DATA VISUALIZATION TO FORECAST WITHDRAWALS IN A BIKE-SHARING SYSTEM: MEXICO CITY CASE

David Edgardo Castillo-Rodríguez[(a)], Luis Antonio Moncayo-Martínez[(a)]

[(a)]Department of Industrial Engineering and Operations, ITAM, Río Hondo 1,
Ciudad de México, 01080, México.

david.castillo@ITAM.mx, luis.moncayo@itam.mx

ABSTRACT

In this paper we predict the aggregated withdrawals of Mexico bike-sharing system, called Ecobici, for a given day. The principal problem adressed was the lack of data available to understand certain behavior related to the system demand at some time of the day. However, with the information available in the Ecobici website we were capable to adjust a time series model that help us forecast the aggregated withdrawals in the short time and midel tiem, this can be used to estimate the overall demand for a given hour in the day in order to identify the time of the day when the demand of bicycles is greater than the availability. The principal motivation of our analysis is to be capable to predict the aggregated demand of bicycles.

Keywords: Ecobici system, aggregated withdrawals, time series.

1. INTRODUCTION

Ecobici is a mobility system of the México´s city that started operations in February of 2010. The Ecobici system consists that through a membership you can use a bike offered by the service and get around over México city. Once you take a bike, you can use it for at most 45 minutes (if not, you must pay an extra charge), and at the end of your travel you must put down the bike in some station. The service and the number of users continue growing, since it started operations, because it has provided an alternative way to move around Mexico city and thus avoid the delay that happens in other mobility systems, such as traffic during crowed hours, for example. As a mobility alternative and the upgrowing number of users have occasioned that in certain hours of the day the demand of bikes is greater that the number of bikes available in some stations or that some stations are saturated and a user can´t put down the bike in that stations, causing the frustration of some users and some extra costs for relocating bikes from saturated stations to the empty ones.

As said above, it is important that the Ecobici system known with anticipation the expected demand of bikes through the day, and thus, make strategies that can afford the demand and shortage of the stations. This is the reason why in this paper we propose a model that help us to forecast the Ecobici aggregated demand for all the stations in one specific day. In section 2 we talk about the literature related to the topic as well as the main approaches proposed to predict demand forecast. In section 3 we talk about the available data and show some figures that help us to understand in deeper the problem. In section 4 we talk about the obtained time series model and present some results. Finally, in section 5 we talk about some conclusions and future work.

2. LITERATURE REVIEW

The study of the bike-sharing systems is growing in popularity around the world (Fishman, Washington and Haworth 2013). In the literature, there are four issues when the researchers and practitioners study these systems (Schuijbroek, Hampshire and Hoeve 2017); those are, the design of the systems, demand analysis, service level analysis, and rebalancing operations.

Most of the studies related to bike-sharing systems are about rebalancing operations. Basically, there are two approaches to rebalance the systems: a) incentives to the users, e.g. monetary (Pfrommer, Warrington, Schildbach and Morari 2014) and b) trucks that move bikes from full stations to empty ones (Schuijbroek, Hampshire and Hoeve 2017).

This work is related to the forecast of withdrawals in the México City bike-sharing systems called Ecobici. Information about demand at any station is input to balance the system. In recent publications, such as Moncayo (2020), and Moncayo and Ramirez (2016) the demand patterns have been studied and some managerial insights have been proposed. Even though, there is not a published study that proposes to forecast the demand in the Ecobici system.

Researchers and practitioners have proposed to predict demand forecast using assumptions about city orography, economic factors, infrastructure, and weather (Faghih, Hampshire, Marla and Eluru 2017).

So as to forecast the demand, the managers want either to know the level of future demand or to understand the behavior to make decision about managerial issues. In transport systems, the first studies to predict demand

were focused on the carsharing problems (Barth and Todd 1999) using a simulation model. In relation to the bike-sharing literature. In Froehlich, Neumann and Oliver (2009), authors tested four forecast methods to predict demand in the Bicing System (Barcelona): last view, historic mean, historic trends and Bayesian networks. The demand forecast was computes in interval times of 10, 20, 30, 60, 90 and 120 minutes. The same data set was used in Kaltenbrunner, Meza, Grivolla, Codina and Banchs (2010) to implement moving average techniques.
The hourly demand forecast of the Vélo'v system (Lyon) was computed in Borgnat, Abry, Flandrin, Robardet, Rouquier and Fleury (2011) using a two stage algorithm. In the first stage, it predicts the non-stationary amplitude during the day; then, the amplitude is modeled as a linear regression.
In this work, we focus on forecast by ARIMA models the demand of the withdrawls of the bikes on the Ecobici System. To the best of our knowledge, this forecast method has not been applied to the data set of this systems.

3. DATA AVAILABLE

The data used in this work is public and can be obtained in the ecobici's system website (see references). This data is reported monthly since February 2010 to June 2019 and consists of the following variables:

- *Genero_Usuario*: this is a nominal variable that takes the value M if the user is male and F if is female.
- *Edad_Usuario*: this is a numeric variable that reports the age of the user.
- *Bici*: this is a nominal variable that defines an id for the bikes.
- *Ciclo_Estacion_Retiro*: this is a nominal
- tion where a bike was taken.
- *Fecha_Retiro*: this is a nominal variable that reports the date when a bike was taken. This variable is reported as month/day/year.
- *Hora_Retiro*: this is a numeric variable that reports the time in which a bike was take. This variable is reported as hour:minute:second.
- *Ciclo_Estacion_Arribo*: this is a nominal variable that defines an id for the station where a bike was leaved.
- *Fecha_Arribo*: this is a nominal variable that reports the date when a bike was leaved. This variable is reported as month/day/year.
- *Hora_Arribo*: this is a numeric variable that reports the time in which a bike was leaved. This variable is reported as hour:minute:second.

3.1 Data visualization

In the Ecobici system website it is possible to consult at the moment the number of bikes per station, this information could be useless to the users because they don´t know if some minutes after their query they would have a chance to withdrawal a bike from a particular station. In Figure 1 we show the information, displayed on the website, of the number of bikes per station when a query is done. The green circles represent the stations with more than 5 bikes, the orange circles represent the stations with less than 5 bikes, the red circles represent the stations with no bikes, and finally the grey circles represent closed stations.

