IEE 579 CASE STUDY 2: TIME SERIES MODELLING OF TEMPERATURE AND ELECTRICITY DATA

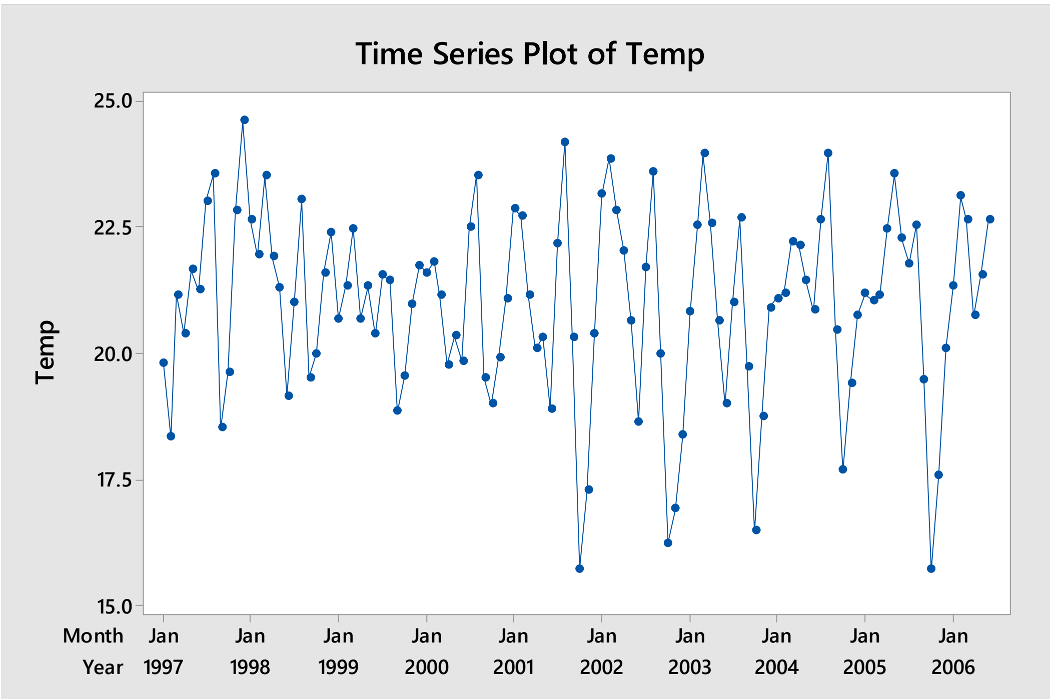
Ashlin Sreedhar

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.

<u>Time Series of the Input Variable: Temperature</u>

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.

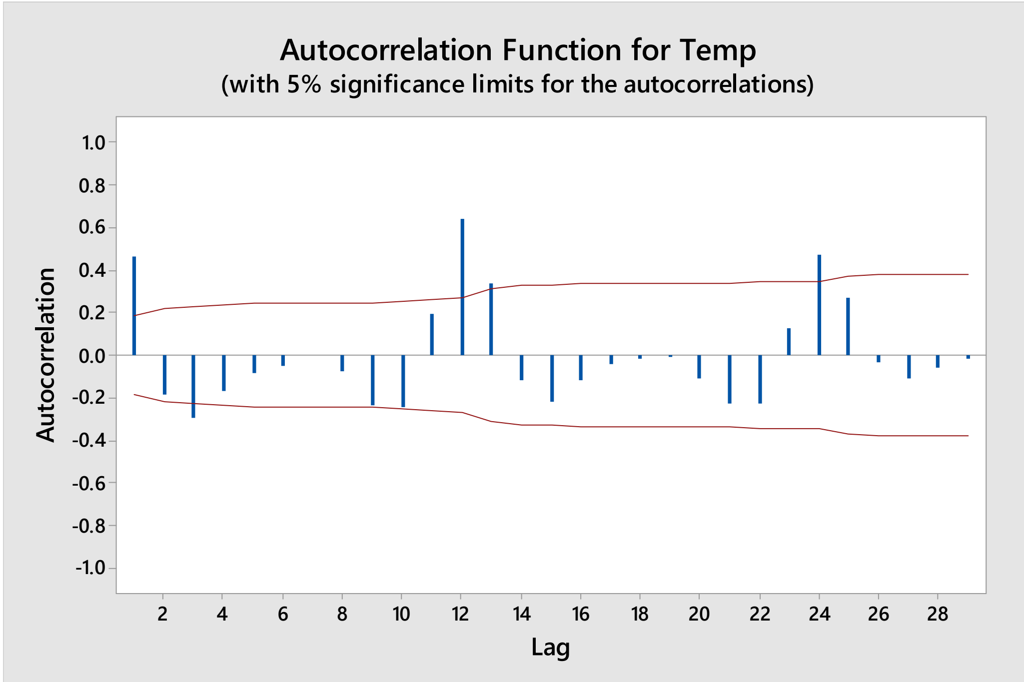


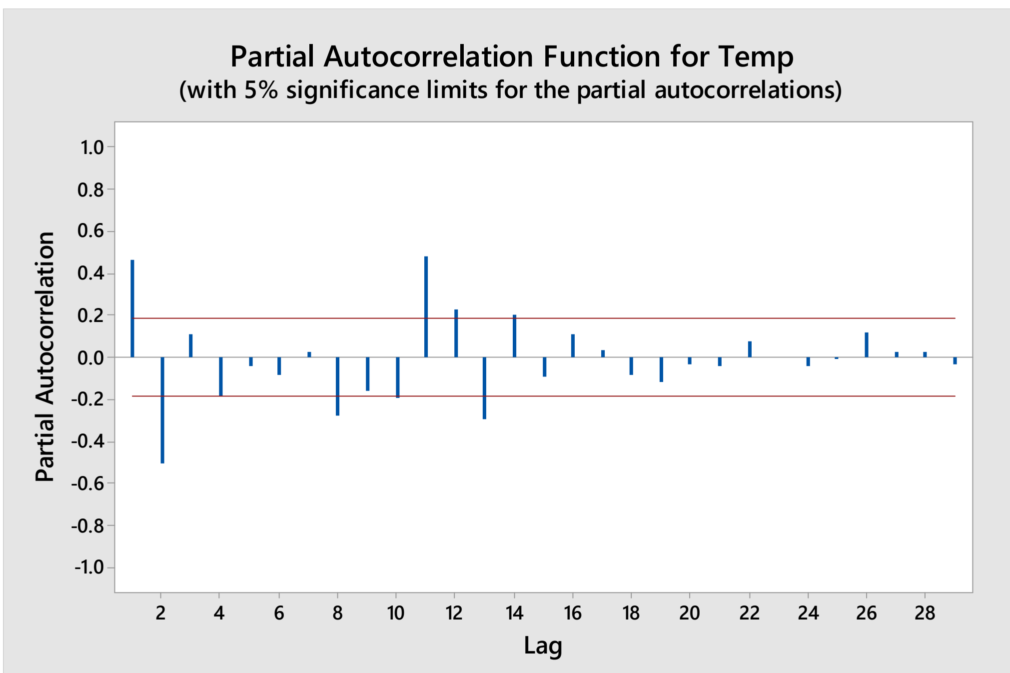
Based on the time series plot of the input variable we can clearly say that the data is seasonal because of the cyclic pattern that the data exhibits. A trend in the data seems to be absent. The constant variance assumption of the data also seems to hold for this time series. Considering that this is 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.

<u>Stationarity Checks</u>

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.

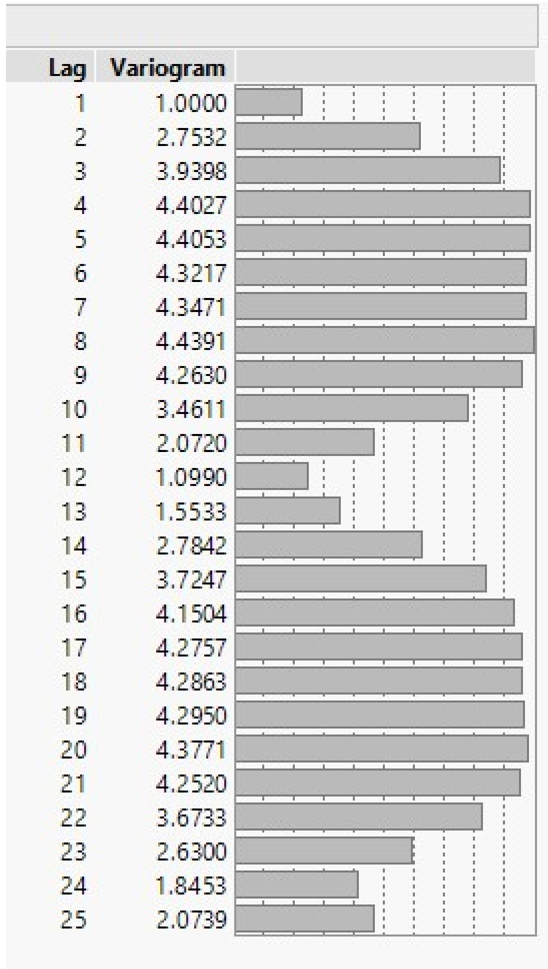
ACF and PACF Plots





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



Lag	Variogram
1	1.0000
2	2.7532
3	3.9398
4	4.4027
5	4.4053
6	4.3217
7	4.3471
8	4.4391
9	4.2630
10	3.4611
11	2.0720
12	1.0990
13	1.5533
14	2.7842
15	3.7247
16	4.1504
17	4.2757
18	4.2863
19	4.2950
20	4.3771
21	4.2520
22	3.6733
23	2.6300
24	1.8453
25	2.0739

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.

