Bike Hires Forecasting

for Santander Cycles and Serco

Christopher A. Swenson

Pennsylvania State University

Executive Summary

Santander Cycles is a bicycle hire ("rental") public service in London, UK, which requires periodic maintenance of the bicycles and docking stations. An (assumed) executive at Serco, the operator of the service, requested an analysis of their daily bicycle hires to investigate the following questions:

1. Are there cyclical patterns in the daily hires?
2. Do weekdays, holidays, and other time-related factors influence daily hires?
3. Can we forecast hires for 2019?

The hires exhibited an annual seasonal trend that resembles a sine wave, with lower counts of hires in winter months and higher counts of hires in summer months. After accounting for this trend, a 7-day seasonal cycle was prominent. Thus, annual and weekly periods are important. Winter days, weekends, and non-Christmas holidays had fewer hires than other days. Trends by month and day of week were accounted for in the annual and weekly seasonal trends. A forecast of 2019, including holidays as important factors, had an average error of about 17%, and this forecast may assist with maintenance, inventory management, and staffing decisions. The forecast was reviewed in early 2020 and had an average error rate of 15.7%.

Bike Hires Forecasting

Santander Cycles is a bicycle hire ("rental") public service operation in London, UK, that provides low-cost electric and non-electric bicycles for rent, including docking stations where the general public are able to rent and return bicycles [0]. Serco, the company that operates the service, conducts maintenance of thousands of bicycles and hundreds of docking stations around London [3], and during maintenance, the bicycles and docks are unavailable. With about 10 million hires per year from 2014–2018, it is critical for the Serco to understand fluctuations in customer demand.

Serco collects and publishes detailed hire data from the docking stations, with some data sets providing hourly counts of hires [1]. In early 2019, an (assumed) executive at Serco requested an analysis of the daily hires to understand demand, to be delivered in the first quarter of 2019. In particular, the executive wished to understand:

1. Are there cyclical patterns in the daily hires?
2. Do weekdays, holidays, and other time-related factors impact the forecasted hires?
3. Can we forecast hires for 2019?

The answers to these questions may help Serco make decisions on maintenance, inventory management, and staffing. Ideally, maintenance would be scheduled during non-peak periods, staffing increased during peak periods, and adjustments made to inventory based on forecasted demand.

Method

Variables

The docking stations automatically track counts of bicycle hires, and Serco provides these counts publicly, aggregated by hour and day. The analysis for the questions from Serco used daily counts of hires from 2010–2018, which included the date and number of hires [1].

Names of days, names of months, and a simple time index (the first date is 1, the next is 2, etc.) were added manually. Two additional variables using the sine and cosine function on the time index were added based on the discovery of a seasonal sine-wave pattern (see *Exploratory Data Analysis*). Holidays were identified from the online database Nager.Date to identify holidays in the UK from 2010–2019 [2]; which included January 2, Battle of Boyne, Christmas, May Bank Day, Easter Monday, Good Friday, New Years, Saint Andrew's Day, Saint Patrick's Day, Spring Bank Day, Saint Stephen's Day, and Summer Bank Day.

Time Series Model

An observational, time series model type called ARIMA ("Auto-Regressive Integrated Moving Average") was used to assess time-related factors and forecast hires for 2019 for questions 2 and 3, respectively. ARIMA has autoregressive ("AR"), "integration" ("I"), and moving average ("MA") components, written as ARIMA(0, 0, 0), where each number represents the value of the component in order. The AR component is the number of prior periods of the measure of interest (e.g., one day before, two days before, etc.), and the MA component is the number of prior periods of random noise (i.e., Gaussian).

Integration. The ARIMA model assumes that the measure of interest is stationary – that it does not have any steady increases or decreases ("trend"), abrupt changes, or long-term patterns. (Additionally, it is sensitive to outliers and changes in variability.) In the presence of trends, the measure of interest must be flattened or "detrended" prior to modeling with ARIMA, and there are a few methods to accomplish this. The most common method is to subtract the current value[1]

1 The "current" value is the measured value on the date of the record of interest.

from the prior value, and this subtraction is called “first differences” or “integration” in ARIMA.

ARIMA Component Values. Using an AR value of 1 would include the prior day of the measure of interest in order to estimate the current day (called “lagging”). Using an AR value of 2 would include two prior days. The MA components are similar but lag random noise instead of an observed value. The integration component usually does not exceed a value of 1 or 2, where 2 would include differences of two prior periods, and as noted are used to detrend the data. For example, a model without integration, AR=2, and MA=1, would be written as ARIMA(2, 0, 1).[2] The ARIMA model can be extended to account for seasonal patterns, written as SARIMA(0, 0, 0) x (0, 0, 0)_S, with the values after the “x” representing the seasonal AR, I, and MA components, respectively, and S representing the seasonal period (e.g., 7 days). The values of the components are similar but applied at the seasonal period (e.g., a seasonal AR component of 1 with a season of 7 would lag last week’s measure).

The non-seasonal values cannot exceed the seasonal period; the seasonal components have an upper limit based on the number of periods in the data. All components can be estimated from the Exploratory Data Analysis (“EDA”) and tested by comparing variations of the model as well as direct statistical tests. High values (called “orders”) for any of these components may lead to models that perform poorly (“over-parameterized” or “over-fit”). Finally, external factors, like the day of the week, can be included as “exogenous” factors (called ARIMAX or SARIMAX, with “X” meaning exogenous).

Assessment. The model variations, including ARIMA components (“parameters”) and exogenous factors, were compared on two model quality measures: BIC (Bayesian Information

2 For this example, the equation would be $y = \beta_0 + \beta_1 \mathrm{lag}(y, 1) + \beta_2 \mathrm{lag}(y, 2) + \beta_3 \epsilon + \beta_4 \mathrm{lag}(\epsilon, 1)$, where y is the outcome, ϵ is the random noise, β_i are the coefficients, and lag(x, n) is a function to find the nth prior value of x.

