ENHANCING AGENT-BASED MODELS WITH ARTIFICIAL INTELLIGENCE FOR COMPLEX DECISION MAKING

وَمَنْ أَحْيَاهَا فَكَأَنَّمَا)
(أَحْيَا النَّاسَ جَمِيعاً

Sheheen Abdullah Abdulkareem

ENHANCING AGENT-BASED MODELS WITH ARTIFICIAL INTELLIGENCE FOR COMPLEX DECISION MAKING

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
Prof. dr. T.T.M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday April 10, 2019 at 14:45 hrs

by

Shaheen Abdullah Abdulkareem

born on April 10, 1983
in Dihouk, Kurdistan region-Iraq

This dissertation has been approved by:

Supervisor: Prof. dr. T. Filatova

Co-Supervisor: Dr. Y. T. Mustafa

Co-Supervisor: Dr. EW. Augustijn

Members of the Graduation Committee:

Chair/Secretary:	Prof. dr. T.A.J. Toonen	University of Twente
Supervisor:	Prof. dr. T. Filatova	University of Twente
Co-Supervisor:	Dr. Y.T. Mustafa	University of Duhok
Co-Supervisor:	Dr. EW. Augustijn	University of Twente
Member:	Prof. dr. N. Numan	University of Duhok
Member:	Prof. dr. A. Heppenstall	University of Leeds
Member:	Dr. D. Karssenberg	Utrecht University
Member:	Prof. dr. H. Bressers	University of Twente
Member:	Prof. dr. R. Sliuzas	University of Twente
Member:	Prof. dr. J. van Hillegersberg	University of Twente

The work described in this thesis was performed at the Department of Governance and Technology for Sustainability, Faculty of Behavioural, Management and Social sciences, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands.

This research project was funded by the split-site PhD program of the Iraqi Kurdistan-region government (KRG) represented by University of Duhok.

Colophon

Cover designed: Haval A. Abdulkareem
Printed by: Ipskamp Printing, Enschede, the Netherlands

ISBN: 978-90-365-4748-2
DOI: 10.3990/1.9789036547482
De URL is: https://doi.org/10.3990/1.9789036547482

UNIVERSITY OF TWENTE.
Faculty of Behavioural, Management and Social sciences (BMS)
Department of Governance and Technology for Sustainability (CSTM)
Enschede, The Netherlands.

UNIVERSITY OF DUHOK (UoD)
College of Science, Department of Computer Science, Duhok, Kurdistan-region, Iraq.

E-mail (for correspondence): s.a.abdulkareem@utwente.nl
Sheheen.abdulkareem@uod.ac

Preface

Every morning we open our eyes, check our phones, browse the social networking sites and read the unfortunate news of many disasters around the world. These disasters vary in their causes, but they are consistent with the results: losses of money and lives. Epidemics are one of these disasters that take hundreds of lives every day throughout the Globe. The spread of infectious diseases does not differentiate between humans in any part of the world they inhabit. What limits their spread is the procedures taken by governments to control them and save the lives of thousands, especially children and elders. Therefore, in order for governments to formulate correct policies in the area of health and prevention, scientific and technical tools such as agent-based models must be available. These simulation models help policymakers to study and analyse past epidemics and their patterns of diffusion, apply different scenarios and prepare for any future emergencies.

My study of Computer Science in the bachelor's degree and then the master's degree in Geoinformatics has provided the fundamentals and important principles in dealing with simulation tools and methods of programming and running them such as artificial intelligence algorithms, coding with different high-level languages, management and processing of spatial database, and data mining techniques. These fundamentals helped me in my doctoral studies and specialize in the application of artificial intelligence algorithms to steer and enhance the behaviour of individuals in simulation models. This in turn will provide decision makers with a tool that simulates the behaviour of individuals during their risk perception and the impact of their spatial and social intelligence on their coping decisions.

Understanding the learning processes of agents in the disease simulation can assist in developing better strategies in health problem-solving and coordination mechanisms. Ideally, the development of policy-oriented agent-based models should go in participatory settings where policymakers could co-design assumptions and develop realistic intervention scenarios. This is the main objective of using and implementing artificial intelligence techniques in these simulation models.

Acknowledgements

Every achievement in this world has a battalion of unknown soldiers besides the financial support, those who stand with all their energies, their efforts, in addition to their love for support in achieving and presenting it with the best image. Now that I have reached this stage of my doctoral study, I should admit that the underlying work would have been impossible to complete without receiving support and help in a number of different ways from groups of wonderful people.

First of all, I want to extend my sincerest thanks to my main supervisor Prof. dr. Tatiana Filatova for her extraordinary scientific guidance, her incredible motivational capabilities, and for showing perpetual confidence in my skills. Thank you, Dr Tatiana, for being a friend, a sister and a great guider. I learnt a lot from you.

I also want to thank my co-supervisors: Dr Ellen-Wien Augustijn and Dr Yaseen Mustafa. Dr Ellen-Wien thank you for hosting me as a daughter then as a student since my very first days in research domain. You were my MSc supervisor and you continued with me to the PhD, your mix of being highly creative combined with heart-warming support have given me great confidence as a researcher, and at the same time made me realise that I am only a beginner in this exciting profession. Dr Yaseen I am so grateful for your engagement in this project and fighting with me to make this dream true. You helped me to digest a lot of concepts that were new for me but with your support I could easily dealt with them. Thank you, Dr Yaseen.

I also want to take a moment to thank my other committee members, Prof. dr Nazar Numan, Prof. dr Allison Heppenstall, Dr. Derek Karssenberg, Prof. dr. Hans Bressers, Prof. dr Richard Sliuzas and Prof. dr. Jos van Hillegersberg. Thank you for investing time reading and evaluating my book. I feel proud and honoured that you have accepted to be on my committee. The thanks are also connected to both Dr Ahmad B. Al-Khalil and Dr Asma' A. Hussein for their scientific and linguistic evaluations on my thesis. Your positive feedbacks made me feel great.

This PhD project connected three departments from two universities. Here, I would take this chance to express my gratefulness and deep thanks to the department of governance and technology for sustainability (CSTM) at University of Twente represented by the Head of the department Prof. dr. Michiel A. Heldeweg and all the staff members who were very helpful and supportive for me all the way of doing this research. Also, I would like to express my special thanks to the active and high responsive secretary Ms Barbera van Dalm-Grobben for all logistic support she used to give with a kind smile.

I would also like to thank the department of Geo-information Processing (GIP) at University of Twente represented by Prof. dr Menno-Jan Kraak and all staff members for hosting me for more than one year and provide me with the facilities that made my work easy to go. In addition, the feedback I had from the department seminar was always helpful to improve my presentations before attending conferences.

My home department, Computer Science at University of Duhok it is your turn to say thank you. You are the home since 2001. Special thanks to the head of the department Dr Luqman M. Qadir for his support and also deep thanks to Prof. dr Ahmed Abdulkhalik Tahir for motivating me not giving up and persistence to finish my PhD successfully. In addition, it is great chance to thank Dr Haval Y. Yacoob for his kind help for preparing and managing all the documents required for the graduation.

I would like to take the advantage of this opportunity to thank the University of Technology Sydney (UTS) represented by the school of Information, Systems and Modelling (ISM) for hosting me from May to June 2018 in Australia. It was a great experience for me to be there. The University and the country itself are wealthy to be there. Special thanks to both Prof Bogdan Gabrys and Dr Katarzyna Musial-Gabrys for hosting meetings and discussing about the motivations behind implementing AI algorithms in ABMs and the main issues regarding data science and complex network systems.

Having distinguished colleagues during the difficult time of work makes the pressure light and sharing experiences together can help to feel you are not alone in the boat. We learn from each other and everyone has a valuable experience that can motivate others positively. Therefore, I would name some of these great colleagues who accompany me in this interesting trip. Thank you Imke, Koen, Leila, Heski, Helmi, Kamia, Karen, Fariba, Monika and Norma for sharing the time at CSTM. Also thank you Tatjana (my first paranymph), Manuel, Hamed and the rest colleagues at GIP. Moreover, thank you Ms Ameera, Mr Majid, Dr Razwan, and the rest colleagues at computer science department. Your advices and support will never be forgotten.

Let me also express my gratefulness and deep thanks to close friends who supported me in all different ways to make this dream true. They kept to listen to me complaining but also feel proud and happy for any achievement I gained during doing this project. Dr Houda El Mustapha, Dr Cesar Casiano, Haidi Abdullah, Samer Karam (my second paranymph), Šaziye Ozge thank you for your support and sharing your research experiences with me. Rasha El Kheshin and Ahmad Fallatah thank you for your generosity of hosting me in Melbourne, Australia. The support of Ahmad Alsamaraai and Sukaina cannot be skipped without saying thank you for being my family overseas.

Beside my great friends in Netherlands, a bunch of wonderful friends in Duhok were also supporting me during this period of more than four years day by day. They were very supportive especially during the discussions on the technical aspects of my work. Their experiences were very helpful to make hints that guided me easily to what I needed to do with my models. I would express my deep thanks to the friend of all times Lamya A. Omar, Ali Shingaly, and Mohammed Hikmat. Thank you for being all the time by my side during this battle. You are real gifts.

Last but not least, the support of all times, my mother (Asyah) and my father (Abdullah). I am so lucky to be your daughter and proud of being a daughter for such wonderful parents. You always do your best to make the life happy and easy for us. You spent your ages to see us in the best positions, safe and happy. I hope I could make you proud of me. I always pray to Allah to keep you healthy and happy and enable me to make you happy and proud. Your always support and encourage to me for gathering knowledge and keep learning are just reflecting how great parents you are. You taught us that by mercy and faith life can better. I wish I am a better writer so I can express all my appreciation and love to you both in little sentences but I totally believe that there is no word to be said. You, my brothers: Reveng, Chalang and Haval (who designed this book's cover), my little girl and lovely sister Dereen and my lovely sister-in-low and love Diljan make me the happiest PhD holder. Ah! Dereen, thank you for feeding me chocolate and orange juice during the months of writing my book. Now you should help me to lose the weight I gained!

These lines cannot be ended without praise and thank Allah Almighty for the His great credit and generosity in giving me the strength, mental and physical abilities, and patience at every moment of my life. Despite all the moments of failure and success that have passed on, I have no choice but to say, Praise be to Allah first and last, with a satisfied heart. Oh Allah, accept this work purely for Your Holy Face and grant me the honesty and sincerity all my way to work satisfying You and serving humanity wherever it is and I am.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABM(s)	Agent-Based Modelling(s)
AI	Artificial Intelligence
BNs	Bayesian Networks
CA	Coping Appraisal
CABM	Cholera spatial Agent-Based Model
GA(s)	Genetic Algorithm(s)
GIS	Geographical Information System
ML	Machine Learning
MOOC	Massive Open Online Course
NNs	Neural Networks
PMT	Protection Motivation Theory
RP	Risk Perception
SABM(s)	Spatial Agent-Based Modelling(s)
SES(s)	Socio-Environmental System(s)
VP	Visual Pollution
ZI	Zero-Intelligence

Chapter 1: Introduction

1.1 Background

Despite the immense progress in science and technology, humanity is still vulnerable to a range of events that disturb the way societies live and develop. In 2018 alone, disasters varied from wildfires in America and Australia (Washington Post, 2018), to epidemics in the developing world (WHO, 2018), as well as to mass migration driven by war or by limited livelihood options in home regions (UNHCR, 2018). These disruptive events continue to generate thousands of human victims and billions of dollars of economic losses annually (IFRC, 2016). Such risk-related problems are complex and involve various actors who participate, interact, learn, and must adapt to constantly changing environments. Effective decisions are made with a short availability of information, under conditions of uncertainty, and limited resources.

Therefore, there is an urgent need for decision-makers and policy-makers to have hands-on scientific tools to help anticipate possible options and develop solutions and interventions, before common risks scale up to become disasters. Policy-makers use applied scientific models to identify and assess possible social and environmental impacts of alternative policies. Simulation tools are particularly prevalent in assessing policy impacts in the domain of sustainable development (Monto, et al., 2005). Simulation models help to identify processes behind unfolding disasters and provide a safe simulated environment to explore managerial strategy responses. Like complex adaptive systems, social-environmental systems (SES) facing risks may exhibit unforeseeable behaviour. Randomness, heterogeneity, and interactions between different entities often make SES mathematically untraceable (Barnes and Chu, 2010; Parunak, et al., 1998; Sun and Cheng, 2005), calling for advanced simulation tools.

In a review of modelling tools for sustainable development, Boulanger and Bréchet (2005) recommend agent-based modelling (ABM) as the most promising approach to support decision-making. An integration of several strengths put ABMs above other methods. ABM is a bottom-up approach that explicitly represents micro/macro relationships and accommodates agent heterogeneity and adaptive behaviour. ABMs allow feedback between the (spatial) environment and cumulative agent behaviours, and are able to integrate a variety of data inputs, such as aggregated and disaggregated data, qualitative information, or even common-sense knowledge (An, 2012; de Marchi and Page, 2014; Filatova, et al., 2013; Fonoberova, et al., 2013; Parker, et al., 2003).

In addition, ABMs serve as a cross-disciplinary platform to integrate advances in social, spatial, and computer sciences when representing behaviour at various temporal and spatial scales, as well as ontological and institutional levels. Further, ABMs of SES often combine elements of other modelling techniques, such as cellular automata, artificial intelligence, and analytical and statistical modelling. In addition, they advance the representation of emergent system properties and phenomena as an outcome of interactions among heterogeneous adaptive agents.

The spatial dimension appears central when studying SES dynamics, especially when risks and societal responses are concerned. ABMs that integrate the model with heterogeneous landscapes are known as spatial agent-based models (SABMs). SABMs explicitly model dynamic environmental processes, ranging from natural environmental processes [e.g., succession of vegetation (Yospin, et al., 2015), flooding (Dubbelboer, et al., 2017), and erosion (Crooks and Castle, 2012))], to the dynamics of a built environment [e.g., growth of cities and settlements (Cantergiani and Delgado, 2016), the construction of roads (Huynh, et al., 2014)], to the emergence of social clusters in space (Sierhuis and Diegelman, 2007). SABMs focus on the "where" question. They often use spatial data from geographic information systems (i.e., GIS data to construct real geographic environments).

Agents are assigned locations in the simulation space, representing their homes, their school or work, or their location during movement. SABMs reflect the richness and variety of the real world that is crucial for an explanation of how spatial structures, such as cities and rivers change and evolve (Crooks, 2010). ABMs in general, and SABMs in particular, are often developed for very specific phenomena or situations with distinct context and data (Simoes, 2012).

SABM have evolved as tools for studying and simulating complex real world processes (Heppenstall, et al., 2012; Borrill and Tesfatsion, 2010). SES applications include agricultural dynamics (Balmann and Happe, 2001; Berger, 2001; Polhill, et al., 2001), land markets (Filatova, 2014; He, et al., 2014; Parker, 2014), and land use in general (Brown, et al., 2005; Matthews, et al., 2007), as well as natural hazards (de Koning, et al., 2017; Magliocca and Walls, 2018), evacuation (Collins, et al., 2014; Li, et al., 2018; Tkachuk, et al., 2018), disaster management (Drakaki, et al., 2018), and the diffusion of infectious diseases (Alshammari and Mikler, 2018; Augustijn, et al., 2016). The rapid evolution and powerful computational abilities on the hardware side enable a large number of mutually interacting spatial agents to be simulated (Husselmann and Hawick, 2011). While agents exhibit adaptive behaviour in SABMs, and in ABMs in general, many models have a simplistic representation of behaviour. A manner in which both social and spatial factors affect agents' learning and, eventually, their adaptive behaviour, varies greatly among models.

1.2 Intelligent Agents in ABMs

To provide a realistic test bed for mimicking human behaviour and societal dynamics, ABMs should consider a range of options to implement learning among agents. The term 'learning' in this thesis refers to activities (processes) of optimising, predicting, decision-making, and adaptation that an agent will execute with the intention to achieve a particular goal. Intelligence and learning are closely related terms (Sen and Weiss, 1999). The ability of a system to learn reflects the intelligence level of that system (Honavar, 2006; Russell and Norvig, 2016). In ABMs, intelligent agents are defined as computational, social interactive, proactive or reactive, and self-directed objects (Macal and North, 2015). They accomplish their internal goals via decisions that are based on strategies or a set of internal rules in dynamic environments due to their ability to learn (Abdou, et al., 2012; Epstein and Axtell, 1996; Gilbert and Terna, 1999; Jennings, 2001; Macal and North, 2015).

Ideally agents should adjust their internal models-- their knowledge about how the world works-- and explore ways to automate the inductive process of generating correct outputs for a large number of input data (Russell and Norvig, 2016). Often, one should develop a measure of success to check if agents have learned correctly about their changing worlds, since learning is defined in terms of improving performance on the basis of some metric (Talwar and Kumar, 2013).

In an SABM, agents operate in a realistic geographically explicit landscape with actual coordinates and can alter this environment or move around over time. Interactions between agents or between agents and their environment have an impact on multiple spatial scales and over various timescales. Therefore, agents must change their behaviour based on their experience over time in response to their environment in a systematic way (North and Macal, 2011). Typically, agents have multiple "options" or "types of behaviour" they can choose to display to reach their internal goal. This can be a one-time decision or when the same decision is repeated. For the latter, the success rate of each attempt is measured so that agents learn to make “smarter” decisions based on experiences. For a one-time decision, agents rely on their prior knowledge and/or the experience of other agents. Smarter decisions can be made by using machine learning techniques, such as genetic algorithms or neural networks, as well as statistical methods, such as regression models (Asadi, et al., 2009; Lorscheid, 2014; Rand, 2006; Sharma, et al., 2012).

1.3 Machine Learning Algorithms

Machine Learning (ML) is a domain of artificial intelligence (AI) that focuses on the development of computer systems, which can learn to improve task performance, and adapt and change when introduced to new data. These systems are able to acquire new knowledge and enhance or refine skills, as well as to recognise prior experience based on newly acquired knowledge. In other words, a computer system learns to improve its predicted future performance (Langley, 1988; Langley and Simon, 1995; Nilsson, 1998).

ML algorithms are useful tools to design intelligent agents with autonomous behaviour (Luger and Stubblefield, 1993; Nilsson, 1998) that have abilities to perceive, reason, and act (Winston, 1992), function more realistically, and perform tasks that require intelligence (Kurzweil, 1990). Researchers in both cognitive science and AI created a wish list of the aspects of intelligence an agent can have (Honavar, 2006). Ideal characteristics of an intelligent agent are perception, action, reasoning, adaptation and learning, communication, autonomy, creativity, awareness, and reflection. Moreover, an intelligent system could also exhibit ingenuity, expressiveness, and curiosity.

Learning processes come in a variety of forms. Two features of the learning processes that are relevant to this research topic are the learning method and learning feedback (Sen and Weiss, 1999). Learning methods vary from rote learning (memorisation), learning from instruction, and learning from examples, to learning by discovery. The main difference between these methods is the amount of learning effort required. For example, the effort required in rote learning is to memorise given facts with no inferences extracted from input information, while learning from instruction requires an instructor that gives new information to be integrated with prior knowledge. Moreover, learning with examples requires the learner to infer and acquire useful examples, from him/herself, a teacher, or the external environment. Finally, learning by discovery requires more effort to perform more inference and learn on the basis of trial-and-error, since no teacher or examples are provided (Michalski and Carbonell, 2013).

To indicate the performance level achieved by agents in the learning process, learning feedback is used. The feedback is provided either by the agents themselves or by the system's environment and comes in one of the following forms:

- **Supervised learning**: the feedback evaluates whether the desired activity of the agent and the objective of the agent match and minimise the differences. The feedback provider acts like a "teacher".
- **Unsupervised learning:** the feedback is not explicitly provided. However, this approach aims to identify useful and desirable activities on the basis of self-organisation and trial-and-error processes. The feedback provider is a passive observer. Agents are left on their own to learn and discover how best to achieve their own goals.
- **Reinforcement learning**: the feedback evaluates the utility of the actual activity of the agent in the current state to maximise the utility value. The feedback provider is the critic. This feedback increases the probability of choices, which has delivered the highest utility in the past, to be chosen more frequently from a set of possible actions.

The three types of learning can be implemented via a range of different algorithms that vary from path finding algorithms (such as, breadth-first search, A*, and hill-climbing algorithms), evolutionary computation (such as, genetic algorithms (GAs)), biological based algorithms (such as, artificial neural networks (NNs)), ML algorithms (such as, decision trees (DT), and random forest (RF)), reinforcement learning algorithms (such as, the Markov decision process), to Bayesian networks (BNs) (see Russell and Norvig, 2016). More details about the implementation of these algorithms in ABMs can be found in *Chapter **2***.

1.4 Implementation of ML in ABMs

Researchers in the field of ABM are aware of the effectiveness of engaging ML algorithms in their models. ML algorithms are employed in ABMs in various ways and for different purposes. ML is considered to be a suitable tool for various steps of ABM design (Oloo and Wallentin, 2017;

van der Hoog, 2016). Currently, ML algorithms are used to improve the performance of the ABM in three ways:

- for processing of input data to calibrate an ABM (Category A in Figure 1-1),
- for employing ML algorithms as agents' brain (Category B in Figure 1-1), and
- for using ML for identifying trends in and visualising of ABM outputs, validating ABMs, and improving their performance (Category C in Figure 1-1).

In this thesis I seek to explore how ML could be used to enhance intelligent behaviour of agents, hence, I focus exclusively on Category B. Out of all the variety of ABMs using ML algorithms to increase the intelligence of agents' brain, this research is confined to ABMs of coupled SES.

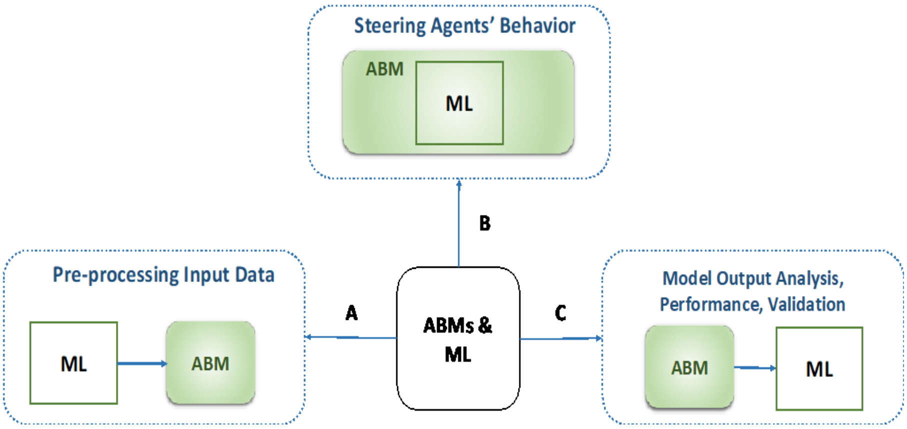


Figure 1-1: Integration of ML algorithms and ABMs

Behaviour of human agents in ABMs may employ various ML algorithms to form expectations and opinions about the environment and future trends of other variables of interests (Balmann and Happe, 2001; Chakraborti, et al., 2011; Happe, 2004; Kirman, 2010). From a social perspective, human agents can be implemented as individuals (Hu, et al., 2017) or as groups (Plikynas, et al., 2014). In addition, an intelligent entity may receive information and exhibit actions either through interactions with other agents 'socially' (Czarnowski and Jędrzejowicz 2018) or with 'environments', which, in socio-economic and spatial ABMs may be spatial (Li, et al., 2018).

One barrier in the use of intelligence in ABMs in general, and spatial ABMs specifically, is that most learning algorithms require extensive training data (Van Der Ploeg, et al., 2014). In the case of data availability, ML algorithms could be trained before they are implemented in ABMs in a supervised learning style (Pope and Gimblett, 2017). When no data is available, the expert may define the parameters of the ML algorithm before implementing it in the ABM. In this case, the training will be done during the simulation (Shen, et al., 2016). In either case, for ABMs to function properly, the behaviour of agents should capture the essential elements of human behaviour.

Models with intelligent agents may help policy-makers to extract important, possibly hidden relationships and correlations among large heaps of data (data mining). The learning capability increases the autonomy of agents that drive unexpected results on micro- and macro-levels (Alonso, et al., 2001; Lorscheid, 2014). Moreover, learning behaviour endows agents with an ability to rationalise in an uncertain and dynamic world (Russell and Norvig, 2016). In summary, the benefits of employing ML in ABMs are vast (Nilsson, 1998; Stone and Veloso, 2000):

- ML enables agents in ABMs to adjust their internal models-- their prior knowledge on how the world works-- and explore ways to automate the inductive process that help them to perform well on their core tasks.
- When designing a model, a system developer has incomplete knowledge about the environment in which the system will be applied. ML provides the ability of using “on-the-job” betterment of existing system designs.
- Certain tasks might require too much knowledge to be explicitly encoded by the developer. Therefore, there is a demand for having systems that gradually extract and learn to use this knowledge to help the developer capture certain behaviours.
- Using ML could help the developer to write less programming codes while handling large knowledge and designing agents’ tasks.
- The dynamic nature of environments requires agents to promptly adapt and respond, thus, ML can be used to avoid static design.

There are two pronounced research gaps in implementing learning in SES ABMs. First, many ABMs use naive deterministic algorithms, which are rule-based or condition-based, to simulate a behavioural change in agents (Heppenstall, et al., 2016). While agents in ABMs are sometimes endowed with memory (prior knowledge), actual learning in an AI style is rarely implemented (Balbi, et al., 2010; van der Hoog, 2016). The study of adaptation, expectations formation, and behavioural changes involves a change in agents' preferences or perceptions within the ABMs and could greatly benefit from the use of state-of-the-art knowledge developed within the AI community (Rand, 2006). Yet, the ML-based endogenous switching of behavioural choices of agents and their expectations formation about consequences or future events is underdeveloped in ABMs of SES (Kocabas and Dragicevic, 2013; van der Hoog, 2016). Moreover, a structured comparison of ABMs with zero-intelligence vs. ML-based learning is missing.

Second, the majority of SES ABMs employ both social and spatial dynamics, implying that both processes may affect agents' decision-making. ML could support an integration of both social and spatial dynamics to assure agents' learning in rich SES environments. Yet, to date, such modelling examples are scarce. A combination of both social and spatial factors influencing individual agent decision-making can be based either on a theoretical model, on data, or on both. However, limited data recording both spatial and social factors is available to guide a data-driven model. The second knowledge gap is in the lack of developments of methods to integrate social and spatial factors, both from a data and from an ML algorithm point of view.

This thesis addresses these gaps at the overlay of the three domains (Figure 1-2). It relies on the methods from ML to enhance the agent's intelligence in a spatial ABM, employing the insights from social sciences on risk perception, in particular, using Protection Motivation Theory (PMT) and data from a social survey on factors affecting risk perception.

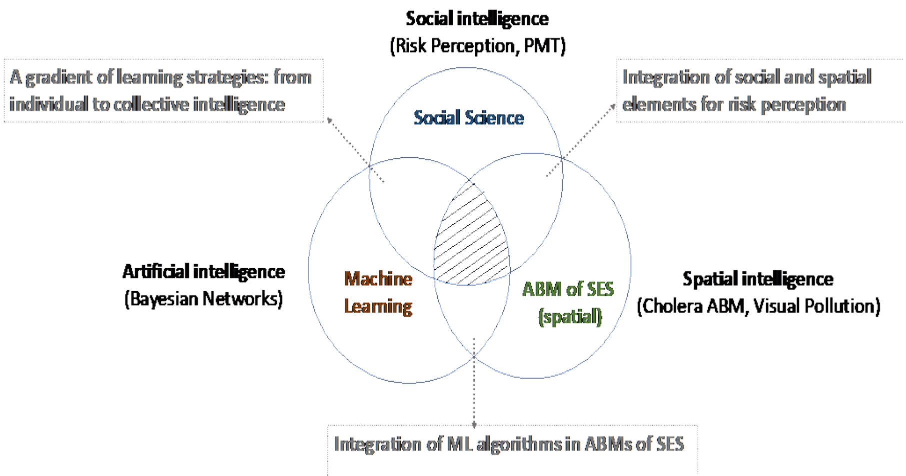


Figure 1-2: This PhD thesis at the intersection of three scientific domains

1.5 Research Objective and Research Questions

The main goal of this research is “to get insights into the implications of ML integration in agent-based simulation models”. It focuses on how ABMs are developed to support realistic policy decisions that may be advanced by enhancing agents’ intelligence with ML algorithms. Specifically, the thesis compares ABMs that pursue the various operationalisation of ML algorithms accounting for spatial and social intelligence that drive agent behaviour. Two sub-objectives serve as stepping stones to achieve this main research goal, in line with the above-mentioned gaps. Each objective is supported by related research questions (RQ):

Sub-Objective 1: to provide insight on how ML algorithms can be integrated into ABMs of SES.

> **RQ1**: What is the state-of-the-art in employing intelligent agent learning into ABMs of SES to enhance agents’ decisions? (Chapter 2)
>
> **RQ2**: How can the spatial and social intelligence driving agents’ decisions under risk be implemented in an ABM? (Chapter 3)
>
> **RQ3**: How can the supervised learning of ML algorithms be implemented in ABM, given scattered micro-level data? (Chapters 4 and 5)

Sub-Objective 2: explore the implications of learning, including social and spatial intelligence, on the behaviour of agents choices in risky contexts.

> **RQ4**: How comparable are results of an ABM with intelligent decision-making agents to the one with zero-intelligent agents (i.e., rule-based learning)? (Chapter 3)
>
> **RQ5**: Given the reliance of ABMs on social interactions, what differences does the level of collective intelligence make when implementing an ML algorithm in an ABM? (Chapter 6)

To achieve this goal, different implementations of BNs are tested in a spatial ABM using a Cholera disease diffusion ABM as an example. A learning target of agents is the risk perception and their behaviour when facing risk.

1.6 Case Study: Modelling the Spread of an Infection Disease

Cholera ABM (CABM) is a geographically explicit model that simulates an environmental reservoir of Cholera bacteria in the urban area of Kumasi, Ghana (Augustijn, et al., 2016). The objective of the original CABM was to test the role of a water runoff from open refuse dumpsites as a pathway for the dissemination of Cholera. CABM simulates two different Cholera infection pathways, via the environment (lower infectiousness) and human-environment-human infection (hyper-infectious). When passing through the digestive system, Cholera bacteria transition to a hyper-infectious state. When faecal materials from Cholera patients are deposited at open dumpsites, runoff during heavy rains can carry the infection to nearby rivers, and as people use the river water for domestic use, this runoff can contribute to the diffusion of the disease. This model incorporates environmental and human behavioural elements, and could be used to explore policy interventions to reduce the spread of Cholera.

There are three agent types in CABM: households, individuals, and rain particles. Household agents are collections of individual agents. The model consists of three sub-models: a hydrological model, a household activity model, and a disease model. Agents are positioned in the geographically explicit environment which consists of different spatial layers of GIS data for the city of Kumasi. Households and individual agents are heterogeneous in terms of their attributes, such as income level, hygiene level, water source, as well as the location for households, and age, educated/not, gender, blood type, and health status (susceptible, infected, and recovered) for individuals. However, in the original version, households were homogeneous in their behaviour and individual behaviour was not explicitly implemented. The agent population (households and individuals) is generated using a synthetic population generator that provides the model with its largest stochastic element.

The study area is 19 km^2 and consists of 21 communities (Figure 1-3, left). There are no administrative boundaries for these communities. However, for this model, the developers determined the boundaries using Thiessen polygons. The spatial environment of the CABM consists of: (1) elevation surface data (DEM) to define the hydrological dynamics that determine the flow direction and flow accumulation of the rain drops, (2) the dumpsites with actual locations gathered using a global positioning system (GPS), (3) the house layer with income levels ranging from high to medium and to low; (4) the river, and (5) the centre and ID of communities (Figure 1-3, right).

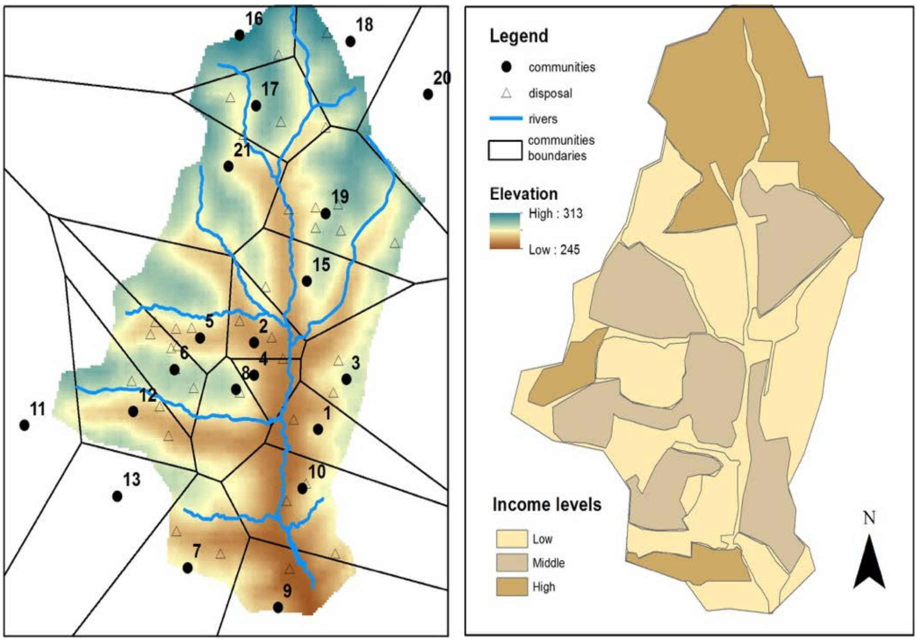


Figure 1-3: CABM study area that is located in the North-East part of Kumasi

The original model simulated the process of Cholera diffusion without elaborating the decision-making of the household and individual agents (Figure 1-4). Agents had a fixed activity pattern. Depending on their income level, household agents obtained water with either a tap, by purchasing water, or by fetching it from the river nearest to their home. They also used the closest dumpsite to their home location. The model was originally developed to advance the spatial dynamics and could be expanded by including change of behaviour among individuals based on disease awareness.

The learning skills of the individuals are missing in the current model. This opens the door to use this SABM to implement different learning strategies. ML algorithms can be employed in the CABM within the two agent types (households and individuals) based on their activities in the model:

- Households and individual agents in the CABM could benefit from intelligent decision-making when:
    - o Searching for the best source of water (river, tap, or buy bottled-water) based on their risk perception (e.g., the number of disease cases the household is aware of). This is

an iterative process where households adjust their behaviour based on constant risk perception during the simulation.

- Changing the place where the water is collected from the river. This requires that the agent has an understanding of spatial patterns and can judge the difference between the various parts of the river.
- Adjusting one's hygiene level (e.g., treating water or not) based on previous experience and awareness of the disease.
- Sensing whether any of the neighbours are infected.
- Drinking/using the water.

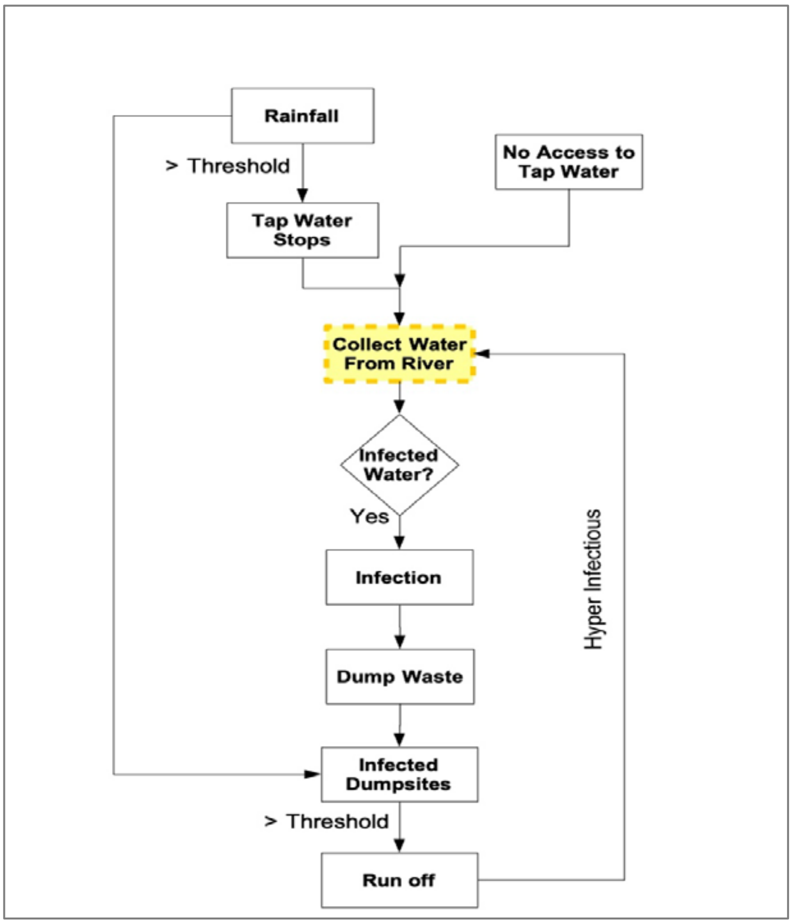


Figure 1-4: Original model scheme where the intelligence elements will be added during the process of collecting water from the river (yellow box)

Consequently, an agent in the case study of SABM will move through a cycle during the simulation time (Figure 1.5). With every timed step, the agents have their own demands and needs to make a decision, however, this decision is governed by the agent's sensitivity to risk perception. The result of the decision the agent makes will be stored in their memory as its learning experience, as shown in the figure below:

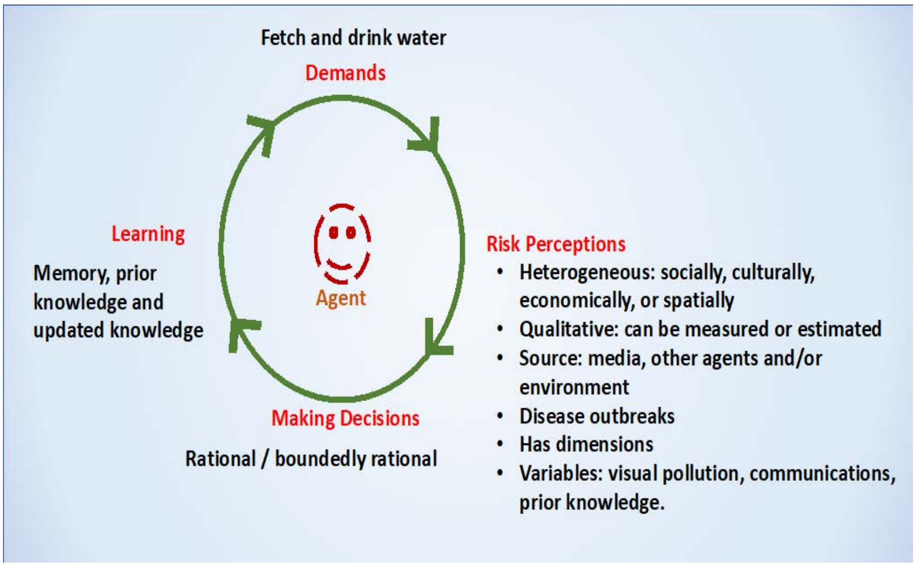


Figure 1-5: Agent's life activity cycle inside CABM

1.7 Modelling Behaviour Changes in Risky Contexts

Adaptive behaviour of individuals is crucial when modelling SES. Understanding factors affecting a shift in behaviour of an individual enables one to trace cumulative consequences for a community, city, and society in an ABM. This is especially important when studying decisions under risk and major events that adversely impact societies. Reconstructing a social phenomenon from the bottom-up in a simulated environment offers the decision-makers an artificial society to test alternative response strategies that minimise losses from disasters and societal costs.

One of the aspects that impacts individual behaviour during a disaster is the perception of risk. Individuals form expectations on how risky a situation is and respond by adapting their behaviour to a new situation.

Risk perception (RP) is an integral part of a decision-making process in uncertain situations. Moreover, RP can be understood as an individual's evaluation of risk in a particular situation. This evaluation is the resolution of an individual uncertainty on how threatening and controllable the situation is. The sufficiency of any risk evaluation is based on the adequacy of accessible risk information (Pablo, et al., 1996). Accordingly, risk impacts the evaluation of available options, the eventual decisions, and perceptions of the decision problem (Williams and Noyes, 2007). Risks that an agent may face can be objective, such as the probability of low rainfall, but they can also be subjective based on individual exposure to various shocks (Doss, et al., 2006). Agents’ subjective assessments associate their expectations about the probability of various events with agents’ beliefs about their capabilities to deal with various emergencies.

Humans have a limited cognitive ability, which affects an individual’s RP evaluation which, in turn, could result in inadequate decisions. The sufficiency of any risk evaluation is based on the sufficiency of the accessible risk information. To realise the effect of RP on the process of decision-making, the way risk information is communicated and received by agents should be understood (Williams and Noyes, 2007). Factors that influence RPs are the message, the source of the message (other agents, and/or the environment), and the target of the message (agents). These factors need to be considered to design effective risk communications and to facilitate decision-making. It is reasonable to conclude then that any effort to understand the effects of exogenous variables on decision-making must consider the role of the RPs (Sitkin and Weingart, 1995).

Social science has a long-standing tradition of studying RP, factors affecting it, and its serving as a trigger for behavioural change (Sjöberg, 2000; Slovic, 2010). Protection motivation theory (PMT) is prominent in conceptualising this process and is used extensively to study health risks (Floyd, et al., 2000). By assuming that decisions in a risky context are made in two steps, as a risk appraisal followed by a coping appraisal, PMT provides a clear link between factors affecting RP and a choice of actions. Besides being used frequently in empirical studies, PMT seems to be a straightforward way to formalise an ABM. Hence, taking the CABM as this case study model, this research further explores how RP is shaped by various factors: whether it triggers a behavioural change among households, and where and how intelligence makes a difference.

1.8 Data

Any model requires a range of input data: spatial data, agent attributes data, and data to formalise learning. For the CABM, data from the Ghana Bureau of Statistics (GSS, 2012) is used to create the synthetic population of individuals and households. Poverty data is derived from literature (Augustijn, et al., 2016). Data on access to tap water was derived from national statistical information from the Ghana Statistical Service (GSS, 2012). The dataset of confirmed Cholera cases for the 2005 epidemics were confirmed by a bacteriological test and were reported to the Disease Control Unit (DCU) by reporting facilities (Osei and Duker, 2008). The DEM was downloaded from CGIAR website as a Geotiff image. Flow direction and flow accumulation layers have been calculated based on this DEM using ArcGIS. Houses were digitised based on the Google image of the area (2006), and refuse dump locations have been collected using GPS (Osei, et al., 2010).

Limited datasets are available about the way the spatial environment influences human decision-making. Most sources discuss RP by evaluating how RP varies in space (e.g., Sridhar, et al., 2016), omitting the role the environment itself plays in the process of shaping RP. Therefore, two online surveys were conducted to gather data on people’s RP for Cholera disease: the MOOC survey (Geohealth online course) and Google survey (an online survey). While most of the questions were identical in the two surveys, there was one difference. In the MOOC survey, participants chose to use or not to use river water for drinking by judging its quality by visual appearance (pictures shown in Chapter 4). The Google survey collected information on the influence of individual risk factors on the willingness to use river water without visuals, using only a textual description of the water quality. The survey data on RP and factors affecting it were used to introduce intelligent judgements about risks and options to cope with disease in the CABM.

1.9 Outline of the Dissertation

The thesis consists of seven chapters (Figure 1-6) that sequentially address the research questions:

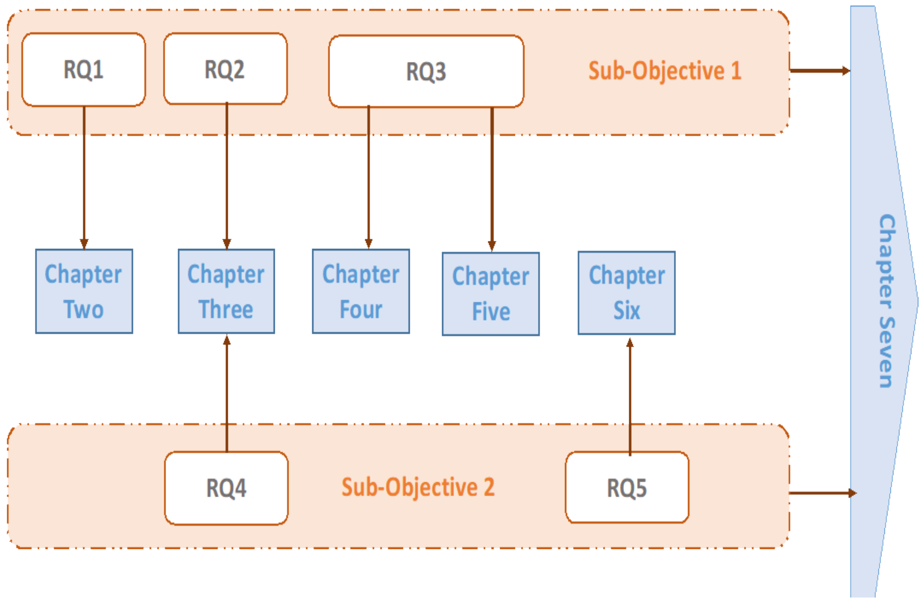


Figure 1-6: The overview of the thesis

Chapter 1 introduces the reader to the field of ABMs and the growing need for developing simulations with intelligent agents. Intelligent agents may use ML algorithms to integrate social and spatial factors to improve their tasks in SES models. The benefits and limitations in the implementation of ML algorithms in ABMs are briefly outlined. The main goal of the thesis, its sub-objectives, and related research questions aim to address the gaps at the intersection of ML, ABM, and social science domains. The chapter presents the case study models to be used and the nature of the datasets employed.

