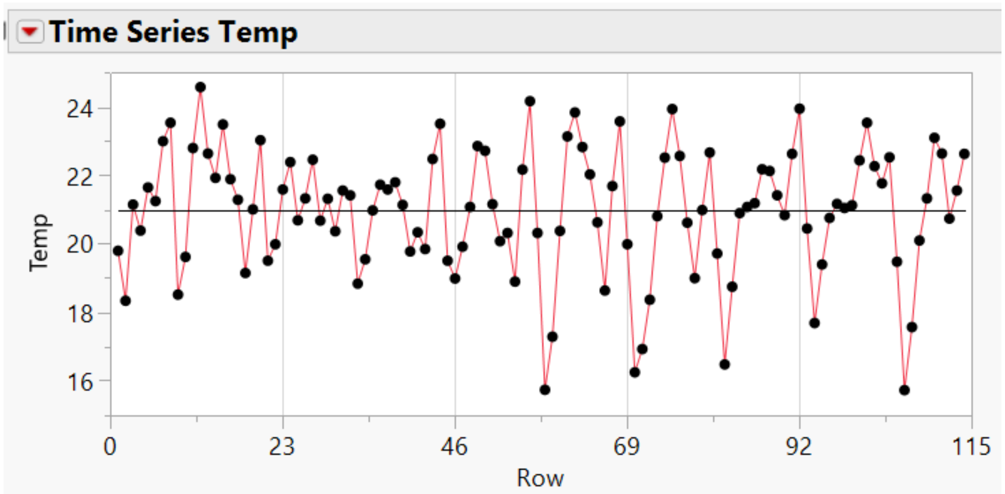
Q1:

The following plot shows the time series plot of the temp data. We can make out from the plot that it is a seasonal data and the period being 12.

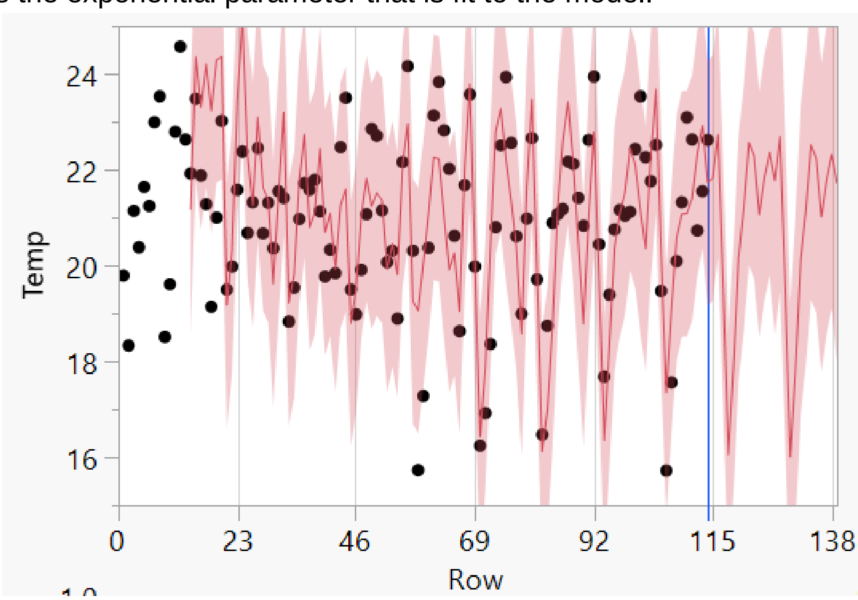


We are using Holt-Winters Additive model because the seasonal pattern remains almost constant in time and remains independent of the average.

The following model was used:

$$y_t = L_t + S_t + \varepsilon_t$$

The following is the exponential parameter that is fit to the model.



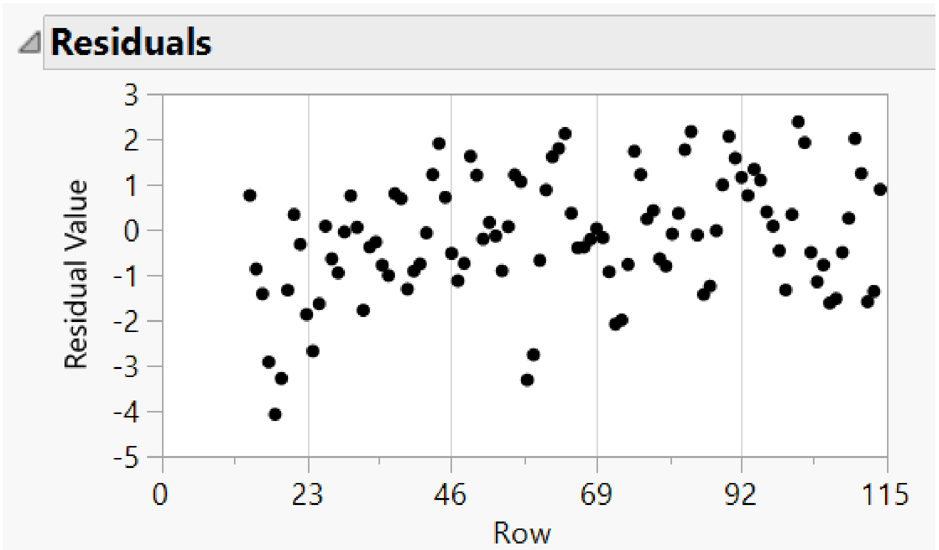
The following are the parameter estimates:

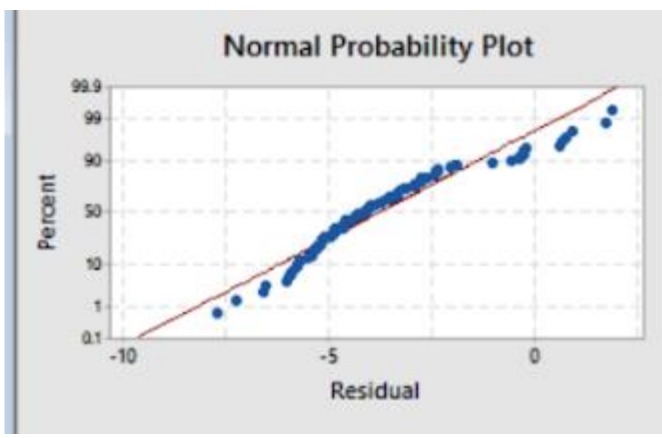
From the parameter estimates we can come to know that the linear and trend are insignificant and can be taken as 0. The seasonal trend has a value of δ=0.7608.

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*

From the above table we can see that only the Seasonal parameter is significant. We check the model for any adequacies





From the residual plot and the normal probability plot for the additive method we do not see unusual pattern. Hence, we can say that the model is close to adequate.

Q2:

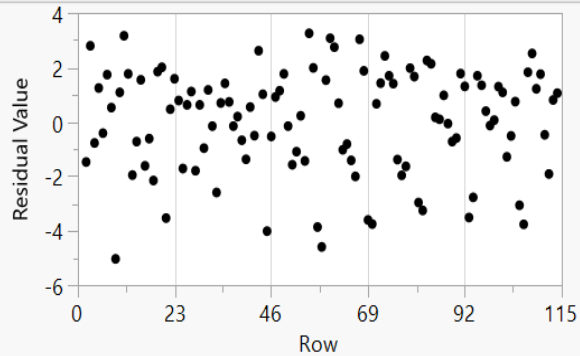
To fit an ARIMA model to the seasonal data we first take the first difference of the data that is: d=1, D=1, d=1 and D=1 together.

The following are the plots and adequacy checks for the three types of differences.

1) d=1

Even though the residuals look normal, the ACF and PACF of d=1 does not capture the seasonal difference. Which shows that just the ARIMA difference will not work for this model.

