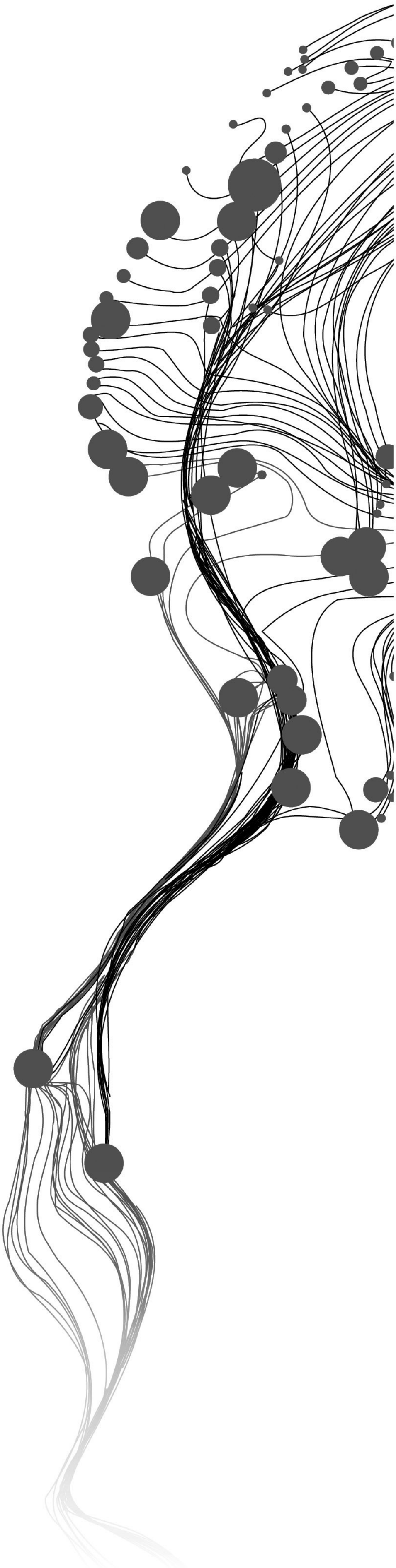
SIMULATING THE SPATIAL HIERARCHICAL SPREAD PATTERN OF PERTUSSIS IN THE TWENTE REGION USING AGENT-BASED MODELING.

ROSEMARY OKLA
March, 2011

SUPERVISORS:
Ms. Ir P.W.M (Ellen-Wien) Augustijn
Dr. R. (Raul) Zurita-Milla



SIMULATING THE SPATIAL HIERARCHICAL SPREAD PATTERN OF PERTUSSIS IN THE TWENTE REGION USING AGENT-BASED MODELING.

ROSEMARY OKLA
Enschede, The Netherlands, March, 2011

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

SUPERVISORS:

Ms. Ir P.W.M (Ellen-Wien) Augustijn
Dr. R. (Raul) Zurita-Milla

THESIS ASSESSMENT BOARD:

Dr.Ir. R. A. de By(chair)
Dr.Ir. P. R. van Oel

ABSTRACT

The spread of diseases is one of the major threats, to the growth and development of a nation. As diseases have a direct impact on the lives of the citizenry and the productivity of the nation in the long-term, diseases have always posed a great challenge to the global community especially infectious diseases.

Pertussis is an airborne infectious disease caused by a bacterial infection. The regular and frequent movement of adolescents and adults contribute to the resurgence of pertussis outbreaks. The means of movement depict different disease diffusion process. Hierarchical spread is a type of disease diffusion processes where the spread of a disease is from a city to its hinterland. The existing pertussis model developed by [1] depict the contagious spread where the disease spreads by direct contact.

The aim of this research is to adjust the existing pertussis model to include the commuting component in relation to the hierarchical disease process for a regional area (Twente region). To achieve this goal hierarchical spread is detected from the empirical pertussis data by adopting the approach used by by Broutin et al. [12]. Furness method is applied to the commuting data to obtain the origin and destination trips for work and education. Lastly, private means of transportation is integrated into existing pertussis model.

From the analysis, hierarchical spread is detected in the empirical pertussis data and the source areas for the spread of pertussis in Twente region are found. It was observed that some cities (like Almelo) that do not have a size corresponding to the critical community size (CCS) also functioned as a source area for the spread of pertussis in Twente region. Cross-commuting behavior in Twente region is detected in the commuting data. The results obtained from the adjusted pertussis model also depicted hierarchical spread from a village to a city.

Keywords

Agent-based simulation, hierarchical spread, pertussis, commuting.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

"Psalm 34: 8 Reads: Taste and see that the Lord is good; blessed is the man who takes refuge in him. "

First and foremost, my thanks goes to God Almighty for his sustenance and successful completion of the M.Sc program. I also appreciate the Netherlands Fellowship Program (NFP) for providing funds needed for the M.Sc.

My thanks also goes to my first and second supervisors, Ms. Ellen-Wien Augustijn (first supervisors) for her motherly advice, friendship, the encouraging words example "go for it you can do it ", the scientific recommendation and suggestion and above all her patience when I was confused. You are indeed a blessing to me, God richly bless you; Dr. R (Raul) Zurita-Milla (second supervisor) for his critical and scientific contribution to this research.

I would also like to appreciate Dr Mark. Zuidgeest from the Urban Planning Department in ITC for his assistance in acquiring the commuting data. Many thanks also to Dr. Kasper van Zuilekonn from the University of Twente for assisting in contacting Mr. Peter Kant from the Ominitrans international BV. I also appreciate Mr. Peter Kant from Ominitrans international BV who helped in retrieving the commuting data from the Ominitrans Model.

Many thanks to Mr Bas Retsios of ITC (Repasts Expert) who assisted me during the implementation stage. I would also like to appreciate Geological Survey of Ghana for the time given to persue the M.Sc study.

Undoubtedly, my sincere gratitude goes to my husband, Mr. Seth Okla and my daughter Acsah Nhyria Okla for their patient and unbroken link during my M.Sc program. I also appreciate all my family members for their support and encouragement.

To all my classmate and Ghanaians colleagues, I say thank you.

Chapter 1
Introduction

1.1 MOTIVATION AND PROBLEM STATEMENT

The spread of diseases is one of the major threats, to the growth and development of a nation. As diseases have a direct impact on the lives of the citizenry and the productivity of the nation in the long-term, diseases have always posed a great challenge to the global community. There are various types of diseases, but infectious diseases have contributed greatly to this threat (of illness or death), as infectious diseases can lead to a pandemic. According to the World Health Organization (WHO), 13.7 million people die every year in the world as a result of an infectious disease [84]. The urgent call for various studies into the patterns and spread of such diseases, with the aim to eradicate them, cannot be over emphasized.

Pertussis (also known as whooping cough) is an airborne infectious disease caused by a bacterial infection. The mode of spread is mainly by direct contact with discharge from the throat or nose of an infected person through coughing and sneezing [8]. As a result of the high transmittable nature of pertussis, the disease spreads easily and faster within urbanized areas due to the densely populated community and the increase in mobility [50, 33].

The vaccination propagation introduced in Europe, since 1950 [58] has controlled the disease for many years, yet there has been a resurgence of pertussis outbreaks that have affected many countries including developed countries with high-vaccination coverage. For example in 1996 there was an outbreak of pertussis in Netherlands and 18 per 100,000 populations cases were reported all over the country [22]. In 2000, 22-58 and 180 per 100,000 populations were reported in Australia and Switzerland respectively [76]. Even though various control measures such as vaccination and medication have been considered, pertussis have not been eradicated but subsided [83]. Literature has proved that, the resurgence of pertussis was due to waning immunity over time which occurs mostly among adolescents and adults. As shown by a number of studies, adolescents and adults are a major source of transmission of pertussis to infants as result of their frequent and regular movement [9, 41, 76].

According to Morse [52] mobility is one of the factors which contribute to the spread of infectious disease, which involves the movement of adolescents and adults to and from locations. The study of Collinge and Ray [19] also indicates that various factors contribute to the transmission of infectious diseases, among these factors are human alternation to ecosystem, mobility and urbanization. The need to study the transmission pattern in relation to mobility as a result of rapid global growth in urbanization which involves commuting has become very relevant in recent times.

Mobility is the ability to move from one place to another by means of air travel or by commuting. The means of movement depicts different disease diffusion processes thus causing mobility to be linked to disease diffusion processes. According to Cliff et al.[17] a disease diffusion process is the process by which an infectious disease is carried through space over a period of time. Air travel depicts relocation diffusion process which is the process where the source of the disease originates to a new location and it is a long distance spread. Commuting depicts an expansion process, where the origin of the disease remains and spread from one place to another. Expansion process occurs in two ways which are contagious and hierarchical spread [4, 16].

Hierarchical spread, is the spread through a sequential manner that is; from city to rural area or vice versa [17]. In this study, the focus is on commuting in relation to hierarchical spread which is a short distance spread. Hierarchical spread is known to occur in many infectious diseases such as influenza, measles, pertussis etc. [12, 80].

As simulation is a computer representation of reality, [34] simulating commuting with regard to hierarchical spread will provide a better understanding of the disease diffusion pattern and assist to relate the pattern to the severity of the disease in relation to the distance and the population size.

Agent-based simulation (ABS) is used to model the commuting behaviour of people which leads to hierarchical spread. Agent-based simulation is a technique which aids to conceptualize reality. It models the individual movement within space. ABS can also be used to simulate the social interaction among people [7, 4]. In disease diffusion an agent could be used to represent a person. These agents are goal oriented, can perceive their environment, communicate extensively, adapt and can move [14].

Several agent-based simulation models have been developed for infectious diseases but validation only focuses on the number of disease cases (epidemic curve) and not on the explicit spatial pattern [50, 33]. Some models are not geographically explicit and can therefore not lead to spatial disease diffusion patterns. Other models are geographically explicit but are for smaller areas which makes it impossible to visualize the emergences of the hierarchical spread.

The existing model presented by Abdulkareem [1] simulates explicitly the location where the social interaction of agents take place in relation to the spread of pertussis in Enschede region. The model does not include the commuting component of the agents and considers a smaller area. This study includes commuting of agents to the existing model and considers a larger area which enhance the emergence of hierarchical spread of pertussis in from a city to its hinterlands.

1.2 RESEARCH IDENTIFICATION

1.2.1 Research objectives

The main objective of this research is to adjust the existing pertussis model to include the commuting component in relation to the hierarchical disease process for a regional area.

1.2.2 Sub-objective

To achieve this, the following specific objectives would be addressed:

1. To analyze the spatial patterns in pertussis data to detect hierarchical spread.
2. To analyze the daily commuting data to gain understanding on the commuting behavior in Twente region.
3. To create a conceptual model to integrate the commuting behavior in Twente region into the existing pertussis model.
4. To test the model and to evaluate if hierarchical spread occurs.

1.2.3 Research questions

The questions that will guide the research and this study will include:

1. What is the commuting behavior in the Twente Region?
2. What impact does commuting have on the spread of pertussis in the Twente Region and over what distance does it cover?
3. How to detect hierarchical spread in the empirical data for historic disease outbreaks?
4. How does the existing model work?
5. How can commuting behavior be integrated in the current pertussis model?
6. Which data will be used to test the adjusted model?

1.2.4 Research Approach

The stated objective is achieve by following the procedures explained below(fig:research).

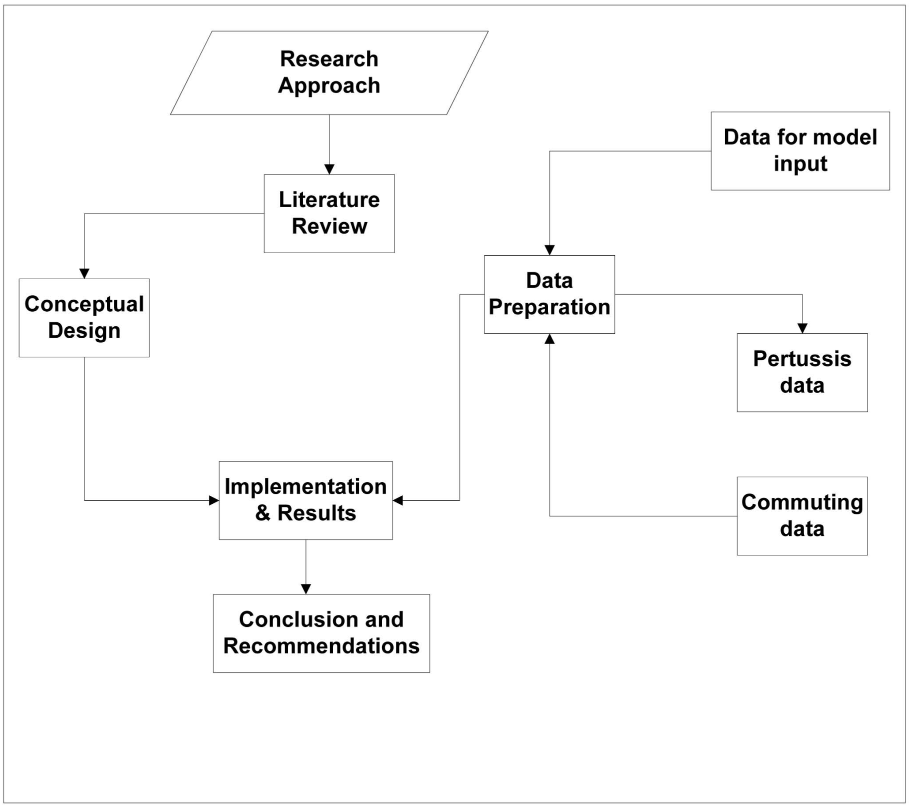


Figure 1.1: Research approach

Literature Study

Literature review was conducted to gain general knowledge regarding the concepts of agent-based simulation. Understanding commuting behavior in the Netherlands and that of Twente region helped in designing the conceptual model. Understanding the existing pertussis model and other epidemiological models has helped to answer some of the research questions in section 1.2.3. Knowledge gain in Repast Simphony has help to understand how the existing pertussis model works.

Conceptual Design

In designing the conceptual model to integrate commuting behavior in Twente region into the existing pertussis model, understanding the conceptual design of the existing model was gained.

Data Preparation

Pertussis and commuting data were explored. The pertussis data was used to identify hierarchical spread in pertussis while as the commuting data was used to detect attraction areas in Twente. Ways to proportional distribute agent over Twente region was acquired. This helped to reduce the actual population to the required model population.

Implementation

At this stage, the results acquired from the pertussis and commuting data were analyzed. In addition, the spatial extent of the existing model was upgraded. This was done by creating synthetic population based on census data acquired and loading spatial layers of the study area. The synthetic agent are distributed randomly in households within the neighborhoods of Twente region. The upgraded model was verified and results from the analyzed data and the upgraded model are presented.

Conclusion and Recommendation

Conclusions are presented and recommendations of further work are suggested.

Chapter 2 describes the modeling approach used to achieve the set out objective of the study as well as the concepts used to address the sub-objectives and theoretical answers regarding the research questions. Moreover related works in epidemiology.

Chapter 3 provides detail description about the conceptual design of existing pertussis model as well as the adjusted pertussis model.

Chapter 4 describes the data set required to detect the hierarchical spread and to detect attraction areas. In addition input data set of the model is addressed. The source of the data and the processes used in generating the required data are discussed in this chapter.

Chapters 5 discuss the implementation approach used to achieve the laid out objective and discuss the outcome of the implemented approach.

Chapter 6 conclusion based on the analysis of the results is discussed in this chapter and recommendation of further works is suggested.

Chapter 2
Background Information

This chapter addresses the relevant information as the bases to achieve the set out objectives. This chapter is divided as follows: section 2.1 describes the components of agent-based simulation modeling and the concepts in developing an agent-based model. Section 2.2 address the concepts of epidemiological model. Section 2.3 addresses the elements of disease diffusion and section 2.4 describes the commuting behavior in the Netherlands and Twente region. In section 2.6 present pertussis incidences in the Netherlands and other relevant issues regarding pertussis.

2.1 CONCEPT OF AGENT-BASED SIMULATION

Since the inception of modeling, various modeling techniques including traditional models using mathematical algorithms have been used to represent the abstraction of reality or systems. The mathematical approach models individual behaviors as a group, which makes it difficult to model the diverse nature of individuals due to the intractability of equation. Modeling the diverse nature of individual behavior helped scientists to understand the emergences phenomenon of reality or systems [34]. This has led to the computerization of modeling approach known as simulation.

Simulation promotes the discovery of complex systems and understands the formalization of these systems which aids in predication [34]. Various simulation models exist; among these are cellular automata and agent-based simulation.

Agent-based simulation (ABS) is one of the simulation options to model and interpret the emergence of complex phenomena of reality or a system. Phenomena or systems that occur due to mobile decision-making entities which are adaptive to spatial and temporal processes [77]. Phenomena or systems are modeled in ABS "as a collection of autonomous decision making entities called agents"[10] within a specific environment. The interaction behavior of individual agent in a dynamic system aids to generate the emergence behavior of a phenomenon or system [10, 14]. Castle et al [14] mentioned three main advantages of agent-based over the traditional models. These are: (i) ABS captures phenomena that grow; (ii) ability to study certain systems in a natural environment and (iii) the flexibility to work with, particularly in relation to the development of geospatial models.

- ABS captures phenomena that grow - the growing phenomena of reality are due to the individual interactions. The spread pattern of a disease, as an example is the interaction of the infected individual with other people who may be susceptible to contract the disease. The interactions of the individual enhance the spread pattern of the disease.

- ABS study certain systems in a natural environment - ABS have the ability to simulate individual behavior based on a set of rules and makes it possible to mimic natural systems. As an example ABS can simulate how and when commuters commute daily to work or school. It could be used to model activities that takes places within an organization [10].

- ABS is flexible to work with - ABS has the ability to tune the dynamic nature of individuals' behavior in relation to changes in reality by changing the characteristics of individuals in the simulation model. For example an agent changes his/her activity based on his/her daily activities.

2.1.1 Components of an agent-based simulation model

The components that make up an agent-based simulation model are as follows

- set of agents
- environment
- time
- the interactions between agents and/their environment

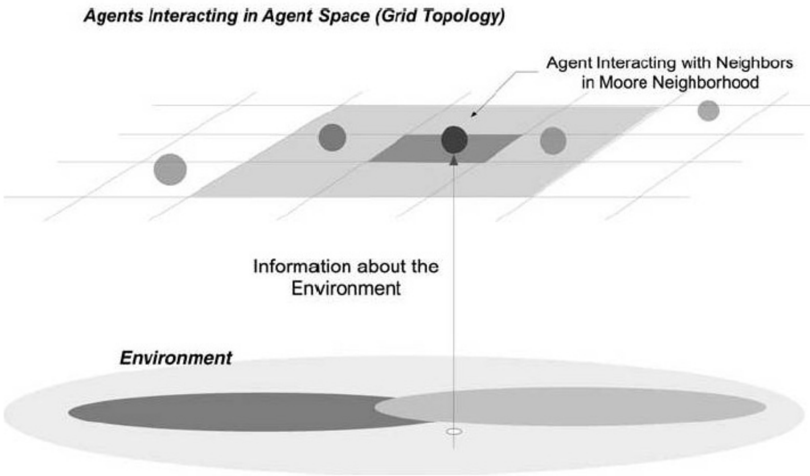


Figure 2.1: The component of a typical agent-based model, as in Sugarscape product by [48]

Agents

Agent has no formal definition but according to Franklin and Graesser [30] an agent is "a system situated within and a part of an environment, that senses the environment and acts on it over a period of time, in the pursuit of its own agenda and so as to affect what it senses in the future". From this definition, an agent has specific features which are described according to the following authors [3, 14, 26, 34]

1. Autonomy - agents are self governed; know what, when and how to act without the influence of a centralized process. Agents have the ability to process information and exchange information with other agents over a limited area of interest.

2. Goal-directed - agents have set of goals to be achieved.

3. Heterogeneity - diverse individuals or groups of agents can develop but spawn up from bottom-up.

4. Active - Agents are independent due to their goal oriented nature within a simulation environment, which makes agents active. The activeness of an agent can be classified into pro-active and reactive.

- Pro-active - the agents have sets of goals to achieve in relation to their behavior.
- Reactive agents have knowledge about, or are sensitive to their environment. Those agents have been programmed with a prior knowledge about their environment.

5. Bounded rationality - agents can have both fettered or unfettered information based on the obstacles in the environment.
6. Interactive/communicative - agents have protocol communication to interact with other agents and/their environment by using an expressive language.
7. Mobility - agents can have the ability to move within geographically or non - geographically environment. This makes it possible for ABS to mimic reality.
8. Adaptation/learning - agents can have the ability to adapt/learn to change based on their status.
9. Flexible - agents learn based on experiences. For agents to learn, they require memory to store their action which is based on sets of rules.
10. Attributes - agents have their own attributes and can also inherit other attributes from higher level agents. For example a child agent has the attribute such as age but inherit the family attribute such as family ID.

Environment

The environment is the spatial element of an agent-based simulation. The environment of ABS "may simply be used to provide information on the spatial location of agent relative to other agents or may provide a rich set of geographic information "[48]. Environment can be geographically explicit which makes the agent have a geographical location or spatially implicit which make the location of the agent not relevant in the simulation model [14]. Environment can be either static or dynamic. With static environment no changes occur when agents interact with the environment while as the dynamic environment changes occur even if the agents do not interact with the environment. Multiple environments can be modeled which may lead to interaction between environments which may influence the behavior of agents within those environments. In ABS the environment is the world where agents act [35]. Below are some of the characteristics of an environment.

- Location - the environment serves as a location for the agent which help to track the movement of agents in a model [48].
- constrain agents - the environment constrain the movement of agents. For instance in modeling the commuting of agents, agents can be restricted to commute through a specific path.
- Interaction - environment can interact with agent by changing when there is an interaction between the agents and the environment. For example the wolf sheep predation in Netlog [53]. As the sheep eat the grass (environment), the grass regrow.

Time

Agent-based simulation models are driven by time. Time can be continuous, discrete and discrete event. In continuous time changes occur continuously across time. Discrete time is split into regular time step while discrete event checks the process when change will occur. Time can be modeled in a synchronous manner where the timing of agents actions are done simultaneously [11]. For instance a time is set for agents in a model to get up at a specific time, at that specific time all agents in the model will get up. In ABS time is modeled as calibrating time. Calibrating time there is a link between real time and simulation time.

Interactions

The interactions of agents is the fundamental concept of an ABS [5]. Interaction in ABS is expressed by behavioral rules which consist of actions and conditions. The specific action (behavior) describes what should be done and when the specific condition actions should be fired. Interactions can be continuous or discrete. Interaction can occur among agent-agent, agent-environment and environment-environment.

2.1.2 Toolkits for agent-based simulation

ABS toolkits provide a conceptual framework for defining the agent, environment and the interaction. ABS toolkits serve as a template to design, visualize and acquire data for simulation/modeling a system [5, 60]. They toolkits provide modeler with available programming tools which aid modelers to focus on the content of the model rather than on the programming aspect [60]. The standard simulation functions assist in the reliability and efficiency of the simulation model. Even though ABS toolkits have numerous advantages, effort is required by the modeler to understand how to use and implement his/her simulation design in the toolkits [60]. Various ABS toolkits such as Repast, Netlog, MASON and others are available for simulation/modeling. These toolkits differ from each other based on the licensing policy either open source, shareware or proprietary -tutorial.

Repast

Recursive Porous Agent Simulation Toolkit (Repast) is an open source software which is based on an object-oriented programming language (Java), thus enabling Repast to be run on different computers platform systems [54, 29, 48]. Repast Organization for Architecture and Development is the organization that manages Repast. Repast *J.*, Repast.*Net* and Repast *py* are the different types of Repast which are implemented in java, Microsoft.net and python respectively. Repast Simphony (RepastS) is the current version of Repast which has the core functions of the older versions. RepastS has the point-and- click interface for certain development interfaces phases and can model both in the 2D and 3D [20].

Repast has the ability to specify the environment and behavior of agents by the use of context and projection. In Repast, the context is a container that host agents. It serves as an environment for the agent [44]. Context has an hierarchical features which allow sub-contexts to inherit features of the main context. Projection is the spatial environment for the agents. It specifics the locations for agents and gives the context a structure [44]. Projection enable agents to interact with each other, thus depicting the behavior of the agent.

Repast has an in built automated Monte Carlo simulation framework which can be used to generate synthetic population. Repast also provides tools to create and visualize agent's environments. The ability of the event schedule in Repast aids to mimic the temporal components in reality. Simulation run in batch-run and non-batch run in Repast, aids to see the effect of the outcome of the model by its initial condition [20]. Repast supports raster and vector data as well as exporting and importing to and from GIS formats. Repast does not provide a set of formulas to create the model. It is the task of the modeler to understand how to design and implement the model [20].

Repast is among the widely used toolkits for modeling and a scalable agent development environment. It has been implemented in various application fields. Yang and Atkinson [86] used repast to model the spread of an infectious disease among students in Southampton University.

2.1.3 Agent-based Simulation modeling protocol

Standard protocol is a lay down procedure which aids to describe a conceptual model of an agent-based simulation. Standard protocol makes it easy to understand and duplicate a simulation model. The standard protocol helps to identify the elements and the goal of the conceptual model. There is no standard protocol for agent-based modeling but various authors has proposal different standard protocol such as the Overview Design concepts and Details(ODD) and Model Driven Development(MDD).

Grimm et al [40] designed, Overview Design concepts and Details (ODD) protocol. The concept of ODD is to structure the information used in an agent-based simulation model in a sequencely manner. The ODD consists of three main groups (Overview, Design concept and Details) with seven elements as shown in figure 2.2 below.

Overview	Purpose
	State variables and scales
	Process overview and scheduling
Design concepts	Design concepts
Details	Initialization
	Input
	Submodels

Figure 2.2: The seven elements of the ODD protocol grouped into three main blocks - overview, design concept and detail [40]

The overview provides the overall purpose, structure and extent of the model. These describe the basic entities of the model that are used to describe a conceptual model. For instance in determining the spread of a disease, the overview in the ODD protocol will consider the purpose of the model and based on the purpose the attribute of the agent such age and health status can be identified. Then the extent of the model. The information gathered from the overview can be used to design the conceptual model.

Design concept "provide a common framework for designing and communicating agent-based simulation model"[40]. The design concept describes the underlying concept of the model. Including what emergence phenomena the model aims at and types of interactions that occur between agents. For example the emergence phenomena of a disease model can be to observed spread pattern in the disease. The design concept consider the processes to achieve the emergence phenomena like interaction of agents enhance disease to spread.

The detail consist of initialization, input and sub-models elements. Initialization describes the initial entity of the model where as the input data describes the data required for the implementation of the model. The sub-models give a detailed description of the processes that make up the model as well as the parameters used.

Model Driven Development (MDD) - is based on the Model Driven Approach (MDA) which uses a modeling programming language to develop a modeling system [13]. The concept of MDA is to use an independent platform to create a conceptual model of a system which generates codes based on a specific platform. MDA can be used for Repast because the codes are generated in Java programming language. INGENIAS [36]is an agent-oriented methodology based on MDD which is established in message/UML. It is an open source software that has the capability to design, implement and analysis systems as well as to automatically generate code by using INGENIAS Development Kit (IDK). The generation of sources code takes several iteration and steps. This is the difficult aspect of INGENIAS because a good knowledge and expert experience is require on how to implement agents in the targeted platform, and is time consuming [57].

2.1.4 Agent-based Conceptual Model

The Conceptual model is the basis for developing a model. The conceptual model aids to identify the goal, elements that make up the system such as the agents, the interactions, the behavior of agents, the environments and the measurable outcome of the system [14]. Conceptual models can be perceived as an imaginary or mental model which can be sketched with a pen and paper or by using a software such as unified modeling language (UML).

Bauer et al and Iba et al [6, 46] describe how UML can be used to generate a conceptual model. UML is a standard modeling language which is used for visualization, specifying the static structure of the model, the dynamic behavior, construct and organize the model by mapping UML to a programming environment which generating automatic code and document each phase of the system [46]. UML consist of a graphical notation which is used to describe the syntax and the meta data of the model. The following are the different types of UML diagram for agent-based model.

1. Conceptual model class diagram
2. Activity diagram
3. Communication - sequence diagram

Conceptual model class diagram

Describes the classes, the attributes of the class, the operation (method) and the association between classes. The class diagram is the static structure of a model which declares the element that makes up the model [6]. A class diagram in ABS is drawn by initially identifying the elements (agents, environment, and relation between agent and / environments) of the model. The identified elements are represented in as classes with their attributes and the association that occur between them. Figure 2.3 is an example of a class diagram.

