A framework for modelling conflict-induced forced migration according to an agent-based approach

by

Christa de Kock

Dissertation presented for the degree of
Doctor of Philosophy in Industrial Engineering
in the Faculty of Engineering at Stellenbosch University

Promoter: Dr Brian J van Vuuren
Co-promoter: Prof Jan H van Vuuren

December 2019

Declaration

By submitting this dissertation 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: December 1, 2019

Dedication

To my dearest grandparents – the late Dr Awie Cronjé, Babs Cronjé,
Ken de Kock and Christa de Kock

But from everlasting to everlasting the Lord's love is with those who fear him, and his righteousness with their children's children. Psalm 103:17–18

Abstract

Over the course of the past decade, numerous calamities worldwide have led to phrases such as 'refugee' and 'undocumented migrant' becoming commonplace in the public discourse. Conflict-induced displacement and the various challenges it creates have received notable attention. A particular challenge posed by the management of sudden migration of large groups of people lies in a general inability to predict the scale and dynamics of such movement accurately. This problem is further complicated by the fact that data associated with such migration are largely incomplete or untrustworthy.

Presently, there is a significant lack of data required to perform strategic, long-term planning related to current and future conflict crisis situations. One of the most fundamental challenges faced by researchers and humanitarian aid organisations when addressing forced displacement is an inability to predict the types of movement and the destinations of those who are forcibly displaced. The provision of a reasonably accurate estimation of the number of forcibly displaced people is potentially a critical input for the planning of logistics and management of structures supporting those fleeing violence and persecution.

A framework is proposed in this dissertation for assisting in the development and application of agent-based simulation models for predicting conflict-induced migration. The framework comprises five phases which encapsulate the formulation, conceptualisation and development of such a model, as well as the associated model execution and documentation. The purpose of the framework is to facilitate the design and development of agent-based models that incorporate the determinant factors of localised decision making and generate the resulting emergent large-scale movement patterns of forced migration.

Collaboration with various subject matter experts throughout the development of this framework allowed for significant insight to be gained from the confluence of research in the fields of forced migration, computer simulation and human decision-making processes, which does not presently appear in the literature. The approach suggested for modelling human decision making is endorsed by knowledge gained from this research confluence, which has been corroborated by expert opinion. To the best of the knowledge of the consulted subject matter experts, no such framework encompassing such a wide variety of factors and implications pertaining to forced migration modelling in the presence of conflict presently exists and, as such, the research has sparked significant interest in the international research community.

A concept demonstrator is furthermore developed according to the proposed framework in an attempt to demonstrate its usefulness and practicability in the context of conflict-induced migration in Syria. The model concept demonstrator is developed in the AnyLogic simulation software environment and allows for an animated output visualisation of the state of conflict pertaining to specified geographic and time scales, with super-imposed agent movement based on localised decision making when confronted with conflict. As per the framework, the model concept demonstrator is subjected to a number of traditional verification and validation techniques which include the calibration of parameters related to the modelling of conflict, the replication of visualised recorded data, the validation of the relevant agent aggregation, a thorough face validation and a parameter variation analysis which, ultimately, facilitates the implementation of a graphical user interface. The model concept demonstrator is thereby deemed capable of modelling specified scenarios of conflict-induced migration when equipped with the correct parameter values, owing to its flexibility. The animation output also allows for easy interpretation of the model output, particularly by parties who are not necessarily scientifically trained.

A framework of this nature naturally presents numerous avenues for future work. By employing machine learning and data mining tools, such follow-up work may, for example, ultimately lead to a framework for assisting in planning and the formulation of logistics strategies in the lieu of identifying required facilities and resources. Such an enhanced framework may prove invaluable in the accommodation of incoming refugees, internally displaced migrants and undocumented migrants in different areas, by predicting the population fluctuations in affected areas during times of conflict, natural disaster, or other forced migration-causing events.

Uittreksel

Met die verloop van die laaste dekade het 'n verskeidenheid rampe wêreldwyd 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. 'n Besondere uitdaging wat met die bestuur van skielike migrasie van groot groepe mense gepaard gaan, lê in 'n algemene onvermoë om die omvang en dinamiek van hierdie beweging akkuraat te voorspel. Hierdie probleem word verder bemoeilik deurdat inligting wat verband hou met sulke migrasie grootliks onvolledig of onbetroubaar is.

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

In hierdie proefskrif word 'n raamwerk voorgestel vir die ontwikkeling en toepassing van agent-baseerde simulasiemodelle wat daarop gemik is om konflik-geïnduseerde migrasie te modelleer. Die raamwerk bestaan uit vyf kernfases wat die formulering, konseptualisering en ontwikkelling van 'n model behels, sowel as die gepaardgaande modeluitvoering en dokumenteringsproses. Die raamwerk is daarop gemik om die ontwerp en ontwikkelling van agent-baseerde modelle te fasiliteer wat die bepalende faktore van gelokaliseerde besluitneming inkorpereer en die gevolglike ontluikende grootskaalse bewegingspatrone van gedwonge migrasie vasvang.

Samewerking met verskeie vakkundiges gedurende die ontwikkeling van hierdie raamwerk het deur die samevloei van navorsing oor gedwonge migrasie, rekenaarsimulasie en menslike besluitnemingsprosesse tot noemenswaardige insig gelei wat nie in die huidige literatuur vervat is nie. Die benadering wat vir die modellering van menslike besluitneming voorgestel word, word deur kennis wat uit hierdie navorsingsamevloei verkry is, ondersteun en is deur vakkundiges bevestig. Sover die kennis van die gekonsulteerde vakkundiges strek, bestaan daar tans geen raamwerk wat so 'n groot verskeidenheid onderliggende faktore en implikasies van vlugteling-modellering in konflik-geteisterde gebiede in ag neem nie. Die navorsing het daarom beduidende belangstelling in die internasionale navorsingsgemeenskap ontlok.

'n Konsepdemonstratormodel word verder deur middel van die voorgestelde raamwerk ontwikkel in 'n poging om die bruikbaarheid en uitvoerbaarheid daarvan in die konteks van konflik-geïnduseerde migrasie in Sirië te demonstreer. Die konsepdemonstratormodel is in die AnyLogic simulasiesagteware-omgewing ontwikkel en is daartoe in staat om geanimeerde afvoer te lewer wat die toestand van konflik oor gespesifiseerde geografiese- en tydskale visualiseer, tesame met gesuperponeerde bewegingspatrone gebaseer op gelokaliseerde besluitneming van agente wat as gevolg van geweld verplaas word. As deel van die raamwerk word die konsepdemon-

stratormodel aan 'n aantal tradisionele verifikasie- en valideringstegnieke onderwerp, insluitend die kalibrasie van parameters wat verband hou met die modellering van konflik, die replikasie van opgetekende visuele data, die validering van die relevante agent-aggregasie, 'n deeglike sigvalidering en 'n parameter variasie-analise wat uiteindelik die implementering van 'n grafiese gebruikerskoppelvlak fasiliteer. Vanweë die buigsaamheid daarvan het die konsepdemonstratormodel die vermoë om gespesifiseerde scenarios van konflik-geïnduseerde migrasie te modelleer indien dit van die toepaslike parameterwaardes voorsien word. Die animasie-afvoer vergemaklik ook die interpretasie van die modelafvoer, veral deur diegene wat nie noodwendig 'n wetenskaplike agtergrond het nie.

'n Raamwerk van hierdie aard bied natuurlik talle moontlikhede vir toekomstige opvolgwerk. Deur gebruikmaking van masjienleer en data-analitiese tegnieke kan sodanige opvolgwerk byvoorbeeld uiteindelik lei na 'n raamwerk wat tydens beplanning en die formulering van logistieke strategieë gebruik kan word om benodigde fasiliteite en hulpbronne te identifiseer. So 'n aangepaste raamwerk mag deurslaggewend wees in die akkommodasie van inkomende vlugtelinge, plaaslik ontworteldes en ongedokumenteerde migrante in verskillende gebiede, asook om die bevolkingsvariasie in geteisterde gebiede tydens tye van konflik, natuurrampe of ander vlugtelingeveroorsakende gebeurtenisse te voorspel.

Acknowledgements

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

- Dr Brian van Vuuren and Prof Jan van Vuuren, for taking me in and giving me the freedom to grow academically, you are both extraordinary supervisors and I am grateful for having had the opportunity to draw from your knowledge and experience;
- 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;
- The Department of Industrial Engineering and, in particular, the *Stellenbosch Unit for Operations Research and Engineering* (SUnORE), for “tea times,” for thought-provoking conversations and for introducing me to Kit Daniel Searle;
- My family, for always encouraging and loving me without having to understand the particulars; and
- God, My Father, for Your goodness and favour are evident in my life.

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

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 organisation 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 in 2017 to be 68.5 million, 40 million of whom were people displaced within their countries of origin, and 28.5 million as refugees fleeing across international borders [169, 176].

The *Global Opportunity Report 2017* [63] identified several topics pertinent to global unstable regions as some of the most important opportunities to pursue in 2017 with respect to social impact. The gross impact of persecution and violent conflicts is 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, and 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 to managing or curbing such realities [63].

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 affect civilians, governments, international humanitarian organisations, as well as global governance [52]. Aiyar *et al.* [2] documented the economic challenges faced by Europe accompanying the influx of refugees, whilst Richmond [138] discussed forced migration from a sociological perspective, identifying determinant factors other than politics which influence forced migration.

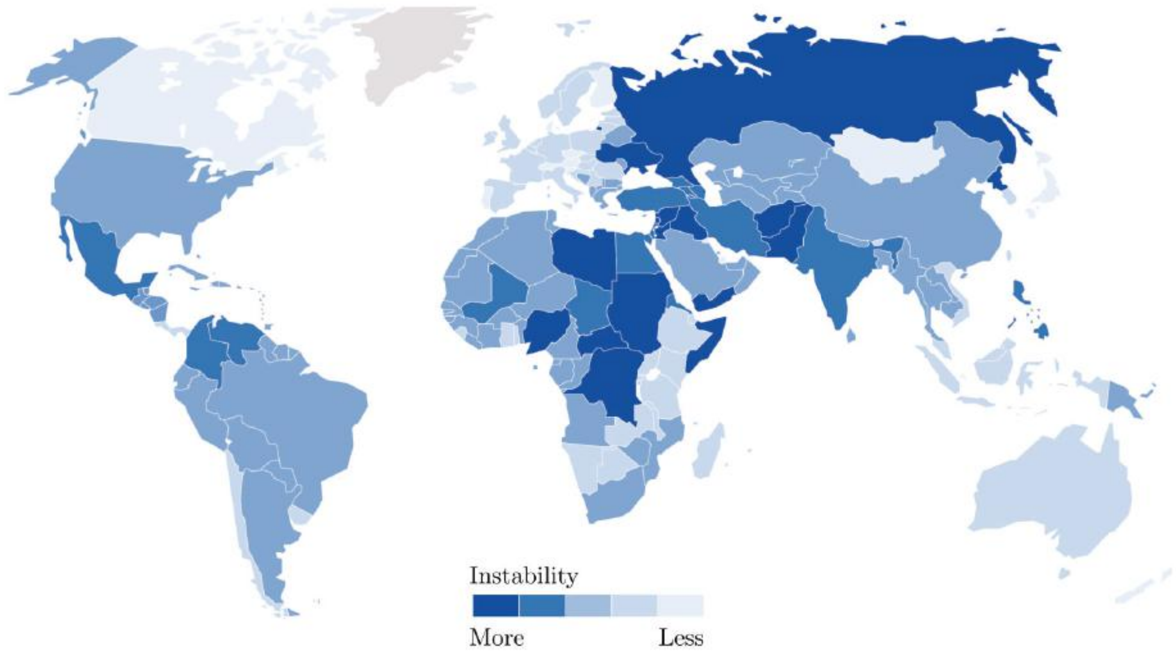


FIGURE 1.1: *A map indicating the global instability in 2015 [63, 79].*

The potential impact of humanitarian intervention efforts in these crises bring about many challenges that require innovative problem-solving approaches. Researchers such as Alhanaee and Csala [6] and Greenwood [65] have attempted to identify the motives of migrants in order to better understand the spatial phenomenon of migration in economic and social contexts. 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 aimed at facilitating their arrival [73]. The lack of adequate and complete data presents a serious problem to humanitarian aid, especially with respect to forced displacement. Efforts towards improving the reliability, quality and scope of data concerning forcibly displaced people are required to address the gaps in data currently available for assisting in long-term development planning during crises [144].

Edwards [47] discussed the potential of using computational models in order to predict the spatial patterns of forcibly displaced individuals, emphasising the assistance that such models may be able to provide humanitarian aid organisations and policy makers. Groen [68] affirmed the importance of utilising such models in order to capture the movements of refugees on a global scale. The notion of numerical modelling is a tool utilised to study the behaviour of large complex systems and, in particular, their dynamic behaviour, when the complexity does not allow for analytical evaluation [153]. One of the most powerful tools available for comprehending the behaviour of complex systems and processes is simulation modelling [147].

Simulation modelling is the computer-based imitation of a real-world system [159]. 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 that guide decision making [20]. *Agent-based modelling* (ABM) allows for the consequences of individual decisions to be modelled locally, taking into account the complexities of social systems such as behaviours, motivations and relationships between individual agents, and provides an aggregated global emergent behaviour as output [8]. 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 [89].

ABM is capable of facilitating a synthetic environment which allows for an understanding of the collective behaviour of forcibly displaced persons through computational experimentation [8]. With knowledge pertaining to the migration behaviour of people, an agent-based model may potentially 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 region. Furthermore, such a model may aid in identifying appropriate locations for aid distribution points, as well as predicting the anticipated demand for aid [47].

Simulation models of this nature presently prevailing in the literature include, amongst others, a model developed by Lemos *et al.* [104] for simulating social conflict and civil violence aimed at capturing the characteristics associated with its spread, a model developed by Crooks and Wise [37] for aiding humanitarian relief after the occurrence of natural disasters, a model simulating the autonomous decision making of environmentally induced migrants developed by Smith *et al.* [155], and a model explicitly incorporating the social network among people migrating between Ecuador and the United States, developed by Rehm [137].

Anderson *et al.* [8] 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 the implementation of policies allows for the evaluation of the potential impact of decisions and the testing of various strategies. Collins and Frydenlund [35] further proposed an agent-based model for simulating strategic group formation of refugees when fleeing. This model is aimed at investigating the evacuation time of refugees, assuming that they tend to form groups when travelling over long distances. Another agent-based model, developed by Orfano [126], simulates empirical economic factors leading to migration and the long-term effects of forced migration.

Johnson *et al.* [89] calibrated an agent-based model for use in the context 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 situations considered with the help of subject matter experts. The spread of disease in refugee camps, another pertinent issue faced by humanitarian agencies at refugee camps, was modelled by Hailegiorgis and Crooks [72], 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.* [156] as a method of investigating the decision making of citizens during forced migration.

Klabunde and Willekens [93] 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 exist, they differ in scale, complexity and documentation 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 a general lack of empirical data, and effectively modelling the manner in which migration decisions are influenced by a human's life course.

Awareness of crises and, in particular, conflict-induced forced displacement, has increased 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. Currently, there are significant gaps in the data required to facilitate long-term development planning in crisis situations. The aggregate number of 68.5 million forcibly displaced people is only an estimate and data concerning those displaced within specific countries of origin are even less reliable [114, 144].

More often than not, research does not focus on those internally displaced or on undocumented migrants, which augments the gap in available data [4, 39]. The unknown number of forcibly displaced people represents an obstacle to humanitarian support organisations, especially in times where the situation is critical in nature [72, 176].

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 main obstacle to predicting such random movement remains the facts that migration is a highly structured process dependent on patterns, historical context and the manner in which an individual's decision-making process develops [44, 47]. The ability to predict the movement of forcibly displaced people with some measure of accuracy is critical to aid organisations in facilitating the planning of logistics and procurement of resources aimed at supporting those fleeing violence and persecution [35]. Groen [68] endorsed 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 view of the aforementioned shortcomings in respect of modelling instances of conflict and the lack of data related to the associated movement patterns of forcibly displaced individuals, a generic framework is proposed in this dissertation for aiding in the design of an agent-based model for simulating conflict instances along with the localised decision-making processes underlying the movement of refugees, undocumented migrants and *internally displaced persons* (IDPs) fleeing conflict-affected areas. The value of such a model lies in the fact that it produces as output the corresponding emergent large-scale migration patterns which may assist in understanding the movement patterns of forcibly displaced people and predicting anticipated destinations of these individuals, and serve as a decision support tool for humanitarian relief.

1.3 Dissertation scope and objectives

In order to facilitate the simulation of forced migration and the subsequent capturing 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 dissertation:

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

(a) historical instances of forced migration and forcibly displacement people,

(b) the factors influencing displacement and the types of movement of forcibly displaced people,

(c) the current state of global conflict-induced forced migration,

(d) computer simulation modelling techniques and, in particular, ABM,

(e) the field of decision making and multi-criteria decision-making methods,

(f) the modelling of human decision making, and

(g) the design process of a generic model design framework.

II To *propose* a generic framework for assisting modellers in the development of an agent-based model in the context of forced migration with the aim of:

(a) gaining an understanding of the numerous determinant factors of forced migrants,

(b) estimating the total number of forcibly displaced persons sufficiently accurately,

(c) estimating the number of refugees, undocumented migrants and IDPs, which compose the forcibly displaced persons group,

(d) predicting population fluctuations within countries affected by a crisis, and

(e) presenting an animation output of a simulation model (*i.e.* the conflict present and the movement of forcibly displaced people) in a user-friendly manner.

III To *implement* the generic framework in the form of a computerised concept demonstrator designed to investigate the movement of forcibly displaced people within a particular conflict-affected area. This model should be capable of:

(a) reading input data pertaining to the initiation of conflict,

(b) allowing the user to manually initiate areas of conflict,

(c) replicating the spread and depletion of conflict in a realistic manner, informed by historical observation,

(d) modelling an aggregated population of agents residing within the conflict-affected area,

(e) capturing an agent's ability to withstand conflict, and

(f) modelling the decision making of agents affected by conflict.

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

V To *recommend* possible future improvements or additions to the generic framework, as well as sensible follow-up work which may stem from this study.

1.4 Research methodology

The methodology followed in this dissertation when designing the generic framework aimed at assisting modellers in developing an agent-based simulation model for investigating the movement patterns of forcibly displaced people, as described in §1.3, is as follows:

I *Consulting* and *reviewing* existing literature on forced migration and, in particular, conflict-induced migration, as well as existing literature in the field of decision making and, in particular, the modelling of human decision making, followed by existing literature on the design of generic frameworks for model design.

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

III *Proposing* the generic framework for guiding the design of an agent-based model which employs the aspects of localised decision making and produces emergent large-scale movement patterns of those forcibly displaced.

IV *Implementing* a computerised concept demonstrator model according to the framework for modelling the movement of forcibly displaced people within a particular conflict-affected region.

V *Verifying* the performance and execution of the concept demonstrator with respect to each aspect modelled.

VI *Validating* the concept demonstrator by means of a face validation with the assistance of subject matter experts.

VII *Calibrating* the concept demonstrator by means of a parameter variation and an analysis of the model output.

VIII *Illustrating* the flexibility, usability and applicability of the concept demonstrator by performing a parameter establishment analysis so as to equip the model in replicating previously encountered conflict situations.

1.5 Dissertation organisation

Apart from this introductory chapter, this dissertation contains ten additional chapters which are partitioned into four separate parts. Part I, comprising Chapters 2–5, encapsulates a comprehensive literature review of the topics stated in Objective I of §1.3. Chapter 2 provides the reader with an understanding of the phenomenon of forced migration, the various factors that result in forced displacement and a typology of forced migration with respect to movement types. A brief history of the notion of refugees is given, after which the current state of conflict-induced forced migration is discussed.

In Chapter 3, information pertaining to the different aspects and considerations of computer simulation modelling is provided. The different types of simulation modelling prevailing in the literature are reviewed and the various levels of model abstraction are discussed. A number of simulation modelling approaches are described, the generic steps typically followed in a simulation study are outlined, and methodologies for the verification and validation of simulation models are reviewed. This is followed by a discussion on ABM, the advantages and disadvantages associated with this simulation modelling approach, and the components of a typical agent-based model. The chapter closes with a review on the application of agent-based modelling with respect to forced displacement, considering the modelling of different movement types, existing migration models for 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 on a discussion on the modelling of human decision making. Thereafter, the modelling of decision making of forcibly displaced individuals is elaborated upon, considering both the choice of a movement type and of a proposed destination of a modelled migrant.

Chapter 5 concludes the literature review with a brief overview of framework design methodologies in the literature. A qualitative approach to developing frameworks, based on the Grounded Theory methodology, is further discussed. The conceptual framework analysis tool, an outcome of this discussion, is outlined.

Part II of this dissertation is devoted to the proposed generic framework, and is the heart of the dissertation. The design of this generic framework for the ABM of forced migration is proposed in Chapter 6, in fulfilment of Objective II. The various phases in the development of an agent-based model capable of simulating conflict instances, along with the localised decision making of an affected population, is described.

Part III, comprising Chapters 7–9, is devoted to a description of a computerised concept demonstrator. The generic framework developed in Chapter 6 is utilised in Chapter 7 to design an agent-based model as concept demonstrator for forced migration in a conflict-affected country, in fulfilment of Objective III. 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 — conflict, the population and individual decision making — are described in some detail, before discussing the graphical user interface created to control these elements.

In Chapter 8, the concept demonstrator is verified, in partial fulfilment of Objective IV. The chapter opens with an explanation of the concept of verification, after which each of the three components modelled is individually verified. The verification process is performed by means of a number of cases tested, analysing the corresponding output in order to determine whether the model is sufficiently accurately developed.

The actual model validation of the concept demonstrator and the analysis of its output are carried out in Chapter 9, in fulfilment of Objective IV. The various concepts pertaining to this analysis are briefly discussed before a calibration is performed of the parameters pertaining to the modelling of conflict. This is followed by a validation of the spread of conflict, as modelled, which includes a comparison of real-world visualised data with the simulated output. Face validation is carried out thereafter, and this is followed by performing a parameter establishment analysis. The chapter closes with a brief discussion on the efficacy of the concept demonstrator as decision support and analysis tool.

The final part of this dissertation, Part IV, comprises Chapters 10 and 11. A short summary of the dissertation and an overview of the contributions made are provided in Chapter 10. Suggestions with respect to future work which may improve or build on the work conducted in this study, or serve as future avenues of investigation, are finally proposed in Chapter 11, in fulfilment of Objective V.

Part I

Literature review

CHAPTER 2

Forced migration

Contents

This chapter contains a discussion on the phenomenon of forced migration by examining various causes of forced displacement and showcasing a typology of forcibly displaced persons. Furthermore, a brief overview of historic refugee movement is given, and this is followed by an overview of current global forced displacement.

2.1 Factors leading to forced displacement

Disasters, either 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 affected areas [130]. The *International Federation of Red Cross and Red Crescent Societies* [77] defines a disaster as a sudden, calamitous event 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 its own resources exclusively [130].

Disastrous events can occur as a result of numerous factors. Green and McGinnis [64] 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* [64].

Natural disasters include any event that results from natural forces in which human intervention is not the primary cause of the forces [64]. Such naturally occurring phenomena, caused either by a rapid or slow onset of events, can be geophysical (earthquakes, tsunamis, volcanic activity), hydrological (avalanches, floods), climatological (droughts, wildfires, extreme temperatures), meteorological (cyclones, storms, wave surges) or biological (disease epidemics, plagues) [77].

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 [64]. 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 [64, 77]. The primary differentiation between human systems failure and conflict-based disasters is the absence or presence, respectively, of intent [64]. 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 assets [64].

Apart from disasters, people may also lose their homes as a result of development. 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, the establishment of industrial parks, infrastructure development or conservation programs [29]. Displacement typically progresses slowly as it usually only occurs after the commencement of a development project [80].

2.2 A typology of forced migration

Migrants are people who leave or flee their homes, relocating either within or across international borders, in search of better or safer surroundings [129]. In the literature, migrants are classified into two distinct groups — those who voluntarily leave their homes and those who are forcibly displaced [11, 80, 129, 138]. *Voluntary migration* often occurs for economic reasons, where a person moves from one place to another in pursuit of better living conditions [44]. *Forced migration*, on the other hand, 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 [138].

Castles [29] reports that forced migration has grown in volume and political significance since the Cold War and has 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 [115].

A diagrammatic illustration of a typology of migration, with an emphasis on forced migration, may be found in Figure 2.1. As indicated, individuals who are forced to relocate, either move within the borders of the country (*i.e.* IDPs) or across its borders (*i.e.* international migrants).

IDPs are forced to leave their places 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 [11, 169]. 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* [80] 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 of people to be less likely to completely abandon their countries of residence as a result of an implicit belief that returning to their home country would then not be possible. Furthermore, immigration is a financially costly exercise [126]. The report also suggests that countries with economic and political instability are more likely to experience a larger number of IDPs resulting from a disaster, especially among people of low-income classes and those subject to socio-economic discrimination [80].

International migrants are migrants who flee their countries of origin owing to natural disasters, human systems failure, conflict-based disasters or persecution. Those fleeing persecution tend

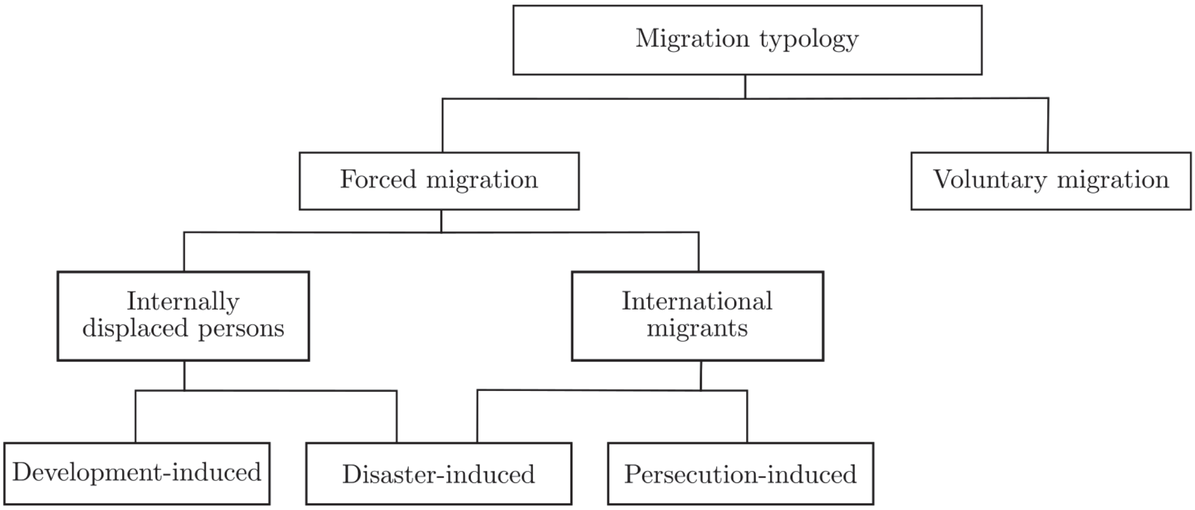


FIGURE 2.1: *A migration typology explaining forced migration.*

to settle indefinitely in a new host country, whereas those who leave their countries for other reasons tend to view the migration as temporary [126]. Individuals who relocate as a result of persecution are explicitly known, under international law, as refugees [167]. 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*" [167]. These individuals are guaranteed, under international law, the right to request asylum in another country.

Asylum-seekers are individuals who have sought international protection in a foreign country and are waiting to find out whether their claims for refugee status have been granted [129, 133]. 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 countries 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 [44].

The options available to a person fleeing persecution are 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 result of his or her application is known. When a persecution-induced migrant chooses not to seek asylum, (s)he becomes an undocumented migrant, thereby living illegally in a foreign country [133]. 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 of a persecution-induced migrant.*

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 [16].

In 2010, the *International Office for Migration* estimated that roughly 20–30 million migrants were undocumented worldwide, which comprises about 10–15% of the global population of migrants [133]. 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 authors, such as Dragostinova [44] and Sarzin [144], informally define refugees as people who attempt to flee their countries due to armed conflict or persecution. Castello [136] 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 [138] 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 [169]. International law, however, has not yet agreed upon a broader definition for refugees [138].

2.3 A brief history of forcibly displaced people

The concept of people being forced to leave their home countries 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 [32]. *The Edict of Fontainebleau* is one of the first recognised reasons for the displacement of people across nations and it is estimated that approximately 200 000 individuals fled France in the two decades that followed its institution [32].

Some time later, during 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 [32]. 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 believed that approximately two million Jews fled Russia, primarily to the United Kingdom, the United States and Europe owing to the Russian anti-Jewish sentiment [32].

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 five million people became refugees as a result of the aforementioned war [25]. As devastating as this refugee crisis appeared at the time, it was simply foreshadowing what was to come. During World War II, more than forty million refugees existed in Europe alone, uprooted and mostly without housing [25, 32]. This unprecedented disaster expedited the formation of international law and organisations tasked with the responsibility of managing 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 [25, 32]. In 1950, the office of the *United Nations High Commissioner for Refugees* was established. A year later, 147 governments signed the international *Convention on the Status of Refugees* [17].

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 five million refugees requiring assistance [32]. Then, in 1979, the Soviet Union occupied Afghanistan and an estimated five million people fled to Pakistan. Since 1990, the number of refugees in Afghanistan has not fallen below two million per annum [32].

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 during the period 1992–1995, resulted in a mass displacement of 2.7 million people (which, at the time, constituted half of the Bosnian population) [32]. African countries were also not excluded from this refugee crises. Rwanda, for example, suffered a mass genocide in 1994 which resulted in more than two million people seeking refuge in neighbouring countries [32].

In 2003, war broke out in the Darfur region of Sudan, causing the internal displacement of more than 2.5 million people [32]. Iraq experienced various humanitarian crises from the 1980s onwards 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 two million of which sought refuge across the border in Jordan, Lebanon and Syria [32].

Less than a decade later, in 2011, civil war broke out in Syria as rebel opposition fought against the Syrian government. At present, after eight years of war, the conflict in Syria still remains a reality between the Syrian government, the rebels, the Islamic State of Iraq and Syria, and the Kurds. As of 2016, an estimated eight million people were internally displaced within Syria, whilst an additional four 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 elicited from governments and partner organisations, in order to track global data on refugees, asylum-seekers and IDPs [170]. At the time of its establishment in 1951, there were an approximate 1.5 million refugees worldwide [118]. 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 [29, 80]. At the end of 2017, however, there existed an estimated 68.5 million forcibly displaced people worldwide, 25.4 million of which are classified as refugees, 3.1 million as asylum seekers and forty million as IDPs. These data are illustrated graphically in Figure 2.3. In comparison to the Earth's population of 7.7 billion, these numbers indicate that, on a global scale, 1 in every 112 people is now either an asylum-seeker, an IDP or a refugee [176].

The primary reason for the sudden increase in forcibly displaced people which occurred over the last decade is threefold. First, conflict situations which initially lead to a significant number of 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) [140, 169].

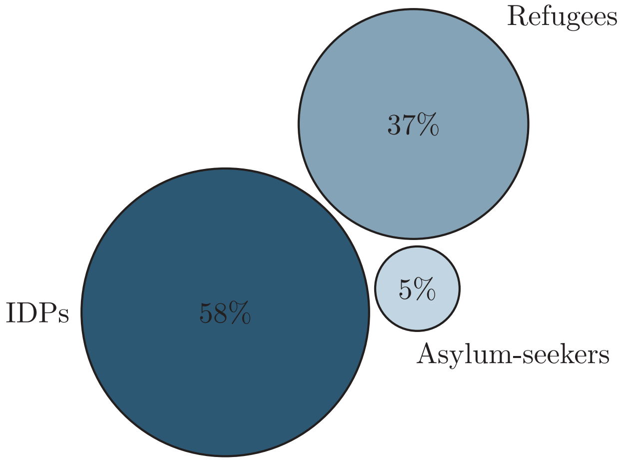


FIGURE 2.3: *The percentages of refugees, asylum-seekers and IDPs composing all forcibly displaced people worldwide, derived from the data reported by UNOCHA [176].*

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 [169]). Finally, feasible solutions on how to manage these crises are not proposed at a rate that 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, this rate had increased to 24 people per minute being forcibly displaced [170, 169].

The UNHCR further estimates that only five countries are responsible for producing more than two-thirds of the world's refugees. These are Syria, Afghanistan, South Sudan, Myanmar and Somalia. Conflicts in Syria, Colombia and the Democratic Republic of Congo have caused the largest number of IDPs, experiencing approximately 6.7 million, 6.5 million and 4.4 million people displaced within each nation's borders, respectively [81, 171]. The number of forcibly displaced people per nation worldwide as of the end of 2015 is shown graphically in Figure 2.4. In the figure, a larger number of displaced people is illustrated by a larger and darker icon.

Although the increasing number of refugees arriving in Europe *via* the Mediterranean Sea has captured the attention of many over the past few years, this accounts for less than two percent of forcibly displaced people worldwide. Most of these individuals (more than 85%) have relocated as refugees in low and middle-income countries which are in close proximity to the conflict. Major refugee hosting countries typically neighbour the countries of refugee 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% [144]. Turkey is currently the country that hosts the largest number of refugees worldwide with 3.48 million refugees living within its borders. Other noteworthy refugee-hosting countries include Jordan (2.89 million), the West Bank and Gaza (2.21 million), Lebanon (1.46 million), Pakistan (1.39 million), Uganda (1.35 million) and Islamic Republic of Iran (979 400). Lebanon, however, hosts the most refugees as a fraction of its total population, with 166 refugees per 1 000 inhabitants [171].

It is estimated that women, youth and children typically account for at least two-thirds of forcibly displaced people worldwide over the last decade [144]. The UNHCR was able to gather data, which showed that children below 18 years of age constituted 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, Syria, Eritrea and Somalia [170, 169].

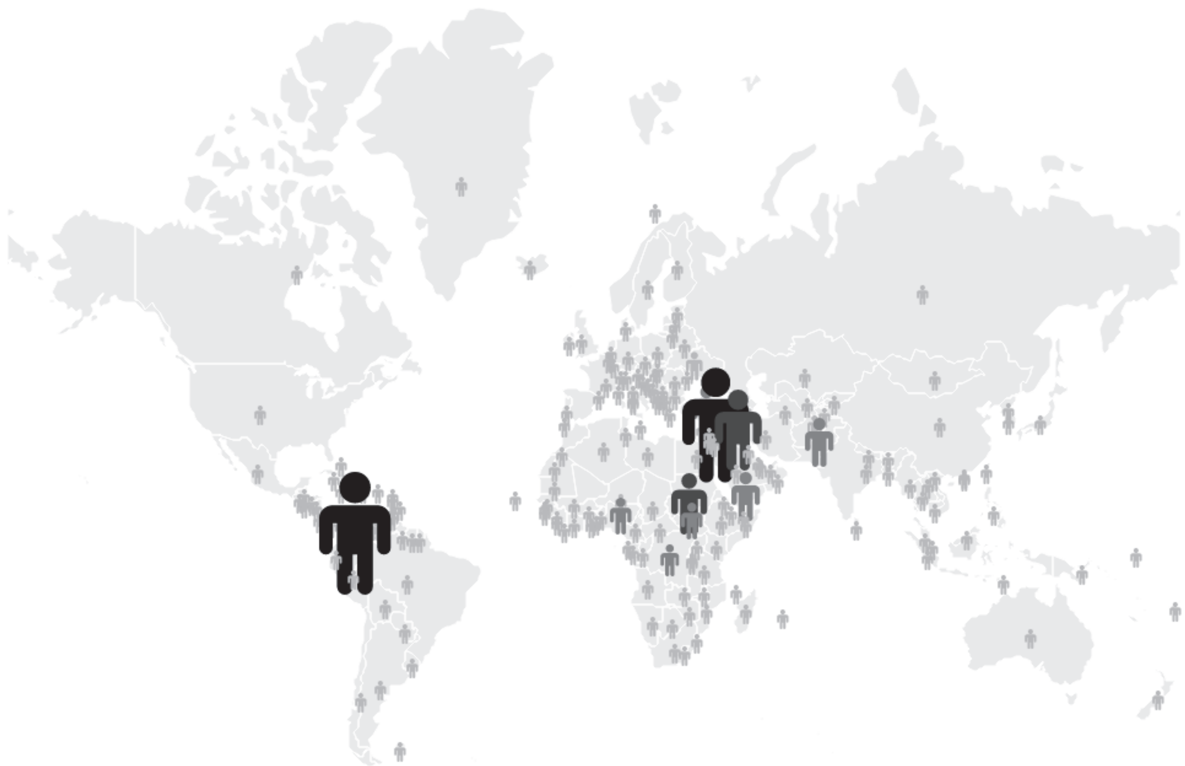


FIGURE 2.4: *Populations of forcibly displaced people worldwide, as adapted from the UNHCR [169].*

In 2019, an estimated 2 billion people are residing in fragile and conflict-affected areas in the world, where they are extremely vulnerable to the impacts of conflict and disasters. Furthermore, one in every four children, on a global scale, is living in a country affected by conflict or disasters where they face violence, hunger and disease [176].

2.5 Chapter summary

In this literature review chapter, an overview was provided of the concept of forced migration. This phenomenon has taken place throughout history and remains topical today, particularly in Syria, where war broke out in 2011. In §2.1, the factors causing 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 reviewed 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, after which worldwide cases of current forced displacement was discussed in §2.4, with respect to refugees, asylum-seekers and IDPs.

CHAPTER 3

Computer simulation modelling

Contents

In this chapter, the literature on computer simulation modelling is reviewed. The chapter opens with an overview of the discipline, as well as an elaboration on the different types of and approaches to simulation modelling. The agent-based approach to simulation modelling is discussed in more detail, and this is followed by a discussion on the use of ABM in simulating refugee movement.

3.1 Simulation modelling

In order to understand a system, the relationships between its various components is required. This also facilitates the ability to predict its behaviour under operating changes [101]. 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 the relative merits of various strategies (within the limits imposed by a criterion or set of criteria) for the operation of the system [146]. Simulation modelling therefore provides a risk-free environment wherein a real-world problem can be addressed first by simplifying the system and then experimenting with the model under various conditions [21]. 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 deciding on a suitable modelling approach [21].

3.1.1 Types of simulation modelling

Simulation models can be characterised according to their time dependency, the nature of their input data and the manner in which time is modelled. Law and Kelton [101] 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 instant into account. In contrast, 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 their behaviour. In a *deterministic* simulation model, the input variables are known and, thus, no probabilistic components are present in the model. It is, however, more likely that a system will involve some form of randomness which will then be modelled 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 that it is most likely only an estimate of a system's true characteristics [101].

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 also be categorised by the quantity of information associated with the model, known as its level of abstraction [22]. Figure 3.1 contains an illustration of the different levels of abstraction and the corresponding systems that can thus be modelled. The amount of information required by and represented within a simulation decreases as the abstraction level increases. Selecting the right level of abstraction is critical to the success of the simulation model [21].

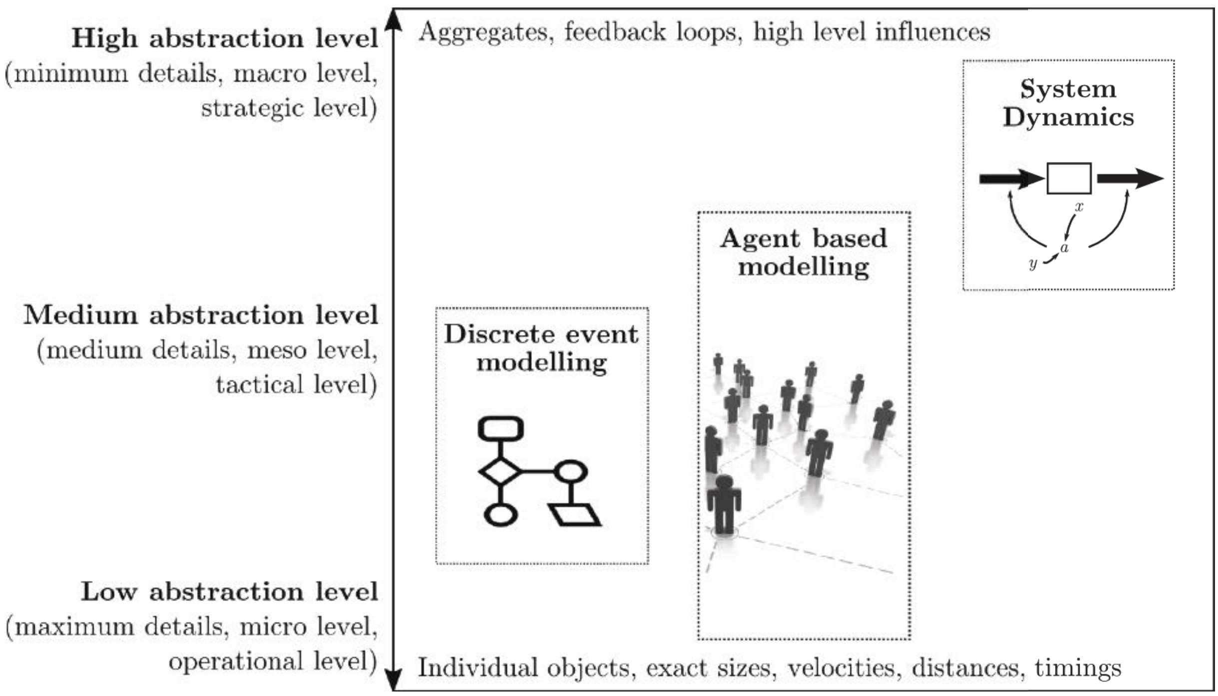


FIGURE 3.1: *The levels of abstraction relevant in modelling systems, adapted from Borshchev [21].*

A high abstraction level refers to a model that has minimal detail, complementing the strategic level at which the model is 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, are highly abstract as they occur on a macro level. In systems where 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 requires 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 that 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, and each is appropriate for a certain range of abstraction level. These methods are elaborated on in §3.1.3.

3.1.3 Simulation modelling approaches

In modern computer simulation, three different approaches prevail in which models can be developed for the purpose of simulating real-world systems, namely discrete-event modelling, system dynamic modelling and ABM [21]. 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 [145]. Whenever entities compete for constrained resources, the formation 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 [21]. Entities can be clients, products, patients or tasks, amongst others, while resources might be doctors, workstations, staff or 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 [145]. Discrete-event models are stochastic as the service and arrival times of the underlying system are drawn from a probability distribution. Due to its stochastic nature, the model requires adequate runtime or a sufficient number of replications before meaningful output can be obtained [21].

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 [22]. Borshchev [21] explains this approach by means of an example: When considering a shop owner serving clients, the queue grows longer as the number of people entering the shop increases. 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 an extended waiting time. As a result, the length of the queue impedes the growth of rate thereof. Causal loops have to be considered in such a scenario, affording a suitable opportunity for the implementation of the system dynamics method.

ABM

Of the three simulation modelling approaches discussed, the ABM approach is the most recent [22]. ABM allows a model developer to capture a large amount of detail within a model which is not typically achievable when adopting the other approaches. The approach 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 the incorporation of 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 or the propensity to adapt their standard behaviour as a result of interaction with external stimuli [92]. 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 [21].

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. The system dynamics and discrete-event modelling are approaches that were developed during the 1950s and 1960s, respectively, while ABM is a fairly novel approach [21], 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 [22].

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 thorough and reliable models [14]. Figure 3.2 is an illustration of the progression of the steps within a typical simulation model study. The framework presented is a consolidation of the research by Banks [14], Law [98], Shannon [147] and Stewart [159].

The generic steps of the framework are as follows:

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

The purpose of the study should first be defined clearly in terms of the project objectives [147]. 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 [14]. 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 resources required, as well as the system configurations that will be modelled [98].

(3) **Conceptual modelling**

A conceptual model can be defined as a non-software description of the proposed simulation model. It should comprise the objectives, assumptions, inputs, outputs, algorithms, data and clarification of the aforementioned [159]. This may be documented either graphically or by means of pseudo-code [147].

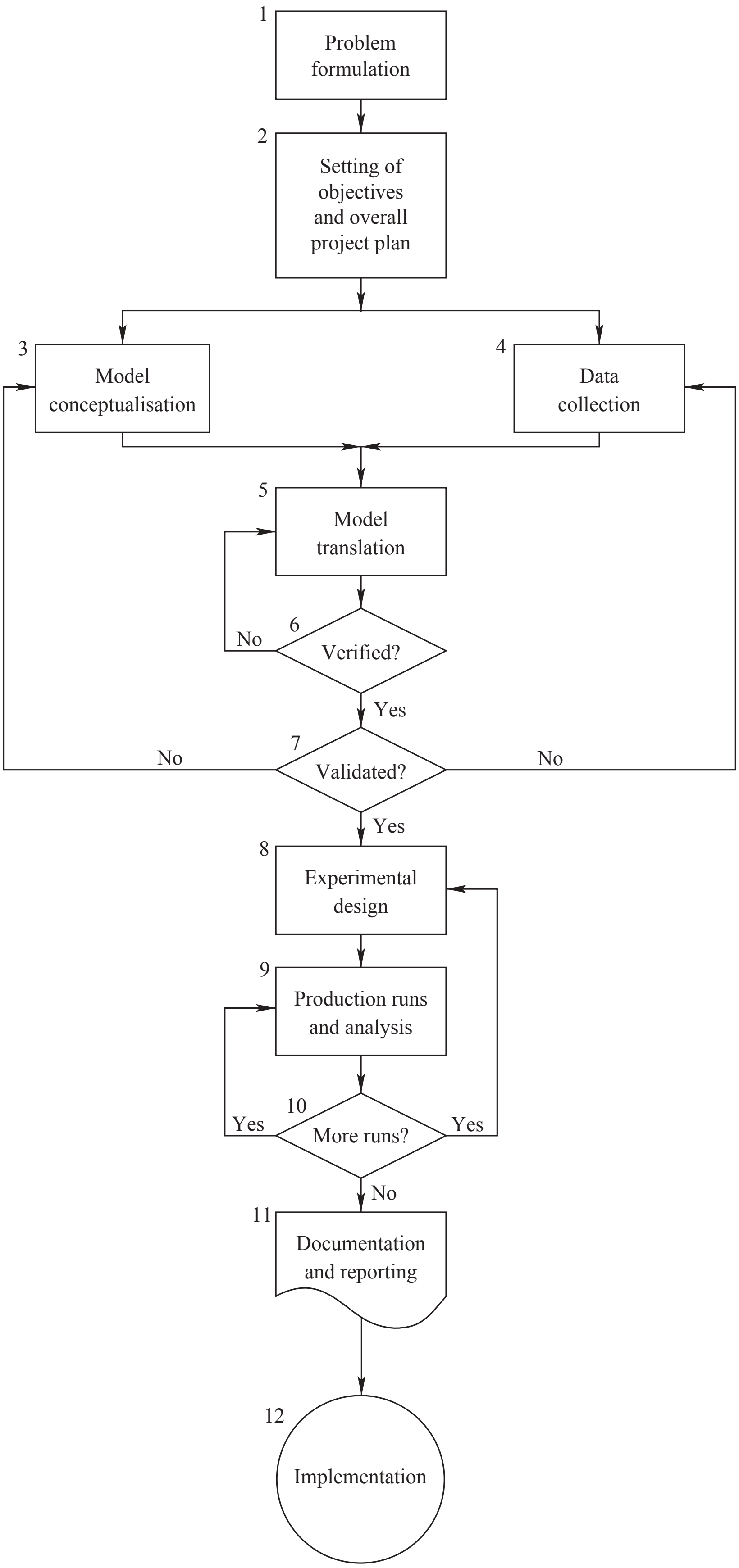


FIGURE 3.2: *The steps in the design and development of a simulation model, adapted from Banks [14].*

Part of designing the conceptual model involves determining whether simulation is indeed a suitable approach toward solving the problem at hand, as well as choosing which of the simulation modelling approaches discussed in §3.1.3 to employ [159]. Before finalising the conceptual model, the model developer should perform a structured walk-through for validation purposes [98].

(4) **Data collection**

Data collection and analysis occurs in conjunction with conceptual modelling, since contextual data are required in order to understand the system. The conceptual model should indicate which detailed data are necessary in terms of developing the simulation model [159]. Information on the system layout and its various operating processes should be collected, along with specific data related to the model parameters and accompanying probability distributions [98]. The model developer should further ensure that the data are adequate in terms of quality, quantity and variety in order to perform reasonable analyses [14].

(5) **Model translation**

The next phase of the simulation study requires a conversion of the conceptual model into an operational model [14]. This is accomplished by programming the model in a commercial simulation software package, such as AnyLogic, Simio or Vensim [98, 159]. Although general-purpose programming languages give the modeller better control when formatting the model into executable code, the use of specialised simulation software packages may reduce the time spent on programming, as well as improve the efficiency and effectiveness of the model [147].

(6) **Model verification**

Verification applies to the operational model and is aimed at confirming whether or not the model is performing properly (*i.e.* as expected and intended) [14]. The model developer should verify the model on a continual basis and it is advisable to make use of an interactive run controller or a debugger to assist in the verification process [14]. 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 is aimed at ensuring that a model accurately represents the real-world system it represents [14]. It allows the model developer to reach an agreeable level of confidence, suggesting that the model conjectures are legitimate and apply to the underlying real-world system [147]. There are many methods of validating a simulation model; these 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, have to be determined for each scenario that is to be simulated [14, 98]. An experimental design should be executed in order to yield valuable information and test run specifications [147]. Once completed, production runs and the results of previous experimental designs may be used to modify the input values for future experiments. This is an iterative process in which "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 ascertain whether the observations in the data set are independent (*i.e.* not correlated), to assist in sensitivity analyses or the like [98, 147]. The key outcomes of this phase are the discovery of sufficiently accurate results, search through the solution space and testing the robustness of solutions [159].

(11) **Documentation and reporting**

At this point in the study, all design, programming and execution of the model should be completed, and an analysis should be carried out of the accompanying results [147]. The next step is to document the study by compiling a report that should include the conceptual model, a detailed description of the simulation model and the results obtained when executing the simulation model [98]. The document should convincingly present the outcome as clearly and concisely as possible [147]. Furthermore, the documentation should ideally include an animation of the simulation model and a proper discussion of the model verification process employed [98]. Documentation is essential for future reuse of the model and holds the added value of actuating the client's confidence in the simulation model [159].

(12) **Implementation**

Shannon [147] argues that a simulation study can only be considered completely successful when its results are understood, accepted and used. Simulation studies may be implemented in various ways. First, the simulation study's solution can be implemented in the real-world system and put into practice [159]. 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 [159].

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 [100, 159, 163]. A simulation model of a complex system can only be an approximation of the actual system. As such, an absolute model validation is not possible, because the model is purely an abstraction and simplification of reality [98].

Stewart [159] suggests some difficulties that may be encountered during the validation process. A model can only be validated according to its objectives. First, a model that is valid for one purpose may not be valid for another. Furthermore, the modelled system does not always exist in the real world, but it is essential that all models are validated, irrespective of whether or not it exists [98]. Another issue that may arise is the accuracy of real-world data, as probabilistic data are not definite. Validation should be considered as essential within the simulation modelling steps of §3.1.4 and it is the model developer's responsibility to ensure that validation is performed.

A three-step approach to validation is discussed by Law *et al.* [100]. This includes high-level face validation, testing model assumptions empirically and establishing whether simulation output data resemble the real system.

Face validation

A model should be developed in such a manner that people who are knowledgeable in respect of the underlying 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 [163].

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 physically exists), although care should be taken when collecting data during observation, ensuring that such data are 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 so as to fully understand the thought processes of these individuals.

Empirical testing

The second step in validating a model involves quantitatively testing of the assumptions made during the model building process. Observed data are often fitted to theoretical probability distributions, which are provided as input to the model. Goodness-of-fit tests, such as the chi-square or Kolmogorov-Smirnov tests [99], can be used to test whether or not the fit is indeed adequate for a particular application of a model. For further details on these and other goodness-of-fit tests, the reader is referred to Law [99]. Another useful tool is a sensitivity analysis, as it identifies parameters that have a greater influence on the output of a model. In this way, the model developer can identify certain parameters to which the output of the model is sensitive and obtain better estimates of these parameters. The model will thus represent the real-world system more accurately, which would add to the perception of its validity.

Comparison of output data

The final step is to establish whether or not the model's output accurately represents that of the underlying real-world 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 with the associated output of the model. If the comparison is favourable, the simulation model of the underlying system is considered 'valid' [98]. The greater the commonality between the real-world system and the simulated system, the greater the confidence in the simulated model. The model developer will never know whether a system is completely valid, as its level of accuracy still remains an estimation. This step in the validation process nevertheless provides a sense of confidence and credibility in respect of the model.

3.2 Agent-based modelling

In its simplest form, ABM entails the construction of a system of autonomous and interacting agents [20]. It is a dynamic simulation technique in which agents, embedded in and interacting with, the environment possess the capacity to learn and adapt their actions in response to changes in the actions of other agents and changes in the environment [7]. ABM follows a so-called bottom-up approach. In an agent-based simulation model the global characteristics of the system are never defined; agent behaviour is instead specified on an individual level [22]. The global system 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 [70]. ABM is well-suited 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 [110]. A broad variety of ABM applications exist and the disciplines in which ABM has been applied also vary. This can range from modelling consumer markets [57] to generating social networks [10] and modelling the diffusion of innovation [92], from coordinating supply chain and inventory control [84] to modelling organisational science [54].

3.2.1 Advantages and disadvantages of agent-based simulation

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

The most integral benefit of simulation as a modelling paradigm is its facilitation of investigations into possible solutions and situations in a virtual reality without the potential of incurring the risks associated with these solutions or situations in real life, such as committing resources or disrupting ongoing operations [21, 147]. Apart from the risks, virtual time allows the model to perform a simulation, which would have taken days or months in reality, in mere minutes [147]. A simulation model can be executed much faster than real time, enabling the model developer to gather results of the system's performance much quicker than would have been possible in reality [147, 159]. Furthermore, simulation affords the model developer the possibility of animating the system, which may be beneficial in conveying the nature of operations or a potential solution to a client [21, 147].

Despite the considerable advantages of adopting a simulation modelling approach, some drawbacks do exist. The process of developing a credible simulation model takes time and requires a specific level of skill and experience from the model developer. Furthermore, although the simulated world is cost-free, specialised simulation software can be expensive [100, 147, 159]. In the same cases, a model may be built accurately, but the model developer must be able to interpret the results of the model constructed in order for it to have any significance [100]. Another challenge for the model developer is acquiring reliable data. This time-consuming process should be handled with care for the model's credibility depends on the reliability of the data [147]. The credibility of a model is important and so a model requires proper validation and verification [7, 147]. This is compounded by the typical limitation of reliable input data to represent the real-world system accurately and the challenge of having sufficient output to compare with reality [147].

Advantages of ABM

The ABM paradigm, in particular, also holds a number of significant advantages when selected as a simulation modelling approach. According to Borschev and Filippov [22], the ABM technique is more general and powerful than other existing paradigms, as it is often capable of effectively capturing real-world systems with their complexities and dynamics. Bonabeau [20] summarises the benefits of ABM in three respects: The manner in which it captures emergent behaviour, the ability of ABM to model the nature of a system and the flexibility of the tool. It has been claimed that the benefit of capturing emergence drives the other advantages of ABM [20].

Emergent phenomena, resulting from the interaction of constituent parts, are difficult to understand or predict and may even be counterintuitive. An example of this is the movement of individuals in the case of an evacuation, where the collective panic behaviour results in emergent behaviour. It is advised that ABM be utilised in cases where emergence may be evident, such as when individual behaviours are non-linear or exhibit memory, as well as when the interactions between agents are heterogeneous [20].

Another favourable aspect of ABM is the fact that agent-based simulation models can be constructed in the absence of information about the system as a whole, by using the details of stakeholders on an individual level. This aspect also makes it easier to maintain the model as changes and refinements to the system will typically occur locally instead of on a global level [22].

