Transportation Research Part C 86 (2018) 147–167

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

To travel or not to travel: 'Weather' is the question. Modelling the effect of local weather conditions on bus ridership

Sui Tao[a,*], Jonathan Corcoran[b], Francisco Rowe[c], Mark Hickman[d]

[a] Institute of Future Cities, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
[b] School of Earth and Environmental Sciences, The University of Queensland, 4072, Australia
[c] Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool L69 7ZT, UK
[d] School of Civil Engineering, The University Queensland, 4072, Australia

ARTICLE INFO

Keywords:
Public transport
Weather
Time-series modelling
Travel behaviour

ABSTRACT

While the influence of weather on public transport performance and ridership has been the topic for some research, the real-time response of transit usage to variations in weather conditions is yet to be fully understood. This paper redresses this gap by modelling the effect that local weather conditions exert on hourly bus ridership in sub-tropical Brisbane, Australia. Drawing on a transit smart card data set and detailed weather measurements, a suite of time-series regression models are computed to capture the concurrent and lagged effects that weather conditions exert on bus ridership. Our findings highlight that changes in particularly temperature and rainfall were found to induce significant hour-to-hour changes in bus ridership, with such effects varying markedly across both a 24 h period and the transit network. These results are important for public transport service operations in their capacity to inform timely responses to real-time changes in passengers' travel demand induced by the onset of particular weather conditions.

1. Introduction

Public transport plays an essential role in maintaining civic and economic activities by providing a mass and sustainable mobility option for urban populations (Schwanen, 2002; Vuchic, 2005; Cervero, 1998). As such, public transport services by and large need to operate in a manner that passengers' travel needs are adequately met ranging from everyday commuting to less routinised, more spontaneous trips. In this regard, the role weather plays in influencing the level of public transport service and ridership has been highlighted as an important issue in transport scholarship (Guo et al., 2007; Böcker et al., 2013). Inclement and extreme weather conditions (e.g., heavy precipitation, low temperatures and strong winds) are known to have the capacity to degrade service quality (e.g., disrupting service schedule) and passenger experience (e.g., prolonged waiting and travel times), with the potential to induce temporary as well as long term declines in ridership (Hofmann and O'Mahony, 2005; Changnon, 1996; Hine and Scott, 2000). As such it is important for us to consider the way in which weather impacts the everyday operation of public transport systems such that its negative effects and potential loss in ridership can be ameliorated. To achieve this, the effects that weather impose on public transport ridership first need to be understood to provide the necessary evidence from which planning and operation strategies can be founded (Guo et al., 2007; Böcker et al., 2013).

Given the need to understand the effects of weather on public transport and its end users, a growing number of recent studies have sought to examine the relationship between weather and ridership, e.g., Changnon (1996), Hofmann and O'Mahony (2005), Guo

* Corresponding author.
E-mail addresses: s.tao@cuhk.edu.hk (S. Tao), jj.corcoran@uq.edu.au (J. Corcoran), F.Rowe-Gonzalez@liverpool.ac.uk (F. Rowe), m.hickman1@uq.edu.au (M. Hickman).

https://doi.org/10.1016/j.trc.2017.11.005
Received 20 April 2017; Received in revised form 30 October 2017; Accepted 5 November 2017

et al. (2007), Kalkstein et al. (2009), Arana et al. (2014), Singhal et al. (2014). Their findings highlight that public transport (e.g., bus and rail transit) ridership is negatively influenced by heavy precipitation, and to a lesser extent, high temperatures, strong winds and high levels of humidity (Böcker et al., 2013; Koetse and Rietveld, 2009). In addition, the effects of weather on ridership have also been shown to significantly vary across different calendar events and transit modes. For example, in a Chicago-based study, Guo and colleagues (2007) found that changes in weather conditions exerted greater impact on metro and bus transit ridership during weekends than weekdays; and metro ridership was less affected by weather than bus ridership, possibly due to the experience of riding a bus is more exposed and vulnerable to inclement weather. Similar findings were reported by Cravo et al. (2009) and Kashfi et al. (2013) in studies of New York City and Brisbane (Australia) respectively. Focusing on two cities in Canada, Trépanier et al. (2012) found that weather had stronger effects on senior passengers; and adverse weather might drive a modal shift from bus to rail transit among public transport passengers. Finally, Singhal et al. (2014) explored hourly relationships between weather and ridership for the New York metro system on weekdays and weekends. Their study revealed that a number of weather variables including the presence of rain, snow and strong winds to be negatively associated with the metro ridership especially during weekends.

While not exclusively focusing on public transport, some other transport studies have also shed light on the impact of weather on people's public transport use under the broader umbrella of travel behaviour. In two linked studies, drawing on the Swedish National Travel Survey data, Liu et al. (2015, 2016) modelled the impacts of weather conditions (in particular, temperature, precipitation and a measure of thermal comfort) on modal choice and trip-chaining behaviour along with a suite of other contextual factors (e.g., household size, income, car ownership and population density). After accounting for the influences of contextual factors, their findings indicate that weather conditions, particularly precipitation exerted significant effects on public transport use, which were shown to vary significantly across different seasons and locations across Sweden. For example, heavy rain was found to discourage bus use in the northern Sweden during summer, autumn and winter, whilst the reverse was shown to be the case for southern locales in the country (Liu et al., 2015). Such findings imply the existence of seasonally varying and localised perceptions of people towards weather, which contribute to a variety of behavioural responses (e.g., whether to take public transport or not) to weather of people across Sweden. In another related study, Creemers et al. (2015) investigated the relationships between modal choice behaviour and an array of hourly as well as lagged weather variables (e.g., temperature, fog, precipitation and a measure of thermal comfort) in the Netherlands, wherein only a thermal comfort, namely, physiologically equivalent temperature (PET), was found to have a significantly negative effect on bus usage.

Despite the accumulating evidence of weather's effect on public transport ridership, their relationships have arguably yet to be fully understood. In particular two research gaps can be identified. First, close scrutiny of the current transport literature reveals that real-time relationships between weather and public transport ridership has seen scant scholarly attention. A commonly adopted approach to investigate the weather-ridership relationship has been the use of daily averages of weather measurements (e.g., the mean daily temperature and main daily rainfall) as exogenous variables on which system-wide daily ridership is modelled, see for example studies by Guo et al.(2007), Kalkstein et al. (2009), Arana et al. (2014). While adopting this analytic strategy is able to establish certain weather-transit associations usually at the system-wide level, this daily-based approach is not able to fully capture the concurrent response of ridership to changes in weather conditions. As weather is known to have the potential to be highly variable over relatively short periods of time (Ephrath et al., 1996; Mapes et al., 2003), the resulting impact on transit ridership may vary accordingly. Only a limited number of studies including Singhal et al. (2014) and Creemers et al. (2015) have begun to examine the real-time impact of weather on transit ridership (e.g., hourly variations in bus ridership). How weather conditions are known to affect people's daily travel behaviour, however, remains to be addressed by transport scholars (Creemers et al., 2015). Furthermore, the lack of evidence on the real-time weather-ridership relationship at finer temporal scales arguably limits their utility for public transit operators in terms of how the findings can be translated to adjustments to service and account for weather induced variations in transit ridership.

Second, although most existing studies that have examined the weather-transit relationship have focused on its temporal variability, the geographic dimension remains largely unexplored. Studies by Liu et al. (2015, 2016) are among the few exceptions that have examined the spatial heterogeneity of weather's influence on travel behaviour. However, these studies both adopted a relatively coarse spatial scale (municipalities) with the effect of limiting the capacity to reveal intra-metropolitan weather-transit ridership patterns. Tao et al. (2016), on the other hand, adopted a suite of geo-visual techniques to unveil spatially varying patterns of bus usage across Brisbane, Australia. This scarcity of the evidence on the spatial variation of weather-travel relationship is despite a growing body of research that shows that people tend to exhibit systematically varying trip-making patterns according to trip distance, frequency and modal choice across urban areas, associated with different physical (e.g., density and design) and social (e.g., income level and household type) structures (Bagley and Mokhtarian, 2002; Wang and Khattak, 2013; Morency et al., 2011). Furthermore, urban spaces are comprised of a mosaic of spatially segregated locations, each with distinct functionalities (e.g., office, commercial and education) and activities with particular patterns (e.g., routinized non-discretionary versus recreational and discretionary activities) (Chapin, 1974; Handy et al., 2002; Ibrahim, 2003). Given what we know of both individual travel behaviours and urban form, it is likely that public transport passengers travelling from and bound for different localities across a city may also exhibit collectively different levels of vulnerability to changing weather conditions. Hence, there is a need for transport scholars to begin to understand the geographic dynamics of trip patterns in particular origins and destinations (Tao et al., 2016; Liu et al., 2016). Furthermore, revealing weather-transit ridership at finer temporal and spatial scales will also provide a necessary evidence base that allows transit operators to impose proactive adjustment in scheduling and resourcing a transit network especially when adverse weather conditions hit. This, in turn, has the potential to enhance transit users' travel experience as well as achieve enhanced performance of transit operators.

This study aims to address these research gaps through a spatio-temporal examination of the weather-transit relationship. To this end, we draw on a three-month smart card data set of bus ridership allied with detailed weather measurements to form an integrated database. Using an autoregressive integrated moving average (ARIMA) with explanatory variables (i.e. ARIMAX) modelling framework and its seasonal extension (i.e. SARIMAX) as our core set of analytical tools, a suite of regression models were estimated to capture hourly relationships between bus ridership and four weather variables, temperature, rainfall, wind and humidity at three different spatial scales: system-wide city level, destination-based and stop level. In transport research, a mounting number of studies have employed ARIMAX and its extensions to model and forecast road traffic demand over time given their ability to handle time-series data. Drawing on seasonal ARIMA (or SARIMA) model, for example, Williams and Hoel (2003) modelled the weekly seasonality of daily traffic counts, particularly vehicular patterns. Williams (2001) utilised ARIMAX model to forecast motorway downstream counts in relation to counter-upstream counts recorded at 30-min intervals. However, to the best of our knowledge, no prior study has employed time-series modelling methods to investigate the weather-transit relationship at relatively fine temporal scales, such as hourly. This study offers a first empirical attempt by addressing three key questions:

(1) To what extent do hourly changes in weather conditions affect bus ridership, and how does this effect vary over weekdays and weekends?
(2) To what extent does the impact of hourly changes in weather conditions on bus ridership vary across destinations?
(3) To what extent do hourly changes in weather conditions affect spatial-temporal patterns of bus ridership across the bus network?

The rest of the paper is structured as follows: the next section presents the study context and data employed. Section 3 presents the methodology for modelling the weather-transit relationship at different spatial scales, before reporting the results in Section 4. Section 5 discusses our findings and presents directions for future research before making some concluding remarks.

