IEE 579 CASE STUDY 2: Time Series Modelling of Temperature and Electricity DATa

Ashlin Sreedhar

**IEE 579 Case Study 2**

**Ashlin Sreedhar**

1)

To model the time series data, two different software were used in conjunction with each other. they are MINITAB and JMP.

**Time Series of the Input Variable: Temperature**

The first step to developing a time series model is plotting the input variable time series data to see if there is any pattern in the data (trend or seasonal behavior) and to check if there are any outliers that are present in the time series data. The time series plot of the input variable (temperature) is given below.



Based on the time series plot of the input variable we can clearly say that the data is seasonal because of the cyclic pattern that the data exhibits. A trend in the data seems to be absent. The constant variance assumption of the data also seems to hold for this time series. Considering that this is temperature data that is recorded every month, seasonal patterns are expected in the data. The next step is to perform stationarity checks on the time series to confirm suspicions that the data is seasonal.

**Stationarity Checks**

To check for time series stationarity, plots such as the ACF, PACF are obtained to evaluate the time series data. If no linearly decreasing or slowly decreasing autocorrelation pattern is observed, the presence of trend in the data can be safely ruled out. Since cyclic patterns were observed in the time series plot, cyclic patterns are expected in

the ACF plot. Significant correlation is expected at higher lags especially around lag 12 and multiples of 12 since the period for monthly data is 12.

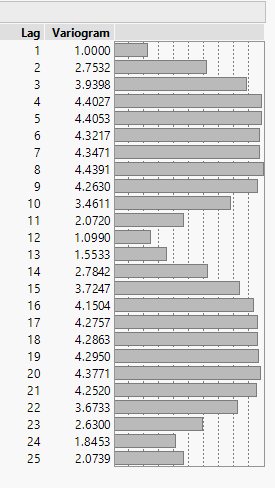
**ACF and PACF Plots**





Looking at the ACF plot we see that there are significant spikes at lag 12 and lag 24 as expected for data with a **seasonal trends**. We next look at the variogram for the input variable to verify the presence of a cyclic like pattern.

**Variogram**



Another plot that can be used to establish stationarity is the variogram plot of the data. If the time series is stationary the variogram should converge to a stable value and then fluctuate around it. This does not happen here. As expected we see a **cyclic like pattern** in the variogram. This is indicative of the seasonal trend in the data.

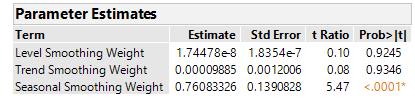
**Time Series Conclusions**

Based on the analysis of the ACF, PACF and the Variogram plots we can safely conclude that the data has a seasonal trend. Based on this conclusion, we need to use suitable models that account for this seasonal trend. Models normally used for seasonal time series data are **Holt-Winters Models** and **Seasonal ARIMA Models**.

**Task 1: Holt-Winters Exponential Smoothing Model**

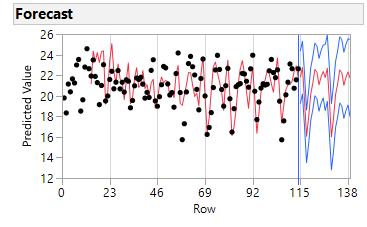
The Holt-Winters method was used to build an exponential smoothing model for the temperature time series. Looking at the time series plot we see that an additive model would be suitable for the times series data since there seems to be no increase in the amplitude of the seasonal trend in the data as the temperature is recorded. For this part of the task we use JMP’s Winters Smoothing option. The optimal smoothing parameters are estimated by JMP’s optimization algorithms. The results for this are displayed below.

**Parameter Estimation**



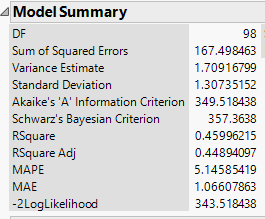
For the parameter estimation table, we can clearly see that the seasonal smoothing parameter is significant. The other parameters seem to be insignificant, this is expected since from the time series plot we neither see a change in level nor a trend-based structure in the data.

**Model Fit**



As seen in the graph above we see that the smoothing model is responsive to the time series data and closely follows the data.

**Performance Metrics**



The Performance metrics of the model especially the Variance, MAPE and MAE look good for the smoothing model. We need to compare these metrics with the seasonal ARIMA to see which model performs better on seasonal data.

**Task 2: Seasonal ARIMA Model**

The seasonal ARIMA model seems quite apt for the temperature time series since there is definite seasonal trend in the data. The first step in the model building process of the seasonal ARIMA model is to take a seasonal difference to observe the underlying patterns that are present in the data. We take a seasonal difference of order 1 (1-B12) and observe the underlying patterns in the data from the ACF and PACF plots of the differenced data. These plots are given below.





Looking at the ACF and PACF of the Seasonal Differenced Data we can say that there appears to be two candidate models that can be used to fit the underlying patterns in the differenced data. The Candidate Models are given below.

|  |  |
| --- | --- |
| **Model Number** | **Candidate Model for Seasonally Differenced Data** |
| **1** | ARIMA(0,0,1) |
| **2** | ARIMA(2,0,0) |

Now that the candidate models have been identified, we proceed with Model Fitting and Parameter Estimation.

**Model Fitting and Parameter Estimation**

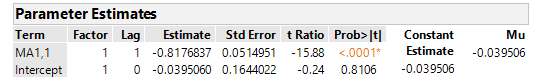
Candidate Model 1: ARIMA (0,0,1)x(0,1,0)12

In this candidate model we have one Normal MA parameter and one order of seasonal difference of lag 12. The equation of the Seasonal ARIMA model used is given below.

**yt = δ +**

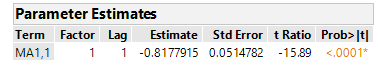
**Parameter Estimation**

We now build a seasonal ARIMA model using the temperate data and estimate the parameters for model using JMP. In the output below the parameter estimates have been specified.



We obtain the parameter estimates for this candidate model built. Based on the above estimates we can see that the Intercept term or the constant term is not significant. And hence we estimate the parameters for the seasonal model once again this time without the intercept term. We re-specify the above Seasonal ARIMA model without the intercept term.

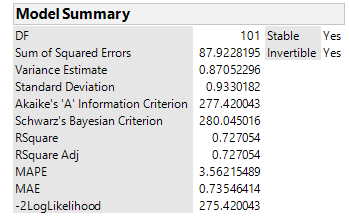
**yt =**



**As seen above the parameter estimates for the model are significant.**

**Performance Metrics**

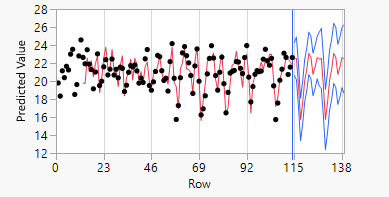
Some of the performance characteristics obtained during the model fitting process are given below.



