Homework#4

Roderick Quiel
STA 4853
Prof. Fred W. Huffer

Problem#1

1(a)

For the series Y = Log(X) (Repair) I decided to use an ARIMA(0,1,1)(0,1,1)$_{12}$ without a constant (NOCONSTANT).
I chose to go with *d = 1* as the raw series X appeared to have nonconstant mean which we can confirmed by visual inspection of the series X plot and of the ACF, as is decaying slowly.

After differencing once the series X appeared to have a constant mean and a rapidly decaying ACF, and as the sample IACF does not appear to decay slowly, the series does not appear to be over differenced. The variability of the series X seemed to increase with the level, so the transformation Y = Log(X) seemed necessary.

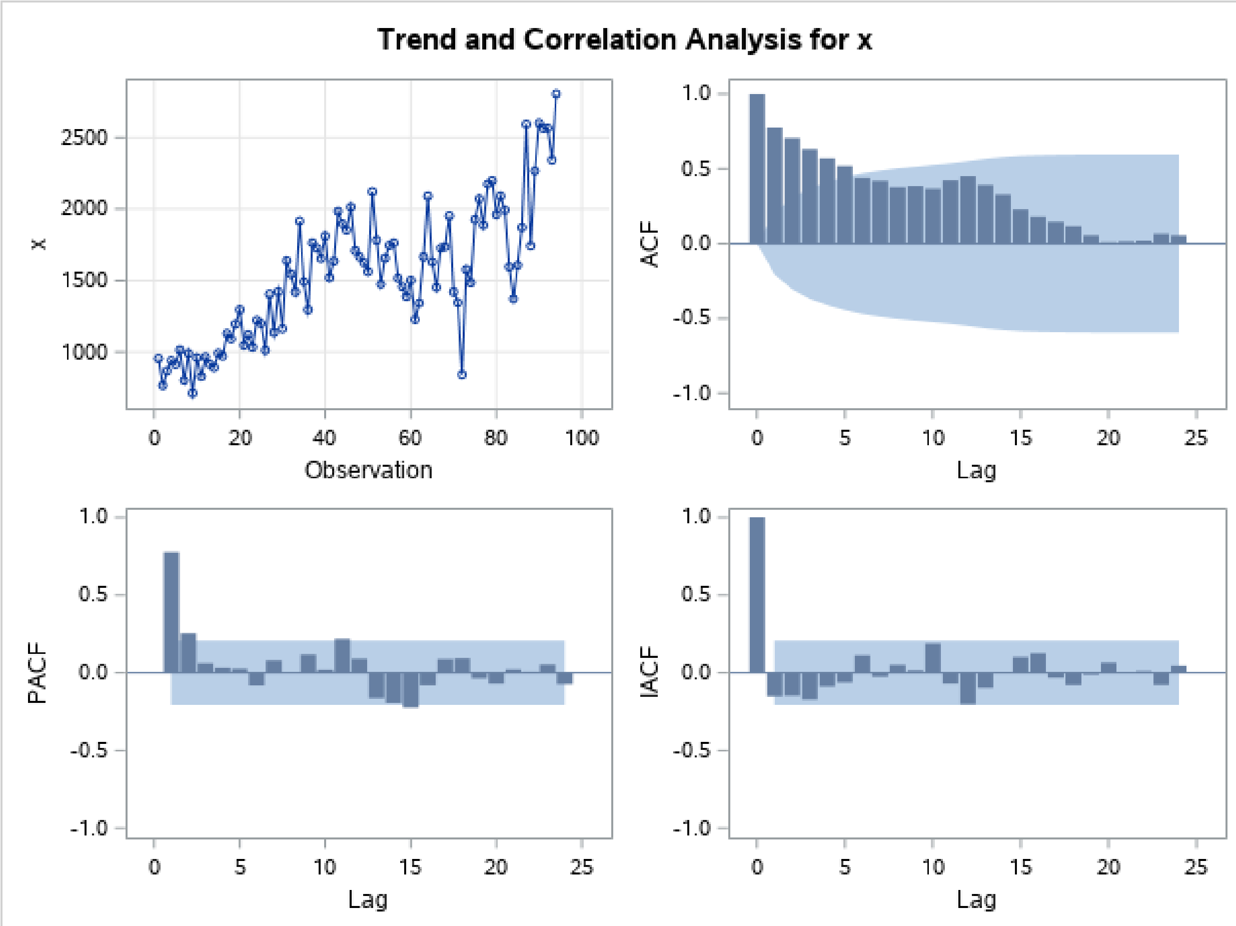
Looking at the IACF we can barely see some seasonal trend, with a spike at lag 12. I decided to do a seasonal difference D = 12 of the series Y, which we already differenced once d = 1. This gave me another stationary series, with a fast-decaying ACF and even better looking IACF.

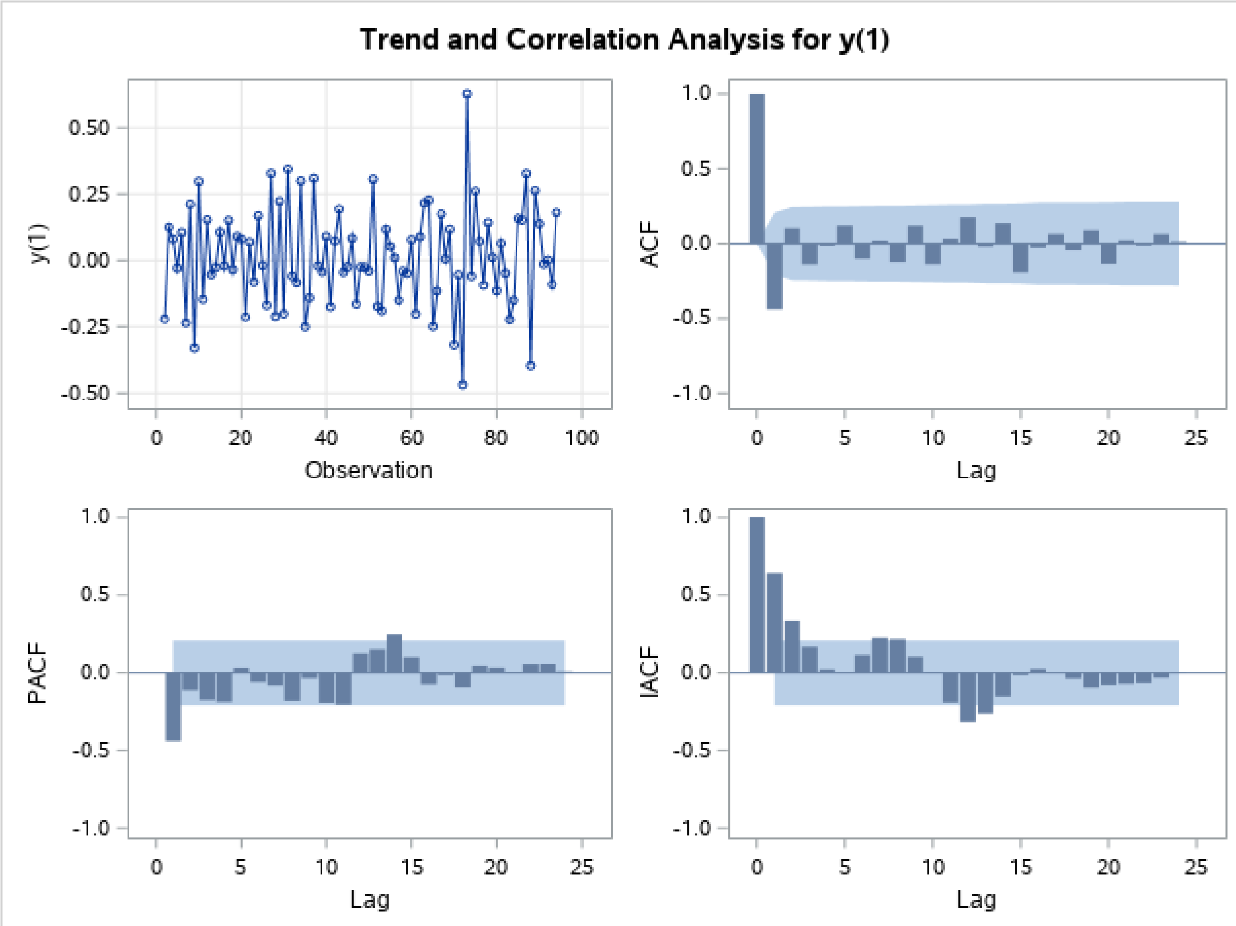
After testing various plausible models for the series Y, I decided that an ARIMA(0,1,1)(0,1,1)$_{12}$ was the most reasonable fitting. I chose this by looking at the ACF of the differenced series Y, with d =1 and D = 12, as it has a cutoff after lag q = 1 and a spike at lag 12 Q = 1.

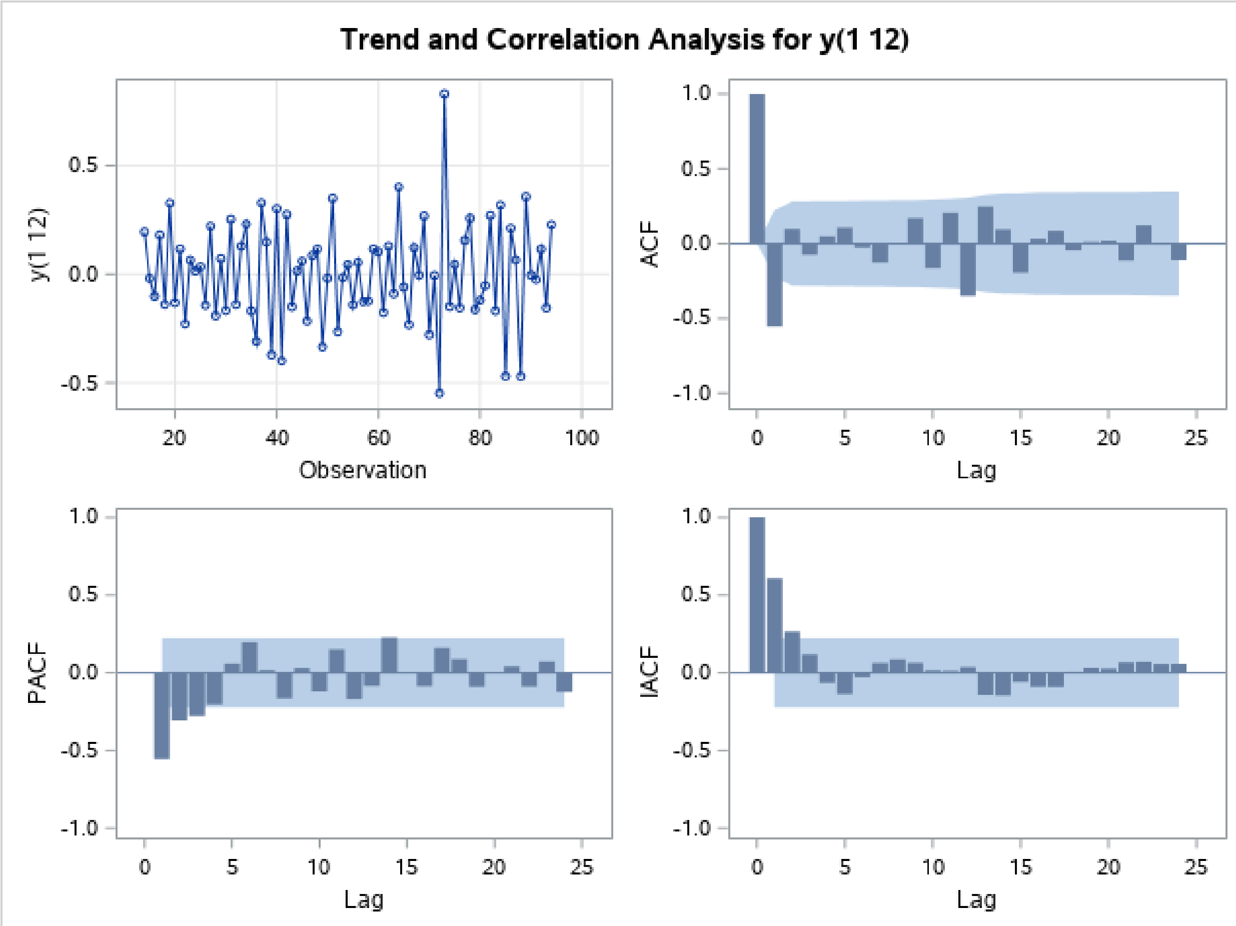
From the **proc arima** result, on the **Maximum Likelihood Estimation** table, we find that the estimate for μ has a p-value > 0.05, so I decided to drop it form the model. Continuing to the **Residual Diagnostics** we can see that both the ACF and PACF decay rapidly to zero and from the **White Noise Probability** graph we see mostly high values. This tells us that the residuals are random shocks, which is what we want.

We also see in the **Residual Normality Diagnostics** that the residuals appear fairly normally distributed, which we can tell from the distribution plot of the residual and the QQ plot.

Finally, we see from **Time series of residuals** plot appears to have a constant mean about zero, as if we drew a line through zero it would be inside our confident limit band, and a constant variance as the variability of plot remain constant about the mean. In **Forecast v.s Residuals** plot we see that, for low values of the forecasts, the confidence limit band does not remain at level zero, but it is still mostly constant about zero.


Trend and Correlation Analysis for x
x
Observation
ACF
Lag
PACF
Lag
IACF
Lag


Trend and Correlation Analysis for y(1)
y(1)
Observation
ACF
Lag
PACF
Lag
IACF
Lag



Maximum Likelihood Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Lag**
MU	0.0002960	0.0032274	0.09	0.9269	0
MA1,1	0.68135	0.08238	8.27	<.0001	1
MA2,1	0.55894	0.12236	4.57	<.0001	12

