Lars Brabyn

Declining Towns and Rapidly Growing Cities in New Zealand
developing an empirically-based model that can inform policy

Introduction

Understanding and predicting spatial patterns in population change has significant implications for infrastructure, property investments, and national spatial planning. It is also at the core of understanding what motivates people to move to different places, and the underlying geographical conditions that are important to people. During recent times, the population growth of large cities in New Zealand (particularly Auckland, but Tauranga has had faster growth) has resulted in severe social and infrastructural problems, such as sky-rocketing house prices, homelessness, and congestion of roads. At the same time, many small towns have had significant population decline, with no proposed solutions apart from acceptance or undertaking so-called "managed decline" (McMillan 2016; Wood 2017). As will

be described in this article, net migration has been a significant component of the spatial variation in population change, while natural change does not have a significant spatial variation and has been generally positive for all urban places. A policy response to the spatial variation of net migration needs to be based on an empirically based understanding of what drives net migration.

Dr Lars Brabyn, Senior Lecturer, School of Geography, University of Waikato (larsb@waikato.ac.nz).

The drivers of net migration are fundamental topics of Human Geography, Regional Economics, and Human Ecology. As a consequence there has been a considerable amount written on this topic (e.g. Lee 1966; Glaeser 2011; and Castles et al 2014), which is beyond the scope of this article to review. Historically the main drivers of net migration have been understood to be economic drivers such as resources and markets for industry and associated employment (Poot 1986; Barreira et al 2017). However, a range of surveys have shown that employment is often not the main motivation for migration (Morrison & Clark 2011). There is increasing interest in the role lifestyle has on migration (McGranaham 2008). Although this is often associated with lifestyle blocks (Andrew & Dymond 2013) and rural living, many small urban towns also provide a rural setting. Lifestyle is often

geographically models net migration in New Zealand, using the substantial amount of information on urban places which has been collected over decades, including both population counts and their environmental and social context. In New Zealand, the quantitative modelling of population change and mobility has been initiated by Grimes and Tarrant (2013) and Grimes et al (2016), which involved developing a data set of population change for 60 towns over the last 80 years. Other studies include Kerr et al (2004), who identified population density and education as pull factors. Our research builds on this research by focusing on the last 40 years, using a substantially greater number of places and more detailed data sets. In particular, net migration (rather than population change) is used as the dependant variable.

Zealand (2017), as the unit of spatial analysis. Urban areas include main urban areas, satellite urban communities, and independent urban communities. The rural centres had a baseline (1976) population between 300 and 999 in a 'reasonably compact area'. For this article, these urban areas and rural centres are all referred to as 'urban places'. In total, Statistics New Zealand identifies 276 urban places, but Waiouru and Burwood were removed from the dataset because they are military bases, and Pegasus was removed because of difficulty in determining early population counts. The choice of spatial units is an important decision in spatial modelling. Results are influenced by the unit of analysis, which is known as the modifiable area unit problem (Gehlke & Biehl 1934). Also, spatial units can have considerable internal heterogeneity, leading to the ecological fallacy (Robinson 1950). The benefit of urban places as spatial units is that they form natural units of analysis that are not biased by processes used to create artificial zones such as Territorial Authority Areas (TAs) or Census Area Units (CAUs). The extent of urban places is based on their natural growth, and they have a relatively consistent population density compared to many other larger units. Most importantly, they are also units that have meaning to people as they are associated with place-names and used in everyday conversations. The disadvantage is that each urban area is treated as one record even though the population count of urban areas can vary significantly. In 2013 the highest count was Central Auckland, with a population of 427,063, and the lowest count was Ohura with 126. Each urban place, regardless of size is a statistical record, and is treated equally. We do this deliberately, as our interest is in what attracts people to both large and small places.

> This article statistically and geographically models net migration in New Zealand, using the substantial amount of information on urban places which has been collected over decades …

linked to landscape setting – Vukomanovic & Orr (2014) have shown that rugged terrain, visual complexity and naturalness are drivers of 'amenity migration'.

Although qualitative surveys, such as used by Sloan and Morrison (2016), provide a picture of the drivers of net migration in New Zealand, it is important that understanding is substantiated through a range of approaches and methods. It is through triangulation of many studies and a range of methods that knowledge is generated. With the availability of geographic information system (GIS) data that describe places, both now and in the past (and combined with population census data that show where people live), it is timely to investigate how well changes in population distribution can be modelled with this data. This paper statistically and

By using just net migration, the model excludes natural change (births and deaths) which has different determinants. This should, in theory, result in a cleaner model. A machine learning technique, as opposed to an Ordinary Least Squared (OLS) regression model, became the statistical modelling technique of choice.

The main motivation of this research was to understand spatial differences in net migration, and to determine the relative influences of the different drivers and how these vary by age. Some drivers of net migration, such as employment, are well known, but others, such as lifestyle, are less well substantiated.