Residuals

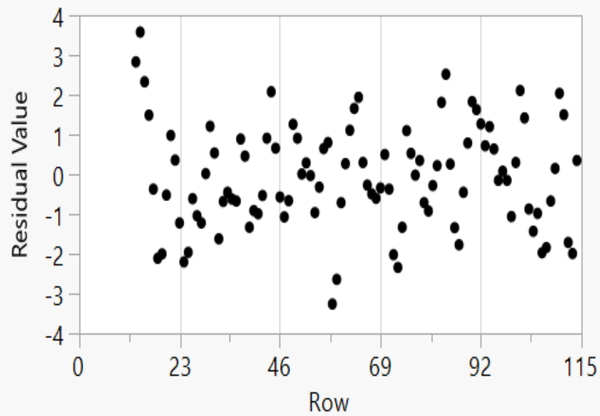


Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	0.1027		1.2247	0.2684	1	0.1027	
2	-0.4906		29.4051	<.0001*	2	-0.5065	
3	-0.2351		35.9327	<.0001*	3	-0.1456	
4	0.0392		36.1155	<.0001*	4	-0.2282	
5	0.0611		36.5652	<.0001*	5	-0.1749	
6	-0.0206		36.6169	<.0001*	6	-0.2133	
7	0.1460		39.2309	<.0001*	7	0.0812	
8	0.0696		39.8301	<.0001*	8	-0.0856	
9	-0.1544		42.8096	<.0001*	9	-0.0790	
10	-0.4326		66.4223	<.0001*	10	-0.5829	
11	0.0083		66.4310	<.0001*	11	-0.2262	
12	0.7001		129.502	<.0001*	12	0.3115	
13	0.1271		131.602	<.0001*	13	-0.2123	
14	-0.3156		144.677	<.0001*	14	0.1075	
15	-0.1912		149.523	<.0001*	15	-0.0641	
16	0.0171		149.562	<.0001*	16	-0.0313	
17	0.0491		149.888	<.0001*	17	0.1251	
18	0.0175		149.931	<.0001*	18	0.0820	
19	0.1098		151.596	<.0001*	19	-0.0177	
20	0.0226		151.668	<.0001*	20	-0.0089	
21	-0.1307		154.080	<.0001*	21	-0.1328	
22	-0.3380		170.393	<.0001*	22	-0.0385	
23	0.0169		170.434	<.0001*	23	-0.0007	
24	0.5190		209.758	<.0001*	24	-0.0294	
25	0.0931		211.039	<.0001*	25	-0.1233	

2) D=1

It can be seen from the below plot that the residuals are not in any pattern. Also, the ACF and PACF are well within limits for the difference part. This is a good candidate for the ARIMA model.

Residuals

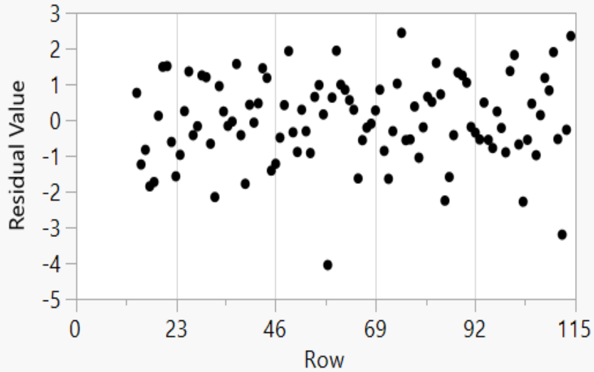


Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	0.5696		34.0795	<.0001*	1	0.5696	
2	-0.0279		34.1620	<.0001*	2	-0.5216	
3	-0.2638		41.6212	<.0001*	3	0.1168	
4	-0.2588		48.8710	<.0001*	4	-0.2020	
5	-0.1389		50.9823	<.0001*	5	0.0714	
6	0.0655		51.4570	<.0001*	6	0.1005	
7	0.2107		56.4132	<.0001*	7	0.0417	
8	0.1462		58.8260	<.0001*	8	-0.0750	
9	-0.0405		59.0131	<.0001*	9	-0.0495	
10	-0.2025		63.7406	<.0001*	10	-0.1238	
11	-0.2261		69.7001	<.0001*	11	0.0034	
12	-0.1028		70.9453	<.0001*	12	0.0065	
13	0.0635		71.4266	<.0001*	13	0.0240	
14	0.0738		72.0838	<.0001*	14	-0.1730	
15	-0.0193		72.1294	<.0001*	15	0.0140	
16	-0.0747		72.8186	<.0001*	16	-0.0378	
17	-0.0742		73.5054	<.0001*	17	0.0226	
18	-0.0105		73.5194	<.0001*	18	0.0663	
19	0.0686		74.1205	<.0001*	19	-0.0177	
20	0.0616		74.6110	<.0001*	20	-0.0946	
21	-0.0302		74.7307	<.0001*	21	-0.0451	
22	-0.1029		76.1351	<.0001*	22	-0.0368	
23	-0.1654		79.8104	<.0001*	23	-0.1430	
24	-0.1931		84.8797	<.0001*	24	-0.0809	
25	-0.1284		87.1514	<.0001*	25	-0.0914	

3) d=1 and D=1 together

Even though this model has ACF and PACF within limits. The initial lags have may insignificant values. An ARIMA model can be fit into this difference but the model will be complex to identify.

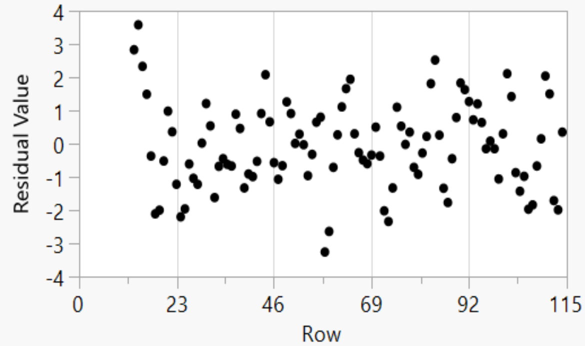
Residuals



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.1965	4.0182	0.0450*	1	0.1965
2	-0.4216	22.6934	<.0001*	2	-0.4787
3	-0.2893	31.5763	<.0001*	3	-0.0981
4	-0.1045	32.7475	<.0001*	4	-0.2787
5	-0.0627	33.1731	<.0001*	5	-0.2431
6	0.0735	33.7652	<.0001*	6	-0.1175
7	0.2311	39.6780	<.0001*	7	0.0112
8	0.1197	41.2818	<.0001*	8	-0.0343
9	-0.0209	41.3313	<.0001*	9	0.0802
10	-0.1369	43.4730	<.0001*	10	-0.0625
11	-0.1625	46.5240	<.0001*	11	-0.0495
12	-0.0641	47.0048	<.0001*	12	-0.0623
13	0.1630	50.1461	<.0001*	13	0.1095
14	0.1419	52.5529	<.0001*	14	-0.0395
15	-0.0389	52.7360	<.0001*	15	-0.0246
16	-0.0972	53.8933	<.0001*	16	-0.0688
17	-0.1032	55.2121	<.0001*	17	-0.1252
18	-0.0048	55.2150	<.0001*	18	0.0003
19	0.1485	58.0128	<.0001*	19	0.1015
20	0.0838	58.9151	<.0001*	20	-0.0423
21	-0.0268	59.0083	<.0001*	21	0.0534
22	-0.0104	59.0226	<.0001*	22	0.0573
23	-0.0428	59.2663	<.0001*	23	0.0107
24	-0.1463	62.1590	<.0001*	24	-0.0617
25	-0.1665	65.9519	<.0001*	25	-0.1924

Hence, we go ahead with using only the **seasonal difference**. The ACF and PACF shown below is of seasonal difference 1 (D=1).

