Modeling and Forecasting Hourly Wind Power Production in Sweden with Time Series Models

Xiangli Meng

Supervisor: Changli He

D-level in Statistics, May 2010

School of Economics and Social Sciences

Högskolan Dalarna, Sweden

Contents

Abstract

In this paper, we model and forecast the hourly wind power production in Sweden with time series models, separately for cold season (from December 1, 2008 to February 28, 2009) and warm season (from June 1, 2009 to August 31, 2009). The seasonality feature in this data is analyzed by spectral analysis and seasonal ARIMA (SARIMA) models are considered for the hourly wind power production data. We extend a HEGY-type seasonal unit roots test to a 24-order autoregressive model and apply it to wind power production data. Testing procedures for our tests are proposed and the distributions of the tests are obtained by simulations. The data reject the presence of unit roots at most of the seasonal frequencies. At last, we use the estimated SARIMA model and ARAR algorithm to forecast the production in cold season and warm season of 2009. For the warm season, the ARAR algorithm outperforms SARIMA, while for the cold season, the ARAR algorithm and SARIMA model have similar forecasting results.

Key Words:

Seasonality, sample periodogram, seasonal unit root, HEGY test, Seasonal ARIMA, short-memory, ARAR algorithm,

1 Introduction

Wind is one of the important renewable resources on the earth. Nowadays, this kind of resources is highly concerned because of their crucial role to the human development. When wind power is collected from reproduced wind resources, there is little additional pollution produced and less money cost. It is an economical energy and is commercially viable. In Sweden, wind is even more applicable than other widely applied renewable energy in many countries, such as hydropower. Since, as you may know, Sweden has a long winter every year, thus, hydropower has to be disturbed in almost half-time a year. Actually, Sweden uses it much in summer, but this is really a waste of hydropower constructions and it is also difficult to take care of the constructions during the long winter period. Furthermore, hydropower has a well-known disadvantage of the limitation to expand its output due to the protection of rivers. As an alternative power system, district heating has been traditionally applied in Sweden for a long time. Although most of the Sweden district heating systems use renewable resources, the advantages of wind resources are still obvious, for example, it is very expensive for the pre-construction of the district heating systems. Wind power does not suffer from those problems, and fortunately Sweden has considerable potential output of wind power by its long coast line and plenty of mountains. That is why we reasonably expect wind power to be one of the main power resources for electricity generation in the future. More details are in Swedish Wind Energy (2010).

Due to the fluctuated behavior of wind power productions, the output predictions of the productions are really helpful to increase the efficiency of wind power. Because “capacity reserve” of a certain amount of electricity will be needed to compensate for the fluctuation if there are some unbalance between the electricity production and consumption. But the capacity reserves raise the economic and environment cost of electricity production because most capacity reserves are generated by conventional power plant. Therefore, the predictions, which are delivered to the Transmission System Operators (TSO) to suggest the electricity system’s adjustment for the reserves, can reduce the cost of the short-term capacity reservations and therefore, make wind power more valuable. The predictions could also be used for other relevant purposes such as generation and transmission maintenance planning, economic dispatch, energy storage optimization and energy trading. See Jursa and Rohrig (2008).

Usually, there are two directions to study wind power and to do the forecasts: physical approach and statistical method. Sometimes, a combination of those two models is also studied. See International Energy Agency (2005) for more discussion. The physical models use the numerical weather prediction (NWP) data as explanatory input variables, while the statistical models use the past values of time series data. Obviously, the physical models have a disadvantage that it is sensitive to the weather of each wind farm. That means, it needs all the NWP data of every included wind farm since different locations of each wind farm have different weathers. That raises several problems such as computational efforts and economic costs. Even more, the problems become more serious when it comes to forecast the production for a whole country. However, the statistical models can ignore those problems since statistical model is based on the real data. See more in Lange and Focken (2005).

The main focus of this paper is on modeling and forecasting the hourly wind power production of whole Sweden by time series models. Our hourly data represent a significant seasonality based on spectral analysis. But there have been few studies of seasonal unit root test in a 24-order autoregressive polynomial model. We extend the seasonal unit root tests and the algorithm in Hylleberg (1994) for order-4 and Franses (1990) for order-12 autoregressive polynomial to be available for our data set of order-24 polynomial. The distributions of the proposed test statistics are studied. Based on the results of the seasonal unit root tests, we do the modeling and evaluation with the data. And thereafter we compare the forecasting results of estimated Seasonal ARIMA (SARIMA) model with that of ARAR algorithm, and find that ARAR algorithm outperforms the SARIMA model especially in warm season in 2009.

The remaining organization in this paper is as follows: the 2nd section presents the wind condition and pre-analysis of the hourly wind power production data in

Sweden, the 3rd section introduces the statistical methodologies used in our modeling and forecasting procedure, the 4th section derives a HEGY type seasonal unit root test for hourly data, the finite-sample and asymptotic distributions are obtained by simulations, the 5th section is the pratical results of our modeling and forecasting procedure on wind power production data, finally, the 6th section provides conclusion and future works.

2 Hourly Wind Power Production Data

In this section, we introduce the wind condition in Sweden and the pre-analysis of our hourly wind power production data.

The data set used in this paper is the total hourly wind power production data in Sweden, starting from January 1 of 2008 and ending on September 30, 2009. The data set is provided by Svensk Energi, the organization that represents companies involved in the production, distribution and trading of electricity in Sweden.

2.1 Wind condition in Sweden

The wind power production is highly affected by the wind conditions. This subsection analyzes the wind condition in Sweden with the wind power production data. We plot the data from January 1, 2008 to December 31, 2008 in Figure 1 and we plot the data in different months during this period by boxplots in Figure 2.

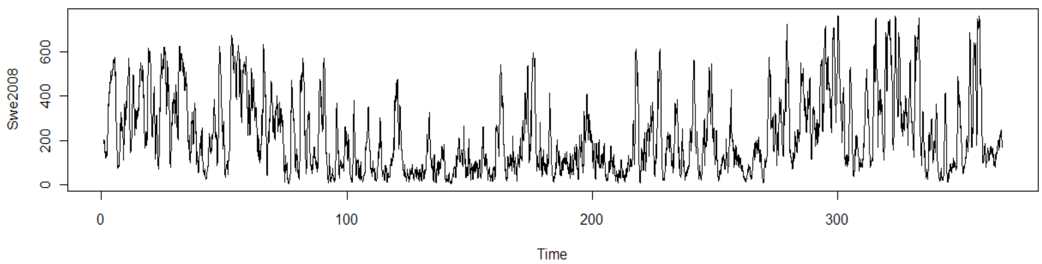

Figure 1: Plot of the data in 2008

From Figure 1 and Figure 2, it can be seen that the data from April to September have smaller median values and smaller inter-quartile range (IQR, difference between upper and lower quartile) than the remaining months, and there are many outliers (data outside 1.5 times IQR of the upper quartile) in the month from April to September and December.

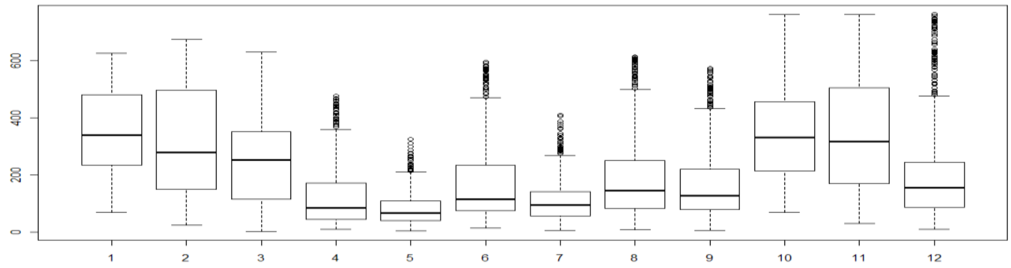

Figure 2 Box plots of data in 12 months in 2008

Based on the box plot, it is informative to split the whole year into warm season and cold season. The time from April to September are treated as warm season and the rest as cold season. The wind conditions of the two seasons are different. The wind is caused by different temperatures of adjacent places. Larger gaps lead to stronger wind. In Sweden, there are two important kinds of temperature gaps in wind generation. The heat capacity of land is smaller than that of sea, so the land is heated faster than the sea. This causes the wind in latitude direction. The different temperatures between subtropics and arctic cause the wind in longitude direction. In warm season, the temperature gap between land and sea is larger than the gap between subtropics and arctic, and therefore the former is the main gap in wind generation. In cold season the condition is opposite, the gap between subtropics and arctic plays a more important role. The changing wind conditions result in different patterns of wind power productions.

2.2 Pre-analysis of the data

For each season, a period of 3 months is chosen for analysis in convenience. For cold season, the data used are from December 1, 2008 to February 28, 2009. For warm season, the data used are from the period June 1 to August 31, 2009. The data from the period February 1 to 29, 2009 and from August 1 to 31, 2009 are plotted in Figure 3. From the plot it can be seen that there are big fluctuations and many waves in both series. Some waves have long periods and some waves have short periods.

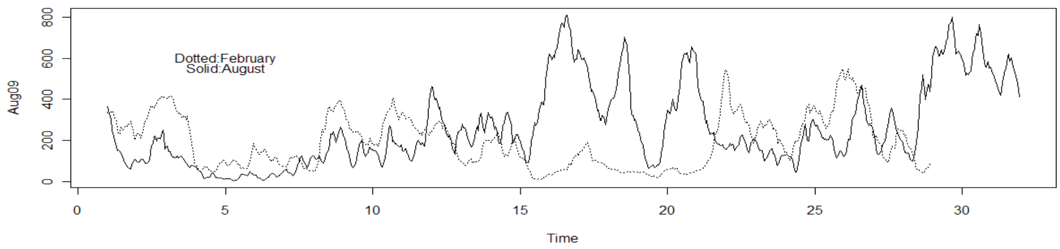

Figure 3: The data plot of February and August in 2009

For the 3 months period in each season defined above, the mean values and variances for the data of each month are given in Table 1. The average means and variances of cold season are higher than those of warm season. For cold season, the mean production in January is much larger than the other two months and it is the largest in all the 6 months. The variance of January is also the largest. The production in February has the smallest mean value and variance. In warm season, the production in August has the largest mean value and variance, but the difference between the largest value and other two values are not as big as that in cold season.

Table1: means and variances for each month.

Cold season					Warm season			
Month	Dec 2008	Jan 2009	Feb 2009	Total	Jun 2009	Jul 2009	Aug 2009	Total
Mean	207.1	289.5	192.1	251.3	236.0	225.3	261.1	240.8
Variance	30461.8	51823.	16512.7	36455.5	27364.6	27301.7	38379.5	31254.5

The mean values of the 24 hours for two seasons are plotted in Figure 4 to show the production in different hours. From the plot, it can be seen that the mean values for 24 hours in warm season has much bigger fluctuation than that in cold season. In cold season, the production at 10 o'clock is the lowest and the production after 15 o'clock is a little higher than the others. In warm season, the differences of the production in different hours are bigger than that in cold season. The production is the lowest at 8 o'clock and the highest at 15 o'clock.