2. Study context and data

2.1. Study context

The bus network in Brisbane, Australia is the study context (Fig. 1). Brisbane is the capital of Queensland and the third most populated city in the country, with around one million population within its local government area (ABS, 2013). Within Brisbane, private cars are the primary trip-making mode, accounting for approximately 85% of all daily trips, followed by public transport (8%)

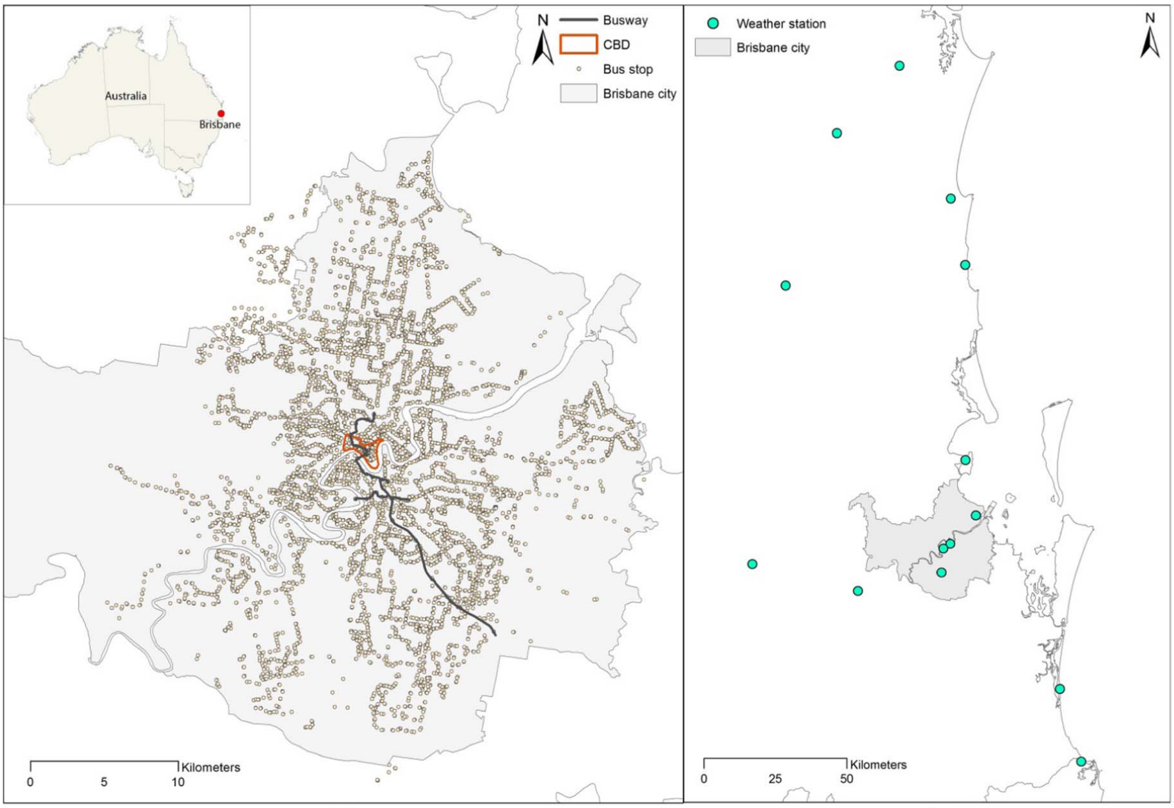

Fig. 1. Study context.

and active transport (7%) (BITRE, 2014a). Despite a strong car-oriented culture, since the new millennium Brisbane's local government has initiated a series of projects (including introducing a new exclusive busway and a series of programs of transit stop upgrading) in order to promote transit usage (Mees and Dodson, 2011; Hoffman, 2008). Currently Brisbane's bus network comprises over 400 bus routes serving for both central and outer suburban locales across the city (Tao et al., 2014). Fig. 1a indicates the locations of bus stops and the busway, while Fig. 1b indicates the weather stations surrounding Brisbane.

The investment by the local government in buses is among the major contributory factors to its notable usage growth over the period between 2004 and 2013, from around 0.7 billion to 1.27 billion passenger-kilometres travelled (BITRE, 2014a). According to recent government reports, Brisbane's bus network has also absorbed more passenger trips than alternative transit modes, with bus transport accounting for over 90 million journeys compared to around 50 million and 5 million journeys captured by rail and ferry respectively in 2011 (BITRE, 2013), and accounted for over 54% of all passenger-kilometres travelled in 2013 (BITRE, 2014a). In addition, given the projected increase of transit usage over the next 15 years (BITRE, 2014b), it is likely that bus transit will assume a more important role in fulfilling the everyday travel needs of Brisbane's population. Given this and the evident vulnerability of bus passengers' travel experience to weather, such as accessing to and waiting at bus stops during rain (Hofmann and O'Mahony, 2005), Brisbane's bus network provides an interesting context for this study.

Brisbane has a subtropical climate. Statistics over the past two decades show that its summers (between December and February) are hot (average monthly maximum temperature close to 30 °C) and wet (mean rainfall over 130 mm per month), whilst winters (between June and August) are mild (average monthly maximum temperature over 20 °C) and dryer (monthly levels rainfall ranging between a quarter to a half of their summer counterpart) (Bureau of Meteorology, 2015). The remainder of the months are warm (average monthly maximum temperature between 23 and 25 °C) with no marked variations in wind throughout the year.

2.2. Data sources

To address our three research questions, transit smart card data and weather measurements are employed as the two principal data sources. A smart card data set covering a three-month period from 4th February to 28th April 2013 was provided by Translink (Brisbane's transit agency) in the form of transaction records that are generated every time a passenger touches on and off public transport. The information contained in a single smart card record includes date, route, direction (i.e., inbound and outbound trips in relation to the city centre), smart card ID, boarding time and stop, and alighting time and stop and journey ID for linked trips made within a one-hour transfer limit.

Weather data were acquired from the Australian Bureau of Meteorology (BOM) for the same period of time as the smart card data. Measurements of four weather variables, i.e., temperature, rainfall, relatively humidity and wind speed on a 30-min interval are included for 14 weather stations located across the study context. Fig. 2 depicts the average hourly patterns of the four weather variables of the 14 stations (refer to Fig. 1b for their locations). The weather conditions captured largely reflect the subtropical climate of Brisbane. Temperature (mean = 23 °C) and wind speed (mean = 3.76 m/s) were relatively stable with small variations over the sampled days. Relative humidity (mean = 73%) saw more noticeable changes, yet remained mostly above the level of 60%. The variable exhibiting the most variations is rainfall, which remained below 1 mm for most hours and entered a relatively intense wet period through mid-February and early March. In addition to the above weather variables, apparent temperature was calculated to capture the collective influence of temperature, wind and relative humidity given their relations to people's subjective heat stress (Bureau of Meteorology, 2010). An Empirical Bayesian Kriging (EBK) tool in ArcGIS was employed to estimate hourly weather conditions for each of the bus stops given its ability to produce robust estimations using relatively sparse spatial data, such as

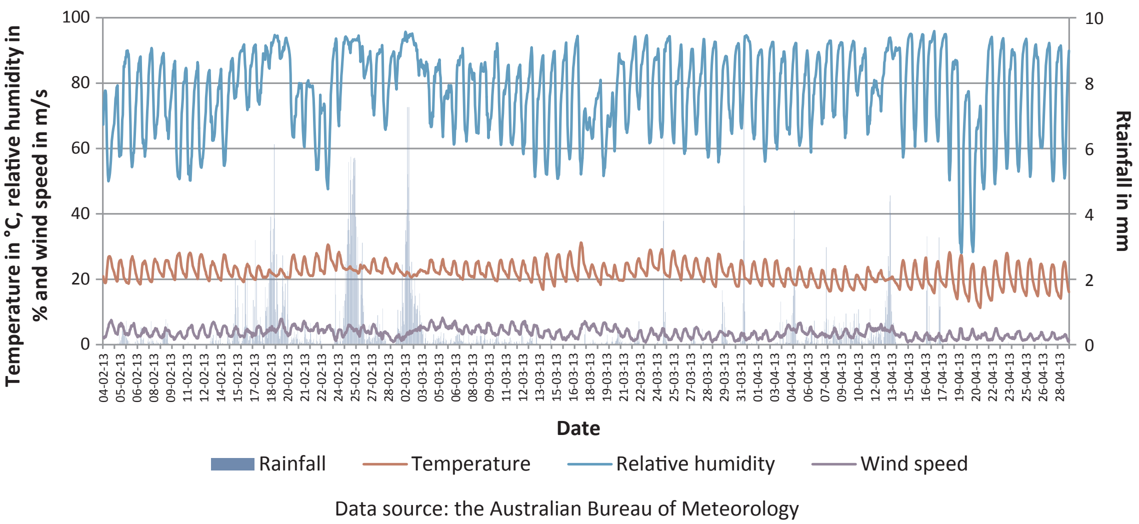

Fig. 2. Hourly weather conditions.

estimating continuous rainfall levels across an urban area drawing on data from several weather stations (Krivoruchko, 2012). General Transit Feed Specification (GTFS) (Google Developers, 2012) was used to obtain detailed geographic information on the bus network, including data on the longitude and latitude of all bus stops; and the Queensland's state calendar (DETA, 2013) was used to link calendar events (e.g., weekdays, weekends, public and school holidays) to smart card and weather data.

3. Methodology

We modelled the hourly weather-ridership relationships at three different spatial scales: (1) system-wide; (2) destination-based and (3) stop level. This section first introduces the class of time-series modelling techniques employed for each of the spatial scales: ARIMAX and its seasonal variant, SARIMAX. Next we describe the analysis strategies for the destination-based and stop-level investigations.

3.1. ARIMAX and SARIMAX models

Previous studies modelling the weather-transit relationship have employed multiple linear regression modelling, e.g., Cravo et al. (2009), Kalkstein et al. (2009), Arana et al. (2014). This approach treats temporally continuous transit ridership as independent incidents and in doing so overlooks the existence of temporal self-dependency. Hourly ridership patterns for a day tend to be associated with hourly ridership of the previous day. Failing to take account of such self-dependency, or temporal autocorrelation is likely to generate biased estimations concerning the effects of weather on transit ridership. Given that we know that public transport ridership normally exhibit systematically recurring temporal patterns in accordance with different times of day (e.g., peak and non-peak hours) and days of week (e.g., weekday and weekends), it is necessary to employ a modelling technique that has the capacity to: minimise autocorrelation while effectively capturing the effects of changing weather conditions on bus ridership.

To meet these requirements, time-series modelling techniques, specifically ARIMAX and SARIAMX models were estimated. To introduce these modelling approaches, basic understanding of some key concepts is first required (see Brockwell and Davis (2002) and Box et al. (2008) for a detailed description). The ARIMAX and SARIAMX models are derived from the ARIMA model (Box and Jenkins, 1970). An ARIMA model encompasses three components, namely an autoregressive (AR) process, a moving average (MA) and an integrated (I) element. The AR and MA parameters control for temporal autocorrelation in a time series resulting from two mechanisms. The first assumes a variable (Y) at time t (Y_t) is explained by its value(s) at previous time point(s) (e.g., Y_{t-1}, $Y_{t-2}\ldots Y_{t-p}$). The second posits Y_t is the function of current and previous moving averages of error terms (e.g., u_{t-1}, $u_{t-2}\ldots u_{t-q}$) (Brockwell and Davis, 2002).

Fitting a time series in a model that contains AR and MA parameters (or an ARMA model) requires the data to be weakly stationary, which is characterised by: (1) constant mean and variance of Y_t over time, and (2) the covariance of Y_t to be time-invariant; that is, it is assumed to only be dependent on the lag between the current and past value and not the actual time at which the covariance is computed (Brockwell and Davis, 2002). However, few time series are weakly stationarity. They have an integrated (I) time series; that is, they have to be differentiated to make them stationarity. An ARIMA model takes the general form of:

$$(1-\phi_1 B^1-\phi_2 B^2-\ldots-\phi_p B^p)(1-B)^d Y_t = (1-\theta_1 B^1-\theta_2 B^2-\ldots-\theta_q B^q)u_t \tag{1}$$

wherein:

ϕ is the autoregressive parameter (e.g., $\phi_1 Y_{t-1}$); θ is the moving average parameter (e.g., $\theta_1 Y_{t-1}$); B is the backshift operator defined by $B^i(Y_t) = Y_{t-i}$; d is the order of differencing (e.g., d_1 indicates $Y_t - Y_{t-1}$); and u_t is the error term.

When a time series exhibits seasonal recurring patterns (e.g., a daily traffic count), an ARIMA model can be expanded to a seasonal ARIMA (or SARIMA) model by adding a differencing operator, AR and/or MA terms at a seasonal lag(s):

$$Y_t = \frac{(1-\theta_1 B^1-\theta_2 B^2-\ldots-\theta_q B^q)(1-\Theta_1 B^{1s}-\Theta_2 B^{2s}-\ldots-\Theta_Q B^{Qs})}{(1-\phi_1 B^1-\phi_2 B^2-\ldots-\phi_p B^p)(1-\Phi_1 B^{1s}-\Phi_2 B^{2s}-\ldots-\Phi_P B^{Ps})(1-B)^d(1-B^s)^D}u_t \tag{2}$$

wherein:

D is the order of seasonal differencing (e.g., for a 24-h seasonal period, D_1 indicates $Y_t - Y_{t-24}$); and Φ is the seasonal autoregressive parameter (e.g., for a 24-h seasonal period, $\Phi_1 Y_{t-24}$); and Θ is seasonal moving average parameter (e.g., for a 24-h seasonal period, $\Theta_1 Y_{t-24}$).