Urban Places Data and the Components of Population Change

This research used urban areas and rural centres, as defined by Statistics New

The development of the New Zealand urban places population data set, which contains natural change and net migration, is described in detail in the Appendix of Jackson & Brabyn (infra). Figure 1 compares natural change, net migration, and total change in population between 1976 and 2013 for the 273 urban places. Although natural change accounts for a substantial proportion of population growth in many places, and all of growth in

Figure 1: Components of population change (number and percentage contribution to change) between 1976 and 2013.
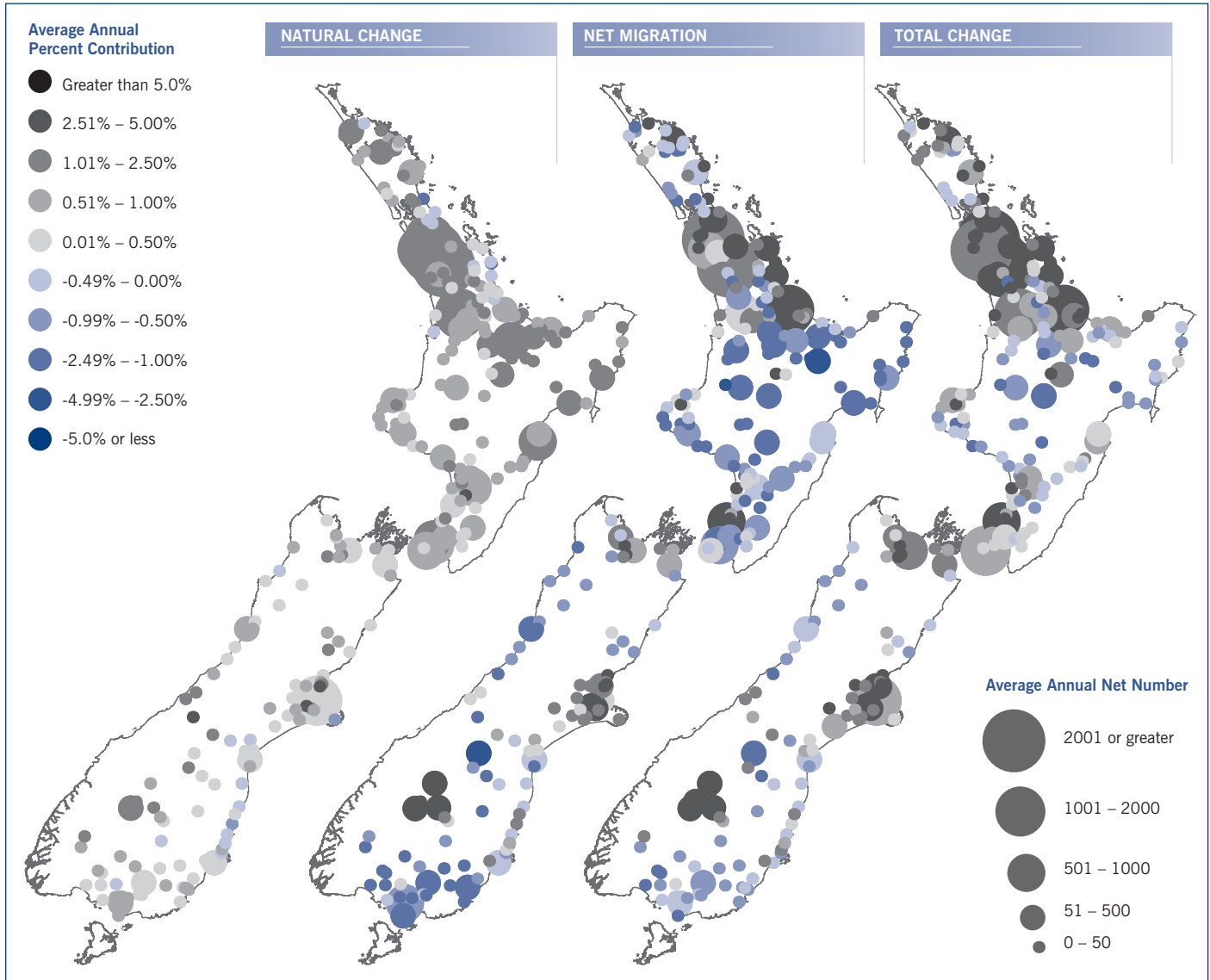
places experiencing severe net migration loss (see Jackson & Brabyn *infra*), this component, as a percentage, is relatively even across all towns. By contrast, net migration has considerable spatial variation, and accounts for over 95 percent of the variation in net population change. Separating natural change and net migration from net population change means that the analysis of drivers can focus entirely on net migration, and is not complicated by the need to consider the determinants of natural change, which include population age structure, and other determinants of birth and survivorship.

