Study in Consumption of Commercial Electricity in Hong Kong

Course Title: Forecasting and Applied Time Series Analysis

08-2012

1 Background

In recent years, energy shortage has become a prominent global issue. One thing is the shortage of electricity. To avoid wasting electricity and achieve efficient allocation of resources, it is important that supply and demand of electricity keep balance. Like economic development with inevitable fluctuations in long-term growth trend, electricity demand also shows a cyclical fluctuation. Therefore, we should try to accurately understand the long term electricity demand trend, find a good model to fit it and finally forecast the demand to avoid power shortage and power surplus situation happens.

As Hong Kong is an international financial center and modern international metropolis, high electricity demand for energy is not in doubt. It is particularly vital for the Hong Kong government to manage electricity supply efficiently and make appropriate plan, which means neither wasting resources due to excess supply of electricity nor hindering the development of economy in Hong Kong attributed to the shortage supply of electricity.

I just focus on the business electricity consumption to understand the regular pattern hint in history data and find future trends by establishing of effective mathematical model using time series analysis.

2 Description of the Dataset

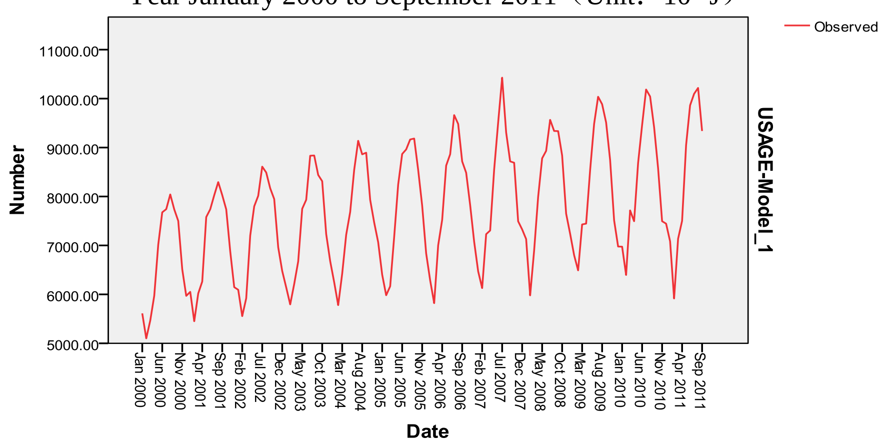
Figure 2.1 provides the data of business electricity consumption in Hong Kong from the year January 2000 to September 2011from Census and Statistics Department of Hong Kong and figure 2.2 is the line chart of it.

It can be seen from the figure 2.2, electricity consumption over time shows a slow upward trend, from 2000 January 5612 September 9338 2011. There is an obvious seasonal trend every year. It keeps high from June to September, when the temperature is very high during these months. After that, the consumption begins to drop slowly until the February in next year, when it starts to rebound.

Figure 2.1 Business Electricity Consumption in Hong Kong from the Year January 2000 to September 2011（Unit：10^{12}J）

Month / Year	1	2	3	4	5	6	7	8	9	10	11	12
2000	5612	5101	6465	5972	7016	7677	7741	8042	7737	7501	6513	5969
2001	6052	5451	6019	6262	7580	7739	8026	8294	8028	7736	6879	6147
2002	6094	5555	5921	7202	7799	8014	8610	8490	8167	7947	6962	6478
2003	6142	5795	6212	6674	7749	7932	8833	8838	8440	8312	7231	6678
2004	6247	5780	6423	7218	7684	8543	9138	8861	8895	7931	7469	7066
2005	6403	5982	6165	7173	8243	8868	8961	9165	9184	8545	7814	6837
2006	6284	5820	7002	7523	8632	8863	9663	9482	8717	8485	7823	7075
2007	6471	6128	7230	7305	8549	9510	10428	9308	8717	8690	7493	7327
2008	7129	5979	6897	7991	8780	8932	9566	9338	9336	8837	7650	7238
2009	6800	6490	7429	7445	8538	9484	10038	9885	9512	8745	7513	6977
2010	6975	6396	7716	7494	8685	9458	10186	10043	9416	8576	7494	7445
2011	7083	5916	7136	7502	9048	9862	10097	10218	9338			

Figure 2.2 Line Chart of Business Electricity Consumption in Hong Kong from the Year January 2000 to September 2011（Unit：10^{12}J）



3 Model Fitting and Diagnostics

In this part, we try to use five methods to analyze the data, including simple exponential smoothing, winter's multiplicative, additive decomposition, multiplicative decomposition and ARIMA model.