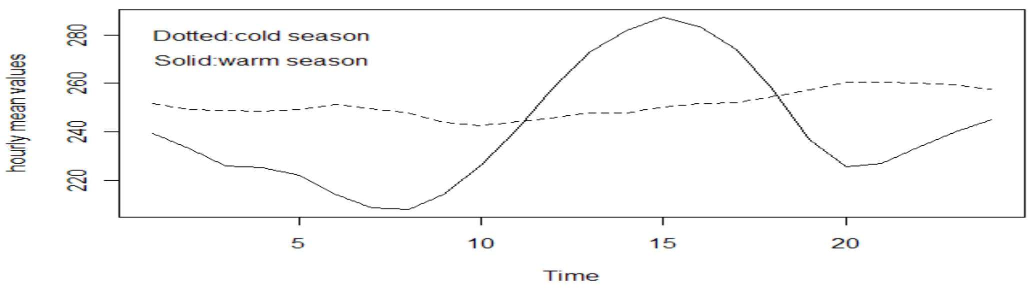

Figure 4: hourly mean values of cold season and warm season

3 Statistical Methodologies

In this section, we introduce several statistical methods which are applied or extended in our procedure of analyzing the hourly wind power production data. The modeling and forecasting procedure for our data is as follows: Spectral analysis is used to analyze the seasonality of the data and to find if periodic patterns exist in the series; a seasonal unit root test is introduced to examines the presence of unit roots; SARIMA model is used to fit the data; ARAR algorithm is considered as an alternative forecasting method at last.

3.1 Spectral analysis

When time series are observed by quarter, month or hour, the data often display periodically feature which we called seasonal pattern or seasonality. The spectral analysis is a useful tool to extract seasonal components and periodicities which are obscure in a given series.

I Definition and interpretation. The sample periodogram is used to infer the portion of variance for series that could be attributed to the circles with frequency ω. The definition of sample periodogram is: For a covariance-stationary process y_t with absolutely summable autocovariances, the sample periodogram at frequency ω is $\hat{S}_y(\omega) = \frac{1}{2\pi} \sum_{j=-T+1}^{T-1} \hat{r}_j e^{-i\omega j}$ where $\hat{r}_j$ is the j-th order sample autocovariance, T is the sample size and $i = \sqrt{-1}$. For a series with T observations where T is assumed to be odd, the possible frequencies are $\omega_j = 2\pi j / T, j = 1,2...,M$, M= (T-1)/2. The corresponding period is $2\pi / \omega_j = T / j$. Hamilton (1994) illustrates in detail how the sample periodogram is related to the portion of the variance for y_t , and we show it briefly. For the sample analog to the Spectral Representation Theorem: Given any T observations on a process $(y_1, y_2..., y_T)$, there exist frequencies $\omega_1, \omega_2, ...\omega_M$ and coefficients $\hat{\mu}, \hat{\alpha}_1, \hat{\alpha}_2, ..., \hat{\alpha}_M$, $\hat{\delta}_1, \hat{\delta}_2 ..., \hat{\delta}_M$ such that the value of y at data t can be expressed as: $y_t = \hat{\mu} + \sum_{j=1}^{M} [\hat{\alpha}_j \cos(\omega_j t) + \hat{\delta}_j \sin(\omega_j t)]$. The magnitude $(1/2)(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$ represents the portion of the sample variance of y that could be attributed to cycles with frequency ω_j. It is proved that the sample periodogram satisfy the equation: $(4\pi / T) \hat{S}_y(\omega_j) = (1/2)(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$. The cycles at the frequencies with larger periodogram represent higher portion of the sample variance of y, and therefore the cycles with these frequencies are important.

II Calculation. To compute the sample periodograms, that the series should be stationary is a basic assumption. Thus the nonstationary series needs to be transformed to be stationary before calculating the sample periodograms.

The sample periodogram can be calculated by $(4\pi / T) \hat{S}_y(\omega_j) = (1/2)(\hat{\alpha}_j^2 + \hat{\delta}_j^2)$, where $\hat{\alpha}_j$ and $\hat{\delta}_j$ are OLS estimate of $y_t = \hat{\mu} + \sum_{j=1}^{M} [\hat{\alpha}_j \cos(\omega_j t) + \hat{\delta}_j \sin(\omega_j t)]$. In practical issue, when sample size T is large, it is very complicated to make the regression above with T parameters. To make the sample periodograms easy to compute, we apply the Discrete Fourier Transformation (DFT) to the series: $d_j = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_t \exp(-\frac{2j\pi t}{T})$, j and T are defined the same as before. It is proved that the periodogram can be calculated by $\hat{S}_y(\omega_j) = \frac{1}{2\pi} \left| d_j \right|^2$, where $\left| d_j \right|$ is the mode of the DFT transformation (See in Stoffer and Shumway (2005), P190). The new expression avoids the complex regression. Because most software has the function to make Fourier transformation, it makes the sample periodograms easy to derive.

3.2 Seasonal unit root and HEGY tests

The seasonal unit roots in series y_t would permanently change the seasonal pattern and seasonally integrated series have variances which increase or decrease linearly, therefore, the investigation of seasonal unit roots logically precedes modeling seasonality. HEGY test is a test for the presence of seasonal and nonseasonal unit roots in quarterly data (Hylleberg, 1990), and monthly data see Franses (1990). Those tests allow us to test for unit roots at some seasonal frequencies without maintaining that unit roots are present at all seasonal frequencies. But it is not available for hourly data. Therefore, in the next section, we shall present details of our extending work for HEGY-type tests to cover the case for hourly data.

3.3 Seasonal ARIMA model

To model seasonality, generally speaking, there are three types of models, which are deterministic seasonal model using dummy variables and stationary/nonstatioanry stochastic seasonal model. The stochastic models allow data at one time point to affect the data circles afterwards, while the deterministic model has fixed seasonality with “no memory”. This characteristic of stochastic models is similar with most society and economic phenomenon that the behavior forward will affect the behavior afterward.

The most frequently used stochastic seasonal model is the Seasonal $\mathrm{ARIMA}(\mathrm{p,d,q})(\mathrm{P,D,Q})_S$ model. It is defined as bellow:

$$(1-B)^d(1-B^S)^D\phi(B)\Phi(B^S)Y_t = \theta(B)\Theta(B^S)Z_t$$

where $\phi(z)=1-\phi_1 z-\cdots-\phi_p z^p$, $\Phi(z^S)=1-\Phi_1 z-\cdots-\Phi_P z^P$, $\theta(z)=1-\theta_1 z-\cdots-\theta_q z^q$, $\Theta(z)=1-\Theta_1 z-\cdots-\Theta_Q z^Q$, and $Z_t \sim WN(0,\sigma^2)$. S is the period of the seasonal ARIMA model as defined before, p, q, P, Q are the order of the model, the values of d and D are often 0, 1 or 2. Some famous seasonal models such as airline model (see in Box and Jenkin (1970), P305) and autoregressive-moving average model (see Granger and Newbold, 1986) are special cases based on the SARIMA model.

There are three steps to identify SARIMA models: 1) Identifying d and D and transform the series $\{Y_t\}$ into stationary series $\{X_t\}$; 2) Identifying p, q, P, Q with the ACF and PACF of $\{X_t\}$. Denote $\hat{\rho}(\bullet)$ is the sample ACF of $\{X_t\}$, then P and Q should be chosen such that $\hat{\rho}(kS), k = 1, 2, ...$ is compatible with the ACF of ARMA(P,Q) process. The orders p and q are selected by matching the sample ACF $\hat{\rho}(1), ... \hat{\rho}(S-1)$ with the ACF of an ARMA (p, q) process (see in Brockwell (2002), P206). The AICc criteria and goodness of fit test are used to select the best model from competing alternatives; 3) Estimate parameters and check diagnostic. The parameters could be estimated by ordinary least square (OLS) or maximum likelihood (ML) method. The diagnostic checking is checking if the residuals are white noise and contain no autocorrelation.

3.4 ARAR algorithm

The estimated SARIMA model could be used to forecast our hourly productions. However, the disadvantage of SARIMA model is that some information must be lost in the differencing transformation. To utilize all the information of data, Newton and Parzen (1984) proposed the ARARMA model. The method firstly transforms the long-memory series to a short-memory series by using a transformation of memory-shorting, and thereafter fits an ARMA model to the new short-memory series.

The ARAR algorithm is an adaption of the ARARMA methodology and is proved to be useful for forecasts in the previous studies, see Makridakis (1982) and Meade (2000). The algorithm is introduced by Brockwell and Davis (2000) and has three steps to do the forecasts.

I Memory shorting process

(1) For each $\tau = 1, 2, ..., 15,$ find the value $\hat{\phi}(\tau)$ that minimizes:

$$ERR(\phi, \tau) = \frac{\sum_{t=\tau+1}^{n} [Y_t - \phi Y_{t-\tau}]^2}{\sum_{t=\tau+1}^{n} Y_t^2}$$

Define $Err(\tau) = ERR(\hat{\phi}(\tau), \tau)$, and choose the lag τ that minimize $Err(\tau)$.

(2) If $Err(\tau) \le 8/n$, where n is the sample size, the transformation is made in the form,

$$\tilde{Y}_t = Y_t - \hat{\phi}(\hat{\tau})Y_{t-\tau}.$$

(3)If $\hat{\phi}(\hat{\tau}) \ge 0.93$ and $\hat{\tau} > 2$, go to the same transformation in (2).

(4)If $\hat{\phi}(\hat{\tau}) \ge 0.93$ and $\hat{\tau} = 1$ or 2, determine the values $\hat{\phi}_1$ and $\hat{\phi}_2$ of ϕ_1 and ϕ_2 that minimize $\sum_{t=3}^{n}[Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2}]^2$, then make the transformation,

$$\tilde{Y}_t = Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2}.$$

(5)If $\hat{\phi}(\hat{\tau}) \le 0.93$, the series is short-memory.

After the above step **I**, the series $\{Y_t\}$ is transformed to a short-memory series:

$$\{S_t, t = k+1, \ldots n\}.$$

II Fit an autoregressive process to the mean-corrected series

$$X_t = S_t - \overline{S}, t = k+1, \ldots n$$

The fitted model has the form:

$$X_t = \phi_1 X_{t-1} + \phi_{l1} X_{t-l1} + \phi_{l2} X_{t-l2} + \phi_{l3} X_{t-l3} + Z_t \quad \text{where} \quad Z_t \sim WN(0, \sigma^2)$$

The lags l_1, l_2, l_3, the coefficient ϕ_j and white noise variance σ^2 are found from the Yule-Walker equation:

$$\begin{bmatrix} 1 & \hat{\rho}(l_1-1) & \hat{\rho}(l_2-1) & \hat{\rho}(l_3-1) \\ \hat{\rho}(l_1-1) & 1 & \hat{\rho}(l_2-l_1) & \hat{\rho}(l_3-l_1) \\ \hat{\rho}(l_2-1) & \hat{\rho}(l_2-l_1) & 1 & \hat{\rho}(l_3-l_2) \\ \hat{\rho}(l_3-1) & \hat{\rho}(l_3-l_1) & \hat{\rho}(l_3-l_2) & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_{l1} \\ \phi_{l2} \\ \phi_{l3} \end{bmatrix} = \begin{bmatrix} \hat{\rho}(1) \\ \hat{\rho}(l_1) \\ \hat{\rho}(l_2) \\ \hat{\rho}(l_3) \end{bmatrix}$$

and $\sigma^2 = \hat{\gamma}(0)[1 - \phi_1 \hat{\rho}(1) - \phi_{l1} \hat{\rho}(l_1) - \phi_{l2} \hat{\rho}(l_2) - \phi_{l3} \hat{\rho}(l_3)]$, where $\hat{\gamma}(j)$ and $\hat{\rho}(j)$ are the sample autocovariances and autocorrelations of the series $\{X_t\}$. The lags l_1, l_2, l_3 are from $1 < l_1 < l_2 < l_3 \le m$ where m=13 or 26, and they are chosen to minimize the white noise variance or maximize the Gaussian likelihood of the observations.