Chapter 2 reviews recent ABMs of SES that employ various learning algorithms to create intelligent agents with a focus on spatial ABMs. Here, a systematic structured analysis is provided of (1) the ways learning algorithms are employed in ABMs for only social, only spatial, or combined social-spatial intelligent decision-making, (2) their specific operationalisation in an agent's decision-making for various tasks: individual versus group learning and the treatment of spatial and social environment in the design of learning algorithms, and (3) the level of empirical data used in ABMs in either the pre-training of the ML algorithm or training during a simulation. This chapter highlights the trends in the current practice of ML algorithms used to enhance ABMs, which social simulation modellers may rely on when designing future simulations.

Chapter 3 presents an innovative approach to extend agent-based disease models by capturing behavioural aspects of decision-making in a risky context using ML techniques. This is illustrated with a case of Cholera in Kumasi, Ghana, accounting for spatial and social risk factors that affect intelligent behaviour and corresponding disease incidents. The thesis discusses the results of computational experiments by comparing spatial and temporal patterns of disease diffusion among zero-intelligent agents with those produced by a population of intelligent agents. A spatial disease ABM is presented with agents' behaviour grounded in PMT from psychology. To introduce agents' intelligence, I designed and coded two BNs in R statistical language, and integrated them with the NetLogo-based CABM. The first BN is a one-tier (BN1), the only RP, and the second is a two-tier (BN2) for risk and coping behaviour.

Chapter 4 is a continuation of the study presented in Chapter 3. It focuses on validating the spatial intelligence by collecting data on people's RP for Cholera via two online surveys: the MOOC and Google surveys. Spatial intelligence refers to the fact that agents sense their environment, perform a judgement on its dynamically changing conditions, and adjust their behaviour based on this judgement. Objectives of this chapter are twofold: to examine the effect of spatial and social RP on disease spread, and to compare the risk awareness of agents with data collected on the RP of the survey participants.

Chapter 5 presents a methodology for training a learning algorithm to guide agent behaviour using limited survey data samples. Various implementation strategies were applied using survey data and BNs. By being grounded in probabilistic directed graphic models, BNs stand out among other learning algorithms in that they can be based on expert knowledge and/or extensive datasets. This chapter presents four alternative implementations of data-driven BNs to support agent decisions in the CABM. Chapter 5 provides a differentiation between training BNs before or during the simulation runs, using only survey data or a combination of survey data and expert knowledge.

Chapter 6 pursues a quantitative test on the influence of agents' ability to learn-- individually or in a group-- on the disease dynamics. The experiments illustrate that individual intelligent judgements about disease risks and the selection of disease coping actions are outperformed by social intelligence (acquired individually or leader-based). The impact of different types of social learning compared to individual learning is an underexplored domain in disease modelling and in ABMs of SES in general.

Chapter 7 provides the conclusion of this research. The conclusion includes the answers to the research questions, the reflection of this PhD project, and the limitations that lead to future work.

Chapter 2: Artificial Intelligence for Enhancing Actors’ Decisions in Agent-Based Models: A Review

2.1 Introduction

Agent-Based Models (ABMs) are indispensable for studying the aggregated impacts of individual actions of heterogeneous agents. Concurrently, Artificial Intelligence (AI) has been employed for decades to simulate autonomous actions of individual entities that react, learn and exchange information with an environment and one another. There are obvious synergies between the two computational approaches – i.e. ABMs and AI – as also discussed in Chapter 1. For example, AI could be used to enhance agents' behaviour in ABMs. Machine learning (ML), as a method to implement intelligence, allows for a richer agents' architecture. ML can help in the operationalization of more realistic learning for reaching decisions beyond a simplistic treatment of agents' cognitive and sensory capacities.

Human beings make decisions both individually and as part of a collective, where an individual could copy a decision taken by a group or a group leader (Carlson et al., 2014). Therefore, ML algorithms can be integrated into ABMs for individual and group learning. In either way, agents may learn in isolation or through interactions with relevant others, e.g. with neighbours or peers within own social networks (Sen and Weiss, 1999). In isolated learning, an agent learns independently without requiring any interaction with other agents. In interactive learning, several agents are engaged in the same process of learning, and they need to communicate and cooperate to learn effectively. Interactive learning can be conducted based on different social learning strategies (Eberlen et al., 2017). In addition to social intelligence – either group or interactive – agents also learn from the environment, which in the case of socio-environmental systems (SES) is often geographically explicit. Hence, spatial intelligence (Gardner, 2006) is also an important aspect to consider in this thesis. In ABMs, spatial intelligence concerns the use of ML algorithms to capture the process of how spatial environments, and especially changes in these environments, influence agents' decisions (Anderson and Dragićević, 2018; Ghnemat et al., 2008; Sahli and Moulin, 2006; Yang et al., 2011).

Besides, intelligence in ABMs can be represented by a fully trained learning algorithm at the start of the simulation, or by training it during a simulation run. When applying supervised learning, ML algorithms need to be trained on data. One limitation in the use of intelligence in ABMs is that most learning algorithms require extensive training data (Van Der Ploeg et al., 2014). Moreover, the scarcity of large sets of micro-level data on human behaviour, which affects the employment of ML algorithms, has been a long standing problem (Kocabas and Dragicevic 2013). The performance of learning algorithms improves with the increasing quantity and quality of training data (Walczak and Walczak, 2001). When no data is available, an expert may define parameters' values, and the training of an algorithm is done during a simulation. However, to our knowledge, no detailed review has been conducted that studies the available data sources and the way they are used in ABMs enhanced with ML.

Complex emerging behaviour can be the result of combinations of previous experiences of an agent (**feedback**), of social interactions with other agents, but also of changes in the agent's environment. ML algorithms can play an important role in combining a large number of different spatial and social variables and in obtaining the social, spatial or social-spatial intelligence level required (Heppenstall et al., 2014; Manson, 2005; Van Der Hoog, 2016).

There are several comprehensive reviews on the integration of learning algorithms in socio-economic and spatial ABMs. They include an overview of ML algorithms in environmental modelling and ecology models (Hamblin, 2013; Chen et al., 2008), early efforts on the usage of AI in socio-economic ABMs (Chattoe, 1998; Gilbert and Troitzsch, 2005; Rennard, 2006) and a general discussion on the use of learning algorithms in ABMs (Lorscheid, 2014). Also worth mentioning is the review of Eshragh et al., (2015) discussing the role of AI in automated negotiation in environmental resource management ABMs. However, these reviews either focus on a particular application domain (e.g. ecology, navigation), not necessary relating to ABMs, or on a particular agents' task (e.g. negotiation). A thorough review of literature on ABMs of SES with intelligent agents is missing. This chapter reviews recent socio and spatial ABMs that employ different learning algorithms to investigate how to create intelligent agents with combined social and spatial intelligence. We do this by conducting a review of literature on ABMs of SES and differentiating between two groups: spatial and non-spatial ABMs.

In this chapter I perform a systematic structured analysis of ABMs of SES using ML algorithms to enhance agents' intelligence. The review focuses on:

1) the way ML algorithms are employed in ABMs for only social, only spatial or combined social-spatial learning of agents,

2) their specific operationalization in the agents' decision-making for various tasks, differentiating among individual versus group learning, and among spatial and social environment in the design of learning algorithms,
3) and the level of empirical information used in ABMs in either a pre-training of an ML algorithm or its training during a simulation run.

This chapter highlights the trends in the current practice of learning algorithms used to enhance ABMs. It also offers 'lessons learned' from this practice, which social simulation modellers may rely on when designing a new generation of ABM simulations.

2.2 Methods

This article surveys ABMs that are designed to explore the dynamics of SES, or at least a subsystem of them. Within this group, I focus on those ABMs that use ML to steer behaviour of agents in either spatial, social or combined socio-spatial learning. Overall, 137 articles are included in this review, of which 60 are non-spatial and the remaining 77 report spatial ABMs. I selected the articles for this review following a number of steps:

- Scientific web engines such as Scopus, and Google scholar (and other relevant databases such as ACM Digital Library, IEEE Xplore, Arxiv.org, and web of Science) were searched looking for different combination of keywords such as agent-based models or agent-based simulation or spatial agent-based modelling or multi-agent systems AND machine learning or artificial intelligence or learning algorithm or intelligence algorithms;
- Using the above-mentioned engines, a search was conducted for ABMs with AI keywords such as genetic algorithms or neural networks or Bayesian networks or fuzzy logic intelligent decision;
- Snowball sampling: recursively finding relevant articles through the reference list of various ABM articles.

I used the following list of criteria to review the ABM literature employing learning algorithms to steer agents' behaviour and enhance agents' decision-making abilities:

1) the way learning algorithms are employed in ABMs to represent social, spatial or socio-spatial intelligence,

2) their specific operationalization in an agent's decision-making for various tasks,
3) the level of empirical information used in ABMs to train an ML algorithm.

Each of these categories will be explained in more detail in the next paragraphs.

2.2.1 *Type of Intelligence*

Wooldridge and Jennings (1995) defined an intelligent agent in spatial settings as being reactive to changes in their environment, pro-active in the sense that they have goal-directed behaviour, and having social abilities. Following this, in my review, I differentiate between three types of intelligence: social, spatial and combined socio-spatial intelligence. Namely:

- *Social intelligence* involves the acquisition of new skills or knowledge by perceiving information, experience, and the performance (actions) of other agents. Agents interact with other agents, for example, using negotiation tasks in order to reach an individual decision (Figure 2-1.a).
- *Spatial intelligence* refers to the process of receiving information, from the spatial landscape. Processes can be based on spatial intelligence in the case of an agent using an algorithm to evaluate its spatial environment to make a decision (Figure 2-1.b). Decisions can be based on recording changes in the environment (e.g. water levels that are rising) or knowledge about locations (e.g. finding an exit or determining the shortest route) or comparing the quality of a spatial location (e.g. finding the best property to buy or avoiding heavily polluted areas).
- In combined *socio-spatial intelligence*, agents have to combine the information from the social interactions with information from their spatial environment to come to a decision. A resolution on how to combine these different factors is challenging in this type of intelligences (Figure 2-1.c).

a. Social Intelligence   b. Spatial Intelligence   c. Socio-spatial intelligence

Figure 2-1: Types of Agents’ Intelligence

Further, I focus on the type of tasks that is being performed by intelligent agents. In ML, tasks are often classified as un-supervised, supervised and reinforcement learning. For this review, this classification is too coarse. In ABM literature, agent tasks vary from predicting a possible future own state or the state of the environment to negotiating with others. Therefore, I want to be more specific, especially since the remainder of this thesis will focus on “risk perception” and “coping appraisal” decisions of agents facing risky choices. Hence, the review differentiates between the following tasks of intelligent agents:

1) *Optimization (OPT)* - concerns the search for the best action or decision from a set of alternatives based on one or several criteria, that might require no prior knowledge to learn a suitable cooperative.
2) *Negotiation (NEG)* - is a dialogue with a purpose of reaching an agreement that may bring mutual advantages to involved actors.
3) *Prediction (PRED)* - Prediction is an attempt to forecast the future.
4) *Adaptation (ADPT)* - Adaptation is an alteration of behaviour or attributes of an agent in response to changing surroundings. The latter in our sample of ABM papers may be represented by a spatial environment, or by a society.

Usually, agents have multiple options regarding the “types of behaviour”, which they can choose from to reach their internal goals. Agents in a model act – i.e. optimize, negotiate, predict and adapt – to eventually make a core decision. The decision can be a one-time or a repeated event. For the latter, a success rate of each attempt is measured, so that agents learn to make “smarter” decisions based on their experiences. In the case of a one-time decision, agents might rely on the experience of other agents. For the spatially explicit cholera diffusion ABM I use as a case-study in the thesis, “risk perception” is an example of a prediction action and “coping appraisal” can be regarded as a form of adaptation. As agents adapt their behaviour according to the risk they perceive, they may choose a different decision with respect to the source of water they use.

2.2.2 Implementation Strategies

When implementing ML to enhance agents’ decision making in a model, a choice has to be made about a specific algorithm. In many papers, no specific motivation is provided for this choice, hence the rationale behind the use of, e.g. Neural Networks (NNs) over Bayesian Networks (BNs) is not transparent. Many ML algorithms are available and preferences for one or another shift over the years as better algorithms are being developed. Therefore, for this review, I took a pragmatic approach and limit the further discussion only to the algorithms used most often in the dataset of papers surveyed. Finally, the following ML algorithms were selected: Bayesian networks, Neural networks, Genetic algorithm, Swarm intelligence, Hybrid algorithms and I group the remaining less frequently used ML methods in the ‘Other algorithms’ category (Figure 2-2).

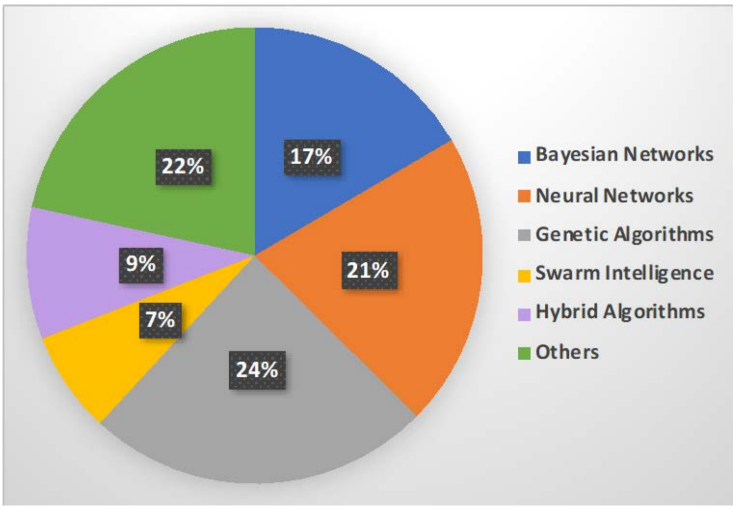


Figure 2-2: Use of different ML Algorithms to enhance agents’ intelligence in ABMs of SES (N=137 reviewed papers).

Further, I distinguish between two different types of strategies to implement an ML algorithm in an ABM with respect to the object of intelligence, an individual agent or a group agent.

Individuals versus groups

An object that pursues intelligent behaviour – an 'individual' or a 'group' – employs various ML methods in the reviewed articles. Here a 'group' should not be mistaken for a set of separate agents connected through social ties. For example, individuals belonging to the same neighbourhood may be influenced by one another but make individual decisions without necessarily having a group goal. With group intelligence, individual agents either individually or jointly use a learning algorithm to support their decision making while striving to achieve a common group goal.

This review further uses the following definitions:

- *Individual Intelligence* refers to the process of gaining skills or knowledge, which an agent pursues individually to support its individual task (Russell and Norvig, 2010). In this case an algorithm is implemented at the agent level and its types or parameterization may vary across agents. Learning depends on prior knowledge such as memory, experience, and/or the perceived knowledge awareness of the environment or actions of others.
- *Group Intelligence* is the process of acquiring new skills or knowledge undertaken collectively in a group of several agents (Sen and Weiss, 1999). In this case, a group performs a unique group task, and ML is used to help the group reach its goal. Group intelligence can be realized by introducing one learner-enhanced agent who learns for the whole group to help it accomplish its group task (e.g. an opinion-leader or a leader-dictator). Alternatively, group intelligence is implemented by using individual ML algorithms for various group members who learn individually or perform specific sub-tasks that support the entire group reaching its goal. A group then makes a decision by combining this individual knowledge, e.g. by majority vote.

Either an individual or a group could make decisions and pursue its goals in isolation or in interactions with other individual or group agents.

Isolated versus Interactive learning

The process of learning includes interactions between the learners – individual or group agents – and the environment or other agents (Shalev-Shwartz and Ben-David, 2013). Social interactions are a core attribute of ABMs. Hence, nearly every model has agents pursuing interactions with other agents to achieve their goals but not every ABM would enhance this process with ML. According to Sen and Weiss (1999) an intelligent agent may pursue isolated (centralized) or interactive (decentralized) learning.

This review further uses the following definitions:

- An *isolated learner* receives information only through the environment without direct interactions with other agents. In this case information comes from own experience, spatial environment, media or institutions.
- *Interactive learning* implies that agents communicate and interact with each other to learn effectively. Such interactions are often implemented by instantiating social networks in ABMs or by connecting agents in spatial neighbourhoods.

When an ABM includes social interactions among agents, we define this model as socially **interactive**, otherwise we call agents **isolated**.

2.2.3 Data for Training ML Algorithms

The lack of empirical data may play a role in the selection and specification of an ML algorithm. The availability of data influences whether an ABM developer can calibrate model parameters, extract and estimate missing information for agents' decision rules, or train an ML algorithm.

Training of the ML Algorithms

Any ML algorithm should have a mechanism to select the rule for achieving the learning task it strives for. Hence, a learning algorithm needs to be trained to learn how and what answer to give using a feedback. The Learning feedback are used to indicate the performance level achieved by the agent. The feedback is provided either by the agent itself or by the system environment. The learning feedback might be realized as a *supervised*, *unsupervised* or *reinforced* learning. The choice of a training method of an algorithm often depends on the availability of data for the feedback. Specifically:

- *Supervised learning* is possible when empirical data is available. Here the algorithms' learning parameter values can be driven directly from the dataset.
- *Unsupervised learning* When no data is available, the expert may define the parameters' values and the algorithm training will be done during the simulation.
- *Reinforced learning* involves data gathering on the go through iterative tasks such as negotiation. It relies on a performance evaluation and requires feedback from the environment.

The ML algorithms require training that can be either pre-training (before implementing the algorithm in the ABM) or training during simulation:

- In the case of *pre-training*, the algorithm learns from the data and it connects every set of inputs with the corresponding outputs. In this case, agents have the answers for every situation and know what to do in which case. The agents learn how to match their prior knowledge with the current status of the environment to react accordingly. Supervised learning relies on pre-training.
- In the case of *training during the simulation*, agents and their ML algorithm learn together and upgrade based on the experience of the agents. In addition, the learning parameters of the ML are adjusted during the simulation using the data from the model itself. Unsupervised and reinforced learning rely on the training that occurs as simulations run.

Data Types

Not all data used to train the ML algorithm is empirical behavioural data on the actual choices. At the same time, a design of ML in ABMs may depend on the type of data that is available. This applies to the learning algorithm that is chosen, or the type of training applied. Not all data used to train the ML algorithm is empirical behavioural data. Many different types of data have been used in ML algorithms. We differentiate between the following data categories: empirical behaviour data, survey data on elicited behaviour in hypothetical situations, simulated data, secondary data from the literature, and expert knowledge.

2.3 Setup of the Research

Given the variety of types of intelligence (Section 2.2.1), alternative ways of implementing it in ABMs (Section 2.2.2) and reliance on data (Section 2.2.3) I have defined three corresponding categories of interest for this review. To understand the status of these different approaches used in the reviewed articles, I report on the trends in the current modelling practice along these categories for spatial and non-spatial ABMs. Given these aspects of intelligence, an overview is provided in Figure 2-3.

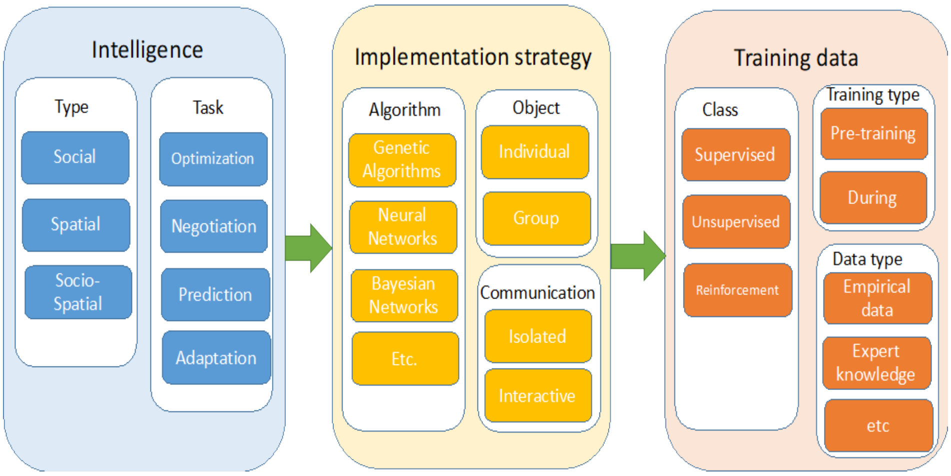


Figure 2-3: A sequence of steps when selecting an ML algorithm for 'agents' brain' in ABMs for various socio-economic and spatial application domains

2.4 Results

Despite an avalanche of ABM papers in the recent years, still only a small group of ABM developers apply ML algorithms to enhance agents' decision-making. This number is even smaller amongst developers of spatial ABMs, while a lot of models focusing on the dynamics of SES are geographically explicit (Figure 2-4). Unfortunately, developers of ABMs often do not provide any explicit motivation for their choices related to the use of ML algorithms. It inhibits learning about the benefits of AI methods and diffusion of advanced practicing within the ABM community.

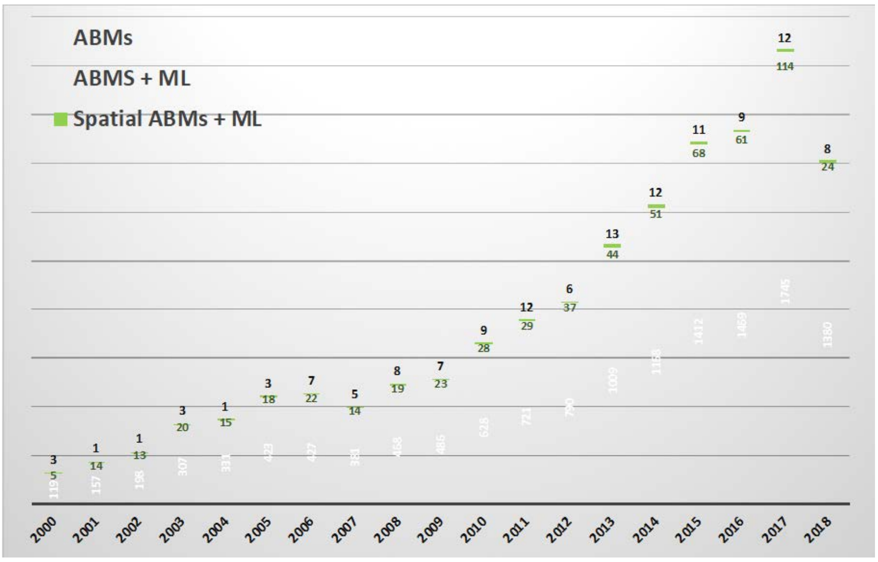


Figure 2-4: Current practice of using ML in ABMs for agents' intelligence (2000 - 2018)

2.4.1 Types of Intelligence

The set of 137 published articles on ABMs of SES using MLs is reviewed differentiating between the type of intelligence implemented: social, spatial or combined social-spatial intelligence. Here I am interested in whether intelligence varies across tasks performed and between spatial and non-spatial ABMs (Table 2-1).

Table 2-1: A frequency of occurrence of social, spatial and combined social-spatial intelligence across agents' tasks in reviewed ABMs (in absolute numbers and in percent from total in the spatial or non-spatial ABMs sub-group).

Spatial ABMs					
Task / Intelligence	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Social	5 (7%)	2 (3%)	3 (4%)	4 (5%)	*14(18%)*
Spatial	23 (30%)		14 (18%)	12 (16%)	*49 (66%)*
Combined	4 (5%)	3 (4%)	5 (7%)	1 (1%)	*13 (17%)*
Total	*32 (42%)*	*5 (7%)*	*22 (29%)*	*18 (23%)*	*77 (100%)*
Non-Spatial ABMs					
Task / Intelligence	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Social	15 (27%)	5 (8%)	11 (18%)	19 (32%)	*50 (83%)*
Spatial	1 (2%)				*1 (2%)*
Combined	4 (7%)		3 (5%)	2 (3%)	*9 (15%)*
Total	*20 (33%)*	*5 (8%)*	*14 (24%)*	*21 (35%)*	*60 (100%)*

What we notice when evaluating the table is that in spatial ABMs, intelligence is used for social learning, spatial learning, and combined socio-spatial learning. This confirms our hypothesis that although models are spatial and agents have intelligent behaviour, this does not automatically mean that this behaviour is also spatially intelligent. One would expect that in the non-spatial group of papers, the cells for spatial and mixed intelligence in Table 2-1 would remain empty. However, it is possible that although the model is not spatial, spatial intelligence is being implemented such as the examples of (Ozik et al. 2018; Rekik et al. 2014; Salle 2015; Sukhbaatar et al. 2016).

It also does not come as a surprise that the number of ABMs combining social and spatial intelligence is low. In only 16% of the cases mixed intelligence was used. For example, ML is used for estimating next state transition probabilities of the environment based on shared learning experience of agents (Barrett et al., 2013); and drives agents negotiations in a land renting auction (Balmann and Happe, 2001). In addition, learning algorithms are used to negotiate over the selected locations to inhabit (Bone et al., 2011).

Forty-two percent (42%) of the spatial ABMs implement ML for optimization. However, in non-spatial ABMs, optimization appears to be only 33% of reviewed articles, indicating a need for more behaviourally rich decisions employed in spatial ABMs beyond the optimization. Examples of using ML algorithms to support agents' optimization decisions in spatial ABMs vary from seeking a land-use allocation that scores as the best on multiple social and environmental criteria (Manson, 2005); to iterative optimization of household travel schedules (Meister et al., 2005); to optimize the performance of battle agents (Lim et al., 2005); or to search for an optimal distribution of various branches of a clinical organization across a country to minimize processing time (Asadi et al., 2009).

In non-spatial ABMs, hybrid approaches of ML algorithms are used for more effective and precise optimization in an electrical power market (Reddy and Veloso, 2011). In addition, ML used to mimic human decisions on performing actions relevant to define exposures stressors (Brandon et al., 2018) and to update of agents belief in opinion dynamic model (Sobkowicz, 2017).

Models, in which learning algorithms are used for negotiation, represent the smallest group in our sample (just about 7%, see Table 2-1) and spatial and non-spatial models score similarly. In spatial ABMs, ML algorithms drive agents negotiations in learning opponent's preferences at an earlier stage in a negotiation (Pooyandeh and Marceau, 2014a) and to reduce prices in land markets (Shen et al., 2016). In non-spatial ABMs, ML algorithms are used to generate proposals with the absence of complete information (Gwak and Sim, 2011); to drive agents negotiations in e-commence (Choi et al., 2001); and to define transaction prices among firms in a supply chain (Russ and Walz, 2009) and game theory (Kattan et al., 2013).

ABMs using ML algorithms for prediction tasks constitute 29% in spatial models and 24% in non-spatial examples (Table 2-1). In spatial ABMs, ML learning used for evaluating agents predict the location of exits in an evacuation ABM (Hajibabai et al., 2007); dynamic response functions such as environmental attributes are estimated based on socio-economic distributions at spatial large scales (Xu et al., 2013); or forecasting the perceptual reasoning of land-user agents (Sun & Müller, 2013). Kaya & Alhajj (2005) augment their hunter agents with a hybrid of ML algorithms to predict actions such as next location of other hunters following the same prey. For non-spatial ABMs, agents use ML algorithms to forecast future stock prices and dividends (Rekik et al., 2014); to make predictions on the monetary policy of banks (Salle, 2015). In addition, ML is implemented to support agents' decision such as forecasting energy demand (Costa et al., 2008).

Adaptation is the second most popular task modelled with ML algorithms among the reviewed ABM papers (23% of the spatial and 35% of the non-spatial papers, Table 2-1). However, this is mostly due to the high use in non-spatial models. An example of spatial ABMs is of Alexandridis and Pijanowski (2006) employ ML to support land-user agents' decisions in response to in- and out-migration. While Verstegen et al. (2010) have agents involved in spatial planning adjusted their spatial preferences based on the level of cooperation with other agents in the past using BN. Lei et al. (2005) design agents that adapt their land purchase decisions and socioeconomic attributes on the basis of ML outcomes as land market and spatial land-use patterns dynamically unfold. For non-spatial ABM ML is used to make adaptive decisions to classify suppliers on the basis of their profitability in the supply chain in order to adapt their choice to the safest supplier (Smeureanu et al., 2012); or agents use ML to adapt their decision whether to obey a social norm based on past actions and corresponding utilities (Nishizaki et al., 2008).

In general, one can conclude that optimization is the most popular task in spatial models and adaptation in non-spatial models. Furthermore, the number of models using mixed spatial and social intelligence is low. Each of the ABMs in this sample has at least one type of intelligent interactions with 48% employing ML to represent social learning, and 37% use ML to implement spatial learning. and 15% of the articles used ML to implement both social and spatial learning in their ABMs for the complete dataset.

2.4.2 Implementation Strategies

Social Intelligence in ABMs

The reviewed papers exhibited a variety of ways intelligence was implemented in an ABM. Figure 2-5 illustrates the trends in the current practice of using ML algorithms to enhance either individual or group decisions, with or without interactions with other agents. When individuals or groups have a social network or engage in interactions within their spatial neighbourhoods, information from peers is integrated in the intelligent decision making process.

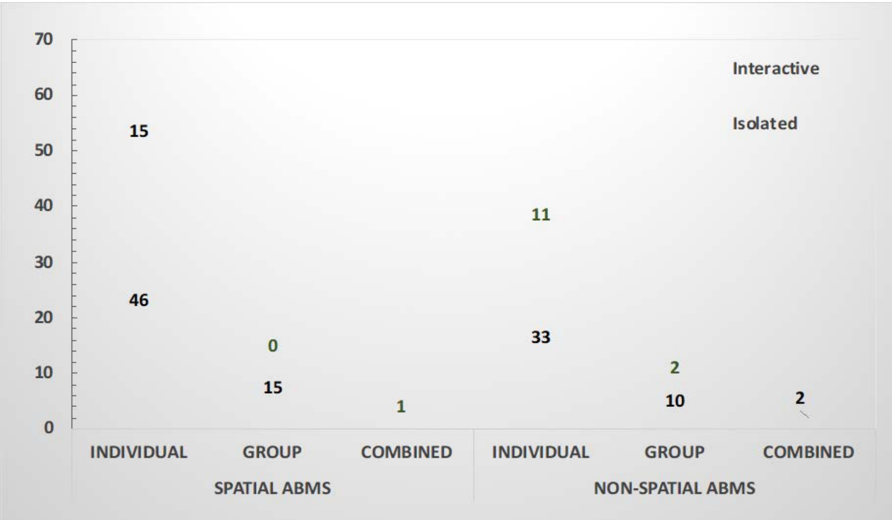


Figure 2-5: A frequency of occurrence of individual, group of combined intelligence among interactive and isolated agents in the reviewed ABMs (in absolute numbers)

Agents, during the process of learning, may be in contact with other agents, i.e. be socially interactive, or not, i.e. be socially isolated. Our sample shows that most papers use individual agents that are interactive. There are few examples of isolated learning, which is expected since social interactions are the core of ABMs. There is not much difference between spatial ABMs (61% individual – interactive) and the non-spatial models (55% individual - interactive).

Interactive individual agents engage in direct interactions with others to perform negotiations (Pooyandeh and Marceau, 2014a; Shen et al., 2016) or to compete in auctions (Graubner et al., 2011; Kellermann and Balmann, 2006; Yuan et al., 2014). Other examples of the use of MLs for individual interactive intelligent agents include regulatory interactions between governments and citizens (Cioffi-Revilla et al., 2012) and diffusion of social norms and information (Sun and Müller, 2012). Individual agents also construct and update a mental model in markets (Manahov et al., 2014).

Agents' own learning and adaptation is based on the experience of interacting with others in their groups (Djennas et al., 2012). Agents also could be updating their beliefs and skills in isolated groups (Nishizaki et al., 2008). To model trust, negotiations and communication in an auction Quteishat et al. (2009) reinforce their agents with neural networks to make individual judgments. Every three agents form a team. Thus, there are individual as well as group level goals, and interactions occur on both levels.

Table 2-2: A frequency of occurrence of interactive versus isolated learning across agents' tasks in the reviewed ABMs (in absolute numbers and in percent from total in the spatial or non-spatial ABMs sub-group).

Spatial ABMs					
Task / Interaction	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Interactive	25 (33%)	5 (7%)	16 (21%)	14 (18%)	60 (79%)
Isolated	7 (9%)		6 (8%)	4 (5%)	17 (22%)
Total	32 (42%)	5 (7%)	22 (29%)	17 (22%)	77 (100%)
Non-Spatial ABMs					
Task / Interaction	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Interactive	17 (28%)	5 (8%)	10 (17%)	14 (23%)	46 (76%)
Isolated	3 (5%)	1 (2%)	4 (7%)	6 (10%)	14 (24%)
Total	*20 (33%)*	*6 (10%)*	*14 (24%)*	*20 (33%)*	*60 (100%)*

When evaluating the tasks against the interactivity, one can see that spatial and non-spatial models have similar patterns (79% against 77%, see Table 2-2). With respect to tasks, agents in spatial ABMs rely on information from social interactions primarily when optimizing and predicting (33% and 21% Table 2-2). Non-spatial ABMs focus on using MLs that rely on inputs from social interactions when agents prsue optimization or adaptation tasks (23% and 23%, Table 2-2). For obvious reasons, intelligent negotiation tasks are never done in isolation in either spatial or non-spatial ABMs.

A Variety ML Algorithms in ABMs

With respect to the choice of an ML algorithm to use for various tasks, there are no clear preferences in the current practice (Table 2-3). One exception is Genetic algorithms, that are dominant for optimization in both spatial and non-spatial models (21% and 13% correspondingly). Interestingly, non-spatial ABMs seem to have a wider variety of ML algorithms used compared to the spatial models, especially for adaptation tasks where 'Other' comprises 17% of the sample. Here besides the common ML methods such as Bayesian or Neural networks, Genetic algorithms and Swarm intelligence, the literature reports.

Table 2-3: Use of ML algorithms for various agents' tasks in the reviewed ABMs (in absolute numbers and in percent from total in the spatial or non-spatial ABMs sub-group).

Spatial ABMs					
Task / ML Algorithms	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Bayesian networks	4 (5%)	2 (3%)	6 (8%)	4 (5%)	16 (21%)
Genetic algorithm	16 (21%)	2 (3%)	2 (3%)	1 (1%)	21 (27%)
Neural networks	3 (4%)		6 (8%)	2 (3%)	11 (14%)
Swarm intelligence	1 (1%)		2 (3%)	2 (3%)	4 (5%)
Hybrid Algorithms	1 (1%)		2 (3%)	4 (5%)	7 (9%)
Other	7 (9%)	1 (1%)	4 (5%)	3 (4%)	15 (20%)
Total	32 (42%)	5 (7%)	22 (29%)	17 (22%)	77 (100%)
Non-Spatial ABMs					
Task / ML Algorithms	**Optimization**	**Negotiation**	**Prediction**	**Adaptation**	***Total***
Bayesian networks	3 (5%)	1 (2%)	1 (2%)	3 (5%)	8 (14%)
Genetic algorithm	8 (14%)	2 (3%)	1 (2%)	1 (2%)	12 (20%)
Neural networks	3 (5%)	2 (3%)	7 (13%)	4 (7%)	16 (28%)
Swarm intelligence	2 (3%)	1 (2%)		1 (3%)	4 (7%)
Hybrid Algorithms	1 (2%)		4 (7%)	5 (8%)	10 (17%)

Other	3 (5%)		1 (2%)	6 (17%)	10 (17%)
Total	*20 (33%)*	*6 (10%)*	*14 (24%)*	*20 (33%)*	*60 (100%)*

2.4.3 Data for Training ML Algorithms

The survey of the ABM literature confirms that the availability of data is a crucial criterion for implementing MLs as drivers of agents' decisions. It influences whether a model developer can calibrate model parameters, extract and estimate missing information for agents' decision rules, or train an ML algorithm.

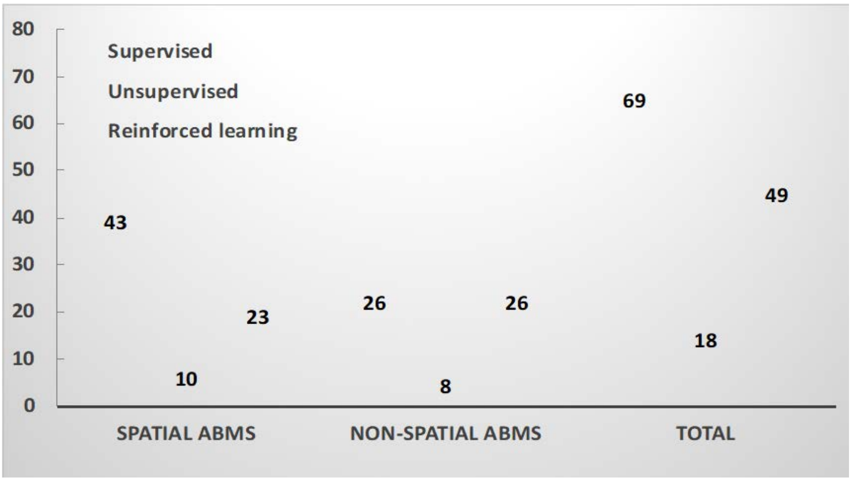


Figure 2-6: Use of supervised, unsupervised and reinforced learning (in absolute numbers)

When supervised learning is used (69 that is 51% of our sample, see Figure 2-6) data to train the ML algorithm often comes from the social or spatial environments. This is especially relevant for Bayesian networks and Genetic algorithms where qualitative data from interviews with experts and stakeholders is used to give an algorithm a direction (Pooyandeh & Marceau, 2014). Where possible, I have made a distinction between the use of empirical data for ML training in spatial ABMs and the three types of learning, though authors of the published articles do not always explicitly discuss the data sources for ML algorithms.

When considering the type of data used for training an ML algorithm, one could see that spatial ABMs use more surveys compared to non-spatial ABMs (15% versus 0%, Table 2-4). This might be due to the limited access spatial scientist have to behavioural data. Simulated data also seems to be more common in spatial ABMs (14%) compared to 8% in non-spatial

ABMs. In spatial ABMs, data used for ML can refer to spatial and behavioural data, implying that different datasets may need to be aligned. Yet, behavioural data may not always be geo-referenced hindering the merge. In non-spatial ABMs I assume the data is used for social intelligence. It would have been interesting to specifically differentiate between the two categories of intelligence: social and spatial. Yet, the published papers in our sample do not always provide enough details to permit it.

Table 2-4: Overview of learning tasks and data sources used to train ML algorithms supporting these tasks (in absolute numbers and in percent from total in the spatial or non-spatial ABMs sub-group).

Spatial ABMs					
Task / Data Types	**OPT**	**NEG**	**PRED**	**ADAP**	***Total***
Survey (hypothetical choices)	5 (8%)	0	3 (4%)	3 (4%)	*11 (15%)*
No data	11 (15%)	0	5 (7%)	6 (8%)	*22 (30%)*
Expert knowledge	1 (1%)	1 (1%)	1 (1%)	3 (4%)	*6 (8%)*
Experimental data	1 (1%)	1 (1%)	4 (5%)	0	*6 (8%)*
Simulation data	6 (8%)	0	3 (4%)	1 (1%)	*10 (14%)*
Empirical data (actual choices)	8 (11%)	2 (3%)	5 (7%)	3 (4%)	*18 (24%)*
Secondary data from the Literature	0	1 (1%)	1 (1%)	1 (1%)	*3 (4%)*
Total	*32 (42%)*	*5 (7%)*	*22 (29%)*	*18 (23%)*	*77 (100%)*
Non-Spatial ABMs					
Task / Data Types	**OPT**	**NEG**	**PRED**	**ADAP**	***Total***
Survey (hypothetical choices)	0	0	0	0	*0*
No data	7 (12%)	2 (3%)	4 (7%)	7 (12)	*20 (33%)*
Expert knowledge	2 (3%)	3 (5%)	0	4 (7%)	*9 (15%)*
Experimental data	5 (8%)	0	2	0	*7 (12%)*
Simulation data	1 (2%)	0	1 (2%)	3 (5%)	*5 (8%)*
Empirical data (actual choices)	4 (7%)	0	6 (10%)	6 (10%)	*16 (27%)*
Secondary data from the Literature	1 (2%)	0	1 (2%)	1 (2%)	*3 (5%)*
Total	*20 (33%)*	*5 (8%)*	*14 (24%)*	*21 (35%)*	*60 (100%)*

When checking for the use of different data types against the ML algorithm we see that Bayesian networks have a high percentage of survey and expert data in the spatial ABMs and that this is not used in the non-spatial models. That neural networks are used mostly with simulated data and that other algorithms use empirical data most. Examples can be presented such as Yuksel (2018) used simulation data to pre-train the hybrid ML algorithms (genetic algorithms and neural networks) to help agents to learn how to change and improve their evacuation behaviours. While Shen et al. (2016) used expert knowledge to identify initial parameters of their Bayesian networks in their land-market negotiation ABM.

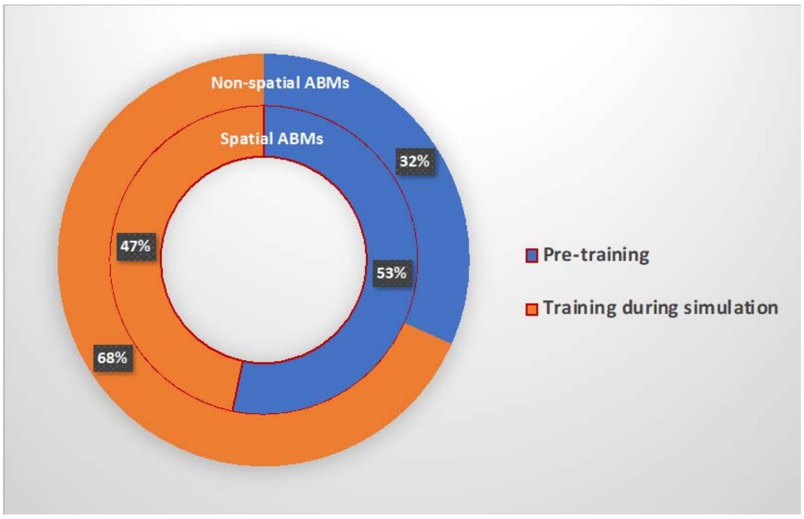


Figure 2-7: Training of ML algorithm. Here the inner ring represents spatial ABMs and the outer ring represents the non-spatial ABMs.

The training of supervised ML algorithms might be performed before integrating an ML in ABMs (average of 43%) or during the integration (average of 58%) (Figure 2-7) based on the availability/absence of data.

In case of data availability, an ML is trained before being integrated in an ABM as in (Asadi et al. 2009; Bennett and Tang 2006; Ghazi and Dugdale 2018). However, when data is not available to train in advance, the ML is trained when it is integrated to the ABM – i.e. during a simulation run. Simulation data could be used to train the ML as in (Caudell et al. 2011; Junges and Klügl 2011; Laite et al. 2016; Yuksel 2018).

Table 2-5: Practice of training ML algorithms before or during their integration in ABMs (in absolute numbers and in percent from total in the spatial or non-spatial ABMs sub-group

Spatial ABMs			
Time of Training / **ML algorithms**	**Pre-training**	**Training during**	***Total***
Bayesian networks	11 (15%)	5 (7%)	16 (21%)

Genetic algorithm	7 (9%)	14 (19%)	21 (27%)
Neural networks	6 (8%)	5 (7%)	11 (14%)
Swarm intelligence	1 (1%)	3 (4%)	4 (5%)
Hybrid Algorithms	6 (8%)	1 (1%)	7 (9%)
Other	8 (10%)	6 (8%)	14 (20%)
Total	32 (42%)	5 (7%)	77 (100%)
Non-Spatial ABMs			
Time of Training / ML algorithms	**Pre-training**	**Training during**	***Total***
Bayesian networks	4 (7%)	4 (7%)	8 (14%)
Genetic algorithm	3 (5%)	10 (17%)	13 (22%)
Neural networks	8 (13%)	8 (13%)	16 (26%)
Swarm intelligence	0	4 (7%)	4 (7%)
Hybrid Algorithms	2 (3%)	4 (7%)	6 (10%)
Other	3 (5%)	10 (17%)	13 (22%)
Total	*20 (33%)*	*40 (66%)*	*60 (100%)*

The process of training an algorithm also depends on other aspects in addition to the availability of data: the nature of the problem and the algorithm itself (Tables 2-4 and Figure 2-7). For example, Genetic algorithms, which are used for optimization, help agents to improve their behaviour on the basis of states of the environment and actions of others (H. H. Zhang et al. 2010). A fitness function might be derived before the simulation but optimizing the behaviour will take place during the simulation. This is regarded as training the algorithm during the simulation.

In addition, Neural networks algorithm are data-based algorithms that requires large datasets to be trained, therefore when there is no data available the simulation may be used as a source of data to train the algorithm. Here, both agents and their learning algorithm learn together (see for example (Caudell et al. 2011)).

2.5 Conclusions

Despite an avalanche of ABM papers in the recent years, still only a small group of ABM developers apply ML algorithms to enhance agents' decision-making. This number is even smaller amongst developers of spatial ABMs, while a lot of models focusing on the dynamics of SES are spatially explicit. Unfortunately, developers of ABMs often do not provide any explicit motivation for their choices related to the use of ML algorithms.

This makes it difficult for other developers to determine if the use of ML would be beneficial in their models.