Based on the performance characteristics we can say that the model performs well in this regard. Before we can conclude about the validity of the model we need to perform adequacy checks on the model using Residual Analysis.

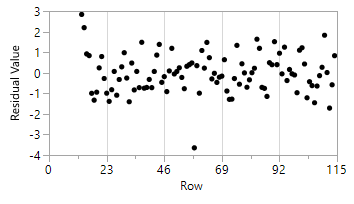
**Model Adequacy**

Actual Vs Fit Plot



Looking at the plot above, we can say that the fit of the model to the actual data looks reasonable.

Residual Plots

****

Looking at the residual plot we can see that the residuals are random. No apparent pattern in the data seems to be present.

ACF and PACF Plots

| **Lag** | **AutoCorr** |  |
| --- | --- | --- |
| 0 | 1.0000 |  |
| 1 | 0.1984 |  |
| 2 | 0.0084 |  |
| 3 | -0.2617 |  |
| 4 | -0.1075 |  |
| 5 | -0.1328 |  |
| 6 | 0.0552 |  |
| 7 | 0.1747 |  |
| 8 | 0.0817 |  |
| 9 | -0.0241 |  |
| 10 | -0.1690 |  |
| 11 | -0.1410 |  |
| 12 | -0.1040 |  |
| 13 | 0.0954 |  |
| 14 | 0.0516 |  |
| 15 | -0.0222 |  |
| 16 | -0.0401 |  |
| 17 | -0.0611 |  |
| 18 | -0.0091 |  |
| 19 | 0.0323 |  |
| 20 | 0.0882 |  |
| 21 | -0.0725 |  |
| 22 | -0.0059 |  |
| 23 | -0.1470 |  |
| 24 | -0.0494 |  |
| 25 | -0.1828 |  |

| **Lag** | **Partial** |  |
| --- | --- | --- |
| 0 | 1.0000 |  |
| 1 | 0.1984 |  |
| 2 | -0.0323 |  |
| 3 | -0.2678 |  |
| 4 | -0.0036 |  |
| 5 | -0.1155 |  |
| 6 | 0.0384 |  |
| 7 | 0.1523 |  |
| 8 | -0.0498 |  |
| 9 | -0.0273 |  |
| 10 | -0.1075 |  |
| 11 | -0.0756 |  |
| 12 | -0.0420 |  |
| 13 | 0.0658 |  |
| 14 | -0.0626 |  |
| 15 | -0.1148 |  |
| 16 | 0.0119 |  |
| 17 | -0.0402 |  |
| 18 | 0.0316 |  |
| 19 | 0.0425 |  |
| 20 | -0.0074 |  |
| 21 | -0.1414 |  |
| 22 | 0.0380 |  |
| 23 | -0.1347 |  |
| 24 | -0.0360 |  |
| 25 | -0.1922 |  |

Looking at the ACF and PACF plots we see that there is still some autocorrelation present in the data. Due to the presence of autocorrelation in the residuals we can conclude and say that this cannot be used as a valid model for the temperature time series. We move on to the next candidate model to evaluate its performance and check its adequacy.

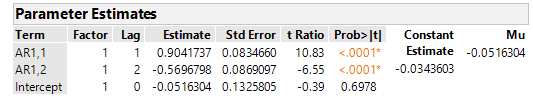
Candidate Model 2: ARIMA (2,0,0)x(0,1,0)12

In this candidate model we have 2 AR parameters and one order of seasonal difference of lag 12. The equation of the Seasonal ARIMA model used is given below.

**yt = δ +**

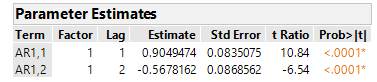
**Parameter Estimates**

We now build a seasonal ARIMA model using the temperate data and estimate the parameters for model using JMP. In the output below the parameter estimates have been specified.

****

We obtain the parameter estimates for this candidate model built. Based on the above estimates we can see that the Intercept term or the constant term is not significant. And hence we estimate the parameters for the seasonal model once again, this time without the intercept term. We re-specify the above Seasonal ARIMA model without the intercept term.

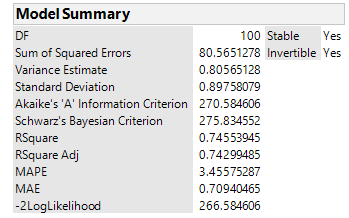
**yt =**



**As seen above the parameter estimates for the model are significant.**

**Performance Metrics**

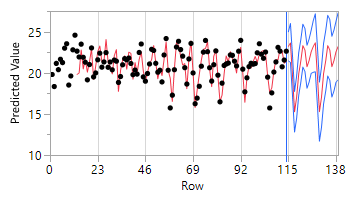
Some of the performance characteristics obtained during the model fitting process are given below.



Based on the performance characteristics we can say that the model performs well in this regard. Before we can conclude about the validity of the model we need to perform adequacy checks on the model using Residual Analysis.

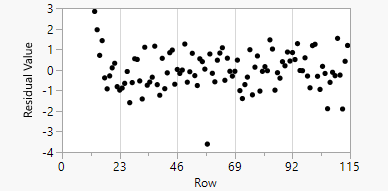
**Model Adequacy**

Actual vs Fit Plot



Looking at the plot above, we can say that the fit of the model to the actual data looks reasonable.

Residual Plot



Looking at the residual plot we can see that the residuals are random. No apparent pattern in the data seems to be present.

ACF and PACF Plots

| **Lag** | **AutoCorr** |  |
| --- | --- | --- |
| 0 | 1.0000 |  |
| 1 | 0.0939 |  |
| 2 | -0.1249 |  |
| 3 | 0.1374 |  |
| 4 | 0.0706 |  |
| 5 | -0.0819 |  |
| 6 | -0.0569 |  |
| 7 | 0.1025 |  |
| 8 | -0.0297 |  |
| 9 | -0.0714 |  |
| 10 | -0.0681 |  |
| 11 | -0.0793 |  |
| 12 | -0.1428 |  |
| 13 | 0.0559 |  |
| 14 | 0.0052 |  |
| 15 | -0.0696 |  |
| 16 | 0.0049 |  |
| 17 | -0.0286 |  |
| 18 | -0.0487 |  |
| 19 | 0.0411 |  |
| 20 | 0.0173 |  |
| 21 | -0.1070 |  |
| 22 | -0.0333 |  |
| 23 | -0.0226 |  |
| 24 | -0.0207 |  |
| 25 | -0.2605 |  |

| **Lag** | **Partial** |  |
| --- | --- | --- |
| 0 | 1.0000 |  |
| 1 | 0.0939 |  |
| 2 | -0.1350 |  |
| 3 | 0.1679 |  |
| 4 | 0.0190 |  |
| 5 | -0.0543 |  |
| 6 | -0.0528 |  |
| 7 | 0.0883 |  |
| 8 | -0.0524 |  |
| 9 | -0.0164 |  |
| 10 | -0.1009 |  |
| 11 | -0.0784 |  |
| 12 | -0.1320 |  |
| 13 | 0.1105 |  |
| 14 | -0.0450 |  |
| 15 | -0.0048 |  |
| 16 | -0.0202 |  |
| 17 | -0.0530 |  |
| 18 | -0.0338 |  |
| 19 | 0.0776 |  |
| 20 | -0.0465 |  |
| 21 | -0.1066 |  |
| 22 | -0.0519 |  |
| 23 | -0.0690 |  |
| 24 | 0.0054 |  |
| 25 | -0.2638 |  |