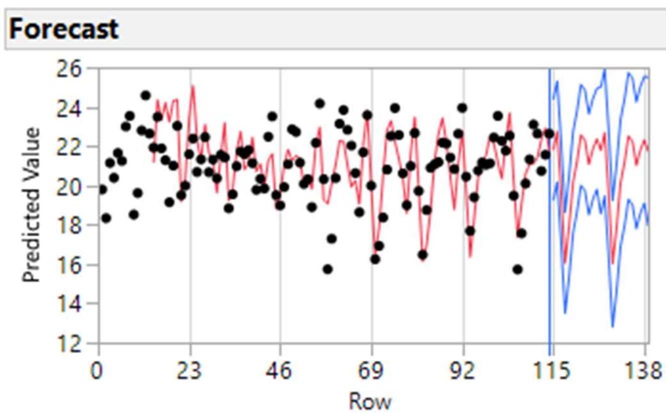
Parameter Estimation

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	1.74478e-8	1.8354e-7	0.10	0.9245
Trend Smoothing Weight	0.00009885	0.0012006	0.08	0.9346
Seasonal Smoothing Weight	0.76083326	0.1390828	5.47	<.0001*

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

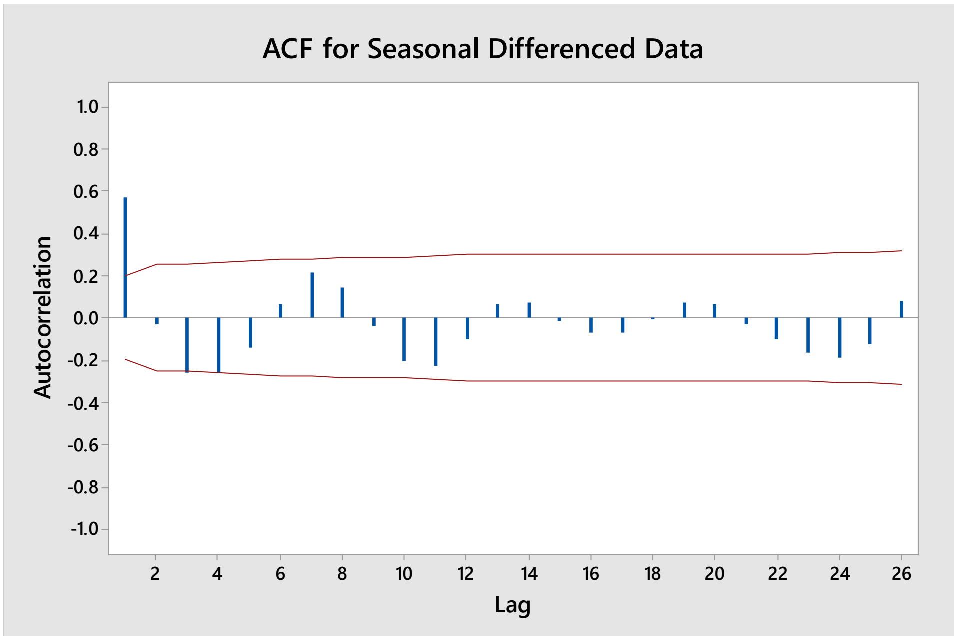
Model Summary

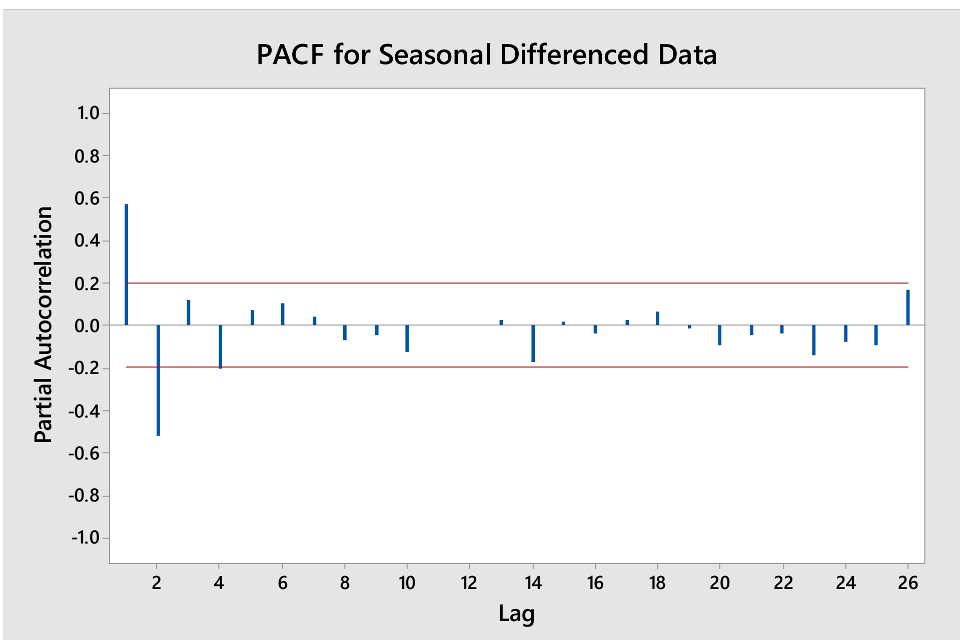
DF	98
Sum of Squared Errors	167.498463
Variance Estimate	1.70916799
Standard Deviation	1.30735152
Akaike's 'A' Information Criterion	349.518438
Schwarz's Bayesian Criterion	357.3638
RSquare	0.45996215
RSquare Adj	0.44894097
MAPE	5.14585419
MAE	1.06607863
-2LogLikelihood	343.518438

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 ($1-B^{12}$) 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 $(0,0,1)\mathrm{x}(0,1,0)_{12}$

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.

$$y_t = \delta + \frac{(1-\theta B)\varepsilon_t}{(1-B^{12})}$$

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.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant	Mu
MA1,1	1	1	-0.8176837	0.0514951	-15.88	<.0001*	Estimate	-0.039506
Intercept	1	0	-0.0395060	0.1644022	-0.24	0.8106	-0.039506	

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.

$$y_t = \frac{(1-\theta B)\varepsilon_t}{(1-B^{12})}$$

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA1,1	1	1	-0.8177915	0.0514782	-15.89	<.0001*

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.

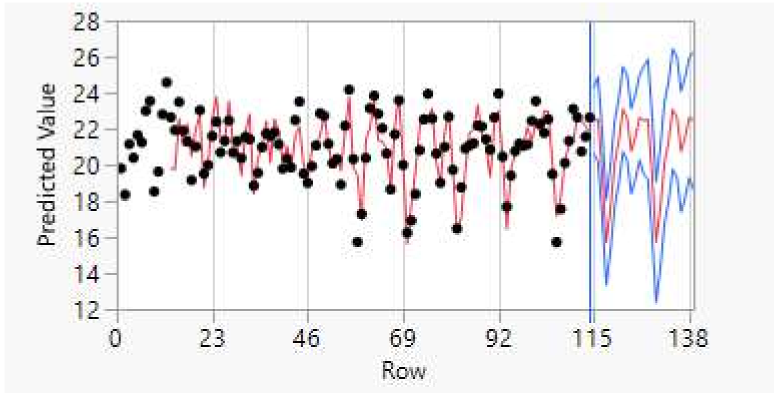
Model Summary

DF	101	Stable	Yes
Sum of Squared Errors	87.9228195	Invertible	Yes
Variance Estimate	0.87052296		
Standard Deviation	0.9330182		
Akaike's 'A' Information Criterion	277.420043		
Schwarz's Bayesian Criterion	280.045016		
RSquare	0.727054		
RSquare Adj	0.727054		
MAPE	3.56215489		
MAE	0.73546414		
-2LogLikelihood	275.420043		

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.

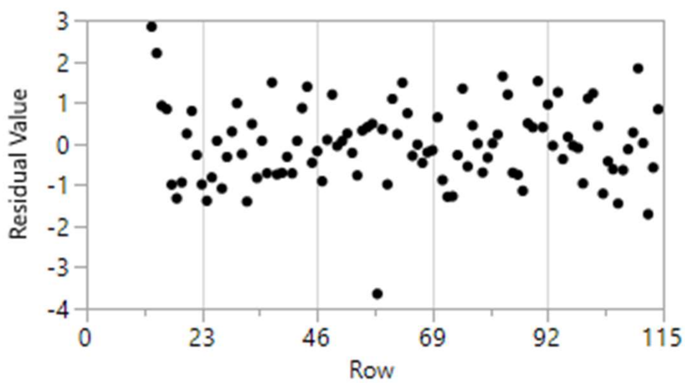
Model Adequacy

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.

$$y_t = \delta + \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

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.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.9041737	0.0834660	10.83	<.0001*
AR1,2	1	2	-0.5696798	0.0869097	-6.55	<.0001*
Intercept	1	0	-0.0516304	0.1325805	-0.39	0.6978

Constant Estimate	Mu
-0.0343603	-0.0516304

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.

$$y_t = \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.9049474	0.0835075	10.84	<.0001*
AR1,2	1	2	-0.5678162	0.0868562	-6.54	<.0001*

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.

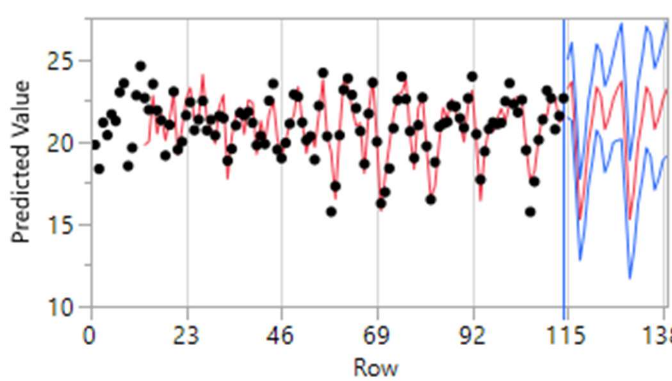
Model Summary

DF	100	Stable	Yes
Sum of Squared Errors	80.5651278	Invertible	Yes
Variance Estimate	0.80565128		
Standard Deviation	0.89758079		
Akaike's 'A' Information Criterion	270.584606		
Schwarz's Bayesian Criterion	275.834552		
RSquare	0.74553945		
RSquare Adj	0.74299485		
MAPE	3.45575287		
MAE	0.70940465		
-2LogLikelihood	266.584606		

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.

