جمهوريـــــــة العراق
وزارة التعليم العالي والبحث العلمي
جـــــامعة
كـــــــــــــــــــــــــــــــربلاء
كلية الادارة و الاقتصـاد
قســـــم الاحصاء

أستعمال بعض طـــــــــــــــــــــــرائق التنبؤ المختلفة
لتحليل اعداد المصابين بالاورام الخبيثة

رسالة مقدمـــة الى مجلس كلية الادارة والاقتصاد في جامعة كربلاء

و هي جزء من متطلبات نيل درجة ماجستير في علوم الاحصاء

تقدم بهـا
صفاء مجيد مطشر الكلابي

بإشـــــــــراف

أ. د. عدنان كريم نجم الدين

1439 هـ **2018 م**

كربلاء المقدسة

بسم الله الرحمن الرحيم

﴿ إِنَّا نَحْنُ نُحْيِي الْمَوْتَى وَنَكْتُبُ مَا قَدَّمُوا وَآثَارَهُمْ وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ فِي إِمَامٍ مُبِينٍ ﴾

صدق الله العلي العظيم

سورة يس

الآية ﴿12﴾

الإهـــــــداء...

الى عائلتي ...

اللهم اسبق عائلتي العافية

دون اكتفاء .....

و راحة قلب لا نهاية لها.....

واطل في اعمارهم...

و لا تريني فيهم مكروهاً ...

الباحث

شكر وتقدير ...

اليك تتسابق الكلمات وتتزاحم العبارات ..... لتنظم عقد الشكر الذي لا يستحقه الا انت

يا من بذلت و لم تنتظر العطاء ...

الى كل من ساندني

اليك اهدي...

عبارات الشكر و التقدير

الباحث

قائمة المحتويات

رقم الفقرة	العنـــوان	رقم الصفحة
	المستخلص	
الفصل الأول :		
1.1	مقدمة عامة General Introduction	3-1
2.1	مشكلة الرسالة Message problem	3
3.1	هدف الرسالة The purpose of the message	4
4.1	اهمية الرسالة The importanceof the message	4
5.1	فرضية الرسالة Message Hypothesis	4
6.1	الاستعراض المرجعي Literature Review	8-5
الفصل الثاني : الجانب النظري		
1.2	مقدمة الفصل الثاني Introduction toChapter II	9
2.2	السلاسل الزمنية Time Series	10-9
1.3.2	الاتجاه العام Trend	10

2.3.2	التغيرات الموسمية Seasonal Variation	10
3.3.2	التغيرات الدورية Cyclical Variation	10
4.3.2	التغيرات العرضية Irregular Variations	11
4.2	نماذج السلاسل الزمنية Time Series Models	11
1.4.2	الأنموذج التجميعي Additive Model	11
2.4.2	نموذج حاصل الضرب Multiplicative Model	12-11
5.2	التنبؤ Forecasting	12
1.5.2	التنبؤ بنقطة Point Forecasting	12
2.5.2	التنبؤ بفترة Forecasting with Interval	12
6.2	اهمية التنبؤ The Importance of Forecasting	13
7.2	خطوات بناء إنموذج تنبؤ Steps of building a forecasting Model	14-13
8.2	انواع التنبؤ الاحصائي Types of Statistical Forecasting	14
1.8.2	التنبؤات طويلة المدى Long-Term Forecasting	15
2.8.2	التنبؤات القصيرة المدى Short-Term	15

	Forecasting	
15	مراحل عملية التنبؤ The steps of theForecasting process	9.2
16	الافتراضات التي يقوم فيها التنبؤ Assumptions of Forcasting	10.2
16	Forecasting Models نماذج التنبؤ	11.2
16	Qualitative النماذج النوعية (الوصفية) (Descriptive) Models	1.12.2
17	Quantitative Models النماذج الكمية	2.12.2
17	Linear Trend Model إنموذج الاتجاه الخطي	1.2.12.2
18	Quadratic Trend إنموذج الاتجاه التربيعي Model	2.2.12.2
18	Power (concave) الأنموذج غير الخطي	3.2.12.2
22-19	Smoothing إنموذج التمهيد الاسي exponential model	4.2.12.2
23-22	Box – Jenkins Models نماذج بوكس جنكينز	13.2
24-23	Autoregressive إنموذج الانحدار الذاتي Model (AR)	1.13.2
24	Moving إنموذج المتوسطات المتحركة	2.13.2

	Average Model (MA)	
26−25	الأنموذج المختلط Auto – Moving Average Model (ARMA) Regressive	3.13.2
27−26	الأنموذج المتكامل المختلط Integrated Mixed Model (ARIMA)	4.13.2
27	منهجية بوكس – جنكينز Box – Jenkins Approach	14.2
27	جمع البيانات Data Collection	1.14.2
29−27	فحص استقراريه السلسلة Check Stationary Of Time Series	2.14.2
30−29	عدم الاستقرارية في السلاسل الزمنية Non-Stationary Time Series	3.14.2
31−30	اختبار جذر الوحدة Unit Root Test	4.14.2
32−31	دالة الارتباط الذاتي Auto Correlation Function (ACF)	5.14.2
34−33	دالة الارتباط الذاتي الجزئي PACF (Partial Auto Correlation Function)	6.14.2
37−35	تحديد رتبة الأنموذج Order Model Selection	15.2
38−37	تقدير معلمات إنموذج الاتجاه العام Estimation Parameters of general trend model	16.2

17.2	فحص مدى ملائمة الأنموذج DiagnosticChecking	38
1.17.2	اختبار ليونغ – بوكس Ljung and Box Test	39-38
2.17.2	اختبار حدي الثقة Confidence Interval Test	39
18.2	التنبؤ Forecasting	40-39
19.2	معايير المفاضلة بين النماذج Criteria for trade-offs between models	41-40
الفصل الثالث : الجانب التطبيقي		
1.3	تمهيد preface	43-42
2.3	تعريف السرطان Definition Cancer	45-43
1.2.3	بعض الاحصاءات العالمية عن مرض السرطان Global statistics	47-45
3.3	بيانات السلسلة Observationsof time Series	47
4.3	التحليل الاحصائي Statistical analyze	48
1.4.3	إنموذج الانحدار الخطي البسيط (الاتجاه العام)	50-48
2.4.3	الأنموذج التربيعي	54-50
3.4.3	الأنموذج غير الخطي power(concave)	55
4.4.3	إنموذج التمهيد الاسي	59-56

5.4.3	استقراريه السلسلة الزمنية **Stationary in time series**	63-59
5.3	تشخيص رتبة الأنموذج والتقدير	67-64
6.3	فحص مدى ملائمة الأنموذج	70-67
الفصل الرابع : الاستنتاجات و التوصيات		
1.4	الاستنتاجات **Conclusions**	72-71
2.4	التوصيات **Recommendations**	72
المصـــادر		75-73
المستخلص باللغة الانكليزية (Abstract)		

قائمة الجداول

الجدول	عنــوان الجـــدول	رقم الصفحة
1-2	يبين التشخيص لنماذج السلاسل الزمنية وفقاً لسلوك دالتي ACF و PACF	35
1-3	أعداد المصابين بالأمراض الخبيثة الشهري في محافظة بابل للسنوات 2011 – 2016	47
2-3	مقاييس الدقة	48
3-3	تحليل التباين للأنموذج الخطي البسيط	49
4-3	مقاييس الدقة	50
5-3	نتائج تحليل التباين للأإنموذج التربيعي	51
6-3	المركبة الخطية والتربيعية للبيانات	52
7-3	نتائج تحليل التباين للأإنموذج التكعيبي	53
8-3	نتائج تحليل مركبات الأنموذج الخطية والتربيعية والتكعيبية	54
9-3	مقدرات معالم الأنموذج غير الخطي power concave	55

3-10	مجموع مربعات الاخطاء النهائية لمرحلة التقدير مع درجات الحرية	55
3-11	مقاييس الدقة	56
3-12	مقاييس الدقة	58
3-13	الاحتمالية لنتائج اختبار دكي- فولير لجذر الوحدة ولبيانات السلسلة الزمنية الاصلية وكذلك نتائج الاختبار بعد الفرق الاول	62
3-14	القيم الاحتمالية لنتائج اختبار دكي – فولير لجذر الوحدة ولبيانات السلسلة الزمنية الاصلية وكذلك نتائج الاختبار بعد الفرق الاول	62
3-15	بعض نماذج بوكس جينكينز المقترحة لحد الدرجة الرابعة ولبيانات السلسلة الزمنية المستقرة (بعد الفرق الاول)	65
3-16	المؤشرات الاحصائية	66
3-17	اختبار كاي سكوير لملائمة الأنموذج	67

قائمة الاشكال

الشكل	عنــوان الشكــل	رقم الصفحة
3-1	معادلة الاتجاه العام الخطية لأعداد المصابين .	50
3-2	معادلة الاتجاه العام التربيعية لأعداد المصابين.	52
3-3	التوزيع الاحتمالي الطبيعي للأخطاء العشوائية.	53
3-4	معادلة الاتجاه العام الخطية التكعيبية .	54
3-5	الأنموذج غير الخطي (power concave).	55
3-6	القيم التنبؤية والقيم الحقيقية لأعداد المصابين بالأورام الخبيثة.	57
3-7	عملية تقدير الأنموذج التمهيدي الاسي المزدوج.	59
3-8	رسم السلسلة الزمنية الاصلية.	60
3-9	قيم ورسم معاملات الارتباط (AC) والارتباط الذاتي الجزئي(PAC) للسلسلة الزمنية لأعداد المصابين بالأورام الخبيثة في محافظة بابل.	61
3-12	قيم ورسم معاملات الارتباط الذاتي (AC) والارتباط الذاتي الجزئي (PAC) للسلسلة الزمنية لأعداد المصابين بالأورام الخبيثة في محافظة بابل بعد اخذ الفرق الاول.	63

3-13	إنموذج معلومات مقياس أكيكي لأفضل 20 إنموذج.	66
3-14	اختبار البواقي عن طريق حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي.	68
3-15	اختبار سلسلة البواقي عن طريق حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي.	69
3-12	التوزيع الاحتمالي الطبيعي للأخطاء العشوائية.	70
3-13	يبين العلاقة ما بين القيم الاصلية لبيانات السلسلة الزمنية والمقدرة من الأنموذج.	70

المستخلص:

ان الزيادة في اعداد المصابين بالاورام الخبيثة حول العالم وما تسببه هذه الظاهرة من خسائر بشرية ومادية وما تولده من أثار اجتماعية ونفسية تشمل فئات المجتمع كافة الدافع الرئيس والهدف الاساس الذي وجهنا لغرض القيام يهذا البحث لهذه الظاهرة واجراء التنبؤات المستقبلية عنها.

تضمنت الرسالة جانبين نظري وتطبيقي، تضمن الجانب النظري كل من بعض طرائق السلاسل الزمنية في التنبؤ منها انموذج ألاتجاه الخطي العام وانموذج الاتجاه التربيعي والانموذج غير الخطي power concave وانموذج التمهيد ألاسي فضلا عن نماذج بوكس جنكينز Box-Jenkins .

في الجانب التطبيقي تم تطبيق الطرائق التي ذكرت في الجانب النظري على عينة من الاشخاص المصابين بالاورام الخبيثة بحجم (72) للفترة من 2011 ولغاية 2016 وحسب الاشهر. وبالاعتماد على نتائج التحليل الاحصائي تم التوصل إلى أن سلسلة أعداد المصابين بالأمراض الخبيثة في محافظة بابل للسنوات (2011 – 2016) تمثل سلسلة زمنية غير موسمية يوجد بها اتجاه عام ، أي أنها تمثل سلسلة زمنية غير مستقرة ، وذلك وفقاً لقيم معاملات الارتباط الذاتي ، ومن ثم فان تطبيق الفرق الاول هو الأنسب من غيره لإيجاد تنبؤات قريبة من القيم الحقيقية . وان المتغير مربع الزمن في الأنموذج التربيعي كان تأثيره عكسياً على أعداد المصابين بالاورام الخبيثة لان إشارته سالبة إذ بلغت المعلمة المقدرة 0.2484 الا ان هذا التأثير طفيف لانخفاض قيمة المعلمة المقدرة وان الأنموذج التربيعي بشكل عام افضل من الإنموذج الخطي البسيط ، كذلك ان الأإنموذج ARIMA(1،1،1) هو المناسب الذي يمكن الاعتماد عليه في اجراء التنبؤات بأعداد المصابين بالأمراض الخبيثة في محافظة بابل لغرض وضع الخطط المستقبلية ، وقد اعطى تنبؤات جيدة وقريبة من القيم الفعلية لسلسلة أعداد المصابين ، وهذا يتضح عن طريق القيم الحقيقية للعام 2016 التي استعملت كسنة ضبط ووقوعها في مدة التنبؤ.

1.1مقدمة عامة General Introduction

شهد العالم خلال مسيرة القرن العشرين وبداية القرن الحالي تطوراً كبيراً في مجال البحث العلمي في الميادين كافة ، ولعل من أهم العوامل التي أدت الى هذا النشاط توجه اهتمام معظم دول العالم الى التخطيط والتنمية في المجالات كافة وذلك بتسخير الطاقات العلمية المبدعة في إعداد البرامج المتكاملة في الميادين كافة لتحقيق أعلى مستوى من التقدم في شتى مجالات الحياة .

إن الهدف الأساس للعلم هو السعي الدائم لمعرفة العلاقات المختلفة بين الظواهر وتفسيرها. وذلك بتحديد العلاقات بين المتغيرات والإفادة من هذه العلاقات المحددة في مجالات البحث العلمي. ويحتل علم الإحصاء الدور الرئيس في وضع الخطط والبرامج . فمن طريق التحليل الإحصائي نعرض سلبيات الماضي ونتنبأ بالمستقبل وحاجاته على وفق الإمكانيات المتيسرة ونظرا" لاختلاف طبيعة الظواهر والاتساع الكبير للمتغيرات والتداخل بين تأثيراتها .

من جانب آخر يعد التنبؤ المستقبلي من الموضوعات الرئيسة التي تمكن أصحاب القرار من إتخاذ القرارات السليمة في المجالات الصحية والاجتماعية والخدمية وغيرها. وتقسم أساليب التنبؤ الى نوعين رئيسين هما: ألاساليب النظامية ، والاساليب غير النظامية ، فالأساليب النظامية تعتمد على طرائق علمية ولا تدخل العوامل الشخصية في التقدير، ومن ثمّ فهي بعيدة عن التحيز ، أما ألاساليب غير النظامية فتعتمد على الخبرة والتجربة والاجتهاد الشخصي ، وتعد السلاسل الزمنية من أهم الاساليب النظامية التي يمكن عن طريقها التنبؤ بما سيحدث بالمستقبل عن قيم الظاهرة في ضوء ما حدث لها في الماضي ، واذ ان الجانب الصحي له اهمية كبيرة في الجانب التنموي ، لأنه يهتم بحياة الانسان والخلاص من الامراض كافة ولاسيما مرض السرطان الذي يعد من الامراض الخطيرة والمميتة إذا ما أهملت دون تخطيط مسبق لمواجهة هذا المرض . لذا فقد اهتمت هذه الرسالة بدراسة هذا المرض اعتمادا على ألاعداد الشهرية للأشخاص المصابين بهذا المرض من الذكور والاناث في محافظة بابل للفترة من (كانون الثاني 2011- كانون الاول

2016) م لغرض وضع انموذج للتنبؤ المستقبلي بأعداد المصابين بهذا المرض ولكلا الجنسين للسنوات القادمة ووضعها بأيدي الجهات المختصة لغرض اتخاذ التدابير اللازمة للوقاية من هذا المرض . ولأهمية موضوع التنبؤ وتأثره بعوامل عدة فقد ظهرت الحاجة الماسة لتطوير الاساليب و الوسائل الاحصائية لزيادة الدقة و المعرفة الشاملة و الواسعة بالتنبؤ بأعداد المصابين لمدة زمنية قادمة ، ومن هذه الاساليب نماذج الانحدار ولكن ليس بشكلها التقليدي ، وانما بشكل متطور بحيث تلائم حالة المتغير التابع .

ولدراسة ظاهرة معينة والبحث عن الاساليب المناسبة للتحليل والمعالجة عن طريق الوسائل الاحصائية المناسبة استعملت السلاسل الزمنية لتحليل التغيرات الحاصلة في دراسة سلسلة زمنية لأعداد المصابين بالاورام الخبيثة ، وقد تزايد الاهتمام بموضوع التنبؤ خلال السنوات الأخيرة وظهرت أساليب حديثة خاصة ، ومنها الشبكات العصبية الاصطناعية ANN (Artificial Neural Networks) .

ويعد مرض السرطان من الامراض الشائعة و واسعة الانتشار في العالم ، وفي دراستنا هذه سيتم التركيز على اعداد المصابين بهذا المرض لخطورته على الانسان و الذي سجلت معدلات الوفيات فيه نسب عالية مقارنة بالامراض الاخرى . و عليه جاءت هذه الرسالة للتعرف على بعض طرائق التنبؤ بأعداد المصابين بالاورام الخبيثة بتطبيق بعض نماذج تحليل السلاسل الزمنية .

اما هيكلة الرسالة فتضمنت اربعة فصول ، تضمن الفصل الاول المقدمة و مشكلة الرسالة و هدف الرسالة و اهميتها و فرضية الرسالة والاستعراض المرجعي ، الفصل الثاني شمل الجانب النظري ، وتضمن بعض المفاهيم الاساسية في التنبؤ. الفصل الثالث فقد خصص للجانب التطبيقي و يتضمن جوانب مهمة عن مرض السرطان من حيث المقدمة و التعريف و عدد من الاحصاءات العالمية و اعراض هذا المرض وكذلك اسباب اختيار هذا المرض في الدراسة وفيه تم وصف

البيانات الخاصة لمتغيرات الدراسة و بعد ذلك تم تحليل بيانات السلسلة الزمنية لأعداد المصابين بالاورام الخبيثة. ويتناول الفصل الرابع اهم الاستنتاجات و التوصيات التي توصلنا اليها في الرسالة.

2.1 مشكلة الرسالة Message problem

السرطان مشكلة صحية عالمية يشمل جميع انحاء العالم ويقدرعدد المصابين سنويا مايقارب 10 ملايين حالة اصابة جديدة وحوالي 7 ملايين حالة وفاة بالسرطان ، أي ان نسبة الوفاة مرتفعة جدا وهي بالتأكيد تشكل عبئا اقتصاديا واجتماعيا كبيرا. اما في العراق هناك ارتفاع ملحوظ بأعداد المصابين بالاورام الخبيثة (السرطان) في السنوات الاخيرة ابتداءا من العام 2003 ، ولاسيما بعد استخدام قوات التحالف مختلف صنوف الأسلحة التي لا تخلو من مسببات لهذا المرض، كذلك الظروف المعيشية والاجتماعية والنفسية التي يعاني منها الافراد . مما يستدعي الامر البحث والتقصي عن هذه الظاهرة والتنبؤ بالمستقبل بعدد الاصابات . علما بان المعدل السنوي للإصابة بهذه الامراض في العراق يبلغ 2500 حالة اصابة بينها 20% اصابات بسرطان الثدي ،وقد اعلنت وزارة الصحة العراقية بانها ستتولى نشر احصاءات دقيقة خلال العام 2018 وعن طريق مجلس السرطان بوزارة الصحة .

3.1 هدف الرسالة The purpose of the message

تهدف الرسالة الى التنبؤ بأعداد المصابين بالاورام الخبيثة في محافظة بابل باستعمال بعض طرائق السلاسل الزمنية في التنبؤ اعتمادا على عينة البحث المأخوذة وايجاد افضل انموذج يمكن استخدامه لأغراض التنبؤ عن طريق المقارنة بين النماذج المدروسة باستعمال بعض المعايير الاحصائية.

1 . 4 أهميـة الرسالة The importance of the message

نظرا لتطور اعداد المصابين بالاورام الخبيثة بشكل ملحوظ وما تسببه هذه الظاهرة من خسائر بشرية ومادية وما تولده من اثار اجتماعية ونفسية تشمل فئات المجتمع كافة. جاءت فكرة كتابة البحث عن التحليل الاحصائي والتنبؤ بأعداد المصابين بالاورام الخبيثة .نقدم ما نتوصل اليه من استنتاجات بعد اتمام الدراسة امام انظار واضعي السياسات وصانعي القرار من مسؤولين في مجال الصحة والاقتصاد لغرض اتخاذ الاجراءات ووضع الخطط والتدابير اللازمة وتقديم ما يرونه مناسباً لمواجهة هذه المشكلة مستقبلاً والسيطرة عليها في الوقت الحاضر.

1 . 5 فرضية الرسالة Message Hypothesis

لغرض اختبار ما هدفت الرسالة له نضع الفرضيات الاحصائية الاتية:

H0 : لا توجد زيادة في عدد المصابين بالاورام الخبيثة في محافظة بابل خلال المدة (2017 – 2018) م مع مثيلاتها للسنوات السابقة .

H1 : توجد زيادة في عدد المصابين بالاورام الخبيثة في محافظة بابل خلال المدة (2017– 2018) مع مثيلاتها للسنوات السابقة.

6.1 الاستعراض المرجعي Literature Review

اهتمت البحوث في الآونة الاخيرة استخدام السلاسل الزمنية والحصول على نماذج يتم عن طريقها التنبؤ بالظواهر لذا ارتأينا اعطاء نبذة بسيطة عن البحوث والدراسات.

- في عام 1970 قام الباحثان بوكس- جينكنز (Box; G.E.P & Jenkins, G.M)[27] بدراسة النماذج المختلطة (Mixed Autoregressive – Moving Average) التي يرمز لها (ARMA) بهدف معالجة مختلف أنواع السلاسل الزمنية (المستقرة وغير المستقرة، الموسمية وغير الموسمية) والتي سميت بالنماذج المختلطة غير المستقرة (Integrated Autoregressive Moving Average Models) اذ كانت طريقة متكاملة لاستخدام هذه النماذج في تحليل السلاسل الزمنية والتنبؤ بها وسميت هذه الطريقة بطريقة بوكس – جينكنز التي تحتوي اربعة مراحل رئيسة هي (تحديد الأنموذج التجريبي ، وتقدير معالم الأانموذج التجريبي ، والتحقق من ملائمة الانموذج التجريبي واخيراً استعمال الانموذج المقدر والملائم في التنبؤ).
- وفي عام 1974 قام الباحث (Akaike, E)[23] باقتراح معياراً سّمي بمعيار خطأ التنبؤ النهائي (Final prediction error criterion) الذي يرمز له (FPE)، وهو تقدير لمتوسط مربعات خطأ التنبؤ لمدة قادمة واحدة، واستعمال هذا المعيار في ايجاد اسلوب جديد لتحديد درجة انموذج الانحدار الذاتي (AR) سمي بخطأ التنبؤ النهائي الأصغر (minimum FPE).
- وفي عام 1985 قام (Liung) [29] بمعالجة مشكلة التشخيص للنماذج المختلطة عن طريق استعمالها تحليل الارتباط القويم (canonical correlation analysis) بين المتغيرات الموسمية Ym.t , Um.t−j اذ Ym.t = (Zt, …, Zt−m) وهذا الاسلوب يتعامل مع النماذج المستقرة وغير المستقرة، فضلاعن اعطاءه المقدرات المتسقة لمعلمات الانحدار الذاتي.
- وفي عام 1994 أشارت (Anna Clara Monti) [24] الى ملاحظات مقترحة لاختبار جودة المطابقة لنماذج السلسلة الزمنية معتمدة اساساً على مجموع المربعات للارتباطات الذاتية الجزئية للأخطاء سميت إحصاءه الاختبار Q تكون تقريباً تتبع توزيع χ^2 بدرجة حرية (n−p)

خلال تجربة مونت كارلو تمت دراسة العينات الصغيرة.إذ ان :n : عدد المشاهداتp : عدد المعلمات.

