# Using agent-based modelling for small area household projections: a New Zealand case study

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

**Masters of Social Science**

at

**The University of Waikato**

by

**Anton Marais**

THE UNIVERSITY OF
WAIKATO
*Te Whare Wānanga o Waikato*

2016

# Abstract

In New Zealand, the Resource Management Act 1991 (RMA) is the guiding legislation that sets out how the environment is managed. The territorial authorities are charged with regulating how the environment is managed in their jurisdiction. In order to effectively manage any resource, good and robust planning processes are required. Population change is the factor that has the greatest impact on the environment, and is one of the most challenging to regulate, thus for the territorial authority planners, knowing where and when infrastructure investment is most likely required is key to fulfilling their statutory requirement.

This line of investigation has been driven by an ever increasing need for more spatially detailed projections of where and when changes are most likely to take place. The investigation and subsequent development of spatially detailed household projections transcends three territorial authorities in the lower Waikato river catchment situated in New Zealand.

An agent-based model was developed in which the individual agents were households. The aim of the model was to produce a simplified simulation of the location choices made by the members of a household in a housing market that is governed by councils' infrastructure provision, household rents, transport costs and the benefit derived from the neighbourhood amenities and the environment. This model simulates the distribution of households over a twelve year period from 2013 to 2025. In the model the households were programmed to move to vacant properties in order to minimise their residential costs. Each time an agent moves this provides new potential options for all other households; thus the simulation runs and the households move until all households settle in their least cost locations, representing the distribution of households in 2025. The incorporation of multiple territorial authorities provides a more holistic approach than the prevailing approach which is based on disaggregating top level projection with the no further account of population movement outside the top-level migration assumption.

The results of the model calibration indicate the model performs well with a 16.6% RMSE at the smallest spatial unit. The projected results produce growth patterns that fall within the expectations of planning staff of the

councils. These staff members have affirmed the model's and input assumptions, and indicate the outcomes to be both useful and plausible.

The results demonstrate the relationship between the councils and their respective growth plan strategies. The scenarios were developed to, demonstrate the relationship between the availability of vacant land and the cost to occupy a property, which ultimately impacts the flux of residents to or away from the city.

With some broad assumptions and limitations, this model is distinctive in its approach as it is developed with the intention of being an applied tool to be used by the three councils. It has the further distinction as modelling the behaviours of individual households rather than the behaviours of an entire population, which is the unique in New Zealand and amongst a few applied models of this nature wide (Triantakonstantis & Mountrakis, 2012).

# Acknowledgment

My interest in this topic started sometime before I embarked on this research. I was involved with planning for future growth at Waikato District Council and became embroiled in a case that was heard in the environment court. Working with Michael Cameron, William Cochrane and Matthew Roskruge inspired me to start thinking about better ways look at population growth and more specifically population distribution patterns. Conversation and debate with Michael and his colleagues inspired me to think about these issues from a geospatial point of view and I thank them for encouraging me to elaborate my thinking.

I would specifically like to thank my supervisor Dr Michael Cameron for the inspiration he provided, the platform for exploring ideas and all round positive support for this work. I am also grateful for the opportunities provide to present this work in the international arena. I would also like to extend thanks and gratitude to Professor Natalie Jackson for assistance and guidance in completing other components of this masters and support in the initial stages of this work.

Funding a substantial research project can be a significant obstacle, particularly at a personal level, thus it is with much appreciation that I thank the senior management at Waikato District Council for providing funding through a personal development award.

I would like to thank Jonathan Hall for a huge amount of moral support, ensuring the work was ticking over with regular catch-ups and progress reports. At Waikato District Council, Julian Hudson assisted with a lot of coding and debugging, and answered hundreds of question that helped me better understand more than just what the code instructions were doing but more about what goes on in the software. Julian provided time, energy and an amazingly positive attitude that helped me turn the concept into a working model.

I recognise and appreciate the contributions of planning staff at the Waikato, Waipa and Hamilton councils, in particular, Vishal Ramduny, Gary Knighton and Michael Spurr. The councils have made invaluable data freely available and answered a number of questions. Thanks also extend to Statistics New

Zealand for their open data policies and well-organised data catalogues and download services, without which work like this would not take place.

I also appreciate the data supplied by Dr Garry McDonald of Market Economics, James Bevan of Latitude Planning, and Dr William Cochrane, the late James Newell.

By convention, last on this list but still the most important are the support, encouragement and occasional ‘hurry up or else threat’ that I have received from my wife, Regina and daughter Gabrielle. I would not have been able to complete (or probably even start) this work without their support. This work has taken a long time to complete and I thank them for the time and activities that they have given up in supporting my endeavour.

# Table of Contents

# List of Figures

# List of Tables

# List of Maps

# 1 Introduction

Population and household projections are a key component of planning for the future and managing the environment. In New Zealand, the Local Government Act 2002 legislates local authorities play a broad role in meeting the current and future needs of their communities for good-quality local infrastructure, local public services, and performance of regulatory functions (Department of Internal Affairs, 2013). To effectively fulfil these legislative requirements, the local authorities require effective methods to project both household and population growth at the local level. Worldwide the need for local level population projections and distribution is recognised by a number of authors (Cai, 2007; Jenner, 2002; Rayer & Smith, 2010; Rees et al., 2004; S. Smith & Cody, 2013; Yang, 2011; Zhan, Tapia Silva, & Santillana, 2010).

The developments in computing and processing (Benenson, 1998; Santé, García, Miranda, & Crecente, 2010) have resulted in the coupling of demographic modelling and urban growth models (Santé et al., 2010; Torrens & Benenson, 2005). This has resulted in an ever-expanding range of methods to simulate changes or events under a wide range of inputs and variables (Triantakonstantis & Mountrakis, 2012). This research investigates the development of a computer simulation that demonstrates the likely household distribution over time at the neighbourhood level.

A number of authors have investigated the accuracy of subnational population projections (Rayer & Smith, 2010; Rees et al., 2004; Statistics New Zealand, 2008; S. K. Smith, Tayman, & Swanson, 2001). They report decreasing accuracy as the population and geographic areas get smaller. There are two main reasons for potentially high error in small populations and or geographic areas. Migration is often a major determinant of growth and is more variable over space and time (S. K. Smith et al., 2001). In small communities, isolated events can alter the trajectory of growth whereas in large populations the effects of a multitude of events tend to offset each other (S. K. Smith et al., 2001). Consequently, population change is more unpredictable and more variable as population and geographic size become smaller.A number of authors have investigated the accuracy of subnational population projections, (Rayer & Smith, 2010; Rees et al., 2004; Statistics New Zealand, 2008; S. K. Smith, Tayman, & Swanson, 2001). They report decreasing accuracy as the population and geographic areas get smaller. There are two main reasons for potentially high error in small populations and or geographic areas. Migration is often a major determinant of growth and is more variable over space and time (S. K. Smith et al., 2001). In small communities, isolated events can alter the trajectory of growth whereas in large populations the effects of a multitude of events tend to offset each other (S. K. Smith et al., 2001). Consequently, population change is more unpredictable and more variable as population and geographic size become smaller.

The field of urban growth modelling comes into consideration at the sub-regional scale. Various forms of Cellular Automata models are the most frequently used algorithms in urban growth models (Triantakonstantis & Mountrakis, 2012) and have been used in areas as small as neighbourhoods or up to whole countries. Agent-based models are another method that is gaining popularity (Axtell, 2000). In the agent-based model an agent represents an individual entity, such as a person or household, and these entities interact with their environment. They are programmed (Macal & North, 2005) such that their collective behaviours result in a change to the environment(s) and it is these emerging trends (Torrens & Benenson, 2005) that provide the valuable insight(s) to the system. Four agent-based models were reviewed in detail and these provided contextual similarities to this New Zealand study and were used in the formulation of a bespoke agent-based model suited to the study area.

The agent-based model that has been developed, simulates residential households' location choices. The household agents primarily use economic based cost assignments and they exhibit the rational behaviour of minimising the cost associated with their residential location. The modelled environment contains a central city with a number of satellite towns. There are three planning authorities and their respective aspirations have impacts on the overall distribution of new household development.

Although collaborative planning is developing between the Territorial Authorities (TA), the existing methods are based on disaggregation of the population at the TA level. Each TA disaggregates their population projection and plans accordingly. From a planning perspective, this approach is certainly a lot easier as it is a closed system. If the population is projected to grow by a certain amount then calculating how much land needs to be zoned and how much infrastructure is required is more straight-forward. The complication arises when a nearby town also has residential land available, possibly at a lower cost. The residents are not constrained by a TA boundary and they can choose where to relocate, this system is open. Holistic planning is significantly more difficult and requires more sophisticated tools. Models, such as this can help to demonstrate the relative benefit citizens receive from living in different areas.

The results of this work demonstrate that even under broad assumptions that there are strong interdependencies between the city and the different towns and villages in the study area. The effects of one authority's regulations have an impact on their neighbours, and this model can be used to identify some of the key drivers, thresholds and impacts of different scenarios. This model has a strong emphasis on the intended housing density in different areas and how much this capacity is utilised in the model time frames.

The research question is "what will the likely household distribution be for Hamilton, Waikato and Waipa in 2025?" The model functions under two broad factors that can be considered to be known; 1) the population growth for the study area, this area has a large enough population for reliable (M. Cameron, Cochrane, & Poot, 2007) cohort component population projections and is accepted as sufficiently accurate by Statistics New Zealand the Territorial Authorities; and 2) Due to the Resource Management Act 1991, land use (zoning) management and planning are relatively stable, as such the quantity of land available for development. The unknown factors are; 1) across the broader geographic study area, which areas are most likely to attract growth; and 2) in what order will residential land use changes take place across the study area. Up to 2025, it is projected that approximately 31,400 new houses will be required to house a growing population (M. Cameron & Cochrane, 2015b) and the purpose of the model is to determine, under given planning objectives, where these houses are most likely to be constructed.

To determine this the model utilises individual households as the primary agent. Each household is attributed with a unique combination of costs associated with their current residential location and cost to travel to their place of employment. The household agents are programmed to simulate rational behaviour with the objective of moving to a different location if their costs can be minimised. The movement of households to a different location has a localised inflationary effect which creates increasing cost in popular areas and a corresponding decrease in costs at the origin. This results in excessive positive and negative growth being arrested and allows the neighbourhoods to have more cyclical trends.

The results are presented in a series of maps showing the overall changes taking place between 2013 and 2025. The results are presented at the neighbourhood level. The model is tested using two planning scenarios. One scenario represents the current planning objectives in which each of the councils has a development plan scheduled up to and beyond 2025. The other scenario represents the same planning objective, however, all the land is available and there are no timing constraints on land development. The household projections of the model are compared with the nationally prepared Statistics New Zealand and a regionally prepared projection series produced by National Institute of Demographic and Economic Analysis (NIDEA). The outcomes of this model demonstrate the relationship between the availability of housing and the impact this has on the relative housing costs and optimum economic locations for citizens.

# 2 Literature Review

The manner in which urban centres grow (in order to accommodate people and industry), although complex, is a field researched by a wide range of authors. This topic covers research transcending the fields of demography, urban planning, mathematics, statistics, computer science and geospatial science. This breadth has resulted in a wide range of techniques and methods to project populations and likely population change.

In New Zealand, the predominant method used to project populations is the cohort component method (Bascand, 2012; Bell, Blick, Parkyn, Rodway, & Vowles, 2010). The bases of a cohort projection are, a known base population, fertility rates, mortality rates and migration assumptions (Bascand, 2012; Bell et al., 2010). In any case where the starting population is not known, the first steps are to produce a method of estimating a base population. In terms of cohort component methods, once the base population is known the next limitations centre on the availability of other parameters such as age structures and birth, death and migration rates. When a base population is not known then a range of population estimation methods have been devised by various authors and authorities in different disciplines, for example (Baker et al., 2008; Jenner, 2002; Wang & Cardenas, 2011; Zhan et al., 2010).

There is an identified need for monitoring and planning for population change, right down to the community level (Baker et al., 2008; English, 2015). There is, however, a range of methods used to quantify populations in smaller areas (Santé et al., 2010; Triantakonstantis & Mountrakis, 2012).

There are however few methods that are widely accepted or adopted by the planning fraternities (Triantakonstantis & Mountrakis, 2012). New Zealand's Area Unit projections are among the smallest in population and geographic size anywhere in the world.

There are two main elements in quantifying populations, the first is estimating population size and the second is estimating a population's size at some point in the future. From the perspective of developing any model, the important criteria are input data quality and suitability for use in the intended context. Basal information can be gathered in a number of different ways. Primary sources are available from censuses, regulatory authorities or commercial data. Secondary data can be derived from the integration of different datasets. Geographically structured data is particularly useful when integrating and interrogating data from different sources and developing derived datasets. Remote sensing is a particular field used to derive new datasets that can be utilised in estimating population sizes. Lizhong Hua, Wang Man, Qiong Wang, & Xiaofeng Zhao, (2012), Yang (2011) and Zhan (2010) describe methods utilising remote sensing to derive new datasets.

## 2.1 Spatially explicit algorithms used to model population or land use changes

The inhabitants of an area are subject to a range of motivations for changing their residential status or location. Such motivations could be 'life stage' (Fontaine & Rounsevell, 2009; Gaube & Remesch, 2013), cultural (Benenson, 1998), economic (Gaube & Remesch, 2013), interactions with natural environment or proximity to infrastructure (Jjumba & Dragićević, 2012). Various algorithms and methods have been developed to simulate population and urban growth. In a comprehensive review, Triantakonstantis and Mountrakis (2012) investigated the main types of models used in urban growth simulation. Models and methods reviewed included:

- housing unit methods,
- cohort component,
- artificial neural network,
- fractal,
- linear and logistic regression,
- decision tree,
- cellular automata, and

- agent-based models.

Housing Unit methodologies bridge the delineation between demographics and urban planning and predominantly apportion population growth based on changes in numbers of housing units or dwelling structures. Static housing units can be quantified more easily than transitory objects such as people. Population estimates can be made by counting households and applying person per household ratios and adding the known population in institutions (Cai, 2007; S. K. Smith & Cody, 2013). Demographic data on household composition provides a way to estimate the number of people per household (Cai, 2007; S. K. Smith & Cody, 2013). Among others, remote sensing disciplines (Yang, 2011), (Zhan et al., 2010), (Jenner, 2002) (Wu & Martin, 2002), building permits (Roskruge, Cameron, & Cochrane, 2010; S. K. Smith & Cody, 2013) or water connections (Jenner, 2002) provide a range of ways in which housing units numbers can be estimated. These methods can in some cases, be utilised to provide time series housing unit counts. When such data are available, methods to establish trends can be developed, for example, both Roskruge et al. (2010) and S. Smith & Cody, (2013) use building permit trends to project changes in housing unit development.

Housing Unit methodologies highlight the intrinsic relationship between infrastructure and residential growth. Residential growth takes place where infrastructure is both available (M. Cameron & Cochrane, 2015a) and most affordable. However, planners look to projections to identify where to establish infrastructure (M. Cameron & Cochrane, 2015a). Urban planning processes provide the location(s) and capacity for housing and urban modelling techniques provide tools to forecast the location of growth and demographics project how much growth is likely to occur. For example, Waikato District Council's 2012 and 2015 Long Term Plan (LTP) population projections are based on a housing unit model (Roskruge et al., 2010). The Housing unit method is used to establish the number of dwellings in each geographical area. The geographic areas are AUs further decomposed into urban and rural components where some portions of the AU are of urban or rural character. This is a further example of an instance where the census tract does not meet the planner’s requirement and the planner has devised a method of proportioning growth projections. The annual increment in dwelling numbers is exogenously determined for the region as a whole. The number of new households is proportionally distributed across the AU’s based on the number of building consents issued per geographic area. The council has also provided their land and infrastructure plans so the maximum number of households per AU is known. When an AU reaches this limit the growth is proportionally redistributed to the next fastest growing AU. The first projections developed in 2012 were based on the 2006 census data. These projections provided suitable outcomes for 2013 such that the 2015 LTP projections required minor adjustments to update the model using 2013 actual census figures.

The cohort component method (CCM) is internationally recognised (Bascand, 2012) and can be used in both deterministically or probabilistically. The cohort component method has been utilised by the three councils in the study area since 2009, (M. Cameron & Cochrane, 2015b; M. Cameron, Cochrane, & Poot, 2008). The US Bureau succinctly describes the CCM as “In the cohort-component method, the components of population change (fertility, mortality, and net migration) are projected separately for each birth cohort (persons born in a given year). The base population is advanced each year by using projected survival rates and net international migration. Each year, a new birth cohort is added to the population by applying the projected fertility rates to the female population.” (US Census Bureau, 2014).

Artificial Neural Networks (ANN) are algorithms based on networks. A complex system is decomposed into simpler interconnected components, referred to as neurons, as in neural networks (Triantakonstantis & Mountrakis, 2012). Data is input at input neurons (or units as described by (Tayyebi, Pijanowski, & Tayyebi, 2011) and transfers to intermediate neurons, sometimes referred to as hidden neurons, the transition from neuron to neuron is measured as a weight. The output, or response, is dependent on the sum of the weights resulting from each transaction course through the network of neurons (Triantakonstantis & Mountrakis, 2012). “Each input unit receives a signal and broadcasts this signal to each of the hidden units while hidden unit sums the signal with different weights, then applies what is called an activation function to compute its output signal and sends this signal to the unit in the output layer. The output unit receives a signal from each hidden layer and sums the signals with corresponding weights and computes the output value which is typically between 0 and 1 (Tayyebi et al., 2011). In development, the ANN must be taught the characteristics of the dataset being processed (Maithani, 2009). A subset of the data is used in the training process, and the weight values of the hidden neurons are adjusted through feedback mechanisms until the error falls below a predefined threshold value (Maithani, 2009).

In this literature review, it appears ANN are suited to identifying where population expansion is likely to take place more so than how much population change is taking place. Maithani (2009) utilised an ANN to identify land suitable for urban development in Saharanpur city India. In this model, the output neuron identifies a 20m sized cell as suitable for urban development or not. Urban growth is a function of distances from major and minor road, nearest built up area, city centre and infrastructure existing in the neighbourhood (Maithani, 2009). Similarly, Tayyebi et al. (2011) investigate the extension of an urban growth boundary. The expansion of the city limits are modelled utilising the evaluation criteria of distance from; roads; built-up areas; service centres; green space; elevation; and aspect and slope.

Fractal models are algorithms that measure repetitive, scale independent patterns (Triantakonstantis & Mountrakis, 2012). Cities can be considered fractal objects conforming to non-linear and self-organising patterns. At each iteration step, the pattern replicates under parameters that promote or inhibit extension of the pattern (Triantakonstantis & Mountrakis, 2012). For example, replication is promoted along the transportation network or inhibited by natural features such as a steep slope or even distance from the transportation network. Internal fluctuations and noise result in variance within the internal self-organisation and result in variance in the output pattern (Triantakonstantis & Mountrakis, 2012).

Fractals' strength in measuring patterns has led to their use being predominantly centred on the analysis of the urban form. Thomas and Frankhauser (2013) compare the fractal dimension measured on built-up spaces with the fractal dimension measured on the street network in an urban environment. They conclude that the indices give a good indication of how regularly buildings are spread along the roads and how well roads serve buildings. Utilising such indices enables urban growth analysis.

Linear and Logistic Regression models estimate the relationship between dependent and independent variables. In urban growth model context, population numbers or land use would be the independent variable. Given particular input variable(s) the population can be projected. Models range from trend extrapolation to more complex modelling techniques (Triantakonstantis & Mountrakis, 2012).

Linear and logistic regression models have been widely used in urban growth modelling, accommodating socio-economic and environmental independent variables (Triantakonstantis & Mountrakis, 2012), with many in particular investigating land use change. A different approach is used by (Bouveyron & Jacques, 2010) who use a linear regression in housing market. Another approach is taken by (Dubovyk, Sliuzas, & Flacke, 2011), who analyse of the driving forces of informal urban development and prediction of probable locations of new informal settlements.

A Decision Tree Model is a top-down classification algorithm (Triantakonstantis & Mountrakis, 2012). The modeller derives a hierarchy of partition rules that are used to split data into sequential segments in a branch like structure (Triantakonstantis & Mountrakis, 2012). These can be represented as rules, for example, if field value x then outcome A, if field value Y then outcome B. In each iteration the decision maker is guided through a series of tests and guided to one of many predetermined outcomes (Triantakonstantis & Mountrakis, 2012). The simulation includes weighting and probability based on observed data at each decision branch. A given input range yields an outcome spectrum (Triantakonstantis & Mountrakis, 2012). In a decision tree, Al-sharif & Pradhan (2015) investigate the factors contributing urban expansion and with these factors, provide a future urban probability map. They identified a series of 12 independent factors, of which two are; distance from coast and distance from the urban centre. In the simulation the study area is described as a cellular network and the state of each cell is run through the set of rules with the dependent outcome being expansion or no expansion.

Decision Tree methods are more frequently used in the realm of data mining and useful in data classification (Kim, Kang, Hong, & Park, 2006). As such this method has been used in remote sensing to classify urban features from satellite data (Lizhong Hua et al., 2012). Kim et al., (2006) applied a decision tree method to extract spatial rules and applied these in a cellular automata model in order to simulate urban changes.

Cellular automata (CA) models are the most popular choice of urban growth model (Triantakonstantis & Mountrakis, 2012). Santé et al., (2010) attributes CA models' popularity to their simplicity, flexibility and intuitiveness, and particularly their ability to incorporate the spatial and temporal dimensions of the processes.

A CA comprises discrete cells, with each cell characterised by a 'state' (Macal & North, 2005; Santé et al., 2010) that changes over time based on internal characteristics, rules and external inputs (Castle & Crooks, 2006).

Cells have strong relationships with their neighbourhood in which local action leads, in many circumstances, to emergent structure (M Batty, 2005). Due to the limitations of the possible statuses of the cell and the relationships with its neighbours the simplicity of CA models may also be one of the main weaknesses (Santé et al., 2010). Santé et al.Santé et al. (2010) look into the advantages and disadvantages of CA in some detail. Relaxations and modifications of CA methods have led to more complexity in the models and hybrid models are evolving using a number of varying algorithms (Santé et al., 2010; Torrens & Benenson, 2005; Wu & Martin, 2002). The lists of identified adaptions to CA models are in Table 2.1. These adaptions increase the complexity of models as a means to overcome the limitations of a pure CA method. There is a wide range of variations and sub-methods used in CA and Santé et al. (2010) described 20 applied models and identified 8 major modifications to CA, as shown in Table 2.1.

Table 2.1. Types of Cellular Automata and relaxations of CA rules

| | |
|---|---|
| Analysis of urban CA models | i) binary values (urban,non-urban) |
| | ii) qualitative values that represent different land use |
| | iii) quantitative values that represent, for example, population density (value of buildings (urban)) |
| | (iv) a vector of several attriubutes |
| Relaxations of CA for urban simulation | 1) Irregular cell space |
| | 2) Non-unifom cell space |
| | 3) Extended neighborhood |
| | 4) Non-stationary neighborhood |
| | 5) More complex transition rules |
| | 6) Non-stationary transition rules |
| | 7) Growth constraints |
| | 8) Irregular time steps |

Source: adapted from Sante et al. (2010)

In a CA, defining an appropriate cell size is critically important. Cells are usually defined as being homogenous. As cell size increases the representation of the real heterogeneous world becomes distorted. For smaller cell sizes the granularity of the input data becomes problematic. Cell size is limited by the spatial resolution of input data. For example, a 10m grid could be smaller than a regular sized house. Santé et al. (2010) identified models with different cell sizes ranging from 10m x 10m up to 1km x 1km. Although Jenner (2002) found no clear benefit of using a 100m and 200m cell in a grid within urban areas, he suggested that larger grids are required as the population becomes sparser. Model outcomes are sensitive to cell size and outcomes vary according to both cell size and neighbourhood configuration (Moreno, Wang, & Marceau, 2009).

One relaxation of CA identified in Santé et al. (2010) is abandoning regular cell size and opting for irregular cells. For example, Moreno et al. (2009) utilise land parcels to define space. The interactions between land parcels are governed by a buffer around each land parcel, defining the neighbourhood of each parcel. Moreno et al. (2009) cover the transition calculations and differing buffer size for two regions, Southern Quebec and Southern Alberta, Canada. They conclude irregular cell space adequately represents the dynamics occurring in both regions.

Further to the identified relaxations, in some applications of CA, limitations may become apparent. Torrens and Benenson (2005) identify the limitation, that the automata (cells) diffuse information through the neighbouring cells, thus limiting the transfer of state information. They also note that the automata have fixed locations within the cellular construct, thus automata have no ability to move, and i.e. they are confined to a change in state. Batty (2005) notes that the automata will usually be limited in the number of states which they can take on, which are most frequently binary states. Beneneson (1998) indicates that although CA utilise bottom-up characteristics they still display the inherent restriction of the top-down approach by using a predetermined set of cell states.

Computational agent models are described in the literature under various terms and acronyms. Commonly, and used in this document, they are termed agent-based models (ABM). Other variations include Agent-Based Computational Modelling (ABCM), Agent-Based Social Simulation (ABSS), Agent Based Computation Simulation (ABCS), Agent-Based Modelling and Simulation (ABMS), Multi-agent Modelling (MA) and Individual-Based Modelling (IBM) (Michael Batty & Torrens, 2005; M. de Smith, Goodchild, & Longley, 2013). Despite a wide range of names and uses in a range of fields, there is no general agreement on the definition of agent-based models (M. de Smith et al., 2013; Macal & North, 2005; Yang, 2011). Even the definition of an agent itself is a topic of debate (Macal & North, 2005). Most authors tend to describe agents by their characteristics, more so, than precisely what an agent is. For example, Fountaine and Rounsevell (2009, p.1238) note: “Agent classes can be natural (plants, animals, etc.), people (individual, households, etc.) or more abstract entities (institution, market stocks, etc.). Agent interactions can be simple queries or complex activities with conditional reactions and feedback mechanisms.”

ABMs are suited to examining human behaviours (Castle & Crooks, 2006), social system and networks (Axtell, 2000; Lei, Pijanowski, Alexandridis, & Olson, 2005) in complex systems or environments (Michael Batty & Torrens, 2005). ABMs are scale independent (Castle & Crooks, 2006; Crooks, Castle, & Batty, 2008; Yang, 2011), and range from micro-scale models of pedestrian movements in a street (M Batty, 2005) to urban simulations of large cities (M Batty, 2005; Yang, 2011). Scale independence extends to agents having the ability to respond to both local and global stimuli. Agents are highly mobile and can be programmed to comprehend distance and direction (Castle & Crooks, 2006). Flexibility in the design of models is improving as higher resolution data become available (Macal & North, 2005).

In developing an ABM the modeller assigns characteristics to the agent(s) which determines how they behave in a simulation. An agent is an identifiable individual with a set of characteristics and rules governing its behaviours (Macal & North, 2005). In the simulation, agents exhibit decision making capability or adaptive behaviour. For example, Fontaine & Rounsevell (2009) define an agent as a household, characterised by the life stage of the household members, and the location choices at different life stages. Agents can assess the traits of the other agents. In Fontaine and Rounsevells (2009) study, household agents seek a desirable location base on proximity to the coast. For example, in their East Anglia study area, retirees prefer coastal areas. Agents interact with each other, these interaction(s) have the capability to induce a response action; often studied are cultural associations (Schelling, 1971). Agents are goal-directed, for example, Fontaine and Rousevells’ (2009) agents evaluate all property location options upon a change in life stage and choose the highest preference location from the properties available. Agents behave autonomously and in a self-directed manner (Macal & North, 2005). Agents are adaptive and may even develop skills. For example, Fontaine and Rousevells’ (2009) agents evolve through demographic life stages accordingly.

Agent-based models have some positive characteristics that make them suited to urban growth modelling. Importantly, these models are visually orientated and people are good at pattern recognition. Presenting results visually can be very effective. This is especially relevant for demonstrating technical results to policy-makers and decision-makers (Axtell, 2000). Nevertheless, the results must be interpreted appropriately (Castle & Crooks, 2006). As Triantakonstantis & Mountrakis (2012) report, in bridging the gap between theoretical and applied models it is important for the audience to be able to grasp the principles of the model, which is easier when the model is represented visually.

There are some limitations associated with ABMs. Most of the limitations or disadvantages are related to computer modelling in general. In relation to agent-based models, the primary criticism is directed at the underlying complexity theory upon which ABMs are based (Couclelis, 2002). Complex psychology, subjective choices and potentially irrational behaviour are human traits that are difficult to account for and complicate the development and implementation of models, and the interpretation of results (Castle & Crooks, 2006). Complex systems such as ABMs are very sensitive to initial conditions and to small variations in the interaction rules (Couclelis, 2002). Consequently, the line between decidedly different outcomes may be very thin (Foss & Couclelis, 2009). As an example of this sensitivity, due to variance in outcomes of model calibration, Fontaine & Rounsevell (2009) conclude that they would require further testing of the input weighting to explore the effect on the results.

Couclelis (2009) cautions that local interaction rules are seldom as simple as the models would have it. Models may reinforce positive feedback loops, thus overemphasising emergence. In contrast, there is more stability in the real world than complexity theory would have us think. Foss & Couclelis (2009) describe agent-based models as metaphors that cannot reproduce reality or pass any statistical tests: “The results of agent-based simulations are thus worth examining carefully not for quantifiable evidence but for the suggestion of interesting patterns, the generation of testable hypotheses, and generally for the exploration of ideas.” (Foss & Couclelis, 2009, p. 139).

It is important that the purpose of the ABM is clearly stated (Crooks et al., 2008; Grimm et al., 2006). Broadly, models can be grouped into either explanatory (where processes within the model are under investigation) or predictive (where the model outcomes are the focus). Management of time should be carefully considered. Agents could operate in synchronous mode, i.e. simultaneously, in which case conflicts can arise when agents compete for limited resources (Torrens & Benenson, 2005). Resolution of conflicts should be handled carefully. Time could be managed asynchronously, whereby agents respond in turn and conflict can be avoided. However, the sequence of actions becomes more important (Torrens & Benenson, 2005).