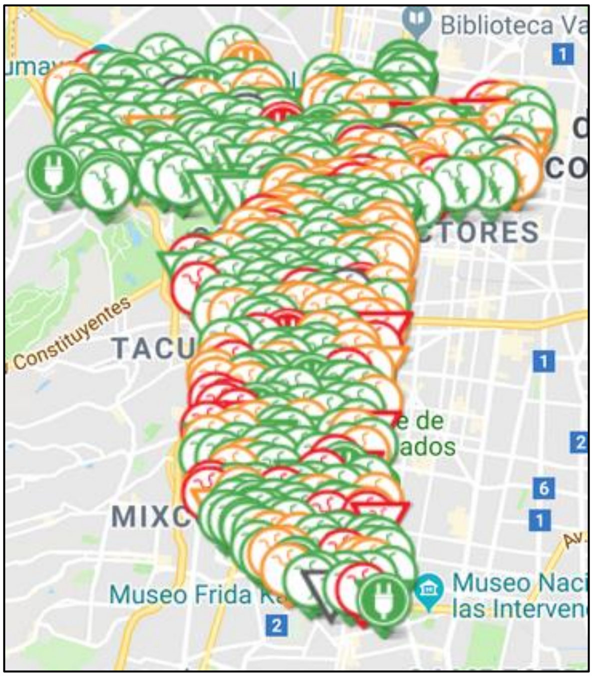

Figure 1: Bikes availability per station

So we start our analysis by monitoring the behavior of the aggregate number of withdrawals and arrivals per stations in a given month (January 2018) by time intervals of 15 minutes, we decided to use this time intervals length because the number of withdrawals and arrivals is almost the same in shorter time intervals in some specific hours as from 6:00 to 6:05 or 12:00 to 12:05, for example. In Figure 2 we illustrate the number of aggregate bikes withdrawals at time intervals 6:00 to 6:15, 9:00 to 9:15 and 18:00 to 18:15 respectively, we decided to start our analysis around this hours because it is the time were student and working activities starts and ends in México City. Also, in figure 3 we illustrate the number of aggregate bikes arrivals at time intervals 6:00 to 6:15, 9:00 to 9:15 and 18:00 to 18:15 respectively. In both cases, each circle represents a station, while each number represents the station id. The size of each circle indicates the number of aggregated withdrawals or arrivals, according each case, while the color is assigned according to the station id. As we expect, these figures give us evidence about the dependency of the demand of bikes and the hour of the day. For example, if we compare the time interval 6:00 to 6:15 with 9:00 to 9:15 we can see a lot of more mobility in the second one. Also, if we compare the retreats and arrivals for interval time interval from 9:00 to 9:15 we can observe that the circles of color yellow and pink are greater in Figure 3 than in Figure 2, in contrast, if now we compare time intervals from 18:00 to 18:15 we can observe that the yellow and pink circles are bigger in Figure 2 than in Figure 3, so we observe the inverse scenario. Also ,this example give us evidence to believe that the geographic zone can influence in the demand of bikes, for example, if some stations are closer to offices, the employees that use the Ecobici system would concentrate their arrivals in this stations by at the time they got to work, and at the end of the workday they are going to withdrawal bikes for a near station (in fact, these circles concerns to Reforma and the center zone, which are laboral areas).