The ARIMA and SARIMA models were developed specifically to model and forecast univariate time series. Given that our study aims to investigate the effects of exogenous variables on a time series, ARIMAX (or SARIMAX) models were developed. These models incorporate the time-series components of an ARIMA (or SARIMA) process into a multiple regression model as follows:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots+\beta_k X_{k,t} + N_t \tag{3}$$

wherein:

Y_t is the time-series dependent variable; $X_{1,t}$ to $X_{k,t}$ are the explanatory variables; and β_0 to β_k are the corresponding regression coefficients; and N_t is the error term, which is next expanded into the following expression:

$$N_t = \frac{(1-\theta_1 B^1-\theta_2 B^2-\ldots-\theta_q B^q)(1-\Theta_1 B^{1s}-\Theta_2 B^{2s}-\ldots-\Theta_Q B^{Qs})}{(1-\phi_1 B^1-\phi_2 B^2-\ldots-\phi_p B^p)(1-\Phi_1 B^{1s}-\Phi_2 B^{2s}-\ldots-\Phi_P B^{Ps})(1-B)^d(1-B^s)^D}e_t \tag{4}$$

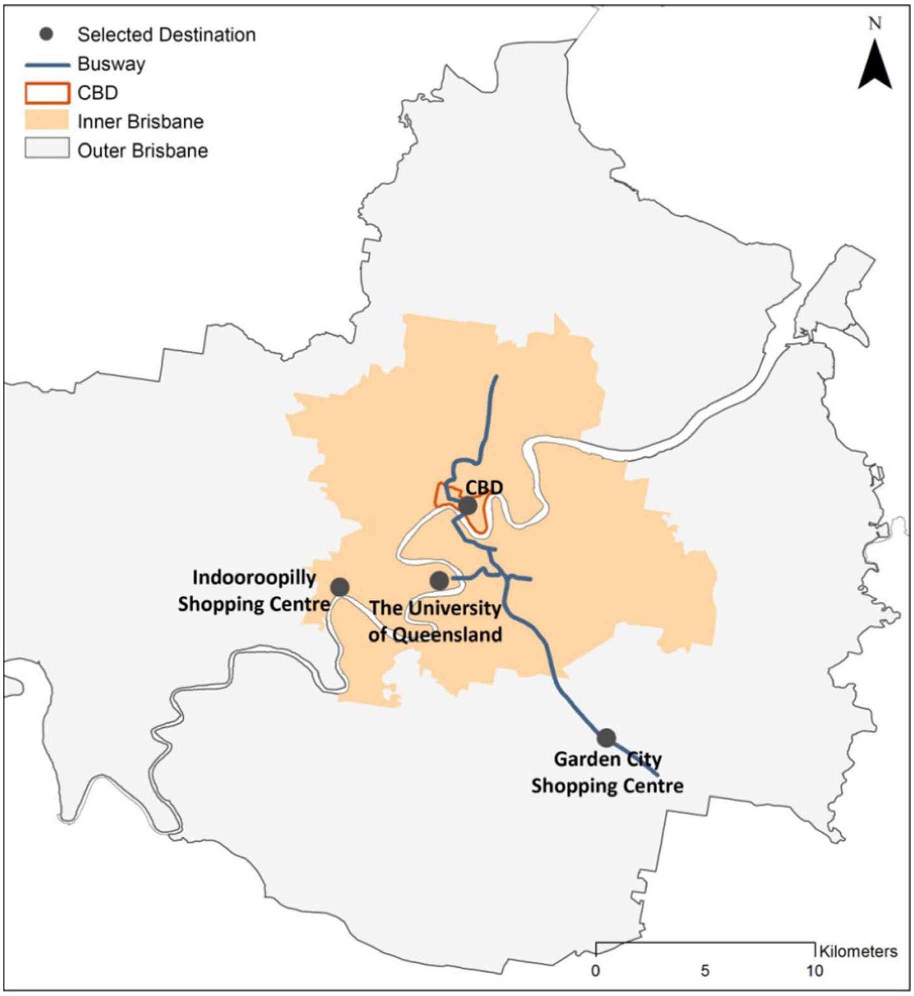

Fig. 3. Major destinations.

where e is the white noise error.

Note that Expression (4) encompasses the AR, MA parameters and the differencing operator components, all of which are central to a SARIMA model in controlling for the self-dependency in time series analysis. Replacing N_t in Eq. (3) with Eq. (4) gives us the SARIMAX (or ARIMAX with the seasonal autoregressive and moving average terms) model. Given their ability to control for seasonal and non-seasonal autocorrelations, ARIMAX and SARIMAX models are well placed to model the effects of weather on hourly transit ridership. The ARIMAX and SARIMAX models introduced above serve as the main analytical tools for system-wide analysis. Following previous studies, e.g., Cools et al. (2009), Chen and Tjandra (2014), Biswas et al. (2014), we carry out the modelling analysis in two steps: (1) fitting hourly bus ridership to univariate ARIMA (or SARIMA) models; and (2) adding weather variables to the fitted ARIMA (or SARIMA) models to investigate their effects on hourly bus ridership. In addition, given we know that weekdays and weekends are typically associated with distinct differences in activities and trip-making patterns, e.g., the former usually involves more non-discretionary trips, as such commuting trips, than the latter, we estimate separate models to capture the effects of changing weather conditions on hourly bus ridership on weekdays and weekends. In addition to the concurrent effects of weather variables, we investigated their lagged effects on hourly ridership by including a variable of weather conditions of the previous hour. A previous study found that hourly bicycle use in Montreal not only varied with concurrent weather conditions, but also weather conditions three hours earlier, suggesting a lagged response of travel patterns to weather (Miranda-Moreno and Nosal, 2011). However, no study to our knowledge has tested the lagged effects of weather within the public transport context.

3.2. Analytic strategy for destination-based analysis

While system-wide ridership patterns are analysed using SARIMAX and ARIMAX models, these modelling approaches are coupled with a kernel density analysis of bus stops to perform a destination-based analysis. To this end, we first selected major destinations (Fig. 3) stratified by the functionality of the locale to model the effects of weather on their hourly ridership (i.e., number of journeys

Table 1
Summary of selected destinations.

Destination	Number of bus stops	Total weekday ridership	Total weekend ridership	Main function
CBD	44	2,625,989	403,829	Office, recreation, retailing
The University of Queensland	11	594,642	29,972	Tertiary education, recreation
Garden City Shopping Centre	15	302,032	45,039	Retailing, recreation
Indooroopilly Shopping Centre	13	165,158	30,931	Retailing, recreation
Total	5970	12,190,494	1,623,526	

bound for each of these destinations). The destination selection process involved the following two steps:

1. First, we calculated the ridership (i.e., number of passengers alighting without further transferring) for all individual bus stops (approximately 6000 in total) over the 3-month period. Using the ridership of bus stops as weights, a kernel density analysis was then conducted to identify hot spots bus stops destinations. The hot spots locations were compared with land use census data that identified four types of activity-intense destinations; (1) the CBD (2) university and (3) shopping centre near the busway and (4) other suburban shopping centre.
2. Next, we calculated the total ridership for each of the major destination from the previous step by summing the ridership of individual bus stops that were located within each cluster. Four destinations with the highest total ridership in each destination category, i.e., the CBD, including its south adjacent areas, one university (the University of Queensland or UQ) and two suburban shopping centres, were selected as our focused destinations. The selection of two shopping centres, the Indooroopilly Shopping Centre and Garden City Shopping Centre, was used to capture different infrastructure settings. The latter is located near the exclusive busway (see Fig. 1) wherein shelters are provided at the busway stations (Currie and Delbosc, 2010), whereas the former is mainly served by ordinary on-road bus services and thus bus ridership involving this destination is expected to be more affected by changes in weather conditions.

Including a total of 83 bus stops, the four activity-intense destinations collectively account for approximately 30 percent of the overall ridership across the network on both weekdays and weekends over the three months (Table 1). For each of the four destinations, hourly ridership was calculated for the destination-based analysis.

3.3. Analytic strategy for stop-level analysis

To enable stop-level analysis, while a geographically weighted ARIMAX model is desirable, to the best of our knowledge, such a tool does not yet exist. Given the large number of bus stops, it would also be infeasible to configure and run individual ARIMAX (or SARIMAX) models for each stop in the network. We developed a methodology able to model the hourly weather-ridership relationships, while capturing autocorrelation and periodicity in ridership for individual bus stops in a logical and computationally manageable manner. The developed methodology involves the following three steps:

1. First, we reduced the number of bus stops to several groups characterised by distinguishable ridership, and to some extent, activity patterns. This was achieved by performing a cluster analysis of bus ridership using three groups of indicators based on the standardised average ridership (number of passengers boarding): (1) in different hours (00:00 to 23:00), (2) days (Monday to Friday, Saturday and Sunday), and (3) weeks (1st week to the 12th week) for weekdays and weekends. To obtain more robust cluster solutions, a hierarchical cluster analysis was first carried out to identify initial cluster centres and numbers, which were then used as input for a k-mean cluster analysis.
2. From the cluster analysis, we found that dividing the bus stops into 2–8 clusters were solutions that captured the majority of the existing ridership patterns for both weekdays and weekends. For each of these cluster solutions, average silhouette widths allied with hourly, daily and weekly ridership patterns were examined to determine the final set of clusters. Through this process, 5 cluster and 7 cluster solutions were determined as the best solution for weekday and weekend respectively given their higher silhouette widths and relatively distinct travel patterns captured.
3. Last, separate univariate ARIMA or SARIMA models (5 for weekdays and 7 for weekends) were configured and fitted to the total hourly ridership for each of the final cluster solutions. These models were utilised to next estimate ARIMAX or SARIMAX models for individual bus stops. For example, if a stop was classified as Cluster 1 on weekdays, the corresponding (S)ARIMA model for that cluster was used, including weather variables and estimate a (S)ARIMAX model.

Table 2 summarises the number of bus stops associated with each of the derived clusters. Figs. 4 and 5 present the hourly ridership patterns across different clusters for weekdays and weekends. Relatively small differences were found for daily and weekly ridership, hence was not reported here. Yet rather distinct ridership patterns were captured by different clusters on an hourly basis.

Table 2
Weekday and weekend clusters of bus stops.

Period	Cluster	Number of stops
Weekday	1	1286
	2	965
	3	2325
	4	750
	5	644
Weekend	1	496
	2	526
	3	696
	4	484
	5	497
	6	1190
	7	607

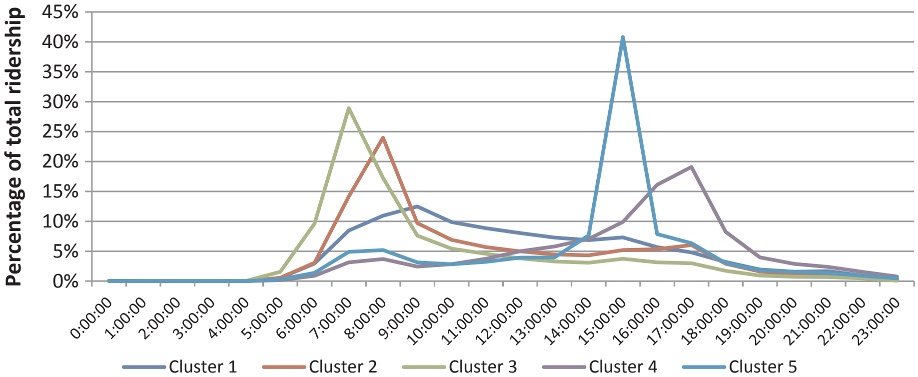

Fig. 4. Average hourly ridership patterns for weekday clusters.

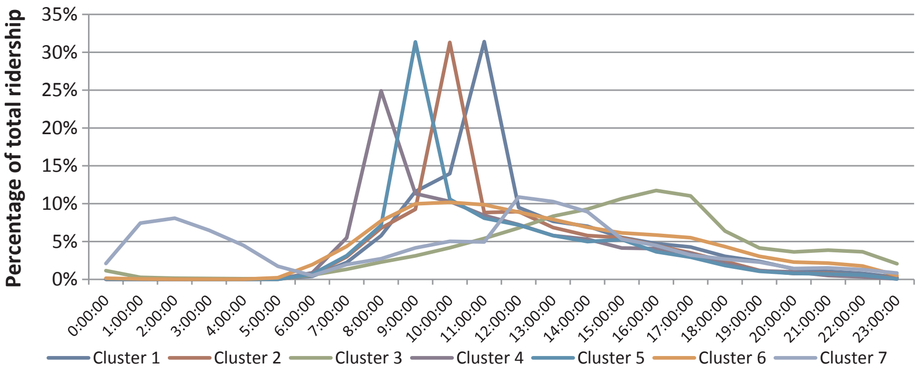

Fig. 5. Average hourly ridership patterns for weekend clusters.