Another of the key challenges identified by Crooks et al. (2008) is making models operational i.e. running them as configurable simulations. Models may be developed but unable to be replicated in other situations or locations. Although generic software is being developed, the extent to which these products can be configured is variable and they will always be limited in their applicability (Crooks et al., 2008). This could also perhaps be one of the factors behind the findings of Triantakonstantis & Mountrakis (2012), who found that ABMs have experienced limited operational adoption. Software model building applications that supply templates and user graphical user interfaces are available. These facilitate easier model development and research rather than producing coding software (Crooks et al., 2008).

# 3 Review of applied agent-based models

This chapter reviews four agent-based models in detail. Each of the models covers different aspects of urban expansion in small geographic areas of similar size to the study area of the Waikato. These models are based on households as the primary agent, share similarities in the types of data and function at the land parcel level. Each of these studies incorporates a household that evaluates the environment in different ways, each providing alternative methods that could be adapted to the study area situated in the Northern area of the Waikato, New Zealand.

Agent-based models can be cumbersome and difficult to explain (Grimm et al., 2006, p. 116). If models are not well presented and not properly understood there is a natural resistance to their adoption into practice. Crooks et al. (2008) express concern that the theoretical implications of many agent-based models remain implicit and hidden. In 2006, Grimm et al. (2006) first proposed the development of a standard protocol (ODD - 'Overview', 'Design concepts', and 'Details') for describing ABMs. A standard protocol helps both researchers and model audience to understand the approach of the modeller. In further addressing the concerns that the underlying theory is not well described, Müller et al. (2013) suggested the ODD protocol would be enhanced by more extensively describing the underlying theory under the design concepts (ODD+). A

Clear understanding of the underlying theory ensures that models and outcomes can be more easily understood and compared to results from other models. The following reviews have been described using the ODD+ to facilitate comparison and contrast.

## 3.1 Multi-agent simulations of residential dynamics in the city (Benenson, 1998).

*Overview*

Benenson's (1998) approach is based on both the economic and cultural attributes of the individual agents. This model is exploratory, investigating the processes of economic and cultural self-organization in the city, with the intention of answering the questions "What are number and the level(s) of [economic and cultural] segregation of the emerging groups, if any?", "Are they fixed or do they [the emerging groups] vanish with time? and What is their [the emerging groups'] 'life-history'?" (Benenson, 1998).

The simulation is run on a theoretical population. The size of the population is not stated in the study. However, based on the grid size the number of agents is 1600. The model investigates individual citizens' economic status, which is uni-dimensional and quantitative, and cultural identity, which is multidimensional and qualitative.

The intended user of the model is not stated. The patterns of sociocultural emergence could be useful for planners, particularly in places where cultural diversity or tensions are high. In the discussion, Benenson questions whether such models can be used in real world situations and concludes that the advances in GIS software, improvements in high resolution data and improvements in computational power allow for the development of multi agent models.

The model time steps (t) are not given any time units. The research questions are based on observation of events at periods in the model. Observations are made and discussed at t= 10 and t=400. Overall the model is run for 2500 time steps.

*Description*

The agents in Benenson's model are individuals, and the theoretical city comprises a grid. Each of the 40x40 grid cells can hold one agent. The model is initiated with a small number of agents at the centre of the grid, while all other cells are vacant.

At each time step, each agent decides whether to stay at the current location or move, based on a 'tension' resulting from an economic status or a 'cognitive dissonance' resulting from differences in cultural identity.

The agents' economic status is a function of a fixed growth rate, mortgage repayment, house value, neighbouring house values and the average city-wide house value. The growth rate is randomly assigned from a normal distribution. Economic tension develops when the economic status of the agent differs from their neighbours. Tension thresholds define the action to move to a suitable vacant cell. Under certain circumstances, the agents can leave the grid (i.e. the city).

Agents are attributed with a multi-dimensional cultural identity, described by a K-dimensional Boolean cultural code comprising n vectors and k dimensions. This identity can give rise to a "cognitive dissonance" as agents seek to associate with agents of a similar identity. In this situation, cognitive dissonance refers to the agent's processes of weighing up the pros and cons of residing in neighbourhoods that share their cultural identity or a different cultural identity. The cognitive dissonance is defined as an average of the differences between an agent's cultural identity and the identities of its neighbours.

The agent's cultural response is influenced by the number of cultural identity dimensions (k) and the global cultural information. The number of dimensions influences the resolution of the cultural identity, and the global cultural information influences the cultural sensitivity to neighbours of different identity. The global cultural information is determined in the model by the value of Lieberson's segregation index (Lieberson, 1981). The higher this index value, the greater the propensity for cultural segregation to persist and for agents to preserve their cultural identity.

Agents move when there is an increase either their economic tension or their local cognitive dissonance. When agents are restricted and no suitable vacant locations are available, the agent can alter their cultural identity in order to reduce the dissonance.

*Initialisation*

The model is run as two versions, the economic and the cultural version of the model. At every cycle, a constant number of individuals try to enter the city and occupy a house (Benenson, 1998). In the economic model, the new agents are randomly and independently assigned an economic status and growth rate. In the cultural version, the identity of the immigrants is assigned at random, in proportion to the current proportions of cultural identities.

*Results and discussion*

The author does not explicitly state other exogenous factors. However, housing supply, ability to migrate and proximity to infrastructure are some factors that could be assumed to be exogenous.

In terms of spatial and temporal resolution, each step in the model represents a 'cycle', which is not explicitly related to a measure of time. The model is run up to t=2500 so it would be highly useful to know how fast the economic and cultural transitions take, it could be months or years. Spatial resolution is not stated, based on a grid where each cell represents a household. The 'city' is only represented as 40 x40 cells so this bears no relationship to a real city.

The results of the economic status version of the model "…converge towards a smooth, slowly varying spatial pattern." (Benenson, 1998, p. 39). In the short and medium term the domains, occupied by individuals of low and high status, can be foreseen.

The cultural identity version shows that from the cultural perspective the distribution can only be foreseen in the short-term. The city can only sustain a limited number of cultural identities, cultural groups can emerge and disappear and overall the city preserves cultural instability.

This study highlights the relative complexity inherent in modelling more complex relationships such as cultural identity and association. Cultural association is recognised as a determining factor in household location choices and urban growth patterns (Schelling, 1971). This work also highlights different approaches using quantitative or qualitative data. The adaptive qualities of the agents provide a good representation of the real world. In the economic model, the agents have an economic growth status which changes over time as the fixed growth factor. On the cultural side, the agents can alter behaviours and reduce dissonance by altering cultural identity.

## 3.2 An agent-based approach to model future residential pressure on a regional landscape (Fontaine and Rounsevell, 2009)

*Overview*

This model is a case study, simulating future demand for residential housing in the Norwich region in the East of England. The number of agents is not stated however, it is indicated that the model will simulate in excess of 700,000 household agents derived from a population of about two million citizens.

The model simulates land-use patterns at the regional scale by integrating qualitative knowledge of agent location preferences with quantitative analysis of urban growth dynamics. The primary driver of residential mobility is a change in the life stage of households.

This model produces results that will inform end users of the land use pressures as increasing quantities of residential land is required to support a growing population. The nature of the life stage approach also provides strong indications as to the types of households that are driving the demand. For example, areas close to the coast experience high demand, primarily driven by retirees.

The intended user of the model is not stated. The outcomes could be used for strategic planning, particularly housing forecast/demand management and environmental impacts of different scenarios.

Each time step represents one year, and the model is simulated for a 25 year period, from the baseline year 2000.

*Description*

There are two agents; members of a household and land units represented in a 250m grid. The household is mobile and chooses to occupy a suitable land unit. The household comprises individuals and is characterised by the numbers and ages of children, adults and family. The land unit, a cell in the grid, is considered to be homogenous and immobile. The land unit's 'potential attractiveness' is assessed by the households on three sets of criteria namely: land tenure, land unit accessibility, and environmental criteria

There is a fixed housing stock (supply) both in quantity and house physical size and household choice indicates the size preference and the preferred location. The agents compete for locations most desirable for their life stage. The model tracks demand for residential properties in land units. Each land unit's residential property count is compared with a hypothetical total property count to produce a 'demand ratio'. A high ratio indicates a greater change in the urban landscape.

Base assumptions stipulated for the modelled environment are that household behaviours don't change over time, demographic trends continue, and no more residential properties are built. Rents and income are kept constant through the model.

*Initialisation*

When a household agent changes from one life cycle stage to another they can change location. The household evaluates the region for a location with their new, preferred characteristics. Each life stage has different 'attractiveness' criteria. Land units' 'potential attractiveness' are classed as 'optimal locations', 'second-best locations' and 'suboptimal locations'. Each household chooses the highest attractiveness class where a vacant land unit exists. If there are no suitable land units available, the household:

- does not relocate but remains unsatisfied (if it already has a land unit);

- chooses any available land unit (if it is new to the region); or
- leaves the region (if it has no land unit and none are available to choose from).

*Results and discussion*

The authors report a low correlation coefficient (<0.5) between the observed and the simulated spatial distributions, although they consider the result to be adequate given the qualitative nature of the exercise. “A better match between data and model would have been surprising as the model is a stylised representation of reality.” (Fontaine & Rounsevell, 2009, p. 1249).

To more comprehensively model the impacts of future growth patterns and the potential environmental impacts agent-based models need to:

- improve on calibration methods that are appropriate to spatial urban data;
- include other agents with a role in urban development and planning;
- adapt the approach to better simulate scenarios of future urban change; and
- consider measures of ecosystem services to better integrate the environmental impacts of residential development.

Fontaine and Rounsevell highlight the role life stage has in the location choice of residential location, particularly in a polycentric region. Their approach to establishing the demand for house types manages the uncertainty of what types of houses would be constructed by developers.

The life stage aspect of this approach could be beneficial in the Waikato region, as broadly speaking there is not a great diversity in the types of houses. The structure of the planning facilitates free standing houses and profit driven maximisation of the floor area. One could rightfully question whether low diversity in house types is beneficial to the future efficiency of a region, particularly when household size is projected to decrease and the composition of households is likely to become more diverse. Fontaine and Rounsevell’s work could be utilised to help guide planners to ensure the right type of houses are built in the right places. It would be very interesting to know if the demand is driving the house type in East Anglia or conversely if planners and developers are driving the house types being constructed.

## 3.3 High Resolution Urban Land-use Change Modelling: Agent iCity Approach (Jjumba and Dragicevic, 2012)

*Overview*

This is a case study and implementation of a software called *Agent iCity*, in which the process of urban land-use change is simulated using irregular spatial units at a cadastral scale. This model simulates the subdivision of land and production of new land parcels. The model is applied in a rapidly developing suburb of the City of Chilliwack, Canada. The study area is approximately 14 $km^2$ covering about 1500 land parcels, many of them undeveloped.

The model was developed for urban planners and this case study demonstrates the outcomes of two policy scenarios.

The different components of the model run using differing time steps:

- the household agents operate on a monthly clock;
- the developer agent operates on a six-monthly clock; and
- the planning agent operates on a yearly clock.

Overall the model runs five years into the future.

*Description*

This model incorporates interactions between five types of agents; the urban planner, housing developers, households, retailers, and industrial manufacturers.

One urban planner agent represents the planning body, with a primary goal to identify the cadastral parcels upon which subdivision can take place. For large enough land parcels the planner agent initiates a subdivision module which simulates the subdivision process. The model user stipulates the planning policies and the planning agent operates within the determined framework.

The housing developer agent searches for land parcels that have been identified by the planning agent. The most profitable parcels have the highest potential land value and are surrounded by desirable land use activities, based on a calculated proximity score. The proximity score is dynamic and is recalculated when new subdivisions occur. The developer introduces new residential units, which are assigned a property value randomly selected from a normal distribution whose mean and standard deviation is specified by the user before the model is initialised.

The household agent is motivated to move when their household income becomes greater than the neighbourhood average income. The household will move to a vacant parcel in a neighbourhood that has an average income and average property value equal or greater than that of their last residence. A household cannot move more than once in 12 months.

Employment is directed by the retailer and industrial manufacturer agents. As the number of households increase, the number of employers increases as do the number of retail and industrial employment locations. Jjumba & Dragićević recognise this representation of the city's retailers and industrialist is over simplified. In this model, employment is responsive to population growth and this too is recognised that employment is a growth driver. The location of employment sites and selection process of the retail and industrialist agents is not elaborated on.

The number of new agents is set by the user. Average income is a fixed input variable. The property value range and normal distribution are set by the user.

*Initialisation*

A neighbourhood is defined as parcels that fall within a specified buffer distance from the boundary of the central parcel. The result is each land parcel has its own unique neighbourhood and the characteristics of this neighbourhood can be determined.

The planning agent is initiated and identifies parcels available for future development. The land subdivision module is used to subdivide land parcels. The initial large land parcels are first subdivided into city blocks which are transected by roads. The final division of the city blocks into new land parcels is done in accordance with predetermined minimum area and frontage rules.

The developer agent follows and adds residential units to the most suited locations and in the final step the household agents evaluate the merits of available vacant residential units and relocate if beneficial. With a number of modules and differing time cycles, an agent module is responsible for coordinating the timing of actions.

*Results and discussion*

The results of this model provide insights into the land use change driven by stakeholder relationships, land value and household income. The interactions of the key stakeholders are a valuable component in providing suitable subdivision and therefore the new options to the household agents.

The authors conclude that the results represent the land use change and subdivisions well, however, more detail could be incorporated into the interactions of the agents as well as the influence of economic factors such as land market dynamics and employment opportunities among others.

The inclusion of a developer agent is particularly relevant and could be a great addition to the Waikato agent-based model. The developers have key drivers that can be in conflict with the planners, thus provide a set of criteria that separate the planners from house owners. Morgan (2010) also identifies a varied approach taken by land developers.

This study also highlights that the complexity of the model increases as a greater number of agents are included and yet there is still a range of very broad assumptions that need to be accounted for.

The approach to the agents that can subdivide the land parcels is particularly interesting and could be a very useful tool to help planners understand the capacity for subdivision in an area or how much impact design standards (e.g. road and sidewalk width) and setback rules have on how many houses can be constructed.

The operation of different time clocks is an interesting feature. In reality, events take place asynchronously, however, the models reviewed treat events synchronously. Castle & Crooks, (2006) discuss managing synchronous and asynchronous time. When time managed synchronously, the agent's behaviour is scheduled in discrete time steps, in which case the order of event needs to be planned carefully. In the real world, behaviour is more commonly asynchronous and events occur on different time schedules. In such cases managing any feedback mechanism require careful planning.

## 3.4 Impact of urban planning on household's residential decisions: An agent-based simulation model for Vienna (Gaube and Remesch, 2013)

*Overview*

This is a model based on the city of Vienna, with roughly 770 000 household agents represented in 59 small city areas. The model simulates the residential patterns of different household types based on migration and demographic life stage progression. One of the focal points centres on sustainability and the implications for energy consumption under different city development scenarios. The scenarios are tabulated as conventional urban planning, sustainable urban planning, expensive centre and no green area.

It is not stated for whom the model is intended. It could very well be used by strategic planners, policy makers or planners interested in energy distribution, including transport energy.

The specific research question is not stated. However, the model outputs firstly demonstrate the distribution of household types depending on the age class and income class to which they belong under different scenarios. Further to this, energy consumption per capita and per household are investigated, highlighting the long-term implications for the city. The four scenarios are conventional urban planning, sustainable urban planning, expensive centre, and no green area preference.

In the model, one time step represents one year and the model runs for 50 time steps.

*Description*

Gaube and Remesch describe the main interaction as that taking place between households and spatial units. The main intention is to investigate energy consumption patterns in response to different city planning policies. Each household agent is classified according to household types, defined by age and family structure. Each household type displays different preferences for residential locations. The residential location characteristics of the spatial units are classified according to environmental amenities, accessibility of high-level public transport, infrastructure, centrality, dwelling cost, dwelling size and social prestige.

This model simulates a synthetic population and utilises a demographic sub-model to simulate biological event, a residential mobility sub-model that calculates a satisfaction/stress factor and an urban development and vacant dwelling sub-model that accounts for changes in the urban cityscape.

*Initialisation*

In the simulation, households transition through different household types (i.e. life stages), which influences their space requirements, income and residential preferences. At each change in household type, the agents assess the residential locations. The residential sub-model establishes the satisfaction of each household and when a stress level is reached, a relocation is triggered. The residential location characteristics are weighted, and scenarios can be run by adjusting the relative weighting.

The city environment is dynamic and the urban development and vacant dwelling sub-model controls the rent, income and the allocation of new residential locations. Once the household type distributions are established, further analysis based on per capita energy consumption is undertaken to demonstrate how different policies have different social impacts which influence energy consumption.

The spatial resolution is at the small city area. The core model has a household type seeking a house (dwelling). There are seven such household type classes establishing the synthetic population. The initial population figures are taken from the Austrian 2001 census and micro-census 2006 and 2008.

Exogenous to the model are the property rents and the urban development scenarios. The scenarios specify the number of newly constructed dwellings per cell and model year and are based on projects already planned or under construction. Expert knowledge provides information for the scenarios identifying suitable areas for large urban projects.

*Results and discussion*

The model successfully presents outcomes for each of the scenarios and demonstrates the effects of different policies on the distribution of households and household types, which in turn affect energy consumption.

In the conclusion, Gaube and Remesch briefly discuss the continuum from simple to complex models. This is relevant in the context of the Waikato study area, as the possible users of this model need to easily understand the model workings and assumptions. “Nevertheless, efforts should aim at developing models [that are] (1) simple in terms of degree of details and technical implications and at the same time (2) complex enough to address socio-ecological issues. In our opinion, this is one of the challenges in using agent-based models for analysing socio-ecological systems.” (Gaube & Remesch, 2013, p. 102).

As with the previous studies that include aspects of the life stage progressions, the inclusion of a demographic component provides an important link into the long-term (50 years) projection, as the population structure is changing in accordance with the other components of the model.

The method of weighting input criteria demonstrates an easily adaptable platform from which scenario testing can take place.

In a similar fashion to Fontaine and Rounsevell (2009), the model projects an outcome of the housing types required. However, there is uncertainty as to whether these housing types actually get constructed. In this particular study, it would be very interesting to know what the implications of a mismatch between households and housing types are in terms of energy consumption. Gaube and Remesch refer to a socioeconomic structure, energy consumption is optimised when the number of occupants' matches the types of houses, i.e. a good socioeconomic structure. Inefficiencies arise when there is a mismatch between the number of occupants and the types of houses, i.e. a poor socioeconomic structure.

# 4 Method and assumptions

## 4.1 Purpose

The purpose of this Waikato agent-based household distribution model is to provide strategic and infrastructure planners with household and population projections that are: i) spatially explicit, independent of predetermined statistical boundaries similarly to Jenner (2002); ii) not necessarily based on the disaggregation of territorial authority population projections; and iii) independent of territorial authority jurisdictions, providing a holistic approach to the study area. Although the model traces individual households, these are not the subjects of the investigation. It is the emergent patterns that are reported and discussed. Figure 4.1 shows the sequence of events undertaken by each individual which determines how (if) they move to different neighbourhoods.

The research question is "what will the likely household distribution be for Hamilton, Waikato and Waipa in 2025?" Up to 2025, it is projected that approximately 31,474 new houses will be required to house a growing population (M. Cameron & Cochrane, 2015b) and the question is, under given planning objectives where are these most likely to be constructed?

The outputs of the model are: a time series of the number of occupied residential households in each meshblock and a series of maps showing net change in occupied residential households. The output data series will include spatial data attributed with the residential household counts that will enable planners to interrogate the projected changes within a meshblock or Area Unit.

## 4.2 Overview

The Waikato agent-based model spatially allocates households within the study area. In the model, each household is an independent agent. Each household agent (HA) pays rent for their current house and incurs an amenity benefit from the neighbourhood they are living in. Each household agent has a single job and incurs a travel cost based on the distance from their home to their job. In each annual time step, agents move such that they minimise their total costs relative to the neighbourhood amenity benefit they receive and the costs of relocating. At each time step, additional land units, HAs and new employment opportunities become available.

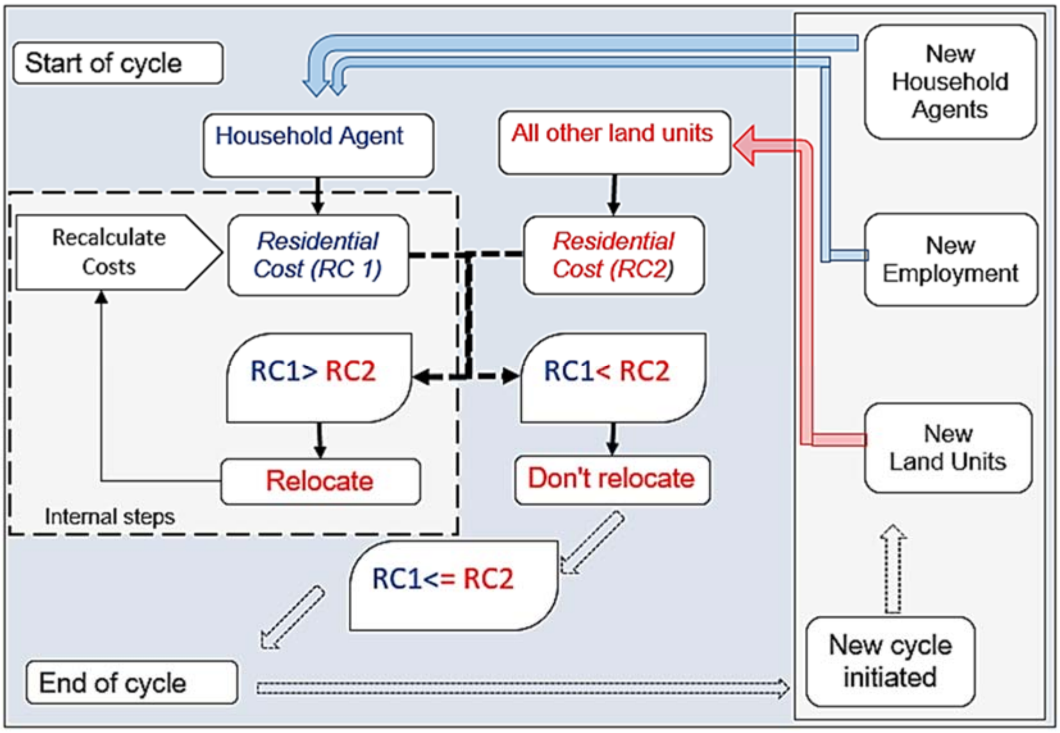



Figure 4.1. Flow Chart of the time cycles and internal process steps.

Measures of cost or satisfaction have been used by authors such as (Benenson, 1998; Fontaine & Rounsevell, 2009; Jjumba & Dragićević, 2012) to model agent responses to residential location choice. HA costs are predominantly based on economic criteria. Social, cultural and other factors such as amenities are collectively accounted for as part of the neighbourhood amenity benefit (NAB). These factors remain constant as the NAB does not change for the duration of the model.

New household agents, employment and land units are added to the modelled environment at the start of each time step, as shown on the right hand side of Figure 4.1. Each cycle represents the annual change. The agents' assessments are represented in Figure 4.1, on the left hand side. The HA's total costs are expressed as residential cost (RC). Agents will initiate a move if their residential cost (RC1) is greater than the residential cost at an alternative location (RC2). The HA with the highest reduction in cost will relocate first. A new cycle is initiated when agents have reached their goal, RC1≤ RC2, and no further cost reductions can occur. The outline of each step in the SQL (Structure Query Language) code, as commented, is provided in appendix 1 for reference.

## 4.3 Overall assumptions

The most significant assumption in this model is that all HAs have the same household composition, which is the same number and age structure of people in the household. Each household has one person employed, five days a week and travelling to a place of employment. In addition, residences within a neighbourhood have the same capital value and thus all HAs have the same imputed rent for each particular neighbourhood.

While there are many motivations for a household to relocate, some authors consider the economic aspects of the household to be more significant (Jjumba & Dragićević, 2012), while some consider life stage of the household and cultural affiliation to be significant (Benenson, 1998; Gaube & Remesch, 2013; Schelling, 1971).

In this model, the primary reason for focusing on the economic aspects is related to the known variables; i.e. population growth rate, planned growth and subdivision, and relevant employment information. In the Waikato study area, whilst cultural and social differentiation does exist, in a national or international context, the dissimilarities are not so great. Any differentiation is accounted for in two ways. The capital value of properties reflects the social status of each neighbourhood and the neighbourhood amenity value will reflect a wide range of other factors, such as landscape, natural features, transportation links, schools or shops, which established the settlement patterns over the period of model calibration (2006 to 2013). In this way, non-economic factors are indirectly taken into account.

It is recognised that it would be ideal to have heterogeneous households. However, the HAs' ability to relocate then requires assessment of housing stock, for which the future provision is required. Forecasting future housing stock is not undertaken by the local authorities and is left to the speculation of developers and likely to be short term focused (Morgan, 2010). From the model development and data requirement point of view, forecasting of housing values, size and residential type is a challenge, and will significantly increase the complexity of the model (Benenson, 1998; Fontaine & Rounsevell, 2009; Morgan, 2010), particularly the calibration, which is already recognised as a challenge in agent-based modelling (Crooks et al., 2008).

The household agent's residential cost (RC) comprises three components, situation costs, relocation costs (RL) and income. The situation costs are imputed rent (LuCost) plus neighbourhood amenity benefit (NAB) plus travel costs (TDC). Net income is calculated as Income (I) subtracting the situation costs; imputed rent (Rent), NAB and travel costs (TC), i.e.:

$$\text{Net Income} = I - \text{rent} - NAB - TC \qquad (1)$$

The comparative Net Income is calculated for all vacant land units and when the change in Net Income is greater than the relocation cost the agent could relocate. Each agent will move in order to achieve the greatest gain in Net Income. The sequence of moves is ordered by the HA with the highest net income differential, moving first.

$$\Delta \text{ Net income} > RL \qquad (2)$$

Relocation costs (RL) comprises a fixed value of $2,000 and an amount that varies according to the distance moved. Relocating incurs a cost, this cost ensures that some incentive is required to undertake a move, and it also limits the model from iterating through movements with low-cost change.

$$RL = 2000 + (0.77 \times td) \qquad (3)$$

The relocation cost comprises a fixed cost of $2000 and a variable distance cost. The distance (td) is added at a cost of 77c per km. The fixed cost of moving is highly variable depending on the service offered by a removal company. A range of values from $500 to $2500 was tested and $2000 was selected as it permitted a high number of agents to relocate as this allowed a sufficient differential between the origin and destination residential costs. The distance component was selected to represent a higher cost of relocating a greater distance. In reality, people are likely to be more amenable to relocate within a town rather more so than relocate to a new town.

Travel costs are based on travel from the place of residence to the place of employment. The TC formula is calculated as:

$$TC = 41.6 \text{ X } ((0.53 \text{ X } tt) + (0.77 \text{ X } td)) \qquad (4)$$

Where 41.6 is the average number of journeys to work per month. Each trip is calculated as the distance to work and back home again on 5 days a week on an average of 4.16 weeks per month. 0.53 is the average wage per minute multiplied by the travel time (*tt*) and 0.77c is the travel cost per km (*td*). The average wage is sourced from Statistics New Zealand household income data (Statistics New Zealand, 2013c) and the travel costs are obtained from Inland Revenue employee reimbursements (Inland Revenue, 2014).

The primary reason one employment unit per household is utilised is because the travel to work information is provided at the area unit level. For this study, the data is proportionally disaggregated to meshblocks. Disaggregating travel to work based on households potentially travelling to more than one destination introduces a complexity that could potentially make calibration of the model more difficult.

The distance to the nearest school travel cost was also investigated but it was not included in the final analysis. The school travel cost resulted in a high number of HAs leaving the rural areas and this out-flux of agents reduced the number of HAs to levels lower than the observed 2013 household count. In contrast, the urban areas filled up to maximum levels quickly and in the latter stages of the model the new agents had no choices but to start filling the rural areas. When the school travel is not included the rural areas are more stable and the new HAs steadily fill the urban area and some growth spills over to the rural areas. Calculating independent school and work travel may not reflect travel patterns accurately as alternative transport options are available for rural schools and many households will route via a school on the way to work. Travel to school would be more applicable in a model incorporating life-stage where agents require (or don't) certain types of schools or other educational facilities.

The imputed rent is calculated as 5% of average property capital value, of properties in each land unit and divided by 12 into a monthly dollar value. HAs rent increases when an agent moves into a land unit – popular land units have increasing costs and rents decline where agents have left. This recalculation is represented in Figure 4.1 by the centre box with the dashed outline. The rate of change in rent is 1% of the LuCost, this rate diminishes by 1% on each HA's move. The first agent causes a 1% change, the subsequent move results in a 0.99% change (relative to the original LuCost), the next a 0.98% change, this creates a diminishing influence on the LuCost. Similarly, the rents increase is larger on the first move and subsequently becomes smaller with each move. Altering the rent by a diminishing amount is necessary in areas such as major subdivisions or growth areas, where high levels of growth are anticipated if the rent changes are not multiplicative the rents can quickly become 0.

## 4.4 Land Units

Land units are selected to represent the smallest geographic area possible and are non-mobile. They could be representative of a census tract, a property or a regular cell. The characteristics of the land unit define how many household agents can reside in the land unit at any one time. Each land unit has a maximum capacity, i.e. the maximum number of household agents that can reside in the land unit. An alternative viewpoint is that the land unit contains a number of dwellings, which could be occupied by household agents. This capacity can change over time as per the town planning rules. Property subdivision or Greenfield developments can result in an increase in the capacity. In this model, land units are represented by the meshblock which typically range in size from 1 to 350 residences. In the study area, the mean number of residences per land unit (meshblock) is 39.4.