3.1 Simple Exponential Smoothing

The simple exponential smoothing method is usually used to fit data on the condition that the average of it is not change dramatically. It supposes that the more recent data has more effects on the current and do the fitting process by continually revising an estimate or forecast by accounting for more recent changes or for fluctuations in the data. The generally form of it is $x_t = b + \varepsilon_t$, where ε_t is the random component having mean 0 and variance σ_a^2. Since the average electricity consumption is changing very slowly over time, we can try to use the simple exponential smoothing method to deal with the data. After using the Simple Exponential Smoothing function of the SPSS, we get the following results.

Figure 3.1.1 Hypothesis Testing Results of Parameter α

	Estimate	SE	t	Sig.
Alpha (Level)	1.000	.085	11.786	.000

Figure 3.1.2 Hypothesis Testing Results of significance of model

Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
RMSE	Statistics	DF	Sig.	
685.763	504.183	17	.000	0

From figure3.1.1, the best value for parameter α is 1, and according to the hypothesis testing, the value is significant, which means the simple exponential smoothing model should be $x_{t+1} = x_t$. Actually it is na ïve model. It shows that the consumption in this month should be predicted as the same the former month usage. From figure 3.1.2, however, the significance of Ljung-Box Q(18), the hypothesis testing statistic of ε_t, is 0, which means ε_t is not with mean 0 and variance σ_a^2. It indicates that the Simple Exponential Smoothing model is not suitable for fitting this time series. Probably because the seasonal effects are very strong compared with the linear trend, and the data fluctuates dramatically from it average value.

3.2 Winter’s Seasonal Multiplicative

From the discussion above, we know it is inappropriate to ignore the seasonal characteristic of the data and, simply base on the linear trend to build the model. As showed in the figure, the data is a seasonal data and as the average level of the series increases, the amplitude of the seasonal pattern also increases. We can consider using the winter's seasonal multiplicative method. Winter's multiplicative method is an exponential smoothing approach to deal with cyclical data and its pattern can be described by $x_t = (\alpha + \beta t)c_t + \varepsilon_t$, where α is the permanent component, β is a linear trend component, c_t is the multiplicative seasonal factor and ε_t is the usual random error component. By using the SPSS with the data, we get the following figures.

Figure 3.2.1 Hypothesis Testing Results of Parameter

Model		Estimate	SE	t	Sig.
USAGE-Model_1	Alpha (Level)	.075	.030	2.512	.013
	Gamma (Trend)	.001	.002	.218	.828
	Delta (Season)	.176	.071	2.486	.014

Figure 3.2.2 Hypothesis Testing Results of significance of model

Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
RMSE	Normalized BIC	Statistics	DF	Sig.	
253.818	11.179	21.743	15	.115	0

The t statistics of α and c_t are 2.512 and 2.468, which conforms that they are both significant compared with the t statistics of β (0.218) that is not significant. The value of α and c_t in this model are 0.075 and 0.176. Moreover, the hypothesis testing of ε_t (sig=0.115>0.05) indicates that ε_t is with constant mean and variance. The RMSE of using this model is 253.818.

3.3 Additive Model

We have used the winter's multiplicative smoothing model to handle the data based on the assumption that the seasonal and trend components are varying over time. From the figure the data changes seasonally over time and maybe the seasonal components are constant. When suppose that the seasonal patterns are stable year after year, we consider the decomposition approach, including additive model and multiplicative model. The additive is consist of the sum of four parts-trend component, seasonal variation, cyclical component and random error component, which can be generally described as following equation, $Y_t = Tr_t + S_t + C_t + \varepsilon_t$, Tr_t is trend component, S_t is seasonal variation, C_t is cyclical component and ε_t is random error component with constant 0 mean and variance. Compared with additive model, the multiplicative model is also composed of the same four parts by multiplying them instead of adding them up, which generally form is $Y_t = Tr_t \times S_t \times C_t \times \varepsilon_t$. Usually we treat the cyclical component as the regular fluctuation in the trend component, however, since it takes 2 to 10 years to complete a cycle.

Firstly, we use the additive model. By the function of Decomposition in SPSS, we can get the 12 seasonal factors S_t in figure3.3.1, since the cycle is 12 predicted from the time plot.