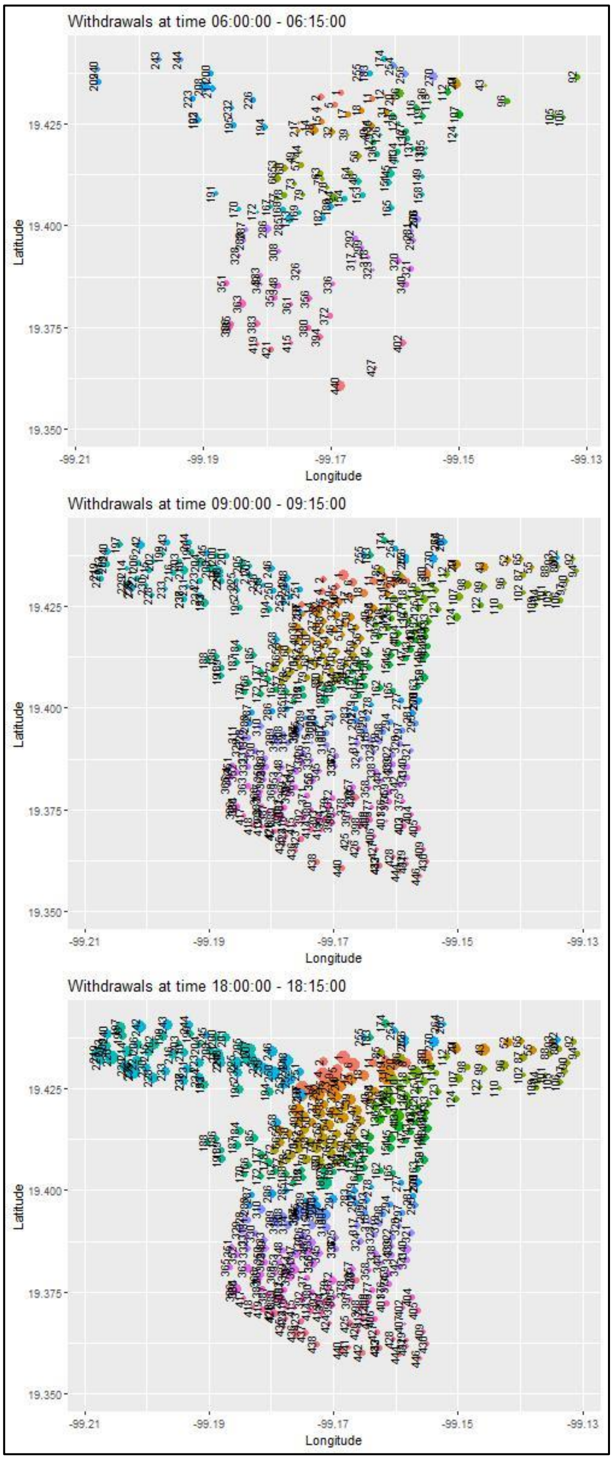

Figure 2: Aggregated withdrawals at January 2018

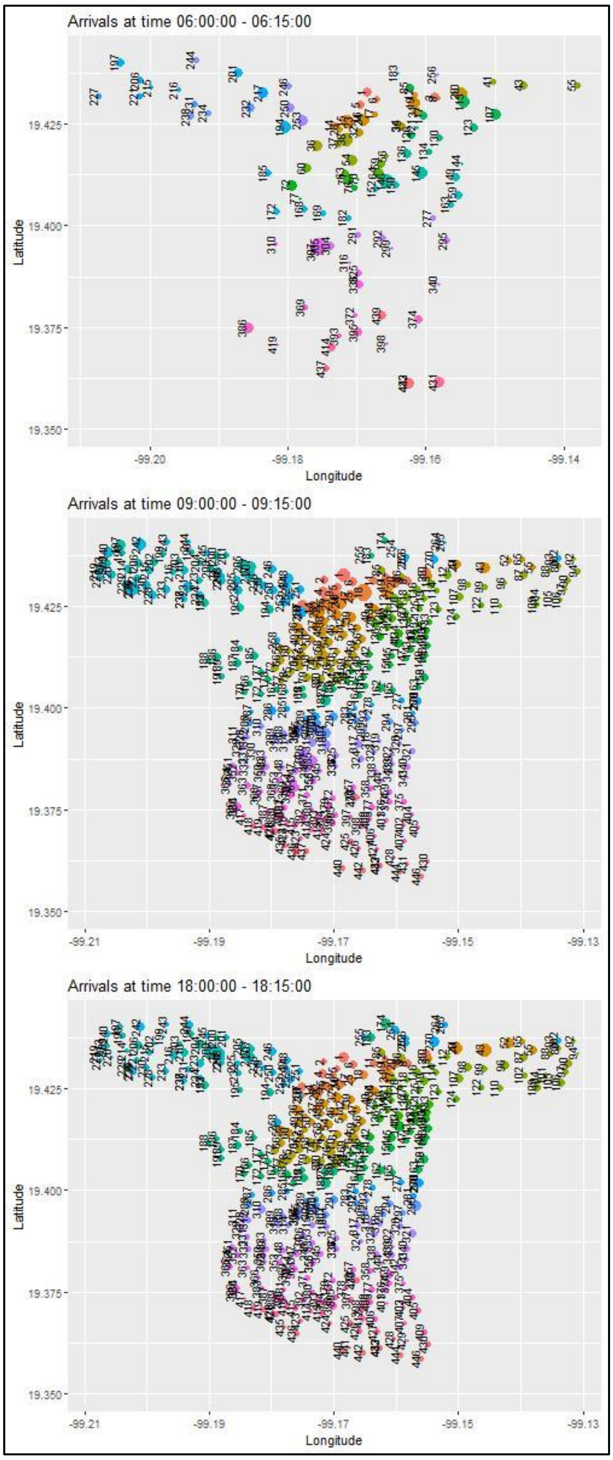

Figure 3: Aggregated arrivals January 2018

Once we explore the aggregate retreats by station in some time intervals, we continue our analysis by exploring the number of arrivals and withdrawals per month, we aggregated the information in time intervals of 5 minutes length because in this way we can fit a time series model that allow us to forecast total withdrawals both in the short time as in the midle time. For example, in the time interval from 5:00 to 5:05 of each day of January 2018, we aggregate the number of arrivals and withdrawals of all of the days. In Figure 4 we show the aggregate arrivals and withdrawals for January, May and December 2018, we can notice that the shape of this graphs is very similar, and it even seems that the aggregate arrivals and withdrawals have a similar behavior that depends strongly on time, for example, for time intervals from 7:00 to 10:00, 14:00 to 16:00 and 18:00 to 19:30 we observe an increase in arrivals and withdrawals. It is interesting to notice that in Mexico´s City these time intervals match with the start of workday, the break time and the end of the workday, respectively. Hence, we have evidence to believe that the demand of bicycles is highly correlated with working times. Also, it is interesting to note that although September is one of the most raining months in Mexico City, the shape of the aggregated withdrawals in this month is very similar to the aggregates withdrawals in January or May (and also in the other months), this fact give us evidence to believe that the whether conditions do not affect in a significant way the aggregated withdrawals, so we can continue our analysis without considering weather conditions.

After we explore the aggregate arrivals and withdrawals, we now analyzed these quantities in a disaggregated