- وفي عام 1997 درس الباحثان (Lai & Lee) [28] معيار المعلومات والتنبؤ لاختبار الانموذج الملائم في نماذج الانحدار ونماذج ARMA التصادفية واقترحا معيارَ خطأ التنبؤ المتراكم المقترح من قبل (Rissanen) وكذلك معيار المعلومات (Fisher Information criterion) (FIC) الذي اقترحه (Wei & Fisher) واثبتا الاتساق الشديد لمعيار اختبار الانموذج تحت شروط معينة.
- في عام (2005) قام الباحث العاني[5] استخدام نماذج (ARIMA) في التنبؤ الاقتصادي وتم أجراء تطبيق عملي على سلسلة زمنية لمنتج طبي (70 مشاهدة) باستعمال البرنامج الإحصائي STATGRAPHICS وتبين من النتائج أن الأنموذجARIMA(1,0,1) حقق قدرة تنبؤية أعلى من الانموذج ARIMA(1,0,0) حسب أختبارات القدرة التنبؤية.

- وفي العام نفسه قامت الباحثة محمود[21] بالمقارنة بين طرائق تحليل السلاسل الزمنية وتطبيقها على مبيعات الشركة العامة لتوزيع الكهرباء في بغداد واستعملت عدة طرائق منها بوكس جنكينز والتمهيد الاسي والتنقية المكيفة وغيرها، اذ تفوقت طريقة التنقية المكيفة (AF) على طريقة التمهيد الأسي (ES) وطريقة Box–Jenkins.

- في عام (2009) قدمت الباحثة رديف[11] بنشر بحثاً أستعملت فيه نماذج ARIMA للتنبؤ عن أنتاج الاسمنت في العراق مستعملة سلسلة زمنية للانتاج العام للاسمنت في العراق للفترة من 1952–2002 ومتنبأة بالإنتاج للفترة من 2003 –2013 وتوصلت الى أن أفضل أانموذج للتنبؤ هو ARIMA(2,1,0) .

- وفي عام (2010) قام(الطائي)[4] بنشر بحثه التنبؤ والتمهيد للسلاسل الزمنية باستعمال التحويلات مع التطبيق وتبين أن انموذج ARIMA(5,0,2) أفضل انموذج للتنبؤ بمعدلات الامطار لامتلاكه اقل قيمة للمعيارين (MSE,AIC) وبعد أجراء التحويل بطريقة الجذر التربيعي كان الأنموذج الافضل هو ARIMA(1,0,0) لأمتلاكه أقل قيمة للمعيارين

(MSE,AIC) ، وعند أجراء التمهيد الأسي كان الأانموذج الأفضل هو ARIMA(1,0,0) لأمتلاكه أقل قيمة للمعيارين (MSE , AIC).

- وفي العام نفسه قامت الباحثة (الجبوري)[1] بنشر بحثها عن التنبؤ بأسعار النفط العراقي لعام 2010 باستعمال السلاسل الزمنية وتبين أن الأنموذج ARIMA(0,2,1) متفوقة على انموذج هولت Holt عن طريق المعايير (MAPE , MAE , RMSE).

- و في العام (2012) استعملت (طعمة) [14] اسلوب تحليل السلاسل الزمنية في بحثها (استخدام تحليل السلاسل الزمنية للتنبؤ بأعداد المصابين بالأورام الخبيثة في محافظة الانبار) وقد تم استخدام طريقة (بوكس جينكنز) في التحليل . وايجاد افضل انموذج للتنبؤ بأعداد المصابين بالأورام الخبيثة في محافظة الانبار وذلك بالاعتماد على البيانات الشهرية للفترة (2010–2016) وقد اظهرت نتائج التحليل ان الانموذج الملائم لها هوه انموذج الانحدار الذاتي المتكامل من الدرجة الثانية ARIMA (2,1,0) مع الاعتماد على هذا الانموذج للتنبؤ بأعداد المصابين بالأورام الخبيثة شهريا ولسنتين قادمتين وقد كانت القيم التنبؤية متناسقة مع قيم السلسلة الاصلية وهذا يدل على ان الانموذج كفؤ .

- في العام (2013) قام الباحث زين العابدين[12] بنشر بحث مقارنة انموذج التدخل مع انموذج ARIMA في السلاسل الزمنية وتبين أن أفضل أانموذج للتنبؤ بأسعار البيوت هو ARIMA(0,0,1) بالاعتماد على أقل قيمة لمعياري (MSE , AIC) .
- وفي العام نفسه قام (الباحثان الصفاوي وغانم)[3] بنشر بحث مقارنة بين نماذج بوكس وجنكينز وبين التمهيد الاسي بالاعتماد على معايير (MSE,MAE,MAPE) للتنبؤ بالقيم المستقبلية لنماذج السلاسل الزمنية وكان الانموذج المقترح والملائم لبيانات السلسلة هو ARIMA(0,1,2)(0,1,1) وتم تطبيقه على بيانات سلسلة تمثل حمل الذروة اليومي للطاقة الكهربائية في الموصل.

- في العام 2015 قام (الباحث فخري)[18] بدراسة التنبؤ بقيم السلاسل الزمنية باستعمال انموذج ARMAX مع تطبيق عملي على درجات الحرارة العظمى في مدينة بغداد واعطى الانموذج ARMAX(4,3,4,1) افضل نتائج للتنبؤ بدرجة الحرارة العظمى لمدينة بغداد وكانت النتائج مقاربة للنتائج الاصلية.
- في عام 2017 طبق (بن مرضي) [9] اسلوب بوكس جنكنز في تحليل السلاسل الزمنية والتنبؤ باعداد الحوادث المرورية واعداد المصابين في المملكة العربية السعودية ، ومعرفة القيمة التنبؤية والسلاسل الزمنية لاعداد الحوادث المرورية واعداد المصابين خلال الستين شهراً القادمة وتوصل الى ان نماذج بوكس جنكنز هي افضل طرائق التنبؤ بالحوادث المرورية وان هنالك انخفاضاً ملحوظاً في قيم الحوادث والمصابين في المستقبل.

1.2 مقدمة الفصل الثاني Introduction to Chapter II

يعد موضوع تحليل السلاسل الزمنية من الموضوعات الاحصائية المهمة التي تدرس سلوك الظواهر، وتفسيرها عبر مدة زمنية معينة. ويمكن تلخيص أهداف تحليل السلاسل الزمنية بالحصول على وصف دقيق للسمات التي تتولد منها السلسلة الزمنية، وبناء إنموذج لتفسير سلوك السلسلة الزمنية واستخدام النتائج للتنبؤ عن القيم المستقبلية للسلسلة، فضلا عن التحكم بالظاهرة اذا امكن ذلك بفحص ما يمكن حدوثه عند تغيير بعض معلمات الانموذج ، ولتحقيق ذلك يتطلب الامر دراسة تحليلية وافية لنماذج السلاسل الزمنية بالاعتماد على الاساليب الاحصائية والرياضية .

2.2 السلاسل الزمنية[7] [16] Time Series

يمكن تعريف السلسلة الزمنية بأنها قيم ظاهرة مرتبة وفقاً للزمن او مجموعة القيم التي تأخذها الظاهرة في مدد زمنية متتالية ومتساوية قد تكون سنوية او فصلية (ربع سنوية) او شهرية او يومية وهكذا، وذلك لمعرفة طبيعة التغيرات التي تحدث لقيم الظاهرة عبر الزمن ليصبح بالامكان القيام بالتقديرات والتنبؤات المستقبلية. وتحليل السلسلة الزمنية يعني تحليل مكوناتها الرئيسة وهيَ الاتجاه العام للتغير في قيمة الظاهرة على المدى الطويل، والاتجاه الموسمي وهي التغيرات التي تحصل على قيم الظاهرة في مدد قصيرة (اقل من سنة) كالتغيرات الفصلية والشهرية واليومية، والتغيرات الدورية التي تحصل على قيمة الظاهرة بشكل دوري في مدد تزيد عن السنة والتي تعود للتطورات الاقتصادية والسياسية، والتغيرات العرضية التي تطرأ على قيمة الظاهرة بشكل غير منتظم نتيجة حدوث الكوارث الطبيعية او الحروب او الاضرابات العمالية.

وتصنف السلاسل الزمنية وفقاً لعدد المتغيرات في السلسلة الزمنية الى صنفين :

1) سلسلة زمنية بمتغير واحد Univariate Time Series

تحتوي السلسلة الزمنية على متغير واحد فقط عندما تستعمل البيانات الحالية والسابقة عن متغير واحد فقط ، فمثلاً للتنبؤ بأعداد المرضى المصابين بالسكر في الشهر القادم أو بعد شهرين من الآن تستعمل فقط البيانات الحالية والسابقة عن اعداد المرضى .

2) سلسلة زمنية متعددة المتغيرات Multiple variables Time Series

تكون السلسلة الزمنية متعددة المتغيرات عندما تستعمل أكثر من متغير واحد في وصف سلوك السلسلة الزمنية وتحليلها فمثلاً للتنبؤ بدرجات الحرارة العظمى اليومية وهنا يتطلب استخدام متغير أو أكثر مثلاً سرعة الرياح اليومية .

3.2 مكونات السلاسل الزمنية Time Series Components

1.3.2 الاتجاه العام [16] Trend

هو التغير في قيمة الظاهرة على المدى الطويل ويرمز لهُ بالرمز (T) والاتجاه العام للسلسلة الزمنية يمكن تمثيله بخط مستقيم إذا كان التغير في قيمة الظاهرة يسير بنسبة ثابتة مع الزمن او بخط غير مستقيم إذا كان التغير في قيمة الظاهرة متغيراً وليس ثابتاً، اذ لا يمكن تمثيله بخط مستقيم وإنما بمنحني، ويكون الاتجاه العام موجباً اذا اتجهت قيم الظاهرة نحو التزايد مدة بعد اخرى ويكون سالباً اذا اتجهت قيم الظاهرة نحو التناقص مدة بعد اخرى، و يكون خطياً او غير خطياً كما في المنحنى الأُسي .

2.3.2 التغيرات الموسمية Seasonal Variation

وهيَ التغيرات التي تحصل على قيمة الظاهرة في مدد زمنية اقل من سنة كالتغيرات الفصلية والشهرية واليومية وتظهر في الموسم نفسه من السنة اللاحقة فالتغيرات المناخية تعتبر من اهم العوامل التي تسبب التغيرات الموسمية فأختلاف المناخ في فصول السنة والعادات الاجتماعية والدينية تُعد اهم الاسباب الرئيسة في التغيرات الموسمية، ويرمز للتغيرات الموسمية بالرمز (S) .

3.3.2 التغيرات الدورية Cyclical Variation

وهيَ التغيرات التي تحصل على قيمة الظاهرة بصورة دورية وتعيد نفسها خلال مدد زمنية تزيد عن السنة الواحدة وتسمى هذه التغيرات بالتذبذبات الدورية وتكون هذه التذبذبات اقل انتظاماً من التغيرات الموسمية اذ ان الذبذبة الواحدة لا تتكرر بالطول نفسه او القوة كمددالركود والازمات الاقتصادية التي تحدث بصورة دورية ويرمز لهذه التغيرات بالرمز (C) . والدورة هيَ المسافة بين التقعرين او التحدبين في منحنى السلسلة الزمنية .

4.3.2 التغيرات العرضية Irregular Variations

هيَ التغيرات التي تحدث بصورة عرضية او عشوائية والتي تنتج عن حوادث غير متوقعة كالأوبئة والزلازل وتحدث حركات واتجاهات لا يمكن تمييزها لأنها لا تحدث بانتظام ويرمز لهذه التغيرات بالرمز (I) .

4.2 نماذج السلسلة الزمنية [15] Time Series Models

هناك إنموذجان للسلسلة الزمنية يوضحان العلاقة بين المكونات الاربعة وتحديد هذه المكونات يعتمد على نوع الانموذج المستعمل في التعبير عن قيمة الظاهرة :

1.4.2 الانموذج التجميعي[16] Additive Model

يعبر الانموذج التجميعي عن قيمة الظاهرة Y_t كحاصل جمع لمكوناتها الرئيسة الاربعة والتي هي : الاتجاه العام والتغيرات الموسمية والتغيرات الدورية والتغيرات العرضية . وهذا الانموذج يفترض ان العوامل الاربعة مستقلة عن بعضها وهذه حالة نادرة في الحياة العملية، اذ ان التغير في احد العوامل لهُ تأثير على العوامل الاخرى في السلسلة الزمنية مما يتناقض مع شرط الاستقلالية .

$$Y(t) = T_t + S_t + C_t + I_t \quad \ldots\ldots\ldots\ldots\ldots\ldots(2\text{-}1)$$

اذ ان :

T_t : الاتجاه العام

S_t : التغيرات الموسمية

C_t : التغيرات الدورية

I_t : التغيرات العرضية

2.4.2 نموذج حاصل الضرب Multiplicative Model

يعبر الانموذج الضربي عن قيمة الظاهر Y_t كحاصل ضرب لمكوناتها الاربعة فهوَ يفترض ان العوامل الاربعة تتفاعل مع بعضها ولا تتحرك مستقلة وهذا اقرب للواقع العملي وهذا الانموذج هوَ الاكثر استعمالاً في السلاسل الزمنية .

$Y_t = T_t \times S_t \times C_t \times I_t$ ……(2-2)

5.2 التنبؤ [7] Forecasting

يعرف التنبؤ بأنه تقدير قيمة الظاهرة بالمستقبل بالاعتماد على بيانات الماضي والحاضر فهو مهم في التخطيط ووضع الافتراضات عن أحداث المستقبل باستعمال الاساليب الاحصائية ذات العلاقة . و يشمل تقدير حجم الظاهرة في المستقبل مع الاخذ بعين الاعتبار اهم العوامل المؤثرة فيها . أن للتنبؤ دوراً مهماً وبارزاً في عملية اتخاذ القرارات ، ذلك أن التنبؤ الذي ما هو الا رحلة سفر عبر الزمن الى المستقبل أي رؤية مستقبلية لما ستكون عليه الظواهر والمتغيرات في المستقبل ، بصيغة اخرى هو اسقاط للماضي على المستقبل عن طريق الحاضر. ولابد ان نميز بين نوعين من التنبؤات هما:

1.5.2 التنبؤ بنقطة[21] : Point Forecasting

وهو التنبؤ بالقيمة المستقبلية للسلسلة الزمنية بقيمة واحدة (مفردة) ، وهذه القيمة عادة ما تكون ذات أقل متوسط مربعات خطأ تنبؤ ، اذ يرمز إلى القيمة التنبؤية إلى Z_{t+1} بـ $Z_t(l)$ اذ يسمى (L) بـ Lead و t بـ Origin .

2.5.2 التنبؤ بفترة[21] : Forecasting with Interval

وهو التنبؤ بالقيم المستقبلية إلى Z_{t+l} بفترة أو مدى من القيم التنبؤية المحسوبة اذ نكون واثقين عند مستوى احتمالية معين أن القيمة الحقيقية للتنبؤ تكون محتواة في هذه الفترة.

6.2 أهمية التنبؤ : The Importance of Forecasting

للتنبؤات اهمية بالغة في التخطيط وصياغة القرارات الاقتصادية والاجتماعية والصحية التي ترسم مسار المنظمات ولاسيما المنظمات او المؤسسات الصحية التي نحن بصدد البحث فيها .

ومن هنا تبرز أهمية ودور التنبؤ والمتمثلة في:

- ✓ يضمن وإلى حد كبير الكفاءة والفاعلية للمؤسسة في المرونة مع البيئة الخارجية .
- ✓ معرفة حاجات المؤسسة في المدى القصير والمتوسط.
- ✓ يسهم في الحد من المخاطر التي قد تواجه المؤسسة.
- ✓ تعطي صورة للمؤسسة عن توجهها المستقبلي.
- ✓ تسهم بقدر كبير في اتخاذ القرارات وترقب آثارها مستقبلا.

هذه الاهمية تقودنا الى طرح التساؤلات الاتية :

- كيف يمكن استخدام النماذج الاحصائية في حساب التنبؤات ؟
- ماهي افضل طرائق التنبؤ الاحصائي ؟
- ما مدى فاعلية نماذج التنبؤ الاحصائي في اتخاذ القرارات ؟

7.2 خطوات بناء انموذج تنبؤ [4] Steps of building a forecasting Model

ان ايجاد إنموذج مناسب تنطبق عليه السلسلة زمنية مشاهدة يعد من المهام الصعبة والتي تحتاج الى الكثير من البحث والخبرة لذلك سنستعرض بعض الخطوات المهمة لبناء إنموذج رياضي للتنبؤ عن متسلسلة زمنية ما :

- تعيين الإنموذج او تحديد الإنموذج : وهذا يتم برسم المتسلسلة الزمنية على شكل بياني اذ يكون الاحداثي الافقي هو الزمن والراسي قيم الظاهرة (المشاهدة) ومن ثم اختيار إنموذج

رياضي معتمدين على بعض المقاييس الاحصائية التي تميز إنموذج عن اخر وعلى الخبرة المستمدة من الدراسات والابحاث .

- تطبيق الإنموذج : بعد ترشيح إنموذج او اكثر كإنموذج مناسب لوصف المتسلسلة المشاهدة نقوم بتقدير معلمات هذا الإنموذج من بيانات المشاهدة باستعمال طرائق التقدير الاحصائي الخاصة بالمتسلسلات الزمنية وهذا الإنموذج المرشح يؤخذ كإنموذج اولي قابل للتعديل لاحقا.
- تشخيص واختبار الإنموذج : اجراء اختبارات تفحصية على اخطاء التطبيق Fitting Errors لمعرفة مدى تطابق المشاهدات مع القيم المحسوبة من الإنموذج المرشح ومدى صحة فرضيات الإنموذج .في حالة اجتياز الإنموذج المرشح لهذه الاختبارات تقوم باعتماده على انه الإنموذج النهائي ويستعمل لحساب تنبؤات للقيم المستقبلية والانعود للخطوة الاولى لتعيين إنموذج جديد .
- حساب التنبؤات : يستعمل الإنموذج النهائي لتوليد تنبؤات عن القيم المستقبلية ومن ثم حساب أخطاء التنبؤ Fitting Errors كلما استحدثت قيم جديدة (مشاهدة) من المتسلسلة الزمنية ومراقبة هذه الاخطاء فيما يسمى باخطاء المراقبة والتي توضع للقبول بنسبة خطأ معين و إذا تجاوزت أخطاء التنبؤ يعاد النظر في الإنموذج وتعاد الدورة من جديد بتحديد إنموذج مرشح اخر .
- استعمال التنبؤات ووضع القرارات : تقدم التنبؤات في تقرير لصانعي القرار للنظر في استعمالها بالشكل المناسب .

8.2 أنواع التنبؤ الاحصائي [7] Types of Statistical forecasting

يصنف التنبؤ الإحصائي وفق مدى أو معايير حساب التنبؤات الى نوعين:

1.8.2 التنبؤ طويل المدى : Long-Term Forecasting

يتميز هذا النوع من التنبؤات بطول وبعد المدة الزمنية للتوقعات ، و اما الطرائق الاحصائية التي تستجيب لهذا النوع من التنبؤات فهي نماذج الانحدار الخطي البسيط والمتعدد وكذلك يمكن استعمال معدلات النمو كطريقة لحساب التنبؤات البعيدة المدى الا على الرغم من انها غير فعالة في حساب التوقعات لان تباين اخطاء التوقع فيها كبيراً (القيم المتوقعة بعيدة عن القيم الحقيقية) وتقل فاعلية هذة النماذج كلما زادت مدة التوقعات ، ومع ذلك فان للتنبؤات البعيدة المدى اهمية بالغة في وضع القرارات البعيدة المدى وتخطيط السياسات الاستراتيجية .

2 .8.2 التنبؤ قصير المدى : Short-Term Forecasting

تكون مدة ومجال التنبؤات القصيرة المدى قصيرة جدا ، ولتطبيق الدراسات التنبؤية القصيرة الاجل نحتاج الى استعمال تقنيات السلاسل الزمنية المتمثلة في نماذج الانحدار الذاتي AR والمتوسطات المتحركة MA وعلى غرار التنبؤات البعيدة المدى تتميز طرائق التنبؤات القصيرة المدى بالفاعلية ولها اهمية كبيرة في اتخاذ القرارات الظرفية والسياسات الانية المستعملة في تسيير المؤسسات ، ومن اشهر الطرائق المستعملة في حساب التنبؤات القصيرة الاجل طريقة بوكس جنكينز ، نماذج الانحدار الذاتي VAR ، نماذج ARCH،.....الخ

9.2 مراحل عملية التنبؤ [20] The steps of the forecasting process

المرحلة الأولى: تحديد الهدف من التنبؤ

المرحلة الثانية: تجميع البيانات اللازمة للظاهرة محل التنبؤ.

المرحلة الثالثة: تحليل البيانات وانتقائها لاستعمالها.

المرحلة الرابعة: اختيار الإنموذج المناسب من أساليب التنبؤ بالظاهرة محل الدراسة.

المرحلة الخامسة: اتخاذ القرار المناسب.

10.2 افتراضات التنبؤ [20] Assumptions of forcasting

يستند التنبؤ الى مجموعة من الافتراضات ملخصها الاتي :

1. ان المستقبل لا يمكن التأكد منه تماما ويبقى عدم التأكد هذا قائما بغض النظر عن الطريقة التي استعملت فيه الى أن يمر الزمن ويمكن حينذاك رؤية الواقع الحقيقي .
2. ان هناك نقاطاً غير واضحة في التنبؤ فنحن على سبيل المثال لا نستطيع التنبؤ بمستجدات التكنلوجيا التي لا تتوفر لدينا معلومات تشير اليها الان .
3. ان التنبؤ يستعمل لوضع السياسات سواء أكانت خدمية أم اجتماعية أم صحية أم اقتصادية ، وان هذه السياسات نفسها اذا ما نفذت ستؤثر في المستقبل وتجري عليه تغيرات لم يتكلم عنها التنبؤ نفسه ما يحدث الافتراق بين ما جاء في التنبؤ وما سيتحقق على أرض الواقع .

