

استخدام التحليل في مجال التردد لنمذجة بيانات مبيعات الغاز الطبيعي في الولايات المتحدة الأمريكية بواسطة شكل الدورة وتحليل سلسلة فوريير

أمل السر الخضر

صفية عبد الله الفاضل

جامعة السودان للعلوم والتكنولوجيا || السودان

عبيد محمود محسن الزوبي

جامعة جيهان || السليمانية || العراق || كوردستان

الملخص: هذه الورقة تأخذ في الاعتبار تطبيق شكل الدورة Periodogram وتحليل سلسلة فوريير Fourier Series Analysis وهي من الطرق اللامعلمية في التحليل في مجال التردد Frequency domain أو التحليل الطيفي spectral analysis للسلاسل الزمنية وذلك باستخدام بيانات مبيعات الغاز في الولايات المتحدة الأمريكية في المدة من 1993-2014 والهدف الأساسي هو تشخيص دورة التقلبات وتقدير النموذج المناسب للبيانات وإيجاد قيم تنبؤات للمستقبل ولتحقيق هذه الأهداف تم الحصول على البيانات ومن ثم استخدمت طريقة شكل الدورة وسلسلة فوريير لتحليل البيانات وبناء على التحليل وجد أن دورة التقلبات خلال المدة قيد الدراسة كانت تقريبا 135 شهراً، وقد تم تقدير نموذج سلسلة فوريير للبيانات ذو كفاءة Accuracy عالية يشتمل على مكونات الاتجاه والموسمية والأخطاء وقد استخدمت معايير RMSE، MASE وMAE للتأكد من كفاءة النموذج ومن ثم تم استخدام النموذج للتنبؤ بأسعار الغاز لمدة ستة أشهر وقد ظهرت جودة التنبؤ أيضا عن طريق إيجاد 90% 95% فترات ثقة للتنبؤات. إضافة إلى ذلك تم تقديم تحليل في مجال الزمن لسلسلة البيانات وقد استخدمت طريقة بوكس جنكيز للحصول على نموذج ARMA المناسب وإيجاد التنبؤات واخيرا تم إجراء مقارنة بين مقاييس الدقة لطريقتي التحليل في مجال الزمن والتحليل في مجال التردد من خلال النتائج المتحصل عليها فوجد أن طريقة مجال التردد تنافس طريقة مجال الزمن والفرق طفيف في الكفاءة.

الكلمات المفتاحية: التحليل الطيفي، تحليل سلسلة فوريير، شكل الدورة، مجال التردد، التنبؤ بالسلسلة الزمنية

المقدمة Introduction {4}

ظهرت الطرق التي تم استخدامها في هذه الدراسة منذ بدايات القرن الماضي حيث إنه في العام 1908 استطاع العالم Joseph Fourier تمثيل السلسلة الزمنية على صورة متسلسلة لا نهائية من دوال جيب Sine وجيب Cosine التمام حيث سمي هذا التحليل بتحليل سلسلة فوريير تقديرا لجهوده. وفي العام 1926 عرض الباحث Yule نماذج الانحدار الذاتي Auto regressive models كما وضع العالم Slutsky في العام 1937 نموذج المتوسطات المتحركة وفي بداية التسعينيات من نفس القرن قدم العالمان بوكس وجنكيز منهجيهما في نمذجة بيانات السلاسل الزمنية والتنبؤ بها والتي أصبحت منذ ذلك الوقت الأداة الأكثر قبولا وشيوعا في الأوساط العلمية والنظرية والتطبيقية واثبتت هذه المنهجية كفاءتها العالية مما أدى إلى اعتبارها المدخل الأساسي لتحليل السلاسل الزمنية وقد ظهرت العديد من الاساليب الحديثة بعد ذلك منها طريقة الشبكات العصبية الاصطناعية وطريقة الجار الأقرب المضطرب للتنبؤ بالسلاسل الزمنية.

الطرق الرئيسية التي تم تطبيقها في هذه الدراسة هي شكل الدورة periodogram وتحليل سلسلة فوريير Fourier series. في حال كانت السلسلة الزمنية تتأثر إلى حد كبير بالموسمية أو الدورية يمكن مباشرة تخمين الدورة period أو التردد frequency للسلسلة الزمنية ولكن إذا كانت الدورة لا يمكن التنبؤ أو تخمينها بصورة دقيقة ذلك يؤدي إلى ضرورة بناء شكل الدورة the periodogram لتحديد الدورة أو التردد للسلسلة الزمنية

الهدف الأساسي في هذه الورقة هو تشخيص دورة التقلبات، تقدير النموذج المناسب للبيانات وإيجاد قيم التنبؤات لبيانات الغاز الطبيعي في الولايات المتحدة الأمريكية باستخدام طريقي شكل الدورة وتحليل سلسلة فورير التي تعتبر من الطرق اللامعلمية لتحليل السلاسل الزمنية وبما أن الأسعار تتأثر بعوامل متعددة كان لابد من استخدام تحليل شكل الدورة لتحديد دورة الأسعار للسلسلة خلال الفترة قيد الدراسة، كما هدفت الورقة إلى عقد مقارنة لنتائج طريقة سلسلة فورير مع نتائج نماذج ARIMA وهي من الطرق المعلمية لتحليل السلاسل الزمنية في مجال الزمن. كما تتمثل أهمية الدراسة في تفرداها باستخدام طريقة نماذج فورير كطريقة معلمية للتنبؤ ببيانات السلاسل الزمنية والتي لم تلق حظها في الاستخدام والتطبيق على مستوى الدراسات باللغة العربية، وكذلك الأمر بالنسبة لطريقة شكل الدورة periodgram كطريقة غير معلمية.

مشكلة البحث

يمكن كتابة مشكلة البحث على هيئة الاسئلة التالية:

ما هو النموذج الاحصائي الأمثل للتنبؤ بمبيعات الغاز الطبيعي في الولايات المتحدة الأمريكية للبيانات المأخوذة على الفترة (1993-2014)

هل يوجد مكون دوري في البيانات ام لا؟؟ وما قيمة مدة التقلبات الدورية إذا وجدت؟؟

فرضية البحث

الفرض الأساسي لهذه الدراسة أن طريقة التحليل في مجال التردد ممثلة في استخدام نموذج سلسلة فورير تعطي جودة أعلى فيما يتعلق بدقة التنبؤات عند مقارنتها مع طريقة التحليل في مجال الزمن ممثلة في استخدام نماذج ARIMA وذلك بناء على مقاييس RMSE ، MAE ، MASE كما وتفترض عدم وجود مكون دوري في البيانات.

