Key Factors Affecting Journey to Work in Melbourne using Geographically Weighted Regression

Daniel Paez*[1] and Graham Currie[2]

*[1]Daniel Paez
Manager Active Transport Policy, Department of Transport, Level 15 - 121 Exhibition Building, Melbourne VIC 3006, Phone: +61 + 4 211 07012 Email: daniel.paez@transport.vic.gov.au
Formerly: Associate, Booz & Company, Victoria

[2]Graham Currie
Professor, Chair in Public Transport, Institute of Transport Studies, Department of Civil Engineering, Building 60, Monash University, Clayton, Victoria 3800, AUSTRALIA and Senior Associate, Booz & Company Phone: +61 3 9905 5574, Fax: +61 3 9905 4944, Email: graham.currie@eng.monash.edu.au

*Corresponding author

Abstract

This paper aims to explore factors affecting private car use for journey to work in Melbourne using a multiple regression analysis and an analysis of the spatial strength of explanatory variables using a geographically weighted regression analysis.

A review of previous research suggests that car ownership, access to transit, distance from the CBD, CBD employee share, and to some extent urban residential and employment density are major factors affecting private vehicle mode share for commuting. There is some variation in specific factors explored and the degree to which each is of influence suggesting that local conditions in each city or study area have an influence in why people chose driving for JTW.

The multiple regression analysis explained 57% of the variance of car use for travel to work in Melbourne. Three variable explained most of the variation in auto use including level of public transport supply (-0.443), distance from the CBD (0.404) and residential density (-0.136). All other variables had very little net influence on car use for JTW (all were below +/- 0.1). In general these findings are much in line with previous research.

GWR increased the explanatory power of the analysis from 57% to 74% of variance and enabled spatial patterns of the three major explanatory variables to be explored. Public transport supply was found to be particularly strong at influencing private vehicle commuter in inner and south western parts of Melbourne. Distance to the city was stronger in inner and northern parts of Melbourne. Residential density was strongest in eastern and outer areas.

.

1 Introduction

Traffic congestion is widely recognised as a growing problem in western economies including Australia (Competition and Regulation Working Group, 2006, Victorian Competition & Efficiency Commission, 2006, Bureau of Transport and Regional Economics, 2007).   Congestion problems are focussed on peak periods and commuting travel by cars. A major objective of approaches to deal with traffic congestion is to reduce car use for journey to work (JTW). However the effectiveness of measures to address work car travel depends greatly on our understanding of the factors driving car use at these times.

This paper presents the results of an empirical analysis - using geographic information systems and census data - aimed at understanding urban form and transport factors that affect car usage for JTW.

The focus of the work is car commuting in Melbourne, Australia. The analysis was undertaken as part of a wider study of car dependence undertaken for the Australian Conservation Foundation (Booz & Co., 2010). The paper presents the approach and findings of the analysis on the empirical analysis of factors affecting car use for JTW.

The paper starts with a brief review of the relevant research literature in this field. It then describes the methodology and data used in the analysis. The results are then presented. The paper concludes with a discussion of the results including some suggestions for how the method might be adopted for trip forecasting. Areas for future research are also suggested.

2 Research context

The major relevant focus of previous research to this paper concerns research related to:
- the factors of journey to work mode choice ;and
- the methodology adopted for the papers analysis; geographically weighted regression (GWR).

Journey to Work Mode Choice Drivers

Commuter mode split factors has been a major focus of research in transport. Much of this has examined links between urban residential densities and commute mode split but with mixed results. Some studies have found minimal or zero impact (Crane and Crepeau, 1998) while others identify more significant links (Geurs et al., 2006).

One analysis explored the relationship between residential patterns and mode choice for a large employment centre in Los Angeles (Zhou et al., 2007). The key factors of mode choice were found to be residence distance from employment and proximity to transit lines.

An exploration of data from the US national transportation survey (Chatman, 2003) also aimed to understand factors affecting commuter mode choice. Higher employment density at the workplace was found to be associated with lower car commuting. A related study explored employment site characteristics and their impact on commuter mode choice using travel survey evidence in Israel (Shiftan and Barlach, 2002). They identified proximity to CBD's as a major driver of transit commuting and also areas with higher quality transit services. Automobile availability was found to be a major driver of car commuting.

In the Australian context a number of studies have sought to better understanding factors affecting commuter mode choice.