Model ARIMA(0,1,1)(0,1,1)_12 for Repair series

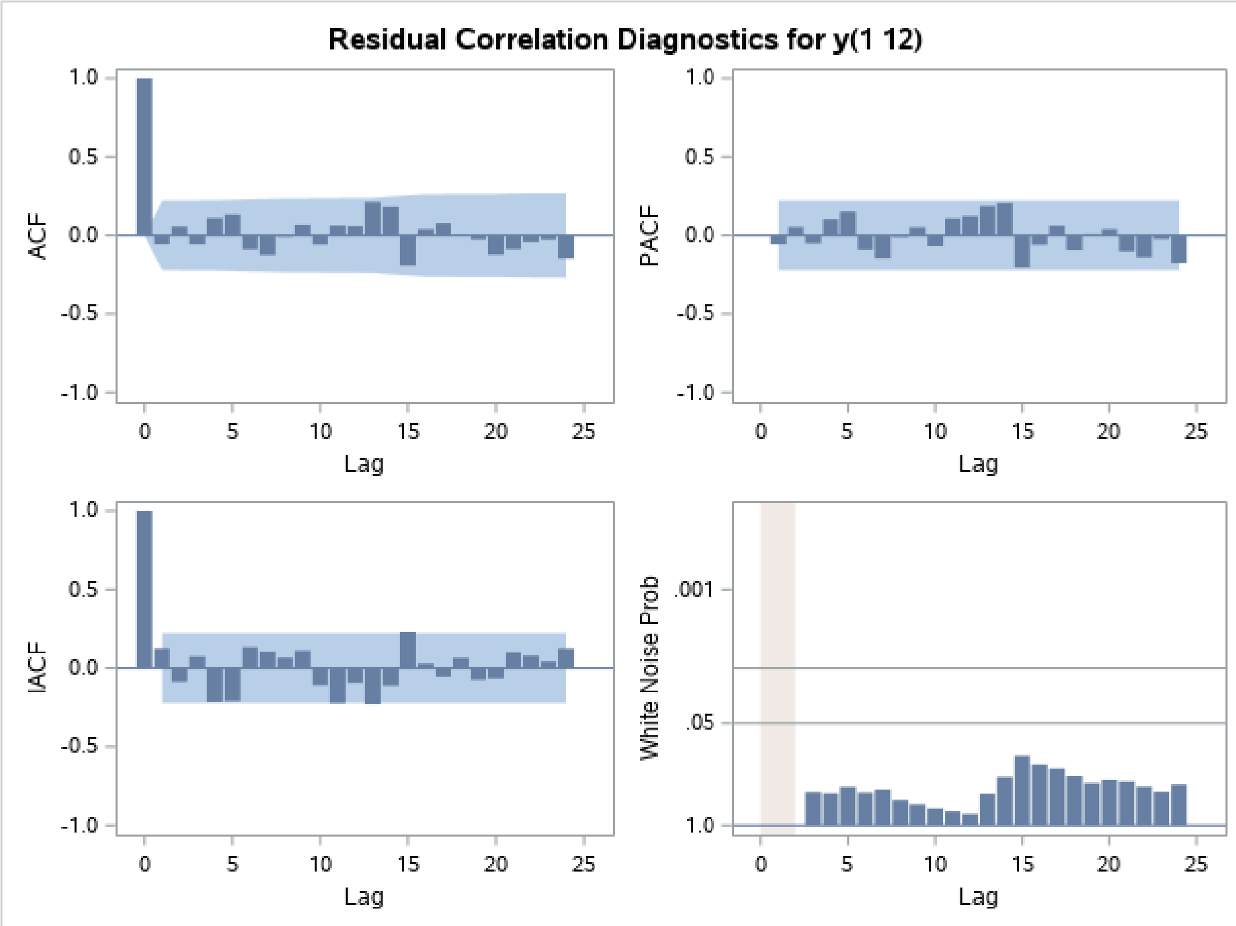
The ARIMA Procedure

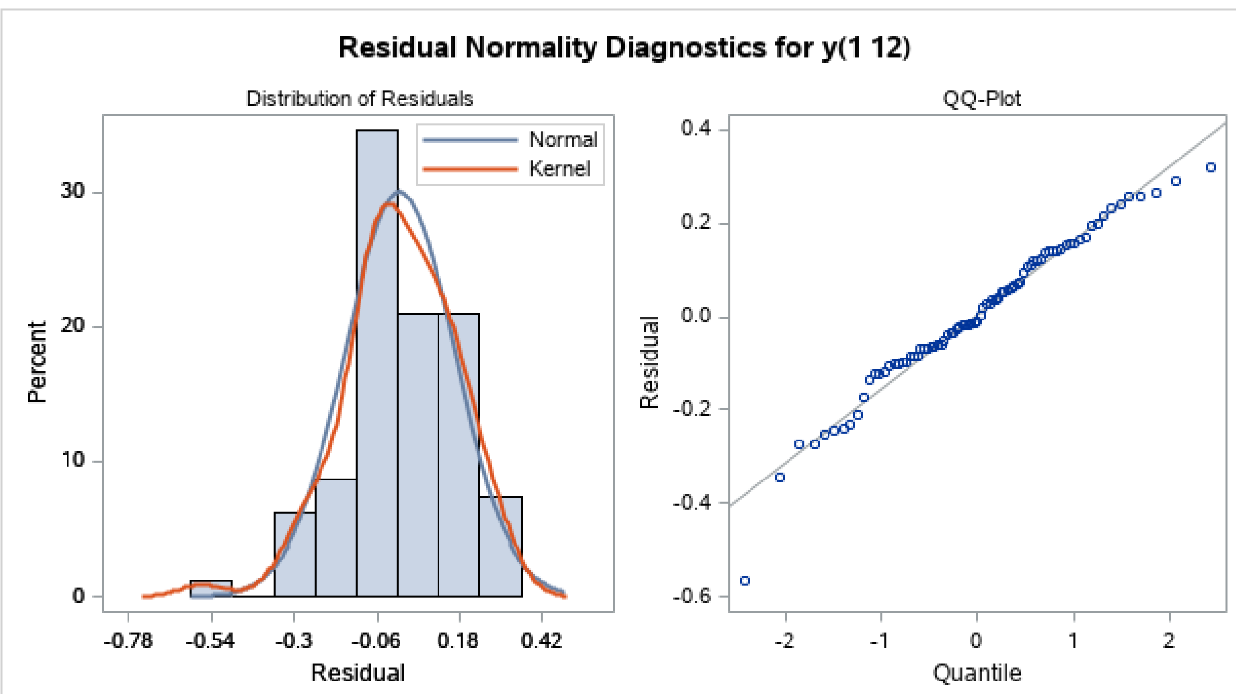
Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MA1,1	0.68190	0.08160	8.36	<.0001	1
MA2,1	0.55991	0.12097	4.63	<.0001	12

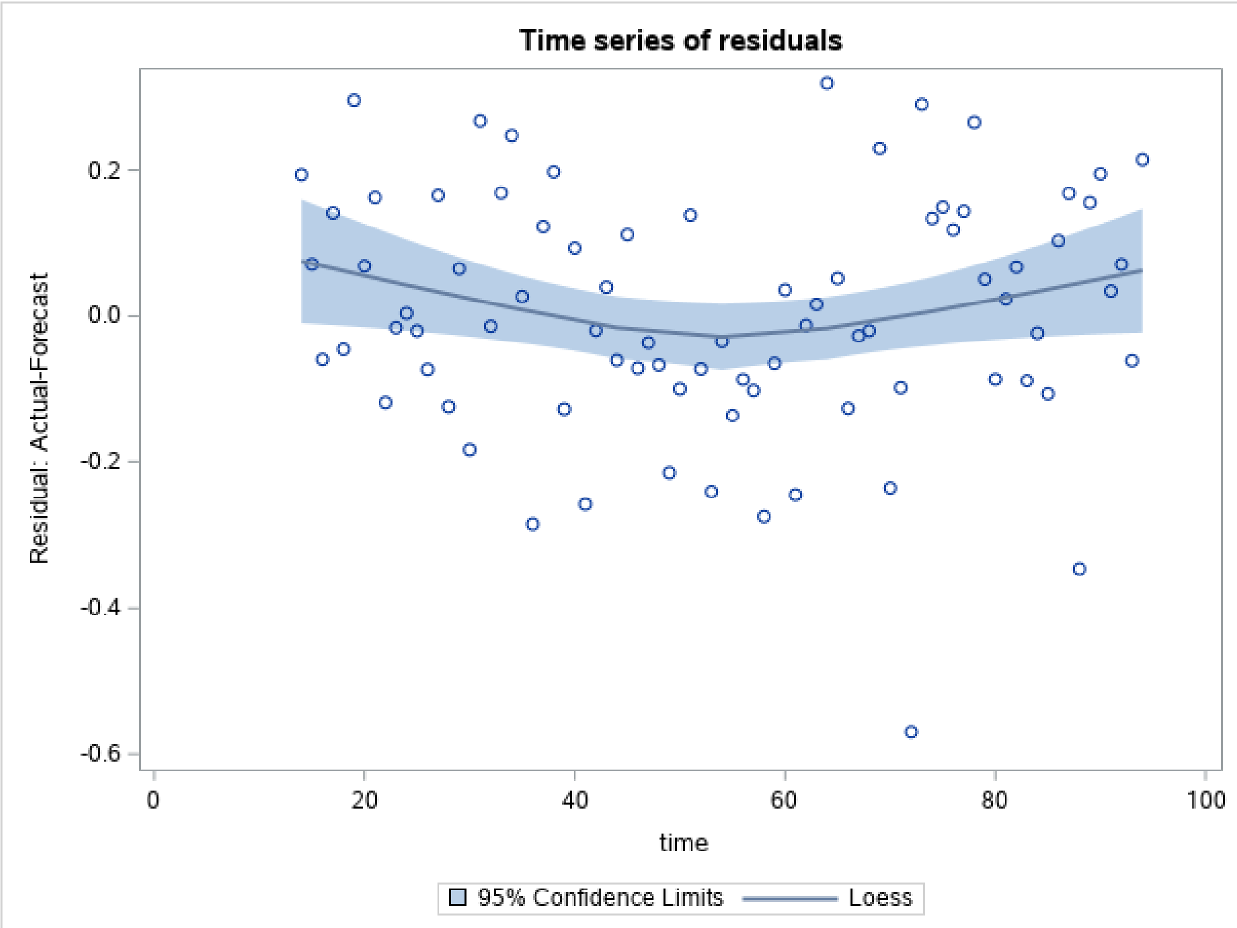
Variance Estimate	0.025771
Std Error Estimate	0.160532
AIC	-59.3481
SBC	-54.5592
Number of Residuals	81

Correlations of Parameter Estimates		
Parameter	MA1,1	MA2,1
MA1,1	1.000	-0.004
MA2,1	-0.004	1.000

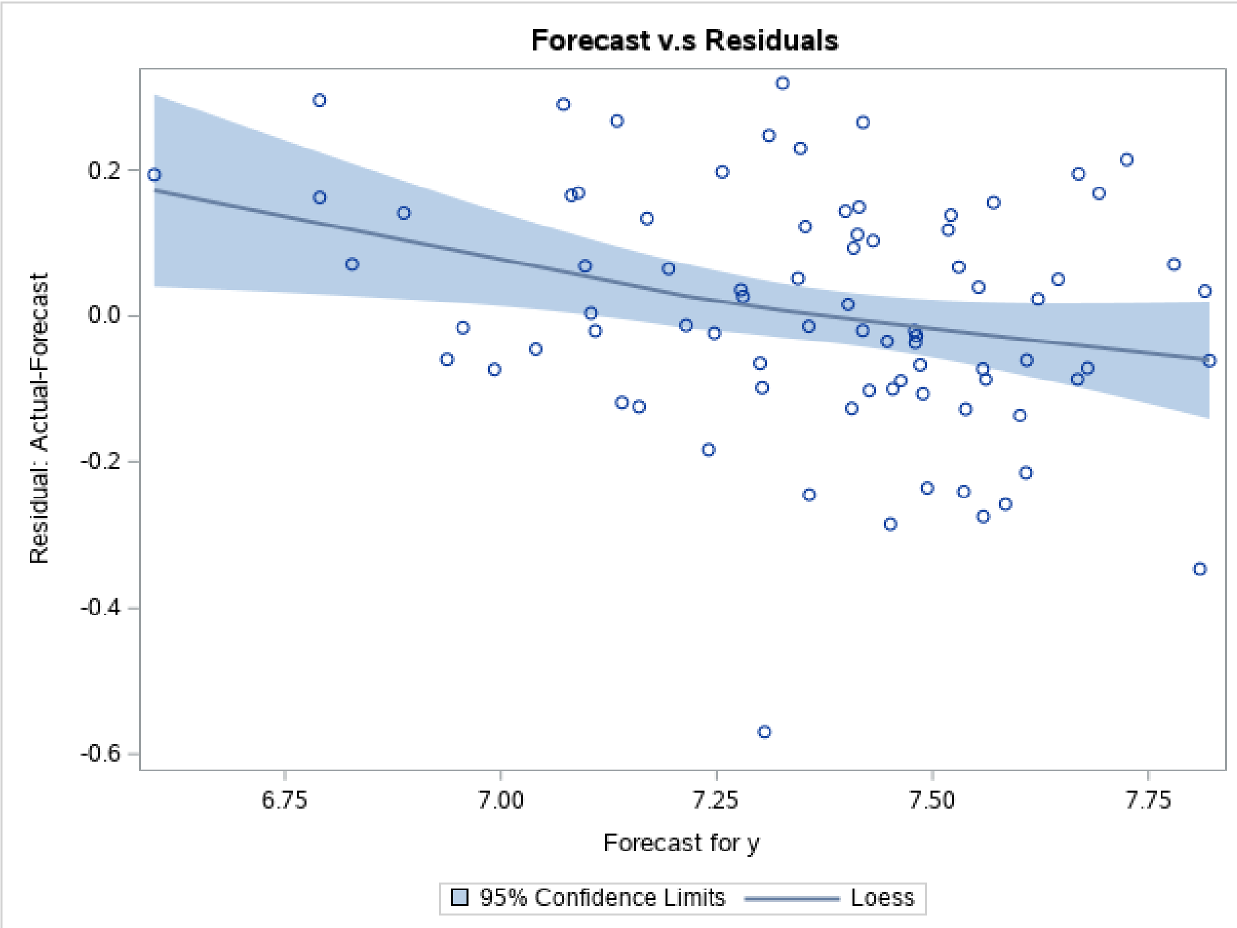
Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	4.17	4	0.3840	-0.055	0.055	-0.055	0.112	0.134	-0.089
12	7.02	10	0.7237	-0.124	-0.011	0.069	-0.058	0.062	0.057
18	19.47	16	0.2451	0.213	0.182	-0.194	0.037	0.079	0.004
24	25.14	22	0.2904	-0.028	-0.124	-0.092	-0.047	-0.030	-0.148


Residual Correlation Diagnostics for y(1 12)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0 5 10 15 20 25


Residual Normality Diagnostics for y(1 12)
Distribution of Residuals
Normal
Kernel
Percent
Residual
-0.78 -0.54 -0.3 -0.06 0.18 0.42
QQ-Plot
Quantile
0.4 0.2 0.0 -0.2 -0.4 -0.6
-2 -1 0 1 2



Time series of residuals
Residual: Actual-Forecast
0.2
0.0
-0.2
-0.4
-0.6
0
20
40
60
80
100
time
95% Confidence Limits
Loess


Forecast v.s Residuals
Residual: Actual-Forecast
0.2
0.0
-0.2
-0.4
-0.6
6.75
7.00
7.25
7.50
7.75
Forecast for y
95% Confidence Limits
Loess

1(b)

Adding the seasonal term clearly affects the residual diagnostics of the model. The model without the seasonal term has a spike in lag 12 in the ACF, PACF and IACF, and we see on the **White Noise Probability** chart that we have multiple significant white noise at different lags, which we don't have in the model with the seasonal term.
We have a bigger magnitude for the AIC/SBC in the model without a seasonal term than the one with the seasonal term.