What Drives Net Migration? – The Conceptual Model

Table 1 summarizes the conceptual drivers of net migration that were used in

this research. In summary, these drivers are based around employment, natural and cultural lifestyle opportunities, access to essential services (hospitals and education facilities), and demographic age (technically this is a factor rather than a driver). Table 1 also includes general information how these drivers are represented by indices. The following section describes how these indices are represented using GIS and available data sets, but for now this section will focus on justifying these conceptual drivers. The drivers used in this study are conventional. For example, Chi and Ventura (2011) use similar categories, but there are some considerations that need further explanation.

The model used in this research uses 'ultimate' (that is, remote) drivers rather than proximate drivers, where possible.

Although the distinction of ultimate and proximate is open to interpretation, these are useful for looking at employment drivers. In this research, the key components of industry are used to represent employment, rather than the number of jobs or the presence of particular industries. The key components of industry are resources, which are linked to access to land-use and population size (and associated agglomeration theory), as well as access to markets, both domestic (large population centres), and international (ports and airports).

The conceptual model shown in Table 1 is based on the New Zealand context. New Zealand has not experienced, during the period being analysed, regional warfare or successive massive natural hazards that have resulted in large scale migration. The only exceptions are the

Table 1: Conceptual model of the drivers of net migration, and the indices that describe these drivers

Main Driver	Sub-driver	General Indices
Economic / Employment / Industry	Productive land-use in surrounding region	% in productive and non-productive land
	International transport hubs	Travel time to ports and airports
	Domestic markets	Travel time to different size population centres
	Critical population size – access to skills	Travel time to different size population centres
Lifestyle - Culture and Nature	Friends and family living close by	Travel time to different size population centres
	Cultural and social entertainment opportunities	Travel time to different size population centres
	Climate – sunny and not too wet	Annual sunshine hours and rainfall
	Aesthetic landscape setting / outdoor recreation	% of land in mountains and nature. Water views and coastal
Essential Services	Healthcare	Travel time to closest hospital
	Education	Travel time to education centres
	Transport, particularly airports	Travel time to airports
Demographic	Age classes	% population in different labour market age groups

Canterbury earthquake events in 2010/11, but to model the influence of earthquakes would require many more records that included other places. New Zealand has experienced many localized floods, but in general these have been ameliorated through flood barriers, or have a long enough reoccurrence interval to not induce significant net migration loss.

The importance of lifestyle does have a specific New Zealand context since it is linked to culture and local environment. Lifestyle choices for New Zealand will be different from other countries and cultures. Lifestyle as a determinant of migration is not as well developed and analysed, compared to the more traditional determinants such as employment and essential services. New Zealand is known for its scenic mountainous areas and coastal landscapes, which is the basis of its tourism industry; this industry also requires servicing, so tourist areas also attract people seeking employment. Anecdotally it is clear that scenic destinations such as Queenstown, and coastal sunny settlements around Nelson, the Coromandel, Bay of Plenty and northern New Zealand are popular places for tourists as well as places to retire. Many landscape studies show that people have a preference for natural, mountainous and coastal landscapes (Swaffield & Foster 2000; Brown & Brabyn 2012). In contrast, people also enjoy the cultural opportunities and diversity provided by large cities. Access to entertainment, such as large concerts, movie theatres, sports clubs and societies, is an important part of people's lifestyle choices. Large cities are also more likely to be where friends and family live, as well as culture peer groups for new immigrants.

Linked to the desirability of cities is improved access to essential services such as hospitals and large tertiary education institutes. Many of these services benefit from economies of scale, as they offer expensive technologies. Over the last 30 years in New Zealand there has been extensive change and consolidation of essential services, driven by the need to rationalize resources and reduce costs, but also by the difficulty of recruiting experts in small provincial towns. International airports are also considered an essential service for many, who want to take advantage of cheaper or more readily accessible international travel. Again, economies of scale are a key determinant of the number and location of international airports in New Zealand, especially given that New Zealand's population is only 4.7 million.

Compounding the complexity of different drivers of migration is age. People's wants and needs change with age. In this research four broad age groups are considered, each based on their labour market roles – labour market entry-age (15-24 years), prime working-age (24-54 years), labour market 'exit-age' (55-64 years), and retirement (65+ years)[1]. The labour market entry-age group may be keen to move away from home, may want to study at a tertiary institution, or move for work reasons and/or start a career. The prime working-age group is primarily focused on work and stable income for raising a family, while the labour market exit-age group (i.e., those in the 'retirement zone' and beginning to leave the labour market) may have considerable savings and be thinking about more desirable lifestyles. The retirement age group may begin with a focus on lifestyle but are more likely to be forced to move closer to tertiary hospitals as they age further.