III Use the model developed from the first 2 steps to forecast

The memory-shorting filter in the first step can be expressed as:

$$S_t = \psi(B)Y_t = (1 + \psi_1 B + \ldots \psi_k B^k)Y_t = Y_t + \psi_1 Y_{t-1} + \ldots + \psi_k Y_{t-k}$$

The AR model in the second step can be expressed as: $\phi(B)X_t = Z_t$, where

$$\phi(B) = 1 - \phi_1 B - \phi_{l1} B^{l1} - \phi_{l2} B^{l2} - \phi_{l3} B^{l3}$$

Combine the two equations and the mean-corrected process together the ARAR model is obtained: $\xi(B)Y_t = \psi(B)\phi(B)Y_t = \phi(1)\overline{S} + Z_t$. Then the final model is used for forecasting.

4 Seasonal Unit Root Tests

This section introduces a test to examine the seasonal unit roots in a 24-order autoregressive model. The first subsection introduces seasonal unit roots. The second subsection introduces the HEGY-type test. In the third subsection we present the test procedure for the HEGY-type test. The fourth subsection gives the finite-sample and asymptotic distributions of the test statistics.

4.1 Seasonal unit roots

Consider an autoregressive model containing seasonal unit roots, such as

$$(1 - B^S)y_t = \varepsilon_t \qquad ①$$

where the autoregressive polynomial $\varphi(z) = 1 - z^S = 0$ has S roots on the unit circle. S is the period of seasonal pattern, usually S=4 for quarterly data, S=12 for monthly data, and S=24 for hourly data.

The root z=1 is called nonseasonal unit root, and the rest of the (S-1) roots are called seasonal unit roots. Each unit root corresponds with a specific frequency and a period. Now consider S=24 in the model ①, the polynomial $1 - z^{24}$ in the equation $1 - z^{24} = 0$ has the decomposition such that

$$\begin{aligned}1 - z^{24} = {} & (1-z)(1+z)(1+iz)(1-iz)[1-(\frac{\sqrt{3}}{2}+\frac{1}{2}i)z][[1-(\frac{\sqrt{3}}{2}-\frac{1}{2}i)z][1-(-\frac{\sqrt{3}}{2}+\frac{1}{2}i)z] \\ & [1-(-\frac{\sqrt{3}}{2}-\frac{1}{2}i)z][1-(\frac{1}{2}-\frac{\sqrt{3}}{2}i)z][1-(\frac{1}{2}+\frac{\sqrt{3}}{2}i)z][1-(-\frac{1}{2}-\frac{\sqrt{3}}{2}i)z][1-(-\frac{1}{2}+\frac{\sqrt{3}}{2}i)z] \\ & [1-(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i)z][1-(\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i)z][1-(-\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i)z][1-(-\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i)z] \qquad ② \\ & [1-(\frac{\sqrt{6}-\sqrt{2}}{4}+\frac{\sqrt{6}+\sqrt{2}}{4}i)z][1-(\frac{\sqrt{6}-\sqrt{2}}{4}-\frac{\sqrt{6}+\sqrt{2}}{4}i)z][1-(-\frac{\sqrt{6}-\sqrt{2}}{4}+\frac{\sqrt{6}+\sqrt{2}}{4}i)z]\end{aligned}$$

$$[1-(-\frac{\sqrt{6}-\sqrt{2}}{4}-\frac{\sqrt{6}+\sqrt{2}}{4}i)z][1-(\frac{\sqrt{6}+\sqrt{2}}{4}+\frac{\sqrt{6}-\sqrt{2}}{4}i)z][1-(\frac{\sqrt{6}+\sqrt{2}}{4}-\frac{\sqrt{6}-\sqrt{2}}{4}i)z]$$
$$[1-(-\frac{\sqrt{6}+\sqrt{2}}{4}+\frac{\sqrt{6}-\sqrt{2}}{4}i)z][1-(-\frac{\sqrt{6}+\sqrt{2}}{4}-\frac{\sqrt{6}-\sqrt{2}}{4}i)z]$$

From the decomposition ②, we derive the unit roots of the equation $1-z^{24}=0$. The unit roots and the corresponding frequencies with circles/day are given in Table 2. The circles/day means the number of circles at that frequency per day (or a 24-hours period). For example, the 3 circles/day for the frequency $\frac{\pi}{4}$ means there are 3 circles in 24 hours, each circle lasts for 8 hours.

Table 2: unit roots and corresponding frequencies with circles/day

Unit root	Frequency	Circles/day	Unit root	Frequency	Circles/day
1	0	0	$\frac{\sqrt{2}}{2}\pm\frac{\sqrt{2}}{2}i$	$\frac{\pi}{4},\frac{7\pi}{4}$	3,21
-1	π	12	$-\frac{\sqrt{2}}{2}\pm\frac{\sqrt{2}}{2}i$	$\frac{3\pi}{4},\frac{5\pi}{4}$	9,15
$\pm i$	$\frac{\pi}{2},\frac{3\pi}{2}$	6,18	$\frac{\sqrt{6}-\sqrt{2}}{4}\pm\frac{\sqrt{6}+\sqrt{2}}{4}i$	$\frac{5\pi}{12},\frac{19\pi}{12}$	5,19
$\frac{1}{2}\pm\frac{\sqrt{3}}{2}i$	$\frac{\pi}{3},\frac{5\pi}{3}$	4,20	$-\frac{\sqrt{6}-\sqrt{2}}{4}\pm\frac{\sqrt{6}+\sqrt{2}}{4}i$	$\frac{7\pi}{12},\frac{17\pi}{12}$	7,17
$-\frac{1}{2}\pm\frac{\sqrt{3}}{2}i$	$\frac{2\pi}{3},\frac{4\pi}{3}$	8,16	$\frac{\sqrt{6}+\sqrt{2}}{4}\pm\frac{\sqrt{6}-\sqrt{2}}{4}i$	$\frac{\pi}{12},\frac{23\pi}{12}$	1,23
$\frac{\sqrt{3}}{2}\pm\frac{1}{2}i$	$\frac{\pi}{6},\frac{11\pi}{6}$	2,22	$-\frac{\sqrt{6}+\sqrt{2}}{4}\pm\frac{\sqrt{6}-\sqrt{2}}{4}i$	$\frac{11\pi}{12},\frac{13\pi}{12}$	11,13
$-\frac{\sqrt{3}}{2}\pm\frac{1}{2}i$	$\frac{5\pi}{6},\frac{7\pi}{6}$	10,14			

4.2 HEGY type test for seasonal unit roots

We extend the HEGY test to the case S=24 in the equation ①. In order to carry out our tests, we derive the auxiliary regression model given in the following Proposition.

Proposition: Consider the auxiliary regression ③,

$$\varphi^*(B)y_{14,t}=\eta_1 y_{1,t-1}+\eta_2 y_{2,t-1}+\eta_3 y_{3,t-1}+\eta_4 y_{3,t-2}+\eta_5 y_{4,t-1}+\eta_6 y_{4,t-2}+\eta_7 y_{5,t-1}+\eta_8 y_{5,t-2}$$
$$+\eta_9 y_{6,t-1}+\eta_{10} y_{6,t-2}+\eta_{11} y_{7,t-1}+\eta_{12} y_{7,t-2}+\eta_{13} y_{8,t-1}+\eta_{14} y_{8,t-2}+\eta_{15} y_{9,t-1}+\eta_{16} y_{9,t-2} \quad ③$$
$$+\eta_{17} y_{10,t-1}+\eta_{18} y_{10,t-2}+\eta_{19} y_{11,t-1}+\eta_{20} y_{11,t-2}+\eta_{21} y_{12,t-1}+\eta_{22} y_{12,t-2}+\eta_{23} y_{13,t-1}+\eta_{24} y_{13,t-2}+\mu_t+\varepsilon_t$$

Then we say that testing for presence of seasonal unit roots in y_t with S=24 is equivalent to testing if the corresponding parameters in ③ are zero. In ③ $\varphi^*(B)$ is an autoregressive polynomial, μ_t is the deterministic component that may consists of intercept, trend or dummy variables, $y_{i,t}$ has expression, for each i:

$$
\begin{aligned}
y_{1,t} &= (1+B)(1+B^2)(1+B^4+B^8)(1+B^{12})y_t \\
y_{2,t} &= (1-B)(1+B^2)(1+B^4+B^8)(1+B^{12})y_t \\
y_{3,t} &= (1-B^2)(1+B^4+B^8)(1+B^{12})y_t \\
y_{4,t} &= (1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})y_t \\
y_{5,t} &= (1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})y_t \\
y_{6,t} &= (1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})y_t \\
y_{7,t} &= (1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})y_t \\
y_{8,t} &= (1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})y_t \\
y_{9,t} &= (1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})y_t \\
y_{10,t} &= (1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t \\
y_{11,t} &= (1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t \\
y_{12,t} &= (1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t \\
y_{13,t} &= (1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t \\
y_{14,t} &= (1-B^{24})y_t.
\end{aligned}
$$

For the proof of the Proposition above, see in Appendix I.

4.3 Testing procedures for HEGY-type test

We propose the testing procedures for our extended HEGY type tests:

(a) The parameters for the auxiliary regression ③ are estimated by ordinary least square method.

(b) The first two parameters η_1 and η_2 correspond to unit roots at frequencies 0 and π, the null hypothesis is $H_0:\eta_k=0$ (k=1, 2) and the alternative hypothesis is: $H_a:\eta_k<0$. The test statistics used are t-statistics: $t_{\eta_k}=\hat{\eta}_k/\hat{\sigma}_{\hat{\eta}_k}$, where $\hat{\eta}_k$ is the estimate of η_k and $\hat{\sigma}_{\hat{\eta}_k}$ is the standard error of $\hat{\eta}_k$. If the null hypothesis is not rejected, the unit root exists at that frequency. If the null hypothesis is rejected, the series is stationary at that frequency.

