Carsharing customer demand forecasting using causal, time series and neural network methods: a case study

Elnaz Moein and Anjali Awasthi*

Faculty of Engineering and Computer Sciences,
Concordia University,
CIISE, Montréal, Canada
Fax: 514-848-7131
Email: el.moein1384@gmail.com
Email: awasthi@ciise.concordia.ca
*Corresponding author

Abstract: Carsharing services are becoming popular in recent times. Deploying right number of fleet at stations is a critical component in assuring high quality service for customers. This can be done efficiently if customer demand is predictable or known in advance. In this paper, we address the problem of customer demand forecasting for improving carsharing operations. Three categories of methods namely causal (regression forecast, regression forecast with seasonality adjustments), time series (exponential smoothing, moving average) and neural networks are evaluated for forecasting customer demand. An application of the proposed methods on demand data from a carsharing organisation called Communauto is provided. The results of our study show that neural network is the best method in this prediction. The proposed work has strong practical applicability. Having an accurate forecast of the customers' demands in different times of the year can help increase customer satisfaction and reach business performance targets. Especially if electric vehicles are used in carsharing companies, since they require special infrastructures.

Keywords: demand forecasting; carsharing; regression; exponential smoothing; moving average; neural networks.

Reference to this paper should be made as follows: Moein, E. and Awasthi, A. (2020) 'Carsharing customer demand forecasting using causal, time series and neural network methods: a case study', *Int. J. Services and Operations Management*, Vol. 35, No. 1, pp.36–57.

Biographical notes: Elnaz Moein is a MASc student in Quality Systems Engineering at the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada. Her areas of interest are quality management, supply chain management, forecasting and scale system optimisation.

Anjali Awasthi is an Associate Professor at the Concordia Institute for Information Systems Engineering (CIISE) in Concordia University, Montreal. Her areas of research are modelling and simulation, data analysis, city logistics and sustainable transportation. She is the author of several journal and conference papers on these topics.

1 Introduction

Carsharing is an alternative to private car usage. A carsharing company usually consists of small to medium sized fleets of vehicles available for members at specific stations spread over the city. The vehicles are reserved and used individually as required. While users are only responsible for the time used and travelled distance and in some cases the fuel, the company pays for the vehicle's expenses, repairs, gas and insurance (Shaheen et al., 2005; Costain et al., 2012). A study by Shaheen and Cohen (2016) shows that around 4.8 million people are sharing more than 104,000 vehicles all over the world. It is expected that there will be over 12 million of carsharing users by 2020 which implies the rapid growth of this business in near future in recent years.

Forecasting of carsharing demand has been addressed by only few researchers in recent years. Jorge and Correia (2013) report that most studies do not, or only insufficiently, address the supply side of carsharing, which directly impacts demand and how to find the right balance between the two. Furthermore, demand models are too context-specific and not widely applicable to other providers and locations. Thus, there is a need for more realistic models, especially for one-way carsharing.

There have been multiple studies on the topic of carsharing in recent years. Weikl and Bogenberger (2013, 2015) study the relocation issues associated with carsharing fleet. El Fassi et al. (2012) evaluate carsharing network's growth strategies through discrete event simulation. Millard-Ball et al. (2005) and Willing et al. (2016) assess its integration with other modes of transport. Martin and Shaheen (2011) assess its potential in providing relief to overburdened urban mobility systems. The environmental impacts of carsharing have been addressed in Awasthi and Omrani (2009), Awasthi and Chauhan (2011), Awasthi et al. (2011) and Firnkorn and Müller (2011). Kopp et al. (2015) study the user behaviour of carsharing customers. Bardhi and Eckhardt (2012) study its role as a prime example of the sharing economy. Awasthi et al. (2007, 2008, 2009) investigate location planning of carsharing stations through multicriteria approaches. Shaheen et al. (2005) study the market growth and future potential of carsharing in North America.

Accurate demand forecasting has a direct impact on the revenues from carsharing industry (Jorge and Correia, 2013). However, it is not easy to get realistic forecasts due to variability in different factors such as number of trips in different time periods, availability of the vehicles and the types of services offered by the carsharing company. Therefore, it is one of the main challenges faced by carsharing industries. An accurate estimation of the demand in short and long term can help planning of operations in advance and ensure that the stations and the vehicles are utilised to their full capacity, thereby fulfilling supply demand inbalance, improving customer satisfaction and retaining their loyalty.

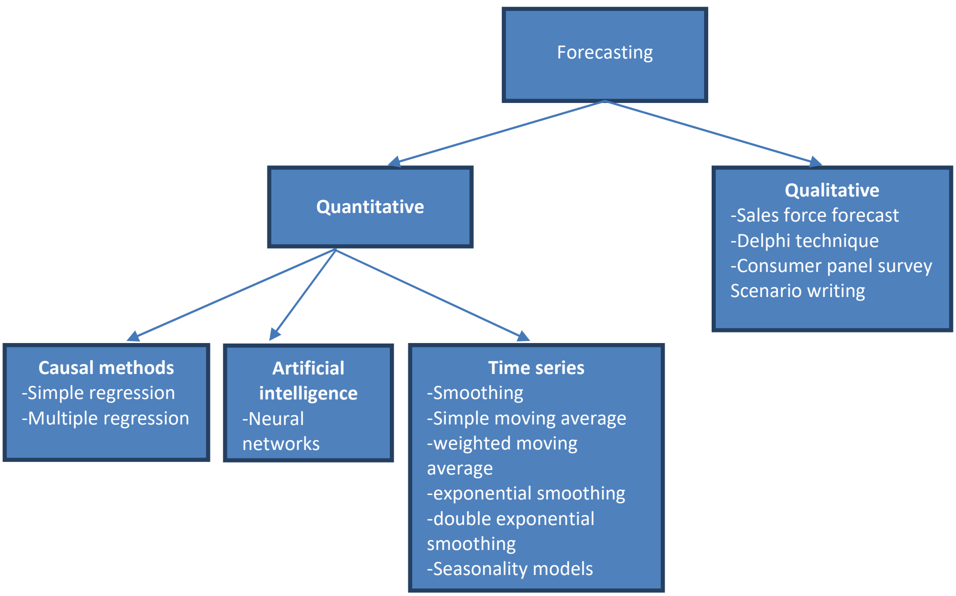
In this paper, we are addressing the problem of customer demand forecasting for improving carsharing operations. Three categories of methods namely causal (regression forecast, regression forecast with seasonality adjustments), time series (exponential smoothing, moving average) and neural networks are evaluated for accurately forecasting customer demand. An application of the proposed methods on demand data from a carsharing organisation called Communauto is provided.