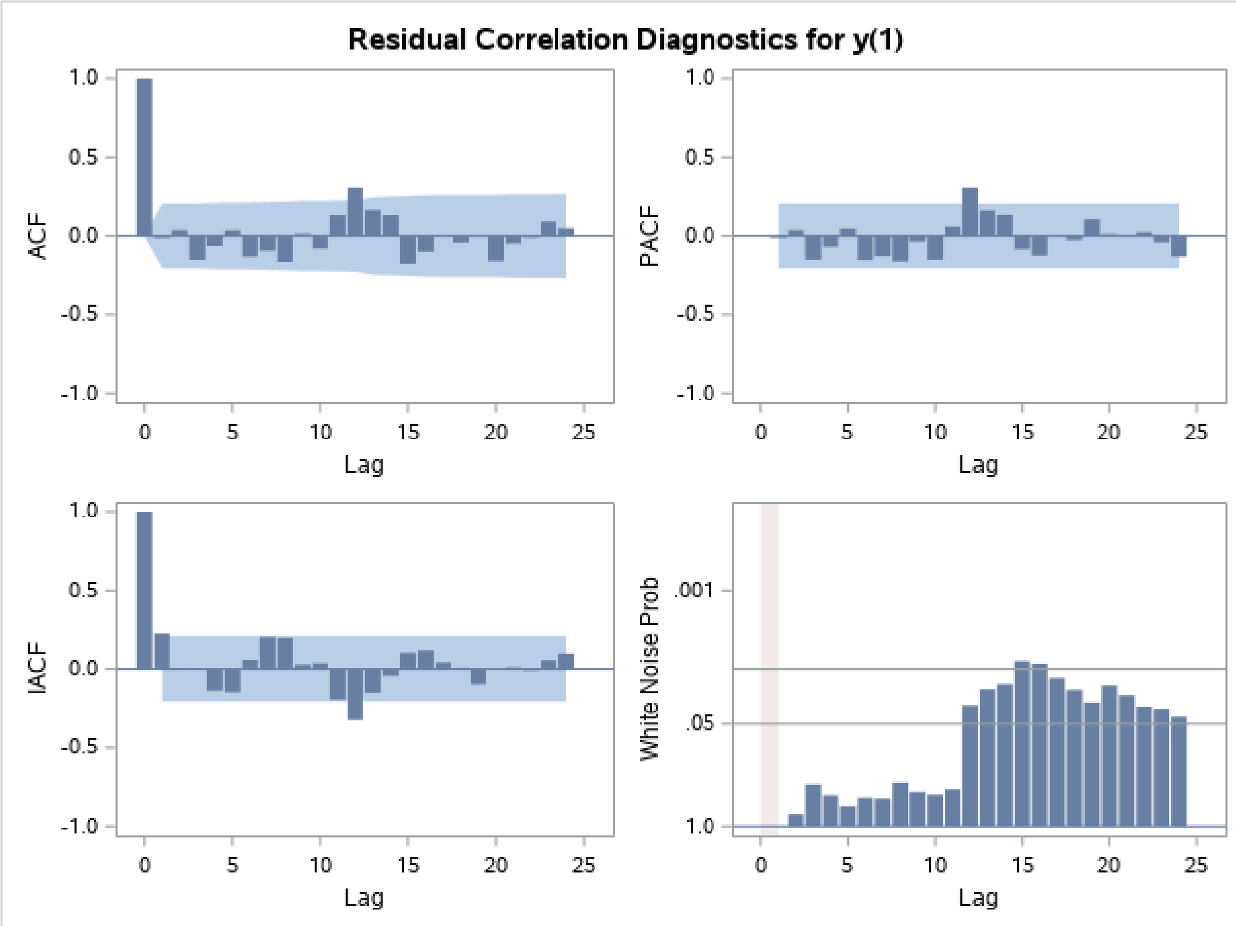
Model without seasonal term

The ARIMA Procedure

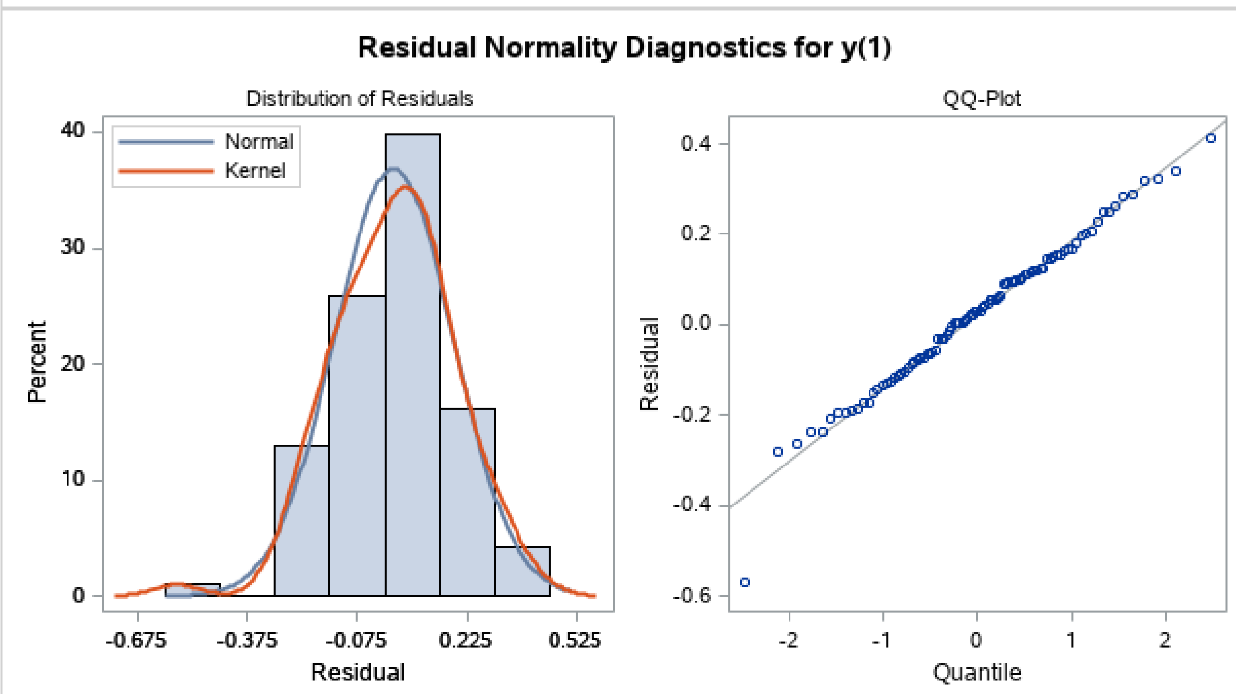
Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MA1,1	0.54508	0.08829	6.17	<.0001	1

Variance Estimate	0.026972
Std Error Estimate	0.164233
AIC	-70.7336
SBC	-68.201
Number of Residuals	93

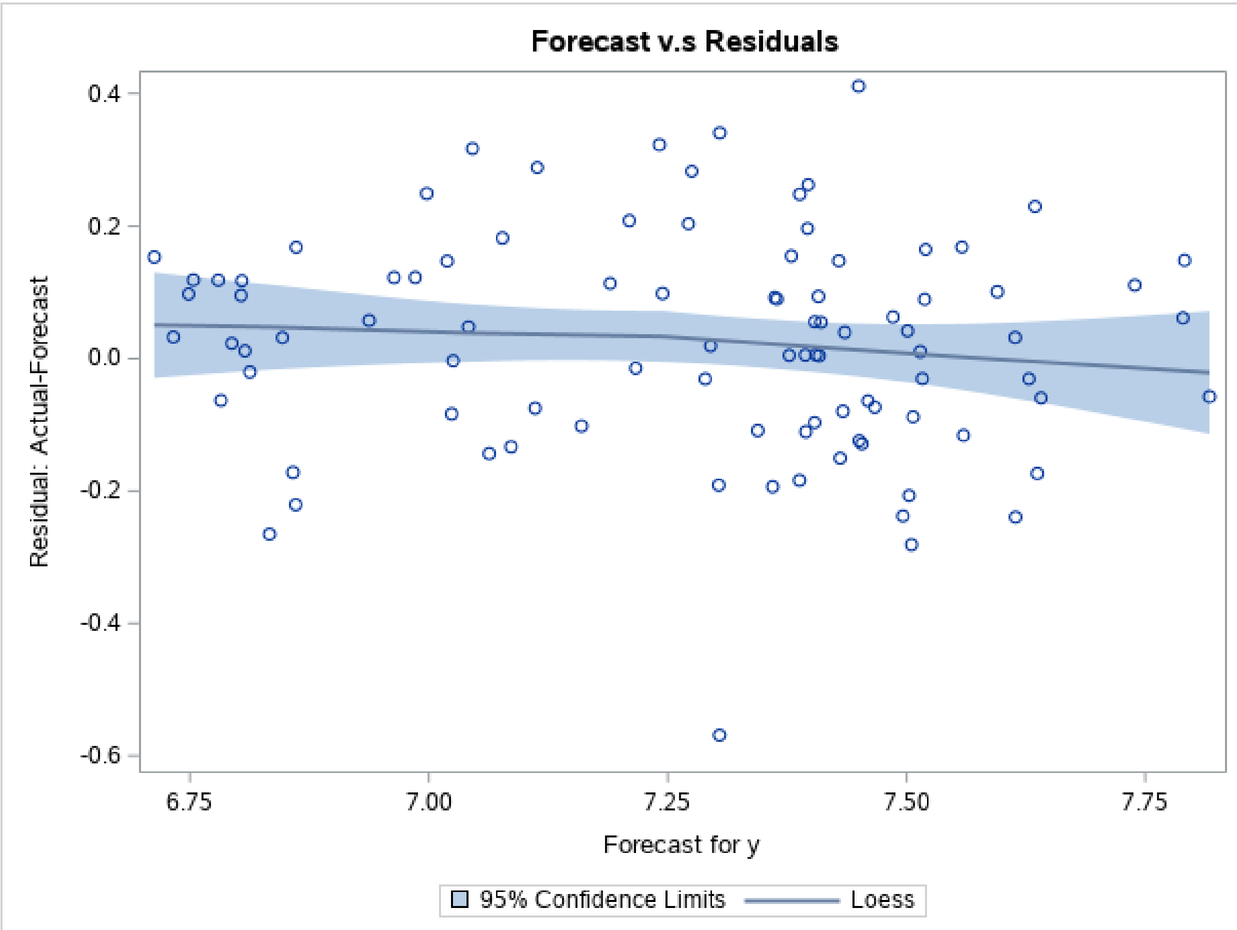
Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	3.70	5	0.5927	0.012	0.059	-0.128	-0.044	0.052	-0.112
12	20.38	11	0.0404	-0.075	-0.145	0.028	-0.065	0.145	0.319
18	29.97	17	0.0266	0.180	0.148	-0.154	-0.082	0.020	-0.025
24	34.55	23	0.0576	0.014	-0.148	-0.037	-0.003	0.100	0.060



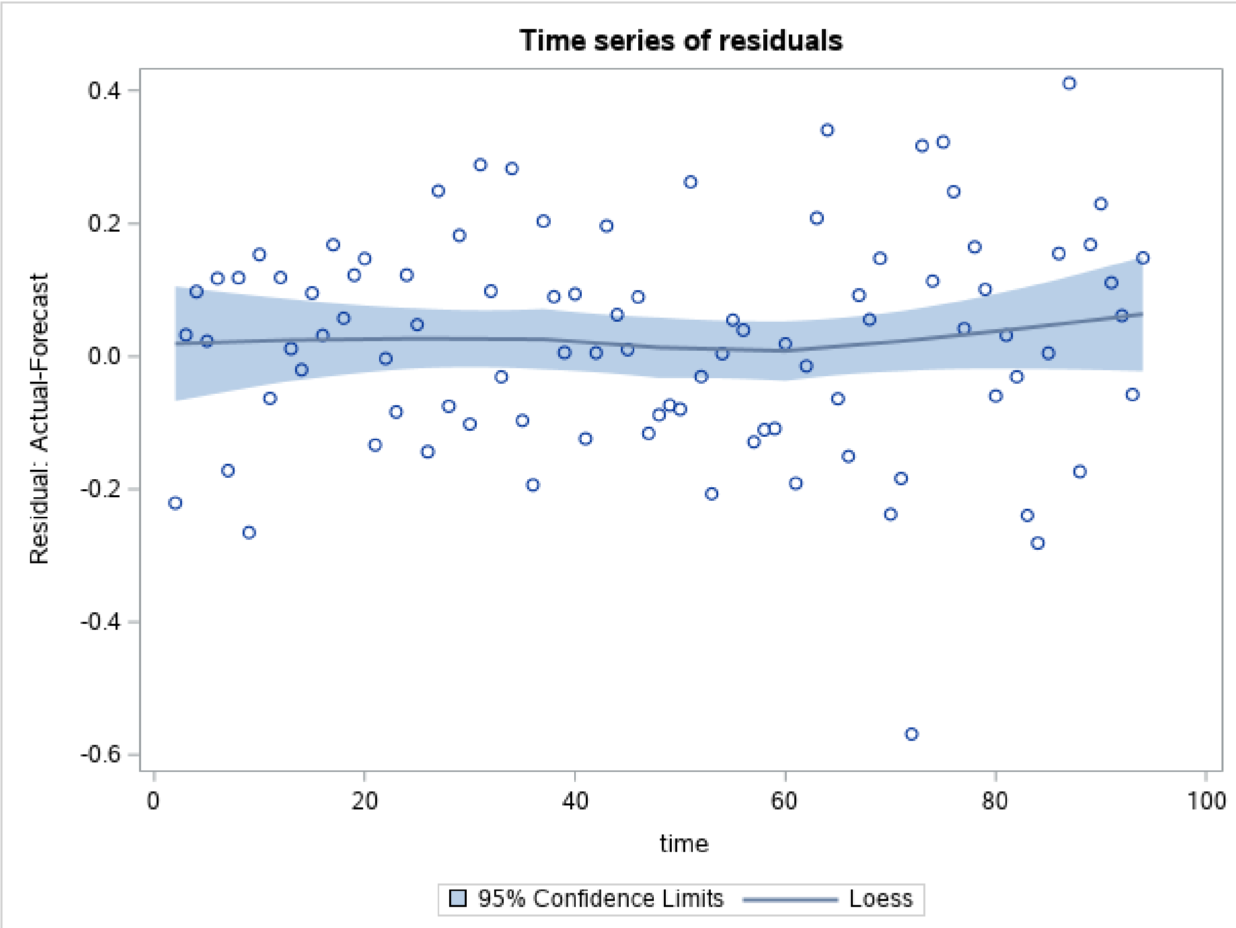
Residual Correlation Diagnostics for y(1)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
0
5
10
15
20
25
.001
.05



Residual Normality Diagnostics for y(1)
Distribution of Residuals
Normal
Kernel
Percent
Residual
-0.675
-0.375
-0.075
0.225
0.525
QQ-Plot
Quantile



Forecast v.s Residuals
Residual: Actual-Forecast
0.4
0.2
0.0
-0.2
-0.4
-0.6
6.75
7.00
7.25
7.50
7.75
Forecast for y
95% Confidence Limits
Loess

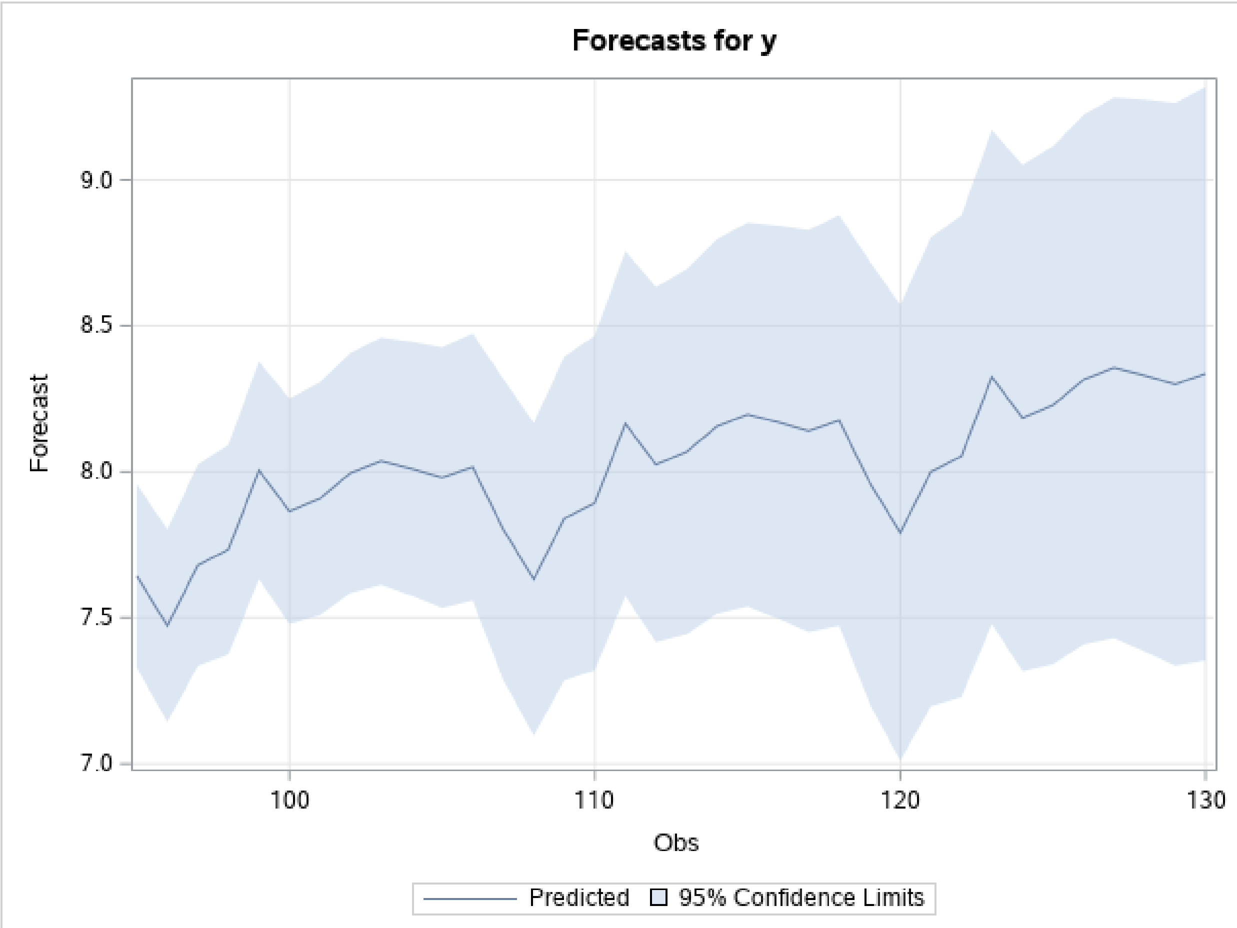

Time series of residuals
Residual: Actual-Forecast
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
20
40
60
80
100
time
95% Confidence Limits
Loess

1(c)

In my case the little difference between the forecasts and the confidence interval widths of the model with a constant and the one without a constant. Most likely because in my original model the constant was not significant, thus plays little to nothing for the forecasts.

