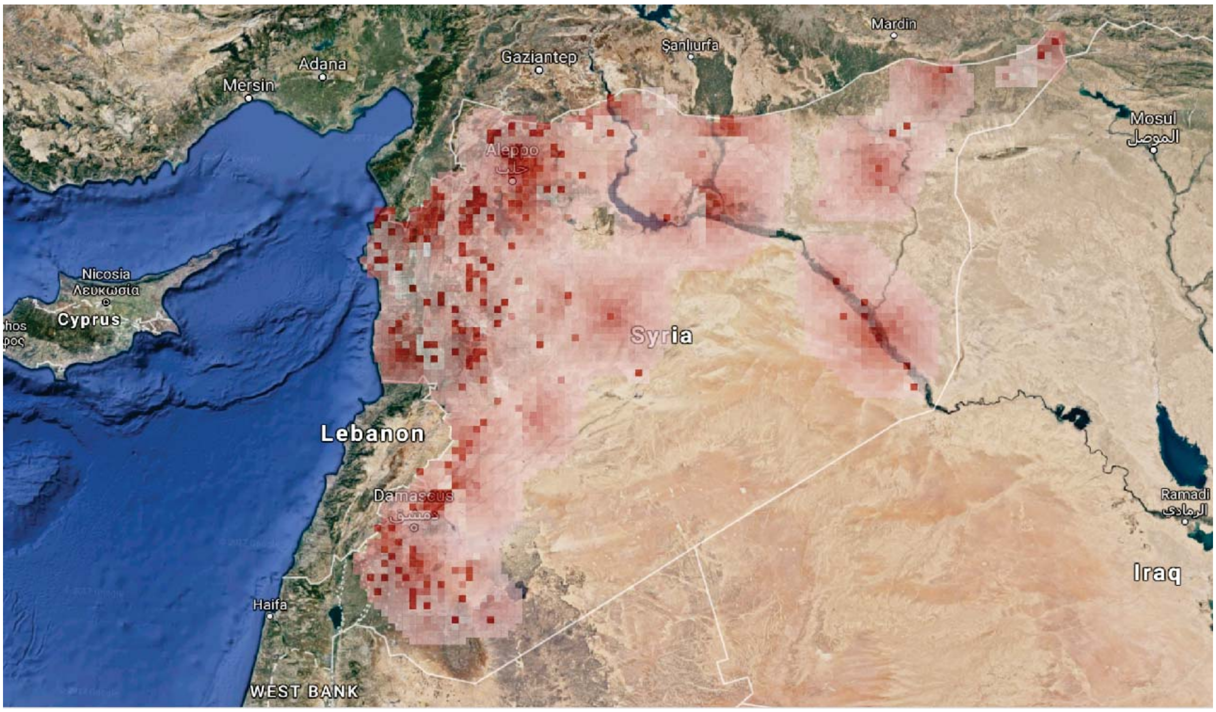


FIGURE A.21: *Experiment 21: The conflict modelled with 0.1 probability of spread and 0.9 probability of depletion.*