The rest of the paper is organised as follows. In Section 2, we present the literature review. Section 3 presents the solution approach and the methods used for forecasting customer demand. Section 4 presents the application of the proposed methods on carsharing demand data of Communauto and compares the achieved results. Finally, we provide the conclusions and future works in Section 5.

2 Literature review

Figure 1 presents the commonly used methods for forecasting (Emam and Min, 2009; Rabbani et al., 2012; Viswanathan et al., 2008; Mady and Khalil, 2006). There are two broad categories: qualitative and quantitative. Qualitative methods are based on experience, judgment and expert knowledge, while the quantitative methods rely on data, statistical and learning based approaches (Karlaftis and Vlahogianni, 2011).

Figure 1 Forecasting methods (see online version for colours)



Lee and Park (2012) forecasted electric vehicle sharing demand using artificial neural network approach. Schmöller et al. (2015) used regression analysis to study the influence of multiple factors on customer demand in free-floating carsharing systems (FFCS). Muller and Bogenberger (2015) applied exponential smoothing with Holt-Winters filter (HWF) and a seasonal ARIMA model to predict the number of bookings in a FFCS. Koegst et al. (2008) studied the impact of demographic changes, socio-economic variability, population and age on customer demand using a multi regression model.

Lorimier and El-Geneidy (2011) proposed regression models for investigating the factors affecting the vehicle usage and availability in Communauto Inc. and concluded that the size of a carsharing station and the customer demand variability have the most impact. Ciari et al. (2014) and Balac et al. (2015) propose a multi-agent simulation tool based on MATSim software to model the spatial and temporal variations in carsharing demand in Zurich. Schmöller et al. (2015) analyse free floating carsharing dynamics in Munich and Berlin and show that demand concentrates around temporal peaks and spatial 'hot spots'. They found that demographic factors influence long-term demand patterns while weather conditions influence short-term demand dynamics. Willing et al. (2017) extended this work by adding the temporal dimension to the estimation of demand and tested in what ways points of interest influence them. Trasarti et al. (2011) proposed a methodology for extracting mobility profiles of individuals from raw digital traces (in particular, GPS traces) and studied criteria to match individuals based on profiles. Correia and Viegas (2011) proposed a stated preference survey to assess value enhancement possibilities in carpooling and carpool clubs in Lisbon, Portugal. Boldrini et al. (2016) study spatio temporal characteristics of carsharing demand and usage patterns in a large station-based car sharing system in France.

Factors affecting carsharing usage have been investigated by Kortum et al. (2016). Their study on 34 cities across nine countries confirms that the household size and residential density are main contributing factors for carsharing use. Schaefers (2013) identified four main motivations for people to turn in to a carsharing user: value-seeking, convenience, lifestyle and environmental motives. Awasthi et al. (2009) analysed carsharing market in La Rochelle, France and found low trip cost, easy access to carsharing station and vehicle size as important factors. In similar research, Ruhrort et al. (2014) studied electric vehicle carsharing (i.e., pre-booking and free-floating system) in Germany and found out that two groups of users are the key groups for electric carsharing. First group are users who have strong environmental orientations and the second group are people who are bike and/or public transport lover.

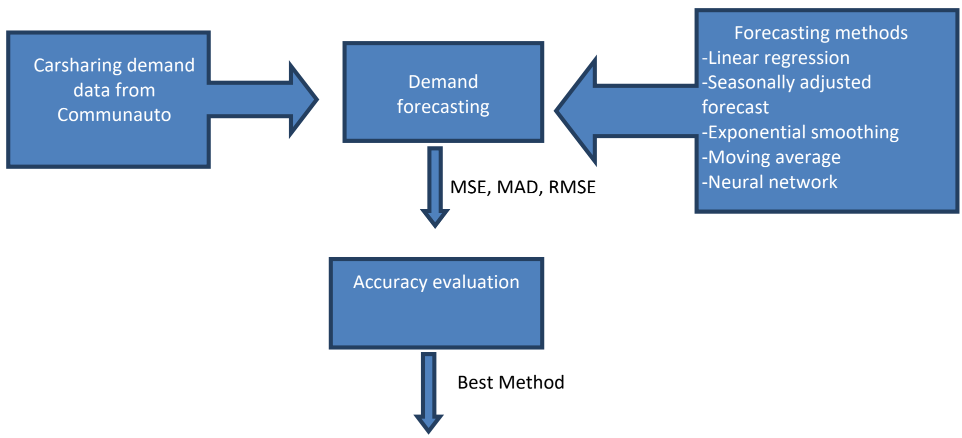
Performing comparison studies for forecasting methods is a very important step in order to find the most efficient ways to forecast customer demands. Without that, optimising the number of infrastructures, facilities and associated cost is not easily achievable and far from execution. It can be seen from above there is a lack of studies in this area. This is the challenge we are addressing in this paper.

3 Solution approach/methodology

The proposed solution approach involves quantitative forecasting methods namely causal (simple linear regression, seasonally adjusted forecast), time series (exponential smoothing, moving average) and neural networks for predicting carsharing customer demand at various locations based on historical data. We apply forecasting techniques, generate results, perform analysis and finally compare the results on a randomly generated data set for a time period between 1994 and 2011. The data set is generated based on real data obtained from a carsharing organisation (Communauto, Montreal) for the first nine months of the years 2011 and 2012. The objective of forecasting is to find the method with least error. The best forecasting approach can be used to get more realistic assessment, regarding demand in each month of the year, specifically for each station and in total. Figure 2 summarises the proposed methodology.

The forecasting methods used in our study are explained as follows.

Figure 2 Methodology (see online version for colours)



3.1 Simple linear regression

Regression analysis is used for investigating the relationship between a set of variables. If we consider y as a dependent variable and x as an independent variable, then the linear equation relating these variables is given by:

$$y = a + bX$$

where a is the intercept and b is the slope. The least squares method is used to estimate slope a and intercept b.

$$b = \frac{\sum_{i=1}^{n} X_i y_i \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}} \quad \text{and} \quad a = \overline{y} - b\overline{X}$$

3.2 Seasonally adjusted forecast

This method considers seasonality changes in forecasting to make more realistic estimation of the future. Many processes exhibit seasonal cycles such as agricultural production and consumer consumption like carsharing business. There is greater demand for renting cars in months leading up to Christmas and in spring when the weather is warm enough for short trips. Seasonal adjustment in regression forecasting follows the usual steps, but except the calculation of seasonal indices which is done as follows:

1 (actual/forecast) for each seasonal cycle for each occurrence

2 average the values calculated from step 1 for each of the seasons

3 Multiply the calculated seasonal index with the forecasted value from the simple regression method to get the seasonality adjusted forecast.