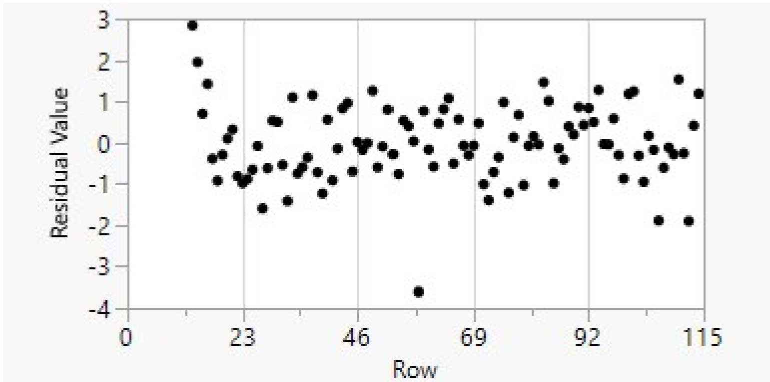
Model Adequacy

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.

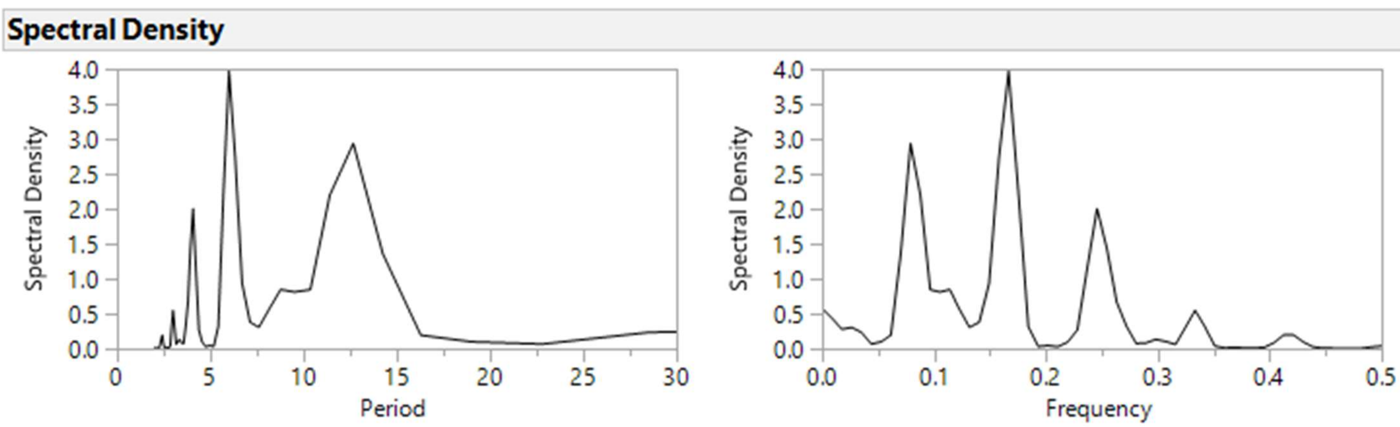
Model Comparison: Holt-Winters Exponential Smoothing Model Vs Seasonal ARIMA $(2,0,0)x(0,1,0)_{12}$

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
R^2 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.

<u>Task 3: Spectral Analysis</u>



White Noise test

Fisher's Kappa	11.752839
Prob > Kappa	0.0001323
Bartlett's Kolmogorov-Smirnov	0.3899686

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.

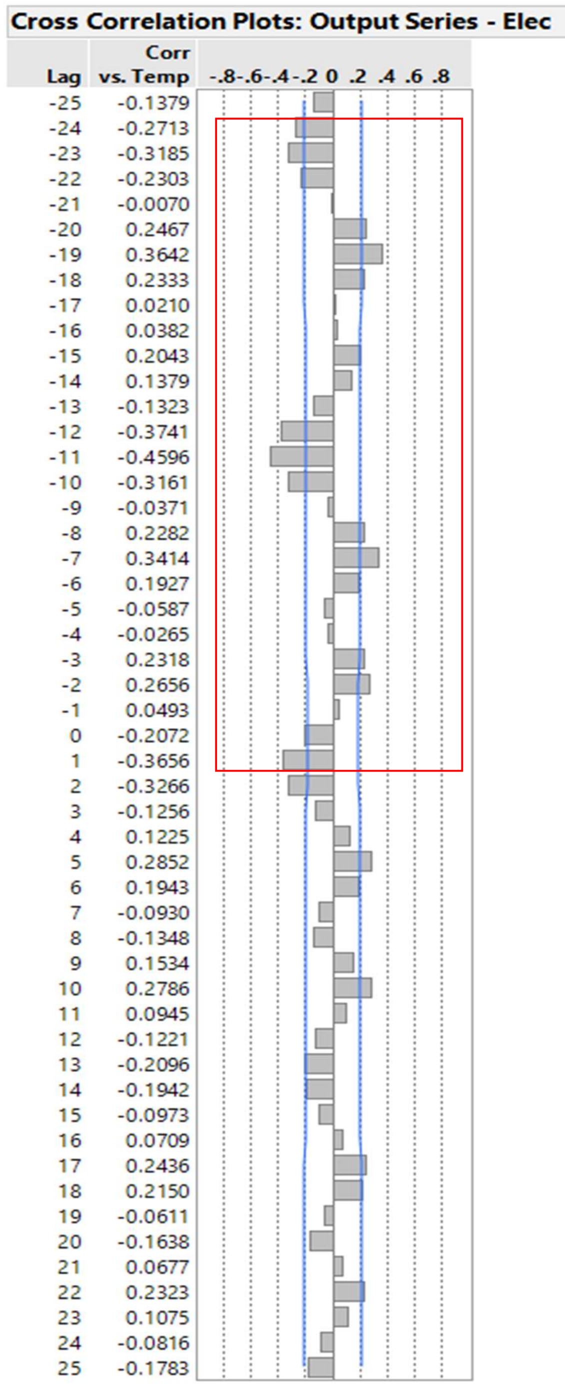
	Period	Frequency	Angular Frequency	Sine	Cosine	Periodogram	Spectral Density
1	•	0	0	0	0	0	0.5737972359
2	114	0.0087719298	0.0551156606	0.2188367252	0.2803771433	7.2105487237	0.4373570825
3	57	0.0175438596	0.1102313212	0.078059956	-0.009363287	0.3523185888	0.2785298681
4	38	0.0263157895	0.1653469818	0.2483136201	-0.212365624	6.0852522984	0.3045790204
5	28.5	0.0350877193	0.2204626424	0.1961964191	-0.101987721	2.786988219	0.2326655386
6	22.8	0.0438596491	0.2755783029	0.0069810432	-0.02407551	0.0358168136	0.0668293335
7	19	0.0526315789	0.3306939635	0.0633784185	-0.069031887	0.5005868422	0.0992107037
8	16.285714286	0.0614035088	0.3858096241	-0.263241661	0.0001666467	3.9498833861	0.1929483797
9	14.25	0.0701754386	0.4409252847	0.1341592044	0.0691254044	1.2982897787	1.3562978157
10	12.666666667	0.0789473684	0.4960409453	-0.542292501	0.887198184	61.628501121	2.9340192783
11	11.4	0.0877192982	0.5511566059	-0.436164199	-0.460376803	22.924602541	2.195479505
12	10.363636364	0.0964912281	0.6062722665	0.0518625048	-0.218680907	2.8791303427	0.8444174746

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

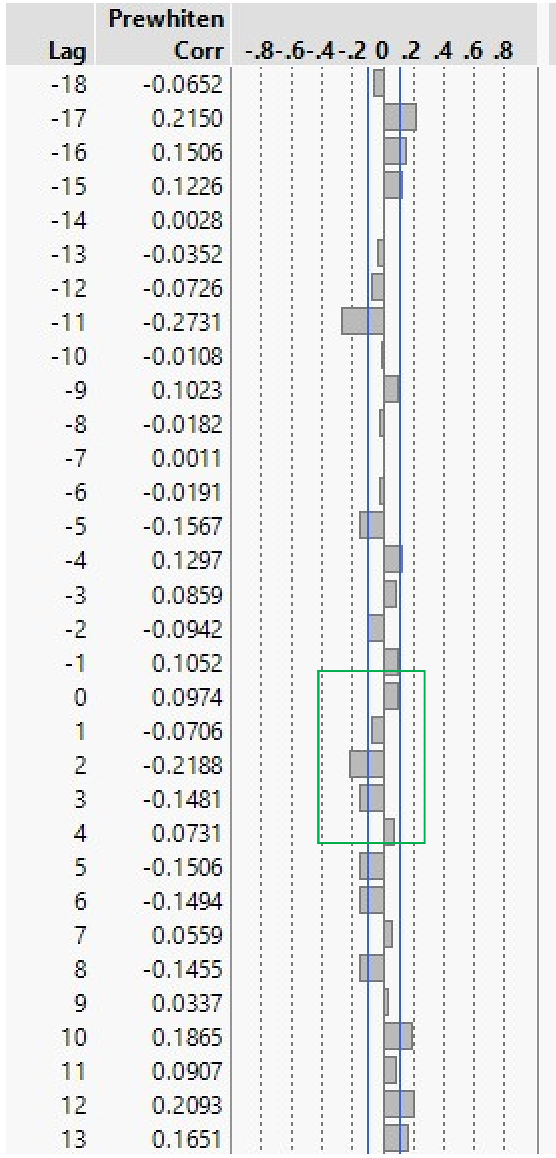
Cross-Correlation Before Pre-Whitening

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-1.3289	0.355760	-3.74	0.0003*
	Intercept	0	0	148.2911	7.483111	19.82	<.0001*