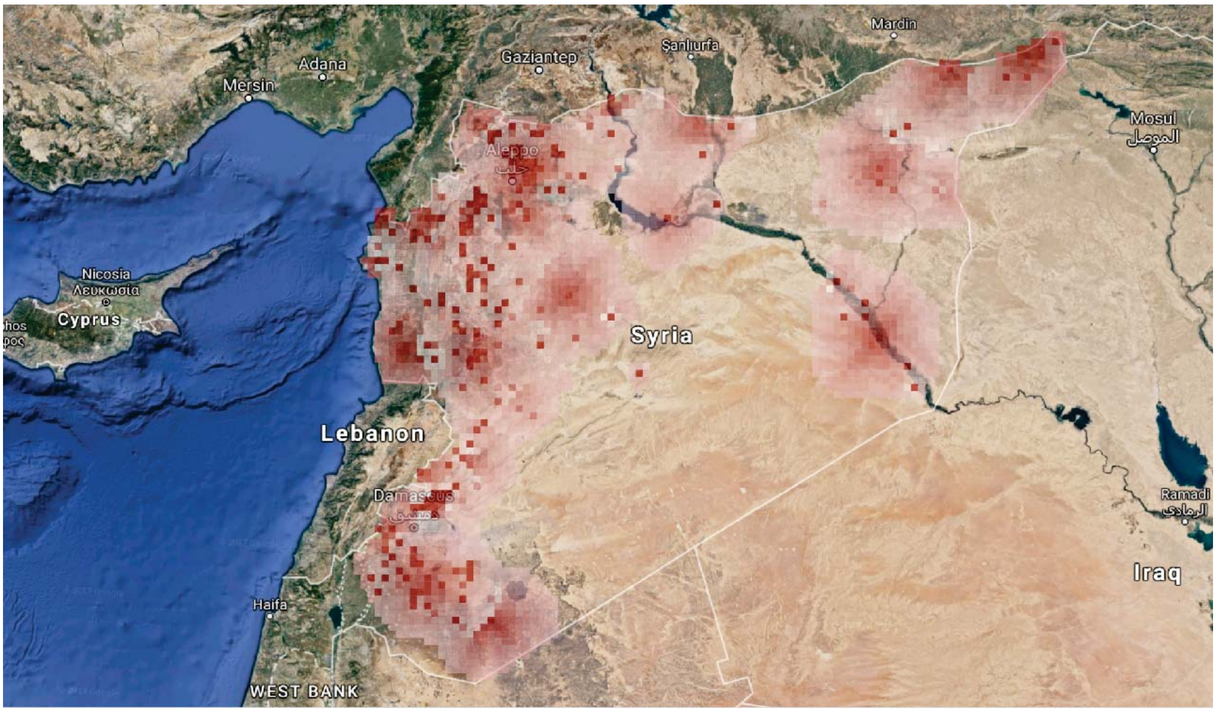


Figure A.22: *Experiment 22: The conflict modelled with 0.3 probability of spread and 0.9 probability of depletion.*