4. Results

Results are presented in three parts: First we report the results for the system-wide models, followed by those for destination-based and stop-level models.

4.1. System-wide models

We first visually inspected hourly system-wide ridership patterns on weekdays (Fig. 6) and weekends (Fig. 7). The marked decline in ridership in late March, early and late April is in parallel with the public holidays during these days. Except for this pattern of low ridership, a strong recurring pattern of hourly ridership persists across both weekdays and weekends. Hence differencing of the hourly ridership is needed to achieve the required stationarity. By plotting the autocorrelation function (ACF) of the hourly ridership, we identified significant autocorrelations (i.e., correlation coefficients above 0.7 at the $p < .05$ level) for daily ridership on 24-h

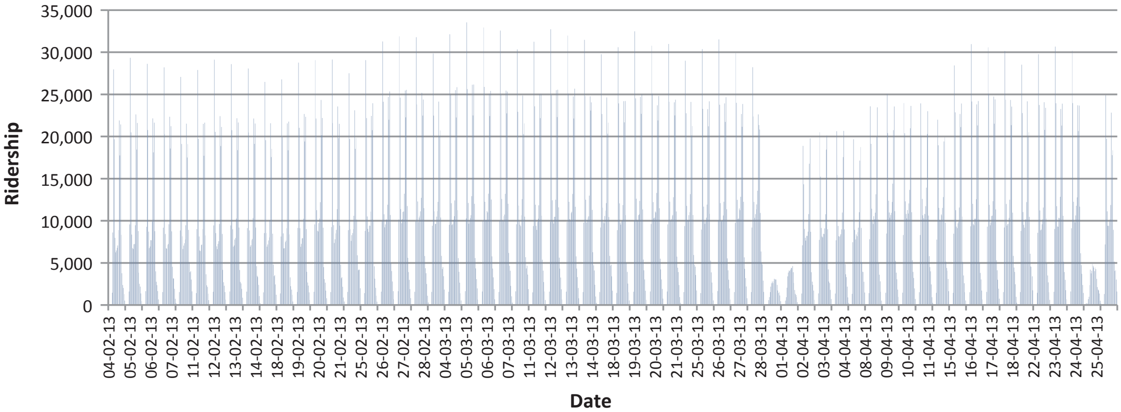

Fig. 6. Weekday hourly ridership.

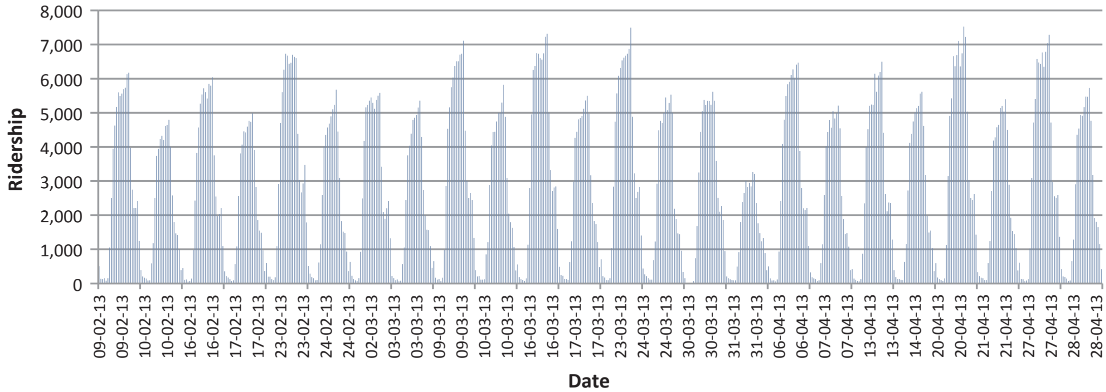

Fig. 7. Weekend hourly ridership.

intervals for weekdays, whilst for weekends significant autocorrelation was found for weekly ridership on 48-h intervals. To address these autocorrelation effects, hourly ridership was differenced to remove the observed periodicity and achieve a more stationary time series. Specifically, times series were differenced at the 24th lag, i.e., $(1\text{-}B)^{24}Y_t$ for weekday hourly ridership, and at the 48th lag, i.e., $(1\text{-}B)^{48}Y_t$, for weekend hourly ridership. After this differencing process, re-examining the ACF of the differenced ridership revealed that most of the temporal dependence was then removed and insignificant at the 0.05 level.

Following the differencing of the hourly ridership, SARIMA models were next estimated including weather conditions as explanatory variables. Seasonal and non-seasonal AR and MA parameters were determined through examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) of model residuals. A range of modelling trials were carried out, which entailed adding statistically significant (i.e., p-value $< .05$) AR and MA parameters to, and excluding statistically insignificant terms from our models with the aim of minimising Akaike's Information Criterion (AIC) values. Through this process, two different models were retained for weekdays and weekends (AIC = 24,529.38 and 7,358.82 respectively). We used the Ljung–Box tests for detection of serial autocorrelation. The results showed that absence of significant autocorrelation in the residuals of the two estimated models, which approximate white noise at the 0.05 level. The estimated model for weekdays was:

$$Y_t = \frac{(1-0.24B^1-0.309B^{10})}{(1-0.671B^1-0.219B^8+0.177B^{10})(1+0.354B^{24}+0.058B^{48})}e_t \tag{5}$$

And, for weekends:

$$Y_t = \frac{(1-1.06B^{48})}{(1-0.909B^1)(1+1.06B^{48})}e_t \tag{6}$$

In Eqs. (5) and (6), B^1, B^8 and B^{10} are the non-seasonal backshift operators of Y_t; B^{24} and B^{48} are the seasonal operators. The numbers (e.g., −0.24, −0.309) in the numerator of Eqs. (5) and (6) are the estimated coefficients associated with the MA terms, while the numbers (e.g., −0.671, −0.219) in the denominator parts are the estimated coefficients for the AR terms. We note that the constant terms of our models were statistically insignificant.[1] The auto-regressive terms at 1st, 8th and 10th lags are also explainable.

[1] For simplicity, we do not report the differencing operation and constant terms. These are available from the authors upon request.

Table 3
Results of modelling system-wide weather-ridership relationships.

Model #	Variable	Coefficient	P-value	AIC	MSE	MAPE
	Weekday					
1	School holiday	477	0.304	24,496.28	1,723,386	10.63%
	Public holiday	−3048	0.000***			
2	Temperature	−82	0.37	24,482.79	1,722,857	10.75%
3	Lagged Temperature	3	0.965			
4	Rainfall	−41	0.691	24,483.7	1,724,178	10.69%
5	Lagged rainfall	−19	0.868			
6	Relative humidity	0	0.993	24,483.34	1,723,629	10.74%
7	Lagged relative humidity	10	0.496			
8	Wind speed	−13	0.886	24,482.97	1,723,395	11.1%
9	Lagged wind speed	−65	0.49			
10	Apparent temperature	−72	0.567	24,482.51	1,722,730	10.76%
11	Lagged apparent temperature	38	0.831			
	Weekend					
12	School holiday	−27	0.891	7,356.72	62,652	2.3%
	Public holiday	−516	0.034**			
13	Temperature	54	0.003***	7,336.66	61,607	2.45%
14	Lagged Temperature	−15	0.445			
15	Rainfall	−83	0.000***	7,297.74	57,113	2.89%
16	Lagged rainfall	−24	0.012**			
17	Relative humidity	−9	0.022**	7,333.198	60,929	2.51%
18	Lagged relative humidity	−3	0.41			
19	Wind speed	−22	0.219	7,335.603	61,184	2.42%
20	Lagged wind speed	53	0.007***			
21	Apparent temperature	43	0.000***	7333, 818	61,116	2.4%
22	Lagged apparent temperature	−28	0.044**			

*** $p < .001$.
** $p < .05$.

The auto-regressive at the 1st lag indicates that travel demand for hours immediately pre and proceeding exhibit a level of interdependence, which relates to its continuous change over time. The terms at the 8th and 10th hours roughly coincide with the commuting hours during the morning and evening, wherein a peak of travel demand re-occurs from day to day. As demonstrated in previous research, such repeated patterns play a key role in governing people's arrangement of daily trip-making (Hanson and Huff, 1988; Schlich and Axhausen, 2003).

Weather conditions variables are likely to correlate with each other. Low temperatures tend to correlate with rainfall and cold wind. To avoid these problems of multicollinearity, for both weekdays and weekends, we computed a SARIMAX model for each of our five weather variables (temperature, rainfall, humidity, wind and apparent temperature) as well as their lagged counterparts. Following the methodology adopted in previous studies, e.g., Van den Bossche et al. (2004), Cools et al. (2009), weather variables were differenced in the same way as for hourly ridership to remove the effects of seasonality and autocorrelation; that is, the weather variables were differenced on a 24 h interval for weekday models, and differenced on a 48 h interval for weekend models. The modelling results are reported in Table 3, with incremental numerical ids assigned to each of the individual models as indicated by the 'Model #' column. In addition, two models (one for weekdays and one for weekends) that only includes dummy variables for public (= 1) and school holidays (= 1) were estimated as baseline models, of which the results are also reported as well (i.e., Models #1 and #12). These two dummy variables were also included in all other models (i.e., Models #2 to #11 and #13 to #22). The effects of these calendar-event variables remained largely comparable across all models. For simplicity, they are not reported repeatedly. Additionally, we report three model fitting indicators, namely AIC, mean square error (MSE) and mean absolute percentage error (MAPE) to provide a measure of fit for our models.

Coefficients for weather variables are interpreted as the predicted change in ridership give an one-unit change in one of our weather variables; that is, a 1 °C change in temperature, a 1 mm change in rainfall, 1% change in relative humidity, 1 m/s change in wind speed and 1 unit change in apparent temperature[2] (Table 3[3]). Examining the estimated coefficients for the holiday dummies reveals that only public holidays has a statistically significant effect on bus ridership. It appeared to induce a marked decrease in bus ridership on both weekdays and weekends, particularly on the former. This may be due to the non-discretionary nature of trips during weekdays: commuting, school- and university-related trips. Examination of the model fit indicators shows marginal improvements in the explanatory power of our models, including weather variables, compared to holiday-only models, especially analysing the AIC and MSE scores. This suggests that such as calendar events, activity patterns and purposes, changes in weather conditions in general

[2] Since apparent temperature is a composite score calculated using temperature, relative humidity and wind speed, it does not actually have a unit. Yet an increase in apparent temperature can be interpreted as an increase in heat stress.

[3] The two dummy variables for public and school holidays were also included in models #2 to #11, and models #13 to #22. Given the largely comparable effects of these variables, they are not reported repeatedly.

Table 4
Univariate models for the four destinations.

Destination	Differencing	Univariate Model	AIC
Weekday			
CBD	$(1-B)^{24}Y_t$	$Y_t = \frac{(1+0.361B^2)}{(1-1.1B^1+0.247B^2)}e_t$	20,830.19
The University of Queensland	$(1-B)^{24}(1-B)^{120}Y_t$	$Y_t = \frac{(1-0.198B^1)(1+0.866B^{120})}{(1-0.722B^1)(1+0.865B^{120})}e_t$	17,215.31
Garden City Shopping Centre	$(1-B)^{24}(1-B)^{120}Y_t$	$Y_t = \frac{(1+0.199B^1)}{(1-0.849B^1)}e_t$	14,290.78
Indooroopilly Shopping Centre	$(1-B)^{24}(1-B)^{120}Y_t$	$Y_t = \frac{(1+0.191B^1)}{(1-0.6B^1-0.185B^2)}e_t$	13,372.51
Weekend			
CBD	$(1-B)^{48}Y_t$	$Y_t = \frac{(1-0.986B^{48})}{(1-0.808B^1)(1+0.987B^{48})}e_t$	6405.2
The University of Queensland	$(1-B)^{48}Y_t$	$Y_t = \frac{(1+0.425B^1)}{(1-0.845B^1)}e_t$	5061.45
Garden City Shopping Centre	$(1-B)^{48}Y_t$	$Y_t = \frac{(1+0.505B^1)}{(1-0.868B^1)}e_t$	4704.46
Indooroopilly Shopping Centre	$(1-B)^{48}Y_t$	$Y_t = \frac{(1+0.574B^1)}{(1-0.884B^1)}e_t$	4454.14

exert important but subtle effects in shaping bus use patterns across the bus network during weekdays and weekends.