3.3 Exponential smoothing

Exponential smoothing is used for time series data, either to forecast or to produce smoothed data. There are three types of methods namely simple exponential smoothing, double exponential smoothing and triple exponential smoothing which are explained as follows.

3.3.1 Simple exponential smoothing

Simple exponential smoothing was proposed by Robert Goodell Brown in 1956. It is given by the formula:

$$F_{t+1} = F_t + \alpha\left(D_t - F_t\right)$$

where:

F_t forecast for the current period *t*

D_t actual demand for the current period *t*

α smoothing factor and $0 < \alpha < 1$.

3.3.2 Double exponential smoothing (Holt's model with trend)

Simple exponential smoothing does not work well when there is a trend in the data. Double exponential smoothing is used in such situation. The basic idea behind this method is to include a term for trends in the basic formula. This method, referred to as 'Holt's double exponential smoothing model', is given as follows:

$$T_{t+1} = \beta * \left(F_{t+1} - F_t\right) + (1-\beta) * T_t$$

where *F* comes from simple exponential smoothing and β is the trend factor.

3.3.3 Triple exponential smoothing (Winter's model with trend and seasonality)

When the data includes both seasonal changes as well as trends, triple exponential smoothing is used. Seasonality is in fact the tendency of time-series to show a behaviour that repeats every *L* period. This method calculates a trend line for the data and also includes seasonal indices. These indices weight the values in a trend line according to the place the time point falls within the length of the cycle.

$$L_t = \alpha\left(\frac{D_t}{S_{t-s}}\right) + (1-\alpha)\left(L_{t-1} + T_{t-1}\right) \qquad S_t = \gamma\left(\frac{D_t}{L_t}\right) + (1-\gamma)S_{t-s}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \qquad F_t + m = \left(L + mT_t\right) * S_{t+m-s}$$

3.4 Moving average

Moving average gives an idea of the recent trends in the data. It is calculated by averaging the values from the previous periods. Simple moving average and weighted moving average are two popular types of moving average methods. The weights in the weighted moving average method can be obtained using expert opinions, multicriteria decision methods such as AHP, Shannon's entropy algorithm, etc.

- Simple moving average: $F_{t+1} = \dfrac{D_t + D_{t-1} + \cdots + D_{t-n+1}}{n}$
- Weighted moving average: $F_{t+1} = \dfrac{W_t * D_t + W_{t-1} * D_{t-1} + \cdots + W_{t-n+1} * D_{t-n+1}}{W_t + W_{t-1} + \cdots + W_{t-n+1}}$

3.5 Artificial neural network

Neural networks are organised though layers including input layers, output layers and hidden layers. They are made up of nodes which are interconnected through an 'activation function'. Input layers present a pattern to the whole network, including the hidden layers. The main processing is done though a system of connections that are weighted in a basic initialisation. ANNs usually contain a form of learning rule, which is used to modify the weights of the links according to the patterns that the input layers present to the network. They learn through following their biological counterparts; which refers to a learning rule or a learning algorithm. Most neural networks are used to solve prediction problems. Its application in forecast various nonlinear time series has been reported in Hill et al. (1996), Tang et al. (1991), Hippert et al. (2001), and Ghiassi et al. (2005).

3.6 Performance evaluation

To evaluate the performance of forecasting methods, mean square error (MSE), mean absolute deviation (MAD) and root mean square error (RMSE) indicators are used. They are computed as follows:

$$MSE = \frac{\sum (E_t)^2}{n}$$

where (E_t) = Actual demand (Dt) – Forecast (Ft)

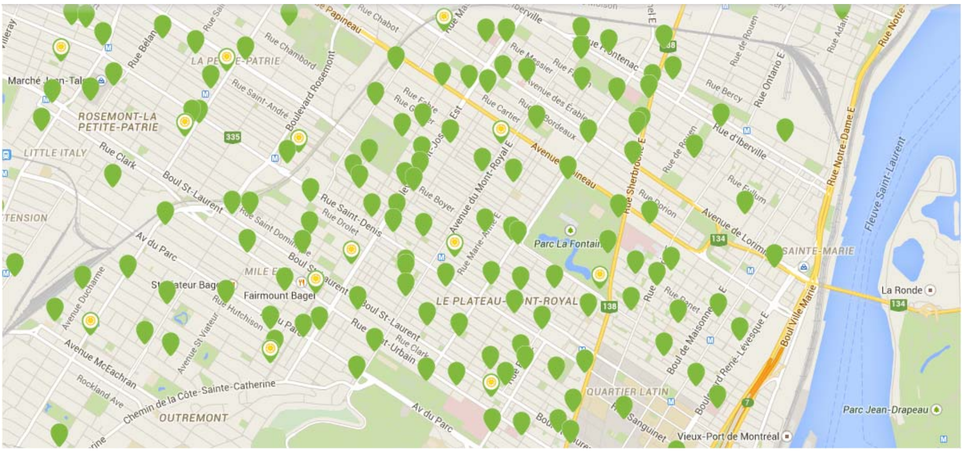
$$MAD = \frac{\sum |E_t|}{n}$$

$$RMSE = \sqrt{\frac{\sum (E_t)^2}{n}}$$

4 Numerical application/case study (performing the forecast)

Communauto Inc. is a carsharing company privately owned and founded in Quebec City in 1994. It was then merged with its competitor, Auto-com in 2000. Communauto operates in four areas of Quebec province namely Québec, Montréal, Gatineau and Sherbrooke. Its membership reached to 36,000 in May 2012 with 883 vehicles in use across 331 stations. Communauto is one of the pioneers in carsharing business, especially in the use of electric vehicles in the carsharing fleet, in Montreal, Canada. The map of its stations across Montreal is shown in Figure 3.

Figure 3 Network map of Communauto's stations across Montreal (see online version for colours)

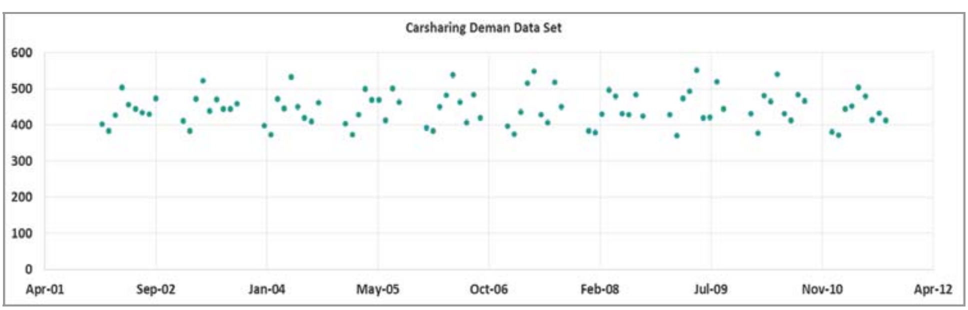


Source: Communauto (2014)

4.1 Input data

The customer demand for a specific station is randomly generated based on actual data from Communauto for the first nine months of the years between 1994 and 2011. The results can be seen in Table 1. Similar demand calculations can be done for other stations.