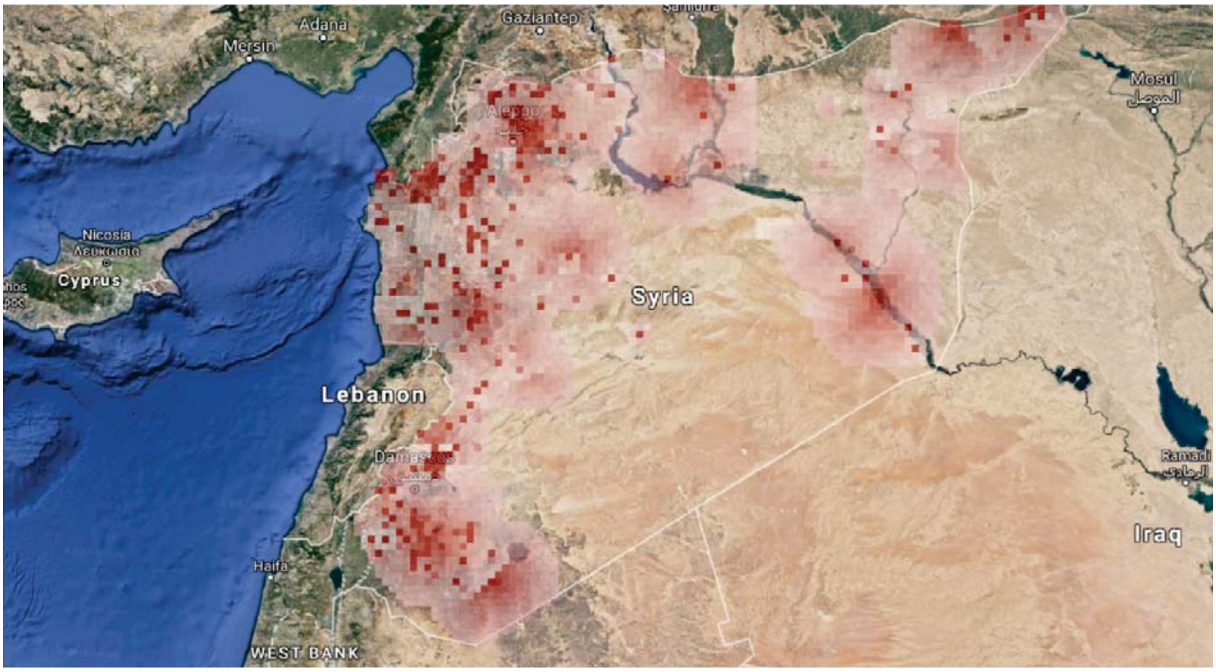


Figure A.23: *Experiment 23: The conflict modelled with 0.5 probability of spread and 0.9 probability of depletion.*

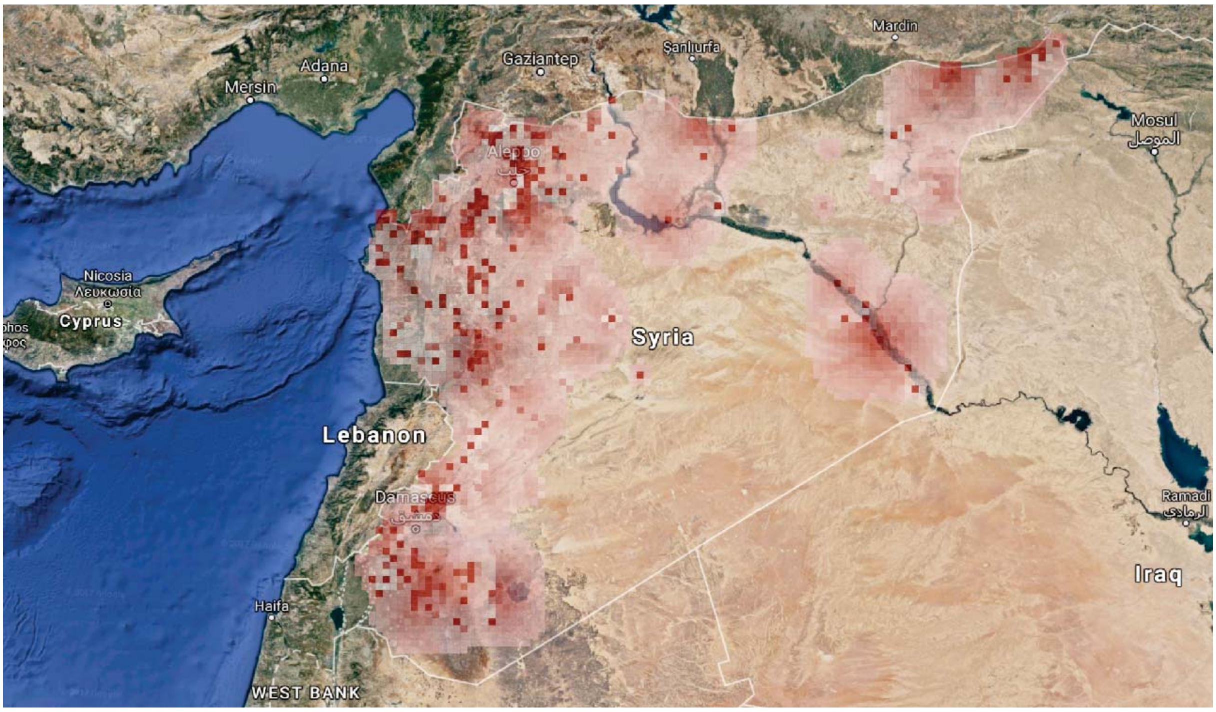


FIGURE A.24: *Experiment 24: The conflict modelled with 0.7 probability of spread and 0.9 probability of depletion.*