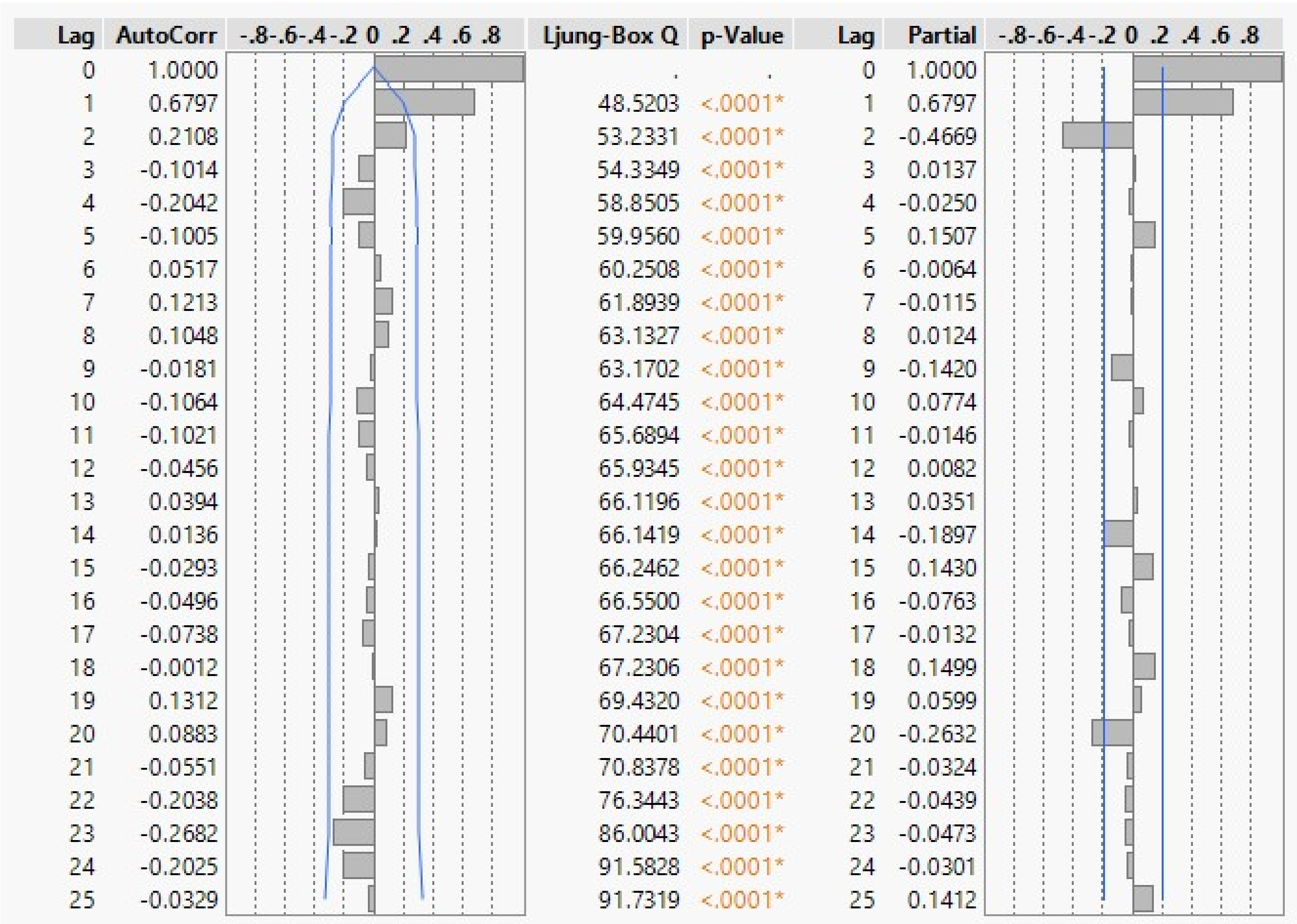
$$Elec_t = \left(148.2911 - 1.3289 \cdot Temp_{t-2} \right) + e_t$$

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

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.6356	47.2791	<.0001*	1	0.6356
2	0.1151	48.8432	<.0001*	2	-0.4847
3	-0.1003	50.0415	<.0001*	3	0.2176
4	-0.1625	53.2171	<.0001*	4	-0.2674
5	-0.2569	61.2237	<.0001*	5	-0.1289
6	-0.3336	74.8520	<.0001*	6	-0.1381
7	-0.2407	82.0123	<.0001*	7	0.0818
8	-0.1315	84.1681	<.0001*	8	-0.2471
9	-0.1140	85.8046	<.0001*	9	-0.0128
10	0.0221	85.8668	<.0001*	10	0.2246
11	0.4093	107.372	<.0001*	11	0.4836
12	0.6731	166.105	<.0001*	12	0.0732
13	0.4275	190.033	<.0001*	13	-0.2870
14	0.0450	190.301	<.0001*	14	0.1197
15	-0.0944	191.491	<.0001*	15	0.0329
16	-0.1115	193.168	<.0001*	16	0.0305
17	-0.2091	199.131	<.0001*	17	-0.1061
18	-0.3044	211.897	<.0001*	18	-0.0042
19	-0.2111	218.100	<.0001*	19	0.0557
20	-0.1077	219.732	<.0001*	20	-0.0880
21	-0.1020	221.211	<.0001*	21	0.0763
22	-0.0396	221.436	<.0001*	22	-0.1817
23	0.2302	229.134	<.0001*	23	0.0914
24	0.4330	256.683	<.0001*	24	-0.1095
25	0.2691	267.445	<.0001*	25	0.0719

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.



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.6797	48.5203	<.0001*	1	0.6797
2	0.2108	53.2331	<.0001*	2	-0.4669
3	-0.1014	54.3349	<.0001*	3	0.0137
4	-0.2042	58.8505	<.0001*	4	-0.0250
5	-0.1005	59.9560	<.0001*	5	0.1507
6	0.0517	60.2508	<.0001*	6	-0.0064
7	0.1213	61.8939	<.0001*	7	-0.0115
8	0.1048	63.1327	<.0001*	8	0.0124
9	-0.0181	63.1702	<.0001*	9	-0.1420
10	-0.1064	64.4745	<.0001*	10	0.0774
11	-0.1021	65.6894	<.0001*	11	-0.0146
12	-0.0456	65.9345	<.0001*	12	0.0082
13	0.0394	66.1196	<.0001*	13	0.0351
14	0.0136	66.1419	<.0001*	14	-0.1897
15	-0.0293	66.2462	<.0001*	15	0.1430
16	-0.0496	66.5500	<.0001*	16	-0.0763
17	-0.0738	67.2304	<.0001*	17	-0.0132
18	-0.0012	67.2306	<.0001*	18	0.1499
19	0.1312	69.4320	<.0001*	19	0.0599
20	0.0883	70.4401	<.0001*	20	-0.2632
21	-0.0551	70.8378	<.0001*	21	-0.0324
22	-0.2038	76.3443	<.0001*	22	-0.0439
23	-0.2682	86.0043	<.0001*	23	-0.0473
24	-0.2025	91.5828	<.0001*	24	-0.0301
25	-0.0329	91.7319	<.0001*	25	0.1412

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.

$$y_t = \omega_0 X_{t-2} + \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.8470830	0.3166323	-2.68	0.0088*
Elec	AR1,1	1	1	0.9954138	0.0891160	11.17	<.0001*
Elec	AR1,2	1	2	-0.4409541	0.0888091	-4.97	<.0001*

$$\left(1-B^{12}\right)\cdot \text{Elec}_t = -\left(0.8471\cdot\left(1-B^{12}\right)\cdot \text{Temp}_{t-2}\right)+\left(\frac{1}{\left(\left(1-0.9954\cdot B\right)+0.441\cdot B^2\right)}\right)\cdot e_t$$

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.

Model Summary

DF	97
Sum of Squared Errors	831.660592
Variance Estimate	8.57381979
Standard Deviation	2.92810857
Akaike's 'A' Information Criterion	502.694127
Schwarz's Bayesian Criterion	510.509638
RSquare	0.6230996
RSquare Adj	0.61548545
MAPE	2.02281247
MAE	2.41195455
-2LogLikelihood	496.694127

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.

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	-0.0299	0.0939	0.7593	1	-0.0299
2	0.0371	0.2397	0.8870	2	0.0362
3	0.0652	0.6947	0.8745	3	0.0675
4	-0.1900	4.6031	0.3305	4	-0.1887
5	0.0505	4.8822	0.4304	5	0.0377
6	0.1307	6.7693	0.3427	6	0.1502
7	-0.0508	7.0577	0.4229	7	-0.0287
8	0.1431	9.3690	0.3121	8	0.0909
9	-0.0677	9.8922	0.3593	9	-0.0630
10	-0.0334	10.0206	0.4387	10	0.0065
11	0.0684	10.5659	0.4803	11	0.0403
12	-0.1595	13.5645	0.3294	12	-0.1352
13	0.1653	16.8225	0.2075	13	0.1532
14	-0.0039	16.8243	0.2657	14	-0.0329
15	-0.0894	17.7985	0.2734	15	-0.0542
16	0.1085	19.2507	0.2559	16	0.0418
17	-0.0659	19.7922	0.2850	17	-0.0063
18	-0.1363	22.1376	0.2259	18	-0.1193
19	0.2850	32.5185	0.0273*	19	0.2312
20	-0.0566	32.9326	0.0343*	20	0.0221
21	0.0200	32.9849	0.0464*	21	-0.0517
22	-0.1001	34.3127	0.0456*	22	-0.1886
23	-0.1154	36.1010	0.0403*	23	0.0238
24	-0.1173	37.9726	0.0349*	24	-0.1634
25	0.0359	38.1506	0.0447*	25	0.0515

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.

$$y_t = \omega_0 X_{t-2} + \frac{(1-\theta_1 B)\varepsilon_t}{(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.8951127	0.3436275	-2.60	0.0106*
Elec	MA1,1	1	1	-0.7234021	0.0576017	-12.56	<.0001*

$$\left(1-B^{12}\right)\cdot Elec_t = -\left(0.8951\cdot\left(1-B^{12}\right)\cdot Temp_{t-2}\right)+\left(1+0.7234\cdot B\right)\cdot e_t$$

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.