3.3.1 Seasonal Factor

t	1	2	3	4	5	6	7	8	9	10	11	12
S_t	-1244.	-1842.	-1051	-523	463	963	1515	1375	1038	586	-378	-903

For the trend part Tr_t, we use the seasonal adjusted data to do the linear regression and the relevant results are showed in the figure3.3.2. From the table, we know the fitting equation is $Tr_t = 6835.24 + 12.576t$. Since the Significance of the equation is almost 0 which is less than 0.05, we reject the hypothesis that the equation is not significant, which means the fitting equation is appropriate and the linear trend is suitable. The figure3.3.3 gives the results of the coefficient hypothesis testing. The significance value of both constant and coefficient of t are 0 which is less than 0.05. It confirms that the constant and coefficient of t can not be ignored. They are both nonzero.

By combining the above two points, we can get the fitting equation of the additive model, $Y_t = 6835.24 + 12.576t + S_t + \varepsilon_t$. The RMSE of the additive model is 267.87.

3.3.2 Results of Hypothesis Testing of Trend Equation

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36946131.156	1	36946131.156	507.591	.000[a]
	Residual	10117413.753	139	72787.149		
	Total	47063544.910	140			

3.3.3 Results of Hypothesis Testing of Constant and Coefficient of t in Trend Equation

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6835.240	45.684		149.621	.000
	t	12.576	.558	.886	22.530	.000

3.4 Multiplicative Model

Now, we consider the multiplicative decomposition model. The ordinary form for this model is $Y_t = Tr_t \times S_t \times C_t \times \varepsilon_t$, Also we neglect the cyclical component and just focus on the simplified multiplicative decomposition $Y_t = Tr_t \times S_t \times \varepsilon_t$. Through the SPSS, we get the results showed in the following figures 3.4.1 and 3.4.2.

3.4.1 Seasonal Factor

t	1	2	3	4	5	6	7	8	9	10	11	12
S_t (%)	84	76.6	86.3	92.9	10.1	112.2	119.1	117.9	113.7	107.6	95.2	88.3

3.4.2 Results of Hypothesis Testing of Trend Equation

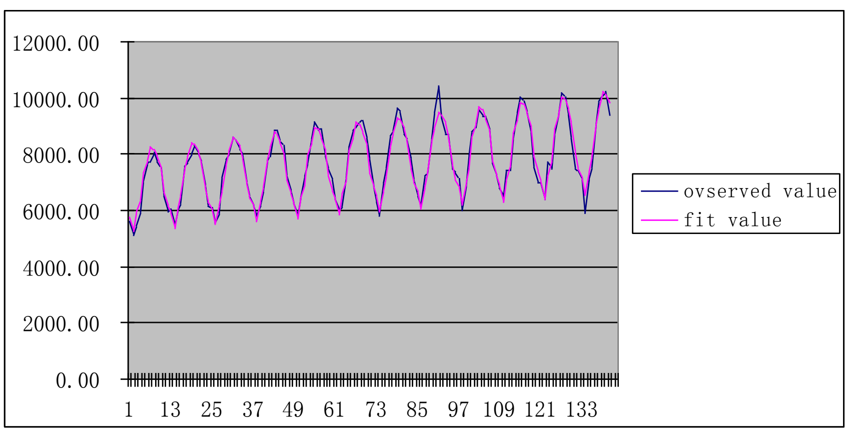
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36379474.457	1	36379474.457	545.062	.000[a]
	Residual	9277375.836	139	66743.711		
	Total	45656850.293	140			

3.4.3 Results of Hypothesis Testing of Constant and Coefficient of t in Trend Equation

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6837.166	43.746		156.292	.000
	t	12.480	.535	.893	23.347	.000

Figure 3.4.1 gives the estimated 12 seasonal factors S_t. From figure3.4.2 and 3.4.3, we know the trend part Tr_t can be fitted as $Tr_t = 6837.166 + 12.48t$ base on the linear regression of seasonal adjusted data. For the significance hypothesis testing of the trend equation, the Sig is approximately equal to 0 which is less than 0.05. It indicates that the trend equation is significant. For the significance hypothesis testing of the constant and coefficient of t in the equation, the Sig for both is almost 0 less than 0.05, which shows that they are significant and can not be ignored. Then after the two testing we can get the significant trend equation $Tr_t = 6837.166 + 12.48t$. Finally, write down the fitting equation, $Y_t = (6837.166 + 12.48t) \times S_t \times \varepsilon_t$. Substitute t and corresponding S_t, we can get the fit value for every month and it is showed in figure 3.4.5. The RMSE of the multiplicative decomposition is 251.62.

