

التنبؤ بسرعة الرياح في ناحية بحركة - اربيل

باستخدام نماذج بوكس - جينكنز*

رفز محمد صالح طاهر

مدرس مساعد، المعهد التقني الإداري، اربيل

المستخلص

الهدف من البحث هو استخدام خطوات التحليل وتوضيحها باستخدام نماذج الانحدار الذاتي - المتوسط المتحرك التكاملية Auto Regressive Integrated Moving Average (ARIMA) وكذلك فحص النموذج المشخص لملاءمته في عمليات التنبؤ.

وتم استخدام السلسلة الزمنية (سرعة الرياح) أسبوعياً بمعدل ١٨٢ مشاهدة وللفترة من ٢٠٠٨/١/١ إلى ٢٠١١/٦/٢٨، وتحليلها باستخدام نماذج بوكس - جينكنز (Box- Jenkins). وبالاعتماد على النماذج النظرية ومقارنتها مع السلسلة الزمنية لسرعة الرياح، وقد تم اختيار نموذج (2,1,1) ARIMA ولكن كانت معلمات النموذج المعتمد نظرياً غير معنوية لذا تم الاعتماد على أربعة معايير هي: MSE, AIC, AICc, BIC لاختيار النموذج الأمثل.

وبعد الحصول على النتائج تبين أن النموذج الملائم هو (0,1,1) ARIMA إذ أمتلك هذا النموذج اقل قيمة للمعايير الأربعة المذكورة.

وللتأكد من مدى ملاءمة النموذج تم استخدام اختبارات مدى ملائمة النماذج وبعد الحصول على النتائج استخدم النموذج للتنبؤ بالقيم المستقبلية للسلسلة الزمنية مع ملاحظة الباحث أن التنبؤ باستخدام نموذج (0,1,1) ARIMA يعطي قيماً متشابهة للفترات الزمنية المتنبأ بها مع اختلاف فترات الثقة.

(*). تم استلام البحث في ٢٣/١٠/٢٠١١ وحصل على قبول النشر في ٢٠/٣/٢٠١٢.

Predicting Wind Speed in the Area of Baharka Using Box- Jenkins Models

Rafaz M. S. Tahir

Asst. Lect., Technical Administrative Institute, Erbil

Abstract

The aim of research is to use the steps of analysis and to clarify them through using Auto Regressive Integrated Moving Average (ARIMA). Also, it aims at examining the suitability of the model diagnosed in predicting.

Time series of wind speed is used weekly with an average of 182 observations, from 1/1/2008 to 28/6/2011. These observations are analyzed via using (Box-Jenkins) models. Depending on the theoretical models and comparing them to time series of wind speed, the model ARIMA (2,1,1) is chosen. However, the parameters of the adopted model were insignificant; therefore, MSE, AIC, AICc, BIC criteria are adopted in order to choose the best model.

The result show that ARIMA (0,1,1) model is the most suitable one for it has the least values of the aforementioned four criteria.

Suitability tests are used to ensure the suitability of ARIMA (0,1,1) model. Then, the latter is used to predict future time series. The researcher notes that this model gives similar values for the predicted periods with trust periods differences.

هدف البحث

يهدف البحث إلى التنبؤ بالسلسلة الزمنية لسرعة الرياح في مدينة أربيل وذلك باستخدام نماذج بوكس-جينكنز (Box- Jenkins) فضلاً عن فحص مدى ملاءمة النموذج المشخص ودقته في التنبؤ.

مشكلة البحث

عند استخدام نماذج بوكس-جينكنز (Box- Jenkins) واختيار النموذج الملائم من خلال الاعتماد على الفحوص التشخيصية نلاحظ أنه يمكن أن يكون هناك أكثر من نموذج ملائم لبيانات السلسلة الزمنية لذا فالمشكلة تكمن في كيفية اختيار النموذج الملائم للسلسلة الزمنية بعد فحص مدى ملاءمة النماذج (الفحوص التشخيصية) لاختياره واستخدامه في عمليات التنبؤ.