Figure 4 Demand data set trend between 1994 and 2011 (see online version for colours)



4.2 Regression forecast

Table 2 represents the regression line including the coefficients *a* and *b* which are the slope and intercept of the related regression line.

Table 1 Demand between 1994 and 2011

	1994	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	406	400	411	409	384	429	414	410	403	411	398	404	392	397	384	428	431	380
Feb	369	381	376	375	373	370	384	379	384	383	374	374	384	375	379	370	377	372
Mar	427	458	436	472	447	462	419	431	427	472	472	428	450	435	430	473	481	444
Apr	467	478	487	479	462	493	433	441	504	522	446	500	482	516	497	494	465	451
May	565	419	485	561	447	470	448	525	456	438	532	469	538	548	479	551	540	504
Jun	441	458	457	460	500	428	435	417	444	470	450	469	463	428	431	420	432	479
Jul	420	416	419	407	415	441	424	403	434	444	419	413	407	407	429	421	412	414
Aug	456	417	438	470	451	519	445	438	430	444	409	501	484	518	483	520	484	433
Sep	418	426	438	472	453	432	463	473	474	459	462	463	420	450	424	444	466	413

Table 2 Slope, intercept and the regression line

a	*b*
416.87	0.32
$y = 416.87 + 0.32X$	

Table 3 presents the results of linear regression for each of the nine months for the 18 years after 2011.

Table 3 Regression forecast results

Month	*Regression forecast*																	
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	417	420	423	426	429	432	434	437	440	443	446	449	452	454	457	460	463	466
Feb	418	420	423	426	429	432	435	438	440	443	446	449	452	455	458	461	463	466
Mar	418	421	424	426	429	432	435	438	441	444	447	449	452	455	458	461	464	467
Apr	418	421	424	427	430	432	435	438	441	444	447	450	453	455	458	461	464	467
May	418	421	424	427	430	433	436	439	441	444	447	450	453	456	459	462	464	467
Jun	419	422	425	427	430	433	436	439	442	445	447	450	453	456	459	462	465	468
Jul	419	422	425	428	431	433	436	439	442	445	448	451	454	456	459	462	465	468
Aug	419	422	425	428	431	434	437	440	442	445	448	451	454	454	460	462	465	468
Sep	420	423	425	428	431	434	437	440	443	446	448	451	454	457	460	463	466	469

The MSE, RMSE and MAD values can be seen in Table 4. It is clear that these values are relatively large.

Table 4 Performance evaluation of the regression forecast method

Month	*MSE*	*RMSE*	*MAD*
Jan	209.02	14.46	11.25
Feb	24.13	4.91	4.23
Mar	365.20	19.11	17.39
Apr	609.92	24.70	20.71
May	1,963.73	44.31	39.34
Jun	457.26	21.38	17.77
Jul	125.87	11.22	8.82
Aug	1,021.64	31.96	26.02
Sep	106.81	20.17	18.11
Final MSE, RMSE and MAD	*1,982.09*	*44.52*	*35.14*

4.3 Linear regression with seasonality adjustment (I)

Table 5 presents the results for linear regression with the seasonality adjustment.

Table 5 Seasonality adjusted forecast (I) results

Month	*Seasonal index*	*Seasonally adjusted forecast (II)*																	
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	0.92	383	386	388	391	394	396	399	402	404	407	410	412	415	417	420	423	425	428
Feb	0.85	356	359	361	364	366	368	371	373	376	378	381	383	386	388	391	393	395	398
Mar	1.01	424	427	429	432	435	438	441	444	447	450	453	456	459	461	464	467	470	473
Apr	1.08	453	456	459	462	465	468	471	474	478	481	484	487	490	493	496	499	502	506
May	1.13	471	475	478	481	484	487	491	494	497	500	504	507	510	513	517	520	523	526
Jun	1.01	425	428	431	434	437	439	442	445	448	451	454	457	460	463	466	469	471	474
Jul	0.95	397	399	402	405	407	410	413	416	418	421	424	426	429	432	435	437	440	443
Aug	1.04	438	441	444	447	450	453	456	459	462	465	468	471	474	477	480	483	486	489
Sep	1.01	423	426	429	432	435	438	440	443	446	449	452	455	458	461	464	467	469	472

The MSE, RMSE and MAD values can be seen in Table 6. It can be seen that they are smaller than the regular regression approach (Table 4) which shows better performance of this method over the former.

Table 6 Performance evaluation of the seasonality adjusted forecast (I)

Seasonally adjusted forecast (I)			
Month	*MSE*	*RMSE*	*MAD*
Jan	209.02	14.46	11.25
Feb	24.13	4.91	4.23
Mar	365.20	19.11	17.39
Apr	609.92	24.70	20.71
May	1,963.73	44.31	39.34
Jun	457.26	21.38	17.77
Jul	125.87	11.22	8.82
Aug	1,021.64	31.96	26.02
Sep	106.81	20.17	18.11
Final MSE, RMSE and MAD	*746.42*	*27.32*	*27.89*

4.4 Linear regression with seasonality adjustment (II)

We propose a second method for calculating seasonal indices. In this method, we calculate the regression line for each month separately and generate the regression line for each of the months, individually. The calculated slopes and intercepts for each month are illustrated in the Table 7.

Table 7 Slopes and intercepts of the seasonality adjusted forecast (II)

Month	*b*	*a*
Jan	–0.26	407.52
Feb	0.01	376.54
Mar	0.70	441.39
Apr	0.62	472.85
May	2.30	476.78
Jun	–0.67	455.39
Jul	–0.21	421.14
Aug	2.46	439.92
Sep	–0.10	448.14

The final forecasted values are illustrated in Table 8.