Looking at the ACF and PACF plots above, we can conclude that no autocorrelation is present in the residuals and hence use of this Seasonal ARIMA model to model the temperature time series is valid. Among the 2 candidate models identified we **choose Candidate Model 2** since its performance and adequacy is better than Candidate Model 1.

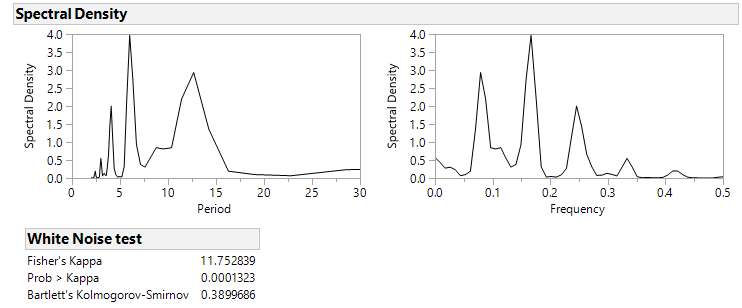
**Model Comparison: Holt-Winters Exponential Smoothing Model Vs Seasonal ARIMA (2,0,0)x(0,1,0)12**

We now compare the 2 models that were used to model the temperature time series to identify which of them performs better. Several performance metrics were used to compare the 2 models. The comparison of the two model is seen below.

|  |  |  |
| --- | --- | --- |
| **Performance Criteria** | **Seasonal ARIMA** | **Seasonal Exp. Smoothing** |
| **Variance** | 0.8056 | 1.7091 |
| **AIC** | 270.584 | 349.518 |
| **BIC** | 275.834 | 357.363 |
| **MAPE** | 3.455 | 5.145 |
| **MAE** | 0.7094 | 1.066 |
| **MSE** | 0.8056 | 1.709 |
| **R2 Adjusted** | 0.7429 | 0.4489 |
| **MAD** | 0.7180 | 1.0458 |

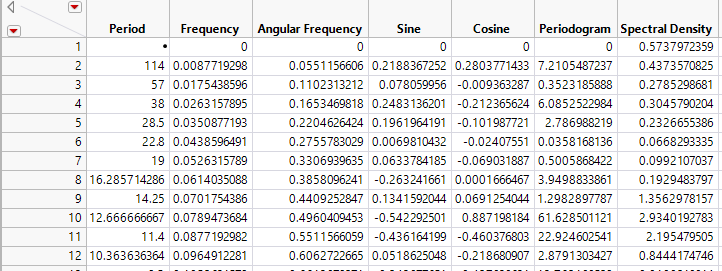
Based on the performance metrics above, we can clearly see that the **Seasonal ARIMA** model clearly outperforms the Exponential Smoothing Model is all the Performance Measures. Because of its superior performance, this model will be used for the pre-whitening phase Transfer Function Model Building process.

**Task 3: Spectral Analysis**



Looking at the spectral density plot, we can see significant spikes on the **Spectral density Vs the Period plot** around Period=12, this is expected since the data is seasonal with s=12. Looking at the plot of **Spectral Density Vs Frequency** we can see that there are significant spikes around the frequencies 0.1, 0.2 and 0.3. This is because these correspond to the periods 12, 24 and 36 respectively and hence show the periodic characteristic of the data with s=12.

The Fisher’s Kappa test has a Null Hypothesis which states that the series has a normal distribution with variance 1 **(White Noise).** The Alternative Hypothesis is that the series has a periodic component present and is not White Noise. If the **Prob>Kappa value** is less than 0.05 we can reject the Null Hypothesis that the series is White Noise and say that the series has a periodic component. In this case we can say that since the **Prob>Kappa Value is lower than 0.05** the series has a periodic component which is reflected in the Spectral Analysis Graphs.

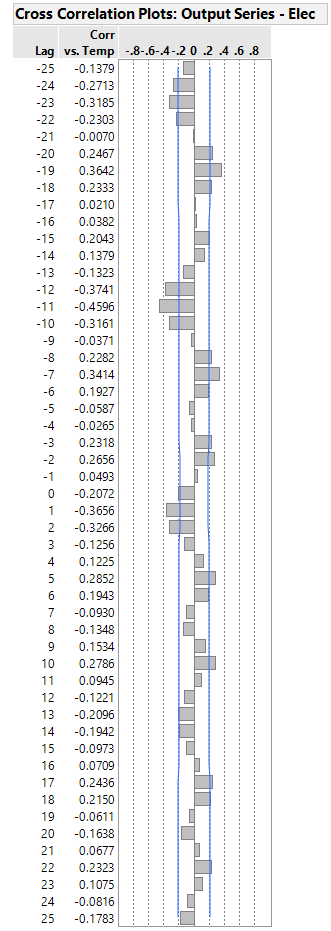


Looking at the spectral density values in the table above we can see that for the period≈12 there is a significant spike in the spectral density. The frequency= 0.070175≈0.1. This confirms our basic assumptions that the period is s=12.

**Task 4: Transfer Function Model**

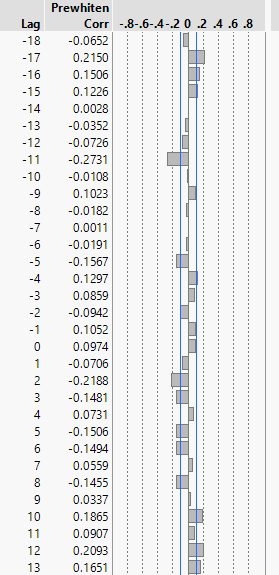
**Cross-Correlation Before Pre-Whitening**

Before building a transfer function model the first step is to check for cross correlation between the input and output time series. During Task 2 it was seen that there was autocorrelation and seasonal trends in the temperature input series. Due to the presence of these properties the cross-correlation between the input and output time series is expected to be contaminated. The plot of this contamination of the cross-correlation plot is given below.