Criterion) and MAPE (Mean Absolute Percent Error), where lower values are better, as well as direct statistical tests on each component (t-tests, in a similar fashion to assessing predictors in linear regression). The BIC measure penalizes complex models, making the BIC measure larger unless the model is more useful when more complicated. The MAPE measure is essentially an average error of the estimated (or forecasted) values over time. The t-tests for each component help to assess whether the parameter or factor is statistically significant within the context of the model, meaning that it has an important relationship with the measure of interest.

Analyses

Question 1. An exploratory data analysis (EDA) was conducted in order to assess the cyclical patterns, with the hires visualized over time. The hires were inspected for seasonal trends, outliers, long-run cycles, changes in the variance, and abrupt changes, which all affect ARIMA models. Four methods were assessed to account for trend: first differences and three models on the time index (linear, polynomial, and trigonometric). The estimated hires from each model was compared with the actual hires to assess how well each detrended the hires.

In order to assess seasonal patterns, the detrended hires were analyzed in a Fourier analysis, which identifies important time components like yearly, monthly, or weekly cycles. The analysis converts data over time into frequency and spectrum components, which are then plotted in a “periodogram” and any spikes in the plot indicate important periods in time. The autocorrelation and partial autocorrelation functions (ACF and PACF, respectively)[3] can also be plotted, and any significant spikes indicate important time periods. From these two analyses, the ARIMA components were estimated, in particular, the seasonal duration identified. Given that

3 Both the ACF and PACF measure correlations between the current and prior values, and the PACF accounts for all prior values in between the current and prior lag (if it is further back than 1 lag).

the hires are daily, multiples of 7-, 28-, 30-, 31-, or 365-day seasons are possible.

Question 2. To assess the impact of holidays and other date-related factors, external factors, including day name, month name, a weekend indicator, and holidays (by an indicator and name), were included in variations of the ARIMA model. The external factors were assessed by comparing model variations and statistical tests (t-tests), along with the ARIMA components. Only select model variations were reported.

Question 3. For forecasting the 2019 hires, the BIC measure was used to select the best model variation, selecting the variation with the smallest BIC.

Results

Exploratory Data Analysis

The data consisted of 2,922 counts of bike hires per day from 2010–2018. Table 1 displays the statistical summary of the bicycle hires. Table 2 (Appendix) displays the hires by year, month, day of week, weekend, holiday, and holiday name.

Table 1. *Summary of the Bicycle Hires*

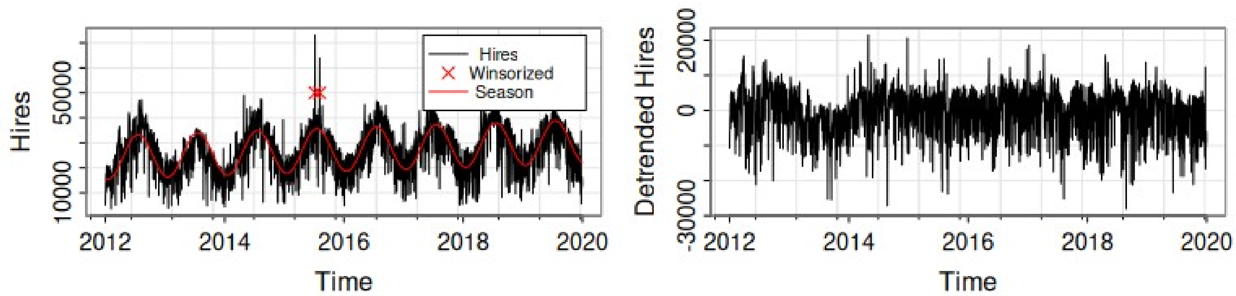
Count	Mean	St. Dev.	Min	Q1	Median	Q3	Max
2,922	27105.63	9007.84	3531	20854.25	27306	33914.75	73094

The years 2010 and 2011 were ramp-up periods with lower counts of hires compared to 2012–2018; thus, these years were excluded from further analysis. Two outlier observations were noted in 2015 with an unusual, excessively large number of hires. These two observations were set to 50,000 hires (i.e., “Winsorized”) in order to avoid difficulties with the ARIMA model. The hires exhibited an annual sine-wave pattern, with fewer hires in winter months and more hires in summer months, and this trend was estimated using the sine and cosine functions[4]. Figure 1

4 The equation is $y = \beta_1 \cos(2\pi t/365) + \beta_2 \sin(2\pi t/365)$, where t is the time index.

displays the hires over time, excluding 2010–2011, with the Winsorized outlier observations and the seasonal sine-wave estimate overlaid. The detrended hires are also displayed.

Figure 1. *Bike Hire Trend: Raw Data (left) and Detrended Hires (right)*



Peaks at 7, 750, and 3.5 days, in order of magnitude were found in the Fourier analysis of the detrended hires (see Figure 2, Appendix). The ACF and PACF of the detrended hires (see Figure 3, Appendix) exhibited a seasonal pattern with moderate ACF values at multiples of 7, beginning moderately and slowly decreasing over time, with similar PACF values tapering to 0. Both the Fourier analysis, ACF, and PACF results indicated a seven-day seasonality that requires a first difference. The ACF in the non-seasonal trend (before lag 7) exhibited a spike that tapers to 0 and the PACF had one single spike at lag 1. Figure 4 (Appendix) displays the same analysis with a first difference applied over seven days. Overall, these patterns support a seasonal moving average component of order 1 and non-seasonal autoregressive component of order 1 or 2; that is, a SARIMA(1 or 2, 0, 0) x (0, 1, 1)_7 model.