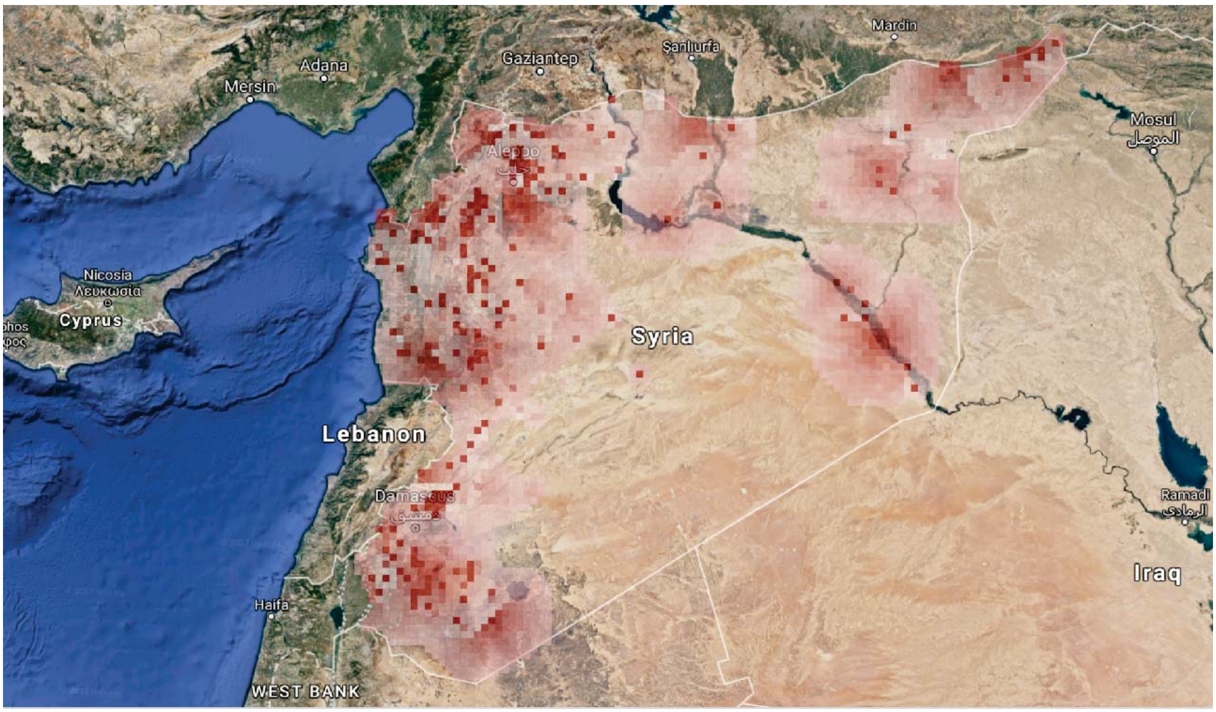


FIGURE A.25: *Experiment 25: The conflict modelled with 0.9 probability of spread and 0.9 probability of depletion.*

APPENDIX B

Probability matrices of experiments

The probability matrices given correlates to the experiments performed in §7.5, where E_i refers to the i^{th} experiment.

$$E_1 = \begin{array}{l} \\ 0 < age < 15 \\ 15 \leq age < 65 \\ age \geq 65 \\ TertiaryEducation \\ NoTertiaryEducation \\ LowEconomicStatus \\ MediumEconomicStatus \\ HighEconomicStatus \\ InternationalFamily \\ NoInternationalFamily \end{array} \begin{array}{c} \begin{array}{ccc} MovementType1 & MovementType2 & MovementType3 \end{array} \\ \begin{pmatrix} 0.25 & 0.40 & 0.35 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.10 & 0.25 \\ 0.20 & 0.20 & 0.60 \\ 0.65 & 0.20 & 0.15 \\ 0.75 & 0.10 & 0.15 \\ 0.20 & 0.45 & 0.35 \\ 0.05 & 0.30 & 0.65 \\ 0.15 & 0.20 & 0.65 \\ 0.85 & 0.05 & 0.10 \end{pmatrix} \end{array}$$