(c) For other parameters, due to the fact that other pairs of unit roots are conjugates, the parameters appear in pairs and correspond with frequencies in pairs. Thus, only that both parameters are zero proof the existence of unit roots. This leads to the joint test of the null hypothesis $H_0 : \eta_{2i} = \eta_{2i-1} = 0$ (i=2,3,…12) against the alternative hypothesis $H_a : \eta_{2i-1} \neq 0 \; or \; \eta_{2i} \neq 0$. The F-statistics

$$F(\eta_{2i-1}, \eta_{2i}) = (1/2)(t^2_{\eta_{2i-1}} + t^2_{\eta_{2i}})$$

are used, along with that t-statistics $t_{\eta_{2i}}$ and $t_{\eta_{2i-1}}$ are derived in the same way as t_{η_1} and t_{η_2}. If the null hypothesis in is not rejected, the unit roots exist at corresponding frequencies, and if the null hypothesis is rejected, the series is stationary at those frequencies.

(d) Another strategy for conjugate unit roots in (c) is to test $H_0 : \eta_{2i} = 0$ against $H_a : \eta_{2i} \neq 0$ by t-statistics $t_{\eta_{2i}}$ at the beginning. If the null hypothesis is not rejected, then examine $H_0 : \eta_{2i-1} = 0$ against $H_a : \eta_{2i-1} \neq 0$ by t-statistics $t_{\eta_{2i-1}}$. If the null hypothesis is not rejected, there are unit roots at the corresponding frequencies.

(e) When the order of $\varphi(B)$ is allowed to be greater than 24 or there is a MA component in the series, the remainder polynomial $\varphi^*(B)$ has order greater than 1 and cannot be treated as constant. If $\varphi^*(B)$ is treated as constant, the residuals of the auxiliary regression would be serially correlated. Therefore certain lags of $y_{14,t}$ are included in the regression to make the residuals white noise. The lags do not affect the asymptotic distribution of the test statistics. When the order of $\varphi(B)$ is much smaller than 24 (e.g. the order is 1, 2 or 3) and no MA component exist in the series, the test loses power when testing the non-seasonal unit root (see in Franses (1998), P115). Therefore, if the null hypothesis $\eta_1 = 0$ is not rejected, an additional Dickey-Fuller test with seasonal dummies is needed to test if the non-seasonal unit root really exists.

4.4 The distributions of HEGY-type tests

In this subsection, we study the finite-sample and asymptotic distributions of the test statistics in Section 4 by Monte Carlo simulations. The data generating process starts from the model $y_t = y_{t-24} + \varepsilon_t$, where $\varepsilon_t \sim N(0,1)$. The sample sizes T=120, 240,480 is designed for studying small-sample distribution of the tests while the T=1200 for asymptotic distribution of the tests. Sequentially, we estimate the auxiliary regression ③ by simulated data. The distributions of the test statistics are different when different deterministic components μ_t are included in ③. Thus we estimate the regression ③ in following four cases: 1) no deterministic included; 2) only intercept included; 3) intercept and trend included; 4) intercept and 23 dummies included. Then the test statistics in (b) and (c) of subsection 4.3 are calculated. This procedure is repeated 10000 times for each case above, yielding the finite-sample and asymptotic distributions of the tests.

The critical values of the statistics for case 1 are presented here, while the results for other three cases are given in Appendix II. Table 3 presents the critical values of t_{η_1}. Table 4 shows the critical values of t_{η_2}. The critical values of F-statistics, $F(\eta_{2i-1}, \eta_{2i})$ i=2,3,…12, are presented in Table 5. It is seen that t_{η_1} and t_{η_2} have distributions skewed to the left.

Table 3 Critical values for the t-statistics t_{η_1}. No deterministic component included

Sample size T	Probability that is less than $t_{\eta_1} = \hat{\eta}_1 / \hat{\sigma}_{\hat{\eta}_1}$ entry					
	0.01	0.05	0.10	0.90	0.95	0.99
T=120	-2.27	-1.68	-1.38	0.88	1.26	1.96
T=240	-2.34	-1.79	-1.51	0.86	1.24	1.97
T=480	-2.47	-1.87	-1.57	0.86	1.25	1.97
T=∞	-2.53	-1.92	-1.61	0.93	1.30	2.01

Table 4 Critical values for the t-statistics t_{η_2} No deterministic component included

Sample size T	Probability that $t_{\eta_2} = \hat{\eta}_2 / \hat{\sigma}_{\hat{\eta}_2}$ is less than entry					
	0.01	0.05	0.1	0.90	0.95	0.99
T=120	-2.28	-1.72	-1.41	0.90	1.28	2.00
T=240	-2.49	-1.83	-1.50	0.91	1.25	1.94
T=480	-2.59	-1.90	-1.57	0.89	1.25	2.00
T=∞	-2.62	-1.91	-1.63	0.88	1.26	2.04

Table 5 Critical values for the F-statistics. No deterministic component included:

Sample size T	Probabilities that $F(\eta_{2i-1}, \eta_{2i}) = (1/2)(t^2_{\eta_{2i-1}} + t^2_{\eta_{2i}})$ is greater than entry, i=2,3,…12.											
	$F(\eta_3,\eta_4)$			$F(\eta_5,\eta_6)$			$F(\eta_7,\eta_8)$			$F(\eta_9,\eta_{10})$		
	0.10	0.05	0.01	0.1	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
T=120	2.08	2.67	4.21	2.12	2.87	4.77	2.03	2.78	4.52	2.37	3.30	5.55
T=240	2.22	2.88	4.23	2.35	3.05	4.97	2.28	3.11	4.96	2.63	3.67	6.24
T=480	2.37	3.09	4.78	2.40	3.22	5.35	2.39	3.82	5.39	2.72	3.83	6.13
T=∞	2.41	3.11	4.80	2.49	3.28	5.23	2.50	3.40	5.51	2.87	3.97	6.32
	$F(\eta_{11},\eta_{12})$			$F(\eta_{13},\eta_{14})$			$F(\eta_{15},\eta_{16})$			$F(\eta_{17},\eta_{18})$		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
T=120	2.34	3.25	5.66	2.11	2.86	4.86	2.24	3.10	5.22	2.00	2.64	4.14
T=240	2.65	3.64	6.07	2.47	3.33	5.35	2.41	3.30	5.65	2.19	2.90	4.56
T=480	2.74	3.48	5.98	2.53	3.48	5.71	2.59	3.61	5.88	2.28	2.95	4.71
T=∞	2.81	3.87	6.03	2.57	3.45	5.78	2.63	3.67	6.16	2.40	3.10	4.98
	$F(\eta_{19},\eta_{20})$			$F(\eta_{21},\eta_{22})$			$F(\eta_{23},\eta_{24})$					
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01			
T=120	2.05	2.70	5.96	2.37	3.30	5.86	2.38	3.32	5.91			
T=240	2.30	3.05	5.97	2.68	3.64	6.11	2.62	3.69	6.31			
T=480	2.35	3.11	6.02	2.83	3.86	6.43	2.76	3.74	6.17			
T=∞	2.37	3.13	4.90	2.89	3.93	6.68	2.91	4.00	6.73			

5 Results

5.1 Spectral analysis

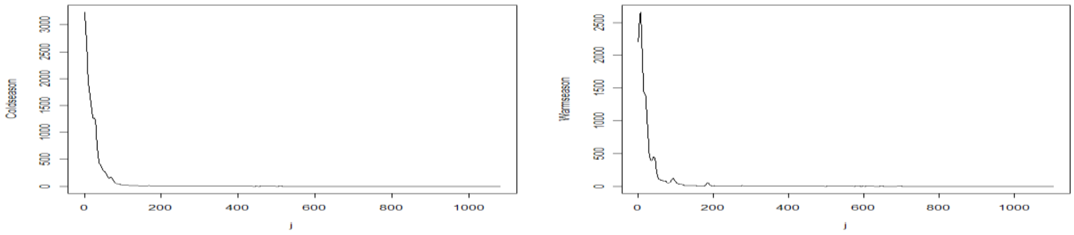

Figure 5 Sample periodogram of original data.

Figure 5 plots the sample periodograms for the data series defined in Section 2. The above sample periodograms suggest that the original series may be not stationary (See in Hamilton (1994), P169). Therefore we analyze the first difference of the series, $\Delta y = y_t - y_{t-1}$. The sample periodograms for the first difference of the data in cold season and warm season are plotted in Figure 6.

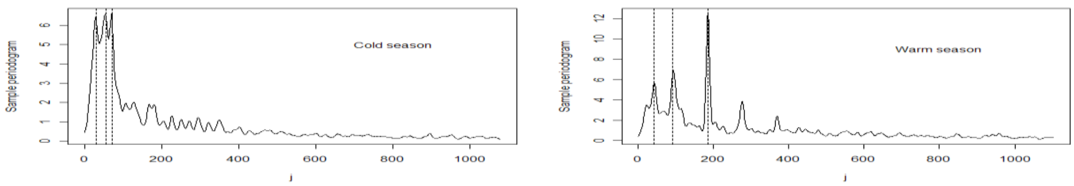

Figure 6: Smoothed periodogram for data in cold season and warm season, with kernel smooth, bandwidth=12

The corresponding lines are j=30,56,72 for cold season, j=(42,91,185) for warm season.

The period of the circle with frequency ω_j is derived by $2\pi / \omega_j = T / j$ as stated in subsection 3.1. j=1,…,M are defined the same as the j in the definition of $\omega_j = 2\pi j / T$. The peak values of sample periodograms stand for important circles. For the peak value in Figure 6, we present corresponding j values and the periods of circles in Table 6.

For cold season, the low frequency components represent most part of the sample variance, and sample variance could be attributed to the circles with the 3 frequencies in the series. However, the periodograms show there are a lot of small circles in the series and it means that there are a lot of noises; the noises take a part of sample variance that could not be neglected. The periodograms for warm season show much better periodical pattern. There are 5 peaks in the Figure 6, and the first 3 peaks are most important. There are not as many small circles in periodograms as in cold season.

Table 6 the j value of peaks and corresponding periods of circles

Season	Cold Season			Warm Season		
j	30	56	72	42	91	185
Period(hours)	72	39	29	52	24	12
Period(days)	3	1.6	1.2	2.2	1	0.5

From the spectral analysis, there are several seasonal patterns in the series. In both warm and cold seasons, the 24 hours period represents a large part of sample variance although in cold season the 29 hours circle is a little longer than 24 hours. In warm season the half-day circle could be covered by the 24 hours circle and are subsequence of the circle. Also considering the fact that the data is hourly data, it is reasonable and necessary to use the 24-hour circle to model the seasonality.

5.2 Test seasonal unit roots

The wind power data for cold season and warm season are tested separately. With different deterministic components included, all the four cases in subsection 4.4 are considered. Then we calculate the statistics in (2) and (3) of subsection 4.3 and compare them with the critical values in Subsection 4.4. We firstly estimate the auxiliary regression with 24 hours lag values of $y_{14,t}$ in the model, which means $\varphi^*(B) = 1 - \varphi_1^* B - \varphi_2^* B^2 - \cdots - \varphi_{24}^* B^{24}$, then we exclude the items whose parameters are not significantly nonzero at 10% level by t-statistic, and at last, the auxiliary regression is estimated again including all remaining lags of $y_{14,t}$. The lags of $y_{14,t}$ are finally chosen. The t-statistics for η_1, η_2 and F-statistics for other pairs of parameters are presented in Table 7 for cold season and Table 8 for warm season, including our four cases. Those four cases are denoted as follows: 1) "n(I,T,D)": no deterministic is included in the regression; 2) "I,nT,nD": only intercept included; 3)

“I,T,nD”: intercept and trend included; 4) “I,nT,D”: intercept and 23 dummies included.

