**Forecasting Utilization Rates**

By Sherri McRae

*In partnership with the Victoria Police Forensic Services Centre*

*Special thanks to Chris Bell, Intelligence & IT Group Manager of the Victoria Police Forensic Services Centre and RMIT University for providing this placement opportunity.*

**TABLE OF CONTENTS**

Executive Summary . . . . . . . . . . . . . . . . . . . . 3

Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . 3

Methods and Materials . . . . . . . . . . . . . . . . . 4

Literature Review . . . . . . . . . . . . . . . . . . . . . . 7

Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16

Recommendations . . . . . . . . . . . . . . . . . . . . . 25

References . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

Appendix . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

**Disclaimer**: I declare the following to be my own work, unless otherwise referenced, as defined by the University’s policy on plagiarism.

**EXECUTIVE SUMMARY**

The Victoria Police Forensic Services Centre maintains a database that contains the case numbers of samples that have been submitted for forensic testing along with their associated offence categories. The purpose of this project was to forecast the utilization rates of different offence categories for the next two years. Utilization rates were calculated by dividing the number of cases in the Forensic Services Centre database with the total number of cases in Victoria. Forecast models were produced using SAS procedures. PROC ARIMA produces ARIMA and seasonal ARIMA models. PROC ESM produces exponential smoothing models. Crime data often contains a lot of fluctuations. This was further complicated by the large number of missing values and an issue with individual cases reoccurring in successive months. Ten offence categories were analysed. Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), AIC and BIC were employed to compare prospective forecast models. Eight of the ten best models were ARIMA or seasonal ARIMA models. The remaining two were exponential smoothing models. MAPE scores of the best models ranged from 6.91 to 21.88. Arson was the only offence category of those studied in which utilization rates are predicted to increase over the next two years. The utilization rates of all other categories studied are expected to remain stable. Recommendations include addressing the issue with reoccurring cases in the data, and testing forecast accuracy by averaging the quality scores from a number of differently sized validation sets.

**INTRODUCTION**

The Victoria Police Forensic Services Centre offers a range of forensic testing to support police investigations. Requests for services are entered onto a database. Select information from these requests are collated into a summary table in SAS. Each observation in the table represents one request and includes information such as; request type, case number, container number, item number, creating unit, request start and end dates, and associated offence category. Each week the table is updated, and the number of observations increases. During this project, there were just over one million rows with requests dating back to 2005. The summary table has the potential to provide insight into trends occurring across different groups in the department and assist in decision making with regards to budgets, planning and resource allocation.

The aim of this project was to develop forecasting models for utilization rates in select offence categories. Utilization rates are a ratio of the number of cases received by the Forensics Services Centre and the number of cases in Victoria as a whole. The Victorian case counts were provided by a SAS query to another Victoria Police database. The forecast models were developed using the SAS procedures PROC ARIMA and PROC ESM, which produce ARIMA and exponential smoothing models respectively. PROC ESM was of particular interest because this procedure is easy to use. The models were compared using quality metrics to determine the best. The offence categories investigated were; homicide, assault, burglary (aggravated), burglary (residential), drugs (possess/use), drugs (cultivation/ manufacture/ traffic), rape, arson, and theft from car. A dataset combining all offence categories into one for overall utilization rates was also analysed.

**LITERATURE REVIEW**

Gorr and Harries (2003) give an overview of crime forecasting with the specific intention of supporting planning decisions and tactical deployment of police resources. They mention a number of techniques that have been used successfully, including; ARIMA models, exponential smoothing, neural networks, regression, naïve time series methods and spatio-temporal models that make use of software such as Geographical Information System (GIS).

Two common times series models are ARIMA and exponential smoothing. Donnelly & Wan (2016) of the NSW Bureau of Crime Statistics and Research found that ARIMA models produced better long-term forecasts than exponential smoothing models when predicting prison populations in NSW. In this instance, long term referred to forecasts of 6-12 months. In short-term forecasting, ARIMA models and exponential smoothing models produced similar results.

Chen, Yuan & Shu (2008) also showed that the ARIMA model had greater accuracy than exponential smoothing when forecasting a week ahead with 50 weeks of historical data. The data was a univariate time series of crime events in one city in China.

Borowik et al. (2018) highlight that crime data presents challenges for analysts. It is often collected for purposes other than data analysis. As a result, the content and structure of such data can be less than ideal. Missing values are a common problem. Crime data analysis therefore tends to involve extensive pre-processing; validation, cleaning, collating and variable selection.

An important consideration in the pre-processing stage is detecting and managing outliers. Agnieszka and Magdalena (2018) advise that visual methods of outlier identification in a univariate time series are the easiest and most effective. These methods include boxplots, histograms and a scatter plot of the variable of interest against time. Arumugam and Saranya (2018) recommend standardizing the residuals to identify outliers. A popular method of fixing problematic outliers is to replace them with the mean of the series. Another method is the use of the Likelihood Ratio Criteria created by A.J. Fox (Agnieszka & Magdalena, 2018).

There is some debate as to how much impact correcting outliers can have on forecast accuracy. Akpan, Lasisi and Adamu (2019) report that detecting and fixing innovation outliers (outliers with lingering effects on the data) results in only a slight improvement in forecasting accuracy. Arumugam and Saranya (2018), however, conclude that SARIMA models achieved significant improvement by replacing outliers with the mean of the series.

Another method for improving accuracy in forecasting models is to take the log of the series. Lütkepohl and Xu (2012) report that that log transformations are good only if they stabilize the variance. If they do not they can damage the forecasting precision. They also point out that applying an exponential function to a log transformed series is not an efficient way to convert the log transformed variables and associated predictions back to their original form. It is recommended that ½σ2 is added to the forecast value before applying an exponential function (Lütkepohl and Xu, 2012). However, caution is advised as this transformation will still not be optimal for some datasets.

Time series crime data may include seasonality. Chen and Lee (2017) report that crime rates are linked to temperature. They found that sexual offences, drug offences and domestic violence related assaults are more likely to occur in the summer. Borowik et al. (2018) showed that ‘hooliganism’ was more common in the summer months. Burglary is more likely on Mondays between 10am and 3pm while traffic violations are most common during peak hour on weekday mornings.

The most important and possibly the most difficult part of developing forecast models is assessing the accuracy of the model. No single metric can provide a clear-cut evaluation of the predictive capabilities of a model (Tularam & Saeed, 2016). As a result, a variety of measures are used in assessing the quality of a forecasting model. Common metrics include the Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE) and the Mean Absolute Error (MAE). Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) scores are also popular as they provide a single value to assess both goodness of fit and the simplicity or parsimony of a model (Awang, Yong & Hoeng, 2017).

Test statistics such as RMSE and MAPE are calculated by setting aside the last few observations of a series and comparing these values to the forecast model predictions. Athanasopoulos and Weatherburn (2018) who studied the prison populations in New South Wales, took a portion of the total data available (36 months) as the initial training set and expanded this set one month at a time while keeping the validation set the same size. As each new data point was added to the training set the MAE was recalculated. Donnelly and Wan (2016) who also studied the prison populations of New South Wales, used a similar rolling validation technique to assess forecasting models. While studying inflows into a reservoir system, Mohan and Vedula (1995) used 27 years of monthly data from a dataset spanning 52 years for validation of their forecasting models.

Model selection involves more than test statistics. Residuals should be normally distributed and uncorrelated. The principle of parsimony should be observed in that simpler models are preferred over more complex models. Overfitting can be avoided by making sure that parameter estimates significantly greater than zero and that there is no correlation between them. If correlation exists the model contains redundant parameters and one or more may be removed (Mohan & Vedula, 1995).

**MATERIALS AND METHODS**

The calculation of utilization rates required that a SAS table from a different branch of the Victoria Police be joined to the summary table from the Forensics Services Centre.

Before this was accomplished, the summary table required processing. It was first joined with two other SAS tables from the Forensic Services Centre, one that linked the requests to offence categories and the other linked requests to the category ‘enquiry’. The tables were joined with a composite key that consisted of profile request type, container number and item number. Before the join, each observation in the summary table represented a single request. After the join, a request could occur more than once because it can be associated with more than one offence category. There is a total of 27 offence categories plus ‘enquiry’ which is an additional category created for the purposes of the forensic department.

Utilization rates are calculated using case counts. However, each case can include more than one request. Some cases include more than one thousand requests. This creates a situation in which a case number can reoccur many times in a single month and across successive months as related samples continue to be collected and submitted by police over the course of an investigation. Data cleaning was performed to address any repetition of cases within each month. However, the work required to prevent a case from re-occurring in successive months would have been time-consuming, and the decision was made to not correct this issue.

There was much missing data in the summary table. Many case listings did not have an offence category entered. An effort was made to mitigate the problem by making assumptions as to what the missing values might be by looking at other information available for each observation. For example: if the offence category was missing and the unit was listed as “Chemical Drug Intelligence” then the offence category was assumed to be “Drugs (possess/use)”. There were 143 instances of request without case numbers.

Originally, the forecast models were to be created using yearly datapoints. As there were only 14 years of data, it was decided that it should be split up into monthly intervals, to increase the amount of historical data available from which to build a forecast. Breaking up the data in this way also had the effect of exposing any seasonality that would have otherwise been obscured.

The data for the overall case counts in Victoria was initially going to be sourced from the Crime Statistics Agency. However, closer inspection of the data revealed that it was not possible to compare the data from the Forensic Services Centre with that of the Crime Statistics Agency because of a difference in the way offence categories where defined. The Victorian case counts data from the Victoria Police uses the same offence category definitions as the Forensic Services Centre. However, the way ‘case’ is defined varies slightly between the two datasets.

The Victorian case counts dataset needed processing before being joined to the other SAS tables. The offence categories were reformatted. The categories burglary (residential) and burglary (other) were combined as burglary (other) did not occur in the Forensic Services Centre summary table. The final resulting table after pre-processing was complete contained month and year, number of cases in the Forensic database, number of cases in Victoria, offence category and utilization rate. This master table was broken down into specific offence categories.

Forecast models were created using the SAS procedures PROC ARIMA and PROC ESM. The standard methodology was used to create ARIMA models. This involved checking for trend, seasonality and stationarity. The data was assessed as to whether a transform was required. Two options; ‘scan’ and ‘minic’, were used to help select viable models along with a visual inspection of ACF and PACF plots. Each prospective model was assessed be checking to see if all autoregressive and moving average components were significant and that the residuals were normally distributed and uncorrelated. Having passed this initial testing, the quality metrics were recorded for use in comparisons with other models. The quality metrics used were root mean square error (RMSE), Mean Absolute Percent Error (MAPE), AIC and BIC. As the SAS procedures for ARIMA and exponential smoothing calculate AIC and BIC in different ways, these metrics were only used to determine the best model of a particular type and not to compare ARIMA and exponential smoothing models. The quality metrics were taken for both the data and a validation set of 10 observations. As time was limited, forecasts were tested with only one validation set, except in the case of residential burglary, where the quality metrics of 5 validation sets of different sizes were recorded for each prospective model.

The exponential smoothing models required a selection of a model type and a particular transform where necessary. The four exponential smoothing models that were considered included ‘simple’, ‘linear’, ‘addwinters’ and ‘winters’. The ‘addwinters’ model creates an additive seasonal model. The ‘winters’ uses the Holt-Winters triple exponential smoothing model. The ‘linear’ model creates a Holt linear model and ‘addwinters’ uses an additive Winters model. The fit of the model was assessed by inspection of the normality plots and residual lag plots. A validation set of 10 observations was used in all cases except residential burglary as noted previously. The same quality metrics were used for the exponential smoothing models as with the ARIMA models.