Disadvantages of ABM

As with many modelling approaches, certain impeding factors apply to ABM. One such factor often mentioned in literature is the difficulty of validating and verifying agent-based simulation models [7, 92]. Before commencing with an 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.* [57] advise the model developer not to overreach in adding details to the model, as it may result in a model with too much sophisticated logic per agent. This in itself is not detrimental, but, as the detail increases, so the resulting increased complexity may cause difficulties due to computational burden of ABM and a lack of adequate data.

3.2.2 Components of an agent-based model

Macal and North [110] describe 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 [86], assert that the main elements of an agent-based model include agents, global interactions and organisational relationships, as depicted in Figure 3.3.

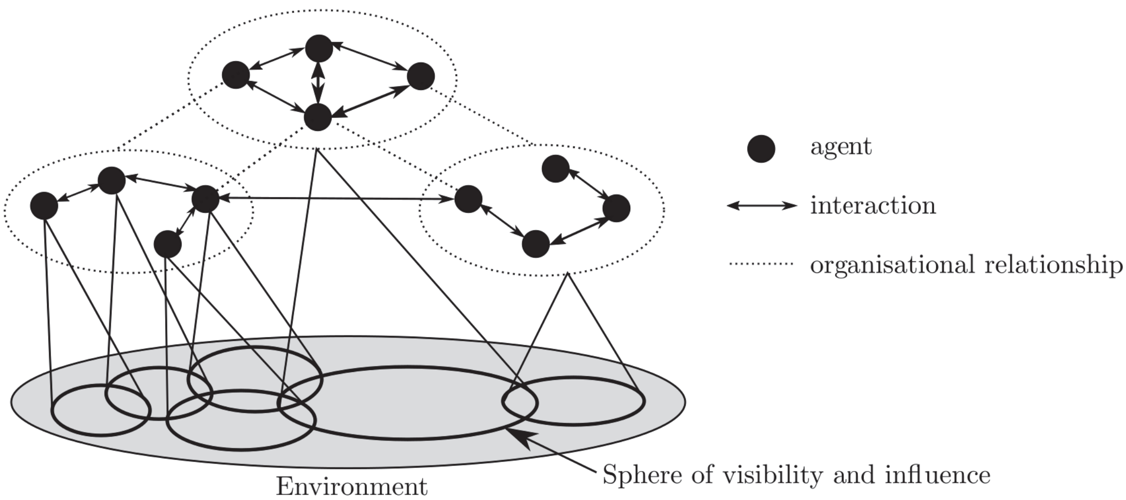


FIGURE 3.3: *The essential components of ABM, adapted from Jennings [86].*

All of the aforementioned components contribute to the structure of an agent-based model. The model developer is required to identify, plan and define these elements in order to construct the simulation model [110]. Agents, their interactions and the environment are now further elaborated upon.

Agents

Beyond the fundamental property of autonomy, the definition of an agent has not been formally agreed upon in the literature. Some authors emphasise that an agent should mainly be defined by its autonomous properties [20], while others argue that an agent is adaptive and defined by its ability to learn from experience. Macal and North [109] define 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 [86, 110, 109]. Furthermore, agents possess the following characteristics:

- Agents are adaptable and uniquely identifiable as individuals [7, 86, 110, 109]. Each agent has a set of characteristics which determines its behaviour and decision-making abilities [109].

- Agents have specific design objectives [86] and are goal-orientated [86, 110, 109]. For this reason, agents can compare the outcomes of certain courses of action in terms of their objectives and alter their future behaviour in order to achieve these goals [110].

- The state of an agent (*i.e.* its subset of attributes) changes over time and affects the behaviour of the agent. Many possible states ensure that a rich set of possible behaviours can be exhibited by an agent [110].

- Agents are dynamic and social individuals which interact with the environment and other agents, thereby impacting on the behaviour of these agents [110, 109]. Certain rules of conduct exist in respect of interactions with other agents in terms of communication, movement, competition for space and the like [110]. Agents are capable of recognising the characteristics of other agents [110].

- Agents are flexible problem solvers in pursuit of achieving their design objectives [86]. They are capable of learning from previous experiences in order to adapt their behaviour as their circumstances change [109].

A typical agent and its interactions are illustrated in Figure 3.4. Agents have certain attributes and methods associated with them. Agent attributes can be of either a static or a dynamic nature [110]. 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) [110].

Agent methods refer to specifications of the behaviour of agents. These specifications can take the form of simple rules or complex models (such as neural networks or genetic algorithms) [20, 110]. In order to model an agent's behaviour, a theory of agent behaviour is required for circumstances that may occur in the model. If such a behavioural theory does not exist, a normative model, in which the agent is required to optimise its design objective, can be utilised as a starting point [110]. Although this may be a simple model, the conclusions drawn from the model may assist in developing an elementary, more detailed (although sensible) heuristic behavioural model [110].

In conclusion, agents are independent decision makers; this allows them to be dynamic entities, rather than mere passive components in the simulation model [109].

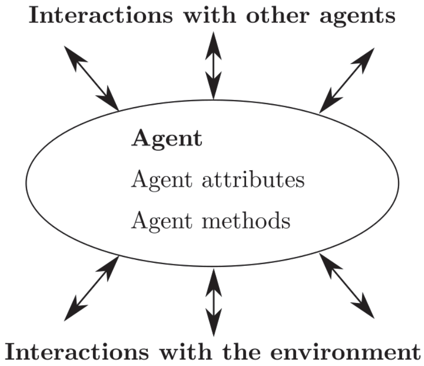


FIGURE 3.4: *A typical agent and its interactions, adapted from Macal and North [110].*

Agent relationships

An agent-based model should be able to imitate the interactions between agents, which may be both repetitive and competitive [20]. Aspects to address when modelling these relationships include identifying which agents are capable of interacting with one another and specifying the dynamics of these interactions [110]. This has to be taken into account when developing the model.

Agent-based systems are decentralised and thus no central authority exists — information rather spreads *via* interactions [110]. An agent would typically interact with its neighbouring agents, but 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 [110]. Interaction with other agents may be crucial for an agent in respect of achieving its objectives. Furthermore, agents interact with one another to manage relationships formed in the common environment [86].

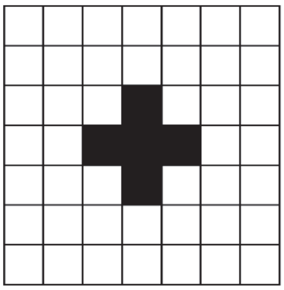
As agents move throughout the simulated environment, the actions and attributes of their neighbours can change rapidly. In order to model agent relationships, a topology of agents has to be determined. The notion of topology in the context of ABM can be defined as the manner in which agents are connected (*i.e.* who transfers information to whom) [110]. The topology selected to model agent relationships can significantly change the outcome of the model in comparison with predicted cooperative behaviour [20]. Therefore, choosing the right topology is necessary in order to model the social influence of the system accurately [92]. Typical topologies include spatial grids as well as more general networks of nodes and links [110].

Figure 3.5 contains illustrations with five different topologies often employed in ABM, as defined by Macal and North [110]. Cellular automata allow agents to move from one cell another in a grid, where no more than one agent can occupy a cell at a time. The von Neumann '5-neighbour', shown in Figure 3.5(a), is an example of one of the many cellular automata neighbourhoods available. A Euclidean space model allows agents to move freely in two-, three- or higher-dimensional spaces, as illustrated in Figure 3.5(b). A *geographical information system* (GIS) topology represents a real geographic landscape which is divided into patches, as illustrated in Figure 3.5(c), and agents can move between these patches.

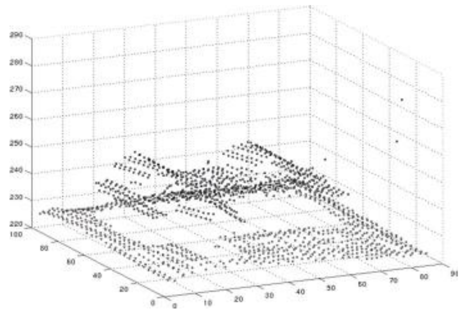
Figure 3.5(d) refers to a network topology, which is a more generally defined agent neighbourhood. Networks can be of either a static or dynamic nature. The links between agents are set and cannot be altered in a static network. In a dynamic network, however, the links and nodes are endogenously decided upon [110]. Lastly, in the 'soup' model, depicted in Figure 3.5(e), location is irrelevant and agents interact by means of a random selection [110].

The environment

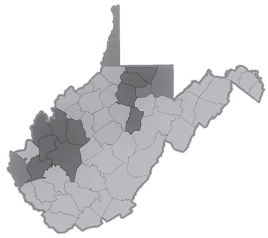
Agents within an agent-based model interact with and are enclosed in a dynamic environment [7]. This environment provides information to the agent which can vary from simply indicating an agent's location relative to other agents, or by providing geographical information in the case of a GIS [110]. The location of an agent is a dynamic attribute of the agent, used to track its movement.



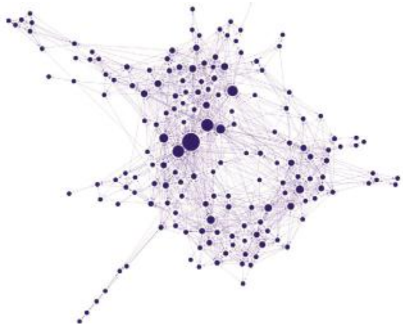
(a) *Cellular Automata*



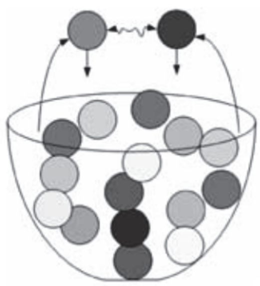
(b) *Euclidean 3D space*



(c) *GIS topology*



(d) *Network topology*



(e) *'Soup' model*

FIGURE 3.5: *The various topologies employed in ABM.*

An agent can learn from its environment and adapt as necessary when changes occur [7]. The environment may also create certain spatial constraints for the agents. A transportation model is an example of such an environment, as it would contain infrastructure and capacity constraints in the links and nodes representing the road network [110].

3.3 ABM in the context of forced displacement

Migration is a complex process which encompasses certain movement patterns, historical context and individual decision making [44]. 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 [72] propose the integration of aggregated statistical data along with a review of the relevant literature in order to accurately capture the complex process of migration within a modelling context.

By using ABM, agent behaviour can potentially be explored in an attempt to better understand the various processes and consequences concerned with refugee movement. Such a model, once sufficiently validated, could possibly aid as a decision support tool for humanitarian relief [37, 68]. ABM is particularly well-suited for such an application for a number of reasons. First, 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 a disaster. Furthermore, the heterogeneity of agent control recognises that decision making may differ according to 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 assists in replicating realistic scenarios [37].

3.3.1 Types of movement modelling

Groen [68] identifies three different ways in which human movement can be modelled, namely evacuation modelling, refugee modelling and migration modelling. The difference between these movement modelling types lies in the variation of the temporal and spatial scales. 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 from within a few metres to no more than a kilometre, while migration modelling could develop over an area of hundreds of kilometres. As may be seen in Figure 3.6, the scale of refugee modelling falls between those of evacuation and migration movement modelling.

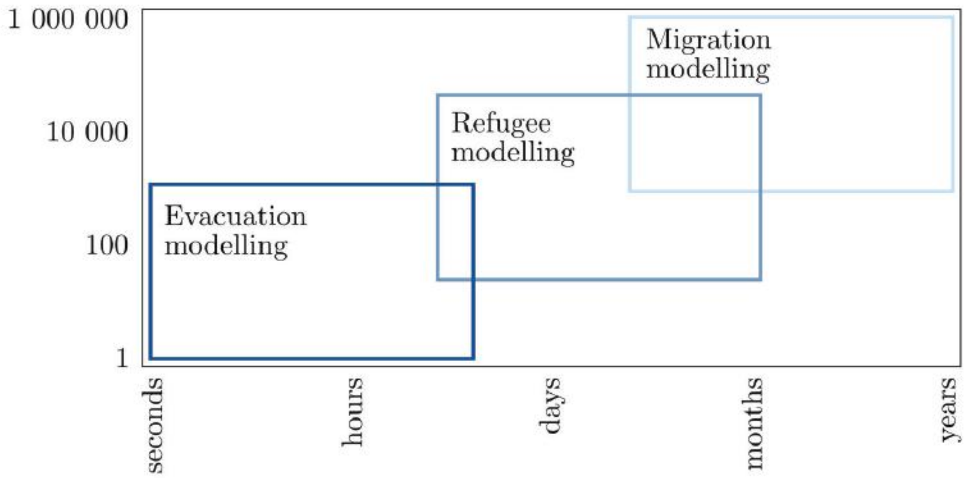


Figure 3.6: *The types of movement modelling discussed by Groen [68].*

Evacuation modelling has been used in the literature to model pedestrian dynamics, crowd movement and urban evacuation [95, 187]. 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 units of seconds or minutes.

Migration modelling has been employed widely in the literature [16, 65, 106, 131, 155], mostly in the context of voluntary migration. Smith *et al.* [155] modelled a particular migration which occurred over a period of thirty years. In their model, migration was driven regionally and internationally due to climate change. Lin *et al.* [106] utilised a country-level agent-based dynamic network model in examining the displacement of populations based on given network relations (such as alliances, shared language, economic influence and proximity) amongst countries. Voluntary migration is primarily driven by socio-economic factors, while refugees are typically forced to relocate for their own safety [138].

Forced migration modelling (*i.e.* refugee modelling) is unique in its spatial and temporal scale as it may span hundreds of 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 [68]. Individuals forced to migrate can relocate within a country's borders (*i.e.* IDPs) or flee across international borders (as discussed in §2.2). It may, however, be argued that the temporal scale of refugee modelling should be extended to years, as it is noted in the *UNOCHA Global Humanitarian Overview of 2019* that the average humanitarian crisis continues for more than nine years [176].

3.3.2 Existing forced migration models using ABM

Agent-based models depicting forced migration do exist in the literature although these studies have mostly been published within the last decade only [8, 35, 37, 47, 67, 68, 72, 106, 155, 156, 164].

In one such study, Collins and Frydenlund [35] investigated strategic group formations which occurs among refugees when migrating. ABM methods were incorporated in this investigation, along with game theoretic 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 size of the group as well as the speed with which it moves in order 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 crowd-sourced data and GIS data, to simulate the situation after a natural disaster in order to explore people's reactions to aid [37]. Geographic information pertaining to population density, existing transportation networks and aid centre locations were included as an input to the model. Crooks and Wise [37] 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.* [8] 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 effect 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 demonstrated by Hailegiorgis and Crooks [72]. 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 accommodated in the model.

Sokolowski *et al.* [156] performed a study on population displacement within the Syrian city of Aleppo using ABM to characterise individuals, entities and the environment by means of an 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, called FLEE, has been developed by Groen [68] for predicting the movement of refugees towards refugee camps, with a particular focus on movement during the Northern Mali conflict in 2012. In another paper, Groen [67] extended the FLEE model to parallel implementation and multi-scale coupling in order to explore larger-scale movement.

Furthermore, a parallel implementation of FLEE, forecasting refugee movements in three major African countries, was described by Suleimenova *et al.* [164].

A more complex approach was adopted by Edwards [47] as counter-intuitive or 'aberrant' patterns of flight were considered in predicting the spatial flow of conflict-induced migrants. According to Edwards, 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 with the use of ABM and knowledge of the appropriate input variables.

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 migrate from and where do they migrate to?*, *When do they migrate?* and *What are the consequences of the movement?* [65, 114]. ABM provides model developers with the ability to simulate the autonomous decision making of agents with respect to these determining factors.

Alhanaee and Csala [6] performed a study on the motives of Syrians when choosing to seek refuge in Lebanon. Upon adopting a gravitational migration model and data from both social media and the UNHCR, it was 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 [6] 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 [121], an individual's preference towards migration was determined by a certain threshold value. Suppose there is a probability $p \in [0, 1]$ that a person will be forcibly removed. As p increases from 0 to 1, there exists an inferred threshold value which, when exceeded, results in a person preferring migration over remaining at his or her current location. Many factors influence an individual's threshold 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 at which violence can be avoided, typically choosing to relocate to places where other refugees have gone before [39, 123].

In attempting to predict displacement or a person's inclination to migrate, the characteristics of the individuals, the environment and the relevant events that trigger displacement need to be considered. Parameters that model intricate considerations (such as linguistic homogeneity, wealth inequality and ethnic chasms), as well as more tangible influences (such as conflict intensity, spatial spread of the conflict and geographical features) should also be considered where possible [47]. In addition, Edwards [47] recommends adoption of the push-pull and 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 [71].

The push-pull approach explains the motivations of an individual to move. The 'push' factor provides a person with causal motivations with respect to leaving their places of residence, while the 'pull' factor of the chosen destination prompts the person to migrate [97]. The decision of an individual to move and their choice of where to go, are formed by a unique combination of circumstances. Smith *et al.* [155] acknowledges the push-pull effect as intervening obstacles which influence an individual's migration aim.

'Push' factors, in the context of 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 [11, 16, 40]. Apart from push-pull factors, Edwards [47] mentions that the decision making of an individual is further influenced by the social or psychological effects of trauma (caused by the conflict situation), as well as other individualistic behaviour. A broader range of factors considered will allow for a model possessing a more accurate predictive value.

Martin and Singh [114] identified big data sources, methodologies and challenges with the aim of improving the effectiveness of early-warning systems with respect to forecasting mass displacement of people in the context of humanitarian crises. They investigated the analytical capability of big data mining in respect of discovering indicators that lead to displacement, including macro drivers (*e.g.* environmental and political indicators), micro drivers (*e.g.* household demographics) and intervening factors (*e.g.* terrain and policies). These drivers were combined with the push-pull approach in pursuit of discovering which of these factors are relevant at different geographical locations and how these drivers may interact with one another.

In another study by Moore and Shellman [122], 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 of 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 employed is Maslow's *hierarchy of needs* [116], which was developed as a theory of human motivation and comprises a five-tiered model of hierarchical levels of motivation. Maslow argued 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.

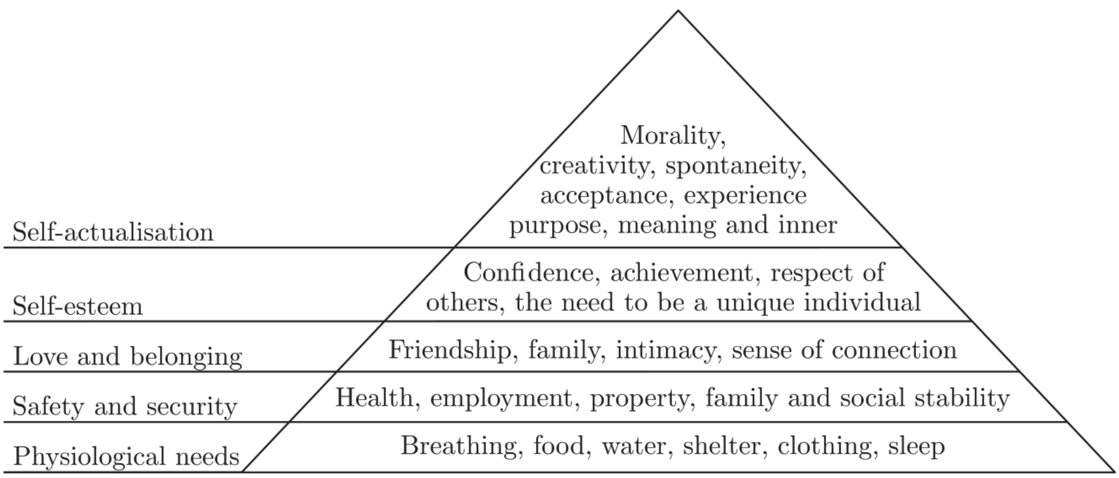


FIGURE 3.7: *The hierarchy of needs proposed by Maslow [116].*

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 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 the literature as determinants of migration [6, 37].

3.4 Chapter summary

Computer simulation modelling, and ABM in particular, were discussed within this chapter. More specifically, ABM was considered in the context of modelling refugee movement.

Simulation modelling was described in §3.1, including aspects such as the different types of modelling, the various levels of abstraction and different modelling approaches. The generic steps typically followed in a simulation study were detailed in §3.1.4, after which various methods for validating 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. In addition to the aforementioned background on 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, and existing studies on forced migration modelling by means of ABM and the factors leading to forced migration were reviewed.

CHAPTER 4

Decision-making methods and human decision behaviour

Contents

Influences and considerations in human decision making are studied within this chapter. Various categorisations of decision-making methods are discussed in §4.1, before considering multi-criteria decision-making methods in §4.2. The chapter closes with a general discussion on human decision making in §4.3.

4.1 Decision making as a field of research

Wang and Ruhe [182] define decision making as a cognitive process during which 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 [42, 43, 160, 182].

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 [160]. Prescriptive decision scientists are therefore concerned with prescribing methods that are anticipated to result in optimal decisions [43]. 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 Pareto optimality and efficiency [31, 42]. Prescriptive decision-making methods are typically used in operations research and the management sciences, where an algorithm recommends some form of decision optimality. A classic example is the *Travelling Salesman Problem*, where a travelling salesman is required to visit a number of cities, each exactly once, while minimising travel distance [15, 142].

Descriptive decision-making theories, on the other hand, imitate the manner in which people make decisions [160]. Such theories are concerned with the actual behaviours and decisions made by humans in an attempt to describe the underlying decision processes [15]. Descriptive decision making is described in this section by means of an example [15, 142]. Consider the case where four vehicles, 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 on the scale of evaluation. Each of these attributes has been allocated a weighting to indicate its relative importance. A cut-off value per attribute is also given. These cut-off values represent aspirational levels which represent minimal scores at which an alternative is considered acceptable.

TABLE 4.1: *An example of descriptive decision making, as presented by Bell et al. [15].*

	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 vehicles provided in the table, one vehicle needs to be chosen. There are various ways in which to arrive at an optimal alternative. For example, when implementing the *lexicographic* rule, the alternative which scores the best on the most important attribute, which in this case is price, will be chosen. Vehicle A would thus be the best alternative according to the lexicographic rule, as it received a score of seven (the highest) in respect of price.

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, Vehicle B fails as an alternative, as it scores below the cut-off value in respect of price. Similarly, Vehicle A's score in respect of safety is below the cut-off value and is thus also eliminated as an option. Vehicle C, however, meets or supersedes the cut-off values associated with all the criteria and should, according to the satisficing rule, be chosen as the best alternative. In this case, different descriptive decision rules lead to completely different results. In other cases, however, different rules may agree on the best alternatives.

A significant gap exists between prescriptive and descriptive decision-making methods [42, 160]. Prescriptive modelling is concerned with optimality and with solving a problem by minimising or maximising some metric, whereas descriptive modelling simply describes human thought processes when making decisions [142]. 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 [33], Gigerenzer and Gaissmaier [61], Simon [151], and Wang and Ruhe [182].

Simon [151, 152] introduced a three-phase decision-making process consisting of the phases *Intelligence*, *Design* and *Choice*, depicted graphically in Figure 4.1. The first phase, intelligence, refers to a process during which the environment is explored and the need for improvement or change is identified. The design phase, following the realisation of such a need, involves investigation of the problem and possible alternatives, thereby broadening the problem domain. During the final phase, the most appropriate alternative is selected from amongst the alternatives generated during the previous phase [42].

The flow between these phases is complex. For example, further intelligence may be required during the design or choice phases. Furthermore, within any phase, a new decision process may arise, itself containing the three-phase decision-making process within the existing phase [42].

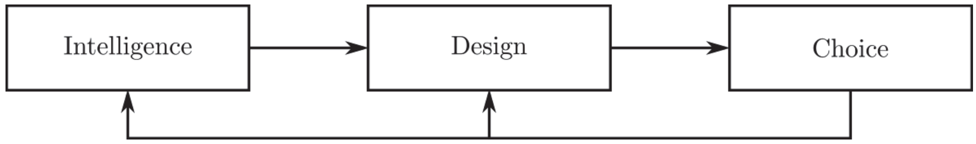


Figure 4.1: *The three-phase decision-making process proposed by Simon [152].*

4.2 Multi-criteria decision-making methods

The study of *multi-criteria decision-making* (MCDM) methods, otherwise known as *multi-criteria decision aid*, relates to generic planning problems in which the most attractive solution is sought from a given set of alternatives. Each of these alternatives is characterised by scores specified in terms of a set of selected criteria. Furthermore, input from a set of interest groups is considered with respect to the selection criteria and associated relative importance. MCDM methods are used to capture the attractiveness of numerous alternatives with the aim of identifying the most attractive solution [117]. As many conflicting priorities are typically considered in making the decision, there is usually no longer a single optimal solution, but rather a set of satisfactory trade-off solutions [69, 180].

MCDM is not restricted to a specific field, but applies to all branches of operations research. Furthermore, it can be prescriptive or descriptive in nature and may apply to many real-life problems [180]. Very few problems have a single criterion determining the choice of solution and decision makers are typically confronted with multiple objectives, attributes and criteria [186].

According to Guitouni and Martel [69], 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 MCDM, however, lies within the multi-criterion aggregation procedures (a combination of steps (ii) and (iii)), considered in more detail 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 [69], Huang *et al.* [74], Hwang and Yoon [76], Malczewski [113], Massam [117], Mota *et al.* [124] and Vincke [180]. Only the classification methods most relevant to this study are considered in this literature review.

One classification entails categorising MCDM methods as continuous or discrete, depending on the nature of the set of alternatives [36, 124, 125, 185]. Hwang and Yoon [76] refer to these categories as *Multi-Attributive Decision-Making* (MADM) and *Multi-Objective Decision-Making* (MODM) methods, respectively. Both methods involve a set of alternatives which are evaluated based on possibly conflicting criteria. These criteria include attributes (the measurements of the system with respect to an objective) and objectives (the desired state of the system under certain conditions, derived from attributes) [48]. Criteria distinguishing typical MADM from MODM methods are listed in Table 4.2.

MADM methods are applicable to decision-making situations in which a finite and feasible set of known alternatives exist and are specified in terms of attributes. These attributes are regarded

TABLE 4.2: *A comparison of MODM and MADM approaches as presented by Malczewski [113].*

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	Large	Small
Decision maker's control	Significant	Limited
Decision-modelling paradigm	Process-orientated	Outcome-oriented
Relevant to	Design/search	Evaluation/choice

as both decision variables and criteria. MODM methods, on the other hand, are applicable to situations in which a large number of feasible alternatives prevail and the objectives functionally relate to the decision variables. MODM methods therefore facilitate the selection of a best solution from a set of high-quality solutions determined from a larger set of initial alternatives, whereas MADM methods commence with a smaller, known set of alternatives [48, 117].

MCDM methods can also be classified according to the quality of available information in view of the complexities associated with modelling a real-world situation which often involves imperfect knowledge and evaluations from humans. It has been suggested that one should categorise information as *Crisp* (*i.e.* pertaining to precise data) or *Fuzzy* (*i.e.* pertaining to incomplete or vague data). MCDM methods can then further be divided into MODM/MADM methods when utilising crisp data, or into *Fuzzy MODM/Fuzzy MADM* when considering fuzzy knowledge [124].

Another classification method involves categorising MCDM methods according the following three categories: (i) *multi-attributive value theory* (MAVT), (ii) *outranking methods* and (iii) *interactive methods* [69, 124, 141, 180].

MAVT is a unique synthesis criterion approach without incomparability which aggregates the different perspectives or points of view into a unique value function. This can then be assessed to find the best solution [3, 180]. It therefore transforms the scores of alternatives given at any level of evaluation into value functions, based on the preferences of an unitary decision maker. The MAVT approach facilitates rational decisions, with highest expected value also being the preferred alternative [74].

Outranking methods are based on the underlying assumption of incomparability among alternatives and involve the development of a relation which represents the strongly established preferences of a decision maker in order to facilitate selection of an alternative [180]. The overall attractiveness of each alternative is determined by weighting the scores given to specific dimensions relative to other alternatives. In contrast to MAVT, the calculated scores are utilised to steer a deliberative process amongst multiple stakeholders, rather than identifying a single solution [74].

Interactive methods utilise a collective local judgement approach involving a trial-error iterations [69]. 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 [180].

4.2.2 Preference modelling

Many preference relations exist, but the focus in this section is only on those utilised in MCDM methods. The preference structures, as discussed by Colson and De Bruyn [36] and by Guitouni and Martel [69] are therefore considered. Assuming that a decision maker must compare two alternatives denoted by a and b, four possible preferences can result in the form of the following binary relations:

An *indifference situation* (denoted by $a\ \mathbf{I}\ b$): The decision maker is indifferent between alternatives a and b. There is thus no evidence that alternative a is preferred to alternative b, and vice versa.

A *preference situation* (denoted by $a\ \mathbf{P}\ b$): The decision maker strictly and strongly prefers alternative a to alternative b, without any hesitation.

A *weak preference situation* (denoted by $a\ \mathbf{Q}\ b$): The decision maker strictly but weakly prefers alternative a to alternative b. There is, however, hesitance in terms of whether to express indifference or preference.

An *incomparability situation* (denoted by $a\ \mathbf{R}\ b$): The decision maker is hesitant between $a\ \mathbf{P}\ b$ and $b\ \mathbf{P}\ a$. These alternatives are seen as incomparable. This may be due to a lack of information.

These elementary preference relations may be combined by means of logical operators to form performance structures (*e.g.* $a\ (\mathbf{P} \cup \mathbf{I})\ b$, which combines the preference and indifference situations). An outranking relation, denoted by $\mathbf{S}$, is obtained by combining $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{I}$ in the form of $\mathbf{S} = \mathbf{P} \cup \mathbf{Q} \cup \mathbf{I}$ [36, 69]. In light of this, the expression $a\ \mathbf{S}\ b$ implies that alternative a is at least as good as alternative b.

4.2.3 MCDM problem formulation

Guitouni and Martel [69], Mota *et al.* [124], Ölçer and Odabaşi [125], and Zanakis *et al.* [185] adopted the following general MCDM problem formulation. Let $\mathcal{A} = \{a_1, \ldots, a_i, \ldots, a_n\}$ denote a set of alternatives when analysing a decision space and let $\mathcal{C} = \{c_1, \ldots, c_j, \ldots, c_u\}$ denote the set of criteria or attributes pertinent to the decision at hand. A rating e_{ij} is associated with the i^{th} alternative in respect of the j^{th} criterion. The MCDM problem is therefore succinctly captured by the following $n \times u$ *performance table* or *decision matrix*

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

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 criterion j, with the sum of the weightings summing to one.

4.2.4 Multi-criteria aggregation procedures

A number of multi-criteria aggregation procedures exist and form part of the various MCDM methods. Guitouni and Martel [69] 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 facilitate selection of a preferred alternative. Due to the simplicity of the general approach and associated analysis, these methods are suited to problems in which a single decision maker prevails [107]. Weightings of criteria are not required, although multiple criteria may be present. 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 and are discussed in detail by Guitouni and Martel [69], Linkov *et al.* [107], MacCrimmon [111], Massam [117] and Steynberg [161].

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 criteria into a single comprehensive index. These methods also admit fuzzy theory, imperfect knowledge and consistency to a certain extent. Prominent single synthesis criterion methods, discussed by Akbulut [3], Guitouni and Martel [69], Steynberg [161], Velasquez and Hester [179], and Zanaksis *et al.* [185], include the *technique for order by similarity to ideal solution* (TOPSIS), MAVT, *utility theory additive* (UTA), MAUT, the *simple multi-attribute rating technique* (SMART), the *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 method and the *preference ranking organization method for enrichment evaluation* (PROMETHEE) followed [69]. Further outranking methods include MELCHOR, ORESTE, REGIME and the *novel approach to imprecise assessment and decision environments* (NAIADE). Akbulut [3], Colson and De Bruyn [36], Guitouni and Martel [69], Velasquez and Hester [179], and Zanakis *et al.* [185] discuss these methods in detail.

The mixed methods, QUALIFLEX [3, 69], the *fuzzy conjunctive and disjunctive method* [69, 107] and the *Martel and Zaras method* [69] cannot be precisely categorised according to the three categories previously mentioned, but are also often used in an MCDM context.

Traditional MCDM methods consider crisp sets of data (*i.e.* ordinal information) as opposed to uncertain or vague data, although ambiguity and vagueness are frequently apparent in real-world decision-making problems. Fuzzy MCDM methods have therefore been developed to accommodate the fuzzy data sets within these problems [69, 125].

Despite the great variety in MCDM methods, no one method can be considered appropriate in all decision-making situations [69]. It is important to apply a multi-criteria aggregation procedure which best suits the problem at hand. Guitouni and Martel [69], Mota et *al.* [124], and Velasquez and Hester [179] provide guidelines for 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 chooses the most appropriate alternative according to a set of criteria influenced by bias and individual preferences [31]. 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.* [43] found that the decision maker's intuition with respect to risk routinely leads to decisions which deviate from rationality.

A central theme in 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 [42]. Simon [150] first proposed this concept, stating that humans have physiological and psychological limitations that influence their rational choice and ability to arrive at optimal solutions. Linked to the concept of bounded rationality is the *satisficing model* first proposed by Simon [150], in which a decision maker (acting as a satisficer) only seeks a satisfactory solution owing to cognitive limitations. The satisficing model posits that the decision maker is satisfied by an alternative, even if it sufficiently meets only some criteria, since decision makers are incapable of maximising in most situations [42, 151].

Conjunctive or disjunctive models are decision-making methods that are also recognised as elementary MCDM methods. The *conjunctive model* is applicable to a set of potential solutions taken from a set of alternatives. The set of potential solutions includes all alternatives that exceed a certain threshold value associated with an evaluation function, while all alternatives which do not qualify are eliminated. Similar to the satisficing model, conjunctive or disjunctive models involve searching for an adequate alternative, rather than for optimal solutions. When minimising the evaluation function, a conjunctive model is used, whereas a disjunctive model is used when maximising the evaluation function [42].

Similar decision-making methods which include some form of rationality are the *additive model* and *additive difference model* regularly mentioned in the literature on descriptive decision-making methods [42]. These models provide good approximations to multi-attribute decision behaviour as they evaluate each multi-dimensional alternative independently. In the additive model, each alternative is considered and rated individually. These ratings are then compared holistically to determine the most appropriate solution. In 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 a 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 comparison purposes [42].

Cervantes *et al.* [31] followed a more multi-disciplinary approach in developing a comprehensive and coherent human decision-making model which encompasses an understanding of decision making within the domains of 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 values of these dimensions in order to identify the best alternative. The model also addresses issues that occur in modelling autonomous agents exhibiting human-like behaviour.

The use of heuristics has been suggested when considering human decision making, instead of strict, rigid rules of optimisation, owing to the complexity of a typical decision situation and an individual's rational inability [150]. Gigerenzer and Gaissmaier [61] define heuristics as a subset of strategies that choose to ignore information partially, with the aim of making quick, frugal and proper decisions. When employing heuristics, there exists a trade-off between experiencing 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 in which only one good reason is considered when making judgements, therefore ignoring other cues [61].

In nature, animal species tend to rely on a single 'clever' cue when searching for food, areas for nesting or mates. An example of such a 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 a constant optical angle between itself and the target is rather maintained. Another example where the one-clever-cue heuristic is employed occurs in geographical profiling [60, 61].

The take-the-best heuristic considers the manner in which one alternative is inferred to have a higher value in respect of some criterion than 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 terminated as soon as the first cue which discriminates between alternatives is found. It is then inferred that the alternative with the positive cue value possesses the higher criterion value [61]. The take-the-best heuristic is similar to a lexicographic model, as it evaluates alternatives considering only one attribute. Furthermore, it may be presented formally as a fast-and-frugal tree [91].

Representing Bayes' rule as a tree, there would be 2^m leaves, where m is the number of binary attributes or cues, which may lead to a computationally unmanageable model. In contrast, a fast-and-frugal tree representation of the same situation will have only $m + 1$ leaves, leading to a model that is more robust and tractable. The foundation rules of this heuristic are similar to those of the take-the-best heuristic, where cues are searched through in a predetermined order, stopping when some cue leads to an exit. Application examples in which fast-and-frugal trees have been utilised include the modelling of physicians' thinking in the field of emergency medicine, as well as in a legal context when having to decide on bail [61].

4.4 Chapter summary

In this chapter, an introduction was given to the field of decision making, as well as to a variety of modelling methods applicable to the context of this dissertation. Different methods for the modelling of human decision-making situations were mentioned. The research area of decision making was discussed in general 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, a variety of MCDM methods were discussed. Classification schemes for these methods in literature were discussed in §4.2.1, before considering the concept of preference modelling pertaining to MCDM and a general formulation of the MCDM problem in §4.2.2 and §4.2.3, respectively. In §4.2.4, various multi-criteria aggregation procedures in literature were mentioned. Finally, the concept of human decision making and the modelling thereof was discussed in §4.3. Various methods, such as the satisficing model, conjunctive or 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.

CHAPTER 5

Conceptual framework analysis

Contents

The notion of a conceptual framework analysis is reviewed in this chapter. A brief overview of existing research paradigms is given in §5.1, and this is followed by an overview of grounded theory in §5.2. In §5.3, the difference between conceptual and theoretical frameworks are highlighted. Lastly, the development of a conceptual framework is discussed in §5.4, which includes a discussion on conceptual framework analysis and various methodologies typically used, as well as the process of verification and validation within the context of a framework.

5.1 Research paradigms

Two dominant research paradigms, quantitative and qualitative, are evident in the literature when considering the different types of research problems, methodologies and data [24, 41, 46, 78, 87, 88]. These research paradigms differ with respect to epistemology, ontology, as well as the relationship between theory and research. Quantitative research approaches adopt a deductive approach as an emphasis is placed on the testing of theories and such approaches therefore use theory to deduce research. Qualitative research approaches, on the other hand, are predominantly inductive as they aim to generate, rather than improve, theory by means of research [24].

As research problems become more interdisciplinary, complex and dynamic, researchers have recognised the need to combine aspects of the aforementioned research paradigms [24]. Johnson and Onwuegbuzie [87], for example, discussed a *mixed methods research* paradigm, formally defined as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, concepts or language in a single study." This paradigm aims to augment the inherent strengths of the qualitative and quantitative research approaches, while also reducing the weaknesses of the two approaches. Mixed methods research allows for a creative approach to research that is expansive, pluralistic and may prompt researchers to be inclusive when selecting methodologies and conducting research [87, 88]. It incorporates several overlapping classes, as illustrated in Figure 5.1.

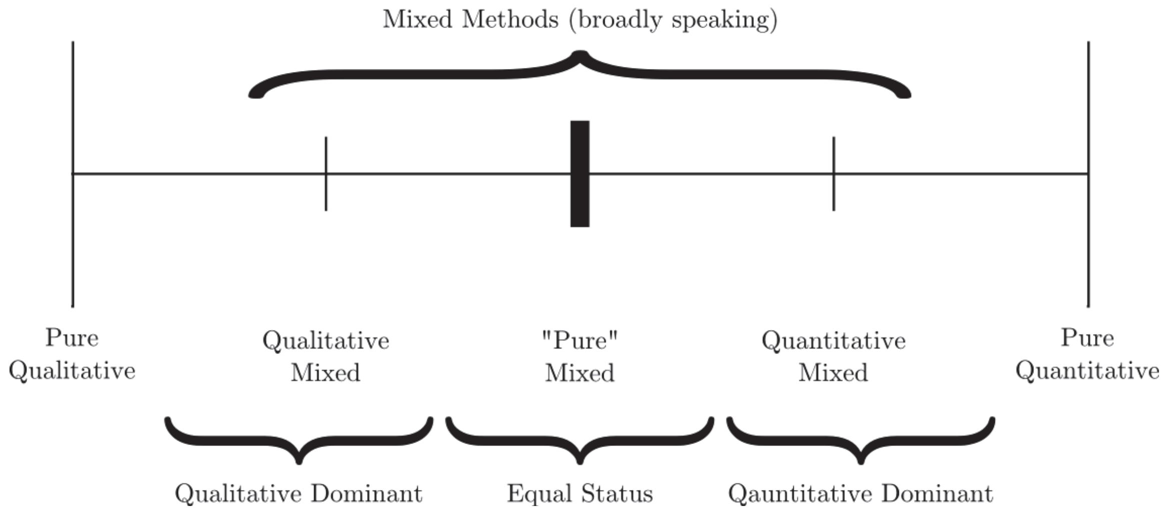


FIGURE 5.1: *Three major research paradigms and subtypes of mixed methods research, as adapted from Johnson et al. [88].*

When referring to the qualitative-quantitative continuum in Figure 5.1, the area within the scope of mixed methods is focussed at the centre and expands towards the qualitative and quantitative sides. On the far-left side of the continuum is the *pure qualitative* approach, while the *pure quantitative* approach is on the far-right side. Around the centre of the continuum is the *pure mixed* approach, also known as the *equal status* approach. A researcher within this domain utilises both qualitative and quantitative approaches equally. The *qualitative dominant* approach is a type of mixed methods research that relies on a qualitative approach while in addition recognising the benefit of utilising quantitative approaches. Another type of mixed method research along the continuum in Figure 5.1 is the *quantitative dominant* approach, which refers to research relying on a quantitative approach while concurrently recognising the benefits of additional qualitative approaches [88]. An example of a methodology within the qualitative dominant spectrum is grounded theory [49].

5.2 Grounded theory

Glaser and Strauss [62] coined the term *grounded theory* as a general methodology that will "fit the situation being researched and work when put into use. By fit we mean that the categories must be readily (not forcibly) applicable to and indicated by the data under study; by work we mean that they must be meaningfully relevant and be able to explain the behaviour under study."

Grounded theory is aimed at investigating and explaining a system or behaviour and can therefore be seen as a methodology for generating theory that is grounded in data gathered and analysed [38, 162]. Social systems are examples of systems which may thus be analysed to discover patterns and processes indicating how individuals define their reality by means of social interactions [158].

It is primarily an inductive methodology as it requires the researcher to investigate the research problem from the research participant's perspective (*i.e.* the data) before generating new theory based on systematic analysis of the data [162]. It has been argued by Elliot and Higgins [49] that grounded theory is both inductive and deductive in nature, although primarily inductive. This is evident in certain methods pertaining to grounded theory, such as theoretical sampling and comparative analysis, where emergent concepts are tested throughout the research process. Comparative analysis of extensive interrelated data collections is performed as a continuous verification technique. This enhances the possibility of developing theory of a suitable conceptual density (*i.e.* richness of concept development and relationships) [162].

5.3 Conceptual and theoretical frameworks

In the context of research, a framework is an outline that serves as a guidance to the researcher with respect to formalising research questions, identifying variables and planning analysis methods [105]. The aim of a structured framework is to serve as a guidance document which can aid and support interdisciplinary role players in undertaking a conceptual design. A shared understanding of the design process amongst the interdisciplinary role players can further aid the decision-making process during the development phase and increase the likelihood of successful collaboration [112].

A theoretical framework refers to the application of a theory or concepts drawn from one theory to explain a certain phenomenon. A conceptual framework, on the other hand, consists of a synthesised version of theoretical and empirical findings concerning a certain phenomenon [78]. Both of these types of frameworks facilitate an understanding of a phenomenon within a particular field of study and serve as a structure for guiding a researcher in investigating a research problem [105]. In general, a study in the field of natural sciences tends to be guided by a particular theory, *e.g.* the theory of evolution, when performing an investigation, whereas studies within the field of social sciences cannot be meaningfully investigated by means of a single theory when researching, for example, the challenges of poverty [78]. Most social phenomena are complex by nature and therefore a conceptual framework may serve as an adequate tool for investigations into these complex systems [85]. The typical relationships that prevail between theoretical and conceptual frameworks relative to qualitative and quantitative research paradigms are illustrated in Figure 5.2.

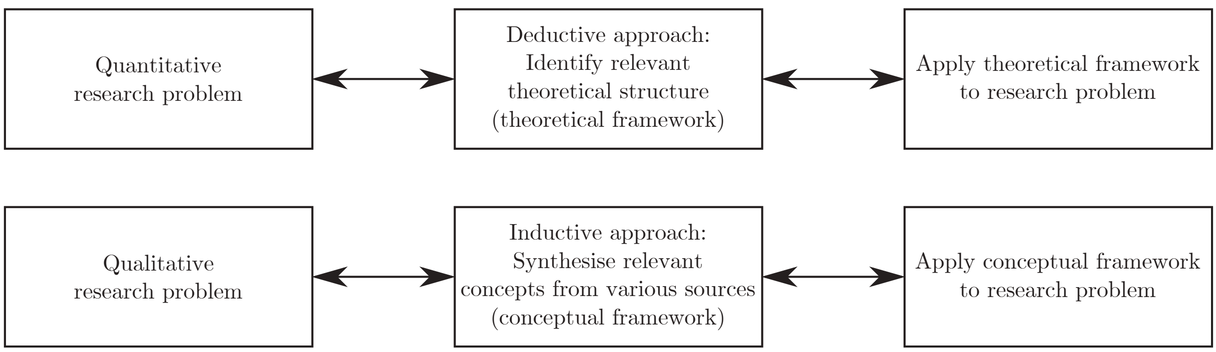


Figure 5.2: *The relationship between theoretical and conceptual frameworks relative to qualitative and quantitative research paradigms, as adapted from Imenda [78].*

A conceptual framework is developed by means of an inductive process during which various concepts are combined to indicate the relationships between concepts [78]. Jabareen [85] defined a conceptual framework as a network of interlinked concepts in which each plays an integral role in providing insight and knowledge into a particular phenomenon. This suggests that although induction takes place during the derivation of concepts from data, deduction is employed to hypothesise the relationships between concepts [128]. Ritchie and Spencer [139] summarised the key features of a conceptual framework as grounded or generative (*i.e.* based upon data, such as suggested by grounded theory), dynamic, systematic (*i.e.* methodological), comprehensive, enabling easy retrieval of original material, allowing between- and within-case analysis (for comparative analysis), and accessible to others.

5.4 Development of a conceptual framework

In developing a conceptual framework, it has been suggested that a framework analysis be employed, which is fairly similar to adopting a grounded theoretic approach [157]. Although there are similarities, a framework analysis differs from grounded theory in that its primary focus is to describe and interpret a particular setting rather than only generating theory. It is therefore better suited for research in which specific research questions are posed and a limited time frame, a pre-designed sample and *a priori* issues are evident [139, 157]. Some of the main considerations when developing a generic framework for conceptual design are flexibility and interdisciplinary interaction, as the framework needs to take into account the dynamic nature of the phenomena investigated [112]. Carroll and Swatman [27] advised that the process of developing theory should not be approached in a linear fashion, as it may entail challenging decisions and an overwhelming amount of data which require a recursive or iterative research approach with continuous data comparison and refinement.

5.4.1 Conceptual framework analysis

Ritchie and Spencer [139] developed the notion of framework analysis as a qualitative methodology in the context of applied policy research, where actionable outcomes are provided as a result of specific information gathered and analysed. A key capability of framework analysis is the flexibility with which data are analysed — a process that may occur after or during collection thereof. Ritchie and Spencer [139] explained the analysis process in terms of five stages: Familiarisation, thematic framework identification, indexing, charting, and, lastly, mapping and interpretation.

The first stage, familiarisation, entails the researcher becoming familiarised with the data collected. A thorough study of a carefully selected data collection will allow for key and recurring themes to be noted which, in turn, will lead the researcher to identify a thematic framework present (the second stage). This thematic framework is only tentative as it is refined throughout the analysis process. The initial development of the thematic framework, as well as the refinement thereof, requires logical and intuitive thinking. Following this stage, is the indexing of data which may be performed using numerical systems. It comprises the identification of clusters of data that correspond to particular themes. During the next stage, charting, the data indexed during the previous stage are arranged according to the thematic framework. Finally, key characteristics of the charts developed are analysed during the mapping and interpretation stage. A schematic diagram of the phenomenon may result from this analysis which will allow for interpretation of the data set.

5.4.2 A conceptual framework analysis with grounded theory

Jabareen [85] proposed a conceptual framework analysis that utilises grounded theory techniques with a dual aim: First, to generate, identify and trace the primary concepts of a phenomenon that form part of the theoretical framework, and, secondly, to develop concepts (including the attributes, assumptions and functions thereof within the conceptual framework) pertaining to the phenomenon under consideration.

The data selected for the framework analysis should appropriately represent the phenomenon and associated behaviour, whether of a social, cultural, political or environmental nature. In the case of multi-disciplinary approaches, the theories developed by specific disciplines may develop into empirical data functional to the conceptual framework analysis. The conceptual framework analysis proposed consists of an iterative process during which the data and developing concept are constantly examined and this comparison controls the conceptual development of the generated theory. This technique relates to grounded theory as it suggests a continuous interaction between the data collected and its analysis.

Jabareen's qualitative methodology for developing a conceptual framework comprises eight phases, as illustrated in Figure 5.3. Phase one requires the researcher to map the sources of data within the multi-disciplinary spectrum within which the phenomenon exists. This requires an extensive collection and review of the data, including interviews with academics and practitioners from the relevant disciplines. The next phase entails the researcher reading and categorising the selected data based on the various disciplines and their representative weights, as well as their levels of importance. During phase three, reviews of the selected data are revised in order to recognise concepts. These concepts may be competing or even contradictory in nature. Phase four aims to deconstruct the concepts previously labelled which will allow the researcher to identify the main attributes and characteristics pertaining to each concept before categorising the concepts according to features and roles. The output of this phase may include a table providing the name, description, role (*i.e.* ontological, epistemological or methodological) and reference of each of these concepts.

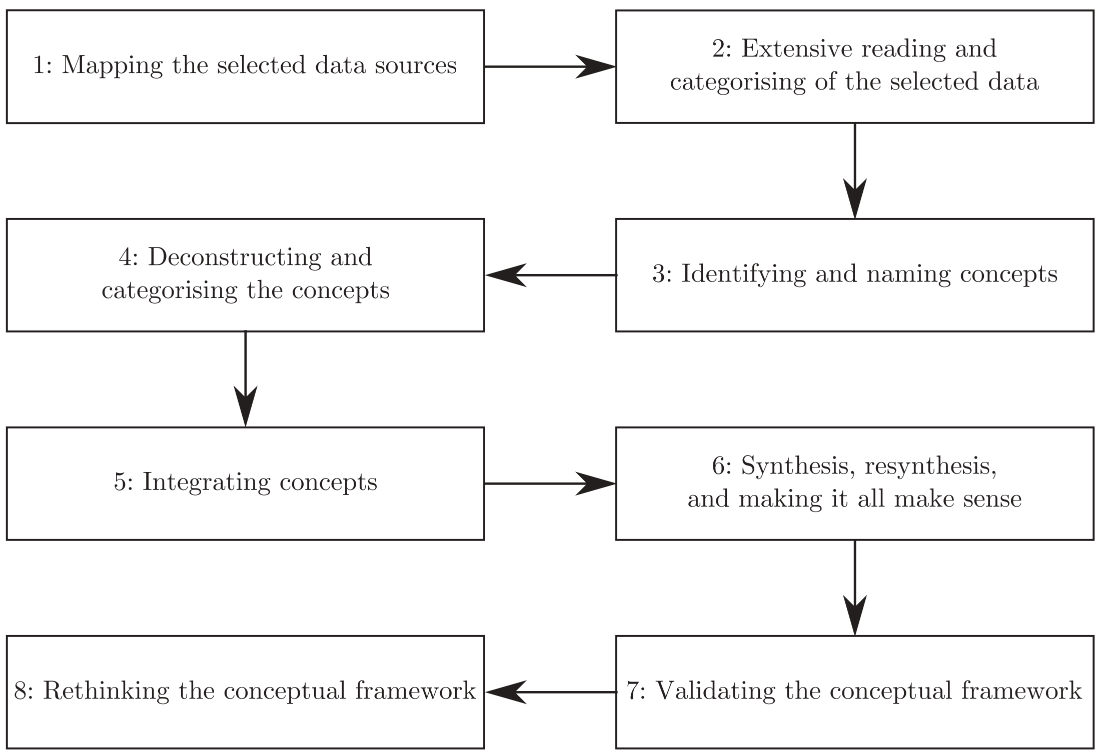


FIGURE 5.3: *The phases pertaining to the conceptual framework analysis proposed by Jabareen [85].*

Phase five aims to integrate and group the various concepts so as to enable the researcher to reduce the number of concepts to a reasonable number. This is followed by phase six, during which a theoretical framework is formed by synthesising the concepts. Jabareen [85] advised that the researcher should be tolerant of the theorisation process during which new theories emerge as it is iterative and repetitive. The concepts should be synthesised and resynthesised until the researcher recognises a generic framework. During phase seven, the conceptual framework developed is validated. The validation should include, in particular, validation by academics and practitioners within the various fields encapsulated by the phenomenon. Finally, phase eight encourages the researcher to revise the conceptual framework developed as the framework should be dynamic, particularly when involving various disciplines, and adaptable to include new insight and literature.

The aforementioned conceptual analysis framework is a qualitative method which uses grounded theory with the aim of providing assistance in developing conceptual frameworks that aid in better understanding complex, multi-disciplinary phenomena.

5.4.3 A structured-case methodological framework

Carroll *et al.* [26] coined the term *structured-case methodological framework* for generating theory, a notion that may enable a deeper understanding of the complex interactions between people, processes and technology within an organisation. The methodology can be classified as *structured* in that it entails three structural components, namely a conceptual framework, a research cycle and a thorough literature review which, in combination, enables the researcher to generate theory.

The conceptual framework component acts as the theoretical foundation of the researcher, while the research cycle governs the collection, analysis and interpretation of data. When combined, these two components allow for the dynamics of a particular phenomenon and the processes pertaining to it to be documented which, in turn, induces theory from data. The final component, which involves a scrutiny of the literature, enables the researcher to compare the outcomes of the first two components in order to support and assert the theory developed [27]. A graphical illustration of the structured-case methodology is provided in Figure 5.4.

The conceptual framework

A key element of the structured-case methodology is the conceptual framework which, in essence, is the researcher's description of the formally defined conceptual structure of the phenomenon [27]. It is formed by inputs such as research themes, the literature and insights pertaining to the phenomenon which are filtered by the theoretical foundations of the researcher. The conceptual framework should portray key concepts and relationships within the given research paradigm and is defined not only at the initiation of the research. The researcher should critically examine the existing conceptual framework at the end of each research cycle and revise it to include information or understanding gained of the research themes. The final conceptual framework is therefore the result of a series of progressing models which were continuously reviewed and revised through the research project.