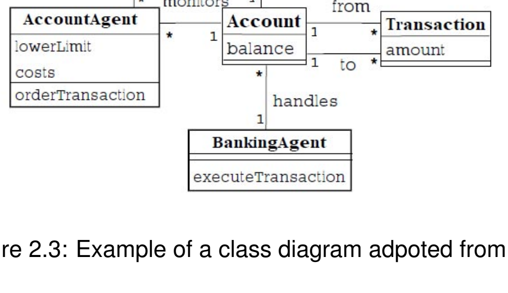


Figure 2.3: Example of a class diagram adpoted from [23]

Activity diagram

Describes the sequences of activities or tasks that occur within a model and are commonly called control flow which is similar to flow charts [46]. The activity diagram in ABS represent the procedures of an action performed by the agent which tells the behavior of the agents as shown in figure 2.4.

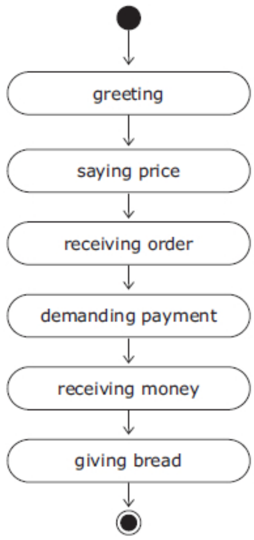


Figure 2.4: Examples of an activity diagram [46]

Communication or sequence diagram

Describe the interaction which emphasis the order of exchanging messages [6]. In ABS the communication or sequence diagram represent the interaction, links as well as the sequence of messages (how messages are sent and received) among agents, groups of agent and/or their environment. Figure 2.5 is an example of a Sequence diagram.

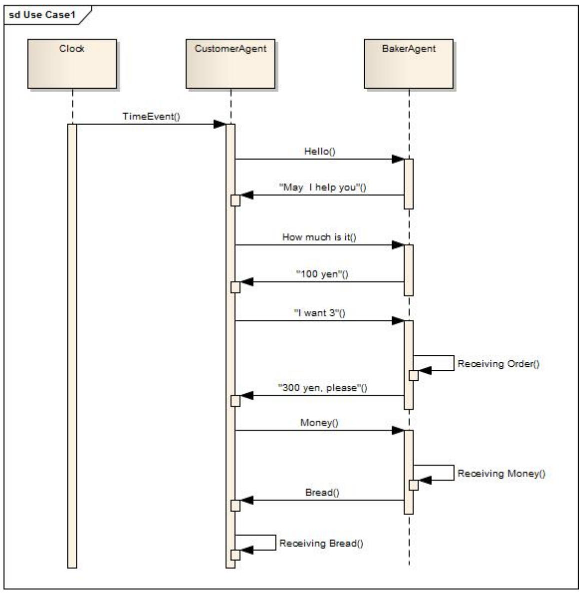


Figure 2.5: Examples of an sequence diagram [46]

2.1.5 Synthetic population

The unavailability of individual data has led to the evolution of synthetic population which is an alternative means of acquiring individual data [63]. Synthetic population is based on the concept of generating artificial set of individuals and households geographical with associated attribute data from an aggregated census dataset by using some algorithm [32, 63]. The aggregated data servers as the input data while the output data is the set of individuals and households population.

There are various techniques used to develop a spatial synthetic population such as iterative proportional fitting, data fusion, synthetic reconstruction and combinatiorial optimization among others. According to Huang and Williamson [45], synthetic reconstruction and combinatiorial optimization approaches are the main techniques for creating synthetic population.

The synthetic reconstruction approach will be the used to generate the synthetic population for Twente region, since the existing pertussis model used this approach to generate the synthetic population for Enschede city. This will aid to gain the same quality, since different techniques differ in quality.

The synthetic reconstruction approach is a sequential approach which uses the Monte Carlo sampling to create a synthetic population based on a set of conditional probabilities. Monte Carlo sampling is used to generate pseudo-random numbers based on relevant conditional probabilities to assign attributes to these generate numbers [63]. The relevant conditional probabilities are obtained by the iterative proportion fitting (IPF) which is a technique that is used to overcome the interdependencies between variables that a are not obtained from the published census data [63]. The conditional probability enhance the possibility to incorporate data from different sources. In generating synthetic population the household heads are the first population to be created with age, sex, marital status and their spatial location as their characteristics. These characteristics are considered because they are the fundamental characteristics of a household structure and are important factors in generating other variables [63]. Below are the sequential procedures used to generate a synthetic population. Each step is to add new attributes by creating conditional probabilities for a given census data [63].

- Step 1: Generating an initial population - the household heads are generated with these characteristics age, sex, marital status and their spatial locations. This generates adult population which is categorized into household heads and non-household heads.

- Step2: Disaggregation of initial population - the initial generated population is disaggregated into sub-groups. For instance the household head with marital status will be split into sub-groups such as single, married, widowed or divorced. The disaggregation is achieved by using Monte Carlo sampling and the conditional probability.
- Step 3: Generating economic status

The generated agents population are randomly assigned to a spatial location within their dwelling.

2.2 EPIDEMIOLOGICAL MODELS

Disease in human beings has been of great concern to scientists, as life is not replaceable. As disease does not occur randomly, the causative factors are the required element which will aid in obtaining the preventive measures. Investigations conducted among different individuals or population has led to the revelation of the causative agents of diseases which has led to the determination of distribution pattern of a disease and the frequency of it occurrences. Investigating the distribution pattern leads to determine the source of infectious and the susceptible class, how and where the disease spreads. The frequency of the occurrence of a disease takes into account when and the number of diseases cases that occur. This has led to the study of occurrence and the distribution pattern of disease in human population which is epidemiology [42]. The great impact of epidemiological models is to improve the health policy decision among a population. Various epidemiological models has been modeled either by the classic or complex approaches. The classic modeling approaches, for example differential equation, models from top to bottom while the complex approaches, for example agent-based model is from bottom-up [10].

The classic approaches are based on mathematical concept which describes the emergent phenomena of diseases with the assumptions that, individuals are a homogeneous population which interact with other individuals at the same rate but do not move within a non-spatial environment [7]. The classic approaches model individuals into groups based on their infectious state. Various infectious state models have been developed, these includes susceptible and infectious model (SI) or susceptible, exposure, infectious and recovery model (SEIR) among others. Based on the stated assumption the classic approach is unable to explicitly model the causative factors of the spread of disease but rather estimate the affected population [7, 78].

The complex approach is the alternative modeling approach. The complex approaches take into account the heterogeneous nature observed in the infection of a disease, the spatial and the temporal nature at which disease spread. Some complex approach models do not explicitly model the mobility and the interaction of agents, an example is the cellular automata. While agent-based simulation explicitly model both the social interaction network and transportation network of agents.

Some examples of agent-based epidemiological models for infectious diseases

Agent-based epidemiological models have been used to simulate various infectious diseases such as influenza, smallpox and pertussis etc. [15, 33, 50].

Viboud et al [80] simulated the spread of influenza for 49 states in USA at a wide spatial wide spatial scale by relating the human movement with the spread of disease. They considered the population size, distance and severity of the disease taken into account the movement of people.

Gravity transportation model and standard stochastic susceptible-infectious model were used for the analysis. They observed the spread of disease in cities with less population is slow as compared to more populated cities. The long distance spread of influenza was attributed to the work flow of adults and also children were attributed to the local spread of the disease within schools and households etc.

Yang, et al [88] simulated the spread of influenza within a population in a city by using individual space-time activity-based model (ISTAM). ISTAM takes into account the individual activities among humans activities bundles which humans' space time dynamics are modeled as changes in activity bundles per individual during one day within a certain area. And by within activities bundles which captures each individual's space time dynamics at fine space time scale which makes ISTAM flexible. The ISTAM model was applied to a city in Netherlands, Eemnes. The result shows that the higher the number of cluster the higher the contact between individuals.

Abdulkareem [1] simulated the spread of pertussis at a city scale which depicts how pertussis spread based on the social interaction of agents. The social model and activities pattern shows where individuals interact and mingling with other individuals which server as a means to transmit the disease. Some of the limitations in the existing pertussis model are agents do not commute. Only families with children are the population use in simulation the existing pertussis model. Also the existing pertussis model can run for a specific number of agents on a single computer. Even though the exist pertussis model has some limitations, this model explicitly simulate the spread of disease base on social interaction.

The studying of these models which simulate the social interaction and the movement of agents for large population size would help to integrate commuting into the existing pertussis model and will help in scaling-up the existing model.

2.3 DISEASE DIFFUSION

Disease diffusion is the process by which disease spread. Diffusion depends on disease occurrences which is "the place where cases are found (i.e. an infected individuals) and the locations where the necessary circumstances for causing illness are prevalent "[64]. Disease diffusion occurs within space and time, it has a temporal and a spatial component.

The temporal component of disease diffusion considers the season, time and duration of the diffusion process. The spatial component of diffusion considers the dispersal and the structure dimension of disease diffusion. The dispersal aspect considers "the routes of commuting "while the structure aspect "refer to the reciprocal relationships between the location along the route"[64].The dispersal and structural component describe the different disease diffusion process.

Disease diffusion processes

The disease diffusion process, as defined in chapter 1 sectionsec:motivation, is the process by which an infectious disease is carried through space over time [17]. The dispersal component describes either contagious or hierarchical diffusions and structural component describe either expansion or relocation processes [64]. The expansion and relocation processes are defined in chapter one. The contagious spread is by direct contact which makes the diffusion process slow while hierarchical diffusion is described in detail in 2.3 since it is the spread pattern considered in this study.

Hierarchical spread

Disease diffusion with spread in sequential manner, such as from a city to its hinterland, is known as hierarchical spread (see figure 2.6). With this type of spread, the source of the disease remains and intensifies within the originated area but spreads out to a new location at a later period and fades away faster. Distance and movement are the factors which influence hierarchical spread [16]. According to Cliff et al [16] and Viboud et al [80] the rate of spread is faster if there is good road network and distance between cities and their hinterlands is small.

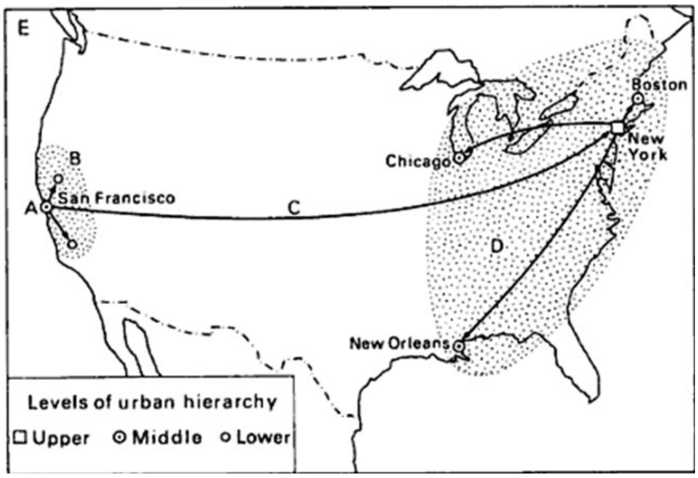


Figure 2.6: The hierarchical spread from town A which spreads down to village B, produce by [17]

2.3.1 Elements of disease diffusion

The element that enhance disease diffusion [64] are

- The original source of the disease.
- The place or places the disease migrates to.
- The pathways of spread and
- The driving force behind the diffusion.

These elements are the underlying elements that contribute to the factors of disease diffusion. The determining factors of disease diffusion [64] are

- The causative agent mode of spread
- Social interaction between infected and susceptible individuals
- Mobility

Social interaction and transportation (mobility) are the origin of disease diffusion which was observed by John Snow, by mapping the relation between the public water-supply system and cholera cases in London [69]. Thus making social interaction and transportation the key factors in disease diffusion [18]

Social interaction

Social interaction is also referred to as the social network, and describes the links or relationship among a group of individuals with a common task [47]. Social interaction is based on individual activities which different based on the purpose, population size, time and space, which makes social interaction to differ. Mao and Bian [50] observed high transmission of influenza in a highly populated area compared to less populated areas due to a lot of social interaction and movement within a highly populated area. This implies that disease diffusion in relation to social interaction is dependent on the population size. This makes critical community size important in the understanding of disease diffusion. The concept of critical community size (CCS) is to "define a population threshold below which disease cannot persist without external input"[12, 38]. The threshold found in studies CCS is 250,000 population [12]. Combining Enschede and Hengelo will sum up to the required threshold for the CCS.

Commuting

Rapid population growth has led to the expansion of cities leading to urbanization and an increase in the distance in commuting. This has caused rapid growth of the transportation network leading to an increase in the establishment of businesses, industries and educational institutions among others within urban cities. As more attention is drawn to economic infrastructure, less housing facilities are available in the urban areas causing a lot of people to commute daily to work or to school [24]. The regular movements between specific locations such as home, work or school by means of transportation is called commuting.

Commuting is characterized by outgoing and incoming commuting activities. The outgoing-commuting are people commuting out of their locality and incoming commuting are people commute into a locality. Time, distance and purpose are the factors that affect the commuting behavior of people [66]. These factors aid to determine the means and the types of commuting.

There are various means of commuting; by private car, public transport, by bicycle, by walking or by boat. Commuting occurs within space and time [59]. The means of commuting varies within various locations, due to the different transportation infrastructure within a locality based on the topology of the area. The means of commuting also determines the types of commuting that occur within a specific area. Some types of commuting are "regular", "random"commuting and "door to door"commuting among others. Regular commuting occurs where the same travel pattern is repeated each day between fixed pair of location [21]. For example commuting from home to work or vice versa. With random or irregular commuting travel pattern occur where the travel pattern is not repeated to a specific location [21]. In this study regular commuting is consider.

2.4 COMMUTING BEHAVIOR IN THE NETHERLANDS

The commuting behavior of a specific group of people depends on the individual activity pattern [66]. Due to the small geographical space of the Netherlands outgoing and incoming commuting behavior within urban areas are feasible. The country has been divided into various provinces, regions and municipalities. Some areas are more urbanized compared to others. Central, decentral, cross and exchange-commuting are the four types of daily urban system as distinguished by Van [79] based on how commuters commute from the urban cities towards the suburban area(outgoing-commuting) and from the suburban areas towards the urban areas(incoming). Below is the description of the four types of daily urban system which is depicted in figure 2.7 by Schwanen et al [65].

1. Central commuting - Many commuters commute mainly towards the urban cities, that is the urban cities attracts more commuters.

2. Decentral commuting - commuters from both the urban cities and other suburbs are oriented towards a particular suburbs areas.

3. Cross-commuting - many commuters in the urban cities commute within the urban areas where as commuters from the suburbs commute toward other suburbs.

4. Exchange - commuting many inhabitants who live in the urban cities commute to work in the suburb while the suburb inhabitants commute to work in the urban cities.

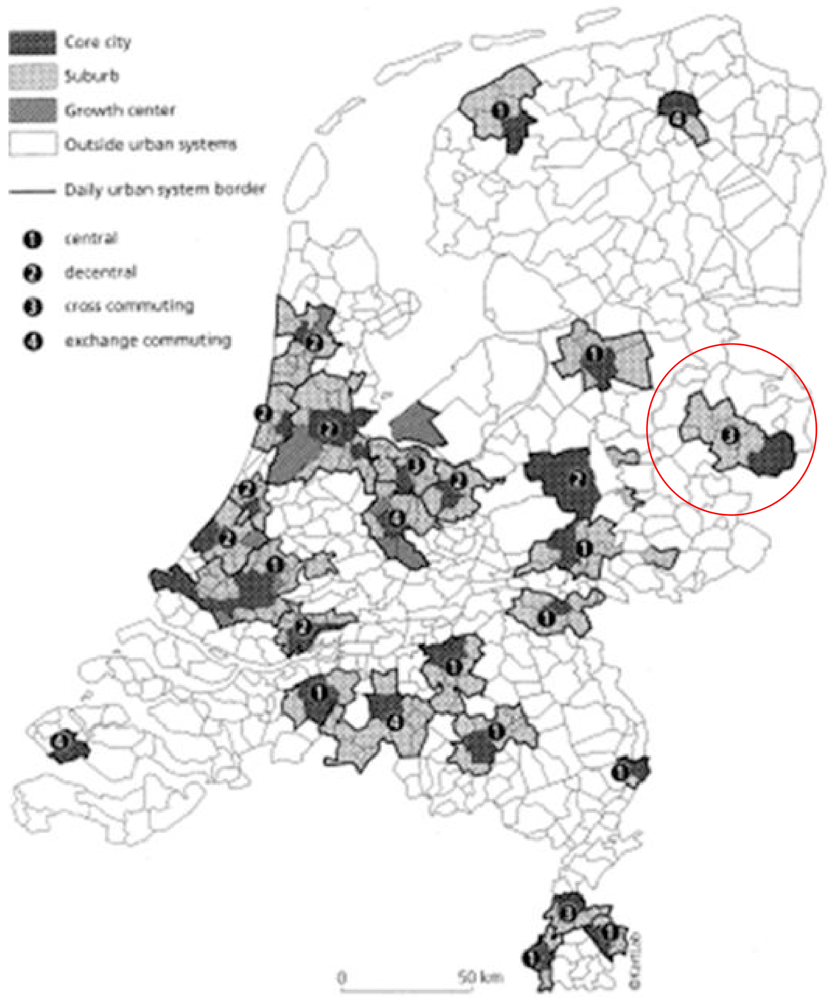


Figure 2.7: Map of the Netherlands shows the daily urban systems in Twente region (the area within the red circle) : produced by Schwanen et al [65](based on Van der laan, 1998)

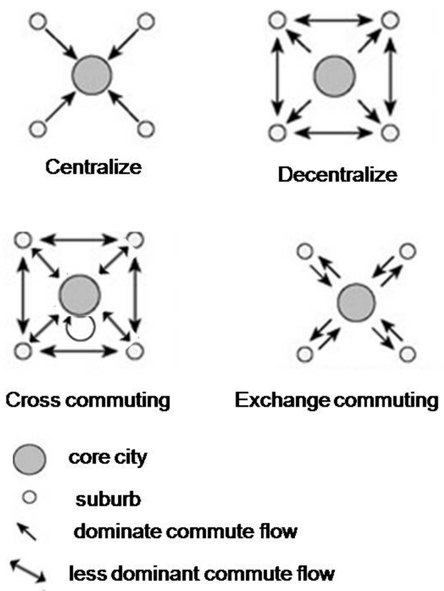


Figure 2.8: Types of commuting behavior : produced by Schwanen et al [65](based on Van der laan, 1998)

There are high numbers of outgoing commuters in the lesser urbanized areas and half of the population in the more urbanized areas are outgoing commuters. Also the highly skilled inhabitants are outgoing commuters which implies that there are more outgoing commuters than incoming commuters in Netherlands [75]. The Netherlands is made up of about 16.3 million inhabitants out of which, 5 million people commute daily to work by various means of transport such as private car, bicycle and public transport among others. Over 60% of the Dutch inhabitants commute by private cars, 25% by bicycle, nearly 10% by public transport and the remaining percentage by walking [71]. The most common means of commuting for less than 3km are by walking and cycling.

The commuting distance within the more urbanized areas are shorter compared to less urbanized areas but more time is spent commuting within more urbanized areas. Short commuting occurs within densely populated areas and vice versa. The commuting time within the more urbanized areas has increased from 2.34 to 2.63 minutes since 1995 to 2005. The departure time for Dutch full-time employees to commute ranges from 7.35 to 8.10 am [75].

The distance and the time of commuting determine the types of commuters that commute. Males are longer distance commuter and are outgoing commuters compared to female commuters. Young workers between the ages of 25 and 34 commute longer distances and are more often outgoing commuters. The older aged workers commute shorter distances because they prefer to work within their locality. Higher income and highly educated workers are long distance commuters. The highly educated commuters prefer other means of commuting than private cars [70].

The Euregio Mass-Rijn which is the cooperation between bordering area in Netherlands, Germany and Belgium has led to a cross-border commuting [27]. In 2009, 45,000 German residents and 45,000 Belgium residents commute daily to work in Netherlands. Most of the cross-border commuters are Dutch nationals who currently reside either in Belgium or Germany [72].

The daily activity pattern of a Dutch worker was depicted by Schwanen [67]. Five periods were defined as shown in figure 2.9. Within these periods two levels of activity patterns occurred. The first level of activities represent bundles of activities that occur in the home, at the work site and the time window which is "the block of time that start when a person leaves either his/her home or work place and arrive at the same or another base location"[67]. The second level differentiate other activities and the travel period within the time window.

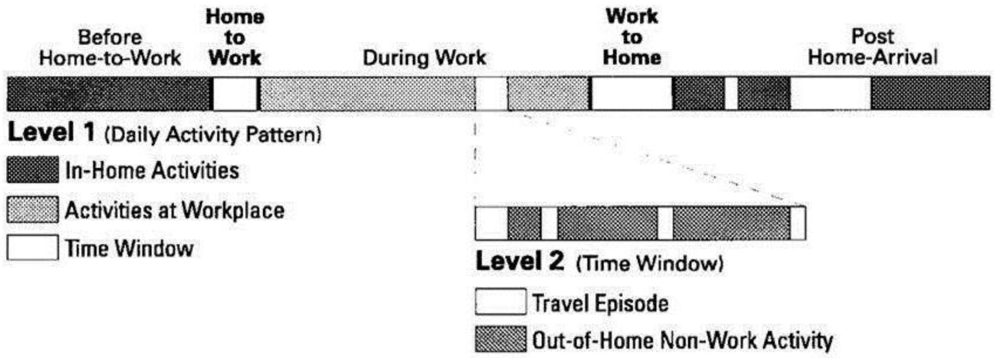


Figure 2.9: Conceptualization of an individual's activity pattern. By Schwanen and Dijst[67]

The individual activity pattern such as activities at home or at work determines the time a person leaves or arrives at home. From Schwann's analysis 70% of Dutch population, are full-time workers, who spend on the average 8.5 hours at work and 30% are part time workers.

Commuting by public transportation

Public transportation is a shared passenger transport service which is available to the general public [82]. There are various means of public transportation such train, bus and trams among other. Public transport is the faster's means of commuting within urbanized areas [75] and also the faster means to outer suburbs and neighboring towns and villages. The low prices and the faster means of public transport serve as the major means of transportation for the general public especial during for peak hours.

According to Schwanen et al [65] more women commute by public transport compared to men, due to the short distance travel by public transport and women prefer short distance commuting because of their responsibility to keep their homes.

2.4.1 Transfer of disease during commuting

Infectious diseases are transferred during commuting due to large number and close proximity of passengers, overcrowding, duration of trip and ventilation system within the public transport [31, 56, 4]. Public transport is the most likely mean by which infectious diseases spread rapidly during commuting [21] due to large numbers and close proximity of passengers. Balcan et al, [4] Mangili et al [49] and Moore et al [51] have proved public transport to be the means by which a passenger contracts infectious diseases whiles traveling. The large numbers of passengers commuting during peak hours leads to congestion and cause poor ventilation within public transports. The large number of passengers during commuting creates a social interaction environment where people can interact. Interaction occurs in various ways; by talking or by direct contract among others, between two or more people. Interaction enhances the spread of infectious diseases because infectious diseases are transmitted by various ways of interaction between people. The dissemination of tiny suspended droplets that spread over a short distance from an infected person, suffering from airborne infectious disease such as pertussis while talking, sneezing or coughing during commuting in a congested and poorly ventilated public transport unit leads to a prolonged suspended droplets exposing of passengers that are susceptible. The duration of commuting thus causes passengers to be more susceptible. Susceptible passengers can be infected by an infectious passenger.

Relationship between commuting and disease spread

Commuting is one of the factors which enhance disease diffusion to occur. Commuting has a link to disease diffusion because it enhance disease to spread rapidly to far and near places [17]. It has been proven by Viboud et al.[80] that long distance spread of influenza was attributed to the commuting behavior of adults where as children attribute to the local spread of the disease within schools and households etc which depict different disease diffusion process. Also the regular and random commuting types of commuters facility the spatial transmission of infectious diseases [21]. “Disease diffusion can only occur between individuals that live at different location due to the discrete movement of the individuals”[21]. This implies that disease diffusion process is dependent on commuting (mobility).

2.5 TWENTE REGION

2.5.1 Study area

The Twente region is located in the eastern part of the Netherlands which is part of the Overijssel province and a non-administrative region [68]. It shares a boundary with Germany and is bordered by the river Regge and Dinkel. Twente region is located within the Geographic coordinates of 52°19’59”and 6°45’0”. Twente region covers 1430 sq km, which consist of 14 municipalities and 119 postal code with a population of about 620,000 which makes up 4% of the Dutch population. The 14 municipalities within the region are Twenterand, Tubbergen, Dinkelland, Hellendoorn, Wierden, Almelo, Borne, Hengelo, Enschede, Oldenzenaal, Losser, Haaksbergen, Hof van Twente and Rijssen-Holten see figure 2.10 of which Enschede is the heart of the Twente region [61].

Enschede, Hengelo and Almelo are the cities within the Twente region with large companies, medical care and industrial estates. Hengelo has a train station which has a direct connection to European rail net (Netherlands: Thalys, Germany: ICE and France: TGV) [68, 61]. The Twente region has good road network with short distances between the cities and its hinterlands.

2.5.2 Commuting behavior in Twente region

Twente region is one of the provinces in the Euregio Maas-Rijn border to Germany. The commuting behavior within the Twente region is cross-commuting where most of the inhabitants within the cities commute within the cities and commuters from the suburb commuter regularly commute towards other suburbs [65].

Figure 2.10: Map of Netherlands on the left,at the upright corner the Overijssel province and the lower right the Twente region with it municipalities.

Commuting behavior in Enschede

Enschede is the core and the biggest city in the Twente region. It is an urban city with various facilities such as the largest goods market in the eastern part of the Netherlands [61]. Enschede does not attract high numbers of incoming commuters and have low outgoing commuters. Only one third of incoming commuters commute into Enschede to work and over one quarter of the inhabitants of Enschede commute out to work [70]. This proves that the daily urban system in Enschede is cross-commuting as shown in figure 2.7 the picture above. The effect of commuting behavior in Enschede in relation to the spread of disease is that the disease remains in Enschede due the large number of people commuting within the city. But the disease later spreads out to its hinterlands due to the small number of incoming and outgoing commuters. Thus the commuting behavior in Enschede depict hierarchical spread.

Understanding the commuting behavior in Netherlands, Twente and Enschede will aid in conceptual design of the adjusted model.

2.6 OVERVIEW OF PERTUSSIS

Pertussis is a respiratory tract bacteria disease caused by Bordetella Pertussis which lives in the throat, mouth and nose of the infected person, making pertussis to be easily transmittable during social interaction. Pertussis is one of the dreadful diseases that occur mostly in infants. It is characterized by prolonged, repeated paroxysmal which last for almost 6-8 weeks with incubation period from 6-20 days. The development stages of pertussis are catarrhal, paroxysmal and convalescent. The catarrhal or cold like stage is the initial stage of the disease which is characterized by runny nose, symptoms of cold, mild coughing and slight fever and sneezing. The catarrhal stage is the most infectious stage and lasts for almost 1-2 weeks which is followed by the paroxysmal or coughing stage. During the paroxysmal stage, severe, prolonged, harsh, dry coughing occur which ends with a whooping sound causing tearing eye and difficulties in breathing. Vomiting or choking occurs due to the coughing bouts. This stage last for 1-4 weeks or more. The convalescence stage is the recovery stage which lasts for several weeks. During this stage mild or less coughing occur and the patient is not infectious [85]. Children are vulnerable, making them more susceptible to pertussis which has led to a high mortally rate [8]. The waning immunity among adolescents and adults, the mild symptoms exhibit leading to late diagnose and treatment and the frequent commuting and interaction behavior causes adolescents and adults to be the major source of transmitting pertussis [41].

2.6.1 Pertussis incidence in the Netherlands

The incidences of pertussis in the Netherlands have occurred for several decades but the introduction of vaccination in 1950 in Europe [58] has caused a drastic decrease in the number of reported cases. The notification, serology data, hospital admission and registration of death are the surveillance source on which the incidences of pertussis were based [37]. From 1993/1994 there was an increase in the number of reported cases even though there was high vaccination coverage. In 1996 there was a rapid increase in the reported cases and the waning immunity in older children and adults was considered as one of the factors that led to the increase in reported cases. There was a decline in reported cases in 1997-1998 but in 1999 reported cases increased. The increase and decline of reported cases shows that the increase in pertussis incidence in Netherlands, since 1995 - 2004 occurs every 2 to 3 years [37]. The exact factors that contributed to the increased number of cases were not identified as at that time [22]. According to Tan et .al [76] the waning immunity among adolescent and adults was one of the factors. The seasonal peak of pertussis in Netherlands was end of summer. The proportion of male incidence cases was lower compared to female [37].