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	0.5696		34.0795	<.0001*	1	0.5696	
2	-0.0279		34.1620	<.0001*	2	-0.5216	
3	-0.2638		41.6212	<.0001*	3	0.1168	
4	-0.2588		48.8710	<.0001*	4	-0.2020	
5	-0.1389		50.9823	<.0001*	5	0.0714	
6	0.0655		51.4570	<.0001*	6	0.1005	
7	0.2107		56.4132	<.0001*	7	0.0417	
8	0.1462		58.8260	<.0001*	8	-0.0750	
9	-0.0405		59.0131	<.0001*	9	-0.0495	
10	-0.2025		63.7406	<.0001*	10	-0.1238	
11	-0.2261		69.7001	<.0001*	11	0.0034	
12	-0.1028		70.9453	<.0001*	12	0.0065	
13	0.0635		71.4266	<.0001*	13	0.0240	
14	0.0738		72.0838	<.0001*	14	-0.1730	
15	-0.0193		72.1294	<.0001*	15	0.0140	
16	-0.0747		72.8186	<.0001*	16	-0.0378	
17	-0.0742		73.5054	<.0001*	17	0.0226	
18	-0.0105		73.5194	<.0001*	18	0.0663	
19	0.0686		74.1205	<.0001*	19	-0.0177	
20	0.0616		74.6110	<.0001*	20	-0.0946	
21	-0.0302		74.7307	<.0001*	21	-0.0451	
22	-0.1029		76.1351	<.0001*	22	-0.0368	
23	-0.1654		79.8104	<.0001*	23	-0.1430	
24	-0.1931		84.8797	<.0001*	24	-0.0809	
25	-0.1284		87.1514	<.0001*	25	-0.0914	

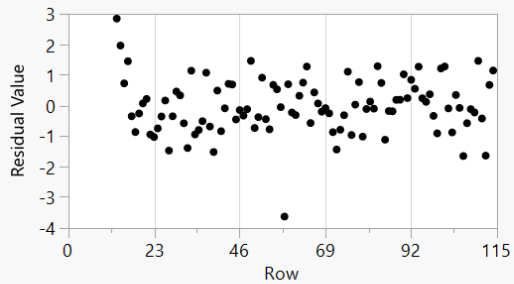
An initial guess can be made an AR2 term is required and an Seasonal MA1 can be fit along with D=1

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.8996852	0.0828486	10.86	<.0001*
AR1,2	1	2	-0.5754422	0.0858781	-6.70	<.0001*
MA2,12	2	12	0.1994701	0.1095395	1.82	0.0716

Forecast

Residuals



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.0979	1.0075	0.3155	1	0.0979
2	-0.1166	2.4497	0.2938	2	-0.1274
3	0.1268	4.1733	0.2433	3	0.1562
4	0.0722	4.7376	0.3153	4	0.0247
5	-0.0630	5.1712	0.3953	5	-0.0408
6	-0.0900	6.0670	0.4157	6	-0.0880
7	0.0974	7.1273	0.4158	7	0.0973
8	-0.0254	7.2001	0.5152	8	-0.0636
9	-0.0754	7.8487	0.5495	9	-0.0125
10	-0.0879	8.7391	0.5570	10	-0.1161
11	-0.0669	9.2615	0.5978	11	-0.0600
12	0.0063	9.2662	0.6800	12	0.0143
13	0.0218	9.3228	0.7482	13	0.0524
14	0.0162	9.3544	0.8077	14	0.0152
15	-0.0264	9.4393	0.8535	15	-0.0296
16	-0.0032	9.4405	0.8942	16	-0.0210
17	-0.0479	9.7268	0.9147	17	-0.0604
18	-0.0581	10.1524	0.9268	18	-0.0339
19	0.0842	11.0590	0.9218	19	0.0868
20	0.0051	11.0623	0.9446	20	-0.0341
21	-0.1165	12.8393	0.9141	21	-0.1055
22	-0.0513	13.1884	0.9278	22	-0.0527
23	-0.0064	13.1939	0.9477	23	-0.0287
24	-0.0134	13.2182	0.9624	24	0.0283
25	-0.2661	22.9721	0.5792	25	-0.2632

Although the residuals are pretty much in limits the Seasonal MA is not significant hence, we do not use this model.

An optimal ARIMA was found using R software. The output was $(2,0,1)x(0,1,1)_{12}$. This model has parameters that are close to the initial guess. Hence, we check the parameter estimates to check to if all parameters are significant.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.6383751	0.1631358	3.91	0.0002*
AR1,2	1	2	-0.4142309	0.1365723	-3.03	0.0031*
MA1,1	1	1	-0.3819288	0.1769813	-2.16	0.0334*
MA2,12	2	12	0.1795443	0.1043152	1.72	0.0884

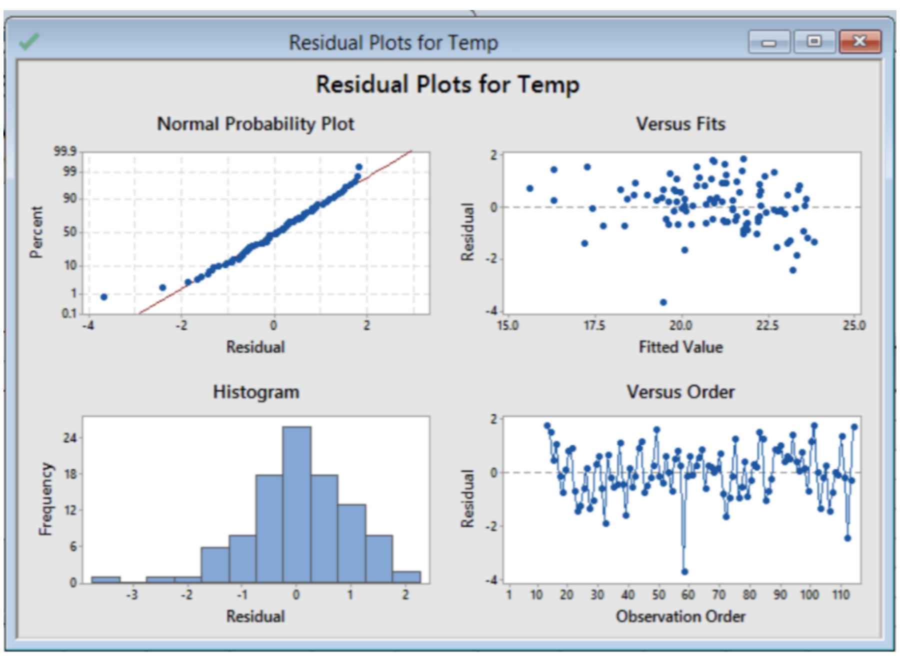
From the above estimates we can see that the final term is not significant. Hence, we do not use this model.