The research cycle

The structured-case research cycle allows for a better understanding of the research themes to be gained by iteratively subjecting the conceptual framework previously defined to four design stages, namely planning, data collection, analysis, and reflection. These stages do not necessarily

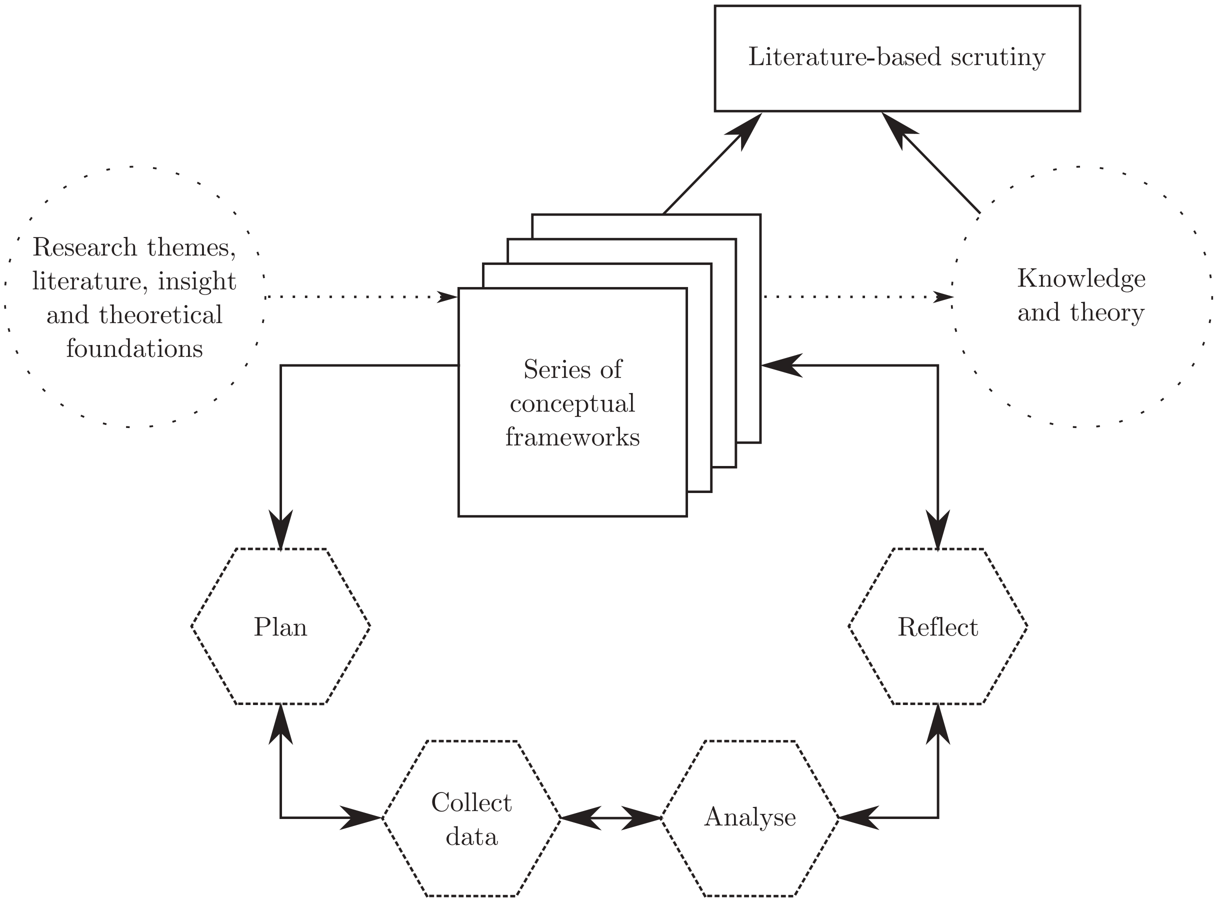


Figure 5.4: *The structured-case methodology, as adapted from Carroll and Swatman [27].*

have to occur in a sequential pattern; it has, in fact, been advised that frequent and continuous movement occurs between these stages [27]. In Figure 5.4, the research cycle is depicted by the four stages which cycle through the series of conceptual frameworks.

The *Plan* stage concerns the investigation of research themes and the selection of a research design by considering concepts and relationships pertaining to the research phenomenon. The use of various data collection and analysis methods is also planned within this stage, as is the method according to which the outcomes should be reported. The *Collect* stage refers to the collection of data, while the *Analyse* stage refers to the analysis of the data collected. These stages may be closely interrelated and, although overlaps between data collection and analysis are evident, the data analysis activities continue even after the data collection activities have been completed. The conceptual framework should guide the collection and analysis of data, while allowing the outcomes thereof to lead to emerging concepts and a better understanding of the phenomenon. The *Reflect* stage involves intentional and critical reflection aimed at reviewing the research process followed, evaluating the outcomes, reviewing the various components within the structured-case, building theory from the findings and revising the conceptual framework to incorporate this theory. The current theory then serves as the basis for another research cycle which may further expand the understanding of the phenomenon.

Reflection during the research cycle enables the researcher to probe further than the data and to consider ideas, relationships between concepts, and emerging patterns. The researcher should be able to clarify and categorise concepts, and indicate the relationships between categories in order to induce theory — the research cycle aids in this process. Although each progressing conceptual framework and research cycle may induce further knowledge and understanding, the researcher will at some point (when a sufficient understanding of the phenomenon has been gained) cease the pursuit.

The literature-based scrutiny of emerging theory

The final step in the structured-case methodology entails comparing and contrasting the conceptual framework developed with the existing literature. This step differs from the reflection stage of the research cycle, in that it involves vigorously and thoroughly comparing the theory generated with an extensive range of literature. Two key aspects to consider are the extent to which correlation and conflict can be found between the literature and the findings. Correlation indicates that the emerging theory is replicating, consolidating or expanding on existing theory, while conflict may arise from differing interpretations and should encourage further investigation. In scrutinising the theory that has emerged, a critical reassessment of the research findings may develop which may, in turn, increase the usability and applicability of the theory to other contexts [27].

The structured-case methodology serves as an effective roadmap to assist researchers in generating theory in respect of phenomena that are complex and difficult to comprehend. It operates at a high level of abstraction and it may be used to extend existing frameworks, which may allow for greater refinement, usability and effective representation of a said framework. The methodology in its entirety aids in constructing preliminary conceptual frameworks, collecting and analysing data, and comparing outcomes with existing literature with the aim of developing theory.

5.4.4 Verification and validation of a conceptual framework

Verification, within the context of a generic framework, refers to the internal evaluation of the framework integrity and logic. It involves evaluating the various components within the framework and determining whether they accurately characterise the elements of the conceptual framework. Validation, on the other hand, involves an external evaluation to determine whether the conceptual framework represents reality by applying the framework to a real-world case. Developing a concept demonstrator may allow for validation to occur, as well as critical feedback for further development and improvements to the preliminary version of the conceptual framework to be received [85, 112].

5.5 Chapter summary

The notion of conceptual framework analysis and various aspects pertaining to it were discussed in this chapter. In §5.1, the research paradigms, qualitative *vs.* quantitative, were discussed, along with the use of mixed methods in which a combination of these two paradigms may co-exist. The notion of grounded theory was discussed in §5.2, and this was followed by a discussion on conceptual and theoretical frameworks in §5.3. The heart of the chapter was §5.4, where the development process of a conceptual framework was discussed. The discussion revolved around conceptual framework analysis as a qualitative methodology and, in particular, a proposed conceptual framework analysis based on grounded theory. This was finally followed by an elucidation of the structured-case methodology, and a discussion on the processes of verification and validation of conceptual frameworks.

Part II

The Conflict-induced Forced Migration Modelling using an Agent-based approach (CoFMMA) framework

CHAPTER 6

A generic framework for modelling conflict-induced forced migration using ABM

Contents

This chapter contains the main contribution of the dissertation — a generic framework proposed for agent-based modelling in the context of conflict-induced forced migration. A background to the framework is given, and this is followed by a detailed discussion on each phase of the framework and its various components.

6.1 The process in developing the generic framework

A generic framework is proposed in this section for aiding in the design and development of agent-based models for simulating conflict instances, populations and the movement patterns of IDPs, refugees and undocumented migrants based on their localised decision-making processes. The framework development process followed is based on the structured-case methodological framework discussed in §5.4.3.

Initially, the main modelling themes within the research problem of forced migration modelling were considered and a first draft conceptual model was created, as illustrated in Figure 6.1. The main modelling themes considered are conflict, population and decision making. This initial framework follows a silo bottom-up approach during which the main conceptual elements are investigated and modelled individually before considering interaction between the primary components. The first component concerns the nature, initiation and spread of the conflict, while the second component is the population density, fluctuation and distribution. Finally, the third component is the decision making of individuals in respect of movement.

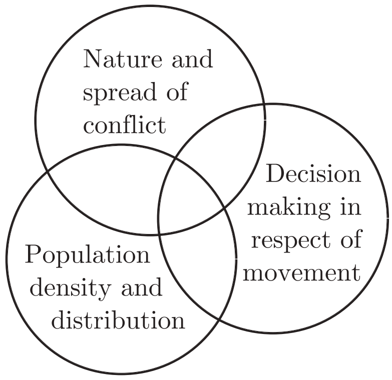


Figure 6.1: *The initial conceptual framework.*

Each of the primary research themes were subjected to a research cycle which entailed further investigation into the particular theme and its relationship with the other themes in the greater modelling paradigm, and the collection and analysis of data pertaining to the theme. This was followed by an overall reflection on the resulting conceptual framework draft. Upon critical reflection, interwoven sub-concepts between the main modelling themes were identified which led to a more adhesive framework.

In order to develop an understanding of the intertwined sub-concepts, further research was conducted which led to a second conceptual framework draft based on the steps proposed for consideration during the design and development of a simulation model by Banks [14], as discussed in §3.1.4. The adapted version of the conceptual framework is illustrated in Figure 6.2. This framework follows a more structured approach in which three main themes, namely conflict, population and decision making, are embedded in and integrated with a number of consecutive framework phases and are addressed continually throughout the framework rather than in a once-off fashion.

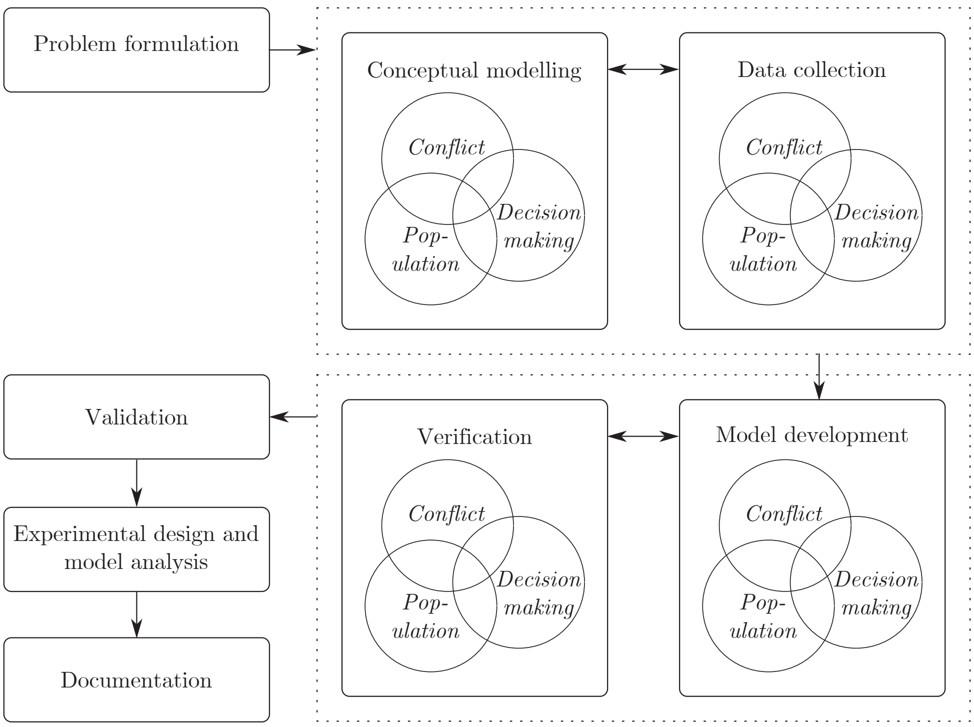


Figure 6.2: *The second conceptual framework.*

Eight phases are considered in this conceptual framework draft, commencing with the phase of problem formulation. The combination of phases two and three, which follow on phase one, allows for an iterative loop between conceptual modelling and data collection before progressing to phase four, namely model development. An iterative loop is again employed between the model development phase and verification phase (phase five). Phase six is validation, followed by experimental design and model analysis (phase seven), and, finally, by documentation (phase eight).

Another research cycle was completed with respect to the second conceptual framework developed. Addressing the research themes simultaneously instead of individually during this cycle, allowed for the capturing of overlaps between concepts during the model development process. A separate phase dedicated to data collection was initially included. After critical reflection it was found, however, that data collection and conceptual modelling should be captured together in one phase, rather than in separate phases, but allowing for iteration within this phase. Similarly, it was found that merging the model development and verification phases, as well as the validation, experimental design and model analysis phases, allows for better integration and model development. The development of a *graphical user interface* (GUI) was further added as a modelling aspect. A diagrammatic illustration of the third conceptual framework draft is shown in Figure 6.3.

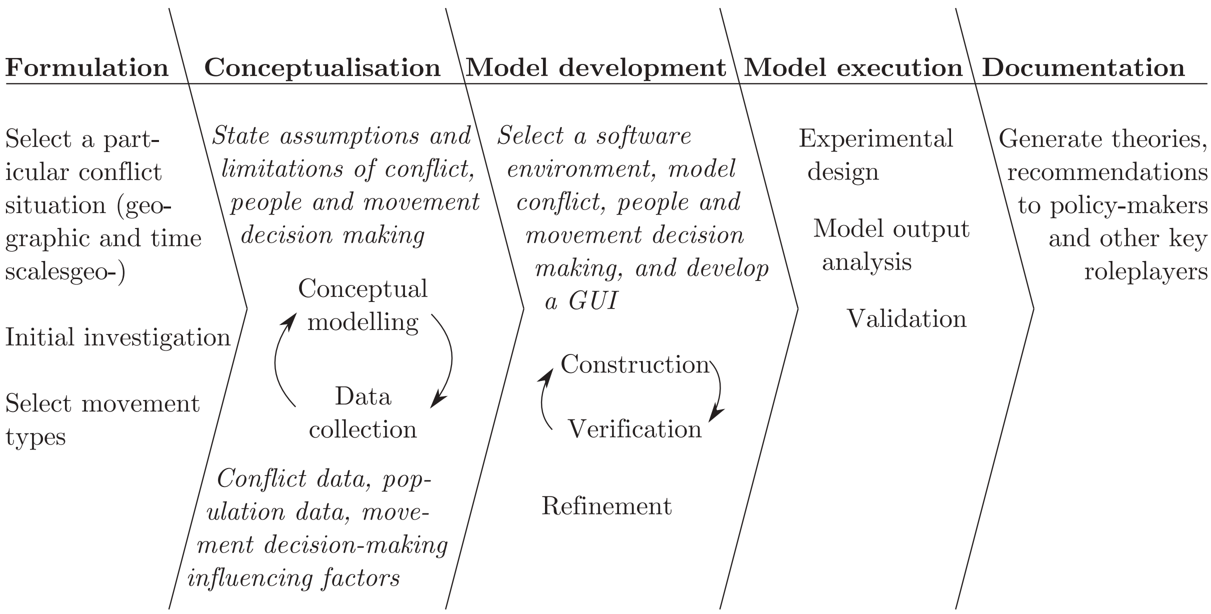


FIGURE 6.3: *The third conceptual framework.*

Still further reflection led to another amendment to the third conceptual framework draft — an iterative cycle which may occur between the conceptualisation and model development phases. This iterative process allows the modeller to add complexity to the model iteratively, as well as to make adjustments to the initial conceptual design if the subsequent model development requires it.

The final framework draft, hereafter called the *Conflict-induced Forced Migration Modelling using an Agent-based approach* (CoFMMA) framework, was subsequently subjected to literature-based scrutiny and the development of concept demonstrators, after which it was found to hold real potential in respect of generating knowledge and theory in the field of forced migration modelling. The CoFMMA framework is discussed in some detail in the remaining sections of this chapter.

6.2 The CoFMMA framework

The CoFMMA framework comprises five phases, namely formulation, conceptualisation, model development, model execution, and documentation. The framework, the primary components within its phases, and the flow between these phases, are illustrated in Figure 6.4.

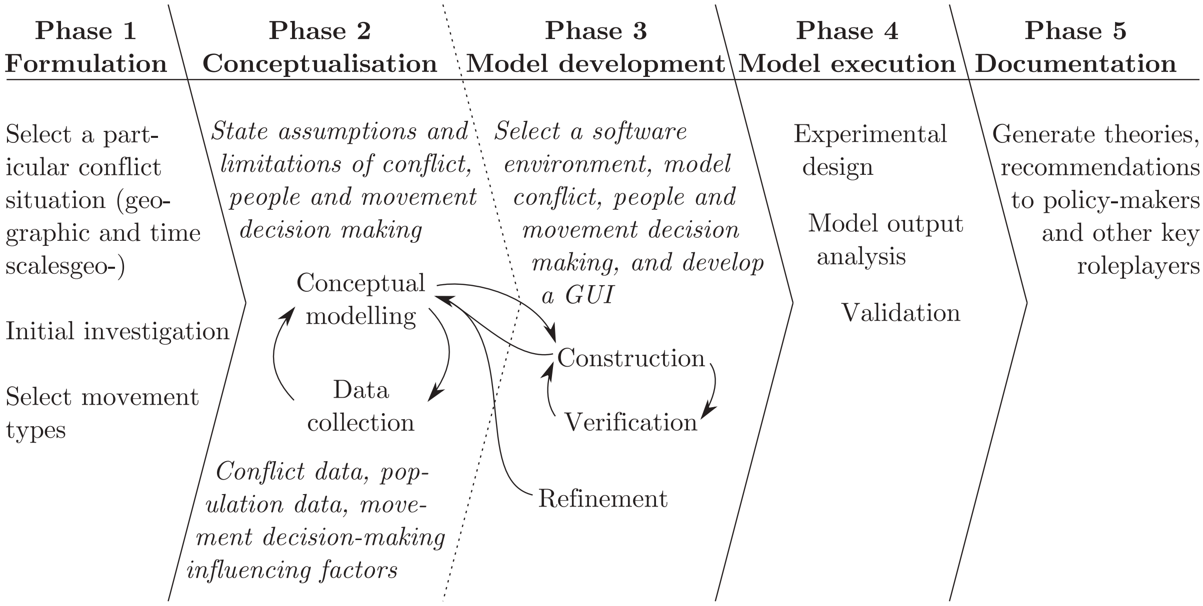


Figure 6.4: *The CoFMMA generic framework.*

The framework is focused on conflict-affected areas in which forced migration is observed and assumes the use of agent-based simulation as the underlying dynamic modelling approach. In utilising the framework, an agent-based model may be developed with the capability of capturing emergent large-scale movement patterns of forcibly displaced people by modelling their individual or group movement decision making on a local level in order to gain an understanding of the phenomenon of migration pattern emergence on a more global level, to predict anticipated destinations of these individuals, and to assist in humanitarian support decisions. Any simulation model to be developed using this framework will be of a dynamic nature in that it will have the ability to simulate conflict instances and associated population movement pattern changes over a given period of time. These changes over time are characteristic of a continuous simulation model, as changes occur continually rather than at specific time instants. The model to be developed may also be stochastic in its behaviour since it is likely that not all input variables will be definitively known and, therefore, calling for probabilistic estimates to be utilised instead.

With respect to the modelled level of detail, a medium level of abstraction is suggested, characteristic of a model that is more tactical than operational or strategic in nature. The level of detail may, however, be determined by the modeller and, as suggested by the characteristics of the agent-based approach, may vary between low and medium levels of abstraction. Agents may, for example, represent individuals when modelling at a lower level of abstraction, or may represent aggregations of individuals with similar attributes when modelling at a medium level of abstraction. The agents defined should be capable of making decisions, in particular, with respect to when and where to move in the presence of conflict. Furthermore, agents may be able to communicate with others and move within the modelled environment based on decisions made in view of their characteristic profiles, decisions made collectively in concert with other agents, and the state of their surroundings.

6.3 Phase 1: Formulation

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 is scoped sufficiently narrowly and is developed at an appropriate level of detail [20]. Within the first phase of the CoFMMA framework, the modeller is required to identify and select a particular conflict situation to which to apply the framework. The modeller should gain a background understanding of the particular conflict situation, including the historical progression of the conflict, as well as the state of the conflict situation during a selected time period. The geographic and time scales of the model also have to be determined as these will influence the level of detail incorporated into the model and, in effect, determine the computational capability required for its implementation. The modeller should keep in mind any limitations with respect to computational power which may lead to either a reduction in geographical scale, or the aggregation of agents. The modeller further has to consider which types of forced migration movement to include in the study. Clarity with respect to the model's purpose and scope will aid in progressing to the next phases of the framework.

6.3.1 Initial investigation

The conflict situation to be modelled should be investigated by the modeller to allow for the development of a better understanding of the situation, the entities involved and the extent of the conflict. The conflict situation selected might be existing, in which case the modeller would need to review the relevant literature and gain insight into the historical characteristics pertaining to the conflict, *e.g.* how it initiated, who the influencing and affected entities are, and what its effect has been. Alternatively, the modeller may want to model potential conflict outbreaks, in which case the factors that may influence the potential outbreak should be considered. This step would assist the modeller in defining the main objectives and scope of the study.

6.3.2 Geographic and time scales

In identifying a particular conflict situation to investigate, the geographic and time scales need to be defined broadly. The geographic area of investigation may entail a city, a governorate or region, a particular country, or even multiple countries. Of course, the level of abstraction of the model will vary based on the extent of the geographic area. Modelling movement in and around a city may, for example, be accomplished at a low level of abstraction, allowing for a considerable amount of detail to be captured, while movement of people between countries might require a higher level of abstraction. The level of detail of the model (refer to §3.1.2) therefore depends on the conflict situation selected and the geographic scale pertaining to it. Furthermore, the particular time scale in which the conflict situation is to be modelled may vary, spanning hours, days, months or even years, and this will affect the required computational ability of any platform on which the model is implemented. The geographic area and time scale would further affect the movement options of the forcibly displaced that can be modelled.

6.3.3 Selecting movement types to investigate

The three primary movement types that can be modelled, are evacuation modelling, refugee modelling, and migration modelling, as discussed in §3.3.1. These movement types differ in their temporal and spatial scales, and the modeller should identify for which movement type the geographic and time scales are selected, as discussed in §6.3.2. Most conflict-induced forced migration studies fall within the realm of refugee modelling, which may span a spatial scale of hundreds of kilometres and days, months or years as temporal scale. The modeller furthermore needs to identify which movement subtypes the agents will be allowed to follow in the model. A typology of these forced migration movement subtypes were discussed in §2.2.

It is proposed that agents fleeing conflict areas should be categorised as doing so according to three movement subtypes: (1) *IDPs*, (2) *Asylum-seekers and refugees* (these may be seen as one or separate categories), and (3) *Undocumented migrants*. IDPs are individuals who were forcibly displaced, but chose to remain within the borders of their countries of origin, whilst refugees are those who chose to cross the borders of their countries of origin legally, either by applying for asylum in some destination country, or being registered as refugees living in refugee camps. The category of undocumented migrants are those people who choose to cross the borders of their countries of origin without following legal and documented procedures. Another movement consideration, in particular when modelling over a larger time scale, is the repatriation of people who have initially fled across the borders and may return to their countries of origin after some time.

6.4 Phase 2: Conceptualisation

There exist many challenges and important considerations when developing a model of social behaviour within the realm of conflict-induced forced migration. These include the gathering of qualitative and quantitative data, the quantification of qualitative data and the understanding of motives and constraints which affect the movement of a population. The second phase of the CoFMMA framework is model conceptualisation which consists primarily of an iterative flow between conceptual modelling and data collection. The modeller is required to model the real-world scenario conceptually by stating assumptions and limitations of the agent-based model to be developed pertaining to the modelling of conflict, the modelling of the particular population, and the modelling of movement decisions made by the people. These assumptions and limitations to consider during the modelling process should be based on qualitative data and insight, as well as the quantitative data available. Within the conceptualisation phase, the modeller should consider various data inputs, agent definitions and attributes, conflict, population and movement considerations, while further detailing the geographic and time scales of the simulation model.

6.4.1 Geographic considerations

The modeller should determine the level of detail to be included within the geographic scale selected, as explained in §6.3.2. The borders of the physical geographic area that are to be considered should be identified and the level of detail that the model has to capture should be specified. In the case where a city is modelled, the suburbs to be included in the model should be identified and the level of detail may allow for GIS incorporation, so as to facilitate the tracking of exact path movements and traffic. In Northern Africa, for example, the modelling area may include a particular country and consideration of the specific terrain of that country. Terrain may include elevation, slopes, dessert areas, rivers or woods. This should, however, only be considered in a model where the terrain might directly influence the movement of the population at hand. It should also be noted that the inclusion of a GIS component may require more computational power.

6.4.2 Time considerations

The assumptions related to the time scale selected for the modelling of a specific scenario (as discussed in §6.3.2) should be detailed by the modeller. The time component of the model may entail a specific date and time, such as when modelling the exact date and time of historical conflict situations. It may, however, also be unrelated to a specific date or time, simply referring to time as a unit, such as when considering a future conflict event which may occur at an uncertain date or time. In the former case, the date and time at which the simulation run is to be executed should be noted by the modeller, as well as the date and time at which the simulation run execution is to be terminated. The period for which the simulation model will be running may be known beforehand, in which case the simulated time period is fixed, or the model may be developed to terminate when certain stopping conditions are satisfied, in which case the simulated time period is unknown. Another consideration with respect to time is selection bias, a phenomenon that can easily undermine the validity of research when unduly restricting or narrowing the range or variation of data [34]. In a case where data might be restricted, causing selection bias, a longer period of simulation is recommended as this may eliminate or minimise the effect of selection bias.

6.4.3 Defining an agent

A critical step within the conceptual modelling process is determining what or who should be represented by an agent and how many agent classes are to be included in the model. In its simplest form, an agent may represent an individual confronted with conflict. In the case of an extended geographic scale, the agent may, however, represent an aggregate of individuals. When adopting an aggregation approach, it is important to determine in which manner aggregation will take place. For example, an agent may represent individuals with similar characteristics, such as age or gender, or it may represent families or people of a similar network. If the modeller intends to follow an aggregation approach, initial investigation and qualitative data collection should assist in deciding upon the aggregation format. In modelling forced migration, an aggregate should consist of individuals with similar behaviour patterns in respect of conflict situations and movement decision making. The modeller may also include different agent classes, such as in the case where the aim is to model distinctly different groups exposed to a conflict situation.

Various attributes of an agent may influence when, why and to what extent an agent will move when confronted with a conflict-affected environment. Edwards [47] suggests considering attributes such as age, gender, wealth, ability to travel, ethnic identity and the like. The model should essentially only include agent attributes that influence agents' decision making with respect to movement. The level of detail may also limit which attributes can be included in agent construction, as social networks, for example, cannot easily be incorporated when one agent represents hundreds of people. The attributes of agents may also be specific to the conflict situation and the affected population. For example, in a conflict situation where parties of various religious convictions are involved, an individual may act differently towards the various parties based on their religions. In such a case, the religion of an agent may be useful to model as an attribute.

6.4.4 Movement considerations

The movement of agents as it will be modelled should be considered. The first consideration is the mode of transport that the agents will make use of, which may include walking, driving, using public transport, or flying. The mode of transport will determine the speed at which an agent will move, as well as the path an agent will take. When modelling a city at a low level of abstraction, for example, the routes and speed of public transport may be useful. Another consideration with respect to agents walking is that people fleeing conflict do not always use public roads as these might be more dangerous.

In modelling movement from an area of conflict, the modeller should consider to what extent the model will include an agent's destination. The destination modelled may be the final destination or it might only be a first destination on a journey towards a final destination. Consider, for example, the movement of refugees from Northern Africa to Europe where a neighbouring country might be a first destination when feeling conflict, while a European country might be the final destination. The modeller should consider the level of detail to be included with respect to the agents' destinations.

Another consideration with respect to the movement of agents is the time an agent might take before reacting to conflict. Agents do not always leave immediately when conflict ensues. The agent's reaction time is typically correlated with the maturity of the conflict [119]. The expected duration of the conflict also has an effect on the likelihood that a person will move [4]. Initially, people might wait longer before deciding to move, whereas if the conflict has been ongoing for a few years, people would be more inclined to move when they feel threatened because they may have an indication of the effect of conflict based on what others have experienced.

6.4.5 Data inputs

The modeller should identify adequate sources of available data to assist during model conceptualisation and, possibly, to use as model input. Lopez-Lucia [108] suggests databases such as the United Nations Global Migration database [166], the Organisation for Economic Cooperation and Development International Migration database [127], the Migration Policy Institute database [120], the International Labour Organization Migration database [83] and data annually compiled by the UNHCR [172] on the stock and flows of refugees in over 150 countries. Other sources that may be considered are the Internal Displacement Monitoring Centre [82], the World Bank [184], the Humanitarian Insight database [75] and the United Nations Office for Coordination of Humanitarian Affairs [177].

Despite the numerous sources mentioned, data estimating migration patterns are very uncertain for a number of reasons, such as the lack of current data, conflicting numbers reported by receiving and sending countries, and different data collection procedures and methodologies employed [108]. In cases where data are not available, the modeller is required to make sensible estimates based on various sources which might include qualitative data and insights gained from subject matter experts.

6.4.6 Conflict considerations

In order to model conflict, the modeller should investigate the various elements pertaining to the conflict, such as the type of conflict, the actors involved, the intensity of the conflict, as well as available data sets that may be used as model input. In conflict situations where people are forced to flee, the type of conflict to be modelled is mainly material conflict (*i.e.* physical conflict) as opposed to verbal conflict. In the case where a city is modelled according to a low level of abstraction, for example, the modeller might include different types of conflict in order to model the escalation of verbal conflict to material conflict.

Another aspect to consider is the entities or actors involved in the conflict. When considering a conflict situation in which various actors are involved, the modeller may consider modelling the interactions between the different classes of actors by representing the various actors by different types of agent classes. ABM allows for agents to interact with one another as well as their environment. Conflict can therefore be modelled as either conflict-causing agents interacting with the civilian agents, or as a dynamic environment with which the civilian agents interact.

In order to initiate conflict events during the simulation, the modeller may either make use of a data set as input to the model, detailing the dates and locations of conflict events, or the modeller may allow the user to input the conflict events manually. The *Global Database of Events, Language, and Tone* (GDELT) [59] monitors news media on a global scale in order to capture a variety of events, including conflict-related events. The modeller may use this or other databases as an input to initiate conflict in the model. The intensity of the conflict is a measure that the modeller should also consider, as it may have a direct correlation with an agent deciding when to move away from the conflict. In case an explicit measure of conflict intensity is not available, the modeller may implicitly derive the conflict intensity from another measurement available. In addition to the intensity of conflict, the modeller should take into account the possibility of conflict spreading further to adjacent areas and also the possibility of conflict depleting after some time.

6.4.7 Population considerations

The model should be a sound representation of the population concerned and capture the fluctuation of that population. When modelling over a time scale of years, for example, the model will have to include annual births and deaths (apart from the deaths caused by the conflict). It is proposed that modellers should integrate data sources in order to account for different increasing and decreasing population factors when considering the fluctuation of populations [144]. The factors to take into account will further depend on the movement types considered in the model, as discussed in §6.3.3. Increasing and decreasing population factors with respect to the primary movement types, namely IDPs, refugees and asylum-seekers, and undocumented migrants, are presented in Table 6.1.

The number of refugees and asylum-seekers increases as individuals stochastically arrive in a country and as they apply for asylum in that 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 countries of origin) and resettlement (individuals who decide to relocate to other countries).

The number of IDPs increases when people decide to flee their places of residence and relocate elsewhere within a country. Similarly, if children are born while the mother is classified as an IDP, or if administrative errors are corrected, the IDP data should be updated. Births are an increasing factor, in the same way that deaths are a decreasing factor, to all the categorised types listed in Table 6.1.

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 the country of origin, the individual should not be recorded for a second time as an IDP. In the case where an IDP decides to return to his or her original place of residence, the individual should then no longer be considered an IDP.

Table 6.1: *Increasing and decreasing population factors of IDPs, refugees, as well as asylum-seekers, adapted from Sarzin [144].*

Type	Increasing factors	Decreasing factors
IDPs	• New internal displacements • Births • Administrative corrections	• Cross-border flight • Return • Settlement elsewhere in the country • Local integration • Deaths • Administrative corrections
Refugees	• Spontaneous arrivals (separately identifying individuals who were previously IDPs) • Births • Administrative corrections	• Repatriation • Resettlement • Cessation • Naturalisation • Deaths • Administrative corrections
Asylum-seekers	• New applications for asylum, separately identifying individuals who were previously IDPs • Births	• Positive decisions (convention status, complementary protection status) • Rejections • Otherwise closed cases • Deaths
Undocume migrants	• Spontaneous arrivals (separately identifying individuals who were previously IDPs) • Resettlement arrivals • Births	• Repatriation • Resettlement • Deaths

6.4.8 Movement decision-making influencing factors

The process of human decision making is extremely complex as it 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 [31]. The CoFMMA framework aims to assist the modeller in emulating the decision-making process of people when confronted with conflict.

One suggestion is to simplify the modelling of the decision-making process of people, in particular, by adopting a medium level of model abstraction in order to grasp only the essential elements related to agent movement [103, 160]. There exists a strong appeal towards the simplification of decision analysis and support, especially by considering the benefit of increased transparency and insight [45, 160]. The research performed by Katsikopoulos *et al.* [91] has, furthermore, 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 sufficient 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 the stochastic nature of the problem. Condition B, on the other hand, implies that the decision is apparent and relatively straightforward. It is thus advised to build simple models rather than complex ones if either of these conditions can be considered to hold true [91]. In considering forced displacement, Condition A typically holds, as data are scarce and, in most cases, not of high quality (especially as most data are only estimates) [4, 55]. The modeller should consider the complexity with which the agent's movement decision making will be modelled.

An individual's motives to move to one place rather than another drives the decision-making process and it is therefore necessary to identify these motives. Alhanaee and Csala [6] investigated the number of incoming refugees at a certain location while taking into consideration various attributes of the location, such as population size, provision of education, protection (safety) and distance from the place of origin to the location. The correlation found between these attractors, using a least squared regression model, are illustrated in Figure 6.5. This research and other similar studies may serve as guidance for the modeller to determine the various factors to consider when modelling the movement decision making of agents.

Population	0.70	0.58	0.029	0.51
	Education	0.83	–0.23	0.76
		Protection	–0.22	0.77
			Distance	–0.11
				Refugee increase

FIGURE 6.5: *A correlation matrix of the main attractors, adopted from Alhanaee and Csala [6].*

Two competing hypotheses of destination selection have been developed in the literature for the case of forced migration, namely a localist model and a rationalist model [23]. In the former case, an individual will move to the nearest safe place, whereas in the latter case, the individual will make a decision based on economic indicators. The factors that influence an individual when moving as an IDP *versus* moving as an undocumented migrant or refugee may also differ.

When an individual has to choose which country to move to, whether as refugee, asylum-seeker or undocumented migrant, it is suggested that the modeller implement an *openness index* for each country. An openness index of a country measures the ease with which a person can relocate to said country. Important factors which may influence the openness index of a country include the economy, popularity, property, and education [126]. An economic factor might, for example, be the GDP per capita of the country, or the employment rate. Popularity may refer to the number of people from the same country, region, family or network group, that have moved to that particular country in the past. Migrants tend to relocate to areas where familiarity is apparent, such as when they know someone there or if the culture or language there is similar to their own. Popularity may therefore include language or cultural similarities. Property refers to the assets of a person — those who own property in their countries of origin are less likely to migrate than those who do not. Similarly, if a person owns property in a neighbouring country, they would be more inclined to move to that particular country. Education is representative of the educational opportunities available in the country, both on secondary and tertiary levels.

Similarly to the implementation of openness indices for the different countries, the modeller may implement openness indices in respect of different regions, cities or areas within a country. For example, if forced migration is modelled on a regional level from a particular city, the other cities (destination options) may each be equipped with an openness index.

Apart from the factors of the destination options, an individual's characteristic may further determine when or where to the individual might flee, as discussed in §6.4.3.

6.5 Phase 3: Model development

Phase 3 of the CoFMMA framework concerns the development of the agent-based model. The modeller is required to select an appropriate software environment in which to develop the model, before modelling the four primary elements — the conflict, the people, the movement decision making of the people, and the GUI which will assist users in utilising the model. The model development phase should follow an iterative process whereby the modeller constructs a portion of the model and verifies that portion before moving on to the construction of the next portion. It is recommended that an agent-based model should be developed iteratively by starting to construct a simple model which captures certain essentials and, only after verification, adding complexity in a systematic manner. Each iteration should add depth and realism to the model, thereby further refining it. If, at some point in this iterative process, the modeller finds a portion developed inadequate or unsuitable, he or she should refer back to Phase 2 (conceptual modelling). Adjustments may be made to the conceptual model if the model development phase requires it. Such changes to the conceptual model should be addressed adequately during the model development phase. The model development may require more than one iteration between Phases 2 and 3, allowing for a reasoned, adequate model to be developed.

6.5.1 Selecting an appropriate software

The model, as conceptually developed during Phase 2, has to be translated by the modeller into an operational model within an appropriate software environment. In selecting a suitable software environment, it is important to consider the modeller's capability of using a particular software suite or the modeller's ability to learn to use that software. As explained in §3.1.4, a commercial simulation software suite (*e.g.* AnyLogic, MASON, NetLogo, Repast or Vensim) or a general-purpose programming language (*e.g.* Java, C, C++ or Python) may be used. Using primarily the latter may allow the modeller more control during the execution of the model, whereas using a dedicated simulation software suite may allow for better effectiveness and efficiency of time use. The modeller should also take into account the capabilities of the software environment selected, as well as the cost involved.

6.5.2 Continuous verification and refinement

Verification, as defined by Stewart [159], is the process of ensuring that the model is accurately developed and that it performs as intended. Verification should occur concurrently with the development of the model in that, as each portion or part of the model is added, it should subsequently be 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 environment and that the input parameters and logical structure of the model are sensible [14]. In essence, verification is an intensive debugging process which confirms that the various parts of the model work independently and jointly under the circumstances as programmed.

Simulation software environments typically have tools available for assisting in model verification, such as tracing the simulation model steps and viewing a detailed breakdown thereof. During verification, the modeller should correct errors with respect to the syntax or the software compiler, as well as confirm the logic underlying the simulation model. Another approach is to conduct a "walk-through" in which the modeller manually simulates the process and confirms whether or not the model acts as anticipated. Dedicated simulation software suites allow for animation, which may be a valuable tool in verifying the execution of a model [14].

6.5.3 Modelling conflict within a particular area

The modeller should model the conflict based on the conceptual model design with consideration of the actual nature of the conflict, as described in §6.4.6. In modelling the conflict of a particular area, the modeller should consider the initiation of conflict instances, as well as the spread and depletion of conflict situations. The modeller should further select a methodology for modelling the conflict. The methodology selected will influence the manner in which the initiation, spread and depletion is modelled, and it may be the case that spread and depletion is not explicitly considered.

One methodology involves employing agents to represent the different actors involved in the conflict, such as rebel forces *versus* a government. These agent classes should be equipped with a statechart which differs from that of the civilian agent class and the behaviour of the conflict agent class should mimic the real-world situation. Another option is to model the spatial spread of the conflict as an aspect of the environment by means of mathematical modelling methods such as cellular automata or reaction-diffusion equations. The method selected by the modeller may influence the type of data required for the modelling process. When considering historical conflict situations, the modeller may be able to use available data to replicate exact conflict instances. The intensities of these conflict instances should also be considered and, if not explicitly available, this measurement may be derived implicitly from another aspect measured, such as media coverage.

6.5.4 Modelling a particular population

In order to model the population of civilians, the modeller should consider the nature of agents as well as how the population is represented in the conceptual model, as described in §6.4.3 and §6.4.7. The modeller should develop the agent class or classes involved and their relevant attributes, such as age, gender and ethnicity. Initial values of these attributes should be populated, such as the population's age or gender distribution which will, most likely, be based upon quantitative data available.

The modeller should also incorporate population characteristics applicable to the model's time and geographic scales. One such characteristic is the initial location of agents. In order to mimic reality, the modeller should initially have the agents distributed across the geographic area in a realistic manner. This may be achieved by referring to census data which indicate the distribution of the population across regions, cities or suburbs, for example. Another population characteristic is the natural fluctuation in the population based on births and natural deaths. The modeller may consider documented natural birth and death rates of the relevant country or region to incorporate this fluctuation in the model. An aspect linked to this is the ageing of agents, which might be necessary to incorporate when modelling over a larger time scale.

In addition to developing agent attributes and population characteristics, the modeller should develop statecharts for the agent classes in order to map the behaviour of these agents. The modeller should consider the agent's perspective and perform a logical "walk-through" of possible states that the agent may experience and the flow between these states. Developing a simplified flowchart of the decision-making stages that an agent may go through will assist the modeller in finally developing the statechart. Primary considerations of the agent's behaviour may include conditions under which the agent decides to flee the particular conflict situation, the decision as to which particular movement type the agent will adopt, and the particular destination to which the agent will move. The modeller should further consider the manner in which an agent progresses or flows between states.

The modeller may include additional agent attributes that are derived from the initial attributes, such as an agent's inclination to flee conflict situations. This attribute may represent an agent's tolerance threshold towards conflict situations and may be based on a combination of factors or other attributes, such as age, gender or location. In deriving additional attributes which rely on initial attributes, the modeller should take into account whether an agent represents one individual or an aggregate, in particular with respect to the derivation of attributes, so as to ensure that the attributes are valid and representative of the population.

6.5.5 Modelling the decision making of forcibly displaced people

The decision-making processes of the agents should be modelled based on the factors identified in the conceptual model, as discussed in §6.4.8. The modeller should identify the various decisions to be modelled, as well as the decision-making methods (as discussed in Chapter 4) that will be used to model them. The process according to which an agent selects a movement type to conform to is an example of a movement decision that may be modelled. The modeller may, for example, consider agent attributes (or derived attributes), characteristics of the conflict, or environmental factors that influence which movement type the agent will choose. The modeller will, most likely, have to code the decision-making method chosen within a function that is activated once an agent is required to choose a movement type to adopt. This example will only be applicable in the case where more than one movement type is considered.

Another critical movement decision to be modelled is the agent's choice of destination. The process of selecting a destination may differ based on the movement type of the agent. In the case where there is list of destinations to choose from, the modeller may implement an openness index (as described in §6.4.8) that allocates a value to each of the destinations. This openness index may be dynamic in order to account for changing factors, such as is the case with economic or political indicators. Careful consideration is required from the modeller in determining the factors influencing an agent when choosing a destination. A decision-making method should be implemented in order to code the calculation describing the decision-making process of an agent which ultimately determines how the agent makes its choice.

6.5.6 Developing a graphic user interface

A GUI acts as a platform for the simulation model user by which to engage with the functionality of the simulation model in an intuitive and informative manner. It typically consists of a configuration screen and a primary screen (*i.e.* the screen on which the simulation output is visible during the simulation run).

The configuration screen prompts the user before the initiation of a simulation run to provide or edit certain user input values, if required. The modeller should be able to choose exactly which parameters to set as user input and it is advised that the modeller pre-populate the model with default values for these parameters in case the user is uncertain or incapable of providing values. The modeller should critically consider which parameters to set as user input values. Changes in values of these parameters typically represent different scenarios or strategies that the user will be able to model. It is important to allow for user input in cases where what-if scenarios or strategies are to be considered. The modeller should, however, only allow for key parameters to be set as user input as too many user input options may confuse the user or cause bias, resulting in a misrepresentation of the real-world system.

The modeller may also include user input capable of changing parameter values during a simulation run. These user input parameters should be included on the primary screen and should accommodate in-run changes facilitating cases where the user wishes to deploy different strategies or change the parameter values of a strategy during the simulation run. The primary screen should further allow for a visual representation of agents and/or model variables during a simulation run. Even if the agents do not physically move on the screen, the inclusion of graphs and the display of important data values may be valuable. The graphs and data representation tools utilised should change dynamically during a simulation run. The modeller should ensure that the GUI is a user-friendly environment that serves the needs of the user and the aim of the study.

6.6 Phase 4: Model execution

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 is to be achieved in Phase 4 by the implementation of an appropriate experimental design, model output analysis and validation of the simulation model developed. *Experimental design* or *design of experiments* is a methodology for designing experiments with the aim of investigating certain scenarios or the optimisation of a system [94]. In a similar vein, *model output analysis* seeks to obtain an accurate estimated performance of a simulation model by determining the simulation model's true parameters and characteristics during model execution [159]. *Validation* 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 [163].

6.6.1 Experimental design

Simulation modelling provides for an environment in which users can experiment with the real-world system in order to gain a better understanding of the system or to answer specific what-if questions [7]. Experimental design is a tool that may assist entities involved, such as humanitarian aid organisations or governments, in making decisions pertaining to some underlying real-world situation. These decisions may be aimed at achieving a specific measurable objective, such as providing aid to an increased percentage of IDPs, for which experimental design may assist in finding the best corresponding strategies or input values so as to achieve this outcome.

The user should be able to experiment with the system and perform numerous simulation runs, each with a different combination of input parameter values, and note the effect of these input combinations on the model output. A known challenge in experimental design, however, is the large number of combinations of parameters available [94]. The modeller should be able to identify which parameters to include as input values for the user, as it is usually not feasible to include all parameters for this purpose. The modeller may design a number of experiments for the user based on the selected input parameters in order to evaluate different scenarios.

6.6.2 Model output analysis

The output of a model depends on the input values and initial configuration adopted by the user which, in turn, determines the values of parameters during and after a simulation run. Unfortunately, detailed data pertaining to real-world elements which are typically modelled by the inclusion of specific input parameters are not widely available in the field of forced migration. In light of this, the modeller may include default values which serve the purpose of 'base case' or 'best guess' values and allow for the attribution of parameter values upon simulation start-up to be left to the discretion of the user.

Output analysis of the model developed may include statistical methods which are implemented to calculate confidence intervals for the performance measures of the simulation model output, comparisons between similar systems and variance reduction methods [5]. During the application of an ABM approach, stochastic elements are often included and the sensitivity of these elements in respect of changing the output of a simulation run has to be tested by means of a sensitivity analysis. Various methodologies may be implemented when performing a sensitivity analysis, such as the Bayesian Analysis of Computer Code Outputs. The modeller can alternatively measure the simulation output while varying each behavioural parameter separately over a realistic range of values [53]. The modeller can decide which combinations of behavioural parameter values to test and performing a full factorial experiment involves considering all possible combinations.

Dedicated simulation software environments typically lend themselves to visual output analyses which allow for easy assessment of the consequent model development, which can contribute to model output analysis as a form of validation. The output of simulation runs, resulting from various input scenarios, may also be examined by experts in the field or compared to real output data, if available, as a form of validation [14].

6.6.3 Validation

Shannon [147] summarised validation as a process in which the modeller needs to reach an acceptable level of confidence that the assumptions and reasoning during the development of a model represents the underlying real-world system. Three critical aspects of validation include whether or not the model mimics the real-world system adequately, whether or not the behavioural aspects of the model relates to the real-world behavioural data, and, finally, whether or not the user of the simulation model has confidence in the results returned by the simulation model [147].

The modeller should note, however, that the validation process will never be able to confirm whether or not a model is absolutely correct or absolutely incorrect as the model remains an abstraction of the real world and will never be able to represent the real underlying system perfectly [13]. The validation process requires creativity on the part of the modeller and he or she should always attempt to remain unbiased when performing validation techniques. In order to perform these techniques it is also important that the modeller should fully understand the simulation model as designed during Phases 1 and 2 of the CoFMMA framework, and as developed during Phase 3. It is also important that the modeller should plan and document the validation process for the model user and for future reference.

Balci [13] lists various validation techniques, some of which include data analysis, interface analysis, comparison testing, graphical comparisons, face validation, Turing tests, execution tracing, sensitivity analysis, field testing, statistical techniques, visualisation and predictive validation.

Statistical techniques for the purpose of validation may include analyses of variance, the establishment of confidence intervals, carrying out factor analyses, goodness-of-fit tests, regression analyses, time-series analyses and performing the t-tests [13]. Due to the stochastic nature of data typically encountered in the field of forced migration, quantitative validation techniques may be less useful, in which case the modeller should consider employing face validation.

Face validation of a simulation model involves having subject matter experts review the model developed in order to ascertain how reasonably it 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 [98] 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?

In conclusion, as the model is purely an abstraction of reality, validation is aimed at establishing whether the model adequately represents reality and whether the assumptions made by the modeller when simplifying the real-world system were acceptable.

6.7 Phase 5: Documentation

The final phase of the CoFMMA framework entails the documentation of the simulation model developed by the modeller as he or she progressed through the aforementioned phases. The modeller should be able to present the model and its results clearly, succinctly and in a compelling manner [147]. This phase should include the various theories generated, recommendations to policy-makers and other key roleplayers based on results obtained during model execution (Phase 4).

Carefully planned and well-written documentation will contribute towards the establishment of a more user-friendly and credible tool for the user, as well as a reference for any future amendments or expansions of the existing model. The modeller should consider using appropriate vocabulary with as little technical jargon as possible in order for the user to understand the model and the issues that are considered as important to the end-user should also be taken into account.

The modeller should further document the model as a decision-support tool to the user. The documentation should facilitate the selection and variation of parameters in a simulation model so as to allow for scenario analysis and experimental design by the user [178]. The model developed will most likely possess the capability of being 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 may facilitate a better understanding of the behaviour and decision making exhibited by people when confronted with conflict. The inclusion of this feature will further increase the usability and flexibility of the model.

6.8 Chapter summary

The generic framework proposed in this dissertation was discussed in this chapter. In §6.1, the process followed during the development of the generic framework was described and initial conceptual framework drafts illustrated the author's transition towards and development of the final framework. The CoFMMA framework and its five phases were introduced in §6.2. Phase 1, the formulation phase, was discussed in §6.3, and this was followed by a description of Phase 2, the conceptualisation phase, in §6.4. In §6.5, the third phase of the CoFMMA framework, namely model development, was discussed, and this was followed by a discussion on the fourth phase, namely model execution, in §6.6. Finally, Phase 5, the documentation phase, was considered in §6.7.

Part III

A concept demonstrator in the context of Syria

CHAPTER 7

The model concept demonstrator

Contents

The CoFMMA framework, discussed in Chapter 6, is implemented for developing a concept demonstrator to model forced migration in Syria. The first three phases, namely formulation, conceptualisation and model development, as implemented in developing this concept demonstrator, are discussed in this chapter. The implementation of the third phase continues in Chapter 8, while the implementation of the fourth and fifth phases, concerning model execution and documentation, are discussed in Chapter 9.

7.1 Phase 1: Formulation

As explained in §6.3, the first phase entails the selection of a particular conflict situation, an initial investigation, as well as identifying appropriate agent movement types. This section concerns the first phase of the CoFMMA framework followed in developing a concept demonstrator to model the movement of forcibly displaced Syrians.

7.1.1 Forced migration within Syria

Since 2011, an estimated five million people have fled Syria to seek refuge in Lebanon, Turkey and beyond. In conjunction, even more Syrians have been displaced internally within the borders of the country. By the end of 2015, an estimated total of 11.7 million Syrians (more than half the Syrian population) had been displaced forcibly as a result of the civil war [90, 169]. The instability in certain areas of Syria has forced families and individuals to abandon their homes to find safekeeping elsewhere. Syrian refugees have typically chosen one of four feasible options to find safety. These are internal displacement, encampment (*i.e.* settlement in refugee camps), self-settlement (settlement in urban areas) or the challenging journey of attempting to settle in European countries [17].

The displacement of Syrians on a national level is illustrated in Figure 7.1, indicating the number of IDPs who have settled in each region of the country. 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 these individuals may have led to significant under-reporting. Furthermore, certain areas in Syria are out of reach of humanitarian agencies [80]. The United Nations Office for the Coordination of Humanitarian Affairs [174] estimates 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 who may be displaced more than once or choose to leave the country in order to seek refuge in neighbouring countries or beyond (thus no longer being classified as internally displaced) [80].

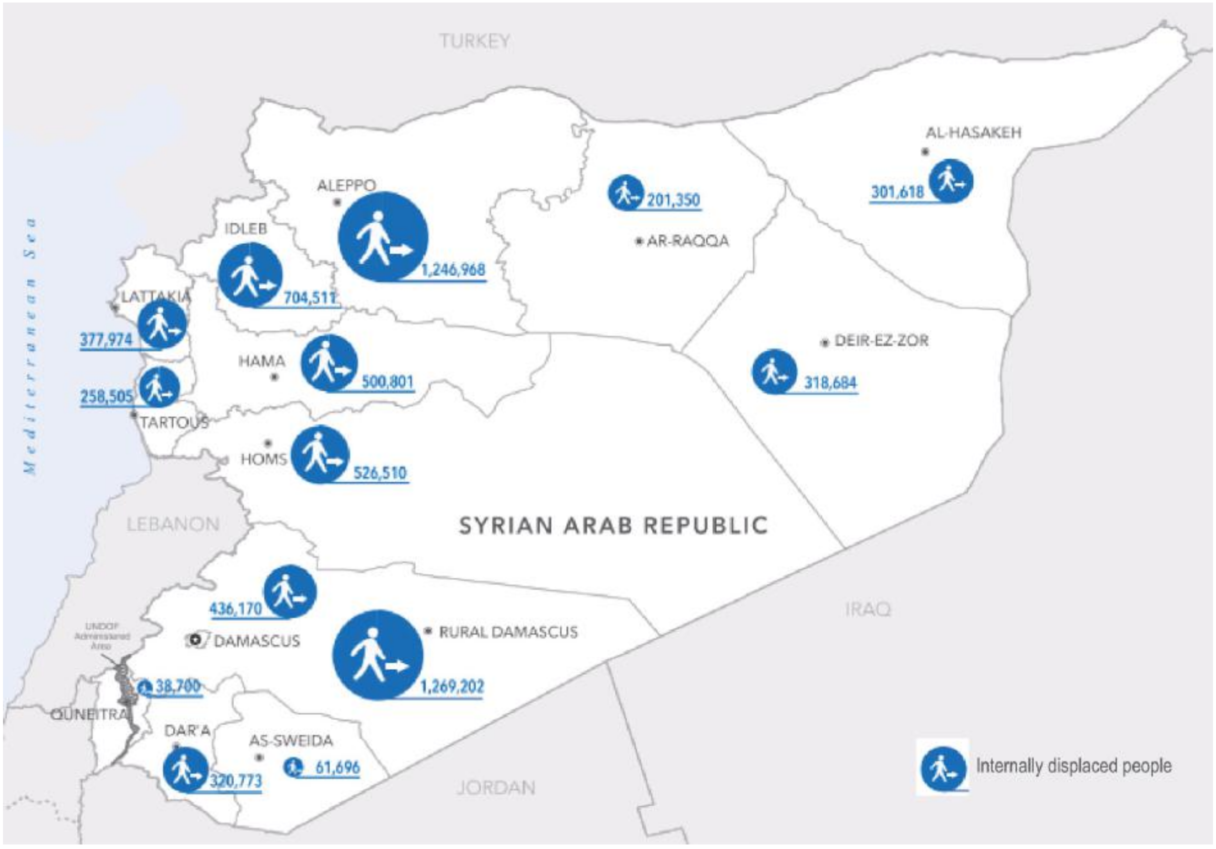


FIGURE 7.1: *A map of the numbers of Syrian IDPs per region at the end of 2015 [80].*

Of Syria's five immediate neighbouring countries, Israel remains inaccessible to refugees. In addition, Iraq, Jordan and Lebanon did not sign the Refugee Convention of 1951 and, whilst providing Syrians with protection, regards them as guests rather than as refugees. Turkey is the only signatory to the 1951 Refugee Convention, although its interpretation of the Convention on the Status of Refugees only applies to Europeans seeking asylum [51]. Turkey and Jordan have set up large refugee camps for those most vulnerable, whilst Lebanon has refused to allow international humanitarian aid to set up such camps [135]. The numbers of Syrians that were registered as refugees in neighbouring countries at the end of 2015 are illustrated graphically in Figure 7.2.

Forcibly displaced Syrians have the option of seeking assistance in refugee camps, but the prospects of living in these camps are dire [17]. Refugee camps are typically densely populated, chaotic settlements owing to overcrowding and a general scarcity of resources [19]. These camps are mostly situated in arid regions where diseases are easily spread due to poor sanitation and housing conditions [17, 72]. Education is typically of poor quality and the economic activity in these 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 [17]. The Refugee Studies Centre [135] corroborates that Syrians who are forcibly displaced prefer self-settlement to encampment.