2.6.2 Diffusion pattern in pertussis

Pertussis among others depict a hierarchical spread pattern. Broutin et.al [12] and Grenfell et al. [39] found out that the first arrival cases of pertussis starts from the urban areas and then spreads out to the hinterlands. Population size and density as well as geographic distance were the parameters used to analyze the hierarchical spread observed in the empirical data that occurs during pertussis in the endemic and epidemic years. The results depicted that large population size and density showed large cluster and longer stay of the disease which is due to lots of social interaction and commuting within large population size. Whereas less social interaction and commuting occur within small population size causes the disease to fadeout faster. Understanding this concepts will help to detect the hierarchical spread in the empirical pertussis data.

Summary

In this chapter the general concept about agent-based simulation, disease diffusion, commuting and pertussis were explained. Major emphasis was on commuting, the types of commuting, it relationship to disease diffusion and the commuting behavior in Netherlands especially the Twente region.

Chapter 3

Conceptual Design

This chapter describes the conceptual design of the adjusted pertussis model. Understanding of the conceptual model of the existing pertussis model helps to integrate commuting behavior into the existing model, thus section 3.1 of this chapter will describe the conceptual model of the existing pertussis model. Section 3.2 describes the conceptual design of the adjusted pertussis model.

3.1 THE EXISTING PERTUSSIS MODEL

The existing pertussis model by Abdulkareem [1] was based on the Individual Space-Time Activity-based Model (ISTAM) which is an agent-based disease model developed by Yang and Atkinson [87]. The model explicitly simulates the social interaction of agents that causes the spread of pertussis in Enschede a city in the Netherlands.

The existing pertussis model generate synthetic population by using the synthetic reconstruction approach described in chapter 2 section 2.1.5. The synthetic reconstruction approach uses the Enschede census data with families with children to create and assign agents. Household heads are first created, then the next family member.For example creating couples with children the next agent to be created is the partner. Children are then created. As agents are created their attributes are assigned. Individuals/agents differ from each other based on some of their attributes. The attribute assist to investigate how individual contract and spread the disease. The attributes of the agent is shown in the class diagram in figure 3.1

The existing pertussis model consist of three main model - the daily activity, social interaction and disease models. Below are the description of these models.

The Daily activity model

The daily activity model, modeled both the individual and group activities. Each individual had an activity to perform within a specific time of the simulation.An example of an agent attending an activity is shown in figure 3.2. Agents with similar activities formed a group. A group is defined according to Ellegård [25] as frequent meetings in a time-space due to the social obligation of the members. Groups are formed based on various purposes. A family group are formed based on the relationship that exists among individuals. Group activities are formed when two or more individuals shared the same activities.

The individual has an activity schedule which is used to determine what activity an individual should perform. The activity schedule assists the individual/agent to change his/her activity after completing another activity. Changing activity causes the individual to be involved in another group which led to between activities interaction [1] making individuals linked to two groups.

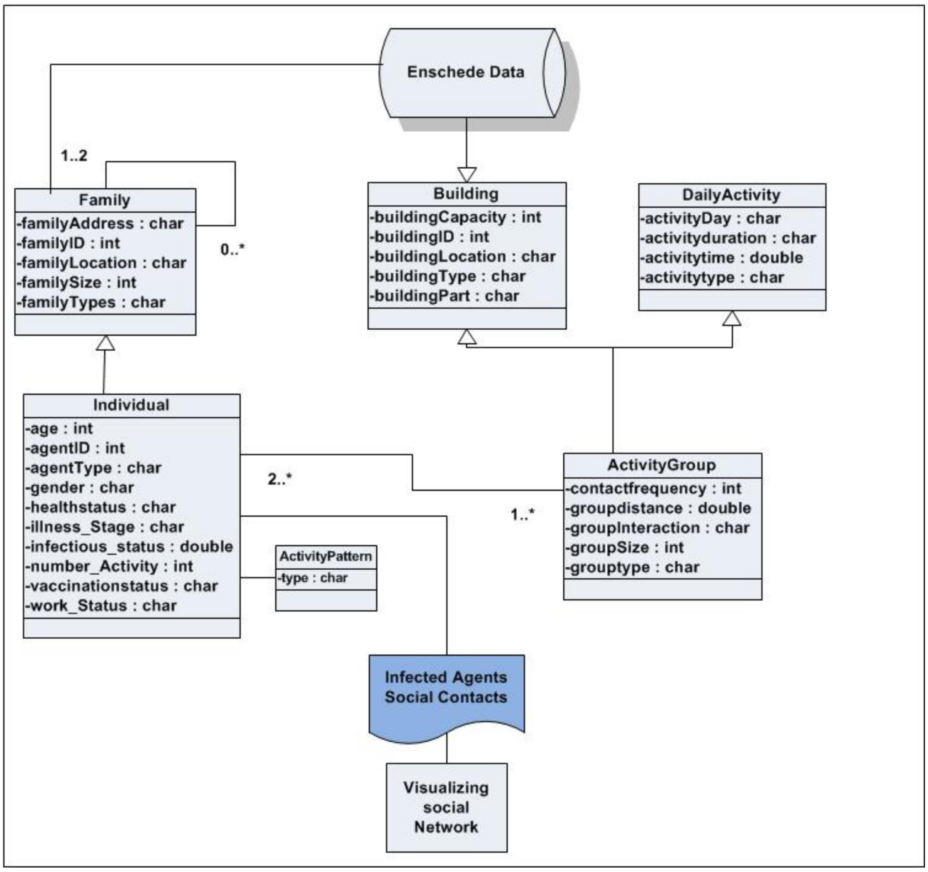


Figure 3.1: The class diagram of the existing pertussis model [1]

Two types of activities are considered static and dynamic. Static activities are the weekly activities such as going to school or work whiles dynamic activities are social activities which vary from day to day like visiting a friend. Figure 3.2 is an example of an individual attending an activity. The activities of various types of individuals/agents is design in a CSV (comma separated value) format which the model read.

Social interaction model

The social interaction model, models the full and partial interactions that occur within groups or between groups, since social interaction is one of the key factors to disease diffusion. Full interaction occurs when all members within a group interact (see fig. 3.3a). For example within a family group, parents interact with all members in the family. The partial interaction occur when some of the members interact within group (see fig.3.3b) for instance a staff agent can only interact with staff in the same department or at the office but not all staff in the institution.

As activities are constraint by both physical and social constraints, basic rules for social interaction such as the size of a group, the distribution of individuals within a group and the distance between individuals are considered. The distance within which an individual will interact was used to determine if an agent is susceptible or not. During group activities social interactions occur which enhance the disease diffusion.

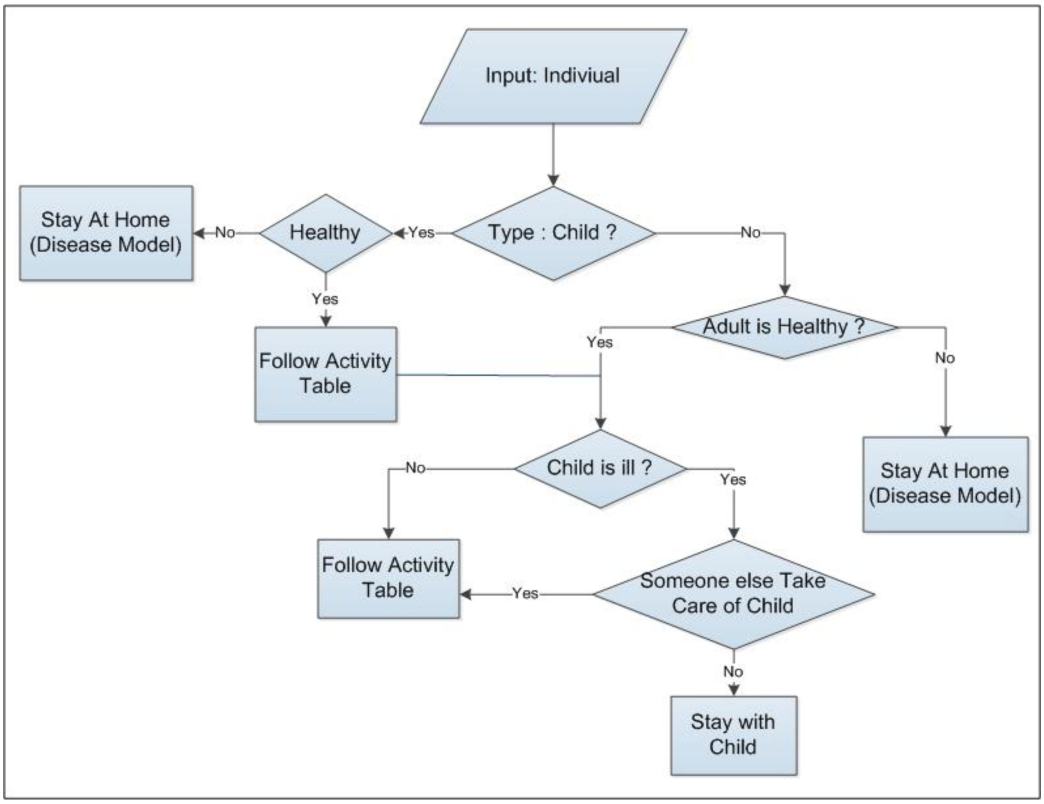


Figure 3.2: Individuals attending an activity [1]

Disease model

The disease model comprises of an infection model and the illness evolution. The infection model evaluates the health status of individuals during interaction. If there are infectious and susceptible individuals in a group activity, disease diffusion occur. For disease diffusion to occur some constraints have to be met, which are;

(a) The distance between individuals during interaction, should be less than or equal to one meter.

(b) A close contact interaction

(c) Within a group there should be an infectious and susceptible individual.

The disease model calculates the contact duration or the frequency of contact per day. For disease to be transferred, constraint are met, which are

- the calculated contact duration should be more than or equal to an hour or
- the calculated frequency of contact per day between two individuals should be more than or equal to one hour for daily activity.

Illness evolution starts when an individual is infected. It models the characteristic stages of pertussis (see 3.4) which are the incubation period (from 6-20 days), catarrhal period (1-2 weeks), paroxysmal period (1-6 weeks) and convalescent period (4 weeks).

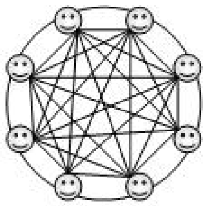


Figure 3.3: shows the full and partial interaction - individuals are represented by the faces, the group boundary is the circle around the individual and the straight line between two individuals represent the contact that occur(Adopted from citeAbdulkareem2010)

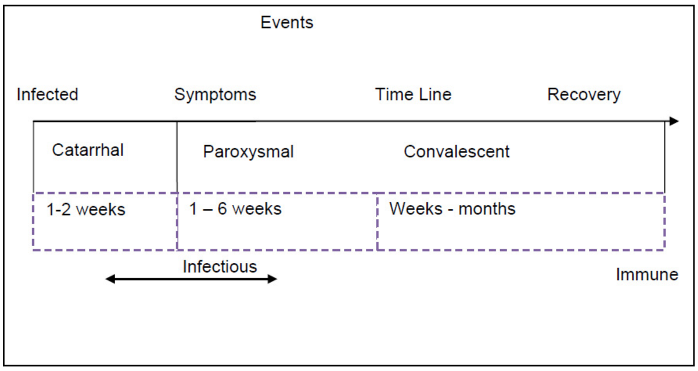


Figure 3.4: Pertussis Illness stage [1]

The infected individual follows these stages and in the infectious period which is "one week of catarrhal period started and continues to 3 weeks after that"[1] the individual is infectious. The health status of the infected individual changes when s/he is in the paroxysmal period and stops every activity to stay home.

An initial infected case is selected at every simulation run randomly, out of the 71 neighborhoods in Enschede, only 36 neighborhoods are simulated. The simulation run for one month with a time interval of 30 minutes.

3.2 ADJUSTMENT TO THE EXISTING PERTUSSIS MODEL

This section describe the conceptual design of the commuting component into the existing pertussis model. The general aim of the conceptual model is to include the commuting behavior among workers and students in Twente region which contributes to the spread of pertussis. Workers or student who commute from Twente region to other area out side the region or from neighboring towns not within Twente region to work or to school, are not consider in the model.

There are various means of transportation in Twente region either by private or public transportation. The private transportation are by car, bike or by walking among others whiles the public transportation are by bus or by train etc. The adjusted model distinguish commuting into two groups by public transport and by private transport.

3.2.1 Commuting Model

Commuting status, commuting types and commuting behavior are added to the individual class in the class diagram of the existing model (see the class diagram for the adjusted model 3.5). The first step is to set a commuting population. The commuting status is use to set the commuting population based on a constraint.

Agents are classifies into various types based on the working status into full worker, part-time worker, jobless and students. Also agents are classified based on age group as infant, child, teenager, adult in the existing model. These classification are use to set a constraint for the commuting status. The constraints is individual/agents whose work status is either a full time worker or part time worker or student with agent type as teenager is classified as a commuter if s/he satisfies this condition. The total number of agents who satisfies the stated condition are calculated.

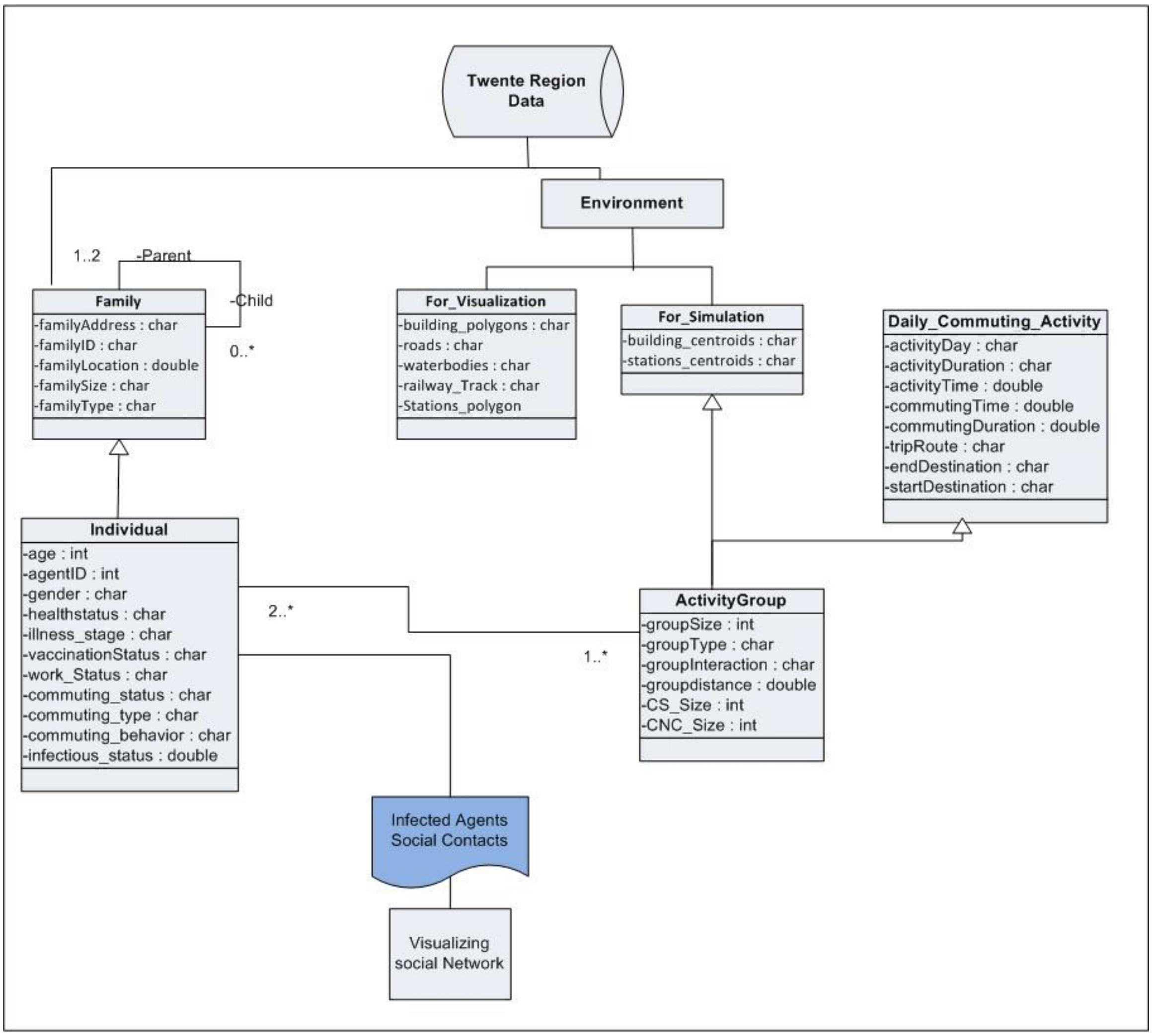


Figure 3.5: The class diagram for the adjusted pertussis model

From the calculated commuting agents, commuting types are classified based on the means of transportation. Commuters commute by various means of transportation either by private or by public transportation. From chapter 2 section 2.4, 60% of commuters in the

Netherlands commuter by private car, 25% by bicycle, 10% by public transport and 5% by walking [71]. These percentages are used to calculate the commuting population that commute by public or private transport. The commuting types are classified by multipying the acquired percentages of commuters by mode of transportation with the total number of commuters created. Commuters that commute by car, bicycle and walking are classified as commuting by private transport. Commuters who commute by private transportation are reclassified as non-commuters. All non-commuters follow thier daily activity (see figure 3.6), modeled in the existing pertussis model.

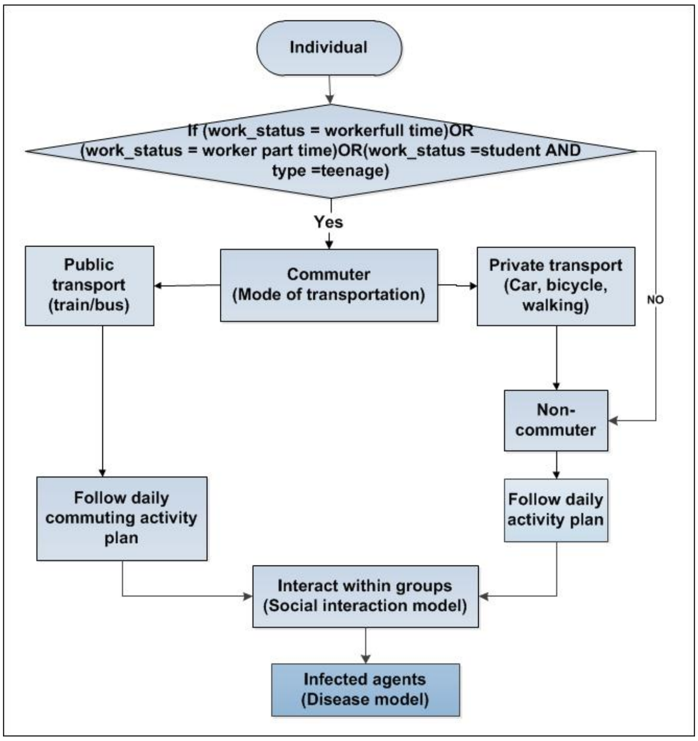


Figure 3.6: Explaining the commuting types

Four types of commuting behavior can be identified from the cross-commuting type of the daily urban systems in the Netherlands from in figure??. There are commuting within the city, commuting out of the city to subareas, commuting from the subareas into the city and commuting from subareas to subareas. From these types of commuting, three types of these commuting behavior are considered. They are commuting within the city, commuting out of the city to subareas, commuting from the subareas into the city because hierarchical spread, spreads from a city to it hinterland or vice versa. Thus the three types of commuting behavior will enhance to detect hierarchical spread.

A commuting agent will have a commuting behavior. Based on the information gathered from 2 in section 2.5.2 33% of commuter who commute to work will be in-coming commuter and 25% of commuters will be out-going commuter. The 42% of commuters will be commute within the city. This fraction is applicable to Enschede but since Enschede is the heart of Twente region, these fractions will be applied to all the localities in Twente region. Also since the fractions of commuting behavior for student is not available, the same fractions is be applied for the student commuter agent. The within commuters are assign to works randomly within their neighborhood by ensuring that such agent have the same home address as that of the workplace address. The in-coming and out-going commuters are assign to works outside their neighborhood by ensuring their home address is not the same as the workplace address.

The commuting agents will identify their start location and their mode of commuting. Within-commuters will commute by bus due to the short travel distance within a neighborhood. The out-coming and in-coming commuters, randomly select their means of transportation. Every neighborhood has one bus station but the original train station in Twente region is use as train stations in the model to depict reality. The agents identify the closes bus/train station by checking if the station is within his/her neighborhood. For commuters that have to commute by train but there is no train station in their neighborhood, commute by bus to the nearest train station. The centroids of stations are used to represent the means of transportation. The explicit movement of the means of transportation is not modeled, since the focus of the model is on how disease spread during commuting. The train is made up several compartments but for modeling purpose only three compartment are considered.

If the agents mean of commuting is by train, the agent is randomly assigned to a train compartment at the train station. The maximum number of agents for each compartment must not exceed 50 agents. The number 50 is a rough estimation of the number of seats in a train. The commuters in each compartment forms a group called compartment commuters. The number of agents within the compartment commuters is calculated to ascertain if commuting agents in each compartment are overcrowded during peak hours or not. This will aid to determine how rapid the disease will spread during commuting because during overcrowding, and the close distance between agents causing the disease to spread rapidly.

Subgroups called close neighborhood commuters (CNC) is formed in buses and in trains. The CNC is formed based on the distance between 2 to 4 commuters. The distance between agents should be greater than 8 meters and less than 15 meters. The 8 and 15 meters are approximate measurement between the seats in a train. Full or partial interaction will takes place within the sub-group. During interaction the disease can be contracted or transmitted based on the health status of the agent. For infection to occur, the calculated contact duration should be equal to hour because the duration cause agent to be susceptible.

The commuter, commute by moving from one station to the other until at their destination, then s/he move to the next activity location. At stations the commuting agent check if it is her/his destination or a transit station for the train commuters (see figure 3.7). If it is the commuters destination s/he end his/her commuting activity. When a commuters alight or board a train/a bus, the CS and CNC is updated. For commuters who commute by train, there is a transit station. Hengelo station is considered as transit station in this model based on the train itinerary in Netherlands [55]. At this station, agents who are commuting end their commuting activity and start again if the agent's commuting schedule state it. Agents who are not commuting will start their commuting activity if only the transit station is their departure station. Agents/individuals who end their commuting activity move to the next activity location by following their daily activity in the daily activity model, then interact with other agent. At the end of a commuter day activity, the agent use the processes described above to commute back to his/her home.

Commuting activity plan

The commuting activity plan determines the commuting activities of a commuter which is determined in space and time. This include the origin and destination location, the trip

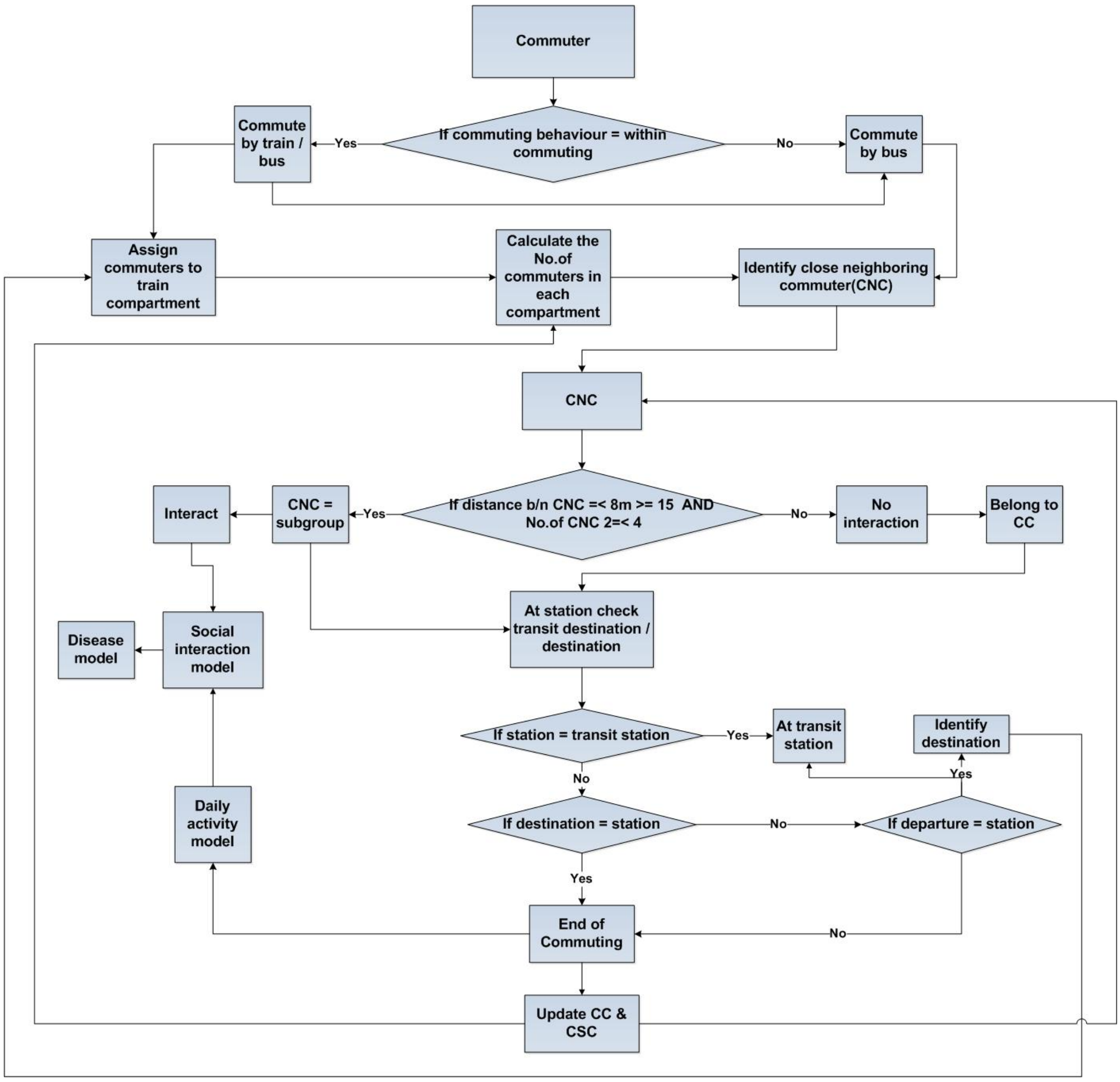


Figure 3.7: Flow chart of the commuting model by public transport

route and duration. The commuter will move to his/her destination based on their commuting activity plan (see Figure 3.8 and 3.9). The commuter agents will move from one station to another until s/he reaches his/her destination. The commuting behavior determines the agents commuting activity plan. The commuting activity pattern use in designing the commuting behavior of Twente region is base on the conceptual individual activity pattern produce by schwanen and Dijst [67] which is shown in chapter 2 in figure 2.9.

The existing activity table consist of the type of agents, the day of the week, the activity time,activity location and the duration of activity. The commuting activity table will be integrated into the daily activity table, which makes it easy to model the changes of commuting agents from commuting activity to his/her daily activity. The commuting activity table will consist of the commuting type, day of activity,origin, destination, time and duration.