Spatial Data Representation
This section outlines how the conceptual model described previously is represented and operationalized through spatial data sets collected from archives and derived from GIS analysis. Table 2 shows how the concepts that form the drivers are represented by 29 variables, and what datasets and GIS analysis functions were used to derive these variables. Key references and web links associated with these layers are also provided. The abbreviated titles of the variables are provided because these are used to present the results. The main driver type that each variable is represented is also listed, and it should be noted that some variables represent more than one driver type. The expected direction of influence for each variable is also provided, which is useful for checking that the statistical models being developed make sense. It is important to note that the travel time variables to essential services and infrastructure all have expected negative influence on net migration. This is because people want or need to be close to

Table 2: Description of independent variables used to model net migration

Abbreviated Variable Title	Full Variable Description	Main Driver Type	Expected Influence of Net Migration	Data Source	Analysis	Reference and Notes
Population size	Population count - all ages	Economic, Lifestyle Cultural	Positive	Statistics NZ Census	Sums	Statistics NZ
Entry-age 15-24	Percent labour market entry age 15-24 yrs	Demographic	Unknown	Statistics NZ Census	Sums	Statistics NZ
Working-age 20-64	Percent prime working-age 55-64 yrs	Demographic	Unknown	Statistics NZ Census	Sums	Statistics NZ
Exit-age 55-64	Percent labour market exit-age 55-64 yrs	Demographic	Unknown	Statistics NZ Census	Sums	Statistics NZ
Retire age 65 +	Percent retirement age 65+ yrs	Demographic	Unknown	Statistics NZ Census	Sums	Statistics NZ
Pop. > 20k	Travel (hrs) to the closest place with a pop. > 20K	Economic, Lifestyle Cultural	Negative	Statistics NZ Census	LCPA *	Statistics NZ
Pop. > 50k	Travel (hrs) to the closest place with a pop. > 50K	Economic, Lifestyle Cultural	Negative	Statistics NZ Census	LCPA	Statistics NZ
Pop. > 100k	Travel (hrs) to the closest place with a pop. > 100K	Economic, Lifestyle Cultural	Negative	Statistics NZ Census	LCPA	Statistics NZ
Pop. > 200k	Travel (hrs) to the closest place with a pop. > 200K	Economic, Lifestyle Cultural	Negative	Statistics NZ Census	LCPA	Statistics NZ
Large port	Travel (hrs) to the closest large container port	Economic	Negative	Prepared for project	LCPA	> 200k containers / yr.
Port	Travel (hrs) to the closest small or large port	Economic	Negative	Prepared for project	LCPA	Container port
Int'l. airport	Travel (hrs) to the closest international airport	Economic, Essential Services	Negative	Prepared for project	LCPA	Just international airports
Airport	Travel (hrs) to the closest airport	Economic, Essential Services	Negative	Prepared for project	LCPA	Regional and international
High producing	Hectares of high producing land-cover within 50km of the centre – dairy and horticulture	Economic	Positive	Land-cover database	Buffer	Newsome, 1987
Low producing	Hectares of low producing land-cover within 50km of the centre – sheep farming and forestry	Economic	Negative	Land-cover database	Buffer	Newsome, 1987
Urban	Hectares of urban land-cover within 50km of the centre	Economic	Positive	Land-cover database	Buffer	Newsome, 1987
Natural	Hectares of natural land-cover within 50km of the centre	Lifestyle Natural	Positive	NZ Landscape Classification	Buffer	Brabyn, 2009
Solar radiation	Solar radiation	Lifestyle Natural	Positive	LENZ dataset	Zonal	Leathwick et al 2002
Temperature	Mean annual temperature	Lifestyle Natural	Positive	LENZ dataset	Zonal	Leathwick et al 2002
Wind	Average wind speed	Lifestyle Natural	Negative	LENZ dataset	Zonal	Leathwick et al 2002
Rainfall	Average rainfall	Lifestyle Natural	Negative	LENZ dataset	Zonal	Leathwick et al 2002
Water view	Percent of urban area that has a water view (lake or sea)	Lifestyle Natural	Positive	NZ Landscape Classification	Sums	Brabyn, 2009

Abbreviated Variable Title	Full Variable Description	Main Driver Type	Expected Influence of Net Migration	Data Source	Analysis	Reference and Notes
Coastline	Total length of coastline within 50km of the centre	Lifestyle Natural	Positive	LINZ topog'l. data	Buffer	Land Information NZ
Mountains	Hectares of mountains within 50km of the centre	Lifestyle Natural	Positive	NZ landscape classification	Buffer	Brabyn, 2009
Hills	Hectares of hills within 50km of the centre	Economic	Negative	NZ Landscape Classification	Buffer	Brabyn, 2009
Major hospital	Travel (hrs) to the closest tertiary hospital	Essential Services	Negative	Prepared for project	LCPA	Tertiary
Hospital	Travel (hrs) to the closest hospital	Essential Services	Negative	Prepared for project	LCPA	Secondary + tertiary
University	Travel (hrs) to the closest university	Essential Services	Negative	Prepared for project	LCPA	Just universities
Tertiary education	Travel (hrs) to the closest tertiary institution	Essential Services	Negative	Prepared for project	LCPA	University, polytechnic or wananga

these services, so a reduced travel time is expected to be more desirable.