This chapter provides a structured review of the state-of-the-art regarding the use of learning algorithms in agents' decision-making among spatial ABMs. I reviewed 137 articles with spatial and non-spatial ABMs. I approached the review by focusing on three aspects: the type of intelligence, the implementation strategies and the data availability. Where the sample of reviewed papers allowed, I identify if any patterns appear on the use of ML in ABMs for enhancing agents' cognition. In spatial ABMs, ML is most often used for optimization tasks to support agents' behaviour, primarily with Genetic algorithms. While non-spatial ABMs use ML mostly for adaptation and optimization tasks. Genetic algorithms and neural networks seem to be preferable ML methods for non-spatial ABMs. Non-spatial models also use a wider variety of algorithms, inclining that spatial ABMs lag behind in experimenting with ML for agents' intelligence.

The number of examples where ML is used for combined social and spatial intelligence is still small. This might partly be due to the non-spatial character of behavioural data (to train the ML algorithm). When the behavioural data does not include spatial aspects, and is not specific to spatial locations, this can hinder the implementation of socio-spatial intelligence in spatial ABMs. Social intelligence is correlated with adaptation tasks of non-spatial ABMs while agents used their spatial intelligence for optimizations in spatial ABMs.

With respect to the implementation strategies, it is clear that ML is most often implemented as individual-interactive learning. Hence, learning at the level of a group is less common, as is isolated learning. Spatial models are just as interactive as non-spatial models. Speaking of trends in the use of ML techniques to enhance agents' decisions in ABMs: the most popular ML algorithm is Genetic algorithms (25% of the reviewed articles) followed by Neural networks (21%) and Bayesian networks (17%) in the sample of reviewed papers.

Initially I thought that a slow progress in the implementation of ML in ABMs might be the lack of social data to train the algorithms. However, most models in our sample used supervised learning (56% of the spatial models and 43% in non-spatial), indicating that this does not seem to be a limitation. Not all data is empirical data, though. For example, Bayesian networks make a lot of use of survey data and expert knowledge.

Simulated data is commonly used for Neural networks, which typically rely on large datasets.

Despite the fact that I found many successful implementations of ML in ABMs this does not mean that that they outscore rule-based implementations. Given the large number of ABM papers in the recent years, it seems that only a small group of ABM developers apply AI techniques to enhance agents' decision-making. This relatively low number can be driven by a desire of every modeller to develop simulations with minimum complexity. ML learning is likely employed only when its benefits – which are vast – prevail. For example, ML learning is very promising to support adaptive agents' behaviour in ABMs as it offers an effective and elegant way of modelling various agents' tasks. Since ABMs are sometimes criticized for making ad-hoc assumptions about micro-rules of behaviour and interactions, ABMs' micro-foundations may benefit from the knowledge base ML offers in this respect.

For the future, it would be interesting to see if there is a more extensive use of ML algorithms in spatial ABMs for processing of input data or model calibration. It is likely that the inclusion of empirical data plays an important role in both pre-processing and calibration of models. In addition, for various tasks and domains, more than one ML algorithm seems to be applicable to support a specific agent's decision. Via a systematic test of several ML algorithms in the same ABM, their impact on the model results and performance can be measured. Moreover, the implementation of ML within an ABM is often insufficiently documented (e.g., no specifications on whether an AI method uses supervised, unsupervised, and reward-based learning, etc.). Transparency in research would require model developers to be very explicit, not only in the description of an ABM itself, but also in respect to an embedded learning algorithm. Perhaps, the ODD protocol (Grimm et al. 2010) can better emphasize the architecture and the implementation of specific ML algorithms elements when discussing agents' learning.

Many new developments seem to occur in the way we collect behavioural data. Serious gaming and Virtual Reality experiments, are promising ways to collect ML datasets. However, it will take a while until we see this in published papers. An advantage of these types of data will be that they are spatially explicit (can be linked to behaviour at a certain location).

Finally, in many cases, ABM toolkits do not provide libraries that contains built-in ML functions that can save developers efforts and time. Therefore, developers should have excellent programming skills to code and implement ML algorithms in ABMs. However, developers of socio-economic and spatial ABMs are not always computer scientists while the implementation of ML algorithm often requires specific computer science knowledge to implement them correctly. Collaboration between the two communities is likely to mitigate the problem.

Chapter 3: Intelligent Judgements over Health Risks in a Spatial Agent-Based Model[1]

[1] This chapter is based on paper (paper 1) which has the same title. The paper is authored by Shaheen A. Abdulkareem (the main writer of the paper and author of this dissertation), Dr Ellen-Wien Augustijn based at the University of Twente, Dr Yaseen T. Mustafa based at the University of Zakho and Professor Tatiana Filatova based at the University of Twente. The paper has been published in International Journal of Health Geographics, Vol 17 No. 8, March 2018, https://doi.org/10.1186/s12942-018-0128-x.

3.1 Background

Globally, millions of individuals are regularly exposed to deadly infectious diseases. For example, news of the Zika virus outbreak was one of the main news stories of the past two years. Perceiving disease risk motivates people to adapt their behaviour toward a safer and more protective lifestyle. Indeed, risk perception (RP) is an integral part of the decision-making process under uncertainty and can be understood as an individual's evaluation of risk in a particular situation. This evaluation includes individual assessments of how severe and controllable a particular situation is. The reliability and effectiveness of any risk evaluation by an individual is based on the risk information available (Pablo et al., 1996). Accordingly, the availability of risk information impacts the perception of a decision problem, the evaluation of available options, and of any risk-coping decisions (Williams and Noyes, 2007). A number of factors related to the design of a risk message influence risk perception: the message, being the source of information (other people, and/or the environment), and the adaptive behaviour in response to that message. These factors need to be considered in order to design effective risk communication strategies and to positively influence health-related decisions (Sitkin and Weingart, 1995).

Numerous examples of human behaviour influencing the spread of infectious diseases are available (Bauch et al., 2013). Namely, Manfredi and D'Onofrio (2013) refer to human behaviour as to the neglected layer of complexity in current epidemiological models (Manfredi and D'Onofrio, 2013). In the latter, the response to risk factors is fixed, and no effect of previous exposure – or learning – is incorporated in most models. This implies that a disease model may underestimate the effectiveness of preventive measures. This can lead to a higher scope of contagion compared to a real situation, consequently leading to an overestimation of the prevalence of disease cases. Instead, employing learning techniques to capture dynamics in RP and corresponding protective behaviour can mimic the complex process of how human beings act upon encountering risk.

Behavioural science has developed various theories to explain, measure, and assess RP. Protection motivation theory (PMT) is one of the dominant approaches in this domain, and has already been applied to the study of health-protective behaviour (Bassett and Ginis, 2011). Originally proposed by Rogers (Rogers, 1983), PMT has been actively applied in health research to study cognitive processes and predict health-related behaviour. Behavioural aspects of decision-making under risk are active with ABMs (Filatova et al., 2011; Haer et al., 2017; van Duinen et al., 2016) outside disease of research, and often without facilitating learning. In fact, ABMs are instrumental in exploring and implementing RP, such as the risk of disease diffusion. Disease ABMs have become significantly sophisticated by integrating rich GIS landscapes with detailed human activities (e.g. mobility and social networks) as well as multi-stage epidemiology models such as the SEIR (Susceptible – Exposed – Infected – Recovered) model. Moreover, ABMs are able to incorporate the social behaviour of individual agents as well as the dynamics of the spatial environment, which also plays an important role in the disease diffusion process. Various infectious diseases have been modelled using ABMs (Crooks and Hailegiorgis, 2014; Gotteland et al., 2014; Perez and Dragicevic, 2009). Wise (2014) provides an extensive review of disease and disaster ABMs (Wise, 2014). Although ABMs are technically suitable for incorporating agents with higher levels of intelligence, this is rarely implemented in disease models. For example, RP typically enters decision-making models either as a variable affecting a decision-making process or as a step within a rule-based procedure (Asgari et al., 2016; Bieberstein, 2014; Kim et al., 2008; Seidl et al., 2016).

In rule–based implementations, behaviour is fixed, meaning that decision-making functions and algorithms remain unchanged. While agents react to changes in their spatial and social environment, they neither adapt their rules in response nor intelligently learn from previous experiences. This is unrealistic, as human beings adjust their behaviour strongly when they perceive a serious risk, which can potentially lead to disease models overestimating risk. Intelligence helps agents assess risks and potentially adapt their behaviour – i.e. learn to reduce or avoid health risks – based on changes in RP.

To test the impact of adaptive RP in human decision-making, we implement PMT in a spatial disease ABM. Namely, we extend the base disease model developed by Augustijn et al. (2016) to the behavioural aspects of decision-making in a risky situation using machine learning (ML) techniques (Augustijn et al., 2016). The spatial agent-based disease model – CABM – is applied to study the spread of cholera in Kumasi, Ghana. In this article, we use Bayesian Networks (BNs) as the learning method to design intelligent agents behaving according to PMT and making decisions on how to cope with cholera in a rich spatial environment. We systematically test the impact of intelligent behaviour on disease spread through a series of simulation experiments: using CABM with zero-intelligent agents, agents enhanced with ML for updating their RP, and agents enhanced with ML for RP and coping appraisal behaviour dynamics. BNs replace ad hoc rule-based schemes for uncertainty reasoning due to their capability for bi-directional inference combined with a strict probabilistic foundation (Heckerman, 1995). They are capable of sensing and reacting to a stochastic environment. In addition, BNs have the ability to constantly adjust to simulate the dynamics of agents' beliefs. Therefore, BNs have been implemented in ABMs as the agents' cognitive model for different purposes, including negotiation (Zhang et al., 2008), prediction (Kocabas and Dragicevic, 2013), and adaptation (Lei, 2005).

3.2 Methods

We start by briefly describing the base ABM and then focus closely on the describing the learning algorithms and their stepwise implementation to support agents' intelligence.

3.2.1 The base Cholera Model and Zero-Intelligence Agents (ZI)

The CABM is used as a testbed for this research. The model was developed to test if runoff water from open dumpsites could have been the diffusion mechanism behind the 2005 cholera outbreak in Kumasi Ghana. This ABM simulates both a hyper-infectious and a low-infectious diffusion route of cholera. It is a spatial ABM with a rich representation of GIS data, including elevation, the location of residential areas, river hydrology, and the location of dumpsites in the study area (Figure 3-1).

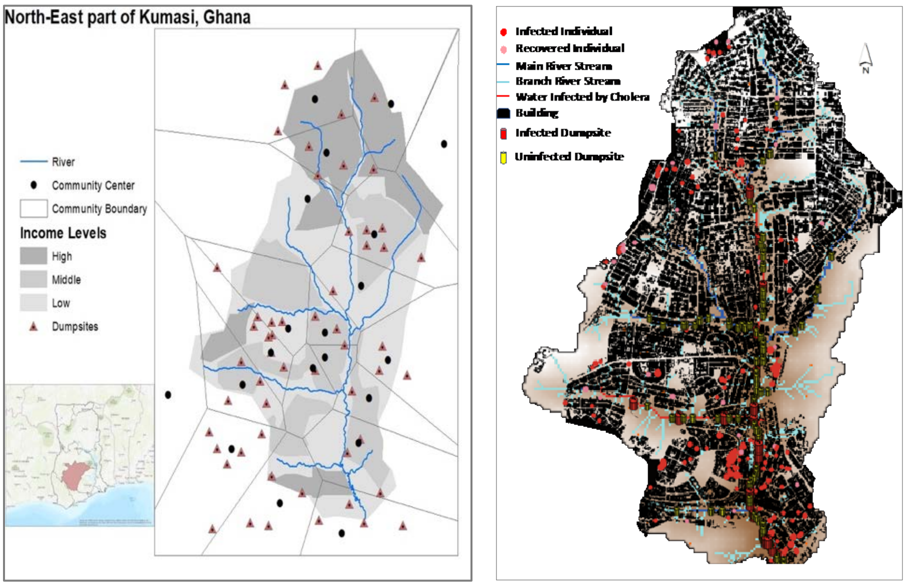


Figure 3-1: Left hand: Study area with community boundaries. We used Thiessen polygons to define the boundaries of communities that were unknown or ill defined. Right hand: Spatial spread of cholera in a typical simulation

The CABM contains three types of agents: households, individuals, and rain particles (Figure 3-2). The model contains three sub-models: a hydrological model, an activity model, and a disease model. The hydrological model moves rain particles over the area. Following heavy rainfall, runoff water can become infected with cholera bacteria when passing through dumpsites, thereby transporting cholera bacteria into the river. Via the activity model, household agents will determine the type of water they should consume (tap water, bottled water, or river water).

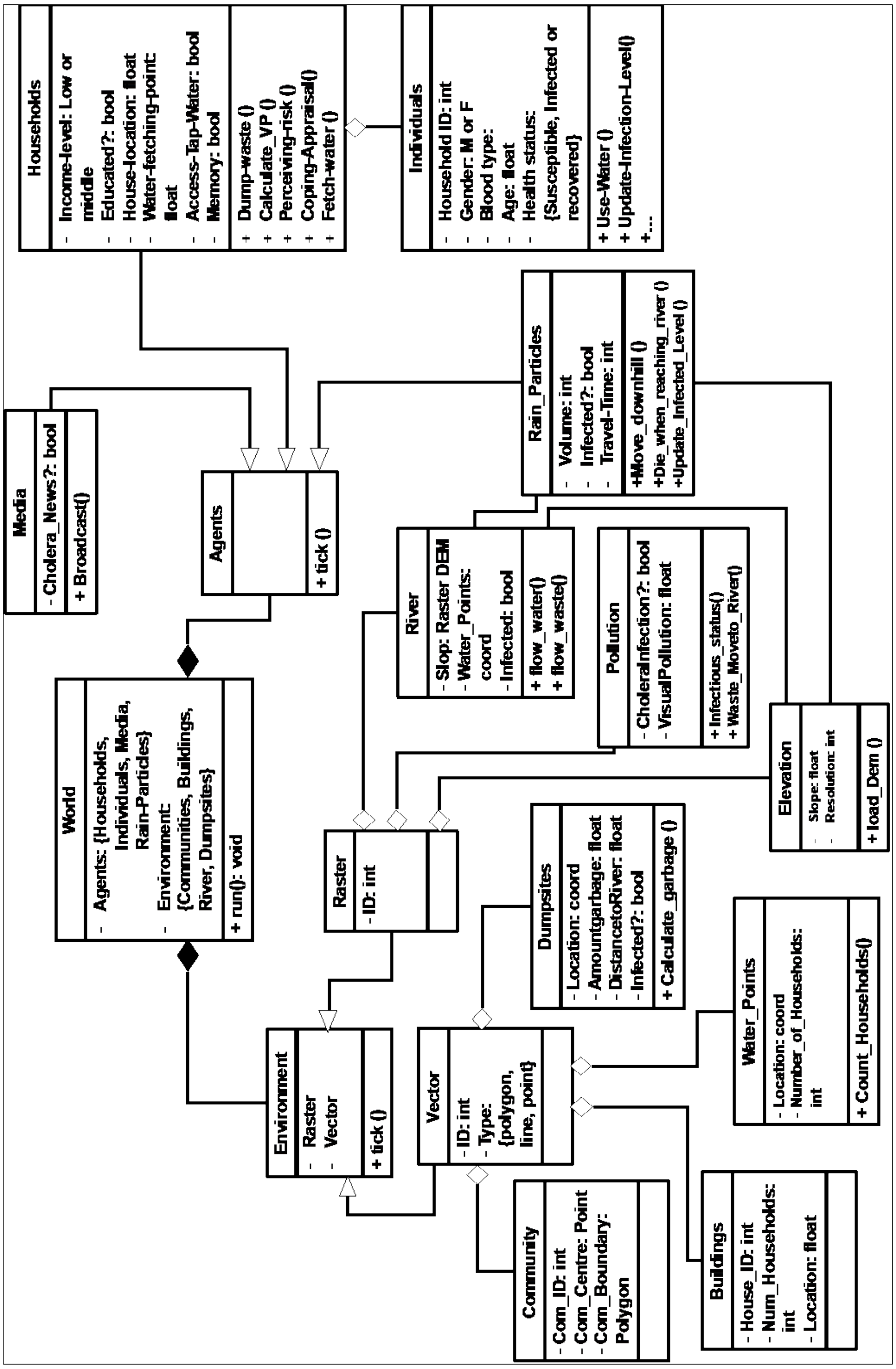


Figure 3-2: The UML diagram of CABM

Household agents use river water when tap water is unavailable. When a household agent uses river water, the model will choose the river location closest to agent's home and determine if the water at this location is infected. Individuals can become infected by using water polluted with cholera and will subsequently shed hyper-infectious materials that will be dumped by the household to the nearest open refuse dumpsite. This increases the infection level of this dumpsite and the probability of rain particles becoming infected. Finally, the disease model will determine the progression of the disease in the individual and the moment of recovery. However, this CABM does not include cholera RP and behavioural change (the selection of another water source) of agents – i.e. the household agents have no intelligence. They follow the same behaviour and activities during the entire simulation period. The time step of the model is 1 hour, with a time horizon of 90 days.

3.2.2 Intelligent Agents: How Do Intelligent Households Make Decisions?

Protection motivation theory (PMT)

PMT is used as the theoretical framework of this paper. PMT considers that, when facing a risky situation, a person goes through two steps: "threat appraisal" and "coping appraisal" (Figure 3-3). Threat appraisal in PMT is the stage at which perceptions of risk are formed. Here, a household agent assesses the probability and consequences of a risky event occurring – i.e. perceived probability and perceived severity, which in fact constitutes the agents RP. Therefore, in the proceeding sections of this paper we refer to threat appraisal as the stage at which RP is developed. The perception of severity enables households to judge how seriously the consequences could be, should they face a threat. Perceived probability measures how susceptible a person is to a given threat. The purpose of this stage is to detect whether a risk is at an acceptable level or not.

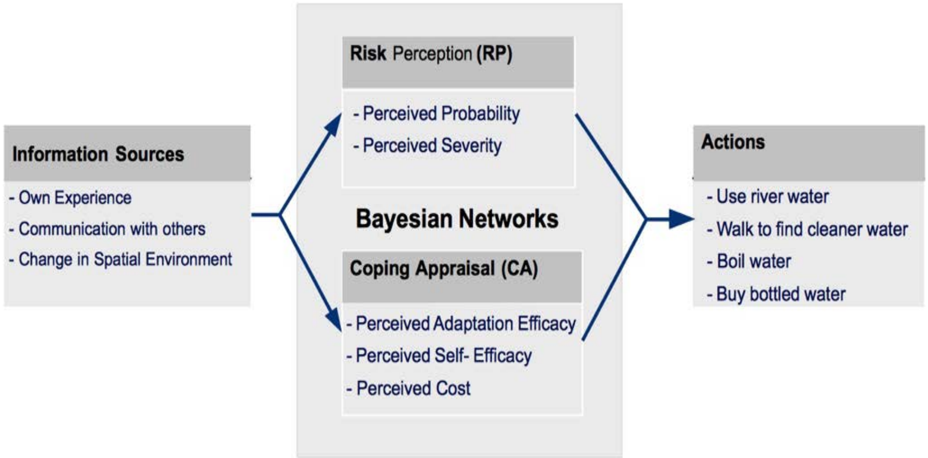


Figure 3-3: Cognitive Process of Protection Motivation Theory

When RP is sufficiently high, household agents consider a number of protective behaviours by passing through the coping appraisal stage. The coping stage consists of two main parts: adaptation–efficacy and self–efficacy. Adaptation-efficacy measures the effectiveness of protective behaviour against a harmful situation – i.e. the beliefs of a person that the recommended behaviour will protect them. Instead, self-efficacy measures the ability of a person to perform the recommended behaviour. In addition, the person must evaluate the cost of coping with the threat. Hence, at this stage, households consider the psychological, physical, and economic consequences of adapting to a particular threat.

CABM – Intelligent Agents

In the intelligent version, the CABM is modified to simulate the RP (threat appraisal) and coping appraisal (CA) processes of household agents – i.e. including the learning technique to create intelligent agents. For this purpose, one extra agent (media) is added to the model (Figure 3-2).

The state variable of the **Household agent** is the type of water they consume, and the infection level of this water. The household agent is responsible for the collection of water, and all household members will use this water for their daily consumption. Learning takes place at the level of the household, as it is directly related to the water source that the household selects. To facilitate this learning, we added memory and education level to the attributes of the household agent.

The state variable of the **Individual agent** is their health status. Individual agents can be susceptible to, infected with, or recovered from cholera.

Some studies have indicated that medical alerts do not have the impact of encouraging people to physically search for medical advice during epidemics (Frias-Martinez et al., 2012). However, information received from different media channels can prevent an epidemic from spreading (Funk et al., 2009). Therefore, **Media** is a new agent that has been introduced to broadcast information about the epidemic in this model. The state variable of the media agent is its activation level, which determines if the media agent has started to broadcast about the epidemic.

The state variable of the **Rain particles** agent is the infection level. While flowing over the terrain, rain particles can acquire the infection (from infected dumpsites) and carry it to the nearest river or tributary.

The processes included in the original model were flow of rain particles, household fetching water, and households dumping their waste. These processes remain unchanged in the version of the model used in the present research. However, in this version of the cholera model, we added the following processes:

1- Activation of the media agent;
2- Clearance of the dumpsites;
3- Calculations of the visual pollution (VP) level;
4- Risk perception
5- Coping appraisal (CA)

Activation of the media agent:

The media agent is deactivated in the beginning of the simulation. It is activated when the number of days exceeds a threshold value (22 days). After activation, the media agent will broadcast news about the cholera epidemic once a day, which all household agents in the simulation will receive. Once the broadcasting has begun, it will continue throughout the remaining part of the simulation. Media information is used in the risk assessment.

Clearance of the dumpsites:

In the original model, dumpsites could be infected with cholera, and when the decay function was activated, this infection would gradually disappear over time. We also introduce the fact that garbage will be removed from dumpsites. This has two separate effects: it will influence the infection and will also have an impact on the visual pollution level.

Clearance of dumpsites will occur randomly. In Kumasi, 85% of household waste is collected by the municipality from the dumpsites twice per week (Asase et al., 2009). Therefore, in this model, a random 85% of simulated dumpsites are discharged twice per week.

Calculation of the Visual Pollution level (VP):

Household agents fetch water from the nearest water collection point on the river, either because they do not have access to tap water, or because their tap water has stopped working due to heavy rain. Open refuse dumpsites are located at varying distances along the river. It is common in Kumasi to observe waste dumps located on riverbanks or in a river's path (Danquah et al., 2011). In the simulation, risk will be assessed based on a combination of factors, including the visual pollution (VP) level of the water collection points. The visual pollution level is calculated based on the combined link order and the number of open refuse dumpsites located within a specific distance from the river. VP is calculated based on the following equation:

$$\mathrm{f(VP)} = \sum_{\mathrm{i=1}}^{\mathrm{N}} \frac{\mathrm{x}_i \, \mathrm{g}_i}{d_i} \qquad 3\text{-}1$$

where $\mathbf{N}$ is the number of dumpsites around the water collection points; $\mathbf{x}$ is the number of households who use the dumpsite; $\mathbf{g}$ is the amount of garbage produced by each household; $\boldsymbol{d}$ is the distance from the dumpsites to the water collection point; and $\mathbf{i}$ represents all dumpsites in $\mathbf{N}$ (either cleared or not). Although the number of dumpsites is fixed throughout the simulation, the amount of garbage remains static, and the number of households will also remain static over a simulation run, while the visual pollution level is dynamic. This dynamic nature is due to the random selection of dumpsites that will be cleared over a simulation run.

Learning – Implementation of Agent's Cognitive Model

The PMT drives the agents' cognitive model. The information sources and the two stages of PMT are illustrated in Figure 3-4. In this model, we used two BNs – BN1 to model the RP, and BN2 to model the CA.

Implementation of Risk perception (RP):

At each time step, the household agent will perceive the risk of cholera infection using the BNs. The following factors are included in the RP: the number of infected individuals in the household, visual pollution level at the water collection point, communication with other agents, media attention, and the memory of the household agent. Together, these factors and the agents' social interactions help agents to assess risk and thus select what decision they could make among several options.

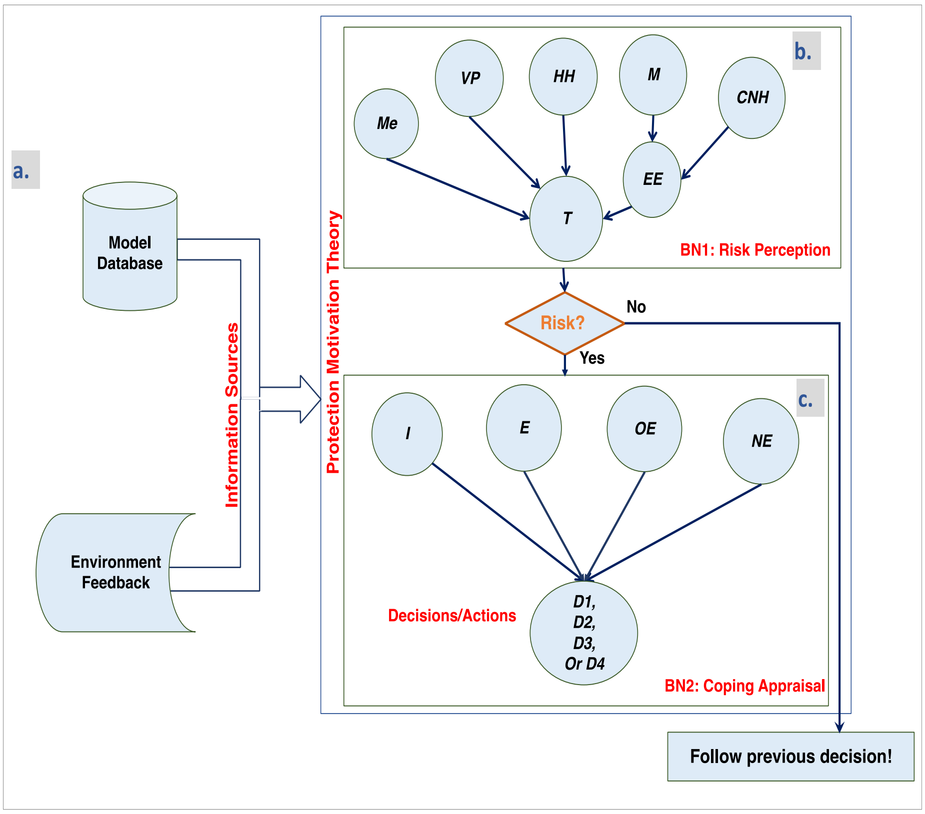


Figure 3-4: Implementation of PMT: a. Information sources; b. BN1 (RP); BN2 (CA)

Communication with other agents (social networks)

Household agents are assumed to have a total awareness of the cholera cases occurring within their neighbours' subset. A neighbour is defined as a household agent, sharing the same water collection point and living in the same community. Interaction with neighbours enables agents to perceive the infection level of the water collection point they use. In addition, household contacts help agents to gain information on adaptive decisions their neighbours took and how effective these decisions were.

No data is available on how many daily contacts Kumasi residents have. However, in a recent study by Melegaro et al. (2017), they conducted a survey of daily contacts in Manicaland, Zimbabwe and reported 10.8 contacts per person/day, including contact with household members (Melegaro et al., 2017). If we consider this rate for our study and exclude the number of household members (average of 3.9), then approximately seven contacts with neighbours per day should be applied. These seven neighbours are chosen randomly every day from the agent's community.

Memory

Agents use their memory to record the RP they experienced during the previous day (the last day they fetched water) and how preventable their last decision was. The feedback of the last decision made is measured by "positive experience" if no illness was observed in the household, otherwise it is a "negative experience".

BN1: Risk Perception

BN1 was designed to represent the RP of PMT in such a way that it answers the question "is there risk?" In the case of a risk being present, agents will proceed to the CA.

Agents with a low- or medium- income level that do not have access to safe water will fetch water from the river. Therefore, they must evaluate the risk of becoming ill with cholera using BN1. In our case, BN1 is formed by the cause-and-effect concept. To design BN1, we derive five nodes from the information sources to evaluate RP (Figure 3-4.b). These nodes include: memory (Me), visual pollution (VP), household health status (HH), media (M) and communication with neighbour households (CNH). Media and communication with neighbour households are combined into "Epidemic Evidence" (EE). EE is a binary measure that indicates to the agents if there are cholera cases outside their own households. The evaluation of infected cases differs by agent due to variations in household income and size, in the health status of different households, in their locations within the city that define VP and their selection of neighbours with whom they communicate, and in the experiences stored in their individual memories.

The reasoning and uncertainty of RP is governed by rules that can be formalized using formula (3-2). For example, we include the states {yes, no} for memory (Me), {yes, no} for threat (T), then the formula of connecting these two variables accordingly was designed as:

$$\mathrm{P}\left(\boldsymbol{T}_{\{yes,no\}} \middle| \boldsymbol{Me}_{\{yes,no\}}\right) = \frac{\mathrm{P}\left(\boldsymbol{Me}_{\{yes,no\}} \middle| \boldsymbol{T}_{\{yes,no\}}\right) \mathrm{P}\left(\boldsymbol{T}_{\{yes,no\}}\right)}{\mathrm{P}\,\boldsymbol{Me}_{\{yes,no\}}} \qquad 3\text{-}2$$

in such a way that each state of Threat is examined with each state of memory.

This was also applicable for computing the probability (P) of threat based on visual pollution (VP) and household health status (HH), as both variables have the states {high, low} and {yes, no}, respectively.

We evaluated the epidemic evidence (EE) that agents record via their communication with neighbour households (CNH) and the media (M) agent.

According to Bayesian rules, the prior probabilities of the nodes should be specified in order to gain the posterior probabilities. These prior probabilities represent the integral part of human reasoning regarding certainty. The prior probabilities will be updated/changed for each agent on the basis of information being passed by each agent to BNs. In BNs, this is called evidence.

The final formula for the threat node (T) that derives the conditional probability table (CPT) will depend on memory (Me), visual pollution (VP), the health status of household (HH), and the severity evidence of epidemic (EE):

$$\boldsymbol{P}(\boldsymbol{T}|\boldsymbol{Me},\boldsymbol{VP},\boldsymbol{HH},\boldsymbol{EE}) = \frac{\boldsymbol{P}(\boldsymbol{Me},\boldsymbol{VP},\boldsymbol{HH},\boldsymbol{EE}|\boldsymbol{T})\,\boldsymbol{P}(\boldsymbol{T})}{\boldsymbol{P}(\boldsymbol{Me},\boldsymbol{VP},\boldsymbol{HH},\boldsymbol{EE})} \qquad 3\text{-}3$$

Intelligent agents in the CABM learn to predict health risks with the help of BN1 (Eq. 3-2). In BN1, the memory node feeds the network with previous information on agents' own RP. Agents learn to revise their beliefs by absorbing other factors from their environment that are updated during the simulation, e.g. currently observed visual pollution, number of illnesses among neighbours, etc. (Eqs. 3-2 and 3-3). Agents conclude the causal relationship between nodes in the BN1 by inference. The output of BN1 would be the probability of high or low risk perception. We consider the agent to be at risk if the probability of RP is greater than or equal to 0.5.

Coping Appraisal (CA):

BN2 was designed to represent the coping appraisal of PMT in such a way that it answers the question “what to do?” In the case of perceiving risk, an agent may either: use the polluted water anyway, walk (find another location to fetch water), boil the fetched water (to increase safety), or purchase bottled water. To select one of these four decisions, a number of variables (nodes) affecting this process were identified and used. These variables include: the income level of the agents (medium or low); their education level (educated or uneducated); and the feedback of their previous and their neighbours’ previous action (positive or negative). Agents cannot learn from their own experience unless they have a feedback on their previous actions (Mitchell, 1997). Together, all of these dynamics guide the decision-making process.

BN2: coping appraisal

BN2 represents the structure of the CA (Figure 3-4.c). The probability of which decision might be chosen by the agent is computed via BN2. The perceived adaptation efficacy will differ per decision. Walking to another location to collect water has a lower efficacy compared to boiling the water, and this has a lower efficacy compared to buying bottled water. Also, perceived self-efficacy (i.e. perceived effectiveness enabling an agent to perform the preventive measure) is varied for each decision. In addition, the perceived costs of the options differ, as river water is free of cost, boiling water has a price tag, and so does the purchase of bottled water. Here, the agents’ income level determines which decision is more likely to be taken.

The formula of BN2 for computing the CPT of a decision can be expressed as:

$$\boldsymbol{P}(\boldsymbol{D}|\boldsymbol{I},\boldsymbol{E},\boldsymbol{OE},\boldsymbol{NE}) = \frac{\boldsymbol{P}(\boldsymbol{I},\boldsymbol{E},\boldsymbol{OE},\boldsymbol{NE}|\boldsymbol{D})\,\boldsymbol{P}(\boldsymbol{D})}{\boldsymbol{P}(\boldsymbol{I},\boldsymbol{E},\boldsymbol{OE},\boldsymbol{NE})} \qquad \text{3-4}$$

where D stands for decision, which can take the form (state) of ‘use water from the same fetching point’ (D1), ‘walk to another fetching point’ (D2), ‘boil water’ (D3), and ‘buy water’ (D4); I denotes an income level, what can be middle or low; E is the education level (educated or not); OE is an agent’s own experience with cholera, which can be either positive (no household member is ill) or negative (at least one household member is ill); and NE is the neighbour’s experience with cholera (anyone ill (negative) or not (positive)).

3.2.3 Model Parametrization

The probability values of both networks' variables are derived from the existing literature and census data for Kumasi. The census data of Kumasi, Ghana includes income distribution. The distribution of the three levels is 19% (low), 52% (medium), and 29% (high). However, we exclude high level incomes since they will not use river water. Therefore, by scaling both medium and low-income levels, we get 73% and 27%, respectively (which represents 71% of the number of simulated households). Additionally, 14% of low and middle-income level households do not have access to tap water. Table 3-1 presents the additional parameters of this cholera model. Naturally, for real policy application, the quality of data regarding initial weights in BN1 (Table 3-2) and the frequency and the extent of information delivery, either via media or through the word-of-mouth across social networks, is essential. We run a sensitivity analysis of final outcomes on the initial weights of both BNs. The results indicate that the model is rather robust, with minimal impact on the final outcomes.

Table 3-1: CABM new Parameters

New Parameters	Value	Description
Literacy rate	74.1 %	(Ghana Statistical Service 2012)
Media	Activation day 22	During the 2005 outbreak, newspapers and TV channels published news about the cholera in the region after about three weeks of epidemic started (visit: Ghana News Archive).
Waste collection	85% of dumpsites	85% of waste is collected by Kumasi municipality (Asase et al., 2009). The rest remain uncollected for a week or more.
Amount of garbage	2.925 Kg/household/day	derived from literature (Miezah et al., 2015)
Number of contacts with neighbours	7 neighbours	Derived from literature (Melegaro et al., 2017)

3.3 Simulation Results

3.3.1 Experiment Setup

To answer the research questions, we have designed three experiments. We systematically vary the cognitive abilities of agents by gradually adding intelligence by means of the two BNs (Figure 3-5).

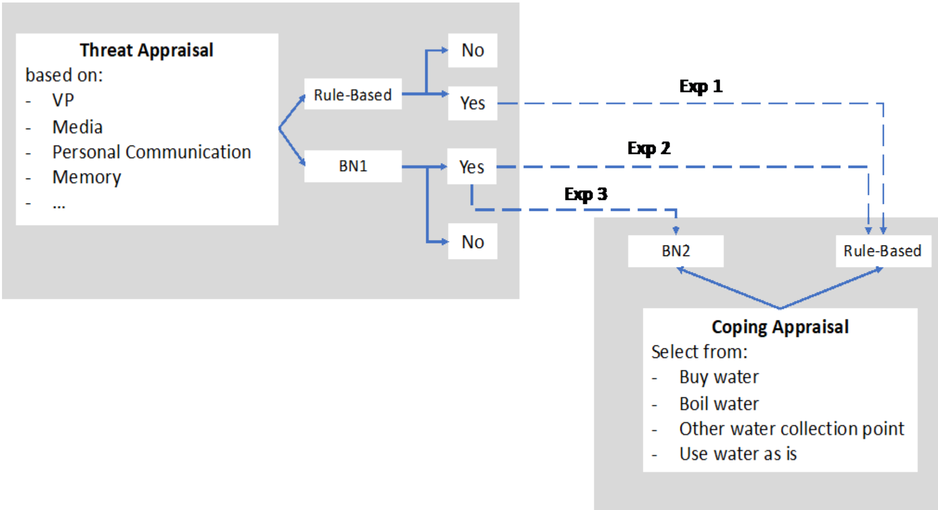


Figure 3-5: Implementation of PMT in CABM where Exp1, Exp2 and Exp3 refers to experiments 1, 2, and 3 respectively

In particular, the first experiment (Exp1) presents a benchmark case to study disease diffusion patterns in a spatial landscape with a population of zero-intelligence agents. Agents are heterogeneous in income, education, and household size but have no cognitive abilities to either perceive risk or act upon it. In the second experiment (Exp2), agents are enhanced with the BN that represents the first stage of decision-making in a risky context: the risk appraisal (BN1). As agents learn and interact with each other, the probabilities of specific factors influencing risk appraisal change. The second stage of decision-making in Exp2 is modelled in a simplistic manner by adopting a rule-based algorithm, which deterministically guides an agent to a specific action if its RP is high. Finally, the third experiment (Exp3) adopts intelligent decision making at both stages of decision making under risk: the risk appraisal (BN1) and the coping appraisal (BN2) both supported by BNs learning algorithms. Thus, if agents begin to perceive risk as an outcome of BN1, they employ BN2 to decide how to act upon it. As agents learn from their own experience and others' through interaction, the probabilities of specific actions to be chosen through BN2 evolve. All other settings among the three experiments remain static (Table 3-2). Each of the experiments is run 100 times to assure the robustness of the results.

Table 3-2: Model settings varied across the three experiments.

Model settings	**Exp1**	**Exp2**	**Exp3**
Threat appraisal	*None*	*BN1*	*BN1*

- Initial weights[2]: *Me*, *VP*, *HH*, *M*, *CNH* - Weights during a simulation - Outcome	 n.a. n.a. n.a.	 (0.1; 0.2; 0.01; 0.01; 0.2) Change as agents learn RP, (0;1)	 (0.1; 0.2; 0.01; 0.01; 0.2) Change as agents learn RP, (0;1)
Coping appraisal - Initial weights: *I*, *E*, *OE*, *NE* - Weights during a simulation - Outcome	*None* n.a. n.a. D1	*Deterministic* Rule based, **Error! Reference source not found.** Static D1-D4: fixed population share	*BN2* (0.52; 0.74; 0,9; 0,6) Change as agents learn D1-D4: adaptive, based on previous experience

We report the results per experiment in terms of several macro metrics of interest: epidemic curve, RP curve, and decision type curve. An epidemic curve is a graphical description of the number of illness cases by date during an outbreak. It illustrates the temporal trend and periods of disease incubation. A RP curve is a graphical description of a number of agents that perceive disease threat, i.e. have their RP equal to 1 in a specific time step. A decision types curve counts the number of agents following a particular decision when deciding on how to cope with cholera risk. In addition, we show several maps illustrating the spatial patterns of RP (Decisions: D1-D4).

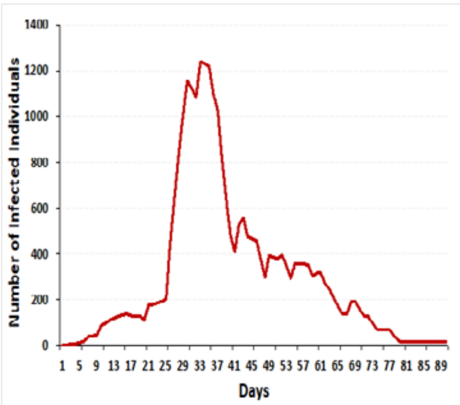
3.3.2 Disease Diffusion in a Population of ZI Agents

The temporal patterns of a cholera epidemic given a population of zero-intelligent (ZI) agents neither perceiving risk nor pursuing any protective measures is presented in Figure 3-6.a. It is evident that, even if a household member becomes ill, media broadcasts cholera being present, and some visual pollution is observed at a water fetching point, a ZI agent will still continue to collect water for its daily needs at the same

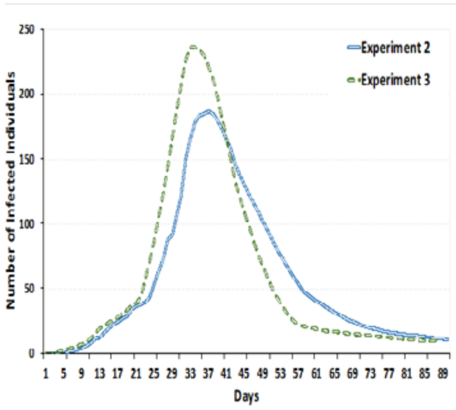
[2] To elicit the factors that may play a role in the context of a water-spread disease in a developing country as well as their relative importance we ran a survey among students. We approached the participants of the Massive Open Online Course (MOOC) on GeoHealth run at ITC (authors host institute) in Sep, 2016. Majority of the participants of this course are from developing countries. Ideally, one would survey real citizens in the case-study area. This was not possible due to the lack of funds and access to the potential respondents.

water fetching point and will use it without precautionary measures. The number of infected agents reaches a maximum between day 28 and day 40 before gradually decreasing towards the end of the epidemic. In total, 81% of the simulation population (27,000 out of 34,000 individuals) is infected with cholera in Exp1. While the ZI CABM succeeds in reproducing the qualitative pattern of this cholera epidemic, it largely overestimates the number of infected individuals. A simulation with non-adaptive ZI agents misrepresents reality, since even middle income and educated people continue to consume potentially contaminated water: 28.6, 64.7, and 6.5% in the low, middle, and high-income categories, respectively.

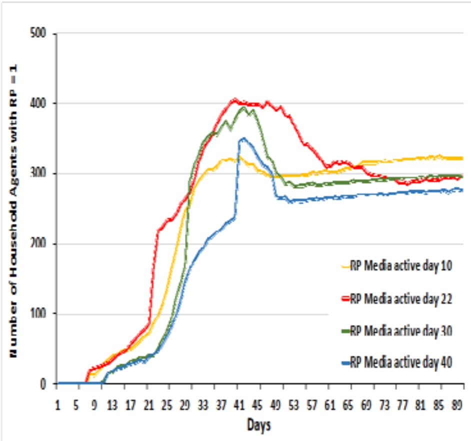
When agents have no cognitive abilities, and are not reactive, then the probability of becoming infected during a rainy period depends on the concentration of infected agents, which may dump infected waste on a dumpsite, leading to flow of cholera-infected rainwater into the river.

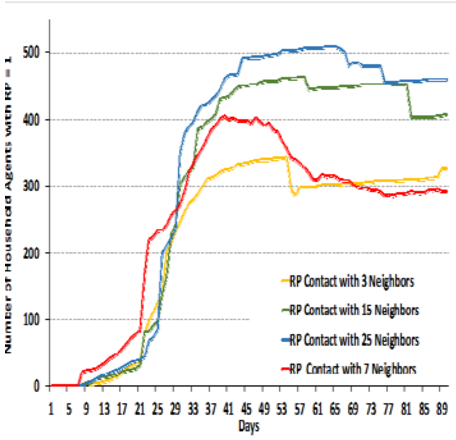


a. Epidemic Curve of Exp1 (Zero-Intelligent Population)



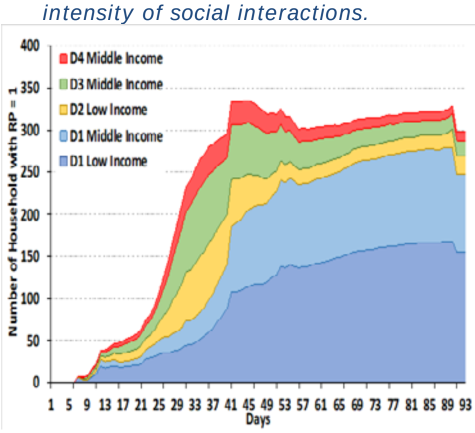
b. Epidemic Curve of Exp2 (blue) and Exp3 (green)

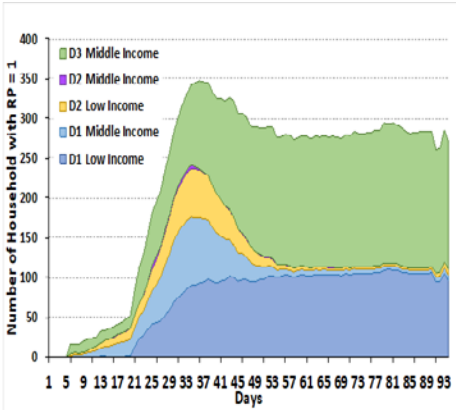




c. *Sensitivity analysis of the risk perception dynamics in a population of BN1 agents (Exp2) depending on the intensity of social interactions.*

d. *Sensitivity analysis of the risk perception dynamics in a population of BN1 agents (Exp2) depending on the timing of the media activation.*





e. *Distribution of preventive actions over time in the population with deterministic CA decision making (Exp2)*

f. *Distribution of preventive actions over time in the population with adaptive BN2 CA decision making (Exp3)*

Figure 3-6: Output measures of the experiments

3.3.3 Intelligent Risk Perception

From a psychological perspective, to be able to act upon risk, people – i.e. agents in the CABM – must first be aware of a risk. Experiment 2 presents the case when intelligence is added in the threat appraisal (BN1) stage. When being aware of risk while fetching water, agents in Exp2 may change their behaviour using a deterministic rule-based algorithm (Table 3-3). Thus, actions that agents select in this CA stage are based on current information, ignoring any previous experiences. Enhancing agents with cognitive abilities for threat appraisal (BN1) reduces the total number of infected agents by 90%. In Exp2, the total number of cholera-infected agents decreases (see the blue epidemic curve of Exp2 in Figure 3-6.b). In other words, information about a disease spreads through different channels – media, own observations, the experience of others, while a simple set of precautionary actions give rise to a steadier epidemic curve. Following the epidemic peak, agents are risk-aware and take a variety of precautionary actions based on their income class and education, ill individuals in their own and/or their neighbours' households; thus, fewer infections occur at the later stages of epidemics. Therefore, the BN1 epidemic curve (in Figure 3-6.b) has a lower peak and a steeper, vanishing tail compared to the ZI epidemic curve (Figure 3-6.a). The first heavy rainfall boosts the spread of cholera and can be detected in the shape of this curve at approximately day 23 in Exp2. Then, the effect of new disease exposure on the number of infected is counterbalanced by the activated risk awareness within the BN1 population. New exposure occurs when agents either lack infection experience in their social network or choose to ignore risks at the coping stage. The CABM enhanced with BN1 for the threat appraisal may be used to explore the spatial and temporal patterns of disease spread depending on varying risk communication strategies. To demonstrate this notion, we run a sensitivity analysis on the main communication channels.