Table 8 Seasonality adjusted forecast (II) results

Month	*Seasonally adjusted forecast (II)*																	
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	407	407	407	406	406	406	406	405	405	405	405	404	404	404	404	403	403	403
Feb	377	377	377	377	377	377	377	377	377	377	377	377	377	377	377	377	377	377
Mar	442	443	443	444	445	446	446	447	448	448	449	450	450	451	452	453	453	454
Apr	473	474	475	475	476	477	477	478	478	479	480	480	481	482	482	483	483	484
May	479	481	484	486	488	491	493	495	497	500	502	504	507	509	511	514	516	518
Jun	455	454	453	453	452	451	451	450	449	449	448	447	447	446	445	445	444	443
Jul	421	421	421	420	420	420	420	419	419	419	419	419	418	418	418	418	418	417
Aug	442	445	447	450	452	455	457	460	462	465	467	469	472	474	477	479	482	484
Sep	448	448	448	448	448	448	447	447	447	447	447	447	447	447	447	447	446	446

Table 9 Performance evaluation of the seasonality adjusted forecast (II)

Seasonally adjusted forecast (II)			
Month	*MSE*	*RMSE*	*MAD*
Jan	209.02	14.46	11.25
Feb	24.13	4.91	4.23
Mar	365.20	19.11	17.39
Apr	609.92	24.70	20.71
May	1,963.73	44.31	39.34
Jun	457.26	21.38	17.77
Jul	125.87	11.22	8.82
Aug	1,021.64	31.96	26.02
Sep	106.81	20.17	18.11
Final MSE, RMSE and MAD	*575.95*	*24.00*	*18.18*

Table 9 presents the MSE, RMSE and MAD values. It can be seen that these values are significantly less than the original seasonality adjustment method (Table 6) which shows the better performance of this method over the previous one.

4.5 Winter's model

4.5.1 Before optimisation

Based on an initial assumption for alpha, beta and gamma ($\alpha = 0.3$, $\beta = 0.3$, $\gamma = 0.3$), Table 10 shows the forecasted values from the Winter's model for the next 17 years. The first year data has been used as the initial data for the Winter's model.

Table 10 Winter's model, initial results

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
Jan	406	387	411	434	404	418	395	400	406	411	398	430	399	421	390	419	413
Feb	367	363	376	381	388	382	369	367	375	373	370	389	373	387	388	396	367
Mar	430	429	436	444	456	461	442	435	434	430	438	456	436	450	449	466	445
Apr	483	468	488	469	487	471	468	456	485	476	449	485	462	483	512	516	470
May	587	542	536	515	534	477	487	517	522	457	480	484	492	497	515	515	491
Jun	410	449	483	418	473	426	475	452	460	466	442	475	471	428	471	445	412
Jul	400	425	447	407	405	391	429	428	445	442	431	441	423	396	431	423	419
Aug	436	452	461	445	450	437	454	470	477	459	444	463	462	463	484	482	489
Sep	390	421	444	431	462	405	424	447	472	450	492	477	473	444	461	434	431

4.5.2 After optimisation

Based on MSE, we optimised alpha, beta and gamma values (α = 0.00, β = 0.51, γ = 0.26). The results are shown in Table 11.

Table 11 Winter's model, optimised results

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	406	404	406	407	401	408	410	410	408	409	406	405	402	401	396	405	411	
Feb	369	372	373	374	373	373	376	376	378	380	378	377	379	378	378	376	376	
Mar	427	435	435	445	445	450	442	439	436	445	452	446	447	444	440	449	457	
Apr	467	470	474	476	472	478	466	459	471	484	474	481	481	490	492	493	485	
May	565	527	516	528	507	497	484	495	485	472	488	483	497	511	502	515	522	
Jun	441	445	448	451	464	455	450	441	442	449	449	455	457	449	444	438	436	
Jul	420	419	419	416	416	422	423	418	422	428	425	422	418	415	419	419	417	
Aug	456	446	444	451	451	469	462	456	449	448	438	454	462	477	478	489	488	
Sep	418	420	425	437	441	439	445	452	458	458	459	460	450	450	443	443	449	

Table 12 Winter's model, Performance evaluation for initial model

Initial results			
Alpha	0.3	MSE	1,167.84
Beta	0.3	RMSE	34.17
Gamma	0.3	MAD	25.66

Table 13 Winter's model, Performance evaluation for optimised model

After optimisation			
Alpha	0.00	MSE	835.71
Beta	30.5	RMSE	28.91
Gamma	0.26	MAD	21.12

It can be seen that the MSE, RMSE and MAD values of the optimised model are significantly less, however still relatively large in comparison with seasonality adjustment methods.

4.6 Artificial neural network

For neural network, MATLAB toolbox was used which employs Levenberg-Marquardt algorithm as a learning rule. Different portions of the training data set as well as the testing and validation data set were defined. We used 90/10, 80/20, 70/30, 75/25, 60/40 and 50/50 randomly divided portions for training, testing and validation data sets respectively. Tables 14–25 show the results from the neural network method for different training, validation and testing ratios.

4.6.1 (90:10)

Tables 14–15 show the neural network results (forecasted y value) for 90:10 ratio (90% training, 5% validation and 5% testing). The final MSE is presented in Table 15. It can be seen that the resulting MSE is fairly small which shows the strength of neural network over other forecasting methods.

Table 14 Neural network 90:10 results

90:10	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	395	403	403	396	402	402	405	395	395	403	399	404	403	400	402	405	401	402
Feb	379	377	376	378	378	378	376	379	379	376	379	377	378	379	378	376	376	378
Mar	446	445	460	452	440	440	454	445	446	461	439	449	443	438	440	454	458	440
Apr	492	480	461	482	488	488	464	493	493	459	494	473	483	494	488	464	465	488
May	437	507	523	456	493	493	538	435	435	527	465	521	502	474	493	538	508	493
Jun	471	434	461	474	431	431	441	470	470	460	441	437	433	434	431	441	464	431
Jul	428	413	412	425	415	415	408	428	428	411	421	411	414	419	415	408	415	415
Aug	438	474	439	432	480	480	463	439	439	440	471	469	476	478	480	462	438	480
Sep	445	441	460	452	436	436	452	445	445	460	436	446	440	434	436	452	458	436

Table 15 Performance evaluation of neural network 90:10

	Target value	*MSE*	*Final MSE*
Training	153	418.14	442.32
Validation	9	446.63	
Testing	9	849.13	

4.6.2 (80:20)

Tables 16–17 present the neural network results for 80:20 ratio. It can be seen that the MSE results for 80:20 ratio is slightly bigger than the 90:10 ratio. However it is still not so large.