Time Factors

Each model variation detrended the hires with two exogenous components to handle the sine-wave annual trend. Table 4 (Appendix) displays select model variations and statistical tests for parameters and factors. Model variation 4 was set up as a SARIMA(2, 0, 0) x (0, 0, 1)_7 with holidays and weekdays, and seasonal integration could not be included due to the use of weekday. All weekdays (aside from Sunday, which was the comparison day) and most holidays were significant ($p < 0.05$). Model variation 5 was set up as a SARIMA(2, 0, 0) x (0, 1, 1)_7 with holidays and month name, including seasonal integration. Most holidays were significant ($p < 0.05$); however, no months were significant. Other time factors, assessed in unreported variations, were not significant.

Model Selection

Table 3 (Appendix) displays a description of select model variations and the quality measures (BIC, MAPE) for each variation. Figure 5 (Appendix) displays each model variation with four measures of model performance, including the BIC, MAPE, root-mean squared error (RMSE), and the mean of the standard errors.[5] Beyond model variation 1, which only included the sine-wave trend, the performance of each model variation was very similar. Model variation 7, SARIMA(3, 0, 0) x (0, 1, 1)_7 with holidays, had the lowest BIC value at 19.830, so this model was selected for forecasting.

Figure 6 (Appendix) displays the diagnostic plots for the selected model, with some concerns: the variability ("variance") of the unexplained counts ("residuals") is slightly large in some areas and it appears there may be significant remaining time components (i.e., significant lags in the Ljung-Box statistic plot). The residuals of the model do not appear to be normally distributed in the QQ plot and histogram (with the Shapiro-Wilks test, we reject the hypothesis that the distribution is normal with $p < 0.001$), which violated an assumption of the ARIMA model. No other model variations had better diagnostics; however, with an average error of about 17% (MAPE), the selected model may still be useful despite these concerns.

5 The RMSE and MAPE measures were split out among the training (<2019), test (2019), and "Total" data.

Discussion

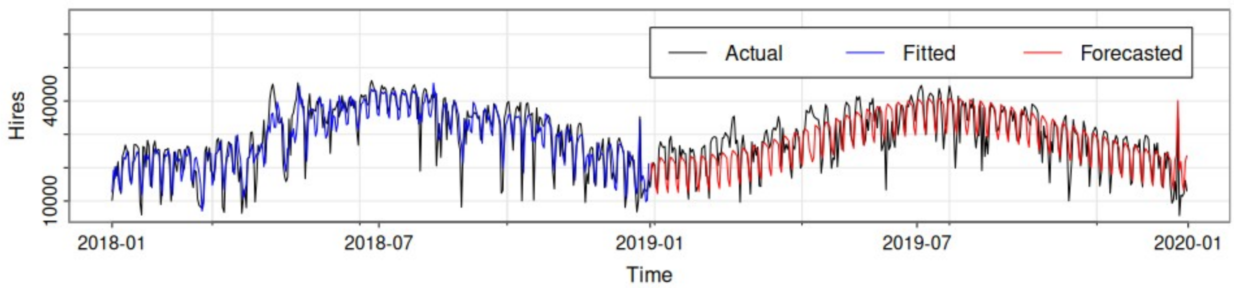
Recommendations

Question 1. Are there cyclical patterns in the daily hires? Yes, annual and weekly patterns influenced daily hires. The hires exhibited a strong annual pattern resembling a sine-wave trend, with the peaks occurring in summer months (July–August) and lows in winter months (December–January). After accounting for the sine-wave pattern in the ARIMA model, a 7-day seasonal cycle was present. Hires were faintly influenced by the prior two or three days, which partially aligned with the 3.5-day cycle identified in the Fourier analysis.

Question 2. Do weekdays, holidays, and other time-related factors influence daily hires? Yes, weekdays and holidays influenced daily hires. The day of the week was significant *without* taking a difference over weeks, with Saturday and Sunday having the lowest number of hires; however, the seasonal integration accounted for weekday variations. Several holidays were significant with a lower number of hires on May Bank Day, Easter Monday, Good Friday, New Year's Day, and Summer Bank Day; Christmas was significant with a *higher* number of hires, exhibiting an unusual spike in hires each year. The annual sine-wave pattern used to detrend the hires accounted for monthly variation; thus, month was not significant. Overall, maintenance on winter days, weekends, and non-Christmas holidays would be generally more acceptable than other days, as there is less demand on these types of days. An increase in inventory and staffing during summer days, weekdays, and Christmas may be help meet customer expectations.

Question 3. Can we forecast hires for 2019? Yes, forecasting 2019 is reasonable with about 17% error using the selected model variation. This indicates that we can estimate 2019 with reasonable accuracy; however, there is room for improvement. Additional, unknown factors may influence the hires, and if possible, the days where large differences occur between estimated and actual hires from 2010–2018 should be investigated for any additional factors. Regardless, I recommend implementing the selected model variation to plan for 2019. The forecasted hires are displayed in Figure 7.

Figure 7. *Forecasted Hires for 2019*



Limitations

An ARIMA model may not be appropriate to estimate or forecast hires, since it is a count measure. Counts are not usually normally distributed (i.e., Gaussian), which is assumed by the ARIMA model. This fact may have contributed to the issues noted in the model diagnostics. ARIMA models are unable to use exogenous factors that also vary over time, which likely influenced the hires (e.g., weather); however, these factors were beyond the scope of this analysis. The ACF and PACF patterns for the seasonal component indicated a fractional difference with moderate ACF values that slowly descend and PACF values that taper to 0. The ARIMA model cannot handle fractional seasonal differences, and variations with seasonal integration may be over-parameterized. Finally, ARIMAs are unable to account for interactions, which likely occurred when holidays and weekends coincided. An lternative model for count data – Poisson regression – was tested, but performed poorly in forecasting.