**RESULTS**

The best forecasts were determined based on Root Mean Square Error (RMSE) and Mean Absolute Percent Error (MAPE) scores for the validation set. Table 1 lists the best model type for each offence category and their respective quality metrics. The best MAPE scores are 6.91 for the drugs (cult/manuf/traf) model, 7.72 for the residential burglary model and 8.4 for the overall model. The worst MAPE scores are for the models for arson (21.88) and rape (18.98). Eight of the models are ARIMA or SARIMA models and two are exponential smoothing models.

The best model was sometimes difficult to determine because the RMSE and MAPE scores were very similar between two models and perhaps the RMSE was slightly better in one model but the MAPE was slightly better in the other. This was the case for residual burglary. To resolve the issue the procedures for each prospective model were run 5 times with validation sets of different sizes and the averages for each metric calculated. The model with the best averaged score was determined to be the best model. The best forecast model for the offence category rape did not have normally distributed residuals as there were some extreme outliers in the data (see appendix 1.H.). The residual lags, however, were uncorrelated.

|  |  |  |  |
| --- | --- | --- | --- |
| **BEST MODELS** | | | |
| **Offence Category** | **Model Type** | **Forecast** | |
| **RMSE** | **MAPE** |
| Overall | SARIMA(2,1,0)x(0,1,1)12 | 0.00497 | 8.4 |
| Burglary (Aggravated) | esm: addwinters | 0.04327 | 14.08 |
| Drugs (Possess/Use) | SARIMA(1,0,1)x(0,1,1)12 | 0.09331 | 16.5 |
| Burglary (Residential) | ARIMA(0,1,1) | 0.00566 | 7.72 |
| Homicide | esm: simple | 0.1104 | 12.94 |
| Theft From Cars | ARIMA(0,1,1) | 0.00169 | 16.67 |
| Drugs (Cult/Manuf/Traff) | ARIMA(1,1,1) | 0.02514 | 6.91 |
| Assault | ARIMA(1,0,1) | 0.00205 | 12.42 |
| Arson | ARIMA(0,1,1) + log transform | 0.06978 | 21.88 |
| Rape | ARIMA(0,1,1) | 0.04038 | 18.98 |

Table 1: Best forecast models for each offence category.

The forecast plots for each of the best models appear below. Circles represent actual data points. The vertical dashed line represents the beginning of the validation set. Circles to the right of the dashed line are actual data points used in the validation sets. The continuous blue line represents the mean of the model and then the forecast. The shading represents the 95% confidence interval.

According to these forecasts, it is expected that utilization rates for all offence categories except for arson will remain steady. Utilization rates for arson are expected to increase in the next two years from mean of approximately 35% to almost 40%.

The 95% confidence intervals for these forecasts can be quite wide. For example; in drugs (possess/use) the confidence interval of the forecast extends from utilization rates of 15% to 65%. The confidence interval for utilization in cases of rape is between 0% and 50%.

As cases can reoccur in subsequent months it is possible for the case count taken from the Forensic Services Centre database to be more than the case count for the same offence category in all of Victoria. As a result, the utilization rates for homicide have a mean of 2, which is a 200% utilization rate. This is merely an effect of the data processing. Samples from homicide cases always receive analytical testing at the Forensic department so the expected utilization rate is 100%.

Seasonal components were found in three categories: overall, drugs (possess/use) and burglary (aggravated). Plots of residuals and raw data can be found in Appendix 1.