The purpose of the land unit is to hold the record of how many HAs occupy properties and how many vacant properties exist. At each time step new vacant properties (subdivisions) are added to each land unit. In addition, the land unit holds the rent and the number of employment opportunities. Employment opportunities are the determinant of travel cost

Imputed rent is a characteristic of each land unit thus a HA assumes rent based on location. In some models, the rent cost is associated with the economic status of the occupier, for example, Benenson (1998) and Gaube & Remesch (2013), whilst in others, such as Jjumba & Dragićević, (2012), the rent is attributed to locations. This model is based on the location holding the rent value. The primary reason for this is that the capital values of all properties are known and in a meshblock, properties are assumed to be identical. Household income data are only available at area unit level. This would not provide sufficient differentiation among the household agents for them to respond individually. Rent associated with the economic status of the household agent or mobile agent is more suited to investigating the selection based on affordability and residential change through life stages, as for example in Benenson (1998) and Gaube & Remesch (2013).

Other than council influenced infrastructure (i.e. major services such as roads, water and sewerage), no account is taken of any restrictions in land supply. If land is required to be developed into residential land parcels they are immediately available, as long as council growth strategies are planning or zoning the release of that land. This contrasts with Fontaine & Rounsevell (2009), who use a fixed number of residences and the model is used to identify where there are pressures for additional residence and the housing type dependent on the life stage of the agents. Jjumba & Dragićević, (2012) and Morgan (2010) use developer agents to manage the supply of residences according to the demand.

## 4.5 Neighbourhood

The number of neighbourhoods is fixed through the duration of the model so in effect, a neighbourhood cannot grow in physical size but may become more densely populated.

Each neighbourhood has its own value (NAB) associated with its characteristics such as; access to amenities, infrastructure, transportation links, natural environment and school zoning. Attractive neighbourhoods have a low value, thus household agents seek to occupy these following their goal of minimising costs. The NAB is derived through the calibration process refer to 4.8. The NAB associated with each neighbourhood does not change over time.

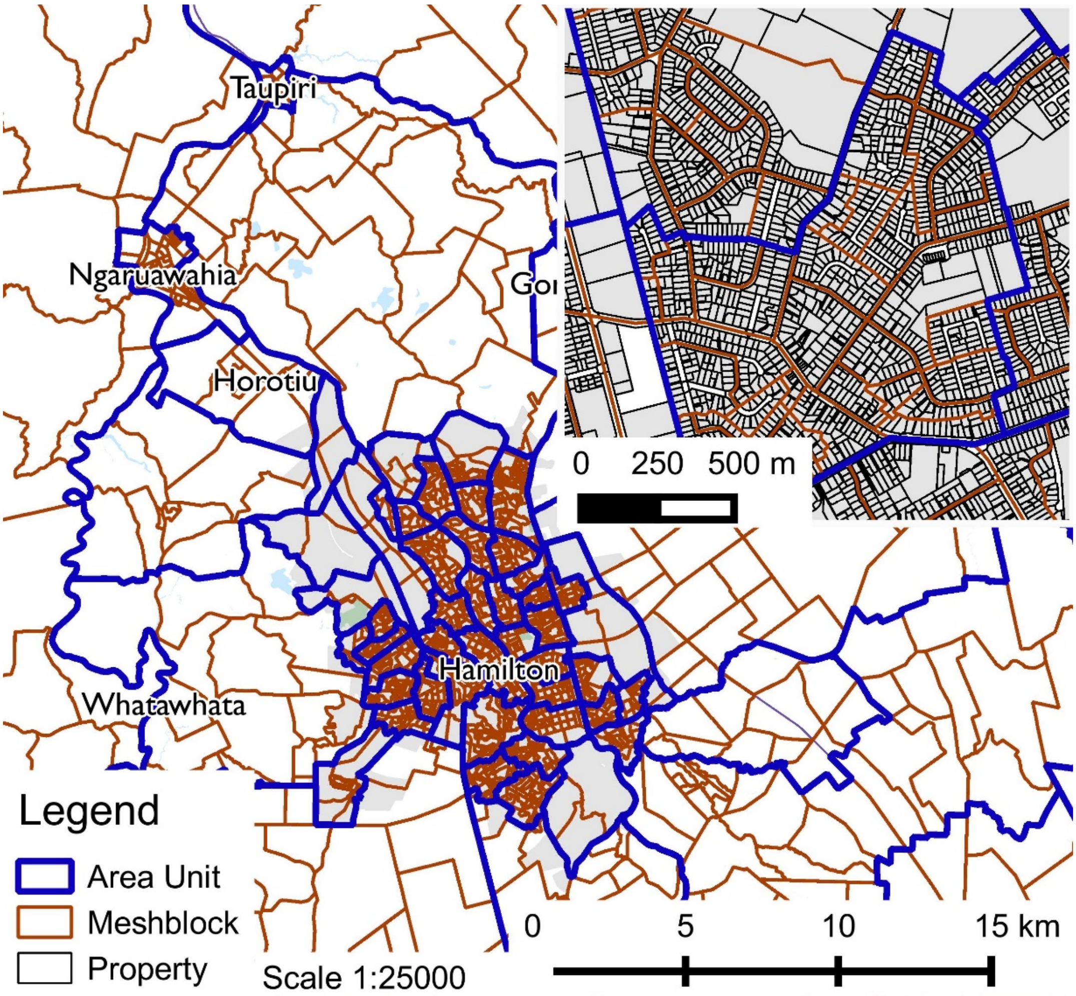
The starting neighbourhood NAB is a predetermined cost. At the commencement of the calibration process, this is arbitrarily set at $2,500 and adjusted accordingly through the calibration. The value $2,500 was selected as being more than the average rent of $1,984. At the starting point of the model neighbourhood NAB has slightly higher weighing as a component of the residential cost. However, rents increase due to the increase in the number of HAs, thus the relative weighting of the NAB decreases as the model progresses. A NAB slightly higher than the average rent allows the neighbourhood amenity value to have a slightly higher proportion of the RC during calibration. This provides slightly more emphasis on the amenity during calibration. When running the model the NAB is fixed and as more agents are added the rent component increases and becomes more significant in the HAs RC.

## 4.6 Data sources

### 4.6.1 Geographical classification

The meshblock is the smallest geographic unit for which statistical data is collected and processed by Statistics New Zealand. A meshblock is a defined geographic area, varying in size from part of a city block to large areas of rural land (Statistics New Zealand, 2013e). Meshblocks contain between 0 and 300 dwellings, with an average of 40 dwellings. There are 2302 meshblocks in the study area. Meshblock spatial extents and unique identification numbers form the basis of land units and neighbourhoods, refer to Map 2.

Area Units (AU) are aggregations of meshblocks. They are non–administrative areas that fall between meshblocks and territorial authorities in size (Statistics New Zealand, 2013a). There are 108 AUs in the study area. AU unique identification numbers have been coded against meshblocks so results may be aggregated at a higher level. Area Units are represented on Map 1




Map 1.Statistics New Zealand Area Units and meshblocks

Note 1. Area Units and meshblocks obtained from Statistics New Zealand Digital Boundaries 2013

A Territorial authority (TA) is defined under the Local Government Act 2002 as an area administered by a city council, district council, or unitary authority (Statistics New Zealand, 2013f). Meshblock and AU unique identification numbers have been coded against TAs so results may be aggregated to a higher level, refer to Map 1. Map 1 shows the relative size of Area Units, meshblocks and property boundaries. Note the relative size difference between the rural areas and the city.

### 4.6.2 Other data classification

Household Agents (HA) are derived from census dwelling count 2006 and 2013 (Statistics New Zealand, 2013b). Refer to map 5 which shows the location of dwelling distribution as at 2013. Each HA has a unique identifier and is assigned a location in a land unit based on the census dwelling count. HAs are assigned a relative income, obtained from one of fifteen income ranges (Statistics New Zealand, 2013e).

The National Institute of Demographic and Economic Analysis (NIDEA) produced regional population and household projections. The household projections are the basis from which household agents are added to the model (M. P. Cameron & Cochrane, 2015).

Employment numbers and projections were obtained and adapted from employment projections developed by Market Economics (McDonald, 2015). Employment was projected at the level of AUs and required converting to household employment number at the meshblock level. The employment projection dataset requires disaggregating from larger AU geographic areas to smaller meshblock geographic area. The disaggregation is done on a proportional basis. This is done, based on the proportion of modified employment count (MEC) per AU to employed people per meshblock (Statistics New Zealand, 2013d). The number of MEC is known for each AU. Each AU contains a number of meshblocks and the utilising the number of people employed in each meshblock, as recorded in the census, the number of MEC can be proportioned to each meshblock. The projections of MEC assumes these proportions remains the same into the future.

The travel and distance matrix was adapted from the Waikato Regional Transport Model (G. Smith & Bevan, 2011). Transport matrix zones are very similar to meshblocks in extent, thus the travel distance between meshblocks can be tabulated. Off peak, 2013 travel distance (td) and travel time (tt) are provided.

Land Units contain a number of land parcels situated in a meshblock. The initial property data are obtained from (Waikato Regional Council, 2013). The number of dwellings (Statistics New Zealand, 2013b) per land unit are calculated as the starting capacity. Each land unit has a maximum capacity regulated by planning policies. Changes in land unit capacity are outputs recorded at each time step.

Land unit cost (rent) is an imputed rent calculated from property value data from the Waikato Shared Valuation Data Service, SVDS (a Local Authority Shared Services Limited (LASS) Council Controlled Organisation CCO) (Shared Valuation Database Service (SVDS), 2014).

The location and number of new land parcels (i.e. land subdivision) that are introduced at each cycle is determined by the respective Councils' long-term growth and development plans. In Waipa District, the sequence of new land units is outlined in the Growth Strategy and Proposed District Plan (Waipa District Council, 2014). Staging schedules and various snippets of local knowledge have been supplied by Gary Knighton, Planning Manager Waipa District Council (Knighton, 2015). Growth cells are spatially split by meshblock boundaries and property counts are used to apportion future growth.

Hamilton City Council provided spatial files detailing the location of brownfield lots (i.e.Hamilton City Council provided spatial files detailing the location of brownfield lots (i.e. land parcels in existing neighbourhoods with potential for further subdivision) within the City as of June 2013 (Hamilton City Council, 2013). Staging schedules, various insights and items of local knowledge have been supplied by Michael Spurr, Senior Strategic Advisor Hamilton City Council (Spurr, 2014). Structure plans outline the development plan for Peacocke (Hamilton City Council, 2014b), Rototuna (Hamilton City Council, 2014d), Rotokauri (Hamilton City Council, 2014c) and Ruakura (Hamilton City Council, 2014e) were accessed from the Hamilton City Council website. Growth cells were spatially split by meshblock boundaries and property counts are used to apportion future growth.

Growth and development nodes in the Waikato District were sourced from structure plans (Waikato District Council, Territorial Authority, 2014a) and growth strategies (Roskruge et al., 2010) published on the Council's website. Staging schedules, various insights and items of local knowledge have been supplied by Vishal Ramduny, Planning and Strategy Manager, Waikato District Council (Ramduny, 2014). Growth cells were spatially split by meshblock boundaries and property counts are used to apportion future growth.

## 4.7 Verification & Validation

A primary component of verification can be described as debugging (Castle & Crooks, 2006), to ensure there are no errors in the code that could be producing an incorrect result. Verification of this model takes place in two stages. The first stage was during the code development. A sample area covering the town of Ngaruawahia with 1,440 HA was selected. Each variable was independently tested, and prior to running the model the predicted outcome was calculated using Excel. These tests on each variable were run to ensure that the model produced the expected outcome. As a result of these tests, some code amendments were required to ensure that the model data and SQL code was suitable for upscaling.

The upscaling of the data being modelled introduced significant and undesirable increases in the time taken to run the simulation. The main issue encountered was the step in the process where the agent with the highest possible gain is selected to move. This process requires the whole population of HA to complete the residential cost calculation prior to relocating. The initial model was written in GAMA (*Gama-platform*, 2014) with data stored in a Microsoft SQL Server Database (Microsoft Corporation, 2012). Upon upscaling it was decided that the model should be exclusively contained in MS SQL. This subsequently reduced cycle run times to less than 45 min per cycle, which is viable for running the model for 13 time steps, or even repeatedly running the model and including a process to automate the calibration (refer to Section 4.8). Following the redevelopment of the code in MS SQL, the calibration model was run using an automated calibration process.

Validating the model means checking against other models or other known or expected data or results (Castle & Crooks, 2006). This model is validated against both Statistics New Zealand and NIDEA household projections, (refer to Section 6). Both of these projections are based on top down population projections. This agent-based model's outputs are aggregated up from the meshblock to the territorial authority level and when compared the outcomes are consistent through the modelled time period.

## 4.8 Calibration

The NAB is a hypothetical cost, thus its value can be used to influence the HA's overall costs. Calibration is based on the period starting with 2006 census and finishes at the 2013 census (New Zealand's scheduled 2011 census was postponed due to Canterbury Earthquakes). In the first run, the model was initiated with all neighbourhoods having the same NAB. Model outputs were compared with the census 2013 results and all neighbourhoods with too many HAs had their NAB increased. Conversely, neighbourhoods with too few agents had a NAB reduced to make them more attractive on the subsequent model runs. The root means square error

(RMSE) was recorded and the adjustment of NAB continued until the RMSE no longer changed significantly.

Localities with the largest differences were amended by the largest quantum, and as the difference between the model output and expected result decreased the alterations to the NAB become smaller. If the auto-calibration overshot the optimal value then the process brought it back to close to the optimal value.

## 4.9 Model scenarios

Two scenarios have been established to demonstrate the effect of council planning strategies on the distribution of households. Each of the councils has a different approach to the location and timing of the new areas for urban expansion and development. Waipa has the most stringent strategy with defined development cells based on a sequential utilisation of these areas and little scope for any subdivision outside these defined cells (Waipa District Council, 2014). Hamilton has five principal development cells as well as development potential within existing areas (Hamilton City Council, 2014a). Due the significantly larger size of the development cells, the infrastructure investment is accordingly more significant. Hamilton has a higher dependency on the role of developers. Waikato district predominantly has a high degree of subdivision potential within its existing land zoning, and it has a less structured or predefined development framework (Waikato District Council, 2014b).

The purpose of these two scenarios is to demonstrate the impact of a relaxation of the growth strategies and to show the inter-relationship between the areas identified as development cells. A relaxation of constraints could also be viewed as increasing infrastructure investment to make more land available. Another interpretation of this scenario is all of the councils are fully providing infrastructure at the same time in all of the planned development areas.

### 4.9.1 Constrained development plan (CDP)

The first scenario is defined as a constrained development plan (CDP). This is the realistic scenario following the councils' respective development plans and constraints imposed by their infrastructure investment. This is termed a constrained scenario because the councils have set investment resources and plans which determine their long-term plans and growth strategies.

Table 4.1. Hamilton, additional residential locations in Land Units considered to be development cells.

| | | **Additional residential locations by (year)** | | | | | |
|---|---|---|---|---|---|---|---|
| **Area Unit** | **Meshblock** | **2013** | **2014** | **2015** | **2016** | **2020** | **Total** |
| Horsham Downs Rototuna | 951911 | | 150 | 400 | 200 | | 750 |
| | 951912 | 91 | | | 60 | 70 | 221 |
| | *sub-total* | *91* | *150* | *400* | *260* | *70* | *971* |
| Huntington Rototuna | 951802 | | | | 82 | | 82 |
| | 952201 | 152 | | | | 25 | 177 |
| | *sub-total* | *152* | | | *82* | *25* | *259* |
| Sylvester Rototuna | 951706 | 525 | 169 | 89 | | | 783 |
| | 951709 | | 68 | 82 | | | 150 |
| | *sub-total* | *525* | *237* | *171* | | | *933* |
| Newstead Ruakura | 954211 | | | | 1,960 | | 1,960 |
| | *sub-total* | | | | *1,960* | | *1,960* |
| Burbush Rotokauri | 976801 | | 1,400 | | | | 1,400 |
| | *sub-total* | | *1,400* | | | | *1,400* |
| Rotokauri | 976701 | | | | | 82 | 82 |
| | 976702 | | | | | 330 | 330 |
| | 976802 | | 1,099 | | | | 1,099 |
| | *sub-total* | | *1,099* | | | *412* | *1,511* |
| Peacocke | 984506 | | 700 | | | | 700 |
| | *sub-total* | | *700* | | | | *700* |
| Temple View | 978900 | | 100 | | | | 100 |
| | 979000 | | 70 | | | | 70 |
| | 979101 | | 0 | | | | 0 |
| | 979102 | | 0 | | | | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 979400 | | 0 | | | | 0 |
| | *sub-total* | | *170* | | | | *170* |
| Hamilton City | Total | 768 | 3,756 | 571 | 2,302 | 507 | 7,904 |

Note 1. Adapted from Hamilton City growth strategy, (Hamilton City Council, 2014a)

Note 2. Land Units are small geographic areas that have an estimated capacity for residential locations, Hamilton City's strategic plans identify development cells for residential growth.

Table 4.1 shows the timing and quantum of land available and potentially developed up to 2025 for Hamilton City. There are a number of residential locations developed on previously undeveloped land. Hamilton has a further 2,825 sites with development potential within existing residential neighbourhoods. The timing of the development of this model coincided with the start of the ten-year planning cycle thus a large quantity of land is available from 2013 and 2014. The development cell of Newstead and Ruakura, are examples of a development led by a private developer with a large quantity of land becoming available from 2016.

Table 4.2. Waikato, additional residential locations in Land Units considered to be development cells.

| | | **Additional residential locations by (year)** | | | | | |
|---|---|---|---|---|---|---|---|
| **Area Unit** | **Meshblock** | **2013** | **2014** | **2015** | **2016** | **2017** | **Total** |
| Buckland South Tuakau | 828400 | | | 240 | | | 240 |
| | 828503 | 22 | | | | | 22 |
| | 828504 | 29 | | | | | 29 |
| | 828505 | 11 | | 123 | | | 134 |
| | *sub-total* | *62* | | *363* | | | *425* |
| Opuawhanga Tuakau | 828602 | | | | 66 | | 66 |
| | 828900 | | | 16 | | | 16 |
| | 829000 | | | | | 26 | 26 |
| | *sub-total* | | | *16* | *66* | *26* | *108* |
| Redoubt Tuakau | 828501 | | | | 82 | | 82 |
| | 828604 | | | | 18 | | 18 |
| | 842700 | | 710 | | | | 710 |
| | 842800 | | | | 70 | | 70 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 829200 | | | 682 | | | 682 |
| | *sub-total* | | | *682* | | | *682* |
| Tuakau | 838102 | 1 | | | | | 1 |
| | 838200 | 1 | | | | | 1 |
| | 838900 | 205 | | | | | 205 |
| | 840102 | 11 | | | | | 11 |
| | 840200 | 29 | | | | | 29 |
| | 840300 | | 61 | | | | 61 |
| | 840400 | | | 17 | | | 17 |
| | 840500 | | | | | 41 | 41 |
| | *sub-total* | *247* | *61* | *17* | | *41* | *366* |
| Te Kauwhata | 937900 | 265 | | | | | 265 |
| | 938000 | 587 | | | | | 587 |
| | 938100 | 530 | | | | | 530 |
| | *sub-total* | *1,382* | | | | | *1,382* |
| **Waikato District** | **Total** | **1,691** | **771** | **1,078** | **236** | **67** | **3,843** |

Note 1. Adapted from Waikato district structure plans, (Waikato District Council, 2014b).

Note 2. Land Units are small geographic areas that have an estimated capacity for residential locations, Waikato's strategic plans identify development cells for residential growth.

Waikato district has less distinctive planned development cells (Table 4.2). Similar to Hamilton, most of the new development land has been serviced with infrastructure and is in a state ready for residential development to take place starting in 2013. The development cells in Te Kauwhata are based on a planned policy intervention designed to accommodate and attract growth into this town. Within the existing neighbourhoods, Waikato has a further 7,700 additional residential locations that could be created through subdividing land in existing neighbourhoods.

Table 4.3. Waipa, additional residential locations in Land Units considered to be development cells

| | | Additional residential locations by (year) | | | | |
|---|---|---|---|---|---|---|
| **Area Unit** | **Meshblock** | **2013** | **2014** | **2016** | **2020** | **Total** |
| Cambridge West | 960007 | | | | 145 | 145 |
| | *sub-total* | | | | *145* | 145 |
| Swayne Cambridge | 959002 | | 272 | | | 272 |
| | *sub-total* | | *272* | | | *272* |
| Hautapu Cambridge | 960003 | | | | 550 | 550 |
| | 960004 | | | | 274 | 274 |
| | 960005 | | | | 887 | 887 |
| | *sub-total* | | | | *1,711* | *1,711* |
| Kihikihi Flat Te Awamutu | 989304 | 210 | | | | 210 |
| | 989306 | 150 | | | | 150 |
| | 989309 | | 180 | | | 180 |
| | *sub-total* | *360* | *180* | | | *540* |
| Te Awamutu West | 970900 | | | 100 | | 100 |
| | *sub-total* | | | *100* | | *100* |
| Waipa District | | 360 | 452 | 100 | 1,856 | 2,768 |

Note 1. Adapted from Waipa growth strategy, (Waipa District Council, 2014).

Note 2. Land Units are small geographic areas that have an estimated capacity for residential locations, Waipa's strategic plans identify development cells for residential growth.

In contrast, Waipa has 16 residential locations that can be created outside the development cells. Waipa (Table 4.3) has a large number of residential sites developed prior to 2013. The development cells that have had subdivision initiated prior to the start point of the model will have the land unit capacity calculated as the number of existing properties and the total planned capacity. The table above only reflects development cells that are in land units with no prior development. Waipa's strategy is clearly set out in sequence and each development cell is initiated upon achieving development thresholds.

Table 4.4 Summary of additional residential locations by territorial authority and year

| | Additional residential locations by (year) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Territorial Authority** | **2013** | **2014** | **2015** | **2016** | **2017** | **2020** | **Total** |
| Hamilton | 768 | 3,756 | 571 | 2,302 | | 507 | 7,904 |
| Waikato | 1,691 | 771 | 1,078 | 236 | 67 | | 3,843 |
| Waipa | 360 | 452 | | 100 | | 1,856 | 2,768 |
| Grand Total | 2,819 | 4,979 | 1,649 | 2,638 | 67 | 2,363 | 14,515 |

Note 1. The number of residential locations are adapted from each councils' growth plans.

Note 2. Land units are small geographic areas that have an estimated capacity for residential locations, each territorial authority's strategic plans identify development cells for residential growth.

Table 4.4 demonstrates that the overall the release of land is not smoothed and under constrained conditions the competition between development cells can limit the option for HAs particular in periods when fewer new Land Units are available, for example, 2017 to 2020 and then again between 2020 and 2025. When the options in locations close to employment are limited the HAs will need to move further from the employment centres.

### 4.9.2 Unconstrained development plan (UDP)

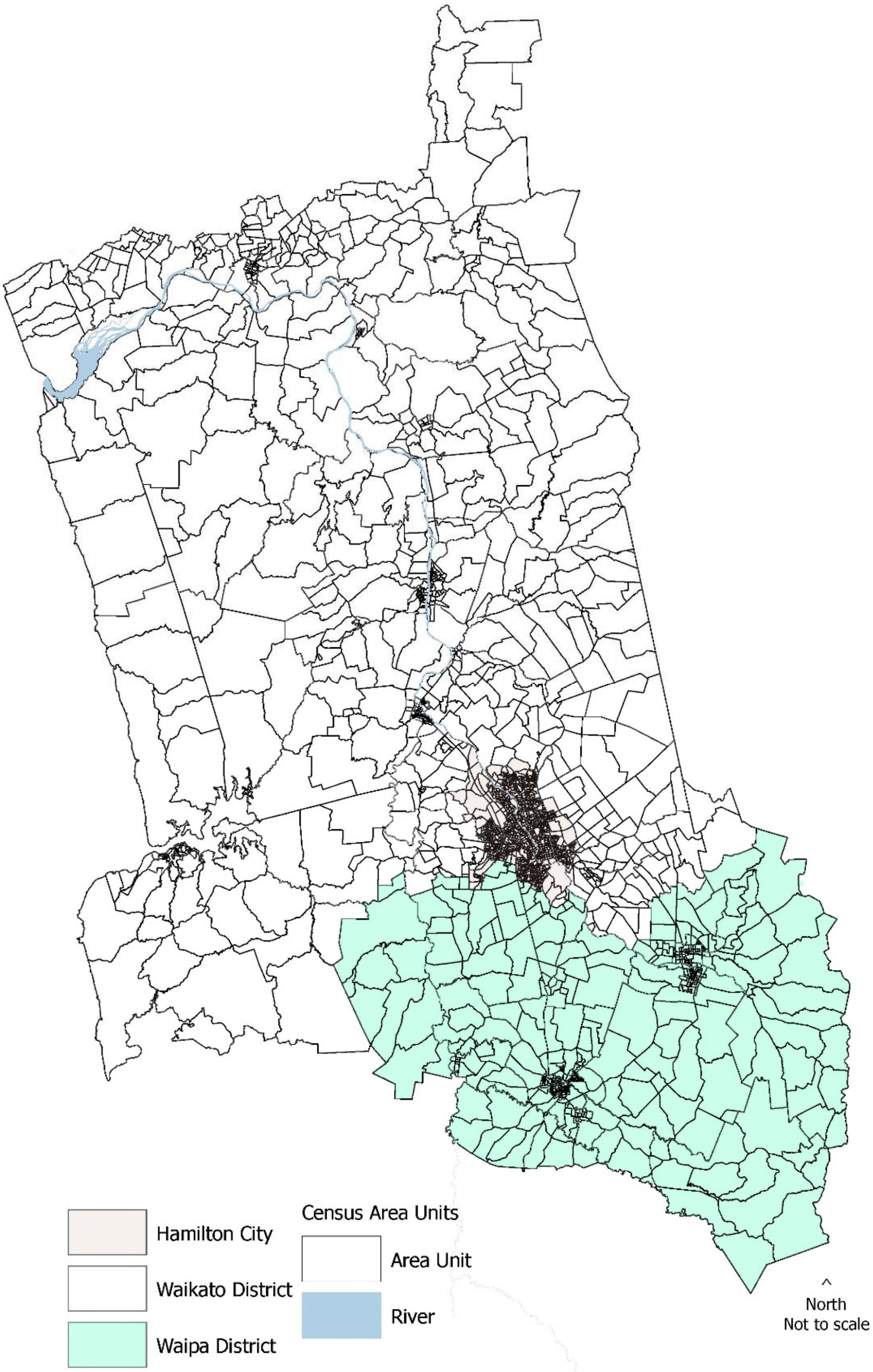
The unconstrained development plan is far more hypothetical and assumes the councils have unlimited capacity to provide infrastructure to support the development cells. In this scenario, all new land units are available for HAs to occupy from the first cycle of the model. In this scenario, the HAs have a higher number of residential location options. When development cells start to fill, the rents increase, however, there are still alternative development cells close to the employment centres.

This scenario is still limited by the quantity of land that the councils plan to develop over the next 10 years. The effect of this scenario would be more significant if 20 years' worth of land supply were introduced into the model instead of only ten years.

# 5 Calibration results

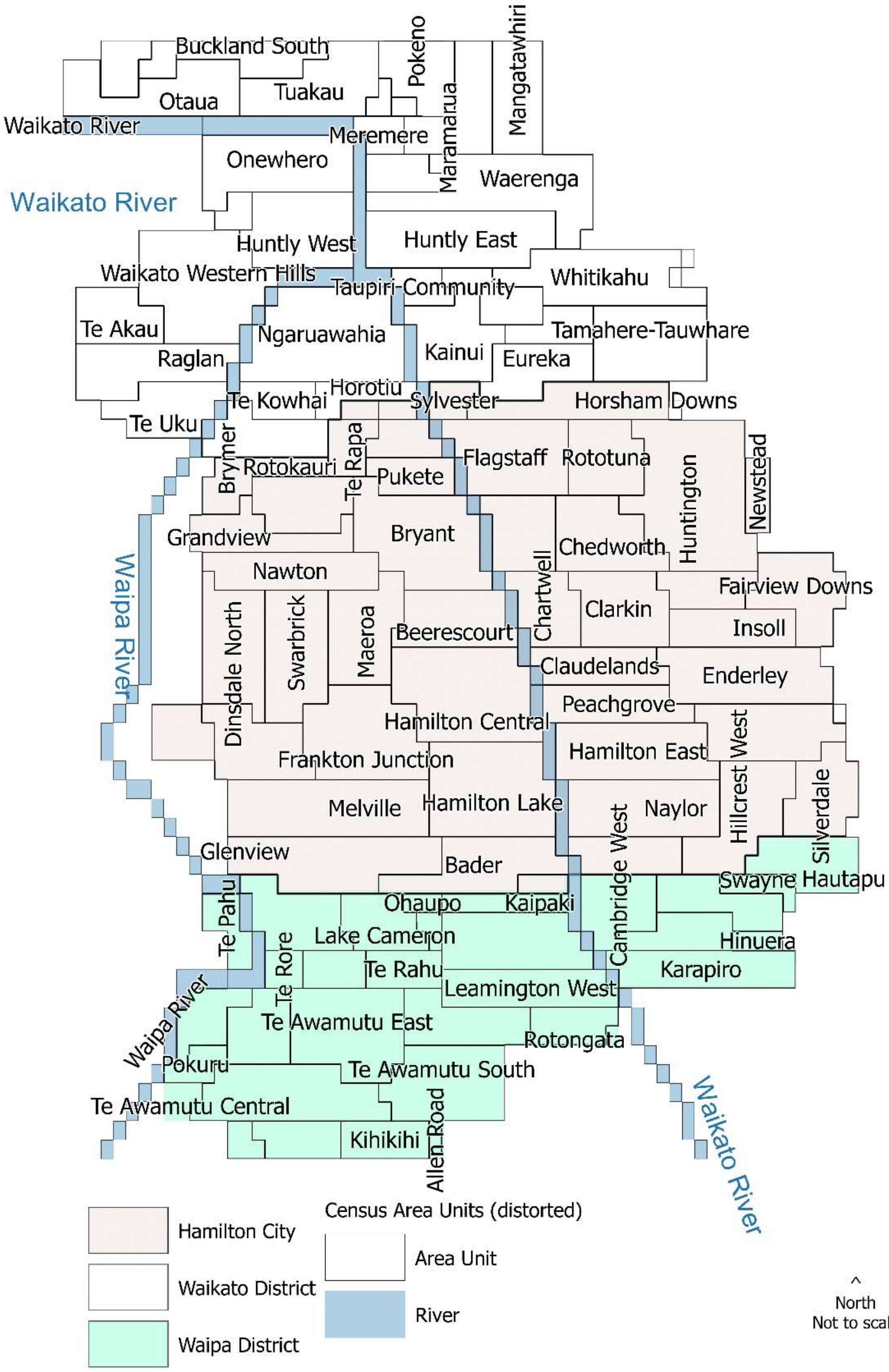
## 5.1 Overview of Area units and Meshblocks

The geographical size of meshblocks and area units varies greatly across the study area, with meshblocks ranging in size from a few hectares to several thousand hectares. Representing the model results graphically becomes a challenge. Showing the extent of the catchment requires a small scale. This thesis displays the results in one of two ways. Comparative and change over time maps are represented by maps that represent each meshblock as an equal size cell. This distorts the geographic relationships between a meshblock and its neighbours. However, care has been taken to represent these in the best possible approximate location as possible. It should be noted that these meshblocks of equal area are not used in the model itself and are only used to present the results. Alternatively, some results are best represented using the unmodified meshblocks and AUs. Map 2 shows the AUs in their unmodified state. At this scale, AUs within urban areas are indistinguishable from each other. When the meshblocks are modified and represented in equal area cells, such as in Map 3, the size of each area unit depends on the number of meshblocks it contains and individual meshblocks become identifiable. Consequently, the size of Hamilton City in Map 3 is larger than the rural Waikato and Waipa councils.