One of the earliest studies in this area was by Abelson and Baker who adopted a conventional multiple regression model to identify factors affecting car ownership in Sydney and how this related to mode choice for journey to work (Abelson and Baker, 1982). Factors identified as affecting commuter mode choice were access to a car, access to public transport, and the degree of congestion on roads. The two latter variables were strongly correlated with residential density and inversely with distance from the CBD. Car ownership was also strongly related to income. All other things being equal, larger household size was also found to increase car ownership and hence impact on auto mode choice.

A more recent analysis of commuter mode choice in Sydney was undertaken using 2001 census data (Longworth and Wilson, 2006). This analysis undertook a simple comparative assessment of mode share for a series of disaggregate groups of data including resident workforce density, proximity to rail, proportion of CBD workers, and proximity to CBD and regional centres. Only a weak association between residential density and transit mode share was identified in inner and middle ring areas however a closer link was found in outer areas. Proximity to rail was much more closely related to rail commute share with a 0.57 R^2 in middle ring areas. An R^2 of 0.63 was found between transit commute share and the proportion of workers who work in the CBD. Overall proximity of home to the CBD was not strongly related to rail mode share.

Another recent study used a multiple regression model exploring auto commute share in Australian major cities including Melbourne (Rickwood and Glazebrook, 2009). This found that distance from the city and car ownership were major factors. Local area density had an indirect affect because it reduced car ownership. Links between local density and transit access were also considered important.

A related recent analysis adopted geographically weighted regression (GWR) to explore factors affecting vehicle kilometres travelled - VKT (Mulley and Tanner, 2009). The aim was to understand for Sydney how GWR could improve quality of modelled VKT compared to traditional linear regression models. The study found a superior fit from a statistical perspective of prediction of VKT when GWR is used, suggesting the method is superior to traditional approaches. This global model identified household car ownership, access to public transport and residential and employment density as key factors of VKT. While this result concerns a broader trip purpose that just JTW, results from Mulley and Tanner (2009) follow the same pattern in relation to key factors as results in previous research.

In summary factors affecting auto choice for JTW; overall previous research suggests car ownership, access to transit, distance from the CBD, CBD employee share, and to some extent urban residential and employment density are major factors of commuter mode share. However there is some variation in specific factors explored and the degree to which each is of influence suggesting that local conditions in each city or study area have an influence in why people chose driving for JTW.

Geographically Weighted Regression

Geographically weighted regression (GWR) is a technique for exploratory spatial data analysis that is normally applied using a geographic information systems (GIS) platform. The main purpose of using GWR is to explore relationships between data incorporating geographic or spatial factors such as proximity.

In standard regression modelled relationships hold no geographic or proximity elements unless explicitly included in the input data. GWR enables exploration of geographical

patterns in the residuals of the models. GWR can improve on conventional multiple regression analysis by modelling the spatial relationships found in regression analysis.

GWR can be used to fit linear, Gaussian, logistic or Poisson multivariable models. In no case would a GWR model produce a worse result or fit than what the non-spatial inputs or standard regression would produce. Indeed generally it improves on those relationships by exploring geographical elements of the data.

GWR has been extensively applied in many natural sciences areas such as meteorology and geology (Fotheringham et al., 2002) where the statistical and mathematical validity of GWR have been detailed including multiple examples of its application.

In the past, the application of GWR required extensive computational processing and the use of custom software not openly available. It is only recently that ESRI (a commercial GIS software producer) has included the methodology as part of the suite of tools available in ArcGIS, a popular GIS platform. This inclusion into commercial software has allowed for an increased used of GWR in many areas including transport and land use planning, which is the area applied in this research.

A detailed description on the mathematical procedures used in ArcGIS has been provided in the literature including a guide to understanding results (Charlton and Fotherringham, 2009b). A complete tutorial on how to use the tool in ArcGIS has also been developed (Charlton and Fotherringham, 2009a).

3 Methodology and model for Melbourne

The objective of the analysis was to identify key factors that explain mode choice of car for journey to work (JTW).

The model was developed around 2006 census data (Australian Bureau of Statistics, 2006) and all variables were analysed and calculated using census collection districts (CDD) as the main geographic unit. There are some 5,506 of these within the Melbourne Statistical District.

In developing the model a simple four step methodology was used (Figure 1). These steps are now described.