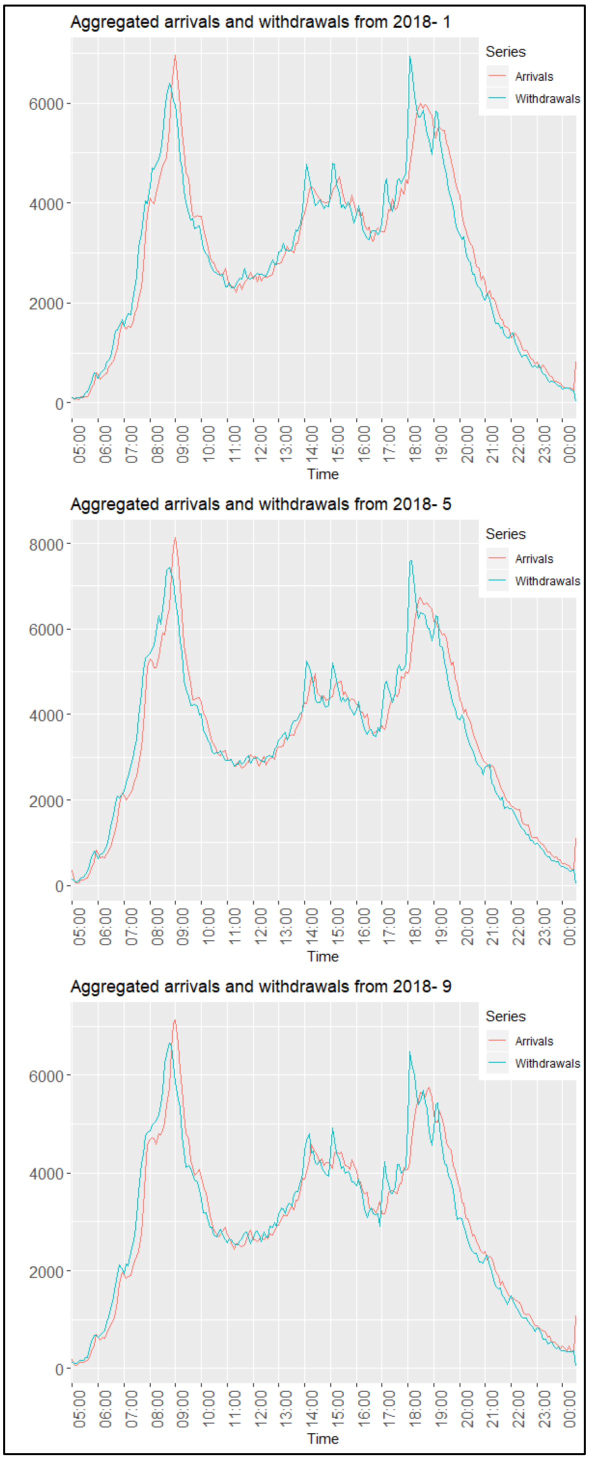

Figure 4: Aggregated arrivals and withdrawals

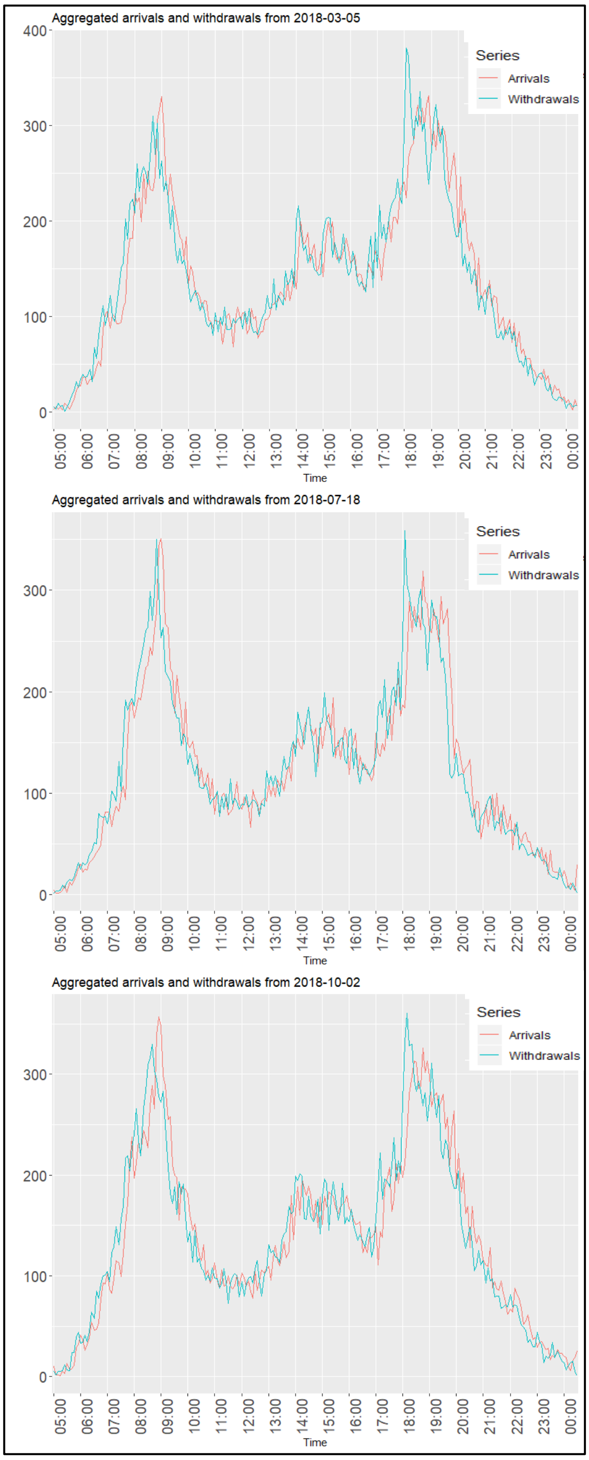

Figure 5: Arrivals and withdrawals by day

way and we found out that in working days (Monday to Friday) the number of arrivals and withdrawals maintain the same shape as the aggregated versions, this is illustrated in Figure 5 for three random working days selected from 2018.

4. TIME SERIES ANALYSIS

A time series is a set of chronological measures registered about the data generated by a random phenomenon of interest. The measures are ordered according to the time in which they were registered, also it is suppose that the time between any two measures is fixed, and that the data generated by the phenome in some period can be correlated with some data generated in the past. A time series model tries to describe through a mathematical expression the phenome that generates the data, and once that we can describe the phenome with a mathematical expression, we can use it to make some forecasts about the data that the phenome could generate in the future. To fit a time series model, it is very common to follow the methodologic proposed by Box and Jenkins (1970), which in general terms consists in the following steps:

1. Apply some statistical exploratory analysis to the measures registered to try to identify a set of suitable time series models.
2. Estimate the parameters of each model.
3. Verify that the assumptions of the model are correct and choose a model.
4. Test of the model.

Once we have explored the available data, we propose to fit a time series model to, forecast the aggregated withdrawals, to do that we will follow the Box and Jenkins methodologic. Intuitively, as we don´t know exactly the number of bicycles withdrawn in any time of any day, we will assuming that this quantity is generated by a random phenomenon, and through a time series model we will describe it, to do this, we decided to work with the aggregated time series, first of all, because although we think that the daily time series are similar in shape it is more difficult that a model fitted to a particular daily data can be used to predict the withdrawals of each day of the year. Instead of that, we prefer a model that give us accurate aggregate predictions and then try to distribute the forecasted quantity among the days of the month. In second, because if we need to adjust a model for each month of the year, it is easier to adjust 12 models than adjust one model for each day of the year. Finally, as we will talk later, for future works it is convenient for us to forecast the aggregate withdrawals.

4.1. Model obtained

To obtain our model we worked with January of 2018 data, and followed the Box and Jenkins methodologic described above in the next steps:

1. We transformed the time series by taking natural logarithm (Box and Jenkins step 1).
2. We take some differences and seasonal differences to the logarithm series to remove the trend and some stational behavior that was present (Box and Jenkins step 1).
3. We use some statistical criterions to identify the model that best suited to our transformed time series and proved that the assumptions of the model were correct (Box and Jenkins steps 2 and 3).
4. After choosing a model, we fit it in other data sets as May and November information, and then make some forecast to the aggregate withdrawals from January, May and November in some time intervals (Box and Jenkins step 4).

All our results were obtained with the statistical software R (R Core Team 2019). The data manipulation were addressed whit the tydiverse package (Hadley Wickham 2017), the figures were obtained in ggplot2 package (H Wickham 2016) and our predictions with the forecast package (Hyndman, Athanasopoulos, Bergmeir, Caceres, Chhay, O'Hara-Wild, Petropoulos, Razbash, Wang and Yasmeen 2019).
Now we illustrate the above steps as well as our results. To stabilize the trend and the variance of the withdrawals series, we used the first difference of the logarithm transformation, then we notice a seasonal pattern in the series so we also take a seasonal difference of order 3, that is, we observed a pattern that was repeated every 3 observations, so we subtract to each observation the observation of three previous times. In Figure 6 we show the resulting time series after taking natural logarithm and differences, the "Withdrawals" series are the series resulting from take natural logarithm and one difference, the "dif" series are the series resulting from take natural logarithm, one no seasonal difference and one seasonal difference of order 3.