Figure 3.4.5 Line Chart of Fitted Value and Observed Value



3.5 ARIMA Model

ARIMA model is a very efficient approach to process data with trend and seasonal factor cyclical feature. When the process is stationary, we can directly use the ARMA model to handle it. While it is not stationary, we have some ways to transform the data to make it stationary and then apply the model to the new stationary data. The methods used for making the data be stationary usually are ordinary difference and seasonal difference. If the time plot of data shows an upward trend, it is obviously not stationary, since the mean is not constant, under which condition we can consider doing one time difference on the original data. Moreover, if the data has the seasonal feature, we can try to do one time seasonal difference to make it stationary. So the general form of ARIMA model for nonstationary data is $\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)a_t$, where $\phi_p(B)$, $\theta_q(B)$ are polynomials with degree p and q in terms of B, B is the backward operator, a_t is the random error. For other typical data, there exists an obviously cycle. In this situation, we should consider the mixed ARIMA model, which can be presented as $\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \Theta_Q(B^s)\theta_q(B)a_t$, where $\phi_p(B)$, $\theta_q(B)$ $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are polynomials with degree p, q, Ps and Qs, respectively, B is the backward operator, a_t is the random error. The key important thing for ARIMA model is to find the best parameters p,d,q,P,D,Q. The rule used for judge the best parameters is selecting the minimum BIC among the several possible models.

Firstly, we get the ACF and PACF of original series in figure 3.5.1 and 3.5.2. Thus, we know the time series is not stationary, since it displays an upward trend and seasonal feature and strongly cyclical (according to the ACF plot). We do an ordinary difference and a seasonal difference, and plot the time plot of the new time series ACF and PACF showed in figure 3.5.3 and 3.5.4.

Figure3.5.1 ACF of Original Data

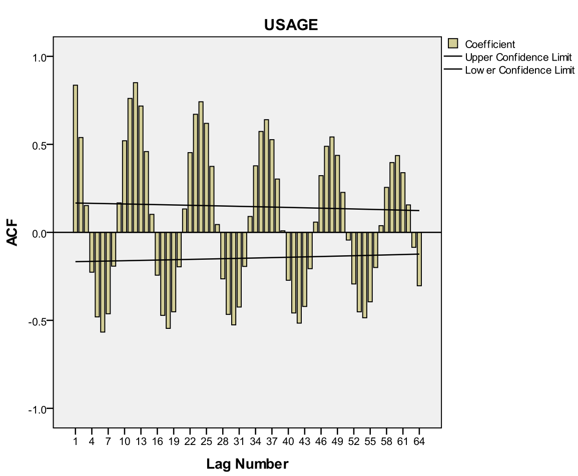
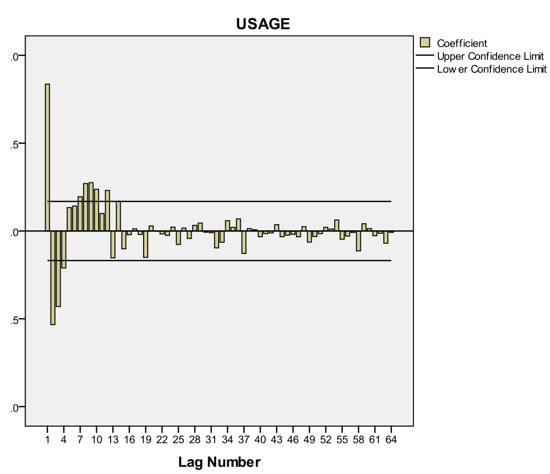
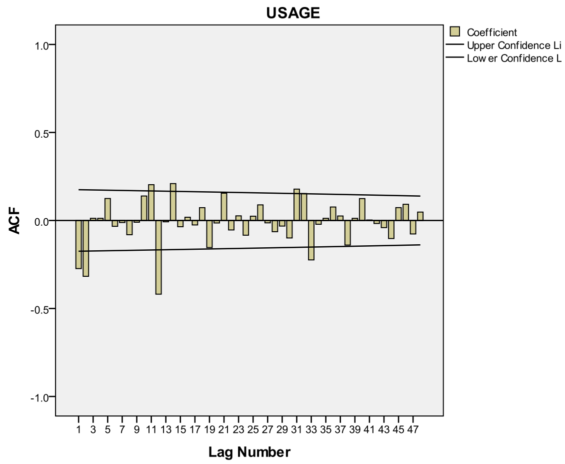
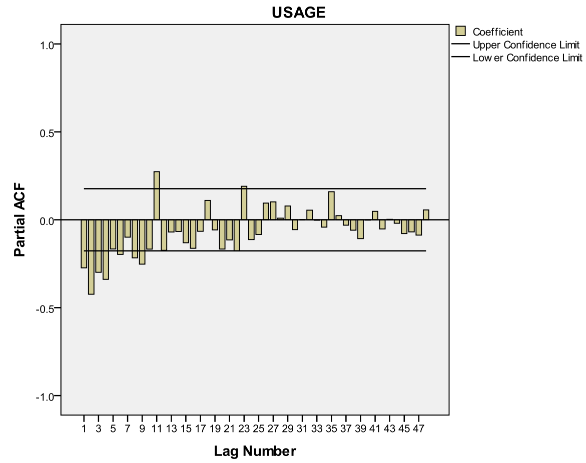