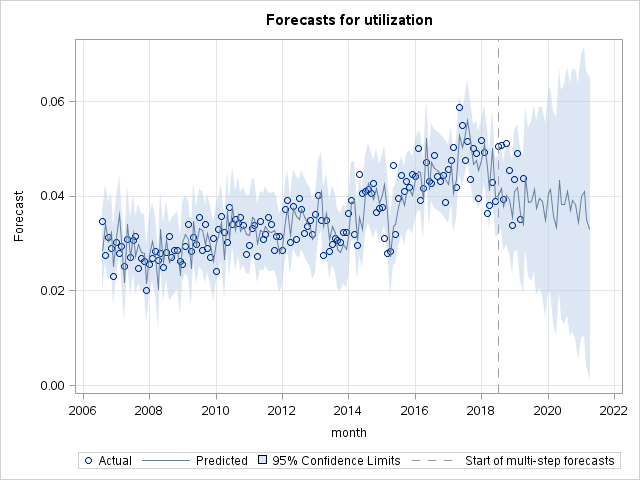


Figure 1: Overall utilization rates forecast model

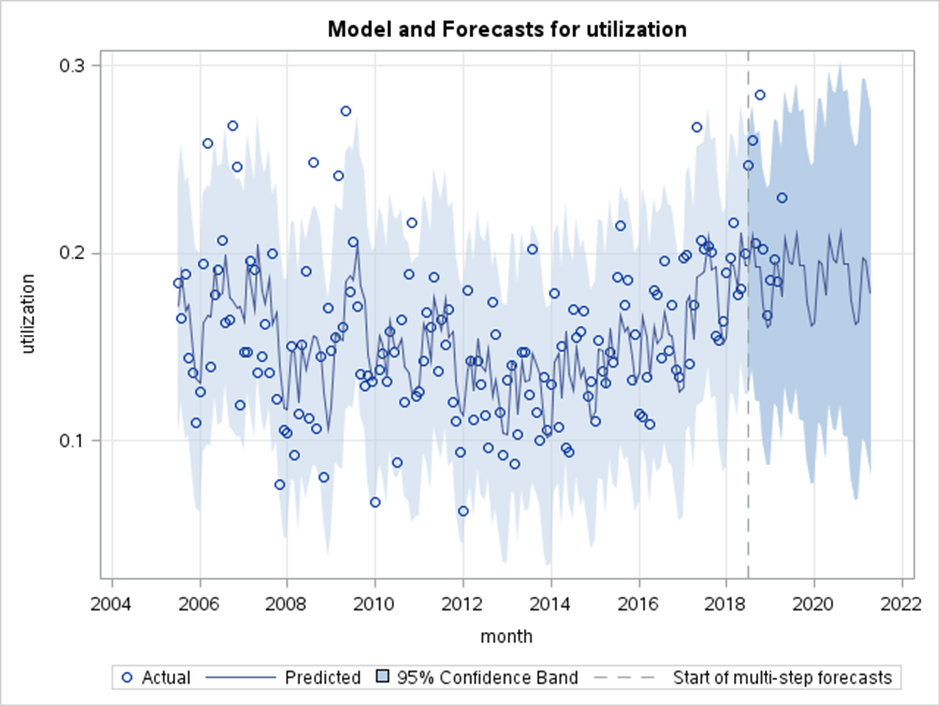


Figure 2: Burglary (aggravated) utilization rates forecast model

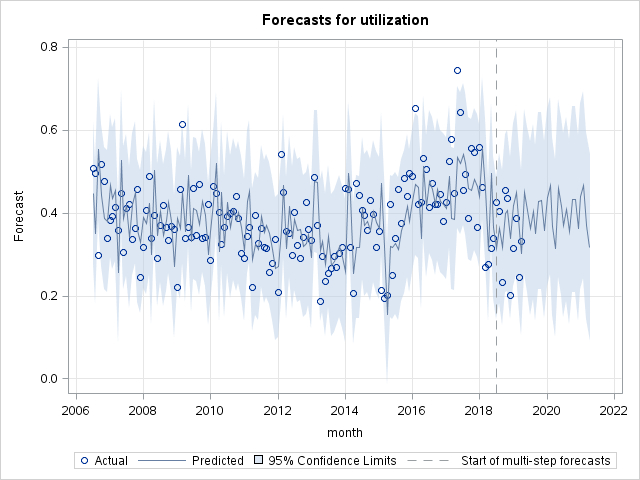


Figure 3: Drugs (possess/use) utilization rates forecast model

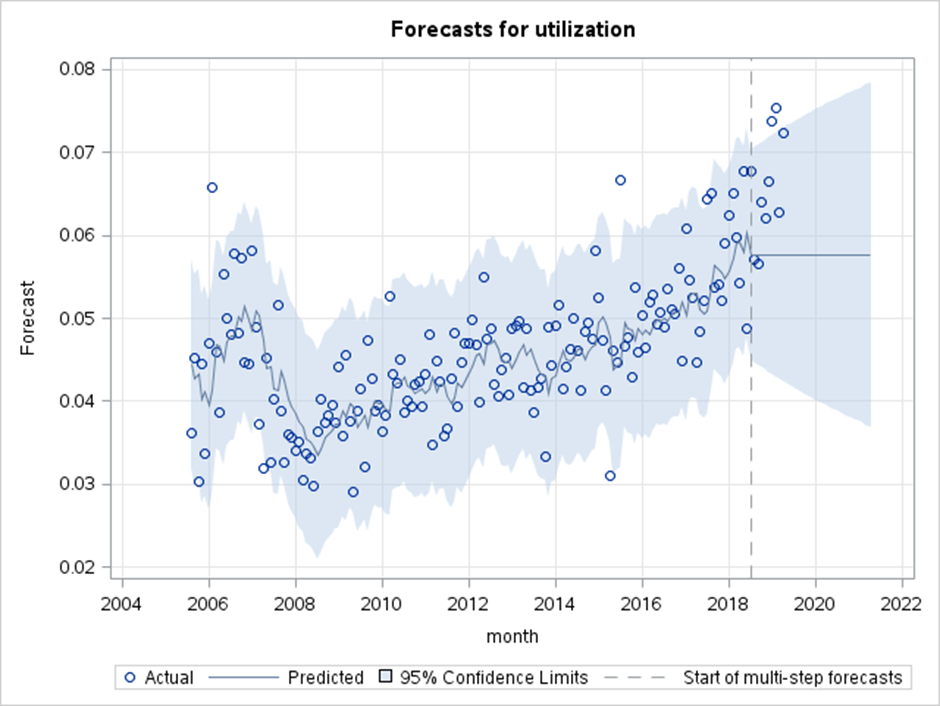


Figure 4: Burglary (residential) utilization rates forecast model

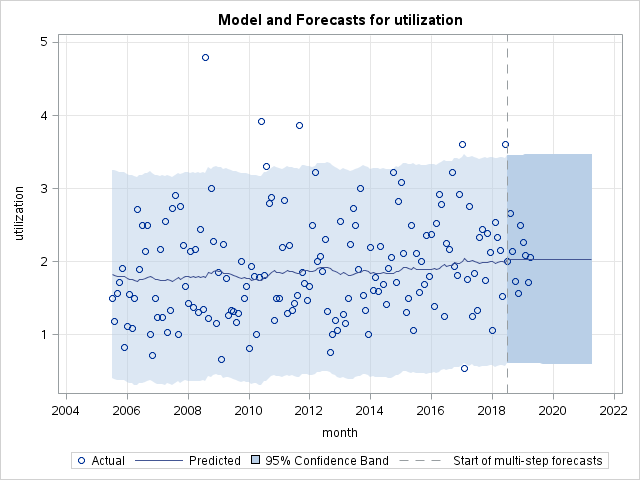


Figure 5: Homicide utilization rates forecast model

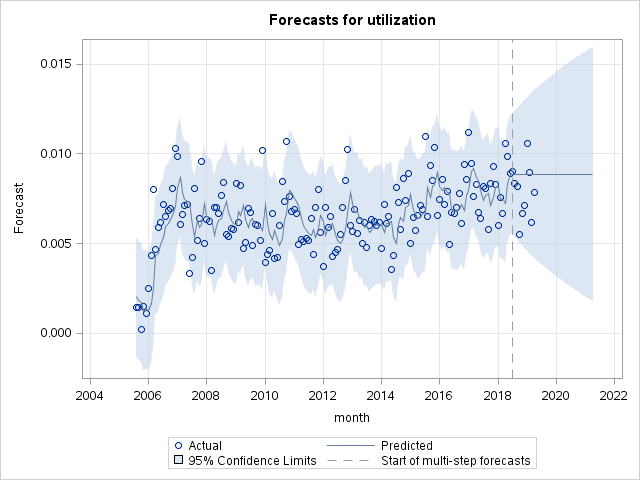


Figure 6: Theft from cars utilization rates forecast model

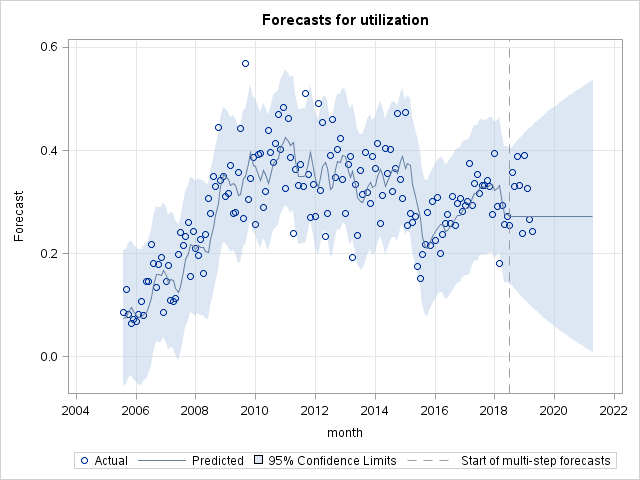


Figure 7: Drugs (cult/manuf/traf) utilization rates forecast model

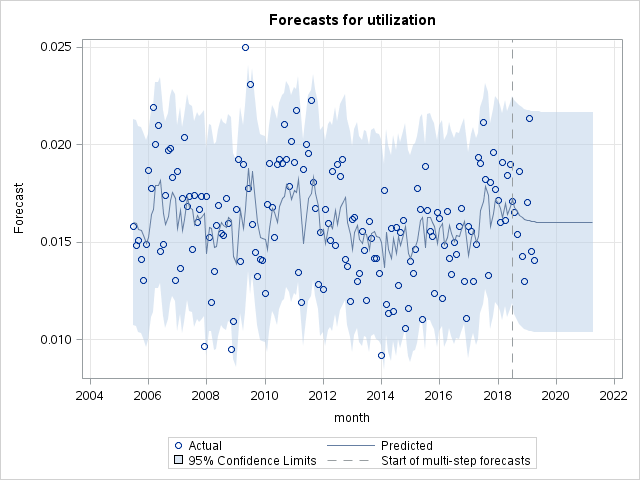


Figure 8: Assault utilization rates forecast model

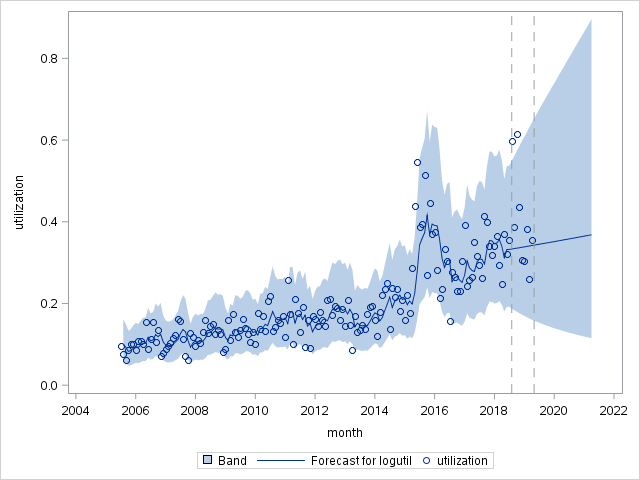


Figure 9: Arson utilization rates forecast model

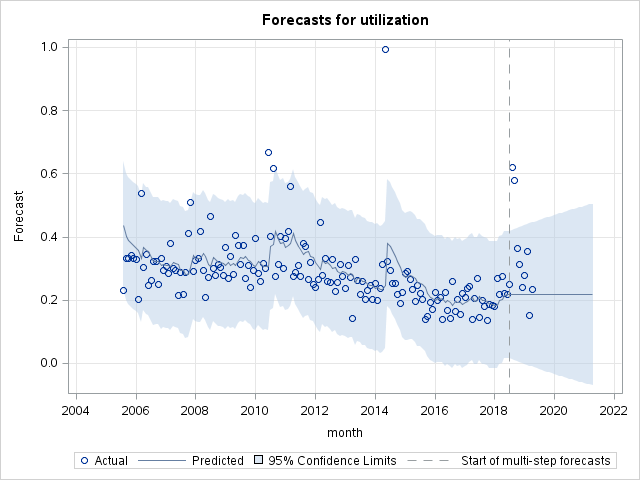


Figure 10: Rape utilization rates forecast model

**DISCUSSION**

As has been mentioned previously, forecast models were created from datasets of monthly data in which an individual case number could reoccur in successive months. This created noise in the data that may have obscured or complicated the patterns in each series. Initially, request counts rather than case counts were used to create the forecast models. Request counts have the benefit of better reflecting the actual work associated with each offence category because each request is for a particular test. Case counts, as noted previously, can include many requests. It is not possible to compare forensic request counts accurately to Victorian case counts. It is also the custom of the Department of Justice and its subsidiary departments to speak in terms of cases. The use of request counts therefore was not continued. Some forecast models were developed using request counts before this change. The models by request count for drugs (cult/manuf/traf), drugs (possess/use), and overall utilization rates were simple ARIMA models. However, the forecast models by case count for the same categories were more complex seasonal models. It is possible that this additional complexity is at least partially the result of repeated cases.

Missing values can also create noisy or inaccurate data. There were many missing values in the summary table. Some information was lost in past transfers from one database system to another. Some departments habitually enter data onto alternative systems. Sometimes only partial information is entered. A good way to visualize the noise in this dataset is to examine the utilization rates for homicide. Samples from all homicide cases are sent to Forensic Services for analysis. Theoretically, the utilization rates for homicide should be a continuous line at 1 (or 100%). The plot below shows that the utilization rates average around 2 but have a range of between just below 1 and almost as much as 5.

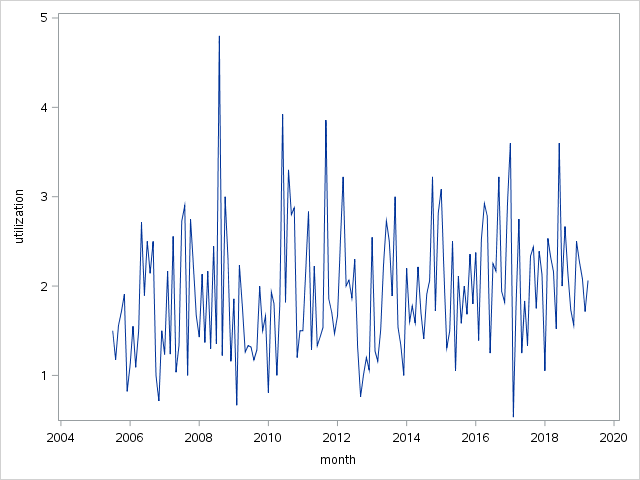


Figure 11: Homicide utilization rates

This noise no doubt occurs in the other offence categories, though it is not possible to see it in the same way. The accuracy of any forecast models will be hindered by it.

The rest of this discussion will focus on examples of the sorts of issues encountered when selecting the best forecast models.

The three forecast models below are for utilization rates for drugs (possess/use). The ARIMA(1,1,1) model (Figure 12) was selected as the best. The residual lags plot for this model shows that lag 24 is significant (Figure 15) which would ordinarily result in this model being disqualified as a contender for best fit. However, the RMSE and MAPE scores for the ARIMA(1,1,1) model are much better than the seasonal ARIMA and exponential smoothing models (Figures 13 & 14). The AIC and BIC scores are better than the SARIMA model (Table 2). The normality plot of residuals shows that the kernel density tightly overlays the normal distribution reference line. The SARIMA and exponential smoothing models are in agreement with the ARIMA(1,1,1) forecast in that they also predict a steady mean utilization rate at around 30% for the next two years. However, they include a seasonal component.

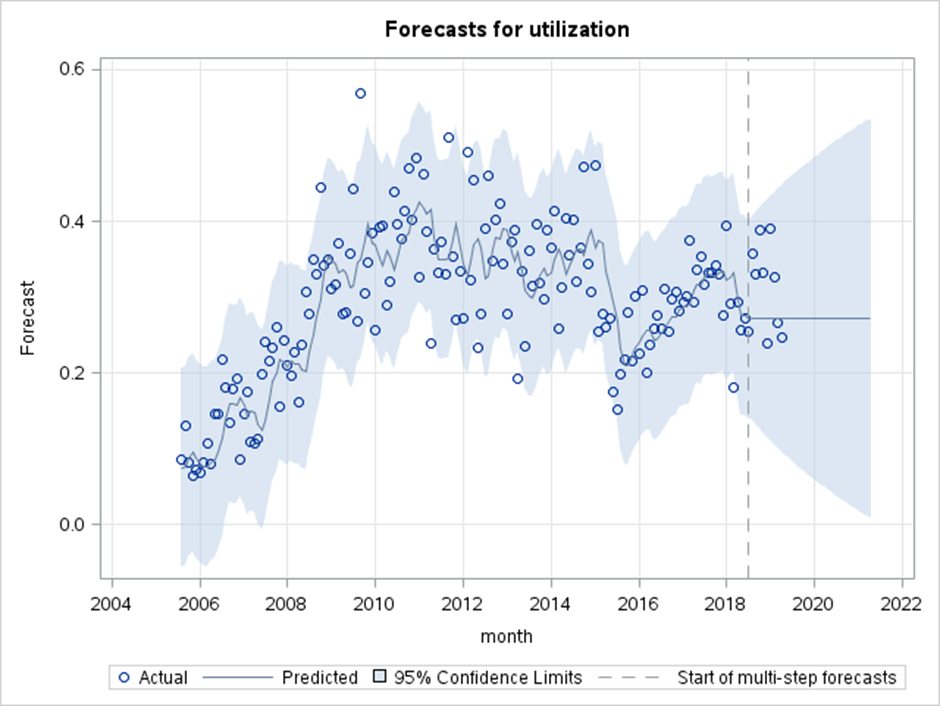


Figure 12:ARIMA (1,1,1) best model for drugs (cult/manuf/traf)

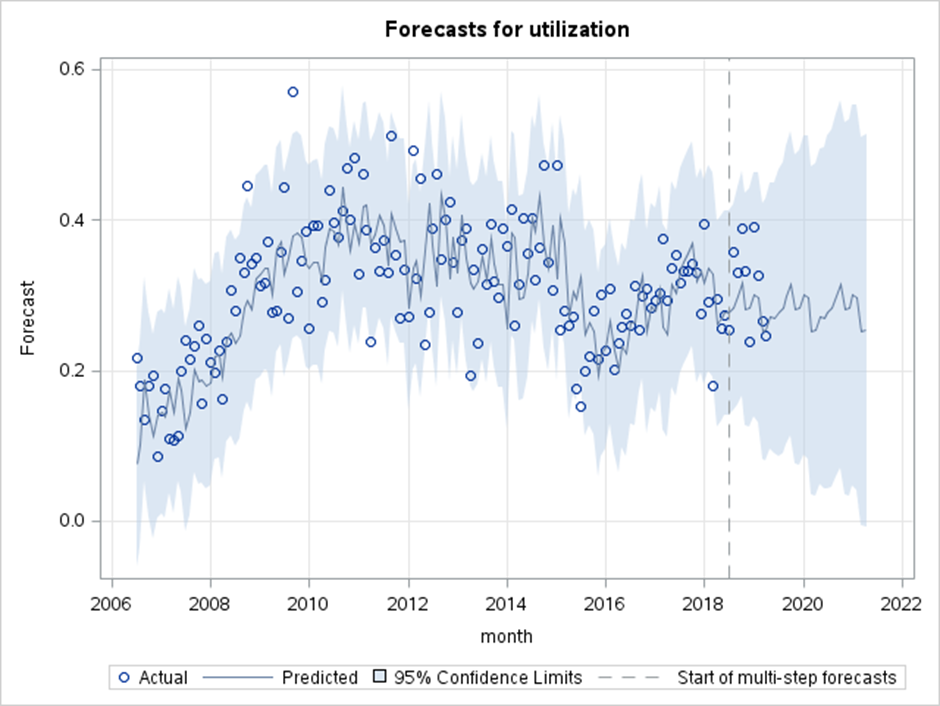


Figure 13: SARIMA(2,0,1)x(0,1,1)12 model for drugs (cult/manuf/traff)