So, we fit another seasonal ARIMA of the order **$(2,0,1)\times(0,1,0)_{12}$**

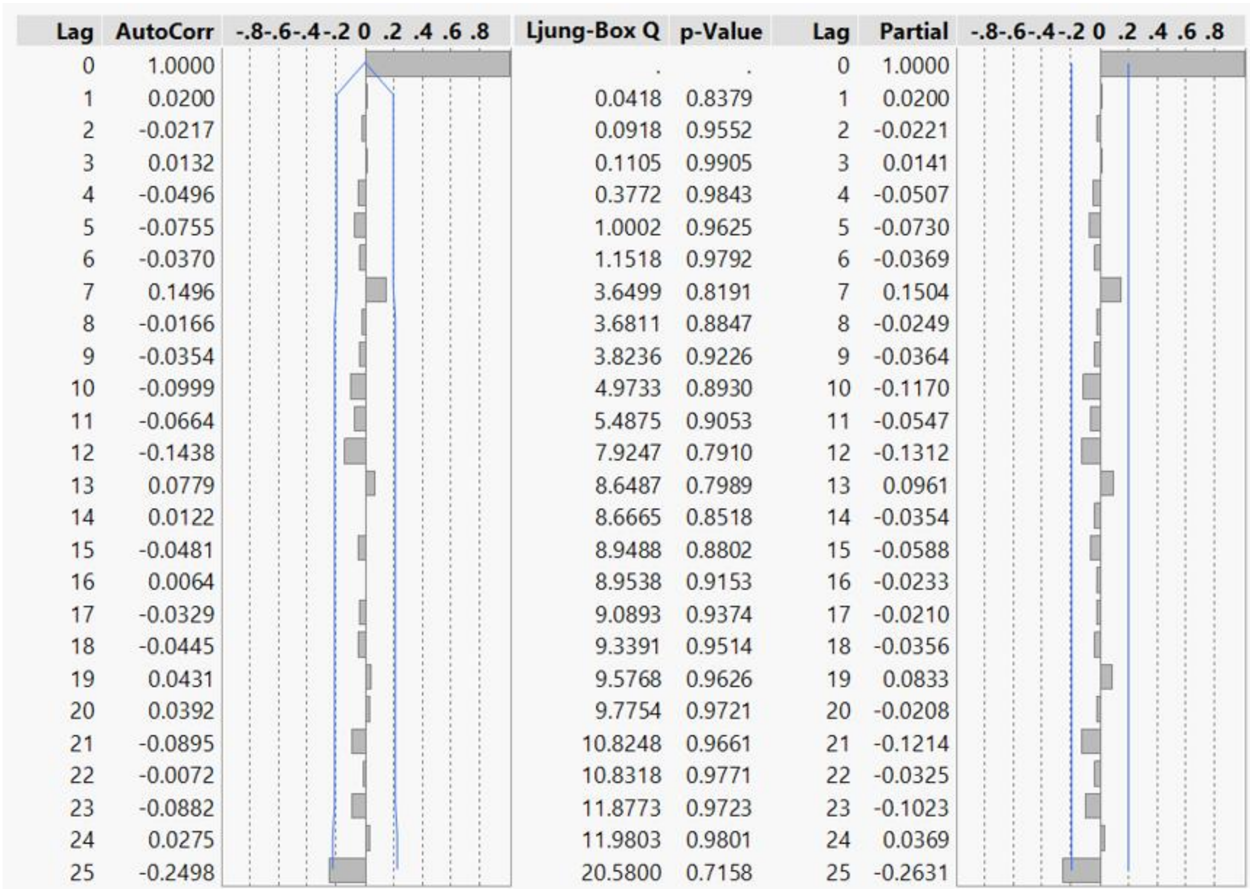
Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.6246971	0.1625715	3.84	0.0002*
AR1,2	1	2	-0.3922484	0.1379396	-2.84	0.0054*
MA1,1	1	1	-0.4022916	0.1713703	-2.35	0.0209*

We can see that all the parameters that were fit into the model are significant. From the residual plot and the normal probability plot we do not see any unusual pattern.



Also, from the ACF and PACF we can see that the values are within limits. Hence, we use this model as the optimal ARIMA.



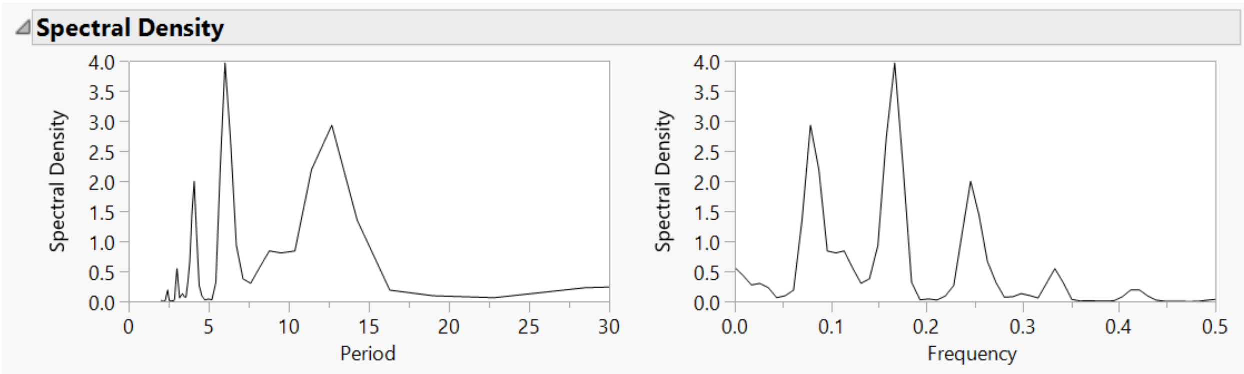
Lag	AutoCorr	-.8-.6-.4-.2 0 .2 .4 .6 .8	Ljung-Box Q	p-Value	Lag	Partial	-.8-.6-.4-.2 0 .2 .4 .6 .8
0	1.0000		.	.	0	1.0000	
1	0.0200		0.0418	0.8379	1	0.0200	
2	-0.0217		0.0918	0.9552	2	-0.0221	
3	0.0132		0.1105	0.9905	3	0.0141	
4	-0.0496		0.3772	0.9843	4	-0.0507	
5	-0.0755		1.0002	0.9625	5	-0.0730	
6	-0.0370		1.1518	0.9792	6	-0.0369	
7	0.1496		3.6499	0.8191	7	0.1504	
8	-0.0166		3.6811	0.8847	8	-0.0249	
9	-0.0354		3.8236	0.9226	9	-0.0364	
10	-0.0999		4.9733	0.8930	10	-0.1170	
11	-0.0664		5.4875	0.9053	11	-0.0547	
12	-0.1438		7.9247	0.7910	12	-0.1312	
13	0.0779		8.6487	0.7989	13	0.0961	
14	0.0122		8.6665	0.8518	14	-0.0354	
15	-0.0481		8.9488	0.8802	15	-0.0588	
16	0.0064		8.9538	0.9153	16	-0.0233	
17	-0.0329		9.0893	0.9374	17	-0.0210	
18	-0.0445		9.3391	0.9514	18	-0.0356	
19	0.0431		9.5768	0.9626	19	0.0833	
20	0.0392		9.7754	0.9721	20	-0.0208	
21	-0.0895		10.8248	0.9661	21	-0.1214	
22	-0.0072		10.8318	0.9771	22	-0.0325	
23	-0.0882		11.8773	0.9723	23	-0.1023	
24	0.0275		11.9803	0.9801	24	0.0369	
25	-0.2498		20.5800	0.7158	25	-0.2631	

When we compare the statistics between Holt-Winters model and Optimal ARIMA we get the following results.

We can clearly see that the ARIMA is a better fit to the temp variable.

Parameter	ARIMA	WINTERS
MAD	0.70	1.056
MAPE	3.43	5.09
MSE	0.81	181.40

Q3:



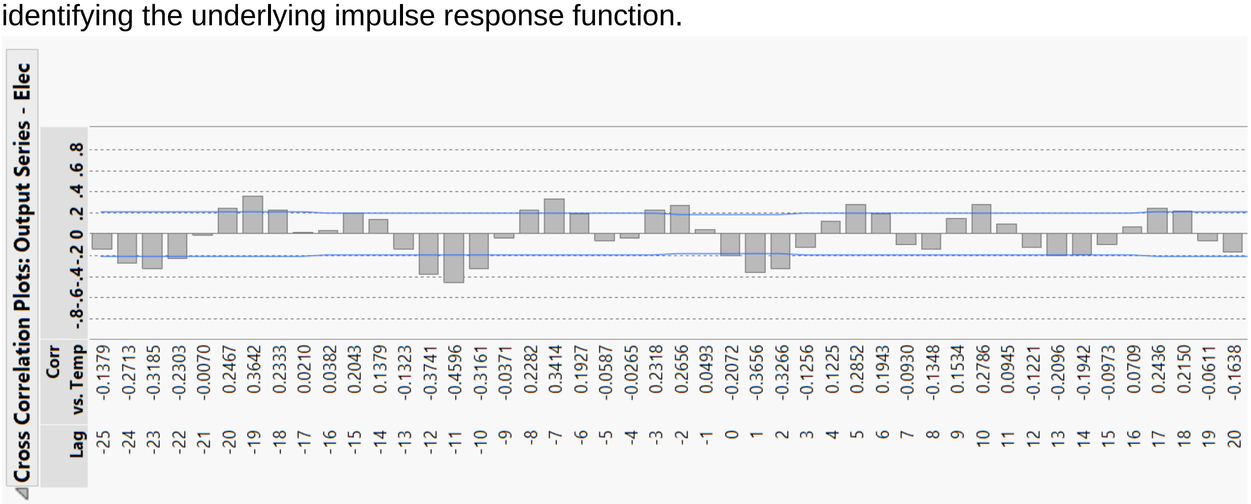
White Noise test	
Fisher's Kappa	11.752839
Prob > Kappa	0.0001323
Bartlett's Kolmogorov-Smirnov	0.3899686

We see that the 1st spike is at frequency of 0.08 ~ 12 months, 2nd is at frequency of 0.16 ~ 4 months and so on. Thus, there is periodicity in the data with period = 12 months.