Model Summary	
DF	98
Sum of Squared Errors	1039.40363
Variance Estimate	10.6061608
Standard Deviation	3.25671013
Akaike's 'A' Information Criterion	522.651814
Schwarz's Bayesian Criterion	527.862154
RSquare	0.52895239
RSquare Adj	0.52424192
MAPE	2.53981689
MAE	3.03055868
-2LogLikelihood	518.651814

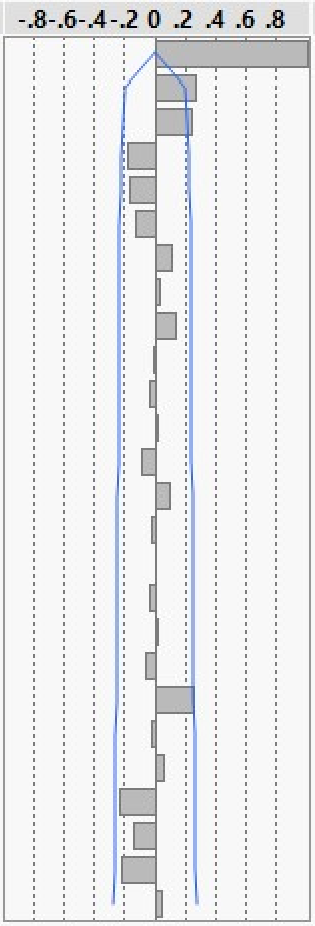
The performance metrics above look reasonable. We still need to compare this model with other candidate models to see if it is the best model among the candidate models. Next we look at the residuals to see if there is any autocorrelation remaining in the residuals.

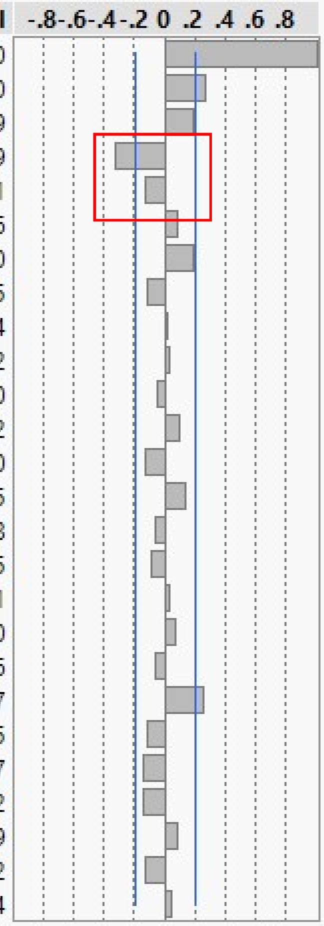
Residual Analysis

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.

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.2690	7.5973	0.0058*	1	0.2690
2	0.2494	14.1964	0.0008*	2	0.1909
3	-0.1768	17.5447	0.0005*	3	-0.3159
4	-0.1668	20.5551	0.0004*	4	-0.1211
5	-0.1226	22.1982	0.0005*	5	0.0906
6	0.1061	23.4412	0.0007*	6	0.1860
7	0.0344	23.5732	0.0014*	7	-0.1145
8	0.1436	25.9008	0.0011*	8	0.0244
9	-0.0072	25.9068	0.0021*	9	0.0272
10	-0.0365	26.0607	0.0037*	10	-0.0510
11	0.0143	26.0847	0.0063*	11	0.1002
12	-0.0909	27.0588	0.0076*	12	-0.1210
13	0.1006	28.2647	0.0083*	13	0.1335
14	-0.0262	28.3474	0.0128*	14	-0.0578
15	0.0011	28.3475	0.0195*	15	-0.0795
16	-0.0387	28.5326	0.0273*	16	0.0341
17	0.0251	28.6113	0.0383*	17	0.0730
18	-0.0534	28.9710	0.0487*	18	-0.0606
19	0.2618	37.7306	0.0064*	19	0.2547
20	-0.0168	37.7670	0.0095*	20	-0.1115
21	0.0587	38.2183	0.0121*	21	-0.1337
22	-0.2311	45.3021	0.0024*	22	-0.1412
23	-0.1377	47.8476	0.0017*	23	0.0879
24	-0.2145	54.1068	0.0004*	24	-0.1252
25	0.0503	54.4547	0.0006*	25	0.0414





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.

$$y_t = (\omega_0\text{-}\omega_1 B)\, X_{t\text{-}2} + \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.832976	0.3156405	-2.64	0.0097*
Temp	Num1,1	1	1	0.390085	0.3326848	1.17	0.2439
Elec	AR1,1	1	1	1.005561	0.0886547	11.34	<.0001*
Elec	AR1,2	1	2	-0.461404	0.0896874	-5.14	<.0001*

$$\left(1-B^{12}\right)\bullet \text{Elec}_t = \left(-0.833 - 0.3901 \bullet B\right)\bullet\left(1-B^{12}\right)\bullet \text{Temp}_{t-2} + \left(\frac{1}{\left(\left(1-1.0056\bullet B\right)+0.4614\bullet B^2\right)}\right)\bullet e_t$$

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.

Model Summary

DF	95
Sum of Squared Errors	819.719275
Variance Estimate	8.62862327
Standard Deviation	2.93745183
Akaike's 'A' Information Criterion	499.340385
Schwarz's Bayesian Criterion	509.720865
RSquare	0.62851128
RSquare Adj	0.61713918
MAPE	1.9298398
MAE	2.31419462
-2LogLikelihood	491.340385

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.

Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	-0.0282		0.0836	0.7725	1	-0.0282	
2	0.0289		0.1724	0.9174	2	0.0282	
3	0.0720		0.7281	0.8666	3	0.0737	
4	-0.1582		3.4359	0.4877	4	-0.1559	
5	0.0267		3.5136	0.6213	5	0.0154	
6	0.1322		5.4460	0.4880	6	0.1425	
7	-0.0058		5.4498	0.6052	7	0.0204	
8	0.1307		7.3761	0.4966	8	0.0962	
9	-0.0631		7.8297	0.5514	9	-0.0746	
10	-0.0371		7.9882	0.6300	10	-0.0089	
11	0.0781		8.6990	0.6497	11	0.0703	
12	-0.1272		10.6068	0.5629	12	-0.1061	
13	0.1893		14.8764	0.3151	13	0.1707	
14	-0.0349		15.0228	0.3766	14	-0.0772	
15	-0.0896		16.0023	0.3819	15	-0.0575	
16	0.1158		17.6567	0.3444	16	0.0721	
17	-0.0693		18.2567	0.3728	17	-0.0214	
18	-0.1323		20.4673	0.3071	18	-0.1247	
19	0.3025		32.1617	0.0300*	19	0.2492	
20	-0.0301		32.2788	0.0404*	20	0.0395	
21	-0.0139		32.3040	0.0545	21	-0.0794	
22	-0.0750		33.0502	0.0612	22	-0.1541	
23	-0.1003		34.4008	0.0596	23	0.0223	
24	-0.1552		37.6766	0.0374*	24	-0.2045	
25	0.0164		37.7134	0.0493*	25	0.0336	

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.

$$y_t = (\omega_0 - \omega_1 B) X_{t-2} + \frac{(1-\theta_1 B)\varepsilon_t}{(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.8176590	0.3601254	-2.27	0.0254*
Temp	Num1,1	1	1	0.3166134	0.3643876	0.87	$\boxed{0.3871}$
Elec	MA1,1	1	1	-0.7299461	0.0579225	-12.60	<.0001*

$$\left(1 - B^{12}\right) \cdot Elec_t = \left(-0.8177 - 0.3166 \cdot B\right) \cdot \left(1 - B^{12}\right) \cdot Temp_{t-2} + \left(1 + 0.7299 \cdot B\right) \cdot e_t$$

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.

Model Summary	
DF	96
Sum of Squared Errors	1029.58372
Variance Estimate	10.7248319
Standard Deviation	3.27487892
Akaike's 'A' Information Criterion	519.548084
Schwarz's Bayesian Criterion	527.333444
RSquare	0.53340268
RSquare Adj	0.52397647
MAPE	2.71571558
MAE	3.22378668
-2LogLikelihood	513.548084

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.

Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	0.2694		7.6217	0.0058*	1	0.2694	
2	0.2390		13.6794	0.0011*	2	0.1794	
3	-0.1910		17.5877	0.0005*	3	-0.3255	
4	-0.1614		20.4060	0.0004*	4	-0.1002	
5	-0.1270		22.1696	0.0005*	5	0.0802	
6	0.1131		23.5835	0.0006*	6	0.1773	
7	0.0672		24.0876	0.0011*	7	-0.0665	
8	0.1410		26.3311	0.0009*	8	0.0056	
9	0.0024		26.3318	0.0018*	9	0.0167	
10	-0.0481		26.5987	0.0030*	10	-0.0441	
11	0.0280		26.6899	0.0051*	11	0.1348	
12	-0.0697		27.2620	0.0071*	12	-0.1065	
13	0.1255		29.1379	0.0063*	13	0.1215	
14	-0.0407		29.3380	0.0094*	14	-0.0877	
15	-0.0095		29.3490	0.0145*	15	-0.0754	
16	-0.0572		29.7532	0.0193*	16	0.0633	
17	0.0165		29.7874	0.0279*	17	0.0424	
18	-0.0403		29.9920	0.0375*	18	-0.0548	
19	0.2805		40.0512	0.0032*	19	0.2798	
20	0.0204		40.1052	0.0048*	20	-0.0990	
21	0.0545		40.4945	0.0065*	21	-0.1524	
22	-0.2178		46.7844	0.0016*	22	-0.1127	
23	-0.1490		49.7666	0.0010*	23	0.0788	
24	-0.2401		57.6078	0.0001*	24	-0.1793	
25	0.0298		57.7298	0.0002*	25	0.0576	

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.

$$y_t = \frac{X_{t-2}}{(1-\delta_1 B)} + \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.8721208	0.3112459	-2.80	0.0061*
Temp	Den1,1	1	1	0.2727271	0.3459629	0.79	$\boxed{0.4325}$
Elec	AR1,1	1	1	0.9976013	0.0881675	11.31	<.0001*
Elec	AR1,2	1	2	-0.4530246	0.0899181	-5.04	<.0001*

$$\left(1-B^{12}\right)\bullet \mathrm{Elec}_t = \left(\frac{-0.8721}{\left(1-0.2727\bullet B\right)}\right)\bullet\left(1-B^{12}\right)\bullet \mathrm{Temp}_{t-2} + \left(\frac{1}{\left(\left(1-0.9976\bullet B\right)+0.453\bullet B^2\right)}\right)\bullet e_t$$

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.