The result of cold season shows that two seasonal unit roots at frequencies $\frac{\pi}{4},\frac{7\pi}{4}$ exist and the unit root at frequency 0 exits. For warm season, the only unit root is the nonseasonal unit root. The roots at frequencies $\frac{\pi}{4},\frac{7\pi}{4}$ exist only when dummies included in the regression. For the other cases there do not exist any seasonal unit roots in the series.

Table 7 Test results for cold season

Case	t_{η_1}	t_{η_2}	$F(\eta_3,\eta_4)$	$F(\eta_5,\eta_6)$	$F(\eta_7,\eta_8)$	$F(\eta_9,\eta_{10})$	$F(\eta_{11},\eta_{12})$
n(I,T,D)	0.08	-13.46*	55.25*	114.04*	53.97*	164.93*	14.52*
I,nT,nD	0.26	-12.92*	55.17*	113.92*	53.95*	164.87*	14.55*
I,T,nD	0.27	-12.91*	55.14*	113.86*	53.92*	164.79*	14.55*
I,nT,D	0.30	-12.98*	53.17*	110.60*	53.75*	165.73*	15.01*
	$F(\eta_{13},\eta_{14})$		$F(\eta_{15},\eta_{16})$	$F(\eta_{17},\eta_{18})$	$F(\eta_{19},\eta_{20})$	$F(\eta_{21},\eta_{22})$	$F(\eta_{23},\eta_{24})$
n(I,T,D)	1.29		168.17*	71.68*	140.66*	41.15*	166.89*
I,nT,nD	1.29		168.13*	71.60*	140.50*	41.16*	166.78*
I,T,nD	1.28		168.05*	71.57*	140.42*	41.15*	166.69*
I,nT,D	1.42		166.92*	69.60*	136.09*	41.41*	163.67*

Note: Lags used are 3,8,9,12,13,20,21,22,23, and they are the same in the 4 models. The parameter of trend is 9.535e-05 and the t-statistic is 0.179. Only 2 parameters of dummy variables are non-zero at 5% level.

Table 8: Test results for warm season

Case	t_{η_1}	t_{η_2}	$F(\eta_3,\eta_4)$	$F(\eta_5,\eta_6)$	$F(\eta_7,\eta_8)$	$F(\eta_9,\eta_{10})$	$F(\eta_{11},\eta_{12})$
n(I,T,D)	0.04	-12.53*	42.43*	79.51*	25.35*	118.86*	6.68*
I,nT,nD	-0.07	-11.79*	42.41*	79.48 *	25.33*	118.81*	6.62*
I,T,nD	-0.02	-11.75*	42.39*	79.43*	25.32*	118.77*	6.63*
I,nT,D	-0.01	-12.06*	50.50*	93.21*	31.53*	128.13*	13.65*
	$F(\eta_{13},\eta_{14})$		$F(\eta_{15},\eta_{16})$	$F(\eta_{17},\eta_{18})$	$F(\eta_{19},\eta_{20})$	$F(\eta_{21},\eta_{22})$	$F(\eta_{23},\eta_{24})$
n(I,T,D)	5.83*		157.19*	58.47*	67.07*	8.94*	131.25*
I,nT,nD	5.83*		157.12*	58.44*	67.03*	8.93*	131.19*
I,T,nD	5.85*		156.99*	58.42*	67.01*	8.92*	131.10*
I,nT,D	6.30		190.51*	58.12*	82.57*	11.38*	158.47*

Note: The lags used for the first 3 models are: (1,3,4,5,6,8,9,10,11,12,13,14,19,20,24). For the last model the lags are:(3,4,5,6,8,9,1011,12,13,14,19,20). The parameter of the trend is: -2.320e-04 and the t-statistic is -0.32. 12 parameters of dummy variable out of 23 are significant non-zero at 10% level. “*”stand for being rejected at 5% level.

As advocated by Box and Jenkin (See in Box and Jenkin (1970), P303), the application of the filter $1-B^d$ to the series requires that the series has unit roots at zero frequency and all of the seasonal frequencies. It means that the series needs the filter $1-B^d$ when all unit roots of $1-Z^d=0$ exist. Consider in an opposite way, due to the fact that each unit root corresponds with a decomposition factor of $1-Z^d$, the multiplication of these factors could compose the filter $1-Z^d$. For example, when d=1, it only needs the unit roots at 0 frequency existing to apply the 1-B filter to the series; when d=2, it needs unit roots exist at frequency 0 and π to apply the filter $1-B^2$ to the series, the corresponding factors of the two roots compose the filter: $(1-\mathrm{B})(1+\mathrm{B})=1-B^2$. For the test results above, in cold season, the unit root at frequency 0 corresponds to the factor 1-B. The unit roots at frequency $\frac{\pi}{4},\frac{7\pi}{4}$, which are $\frac{\sqrt{2}}{2}\pm\frac{\sqrt{2}}{2}i$, correspond to $[1-(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i)B][1-(\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i)B]=(1-\sqrt{2}B+B^2)$. The application of the filter $1-B^{24}$ on series y_t needs 24 unit roots existing in the series, so the series do not need that filter. The multiplication of the factors above $(1-B)(1-\sqrt{2}B+B^2)$ does not compose any other filter except (1-B). Therefore the results tell us that the filter (1-B) is needed for the series, the filter $1-B^{24}$ for hourly data will cause overdifferenciate problem and the forecast will be severely biased. In warm season the result is similar and (1-B) is needed for the series. Therefore the series of warm season and cold season are not seasonally integrated.

The deterministic parts have little influence on the results of the tests; the intercept only influences the value of the first statistics. The estimated values of trend parameters and t-statistics in both seasons are close to 0, and then there is no evidence of including trend in the two series. The dummies have only a little influence on the values of the test statistics. Most of the estimated parameters of dummy variables in cold season and half of them in warm season are zero at 10% level. It suggests that trend and dummy variables are not necessary included in our model.

5.3 SARIMA Modeling

We model the data from December 1 to December 30, 2008; According to the results of testing seasonal unit roots, the difference filter (1-B) is needed to transform the series to be stationary and there is not any other filter needed, thus, D equals to 0 and d=1. Then we obtain a stationary series $\{X_t\}$, $X_t = y_t - y_{t-1}$ where y_t is the raw data.

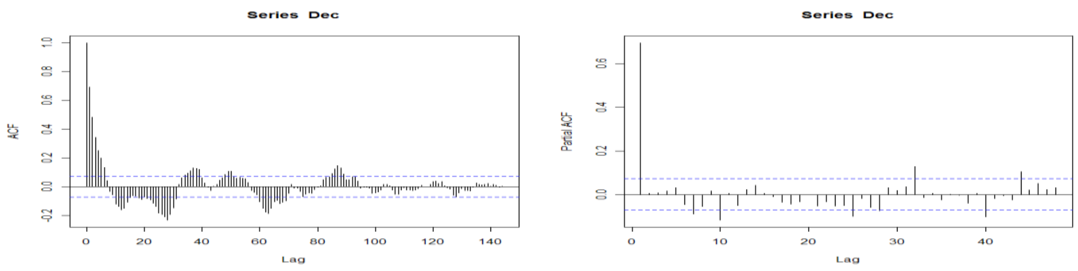

Figure 7: ACF and PACF of the first difference of the data in December, 2008

The sample autocorrelation function and partial autocorrelation function for $\{X_t\}$ are given in Figure 7. In ACF plot of Figure 7, the values of $\hat{\rho}(S), \hat{\rho}(2S), \hat{\rho}(3S)$ fall outside the bounds $\pm 1.96/\sqrt{T}$ (the 2 dotted horizontal lines in the figure, T is sample size) and the value of $\hat{\rho}(4S)$ does not. This suggests Q equal to 3 or 4 and P equal to 0. The values of $\hat{\rho}(1), \ldots, \hat{\rho}(6)$ fall outside the same bounds, so q is around 6 and p is 0. The estimation method is maximum likelihood with starting values determined by conditional sum of squares. The model with order (0,1,6)(0,0,3) is preferred by lowest AICc value of 6111 and there is no residual autocorrelation in it. The p-value of Ljung-Box test statistic for independence is 0.760 (null hypothesis is independence). The residuals are independent. Figure 8 shows the autocorrelation of the residuals.

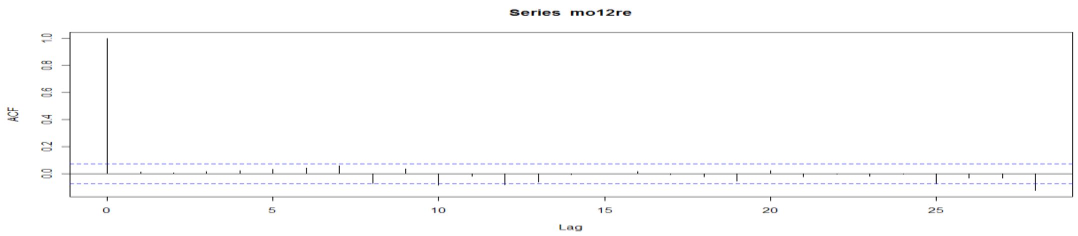

Figure 8 autocorrelation of the model residual

5.4 Forecasting

First the forecasting horizons need to be determined. Different horizons have different usages. The forecasts with horizons between 1-24 hours are used for intra-day trade and transportation, and the forecasts with horizons between 24-48 hours are used in day-ahead trade. In this paper we focus on 1 hour ahead forecast for intra-day trade and 36 hours ahead forecast for day-ahead trade. In Sweden, the gate closure time for intra-day trade is 1 minute before next hour (10:59 for 11:00-12:00) and 3:30 pm for the next day trade, therefore the 1 hour prediction and 36 hour prediction are useful for both trade and transmission. The forecasts for other horizons could be investigated in the same way.

The cost function used is root mean square error (RMSE) and mean absolute error (MAE):

$$\text{RMSE:} \quad \left(\frac{1}{T}\sum_{1}^{T}(y_t - \hat{y}_t)^2\right)^{1/2} \qquad \text{MAE:} \quad \frac{1}{T}\sum_{1}^{T}\left|y_t - \hat{y}_t\right|$$

where y_t is the true production and $\hat{y}_t$ is the predicted production. Both the two cost functions measure the deviation of estimates. The difference is that RMSE put more weight on large deviations while MAE put equal weight on each deviation. The prediction method is used to predict the data of cold season (from 1a.m January 1, 2009) and warm season (from 1a.m July 30, 2009). We predict 180 hours production for each season. The 180 hours period is divided into 5 periods, each period lasts for 36 hours. For each period, the 720 observations (30 days) before the period are used for model estimation. The RMSE and MAE for SARIMA and ARAR are presented in Table 9. Figure 9 gives the number of smaller values for the two methods by bar chart.