أهمية البحث

للرياح أهمية كبيرة في الحياة العملية لاسيما عملية توليد الطاقة الكهربائية باستخدام التوربينات الكهربائية (Martin, 2003) ومعرفة سقوط الأمطار. لذلك فالتنبؤ بالمتغير الموسوم مهم جداً لحكومة إقليم كردستان، فضلاً عن ذلك تكمن أهمية البحث في كيفية فحص النموذج المشخص عن طريق استخدام الفحوص التشخيصية.

١- الجانب النظري:

١-١: تعريف السلسلة الزمنية

هي عبارة عن مجموعة من المشاهدات لظاهرة معينة في فترات محددة، أو بمعنى آخر هي عبارة عن قيم أو مقادير هذه الظاهرة في سلسلة من التواريخ المتتابعة مثل الأشهر أو السنين وعادة تكون الفترات بين التواريخ متساوية (البري، ٢٠٠٢).

٢-١: دالة التغاير الذاتي Autocovariance Function

تعرف دالة التغاير الذاتي (البري، ٢٠٠٢) كما في الصيغة الآتية:

$$\begin{aligned} \gamma_{t,s} &= \text{COV}(y_t, y_s), \forall t, \forall s \\ &= E[(y_t - \mu)(y_s - \mu)] \end{aligned} \quad (1)$$

وإذا عرفنا k على أنه الفترة الزمنية التي تفصل بين y_t و y_s فإن دالة التغاير

الذاتي تكون:

$$\begin{aligned} \gamma_k &= \text{COV}(y_t, y_{t-k}), k = \pm 0, \pm 1, \pm 2, \dots \\ &= E[(y_t - \mu)(y_{t-k} - \mu)] \end{aligned} \quad (2)$$

٢-١: دالة الارتباط الذاتي Autocorrelation Function

تعرف دالة الارتباط الذاتي ACF (Box & Jenkins, 1970) كما في الصيغة

الآتية:

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = \pm 0, \pm 1, \pm 2, \dots \quad (3)$$

إذ إن:

γ_k : تمثل التغاير الذاتي للسلسلة الزمنية y_t بين الفترة t و $t-k$.

γ_0 : تمثل تباين السلسلة الزمنية y_t .

أما دالة الارتباط الذاتي للعينة Sample Autocorrelation Function

ويرمز لها SACF فتعرف بالصيغة الآتية:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (4)$$

٤-١: دالة الارتباط الذاتي الجزئي Partial Autocorrelation Function

تستخدم الارتباطات الذاتية الجزئية لقياس درجة الاقتران بين y_t و y_{t-k} عندما يتم تثبيت تأثير فترات التخلف الأخرى على المتغير y .

إن الغرض من استخدام هذه الارتباطات مع تحليل السلاسل الزمنية هو المساعدة في تحديد النموذج المناسب من نوع ARMA. في الحقيقة يتم تناول الارتباطات الذاتية الجزئية لهذا الغرض فقط. (الوردي، ١٩٨٢)

ويعبر عن معاملات الارتباط الذاتي الجزئي ϕ_{KK} (البري، ٢٠٠٢) على النحو

الآتي:

$$\phi_{kk} = \begin{cases} 1, & k = 0 \\ \rho_1, & k = 1 \\ \left| \begin{array}{ccccc} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & \rho_k \end{array} \right|, & k = 2, 3, \dots \\ \left| \begin{array}{ccccc} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & 1 \end{array} \right|, & k = 2, 3, \dots \end{cases}$$

أما تقدير معاملات الارتباط الذاتي الجزئي r_{kk} فيسمى بمعامل الارتباط الذاتي الجزئي للعيننة SPACF وتحسب تكرارياً من العلاقات الآتية:

$$\begin{aligned} r_{00} &= 1 \\ r_{11} &= r_1 \\ r_{kk} &= \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, \quad k = 2, 3, \dots \end{aligned} \quad (5)$$

إذ إن:

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-1}, \quad j = 1, 2, \dots, k-1$$

١-٥: تحليل السلاسل الزمنية

هناك عدد من الطرق المستخدمة في تحليل السلاسل الزمنية ومنها طريقة الانحدار الذاتي- المتوسط المتحرك المتكامل وهذه الطريقة تسمى بنماذج بوكس-جينكنز (Box - Jenkins) في تحليل السلاسل الزمنية.