Model Summary

DF	96
Sum of Squared Errors	825.239423
Variance Estimate	8.5962433
Standard Deviation	2.93193508
Akaike's 'A' Information Criterion	503.934923
Schwarz's Bayesian Criterion	514.355604
RSquare	0.62600961
RSquare Adj	0.61456092
MAPE	2.1614116
MAE	2.55129028
-2LogLikelihood	495.934923

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.

Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	-0.0277		0.0804	0.7767	1	-0.0277	
2	0.0344		0.2062	0.9020	2	0.0337	
3	0.0683		0.7061	0.8718	3	0.0703	
4	-0.1769		4.0947	0.3933	4	-0.1754	
5	0.0423		4.2901	0.5084	5	0.0305	
6	0.1410		6.4879	0.3708	6	0.1571	
7	-0.0290		6.5821	0.4736	7	-0.0048	
8	0.1415		8.8406	0.3559	8	0.0954	
9	-0.0676		9.3622	0.4045	9	-0.0718	
10	-0.0326		9.4844	0.4868	10	0.0058	
11	0.0786		10.2044	0.5121	11	0.0606	
12	-0.1362		12.3909	0.4148	12	-0.1155	
13	0.1848		16.4593	0.2252	13	0.1678	
14	-0.0146		16.4851	0.2847	14	-0.0528	
15	-0.0914		17.5032	0.2897	15	-0.0552	
16	0.1077		18.9347	0.2721	16	0.0513	
17	-0.0704		19.5528	0.2977	17	-0.0166	
18	-0.1333		21.7954	0.2412	18	-0.1215	
19	0.3015		33.4110	0.0215*	19	0.2448	
20	-0.0405		33.6231	0.0288*	20	0.0344	
21	0.0116		33.6409	0.0396*	21	-0.0617	
22	-0.0912		34.7439	0.0412*	22	-0.1768	
23	-0.1134		36.4711	0.0369*	23	0.0225	
24	-0.1412		39.1828	0.0261*	24	-0.1877	
25	0.0215		39.2466	0.0348*	25	0.0315	

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.

$$y_t = \frac{X_{t-2}}{(1-\delta_1 B)} + \frac{(1-\theta_1 B)\varepsilon_t}{(1-B^{12})}$$

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.

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.010471	0.1947393	-0.05	0.9572
Temp	Den1,1	1	1	-0.733807	0.9149150	-0.80	0.4245
Elec	MA1,1	1	1	-1.000000	.	.	.

$$\left(1 - B^{12}\right)\cdot \text{Elec}_t = \left(\frac{-0.0105}{\left(1 + 0.7338 \cdot B\right)}\right)\cdot\left(1 - B^{12}\right)\cdot \text{Temp}_t + \left(1 + B\right)\cdot e_t$$

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.

Model Summary	
DF	98
Sum of Squared Errors	1345.87899
Variance Estimate	13.7335423
Standard Deviation	3.70587942
Akaike's 'A' Information Criterion	558.809059
Schwarz's Bayesian Criterion	566.654421
RSquare	0.39005713
RSquare Adj	0.37773505
MAPE	3.17530068
MAE	3.8073501
-2LogLikelihood	552.809059

Hessian is not positive definite.

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.

Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	-0.1883		3.7258	0.0536	1	-0.1883	
2	0.4864		28.8201	<.0001*	2	0.4675	
3	-0.4463		50.1644	<.0001*	3	-0.4106	
4	0.2141		55.1243	<.0001*	4	-0.0474	
5	-0.3491		68.4539	<.0001*	5	0.0374	
6	0.3245		80.0917	<.0001*	6	0.1508	
7	-0.1524		82.6863	<.0001*	7	0.0464	
8	0.2496		89.7162	<.0001*	8	-0.0910	
9	-0.1622		92.7173	<.0001*	9	0.0308	
10	0.0209		92.7676	<.0001*	10	-0.1495	
11	-0.0963		93.8489	<.0001*	11	0.1236	
12	-0.0380		94.0189	<.0001*	12	-0.1009	
13	0.0923		95.0334	<.0001*	13	0.1122	
14	-0.0268		95.1199	<.0001*	14	-0.0201	
15	0.0661		95.6533	<.0001*	15	-0.1341	
16	-0.1218		97.4825	<.0001*	16	0.0443	
17	0.0615		97.9552	<.0001*	17	0.0091	
18	-0.1842		102.241	<.0001*	18	-0.0725	
19	0.2614		110.972	<.0001*	19	0.2260	
20	-0.1312		113.198	<.0001*	20	-0.0437	
21	0.1852		117.690	<.0001*	21	-0.1751	
22	-0.2741		127.650	<.0001*	22	-0.0981	
23	0.0626		128.176	<.0001*	23	0.0031	
24	-0.2998		140.397	<.0001*	24	-0.0786	
25	0.1841		145.064	<.0001*	25	0.0551	

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

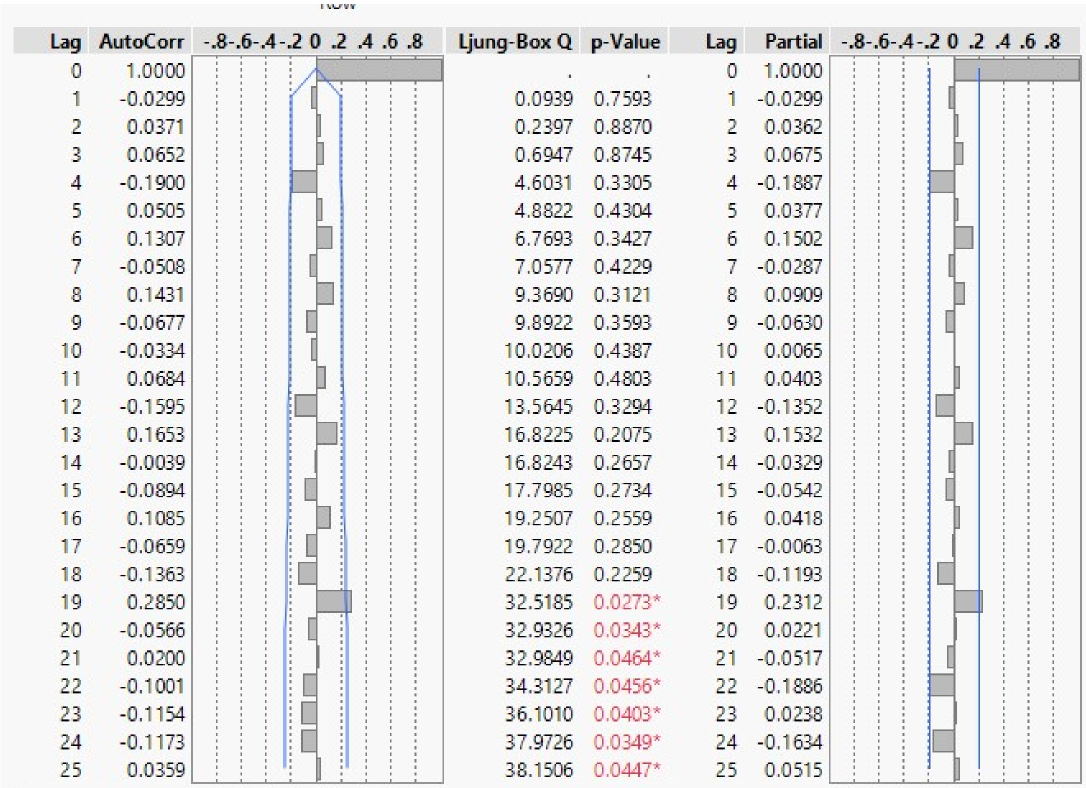
Final Model Building and Model Adequacy Checks

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.