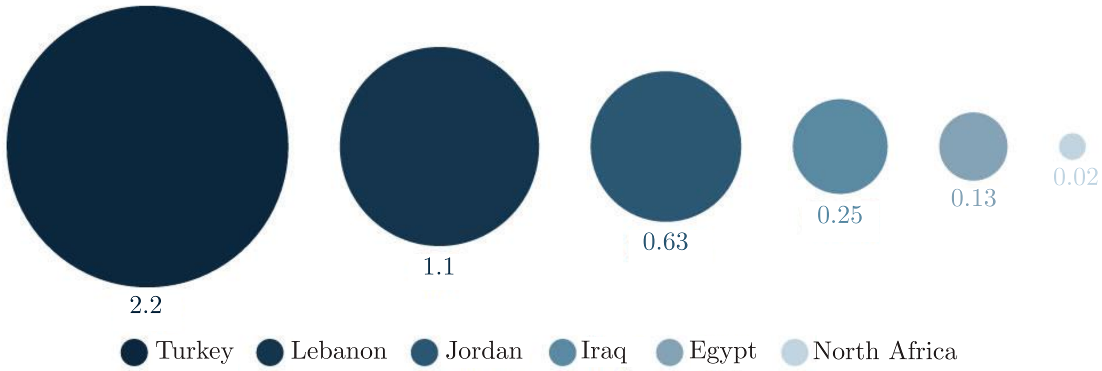


Figure 7.2: *The estimated number of Syrian refugees (in millions) in neighbouring countries at the end of 2015, as derived from data gathered by the UNHCR [169].*

It is further estimated that more than 60% of refugees fleeing across the Syrian border have settled in urban areas where they have family or established social networks. Individuals who chose to settle in other countries without applying for asylum are not included in the United Nations' statistics, as it is assumed that such people are not in need of financial support [51]. Initially, neighbouring countries welcomed large numbers of Syrians, but, as the war intensified, these countries began restricting the influx of refugees from Syria, with some borders being closed altogether [80]. Most official border crossings from Syria into 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 entering into Turkey, Lebanon, Jordan and Egypt by air. Apart from migration restrictions, lawful employment and poor housing in neighbouring countries have led to urban destitution and many have had to consider other alternatives [181].

Syrians hoping to seek asylum in Europe are required to migrate to Europe physically 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 [181]. In the absence of safe, orderly and reliable pathways to European countries, migrants have often been 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 [44, 90]. The aggregated flow of Syrian refugees by January 2016 is depicted in Figure 7.3. This figure does not reflect exact routes followed by refugees [143].

In light of the fact that Syrians can enter Turkey without a visa, many refugees have chosen to travel through that country *en route* to Greece or Bulgaria. From there, they have entered the European Union as ordinary travellers or illegal immigrants, depending on whether or not they had visas. Individuals without visas have the option of applying for asylum [51]. There are also individuals who leave Syria *via* the Mediterranean to Greece, Cyprus, Malta or other European countries situated along the 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 have formally sought asylum in a European country [51].

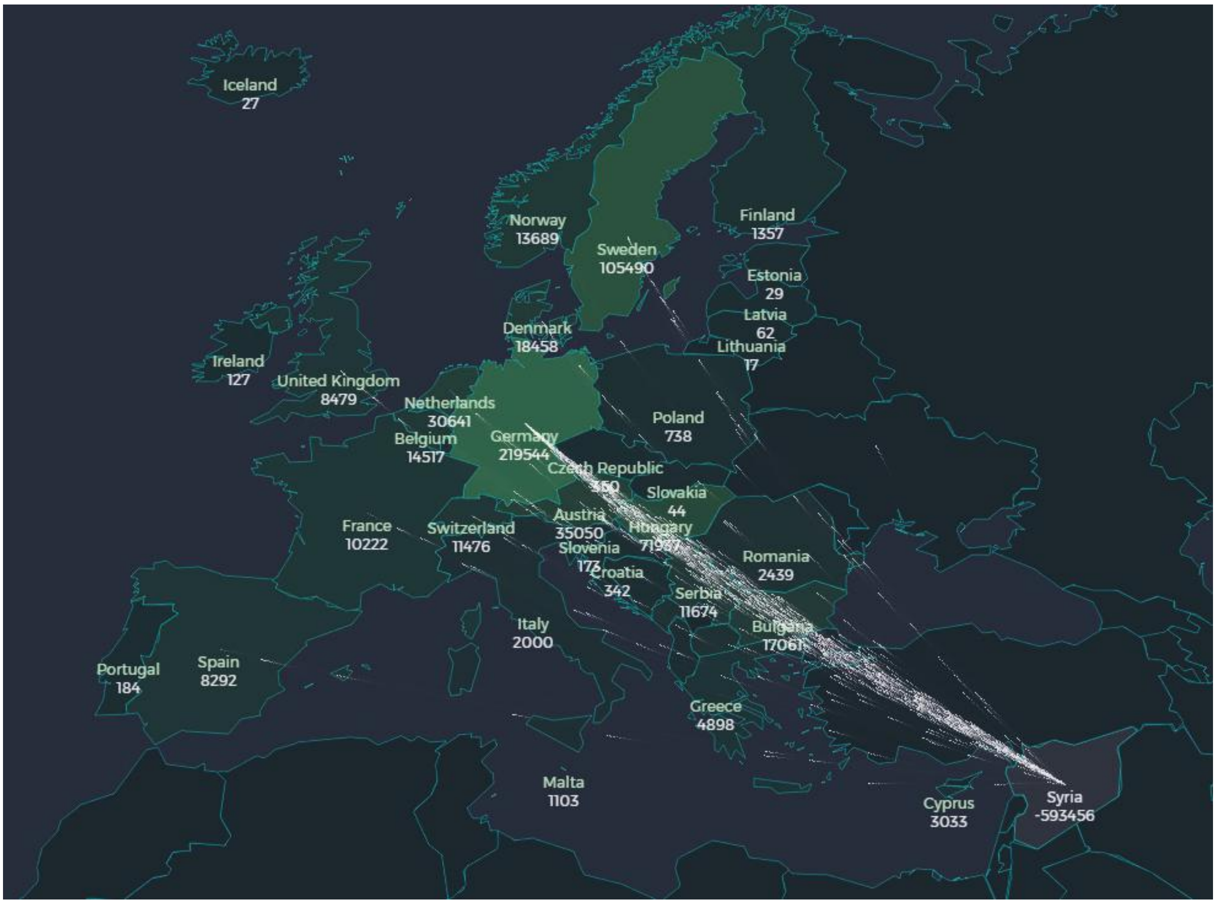


FIGURE 7.3: *The flow of asylum seekers from Syria to European countries as of the beginning of 2016, adapted from Saarinen [143] and the UNHCR [169].*

7.1.2 Geographic and time scales

The aim of the model concept demonstrator proposed in this part of the dissertation is to represent the movement of forcibly displaced Syrians during the Syrian war. On a geographic scale, the model includes the Syrian Republic and its neighbouring countries in order to model the entire Syrian population (approximately 23.7 million people). This calls for a medium level of model abstraction, as a model with millions of agents would be computationally too expensive. A medium level of model abstraction requires the aggregation of agents. The exact scale (*i.e.* the number of people represented by one agent) and manner (*i.e.* the way in which aggregate groups are formed) of aggregation will be determined during Phase 2 of the model concept demonstrator development. In order to model the Syrian war, a time scale of years is considered, as the war has been ongoing since 2011.

7.1.3 Movement types

Since the model abstraction is restricted to a medium level, the exact paths along which the people move does not necessarily have to be modelled. The focus will be on which movement type an agent chooses as well as the anticipated destination of this move.

Agents flee conflict according to three movement types in the model concept demonstrator: (1) IDPs, (2) asylum-seekers/refugees (who are henceforth referred to as refugees) and (3) undocumented migrants. IDPs are Syrians who are forcibly displaced, but choose to stay within the borders of Syria, whilst refugees are those who choose to cross the Syrian border legally, either by applying for asylum in the destination country, or being registered as refugees living in refugee camps. The category of undocumented migrants represents those people who choose to cross the Syrian border without following the required legal and documented procedures.

7.2 Phase 2: Conceptualisation

Phase two of the CoFMMA framework concerns the conceptualisation of the model, as discussed in §6.4. The primary assumptions made and limitations considered in the model are related to the geography, time, data and agent attributes within the simulation model, as well as the conflict, population and movement decisions made by the population. The conceptualisation phase is applied to the model concept demonstrator in this section.

7.2.1 Geographic considerations

The model considers only neighbouring countries of Syria as plausible destinations for immigration by forcibly displaced people as, according to data from the UNHCR [169], more than 85% of Syrians who had fled their country have moved to a neighbouring country. These countries include Turkey, Iraq, Jordan, Lebanon and Greece, as shown in Figure 7.4 (Cyprus is representative of Greece). Other countries are not considered in the model concept demonstrator owing to modelling complexities.

FIGURE 7.4: *A 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, rendering the border between these two countries sealed [51].

The model concept demonstrator will capture 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 is no longer required to keep track of that agent's subsequent 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 concept demonstrator.

7.2.2 Time considerations

The 'Arab Spring'[1] was initiated in Syria at the end of January 2011 [119, 169, 168]. The simulation will therefore commence at the start of 2011, just before the occurrence of 'Arab Springs', and run until the end of 2016. This aims to minimise the effect of selection bias which occurs when more than five years are considered. This is further mitigated by initiating the simulation before the occurrence of the Syrian war and continuing the simulated period over five years.

7.2.3 The agents

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 [103], a subject matter expert, affirms this decision, advocating that agent attributes be restricted to only essential inclusions.

7.2.4 Movement considerations

To replicate the movement of forcibly displaced Syrians, who usually flee by foot, a fixed average movement speed of $4\,km/h$ is employed [143]. Agents do not always leave immediately and the waiting period correlates to the maturity of the conflict [119]. 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.

The geographic area of neighbouring countries utilised in the model, as illustrated in Figure 7.4, 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.

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

7.2.5 Data inputs

The quantitative data employed by the model is available to the public and was gathered from various sources, such as the GDELT [59], the United Nations Department of Economics and Social Affairs Population Division [165], the World Bank [183] and the Central Intelligence Agency [30]. 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.

7.2.6 Conflict considerations

Material conflict (according to the GDELT [59] classification of conflict events) is modelled in Syria over the aforementioned timeline, taking into account the intensity of these conflict occurrences. Other attributes of these conflict events, such as the type of actors involved (*e.g.* police forces, government troops or rebels) or the reason for occurrence are not considered as the modelled outcome is not effected by these factors. Only conflict events initiated within Syria are captured by the model.

7.2.7 Population considerations

Various factors may lead to a fluctuation in a population, as discussed in §6.4.7. Internal displacement, new asylum applications and impetuous arrivals of migrants across the border are considered. The aim is to eliminate reliance on administrative data, thereby effectively removing 'administrative corrections' as an increasing or decreasing factor of population size. Furthermore, the model accounts for births and deaths, as well as IDPs fleeing across the border, but it does not accommodate the other decreasing factors of refugees and undocumented migrants listed in Table 6.1, owing to the associated modelling complexities.

One noteworthy population size influencing factor excluded from the model is the voluntary and impetuous repatriation of Syrians. As noted by Koepke [96] in a study on Afghans in the Islamic Republic of Iran nine years after the conflict in Afghanistan, the decision Afghans had to make in terms of whether or not to return to their country of origin was somewhat startling, as most areas in Afghanistan are rural and offer very basic infrastructure, social services and employment opportunities after the conflict. Younger people preferred to remain in Iran due to the economic and educational opportunities there. In the same vein, 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 stated reasons which limit repatriation in respect of Afghans in Iran, although these reasons 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 healthcare, education and other humanitarian support upon return.

7.2.8 Factors influencing movement decision making

In an attempt to understand the process of modelling the decision making of forcibly displaced people, advice and insights were sought from various subject matter experts, including Aksel [4] from Koç University, Frydenlund [55] from the Old Dominion University, Groen [66] from Brunell University, Lemos [103] from the University of Agder, Shomary [149] from Stockholm University, Smith [154] from the University of Sussex, and Stewart [160] from the University of Cape Town.

Their fields of expertise range from simulation and conflict modelling, to global studies, decision-making theories and sub-fields of the social sciences.

Choosing a movement type

Alhanaee and Csala [6] analysed the motives of forcibly displaced people and concluded that safety and the proposition of a better life appeared to be their strongest motivators. People's attributes in totality, however, affect their decisions when it comes to leaving their countries or relocating within their own borders.

Shomary [149] explained that the people with little wealth, for example, will either stay within Syria or move to refugee camps in neighbouring countries, as they cannot afford to pay smugglers to travel as undocumented migrants. Aksel [4] mentioned, 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 not to experience financial need [51, 181]. Shomary [149] 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 reside in a medium economic class and posses tertiary education, knowing that they will be able to look after themselves once granted asylum [4, 12, 134].

Furthermore, while some younger Syrians wish to relocate and seek for job opportunities, having the necessary funds available to do so, others are reluctant to move. Shomary [149] spoke of her family who, like many others, chose to remain within the country, while only two of the younger family 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 aspiration 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 [149]. Many Syrians who originally moved to other countries as refugees or undocumented migrants have subsequently returned to Syria owing to difficulty adapting. The return of refugees or undocumented migrants is, however, not taken into account within the model concept demonstrator.

Fargues and Fandrich [51] 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 apply for asylum or stay in the refugee camps. The Refugee Studies Centre [135] has estimated that more than 60% of people crossing the Syrian border have settled with families or connections, preferring self-settlement over encampment, as discussed in §7.1.1.

Making a choice between the different movement types is mainly influenced by a person's characteristics and attributes [11]. Initially, a fast-and-frugal tree (described in §4.3) was considered for use in modelling this decision-making process, but predetermining the order in which to present the alternatives based on the person's attributes was found to be infeasible, since none of the attributes are predominant over the rest in determining the movement type of a person. As a result, the additive model (also described in §4.3) 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 may be considered to correlate with the probability of that alternative being selected. Klabunde and Willekens [93] agreed that the most apparent manner in which to evaluate such choices is to enumerate the alternatives and select the option that achieves the highest valuation.

Choosing a destination

The decision of where to move to depends on the movement type of a person. Factors such as safety, distance and population density should be taken into consideration. Shomary [149] agreed that people fleeing will take the population density of possible destinations into account when deciding where to move. In the case of IDPs, an exact location within the country needs to be identified as a new destination for the person, whereas refugees and undocumented migrants simply have to choose which destination country to move to (not necessarily a specific location within that country).

The decision-making method deemed appropriate for modelling the selection of a new destination for an IDP is the conjunctive model. A set of potential solutions has to be identified based on certain criteria, such as safety and population density. The subject matter experts consulted, along with various sources from literature, agreed with these selected criteria [39, 55, 103, 123, 149, 154]. The weighted sum MCDM method may be employed to identify the group of potential solutions by allocating weights to the aforementioned criteria. Then, further utilising the conjunctive model, the purpose of the evaluation function is to minimise the distance between the person's current location and his or her new destination. The set of potential solutions is then explored to find an alternative which minimises this distance. This allows for the selection of an adequate, although not necessarily ideal, alternative. This feasibly replicates the decision making of a person who, due to bounded rationality, also tends not necessarily to pursue optimality but to satisfice instead [42].

Refugees and undocumented migrants are required to choose a destination country. In order to model this decision-making process, the weighted sum MCDM method is again employed. Each alternative (destination country) is scored, based on weighted factors such as the distance from the person's current location to the closest point along the border of that country, the ease with which a person will be able to enter that country (referred to as an openness index) and the number of people from the country of origin already in that country at a given time instant (referred to as popularity). The total score calculated per country is then expected to correlate with the probability of a person selecting that country as destination. Lemos [103] agreed with this modelling approach and also with the use of openness indices and further suggested the allocation of different openness values for the various neighbouring countries for refugees and undocumented migrants. In view of the fact that the openness index for these two groups of people would not typically differ significantly for the purpose of this model concept demonstrator, this suggestion is neglected so as to avoid unnecessarily complicating the model concept demonstrator.

7.3 Phase 3: Model development

The model development phase of the CoFMMA framework involves selecting an appropriate software environment in which to implement the four primary modelling aspects of the model concept demonstrator — the conflict, the population, the movement decision making of the agents, and the GUI. In developing this concept demonstrator, iterative alteration between construction and verification is key, although the intermittent verification process followed is only discussed in the chapter that follows this one.

A map of Syria and its surrounding countries and ocean is employed within the simulation model concept demonstrator, based on a scale of $1\,km : 1\,pixel$, 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 realistic conflict initiation and spread to be replicated while considering the entire Syria. When confronted with conflict, agents have to decide whether to move or not, as well as where to move. The manner in which this decision-making process is captured in the model concept demonstrator is based on decision-making factors described in §7.2.8 so as to incorporate the qualitative data which form part of this decision-making process. Finally, the GUI is developed to assist users in the model application.

7.3.1 The AnyLogic Simulation Software Suite

The design and development of the agent-based model concept demonstrator is conducted in the AnyLogic 8 Personal Learning Edition 8.4.0 [9] software suite. This multi-method simulation modelling tool enables a modeller to gain deeper insights into complex systems and processes across various industries [9]. AnyLogic includes a sophisticated suite of model development tools which, along with the Java modelling language, allows modellers to create complex graphical simulation models [178].

AnyLogic was found to be a suitable choice in facilitating the development of the model concept demonstrator 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 the various concepts of consideration (the nature and spread of conflict, the population density and distribution, as well as the decision making in respect of movement) and also the interaction between them. Furthermore, the GIS animation capabilities supported by AnyLogic will be useful for future work proposed to follow on this study.

Various AnyLogic components were employed in order to construct a model framework and code sequences which appropriately represent the conflict, people within the conflict-affected area and their decision making. These components are elucidated in Table 7.1.

TABLE 7.1: *The different components employed in the simulation environment [21, 178].*

Component	**Application in simulation**	**AnyLogic icon**
Object class	Represents a separate agent which contains an own internal logic structure governing its behaviour and decisions	
Parameter	Describes an agent characteristic of an object class which only changes as the behaviour of the agent changes	
Variable	Stores a model characteristic during a simulation run which changes 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	

7.3.2 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 model concept demonstrator, historical data pertaining to conflict incidents that have occurred in Syria during the specified timeline were acquired using the event data analysis service of the GDELT Project [59].

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

The initiation of conflict

The data gathered on conflict events that occurred in Syria during the period 2011–2016 are used as input to model the conflict. For each occurrence, the processed data file, called ConflictData, includes the date, *Global Positioning System* (GPS) coordinates and total number of information sources citing this event. The total number of information sources citing an event is normalised so as to determine a significance or intensity rating for each conflict event.

The array `GroundState[][]` contains the state of conflict within each cell as a numerical value in the interval $[0, 100]$, where 0 denotes 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 a simulation run.

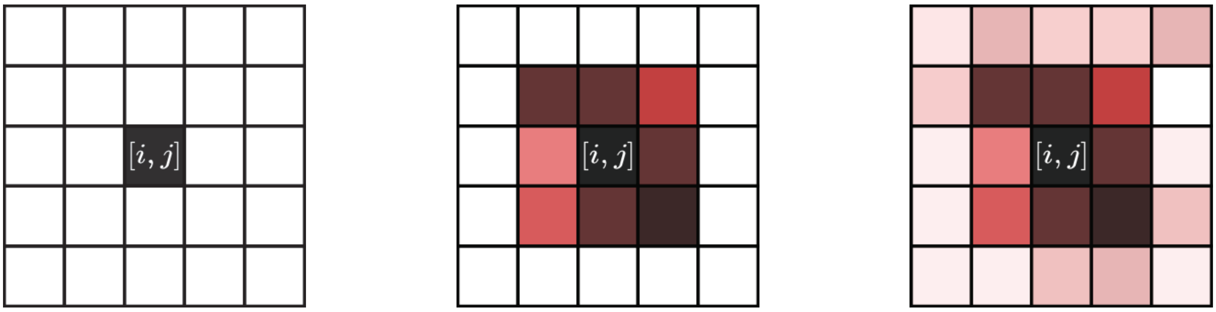
The event `ConflictInitiate` is set to be triggered once per simulation day (*i.e.* a day within the simulation runtime). At every trigger of the `ConflictInitiate` event, the current simulation date is compared with the list of recorded dates in the data file. When the simulation date matches a date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

Before concluding a model representation of the event, the function `UpdateGroundState` is called. This function initiates a search within the `GroundState[][]` array to identify cells in the array which contain a value greater than zero (*i.e.* the cells in which conflict occurs). 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.

The spread of conflict

The modelling of the spread of conflict is based on a combination of the concepts of *reaction-diffusion theory* (*i.e.* the uniform distribution of constituents within a space diffusing from a point of initial action) 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 states of the surrounding cells affect the state of the current cell.

The spread of conflict, as modelled in the model concept demonstrator, is illustrated in Figure 7.5. The intensity of conflict is represented by the darkness (or intensity) of colour: A more intense conflict is represented by a darker colour of the cell. Figure 7.5(a) contains a cell in which conflict originates at position $[i, j]$ (where i represents the row and j the column of `GroundState [i][j]`). The spread of conflict from this cell to its immediate neighbours is depicted in Figure 7.5(b). In Figure 7.5(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 7.5: *A graphical illustration of the spread of conflict.*

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

The cells immediately surrounding the cell in which conflict initiated (depicted in Figure 7.5(b) as $[i, j]$), will then, with probability `ProbabilityofInfection` be propagated. The conflict intensity realised in each of these cells is calculated as a fraction of the conflict intensity 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 7.5(b)) where conflict originated.

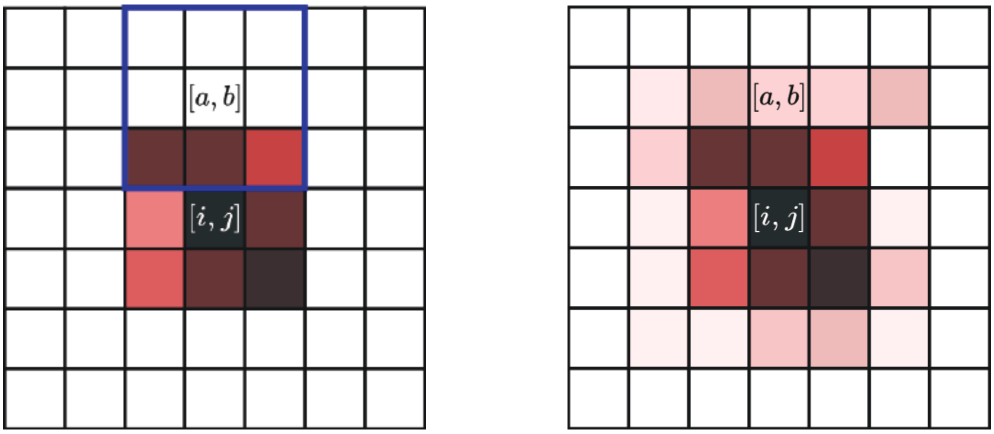
Furthermore, conflict continues to spread to cells beyond the immediate neighbouring cells of the point of origin, as indicated in Figures 7.5(c) and 7.6. The `ConflictSpreading` event begins by searching through cells in the `GroundState[][]` array which fall within Syria, identifying cells with a zero conflict value. Suppose, for example, that a cell thus identified has position $[a, b]$ (as shown in Figure 7.6(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 `GroundState[][]` value of these cells is calculated as

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

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

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

then V GroundState[a][b] (*i.e.* the value of conflict at cell $[a, b]$) will, with probability Prob-abilityofInfection, gain a conflict intensity value. In such a case, the value allocated to V GroundState[a][b] will be a fraction of the maximum conflict value identified amongst the immediate neighbouring values, as illustrated in Figure 7.6(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* $[a, b]$ *has gained a conflict intensity value*

FIGURE 7.6: *An illustration of the spread of conflict to other neighbouring cells.*

Finally, the ConflictSpreading event activates the F UpdateGroundState function which identifies cells with a V GroundState[][] value greater than zero and assigns them a red shading according to the intensity of the conflict relative to those of other areas of existing conflict.

The method employed to model the spread of conflict allows for the spread to occur organically, in a reaction-diffusion-like manner, utilising the rules of a typical cellular automata, as can be seen in Figures 7.5 and 7.6. The process detailed within this section occurs on a daily basis within the simulation runtime.

The depletion of conflict

The depletion of conflict is modelled in a similar fashion, but instead decreases overall conflict rather than propagating it, as illustrated in Figure 7.7. As explained above, the darker the shade of the cell, the higher its conflict intensity value. In Figure 7.7(a), a cell $[c, d]$ with an existing conflict intensity value and a certain average neighbouring conflict intensity value is identified. The depletion of conflict within this identified cell is illustrated in Figure 7.7(b).

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

This event begins by first searching through the V GroundState[][] array to identify cells within Syria which have a V GroundState[][] value greater than zero (*i.e.* conflict exists within these cells). Consider, for example, cell $[c, d]$ in Figure 7.7(a) as having been identified with a V GroundState[][] value greater than zero.

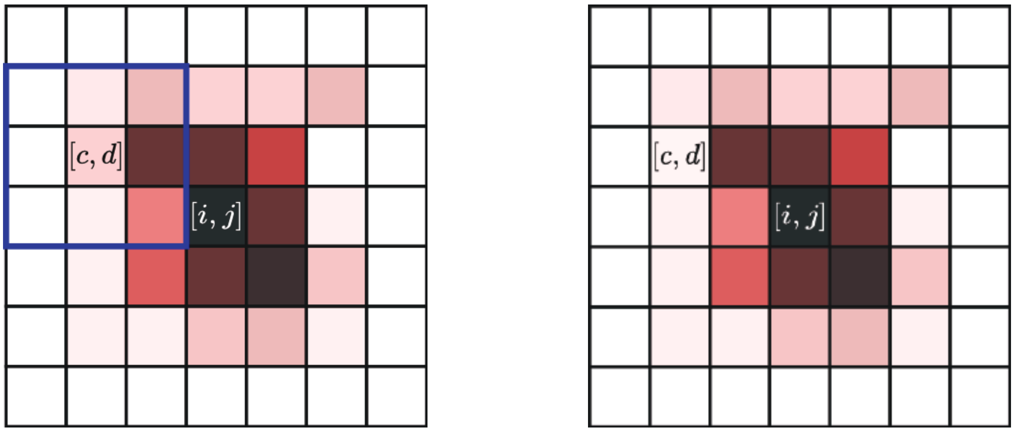
In a manner similar to the way in which conflict spreads, the average of the `GroundState[][]` values of the immediate neighbouring cells is, once again, computed as

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

If this average conflict value is less than the variable value `SetNeighbourAverage`, that is, if

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

and the conflict has been present in cell $[c,d]$ for more than a specified number of days, the `GroundState[c][d]` value is, with a probability `ProbabilityofDepletion`, decreased. The `GroundState[c][d]` value is decreased by a fraction of the maximum conflict value identified amongst the neighbouring cells (as illustrated in Figure 7.7(b)), thereby ensuring that the conflict within that cell dissipates.



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

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

FIGURE 7.7: *An illustration of the conflict depletion.*

The `UpdateGroundState` function is then activated to re-shade the cells with `GroundState[][]` values greater than zero appropriately.

7.3.3 Modelling of people

As mentioned in §7.2.3, 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, but Aksel [4] suggested rather grouping individuals of similar ages and genders as, for example, men exhibit different migrating patterns than women, and different age groups also exhibit different behaviours.

The initial distribution 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 [173] and the natural fluctuation in population owing to births and deaths is also taken into account, as described in the following section.

The different states and subsequent decisions an agent may encounter in the model concept demonstrator are illustrated in Figure 7.8 by means of a flowchart.

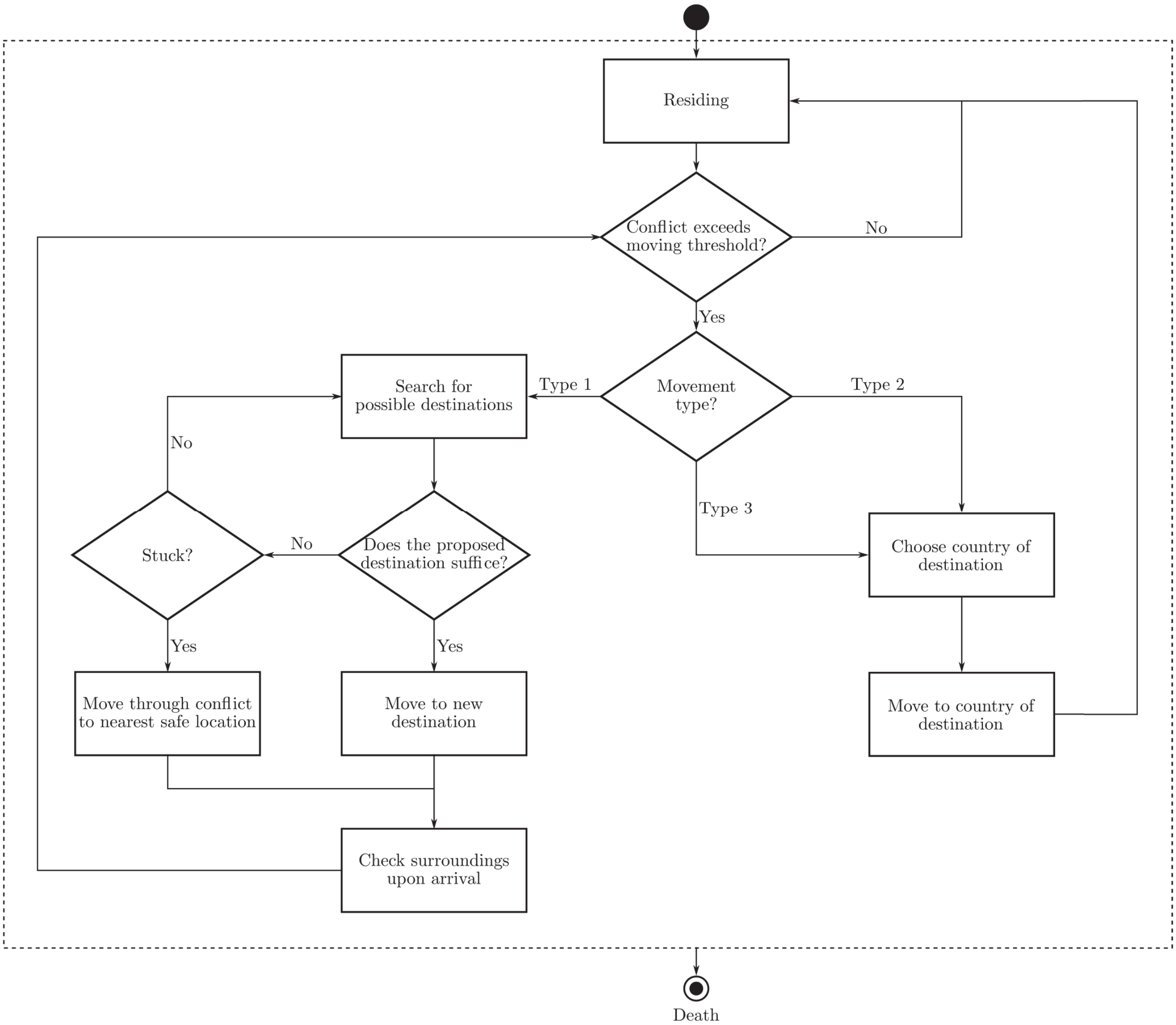


FIGURE 7.8: *A simplified flowchart of the decision making process adopted by an agent.*

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 of tolerating 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 remain within Syria as an IDP (Type 1), a new possible destination in Syria is sought. If the newly discovered destination is sufficiently free of conflict, the agent will move there. If, however, the agent cannot seem to find a destination which suffices, and has been iteratively searching for more than a certain set time period, the agent is considered stuck. In this case, the agent will choose to move through conflict to the nearest location of relative safety, although this may lead to death. If the agent is not stuck, but has failed to find a suitable destination, another search is launched 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 evaluate its immediate surroundings for the presence of conflict in case the conflict has progressed during its journey. Multiple iterations of an agent failing to find a suitable destination may therefore lead to the agent being stuck.

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 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.

The primary decisions to be made by each fleeing agent in the model concept demonstrator are (i) when to move and (ii) where to move to. These decision-making processes and the modelling thereof are described in more detail in the following section. Before discussing the decision making of an agent, however, it is necessary to contextualise agents within the model space as exhibiting certain combinations of attributes in an attempt to replicate reality to some extent.

Agent attributes

Each agent in the simulation model concept demonstrator is characterised by six parameters — gender, age, whether or not it has received tertiary education, economic status, whether or not it 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 initialised according to the probabilities distribution indicated in Table 7.2.

TABLE 7.2: *Agent attributes.*

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

The "Syrian population by sex and governorate according to civil affairs records on 1/2010" [173] 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 male or female with equal probabilities. The model incorporates a boolean parameter called `Male` to indicate an agent's gender. The parameter will, according to the above probability, be initialised to true.

The next agent attribute, stored in the `Age` parameter, specifies the age of the agent. In 2010, the percentage of the 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 (inclusive) and 3.5% were older than 65 years of age [165]. The function `AgeDistribution` ensures a similar initial age distribution of agents as suggested by the data. The ageing of an agent is also taken into account in the model concept demonstrator 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 [30]. Further information suggests that an estimated 28.1% of the Syrian population were using the internet in 2014 [165] and, for modelling purposes, these individuals are considered as those of high economic status. For model simplification purposes, the parameter `EconomicStatus` can assume one of three categories: Low, medium or high economic status. The probability of a person being of low, medium or high economic status is implemented in the model concept demonstrator based on the aforementioned data as 11.9%, 60% and 28.1%, respectively.

The boolean parameter `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 would have received tertiary education [30]. It is therefore modelled that, if a person is older than 18 years of age, there is a 33% probability of him or her having been educated at tertiary level.

Another attribute considered in the model concept demonstrator is the whether or not an agent has family living outside of Syria. According to Fargues [50], an estimated twenty million individuals from Arab countries were not living in their countries of origin by 2011. Syria, along with eight other countries, were identified as 'major senders' and it is further estimated that between 5% and 20% of nationals from these countries were living abroad in 2011. To simplify the incorporation of this aspect in the model concept demonstrator, the boolean agent parameter `InternationalFamily` is initialised randomly according to a uniform distribution 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 [183] has estimated 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 7.3, and the average values are utilised in the model concept demonstrator.

TABLE 7.3: *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

An agent's anticipated age at death is stored in the parameter `AgeAtDeath` and is determined by generating a random number from a truncated normal distribution with mean the average anticipated age at death over the six years and with default standard deviation of 12. The parameter value is clipped to have a lower bound of 0 and an upper bound of 100.

Simulating agent population

Certain variables, such as `BirthRate` and `InitialPopulation`, are utilised to control the initialisation and growth of the population. According to the data from the *United Nations Office for the Coordination of Humanitarian Affairs* (UNOCHA) [173], 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 [165].

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, shown in Figure 7.9, allows for interaction between the environment and the agents.

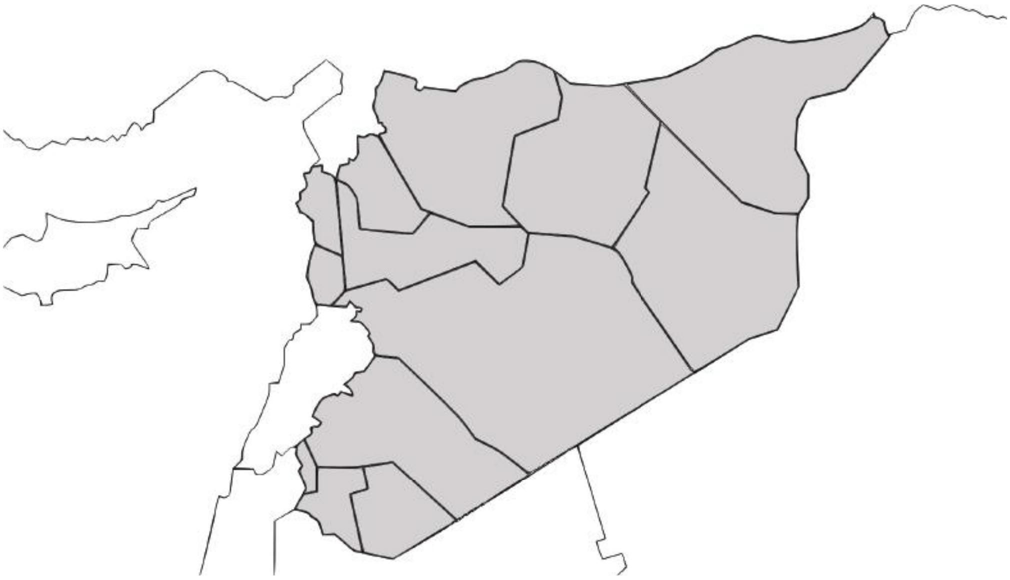


FIGURE 7.9: *Shape files representing the Syrian governorates.*

At the start of a simulation run, a function `Populate` is called to populate the modelled area of Syria with agents. The percentage of agents populated per governorate correlates with the percentage of the total Syrian population per governorate as recorded in 2010 [173]. These percentages are stored as variables within the model concept demonstrator as a set of variables shown in Table 7.4. Within each governorate the agents are, however, placed randomly.

The `UpdatePopulation` event calls a function `Births` 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 are assigned an an age of zero, with the remaining agent attributes being allocated according to Table §7.2. Instead of implementing the death rate in this manner, the age at which each agent dies is determined as a local parameter of each agent, called `AgeAtDeath`. 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.

TABLE 7.4: *Variables containing the percentages of the population per governorate.*

DamascusPop	DeirEzZorPop	DaraaPop
AleppoPop	IdlibPop	TartusPop
HomsPop	AlHasakahPop	QuneitraPop
HamaPop	RaqqaPop	
LatakiaPop	AsSuwaydaPop	

Data analyses conducted by Alhanaee and Csala [6] 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 died fighting in the war. These aspects are not modelled explicitly, as they are encompassed in the data pertaining to the life expectancy previously mentioned.

Another annual event 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 influencing a person's movement choices and, as this simulation takes place over a number of years, it is necessary to account for ageing.

The agent statechart

The behaviour exhibited by an agent in an agent-based model is governed by the statechart of that agent [22]. This statechart portrays the actions taken by the agent and contains the different states in which an agent can reside. The statechart constructed within the `people` agent class environment of the simulation model concept demonstrator is shown in Figure 7.10. The changing of states is either condition-triggered (?), arrival-triggered () or timeout-triggered ().

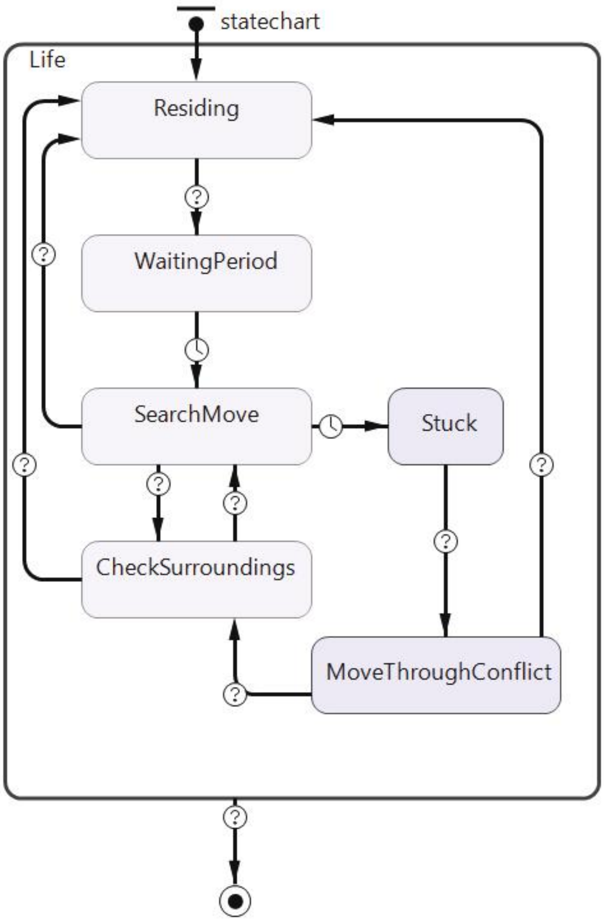


FIGURE 7.10: *The statechart of a person as modelled.*

This statechart correlates with the flowchart depicting the agent states in Figure 7.8. At the point where an agent enters the system, attributes are assigned to the agent according to the description above and the agent is located within Syria according to the geographic distribution of people in Syria, also described above. 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 the `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, however, 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 `Residing` state.

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 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 transitions to the `CheckSurroundings` state and, if its current location is indeed a safe zone, it progresses further to the `Residing` state. At any moment during a simulation run, an agent may die, either because it reached its anticipated age at death, or due to being stuck amid conflict. If this occurs, the agent simply exits the system. The probability value is by default set to 50% as an exact probability cannot be determined and the stochastic nature of the decision-making process should allow for sufficient randomness.

The tolerance threshold of an agent

Agents may choose to leave their current places of residence and migrate to other locations. It is assumed that agents value their physical security and will relocate when they feel 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. In the model concept demonstrator, conflict is modelled as an 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 intensities may react differently. As mentioned, the influencing attributes accounted for in the model concept demonstrator 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 the variable `MovingThreshold` within the `people` agent class. Each agent's `MovingThreshold`, defined as a value between 0 and 1, is empirically calculated by a function called `DetMovingThreshold`. This function is called 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 [175] 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 relocation, 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 [39]. People who have tertiary education are also more inclined to relocate in order to find new jobs in a better economy [149]. It is also assumed that, once members within an agent's social network relocate, that the agent itself is then more likely to relocate, as the cost and risks involved in moving is decreased when more information pertaining to the experience or opportunity at the destination is available directly from someone within a person's social network [39]. Although the simulation model concept demonstrator does not explicitly 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, he or she is likely to be less reluctant towards moving.

By the end of 2015, the UNHCR estimated that more than half of the Syrian population had been forcibly displaced [169]. 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 V `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. In the F `DetMovingThreshold` function, this initial value is then increased or decreased, according to the attributes of the agent, in order to determine the final V `MovingThreshold` value of each agent.

The simulation daily initiates an event called `CheckForConflict` which, in turn, activates a function called F `CheckGroundState`. This function calculates the average V `GroundState` value of the cell in which the agent finds itself as well as in the immediate neighbouring cells. If this average value exceeds the agent's V `MovingThreshold` value, the agent transitions from a *residing* state to a state of *searching* for a new destination. This process depends on the V `MovementType` of the agent, which will is described in the next section.

7.3.4 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 §7.2. Movement type 1 pertains to those individuals who decide to leave their places 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 some country of destination (referred to as refugees). Movement type 3 is relevant to those individuals who cross the Syrian border and relocate to other countries 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 locations 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 [93].

Choosing a movement type

Within the simulation model concept demonstrator, an agent's V `MovementType` is determined as it enters the `WaitingPeriod` state. As discussed in §7.2.8, an additive model is employed to determine the movement type of each agent. Various agent attributes influence this decision, as described above.

A probability matrix is populated 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 $\mathcal{C} = \{c_1, \ldots, c_i, \ldots, c_u\}$ denote the set of criteria constituting to each attribute. Let $\mathcal{A} = \{a_1, a_2, a_3\}$ denote the set of alternatives where a_j represents movement type j. A probability p_{ij} is assigned to the i^{th} criterion for the j^{th} movement type. Let

$$\boldsymbol{E}_T = \begin{array}{c} \\ c_1 \\ \vdots \\ c_i \\ \vdots \\ c_u \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ \vdots & \vdots & \vdots \\ p_{i1} & p_{i2} & p_{i3} \\ \vdots & \vdots & \vdots \\ p_{u1} & p_{u2} & p_{u3} \end{pmatrix} \end{array},$$

where $T \in \{1, \ldots, n\}$ is an attribute label. These probabilities are populated by numerical values derived or inferred from qualitative data pertaining to the attributes of forced migrants within each movement type. For each movement type or alternative, a_j, the probability that an agent will move according to that movement type is given by

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

For each attribute probability matrix, $\boldsymbol{E}_T$, the criterion i corresponds to the specific criterion of the attribute associated with the specific agent. According to the probability associated with each movement type $p(a_j(i))$, the agent is 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`.

Choosing a destination

The actions of an agent correlate directly with the movement type attributed to it. Distance and population density (or popularity) are important criteria in the respect, as individuals tend to move to the nearest location when fleeing conflict, typically choosing a destination where others have gone before [39, 123]. When an agent has been assigned 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 index of each neighbouring country, the distance from its current location to the neighbouring country borders, and the number of Syrians who have already relocated to the countries under consideration.

Modelling movement type 1

Harrison [73] claimed that IDPs generally move within the same area in which the conflict or disaster occurred (trying to minimise distance), although Alhanaee and Csala [6] have stated that these individuals prioritise safety. Both authors agree that the destinations to which IDPs move correlate strongly with areas of high population density [6, 73]. This corresponds with a finding by Zipf [188] which states that the inter-community movement of people between two communities, P_1 and P_2, separated by a distance D, is 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 model concept demonstrator, arrays are superimposed over the physical modelling space, each with a granularity of 100×60 cells, in order to facilitate the modelling of conflict, the population density, and the attractiveness of destinations within Syria, based on these two weighted criteria. Each of these cells represent a $100\,km^2$ area.

The array, `GS`, translates the array `GroundState` (which indicates the conflict intensity within Syria over a 200×120 cellular space, as explained in 7.3.2) into a 100×60 array. In the simulation model concept demonstrator, an event called `UpdateAttrZones` calls the function `PopulateGS` on a monthly cycle which is responsible for this translation. Another function, `UpdatePopDensity`, is simultaneously executed to count the number of agents in each of the simulated cells covering Syria and then to populate the `PopDensity` array with these values. Furthermore, these population density values are normalised to values between 0 and 100, and stored within another array called `PopDensityNorm`, for further calculation purposes. For visualisation purposes, the cells are shaded 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 a `CombinePopGS` array so as to consider all attractive zones within Syria. Cell $[i, j]$ is allocated a weighted sum of the 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

$$w_1(100 - GS[i][j]) + w_2(\text{PopDensityNorm}[i][j]),$$

where w_1 and w_2 denote the weights allocated to the conflict and population density, respectively. Both of these weights are user inputs with default values of 0.6 and 0.4, respectively. Furthermore, these attractiveness values are normalised and stored within the array `CombinePopGSNorm` as values between 0 and 100. These normalised values represent the attractiveness of cells and are illustrated visually by applying a purple hue to the cells. The darker the hue of a cell, the more attractive it is as potential destination for IDPs.

The `MovementType1` function in 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, attempting to minimise the distance between its current location and the various destination alternatives before finally selecting a location as its destination.

Modelling movement types 2 and 3

When considering cross-border movement, the model concept demonstrator only takes into account Syrians moving to neighbouring countries, either as refugees or undocumented migrants. The modelled decision-making process, as discussed earlier, 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 approach is similar for both movement types.

As in the case of the decision-making process of IDPs, distance and population are taken into account when choosing a destination. Refugees and undocumented migrants consider the distance between their current locations and the borders of the various neighbouring countries, as well as the popularities of these countries. As mentioned, 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 index of each neighbouring country (*i.e.* the ease with which a person would be able to enter the country). The objective when making a movement decision is to maximise the openness index and popularity, whilst minimising the distance.

When an agent is classified as either of movement type 2 or 3, the F `MovementType23` function is called. User input variables, each representing the weighting factor of a certain criterion, are considered within this function. These variables are the weight factors corresponding to the distance V `MT23_factDist`, the openness index V `MT23_factOS`, and the popularity V `MT23_factPop` of the countries. The default values for these weights are 0.55, 0.35 and 0.1, respectively.

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 representing the different Syrian governorates, as explained in §7.3.3, shape files representing the different neighbouring countries are also constructed, as illustrated in Figure 7.11. The inverse values of the various distances to each country are normalised over the maximum inverse distance, in order to transform all distances to values between 0 and 100.

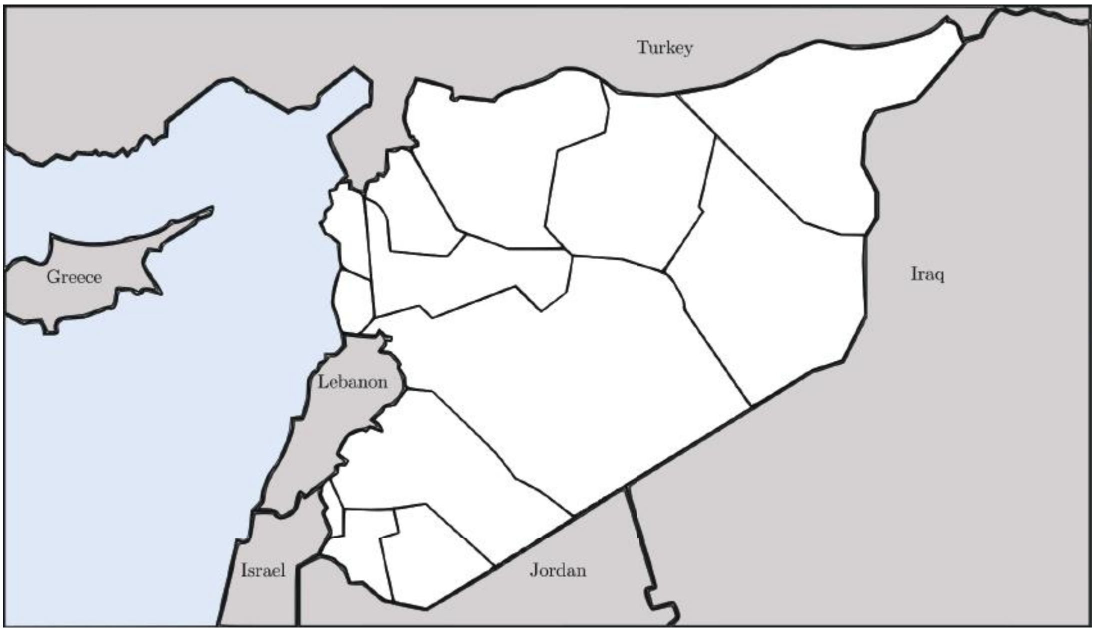


FIGURE 7.11: *Shape files representing the countries neighbouring Syria.*

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

These data are stored in the `Main` object class within a series of data sets which are accessed *via* a series of variables, listed in Table 7.5. The values representing the population sizes of Syrians in the respective neighbouring countries are normalised in respect of the maximum population in order to transform the popularities of these countries to values between 0 and 100.

The openness indices of the neighbouring countries are stored in parameters, as listed in Table 7.6. Each country has an initial openness index which can be specified by the user. Within

TABLE 7.5: *The series of data sets and corresponding variables which stores the population data per country.*

DSyrianPop	dataSyrianPop
DTurkeyPop	dataTurkeyPop
DGreecePop	dataGreecePop
DLebanonPop	dataLebanonPop
DIraqPop	dataIraqPop
DJordanPop	dataJordanPop

the `Main` object class, an event called `AdjustOpenScores` calls a function `OpennessScores` on a monthly basis to update the values of the openness indices.

TABLE 7.6: *Openness indices stored as parameters within the `Main` object class.*

TurkeyOS	LebanonOS	IraqOS
GreeceOS	JordanOS	

The Refugee Studies Centre [135] regards Turkey as being far more humane and pragmatic than the other countries neighbouring Syria in respect of its approach towards handling the mass influx of Syrians. It is therefore assumed that Turkey initially possesses the highest openness index. Large refugee camps have been set up in Turkey and Jordan to accommodate those most vulnerable and, although Lebanon refused to allow international humanitarian aid to set up such camps, Syrians were still granted access into Lebanon [135]. In light of this, Turkey and Jordan have initial default openness indices that are larger than that of Lebanon.

As the war intensified, many neighbouring countries began restricting the influx of Syrians and closing borders, as mentioned before [80]. With this occurrence in 2013, Turkey became increasingly popular as it was still open to Syrians, as was Iraq which experienced an influx owing to a number of camps that had been set up for Syrian refugees [119, 181]. Towards the end of 2014, Jordan's attitude towards the Syrians started changing, as the government realised that its approach was unsustainable [1].

In January 2015, the Lebanese government introduced new regulations which required Syrians to apply for visas 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 into Jordan and Turkey were strictly controlled. In January 2016, Turkey began requiring Syrians to obtain visas before entering the country and Iraq started closing most of its borders which resulted in a decreased openness for both these countries [181]. Then, when Germany opened its borders to refugees in June 2015, this resulted in an increased influx of Syrians into Greece, as they travelled *via* this country towards other European countries [12].

The `OpennessIndices` function compares the time within a simulation run to the timeline of openness adjustments described above and adjusts the openness index of each neighbouring country accordingly. The fluctuation is managed by employing a certain percentage increase or decrease, rather than simply adding or subtracting values. Figure 7.12 contains a timeline showing when changes to neighbouring countries' openness indices occurred. The notation ↑ and ↓ indicate an increase and a decrease in openness index, respectively.

July 2013	August 2013	September 2014	January 2015	March 2015	June 2015	January 2016
Turkey ↑	Iraq ↑	Jordan ↓	Lebanon ↓	Turkey ↓	Greece ↑	Turkey ↓

FIGURE 7.12: *A timeline illustrating the fluctuation of the openness indices of neighbouring countries.*

When an agent adopting movement types 2 or 3 selects a country as proposed destination, the distance, popularity and openness index are taken into account within the `MovementType23` function. Let $\mathcal{N} = \{n_1, \ldots, n_i, \ldots, n_I\}$ denote the set of alternatives or neighbouring countries. $\mathrm{Dist_i}$ denotes the normalised inverse value of the distance between an agent and the border of country i, whilst $\mathrm{Popularity}_i$ represents the popularity of country i and $\mathrm{OpennessIndex}_i$ denotes the openness index of country i. The overall score of country i is calculated as

$$S_{n_i} = w_1(\mathrm{Dist}_i) + w_2(\mathrm{Popularity}_i) + w_3(\mathrm{OpennessIndex}_i),$$

where w_1, w_2, and w_3 are the weight factors associated with the criteria. This overall score is stored as a variable within the `people` agent's modelling environment, shown in Figure 7.7. The probability of an agent selecting a specific country as proposed destination directly correlates with the overall score of that country. According to this probability distribution, an agent then selects a country which results in the agent moving to a random point within that country.

TABLE 7.7: *The overall scores calculated per neighbouring country.*

TurkeyScore	LebanonScore	IraqScore
GreeceScore	JordanScore	

7.3.5 The graphical user interface

The GUI employed in the model concept demonstrator consists of a configuration screen and a primary screen which facilitate interactions with the user. The GUI of the simulation model concept demonstrator is described in this section, along with the underlying modelling infrastructure which regulates and facilitates access to the various model components.

The configuration screen

The configuration screen, shown in Figure 7.13, is constructed within the `Simulation:Main` tab of the model concept demonstrator. Upon initiation of the simulation model concept demonstrator, the configuration screen will appear. It is designed to prompt the user for certain user inputs required before the model is executed. If the user chooses not to enter any inputs, the model concept demonstrator can still be executed by employing default values, as described above. The user inputs have been labelled in Figure 7.13, and are discussed below.