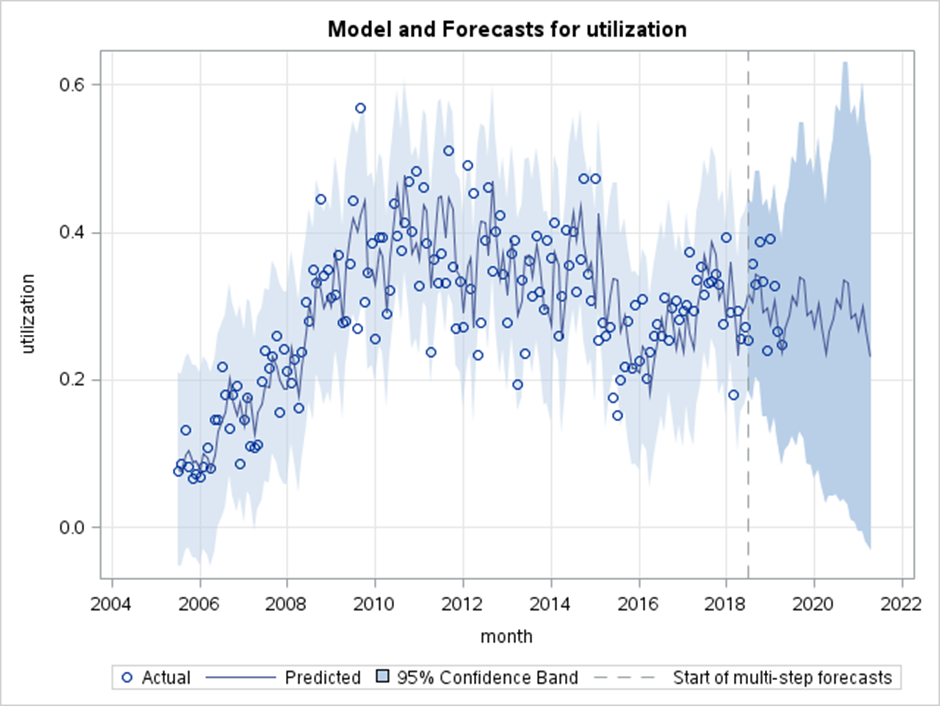


Figure 14: Holt Winters exponential smoothing model for drugs (cult/manuf/traf)

Given that the ARIMA(1,1,1) model is a good model in all but one significant residual lag, that its forecast agrees with the other models and that it is a simpler model which is in keeping with the *Principle of Parsimony*, it was selected as the best model.

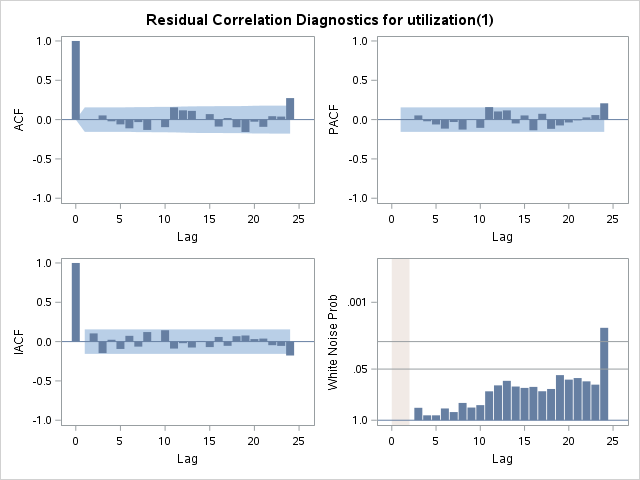


Figure 15: Residual correlations for drugs (cult/manuf/traf)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Forecast Models for Drugs (cult/manuf/traf) Utilization Rates | | | | | | |
| Model Type | Data Fit | | | | Forecast | |
| RMSE | MAPE | AIC | BIC | RMSE | MAPE |
| ARIMA(1,1,1) | 0.06984 | 20.84 | -422.13 | -415.92 | 0.02514 | 6.91 |
| SARIMA(2,0,1)x(0,1,1)12 | 0.06915 | 18.25 | -296.24 | -290.17 | 0.07673 | 30.31 |
| ESM: Winters | 0.0661 | 19.35 | n/a | n/a | 0.05314 | 13.46 |

Table 2: Prospective forecast models for drugs (cult/manuf/traf)

Two prospective forecast models for the offence category arson, one an ARIMA model (Figure 16) and the other a Holt Linear model (Figure 17), produced a very similar forecast but had very different quality scores. In this case, the model with the worse RMSE and MAPE scores was selected as the best model, based on the residual normality plots and lags. While the Holt Linear model had a very low MAPE score of 6.27 for the validation set and 3.02 for the training set (table 3), the kernel density in the residual normality plot did not overlay the reference line well and there were significant lags in the autocorrelation plot (see appendix 1.J.).

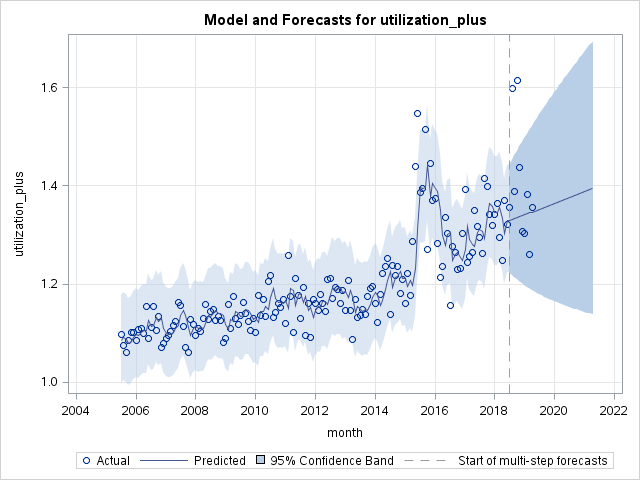
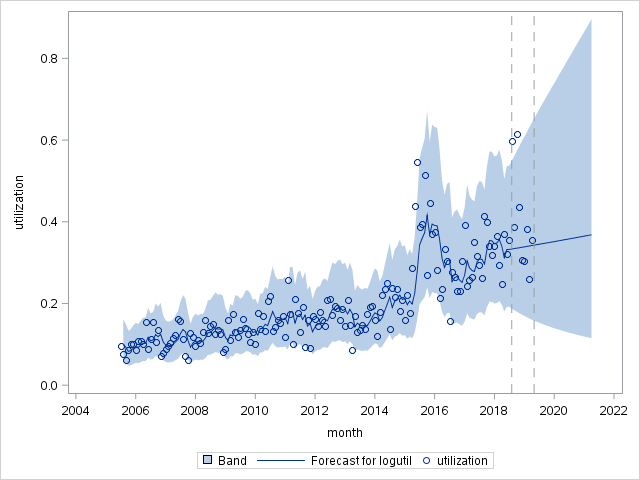


Figure 16: Arson model ARIMA(0,1,1) with log transform

Figure 17: Arson model Holt Linear with log transform

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Forecast Models for Arson Utilization Rates | | | | |
| Model Type | Data Fit | | Forecast | |
| RMSE | MAPE | RMSE | MAPE |
| ESM: linear + log tranform | 0.05276 | 3.02 | 0.13233 | 6.27 |
| ARIMA(0,1,1) + log transform | 0.05222 | 21.47 | 0.06978 | 21.88 |

Table 3: Quality scores for two prospective forecasting models for arson utilization rates

The best model for residential burglary was taken to be an ARIMA(0,1,1) (Figure 18), however it seems to be ignoring an increasing trend from 2008 onwards. Another prospective model was a Holt Linear model (Figure 19) that included a trend line that seems to make sense to the eye. The residual plots for both models were equally good in that they showed that the errors were normally distribute and uncorrelated. However, the quality scores for the Holt Linear model were worse than the ARIMA model. In order to be sure that this was not an effect of the validation set size, the procedures for both of these models were run five times, the quality scores recorded and then averaged (see appendix 2.A.). The best was still the ARIMA model.

If the data had been log transformed an ARIMA(0,1,1) would have resulted in an increasing trend. The process of determining whether a log transform was necessary involved a number of tests. Firstly, a visual comparison of the log transformed data and the original data was conducted to determine if the variance of the log transformed data seemed more constant. SAS includes a function called ‘boxcoxar’ that allows a comparison of AIC and BIC scores between the log transformed data and the original. The TRANSREG procedure suggests a best value for lambda if a Box Cox transformation is appropriate. It was common that these tests did not agree. Log transformations were avoided unless they were obviously effective at stabilizing the variance as suggested Lütkepohl and Xu (2012).

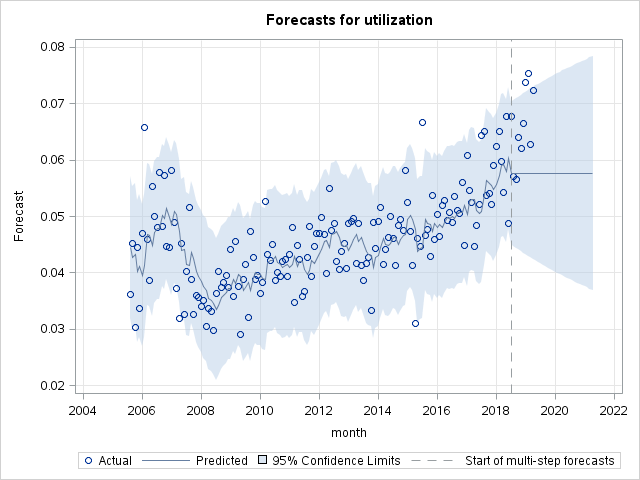
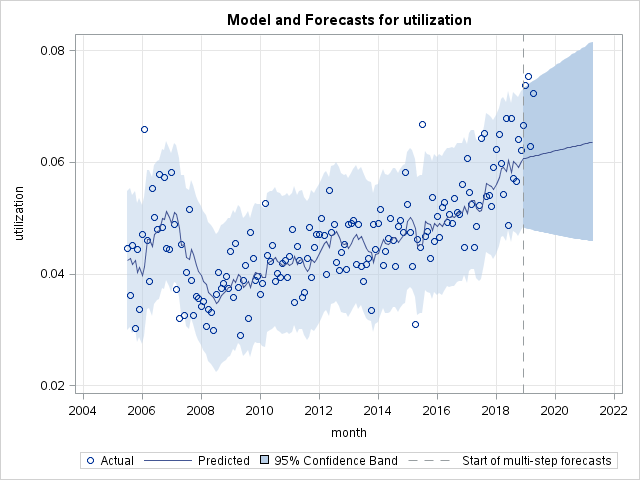


Figure 18: Holt Linear model for residential burglary Figure 19: Best forecast model for residential burglary

The forecast model for the overall utilization rates was difficult to determine. Ultimately, there was only one model with residuals that were normally distributed and uncorrelated, SARIMA(2,1,0)x(0,1,1)12 (Figure 20). This model did not include a log transform.