Map 2.Statistics New Zealand Area Units (2013).

Note. Area Units sourced from Statistics New Zealand geographic boundary files.



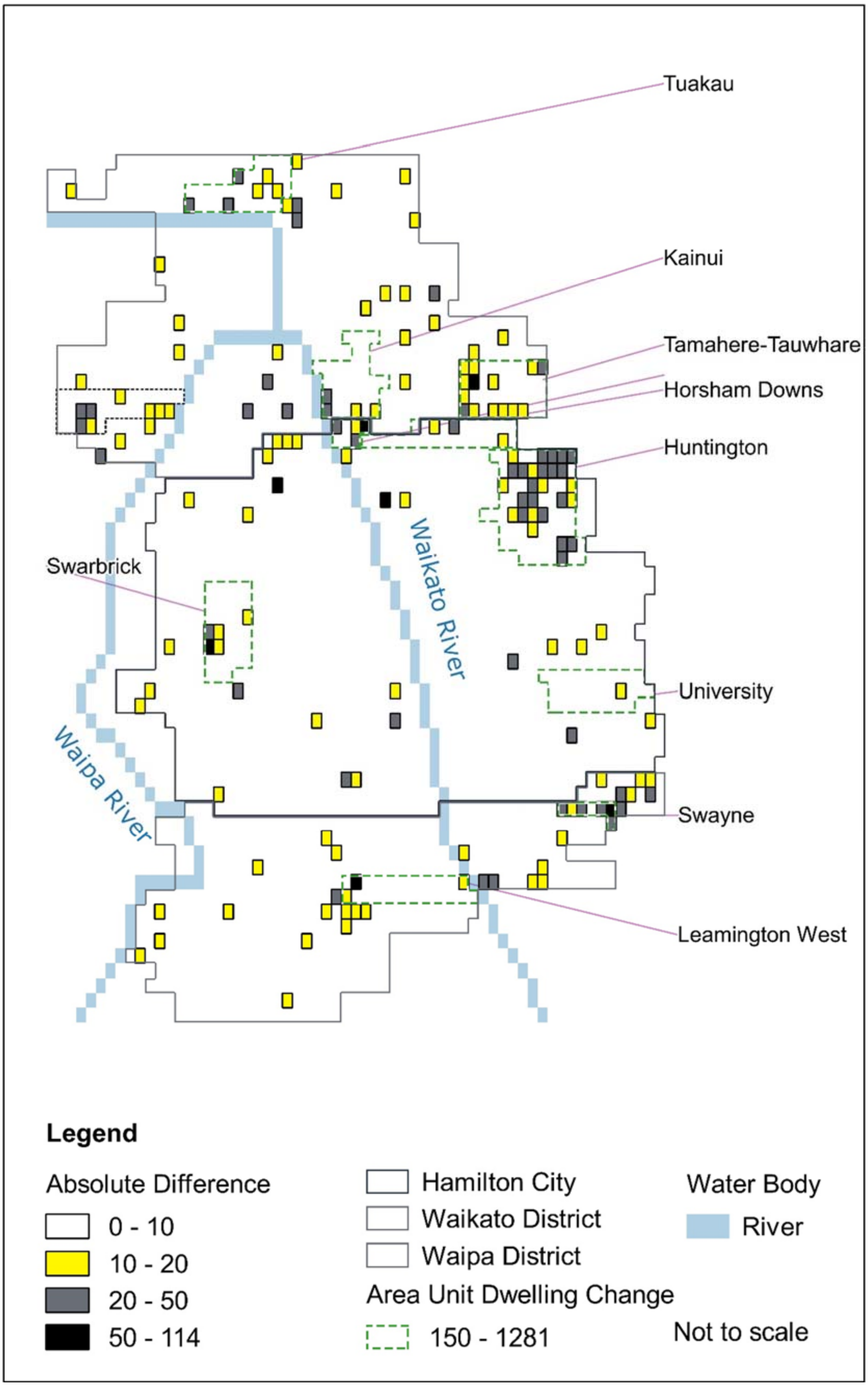

Map 2. Modified Area Units.

Note. Area Unit data sourced from Statistics New Zealand and modified for the purpose of result presentation.

## 5.2 Calibration results

The model was calibrated using the 2006 census dwelling counts and run for seven time steps to 2013. On the first run of the calibration model, the Root Mean Square Error (RMSE) at the meshblock level was 9.71. This could be interpreted as the average difference between the number of households in each meshblock from the model, and the actual number of households observed in 2013 before the model is run. Adjusting the NAB resulted in the RMSE steadily decreasing to 6.47 after 20 model runs. In 2013, the average number of dwellings per neighbourhood (meshblock) is 37. The resultant meshblock error as a proportion of the average number of households in a meshblock is 16.6%.The error can also be calculated at the Area Unit level, where the AU RMSE is 90.36 with an average of 855.5 dwellings per AU, leading to a 10.56% average error at the AU level.

The land unit with the highest difference had 114 too few HAs. This land unit is located in the AU of Swayne, Cambridge, in Waipa District. In 2006 this land unit had 18 existing dwellings. It had the capacity to hold 213 dwellings for HAs to occupy. In 2013 the census result was 168 households; however, the modelled outcome was 54 HAs. Swayne corresponds to one of Waipa's identified growth areas and is one of the AUs that experienced a high growth rate between 2006 and 2013 (refer to Map 4). On Map 4, the ten AUs with the highest growth rate between the 2006 and 2013 censuses are labelled and shown in green dotted outline.

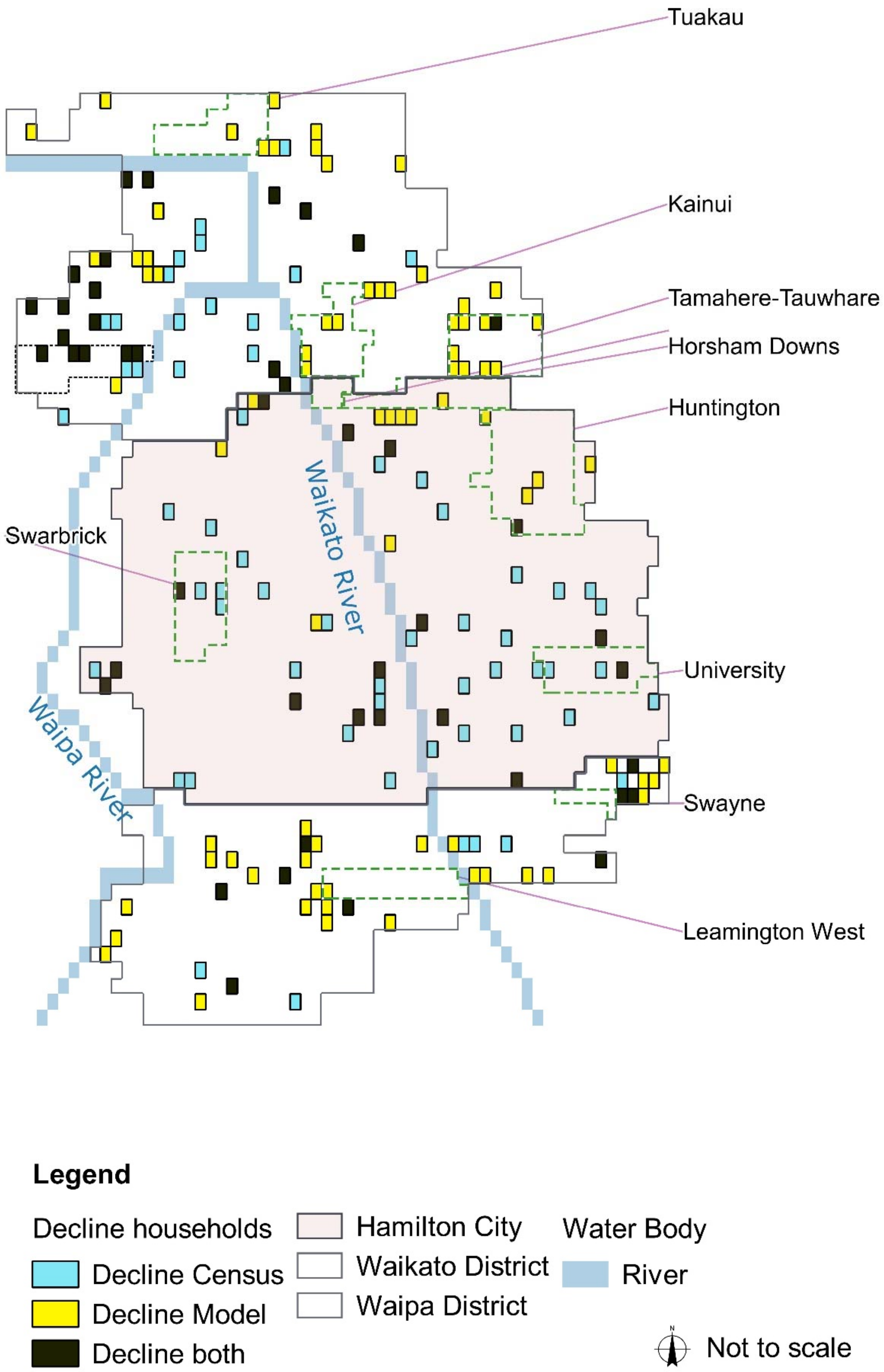



Map 3. Calibration results, 2006 - 2013.

Note. Statistics New Zealand meshblock numbers are rounded to base 3.

It is reported that population or household projection accuracy decreases as population and geographic areas decline (Rayer & Smith, 2010; Statistics New Zealand, 2008). Rapid increase or decrease in populations in the base period are also shown to have a negative association with accuracy (Rayer & Smith, 2010). The meshblocks with highest absolute error are shown on Map 4 as black cells. This shows that areas with high errors are associated with high growth in the basal years 2006 to 2013. The calibration period also coincides with some significant amenity and infrastructure changes taking place, in North Hamilton in particular. This is a challenge as the model is calibrated using an amenity value, yet over the calibration period, the real amenity is likely to be changing. Huntington and Horsham Downs AUs experienced retail growth, significant suburban expansion and the opening of a new primary school. The area with the highest amenity change is to the West of Huntington and Horsham Downs in the form of a large retail complex, light industrial expansion and new employment opportunities.

One of the unique attributes of agent-based models is that outputs are non-linear (Michael Batty & Torrens, 2005; Castle & Crooks, 2006). This is particularly useful when dealing with small areas that experience marginal or negative growth. In the cases of negative growth, linear household projections have the potential to continue decreasing household numbers. In such cases, in reality, we can expect that property prices would decrease and the rate of decrease can be arrested or even reversed as neighbourhood perceptions can change quite readily. In the calibration process, it is of interest to know how the agent-based model has managed the case of negative growth.



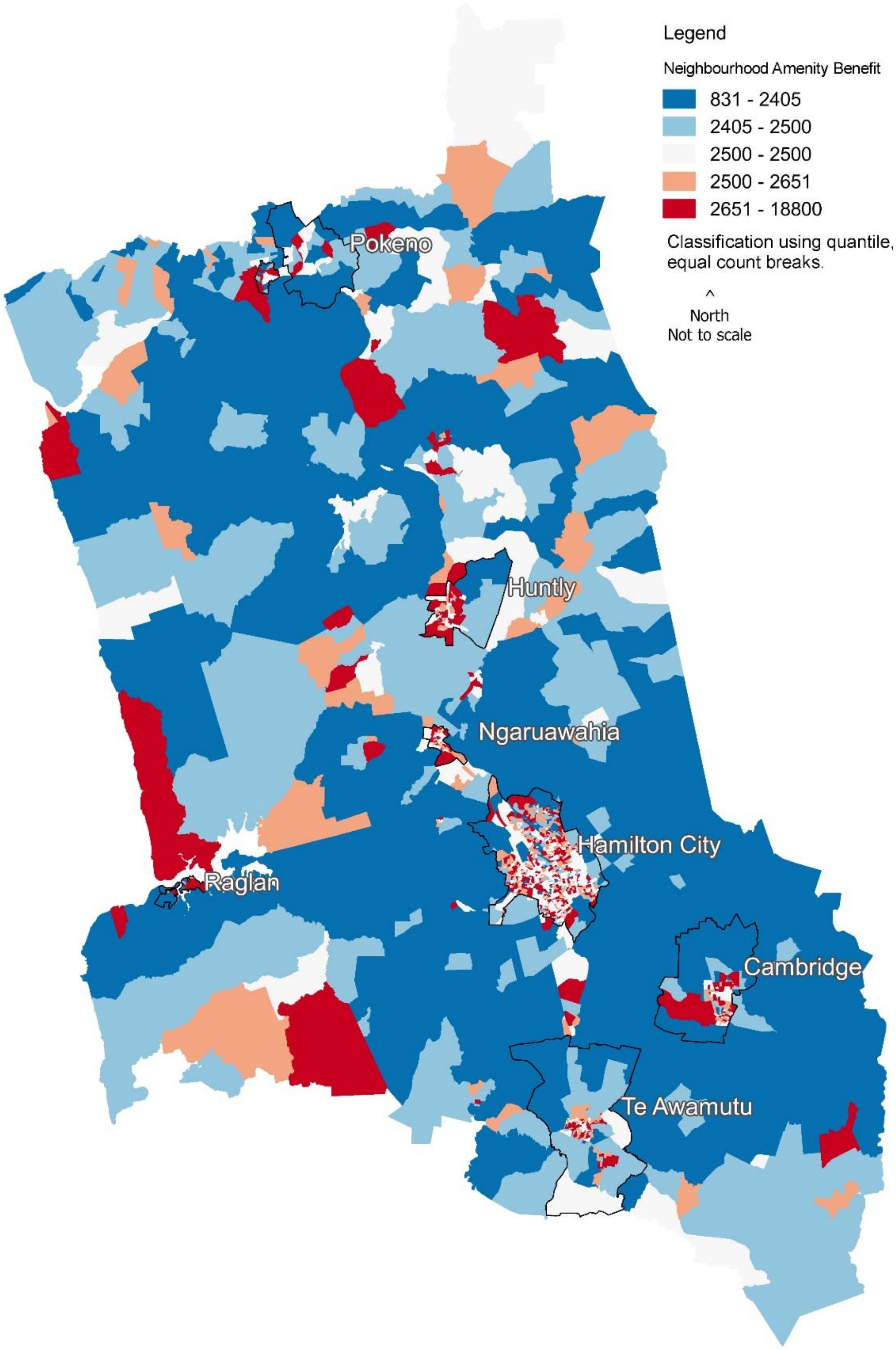

Map 4. Meshblocks with household decline, in the calibration period 2006 to 2013.

In the census period, 2006 to 2013, 105 meshblocks experienced a decline in household numbers. The meshblock that experienced the largest decline had a total of 735 fewer households in 2013 than 2006. The model results showed declines in 122 meshblocks with a total reduction of 868 households. The model is accounting for household declines in approximately the right quantities, however, the spatial distribution of the declines differ from the observed distribution. There are 45 meshblocks where both the model and the census count declined, as shown in Map 5. The meshblocks that had declined in both the model and the census are shown as black cells. The black cells are distributed across all three councils. However, there are more meshblocks with declines in the Western Waikato. These are remote areas with geographically large meshblocks and small populations. The meshblocks with a decline recorded in the model only, shown as yellow cells, are more prevalent in the rural areas. The meshblocks with declines recorded in the census, but not by the model are shown as light blue cells, are predominantly distributed across Hamilton City.

Although the average error is 16.6% at the meshblock level, the distribution of the errors indicates that the areas of high growth are more difficult to calibrate, resulting in clusters of meshblocks with high error (Map 4). This is, however, consistent with other studies (Rayer & Smith, 2010; Statistics New Zealand, 2008). Areas of high growth are likely to be associated with changes in the neighbourhood amenity value. High growth areas could be the new and fashionable places to live and move to, or in contrast could be lower cost, mass housing developments with an appeal to lower income families. The resultant change in amenity value from 2006 to 2013 cannot be calculated and thus is not incorporated in the calibration process. The change in property values can be considered an indication of the change in neighbourhood appeal, with areas of higher gain in property value associated with positive amenity changes. Property value change was not taken into account in the calibration process.

In this study, it is noticeable that the areas of high growth and highest error are in close proximity to areas that have experienced development in the employment and retail sectors in the northern part of Hamilton. The other areas of high growth are situated to the South East of Hamilton in Tamahere, Waikato and Swayne, North Cambridge. These are characterised by good access to the city, large properties with semi-rural outlook and are some of the most sought after residential areas in the region.

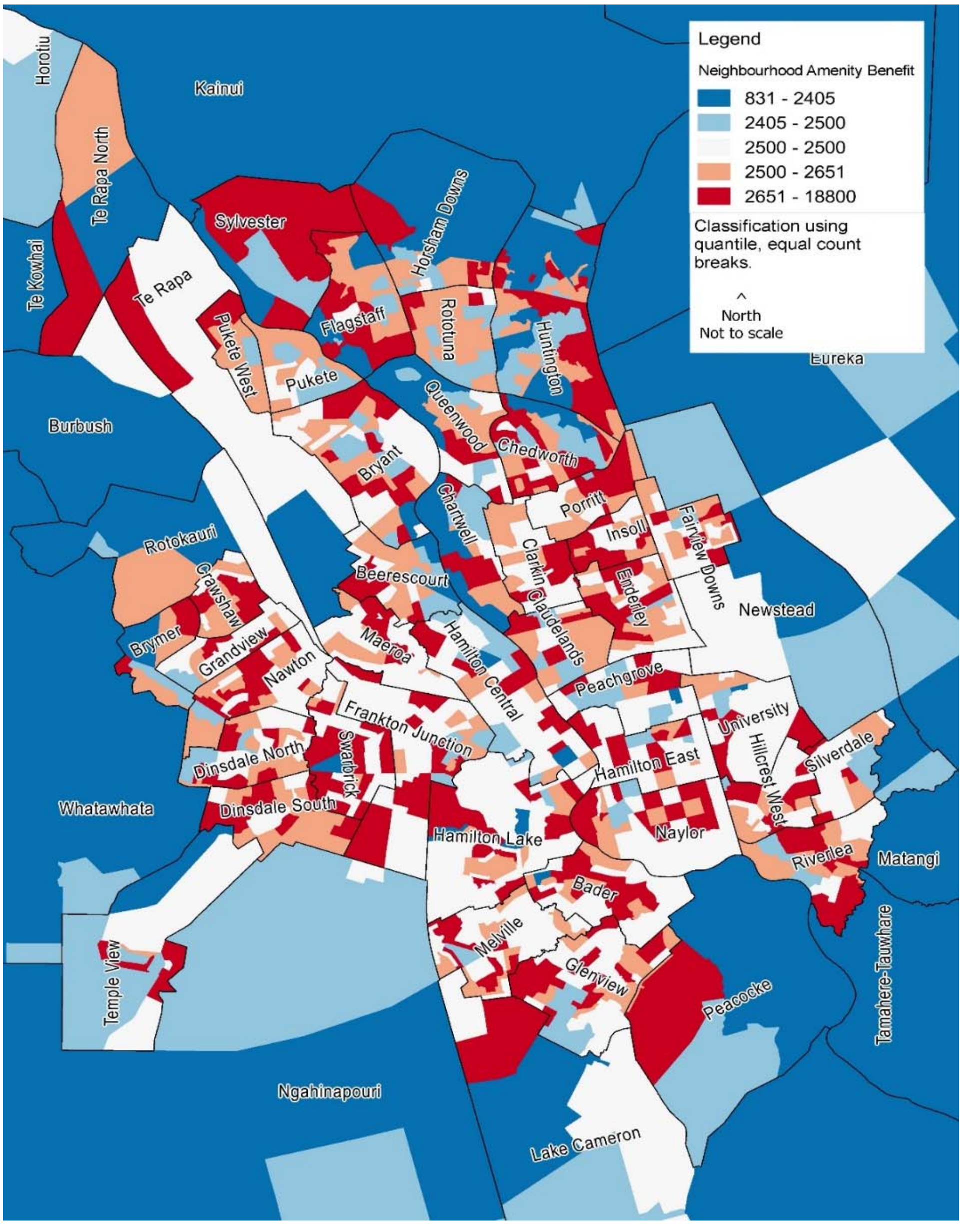
The calibration could be improved through analysis of the amount that each locality's NAB was adjusted. The absolute differences between the model result and the census result have a few outlier values and there is potential to reduce the error by investigating and manually adjusting the NAB of the outliers. It is likely that the automated adjustment is too coarse and a few particular localities consistently overshoot and then undershoot the optimal value on each subsequent run of the auto-calibration calculation.




Map 5. Neighbourhood amenity benefit, at the end of calibration, 2013.

Note. Colour ranges are based on quantile (equal count) classification.

On Map 6 and Map 7, the calibration sequence resulted in increasing the NAB values in areas where more household agents located than were counted in the census. Relative to the initial starting value of 2500, the red areas have an increased NAB, and are mostly situated in or adjacent to the major urban areas. The calibration sequence did not alter any areas shown as white, as in each run of the calibration the ‘correct’ number of household agents relocated to these areas. The areas indicated in blue colour are areas with NAB values that were reduced during calibration. These are areas that required decreasing NAB value in order to ‘attract’ the correct number of household agents. The blue areas are predominantly in the rural areas. It is interesting to note that the darker blue areas, with higher NAB, tend to be in the areas surrounding the main towns.

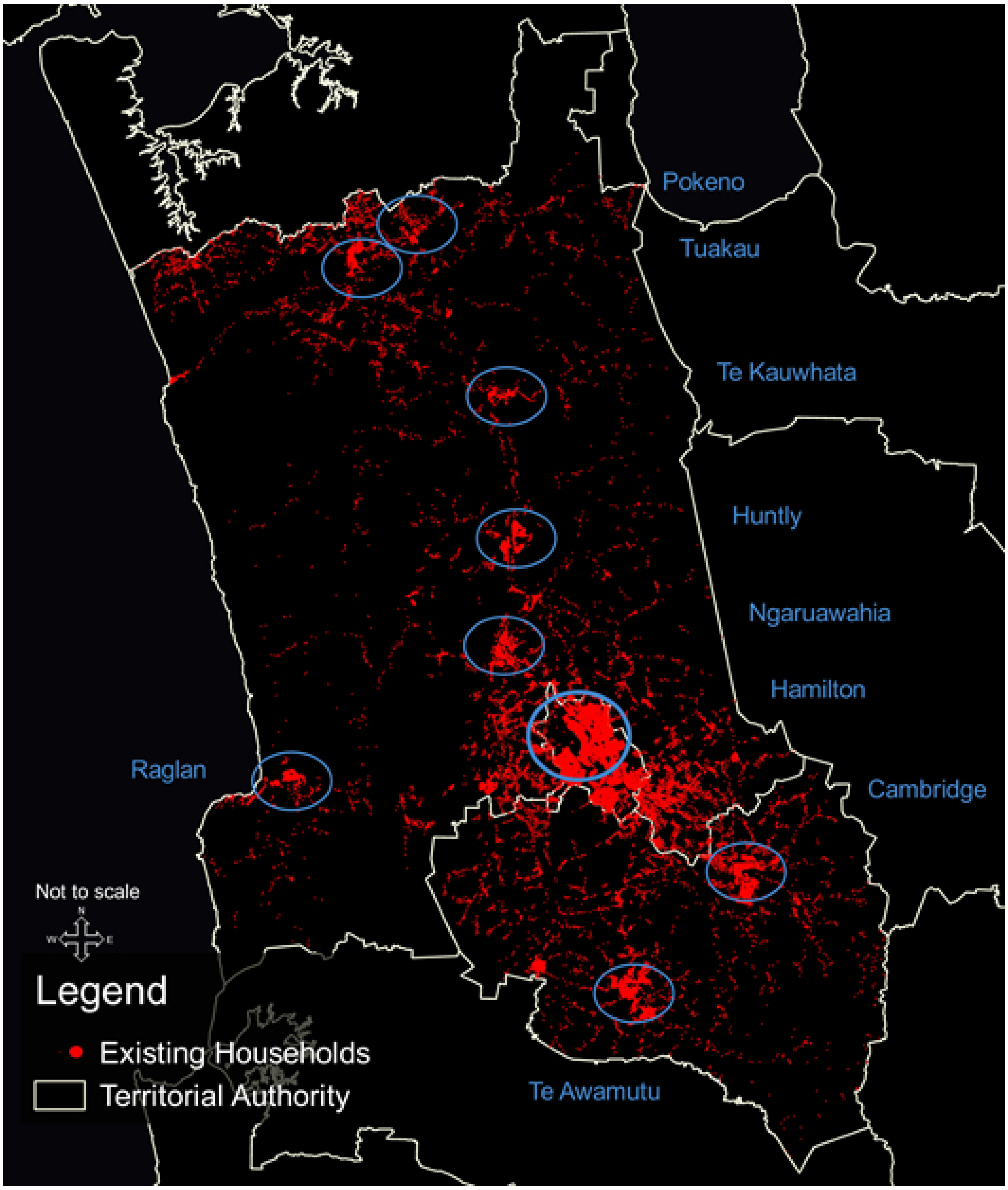



Map 6. Neighbourhood amenity benefit for Hamilton City, at the end of the calibration.

Note. Colour ranges are based on quantile (equal count) classification.

# 6 Results

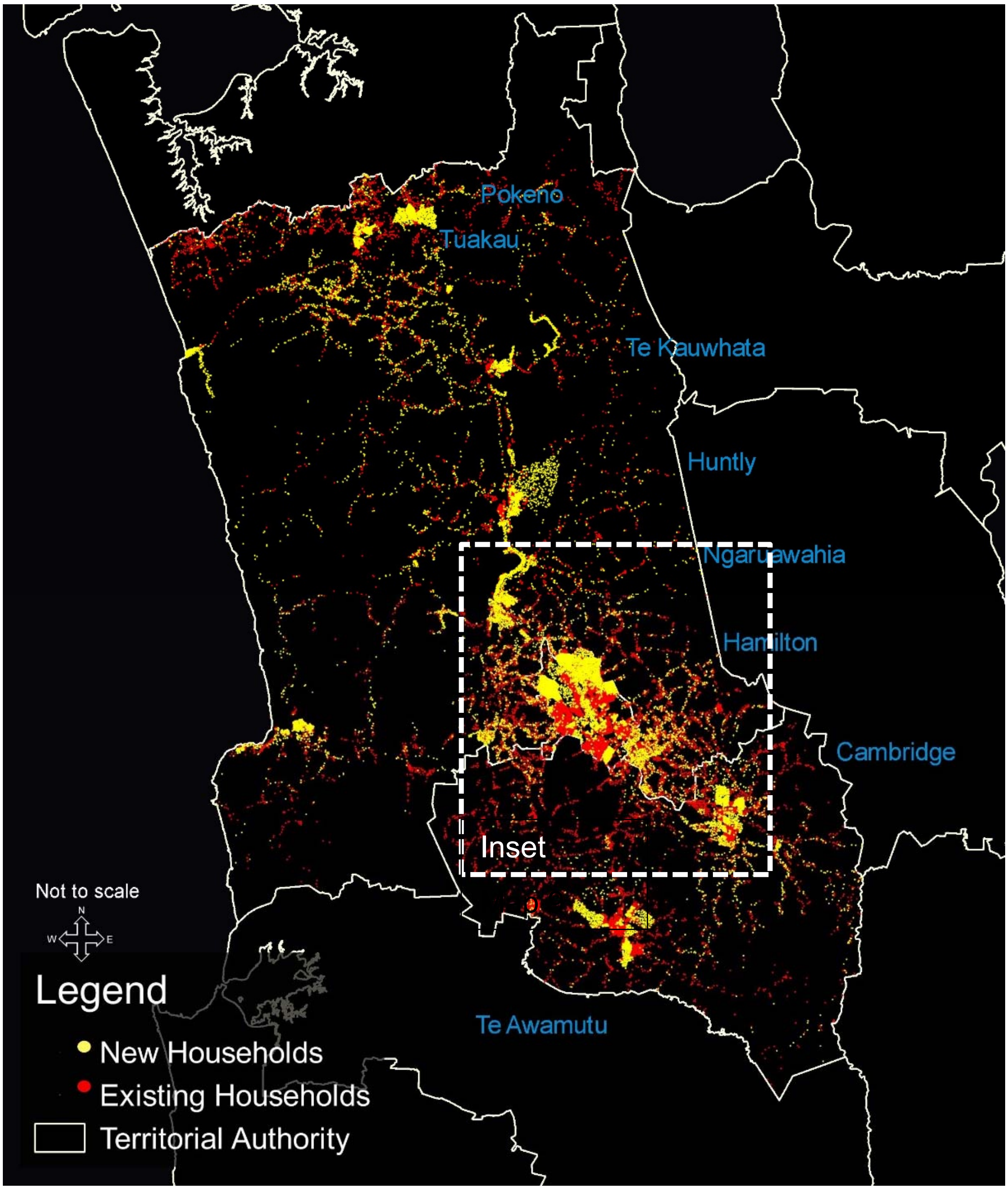
The model produces a file holding the record of each HAs location at each time step of the model. This file is used to calculate how many HAs are located in each meshblock. Net changes and differences between input variations have been analysed below.