الإطار النظري Theoretical framework { [3] ، [4] [5] }

من أجل بناء شكل الدورة يتم عرض السلسلة الزمنية في مجال التردد frequency-domain بمعنى عرض النقاط وكأنها مكونة من موجات جيب sine وجيب تمام cosine عند ترددات frequencies مختلفة فيكون النموذج على الصورة

$$X_t = T_t + \sum_{k=1}^{N/2} [\alpha_k \cos(2\pi f_k t) + \beta_k \sin(2\pi f_k t) + \epsilon_t] \quad t = 1, 2, \dots, N \quad (1)$$

حيث أن

T_t تمثل معادلة مكون الاتجاه

X_t قيمة أسعار الغاز عند الزمن t

$2\pi f_k$ تمثل التردد الزاوي الذي يقاس بالراديان radians

ϵ_t عبارة عن الخطأ عند الزمن t

α_k, β_k عبارة عن المعاملات coefficients التي يجب تقديرها

اذن فالسلسلة تتكون من مكونات الاتجاه والموسمية والأخطاء، يتم أولاً استبعاد الاتجاه ويتم تطبيق طريقة

المربعات الصغرى لتقدير المعلمات للسلسلة بعد ابعاد الاتجاه

حيث أن

$$\hat{\alpha}_k = \frac{2}{N} \sum_{t=1}^N X_t \cos(2\pi f_k t) \quad k = 1, 2, \dots, M \quad (2)$$

$$\hat{\beta}_k = \frac{2}{N} \sum_{t=1}^N X_t \sin(2\pi f_k t) \quad k = 1, 2, \dots, (3) \text{ و}$$

وبذلك يمكن تعريف شكل الدورة للسلسلة الزمنية ΔX_t حيث $N = 2q + 1$ (القيم الفردية من المشاهدات) يعرف كدالة من الكثافة intensities يرمز لها بالرمز $I(f_i)$ عند التردد f_i بالصيغة الآتية

$$I(f_i) = N/2 [\alpha_i^2 + \beta_i^2] \quad i = 1, 2, \dots, q \quad (4)$$

حيث إن f_i عبارة عن التوافق رقم i للتردد $\frac{1}{N}$ حيث أن $0 \leq f_i \leq 0.5$

وبالنسبة للملاحظات الزوجية يكون لدينا

$$I(f_{0.5}) = N\alpha_q^2 \quad (5)$$

والشكل البياني للكثافة يتم برسم $I(f_i)$ على محور X مع التردد f_i على محور Y أو مع الدورات periods وهي تساوي $\frac{1}{f_i}$ هي ما يعرف بشكل الدورة.

يعمل شكل الدورة على قياس القدرة power للسلسلة الزمنية لكل الترددات الممكنة واطوال الموجات. ويمكن تفسير ذلك على أنه مقدار مجموع المربعات الكلي التي يتم تفسيرها من قبل ترددات محددة، الدورة أو period التردد frequency للسلسلة الزمنية يتم تشخيصه بواسطة الكثافة intensity الأكبر وبعد ذلك يمكن استخدام تردد فورير للحصول على تقديرات المعلمة للنموذج. ونلاحظ هنا أن الدورة المتحصل عليها من شكل الدورة تعطي دورة الأسعار.

يتم تقدير النموذج فيكون على الصورة

$$\hat{X}_t = \hat{T}_t + \sum_{i=1}^k [\hat{\alpha}_k \cos(\omega_i t) + \hat{\beta}_k \sin(\omega_i t) +] \quad i = 1, 2, \dots, N \quad (6)$$

حيث أن

\hat{X}_t عبارة عن القيمة المقدرة للسلسلة الزمنية

\hat{T}_t عبارة عن معادلة الاتجاه المقدرة و $\hat{\alpha}_k, \hat{\beta}_k$ عبارة عن معلمة النموذج المقدرة

$\omega_i = 2\pi f_i$ تمثل التردد الزاوي

K تمثل التوافق harmonic الأعلى ل ω

أعلى توافق K في تحليل سلسلة فورير هو عدد المشاهدات في الموسم مقسوم على 2 بالنسبة للملاحظات

الزوجية أو مقسوما على $(n - 1)/2$ بالنسبة للملاحظات الفردية.

أولا يتم فصل أو استبعاد مكون الاتجاه عن طريق تقدير نموذج الاتجاه الخطي أو التربيعي أو المتوسط

العام للبيانات باستخدام طريقة المربعات الصغرى least squares ويتم تقدير معادلة الاتجاه بناءً على معنويتها

significance في النموذج لذا فإن معادلة الاتجاه تشتمل على

$$T_t = a_0 + a_1 t + a_2 t^2 + \dots + a_p t^p \quad (7)$$

حيث أن $a_0 = \mu = \frac{\sum X_t}{N}$ وبعد تقدير مكون الاتجاه يتم استبعاده من السلسلة الزمنية وتستخدم

السلسلة الزمنية الخالية من الاتجاه detrended في تقدير التغيرات الموسمية seasonal variation.

دوال الجيب وجيب التمام المعطاة في المعادلة (1) تعطي النموذج المقدر للمكون الموسمي للنموذج كالاتي

$$\Delta \hat{X}_t = \hat{T}_t + \sum_{i=1}^k [\hat{\alpha}_k \cos(\omega_i t) + \hat{\beta}_k \sin(\omega_i t) +] \quad i = 1, 2, \dots, N \quad (8)$$

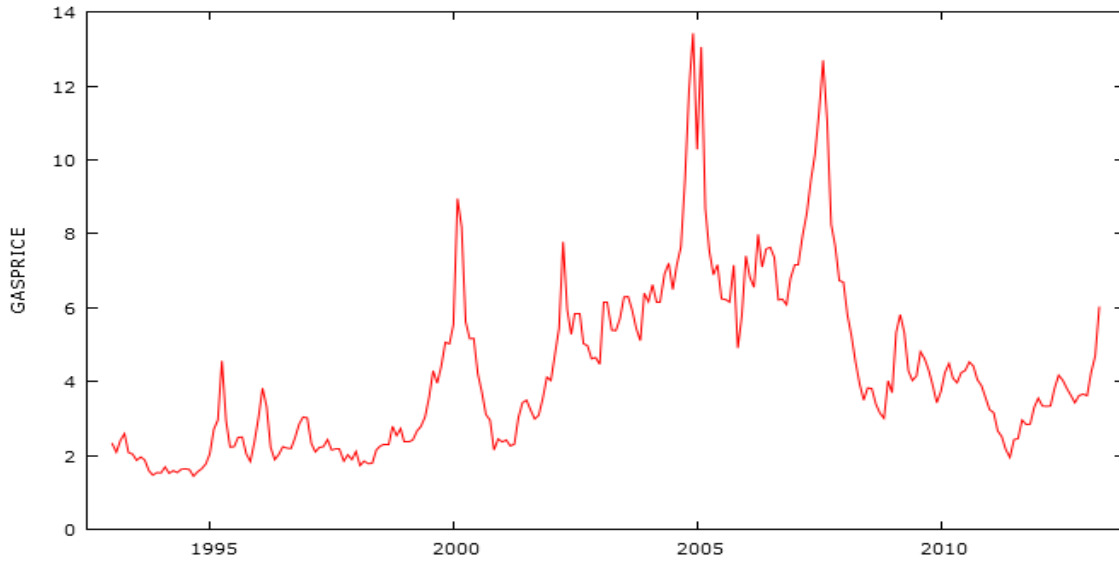
حيث إن $\Delta \hat{X}_t = \hat{X}_t - \hat{T}_t$ المعادلة اعلاه عبارة عن نموذج انحدار خطي متعدد من أجل الحصول على التقديرات $\hat{\alpha}_k, \hat{\beta}_k$.