Table 3-3: Rule - Based algorithm (CA) for Experiment 2 where agents select a static decision to take based on their characteristics

Household Characteristics				
Income	Educated	Infection in household	Infection in neighbour households	Decision
Low	No	No	No	D1 (same)
Low	No	No	Yes	D1
Low	No	Yes	No	D2 (walk)
Low	No	Yes	Yes	D2
Low	Yes	No	No	D1
Low	Yes	No	Yes	D2
Low	Yes	Yes	No	D2
Low	Yes	Yes	Yes	D2
Middle	No	No	No	D1
Middle	No	No	Yes	D2
Middle	No	Yes	No	D4 (buy)
Middle	No	Yes	Yes	D4
Middle	Yes	No	No	D1
Middle	Yes	No	Yes	D3 (boil)
Middle	Yes	Yes	No	D3

Middle	Yes	Yes	Yes	D4

Sensitivity analysis on the number of social interactions

A diffusion of information about disease risk and the effectiveness of risk-coping measures occur through social interactions. Their intensity impacts the spread of awareness about cholera risk in the study area as well as the number of infected individuals. Following Melegaro et al. (2017), the base scenario of Exp2 (and Exp3) assumes that when fetching water, agents exchange information daily with seven agents from their social network. These social links are set up randomly among households in the same community using the same water collection point. In addition, we run sensitivity analysis considering 3, 15, and 25 unique social interactions with individuals outside their own household per day. Figure 3-6.c and Table 3-3 illustrate the sensitivity of the number of individuals perceiving cholera risk and the resulting number of infections under various assumptions regarding social interaction.

All curves in Figure 3-6.c demonstrate a steep increase in risk perception around day 23 of the simulations. This point indicates the first heavy rainfall, when the population of agents depending on river water increases, and the disease diffusion via the dumpsites begins. During this first period, all scenarios exhibit the same pattern. However, after day 40, a clear difference is observed between the four scenarios. As expected, the higher the number of daily contacts (with which intelligent BN1-agents exchange information), the higher the number of households who perceived risk. Higher levels of cholera risk awareness trigger agents to make alternative decisions regarding water use (D2-D4 instead of D1), following the deterministic rule-based algorithm, and thus leads to a reduction in the number of infected individuals (Table 3-4).

Table 3-4: Sensitivity of the extent of an epidemic on the intensity of social interactions and information exchange among intelligent agents (Exp2)

No. of Contacts	RP Peak Day	Epidemic Peak Day	Percentage of Total Population Infected from the base
Three	83	35	103 %
Seven (base)	40	36	100 %
Fifteen	71	35	75 %
Twenty-Five	66	36	74 %

With fewer social interactions, BN1-agents are less likely to be aware of any cholera cases in their neighbourhood. Therefore, they will use the usual water fetching point, causing more individuals to be infected with cholera. As the speed of information exchange increases, agents learn from the experience of a larger group of individuals with respect to safety of alternative water fetching points and potential preventive behaviours. Since communication with neighbours is not the sole information source influencing the formation of RP in intelligent BN1-agents, the relation between the number of daily contacts and the resulting number of infected is non-linear: when interaction intensity changes from 7 to 15 people, the number of disease cases decreases by only 25% (Table 3-4).

Sensitivity analysis with respect to the timing of media broadcasting

During the 2005 cholera epidemic in Kumasi, the media began to widely broadcast epidemic information 21 days after the first infected case. We test the sensitivity of risk perception dynamics and the number of infected in response to the different media broadcasting timings. Thus, we ran the CABM with different media activation dates – 10, 30, and 40 days post-infection – in addition to day 22 (the base case of Exp2).

Figure 3-6.d illustrates that, generally, when the media reports on the cholera outbreak, the number of BN1-agents perceiving risk increases abruptly. This is true for the media activation scenarios on day 22, 30, and 40; however, this does not hold true for early activation (at day 10). The BN learning algorithm considers several factors at the threat appraisal stage. Thus, although BN1-agents have been alerted about cholera by the media on day 10, they did not yet observe any cholera cases in their household or neighbourhood. In addition, depending on the rainfall intensity, they may still have access to safe tap water that will only stop working following heavy rainfall on day 23. This combination of observations within their household and social network triggers BN1-agents to discard media messages and conclude BN1 simulations with low RP.

Table 3-5: Sensitivity of the extent of an epidemic on the timing of Media broadcasting in the population of Intelligent agents (Exp2)

Day of Media Activation	Percentage of Total Population Perceived Risk	Epidemic Peak Day	Percentage of Total Population Infected from the base
Tenth	83.7 %	36	89.4 %

Twenty Second (base)	100 %	36	100 %
Thirtieth	87.8 %	35	106.1 %
Fortieth	75.2 %	35	108.3 %

The timing of media messages does not affect the peak day of an epidemic, but impacts the resulting number of infected individuals (Table 3-5). It seems that early media attention (day 10) increases public awareness, resulting in individuals taking precautionary measures at a later stage, when other factors contributing to thread appraisal become evident (the yellow RP curve above others at the second half of the epidemics in Figure 3-6.d). Yet, the relationship is non-linear: the later the announcement, the smaller the marginal impact. Namely, postponing the broadcast for 10 additional days (e.g. day 22 vs. day 30) results in 6% more infected individuals, while another 10 days of delay results in only 2% more infected (day 30 vs. day 40). It is evident that announcing the epidemic 10 days earlier than the base scenario (day 22) reduces infections by over 10%.

3.3.4 Disease Coping Strategies: Rule-Base vs. Intelligent Risk Perception

According to PMT, when individuals are aware of risks, they choose actions based on their response efficacy and self-efficacy (positive influence) and the response costs (negative influence). The population of agents in Exp2 is intelligent in their risk appraisal, but pursue simple, rule-based decision- making (Table 3-3) at the CA stage.

Following the heavy rainfall (between days 23-50), BN1 agents begin to explore alternative options to drawing water from their normal nearest fetching point (D1). The latter is almost equally chosen by low and middle-income households throughout the entire simulation (Figure 3-6.e.). As cholera risk awareness spreads, the proportion of agents deciding to walk to an alternative fetching point (D2, only low-income households) and to boil water (D3, only middle-income households) increases. Some middle-income households also decide to purchase water (D4). However, since all three alternatives – walk, boil, and purchase – infer additional costs, households shift back to the default D1 option as soon as heavy rainfall ceases, and the number of disease cases decreases. As Figure 3-6.e. illustrates, a difference also exists in the distribution of preventive actions across income classes. However, the action choice remains deterministic: it depends only on the characteristics of agents at initialization such as income and education. There is no feedback between the effectiveness of previous actions taken by BN1 households or their peers and current agents' choices regarding water use. Thus, BN1 agents in Exp2 do not learn at the CA stage.

Experiment 3 is run in order explore how the learning process on precautionary measures is reinforced based on previous experiences. Here, agents employ two BN learning algorithms: BN1 for the threat appraisal and BN2 for the CA. When facing cholera risk, agents in Exp3 learn to perceive risk and subsequently learn to protect themselves by making adaptive decisions based on their own previous experience and their neighbours' experience. The epidemic curves of Exp2 and Exp3 fall within a similar range (Figure 3-6.b), with one important difference; namely, BN2-agents seem to be over-confident about their disease prevention choices at the epidemic's onset (approx. day 23), but quickly learn to alter strategies immediately after the peak (Figure 3-6.f).

Cholera begins to spread from the first few days of the simulation in both Exp2 and Exp3. The total number of infected agents during the cholera epidemic is approximately the same: on average, 14.7% of the simulation's population (5000 individuals) in both Exp2 and Exp3. However, a qualitative difference exists in the type and dynamics of preventive actions. Figure 3-6.f demonstrates that, over time, agents driven by growing RP learn to boil water based on the previous experience, which leads to a steady increase of D3 strategy use in the BN2 agent population. Among middle and low-income household agents enhanced with BN2, no agents purchase water. Instead, they switch to boiling water (see green D3 zone in Figure 3-6.f). Simultaneously, the number of middle-income households taking water from their usual, now suspicious-looking fetching point is nearly reduced to zero over time (see the light blue zone in Figure 3-6.f). BN2-agents also learn that walking to another water collection point still may result in a negative outcome.

The distribution of coping strategies between Exp2 and Exp3 also varies in space and by income class (Figure 3-7). When low-income BN2-agents learn to compare efficacy and costs based on past experience in Exp3, they realize that walking to another fetching point may not be worth the effort. Instead, in Exp2, low-income agents basing their CA decision on the deterministic rule-based process continue to walk alternate fetching points (compare left-hand side maps in Figure 3-7). Non-adaptive middle-income households in Exp2 continue to use a combination of the three strategies provided at initialization. Yet, intelligent BN2 individuals in Exp3 converge to using boiled water in the majority of the cases (right-hand side maps of Figure 3-7), as it proved to be most rewarding alternative to D1.

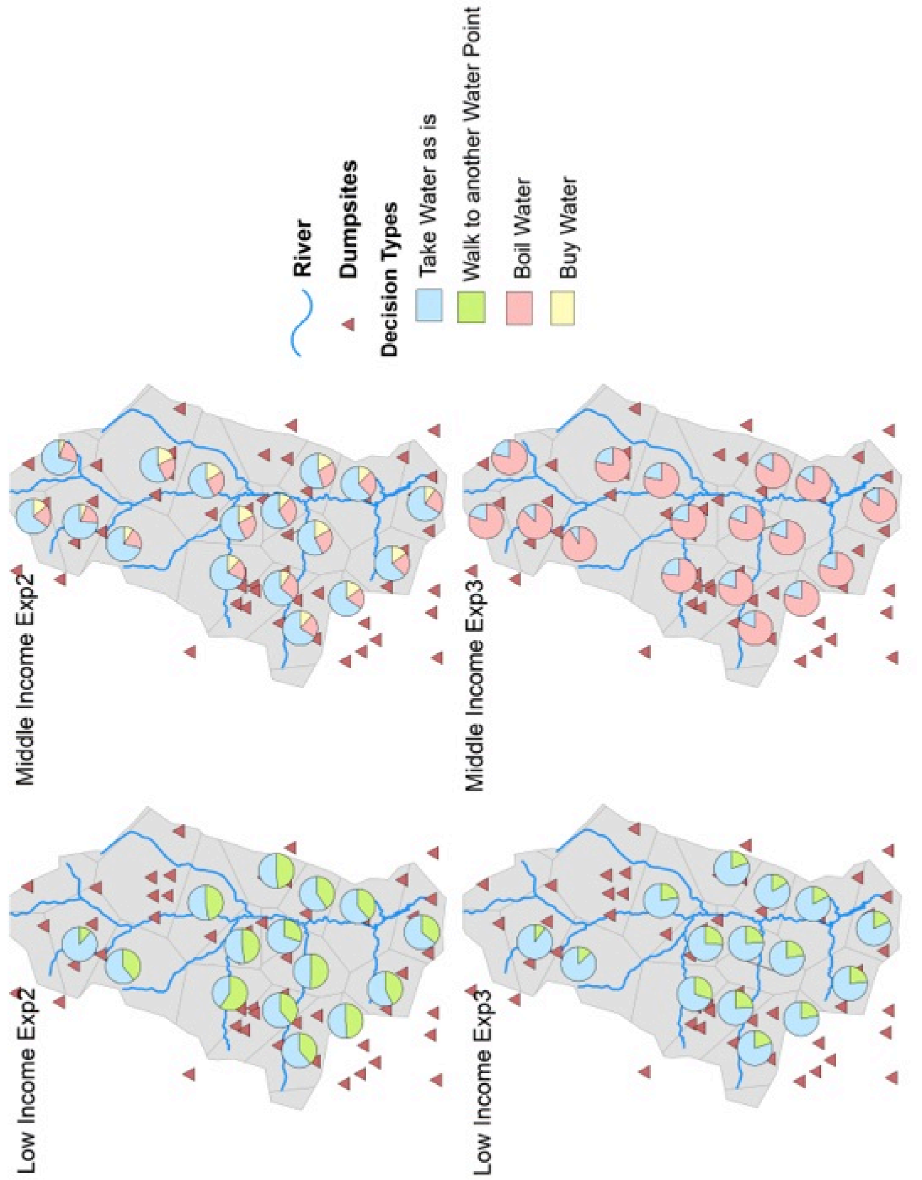


Figure 3-7: Distribution of preventive actions across space and income groups in Exp2 and Exp3

3.4 Discussion and Conclusions

Risk awareness and risk prevention behaviour can have a major impact on the number of disease cases during an epidemic. Models ignoring these elements of human behaviour may overestimate the expected number of disease cases. While a number of comprehensive disease ABMs have been developed, few explore the implications of these behavioural aspects and learning. This article introduces an innovative contribution by integrating psychological aspects of decision-making under risk into a spatial ABM using BNs learning algorithms.

We use an empirical spatial ABM of cholera diffusion (Augustijn et al., 2016) as a baseline model to test the impact of a multi-stage intelligent decision-making in a risk context. Two sets of BN learning algorithms are designed and coded using R, and are further integrated with the NetLogo-based CABM. Protection motivation theory from psychology lays the foundation for designing BN learning in two stages: one for RP appraisal and another for coping appraisal. We compare the results of the spatial agent-based disease model without intelligence (zero-intelligence), with an implementation of one-stage BN1 (only RP), and a two-stage BN2 (risk and coping behaviour) intelligence. Learning allows a population of heterogeneous and spatially distributed agents to perceive risk and acquire and share knowledge via a social network about the effectiveness of various disease protection actions. This spatial ABM enhanced with BNs allows us to explore the emergence of disease diffusion patterns tracing both geographic, educational, and income inequalities. The implementation strategy, in which we apply both BN1 for risk awareness and BN2 for risk appraisal, seems to outperform an implementation with a single BN. As agents learn about the effectiveness of preventive measures in addition to learning to recognize risks, the society as a whole makes healthier and more cost-effective choices. The sensitivity analysis on the behavioural assumptions indicates that the model is rather robust, with minimal impact on the final outcomes.

While this research presents a step forward in ABMs of disease diffusion by integrating psychology-based intelligence the context of risk, it can be further developed in a number of directions. Firstly, in addition to spatial, hydrological, and socio-economic data, this modelling effort could benefit from disaggregated behavioural data. Currently, our BN1–RP model is updated based on information obtained via personal communication, media, and visual observations of the environment. While we use data from the survey among students from developing countries to parameterize initial weights for RP factors, this may not be fully representative of the population in Kumasi. Disaggregated data on socio-demographic and behavioural characteristics of a target population is in demand to gain better insights on the interplay of factors influencing human behaviour during a disease outbreak. This is especially true for visual perception of the environment, as a current lack of information exists on how this factor influences total RP. In addition, a survey to collect data on how media affects people would improve the simulation. Model runs with richer datasets is within the scope of our future work.

Secondly, individual RP and coping appraisal can be implemented in disease ABMs using different AI algorithms. Besides BNs, genetic algorithms or neural networks might also prove useful. Further research is needed to explore the impact of various AI algorithms within the same base ABM. In addition, a systematic study on the performance of one AI algorithm across multiple ABMs for different types of risks in various geographic environments will provide a comprehensive understanding of the implications of introducing intelligence to agent-based modelling will have.

The implementation of risk and coping appraisals in disease ABMs will ultimately aid in supporting decisions regarding the timing of media attention to societal risks, and on the information that must be communicated to the public in order to prevent as many disease cases as possible.

Chapter 4: Spatial Intelligence in a Risky Context: Comparing Artificial and Real Actors[3]

[3] This chapter is based on Geocomputation 2017 conference paper that has the title “Integrating Spatial Intelligence for Risk Perception in an Agent-Based Disease Model”. The conference paper is authored by Shaheen A. Abdulkareem (the main writer of the paper and author of this dissertation), Dr Ellen-Wien Augustijn based at the University of Twente, Dr Yaseen T. Mustafa based at the University of Zakho and Professor Tatiana Filatova based at the University of Twente. The 2017 International Conference on GeoComputation held in 4th to 7th of September, 2017 in Leeds, United Kingdom.

4.1 Background

Spatial intelligence is one of the elements of the multiple intelligences theory developed by Gardner (2006). In ABMs, spatial intelligence is often applied for navigation (human or animal) or adaptation to land cover change (Kocabas and Dragicevic 2013). Few examples exist in which spatial intelligence is associated with spatial risk perception (RP) (Rufat and Samuel 2015). Thus, how does the spatial environment, and especially changes in the environment, influence individual risk perception?

RP is often the result of a combination of signals that a person receives. It may result from information received via (social) media, direct communication or their own observations made in the spatial environment (change detection). The judgement of all of these signals may differ per individual based on four factors, including the type of risk, the context in which the risk is perceived, the personality of the person and the social context (Wachinger and Renn 2010). However, risk appraisal based on spatial intelligence is not easy to measure. Limited data are available about the way the spatial environment impacts human decision making (Rufat and Samuel 2015). Most sources discussing risk perception will evaluate how risk perception varies in space but not which role the environment itself plays in the process of feeling scared. Implementation of spatial intelligence in agent-based models is relatively straightforward. Thus, finding suitable behavioural data to validate the risk perception implemented in the ABM remains a challenge.

Psychology approaches this subject using the Protection Motivation Theory (PMT), which is often applied in the health domain (Floyd et al., 2000). PMT assumes that a person facing a risky situation goes through a two-stage cognitive process: risk appraisal followed by a coping appraisal. The former is about checking risk and evaluating if RP is high enough to take an action. The latter stage concerns possible options and an intention to take an action. Moreover, RP can greatly impact the spread of diseases (Kitchovitch et al. 2010). When individuals are aware of risk, they may change their behaviour to prevent infection. Often, the risk awareness is modelled at two levels: the global level and location or personal level (Kitchovitch et al. 2010). The global level focuses on any media or government attention that increases the RP of individuals, while the location level is concerned about the observation of illness in neighbours that may have less impact – i.e. it depends on the number of infectious neighbours. By incorporating RP in the disease model, the diffusion of disease will decrease, which leads to a significant reduction in the number of infected cases.

RP is complex and therefore, can best be implemented using ML. In the previous chapter, RP was integrated in the disease model of cholera using BNs. The integration of RP in cholera ABM (CABM) led to reducing the number of infected cases by 90%. The spatial and social factors of RP were combined using BNs, which drive the agents' cognitive model to steer their behaviour. In this chapter, we focus explicitly on spatial risk perception and, as before, use CABM as a testbed to evaluate the spatial intelligence of household agents. Since little data is available on the influence of spatial factors on RP, we collected behavioural data using an online survey among students, mostly from developing countries. Therefore, the objectives of this chapter are twofold. Firstly, we seek to elicit the impact of spatial factors on people's risk judgements by collecting behavioural data on spatial intelligence. Secondly, we compare the risk awareness of agents in CABM with the data collected on RP of MOOC, as well as Google participants.

4.2 Methods

4.2.1 Visual Pollution in Spatial ABM

For this study, we used the cholera model for Kumasi Ghana, which was developed by Augustijn et al. (2016). Figure 4-1 illustrates the processes included in CABM that every agent passes through during simulation. As it is impossible to visually detect the presence of cholera bacteria in water, we assume that the safety of drinking water is assessed via the level of visual pollution at water collection points. The fact that individuals rely on personal observations when assessing the quality of drinking water is also supported by literature (Crampton and Ragusa 2016).

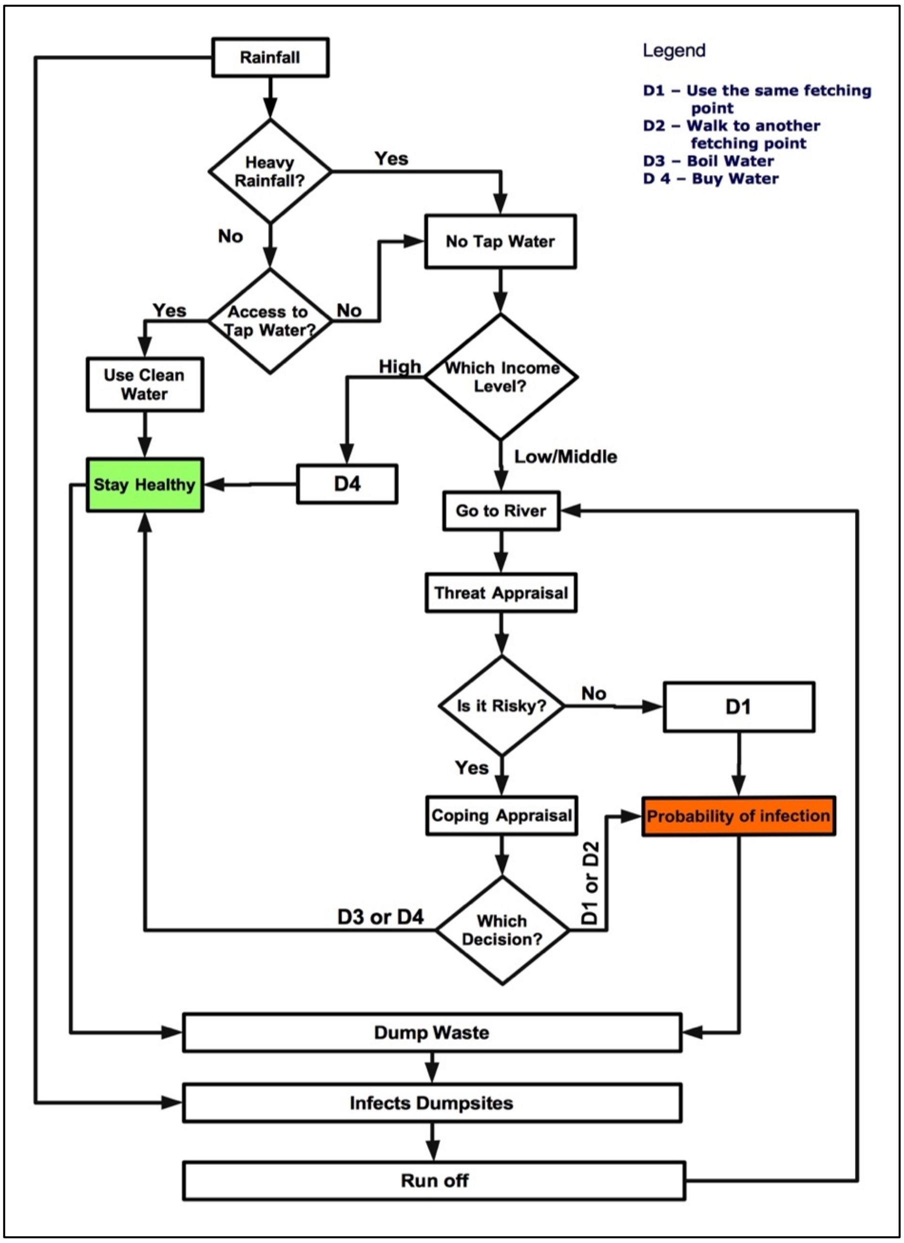


Figure 4-1: The conceptual flow of decision making, including two cognitive stages from PMT, in the spatial ABM of cholera diffusion

We model floating garbage in river water and on riverbanks and use the perception of pollution as an indicator for the safety of the drinking water. When dumpsites are located on the riverbanks, it is likely that some dumped materials will end up in the river. We refer to this as primary visual pollution (VP1), which can also be interpreted as the rate at which garbage enters the river. VP1 is calculated for every water collection point, once a day, based on the number of open refuse dumpsites that are located within a distance of 200 meters from the river. Thus, the primary visual pollution is estimated as:

$$\mathbf{f(VP1)} = \sum_{\mathbf{i=1}}^{\mathbf{N}} \frac{\mathbf{x\,g}}{\boldsymbol{d}} \qquad 4.1$$

where N is the number of dumpsites around the river water collection points; x is the number of households who use the dumpsite; g is the amount of garbage produced by each household; and d is the distance from the dumpsites to the water point (1 m $\leq d \leq$ 200 m).

We assume that during dry days, the garbage in the river will remain relatively static; i.e. a limited amount of new garbage will be deposited into the water. However, during heavy rainfall events, the river will carry floating garbage to downstream areas. This process will lead to a variability in visual pollution levels at all water collection points over time. Although floating debris in river waters is a fast developing research field (Gasperi et al. 2014), as it is relevant for the diffusion of plastic in seas and oceans, generally accepted models, especially for rivers, are still lacking. Therefore, we make a simplification by assuming that the speed of plastic debris is equal to the simulated flow velocity of water. In reality, other factors, such as wind direction and velocity, also impact the flow. The total visual pollution (VP) at any water collection point (w) can then be calculated as:

$$\boldsymbol{VP_w = VP_{in} + VP1 - (VP_{in} + VP1) * f} \qquad 4.2$$

where VP_{in} is the amount of garbage reaching the point from upstream water points and f is a random variable (0-1) that represents the fraction of the pollution floating downstream. As such, this total VP will only be calculated for days with heavy rainfall. In 2005, the epidemic of cholera started from September to December, which is known as the rainy season there in Kumasi. Thus, the model has real data recorded of rain, which contains data of heavy rainfall.

Further, we combine the risk awareness raised by VP via spatial cognition along with other factors that induce RP, including media, memory and personal communication.

4.2.2 ***Behavioural Data Collection***

To derive a BN from data, micro-level behavioural datasets describing the relation between risk perception and the model variables are needed, including visual pollution, media, neighbour, and memory.

This data was collected during two surveys: a survey among international participants of the Massive Open Online Course (MOOC) Geohealth and an online survey (Google survey). The MOOC Geohealth was organised by the Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente, in the Netherlands, during 2016 and 2017, with 194 and 235 participants from 92 countries (54% were from Africa, including Ghana) completing the survey. MOOC participants were split randomly into four equal subgroups, and were then shown pictures of rivers with different levels of visual pollution (Figure 4-2). All subgroups answered the same set of survey questions, testing their willingness to use the river water for drinking and cooking purposes. In every subsequent question, additional information on other factors, such as memory, media attention, and communication with neighbours, was provided.

Examples of the questions are shown below, and the answers to the questions are either yes or no:

Q 1.A *You have never used water from this source for cooking food. Would you use the water shown in this picture (one of the pictures from Figure 4-2) for cooking food?*

Q 1.B *You have previously used water from this source (one of the pictures from Figure 4-2) for a period of time in your cooking. Would you use this water for cooking food again?*

Question Q 1.A is an example of a question were only VP is being assessed, based on the picture shown to the survey participant. Question Q 1.B checks the combination of visual pollution (shown in the picture) with another factor; in this case, memory.

Figure 4-2: Four different pictures of rivers with pollution of various intensity. These pictures differ in colour of river water and level of floating debris (only on banks – in banks and in water). Photos source: https://www.shutterstock.com/

Combinations that have been questioned are: VP and memory (1), VP and media (2), VP, memory and media (3), VP and contact with neighbours (4), and VP, memory and contact with neighbours (5).

Information on the influence of individual risk factors on water use was collected during a separate survey implemented using a Google form. The questions were divided into two categories: risk perception based on individual factors and risk perception based on a combination of two factors. Participants were asked to indicate their risk perception for all combinations. The importance of factors including 'Visual pollution', 'Media', and 'Contact with neighbours' was surveyed both individually and in combination with other factors. This survey was distributed to students enrolled in Master of Science courses in Geoinformatics and urban planning and management at the Faculty of ITC, University of Twente. In total, this led to 125 survey participants from 33 countries (35% of them were from Africa including Ghana).

The participants were asked to indicate their risk perception for all combinations. The survey aimed to collect information on the degree to which RP related to cholera varies from person to person. There are different factors that influence the risk perception. In addition, the participants were informed that from literature three factors were selected: visual appearance of the water (is it visually polluted) (1), the fact that media is reporting on cholera cases in your area (2), and the fact that you hear about cholera in the neighbourhood from neighbours (3). The participants indicate for each individual factor how it influences their RP on a scaled of 1 to 10 where 1 indicates very low influence and 10 a very high influence. This is to test how strong each factor weights in the process of cholera risk perception.

The combination of two factors were proposed to participant as three multiple-choice questions where participants check boxes that indicates all situations where they might perceive risk. For example, you hear from the media that there is illness and your water look polluted:

	Media reports illness? Yes	No
Visual Pollution of the river and river w...	☐	☐
No	☐	☐

In the MOOC survey, four levels of pollution were shown through pictures: no visual pollution, brown water, low visual pollution and high visual pollution. However, in the Google survey, participants only knew that the water was polluted, without any indication of the level of pollution. Thus, we combined the data gathered from these two surveys into one dataset to ensure that all possible combinations of factors were stated and RP responses were included.

4.2.3 Experiments in CABM with the Spatially-Intelligent Agents

To instantiate each BN, one needs to define factors affecting an agent's choice and specify initial weights that will be further updated during the learning process. We used the factors that our respondents found important to parameterise the weights in a risk appraisal BN1 and a coping appraisal BN2. Here, we first focus our analysis on the spread of the visual pollution and the differences in VP at different study area locations (1). We then compare the visual pollution with the actual infection to determine if VP is a good indicator for cholera bacteria (2). Eventually we focus on the risk perception of agents as implemented via the BNs (3). For this last step we conducted two experiments, one with only VP1 (visual pollution around dumpsites) (Figure 4-3) and another experiment with total VP (down flow of plastic debris during heavy rainfall). We compare the impact of various factors on the dynamics of risk perception in agents' population as well as on the diffusion of different water use practices over space and time.

4.3 Results and Discussion

4.3.1 Data Collection

The observed pictures had a strong influence on the risk judgements of the MOOC participants (Table 4-1). Only 39% (world) and 29% (Africa) of the respondents perceived cholera risk in clear water, while 84% and 80% perceived risk when water and river banks appeared visually polluted. However, when other RP factors, such as memory, neighbours, and media were added, this pattern changed. For clear water, RP increased to 78% with memory and neighbours, and 80% when adding memory and media (world). A similar increase (84%) was observed in the African subsample.

Table 4-1: Percentage of positive responses relating VP to individual risk perception in the MOOC survey

Risk Factor(s)	Percentage of Risk Perception = Yes	
	World (%)	Africa (%)
Clear Water	39	29
Clear Water + Memory	30	30
Clear Water + Memory + Media	80	84
Clear Water + Memory + Neighbours	78	76
Water with brown colour	53	66
Water with brown colour + Memory	49	64
Water with brown colour + Memory + Media	81	76
Water with brown colour + Memory + Neighbours	82	79
Visually low polluted Water	79	86
Visually low polluted Water + Memory	56	55
Visually low polluted Water + Memory + Media	82	85
Visually low polluted Water + Neighbours	82	79
Visually high polluted Water	84	80
Visually high polluted Water + Memory	65	75
Visually high polluted Water + Memory + Media	95	94
Visually high polluted Water + Neighbours	92	92

This trend of increasing risk awareness was also observed for water with lower or higher levels of pollution. However, previous experience of safely using a water source (memory) decreases RP. Thus, when respondents were told that they previously used a water source, illustrated on the picture, the percentage of people perceiving risk decreased. The communication channels – either media or talking to neighbours – had a positive impact on RP, meaning that information about cholera cases confirms the agent's perception of risk. Notably, media had a stronger effect on risk perception than contact with neighbours. This is confirmed by the literature, as McClusky and Swinnen found that media can have a substantial effect on the opinion of an audience, either positively or negatively, through scary stories (McCluskey and Swinnen 2011).

4.3.2 Google Survey Results

In the Google survey, the respondents reported that for single risk factors (Table 4-2), a high level of VP leads to a high RP (86% world and 83% for Africa). Intensive media reporting particularly influences RP for African participants (77%). Contact with neighbours is less influential on perceiving cholera risk compared to VP and media (64% world and 66% for Africa.

Table 4-2: The percentage of individual risk factor influencing risk perception of participants in Google form survey

Risk Factor	**Influence Risk Perception (World)**		**Influence Risk Perception (Africa)**	
	Low (%)	**High (%)**	**Low (%)**	**High (%)**
Visual Pollution	14	86	17	83
Media	30	70	23	77
Contact Neighbours	36	64	34	66

Results for the combination of two risk factors can be found in Table 4-3. For all responses, including the world and Africa, risk perception based on multiple sources leads to a higher perception of risk. For the Google survey, combinations, including visual pollution, scored higher, though the differences were smaller (Table 4-3). When visual pollution is supported by other factors, respondents increased their perception of risk. This might be explained by the fact that cholera bacteria cannot be detected visually in the water but a confirmation by media or via contacts with neighbours can play a role in updating people's RP.

Table 4-3: Percentage of positive responses to risk perception with a combination of two risk factors

Combination of two factors	Percentage of Risk Perception = Yes	
	World (%)	Africa (%)
Visual Pollution + Media	88	83
Visual Pollution + Contact Neighbours	86	86
Media + Contact Neighbours	76	80

4.3.3 Experiments

Visual Pollution

Visual pollution differs considerably over the study area and over time. Thus, visual pollution is measured from 0 (no pollution) to 1 (high level of pollution). As can be seen in Figure 4-3, VP is higher at downstream locations (south) compared to upstream locations. It is also notable from Figure 4-3 that the level of VP increases when there are dumpsites located close to the river. As such, the three circles indicate locations that show that dumpsites effect the VP downstream.

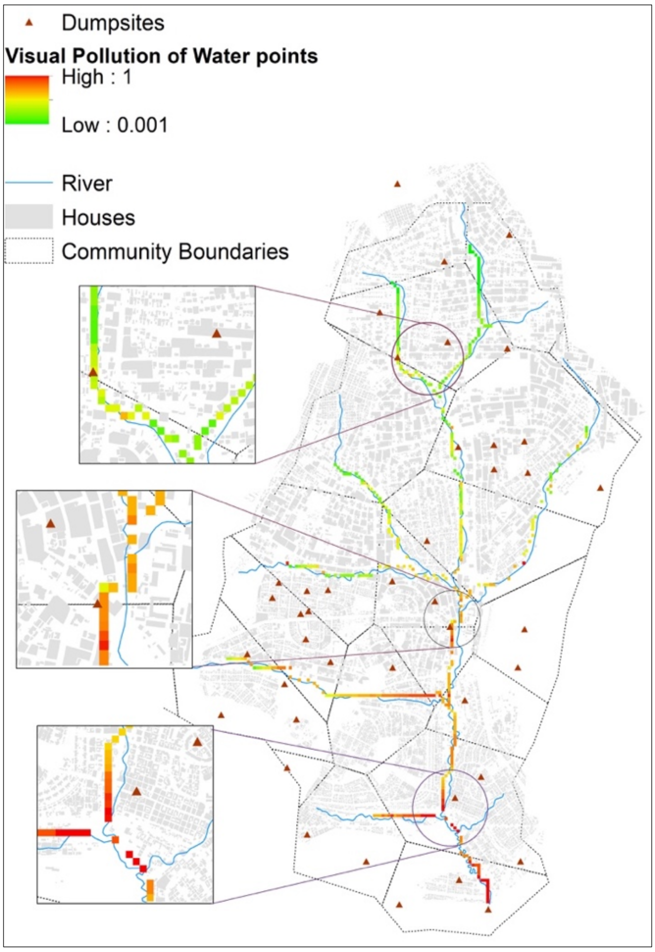
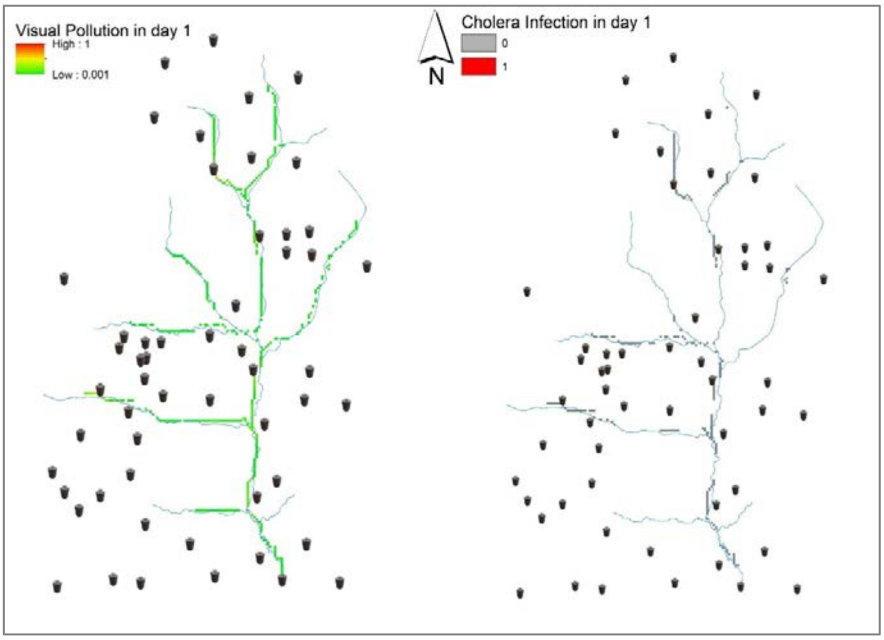
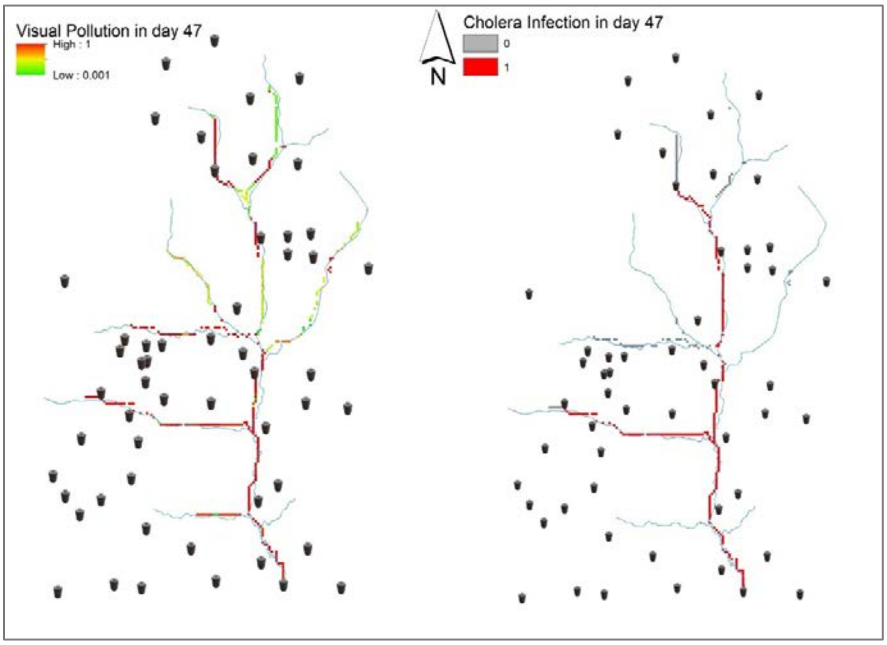


Figure 4-3 Simulated levels of visual pollution (VP) around open dumpsites. Higher levels of VP are observed for dumpsites closer to the river.

When we compare the VP levels over time, we see that at the beginning of the simulation, the VP level was low (Figure 4-4), though it increases with the number of disease cases and remains high towards the end of the simulation.

We can also compare VP with the actual level of infection at the water collection points. This makes it possible to detect if VP is a valid predictor for the risk of infection. We can differentiate between true positives (there is VP and there is infection), true negatives (no VP and no infection), false positives (there is VP but no infection) and false negatives (no VP but there is infection).


Visual Pollution in day 1
High : 1
Low : 0.001
N
Cholera Infection in day 1
0
1


Visual Pollution in day 47
High : 1
Low : 0.001
N
Cholera Infection in day 47
0
1

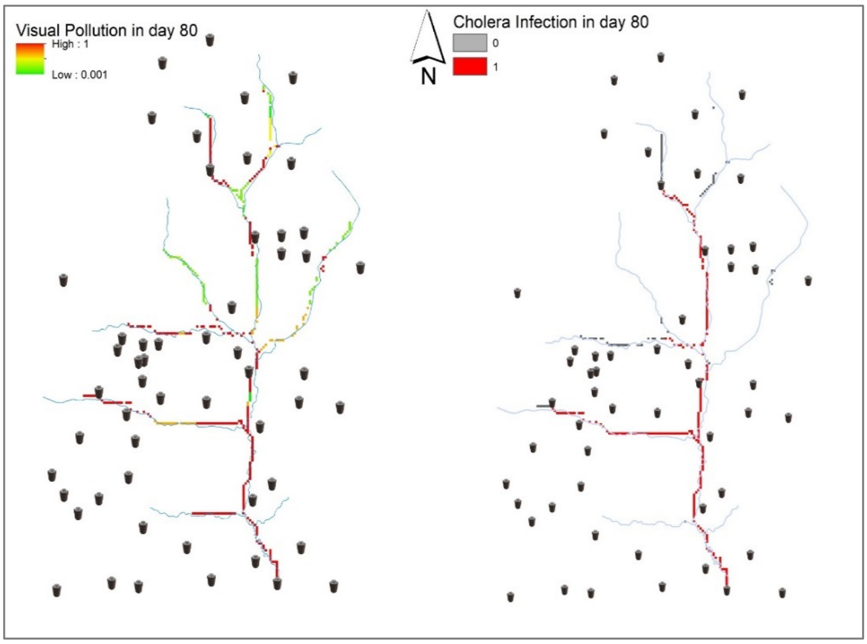


Figure 4-4: Spatial comparison between VP (left hand) and cholera infection per water point at the beginning, middle and end of the simulation

The predictive value of VP can be measured per time step. Figure 4-5 shows an aggregation per 10 days. In the time period between day 31 and day 40 (31-40), we see that in 53% of the cases, there is no VP but there is infection. This high percentage is probably due to the time it takes for VP to diffuse over the study area. Especially in the beginning of the epidemic, VP is not a good indicator.

	False predictions		Correct predictions	
Days	**VP No Infection %**	**No VP Infection %**	**No VP No Infection %**	**VP Infection %**
1 - 10	7.05	0	23.2	0
11 - 20	17.4	0	78.7	4.08
21 - 30	24.2	22.01	39.8	15.2
31 - 40	2.7	52.6	23.2	21.3
41 - 50	28.5	17.9	10	44.3
51 - 60	39.7	1.8	0.6	58.9
61 - 70	39.1	1.5	1.2	56.2
71 - 80	59.6	1.4	2.7	37.6
81 - 90	96.2	0.001	1.4	0.6

Figure 4-5: False vs true prediction of cholera infection through visual pollution

Agents learning and survey data

When PMT is integrated into CABM and represented by two BNs, we see a drop in the total number of disease cases to a level of approximately 10% of the original numbers (see Chapter 3, Figure 3-6). This confirms the findings of Kitchovitch et al. (2010), who argue that incorporating RP in disease models greatly decrease the number of transmissions.

Table 4-4 presents the percentage of agents in the CABM who perceive risk, and compares it to the percentage of survey participants. Table 4-4 includes individual RP factors, as well as combinations of them (two or three) and follows the setup of both surveys (MOOC and Google).

Table 4-4: Comparison of agent risk perception per risk factor with the original survey data

Risk Perception Factors	Survey	CABM
VP	49.1	46.2
Memory	42.7	15.9
Media	63.8	59.2
Neighbours	61.3	56.7
VP + Memory	53.9	58.9
VP + Media	90.1	72.8
VP + Neighbours	86.3	68.7
Memory + Media	81.1	63.6
Memory + Neighbours	80.0	65.1
Media + Neighbours	75.8	72.5
VP + Memory + Media	87.1	79.2
VP + Memory + Neighbours	83.7	73.7
VP + Media + Neighbours	78.9	78.2
All factors	NA	52.1

When agents in CABM base their risk perception only on VP (no media, memory, nor neighbour information is included), which is the case in 46% of household agents, their risk perception is in line with real values (49%). For memory, the opposite result was evident. In the survey, 42% of the participants perceived risk based only on memory. When participants had been informed that they used this water before, 7% changed their mind and indicated that they would use this water again. The same is true for agents in CABM; although they may know there is a risk, they do not have any other choice and will still have to use the water. A value of 16 % for

CABM indicates that agents who had been using river water before the cholera outbreak will continue doing so during the cholera outbreak because they trust the source. Therefore, we found that memory has a negative impact on the risk perception process. Whenever memory exists, either alone or in combination with other factor(s), it will lower the RP, reflecting the trust of people on the source of a water that might be polluted with cholera.

The effect of neighbours communicating experiences of illness leads to 57% of the agents having RP in the ABM, which is close to the 62% in the survey. This is similar for the media's impact on agents' RP in the cholera model, which accounts for 59% of agents compared to 64% in the combined surveys dataset.

The results of the survey also confirm that the level of trust in boiled water is much higher compared to un-boiled water, as agents also change their behaviour to boil water in the model (Table 4-5).