The sources of the data sets are also listed as well as the GIS analysis functions applied to these data sets and associated references for more information. The GIS analysis used are generally routine functions. Buffer (also known as proximity analysis) measures a set distance from a point or object. Zonal functions calculate the mean value for a region, and Least Cost Path Analysis (LCPA) calculates the shortest path via a network. The LINZ topographic road layer was used for the road network and the shortage path was based on time rather than distance – see Brabyn and Skelly (2002) for an example of how this is calculated.

Where applicable and available, the values for each variable were derived for the base year of each period. For example, the 1981 (Newsome, 1987) landcover layer was used to determine the extent of productive, natural, and urban land uses. Some variables like climate do not change significantly and 2002 data was used. The location of essential services and infrastructure (hospitals, tertiary education, ports and airports) does change over time and compiling this information was a major task. Archived information available on the internet was often the best source of this information. Hospitals reduced in number over the analysis period, primarily due to costs and government restructuring. Tertiary institutions combined but created many

learning outposts in smaller centres – primarily due to costs and government policy to upskill the population. Airports fluctuated from national carrier to regional carriers and vice versa – primarily due to demand and costs. Ports increased the volume of cargo moving through their facilities over the period - due to increased trade and population demand.

Statistical modelling
Ideally, a statistical model will explain all the variation in net migration and provide the relative importance that each variable makes to the model. For example, if a new hospital was built in a town, the perfect model would be able to predict the net migration resulting from this change. Ordinary Least Squared (OLS) regression models provide coefficient values that give an exact break down of what each variable contributes to the model; however there are significant stability issues when the variables are not independent (have collinearity). This is a problem with the variables in this study as well as other studies (see Grimes et al 2016), and is to be expected given that airports, universities and hospitals are often located in major cities, so the travel time to these services will often be similar for different urban places. When two or more variables correlate, it is impossible to know which variable is the actual driver. The other concern with regressions is that outliers often have a significant weighting on coefficient values, and

the removal of one or two records can have a huge change to the coefficient value. This research investigated the use of stepwise OLS regression with standardized variables, but soon realized that the significant coefficient values were meaningless, as they change dramatically when one variable is dropped or added. The use of Principle Component Analysis was also used to reduce the number of variables and collinearity, but it became impossible to interpret the results and what each component actually meant. The significant coefficients were also highly unstable.

Random Forest (RF) is a machine learning technique that resolves many of the instability problems with regression models (Rodriguez-Galiano et al 2012), and was one of the adopted methods used in the analysis – the other method was Pearson's 'r'[2]. The output from a Random Forest model is an assessment of the variation in the dependant variable that the model can explain, as well as a set of importance values associated with each variable. By comparing the importance values for each variable, the relative importance (RI) of each variable is determined. Random Forest uses decision trees to produce a model. The variables used at each branch in a tree are selected randomly based on a set number of variables specified by the user (set to 4 for this study), and the number of branches used is determined by the number of variables (Breiman 2001). The random

Table 3: Net migration by different age groups - model description and performance

Dependent Variables % Net Migration	All Ages Combined			Entry Age Group (15-24 years)			Prime Age Group (25-54 years)			Exit Age Group (55-64 years)			Retirement Age Group (65+ years)		
		RI	ρ		RI	ρ		RI	ρ		RI	ρ		RI	ρ
	Entry-age 15-24	1.00	-0.35	Population size	1	0.313	Population size	1.00	-0.05	Population size	1.00	-0.11	Population size	1.00	0.017
	Exit-age 55-64	0.99	0.46	Temperature	0.92	-0.36	Water view	0.53	0.402	Water view	0.86	0.375	Int'l airport	0.69	-0.25
	Mountains	0.65	0.04	Pop. > 50k	0.92	0.051	Hills	0.36	-0.31	Hills	0.53	-0.19	Population > 200k	0.56	-0.24
	Population size	0.62	-0.03	Natural	0.89	0.251	Mountains	0.30	0.054	Solar radiation	0.52	0.156	Low producing	0.45	-0.15
Top 10 Explanatory Variables Listed by Order of Importance	Retire age 65+	0.57	0.36	Hospital	0.88	0.048	Solar radiation	0.27	0.191	Temperature	0.49	0.158	Temperature	0.40	0.196
	Hills	0.43	-0.32	Hills	0.89	-0.2	Coastline	0.25	0.304	Int'l airport	0.49	-0.21	Coastline	0.38	0.203
	Working-age 20-64	0.37	0.27	Rainfall	0.88	0.103	Natural	0.21	-0.03	Low producing	0.48	-0.03	Major hospital	0.37	-0.2
	Coastline	0.33	0.32	University	0.87	-0.04	Temperature	0.20	0.168	Rainfall	0.48	-0.1	Large port	0.32	-0.08
	Solar radiation	0.31	0.20	Population > 20k	0.73	0.133	Hospital	0.20	0.13	Hospital	0.39	0.127	Hills	0.30	-0.11
	Int'l airport	0.28	-0.31	Tertiary education	0.6	-0.03	Low producing	0.18	-0.04	Mountains	0.39	0.049	Airport	0.30	-0.15
Variability Explained	0.47			0.43			0.41			0.24			0.08		