أدوات وطرق البحث

تم استخدام برنامج R وبرنامج GRETEL كأدوات لإجراء طريقة التحليل في مجال التردد للبيانات والتوليفة من المكونات الثلاثة للسلسلة الزمنية - هي الاتجاه والموسمية والمكون العشوائي - نتج عنها النموذج العام لسلسلة فوريير لتقدير أسعار المبيعات ومن ثم تم استخدام النموذج لتوليد تنبؤات لأسعار الغاز الطبيعي في الولايات المتحدة كما تم إيجاد شكل الدورة وتحديد دورة المبيعات واخيرا تم الحصول على نتائج طريقة التحليل في مجال الزمن ولاختبار دقة النماذج والمقارنة بينها تم استخدام معايير RMSE، MAE، MASE.

تحليل البيانات { Data Analysis [2]، [1] [5] }

يبدأ تحليل السلسلة الزمنية عادة بالفحص البصري للرسم البياني للقراءات الأصلية مع الزمن كما يتضح ذلك في شكل (1)



شكل (1) رسم بياني لسلسلة بيانات مبيعات الغاز الطبيعي GASPRICE

يحتوى الشكل على مكون الاتجاه العام والذي يظهر على شكل تزايد مطرد في المبيعات حتى العام 2008 تقريبا ومكون التغيرات الموسمية وتظهر على شكل نتوءات خلال السنة الواحدة إضافة إلى مكون الدورية والتي تبدو على هيئة قمم peaks كل خمسة سنوات وبسبب وجود هذه المكونات يتضح أن البيانات غير ساكنة فلا بد من إجراء تحويله للبيانات للحصول على السكون وتم استخدام التحويلة اللوغرتمية كما يظهر في الشكل (2) كما تم استخدام اختبار السكون على البيانات فكانت النتائج كالآتي:

data: GASPRICE

p-value = 0.1433، Lag order = 6، Dickey-Fuller = -3.0281

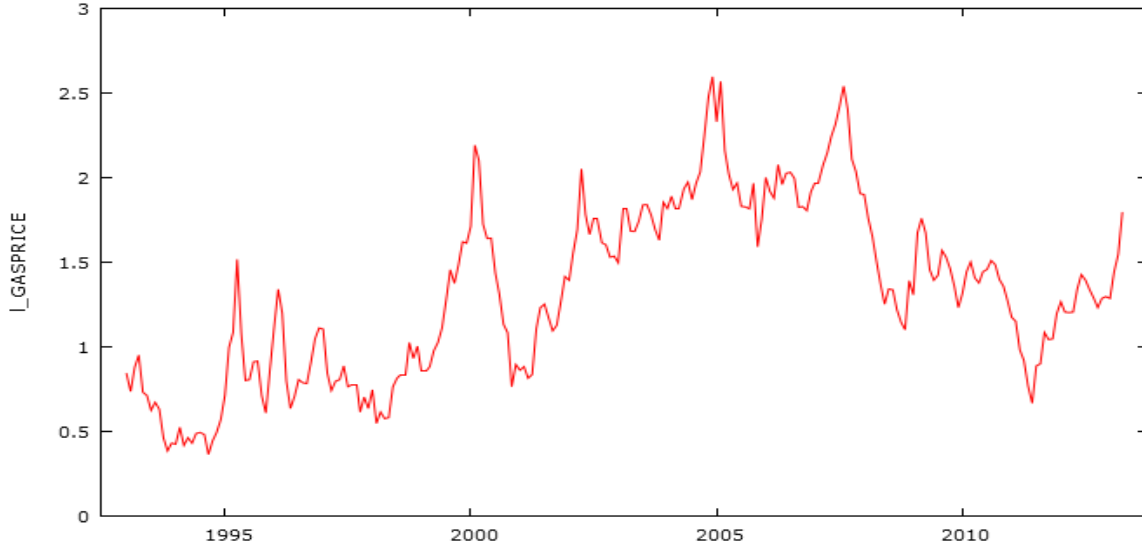
alternative hypothesis: stationary

وأيضاً تم استخدام اختبار KPSS

data: GASPRICE

KPSS Level = 2.4424 ، Truncation lag parameter = 3 ، p-value = 0.01

ويظهر بوضوح حسب قيمة p-value بالنسبة للاختبارين عدم سكون السلسلة الزمنية لان الفرض البديل في اختبار KPSS هو عدم سكون البيانات بعكس إختبار Dickey-Fuller الذي يكون الفرض البديل فيه هو سكون السلسلة الزمنية.



شكل (2) رسم للسلسلة الزمنية L-GASPRICE بعد إجراء التحويلة اللوغرتمية على البيانات الأصلية (إعداد الباحثين باستخدام برنامج GRETEL) حيث نلاحظ أن تطبيق تحويلة اللوغرثم على البيانات قد ساعد على التخلص من مشكلة عدم السكون الذي كانت البيانات الأصلية تعاني منه وضح ذلك في الرسم البياني بالتخلص من الاتجاه العام والتقلص الواضح في التواءات والقمم مقارنة بما كان في شكل (1) وقد تم استخدام الاختبارات مرة ثانية للتأكد من سكون البيانات بعد التحويل فكانت النتائج

Dickey-Fuller = -5.8028 ، Lag order = 6 ، p-value = 0.01

alternative hypothesis: stationary

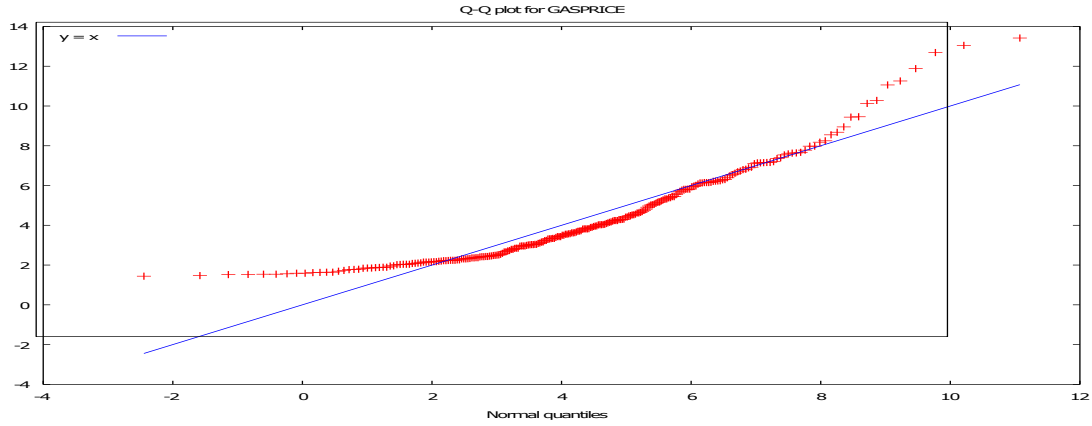
أما اختبار KPSS فيعطي الناتج

KPSS Level = 0.10651 ، Truncation lag parameter = 3 ، p-value = 0.1

(إعداد الباحثين باستخدام برنامج R)