A few decisions in the analysis were somewhat arbitrary, for example, the exclusion of the first two years and the method of Winsorization for the outliers. A more complex, nuanced estimate of the sine-wave pattern is possible, especially to account for leap years. The 750- and 3.5-day periods from the Fourier analysis were ignored since the ARIMA model can only handle one seasonal period. Although many of these decisions were tested in unreported model variations, these decisions may have impacted the final results.

Even with these concerns, the forecast from the selected model should help Serco plan maintenance, adjust inventory, and meet staffing expectations.

Addendum

Retrospective Analysis. In early 2020, a retrospective analysis was conducted to compare the forecasted and actual hires for 2019 and confirm the model accuracy. Figures 5 and 7 were updated to include the actual hires from 2019, collected in early 2020. The selected variation had a 15.7% average error (MAPE) compared to the actual hires in 2019. Most other model variations performed worse, except for variation 5 with 15.5% average error. Variation 5 used AR=2 with holidays and months, compared to the selected model that used AR=3 with holidays only. Thus, the retrospective analysis confirmed that the selected model variation was over-parameterized in the non-seasonal AR component. Using the BIC measure to select the variation may have led to this result, and had the MAPE measure been used instead, variation 5 would have been selected for forecasting, as it was the simplest model with the smallest MAPE. (Variation 3 with an error of 17.1% on estimated hires and and error of 15.7% on forecasted hires could also have been a candidate.) The alternative forecasts were indiscernible in a visual comparison (see Figure 8, Appendix); thus, the selected variation was still a reasonable choice.

References

[0] Transport for London: Santander Cycles https://tfl.gov.uk/modes/cycling/santander-cycles

[1] London Datastore: Number of Bicycle Hires https://data.london.gov.uk/dataset/number-bicycle-hires

[2] Nager.Date: Public Holidays in United Kingdom https://date.nager.at/PublicHoliday/United-Kingdom/

[3] London cycle hire scheme: Santander Cycles https://www.visitlondon.com/traveller-information/getting-around-london/london-cycle-hire-scheme

Resources

This analysis used the free, open-source software R version 4.1 on Ubuntu GNU/Linux 24.04 LTS. The specific implementation of ARIMA was provided by the package "astsa" version 2.1 downloaded from the CRAN R package repository.[6] In R, ARIMA models with exogenous factors are treated as linear regressions with error terms having an ARIMA structure. Alterative software packages must treat exogenous factors in the same manner in order to obtain similar (if not identical) results, and may vary from the reported model variations due to the use of non-deterministic methods like maximum likelihood estimation. Using the "statsmodels" package in Python should provide a similar results.[7]

6 https://cran.r-project.org/web/packages/astsa/index.html
7 Specifically: statsmodels.tsa.arima.model.ARIMA or statsmodels.tsa.statespace.sarimax.SARIMAX (MLE).

Appendix

Tables

Table 2.

Total Cycle Hires by Time Factors

Factor	Factor Value	Count	Unique	Total	Mean	St. Dev.	Min	Median	Max
Year	2014	365	363	10023897	27462.73	9065.09	4327	27676	4327
	2015	365	365	9871839	27046.13	8547.89	5779	26618	5779
	2016	366	364	10303637	28152.01	8850.96	4894	27881.5	4894
	2017	365	362	10446044	28619.3	8376.55	5143	29490	5143
	2018	365	363	10567540	28952.16	10174.35	5859	29190	5859
Month	Jan	155	152	2948250	19020.97	6012.81	4327	19939	4327
	Feb	141	140	2864834	20317.97	5006.3	8975	21217	8975
	Mar	155	155	3548406	22892.94	6260.11	6362	24230	6362
	Apr	150	149	4149875	27665.83	7084.49	7188	28433	7188
	May	155	155	4909631	31675.04	6929.78	12049	32618	12049
	Jun	150	150	5335509	35570.06	5657.39	17387	36579	17387
	Jul	155	155	5891366	38008.81	6619.05	12497	38879	12497
	Aug	155	155	5308911	34251.04	7611.35	6175	34862	6175
	Sep	150	149	4958104	33054.03	5486.84	12376	34577	12376
	Oct	155	155	4690488	30261.21	5761.4	10043	31364	10043
	Nov	150	148	3641073	24273.82	6050.81	6030	25935	6030
	Dec	155	155	2966510	19138.77	7087.8	5143	18883	5143
Weekday	Mon	261	258	6411029	24563.33	9981.84	5779	24844	5779
	Tue	261	260	6113685	23424.08	10305.63	4894	23063	4894
	Wed	261	261	7348048	28153.44	7885.7	6175	27880	6175
	Thu	260	259	7921619	30467.77	7666.17	7149	30184	7149
	Fri	261	261	7869118	30149.88	8164.78	4327	29455	4327
	Sat	261	259	7963618	30511.95	8314.09	8744	30856	8744
	Sun	261	259	7585840	29064.52	7780.73	6362	28782	6362
Weekend	No	1304	1266	38688243	29668.9	8007.15	4327	29413.5	4327
	Yes	522	514	12524714	23993.7	10151.31	4894	23789	4894
Holiday	No	65	65	1461203	22480.05	11168.56	4327	23498	4327
	Yes	1761	1701	49751754	28251.99	8892.17	4894	28401	4894
Holiday Name	New Year's Day	5	5	64803	12960.6	3528.38	7246	14394	7246
	2 January	5	5	172218	34443.6	7466.87	23498	38331	23498
	Saint Patrick's Day	5	5	161989	32397.8	6990.81	22871	35267	22871
	Good Friday	5	5	144207	28841.4	11130.97	18493	26933	18493
	Easter Monday	5	5	84700	16940	7484.55	7998	20130	7998
	May Bank Day	5	5	87986	17597.2	8231.16	6362	22152	6362
	Spring Bank Day	5	5	40633	8126.6	2591.21	4327	9615	4327
	Battle of the Boyne	5	5	123604	24720.8	3635.17	19998	25986	19998
	Summer Bank Day	5	5	114206	22841.2	8249.09	8106	26465	8106
	Saint Andrew's Day	5	5	128915	25783	9039.64	12049	26476	12049
	Christmas Day	5	5	47667	9533.4	1996.84	7149	10362	7149
	St. Stephen's Day	10	10	290275	29027.5	12095.62	6175	33484	6175

Table 3.