Figure 1: Model Development Method

Variable design and calculation

In the model, the dependant variable (or variable to be explained) is the proportion of people that choose private vehicle (as driver or passenger) for JTW in the 2006 census.

The independent or explanatory variables were calculated based on urban form and transport factors.

The first step was to decide on the independent variables to be explored in the model and then to calculate them for each CCD. Variables to be explored were selected based on the availability of information and their relevance according to previous literature.

Table 1 shows the independent or explanatory variables explored in the analysis and indicates their source.

Table 1: Independent Variables Explored in the Analysis

Variable	Description	Data source	GIS calculation
Public Transport Supply	Estimation of the level of service of Public Transport in each CCD based on access, frequency and spam	Currie (2010)	Not required as already in the geographic unit
Public Transport Ranking	Relative ranking of each CCD in relation to PT supply	Currie (2010)	Not required as already in the geographic unit
Residential Density	Total number of people divided by area	Census 2006	Not required as Census 2006 data already in the geographic unit
Distance to Business Zone	Linear distance to the closest business 1 planning zone	Victorian Planning Schemes	Distance calculated from the centre of each polygon
Distance to Rail Station	Linear distance to the closest metropolitan train station	DOT point data	Distance calculated from centre of polygon to station (point data)
Distance to Melbourne CBD	Total distance using the shortest path in the road network	Vicdata Road dataset.	Shortest path algorithm from centre of polygon to corner of Collins and Swanston as the CBD centre
Distance to Local Activity Centre	Linear distance to the closest activity centre (CAD, PAC or MAC)	Melbourne 2030	Centre of polygon to activity centre (point data)
Distance to Arterial Road	Linear distance to the closest arterial road	Vicdata road dataset	Shortest distance from polygon to line
Distance to Highway	Linear distance to the closest Highway	Vicdata road dataset	Shortest distance from polygon to line
Provision of Roads	Level of service in relation to roads within the CCD	Vicdata Road dataset.	Sum of the total length of arterial and highways within the polygon
Provision of Cycling	Ratio between the Cycling network in the area and total number of persons	VicRoads Principal Bicycle Network (PBN)	Sum of total length of sections of the PBN within the polygon

Two measures of public transport supply were explored both derived from previous recent research (Currie, 2010). Here an index of public transport service level was

derived from the frequency of services each week supplied at individual bus stops/stations factored for walk access distances to stops/stations within each CCD. The final index measure and the relative ranking of each CCD were explored in the analysis.

Standard Regression Model

Once the variables were calculated, a series of standard linear multiple regression models were developed. Over 15 combinations of the independent variables were processed using a trial and error methodology. In this forward exploration and backward exploration were used. Forward exploration involves starting with no variables in the model and trying them one by one and including them if they appear as 'statistically significant'. Backward exploration starts with all candidate variables testing the significant of the model when the variables are removed.

The final model - which met the criteria of being statistically viable and at the same time with the highest level of significance - was ultimately developed combining both backward and forward exploration.

Geographically Weighted Regression Model

The GRW model aimed to explore the spatial aspects of the multiple regression outputs rather than generating new aggregate measures of the links between the dependent and independent variables. The more significant variables are included in the GRW model using ArcGIS 9.3. The model was run using an adaptive kernel (Charlton and Fotherringham, 2009a) as it provided a better representation of the spatial interaction between variables.

Spatial Analysis and Results

Standard regression results involve the assessment of overall model statistical fit and the significance and weighting of the links between individual independent variables and the dependent variable. The GWR outputs are primarily spatial maps of the relative strength of links between individual independent variables and the dependent variable.

4 Results

Table 2 shows the results of the multiple regression analysis of the variables explored. The overall fit of the model (adjusted R^2) was 0.57 suggesting that the model explained 57% of the variance of car use for travel to work. All variables examined were significant at the 99% level however 3 main variables explained most of the variance in JTW car use. These were:

- public transport supply (standardised beta = -0.443),

- proximity from the CBD (standardised beta = 0.404) ; and

- residential density (standardised beta = -0.136).

All other variables had very little net influence on car use for JTW (all were below a standardised beta of +/- 0.1. Tests on the data set established little concern about multi-colinearity in the variables.