As expected significant cross-correlation exists for the negative lags, which does not make sense since the relationship between the input series temperature and the output series is supposed to be causal like in nature. We also see significant cyclic cross-correlation patterns in the plot which indicate that pre-whitening is necessary to estimate the order of the transfer function.

**Cross-Correlation After Pre-Whitening**



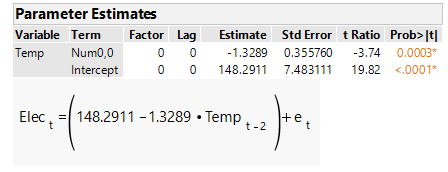
Based on the cross-correlation plot obtained after pre-whitening, we see that the cross-correlation present over the negative lags has significantly gone down. Using the above plot the order of the transfer function can be identified. Since this is only an estimate of what the order of the transfer function should be, it is important that we evaluate several different candidates before a decision is made. The Transfer Function models are compared based on their AIC and BIC scores and their adequacy checks to see if the residuals and the input series are independent. The candidate ARMAX models are listed below.

|  |  |
| --- | --- |
| **Model Number** | **Transfer Function Model ARMAX (b,r,s)** |
| **1** | ARMAX (2,0,0) |
| **2** | ARMAX (2,1,0) |
| **3** | ARMAX (2,0,1) |

**Transfer Function Model Building**

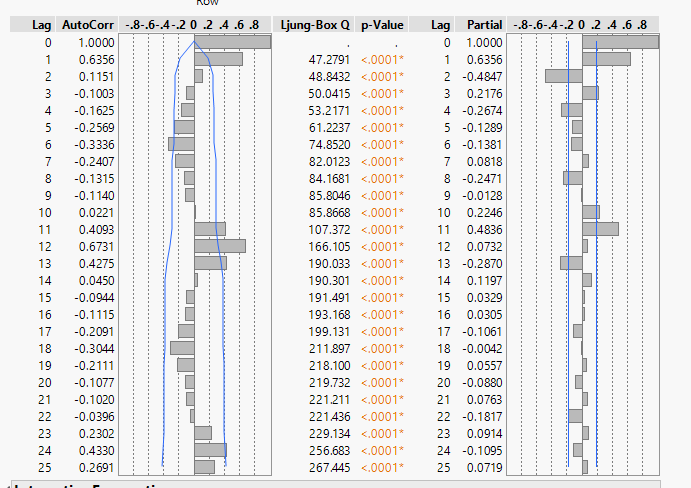
Preliminary Model Fitting: ARMAX (2,0,0)

We fit a preliminary model for the transfer function just to identify the order of the noise model to fit an overall **Transfer Function Noise Model.** After fitting the preliminary Transfer Function Model, we analyze the PACF and ACF plots of the residuals so as to estimate the parameters for the noise model. The parameter estimates for the preliminary transfer function model are given below.

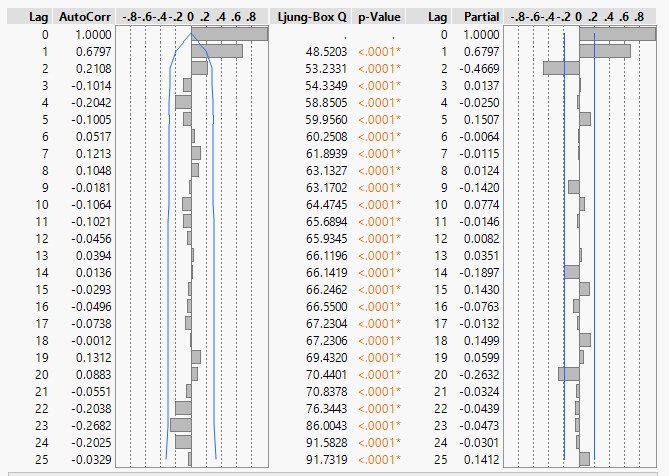


Based on the parameter estimates obtained we can say that the model specification seems to be correct since as the terms look to be significant. Next we look at the ACF and PACF plots of the residuals to find the parameters for the noise model.

ACF and PACF Plots



Based on the plots above, we estimate that the noise model would need a seasonal difference, because of the seasonal difference present in the data. We perform one Seasonal Difference for the Noise Model and re-examine the ACF and PACF plots are the seasonal patterns have been removed.



After **adding seasonal difference** to the noise model, the above ACF and PACF plots are obtained. Based on the plots, candidate models for Noise Model are obtained. They are mentioned below.

|  |  |
| --- | --- |
| **Model Number** | **Noise Model ARIMA (p,d,q))** |
| **1** | ARIMA (2,0,0)x(0,1,0) |
| **2** | ARIMA (0,0,1)x(0,1,0) |

**Transfer Function-Noise Model Fitting**

The candidate models identified for the transfer function and the noise model are combined and each of these overall models are evaluated. The list of the candidate overall models that are to be evaluated are given in the table below.

|  |  |
| --- | --- |
| **Model Number** | **Transfer Function - Noise Model** |
| **1** | ARMAX (2,0,0) + ARIMA (2,0,0)x(0,1,0) |
| **2** | ARMAX (2,0,0) + ARIMA (0,0,1)x(0,1,0) |
| **3** | ARMAX (2,0,1) + ARIMA (2,0,0)x(0,1,0) |
| **4** | ARMAX (2,0,1) + ARIMA (0,0,1)x(0,1,0) |
| **5** | ARMAX (2,1,0) + ARIMA (2,0,0)x(0,1,0) |
| **6** | ARMAX (2,1,0) + ARIMA (0,0,1)x(0,1,0) |

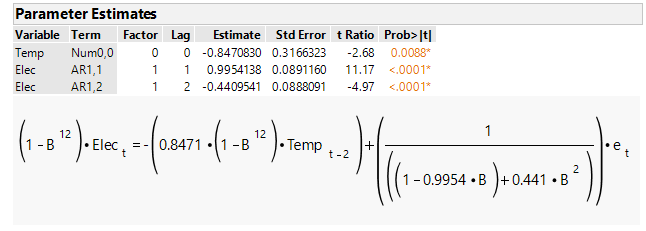
Candidate Model 1: ARMAX (2,0,0) + ARIMA (2,0,0)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, a Seasonal difference of order 1 and Autoregressive Term of order 2 for the Noise Model. The equation of the model is given below.

**yt = ω0Xt-2 +**

**Parameter Estimates**

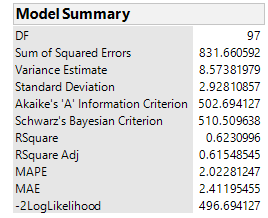
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for model. In the output below the parameter estimates have been specified.