<u>Model Adequacy and Final Model Performance Metrics</u>

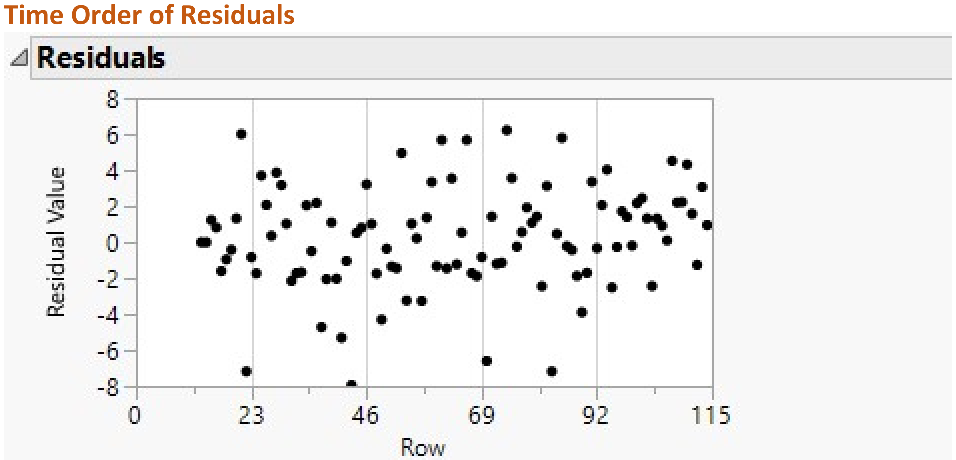
ACF and PACF Plots



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	-0.0299	0.0939	0.7593	1	-0.0299
2	0.0371	0.2397	0.8870	2	0.0362
3	0.0652	0.6947	0.8745	3	0.0675
4	-0.1900	4.6031	0.3305	4	-0.1887
5	0.0505	4.8822	0.4304	5	0.0377
6	0.1307	6.7693	0.3427	6	0.1502
7	-0.0508	7.0577	0.4229	7	-0.0287
8	0.1431	9.3690	0.3121	8	0.0909
9	-0.0677	9.8922	0.3593	9	-0.0630
10	-0.0334	10.0206	0.4387	10	0.0065
11	0.0684	10.5659	0.4803	11	0.0403
12	-0.1595	13.5645	0.3294	12	-0.1352
13	0.1653	16.8225	0.2075	13	0.1532
14	-0.0039	16.8243	0.2657	14	-0.0329
15	-0.0894	17.7985	0.2734	15	-0.0542
16	0.1085	19.2507	0.2559	16	0.0418
17	-0.0659	19.7922	0.2850	17	-0.0063
18	-0.1363	22.1376	0.2259	18	-0.1193
19	0.2850	32.5185	0.0273*	19	0.2312
20	-0.0566	32.9326	0.0343*	20	0.0221
21	0.0200	32.9849	0.0464*	21	-0.0517
22	-0.1001	34.3127	0.0456*	22	-0.1886
23	-0.1154	36.1010	0.0403*	23	0.0238
24	-0.1173	37.9726	0.0349*	24	-0.1634
25	0.0359	38.1506	0.0447*	25	0.0515

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

Time Order of Residuals

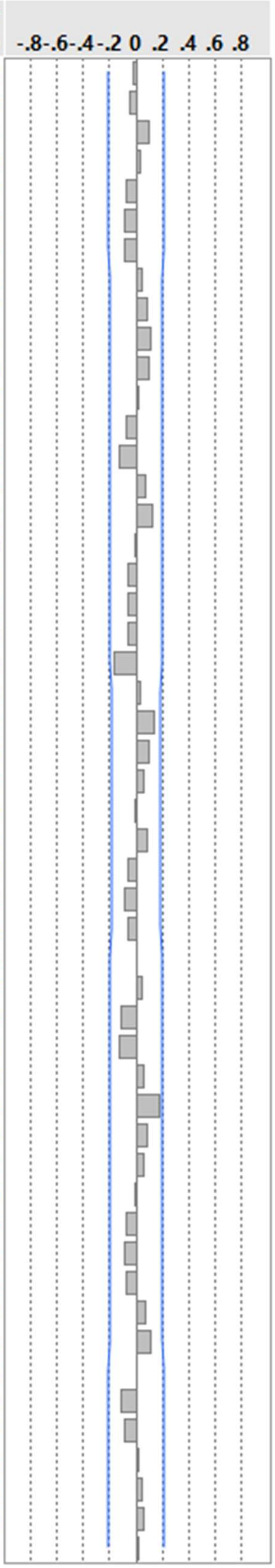


The pattern in the residuals appears to be random.

Independence of Input Time Series and Residuals

Cross Correlation Plots:
Output Series - Residual Elec

Lag	Corr vs. Temp
-25	-0.0228
-24	-0.0510
-23	0.0945
-22	0.0320
-21	-0.0702
-20	-0.0857
-19	-0.0845
-18	0.0415
-17	0.0817
-16	0.1053
-15	0.1030
-14	0.0163
-13	-0.0667
-12	-0.1272
-11	0.0666
-10	0.1194
-9	-0.0060
-8	-0.0611
-7	-0.0634
-6	-0.0589
-5	-0.1658
-4	0.0309
-3	0.1387
-2	0.0968
-1	0.0576
0	-0.0113
1	0.0805
2	-0.0592
3	-0.0904
4	-0.0528
5	0.0066
6	0.0398
7	-0.1119
8	-0.1241
9	0.0629
10	0.1772
11	0.0880
12	0.0627
13	-0.0010
14	-0.0696
15	-0.0845
16	-0.0685
17	0.0757
18	0.1066
19	0.0114
20	-0.1078
21	-0.0847
22	0.0203
23	0.0415
24	0.0530
25	0.0201



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
R^2	0.623
R^2 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.

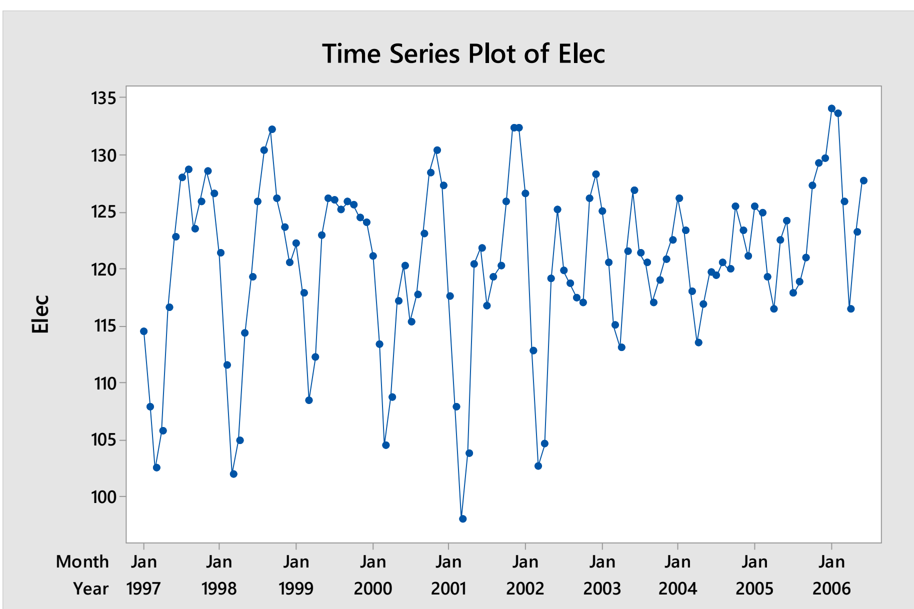
Input Series Temperature Forecasts

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

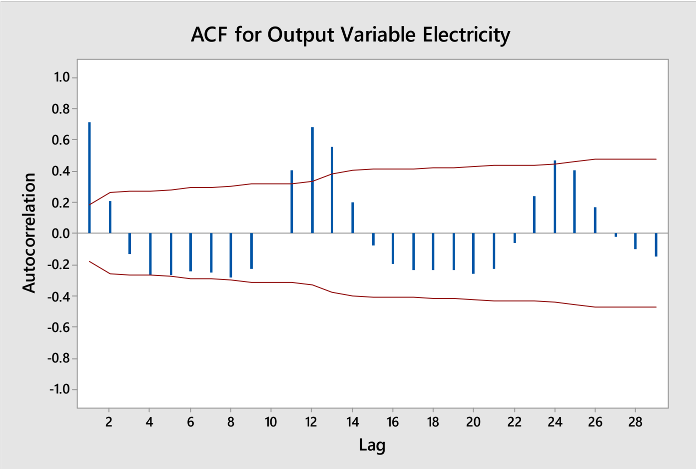
Output Series Electricity Forecasts

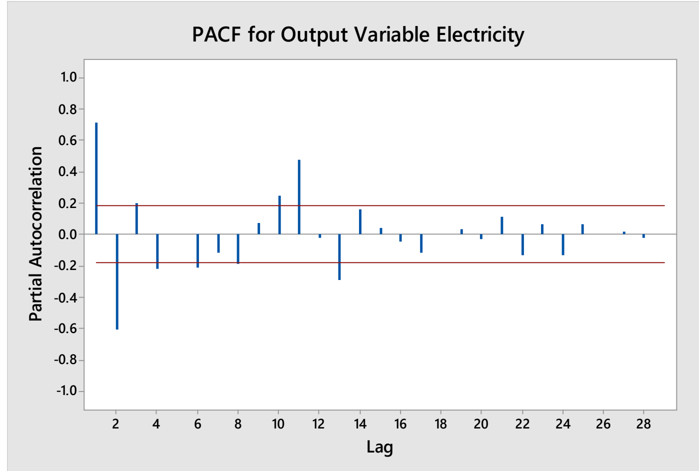
For the output series Electricity, we first look at the time series plot to check for seasonal patterns in the data.



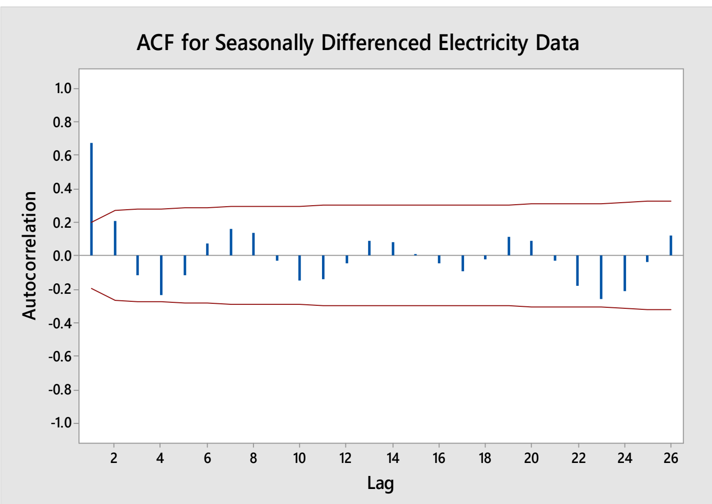
Based on the time series plot the output variable Electricity looks seasonal in nature. We look at the ACF and PACF plots to confirm this.