Another prospective model that did involve a log transform, SARIMA(3,1,1)x(0,0,1)12 (Figure 21), achieved better RMSE and MAPE scores (0.003 and 5.85 respectively, see Table 4), and had normally distributed residuals. This model was rejected because the 10th and 24th lags were significant (Figure 22). It is worth noting that the ARIMA model that was selected as the best for drugs(cult/manuf/traf) also had a significant 24th lag. It was chosen as the best regardless because it had excellent quality scores and its forecast was in keeping with other models that had normally distributed and uncorrelated residuals. While the SARIMA(3,1,1)x(0,0,1)12 model for overall utilization rates (Figure 21) predicted an increase similar to the other log transformed models, none of these models could pass the residual analysis and were therefore deemed untrustworthy. An attempt was made to reduce the correlation at lag 24 by changing the model to include seasonal differencing but no non-seasonal differencing. The resulting model, SARIMA(3,0,1)x(0,1,1)12, was rejected because the residuals had many significant lags and were not normally distributed.

The best model for overall utilization rates, SARIMA(2,1,0)x(0,1,1)12 , had much lower AIC and BIC scores than the other prospective SARIMA models (Table 4). Had there been more time available, the procedures for these models (Figures 20 and 21) would have been run multiple times with different validation set sizes and the quality metrics of these runs would have been averaged.

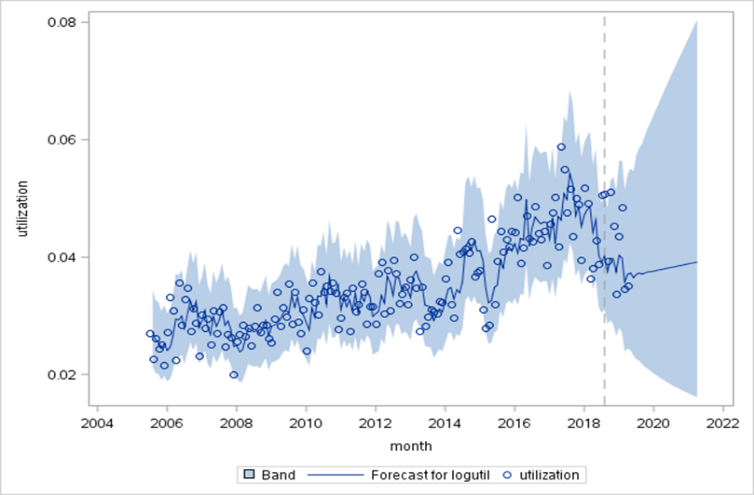
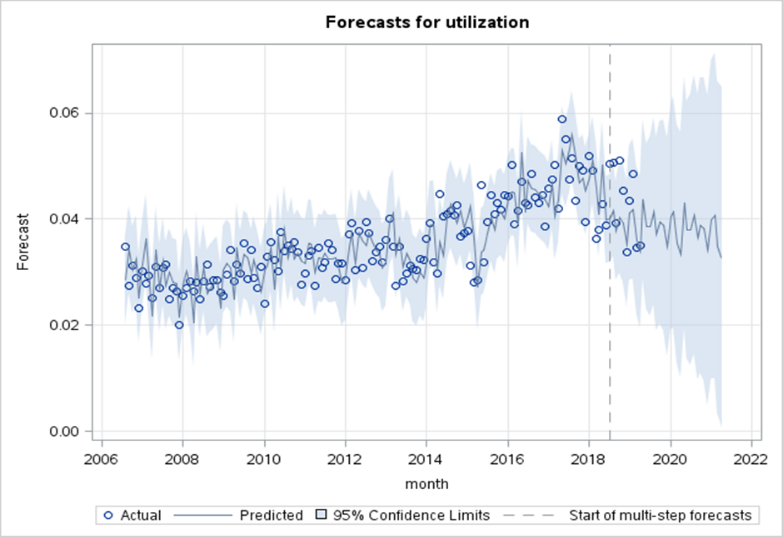


Figure 20: Best model for Overall utilization rates, SARIMA(2,1,0)x(0,1,1)12

Figure 21: Model for Overall utilization rates with best quality scores, SARIMA(3,1,1)x(0,0,1)12

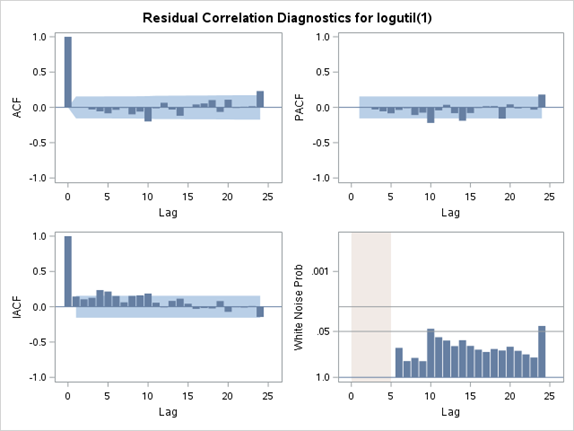


Figure 22: Residual lags for SARIMA(3,1,1)x(0,0,1)12 for Overall Utilization

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Forecast Models for Overall Utilization Rates | | | | | | |
| Model Type | Data Fit | | | | Forecast | |
| RMSE | MAPE | AIC | BIC | RMSE | MAPE |
| SARIMA(3,1,1)x(0,0,1)12 + log | 0.003874 | 9.45 | -220.7 | -205.17 | 0.00302 | 5.85 |
| SARIMA(3,1,1)x(0,1,1)12 + log | 0.003653 | 8.76 | -211 | -196.4 | 0.004845 | 8.71 |
| SARIMA(2,1,0)x(0,1,1)12 | 0.003522 | 8.5 | -1237.8 | -1228.7 | 0.004969 | 8.4 |

Table 4: Prospective models for Overall Utilization rates

**RECOMMENDATIONS**

This project only scratched the surface of the sorts of insights that an analysis of forensic request and case data could produce. Not covered in the scope of this project, but possible avenues of exploration include; an analysis of utilization rates by location, and an analysis of requests (rather than utilization rates) per offence category.

Before any further forecast models are created, it is recommended that the issue of cases reoccurring in successive months is resolved. Removing this problem is likely to result in simpler forecast models and tighter 95% confidence intervals. In terms of counting cases, it would also be useful to make a determination as to how to deal with cases that have been closed, or gone cold, and reopened later as new evidence emerges. Should a reopened case be counted as a new case?

It is recommended that forecast models are tested for accuracy by collecting the quality scores from a number of validation sets of different sizes and averaging the results. Also, the PROC ARIMA procedure can be set to maximum likelihood method (the default is conditional likelihood) which should allow some comparison between the AIC and BIC scores produced by PROC ESM models and PROC ARIMA models.

**REFERENCES**

Chen, C. & Lee, S. (2017). Bayesian causality test for integer-valued time series models with applications to climate and crime data. *Applied Statistics*, 66(4), 797-814. doi/pdf/10.1111/rssc.12200

Borowik, G., Wawrzyniak,Z.M., & Cichosz, P. (2018). Time series analysis for crime forecasting. *26th International Conference on Systems Engineering*. Sydney, Australia: IEEE. doi:[10.1109/ICSENG.2018.8638179](https://doi-org.ezproxy.lib.rmit.edu.au/10.1109/ICSENG.2018.8638179)

Tularam, G.A. & Saeed, T. (2016). The use of exponential smoothing (ES), Holts and Winter (HW) and ARIMA models in oil price analysis. International Journal of Mathematics, Game Theory and Algebra, 25(1), 13-21. Retrieved from <https://search-proquest-com.ezproxy.lib.rmit.edu.au/docview/1864051412?accountid=13552>

Awang, N., Yong, N.K. & Hoeng, S.Y. (2017). Forecasting ozone concentration levels using Box-Jenkins ARIMA modelling and artificial neural networks: a comparative study. *Matematika,* 33(2), 119-130. doi.org/10.11113/matematika.v33.n2.900.

Athanasopoulos, G. & Weatherburn, D, NSW Bureau of Crime Statistics and Research. (2018). Forecasting male and female inmate numbers: A comparison of ARIMA and ETS modelling results. Crime and Justice Bulletin, 219. Retrieved from https://search-informit-com-au.ezproxy.lib.rmit.edu.au/fullText;dn=251904843256192;res=IELHSS

Donnelly, N., & Wan, W. Y., NSW Bureau of Crime Statistics and Research. (2016). Adult prison population size in New South Wales: comparative forecasts*. Crime and Justice Bulletin*, 199. Retrieved from https://search-informit-com-au.ezproxy.lib.rmit.edu.au/documentSummary;dn=042680895187999;res=IELHSS

Gorr, W., & Harries, R. (2003). Introduction to crime forecasting. *International Journal of Forecasting*, 19, 551-555. doi:10.1016/S0169-2070(03)00089-X

Mohan, S. & Vedula, S. (1995). *Multiplicative Seasonal ARIMA Model for Longterm Forecasting of Inflows*. Water Resource Management, 9, 115-125. https://doi-org.ezproxy.lib.rmit.edu.au/10.1007/BF00872463

Agnieszka, D. & Magdalena, L. (2018). Detection of outliers in the financial time series using ARIMA models. *Applications of Electromagnetics in Modern Techniques and Medicine (PTZE)*. Poland: IEEE. doi: [10.1109/PTZE.2018.8503260](https://doi-org.ezproxy.lib.rmit.edu.au/10.1109/PTZE.2018.8503260)

Akpan, E.A., Lasisi, K.E. & Adamu, A. (2019). Evaluation of Forecasts Performance of ARIMA-GARCH-type Models in the Light of Outliers. *World Science News*, 119, 68-84. Retrieved from <http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.psjd-7420d652-c1cf-48c3-8bfe-20bb4b87e1a0>

Arumugam, P. & Saranya, R. (2018).Outlier Detection and Missing Value in Seasonal ARIMA Model Using Rainfall Data. *Materials Today: Proceedings*, 5(1), 1791-1799. [doi.org/10.1016/j.matpr.2017.11.277](https://doi-org.ezproxy.lib.rmit.edu.au/10.1016/j.matpr.2017.11.277)

Lütkepohl, H. & Xu, F. (2012). The role of the log transformation in forecasting economic variables. *Empirical Economics*, 42(3), 619-638. doi-org.ezproxy.lib.rmit.edu.au/10.1007/s00181-010-0440-1

Chen, P., Yuan, H., & Shu, X. (2008) Forecasting Crime Using the ARIMA Model*. Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. (pp. 627-630). Shandong, China: IEEE. doi: 10.1109/FSKD.2008.222

**APPENDIX**

**1.A. Overall Utilization Rates:**

| **Autoregressive Factors** | |
| --- | --- |
| Factor 1: | 1 + 0.65976 B\*\*(1) + 0.38147 B\*\*(2) |
| **Moving Average Factors** | |
| Factor 1: | 1 - 0.70716 B\*\*(12) |