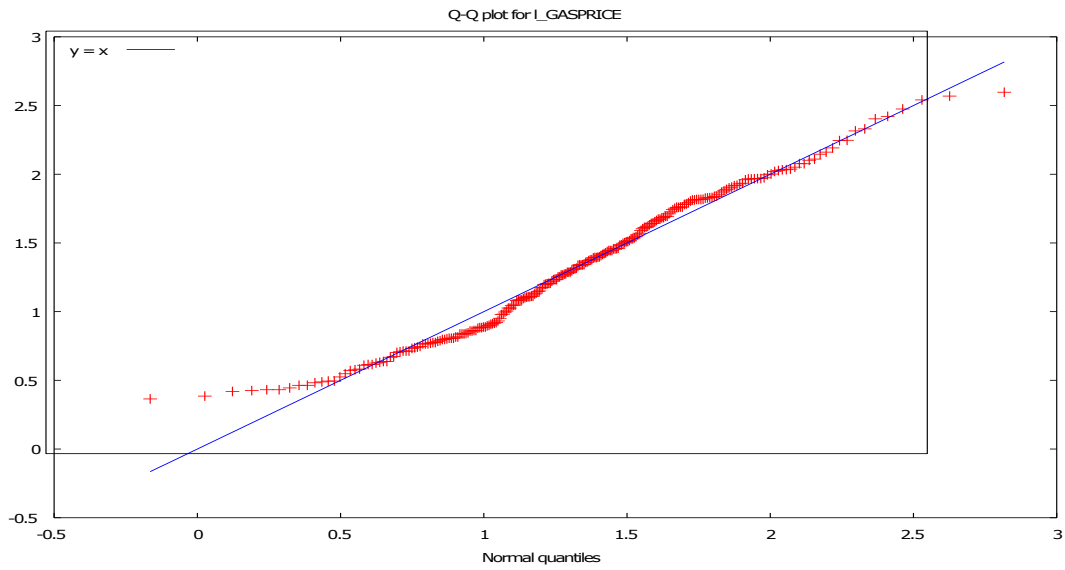
وهنا يتفق الاختباران على سكون بيانات السلسلة الزمنية بعد تطبيق تحويلة اللوغرثم عليها. وقد استخدم

الرسم البياني لأشكال Q-Q plot لفحص التوزيع الطبيعي للبيانات كما يظهر في شكل (3)



شكل (3) يمثل رسم بياني Q-Q plot البيانات الأصلية

(إعداد الباحثين باستخدام برنامج GRETEL) وحيث أن أشكال Q-Q plot يتم استخدامها للتأكد من أن البيانات تتوزع توزيع طبيعي Normal distribution يتضح في الشكل أن نقاط البيانات الأصلية تقع تقريبا على شكل خط مستقيم مما يعني أن البيانات تنحرف قليلا عن التوزيع الطبيعي. وبعد تطبيق التحويلة اللوغارتمية تم الحصول على شكل (4)



شكل (4) يوضح الرسم البياني للتوزيع الطبيعي Q-Q plot بعد إجراء تحويل البيانات

(إعداد الباحثين باستخدام برنامج GRETEL) ويظهر الشكل انتظام البيانات حول الخط المستقيم بصورة أكثر دقة من الشكل (3) مما يدل على أن استخدام تحويل اللوغارتم نتج عنه أن البيانات تتوزع توزيع طبيعي

نتائج التحليل في مجال التردد $\{[1],[2]\}$

أولاً: تحليل شكل الدورة Periodogram Analysis

تم الحصول على الكثافة intensities الممثلة بـ $I(f_i)$ عند مختلف الترددات باستخدام المعادلات 2، 3، و4 كما في (مخرجات 1). والرسم البياني للكثافة مقابل التردد frequencies ومقابل الدورات periods كما في الشكلين 4 و

5 ويتضح من الشكل أن الكثافة الاعلى حدثت عند 135 شهر ذلك يعني أن دورة المدة للبيانات تساوي 135 شهر وذلك ما يتم استخدامه لتقدير نموذج البيانات
مخرجات (1) التالية توضح جانب من قيم شكل الدورة لتحليل بيانات مبيعات الغاز بعد إجراء تحويلة اللوغرثم (I_GASPRICE)

مخرجات (1)