Model Variation Descriptions and Training Performance

Variation	Description*	Parameters‡	BIC	MAPE
1	sine(time index) + cosine(time index)	2	20.438	27.30%
2	SARIMA(2,0,0)x(0,1,1)_7	5	19.878	17.80%
3	SARIMA(2,0,0)x(0,1,1)_7 + Holidays	13	19.832	17.10%
4	SARIMA(2,0,0)x(0,**0**,1)_7 + Holidays + Weekday†	20	19.853	17.20%
5	SARIMA(2,0,0)x(0,1,1)_7 + Holidays + Month	24	19.860	17.00%
6	SARIMA(2,0,0)x(0,1,**2**)_7 + Holidays	14	19.833	17.00%
7	SARIMA(**3**,0,0)x(0,1,1)_7 + Holidays	14	19.830	17.00%
8	SARIMA(**3**,0,0)x(0,**0**,1)_7 + Holidays + Weekday†	22	19.840	17.00%

* Each model includes two variables calculated with sine and cosine as noted in Variation 1.
† Models with Weekday could not be estimated with seasonal integration.
‡ Integration component values are not included in the number of parameters.

Table 4.

Statistical Tests by Model Variation

Variation	Parameter*	Estimate	Standard Error	Statistic (t)	Significance (p)
1	intercept	26889.808	130.435	206.156	< 0.001 *
1	xcos	-8221.514	184.392	-44.587	< 0.001 *
1	xsin	-3077.516	184.536	-16.677	< 0.001 *
2	ar1	0.404	0.020	20.246	< 0.001 *
2	ar2	0.159	0.020	8.121	< 0.001 *
2	sma1	-0.937	0.014	-66.980	< 0.001 *
2	xcos	-8233.444	347.994	-23.660	< 0.001 *
2	xsin	-2935.090	355.191	-8.263	< 0.001 *
3	ar1	0.433	0.020	21.456	< 0.001 *
3	ar2	0.148	0.020	7.537	< 0.001 *
3	sma1	-0.938	0.014	-66.585	< 0.001 *
3	xcos	-8349.112	350.184	-23.842	< 0.001 *
3	xsin	-2886.126	357.053	-8.083	< 0.001 *
3	Christmas	16898.409	1716.466	9.845	< 0.001 *
3	MayBank	-3618.245	1625.498	-2.226	0.026 *
3	EasterMonday	-7462.011	1624.024	-4.595	< 0.001 *
3	GoodFriday	-4685.367	1622.325	-2.888	0.004 *
3	NewYears	-4568.027	1645.192	-2.777	0.006 *
3	SpringBank	-2880.355	1623.757	-1.774	0.076
3	SaintStephen	-3220.402	1717.179	-1.875	0.061
3	SummerBank	-5062.026	1170.476	-4.325	< 0.001 *
4	ar1	0.458	0.020	23.093	< 0.001 *
4	ar2	0.164	0.020	8.365	< 0.001 *
4	sma1	0.085	0.018	4.667	< 0.001 *
4	intercept	22152.586	352.008	62.932	< 0.001 *
4	xcos	-8343.815	383.842	-21.738	< 0.001 *
4	xsin	-3018.203	384.571	-7.848	< 0.001 *

Variation	Parameter*	Estimate	Standard Error	Statistic (t)	Significance (p)	
4	Christmas	17648.275	1722.997	10.243	< 0.001	*
4	MayBank	-3708.100	1641.147	-2.260	0.024	*
4	EasterMonday	-7358.955	1640.072	-4.487	< 0.001	*
4	GoodFriday	-4598.335	1639.183	-2.805	0.005	*
4	NewYears	-4218.930	1656.193	-2.547	0.011	*
4	SpringBank	-2940.528	1639.025	-1.794	0.073	
4	SaintStephen	-2452.233	1717.505	-1.428	0.154	
4	SummerBank	-4992.093	1172.091	-4.259	< 0.001	*
4	Mon	5267.921	318.385	16.546	< 0.001	*
4	Tue	6978.052	344.579	20.251	< 0.001	*
4	Wed	7009.464	369.001	18.996	< 0.001	*
4	Thu	7072.915	368.906	19.173	< 0.001	*
4	Fri	5877.759	345.973	16.989	< 0.001	*
4	Sat	1316.236	310.363	4.241	< 0.001	*
5	ar1	0.428	0.020	21.108	< 0.001	*
5	ar2	0.144	0.020	7.274	< 0.001	*
5	sma1	-0.938	0.014	-68.018	< 0.001	*
5	xcos	-7587.576	1359.727	-5.580	< 0.001	*
5	xsin	-3121.917	1386.168	-2.252	0.024	*
5	Christmas	17279.933	1716.840	10.065	< 0.001	*
5	MayBank	-3390.594	1635.205	-2.074	0.038	*
5	EasterMonday	-7601.038	1629.560	-4.665	< 0.001	*
5	GoodFriday	-4768.319	1622.341	-2.939	0.003	*
5	NewYears	-5685.336	1740.425	-3.267	0.001	*
5	SpringBank	-2729.480	1626.674	-1.678	0.094	
5	SaintStephen	-2854.178	1717.638	-1.662	0.097	
5	SummerBank	-4931.833	1178.408	-4.185	< 0.001	*
5	Jan	-431.076	1131.334	-0.381	0.703	
5	Mar	160.115	1134.371	0.141	0.888	
5	Apr	771.936	1618.989	0.477	0.634	
5	May	-409.527	2071.571	-0.198	0.843	
5	Jun	108.900	2448.365	0.045	0.965	
5	Jul	725.752	2692.655	0.270	0.788	
5	Aug	-395.573	2792.969	-0.142	0.887	
5	Sep	280.040	2723.566	0.103	0.918	
5	Oct	130.054	2487.129	0.052	0.958	
5	Nov	-1049.940	2116.922	-0.496	0.620	
5	Dec	-3019.051	1673.129	-1.804	0.071	
6	ar1	0.433	0.020	21.452	< 0.001	*
6	ar2	0.149	0.020	7.546	< 0.001	*
6	sma1	-0.906	0.019	-47.827	< 0.001	*
6	sma2	-0.048	0.020	-2.348	0.019	*
6	xcos	-8353.767	349.218	-23.921	< 0.001	*
6	xsin	-2877.137	354.852	-8.108	< 0.001	*
6	Christmas	16936.373	1720.125	9.846	< 0.001	*
6	MayBank	-3636.140	1627.477	-2.234	0.026	*
6	EasterMonday	-7437.722	1626.699	-4.572	< 0.001	*
6	GoodFriday	-4711.016	1625.385	-2.898	0.004	*
6	NewYears	-4496.349	1649.277	-2.726	0.006	*
6	SpringBank	-2917.319	1625.521	-1.795	0.073	
6	SaintStephen	-3097.927	1717.723	-1.804	0.071	
6	SummerBank	-5028.988	1166.150	-4.313	< 0.001	*