Forecasts for variable y				
Obs	Forecast	Std Error	95% Confidence Limits	
95	7.6436	0.1605	7.3290	7.9582
96	7.4721	0.1685	7.1419	7.8023
97	7.6804	0.1760	7.3354	8.0254
98	7.7337	0.1833	7.3745	8.0929
99	8.0052	0.1903	7.6323	8.3781
100	7.8651	0.1970	7.4790	8.2512
101	7.9085	0.2035	7.5096	8.3074
102	7.9961	0.2098	7.5848	8.4073
103	8.0365	0.2159	7.6133	8.4598
104	8.0104	0.2219	7.5755	8.4454
105	7.9803	0.2277	7.5340	8.4266
106	8.0163	0.2334	7.5590	8.4737
107	7.8033	0.2632	7.2875	8.3192
108	7.6319	0.2733	7.0963	8.1675
109	7.8402	0.2830	7.2855	8.3948
110	7.8935	0.2924	7.3204	8.4665
111	8.1650	0.3015	7.5740	8.7559
112	8.0248	0.3103	7.4166	8.6331
113	8.0682	0.3189	7.4432	8.6933
114	8.1558	0.3273	7.5143	8.7973
115	8.1963	0.3355	7.5388	8.8538
116	8.1702	0.3434	7.4971	8.8433
117	8.1400	0.3512	7.4517	8.8284
118	8.1761	0.3588	7.4728	8.8794
119	7.9631	0.3867	7.2051	8.7210
120	7.7916	0.3985	7.0107	8.5726
121	7.9999	0.4099	7.1966	8.8032
122	8.0532	0.4210	7.2282	8.8783
123	8.3247	0.4318	7.4785	9.1709
124	8.1846	0.4423	7.3177	9.0515
125	8.2280	0.4526	7.3409	9.1151
126	8.3156	0.4627	7.4087	9.2224
127	8.3560	0.4725	7.4299	9.2822
128	8.3299	0.4822	7.3849	9.2750
129	8.2998	0.4917	7.3361	9.2634
130	8.3358	0.5009	7.3540	9.3177

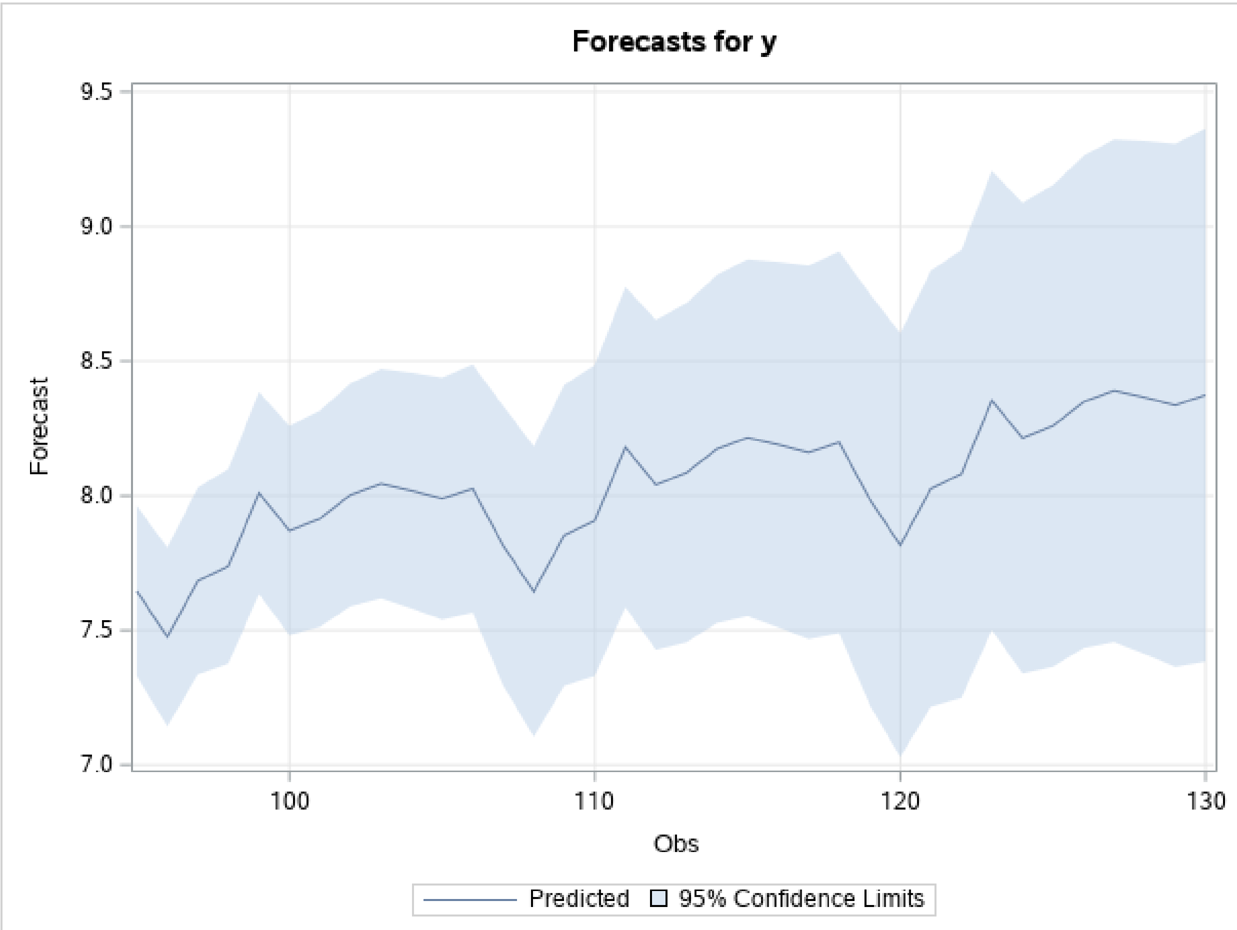
Forecast for model without constant.



Forecast for model without constant.

Forecasts for variable y				
Obs	Forecast	Std Error	95% Confidence Limits	
95	7.6456	0.1616	7.3290	7.9623
96	7.4746	0.1696	7.1422	7.8069
97	7.6837	0.1772	7.3364	8.0310
98	7.7379	0.1845	7.3762	8.0996
99	8.0102	0.1916	7.6346	8.3857
100	7.8704	0.1984	7.4815	8.2592
101	7.9147	0.2050	7.5130	8.3164
102	8.0031	0.2113	7.5889	8.4173
103	8.0441	0.2175	7.6178	8.4704
104	8.0186	0.2235	7.5806	8.4567
105	7.9892	0.2294	7.5396	8.4387
106	8.0260	0.2351	7.5653	8.4868
107	7.8144	0.2652	7.2946	8.3342
108	7.6436	0.2754	7.1039	8.1834
109	7.8530	0.2852	7.2941	8.4120
110	7.9075	0.2947	7.3299	8.4851
111	8.1801	0.3039	7.5845	8.7757
112	8.0406	0.3128	7.4275	8.6537
113	8.0853	0.3215	7.4551	8.7154
114	8.1739	0.3299	7.5272	8.8206
115	8.2152	0.3382	7.5524	8.8781
116	8.1901	0.3462	7.5115	8.8686
117	8.1609	0.3541	7.4669	8.8549
118	8.1980	0.3618	7.4889	8.9071
119	7.9867	0.3899	7.2225	8.7509
120	7.8162	0.4018	7.0287	8.6037
121	8.0259	0.4133	7.2159	8.8360
122	8.0807	0.4245	7.2487	8.9127
123	8.3536	0.4354	7.5002	9.2070
124	8.2144	0.4461	7.3401	9.0887
125	8.2593	0.4565	7.3646	9.1540
126	8.3483	0.4667	7.4337	9.2629
127	8.3899	0.4766	7.4558	9.3240
128	8.3650	0.4864	7.4118	9.3183
129	8.3361	0.4959	7.3642	9.3081
130	8.3736	0.5053	7.3832	9.3639

Forecast for model constant.



Forecast for model constant.

Problem#2

2(a)

For the series Y = Log(X) (Bus) I decided to use an ARIMA$(2,0,0)(1,1,0)_{12}$ with a constant
The raw series X appeared to have nonconstant mean which we can confirmed by visual inspection of the series X plot and of the ACF, as is decaying slowly.

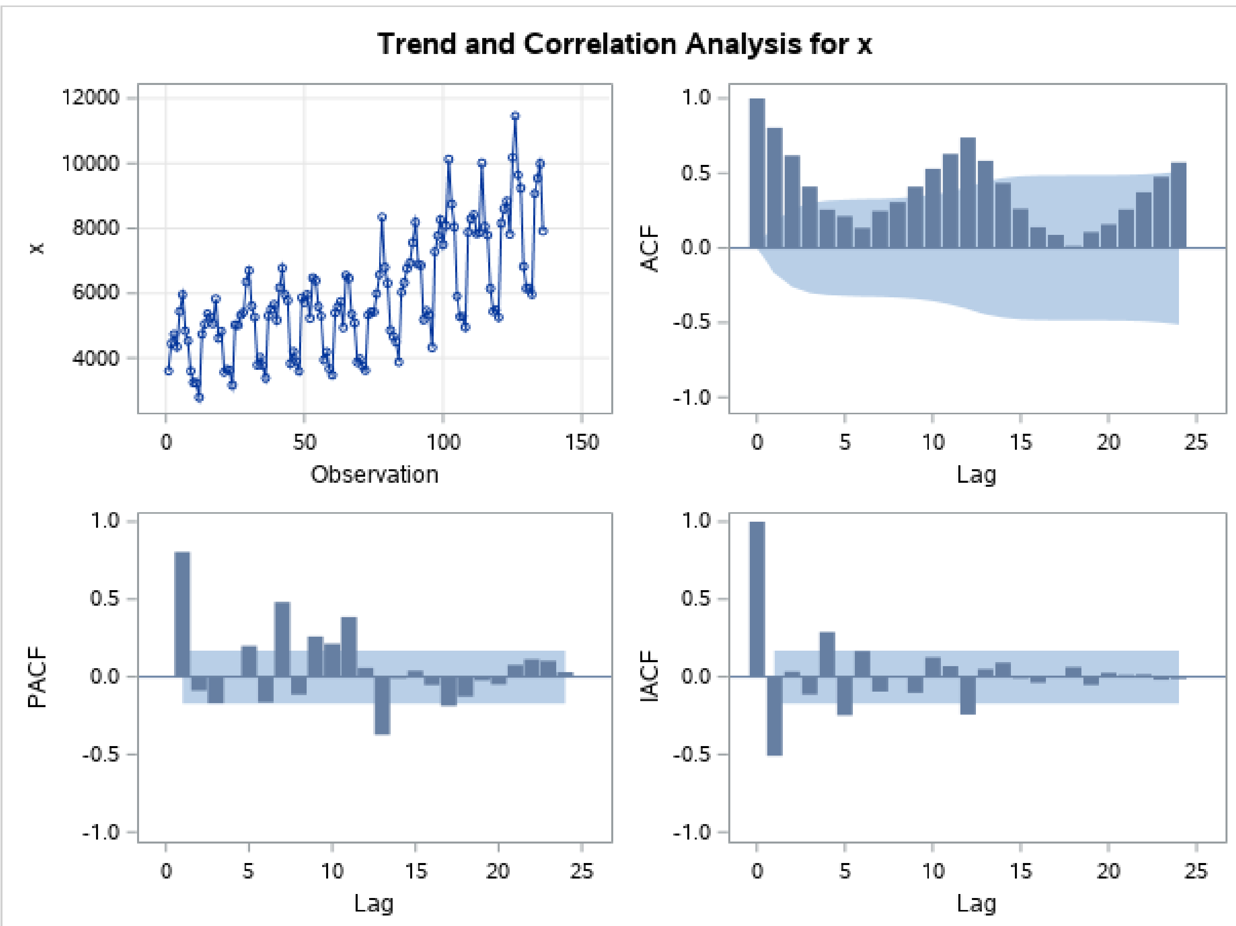
After differencing once the series X appeared to have a constant mean but looking at the ACF we can see it is slowly decaying with some seasonal trend spikes at lag 12 and 24. The variability of the series X seemed to increase with the level, so the transformation Y = Log(X) seemed necessary. I decided to do a seasonal difference D = 12 of the series Y. This gave me a roughly stationary series, with a fast-decaying ACF and a sample IACF that does not appear to decay slowly.