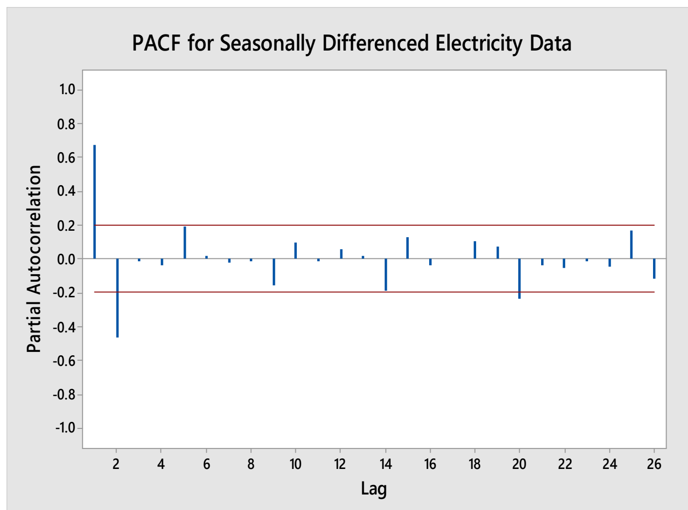
ACF and PACF Plots





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.

$$y_t = \frac{(1-\theta B)\varepsilon_t}{(1-B^{12})}$$

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.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA1,1	1	1	-0.7486430	0.0551879	-13.57	<.0001*

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.

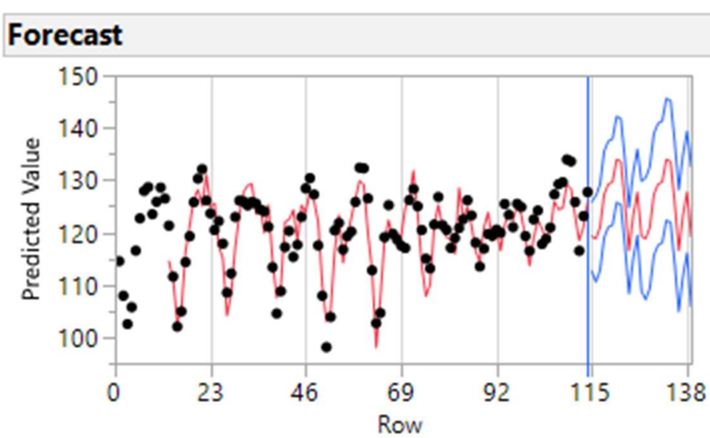
Model Summary

DF	101	Stable	Yes
Sum of Squared Errors	1138.14092	Invertible	Yes
Variance Estimate	11.268722		
Standard Deviation	3.35689172		
Akaike's 'A' Information Criterion	538.327705		
Schwarz's Bayesian Criterion	540.952678		
RSquare	0.77923264		
RSquare Adj	0.77923264		
MAPE	2.25159772		
MAE	2.70848154		
-2LogLikelihood	536.327705		

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

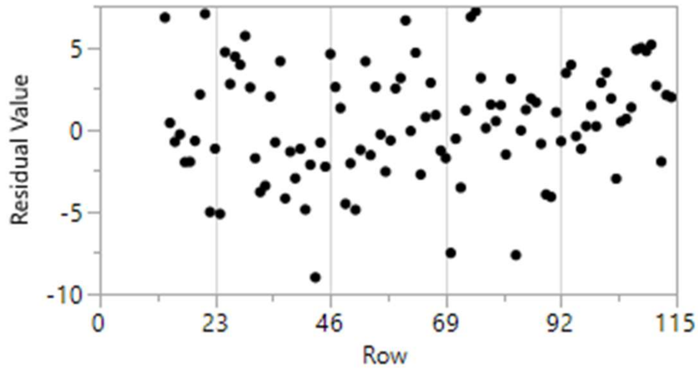
Actual Vs Fit Plot



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

Residual Plots

Residuals



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.2665	7.4571	0.0063*	1	0.2665
2	0.2139	12.3100	0.0021*	2	0.1538
3	-0.1887	16.1248	0.0011*	3	-0.3064
4	-0.1407	18.2690	0.0011*	4	-0.0673
5	-0.1334	20.2165	0.0011*	5	0.0323
6	0.1208	21.8294	0.0013*	6	0.1728
7	0.0809	22.5597	0.0020*	7	-0.0186
8	0.1604	25.4625	0.0013*	8	0.0304
9	-0.0632	25.9178	0.0021*	9	-0.1019
10	-0.1242	27.6972	0.0020*	10	-0.1042
11	-0.0798	28.4390	0.0028*	11	0.1126
12	-0.0879	29.3504	0.0035*	12	-0.0720
13	0.1532	32.1471	0.0023*	13	0.1595
14	0.0119	32.1641	0.0038*	14	-0.1329
15	0.0422	32.3811	0.0057*	15	-0.0448
16	-0.0517	32.7110	0.0081*	16	0.0574
17	-0.0308	32.8298	0.0118*	17	0.0112
18	-0.1156	34.5162	0.0109*	18	-0.0675
19	0.1876	39.0119	0.0044*	19	0.2099
20	-0.0105	39.0261	0.0066*	20	-0.0981
21	0.0687	39.6448	0.0082*	21	-0.1215
22	-0.1960	44.7404	0.0029*	22	-0.1161
23	-0.1157	46.5378	0.0026*	23	0.0329
24	-0.2406	54.4086	0.0004*	24	-0.1463
25	0.0407	54.6364	0.0005*	25	0.1258

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.

$$y_t = \frac{\varepsilon_t}{(1-\phi_1 B-\phi_2 B)(1-B^{12})}$$

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.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	1.022627	0.0871317	11.74	<.0001*
AR1,2	1	2	-0.478420	0.0867583	-5.51	<.0001*

Performance Metrics

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

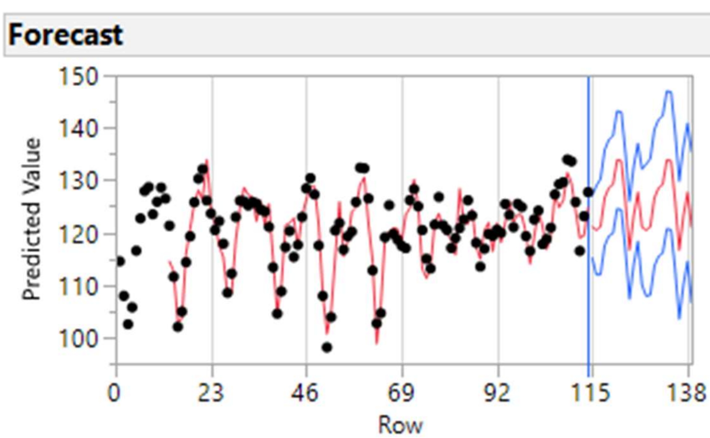
Model Summary

DF	100	Stable	Yes
Sum of Squared Errors	912.477174	Invertible	Yes
Variance Estimate	9.12477174		
Standard Deviation	3.02072371		
Akaike's 'A' Information Criterion	518.13566		
Schwarz's Bayesian Criterion	523.385605		
RSquare	0.82028474		
RSquare Adj	0.81848758		
MAPE	2.02077163		
MAE	2.44404811		
-2LogLikelihood	514.13566		

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

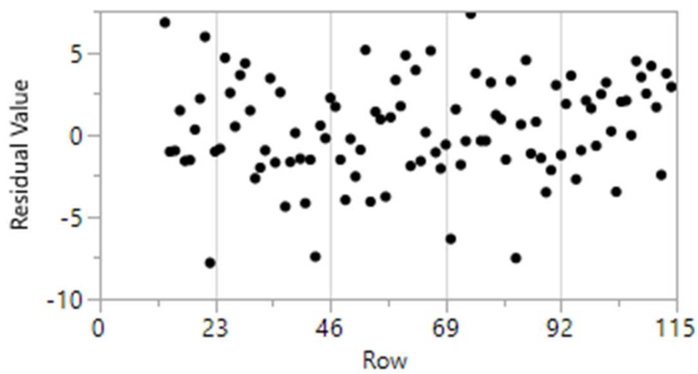
Actual Vs Fit Plot



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

Residual Plots

Residuals



Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	-0.0185		0.0359	0.8498	1	-0.0185	
2	-0.0197		0.0769	0.9623	2	-0.0200	
3	0.0887		0.9197	0.8207	3	0.0880	
4	-0.1125		2.2892	0.6827	4	-0.1106	
5	0.0203		2.3342	0.8012	5	0.0213	
6	0.1102		3.6761	0.7204	6	0.1005	
7	-0.0057		3.6798	0.8158	7	0.0164	
8	0.1561		6.4303	0.5991	8	0.1484	
9	-0.1089		7.7836	0.5561	9	-0.1253	
10	-0.0683		8.3216	0.5975	10	-0.0441	
11	-0.0061		8.3259	0.6838	11	-0.0415	
12	-0.1312		10.3548	0.5849	12	-0.1019	
13	0.2123		15.7258	0.2643	13	0.2103	
14	-0.0196		15.7720	0.3275	14	-0.0735	
15	-0.0554		16.1460	0.3724	15	-0.0027	
16	0.1073		17.5671	0.3498	16	0.0487	
17	-0.0980		18.7658	0.3421	17	-0.0327	
18	-0.1629		22.1162	0.2269	18	-0.1372	
19	0.2185		28.2168	0.0793	19	0.1737	
20	-0.0376		28.3999	0.1003	20	-0.0007	
21	0.0091		28.4107	0.1289	21	-0.0578	
22	-0.0743		29.1431	0.1408	22	-0.1127	
23	-0.0714		29.8268	0.1544	23	-0.0085	
24	-0.1682		33.6742	0.0906	24	-0.2065	
25	-0.0080		33.6830	0.1149	25	0.0644	

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.).