11.2 نماذج التنبؤ [20] Forecasting Models

لا يمكن القول بأن هناك تقنية من بين الطرائق الخاصة بالتنبؤ تكون فعالة إلا إذا حققت مجموعة من الشروط منها الكلفة ، الدقة ، توفير البيانات اللازمة ، الوقت المحدد لجمع المعلومات ، وتوفر الإمكانيات اللازمة المادية والبشرية والمعنوية للقيام بعملية التنبؤ. ويمكن تقسيم نماذج التنبؤ الى قسمين كالاتي :

1.12.2 النماذج النوعية (الوصفية) [7] Qualitative (Descriptive) Models

وهي النماذج التي يعتمد بناءُها على الخبرة ورأي الأفراد داخل وخارج المؤسسة وحسب المستوى الهرمي للقرارات ومنها إنموذج الحدس والخبرة وإنموذج دلفي وأسلوب لجنة الخبراء.

2.12.2 النماذج الكمية [7] Quantitative Models

تعتمد الأساليب الكمية على استعمال النماذج الرياضية في تحليل المتغيرات الخاصة بالمؤسسة وحسب توفر البيانات اللازمة عن الظاهرة . واستخدام الطرائق الإحصائية ومنها إنموذج المتوسطات المتحركة البسيطة وإنموذج التمهيد الاسي البسيط وإنموذج تحليل الانحدار ... الخ والتي هي موضوع دراستنا في هذه الرسالة وهي على انواع منها :

1.2.12.2 انموذج الاتجاه الخطي Linear Trend Model

ويسمى بإنموذج الانحدار الخطي البسيط ، وهو من احدى الطرائق الواسعة الاستعمال لملائمة خط الاتجاه العام ، فإذا كانت الظاهرة تزيد او تنقص بمقدار ثابت لكل مدة زمنية فان خط الاتجاه العام يكون على صورة خط مستقيم ويأخذ الصيغة الاتية :

$$Y_t = ß_0 + ß_1 x_t + e_t \quad \text{.....(2-3)}$$

اذ ان :

Y_t : المتغير المعتمد الذي يمثل قيم الظاهرة .

t : المتغير المستقل الذي يمثل الزمن .

($ß_0$ و $ß_1$) يمثلان معلمات الانموذج : $ß_0$ هي نقطة تقاطع خط الاتجاه العام مع المحور الصادي ، $ß_1$ ميل خط الاتجاه العام الخطي .

e_t : الخطأ العشوائي يتوزع طبيعياً بمتوسط مقداره صفر وتباين σ_e^2 وتغاير ذاتي مقداره صفراً.

2.2.12.2 انموذج الاتجاه التربيعي Quadratic Trend Model

علـى الاغلـب يكـون الخـط المسـتقيم مناسـباً لطبيعـة البيانـات فـي تمثيـل الاتجـاه العـام، وفـي بعـض الاحيـان يكـون غيـر مناسـب وفـي مثـل هـذه الحالـة يكـون مـن الافضـل ان نقـوم بتوفيـق منحنـى مـن الدرجـة الثانيـة والـذي يسـمى انمـوذج الاتجـاه العـام التربيعـي وتكـون علـى صورة قطع مكافئ وتأخذ الصيغة الاتية :

$Y_t = ß_0 + ß_1 t + ß_2 t^2 + e_t$ ..... (2-4)

اذ ان :

$ß_0$ و $ß_1$ و $ß_2$: تمثل معلمات الانموذج .

3.2.12.2 الانموذج غير الخطي Power (concav)

وهو احد نماذج الانحدار غير الخطية الذي يستند على المعادلة الاتية :

$Y_t = ß_0 x^{ß_1}$……..(2-5)

وهـي معادلـة غيـر خطيـة ، لتحويلهـا الـى معادلـة خطيـة ناخـذ اللوغـاريتم الطبيعـي للطـرفين ينتج :

$lnY_t = \ln ß_0 + ß_1 \ln x$ ……..(2-6)

اذ ان المعادلـة (6–2) اصـبحت معادلـة انحـدار خطـي بسـيط باضـافة حـد الخطـا اليهـا تصبح :

$Y'_t = ß_0' + ß_1 + \boldsymbol{e_t}$ ……..(2-7)

4.2.12.2 انمــوذج التمهيــد الاســي [8] Smoothingexponentiamodel

يقصد بالتمهيد محاولة تقليل التغايرات في قيم السلسلة عن خط المنحنى الذي يمثل النمط العام للسلسلة. ويعتمد التمهيد الأسي على حساب مايعرف باسم المتوسط المتحرك Moving Average كوسيلة تمهيد؛ وهو عبارة عن سلسلة من المتوسطات الحسابية لقيم متتالية من السلسلة تتحرك على طول زمن السلسلة لتشكل سلسلة جديدة. ويحسب المتوسط المتحرك وفق ترتيب معين يتمثل في طول الفترة أو عدد قيم السلسلة المتتالية التي تدخل في الحساب. فمثلاً، باعتبار Y_t يمثل قيمة السلسلة، فإن القيمتين الأولى والثانية في المتوسط المتحرك من الترتيب الثالث (طول الفترة ثلاث سنوات) $MA(3)$، يتم حسابهما كمايلي:

المتوسط الأول= $$MA_1(3) = \frac{Y_1 + Y_2 + Y_3}{3} \quad \text{.....(2-8)}$$

المتوسط الثاني = $$MA_2(3) = \frac{Y_2 + Y_3 + Y_4}{3} \quad \text{.....(2-9)}$$

وتأثير المتوسط المتحرك التمهيدي يعتمد على ترتيبه؛ فكلما زاد ترتيبه أنتج سلسلة أفضل تمهيداً. ولكن زيادة ترتيب المتوسط تأتي على حساب طول السلسلة ؛ فكلما زادت قيمة الترتيب تقلص طول السلسلة المتكونة من قيم المتوسط المتحرك.

وربما يكون المتوسط المتحرك المنفرد Single Moving Average غير كاف لتحقيق تمهيداً مقبولاً، فيلجأ إلى حساب مايسمى التمهيد الأسي المضاعف Double Moving Average، وهو المتوسط المتحرك للمتوسط المتحرك، ولكن هذا سيكون على حساب طول السلسلة.

وخلاف وظيفته التمهيدية، فإن المتوسط المتحرك يمكن استخدامه للتوقع لنقطة زمنية واحدة للأمام، باستعمال عدد من النقاط الزمنية الماضية. فباختيار عدد k من النقاط الزمنية، يمكن التوقع لقيمة السلسلة عند النقطة المستقبلية t بأنها:

$$\hat{y}_t = \frac{y_{t-1} + y_{t-2} + \dots + y_{t-k}}{k} \quad \text{.....(2-10)}$$

هذا لايعني أن آخر قيمة في السلسلة يتم التوقع لها بمتوسط جميع قيم السلسلة، ولكن k هنا يمثل ترتيب المتوسط المتحرك، ومن ثم سوف يكون هناك سلسلة من المتوسطات، كل منها

يتوقع لنقطة زمنية معينة بناءً على عدد k من قيم السلسلة السابقة للنقطة. فمثلاً، إذا كان لدينا سلسلة بطول 24، وباعتبار $k = 4$، فإن القيمتين رقم 24 و 25 يتم التوقع لهما كما يأتي:

$$\hat{y}_{24} = \frac{y_{23} + y_{22} + y_{21} + y_{20}}{4} \quad \text{.....(2-11)}$$

$$\hat{y}_{25} = \frac{y_{24} + y_{23} + y_{22} + y_{21}}{4} \quad \text{.....(2-12)}$$

ومن المحددات المهمة للقيمة المتوقعة بالمتوسط المتحرك طول المدة المستعملة في الحساب (قيمة k)؛ كلما زاد طول الفترة زادت درجة التمهيد، ولكن ربما على حساب درجة خطأ التوقع.

أيضاً، توقع المتوسط المتحرك لايعمل اعتبار إلا لقيم المدة الأخيرة في السلسلة، مهملاً قيم جميع المدد السابقة، ثم إنه يعطي القيم الماضية وزناً واحداً، بينما القيمة المستقبلية غالباً ماتتأثر بشكل أكبر بالقيمة الأحدث (الأقرب لها).

وللتغلب على عيوب المتوسط المتحرك ظهر أسلوب التمهيد الأسي، والذي يعمل اعتباراُ لجميع القيم السابقة بناء على مايسمى المتوسط المتحرك الموزون Weighted Moving Average، والذي يستعمل معامل معين (يطلق عليه غالباً أسم ألفا Alpha، ويرمز له بـ α)، لوزن تأثير القيم السابقة، تتراوح قيمته بين واحد صحيح وصفر: واحد صحيح يعني أن القيمة المتوقعة حسبت على أساس آخر قيمة في السلسلة، وصفر تعني أن القيم السابقة جميعاً أعطيت وزناً متساوياً في التوقع للمستقبل.

وتكتب صيغة التوقع باستعمال الإنموذج الأسي رياضياً كمايأتي:

$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)\hat{y}_{t-1} \quad \text{.....(2-13)}$$

اذ أن $\hat{y}_t$ تمثل القيمة المتوقعة للسلسلة عند الزمن t، و y_{t-1} تمثل قيمة السلسلة عند الزمن $t-1$، و $\hat{y}_{t-1}$ تمثل القيمة المتوقعة للسلسلة عند $t-1$، و α يمثل معامل لحساب المتوسط المتحرك الموزون.

ويتميز إنموذج التمهيد الأسي عن المتوسط المتحرك باشتراك جميع القيم السابقة للسلسلة في التوقع ، لأن الجزء الثاني من الطرف الأيمن من المعادلة المذكورة أنفا عبارة عن التوقع للقيمة السابقة للسلسلة ، والذي بدوره يتكون من حاصل التوقع للقيمة السابقة لها وقيمة السلسلة الفعلية السابقة لها، وهكذا تباعاً لبقية قيم السلسلة كما يتضح مماياتي:

$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)\hat{y}_{t-1}$$
$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)[\alpha y_{t-2} + (1-\alpha)\hat{y}_{t-2}]$$
$$\hat{y}_t = \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + (1-\alpha)^2 \hat{y}_{t-2}$$
$$\hat{y}_t = \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + (1-\alpha)^2[\alpha y_{t-3} + (1-\alpha)\hat{y}_{t-3}]$$
$$\hat{y}_t = \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + \alpha(1-\alpha)^2 y_{t-3} + (1-\alpha)^3 \hat{y}_{t-3}$$
$$\vdots$$
$$\hat{y}_t = \alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} + \ldots + (1-\alpha)^{t-2} y_1 \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2\text{-}14)$$

أما معنى " الأسي" فانه يعني ان المعاملات α، $\alpha(1-\alpha)$، $(1-\alpha)^2$، إلخ، تتناقص بحسب تأثير قوة أسية. وصغر قيمة ألفا مثلها مثل طول الفترة المستعملة في حساب المتوسط المتحرك، يعني ضمان درجة أكبر من التمهيد، ولكن ربما على حساب دقة التوقع. وعموماً، فاختيار قيمة ألفا مسألة تخضع للحكم الذاتي (Subjective Choice)، ولذلك كثيراً مايترك للبرنامج الإحصائي اختيارها على أساس أقل درجة خطأ، عن طريق طريقة "بحث الشبكة Grid Search"

وهذه الصيغة يطلق عليها اسم الإنموذج الأسي البسيط Simple Exponential Smoothing Model. وبالرغم من أنها تمكن من التغلب على مشكلتين مهمتين من مشاكل التوقع الناتج من المتوسط المتحرك، كما أشير إلى ذلك أنفاً ، إلا إنها مثلها مثل المتوسط المتحرك، لاتتوقع إلا لقيمة مستقبلية واحدة؛ أي أنها تفترض عدم وجود نمطية في السلسلة، وبالتالي فجميع القيم المستقبلية تساوي القيمة المستقبلية الأولى ($\hat{y}_t = \hat{y}_{t+1} = \hat{y}_{t+2} = \cdots$).

ولذلك ظهرت نماذج تمهيد أسي أكثر تعقيداً، تعمل اعتباراً لنمط السلسلة (شكل المنحنى العام الذي تتبعه السلسلة)؛ فهناك إنموذج التمهيد الأسي ذو النمط الخطي Linear Model، وإنموذج التمهيد الأسي ذو النمط الأسي Exponential Model، وإنموذج التمهيد الأسي ذو النمط الهابط Dampened Model.

ويعبر كل من النمط الخطي والأسي عن حالة النمو (الزيادة) في قيمة السلسلة، ولكن النمط الخطي يعبر عن الزيادة الثابتة بينما يعبر النمط الأسي عن الزيادة المتزايدة؛ بمعنى أن

النمط الخطي يعبر عن الحالة التي تنمو فيها قيم السلسلة بمقدار شبه ثابت، أما النمط الأسي فيعبر عن الحالة التي تنمو فيها قيم السلسلة بمعدل (نسبة) شبه ثابتة. وعن طريق شكل السلسلة بيانياً، يتبين النمط الخطي من كون السلسلة تتبع تقريباً خطاً مستقيماً صاعداً، في حين يتبين النمط الأسي من كون السلسلة تتبع خط صاعد منحنٍ إلى أعلى. وفضلا عن ثابت التمهيد (ألفا)، يتأثر التوقع المستقبلي للنمطين الخطي والأسي بثابت النمط (يطلق عليه اسم جاما Gamma،Γ)؛ تتراوح قيمته بين الصفر والواحد: صفر يعني أن النمط بني على جميع قيم السلسلة بشكل متساوٍ، وواحد يعني أن النمط بني فقط على القراءات الأخيرة في السلسلة). أما النمط الهابط فيشير إلى الحالة التي تتناقص فيها قيم السلسلة في طريقها إلى أن تتلاشى. وهنا، وبالإضافة إلى معامل التمهيد، يحكم التوقع المستقبلي ثابت آخر يطلق عليه اسم معامل الهبوط (يعرف باسم فاي، ϕ)، تتراوح قيمته كذلك بين الواحد والصفر؛ القيمة القريبة من الصفر تعمل على إدخال جميع قيم السلسلة في تقدير الهبوط المستقبلي، بينما القيمة القريبة من الواحد تتأثر بشكل أكبر بالقيم الأحدث في تقدير مدى هبوط قيم السلسلة مستقبلاً.

وجملة القول أن نماذج التمهيد الأسي تتميز عن المتوسط المتحرك بأنها تأخذ في الاعتبار جميع قيم المددالسابقة، ولكن بحسب وزن معين لتقدير تأثيرها على التوقع المستقبلي، إضافة إلى إمكانية نمذجة النمط فيما لوظهر في السلسلة، سواء أكان ذلك نمطياً خطياً أم أسياً أم هابطاً.

13.2 نمـــاذج بـــوكس جنكينـــز [4],[25] Box–JenkinsModels

تعتمد النماذج الاحصائية للسلسلة الزمنية بالدرجة الاولى على طبيعة الظاهرة المدروسة وعن طريق الاعتماد على الاساليب الاحصائية المتقدمة يمكن التوصل الى الصيغ الدقيقة لمثل هذه الظواهر، ويعد أسلوب تحليل السلاسل الزمنية الذي قدمه الباحثان بوكس وجنكينز في العام (1970) أحد الأدوات الإحصائية التي يمكن الاعتماد فيها في تحليل السلاسل الزمنية واستعمالها لغرض التنبؤ، ويتمثل هذا الأسلوب في المراحل (التعريف – التشخيص – التقدير – التنبؤ) وقد أحدث هذا الأسلوب نقلة نوعية متميزة في مجال تحليل السلاسل الزمنية إذ أصبح أكثر الطرائق انتشاراً واستعمالاً من قبل الباحثين لِما يتصف به من مزايا متعددة تؤدي في النهاية إلى الحصول على نظام نمذجة وتنبؤ موثوقاً بهِ في معظم السلاسل الزمنية التي تحدث في المجالات المختلفة.

وان هذه الطريقة لا تفرض نمط معين في بيانات السلسلة الزمنية قبل تطبيقها وإنما تبدأ بأنموذج تجريبي يحدد بالاعتماد على كل من دالة الارتباط الذاتي Autocorrelation

Function ودالة الارتباط الذاتي الجزئي Partial Autocorrelation function يتم تقدير معلمات الأنموذج بالاعتماد على قيم السلسلة الزمنية لاسيما التي تجعل أخطاء التنبؤ أقل ما يمكن ويستعمل في هذه الطريقة عدد من المؤشرات والتي تجعل الباحث قادراً على الحكم فيما إذا كان الأنموذج ملائماً أم لا، فإذا كان الأنموذج ملائماً فانه يستعمل مباشرة في التنبؤ أما إذا لم يكن كذلك فان هذه الطريقة توفر للباحث معلومات عن كيفية الحصول على الأنموذج الملائم. ومن ابرز اسباب استعمال نماذج بوكس جنكينز هي :

1) تستخدم في حالات السلسلة الزمنية المستقرة اما غير المستقرة يتم تحويلها الى سلسلة مستقرة.
2) لا تكتفي هذه الطريقة بمعالجة النماذج ذات المتغير الواحد فقط وانما تقوم ايضاً بمعالجة النماذج ذات المتغيرات المتعددة.
3) الدقة الكبيرة لهذه الطريقة لكونها تمر بمراحل مهمة بغية استخلاص إنموذج التنبؤ المناسب.
4) يعد الافق الزمني المطلوب للتنبؤ بسبب قلة خطأ التنبؤ.

ان منهجية هذا الاسلوب تعتمد في صياغتها على ثلاثة نماذج هيَ إنموذج الانحدار الذاتي Autoregressive Model وإنموذج المتوسطات المتحركة Moving Average Model والإنموذج المختلط Auto Regressive – Moving Average Model .

1.13.2 إنموذج الانحدار الذاتي[4],[19] Autoregressive Model (AR)

يعرف على انه القيمة الحالية للسلسلة الزمنية في إنموذج الانحدار الذاتي بدلالة المجموع الموزون للقيم السابقة للسلسلة الزمنية (Y_{t-1} , Y_{t-2},…, Y_{t-p}) مضافا اليه الخطأ العشوائي، فأن الصيغة العامة للانموذج من الرتبة p الذي يرمز له AR(p) وان AR(p)=ARMA(p,0) :

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \dots + \emptyset_p Y_{t-p} + a_t \qquad \dots\dots\dots\dots(2-15)$$

ويمكن كتابة المعادلة اعلاه بشكل عامل الازاحة الخلفي :

$$\left(1 - \emptyset_1 B^1 - \dots - \emptyset_p B^p\right) Y_t = a_t$$

$$\emptyset_p(B) Y_t = a_t$$

اذ ان

Y_t: يمثل قيم الظاهرة المدروسة في الزمن t

$\emptyset_1 , \emptyset_2, \ldots \emptyset_p$، : هي معلمات الانموذج .

$Y_{t-1}, \ldots\ldots, Y_{t-p}$: هي قيم الظاهرة المرتدة زمنيا خلال الزمن t.

a_t: الخطأ العشوائي المستقل ويسمى التشويش الابيض (White Noise) بمتوسط مقداره صفر وتباين σ_a^2 وتغاير ذاتي مقداره صفراً. اي ان $a_t \sim N(0, \sigma_a^2)$

$$E(a_t) = 0$$

$$E(a_t^2) = \sigma_a^2$$

$$E(a_t, a_j) = 0 \quad for \;\; t \neq j$$

إن دالة الارتباط الذاتي للأإنموذج AR(p) تكون متناقصة بشكل أسي أو بشكل موجات جيبيه متضائلة ، أما دالة الارتباط الذاتي الجزئي فإنها تنقطع بعد الإزاحة p .

2.13.2 انموذج المتوسطات المتحركة[4] Moving Average Model (MA)

في هذا الانموذج يُعبر عن القيمة الحالية للسلسلة الزمنية Y_t بدلالة المجموع الموزون للقيم السابقة للأخطاء $(\boldsymbol{a_t}, \boldsymbol{a_{t-1}}, \ldots, \boldsymbol{a_{t-q}})$ والصيغة العامة لهذا الإنموذج من الرتبة q الذي يرمز له MA(q) وان MA(q)=ARMA(0,q):

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-16)$$

او

$$Y_t = \emptyset(B) a_t$$

اذ ان

$\theta_1 , \theta_2, \ldots \theta_p$: معلمات انموذج المتوسطات المتحركة .

$a_{t-1} , a_{t-2} , \ldots , a_{t-q}$: تمثل الاخطاء السابقة في الأوقات t−q , … , t−2 , t−1.

إن دالة الارتباط الذاتي لأنموذج MA(q) تنقطع بعد الإزاحة q ، أما دالة الارتباط الذاتي الجزئي فإنها تتناقص بشكل أسي أو بشكل موجات جيبيه متضائلة.

3.13.2 الانموذج المختلط[10] (ARMA) Auto – Moving Average Model Regressive

ان الكثير من السلاسل الزمنية لا يمكن تمثيلها بأنموذج انحدار ذاتي AR فقط او إنموذج وسط متحرك MA فقط ، لأنه غالباً ما يكون للسلسلة خواص كلا الإنموذجين وبذلك تمثل بالإنموذج المختلط(انحدار ذاتي – متوسطات متحركة) ويكتب باختصار ARMA(p,q) اذ p تمثل رتبة الانحدار الذاتي و q تمثل رتبة الوسط المتحرك .

فاذا كانت Y_t سلسلة زمنية مستقرة احادية المتغير فان صيغة الانموذج المختلط تكون بالصيغة الاتية :

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \qquad \ldots\ldots\ldots\ldots\ldots(2-17)$$

نلاحظ من المعادلة(2−17) ان الإنموذج هو انحدار ذاتي من رتبة p وان الخطأ له يتبع إنموذج الأوساط المتحركة من الرتبة q لذلك نسميه بالإنموذج المختلط ، فعندما يكون q=0 نحصل على إنموذج انحدار ذاتي من رتبة p و يرمز له AR(p) وعندما يكون p=0 نحصل على إنموذج وسط متحرك من رتبة q ويرمز له MA(q) .