(A) **Data input**

The simulation model allows for data-based or manual initialisation of instances of conflict. A user can choose whether to initialise conflict manually by clicking on specific locations during a simulation run, or to initialise instances of conflict based on historical data, as described above. These data should 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

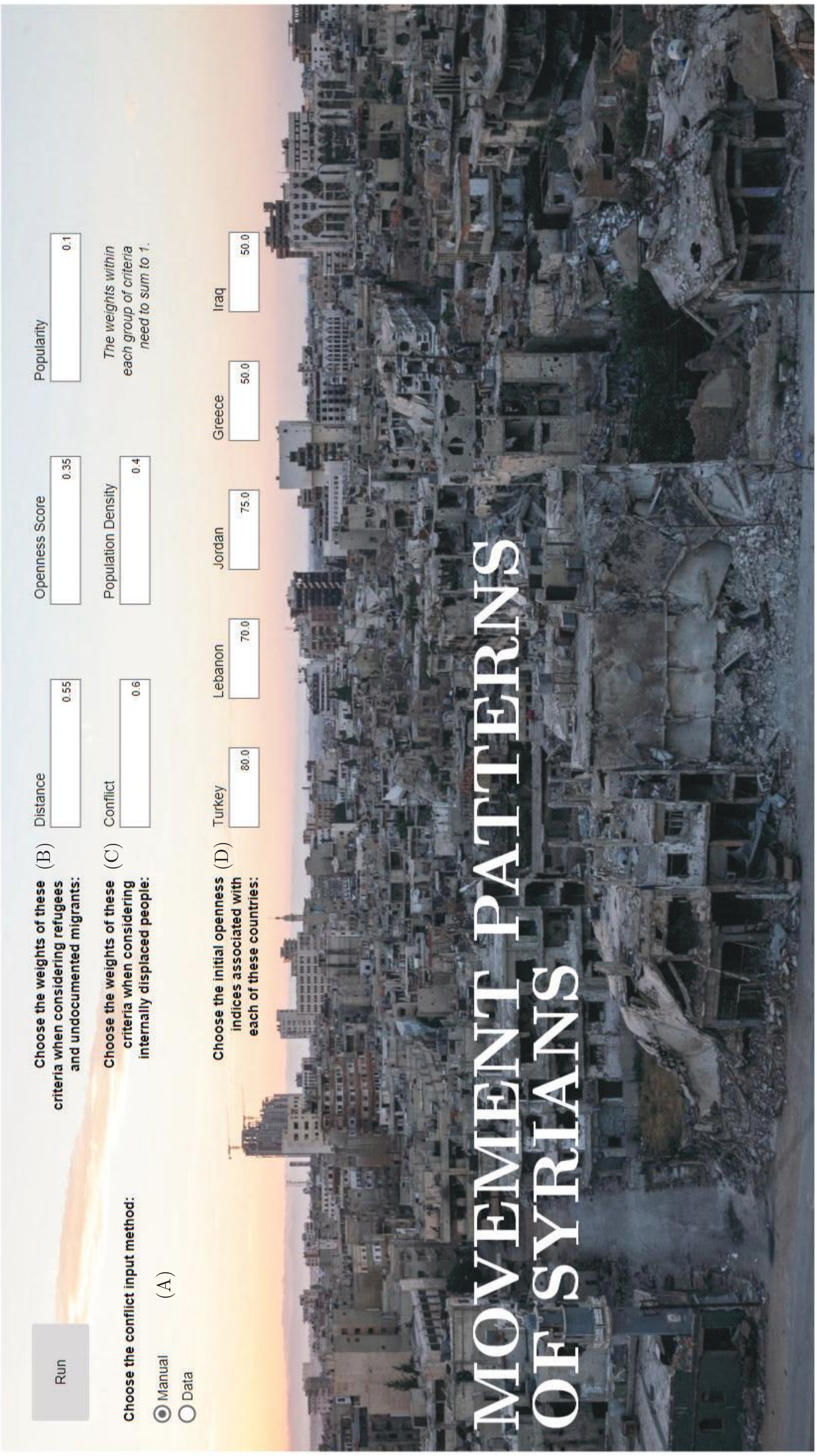


FIGURE 7.13: *A screenshot of the simulation model concept demonstrator's configuration screen.*

determines whether to activate the manual conflict initialisation environment or to initiate the `ConflictInitiate` event, which reads the relevant data from an EXCEL file linked to the EXCEL file object, `ConflictData`.

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

When classified as either a refugee or an undocumented migrant, an agent is required to decide on a country as a proposed destination. This decision-making process, as explained in §7.3.4, involves three criteria (distance, openness index and popularity), which are weighted in order to model the decision. These variables are employed whenever an agent has to choose 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 index 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 a simulation run in order to determine the attractive destination zones.

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

When an agent classified as a refugee or an undocumented migrant has to select a neighbouring country as its new destination, one of the criteria taken into account is the openness index values of the various countries. These are scores out of 100, where 0 denotes extreme reluctance of a country to accept migrants, while 100 indicates that the country is open and welcoming towards migrants. The openness indices fluctuate as functions of time during the course of a simulation run, as described above. The initial openness index of each neighbouring country at the start of the war may be controller as user input.

When the user has completed the model concept demonstrator configuration, its execution can be initiated by the user by clicking on the "Run" button.

The primary screen

The primary screen of the simulation model concept demonstrator, shown in Figure 7.14, has been developed in the `Main` object class and this screen is visible to the user during the execution of the model concept demonstrator. The user-specified values on the primary screen may be altered during a simulation run, while the user-input settings on the configuration screen may only be set before execution of a simulation run. The user inputs included in the primary screen, as labelled in Figure 7.14, 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 is superimposed as a display over the primary modelling space. Three layers are constructed

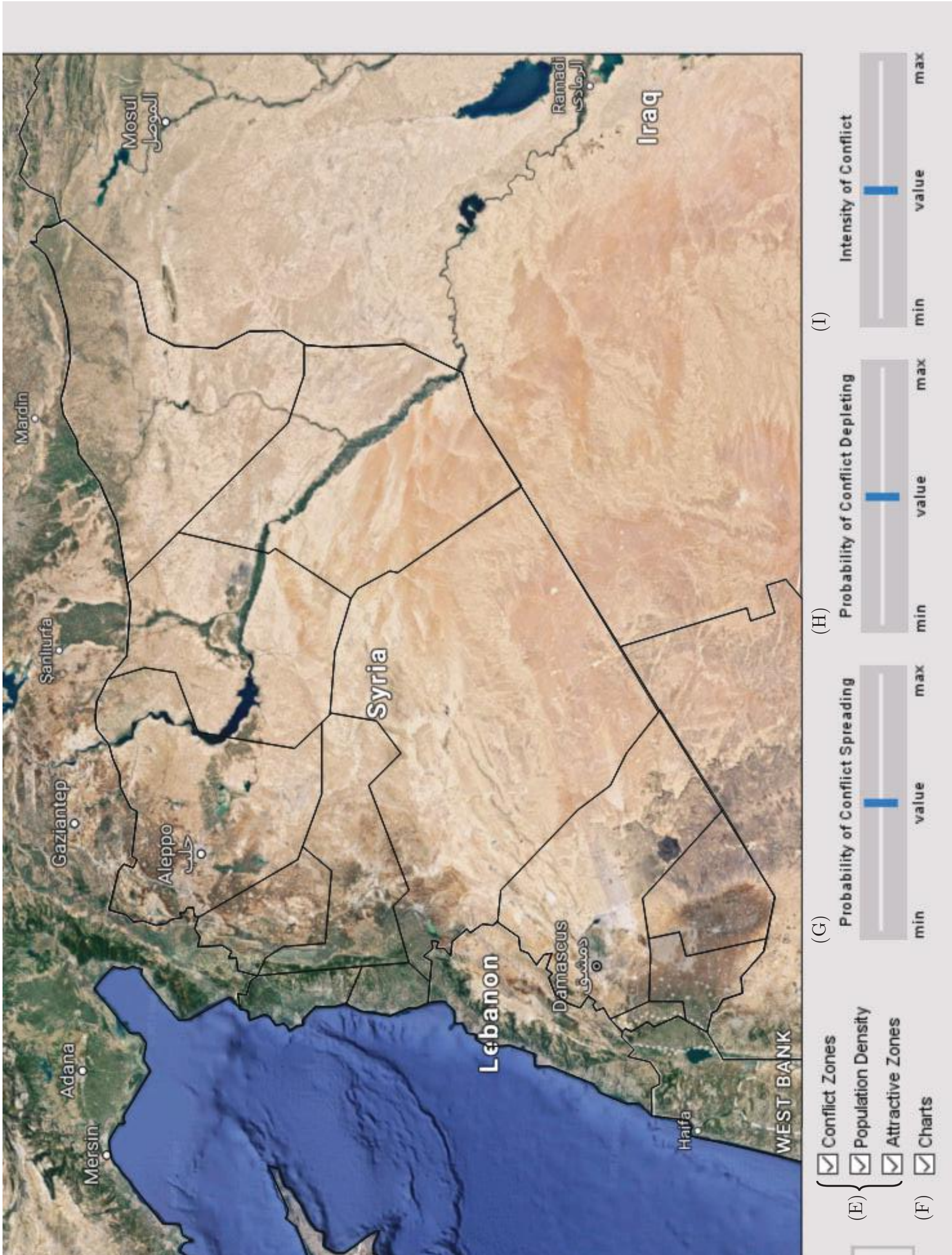


FIGURE 7.14: *A screenshot of the simulation model concept demonstrator's primary screen.*

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 darker hue. The "Population Density" layer illustrates the different population densities across Syria by means of a blue hue. Similarly, cells with a darker blue hue are indicative of higher population densities within those cells. Finally, the "Attractive Zones" layer gives an indication of the attractiveness of an area with respect to locations proposed for IDPs and, as previously explained, the attractiveness is visually illustrated by means of 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 emanating from a simulation run, plotted against the simulation timeline. These graphs, shown in Figure 7.15, include the number of Syrians who have moved to each of the neighbouring countries, the fluctuation of the Syrian population living in Syria over time, the total number of Syrian IDPs as a function of time, the total number of Syrians who have fled as refugees, the total number of Syrians who have fled as undocumented migrants, the number of Syrian refugees per neighbouring country as a function of time, and the number of Syrian undocumented migrants per neighbouring country as a function of time. The data pertaining to these graphs are stored in data sets by an event, `UpdatePopulationData`, which retrieves the data per simulated year from a simulation run by means of the `PopData` function.

(G) **Conflict spread adjustment bar**

The probability of conflict spread is another input the user may alter before or during simulation model concept demonstrator execution. This value links to the `ProbabilityofInfection` parameter. The `ConflictSpreading` event employs this variable in modelling the spread of conflict.

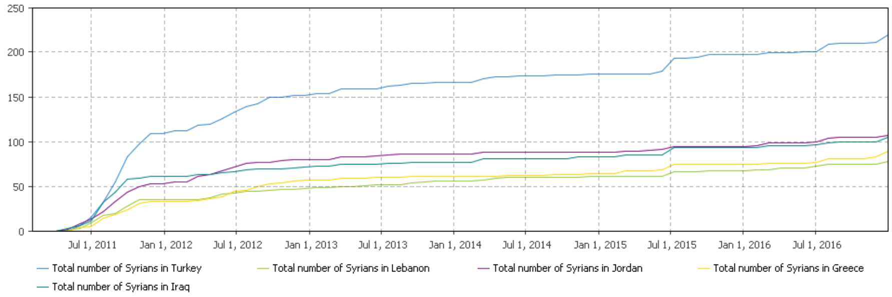
(H) **Conflict depletion adjustment bar**

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 utilising this input variable.

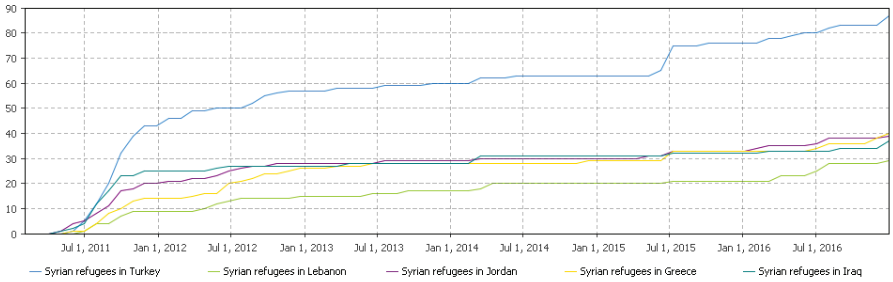
(I) **Conflict intensity adjustment bar**

The conflict intensity adjustment bar is only available for user manipulation when conflict data are initialised manually. Conflict is manually initialised by the user who 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 `ConflictIntensity`.

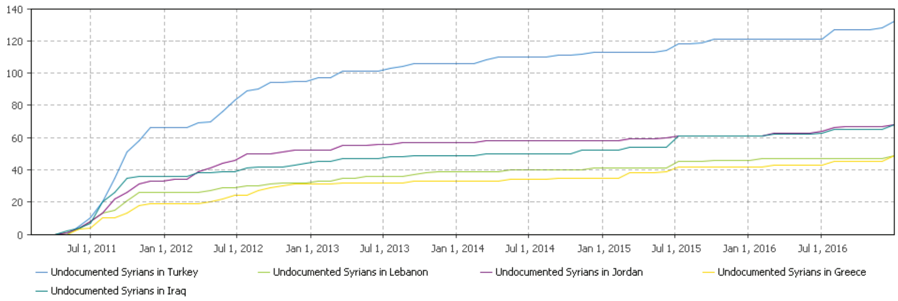
The primary screen appears with all the checkboxes unchecked and default values (as determined through model testing) set to the adjustment bars. The user can therefore choose whether or not to exert any further influence, before clicking on the "Start" button for the simulation time to start. The user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



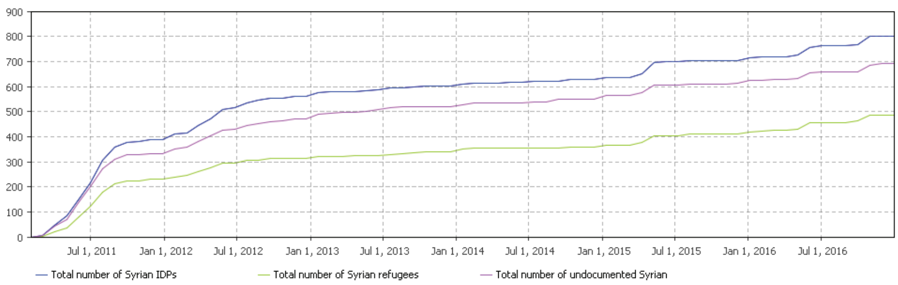
(a) *Total number of Syrians per neighbouring country*



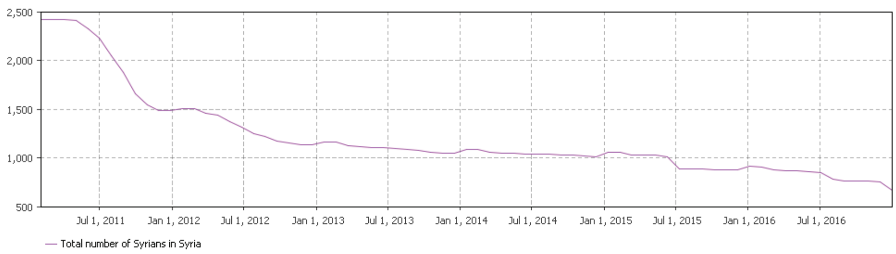
(b) *Number of Syrian refugees per neighbouring country*



(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 7.15: *A screenshot of the graphs displayed during each simulation run.*

7.4 Chapter summary

In this chapter, the first three phases of the CoFMMA framework were implemented in the development of a model concept demonstrator aimed at modelling the conflict-induced forced migration in Syria. In Phase 1, the objective of the concept demonstrator was clearly formulated with a specific focus on the geographic and time scales that apply, as well as the movement types modelled. The conceptualisation of the model concept demonstrator, as performed in Phase 2, was documented in §7.2 and entailed various considerations and assumptions made with respect to the geography, time, data, and agents, as well as various aspects modelled, such as the population, conflict and the decision making of agents.

The model development of this concept demonstrator was discussed in §7.3 as part of carrying out Phase 3 of the CoFMMA framework. The discussion included a background to the AnyLogic simulation environment in which the concept demonstrator was developed, the four components modelled, namely conflict, people, an agent's decision-making process, and the GUI implementation.

The modelling of conflict encompassed the initialisation of conflict based on existing data or manual input, and the spread and depletion of conflict based on a combination of the theories of reaction-diffusion and cellular automata. The modelling of the people involved the attributes allocated to agents, the births and deaths which governed the fluctuation of the population size, the different states in which an agent may reside and the tolerance threshold of each agent before triggering movement. The modelled decision-making processes of the agents included the allocation of an agent's movement type along with the choice of a suitable destination. The GUI, allowing user-specified input, was finally also discussed along with the various elements which form part thereof.

CHAPTER 8

Verification of the model concept demonstrator

Contents

The third phase of the CoFMMA framework entails model development and continual verification of the model developed. The four fundamental building blocks of the simulation model concept demonstrator, as described in §7.3, are verified individually before considering the model concept demonstrator as a whole, when all of these building blocks are implemented together. These elements include the modelling of conflict, the population, the movement decision-making process and the GUI. A number of cases are tested within each of the modelling divisions for verification purposes.

The manner in which the conflict modelling in the simulation model concept demonstrator is verified is explained by means of a number of tests in §8.1. 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 §8.2 through a series of experiments in order to verify that the model concept demonstrator performs as intended. Finally, verification is performed in respect of the decision making of the agents in the model concept demonstrator in §8.3.

8.1 Phase 3: Verification of the modelled conflict

The process followed in modelling the initiation, spread and depletion of conflict in the model concept demonstrator is discussed in §7.3.2. In order to verify the implementation of this modelling approach, the following cases were tested during simulation experiments:

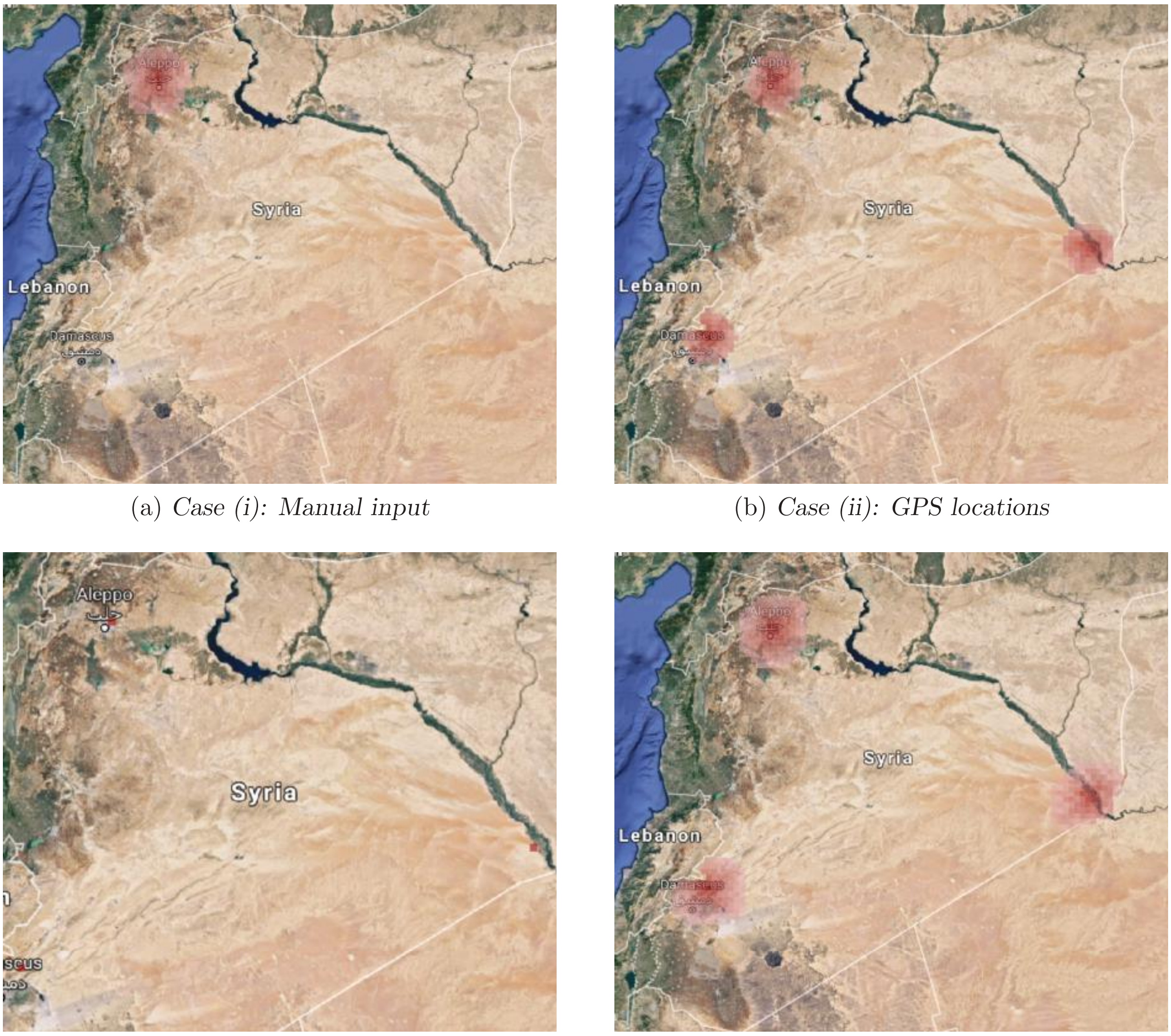
Case (i): Manual conflict initiation,

Case (ii): A small set of input data points aimed at testing the accuracy of GPS positioning, Case (iii): The effect of varying probability of conflict spread,

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

Case (v): The effect of varying conflict intensity.

These cases were implemented in the model concept demonstrator and subsequently reflected upon. The results achieved for the different cases implemented are illustrated in Figure 8.1. For each of these cases, the *Conflict Zones* layer checkbox within the primary screen was selected to indicate the presence of conflict.



(a) *Case (i): Manual input* (b) *Case (ii): GPS locations*

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

FIGURE 8.1: *Cases (i)–(iii) of verifying the modelling 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 initialised in the city of Aleppo by clicking on the specific location on the map. The outbreak of conflict was indicated by the red area shown in Figure 8.1(a), which is in accordance with the manner in which the conflict initialisation is modelled.

A set of four data points, chosen by the model concept demonstrator developer, were used as conflict input data to ensure that the latitude and longitude given as input are correctly superimposed on the Syrian map and that the associated conflict intensity is correctly read in. The data set to be used as input for Case (ii) is given in Table 8.1. 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 in the data set are located in Syria and, as a result, the model concept demonstrator should only show the conflict associated with those GPS coordinates, as conflict that occurs outside of Syria is not considered in the simulation model concept demonstrator. The conflict intensity was also fixed at 100 for the purpose consistency.

TABLE 8.1: *The data set 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

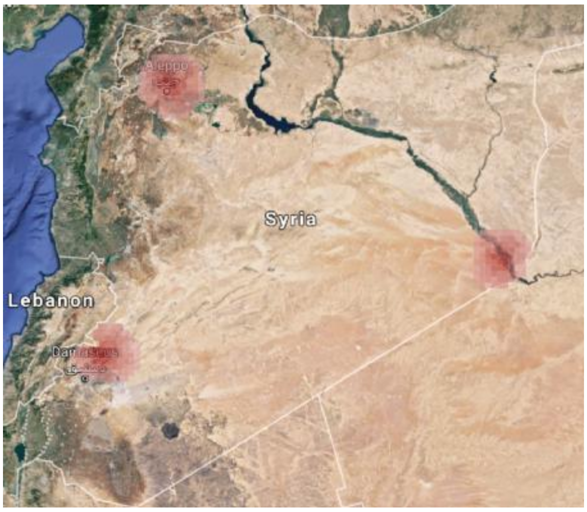
The *Data* conflict input method was chosen from the configuration page and the *Conflict Zones* layer checkbox was again selected to show the presence of conflict. The result from the simulation run at the end of January 2011, depicting the conflict within Syria, is illustrated in Figure 8.1(b). The output of the simulation run concurs with the input data given and the data input method is therefore considered verified.

In Case (iii), the *Probability of Conflict Spreading* was varied to ascertain the degree to which the model concept demonstrator output accurately captures this phenomenon. The data set specified in Table 8.1 was utilised as conflict input and the *Probability of Conflict Depleting* was set to 0 in order to determine exclusively the effect of conflict spreading. Two tests were conducted: 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. The respective simulation model concept demonstrator outputs for Test A and Test B at the end of January 2011 are shown in Figures 8.1(c) and 8.1(d), respectively.

The output of Test A shows only the point where the conflict was initialised as being affected, while Test B correlates with Case (ii), where the data set was simulated at a probability spread value of 1, and a definite spread of conflict is visible. The output of these tests therefore conform to the manner in which conflict spreading was modelled.

Case (iv) is aimed at testing 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 data set provided in Table 8.1 was implemented. As for 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 the effect of depletion. The output of the simulation runs at the end of December 2011 are shown in Figures 8.2(a) and 8.2(b). Test A (Figure 8.2(a)) shows no indication of the conflict depleting, whereas in Test B (Figure 8.2(b)) a notable decrease in conflict is visible.

Finally, the effect of varying the *Intensity of Conflict* was tested in Case (v). For the purpose of consistency, the *Probability of Conflict Spreading* was set to 1 and the *Probability of Conflict Depleting* was set to 0. The data provided as input to the conflict were adjusted to accommodate a varying conflict intensity, as shown in Table 8.2.



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

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

(c) *Case (v)*

FIGURE 8.2: *Cases (iv)–(v) of verifying the modelling of conflict.*

TABLE 8.2: *The data set 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

The simulation model concept demonstrator output at the end of January 2011 for Case (v) is illustrated in Figure 8.2(c). The differences between conflict intensity values of 100, 50 and 10, are notable and the test case verifies the application of this input variable.

Based on the results of Cases (i)–(v) above, the manner in which conflict was modelled in the model concept demonstrator is considered to be successfully verified. A number of input measures which relate to the modelling of conflict were varied and the outputs were compared with logical reasoning during the process of verifying the manner in which conflict was modelled.

8.2 Phase 3: Verification of the modelled population

The manner in which people were modelled 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 states an agent can be in, and an agent's ability to tolerate conflict based on its attributes. A number of cases were considered during the model concept demonstrator implementation and the output of the simulation runs are explored in this section:

Case (vi): Agent attribute statistics,

Case (vii): The fluctuation of the agent population,

Case (viii): The ageing of a single agent,

Case (ix): An agent changing states,

Case (x): The tolerance threshold of agents, and

Case (xi): The geographic distribution of agents.

In Case (vi), the modelled agent attributes were investigated to verify their correlation with the agent attributes introduced in §7.3.3. A series of 30 simulation runs were executed and data pertaining to the current population were recorded in order to determine a mean distribution per attribute for the different criteria. The correlation was confirmed by ensuring that the percentage error between the simulated and actual distributions was below one percent, as can be seen in Table 8.3 (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. Every agent has a probability p of having international family, where p is drawn from a uniform distribution between 5% and 20%. The output indicates that, over the 30 runs executed, the average probability of an agent having international family was 12.57%, which aligns with the input data.

Table 8.3: *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%

In Case (vii), the fluctuation of population size over the 6-year simulated period, taking births and deaths into account, were investigated. No conflict was initiated during the simulation experiment in order to disregard deaths which might occur as a result of conflict. As the simulation model concept demonstrator accounts for the annual births at a single annual time instant, the graph shown in Figure 8.3 indicates a sharp increase in population size annually, followed by a slight decrease which accounts for the occurrence of deaths over time. During 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 concept demonstrator. The reason why births are modelled to occur annually at a single time instant is that this approach is computationally less expensive than creating various events which would spread the occurrence of births over the year.

Another modelled element which requires verification is the ageing of agents, as addressed in Case (viii). In order to test whether the model concept demonstrator incrementally increases the

FIGURE 8.3: *The fluctuation in population size verified.*

age of an agent on an annual basis, a simulation experiment was executed with only one agent in the population. A dynamic textbox indicating the age of the agent was implemented to show the age of the agent during the simulation so as to assist the modeller in verifying the increase in age as depicted in Figure 8.4. The agent had an initial age of 24 years which increased to an age of 29 years over the 6-year modelled period.

(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 8.4: *The verification of an agent ageing.*

In Case (ix), the simulation model concept demonstrator was altered to change the colour of an agent as the agent changes states. Each state was associated with a specific colour in order to ascertain whether or not the agent progressed to the correct state according to its modelled action. An agent in the `Residing` state is coloured black (■) and as soon as it progress to the

`SearchMove` state, where it searches for a safer destination, it is coloured gold (). An agent that has selected a new destination is coloured light blue () as it begins to move towards that destination. When an agent arrives at its destination, whether within Syria or in another country, it is coloured green () as it enters the `CheckSurroundings` state. If, however, an agent fails to find a destination for a certain period of time searching, thereby directing it to the `Stuck` state, it is coloured pink (). If it were to escape the stuck state and progress to the `MoveThroughConflict` state, it is coloured purple (). In the test simulation experiment executed, a logical walk-through approach was taken and the visual animation of agents changing states conformed to the imposed modelling rules. A screenshot of the simulation experiment, shown in Figure 8.5, illustrates agents active in some of the above-mentioned states by means of the colour variation of their presentation element.

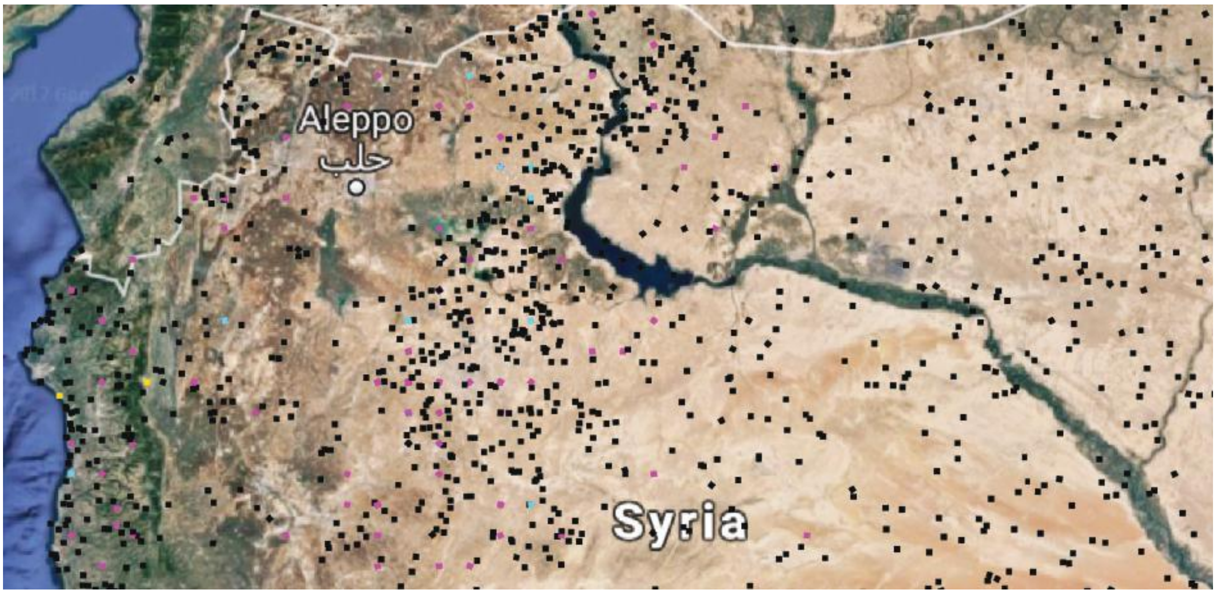


FIGURE 8.5: *An example of the agents' states and associated colours during a simulation experiment.*

Case (x) is aimed at verifying the `ToleranceThreshold` variable allocated per agent, modelled according to the discussion in §7.3.3, which indicates an agent's ability to withstand conflict. A number of scenarios were considered in which the average tolerance threshold of agents sharing similar attributes were compared with reality (as summarised in §7.3.4). In Scenario A, people above the age of 65 who do not have any family internationally are considered. These agents are expected to have a relatively high tolerance threshold. In Scenario B, adults between the ages of 20 and 40, who received tertiary education, are considered. These people should have a relatively low tolerance threshold. In Scenario C, people of low economic status who do not have tertiary education are considered. These individuals should also have a relatively high tolerance threshold. Thirty simulation runs were executed in order to calculate the average `ToleranceThreshold` value of the agents and the associated standard deviation in each scenario. The model concept demonstrator output is shown in Table 8.4.

TABLE 8.4: *The output from Case (x) verifying the allocation of a tolerance threshold to agents.*

	Average	Standard Deviation
Scenario A	0.672	0.029
Scenario B	0.425	0.048
Scenario C	0.671	0.014

The output indicates that Scenarios A and C have equally high tolerance threshold values, while Scenario B's tolerance threshold is less than two-thirds of the values associated with Scenarios A and C. The output correlate with the assumptions made in §7.3.3 with respect to certain characteristics influencing a person's inclination to move and the small associated standard deviations prove that the V `ToleranceThreshold` values do not deviate too much from this norm.

The geographic distribution of people is considered in Case (xi). The modelling thereof, as described in §7.3.3, ensures that the initial population of Syrians is dispersed across the country, mimicking reality. The spread of agents (marked as small black squares) over Syria is illustrated in Figure 8.6. 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 where the major cities in Syria are located.

FIGURE 8.6: *The geographical distribution of people in Syria.*

8.3 Phase 3: Verification of the modelling of decision making

As described in §7.3.4, modelling the decision-making process of agents includes implementation of the logic related to an agent's decision in terms of which movement type to adopt and then, depending on its choice, the selection of an alternative 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 adopted, based on attributes,

Case (xiii): The attractive zones of IDPs and the associated weighted criteria,

Case (xiv): The effect of varying openness scores, and

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

Case (xii) is aimed at investigating 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 with no tertiary education, who 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 §7.3.4, a person of Type A is likely to choose to move as an IDP (movement type 1), while a person of Type B is likely 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 concept demonstrator was adapted for the testing of this case in order to calculate an agent's choice of movement type at initiation of a simulation run. Three agents, each subscribing to the characteristics of the above-mentioned three types, were selected randomly during each run and their movement types were noted.

After 30 simulation runs had been executed, the movement types most commonly associated with the three different types of agents proved to be similar to what had been assumed in modelling this decision. The output of the simulation experiment is documented in Table 8.5. Cases A and C associate predominantly with movement types 1 and 3, respectively, and, although the mode of Case B associates with movement type 2, it can be seen that the set of characteristics portrayed by Case B associate predominantly with both movement types 1 and 2.

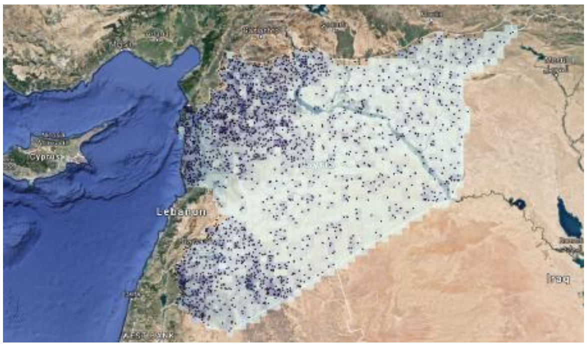
Table 8.5: *Verification of the allocation of movement types.*

	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

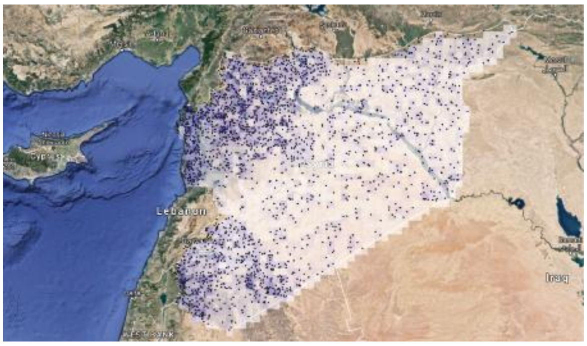
In order to verify the inclusion of so-called *Attractive Zones*, Case (xiii) employs existing data portraying the conflict in Syria during the simulated period in order to compare the changes in *Attractive Zones*, based on conflict zones and population density, as a simulation experiment progresses.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

At the initiation of a simulation experiment when no conflict has been initialised according to the manner in which the *Attractive Zones* are calculated (as explained in §7.3.4), the attractiveness per cell should equal the normalised population density value per cell. The *Population Density* layer is graphically illustrated in Figure 8.7(a) and the *Attractive Zones* at the initiation of the simulation run is illustrated in Figure 8.7(b). 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 with 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 contributes 0.9 towards the weight and the population density factor contributes the balance. The spread of conflict at the end of the simulated period, illustrated in Figure 8.8(a), is the same for all three scenarios owing to the fixed seed (configured for this case) and the other parameters being kept constant. For this scenario, the *Population Density* is shown in Figure 8.8(b) and the *Attractive Zones*



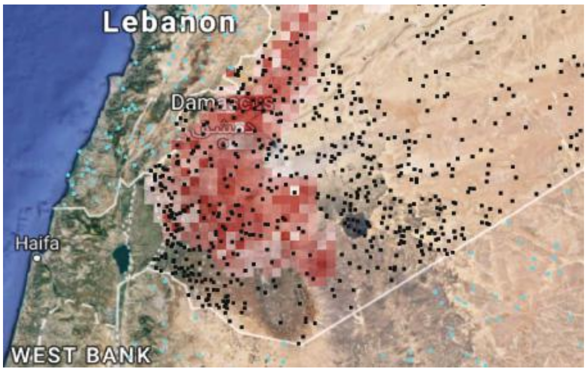
(a) *Case (xiii) Population Density* at the start of the simulation run



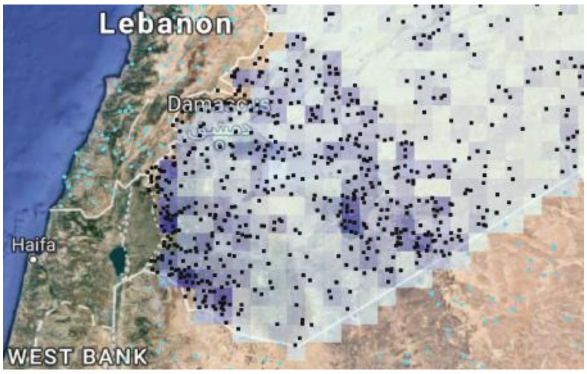
(b) *Case (xiii) Attractive Zones* at the start of the simulation run

Figure 8.7: *The Population Density and Attractive Zones layers as shown at the start of a simulation run.*

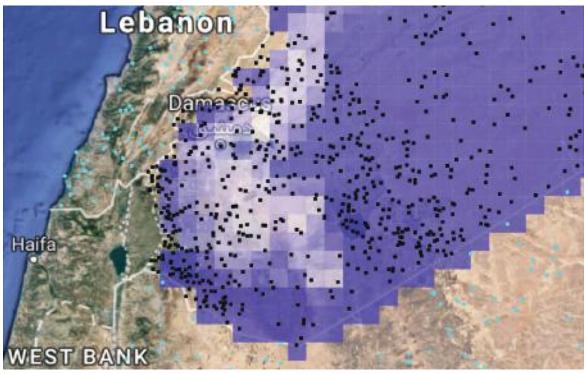
are shown in Figure 8.8(c). Both the *Attractive Zones* and the *Conflict* are superimposed over the simulated area within Syria in Figure 8.8(d).



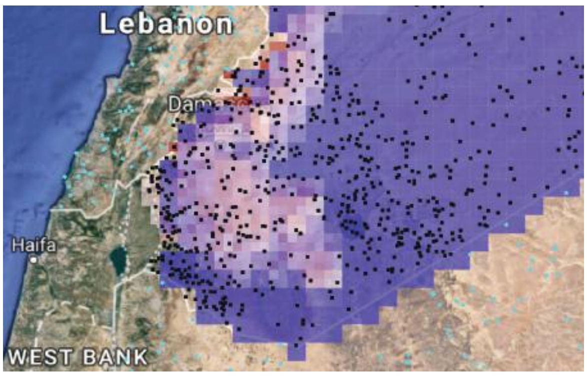
(a) *Case (xiii): Scenario I Conflict*



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



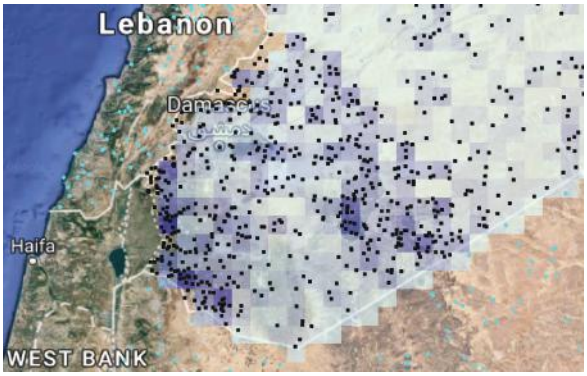
(c) *Case (xiii): Scenario I Attractive Zones*



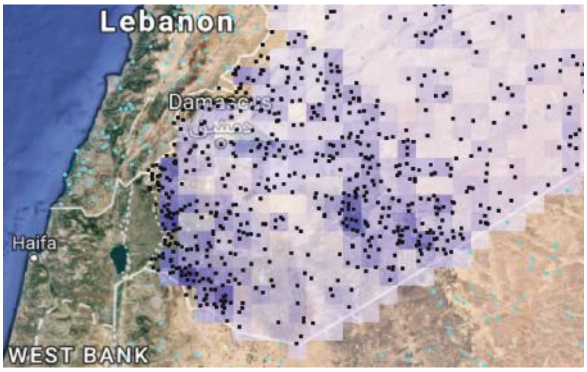
(d) *Case (xiii): Scenario I Attractive Zones and Conflict*

Figure 8.8: *Scenario I considered during the verification of Attractive Zones.*

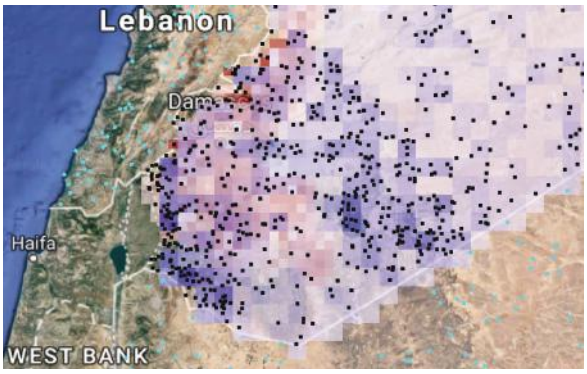
The areas which are more attractive to IDPs are identified by the dark purple hue, as shown in Figures 8.8(c) and 8.8(d), and it is apparent that the location of these areas is almost an exact inverse of areas where conflict is present. In Scenario II, the criteria weights are altered such that the conflict factor contributes 0.1 towards the total weight and the population density factor contributes to the balance. The *Population Density* as at the end of the simulation experiment is shown in Figure 8.9(a), with the *Attractive Zones* illustrated in Figure 8.9(b). The effect when the *Conflict* and *Attractive Zones* layers are superimposed over the simulated area is illustrated in Figure 8.9(c). 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 8.9(b), are notably similar to the *Population Density* shown in Figure 8.9(a), indicating that the conflict imposes a significant effect.



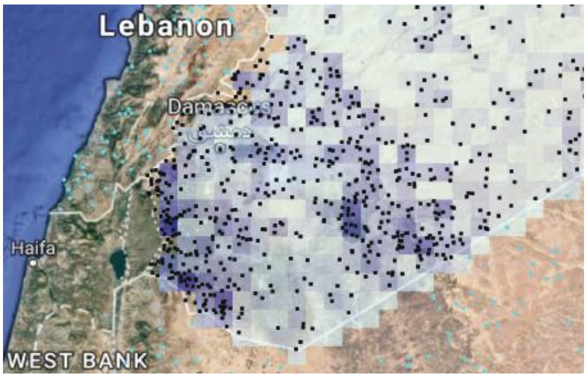
(a) *Case (xiii): Scenario II Population Density*



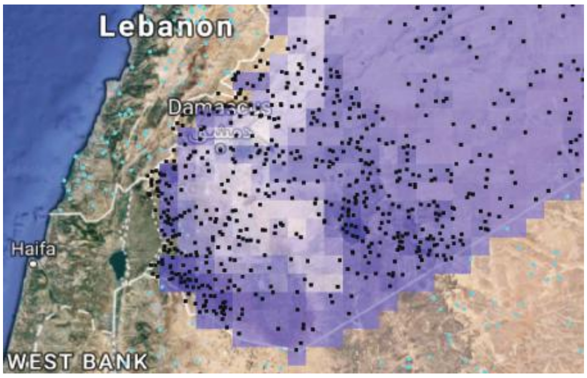
(b) *Case (xiii): Scenario II Attractive Zones*



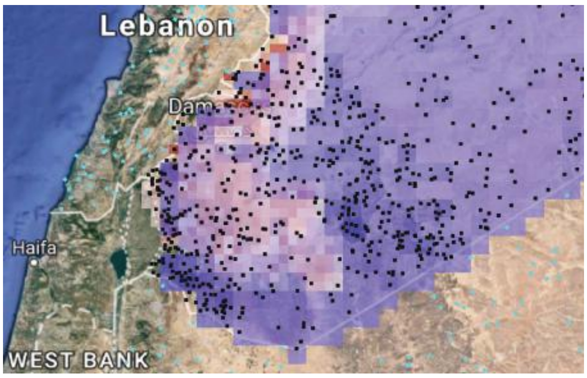
(c) *Case (xiii): Scenario II Attractive Zones and Conflict*



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



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



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

FIGURE 8.9: *Scenarios II and III considered during the verification of Attractive Zones.*

Finally, in Scenario III, the weights for the criteria influencing the *Attractive Zones* are adjusted to prescribe to the default model concept demonstrator values so that conflict contributes 0.6 towards the weighting and population density contributes the balance. The *Population Density*, *Attractive Zones* and a combination of both these layers for this scenario are illustrated in Figures 8.9(d), 8.9(e) and 8.9(f), 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. As a result of the three scenarios executed, the process determining the *Attractive Zones* and the effect of the weighted criteria are considered successfully verified.

The effect of the indices representing the openness of neighbouring countries towards refugees and undocumented migrants are considered in Case (xiv). In order to verify this, the initial openness index of 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 would select as destination are set to a distance, popularity and openness score ratio of 0.05:0.05:0.9, respectively, and kept constant for purposes of consistency.

The fluctuation of openness indices over the simulated time was disabled in this case in order to focus on the effect of the openness index itself and the simulation runs were executed over the full 6-year period. The output yielded when varying these indices and allocating each neighbouring country the significantly higher openness index value in turn, are illustrated in Figures 8.10–8.14. In each case, it is apparent that the neighbouring country with an allocated openness index of 80 has a significantly higher influx of Syrians over the simulated period than all the other neighbouring countries.

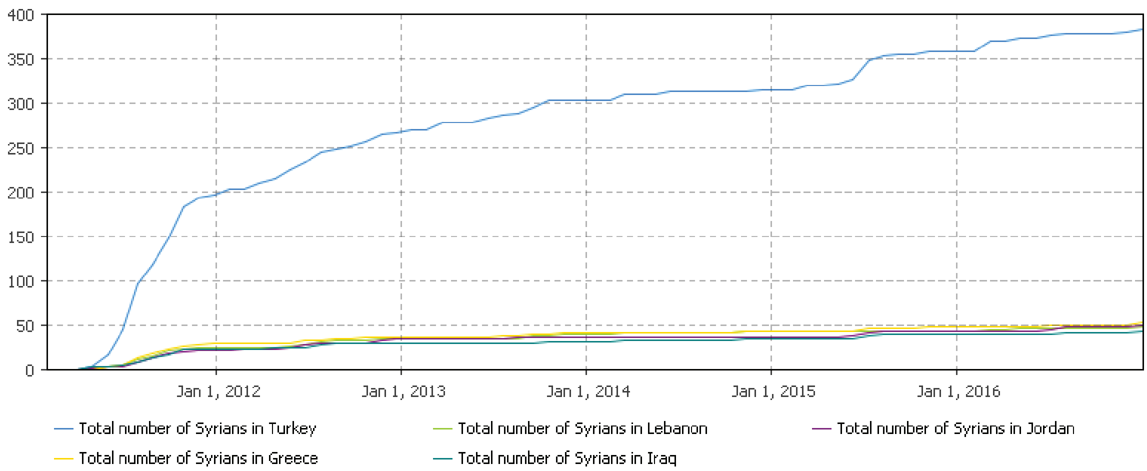


FIGURE 8.10: *Allocation of an Openness Index of 80 to Turkey.*

The purpose of Case (xv) is to investigate the movement of an agent when adopting different movement types. Scenarios considering IDPs, refugees and undocumented migrants where therefore investigated. The colour of an agent identified in each scenario was changed to white in order to make the agent more distinguishable for movement tracking purposes. All parameters were set to default values and the conflict was initialised manually at the location of the agent in order to investigate its reaction. In each scenario, a random agent was chosen and the progression through its statechart was documented. The output yielded for these scenarios are depicted in Figures 8.15, 8.16 and 8.17, where a white circle has been placed over the agent under investigation.

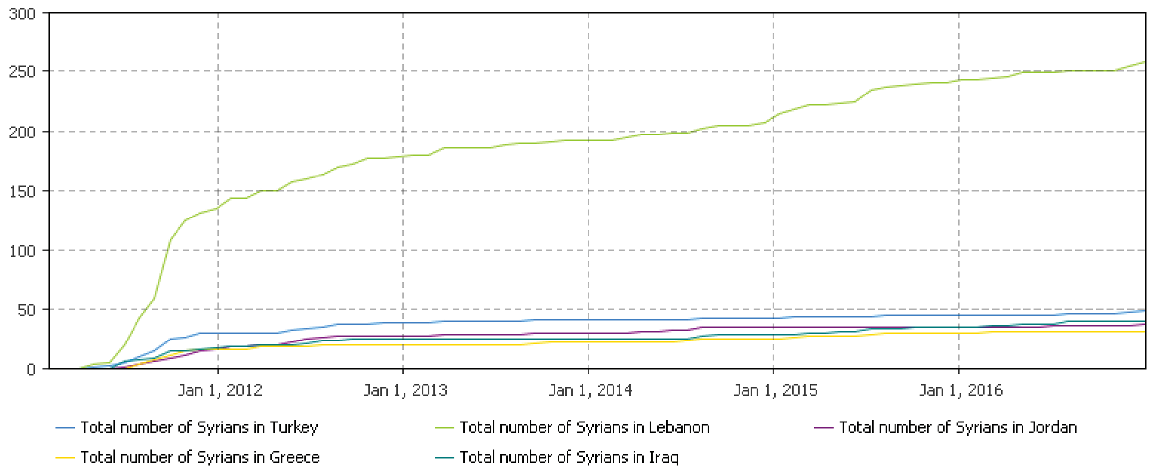


FIGURE 8.11: *Allocation of an Openness Index of 80 to Lebanon.*

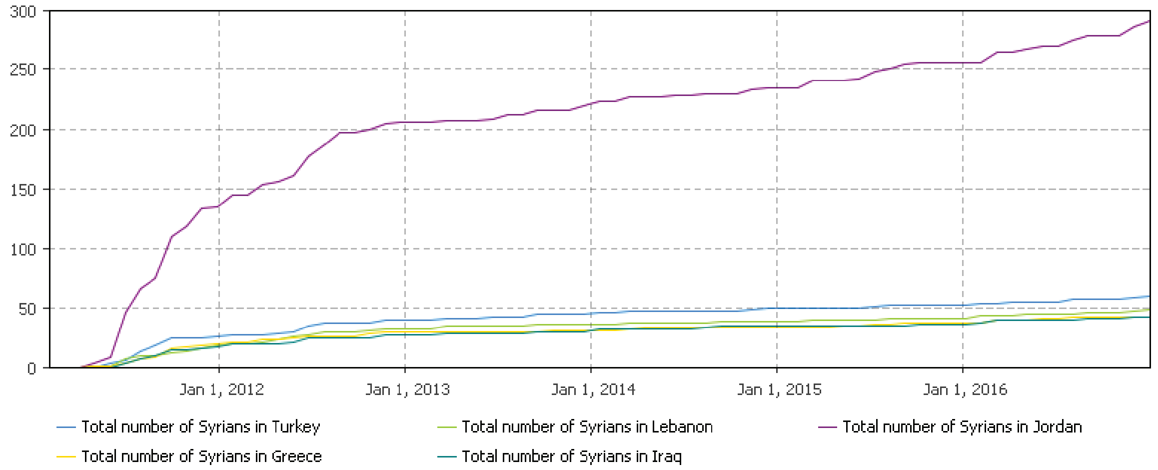


FIGURE 8.12: *Allocation of an Openness Index of 80 to Jordan.*

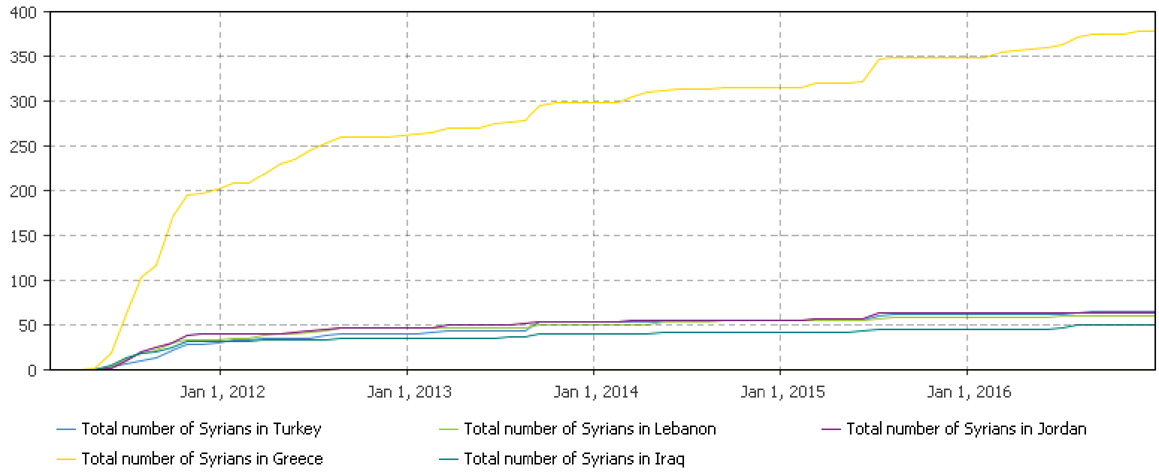


FIGURE 8.13: *Allocation of an Openness Index of 80 to Greece.*

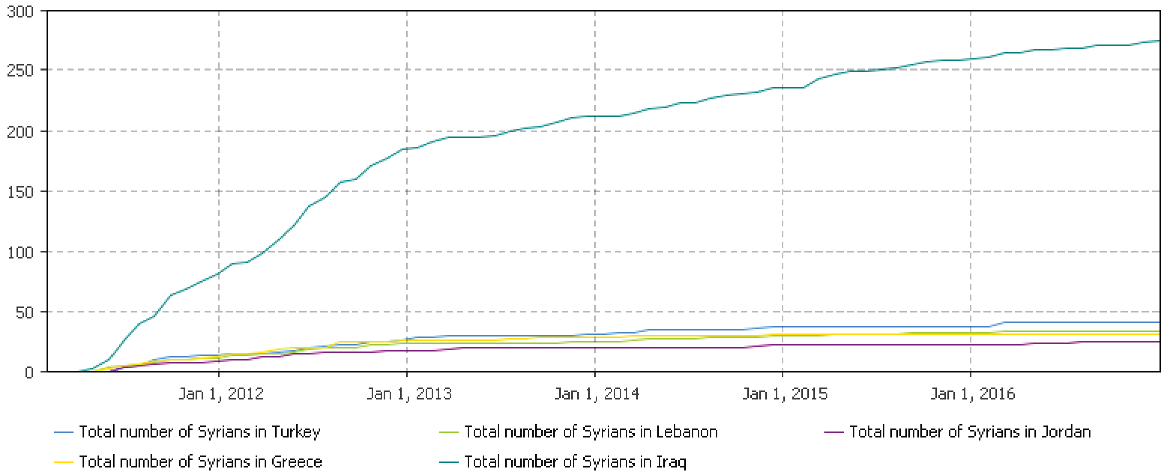
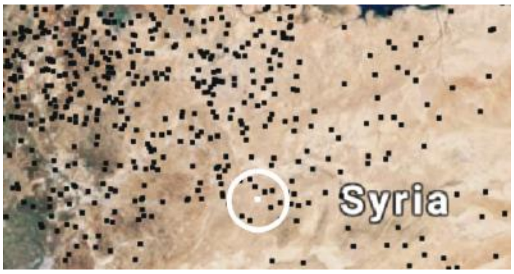


FIGURE 8.14: *Allocation of an Openness Index of 80 to Iraq.*



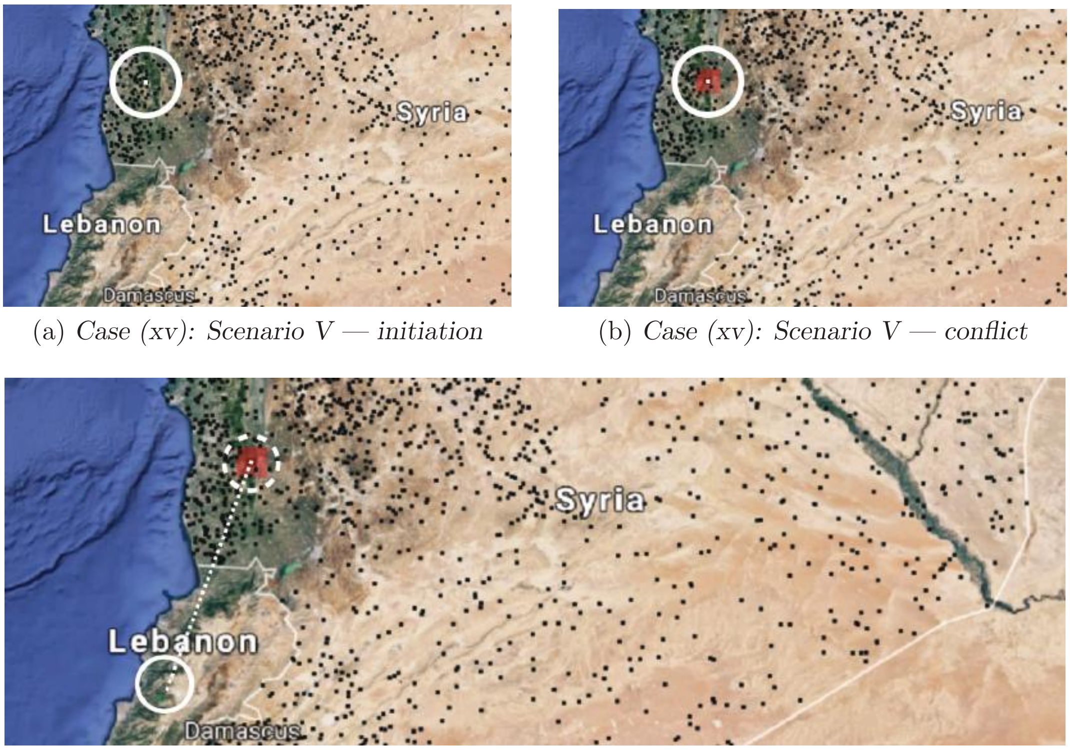
(a) *Case (xv): Scenario IV — initiation*

(b) *Case (xv): Scenario IV — conflict*

(c) *Case (xv): Scenario IV — tracking movement*

FIGURE 8.15: *Verifying the movement of Syrian IDPs.*

Scenario IV represents the movement of an agent who, when exposed to conflict exceeding its threshold, chooses to relocate within Syria. The movement of the observed agent from the initiation of the simulation experiment to the point of relocation is shown in Figure 8.15. The agent had a tolerance threshold of 0.49 (49%), which was exceeded by conflict, as shown in Figure 8.15(b), with an intensity value of 80. The agent decided to adopt movement type 1 and identified a destination within Syria in a north-westerly direction from its original location. The agent then moved there, as shown in Figure 8.15(c). The agent subsequently checked its surroundings upon arrival and deemed the area to be safe before progressing to the state of `Residing`.