Table 2: Results of the Multiple Regression Analysis

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.756081099
R Square	0.571658628
Adjusted R Square	0.570957196
Standard Error	0.083805347
Observations	5506

	Unstandardised		Standardised		
	Coefficients	Standard Error	Beta Coefficients	t Stat	P-value
Intercept	0.638430345	0.0035398		180.3577	0
Density	-1.46101E-05	1.13693E-06	-.136	-12.8506	2.92388E-37
Total Supply	-1.10632E-05	2.55574E-07	-.443	-43.2877	0
Dist_to_Business	-0.012772074	0.001818474	-.076	-7.02351	2.42599E-12
Dist_to_Activity Centre	-0.004401445	0.000642713	-.093	-6.84823	8.294E-12
Dist_to_rail	-0.001367273	0.000407942	-.042	-3.35163	0.000808784
DisCBD	0.377274374	0.012635193	.404	29.85901	1.1136E-181
DisHwy	0.149029249	0.077255562	.020	1.929042	0.053777026
DisArt	0.270160363	0.133644902	.020	2.021479	0.043278569
DensPBN	-1.1425E-07	1.41038E-08	-.075	-8.10065	6.67838E-16

Figure 2 presents the resulting simplified illustration used to simplify the results for a wider non-technical readership (this was a major aim of the analysis).
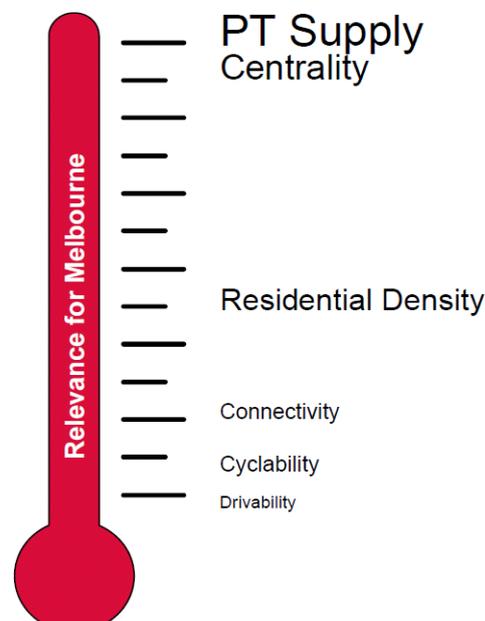

Figure 2: Simplified Regression Output – Key Drivers of Car JTW in Melbourne

Exploring Regression Results Using GRW

Table 3 illustrates the key statistical results from the GRW model after the regression results were input to the system. The Adjusted R^2 was 0.75 suggesting an improvement on the non-spatial multiple regression model.

Table 3: Statistical Fit of the Geographically Weighted Regression Model

NAME	VALUE
Bandwidth	0.27319257992
ResidualSquares	22.34178023090
EffectiveNumber	58.97956002890
Sigma	0.06404413809
AICc	-14590.88225500000
R2	0.75207654510
R2Adjusted	0.74943758074

The spatial aspects of the top three explanatory variables were explored using GWR. Outputs mapped the spatial strength of their strength in explaining private vehicle use for JTW.