١-٦: الاستقرار وعدم الاستقرار في السلاسل الزمنية

تفترض التطبيقات جميعها أن السلاسل الزمنية تتمتع بخاصية الاستقرار أو السكون Stationary، وأن الخطوة الأولى في تحليل نماذج بوكس- جينكنز (Box- Jenkins) هو التأكد من كون السلسلة الزمنية مستقرة. ويقصد بالاستقرار من الناحية الإحصائية أن يكون الوسط الحسابي وتباين السلسلة ثابتين.

ويمكن التعرف على كون السلسلة الزمنية مستقرة أو غير مستقرة من خلال الرسم البياني للظاهرة المدروسة أو من خلال مشاهدة دالة الارتباط الذاتي ACF ودالة الارتباط الذاتي الجزئي PACF إذ تقترب قيمها من الصفر بعد الإزاحة الثانية أو الثالثة.

وتكون السلسلة الزمنية مستقرة عند توفر الشروط الثلاثة الآتية:

(١) أن يكون الوسط الحسابي للسلسلة كمية ثابتة لا تعتمد على الزمن وكما مبين أدناه:

$$E(y_t) = \mu$$

(٢) أن يكون تباين السلسلة كمية ثابتة لا يعتمد على الزمن وكما مبين أدناه:

$$Var(y_t) = \sigma^2$$

(٣) أن يكون التباين المشترك للسلسلة كمية ثابتة لا يعتمد على الزمن، بل يعتمد على

الفرق بين الزمنين Lag Time وكما مبين أدناه:

$$Cov(y_t, y_{t-k}) = E(y_t - \mu)(y_{t-k} - \mu)$$

أما إذا كانت السلسلة الزمنية لها اتجاه عام فيجب إزالة هذا الاتجاه للحصول على سلسلة زمنية مستقرة إذ تستخدم طريقة الفروق لإزالة الاتجاه والحصول على سلسلة زمنية مستقرة، فإذا كان الفرق الأول لا يكفي فيؤخذ الفرق الثاني ثم الفرق الثالث حتى نحصل على سلسلة زمنية مستقرة، وغالباً ما نحصل على الاستقرار بعد أخذ الفرق الأول.

بافتراض أن y_t هي السلسلة الزمنية الأصلية وأن x_t تمثل السلسلة الزمنية المحسوبة بعد أخذ الفرق الأول إذ تكون معادلة أخذ الفرق الأول كما يأتي:

$$\begin{aligned} x_t = \nabla y_t &= y_t - y_{t-1} \\ &= y_t - B y_t \\ &= (1 - B) y_t \end{aligned}$$

إذ إن:

y_t تمثل بيانات السلسلة الزمنية غير المستقرة.
 x_t و y_t تمثل بيانات السلسلة الزمنية الجديدة.
 و B تمثل معامل الارتداد الخلفي (معامل التأخير).
 $\nabla = 1 - B$ و

أما الفروق من الرتبة الثانية فتكون كما يأتي:

$$\begin{aligned} x_t = \nabla^2 y_t &= \nabla y_t - \nabla y_{t-1} \Rightarrow y_t - y_{t-1} - y_{t-1} + y_{t-2} = y_t - 2y_{t-1} + y_{t-2} \\ &= y_t - 2By_t + B^2 y_t \Rightarrow (1 - 2B + B^2) y_t = (1 - B)^2 y_t \end{aligned}$$

وإذا كانت d تمثل عدد الفروق للسلسلة الزمنية الاصلية فإن معادلتها تكون كالاتي:

$$x_t = \nabla^d y_t = (1-B)^d y_t$$

أما إذا كانت السلسلة غير مستقرة حول التباين فإنه يتم تحويل البيانات إلى الصيغة اللوغارتمية أو يؤخذ لها الجذر التربيعي وهذا قد يؤدي إلى ثبات التباين (Wei, 1990).