Table 4-5: Percentage of individuals decision type in both survey and CABM

Decision type	MOOC (all participants)	MOOC (Participants from Africa)	CABM
No Risk - Use this water	42 %	56 %	42 %
RP - Walk to another water point (source)	84 %	77 %	30 %
RP - Boil water	72 %	75 %	57 %

4.4 Conclusions

There are clear indications that the spatial environment plays a role in the risk perception of people. However, risk appraisal based on spatial intelligence is not easy to measure. Limited data are available about the way the spatial environment impacts human decision making. Most sources discussing RP will evaluate how risk perception varies in space (Rufat and Samuel 2015) but not which role the environment itself plays in the process of feeling scared. By implementing spatial and social cognition for risk appraisal and coping appraisal, we attempted to mimic this behaviour and evaluate its impact on disease diffusion.

The data collected in the surveys showed that visual observations of the spatial environment impact the perceptions and decisions of people. In the MOOC survey, different results were found for pictures showing different levels of visual pollution. This confirms the fact that people judge by appearance.

When the observation of visual pollution was supported by other elements, such as communication with other people or hearing news from the media, people become more aware of health risks. The results of our surveys also show that every combination of risk factors have their own RP values. Adding one extra factor may change the risk perception, either increasing or decreasing it. This underlines the fact that combining different factors into one total risk perception is a very difficult task. As such, a data driven approach using ML can be very helpful in this respect.

Our results show that risk perception via VP modelling does not always match cholera infection levels. This is not a problem except when we are dealing with false negatives (no pollution and infection). The model revealed that many false negatives occurred in the most crucial phase of the simulation (peak period). In this case, the absence of pollution is taken as a signal that the water is safe to use. The same result was also found in the surveys. In the MOOC, the pictures showing clean water tricked people into trusting that the water was clean. This is important information that can be used by managers to take intervention actions. Making the public aware of the fact that they should not judge the water by its appearance (at least, not to trust clean water), can help to prevent disease cases.

Our model of VP was rather simple and further research is needed to improve this model. This will also require the validation of VP, which was not possible in this study, as we did not have data of the location of garbage during the 2005 epidemic. Running experiments that could give measurements on the way and amount of floating garbage in the river can help for validation. Flying drones along the river in Kumasi during the rainy season and capturing video and photos can also help to validate the VP sub-model of CABM. In addition, other factors, like wind direction and velocity, also impact the flow, but they are not included in the model. These factors also need to be considered in the process of improving the visual pollution model. This improvement is necessary because it is relevant for the diffusion of plastic in seas and oceans, as well as for rivers.

Research on risk perception during epidemics is often conducted too late when the peak is over, or in distant geographic locations outside of the epidemic area. Hence, it provides little empirical evidence on the dynamics of people’s behaviour and risk perceptions. More research on risk perception during epidemics, including other related variables, such as disaggregated data on socio-demographic and behavioural characteristics of a target population, is in demand. This will help to gain better insights into the interplay of factors influencing human behaviour during a disease outbreak, which is especially true for a visual perception of the environment.

Notably, the survey participants employed in this study were well-educated individuals from a variety of nations. This imposes limitations to make policy-relevant conclusions, though it allows us to test the fitness of ML algorithms implemented within a spatial ABM. Still, we recognise that in communities with low income and marginal education, the response of individuals might be different. In addition, the response of people might also differ whether or not they have an alternative water source to use. Therefore, they might know there is a risk but since they do not have another choice, they still may use the water. This could be a point to emphasize the spatial ABM part too, where different agents in different locations have access to certain water sources.

Chapter 5: A Workflow on Using Limited Survey Data for Training Bayesian Networks for Spatial Learning[4]

[4] This chapter is based on journal paper (paper 2) that has the title "Bayesian Networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models". The paper is authored by Shaheen A. Abdulkareem (the main writer of the paper and author of this dissertation), Dr Yaseen T. Mustafa based at the University of Zakho, Dr Ellen-Wien Augustijn based at the University of Twente, and Professor Tatiana Filatova based at the University of Twente. The paper has been submitted to the special issue of an International Journal on Advances of Computer Science for Geographic Information Systems (Geoinformatica) under the theme "Spatial Agent-Based Models: Current Practices and Future Trends". DOI: 10.1007/s10707-019-00347-0

5.1 Introduction

The proliferation of agent-based models (ABMs) as a research method calls for advancements in how agents learn and adapt. In ABMs, agents can possess their own cognitive model that can be trained using real data. Machine learning (ML) algorithms are used increasingly to enhance agent learning abilities and implement autonomous smart behaviour. Notably, ABMs with intelligent agents are argued to capture complex real-world phenomena more realistically (Asadi et al., 2009). Therefore, the behavioural rules guiding an individual agent's decisions, and the interactions between agents and environments, significantly affects the macro-patterns emerging from a model (Alonso et al., 2001).

In ABMs, ML algorithms are applied in one of the following phases: (1) prior running the ABM, (2) during the run of the ABM, or (3) after running the ABM to analyse model output. Before running an ABM, ML algorithms can be used to derive parameter values based on empirical data or to prepare input data (Besaw et al., 2010). During the simulation, they are used to capture ABM outputs to train a learning algorithm during the simulation (Kocabas and Dragicevic, 2013). After running the simulation, ML algorithms are applied for calibrating and validating ABM output (Heppenstall et al., 2007). The algorithms used for these three types of application vary from neural networks, to Bayesian networks (BNs), to genetic algorithms, and other ML algorithms.

ML algorithms, used during the run of the ABM, are often implemented to steer agent behaviour. Complex emerging behaviour can be the result of combinations of previous experiences of an agent (feedback), of social interactions with other agents, but also of changes in the agent's environment. ML algorithms can play an important role in combining a large number of different variables (spatial and non-spatial variables) and obtaining the social- and spatial intelligence level required. Learning is achieved when agents can use their observations to solve complex problems or derive smart solutions. Examples of ML implemented in spatial ABMs are the implementation of ML for spatial optimization of land-use allocation (Vallejo et al., 2013), searching for the best location and pricing strategies in a competitive business environment (He et al., 2014), and optimizing migration patterns (Heppenstall et al., 2007). Other researchers use spatial learning to spread spatial externalities and use spatial neighbourhood to diffuse information or strategies (Verstegen et al., 2010; Heinonen et al., 2012), in other words for the explicit modelling of spatially-correlated phenomena. Spatial learning can be achieved using rule-based modelling, but via ML it is easier to include spatial factors in an agent's decision-making (Pooyandeh and Marceau, 2014b). In addition, ML is used to obtain spatial knowledge of a resource and a quality of the environment (Kocabas and Dragicevic, 2013); or a threat/obstacle (Sharma et al., 2012) in a specific location.

One limitation in the use of intelligence in ABMs in general, and spatial ABMs specifically, is that most learning algorithms require extensive training data (Van Der Ploeg et al., 2014). Although massive geo-data is becoming increasingly available, data on human choices and rules that guide behaviour remain scarce. Moreover, the reduced availability of large sets of micro-level data on human behaviour influences the use of learning algorithms (Bratko, 1994). The performance of learning algorithms improves with the increasing quantity and quality of training data (Walczak and Walczak, 2001). However, in many domains, the problem of obtaining such qualitative large sets of empirical data remains (Karr, 2014).

It is important to note that ML and big data are not synonymous. The amount of data necessary to train and test the algorithm relates to the complexity of the problem at hand and the nature of the learning algorithm. A good addition to any dataset is domain expertise. BNs are not data intensive and represent a viable alternative for small training datasets; also, they tend to exhibit good prediction accuracy – even with small sample sizes (Kontkanen et al., 1997; Uusitalo 2007). BNs are particularly useful for the simulation of processes such as decisions under risk, which are characterized by multiple related (and uncertain) variables (Constantinou et al., 2016). This is primarily due to BNs being suitable for the reduction of complex domains into computationally manageable models (Weber et al., 2012). Another advantageous feature of BNs for such applications includes their capability in managing incomplete data and uncertain information.

Learning in BNs consists of two different tasks: design of a network structure with network probability values (1) and defining the conditional probability tables (CPTs)(2) (Bidyuk et al., 2005). The design of the BN determines the variables (nodes) used and the way these variables are combined (linked) to derive a decision. A design can be based on expert knowledge or derived from data. The role of the expert is to resort to subjective assessment of the network design and probabilities and make use of their experience and literature published in the field (Diez, 2003). The next step is the construction and adjustment of conditional probability distributions. Quantitative information for the BN must be obtained in the form of conditional probabilities (i.e. CPTs). In case of data availability, the CPT values are driven directly from the dataset during the process of constructing the network structure. When no data is available, the expert defines the CPT using marginal likelihood for parameter learning to cover the uncertainty in the values of the parameters. This training will be done during the simulation.

The challenge in constructing BNs from data relates to finding a network that best fits the available data (Campos, 2006). Expert-driven BNs are entirely reliant on experts with full knowledge about the domain (Julia Flores et al., 2011). However, it remains unclear how the outcomes of a model are impacted by variations in constructed BNs or the way in which they are implemented in the ABM (e.g., (Pooyandeh and Marceau, 2014b; Shen et al., 2011)).

Several studies have combined BNs with ABMs; for example, Kocabase and Dragicevic derived BN structures for different agent types and obtained their CPT values from census data before using the BN in a land-use change model (Kocabas and Dragicevic, 2013). In another example, both Ren and Anumba and Ma et al. used a simple BN structure, utilised experts to derive their CPT values, and trained during the simulation (Ma et al., 2004; Ren and Anumba, 2002). While Matsumoto et al. constructed a data-driven BN using a survey, they simultaneously trained their network to estimate internal parameters (Matsumoto et al., 2017). Furthermore, Pope & Gimblett used stakeholders to design their BN and CPT values (Pope and Gimblett, 2015).

The present study contributes to this literature by exploring: (1) the possibility of implementing learning in spatial ABMs with a small behavioural dataset, (2) the extent to which supervised learning of ML algorithms should depend exclusively on data, and (3) the level of intelligence necessary for agents to simulate realistic risk perception. To address this aim, we test alternative methods of designing a BN from a small sample of micro-level behavioural data.

The next section (5.2) describes how alternative, empirically-driven BNs are integrated in the spatial ABM. Section 5.3 discusses the results of the simulation experiments, and Section 5.4 concludes by discussing the main findings, advantages, and limitations of our approach, as well as possible directions for future work.

5.2 Methodology

5.2.1 Integration of Empirically – Driven BNs in the Spatial ABM

All BNs consist of a number of nodes connected by links in the form of a directed acyclic graph (DAG) (Heckerman, 1995). When integrated into a spatial ABM to enhance agent intelligence, each node represents a variable in the agent decision-making process simulated in that model. In the present study, these variables include: the observation of visual pollution at water collection points (VP), the reporting of media on cholera cases (Media), communication with neighbours that may or may not have cholera in their household (Neighbours), and updating and retrieving memory representing a household's previous use of the current water source (Memory). The BN supports agent decisions on assessing water infection level (i.e. Risk node) based on VP, Media, Neighbours, and Memory. The latter nodes (except VP) have a Boolean value (true or false) indicating the presence or absence of their relationship to risk. The VP node has three states: no, low, and high. These states indicate the level of visual pollution at water collection points. In the survey there were four different states, which were mapped as follows: clean water (no risk), brown water or a small amount of garbage around water collection points (low risk), and garbage on the river banks and in the river (high risk). In this paper, we explore four combinations for the specification of either BNs network structure or CPT. Namely, we run our spatial ABM with BNs designed (i): based on data only (both structure design and CPT is derived from data); (ii) based on data complemented with expert knowledge (structure is data-driven but CPT is expert-driven; (iii) on structure that is expert-driven but CPT is data-driven; (iv) based on expert knowledge only (structure and CPT is derived by the expert knowledge).

Any node can be updated upon new evidence, even when they are related to multiple variables. The evidence acquired about a state variable should propagate to update states in the rest of the network, and this process requires network training (learning). The training of BNs can be either by using data or by eliciting expert knowledge (Flores et al., 2011). BN training is performed via a flow of information through the network, and it can take place prior to using the network (i.e., before implementation within the ABM) with the availability of data or continue during the simulation runs (when it is an expert-driven network). In the first example, the training process of BNs ends with final probabilities (posterior probabilities) that the network will continue to produce every time it is consulted by the ABM. In the second case, the BN model needs improvement since it is not fully trained at the start of the simulation, though will be trained using data (agent decisions) generated during the simulation. This process is called sequential learning. Usually when no data are available to construct BNs, the adjustment of parameters (nodes) takes place when the network performs identification based on new evidences.

We compare four different BNs (DDBN, DEBN, EDBN, and EEBN). The first letter in the BN acronyms refers to the information source – data-driven (D) or expert-driven (E) – for their derived structure, while the second letter refers to the estimation of probabilistic parameters. When a BN is expert-driven, we either designed the structure and/or retrieved parameters based on census data of the case study area or literature dealing with risk perception of waterborne diseases such (e.g., (Doria, 2010; Driedger, 2007; Hedman and Lindberg, n.d.)). An overview of the networks is provided below.

DDBN

In DDBN, both the BN structure and its parameter values were driven by survey data. The scored-based algorithm “Tabu search” is used to construct the BN (Beretta et al., 2017). This algorithm makes use of a goodness-of-fit score function for evaluating graphical structures with regard to a dataset. Tabu search is a metaheuristic algorithm using short-term memory to ensure that the search explores new areas without remaining in a local optimum. In this algorithm, the fitting function is used to score a network/DAG with respect to the training data, and a search method is used to determine the highest-scoring network structure. This algorithm continually improves scores until converging at optimal results. DDBN was trained prior to the simulation.

EDBN

The structure of this BN follows the same approach as in the MOOC surveys. In the survey, we first showed participants a picture of water with a specific level of visual pollution, followed by questions related to the other factor(s). In this approach, VP is assumed to be the parent of the other factors. Then, we derive the probabilities of nodes states and CPT from the survey data and train it prior to simulation.

DEBN

The structure of DEBN is identical to that of DDBN. However, an expert assigned probabilities to node states and the related CPTs to formulate logical scenarios. These values were driven from the literature and used in the BNs.

EEBN

This is a fully expert-driven network adopted from Abdulkareem et al., (2018). The probability values of these network variables were derived from the available literature and census data for Kumasi, Ghana. Here, EEBN settings reproduce the original setup and serve as a benchmark to compare the three alternative combinations between survey and expert data for BNs.

Goodness of fit of BNs and model output:

BN validation was conducted using two steps: validation of the network structure using scored functions and validation of the learning parameters (CPT) (Figure 5-1). We also compared the outcome of integrating the four BNs into the CABM with the survey data to validate the realism of agent risk perception.

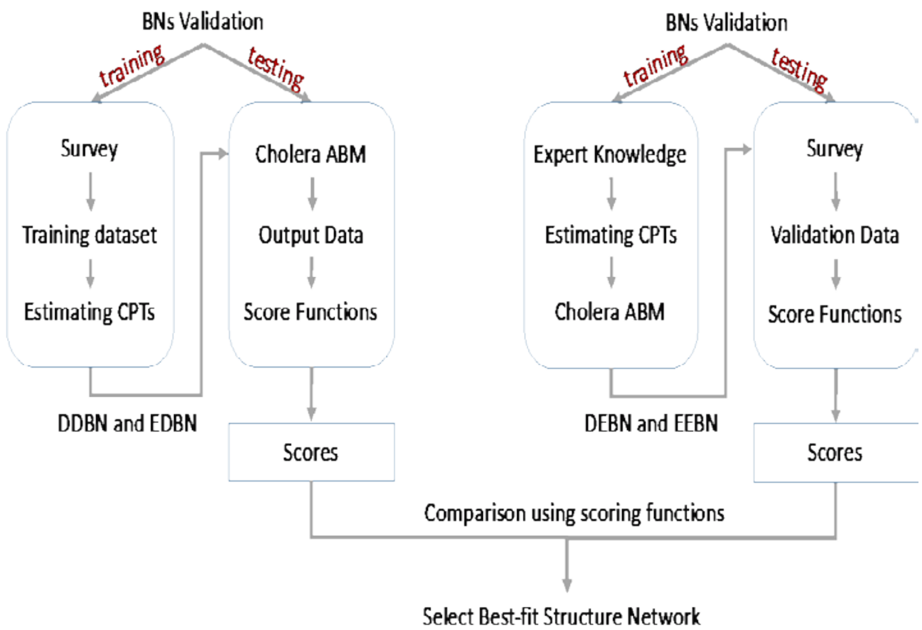


Figure 5-1: BN models' validation

There are two approaches that are commonly used to measure the goodness of fit of BN models (Needham et al., 2007). The first is to test if the conditional independence assertions involved by the structure of the BN model are satisfied by the training dataset. The second method is to evaluate the degree to which the resulted structure describes the data. To achieve this, we use scoring functions. Many scoring functions exist, and the most popular are AIC (Akaike information criterion), BIC (Bayesian information criterion), and Bayesian Dirichlet with likelihood equivalence (BDe) (Liu et al., 2012). The primary issue with scoring functions is the absence of an objective method to determine which function is optimal (Liu et al., 2012). AIC provides a relative measure of the information lost when a given BN model is used to represent reality, while BIC is an example of penalised likelihood and it selects the true model that fits the data. Moreover, BDe calculates the joint probability of a BN model for a given dataset. Overall, the optimal model from the set of BN models is the one with the higher absolute AIC, BIC, and BDe values (Carvalho, 2009).

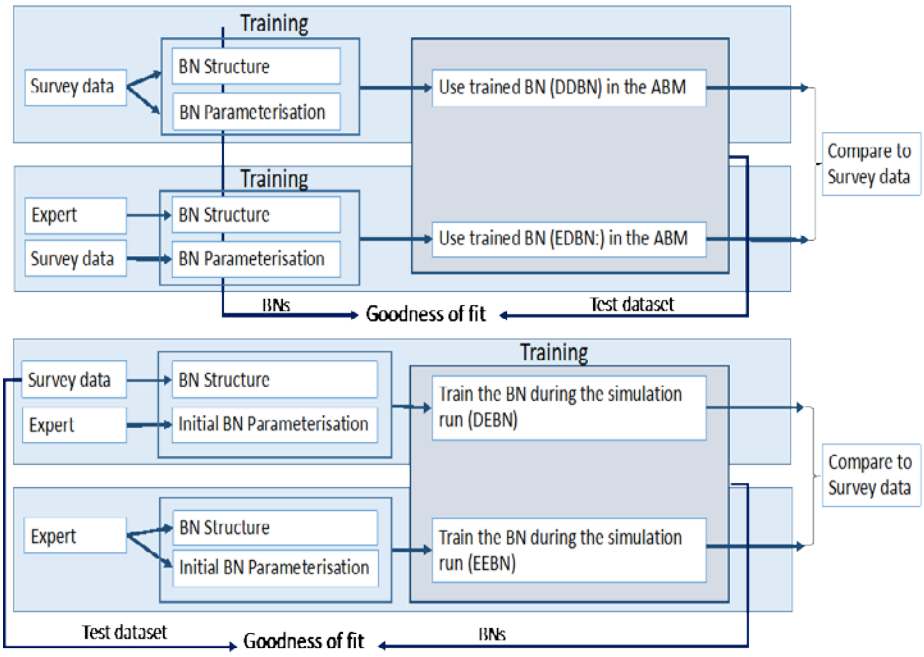


Figure 5-2: Methodological workflow used in this article

To address the main research questions, we follow a number of steps (Figure 5-2). The primary elements of this workflow are explained in the following sections. We conduct a total of four experiments, in which we run the CABM with all four BNs for 100 random seed runs, creating a new synthetic population every 5 runs[5]. We provide the mean values across 100 sets of runs for all output metrics. In the first set of experiments, we run the CABM with DDBN and EDBN, training them prior to the simulations. Then we run the CABM with DEBN and EEBN, training them during the simulations to adjust the initial values of CPT of both networks proposed by the expert knowledge. Since DDBN and EDBN are trained prior to implementation of the simulation, testing data for the goodness of fit of these two BNs arises from the ABM. Additionally, since DEBN and EEBN are trained while running the simulation, the survey data serves as the goodness of fit data. The sample size of all validation dataset is equal (i.e., the size of the empirical data) to balance the scores.

To compare model outcomes with the survey results, we calculated the average number of agents that perceived risk during the simulation. These percentages were computed for each risk factor and combinations of factors. In addition, mean epidemic curves and risk perception curves were obtained and compared.

BN models and resulting spatial patterns:

The implementation of different BN models may impact the behaviour of agents based on their location in space. To evaluate this impact, we present a set of maps that show the spatial distribution of risk perception variables. Risk perception factors: VP alone, VP with media, VP with contact with neighbours, VP combined with media and contact with neighbours, and combination of media and contact with neighbours are displayed per community with the implementation with each BN model. Risk perception based on media or contact with neighbours only does not occur.

[5] The agents in the model are created using a synthetic population in which household agents and individuals are distributed spatially over communities on the basis of the data from Osei (2010). For five runs, we have a fixed synthetic population, in which agents' initialization attributes such as the number of individuals per household, house location, river fetching water points, and accessibility to tap water remain constant. After five runs new attribute values are created. Stochastic events that influence agents' behaviour during the simulation include rain time, time to fetch water, spatio-temporal location of VP and cholera bacteria, and selection of dumpsites to be cleaned by the municipality.

5.3 Results and Discussion

5.3.1 BNs Structures and Parameters Validation

To select the best-fit network among our four BNs, we tested their structure and parameter values. The graphical structures (DAGs) of these BNs are illustrated in Figure 5-3.

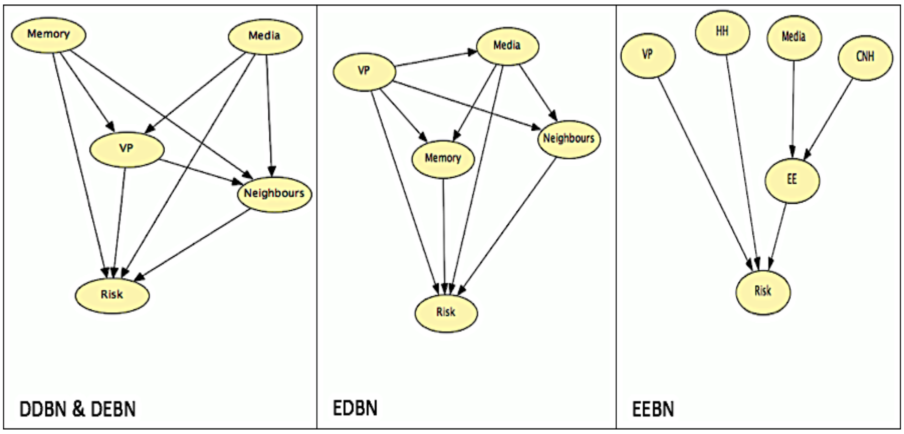


Figure 5-3: Graphical structures (DAGs) of Bayesian Networks

We validate DDBN and EDBN using the outcome of CABM, while DEBN and EEBN were validated using the survey data. The validation results are provided in Table 5-1. The numbers between brackets represent the relative metric of each BN models with (1) corresponding to the best value and (4) to the worst one.

Table 5-1: Scores of the four BN models. The number in brackets indicates how well each BN model does relative to the others on each scoring function, with (1) corresponding to the best value and (4) to the worst one.

BN Type	Entire survey sample			Survey subset with African participants only		
	AIC	BIC	BDe	AIC	BIC	BDe
DDBN	-13110 (3)	-13235 (3)	-13158 (4)	**-2589 (1)**	**-3173 (1)**	**-2663 (1)**
EDBN	**-13541 (1)**	**-13681 (1)**	**-13599 (1)**	-2488 (3)	-3140 (2)	-2598 (3)
DEBN	-13046 (4)	-13186 (4)	-13195 (3)	-2468 (4)	-2925 (4)	-2576 (4)
EEBN	*-13213 (2)*	*-13358 (2)*	*-13249 (2)*	*-2566 (2)*	*-2976 (3)*	*-2638 (2)*

As illustrated in the first half of Table 5-1 (Entire survey sample), the best-fit network is EDBN, since it exhibits the highest absolute scores for AIC, BIC, and BDe. The AIC score of its absolute value is higher than the other three scores. It also scores best in the BIC and BDe values. EEBN (the original expert network (Weber et al., 2012)) has the second-best scores, confirming that the involvement of experts leads to a better fit than using data only. This holds true for both the structure of the BN and estimation of probabilistic parameters and CPTs for the network, as indicated by the scores for EEBN.

The second half of Table 5-1 (subset African Participants) presents fitness scores for the four BNs with respect to responses of African participants. With the change in the sample from the full sample set to African respondents only, measured values of the goodness of fit of BNs change. For the subsample, DDBN exhibits the best fit. African participants seem to be more sensitive to media reporting and less sensitive to visual pollution. EEBN remains the second-best fit BN model. This again supports the usefulness of expert knowledge in implementing ML. in addition, this demonstrates the sensitivity of the metrics (AIC, BIC, and BDe) to the survey sample used.

5.3.2 Implementation of BN models in CABM

After testing the performance of the four BNs, we integrated them for agent decision making in the ABM. The DDBN and EDBN are trained then linked with the CABM. The DEBN and EEBN are trained during the simulation. We report average results across 100 runs for each BN model for several macro metrics of interest, including: epidemic curve[6], risk perception curve[7], and percentage of agents that perceived risk.

Epidemic Curve and Risk Perception

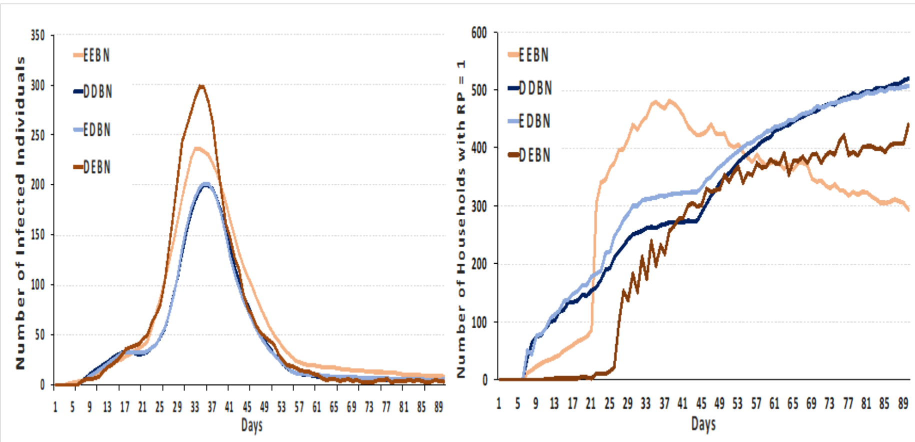
It is evident that agents with intelligence are less susceptible to cholera compared to those agents who are not using BN for their coping appraisal (Abdulkareem et al., 2018). Since BNs steer their interactions during

[6] Epidemic curves count the number of disease cases in each particular time step.

[7] Risk perception curves plot the number of individuals who perceived cholera disease risk – i.e. the outcome of their risk appraisal, BN1, represented here either as DDBN, EDBN, DEBN, or EEBN becomes 1 – over time.

simulations, they learn how to protect themselves. This can be observed in the epidemic curve (Figure 5-4 – left) and risk perception over time (Figure 5-4 – right).

Epidemic curves resulting from DDBN and EDBN (Figure 5-4.a) are approximately the same, and both provide lower peaks compared to the two alternatives (average peak value of 200 disease cases). The risk perception curves (Figure 5-4.b) of both BNs develop in a similar manner Agents are intelligent enough to perceive cholera risk from the start of the simulation, though EDBN exhibits a higher peak in the period between days 25 and 58. The risk perception curves continue to increase until the end of the simulation because this BN is pre-trained and does not change during the simulation. As such, agents do not realise that the epidemic is over and risk is decreasing.



a. Epidemic curves *b. Risk perception curves*

Figure 5-4: Results for the four BNs models of running CABM 100 times per BN integration

Running the CABM with DEBN and EEBN produces average epidemic curves with higher peaks of 230 disease cases for EEBN and 300 cases for DEBN (Figure 5-4.a). The risk perception curve for EEBN (Figure 5-4.b) exhibits a very steep increase in risk perception at approximately day 21, when media begins reporting. The same increase is observed for DEBN, though later in the simulation (day 25), and with a less dramatic increase in the number of individuals perceiving risk. It should be noted that the BN models are trained endogenously during the simulation, and agents learn to perceive risk and cope with cholera diffusion. This implies that they also learn (especially for EEBN) that the epidemic is over and risk levels are declining. This is evident in the decrease in risk perception after day 45.

Agents learning and Survey data

After running the experiments, we evaluated the outcomes of the CABM by dividing agents into groups according to their level of risk perception.

Table 5-2 presents the percentage of agents in the CABM who perceive risk, and compares it to the percentage of survey participants who reported perceived risk in a similar situation. Both DDBN (85%) and EDBN (95%) overestimated the risk perception for VP only, which was 49% in the survey. This disjunction between the actual percentage and simulated percentage is likely due to the BNs being trained prior to the start of the simulation. Based on the training data, agents learned that when visual pollution exists, the RP should be positive. The values of EEBN (46%) and DEBN (53%) for only VP are more in line with the real values (49%).

Table 5-2: Comparison of agent risk perception per risk factor with the original survey data

Risk Factors	Survey				
	% RP	% Risk Perception		% Risk Perception	
		DDBN	**EDBN**	**EEBN**	**DEBN**
VP	49,1	85,3	95,3	46,2	53,3
Memory	42,7	0	0	15,9	0
Media	63,8	60,2	62,5	59,2	44,5
Neighbours	61,3	100,0	100,0	56,7	78,3
VP + Memory	53,9	17,1	25,2	58,9	14,6
VP + Media	90,1	89,6	82,3	72,8	37,5
VP + Neighbours	86,3	66,7	71,9	68,7	66,5
Memory + Media	81,1	76,0	60,9	63,6	64,7
Memory + Neighbours	80,0	98,0	100,0	65,1	68,5
Media + Neighbours	75,8	100,0	100,0	72,5	100,0
VP + Memory + Media	87,1	73,5	77,7	79,2	54,7
VP + Memory + Neighbours	83,7	94,3	90,0	73,7	70,0
VP + Media + Neighbours	78,9	73,8	73,5	78,2	100,0
All factors	NA	92,3	69,7	52,1	66,7

For Memory, the opposite result was evident. In the survey, 42% of participants perceived risk; however, no risk was observed due to these

BNs. In this case, a value of zero percent indicates that when agents have been using river water before the cholera outbreak, they will continue doing so during the cholera outbreak because they trust the source (no perceived risk).

The effect of neighbours communicating experiences of illness in the BNs leads to 100% risk perception in the ABM, while risk perception was only 62% in the survey. For the two BNs – DDBN and EDBN – illness in neighbouring households was a strong indication of possible risk, which explains the high score. The scores for neighbours for EEBN (57%) was much closer to the actual scores of 61%. The scores for VP combined with Memory of DDBN and EDBN are much lower than the survey scores. This is explained in the same way as the score for Memory only. When people have been using the river water, they trust the water and continue using it. Again, EDBN and EEBN score much better.

It is evident that memory (remembering the prior use of a water collection point) has a negative impact on the risk perception. Agents perceive risk during their communication with neighbours and via reporting of media, though they might change their mind via their memory, and continue to use water from a particular location. In addition, during the simulation runs, communication with neighbours had a strong impact on risk perception, although media had a stronger effect in the survey.

For media, the results for DDBN and EDBN are good. Media scores were less robust for DEBN (44% versus 64% in the survey), but were good for EEBN.

Overall, the results obtained from the implementation of EEBN were closer to the data from the surveys than the results of DEBN. This can be explained because the values of CPT in EEBN have been driven by the literature and census data for Kumasi, Ghana. Additionally, the network structure is driven logically from expert knowledge. Again, in both implementations, Memory has a negative impact on the risk perception process. Whenever Memory exists, either alone or in combination with other factor(s), it will lower the RP. Furthermore, VP may also play a negative role in DEBN implementation. This confirms that survey data helps in training BNs, though the agent learning process can be better controlled with the support of expert knowledge.

An implementation using different training datasets per community or water collection point may reveal divergent learning patterns and risk perception throughout the area. Furthermore, in the CABM, each agent calculates their risk level and compares it to a threshold value of 0.5, which determines if risk is perceived or not. In reality, the threshold level will vary by individual, as some are more sensitive to risk than others. Introducing greater heterogeneity among agents by varying the risk threshold value per household may influence the results of the simulation. Such combinations of agents pursuing intelligent decisions in spatial ABMs based on ML algorithms opens opportunities for overlaying spatial, socio-economic, and cognitive heterogeneity in a range of applications.

Spatial Patterns

To assess the influence of running CABM with different BN models, the outcome of the spatial patterns of each simulation is presented.

Figure 5-5 shows the risk perception per community (size of the circles) and the variables the risk perception is based on (colours). The river shows the average VP value scaled from low (green) to high (red). The total risk perception is highest in the EEBN model, however, risk perception in the northern communities in this model is low. This might be due to the fact that most of the household agents that live there are of high-income and can buy bottled-water during the epidemic, or to the fact that the VP level is low.

There is considerable variation in risk perception between the different communities. Risk perception is not always related to the income levels. For example, in the DEBN model, the northern community (18) has a relatively high risk perception but is completely located in a high income and low VP zone. Communities score differently in different models. For community 17, we see low risk perception in DEBN and EDBN but higher risk perception in DDBN and EEBN. In this case, this is not steered by data-driven versus expert-driven models. Community 5 scores a very high risk perception in the EEBN model, has a high score in the DEBN model, but scores lower in risk perception in the DDBN and EDBN. In general, a fully expert based model (EEBN) leads to the highest risk perception followed by the model that is based on expert probabilities for node states. From this we can derive that the design of the model has less impact on the total risk perception compared to the node state probabilities.

The second aspect are the sources the risk perception is based on. This is indicated as colours in the pie-charts of Figure 5-5. Household agents in the DDBN model pay more attention to the combination of media and contacts with neighbours (green colour). This also applies to the two communities in the north (16 and 18) and the community in the south-east (13). This can be explained by the fact that in these upstream locations, the VP is low. CABM combined with an EEBN model consider VP (alone or combined) to be more effective as a household agents risk indicator. In addition, risk perception with EEBN represents a more balanced risk perception between all factors (Figure 5-5).

Household agents who live in communities 7 and 9 which are located downstream perceive more risk based on VP (red colour of the river represents high VP). This is because more waste accumulates in the river and on the river banks. This also applies to communities 5 and 11 in the EEBN model.

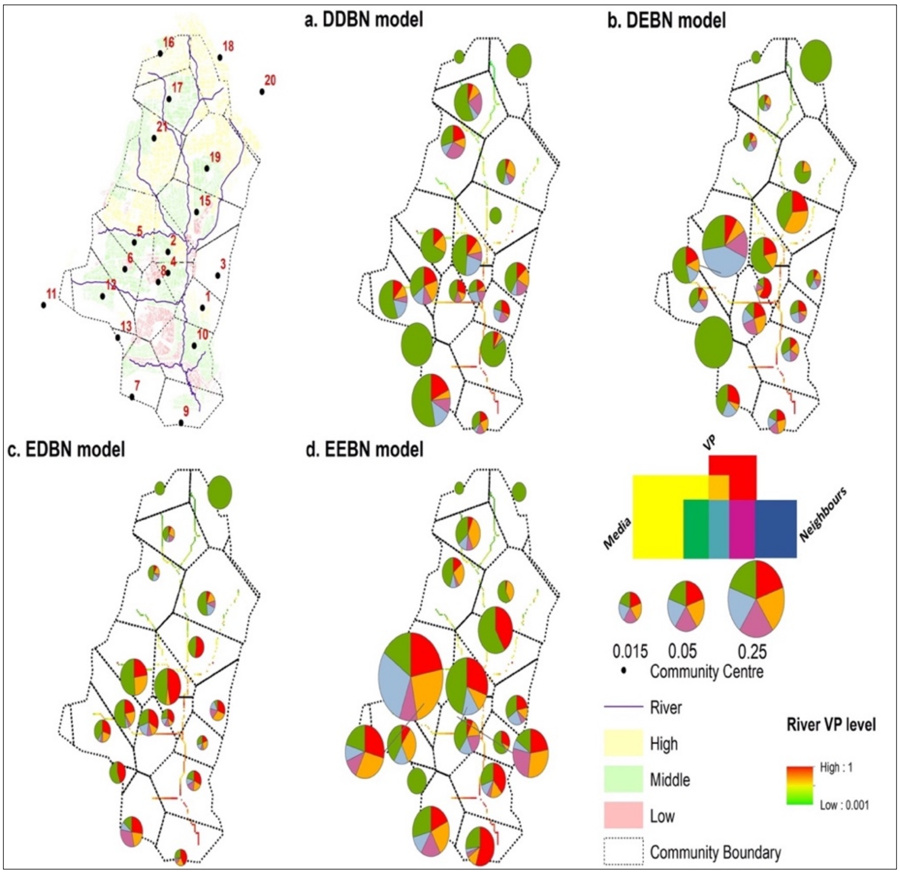


Figure 5-5: Spatial distribution of risk perception based on different variables

5.4 Conclusions

Although several studies have combined ABMs and BNs, no comprehensive overview exists on the advantages and disadvantages of different integration approaches and their impact on the output of the models. To explore how learning in spatial ABMs could be realized in the absence of large behavioural datasets, we tested four different ways to design and train BNs to enhance agent cognitive abilities in a spatial ABM. We constructed BN fully data driven, fully expert-knowledge driven and as a mix of data and expert-knowledge, trained prior to running BNs in the ABM or during a simulation run. For models that are data-driven (DDBN and EDBN), the results of both prior trained models show similar percentage of risk perception. Expert-driven models (DEBN and EEBN) outperformed the data-driven ones. It indicates, that supervised learning, which aids training of a BN algorithm with the support of expert knowledge, provides more control over the learning process and offers a logical framework. Expert-knowledge method helps to avoid intense overfitting and enables direct model comparison since it computes a full posterior distribution of the BN. A development of an efficient BN requires a combination of data and expert knowledge (Fenton and Neil, 2012), as has been also illustrated in other applications in ecology (De Waal et al., 2016), sports (Constantinou et al., 2012), robotics (Park and Cho, 2012) and medicine (Yet et al., 2013). Hence, instead of deriving BNs directly from data, we advise an expert interpretation and construction of BNs based on expert logic.

We measured the goodness of fit for the four BNs using survey results and the output of CABM as test data. However, goodness of fit scores for these BNs did not differ greatly.

The highest-scoring BN according to goodness of fit is not necessarily the same as the highest-scoring based on e.g. the risk perception curve. We also had two networks with the same structure (DDBN and DEBN), yet leading to different results when combined with our ABM. This confirms that the structure of the BN is the factor that least impacts the final outcomes. Parameterisation and the way the network is trained (prior or during simulation) play a more important role and have more impact on the final model outcomes. We also observe that training prior to a simulation run leads to “overly intelligent agents”, with high risk perception at initialization that does not decline even in the absence of cholera reports. The choice between prior training and training during simulation runs is individual for each application. In our case, although Kumasi citizens had previous experience with cholera before 2005, hence they were not prepared. For other applications, a certain level of risk awareness may be essential at the start of the simulation, demanding prior training of BNs.

As conditions differ per community, it is logical that risk perception differs spatially. This applies to the level of risk perception, but also to the factors contributing to this risk perception. The expert driven model, DEBN and EEBN, provided the most balanced risk perception (based on all risk factors). Risk perception and the process of making a decision are complex processes combining spatial and social factors. However, less implementations are available integrating ML for assessing risky situation engaging agents' risk perception due to the lack of gather data on people's risk perception and little is known about spatial risk detection especially in developing countries. Furthermore, risk perception based on spatial learning is not easy to measure. Limited datasets are available about the way the spatial environment influences the human decision making. Most researches that discuss risk perception will evaluate how risk perception varies in space but not which role the environment itself plays in the process of feeling scared.

Our research uses BNs as the ML algorithm because this method can combine expert knowledge and use small behavioural datasets. However, there are other ML methods with the same characteristics. The workflow presented in this paper can also be used for example using, Genetic Algorithms, Decision Trees, linear SVM, and Naive Bayes.

Chapter 6: Risk perception and behavioural change during epidemics: comparing models of individual and collective learning[8]

[8] This chapter is based on an under-review paper (paper 3) under the same title. The paper is authored by Shaheen A. Abdulkareem (the main writer of the paper and author of this dissertation), Dr Ellen-Wien Augustijn based at the University of Twente, Professor Tatiana Filatova based at the University of Twente, Katarzyna Musial at the University of Technology Sydney, and Dr Yaseen T. Mustafa based at the University of Zakho. The paper has been submitted to PLOS ONE journal

6.1 Introduction

Epidemics have always been a source of concern worldwide, especially in developing countries. Therefore, good responsive and preventive strategies both at the individual and government level are vital to saving lives. Most of these strategies depend on the behavioural aspects of choice and complex interactions among people (Ruland et al., 2015). Perceiving the risk of infectious diseases may lead people to change their behaviour spontaneously, as shown during the epidemic of SARS in 2003 (Tan et al., 2004). People change behaviour and adapt to protect themselves based on the information they receive about the disease (Zhao et al., 2015). Gathering information and experience through multiple sources is essential for increasing disease risk awareness and taking protective measures (Williams et al., 2010). To fight epidemics effectively, we need advanced tools that enable us to understand the factors that are contributing to the spread of information about life-threatening diseases and influencing changes in individual behaviour that curbs a disease's diffusion.

Simulation tools are commonly used in assessing policy impacts in the health domain. Different types of models are used, varying from mathematical models (Kerkhove and Ferguson, 2012) to spatial games (Zhao et al., 2015). Boulanger and Bréchet in their evaluation of six paradigm models for policy-making recommended agent-based modelling (ABM) as the most promising modelling approach (Boulanger and Bréchet, 2005). ABM is widely used to understand the dynamics of epidemics (Pizzitutti et al., 2018; Venkatramanan et al., 2018; Tang et al., 2017). ABMs can be used to study the dynamics of complex systems, where many heterogeneous individuals learn from their experience and their environments, interact with each other and make decisions. Being a primary bottom-up method, ABM can represent micro/macro relationships, accommodate agents' heterogeneity and their adaptive behaviour. As such, ABMs assure explicit feedbacks between the spatial environment and cumulative agents' behaviour and can integrate a variety of data inputs including aggregated, disaggregated and qualitative information and data (An, 2012; de Marchi and Page, 2014; Filatova et al., 2013; Fonoberova et al., 2013).

In disease modelling, two elements are essential in representing agents' health behaviour: (i) evolution of risk perception, and (ii) selection of a response strategy. Hence, the core of a disease ABMs is in defining a learning method that is used to steer the risk perception and risk coping behaviour of agents. In both, the sensing of information (global, from the environment and from other agents), exchange of information (between agents) and processing of information (decision making) are central. Machine learning (ML) techniques are good in all three, and offer a more realistic way of adjusting agents' behaviours in ABMs (Abdulkareem et al., 2018). As more data becomes available in the area of disease spread analysis, a new research direction has emerged – supporting ABMs with data-driven approaches. ML has a potential to enhance the performance of ABMs, especially when the number of agents is large (e.g., pandemics) and the decision making of the agents is complex (e.g. depending on past experiences and new information from the environment and peers).

The purpose of using ML approaches in the context of ABM is to provide agents with the ability to learn by enabling them to adapt the decision-making process according to the available information. Human beings make decisions both individually and as part of a collective, where the individual copies the decision taken from a group or group leader (Carlson et al., 2014).

Further, information about social networks formed by people is becoming more readily available, e.g. by extracting social connections based on social media, and reveals collective behaviour in many application domains, including health (Tang and Liu, 2009). For example, concerning vaccinations, people are not entirely rational but imitate others, leading to group behaviour (Mbah et al., 2012). Many ABM models rely solely on individual behaviour, yet group emotions and group behaviour also need to be captured (Li et al., 2014). Bosse et al. presented an ABM that models collective decision making in crowds and groups, in which they integrated interacting emotions, beliefs and intension and social contagion (Bosse et al., 2013). The purpose of their simulation was to discover the impact of mirroring emotions, beliefs and intention on individuals' behaviour. Agents may learn in isolation or through interactions e.g. with neighbours (Sen and Weiss, 1999). In isolated learning, the agent learns independently without requiring any interaction with other agents. In interactive learning, several agents are engaged in the same process of learning, and they need to communicate and cooperate to learn effectively. Interactive learning can be conducted in multiple ways (based on different social learning strategies) (Eberlen et al., 2017). Agents might be represented as members of local groups (small social networks), learning together and copying behaviour from other group members (Collins et al., 2014). The impact of different types of group learning compared to individual learning is an underexplored domain in the development of ABMs. Most developers are not aware of the impact of their choices on the model's results, at times assuming that group learning is computationally attractive.

This article evaluates the influence of individual vs. collective learning on an epidemic's dynamics within a disease ABM. We will pursue a quantitative test on the influence of agents' ability to learn - individually or in a group - on the dynamics of a disease.

The main goals of this article are to (1) simulate the learning processes in agent groups that reflect a gradient of learning (from individual to collective), and (2) understand how these learning processes can help in obtaining more insights into the dynamics of social interactions and their emergent features during an epidemic. To address these objectives, the article aims to answer a number of research questions: (1) What is the impact of social interactions on individuals' decisions and behaviour? (2) How do individuals learn when they are in groups? (3) How do different forms of implementing ML in groups affect the group's learning processes? (4) Is there an impact of implementing group learning for risk evaluation only or for also coping and self-protection during epidemics? And (5) do individuals perform better at coping and self-protection during epidemics compared to groups?