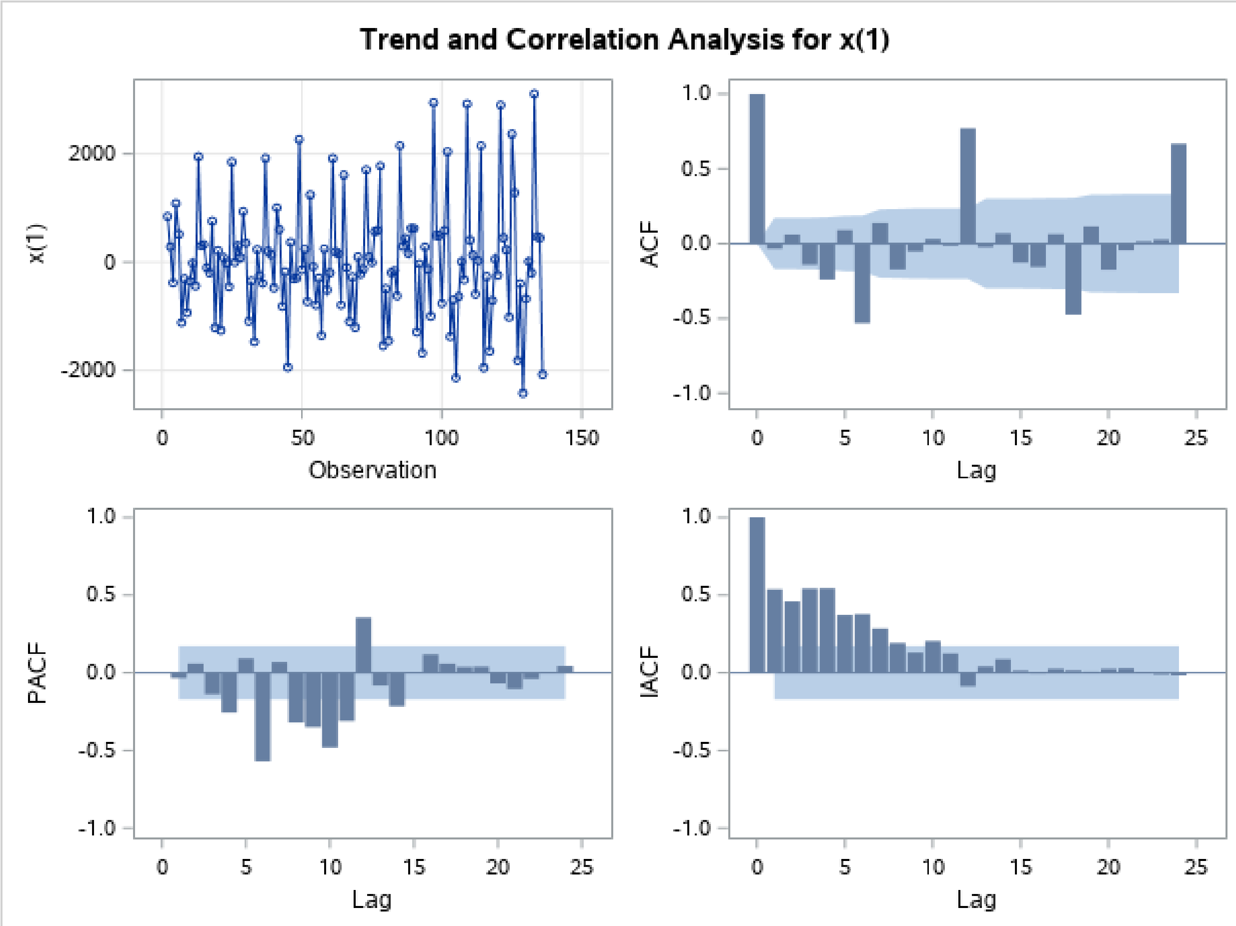
After testing various plausible models for the series Y, I decided that an ARIMA$(2,0,0)(1,1,0)_{12}$ was the most reasonable fitting. I chose this by looking at the PACF of the differenced series Y, with D = 12, as it has a cutoff after lag p = 2 and a spike at lag 12 P = 1.

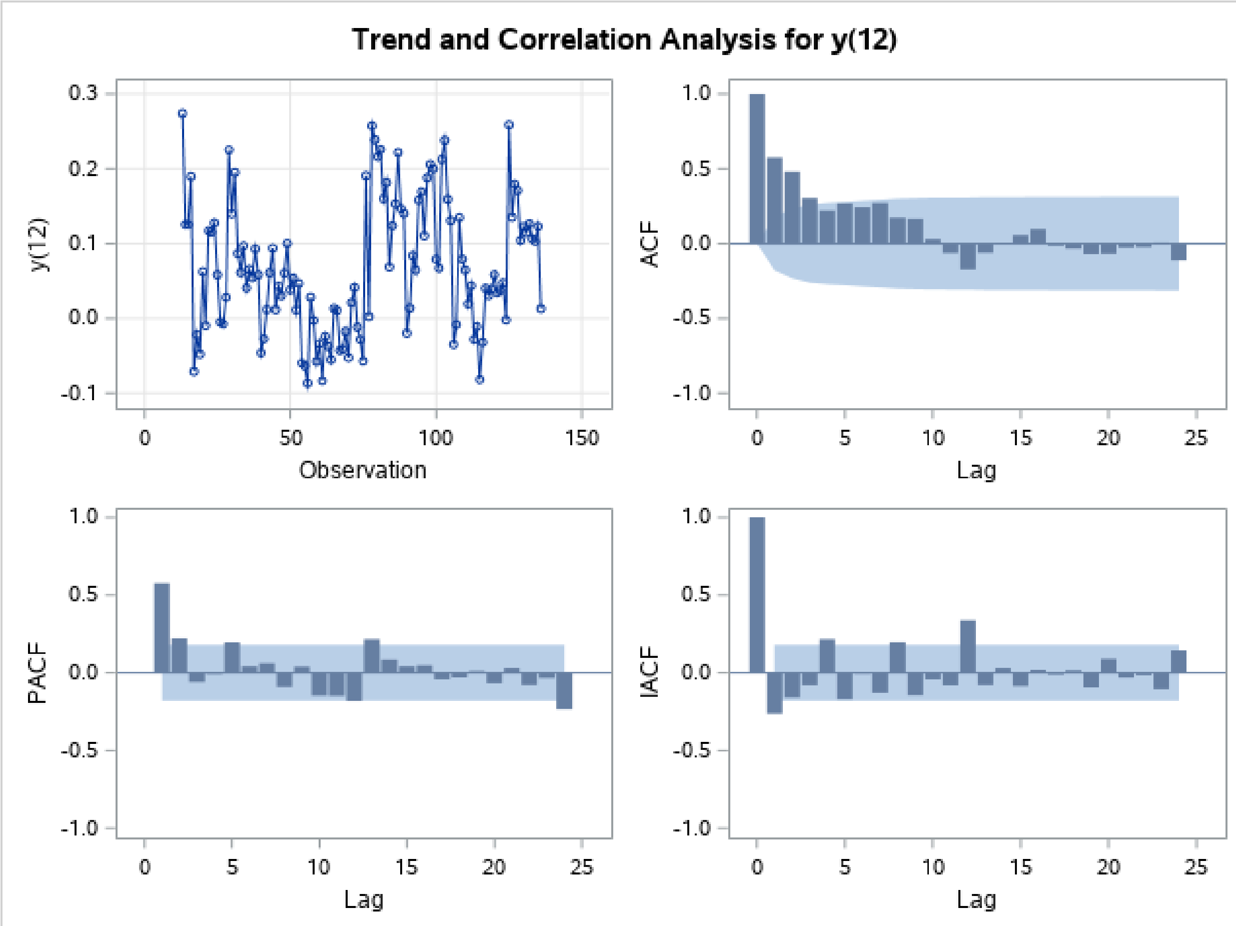
From the **proc arima** result, on the **Maximum Likelihood Estimation** table, we find that all estimates are significant in the model. Continuing to the **Residual Diagnostics** we can see that both the ACF and PACF decay rapidly to zero and from the **White Noise Probability** graph we see mostly high values. This tells us that the residuals are random shocks, which is what we want.

We also see in the **Residual Normality Diagnostics** that the residuals appear fairly normally distributed, which we can tell from the distribution plot of the residual and the QQ plot.

Finally, we see from both **Time series of residuals** and the **Forecast v.s Residuals** plots appear to have a constant mean about zero, as if we drew a line through zero it would be inside our confident limit band, and a constant variance as the variability of plot remain constant about the mean.


Trend and Correlation Analysis for x
x
Observation
ACF
Lag
PACF
Lag
IACF
Lag


Trend and Correlation Analysis for x(1)
x(1)
Observation
ACF
Lag
PACF
Lag
IACF
Lag



Trend and Correlation Analysis for y(12)
y(12)
Observation
ACF
Lag
PACF
Lag
IACF
Lag

Model ARIMA(2,0,0)(0,1,1)_12 for Bus series

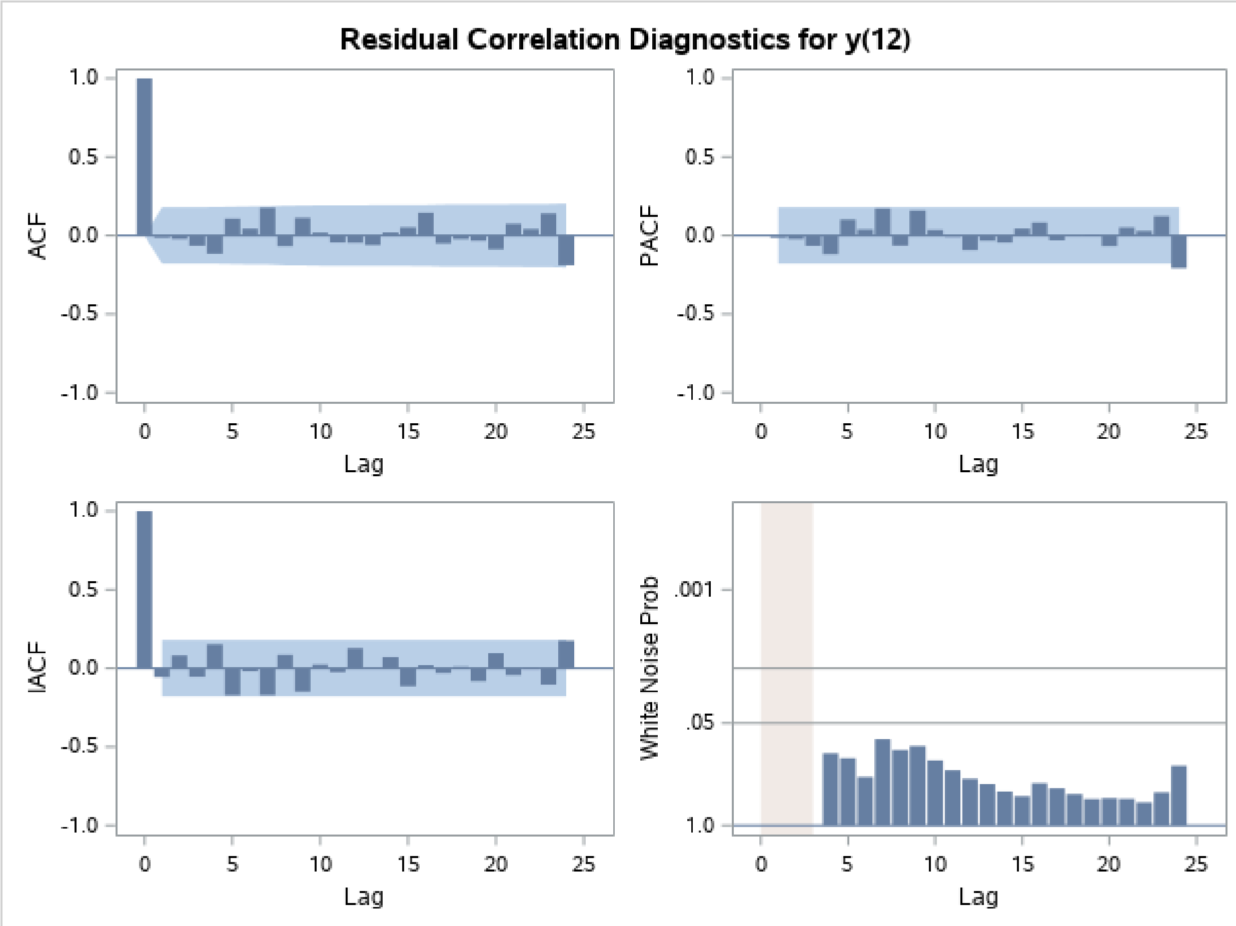
The ARIMA Procedure

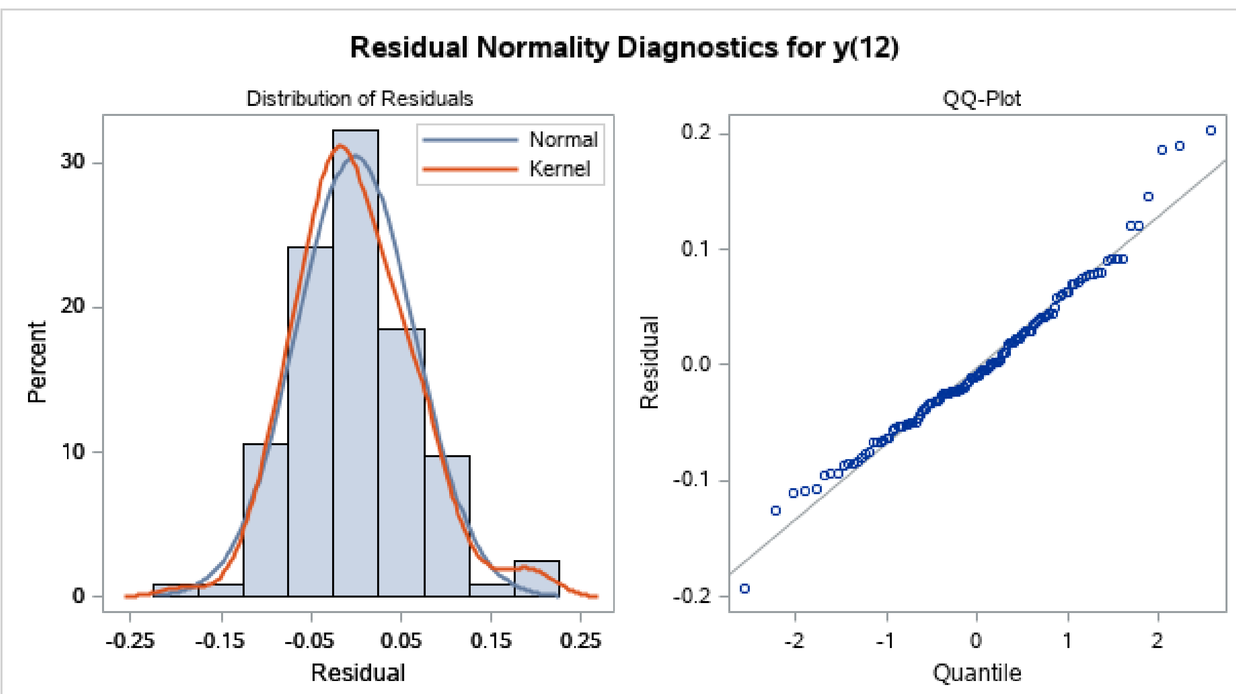
Maximum Likelihood Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Lag**
MU	0.06832	0.01652	4.14	<.0001	0
AR1,1	0.45724	0.08649	5.29	<.0001	1
AR1,2	0.28141	0.08802	3.20	0.0014	2
AR2,1	-0.35994	0.08915	-4.04	<.0001	12