Figure 3 shows the results for the analysis of public transport supply.
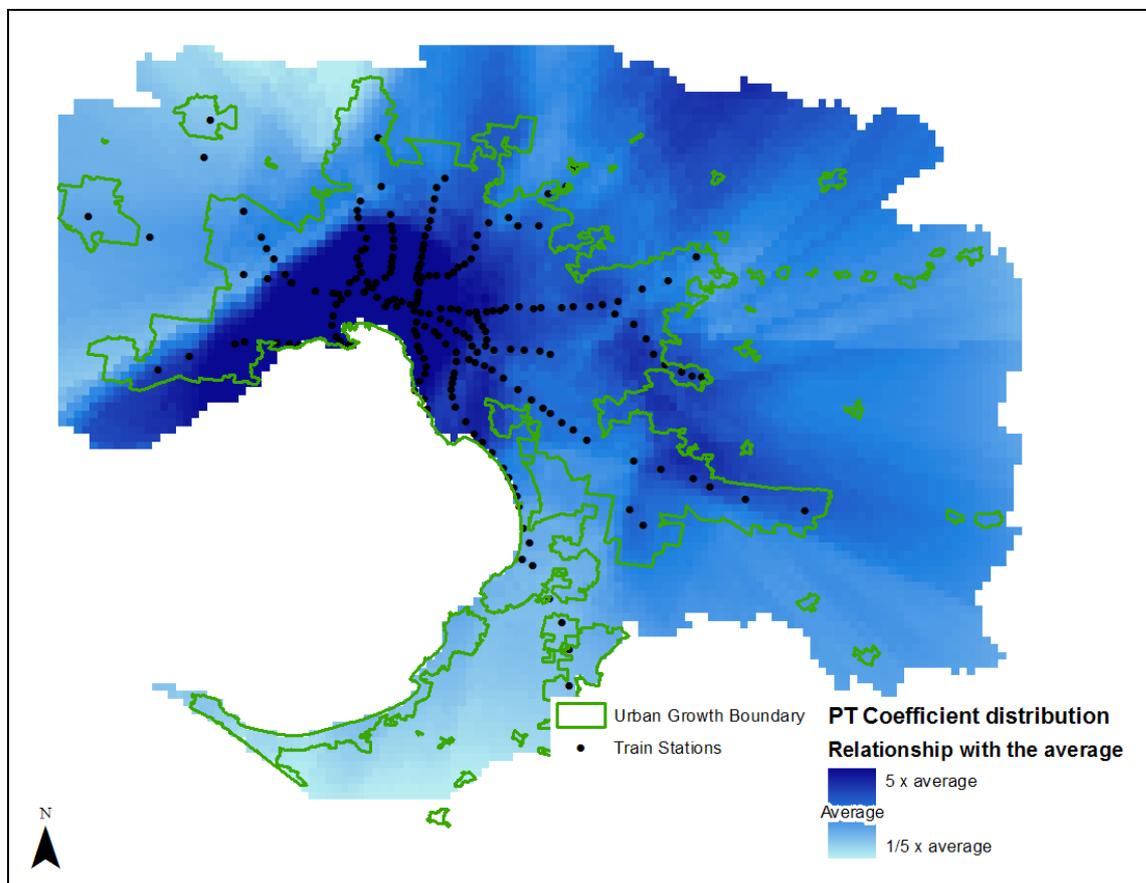
Figure 3: Spatial Strength of Links between Public Transport Supply and Car JTW in Melbourne