وبكتابة الإنموذج في المعادلة (2−13) بصيغة متعددة حدود Polynomial Form يكون كالاتي:

$$(Y_t - \varphi_1 Y_{t-1} - \varphi_2 Y_{t-2} - \ldots\ldots - \varphi_p Y_{t-p}) = (a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots\ldots - \theta_q a_{t-q})$$

$$(1 - \varphi_1 B - \varphi_2 B^2 - \ldots\ldots - \varphi_p B^p) Y_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots\ldots - \theta_q B^q) a$$

$$\therefore \phi_p(B) Y_t = \theta_q(B) a_t \qquad \ldots\ldots\ldots\ldots(2-18)$$

اذ ان

$$\emptyset_p(B) = 1 - \emptyset_1 B^1 - \emptyset_2 B^2 - \cdots - \emptyset_p B^p \qquad \ldots\ldots\ldots (2-19)$$

$$\theta_q(B) = 1 - \theta_1 B^1 - \theta_2 B^2 - \cdots - \theta_q B^q \qquad \ldots\ldots\ldots\ldots(2-20)$$

وان B يمثل عامل الازاحة الخلفي (Backshift Operator) الذي يعرف بالصيغة الاتية :

$$B^j Y_t = Y_{t-j} \quad \forall\, j = (1,2,\dots) \quad \dots\dots\dots\dots\dots(2-21)$$

اي ان عامل الازاحة B يعيد السلسلة Y_t الى Y_{t-1} (قيمة واحدة الى الخلف) اذ أن

$$BY_t = Y_{t-1} \quad \dots\dots\dots(2-22)$$

$\emptyset_p(B)$: عبارة عن متعددة حدود في (B) من الرتبة p لمعلمات إنموذج الأنحدار الذاتي $(\emptyset_1, \emptyset_2, \dots . \emptyset_p)$

$\theta_q(B)$: عبارة عن متعددة حدود في (B) من الرتبة q لمعلمات إنموذج الوسط المتحرك $(\theta_1, \theta_2, \dots . \theta_q)$

a_t : عبارة عن سلسلة من الأخطاء العشوائية المستقلة (التشويش الابيض) وبتوزيع $N(0, \sigma_a^2)$

ان دالة الارتباط الذاتي للانموذج بعد الازاحة (q−p) تسلك سلوكا مشابها الى دالة الارتباط الذاتي للإنموذج AR(p) (تتناقص بشكل أسي أو بشكل موجات جيبية). اما دالة الارتباط الذاتي الجزئي له بعد الازاحة (p−q) فأنها تسلك سلوكا مشابها الى دالة الارتباط الذاتي الجزئي للإنموذج MA(q) (تتناقص بشكل أسي أو بشكل موجات جيبية).

ان هذه النماذج التي ذكرناها لا يمكن تطبيقها لمعالجة السلاسل الزمنية غير المستقرة الا بعد تحويلها الى سلاسل زمنية مستقرة.

4.13.2 الانموذج المتكامل المختلط[25] Integrated Mixed Model (ARIMA)

وصف الباحثان بوكس وجنكينز الانموذج المختلط (ARMA) على نحو اشمل ووضعا اسلوباً لفهم عدم الاستقرارية ومعالجتها في حال كون السلسلة الزمنية غير مستقرة عن المتوسط أذ يتم تحويلها الى سلسلة زمنية مستقرة وذلك بأخذ الفروق اللازمة لها ويرمز لإنموذج (ARMA) بعد اخذ الفروق للسلسلة الزمنية بالرمز (ARIMA) ، اذ يتم اخذ الفروق باعداد صحيحة موجبة من الدرجة d فأن Y_t هيَ عملية (ARIMA) اذا كان:

$$\nabla^d Y_t = (1-B)^d Y_t \quad \dots\dots\dots\dots\dots\dots(2-23)$$

اذ ان Y_t هي عملية ARMA وان $(1-B)^d$ يمثل عامل الفروق الخلفية (Backward Difference Operation) . اذ ان:

$\nabla Y_t = (1-B)Y_t = Y_t - Y_{t-1}$ .....(2−24)

$\nabla^2 Y_t = (1-B)^2 Y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$ .....(2−25)

.

.

$\nabla^d Y_t = (1-B)^d Y_t$; d=1,2,3,...

والإنموذج (ARIMA) للسلسلة الزمنية Y_t يكون بالشكل الاتي :

$\emptyset_p(B)(1-B)^d Y_t = \theta_q(B) a_t$ .....(2−26)

حيث ان كل المفاهيم النظرية لنماذج ARMA(p,q) يمكن تطبيقها على نماذج ARIMA(p,d,q) .

14.2 منهجية بوكس – جنكينز[26] Box – Jenkins Approach

تتضمن طريقة بوكس وجنكينز عدد من الخطوات كما يلي :

1.14.2 جمع البيانات[8] Data Collection

افترض الباحثان بوكس وجنكينز (1970) توفر على الأقل (50) مشاهدة هي الحد الأدنى في السلسلة وهناك من يرى استخدام حجم عينة اقل يصل الى (30) مشاهدة كحد أدنى لتحليل وبناءإنموذج للسلسلة الزمنية وكلما كان عدد المشاهدات كبير كلما كان الإنموذج المتوقع الحصول عليه للتنبؤ أفضل وادق .

2.14.2 فحص استقرارية السلسلة[25],[12]

Check Stationary Of Time Series

المقصود باستقرارية السلسلة الزمنية أن يكون المتوسط والتباين و دالة الارتباط الذاتي لا تتغير مع الزمن ويمكن القول أن السلسلة الزمنية مستقرة هي التي تكون متجانسة زمنياً أي أن رسم السلسلة في المدة (t,t+h) يكون مطابقاً لرسم السلسلة في مدة أخرى (s,s+h) وفي الحقيقة أن حالة الاستقرارية هي نادرة جداً وتوضع لغرض تسهيل التعامل الرياضي مع السلاسل الزمنية وهي على نوعين:

1) استقرارية تامة Strictly Stationary

تكون السلسلة الزمنية (Y_t ; t=1,2,…,n) مستقرة بشكل تام عندما يكون التوزيع المشترك لأية مجموعة من المشاهدات لا يتأثر بإزاحة كل المددالزمنية للمشاهدات الى الأمام أو الى الخلف بأية كمية صحيحة ويمكن التعبير عن ذلك بطريقة الاحتمالات بالصيغة الاتية:

$$\Pr(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_m}) = \Pr(Y_{t_1+k}, Y_{t_2+k}, \ldots, Y_{t_m+k}) \qquad \ldots\ldots\ldots\ldots(2-27)$$

اذ أن

t : تمثل أي مدة زمنية.

m , k : تمثل أي زوج من القيم الصحيحة.

2) استقرارية من الدرجة الثانية[12] Second - order Stationary

تكون السلسلة الزمنية (Y_t ; t=1,2,…,n) ذات استقرارية من الدرجة الثانية أذا تحققت الشروط الاتية :

i) $E(Y_t) = \mu \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2-28)$

اذ أن μ ثابت الانموذج .

ويكون ثابت لجميع قيم t ويقدر من مشاهدات السلسلة الزمنية كالاتي :

$$\hat{\mu} = \bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2-29)$$

ii) $\operatorname{var}(Y_t) = E(Y_t - \mu)^2 = \sigma_y^2 = \gamma_o \ldots\ldots\ldots\ldots(2-30)$

اذ ان

$\hat{\gamma}_o$: تباين العملية العشوائية ويكون ثابت لجميع قيم t ويقدر من مشاهدات السلسلة الزمنية بالصيغة الاتية :

$$\hat{\gamma}_o = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \bar{Y})^2 \quad .....(2-31)$$

$$\mathrm{cov}(Y_t, Y_{t+k}) = \mathrm{E}(Y_t - \mu)(Y_{t+k} - \mu) = \gamma_k \quad .....(2-32)$$

اذ ان :

γ_k : هو التغاير الذاتي (Auto covariance) للعملية العشوائية عند الإزاحة k اي عند (lag k) ويكون ثابتا لجميع القيم الصحيحة الى k والذي يقدر من مشاهدات السلسلة الزمنية كالاتي:

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)\left(Y_{t+k} - \bar{Y}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-33)$$

3.14.2 عدم الاستقرارية في السلاسل الزمنية[25]

Non-Stationary Time Series

ان حالة الاستقرارية هي حالة نادرة جداً ، فكثير من السلاسل الزمنية التي تواجهنا في التطبيق العملي تكون غير مستقرة وللتعرف على حالة عدم الاستقرارية في السلاسل الزمنية يمكن استعمال اختبار جذر الوحدة Uint Root Test وكذلك دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) ، وهناك نوعين من عدم الاستقرارية في السلسلة الزمنية:

الأول: عدم الاستقرارية حول المتوسط Non-Stationary about the mean

عندما لاتتذبذب السلسلة الزمنية عن متوسط ثابت فالسلسلة الزمنية تكون غير مستقرة عن المتوسط وفي هذه الحالة يتم أخذ الفروق المناسبة لتحقيق الاستقرارية في السلسلة الزمنية.

$$(1-B)\mathrm{Y_t} = \mathrm{Y_t} - \mathrm{Y_{t-1}} = \nabla \mathrm{Y_t} \quad ..... (2-34)$$

$$(1-B)^2\mathrm{Y_t} = (\mathrm{Y_t} - \mathrm{Y_{t-1}}) - (\mathrm{Y_{t-1}} - \mathrm{Y_{t-2}}) = \nabla^2 \mathrm{Y_t} \quad .....(2-35)$$

$$(1-B)^{\mathrm{d}}\mathrm{Y_t} = \nabla^d \mathrm{Y_t} \quad ..... (2-36)$$

اذ d=1,2,3 ,...

وعملية الفروق والتي يرمزلها (d) للبيانات جعل السلسلة الجديدة أقل من السلسلة القديمة بنقطة واحدة والتي تكون اول قراءة ، وتم استخدام الاستقرارية عن المتوسط في الجانب العملي .

الثاني: عدم الاستقرارية حول التباين

Non-Stationary around the variance

وفي هذه الحالة نجد أن السلسلة الزمنية لا تتذبذب حول تباين ثابت ومن أجل تحقيق الاستقرارية في السلسلة الزمنية يتم إجراء التحويل اللوغاريتمي أو الجذر التربيعي أو الأُسي فضلاً عن طرائق اخرى للسلسة الزمنية.

ويوجد عدة اختبارات للكشف عن أستقرارية أو عدم أستقرارية السلسلة الزمنية منها :

4.14.2 اختبار جذر الوحدة [4],[21] Unit Root Test

هوَ احد الأساليب الحديثة في تحديد الاستقرارية اذ ان السلاسل الزمنية المستقرة تقع جذورها خارج دائرة الوحدة. ومن بين الاختبارات المهمة للكشف عن استقرارية السلسلة هو اختبار ديكي فولر الموسع(Augmented Dickey–Fuller) اذ يعتمد على تقدير الإنموذج الاتي:

$$\nabla Y_t = \gamma Y_{t-1} + \sum_{j=1}^{k} \emptyset_j \, \nabla Y_{t-j} + a_t \quad \ldots\ldots\ldots\ldots\ldots(2-37)$$

اذ ان

Y_t: يمثل قيمة الظاهرة في الزمن t .

γ : يمثل معامل التغاير الذاتي للعملية العشوائية .

a_t: الخطأ العشوائي او الضوضاء الابيض (White Noise) .

∇ : يمثل معامل الفروق.

اذ يقوم اختبار ديكي فولر على الفرضيتين الأتيتين:

$H_0 : \emptyset_j = 0$ (وجود جذر الوحدة اي عدم استقرار السلسلة)

$H_1 : \emptyset_j \neq 0$ (عدم وجود جذر الوحدة اي استقرار السلسلة)

عندما تكون $\emptyset_j = 0$ فذلك يدل على عدم الاستقرارية وأن البيانات تعاني من الجذر الأحادي ، وقد درس ديكي فولر التوزيع الأحصائي للمقدر $\emptyset_j$ وأعد جدولاً للقيم الحرجة أي T_{tab} ليتم مقارنتها مع T_{cal} المحسوبة اذ أن :

$$T_{cal} = \frac{\emptyset_j}{\sigma_{\emptyset_j}} \quad \text{.....(2-38)} \qquad ; j = 1 , 2 ,\ldots, k$$

يتم رفض فرضية العدم H_0 اذا كانت القيمة المحسوبة اكبر او تساوي القيمة الجدولية اي ان $(T_{cal} \geq T_{tab})$ وقبول الفرضية البديلة H_1 اي عدم وجود جذر الوحدة وبالتالي استقرار السلسلة، أما إذا كانت $(T_{cal} < T_{tab})$ فهذا يعني قبول فرضية العدم H_0 وفي هذه الحالة يوجد جذور وحدة وان السلسلة غير مستقرة .

5.14.2 دالة الارتباط الذاتي[12],[25] Auto Correlation Function (ACF)

تعد واحدة من الاساليب المهمة في معرفة استقرارية السلسلة الزمنية فضلاً عن الوسط الحسابي والتباين الثابتين وكذلك لها دور في تشخيص نماذج السلسلة الزمنية عن طريقها وعن طريق دالة الارتباط الذاتي الجزئي وتعرف دالة الارتباط الذاتي Auto Correlation Function والذي يرمز لها بالرمز (ACF) بأنها مقياس لدرجة العلاقة بين قيم السلسلة مع نفسها عند المددالزمنية المزاحة ، و تعد الارتباطات الذاتية صفة مميزة للعملية العشوائية فلها اهمية كبيرة لأنها احدى اساليب تحديد فيما اذا كانت العملية العشوائية مستقرة أو لا فإذا كانت كذلك فيتم اختيار احد النماذج المناسبة من مجموع نماذج العمليات العشوائية المستقرة ، والتعريف الرياضي لدالة الارتباط الذاتي هوَ قسمة التباين المشترك على التباين وان الصيغة الرياضية لإيجاد دالة الارتباط الذاتي (ACF) للعملية العشوائية عند الازاحة k هي:

$$\rho_k = \frac{\operatorname{cov}(Y_t , Y_{t+k})}{\sqrt{\operatorname{var}(Y_t)\ \operatorname{var}(Y_{t+k})}} = \frac{E(Y_t - \mu)(Y_{t+k} - \mu)}{\sqrt{E(Y_t - \mu)^2\ E(Y_{t+k} - \mu)^2}} \quad \text{.....(2-39)}$$

وعندما تكون العملية العشوائية مستقرة فإنّ المتوسط والتباين يكونان ثابتين ومتساويين لكل المدد

اذ: $\sigma_{Y_t} = \sigma_{Y_{t+k}}$

فان :

$$\rho_k = \frac{E(Y_t - \mu)(Y_{t+k} - \mu)}{\sigma_Y^2} = \frac{\gamma_k}{\gamma_0} \qquad \text{.....} (2-40)$$

ويقدر معامل الارتباط الذاتي للعينة $\hat{\rho}_k$ عند الإزاحة k من مشاهدات السلسلة الزمنية كالاتي:

$$\hat{\rho}_k = r_k = \frac{\sum_{t=1}^{n-K} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \qquad \text{.....} (2-41)$$

وان لدالة الارتباط الذاتي الخواص الاتية :

1. $r_0 = 1$
2. $r_{-k} = r_k$
3. $|r_k| \leq 1$

ومن اجل الحصول على تقديرات دقيقة فان عدد قيم السلسلة الزمنية يجب ان يكون كبيراً نسبياً ولاستقرار السلسلة يجب ان يكون ($r_k = 0$) أو ان لا يختلف جوهريا عن الصفر بالنسبة لأي ازاحة اي ان تقع معاملات الارتباط الذاتي داخل حدود فترة الثقة بمستوى 95% فإذا وقع خارج حدود الثقة لفترة طويلة فأن معاملات الارتباط الذاتي (ACF) تختلف عن الصفر معنويا لعدد كبير نسبيا من الازاحات لذا يقال ان السلسلة غير مستقرة. أذ أن معاملات الارتباط الذاتي للسلسلة المستقرة لها توزيع طبيعي غالبا وسطه الحسابي صفر وتباينه $(1/n)$ فحدود فترة الثقة عند مستوى معنوية (0.05) تتمثل بالصيغة الاتية:

$$C.I = \pm 1.96 * 1/\sqrt{n} \qquad \text{.....}(2-42)$$

فأن كان r_k يقع داخل حدود الثقة يتم قبول فرضية العدم $H_\circ : \ r_k = 0$

وان كان r_k يقع خارج حدود الثقة يتم قبول الفرضية البديلة $H_1 : \ r_k \neq 0$

6.14.2 دالة الارتباط الذاتي الجزئي[25]

PACF (Partial Auto Correlation Function)

توجد هناك دالة أخرى تسمى دالة الارتباط الذاتي الجزئي ويشار اليها أختصاراً (PACF) وتسهم Y_{t+k} و Y_t في تشخيص الأنموذج الانسب ، وتعتبر هذه الدالة مؤشراً جيداً لقياس العلاقة بين للسلسلة نفسها مع أفتراض ثبوت بقية قيم السلسلة الزمنية وتقوم أيضاً في عملية تحديد رتبة الأنموذج ونوعه لذلك فهي إحدى الأدوات المستعملة في تحليل السلاسل الزمنية ويمكن تقديره بواسطة المربعات الصغرى أو مجموعة المعادلات التقريبية لتقدير الارتباط الذاتي الجزئي ويتم حسابها وفق الصيغة :

$$\rho_i = \emptyset_{k1}\,\rho_{i-1} + \emptyset_{k2}\,\rho_{i-2} + \cdots + \emptyset_{kk}\,\rho_{i-k} \qquad ; i = 1, \ldots, k$$

ويتم الحصول على مجموعة المعاملات باستعمال قاعدة كَريمير (Gramer's) للمعاملات الخاصة بالارتباطات الذاتية وتكون معاملات الارتباط الذاتي الجزئي بالشكل $\emptyset_{kk} = \rho_k^{-1}\,\rho_k$

اذ ان :

$\emptyset_{kk}$: متجه معاملات (PACF) ذات الرتبة $k\times 1$

ρ_k : مصفوفة معاملات (ACF) ذات الرتبة $k\times k$ وعناصر قطرها مساوية للواحد .

ولحساب $\widehat{\emptyset}_{kk}$ نستعمل الصيغة الاتية :

$$\widehat{\emptyset}_{k+1,k-1} = \frac{\widehat{\rho}_{k+1} - \sum_{j=1}^{k} \widehat{\emptyset}_{kj}\,\widehat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \widehat{\emptyset}_{kj}\,\widehat{\rho}_j} \quad \ldots\ldots\ldots\ldots \quad (2-43) \qquad ; \; j = 1,2,\ldots,k$$

$$\widehat{\emptyset}_{k+1,j} = \widehat{\emptyset}_{kj} - \widehat{\emptyset}_{k+1,k+1}\widehat{\emptyset}_{k,k+1-j}$$

وباستعمال قاعدة كريمير نحصل على:

$\emptyset_{11} = \rho_1$

$$\emptyset_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

اذ أن الحالة العامة لمعادلات الارتباط الذاتي الجزئي $\emptyset_{kk}$ وكما يأتي:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdot & \cdot & \cdot & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{k-3} & \rho_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \cdot & \cdot & \cdot & \cdot & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdot & \cdot & \cdot & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdot & \cdot & \cdot & \rho_{k-3} & \rho_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \cdot & \cdot & \cdot & \cdot & \rho_1 & 1 \end{vmatrix}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2-44)$$

إن المعلمات($\emptyset_{kk}$; k: 1,2, ...) تسمى بدوال الارتباط الذاتي الجزئي(PACF)، وهيَ أداة أخرى مهمة إضافة الى دالة الارتباط الذاتي(ACF) في تحليل السلاسل الزمنية إذ تساعدنا في تحديد درجة الأنموذج AR(p) أو MA(q) وكذلك الأنموذج المختلط ARMA(p,q) الملائم لتمثيل بيانات السلسلة الزمنية.

ويمكن الافادة من تحليل دالتي الارتباط الذاتي والارتباط الذاتي الجزئي في هذه المجالات :

1) بواسطة معاملات دالتي الارتباط الذاتي والارتباط الذاتي الجزئي يمكن تحديد فيما إذا كانت مجموعة المشاهدات عشوائية ام غير عشوائية إذ يتم اختبار معاملات الارتباط الذاتي لعدد من الازاحات لمعرفة مدى اختلافها معنوياً عن الصفر وكذلك عن طريق رسم معاملات الارتباط الذاتي لتحديد وجود الاتجاه (Trend) أو عدم وجوده في المشاهدات وبعد إيجاد الإنموذج الملائم للسلسلة، يمكن حساب معاملات الارتباط للدالتين لسلسلة البواقي (الأخطاء) (Residual) لتحديد مدى عشوائيتها.
2) وجود الاستقرارية أو عدم الاستقرارية في السلسلة، والمقصود بالاستقرارية عدم وجود نمو أو هبوط في المشاهدات بعبارة أخرى إنّ المشاهدات تكون أفقية عن المحور السيني .
3) تحديد الموسمية في السلسلة الزمنية ، ويقصد بالموسمية وجود نمط معين في السلسلة يعيد نفسه خلال مدد زمنية محددة، مثل ازدياد مبيعات سلعة معينة في الشتاء وانخفاضها في الصيف .
4) تحديد رتبة الأنموذج وذلك عن طريق سلوك دالتي الارتباط الذاتي والارتباط الذاتي الجزئي.

15.2 تحديد رتبة الانموذج[2],[6],[15] Order Model selection

بعد فحص استقرار السلسلة الزمنية تبدأ مرحلة تشخيص الإنموذج إذ يتم تحديد الإنموذج الملائم ورتبته ويتم ذلك عن طريق دراسة السلوك العملي لدالتي الارتباط الذاتي (ACF) ودالة الارتباط الجزئي (PACF) للسلسلة المدروسة، فأذا تناقصت معاملات (ACF) اسياً وانقطعت معاملات (PACF) بعد الازاحة (P) في الوقت نفسه فأن الإنموذج الملائم هو AR(P) ، أما اذا كانت معاملات (PACF) تتناقص اسياً ومعاملات (ACF) تنقطع بعد الازاحة (q) في الوقت نفسه فأن الإنموذج الملائم هو MA(q) ، أما في حالة الإنموذج المختلط ARMA(p,q) فأن معاملات (ACF) و(PACF) تتناقصان اسياً، ويتم حساب الرتبتين للإنموذج وفقاً للقطع الحاصل بعد الازاحتين(p,q)، والجدول رقم (2−1) ملخص للأنماط المختلفة لدالتي الارتباط الذاتي والارتباط الذاتي الجزئي في تحديد النماذج .