Smaller values of cost functions stand for better forecast. In Figure 9, the results show us that for cold season, the ARAR algorithm has 6 smaller values which is slightly better than warm season (4 smaller values) in RMSE. The two methods have the same number of smaller values in MAE. For warm season, ARAR (7 smaller values) outperforms ARIMA (3 small values), and ARAR is better at both 1 hour prediction and 36 hours prediction. As a conclusion, for cold season, both SARIMA and ARAR algorithm could be used to forecast. For warm season, it is much better to use ARAR algorithm to forecast.

The results of RMSE and MAE do not have much difference except for 1 period, so the two methods do not perform differently between large deviations and small deviations. The results between the 2 horizons do not vary a lot; so for the two horizons, the SARIMA model and ARAR algorithm do not perform differently.

Table 9: RMSE and MAE of predictions

RMSE		Cold Season		Warm Season	
		ARAR	SARIMA	ARAR	SARIMA
1-36h	1 h	12.07*	12.08	21.8	21.0*
	36 h	27.5*	29.3	176.0	175.8*
37-72h	1h	22.4	21.8*	16.7*	17.0
	36h	187.8	183*	233*	241.1
73-108h	1h	13.7*	13.8	11.0*	11.5
	36h	51.5*	81.5	46.0	30.6*
109-144h	1h	14.8	14.3*	5.2*	5.3
	36h	386*	426.6	11.7*	17.7
145-180h	1h	20.4*	20.8	5.5*	5.6
	36h	266.3	243.0*	13.7*	15.6
MAE		ARAR	SARIMA	ARAR	ARIMA
1-36h	1 h	10.0	9.9*	16.7	16.6*
	36 h	24.4*	25.9	140.2	136.7*
37-72h	1h	19.2	18.6*	11.45*	11.49
	36h	161.3	157*	225*	234.5
73-108h	1h	9.7*	10.0	7.7*	7.8
	36h	45.6*	59.0	39.0	26.4*
109-144h	1h	12.0	11.1*	3.9*	3.9
	36h	332.4*	372.0	9.3*	12.8
145-180h	1h	16.1*	16.3	4.2*	4.4
	36h	235.8	217.0*	11.0*	12.3

Note: The number of smaller values for each method is given below where ARAR: 5(2/36.3/1) means ARAR has 5 samller values, 2 at 36 hours ahead and 3 at 1 hour ahead. “*” stand for smaller value of cost function.

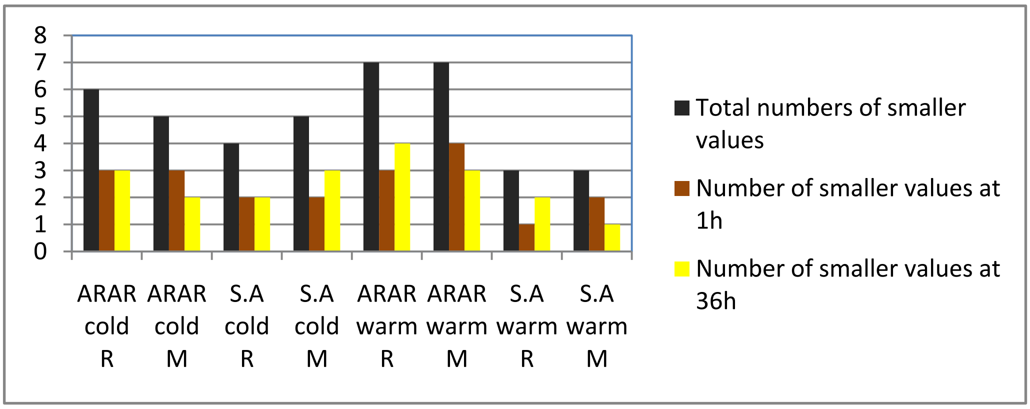

Figure 9: The number of smaller value in RMSE and MAE. "S.A" stand for SARIMA, "R" stands for RMSE, "M" stands for MAE.

6 Conclusions and Future work

In this paper, we model and forecast the hourly wind power production in Sweden. The spectral analysis is used to investigate the seasonality of the series. We extend the HEGY seasonal unit root test to hourly data and apply the test to wind power productions. We find the filter (1-B) is needed to transform the series to be stationary. Based on the unit root test results, the Seasonal ARIMA model is used to fit the wind power production data. To forecast the productions, we use the ARAR algorithm and the estimated SARIMA model. The ARAR algorithm yields better prediction than estimated SARIMA model, especially in warm season.

The SARIMA model is also called "green box" models in wind power production forecasting. There is still a lot of future work to do with the forecast of the hourly wind power production. The future work may contain further investigation of the seasonality of the series, interval forecasting and different horizon forecasting, using other statistical models such as threshold autoregressive model and regime-switching model to forecast, wavelet analysis and transformation, "Black box" models such as neutral network and nearest neighbor search.

References

Box, G.E.P and Jenkins, G.M., 1970. *Time Series Analysis,* Forecasting and Control. Holden-Day, San Francisco, CA.

Franses, P.H., 1990. Testing for seasonal unit roots in monthly data. T*echnical Report 9032*, Econometric Institute.

Franses, P.H, 1998. *Time series models for business and economic forecasting.* Cambridge: Cambridge University Press

Hamilton J.D., 1994. *Time Series analysis*. Princeton: Princeton University Press.

Hylleberg, Engle, R., Granger, C., and Yoo, B., 1990. *Seasonal integration and cointegration*. Journal of Econometrics, 44, 215-238.

International Energy Agency, 2005 Variability of Wind Power and other Renewables: Management Options and Strategies

Lange, M. and Focken, U., 2005. *State-of-the-Art in Wind Power Prediction in Germany and International Developments,* Second Workshop of International Feed-In Cooperation. Berlin, Germany.

Makridakis,S., et al., The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting.* 1, pp.111-153

Meade,N., 2000. A note on the Robust Trend and ARARMA methodologies used in the M3 Competition. *International Journal of Forecasting.* 16, pp.517-519

Newton, H.J. & Parzen, E., 1984. Forecasting and time series model types of 111 economic time series. *In S. Makridakis, etal. (Eds.), The forecasting accuracy of major time series methods*. Chichester: Wiley.

Peter J. Brockwell and Richard A. Davis, 2000. *Introduction to Time Series and Forecasting*. NewYork, Inc: Springer-Verlag.

Shumway, R.H., Stoffer, D.S., 2005. *Time Series Analysis and Its Applications With R Examples* NewYork, Inc: Springer-Verlag.

Appendix I

The basic assumption of HEGY procedure is given below:

1) Basic proposition: For any polynomial $\varphi(B)$ (possibly infinite or rational), which is finite-valued at the distinct, nonzero, possibly complex points $\theta_1,...\theta_p$, can be expressed as:

$$\varphi(B)=\sum_{k=1}^{p}\lambda_k\Delta(B)(1-\delta_k(B))/\delta_k(B)+\Delta(B)\varphi^*(B)\,,\quad ④$$

where $\delta_k(B)=1-(1/\theta_k)B,\Delta(B)=\prod_{k=1}^{p}\delta_k(B)$, $\varphi^*(B)$ is a remainder polynomial. If $\theta_t\neq 0$ is the root of $\varphi(B)=0$, then $\varphi(\theta_t)=0,\Delta(\theta_t)=\delta_t(\theta_t)=0$, with ④:

$$\varphi(\theta_t)=\sum_{k=1}^{p}\lambda_k\Delta(\theta_t)(1-\delta_k(\theta_t))/\delta_k(\theta_t)+\Delta(\theta_t)\varphi^*(\theta_t)=\lambda_t\Delta(\theta_t)(1-\delta_t(\theta_t))/\delta_t(\theta_t)$$

$$=\lambda_t\theta_t\prod_{k=1,k\neq t}^{p}\delta_k(\theta_t)=0$$

Because $\theta_t\prod_{k=1,k\neq t}^{p}\delta_k(\theta_t)\neq 0$ (if θ_t is not a repeated root), so λ_t must be 0. Then with the expansion in ④, if a specific parameter λ_k=0, then we could infer that the corresponding θ_k is the root of $\varphi(B)=0$.

2) Seasonal unit roots in hourly data:

To apply this proposition to testing for seasonal unit root, we make θ_k as the unit roots in hourly data. For hourly data, the basic polynomial is $1-z^{24}=0$ can be factored as follows:

$$1-B^{24}=(1-B)(1+B)(1+B^2)(1-B+B^2)(1+B+B^2)(1-\sqrt{3}B+B^2)$$

$$(1+\sqrt{3}B+B^2)(1-\sqrt{2}B+B^2)(1+\sqrt{2}B+B^2)(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)$$

$$(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)$$

With the unit roots, the autoregressive polynomial $\varphi(B)$ can be decomposed as follows :

$$
\begin{aligned}
\varphi(B) &= \lambda_1 B(1+B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&\lambda_2(-B)(1-B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&\lambda_3(-iB)(1-iB)(1-B^2)(1+B^4+B^8)(1+B^{12})+\\
&\lambda_4(iB)(1+iB)(1-B^2)(1+B^4+B^8)(1+B^{12})+\\
&\lambda_5(\frac{1}{2}B)(1-\sqrt{3}i-2B)(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&\lambda_6(\frac{1}{2}B)(1+\sqrt{3}i-2B)(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&\lambda_7(-\frac{1}{2}B)(1-\sqrt{3}i+2B)(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&\lambda_8(-\frac{1}{2}B)(1+\sqrt{3}i+2B)(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&\lambda_9(\frac{1}{2}B)(\sqrt{3}-i-2B)(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&\lambda_{10}(\frac{1}{2}B)(\sqrt{3}+i-2B)(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&\lambda_{11}(-\frac{1}{2}B)(\sqrt{3}-i+2B)(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+ \qquad ⑤\\
&\lambda_{12}(-\frac{1}{2}B)(\sqrt{3}+i+2B)(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&\lambda_{13}B(\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i-B)(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&\lambda_{14}B(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i-B)(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&\lambda_{15}B(-\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i-B)(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&\lambda_{16}B(-\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}i-B)(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&\lambda_{17}B(\frac{\sqrt{6}-\sqrt{2}}{4}-\frac{\sqrt{6}+\sqrt{2}}{4}i-B)(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{18}B(\frac{\sqrt{6}-\sqrt{2}}{4}+\frac{\sqrt{6}+\sqrt{2}}{4}i-B)(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{19}B(-\frac{\sqrt{6}-\sqrt{2}}{4}-\frac{\sqrt{6}+\sqrt{2}}{4}i-B)(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{20}B(-\frac{\sqrt{6}-\sqrt{2}}{4}+\frac{\sqrt{6}+\sqrt{2}}{4}i-B)(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{21}B(\frac{\sqrt{6}+\sqrt{2}}{4}-\frac{\sqrt{6}-\sqrt{2}}{4}i-B)(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{22}B(\frac{\sqrt{6}+\sqrt{2}}{4}+\frac{\sqrt{6}-\sqrt{2}}{4}i-B)(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{23}B(-\frac{\sqrt{6}+\sqrt{2}}{4}-\frac{\sqrt{6}-\sqrt{2}}{4}i-B)(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\lambda_{24}B(-\frac{\sqrt{6}+\sqrt{2}}{4}+\frac{\sqrt{6}-\sqrt{2}}{4}i-B)(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\varphi^*(B)(1-B^{24})
\end{aligned}
$$