The commuting time depends on the type of commuter, for instance a student commuter

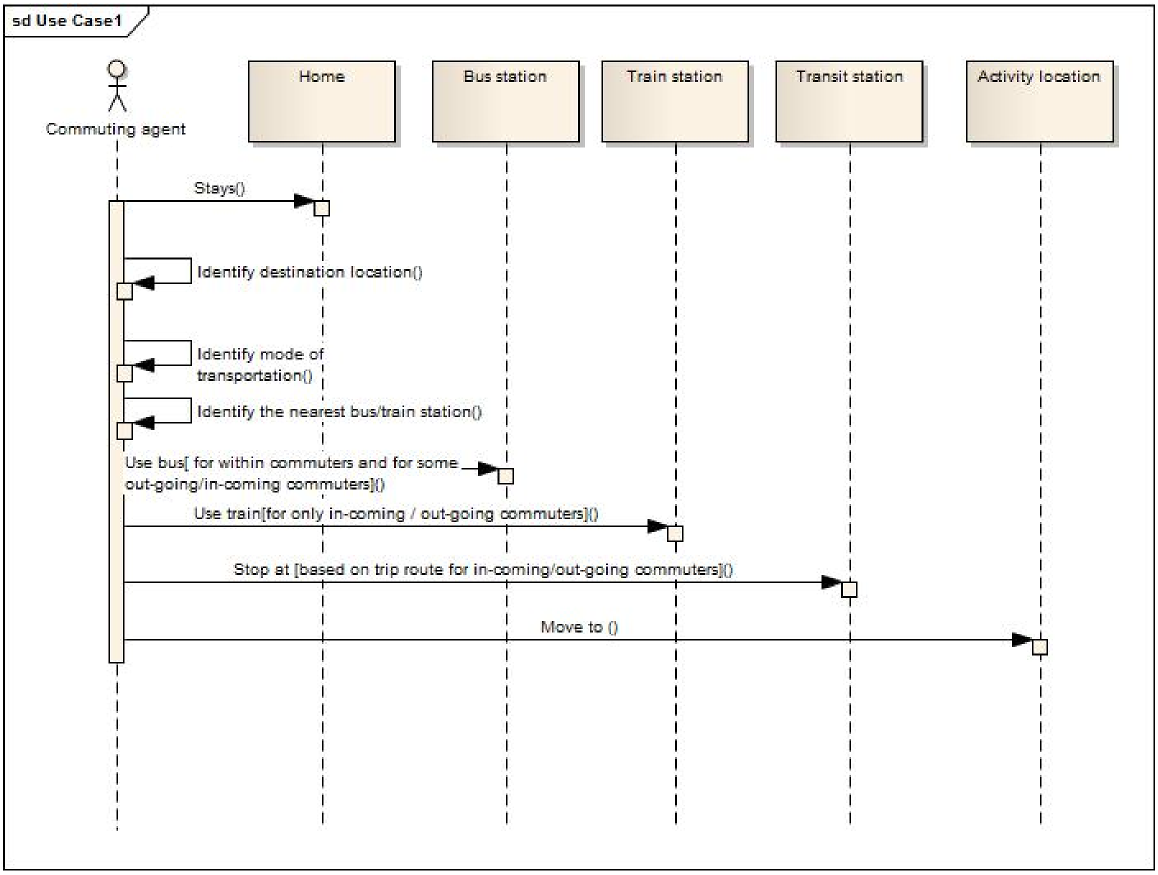


Figure 3.8: Sequences diagram, showing a commuting agent commuting from home to the next activity location

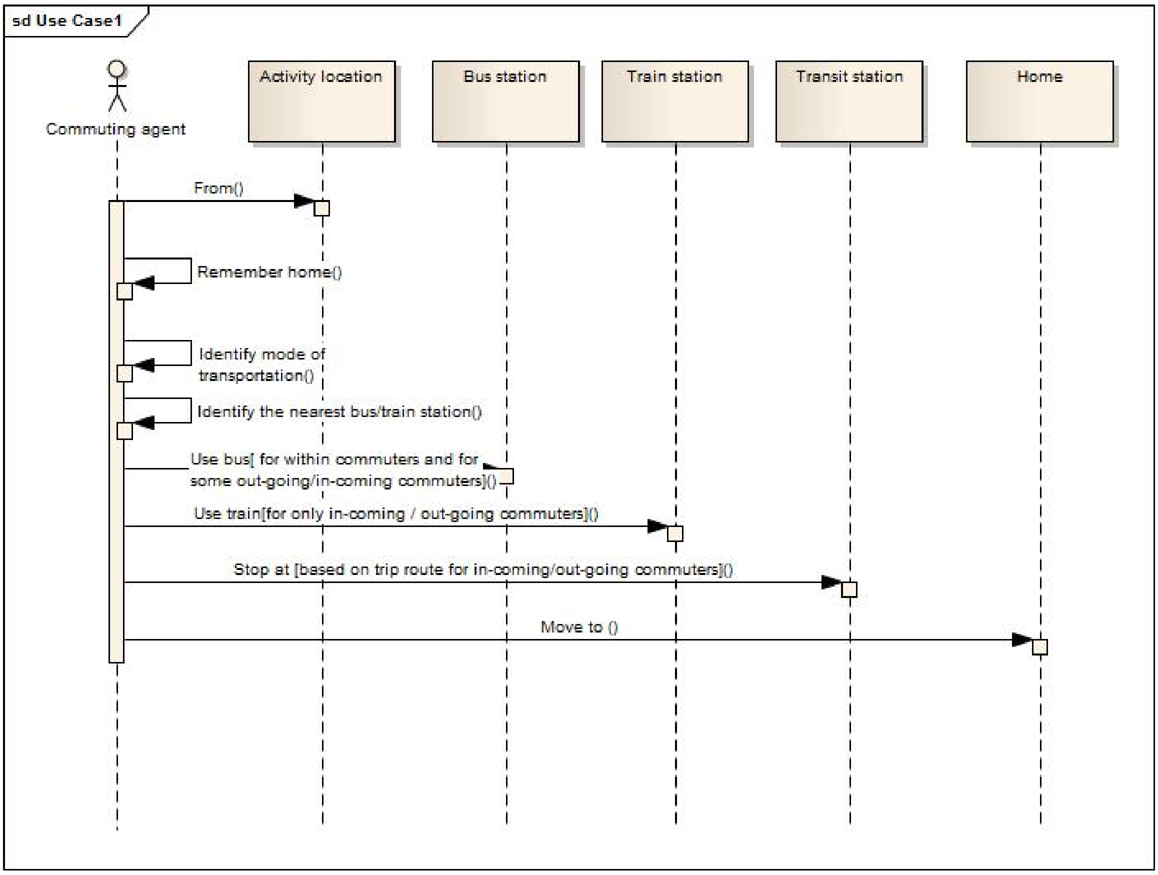


Figure 3.9: Sequences diagram, showing a commuting agent commuting from an activity location to his/her

start to commute to school at 8:00 am in the morning and commute from school to the house at 2:00 pm. While a worker commuter start to commute to work at 8:00 am and commute back to the house at 5:00 pm. The time agents get up in the existing will be change, since commuting agents have to commute before starting their daily activity. The simulation time (tick) in the existing pertussis model is maintained, which is 30 minutes in real time but simulation period is increased from one month to two months to make infected agents to recovery.

Summary

In this chapter, the existing pertussis model is describe. The conceptual design to integrate the commuting behavior in Twente region into the existing pertussis model is explained.

Chapter 4

Data preparation

Chapter 3 described the conceptual design for the adjusted pertussis model. This chapter describe the data preparation and the process used to obtain the required input data for the model. This chapter is divided into three main sections. Section 4.1 describes the data preparation for the empirical disease data, section 4.2 the preparation of commuting data is presented in this section. Lastly, section 4.3 describe the data preparation used to acquire the input data for the adjusted pertussis model. Figure 4.1 shows data preparation structure.

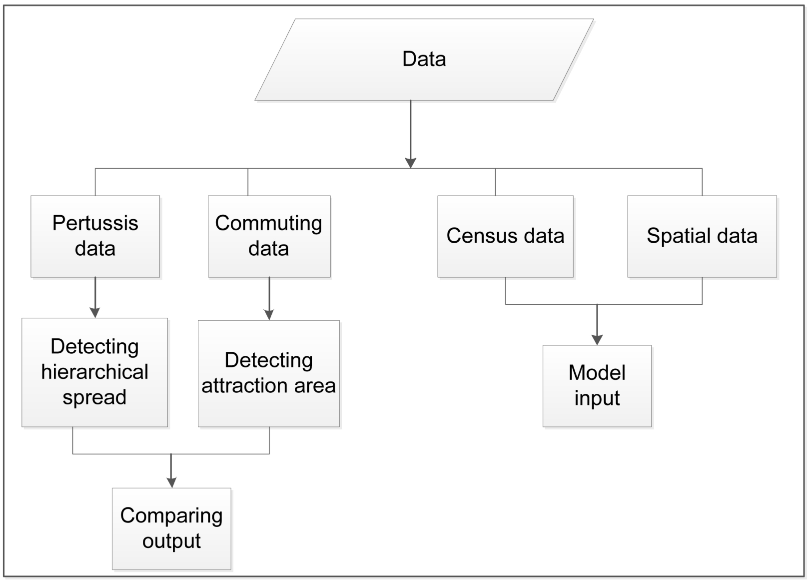


Figure 4.1: Data structure

4.1 PERTUSSIS DATA

Pertussis data was received from ITC Faculty of Geo-information science and earth observation). Pertussis data consist of empirical pertussis reported cases in Microsoft access database format and spatial data of Netherlands which contain shapefiles of the province, municipalities and postal codes of Netherlands. The empirical pertussis data contained reported cases from January, 1st 1993 to December 2004 and the postal code of the reported cases. The empirical data does not contain information regarding the age of the patient, since the age of the patient relate to the transmission of the disease [2] which will be useful in the implementation of the model. From Greff et al. [37] it was found out that the number of pertussis incidence relative to age greater than or equal to 20 years was less compared to the ages below 20 year. It can be implied that only a small fraction of adult reported cases are present in the acquired pertussis data. This is because adult exhibit mild symptoms which make it difficult to be diagnose and due to the mild symptoms, most infected adults do not report to the hospitals [43].

Source of pertussis data

The Center for Infectious Disease Control under the National Institute for Public Health and Environment is responsible for the control and prevention of infectious disease such as pertussis. The institute is the custodian and monitoring agency of infectious disease data in Netherlands [81]. This makes the available pertussis data authentic but there are some disparity in the data.

Extraction of Pertussis data for Twente region

The study area covers only the Twente region but the data received covers the entire Netherlands. The study area is extracted from the received data. The Twente region is extracted by loading the municipal and postal code shapefiles of Netherlands in ArcMap. The 14 municipalities within the Twente region are selected and merged. The generated Twente region shapefile is use to generate the postal code shapefile for Twente region by using the clip tool. The attribute of the Twente region postal code shapefile is use to extract the yearly reported pertussis cases for Twente region.

4.2 COMMUTING DATA

The commuting data was acquired from Omintrans, a transportation model for Twente region. The commuting data contains area and centroid shapefiles of Twente region as well as the data for two skim matrices. The area shapefile consist of five zoning postal code for Twente region with unique identification for each area and unique identification number of centroid in each area.

The centroid consist of the centroid of each zone, unique identification number of centroid, the corresponding postal code four number and code representing municipalities where each zone belong to and the X and Y coordinate of the centroids. Other attributes of the centroid shapefiles are the production and attraction trips for various purpose such as education, work, shopping and business trips among others. Only production and attraction trips for educational and work purposes are considered. Because workplaces and school are places where people spend most of their time apart from home and interact with others. Thus workplaces and educational institution are areas people can easily contract disease.

Production trip is the number of people moving from a postal code (home) for a particular purpose(work or education). While attraction trip is the number of people moving towards a particular postal code within Twente region. The number of production and attraction are based on the mode of transportation such as those who posses a car, those who do not posses a car and unknown possession of car. Table 4.1 is an example of the commuting data. In the table P represent production trip. ABNB represent those who's car possession is unknown, AB - for those who possess a car and NB is for those who do not possess a car to commute.

The data for the skim matrices are in comma separated value (CSV) format. The data consists of the total travel time from one zone to another by means of public transport and in-vehicle travel time for these trips for all combinations from one postal code to another. But the zones was represented by the centroids numbers of each area which was based on five zone postal codes.

Table 4.1: Example of commuting data

PC	*Centriod*	*x*	*y*	*Pwork-ABNB*	*Pwork-AB*	*Pwork-NB*
7511	1	258095	471345	69	70	65
7511	2	258277	471387	16	0	40
7511	3	258271	471128	70	40	50
7511	4	258051	471096	2	200	47
7511	5	257834	471104	17	72	10
7512	6	258603	470732	41	0	100

Processing commuting data

In order to acquire the skim matrices from the skim data, the following procedures was done. The total production and attraction trip for work and educational purpose are acquired by summing up each of the production and attraction trips per postal codes. The skim data is loaded in a database (Microsoft Access). The attribute of the centroid shapefile is exported into the database format and loaded into the same database as the skim data. The purpose of loading the centroid attribute into the same database is to able to assign the postal code number to the centroid number in the skim data. That was use to aggregate all internal trips within a postal code. The centroid numbers are assigned to the four postal codes numbers. An aggregation is done based all internal trips within a postal code and for all combinations between postal codes and averaging the travel time and and in-vehicle travel time is calculated. The results obtained from the aggregation, is use to generate two matrices for the total travel time and in vehicle travel time.

4.3 DATA PREPARATION FOR THE ADJUSTED PERTUSSIS MODEL

Scaling-up the existing pertussis model at a city scale to a region scale require input data for to scale-up the adjusted pertussis model. This section describes the into data to scale-up the adjusted pertussis model.

The total population of Twente region for 2009 is approximately 620,000 but this population is not the required population needed to simulate the existing pertussis model. The model simulate population with children, even with the population with children in Twente region the existing pertussis model is unable to simulate such population. Because the existing pertussis model is able to simulate approximately 50,000 agents on a single computer [1, 28]. Due to this, the actual population for Twente region has to be reduced to achieve the required population for the model. To reduce the actual population with children for Twente region to the required population the following criteria are considered.

(a) How to distribute agents over Twente region.

(b) How to distribute workplaces, school and other social centers for the required population that will enhance social interaction to take place.

Based on these criteria a method to reduce the actual population with children to the required population for the model has to be determined. To determine which method to use to reduce the actual population, the way the model distribute agents, workplaces, schools and other facilities are analyzed.

From the analysis, the model randomly distributes agents to houses within a neighborhoods. Kindergarten and elementary schools are selected by children (agent type) based on the Euclidean distance. The existing pertussis model identify children within a zone making children to be assigned to school within their zone. But if there are more than one school for the child within the computed distance, the child randomly selects one of the schools and save the coordinate of the school in his/her memory. The same procedure is used by children to select social facilities such as music centers and sport centers etc. Teenagers select their high school and social facilities randomly and save the coordinate in memory. Workers are assign to their workplace randomly within the city and save the coordinate for the. Agents select shopping centers randomly, also the agents are assigned to religious building based on distance.

Based on how the model assign agents to various location especial to school and workplace, a method of reducing the population is required.There are several possible methods to reduce the actual population to the required model population, some of the possible methods are

(a) To reduce the actual population proportionally over the entire whole area.
(b) To reduce the actual population for a selected number of area.
(c) To omit some of the rural areas in all direction North, South, East and West and then reduce the population for the remaining areas to the required number.
(d) To reduce the actual population based on areas where people are attracted to and originate from in Twente region that is identify in the commuting data.

Each of the possible methods will produce a different outcome. From the list of possible methods, the first and the fourth option are used to reduce the actual population to the reduced modeling population. The first option is considered because proportional distributing agent over Twente region helps to identify the spread pattern of pertussis in Twente region. Also using commuting behavior detected in the commuting data helps to test the hypothesis of this research. That daily commuting contribute to the spread of infectious disease from a city to it hinterland. Thus reducing the population based on the commuting behavior in Twente region helps to achieve the hierarchical spread in pertussis for Twente region. But due to error of data in the commuting data the reduction of the actual population based on the commuting data was not done.

Twente census dataset

The census dataset which was acquired from the Central Bureau of statistics website [73] for all the fourteen municipalities within the Twente region. The census data for 2009 was downloaded and translated to English with Google translator. The population data for 2009 was chosen, to be able to compare the spread of pertussis in Enschede in the exist pertussis model to the adjusted model. The derived data was in a tabular form containing population data per postal code for all the age groups ranging from 0 - 5 to 95 and above and also the average household size. The postal code data is an essential data required, since that is used to represent the neighborhood of the agent to detect the location of the agent.

Calculation for the estimated population

To proportional reduce the actual population to the required model population, an estimated population per municipality is calculated which is used to generate the estimated population per postal code per municipality based on the actual population per postal code. *Pm* is the estimated population per municipality, *Pam* is the actual population per municipality, *Pe* is the estimated total population for Twente region(50,000) and *Pat* is the actual population of Twente region. Equation 4.1 was used to the estimated population per municipality.

$$Pm = \frac{Pam * Pe}{Pat} \tag{4.1}$$

$$Pc = \frac{Pa * Pm}{Pam} \tag{4.2}$$

Based on the result acquired from equation 4.1, the estimated population per postal code is obtained based on equation 4.2. With *Pc* as the estimated population per postal code and *Pa* is the actual population per postcode. The table below shows the actual population and the estimated population per municipality.

Table 4.2: Actual and estimated population of Twente region,2009

Municipal name	Actual population per Municipality	Estimated population per municipality
Almelo	72428	5829
Borne	21106	1699
Dinkelland	26066	2098
Enschede	156071	12560
Haaksbergen	24495	1971
Hellendoorn	35846	2885
Hengelo	80925	6513
Hof van Twente	35151	2829
Rijssen-Holten	36787	2961
Losser	22589	1818
Oldenzaal	31764	2556
Tubbergen	20992	1689
Twenterand	33605	2704
Wierden	23467	1889
Total	621292	50000

Since pertussis is related to children, families with children such as couples with children, single mothers with children and single fathers with children is the fraction of population considered for the simulation. The population data for the families with children per postal code was not accessible from the Central Bureau of statistics website for all the fourteen municipalities within the Twente region. Thus the population for families is retrieve for the whole of Netherlands which contains population of single men with children, single women with children and couples with children for years ranging from 1915 - 2010 [74]. The proportion of these family types with children was calculated from the census data for Netherlands. From the calculation the proportion of single men with children was approximately 0.01, single women was approximately 0.23 and with couples with children the proportion of single men and single women with children was added up to obtain 0.42 to represent couples with children. The obtained proportion was applied to the summed population per postal code age ranging from 25 - 60 years by assuming that, the population with children ranges 25 - 60 years. The estimation of population of single men with children, single women with children and couples with children was obtained by using equation 4.2 above.

Twente Spatial Data

To apply the behavior of agent during the simulation an environment is needed which will assist the agents to interact. The spatial data of Twente spatial data of Twente region is the environment used. As the model tries to replicate the spatial objects in real world the spatial objects (building locations) was considered because they are the location where agents (individuals) meet and interact with other agent which will enhance disease diffusion. Two main purpose of spatial data used for the model are for the model environment and for the visualization of the model.

(a) The spatial data used for building the model environment are:

- Municipal
- City neighborhoods
- Buildings (points(centroids of buildings)):
- Residential building
- Schools (kindergarten, elementary, high schools, other higher institutions)
- Work locations
- Stations(train/bus) item Other (Departmental stores, sport center, Churches etc.)
- City business areas

This spatial data give the agents their geographic location where their activities take place.

(b) The spatial data that will be used for visualization of the model are:

- Buildings (polygon)
- Main roads within the municipality
- Water bodies within the municipality
- Railway tracks

This spatial data give the pictorial view of the model and the study area.

The spatial data for Twente region was acquired from ITC (Faculty of Geo-information science and earth observation), consisting of all the topographic features of Netherlands in a vector format at a scale of 1:10,000 which is based on the sheet index of Netherlands for 2004 - 2005.

The sheets index for Twente is identified and used to download all the shapefiles that covers Twente region. The naming of the shapefiles as well as their attributes are translated to English with Google translator. The shapefiles that are downloaded are the houses (huizen), areas (vlakken) and lines (lijnen). The houses consist of building that will serve as the environment for the model. The areas contains the roads, water bodies, large buildings and departmental stores in a polygon form. The roads and the water bodies are for visualization purpose, where as the large buildings and departmental stores will serve as part of environments of the model. The line layer contains all the line features including the railway tracks which is also used to visualize the model. These individual shapefiles are merged by using the merge tool in ArcGIS. The merged shapefiles are clipped by using the clip tool to obtain only the features within the Twente region extent.

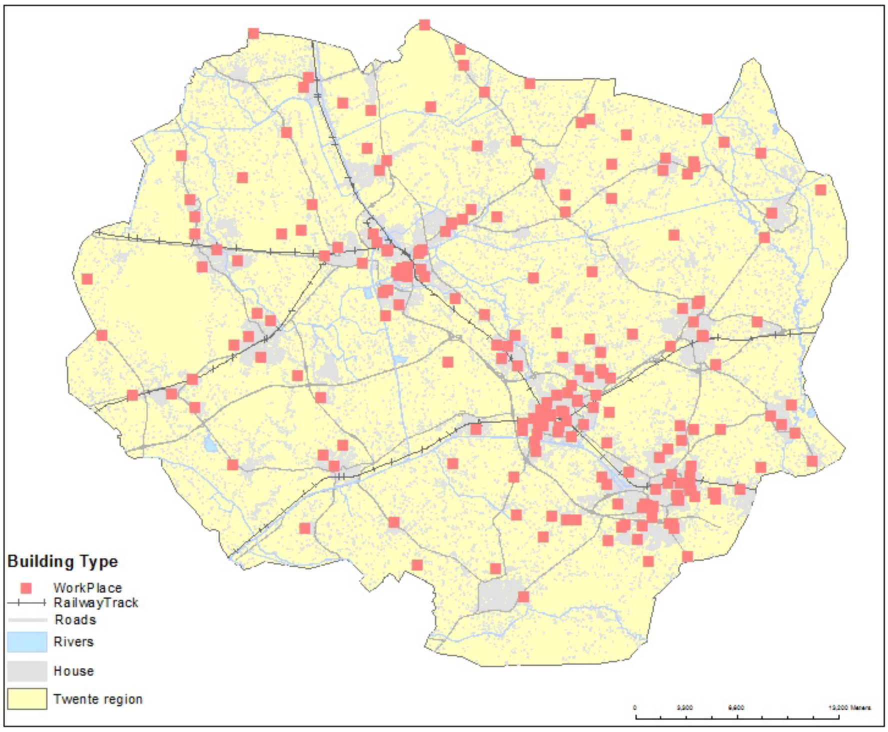


Figure 4.2: The geographical distribution of workplaces in Twente region

The building shapefile did not have the exact attributes of the building type and the address of the building. These attributes are crucial attributes needed for the model to identify the type of building which tells the model the function of the building when agents are assigned to their activities. The building address is created based on the buildings within a postal code. To determine how to assign the building types, the building types in the existing pertussis model are analysied to find out how the building types are assigned. The building types in the existing model are assign by finding the location of the building type in a telephone directory and cross checking it in Google Earth. Due to the large spatial extent and the large number of buildings within Twente region, this procedure is not applied. The building types are assigned by using a command that divide the unique identification number (fid) for each building by 50 and the remainder is use to classify the building types. For example the building type is a school if remainder is 1. The number 50 is an arbitrary number chosen to have more house than other types of building. The assigned building types is cross check to verify if at least each neighborhood has most of the building types. Figure 4.2, 4.3 and 4.4 shows the distribution of the building types.

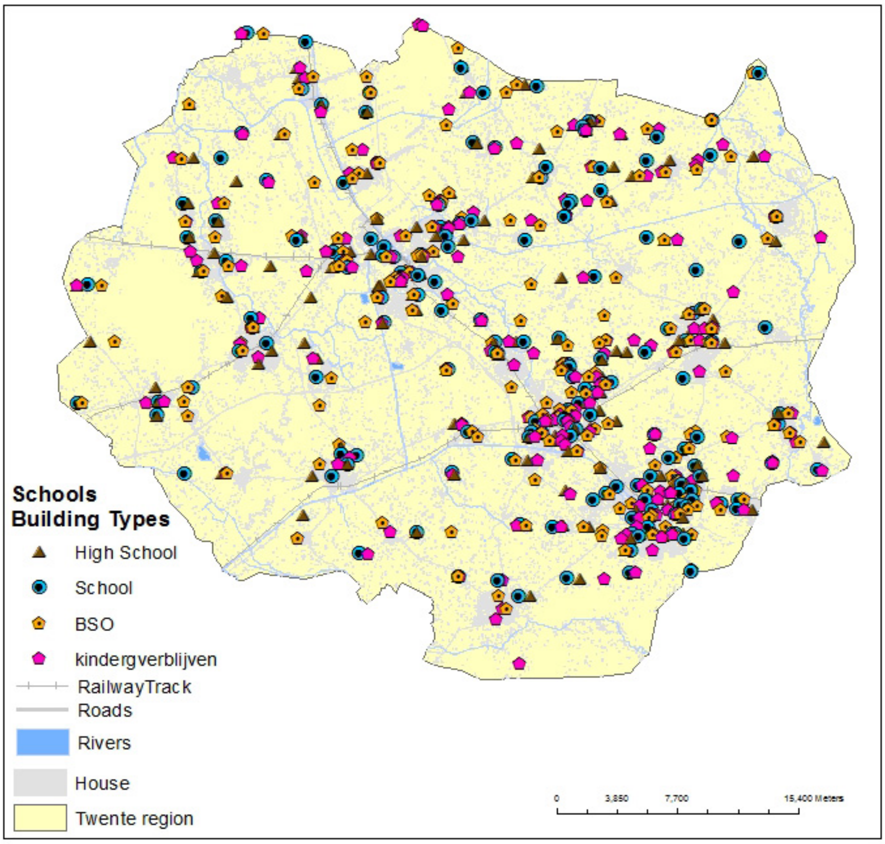


Figure 4.3: The geographical distribution of schools in Twente region

Data limitation

All the data acquired has some limitation. In the pertussis data the age of the patient was not reported. This data will have been use to determine the reported cases per age group per areas (postal code). This will have been used to set the adjusted model. Acquiring the commuting data was difficult, even though this data was acquired, the data has an error. Total production and attraction trips was almost the same. Due to this the areas where people commute from was not found. It will have been used in analyzing pertussis data. With the input data, the attribute of the building type and address was lacking. This will have been use to redistribute the building types in the model.

Summary

In this chapter, pertussis, commuting and required input data for the adjusted pertussis model are prepared.

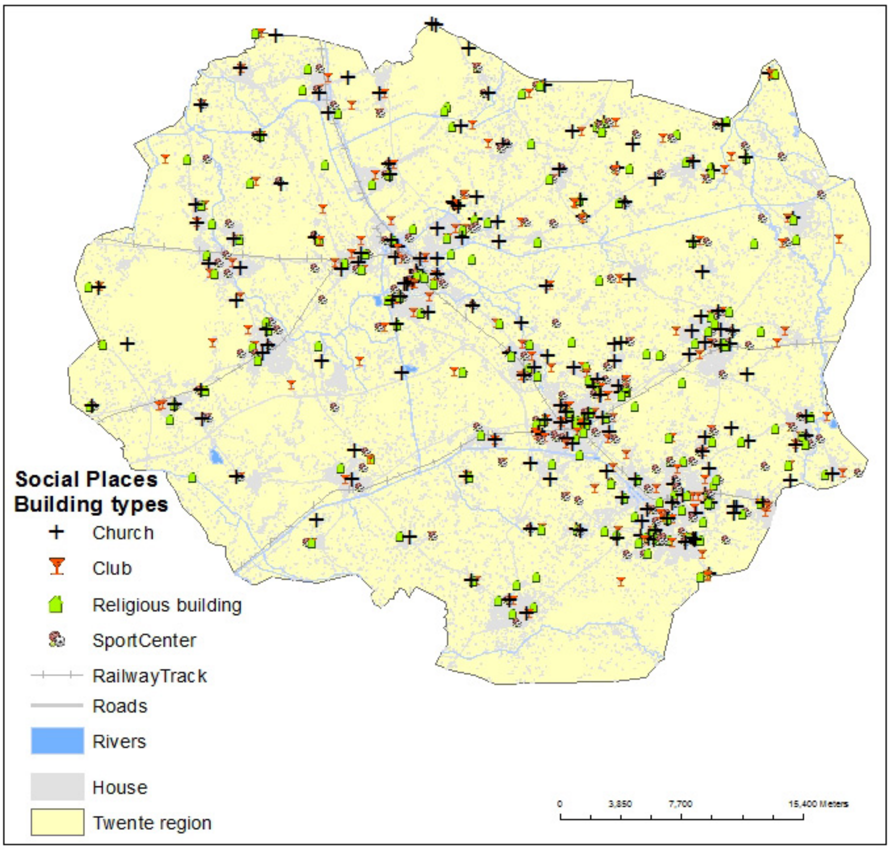


Figure 4.4: The geographical distribution of social places in Twente region

Chapter 5

Implementation and Results

This chapter, present the approach use to analysis the pertussis and commuting data and also the scaling-up of the existing pertussis model. Section 5.1 present the approach used to detect hierarchical spread in the empirical data of pertussis and the obtained result. Section 5.2 addresses the approach used to detect the commuting pattern from the commuting data and the obtained results. Lastly, section 5.3 present the implementation approach used in scaling-up the existing pertussis model.