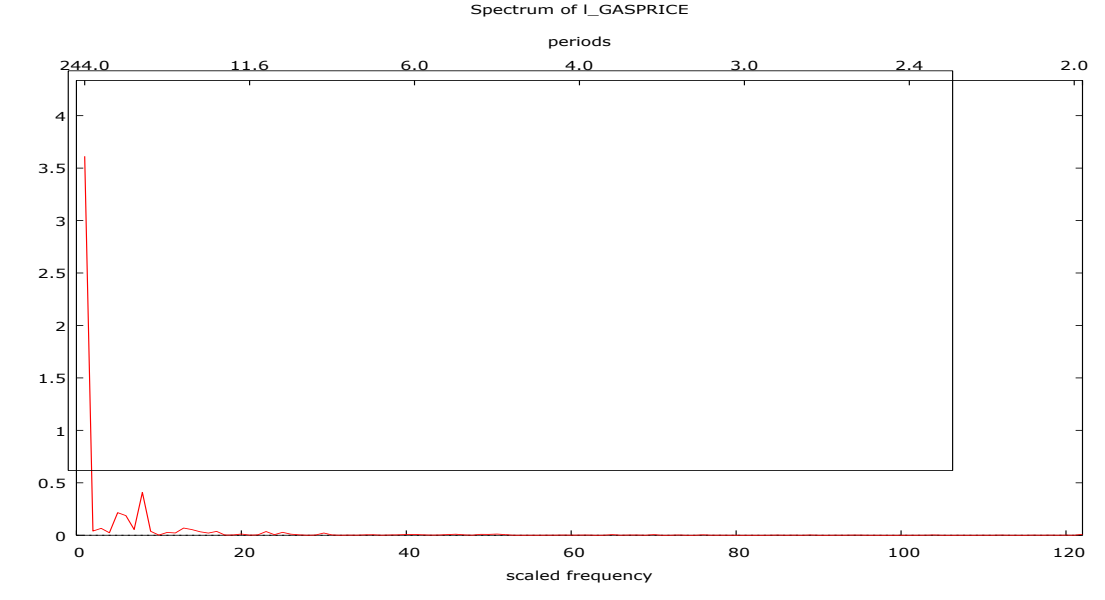
Periodogram for I_GASPRICE

Number of observations = 244

omega scaled frequency periods spectral density

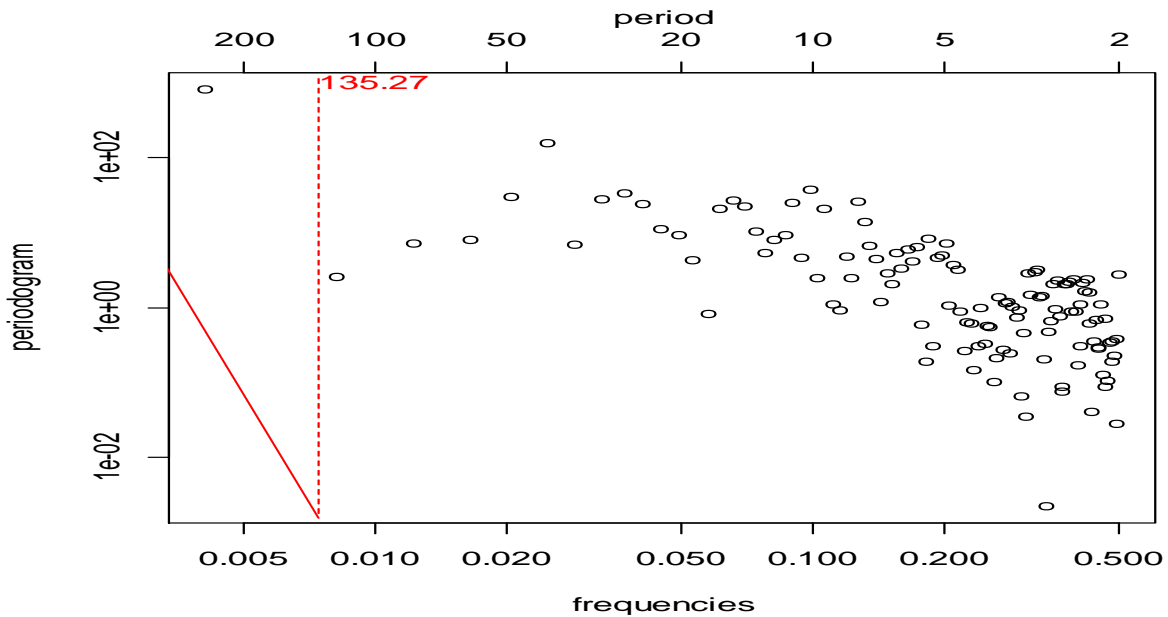
0.02575	1	244.00	3.6129
0.05150	2	122.00	0.039804
0.07725	3	81.33	0.065777
0.10300	4	61.00	0.023724
0.12875	5	48.80	0.21550
0.15450	6	40.67	0.18699
0.18026	7	34.86	0.055010
0.20601	8	30.50	0.40853
0.23176	9	27.11	0.037735
0.25751	10	24.40	0.0017846
0.28326	11	22.18	0.026972
0.30901	12	20.33	0.021739
0.33476	13	18.77	0.068806
0.36051	14	17.43	0.054059
0.38626	15	16.27	0.033043
0.41201	16	15.25	0.020810
0.43776	17	14.35	0.036791
0.46351	18	13.56	0.0019022
0.48926	19	12.84	0.0029862
0.51502	20	12.20	0.011667

(إعداد الباحثين باستخدام برنامج R) واضح من المخرجات أن لدينا 244 مشاهدة وتم حساب دالة الكثافة للطيف spectral density لعدد من الترددات scaled frequency المرقمة من 1 إلى 20 والعمود الأول يمثل التردد الزاوي ω بمعنى التردد بمقياس الراديان والعمود period يعطي الدورات المقابلة لكل تردد مقاسة بالشهور.



شكل (5) يمثل periodgram لبيانات مبيعات الغاز المحولة (frequency/period)

(إعداد الباحثين باستخدام برنامج R) وقد تم استخدام نتائج (مخرجات 1) للحصول على الرسم البياني في شكل (5) حيث الخط الأفقي الأسفل يمثل الترددات والخط الأفقي الأعلى يمثل الدورات periods كما تمثل القراءة على الخط الرأسي قيم كثافة الطيف spectral density عند كل تردد والنتوء البارز أو القمة peak على الرسم يوضح معنوية المكون الدوري عند هذه التردد.



شكل (6) يوضح التردد المهيمن على السلسلة يساوي تقريبا 0.007 ويوافق الفترة 135.25 شهر

(إعداد الباحثين باستخدام برنامج R) باستخدام مخرجات (1) في هذا الرسم تم تحديد المدة الزمنية التي تحدث عندها الدورية في بيانات السلسلة بوضع الخط الرأسي باللون الاحمر علي قيمة الدورة والتي تحدث كل 135 شهر.

ثانيا: التقدير والتنبؤ لنموذج العام لسلسلة فوريير Fitting the General Fourier Series Model

مخرجات 2

Forecast method: Linear regression model

Model Information:

`tslm(formula = IGASPRICE ~ trend + fourier(1GASPRICE, 1))`

Coefficients:

(Intercept)	trend
3.371823	0.007308
fourierS1-12	fourierC1-12
0.094388	-0.059535

Error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-9.426299e-17	0.7652809	0.5989277	-Inf	Inf	0.8774864	0.8323316

Forecasts:

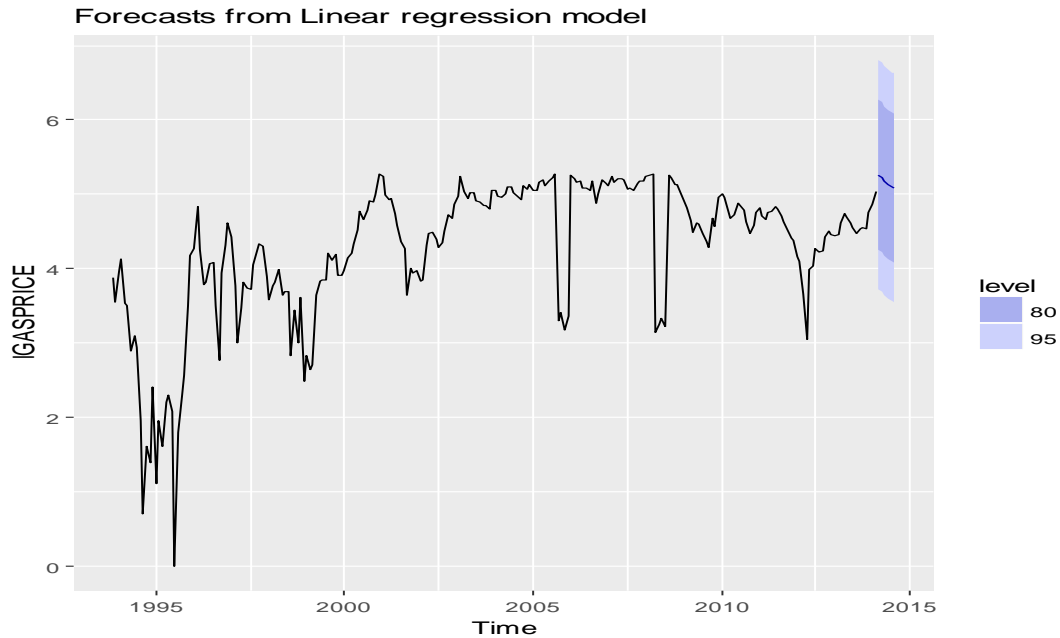
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Mar 2014	5.260929	4.256963	6.264894	3.721962	6.799895
Apr 2014	5.229018	4.224895	6.233141	3.689810	6.768227
May 2014	5.181155	4.176955	6.185355	3.641829	6.720481
Jun 2014	5.132123	4.127903	6.136342	3.592767	6.671479
Jul 2014	5.097017	4.092796	6.101239	3.557658	6.636377
Aug 2014	5.087203	4.082970	6.091436	3.547827	6.626579