The supply of public transport has the strongest link with private vehicle use for JTW in the analysis. The spatial analysis in Figure 1 suggests that this is particularly important in the inner areas (10km from CBD) and also along a southwest corridor. For the eastern and southern sections of Melbourne the importance of public transport supply is less critical as a means of explaining private vehicle use.
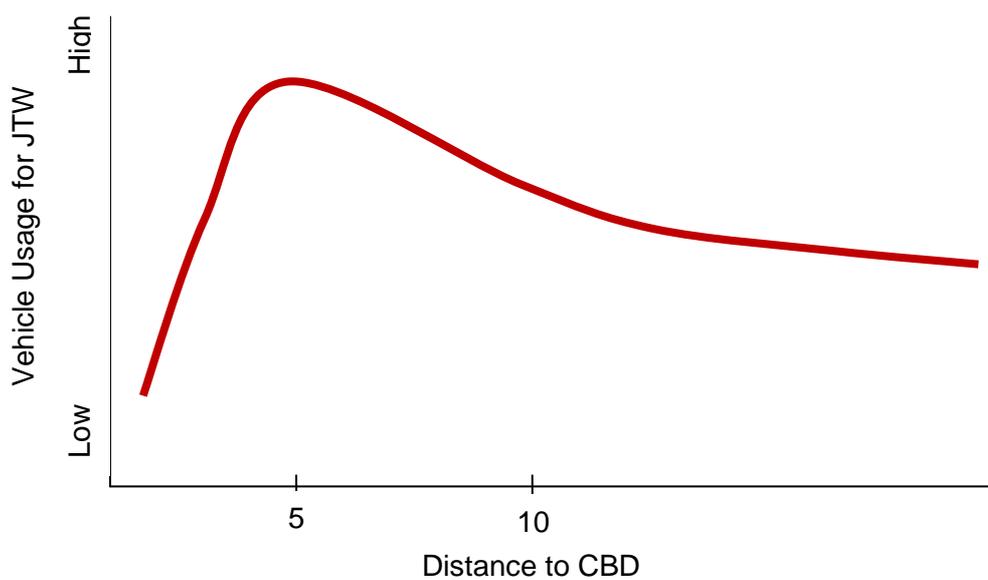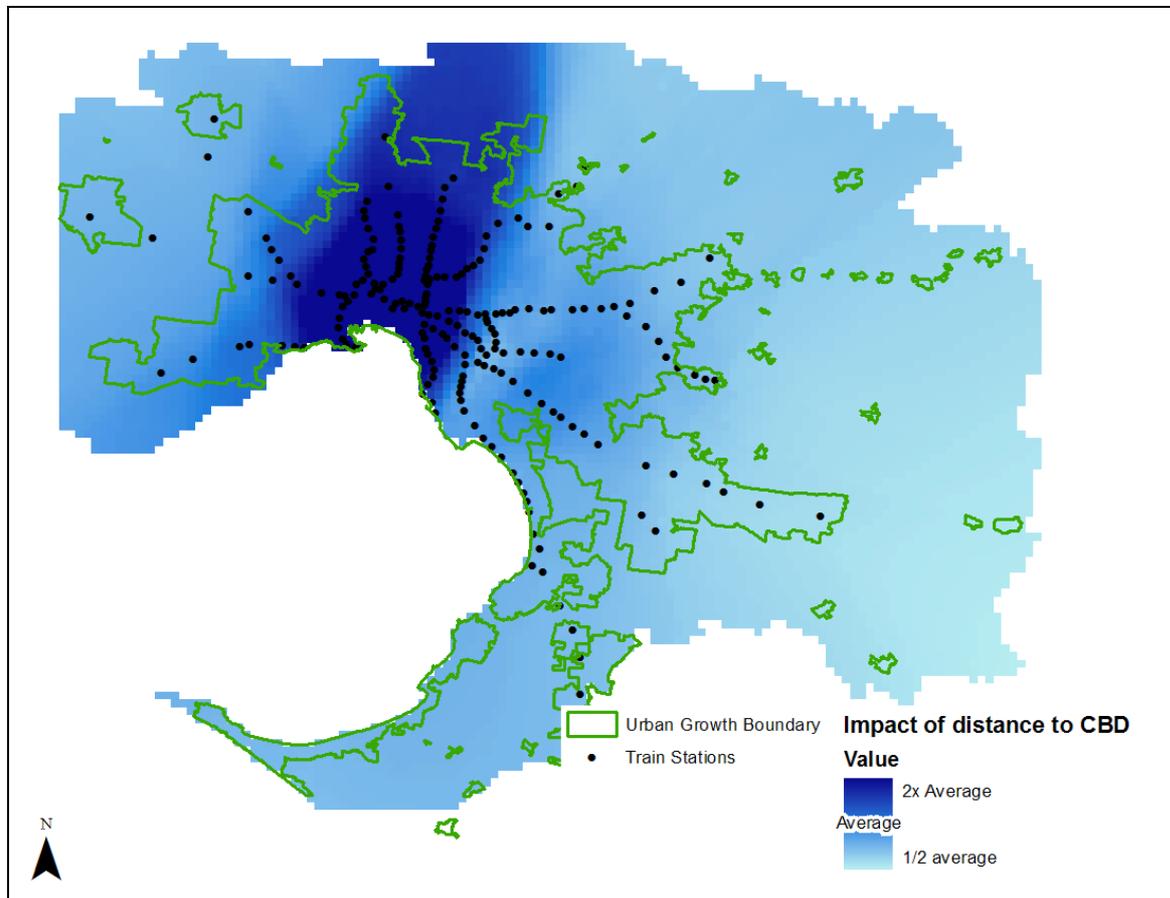

Figure 4: Spatial Strength of Links between Distance from the CBD and Car JTW in Melbourne

Figure 4 shows the spatial strength of the variable distance to the CBD and its link to private vehicle usage for JTW. This suggests that CBD proximity is very important in the very inner areas but decreases from around 10 to 15 km from the City. Then the importance of this variable appears again in outer areas (30-40km) particularly in the north. In the main growth areas the impact of the variable appears as low.