In this research, we employ a spatially-explicit ABM of cholera diffusion (Augustijn et al., 2016) as our case study to show the implications of our research. We use BNs to steer the behaviour of agents by representing risk perception and coping appraisal using a cholera model for Kumasi, a large city in Ghana (Abdulkareem et al., 2018) . We extend previous work by conducting eight scenarios, in which we run eight models to test individual versus group learning in combination with different information sources (including contact with other agents) as factors in the BN. We investigate how the epidemic spread depends on different learning approaches used for risk perception and coping decisions in the face of outbreaks.

6.2 Methods

This chapter aims to compare the models of individual and collective learning during the processes of risk perception and making decisions about how to cope with the situation during an outbreak. The next subsections present the principles of individual and collective intelligence used here. In addition, the description of the CABM and models' setup and measures used to evaluate the outcomes are presented.

6.2.1 From Individual to Collective Intelligence: Defining the Gradient of Learning Strategies

A feeling of risk in people is triggered by the amount of information communicated, its type and the attention to specific information (fear) that stimulate to produce the learning of new responses (Rogers, 1983). Communication with information sources helps individuals to estimate the severity of the emerging event, the probability of being exposed to infection, as well as evaluating the efficiency of their coping responses.

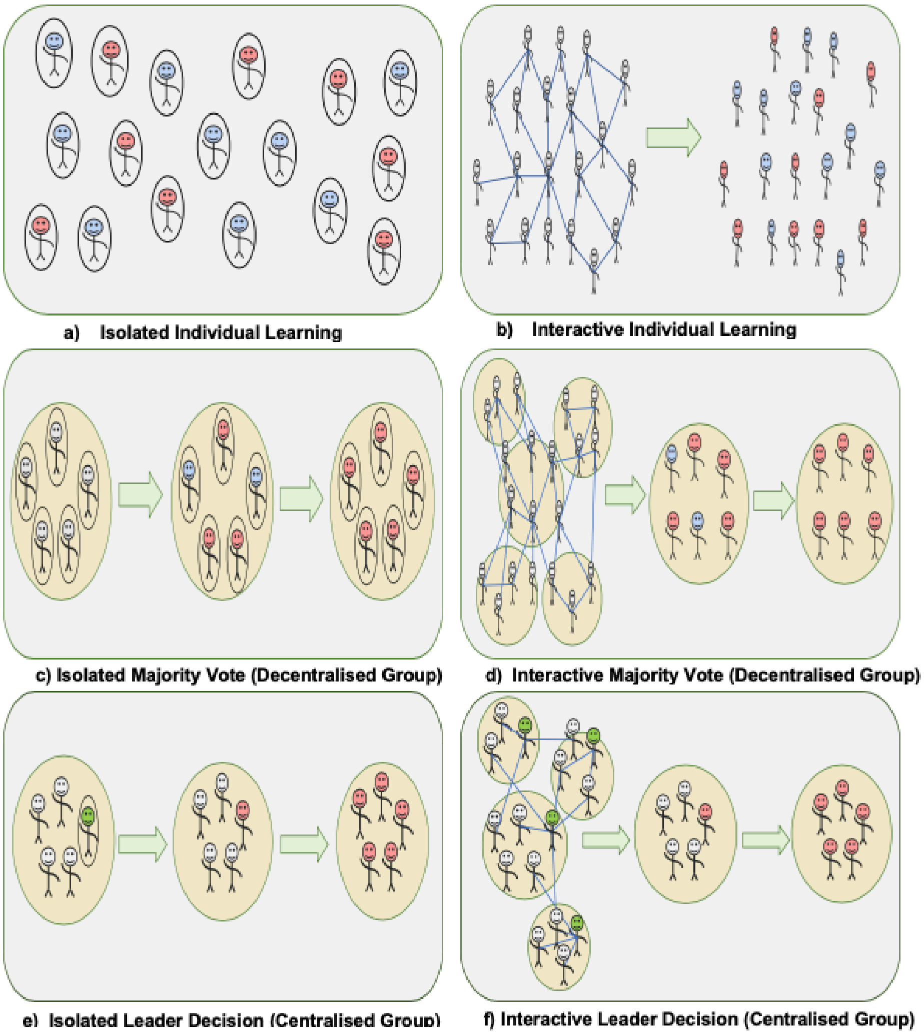


Figure 6-1: Agents' learning types in agent-based models. Blue ties, wherever exist, refer to communication with other agents who might be either in the same or in another community;

leaders in centralized groups are marked in green, and red/ blue colour denote different learning outcomes (i.e. Risk/no Risk in Risk Perception stage, and any of the Coping Appraisal stage decisions)

When agents **learn individually** (Figure 6-1.a and b), their learning depends on their prior knowledge (memory, experience, and/or the perceived risk awareness of the environment, such as visual pollution). The learning, in this case, is the process of gaining skills or knowledge, which an agent pursues individually to support its individual task (Russell and Norvig, 2010). **Group learning** is the process of acquiring new skills or knowledge that is undertaken collectively in a group of several individual agents and driven by a common goal (Sen and Weiss, 1999). Group learning can be realised by making all group members use their own ML algorithms to gather information to perform a specific sub-task (**decentralised**, Figure 6-1.c and d), and then pool their opinions collectively by making one decision for the entire group. Here, we adopt a 'majority vote' as the resolution mechanism in decentralized group decision-making. Alternatively, group learning can also be realised by introducing one agent (**leader**) who uses ML to learn for the whole group to help it accomplish its group task (**centralised**, Figure 6-1.e and f). In centralized group learning, agents in the group copy the decisions of their leader. In both cases, all agents that belong to a group will share the same decision, but the information this decision is based on varies considerably.

Both individuals and groups may learn by either taking information from their social networks - i.e. have it as an additional source of information in their ML algorithms - or not. When individual agents are **isolated** learners (Figure 6-1.a) they do not have a social network but use only their own information to make a decision. When individuals learn in an **interactive way** (Figure 6-1.b), they acquire new skills or knowledge by perceiving information, experience, and the performance of other agents via their social network. Like individual agents, groups can also learn in isolation or interactively. In **isolated learning**, a group of agents learns independently without exchanging any information with others (Figure 6-1.c and 6-1.e). In **interactive learning**, groups of agents communicate with their neighbours to learn effectively. Here, the group members interact together and with agents outside of their groups and share their experiences to improve their common group skills (Figure 6-1.d and 6-1.f).

6.2.2 Simulation Scenarios: Individual vs Group Learning

We designed eight simulation scenarios to answer the research questions: explore the influence of individual vs. group (1), centralized vs. decentralized (2) and isolated vs. interactive (3) learning in processes – during both the risk perception (RP, BN1) and coping appraisal (CA, BN2) stages (4 and 5) - on the epidemic's dynamics and the model's performance (Table 6-1). Given the importance of communication with neighbours on the geographic basis in cholera diffusion, the groups here are spatial: household agents living in the same community, with the same education and income levels, belonging to the same group. We systematically vary the ABM settings following the steps in Figure 6-2 to change the gradient of intelligent learning (Steps 2 and 3) in different cognitive stages corresponding to our decisions of interest: risk and coping appraisal (Step 1).

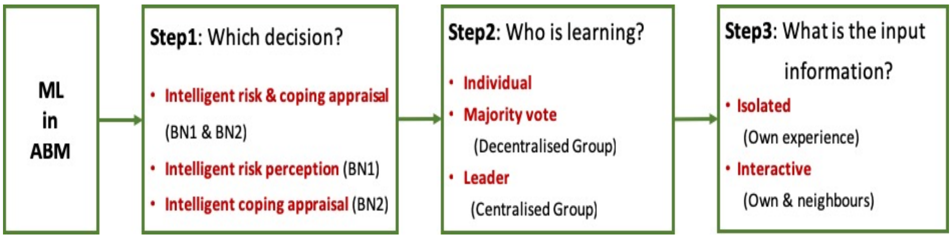


Figure 6-2: From individual to collective intelligence in ML-based ABMs

Table 6-1 shows the setup of the eight scenarios that reflects the three stages shown in Fig 6-2.

Table 6-1: Simulation scenarios

Scenario	**Decision that relies on ML**	**Agent that employs ML**	**Influence of others on ML input information**	**Commentary**
M1: **RP&CA (In-I)**	RP and CA (BN1 & BN2)	Individual (In)	Isolated (I)	An individual uses ML to update her risk perception and to take protective actions only based on his individual

				experience neglecting any communication with others (Fig 1.a)
M2: **RP&CA (In-N)**	RP and CA (BN1 & BN2)	Individual (In)	Interactive with neighbours (N)	An individual uses ML to update her risk perception and to take protective actions based on his individual experience as well as based on past disease experiences of peers (Fig 1.b).
M3: **RP&CA (D-I)**	RP and CA (BN1 & BN2)	Majority vote (M) (decentralized group)	Isolated (I)	All agents in a group use MLs to make decisions without taking experience of others into account. The final decision on RP and CA is defined through the majority vote (Fig 1.c).
M4: **RP&CA (D-N)**	RP and CA (BN1 & BN2)	Majority vote (decentralized) (M)	Interactive with neighbours (N)	All agents in a group use MLs to make decisions taking experience of others into account. The final decision on RP and CA is defined

				through the majority vote (Fig 1.d)
M5: **RP&CA (L-I)**	RP and CA (BN1 & BN2)	Leader (L) (centralized group)	Isolated (I)	Each agent group randomly chooses a leader who uses ML to make a decision. The leader decides in isolation without communicating with others; all group members mimic his decisions (Fig 1.e)
M6: **RP&CA (L-N)**	RP and CA (BN1 & BN2)	Leader (L) (centralized group)	Interactive with neighbours (N)	Each agent group randomly chooses a leader who uses ML to make a decision. The leader considers disease experience of others in his group and outside; all group members mimic leader's decisions (Fig 1.f).

M7: **RP(D-N), CA (In-N)**	RP (BN1) as in M6 CA (BN2) as in M2	RP: Majority vote (M) (decentralized group) CA: Individual (In)	For both RP and CA: Interactive with neighbours (N)	Taking past experience of others into account, all agents in a group use own BN1 to decide if disease risks are real. The group members vote to evaluate the final risk perception for all group members (RP as in Fig 1.d). Everyone individually assesses own self-efficacy regarding disease prevention actions (CA). They run individual BN2 while considering past experience of others (CA as in Fig 1.b).
M8: **RP(L-N), CA (In-N)**	RP (BN1) as in M4 CA (BN2) as in M2	RP: Leader (L) (centralized group) CA: Individual (In)	For both RP and CA: Interactive with neighbours (N)	Each agent group randomly chooses a leader who uses ML to decide whether the disease risk is real (RP). The leader considers disease experience of others in his group and outside; all group members mimic the leader's

				RP decision (RP as in Fig 1.f). Everyone individually assesses own self-efficacy regarding disease prevention actions (CA). They run individual BN2 while considering past experience of others (CA as in Fig 1.b).

M7 and M8 are suggested here to evaluate the impact of group learning for coping decisions and whether there are any consequences or advantages on risk perception.

6.2.3 Output Measures of the CABM Enhanced with Bayesian Networks

To evaluate the impact of different types of individual and social intelligence on agents' learning processes regarding risk perception and coping appraisal and the resulting disease spread patterns, we use four measurements: disease diffusion, risk perception, spatial patterns, and model performance. All will be discussed below.

Disease Diffusion

The epidemic curves, duration of infection, total cases, and peak days are the most common measures of disease diffusion.

Epidemic curve is a graph representation of the distribution of infected cases over an epidemic period (Wilson and Burke, 1943). It is a useful way to assign the type of epidemic, calculate the difference between the minimum and maximum incubation period and determine the possible time of exposure.

Having the real data, one can validate the simulations. The real dataset available for this study belongs to the outbreak of cholera in 2005. The cholera epidemic started in September 2005 and lasted three months. The period from September to December is the rainy season in Kumasi. We don't know exactly on what day cholera emergent and ended but the first recorded date was 26.09.2005 and the last recorded cases were on 12.12.2005. In addition, data were recorded in discrete time steps during that period, and contains cases of those whom visited the hospital and health centres in the region. However, if we compare the date of recorded cases in reality to the date in the simulated models, we will find that earlier dates are recorded in the simulation, since the model immediately registers the ill agents at the moment they caught the disease. Nonetheless, in reality, in most countries, the reporting systems of infectious diseases are deficient in their infrastructure, as well as time accuracy that requires attention and efficient care to improve them (Janati et al., 2015). Moreover, the delay in recording infected cases might be due to the ill people themselves, who are not going to hospitals or to physicians who are not aware of the importance of reporting cases accurately and rapidly (Reijn et al., 2011). This can also be the explanation for the rest of the values regarding the duration of infections and total cases.

Risk Perception

There is less known about the risk perception of infectious disease compared to other research fields, such as environmental risks (De Zwart et al., 2009). Data on risk perception is rarely collected in early stages of epidemics, especially in developing countries (Liao et al., 2017). Literature often reports the percentage of people who perceive risk after an epidemic is over (Yang and Cho, 2017). Some literature focused on other infectious diseases (Fritzell et al., 2018) and/or pays more attention to other factors of risk perception (Kim and Kim, 2018), while others are using online games to test the responses of people during epidemics (Chen et al., 2013). Here, using the simulated data, we measure risk perception as the percentage of agents who perceived risk on a given day and plot this as a risk perception curve. We propose an assumption that the risk perception peak and epidemic peak should be correlated, which has been proven by literature even if the epidemic occurs outside the country (De Zwart et al., 2009; Rübsamen et al., 2015; Zhao et al., 2018; Sridhar et al., 2016).

Spatial Patterns

We assess the accuracy of our eight models using R^2, which calculates the spatial distribution of infected cases in both real dataset and the outcomes of the simulations (Augustijn et al., 2016) to understand how well each model (with each type of learning) complies with the real number of cholera cases per community. These were confirmed by bacteriological tests and were registered by the Disease Control Unit (DCU) in Kumasi Ghana during the epidemic of 2005. A simulation with $R^2 = 1$ indicates a perfect reproduction of the actual epidemic[9]. R^2 is calculated by:

$$\boldsymbol{R^2} = \frac{\sum_{i=0}^{n}\left(\left(GD_{s,i} - \overline{GD_s}\right)\left(GD_{d,i} - \overline{GD_d}\right)\right)}{\sqrt{\sum_{i=1}^{n}(GD_{s,i} - \overline{GD_s})^2 \sum_{i=1}^{n}(GD_{d,i} - \overline{GD_d})^2}} \qquad \textbf{6-1}$$

Where GD is the relative percentage of diagnosed diseases per community, s refers to the simulation scenario and d to the real cases of 2005. The index i refers to community i and n is the total value of communities. The GD is calculated by:

$$GD = \frac{Number\ of\ cases\ in\ community}{Total\ number\ of\ cases} \times 100 \qquad \textbf{6-2}$$

The study area in our CABM consists of 21 communities. Eleven communities are completely inside the study area and the other ten are partially included.

In addition, we will show maps, in which we present the spatial distribution of coping appraisal decisions type per community per model. This is to show the impact of input information and learning type on the decision that household agents take during the process of coping appraisal.

Model Performance

The CABM is implemented in Netlogo (version 5.2.0), BN1 and BN2 are coded using R statistical language. During each model run, household agents in Netlogo collect their BNs' inputs (e.g. estimate pollution levels, retrieve own memory, check own health status, etc.), and call their BNs via the R extension of Netlogo. After processing the BNs, R returns the information on risk perception and disease coping strategies back to the agents in Netlogo. Here, we measure the time required for each model run to be completed as a measure of its performance.

[9] While we do have real data on recorded cholera cases in Kumasi's 2005 epidemics, the data quality is far from perfect. It is likely to under-represent the extent of the epidemic. Still, this is the best of what exists in the context of a developing country.

6.3 Results and Discussion

Given the stochastic nature of ABMs, we ran each of the eight models 100 times, and for every 10 runs, a new synthetic population was generated. We report the average and standard deviation of the results across 100 runs for each model for our metrics (Section 6.2.3) in Table 6-2.

Models M5, M6, M7 and M8 record a longer duration of active infection during the epidemic (75-79 days), which is closer to the real duration of the epidemic in 2005 (75 days). M5, M6 and M8 apply centralized learning, while M7 applies decentralized learning, but only for risk perception. M2, which is individual learning with social interactions, also records a long duration when compared to the real data of 2005 (68 days in M2). However, isolated learning and decentralized learning for both risk perception and coping appraisal records lower values for duration of the epidemic, with an average difference of -25% of the real duration.

All eight scenarios give more infected cases than the empirical data. This is because infection with cholera bacteria leads to a clinical spectrum that ranges from asymptomatic cases to symptomatic cholera cases. Asymptomatic cases are not reported but represent roughly half of all cases (Harris et al., 2008). In our simulations, we are not differentiating between symptomatic and asymptomatic, all infected cases are considered to be symptomatic cases. Therefore, for validation purposes, in Table 6-2, we reported that 57% of the total infected cases occurred in running the eight models, as has been assumed in (Harris et al., 2008).

Table 6-2: Validation measures of the eight scenarios

Scenarios \ Measure		Duration (days)	Total of infected	Peak day - Epidemic	Peak value - Epidemic	Peak day - Risk	Peak value - Risk	Run time (minutes)	R2
Real data (2005)		75	1621	42	181	N/A	N/A	90 days	1
M1: RP&CA (In-I)	value	55	2457*	35	232	88	501	85	0.65
	SD	2	195	1.3	30.12	1.9	103	3.1	
M2: RP&CA (In-N)	value	68	2279*	35	209	38	481	95	
	SD	0.6	113	0.96	18.4	2.3	98	1.1	0.66
M3: RP&CA (D-I)	value	58	3355*	37	345	90	501	90	
	SD	3	402	2.5	83.2	0.4	233	2.7	0.62
	value	55	3149*	36	320	85	708	125	0.61

M4: RP&CA (D-N)	SD	1.8	268	0.97	60.4	1.7	265	4.2	
M5: RP&CA (C-I)	value	79	2851*	37	215	44	676	26	0.7
	SD	2.13	243	1.5	26.8	1.2	114	0.15	
M6: RP&CA (C-N)	value	79	3071*	38	210	44	456	35	0.64
	SD	3.8	105	2.4	41.7	0.96	198	2.1	
M7: RP(D-N), CA (In-N)	value	77	2911*	37	307	89	610	72	0.61
	SD	1.65	78	1.62	14.5	0.92	122	3.2	
M8: RP(C-N), CA (In-N)	value	75	2107*	37	136	44	462	45	0.75
	SD	0.64	129	1.6	22	1.2	221	1.7	

(*) representing 57% of total infected cases

M8, which uses centralized learning for risk perception and individual interactive learning for coping appraisal, reports the least number of infected cases (1840 against 1621 in reality), followed by M2 (individual social learning) with 2000 cases and individual isolated learning (M1) with 2156 occurrences. These three values reflected the fact that when household agents learn to cope and make decisions individually, this is more efficient than being in groups. Moreover, when these decisions are combined with social interactions, they lead to better protection (M2 and M8). In general, group behaviour has a negative effect, although centralized groups have a less negative impact compared to decentralised ones. Finally, in M7, where household agents learn in decentralized groups for risk perception and individually, learn to cope, 2554 infected cases are recorded (Table 6-2). This is because their engagement in decentralized groups for risk perception guides them to less risk perception, which does not motivate them to change their behaviour to more protective alternatives.

The R2 of M8 reports the closest spatial distribution of the infected cases over the communities (0.75) compared to the real data, followed by M5, with 0,7.

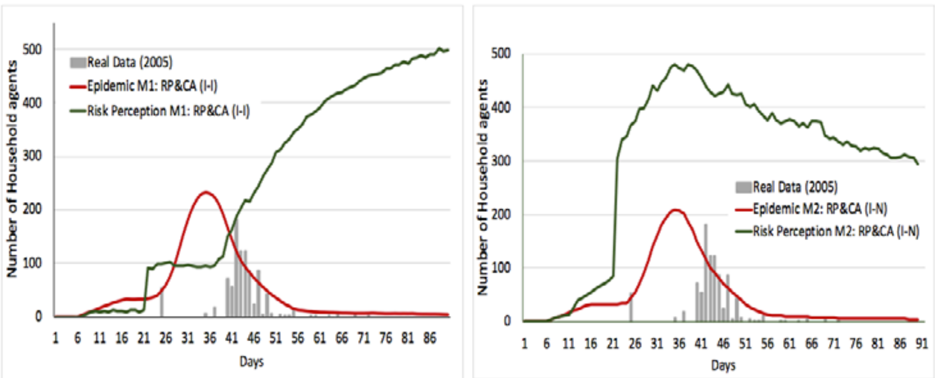
The correlation between the peak of the epidemic and the peak of risk perception reflects on the responsiveness of the household agents' risk perception to the epidemic. Scenarios M2, M5, M6, and M8 are more responsive. The peak of RP in M2 comes after three days of its epidemic peak, and the peaks in M5, M6, and M8 come after seven days of their epidemic peaks (Table 6-2). M1, M3, M4 and M7 show their peaks for risk perception approximately at the end of the simulation time. Individuals in

M1 are isolated, along with individuals in M3; therefore, they keep following their normal behaviour of fetching water and use it as it is. In M4 and M7, household agents depend on majority votes in their groups to make their decision on risk and changing behaviour. More explanations can be found visually in next sections.

In general, models with centralized learning require the shortest computation times. M5 records the shortest runtime (best model performance). This is because only the leaders in the centralised learning consult their BNs and they are isolated, which reduces the time required for calculations. On the other hand, M4 records the highest computational time due to intensive computations needed in two layers: individual agents' layer and the decentralized group layer.

6.3.1 Making Decisions Individually does not pay off

In the M1 scenario, individual household agents evaluate the risk of getting cholera and make a decision relying only on their own experience (i.e. each has individual BN1 and BN2 and does not communicate with neighbours). Scenario M2 extends this stylized isolated benchmark case by assuming that while agents continue to make decisions individually, they do share information with neighbours about the perception of risk and protective behaviour (both BN1 and BN2 take the experience of neighbours as one of the information input nodes). Figure 6-3 shows the epidemic curves and the dynamics of risk perception in both scenarios.



a) Epidemic vs Risk Perception Curves (M1) b) Epidemic vs Risk Perception Curves (M2)

Figure 6-3: Epidemic curves (in red) and Risk perception curves (in green) for scenarios M1 and M2

In the absence of social interactions, more agents became infected with cholera. The peak of the epidemic curve in M1 (In-I) is higher than that of

M2 (In-N), leading to 11% more disease cases (Figure 6-3 and Table 6-2). Overlaying risk perception and epidemic curves suggest that when agents make decisions in isolation (M1: In-I), the dynamics of risk perception is hardly realistic (Figure 6-3.a). Namely, when the epidemic is at its peak, household agents in M1 respond very slowly, with BN1 delivering a wrong evaluation of risk perception (Figure 6-3.a). However, they start to be aware of the risks very late: when the epidemic vanishes, the number of agents with risk perception = 1 keeps increasing. In the absence of communication and experience sharing among peers (In-I), information about disease spreads slowly and there is a significant time-lag between the occurrence of the disease and people's awareness. The small stepwise increase, around day 21, is due to the fact that the media starts to broadcast information about the epidemic and that day.

In M2, household agents behave according to the expected pattern, when RP first becomes amplified by the media and social interactions and then vanishes as disease cases become rare (Figure 6-3.b). Only those who experience cholera infection in their households remain alert. Household agents in M2 after day 21 have more responses to the media's news compared to isolated agents. Media support their social interactions with their neighbours, which leads to more agents perceiving risk, especially when the number of infected cases increases and reaches its peak (Figure 6-3.b). Still, even in M2, the limitations of making decisions about risk perceptions individually remain: RP falls too quickly, implying that people stop worrying about the epidemics despite the fact that it continues.

Since household agents in M1 do not have interactions with other agents, running the model requires less time to complete compared to M2 (10% increase in performance, Table 6-2). The interaction between household agents requires time to process the information exchange between agents. In addition, both (In-I) and (In-N) are approximately the same in terms of realistic spatial distribution of infected cases over the communities, with values of 0,65 and 0,66, respectively (Table 6-2).

In Figure 6-4, we presented the spatial distribution of decision types over the study area in both M1 (In-I) and M2 (In-N). The household agents in isolated learning are not aware of the cholera-infected cases in their neighbours' household. Household agents in M1 are taking an unsecured decision and trust more on using the water fetched from the river as it is (D1 in Figure 6-4.a). Household agents in M2 are more rational and mostly go for boiling the water that they fetched from the river before using it (D3 in Figure 6-4.b).

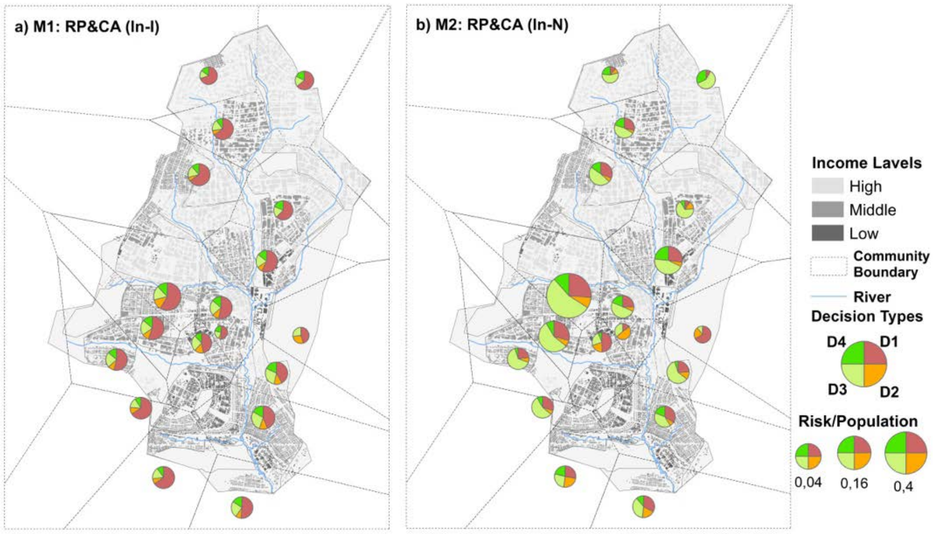


Figure 6-4: Spatial distribution of different coping appraisal decisions of scenarios M1 and M2; the size of the pie represents the size of household agents with RP = 1 over the community population

6.3.2 Majority Votes is Imperfect

In decentralized learning, groups of household agents vote for risk perception and coping appraisal. The final decision of the group is the output of the majority votes. Thus, all group members follow the final decision of the group. These groups represent the democratic system, which depends very much on the composition of the group. The decentralised groups with majority vote can lead to negative risk perception. Besides, the coping appraisal that depends on a majority vote leads to inappropriate decisions regarding protection from cholera. When individuals are engaged in social groups, their behaviours were not independent anymore (Zacharias et al., 2008). This leads to an increase in the randomization of decentralized learning models (M3 and M4), which can be seen in Table 6-2, with a higher standard deviation of these two models in all measures.

The qualitative patterns of the three scenarios (M3, M4, and M7) is the same irrespectively of the social interactions that add new information to ML (Figure 6-5). For the development of the disease, the voting mechanisms seem to overwrite individual judgments. The M3 scenario assumed that household agents are isolated during the processes of risk perception and coping appraisal. While both M4 and M7 are assuming that household agents communicate with neighbours during the process of risk perception and before making a decision, M4 and M7 provide higher risk perception compared to M3, as shown in Figure 6-5.b. This reflects that the social interactions still amplify the processes.

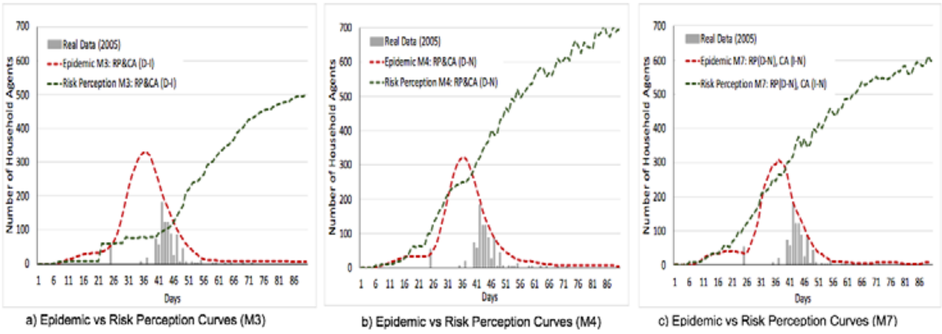


Figure 6-5: Epidemic curves (in red) and Risk perception curves (in green) for scenarios M3, M4 and M7

The epidemic curves in the three scenarios report more infected cases with approximately the same peak heights. However, M7 reports less infected cases, since household agents in their coping appraisal depend on themselves rather than their decentralised groups. Overall, all three models (M3, M4, and M7) during the process of evaluating the disease risk seem to get it wrong: risk perception slowly grows in the days when the epidemic is peaking (Figure 6-5.a, b, and c). It seems they are not reacting to the peak in any way, which looks unrealistic. Moreover, RP in the three models is continuously growing after the epidemics are almost over. The risk perception peaks when there is no longer a risk (last days as shown in Table 6-2). This is due to the impact of group members of household agents who experienced cholera in their household. Thus, their number increases over time, and they keep on voting for risk perception even when the epidemic is over.

In M3, the small stepwise increase in risk perception represents the response to media, and it is similar in its developing to M1 (In-I) (Figure 6-5.a). The household agents in their decentralized groups do not have contact with neighbours, and therefore, no cases are reported to them from their neighbourhoods. As such, they are disconnected from what is happening around them.

In M4 and M7, where social interactions are included, the developing of risk perception seems more responsive, especially after day 21 and the activation of media, although their response time is still slow (Figure 6-5.b and c). Here, in these models, the group decisions are very much dependent on the composition of the group member's opinions, which vary from one another and have different information sources to connect for one final decision regarding risk perception (in both M4 and M7) and coping appraisal (M4).

Therefore, the majority vote leads to unsecured decisions. Groups in these models are heterogeneous in terms of household agents having different opinions from the group members they vote with. Decentralized groups with isolated input information (M3) lead household agents to vote to use the water fetched from the river (D1) most of the time (Figure 6-6 map a). Because of their lack of communication with neighbours, household agents miss the opportunity of getting information about the infection in their neighbourhoods. This explains the high number of infected cases of these models compared to others.

Social interactions in both M4 and M7 helped agents to make better decisions, although following the majority still impacts their choices. For instance, in high-income communities (upper communities in Maps b and c in Figure 6-6), household agents mostly use the water as it is even though they are rich enough to boil water before using it (D3) or buy bottled water (D4). Besides, the opposite also occurs when household agents with low incomes buy bottled water, which is an expensive decision for them.

Thus, the process of coping appraisal in M4 may lead to the inconvenience of individual members. As such, all members should follow the final decision of their groups even though these decisions might not be good enough to protect them. The household agents need to find a balance between preventive behaviour and their capability to implement it. Moreover, there is always a possibility of routinely changing one's mind based on daily updates of information regarding the epidemic and the status of the surrounding people. However, when following the majority in the groups, this possibility becomes less.

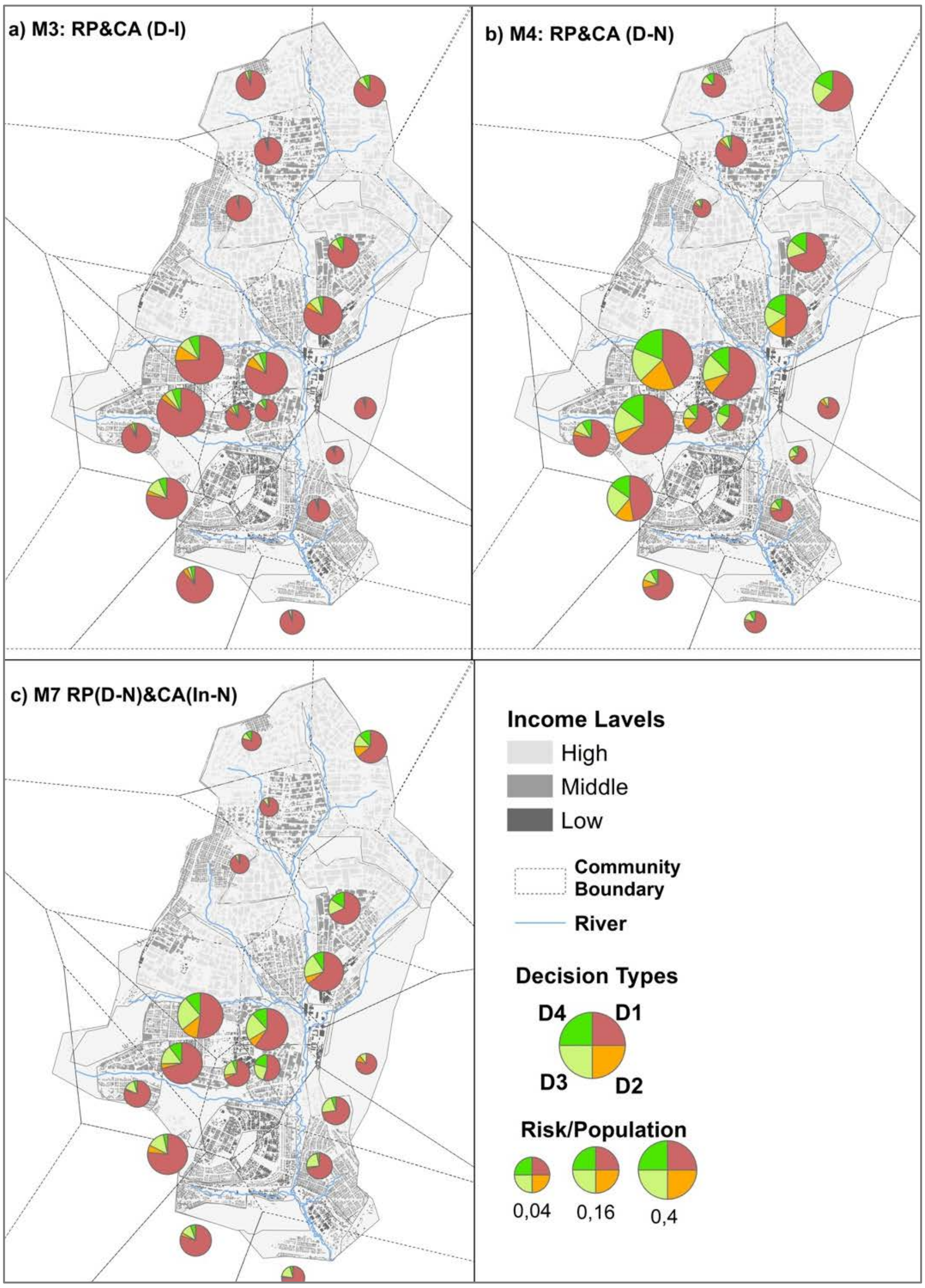


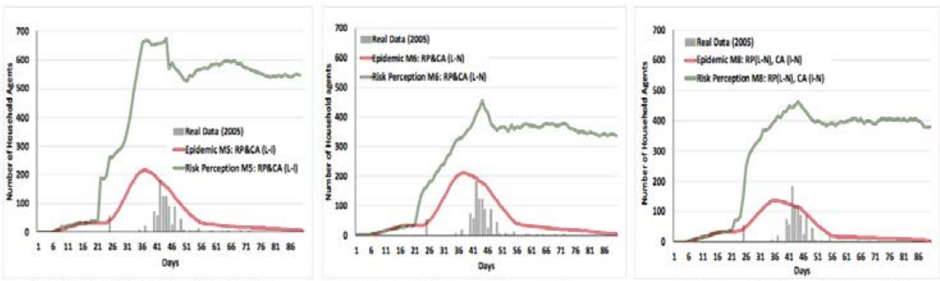
Figure 6-6 Spatial distribution of different coping appraisal decisions of scenarios M3, M4 and M7; the size of the pie represents the size of household agents with RP = 1

In M7, the household agents depend on risk perception for their decentralized groups, which often leads to no risk perception (Figure 6-

5.c). When they go for coping appraisal individually, more agents make a decision D1 (Figure 6-6.c). When they start to perceive risk during the last days of the epidemic, household agents make decisions D3 and D4 in the middle-income level and D2 for the low-income level, as can be seen in Figure 6-6.c.

6.3.3 Impact of the Leaders

The centralised groups represent the top-down scenario. In centralized groups, one household agent is randomly selected to be the group leader. The leader is responsible for risk perception and the coping appraisal of the group. The groups' members copy the risk perception and decisions of the coping appraisal of the leaders. The groups' leaders can help their groups to improve their performance if they model the appropriate responses to the situation their group faces (Zhao et al. 2018). In this article, we simulated the leader in two ways: as a dictator guiding the group (M5 and M6), and as an opinion leader who evaluates the risk of cholera and gives freedom to group members to select their own coping appraisal (M8). The three models have the same qualitative trends and the trends coincide with what is expected: peak due to amplification and gradual decrease (plateau) (Figure 6-7). The centralized group learning seems to represent the processes well, as the leader guides the group members and might help them to increase their effectiveness. Further, the leader brings them together to behave protectively. However, since no real data existed for risk perception and the correct behaviour of people during the epidemic, we cannot judge which model (M5, M6, and M8) is the best. In the following subsections, we will evaluate the three models if the leader is a dictator (M5, and M6) and if s/he is an opinion leader (M8).



a) Epidemic vs Risk Perception Curves (M5) b) Epidemic vs Risk Perception Curves (M6) c) Epidemic vs Risk Perception Curves (M8)

Figure 6-7: Epidemic curves and Risk perception curves for scenarios M5, M6 and M8

a. *As a Dictator (M5 and M6)*

When the leader is a dictator, s/he gives the group members her/his risk perception value and tells them what to do for coping appraisal. The leader learns either isolated (M5) or interacting with her/his neighbour's household agents (M6). The leaders in M5 are overestimating the risk perception (Figure 6-7.a). This is perhaps because the leader might have a bad experience with cholera, so s/he keeps warning the group. In either case, in the presence of social interactions (interactive with neighbours in M6), the uncertainty in the process of RP update is lower, i.e. around the epidemic peak compared to M5 (Figure 6-7.b). However, it is still responsive to the development of the number of infected cases.

By examining Figure 6-8, we see the impact of dictator leaders concerning coping appraisal. Isolated leaders guide their groups to different types of decisions (Figure 6-8.a), and sometimes, to less secure decisions (D1). With interactive leaders, leaders seem to have more trust in their neighbours and to make the decision of walking to a cleaner water point over the river (D2) more often. Besides, dictator leaders also guide their groups to use the river water as it is (D1). Very few leaders direct their groups to boiling the fetched water (D3) and buying bottled water (D4) (Figure 6-8.b).

b. *As an Opinion Leader (M8)*

In M8, the leader in the centralised groups is responsible for evaluating the risk perception of their groups. The leaders contact their neighbours during the process of risk perception. For the coping appraisal, the group members make their own decisions, including input information from their social interactions. With this model, the least number of infections occur. The shape of the epidemic curve (except for its height) is very close to the real data of 2005 (Figure 6-7.c.). As in M6, in M8, the uncertainty in the process of RP is lower (Figure 6-7.c). The risk perception curve develops around the epidemic peak (Figure 6-7.c).

When giving group members the opportunity to make their own decision for their coping appraisal, this leads to better performance and more preventive decisions. Figure 6-8.c shows the spatial distribution of different types of decisions during the simulation. More household agents go for D3 and D4, which are considered to be the most protective decisions. In addition, communities have at least three types of decisions that reflect their heterogeneity.

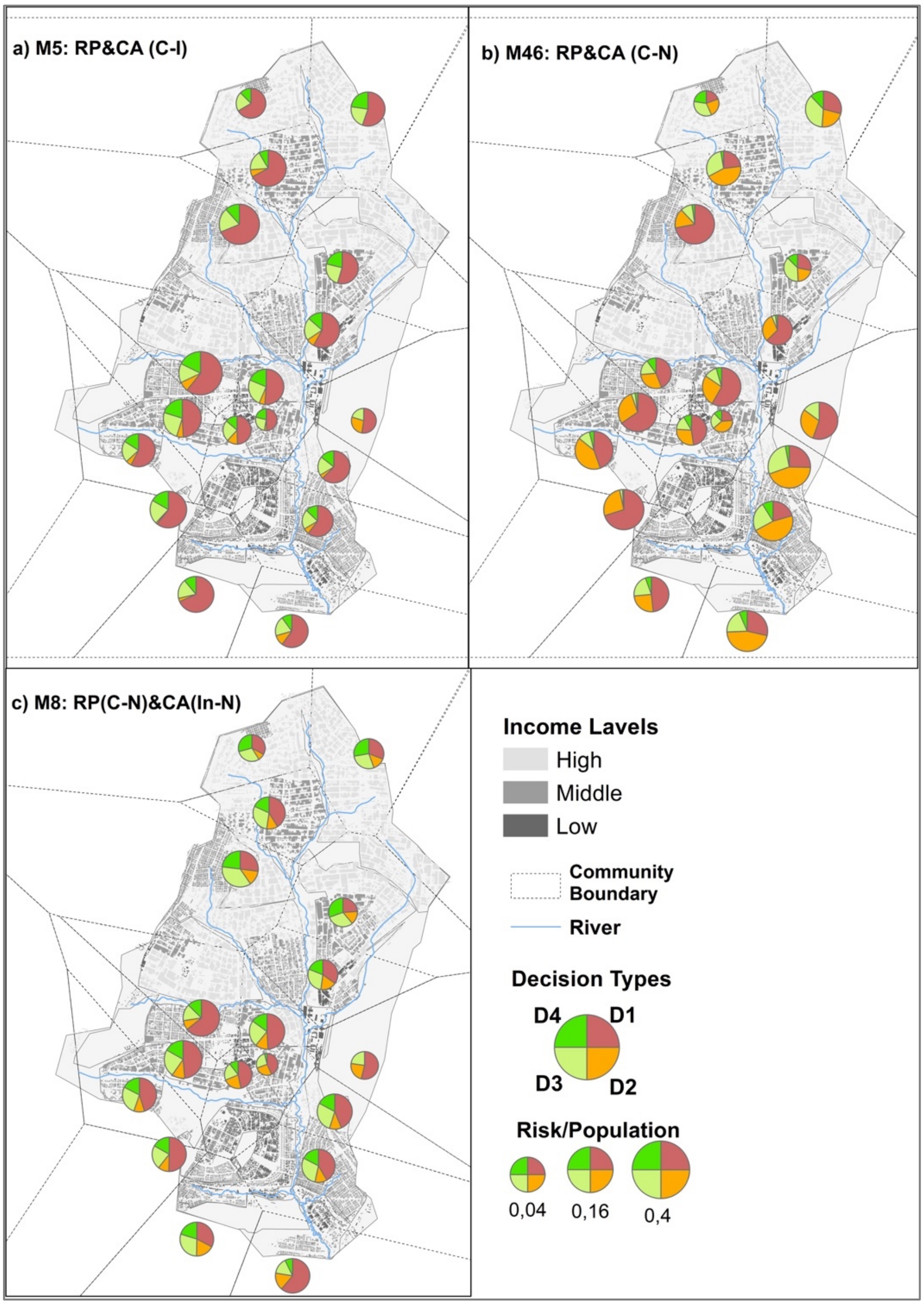


Figure 6-8: Spatial distribution of different coping appraisal decisions of scenarios M5, M6 and M8; the size of the pie represents the size of household agents with RP = 1

6.4 Conclusions

The decision to integrate learning in an ABM is often obvious. Yet, the way this learning is implemented receives less consideration. This paper illustrates that different implementations of individual and collective intelligence in agents' behaviour lead to different model outcomes. Interactive learning, which assumes that agents share information about risks and potential protective actions, outperformed isolated learning for both individuals and in groups. This underlines the fact that the integration of social learning is very important in ABMs.

We also saw that decentralised groups with majority votes were less successful compared to groups with leaders. When deciding about current risk perceptions, majority votes may not be the best mechanism of group decisions. Perceiving risk is a very personal decision making process (Brown, 2014). Therefore, when group members vote on their personal evaluation of risk, the majority arrives at a wrong decision. In contrast, when leaders give their opinions on risk perception, such groups perform better in terms of risk appraisal. Moreover, opinion leaders are even more effective and help their group members to make better coping decisions by giving them the freedom of making appropriate decisions compared to leaders-dictators or majority votes who impose a decision that all group members should follow.

In our experiments, the structure of the groups is simple and is formed on the basis of spatial and socio-demographic characteristics of agents. Future research may focus on constructing groups based on different variables (family ties, religion, tribes). In our ABM the leaders have no particular knowledge but are randomly selected and assigned to the groups. In reality, this may not be the case, e.g. leaders may have access to better information or have already earned group's trust and respect. In addition, decentralized groups can be improved by giving more weights to more trustable partners who make wiser decisions.

We conducted experiments using ML for both risk perception and coping appraisal (M1 – M6) and for risk perception only (M7 and M8). The two-tier social learning leads to very different results compared to the latter, indicating that the way learning is implemented is crucial. The implementation of BNs with interactive mode (social learning) improved the results in terms of total infected cases, type of coping decision, and the response to the development of the epidemic. This seems to be the most realistic strategy as decisions, in reality, are often made at interactive individual level (in this model – household agent).

The model's performance can be a strong argument when the number of agents is massive, e.g. when simulating a pandemic and a very large population is needed to detect a worldwide diffusion mechanism. In this case, social group learning, as described in model M5, is a very good alternative to individual interactive behaviour. Further, it shortens the computation time by 73 %, while maintaining a good quality model output.

The ultimate decision on which type of social behaviour to use can be steered by different considerations. In addition to the technical model performance metrics discussed here, the choice of a particular type of social behaviour can also be based on the type of society that is being modelled. Different political systems, the presence of tribes, different ethnic groups or religious leaders can be reasons to carefully consider the social interactions in a model. One should make sure that the actual situation regarding social learning is represented in line with existing cultural and social norms of the society being modelled.