decision tree is repeated several hundred times (set to 500 for this study) to produce a 'random forest'. Generally, after 100 repeats, the results are consistent. The data records were divided into a bootstrap sample ('in bag' data) and 'out-of-bag' (OOB) data using a split of 2/3 and 1/3 respectively (Rodriguez-Galiano et al 2012). The OOB data are used to calculate a prediction error by comparing the predictions from the in-bag data with the OOB data. The OOB samples are also used to calculate the variable importance (the prediction strength of each variable) by changing randomly the values of a given variable in the OOB samples. The increase of OOB error from these changes are averaged over all trees and is a measure of the importance of the variable.

As noted previously, when several hundred random trees are used, the relative importance of each variable becomes stable, but the value is relevant only to that particular model and is useful only as a relative measure. For this reason it is not valid to compare the magnitude of the importance value of a variable with another model generated with different data, such as another year; however, it is valid to compare the relative position of the importance value of a variable across different models. The importance values provide an estimate of the significance the variable has to the model. The lower ranked variables generally make only a small contribution to the model. In order to determine the relative importance (RI) of each variable the importance values were normalised based on their position relative to the maximum and minimum values ((value – min)/(max-min)).

The advantage of RF is that it is non-parametric, is relatively quick at processing large data sets with many variables, and is relatively stable with outliers and collinearity. The disadvantage with RF is that it needs many records, and provides no indication on the direction (positive or negative) of the relationship between the dependent and independent variables. For example, if mean annual temperature of a town is identified as an important variable, it is not known from RF whether there is a positive or negative influence on net migration. To get around this, Pearson's r correlation was used in conjunction with Random Forest to provide an indication of the direction.

RF was used to model net migration between 1976 and 2013. When all age groups were combined the 29 variables listed in Table 2 were used. When the net migration of each individual age group was modelled, all four age variables were not used, so the model was based on 25 variables.

Results

Table 3 shows the results in terms of percent net migration for the different age groups. This table lists the top ten variables in order of their importance values for the different periods. Each variable contains the relative importance values (normalised) as well as its Pearson's r (ρ) value to show the direction of the influence. In general, the directions of influence are consistent with the expected directions. As previously stated, travel time to essential services is negative because increased travel time is not favoured by people. Interpretation of the magnitude of the Pearson's r values requires care because of the collinearity between values, but it still provides additional information. Pearson's r (ρ) values between -0.1 and +0.1 were not significant at p < 0.10, since

n=273. This table also shows how much of the variation is explained by the RF models.

Observation 1: Overall model performance is good.

The best RF model explains 47 percent of the variation in net migration. Although this can be considered low, it is common in predictive spatial modelling at an ecological level for model performance to be around 20-30 percent (Jansen, Judas & Saborowski 2002). Given the complex nature of the human world and the diverse range of values among people, a model that predicts 47 percent of the spatial variability is a good achievement. However, this model performance is less than that achieved by Grimes et al (2016), which was close to 60 percent. An explanation for this result could be because

have an influence on the net migration of another group. The exception is children moving with their parents, but children were not analysed separately in this study.

Observation 2: Age is a significant factor that determines net migration.

The results show that labour market age groups are a significant factor that determines net migration. The results for all ages combined show that the labour market entry-age group (%15-24) years and the labour market exit-age group (%55-64) were the top two most important variables. The Pearson's r shows a significant positive correlation for the exit-age group and a significant negative correlation for the entry-age. The positive correlation for the exit age group implies that places with high percentages of labour market exit-age

Observation 3: Population size is important but varies with age group.

The high RI value for population size for all age groups implies that the population size of the urban place had an influence, but the Pearson's r does not show a clear direction. It was expected that many small towns would have had net migration loss, which is the case, but there are also many that have had a gain for lifestyle reasons. To understand the influence of population size, it is necessary to examine this by the different labour market age groups. For the labour market entry-age group, the Pearson's r is positive and significant, indicating that the entry-age group is moving from small towns to large cities. With the prime working-age, population size is important but the direction is not clear because the Pearson's r value is not significant. The exit-age group has a significant negative correlation with population size, indicating a move from large urban places to small urban places. This is consistent with the idea that people close to retirement are seeking places with high natural lifestyle values, which, judging by the high RI values for water views, appears to be small coastal and lakeside towns. The retirement-zone age group has a significant positive correlation with population size, indicating that this group has a preference for larger cities, where hospital care is available. Related to population size is distance to different size cities. People may seek the benefits of a large population centre, but choose to live in satellite towns. These variables do appear in the list of top ten variables but do not have particularly high RI values compared to population size. Distance to cities with a population greater than 200,000 is significant for the retirement group. The correlation is negative indicating this groups prefers being close to large cities, rather than in them.