No statistically significant concurrent or lagged effects were found on ridership on weekdays. This reinforces our interpretation that bus public users appear to have little discretion on their trips during weekdays, reflecting the nature of activities underpinning these trips. Nonetheless, particular weather variables appeared to exert significant effects on weekends. Specifically, temperature and lagged wind speed seemed to have a positive influence on ridership, while negative concurrent and lagged effects were found for rainfall. Relative humidity was found to have a small yet significant negative impact on ridership. These findings in general affirms previous studies that detected weekend trips being more subject to the influence of changing weather conditions than weekdays on an hourly basis, in correspondence with Guo et al. (2007), Singhal et al. (2014). In particular, wet periods appeared to discourage bus use whereas more pleasant weather characterised by warmer temperatures encouraged bus use on weekends. The positive effect of lagged wind speed appears to be contradictory to previous studies that found negative effects of wind on active transport, such as cycling e.g., Miranda-Moreno and Nosal (2011). This in part may be due to the fact that the wind speeds captured in this study are relatively light (around 5 m/s), which might be considered to be pleasant within our study context. Moreover, the positive effect detected for apparent temperature lends insights into the combined effects when different weather variables (i.e., temperature, wind and humidity) are concurrent. In particular, it appears that the negative effect of relative humidity was largely repressed when there was increase in both temperature and wind speed. Hence this variable (relative humidity) may be of less practical importance relative to the other two variables (temperature and wind).

4.2. Destination-based models

We next modelled the hourly weather-ridership relationships for our four selected destinations (in Fig. 3). Following the system-wide analysis, univariate ARIMA (or SARIMA) models were separately estimated for each destination based on a thorough examination of ACFs, PACFs of residuals and AICs. Residuals for these models were found to be close to white noise based on the results of Ljung–Box tests. Table 4 reports the estimated models for weekdays and weekends.

Similar to the system-wide analysis, for destination-based analysis, we report coefficients from 11 different models that were separately estimated: a model including only dummy variables for school and public holidays, and 10 separate models including each of our weather (concurrent and lagged) variables. Constant terms were statistically insignificant and are not reported. Tables 5 and 6[4] show the estimated coefficients for weekdays and weekends respectively. In line with the system-wide analysis, a marked negative effect was found for public holidays across each of our selected destinations on both weekdays and weekends. School holidays were found to have a positive effect for weekend ridership associated with university-related trips, possibly because of extracurricular activities during such periods, such as workshops, open days and marathons.

On weekdays, no significant effects were found for changes in weather conditions on bus ridership bound to the CBD. This finding might be attributed to likelihood that most trips involving this destination were routine commute trips. For university-bound ridership, a notable positive effect was detected for rainfall, suggesting the possibility of a modal shift among certain trip-makers. For example, more passengers, particularly tertiary students, might prefer bus over alternative travel modes, such as car, walking and cycling during wet weather conditions. In addition, wind speed was found to exert a small but significant negative effect on ridership (p-value $< .1$). However, given the insignificant effect of apparent temperature, it appears that the effect of wind on ridership for this destination was largely mitigated as a consequence of a concurrent shift in the other weather variables (e.g., an increase in

[4] Again, the effects of the school and public holidays were only reported for the baseline models (models #1, #12, #23 and #34), but not for the remaining models given their largely similar effects across all models. This is the same for Table 6.

Table 5
Results of modelling destination-based weather-ridership relationships on weekdays.

Model #	Variable	Coefficient	P-value	AIC	MSE	MAPE
	CBD					
1	School holiday	109	0.275	20,813.87	139,948	3.18%
	Public holiday	−648	0.000***			
2	Temperature	−1	0.948	20,804.17	140,046	3.18%
3	Lagged Temperature	1	0.955			
4	Rainfall	−9	0.804	20,803.72	140,002	3.2%
5	Lagged rainfall	−10	0.801			
6	Relative humidity	−3	0.547	20,802.88	139,919	3.23%
7	Lagged relative humidity	−1	0.894			
8	Wind speed	−12	0.603	20,803.34	139,964	3.14%
9	Lagged wind speed	−8	0.72			
10	Apparent temperature	−1	0.967	20,804.17	140,046	3.18%
11	Lagged apparent temperature	1	0.785			
	The University of Queensland					
12	School holiday	79	0.199	17,209.98	33,254	2.4%
	Public holiday	−181	0.005***			
13	Temperature	−9	0.282	17,198.71	33,213	2.52%
14	Lagged Temperature	16	0.147			
15	Rainfall	18	0.036**	17,199.93	33,308	2.65%
16	Lagged rainfall	1	0.911			
17	Relative humidity	2	0.25	17,199.46	33,233	2.59%
18	Lagged relative humidity	−1	0.751			
19	Wind speed	−12	0.072*	17,197.26	33,167	2.65%
20	Lagged wind speed	5	0.555			
21	Apparent temperature	6	0.262	17,204.76	33,433	2.53%
22	Lagged apparent temperature	3	0.62			
	Garden City Shopping Centre					
23	School holiday	8	0.694	14,290.92	3560	1.6%
	Public holiday	−59	0.007***			
24	Temperature	−1	0.761	14,284.67	3562	1.6%
25	Lagged Temperature	1	0.637			
26	Rainfall	2	0.396	14,282.03	3555	1.6%
27	Lagged rainfall	−3	0.248			
28	Relative humidity	0	0.679	14,284.67	3562	1.61%
29	Lagged relative humidity	0	0.765			
30	Wind speed	−1	0.812	14,284.77	3562	1.61%
31	Lagged wind speed	0	0.925			
32	Apparent temperature	1	0.583	14,284.23	3561	1.61%
33	Lagged apparent temperature	1	0.778			
	Indooroopilly Shopping Centre					
34	School holiday	7	0.531	13,370.33	1747	1.3%
	Public holiday	−40	0.001***			
35	Temperature	−4	0.039**	13,361.54	1744	1.3%
36	Lagged Temperature	3	0.042**			
37	Rainfall	5	0.012**	13,351.15	1742	1.37%
38	Lagged rainfall	−5	0.007***			
39	Relative humidity	1	0.065*	13,361.65	1744	1.32%
40	Lagged relative humidity	0	0.187			
41	Wind speed	0	0.835	13,360.94	1743	1.3%
42	Lagged wind speed	−2	0.12			
43	Apparent temperature	−1	0.369	13,360.1	1742	1.3%
44	Lagged apparent temperature	3	0.03**			

*** $p < .001$.
** $p < .05$.
* $p < .1$.

temperature and wind speed).

Bus ridership for Indooroopilly Shopping Centre was found to be significantly influenced by changes in weather conditions. Specifically, positive effects were found for rainfall, relative humidity; and negative effects were found for temperature. Hence it appears that under less pleasant weather conditions, such as rainy hours coupled with lower temperature, more passengers travelled to this destination. A modal shift for trips (e.g., from walking to bus) to this particular shopping centre may also contribute to this finding. Further examination of the temporal distribution of ridership associated to this location indicates that Indooroopilly Shopping Centre mainly serves as a trip destination during afternoon peak hours. Given this, a possibility is that during inclement weather, some passengers (e.g., those just get off work) might choose to stop at the shopping centre for temporary shelter and shopping before heading towards their actual destinations (e.g., home). Temperature, apparent temperature and rainfall in the

Table 6
Results of modelling destination-based weather-ridership relationships on weekends.

Model #	Variable	Coefficient	P-value	AIC	MSE	MAPE
	CBD					
1	School holiday	28	0.53	6401.49	10,149	2.24%
	Public holiday	−158	0.001***			
2	Temperature	18	0.013**	6387.04	10,029	2.44%
3	Lagged Temperature	−9	0.211			
4	Rainfall	−21	0.000***	6372.29	9746	2.63%
5	Lagged rainfall	−7	0.284			
6	Relative humidity	−4	0.012**	6408.08	10,615	2.57%
7	Lagged relative humidity	1	0.378			
8	Wind speed	−7	0.289	6388.63	10,062	2.3%
9	Lagged wind speed	15	0.069*			
10	Apparent temperature	12	0.014**	6387.86	10,046	2.38%
11	Lagged apparent temperature	−9	0.105			
	The University of Queensland					
12	School holiday	23	0.01***	5052.76	816	1.91%
	Public holiday	−31	0.002***			
13	Temperature	−1	0.683	5047.81	817	1.94%
14	Lagged Temperature	1	0.6			
15	Rainfall	−1	0.39	5046.3	814	1.93%
16	Lagged rainfall	1	0.236			
17	Relative humidity	0	0.614	5047.99	817	1.92%
18	Lagged relative humidity	0	0.687			
19	Wind speed	−1	0.647	5047.79	817	1.9%
20	Lagged wind speed	0	0.913			
21	Apparent temperature	1	0.629	5047.64	816	1.92%
22	Lagged apparent temperature	0	0.811			
	Garden City Shopping Centre					
23	School holiday	3	0.637	4678	401	1.51%
	Public holiday	−21	0.007***			
24	Temperature	1	0.747	4670	398	1.55%
25	Lagged Temperature	1	0.717			
26	Rainfall	−2	0.063*	4669	398	1.56%
27	Lagged rainfall	1	0.427			
28	Relative humidity	0	0.807	4663	394	1.55%
29	Lagged relative humidity	0	0.295			
30	Wind speed	2	0.137	4670	399	1.52%
31	Lagged wind speed	−1	0.48			
32	Apparent temperature	0	0.839	4672	401	1.54%
33	Lagged apparent temperature	1	0.645			
	Indooroopilly Shopping Centre					
34	School holiday	12	0.003***	4446.21	258	1.37%
	Public holiday	−14	0.018**			
35	Temperature	1	0.323	4440.47	258	1.39%
36	Lagged Temperature	0	0.687			
37	Rainfall	−1	0.316	4433.2	254	1.39%
38	Lagged rainfall	−2	0.023**			
39	Relative humidity	0	0.792	4441.28	258	1.36%
40	Lagged relative humidity	0	0.48			
41	Wind speed	−1	0.636	4441.83	258	1.36%
42	Lagged wind speed	1	0.404			
43	Apparent temperature	1	0.162	4439.74	257	1.39%
44	Lagged apparent temperature	−1	0.387			

*** $p < .001$.
** $p < .05$.
* $p < .1$.

previous hour were, however, found to have positive effects. The reasons for these findings, however, are not readily identifiable and would require additional data and analyses (e.g., analysis of survey data on bus users' trip-making and activity change in response to weather). Such exercise, while calling for further attention, is beyond the scope of the current paper.

By comparison, bus ridership to the Garden City Shopping Centre does not appear to be significantly impacted by changes in weather conditions, with no coefficient being statistically significant. This might be attributed to that this shopping centre is mainly served by the bus services operating on the busway, wherein rail-like shelters exist to more comprehensively shield passengers from the prevailing weather conditions.

On weekends, weather variables were found to have larger and statistically significant impacts on CBD-bound ridership than weekdays. Specifically, rainfall was found to have negative effects, whilst positive effects were found for temperature and lagged

Table 7
Univariate models for different clusters.