5.1 DETECTING THE HIERARCHICAL SPREAD IN THE EMPIRICAL DATA OF PERTUSSIS

The pertussis empirical data is analysis to detect the hierarchical spread by adopting the approach used to detect hierarchical spread in pertussis by Broutin et al. [12]. They determined the number of pertussis cases in weeks for each endemic and epidemic year for each locality.

The number of pertussis reported cases in Twente region per month per postal code for both endemic and epidemic years is obtained. The obtained data is joined to a postal code shapefile which is used to display the spatial pattern in each month for each postal code in Twente region for the endemic and epidemic year from 1996-2004. The obtained spatial pattern depicted hierarchical spread in the empirical pertussis data. Railway tracks and train stations are overlaid on the observed hierarchical pattern in the empirical pertussis data. The railway tracks and train stations helped to understand the frequent incidences of reported cases in some areas. In appendix 6.1.1 the observed hierarchical pattern of pertussis from January to December for 1996-2004 is shown.

5.1.1 Result from the empirical pertussis data

Analyzing the pertussis data, it shows that the incidences of pertussis start from the three cities in Twente region - Hengelo, Enschede and Almelo. From figures 5.1 5.2 5.3 it shows that, the incidence of pertussis from 1996-2004 in the three cities which source of the disease. From the identified sources, the disease spread out to the hinterlands as identified by Broutin et al. [12] and Grenfell et al. [39] in chapter 2 section 2.6.2. This is prove by plotting of weeks of reported cases in 1996 in each postal code in Twente region against the distance from Enschede as the source in Microsoft excel as shown in figure 5.4

From the plot it is observed that pertussis start from the cities. The pertussis remain in the cities than spread out to the hinterlands. From the spatial distribution map 6.1.1in appendix some years for example 2004 depict that transportation is a contribution factor to the hierarchical spread observed in the empirical data.

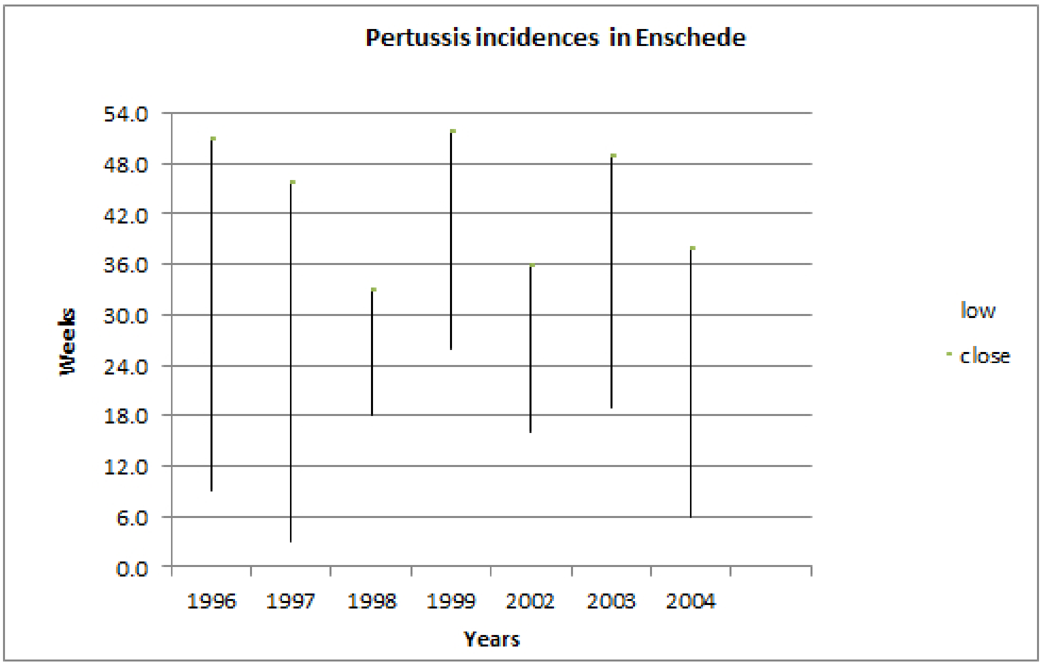


Figure 5.1: Pertussis incidences in Enschede from 1996 - 2004

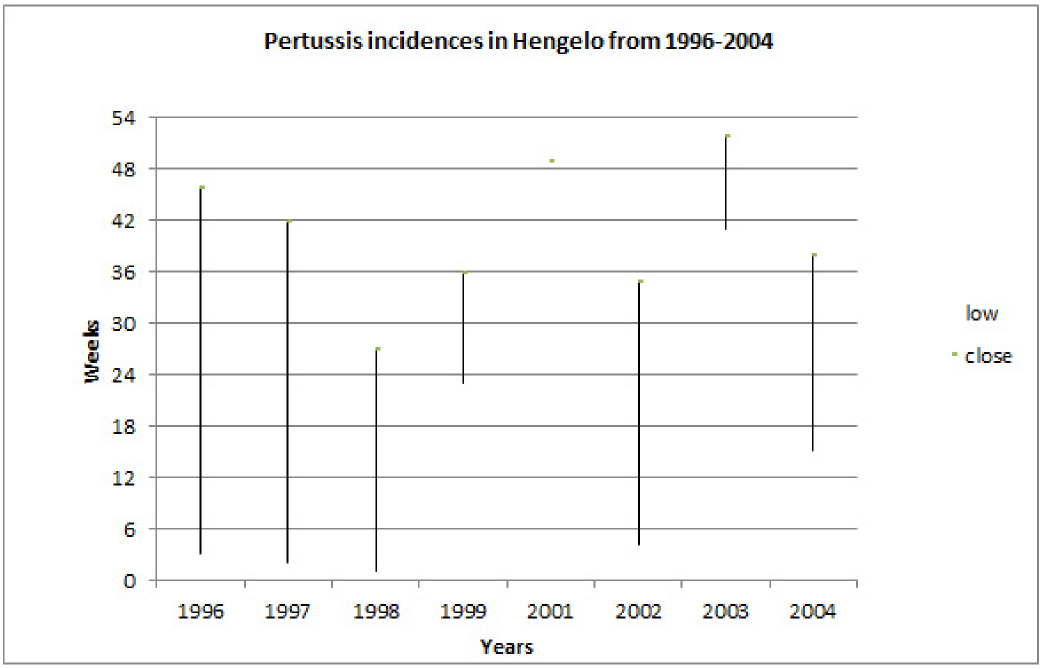


Figure 5.2: Pertussis incidences in Hengelo from 1996 - 2004

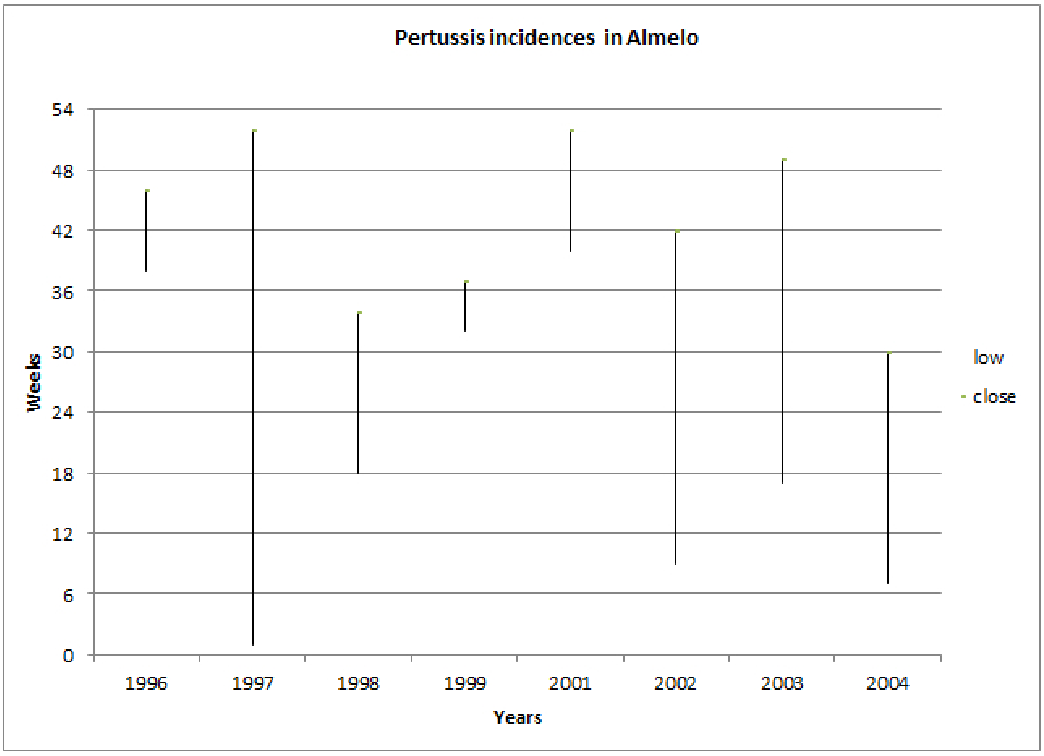


Figure 5.3: Pertussis incidences in Almelo from 1996 - 2004

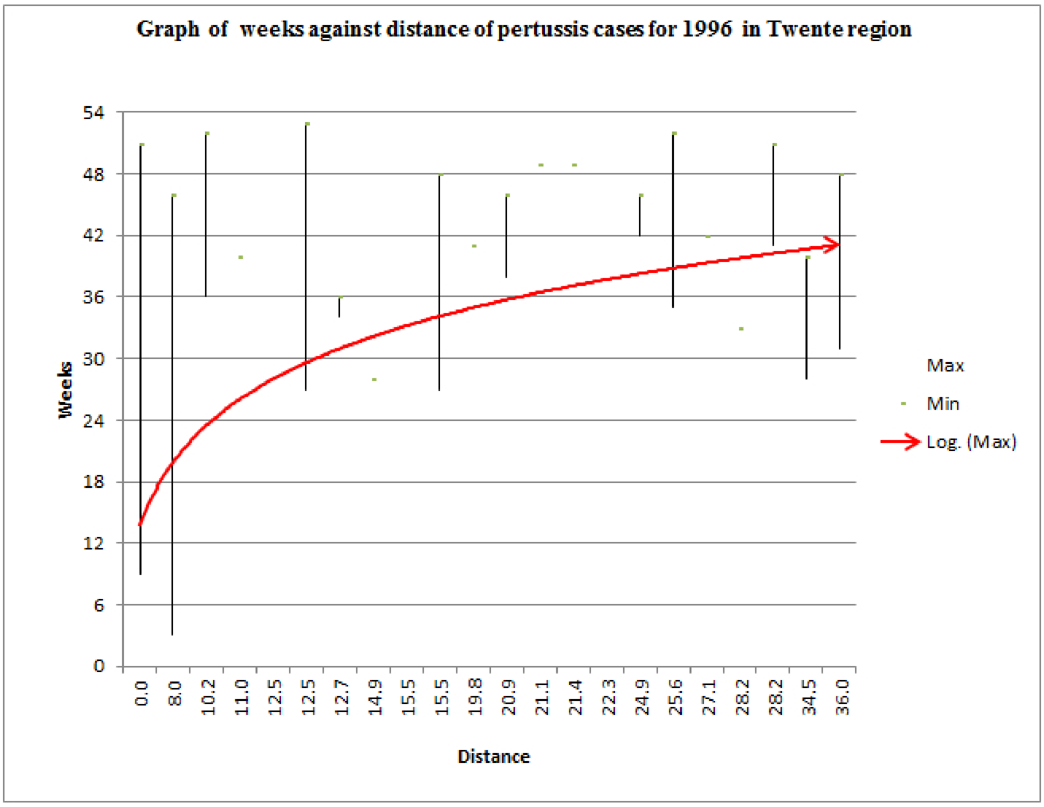


Figure 5.4: Showing the hierarchical spread in the empirical data

5.2 DETECTING THE COMMUTING PATTERN IN TWENTE REGION FROM THE COMMUTING DATA

The commuting pattern in Twente region is obtained from the commuting data by applying the furness method. The furness method is use to determine the number of trips per unit time between two zones [62]. Before applying the the furness method the skim matrix has to be obtained. Below are the processes used to obtain the skim matrix.

The processes used to obtain the matrix of travel time from the commuting data is described in chapter 4 in section 4.2. Matrix of travel time is used to produce the skim matrix by dividing the travel time per postal code by 1. The sum of production and attraction trips for work is calculated and a cross check is done if sum production and attraction trips for work is balanced per postal code. The total sum of production and attraction trips for work was not balance. A balancing factor is applied to the attraction trip values to obtain an adjusted attraction trip. The balancing factor is obtain by dividing the production trips for work by attraction trips for work. The furness method is then applied to the skim matrix to obtain the origin destination matrix which is use to determine areas where people are attracted to. Since there was an error in the production trips, the areas where people originate from was not obtained.

The following are the procedure for the furness method that was used.

(a) The sum production trip of travel time per postal code is calculated and the zonal growth factor is calculated. The zonal growth factor is obtained by dividing the sum of production for work per postal code by the sum of production trips of travel time per postal code .

(b) The zonal growth factor is use to multiply all the columns in the matrix.

(c) the sum of all attraction trips of travel time per postal code is obtain and the destination factor is obtained by dividing the sum of an adjusted attraction trips for work by the sum of attraction trips of travel time per postal code.

(d) The destination factor is multiplied by the all the rows in the matrix.

(e) Iteration processes of step 1 - 4 is repeated until the calculated growth factor and destination factor of 2% is obtained. This procedure is also applied to production and attraction trips for educational purpose. The result is discussed in the next section.

5.2.1 Results from commuting data

The finial product for attraction trips for work and education is loaded into ArcMap and flowmap to visualize the areas with more attraction trips. In ArcMap the final matrix is joined to the spatial postcode layer to visualize the attraction areas. The railway track and the train station are overlaid on the spatial attraction areas. The minimum travel time from one area to the other is 27 minutes. Figure 5.5 shows an attraction map, showing areas where people are attracted to for work. Figure 5.6 shows an attraction map showing areas where people are attracted to for educational purpose.

visualizing using flowmap

The final production and attraction trip and the spatial postal code layer is loaded into flowmap. The distance matrix is created based on airline distance, since the data

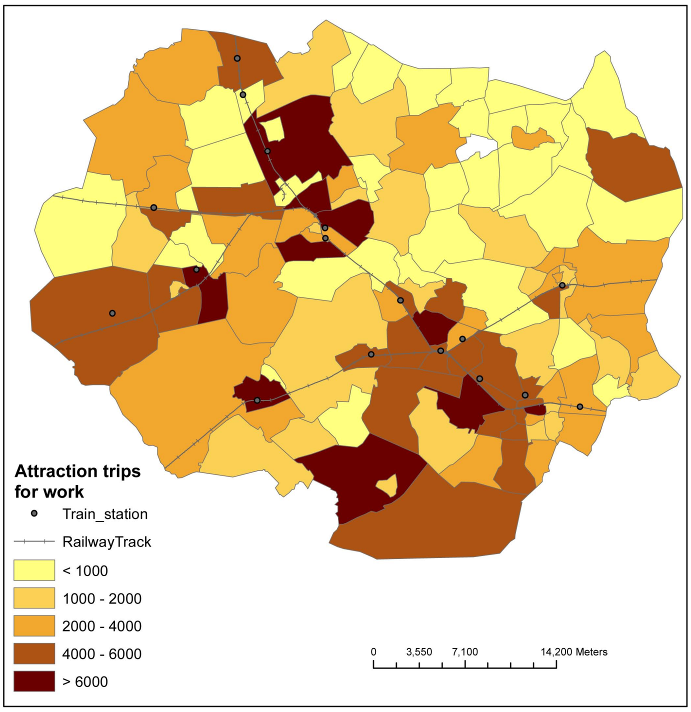


Figure 5.5: Map showing attraction areas for work

for the correct road network was not available. Double-constrain gravity model is performed. The double constrain gravity model is use to fit the distribution function with the production and attraction trips. The gravity model is performed and desire lines based on a threshold from 50 is set. Figure 5.7 shown the desire lines for work trips. Figure 5.8 shown the desire lines for educational trips.

From the commuting data, the areas where people are attracted to are also the areas where reported cases has been observed. For instance Buurtschap Broekheurne is an area with cluster of basic schools where more attraction trips for both work and for educational purpose is observed. Also this area has reported cases in all the observed years(from 1996-2004) apart from 2001. This confirms that there is a relationship between commuting and the empirical pertussis data. The number of disease cases in 2004 is plotted against the log of the number of attraction trips (see figure 5.9). Comparing the spatial distribution of pertussis with the attraction trips, some areas shows a relation between pertussis cases and attraction trips.

5.2.2 Comparing the results of pertussis and commuting data

Analyzing the outcome from both the pertussis and commuting data, it is observed the more attracted areas for educational purpose are in the cities specially Enschede and Almelo. Also in the attraction areas for work, the same cities are among the more attracted areas. The cities serves as the source areas of pertussis spread. The attraction areas in the cities confirm why the cities are the source areas of the pertussis.

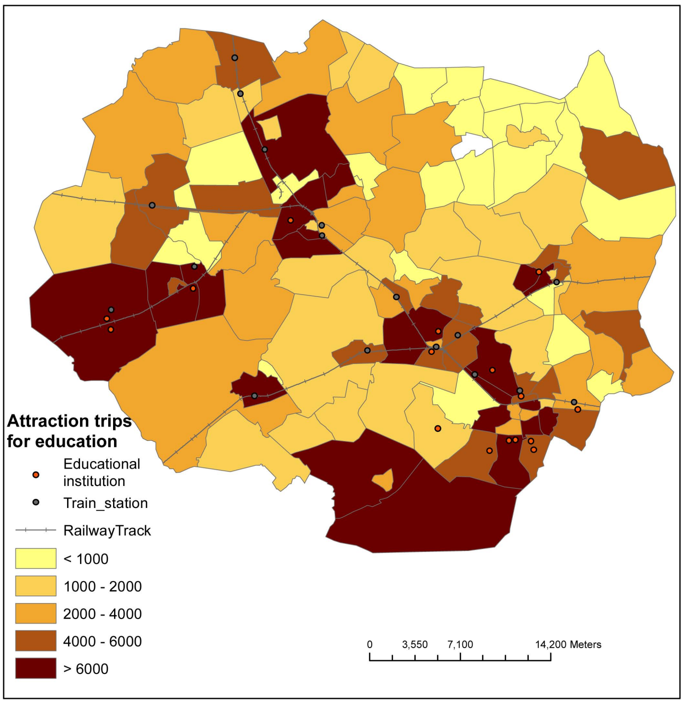


Figure 5.6: Map showing attraction areas for educational purpose

5.3 SCALING-UP THE EXISTING PERTUSSIS MODEL

Implementing the adjusted pertussis model is to scale-up the existing pertussis model. The scaling-up takes into accounts the extent of the area and the number of agents due to the large spatial extent of Twente region and the large number of inhabitance.

In relation to the large number of agents, the existing pertussis model crashes when running for a large number of agent above 50,000. The model is able to initialize large number of agent but is unable to run to find out the number of infected agent. This is because the disease model in the existing pertussis model compute, who and where the infected agent contracted the disease from by finding out who and where the infected agent visited and interacted with. This process require a lot of memory space which causes the model to crash. Thus for scaling-up the existing model for a large spatial extent, in a systematic approach is required.

The following are the approach use in implementing the adjusted pertussis model.

i. Upgrading the spatial extent.
ii. Including public transportation into the existing model.
iii. Upgrading the total number of agents.

5.4 UPGRADING THE SPATIAL EXTENT

Upgrading the spatial extent of the existing pertussis model consist of

- Creating synthetic population.
- Loading the Twente region spatial layers.
- Verification of model.

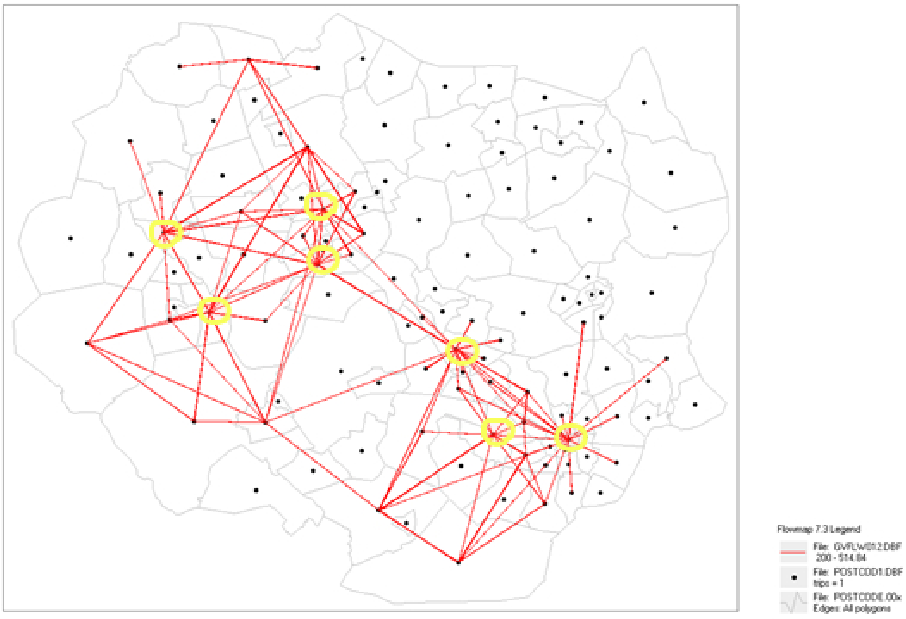


Figure 5.7: Map showing the desire line for attraction trips for work

Figure 5.8: Map showing the desire line for attraction trips educational

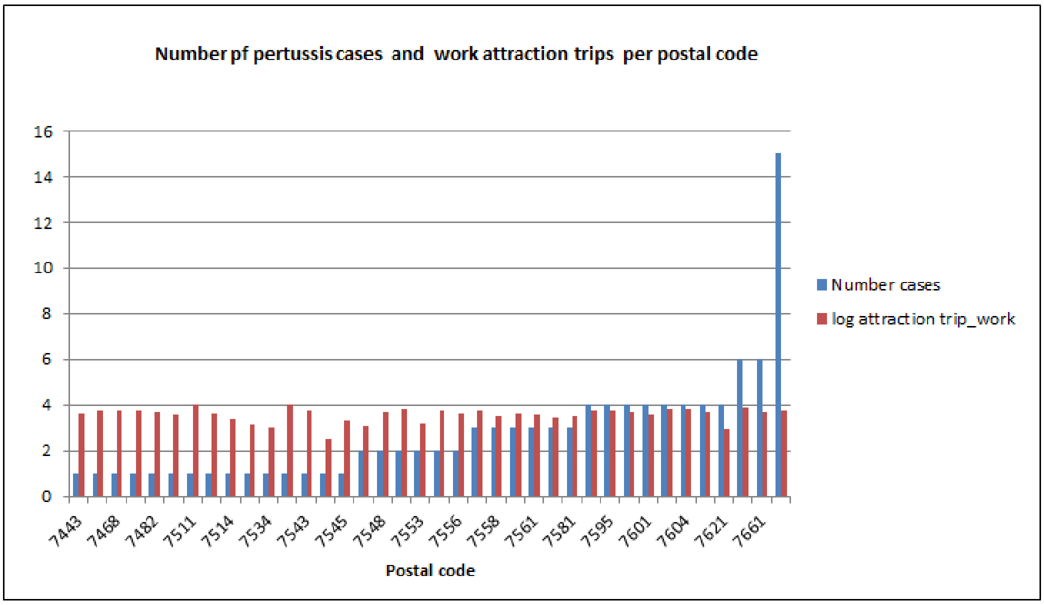


Figure 5.9: Plot for attraction trips against number of disease cases in 2004

5.4.1 Creating synthetic population

This section discuss the procedure used in creating the synthetic population for simulation the model and the types of attribute assigned to the created agents. The prepared census data in a CSV (comma separated value) format is used to create synthetic population randomly over Twente region by using the synthetic reconstruction approach describe in chapter 2 in section ??. The following procedures are used

- The model read the file which contains the neighborhood name, code, the number of families' types and the number of children ranging from 0-19 for each neighborhood.
- Based on the data in the CSV format, household head are created for each neighborhoods and family members are created.
- The final populations to be created are children. The children are distributed to the created families.

The created synthetic populations are assigned with three main types of attributes, which are demographic, changeable, and reaction attributes. Demographic attributes, identifies the agent, example agent ID. Demographic attributes do not change during the simulation run. With changeable attribute, the values of the attribute change during simulation run which can have an effect on the model. For instances an agent change his/her attribute when s/he is sick. The reaction attribute is based on the reaction of the agents response which there is an effect of change. Other attributes the agents can have is to inherited some attribute from their parent or their partner.

5.4.2 Loading the Twente region environment

The spatial layers serves as the environment for the model. It consists of different building types where activities of agents take place. The buildings types deter-

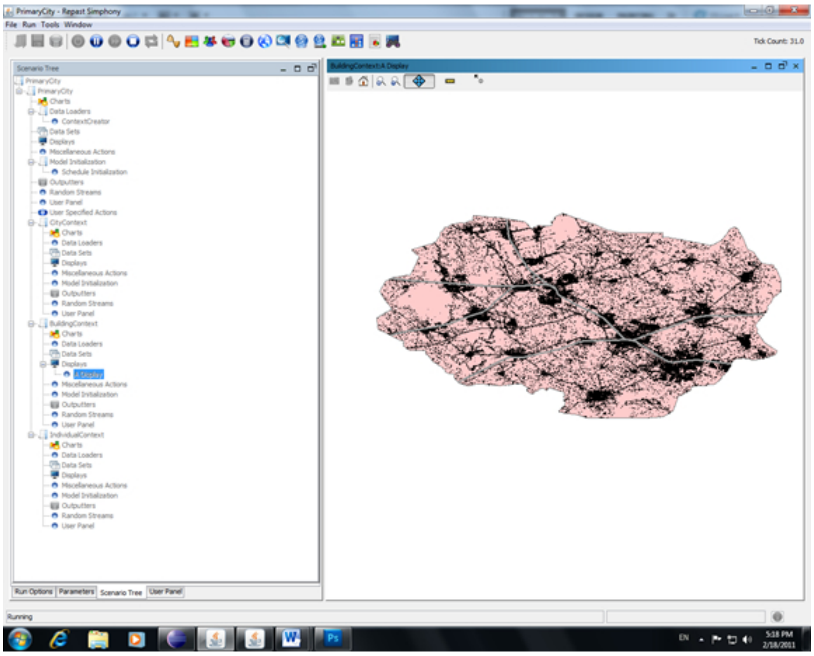


Figure 5.10: The spatial layers of Twente region in Repast

mine the types of activity that take place at a location.
The spatial layer of Twente region is loaded into Repasts 2.0 (see figure 5.10) by editing the *Building context* in the existing model. The *Building context*is a container for the spatial layers for the existing pertussis model. It has a geographic projection which identify the location of the buildings where agents are located. The created agents are randomly assigned to the created centroids of the building. After the creation of synthetic population and the loading of spatial layers for Twente region into the existing pertussis model, the model is run to test if the created synthetic population are spatially distributed correctly. Several test runs was done but no result was acquired. This shows that agent do not form groups and no interaction occur among agents. This imply that an agent is the only one in an activity location apart from home. This implies that, there are more activity location compare to the estimated population in the model. To solve this problem two options are identified below, they are

i. To increase the arbitrary number in the algorithm used to generate the building types. To generate more of the building type for house and less building types for other activity location.

ii. To reduce the building in Twente region to correspond to the estimated population. This is done by including null to the algorithm used to generate the building types and the null attribute is then deleted. After deleting about 50,000 building, neighborhoods with less building had no buildings caused an error while initializing the model.

Based on the outcome, the first option is used. During the initial run agents look for their initial environment (building) and then look for the next location to visit which is done either randomly or based on a constrain.