الجدول (2−1) يبين التشخيص لنماذج السلاسل الزمنية وفقاً لسلوك دالتي ACF و PACF

الإنموذج	ACF	PACF
AR (p)	تتناقص تدريجاً بشكل أسّي او تأخذ موجات الجيب المتضائلة	يقطع بعد p من الازاحات
MA (q)	يقطع بعد q من الإزاحات	تتناقص تدريجاً بشكل أسّي او تأخذ موجات الجيب المتضائلة
ARMA(p,q)	تتناقص تدريجاً بشكل أُسّي	تتناقص تدريجاً بشكل أُسّي

وفي اغلب التطبيقات هذا الاسلوب غير معتمد كثيرا وذلك لعدم تطابق السلوك النظري مع السلوك العملي للدالتين اذ ان معاملات ACF وPACF لا تعطي حكماً مطلقاً في تحديد الإنموذج ورتبته وذلك لعدم سلوك هذه المعاملات السلوك النظري اذ يؤدي الى تمثيل السلسلة الزمنية لأكثر من إنموذج . اذ يتم الاعتماد في تشخيص الإنموذج الملائم للسلسلة الزمنية وتحديد الرتبة على بعض المعايير اذ تستند اختيار رتبة الانموذج على احصاءات البواقي Residuals الناتجة من

مطابقة الإنموذج غير المتحيز unbiased، ويتم تحديد الرتبة بأختيار الإنموذج الذي يقابل أقل قيمة من هذه المعايير، وبعض هذه المعايير هي كماياتي:

1) معيار معلومة اكيكي[6] Akaike Information Criterion (AIC)

في عام 1973 اقترح العالم الياباني اكيكي (Akaike) معياراً لتحديد رتبة الانموذج واطلق عليه معيار معلومة أكيكي (Akaike Information Criterion) ويرمز له (AIC) وصيغته الرياضية هي:

$$\mathrm{AIC} = n\, Ln\,(\hat{\sigma}_a^2) + 2m \quad \dots\dots\dots\dots\dots\dots\dots\dots (2-45)$$

اذ ان

n : عدد المشاهدات

$\hat{\sigma}_a^2$: مُقدّر تباين الخطأ وصيغته ، $\hat{\sigma}_a^2 = \frac{\sum(Y_t - \hat{Y}_t)^2}{n-k}$

m : رتبة الإنموذج المختار (عدد معلمات الانموذج)

وبالنسبة لانموذج ARMA(p,q) لعدد n من المشاهدات فان معيار AIC يكون :

$$\mathrm{AIC} = n\, Ln\,(\hat{\sigma}_a^2) + 2(p+q)$$

2) معيار شوارتز[17] Schwartz Bayesian Criterion (SBC)

اقترح الباحث (Schwartz) عام 1978 معياراً مشابهاً لمعيار Akaike يسمى معيار Bayesian Criterion(Schwartz) ويرمز لهُ (SBC) وصيغته :

$$SBC_{(m)} = n\,\mathrm{Ln}\,(\hat{\sigma}_a^2) + m\, Ln(n) \quad \dots\dots\dots\dots\dots\dots\dots\dots (2-46)$$

3) معيار حنان كوين[6] Hannan – Quinn Criterion (H-Q)

في العام (1979) أقترح الباحثان (Hannan & Quinn) معياراً جديد سمي معيار (حنان – كوين) (Hannan – Quinn) ومختصره (H–Q) لتحديد الرتبة للأنموذج المدروس وصيغته كالاتي:

$$H-Q(m) = \mathrm{Ln}(\hat{\sigma}_a^2) + \frac{2m\ c\ Ln(Ln\, n)}{n} \quad ; c > 2 \quad \dots\dots\dots\dots\dots (2-47)$$

اذ ان

c : ثابت ، $c > 2$

16.2تقدير معلمات انموذج الاتجاه العام [19]

Estimation Parameters of general trend model

ان تقدير معلمات إنموذج الاتجاه العام لا تطرح اي مشكلة، ويمكن تقديرها باستعمال الكثير من الطرائق المعروفة منها طريقة التمهيد باليد وطريقة المربعات الصغرى وطريقة الاوساط المتحركة وغيرها، وسنقدر معلمات إنموذج الاتجاه العام بطريقة المربعات الصغرى وذلك لدقتها وكونها من اكثر الطرائق شيوعاً وهيَ الطريقة المستعملة في البرنامج الاحصائي Minitab كالاتي:

تعتبر من ادق الطرائق لقياس خط الاتجاه العام لتثبيت خط او منحني يُمثل نقاط قيم الظاهرة للمدد الزمنية المناظرة لها بشكل دقيق، وتنص على ان افضل خط يطابق نقاط شكل الانتشار هوَ ذلك الخط الذي تكون أبعاد نقاط شكل الانتشار عنهُ اصغر ما يمكن ويمكن تحقيق ذلك بالاشتقاق الجزئي لمجموع مربعات الاخطاء $\sum e_t^2$ نسبة لكل من قيم الثوابت ومساواة المشتقة للصفر نحصل على القيمة الصغرى لمجموع مربعات الاخطاء، وعلى فرض ان خط الاتجاه العام مستقيم فأن معادلته :

$$Y_t = ß_0 + ß_1 t + e_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2\text{-}48) \qquad ; \; t=1,2,3,\ldots\ldots\ldots,n$$

وبالتبسيط نجد ان :

$$b_0 = \bar{y} - b_1\bar{t} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-49)$$

وبالاشتقاق الجزئي نسبة للثابت b_1 نحصل على :

$$b_1 = \frac{\sum ty - n\bar{t}\bar{y}}{\sum t^2 - n\bar{t}^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-50)$$

ومعادلة الاتجاه العام التقديرية تكون :

$$\hat{y}_t = b_0 + b_1 t \qquad \dots \dots \dots \dots \dots \dots (2-51)$$

17.2 فحص مدى ملائمة الانموذج [14],[26] Diagnostic Checking

بعد إيجاد معلمات الإنموذج الملائم تأتي مرحلة فحص مدى ملائمة الإنموذج اذ يتم حساب البواقي Residuals للإنموذج الملائم (الفرق بين القيم الفعلية والقيم التقديرية) ونختبر البواقي للتأكد من صحة مدى ملائمة الإنموذج المختار وتستخدم في ذلك الفرضية الاتية :

$$H_\circ: \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

$$H_1: \rho_i \neq \rho_L \quad ; for \quad i \neq L$$

ومن الاختبارات المستعملة في التحقق من دقة الانموذج المشخص اختبار Ljung–Box وأختبار البواقي .

1.17.2 اختبار ليونغ – بوكس[14] Ljung and Box Test

قـــام الباحثـان Ljung و Box عـام 1978 بإجراء اختبار ويعتمد على توزيع الارتباطات الذاتية لبواقي الإنموذج المحدد بانه يتوزع توزيعاً طبيعياً بوسط حسابي يساوي صفراً وتباين مقداره $\frac{n-k}{n(n+2)}$ عندما يزداد حجم العينة n اي:

$$\hat{r}_k(a) \sim N\left(0 \ , \frac{n-k}{n(n+2)}\right)$$

وعليه فان صيغة إحصاءة الاختبار تكون وفق الصيغة الاتية :

$$Q_{L-B} = n(n+2) \sum_{i=1}^{k} \frac{\hat{r}_i(a)}{n-i} \sim \chi^2_{(k-p-q\ ,\alpha)} \qquad \dots\dots\dots\dots (2-52)$$

اذ يتم مقارنة قيم Q_{L-B} المحسوبة مع قيم χ^2 الجدولية بدرجة حرية (k–p–q) ومستوى معنوية α فإذا كانت القيمة المحسوبة اصغر من الجدولية دل ذلك على ان معاملات الارتباط الذاتي

للأخطاء المقدرة تتوزع عشوائيا وان الإنموذج المختار ملائم اما اذا كان العكس فأن الإنموذج غير ملائم ويتم اعادة التشخيص.

2.17.2 اختبار حدي الثقة[26] Confidence Interval Test

ويسمى ايضاً باختبار بارتلت (Bartlett Test) ، لنفرض ان سلسلة البواقي Residuals (a_t) للإنموذج المختار تتوزع توزيعاً طبيعياً وان $E(a_t\ a_{t+k}) = 0$ واختبارها بعد تقدير الارتباطات الذاتية لها $r_k\ (a)$ وان تلك المعاملات تتوزع توزيعاً طبيعياً بوسط حسابي مقداره صفراً وتباين مقداره $(\frac{1}{n})$ عندما يزداد حجم العينة n اذ ان :

$$r_k(a) \sim N(0\ , \frac{1}{n})$$

وان تقدير معامل الارتباط الذاتي للبواقي هوَ :

$$\hat{r}_k(a) = \frac{\sum_{t=1}^{n-k} a_t a_{t+1}}{\sum_{t=1}^{n} a_t^2} \qquad \dots\dots\dots\dots (2-53)$$

فاذا كانت معاملات الارتباطات الذاتية لبواقي الإنموذج المختار واقعة ضمن حدي الثقة وبمستوى 95% اي ان الأخطاء عشوائية لمعاملات الارتباط ومن ثم فان الإنموذج المشخص يكون ملائماً لتمثيل بيانات السلسلة الزمنية. اي ان :

$$pr(-1.96 * \frac{1}{\sqrt{n}} \leq \hat{r}_k(a) \leq +1.96 * \frac{1}{\sqrt{n}}) = 1 - \alpha \qquad \dots\dots\dots\dots(2-54)$$

18.2 التنبؤ[26],[8], [13] Forecasting

التنبؤ هو المرحلة الاخيرة من مراحل تحليل السلاسل الزمنية وانه لا يُمكن الانتقال الى هذه المرحلة الا بعد الانتهاء من اجراء جميع الفحوص والاختبارات الاحصائية الضرورية لتشخيص الإنموذج الذي تم اختياره اذ تمثل هذه المرحلة التطبيق العملي للإنموذج المقترح عن طريق الحصول على القيم المتوقعة للسلسلة محل الدراسة، والتنبؤ للقيم المستقبلية لمشاهدات السلسلة الزمنية هو عبارة عن التوقع الشرطي في المدة (t+L) عند الزمن (t) اذ:

$$\hat{Y}_t(L) = E[Y_{t+L} | Y_t, Y_{t-1}, Y_{t-2}, \dots]$$

فإذا كان الإنموذج الملائم والذي تم التوصل اليه عن طريق المراحل السابقة هو ARIMA(p,d,q) فأننا نستطيع التنبؤ بالقيم المستقبلية الى (L=1,2,..) مدة قادمة باستعمال معادلة $(2-17)$.

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

وللمدة L المستقبلية فان:

$$Y_{t+L} = \emptyset_1 y_{t+L-1} + \emptyset_2 y_{t+L-2} + \cdots + \emptyset_p y_{t+L-p} + a_{t+L} - \theta_1 a_{t+L-1} - \cdots - \theta_q a_{t+L-q} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-55)$$

ومن ثم التنبؤ بالإنموذج يكون كالاتي :

$$E\,Y_{t+L} = \emptyset_1 E y_{t+L-1} + \emptyset_2 E y_{t+L-2} + \cdots + \emptyset_p E y_{t+L-p} + a_{t+L} - \theta_1 E a_{t+L-1} - \cdots - \theta_q E a_{t+L-q} \quad \ldots\ldots\ldots (2-56)$$

19.2 معايير المفاضلة بين النماذج [26]

Criteria for trade-offs between models

هناك معايير متعددة يمكن عن طريقها تقييم الافضلية بين النماذج المتنبأ بها ومنها الاتي :

1- متوسط الاخطاء : Mean Error (ME)

$$\mathrm{ME} = \frac{1}{n}\sum_{t=1}^{n} e_t \quad \ldots \quad \ldots \quad \ldots \quad (2-57)$$

2- متوسط مربعات الاخطاء : Mean Sguared Error(MSE)

$$\mathbf{MES} = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \quad \ldots \ldots \ldots \quad (2-58)$$

3- متوسط الخطأ المطلق: Mean Absolute Error (MAE)

هو مؤشر يجعل الأخطاء موجبة بأخذ القيمة المطلقة لها ثم يأخذ لها المعدل وهو عبارة عن معدل الانحرافات عن القيم الحقيقية والصيغة المستعملة لمتوسط مطلق الخطأ :

$$\mathbf{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t| \qquad \dots\dots\dots\dots (2-59)$$

4- متوسط مطلق الخطأ النسبي

Mean Absolute Percentage Error(MAPE)

ويكتب وفق الصيغة الاتية :

$$\mathbf{MAPE} = \frac{\sum_{t=1}^{n} \left|\frac{y_t - \hat{y}_t}{y_t}\right|}{n} * 100 \qquad \dots\dots\dots\dots (2-60)$$

5- جذر متوسط مربعات الاخطاء :

Root mean Squared Error (RMSE)

$$\mathbf{RMES} = \sqrt[2]{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \qquad \dots\dots\dots\dots (2-61)$$

6- متوسط النسبة المطلقة للأخطاء: Mean Absolute percentage Error (MAPE)

$$\mathbf{MAPE} = \frac{1}{n}\sum_{t=1}^{n} \frac{|e_t|}{yt} * 100 \qquad \dots\dots (2-62)$$

1.3 تمهيد preface

يعد تطوير وتوسيع الخدمات الصحية أحد الأركان الأساسية للتنمية الاقتصادية والاجتماعية ، والتي تستهدف تطويق الأوبئة والأمراض ، والاهتمام بتطوير الموارد البشرية ، ولاسيما في مرحلة التحولات والتغيرات الجذرية الواسعة التي يشهدها المجتمع ، وان المعالجة الطبية الفاعلة للكثير من الامراض من شانها أن ترفد المجتمع بعناصر منتجة تسهم في تطوير الناتج القومي، و الحركة التنموية في المجتمعات ، فضلاً عن الارتقاء بواقع الخدمات للأفراد، ومن هنا نرى ان حاجة الانسان الى استمرار الحياة بشكل افضل ، كانت الدافع و الاساس لبداية الدراسات و البحوث التي تتعلق بطرائق التنبؤ بأعداد المصابين بالاورام الخبيثة. سنتناول في هذا الفصل مبحثين . يتطرق الاول الى مقدمة وتعريف بالاورام الخبيثة وانواعها والفئات العمرية التي يصيبها المرض وبعض الاحصاءات العالمية ، كذلك مسببات المرض واعراضه واماكن الانتشار بالإضافة الى طرائق علاجه والوقاية منه.. اما الثاني في هذا الفصل سيتم تطبيق ما تم ذكره في الفصل الثاني (الجانب النظري) ، اذ ان الجانب التطبيقي يعد من اساسيات البحث العلمي وعن طريقه يتسنى للباحث والقارئ في الوقت نفسه بالتعرف على ابعاد الظاهرة المدروسة والوصول الى النتائج الحقيقية ومناقشتها ، انسجاماً مع هدف ومتطلبات البحث تم التطبيق على سلسلة زمنية للمدة من كانون الاول 2011 م الى كانون الثاني 2016 م لأعداد المصابين بالاورام الخبيثة وبواقع (72) مشاهدة مأخوذة من سجلات دائرة صحة بابل. طبقت الطرائق والاختبارات الاحصائية التي وردت في الجانب النظري على بيانات السلسلة الزمنية ، ومن اجل بيان وتوضيح اثر بعض العوامل المسلطة على معايير السلسلة .

تم تحليل البيانات وتقدير معلمات الانموذج المستعمل باستعمال البرامج الاتية :-

Excel (2007)

Minitab18.

E-views (8).

وبعد هذه التمهيد لا بد لنا من تناول بعض التعاريف المهمة والتي وضعت من قبل جهات رسمية مختصة لتوضيح ماهية هذا المرض الفتاك .

2.3. تعريف مرض السرطان [30] [31] Definition Cancer

هناك تعاريف متعددة ومتنوعه للسرطان نذكر ماهو مهم منها اذ اشارت منظمة الصحة العالمية (WHO) ضمن اوراق الصحة العامة الرقم (34) ، بان :

مرض السرطان هو مجموعة من الأمراض التي تتميز خلاياها بالعدائية Aggressive (هو نمو وانقسام الخلية غيرالمحدود)، وقدرة هذه الخلايا المنقسمة على غزو Invasion الأنسجة المجاورة وتدميرها، أو الانتقال إلى أنسجة بعيدة في عملية نطلق فيها اسم النقلية.

وهذه القدرات هي صفات الورم الخبيث على عكس الورم الحميد، والذي يتميز بنمو محدد وعدم القدرة على الغزو وليس لهُ القدرة على الانتقال أو النقلية. كما يمكن أن يتطور الورم الحميد إلى سرطان خبيث في بعض الأحيان.

وقد تم تعريف مرض السرطان من قبل (منظمة الصحة العالمية):

هو نمو الخلايا وانتشارها بشكل لا يمكن التحكّم فيه. وبإمكان هذا المرض إصابة كل أعضاء الجسم تقريباً. وغالباً ما تغزو الخلايا المتنامية النُسج التي تحيط بها ويمكنها أن تتسبّب في

نقائل تظهر في مواضع أخرى بعيدة عن الموضع المصاب. ويمكن توقي العديد من الافراد من مرض السرطان بتجنّب التعرّض لعوامل الاخطار الشائعة، مثل دخان التبغ. كما يمكن علاج نسبة كبيرة من السرطانات عن طريق الجراحة أو المعالجة الإشعاعية أو المعالجة الكيميائية، لاسيما إذا تم الكشف عنها في مراحل مبكّرة.

ويعرف السرطان بأنه هو عبارة عن نمو غير طبيعي لنسيج من أنسجة الجسم لذا فهو يصيب أنواعا مختلفة من الأعضاء وتختلف الأعراض عادة باختلاف العضو أو النسيج المصاب وهناك أعراض عامة قد تظهر مع وجود المرض مثل فقدان الوزن وفقدان الشهية والارتفاع في درجة الحرارة إلا أن هذه الأعراض بالطبع لا تعني وجود مرض السرطان فهي قد تصاحب أي مرض آخر وأحيانا حتى الحالة النفسية قد تتسبب في أعراض مشابهة لذا لا يمكن تشخيص المرض إلا بعد الفحص الطبي الشامل وإجراء الفحوصات الطبية والتأكد من الأنسجة بأخذ عينة أو خزعة من الأنسجة لفحصها تحت المجهر.

يستطيع السرطان أن يصيب كل المراحل العمرية عند الإنسان حتى الأجنة، ولكن تزيد مخاطر الإصابة به كلما تقدم الإنسان في العمر. ويسبب السرطان الوفاة بنسبة 13% من جميع حالات الوفاة. ويشير مجتمع السرطان الأمريكي ACS إلى موت 7.6 مليون شخص مريض بالسرطان في العالم في عام 2007. كما يصيب السرطان الإنسان فإن أشكال منه تصيب الحيوان والنبات على حد سواء.

في الأغلب، يعزى تحول الخلايا السليمة إلى خلايا سرطانية إلى حدوث تغييرات في المادة الجينية المورثة. وقد يكون سبب هذه التغيرات عوامل مسرطنة مثل التدخين، أو الأشعة أو مواد كيميائية أو أمراض مُعدية (كالإصابة بالفيروسات). وهناك أيضا عوامل مشجعة لحدوث السرطان مثل حدوث

خطأ عشوائي أو طفرة في نسخة الحمض النووي ال DNA عند انقسام الخلية، أو بسبب توريث هذا الخطأ أو الطفرة من الخلية الأم.

في الوقت الحالي يتم معالجة معظم أمراض السرطان وقد يتم الشفاء منها، وهذا يعتمد على نوع السرطان، وموقعه، ومرحلته. وعند اكتشاف السرطان، تبدأ معالجته بالجراحة Surgery أو بالعلاج الكيماوي Chemotherapy والإشعاعي Radiotherapy. بفضل التطورات البحثية، أمكن إنتاج أدوية قادرة على استهداف الخلايا السرطانية بتمييزها على المستوى الجزيئي، مما يقلل من احتمال استهداف الخلايا السليمة.

2.2.3 بعض الاحصاءات العالمية عن مرض السرطان [31] Global statistics

يعد السرطان من أهمّ أسباب الوفاة في أرجاء العالم كافة ، فقد تسبّب هذا المرض في وفاة 7.6 مليون نسمة على المستوى العالمي (نحو 13% من مجموع الوفيات) في عام 2008

وتحدث 30% من وفيات السرطان بسبب خمسة عوامل خطر سلوكية وغذائية رئيسة هي ارتفاع منسب كتلة الجسم، وعدم تناول الفواكه والخضر بشكل كاف، وقلّة النشاط البدني، وتعاطي التبغ، وتعاطي الكحول.

ويمثّل تعاطي التبغ أهم العوامل المرتبطة بمرض السرطان ، إذ يقف وراء 22% من وفيات السرطان العالمية و71% من الوفيات الناجمة عن سرطان الرئة.

وان العدوى التي تسبّب السرطان، مثل العدوى الناجمة عن فيروس التهاب الكبد B أو C وفيروس الورم الحليمي البشري، مسؤولة عن نحو 20% من وفيات السرطان التي تحدث في البلدان المنخفضة الدخل والبلدان المتوسطة الدخل.

ومن الملاحظ أنّ نحو 70% من مجمل وفيات السرطان التي سُجلت في عام 2008 حدثت في البلدان المنخفضة الدخل والبلدان المتوسطة الدخل.

كما ان من المتوقع أن يتواصل ارتفاع عدد الوفيات الناجمة عن السرطان على الصعيد العالمي، وأن يناهز 13.1 مليون وفاة في عام 2030 .

تزداد مخاطر الإصابة بهذا المرض لدى الأطفال الرضع وتقل مع تقدم العمر . ويعد مرض سرطان ابيضاض الدم من أكثر السرطانات الشائعة في الأطفال ثم تليها الليمفوما ثم سرطان الدماغ والجهاز العصبي.

كما تشير الإحصاءات إلى أن 70% من الاصابة بمرض السرطان تعود إلى عادات غير صحية مثل التدخين وعدم الاهتمام بمأكولات طازجة ومفيدة. وتبين الإحصائية التوزيع الاتي بين المصابين بمرض سرطاني :

35% منهم بسبب سوء التغذية.

30 % منهم بسبب التدخين.

15 % بسبب عيب وراثي عن الأب أو الأم.

3% بسبب شرب الكحول.

2 % بسبب التعرض للأشعة الفوق بنفسجية.

والباقي لأسباب مختلفة نادرة.

يتبين من تلك القائمة أن الإنسان يستطيع خفض احتمال تعرضه لمرض السرطان عن طريق : أولا، الامتناع عن التدخين، وثانيا تناول أطعمة مفيدة ينصح بها الأطباء تقي من هذا المرض.

3.3 بيانات السلسلة Observations of time series

تم الحصول على البيانات الخاصة بالبحث من دائرة صحة بابل– قسم التخطيط وتنمية الموارد – شعبة الاحصاء الصحي والحياتي والتي تمثل اعداد المصابين بالاورام الخبيثة وحسب الاشهر للمدة من كانون الثاني2011 م ولغاية كانون الاول 2016 م وبواقع (72) مشاهدة والمبينة في الجدول رقم (3–1).

جدول (3–1) لاعداد المصابين بالاورام الخبيثة الشهري في محافظة بابل للسنوات (2011–2016)

السنة	2011		2012		2013		2014		2015		2016	
الشهر	ذكر	انثى	ذكر	انثى	ذكر	انثى	ذكر	انثى	ذكر	انثى	ذكر	انثى
كانون 2	145	409	193	525	358	711	545	682	330	984	463	1114
شباط	171	461	216	573	365	754	466	955	393	921	532	1192
اذار	194	569	184	573	400	786	499	1161	477	1120	546	1216
نيسان	288	530	213	543	393	820	468	1026	436	1134	490	1149
ايار	168	483	187	573	431	799	510	1223	531	1311	431	1123
حزيران	158	496	310	706	466	870	469	1177	528	1256	462	993
تموز	190	511	270	743	409	878	452	1076	497	1186	461	1063
اب	207	491	238	632	493	1018	532	1254	478	1170	514	1193
ايلول	158	468	283	709	443	1054	495	1150	636	1293	413	1032
تشرين1	231	589	297	778	462	1056	410	995	346	1041	429	1199
تشرين2	176	437	259	732	450	797	275	988	312	1017	381	1367
كانون1	205	545	354	807	394	815	334	1039	456	1140	412	1356
المجموع	2291	5989	3004	7894	5064	10358	5455	12726	5420	13573	5534	13997

المصدر/ دائرة صحة بابل / قسم التخطيط / الاحصاء الصحي الحياتي .