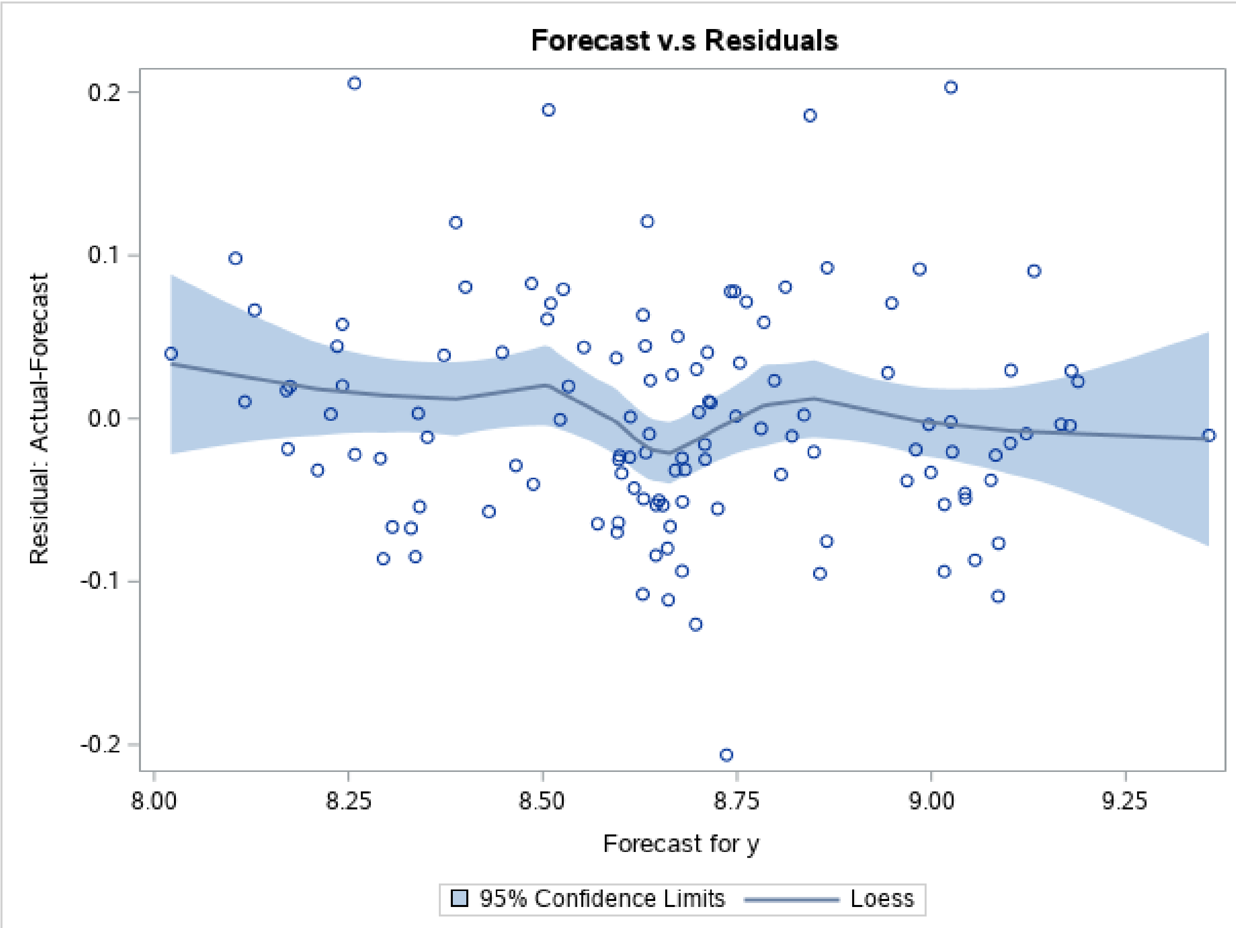
Constant Estimate	0.024281
Variance Estimate	0.004401
Std Error Estimate	0.066342
AIC	-314.681
SBC	-303.4
Number of Residuals	124

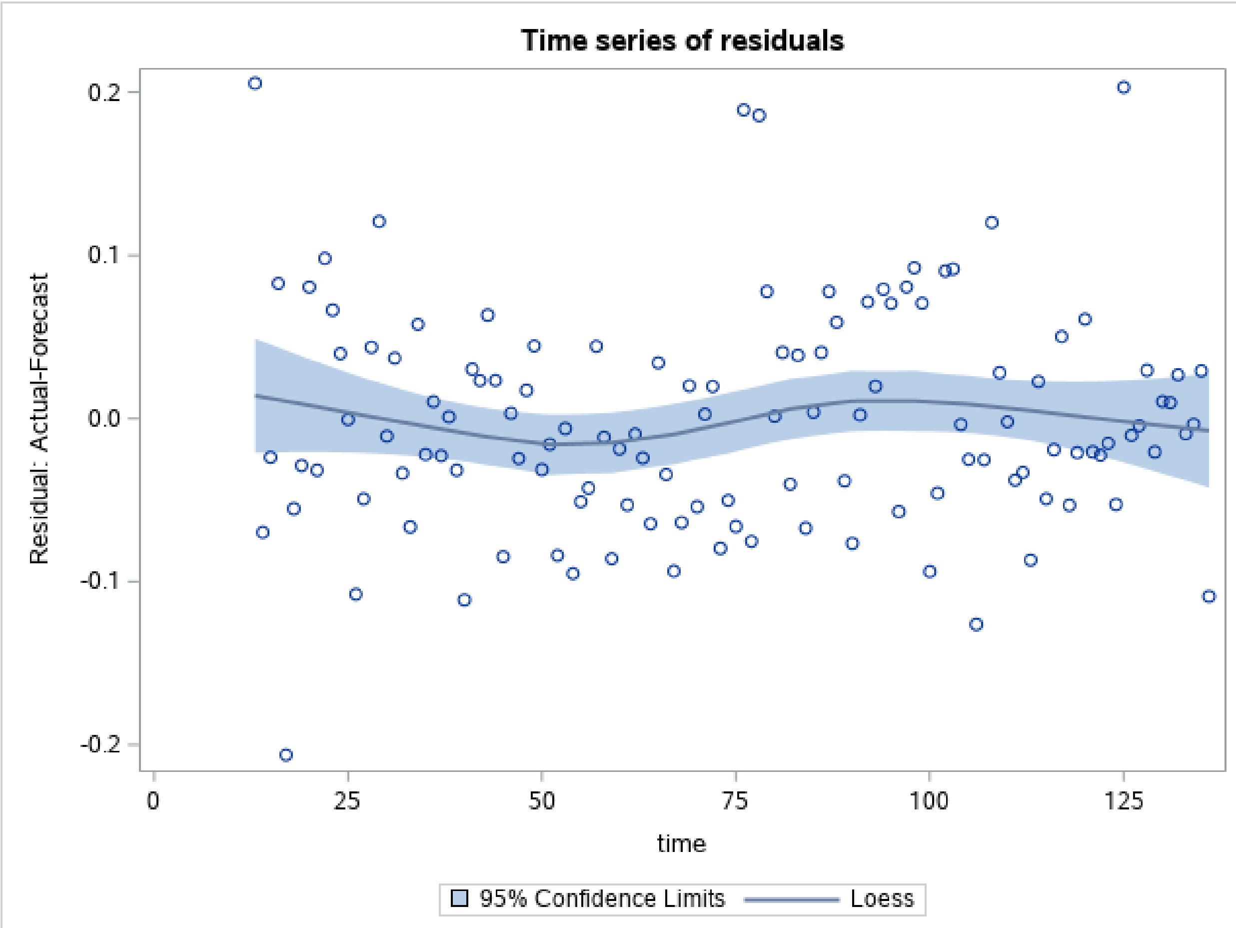
Correlations of Parameter Estimates				
Parameter	**MU**	**AR1,1**	**AR1,2**	**AR2,1**
MU	1.000	-0.019	0.001	0.019
AR1,1	-0.019	1.000	-0.654	0.000
AR1,2	0.001	-0.654	1.000	-0.079
AR2,1	0.019	0.000	-0.079	1.000

Autocorrelation Check of Residuals									
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Autocorrelations**					
6	4.15	3	0.2458	-0.012	-0.021	-0.066	-0.115	0.108	0.043
12	11.25	9	0.2590	0.176	-0.069	0.113	0.019	-0.042	-0.044
18	15.65	15	0.4057	-0.060	0.020	0.051	0.145	-0.048	-0.021
24	26.84	21	0.1763	-0.031	-0.088	0.077	0.041	0.141	-0.190


Residual Correlation Diagnostics for y(12)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0
5
10
15
20
25


Residual Normality Diagnostics for y(12)
Distribution of Residuals
Normal
Kernel
Percent
Residual
-0.25
-0.15
-0.05
0.05
0.15
0.25
0
10
20
30
QQ-Plot
Quantile
0.2
0.1
0.0
-0.1
-0.2
-2
-1
0
1
2


Forecast v.s Residuals
Residual: Actual-Forecast
0.2
0.1
0.0
-0.1
-0.2
8.00
8.25
8.50
8.75
9.00
9.25
Forecast for y
95% Confidence Limits
Loess


Time series of residuals
Residual: Actual-Forecast
0.2
0.1
0.0
-0.1
-0.2
0
25
50
75
100
125
time
95% Confidence Limits
Loess

2(b)

Future 36 Forecasts

Obs	osforecast	osL95	osU95
137	9937.89	8726.18	11317.87
138	11581.70	10038.73	13361.83
139	9701.79	8295.85	11346.00
140	9353.44	7945.47	11010.90
141	7103.77	6006.86	8400.98
142	6376.13	5376.57	7561.53
143	6423.58	5406.71	7631.70
144	6201.01	5213.29	7375.86
145	9515.79	7994.04	11327.22
146	10033.01	8424.43	11948.74
147	10441.77	8764.87	12439.49
148	8617.38	7231.98	10268.17
149	10969.92	9007.86	13359.35
150	12633.19	10324.38	15458.31
151	10603.44	8622.23	13039.89
152	10205.67	8277.14	12583.52
153	7674.59	6212.93	9480.14
154	6899.73	5579.26	8532.71
155	6938.67	5606.54	8587.31
156	6705.08	5415.18	8302.25

Obs	osforecast	osL95	osU95
157	10257.63	8281.69	12705.01
158	10806.88	8723.36	13388.03
159	11273.01	9098.42	13967.36
160	9169.43	7399.98	11361.98
161	11614.98	9148.40	14746.58
162	13434.00	10527.10	17143.61
163	11267.97	8780.99	14459.33
164	10852.07	8432.71	13965.54
165	8190.02	6351.26	10561.11
166	7358.92	5699.58	9501.33
167	7405.31	5730.77	9569.15
168	7153.47	5532.93	9248.64
169	10955.83	8470.97	14169.60
170	11545.73	8925.07	14935.90
171	12033.86	9301.04	15569.63
172	9839.58	7604.35	12731.84

Problem#3

1(a)

For the series X (Eureka) I decided to use an ARIMA(0,1,1) + the trend **xsine1, xcos1, xcos2,** and **xcos3**.

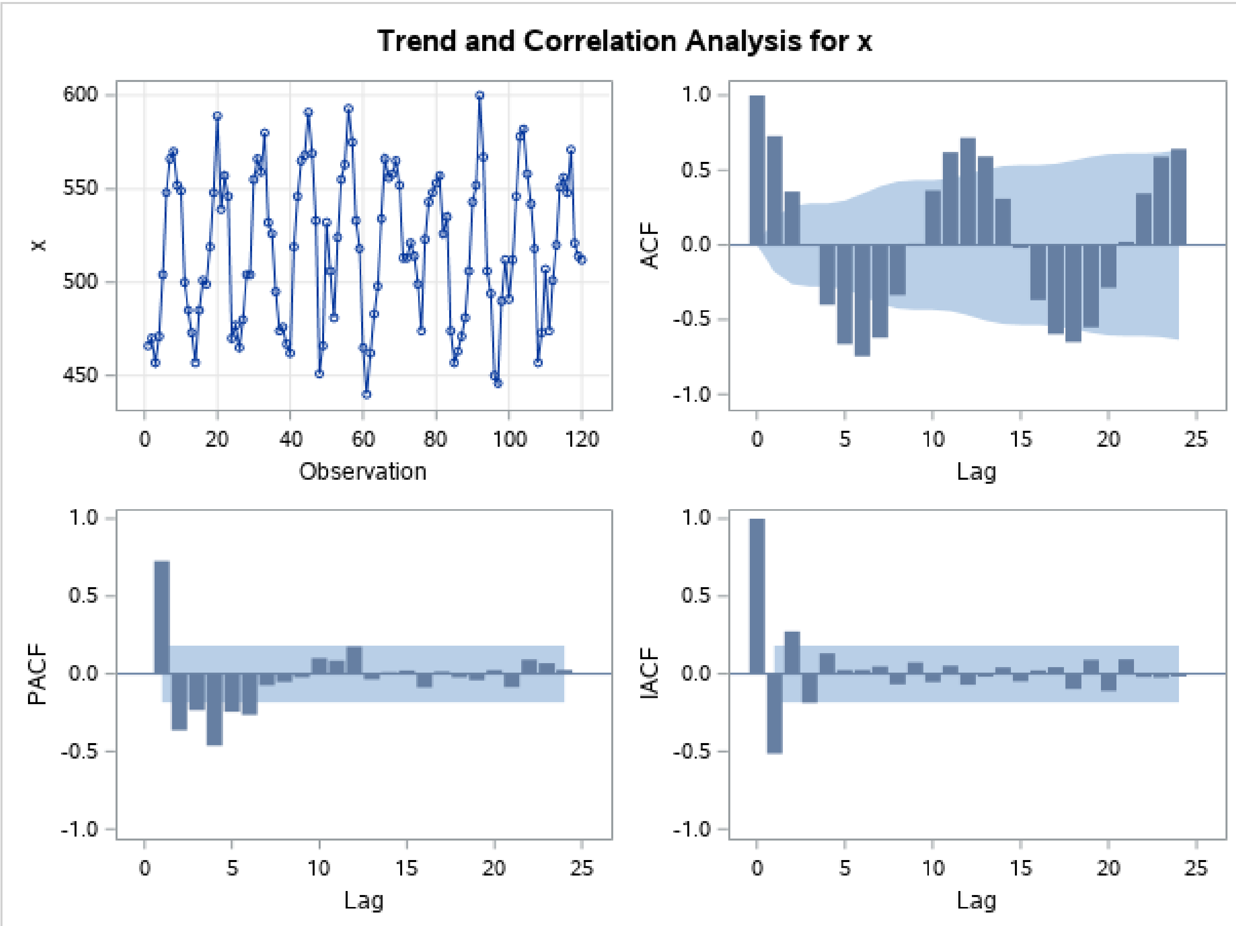
Series X already looks stationary, so we only needed to choose the trend. For this I chose to try an MA(1) for the series as the PACF has a cutoff after lag 1.

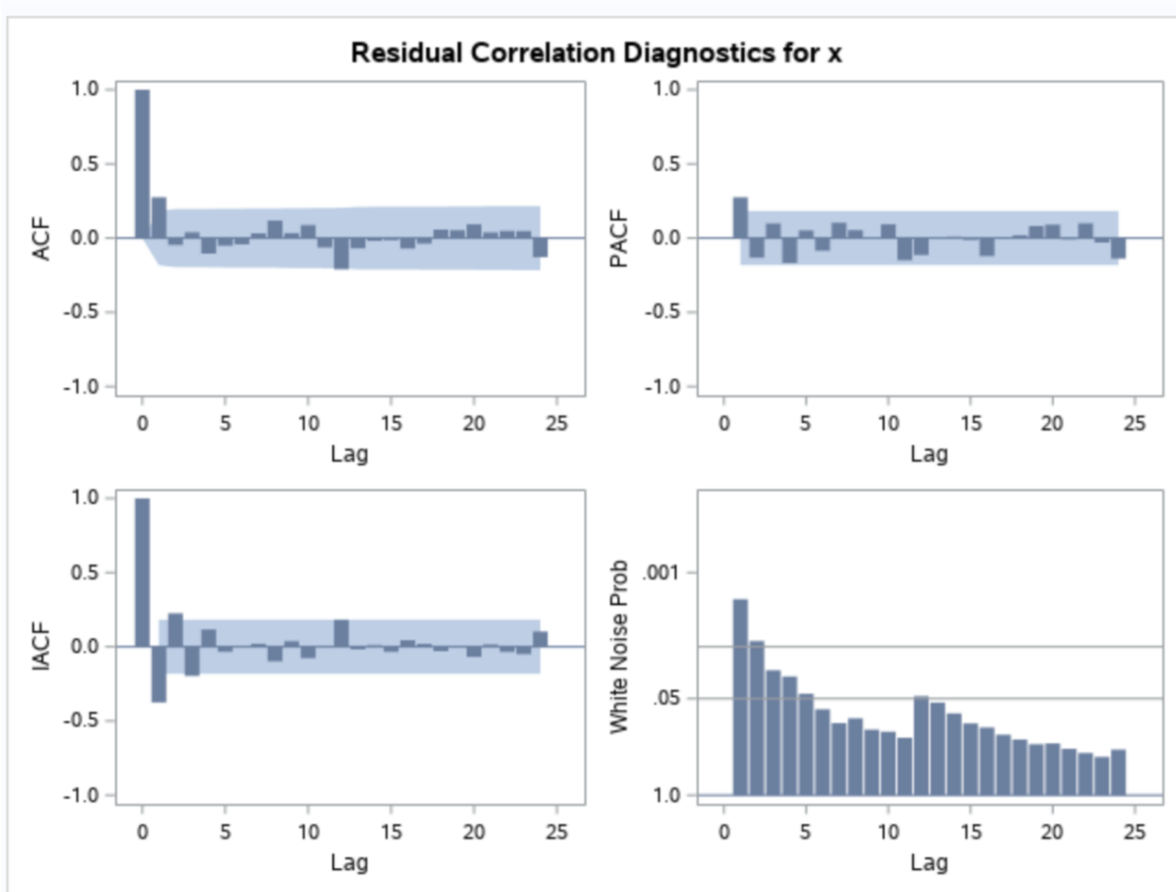
Looking at the output of **proc arima,** on the **Maximum Likelihood Estimation** table, we find that **xsin2, xsin3, xsin4, xcos4, xsin5, xcos5,** and **xcos6** have all p-values>0.05 so I removed from the model. After doing so all our estimates and trend remained significant to our model.