Table 5.1: Result of simulation run with an infant as the source of infection

Simulation runs	*Total population created*	*No.of infected agents*	*Initial Infected neighborhood*
Initial	46318	10	Geesteren (Kern)
Run 1	46318	8	Geesteren (Kern)
Run 2	46318	3	Vriezenveensewijk
Run 3	46318	16	Vriezenveensewijk
Run 4	46318	1	Geesteren (Kern)

Verification of model

After redistributing the spatial location to correspond to the estimated model population, the existing pertussis model is verify to detect how the model spread the disease over the extended area (Twente region). This was done by first using the initial setting in the model and changing the initial infected agent to a teenager or an adult. Since they are the major source of pertussis [41].
After several simulation runs with no infected case reported for an infant as the initial infected case, the initial setting of the model is adjusted. The initial settings of the existing pertussis model are 96% of infant population are immune. The children between 4 to 9 years of the synthetic population are 96% immune. 95% of the teenage population are immune and 70% of adult population are immune. The percentage of immune adult is reduce to 50% and teenage population immune to 30%. The results of the infant as the inital source of infection are shown in table ??, 5.4.2 and Figure 5.11

Table 5.2: Showing the locations of infection

[ht]

Where agents are infected	Initial run	Run 1	Run 2	Run 3	Run 4
Club	1	2	1	3	0
Youth center	1	3	0	3	0
Shopping center	0	1	0	0	0
Religious building	2	0	0	0	1
Kinderdagverblijver	6	2	2	8	0

The results from the model show that hierarchical spread occurs but the source of the spread is from a village to city. This imply commuting work. No results was obtained for the adult as an initial infected case.Due to time constrain including public transportation into the existing model and upgrading the total number of agents was not done.

Summary

In this chapter, methods used to detect hierarchical spread in the empirical pertussis data as well method used to detect commuting pattern from the commuting data discussed. In addition, the scaling-up of the existing pertussis model is presented.

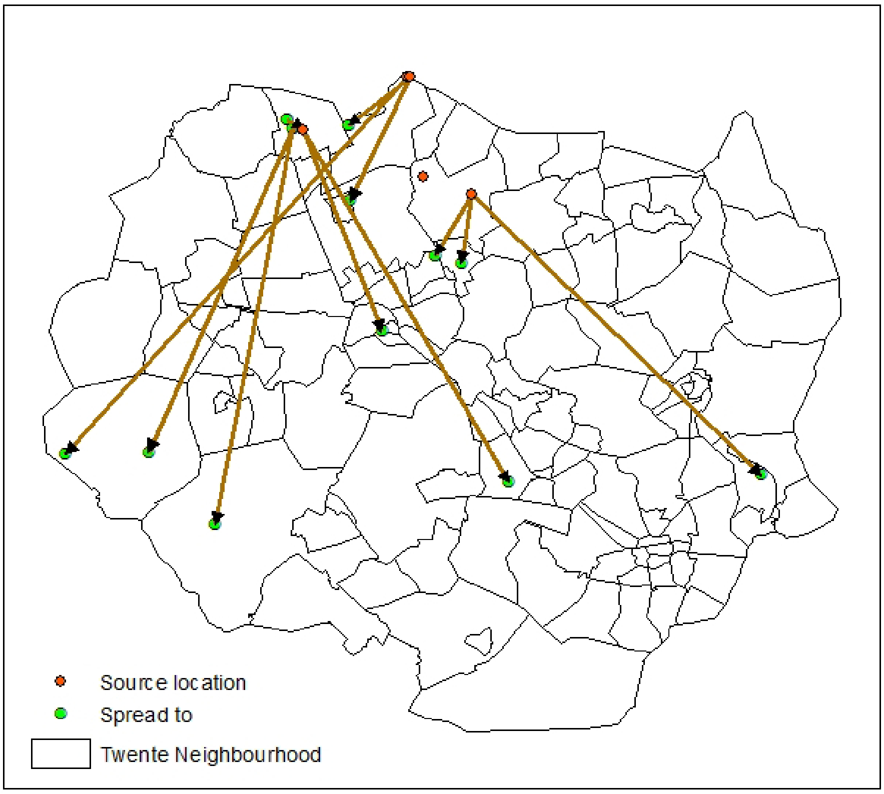


Figure 5.11: Shows hierarchical spread obtained from the adjusted pertussis model

Chapter 6

Conclusion and Recommendation

6.1 CONCLUSION

The objective of this research was to adjust the existing pertussis model to include the commuting component in relation to the hierarchical process for a regional area. In order to do this the following sub-objectives needed to be addressed:

i. To analyzes the spatial patterns in pertussis data to detect hierarchical spread.
ii. To analyzes the daily commuting data to gain understanding on the commuting behavior in Twente region.
iii. To create a conceptual model to integrate the commuting behavior in Twente region into the existing pertussis model.
iv. To test the model and to evaluate if hierarchical spread occurs.

Based on these sub-objectives following processes were done to achieve the set out objectives

i. To analyzes the disease data to confirm that hierarchical spread takes place in Twente region.
ii. To analyzes the commuting data in order to detect the amount of people who commute and the attraction areas in Twente region.
iii. To compare the disease data and the commuting data in order to detect the relationships between the observed patterns.
iv. To develop an adjusted version of the existing pertussis model in which different types of commuting were integrated.

Indeed it was found that hierarchical patterns exist in empirical pertussis data. The spread was from a city to it hinterlands. It was found out that some of the smaller cities (like Almelo) that do not have a size corresponding to the critical community size (CCS) also functioned as a center of spread. This can be seen from the fact that, pertussis is contained in these areas over a longer period of time (months).

We expected to find a cross commuting pattern in Twente region as this was indicated from literature in chapter 2 section 2.5.2. The analysis of the commuting data confirmed this. We were able to find the areas of attractions (like the zone of the University of Twente) and areas of lesser attraction. It also proved possible to determine patterns for both attraction trips for work and attraction trips for educational purpose. These patterns partially overlapped.

Commuting patterns were compared to the disease patterns. The results of the commuting data confirmed the hierarchical spread in the empirical pertussis data. It gave a better explanation to areas with occurring disease case. Some of the areas with reported case showed a correlation to the disease date. For instance in 2004 the distribution pattern observed can be explained by the outcome from

the commuting data. From the analysis, Almelo served as the source of pertussis disease in 1997 and it remains for a long period. Also from the commuting data, Alemlo serve as a source of attract for both work and educational purpose. Even though Alemlo does not meet the critical community size, pertussis remains for a long time due to the commuting behavior in that area.
A conceptual model was developed that includes both commuting by public transport (allowing disease transfer during commuting) and private transport (car,bicycle and walking). The development of the model was done in phases:

i. Upscale to include the total spatial extent (Twente Region).
ii. Include the commuting with public transport.
iii. Increase the number of agents.

Due to time constraints only the first step was achieved. It turned out to be rather difficult to conduct a large number of runs. The runs that were conducted experimented with different initial infected agents. For children of low age it was proven that the disease spreads from village to village due to visiting the same daycare centre. The outcome of the model shows hierarchical spread from a village to a city. This implies that the private means of transportation described in chapter 3 worked but low number of disease cases was acquired. Modeling the public mean of transportation into the existing pertussis model would enhance the number of disease cases, since public transport serves as a mean for contracting disease. The low number of disease case leads to the fact that small number of synthetic population was create for a large spatial extent. Increasing the number of agents will contribute greatly to the number of disease cases.
Limitations:

- Data for commuting. Travel time from one postal code to another was acquired instead of the number of people commuting. In preparing the commuting data, the travel time was not split out per mode of transport, but this can be done. Bicycles should be included at least for education.
- The adjusted model, turned out to be very difficult to run with required model population. The fact that the model is not really running for 300000 or 400000 agents but only for 50000 is a true limitation in this type of work and should be addressed.

6.1.1 Recommendation

This research has implemented the private means of transportation into the existing pertussis model. Integrating public means of transportation into the existing pertussis model would help to detect hierarchical spread in model and increase the number of infected cases. Since hierarchical spread has been detected rather further work should be conducted to upgrade the total number of agent which would lead to more disease cases. Significant number of infected agents can not be achieved because only a fraction of the created population would commute.
The delay in acquiring commuting data for Twente region and raw data acquired restricted us from determining the exact correlation between pertussis data and commuting data. Thus further work should be conducted to ascertain correlation type.
Finally we recommend if the model is tune to adjust the immunity level based on specific location this would enhance the model to replicate reality and depict the

actual spread pattern in Twente region. Tuning the adjusted model to stop when no infected case are detected and also to stop when a specific number of infected cases are reported before the model crashes is also highly recommended.

Bibliography

[1] S.A. Abdulkareem. Simulating the spread of pertussis in enschede region using agent - based modelling. Master's thesis, ITC, 2010.

[2] R. AGUAS, G. Gonçalves, and M.G.M. Gomes. Pertussis: increasing disease as a consequence of reducing transmission. *Lancet. Infectious diseases*, 6(2):112–117, 2006.

[3] Robert Axelrod. Advancing the art of simulation in the social sciences, 1997.

[4] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, and A. Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484, 2009.

[5] S. Bandini, S. Manzoni, and G. Vizzari. Agent based modeling and simulation: an informatics perspective. *Journal of Artificial Societies and Social Simulation*, 12(4):4, 2009.

[6] B. Bauer and J. Odell. UML 2.0 and agents: how to build agent-based systems with the new UML standard. *Engineering applications of artificial intelligence*, 18(2):141–157, 2005.

[7] L. Bian and D. Liebner. A network model for dispersion of communicable diseases. *Transactions in GIS*, 11:155–173, 2007.

[8] K.M. Bisgard, F.B. Pascual, K.R. Ehresmann, C.A. Miller, C. Cianfrini, C.E. Jennings, C.A. Rebmann, J. Gabel, S.L. Schauer, and S.M. Lett. Infant pertussis: who was the source? *The Pediatric infectious disease journal*, 23(11):985, 2004.

[9] S. Black. Epidemiology of pertussis. *The Pediatric infectious disease journal*, 16(4):S85, 1997.

[10] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280, 2002.

[11] K.H. Brassel, M. M
"ohring, E. Schumacher, and K.G. Troitzsch. Can agents cover all the world? *LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS*, pages 55–72, 1997.

[12] H. Broutin, E. Elguero, F. Simondon, and J. Guégan. Spatial dynamics of pertussis in a small region of Senegal. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1553):2091, 2004.

[13] A.W. Brown. Model driven architecture: Principles and practice. *Software and Systems Modeling*, 3(4):314–327, 2004.

[14] C.J.E. Castle and A.T. Crooks. Principles and concepts of agent-based modelling for developing geospatial simulations. *Centre for Advanced Spatial Analysis (University College London): Working Paper*, 110, 2006.

[15] L.C. Chen, B. Kaminsky, T. Tummino, K.M. Carley, E. Casman, D. Fridsma, and A. Yahja. Aligning simulation models of smallpox outbreaks. *Intelligence and Security Informatics*, pages 1–16, 2004.

[16] A. Cliff and P. Haggett. Time, travel and infection. *British medical bulletin*, 69(1):87, 2004.

[17] A. Cliff, P. Haggett, J.K. Ord, and G.R. Versey. *Spatial Diffusion*. Press Syndicate of the University of Cambridge, 1981.

[18] V. Colizza, M. Barthélemy, A. Barrat, and A. Vespignani. Epidemic modeling in complex realities. *Comptes Rendus Biologies*, 330(4):364–374, 2007.

[19] Chris. Collinge, Sharon K.. Ray. Disease ecology: community structure and pathogen dynamics, 2006. page 153-164.

[20] A.T. Crooks. The repast simulation/modelling system for geospatial simulation. In *Agent-Based Models for Spatial Systems in Social Sciences & Economic Science with Heterogeneous Interacting Agents (ABM-S4-ESHIA) workshop in Agelonde, La Londe les Maures (France)*. Citeseer, 2007.

[21] L. Danon, T. House, and M.J. Keeling. The role of routine versus random movements on the spread of disease in Great Britain. *Epidemics*, 1(4):250–258, 2009.

[22] HE De Melker, MA Conyn-van Spaendonck, HC Rümke, JK Van Wijngaarden, FR Mooi, and JF Schellekens. Pertussis in The Netherlands: an outbreak despite high levels of immunization with whole-cell vaccine. *Emerging Infectious Diseases*, 3(2):175, 1997.

[23] R. Depke, R. Heckel, and J.M. K
"uster. Formal agent-oriented modeling with UML and graph transformation* 1. *Science of Computer Programming*, 44(2):229–252, 2002.

[24] R.E. Dickinson. The geography of commuting: the Netherlands and Belgium. *Geographical Review*, 47(4):521–538, 1957.

[25] K. Ellegård. A time-geographical approach to the study of everyday life of individuals–a challenge of complexity. *GeoJournal*, 48(3):167–175, 1999.

[26] J.M. Epstein. Agent-based computational models and generative social science. *Complexity*, 4(5):41–60, 1999.

[27] euregio. Das interaktive zukuuttsportal der euregio mass- rhein. URL:`http://translate.google.com/translate?hl=en&sl=de&u=http://www.euregio-mr.org/&ei=dmTATIzfCMPpOfaAjdkL&sa=X&oi=translate&ct=result&resnum=1&ved=0CBoQ7gEwAA&prev=`

/search%3Fq%3DEuregio%2BMaas-Rijn%26hl%3Den%26prmd%3Div. Access Date 21-10-2010.

[28] Sugi Felix. Analysis of an existing agent-based simulation model for the spread of pertussis(disease:whooping cough) in the enschede area. 2010.

[29] H.J. Foxwell. Java 2 software development kit. *Linux Journal*, 1999(66es):30, 1999.

[30] S. Franklin and A. Graesser. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. *Intelligent Agents III Agent Theories, Architectures, and Languages*, pages 21–35, 1997.

[31] H. Furuya. Risk of transmission of airborne infection during train commute based on mathematical model. *Environmental Health and Preventive Medicine*, 12(2):78–83, 2007.

[32] F. Gargiulo, S. Ternes, S. Huet, and G. Deffuant. An iterative approach for generating statistically realistic populations of households. *PloS one*, 5(1):95–184, 2010.

[33] T.C. Germann, K. Kadau, I.M. Longini, and C.A. Macken. Mitigation strategies for pandemic influenza in the United States. In *vol. 103 no. 15 page 5935Ű5940*. National Acad Sciences, 2006.

[34] G.N. Gilbert and K.G. Troitzsch. *Simulation for the social scientist*. Open Univ Pr, 2005.

[35] N. Gilbert and P. Terna. How to build and use agent-based models in social science. *Mind & Society*, 1(1):57–72, 2000.

[36] J. Gomez-Sanz and J. Pavon. Agent oriented software engineering with INGENIAS. In *Proceedings of the 3rd Central and Eastern Europe Conference on Multiagent Systems, Springer Verlag, LNCS*, volume 2691, pages 394–403. Citeseer, 2003.

[37] S.C. de Greeff, J.F.P. Schellekens, F. R. Mooi, and H.E. de Melker. resurgence of pertussis outbreaks that have affected many countries including developed countries with high-vaccination coverage. Technical report, RIVM report, 2002.

[38] B. Grenfell. Cities and villages: infection hierarchies in a measles metapopulation. *Ecology letters*, 1(1):63–70, 1998.

[39] B. Grenfell and J. Harwood. (Meta) population dynamics of infectious diseases. *Trends in Ecology & Evolution*, 12(10):395–399, 1997.

[40] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S.K. Heinz, G. Huse, et al. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, 2006.

[41] Q. He and J. Mertsola. Factors contributing to pertussis resurgence. *Future Microbiology*, 3(3):329–339, 2008.

[42] C.H. Hennekens, J.E. Buring, and S.L. Mayrent. *Epidemiology in medicine*. Lippincott Williams & Wilkins, 1987.

[43] J. Hoey. Pertussis in adults. *Canadian Medical Association Journal*, 168(4):453, 2003.

[44] S. Holzhauer. *Developing a Social Network Analysis and Visualization Module for Repast Models*. kassel university press GmbH.

[45] Z. Huang and P. Williamson. A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Technical report, Working Paper 2001/2, Population Microdata Unit, Department of Geography, University of Liverpool, Liverpool L69, 2001.

[46] T. Iba, Y. Matsuzawa, and N. Aoyama. From conceptual models to simulation models: Model driven development of agent-based simulations. In *9th Workshop on Economics and Heterogeneous Interacting Agents*. Citeseer, 2004.

[47] A.S. Klovdahl. Social networks and the spread of infectious diseases: the AIDS example. *Social Science & Medicine*, 21(11):1203–1216, 1985.

[48] C.M. Macal and M.J. North. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3):151–162, 2010.

[49] A. Mangili and M.A. Gendreau. Transmission of infectious diseases during commercial air travel. *The Lancet*, 365(9463):989–996, 2005.

[50] L. Mao and L. Bian. Spatial-temporal transmission of influenza and its health risks in an urbanized area. *Computers, Environment and Urban Systems*, pages 204 – 215, 2010.

[51] M. Moore, S.E. Valway, W. Ihle, and I.M. Onorato. A train passenger with pulmonary tuberculosis: evidence of limited transmission during travel. *Clinical Infectious Diseases*, pages 52–56, 1999.

[52] S.S. Morse. Factors in the emergence of infectious diseases. *Emerging Infectious Diseases*, 1(1):7, 1995.

[53] Netlogo. URL:`http://ccl.northwestern.edu/netlogo/`, 2010.

[54] M.J. North, N.T. Collier, and J.R. Vos. Experiences creating three implementations of the repast agent modeling toolkit. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 16(1):1–25, 2006.

[55] NV Nederlandse Spoorwegen. Journey planner. URL:`http://www.ns.nl/`,, 2008.

[56] Y. Ohkusa and T. Sugawara. Application of an individual-based model with real data for transportation mode and location to pandemic influenza. *Journal of Infection and Chemotherapy*, 13(6):380–389, 2007.

[57] J. Pavón, J. Gómez-Sanz, and R. Fuentes. Model driven development of multi-agent systems. In *Model Driven Architecture–Foundations and Applications*, pages 284–298. Springer, 2006.

[58] RG Pebody, NJ Gay, A. Giammanco, S. Baron, J. Schellekens, A. Tischer, R.M. OeLANDER, NJ Andrews, WJ Edmunds, H. Lecoeur, et al. The seroepidemiology of Bordetella pertussis infection in Western Europe. *Epidemiology and infection*, 133(01):159–171, 2004.

[59] R.M. Pendyala, T. Yamamoto, and R. Kitamura. On the formulation of time-space prisms to model constraints on personal activity-travel engagement. *Transportation*, 29(1):73–94, 2002.

[60] LF Perrone, FP Wieland, J. Liu, BG Lawson, DM Nicol, and RM Fujimoto. TUTORIAL ON AGENT-BASED MODELING AND SIMULATION PART 2: HOW TO MODEL WITH AGENTS.

[61] regio twente. Network city twenty. URL:`http://translate.google.com/translate?js=y&prev=_t&hl=en&ie=UTF-8&layout=1&eotf=1&u=http%3A%2F%2Fwww.regiotwente.nl%2F&sl=nl&tl=en`,. Access Date 7-7-2010.

[62] P. Robillard and N.F. Stewart. Iterative numerical methods for trip distribution problems. *Transportation Research*, 8(6):575–582, 1974.

[63] J. Ryan, H. Maoh, and P. Kanaroglou. Population Synthesis: Comparing the Major Techniques Using a Small, Complete Population of Firms. *Geographical Analysis*, 41(2):181–203, 2009.

[64] A. Schaerstrom. Disease diffusion. *International Encyclopedia of Human Geography,Elsevier*, pages 222–233, 2009,.

[65] T. Schwanen, F.M. Dieleman, and M. Dijst. Travel behaviour in Dutch monocentric and policentric urban systems. *Journal of Transport Geography*, 9(3):173–186, 2001.

[66] T. Schwanen, F.M. Dieleman, and M. Dijst. The impact of metropolitan structure on commute behavior in the Netherlands: a multilevel approach. *Growth and Change*, 35(3):304–333, 2004.

[67] T. Schwanen and M. Dijst. Time windows in workers' activity patterns: Empirical evidence from the Netherlands. *Transportation*, 30(3):261–283, 2003.

[68] I. Sijgers, M. Hammer, W. Ter Horst, P. Nieuwenhuis, and P. Van Der Sijde. Supporting the contribution of Higher Education Institutes to regional development. 2008.

[69] J. Snow. *On the mode of communication of cholera*. John Churchill, 1855.

[70] Statistics Netherlands. The dutch virtual census of 2001 analysis and methodology. URL:`http://www.cbs.nl/NR/rdonlyres/D1716A60-0D13-4281-BED6-3607514888AD/0/b572001.pdf`, 2001. Access Date 16-10-2010.

[71] Statistics Netherlands. the netherlands in figures. URL:`http://www.cbs.nl/NR/rdonlyres/1351F633-78C3-4339-AAA5-D377A8A6C8D7/0/thenetherlandsinfigures.pdf`, 2004. Access Date 16/10/2010.

[72] Statistics Netherlands. Cross-border commuting. URL:http://www.cbs.nl/en-GB/menu/themas/arbeid-sociale-zekerheid/publicaties/artikelen/archief/2009/2009-2996-wm.htm,, 2009.

[73] Statistics Netherlands. Population and households, four-digit code, january 1, 2009. URL:http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=80280ned&D1=0-62,66-72&D2=2770-2780,2788,4034,4267&VW=T, 2009.

[74] Statistics Netherlands. Size and composition household, position in the household, 1 january. URL:http://statline.cbs.nl/StatWeb/publication/?DM=SLEN&PA=37312eng&D1=0-69&D2=a,!1-4,!6-7&LA=EN&VW=T, 2010.

[75] Y.O. Susilo and K. Maat. The influence of built environment to the trends in commuting journeys in the Netherlands. *Transportation*, 34(5):589–609, 2007.

[76] T. Tan, E. Trindade, and D. Skowronski. Epidemiology of pertussis. *The Pediatric infectious disease journal*, 24(5):S10, 2005.

[77] W. Tang. Development of a Spatially-explicit Agent-based Simulation Package for Modeling Complex Adaptive Geographic Systems. *Proceedings of the UCGIS Summer Assembly*, 2007.

[78] H. Trottier and P. Philippe. Deterministic modeling of infectious diseases: theory and methods. *The Internet Journal of Infectious Diseases*, 1(2), 2001.

[79] L. Van der Laan. Changing urban systems: an empirical analysis at two spatial levels. *Regional studies*, 32(3):235–247, 1998.

[80] C. Viboud, O.N. Bjornstad, D.L. Smith, L. Simonsen, M.A. Miller, and B.T. Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312(5772):447, 2006.

[81] Rijksinstituut voor Volksgezondheid en Milieu (RIVM). About CIb. URL:http://www.rivm.nl/en/infectious-diseases/, November 2009.

[82] Wikimedia Foundation Inc. Public transport. URL:http://en.wikipedia.org/wiki/Public_transport,, 1 2011. Assess Date 12/1/2011.

[83] World Health Organization. Weekly epidemiological record. URL:http://www.who.int/immunization/topics/wer8004pertussis_Jan_2005.pdf,, 2005. Access Date 3-7-2010.

[84] World Health Organization. Global health risk mortality & burdern of disease attributable to selected risk. URL:http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf, 2010. Access date: 3-7-2010.

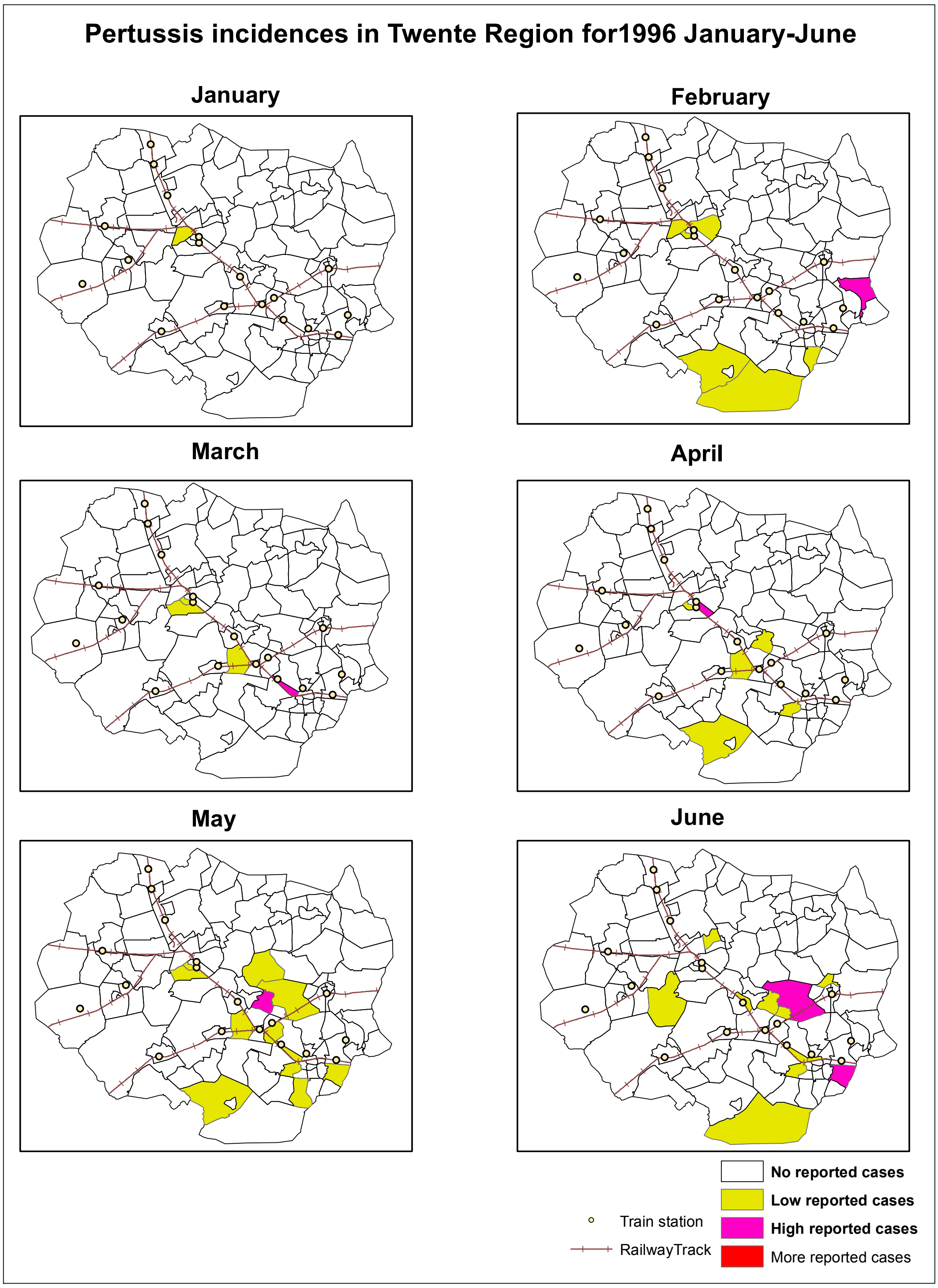
[85] E. Yaari, Y. Yafe-Zimerman, S.B. Schwartz, P.E. Slater, P. Shvartzman, N. Andoren, D. Branski, and E. Kerem. Clinical Manifestations of Bordetella pertussis Infection in Immunized Children and Young Adults*. *Chest*, 115(5):1254, 1999.

[86] Y. Yang and PM Atkinson. An Integrated ABM and GIS Model of Infectious Disease Transmission. *Computers in Urban Planning and Urban Management–CUPUM'05; 29 June–1 July; London, England*, 2005.