Table 16 Neural network 80:20 results

80:20	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	398	407	410	399	401	401	409	406	406	410	401	401	406	405	406	401	401	406
Feb	379	378	377	378	378	378	377	378	378	377	378	378	378	378	378	378	378	378
Mar	447	450	449	447	446	446	448	450	450	449	446	446	450	449	450	446	446	450
Apr	494	491	460	482	459	459	460	496	496	460	459	459	496	496	496	459	459	495
May	436	485	538	456	495	495	534	478	477	539	495	495	477	470	477	495	495	479
Jun	471	443	432	467	459	459	435	445	445	432	459	459	445	450	445	459	459	445
Jul	423	426	417	419	412	412	416	428	428	417	412	412	428	427	428	412	412	428
Aug	464	452	447	462	459	459	448	452	452	447	459	459	452	454	452	459	459	452
Sep	431	452	466	436	447	447	464	450	450	466	447	447	449	446	449	447	447	450

Table 17 Performance evaluation of neural network 80:20

	Target value	*MSE*	*Final MSE*
Training	137	422.62	552.64
Validation	17	1,082.13	
Testing	17	1,071.00	

4.6.3 (75:25)

Tables 18–19 present the neural network results for 75:25 ratio. For the ratio of 75:25 (75% training data set, 15% validation data set and 10% testing data set), the resulting MSE is slightly more than 80:20 and significantly more than 90:10 ratio.

Table 18 Neural network 75:25 results

75:25	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	394	399	410	385	385	412	385	385	413	413	410	385	413	412	413	413	387	385
Feb	378	378	377	378	378	377	378	378	377	377	377	378	377	377	377	377	378	378
Mar	443	447	458	432	432	460	432	433	461	461	458	432	461	459	461	461	434	432
Apr	476	475	470	478	478	469	478	478	469	469	471	478	469	469	469	469	478	478
May	477	477	494	473	473	496	473	471	496	496	491	473	496	495	496	496	474	473
Jun	438	441	446	434	434	447	434	434	448	448	446	434	448	447	448	448	435	434
Jul	419	420	418	419	419	418	419	419	418	418	418	419	418	418	418	418	419	419
Aug	456	456	465	454	454	466	454	454	466	466	464	454	466	466	466	466	455	454
Sep	431	432	439	427	427	440	427	427	441	441	439	427	441	440	441	441	428	427

Table 19 Performance evaluation of neural network 75:25

	Target value	*MSE*	*Final MSE*
Training	128	670.29	679.94
Validation	26	673.36	
Testing	17	762.62	

4.6.4 (70:30)

Tables 20–21 present the results for 70:30 ratio (70% training, 15% validation and 15% testing). It can be seen that the final MSE in this ratio is almost the same as 75:25 ratio.

Table 20 Neural network 70:30 results

70:30	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	402	402	402	395	393	402	400	396	398	401	394	396	402	399	401	402	402	402
Feb	374	374	374	380	381	374	375	378	377	375	380	378	374	376	374	374	375	374
Mar	458	458	458	446	442	459	454	446	450	456	445	447	457	452	456	460	459	457
Apr	467	466	466	504	517	466	479	502	491	474	507	500	470	484	472	463	466	469
May	495	494	491	436	442	492	496	478	489	493	442	474	501	492	501	480	463	490
Jun	443	443	444	476	477	444	446	459	452	446	474	460	442	449	442	448	455	446
Jul	419	419	419	420	419	419	418	418	418	418	419	418	418	418	418	420	421	419
Aug	451	451	451	468	469	451	453	460	456	452	467	461	450	454	451	452	456	452
Sep	436	436	436	457	462	436	440	451	446	439	458	451	436	443	437	436	439	437

Table 21 Performance evaluation of neural network 70:30

	Target value	*MSE*	*Final MSE*
Training	119	657.57	677.97
Validation	26	631.30	
Testing	26	817.99	

4.6.5 (60:40)

Tables 22–23 present the neural network forecasted results for 60:40 ratio (60% fir training, 20% validation and 20% testing). It can be seen that the MSE is slightly bigger than the 70:30 and 75:25 ratios.

Table 22 Neural network 60:40 results

60:40	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	402	402	408	391	402	397	397	408	397	397	402	397	391	397	397	397	391	402
Feb	377	377	378	376	377	377	377	378	377	377	377	377	376	377	377	377	376	377
Mar	444	444	460	433	444	449	449	460	449	449	444	449	433	449	449	449	433	444
Apr	478	478	465	479	478	466	466	465	466	466	478	466	479	466	466	466	479	478
May	470	470	506	493	470	530	530	506	530	530	470	530	493	530	530	530	493	470
Jun	440	440	437	443	440	441	441	437	441	441	440	441	443	441	441	441	443	440
Jul	416	416	415	415	416	414	414	415	414	414	416	414	415	414	414	414	415	416
Aug	466	466	442	473	466	449	449	442	449	449	466	449	473	449	449	449	473	466
Sep	441	441	453	447	441	459	459	453	459	459	441	459	447	459	459	459	447	441

Table 23 Performance evaluation of neural network 60:40

	Target value	*MSE*	*Final MSE*
Training	103	633.50	669.23
Validation	34	569.51	
Testing	34	877.20	

4.6.6 (50:50)

Tables 24–25 present the neural network forecasted results for 50:50 ratio (50% of training, 25% validation and 25% testing). For 50:50 ratio, the MSE value increases significantly.

Table 24 Neural network 50:50 results

50:50	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
Jan	398	399	398	398	407	399	407	407	407	407	407	403	399	398	399	399	398	399
Feb	375	375	375	375	377	375	376	376	376	377	377	376	376	375	375	376	375	376
Mar	452	451	455	452	448	448	448	448	448	447	441	451	443	452	445	444	453	443
Apr	480	481	477	480	491	484	491	491	491	492	499	485	490	480	488	489	479	490
May	459	463	442	458	509	481	507	507	507	513	551	479	513	457	502	507	453	512
Jun	469	468	480	470	480	457	481	481	481	477	456	480	440	470	445	443	472	440
Jul	414	414	415	414	414	413	414	414	414	413	411	414	410	414	411	411	415	411
Aug	428	432	405	426	478	456	475	475	475	483	532	445	497	425	483	490	420	496
Sep	438	439	436	437	464	441	463	463	463	464	469	452	445	437	444	445	437	445

Table 25 Performance evaluation of neural network 50:50

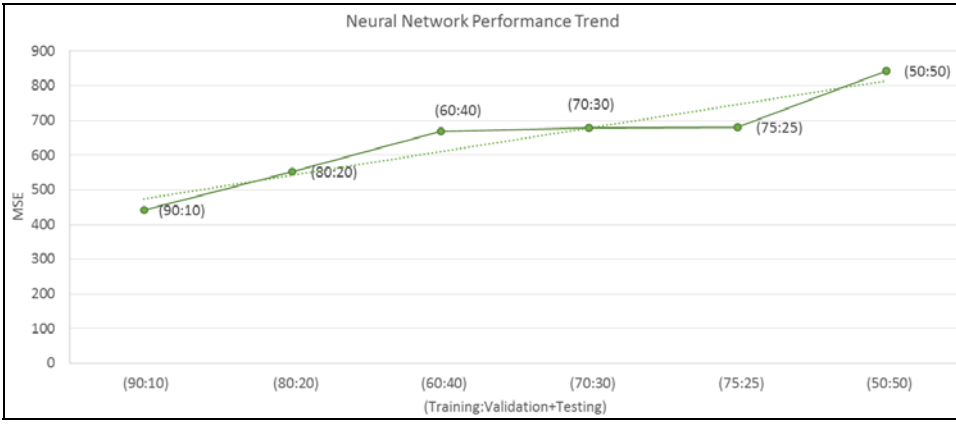
	Target value	*MSE*	*Final MSE*
Training	85	639.92	841.81
Validation	43	1,203.97	
Testing	43	878.73	