Looking at the output above we see that all the parameter estimates are significant.

**Performance Metrics**

Some of the performance characteristics obtained during the model fitting process are given below.

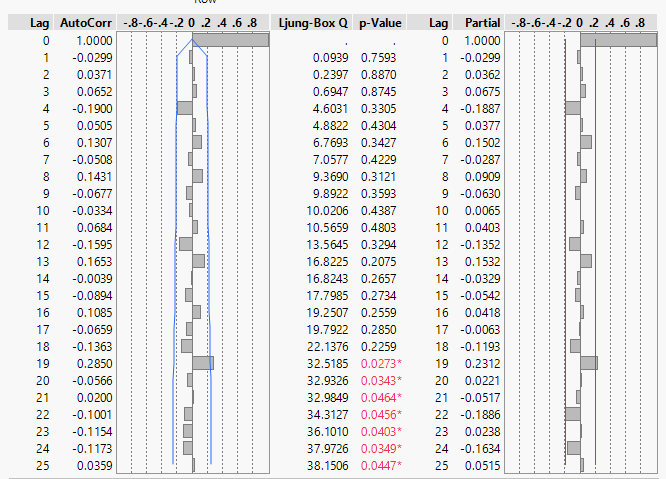
****

The performance metrics above look reasonable. We still need to compare this model with other candidate models to see if it is the best model among the candidate models. Next we look at the residuals to see if there is any autocorrelation remaining in the residuals.

**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



The residuals seem to be behaving like white noise, the model seems reasonable.

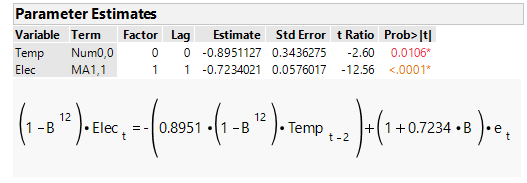
Candidate Model 2: ARMAX (2,0,0) + ARIMA (0,0,1)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, a Seasonal difference of order 1 and Moving Average Term of order 1 for the Noise Model. The equation of the model is given below.

**yt = ω0Xt-2 +**

**Parameter Estimates**

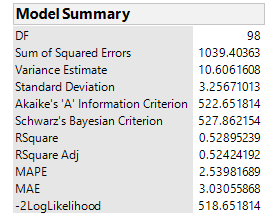
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for model. In the output below the parameter estimates have been specified.



Looking at the output above we see that all the parameter estimates are significant.

**Performance Metrics**

Some of the performance characteristics obtained during the model fitting process are given below.

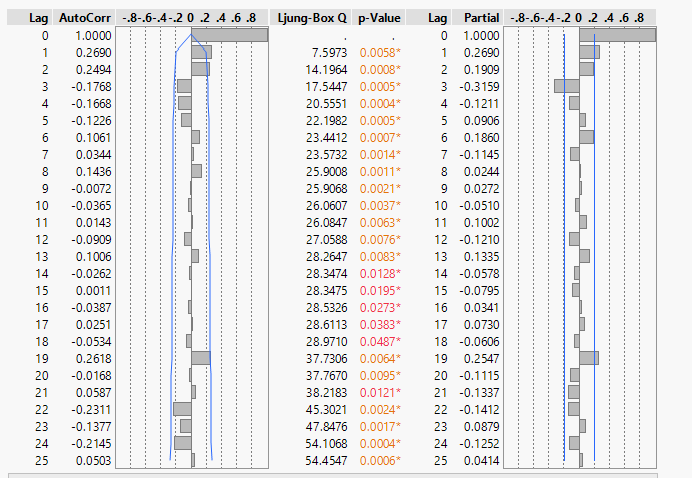


The performance metrics above look reasonable. We still need to compare this model with other candidate models to see if it is the best model among the candidate models. Next we look at the residuals to see if there is any autocorrelation remaining in the residuals.

**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



There still seems to be autocorrelation in the residuals which are usually indicative that the model is **not appropriate**.

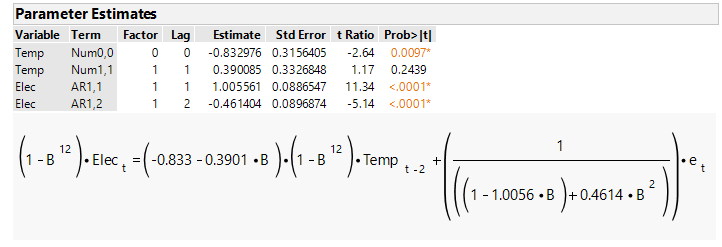
Candidate Model 3: ARMAX (2,0,1) + ARIMA (2,0,0)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, an ARMA MA term s=1, a Seasonal Difference of order 1 and an Autoregressive Term of order 2 for the Noise Model. The equation of the model is given below.

**yt = (ω0-ω1B)Xt-2 +**

**Parameter Estimates**

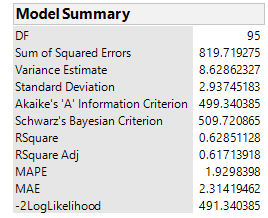
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for the model. In the output below the parameter estimates have been specified.



The ARMAX MA parameter s=1 does not appear to be significant. The model is probably **not appropriate.**

**Performance Metrics**

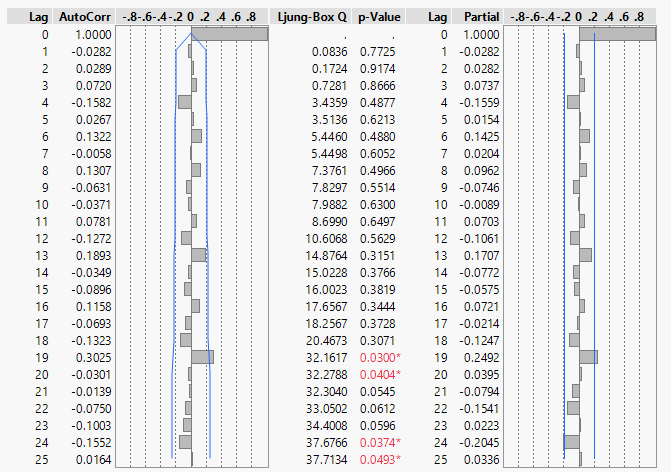
Some of the performance characteristics obtained during the model fitting process are given below.



**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



The residuals seem to behave like white noise.