Then testing seasonal unit roots is turned into testing whether the parameters λ_k equal to 0. $\lambda_k = 0$ implies the existence of seasonal unit root. However there are complex parts in the expansion which raise trouble in getting parameters. To remove the complex parts we make a couple of substitutions.

$$\pi_1 = \lambda_1, \pi_2 = -\lambda_2, 2\lambda_3 = -\pi_3 + i\pi_4, 2\lambda_4 = -\pi_3 - i\pi_4$$
$$2\lambda_5 = -\pi_5 + i\pi_6, 2\lambda_6 = -\pi_5 - i\pi_6, 2\lambda_7 = -\pi_7 + i\pi_8, 2\lambda_8 = -\pi_7 - i\pi_8$$
$$2\lambda_9 = -\pi_9 + i\pi_{10}, 2\lambda_{10} = -\pi_9 - i\pi_{10}, 2\lambda_{11} = -\pi_{11} + i\pi_{12}, 2\lambda_{12} = -\pi_{11} - i\pi_{12}$$
$$2\lambda_{13} = -\pi_{13} + i\pi_{14}, 2\lambda_{14} = -\pi_{13} - i\pi_{14}, 2\lambda_{15} = -\pi_{15} + i\pi_{16}, 2\lambda_{16} = -\pi_{15} - i\pi_{16}$$
$$2\lambda_{17} = -\pi_{17} + i\pi_{18}, 2\lambda_{18} = -\pi_{17} - i\pi_{18}, 2\lambda_{19} = -\pi_{19} + i\pi_{20}, 2\lambda_{20} = -\pi_{19} - i\pi_{20}$$
$$2\lambda_{21} = -\pi_{21} + i\pi_{22}, 2\lambda_{22} = -\pi_{21} - i\pi_{22}, 2\lambda_{23} = -\pi_{23} + i\pi_{24}, 2\lambda_{24} = -\pi_{23} - i\pi_{24},$$

Then the decomposition above becomes:

$$\varphi(B) = \pi_1 B(1+B)(1+B^2)(1+B^4+B^8)(1+B^{12}) +$$
$$\pi_2 B(1-B)(1+B^2)(1+B^4+B^8)(1+B^{12}) +$$
$$(\pi_3 B + \pi_4) B(1-B^2)(1+B^4+B^8)(1+B^{12}) +$$
$$[(-\frac{1}{2}+B)\pi_5 + \frac{\sqrt{3}}{2}\pi_6] B(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12}) + \quad ⑥$$
$$[(\frac{1}{2}+B)\pi_7 + \frac{\sqrt{3}}{2}\pi_8] B(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12}) +$$

$$[(-\frac{\sqrt{3}}{2}+B)\pi_9 + \frac{1}{2}\pi_{10}] B(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12}) +$$
$$[(\frac{\sqrt{3}}{2}+B)\pi_{11} + \frac{1}{2}\pi_{12}] B(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12}) +$$
$$[(-\frac{\sqrt{2}}{2}+B)\pi_{13} + \frac{\sqrt{2}}{2}\pi_{14}] B(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12}) +$$
$$[(\frac{\sqrt{2}}{2}+B)\pi_{15} + \frac{\sqrt{2}}{2}\pi_{16}] B(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12}) +$$
$$[(-\frac{\sqrt{6}-\sqrt{2}}{4}+B)\pi_{17} + \frac{\sqrt{6}+\sqrt{2}}{4}\pi_{18}] B(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12}) +$$
$$[(\frac{\sqrt{6}-\sqrt{2}}{4}+B)\pi_{19} + \frac{\sqrt{6}+\sqrt{2}}{4}\pi_{20}] B(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12}) +$$
$$[(-\frac{\sqrt{6}+\sqrt{2}}{4}+B)\pi_{21} + \frac{\sqrt{6}-\sqrt{2}}{4}\pi_{22}] B(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12}) +$$
$$[(\frac{\sqrt{6}+\sqrt{2}}{4}+B)\pi_{23} + \frac{\sqrt{6}-\sqrt{2}}{4}\pi_{24}] B(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12}) +$$
$$\varphi^*(B)(1-B^{24})$$

With this substitution testing $\lambda_k = 0$ is turned to testing if parameters of ⑥ are 0.

Because of the substitutions are taken in pairs, π_{2k-1} and π_{2k} (k=2,3,…12) should equal to 0 simultaneously to get conclusion $\lambda_{2k-1}=\lambda_{2k}=0$, except for λ_1 and λ_2 which can be tested separately by π_1 and π_2. The expression above is still complicated to get the practical auxiliary regression.

To make the function much easier to test another small substitution is needed. Rewrite ⑥ as:

$$
\begin{aligned}
&\varphi(B)=\pi_1 B(1+B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&\pi_2 B(1-B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&(\pi_3 B+\pi_4)B(1-B^2)(1+B^4+B^8)(1+B^{12})+\\
&[(-\frac{1}{2}\pi_5+\frac{\sqrt{3}}{2}\pi_6)+\pi_5 B]B(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&[(\frac{1}{2}\pi_7+\frac{\sqrt{3}}{2}\pi_8)+\pi_7 B]B(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&[(-\frac{\sqrt{3}}{2}\pi_9+\frac{1}{2}\pi_{10})+\pi_9 B]B(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&[(\frac{\sqrt{3}}{2}\pi_{11}+\frac{1}{2}\pi_{12})+\pi_{11} B]B(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&[(-\frac{\sqrt{2}}{2}\pi_{13}+\frac{\sqrt{2}}{2}\pi_{14})+\pi_{13} B]B(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&[(\frac{\sqrt{2}}{2}\pi_{15}+\frac{\sqrt{2}}{2}\pi_{16})+\pi_{15} B]B(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&[(-\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{17}+\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{18})+\pi_{17} B]B(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&[(\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{19}+\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{20})+\pi_{19} B]B(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&[(-\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{21}+\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{22})+\pi_{21} B]B(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&[(\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{23}+\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{24})+\pi_{23} B]B(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\varphi^*(B)(1-B^{24})
\end{aligned}
\qquad ⑦
$$

Use the substitution:

$$\pi_1=\eta_1,\pi_2=\eta_{2,}\pi_4=\eta_{3,}\pi_3=\eta_{4,}(-\frac{1}{2}\pi_5+\frac{\sqrt{3}}{2}\pi_6)=\eta_{5,}\pi_5=\eta_{6,}(\frac{1}{2}\pi_7+\frac{\sqrt{3}}{2}\pi_8)=\eta_{7,}\pi_7=\eta_{8,}$$
$$(-\frac{\sqrt{3}}{2}\pi_9+\frac{1}{2}\pi_{10})=\eta_{9,}\pi_9=\eta_{10,}(\frac{\sqrt{3}}{2}\pi_{11}+\frac{1}{2}\pi_{12})=\eta_{11,}\pi_{11}=\eta_{12,}(-\frac{\sqrt{2}}{2}\pi_{13}+\frac{\sqrt{2}}{2}\pi_{14})=\eta_{13,}\pi_{13}=\eta_{14,}$$
$$(\frac{\sqrt{2}}{2}\pi_{15}+\frac{\sqrt{2}}{2}\pi_{16})=\eta_{15,}\pi_{15}=\eta_{16,}(-\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{17}+\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{18})=\eta_{17,}\pi_{17}=\eta_{18,}$$
$$(\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{19}+\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{20})=\eta_{19,}\pi_{19}=\eta_{20,}(-\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{21}+\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{22})=\eta_{21,}\pi_{21}=\eta_{22,}$$
$$(\frac{\sqrt{6}+\sqrt{2}}{4}\pi_{23}+\frac{\sqrt{6}-\sqrt{2}}{4}\pi_{24})=\eta_{23,}\pi_{23}=\eta_{24,}$$

Then the final expansion is achieved:

Derive the final decomposition that could be used in practical test:

$$\begin{aligned}
\varphi(B)=\;&\eta_1B(1+B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&\eta_2B(1-B)(1+B^2)(1+B^4+B^8)(1+B^{12})+\\
&(\eta_4B+\eta_3)B(1-B^2)(1+B^4+B^8)(1+B^{12})+\\
&(\eta_5+\eta_6B)B(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&(\eta_7+\eta_8B)B(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})+\\
&(\eta_9+\eta_{10}B)B(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&(\eta_{11}+\eta_{12}B)B(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})+\\
&(\eta_{13}+\eta_{14}B)B(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&(\eta_{15}+\eta_{16}B)B(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})+\\
&(\eta_{17}+\eta_{18}B)B(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&(\eta_{19}+\eta_{20}B)B(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&(\eta_{21}+\eta_{22}B)B(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&(\eta_{23}+\eta_{24}B)B(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})+\\
&\varphi^*(B)(1-B^{24})
\end{aligned}$$

Then , denote $y_{i,t}$ for each i=1,…14

$$y_{1,t}=(1+B)(1+B^2)(1+B^4+B^8)(1+B^{12})y_t$$
$$y_{2,t}=(1-B)(1+B^2)(1+B^4+B^8)(1+B^{12})y_t$$
$$y_{3,t}=(1-B^2)(1+B^4+B^8)(1+B^{12})y_t$$
$$y_{4,t}=(1+B+B^2)(1-B^2+B^6-B^8)(1+B^{12})y_t$$
$$y_{5,t}=(1-B+B^2)(1-B^2+B^6-B^8)(1+B^{12})y_t$$
$$y_{6,t}=(1+\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})y_t$$
$$y_{7,t}=(1-\sqrt{3}B+B^2)(1+B^2-B^6-B^8)(1+B^{12})y_t$$
$$y_{8,t}=(1+\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})y_t$$
$$y_{9,t}=(1-\sqrt{2}B+B^2)(1-B^4+B^8)(1-B^{12})y_t$$
$$y_{10,t}=(1+\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t$$
$$y_{11,t}=(1-\frac{\sqrt{6}-\sqrt{2}}{2}B+B^2)(1-\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t$$
$$y_{12,t}=(1+\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t$$
$$y_{13,t}=(1-\frac{\sqrt{6}+\sqrt{2}}{2}B+B^2)(1+\sqrt{3}B^2+B^4)(1+B^4)(1-B^{12})y_t$$
$$y_{14,t}=(1-B^{24})y_t.$$

We get the auxiliary regression model:

$$\varphi^*(B)y_{14,t}=\eta_1 y_{1,t-1}+\eta_2 y_{2,t-1}+\eta_3 y_{3,t-1}+\eta_4 y_{3,t-2}+\eta_5 y_{4,t-1}+\eta_6 y_{4,t-2}+\eta_7 y_{5,t-1}+\eta_8 y_{5,t-2}$$
$$+\eta_9 y_{6,t-1}+\eta_{10} y_{6,t-2}+\eta_{11} y_{7,t-1}+\eta_{12} y_{7,t-2}+\eta_{13} y_{8,t-1}+\eta_{14} y_{8,t-2}+\eta_{15} y_{9,t-1}+\eta_{16} y_{9,t-2}$$
$$+\eta_{17} y_{10,t-1}+\eta_{18} y_{10,t-2}+\eta_{19} y_{11,t-1}+\eta_{20} y_{11,t-2}+\eta_{21} y_{12,t-1}+\eta_{22} y_{12,t-2}+\eta_{23} y_{13,t-1}+\eta_{24} y_{13,t-2}+\varepsilon_t$$