٧-١: نماذج الانحدار الذاتي - المتوسط المتحرك المتكامل

Auto Regressive Integrated Moving Average (ARIMA) □

في كثير الأحيان تكون مشاهدات السلسلة الزمنية المتتالية غير مستقلة عبر الزمن، من أجل ذلك يجب استخدام أساليب إحصائية خاصة لدراسة مثل هذه السلاسل والتنبؤ بقيمتها في المستقبل.

ويعد أسلوب تحليل السلاسل الزمنية الذي قدمه العالمان Box and Jenkins (1976) الأدوات الإحصائية التي يمكن الاعتماد عليها في تحليل السلاسل الزمنية (الوردي، ١٩٨٢).

ويمكن تمثيل هذا الأسلوب بأربع مراحل:

١. مرحلة التعرف: في هذه المرحلة يتم تحديد النموذج المبدئي، وتعد كل من دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي أداتين رئيسيتين في تحديد النموذج ورتبته.
٢. مرحلة التقدير: في هذه المرحلة يتم تقدير معالم النموذج الذي تم اختياره في المرحلة الأولى.
٣. مرحلة التشخيص: في هذه المرحلة يتم دراسة كفاءة النموذج المبدئي من خلال الفحوص التشخيصية التي إذا اجتازها النموذج أصبح ملائماً والعكس صحيح.

٤. **مرحلة التنبؤ:** في هذه المرحلة يتم التنبؤ من خلال النموذج الملائم بالقيم الظاهرة في المستقبل، ولا يتم الانتقال إلى هذه المرحلة إلا بعد التأكد من أن النموذج المبدئي الذي تم اختياره في المرحلة الأولى وتم تقدير معلماته في المرحلة الثانية قد اجتاز كل الفحوصات والاختبارات التي أجريت عليه بكفاءة عالية.

إن نماذج ARIMA تتكون من مركبتين هما الانحدار الذاتي AR والوسط المتحرك MA فضلاً عن الفروق (difference) والتي يرمز لها بالرمز (Integrated) وهو رمز التكامل.

نفرض أن $y_t = 0, \mp 1, \mp 2, \dots$ سلسلة زمنية مستقرة فتكون صيغة نماذج الانحدار الذاتي والمتوسط المتحرك من الرتبة (p,q) (البري، ٢٠٠٢) وكما يأتي:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \dots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \dots - \theta_q y_{t-q} \quad (6)$$

$$\Rightarrow y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \phi_3 y_{t-3} - \dots - \phi_p y_{t-p} = a_t + \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \dots - \theta_q y_{t-q} \quad (6)$$

ويمكن إعادة كتابة المعادلة (٦) باستخدام مشغل الفروق وكما يأتي:

$$\phi(B) y_t = \theta(B) a_t$$

إذ إن:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \dots - \phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \dots - \theta_q B^q$$

$B^j = y_{t-j}$ Backshift Operator يمثل عامل الإزاحة الخلفي بحيث أن

a_t : هي عبارة عن سلسلة من المشاهدات العشوائية غير المترابطة وتسمى الضجة البيضاء White Noise ونفترض أنها تتبع توزيعاً طبيعياً بمتوسط يساوي

صفرًا وتباين يساوي σ^2 .

$\phi(B), \theta(B)$: تمثل كل منهما متعددًا محدوداً.