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

(b) *Case (xv): Scenario V — conflict*

(c) *Case (xv): Scenario V — tracking movement*

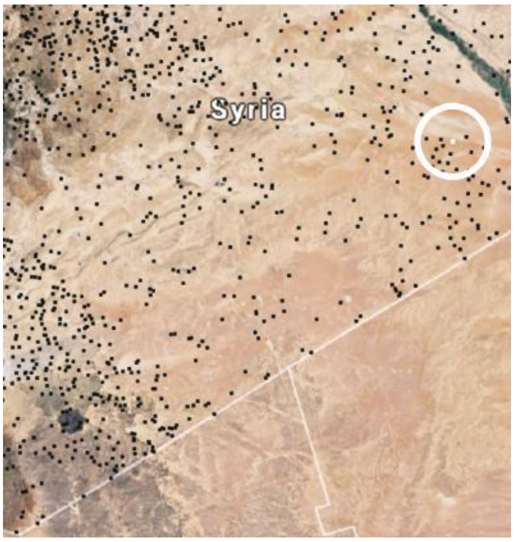
FIGURE 8.16: *Verifying the movement of a Syrian refugee.*

In Scenario V, the movement of a refugee is considered and illustrated in Figure 8.16. The agent had a tolerance 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 8.16(b). The agent decided to adopt movement type 2 and chose Lebanon as the neighbouring country of destination in which to seek refuge. The movement of the agent from its original location to Lebanon is illustrated in Figure 8.16(c). Upon arrival, the agent immediately entered the `Residing` state, as conflict outside of Syria is not considered in the model concept demonstrator.

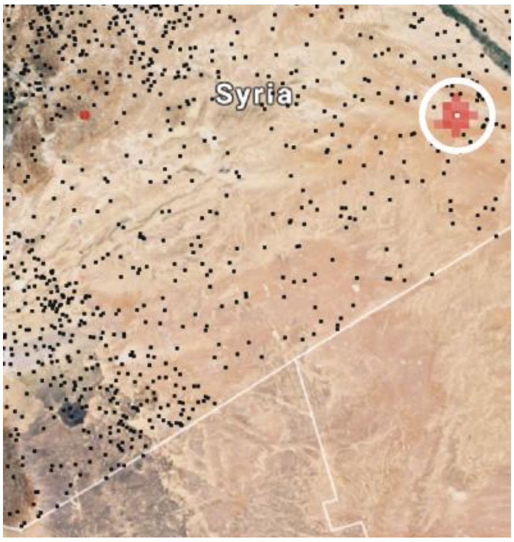
The movement of an undocumented migrant exhibited by a random agent is followed in Scenario VI and depicted in Figure 8.17. The agent possessed a tolerance threshold of 0.15 (15%) which was exceeded by a conflict intensity of value 62, as shown in Figure 8.17(b). After adopting movement type 3, the agent selected Jordan as its destination country and moved there, as illustrated in Figure 8.17(c). The output of all three scenarios investigated in Case (xv) indicate that the behaviour of agents randomly chosen to represent each of the movement types conform to the predicted behaviour as modelled.

8.4 Chapter summary

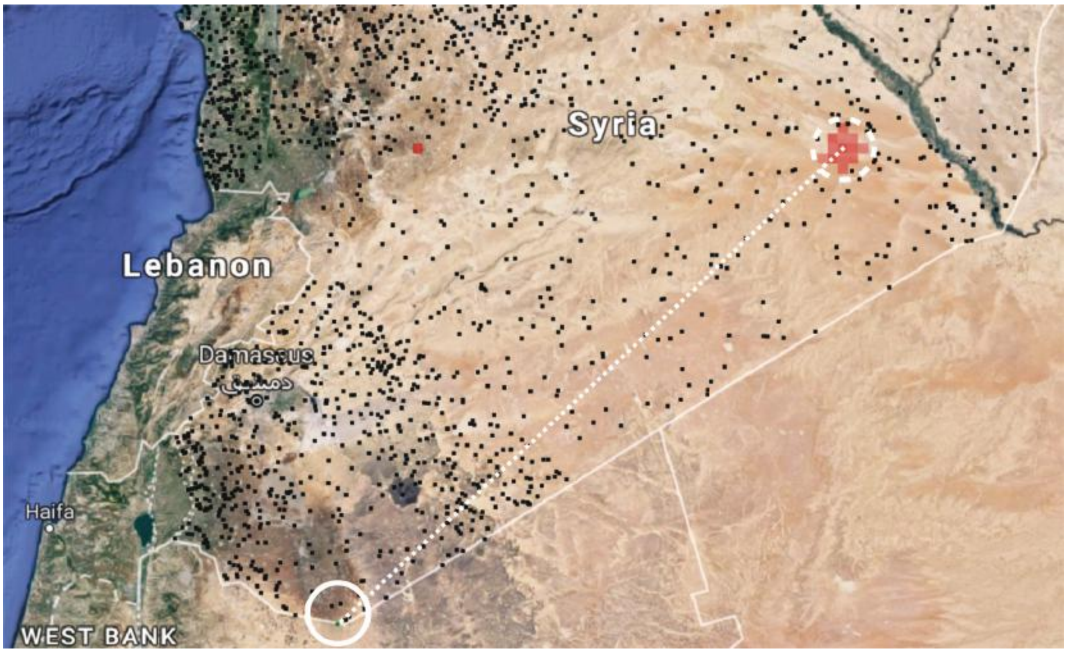
The verification processes followed in order to ensure the correct operation of the different components of the model concept demonstrator were described in this chapter. This included a verification of the modelling of conflict, the population of agents residing in the simulation and the subsequent decision making of these agents when confronted with conflict. Illustrative and numeric outputs were gathered from the model concept demonstrator during specific tests in order to assess the degree to which the aforementioned model concept demonstrator components



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



(b) *Case (xv): Scenario VI — exposure to conflict*



(c) *Case (xv): Scenario VI — tracking movement*

FIGURE 8.17: *Verifying the movement of an undocumented migrant from Syria.*

perform as intended. In all cases, the implementation was found to be sound and accurate, and the response of the model concept demonstrator appears to be in line with what was expected from its construction. In light of this, the model concept demonstrator is deemed to have been verified.

CHAPTER 9

Validation and analysis of the model concept demonstrator

Contents

Phase 4 of the CoFMMA framework entails the model execution. Validation and output analysis, also included in this phase, are performed in this chapter on the model concept demonstrator developed and discussed in Chapter 7. The manner in which conflict is modelled in the model concept demonstrator is calibrated and validated in §9.1 and §9.2, respectively. This is followed by a validation of the agent aggregation in §9.3 and a face validation in §9.4, based on corroboration by subject matter experts in respect of various important components of the model concept demonstrator. Existing data are then considered by means of a parameter establishment analysis in §9.5 where suitable model parameters are sought to recreate the documented scenarios. In light of the model concept demonstrator's capability to recreate specific scenarios pertaining to conflict outbreak and the associated flight of forcibly displaced people, parameter variation is employed in §9.6 to investigate the efficacy of the model concept demonstrator as a decision support and analysis tool.

9.1 Phase 4: Calibration of parameters used to model conflict

In an attempt to validate the spread of conflict as modelled within the model concept demonstrator, the population of agents was temporarily disregarded. The primary parameters associated with the modelling of conflict, `ProbabilityofInfection` and `ProbabilityofDepletion`, were varied in an attempt to establish values which best replicate the historical spread of conflict in Syria based on historical data that exist on record in the form of graphical mapping. The other parameters and variables of the model concept demonstrator were kept constant and a fixed seed random number generator was implemented to ensure consistency between 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 9.1. The visual model concept demonstrator outputs for each of these experiments are documented in Appendix A.

TABLE 9.1: *Experiments conducted and corresponding parameter values considered 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

The general conflict distribution and spread are relatively similar in all of the experiments performed, although, as expected, the extent of the spread of conflict is greatest when the `ProbabilityofInfection` tends towards 1 and the `ProbabilityofDepletion` tends towards 0. The model concept demonstrator is therefore configured as having a `ProbabilityofInfection` and a `ProbabilityofDepletion` drawn from a triangular distribution with 0 as minimum, 1 as maximum and 0.75 as mode. The addition of a mode value is rather trivial, but experimental configurations in which the two parameter values were varied between 0.5 and 1 indicated an overall slightly better presence of conflict. These distributions were next subjected to validation, as described in the following section.

9.2 Phase 4: Validation of the modelling of conflict

Validation is commonly achieved by utilising a simulation model to replicate empirical real-world data. In this case, numerical data with respect to the spread and depletion of conflict in Syria do not exist, but graphical data indicating the state of conflict at various time instants throughout the conflict are available. AnyLogic facilitates animation of the model concept demonstrator 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 were then compared with this simulated output and, in doing so, the degree of visual replication contributed to confidence in the model concept demonstrator's accuracy.

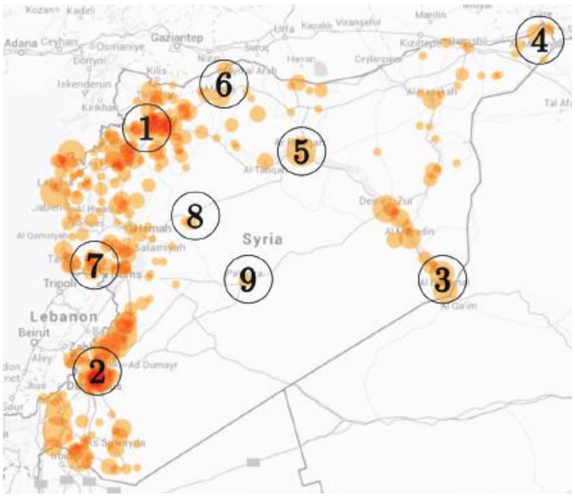
Visual data pertaining to the state of conflict in Syria during specific time instants were gathered from The Carter Center [28], GDELT [59] and Palantir Technologies [58]. Simulation experiments were conducted using a fixed seed for consistency and the `ProbabilityofInfec- tion` and `ProbabilityofDepletion` parameters were drawn from a triangular distribution, as described in §9.1, while all other parameters were held constant with no agents included in the simulation runs. The model concept demonstrator output for three specific time instants were compared to the visual data available, as shown in Figures 9.1, 9.2 and 9.3. In each of these visual comparisons, numerical annotations are superimposed on the images for the purpose of analytical discussion.

Figure 9.1 facilitates a comparison between two sources who recorded 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 9.1(a) by means of a single colour intensity, whereas in Figure 9.1(b) the intensity of conflict ranges over a colour spectrum from dark blue to red, where the latter suggests a high intensity. High conflict intensity is present at points (1) and (2) on the visual data, as recorded by GDELT and Palantir Technologies, as well as in the simulated output. Furthermore, at point (3) conflict is present along the Euphrates river in both the simulated and recorded graphical representations of the conflict. The model concept demonstrator further recorded conflict with relatively high intensity at point (4), whereas both simulated visualisations, although agreeing with the presence of conflict at this point, presented it as having a lower intensity. Although the simulated model concept demonstrator emulated conflict at and immediately surrounding point (4), it did not replicate the spread of this conflict inland in the same manner that the records of Figures 9.1(a) and 9.1(b) do.

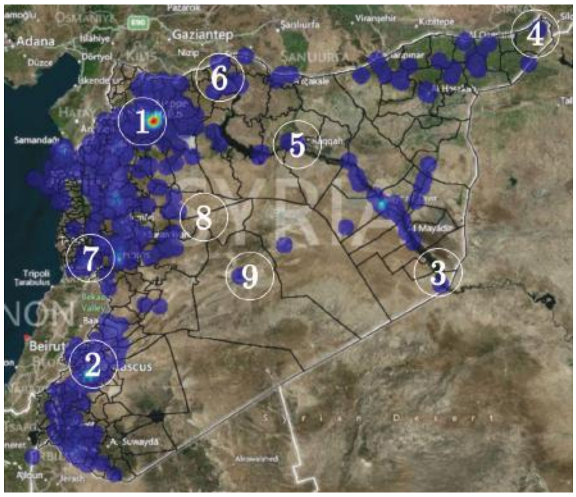
Along the Euphrates river, at points (5) and (6), correlation between the simulated and recorded output is evident, although the simulated version of conflict at point (6) is less intense and not as far-reaching than that of the true data. The presence of conflict between points (1) and (7), spreading along the Syrian border and further inland, is similar in both the simulated output and the visualised data. At point (8), the presence of conflict in central Syria is shown 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, conflict was present at point (9) as recorded in Figures 9.1(a) and 9.1(b), although this area of conflict was not replicated in the simulated model concept demonstrator output. In general, the spread of conflict at this time instant, as recorded, correlates well with the simulated output, especially along 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 further compared with the conflict as recorded for the same time period by the UNHCR in Figure 9.2. The visual record of the UNHCR shown in Figure 9.2(a) did not, however, indicate the intensity of conflict, but only its presence in certain regions.

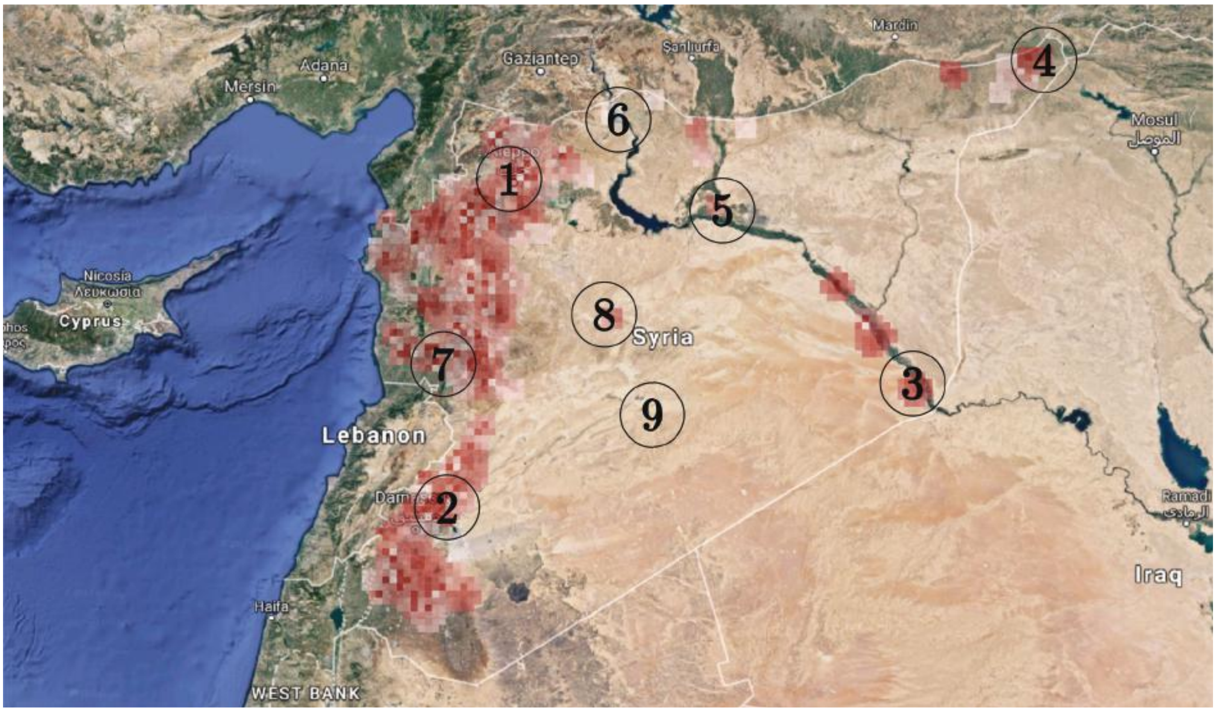
Similarly to the previous time instant analysed, the majority of conflict occurred along the northern and north-western borders of Syria, stretching between points (11), (10) and (14). The



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



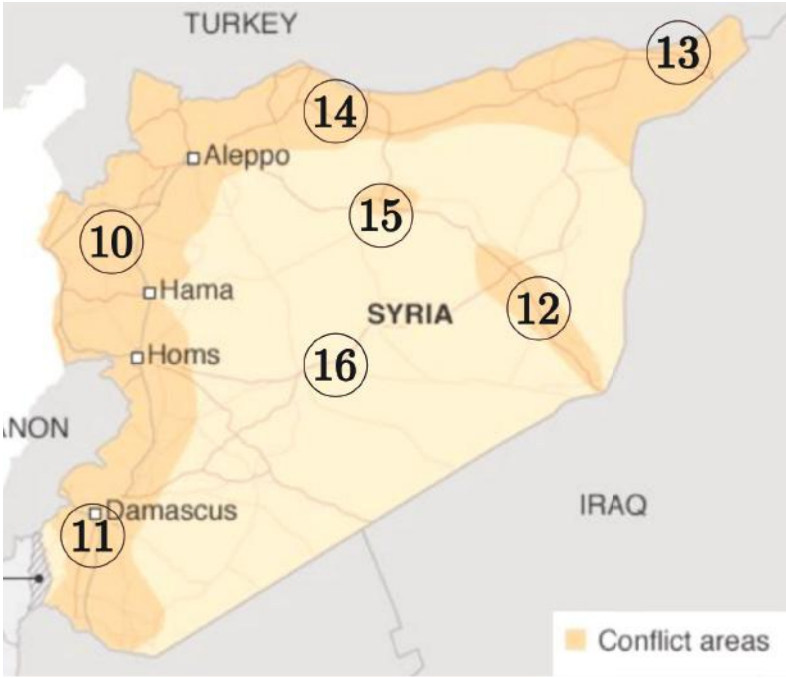
(b) *The spread of conflict as recorded in June 2014 by Palantir Technologies*



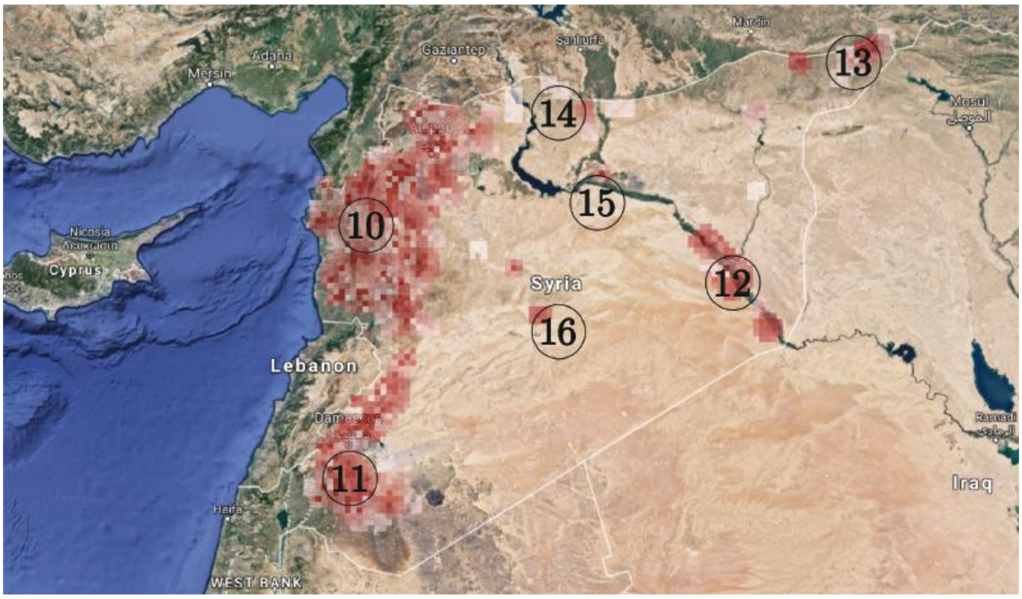
(c) *The spread of conflict as simulated in June 2014*

FIGURE 9.1: *A comparison of the spread of conflict simulated and that recorded in June 2014 by GDELT and Palantir Technologies [58, 59].*

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) was present in the recorded data from the UNHCR, while the simulated output indicated 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 was aggregated and excludes consideration of the intensity of the conflict. A correlation in the conflict along the Euphrates river, at point (12), is evident between both the simulated and recorded visualisations. Similarly, at point (15) in Figure 9.2(a), conflict was present at a specific location with little spread, which was accurately replicated by the simulated output. At point (16), however, the simulated output, shown in Figure 9.2(b), indicated the presence of conflict, whereas this was not the case in the recorded conflict from the UNHCR.



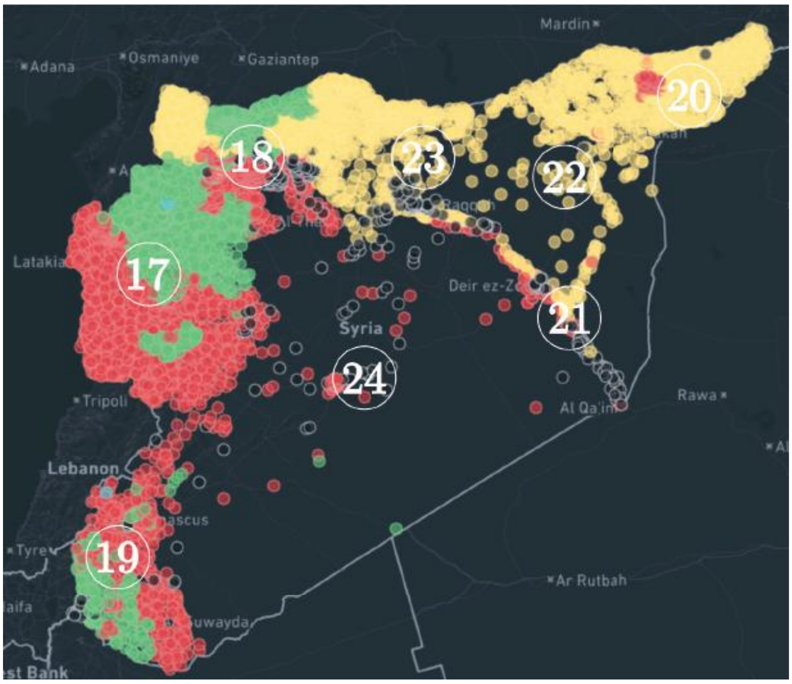
(a) *The spread of conflict as recorded in July 2015 by the UNHCR*



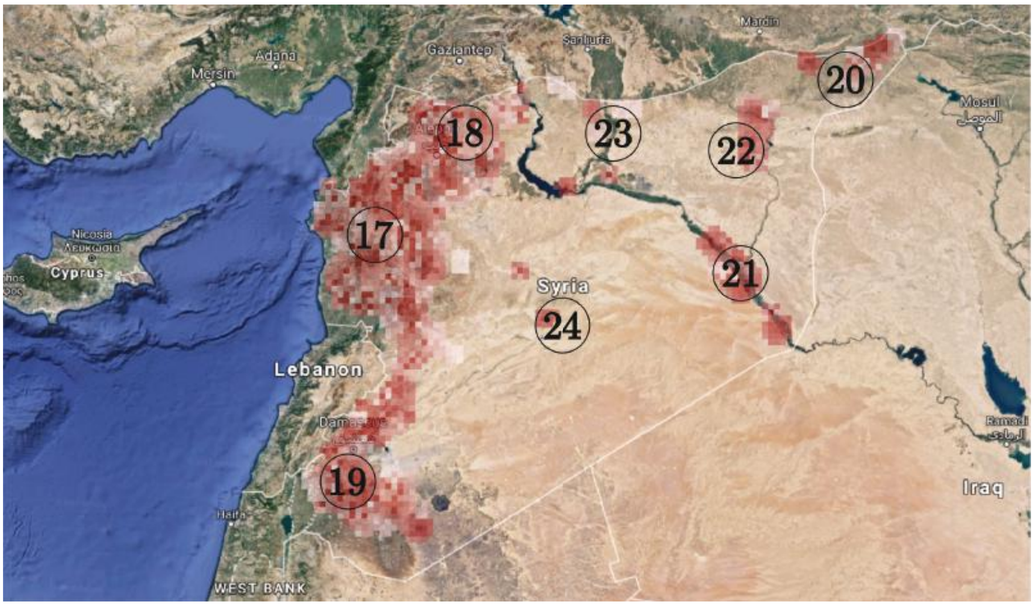
(b) *The spread of conflict as simulated in July 2015*

Figure 9.2: *A comparison of the spread of conflict simulated and that recorded in July 2015 by the UNHCR [169].*

Finally, the conflict as simulated in December 2016 was compared with the conflict recorded by The Carter Center, at the same time instant, as shown in Figure 9.3. The illustration of The Carter Center data shown in Figure 9.3(a) did not include the intensity of conflict, although the different parties involved in the conflict were included, indicated here by means of colour. Red refers to the government, green refers to the opposition, black refers to ISIS, yellow refers to YPG and its allies, and blue refers to ceasefires.



(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 9.3: *A comparison of the spread of conflict simulated and that recorded in December 2016 by The Carter Center [28].*

The conflict recorded in Figure 9.3(a) stretches from point (19), past points (17) and (18), towards point (23). The simulated output showed a similar spread of conflict between points (19) and (18), although it did not indicate such a wide spread of conflict between points (18) and (23). The Carter Center data furthermore indicated the presence of conflict 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 concept demonstrator accounted 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 did not 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 mimicked the scarcely dispersed conflict shown in Figure 9.3(a) in the triangular area between points (17), (23) and (24).

In general, the data on record for the three time instants considered indicated high density of conflict occurring along the northern and north-western borders of Syria, in the vicinity of Aleppo and Damascus. Conflict was also present at the north-eastern corner of the country, where it borders Turkey and Iraq. Furthermore, notable conflict occurred along the Euphrates river from where it enters Syria at the Iraqi border. The conflict simulated by the model concept demonstrator indicated a reasonable replication of the actual spread of conflict as described, and visually correlated with the recorded data. This indicates that the manner in which conflict has been modelled in the agent-based model concept demonstrator developed, is capable of producing meaningful similarities to the manner in which it develops and depletes in reality.

9.3 Phase 4: Validation of agent aggregation

The model concept demonstrator employed an aggregation in which one agent represents 10 000 people of similar attributes, as discussed in §7.2.3 and §7.3.3. In order to validate the assumption of aggregation and ensure that it does not compromise the model concept demonstrator in representing reality, a statistical analysis was performed.

During this analysis, the model concept demonstrator was executed at various levels of granularity in order to determine where a statistically significant difference in output results. In Table 9.2, nine experiments are provided, as well as the corresponding level of granularity at which the model was executed for each experiment. The parameter associated with the granularity of the model is the `InitialPopulation` which increases as the granularity increases.

Table 9.2: *Experiments conducted during the validation of the granularity along with the corresponding levels of granularity.*

Experiment	Ratio	Initial agent population
26	1 : 13 953	1 700
27	1 : 10 000	2 372
28	1 : 9 123	2 600
29	1 : 7 413	3 200
30	1 : 6 242	3 800
31	1 : 4 744	5 000
32	1 : 2 965	8 000
33	1 : 2 372	10 000
34	1 : 2 000	11 860

In Experiment 26, for example, a low granularity is chosen where one agent represents 13 953 people and, as a result, the `InitialPopulation` value is 1 700 (as the actual initial population at the initialisation of the simulation should amount to 23 720 000 people). Other parameters in the model concept demonstrator were kept constant and a fixed seed random number generator was employed to ensure consistency between experiments. The `InitialPopulation` parameter value was varied from 1 700, which is representative of a low granularity, to 11 860, which represents a higher granularity.

For each experiment various output parameters were considered, such as the total population of people, the total number of Syrians in each of the respective countries, and the total number of Syrians per movement type, as shown in Tables 9.3 and 9.4. In order to determine whether there exists a statistically significant difference in the output data across the levels of granularity, two t-tests were conducted. The first t-test pertained to a comparison between the output of a particular data set and that of Experiment 27, where one agent represents 10 000 people, which is the default value implemented in the model concept demonstrator, while the second t-test was carried out in respect of a comparison between the set of output data and Experiment 34, where one agent represents merely 2 000 people.

TABLE 9.3: *The output of the t-tests conducted when considering the total Syrian population size as well as the number of Syrians residing in Syria and each of its neighbouring countries.*

Experiment	Population	Syria	Turkey	Greece	Lebanon	Iraq	Jordan
26	24 319 976	19 422 494	1 841 788	990 659	795 318	697 647	572 071
27	25 220 000	19 860 000	2 300 000	990 000	880 000	590 000	600 000
28	24 896 877	19 194 954	2 207 785	1 012 662	1 012 662	766 338	702 477
29	25 387 813	20 124 938	2 060 675	963 625	1 059 988	526 288	652 300
30	25 018 358	20 087 095	1 910 084	936 316	898 863	555 547	630 453
31	25 731 456	20 266 368	2 035 176	1 162 280	825 456	692 624	749 552
32	25 709 515	20 689 770	2 007 305	996 240	883 570	501 085	631 545
33	25 672 156	20 444 268	2 020 944	1 012 844	882 384	630 952	680 764
34	25 398 000	20 118 000	2 086 000	1 032 000	890 000	504 000	768 000
μ	25 261 572	20 023 098	2 052 195	1 010 736	903 138	607 165	665 240
S_x	461 663	471 386	139 005	63 469	83 514	95 386	65 903
t_{27}	0.27	1.04	−5.35	0.98	0.83	0.54	2.97
t_{34}	−0.89	−0.60	−0.73	−1.01	0.47	3.24	−4.68

In Tables 9.3 and 9.4, the test statistic was calculated as

$$t_i = \frac{x_i - \mu}{S_x/\sqrt{n}},$$

where i is the experiment index, x_i is the output data for experiment i and n denotes the total number of experiments compared. Furthermore, μ and S_x denote the mean and standard deviation of each data set, respectively. The degrees of freedom, denoted by d_f, was taken as

$$d_f = n - 1,$$

while the critical t-value was ascertained from tabulated values. More specifically, for a two-tailed test with statistical significance levels of 1%, 5% and 10%, the critical t-values, denoted by $t_{\frac{\alpha}{2},d_f}$, are

$$t_{0.005,8} = 3.833, \quad t_{0.025,8} = 2.752, \quad \text{and} \quad t_{0.050,8} = 2.306.$$

In comparing the test statistics in Table 9.3 with the aforementioned critical t-values, it was noted that the data output with respect to the population of Syrians in Turkey differs statistically significantly when compared with Experiment 27, while the data output concerning the population of Syrians in Jordan differs statistically significantly when compared with the corresponding value in Experiment 34. It seems that the level of granularity may indeed have an impact with respect to the Syrian populations in some of the neighbouring countries. From Table 9.4 it was, however, evident that there is no statistical significance between the different levels of granularity when considering the number of individuals per movement type. The level of statistical significance can increase up to 24%, where $t_{0.12,8} = 1.740$, before any significant difference is detected.

Table 9.4: *The output of the t-tests conducted when considering the number of people per movement type.*

Experiment	IDPs	Refugees	Undocumented migrants
26	4 646 329	2 232 471	2 678 965
27	4 470 000	2 550 000	2 720 000
28	4 534 169	2 773 415	2 928 508
29	4 291 838	2 379 413	2 801 925
30	4 675 337	2 484 358	2 365 758
31	4 634 888	2 822 680	2 542 784
32	4 308 145	2 626 990	2 330 490
33	4 276 716	2 656 640	2 521 436
34	4 494 000	2 620 000	2 572 000
μ	4 481 269	2 571 774	2 606 874
S_x	157 834	185 888	195 779
t_{27}	0.21	0.35	−1.73
t_{34}	−0.24	−0.78	0.53

9.4 Phase 4: Face validation

Subject matter experts, such as Aksel [4] from Koç University, Frydenlund [55] from the Old Dominion University, Groen [66] from Brunell University, Lemos [103] from the University of Agder, Shomary [149] from Stockholm University, Smith [154] from the University of Sussex and Stewart [160] from the University of Cape Town, were contacted throughout the progression of this study for input and expert recommendations in terms of the implementation of certain aspects of the model concept demonstrator. As a final validation, the expertise of Lemos [103] was utilised to validate the modelling of conflict, while the expertise of Aksel [4] and Frydenlund [55] were utilised in validating the simulated Syrian people and their associated decision-making process.

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 published a book entitled “Agent-based modelling of social conflict” [102] 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 simulation models focussing on protracted refugee situations. Frydenlund’s insight and expertise were sought throughout the development of the simulation model concept demonstrator described in this dissertation. 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 described in Chapter 7.

Stewart is a professor emeritus at the University of Cape Town in South Africa and is an expert in the field of MCDM, having published books on the topic. Groen is a lecturer of simulation and modelling at the Brunel University in England, with a research focus 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 their personal characteristics play in this process.

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

9.4.1 Model 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 during the development of the simulation model concept demonstrator, as described in §7.2. These assumptions are related to geography, time, data and agent attributes. With respect to geography, Groen advised that only the neighbouring countries of Syria should be considered for the sake of computational efficiency. Movement between these neighbouring countries, or possible movement from these countries into other European countries or back to Syria are, however, not considered in the model concept demonstrator. 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 discussed by Shomary was the repatriation of Syrians, where Syrians who had previously relocated to another country chose to return to their places of origin. According to Shomary, this phenomenon is evident, even in some European countries. For the purpose of limiting the complexity of the model concept demonstrator, however, it was agreed that neglecting both the incorporation of transit countries and the aspect of repatriation in the simulation model concept demonstrator 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 affect the manner in which a person reacts towards the conflict. Aksel further recommended referring to studies completed on Afghan forced migrants in an attempt to better understand how the people react to conflict situations. In terms of data with respect to conflict, a suggestion was made by Lemos to utilise the Social Conflict Analysis Database but, by this stage in the model concept demonstrator'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 concept demonstrator 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 supported this recommendation and further recommended only to account for the most essential agent characteristics so as to decrease the model complexity, as well as to minimise the number of assumptions required in modelling the Syrian population.

9.4.2 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 was modelled in the simulation model concept demonstrator is described in §7.3.2. After reviewing this implementation of modelled conflict, Lemos confirmed it to be a sensible and acceptable implementation. Furthermore, he commented on the systematic approach utilised in the construction of the model concept demonstrator, especially with respect to the conflict modelling, noting that it encompasses all fundamental issues without adding unnecessary complexity. Although the model concept demonstrator 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 purposes of this model concept demonstrator. Lemos and Frydenlund agreed that for models with application in an environment other than Syria it may be applicable to include modelling instances of conflict as 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.

9.4.3 The agent population

The manner in which people are modelled in the simulation model concept demonstrator was described in §7.3.3. 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 was, however, suggested by Frydenlund to increase the granularity of the model concept demonstrator in future improvements with the use of supercomputers. Aksel and Frydenlund also recommended to include an agent's social network in future work, which may be captured by adopting a lower level of abstraction in the model concept demonstrator.

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 tolerance 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.

9.4.4 The decision-making process of agents

The decision-making process of an agent in the simulation model concept demonstrator, as described in §7.3.4, comprises an agent being required to adopt a movement type and select a suitable alternative destination. Initially, a decision was made to employ MCDM methods to model these decisions. Stewart, however, explained the implications of utilising prescriptive decision-making methods in an environment which requires descriptive methods, as explained in §4.1 and §7.2.8. It was agreed that employing a simplified modelling approach would more effectively reflect the reality of the decision-making process.

An agent's characteristics govern the movement type it adopts. Lemos suggested utilising exploratory factor analysis to identify 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 his or her proposed movement type. A phenomenon commonly witnessed is people paying smugglers for access across the Syrian border, travelling towards Europe. These individuals usually fall within 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 of people of an advanced age and in a low income bracket or who do not have tertiary education with respect to relocating.

The decision making of an individual with respect to selecting a 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 its place of origin. Aksel, Frydenlund, Lemos, Shomary and Smith explicitly affirmed the manner in which this was modelled. Refugees and undocumented migrants take into account the distances to neighbouring countries, the popularities of these countries (*i.e.* the number of Syrians who have moved 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 approved of the simple incorporation thereof. Lemos agreed with the implementation of openness indices, but suggested considering a country to have a different openness indices 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 concept demonstrator 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 of the simulation model concept demonstrator 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 concept demonstrator such as the one developed in this dissertation and, to the best of the knowledge of this expert panel, no such model exists which considers such a large collection of factors and implications pertaining to forced migration modelling in the presence of conflict. Frydenlund further commented that this model concept demonstrator allows for a serious attempt at good estimates of 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 contributed largely to the validation of the constructed simulation model concept demonstrator in light of the numerous assumptions and simplifications required during its construction.

9.5 Phase 4: 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 was performed in an attempt to fit the model concept demonstrator to existing partial data and, by implication, illustrate the model concept demonstrator's capability to be used to simulate specific scenarios. The model concept demonstrator 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 suitable 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 concept demonstrator in replicating the outputs as recorded.

The first set of experiments is aimed at demonstrating replication of 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 was applied to the relationship between IDPs and refugees exclusively, allowing inference data pertaining to undocumented migrants subsequently to be generated. It is reported by the UNHCR [169] 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 $1 : 0.8182$ when considering only these two movement types.

The function which determines the movement type of an agent, `DetMovementType`, utilises a probability matrix which associates the various agent characteristics with a probability of adopting a certain movement type, as discussed in §7.3.4. The probability matrices for the experiments performed in respect of 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 the results pertaining to the ratio of the number of IDPs to the number of refugees per experiment are listed in Table 9.5.

TABLE 9.5: *The ratios of movement type 1 to movement type 2 produced by model concept demonstrator output.*

Experiment	Movement Type 1 to Movement Type 2 Ratio
35	1.000 : 0.3249
36	1.000 : 0.2953
37	1.000 : 0.3689
38	1.000 : 0.3515
39	1.000 : 0.4132
40	1.000 : 0.4376
41	1.000 : 0.6015

In Experiment 41, the best alternative probability matrix was determined which allows for a ratio that best matches the real-world approximation. The simulation model concept demonstrator, configured in Experiment 41, modelled the number of undocumented migrants at the end of 2015 to account for 2.49 million people. Furthermore, the model concept demonstrator suggested that

460 000 deaths resulted from people being trapped in conflict-effected areas. This correlates with results documented in the literature which estimate that, by the end of 2015, more than 400 000 Syrians had died in the civil war [90, 119].

The second series of experiments conducted concerned the weighted criteria and initial openness indices of countries neighbouring Syria influencing an agent's decision with respect to selecting a country as destination. An existing data set of the UNHCR [169] containing the number of refugees within each neighbouring country at the end of 2015, is shown in Table 9.6. Owing to the sporadic estimations of data with respect to undocumented migrants and the complexities associated with the consideration of specific destinations for 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 9.6: *The data recorded by the UNHCR [169] in respect of the number of Syrian refugees within each of the neighbouring countries of Syria 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
Proportion (as a percentage)	42.1	21.0	12.0	22.9	1.9

Simulation experiments were performed in an attempt to fit the model concept demonstrator's 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 with the data set, which was recorded at 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 destination country, as well as the neighbouring countries' initial openness index values towards these individuals. A list of the experiments performed and their associated parameter values are given in Table 9.7. The simulated output for these experiments with respect to the number of refugees per neighbouring country of Syria are noted in Table 9.8. Furthermore, the ratio of refugees and undocumented migrants as per neighbouring country and per experiment are illustrated in Figure 9.4.

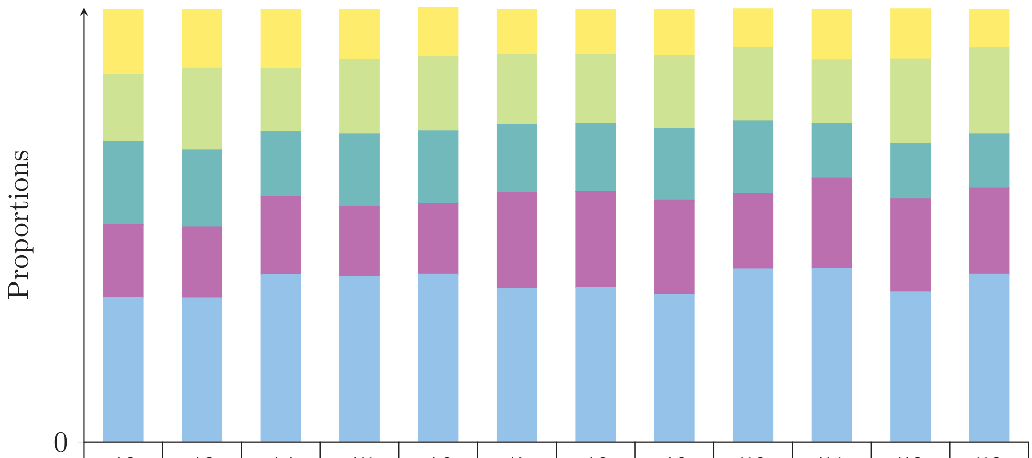
Table 9.7: *Parameter values for the sensitivity analysis experiment pertaining to choosing a destination among neighbouring countries of Syria.*

Experiment	42	43	44	45	46	47	48	49	50	51	52	53
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 Index	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 indices 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

When comparing the experimental outputs with the ratios recorded by the UNHCR [169], shown in Table 9.6, a noticeable difference is visible in respect of 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 weights

Table 9.8: *The model concept demonstrator output (in millions) for the experiments performed in respect of choosing a destination among neighbouring countries of Syria.*

Experiment	42	43	44	45	46	47	48	49	50	51	52	53
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



	42	43	44	45	46	47	48	49	50	51	52	53
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

Proportions of refugees and undocumented migrants per experiment

Figure 9.4: *The model concept demonstrator output (ratios) of the experiments performed in respect of choosing a destination among neighbouring countries of Syria.*

of criteria did not have as significant an effect as changing the initial openness index values. The openness indices were therefore adjusted for each iterative experiment until the output was deemed sensible in its correlation with the existing data. Experiment 53 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 53 in respect of the number of forcibly displaced people per movement type and per destination is given in Appendix C.

It is concluded that by implementing an iterative approach towards generating output which correlates with recorded data, the model concept demonstrator may be effectively calibrated for any given scenario. The user may then decide to either implement these parameters in an attempt to infer future outcomes with respect to future forced migration under similar conditions, or adjust the parameters slightly in order to compare an expected simulated output with an actual simulated output, allowing for further insights to be gained *via* what-if analyses.

9.6 Phase 5: Documentation of the decision-support analysis tool

Phase 5 of the CoFMMA framework entails the documentation of the model concept demonstrator developed. The CoFMMA framework followed in developing the model concept demonstrator was described in Chapters 7–9. The model concept demonstrator is discussed in this section as a decision-support and analysis tool.

The agent-based model concept demonstrator developed is equipped with a GUI, detailed in §7.3.5, for the purpose of utilising the model concept demonstrator as a decision support and analysis tool. The GUI includes a configuration screen designed to prompt the user for various selections and parameter input values, such as the criteria weights pertaining to refugees and undocumented migrants, the criteria weights concerning IDPs, as well as the initial openness index values of countries neighbouring Syria. A brief explanation of each of the available choices and configurable parameters is given below, as well as the default configuration of the model concept demonstrator 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 weights of these criteria when considering refugees and undocumented migrants

Input: Distance, Openness Index and Popularity.

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

Default values: Distance = 0.55, Openness Index = 0.35 and Popularity = 0.1.

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

Effect of decrease: An agent would be less significantly influenced by the decreased factor when having to select a destination.

Choose the weights of these criteria when considering internally displaced people

Input: Conflict and Population Density.

Explanation: Sets the weights 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 values: Conflict = 0.6 and Population Density = 0.4.

Effect of increase: The increased factor will have a greater effect influencing an agent when selecting a destination.

Effect of decrease: The decreased factor will have a lesser effect influencing an agent when selecting a destination.

Choose the initial openness index 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 values: 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 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 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 with which conflict will spread from conflict-affected 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 with which conflict will deplete from the edges of conflict-affected 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 present.

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

Intensity of Conflict

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

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

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

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

9.7 Chapter summary

Phases 4 and 5 of the CoFMMA framework of Chapter 6 were implemented and discussed in respect of executing and documenting the model concept demonstrator in this chapter. In §9.2–§9.5, the model execution phase was discussed as applied to the model concept demonstrator. Calibration of parameters pertaining to the modelling of conflict was performed in §9.1, and this was followed in §9.2 by the validation of the method according to which conflict was modelled and the manner in which it spreads within the simulated area. In §9.3, validation was performed on the method of agent aggregation employed. Face validation was next performed in consultation with international leaders in the field of forced migration and conflict modelling, as discussed in §9.4. This was followed by a parameter variation analysis in §9.5 aimed at illustrating the flexibility of the model as well as its ability to simulate a desired situation as required by a user. The implementation of the model concept demonstrator as a decision-support and analysis tool was documented in §9.6. Each of these steps were discussed in detail within this chapter in an attempt to validate the model concept demonstrator and increase the credibility of its outputs in view of the restrictions that exist when attempting to validate the model concept demonstrator according to traditional methods.

Part IV

Conclusion

CHAPTER 10

Summary and contributions

Contents

A summary of the contents of this dissertation is provided in §10.1 with reference to the dissertation objectives outlined in Chapter 1. This is followed in §10.2 by an overview of the contributions made in this dissertation.

10.1 Dissertation summary

This dissertation consists of eleven chapters, ten of which are partitioned into four separate parts. In the introductory chapter of the dissertation, Chapter 1, which is not included in any of these parts, brief descriptions of global forced migration, the potential impact of humanitarian crises and the application of simulation models within this context were provided. The research problem considered in the dissertation was presented 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 dissertation organisation.

In Part I, consisting of Chapters 2–5, a review of the existing literature was given within the context of forced migration, computer simulation modelling, human decision making and conceptual framework analysis.

The aim of Chapter 2 was to introduce the reader to the concept of forced migration, in partial fulfilment of Dissertation Objective I(a). The review in this introduction covered various causes of forced displacement and expanded on the typology of forced migration, elucidating the various movement types of forcibly displaced people, in accordance with Dissertation Objective I(b). This was followed by a brief overview of forced displacement throughout history, in further fulfilment of Dissertation Objective I(a). The chapter closed with a discussion on the forced migration currently evident on a global scale, in fulfilment of Dissertation Objective I(c).

A number of different considerations related to computer simulation modelling were reviewed in Chapter 3, in partial fulfilment of Dissertation Objective I(e). This discussion included topics such as the various simulation modelling types, levels of model abstraction, various simulation modelling paradigms, generic steps to be followed when carrying out a simulation study, as well as validation and verification procedures pertaining to simulation modelling. In further fulfilment of Dissertation Objective I(e), agent-based simulation modelling was discussed in §3.2 with particular reference to the various advantages and disadvantages of this simulation paradigm and the various components typically comprising such a model. Furthermore, specific existing examples of the application of ABM were discussed in §3.3 with respect to the modelling of forced displacement.

Chapter 4 was an introduction to the field of decision making, in accordance with Dissertation Objective I(f). The area of MCDM was elaborated upon. A classification of MCDM methods was reviewed, and an overview of the available multi-criteria aggregation procedures was provided. In §4.3, the focus of the discussion shifted to the modelling of human decision making. Various methods previously employed to model decision making of humans were described.

The design process of a generic model framework was considered in Chapter 5, in fulfilment of Dissertation Objective I(g). A brief overview of the two dominant research paradigms, as well as of the notion of grounded theory, was given. This led to a discussion on the difference between conceptual and theoretical frameworks. The chapter closed with an overview of the development process of a conceptual model framework and various methodologies employable, as well the processes of verification and validation in the context of such a framework.

Part II was dedicated to the generic framework proposed in this dissertation for assisting modellers in developing agent-based simulation models of forced migration, in accordance with Dissertation Objective II, which is the primary contribution of this dissertation. In Chapter 6, the CoFMMA framework was introduced, commencing in §6.1 with a discussion on the process undertaken in developing this generic framework. The various phases within the CoFMMA framework were elaborated upon in §6.3–6.7. The aim of the framework is to facilitate the process of model development with the objective of understanding the various determinant factors of forced migration, of generating data in an effort to estimate the number of forcibly displaced people and the number of people per movement type in a given geographic area with sufficient accuracy, of predicting the fluctuations in populations in the affected areas and of presenting an animated output of the model results in a user-friendly manner.

In Part III, comprising Chapters 7–9, a model concept demonstrator was developed based on the framework proposed, in fulfilment of Dissertation Objectives III and IV. The aim of this model concept demonstrator was to demonstrate the usefulness and practicability of the CoFMMA framework in the context of conflict-induced forced migration in Syria.

Chapter 7 contained a description of the implementation of the first three phases of the CoFMMA framework, namely formulation, conceptualisation and model development in the context of the concept demonstrator. The characteristics of this concept demonstrator, with a particular focus on the geographic and time scales, as well as the movement types considered, were discussed in §7.1, as per Phase 1 of the CoFMMA framework. Various considerations and assumptions made during Phase 2 were highlighted in §7.2. The conceptualisation phase involved consideration of the geography, time, data and agents, as well as other factors modelled.

In §7.3, the third phase of the CoFMMA framework, concerned with the model development, was discussed. After providing a background to the simulation software environment utilised, Dissertation Objectives III(a), III(b) and III(c) were achieved in §7.3.2, which contained a detailed description of the manner in which conflict was modelled, including the initiation, spread and depletion of conflict. The manner in which the actions of people were modelled was explained in §7.3.3, in fulfilment of Dissertation Objectives III(d) and III(e), and this explanation included the agent attributes considered, the method of simulation of the agent population, particulars of the agent statechart and the incorporation of a tolerance threshold for an agent. In §7.3.4, a discussion followed on the manner in which human decision making was modelled in accordance with Dissertation Objective III(f). The chapter concluded in §7.3.5 with an explanation of the GUI in which the model concept demonstrator is embedded, allowing for user-specified input.

The discussion on Phase 3 of the CoFMMA framework, as implemented for the concept model demonstrator, continued in Chapter 8, focussing on the verification process in partial fulfilment of Dissertation Objective IV. This included descriptions of the verification process followed in respect of the modelling of conflict, the population of agents considered during the simulation process and the decision making of the agents when confronted with conflict. Verification was achieved by the application of a number of test cases.

In accordance with Phase 4 of the CoFMMA framework, a structured analysis was performed in Chapter 9 in respect of the model concept demonstrator. This included model validation, as well as parameter calibration and variation, in further fulfilment Dissertation Objective IV. The parameters pertaining to the modelling of conflict were calibrated in §9.1 and the method of modelling conflict was further validated in §9.2 by means of a number of experiments. Validation of the level of agent aggregation adopted in the model concept demonstrator was performed in §9.3, and this was followed by the face validation in §9.4 of the model concept demonstrator, which included the research opinions of subject matter experts. A parameter establishment analysis was performed in §9.5, after which the possibility of implementing the model as a decision support tool was investigated in §9.6.

Part IV, the present chapter and the following chapter, is the conclusion of the dissertation. Dissertation Objective V is to be achieved in the following chapter where potential future improvements, as well as sensible follow-up work which may stem from this study, are discussed.

10.2 The contributions of this study

The primary contribution of this dissertation was the generic framework proposed in Part II for assisting in the design and development of agent-based models within the context of conflict-induced forced migration towards understanding the behaviour and actions of people when confronted with conflict and, in light of this, facilitating the modelling of the decision making and movement of forcibly displaced people, such as IDPs, refugees and undocumented migrants. Furthermore, a model concept demonstrator was designed and developed in Part III by employing the proposed framework. The individual contributions of this dissertation are presented in this section.

Contribution 10.1 *The development of a novel framework, called the CoFMMA framework, for facilitating the design and development of agent-based simulation models that employ the aspects of localised decision making to produce emergent large-scale movement patterns of those forcibly displaced.*

The CoFMMA framework comprises five phases which encapsulate the components required when developing agent-based models of forcibly displaced migrants. The framework applies to models that pertain not only to information typically recorded by available data sources, but also recognises various movement types. Both Aksel [4] and Groen [66] emphasised that, aside from refugees and asylum-seekers, models of IDP and undocumented migrant movement are important, because there are at present no adequate data sources of this movement. In light of this, the proposed framework has the potential to contribute to research in this respect by providing the possibility of modelling conflict-induced forced migration and generating inference data for further investigation.

Contribution 10.2 *The formal, comprehensive confluence of research opinions and insights, in* *collaboration with subject matter experts, pertaining to the modelling of conflict-induced forced migration.*

A significant effort was made to perform 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 [55] from the Old Dominion University, Groen [66] from Brunell University, Lemos [103] from the University of Agder, Shomary [149] from Stockholm University, Smith [154] from the University of Sussex, and Stewart [160] 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 presently appear in the literature. This process facilitated a collated understanding of social concepts, such as the human characteristics influencing decision making of forcibly displace people, along with a practical understanding of ABM and the various components in the modelling of forced migration. The various opinions and insights were synthesised and formally documented to facilitate future development of models of conflict-induced forced migration to the most realistic degree of accuracy presently possible.