Figure3.5.1 PACF of Original Data







According to the new ACF and PACF, it is clear that the new time series is stationary. Therefore, it is reasonable to determine $d = 1, D = 1$ and $p \leq 4, q \leq 2, Q = 1, P \leq 1$, and consider begin with the model $ARIMA(4,1,2)\times(1,1,1)_{12}$. Through changing the value of p, q, P, Q in the model, we select the eligible models which should meet the condition that its fit residual is white noise, and model parameters are significant. After several trials, we get three eligible models, and then do a square root and a natural log transformation, respectively and summarize the model in the figure 3.5.5 below.

Figure 3.5.5 Summarize of Several Appropriate Model

ARIMA	$(2,1,1)\times(0,1,1)_{12}$	$(2,1,2)\times(0,1,1)_{12}$	$(1,1,2)\times(0,1,1)_{12}$	$(2,1,1)\times(1,1,0)_{12}$
BIC	11.379	11.388	11.378	11.449
(SR)BIC*	11.362	11.369	11.371	11.455
(LOG)BIC*	11.377	11.377	11.390	11.483

*(SR)BIC and (Log)BIC represent the BIC of ARIMA model with original data transformed by square root and natural log, respectively.

From the table, we select the model $ARIMA(2,1,1)\times(0,1,1)_{12}$ with square root transformation of its original data, whose BIC 11.362 is least compared with all other model. Then display the hypothesis testing results of residual significance, parameters significance and residual ACF and PACF in figure 3.5.6, 3.5.7 and 3.5.8.

Figure 3.5.6 Hypothesis Testing Results of significance of model

Model Fit statistics		Ljung-Box Q(18)			Number of Outliers
RMSE	Normalized BIC	Statistics	DF	Sig.	
271.877	11.362	14.992	14	.379	0

Figure 3.5.7 Hypothesis testing Results of parameters of ARIMA model

				Estimate	SE	t	Sig.
USAGE-Model_1	Square Root	AR	Lag 1	.113	.095	1.190	.236
			Lag 2	-.243	.094	-2.596	.011
		Difference		1			
		MA	Lag 1	.903	.053	17.100	.000
		Seasonal Difference		1			
		MA, Seasonal	Lag 1	.585	.095	6.134	.000

From figure3.5.6, significance of Ljung-Box Q(18) statistic is 0.379 greater than 0.05, which means we should accept the hypothesis that the transformed time series is stationary. From figure 3.5.7, we can get the significant value of the model parameters. Finally, we can determine the fitted equation of $ARIMA(2,1,1)\times(0,1,1)_{12}$,

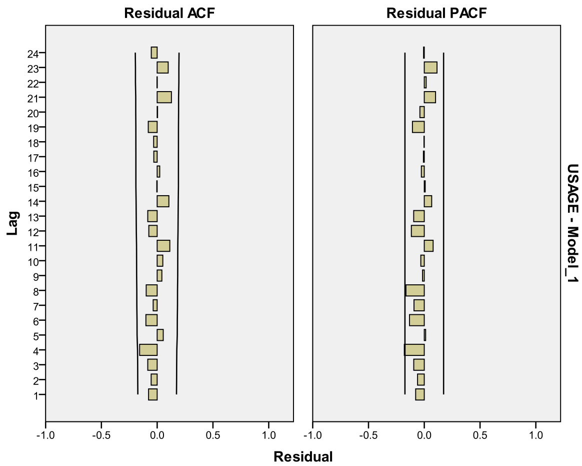
$$\left(1+0.243B^2\right)\left(1-B\right)\left(1-B^{12}\right)\sqrt{Z_t}=(1-0.585B^{12})\left(1-0.903B\right)a_t,$$