Continuing to the **Residual Diagnostics** we can see that both the ACF and PACF decay rapidly to zero and from the **White Noise Probability** graph we see mostly high values. This tells us that the residuals are random shocks, which is what we want.

We also see in the **Residual Normality Diagnostics** that the residuals appear fairly normally distributed, which we can tell from the distribution plot of the residual and the QQ plot.

Finally, we see from both **Time series of residuals** and the **Forecast v.s Residuals** plots appear to have a constant mean about zero, as if we drew a line through zero it would be inside our confident limit band, and a constant variance as the variability of plot remain constant about the mean.


Trend and Correlation Analysis for x
600
550
500
450
x
0
20
40
60
80
100
120
Observation
ACF
Lag
PACF
IACF
1.0
0.5
0.0
-0.5
-1.0
0
5
10
15
20
25



Residual Correlation Diagnostics for x
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0
5
10
15
20
25

Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	518.21229	2.14825	241.23	<.0001	0	x	0
MA1,1	-0.40809	0.08954	-4.56	<.0001	1	x	0
NUM1	-41.85332	2.95945	-14.14	<.0001	0	xsin1	0
NUM2	-26.80944	2.94742	-9.10	<.0001	0	xcos1	0
NUM3	3.86242	2.71224	1.42	0.1544	0	xsin2	0
NUM4	-5.43872	2.70521	-2.01	0.0444	0	xcos2	0
NUM5	-2.04006	2.33350	-0.87	0.3820	0	xsin3	0
NUM6	-5.66869	2.33316	-2.43	0.0151	0	xcos3	0
NUM7	-4.94216	1.88129	-2.63	0.0086	0	xsin4	0
NUM8	0.08467	1.88787	0.04	0.9642	0	xcos4	0
NUM9	-1.23673	1.46534	-0.84	0.3987	0	xsin5	0
NUM10	-2.57796	1.47812	-1.74	0.0811	0	xcos5	0
NUM11	-0.71432	0.91073	-0.78	0.4328	0	xcos6	0

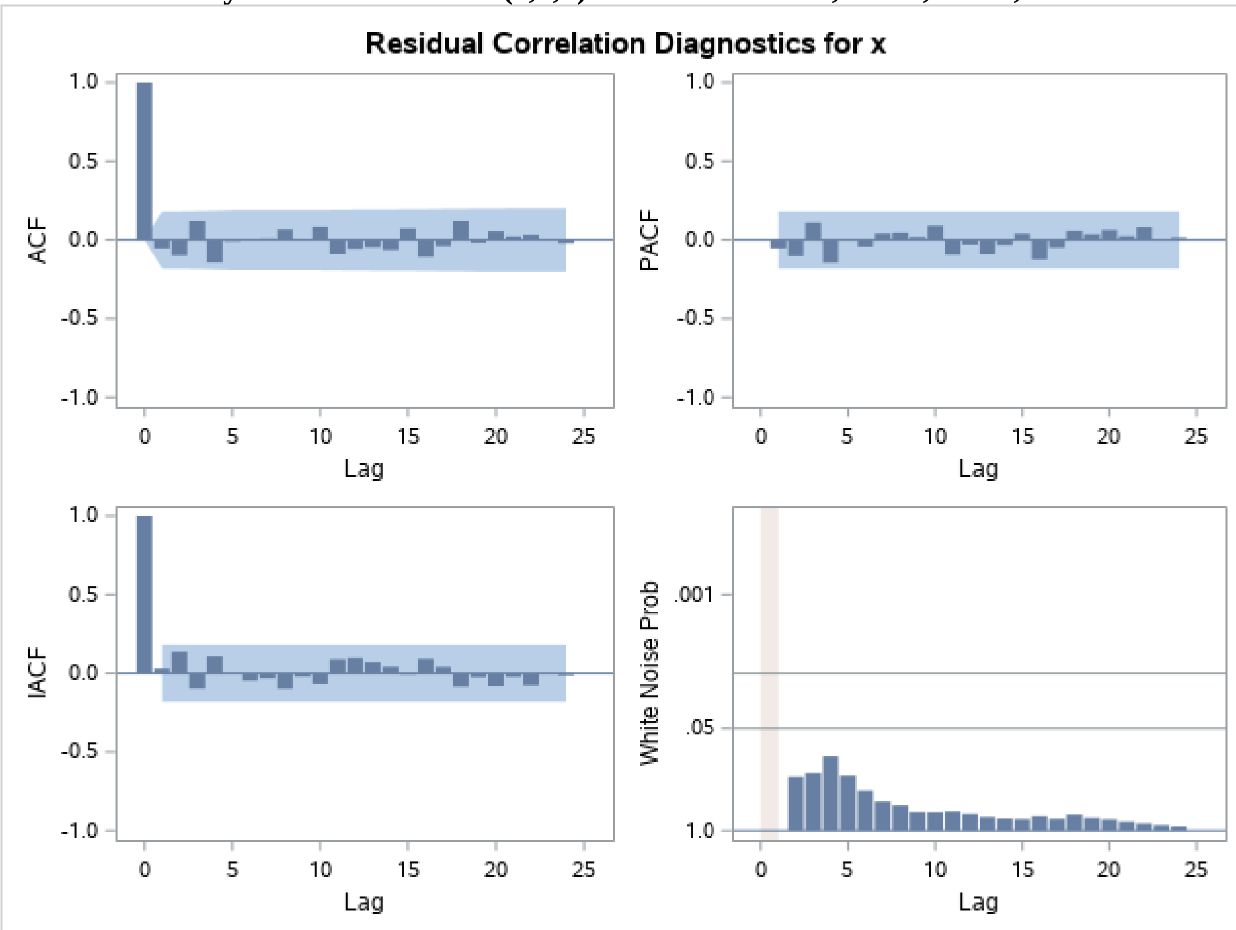
Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	518.17435	2.13005	243.27	<.0001	0	x	0
MA1,1	-0.35509	0.08857	-4.01	<.0001	1	x	0
NUM1	-41.88530	2.93910	-14.25	<.0001	0	xsin1	0
NUM2	-26.87676	2.92805	-9.18	<.0001	0	xcos1	0
NUM3	-5.48367	2.71166	-2.02	0.0431	0	xcos2	0
NUM4	-5.67974	2.36767	-2.40	0.0164	0	xcos3	0

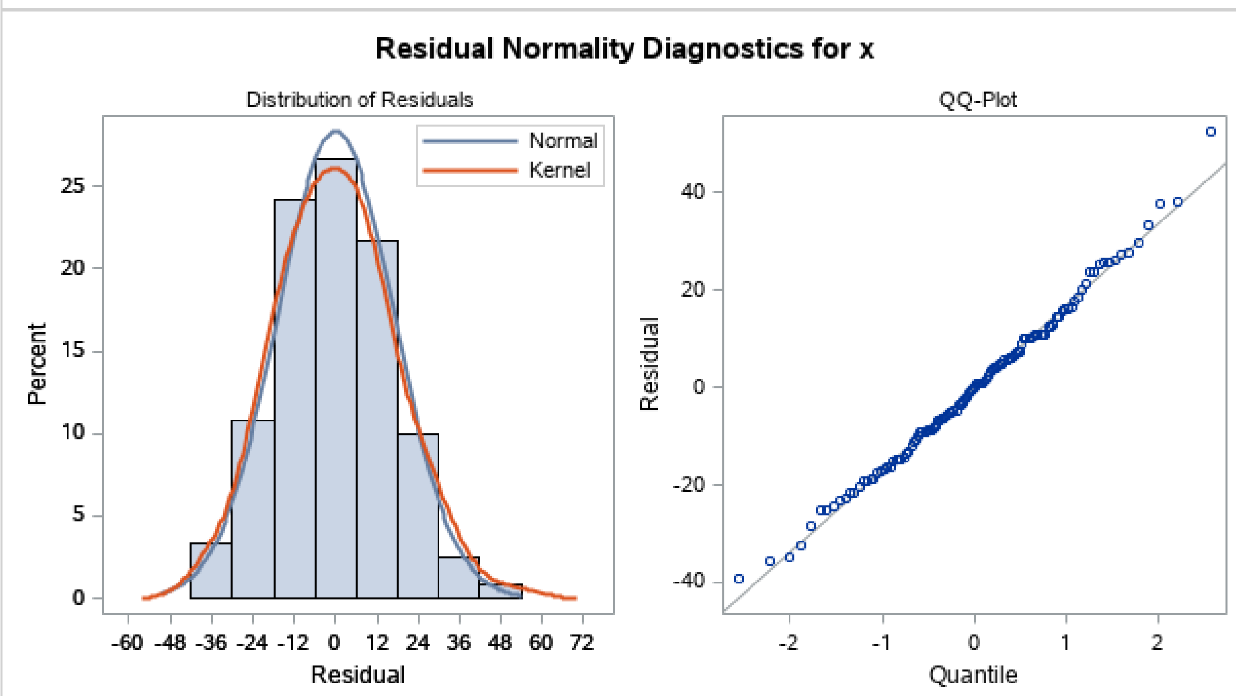
Constant Estimate	518.1743
Variance Estimate	298.0842
Std Error Estimate	17.26511
AIC	1030.21
SBC	1046.935
Number of Residuals	120

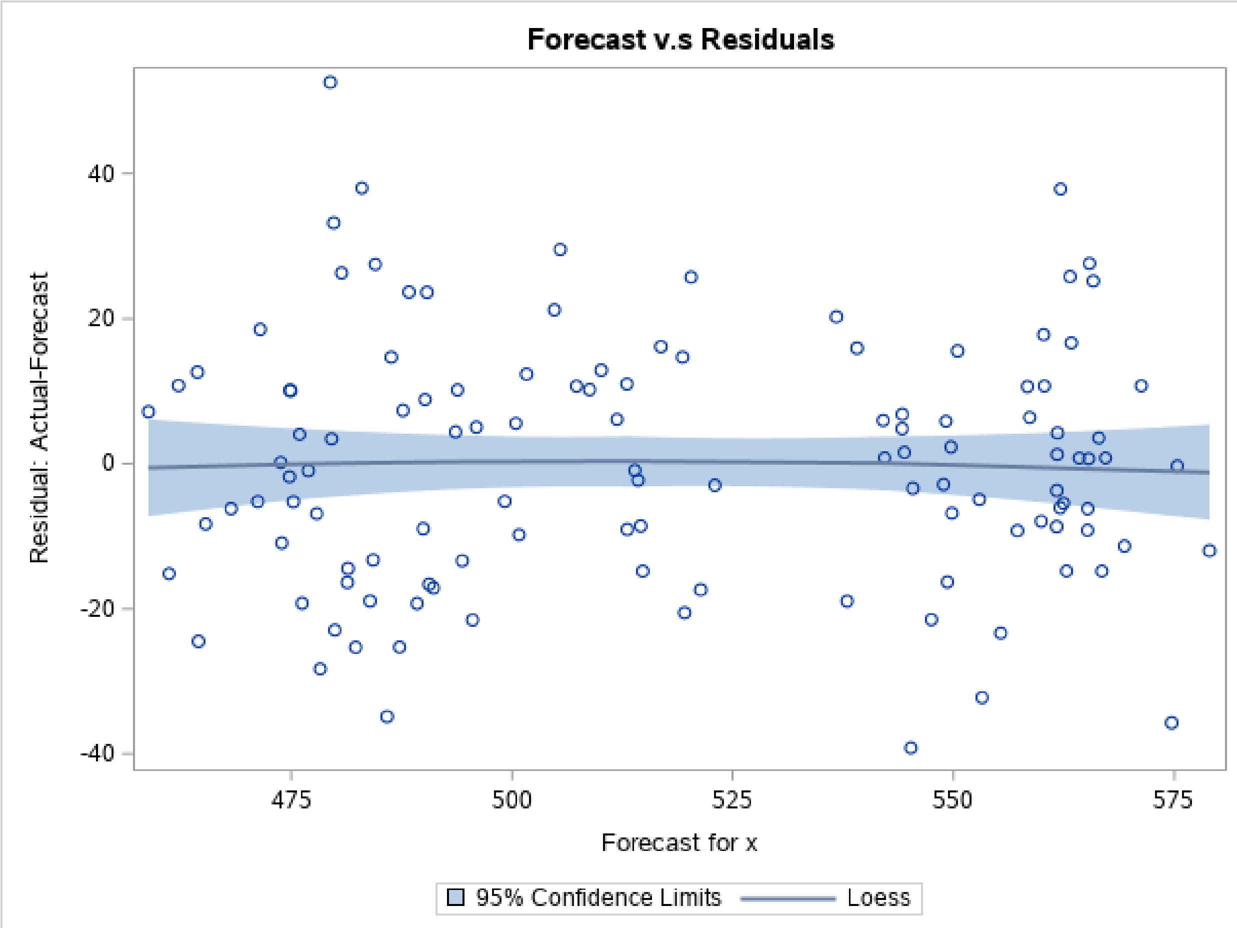
Correlations of Parameter Estimates						
Variable Parameter	x MU	x MA1,1	xsin1 NUM1	xcos1 NUM2	xcos2 NUM3	xcos3 NUM4
x MU	1.000	-0.006	-0.002	-0.006	-0.005	-0.004
x MA1,1	-0.006	1.000	-0.005	-0.008	-0.077	0.070
xsin1 NUM1	-0.002	-0.005	1.000	-0.002	-0.002	-0.002
xcos1 NUM2	-0.006	-0.008	-0.002	1.000	-0.006	-0.006
xcos2 NUM3	-0.005	-0.077	-0.002	-0.006	1.000	-0.009
xcos3 NUM4	-0.004	0.070	-0.002	-0.006	-0.009	1.000