Within the scope of this article, it was not possible to define, which implementation (M1-M8) represents the situation in Kumasi most closely. To validate the risk perception-behaviour, one would need risk perception data for this area for the time of the epidemic. During this study, we discovered that this type of data is very scattered, not only for Kumasi but in general. As we illustrated in this study, technically, many different implementations of social behaviour using ML are possible, but data is needed to validate alternative implementations. One of the major limitations of this research is the lack of empirical data on risk perception. We assumed that risk perception will grow during a disease outbreak. For this reason, we assume that the risk perception peak should proceed the fading out of the disease cases. Research on risk perception during epidemics is often conducted too late (when the peak is over) or at distant geographic locations (not in the area where the disease spreads). Hence, it provides little empirical proofs on people's behaviour and risk perception. More research on risk perception during epidemics, including other related variables, such as cultural aspects and group behaviour, can be very helpful in generating a model that represent a specific society realistically.

Moreover, agent-based modelling software does not always include ML toolkits and libraries. This complicates the implementation of different types of social intelligence. Hence, a better integration of the domains of ABM and ML in one software package or linkable libraries can eliminate this problem in the future.

Finally, another important direction of future research is to implement other ML techniques besides BNs, such as decision trees or genetic algorithms. In addition, implementing groups with different ML algorithms may lead to different results since groups will be heterogenous in terms of members' learning algorithms. Moreover, several developments in health research draw our attention to the implementation of learning in disease models. One is the fact that fake news and the impact it has on the behaviour of people. The other is the fact that human behaviour towards vaccination can change radically based on (fake) news items. Therefore, including these factors and testing the impact of them on behaviour of agents may help to draw more conclusion policy makers may consider to control epidemics.

Chapter 7: Synthesis, Conclusions and Future work

7.1 Synthesis and Conclusions

Despite all the progress, modern societies are still vulnerable to a variety of threats: natural hazards, diseases turning into epidemics and technological disasters to name a few. Effective decisions have to be made under the conditions of uncertainty, with partial availability of information, and valuable resources or even lives at stake. Risk-related problems are complex and involve different actors who participate, interact, learn and need to adapt to constantly changing environments. While simulation models are often used to support policy decisions in risky contexts, many have a simplistic representation of behaviour and learning. This thesis focuses on agent-based models of socio-environmental systems and aims to explore the implications of a machine learning integration to represent adaptive behaviour on an agent level. Two sub-objectives serve as stepping stones to achieve this main research goal. Specifically, I seek (1) to provide insights into how ML algorithms can be integrated into ABMs developed to study SES dynamics, and (2) to explore the implications of learning, including social and spatial intelligence on the behaviour of agents facing risky choices. To address these sub-objectives, I performed a thorough literature review on the state-of-the-art practices of using ML algorithms in ABMs and systematically tested different implementations of BNs in a spatial ABM taking a cholera disease diffusion ABM as an example.

Globally, millions of individuals are regularly exposed to deadly infectious diseases, which at times cause epidemics. Governments and international organizations do everything they can to curb epidemics, but they also rely on individuals for protective actions. Perceiving disease risk motivates people to adapt their behaviour towards a safer and more protective lifestyle. Indeed, risk perception is an integral part of the decision-making process under uncertainty and can be understood as an individual's evaluation of risk in a particular situation. In this thesis, I use a geographically explicit simulation model—the CABM (CABM)—which was developed initially with zero-intelligent agents (Augustijn et al. 2016). I systematically adjust the model to enhance it with ML and enable agents to learn about the disease risks and effectiveness of protective actions. Specifically, agents deal with uncertainty by assessing the risk of being infected with cholera using both spatial and social information and change their behaviour accordingly. Agents' risk perception and coping appraisal are updated using ML algorithms, which continuously adjust the dynamics of agents' beliefs and actions during a simulation. Given the type of decisions, availability of data and a variety of sources of information that agents use to learn about risks, I chose BNs to advance the agents' cognitive model in the CABM. I outline the main findings in line with the research questions posted earlier in this thesis.

7.2 Answers to Research Questions

Research Question 1: What is the state-of-the-art in employing intelligent agent learning in ABMs of SES?

In *Chapter **2***, I reviewed 137 published articles presenting ABMs that used ML algorithms to enhance agents' cognition. As many authors do not explicitly state the reason or added value of using ML algorithms, I wanted to reveal if this choice was perhaps driven by the type of intelligence they were seeking, the tasks agents had to perform or the fact that empirical data were available to be used in supervised learning. I differentiated among various tasks for which ML algorithms could be used: optimization, adaptation, negotiation and prediction. Learning related to risk perception and coping appraisal of protective actions falls under prediction and adaptation in this classification. Although spatial models had a small bias towards optimization and non-spatial models to adaptation, ML algorithms seem to be useful for a wide range of learning activities. The same conclusions could be drawn for the algorithms themselves. Different ML algorithms were used for the same type of tasks while the same algorithm was employed to support an intelligent judgement for different tasks. An object of learning in the reviewed ABM literature was mostly an individual agent rather than a collective one: examples of group learning were scarce. While I anticipated that data could be a limitation for the implementation of ML in ABMs, this seemed less problematic than expected. Modellers employing ML algorithms use a wide range of data sources, including survey data on stated preferences (a hypothetical choice), simulation data or expert knowledge. Interestingly, there were very few examples of mixed social and spatial intelligence when implementing learning at the agent level in ABMs of SES.

Research Question 2: How can spatial and social intelligence driving risk choices be implemented in an ABM?

To address this research question, I integrated psychological aspects of decision-making under risk into a spatial ABM using ML, as demonstrated in *Chapter **3***. The spatial CABM (Augustijn et al., 2016) used as a case study in this thesis initially had household agents that were not learning at all. First, I introduced simple rule-based learning, creating a benchmark spatial ABM that did *not use machine learning*. Second, I implemented BNs algorithms for both risk perception and coping appraisal decisions of household agents to steer their adaptive behaviour. The BNs have replaced *ad hoc* rule-based schemes for individual reasoning under uncertainty. Hence, the intelligent agents became capable of sensing and reacting to the stochastic spatial and social environment. In addition, BNs constantly adjusted to the dynamics of agents' own beliefs.

CABM enhanced with BN1 for the threat appraisal was used to explore the spatial and temporal patterns of disease spread depending on varying risk-communication strategies. What individual agents saw in their environment impacted their willingness to perceive risks and adapt to it. Because it is impossible to detect the presence of cholera bacteria in water visually, we assume that the safety of drinking water is assessed via the level of visual pollution at water collection points. The process of evaluating the visual pollution of the river water represented the spatial intelligence of household agents in CABM. Social intelligence was activated when combining information from various information sources, such as media or neighbours. Information about emerging disease risks (regarding the density of infected cases) and the effectiveness of risk-coping measures (how effective is the decision of treating river water) is transferred via social interactions. Their intensity impacts the awareness of cholera risk in the study area as well as the number of infected individuals. This was shown by running a sensitivity analysis on the number of contacts an agent has. With fewer social interactions, agents with intelligent risk perception are less likely to be aware of any cholera cases in their neighbourhood.

The structure and the data used for the BN1 came from expert knowledge and a small survey to parameterize initial weights of the risk perception factors. Risk perception and the process of making a decision are complex processes combining spatial and social factors. However, because of the lack of data on individual risk perception of disease, few implementations are available that integrate ML for agents' risk perception in ABMs. Little is known about spatial health risk detection and corresponding data records, especially in developing countries. This limitation is addressed by answering the next research question.

Research Question 3: How can supervised learning of ML algorithms be implemented in ABM, given scattered micro-level data?

In my thesis, the factors that influence individual risk perception were assumed to include: visual pollution of river water, memory on using a particular water source before, media broadcasting of cholera news, and communication with neighbours living in the same community and fetching water from the same water point. These factors were represented as input nodes of BN1. I collected behavioural data on cholera risk perception to parameterize these factors in BNs in two ways: using a Massive Open Online Course (MOOC) and an online Google survey. In the MOOC, survey participants chose to use or not to use river water for drinking through judging its quality by the visual appearance (pictures). The Google survey collected information on the influence of individual risk factors on the willingness to use the river water without visuals using only a textual description of the water quality situation. The dataset coming from the surveys was presented in two chapters. In *Chapter **4***, the data were used to validate the outcome of BN1; and in *Chapter **5*** the data were used to construct and train BN1.

The results of the model with BNs designed based on expert information resemble the data gathered from the surveys. This applies to agents who predicted risk through individual risk factors (e.g. only media attention, or only visual pollution) and also for agents that predict risk based on a combination of two or more factors (e.g. media attention and neighbour communication). In particular, communication, either with neighbours or media, leads to increased awareness for the survey participants as well as for the CABM agents. The similarity in trends indicates that ABM with the BNs, which are designed based on the expert knowledge and parameterized with data from the literature and the census data for Kumasi, is in line with the patterns observed in the survey.

In *Chapter **5***, I demonstrated that the structure of a BN is the factor that least impacts the final outcomes. Parameterisation and the way the network was trained (before or during a simulation) played a more important role and had more impact on the final model outcomes. I also observed that training a BN prior to an ABM simulation run led to ‘overly intelligent agents’, with high risk perception at initialization that did not decline even in the absence of cholera reports. The choice between prior training and training during simulation runs is specific to the application. In my case, as the citizens of the study area had no previous experience with cholera, they were not prepared. For other applications, a certain level of risk awareness may be essential at the start of the simulation, demanding prior training of the BNs. Moreover, I found that risk perception differs spatially within the simulated city. This applies to the level of risk perception but also to the factors affecting it. This stresses the importance of treating risk perception as a spatially heterogeneous factor. Agents with the expert-driven BN model guiding their beliefs (including risk perception and coping appraisal) provided the most balanced risk perception. The use of expert knowledge in the design and parameterization of the BN helps to avoid overfitting to a specific training dataset. It also enables direct model comparison because it computes a full posterior distribution of the BN of all risk factors (nodes).

Research Question 4: How comparable are the results of an ABM with intelligent decision-makers to the one with zero-intelligent agents (i.e. rule-based learning)?

To evaluate the difference in the integration of BNs in the CABM, I compared the outcome of the intelligent ABM with the result of the rule-based cholera model (*Chapter **3***). I increased the level of intelligence gradually, moving from zero-intelligence agents to the intelligent judgements about risks (i.e. BN1 for risk perception only), and eventually to the intelligent judgements on both risk perception and coping appraisal (BN1 and BN2). When agents had no cognitive abilities and were not reactive, then the probability of becoming infected during a rainy period depended on the density of infected agents. However, enhancing agents with cognitive abilities for risk appraisal (BN1) reduced the total number of infected agents considerably. Agents were risk-aware and took a variety of precautionary actions based on their income class and education, ill individuals in their own and/or their neighbours' households. Hence, fewer cases of infection occurred at the later stages of epidemics. The CABM enhanced with BN1 for the threat appraisal could be used to explore the spatial and temporal patterns of disease spread depending on different risk-communication strategies.

The two-tier learning on both risk perception and coping appraisal (BN1 and BN2) in CABM enabled agents to perceive risk, to acquire and to share knowledge via a social network about the effectiveness of various disease protection actions. In addition, it allows exploring the emergence of disease diffusion patterns tracing geographic, educational and income inequalities. Agents with two-tier BNs performed better than agents with one BN. Agents learn about the effectiveness of preventive measures and learn to recognize risks. The society as a whole makes healthier and more cost-effective choices. The total number of disease cases dropped by 90% of the original number of cases. The implementation strategy, in which we apply both BN1 for risk awareness and BN2 for risk appraisal, outperformed the implementation with a single BN.

Research Question 5: Given the reliance of ABMs on social interactions, what difference does the level of collective intelligence make when implementing ML algorithms in an ABM?

In *Chapter* ***6***, I tested the influence of social learning on agents' behaviour. Adaptive behaviour of agents was contingent on how well they learn about changes in disease risks and coping options, individually or in interactions with others. ML techniques could be instrumental for modelling risk and coping appraisal processes either as an individual or collective intelligence. The impact of different types of group learning compared with individual learning is an underexplored domain in disease modelling, and in ABMs of SES in general.

By achieving this objective, *Chapter* ***6*** illustrated that different implementations of individual and collective intelligence in agents' behaviour led to different model outcomes. Interactive learning, which assumed that agents share information about risks and protective actions, outperformed isolated learning for both individuals and groups. This underlined the fact that the integration of social learning is essential in ABMs.

In addition, *Chapter* ***6*** showed that agents might be represented as members of local groups (small social networks), learning together and copying behaviour from other group members. Group learning can be realised by making all group members use their own ML algorithms to gather information to perform a specific task (decentralised), and then pool their opinions collectively by making one decision adopted by the entire group. Here, I used a 'majority vote' as the resolution mechanism in decentralised group decision-making. Alternatively, group learning could be realised by introducing one agent (an opinion leader or a dictator) who uses ML to learn for the whole group to help it accomplish its group task (centralised). Hence, I implemented the centralised group learning in CABM where agents in the group copied the decisions of their leader. In both cases, all agents that belonged to a group shared the same decision, but the information this decision was based on varied considerably.

The decentralised groups with majority votes were less successful compared with groups with leaders. When deciding about risk perceptions, which could vary a lot across heterogeneous households, the majority vote seemed to average potentially polarized opinions and arrived at a wrong decision. In addition, two types of leaders were presented: dictator leaders and opinion leaders. In groups with dictator leaders, the leader is responsible for risk perception and the coping appraisal of the group. The groups' members copy the risk perception and decisions of the coping appraisal of the leaders. The leader guides the group members and might help them to increase their effectiveness and brings them to behave protectively. When opinion leaders share their assessments of risk perception with a group, such groups performed better regarding risk appraisal. Moreover, opinion leaders were even more effective than dictator leader groups in the coping appraisal. They helped their group members to make better coping-decisions by giving them the freedom to make appropriate decisions. Hence, collective intelligence implemented as a group with an opinion leader performed better than individual intelligence, leader-dictators or majority votes that imposed a decision that all group members should follow. We highlight that the implementation of collective intelligence should be aligned with the cultural norms and a possible hierarchy in a society that is being modelled.

7.3 Innovative Contributions to Science

This thesis contributes to the scientific efforts to integrate an ML algorithm into ABMs designed to study SES dynamics. The presented models, data and insights make a number of innovative contributions to science:

Methodologically, this thesis for the first time provides a systematic test on the implications of alternative implementations of agents' intelligence in ABMs of SES. First, I developed BNs that drive intelligent decisions of household agents regarding risk appraisal and behavioural change in coping strategies in a spatial ABM of cholera diffusion. Second, I presented the possibility of implementing learning in spatial ABMs with a small behavioural dataset via an innovative combination of implementing **a double BN** (one for risk assessment and one for coping appraisal) instead of a single BN driving the complete decision-making process. This allowed us to assess individually the factors included in the risk perception, and the decision agents made. In addition, this double implementation also has other advantages. It will allow the developer to replace one of the BNs with another ML algorithm, for example when either of the two steps is computationally more demanding or when a dataset is available for either of the two steps. In our example, this would most likely be the coping appraisal step in the process. If empirical data are found of the preferences of individuals for certain water sources, it can feed an alternative coping appraisal. In addition, I showed the extent to which supervised learning of ML algorithms should depend exclusively on data, and the level of intelligence necessary for agents to mimic realistic risk perceptions. Third, I explored the impact of different types of group learning compared with individual learning on SES-ABM dynamics. Rarely are the effects of alternative implementations of collective intelligence in ABMs on modelling results assessed, usually assuming that the group learning is computationally attractive. I evaluated the influence of individual vs. collective learning on an epidemic's dynamics within a disease ABM by pursuing a quantitative test on the influence of agents' ability to learn, individually or in a group.

With respect to the application domain, the thesis went beyond the traditional representation of fixed behaviour responses in the risky context common for disease ABMs by explicitly modelling learning. In particular, I included dynamic risk perception in the CABM benefiting from the protection motivation theory from psychology that has been actively applied in health research to study cognitive processes and to predict health-related behaviour. Employing learning techniques to capture dynamics in risk perception and corresponding protective behaviour mimic the complex process of how humans act upon encountering risk. I illustrated that our spatial ABM with static behaviour and zero-intelligent agents led to a higher scope of contagion compared with the real situation, consequently leading to an overestimation of the prevalence of disease cases. Omitting the dynamics of cognitive processes on the agents' side may lead to misguiding conclusions on the effectiveness of preventive measures.

Bridging interdisciplinary gaps in understanding the learning processes—including both social and spatial intelligence—offers better modelling tools that could support policy decisions with a mix of response strategies that account for adaptive behaviour. The learning method developed within this thesis steers risk perceptions and risk-coping behaviour of household agents relying on a range of information sources and social interactions. In both, the sensing of information (global, from the spatial environment and from other agents), exchange of information (between agents), and processing of information (intelligent decision-making) are central. As the speed of information exchange increases, agents use social intelligence to learn from the experience of a larger group of individuals with respect to the safety of alternative water fetching points and potential preventive behaviours. Besides communication between agents, the timing of the media reports impacts public awareness and individuals' precautionary measures. In addition, the thesis advances the implementation of spatial intelligence. Namely, it analyses the role of visual pollution in a particular location, which agents process using BN and which is supported by the original survey data. The spatial character of the judgement processes was also important when considering risk-coping alternatives because agents in different locations had access to certain water sources.

7.4 Implications for Policy and Society

This thesis focused on resolving methodological issues when integrating ABMs and ML algorithms. Even so, there are a number of implications of practical value that can be drawn from its conclusions. First, the experiments conducted with the CABM with intelligent agents test the impact of the timing of media attention (controllable by decision-makers) on the spread of risk awareness and epidemics dynamics. Because news media are among factors affecting risk perception and protective actions, the model could be further used to explore alternative communication strategies. This centralised communication strategy could be tested in the presence or absence of partial information–or even fake news–often spread via social networks. Other intervention methods, like garbage collection strategies, handing out bottled water etc., could also be considered. Ideally, the development of policy-oriented ABMs should go in participatory settings where policymakers could co-design assumptions and develop realistic intervention scenarios.

Second, this thesis illustrated that alternative implementations of social intelligence would influence the validity of the disease ABM enhanced with ML. Specifically, majority vote and a leader-dictator CABM underperform compared with the opinion-leader implementation. This calls for the critical assessment on a model developer side when selecting an implementation for a particular context. Social and cultural norms prevailing in the society under study are crucial. Understanding this 'soft' aspect of society is necessary to select the correct individual or group implementation. Different political systems, the presence of tribes, different ethnic groups or religious leaders can be reasons to consider a particular type of social learning when formalizing it in a model. One should make sure that the actual situation regarding social learning is represented in line with existing cultural and social norms of the society being modelled.

7.5 Limitations and Future Work

With all efforts spent on any project, there is always space to develop it further. The research questions of this PhD project have been answered by integrating BNs in the spatial ABM of disease diffusion. At the beginning of this project, there was little empirical data available on risk perception during cholera epidemics and corresponding behaviour. Given the nature of the learning tasks that agents pursued and the data, BNs were the best candidate for the implementation of risk perception and coping appraisal in the CABM. BNs can continuously adjust to the dynamics of agents' beliefs. During this PhD research, I tested several different ways of implementing the BNs, including the impact of spatial learning and the impact of social learning via individual or collective intelligence. In addition to spatial, hydrological and socio-economic data used in our case study, this modelling effort could benefit further from disaggregated behavioural data. Currently, the BNs implemented in the CABM were updated based on information obtained via personal communication, media and visual observations of the environment. While we use data from the survey among students from developing countries to parameterize initial weights of BN nodes, this may not be fully representative of the population in the case-study area where the ABM was applied. Notably, the survey participants employed in this study were well-educated individuals from a variety of nations. This imposes limitations on making policy-relevant conclusions, although it allows us to test the fitness of ML algorithms implemented within a spatial ABM.

Another interesting extension could be to test the learning behaviour of agents using another ML algorithm. ML algorithms that do not require extensive data for training and testing could replace BNs; for example, genetic algorithms, decision trees and Random Forest. Algorithms can differ in terms of intensity of training data, mathematical nature and the time required to learn and reach a stable state. Integrating different algorithms will help to evaluate which of the algorithms performs better within the same model.

In all experiments, agents were heterogeneous in location and socio-economic characteristics, which influenced inputs to the agents' cognitive models. However, the cognitive model itself was homogeneous. In reality, some individuals are more risk averse than others, or more reluctant to change their behaviour. This can be represented by different behavioural models. It would be very interesting to test different risk threshold levels and evaluate the impact on the risk perception curves. Currently, it is unclear if agents with a lower risk perception threshold learn faster and are more efficient in their coping appraisal.

Moreover, in our experiments, we considered the media to be a trustworthy source of information. This might not be realistic. Fake news also has an impact on human behaviour. Human attitudes towards vaccination, for example, can change radically based on (fake) news items. Media can have a substantial effect on the opinion of an audience, either positively or negatively through scary stories (McCluskey and Swinnen, 2011). People can easily be influenced by what is broadcast from media. Therefore, research on how media affects people is important to improve ABMs further. Including these factors and testing their impact on the behaviour of agents may help policymakers in their efforts to control epidemics and other disasters.

Finally, ABM software does not always include ML toolkits and libraries. This complicates the implementation of different types of ML algorithms for agents' intelligence. Better integration of the domains of ABM and ML in one software package or linkable libraries can eliminate this problem in the future.

Bibliography

Abdou, M., Hamill, L., & Gilbert, N. (2012). Designing and Building an Agent-Based Model. In *Agent-Based Models of Geographical Systems* (pp. 141–165). http://doi.org/10.1007/978-90-481-8927-4

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

Alexandridis, K., & Pijanowski, B. C. (2006). Modular Bayesian Inference and Learning of Decision Networks as Stand-Alone Mechanisms of the MABEL Model: Implications for Visualization, Comprehension, and Policy-Making. *Agent 2005 Conference on: Generative Social Processes, Models, and Mechanisms*, 419–438.

Alonso, E., D'Inverno, M., Kudenko, D., Luck, M., & Noble, J. (2001). Learning in multi-agent systems. *Knowledge Engineering Review*, *16*(3), 277–284. http://doi.org/10.1017/S0269888901000170

Alshammari, S. M., & Mikler, A. R. (2018). Modeling Spread of Infectious Diseases at the Arrival Stage of Hajj. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10814 LNBI*, 430–442. http://doi.org/10.1007/978-3-319-78759-6_39

An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, *229*, 25–36. http://doi.org/10.1016/j.ecolmodel.2011.07.010

Anderson, T. M., & Dragićević, S. (2018). Network-agent based model for simulating the dynamic spatial network structure of complex ecological systems. *Ecological Modelling*, *389*, 19–32. http://doi.org/10.1016/j.ecolmodel.2018.10.008

Asadi, R., Mustapha, N., & Sulaiman, N. (2009). A Framework For Intelligent Multi Agent System Based Neural Network Classification Model. *International Journal of Computer Science and Information Security*, *5*(1), 168–174.

Asase, M., Yanful, E. K., Mensah, M., Stanford, J., & Amponsah, S. (2009). Comparison of municipal solid waste management systems in Canada and Ghana: A case study of the cities of London, Ontario, and Kumasi, Ghana. *Waste Management*. http://doi.org/10.1016/j.wasman.2009.06.019

Asgari, S., Awwad, R., Kandil, A., & Odeh, I. (2016). Impact of considering need for work and risk on performance of construction contractors:

An agent-based approach. *Automation in Construction*, *65*, 9–20. http://doi.org/10.1016/j.autcon.2016.01.004

Augustijn, E. W., Doldersum, T., Useya, J., & Augustijn, D. (2016). Agent-based modelling of cholera. *Stochastic Environmental Research and Risk Assessment*, *30*(8), 2079–2095. http://doi.org/10.1007/s00477-015-1199-x

Balbi, S., Perez, P. . D., & Giupponi, C. . C. (2010). A spatial agent-based model to explore scenarios of adaptation to climate change in an alpine tourism destination. *Modelling for Environment's Sake: Proceedings of the 5th Biennial Conference of the International Environmental Modelling and Software Society, IEMSs 2010*, *1*, 700–707. http://doi.org/10.2139/ssrn.1566861

Balmann, A., & Happe, K. (2001). Applying parallel genetic algorithms to economic problems: The case of agricultural land markets. *IIFET 2000 Proceedings*.

Barnes, D. J., & Chu, D. (2010). *Introduction to modeling for biosciences*. *Introduction to Modeling for Biosciences*. London: London: Springer. http://doi.org/10.1007/978-1-84996-326-8

Barrett, E., Howley, E., & Duggan, J. (2013). Parallel Bayesian Model Learning. *Swarmlab.Unimaas.Nl*.

Bassett, R. L., & Ginis, K. A. M. (2011). Risky business: the effects of an individualized health information intervention on health risk perceptions and leisure time physical activity among people with spinal cord injury. *Disability and Health Journal*, *4*(3), 165–76. http://doi.org/10.1016/j.dhjo.2010.12.001

Bauch, C. T., D'Onofrio, A., & Manfredi, P. (2013). *Behavioral Epidemiology of Infectious Diseases: An Overview*. *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*. http://doi.org/10.1007/978-1-4614-5474-8

Bennett, D. A., & Tang, W. (2006). Modelling adaptive, spatially aware, and mobile agents: Elk migration in Yellowstone. *International Journal of Geographical Information Science*, *20*(9), 1039–1066. http://doi.org/10.1080/13658810600830806

Beretta, S., Castelli, M., Gonçalves, I., & Ramazzotti, D. (2017). A quantitative assessment of the effect of different algorithmic schemes to the task of learning the structure of Bayesian Networks.

Berger, T. (2001). Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes and policy analysis. In *Agricultural Economics* (Vol. 25, pp. 245–260). John Wiley & Sons, Ltd (10.1111). http://doi.org/10.1016/S0169-

5150(01)00082-2

Besaw, L. E., Pellon, M. B., Rizzo, D. M., Grover, D. K., Eppstein, M. J., & Marshall, J. S. (2010). Up - scaling Agent - Based Discrete - Choice Transportation Models using Artificial Neural Networks. In *Transportation Research Board 89th Annual Meeting*.

Bidyuk, P. I., Terent'ev, A. N., & Gasanov, A. S. (2005). Construction and Methods of Learning of Bayesian Networks. *Cybernetics and Systems Analysis*, *41*(4), 587–598. http://doi.org/10.1007/s10559-005-0094-8

Bieberstein, A. (2014). Background to Risk Perception. In *An Investigation of Women's and Men's Perceptions and Meanings Associated with Food Risks* (pp. 25–62). Wiesbaden: Springer Fachmedien Wiesbaden. http://doi.org/10.1007/978-3-658-03275-3_2

Bone, C., Dragicevic, S., & White, R. (2011). Modeling-in-the-middle: bridging the gap between agent-based modeling and multi-objective decision-making for land use change. *International Journal of Geographical Information Science*, *25*(5), 717–737. http://doi.org/10.1080/13658816.2010.495076

Borrill, P. L., & Tesfatsion, L. (2010). Agent-Based Modeling: The Right Mathematics for the Social Sciences? In J. B. Davis & D. W. Hands (Eds.), *Director* (pp. 1–28). Edward Elgar Publishers. http://doi.org/10.4337/9780857938077.00018

Bosse, T., Hoogendoorn, M., Klein, M. C. A., Treur, J., van der Wal, C. N., & van Wissen, A. (2013). Modelling collective decision making in groups and crowds: Integrating social contagion and interacting emotions, beliefs and intentions. *Autonomous Agents and Multi-Agent Systems*, *27*(1), 52–84. http://doi.org/10.1007/s10458-012-9201-1

Boulanger, P.-M., & Bréchet, T. (2005). Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics*, *55*(3), 337–350. http://doi.org/10.1016/j.ecolecon.2005.07.033

Brandon, N., Dionisio, K. L., Isaacs, K., Tornero-Velez, R., Kapraun, D., Setzer, R. W., & Price, P. S. (2018). Simulating exposure-related behaviors using agent-based models embedded with needs-based artificial intelligence. *Journal of Exposure Science and Environmental Epidemiology*, 1. http://doi.org/10.1038/s41370-018-0052-y

Bratko, I. (1994). Machine learning in artificial intelligence. *Artificial Intelligence in Engineering*, *8*(1993).

Brown, D., Page, S., & Riolo, R. (2005). Path dependence and the validation of agentâ€ based spatial models of land use.

Journal of …, 1–42. http://doi.org/10.1080/13658810410001713399

Brown, V. J. (2014, October). Risk perception it's personal. *Environmental Health Perspectives*. National Institute of Environmental Health Science. http://doi.org/10.1289/ehp.122-A276

Campos, L. M. De. (2006). A Scoring Function for Learning Bayesian Networks based on Mutual Information and Conditional Independence Tests. *Journal of Machine Learning Research*, *7*, 2149--2187.

Cantergiani, C., & Delgado, M. G. (2016). Design of an agent-based model to simulate urban growth at Corredor del Henares (region of Madrid) . *Boletin de La Asociacion de Geografos Espanoles*, *70*, 259–503. http://doi.org/10.21138/bage.2171

Carlson, S. M., White, R. E., & Davis-Unger, A. C. (2014). Evidence for a relation between executive function and pretense representation in preschool children. *Cognitive Development*, *29*(1), 1–16. http://doi.org/10.1016/j.cogdev.2013.09.001

Carvalho, A. M. (2009). Scoring functions for learning Bayesian networks.

Caudell, T. P., Burch, C. T., Zengin, M., Gauntt, N., & Healy, M. J. (2011). Retrospective learning of spatial invariants during object classification by embodied autonomous neural agents. In *The 2011 International Joint Conference on Neural Networks* (pp. 2135–2142). IEEE. http://doi.org/10.1109/IJCNN.2011.6033492

Chakraborti, A., Toke, I. M., Patriarca, M., & Abergel, F. (2011). Econophysics review: II. Agent-based models. *Quantitative Finance*, *11*(7), 1013–1041. http://doi.org/10.1080/14697688.2010.539249

Chattoe, E. (1998, June 30). Just How (Un)Realistic Are Evolutionary Algorithms as Representations of Social Processes? *Journal of Artificial Societies and Social Simulation*. JASSS.

Chen, F., Griffith, A., Cottrell, A., & Wong, Y.-L. (2013). Behavioral responses to epidemics in an online experiment: using virtual diseases to study human behavior. *PloS One*, *8*(1), e52814. http://doi.org/10.1371/journal.pone.0052814

Chen, S. H., Jakeman, A. J., & Norton, J. P. (2008). Artificial Intelligence techniques: An introduction to their use for modelling environmental systems. *Mathematics and Computers in Simulation*, *78*(2–3), 379–400. http://doi.org/10.1016/j.matcom.2008.01.028

Choi, S. P. M., Liu, J., & Chan, S. P. (2001). A genetic agent-based negotiation system. *Computer Networks*, *37*(2), 195–204. http://doi.org/10.1016/S1389-1286(01)00215-8

Cioffi-Revilla, C., de Jong, K., & Bassett, J. K. (2012). Evolutionary computation and agent-based modeling: Biologically-inspired approaches for understanding complex social systems. *Computational and Mathematical Organization Theory*, *18*(3), 356–373. http://doi.org/10.1007/s10588-012-9129-7

Collins, A., Elzie, T., Frydenlund, E., & Robinson, R. M. (2014). Do groups matter? an agent-based modeling approach to pedestrian egress. *Transportation Research Procedia*, *2*, 430–435. http://doi.org/10.1016/j.trpro.2014.09.051

Constantinou, A. C., Fenton, N. E., & Neil, M. (2012). Pi-football: A Bayesian network model for forecasting Association Football match outcomes. *Knowledge-Based Systems*, *36*, 322–339. http://doi.org/10.1016/j.knosys.2012.07.008

Constantinou, A. C., Fenton, N., Marsh, W., Radlinski, L., & Constantinou, A. C. (2016). From complex questionnaire and interviewing data to intelligent Bayesian Network models for medical decision support Europe PMC Funders Group. *Artificial Intelligence in Medicine*, *67*, 75–93. http://doi.org/10.1016/j.artmed.2016.01.002

Costa, M. L., Oliveira, F. S., Economia, F. De, & Porto, U. (2008). A Reinforcement Learning Algorithm for Agent-Based Modeling of Investment in Electricity Markets. *Researchgate.Net*, 1–30.

Crampton, A., & Ragusa, A. (2016). Exploring Perceptions and Behaviors about Drinking Water in Australia and New Zealand: Is It Risky to Drink Water, When and Why? *Hydrology*, *3*(1), 8. http://doi.org/10.3390/hydrology3010008

Crooks, A. T. (2010). Constructing and Implementing an Agent-Based Model of Residential Segregation through Vector GIS. *International Journal of Geographical Information Science*, *24*(5), 661–675.

Crooks, A. T., & Castle, C. J. (2012). The integration of agent-based modelling and geographical information for geospatial simulation. In *Agent-Based Models of Geographical Systems* (pp. 219–251). Springer Netherlands.

Crooks, A. T., & Hailegiorgis, A. B. (2014). An agent-based modeling approach applied to the spread of cholera. *Environmental Modelling & Software*, *62*, 164–177. http://doi.org/10.1016/j.envsoft.2014.08.027

Czarnowski, I., & Jędrzejowicz, P. (2018). An Approach to Data Reduction for Learning from Big Datasets: Integrating Stacking, Rotation, and Agent Population Learning Techniques. *Complexity*, *2018*, 1–13. http://doi.org/10.1155/2018/7404627

Danquah, L., Abass, K., & Nikoi, A. A. (2011). Anthropogenic Pollution of Inland Waters: the Case of the Aboabo River in Kumasi, Ghana.

Journal of Sustainable Development, *4*(6), 103–115. http://doi.org/10.5539/jsd.v4n6p103

de Koning, K., Filatova, T., & Bin, O. (2017). Bridging the Gap Between Revealed and Stated Preferences in Flood-prone Housing Markets. *Ecological Economics*, *136*, 1–13. http://doi.org/10.1016/j.ecolecon.2017.01.022

de Marchi, S., & Page, S. E. (2014). Agent-Based Models. *Annual Review of Political Science*, *17*(1), 1–20. http://doi.org/10.1146/annurev-polisci-080812-191558

De Waal, A., Koen, H., De Villiers, P., Roodt, H., Moorosi, N., & Pavlin, G. (2016). Construction and Evaluation of Bayesian Networks with Expert-Defined Latent Variables. *19th International Conference on Information Fusion*, (July), 2337.

De Zwart, O., Veldhuijzen, I. K., Elam, G., Aro, A. R., Abraham, T., Bishop, G. D., … Brug, J. (2009). Perceived threat, risk perception, and efficacy beliefs related to SARS and other (emerging) infectious diseases: Results of an international survey. *International Journal of Behavioral Medicine*, *16*(1), 30–40. http://doi.org/10.1007/s12529-008-9008-2

Diez, F. J. (2003). Parameter adjustement in Bayes networks. The generalized noisy OR-gate. *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, *Abs/1303,* 99–105.

Djennas, M., Benbouziane, M., & Djennas, M. (2012). Agent-Based Modeling in Supply Chain Management: a Genetic Algorithm and Fuzzy Logic Approach. *International Journal of Artificial Intelligence & Applications (IJAIA)*, *3*(5), 13–30.

Doria, M. de F. (2010). Improving effectiveness and efficiency in the water sector: Institutions, infrastructure and indicators. *Water Policy*, *12*(SUPPL. 1), 1–7. http://doi.org/10.2166/wp.2009.051

Doss, C. R., McPeak, J. G., & Barrett, C. B. (2006). Interpersonal, Intertemporal and Spatial Variation in Risk Perceptions: Evidence from East Africa.

Drakaki, M., Gören, H. G., & Tzionas, P. (2018). An intelligent multi-agent based decision support system for refugee settlement siting. *International Journal of Disaster Risk Reduction*, *31*, 576–588. http://doi.org/10.1016/j.ijdrr.2018.06.013

Driedger, S. M. (2007). Risk and the media: A comparison of print and televised news stories of a Canadian drinking water risk event. *Risk Analysis*, *27*(3), 775–786. http://doi.org/10.1111/j.1539-6924.2007.00922.x

Dubbelboer, J., Nikolic, I., Jenkins, K., & Hall, J. (2017). An agent-based model of flood risk and insurance. *JASSS*, *20*(1), 6. http://doi.org/10.18564/jasss.3135

Eberlen, J., Scholz, G., & Gagliolo, M. (2017). Simulate this! An Introduction to Agent-Based Models and their Power to Improve your Research Practice. *International Review of Social Psychology*, *30*(1), 149. http://doi.org/10.5334/irsp.115

Epstein, J. M., & Axtell, R. L. (1996). Growing Artificial Societies: Social Science from the Bottom Up (Complex adaptive Systems).

Eshragh, F., Pooyandeh, M., & Marceau, D. J. (2015, October 1). Automated negotiation in environmental resource management: Review and assessment. *Journal of Environmental Management*. Academic Press. http://doi.org/10.1016/j.jenvman.2015.07.051

Fenton, N. E., & Neil, M. (Martin D. . (2012). *Risk assessment and decision analysis with Bayesian networks*. Taylor & Francis.

Filatova, T. (2014). Market-based instruments for flood risk management: A review of theory, practice and perspectives for climate adaptation policy. *Environmental Science and Policy*, *37*, 227–242. http://doi.org/10.1016/j.envsci.2013.09.005

Filatova, T., Parker, D. C., & Veen, A. van der. (2011). The Implications of Skewed Risk Perception for a Dutch Coastal Land Market: Insights from an Agent-Based Computational Economics Model. *Agricultural and Resource Economics Review*, *40*(3), 405–423.

Filatova, T., Verburg, P. H., Parker, D. C., & Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: Challenges and prospects. *Environmental Modelling & Software*, *45*, 1–7. http://doi.org/10.1016/j.envsoft.2013.03.017

Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). A Meta-Analysis of Research on Protection Motivation Theory. *Journal of Applied Social Psychology*, *30*(2), 407–429. http://doi.org/10.1111/j.1559-1816.2000.tb02323.x

Fonoberova, M., Fonoberov, V. A., & Mezić, I. (2013). Global sensitivity/uncertainty analysis for agent-based models. *Reliability Engineering & System Safety*, *118*, 8–17. http://doi.org/10.1016/j.ress.2013.04.004

Frias-Martinez, V., Rubio, A., & Frias-Martinez, E. (2012). Measuring the impact of epidemic alerts on human mobility. *Pervasive Urban Applications--PURBA*, 18.

Fritzell, C., Raude, J., Kazanji, M., & Flamand, C. (2018). Emerging trends of Zika apprehension in an epidemic setting. *PLoS Neglected Tropical*

Diseases, *12*(1), e0006167. http://doi.org/10.1371/journal.pntd.0006167

Funk, S., Gilad, E., Watkins, C., & Jansen, V. A. A. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(16), 6872–7. http://doi.org/10.1073/pnas.0810762106

Gardner, H. (2006). *Multiple intelligences : new horizons*. BasicBooks.

Gasperi, J., Dris, R., Bonin, T., Rocher, V., & Tassin, B. (2014). Assessment of floating plastic debris in surface water along the Seine River. *Environmental Pollution*, *195*, 163–166. http://doi.org/10.1016/J.ENVPOL.2014.09.001

Ghana Statistical Service. (2012). 2010 Population and Housing Census. *Ghana Statistical Service*, 1–117. http://doi.org/10.1371/journal.pone.0104053

Ghazi, S., & Dugdale, J. (2018). A Multi-Agent based Approach for Simulating the Impact of Human Behaviours on Air Pollution | Ghazi | Informatica, *42*, 199–209.

Ghnemat, R., Bertelle, C., & Duchamp, Ǵerard H.E. (2008). Agent-based modeling using swarm intelligence in geographical information systems. In *2008 International Conference on Innovations in Information Technology, IIT 2008* (pp. 69–73). http://doi.org/10.1109/INNOVATIONS.2008.4781737

Gilbert, N., & Terna, P. (1999). How to build and use agent-based models in social science.

Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the Social Scientist (second edition)* (Second). Open University Press.

Gotteland, C., McFerrin, B. M., Zhao, X., Gilot-Fromont, E., & Lélu, M. (2014). Agricultural landscape and spatial distribution of Toxoplasma gondii in rural environment: an agent-based model. *International Journal of Health Geographics*, *13*(1), 45. http://doi.org/10.1186/1476-072X-13-45

Graubner, M., Balmann, A., & Sexton, R. J. (2011). Spatial price discrimination in agricultural product procurement markets: A computational economics approach. *American Journal of Agricultural Economics*, *93*(4), 949–967. http://doi.org/10.1093/ajae/aar035

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. http://doi.org/10.1016/j.ecolmodel.2010.08.019

Gwak, J., & Sim, K. M. (2011). Bayesian Learning based Negotiation

Agents for Supporting Negotiation with Incomplete Information. *Proceedings of the International MultiConference of Engineers and Computer Scientists*, *1*.

Haer, T., Botzen, W. J. W., de Moel, H., & Aerts, J. C. J. H. (2017). Integrating Household Risk Mitigation Behavior in Flood Risk Analysis: An Agent-Based Model Approach. *Risk Analysis*, *37*(10), 1977–1992. http://doi.org/10.1111/risa.12740

Hajibabai, L., Delavar, M. R., Malek, M. R., & Frank, A. U. (2007). Agent-Based Simulation of Spatial Cognition and Wayfinding in Building Fire Emergency Evacuation. *Geomatics Solutions for Disaster Management*, 255–270. http://doi.org/10.1007/978-3-540-72108-6_17

Hamblin, S. (2013). On the practical usage of genetic algorithms in ecology and evolution. *Methods in Ecology and Evolution*, *4*(2), 184–194. http://doi.org/10.1111/2041-210X.12000

Happe, K. (2004). Agricultural policies and farm structures: agent-based modelling and application to EU policy reform. *… and Application to EU-Policy Reform, Institut Fur …*, 291.

Harris, J. B., LaRocque, R. C., Chowdhury, F., Khan, A. I., Logvinenko, T., Faruque, A. S. G., … Calderwood, S. B. (2008). Susceptibility to Vibrio cholerae infection in a cohort of household contacts of patients with cholera in Bangladesh. *PLoS Neglected Tropical Diseases*, *2*(4), e221. http://doi.org/10.1371/journal.pntd.0000221

He, Z., Cheng, T. C. E., Dong, J., & Wang, S. (2014). Evolutionary location and pricing strategies in competitive hierarchical distribution systems: A spatial agent-based model. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *44*(7), 822–833. http://doi.org/10.1109/TSMC.2013.2290506

Heckerman, D. (1995). *A tutorial on learning Bayesian networks*. *Technical Report MSR-TR-95-6*.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. *Innovations in Bayesian Networks*, (November).

Hedman, M., & Lindberg, C. (2009). Women , Water , and Perceptions of Risk.

Heinonen, J., Travis, J., Redpath, S., & Pinard, M. (2012). Combining socio-economic and ecological modelling to inform natural resource management strategies. *Iemss.Org*.

Heppenstall, A., Crooks, A. T., See, L. M., & Batty, M. (2012). Agent-based models of geographical systems. *Agent-Based Models of Geographical Systems*, *164*(2011), 1–759. http://doi.org/10.1007/978-90-481-

8927-4

Heppenstall, A., Evans, A. J., & Birkin, M. H. (2007). Genetic algorithm optimisation of an agent-based model for simulating a retail market. *Environment and Planning B: Planning and Design*, *34*(6), 1051–1070. http://doi.org/10.1068/b32068

Heppenstall, A., Malleson, N., & Crooks, A. (2016). "Space, the Final Frontier": How Good are Agent-Based Models at Simulating Individuals and Space in Cities? *Systems*, *4*(1), 9. http://doi.org/10.3390/systems4010009

Heppenstall, A., Malleson, N., See, L., & Evans, A. (2014). Optimising an Agent-Based Model to Explore the Behaviour of Simulated Burglars. *Intelligent Systems Reference Library*, *52*, 179–204. http://doi.org/10.1007/978-3-642-39149-1

Honavar, V. (2006). Artificial intelligence: An overview. *Iowa State University*, 1–14.

Hu, Y., Quinn, C. J., Cai, X., & Garfinkle, N. W. (2017). Combining human and machine intelligence to derive agents' behavioral rules for groundwater irrigation. *Advances in Water Resources*, *109*, 29–40. http://doi.org/10.1016/j.advwatres.2017.08.009

Husselmann, A., & Hawick, K. (2011). Spatial agent-based modelling and simulations-a review. *Computational Science Technical*.

Huynh, N., Cao, V. L., Wickramasuriya, R., Berryman, M., Perez, P., & Barthelemy, J. (2014). An Agent Based Model for the Simulation of Road Traffic and Transport Demand in A Sydney Metropolitan Area. *Eighth International Workshop on Agents in Traffic and Transportation*, (2005), 1–7. http://doi.org/10.13140/2.1.4023.2961

IFRC. (2016). *World Disasters Report 2016. Resilience: saving lives today, investing for tomorrow*. *International Federation of Red Cross and Red Crescent Societies*.

Janati, A., Hosseiny, M., Gouya, M. M., Moradi, G., & Ghaderi, E. (2015). Communicable Disease Reporting Systems in the World: A Systematic Review Article. *Iranian Journal of Public Health*, *44*(11), 1453–65.