Autoregressive Parameters : تمثل معلمات نموذج الانحدار الذاتي $\phi_1, \phi_2, \phi_3, \dots, \phi_p$

Moving Average Parameters : تمثل معلمات نموذج المتوسطات المتحركة $\theta_1, \theta_2, \theta_3, \dots, \theta_p$

١-٨: تشخيص رتبة نموذج الانحدار الذاتي AR والمتوسط المتحرك MA

يتم تحديد رتبة كل من AR و MA من خلال رسم دالة الارتباط الذاتي ACF ودالة الارتباط الذاتي الجزئي PACF، فإذا تنازلت دالة الارتباط الذاتي بسرعة مع زيادة درجات التخلف (Lag k) فهذا يعني أن السلسلة الزمنية مستقرة وإذا لم تكن السلسلة الزمنية مستقرة في الوسط والتباين فعليه يتم استخدام أحد التحويلات والفروقات لتتحول السلسلة إلى سلسلة زمنية مستقرة (الوردي، ١٩٨٢).

يمكن الاعتماد على الجدول الآتي في تحديد رتبة كل من AR و MA (ARMA):
(Makridakis, 1998) و(البري، ٢٠٠٢).

جدول (١)

دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي لنماذج Box-Jenkins

ACF	PACF	النماذج
$\phi_{kk} = 0, k > 1$	تخامد آسي أو آسي متردد	AR(1)
$\phi_{kk} = 0, k > 2$	تخامد آسي أو تخامد جيبي	AR(2)
$\phi_{kk} = 0, k > p$	تخامد آسي أو تخامد جيبي	AR(p)
يسيطر عليها تخامد آسي	$\rho_k = 0, k > 1$	MA(1)
يسيطر عليها تخامد آسي أو جيبي	$\rho_k = 0, k > 2$	MA(2)
يسيطر عليها تخامد آسي أو جيبي	$\rho_k = 0, k > q$	MA(q)
تتناقص ويسيطر عليها تخامد آسي من التخلف ١ (k=1)	تتناقص وتتخامد آسيًا من التخلف ١ (k=1)	ARMA(1,1)
تتناقص بعد التخلف (p-q) ويسيطر عليها تخامد آسي أو جيبي بعد التخلف (p-q)	تتناقص بعد التخلف (q-p) وتتخامد آسيًا أو جيبيًا بعد التخلف (q-p)	ARMA(p,q)

المصدر: البري، عدنان ماجد (٢٠٠٢)، طرق التنبؤ الإحصائي، الجزء ١، كلية العلوم، جامعة الملك سعود، السعودية، ص ٨٤.

إذ إن:

معاملات الارتباطات الذاتية: $\rho_1, \rho_2, \dots, \rho_k$

معاملات الارتباطات الذاتية الجزئية: $\phi_{11}, \phi_{22}, \dots, \phi_{kk}$

٩-١: المفاضلة بين النماذج

يتم المفاضلة بين النماذج باستخدام المعايير الآتية:

١- متوسط مربع الأخطاء (Mean Square Error MSE):

ويمكن حسابه باستخدام الصيغة الآتية: (Box & Jenkins, 1970)

$$MSE = \frac{\sum_{t=1}^n a_t^2}{n} \quad (7)$$

٢- معيار (Akaike Information Criterion AIC):

ويمكن حسابه باستخدام الصيغة الآتية: (Akaike, 1973)

$$AIC = n \ln(MSE) + n(\ln 2\Pi + 1) + 2[(p+q) + 1] \quad (8)$$

٣- معيار (Akaike Information Criterion AICc) المعدل:

ويمكن حسابه باستخدام الصيغة الآتية: (Akaike, 1973)

$$AIC_c = n \ln(MSE) + n(\ln 2\Pi + 1) + \frac{2[(p+q) + 1]n}{n - (p+q) - 2} \quad (9)$$

٤- معيار (Bayesian Information Criterion BIC):

ويمكن حسابه باستخدام الصيغة الآتية: (Schwarz, 1978)

$$BIC = n \ln(MSE) + n(\ln 2\Pi + 1) + [(p+q) + 1] \ln(n) \quad (10)$$

إذ إن:

a_t : تمثل الخطأ أو البواقي.

n : تمثل عدد مشاهدات السلسلة الزمنية.