Map 7. Household distribution in 2013 and location of main towns prior to initiating the model.

Note. Starting households locations are derived from property information supplied by Waikato Regional Council 2013.

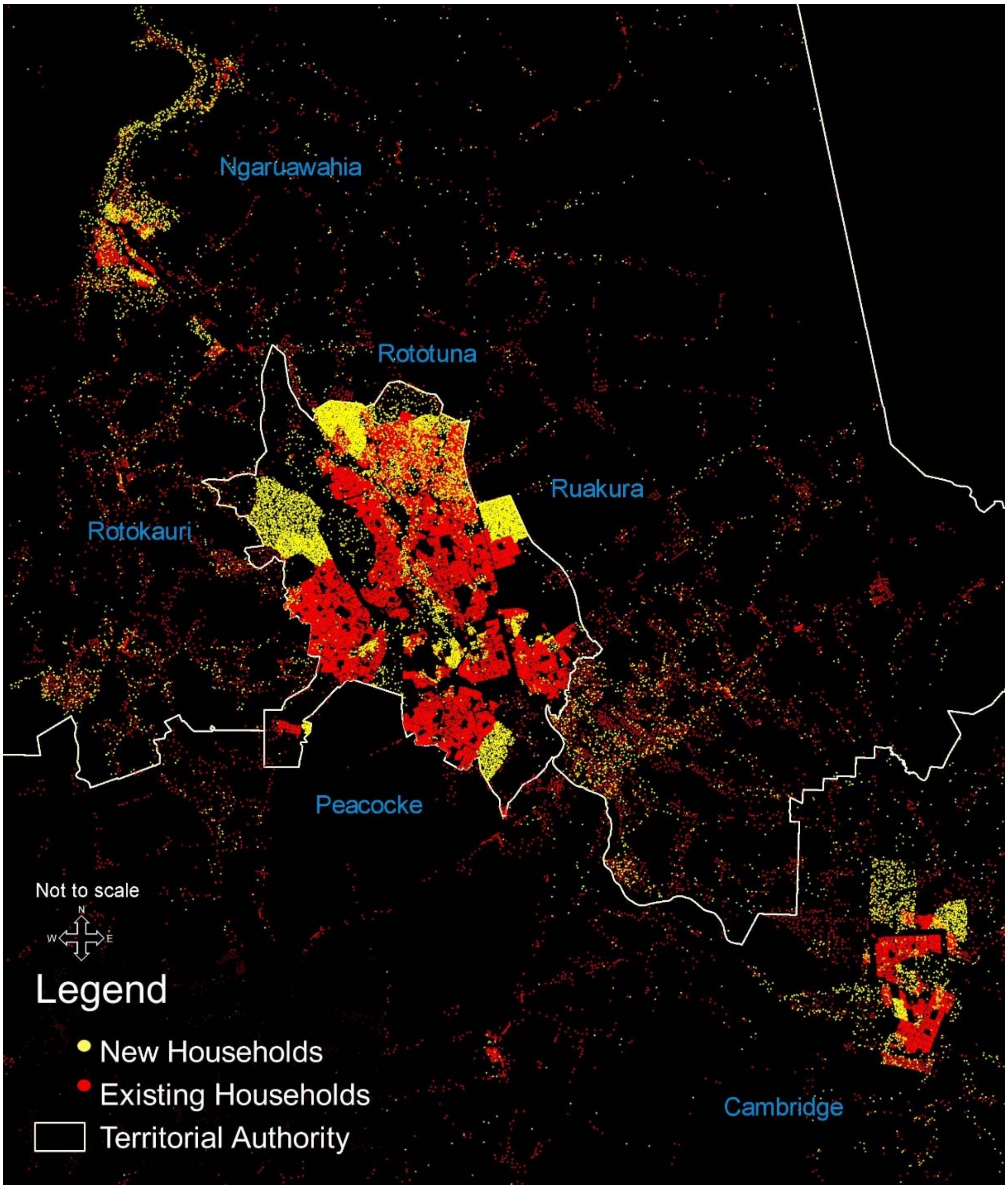
Map 9 and Map 10 are used to present results of this agent-based model and illustrate the distribution of households in 2013 and 2025 respectively, with map 9 showing results of the CDP scenario. Map 8 shows the starting point of the model (2013) and all existing households are shown as red coloured dots. The distribution pattern shows a concentration of households in Hamilton City, a number of towns and numerous small settlements. The inter-relationships of these satellite towns are visible through the linkages and developments along the key transport networks, looking like threads interconnecting the towns. The main towns are circled in blue.




Map 8. Distribution of new household agents in 2025 under the constrained development plan scenario.

Note. The precise location of the new HAs is not known. In urban areas, each HA's pseudo location is a randomly calculated point within the meshblock. In the rural areas, the pseudo HA locations are randomly located within the meshblock and within 100m of a road.

Note 2. The yellow dots are not indicating net change, it is possible that a new agent takes the place of an existing agent and as such, there is no net gain.




Map 9. New household agents located in Hamilton and surrounds in 2025 under the constrained development plan scenario.

Note. This map is an inset of Map 9.

Note 2. The precise location of the new HAs is not known. In urban areas, each HA's pseudo location is a randomly calculated point within the meshblock. In the rural areas, the pseudo HA locations are randomly located within the meshblock and within 100m of a road.

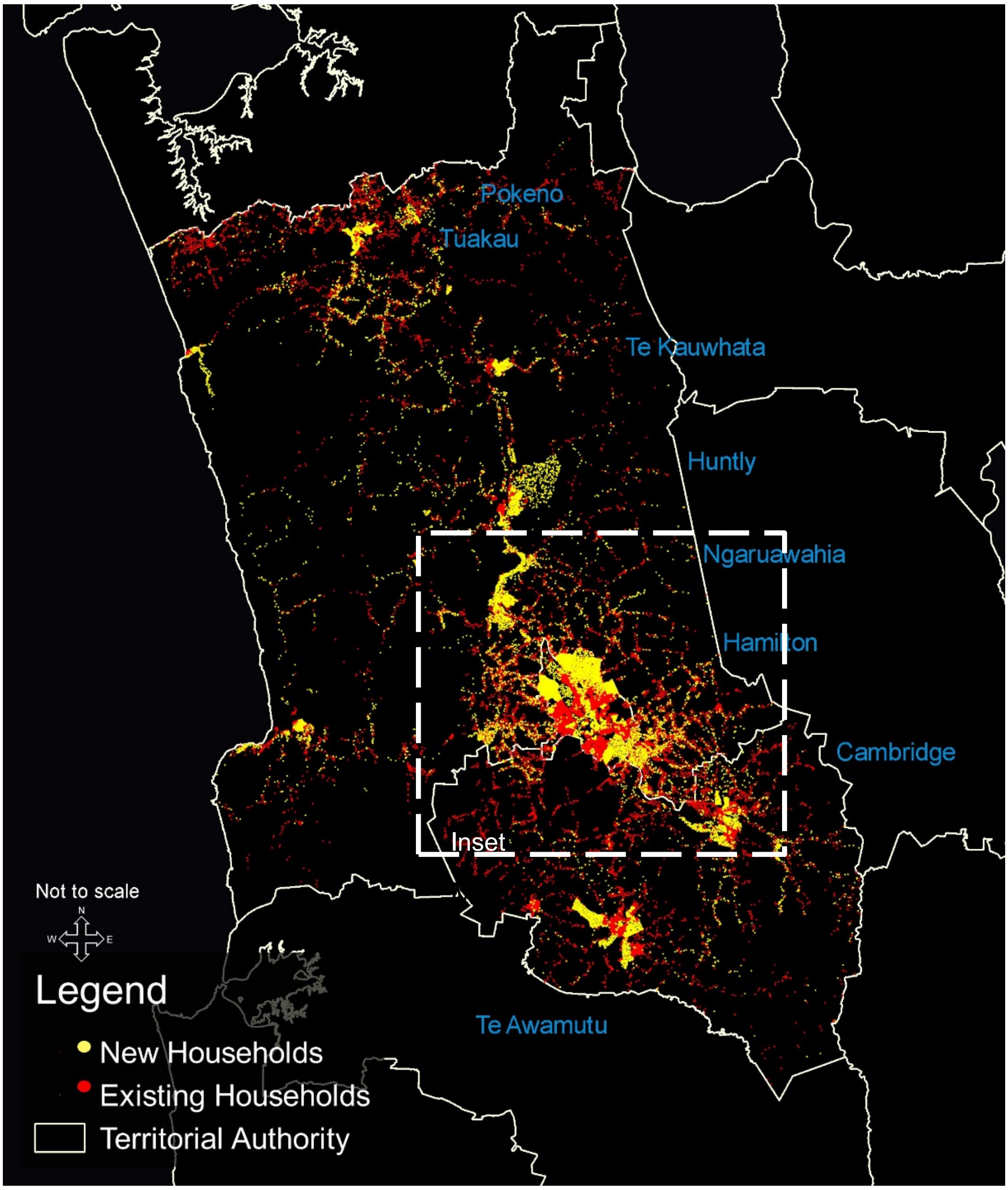
Note 3. The yellow dots are not indicating net change, it is possible that a new agent takes the place of an existing agent and as such, there is no net gain.

Map 9 and Map 10 show the end-point results of the constrained development plan scenario (in 2025). The location of all the new HAs are shown as the yellow dots, shown superimposed on top of the red existing HAs from Map 8. The intensity of development in this scenario can be clearly seen on Map 9 as bright yellow areas particularly in the Hamilton area highlighted in the Map 10 inset. The areas of most intense settlement are located in the areas identified by the councils as development cells. The development cells are all located on the margins of the urban areas and allow for the expansion of the residential areas. This is an expected outcome as these areas will present the household agents with good neighbourhood values, comparatively low rent and good transport routes to the places of employment which are predominantly located in Hamilton and the larger towns. The development cells are also particularly noticeable in the towns of Te Awamutu and Cambridge in the South.

Intensification within existing urban areas can be identified as a mixture of existing (red) and new (yellow) HAs, more distinctly noticeable on the larger scale Map 10. Ngaruawahia and the northeastern suburbs of Hamilton display this pattern of intensification. This follows an expected suburban development life cycle where initially a lot of development takes place in a suburb and the development rate progressively slows. Intensification can also occur where the initial development results in large residential sections following which these can be progressively subdivided and a more intensive urban form develops. Ngaruawahia is a good example of this intensification process.

The increase in HA numbers in the rural areas is more widespread. There are distinguishable settlement areas in the area between Tuakau, Pokeno and Te Kauwhata as well as south east of Hamilton through the Cambridge area. These are expected development patterns as they are close to both employment centres and transport. These areas also correspond to productive agricultural areas. These areas in the north and surrounding Cambridge have more intensive agriculture on smaller farms.

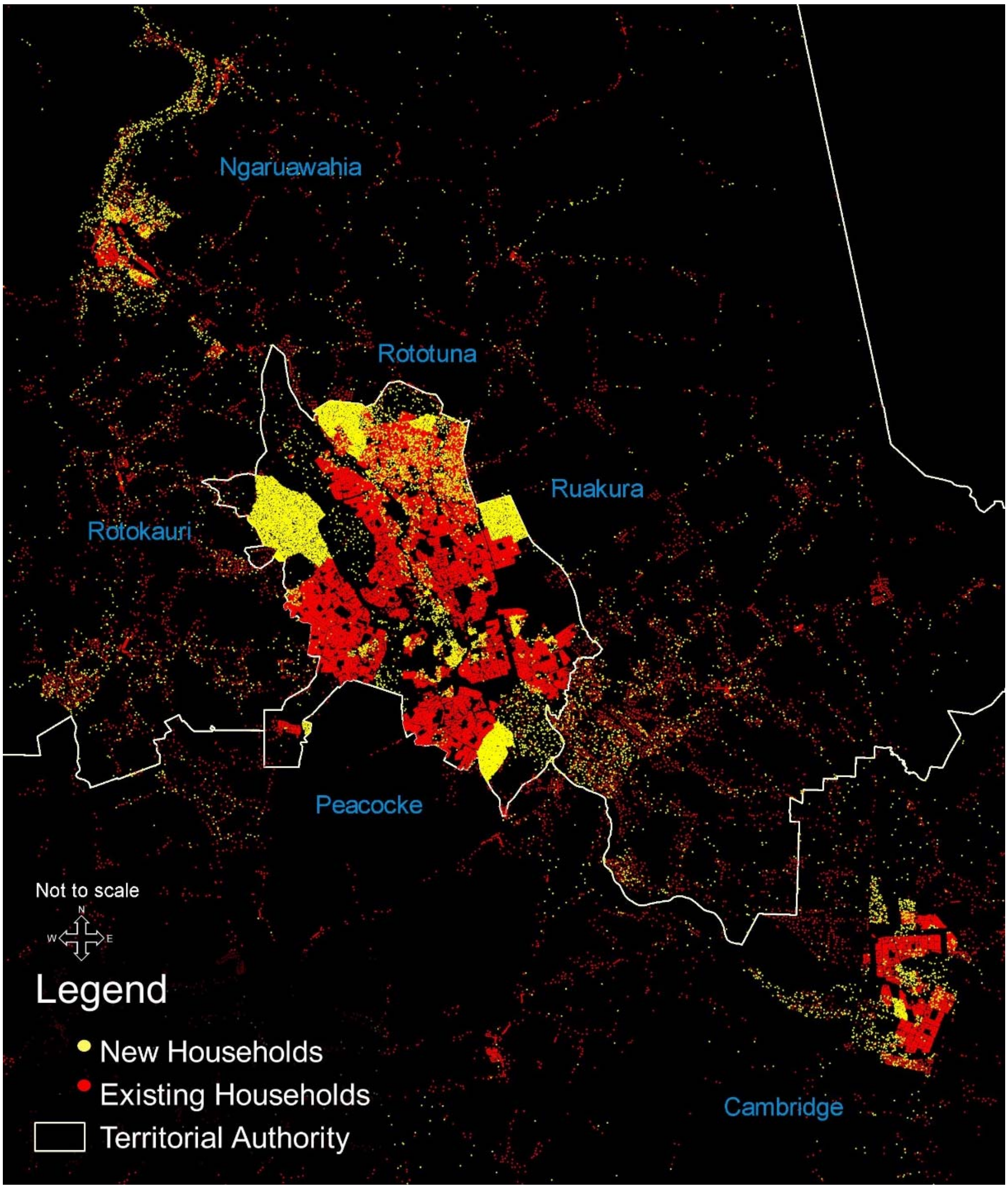
The scenario based on unconstrained development plans produces a variation in the overall settlement pattern, Map 11 and Map 12.




Map 10. Distribution of household agents in 2015 under the unconstrained development plan scenario.

Note 1. The precise location of the new HAs is not known. In urban areas, each HA's pseudo location is a randomly calculated point within the meshblock. In the rural areas, the pseudo HA locations are randomly located within the meshblock and within 100m of a road.

Note 2. The yellow dots are not indicating net change, it is possible that a new agent takes the place of an existing agent and as such, there is no net gain.




Map 11. New household agent located in Hamilton and surrounds in 2025 under the unconstrained development plan scenario.

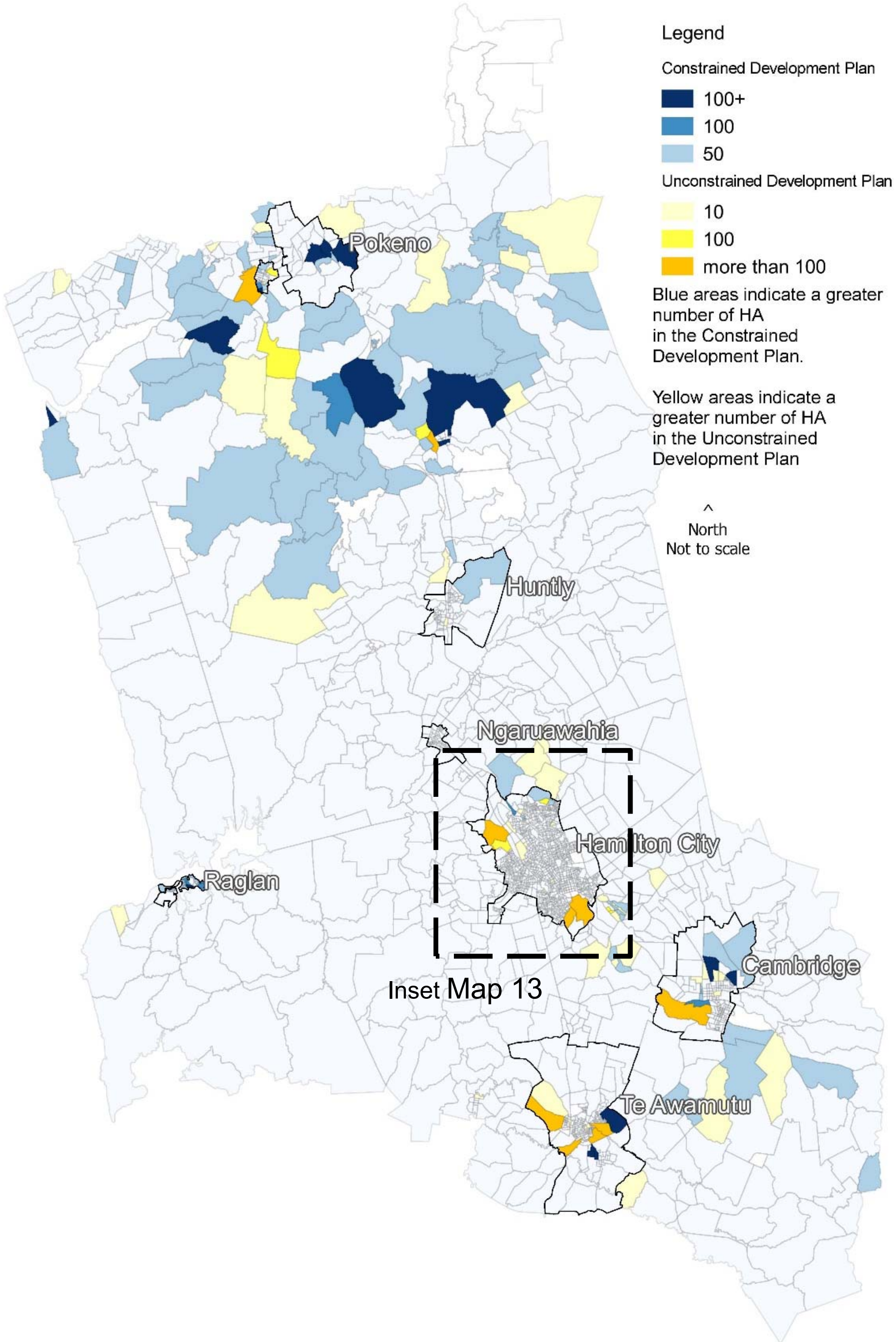
Note1. This map is an inset of Map 11.

Note 2. The precise location of the new HAs is not known. In urban areas, each HA's pseudo location is a randomly calculated point within the meshblock. In the rural areas, the pseudo HA locations are randomly located within the meshblock and within 100m of a road.

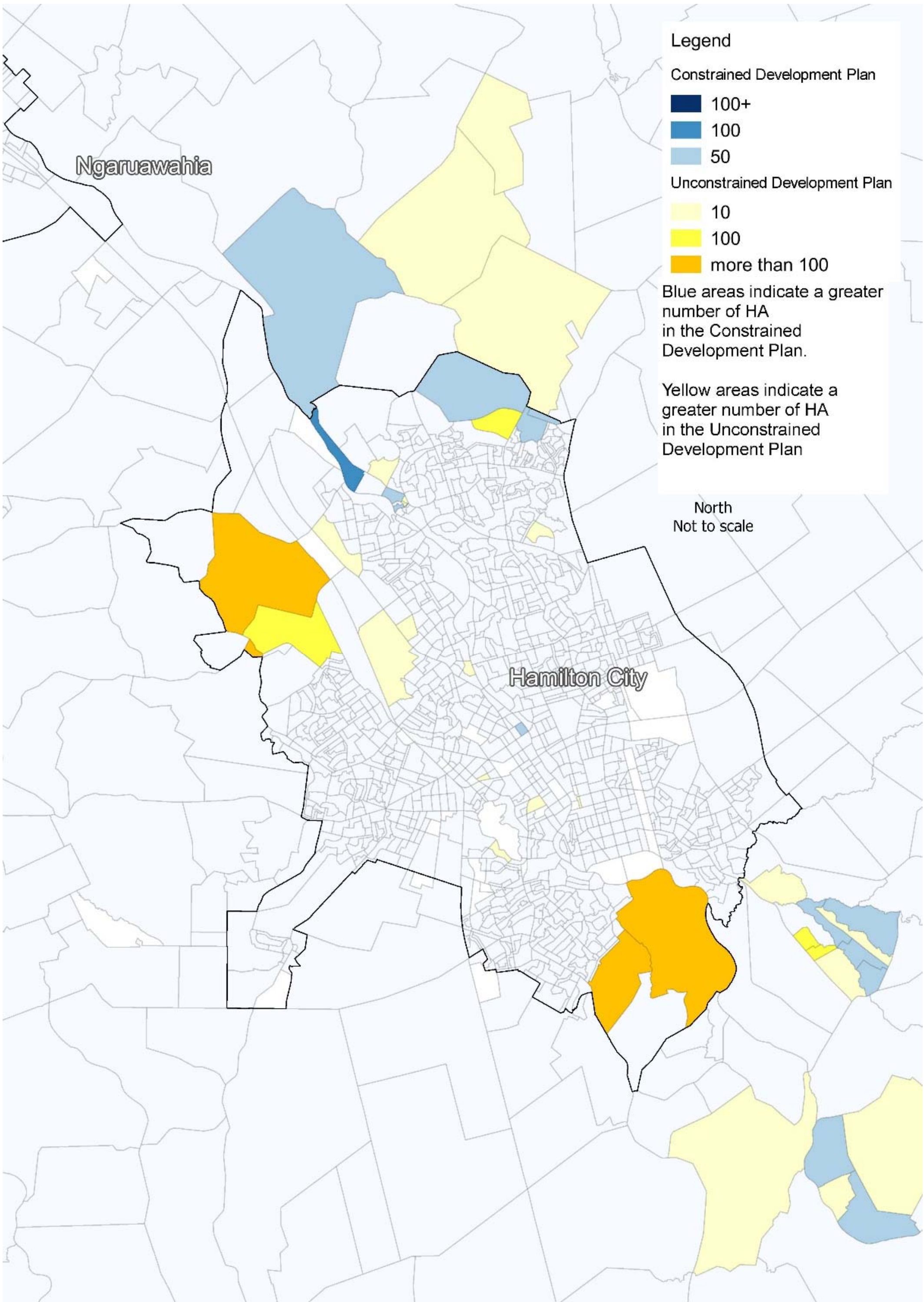
Note 3. The yellow dots are not indicating net change, it is possible that a new agent takes the place of an existing agent and as such, there is no net gain.

As the predominant areas of employment are situated in Hamilton, the motivations are for HAs to minimise costs and seek residential locations in or near to Hamilton. Providing more residential locations from the start of the model allows agents greater number of choices of residential location. Locations in the development cells and close to areas of employment offer opportunities for HAs to reduce costs. The result of this is a higher concentration of HA in the development cells and lower concentrations in the outer rural areas.

Choropleth Maps 14 and 15 accentuate the differences between the CDP and UDP results. The blue shaded meshblocks are the locations with higher numbers of HAs in the CDP scenario. Light blue indicates a difference of fewer than 50 HAs. Mid blue areas have between 50 and 100 more HAs than the UDP. The dark areas have a difference between the CDP and UDP of more than 100 HAs. The difference in population distribution is predominant in the North. Under the UDP scenario, the development cells of Hamilton have a higher number of HAs than the CDP. These areas are shaded light yellow to a darker yellow/orange indicating an increasing range of 0 – 50, 50 – 100 and over 100 HA.




Map 12. Difference between CDP and UDP, choropleth map showing the difference in the number of HA in each scenario.




Map 13. Hamilton area CDP vs UDP, choropleth map showing the difference in number of HA in each scenario

Note. This map is an inset of Map 13

Maps 10 through to 15 show the differences in the distribution patterns associated with different planning strategies. In the following section table, 10.1 and figures 10.1 through to 10.3 outline the numerical difference between the CDP and UDP scenarios.

Table 6.1. Number of household agents in 2025 under CDP and UDP scenarios.

| | Number of HAs in 2025 | | |
|---|---|---|---|
| | Waikato | Hamilton | Waipa |
| CDP | 36,042 | 64,752 | 21,370 |
| UDP | 31,170 | 70,259 | 20,735 |
| *Difference* | *4,872* | *-5,507* | *635* |

Note. Total household projection sourced from NIDEA

Table 6.1 shows that the relative movement of HAs is more towards Hamilton City under the UDP scenario than under the CDP scenario. Conversely, in the CDP scenario, the net change in HA distribution is mostly towards the Waikato District and a lesser extent to Waipa. It should be noted that, relative to the existing population size and expected growth, the impact of an additional four to five thousand households is substantial for the Waikato District.

For results validation, the model output results were compared against both Statistics New Zealand and NIDEA 2015 base household projections. NIDEA published household projections in April 2015. Statistics New Zealand released household projections in December 2015. There is uncertainty as to why the Statistics New Zealand starting household counts are higher than the projections produced by both NIDEA and this model. Statistics New Zealand produces household projections for the whole of New Zealand. In doing so, there is limited scope to consider potential policy interventions or council infrastructure investment into account (Statistics New Zealand, 2008). As such both private and council development plans could avert population growth (Statistics New Zealand, 2008). NIDEA household projections are produced at a regional level and have a significantly higher degree of input from the council planning departments. This agent-based model also has a high degree of input based on different planning policies. In all three councils, Figures 10.1 to 10.3, the Statistics New Zealand household projection starts at a higher number in 2013 with a slower growth rate and a lower household count in 2025 than the NIDEA household projection.

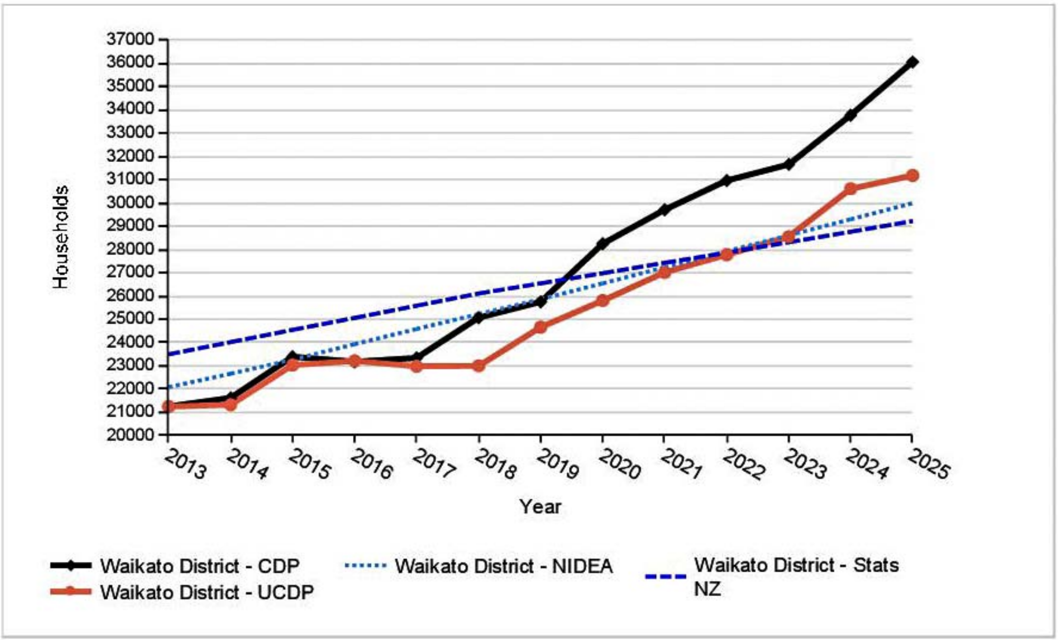



Figure 6.1. Household agent counts for CDP and UDP scenarios for the Waikato District, from 2013 to 2025.

Note 1. NIDEA projections sourced from Cameron (2015)

Note 2. Stats NZ household projection sourced from Statistics New Zealand 2015

Figure 6.1 charts the model results based on the CDP and UDP scenarios against the NIDEA and Statistics New Zealand household projection for the Waikato District. From 2013 to 2017 the outputs from the model are very similar under both of the scenarios. From 2017 onward the CDP scenario has a higher growth rate. The CDP growth rate is lower than both the Statistics New Zealand and NIDEA projections up until 2020. After 2020 the CDP growth rate increases and by 2025 there are about 7000 more households in the Waikato District than projected by either Statistics New Zealand or NIDEA.

The UDP scenario shows a similar increase in the growth rate to the CDP scenario, however, this increase does not occur until after 2023. What these results indicate is that under the CDP, the residential cost for HAs increases in Hamilton City and the HAs relocate predominantly to the Waikato and to a lesser degree Waipa.