(إعداد الباحثين باستخدام برنامج R) مخرجات (2) عبارة عن نتائج تحليل نموذج من نوع الانحدار الخطي على شكل سلسلة فوريير المكون من القاطع $intersept$ و الاتجاه $trend$ وتظهر قيم معاملات فوريير $fourierS1-12$ ، الجزء الاخير من المخرجات على قيم التنبؤ للمبيعات لخمس فترات زمنية وقد تم الحصول عليها بتطبيق نموذج سلسلة فوريير

يمكن كتابة النموذج في صورته النهائية بعد الاستفادة من (مخرجات 2) بتعويض قيم المعلمات $\hat{\alpha}_1, \hat{\beta}_1$ التي تم الحصول عليها من نتائج التحليل كالآتي

$$\log X_t = 3.3718 + 0.0073t + 0.094\cos\omega t - 0.059\sin\omega t \quad (9)$$

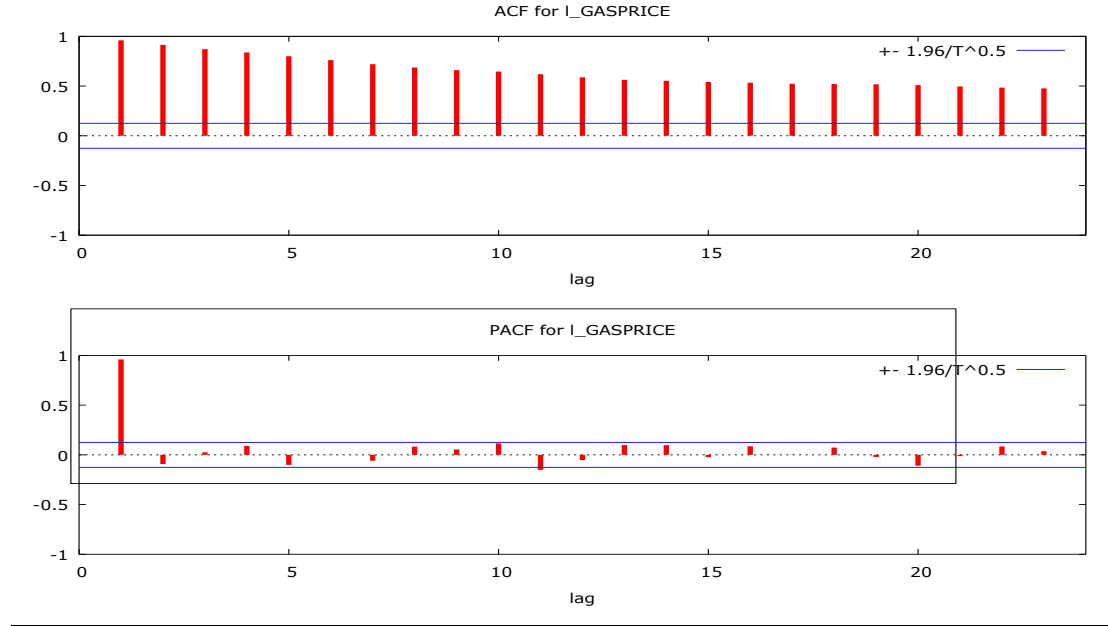
والشكل (7) التالي يوضح الرسم للسلسلة الزمنية والتنبؤات مع 95% و 80% فترات الثقة كالآتي



الشكل (7) الرسم للسلسلة الزمنية والتنبؤات مع 95% و 80% فترات الثقة

(إعداد الباحثين باستخدام برنامج R) حيث يظهر اللون الازرق فترة ثقة 80% للتنبؤات الست التي تم الحصول عليها في (مخرجات 2) كما يوضح اللون الرمادي بوضوح 90% فترة ثقة لنفس التنبؤات في مخرجات 3 ونلاحظ انخفاض المبيعات الشهرية خلال الستة اشهر التي تم التنبؤ بها.

نتائج طريقة التحليل في مجال الزمن $\{[1],[2]\}$



شكل (8) رسم دوال ACF و pacf بعد إجراء التحويلة اللوغرتمية

(إعداد الباحثين باستخدام برنامج GRETEL) بناءً على دالتي الارتباط الذاتي في شكل ACF(9) التي تستخدم لتشخيص نموذج المتوسطات المتحركة MA ودالة PACF التي تستخدم لتشخيص نموذج الانحدار الذاتي AR يتم اقتراح نموذج $ARMA(1,0,1)$.

تحليل البيانات باستخدام نموذج $ARMA(1,0,1)$

مخرجات 3

`arma(x = IGASPRICE ,order = c(1 ,0 ,1))`

Coefficients:

ar1 ma1 intercept

0. 9026 -0. 0906 4. 2793

s. e. 0. 0301 0. 0689 0. 2475

sigma^2 estimated as 0. 1843: log likelihood = -140. 68 , aic = 289. 36

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

0. 002798724 0. 4293453 0. 2432592 -Inf Inf 0. 9996888 -0. 0006598921

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Mar 2014 4. 940534 4. 390306 5. 490762 4. 099033 5. 782035