Π : كمية ثابتة وتساوي (٣,١٤١٥٩٢٦٥٤).

p : عدد المعلمات المقدرة في نموذج الانحدار الذاتي.

q : عدد المعلمات المقدرة في نموذج المتوسط المتحرك.

١-١٠: الفحوص التشخيصية Diagnostic Checking

اختبارات Portmanteau:

يستخدم اختبار Portmanteau للتحقق من مدى ملاءمة النموذج المشخص

وتكون صيغة الاختبار كما يأتي:

$$H_0: Corr_1(a) = Corr_2(a) = Corr_3(a) = \dots = Corr_k(a) = 0 \quad (*)$$

$$H_1: Corr_k(a) \neq 0 \quad \text{for some value of } k$$

إذ إن:

Corr: تمثل معاملات دوال الارتباطات الذاتية والجزئية.

ρ_k, ϕ_{kk} : وأن كلا الدالتين تتمتعان بخصائص متماثلة (Wei,1990).

وقد وجدت عدد من الاختبارات وكما يأتي:

A- أن معاملات دوال الارتباط الذاتي والجزئي تكون مستقلة فيما بينها وتتبع توزيعاً طبيعياً بمتوسط يساوي صفراً وتباين يساوي $1/n$ وكما يأتي:

$$Corr_k(\hat{a}) \sim N(0, \frac{1}{n}) \quad k=0,1,2,\dots,m \quad (11)$$

واعتماداً على هذه الخاصية فقد توصل إلى:

A-١: اقترح الباحثان Pierce & Box عام ١٩٧٠ (Box & Pierce, 1970) احصاءة لاختبار الفرضية السابقة (*) وبيان مدى كفاءة النموذج المشخص بالاعتماد على توزيع معاملات دالة الارتباط الذاتي في المعادلة ذات الرقم (١١) وكما مبين أدناه:

$$Q_{B-P} = n \sum_{k=1}^m \hat{\rho}_k^2(a) \sim \chi^2 [m-(p+q)] \quad (12)$$

إذ إن:

n: يمثل عدد مشاهدات السلسلة الزمنية.

k: يمثل أكبر إزاحة لـ m.

$\hat{\rho}_k(a)$: تقدير الارتباط الذاتي للبواقي وصيغته:

$$\hat{\rho}_k(a) = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2} \quad k=1,2,3,\dots,m$$

ونقارن قيمة Q_{B-P} مع قيمة χ^2 الجدولية بدرجة الحرية $[m-(p+q)]$ ومستوى دلالة (α):

فإذا كانت قيمة Q_{B-P} أكبر من القيمة الجدولية دل ذلك على رفض فرضية العدم المثبتة اعلاه (*) أي أن النموذج المشخص غير ملائم.

أما إذا كانت قيمة Q_{B-P} اصغر من القيمة الجدولية دل ذلك على عدم رفض فرضية العدم أي إن النموذج المشخص ملائم.

A-٢: اقترح الباحثان ناصر واحلام (ناصر واحلام، ٢٠٠٩) احصاءة لاختبار الفرضية السابقة وبيان مدى كفاءة النموذج المشخص بالاعتماد على توزيع معاملات دالة الارتباط الذاتي الجزئي في المعادلة ذات الرقم (١١) وكما مبين أدناه:

$$\hat{\theta}_{kk}(a) \sim N(0, 1/\sqrt{n})$$

$$\frac{\hat{\theta}_{kk}(a) - 0}{1/\sqrt{n}} \sim N(0, 1)$$

بعد تعويض $k=1$ نحصل على:

$$\frac{\hat{\theta}_{11}(a) - 0}{1/\sqrt{n}} \sim N(0, 1)$$

وبالتربيع نحصل على:

$$Z_1 = Z_{\hat{\phi}_{11}(a)} = n \hat{\theta}_{11}^2(a) \sim \chi^2(1)$$

وبأخذ المجموع للطرفين نحصل على:

$$\sum_{k=1}^m Z_k^2 = n \sum_{k=1}^m \hat{\theta}_{kk}^2(a) \sim \chi^2(m - (p + q))$$