$$E_2 = \begin{array}{l} \\ 0 < age < 15 \\ 15 \leq age < 65 \\ age \geq 65 \\ TertiaryEducation \\ NoTertiaryEducation \\ LowEconomicStatus \\ MediumEconomicStatus \\ HighEconomicStatus \\ InternationalFamily \\ NoInternationalFamily \end{array} \begin{array}{c} \begin{array}{ccc} MovementType1 & MovementType2 & MovementType3 \end{array} \\ \begin{pmatrix} 0.25 & 0.40 & 0.35 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.10 & 0.25 \\ 0.20 & 0.20 & 0.60 \\ 0.65 & 0.20 & 0.15 \\ 0.75 & 0.15 & 0.10 \\ 0.15 & 0.50 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.15 & 0.20 & 0.65 \\ 0.85 & 0.05 & 0.10 \end{pmatrix} \end{array}$$

$E_3 =$

	$MovementType1$	$MovementType2$	$MovementType3$
$0 < age < 15$	0.25	0.45	0.30
$15 \leq age < 65$	0.10	0.35	0.55
$age \geq 65$	0.65	0.20	0.15
$TertiaryEducation$	0.15	0.25	0.60
$NoTertiaryEducation$	0.65	0.20	0.15
$LowEconomicStatus$	0.75	0.15	0.10
$MediumEconomicStatus$	0.15	0.50	0.35
$HighEconomicStatus$	0.05	0.15	0.80
$InternationalFamily$	0.10	0.20	0.70
$NoInternationalFamily$	0.85	0.10	0.05

$E_4 =$

	$MovementType1$	$MovementType2$	$MovementType3$
$0 < age < 15$	0.25	0.45	0.30
$15 \leq age < 65$	0.10	0.35	0.55
$age \geq 65$	0.65	0.20	0.15
$TertiaryEducation$	0.15	0.25	0.60
$NoTertiaryEducation$	0.65	0.25	0.10
$LowEconomicStatus$	0.80	0.15	0.05
$MediumEconomicStatus$	0.10	0.55	0.35
$HighEconomicStatus$	0.05	0.15	0.80
$InternationalFamily$	0.05	0.25	0.70
$NoInternationalFamily$	0.85	0.10	0.05

$E_5 =$

	$MovementType1$	$MovementType2$	$MovementType3$
$0 < age < 15$	0.20	0.50	0.30
$15 \leq age < 65$	0.05	0.40	0.55
$age \geq 65$	0.65	0.20	0.15
$TertiaryEducation$	0.10	0.30	0.60
$NoTertiaryEducation$	0.65	0.25	0.10
$LowEconomicStatus$	0.80	0.15	0.05
$MediumEconomicStatus$	0.10	0.55	0.35
$HighEconomicStatus$	0.05	0.15	0.80
$InternationalFamily$	0.05	0.25	0.70
$NoInternationalFamily$	0.85	0.10	0.05

$E_6 =$

	$MovementType1$	$MovementType2$	$MovementType3$
$0 < age < 15$	0.15	0.55	0.30
$15 \leq age < 65$	0.05	0.40	0.55
$age \geq 65$	0.65	0.20	0.15
$TertiaryEducation$	0.10	0.30	0.60
$NoTertiaryEducation$	0.65	0.25	0.10
$LowEconomicStatus$	0.80	0.15	0.05
$MediumEconomicStatus$	0.05	0.60	0.35
$HighEconomicStatus$	0.05	0.15	0.80
$InternationalFamily$	0.05	0.25	0.70
$NoInternationalFamily$	0.85	0.10	0.05

$$E_7 = \begin{array}{l} \\ 0 < age < 15 \\ 15 \leq age < 65 \\ age \geq 65 \\ TertiaryEducation \\ NoTertiaryEducation \\ LowEconomicStatus \\ MediumEconomicStatus \\ HighEconomicStatus \\ InternationalFamily \\ NoInternationalFamily \end{array} \begin{array}{c} \begin{array}{ccc} MovementType1 & MovementType2 & MovementType3 \end{array} \\ \begin{pmatrix} 0.15 & 0.55 & 0.30 \\ 0.05 & 0.40 & 0.55 \\ 0.65 & 0.20 & 0.15 \\ 0.10 & 0.30 & 0.60 \\ 0.65 & 0.25 & 0.10 \\ 0.80 & 0.15 & 0.05 \\ 0.05 & 0.60 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.05 & 0.25 & 0.70 \\ 0.75 & 0.20 & 0.05 \end{pmatrix} \end{array}$$

APPENDIX C

Graphical model outputs

Figures C.1– C.5 shows the output of the simulation model when employing the parameters estimated as best fit, as determined in §7.5.