Apr 2014 4. 876126 4. 167366 5. 584887 3. 792170 5. 960082

May 2014 4. 817992 4. 002551 5. 633433 3. 570883 6. 065101

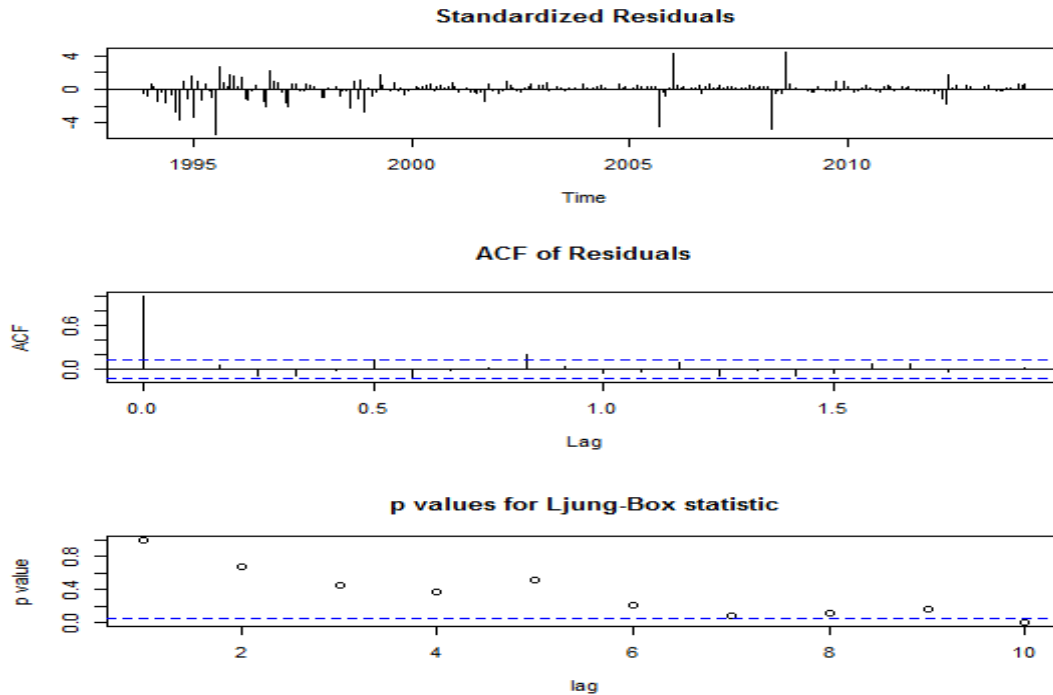
Jun 2014 4. 765521 3. 872542 5. 658500 3. 399827 6. 131214

Jul 2014 4. 718160 3. 766672 5. 669649 3. 262984 6. 173336

Aug 2014 4. 675413 3. 678794 5. 672032 3. 151216 6. 199610

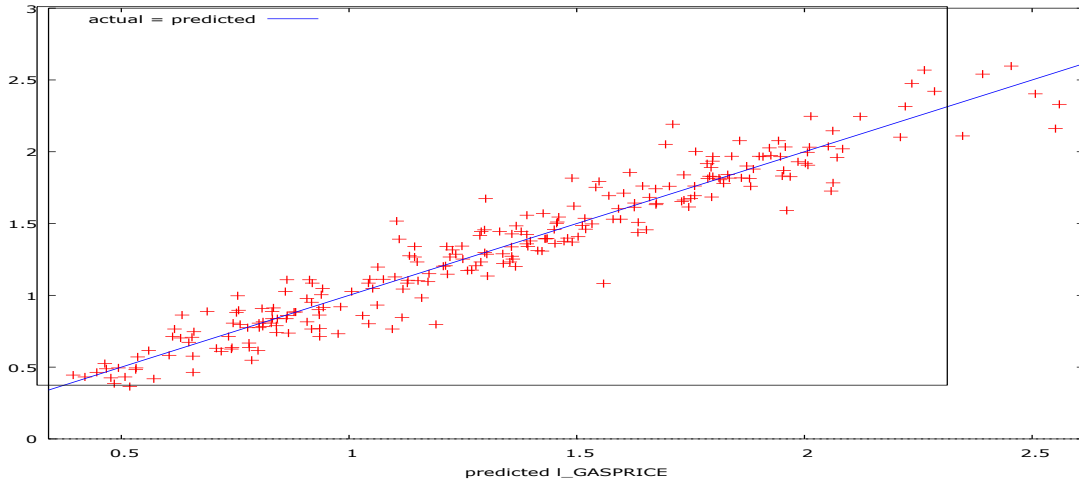
Feb 2016 4. 341916 3. 171411 5. 512422 2. 551783 6. 132050

(إعداد الباحثين باستخدام برنامج R) توضح النتائج في مخرجات (3) انه تم تطبيق نموذج (1) ARIMA، 0،
1) ومعاملات Coefficients النموذج تظهر مع قيم الأخطاء المعيارية التابعة لها، وكانت قيمة التباين للنموذج
 $\sigma^2 = 0.1843$ كما تم الحصول على قيم معايير الكفاءة، RMSE، MAE، MAPE، MPE، MASE وتدل القيم
لهذه المعايير على جودة النموذج لأنها جميعا اصغر من الواحد.



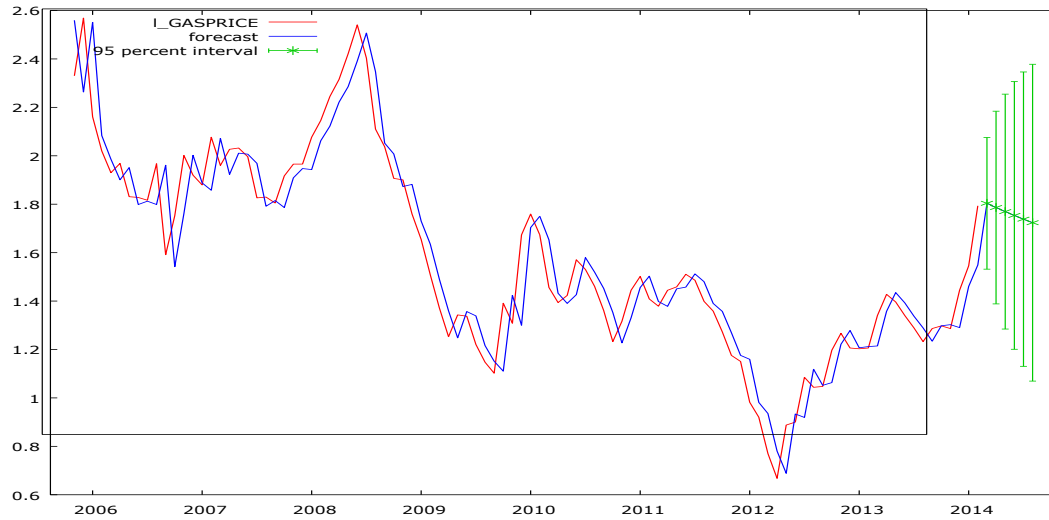
الشكل (9) يعطي الفحص التشخيصي Diagnostic Checking للتحليل باستخدام البواقي

(إعداد الباحثين باستخدام برنامج R) يتضح من الشكل (9) أن الأخطاء المعيارية Standardized Residuals تتوزع توزيع عشوائي، أما الارتباط الذاتي للبواقي ACF of Residuals والتي تعتبر وسيلة مهمة لفحص ملائمة النموذج عن طريق اختبار عشوائية الأخطاء للبواقي والرسم يؤكد جودة النموذج حيث $\rho=1$ عند $k=0$ وتساوي 0 عدا ذلك كما يظهر قيم p (لاحصاءة Ljung-Box للتباطؤات الـ 10 الأولى على التوالي