Variation	Parameter*	Estimate	Standard Error	Statistic (t)	Significance (p)
7	ar1	0.425	0.020	20.882	< 0.001 *
7	ar2	0.119	0.022	5.520	< 0.001 *
7	ar3	0.069	0.020	3.422	0.001 *
7	sma1	-0.946	0.014	-66.553	< 0.001 *
7	xcos	-8353.084	368.114	-22.692	< 0.001 *
7	xsin	-2881.888	374.920	-7.687	< 0.001 *
7	Christmas	16707.240	1722.307	9.701	< 0.001 *
7	MayBank	-3612.078	1632.272	-2.213	0.027 *
7	EasterMonday	-7597.103	1634.320	-4.649	< 0.001 *
7	GoodFriday	-4827.817	1632.500	-2.957	0.003 *
7	NewYears	-4587.981	1649.838	-2.781	0.006 *
7	SpringBank	-2997.154	1630.235	-1.839	0.066
7	SaintStephen	-3436.734	1723.689	-1.994	0.046
7	SummerBank	-5269.443	1174.276	-4.487	< 0.001 *
8	ar1	0.437	0.020	21.822	< 0.001 *
8	ar2	0.124	0.022	5.791	< 0.001 *
8	ar3	0.082	0.020	4.089	< 0.001 *
8	sma1	0.071	0.020	3.502	0.001 *
8	sma2	0.106	0.019	5.588	< 0.001 *
8	intercept	22154.343	390.022	56.803	< 0.001 *
8	xcos	-8333.066	435.242	-19.146	< 0.001 *
8	xsin	-3018.935	436.838	-6.911	< 0.001 *
8	Christmas	17188.106	1714.844	10.023	< 0.001 *
8	MayBank	-3462.025	1632.721	-2.120	0.034 *
8	EasterMonday	-7390.955	1636.999	-4.515	< 0.001 *
8	GoodFriday	-4828.820	1634.344	-2.955	0.003 *
8	NewYears	-4437.680	1646.267	-2.696	0.007 *
8	SpringBank	-2661.324	1630.616	-1.632	0.103
8	SaintStephen	-3181.705	1712.365	-1.858	0.063
8	SummerBank	-4984.269	1157.715	-4.305	< 0.001 *
8	Mon	5260.518	342.636	15.353	< 0.001 *
8	Tue	6981.580	372.153	18.760	< 0.001 *
8	Wed	7015.410	380.787	18.423	< 0.001 *
8	Thu	7076.086	380.688	18.588	< 0.001 *
8	Fri	5884.670	373.447	15.758	< 0.001 *
8	Sat	1317.072	335.375	3.927	< 0.001 *

* Sun and Feb were excluded as baseline values in the model for day of week and month.

Figures

Figure 2.

Fourier Analysis of the Detrended Bike Hires

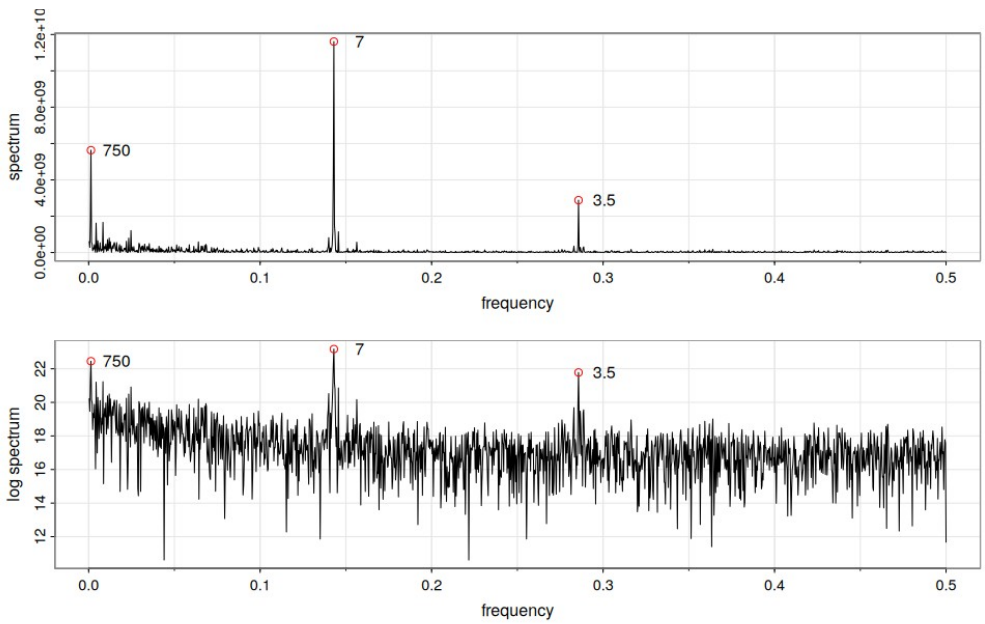


Figure 3.