We can also look at the kappa value which is significant. Meaning that there is at least one periodic component.

Q4:

We begin by looking at the cross correlation of electricity and temperature. It shows there is autocorrelation between the two variables. A process called pre-whitening must be done to remove the auto correlation and introduce a noise term in the model. It will also help us in identifying the underlying impulse response function.

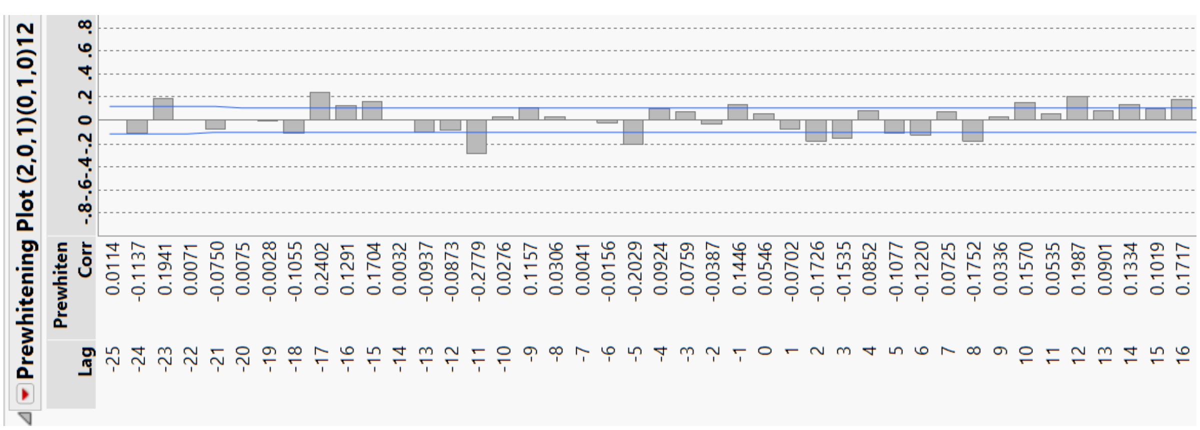


We use the ARIMA model got from question 2. The ARIMA model (2,0,1)x(0,1,0)12 for the temp data to find out the underlying impulse response function.

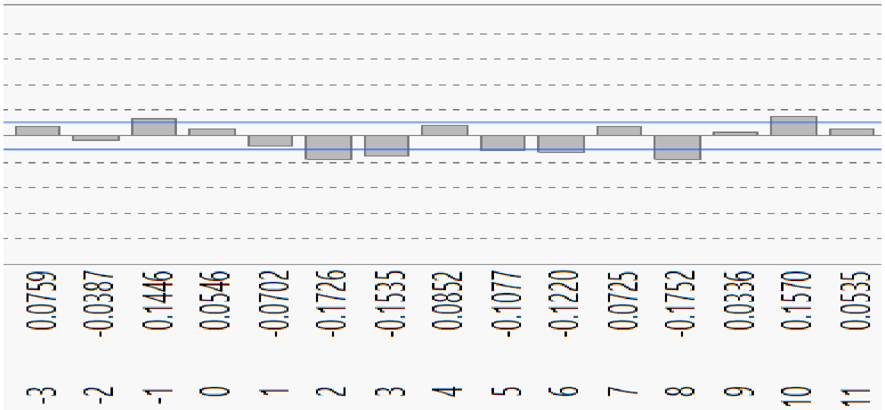
The following is the ACF and PACF of temp using the above ARIMA model.

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	0.0200		0.0418	0.8379	1	0.0200	
2	-0.0217		0.0918	0.9552	2	-0.0221	
3	0.0132		0.1105	0.9905	3	0.0141	
4	-0.0496		0.3772	0.9843	4	-0.0507	
5	-0.0755		1.0002	0.9625	5	-0.0730	
6	-0.0370		1.1518	0.9792	6	-0.0369	
7	0.1496		3.6499	0.8191	7	0.1504	
8	-0.0166		3.6811	0.8847	8	-0.0249	
9	-0.0354		3.8236	0.9226	9	-0.0364	
10	-0.0999		4.9733	0.8930	10	-0.1170	
11	-0.0664		5.4875	0.9053	11	-0.0547	
12	-0.1438		7.9247	0.7910	12	-0.1312	
13	0.0779		8.6487	0.7989	13	0.0961	
14	0.0122		8.6665	0.8518	14	-0.0354	
15	-0.0481		8.9488	0.8802	15	-0.0588	
16	0.0064		8.9538	0.9153	16	-0.0233	
17	-0.0329		9.0893	0.9374	17	-0.0210	
18	-0.0445		9.3391	0.9514	18	-0.0356	
19	0.0431		9.5768	0.9626	19	0.0833	
20	0.0392		9.7754	0.9721	20	-0.0208	
21	-0.0895		10.8248	0.9661	21	-0.1214	
22	-0.0072		10.8318	0.9771	22	-0.0325	
23	-0.0882		11.8773	0.9723	23	-0.1023	
24	0.0275		11.9803	0.9801	24	0.0369	
25	-0.2498		20.5800	0.7158	25	-0.2631	

Hence, we use the above model to pre-white the temp variable.



Zoomed in version of the impulse response function from lag 0.



From the above plot we can conclude the following models for the impulse response.

1) b = 2, r = 1, s = 1

2) b = 2, r = 1, s = 2

We compare all the two above functions and look for parameters that can help us decide which impulse response function is a better performing one.

1) b = 2, r = 1, s = 1

Model Summary

DF	107
Sum of Squared Errors	4843.74737
Variance Estimate	45.2686664
Standard Deviation	6.72819934
Akaike's 'A' Information Criterion	742.130779
Schwarz's Bayesian Criterion	752.9689
RSquare	0.22439723
RSquare Adj	0.20324443
MAPE	4.07791606
MAE	4.79979714
-2LogLikelihood	734.13078

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-1.3865	0.328937	-4.22	<.0001*
Temp	Num1,1	1	1	-1.4306	0.318899	-4.49	<.0001*
Temp	Den1,1	1	1	0.7532	0.120491	6.25	<.0001*
	Intercept	0	0	117.0371	6.397970	18.29	<.0001*

$$Elec_t = 117.0371 + \left(\frac{\left(-1.3865 + 1.4306 \bullet B \right)}{\left(1 - 0.7532 \bullet B \right)} \right) \bullet Temp_{t-2} + e_t$$

2) b = 2, r = 1, s = 2

Model Summary

DF	106
Sum of Squared Errors	10744.9801
Variance Estimate	101.367737
Standard Deviation	10.0681546
Akaike's 'A' Information Criterion	824.154972
Schwarz's Bayesian Criterion	834.956894
RSquare	-0.7205349
RSquare Adj	-0.7674586
MAPE	5.9832423
MAE	7.10472166
-2LogLikelihood	816.154973

Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	2.125054	0.5048264	4.21	<.0001*
Temp	Num1,1	1	1	1.559626	0.7751758	2.01	0.0468*
Temp	Num1,2	1	2	-3.768415	0.5132162	-7.34	<.0001*
Temp	Den1,1	1	1	0.247684	0.0638740	3.88	0.0002*

$$Elec_t = \left(\frac{\left(\left(2.1251 - 1.5596 \cdot B \right) + 3.7684 \cdot B^2 \right)}{\left(1 - 0.2477 \cdot B \right)} \right) \cdot Temp_{t-2} + e_t$$

From the above two functions we can see Model (1) b = 2, r = 1, s = 1 is the best one. Also, model (2) has close parameters we will try fitting a model in those ones also to see if we are able to get an adequate transfer function model.