Jennings, N. (2001). Agent-Oriented Software Engineering, *1957*. http://doi.org/10.1007/3-540-44564-1

Julia Flores, M., Nicholson, A. E., Brunskill, A., Korb, K. B., Mascaro, S., Flores, J., … Mascaro, S. (2011). Incorporating expert knowledge when learning Bayesian network structure: A medical case study. *Artificial Intelligence in Medicine*, *53*(3), 181–204. http://doi.org/10.1016/j.artmed.2011.08.004

Junges, R., & Klügl, F. (2011). Evolution for modeling - A Genetic

Programming Framework for SeSAm. In *Proceedings of the 13th annual conference companion on Genetic and evolutionary computation - GECCO ’11* (p. 551). New York, New York, USA: ACM Press. http://doi.org/10.1145/2001858.2002047

Karr, A. F. (2014). Why data availability is such a hard problem. *Statistical Journal of the IAOS*, *30*(2), 101–107. http://doi.org/10.3233/SJI-140809

Kattan, A., Ong, Y. S., Galvan-Lopez, E., & Ieee. (2013). Multi-Agent Multi-Issue Negotiations with Incomplete Information: A Genetic Algorithm Based on Discrete Surrogate Approach. In *2013 Ieee Congress on Evolutionary Computation* (pp. 2556–2563). IEEE. http://doi.org/10.1109/CEC.2013.6557877

Kaya, M., & Alhajj, R. (2005). Fuzzy OLAP Association Rules Mining-Based Modular Reinforcement Learning Approach for Multiagent Systems. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *35*(2), 326–338. http://doi.org/10.1109/TSMCB.2004.843278

Kellermann, K., & Balmann, A. (2006). How Smart Should Farms be Modeled? Behavioral Foundation of Bidding Strategies in Agent-Based Land Market Models. *Contributed Paper to the 26th Conference of the International Association of Agricultural Economists Conference,*.

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*(2), 544–564. http://doi.org/10.1016/j.dss.2007.07.001

Kim, S., & Kim, S. (2018). Exploring the determinants of perceived risk of middle east respiratory syndrome (MERS) in Korea. *International Journal of Environmental Research and Public Health*, *15*(6). http://doi.org/10.3390/ijerph15061168

Kirman, A. (2010). Learning in agent based models. *Eastern Economic Journal*, *37*(December), 20–27.

Kitchovitch, S., Liò, P., Li, P., & Liò, P. (2010). Risk perception and disease spread on social networks. *Procedia Computer Science*, *1*(1), 2345–2354. http://doi.org/10.1016/j.procs.2010.04.264

Kocabas, V., & Dragicevic, S. (2013). Bayesian networks and agent-based modeling approach for urban land-use and population density change: a BNAS model. *Journal of Geographical Systems*, *15*(4), 403–426. http://doi.org/10.1007/s10109-012-0171-2

Kontkanen, P., Kontkanen, P., Myllymäki, P., Silander, T., & Tirri, H. (1997). A Bayesian Approach to Discretization. *PROCEEDINGS OF*

THE EUROPEAN SYMPOSIUM ON INTELLIGENT TECHNIQUES, 265--268.

Kurzweil, R. (1990). The Age of Intelligent Machines. *Cambridge: MIT Press*, *579*.

Laite, R., Potman, N., Sankaranarayanan, K., Portman, N., & Sankaranarayanan, K. (2016). Behavioral Analysis of Agent Based Service Channel Design Using Neural Networks. In *Proceedings of the 2016 Winter Simulation Conference* (pp. 3694–3695). IEEE. http://doi.org/10.1109/WSC.2016.7822404

Langley, P. (1988). *Elements of Machine Learning*. 1996 Morgan Kaufmann Publishers, Inc.

Langley, P., & Simon, H. A. (1995). Applications of machine learning and rule induction. *Institute for the Study of Learning and Expertise Technical*.

Lei, Z. (2005). Distributed Modeling Architecture of a Multi-Agent-Based Behavioral Economic Landscape (MABEL) Model. *Simulation*, *81*(7), 503–515. http://doi.org/10.1177/0037549705058067

Li, B., Sun, D., Guo, S., & Lin, Z. (2014). Agent Based Simulation of Group Emotions Evolution and Strategy Intervention in Extreme Events. *Discrete Dynamics in Nature and Society*, *2014*, 1–17. http://doi.org/10.1155/2014/464190

Li, L., Xiao, L., Zhang, M., Kong, J., & Sun, H. (2018). Agent-based simulation of evacuation considering extreme weather in tourist areas. In *Proceedings - 13th International Conference on Computational Intelligence and Security, CIS 2017* (Vol. 2018–Janua, pp. 431–434). IEEE. http://doi.org/10.1109/CIS.2017.00100

Liao, Y., Xu, B., Wang, J., & Liu, X. (2017). A new method for assessing the risk of infectious disease outbreak. *Scientific Reports*, *7*(1), 40084. http://doi.org/10.1038/srep40084

Lim, K. L., Mann, I., Santos, R., Tobin, B., Berryman, M. J., Abbott, D., & Ryan, A. (2005). Adaptive battle agents: complex adaptive combat models, *6039*, 603907-603907–13. http://doi.org/10.1117/12.638877

Liu, Z., Malone, B., & Yuan, C. (2012). Empirical evaluation of scoring functions for Bayesian network model selection. *BMC Bioinformatics*, *13*(Suppl 15), S14. http://doi.org/10.1186/1471-2105-13-S15-S14

Lorscheid, I. (2014). Learning Agents for Human Complex Systems. *2014 IEEE 38th International Computer Software and Applications Conference Workshops*, 432–437. http://doi.org/10.1109/COMPSACW.2014.73

Luger, G. F., & Stubblefield, W. . (1993). Artificial Intelligence: Its Roots and Scope. *Artificial Intelligence: Structures and Strategies For*, 1–34.

Ma, L., Arentze, T., Borgers, A., & Timmermans, H. (2004). Using Bayesian Decision Networks for Knowledge Representation under Conditions of Uncertainty in Multi-Agent Land Use Simulation Models. *Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning*.

Macal, C., & North, M. (2015). Introductory tutorial: Agent-based modeling and simulation. In *Proceedings - Winter Simulation Conference* (Vol. 2015–Janua, pp. 6–20). http://doi.org/10.1109/WSC.2014.7019874

Magliocca, N. R., & Walls, M. (2018). The role of subjective risk perceptions in shaping coastal development dynamics. *Computers, Environment and Urban Systems*, *71*, 1–13. http://doi.org/10.1016/j.compenvurbsys.2018.03.009

Manahov, V., Hudson, R., & Linsley, P. (2014). New evidence about the profitability of small and large stocks and the role of volume obtained using Strongly Typed Genetic Programming. *Journal of International Financial Markets, Institutions and Money*, *33*, 299–316. http://doi.org/10.1016/j.intfin.2014.08.007

Manfredi, P., & D'Onofrio, A. (2013). *Modeling the interplay between human behavior and the spread of infectious diseases*. *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*. Springer. http://doi.org/10.1007/978-1-4614-5474-8

Manson, S. M. (2005). Agent-based modeling and genetic programming for modeling land change in the Southern Yucatán Peninsular Region of Mexico. *Agriculture, Ecosystems & Environment*, *111*(1–4), 47–62. http://doi.org/10.1016/j.agee.2005.04.024

Matsumoto, O., Miyazaki, M., Ishino, Y., & Takahashi, S. (2017). Method for Getting Parameters of Agent-Based Modeling Using Bayesian Network: A Case of Medical Insurance Market. *Agent-Based Approaches in Economics and Social Complex Systems*, 45–57. http://doi.org/10.1007/978-981-10-3662-0_4

Matthews, R. R., Gilbert, N., Roach, A., Polhill, J. G., Gotts, N., & Polhill, J. G. (2007). Agent-based land-use models : a review of applications. *Landscape …*, 1–32.

McCluskey, J., & Swinnen, J. (2011). The media and food-risk perceptions. *EMBO Reports*, *12*(7), 624–629. http://doi.org/10.1038/embor.2011.118

Meister, K., Frick, M., & Axhausen, K. W. (2005). A GA-based household scheduler. *Transportation*, *32*, 473–494. http://doi.org/10.1007/s11116-005-5325-3

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

Michalski, R. S., & Carbonell, J. G. (2013). Machine Learning: An Artificial Intelligence Approach (Symbolic Computation). *Springler*, 572.

Miezah, K., Obiri-Danso, K., Kadar, Z., Fei-Baffoe, B., & Mensah, M. Y. (2015). Municipal solid waste characterization and quantification as a measure towards effective waste management in Ghana. *Waste Management*, *46*, 15–27. http://doi.org/10.1016/j.wasman.2015.09.009

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Science/Engineering/Math.

Monto, M. (Mani), Ganesh, L. S., & Varghese, K. (2005). *Sustainability and human settlements : fundamental issues, modeling and simulations*. Sage Publications.

Ndeffo Mbah, M. L., Liu, J., Bauch, C. T., Tekel, Y. I., Medlock, J., Meyers, L. A., & Galvani, A. P. (2012). The impact of imitation on vaccination behavior in social contact networks. *PLoS Computational Biology*, *8*(4), e1002469. http://doi.org/10.1371/journal.pcbi.1002469

Needham, C. J., Bradford, J. R., Bulpitt, A. J., & Westhead, D. R. (2007, August). A primer on learning in Bayesian networks for computational biology. *PLoS Computational Biology*. Public Library of Science. http://doi.org/10.1371/journal.pcbi.0030129

Nilsson, N. J. (1998). Artificial Intelligence: A New Synthesis. *Morgan Kaufmann Publishers Inc.*

Nishizaki, I., Katagiri, H., Oyama, T., & Hayashida, T. (2008). Agent-based simulation analysis for social norms. *2008 IEEE International Conference on Systems, Man and Cybernetics*, (i), 2105–2110. http://doi.org/10.1109/ICSMC.2008.4811602

North, M. J., & Macal, C. M. (2011). Product Design Patterns for Agent-based Modeling. *Proceedings of the 2011 Winter Simulation Conference*, (1996), 3087–3098.

Oloo, F., & Wallentin, G. (2017). An Adaptive Agent-Based Model of Homing Pigeons: A Genetic Algorithm Approach. *ISPRS International Journal of Geo-Information*, *6*(1), 27. http://doi.org/10.3390/ijgi6010027

Osei, F. B., & Duker, A. A. (2008). Spatial dependency of V. cholera prevalence on open space refuse dumps in Kumasi, Ghana: a spatial statistical modelling. *International Journal of Health Geographics*, *7*(1), 62. http://doi.org/10.1186/1476-072X-7-62

Osei, F. B., Duker, A. A., Augustijn, E. W., & Stein, A. (2010). Spatial dependency of cholera prevalence on potential cholera reservoirs in an urban area, Kumasi, Ghana. *International Journal of Applied Earth Observation and Geoinformation*, *12*(5), 331–339. http://doi.org/10.1016/j.jag.2010.04.005

Ozik, J., Collier, N. T., Wozniak, J. M., Macal, C. M., & An, G. (2018). Extreme-scale dynamic exploration of a distributed agent-based model with the EMEWS framework. *IEEE Transactions on Computational Social Systems*, *5*(3), 884–895. http://doi.org/10.1109/TCSS.2018.2859189

Pablo, A. L., Sitkin, S. B., & Jemison, D. B. (1996). Acquisition decision-making processes: The central role of risk. *Journal of Management*, *22*(5), 723–746. http://doi.org/10.1016/S0149-2063(96)90020-3

Park, H.-S. S., & Cho, S.-B. B. (2012). A modular design of Bayesian networks using expert knowledge: Context-aware home service robot. *Expert Systems with Applications*, *39*(3), 2629–2642. http://doi.org/10.1016/j.eswa.2011.08.118

Parker, D. C. (2014). *Oxford Handbooks Online An Economic Perspective on Agent-Based Models of Land Use and Land Cover Change*. *The Oxford Handbook of Land Economics*. Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199763740.013.014

Parker, D. C., Manson, S. M., Janssen, M. a., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers*, *93*(2), 314–337. http://doi.org/10.1111/1467-8306.9302004

Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998). Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide. *Proceedings of the First International Workshop on Multi-Agent Systems and Agent-Based Simulation*, 10–25.

Perez, L., & Dragicevic, S. (2009). An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, *8*(1), 50. http://doi.org/10.1186/1476-072X-8-50

Pizzitutti, F., Pan, W., Feingold, B., Zaitchik, B., Álvarez, C. A., & Mena, C. F. (2018). Out of the net: An agent-based model to study human movements influence on local-scale malaria transmission. *PLoS ONE*, *13*(3), e0193493. http://doi.org/10.1371/journal.pone.0193493

Plikynas, D., Raudys, A., & Raudys, S. (2014). Agent-based modelling of excitation propagation in social media groups. *Journal of Experimental & Theoretical Artificial Intelligence*.

Polhill, J. G., Gotts, N. M., & Law, A. N. R. (2001). Imitative versus nonimitative strategies in a land-use simulation. *Cybernetics and Systems*, *32*(1–2), 285–307. http://doi.org/10.1080/019697201300001885

Pooyandeh, M., & Marceau, D. J. (2014a). Incorporating Bayesian learning in agent-based simulation of stakeholders' negotiation. *Computers, Environment and Urban Systems*, *48*, 73–85. http://doi.org/10.1016/j.compenvurbsys.2014.07.003

Pooyandeh, M., & Marceau, D. J. (2014b). Incorporating Bayesian learning in agent-based simulation of stakeholders' negotiation. *Computers, Environment and Urban Systems*, *48*, 73–85. http://doi.org/10.1016/j.compenvurbsys.2014.07.003

Pope, A., & Gimblett, R. (2017). Linking Participatory, Bayesian, and Agent- Based Modeling Techniques to Simulate Coupled Natural-Human System: A Case Study with Ranchers in Sonora, Mexico. In S. Gray (Ed.), *Environmental Modeling with Stakeholders* (pp. 345–359). http://doi.org/10.1007/978-3-319-25053-3

Pope, A. J., & Gimblett, R. (2015). Linking Bayesian and agent-based models to simulate complex social-ecological systems in semi-arid regions. *Frontiers in Environmental Science*, *3*, 55. http://doi.org/10.3389/fenvs.2015.00055

Quteishat, A., Peng Lim, C., Tweedale, J., & Jain, L. C. (2009). A neural network-based multi-agent classifier system. *Neurocomputing*, *72*(7–9), 1639–1647. http://doi.org/10.1016/j.neucom.2008.08.012

Rand, W. (2006). Machine learning meets agent-based modeling: When not to go to a bar. *The Proceedings of Agent*.

Reddy, P. P., & Veloso, M. M. (2011). Strategy learning for autonomous agents in smart grid markets. http://doi.org/10.5591/978-1-57735-516-8/IJCAI11-244

Reijn, E., Swaan, C. M., Kretzschmar, M. E. E., & Van Steenbergen, J. E. (2011). Analysis of timeliness of infectious disease reporting in the Netherlands. *BMC Public Health*, *11*(1), 409. http://doi.org/10.1186/1471-2458-11-409

Rekik, Y. M., Hachicha, W., & Boujelbene, Y. (2014). Agent-based Modeling and Investors' Behavior Explanation of Asset Price Dynamics on Artificial Financial Markets. *Procedia Economics and Finance*, *13*, 30–

46. http://doi.org/10.1016/S2212-5671(14)00428-6

Ren, Z., & Anumba, C. . (2002). Learning in multi-agent systems: a case study of construction claims negotiation. *Advanced Engineering Informatics*, *16*(4), 265–275. http://doi.org/10.1016/S1474-0346(03)00015-6

Rennard, J. P. (2006). *Handbook of Research on Nature-Inspired Computing for Economics and Management*. *Social Sciences* (Vol. 1). http://doi.org/10.4018/978-1-59140-984-7

Rogers, R. W. (1983). Cognitive and Physiological Processes in Fear Appeals and Attitute Change: A Revised Theory of Porotection Motivation. In *Social Psychophysiology: A Sourcebook* (pp. 153–177). http://doi.org/10.1093/deafed/ent031

Rübsamen, N., Castell, S., Horn, J., Karch, A., Ott, J. J., Raupach-Rosin, H., … Mikolajczyk, R. T. (2015). Ebola risk perception in Germany, 2014. *Emerging Infectious Diseases*, *21*(6), 1012–1018. http://doi.org/10.3201/eid2106.150013

Rufat, S., & Samuel. (2015). Towards a Social and Spatial Risk Perception Framework. *Cybergeo*. http://doi.org/10.4000/cybergeo.27010

Ruland, B. E. C., Dinca, I., Curtis, V., Barry, M. M., Ekdahl, K., & Timen, A. (2015). Learning from each other: where health promotion meets infectious diseases. *Eurohealth INTERNATIONAL Eurohealth Incorporating Euro Observer*, *21*(1), 13–17.

Russ, C., & Walz, A. (2009). MACSIMA: on the effects of adaptive negotiation behavior in agent-based supply networks. *Multiagent System Technologies*, 128–140. http://doi.org/10.1007/978-3-642-04143-3_12

Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (Third). PEARSON. http://doi.org/10.1016/B978-012161964-0/50009-1

Russell, S., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. *Pearson Education, Inc., Upper Saddle River* (Third Edit). Pearson; 3 edition (December 11, 2009).

Sahli, N., & Moulin, B. (2006). Agent-based geo-simulation to support human planning and spatial cognition. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 3891 LNAI, pp. 115–132). http://doi.org/10.1007/11734680_9

Salle, I. L. (2015). Modeling expectations in agent-based models — an application to central bank's communication and monetary policy. *Economic Modelling*, *46*, 130–141.

http://doi.org/10.1016/j.econmod.2014.12.040

Seidl, R., Kuhn, S., Elbers, M., Ernst, A., & Klemm, D. (2016). Modelling Risk Perception and Indicators of Psychosocial Sustainability in Private Households: The Risk Perception Module in DeepHousehold. In *Regional Assessment of Global Change Impacts* (pp. 347–353). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-16751-0_44

Sen, S., & Weiss, G. (1999). Learning in multi-agent systems. In G. Weiss (Ed.), *Multiagent Systems* (pp. 259–298). MIT Press.

Shalev-Shwartz, S., & Ben-David, S. (2013). *Understanding machine learning: From theory to algorithms*. *Understanding Machine Learning: From Theory to Algorithms* (Vol. 9781107057). http://doi.org/10.1017/CBO9781107298019

Sharma, S., Otunba, S., Ogunlana, K., & Tripathy, T. (2012). Intelligent agents in a goal finding application for homeland security. *2012 Proceedings of IEEE Southeastcon*, 1–5. http://doi.org/10.1109/SECon.2012.6197067

Shen, Y., Guo, Y., Fang, Z., & Ma, J. (2016). Multi-agent simulation model of China's real estate market based on Bayesian network decision making. *Journal of Grey System*, *28*(1), 127–134. http://doi.org/10.1109/GSIS.2015.7301883

Shen, Y., Liu, S. F., Fang, Z. G., & Hu, M. L. (2011). Modeling and simulation of stranded passengers' transferring decision-making on the basis of herd behavior. *Proceedings of 2011 IEEE International Conference on Grey Systems and Intelligent Services, GSIS'11 - Joint with the 15th WOSC International Congress on Cybernetics and Systems*, (90924022), 795–802. http://doi.org/10.1109/GSIS.2011.6043958

Sierhuis, M., & Diegelman, T. E. (2007). Agent-based simulation of shuttle mission operations. *Proceedings of the 2007 Spring Simulation Multiconference - Volume 2*, 53–60.

Simoes, J. A. (2012). An Agent-Based/Network Approach to Spatial Epidemics. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.), *Agent-Based Models of Geographical Systems* (pp. 591–610). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-90-481-8927-4

Sitkin, S. B., & Weingart, L. R. (1995). Determinants of Risky Decision-Making Behavior: a Test of the Mediating Role of Risk Perceptions and Propensity. *Academy of Management Journal*, *38*(6), 1573–1592. http://doi.org/10.2307/256844

Sjöberg, L. (2000). The methodology of risk perception research. *Quality and Quantity*, *34*(4), 407–418. http://doi.org/10.1023/A:1004838806793

Slovic, P. (2010). *The feeling of risk : new perspectives on risk perception*. Earthscan.

Smeureanu, I., Ruxanda, G., Diosteanu, A., Delcea, C., & Cotfas, L. A. (2012). Intelligent agents and risk based model for supply chain management. *Technological and Economic Development of Economy*, *18*(3), 452–469. http://doi.org/10.3846/20294913.2012.702696

Sobkowicz, P. (2017). Opinion dynamics model based on cognitive biases.

Sridhar, S., Régner, I., Brouqui, P., & Gautret, P. (2016, July 1). Methodologies for measuring travelers' risk perception of infectious diseases: A systematic review. *Travel Medicine and Infectious Disease*. Elsevier. http://doi.org/10.1016/j.tmaid.2016.05.012

Stone, P., & Veloso, M. M. (2000). MultiAgent Systems: A Survey from a Machine Learning Perpective. *Autonomous Robots*, *8*(3), 345–383. http://doi.org/10.1023/A:1008942012299

Sukhbaatar, S., Szlam, A., & Fergus, R. (2016). Learning Multiagent Communication with Backpropagation.

Sun, Y., & Cheng, L. (2005). A Survey on Agent-Based Modelling and Equation-based Modelling. *Citeseerx*.

Sun, Z., & Müller, D. (2012). A framework for modeling payments for ecosystem services with agent-based models, Bayesian belief networks and opinion dynamics models. *Environmental Modelling & Software*, *45*, 1–14. http://doi.org/10.1016/j.envsoft.2012.06.007

Talwar, A., & Kumar, Y. (2013). Machine Learning: An artificial intelligence methodology. *International Journal of Engineering and Computer Science*, *2*(12), 3400–3405.

Tan, X., Li, S., Wang, C., Chen, X., & Wu, X. (2004). Severe Acute Respiratory Syndrome epidemic and change of people's health behavior in China. *Health Education Research*, *19*(5), 576–580. http://doi.org/10.1093/her/cyg074

Tang, L., & Liu, H. (2009). Scalable learning of collective behavior based on sparse social dimensions. In *Proceeding of the 18th ACM conference on Information and knowledge management - CIKM '09* (p. 1107). New York, New York, USA: ACM Press. http://doi.org/10.1145/1645953.1646094

Tang, X., Zhao, S., Chiu, A. P. Y., Ma, H., Xie, X., Mei, S., ... He, D. (2017). Modelling the transmission and control strategies of varicella among school children in Shenzhen, China. *PLoS ONE*, *12*(5), e0177514.

http://doi.org/10.1371/journal.pone.0177514

Tkachuk, K., Song, X., & Maltseva, I. (2018). Application of artificial neural networks for agent-based simulation of emergency evacuation from buildings for various purpose. In *IOP Conference Series: Materials Science and Engineering* (Vol. 365, p. 042064). IOP Publishing. http://doi.org/10.1088/1757-899X/365/4/042064

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, *203*(3–4), 312–318. http://doi.org/10.1016/J.ECOLMODEL.2006.11.033

Vallejo, M., Corne, D. W., & Rieser, V. (2013). Evolving urbanisation policies: Using a statistical model to accelerate optimisation over agent-based simulations. In *ICAART 2013 - Proceedings of the 5th International Conference on Agents and Artificial Intelligence* (Vol. 2, pp. 171–181).

van der Hoog, S. (2016). Deep Learning in Agent-Based Models: A Prospectus. *Ssrn*. http://doi.org/10.2139/ssrn.2711216

Van Der Hoog, S. (2016). *Deep Learning in Agent-Based Models: A Prospectus Deep Learning in Agent-Based Models: A Prospectus*.

Van Der Ploeg, T., Austin, P. C., & Steyerberg, E. W. (2014). Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. *BMC Medical Research Methodology*, *14*(1), 137. http://doi.org/10.1186/1471-2288-14-137

van Duinen, R., Filatova, T., Jager, W., & van der Veen, A. (2016). Going beyond perfect rationality: drought risk, economic choices and the influence of social networks. *Annals of Regional Science*, *57*(2–3), 335–369. http://doi.org/10.1007/s00168-015-0699-4

Van Kerkhove, M. D., & Ferguson, N. M. (2012). Epidemic and intervention modelling – a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bulletin of the World Health Organization*, *90*(4), 306–310. http://doi.org/10.2471/BLT.11.097949

Venkatramanan, S., Lewis, B., Chen, J., Higdon, D., Vullikanti, A., & Marathe, M. (2018, February 1). Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics*, pp. 43–49. Elsevier. http://doi.org/10.1016/j.epidem.2017.02.010

Verstegen, J. a, Ligtenberg, A., & Wachowicz, M. (2010). Learning Actors in Spatial Planning: Incorporating Bayesian Networks in an Agent Based Model. In *Pwktech.Info* (pp. 1–10).

Wachinger, G., & Renn, O. (2010). *Risk perception of natural hazards Chiara Bianchizza*.

Walczak, S., & Walczak, S. (2001). An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks. *Journal of Management Information Systems*, *1222*(August), 203–222. http://doi.org/10.1080/07421222.2001.11045659

Weber, P., Medina-Oliva, G., Simon, C., & Iung, B. (2012). Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*. http://doi.org/10.1016/j.engappai.2010.06.002

Williams, D. J., & Noyes, J. M. (2007). How does our perception of risk influence decision-making? Implications for the design of risk information. *Theoretical Issues in Ergonomics Science*, *8*(1), 1–35. http://doi.org/10.1080/14639220500484419

Williams, L., Collins, A. E., Bauaze, A., & Edgeworth, R. (2010). The role of risk perception in reducing cholera vulnerability. *Risk Management*, *12*(3), 163–184. http://doi.org/10.1057/rm.2010.1

Wilson, E. B., & Burke, M. H. (1943). The Epidemic Curve. *Proceedings of the National Academy of Sciences of the United States of America*, *29*(1), 43–48. http://doi.org/10.1073/pnas.29.1.43

Winston, P. H. (1992). Artificial Intelligence.

Wise, S. (2014). Using Social Media Content to Inform Agent Based Models for Humanitarian Crisis Response. http://doi.org/10.1017/CBO9781107415324.004

Xu, F. Y., Wang, X., Lai, L. L., & Lai, C. S. (2013). Agent-Based Modeling and Neural Network for Residential Customer Demand Response. In *2013 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1312–1316). IEEE. http://doi.org/10.1109/SMC.2013.227

Yang, S., & Cho, S. Il. (2017). Middle East respiratory syndrome risk perception among students at a university in South Korea, 2015. *American Journal of Infection Control*, *45*(6), e53–e60. http://doi.org/10.1016/j.ajic.2017.02.013

Yang, Y., Diez Roux, A. V, Auchincloss, A. H., Rodriguez, D. a, & Brown, D. G. (2011). A spatial agent-based model for the simulation of adults' daily walking within a city. *American Journal of Preventive Medicine*, *40*(3), 353–61. http://doi.org/10.1016/j.amepre.2010.11.017

Yet, B., Bastani, K., Raharjo, H., Lifvergren, S., Marsh, W., & Bergman, B. (2013). Decision support system for Warfarin therapy management using Bayesian networks. *Decision Support Systems*, *55*(2), 488–498. http://doi.org/10.1016/j.dss.2012.10.007

Yospin, G. I., Bridgham, S. D., Neilson, R. P., Bolte, J. P., Bachelet, D. M., Gould, P. J., … Johnson, B. R. (2015). A new model to simulate climate-change impacts on forest succession for local land management. *Ecological Applications*, *25*(1), 226–242. http://doi.org/10.1890/13-0906.1

Yuan, M., Liu, Y., He, J., & Liu, D. (2014). Regional land-use allocation using a coupled MAS and GA model: from local simulation to global optimization, a case study in Caidian District, Wuhan, China. *Cartography and Geographic Information Science41*, *41*(4), 363–378. http://doi.org/10.1080/15230406.2014.931251

Yuksel, M. E. (2018). Agent-based evacuation modeling with multiple exits using NeuroEvolution of Augmenting Topologies. *Advanced Engineering Informatics*, *35*, 30–55. http://doi.org/10.1016/j.aei.2017.11.003

Zacharias, G. L., MacMillan, J., & Hemel, S. B. Van. (2008). *Behavioral Modeling and Simulation: From Individuals to Societies*. *National Academies Press*. National Research Council. http://doi.org/https://doi.org/10.17226/12169

Zhang, H. H., Zeng, Y. N., & Bian, L. (2010). Simulating multi-objective spatial optimization allocation of land use based on the integration of multi-agent system and genetic algorithm. *International Journal of Environmental Research*, *4*(4), 765–776.

Zhang, M., Tan, Z., Zhao, J., & Li, L. (2008). A Bayesian Learning Model in the Agent-based Bilateral Negotiation between the Coal Producers and Electric Power Generators. *2008 International Symposium on Intelligent Information Technology Application Workshops*, (2), 859–862. http://doi.org/10.1109/IITA.Workshops.2008.144

Zhao, S., Kuang, Y., Wu, C.-H. C.-H., Bi, K., & Ben-Arieh, D. (2018). Risk perception and human behaviors in epidemics. *IISE Transactions on Healthcare Systems Engineering*, *0*(0), 1–14. http://doi.org/10.1080/24725579.2018.1464085

Zhao, S., Wu, J., & Ben-Arieh, D. (2015). Modeling infection spread and behavioral change using spatial games. *Health Systems*, *4*(October), 41–53. http://doi.org/10.1057/hs.2014.22

Summary

Complexity in human behaviour can play a crucial role in socio-environmental processes like disease diffusion. An example of such complex behaviour is risk perception, and behavioural change due to perceived risk. Computational models, and in particular Agent-based models (ABMs), have evolved as tools for simulating complex real-world processes.

ABMs for describing and simulating a system composed of behavioural entities, ABMs provide the most natural environment. ABMs often use naive deterministic algorithms, which are rule-based, to simulate behavioural change in agents. While agents in ABMs are sometimes endowed with memory, the actual learning in machine learning style is rarely implemented. The endogenous switching of expectations formation strategies using learning algorithm is underdeveloped in ABMs.

The goal of my PhD research is to systematically test the effects of implementing social and environmental intelligence on the dynamics and emergent outcomes of spatial ABM. Spatial ABMs often use spatial data (GIS data) to construct real geographic environments in which agents are situated. Agents need to take changes in the spatial environment into account and adjust their behaviour accordingly. In this PhD research, intelligence, rational, and risk perception are playing an important role in the decision making of agents. Understanding the learning processes of agents in the spatial ABM can assist developing better strategies in problem-solving and coordination mechanisms.

Learning algorithms allow for a richer agents' architecture for operationalization of more realistic learning decisions beyond a simplistic treatment of agents' cognitive and sensory capacities. Chapter two reviews recent spatial ABMs that employ different learning algorithms to create intelligent agents and steer their behaviour. We provide a systematic structured analysis of 1) the growth rate of integrating learning algorithms with spatial ABMs, 2) the reasons that motivate researchers to use learning algorithms in their models, and 3) the specific operationalization of agent's decision-making for various tasks, and treatment of spatial environment in the design of learning algorithms. This chapter highlights the trends in the current practice of learning algorithms used to enhance ABMs, which social simulation modellers may rely on when designing their spatial ABM simulations.

Chapter three presents an innovative approach to extend agent-based disease models by capturing behavioural aspects of decision-making in a risky context using machine learning techniques. We illustrate it with a case of cholera in Kumasi, Ghana, accounting for spatial and social risk factors that affect intelligent behaviour and corresponding disease incidents. The results of computational experiments comparing intelligent with zero-intelligent representations of agents in a spatial disease agent-based model are discussed. We present a spatial disease ABM with agents' behavior grounded in Protection Motivation Theory (PMT). Spatial and temporal patterns of disease diffusion among zero-intelligent agents are compared to those produced by a population of intelligent agents. Two Bayesian Networks (BNs) designed and coded using R and are further integrated with the NetLogo-based CABM. The first is a one-tier BN1 (only risk perception), the second is a two-tier BN2 (risk and coping behavior). Our results emphasize the importance of integrating behavioural aspects of decision making under risk into spatial disease ABMs using machine learning algorithms. This is especially relevant when studying cumulative impacts of behavioural changes and possible intervention strategies.

There is a difference between ABMs with pure social intelligence based on information exchange among agents and ABMs with integrated spatial intelligence. Spatial intelligence refers to the fact that agents sense their environment, perform a judgement on the condition of this environment, and change their behaviour based on this judgement. When spatial intelligence is used in ABMs, it often facilitates navigation (human or animal) or adaptation to land cover change. Less implementations are available for assessing risky situation engaging agents' risk perception. In chapter three, agents evaluate changes in floating plastic debris in a river combined with personal information and media attention on cholera to decide which water source to use. In chapter four, data to validate the spatial intelligence was collected via two online surveys were run to gather data on people's risk perception for cholera: MOOC survey (Geohealth online course) and Google survey (an online survey). While most of the questions were identical in the two surveys, there was one difference. In the MOOC survey participants chose to use or not to use river water for drinking through judging about its quality by the visual appearance (pictures shown). The Google survey collected information on the influence of individual risk factors on the willingness to use the river water without visuals using only textual description of the water quality situation. The risk perception of participants is questioned based on one factor and a combination of factors. Results from the survey confirm the fact that people judge quality of water visually, but also show the strong influence of media on risk perception.

Learning algorithms steer agent decisions in ABMs, serving as a vehicle for implementing behaviour changes during simulation runs. However, when training an ML algorithm, obtaining large sets of micro-level human behaviour data is often problematic. Information on human behaviour is often collected via surveys of relatively small sample sizes. Chapter five presents a methodology for training a learning algorithm to guide agent behaviour in CABM using a limited survey data sample. We apply different implementation strategies using survey data and BNs. By being grounded in probabilistic directed graphical models, BNs stand out among other learning algorithms in that they can be based on expert knowledge and/or known datasets. This chapter presents four alternative implementations of data-driven BNs to support agent decisions in CABM. We differentiate between training BNs prior to, or during the simulation runs, using only survey data or a combination of survey data and expert knowledge. The four different implementations are then illustrated using the CABM. The results indicate that a balance between expert knowledge and survey data provides the best control over the learning process of the agents and produces the most realistic agent behaviour.

Adaptive behaviour of agents is contingent on how well they learn about changes in disease risks and about coping options, individually or in interactions with others. The impact of different types of group learning compared to individual learning is an underexplored domain in disease modelling, and in agent-based models of socio-environmental systems in general. Chapter six pursues a quantitative test on the influence of agents' ability to learn – individually or in a group – on the disease dynamics. Our experiments illustrate that individual intelligent judgements about disease risks and the selection of disease coping actions are outperformed by social intelligence (individually or leader-based). While the majority vote performs poorly here. Importantly, the choice of a particular type of individual or group learning in agents-based models should account for the nature and cultural norms of the society, for which epidemics prevention strategies are being tested.

Samenvatting

Complexiteit in menselijk gedrag kan een cruciale rol spelen in sociaalecologische processen zoals ziekteverspreiding. Een voorbeeld van een dergelijk complex gedrag is risicoperceptie, en gedragsverandering als gevolg van dit waargenomen risico. Computationele modellen, en met name op agenten gebaseerde modellen (ABM's), zijn geëvolueerd als hulpmiddelen voor het simuleren van complexe, realistische processen.

ABM's gebruiken vaak naïeve deterministische algoritmen, om gedragsverandering te simuleren. Hoewel agenten in ABM's soms over geheugen beschikken, wordt het feitelijke leren in de vorm van zelflerende systemen zelden geïmplementeerd. Het endogene omschakelen van strategieën voor het formuleren van verwachtingen met behulp van leeralgoritmen is onderontwikkeld in ABM's.

Het doel van mijn promotieonderzoek is om de effecten van het implementeren van sociale en ruimtelijke-informatie op de dynamiek en de emergente uitkomsten van ABMs systematisch te testen. Ruimtelijke ABM's gebruiken vaak omgevingsinformatie (GIS-gegevens) om reële geografische omgevingen te bouwen waarin agenten zich bevinden. Agenten moeten rekening houden met veranderingen in de ruimtelijke omgeving en hun gedrag dienovereenkomstig aanpassen. In dit doctoraatsonderzoek spelen intelligentie, ratio en risicoperceptie een belangrijke rol bij de besluitvorming van agenten. Het begrijpen van de leerprocessen van agenten in een ruimtelijke ABM kan helpen bij het ontwikkelen van betere strategieën in probleemoplossende en coördinatiemechanismen.

Leeralgoritmen maken een rijkere agentenarchitectuur mogelijk voor het operationaliseren van meer realistische beslissingen dan een simplistische behandeling van de cognitieve en zintuiglijke capaciteiten van agenten. Hoofdstuk twee geeft een overzicht van recente ruimtelijke ABM's die gebruikmaken van verschillende leeralgoritmen om intelligente agenten te creëren en hun gedrag te sturen. We bieden een systematische gestructureerde analyse van 1) trends in het integreren van leeralgoritmen in ruimtelijke ABM's, 2) motiveren om leeralgoritmen in modellen te gebruiken, en 3) de specifieke operationalisering van de besluitvorming door agenten voor verschillende taken, en behandeling van ruimtelijke omgeving bij het ontwerp van leeralgoritmen. Dit hoofdstuk belicht hoe leeralgoritmen worden gebruikt om ABM's te verbeteren, en geeft een overzicht van de technieken die ontwikkelaar van sociale simulatiemodellen kunnen gebruiken bij het ontwerpen van hun ruimtelijke ABM-simulaties.

Hoofdstuk drie presenteert een innovatieve benadering om agent-gebaseerde ziektemodellen uit te breiden door gedragsaspecten van besluitvorming vast te leggen in een risicovolle context met behulp van machinale leertechnieken. We illustreren dit aan de hand van een model voor het simuleren van cholera in Kumasi, Ghana. Dit model houdt rekening met ruimtelijke en sociale risicofactoren die van invloed zijn op intelligent gedrag en overeenkomstige ziekte-incidenten. De resultaten van computationele experimenten waarin een intelligent model wordt vergeleken met een model zonder leeralgoritmen worden besproken. We presenteren een ruimtelijke ziekte-ABM met agentengedrag, gefundeerd in Protection Motivation Theory (PMT). Ruimtelijke en temporele patronen van ziekteverspreiding onder nul-intelligente agentia worden vergeleken met die van een populatie van intelligente agenten. Twee Bayesiaanse netwerken (BN's), ontworpen en gecodeerd met R, werden verder geïntegreerd met het op NetLogo gebaseerde CABM model. Het eerste BN is een one-tier netwerk (BN1 - risicoperceptie), de tweede is een two-tier Bayesiaans netwerk (BN2 risico- en coping-gedrag). Onze resultaten benadrukken het belang van het integreren van gedragsaspecten van besluitvorming onder risico in ruimtelijke ziekte-ABM's met behulp van machine learning-algoritmen. Dit is vooral relevant bij het bestuderen van cumulatieve effecten van gedragsveranderingen en mogelijke interventiestrategieën.

Er is een verschil tussen ABM's met pure sociale intelligentie op basis van informatie-uitwisseling tussen agenten en ABM's met geïntegreerde ruimtelijke intelligentie. Ruimtelijke intelligentie verwijst naar het feit dat agenten hun omgeving waarnemen, een oordeel vellen over de toestand van deze omgeving en hun gedrag veranderen op basis van dit oordeel. Wanneer ruimtelijke intelligentie wordt gebruikt in ABM's, vergemakkelijkt het vaak de navigatie (mens of dier) of aanpassing aan de verandering van landbedekking. Er zijn minder implementaties beschikbaar om de risicoperceptie van risicovolle situaties te beoordelen. In hoofdstuk drie evalueren agenten veranderingen in drijvend plastic afval in een rivier in combinatie met persoonlijke informatie en media-aandacht voor cholera om te bepalen welke waterbron te gebruiken. In hoofdstuk vier wordt ruimtelijke intelligentie gevalideerd door informatie te verzamelen via twee online enquêtes over cholera risicoperceptie : MOOC survey (Geohealth online course) en Google survey (een online survey). Hoewel de meeste vragen in de twee enquêtes identiek waren, was er één verschil. In de

MOOC-enquête kozen de deelnemers om rivierwater al dan niet te gebruiken door de kwaliteit van water te oordelen op basis van getoonde foto's. De Google-enquête verzamelde informatie over de invloed van individuele risicofactoren op de bereidheid om het rivierwater te gebruiken zonder beelden, waarbij alleen de tekstuele beschrijving van de waterkwaliteitssituatie werd gebruikt. De risicoperceptie van deelnemers wordt getoetst op basis van één factor en een combinatie van factoren. Resultaten van de enquête bevestigen dat mensen de kwaliteit van het water visueel beoordelen, maar laten ook de sterke invloed van media op risicoperceptie zien.

Leeralgoritmen sturen agentbeslissingen in ABM's, en dienen als een hulpmiddel voor het implementeren van gedragsveranderingen tijdens simulatieruns. Bij het trainen van een ML-algoritme is het verkrijgen van grote sets gegevens over menselijk gedrag op microniveau echter vaak een probleem. Informatie over menselijk gedrag wordt vaak verzameld via enquêtes van relatief kleine steekproefgroottes. Hoofdstuk vijf presenteert een methodologie voor het trainen van een leeralgoritme voor het gedrag van agenten met behulp van een steekproef met beperkte enquêtegegevens. We passen verschillende implementatiestrategieën toe met behulp van enquêtegegevens en BN's. Door gegrond te zijn in probabilistisch gestuurde grafische modellen vallen BN's op tussen andere leeralgoritmen omdat ze gebaseerd kunnen zijn op expertkennis en / of bekende datasets. Dit hoofdstuk presenteert vier alternatieve implementaties van data-gebaseerde BN's ter ondersteuning van agentbeslissingen in CABM. We maken onderscheid tussen training van de BN's voorafgaand aan, of tijdens de simulatieruns, met alleen enquêtegegevens of een combinatie van onderzoeksgegevens en expertkennis. De vier verschillende implementaties worden vervolgens geïllustreerd met behulp van CABM. De resultaten geven aan dat een balans tussen expertkennis en onderzoeksgegevens de beste controle biedt over het leerproces van de agenten en het meest realistische gedrag van de agent oplevert.

Adaptief gedrag van agenten hangt af van hoe goed ze leren over veranderingen in ziekterisico's en over coping-opties, individueel of in interacties met anderen. De impact van verschillende soorten groepsleren in vergelijking met individueel leren is een onderbelicht domein in ziektemodellering en in agent-gebaseerde modellen van sociaal-ecologische systemen in het algemeen. Hoofdstuk zes voert een kwantitatieve test uit op de invloed van het vermogen van agenten om individueel of in groep te leren over de ziektedynamiek. Onze experimenten illustreren dat individuele intelligente oordelen over ziekterisico's en de selectie van handelingsacties ter voorkoming van de ziekte worden overtroffen door sociale intelligentie (individueel of op leiding gebaseerd). Terwijl risico evaluatie en de keuze voor handelingsacties op basis van meerderheid van stemmen hier slecht presteert. Belangrijk is dat de keuze voor een bepaald type individu of groepsleren in agents-gebaseerde modellen rekening moet houden met de aard en culturele normen van de samenleving, waarvoor epidemiologische preventiestrategieën worden getest.

ب کوردی و ب کورتی

روژانە چاڤێن خو ڤەدکەین تەماشەی مۆبایێلێن خو دکەین و ب رێکا مالپەرێن میدیا جڤاکی دەهان دەنگ و باسێن سەر کارەساتێن جودا جودا لسەرانسەری جیهانێ دخوینین. ئەگەرێن ڤان کارەساتا ژێك جودانە لێ هەمی بۆ ئێك ئەنجام دچن ئەو ژی زیانێن گیانی وئابوری. نەخوشیێن زوبیەلاڤ ئێك ژ وان کارەساتێن هەرە مەزنن یێن ب سەدان دکۆژییت رۆژانە ل ولاتێن جیهانێ. بەلاڤ بونا نەخوشیان چ جوداهیێ ناکەت ناڤ بەرا مرۆڤان ل هەر جهەکی ئاکنجی بن. پێت بیتە ئاستەنگ درێکا بەلاڤ بونێ دا رێکێ لبەر ڤان نەخۆشیا دگریت ئەو ئیجرائاتێن پاراستنێ نە کو دام و دەزگەهێن حکومی پێ رادبیت ژبو کونترول کرنا بەلاڤ بوونێ و قورتال کرنا خەلکی تایبەت بچیك و دان عەمران. لەورا، و ژبو حکومەت بشێن رێنما و یاساییێن دروست دارێژیت، سۆفتوێر و مودێلێن زانستی و تەکنیکی پێتڤیە بهێنە دابین کرن.

ئەڤ مودێلێن هە هاریکاریا جهێن بەرپرس دکەن ژبو تێگەهشتن و شروڤە کرنا رویدانێن بوری، چەوانیا بەلاڤ بوونا نەخوشیێن بەربلاڤ، و بجهئینانا سیناریوییێن بۆ هەر رویدانەکا داهاتی.

خاندنا من بو بەکالوریوسا زانستێن کۆمپیوتەری سەرەرای خاندنا من بو ماستەرێ ل بوارێ پێزانییێن ئەردی هاریکاریا من کریە ژبو دابیین کرنا تێگەهشتن و بنەرەتێن سەرەکی و گرنگ یێن پێتڤی بو چێکرن و پێشئخستن ڤان جورە سێمیولەیشنا وەك ئەلگۆرتمێن زیرەك، پرۆگرامکرن ب کومەکا زمانێن کۆمپیوتەری، چێکرن و رێڤەبەریا داتابەیسان، دگەل تێکنیکێن دن یێن پێتڤی. ڤان شیانێن هە هاریکاریا من کرینە ژبو پێشڤەبرنا مودێلەکێ نساخیا کۆلێرا ل باژێرەکێ ۆلاتێ غانا ل ئەفریقیا. ئەڤ مودێلێ هە هاتیە پێش ئێخستن ب رێکەکێ کۆ خەلکێ وی باژێری هاتیە نوینەرکرن ڤی خەلکی شیانێن هەین وەك مرۆڤان ل واقعی ژیانا خو برێڤە ببەن و نساخی بەلاڤە ببیت.