Although the determinants of net migration are highly complex, this research has produced models of net migration that explain close to 50 percent of the variation across 273 urban places.

our study used 273 towns compared to 56 towns for the study by Grimes et al. These 274 towns included many small rural centres, and as a consequence there would be greater diversity in characteristics, making it more difficult to model. Small towns also have greater data variability for percent net migration because of the smaller denominator values.

The performance of the individual RF models for the different labour market age groups is less than when all age groups are combined. This is surprising because it was expected that there would be less variation within an age group and therefore be easier to model. The age variables were not included in these models, which reduced the model performance. Age variables do not have a logical explanation for improving the model. In general, one age group will not

people had high levels of net migration gain, and the negative correlation for the entry-age group implies that places with high percentages of entry age people had net migration loss. The likely explanation for this is that many retirement (lifestyle) destinations had a high percentage of older people and these places have high positive net migration, which is likely to be more retirees seeking a retirement destination. Many small farming towns had a high percentage of young people and these towns have net migration loss. This is likely to be the result of young people moving from these small towns to cities, and retirees moving to retirement destinations. The importance of age prompted separate modelling of each age group.

Observation 4: Economic and lifestyle are both important drivers.

As discussed previously, lifestyle drivers such as water views are important, but variables associated with employment and the economy are also consistently important. Surrounding mountains and natural land-cover are important variables, but the correlation is not

consistently significant. This could reflect the boom in tourism during these years, and many 'alpine' towns such as Queenstown, Wanaka, and Fox Glacier had high net migration gains. Being close to international airports also had a highly-ranked RI value. This is consistent with lifestyle values, since airports give access to international travel; however, it could also be interpreted as an employment driver as airports provide access to international markets and this helps businesses. Surrounding hills had a negative influence on net migration and this is linked to employment. In the New Zealand context, hill country is associated with sheep farming, which has been in decline, resulting in towns servicing these regions having net migration loss.

Observation 5. Access to essential services is important but not highly ranked.

There is a general trend of health services being consolidated in the larger cities, and this is an important consideration for many people. The variable, 'travel time to a hospital' appears relatively frequently in the top 10 list, but is mainly ranked as a mid-range variable. The other hospital variable 'travel time to a tertiary hospital' only appears for the 65+ age group (we might expect this to be higher again at 75+ years). Tertiary education appears to be important for the labour market entry-age group only, which is to be expected.

Conclusion

Although the determinants of net migration are highly complex, this research has produced models of net migration that explain close to 50 percent of the variation across 273 urban places. The models that have been developed show that the determinants of net migration are considerably influenced by age. Labour market entry and exit ages are primary factors that had a strong influence on net migration. Further analysis of these groups showed that labour market entry-age people are moving to the more populated places with access to tertiary education and different types of work, while the labour market exit-age group is moving to places with high natural lifestyle value - water views, mountains, and warm temperatures. The retirement-zone age group was also an influential group, and there appears to be a movement away from small farming towns to slightly larger towns, with access to international airports and warmer temperatures.

Although the drivers of net migration are reasonably well understood, much of this understanding has been based on traditional drivers of employment and essential services. This research has substantiated this evidence but also modelled the importance of lifestyle drivers, which are increasingly being recognised. Lifestyle choice, both nature based (mountains, climate and water views) and, possibly, cultural (access to large cities) are just as influential as employment drivers and access to essential services. This research, however, has not used indices that distinguish cultural lifestyle drivers from economic drivers, as both are based on access to large cities. It was therefore inconclusive whether cultural lifestyle is important, although anecdotal evidence would suggest it is.

In New Zealand there is general concern over the future of many towns that are consistently declining in population due to net migration loss. A policy suggestion for these towns has been 'managed decline' (McMillan 2016; Wood 2017); however this research indicates that towns close to airports and with natural amenity value are especially favoured by the labour market exit age group. With population ageing this group's size will increase and there will be increased movement of people to such towns. New Zealand has many towns with high natural amenity value that are close to the coast and/or among the mountains. Rather than having a 'managed decline' policy response to many towns in decline, this research has shown that a potential policy response is to make such towns more attractive to people nearing retirement, especially towns with regional airports and access to medical facilities. This research has also emphasised the importance of having quality empirical data and analysis when developing policy responses to net migration within New Zealand.

---

1   New Zealand does not have a formal retirement age. These groupings reflect observed labour force movements by age, with male full-time employment at the 2013 Census dropping from 65 per cent at 60-64 years to just on 18 per cent at 65+ years, and for females, from 39 per cent to 7 per cent. In 2013, 11 per cent of males and 10 per cent of females aged 65+ years were also employed part-time; for males this was similar to levels at 60-64 years of age, while for females the level had halved (Statistics New Zealand 2015).

2   The Pearson product moment correlation coefficient 'r' measures the linear strength of the relationship between two variables, with +1 meaning that both variables increase and there is a perfect positive correlation, 0 meaning there is no linear relationship, and -1 meaning there is a perfect negative correlation (one variable increases while the other decreases).