Figure 23: Parameter estimates for overall utilization rates forecast model

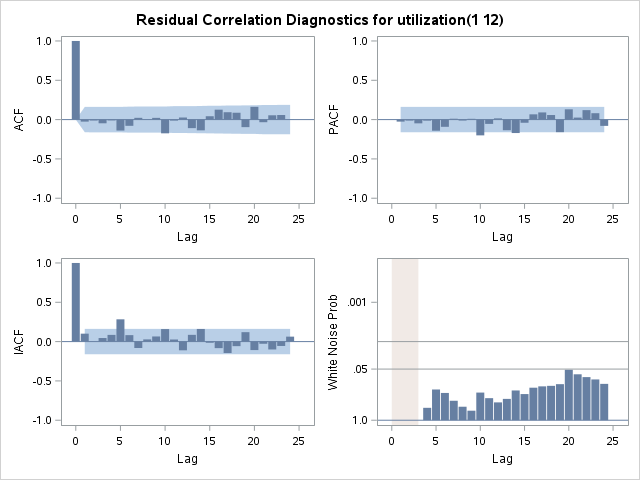
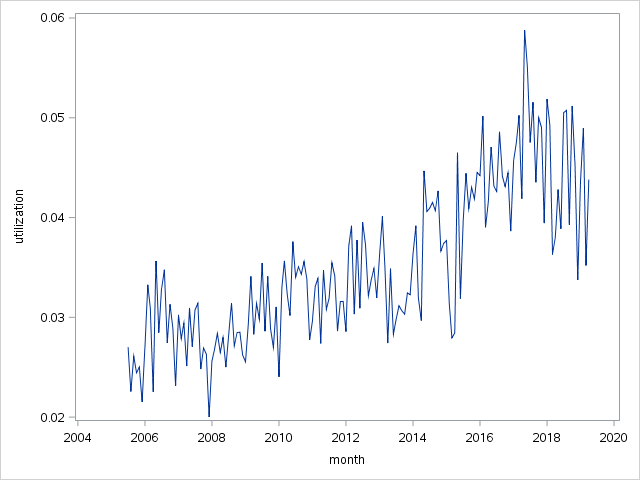


Figure 24: Overall Utilization Rates

Figure 25: Residual lags for overall utilization rates forecast model

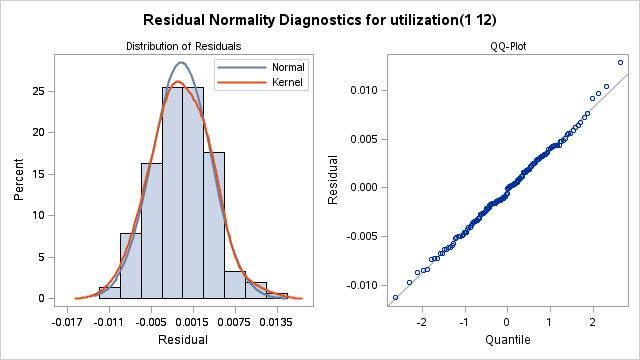


Figure 26: Kernel density plot for overall utilization rates forecast model

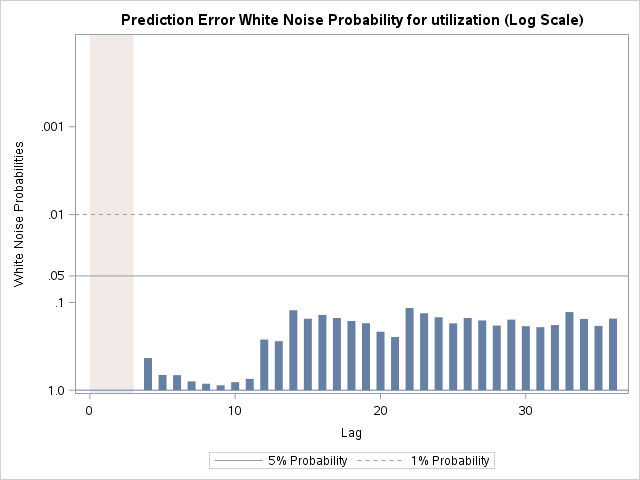
**1.B. Burglary Aggravated**:

Figure 27: Burglary (aggravated) utilization rates forecast model residual lags

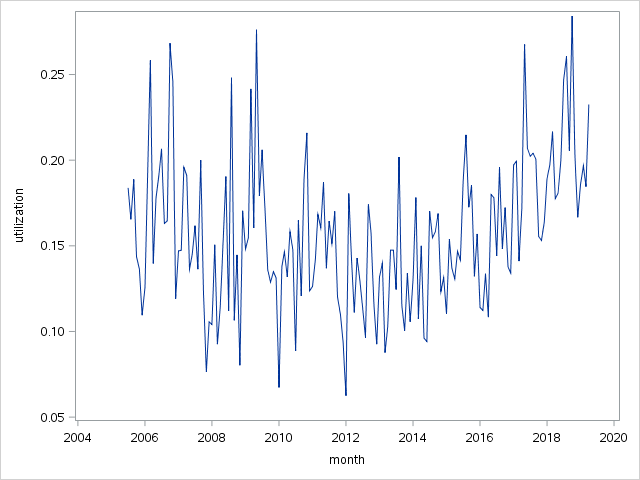


Figure 28: Burglary (aggravated) utilization rates

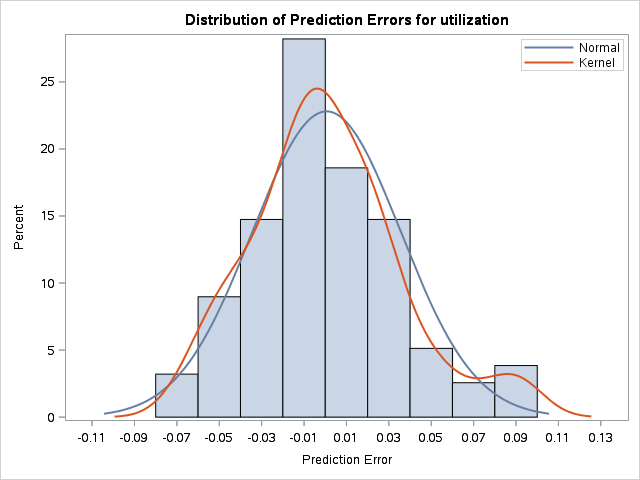


Figure 29: Burglary (aggravated) utilization rates forecast model kernel density plot

**1.C. Drugs (cult/ manuf/ traff):**

| **Autoregressive Factors** | |
| --- | --- |
| Factor 1: | 1 + 0.17895 B\*\*(1) |
| **Moving Average Factors** | |
| Factor 1: | 1 - 0.64035 B\*\*(1) |

Figure 30: Parameter estimates for drugs(cult/manuf/traf) forecast model

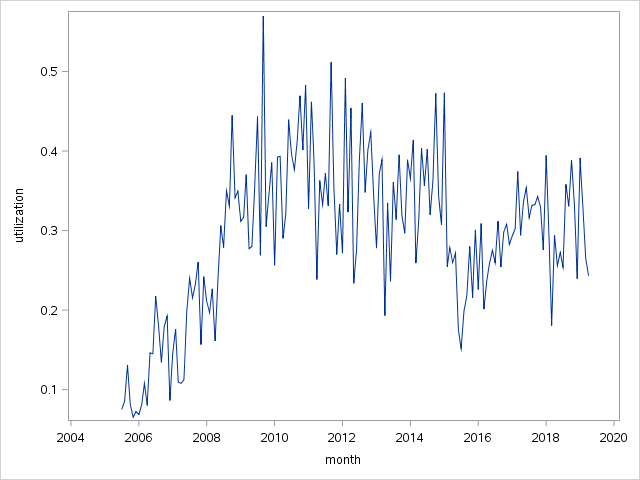


Figure 31: Drugs (cult/manuf/traf) utilization rates

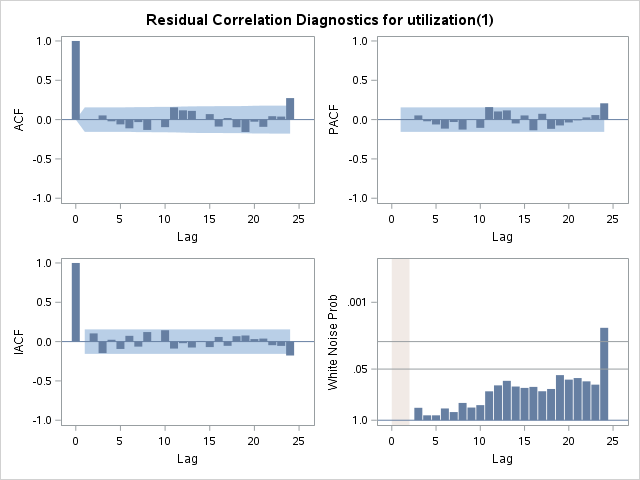
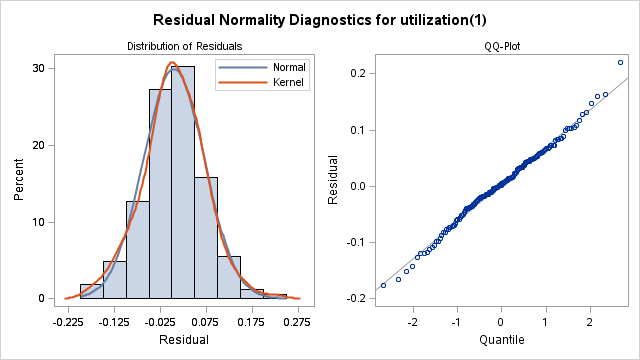


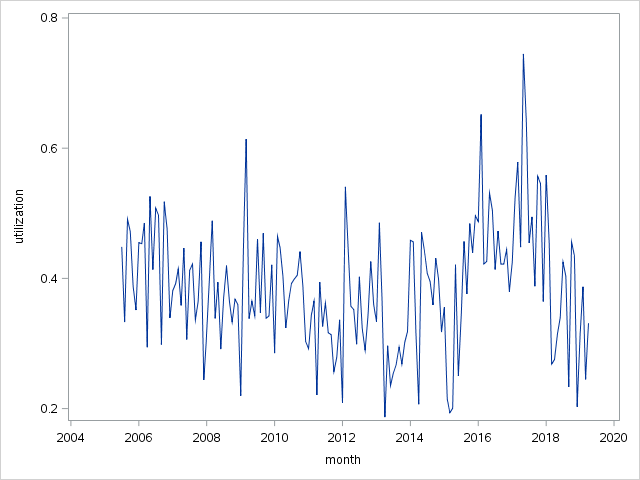
Figure 32: Kernel density plot for drugs(cult/manuf/traf) forecast model

Figure 33: Residual lags for drugs(cult/manuf/traf) forecast model

**1.D. Drugs (possess/use):**

| **Autoregressive Factors** | |
| --- | --- |
| Factor 1: | 1 - 0.84658 B\*\*(1) |
| **Moving Average Factors** | |
| Factor 1: | 1 - 0.52553 B\*\*(1) |
| Factor 2: | 1 - 0.68685 B\*\*(12) |

Figure 34: Parameter estimates for drugs (possess/use) forecast model



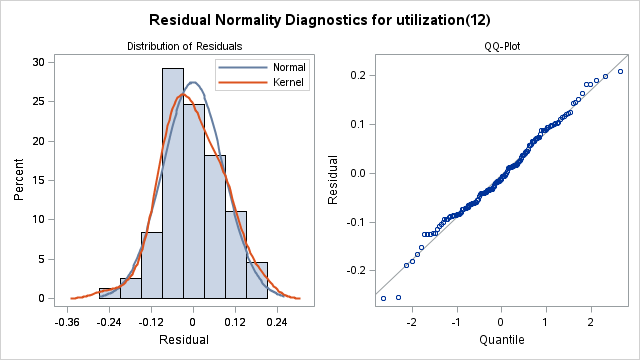
Figure 36: Drugs (possess/use) utilization rates