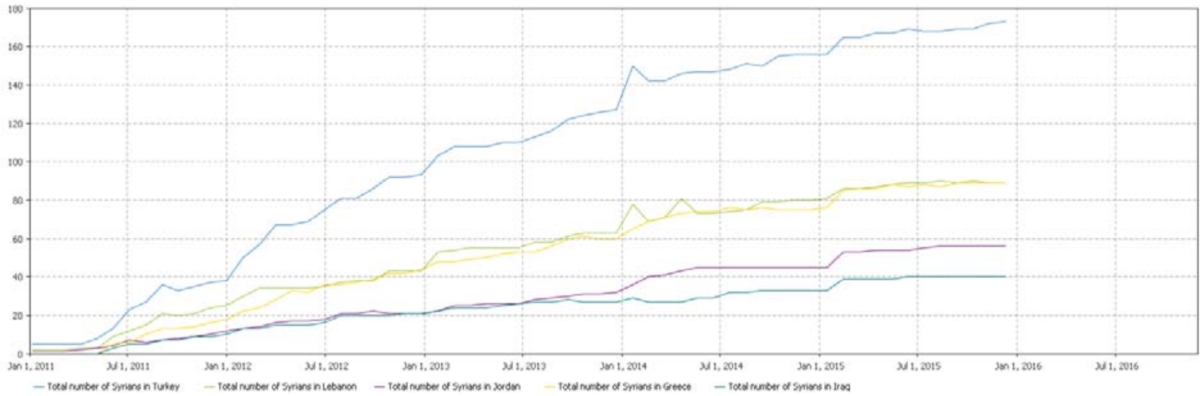


Figure C.1: *The total number of Syrians per neighbouring country as simulated.*

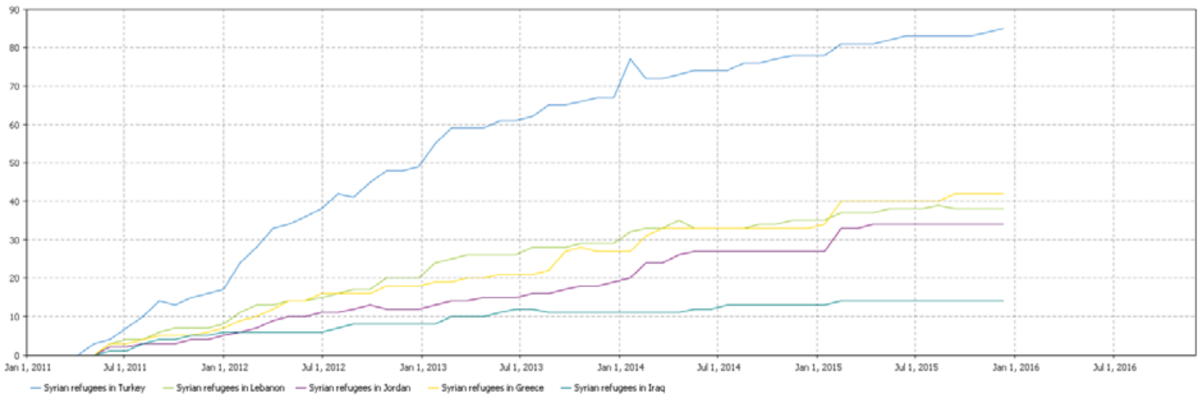


Figure C.2: *The number of Syrians refugees per neighbouring country as simulated.*

FIGURE C.3: *The number of undocumented Syrians per neighbouring country as simulated.*