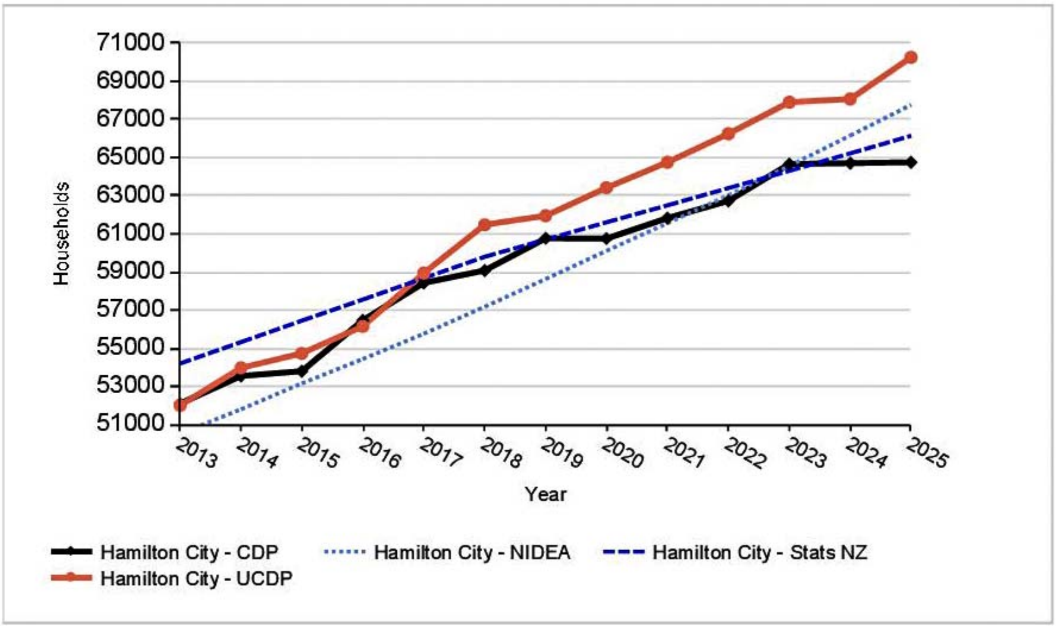



Figure 6.2. Household agent counts for CDP and UDP scenarios for Hamilton City, from 2013 to 2025.

Note 1. NIDEA projections sourced from Cameron (2015)

Note 2. Stats NZ household projection sourced from Statistics New Zealand 2015

Figure 6.2 charts the model results based on CDP and the UDP scenarios against the NIDEA and Statistics New Zealand household projections for Hamilton City. Hamilton initially tracks between the NIDEA and StatsNZ growth rates. 2017 marks the year when the effects of constraints become effective and the growth rate starts to slow down for the CDP. Under the UDP scenario, more household locations are available within the city from the starting point in the model, and a sustained growth rate is maintained up to 2025. Under the UDP scenario, around 5,500 more household agents locate in Hamilton than under the CDP scenario. The UDP scenario has about 2,300 more households than the NIDEA projection.

The CDP scenario has a growth rate marginally lower than the Statistics New Zealand household projection.  When the constraints are imposed in the city, the HA can reduce cost by locating in areas of the Waikato. When there are fewer constraints the HAs will locate in the city development cells.

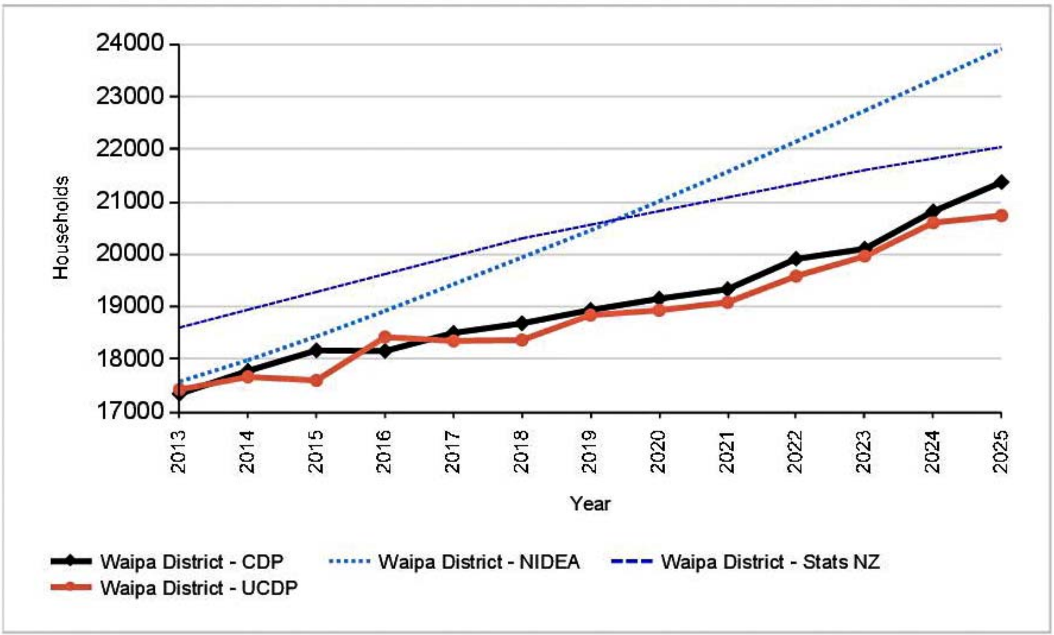



Figure 6.3. Household agent counts for CDP and UDP scenarios for Waipa district from 2013 to 2025.

Note 1. NIDEA projections sourced from Cameron (2015)

Note 2. Stats NZ household projection sourced from Statistics New Zealand 2015

Waipa, Figure 6.3, tracks at a growth rate substantially lower than the NIDEA household projection. The growth rate, i.e. the slope of the line, is similar to that of Statistics New Zealand's projection, however as noted above the two projections are not originating at the same household count. There is only a difference of 635 HA between the CDP and UDP scenarios by 2025. This indicates that Waipa's constrained development plan is not actually restrictive and does not result in HA having any cost benefit elsewhere. Further scenario testing could be useful for the Waipa planners to see if there are significant impacts if a greater number of HAs are introduced into the early time steps, particularly with constrained land supply in Hamilton City.

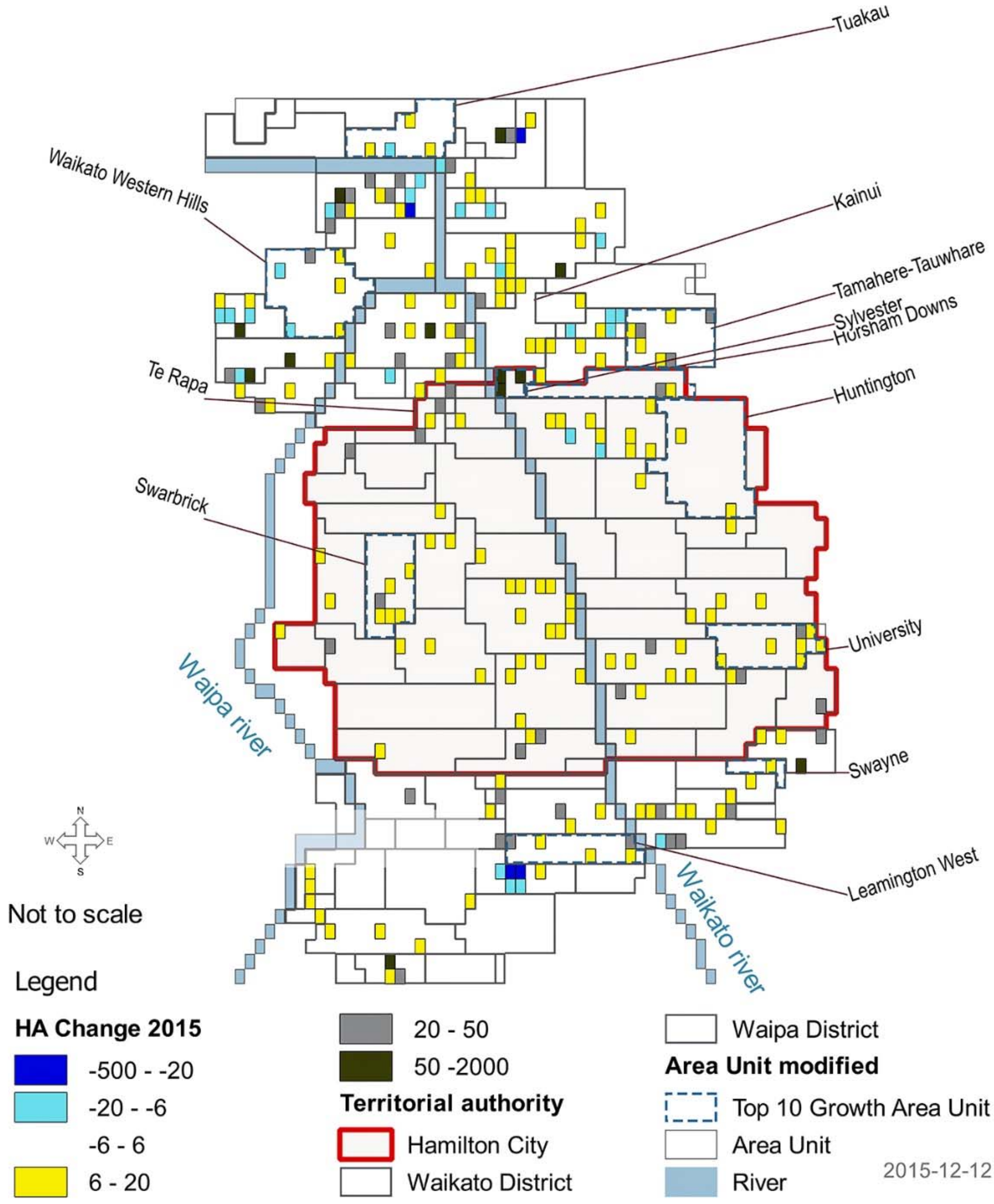
Two mechanisms are operating in the model to result in divergence of growth from the Hamilton centre. The unconstrained development plan results in a higher number of neighbourhoods from which the agents can choose residential locations. As the HAs start to move into these neighbourhoods the result is an increase in rent. If there are alternative lower cost neighbourhoods still within the city, then HAs remain near to the city. The relationship between neighbourhood capacity and the change in rent determines the flow into or away from the city.

Constraints develop as the supply of vacant land diminishes and the rents increase. The agents display a tendency to follow the path of least resistance or lowest constraints, i.e. costs. Waikato has both a high amount of vacant land and a less stringent growth planning, thus there is relatively little difference for the Waikato between CDP and UDP. Under both scenarios, Waikato can be seen as having low constraints. When Hamilton has greater constraints the HAs seek lower constraints and re-locate to Waikato District. When constraints reduce in Hamilton, the balance shifts away from Waikato District.

Overall there is a stronger relationship between Hamilton and Waikato than between either of these areas and Waipa. The differences in results between the scenarios are nearly equal and opposite for Hamilton and Waikato. Waipa has a much more slight variation between CDP and UDP. Waipa's more stringent growth strategy retains a relatively high level of constraint under either scenario and the growth rate is not greatly affected until such point as Hamilton's constraints are low enough to create a differential. A hypothetical example of this might be Hamilton providing 20 years' worth of land supply in the model time frame with Waipa not making any changes.

## 6.1 CDP change analysis

Map 8 and Map 9 suitably present the overall distribution trends for the CDP scenario. However, one purpose of this model is to drill down to the changes taking place at the neighbourhood and meshblock level.




Map 14. Short-term Difference in the number of HAs between 2013 and 2015, in the CDP scenario.

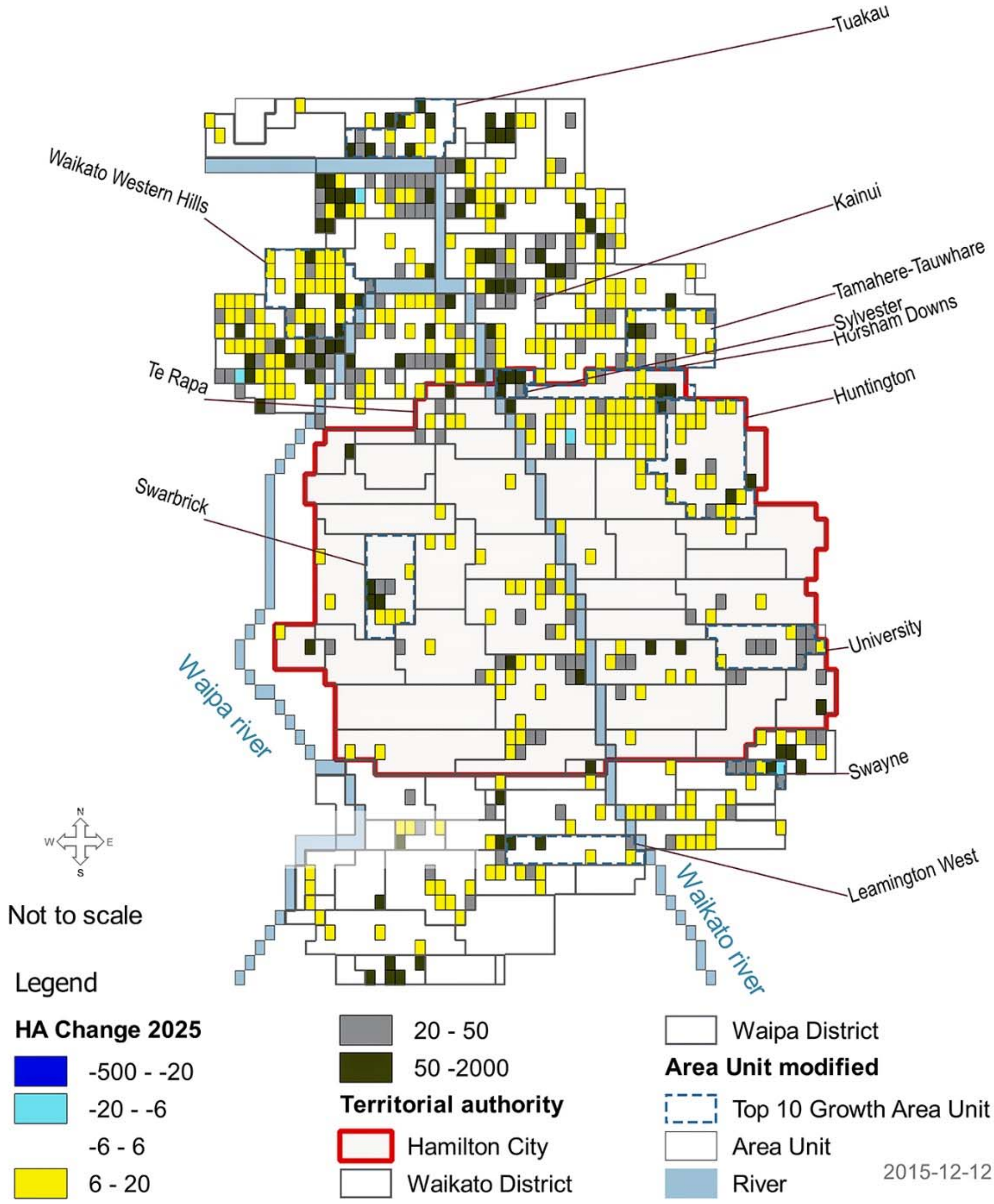
Note 1. White areas have small or no change.

Note 2. Meshblocks are modified and represented as equal area cells in order to ensure all meshblocks are visible and to ease change over time comparisons.

Note 3. Area units are constructed from the modified meshblocks, large AUs have a higher number of meshblocks.

In the period 2013 through to 2015, Hamilton experiences a net gain in HA numbers and these areas are presented as grey and yellow cells in Map 15. These yellow and grey cells are fairly evenly distributed. There are a few key meshblocks that have the highest growth (dark green). A number of these changes correspond to the top 10 AUs for growth between the 2006 and 2013 census. Most of the districts' growth over the period 2006 to 2013 took place in Sylvester, North Hamilton, and Map 10 shows a continuation of this growth up to 2015. The model outputs also indicate growth in other key areas, including Pokeno to the east of Tuakau, Ngaruawahia just to the north of Sylvester, Swayne and Leamington (neighbourhoods of Cambridge) and Kihikihi Flat (Te Awamutu).

Up to 2015, some areas show marginal declines. These are not showing signs of significant clustering and most of the light blue cells are located in areas some distance from employment centres. Thus it is likely that the HAs have opportunities closer to the centres, as at this point in the model there are a high number of vacant locations for the HAs to reduce their costs.



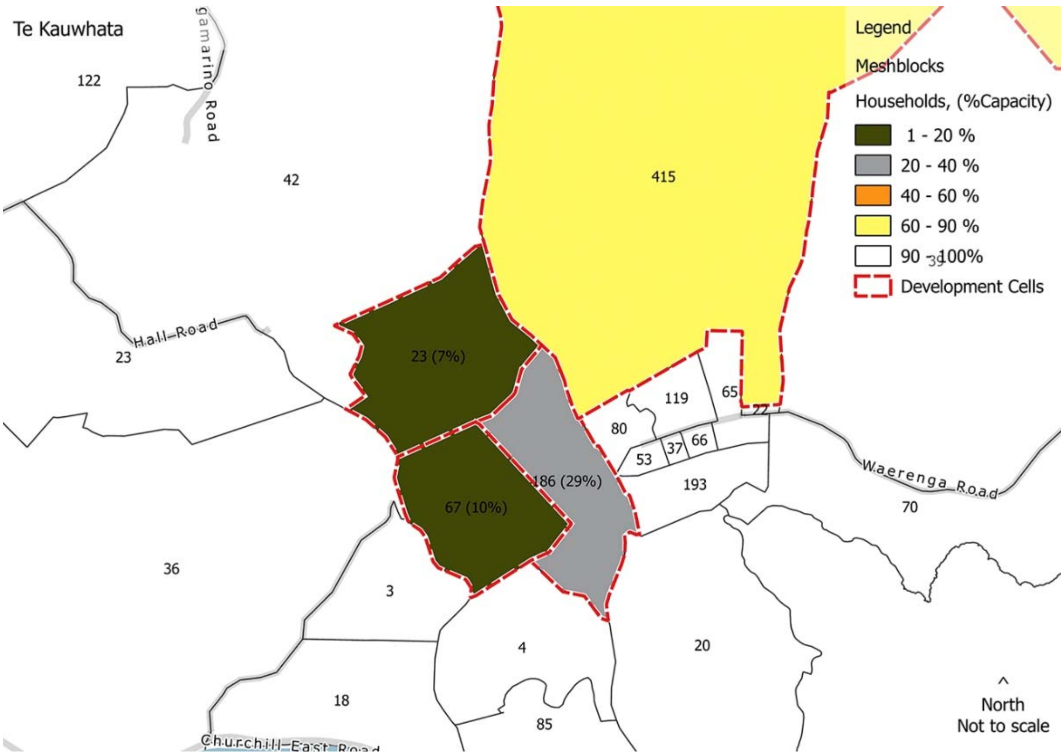

Map 15. Long-term difference in the number of HA between 2013 and 2025, in the CDP scenario.

Map 12 shows the change in HAs for each meshblock over the period 2013 to 2025. On Map 16 the general settlement is to the north of the study catchment. The meshblocks with a gain of more than 500 HA are located in Tuakau, Pokeno, Rotokauri, Newstead (Ruakura), and Peacocke, which are all areas of planned development. As previously noted, the overall net gain of agents is higher than expected in Waikato District, as seen in the difference between the CDP and the UDP. The quantity of available residential land and new houses in Hamilton is driving this growth pattern. In the Waikato, the areas of greatest gain are mostly in close proximity to the Hamilton boundary or the two fast growing towns of Tuakau and Pokeno in the North.

## 6.2 Detailed investigation of growth areas

Under the constrained development plan most of the meshblocks in the councils' development cells fill to more than 90% capacity. The following series of maps will outline areas that did not reach capacity. Appendix 2 shows the remaining development cells in more detail. Throughout Waikato and Hamilton, many of the existing neighbourhoods have the potential to subdivide and create new land units. The subdivision of land can take place, either in older neighbourhoods that historically have larger land parcels that can be subdivided into smaller properties or in other places where some land parcels are not fully utilised. Over the model time period, these areas are in demand from the agents and they become occupied. Waipa's subdivision and planning rules allow for much less intensification and subdivision.

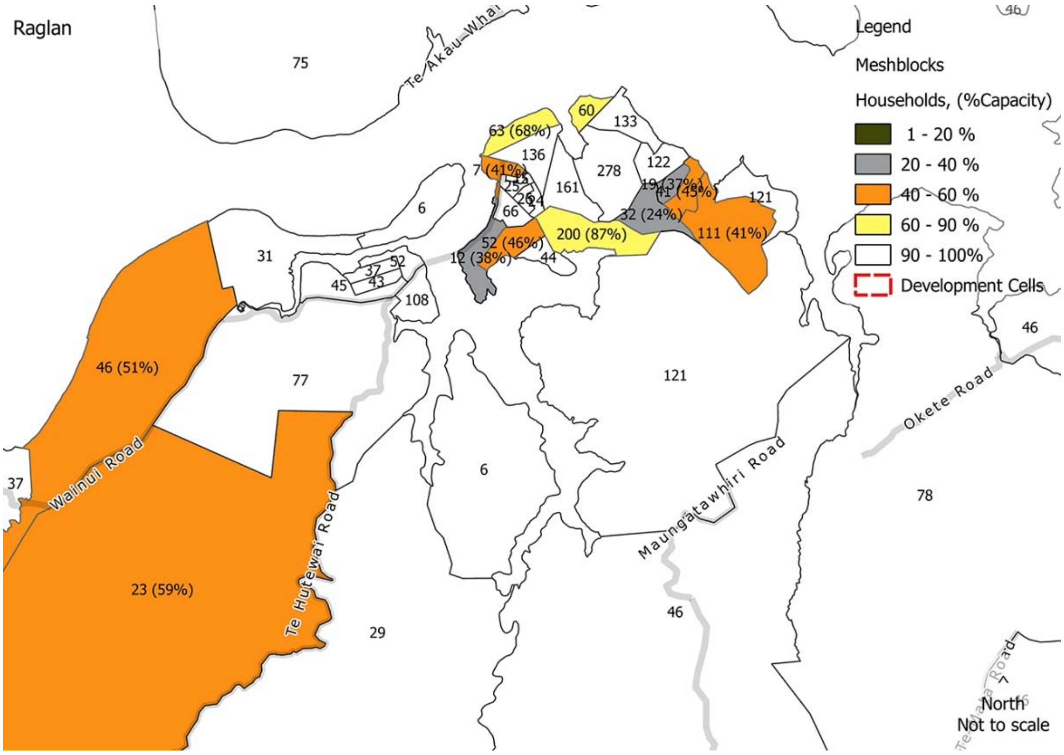
### 6.2.1 Waikato




Map 16. Te Kauwhata, meshblock percentage utilised by 2025 in the CDP scenario.

Note 1. When the meshblocks are less than 90% filled, then the number in brackets indicated the actual percentage filled.

Te Kauwhata (Map 17) is one of the areas with the most unexpected results. This is a pleasant country town, bordering a Ramsar site wetland. The town has a relatively high level of amenities with a library, fully serviced water networks and schools. The proximity of the town to the state highway also places the town on good access routes to employment centres to the north. The employment projection for Te Kauwhata is for 122 future employment opportunities by 2025. This may account for the limited movement of HAs into this village. Employment projections (McDonald, 2015) show an increase between 2021 and 2031 indicating this capacity is more likely to be utilised after 2025.

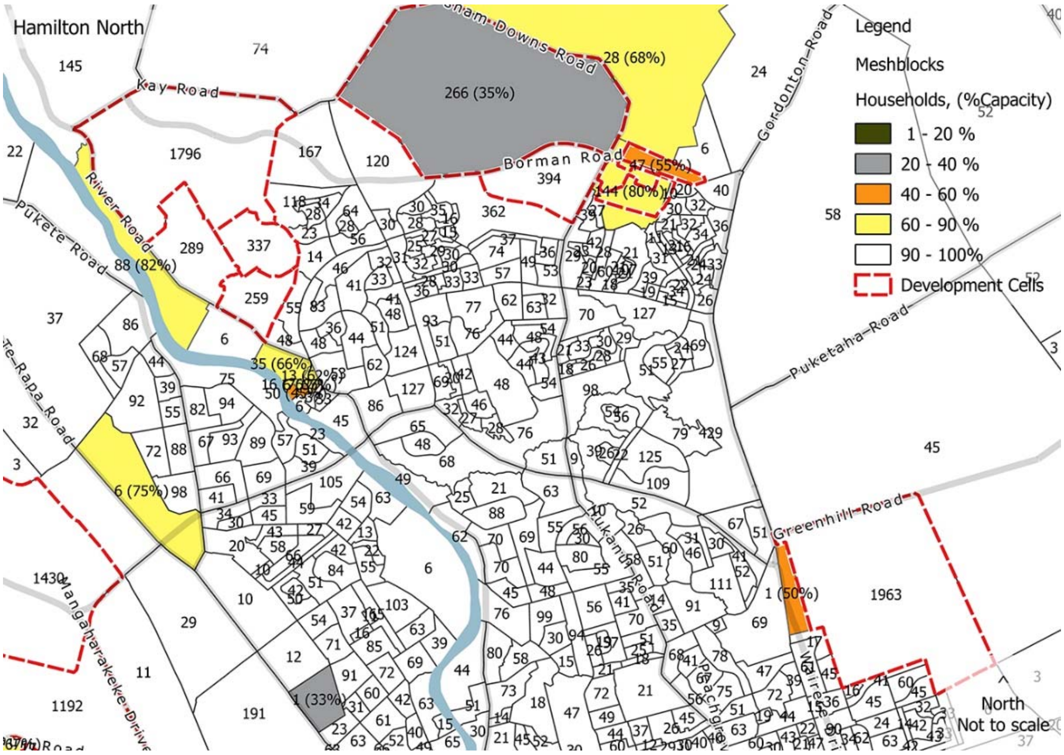



Map 17. Raglan, meshblock percentage utilised by 2025 in the CDP scenario.

Note 1. When the meshblocks are less than 90% filled, then the number in brackets indicated the actual percentage filled.

Raglan (Map 18) shows a number of meshblocks in the 40% to 60% fill. The Waikato council has provided estimates of the potential number of new land parcels. From observations using aerial photos, a number of the properties in these meshblocks are larger properties which the model interprets as having potential to subdivide. However, many of these are either on steep slopes or holiday homes in this coastal town and as such, they are actually less likely to be subdivided. It is possible that the capacity for subdivision is overestimated in Raglan. If the capacity is overestimated the model results could actually have a much higher percentage of the total capacity utilised. Raglan provides a challenge in projecting household and population changes, as do many of New Zealand’s coastal towns, because of the seasonal nature of the inhabitants and the distortion in the ratio between dwellings and the number of people who are usually resident.

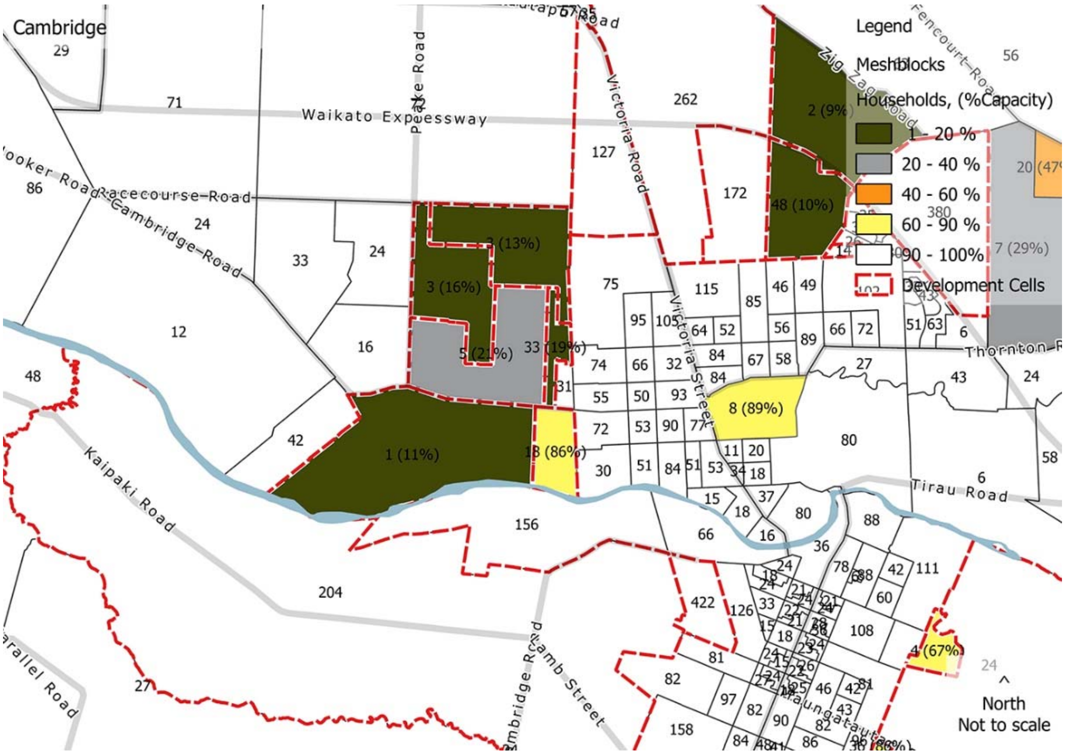
### 6.2.2 Hamilton




Map 18. North Hamilton, meshblock percentage utilised by 2025 in the CDP scenario.