Using (1) b = 2, r = 1, s = 1 we fit the noise term.

The ACF and PACF of the response function is shown below:

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	0.6097		43.5050	<.0001*	1	0.6097	
2	0.1198		45.2004	<.0001*	2	-0.4010	
3	-0.0445		45.4368	<.0001*	3	0.1851	
4	-0.0720		46.0602	<.0001*	4	-0.1503	
5	-0.2389		52.9832	<.0001*	5	-0.2932	
6	-0.3961		72.1987	<.0001*	6	-0.1065	
7	-0.2567		80.3445	<.0001*	7	0.1493	
8	-0.1084		81.8105	<.0001*	8	-0.2476	
9	-0.1096		83.3222	<.0001*	9	-0.0242	
10	-0.0188		83.3672	<.0001*	10	0.2062	
11	0.3593		99.9421	<.0001*	11	0.4147	
12	0.6208		149.902	<.0001*	12	0.0889	
13	0.3737		168.186	<.0001*	13	-0.2451	
14	0.0353		168.350	<.0001*	14	0.0270	
15	-0.0410		168.574	<.0001*	15	0.0328	
16	0.0007		168.575	<.0001*	16	0.0849	
17	-0.1425		171.345	<.0001*	17	-0.0415	
18	-0.2958		183.401	<.0001*	18	-0.0147	
19	-0.2030		189.138	<.0001*	19	0.0049	
20	-0.0851		190.158	<.0001*	20	-0.0046	
21	-0.0980		191.524	<.0001*	21	0.0285	
22	-0.0565		191.983	<.0001*	22	-0.0068	
23	0.2173		198.844	<.0001*	23	-0.0045	
24	0.4176		224.465	<.0001*	24	0.0348	
25	0.2651		234.908	<.0001*	25	0.0356	

Looking at the above pattern we can try:

Seasonal ARIMA $(2,1,1)x(1,1,0)_{12}$

We can see that all the parameters are significant. The intercept term can be ignored as it is insignificant.

Model Summary

DF	90
Sum of Squared Errors	780.494935
Variance Estimate	8.67211342
Standard Deviation	2.94484523
Akaike's 'A' Information Criterion	503.120047
Schwarz's Bayesian Criterion	523.799787
RSquare	0.43730574
RSquare Adj	0.39495241
MAPE	1.61009529
MAE	1.93621525
-2LogLikelihood	487.120047

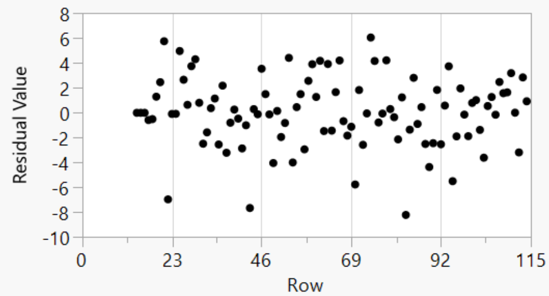
Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.135726	0.0134464	-10.09	<.0001*
Temp	Num1,1	1	1	-0.136886	0.0135261	-10.12	<.0001*
Temp	Den1,1	1	1	0.952139	0.0995505	9.56	<.0001*
Elec	AR1,1	1	1	1.002047	0.0848539	11.81	<.0001*
Elec	AR1,2	1	2	-0.526135	0.0855363	-6.15	<.0001*
Elec	AR2,12	2	12	-0.244629	0.1078992	-2.27	0.0258*
Elec	MA1,1	1	1	1.000000	0.0314689	31.78	<.0001*
	Intercept	0	0	-0.424804	0.6036267	-0.70	0.4834

$$(1-B)\cdot\left(1-B^{12}\right)\cdot Elec_t = -0.4248 + \left(\frac{\left(-0.1357+0.1369\cdot B\right)}{\left(1-0.9521\cdot B\right)}\right)\cdot Temp_{t-2} + \left(\frac{\left(1-B\right)}{\left(\left(\left(1-1.002\cdot B\right)+0.5261\cdot B^{2}\right)\cdot\left(1+0.2446\cdot B^{12}\right)\right)}\right)\cdot e_t$$

Model Adequacy checks:
If we look at the residual plot, we do not see any specific pattern. Also, the ACF and PACF plots are well within limits. We can conclude that the model is adequate.

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	0.0026		0.0007	0.9786	1	0.0026	
2	-0.0202		0.0436	0.9784	2	-0.0202	
3	0.0674		0.5258	0.9132	3	0.0675	
4	-0.0668		1.0040	0.9092	4	-0.0680	
5	-0.0377		1.1578	0.9489	5	-0.0344	
6	0.0837		1.9249	0.9265	6	0.0777	
7	-0.0044		1.9271	0.9638	7	0.0020	
8	0.0779		2.6066	0.9566	8	0.0824	
9	-0.0866		3.4544	0.9435	9	-0.1057	
10	-0.0959		4.5058	0.9217	10	-0.0839	
11	-0.0279		4.5959	0.9491	11	-0.0355	
12	-0.0349		4.7381	0.9662	12	-0.0218	
13	0.1725		8.2557	0.8266	13	0.1875	
14	-0.0755		8.9366	0.8351	14	-0.1155	
15	-0.0623		9.4068	0.8553	15	-0.0490	
16	0.0411		9.6137	0.8860	16	0.0161	
17	-0.1214		11.4382	0.8329	17	-0.0808	
18	-0.1968		16.2922	0.5722	18	-0.1769	
19	0.2273		22.8461	0.2442	19	0.1918	
20	-0.0279		22.9458	0.2915	20	-0.0256	
21	-0.0528		23.3083	0.3278	21	-0.0762	
22	-0.1145		25.0358	0.2954	22	-0.1523	
23	-0.1055		26.5212	0.2769	23	-0.0600	
24	-0.2556		35.3460	0.0634	24	-0.2733	
25	-0.0397		35.5612	0.0785	25	-0.0494	

(2) b = 2, r = 1, s = 2 we try to fit an ARIMA model:
We need to fit a noise model to the following model.

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.6693	52.4304	<.0001*	1	0.6693
2	0.3387	65.9770	<.0001*	2	-0.1980
3	0.1391	68.2807	<.0001*	3	-0.0000
4	-0.1546	71.1536	<.0001*	4	-0.3752
5	-0.3295	84.3216	<.0001*	5	-0.0403
6	-0.3817	102.163	<.0001*	6	-0.1317
7	-0.3737	119.421	<.0001*	7	-0.0137
8	-0.2845	129.515	<.0001*	8	-0.0417
9	-0.1504	132.363	<.0001*	9	0.0072
10	-0.0783	133.142	<.0001*	10	-0.1423
11	0.1276	135.233	<.0001*	11	0.2986
12	0.4137	157.421	<.0001*	12	0.2874
13	0.3080	169.839	<.0001*	13	-0.4246
14	0.1960	174.921	<.0001*	14	0.0780
15	0.1793	179.214	<.0001*	15	0.0490
16	0.0186	179.261	<.0001*	16	0.0473
17	-0.1566	182.606	<.0001*	17	-0.2136
18	-0.2703	192.668	<.0001*	18	-0.0936
19	-0.3481	209.539	<.0001*	19	-0.0943
20	-0.3218	224.108	<.0001*	20	0.0885
21	-0.2103	230.397	<.0001*	21	0.0763
22	-0.1173	232.376	<.0001*	22	0.0476
23	0.1093	234.110	<.0001*	23	0.0967
24	0.3935	256.866	<.0001*	24	0.0379
25	0.3411	274.155	<.0001*	25	-0.0914

We try to fit $(2,1,1)x(0,1,0)_{12}$ as a noise model.