لهذا فإن الصيغة النهائية تكون:

$$AQ_1 = n \sum_{k=1}^m \hat{\theta}_{kk}^2(a) \sim \chi^2[m - (p + q)] \quad (13) \quad \square \quad \square \quad \square$$

ونقارن قيمة AQ_1 مع قيمة χ^2 الجدولية وكما ذكرنا سابقاً. □

□

B- نذكر أن (Davies & Newbold, 1979) قد أشارا إلى أن:

$$\text{var}(\hat{corr}_k) = (n-k)/[n(n+2)]$$

وأن:

$$\text{Cov}(\hat{corr}_k, \hat{corr}_j) = 0 \quad ; \quad \forall (k \neq j)$$

أي إن:

$$\hat{corr}_k(a) \sim N(0, (n-k)/[n(n+2)]) \quad (14) \quad \square$$

واعتماداً على هذه الخاصية فقد تم صياغة إحصاءات اختبار العشوائية لكل من معاملات الارتباطات الذاتية ومعاملات الارتباطات الذاتية الجزئية وكما يأتي:

B-1: توصل الباحثان Box & Ljung (1978) Box & Ljung إلى اختبار أكثر كفاءة وتطوراً في اختبار الفرضية السابقة (*) وبالاعتماد على توزيع معاملات الارتباط الذاتي للبواقي في المعادلة ذات الرقم (١٤) على وفق إحصائية Q المعدلة Modified Q Statistic الآتية:

$$Q_{L-B} = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{\rho}_k^2 \sim \chi^2 [m - (p+q)] \quad (15)$$

نقارن القيمة المحتسبة لـ Q_{L-B} مع قيمة χ^2 الجدولية بدرجة الحرية $[m - (p+q)]$ ومستوى المعنوية (α) إذ نرفض فرضية العدم المثبتة السابقة (*). إذا كانت القيمة المحتسبة أكبر من القيمة الجدولية، ولا نستطيع أن نرفض فرضية العدم إذا كانت القيمة المحتسبة أقل من القيمة الجدولية بمعنى آخر عشوائية وأن النموذج المشخص ملائم.

B-٢: اقترحت الباحثة Monti (1994) اختبار لمدى ملاءمة النموذج المشخص بالاعتماد على معاملات الارتباطات الذاتية الجزئية على وفق الصيغة الآتية:

$$Q_{mon.} = n(n+2) \sum_{k=1}^m (n-k)^{-1} \hat{\theta}_{kk}^2 \sim \chi^2 [m - (p+q)] \quad (16)$$

وبمقارنة قيمة $Q_{mon.}$ المحسوبة مع قيمة χ^2 الجدولية بدرجة حرية $[m - (p+q)]$ ومستوى معنوية (α) معين يتم التوصل إلى قرار وبالأسلوب الذي ذكر سابقاً.

C- اقترح الباحثان ناصر وأحلام عام ٢٠٠٩ (ناصر وأحلام، ٢٠٠٩) بتطوير احصاءة اختبار حدي الثقة Confidence Interval Test أو احصاءة بارتلت Bartlett Test وذلك آخذين الإشارة الخاصة بمعاملات الدوال بنظر الاعتبار لاختبار عشوائية الأخطاء (البواقى).