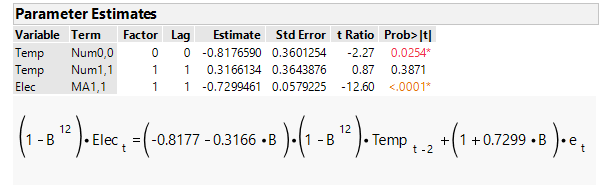
Candidate Model 4: ARMAX (2,0,1) + ARIMA (0,0,1)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, an ARMAX MA term s=1, a Seasonal difference of order 1 and Moving Average Term of order 1 for the Noise Model. The equation of the model is given below.

**yt = (ω0-ω1B)Xt-2 +**

**Parameter Estimates**

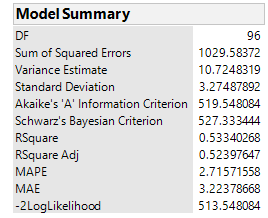
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for model. In the output below the parameter estimates have been specified.



Not all parameters of the model seem to be significant. The model **is probably not appropriate.**

**Performance Metrics**

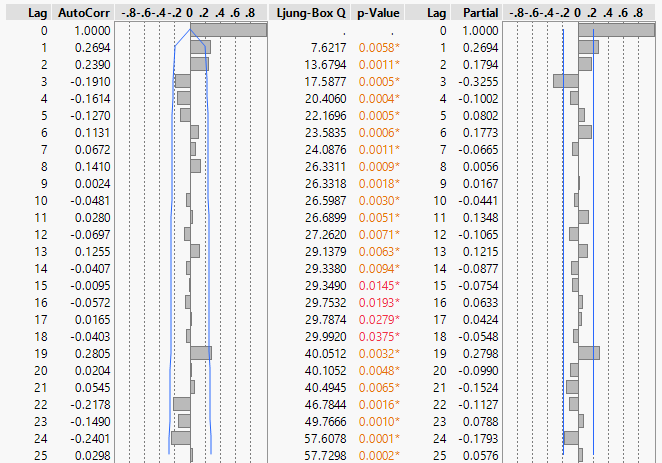
Some of the performance characteristics obtained during the model fitting process are given below.



**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



The residuals do not behave like white noise and hence it is further confirmed that the model is **inappropriate**.

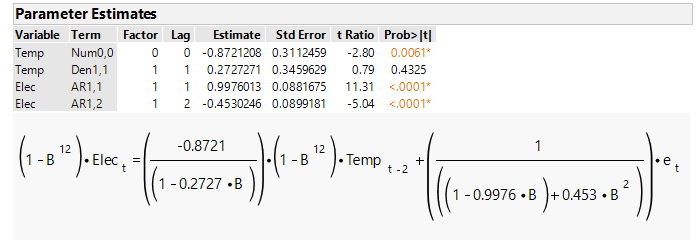
Candidate Model 5: ARMAX (2,1,0) + ARIMA (2,0,0)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, an ARMA AR term r=1, a Seasonal Difference of order 1 and an Autoregressive Term of order 2 for the Noise Model. The equation of the model is given below.

**yt = +**

**Parameter Estimates**

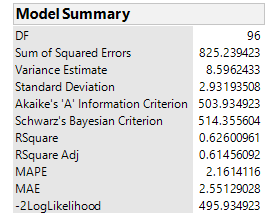
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for model. In the output below the parameter estimates have been specified.



Not all parameters of the model seem to be significant. The model **is probably not appropriate.**

**Performance Metrics**

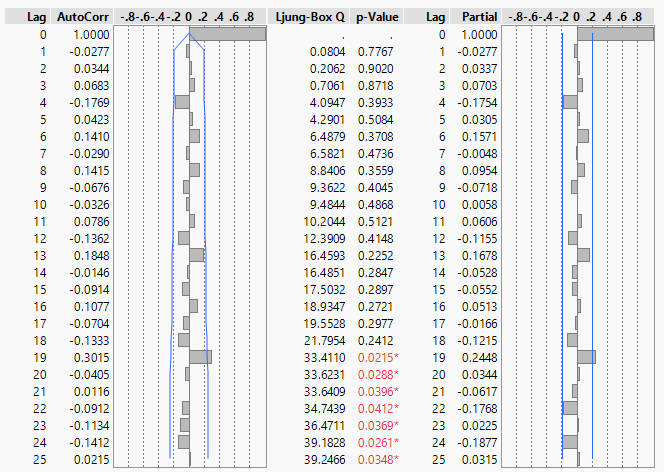
Some of the performance characteristics obtained during the model fitting process are given below.



**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



Residuals seem to resemble white noise.

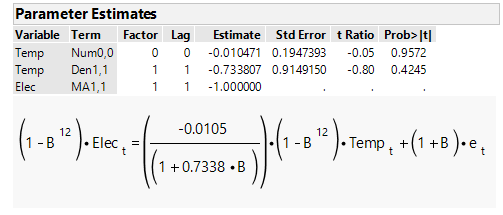
Candidate Model 6: ARMAX (2,1,0) + ARIMA (0,0,1)x(0,1,0)

In this candidate model we have a time delay of 2 for the Transfer Function, an ARMAX AR term r=1, a Seasonal difference of order 1 and Moving Average Term of order 1 for the Noise Model. The equation of the model is given below.

**yt = +**

**Parameter Estimates**

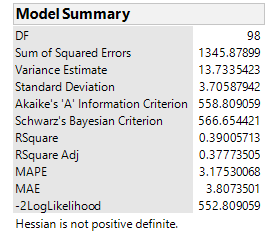
We now build the overall model using the transfer function functionality in JMP and estimate the parameters for model. In the output below the parameter estimates have been specified.



None of the parameters estimated appear to be significant. The model is probably **inappropriate**.

**Performance Metrics**

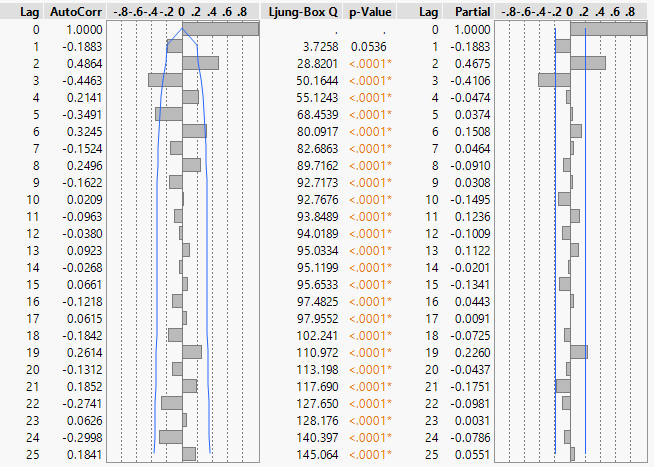
Some of the performance characteristics obtained during the model fitting process are given below.