Appendix II Critical values of HEGY test

Table 10 Critical values for the t-statistics. t_{η_1} . Only intercept included: I T nDI

Sample size T	Probability that is less than $t_{\eta_1}=\hat{\eta}_1/\hat{\sigma}_{\hat{\eta}_1}$ entry					
	0.01	0.05	0.10	0.90	0.95	0.99
T=120	-3.15	-2.56	-2.27	-0.29	0.06	0.76
T=240	-3.25	-2.70	-2.41	-0.38	-0.04	0.64
T=480	-3.31	-2.77	-2.49	-0.40	-0.06	0.62
T= ∞	-3.35	-2.83	-2.54	-0.43	-0.03	0.64

Table 11 Critical values for the t-statistics. t_{η_2} . Only intercept included: I T nDI

Sample size T	Probability that is less than $t_{\eta_2} = \hat{\eta}_2 / \hat{\sigma}_{\hat{\eta}_2}$ entry					
n(I T D)	0.01	0.05	0.1	0.90	0.95	0.99
T=120	-2.25	-1.66	-1.37	0.91	1.24	1.98
T=240	-2.49	-1.84	-1.51	0.87	1.27	1.97
T=480	-2.53	-1.87	-1.54	0.89	1.28	2.01
T= ∞	-2.61	-1.97	-1.62	0.89	1.26	2.04

Table 12 Critical values for the F-statistics. Only intercept included: I T nDI

Sample Size T	Probabilities that $F(\eta_{2i-1}, \eta_{2i}) = (1/2)(t^2_{\eta_{2i-1}} + t^2_{\eta_{2i}})$ is greater than entry, i=2,3,…12.											
	$F(\eta_3, \eta_4)$			$F(\eta_5, \eta_6)$			$F(\eta_7, \eta_8)$			$F(\eta_9, \eta_{10})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	2.05	2.61	4.06	2.02	2.72	4.31	2.15	2.89	4.68	2.27	3.20	5.54
T=240	2.18	2.87	4.30	2.28	3.09	4.96	2.29	3.16	5.28	2.57	3.52	5.95
T=480	2.32	2.96	4.63	2.42	3.25	5.32	2.44	3.28	5.25	2.67	3.68	6.20
T= ∞	2.41	3.12	4.79	2.50	3.39	5.52	2.49	3.28	5.23	2.81	3.87	6.02
	$F(\eta_{11}, \eta_{12})$			$F(\eta_{13}, \eta_{14})$			$F(\eta_{15}, \eta_{16})$			$F(\eta_{17}, \eta_{18})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	2.33	3.23	5.58	2.47	3.53	5.81	2.33	3.29	5.78	2.01	2.67	4.22
T=240	2.56	3.49	5.77	2.67	3.75	6.15	2.61	3.75	6.27	2.23	2.88	4.59
T=480	2.83	3.76	6.49	2.82	3.86	6.16	2.74	3.80	6.49	2.35	3.08	4.73
T= ∞	2.86	3.97	6.32	2.89	3.92	6.71	2.91	4.02	6.72	2.39	3.10	4.96
	$F(\eta_{19}, \eta_{20})$			$F(\eta_{21}, \eta_{22})$			$F(\eta_{23}, \eta_{24})$					
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01			
T=120	2.03	2.67	4.19	2.21	3.04	5.27	2.18	2.97	5.23			
T=240	2.28	2.91	4.43	2.38	3.42	5.40	2.38	3.32	5.57			
T=480	2.38	3.18	4.91	2.60	3.54	5.74	2.65	3.52	5.78			
T= ∞	2.37	3.14	4.90	2.57	3.45	5.80	2.63	3.66	6.16			

Table 13 Critical values for the t-statistics. t_{η_1}. Intercept and trend included: I T nD

Sample size T	Probability that is less than $t_{\eta_1} = \hat{\eta}_1 / \hat{\sigma}_{\hat{\eta}_1}$ entry					
	0.01	0.05	0.10	0.90	0.95	0.99
T=120	-3.60	-3.09	-2.80	-1.00	-0.70	-0.12
T=240	-3.76	-3.27	-2.97	-1.11	-0.82	-0.21
T=480	-3.93	-3.34	-3.06	-1.22	-0.89	-0.28
T= ∞	-3.89	-3.36	-3.08	-1.24	-0.93	-0.33

Table 14 Critical values for the t-statistics. t_{η_2} Intercept and trend included: I T nD

Sample size T	Probability that $t_{\eta_2} = \hat{\eta}_2 / \hat{\sigma}_{\hat{\eta}_2}$ is less than entry					
	0.01	0.05	0.1	0.90	0.95	0.99
T=120	-2.28	-1.65	-1.34	0.90	1.29	1.94
T=240	-2.37	-1.79	-1.47	0.90	1.27	1.96
T=480	-2.52	-1.89	-1.56	0.88	1.26	2.00
T= ∞	-2.60	-1.97	-1.61	0.88	1.26	2.04

Table 15 Critical values for the F-statistics. Intercept and trend included: I T nD

	P robabilities that $F(\eta_{2i-1}, \eta_{2i}) = (1/2)(t^2_{\eta_{2i-1}} + t^2_{\eta_{2i}})$ is greater than entry, i=2,3,…12											
	$F(\eta_3, \eta_4)$			$F(\eta_5, \eta_6)$			$F(\eta_7, \eta_8)$			$F(\eta_9, \eta_{10})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	1.98	2.54	3.93	2.01	2.72	4.49	2.04	2.71	4.48	2.17	3.01	5.32
T=240	2.16	2.84	4.27	2.25	3.04	4.95	2.28	3.10	4.86	2.49	3.56	5.80
T=480	2.33	3.02	4.57	2.34	3.21	5.34	2.43	3.26	5.51	2.71	3.74	6.07
T= ∞	2.40	3.11	4.79	2.50	3.38	5.50	2.49	3.27	5.21	2.81	3.86	6.02
	$F(\eta_{11}, \eta_{12})$			$F(\eta_{13}, \eta_{14})$			$F(\eta_{15}, \eta_{16})$			$F(\eta_{17}, \eta_{18})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	2.25	3.15	5.15	2.57	3.61	6.21	2.33	3.22	5.60	1.97	2.60	4.05
T=240	2.56	3.67	6.14	2.68	3.70	6.46	2.62	3.64	6.22	2.56	2.85	6.14
T=480	2.68	3.73	6.47	2.75	3.79	6.39	2.79	3.89	6.52	2.38	3.08	4.76
T= ∞	2.86	3.94	6.33	2.89	3.91	6.65	2.90	4.00	6.72	2.39	3.09	4.95
	$F(\eta_{19}, \eta_{20})$			$F(\eta_{21}, \eta_{22})$			$F(\eta_{23}, \eta_{24})$					
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01			
T=120	2.04	2.72	4.32	2.13	2.92	4.84	2.17	2.96	5.02			
T=240	2.22	2.93	4.42	2.38	3.11	5.43	2.43	3.36	5.58			
T=480	2.34	3.10	4.90	2.49	3.43	5.52	2.62	3.62	5.91			
T= ∞	2.37	3.13	4.89	2.56	3.45	5.80	2.63	3.66	6.01			

Table 16 Critical values for the t-statistics. t_{η_1}. Intercept and 23 dummy variables included: I nT D

Sample size T	Probability that is less than $t_{\eta_1} = \hat{\eta}_1 / \hat{\sigma}_{\hat{\eta}_1}$ entry					
	0.01	0.05	0.10	0.90	0.95	0.99
T=120	-2.99	-2.34	-2.04	-0.23	0.10	0.69
T=240	-3.12	-2.62	-2.31	-0.39	-0.06	0.59
T=480	-3.25	-2.69	-2.42	-0.40	-0.07	0.57
T= ∞	-3.30	-2.78	-2.51	-0.43	-0.05	0.57

Table 17 Critical values for the t-statistics. t_{η_2} Intercept and 23 dummy variables included: I nT D

Sample size T	Probability that $t_{\eta_2} = \hat{\eta}_2 / \hat{\sigma}_{\hat{\eta}_2}$ is less than entry					
	0.01	0.05	0.1	0.90	0.95	0.99
T=120	-2.94	-2.39	-2.11	-0.20	0.12	0.78
T=240	-3.23	-2.63	-2.33	-0.35	-0.01	0.61
T=480	-3.34	-2.79	-2.50	-0.38	-0.03	0.60
T= ∞	-3.43	-2.82	-2.51	-0.41	-0.06	0.64

Table 18 Critical values for the F-statistics. Intercept and 23 dummy variables included: I nT D

Sample Size T	P robabilities that $F(\eta_{2i-1}, \eta_{2i}) = (1/2)(t^2_{\eta_{2i-1}} + t^2_{\eta_{2i}})$ is greater than entry, i=2,3,…12.											
	$F(\eta_3, \eta_4)$			$F(\eta_5, \eta_6)$			$F(\eta_7, \eta_8)$			$F(\eta_9, \eta_{10})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	3.65	4.49	6.21	4.18	5.15	7.54	4.60	5.66	8.17	5.77	7.13	9.99
T=240	4.69	5.50	7.34	5.35	6.46	8.97	5.79	7.15	10.06	7.45	8.91	12.30
T=480	5.15	6.10	8.16	7.35	8.67	10.12	6.28	7.65	10.60	8.14	9.65	13.09
T= ∞	5.42	6.42	8.48	6.47	7.77	10.66	6.58	7.98	10.99	8.57	10.37	14.52
	$F(\eta_{11}, \eta_{12})$			$F(\eta_{13}, \eta_{14})$			$F(\eta_{15}, \eta_{16})$			$F(\eta_{17}, \eta_{18})$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
T=120	5.32	6.62	9.65	6.16	7.60	10.76	5.66	7.04	10.20	3.64	4.46	6.39
T=240	6.94	8.57	11.99	7.78	9.44	13.10	7.43	8.84	12.51	4.77	5.76	7.89
T=480	7.69	9.27	12.73	8.54	10.14	13.82	7.98	9.72	12.97	5.14	6.20	8.51
T= ∞	8.03	9.76	13.46	9.00	11.02	15.10	8.58	10.31	14.51	5.50	6.60	8.99
	$F(\eta_{19}, \eta_{20})$			$F(\eta_{21}, \eta_{22})$			$F(\eta_{23}, \eta_{24})$					
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01			
T=120	4.18	5.07	7.04	4.90	6.11	8.79	5.17	6.40	9.48			
T=240	5.15	6.16	8.36	6.44	7.85	10.68	6.32	7.63	10.97			
T=480	5.77	6.93	9.58	7.01	8.48	11.32	6.96	8.51	12.09			
T= ∞	6.06	7.25	9.78	7.61	9.07	12.48	7.39	8.94	12.33			