That is

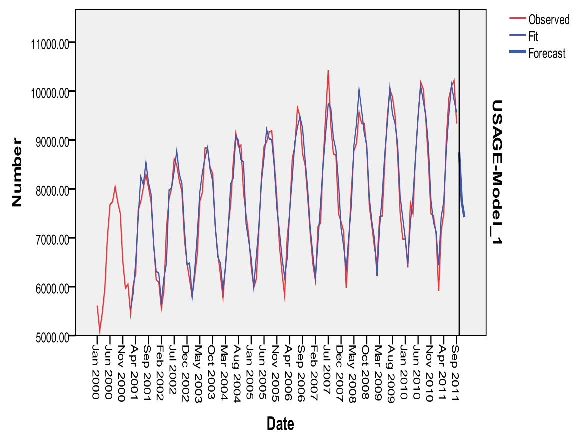
$$(1-B+0.243B^2-0.243B^3-B^{12}+B^{13}-0.243B^{14}+0.243B^{15})\sqrt{Z_t}=(1-0.903B-0.585B^{12}+0.528B^{13})a_t$$

The figure 3.5.9 shows the fit value and observed value under this optimum model. And the RMSE of this model is 271.88. We use the line chart to draw the observed value and fitted value in the figure in the figure 3.5.9.

3.5.8 Redidual ACF and PACF of Preferred Model



3.5.9 Observed Value and Fitted Value



4 Selection of the Best Model

From discussion above, we used five models to fit usage of commercial electricity over the January 2000 to September 2011. Now we choose the most preferred model from the efficient four acquired models based on the minimum RMSE criteria. According to figure 4.1 below, multiplicative decomposition has the least minimum RMSE 251.62 compared with the other three models. We can conclude that multiplicative decomposition is the best model for fitting the practical time series.

Figure 4.1 Comparisons of Four Models' RMSE

MODEL	Winter' Multiplicative	Additive Decomposition	Multiplicative Decomposition	$ARIMA(2,1,1)\times(0,1,1)_{12}$
RMSE	253.82	267.87	251.62	271.88

5 Summary

In this essay, we first use the simple exponential model to deal the data and find that it is not suitable, since this model ignores the seasonal effect. Then we use the winter's multiplicative, additive decomposition, multiplicative decomposition and ARIMA model, all of which consider the seasonal impact. Finally, based on the minimum RMSE criteria, we choose the multiplicative decomposition model as the best model. Its expression is

$$Y_t = (6837.166 + 12.48t) \times S_t \times \varepsilon_t$$

S_t represents the seasonal variation and ε_t is random error component.

From the expression we know the fitted usage of electricity is determined by the product of two parts-linear trend and seasonal component. The trend part is an increasing function in terms of time t, which indicates that with the increasing of t, the business electricity consumption will increase as well. When considering effect of the seasonal part executed on the consumption, it is clear that the appearance of the plot will like a wave. So integrating the two factors, we can conclude that the usage will keeping on an upward trend with fluctuation in the near future.

When comparing and contrasting the fitted value to the observed value from the figure 3.4.5, we find that it fits pretty well except one point-January 2008. The reason for the difference on January 2008 is the impact of continued extreme cold weather, compared with the same month in previous, began from 24th January. Since the lasting colder weather, people need to consume much more electricity to keep warm and comfortable, which leads the actual usage is larger than the fitted value.

So in short, the multiple composition model can depict the data efficiently, and has accurate prediction in most cases. The bias between fitted value and actual value appears when some unusual events happen like the continuous abnormal weather. When the government makes the power plan for the following month or year, they can apply this model to forecast the future electricity consumption first.

6 Data Resource

The data of business electricity consumption in Hong Kong from the year January 2000 to September 2011analyzed in this essay can be downloaded from the website of Census and Statistics Department of Hong Kong as following, http://www.censtatd.gov.hk/hong_kong_statistics/statistical_tables/index_tc.jsp?charsetID=2