**Residual Analysis**

ACF and PACF Plots

We look at the ACF and PACF plots of the transfer function model to see if the residuals behave like white noise.



The residuals don’t seem to be behaving like white noise. Autocorrelation still exists in the residuals.

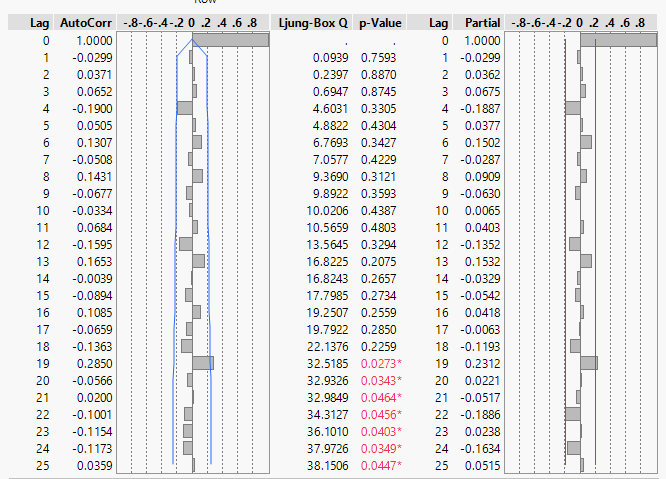
**Final Model Building and Model Adequacy Checks**

|  |  |  |
| --- | --- | --- |
| **Model Number** | **Transfer Function - Noise Model** | **Validity of Candidates** |
| **1** | ARMAX (2,0,0) + ARIMA (2,0,0)x(0,1,0) | Valid |
| **2** | ARMAX (2,0,0) + ARIMA (0,0,1)x(0,1,0) | Not Valid. Autocorrelation in residuals |
| **3** | ARMAX (2,0,1) + ARIMA (2,0,0)x(0,1,0) | Not Valid. Parameter Insignificant |
| **4** | ARMAX (2,0,1) + ARIMA (0,0,1)x(0,1,0) | Not Valid. Parameter Insignificant |
| **5** | ARMAX (2,1,0) + ARIMA (2,0,0)x(0,1,0) | Not Valid. Parameter Insignificant |
| **6** | ARMAX (2,1,0) + ARIMA (0,0,1)x(0,1,0) | Not Valid. Parameter Insignificant |

The Final Model has been identified and is **highlighted in green** above. We proceed to perform Final Adequacy checks on the model and report the performance metrics for the final model.

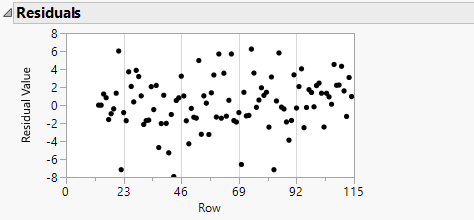
Model Adequacy and Final Model Performance Metrics

**ACF and PACF Plots**



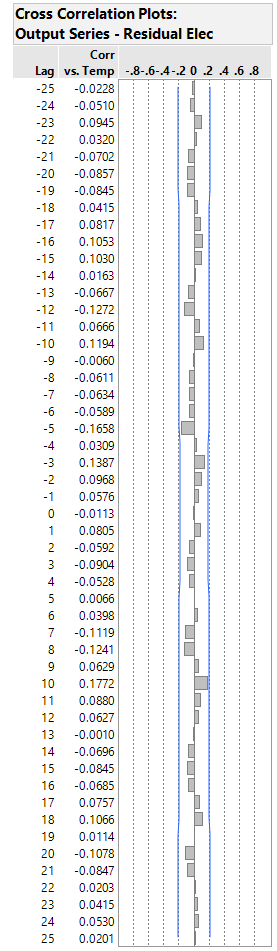
ACF and PACF plots do not show any autocorrelation and resemble white noise.

**Time Order of Residuals**



The pattern in the residuals appears to be random.

**Independence of Input Time Series and Residuals**

****

Looking at the cross-correlation plot between the input time series and the residuals, it is clear that they are independent of each other and hence the validity of the model is confirmed. The final model selected is **ARMAX (2,0,0) + ARIMA (2,0,0)x(0,1,0)**.

**Final Model Performance Metrics**

|  |  |
| --- | --- |
| **Performance Criteria** | **Final Transfer Function-Noise Model** |
| **Variance** | 8.573 |
| **AIC** | 502.694 |
| **BIC** | 510.509 |
| **MAPE** | 2.022 |
| **MAE** | 2.411 |
| **MSE** | 8.573 |
| **R2** | 0.623 |
| **R2 Adjusted** | 0.615 |
| **MAD** | 2.342 |
| **-2 Log Likelihood** | 496.694 |

**Task 5: Forecasting**

Now that the model has been built, we use the above models to forecast the input and output series for the next 6 time periods. The forecasts for the next 6 periods are given below and are also recorded in the csv file provided.

Input Series Temperature Forecasts

We first look at the forecasts for the input series temperature. We use the pre-whitening model **ARIMA(2,0,0)x(0,1,0)12** in previous sections to estimate the forecasted values.

|  |  |  |
| --- | --- | --- |
| **Year** | **Month** | **Temp Forecast** |
| **2006** | **7** | 23.2200572 |
| **2006** | **8** | 23.6378116 |
| **2006** | **9** | 19.65914518 |
| **2006** | **10** | 15.26308355 |
| **2006** | **11** | 17.04574364 |
| **2006** | **12** | 19.89069833 |

Output Series Electricity Forecasts

For the output series Electricity, we first look at the time series plot to check for seasonal patterns in the data.



Based on the time series plot the output variable Electricity looks seasonal in nature. We look at the ACF and PACF plots to confirm this.

**ACF and PACF Plots**

** **

Based on the ACF and PACF plots we see obvious seasonal patterns in the data. So, we look at the ACF and PACF plots of the Seasonally Differenced Data to identify patterns.

Based on the ACF and PACF plots for Seasonally Differenced Electricity Data we have the below candidate models.

|  |  |
| --- | --- |
| **Model Number** | **Candidate Model for Seasonally Differenced Data** |
| **1** | ARIMA(0,0,1) |
| **2** | ARIMA(2,0,0) |

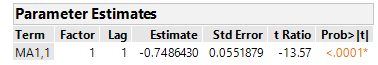
Candidate Model 1: ARIMA (0,0,1)x(0,1,0)12

In this candidate model we have one Normal MA parameter and one order of seasonal difference of lag 12. The equation of the Seasonal ARIMA model used is given below.