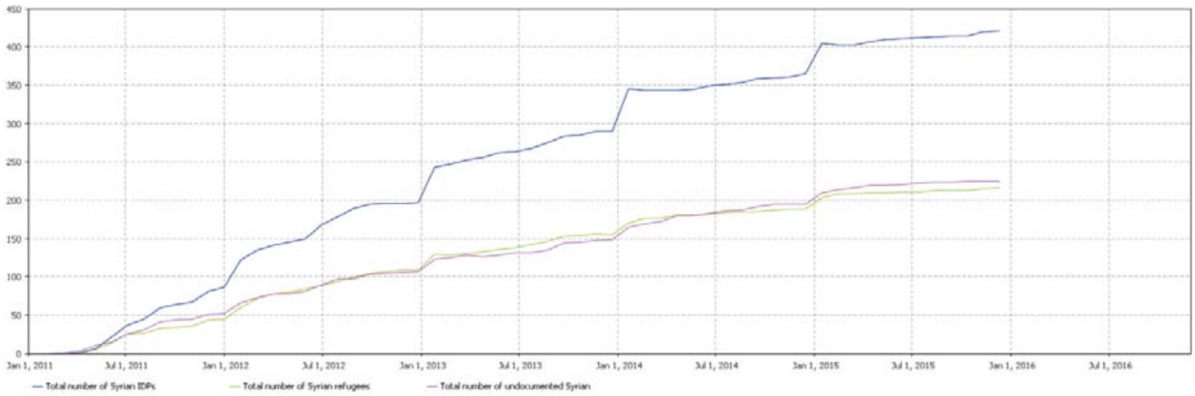


FIGURE C.4: *The total number of Syrians per movement type as simulated.*

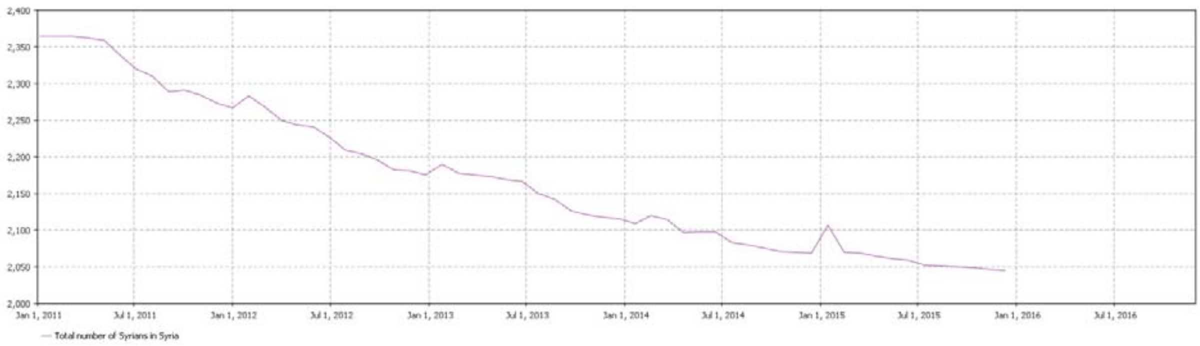


FIGURE C.5: *The total number of Syrians as simulated to remain in Syria.*

APPENDIX D

Contents of the accompanying compact disc

This appendix contains a brief description of the compact disc included with this thesis. The compact disc contains an electronic version of the thesis itself in ".pdf" format, as well as the AnyLogic project file of the final, revised agent-based model, described in Chapters 5, 6 and 7. The model was created in AnyLogic version 7.3.6 and may be executed in this version of the software or later. There are two directories on the compact disc and their contents are described here by their directory names.

Thesis. This directory contains an electronic copy of this thesis in ".pdf" format.

ThesisSimulation. This directory contains the complete agent-based simulation model described in Chapters 5, 6 and 7 as AnyLogic script files (".alp" format). The simulation model is labelled "ThesisSimulation.alp". To execute this simulation model, the file should be opened from AnyLogic. Once opened, the user is required to run the model by either clicking the "Run" button in AnyLogic, or by pressing F5. Following this, the window shown in Figure 5.14 will appear from which the user can either alter the default parameter values or run the model by clicking the "Run" button.