References

Andrew, R & J Dymond (2013) 'Expansion of lifestyle blocks and urban areas onto high-class land: an update for planning and policy' *Journal of the Royal Society of New Zealand* 43(3), pp.128-140

Barreira, AP, S Ramalho, JJ Panagopoulos & MH Guimarães (2017) 'Factors driving the population growth and decline of Portuguese cities' *Growth and change* doi:10.1111/grow.12205

Brabyn, L (2009) 'Classifying landscape character' *Landscape Research* 34(3), pp.299-321

Brabyn, L & C Skelly (2002) 'Modeling population access to New Zealand public hospitals' *International Journal of Health Geographics* 1(3) DOI: 10.1186/1476-072X-1-3

Breiman, L (2001) 'Random forests' *Machine Learning* 45(1), pp.5-32

Brown, G & L Brabyn (2012) 'An analysis of the relationships between multiple values and physical landscapes at a regional scale using public participation GIS and landscape character classification' *Landscape and Urban Planning* 107(3), pp.317-331

Castles, S, de Haas H & M J Miller (2014) *The Age of Migration: International Population Movements in the Modern World* (5th ed.) London: Palgrave Macmillan

Chi, G & SJ Ventura (2011) 'An Integrated Framework of Population Change: Influential Factors, Spatial Dynamics, and Temporal Variation' *Growth and Change – A journal of Urban and Regional Policy* 42(4), pp.549-570

Gehlke, CE & K Biehl (1934) 'Certain effects of grouping upon the size of the correlation coefficient in census tract material' *Journal of the American Statistical Association* 29(185A), pp.169–170

Glaeser, E (2011) *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier* New York: Penguin Press

Grimes, A & N Tarrant (2013) 'A New Zealand urban population database' *Motu Working Paper* 13-07 Wellington: Motu Economic and Public Policy Research

Grimes, A, E Apatov, L Lutchman & A Robinson (2016) 'Eighty years of urban development in New Zealand: impacts of economic and natural factors' *New Zealand Economic Papers* 50(3), pp.303-322

Jansen, M, M Judas, & J Saborowski (2002) *Spatial modelling in Forest Ecology and Management.* New York: Springer

Kerr, S, D Maré, W Power & J Timmins (2004) 'Internal mobility in New Zealand' *Treasury Working Paper 01/04* Wellington: Motu Economic and Public Policy Research Trust

Leathwick, J, F Morgan, G Wilson, D Rutledge, M McLeod & K Johnson (2002) 'Land environments of New Zealand: a technical guide' Ministry for the Environment, New Zealand

Lee, ES (1966) 'A theory of migration' *Demography* 3, pp.47-57

McGranahan, DA (2008) 'Landscape influence on recent rural migration in the U.S.' *Landscape and Urban Planning* 30, pp.228–224

McMillan, R (2016) 'The shrinkage pathway: managing regional depopulation', in P Spoonley (ed) *Rebooting the Regions: How low or zero growth needn't mean the end of prosperity* pp.213-237, Albany, NZ: Massey University Press

Morrison, PS & WAV Clark (2011) 'Internal migration and employment: Macro flows and micro motives' *Environment and Planning A* 43(8), 1948-1964

Newsome, P (1987) 'The vegetation cover of New Zealand' Wellington: Ministry of Works and Development

Poot, J (1986) 'A system approach to modelling the inter-urban exchange of workers in New Zealand' *Scottish Journal of Political Economy,* 33(3), pp.249-274

Robinson, WS (1950) 'Ecological correlations and the behavior of Individuals' *American Sociological Review* 15(3), pp.351-357

Rodriguez-Galiano, VF, B Ghimire, J Rogan, M Chica-Olmo & J P Rigol-Sanchez (2012) 'An assessment of the effectiveness of a random forest classifier for land-cover classification' *ISPRS Journal of Photogrammetry and Remote Sensing* 67(0), pp.93-104

Sloan, M & PS Morrison (2016) 'The post-move satisfaction of internal migrants' *Migration Studies* 4(1), pp.97-115

Statistics New Zealand (2015) 'Area of usual residence, work and labour force status for the census usual resident population count aged 15 years and over' Customised database

Statistics New Zealand (2017) 'Classification of Urban Area' www.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/urban-area.aspx

Swaffield, SR & RJ Foster (2000) 'Community perceptions of landscape values in the South Island high country: a literature review of current knowledge and evaluation of survey methods' Wellington, NZ: Department of Conservation

Vukomanovic, J & BJ Orr (2014) 'Landscape aesthetics and the scenic drivers of amenity migration in the New West: Naturalness, visual scale, and complexity' *Land* 3(2), pp.390-413

Wood, J (2017) 'Growing beyond growth. Rethinking the goals of regional development in New Zealand' *Maxim Institute Discussion Paper* Auckland, NZ: Maxim Institute