Figure 5 shows the spatial strength of the variable residential density and how it acts to explain private vehicle use for JTW.
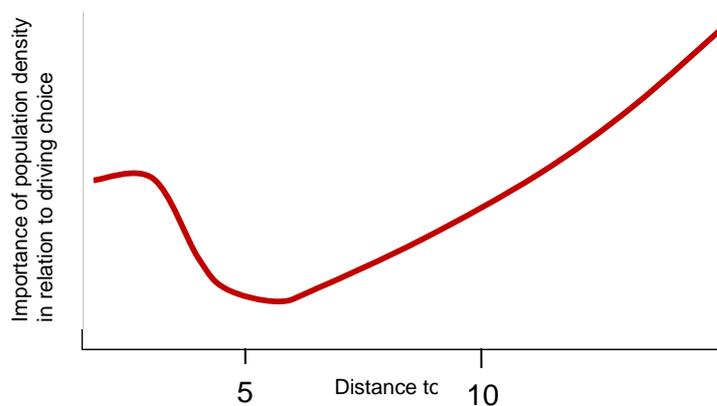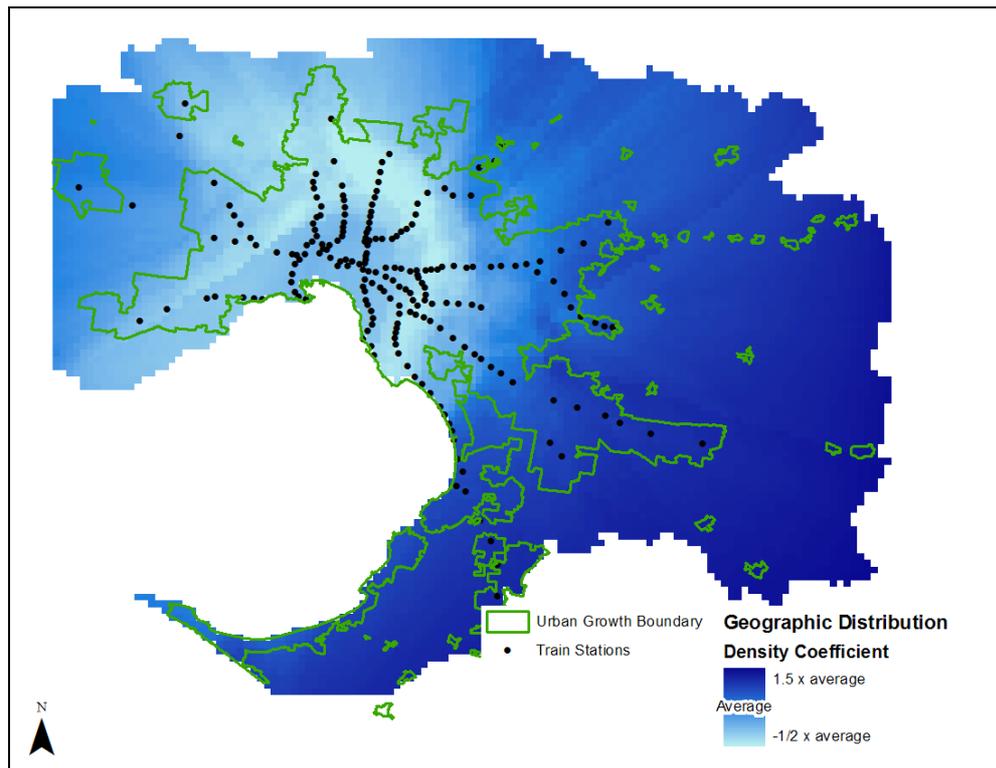

Figure 5: Spatial Strength of Links between Residential Density and Car JTW in Melbourne

Population density power in explaining private vehicle JTW appears low in inner areas. This reduces further in areas that area around 15 to 20 kilometres from the CBD. In outer and growth areas, density appears a more significant factor.

5 Discussion and Conclusions

This paper aims to explore factors affecting private car use for journey to work in Melbourne using a multiple regression analysis and an analysis of the spatial strength of explanatory variables using a geographically weighted regression analysis.

The multiple regression analysis explained 57% of the variance of car use for travel to work in Melbourne. Three variable explained most of the variation in auto use including level of public transport supply (-0.443), distance from the CBD (0.404) and residential density (-0.136).

GWR increased the explanatory power of the analysis from 57% to 74% of variance and enabled spatial patterns of the three major explanatory variables to be explored. Public transport supply was found to be particularly strong at influencing private vehicle commuting in inner and south western parts of Melbourne. Distance to the city was stronger in inner and northern parts of Melbourne. Residential density was strongest in eastern and outer areas.

In this research we found that GRW provided us with an alternative to improve standard regression models and identify key factors affecting journey to work for Metropolitan Melbourne. In this process, and due to the fact that travel choice has clear spatial patterns, GWR was very appropriate to not only make a standard model viable but to understand the variability of factors in the space.