Map 19 of North Hamilton shows the highest number of HAs settling in the development cells and filling both the development cells and the existing neighbourhoods to capacity. The two largest are the cells in the far north, Sylvester and across towards the east, in Ruakura with 1,795 and 1,963 HAs respectively. The grey area between Horsham Downs Road and Borman Road is planned to develop between 2014 and 2020, and it is expected that this area is likely to have filled by 2025. The areas in and adjacent to this grey coloured meshblock will experience a significant shift in neighbourhood amenity benefit as a new village centre and sports and education precincts are being developed between 2015 and 2016. The remaining development cells of Ruakura, Rotokauri and Peacocke fill to the planned capacity by 2025 (refer to Appendix 2).

### 6.2.3 Waipa



Map 19. Cambridge, meshblock percentage utilised by 2025 in the CDP scenario.

In a similar fashion to Te Kauwhata, the model outputs for Cambridge (Map 20) reflect an unexpected outcome. Cambridge is an area that showed a high error on the calibration output. Employment is projected to increase by slightly over 700 jobs up to 2025. Relative to the size of Cambridge, this amount of new employment should provide the potential for new agents to relocate to the development cells.

The development cells to the west are in the Cambridge West AU. These cells had 54 new dwellings between 2006 and 2013. The calibration only altered the NAB marginally, the starting NAB was 2500 and after calibration, the average NAB for land units in Cambridge West was 2558. With a possibly incorrect NAB a lower than anticipated number of HA located in these development cells. Further manual NAB adjustment is likely to resolve this inaccuracy. The development cells located south of the Waikato River have all filled to 100% of the planned capacity. Cambridge will experience a change in both NAB and travel cost as in 2016 an expressway bypass was opened.

# 7 Discussion

The key strength of this Waikato agent-based model is in its design that allows interactions that are independent of TA jurisdiction. Each of the council's strategic planning is centred on accommodating their own projected growth. There are, however, no tools to test the impacts of policy decisions and possible outward or inward impact of these policies or the spill overs of planning decisions in one TA on the surrounding TAs. Although in a simplified manner, this model has highlighted the relationship between how much land is available for development and the impact this has on rent, which influences the intra-regional household distribution.

The results show that even under broad assumptions there are intricate interdependencies between the city, the satellite towns and rural areas. In a complex world these inter-relationships are exponentially more complex and beyond the possibility of models to adequately replicate. The behaviour of the HAs clearly reflects the rational behaviour to minimise costs, as embedded in the agent-based model. The different planning approaches and different constraints imposed by the planning authorities have varied impacts which influence the neighbours to a greater or lesser degree.

The results show that constraints in land availability are detrimental to the city's growth. The complex relationship between council policy and land developers underlie the supply of residential land parcels. As Morgan's (2010) research shows, land developers play a central role in the provision of new properties to support residential developments. This is an area that is of particular interest to council policy planners as they decide on the areas of residential and business zoning and the funding of the support infrastructure. The councils collectively need to act conservatively and not over-invest in infrastructure. Similarly, the developers will seek the areas where they can return the highest margins. They may not develop land where a council intends for development to take place or through their profit motivations may slow down the supply of land to drive higher prices. These are some of the actions that constitute the parameters of the land availability constraint.

This model produces results at a meshblock level. An important question to ask is “Can the change in households for an individual meshblock be used in a planning context?” The model reflected the changes between 2006 and 2013 well with a relatively low error. This indicates that the projected changes at the meshblock do carry value. In order to really utilise and the output at the individual meshblock level, the model should be run and tested with a range of input values. Some areas are sensitive to variance in the input values and others are less sensitive. Two examples are represented in Cambridge West (Map 20) and the western side of Te Kauwhata Map 17. It has proved to be more difficult to establish the optimal calibration value for the NAB in areas that experienced high growth rates or those that start with sparse populations. Testing scenarios will help to establish if these areas are sensitive to the calibration NAB or other development cells that are directly ‘competing’ with these areas.

Although not prevalent in this catchment, declining population was captured well and the calibration results show that this model provides a good job of accounting for past population decline in this study area. Dealing with declining growth in other household projection methods can be challenging. The challenge is greater for top down approaches that disaggregate the number of projected households from larger to smaller areas, especially where households are disaggregated proportionally and areas with declining household counts can’t be adequately accounted for. This model lends itself well to a polycentric environment with agents being able to move between centres and capitalises on the highly mobile nature of the household agents in this agent-based model. The model also currently allows HAs to lose a job and become unemployed. This feature is not specifically required in this study area as both employment and population projections increase. The model’s calibration results have been tested for meshblocks that experience decline and further studies in areas with declines in employment will be useful for further development and improvement of the method.

The spatial distribution of the error was of interest. That is, would an agent-based model provide a consistent outcome and would the outliers in error be evenly distributed? Mapping the error showed a relatively even distribution of error (see Map 4). The meshblocks with higher error could distinctly be identified in areas of high growth over the calibration period. Some of the low-density areas situated in the North also showed error in the medium to high range. These types of error have been experienced by a range of authors such as Benenson (1998), Rayer & Smith (2010), and Statistics New Zealand (2008). Fontaine & Rounsevell (2009) produced similar maps highlighting the spatial distribution of error, although they only displayed a map showing model results alongside a map showing known results.

As recognised by Axtell (2000), people are good at pattern recognition and analogical reasoning, and frequently agent-based models can represent complex relationships in visual formats that are much easier to interpret. The visual outputs of this agent-based model present well and the results are not overly complex for a general audience to interpret. The pseudo representation of individual households provides a clearer picture of how property development and household locations might change in the real world. Presenting results in such a manner will hopefully be clearer for planners, senior management and politicians. Choropleth mapping is frequently used to show household or population distribution and can be misleading in the visual presentation of outcomes. First, it can mislead in terms of distortion due to the relative size of the feature being represented, where a less significant outcome can dominate the map due to its geographical size. For example, in this study area, some of the meshblocks with the highest household counts are too small to distinguish on small scale printed maps, while some with minimal populations are geographically cover a large area. Second, features such as houses will not have an even distribution over the mapped areas, so presenting these areas as uniform colour misdirects the attention of the reader. This model addresses these issues by representing the meshblocks as equal area cells and presenting a pseudo location of households within the meshblock.

The calibration of the model produced better than expected results with a Root Mean Squared Error of 16.6% at the meshblock level. By international standards, meshblocks are very small and represent the census data at a high resolution. Achieving this level of error at this spatial resolution is positive. The error reduces to 10.56% at the area unit level, which reflects reducing error with an increase in the geographic size of an area and is consistent with findings in other studies (Rayer & Smith, 2010; S. K. Smith & Cody, 2013; Statistics New Zealand, 2008). The areas with the highest measured error correspond with the fastest growing areas. A number of other investigations have also found the faster growing areas are subject to poorer calibration results (Rayer & Smith, 2010; Statistics New Zealand, 2008).

A further method of validation of the model was to compare the outputs against other regionally and nationally recognised household projections.

Overall the model outputs do not have any significant or unexplainable difference from either Statistics New Zealand or NIDEA projections. The model results show a more irregular projection than the smooth lines of either the Statistics New Zealand or the NIDEA household projections. The more irregular pattern of the agent-based model reflects emergent behaviour and the dynamics between the TAs. Statistics New Zealand projections have some discrepancy and the starting household counts are not consistent with the census household and dwelling counts, which unfortunately diminishes some of the comparative value of this dataset.

The differences between the model outputs under the two scenarios were greater than expected. The disproportional impact of the scenarios is also unexpected. Under both scenarios Waikato district is projected to accommodate significantly more households than the NIDEA and Statistics New Zealand models show. Being smaller than Hamilton City but covering a much larger geographic area, an additional four to five thousand additional houses represents about 50% more growth than planned. Potentially, having this much additional growth will put a lot of pressure on the existing resources and infrastructure, particularly if it is spread over a wide area. Waikato District will have little influence over this growth as the policies of Hamilton are a major contributing factor. Waikato District faces the same pressure from Auckland City located on its northern boundary, with Auckland being an order of magnitude larger than Hamilton.

Waipa HA change is relatively insensitive to the supply of new land parcels in Hamilton. An interesting exercise would be finding the thresholds where HAs begin to move towards or away from Waipa towns. If development cells in Hamilton become more constrained there may be a point at which the movement shifts to Waipa. One example might be a high influx of new HAs into the region. If more land was to become available in Hamilton, would this impact Waipa?

As identified by authors such as Castle & Crooks (2006), Couclelis (2005), Fontaine and Rounsevell (2009), and Foss and Couclelis (2009), ABMs are sensitive to small variations in input. This model also displays sensitivity to variance in the input values. Some monitoring of sensitivity was undertaken. However, further work in this area would be beneficial, particularly as varying input values can provide planners with valuable information about the relationships between different meshblocks.

The results of the Waikato Agent-based Model have been specifically compared against the 2015 household projections of NIDEA and Statistics New Zealand. The most notable difference in the results is the Waikato Agent-based Model graphed growth rate is not a smooth line and each TA has ‘bumps’ and ‘dips’. This is due to the movement of households between different areas and the changes induced in the establishment of new subdivision developments. The models differ in that the Waikato Agent-based model has an increasing growth rate for the Waikato district. For NIDEA and Statistics New Zealand projections the disaggregation of each set of TA data are treated independently, thus cross boundary effects are not taken into account in the disaggregation procedure. The Waikato Agent-based model procedures are independent of the TA boundaries and reflect internal migration effects based on zoning capacity and economic drivers. This study demonstrates a changing growth rate whereas the NIDEA and Statistics New Zealand have a more constant rate and produce smooth growth projections.

The Waikato Agent-based Model sums household numbers to meshblock and then again up to the Area Unit. The NIDEA and Statistics New Zealand models are based on a disaggregation of TA projections down to the area unit. Overall the two approaches converge and produce similar results at the area unit.

Improving models such as this one pivot on achieving the right level of complexity. The complexity is a composite of the number of agents that operate in the environment and how complex their decision making needs to be. The more complex the model is the more challenging the calibration process becomes (Castle & Crooks, 2006). Further to this, increasing the complexity, in either the number of variables or the decision-making process, could result in quite different model outcomes, as outlined by Couclelis (2002) and Foss and Couclelis (2009). The last consequence of complexity depends on the purpose of the model. If the model is to be used by non-technical users then it needs to be easily explainable (Grimm et al., 2010) and justifiable.

There are a few areas where this model can be improved. The calibration process can be improved and the recorded RMSE could potentially be reduced further, especially if a few of the outliers can be optimised. Improving the calibrating was also suggested by Fontaine and Rounsevell (2009) for their model. Calibrating the model over a longer period of time might result in a reduction in the calibration errors. However this model specifically uses the neighbourhood amenity benefit, and this in itself can be quite dynamic over even a short period of time. A good example of this is between 2006 and 2013 significant developments in the retail and schools has taken place in North Hamilton and consequently the focus of the city has changed due to this amenity change. A longer calibration period will expose more of these types of change.

As outlined in the literature review, and highlighted by a number of authors, for example, Fontaine and Rounsevell (2009), life stage is a significant driver for relocating to a new house. An area of complexity for a model designer is to project household types into the future. Incorporating what type and size of houses will be constructed in the future would be challenging to provide a model and potentially highly speculative.

Having a demographic module in the model, such as Fontaine and Rounsevell (2009) Jjumba and Dragicevic (2012), and Gaube and Remesch (2013), has some distinct advantages. These advantages are centred around simulating population growth, with data on ages as well as household composition. Household data from the model could be used to differentiate household travel patterns. For example households with school age children can incur costs to travel to a school. The neighbourhood amenity benefit could be calculated differently for different life stages of the household, in a manner similar to Fontaine and Rounsevell (2009) and Gaube and Remesch (2013).

This model runs in synchronous time and runs until spatial equilibrium is reached when the agents can no longer reduce costs. A potential step to bring the model closer to reality would be to run in asynchronous time. No point of equilibrium would be achieved and time could be managed on multiple time clocks such as used by Jjumba and Dragićević (2012). In this case, asynchronous time would represent the real world better as generally households choose when they would move. At this point in time there a significantly fewer location options than in synchronous time, where all options are available at the start of the time step.

An area of further consideration is based on other parameters of the model's data limitations. The data suitability and applicable data scales must be considered for appropriate data quality and assessment of 'fit for purpose' for use in the intended model (Castle & Crooks, 2006). This model utilises datasets that are constructed at a scale (AU) that is greater than the scale at which the results are presented (meshblock). The data have been reviewed and data handling has been outlined in 4.6. The datasets that have been downscaled have been identified. Input data, specifically Statistics New Zealand census data, are subject to inherent collection error, rounding error and in some cases data are censored and not presented. Such variations in the input data have the potential to generate variable outputs as cautioned by (Foss & Couclelis, 2009; Couclelis, 2002).

Employment location provides a significant component of the travel cost calculation for each HA. The model, therefore, has a high level of dependency in the temporal, spatial and quantitative accuracy of employment projections produced by Market Economics (McDonald, 2015). In this model, the employment projections proceed at a slightly higher growth rate than the household projections. As such there are no constraints for HAs in terms of future employment opportunities. In reality, there are likely to be constraints within industry sectors, but these have not been considered in this model. When new HAs are introduced, they are randomly assigned an employment location. This introduces a small amount of random variation which results in different employment opportunities being taken up on each model run. Some limited testing on this variation has taken place. At the TA level there is not much variation; however, the spatial variation at the meshblock level has not been fully investigated and requires further analysis.

Through the duration of running this Waikato Agent-based Model, many thousands of households are relocated and this would be many times more than the relocations that take place in reality. Foss and Couclelis (2009) highlight that agent-based models may have a tendency to over-emphasise emergent behaviour. In the model used here, the agent is free to relocate in the subsequent cycle. There are instances where a land unit steadily gains household agents and then suddenly a high number of HAs leave. Creating a lag in this behaviour may buffer spontaneous relocations and address potential over-emphasis of household relocations. Jjumba & Dragićević (2012) have a 12 month lag preventing and household from moving again. An example is found in Pokeno. In the CDP, the land unit containing meshblock number 842200 is projected to have 1,168 HAs in 2019, but only 9 in the following year. Overall the Pokeno AU declines by 881 HAs in this time step. This indicates these are not local shifts taking place. Another example is Huntington, which declines by 213 HAs in one time step.

A key reaction that the HA induces when they move is the change in rent. This was set to have a diminishing increment and respective decrement in order to prevent rent values exceeding reasonable values. The amount that rents change was not based on any observed price changes and having this as an observed value could provide valuable analysis of the future value of land and return on infrastructure investment. In the study areas, property owners pay proportional council tax based on land and capital value. It is also recognised that property value inflation is not uniform and higher value properties are more inclined to have higher inflation rates and vice versa. This model does not address any of these complex housing market issues. For that matter, this issue was not addressed in any of the research reviewed as these considerations would require modelling a housing market which is not straightforward by any means.

The total capacity of land units is defined as the maximum number of HAs that can reside in the land unit. This capacity can increase, either through subdivision of existing land parcels (brownfield development), or new land parcels introduced in development cells (greenfield development). It is very difficult to make assumptions about the rate at which brownfield land will develop. In this model the quantities of brownfield land are estimated by the councils and in the first ten years this land is completely utilised by the HAs. A significant component of Morgan's (2010) model is devoted to the land developers' actions. The timing and release of brownfield infill has an impact on the time based results of the model. The uncertainties associated with brownfield development should be considered in areas where the brownfield development constitutes a meaningful portion of the land unit capacity.

Household projections are an exogenous input to this model. The household projections determine how many HAs are introduced in each time step. The model's temporal outputs, therefore, have a high dependency on the temporal certainty of the input employment and household projections. Interpreting the outcomes of the model should, therefore, have careful consideration of the temporal constraints of both of these input projection series. The further out in time the results are produced the higher the uncertainty of the specific time related event becomes. As discussed earlier, there is well-documented evidence that accuracy of population and household projections decreases as geographic size decreases (Rayer & Smith, 2010; Rees et al., 2004; S. K. Smith et al., 2001; Statistics New Zealand, 2008) . This concurs with this model that has higher calibration error when results are measured at smaller area i.e. meshblocks than at the larger area unit. Following from this, the results should be interpreted with a greater consideration of higher error relating to the timing of events at smaller geographic and or population size.

One of the issues relating to the relationship between model producers and planners is endogeneity (Cameron & Cochrane, 2015; Statistics New Zealand, 2008). Cameron and Cochrane (2015) identify the self-fulfilling nature of planning in that planners seek population projection to identify where to provide infrastructure, development most likely takes place in the areas where planners provide infrastructure. This Waikato Agent-based Model allows testing of how residents are likely to respond to the development plans and location of infrastructure. Waipa is providing infrastructure to two primary areas, Hamilton to five and Waikato more than six. As constraints exist in delivering this infrastructure, it is a serious question that the council are collectively trying to address, "what is the optimum infrastructure investment plan?" This is further compounded by relatively low population densities spread over a high number of towns and settlements resulting in a per capita high cost for services such as piped water, public transport, libraries and other amenities.

The location and timing of new development cells and associated infrastructure is a primary input to the model. From the model designers perspective, they have to ensure they are not reinforcing such behaviour through feedback loops, as identified by Foss and Couclelis (2009) as potential negative aspects of agent-based models. Similarly, the model designer has to ensure the model results are not individually adopted as the answer to infrastructure development. Utilisation of models similar to this one can feasibly provide planners with the context to support their development planning and as such reinforce the intrinsic relationship between model outcomes and the planner's response.

The model itself and its usage can, however, be used as a tool to mitigate these types of endogenous relationships. As demonstrated, this model can be used to test more than one scenario. Running a range of scenarios will enable the emergent trends to be identified and explored, and it is foreseeable that analysing the model's outputs will show areas that have a consistent outcome and other areas with higher variance. Identifying which variable causes the variance can, in turn, helps planners with their policy objectives. Varying the neighbourhood amenity benefit, rent or travel time would show that above or below certain thresholds HAs will start to populate an area. As identified in the results, both Te Kauwhata and Cambridge West don't experience as much growth as anticipated, thus altering the costs and NAB could indicate the criteria required to influence more growth. Planners could identify possible policy measures to influence growth patterns. In conjunction, it is valuable to identify where the opposite reaction is taking place. Planners would need to ensure that the opposite effect is widespread and not a direct depletion of HAs from a single or a small number of sources. For example in Te Kauwhata and Cambridge, measures might be taken to make these destinations more cost beneficial, but in doing so attract the HA neighbouring towns such as Huntly or Te Awamutu.

# 8 Conclusion

The cohort component method is one of the most widely used methods used for projecting population change, (Bascand, 2012, p. 9; Bell et al., 2010, p. 17) and is generally well accepted as a robust method. The challenge for demographers is to not only project how much a population is going to change but also where changes most likely to occur. This becomes increasingly difficult as population and or geographic size decrease (Jenner, 2002; Rayer & Smith, 2010; Statistics New Zealand, 2008). Urban planners, on the other hand, have a range of tools used to simulate land use change and city growth. This research presents an agent-based model that uses the household projections obtained from conventional methods (NIDEA 2015) and applies an urban growth model to investigate changes in the geographic distribution of households.

A wide range of urban growth models may be coupled with demographic methods in a wide range of ways (Triantakonstantis & Mountrakis, 2012). The literature review established CA and agent-based models (ABM) as the most likely candidate algorithms to use in the Waikato study. Both of these methods have previously been utilised in describing actions of households and how their location is affected by various influences. Households and/or individual citizens have clear links with demographic processes which can be utilised in simulating change in demographic and locations over time.

The study area can be described as poly-centric with a central city and a number of satellite towns, villages and small settlements. This agent-based model's results are a detail map series of the household distributions over a 12 year period, showing the effects of two different scenarios. This model adequately addresses the question as to "What will the likely population distribution be for Hamilton City and the Waikato and Waipa districts in 2025?" Over this time period, around 31,000 houses need to be constructed and this model produces a detailed picture of what household distribution might look like. Furthermore, the model demonstrates the complexities of the relationship between the three councils and their development aspirations.

The results have been presented to staff at each of the councils. The assumptions, method and selection of primary inputs have been acknowledged and recognised as appropriate in this context. The three councils are working together closely and there is a far greater alignment of strategic planning than there has been in the past. Further development of this tool would enable more in-depth scenario testing. In conjunction with other tools such as Waikato Integrated Scenario Explorer (WISE) (Waikato Regional Council, 2014), Statistics New Zealand and NIDEA household and population projections, this tool could help in developing a better understanding of how sensitive different areas are to variations in the external environment. Along with having projections of households, population and land use, future planning is equally reliant on gauging certainty. Some areas will develop under most circumstances whilst under certain constraints, some areas will be more unpredictable and pose a higher risk for councils.

# 9 References


Al-sharif, A. A. A., & Pradhan, B. (2015). A novel approach for predicting the spatial patterns of urban expansion by combining the chi-squared automatic integration detection decision tree, Markov chain and cellular automata models in GIS. *Geocarto International*, *30*(8), 858–881. http://doi.org/10.1080/10106049.2014.997308

Axtell, R. (2000). Why Agents? On the varied motivations for agent computing in the social sciences [Publisher]. Retrieved December 10, 2013, from http://www.brookings.edu/research/reports/2000/11/technology-axtell

Baker, J., Ruan, X., Alcantara, A., Jones, T., Watkins, K., McDaniel, M., … Stallcup, W. (2008). Density-dependence in urban housing unit growth: An evaluation of the Pearl-Reed model for predicting housing unit stock at the census tract level. *Journal of Economic & Social Measurement*, *33*(2/3), 155–163.

Bascand, G. (2012, August). Demographic projections from Statistics New Zealand: Aims, methods, and results. Statistics New Zealand.

Batty, M. (2005). Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. *Environment and Planning A*, *37*(8), 1373–1394.

Batty, M., & Torrens, P. M. (2005). Modelling and prediction in a complex world. *Complexity and the Limits of Knowledge*, *37*(7), 745–766. http://doi.org/10.1016/j.futures.2004.11.003

Bell, M., Blick, G., Parkyn, O., Rodway, P., & Vowles, P. (2010, January). Challenges and choices: modelling New Zealand's long-term fiscal position. The Treasury. Retrieved from http://www.treasury.govt.nz/publications/research-policy/wp/2010/10-01/09.htm

Benenson, I. (1998). Multi-agent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems*, *22*(1), 25–42. http://doi.org/10.1016/S0198-9715(98)00017-9

Bouveyron, C., & Jacques, J. (2010). Adaptive linear models for regression: Improving prediction when population has changed. *Pattern Recognition Letters*, *31*(14), 2237–2247. http://doi.org/10.1016/j.patrec.2010.03.021

Cai, Q. (2007). New techniques in small area population estimates by demographic characteristics. *Population Research and Policy Review*, *26*(2), 203–218. http://doi.org/10.1007/s11113-007-9028-7

Cameron, M., & Cochrane, W. (2015a). Using land-use modelling to statistically downscale population projections to small areas. Department of Economics, Waikato University.

Cameron, M., & Cochrane, W. (2015b, July). Population, family and household, and labour force projections for the Waikato region, 2013 - 2063 (2015 update). Waikato Regional Council. Retrieved from http://www.waikatoregion.govt.nz/TR201528/

Cameron, M., Cochrane, W., & Poot, J. (2007). *End-user informed demographic projections for Hamilton up to 2041* (Working Paper). Retrieved from http://researchcommons.waikato.ac.nz/handle/10289/805

Cameron, M., Cochrane, W., & Poot, J. (2008). *Population projections until 2061 for FutureProof - the Hamilton sub-regional growth strategy.* Population Studies Centre, The University of Waikato. Retrieved from http://www.waikato.ac.nz/php/research.php?skin=2014&year=2008&mode=show&author=62015

Cameron, M. P., & Cochrane, W. (2015). *Population, family and household, and labour force projections for the Waikato region, 2013-2063 (2015 updated)* (Vol. 2015/28;2015/28.;). Hamilton [New Zealand]: Waikato Regional Council. Retrieved from http://waikato.lconz.ac.nz/vwebv/holdingsInfo?bibId=1664218

Castle, C., & Crooks. (2006). Principles and concepts of agent-based modelling for developing geospatial simulations. Retrieved February

11, 2013, from http://www.bartlett.ucl.ac.uk/casa/latest/publications/working-papers

Couclelis, H. (2002). Chapter 2: Modeling frameworks, paradigms, and approaches In K. C. Clarke, B. E. Parks, & M. P. Crane (Eds.). In *Geographic information systems and environmental modelling.* London: Prentice Hall. Retrieved from http://www.geog.ucsb.edu/~kclarke/ucime/Helens-Sem/seminar2001/Banff-book-chapter2.pdf

Couclelis, H. (2005). Where has the future gone? Rethinking the role of integrated land-use models in spatial planning. *Environment and Planning A*, *37*(8), 1353–1371.

Crooks, A., Castle, C., & Batty, M. (2008). Key challenges in agent-based modelling for geo-spatial simulation. *GeoComputation: Modeling with Spatial Agents*, *32*(6), 417–430. http://doi.org/10.1016/j.compenvurbsys.2008.09.004

Department of Internal Affairs. (2013). Local Government Act 2002 [Government Department]. Retrieved December 29, 2013, from http://www.localcouncils.govt.nz/lgip.nsf/wpg_url/Policy-Local-Government-Legislation-Local-Government-Act-2002?OpenDocument

de Smith, M., Goodchild, M. F., & Longley, P. A. (2013). *Geospatial analysis a comprehensive guide to principles, techniques and software tools* (4th Edition). Retrieved from http://www.spatialanalysisonline.com/HTML/index.html

Dubovyk, O., Sliuzas, R., & Flacke, J. (2011). Spatio-temporal modelling of informal settlement development in Sancaktepe district, Istanbul, Turkey. *Quality, Scale and Analysis Aspects of Urban City Models*, *66*(2), 235–246. http://doi.org/10.1016/j.isprsjprs.2010.10.002

English, B. (2015, August 20). Launch of the thirty year New Zealand infrastructure plan 2015 [Government Department]. Retrieved September 20, 2015, from https://www.beehive.govt.nz/speech/launch-thirty-year-new-zealand-infrastructure-plan-2015

Fontaine, C. M., & Rounsevell, M. D. A. (2009). An agent-based approach to model future residential pressure on a regional landscape. *Landscape Ecology*, *24*(9), 1237–1254. http://doi.org/10.1007/s10980-009-9378-0

Foss, S., & Couclelis, H. (2009). Throwing space back in: a tale of Indian fishermen, ICT and travel behavior. *ICT and the Shaping of Access, Mobility and Everyday Life*, *17*(2), 134–140. http://doi.org/10.1016/j.jtrangeo.2008.11.011

*Gama-platform*. (2014). IRD/UPMC international research unit UMMISCO. Retrieved from https://github.com/gama-platform/gama/wiki

Gaube, V., & Remesch, A. (2013). Impact of urban planning on household's residential decisions: An agent-based simulation model for Vienna. *Thematic Issue on Spatial Agent-Based Models for Socio-Ecological Systems*, *45*(0), 92–103. http://doi.org/10.1016/j.envsoft.2012.11.012

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., … DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. http://doi.org/10.1016/j.ecolmodel.2006.04.023

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

Hamilton City Council. (2013). Property Subdivisions. ESRI Shapefile, Hamilton City Council.

Hamilton City Council. (2014a). Hamilton City StructurePlans [Territorial Authority]. Retrieved January 7, 2014, from http://www.hamilton.govt.nz/our-council/council-publications/districtplans/proposeddistrictplan/appendix2/Pages/default.aspx

Hamilton City Council. (2014b). Peacocke Structure Plan [Territorial Authority]. Retrieved July 1, 2014, from http://www.hamilton.govt.nz/our-council/council-publications/districtplans/proposeddistrictplan/appendix2/Pages/Peacocke-Structure-Plan.aspx

Hamilton City Council. (2014c). Rotokauri Structure Plan [Territorial Authority]. Retrieved July 1, 2014, from http://www.hamilton.govt.nz/our-council/council-publications/districtplans/proposeddistrictplan/appendix2/Pages/Rotokauri-Structure-Plan.aspx

Hamilton City Council. (2014d). Rototuna Structure Plan [Territorial Authority]. Retrieved July 1, 2014, from http://www.hamilton.govt.nz/our-council/council-publications/districtplans/proposeddistrictplan/appendix2/Pages/Rototuna-Structure-Plan.aspx

Hamilton City Council. (2014e). Ruakura Structure Plan [Territorial Authority]. Retrieved July 1, 2014, from http://www.hamilton.govt.nz/our-council/council-publications/districtplans/proposeddistrictplan/appendix2/Pages/Ruakura%20Structure%20Plan.aspx

Inland Revenue. (2014). Business income tax, claiming business expenses, mileage rate [Government Department]. Retrieved May 10, 2014, from http://www.ird.govt.nz/business-income-tax/expenses/mileage-rates/emp-deductions-allowances-mileage.html

Jenner, A. (2002). Population estimates for non-standard geographical areas. The small area pilot project [Conferences 2002]. Retrieved September 23, 2013, from http://www.apa.org.au/upload/2002-5C_Jenner.pdf

Jjumba, A., & Dragićević, S. (2012). High resolution urban land-use change modeling: Agent iCity approach. *Applied Spatial Analysis and Policy*, *5*(4), 291–315. http://doi.org/10.1007/s12061-011-9071-y

Kim, J., Kang, Y., Hong, S., & Park, S. (2006). Extraction of spatial rules using a decision tree method: A case study in urban growth modeling. In B. Gabrys, R. Howlett, & L. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems* (Vol. 4251, pp. 203–211). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/11892960_25

Knighton, G. (2015). Waipa District Council, growth strategies and development plans [Personal Conversation].

Lei, Z., Pijanowski, B. C., Alexandridis, K. T., & Olson, J. (2005). Distributed modeling architecture of a Multi-Agent-based Behavioral Economic Landscape (MABEL) Model. *Simulation*, *81*(7), 503–515. http://doi.org/10.1177/0037549705058067

Lieberson, S. (1981). An asymmetrical approach to segregation. In C. Peach, V. Robinson, & S. Smith (Eds.), *Ethnic Segregation in Cities* (pp. 61–82). London: Croom Helm.