Cluster	Differencing	Univariate model	AIC
Weekday			
Cluster 1	$(1\text{-B})^{24}\mathrm{Y_t}$	$Y_t=\frac{(1+0.112B^1)}{(1-0.869B^1)}e_t$	14,984.15
Cluster 2	$(1\text{-B})^{24}\mathrm{Y_t}$	$Y_t=\frac{(1-0.206B^1)}{(1-0.595B^1)}e_t$	18,637.24
Cluster 3	$(1\text{-B})^{24}\mathrm{Y_t}$	$Y_t=\frac{(1-0.434B^1)(1+0.905B^{24})}{(1-0.511B^1)(1-0.487B^{24})}e_t$	23,095.87
Cluster 4	$(1\text{-B})^{24}(1\text{-B})^{120}\mathrm{Y_t}$	$Y_t=\frac{(1-0.371B^1)(1+0.993B^{24})}{(1-0.77B^1)(1-0.505B^{24})}e_t$	20,780.49
Cluster 5	$(1\text{-B})^{24}(1\text{-B})^{120}\mathrm{Y_t}$	$Y_t=\frac{(1+0.539B^1)}{(1-0.754B^1)}e_t$	20,061.68
Weekend			
Cluster 1	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{(1+0.302B^1-0.158B^3)}{(1-0.557B^1)}e_t$	4250.98
Cluster 2	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{(1-0.122B^3)}{(1-0.214B^2)}e_t$	3957.98
Cluster 3	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{(1-1.096B^{48})}{(1-0.754B^1-0.121B^2)(1+1.097B^{48})}e_t$	6965.07
Cluster 4	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{1}{(1-0.368B^1)}e_t$	4050.34
Cluster 5	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{(1+0.65B^1)}{(1-0.802B^1)}e_t$	3770.67
Cluster 6	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{1}{(1-0.873B^1)}e_t$	6733.5
Cluster 7	$(1\text{-B})^{48}\mathrm{Y_t}$	$Y_t=\frac{(1+0.283B^1)}{(1-0.705B^1)}e_t$	4465.39

wind speed. In line with the system-wide model, the positive effect of apparent temperature again, suggests that the effect of relative humidity was possibly overshadowed by the changes in other weather variables. These findings also suggest that: (1) compared to weekdays, more CBD-bound trips might be associated with recreational purposes, such as going to the parks, theatres or dinning out; and (2) echoing the finding from the system-wide analysis, less pleasant weather conditions tended to discourage people from using the bus for recreational trips to the CBD, which might be a function of both a modal shift (from bus to other travel modes such as cars) and trip cancellation. This also appears to be the case for Garden City and Indooroopilly Shopping Centres as rainfall and lagged rainfall were found to have significant negative effects on ridership connected with these two destinations. In contrast, changes in weather conditions did not lead to significant changes in university-related ridership. This possibly relates to the fact that except for those who reside nearby, few users actually need to travel to the university on weekends, hence the insignificant effects of weather on bus ridership.

4.3. Stop-level models

Drawing on the methodology described in Section 3.3, we also conducted stop-level modelling in order to further reveal the spatial variability of the hourly weather-ridership relationship across Brisbane. Various univariate ARIMA and SARIMA models were first developed and estimated for each of the clusters identified in Section 3.3. Models with relatively smaller AIC values were retained, and are summarised in Table 7.

Drawing on the above cluster-based models, ARIMAX and SARIMAX models were next separately estimated for each weather variable (including concurrent and lagged terms) and individual stops. A total of over 100,000 models (over 60,000 for weekdays and over 40,000 for weekends) were estimated. For the sake of display and practical reasons, only stops with statistically significant coefficients were retained. Based on these stops, a spatially continuous surface of the effect of changes in weather variables on bus boarding –as captured by the model coefficients- were generated and visualised using the inverse distance weighted (IDW) interpolation tool in the ESRI ArcGIS software package. For each map (Figs. 8–11), we report the number of stops that displayed a significant influence of changes in weather conditions on the bottom right, and the ranges of coefficients on the upper left, which correspond to natural breaks in their distribution. These coefficients are interpreted as the estimated change in bus ridership associated with one unit increase in a given weather variable.

On weekdays (Figs. 8 and 9), compared to other weather variables, our model results indicate that changes in temperature tend to affect the largest number of bus stops. In contrast and somewhat unexpectedly, rainfall appears to affect the least number of bus stops, which is true for both concurrent and lagged effects. The individual concurrent effects of weather conditions on ridership appear to be highly geographically localised (Fig. 8). Extreme changes in ridership under the influence of changes in weather conditions are concentrated in certain locations, including our selected destinations: the university and the CBD. This, to some extent, reinforces that these activity-intense areas are more subject to the effects of changing weather patterns than bus stops across the bus network.

In addition, the concurrent effects of weather appear to also incur changes in bus use in the opposite ways across bus stops. Such variance of effects may have to do with the local built environment, demographic and activity profiles. For example, higher

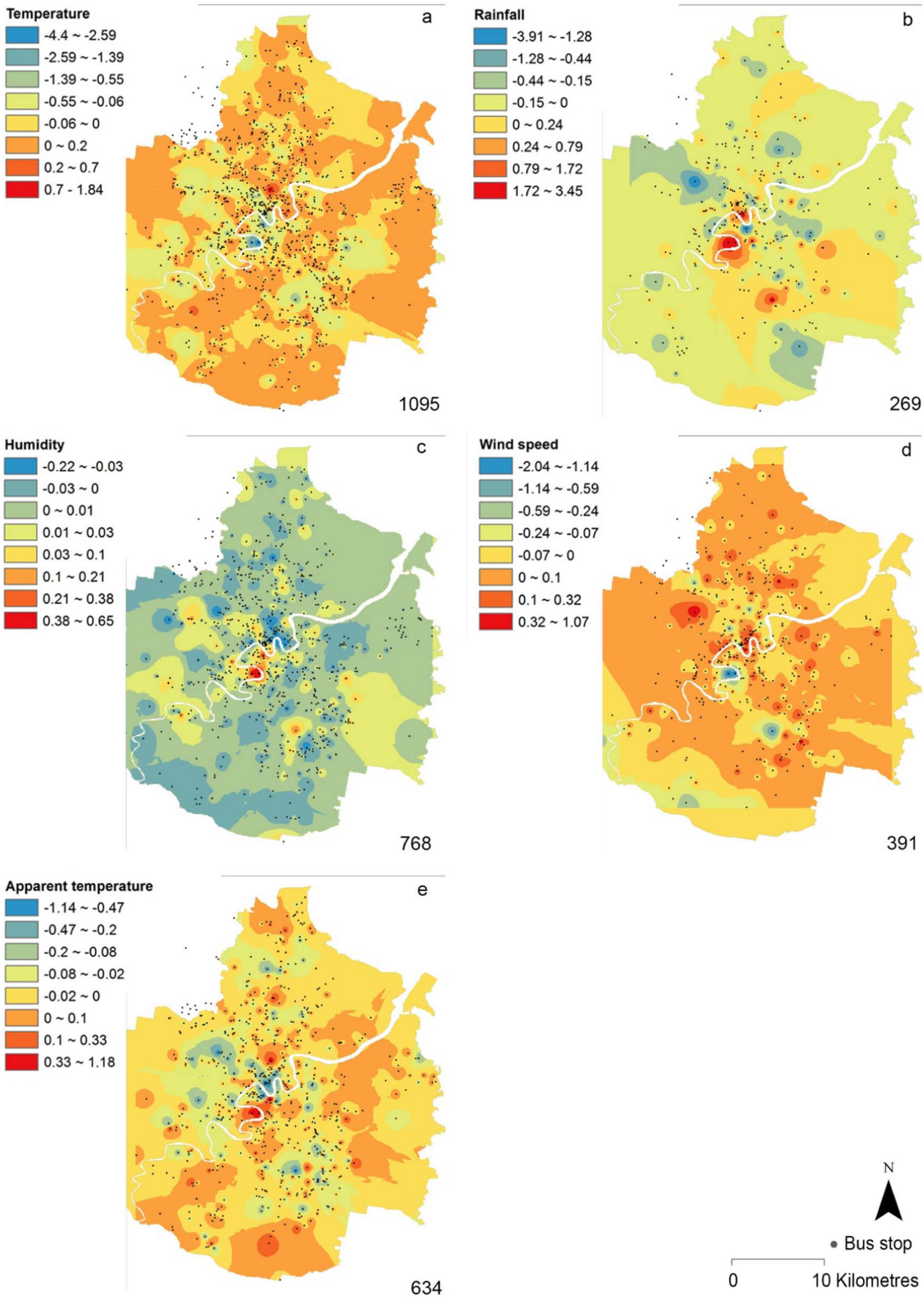

Fig. 8. The concurrent effects of weather on weekdays.

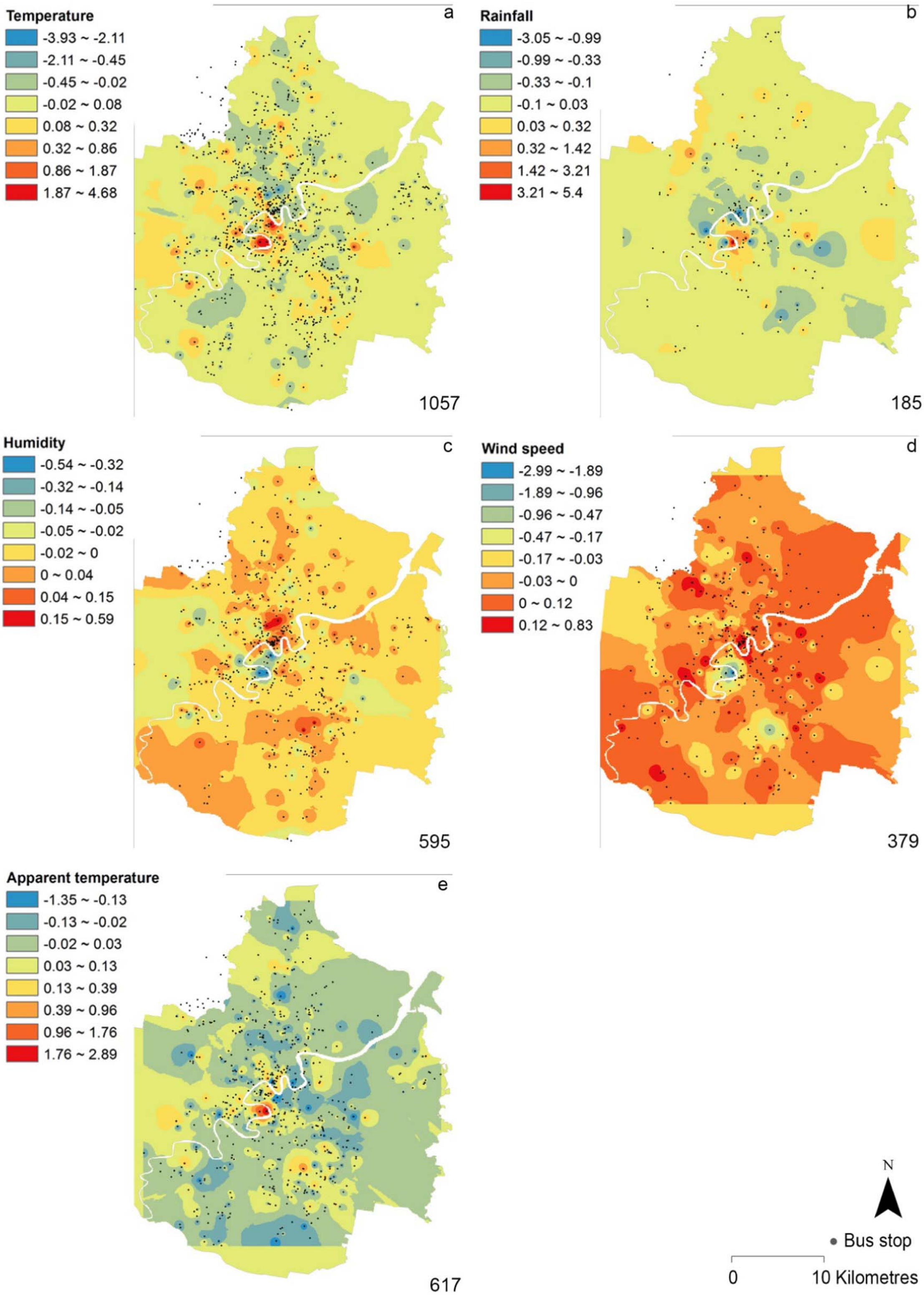

Fig. 9. The lagged effects of weather on weekdays.