The methodology used appears a good complement to existing transport analysis as it focuses attention in objectives (e.g. reduced car dependency) rather than in network capacity or operational factors.

The analysis conducted only covered JTW and, therefore, factors identified cannot be fully assumed by all trips. In any case, and considering the consistency of our results with other studies that analysed many modes, it appears that in Melbourne the provision of public transport is the most important factor over any other land use or transport driver.

There is the possibility for the methodology developed to be advanced to allow prediction of travel patterns in future areas to be developed. For this, data input would have to be expanded to include all trips and more detailed spatial calculations would be required to represent better factors such as access to cycling, walking and the vehicular network.

Further development of a statistical model for Melbourne could potentially allow GWR results to be used to not only identify factors more accurately for all trips but support standard four steps transport models in the prediction of behaviours in new areas based on land use and transport plans.

6 References

Abelson, P. W. and S. Baker (1982). <u>Household Car Ownership and Mode Choice for the Journey to Work</u>. Australasian Transport Research Forum, Hobart.

Australian Bureau of Statistics (2006). 2006 Census Tables.

Booz & Co. (2010). Werribee Plains - Reducing Car Dependency: Transport and Urban Design Solutions. Melbourne, Australia, for the Australian Conservation Foundation.

Bureau of Transport and Regional Economics (2007). Estimating urban traffic and congestion cost trends for Australian cities - Working Paper 71. Canberra, Australia, Bureau of Transport and Regional Economics.

Charlton, M. E. and A. S. Fotherringham (2009a). Geographically Weighted Regression - A Tutorial on using GWR in ArcGIS 9.3, National Centre for Geocomputation, National University of Ireland Maynooth Maynooth, Co Kildare, IRELAND March 3 2009.

Charlton, M. E. and A. S. Fotherringham (2009b). Geographically Weighted Regression, White Paper National Centre for Geocomputation, National University of Ireland Maynooth Maynooth, Co Kildare, IRELAND March 3 2009.

Chatman, D. (2003). "How Density and Mixed Uses at the Workplace Affect Personal Commercial Travel and Commute Mode Choice." <u>Transportation Research Record</u> **Vol 1831**: pp 193-201.

Competition and Regulation Working Group (2006). Review of urban congestion trends, impacts and solutions, Council of Australian Governments.

Crane, R. and R. Crepeau (1998). "Does Neighbourhood Design Influence Travel? A Behavioural Analysis of Travel Diary and GIS Data " <u>Transportation Research Part D</u> **Vol 3**: pp 225-238.

Currie, G. (2010). "Quantifying Spatial Gaps in Public Transport Supply Based on Social Needs." <u>Journal of Transport Geography</u> **Volume 18 (2010)** pp 31–41

Fotheringham, A. S., C. Brunsdon and M. E. Charlton (2002). <u>Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, </u>. Chichester, UK, Wiley.

Geurs, K., B. van Wee and P. Rietveld (2006). "Accessibility appraisal and integrated land use- transport strategies: methodology and case study for te Netherlands Reinstadt area." <u>Environment and Planning B</u> **Vol 33**: pp 639-660.

Longworth, T. and C. Wilson (2006). <u>An Exploration of Mode Choice in Sydney Using Journey to Work Data</u>. 28th Australasian Transport Research Forum, CHECK LOCATION AND DATE.

Mulley, C. and M. Tanner (2009). <u>The Vehicle Kilometres ravelled (VKT) by Private Car: A Spatially Analysis Using Geographically Weighted Regression</u>. Australasian Transport Research Forum, Auckland New Zealand.

Rickwood, P. and G. Glazebrook (2009). "Urban Structure and Commuting in Australian Cities." <u>Urban Policy and Research</u> **Vol 00 No 0**: pp 1-18.

Shiftan, Y. and Y. Barlach (2002). "Effect on Employment Site Characteristics on Commute Mode Choice." <u>Transportation Research Record</u> **Vol 1781**: pp 19-25.

Victorian Competition & Efficiency Commission (2006). "Making the right choices: Options for managing transport congestion - Final Report."

Zhou, J., J. Gould and S. O'Flaherty (2007). <u>Commuter Travel: How Does Proximity Influence Mode Choice? GIS Analysis of a Large Urban University</u>. Transportation esearch Board Annual Meeting, Washginton DC, USA.