3.4 التحليل الاحصائي Statistical analyze

تم تحليل بيانات السلاسل الزمنية التي تمثل اعداد المصابين في محافظة بابل باستعمال نماذج الاتجاه العام الخطية والانموذج التربيعي والانموذج غيرخطي non-linea باستعمال طريقة Gauss-Newton ، وإنموذج التمهيد الاسي ونماذج بوكس جنكينز لبيانات السلسلة من علاقتها بالزمن كمتغير مستقل يعبر عن اتجاه أعداد المصابين خلال المدة وكالاتي:

3.4. 1 إنموذج الانحدار الخطي البسيط (الاتجاه العام)

بين اجمالي أعداد المصابين بالاورام الخبيثة كمتغير معتمد y_i والزمن t_i كمتغير توضيحي مستقل يوثر في أعداد المصابين (بافتراض ان الاخطاء للإنموذج مستقلة ولا تتبع اي إنموذج احصائي يربط الاخطاء بعضها ببعض) ، اذ تم تقدير الإنموذج بالاعتماد على طريقة المربعات الصغرى مع استخراج الاختبارات الاحصائية للمعالم والإنموذج مع مقاييس الكفاءة للإنموذج وكانت النتائج كما يلي:-

$$y = 668.4 + 16.43t$$

t (15.432) (15.286) or P-value (.0000) (.0000)

S = 178.325 R-Sq = 79.0% R-Sq (adj) = 78.7%

وان مقاييس كفاءة الإنموذج والمتمثلة بمتوسط النسبة المطلقة للأخطاء ومتوسط الانحرافات المطلقة ومتوسط مربعات الاخطاء وفق الاتي :

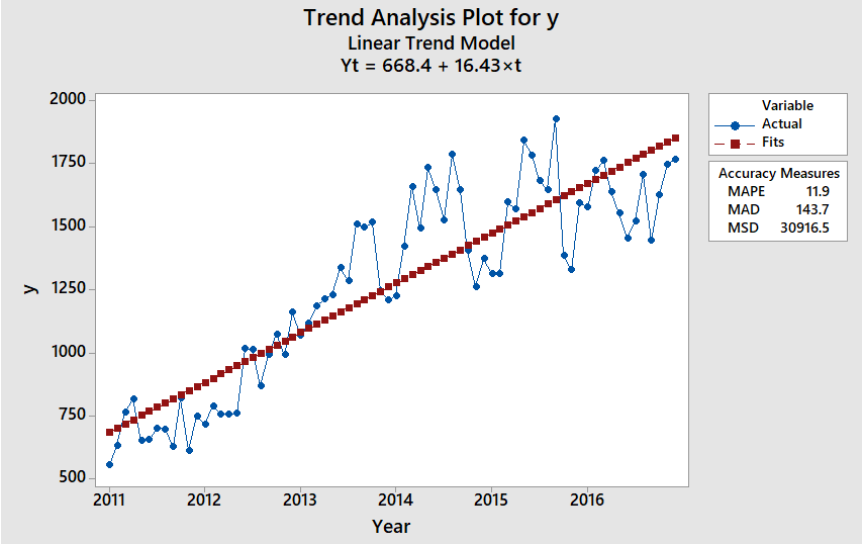
الجدول (3-2) مقاييس الدقة

المقياس	القيمة
MAPE	11.9
MAD	143.7
MSE	30916.5

جدول(3-3) جدول تحليل التباين للإنموذج الخطي البسيط.

Source	DF	SS	MS	F	P
Regression	1	8394495	8394495	263.98	0.000
Error	70	2225986	31800		
Total	71	10620482			

وتشير النتائج الى وجود تأثير ما بين المتغير التابع أعداد المصابين مع المتغير التوضيحي الزمن مع وجود علاقة التأثير الطردية وهذا يعني انه بمرور الزمن يرتفع اجمالي اعداد المصابين بالاورام الخبيثة ، كما ان قيم معامل التحديد تفيد بان القوة التفسيرية للإنموذج مرتفعة اذ بلغت قيمه معامل التحديد 79% وتدل ان المتغير المستقل الزمن قد فسر التغيرات الحاصلة في متغير أعداد المصابين بنسبة 79% ، كما ان الإنموذج المقدر ككل مقبول من الناحية الاحصائية وحسب قيمة اختبار F التي بلغت 263.98 واحتمالها ، p =0.000 للإنموذج المقدر والذي قيمته أقل من 0.05 مما يؤكد معنوية إنموذج الانحدار الخطي ككل ، ويفسر ذلك امكانية الاعتماد على نتائجه في تفسير أعداد المصابين بالاورام الخبيثة من علاقتها بالزمن ، بالإضافة الى ان معلمة المتغير المستقل (الزمن) موجبة ومعنوية وتبلغ قيمتها 16.43 وتعني ان زيادة الزمن بمقدار وحدة واحدة (شهر واحد) تودي الى زيادة اجمالي أعداد المصابين بالاورام الخبيثة بمقدار مساوى الى معلمة الإنموذج المقدرة وهي 16.43 ، والشكل (3-1) يوضح نتائج الإنموذج المقدر .



شكل (3–1) معادلة الاتجاه العام الخطية لأعداد المصابين

3.4. 2 الأنموذج التربيعي

تم بناء إنموذج تربيعي من الدرجة الثانية لوصف العلاقة ما بين المتغير المعتمد أعداد المصابين بالاورام الخبيثة والزمن وكانت نتائج الإنموذج كالاتي :

$$y = 444.8 + 34.56\, t - 0.2484\, t^2$$

t (7.852) (10.327) (−5.194) or P-value (.000) (.000) (.000)

S = 150.524 R-Sq. = 85.3% R-Sq. (adj) = 84.9%

وان مقاييس الكفاءة كانت كالاتي :

الجدول (3–4) مقاييس الدقة

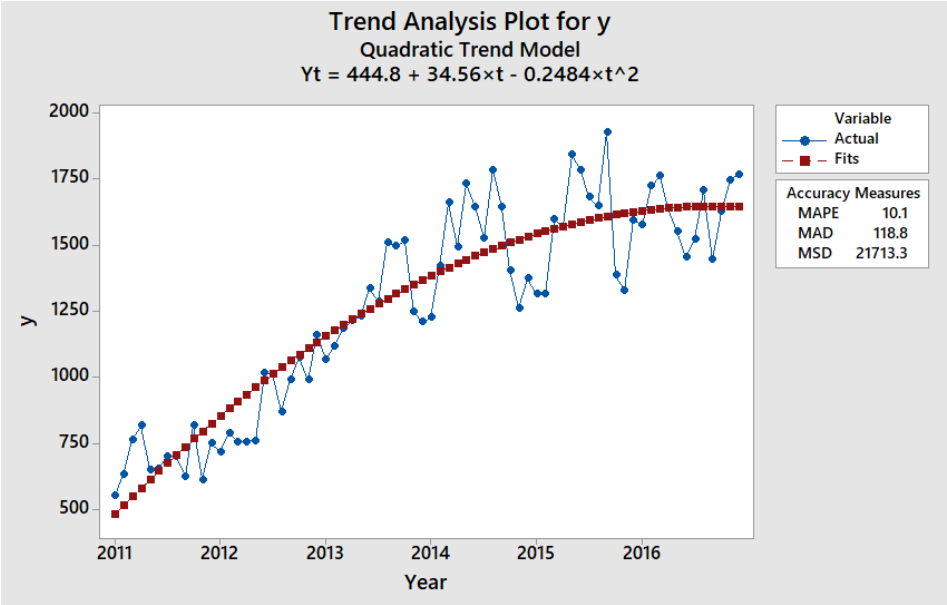
المقياس	القيمة
MAPE	10.1
MAD	118.8
MSD	1713.3

وهي اقل من نظيراتها في الإنموذج الخطي.

جدول(3-5) نتائج تحليل التباين للإنموذج التربيعي

Source	DF	SS	F	MS	P
Regression	2	9057126	4528563	199.87	0.000
Error	69	1563356	22657		
Total	71	10620482			

وتدل النتائج الى تأثير المتغير التابع أعداد المصابين مع المتغير التوضيحي الزمن مع وجود علاقة التأثير الطردية وهذا يعني ان زيادة الزمن تؤدي الى ارتفاع اجمالي أعداد المصابين بالاورام الخبيثة ، كما ان قيم معامل التحديد تفيد بان القوة التفسيرية للإنموذج مرتفعة اذ بلغت قيمه معامل التحديد 85.3% وتدل ان المتغير المستقل (الزمن) قد فسر التغيرات الحاصلة في متغير اعداد المصابين بنسبة 85.3% ، كما ان الإنموذج المقدر ككل مقبول من الناحية الاحصائية وحسب قيمة اختبار F التي بلغت 199.87 واحتمالها ، p-value =0.000 للإنموذج المقدر والذي قيمته أقل من 0.05 مما يؤكد معنوية إنموذج الانحدار الخطي ككل ، ويفسر ذلك امكانية الاعتماد على نتائجه في تفسير أعداد المصابين بالاورام الخبيثة من علاقتها ب (الزمن) ، بالإضافة الى ان معلمة المتغير المستقل (الزمن) موجبة ومعنوية وتبلغ قيمتها 34.56 وتعني ان زيادة الزمن بمقدار وحدة واحدة (شهر واحد) تودي الى زيادة اجمالي أعداد المصابين بالاورام الخبيثة بمقدار مساوى الى معلمة الإنموذج المقدرة وهي 34.56 اما المتغير الاخر والذي يمثل مربع (الزمن) فان تأثيره عكسي على أعداد المصابين بالاورام الخبيثة لان اشارته سالبة اذ بلغت المعلمة المقدرة 0.2484 الا ان هذا التأثير طفيف لانخفاض قيمة المعلمة المقدرة وان الإنموذج التربيعي بشكل عام افضل من الإنموذج الخطي البسيط ، والشكل (3-2) يوضح نتائج الإنموذج المقدر عن طريق مقارنة الإنموذج المقدر مع البيانات الاصلية للسلسلة الزمنية لأعداد المصابين بالاورام الخبيثة .



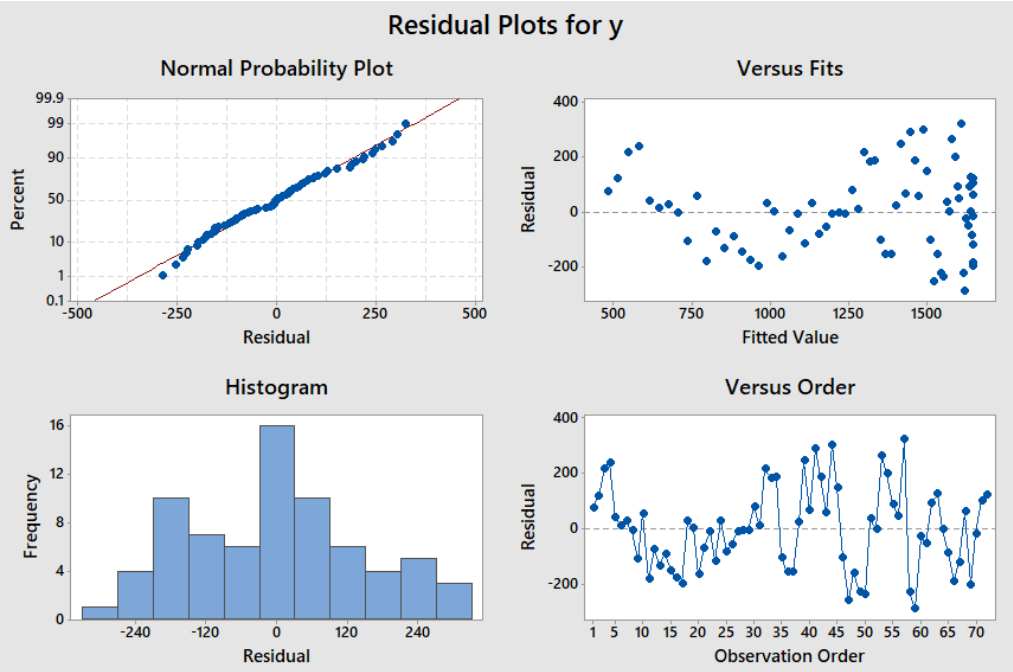
شكل رقم (3-2) معادلة الاتجاه العام التربيعية لأعداد المصابين

كذلك تم اجراء تحليل متتابع لتباين الإنموذج التربيعي وفقا الى مركباته الخطية والتربيعية وكانت النتائج وفق الاتي :

جدول(3–6) المركبة الخطية والتربيعية للبيانات

Source	DF	SS	F	P
Linear	1	8394495	263.98	0.000
Quadratic	1	662631	29.25	0.000

ومنها نجد معنوية تأثير المركبات الخطية والتربيعية للإنموذج المقدر، اذ بينت قيم (p_(value بان قيمها اقل من 0.05 مما يؤكد معنوية المركبات وجودة الإنموذج الموفق ، ولأجل التأكد من كفاءة الإنموذج تم التعامل مع اخطاء الانموذج ورسمها بعدة صيغ فعن طريق (Normal probability plot) . نجد اخطاء الانموذج (بواقي الانموذج) تكون حول خط الاتجاه ولا تبتعد عنه كثيرا مما يدل على عشوائيتها ومن ثم جودة الانموذج كما ان توزيع الاخطاء عن طريق المدرج التكراري (Histogram) يكون اقرب الى التوزيع الطبيعي .



شكل (3-3) التوزيع الاحتمالي الطبيعي للأخطاء العشوائية

ومن كل ذلك نستدل على عدم وجود إنموذج معنوي متبقي في اخطاء الإنموذج المقدر.

كذلك تم بناء إنموذج تكعيبي من الدرجة الثالثة لغرض زيادة دقة الإنموذج وكانت النتائج كـالاتي :

y = 543.4 + 18.90 t + 0.2845 t^2 - 0.004867 t^3

t (7.747) (1.832) (1.548) (-2.595) or P-value (.000) (.071) (.162) (.012)

S = 147.466 R-Sq. = 86.1% R-Sq.(adj) = 85.5%

الجدول (3–7) نتائج تحليل التباين للإنموذج التكعيبي.

Source	DF	SS	MS	F	P
Regression	3	9141738	3047246	140.13	0.000
Error	68	1478744	21746		
Total	71	10620482			

والانموذج بشكل عام مقبول من الناحية الاحصائية إلا انه بالمقارنة مع الإنموذج التربيعي فان الإنموذج المقدر لم يحسن النتائج كثيرا لاقتراب قيم معامل التحديد من بعضها ولأجل اختبار معنوية مركبات الإنموذج كانت النتائج كالاتي :

الجدول (3-8) نتائج تحليل مركبات الإنموذج الخطية والتربيعية والتكعيبية

Source	DF	SS	MS	F	P
Linear	1	8394495	8394495	263.98	0.000
Quadratic	1	662631	662631	29.25	0.000
Cubic	1	84612	84612	3.89	0.053

والتي تشير الى معنوية المركبة الخطية والتربيعية وعدم معنوية المركبة التكعيبية ، والشكل (3-4) يوضح النتائج :

شكل رقم (3-4) يوضح معادلة اتجاه العام الخطية التكعيبية

3.4.3 الإنموذج غير الخطي power(concave)

وقد تم استخدام طريقة (Gauss-Newton)لاجل تقدير الإنموذج وهي طريقة تكرارية بعدد (8) تكرارات وكان الإنموذج الموفق كالاتي :

$$y = 333.798 * t^{0.386035}$$

الجدول (3-9) مقدرات معالم الإنموذج غير الخطي **power(concave)**

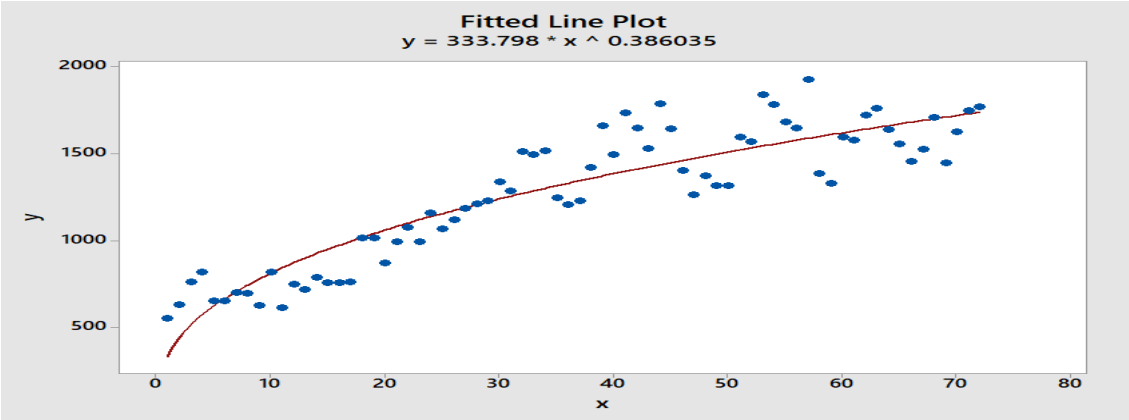
Parameter	Estimate	SE Estimate
Theta1	333.798	33.3147
Theta2	0.386	0.0265

وبالمقاييس الاحصائية التي تمثل مجموع مربعات الاخطاء لمرحلة التقدير النهائية (Final SSE) ومتوسط مربعات الاخطاء بقيمة بلغت (27596.6 MSE) كما في الجدول(3-10)الاتي:

الجدول (3-10) مجموع مربعات الاخطاء النهائية لمرحلة التقدير مع درجات الحرية

Final SSE	1931763
DFE	70
MSE	27596.6
S	166.122

والشكل (3-5) الاتي يوضح الإنموذج المقدر :



شكل رقم (3-5) الإنموذج غيرالخطي (power concave) .

3.4. 4 انموذج التمهيد الاسي

1–تم تطبيق إنموذج التمهيد الاسي المفرد وبمعاملات تمهيد مختلفة وكانت افضل النتائج من ناحية المقاييس الاحصائية للأخطاء عندما تكون قيمة معامل التمهيد ($\alpha = 0.658156$) ووفق الاتي :

Smoothing Constant

$\alpha = 0.650847$

الجدول (3–11) مقاييس الدقة

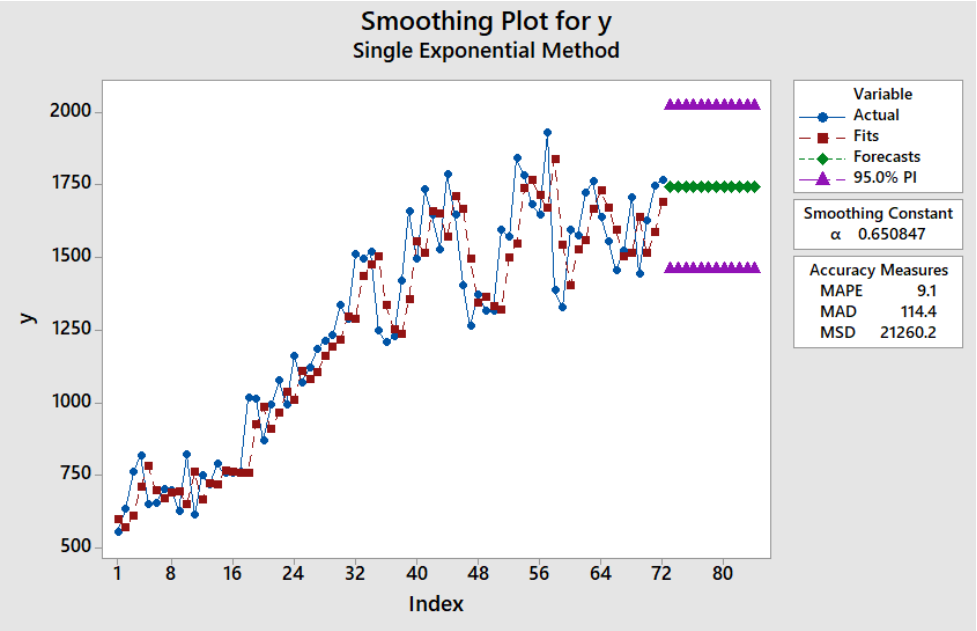
المقياس	القيمة
MAPE	9.1
MAD	114.4
MSD	21260.2

وبالقيم التنبؤية الاتية :

Period	Forecast	Lower	Upper
73	1741.51	1461.33	2021.68
74	1741.51	1461.33	2021.68
75	1741.51	1461.33	2021.68
76	1741.51	1461.33	2021.68
77	1741.51	1461.33	2021.68
78	1741.51	1461.33	2021.68
79	1741.51	1461.33	2021.68
80	1741.51	1461.33	2021.68
81	1741.51	1461.33	2021.68

82	1741.51	1461.33	2021.68
83	1741.51	1461.33	2021.68
84	1741.51	1461.33	2021.68

اذ ان الشكل (3-6) يبين العلاقة ما بين القيم الحقيقية لسلسلة اعداد المصابين بالاورام الخبيثة مع القيم التقديرية والتنبؤية مع بعض المقاييس الاحصائية الخاصة بكفاءة الإنموذج .