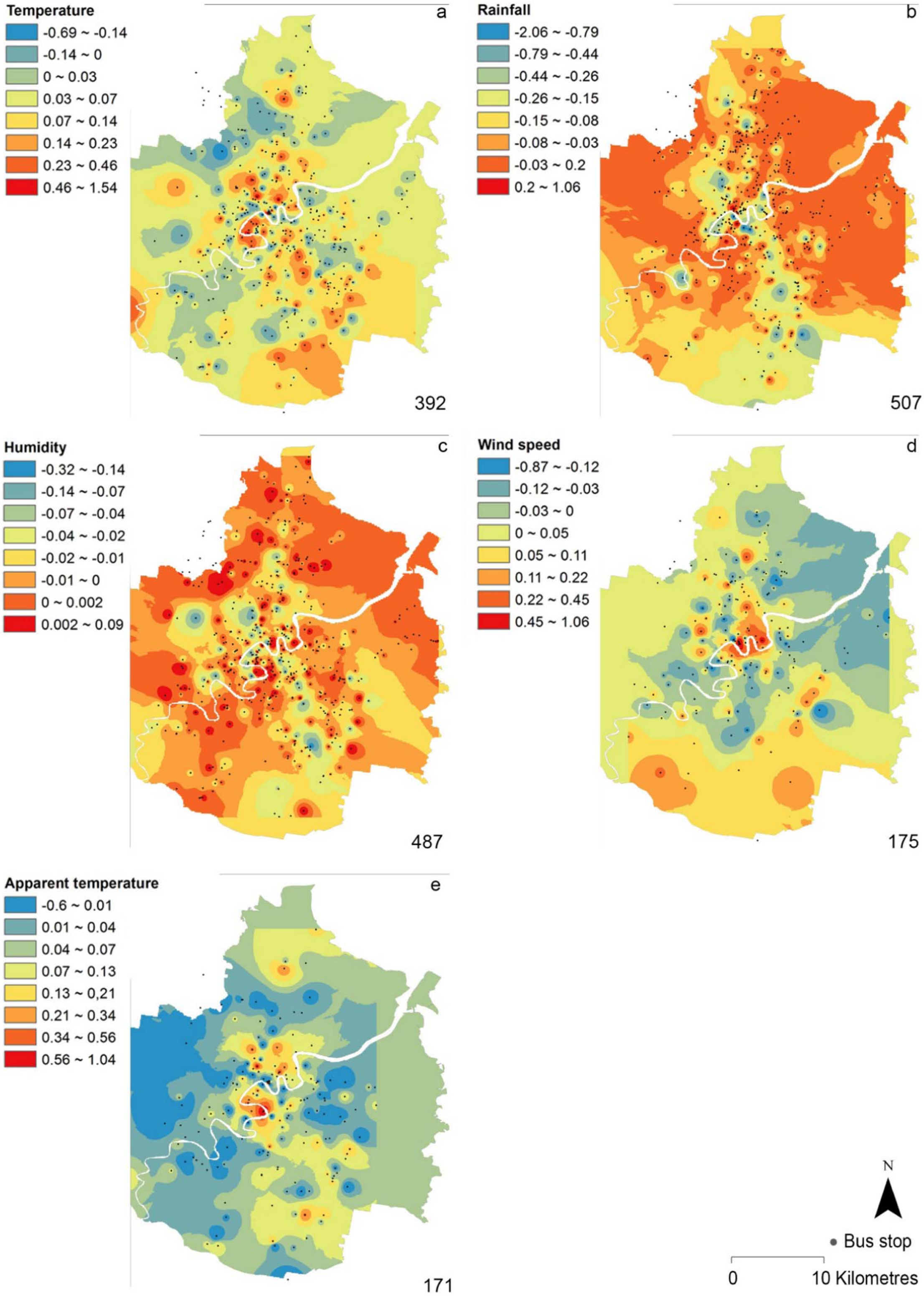

Fig. 10. The concurrent effects of weather on weekends.

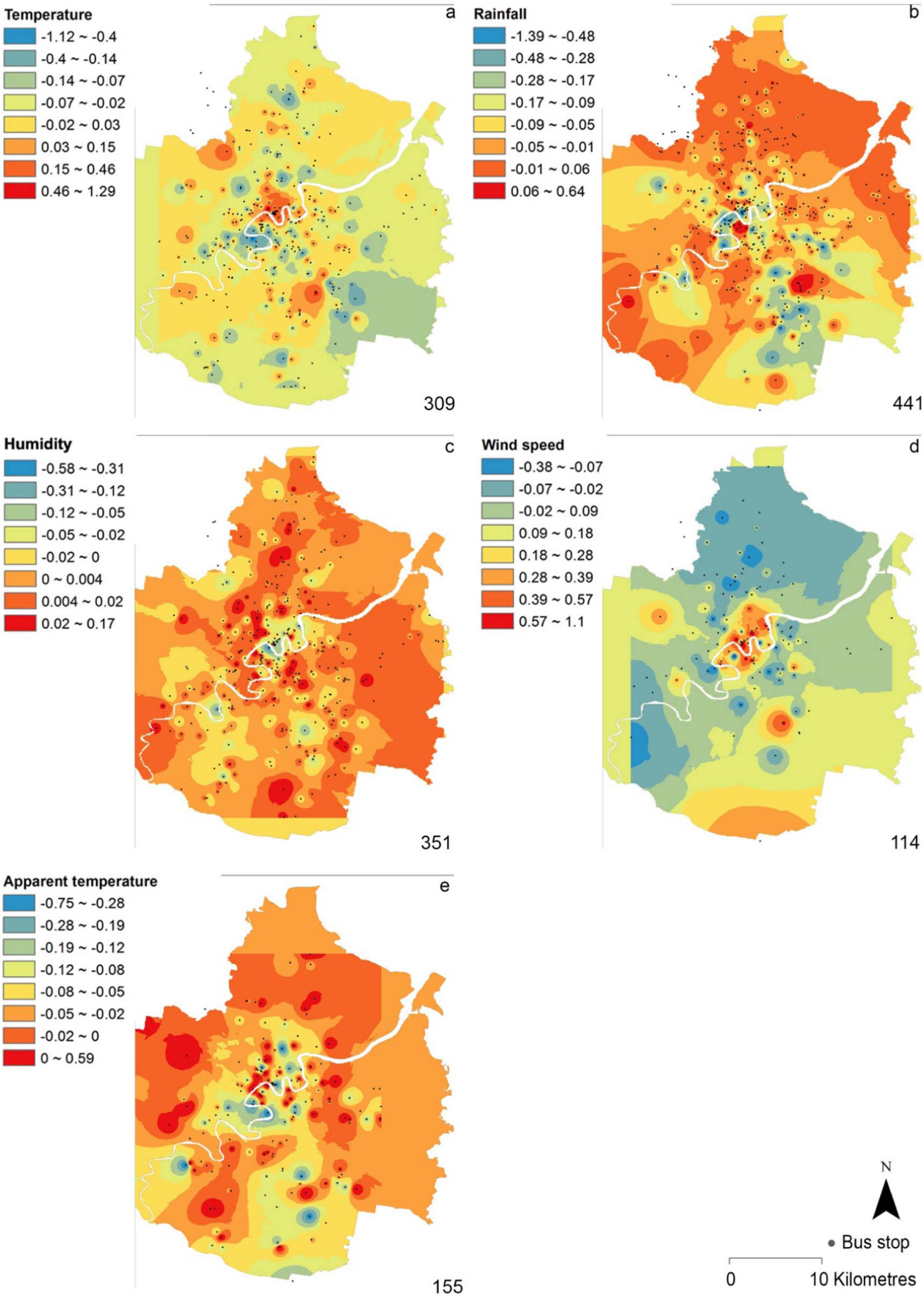

Fig. 11. The lagged effects of weather on weekends.

temperature and stronger wind were found to have a deterring effect on bus use at the university area, while other places experienced such influence to a lesser extent or even leading to a slight increase in bus ridership (Fig. 8a and c). This may in part be attributed to that the university area is relatively less sheltered compared to some other locales (e.g., the CBD), hence more trips cancelled at this location during less pleasant. Increase in rainfall was found to be associated with increased bus use at the CBD, shopping centre and university areas. The effects of relatively humidity, while significant at many locations, are marginal compared to rainfall and wind (i.e., the changes in bus ridership were largely between −0.2 and 0.1) (Fig. 8c). Last, examining the effects of apparent temperature (Fig. 8e) indicates that the combined effects of weather variables (particularly temperature, wind speed and relative humidity) affected fewer bus stops. And for some locations such as the university (i.e., UQ), apparent temperature was found to exert an inverse effect to that of the individual variables (e.g., temperature). These findings again suggest the effects of individual weather variables tend to either suppress or indeed cancel out or suppress one another. Compared to the concurrent effects of weather conditions, their lagged effects appear to be largely modest, with associated changes in bus use between −0.03 and 0.03 (Fig. 9). A closer look shows that the spatial patterns of the effects of rainfall, wind speed and apparent temperature are in large part consistent with their concurrent counterparts.

On weekends (Figs. 10 and 11), weather conditions appear to affect considerably fewer stops (and the combined effects captured by apparent temperature affected even fewer stops) than on weekdays, except for rainfall. Rainfall seems to affects the larger number of bus stops on weekends than weekdays, suggesting that it plays a more important role in influencing bus riders' trip-making decisions on weekends than weekdays. Concerning the spatial patterns of weather influencing bus use (Fig. 10), although still largely varied across the study context, some discernible patterns emerge. With regard to concurrent effects of weather, some locations showing evidence of significant weather-ridership relationships appear to form a corridor that aligns with Brisbane's busway. In particular, wet weather was found to decrease bus use along the busway corridor (Fig. 10b and c), although the effects of relative humidity was rather marginal. The reason for decreased bus use along the busway during wet weather may indicate a modal shift from bus to other transport modes, such as cars for leisure trips on weekends. The effects of wind speed is, however, less spatially discernible by comparison.

The spatial patterns of the lagged effects of weather (Fig. 11) are somewhat similar to the concurrent effects. This is especially the case for rainfall. Yet, spatial shifts were observed for wind speed and apparent temperature at certain stops. For example, the lagged effects of wind speed appear to incur decreased bus use in more places in the north of Brisbane than its concurrent effects. Additionally, the effects of apparent temperature of the previous hour were not spatially systematic, with the university area, experiencing contrasting lagged (negative) and concurrent (positive) effects. While the reasons underpinning these observed effects are difficult to extract, this may reflect the differentiated relative importance of different weather variables affecting bus ridership at particular locations across Brisbane.

5. Discussion and conclusions

A growing interest in transport studies has been to unveil the dynamics between weather and public transport use given that changes in weather conditions have the potential to influence public transport services and its users in a multitude of ways (Arana et al., 2014; Böcker et al., 2013; Guo et al., 2007). However, this relationship has arguably yet to be fully understood especially with regard to the concurrent effects weather exerts on public transport ridership and their variability across urban area. Through estimating ARIMAX and SARIMAX models on an integrated data set of bus ridership and weather measurements, this paper aimed to investigate the hourly effects of weather on bus ridership at three different spatial scales: system-wide, destination-based and stop-level. A series of meaningful insights with implications for policy were derived from our analyses, each of which is now discussed.

First, in line with previous studies that have examined daily weather-ridership relationships (e.g., Guo et al. 2007; Kalkstein et al. 2009), our system-wide modelling analysis revealed that hourly bus ridership on weekends was considerably more affected by changing weather conditions than weekdays. This suggests that even at a fine temporal scale, weekday bus use across Brisbane is predominantly shaped by people's routinised behavioural patterns (commuting), and is less governed by weather. Yet, on weekends, hourly bus ridership was found to be promoted by warmer weather, coupled with the presence of a light breeze, with the combined effect of reducing the negative effects of higher temperatures and humidity, arguably linked to the more discretionary nature of weekend trip patterns.

Second, our destination-based models highlighted that the effects of weather on hourly bus ridership varied not only between weekdays and weekends, but also across trip destinations that we argue can be explained by distinctions between their function and associated infrastructure. More specifically, on weekdays, Brisbane's CBD and Garden City Shopping Centre both located adjacent to the busway were found not to be affected by changes in weather conditions. A major university was found to be influenced by rainfall and to a lesser extent by wind speed, possibly due to the behavioural change of tertiary student bus riders especially during inclement weather, such as taking the bus instead of other travel options such as walking, cycling and use of private cars. The Indooroopilly Shopping Centre that is served mainly by an on-road bus services was found to be influenced by a number of weather variables including temperature, rainfall and their lagged counterparts. This suggests that ridership bound to this destination is more subject to the influence of weather compared to Garden City Shopping Centre because of the relatively limited shelter along bus routes connecting the Indooroopilly Shopping Centre. On weekends, the CBD experienced increase in ridership during warmer and slightly windy conditions and decreased ridership during wet periods. Ridership bound to the two shopping centres also experienced a decrease in ridership during periods of rainfall, whereas the university was not influenced by weather on weekends, probably because of the reduced bus use during this period.