Model Summary

Sum of Squared Errors	829.587471
Variance Estimate	9.32115109
Standard Deviation	3.05305603
Akaike's 'A' Information Criterion	503.788569
Schwarz's Bayesian Criterion	524.386257
RSquare	0.4019127
RSquare Adj	0.35689538
MAPE	2.45014554
MAE	2.88272216
-2LogLikelihood	487.788569

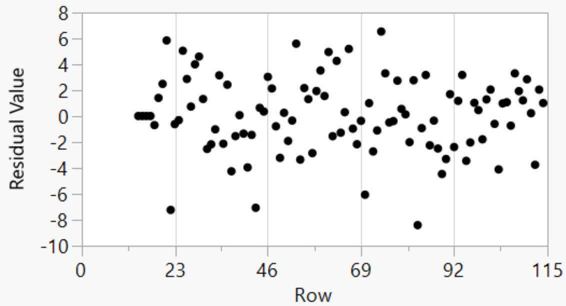
Parameter Estimates

Variable	Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
Temp	Num0,0	0	0	-0.170920	0.0137270	-12.45	<.0001*
Temp	Num1,1	1	1	-0.176977	0.0132907	-13.32	<.0001*
Temp	Num1,2	1	2	0.004910	0.0201009	0.24	0.8076
Temp	Den1,1	1	1	0.968247	0.1369218	7.07	<.0001*
Elec	AR1,1	1	1	0.981367	0.0878024	11.18	<.0001*
Elec	AR1,2	1	2	-0.486490	0.0880500	-5.53	<.0001*
Elec	MA1,1	1	1	1.000000	0.0298910	33.45	<.0001*
	Intercept	0	0	-0.647346	0.6887430	-0.94	0.3498

$$(1-B)\cdot(1-B^{12})\cdot Elec_t = -0.6473 + \left(\frac{\left(\left(-0.1709 + 0.177\cdot B\right) - 0.0049\cdot B^2\right)}{\left(1 - 0.9682\cdot B\right)}\right)\cdot Temp_{t-2} + \left(\frac{(1-B)}{\left(\left(1 - 0.9814\cdot B\right) + 0.4865\cdot B^2\right)}\right)\cdot e_t$$

The following plot shows that the residuals do not have any pattern also the ACF and PACF are in limits there is an insignificant term in the model.

Residuals



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	-0.0133	0.0184	0.8922	1	-0.0133
2	0.0018	0.0187	0.9907	2	0.0017
3	0.0792	0.6854	0.8766	3	0.0793
4	-0.1106	1.9970	0.7363	4	-0.1092
5	-0.0001	1.9970	0.8496	5	-0.0025
6	0.1049	3.2016	0.7832	6	0.1011
7	-0.0425	3.4013	0.8456	7	-0.0249
8	0.1225	5.0792	0.7491	8	0.1118
9	-0.0782	5.7707	0.7626	9	-0.0965
10	-0.0730	6.3796	0.7824	10	-0.0491
11	0.0077	6.3865	0.8464	11	-0.0149
12	-0.1959	10.8741	0.5397	12	-0.1791
13	0.1640	14.0553	0.3700	13	0.1813
14	-0.0709	14.6557	0.4021	14	-0.1299
15	-0.0993	15.8480	0.3922	15	-0.0449
16	0.0757	16.5497	0.4153	16	0.0226
17	-0.1272	18.5520	0.3548	17	-0.1012
18	-0.1913	23.1402	0.1853	18	-0.1424
19	0.2370	30.2685	0.0485*	19	0.1905
20	-0.0669	30.8433	0.0573	20	-0.0059
21	-0.0354	31.0065	0.0736	21	-0.0996
22	-0.0892	32.0549	0.0765	22	-0.1523
23	-0.1373	34.5684	0.0574	23	-0.0664
24	-0.1846	39.1722	0.0262*	24	-0.2526
25	-0.0257	39.2624	0.0347*	25	0.0205

We compare the AIC BIC of the of both the models and select which is a better model.

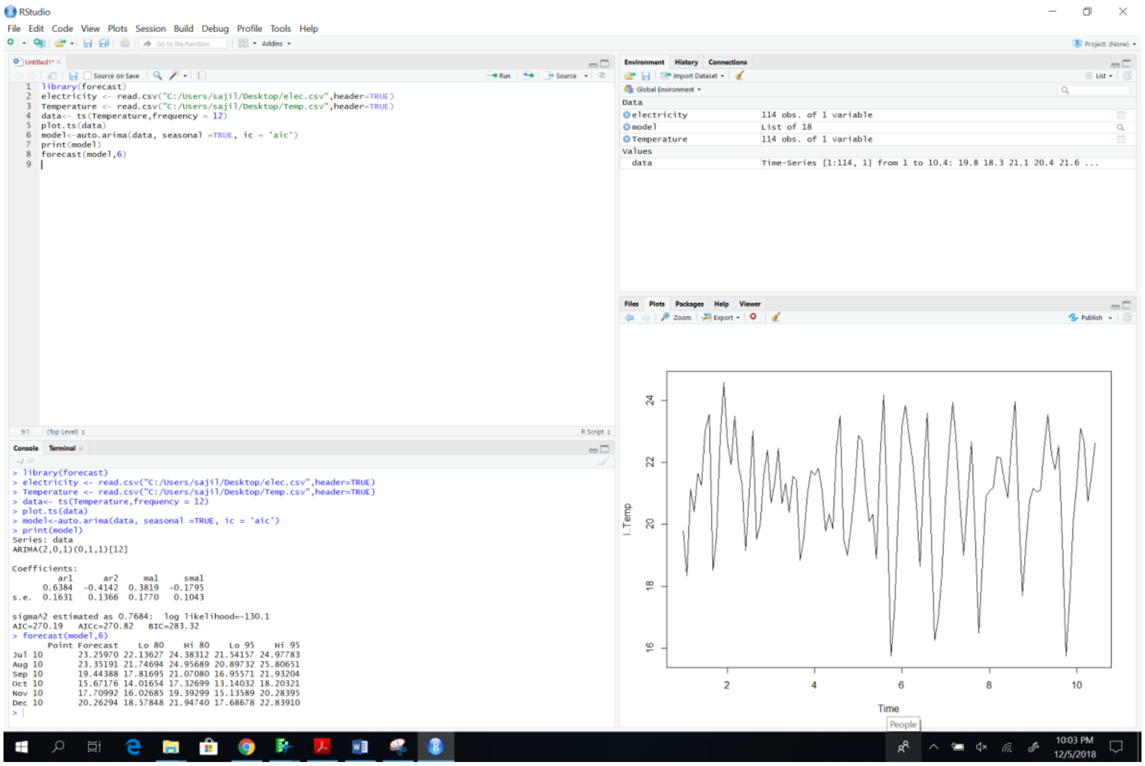
No.	Model (b,r,s) (p,d,q)(P,D,Q)$_{12}$	AIC	BIC	MAPE	MAE
1	(2,1,1) (2,1,1)(1,1,0)$_{12}$	503.12	523.8	1.61	1.94
2	(2,0,2) (2,1,1)(0,1,0)$_{12}$	503.79	524.38	2.45	2.88