شكل (3-6) القيم التنبؤية و القيم الحقيقية لأعداد المصابين بالاورام الخبيثة

2- تم تقدير إنموذج التمهيد الاسي المزدوج وبمعاملات تمهيد مختلفة وكانت النتائج لأفضل القيم كالاتي :

Smoothing Constants

α (level) 0.945679

γ (trend) 0.032345

α = 0.650847

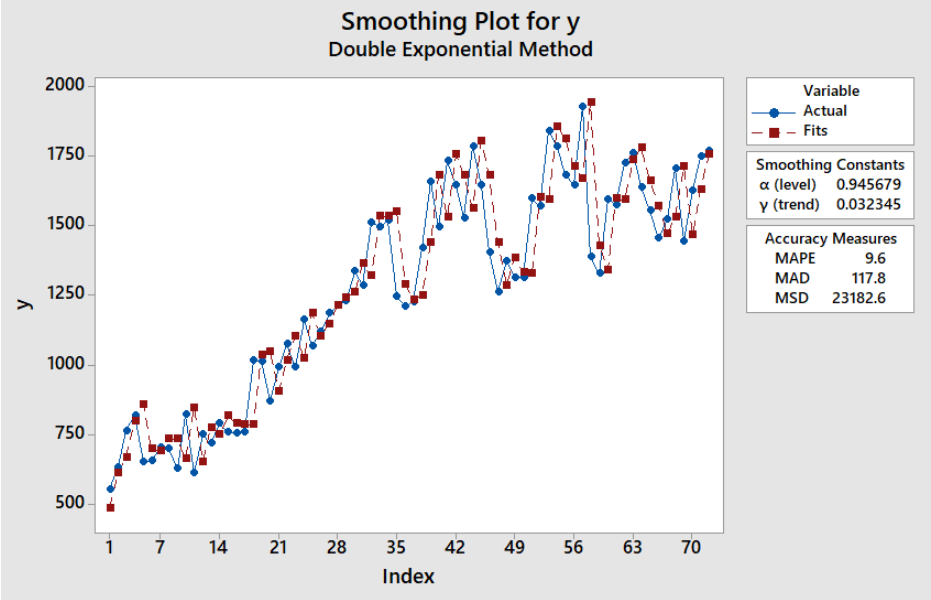
الجدول (3-12) مقاييس الدقة

المقياس	القيمة
MAPE	9.6
MAD	117.8
MSD	23182.6

وبالقيم التنبؤية الاتية :-

Period	Forecast	Lower	Upper
73	1784.00	1495.28	2072.72
74	1800.55	1375.12	2225.98
75	1817.10	1247.20	2387.00
76	1833.65	1116.19	2551.11
77	1850.20	983.67	2716.72
78	1866.75	850.30	2883.19
79	1883.30	716.42	3050.17
80	1899.85	582.19	3217.50
81	1916.39	447.72	3385.07
82	1932.94	313.08	3552.80
83	1949.49	178.32	3720.66
84	1966.04	43.46	3888.62

والشكل (3-7) يوضح نتائج عملية تقدير الإنموذج

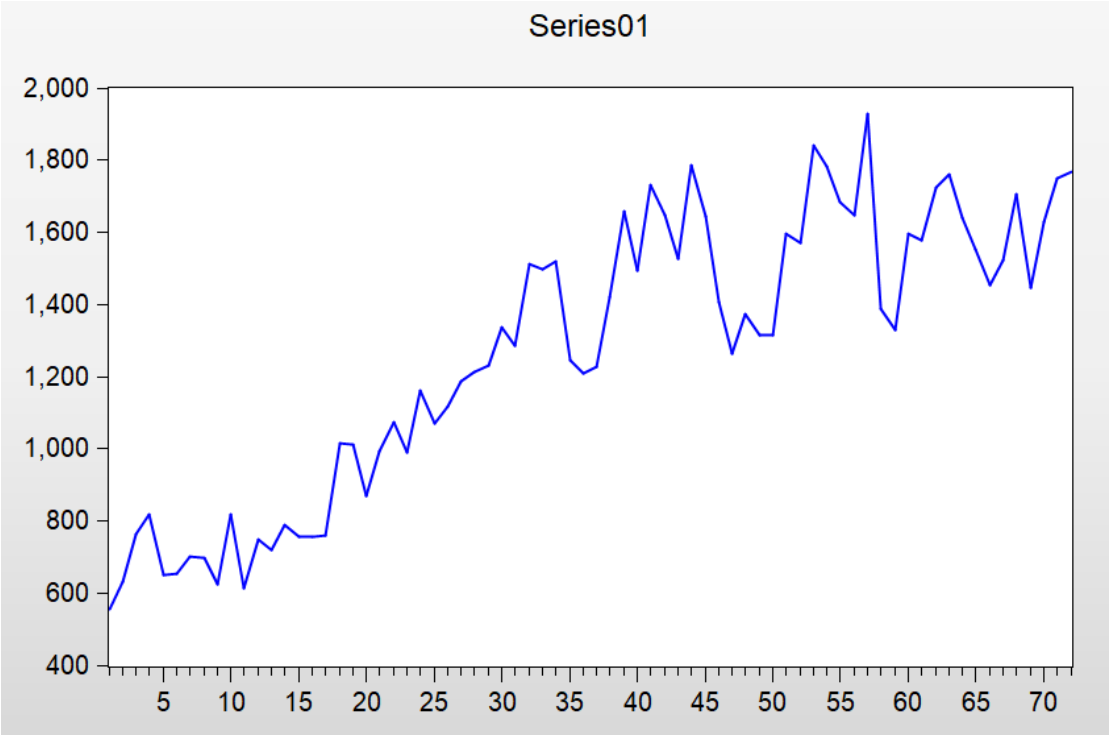


شكل رقم (3-7) عملية تقدير الإنموذج التمهيدي الاسي المزدوج

ان تأثير المستوى والمعبر عنه بقيمة المعلمة (α) بلغ 0.945679 (level) وهو تأثير عالي يشير الى اهمية القيمة السابقة ومساهمتها في إنموذج التنبؤ لارتفاع قيمتها واقترابها من الواحد الصحيح ، اما معلمة الاتجاه بلغت 0.032345 (trend) γ وهي قيمة قريبة من الصفر مما يعني ان تأثيرها ضعيف وفق الإنموذج المقدر ، وعند المقارنة ما بين نماذج التمهيد نجد ان إنموذج التمهيد الاسي المفرد افضل من إنموذج التمهيد الاسي المزدوج لانخفاض قيم المقاييس الاحصائية عند المقارنة فيما بينها.

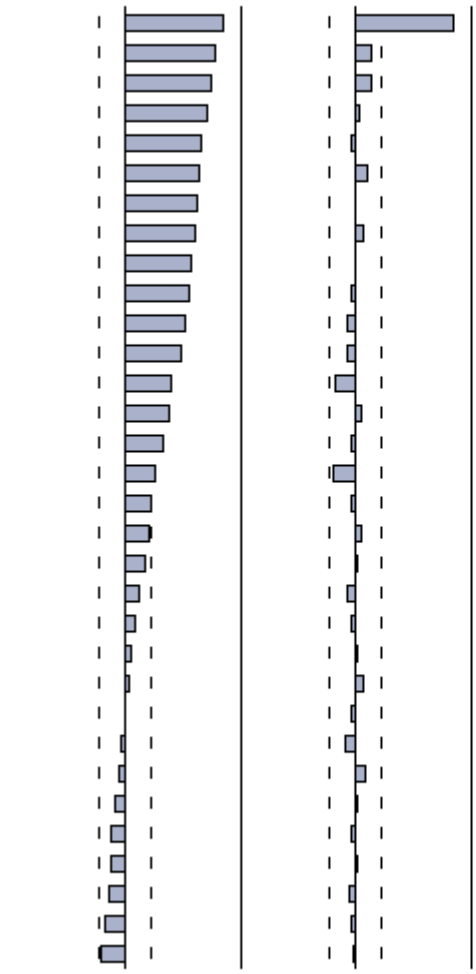
3.4. 5 استقرارية السلسلة الزمنية (stationary in time series)

ان الخطوة الاولى من تحليل السلاسل الزمنية يكون عن طريق رسمها للتعرف على بعض مكوناتها كالاتجاه العام و التغيرات الدورية والتغيرات الموسمية واخيرا العرضـية ، والشكل (3-8) يوضح رسم السلسلة الزمنية والذي يبين وجود الاتجاه العام بالزيادة الموجبة في السلسلة الزمنية كما ان البيانات لا تتذبذب عن متوسط ثابت ومن ذلك نستدل على عدم استقرارية في المتوسط والتباين اي عدم تحقق الاستقرارية الضعيفة الى السلسلة الزمنية لأعداد المرضى المصابين بالاورام الخبيثة ، مما يستوجب العلاج .



شكل (3-8) رسم السلسلة الزمنية الاصلية

ويتم التحقق من الاستقرارية للسلسلة الزمنية فيما اذا كانت استقرارية قوية او استقرارية ضعيفة ونعني بذلك استقرارية في المتوسط واستقرارية في التباين ، ولمزيد من الدقة تم حساب ورسم معاملات دالة الارتباط الذاتي (ACF) ومعاملات دالة الارتباط الذاتي الجزئي(PACF) للسلسلة الاصلية وكانت النتائج وفق الاتي :



	AC	PAC	Q-Stat	Prob
1	0.885	0.885	58.816	0.000
2	0.817	0.152	109.57	0.000
3	0.778	0.146	156.36	0.000
4	0.737	0.030	198.93	0.000
5	0.688	-0.029	236.54	0.000
6	0.668	0.111	272.56	0.000
7	0.641	0.005	306.26	0.000
8	0.625	0.077	338.83	0.000
9	0.605	0.005	369.81	0.000
10	0.577	-0.034	398.43	0.000
11	0.537	-0.070	423.59	0.000
12	0.493	-0.079	445.19	0.000
13	0.422	-0.184	461.29	0.000
14	0.389	0.057	475.16	0.000
15	0.351	-0.039	486.70	0.000
16	0.279	-0.197	494.08	0.000
17	0.226	-0.035	499.05	0.000
18	0.206	0.043	503.24	0.000
19	0.175	0.013	506.32	0.000
20	0.128	-0.078	508.00	0.000
21	0.091	-0.025	508.86	0.000
22	0.057	0.009	509.22	0.000
23	0.036	0.076	509.36	0.000
24	0.002	-0.034	509.36	0.000
25	-0.048	-0.092	509.62	0.000
26	-0.064	0.096	510.10	0.000
27	-0.083	0.023	510.92	0.000
28	-0.118	-0.038	512.62	0.000
29	-0.130	0.014	514.73	0.000
30	-0.150	-0.051	517.57	0.000
31	-0.185	-0.029	522.01	0.000
32	-0.208	-0.024	527.79	0.000

شكل (3-9) قيم ورسم معاملات الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) للسلسلة الزمنية لأعداد المصابين بالاورام الخبيثة في محافظة بابل

وان معاملات الارتباط الذاتي تتناقص اسيا من علاقتها بالإزاحة الزمنية ومعاملات الارتباط الذاتي الجزئي تنقطع بعد الازاحة الاولى وتدخل ضمن حدود القبول لقيم المعاملات ، وللكشف عن السلاسل غير المستقرة يكون عن طريق استخدام الاختبارات الاحصائية ومنها اختبار دكي – فولير الموسع المبني على إنموذج الانحدار الذاتي من الدرجة الاولى ، وقد يكون الاختبار للحد الثابت ، حد ثابت واتجاه عام ، او بدون حد ثابت ولا اتجاه عام واختبار اخر وهو اختبار فيلبس – بيرون وهو تطوير لاختبار دكي – فولير ، وقد تم حساب الاختبارين لمتغير الدراسة وقد وضحت النتائج في جدول (3–13) و (3–14) نتائج اختبار دكي – فولير الموسع لجذر الوحدة ولبيانات السلسلة الزمنية الاصلية وكذلك نتائج الاختبار بعد الفرق الاول.

جدول (3-13) الاحتمالية لنتائج اختبار دكي – فولير لجذر الوحدة ولبيانات السلسلة الزمنية الاصلية وكذلك نتائج الاختبار بعد الفرق الاول

الفرق الاول (1 ST DIFFERENCE)			المستوي (LEVEL)			
بدون حد ثابت ولا اتجاه عام	حد ثابت واتجاه عام	حد ثابت	بدون حد ثابت ولا اتجاه عام	حد ثابت واتجاه عام	حد ثابت	رموز المتغيرات
-10.7656	-10.7827	-10.8519	0.3327	-3.9775	-1.9779	t-statistic
0.0000	0.0000	0.0001	0.7788	0.0138	0.2958	p-value

جدول رقم (3-14) القيم الاحتمالية لنتائج اختبار فيليبس – بيرون لجذر الوحدة ولبيانات السلسلة الزمنية الاصلية وكذلك نتائج الاختبار بعد الفرق الاول .

بدون حد ثابت ولا اتجاه عام	حد ثابت واتجاه عام	حد ثابت	بدون حد ثابت ولا اتجاه عام	حد ثابت واتجاه عام	حد ثابت	رموز المتغيرات
-11.1556	-11.5735	-11.6374	0.8353	-3.9885	-1.9779	t-statistic
0.0000	0.0000	0.0001	08893	0.0134	0.2958	p-value

تشير نتائج اختبار جذر الوحدة في الجدول (3-13) و(3-14) ان بيانات السلسلة الزمنية الاصلية تحتوي على جذر الوحدة وهذا يعني عدم رفض فرض العدم الخاص بوجود جذر

الوحدة ولحالات الحد الثابت ، الحد الثابت واتجاه عام ، وبدون حد ثابت واتجاه عام لان اغلب قيم p-value اكبر من مستوى المعنوية 0.05 . ومنه نستدل ان السلسلة الزمنية الاصلية غير مستقرة ويجب اخذ الفروق لها لتحقيق الاستقرارية . وعند اعادة الاختبارات على بيانات السلسلة المحولة بعد اخذ الفرق الاول وفي كلا الاختبارين نجد ان قيم p-value اقل من 0.05 وهذا يعني رفض فرضية العدم وقبول الفرض البديل الذي يشير الى عدم وجود جذر الوحدة وبذلك تحقق الاستقرارية في السلسلة الزمنية وفي المتوسط .

وقد تم حساب معاملات دالة الارتباط الذاتي والارتباط الذاتي الجزئي للسلسلة الزمنية بعد حساب الفرق الاول لها اذ نجد ان كل المعاملات تقع ضمن حدود الثقة مما يدل على استقرارية السلسلة الزمنية بعد الفرق الاول.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.267	-0.267	5.2685	0.022
		2	-0.113	-0.198	6.2259	0.044
		3	0.092	0.004	6.8749	0.076
		4	-0.111	-0.115	7.8276	0.098
		5	-0.143	-0.219	9.4372	0.093
		6	0.077	-0.082	9.9120	0.128
		7	-0.112	-0.192	10.932	0.142
		8	-0.040	-0.179	11.063	0.198
		9	0.087	-0.106	11.693	0.231
		10	-0.014	-0.118	11.709	0.305
		11	0.115	0.038	12.844	0.304
		12	0.144	0.150	14.666	0.260
		13	-0.136	-0.019	16.310	0.233
		14	-0.017	-0.025	16.337	0.293
		15	0.049	0.032	16.561	0.346
		16	-0.058	0.065	16.880	0.393
		17	-0.114	-0.068	18.119	0.381
		18	0.025	-0.057	18.180	0.444
		19	0.010	0.012	18.189	0.510
		20	0.065	0.090	18.616	0.547
		21	-0.009	-0.014	18.625	0.609
		22	-0.033	-0.102	18.739	0.661
		23	0.147	0.083	21.060	0.577
		24	0.040	0.148	21.238	0.625
		25	-0.170	-0.052	24.484	0.492
		26	-0.002	-0.104	24.484	0.548
		27	0.033	-0.005	24.611	0.596
		28	-0.195	-0.129	29.196	0.403
		29	0.165	0.079	32.560	0.296
		30	0.035	0.003	32.715	0.335
		31	-0.013	0.036	32.736	0.382
		32	-0.011	-0.050	32.753	0.430

شكل (3-12) قيم ورسم معاملات الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) للسلسلة الزمنية لأعداد المصابين بالاورام الخبيثة في محافظة بابل بعد اخذ الفرق الاول

3-5 تشخيص رتبة الإنموذج والتقدير

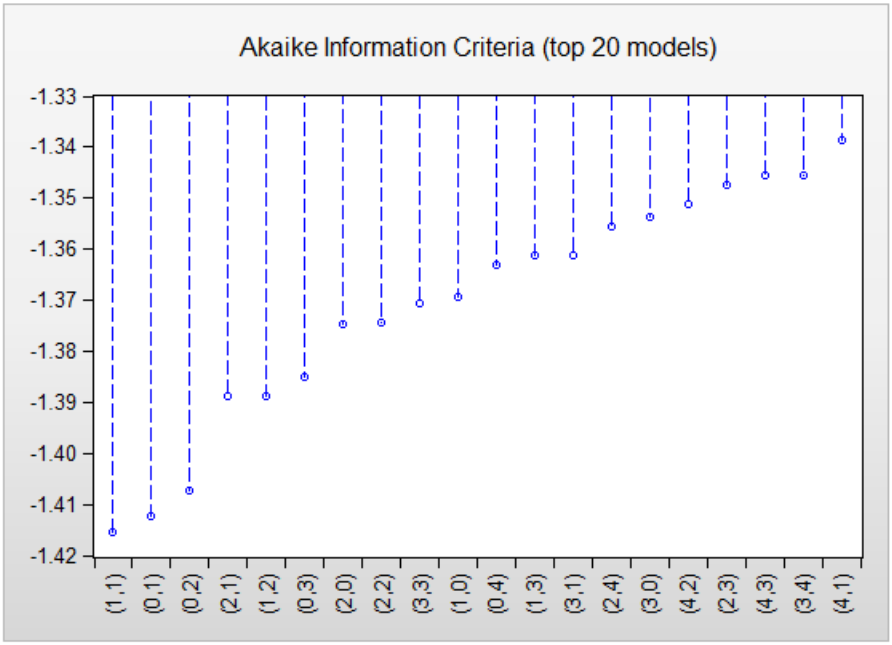
لأجل تشخيص الإنموذج بالاعتماد على اسلوب بوكس – جينكز يتم بعد تحقيق الاستقرارية في السلسلة الزمنية عن طريق النظر الى العلاقة ما بين معاملات الارتباط الذاتي والارتباط الذاتي الجزئي وكقاعدة عامة نظرية عندما تنحدر الارتباطات الذاتية بصورة اسية وتتقارب الى الصفر فان هذا يعني وجود إنموذج انحدار ذاتي (AR) تحدد درجته عن طريق عدد الارتباطات الذاتية الجزئية التي تختلف معنويا عن الصفر ، اما لو كان التناقص اسيا في الارتباطات الذاتية الجزئية فان الإنموذج من نوع (MA) وتحدد درجته عن طريق عدد الارتباطات الذاتية ذات الدلالة الاحصائية ، اما عندما تسلك الارتباطات الذاتية والارتباطات الذاتية الجزئية سلوكا اسيا في انحدارهما واقترابهما من الصفر فان الإنموذج من نوع الإنموذج المختلط (ARMA) ، ولكن في الجانب العملي غالبا لا تتحقق هذه المواصفات في البيانات المحللة لذلك نلجأ الى توفيق مجموعة من النماذج واختيار افضل إنموذج بالاعتماد على بعض المقاييس الاحصائية ومن هذه الاختبارات مقياس معلومة اكيكي (AIC) ومقياس حنان كوين (HQ) ومقيـاس شوارتز (BIC) ومقياس لوغاريتم الامكان الاعظم (LogL) واختيار اقل القيم وقد وصفت النتائج لحد الدرجة الرابعة لبعض النماذج المختارة كما في جدول (3-15).

جدول رقم (3-15) بعض نماذج بوكس جينكينز المقترحة لحد الدرجة الرابعة ولبيانات السلسلة الزمنية المستقرة (بعد الفرق الاول) .

Model Selection Criteria Table
Dependent Variable: DLOG(SERIES01)
Date: 06/01/18 Time: 22:27
Sample: 1 72
Included observations: 71

Model	LogL	AIC*	BIC	HQ
(1,1)	54.955085	-1.415419	-1.288938	-1.365066
(0,1)	53.840423	-1.412234	-1.317373	-1.374470
(0,2)	54.664657	-1.407352	-1.280870	-1.356999
(2,1)	55.000376	-1.388899	-1.230798	-1.325959
(1,2)	54.999149	-1.388865	-1.230763	-1.325924
(0,3)	54.859501	-1.384986	-1.226884	-1.322045
(2,0)	53.486675	-1.374630	-1.248148	-1.324277
(2,2)	55.473547	-1.374265	-1.184543	-1.298736
(3,3)	57.346432	-1.370734	-1.117771	-1.270029
(1,0)	52.298270	-1.369396	-1.274535	-1.331632
(0,4)	55.070587	-1.363072	-1.173350	-1.287543
(1,3)	55.003487	-1.361208	-1.171486	-1.285679
(3,1)	55.002474	-1.361180	-1.171458	-1.285651
(2,4)	56.803581	-1.355655	-1.102692	-1.254950
(3,0)	53.734249	-1.353729	-1.195627	-1.290788
(4,2)	56.640665	-1.351130	-1.098167	-1.250424
(2,3)	55.502788	-1.347300	-1.125957	-1.259183
(4,3)	57.438994	-1.345528	-1.060944	-1.232234
(3,4)	57.438000	-1.345500	-1.060917	-1.232207
(4,1)	55.191348	-1.338649	-1.117306	-1.250531
(4,4)	58.127459	-1.336874	-1.020670	-1.210992
(1,4)	55.117204	-1.336589	-1.115246	-1.248472
(3,2)	55.009157	-1.333588	-1.112245	-1.245471
(4,0)	53.814743	-1.328187	-1.138465	-1.252658
(0,0)	48.771732	-1.299215	-1.235974	-1.274038

وبالمقارنة ما بين النماذج بالرتب المختلفة حتى الرتبة الرابعة ،ووفقا الى مقياس اكيكي وعن طريق اختيار اقل القيم لأقل 20 (إنموذجاُ) نجد ان النتائج تمثلت بالجدول (1-15) الذي يوضح مدى الانخفاض في المقياس من تغير رتب النماذج المقارنة ومنه فان افضل إنموذج هو ARIMA (1, 1,1) والانموذج وفق هذه الحالة يمتاز بالتعقيد لكثرة عدد المعالم التي يحتويها.



شكل رقم (3-13) اإنموذج معلومات مقياس أكيكي لأفضل 20 إنموذج .

وبتقدير الإنموذج كانت معالمه وبعض المؤشرات الاحصائية وفق الجدول (3–16) الاتي:

جدول (3–16) المؤشرات الاحصائية

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.014469	0.005185	2.790722	0.0068
AR(1)	0.343745	0.196646	1.748045	0.0850
MA(1)	-0.761765	0.159266	-4.782972	0.0000
SIGMASQ	0.012384	0.002272	5.451194	0.0000

R-squared	0.164417	Mean dependent var	0.016344
Adjusted R-squared	0.127003	S.D. dependent var	0.122607
S.E. of regression	0.114557	Akaike info criterion	-1.435355
Sum squared resid	0.879260	Schwarz criterion	-1.307880
Log likelihood	54.95509	Hannan-Quinn criter.	-1.384662
F-statistic	4.394508	Durbin-Watson stat	1.993158
Prob(F-statistic)	0.006974		

Inverted AR Roots	.34
Inverted MA Roots	.76

ان الإنموذج المقدر معنوي لان قيمة اختبار F بلغت 4.3945 وبقيمة احتمالية للاختبار بلغت (0.006974) وذلك دليل على معنوية الإنموذج المقدر ، كما ان الإنموذج فسر التباين الكلي بنسبة 16.4%، اما المعالم المقدرة وحسب اختبار T وقيم الاحتمال Prob المرافقة لكل

معلمة نجد ان اغلب المعالم للإنموذج معنوية ومؤثرة في الإنموذج وهي ذات دلالة احصائية ، ويمكن وصف الإنموذج بالاتي :

$$Z_t = 0.014469 + 0.343745\, Z_{t-1} - 0.761765\, a_t + a_{t-1} + a_t$$

(0.05) (−4.78297) (1.748045) (2.790722) t

6-3 فحص مدى ملائمة الإنموذج

تم اختبار الإنموذج المقدر بالاعتماد على بعض الاختبارات الاحصائية لبواقي الإنموذج مثل احصاءة (Ljung-box) لاختبار مدى ملائمة الإنموذج المقدر وكانت النتائج عند بعض الأزاحات كما يلي :

جدول (3-17) اختبار كاي سكوير لملائمة الإنموذج

Lag	12	24	36	48
Chi-Square	11.0	17.7	27.2	44.1
DF	9	21	33	45
P-Value	0.279	0.668	0.751	0.510

واذ ان قيم P-Value اكبر من 0.05 عند اغلب الازاحات عندها يتم قبول فرضية العدم والتي تشير الى ان معاملات الارتباط الذاتي للاخطاء غير مرتبطة مع بعضها .