ACF and PACF of the Detrended Bike Hires

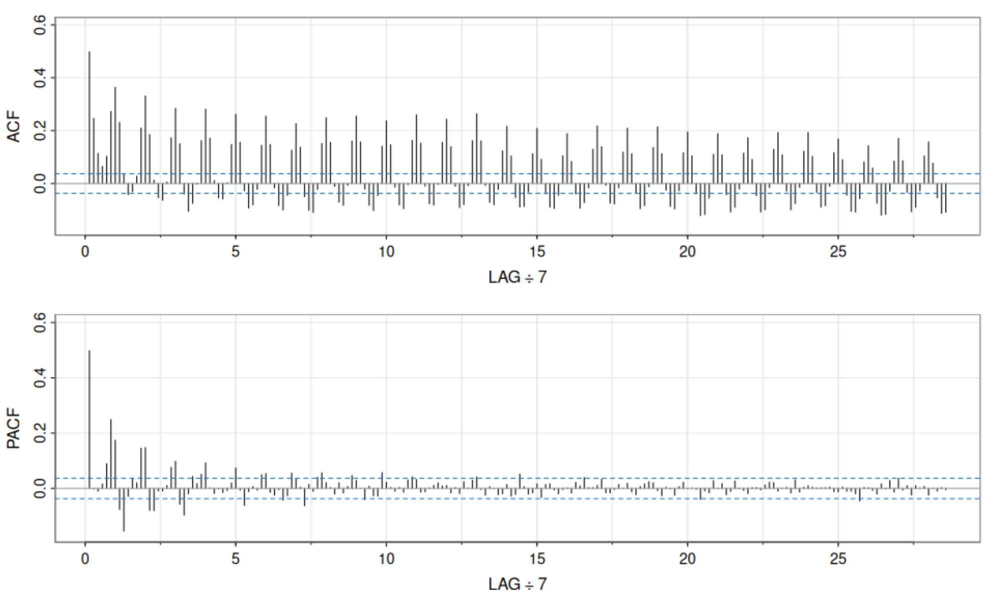


Figure 4.

ACF and PACF with Seasonal Integration of the Detrended Hires

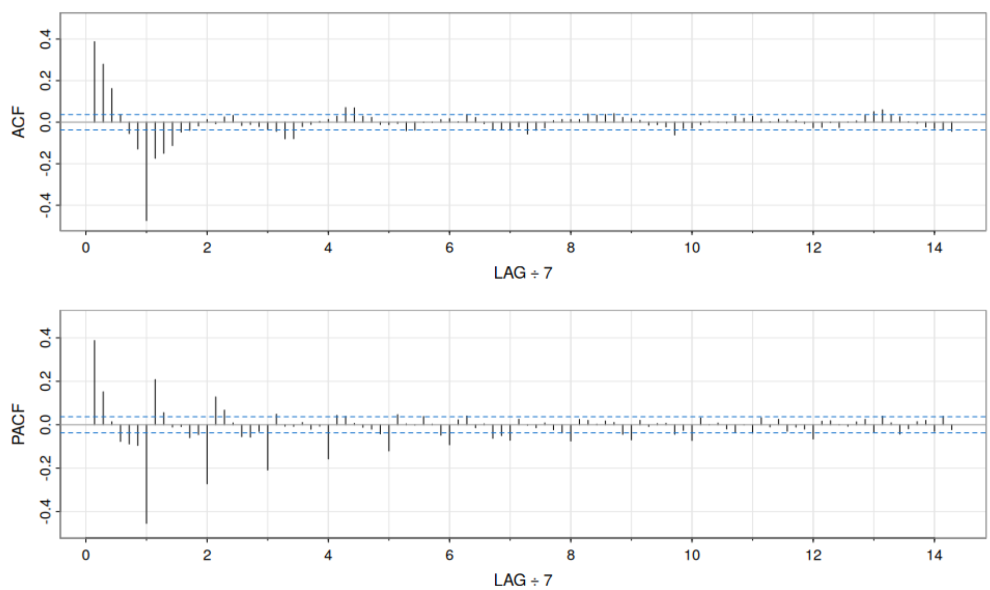
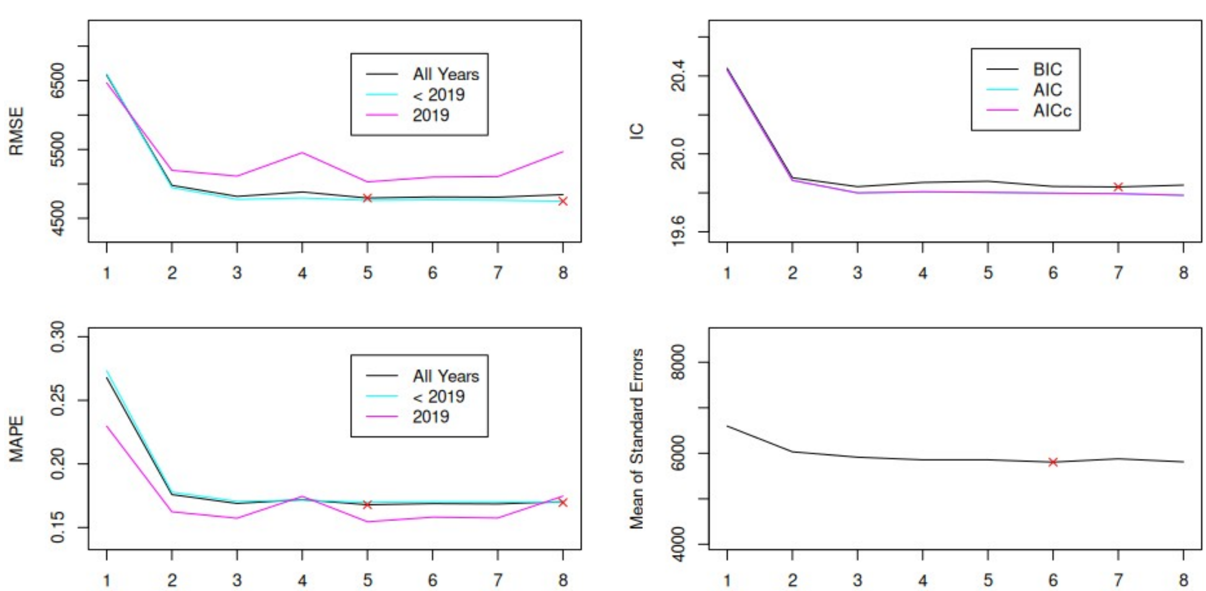