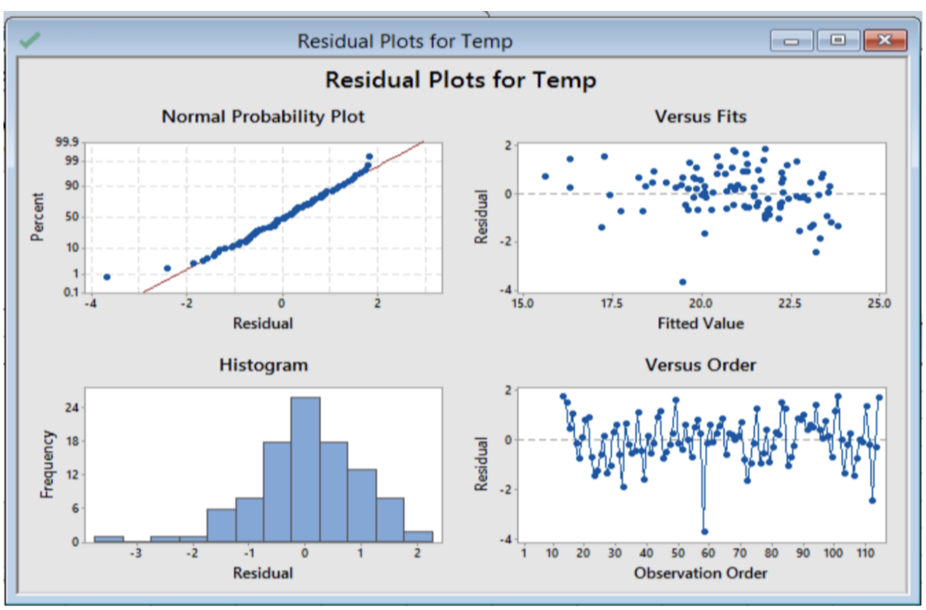
From the above comparisons we can see that model one has better statistics. Hence, we choose model 1 as the transfer function model.

Q5:

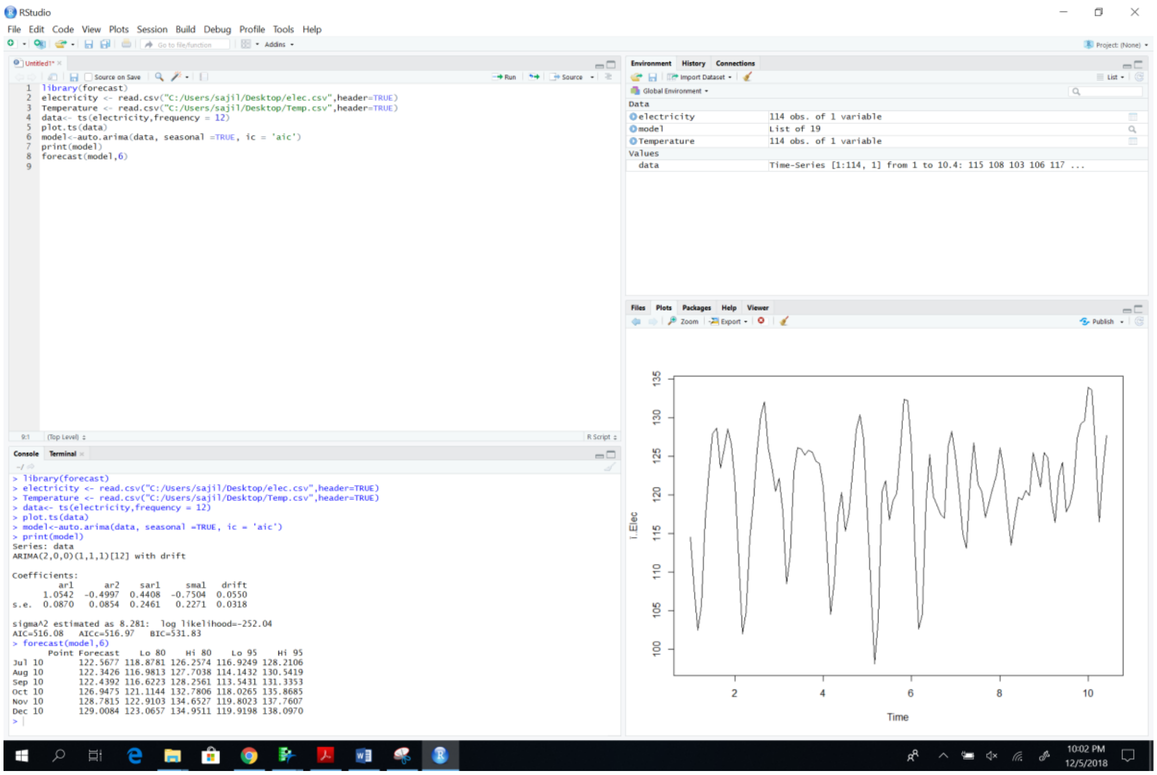
Optimal ARIMA from R software has been used to find the forecast values of the next 6 data points.

Optimal ARIMA for temp: The optimal ARIMA for temperature is $(2,0,1)(0,1,1)_{12}$. The model seems to be adequate, which can be seen from the measures given below.





Optimal ARIMA for elec: The optimal ARIMA for electricity is $(2,0,0)(1,1,1)_{12}$. The model seems to be adequate, which can be seen from the measures given below.



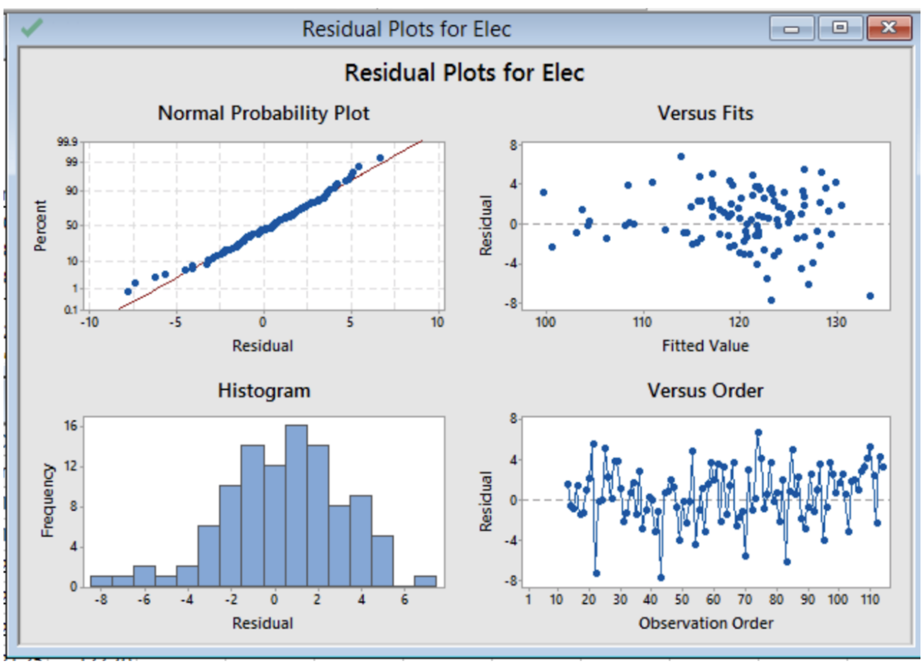


Table-

Forecast for temperature: $(2,0,1)(0,1,1)_{12}$

Temp	Forecast
Jul 10	23.2597
Aug 10	23.35191
Sep 10	19.44388
Oct 10	15.67176
Nov 10	17.70992
Dec 10	20.26294

Forecast for electricity: $(2,0,0)(1,1,1)_{12}$

Elec	Forecast
Jul 10	122.5677
Aug 10	122.3426
Sep 10	122.4392
Oct 10	126.9475
Nov 10	128.7815
Dec 10	129.0084