Figure 35: Kernel density plot for drugs (possess/ use) forecast model

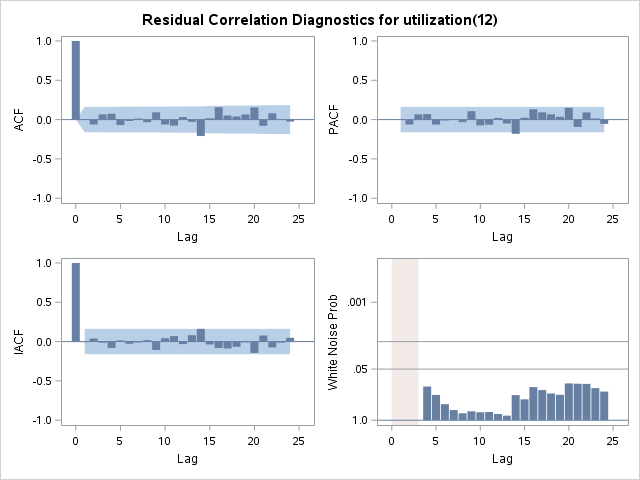


Figure 37: Residual lags for drugs (possess/use) forecast model

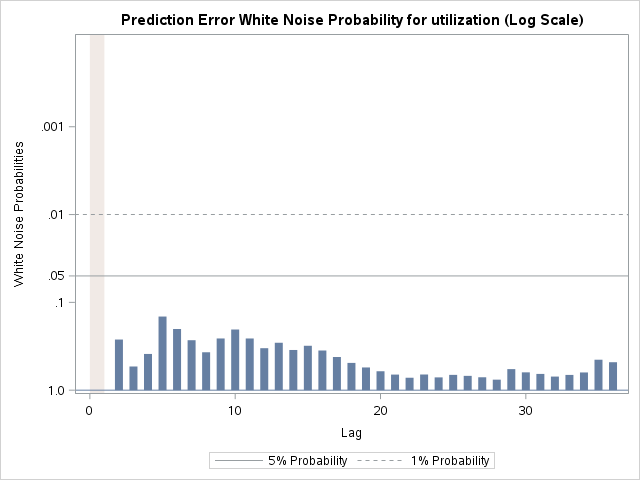
**1.E. Homicde:**

Figure 38: Residual lags for homicide forecast model

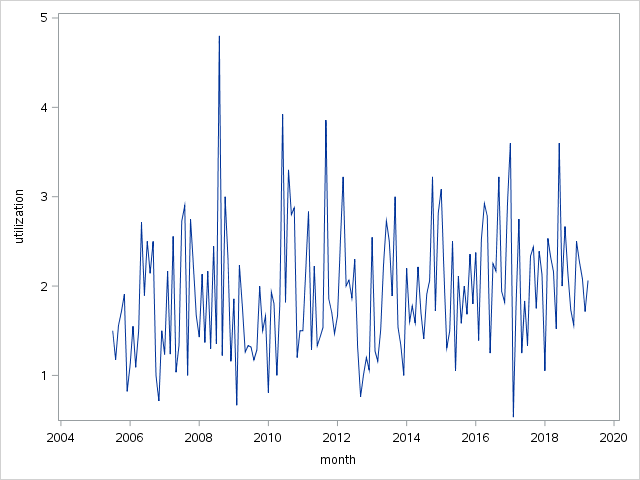


Figure 39: Homicide Utilization Rates

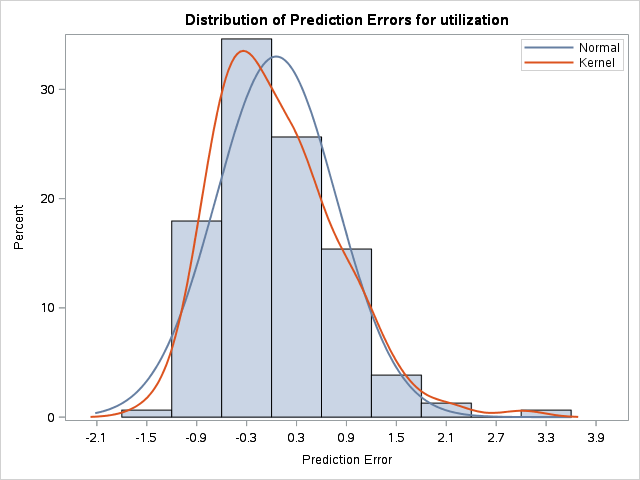


Figure 40: Kernel density plot for homicide forecast model

**1.F. Theft from Cars:**

| **Moving Average Factors** | |
| --- | --- |
| Factor 1: | 1 - 0.67174 B\*\*(1) |

Figure 41: Parameter estimates for theft from cars forecast model

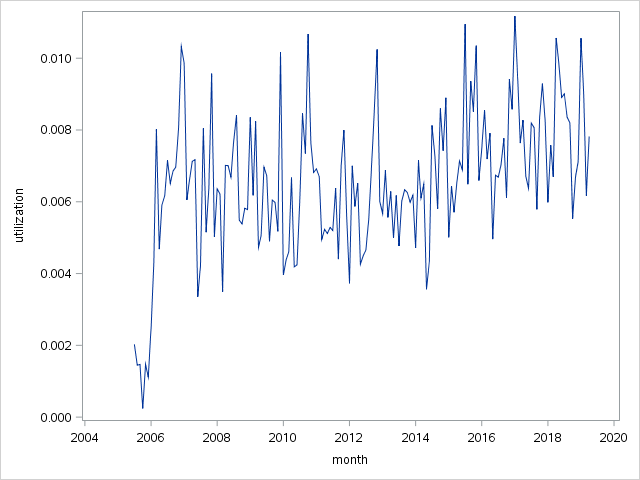


Figure 42: Theft from Cars Utilization Rates

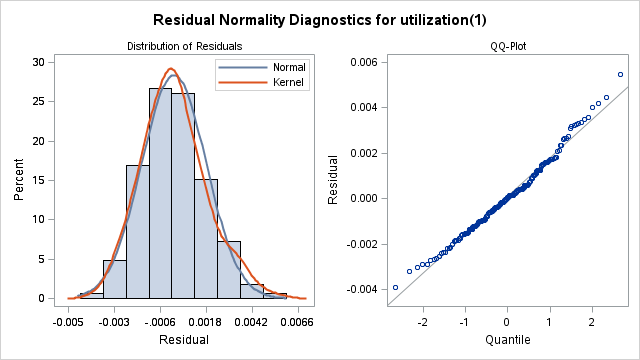
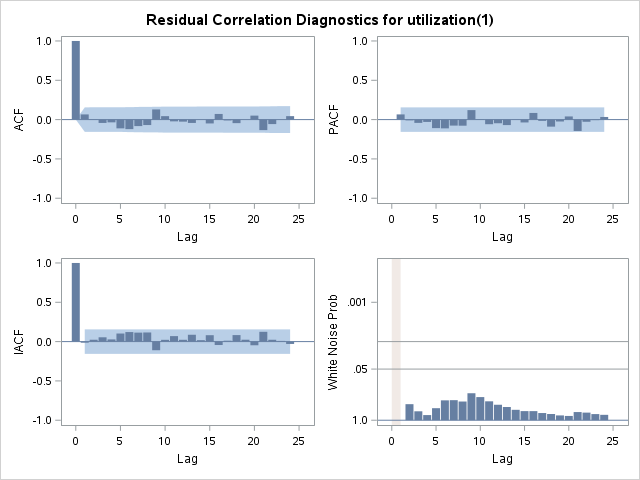


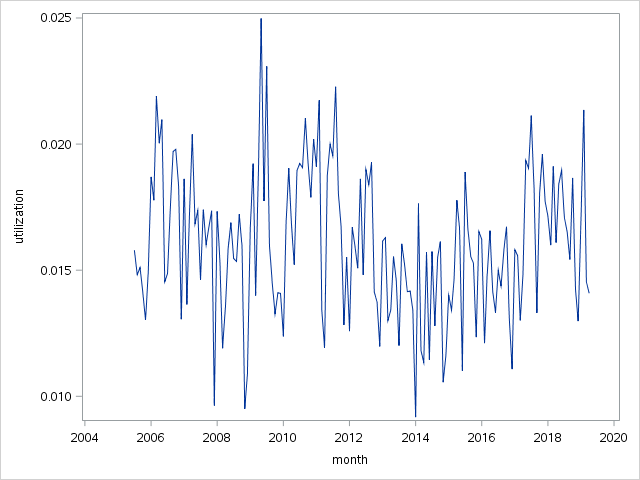
Figure 43: Residual lags for theft from cars forecast model

Figure 44:Kernel density plot for theft from cars forecast model

**1.G. Assault:**

| **Model for variable utilization** | |
| --- | --- |
| Estimated Mean | 0.016014 |
| **Autoregressive Factors** | |
| Factor 1: | 1 - 0.68476 B\*\*(1) |
| **Moving Average Factors** | |
| Factor 1: | 1 - 0.40876 B\*\*(1) |

Figure 45: Parameter estimates for assault forecast model



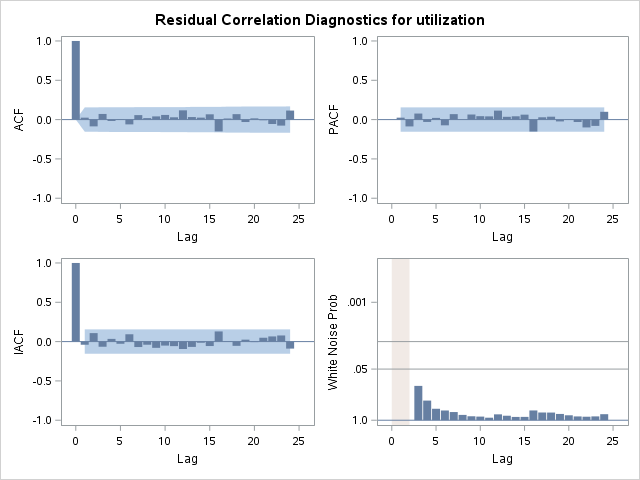
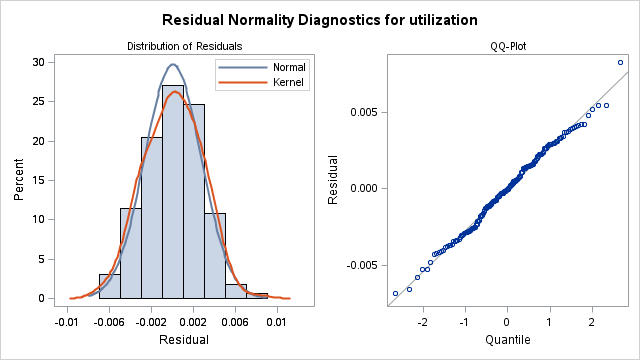
Figure 48: Assault Utilization Rates

Figure 46: Residual lags for assault forecast model

Figure 47: Kernel denisty plot for assault forecast model residuals

**1.H. Rape:**

| **Moving Average Factors** | |
| --- | --- |
| Factor 1: | 1 - 0.82257 B\*\*(1) |

Figure 49: Parameter estimates for rape forecast model

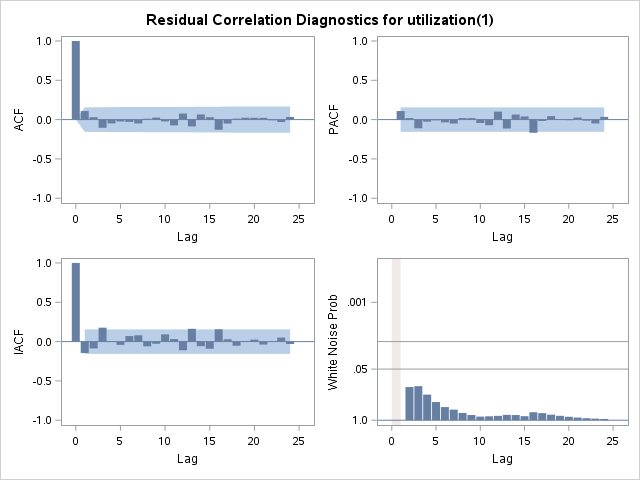
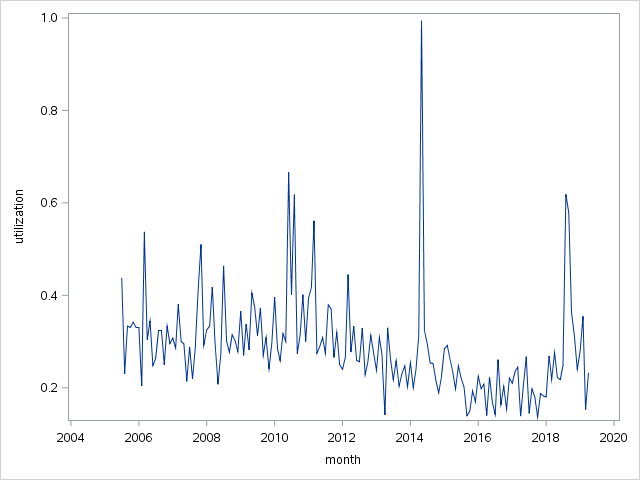