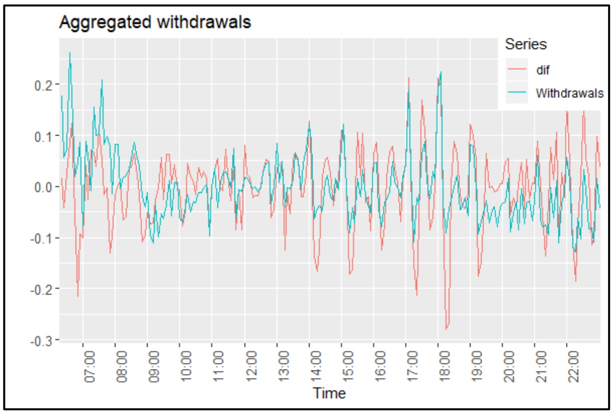

Figure 6: Transformed series

Once we have build the "dif" series, we identified and fit some models with the help of the autocorrelogram (ACF) and partial autocorrelogram (Partial ACF) graphs, this graphs are showed in Figure 7. The partial ACF help us to identify the autoregressive component order k, for example, if the value of the series at some time t has any correlation with the value of the same series at some time t-k, we will see a great value at lag k in the partial ACF (we mean a great lag value as a value that is above or below of the blue dashed lines of the ACF and partial ACF, respectively). The ACF help us to identify a moving average component of order k or a seasonal component of order k. In the first case, for example, if we observe a great value in the first k lags of the ACF, this give us evidence to believe that the value of the series at time t can be explained by k random components observed from time t-1 to time t-k. In the second case, for example, if we note a great lag that follows a pattern, this is, that is present each k periods we have evidence to believe that the value of the series at some time t can be explained by the values of the series at time t-k, t-2k, t-3k, etc. From figure 7, we used the autocorrelogram (ACF) and partial autocorrelogram (PACF) to identify model candidates. The great correlation at lags 3, 6 and 9 of the ACF suggest a seasonal component AR(3) (autoregressive component of order 3), while the great correlation at lags 1 and 2 of the partial ACF suggest a non-seasonal component AR(2). Also, the great correlation at lags 3,6 and 9 of the ACF suggest a seasonal MA(3) component, and the great correlation at lag 1 suggest a non-seasonal MA(1) component. After we observed the ACF and PACF we try many combinations of the suggested seasonal and non-seasonal components to fit the model, in Table 1 we report the models ARIMA(p,d,q)(P,D,Q)$_{[s]}$ that best fitted the data. The acronyms p, d and q refer to the non-seasonal component of the model, where p is the order of the autoregressive component, q the order of the moving average component and d refers to the number of non-seasonal differences. The acronyms P, D, Q and s refer to the seasonal component of the model, where P is the order of the seasonal autoregressive component, Q refers to the seasonal moving average component, D refers to the number of seasonal differences applied, and s refers to the seasonal pattern, more details about time series seasonal models can be found in Box and Jenkins (1970). For each proposed model, we notice that the increase of withdrawals at the time period from 5:30 to 6:00, and the pronounced decrease from 23:00 to 00:30 generated great outliers in their series of errors, these outliers were misleading, because if we removed them the series of errors was so correlated that the Ljung-Box (L-B) test failed (this test tell us if there is enough evidence to believe that the series of errors is not correlated, this is an assumption that must be met in a time series model, for more details see Box and Jenkins 1970). To overcome this problem, we fitted our model to the data belonging to time interval from 06:00 to 23:00.

The AIC statistic (Akaike 1974) give us a criterion to choose among different time series models fitted to a data set. In general, following this criteria, we choose the model with the lowest AIC. Also, looking for a parsimonious model, this is, a model with a low number of parameters, we decided to choice an ARIMA (1,1,0)(3,1,0)$_{[3]}$ model, which respective residuals are showed in Figure 8. The Ljung-Box test (Ljung and Box 1978) associated to the residuals of our chosen model has

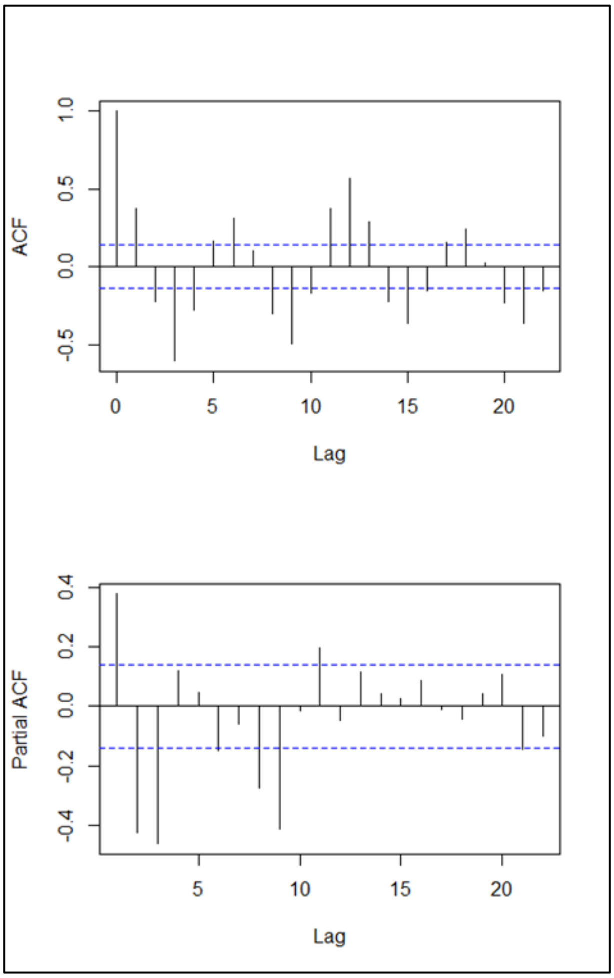

Figure 7: ACF and PACF

Table 1: Models comparison

ARIMA (2,1,0)(3,0,0)$_{[3]}$	
Coefficients	ar1=0.41 ar2=0.01 sar1=0.17 sar2=0.07 sar3=-0.12
AIC	-594.15
L-B Test	p-value=0.2084
ARIMA (2,1,0)(3,1,0)$_{[3]}$	
Coefficients	ar1=0.20 ar2=0.09 sar1=-0.72 sar2=-0.46 sar3=-0.56
AIC	-613.84
L-B Test	p-value=0.84
ARIMA (1,1,0)(3,1,0)$_{[3]}$	
Coefficients	ar1=0.22 sar1=-0.7 sar2=-0.44 sar3=-0.55
AIC	-614.13
L-B Test	p-value=0.79
ARIMA (1,1,0)(3,0,0)$_{[3]}$	
Coefficients	ar1=0.41 sar1=0.17 sar2=0.33 sar3=-0.12
AIC	-596.13
L-B Test	p-value=0.229

a p-value 0.79 so we can accept the statistical hypothesis that the residuals are not correlated (this is also an assumption that must meet a time series model). Note that the histogram of the residuals seem to be normal distributed, also, if we visualize the qq-plot of the residuals as in Figure 9, and we perform the Shapiro-Wilk test (Shapiro and Wilk 1965), we got a p-value equal to 0.84, so we can accept the statistical hypothesis that the residuals are generated by a normal distribution. In summary, the residuals can be seen as independent realizations of a normal distribution with constant variance.