ثم تم اجراء اختبار سلسلة البواقي عن طريق حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي وكانت النتائج كما ياتي :

Correlogram of Residuals

Date: 06/01/18 Time: 23:59
Sample: 1 72
Included observations: 71
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.003	-0.003	0.0008	
		2	0.045	0.045	0.1498	
		3	0.041	0.042	0.2802	0.597
		4	-0.036	-0.038	0.3829	0.826
		5	-0.152	-0.157	2.2052	0.531
		6	0.071	0.072	2.6021	0.626
		7	-0.139	-0.125	4.1722	0.525
		8	0.016	0.023	4.1933	0.651
		9	0.081	0.078	4.7440	0.691
		10	0.072	0.066	5.1893	0.737
		11	0.168	0.182	7.6357	0.571
		12	0.185	0.142	10.641	0.386
		13	-0.139	-0.139	12.375	0.336
		14	0.061	0.045	12.715	0.390
		15	0.056	0.083	13.002	0.448
		16	-0.058	0.014	13.320	0.501
		17	-0.111	-0.104	14.509	0.487
		18	-0.010	-0.046	14.519	0.560
		19	-0.011	0.069	14.532	0.629
		20	0.127	0.099	16.175	0.580
		21	-0.002	-0.068	16.175	0.646
		22	0.037	-0.030	16.316	0.697
		23	0.090	0.061	17.191	0.699
		24	0.037	0.057	17.342	0.744
		25	-0.145	-0.143	19.698	0.660
		26	-0.023	-0.128	19.761	0.710
		27	-0.035	0.041	19.905	0.752
		28	-0.169	-0.078	23.351	0.613
		29	0.097	0.114	24.509	0.602
		30	0.037	-0.041	24.683	0.645
		31	-0.013	-0.049	24.706	0.693
		32	-0.046	-0.083	24.982	0.726

شكل (3–14) اختبار سلسلة البواقي عن طريق حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي

اذ ان اغلب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي تقع ضمن حدود الثقة مما يعني ان سلسلة البواقي عشوائية وان الإنموذج الموفق جيد ومعنوي من الناحية الاحصائية ، كما ان اغلب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي تقع ضمن حدود الثقة مما يعني ان سلسلة مربع البواقي عشوائية ، ووفق اختبار (Q-stat) اشار الى عدم المعنوية عند كل ازاحة وبذلك يتم قبول فرض العدم .

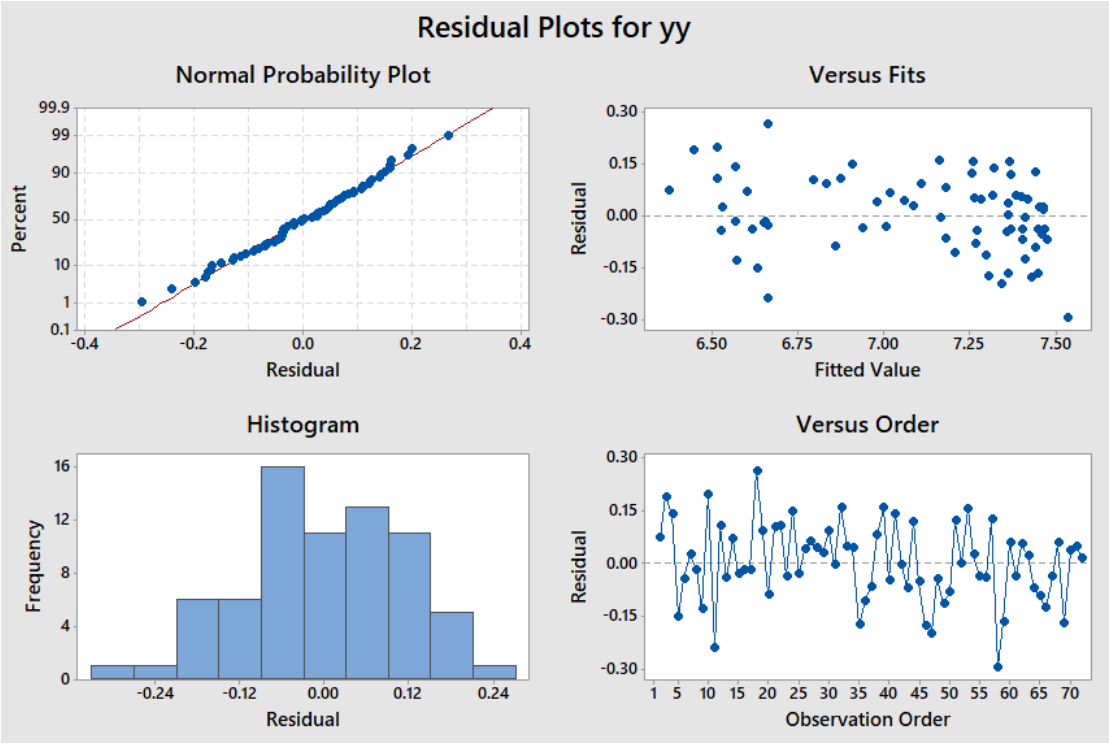
Correlogram of Residuals Squared

Date: 06/01/18 Time: 23:59
Sample: 1 72
Included observations: 71

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.124	0.124	1.1312	0.288
		2	-0.109	-0.126	2.0226	0.364
		3	-0.183	-0.157	4.5707	0.206
		4	-0.178	-0.156	7.0104	0.135
		5	-0.126	-0.137	8.2500	0.143
		6	0.023	-0.021	8.2940	0.217
		7	0.173	0.096	10.732	0.151
		8	0.128	0.045	12.089	0.147
		9	-0.057	-0.084	12.357	0.194
		10	-0.136	-0.090	13.928	0.176
		11	0.151	0.243	15.910	0.145
		12	0.057	0.048	16.190	0.183
		13	-0.093	-0.111	16.970	0.201
		14	0.014	0.040	16.989	0.257
		15	0.009	0.020	16.997	0.319
		16	-0.130	-0.105	18.580	0.291
		17	0.033	0.100	18.686	0.347
		18	-0.067	-0.176	19.126	0.384
		19	-0.005	-0.088	19.129	0.449
		20	-0.078	-0.092	19.742	0.474
		21	0.008	0.052	19.749	0.537
		22	-0.026	-0.160	19.819	0.594
		23	0.092	0.042	20.726	0.598
		24	-0.091	-0.129	21.642	0.601
		25	-0.045	-0.046	21.868	0.643
		26	0.008	-0.008	21.875	0.696
		27	-0.055	0.010	22.238	0.725
		28	0.042	-0.039	22.447	0.760
		29	0.055	0.040	22.824	0.785
		30	-0.057	-0.082	23.229	0.806
		31	-0.073	-0.032	23.920	0.814
		32	-0.084	-0.112	24.847	0.812

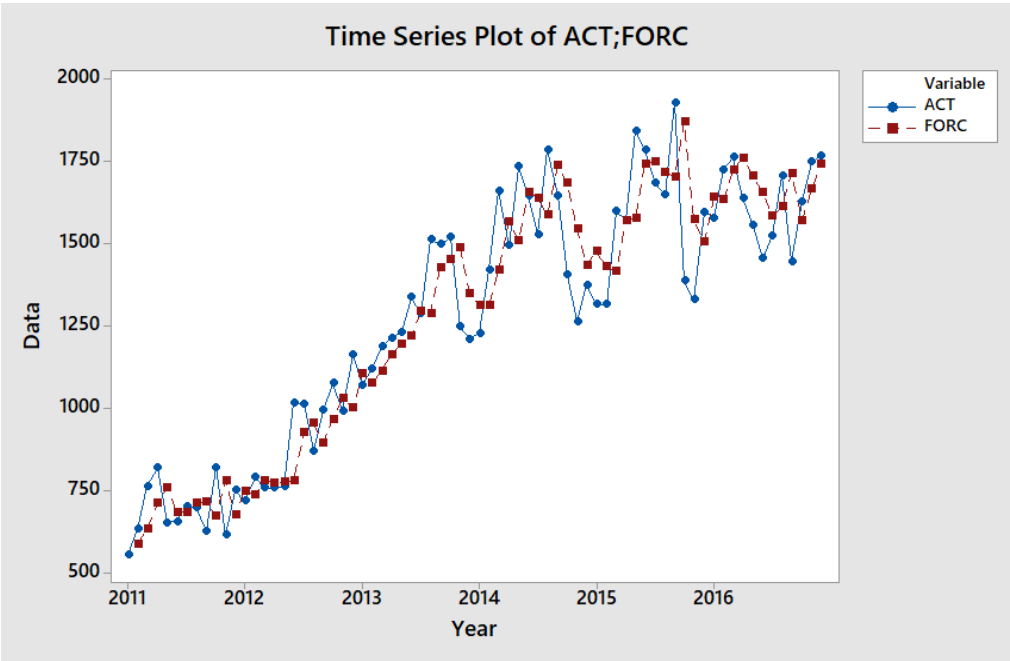
شكل (3–15) اختبار سلسلة البواقي عن طريق حساب معاملات الارتباط الذاتي والارتباط الذاتي الجزئي

كذلك تم اجراء مجموعة من الاختبارات ومنها ((Normal probability plot) اذ نجد ان اغلب البواقي تكون قريبة من الخط القطري الموفق وعدم وجود قيم متطرفة (قيم شاذة في الإنموذج) بعيدة عنه وهذا مؤشر على جودة الإنموذج كما تم رسم المدرج التكراري (Histogram) لاخطاء الإنموذج اقرب الى التوزيع الطبيعي مما يدل على عشوائيتها وهذا تأكيد على جودة الإنموذج ، والاشكال الاخرى تشير الى عدم وجود إنموذج متبقي في بواقي الإنموذج سواء للعلاقة ما بين الاخطاء مع القيم الموفقة وكذلك العلاقة ما بين الاخطاء وترتيب المشاهدة .



شكل رقم (3–12) التوزيع الاحتمالي الطبيعي للأخطاء العشوائية .

وبعد ذلك تم حساب القيم التنبوية ومقارنتها مع القيم الاصلية للبيانات ووفق الشكل (3–13) الذي يوضح ذلك



شكل رقم (3–13) يبين العلاقة ما بين القيم الاصلية لبيانات السلسلة الزمنية والمقدرة من الإنموذج

1.4 الاستنتاجات Conclusions

عن طريق ما تم التوصل اليه من نتائج في الجانب التطبيقي تم التوصل الى الاستنتاجات الاتية:

1. عند المقارنة بين نماذج التمهيد نجد ان انموذج التمهيد الاسي المفرد افضل من انموذج التمهيد الاسي المزدوج لانخفاض قيم المقاييس الاحصائية عند المقارنة في ما بينها . MAE=(114.4) و MSE=(21260.2) و MAPE=(9.1) .
2. على الرغم من ان طريقة التمهيد الاسي المفرد اعطت اقل قيم للمعايير الاحصائية من طرائق التنبؤ الاخرى لكن لا يمكن اعتبارها افضل الطرائق لان القيم المتنبأ بها عبارة عن القيمة (1741.51) نفسها لجميع الفترات الحقيقية اي ان القيم المتنبأ بها تكون على شكل مستقيم و لا تسلك سلوك القيم الحقيقية نهائياً .
3. ان طريقة انموذج الاتجاه العام الخطي اعطت اكبر قيم للمعايير الاحصائية من الطرائق الاخرى المعتمدة للتنبؤ و ذلك لوجود اتجاه عام واضح عند رسم السلسلة الاصلية للبيانات .
4. سلسلة اعداد المصابين بالأمراض الخبيثة في محافظة بابل للسنوات (2011 – 2016) تمثل سلسلة زمنية غير موسمية وغير مستقرة ، وذلك وفقاً لقيم معاملات الارتباط الذاتي ، ومن ثم تم اخذ الفرق الاول لمعالجة حالة عدم الاستقرارية في المتوسط .
5. ان المتغير مربع الزمن في الإنموذج التربيعي كان تأثيره عكسي على اعداد المصابين بالاورام الخبيثة لان اشارته سالبة اذ بلغت المعلمة المقدرة 0.2484 الا ان هذا التأثير طفيف لانخفاض قيمة المعلمة المقدرة وان الإنموذج التربيعي بشكل عام افضل من الإنموذج الخطي البسيط .
6. ان الانموذج ARIMA(1،1،1) هو المناسب الذي يمكن الاعتماد عليه في اجراء التنبؤات بأعداد المصابين بالأمراض الخبيثة في محافظة بابل لغرض وضع الخطط المستقبلية ، وقد اعطى تنبؤات جيدة وقريبة من القيم الفعلية لسلسلة أعداد المصابين ،

وهذا يتضح عن طريق القيمة الحقيقية للعام 2016 التي استعملت كسنة ضبط ووقوعها في فترة التنبؤ.

2.4 التوصيات Recommendations

بناءاً على تم التوصل اليه استنتاجات، ندرج بعض التوصيات وكالاتي :

1. يوصى باستخدام طرائق التقنيات الذكية كالشبكات العصبية والخوارزمية الجينية في التعامل مع سلسلة البيانات المستخدمة في البحث وامكانية الحصول على نتائج افضل ومقارنة النتائج مع طريقة بوكس– جينكنز وطرائق اخرى.
2. امكانية دراسة السلاسل الزمنية متعددة المتغيرات في الدراسات المستقبلية لحالات قد تتطلبها بحوث التنبؤ خاصة في المجالات الطبية ودراسة الامراض التي قد تتأثر بأكثر من متغير واحد مثل العمر والجنس وغيرها.
3. يوصي الباحث دائرة صحة بابل ، بالاهتمام بالجنب التخطيطي وعدم الاكتفاء بالجانب الفني لما لهذا الجانب من اهمية في اساليب مواجهة هذا المرض الفتاك و دفع عملية التنمية في القطاع الصحي.

اولاً: المصادر العربية :

1. الجبوري ، عبير حسن علي (2010) "التنبؤ بأسعار النفط العراقي لعام 2010"، مجلة جامعة بابل ، العلوم الإنسانية ، المجلد 18 العدد 1.
2. الزوبعي، عبيد محمود والتكريتي، محمد سمير (2013) "طريقة مقترحة لتحديد النماذج الكفوءة للسلاسل الزمنية"، مجلة جامعة دمشق للعلوم الاساسية، المجلد 29، العدد 1.
3. الصفاوي، صفاء يونس وغانم ، ايمان ، (2013) "مقارنة بين طرائق التنبؤ لسلسلة حمل الذروة اليومي للطاقة الكهربائية في مدينة الموصل"، بحث منشور في المجلة العراقية للعلوم الاحصائية، العدد 25.
4. الطائي ، فاضل عباس (2010) "التنبؤ والتمهيد للسلاسل الزمنية باستعمال التحويلات مع التطبيق "بحث منشور في مجلة العراقية للعلوم الاحصائية ، العدد 17.
5. العاني ، احمد حسين بتال (2005) "استخدام نماذج ARIMA في التنبؤ الاقتصادي" ، بحث منشور في مجلة كلية الادارة والاقتصاد / جامعة الانبار .
6. الوكيل ، خولة حسين (1981) " التنبؤ بالولادات في العراق باستعمال إنموذج الانحدار الذاتي " رسالة ماجستير إحصاء ، كلية الإدارة والاقتصاد / جامعة بغداد.
7. بري ، عدنان ماجد،(2002م) "طرائق التنبؤ الاحصائي" الجزء الاول ، كتاب ، جامعة الملك سعود / الرياض .
8. بن رشيد ، نياف الجابري' أ.د. كمال حسني بيومي، أ.د. إبراهيم بن عبدالله المحيسن ، (2004) " استشراف مستقبل التعليم بمنطقة المدينة المنورة:تطبيق السلاسل الزمنية" ، المدينة المنورة . بحث منشور في مجلة العلوم التربوية – الكويت / جامعة الكويت .
9. بن مرضي ، (2017) ، " طريقة بوكس جنكينز Box- Jenkins في تحليل السلاسل الزمنية والتنبؤ" ،جامعة نايف العربية للعلوم الامنية ، السعودية . DOI: 10.12816/0037211
10. جوجارت، 2015، " الاقتصاد القياسي " ، تعريب ومراجعة ا م د هند عبد الغفار عودة، دار المريخ للطباعة والنشر ، الجزء الثاني، مصر .
11. رديف ، نرجس هادي (2009) " أستخدام نماذج بوكس – جنكينز للتنبؤ بإنتاج سمنت العراق وبيان مدى كفايتها في ظل مشاريع الاعمار المستقبلية " ، بحث منشور ، جامعة بغداد ، كلية الادارة والاقتصاد ، مجلة العلوم الاقتصادية والادارية ، المجلد 16 ، العدد 58.

12. زين العابدين ، نورسل احمد (2013) "مقارنة إنموذج التدخل مع إنموذج ARIMA في السلاسل الزمنية " ، مجلة تكريت للعلوم الصرفة العدد 18.
13. شعراوي ، د. سمير مصطفى ، (2005م) ، "مقدمة في التحليل للسلاسل الزمنية "، كتاب ، جامعة الملك عبد العزيز / السعودية .
14. طعمة ، سعدية عبد الكريم (2012) " أستخدام السلاسل الزمنية للتنبؤ بأعداد المصابين بالأورام الخبيثة في محافظة الانبار" ، مجلة جامعة الانبار للعلوم الاقتصادية والادارية ، المجلد 4، العدد 8 .
15. عبد الحافظ ، علي سيف الدين (2006) " تحليل السلاسل الزمنية لبعض انواع مرض السرطان في العراق " رسالة ماجستير علوم في الاحصاء ، كلية الادارة و الاقتصاد / جامعة المستنصرية .
16. عبد الرسول ، محمود جواد (1981) " دراسة أحصائية تطبيقية للمقارنة بين النماذج الاسية ونماذج بوكس جنكينز في التوقعات المستقبلية مع تطبيق عملي " ، رسالة ماجستير أحصاء ، كلية الإدارة والاقتصاد/ جامعة بغداد .
17. عبيد، حميد، (2017)، " الاقتصاد القياسي" ، دار الكتب ، موزعون وناشرون، العراق، كربلاء .
18. فخري ، مصطفى علي (2015) "التنبؤ بقيم السلاسل الزمنية باستعمال إنموذج ARMAX مع تطبيق عملي على درجات الحرارة العظمى في مدينة بغداد" ، رسالة ماجستير ، كلية الادارة والاقتصاد / جامعة بغداد.
19. كنيهر ، عباس لفتة (2010) "بناء إنموذج سلاسل زمنية للتنبؤ بإعداد المرضى المراجعين لمستشفى الزهراء في محافظة واسط" بحث منشور في مجلة كلية الادارة والاقتصاد / جامعة واسط .
20. نادرة ، أيوب ، نظرية القرارات الادارية ، دار زهران للنشر ، الاردن ، 1997 ، ص:177
21. محمود ، أسيل سمير محمد ، 2005 ، " مقارنة بين طرائق تحليل وتنبؤ السلاسل الزمنية وتطبيقها على مبيعات الشركة العامة لتوزيع كهرباء بغداد" ، اطروحة دكتوراه ، كلية الإدارة والاقتصاد / جامعة بغداد .
22. .

ثانياً: المصادر الاجنبية :

23. Akaike, (1974), "A new look at the statical model Identification", IEEE, transaction on automatic control, vol. (AC-19), No. (6).

24. Anna Clara Monti (1994), "Aproposal for A Residual Autocorrelation test in Linear models", Biometrika, vol. (81), No. (4), pp. (776-780).

25. Box G, E.P & Jenkins, G. M., "Time series analysis forcasting and control sanfrancisco Helden-day", 1976.

26. Box, G. E. and Tiao, "Distribution of Residual Autocorrelation in Multiple Autoregressive, (1979),", JASA, Vol. (74), pp. (928-934).

27. Box, G.E.P., & Jenkins, G.M. (1970), "Time series analysis: forecasting and control", Holden-Day, San Francisco.

28. Lai T.L. & Lee C. P. 1997, "Information and prediction criteria for model selection in stochastic regression and ARMA models", statistic a sinica, vol. (7), pp. (285-309).

29. Ljung G. M., (1985), "Diagnostic testing of univariate time series model identification", Biometrica, vol. (72), No. (2), pp. (299-315).

30. J. M. G. WILSON & G. JUNGNER , (1968) , " Principles And Practice Of Screening For Disease".

31. https://ar.wikipedia.org/wiki/%D8%B3%D8%B1%D8%B7%D8%A7%D9%86 .

Abstract:

The huge development in the number of people infected with malignant tumors around the world and the causes of this phenomenon of human and material losses and the effects of social and psychological, including all groups of society the main motive and the main objective we have for the purpose of statistical analysis of this phenomenon and make forecast about the future.

In the theoretical side, some time series methods were used to predict the general linear trend pattern, the quadratic direction model, the non-linear model power (cocave) and the smoothing exponential model in addition to the Box-Jenkins models.

In the applied side, the methods mentioned in the theoretical side were applied to a sample of 72 people with malignant tumors for the period from 2011 to 2016 according to the months. The results of the statistical analysis found that the series of numbers of patients with malignant diseases in the province of Babylon for the years (2011 - 2016) represent a series of time, not seasonal, which has a general trend, that is an unstable time series, according to the values of the coefficients of self-correlation, the first difference was the most suitable for finding predictions close to the real values. The time square variable in the quadratic model had an adverse effect on the number of people with malignant tumors because its negative sign was estimated at 0.2484 but this effect was slight for the estimated parameter decrease and the growth The ARIMA(1،1،1) model is the appropriate model for predicting the numbers of patients with malignant diseases in Babil Province for the purpose of developing future plans. It has given good predictions and is close to the actual values of the number of casualties, which is illustrated by the real value of 2016, which was used as a control year and in the forecast period.

Republic of Iraq
Ministry of Higher Education
And Scientific Research
University of Karbala
Faculty of Management
And Economics
Department of Statistics

Use of different prediction methods To analyze the numbers of patients with malignant tumors

A thesis submitted to the council of the college of Administration &Economics\ University of Karbala as partial fulfillment of the requirements for the degree of Master of Statistics Sciences

By
Safaa Majeed Mutashar Al - Kalabi

Under supervision
Prof. Dr. Adnan Karim Najmaldin

2018 AD Holy Karbala 1440 AH