**yt =**

**Parameter Estimation**

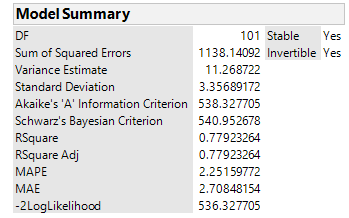
We now build a seasonal ARIMA model using the temperate data and estimate the parameters for model using JMP. In the output below the parameter estimates have been specified.



**As seen above the parameter estimates for the model are significant.**

**Performance Metrics**

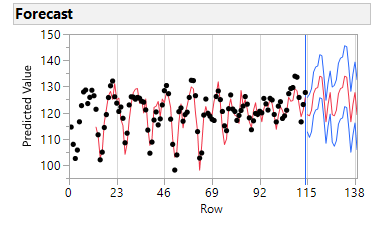
Some of the performance characteristics obtained during the model fitting process are given below.



Based on the performance characteristics we can say that the model performs well in this regard. Before we can conclude about the validity of the model we need to perform adequacy checks on the model using Residual Analysis.

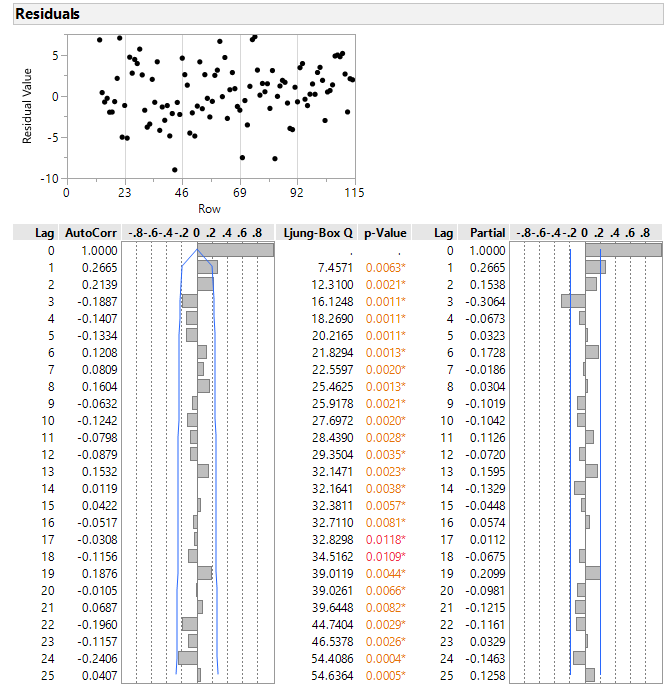
**Model Adequacy**

Actual Vs Fit Plot



Looking at the plot above, we can say that the fit of the model to the actual data looks reasonable.

Residual Plots



Residuals look random. But autocorrelation can still be seen in the Residual ACF and PACF plots. This model **does not seem appropriate** because of the presence of autocorrelation in the residuals

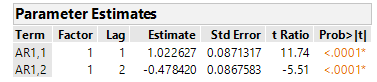
Candidate Model 2: ARIMA (2,0,0)x(0,1,0)12

In this candidate model we have 2 AR parameters and one order of seasonal difference of lag 12. The equation of the Seasonal ARIMA model used is given below.

**yt =**

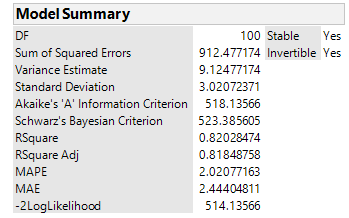
**Parameter Estimates**

We now build a seasonal ARIMA model using the temperate data and estimate the parameters for model using JMP. In the output below the parameter estimates have been specified.



**Performance Metrics**

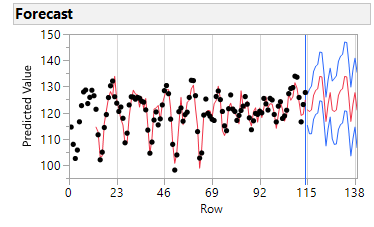
Some of the performance characteristics obtained during the model fitting process are given below.



Based on the performance characteristics we can say that the model performs well in this regard. Before we can conclude about the validity of the model we need to perform adequacy checks on the model using Residual Analysis.

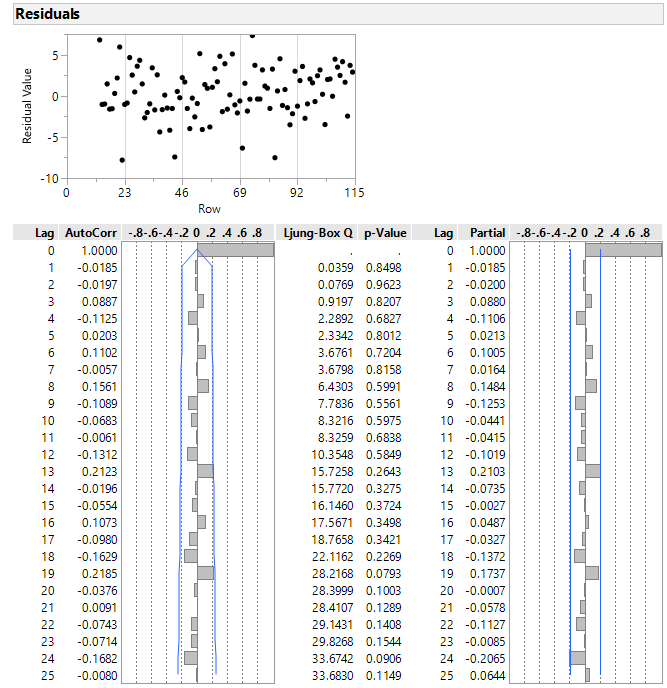
**Model Adequacy**

Actual Vs Fit Plot



Looking at the plot above, we can say that the fit of the model to the actual data looks reasonable.

Residual Plots



Looking at the ACF and PACF plots above, we can conclude that no autocorrelation is present in the residuals and hence use of this Seasonal ARIMA model to model the Electricity time series is valid. Among the 2 candidate models identified we **choose Candidate Model 2** **( ARIMA (2,0,0)x(0,1,0)12 )** since its performance and adequacy is better than Candidate Model 1. We use this model to forecast the Electricity Consumption for the next 6 periods.

|  |  |  |
| --- | --- | --- |
| **Year** | **Month** | **Elec Forecast** |
| **2006** | **7** | 121.0728 |
| **2006** | **8** | 120.4681 |
| **2006** | **9** | 121.0909 |
| **2006** | **10** | 126.6167 |
| **2006** | **11** | 128.4797 |
| **2006** | **12** | 129.1655 |

**References**

**1.** Montgomery, D., Kulahci, M., & Jennings, C. (2016). *Introduction to time series analysis and forecasting* (2nd ed.).