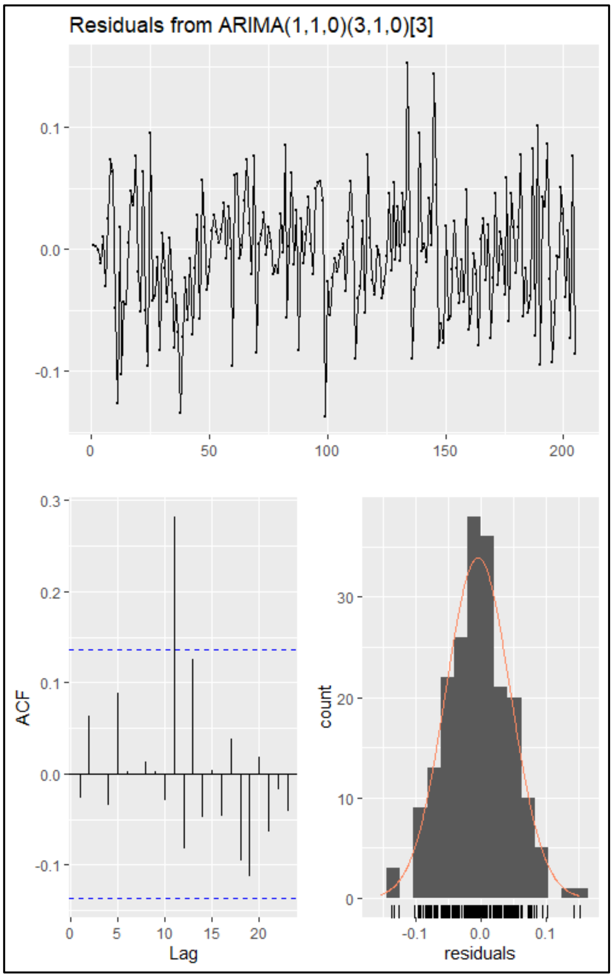

Figure 8: Model residuals

Once we proved that our model is adequate to fit the January data, we also used it to fit the data of all the other months of 2018, in general we got good results. In Figure 10 we show the fitted series for January data of the ARIMA(1,1,0)(3,1,0)[3] model obtained, we also show the one-step forecast of our model applied to May and November data, the red series refers to the fitted values, while the black series refers to the observed data.
Finally, to complete our analysis we got some forecast of our model, in Figures 11 we show the forecast of our model in the data series of January, May and November of 2018 for time interval from 11:00 to 12:15, in Figure 12 from 13:20 to 14:05 and in figure 13 from 20:00 to 20:45, we also show confidence intervals at 85% and 95% in all Figures.

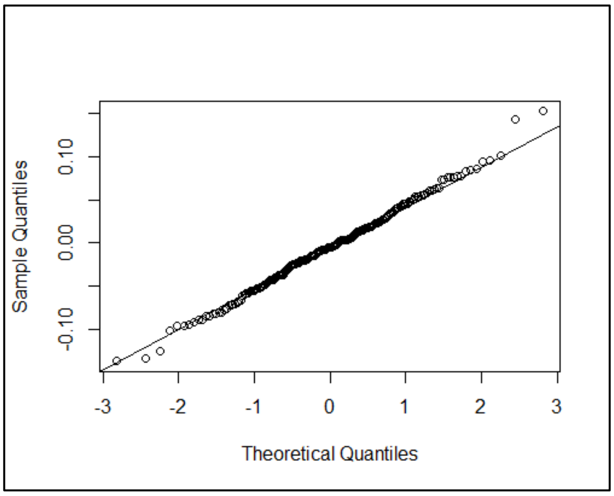

Figure 9: Residuals Q-Q Plot

5. CONCLUSIONS

To estimate the demand of Ecobici system we focused in estimate the aggregate number of withdrawals in time intervals of 5 minutes. These quantities can be visualized as a surplus or deficit of bikes if we subtract the overall quantity of bikes that all the stations could store. The predictions that we showed can help us to estimate the number of withdrawals in the short term as well as in the middle term, we showed some time intervals of 45 minutes with confidence intervals at 85% and 95%.
It is important to mention that the model showed here seems to fit adequately the withdrawals series, and also that the model was built just with historical withdrawals data, so one of its strengths is that it can give accurate forecast without using extra information as weather conditions, or geospatial data. This last point also give us an area of opportunity to make more robust the model by adding another sources of information, we believe that we can get better results if we incorporate some geospatial information related to the surroundings of the stations, for example incorporate the information related to offices, schools, hospitals, etc. Also, we can try to incorporate some covariables, a natural choice is to incorporate the aggregate number of arrivals, but due to its great similitude to the number of withdrawals we don´t know yet if this information will add useful information. Hence, as future analysis we will try to make more robust our model by exploring the points mentioned before, and also find and add more sources of information as covariables, this last point can be a little difficult because we must find information that is reported and stored through the hours and days. Another point in which we are still working is how to predict the individual number of withdrawals of each station through the predicted aggregate demand. We have though to search for graph structures among the stations and investigate if these structures change during the day, so we could estimate the number of withdrawals that each