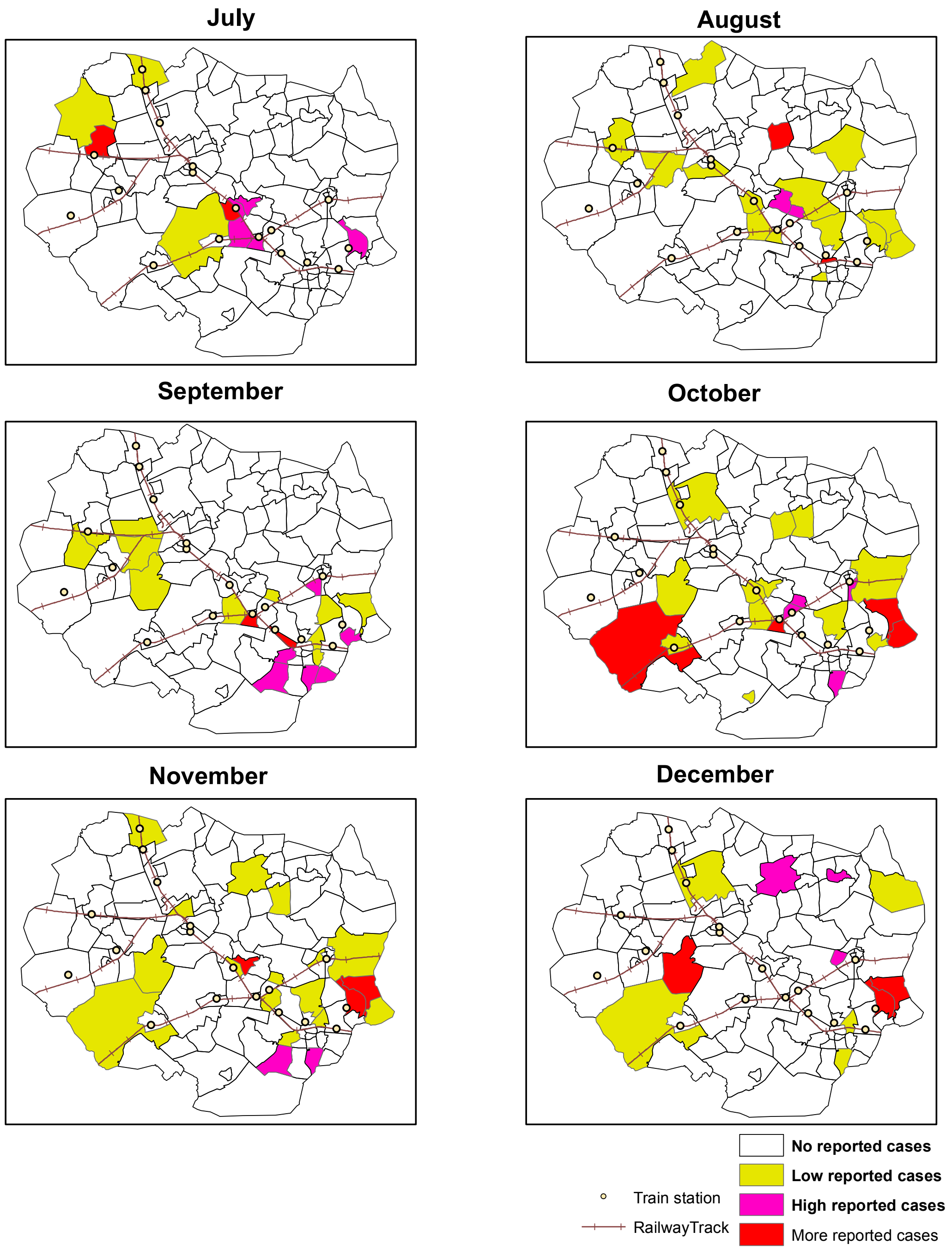
[87] Y. Yang and P.M. Atkinson. Individual space -time activity-based model: a model for the simulation of airborne infectious-disease transmission by activity-bundle simulation. *Environment and Planning B: Planning and Design*, 2007.

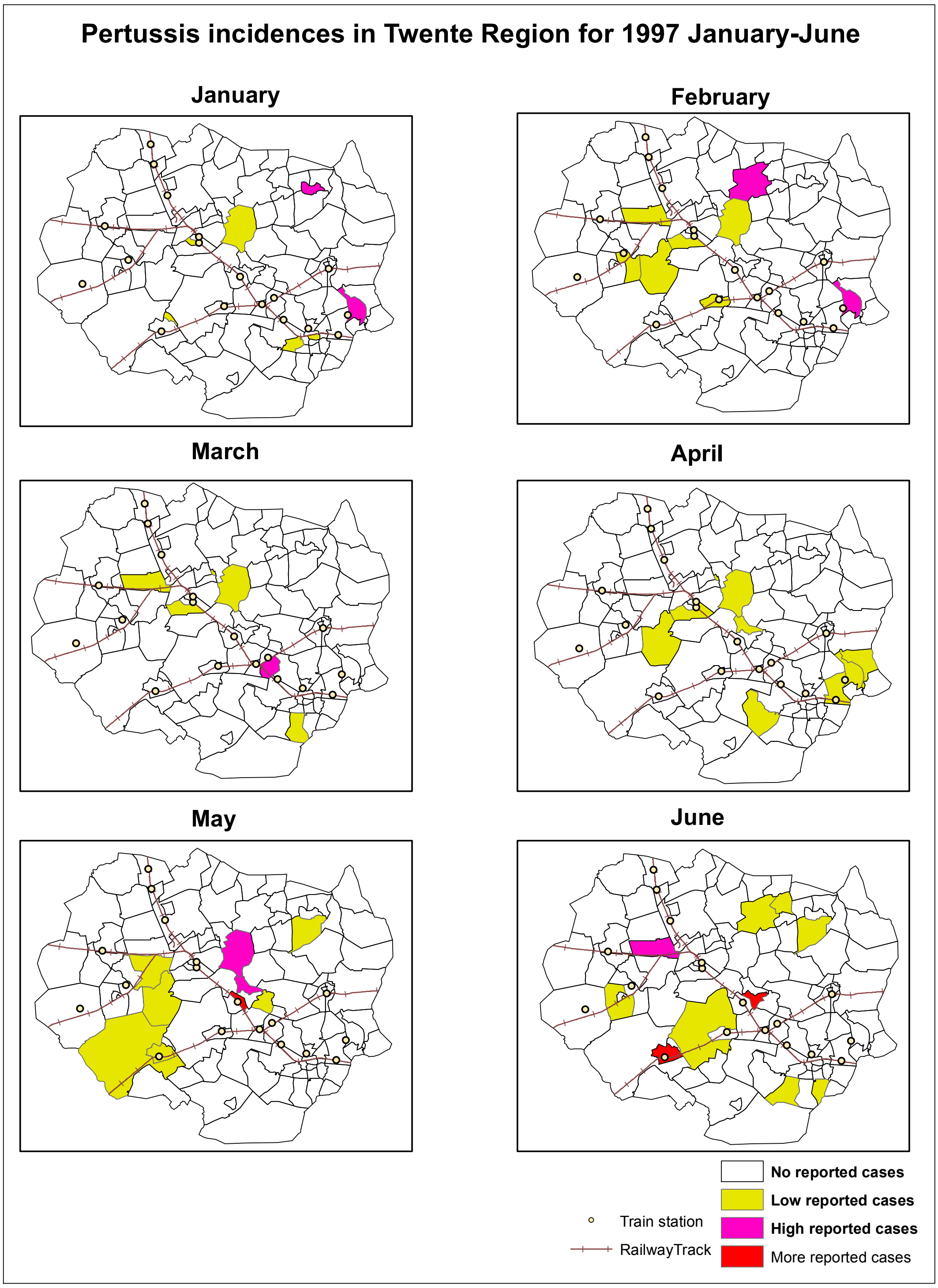
[88] Y. Yang and P.M. Atkinson. Individual space- time activity-based model: a model for the simulation of airborne infectious-disease transmission by activity-bundle simulation. *Environment and Planning B: Planning and Design*, 35:80–99, 2008.

Spatial distribution of pertussis incidences in Twente region from 1996-2004

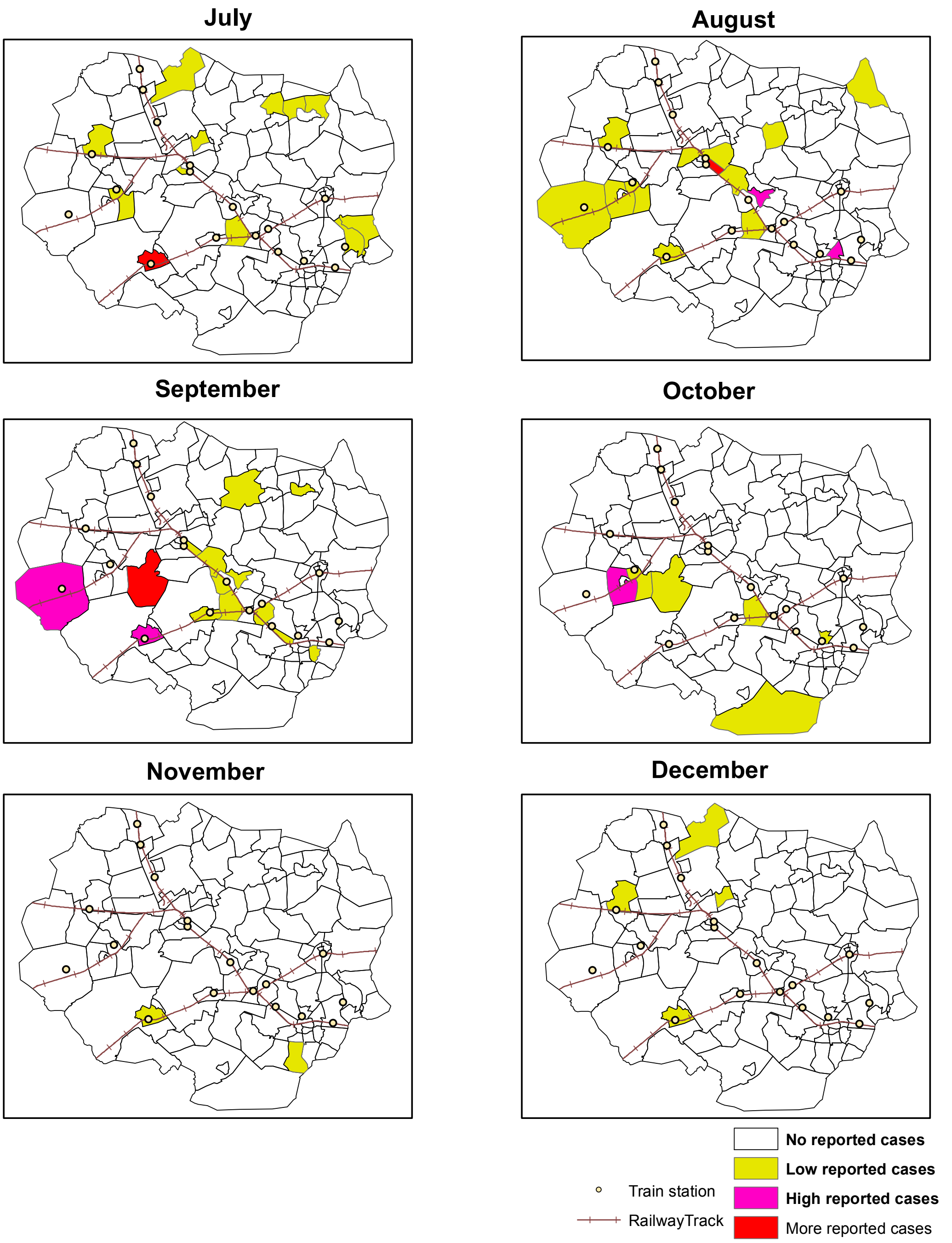

Pertussis incidences in Twente Region for1996 January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

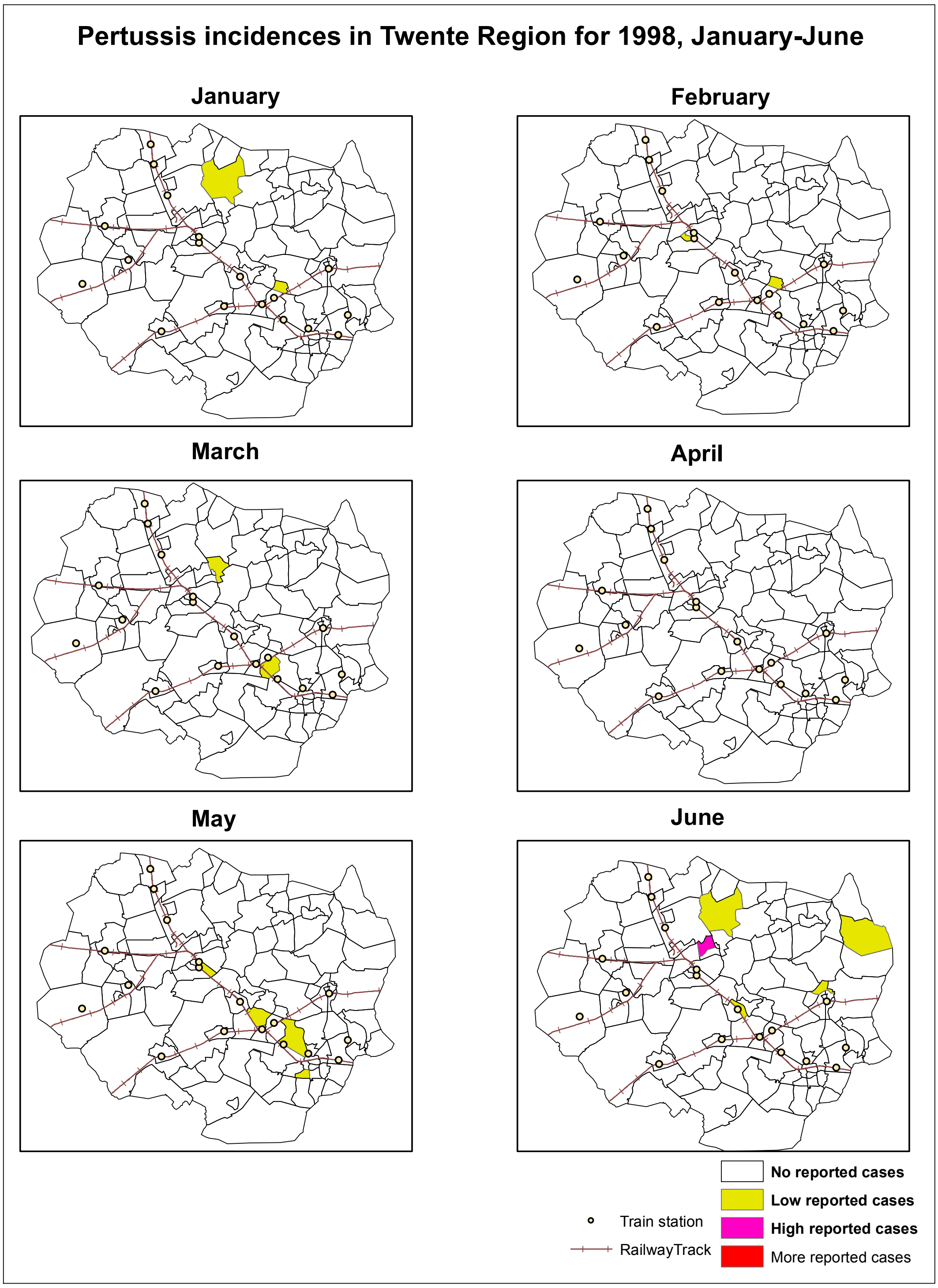
Pertussis incidences in Twente Region for1996 July-December

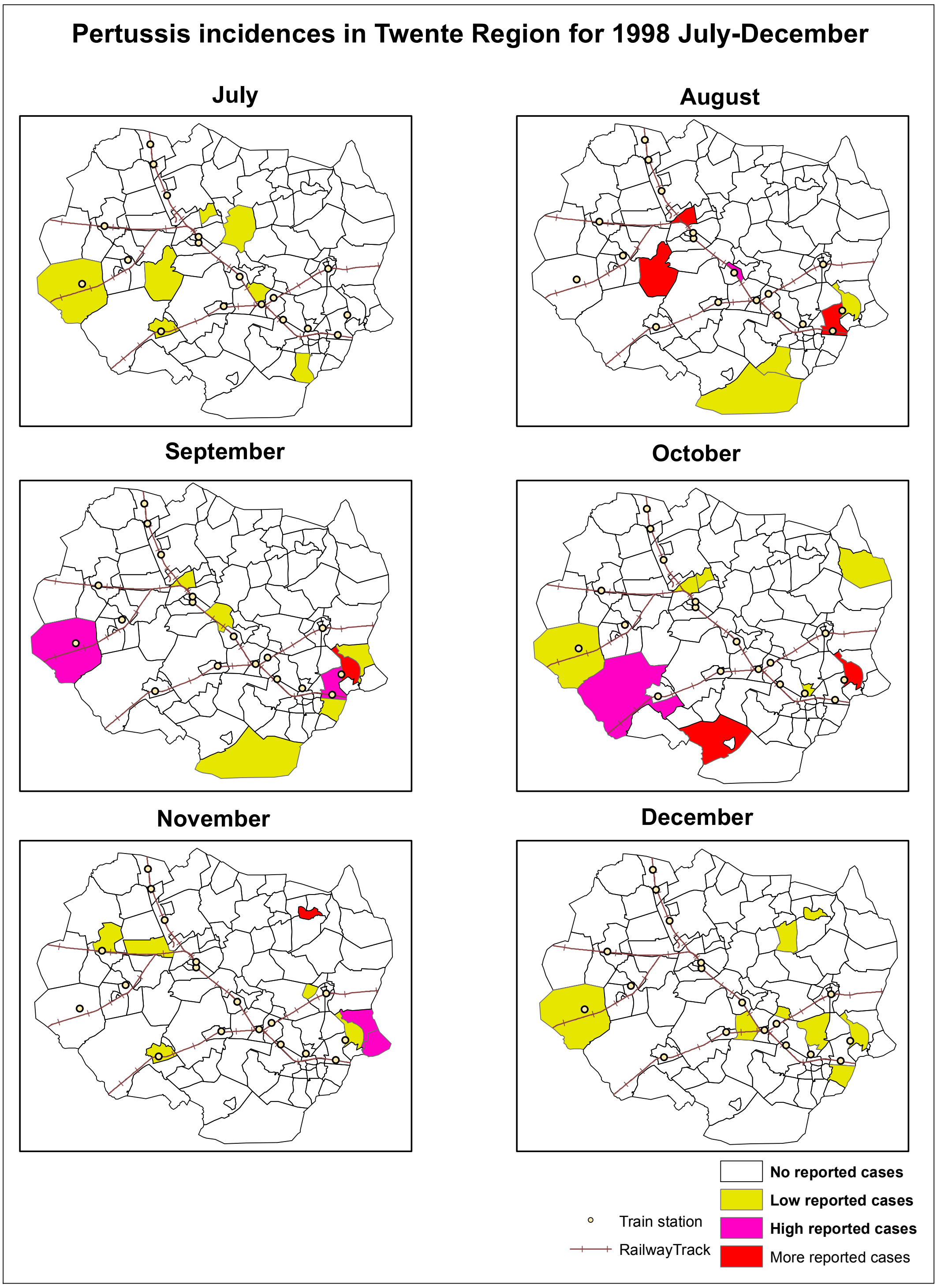


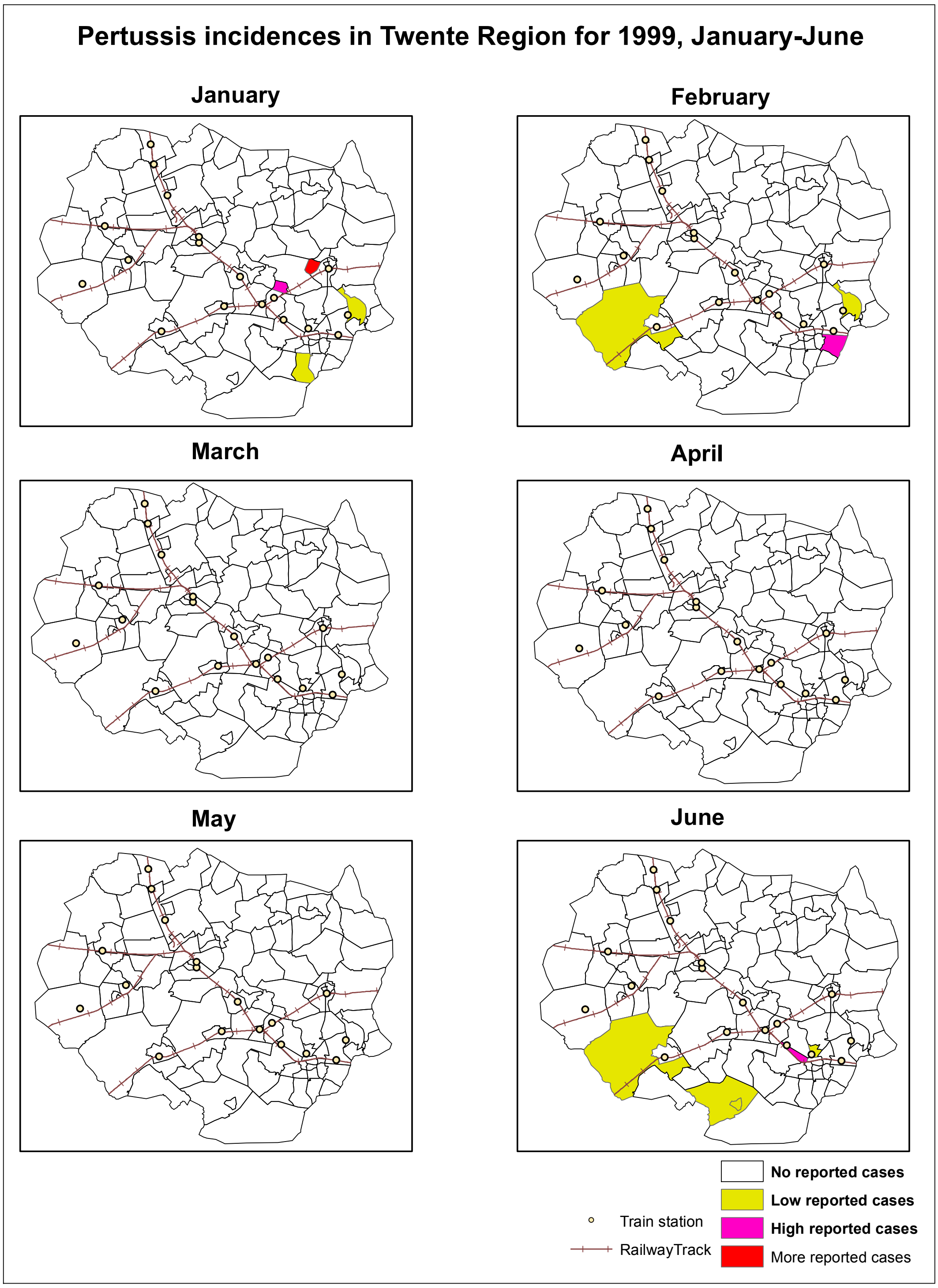

Pertussis incidences in Twente Region for 1997 January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

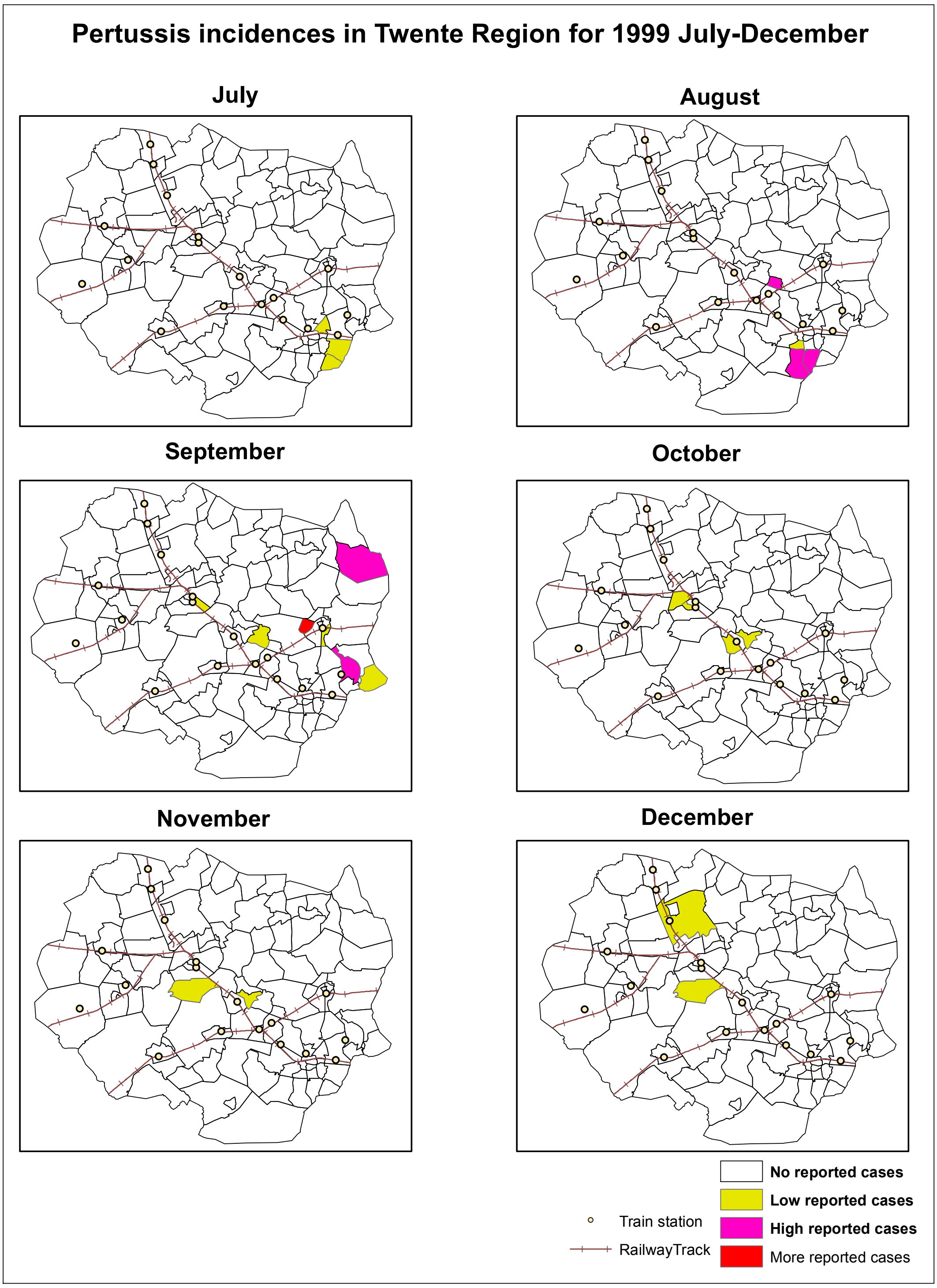
Pertussis incidences in Twente Region for 1997 July-December

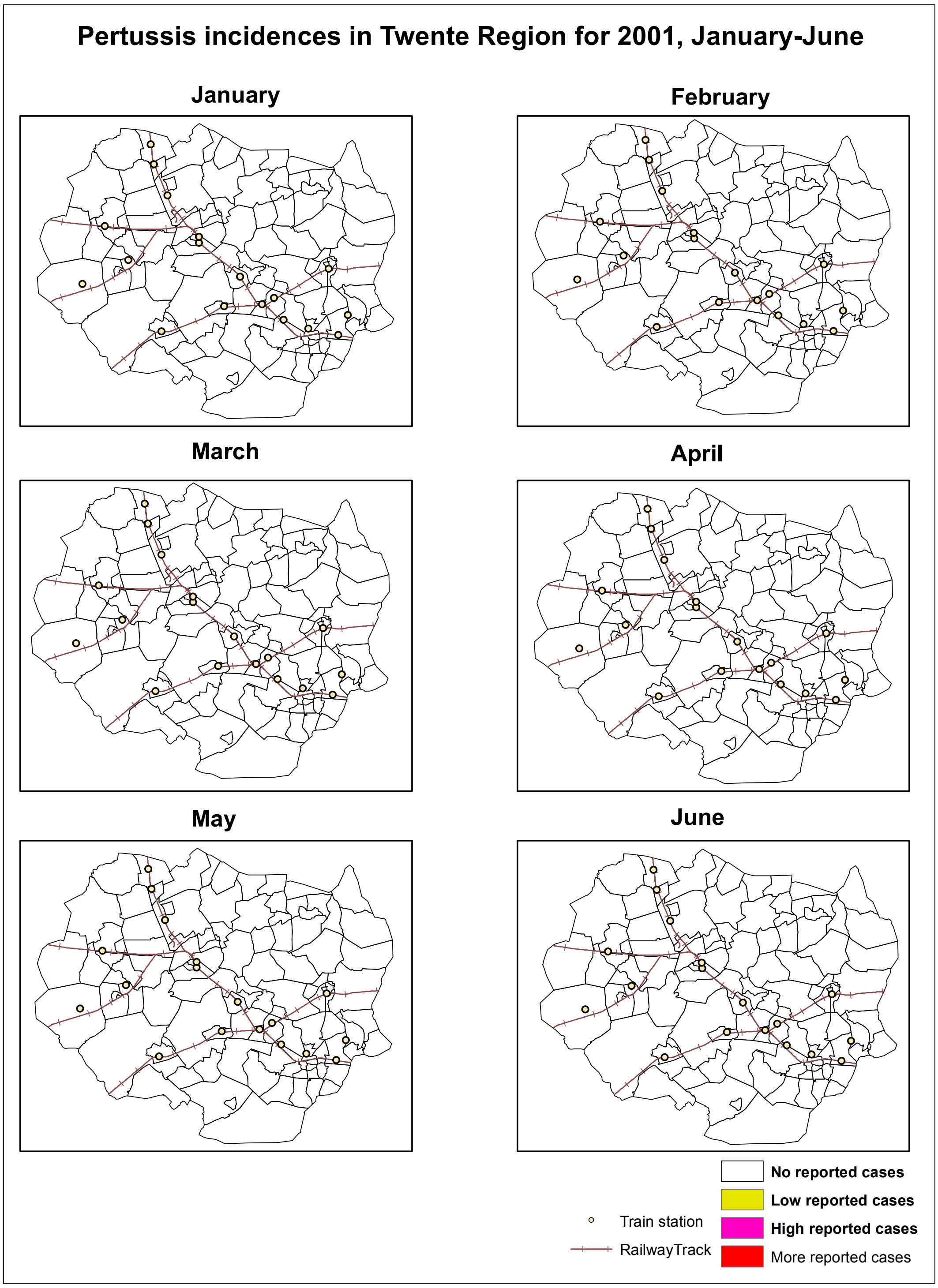



Pertussis incidences in Twente Region for 1998, January-June
January
February
March
April
May
June
Train station
RailwayTrack
No reported cases
Low reported cases
High reported cases
More reported cases