Third, through a methodology that combines cluster analysis with time-series modelling, we modelled and visualised stop-level trip generation vis-à-vis weather conditions. This exercise highlighted that on both weekdays and weekends, different locales across the study context experienced effects of weather on bus use at differing levels or even in opposite directions, for instance, increases in temperature resulted in increased bus use in some places while decreased bus use in others. The activity-intense locales, including the CBD, university and shopping centres, were found to be more susceptible to changes in weather conditions than other places. These highly localised patterns of bus use in response to weather, we contend, are partially a function of local built environment (e.g., sheltered versus less sheltered places) in conjunction with demographic and activity profiles of bus passengers (e.g., tertiary students versus workers, flexible versus rigid schedules). Furthermore, contrasting concurrent and lagged (i.e., positive versus negative) effects of weather variables was detected for particular areas, which may reflect certain coping mechanisms to changing weather conditions adopted by bus riders, such as delaying trip-making to a slightly later time when there is onset of high temperatures or heavy rainfall.

To this end, our findings indicate that, while less dominant than other factors such as calendar events, temporary fluctuations of weather conditions indeed may induce concurrent behavioural changes of bus passengers, which vary markedly across urban space. The implications of these findings, we argue, are twofold. First, in addition to other conventionally relevant information such as traffic volume and time of day, real-time weather information should be taken into account in the monitoring of the demand for bus transit over the course of the day. While this may not always mean instant adjustment of transit service associated with changes in weather conditions, such a constantly updated information base may better equip transit operators with the ability to make timely adjustments especially in the face of sudden changes in weather, like the onset of high temperatures or spells of heavy rainfall. For example, for bus stops that were found to experience decreases in ridership under adverse weather conditions, transit operators may consider upgrades to bus stops to provide improved shelter for passengers from weather elements. For bus stops that are associated with increases in passenger demand, on the basis that the particular service capacity is reached or exceeded, more services might be considered in order to cope with the additional demand. Second, considering the spatial variation in the weather-bus transit relationship, monitoring of bus demand should also be implemented across different areas of the city. Although it is not realistic to monitor every bus stop, a worth-trying start may be to focus on certain major trip destinations and origins that are more subject to the influence of weather, such as the CBD, the university and shopping centres in Brisbane. This will also help the transit operators make more localised and targeted adjustment when necessary.

There are at least four avenues for future research to build on our study. First, given the availability of transit smart card data, we were only able to investigate the effects on bus use of weather over a relatively short period of time (three months). Drawing on larger data sets, future research may examine this relationship over longer periods (e.g., one to two years) and also, how replicable our findings are across other situational contexts. Second, our findings suggest that the behavioural response of bus passengers is highly localised and complex across the study context. Some patterns appear to be the function of local land use, built environment and socio-demographic characteristics of local neighbourhoods, while for others the underpinning mechanisms are less readily identifiable. Given this, explicitly incorporating the aforementioned factors in the modelling of weather influencing bus use will likely provide a more thorough understanding of the decision process of people's use of bus transit under the influence of weather. Third, at the time of writing, smart card records of other transit modes were not available to the researchers. Given the availability of smart card data for all public transit modes in Brisbane, it would be worthwhile to build upon the present study to examine the weather effects across public transport modes, such as shared bicycles, ferries, buses and trains, to reveal the dynamics of modal substitution during inclement weather. Last, as previously noted, in the stop-level models, potential spatial autocorrelation present in the weather-ridership relationship was not factored out, due to that a geographically weighted (S)ARIMAX model does not yet exist. Future research may tackle this technical problem. Yet, controlling for both spatial and temporal autocorrelation within the framework of (S)ARIMAX model would certainly not be a trivial task, given the potentially substantial alteration of the original model.

To conclude, this study has examined the spatial and temporal dynamics of the weather-ridership relationship at a level of detail previously unexplored. This adds a more complete understanding of how weather shapes individual's use of public transport (bus), and provides some meaningful implications for the management of transit services in response to dynamic changes in weather conditions. It is hoped that this study stimulates further research in this area in order to help develop more weather-responsive transit services.

Acknowledgements

We would like to thank Mr. Anthony Kimpton for his help in the coding process. We are also grateful for the valuable comments of three anonymous reviewers that helped substantially improve the manuscript.

References

ABS, 2013. 2011 Census QuickStats, Canberra: ABS. Available at: http://www.censusdata.abs.gov.au/census_services/getproduct/census/2011/quickstat/0?opendocument&navpos=220 Accessed 18th, November 2013.

Arana, P., Cabezudo, S., Peñalba, M., 2014. Influence of weather conditions on transit ridership: A statistical study using data from Smartcards. Transportation Res. Part A: Policy Pract. 59, 1–12.

Bagley, M.N., Mokhtarian, P.L., 2002. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. Ann. Reg. Sci. 36 (2), 279–297.

Biswas, P.K., Islam, M.Z., Debnath, N.C., Yamage, M., 2014. Modeling and roles of meteorological factors in outbreaks of highly pathogenic avian influenza H5N1. PLoS One 9 (6), e98471.

BITRE, 2013. Population growth, jobs growth and commuting flows in South East Queensland. (Report 134). Canberra, Australia.

BITRE, 2014. Urban public transport: updated trends (Information Sheet 59), Canberra, Australia.

BITRE, 2014. Public transport use in Australia's capital cities: modelling and forecasting (Report 129). Canberra, Australia.

Böcker, L., Dijst, M., Prillwitz, J., 2013. Impact of everyday weather on individual daily travel behaviours in perspective: a literature review. Transport Rev. 33 (1), 71–91.
Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, US.
Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 2008. Time Series Analysis: Forecasting and Control. John Wiley, Hoboken, N.J.
Brockwell, P.J., Davis, R.A., 2002. Introduction to Time Series and Forecasting. Springer, New York.
Bureau of Meteorology, 2015. Climate statistics for Australian locations. Canberra, Australia: Bureau of Meteorology, Australia Government. Available at: http://www.bom.gov.au/climate/averages/tables/cw_040842_All.shtml (Accessed December 2015).
Bureau of Meteorology, 2010. Thermal Comfort observations. Australian Government. Available at: http://www.bom.gov.au/info/thermal_stress/ (Accessed 7th, April 2015).
Cervero, R., 1998. The Transit Metropolis: A Global Inquiry. Island Press, Washington, DC.
Changnon, S.A., 1996. Effects of summer precipitation on urban transportation. Climatic Change 32 (4), 481–494.
Chapin, F.S., 1974. Human activity patterns in the city: things people do in time and space, New York, s.n.
Chen, Y., Tjandra, S., 2014. Daily collision prediction with SARIMAX and generalized linear models on the basis of temporal and weather variables. Transportation Res. Rec.: J. Transportation Res. Board 2432, 26–36.
Cools, M., Moons, E., Wets, G., 2009. Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: assessing the effect of holidays on two site locations. Transportation Res. Rec.: J. Transportation Res. Board 2136, 57–66.
Cravo, V.S., Cohen, J., Williams, A., 2009. Impact of weather on transit revenue in New York city. In: Transportation Research Board 88th Annual Meeting (No. 09–3036).
Creemers, L., Wets, G., Cools, M., 2015. Meteorological variation in daily travel behaviour: evidence from revealed preference data from the Netherlands. Theoret. Appl. Climatol. 120 (1–2), 183–194.
Currie, G., Delbosc, A., 2010. Bus rapid transit in Australasia an update on progress. Built Environ. 36 (3), 328–343.
DETA, 2013. School holidays calendar. Available at: http://education.qld.gov.au/public_media/calendar/holidays.html Accessed 14th August 2013.
Ephrath, J.E., Goudriaan, J., Marani, A., 1996. Modelling diurnal patterns of air temperature, radiation wind speed and relative humidity by equations from daily characteristics. Agric. Syst. 51 (4), 377–393.
Google Developers, 2012. Transit. Available at: https://developers.google.com/transit/gtfs Accessed April 2013.
Guo, Z., Wilson, N., Rahbee, A., 2007. Impact of weather on transit ridership in Chicago, Illinois. Transportation Res. Rec.: J. Transportation Res. Board 2034, 3–10.
Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity: views from urban planning. Am. J. Prev. Med. 23 (2), 64–73.
Hanson, S., Huff, O.J., 1988. Systematic variability in repetitious travel. Transportation 15 (1), 111–135.
Hine, J., Scott, J., 2000. Seamless, accessible travel: users' views of the public transport journey and interchange. Transp. Policy 7 (3), 217–226.
Hoffman, A., 2008. Advanced Network Planning for Bus Rapid Transit: The 'Quickway' Model as a Modal Alternative to 'Light Rail Lite'. Federal Transit Administration, U.S. Department of Transportation, Washington, DC.
Hofmann, M., O'Mahony, M., 2005. The impact of adverse weather conditions on urban bus performance measures an analysis using ITS technology. IEEE Conference on Intelligent Transport Systems, Vienna.
Ibrahim, M.F., 2003. Car ownership and attitudes towards transport modes for shopping purposes in Singapore. Transportation 30 (4), 435–457.
Kalkstein, A.J., Kuby, M., Gerrity, D., Clancy, J.J., 2009. An analysis of air mass effects on rail ridership in three US cities. J. Transp. Geogr. 17 (3), 198–207.
Kashfi, S., Lee, B., Bunker, J.M., 2013, October. Impact of rain on daily bus ridership: a Brisbane case study. In: Australasian Transport Research Forum 2013 Proceedings. Australian Transportation Research Forum.
Koetse, M.J., Rietveld, P., 2009. The impact of climate change and weather on transport: An overview of empirical findings. Transportation Res. Part D: Transport Environ. 14 (3), 205–221.
Krivoruchko, K., 2012. Empirical bayesian kriging implemented in ArcGIS geostatistical analyst. ArcUser, The Magazine for Esri Software Users. ESRI, pp. 6–10.
Liu, C., Susilo, Y.O., Karlström, A., 2015. The influence of weather characteristics variability on individual's travel mode choice in different seasons and regions in Sweden. Transp. Policy 41, 147–158.
Liu, C., Susilo, Y.O., Karlström, A., 2016. Measuring the impacts of weather variability on home-based trip chaining behaviour: a focus on spatial heterogeneity. Transportation 43 (5), 843–867.
Mapes, B.E., Warner, T.T., Xu, M., Negri, A.J., 2003. Diurnal patterns of rainfall in northwestern South America. Part I: Observations and context. Mon. Weather Rev. 131 (5), 799–812.
Mees, P., Dodson, J., 2011. Public transport network planning in Australia: Assessing current practice in Australia's five largest cities. Research paper, pp. 34.
Miranda-Moreno, L., Nosal, T., 2011. Weather or not to cycle: Temporal trends and impact of weather on cycling in an urban environment. Transportation Res. Rec.: J. Transportation Res. Board 2247, 42–52.
Morency, C., Paez, A., Roorda, M.J., Mercado, R., Farber, S., 2011. Distance traveled in three Canadian cities: Spatial analysis from the perspective of vulnerable population segments. J. Transp. Geogr. 19 (1), 39–50.
Schlich, R., Axhausen, K.W., 2003. Habitual travel behaviour: evidence from a six-week travel diary. Transportation 30 (1), 13–36.
Schwanen, T., 2002. Urban form and commuting behaviour: a cross-European perspective. Tijdschrift voor economische en sociale geografie 93 (3), 336–343.
Singhal, A., Kamga, C., Yazici, A., 2014. Impact of weather on urban transit ridership. Transportation Res. Part A: Policy Practice 69, 379–391.
Tao, S., Corcoran, J., Hickman, M., Stimson, R., 2016. The influence of weather on local geographical patterns of bus usage. J. Transp. Geogr. 54, 66–80.
Tao, S., Corcoran, J., Mateo-Babiano, I., Rohde, D., 2014. Exploring Bus Rapid Transit passenger travel behaviour using big data. Appl. Geogr. 53, 90–104.
Trépanier, M., Morency, C., Agard, B., Descoimps, E., Marcotte, J.S., 2012, July. Using smart card data to assess the impacts of weather on public transport user behavior. In: Conference on Advanced Systems for Public Transit-CASPT12, Santiago, Chile.
Van den Bossche, F., Wets, G., Brijs, T., 2004. A regression model with ARIMA errors to investigate the frequency and severity of road traffic accidents.
Vuchic, V.R., 2005. Urban Transit : Operations, Planning and Economics. Wiley, Hoboken, N.J.
Wang, X., Khattak, A., 2013. Role of travel information in supporting travel decision adaption: exploring spatial patterns. Transportmetrica A: Transport Sci. 9 (4), 316–334.
Williams, B., 2001. Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling. Transportation Res. Rec.: J. Transportation Res. Board 1776, 194–200.
Williams, B.M., Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. J. Transportation Eng. 129 (6), 664–672.