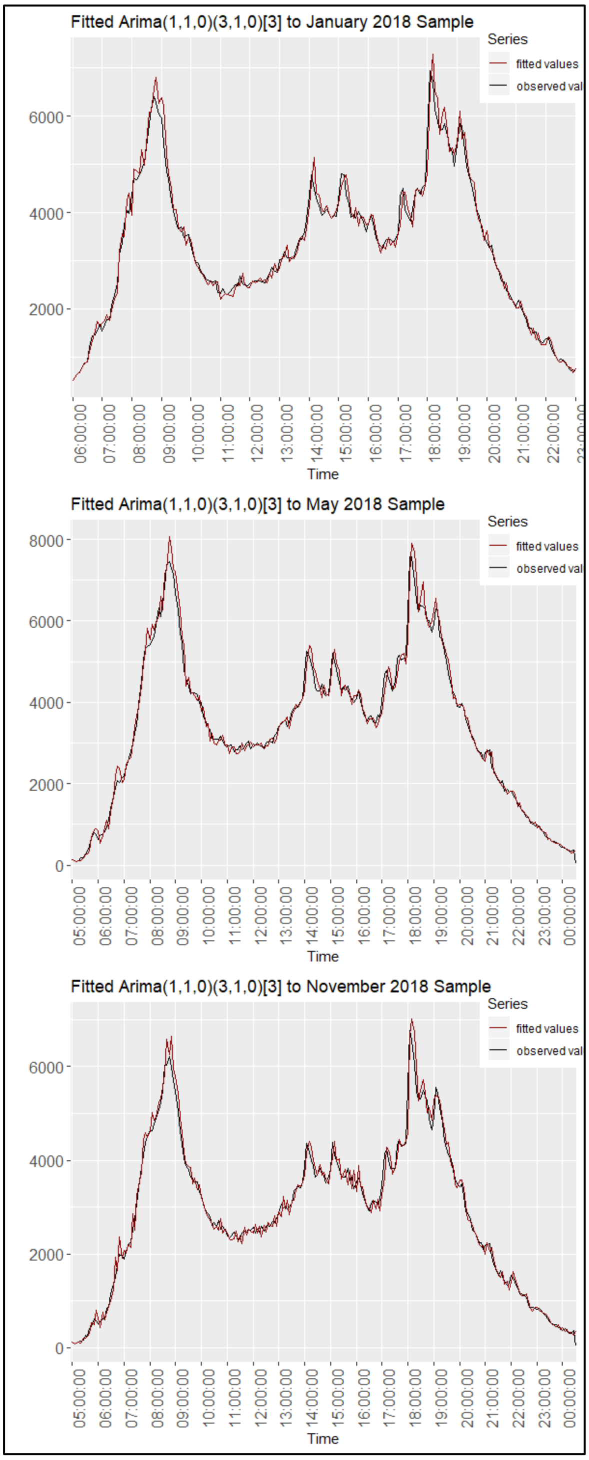

Figure 10: Fitted model

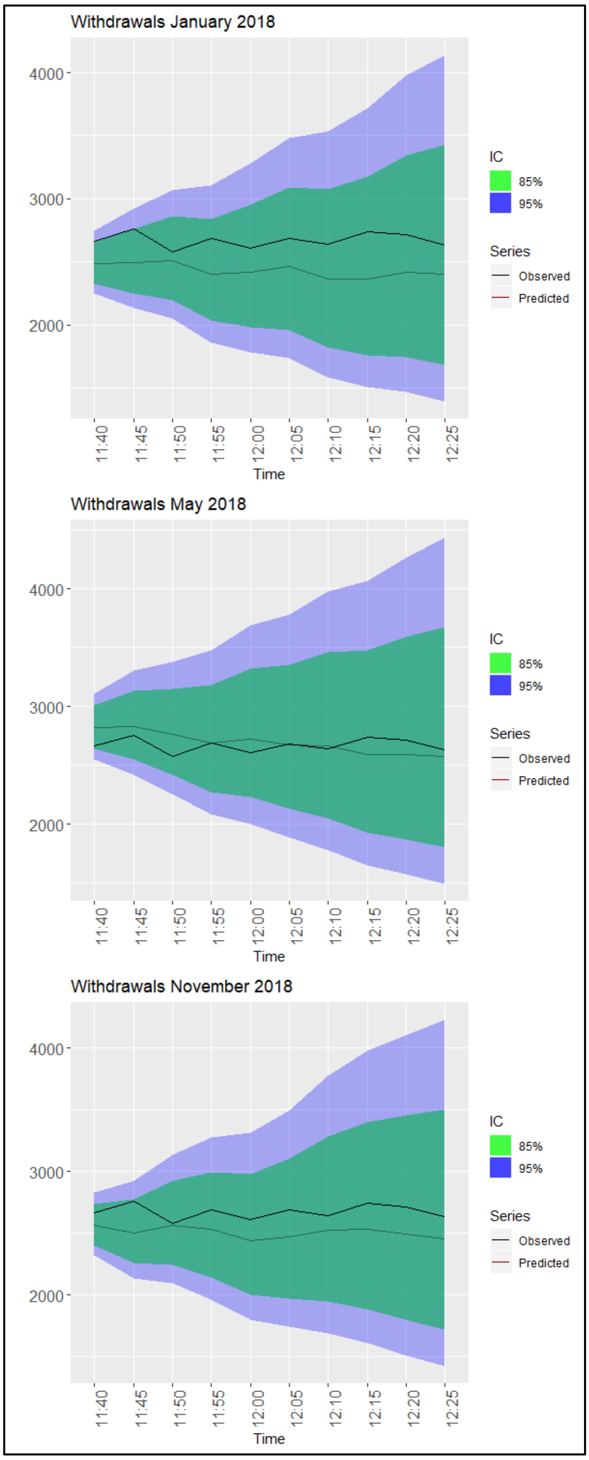

Figure 11: Forecasts from 11:40 to 12:25

component of the graph demands conditioned in the overall aggregate withdrawals.

REFERENCES

Akaike H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

Barth M., and Todd M., 1999. Simulation model performance analysis of a multiple station shared vehicle system, Transp. Res. Part C Emerg. Technol., 7 (4), 237–259.

Borgnat P., Abry P., Flandrin P., Robardet C., Rouquier J.-B., and Fleury E., 2011. Shared bicycles in a city: A signal processing and data analysis perspective, Adv. Complex Syst., 14 (3), 415–438.

Box G.E.P., Jenkins G., 1970. Time Series Analysis, Forecasting and Control. San Francisco:Holden-Day.

Faghih-Imani A., Hampshire R., Marla L., and Eluru N.,2017. An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona

.