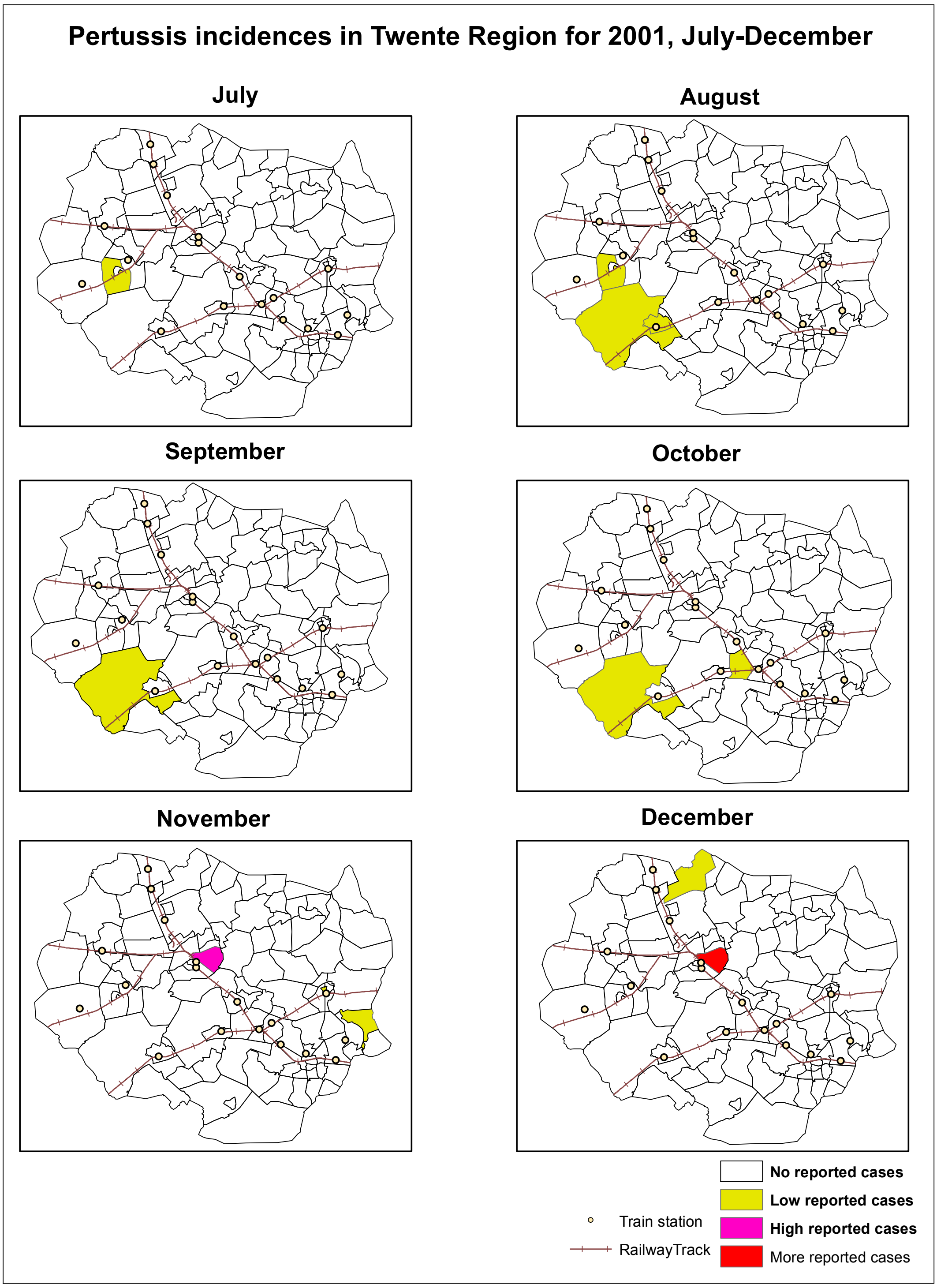

Pertussis incidences in Twente Region for 1998 July-December
July
August
September
October
November
December
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

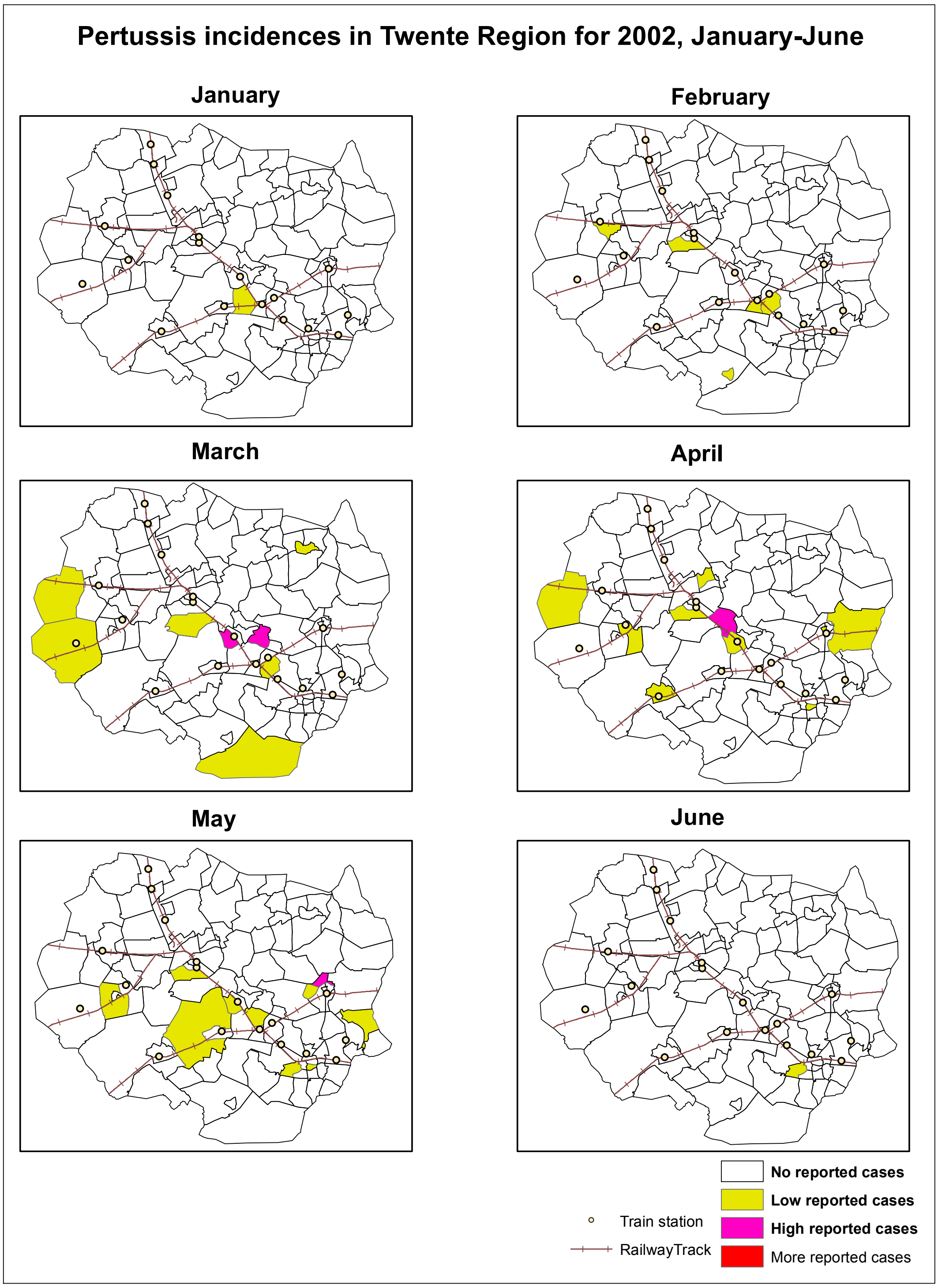

Pertussis incidences in Twente Region for 1999, January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack


Pertussis incidences in Twente Region for 1999 July-December
July
August
September
October
November
December
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

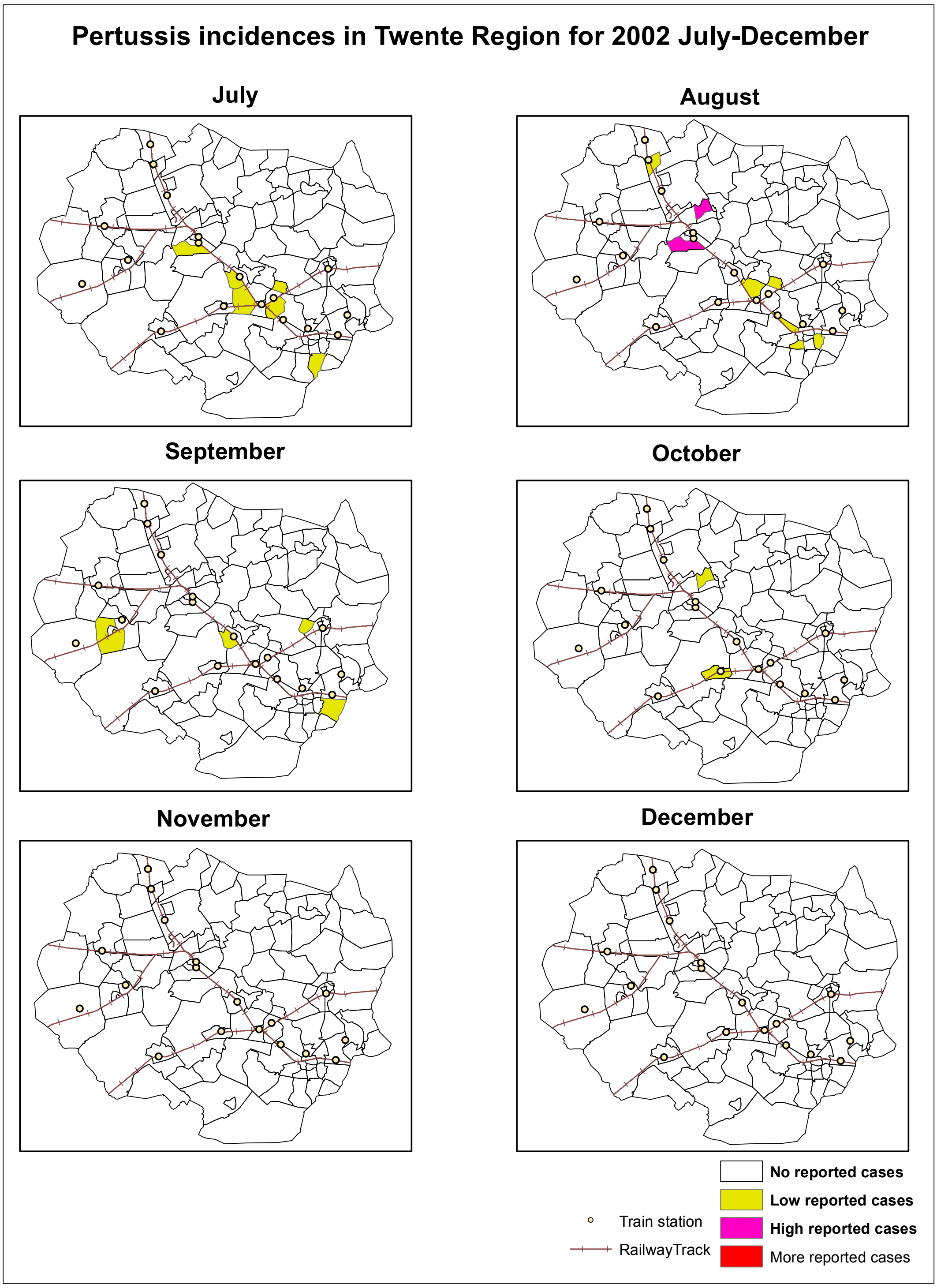


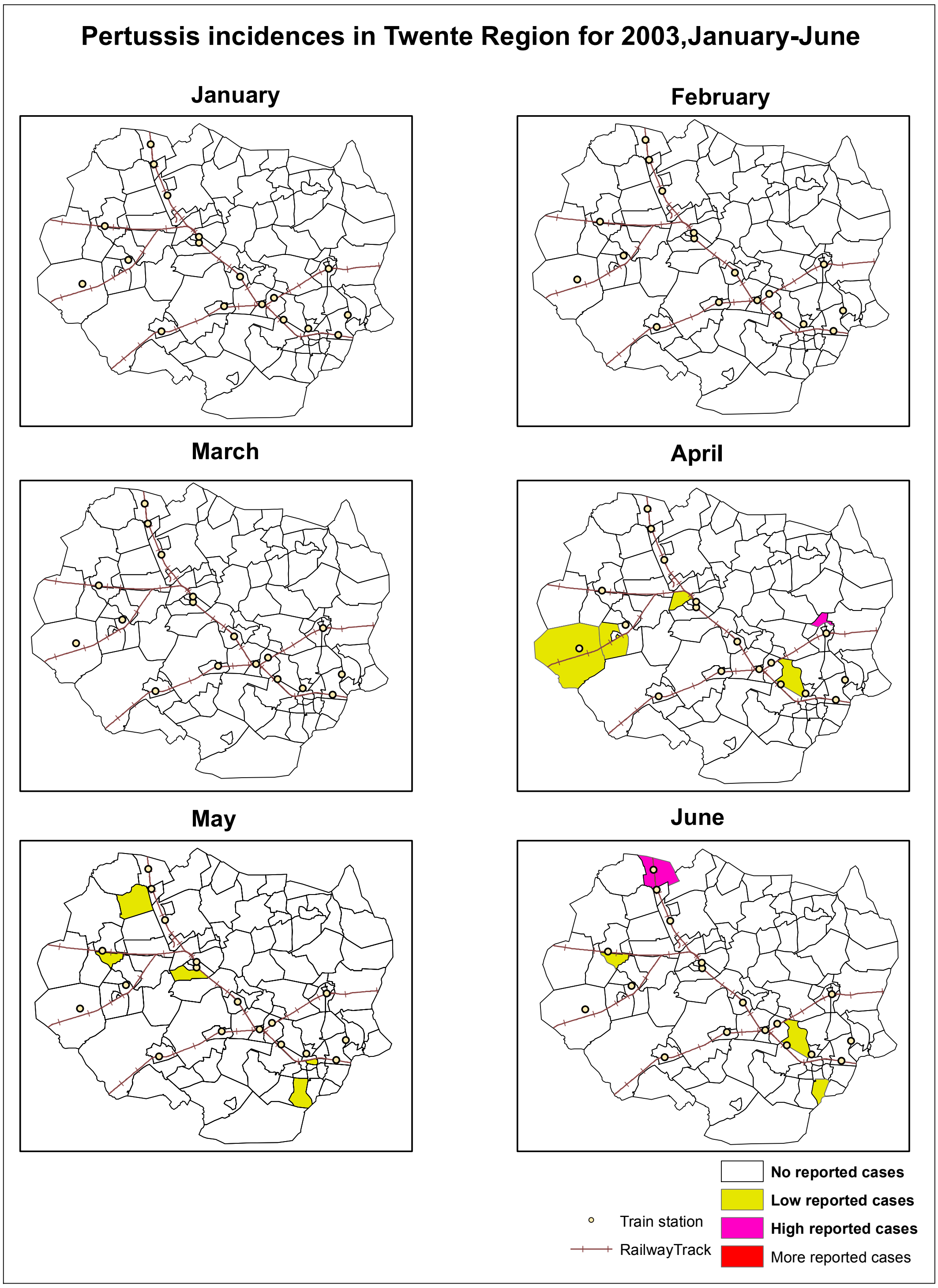
Pertussis incidences in Twente Region for 2001, January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack


Pertussis incidences in Twente Region for 2001, July-December
July
August
September
October
November
December
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

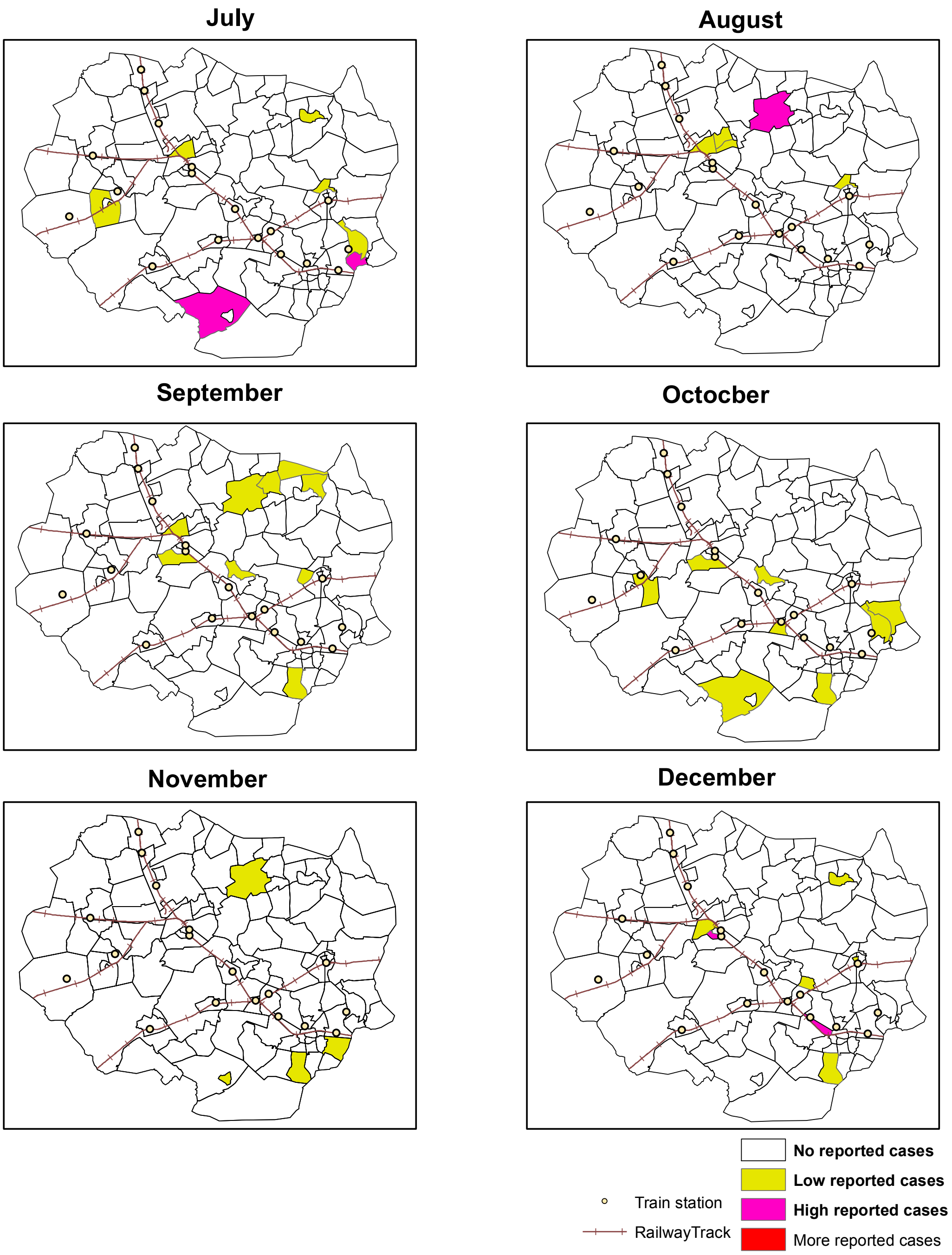


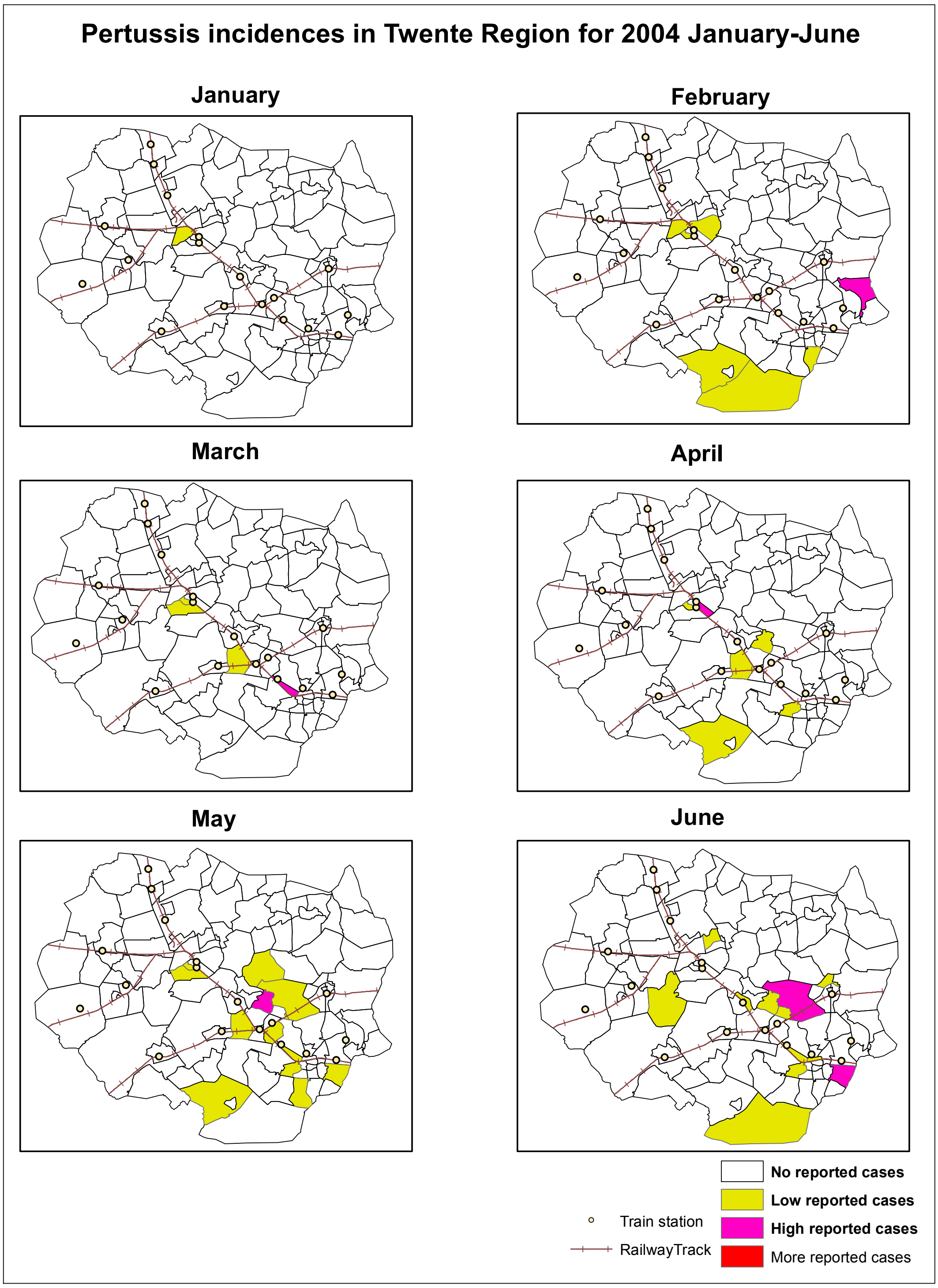
Pertussis incidences in Twente Region for 2002, January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

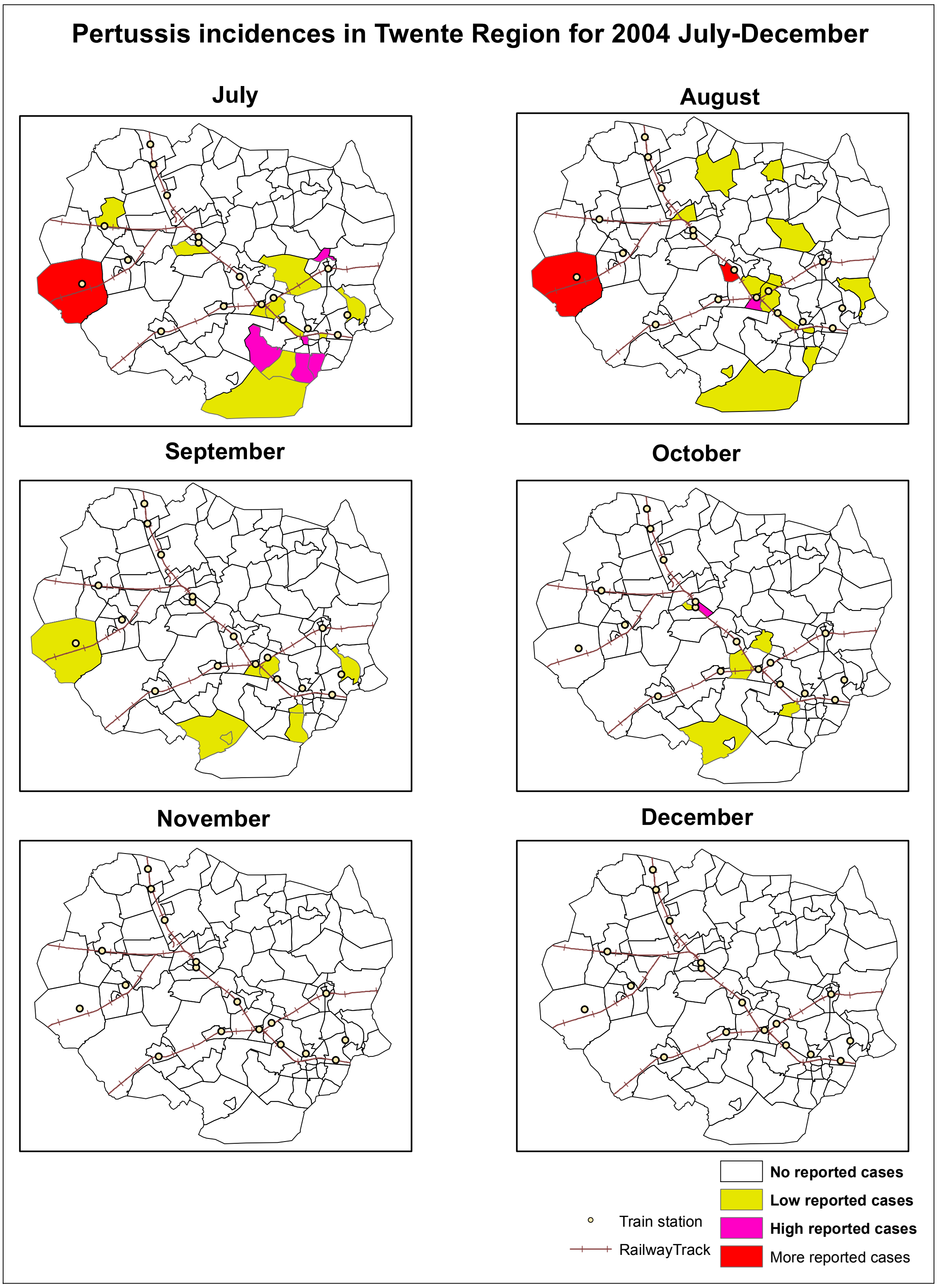

Pertussis incidences in Twente Region for 2002 July-December
July
August
September
October
November
December
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack


Pertussis incidences in Twente Region for 2003,January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack

Pertussis incidences in Twente Region for 2003,July-December




Pertussis incidences in Twente Region for 2004 January-June
January
February
March
April
May
June
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack


Pertussis incidences in Twente Region for 2004 July-December
July
August
September
October
November
December
No reported cases
Low reported cases
High reported cases
More reported cases
Train station
RailwayTrack