Figure 50: Rape Utilization Rates

Figure 51: Residual lags for rape forecast model

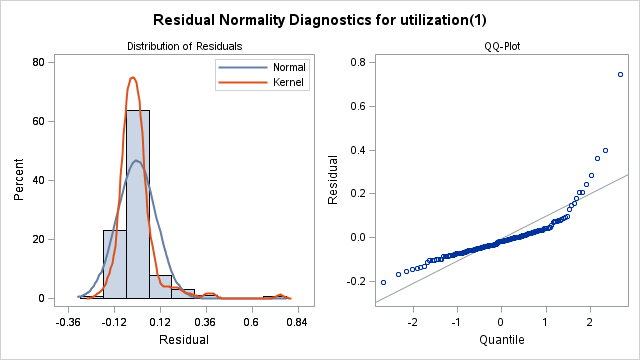
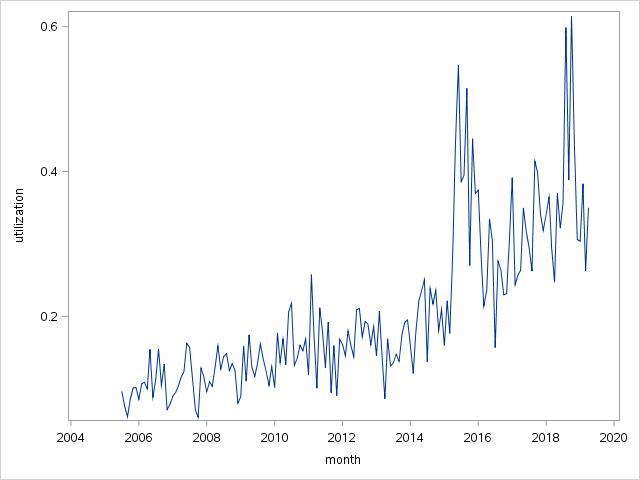


Figure 52: Kernel density plot for residuals of rape forecast model

**1.I. Arson:**

| **Moving Average Factors** | |
| --- | --- |
| Factor 1: | 1 - 0.69966 B\*\*(1) |

Figure 53: Parameter estimates for arson forecast model

Figure 54: Arson Utilization Rates

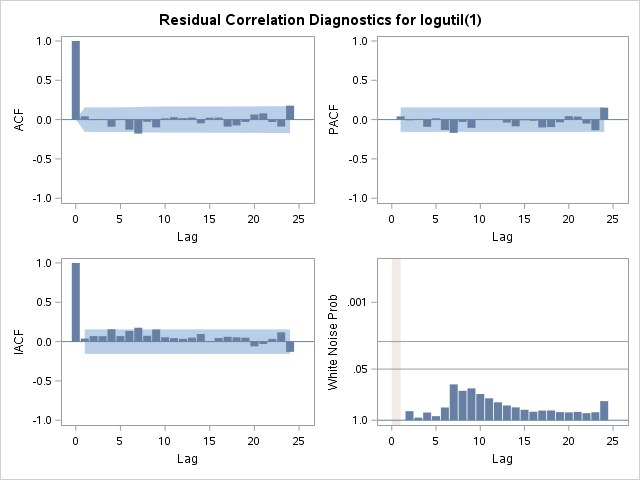
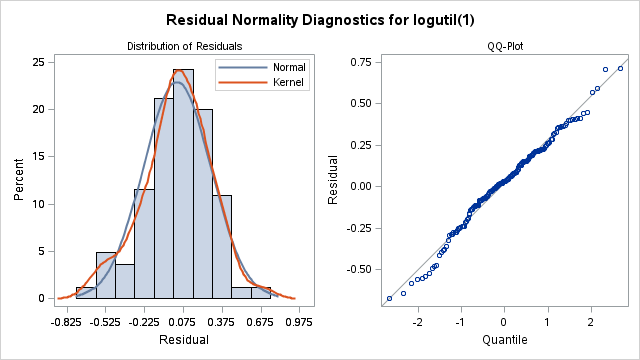


Figure 55: Kernel density for arson forecast model

Figure 56: Residual lags for arson forecast model

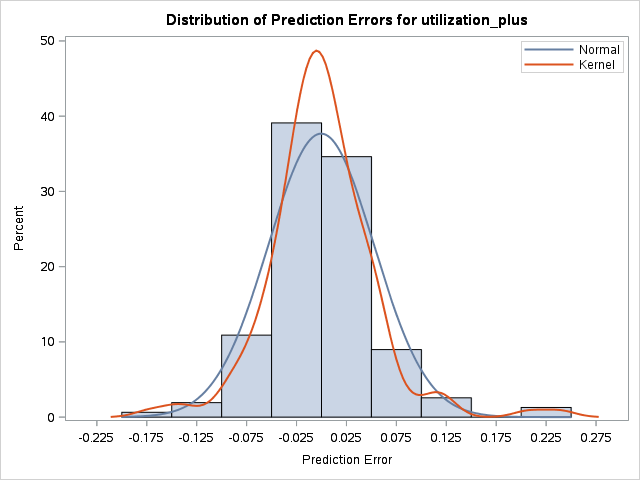
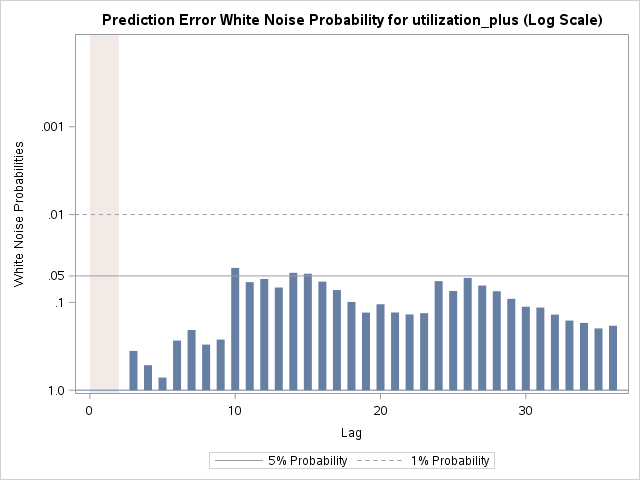
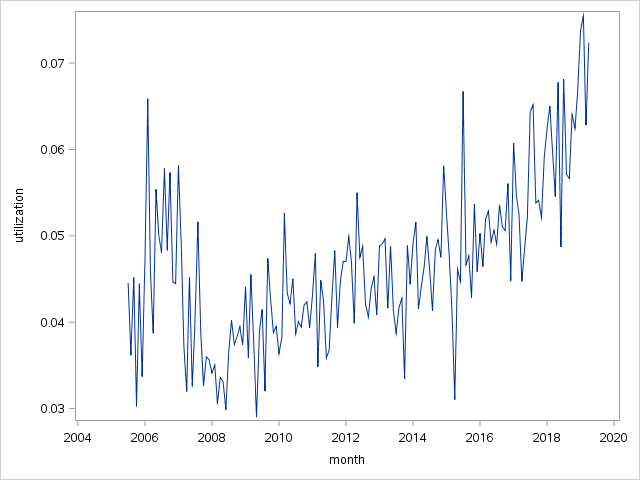
**1.J.Rejected ARSON model (esm: linear):**

Figure 57: Residual lags for rejected arson forecast model

Figure 58: Kernel density plot for rejected arson forecast model

**1.K. Burglary (residential):**

| **Moving Average Factors** | |
| --- | --- |
| Factor 1: | 1 - 0.7739 B\*\*(1) |

Figure 59: Parameter estimates for burglary (residential) forecast model

Figure 60: Burglary Utilization Rates

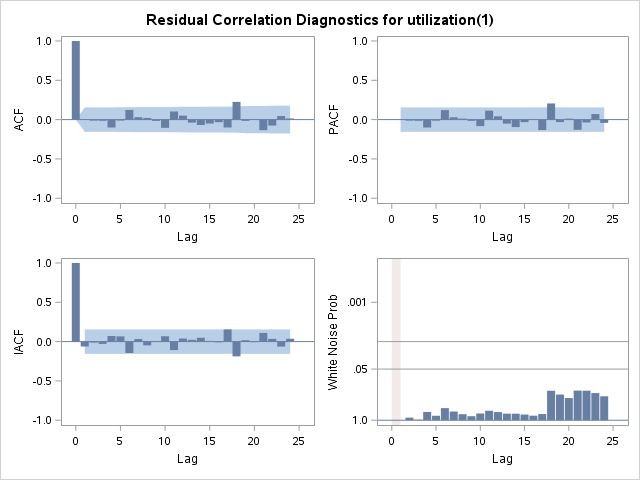
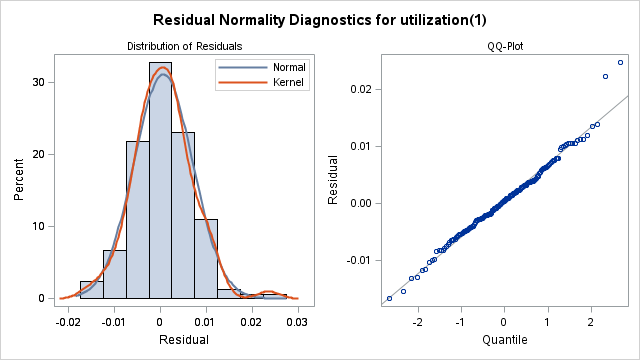


Figure 61: Kernel density plot for burglary (residential) forecast model residuals

Figure 62: Residual lags for burglary (residential) forecast model

**2. A. Validation results for Burglary (residential):**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Forecast Models for Offence Category: Residential Burglary | | | | | | | |
| Model | Validation Set Size | Data Fit | | | | Forecast | |
| RMSE | MAPE | AIC | BIC | RMSE | MAPE |
| ARIMA(0,1,1) | 10 | 0.00636 | 11.28 | -1195.65 | -1192.55 | 0.00458 | 5.92 |
| 15 | 0.006458 | 11.4953 | -1195.59 | -1192.48 | 0.004141 | 5.94 |
| 20 | 0.006217 | 11.42 | -1195.59 | -1192.48 | 0.006462 | 7.814 |
| 25 | 0.006103 | 11.18 | -1195.59 | -1192.48 | 0.006902 | 9.6439 |
| 5 | 0.006254 | 10.9645 | -1195.59 | -1192.48 | 0.006209 | 9.28 |
| Average | | | | | | 0.005659 | 7.71958 |
| ESM: linear | 10 | 0.00632 | 10.93 | -1575.82 | -1569.72 | 0.00965 | 11.51 |
| 15 | 0.006269 | 10.907 | -1527.77 | -1521.73 | 0.009032 | 11.4389 |
| 20 | 0.006332 | 11.0566 | -1474.11 | -1468.15 | 0.008323 | 10.357 |
| 25 | 0.006234 | 10.97 | -1427.91 | -1422.01 | 0.010049 | 12.75 |
| 5 | 0.006296 | 10.83999 | -1627.87 | -1621.71 | 0.01044 | 12.85 |
| Average | | | | | | 0.009499 | 11.78118 |