Contribution 10.3 *The design of a user-friendly, agent-based simulation model concept demonstrator of the initiation, spread and depletion of conflict in Syria.*

An agent-based model which illustrates the value of the CoFMMA framework in the modelling of conflict was developed in the AnyLogic Simulation Software Suite [9] as a model concept demonstrator. The initiation of conflict, as well as the spread and depletion of conflict, was modelled according to the description in §7.3.2. The demonstrator's GUI accommodates the manual initiation of conflict, as well as the use of data input values for this purpose, while taking into account the conflict intensity, its probability of spread and its probability of depletion. The model concept demonstrator further allows for an animated model 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 10.4 *A simulation model concept demonstrator encapsulating decision making of agents pertaining to the classification of movement types and destinations in a novel manner.*

The aforementioned model concept demonstrator incorporates notable aspects in the decision-making process of forcibly displaced people. It allows for the calculation of a person's ability to withstand conflict, based on personal characteristics, as described in §7.3.3, as well as a quantification of the decision making of a person in selecting a movement type, as explained in §7.3.4. The decision-making process modelled also accounts for the selection of a destination based on the movement type adopted by a person, his or her characteristics, and other external factors. The GUI accommodates user inputs with respect to the weights of the various decision-making criteria which influence the destination selection process. The modelling of people and their decision making is endorsed by research in social sciences, as well as by the opinions and knowledge gathered from the aforementioned subject matter experts.

Contribution 10.5 *The design and implementation of a decision-support tool which provides information on the influx of refugees and undocumented migrants into regions neighbouring a selected region.*

The GUI of the model concept demonstrator, described in §7.3.5, may be utilised as a decision-support tool. As mentioned above, the interface accommodates user input with respect to the weights of decision-making criteria. It also allows for user input with respect to the initial openness of neighbouring countries towards refugees and undocumented migrants. The decision making of agents pertaining to the selection of a destination was described in detail in §7.3.4. The decision-support tool therefore facilitates the generation of model output which provides graphical data pertaining to the number of refugees and undocumented migrants who flee to identified regions neighbouring a selected region. This has the potential to assist governments and humanitarian support organisations in preparing for the influx of people requiring aid.

Contribution 10.6 *A novel contribution to the global research effort in forced migration simulation modelling.*

During the development of the agent-based modelling framework and subsequent model concept demonstrator presented in this dissertation, regular contact was made with subject matter experts in order to gain insight and knowledge about the field of forced migration so as 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. The study provided viable context to the state of computer simulation-based research with respect to its application in the field of forced migration.

In view of the outcomes of this study, the author was invited by Frydenlund [55] to partake as a member in an expert panel discussion on the theme of "Changing durable solutions" at the 17^{th} *International Association for the Study of Forced Migration* conference hosted by the University of Macedonia in Thessaloniki, Greece in July 2018. The Emerging Scholars and Practitioners on Migration Issues Network also invited the author to partake in a roundtable discussion on "Methodological challenges in forced migration research," which occurred at the same conference. Furthermore, the author was invited to participate in a study on the analysis of mobile telephone data and twitter sentiment in evaluating the integration of refugees in hosting communities, which was published in 2019 by Springer in a book titled *Guide to mobile data analytics in refugee scenarios*. 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 with respect to improving and extending the framework developed in this dissertation.

Contribution 10.7 *The recommendation of sensible avenues for follow-up work which may stem from this study.*

The CoFMMA framework proposed in this dissertation allows for various further extensions to be performed. Recommendations of specific future endeavours are made in the chapter that follows. These recommendations include further novel areas in the area of intersection between the research fields of computer simulation and forced migration which may be investigated in future.

CHAPTER 11

Future Work

Contents

The final chapter of this dissertation contains a number of suggestions for possible future work stemming from this study.

11.1 Extension recommendations for the CoFMMA framework

Plausible extensions to the CoFMMA framework proposed in this dissertation may be pursued in order to increase its versatility and level of realism. The following recommendations are made in respect of extensions to the current CoFMMA framework for possible future research.

Proposal 11.1 *The inclusion of non-conflict factors leading to forced migration.*

The CoFMMA framework is specific to situations where forced migration is induced by conflict. An extension to the framework may include allowing for the modelling of forced migration caused by non-conflict situations, such as human systems failure, climate migration or natural disasters. The concept of *climate refugees*, in particular, is becoming more popular in the literature, as climate change is expected to become a more prominent cause of forced migration in future [18]. This extension to the framework could allow for versatile, forced migration modelling according to an agent-based approach framework which is applicable in the general field of forced migration.

Proposal 11.2 *The development of generic agent-based models which may form a basis for modellers to further develop models based on particular situations.*

The current CoFMMA framework allows for the design and development of dynamic models aimed at simulating conflict-induced forced migration. A number of general sub-approaches may exist according to which standardised generic sub-models can be developed for different situations, such as modelling of conflict in developing countries where more than two parties are involved. The aim would be for several generic sub-models to be developed for simulating conflict and decision making with respect to various movement types. A modeller would then be able to configure a model by integrating a selection of the generic sub-models deemed to be relevant to a particular situation under consideration, followed by further model refinement. This would assist modellers to focus on designing and developing further intricacies underlying the real-word system and, in the process, developing more encompassing models.

Proposal 11.3 *The utilisation of machine learning and data mining for generating model input.*

The application of machine learning and data mining techniques is rapidly gaining traction due to the availability of data and the accessibility of data analytical tools. The CoFMMA framework supports the modeller in gathering qualitative and quantitative data pertaining to a particular conflict situation under investigation in order to conceptualise and develop an appropriate model, as well as performing experiments to analyse different scenarios. An extension to the current framework may include the use of machine learning and data mining tools in gathering input data for models developed.

Various academics, such as Frydenlund *et al.* [56], Groen [66], Martin and Singh [114], and Ruhil [132] have advocated the use of data analytic tools in forced migration research. An example of a recent data mining initiative is the establishment of a Joint Data Center on Forced Displacement which builds on previous work from the UNHCR and the World Bank with the aim of providing a standardised system for analysing population and socio-economic data in support of forced migration studies [132]. The inclusion of machine learning and data mining as an extension to the current CoFMMA framework may allow for more comprehensive data to be used as model input. This has the potential of increasing the level of accuracy with which a model may represent the underlying real-world system.

11.2 Application of the CoFMMA framework for generating recommended courses of action

Proposal 11.4 *The application of socio-economic analyses in respect of humanitarian logistical and disaster relief.*

The framework proposed in this dissertation was developed from academic knowledge as well as practical insights and the opinions of experts, although the author did not have the opportunity to collaborate with a humanitarian organisation. Peres *et al.* [130] studied trends in humanitarian logistics and disaster relief research, and concluded that close collaboration between theory and practice is required, 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.

The framework allows for the development of agent-based models with a graphically recorded output to provide users with model-generated data pertaining to the number of people forcibly displaced exhibiting different movement types and various destinations. A useful application of the framework could be the facilitation of socio-economic analyses based on the model output by investigating certain challenges in humanitarian logistics and disaster relief, such as resource allocations or the identification of locations at which to station refugee camps in an attempt to better cater for the anticipated number of forcibly displaced individuals who would require assistance.

Proposal 11.5 *The application of early warning tools.*

The international community has placed some focus on the anticipation of crises, for actions to be taken swiftly and preventative measures to be set in place [176]. Early warning tools allow for the prediction of risk factors which may cause forced migration, such as weather forecasts, civil violence, poor economic environments and foreign impedance [148]. Many researchers in the field of forced migration, such as Lopez-Lucia [108], Martin and Singh [114], Shellman and Stewart [148], and Suleimenova *et al.* [164] are investigating the development of early warning tools. A recent collaboration between the World Bank and the United Nations further aims to incentivise the development of a link between early warning systems and early response actions [176]. According to Suleimenova *et al.* [164], existing early warning models in the context of forced migration are centred around the prediction of risk factors that result in forced migration rather than the movement patterns of those potentially displaced. Martin and Singh [114] have confirmed the urgent need for evidence-based early warning tools.

This serves as an opportunity for an extension to the current CoFMMA framework by including an early warning tool in the development process of models. Such an extension will allow for the development of effective early warning tools in addition to the modelling for forced migration resulting from the predicted disasters. The implementation of such an application has the potential to protect development gains and aid affected communities more effectively, as well as assist policy makers in their decision making as to which strategies to employ.

References

[1] Achilli L, 2015, *Syrian refugees in Jordan: A reality check*, (Unpublished) Technical Report, European University Institute Policy Brief, 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] Alexopoulos A & Seila AF, 1998, *Output data analysis*, pp. 225–272 in Banks J (Ed), *Handbook of simulation*, John Wiley & Sons, New York (NY).

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

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

[8] 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.

[9] AnyLogic Company, 2016, *Multimethod simulation software and solutions*, [Online], [Cited October 2016], Available from `http://www.anylogic.com`.

[10] 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.

[11] 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.

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

[13] Balci O, 1998, *Verification, validation, and testing*, pp. 335–393 in Banks J (Ed), *Handbook of simulation*, John Wiley & Sons, New York (NY).

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

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

[16] 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)**, pp. 5–49.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] Borshchev A, 2013, *The big book of simulation modeling: Multimethod modeling with AnyLogic 6*, AnyLogic North America, Chicago (IL).

[22] 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.

[23] Brück T, Dunker KM, Fergusson NTN, Meysonnat A & Nillesen E, 2018, *Determinants and dynamics of forced migration to Europe: Evidence from a 3-D model of flows and stocks*, IZA Discussion Paper Series, **11834**, pp. 1–31.

[24] Bryman A & Bell E, 2014, *Research methodology: Business and management contexts*, Oxford University Press, Oxford.

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

[26] Carroll JM, Dawson LL & Swatman PA, 1998, *Using case studies to build theory: Structure and rigour*, Proceedings of the 9^{th} Australasian Conference on Information Systems, Sydney, pp. 64–76.

[27] Carroll JM & Swatman PA, 2000, *Structured-case: A methodological framework for building theory in information systems research*, European Journal of Information Systems, **9**, pp. 235–242.

[28] Carter Center, 2017, *Tracking the front lines in Syria*, [Online], [Cited November 2017], Available from https://www.cartercenter.org/syria-conflict-map/.

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

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

[31] 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.

[32] 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.

[33] 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.

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

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

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

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

[38] Cutcliffe JR, 2000, *Methodological issues in grounded theory*, Journal of Advanced Nursing, **31(6)**, pp. 1476–1484.

[39] 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.

[40] 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.

[41] Denzin N, 1978, *The research act: A theoretical introduction to sociological methods*, Praeger, New York (NY).

[42] Dillon SM, 1998, *Descriptive decision making: Comparing theory with practice*, Proceedings of the 33^{rd} ORSNZ Conference, University of Auckland, Auckland, pp. 1–10.

[43] 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.

[44] Dragostinova T, 2016, *Refugees or immigrants? The migration crisis in Europe in historical Perspective*, Origins, **9(4)**, pp. 1–16.

[45] Durbach IN & Stewart TJ, 2011, *An experimental study of the effect of uncertainty representation on decision making*, European Journal of Operational Research, **214(2)**, pp. 380–392.

[46] Dzurec L & Abraham J, 1993, *The nature of inquiry: Linking quantitative and qualitative research*, Advances in Nursing Science, **16**, pp. 73–79.

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

[48] 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, pp. 1–10.

[49] Elliott N & Higgins A, 2012, *Surviving grounded theory research method in an academic world: Proposal writing and theoretical frameworks*, The Grounded Theory Review, **11(2)**, pp. 1–12.

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

[51] Fargues P & Fandrich C, 2012, *The European response to the Syrian refugee crisis — What next?*, (Unpublished) Technical Report, Robert Schuman Centre for Advanced Studies, European University Institute, Florence.

[52] 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).

[53] Filatova T, Verburg PH, Parker DC & Stannard CA, 2013, *Spatial agent-based models for socio-ecological systems: Challenges and prospects*, Environmental Modelling and Software, **45**, pp. 1–7.

[54] Fioretti G, 2013, *Agent-based simulation models in organization science*, Organizational Research Methods, **16(2)**, pp. 227–242.

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

[56] Frydenlund E, Şener ME, Gore R, Boshuijzen-van Burken C, Bozdag E & De Kock C, 2019, *Characterizing the mobile phone use patterns of refugee hosting provinces in Turkey*, pp. 419–434 in Salah AA, Pentland A, Lepri B & Letouzé E (Eds), *Guide to mobile data analytics in refugee scenarios: The 'Data for Refugees Challenge' study*, Springer, Geneva.

[57] 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.

[58] GDELT Project, 2014, *Mapping the Syrian conflict*, [Online], [Cited September 2017], Available from `https://blog.gdeltproject.org/mapping-the-syrian-conflict/`.

[59] 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`.

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

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

[62] Glaser BG & Strauss AL, 1967, *Discovery of grounded theory: Strategies for qualitative research*, Aldine, Chicago (IL).

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

[64] Green WG & Mcginnis SR, 2002, *Thoughts on the higher order taxonomy of disasters*, Notes on the Science of Extreme Situations, **(7)**, pp. 1–6.

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

[66] 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`.

[67] Groen D, 2018, *Development of a multiscale simulation approach for forced migration*, Proceedings of the International Conference on Computational Science, Wuxi, pp. 869–875.

[68] 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.

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

[70] Gulden T, Harrison JF & Crooks AT, 2011, *Modeling cities and displacement through an agent-based spatial interaction model*, Proceedings of the Computational Social Science Society of America Conference, Santa Fe (NM), pp. 1–9.

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

[72] Hailegiorgis A & Crooks AT, 2012, *Agent-based modeling for humanitarian issues: Disease and refugee camps*, Proceedings of the Computational Social Science Society of America Conference, Santa Fe (NM), pp. 1–23.

[73] 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.

[74] 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.

[75] Humanitarian InSight, 2019, *Global humanitarian overview 2019*, [Online], [Cited March 2019], Available from `https://hum-insight.info/`.

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

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

[78] Imenda S, 2014, *Is there a conceptual difference between theoretical and conceptual frameworks?*, Journal of Social Sciences, **38(2)**, pp. 185–195.

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

[80] Internal Displacement Monitoring Centre, 2016, *Global report on internal displacement*, (Unpublished) Technical Report, Norwegian Refugee Council, Geneva.

[81] Internal Displacement Monitoring Centre, 2017, *Global internal displacement database*, [Online], [Cited March 2019], Available from `http://www.internal-displacement.org/database/displacement-data`.

[82] Internal Displacement Monitoring Centre, 2019, *Global picture*, [Online], [Cited March 2019], Available from `http://www.internal-displacement.org/`.

[83] International Labour Organization, 2019, *ILOSTAT*, [Online], [Cited May 2019], Available from `http://www.ilo.org/ilostat/faces/ilostat-home/download?_adf.ctrl-state=lq39dx3h3_4&_afrLoop=188724284597363#!`.

[84] 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.

[85] Jabareen Y, 2009, *Building a conceptual framework: Philosophy, definitions, and procedure*, International Journal of Qualitative Methods, **8(4)**, pp. 49–62.

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

[87] Johnson RB & Onwuegbuzie AJ, 2004, *Mixed methods research: A research paradigm whose time has come*, Educational Researcher, **33(7)**, pp. 14–26.

[88] Johnson RB, Onwuegbuzie AJ & Turner LA, 2007, *Toward a definition of mixed methods research*, Journal of Mixed Methods Research, **1(2)**, pp. 112–133.

[89] 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, Austin (TX), pp. 1778–1786.

[90] Jonson B, Mouamar P, Huber C, Reid K & Koenig DC, 2017, *Syria refugee crisis: Facts you need to know*, [Online], [Cited October 2017], Available from `https://www.worldvision.org/refugees-news-stories/syria-refugee-crisis-war-facts`.

[91] Katsikopoulos KV, Durbach IN & Stewart TJ, 2017, *When should we use simple decision models? A synthesis of various research strands*, Omega, **81**, pp. 17–25.

[92] 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.

[93] 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.

[94] Kleijnen JPC, 1998, *Experimental design for sensitivity analysis, optimisation, and validation of simulation models*, pp. 173–223 in Banks J (Ed), *Handbook of simulation*, John Wiley & Sons, New York (NY).

[95] Klüpfel H, Schreckenberg M & Meyer-König T, 2005, *Models for crowd movement and egress simulation*, Proceedings of the Traffic and Granular Flow '03, Berlin, pp. 357–372.

[96] 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: Refugee Cooperation, **4**, pp. 1–13.

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

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

[99] 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.

[100] Law AM & Kelton WD, 1982, *Simulation modeling & analysis*, 1$^{\text{st}}$ Edition, McGraw-Hill, New York (NY).

[101] Law AM & Kelton WD, 1991, *Simulation modeling & analysis*, 2$^{\text{nd}}$ Edition, McGraw-Hill, Singapore.

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

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

[104] 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*, Proceedings of the 11^{th} European Conference on Multi-Agent Systems, Toulouse, pp. 124–138.

[105] Liehr P & Smith MJ, 1999, *Middle range theory: Spinning research and practice to create knowledge for the new millennium*, Advances in Nursing Science, **21(4)**, pp. 81–91.

[106] Lin L, Carley KM & Cheng SF, 2016, *An agent-based approach to human migration movement*, Proceedings of the 2016 Winter Simulation Conference, Arlington (VA), pp. 3510–3520.

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

[108] Lopez-Lucia E, 2015, *Early warning models for irregular migration*, (Unpublished) Technical Report, Governance, Social Development, Humanitaruan & Conflict, University of Birmingham, Birmingham.

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

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

[111] MacCrimmon KR, 1968, *Decision making among multiple-attribute alternatives: A survey and consolidated approach*, (Unpublished) Technical Report, Advanced Research Projects Agency, Santa Monica (CA).

[112] Macmillan S, Steele J, Austin S, Kirby P & Spence R, 2001, *Development and verification of a generic framework for conceptual design*, Design Studies, **22**, pp. 169–191.

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

[114] Martin SF & Singh L, 2018, *Data analytics and displacement: Using big data to forecast mass movements of people*, pp. 185–206 in Maitland CF (Ed), *Digital lifeline? — ICTs for refugees and displaced persons*, MIT Press, Cambridge (MA).

[115] 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/`.

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

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

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

[119] Mercy Corps, 2017, *Quick facts: What you need to know about the Syria crisis*, [Online], [Cited October 2017], Available from `https://www.mercycorps.org/articles/south-sudan/quick-facts-what-you-need-know-about-south-sudan-crisis`.

[120] Migration policy Institute, 2019, *Migration data hub*, [Online], [Cited May 2019], Available from `https://www.migrationpolicy.org/programs/migration-data-hub`.

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

[122] Moore WH & Shellman SM, 2006, *Refugee or internally displaced person? To where should one flee?*, Comparative Political Studies, **39(5)**, pp. 599–622.

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

[124] 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.

[125] 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.

[126] 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.

[127] Organisation for Economic Cooperation and Development, 2019, *International migration database and labour market outcomes of immigrants*, [Online], [Cited May 2019], Available from `http://www.oecd.org/els/mig/keystat.htm`.

[128] Patton MQ, 2002, *Qualitative research and evaluation methods*, Sage Publications, Thousand Oaks (CA).

[129] 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.

[130] 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 (QC).

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

[132] Petronzio M, 2018, *Big data's potential to help solve the refugee crisis*, [Online], [Cited January 2019], Available from `https://mashable.com/2018/04/24/big-data-refugees/#5wVsUz85fiq9`.

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

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

[135] 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.

[136] Refugee Studies Centre, 2016, *Annual report 2015–2016*, (Unpublished) Technical Report, University of Oxford, Oxford.

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

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

[139] Ritchie J & Spencer L, 1994, *Qualitative data analysis for applied policy research*, pp. 173–194 in Bryman A & Burgess B (Eds), *Analyzing qualitative data*, Routledge, London.

[140] Rizvi S, 2010, *Internal displacement in Afghanistan*, Middle East Institute-Fondation pour la Recherche Stratégique, **25**, pp. 1–5.

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

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

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

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

[145] 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.

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

[147] Shannon R, 1998, *Introduction to the art and science of simulation*, Proceedings of the 1998 Winter Simulation Conference, Washington (DC), pp. 7–14.

[148] Shellman SM & Stewart BM, 2007, *Predicting risk factors associated with forced migration: An early warning model of Haitian flight*, Civil Wars, **9(2)**, pp. 174–199.

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

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

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

[152] Simon HA, 1960, *The new science of management decision*, Prentice Hall, New York (NY).

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

[154] Smith C, 2017, Research Fellow at *School of Global Studies, University of Sussex*, [Personal Communication], Contactable at `c.d.smith@sussex.ac.uk`.

[155] 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, pp. 1–15.

[156] 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.

[157] Srivastava A & Thomson SB, 2009, *Framework analysis: A qualitative methodology for applied policy research*, Journal of Administration and Governance, **4(2)**, pp. 72–79.

[158] Stern PN, Allen L & Moxley P, 1982, *The nurse as a grounded theorist: History, process and uses*, Review Journal of Philosophy and Social Science, **7**, pp. 200–215.

[159] Stewart R, 2004, *Simulation: The practice of model development and use*, John Wiley & Sons, Chichester.

[160] Stewart TJ, 2017, Professor at *Department of Statistical Sciences, University of Cape Town*, [Personal Communication], Contactable at `theodor.stewart@uct.ac.za`.

[161] Steynberg R, 2016, *A framework for identifying the most likely successful underprivileged tertiary bursary applicants*, MEng thesis, Stellenbosch University, Stellenbosch.

[162] Strauss A & Corbin J, 1994, *Grounded theory methodology*, pp. 273–285 in Denzin NK & Lincoln YS (Eds), *The SAGE handbook of qualitative research*, SAGE, Thousand Oaks (CA).

[163] Sturrock DT, 2013, *Tutorial: Tips for successful practice of simulation*, Proceedings of the 2013 Winter Simulation Conference, Washington (DC), pp. 1756–1764.

[164] Suleimenova D, Bell D & Groen D, 2017, *A generalized simulation development approach for predicting refugee destinations*, Scientific Reports, **7**, pp. 1–13.

[165] United Nations Department of Economics and Social Affairs, 2017, *World population prospects: The 2017 revision*, [Online], [Cited July 2017], Available from `https://esa.un.org/unpd/wpp/DataQuery/`.

[166] United Nations Department of Economics and Social Affairs, 2019, *Population division: International migration*, [Online], [Cited May 2019], Available from `http://www.unmigration.org`.

[167] United Nations High Commissioner for Refugees, 1961, *Convention and protocol relating to the status of refugees*, (Unpublished) Technical Report, United Nations, Geneva.

[168] United Nations High Commissioner for Refugees, 2014, *World at war: Global trends 2014 in review*, (Unpublished) Technical Report, United Nations, Geneva.

[169] United Nations High Commissioner for Refugees, 2015, *Global trends: Forced displacement in 2015*, (Unpublished) Technical Report, United Nations, Geneva.

[170] United Nations High Commissioner for Refugees, 2016, *Fact sheet: Global forced displacement*, (Unpublished) Technical Report, United Nations, Geneva.

[171] United Nations High Commissioner for Refugees, 2018, *Global trends: Forced displacement in 2017*, (Unpublished) Technical Report, United Nations, Geneva.

[172] United Nations High Commissioner for Refugees, 2019, *Figures at a glance: Statistical yearbooks*, [Online], [Cited May 2019], Available from `https://www.unhcr.org/figures-at-a-glance.html`.

[173] 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`.

[174] United Nations Office for the Coordination of Humanitarian Affairs, 2016, *Syrian Arab Republic — Humanitarian needs overview*, (Unpublished) Technical Report, United Nations, New York (NY).

[175] United Nations Office for the Coordination of Humanitarian Affairs, 2017, *Syrian Arab Republic — Humanitarian needs overview*, (Unpublished) Technical Report, United Nations, New York (NY).

[176] United Nations Office for the Coordination of Humanitarian Affairs, 2019, *Global humanitarian overview*, (Unpublished) Technical Report, United Nations, New York (NY).

[177] United Nations Office for the Coordination of Humanitarian Affairs, 2019, *The humanitarian data exchange*, [Online], [Cited March 2019], Available from `https://data.humdata.org/`.

[178] van Vuuren BJ, 2016, *An agent-based simulation model of Eldana saccharina Walker*, PhD dissertation, Stellenbosch University, Stellenbosch.

[179] Velasquez M & Hester PT, 2013, *An analysis of multi-criteria decision making methods*, International Journal of Operations Research, **10(2)**, pp. 56–66.

[180] Vincke P, 1992, *Multicriteria decision-aid*, John Wiley & Sons, Chichester.

[181] 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`.

[182] Wang Y & Ruhe G, 2007, *The cognitive process of decision making*, International Journal of Cognitive Informatics and Natural Intelligence, **1(2)**, pp. 73–85.

[183] World Bank, 2017, *World development indicators*, (Unpublished) Technical Report, The World Bank, Washington (DC).

[184] World Bank, 2019, *World Bank open data*, [Online], [Cited March 2019], Available from `https://data.worldbank.org`.

[185] 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.

[186] Zeleny M, 1984, *MCDM: Past decade and future trends: A source book of multiple criteria decision making*, Jai Pr, London.

[187] 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 26^{th} Workshop on Principles of Advanced and Distributed Simulation, Zhangjiajie, pp. 233–242.

[188] 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

The material presented in this appendix should be interpreted in conjunction with the discussion in §9.1. The spread of conflict across Syria at the end of December 2016 is illustrated in Figures A.1–A.25, modelled with different probabilities of conflict spreading and depleting. These 25 sets of results correspond to the 25 scenarios in Table 9.1.

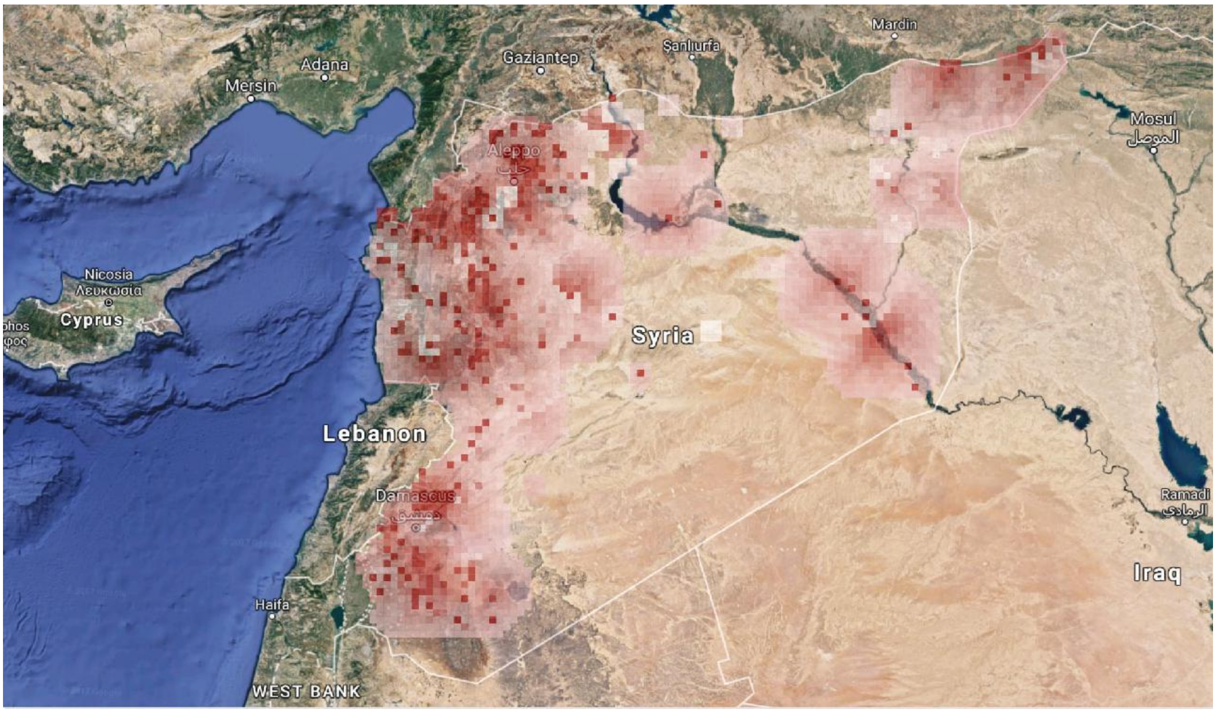


Figure A.1: *Experiment 1: Conflict modelled with a 0.1 probability of spread and a 0.1 probability of depletion.*

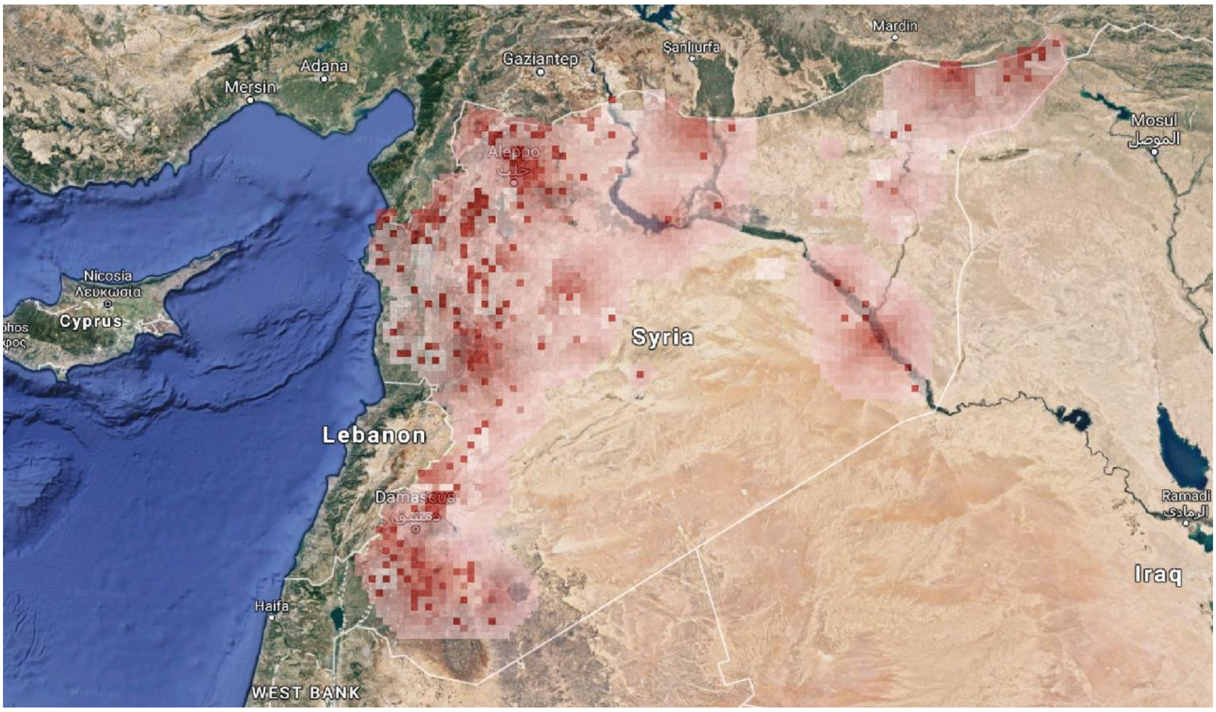


FIGURE A.2: *Experiment 2: Conflict modelled with a 0.3 probability of spread and a 0.1 probability of depletion.*

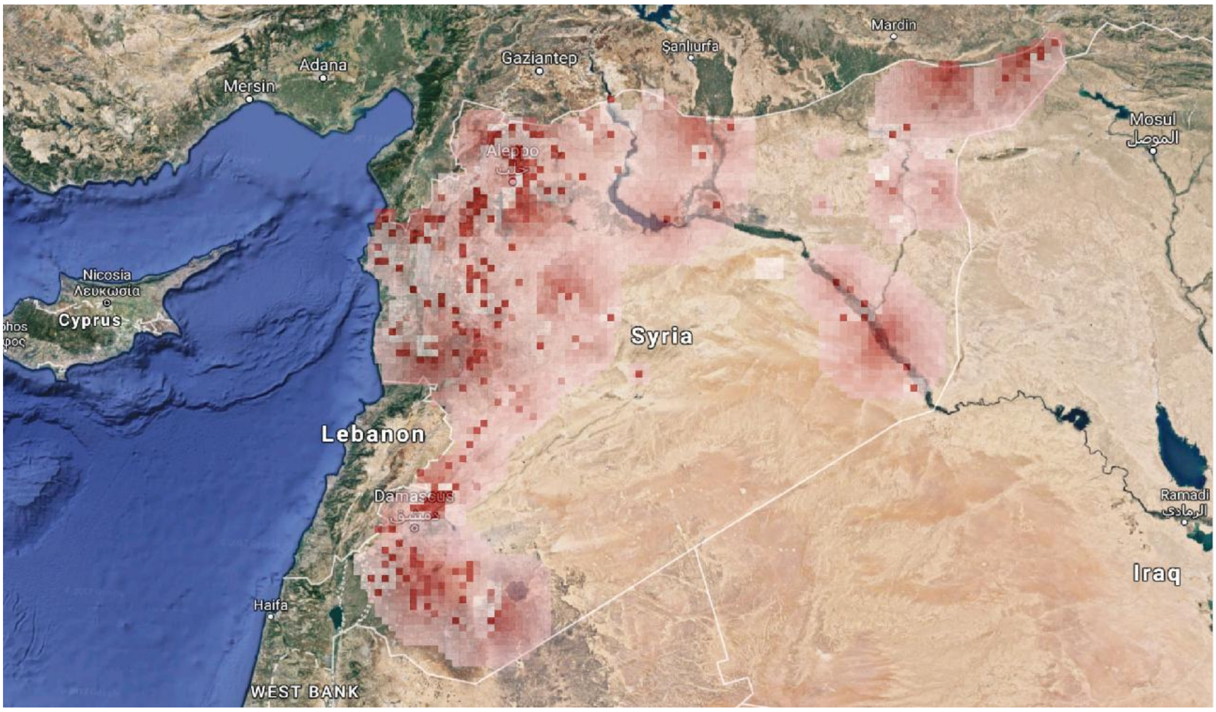


FIGURE A.3: *Experiment 3: Conflict modelled with a 0.5 probability of spread and a 0.1 probability of depletion.*

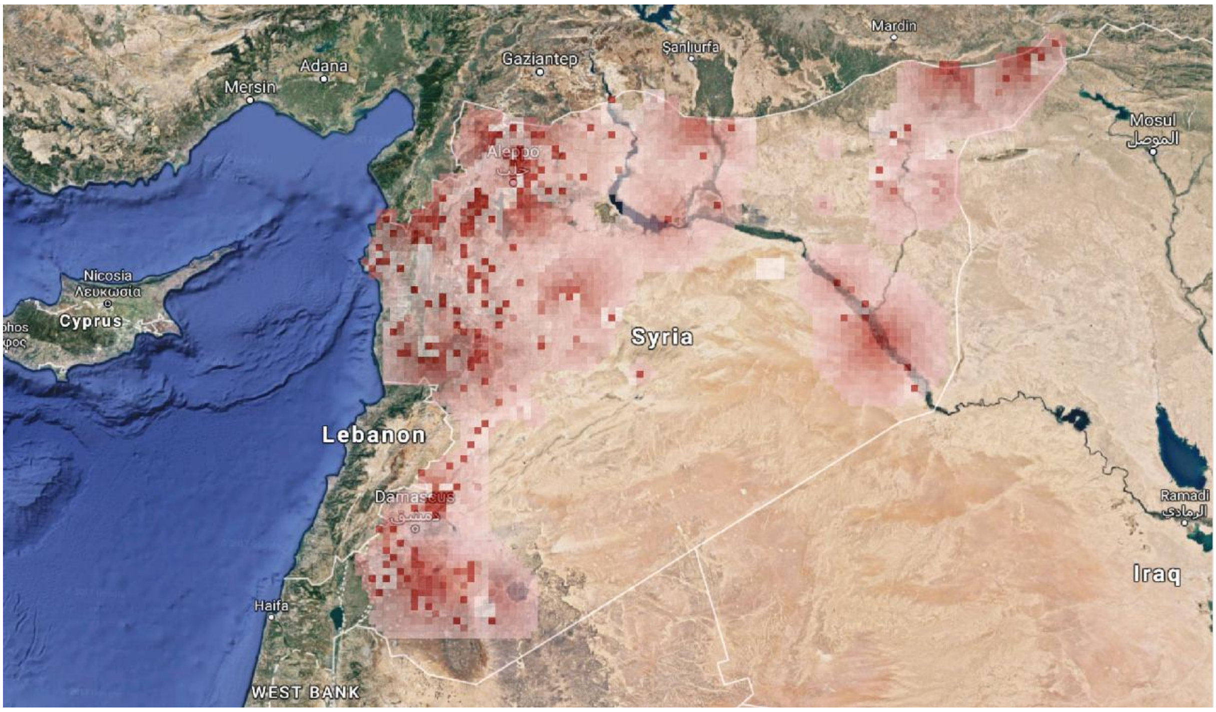


FIGURE A.4: *Experiment 4: Conflict modelled with a 0.7 probability of spread and a 0.1 probability of depletion.*

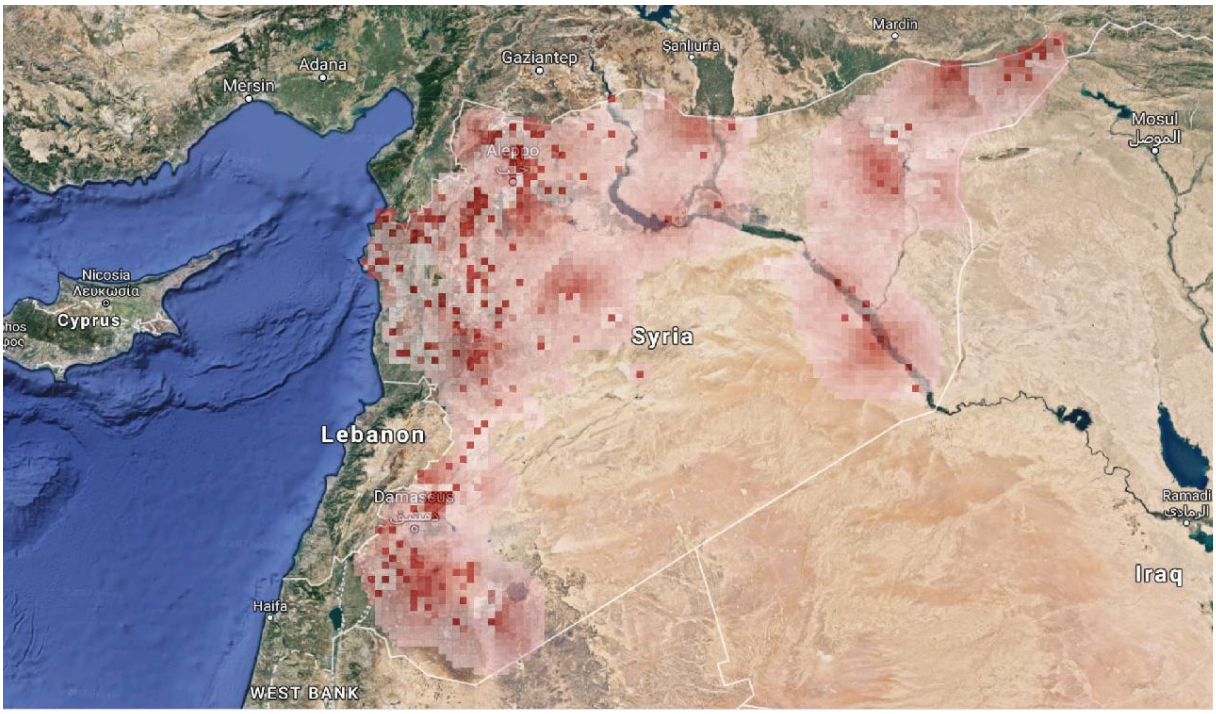


FIGURE A.5: *Experiment 5: Conflict modelled with a 0.9 probability of spread and a 0.1 probability of depletion.*

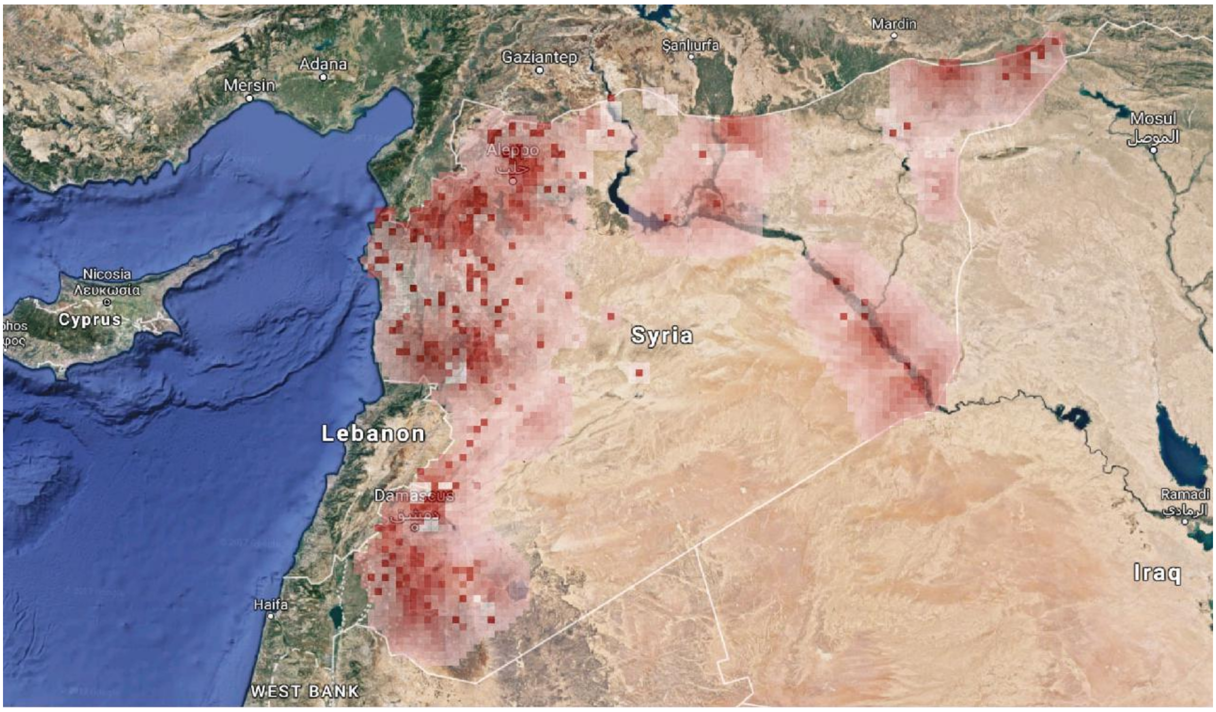


Figure A.6: *Experiment 6: Conflict modelled with a 0.1 probability of spread and a 0.3 probability of depletion.*

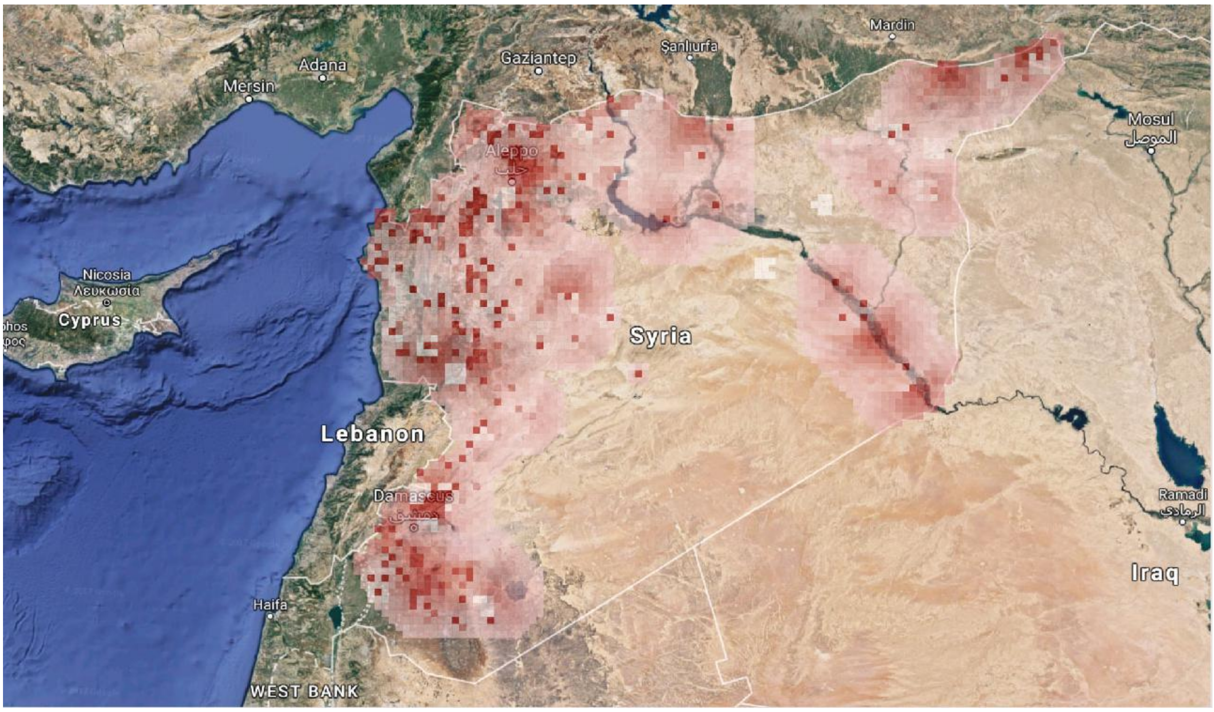


Figure A.7: *Experiment 7: Conflict modelled with a 0.3 probability of spread and a 0.3 probability of depletion.*

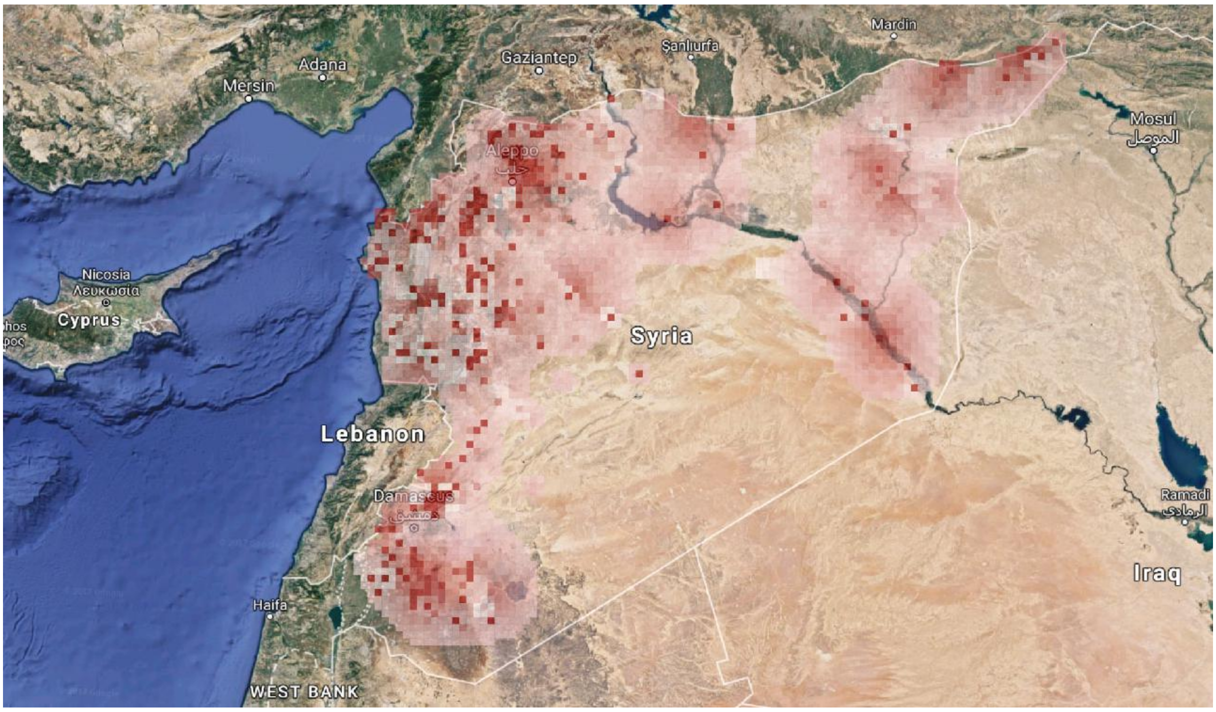


Figure A.8: *Experiment 8: Conflict modelled with a 0.5 probability of spread and a 0.3 probability of depletion.*

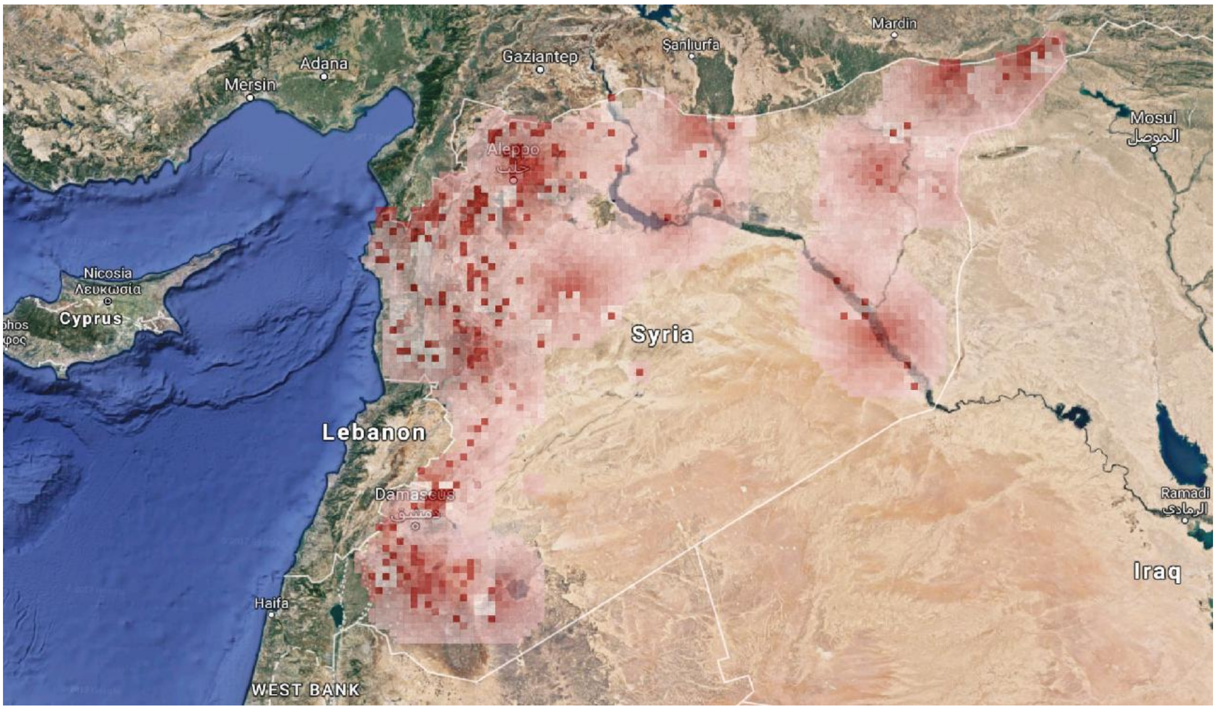


Figure A.9: *Experiment 9: Conflict modelled with a 0.7 probability of spread and a 0.3 probability of depletion.*

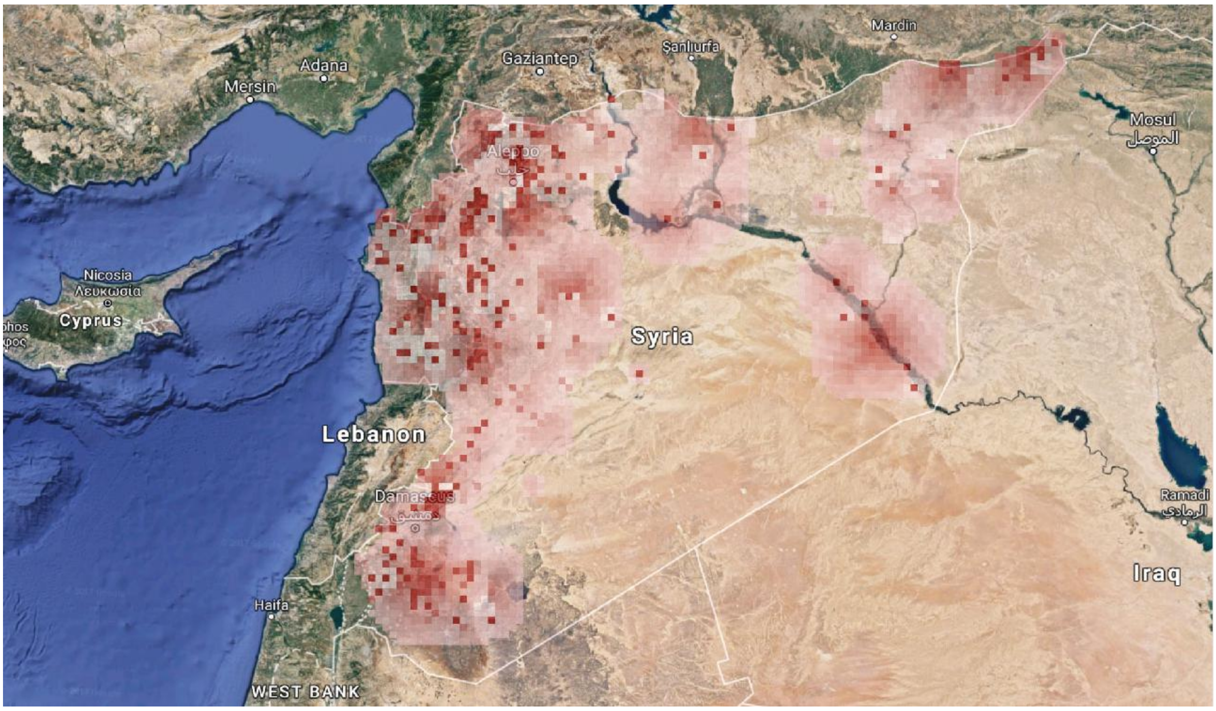


FIGURE A.10: *Experiment 10: Conflict modelled with a 0.9 probability of spread and a 0.3 probability of depletion.*