4.7 Performance comparison

4.7.1 Accuracy results for neural networks

Figure 5 depicts the change in MSE values with change in the ratios of training, validation and testing data sets for neural network. It can be seen that there is an obvious increasing trend when we decrease the ratio of the training data set in comparison to validation and testing. While 90:10 has the least MSE, 50:50 has the biggest MSE among all others.

Figure 5 Neural network performance evaluation (see online version for colours)



4.7.2 Accuracy results for all methods

Table 26 presents the accuracy results for all forecasting methods. It can be seen that the 90:10 ratio gives the lowest MSE and performs better than other ratios in neural network method. The proposed seasonally adjusted method (II) gives the second lowest MSE and RMSE in comparison to 80:20, 75:25, 70:30, 60:40 and 50:50 ratios and hence assumed as the second best method.

Table 26 Results evaluation (see online version for colours)

Forecasting method	*MSE*	*RMSE*	*MAD*
Neural network (90:10)	442.32	21.03	-
Neural network (80:20)	552.64	23.51	-
Regression forecast with seasonality adjustment (II)	575.95	24	-
Neural network (60:40)	669.23	25.87	-
Neural network (70:30)	677.97	26.04	-
Neural network (75:25)	679.94	26.08	-
Regression forecast with seasonality adjustment (I)	746.42	27.32	-
Winter's model (optimized version)	835.71	28.91	-
Neural network (50:50)	841.81	29.01	-
Linear regression	1,982.09	44.52	-

5 Conclusions and future works

Customer demand forecasting is vital to efficient management of fleet and improving carsharing operations. In this paper, we investigate three categories of forecasting methods namely causal (regression forecast, regression forecast with seasonality adjustments), time series (exponential smoothing, moving average) and neural networks for predicting carsharing customer demands for a carsharing organisation called

Communauto in Quebec, Montreal. The results of our study show that neural network is the best method in this prediction.

The proposed work can be of interest to companies interested in efficiently managing customers' demands and vehicle fleets at different times of the year. This will in turn have a direct impact on their customer satisfaction and business market performance. If electric vehicles are used in carsharing fleet, then it will be even more important since they require special infrastructures.

As a future work, we can consider the effects of other factors like customer behaviour and/or car rental costs as well as climate impacts on customer demand in carsharing industry. Also, integration with other methods like heuristics, meta-heuristics and NARX method in neural networks for the forecasting problem can bring more interesting results for solving these mathematical problems.

References

Awasthi, A. and Chauhan, S.S. (2011) 'Using AHP and Dempster-Shafer theory for evaluating sustainable transport solutions', *Environmental Modelling and Software*, Vol. 26, No. 6, pp.787–796.

Awasthi, A. and Omrani, H. (2009) 'A hybrid approach based on AHP and belief theory for evaluating sustainable transportation solutions', *International Journal of Global Environmental Issues*, Vol. 9, No. 3, pp.212–226.

Awasthi, A., Breuil, D., Chauhan, S.S., Parent, M. and Reveillere, T. (2007) 'A multicriteria decision making approach for carsharing stations selection', *Journal of Decision Systems*, Vol. 16, No. 1, pp.57–78.

Awasthi, A., Chauhan, S.S. and Breuil, D. (2009) 'Sustainable mobility solutions: a pre-implementation questionnaire study for carsharing', *International Journal of Services Sciences*, Vol. 2, Nos. 3/4, pp.242–264.

Awasthi, A., Chauhan, S.S. and Omrani, H. (2011) 'Application of fuzzy TOPSIS in evaluating sustainable transportation systems', *Expert systems with Applications*, Vol. 38, No. 10, pp.12270–12280.

Awasthi, A., Chauhan, S.S., Hurteau, X. and Breuil, D. (2008) 'An analytical hierarchical process-based decision-making approach for selecting car-sharing stations in medium size agglomerations', *International Journal of Information and Decision Sciences*, Vol. 1, No. 1, pp.66–97.

Balac, M., Ciari, F. and Axhausen, K.W. (2015) 'Carsharing demand estimation', *Transportation Research Record, J. Transp. Res. Board*, Vol. 2536, pp.10–18.

Bardhi, F. and Eckhardt, G.M. (2012) 'Access-based consumption: the case of car sharing', *J. Consum. Res.*, Vol. 39, No. 4, pp.881–898.

Boldrini, C., Bruno, R. and Conti, M. (2016) 'Characterising demand and usage patterns in a large station-based car sharing system', in *The 2nd IEEE INFOCOM Workshop on Smart Cities and Urban Computing*, pp.1–6.

Celsor, C. and Millard-Ball, A. (2007) 'Where does carsharing work?: using geographic information systems to assess market potential', *Transportation Research Record, J. Transp. Res. Board*, pp.61–69.

Ciari, F., Bock, B. and Balmer, M. (2014) 'Modeling station-based and free-floating carsharing demand', *Transportation Research Record, J. Transp. Res. Board*, Vol. 2416, pp.37–47.

Communauto (2014) [online] http://www.communauto.com/ (accessed 14 February 2014).

Correia, G. and Viegas, J. (2011) 'Carpooling and carpool clubs: clarifying concepts and assessing value enhancement possibilities through a stated preference web survey in Lisbon, Portugal', *Transportation Research A*, Vol. 45, No. 2, pp.81–90.

Costain, C., Ardron, C. and Habib, K.N. (2012) 'Synopsis of users' behaviour of a carsharing program: a case study in Toronto', *Transportation Research A*, Vol. 46, No. 3, pp.421–434.

El Fassi, A., Awasthi, A. and Viviani, M. (2012) 'Evaluation of carsharing network's growth strategies through discrete event simulation', *Expert Systems with Applications*, Vol. 39, No. 8, pp.6692–6705.

Emam, A. and Min, H. (2009) 'The artificial neural network for forecasting foreign exchange rates', *Int. J. of Services and Operations Management*, Vol. 5, No. 6, pp.740–757.