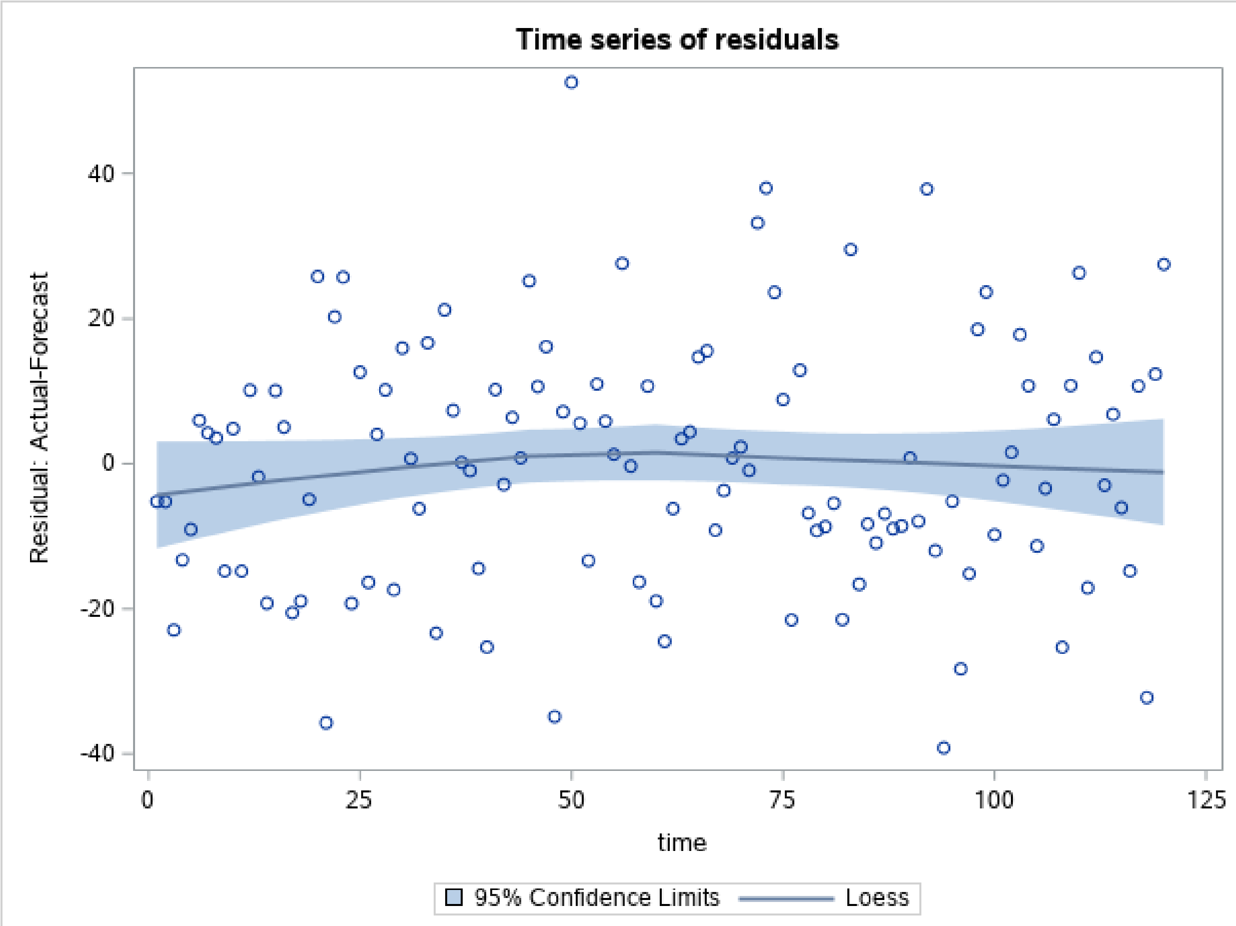
Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	5.98	5	0.3086	-0.054	-0.100	0.119	-0.143	-0.009	0.003
12	9.05	11	0.6170	0.009	0.067	0.004	0.083	-0.091	-0.057
18	14.54	17	0.6285	-0.046	-0.063	0.073	-0.108	-0.039	0.119
24	15.32	23	0.8827	-0.018	0.056	0.021	0.032	0.003	-0.021

Here we try the model ARIMA(1,0,0) with trend xsine1, xcos1, xcos2, and xcos3.






Forecast v.s Residuals
Residual: Actual-Forecast
40
20
0
-20
-40
475
500
525
550
575
Forecast for x
95% Confidence Limits
Loess

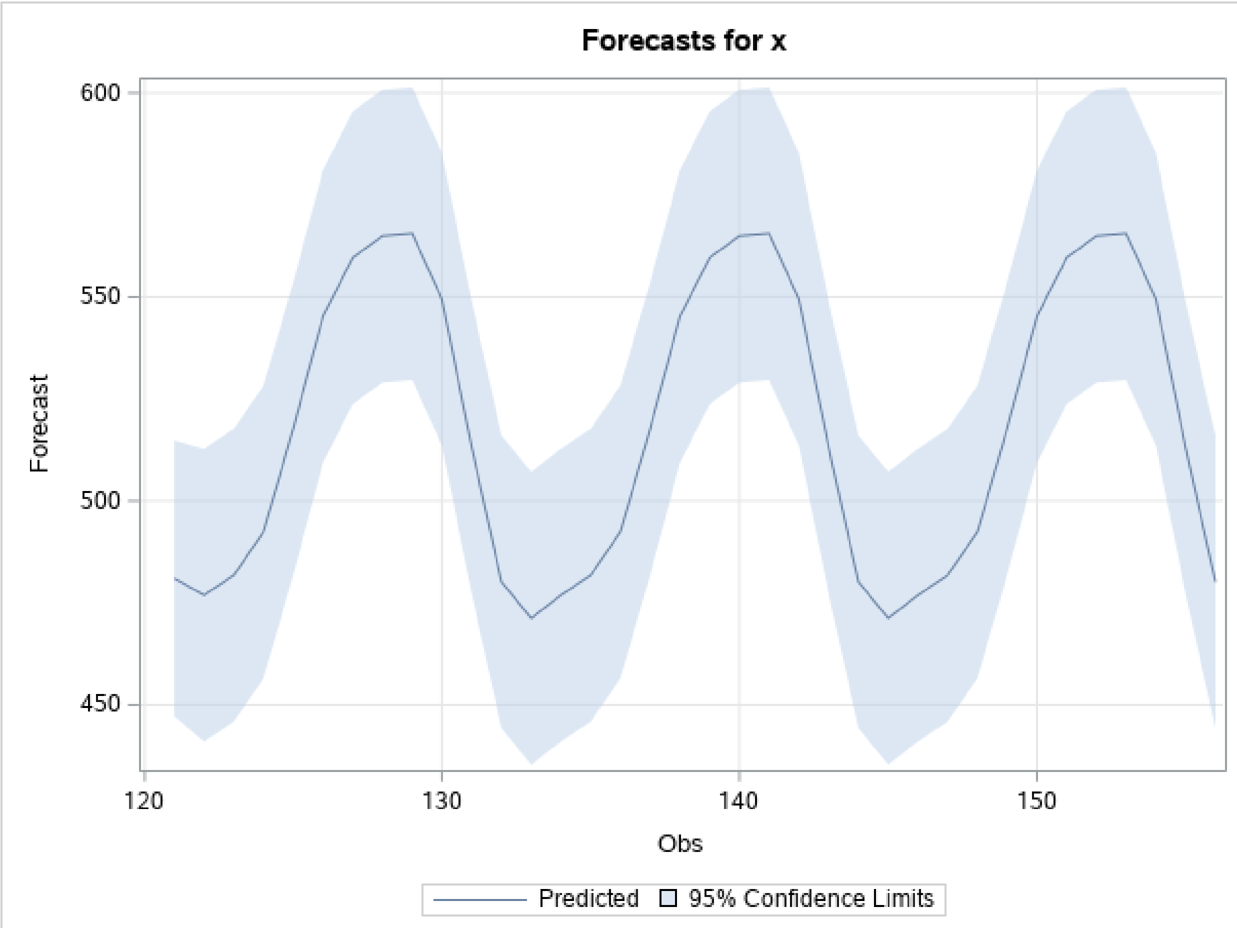


Time series of residuals
Residual: Actual-Forecast
40
20
0
-20
-40
0
25
50
75
100
125
time
95% Confidence Limits
Loess

3(b)

Since we know that we have an ARIMA(1,0,0) is a good fit to our model and can observed all the values z_t and a_t up to n, which in our case is n = 120, we can know future forecast $z_{n+k.}$
Let I_n be the information of observed all the values z_t and a_t up to n, then the future forecast will be $z_{n+k} = E(z_{n+k} \mid I_n)$.
We also know that standard error for these future forecasts will be the sqr(Var(a_{n+1})).
Finally, because we chose an MA process, we know that the confidence interval width will converge to a limiting value on the long run.

Forecasts for variable x				
Obs	**Forecast**	**Std Error**	**95% Confidence Limits**	
121	480.9724	17.2651	447.1334	514.8114
122	476.8838	18.3213	440.9747	512.7929
123	481.7727	18.3213	445.8636	517.6818
124	492.4011	18.3213	456.4920	528.3102
125	517.7658	18.3213	481.8567	553.6749
126	545.2472	18.3213	509.3381	581.1562
127	559.6511	18.3213	523.7420	595.5602
128	564.9486	18.3213	529.0395	600.8576
129	565.5433	18.3213	529.6343	601.4524
130	549.4313	18.3213	513.5222	585.3404
131	513.0992	18.3213	477.1901	549.0083
132	480.1342	18.3213	444.2251	516.0433
133	471.2139	18.3213	435.3048	507.1230
134	476.8838	18.3213	440.9747	512.7929
135	481.7727	18.3213	445.8636	517.6818
136	492.4011	18.3213	456.4920	528.3102
137	517.7658	18.3213	481.8567	553.6749
138	545.2472	18.3213	509.3381	581.1562
139	559.6511	18.3213	523.7420	595.5602
140	564.9486	18.3213	529.0395	600.8576
141	565.5433	18.3213	529.6343	601.4524
142	549.4313	18.3213	513.5222	585.3404
143	513.0992	18.3213	477.1901	549.0083
144	480.1342	18.3213	444.2251	516.0433
145	471.2139	18.3213	435.3048	507.1230
146	476.8838	18.3213	440.9747	512.7929
147	481.7727	18.3213	445.8636	517.6818
148	492.4011	18.3213	456.4920	528.3102
149	517.7658	18.3213	481.8567	553.6749
150	545.2472	18.3213	509.3381	581.1562
151	559.6511	18.3213	523.7420	595.5602
152	564.9486	18.3213	529.0395	600.8576
153	565.5433	18.3213	529.6343	601.4524
154	549.4313	18.3213	513.5222	585.3404
155	513.0992	18.3213	477.1901	549.0083
156	480.1342	18.3213	444.2251	516.0433


Forecasts for x
600
550
500
450
Forecast
120
130
140
150
Obs
Predicted
95% Confidence Limits

SAS CODE

```
/* Name: Roderick Quiel */
/* Course: STA 4853 */
/* Assignment: Homework#3 */
/* Date: 04/20/2022 */

/*Problem#1*/

/*Data Step*/
filename Repair "~/my_shared_file_links/huffer/repair.txt";
data data1;
time=_n_;
infile Repair;
input x;
y=log(x);
run;

/*Part 1a. */
proc arima data=data1;
identify var=x nlag=24;
identify var=y(1) nlag=24;
identify var=y(1,12) nlag=24;
run;

proc arima data=data1 plots=all;
identify var=y(1,12) nlag=24 noprint;
estimate q=(1)(12) method=ml;/*ARIMA(0,1,1)(0,1,1)_12*/
run;

Title "Model ARIMA(0,1,1)(0,1,1)_12 for Repair series";
proc arima data=data1;
identify var=y(1,12) nlag=24;
estimate q=(1)(12) noconstant method=ml; /*ARIMA(0,1,1)(0,1,1)_12 NOCONSTANT*/
forecast out=resids id=time lead=0 noprint;
run;


title "Forecast v.s Residuals";
proc sgplot data=resids;
loess x=forecast y=residual / clm;
run;

title "Time series of residuals";
proc sgplot data=resids;
loess x=time y=residual / clm;
```

```
run;

/* Part 1b. */
title "Model without seasonal term";
proc arima data=data1;
identify var=y(1) nlag=24 noprint;
estimate q=1 noconstant method=ml;
forecast out=resids id=time lead=0 noprint;
run;

title "Forecast v.s Residuals";
proc sgplot data=resids;
loess x=forecast y=residual / clm;
run;

title "Time series of residuals";
proc sgplot data=resids;
loess x=time y=residual / clm;
run;

/*Part 1c. */

proc arima data=data1;
identify var=y(1,12) nlag=24 noprint;
estimate q=(1)(12) noconstant method=ml noprint; /*ARIMA(0,1,1)(0,1,1)_12
NOCONSTANT*/
forecast out=resids lead=36;
run;


proc arima data=data1;
identify var=y(1,12) nlag=24 noprint;
estimate q=(1)(12) method=ml noprint; /*ARIMA(0,1,1)(0,1,1)_12 NOCONSTANT*/
forecast out=resids lead=36;
run;





/*Problem #2*/

/*Data Step*/
filename Bus "~/my_shared_file_links/huffer/bus.txt";
data data2;
time=_n_;
```

```
infile Bus;
input x;
y=log(x);
run;

/*Part 2a. */
proc arima data=data2;
identify var=x nlag=24;
identify var=x(1) nlag=24;
identify var=y(12) nlag=24;
run;

Title "Model ARIMA(2,0,0)(0,1,1)_12 for Bus series";
proc arima data=data2;
identify var=y(12) nlag=24 noprint;
estimate p=(1,2)(12) method=ml;
forecast out=resids id=time lead=0 noprint;
run;


title "Forecast v.s Residuals";
proc sgplot data=resids;
loess x=forecast y=residual / clm;
run;

title "Time series of residuals";
proc sgplot data=resids;
loess x=time y=residual / clm;
run;

/*Part 2b. */
proc arima data=data2;
identify var=y(12) nlag=24 noprint;
estimate p=(1,2)(12) method=ml noprint; /*ARIMA(0,1,1)(0,1,1)_12 NOCONSTANT*/
forecast out=resids lead=36 noprint;
run;

data original;
set resids;
keep osforecast osL95 osU95;
osforecast=exp(forecast);
osL95=exp(L95);
osU95=exp(U95);
run;

title "Future 36 Forecasts";
```

```
proc print data=original (firstobs=137);
run;




/*Problem#3*/

/*Data Step*/
filename Eureka "~/my_shared_file_links/huffer/eureka.txt";
data data3;
infile Eureka;
input x;
run;


proc arima data=data3;
identify var=x;
run;

data test;
tpi=2*arcos(-1);
drop tpi;
do time=1 to nobs+36;
if time <= nobs then set data3 nobs=nobs;
else x=. ;
xsin1=sin(tpi*time/12);
xcos1=cos(tpi*time/12);
xsin2=sin(2*tpi*time/12);
xcos2=cos(2*tpi*time/12);
xsin3=sin(3*tpi*time/12);
xcos3=cos(3*tpi*time/12);
xsin4=sin(4*tpi*time/12);
xcos4=cos(4*tpi*time/12);
xsin5=sin(5*tpi*time/12);
xcos5=cos(5*tpi*time/12);
xcos6=cos(6*tpi*time/12);
output;
end;
run;

proc arima data=test;
identify var=x crosscor=(xsin1 xcos1 xsin2 xcos2
        xsin3 xcos3 xsin4 xcos4 xsin5 xcos5 xcos6);
estimate input=(xsin1 xcos1 xsin2 xcos2
        xsin3 xcos3 xsin4 xcos4 xsin5 xcos5 xcos6) method=ml;
```

```
estimate q=1 input=(xsin1 xcos1 xsin2 xcos2
        xsin3 xcos3 xsin4 xcos4 xsin5 xcos5 xcos6) method=ml;
quit;

/*Model for series */
proc arima data=test;
identify var=x crosscor=(xsin1 xcos1 xcos2 xcos3) noprint;
estimate q=1 input=(xsin1 xcos1 xcos2 xcos3) method=ml;
forecast lead=0 id=time out=resids noprint;
quit;

title "Forecast v.s Residuals";
proc sgplot data=resids;
loess x=forecast y=residual / clm;
run;

title "Time series of residuals";
proc sgplot data=resids;
loess x=time y=residual / clm;
run;



/*Part 3b. */
proc arima data=test;
identify var=x crosscor=(xsin1 xcos1 xcos2 xcos3) noprint;
estimate q=1 input=(xsin1 xcos1 xcos2 xcos3) noprint method=ml;
forecast lead=36 out=resids;
quit;
```