Lizhong Hua, Wang Man, Qiong Wang, & Xiaofeng Zhao. (2012). A new Decision Tree Classification approach for extracting urban land from Landsat TM in a coastal city, China. In *Information Science and Engineering (ISISE), 2012 International Symposium on* (pp. 282–286). http://doi.org/10.1109/ISISE.2012.71

Macal, C. M., & North, M. J. (2005). Tutorial on agent-based modeling and simulation. In *Proceedings of the 37th conference on Winter simulation* (pp. 2–15). Orlando, Florida: Winter Simulation Conference. Retrieved from http://dl.acm.org.ezproxy.waikato.ac.nz/citation.cfm?id=1162708.1162712

Maithani, S. (2009). A neural network based urban growth model of an Indian city. *Journal of the Indian Society of Remote Sensing*, *37*(3), 363–376. http://doi.org/10.1007/s12524-009-0041-7

McDonald, G. (2015). Employment projections for the Waikato region, 2013-2063 (2015 updated) employment projection. Waikato Regional

Council. Retrieved from http://waikato.lconz.ac.nz/vwebv/holdingsInfo?bibId=1664218

Microsoft Corporation. (2012). Servers and Enterprise Development, SQL Server 2012 (Version 2012 R2). Microsoft. Retrieved from https://msdn.microsoft.com/en-us/library/dd206988%28v=sql.110%29.aspx

Moreno, N., Wang, F., & Marceau, D. J. (2009). Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model. *Computers, Environment and Urban Systems*, *33*(1), 44–54. http://doi.org/10.1016/j.compenvurbsys.2008.09.008

Morgan, F. J. (2010). *Residential property developers in urban agent-based models: competition behaviour and the resulting spatial landscape*. Auckland University. Retrieved from http://hdl.handle.net/2292/6375

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., … Schwarz, N. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. *Environmental Modelling & Software*, *48*, 37–48. http://doi.org/10.1016/j.envsoft.2013.06.003

Ramduny, V. (2014). Waikato District Council, growth strategies and development plans [Personal Conversation].

Rayer, S., & Smith, S. K. (2010). Factors affecting the accuracy of subcounty population forecasts. *Journal of Planning Education and Research*, *30*(2), 147–161. http://doi.org/10.1177/0739456X10380056

Rees, P., Norman, P., & Brown, D. (2004). A framework for progressively improving small area population estimates. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *167*(1), 5–36. http://doi.org/10.2307/3559797

Roskruge, M., Cameron, M., & Cochrane, W. (2010). Waikato District sub-district population projections, 2011-2031.

Santé, I., García, A. M., Miranda, D., & Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*, *96*(2), 108–122. http://doi.org/10.1016/j.landurbplan.2010.03.001

Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, *1*(2), 143–186. http://doi.org/10.1080/0022250X.1971.9989794

Shared Valuation Database Service (SVDS). (2014). Property. Local Authority Shared Services Ltd.

Smith, G., & Bevan, J. (2011). The Waikato Regional Transportation Model. Gabites Porter Consultants Ltd & Latitude Planning Ltd. Retrieved from http://www.hardingconsultants.co.nz/ipenz2011/downloads/Smith_Grant.pdf

Smith, S. K., & Cody, S. (2013). Making the housing unit method work: An Evaluation of 2010 Population Estimates in Florida. *Population Research and Policy Review*, *32*(2), 221–242. http://doi.org/10.1007/s11113-012-9265-2

Smith, S. K., Tayman, J., & Swanson, D. A. (2001). *State and Local Population Projections: Methodology and Analysis*. New York: Kluwer Academic Publishers.

Spurr, M. (2014). Hamilton City Council, growth strategies and development plans [Personal Conversation].

Statistics New Zealand. (2013e). Total personal income by age group, 2013 Census (RC, TA, AU). Retrieved from http://nzdotstat.stats.govt.nz/OECDStat_Metadata/ShowMetadata.ashx?Dataset=TABLECODE8138&Coords=[INCOME]&ShowOnWeb=true&Lang=en

Statistics New Zealand. (2008). How accurate are population projections? An evaluation of Statistics New Zealand population projections, 1991-2006. *Statistics New Zealand*. Retrieved from ISBN 978-0-478-31533-2

Statistics New Zealand. (2013a). Area unit definition [Government Department]. Retrieved November 4, 2013, from http://www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/area-unit/definition.aspx

Statistics New Zealand. (2013b). Dwelling Count [Government Department]. Retrieved November 2, 2015, from http://www.stats.govt.nz/Census/2013-census/data-tables/meshblock-dataset.aspx#excelfiles 11/02/2015

Statistics New Zealand. (2013c). Household income by region [Government Department]. Retrieved October 5, 2015, from http://nzdotstat.stats.govt.nz/wbos/Index.aspx?DataSetCode=TABLECODE7464

Statistics New Zealand. (2013d). Meshblock Dataset Census 2013 [Government Department]. Retrieved from http://www.stats.govt.nz/Census/2013-census/data-tables/meshblock-dataset.aspx

Statistics New Zealand. (2013e). Meshblock definition [Government Department]. Retrieved November 4, 2015, from http://www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/meshblock/definition.aspx

Statistics New Zealand. (2013f). Territorial authority definition [Government Department]. Retrieved January 6, 2013, from http://www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/territorial-authority/definition.aspx

Tayyebi, A., Pijanowski, B. C., & Tayyebi, A. H. (2011). An urban growth boundary model using neural networks, GIS and radial parameterization: An application to Tehran, Iran. *Landscape and Urban Planning*, *100*(1–2), 35–44. http://doi.org/10.1016/j.landurbplan.2010.10.007

Thomas, I., & Frankhauser, P. (2013). Fractal Dimensions of the built-up footprint: buildings versus roads. Fractal evidence from Antwerp (Belgium). *Environment and Planning B: Planning and Design*, *40*(2), 310–329. http://doi.org/10.1068/b38218

Torrens, P. M., & Benenson, I. (2005). Geographic Automata Systems. *International Journal of Geographical Information Science*, *19*(4), 385–412. http://doi.org/10.1080/13658810512331325139

Triantakonstantis, D., & Mountrakis, G. (2012). Urban growth prediction: A review of computational models and human perceptions. *Journal of Geographic Information System*, *04*(06), 555–587. http://doi.org/10.4236/jgis.2012.46060

US Census Bureau. (2014). About population projections. What is the Cohort Component Method [Government Department]. Retrieved January 1, 2016, from http://www.census.gov/population/projections/about/

Waikato District Council. (Territorial Authority). Te Kauwhata Structure Plan [Government Department]. Retrieved January 6, 2014, from

Waikato District Council. (2014a). Tuakau and Ngaruawahia Structure plans [Government Department]. Retrieved from https://www.waikatodistrict.govt.nz/Documents/Plans/Structure-plans.aspx

Waikato District Council. (2014b). Waikato District Structure Plans [Territorial Authority]. Retrieved from http://www.waikatodistrict.govt.nz/Documents/Plans/Structure-plans.aspx

Waikato Regional Council. (2013, July 31). Property. Waikato Regional Council.

Waikato Regional Council. (2014). Creating Futures - WISE [Regional Authority]. Retrieved from http://www.creatingfutures.org.nz/

Waipa District Council. (2014). Waipa District Plan and Growth Strategy [Territorial Authority]. Retrieved May 20, 2014, from http://www.waipadc.govt.nz/our-council/Waipa2050/Pages/default.aspx

Wang, L., & Cardenas, J.-S. (2011). Small area population estimation with high-resolution remote sensing and LIDAR. In *Urban Remote Sensing* (pp. 183–193). John Wiley & Sons, Ltd. Retrieved from http://dx.doi.org/10.1002/9780470979563.ch13

Wu, F., & Martin, D. (2002). Urban expansion simulation of Southeast England using population surface modelling and cellular automata. *Environment and Planning A*, *34*(10), 1855–1876.

Yang, X. Y. (2011). Urban remote sensing: monitoring, synthesis and modeling in the urban environment. In *Urban Remote Sensing* (pp. i–xx). John Wiley & Sons, Ltd. Retrieved from http://dx.doi.org/10.1002/9780470979563.fmatter

Zhan, F. B., Tapia Silva, F. O., & Santillana, M. (2010). Estimating small-area population growth using geographic-knowledge-guided cellular automata. *International Journal of Remote Sensing*, *31*(21), 5689–5707. http://doi.org/10.1080/01431161.2010.496802

# 10 Appendix 1

## 10.1 SQL Sequence

The following is a description of the code sequence taking place. The code is run in a Microsoft SQL Server database with a series of stored procedure initiated in sequence.

*Procedure 1* - all base tables are generated from the source files, enabling the model to be repopulated.

*Procedure 2* - the modeller sets the required number of cycles for which to run the model.

*Procedure 3* - the base tables are replicated into a set of working tables which are updated as the simulation progresses. Base tables hold starting values and working tables hold values at each time step.

*Procedure 4* - the input cost of travel is set by the modeller (time cost and distance cost), the time and distance costs are calculated in the travel matrix.

*Procedure 5* - the total cost for all HAs is calculated and the cost field is updated for the HAs at their current location and current employment.

*Procedure 6* – new employment, new land units and new employment locations are added to the respective working tables.

*Procedure 7* – if any land units have a decrease in employment a corresponding number of randomly selected HAs from that land unit lose their employment. (In the next cycle unemployed Household Agents are randomly selected and may acquire new employment location).

*Procedure 8* – New HAs are added and randomly assigned employment. All new HAs are assumed to start with a residential cost of $99999. This ensures that the new HAs will be first to occupy a suitable vacant location with the lowest possible residential cost.

*Procedure 9* – a SQL cursor runs in which the HA with the highest residential cost is selected.

The vacant location with the lowest residential cost for that agent is identified and the HA's land unit identifier is updated. The originator land unit capacity decreases by 1 and the rent reduction calculated. The destination land unit capacity is increased by 1 and the rent inflates accordingly. All agents in the originator and destination land unit have their residential cost recalculated for the change in rent.

This cursor runs until there are no longer agents that can reduce their residential cost by relocating.

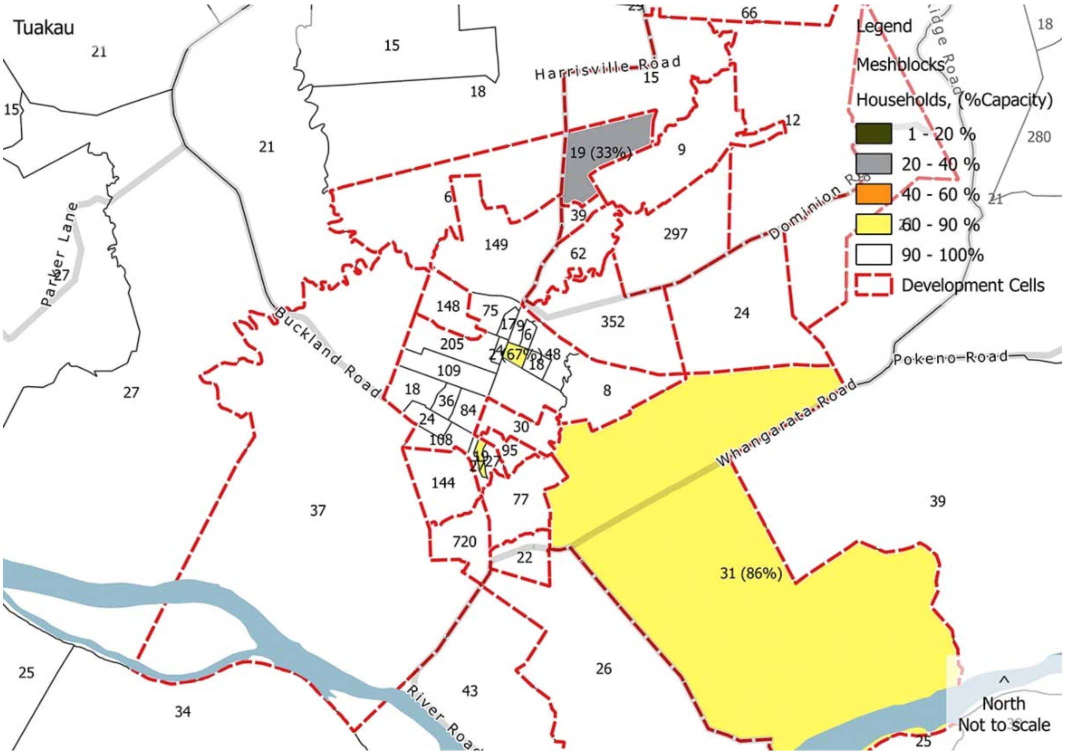
Procedure 10 – all agent movements are logged.

The simulation then enters the next cycle and procedures 5 to 11 are repeated for the number of cycles stipulated in procedure 2.

# 11 Appendix 2

Meshblock Capacity in development cells

## 11.1 Details of towns with 90% or more capacity

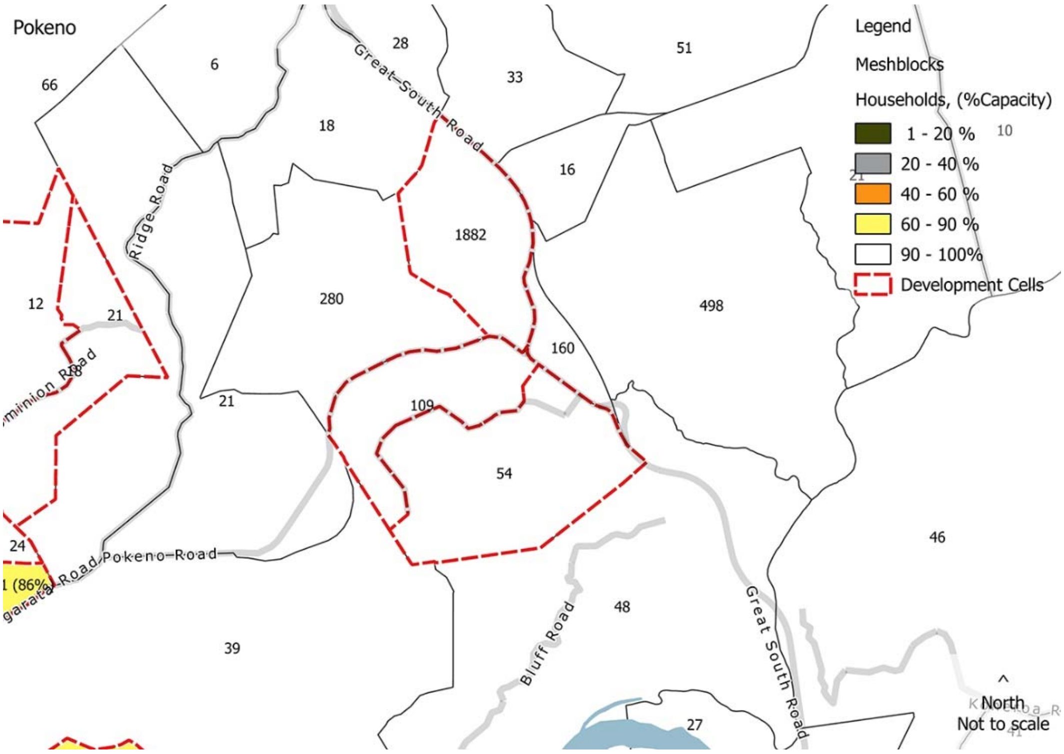



Map 21. Tuakau, meshblock percentage utilised, 2025

Note 1. Meshblocks coloured white are more than 90% fill and the number in the cell is the number of HAs.

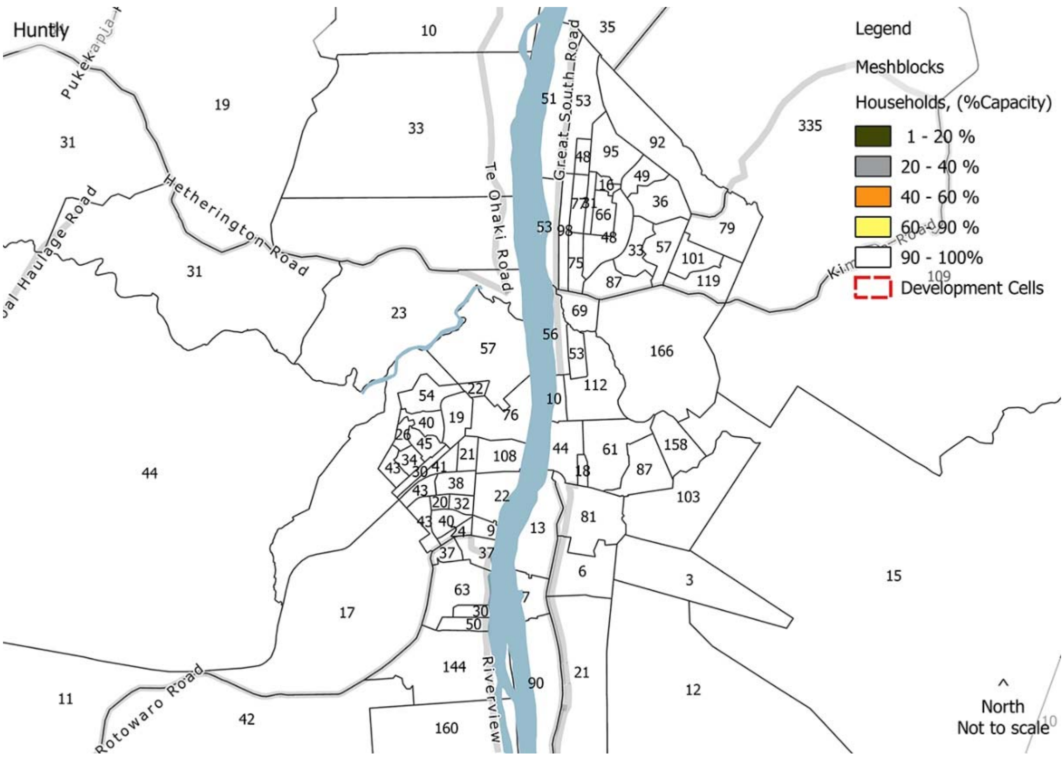
Note 2. When the meshblocks are less than 90% filled, then the number in brackets indicated the actual percentage filled.

On Map 22, the development cells, both to the north and the south of the town centre each fill up to around 900 households. The meshblock to the south east fills to 86%, this is not unexpected as this is a rural block with intensive agricultural land use. The light grey meshblock in the North is outside of the town boundary and is in the longer term growth plan. Two small meshblocks in the centre of the town don’t fill, one of these is filled with commercial buildings.

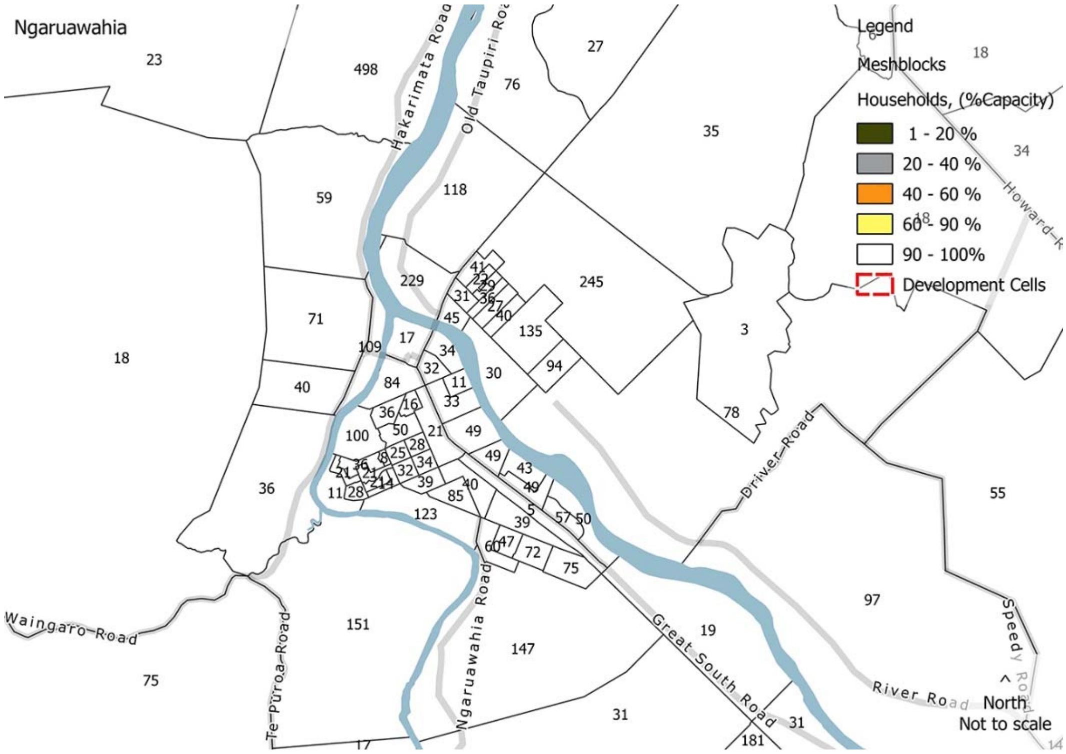



Map 20. Pokeno, meshblock percentage utilised, 2025

Pokeno (Map 22), is a primary growth area in this catchment, and all areas fill to planned capacity within the 2025 timeframe.



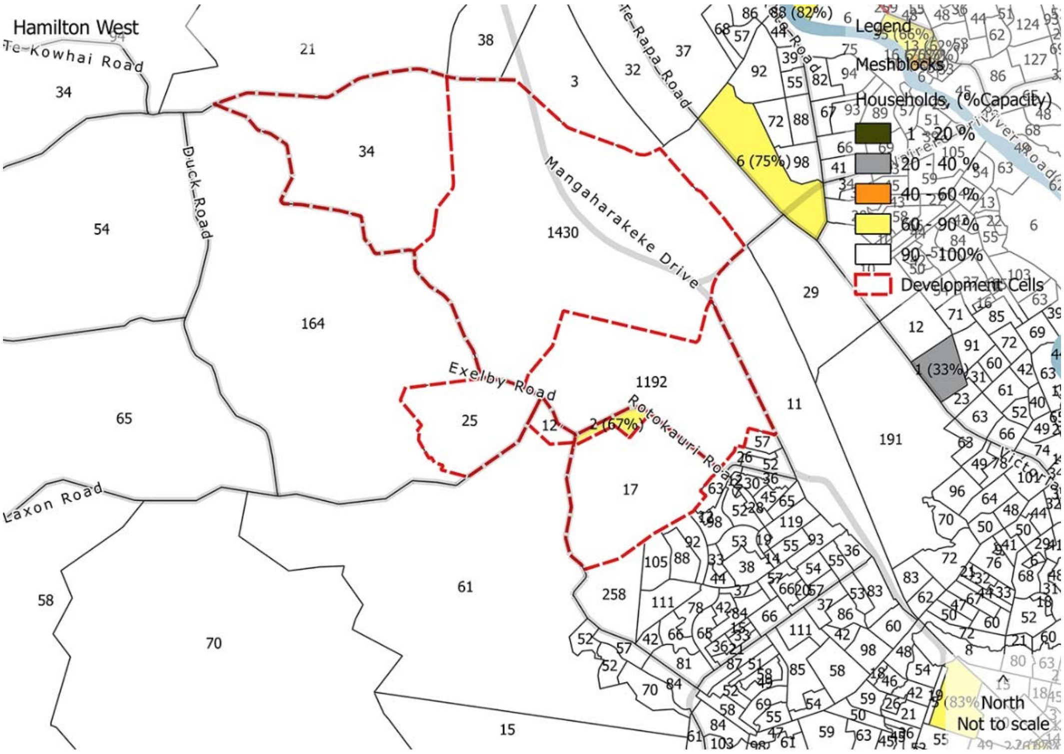
Map 21. Huntly, meshblock percentage utilised, 2025




Map 22. Ngaruawahia, meshblock percentage utilised, 2025

Map 23 and Map 24, for Huntly and Ngaruawahia, show the same trend of all available residential locations being utilised. These two towns do not have development cells and their capacity is based on existing properties and their potential to subdivide. This is particularly challenging in modelling as it is not possible to know when the land owner decides to subdivide. The outputs from this model show that all the landowners have chosen to subdivide within the 2025 timeframe.

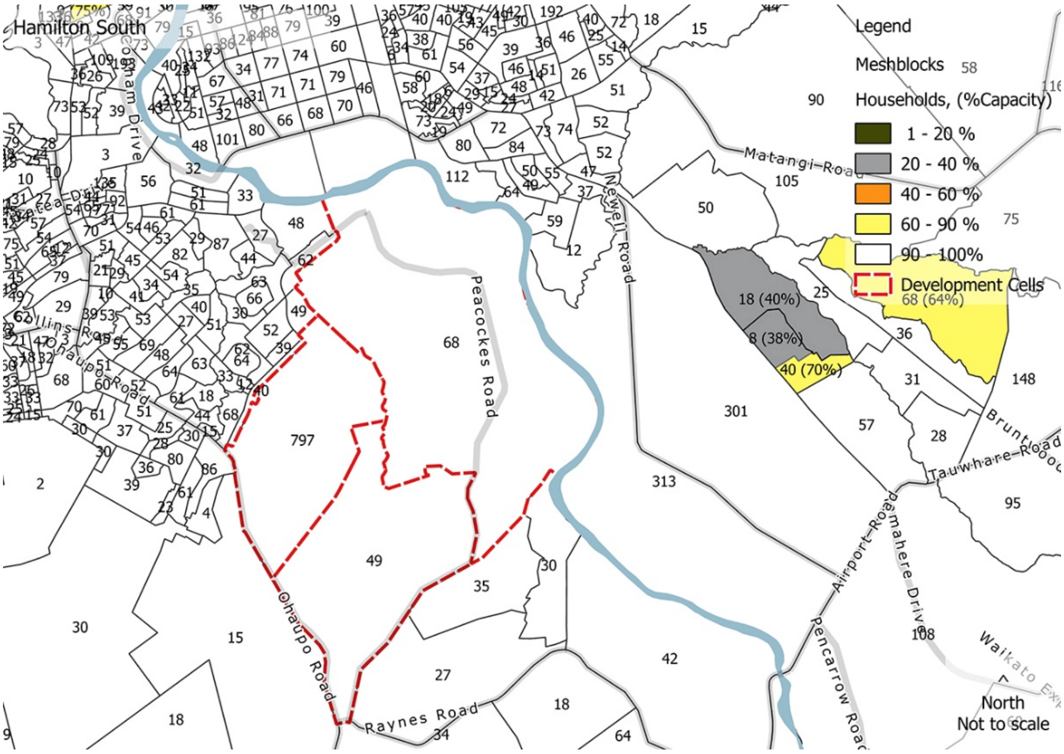
### 11.1.1 Hamilton




Map 23. Hamilton West, meshblock percentage utilised, 2025

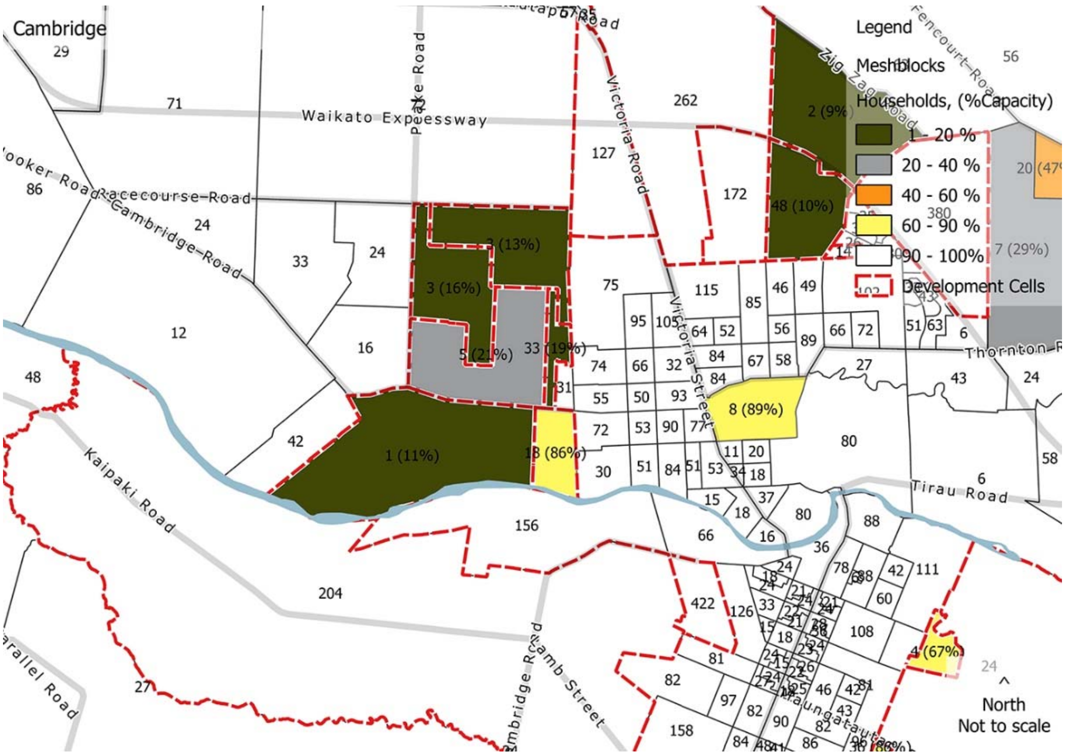
Map 25, the Rotokauri development cells to the west of Hamilton fill to the planned capacity. The two main development cells are to the west of Mangaharakeke Drive, and part of the services and infrastructure investment is planned within the model time period. These two meshblocks have 1,430 and 1,192 household agents locating by 2025. There is further planned development that will nearly double the capacity in timeframes beyond the model time period.

Map 26, Southern part of Hamilton, has a very large development cell with potential for about 8,500 residential properties. A small portion of this area has infrastructure in place and is utilised 797 HAs. Further infrastructure investment and development is planned beyond the time period of the model




Map 24. Hamilton South, meshblock percentage utilised, 2025

### 11.1.2 Waipa




Map 25. Cambridge, meshblock percentage utilised, 2025