Firnkorn, J. and Müller, M. (2011) 'What will be the environmental effects of new free floating car-sharing systems? The case of car2go in Ulm', *Ecol. Econ.*, Vol. 70, No. 8, pp.1519–1528.

Ghiassi, M., Saidane, H. and Zimbra, D.K. (2005) 'A dynamic artificial neural network model for forecasting time series events', *Int. J. Forecasting*, Vol. 21, No. 2, pp.341–362.

Hill, T., O'Connor, M. and Remus, W. (1996) 'Neural network models for time series forecasts', *Manag. Sci.*, Vol. 42, No. 7, pp.1082–1092.

Hippert, H.S., Pedreira, C.E. and Souza, R.C. (2001) 'Neural networks for short-term load forecasting: a review and evaluation', *IEEE Trans Power System*, Vol. 16, No. 1, pp.44–55.

Jorge, D. and Correia, G. (2013) 'Carsharing systems demand estimation and defined operations: a literature review', *Eur. J. Transp. Infrastruct. Res.*, Vol. 13, No. 3, pp.201–220.

Jorge, D. and Homem de Almeida Correia, G. (2013) 'Carsharing systems demand estimation and defined operations: a literature review', *European Journal of Transport and Infrastructure Research*, Vol. 13, No. 3, pp.201–220

Karlaftis, M.G. and Vlahogianni, E. (2011) 'Statistical methods versus neural networks in transportation research: differences, similarities and some insights', *Transportation Research C*, Vol. 19, No. 3, pp.387–399.

Koegst, T., Tranckner, J., Franz, T. and Krebs, P. (2008) 'Multi-regression analysis in forecasting water demand based on population age structure', *11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK.

Kopp, J., Gerike, R. and Axhausen, K.W. (2015) 'Do sharing people behave differently?: an empirical evaluation of the distinctive mobility patterns of free-floating car-sharing members', *Transportation*, Vol. 42, No. 3, pp.449–469.

Kortum, K., Schonduwe, R., Stolte, B. and Bock, B. (2016) 'Free-floating carsharing: city specific growth rates and success factors', *Transportation Research Procedia*, Vol. 19, pp.328–340.

Lee, J. and Park, G.L. (2012) 'Demand forecast for electric vehicle sharing systems using movement history archive', in Kim, T., Ramos, C., Abawajy, J., Kang, B.H., Ślęzak, D. and Adeli, H. (Eds.): *Computer Applications for Modeling, Simulation and Automobile, Communications in Computer and Information Science*, Vol. 341, Springer, Berlin, Heidelberg.

Lorimier, A. and El-Geneidy, A. (2011) 'Understanding the factors affecting vehicle usage and availability in carsharing networks: a case study of Communauto carsharing system from Montréal, Canada', *International Journal of Sustainable Transportation*, Vol. 7, No. 1, pp.35–51.

Mady, M.T. and Khalil, O. (2006) 'IT adoption and manufacturing performance in Kuwaiti industrial corporations', *Int. J. of Services and Operations Management*, Vol. 2, No. 1, pp.60–77.

Martin, E.W. and Shaheen, S.A. (2011) 'Greenhouse gas emission impacts of carsharing in North America', *IEEE Trans. Intell. Transp. Syst.*, Vol.12, No. 4, pp.1074–1086.

Millard-Ball, A., Murray, G., Schure, J., Fox, C. and Burkhardt, J. (2005) 'Car-sharing: where and how it succeeds', *TCRP Report*, Vol. 108, Transportation Research Board, Washington, D.C.

Muller, J. and Bogenberger, K. (2015) 'Time series analysis of booking data of a free floating carsharing system in Berlin', *Transportation Research Procedia*, Vol. 10, pp.345–354.

Rabbani, M., Ghoreyshi, S.M., Rafiei, H. and Ghazanfari, M. (2012) 'Energy consumption forecasting using a bi-objective fuzzy linear regression model', *Int. J. of Services and Operations Management*, Vol. 13, No. 1, pp.1–18.

Ruhrort, L., Steiner, J., Graff, A., Hinkeldein, D. and Hoffmann, C. (2014) 'Carsharing with electric vehicles in the context of users' mobility needs – results from user-centred research from the BeMobility field trial (Berlin)', *Int. J. Automotive Technology and Management*, Vol. 14, Nos. 3/4, pp.286–305.

Schaefers, T. (2013) 'Exploring carsharing usage motives: a hierarchical means-end chain analysis', *Transportation Research Part A: Policy and Practice*, Vol. 47, pp.69–77.

Schmöller, S., Weikl, S., Muller, J. and Bogenberger, K. (2015) 'Empirical analysis of free-floating carsharing usage: the munich and Berlin case', *Transportation Research Part C: Emerging Technologies*, July, Vol. 56, pp.34–51.

Shaheen, S. and Cohen, A. (2016) *Carsharing Market Overview, Analysis and Trends, Innovative Mobility Industry Outlook* [online] http://innovativemobility.org/wp-content/uploads/2016/02/Innovative-Mobility-Industry-Outlook_World-2016-Final.pdf (accessed 31 October 2017).

Shaheen, S., Cohen, A. and Roberts. J.D. (2005) 'Carsharing in North America: market growth, current developments and future potential', *Transportation Research Board*.

Tang, Z., Almeida, C. and Fishwick, P.A. (1991) 'Time series forecasting using neural networks vs. Box-Jenkins methodology', *Simulation*, Vol. 57, No. 5, pp.303–310.

Trasarti, R., Pinelli, F., Nanni, M. and Giannotti, F. (2011) 'Mining mobility user profiles for carpooling', in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.1190–1198.

Viswanathan, S., Widiarta, H. and Piplani, R. (2008) 'Evaluation of hierarchical forecasting for substitutable products', *Int. J. of Services and Operations Management*, Vol. 4, No. 3, pp.277–295.

Weikl, S. and Bogenberger, K. (2013) 'Relocation strategies and algorithms for free-floating car sharing systems', *IEEE Intell. Transp. Syst. Mag.*, Vol. 5, No. 4, pp.100–111.

Weikl, S. and Bogenberger, K. (2015) 'A practice-ready relocation model for free-floating carsharing systems with electric vehicles – mesoscopic approach and field trial results', *Trans. Res. C Emerg. Technol.*, Vol. 57, pp.206–223

Willing, C., Gust, G., Brandt, G., Schmidt, S. and Neumann D. (2016) 'Enhancing municipal analytics capabilities to enable sustainable urban transportation', *Proceedings of the 24th European Conference on Information Systems (ECIS).*

Willing, C., Klemmer, K., Brandt, T. and Neumann, D. (2017) 'Moving in time and space – location intelligence for carsharing decision support', *Decision Support Systems*, Vol. 99, pp.75–85.