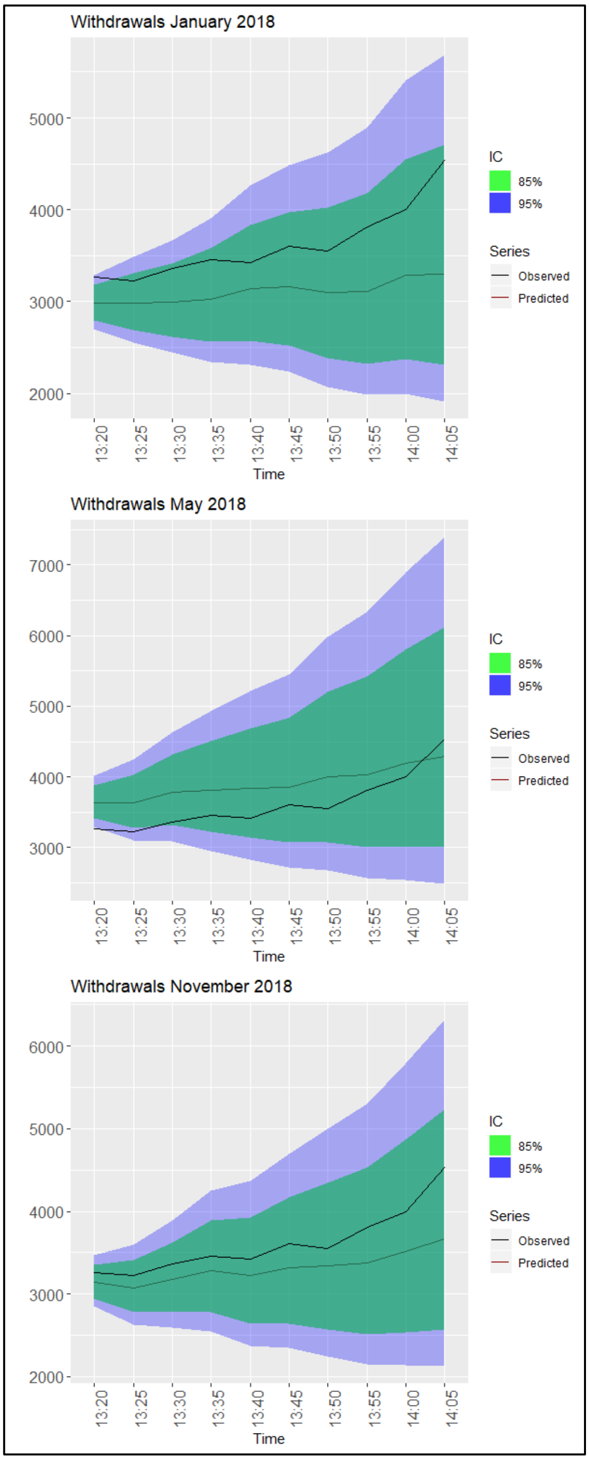

Figure 12: Forecasts from 13:20 to 14:05

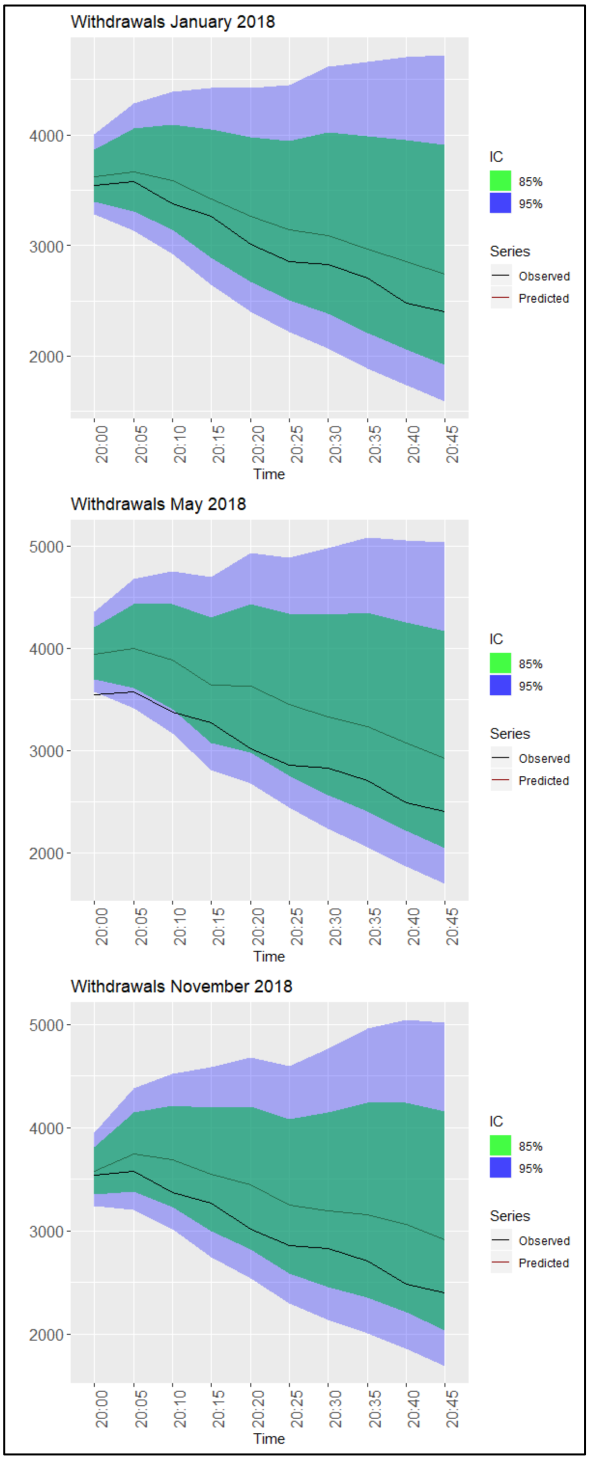

Figure 13: Forecasts from 20:00 to 20:45

and Seville, Transp. Res. Part A Policy Pract., 97, 177–191.

Fishman E., Washington S., and Haworth L.N., 2013 Bike Share: A Synthesis of the Literature. Transp. Rev., 33 (2), 148–165.

Froehlich J. E., Neumann J., and Oliver N., 2009. Sensing and Predicting the Pulse of the City through Shared Bicycling. Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence, Pasadena (California, USA).

Hyndman R.J., Athanasopoulos G., Bergmeir C., Caceres G., Chhay L., O'Hara-Wild M., Petropoulos F., Razbash S., Wang E., Yasmeen F., 2019._forecast: Forecasting functions for time series and linear models. R package version 8.9, URL:http://pkg.robjhyndman.com/forecast.

Hyndman R.J., Khandakar Y., 2008. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software, 26(3), 1-22. URL: http://www.jstatsoft.org/article/view/v027i03.

Kaltenbrunner A., Meza R., Grivolla J., Codina J., and

Banchs R., 2010. Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system, Pervasive Mob. Comput., 6 (4), 455–466.

Ljung G.M., Box G.E.P., 1978. On a mueasure of lack of fit in time series models. Biometrika 65:297-303.

Moncayo–Martínez L. A., 2020. Analysing Data Set of the Bike–Sharing System Demand with R Scripts: Mexico City Case. In: Bi Y., Bhatia R., Kapoor S. (eds) Intelligent Systems and Applications, 1038, 90–105.

Moncayo-Martinez L. A., and Ramirez-Nafarrate A., 2016. Visualization of the mobility patterns in the bike-sharing transport systems in Mexico City. Proceedings of the International Conference on Industrial Engineering and Engineering Management (IEEM), 1851–1855, Bali (Indonesia).

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Pfrommer J., Warrington J., Schildbach G., and Morari M., 2014. Dynamic Vehicle Redistribution and Online Price Incentives in Shared Mobility Systems, IEEE Trans. Intell. Transp. Syst., 15 (4), 1567–1578.

Schuijbroek J., Hampshire R. C., and van Hoeve W.-J., 2017. Inventory rebalancing and vehicle routing in bike sharing systems, Eur. J. Oper. Res., 257 (3), 992–1004.

Shapiro S.S., Wilk M.B., 1965. An Analysis of Variance Test for Normality (Complete Samples). Biometrika 52, 591-611.

Wickham H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham H., 2017. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse.

AUTHORS BIOGRAPHY

David E. Castillo-Rodríguez earned his bachelor degree in the Actuarial science program provided by the Instituto Tecnológico Autónomo de México (ITAM). He will graduate in the master's in Computer Science program provided also by the ITAM in 2020. He has participated as assistant professor in the Universidad Nacional Autónoma de México (UNAM), and nowadays he is partial time professor at the department of statistics in the ITAM. He has experience working and developing statistical models both in the private and public sector. David Castillo is very passionate and interested in topics related to statistical models, computer and data science.

Luis A. Moncayo-Martínez Autónomo de México (ITAM) and member of the Mexican National Researchers System (NRS). Luis A. earned a PhD in Engineering from the University of Exeter (UK). His research focuses on developing algorithms to optimise logistics and manufacturing systems. The results of his research have been published in peer-reviewed journals e.g. Int. J. Production Economics, Expert Systems with Applications, J. of Manufacturing Systems, and Int. J. of Advanced Manufacturing Technology.He has developed several projects about process improvement financed by the CONACYT and private companies. He is a member of INFORMS, IISE, and the Int. Association for Statistical Computing.