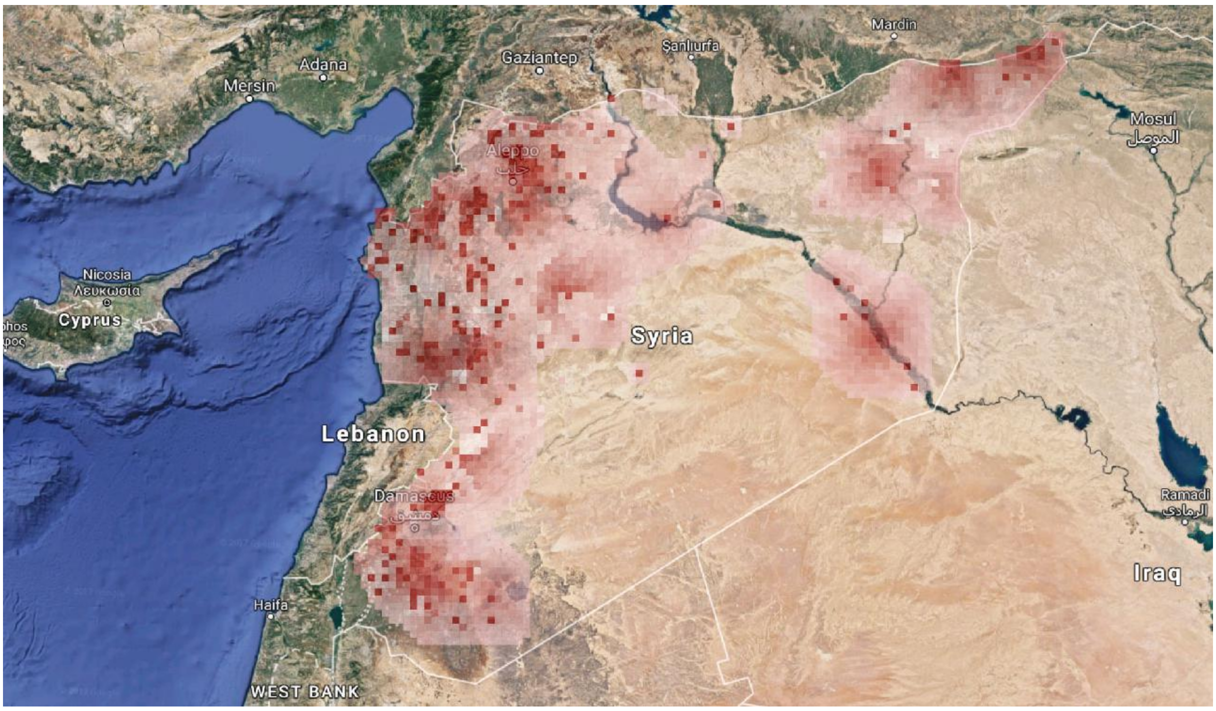


FIGURE A.11: *Experiment 11: Conflict modelled with a 0.1 probability of spread and a 0.5 probability of depletion.*

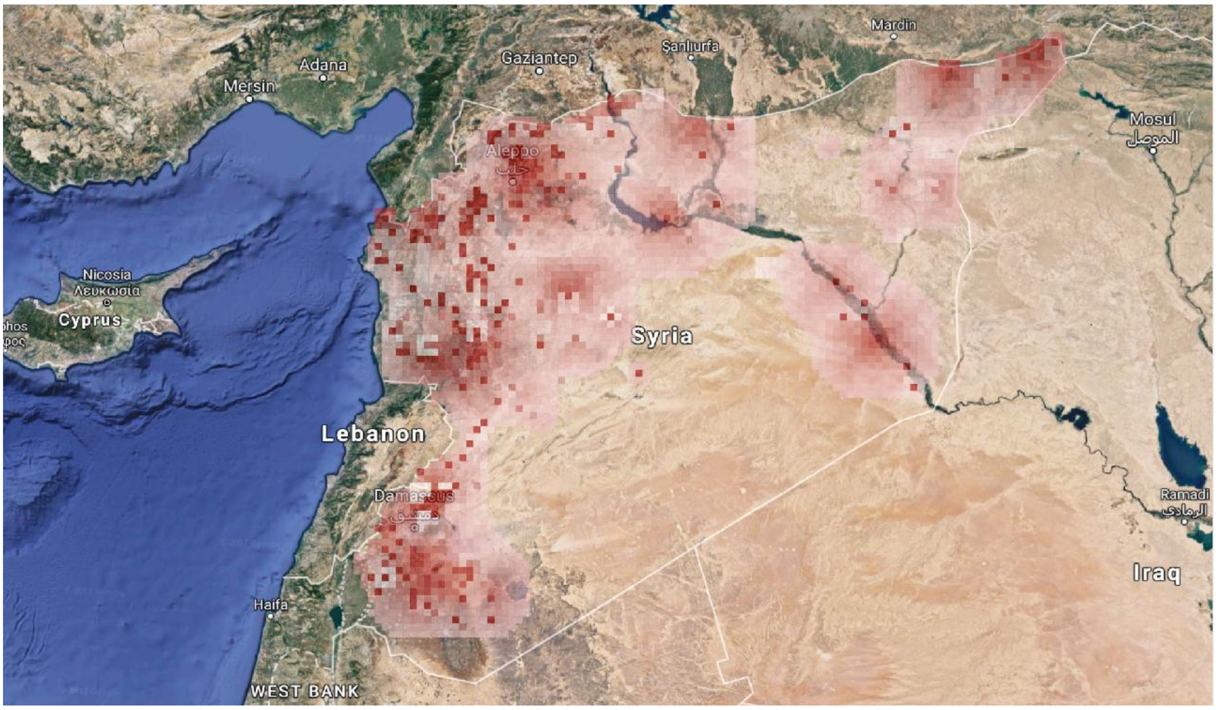


Figure A.12: *Experiment 12: Conflict modelled with a 0.3 probability of spread and a 0.5 probability of depletion.*

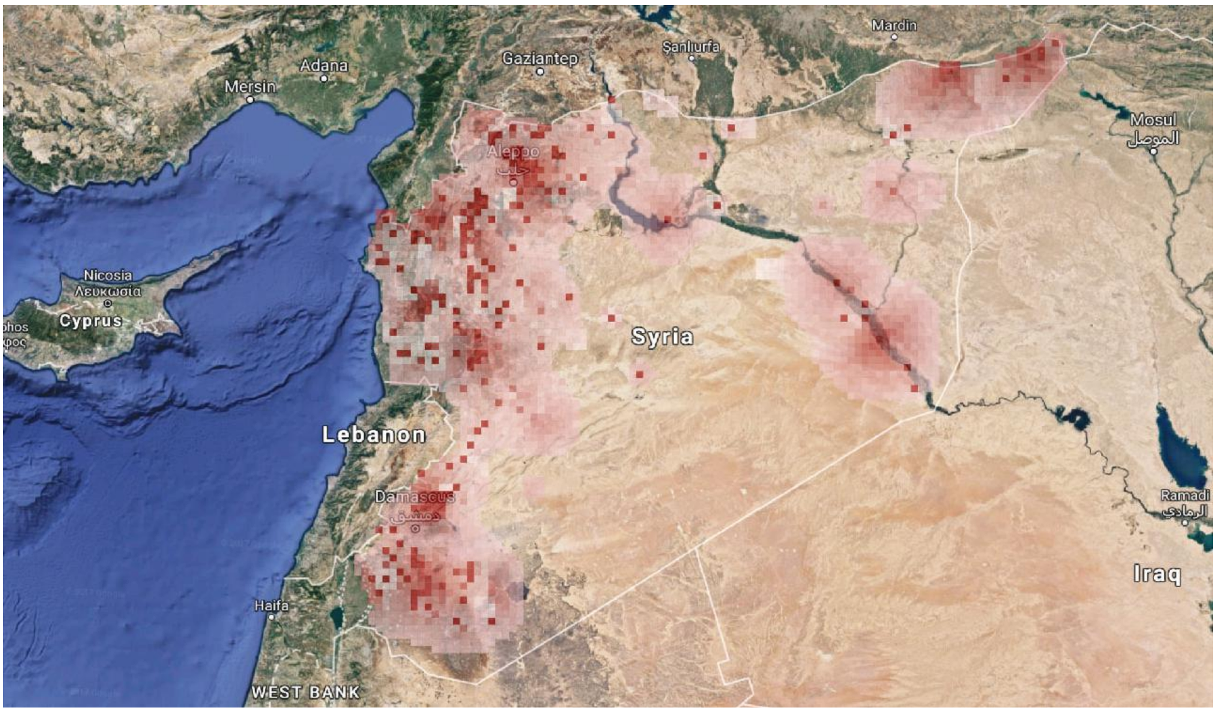


Figure A.13: *Experiment 13: Conflict modelled with a 0.5 probability of spread and a 0.5 probability of depletion.*

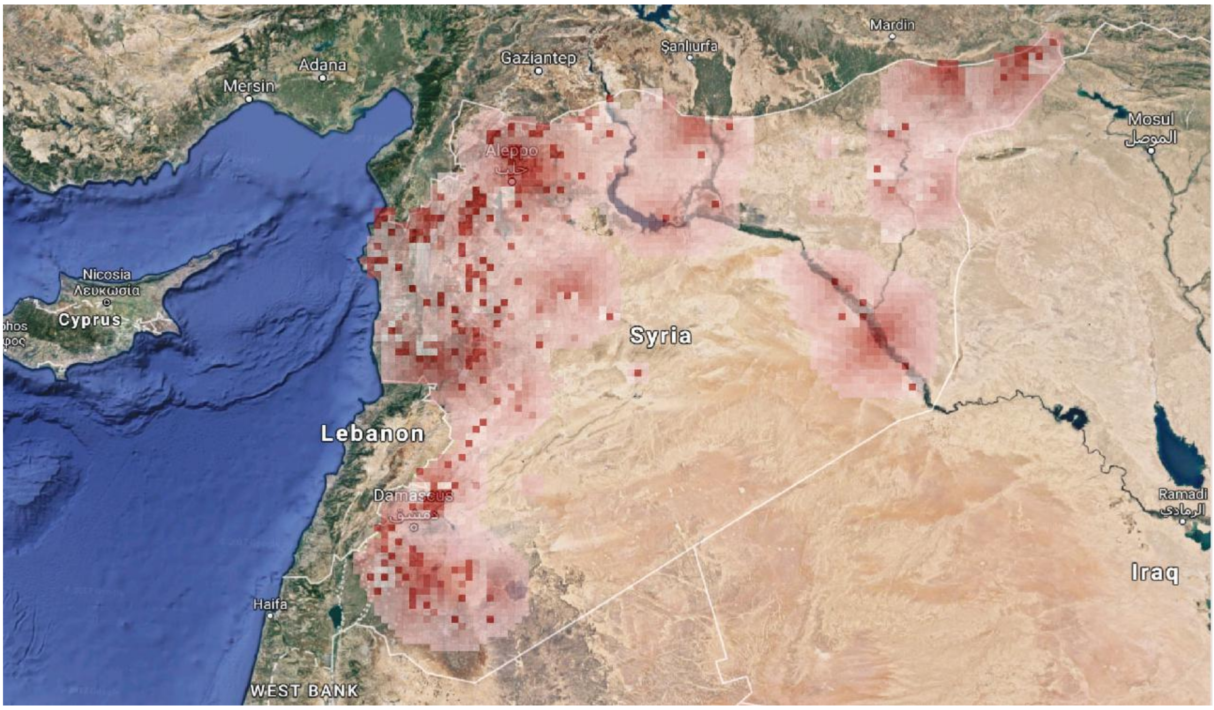


FIGURE A.14: *Experiment 14: Conflict modelled with a 0.7 probability of spread and a 0.5 probability of depletion.*

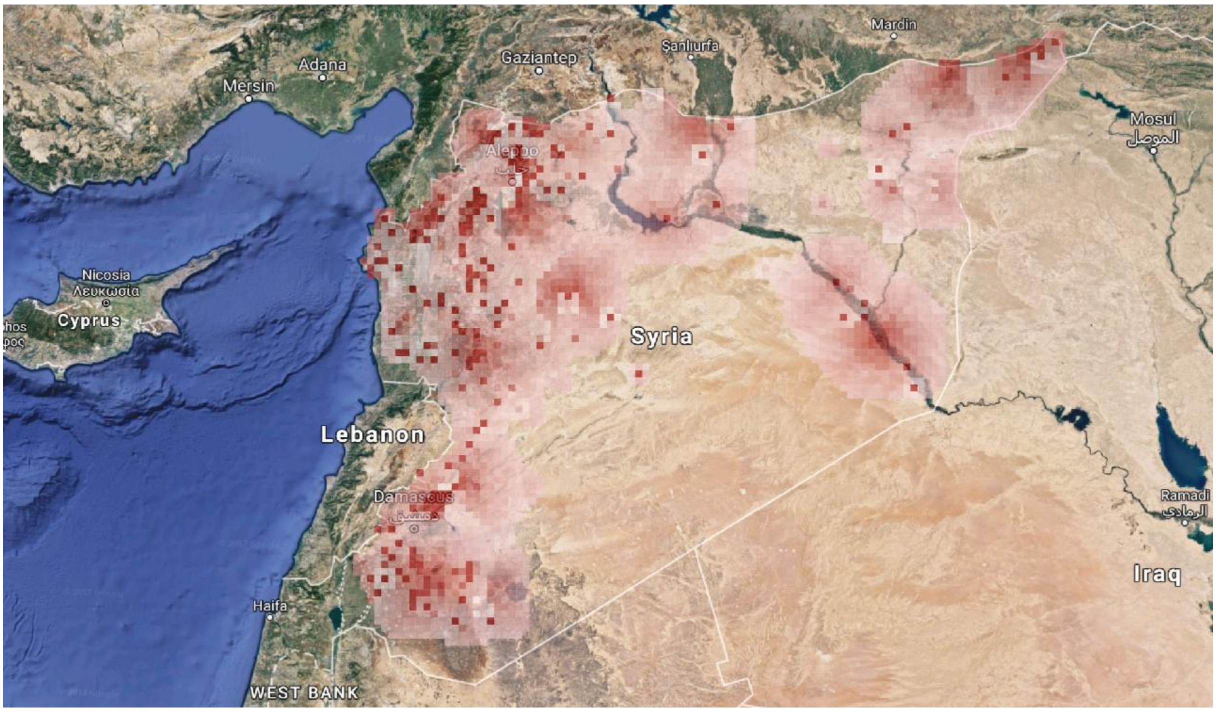


FIGURE A.15: *Experiment 15: Conflict modelled with a 0.9 probability of spread and a 0.5 probability of depletion.*

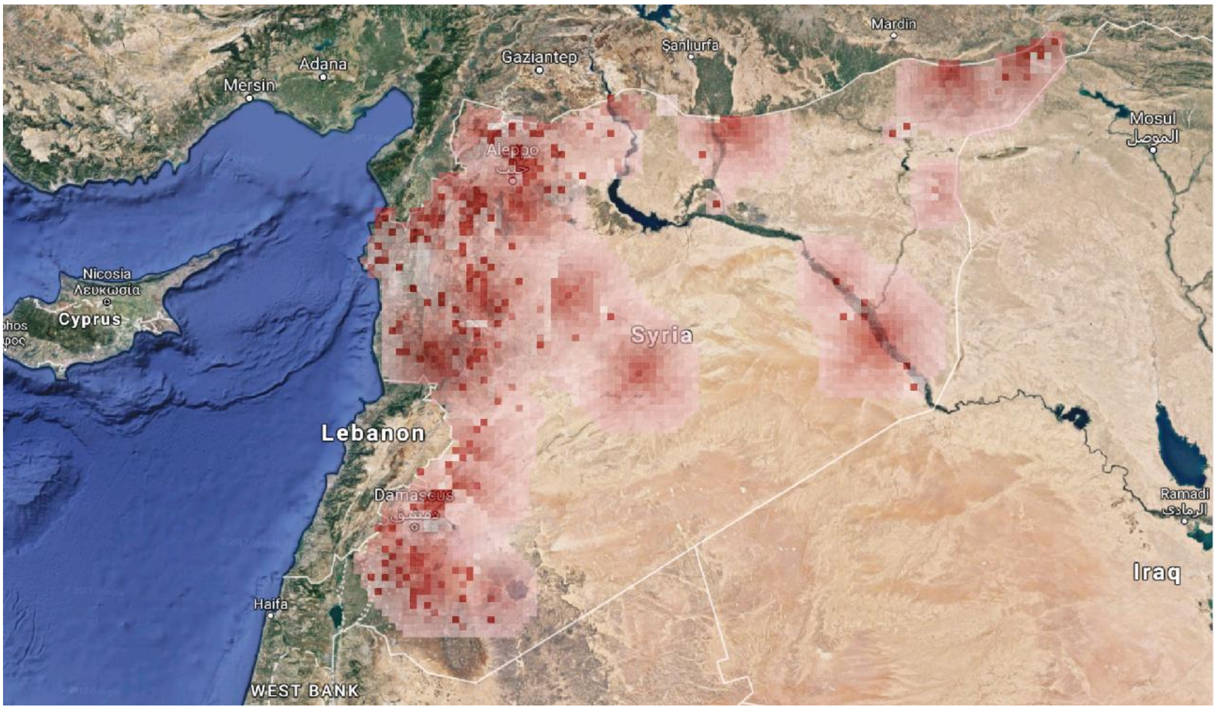


FIGURE A.16: *Experiment 16: Conflict modelled with a 0.1 probability of spread and a 0.7 probability of depletion.*

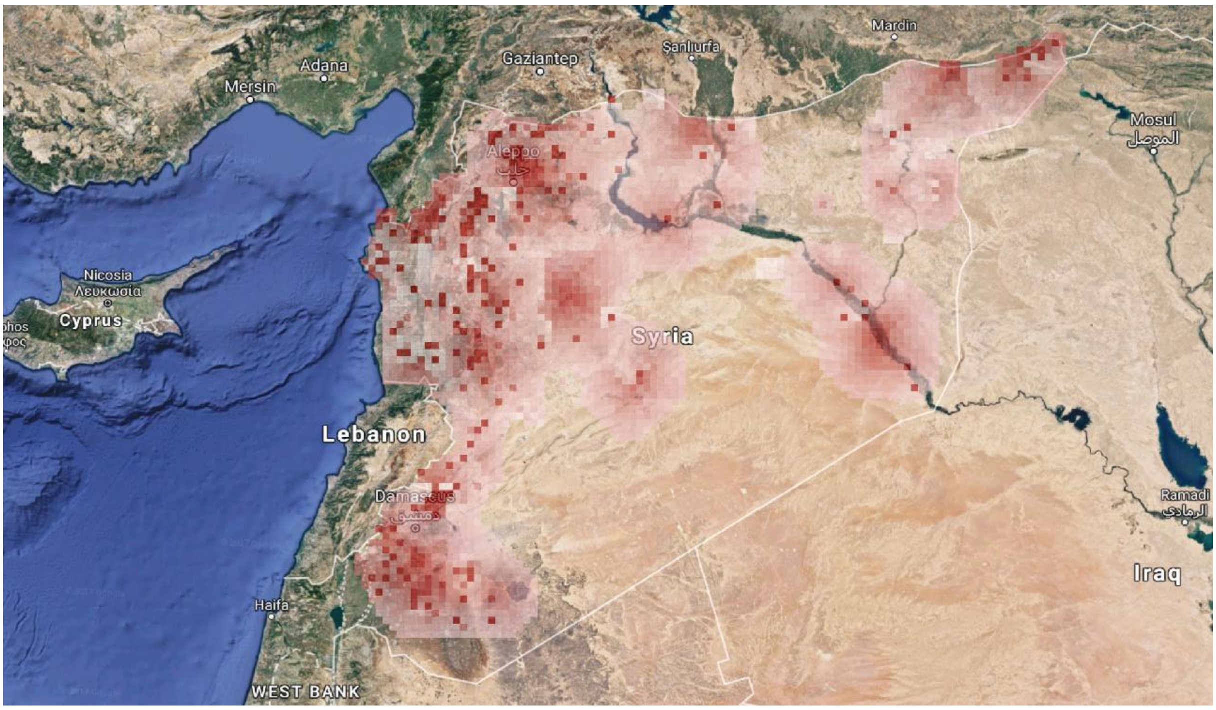


FIGURE A.17: *Experiment 17: Conflict modelled with a 0.3 probability of spread and a 0.7 probability of depletion.*

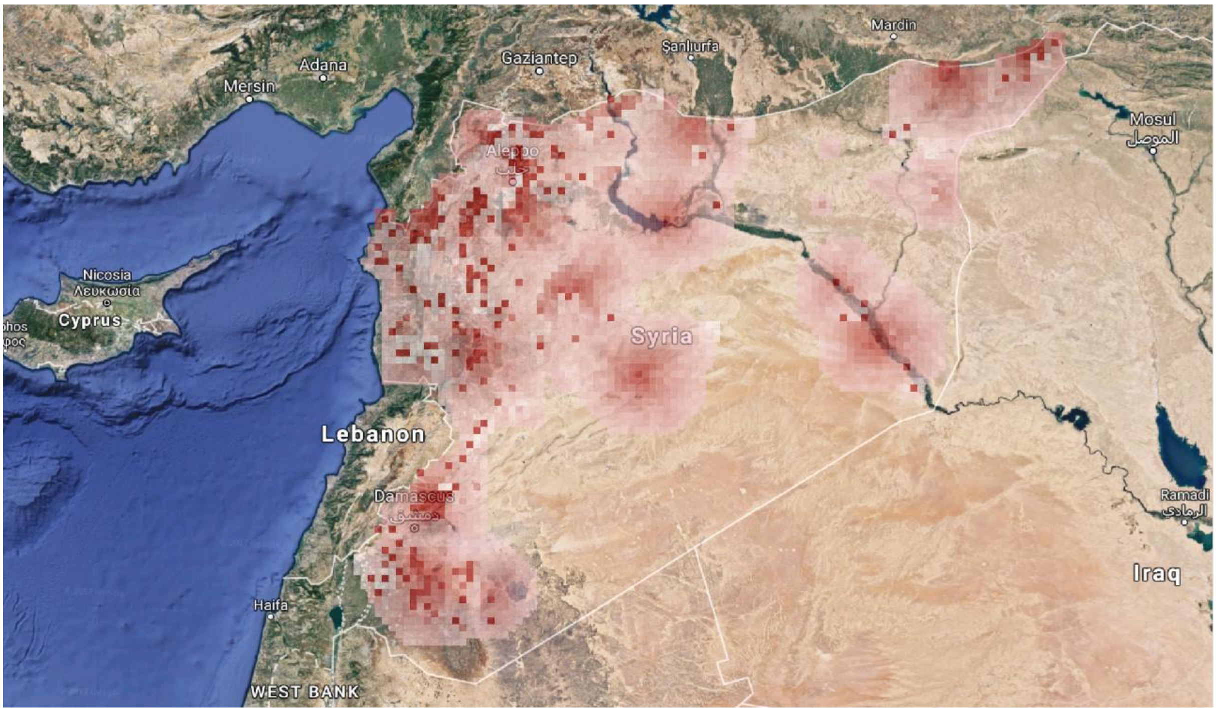


FIGURE A.18: *Experiment 18: Conflict modelled with a 0.5 probability of spread and a 0.7 probability of depletion.*

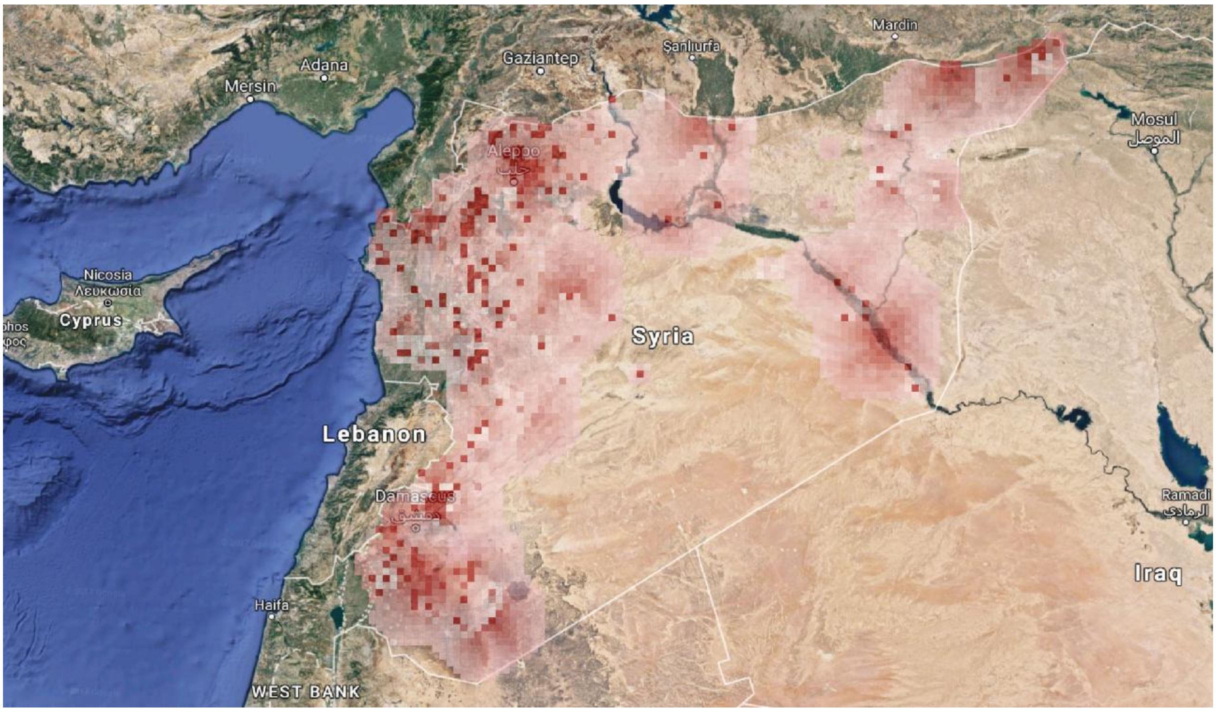


FIGURE A.19: *Experiment 19: Conflict modelled with a 0.7 probability of spread and a 0.7 probability of depletion.*

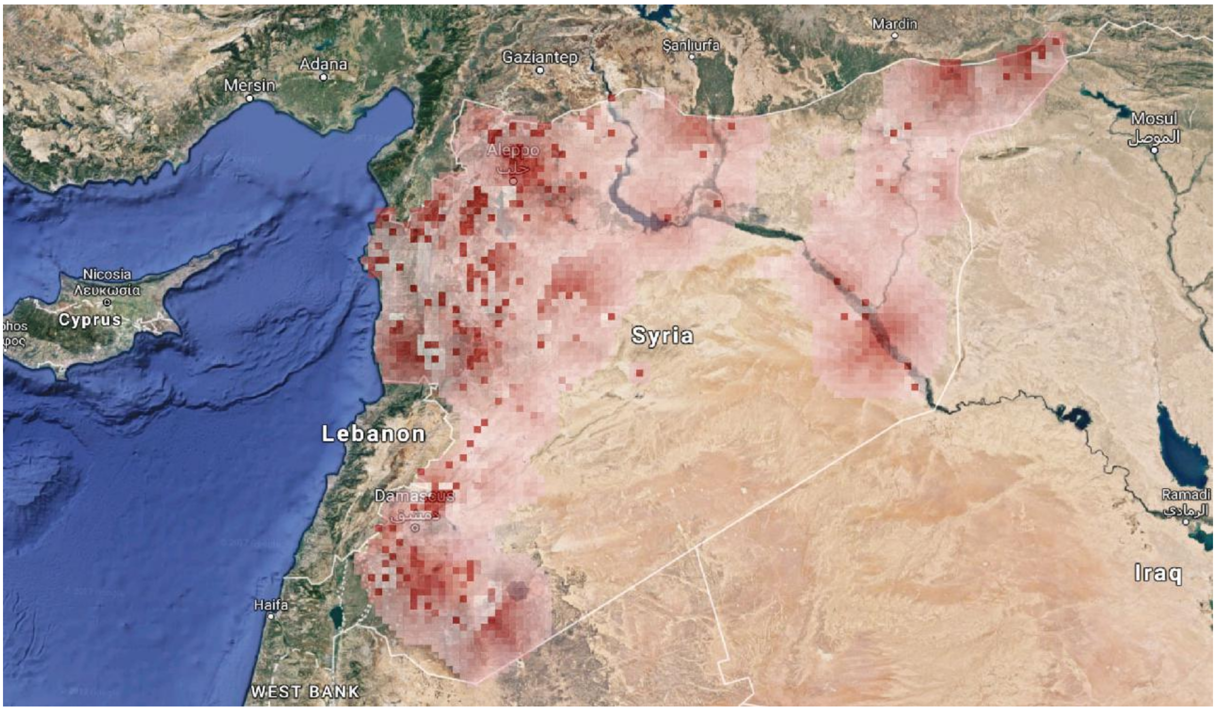


FIGURE A.20: *Experiment 20: Conflict modelled with a 0.9 probability of spread and a 0.7 probability of depletion.*

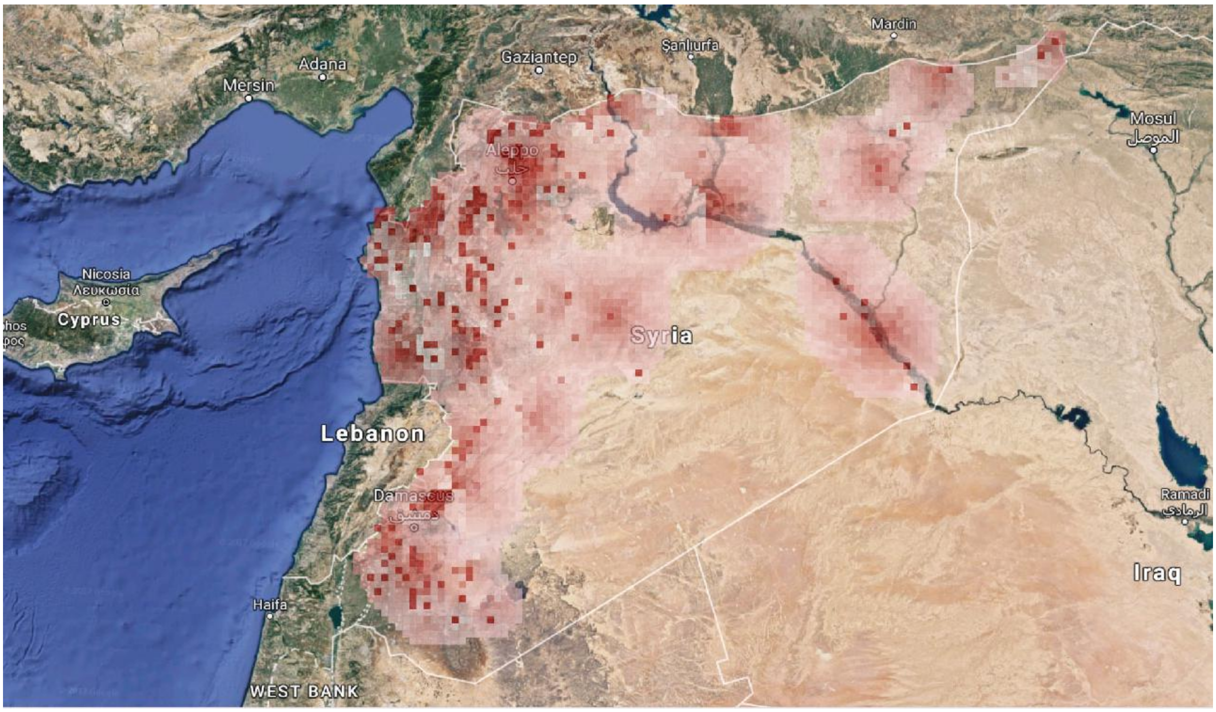


FIGURE A.21: *Experiment 21: Conflict modelled with a 0.1 probability of spread and a 0.9 probability of depletion.*

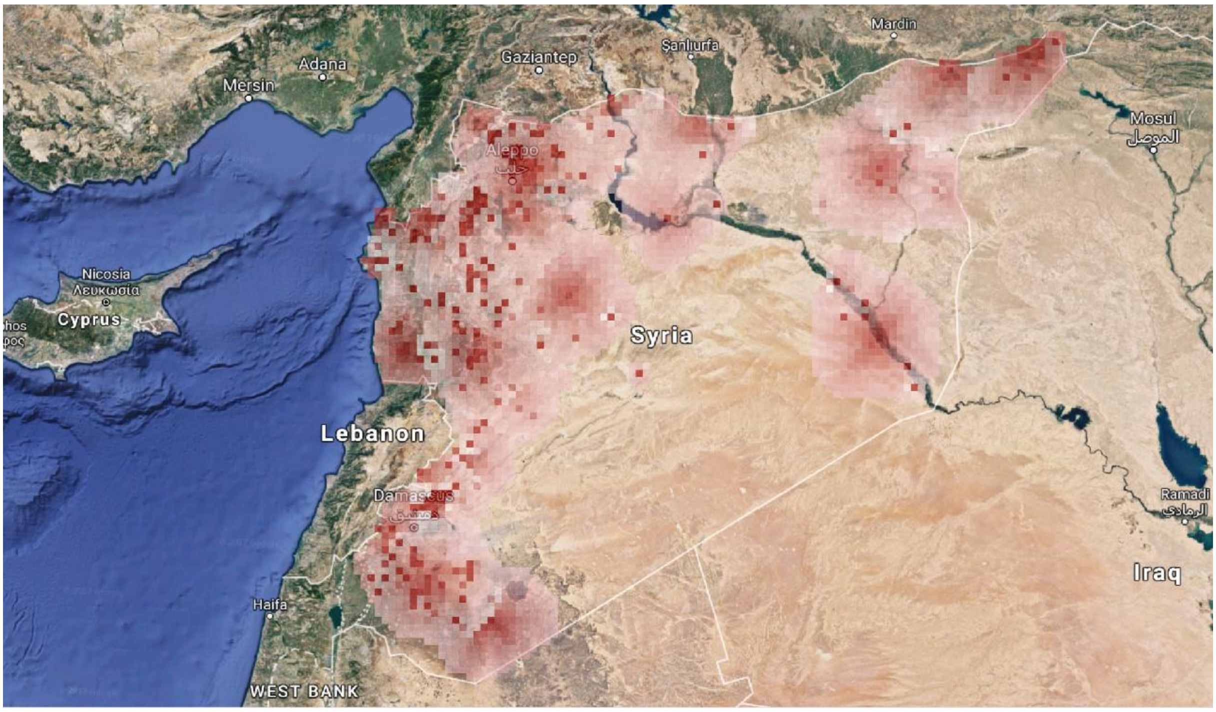


FIGURE A.22: *Experiment 22: Conflict modelled with a 0.3 probability of spread and a 0.9 probability of depletion.*

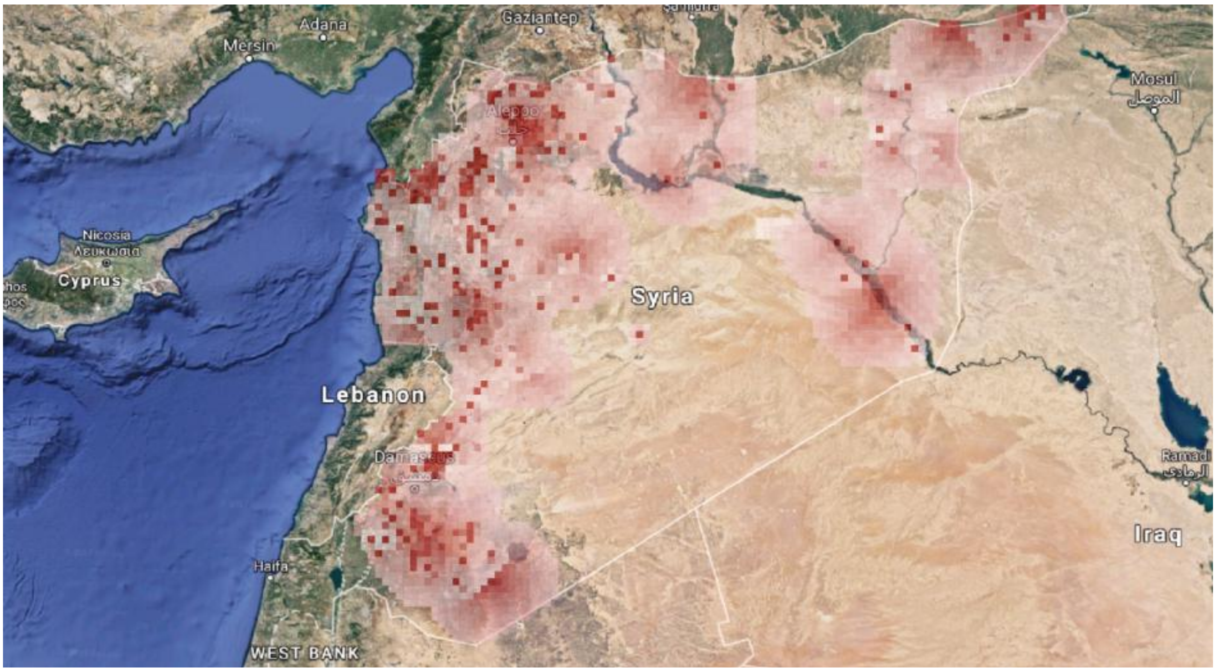


FIGURE A.23: *Experiment 23: Conflict modelled with a 0.5 probability of spread and a 0.9 probability of depletion.*

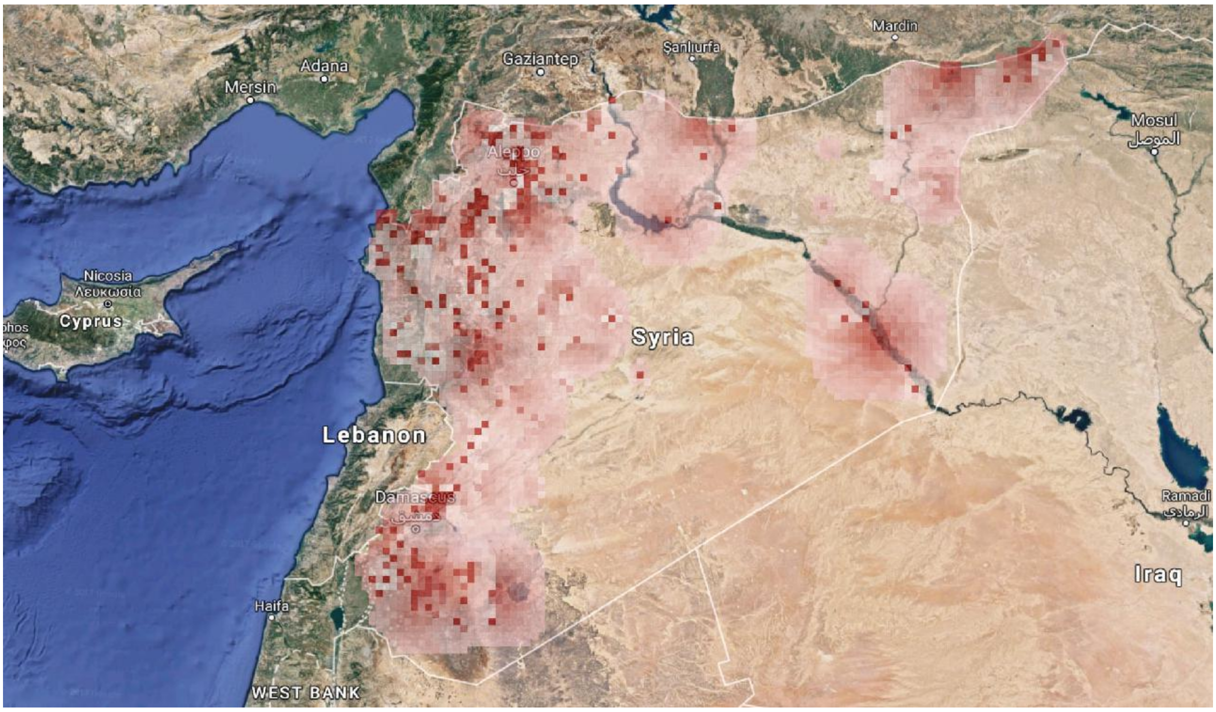


Figure A.24: *Experiment 24: Conflict modelled with a 0.7 probability of spread and a 0.9 probability of depletion.*

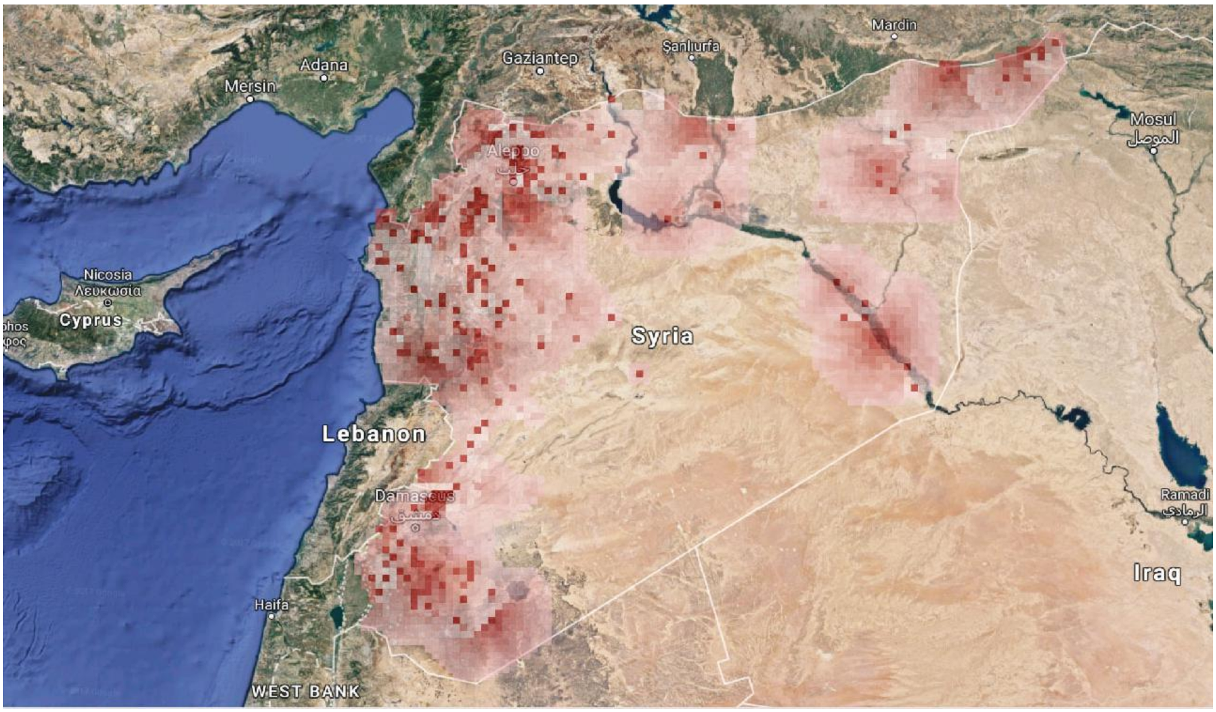


Figure A.25: *Experiment 25: Conflict modelled with a 0.9 probability of spread and a 0.9 probability of depletion.*

APPENDIX B

Probability matrices of experiments

The probability matrices given in this appendix correlate with the experiments performed in §9.5, where E_i refers to the matrix for the i^{th} experiment.

$$\boldsymbol{E}_{35} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \left(\begin{array}{ccc} 0.25 & 0.40 & 0.35 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.10 & 0.25 \\ 0.20 & 0.20 & 0.60 \\ 0.65 & 0.20 & 0.15 \\ 0.75 & 0.10 & 0.15 \\ 0.20 & 0.45 & 0.35 \\ 0.05 & 0.30 & 0.65 \\ 0.15 & 0.20 & 0.65 \\ 0.85 & 0.05 & 0.10 \end{array} \right) \end{array}$$

$$\boldsymbol{E}_{36} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \left(\begin{array}{ccc} 0.25 & 0.40 & 0.35 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.10 & 0.25 \\ 0.20 & 0.20 & 0.60 \\ 0.65 & 0.20 & 0.15 \\ 0.75 & 0.15 & 0.10 \\ 0.15 & 0.50 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.15 & 0.20 & 0.65 \\ 0.85 & 0.05 & 0.10 \end{array} \right) \end{array}$$

$$\boldsymbol{E}_{37} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \begin{pmatrix} 0.25 & 0.45 & 0.30 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.20 & 0.15 \\ 0.15 & 0.25 & 0.60 \\ 0.65 & 0.20 & 0.15 \\ 0.75 & 0.15 & 0.10 \\ 0.15 & 0.50 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.10 & 0.20 & 0.70 \\ 0.85 & 0.10 & 0.05 \end{pmatrix} \end{array}$$

$$\boldsymbol{E}_{38} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \begin{pmatrix} 0.25 & 0.45 & 0.30 \\ 0.10 & 0.35 & 0.55 \\ 0.65 & 0.20 & 0.15 \\ 0.15 & 0.25 & 0.60 \\ 0.65 & 0.25 & 0.10 \\ 0.80 & 0.15 & 0.05 \\ 0.10 & 0.55 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.05 & 0.25 & 0.70 \\ 0.85 & 0.10 & 0.05 \end{pmatrix} \end{array}$$

$$\boldsymbol{E}_{39} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \begin{pmatrix} 0.20 & 0.50 & 0.30 \\ 0.05 & 0.40 & 0.55 \\ 0.65 & 0.20 & 0.15 \\ 0.10 & 0.30 & 0.60 \\ 0.65 & 0.25 & 0.10 \\ 0.80 & 0.15 & 0.05 \\ 0.10 & 0.55 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.05 & 0.25 & 0.70 \\ 0.85 & 0.10 & 0.05 \end{pmatrix} \end{array}$$

$$\boldsymbol{E}_{40} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array} \begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\ \begin{pmatrix} 0.15 & 0.55 & 0.30 \\ 0.05 & 0.40 & 0.55 \\ 0.65 & 0.20 & 0.15 \\ 0.10 & 0.30 & 0.60 \\ 0.65 & 0.25 & 0.10 \\ 0.80 & 0.15 & 0.05 \\ 0.05 & 0.60 & 0.35 \\ 0.05 & 0.15 & 0.80 \\ 0.05 & 0.25 & 0.70 \\ 0.85 & 0.10 & 0.05 \end{pmatrix} \end{array}$$

$$
\boldsymbol{E}_{41} = \begin{array}{l} \\ 0 < \text{age} < 15 \\ 15 \leq \text{age} < 65 \\ \text{age} \geq 65 \\ \text{TertiaryEducation} \\ \text{NoTertiaryEducation} \\ \text{LowEconomicStatus} \\ \text{MediumEconomicStatus} \\ \text{HighEconomicStatus} \\ \text{InternationalFamily} \\ \text{NoInternationalFamily} \end{array}
\begin{array}{c} \begin{array}{ccc} \text{MovementType1} & \text{MovementType2} & \text{MovementType3} \end{array} \\
\left(\begin{array}{ccc}
0.15 & 0.55 & 0.30 \\
0.05 & 0.40 & 0.55 \\
0.65 & 0.20 & 0.15 \\
0.10 & 0.30 & 0.60 \\
0.65 & 0.25 & 0.10 \\
0.80 & 0.15 & 0.05 \\
0.05 & 0.60 & 0.35 \\
0.05 & 0.15 & 0.80 \\
0.05 & 0.25 & 0.70 \\
0.75 & 0.20 & 0.05
\end{array} \right) \end{array}
$$

APPENDIX C

Graphical model outputs

The output of the simulation model, when employing the parameters estimated as the best-fit values, as determined in §9.5, are shown in Figures C.1–C.5.

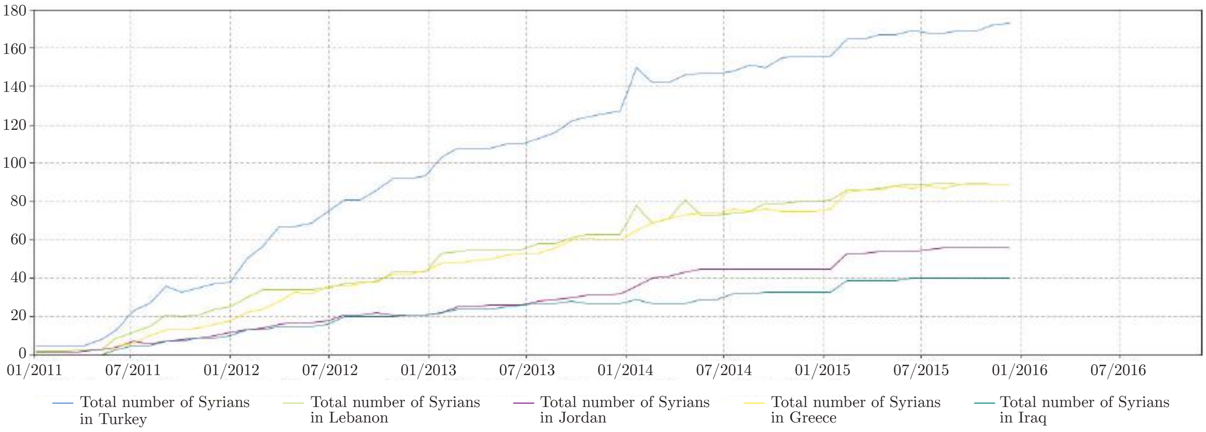


FIGURE C.1: *The total number of Syrians per neighbouring country as simulated.*

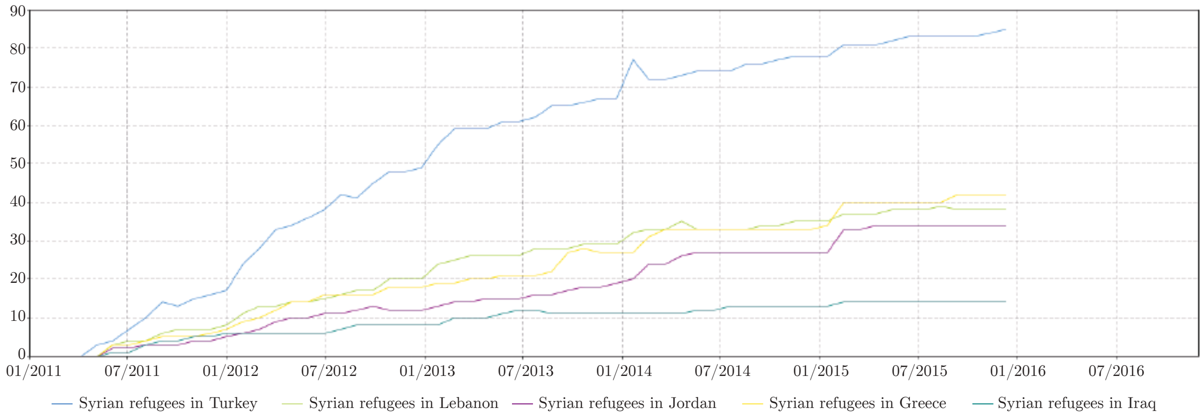


FIGURE C.2: *The number of Syrians refugees per neighbouring country as simulated.*

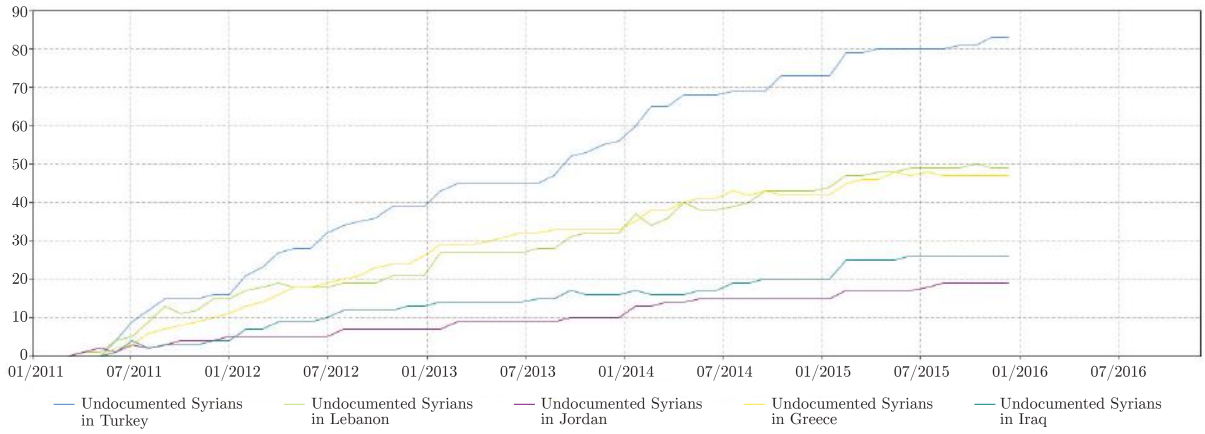


FIGURE C.3: *The number of undocumented Syrians per neighbouring country as simulated.*

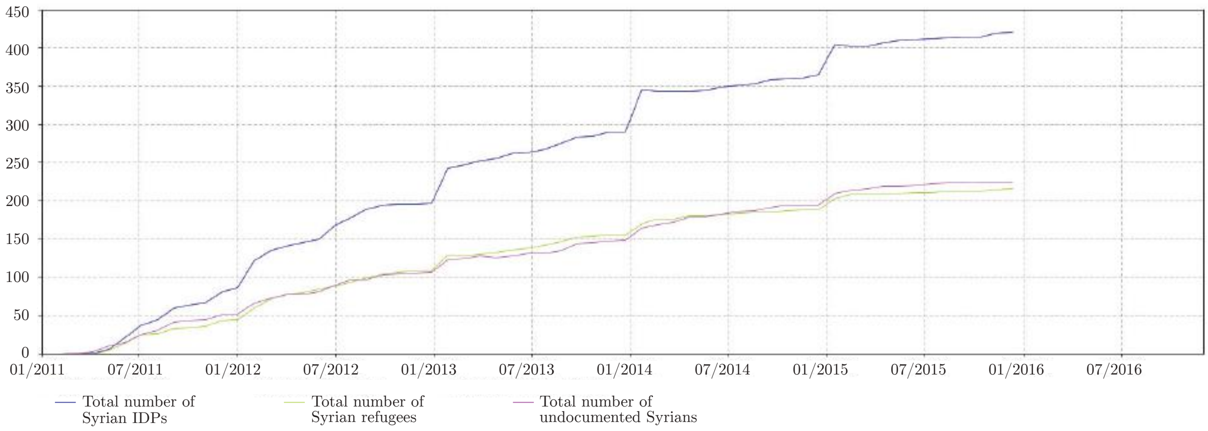


FIGURE C.4: *The total number of Syrians per movement type as simulated.*

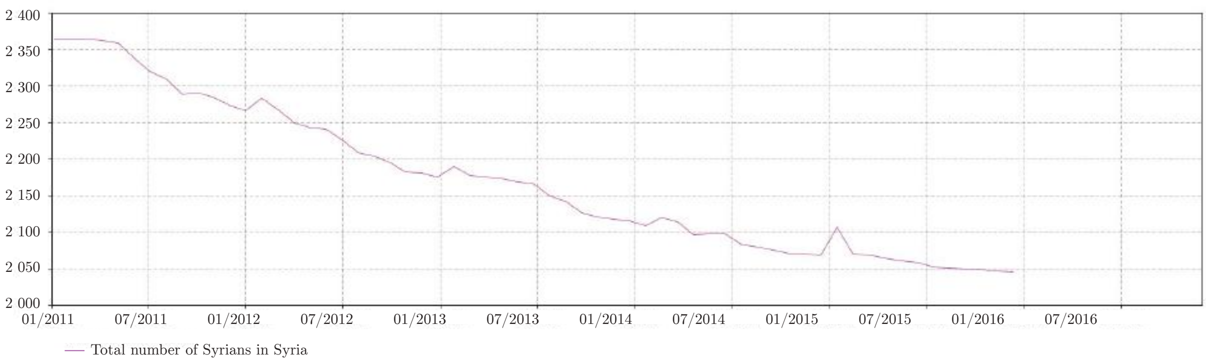


FIGURE C.5: *The total number of Syrians remaining in Syria as simulated.*

APPENDIX D

Contents on a GitHub repository

This appendix contains a brief description of the contents on the GitHub repository, titled "christadk/CoFMMA_framework," for which a link is provided. The GitHub repository contains an electronic copy of the dissertation itself as well as the AnyLogic project file of the final, revised agent-based model concept demonstrator, discussed in Chapters 7, 8 and 9. The model concept demonstrator was created in AnyLogic version 8.4.0 and may be executed in this or a later version of the software. There are two directories on the GitHub repository and their contents are as follows:

Dissertation. This directory contains an electronic copy of this dissertation in ".pdf" format.

ConceptModelDemonstrator. This directory contains the complete agent-based simulation model concept demonstrator described in Chapters 7, 8 and 9 as AnyLogic script files (".alp" format). The simulation model is labelled "ConceptModelDemonstrator.alp." To execute this simulation model, the file should be opened from within 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 7.13 will appear from which the user can either alter the default parameter values or run the model by clicking the "Run" button.

The link to the GitHub repository is as follows:

https://github.com/christadk/CoFMMA_framework.git