شكل (10) رسم يوضح القيم الفعلية والمقدرة باستخدام نموذج ARMA(1,0,1)

(إعداد الباحثين باستخدام برنامج R) يتضح في شكل (10) اقتراب القيم الفعلية من القيم المقدرة



شكل (11) رسم يوضح القيم الفعلية مع القيم المتنبأ بها مع فترة ثقة 95%

(إعداد الباحثين باستخدام برنامج R) حيث يمكن ملاحظة في شكل (11) تقارب القيم الأصلية (ويمثلها الخط الاحمر) من القيم المتنبأ بها باستخدام نموذج سلسلة فوريير والتي تظهر في الرسم باللون الازرق كما تم حساب 95% فترة ثقة بين القيم الأصلية والمتنبأ بها تظهر باللون الأخضر ونلاحظ انخفاض المبيعات للفترات الست المتنبأ بها.

الاستنتاجات والتوصيات Conclusions and recommendations

عملت طريقة تحليل الطيف أو التحليل في مجال التردد على تشخيص دورة بيانات مبيعات الأسعار للغاز الطبيعي باستخدام شكل الدورة periodogram وهي تساوي تقريبا (135) شهر وكان النموذج المقدر لتحليل سلسلة فورير الذي تم تطبيقه على السلسلة الزمنية كما ذكر في الصيغة (9)

$$\log X_t = 3.3718 + 0.0073t + 0.094\cos\omega t - 0.059\sin\omega t$$

وتم استخدام هذا النموذج لتوليد التنبؤات لسته من الفترات الزمنية التالية وكانت ذات كفاءة عالية حسب معايير الدقة المحسوبة مما يعني أن نموذج سلسلة فورير يعمل بصورة جيدة لبناء نموذج مناسب لبيانات السلسلة الزمنية مما يشجع على استخدام هذه الطريقة في التنبؤ.

ايضا تم وضع نموذج ARMA(1,0,1) بناء على نتائج مخرجات (3) لتحليل السلسلة في مجال الزمن على

الصورة

$$X_t = 4.2793 + 0.9026X_{t-1} - 0.0906 \epsilon_{t-1} \quad (10)$$

وبمقارنة مقاييس الدقة للنموذجين وعلى الرغم من تفوق طريقة التحليل في مجال الزمن الا انه من الواضح أن الفرق ليس كبيرا، فيتم اعتماد النموذج في الصيغة (10) على انه النموذج الامثل للتنبؤ بالبيانات بناءً على نتائج مقاييس الدقة.

جدول (1) مقارنة مقاييس الكفاءة لنموذجي سلسلة فورير ونماذج ARIMA

معايير الدقة	RMSE	MAE	MASE
طريقة مجال التردد (سلسلة فورير)	0.7652809	0.5989277	0.8774864
طريقة مجال الزمن (نماذج ARIMA)	0.4293453	0.2432592	0.9996888

واضح عند فحص جدول (1) الذي تم تكوينه من نتائج مخرجات (2) ومخرجات (3) ان كل من طريقي سلسلة فورير ونماذج اربما ARIMA قد عملتا بصورة ممتازة في ما يخص كفاءة النماذج حسب قيم المعايير MASE، MAE، RMSE وعلى الرغم من أن طريقة التحليل في مجال الزمن أظهرت دقة أكبر فإن الفارق ليس كبيرا فيتم بذلك قبول الفرض البديل للفرض الذي تم وضعه في بداية الدراسة وهو أن نماذج ARIMA تظل الافضل من حيث كفاءة قيم معاييرها مقارنة مع قيم نموذج سلسلة فورير كما ويتم قبول الفرض البديل للفرض الذي تم وضعه في بداية الدراسة بوجود تقلبات دورية تحدث تقريبا كل 135 شهر في المبيعات والتي أظهرها تحليل شكل الدورة وهنا تكون التوصيات كالآتي

1. استخدام طريقة سلسلة فورير للتنبؤ بالسلاسل الزمنية كطريقة (مرافقة/ بديلة) لطريقة نماذج اربما حتى يستطيع الباحث أن يؤكد أو ينفي افتراضه أي من النموذجين أفضل لتمثيل البيانات قيد الدراسة.
2. عند تحليل بيانات السلاسل الزمنية ينصح باستخدام طريقة شكل الدورة Periodogram كأداة تشخيص لفحص ما إذا كان هنالك مكون دوري في البيانات أم لا.
3. فيما يتعلق بتحليل السلسلة الزمنية في مجال التردد يمكن تشجيع الباحثين على استخدام طرق معلمية مثل طريقة Burg وطرق لا معلمية مثل Welch's method. والتي تعتبر اصدارة مطورة من طريقة شكل الدورة periodogram. أو استخدام ما يسمى بطرق الفضاء الجزئي Subspace methods لتحليل السلاسل الزمنية

المراجع Bibliography

- Adkins ،L. C. (2010). Using gretl for Principles of Econometrics ،Version 1. 3131.
- Meko ،D. M. (2015). Applied Time Series Analysis ،Lesson notes published on internet ،GEOS 585A ، Spring 2015.
- Omekara ،C. O. ،Ekpenyong ،E. J. ،& Ekerete ،M. P. (2013). Modeling the Nigerian Inflation Rates Using Periodogram and Fourier Series Analysis. CBN journal of Applied Statistics ،4(2) ،51-68.
- Pollock ،D. S. G. ،Green ،R. C. ،& Nguyen ،T. (Eds.). (1999). Handbook of time series analysis ،signal processing ،and dynamics. Elsevier.
- Coghlan ،A. (2014). Using R for time series analysis. 2013-04-05]. <http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/src/timeseries.html>.

Modeling the Gas price data Using Periodogram and Fourier Series Analysis

Abstract: This paper takes into account the application of the Periodogram and the Fourier Series Analysis. It is one of the non-parametric methods of Frequency domain analysis or spectral analysis of time series using Gas sales data in the United States of America from 1993-2014. In order to achieve these objectives ،the data were obtained and then the Periodogram and the Fourier series methods were used to analyze the data. Based on the analysis ،the cycle of variability within the period under study was 135 months ،and a high Accuracy data model was estimated for the Fourier series which included trend ،seasonal and error components. The RMSE ،MASE and MAE standards were used to confirm the efficiency of the model and the model was then used to predict gas sales for six months ،and we have 90% 95% confidence intervals for predictions. In addition ،a time domain analysis was provided for the data series using Bok Jenkins method to obtain the appropriate ARMA model and to generate Predictions. Finally ،a comparison was made between the accuracy measures of the time domain and frequency domain methods The frequency domain method competed with the time domain and the slight difference in efficiency.

Keywords: spectral analysis ،Fourier series analysis ،Periodogram ،frequency domain and time series forecasting