Figure 5.

Model Variation Comparisons



Note: The red X indicates the lowest value for the measure in the training data set (2012–2018).

Figure 6.

Diagnostic Plots for Selected Model

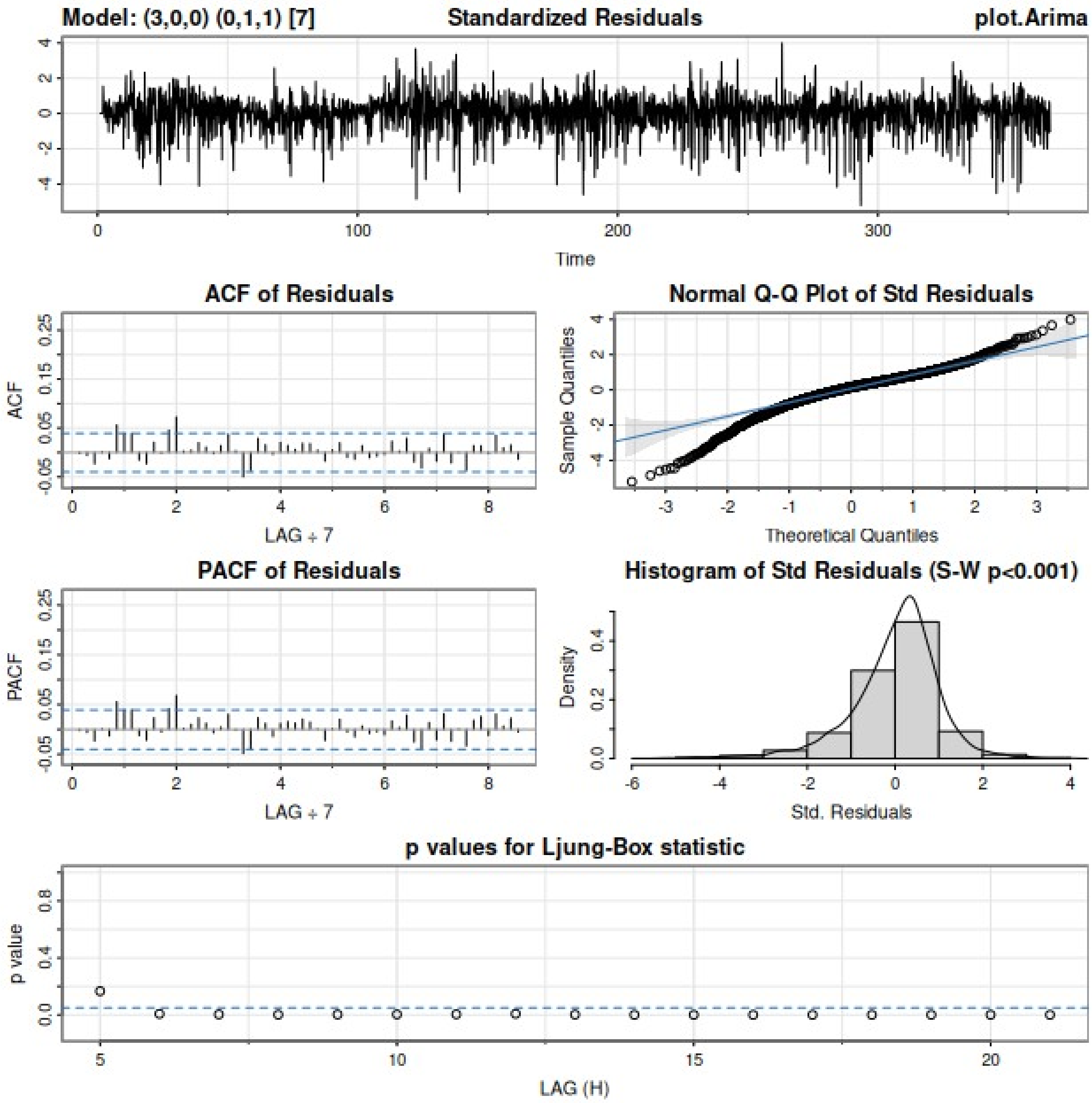


Figure 8.

Model Variation Forecasts Compared to Actual