وبما أن دوال كل من الارتباطات الذاتية والارتباطات الذاتية الجزئية تتوزعان توزيعاً طبيعياً ومستقلاً تقريباً Approximately Independent Normally Distribution أي إن:

$$\text{Corr}_k(\hat{a}) \sim N(0, 1/n) \quad k=0,1,2,\dots,m$$

وبأخذ المجموع نحصل على:

$$\sum_{k=1}^m \hat{corr}_k(\hat{a}) \sim N(0, m/n) \quad k=0,1,2,\dots,m$$

وإن صيغة الاقتراح تكون:

$$\Pr\left\{-1.96\left(\frac{m}{n}\right)^{\frac{1}{2}} \leq \sum_{k=1}^m \hat{corr}_k(\hat{a}) \leq 1.96\left(\frac{m}{n}\right)^{\frac{1}{2}}\right\} = 0.95$$

وبالتفصيل فإن:

$$A_1 = \Pr\left\{-1.96\left(\frac{m}{n}\right)^{\frac{1}{2}} \leq \sum_{k=1}^m \hat{\rho}_k(\hat{a}) \leq 1.96\left(\frac{m}{n}\right)^{\frac{1}{2}}\right\} = 0.95 \quad (17)$$

$$A_1 = \Pr\left\{-1.96\left(\frac{m}{n}\right)^{\frac{1}{2}} \leq \sum_{k=1}^m \hat{\theta}_k(\hat{a}) \leq 1.96\left(\frac{m}{n}\right)^{\frac{1}{2}}\right\} = 0.95 \quad (18)$$

يكون النموذج المشخص ملائماً عندما تتحقق كل من الصيغتين السابقتين الخاصتين بمجموع حدي الثقة تحت مستوى الثقة ٩٥٪ كما يتم اختبارها أيضاً عند مستوى الثقة ٩٩٪ أو ٩٠٪.

٢: الجانب العملي

١-٢: حالة الدراسة

تم جمع البيانات عن سرعة الرياح من المديرية العامة للزراعة - مركز الأنواء الجوية في محافظة أربيل واعتمدت ناحية (بحركة) نظراً لتوفر البيانات عن سرعة الرياح (نشرات منشورة، ٢٠٠٨-٢٠١١)، وتتألف عدد المشاهدات من ١٨٢ مشاهدة مأخوذة أسبوعياً كمعدل من تاريخ ٢٠٠٨/١/١ إلى ٢٠١١/٦/٢٨، وتم تحليل البيانات باستخدام برنامج Statigraphics وأدناه نتائج التحليل:

٢-٢: رسم السلسلة

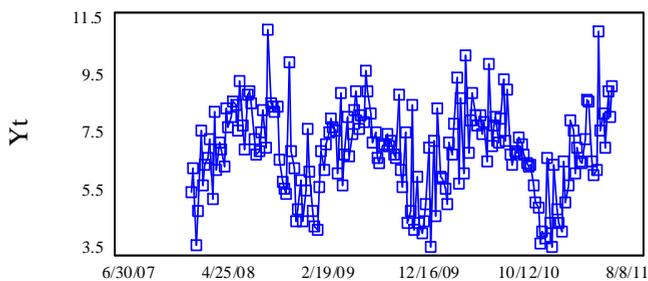
بالنظر إلى السلسلة الزمنية في الشكل ذي الرقم (١) نلاحظ وجود تذبذبات متزايدة ومتناقصة ويدل على اختلاف سرعة الرياح بين الأسابيع مع وجود اتجاه متزايد ومتناقص وأن أقل سرعة كانت (٣,٥٨٢٧١) وأعلى سرعة كانت (١١,١٥٧١) متر/ ثانية وكما مبين أدناه:

جدول (٢)

المعدلات الأسبوعية لسرعة الرياح في ناحية بحركة- أربيل

Week	Yt	Week	Yt	Week	Yt	Week	Yt	Week	Yt	Week	Yt
1	5.491	32	8.329	63	7.649	94	4.393	125	8.144	156	3.583
2	6.330	33	7.000	64	6.127	95	4.817	126	7.487	157	6.456
3	3.640	34	11.157	65	8.957	96	8.484	127	7.917	158	4.903
4	4.826	35	8.557	66	5.739	97	4.187	128	6.574	159	4.427
5	7.639	36	8.457	67	6.780	98	5.986	129	9.966	160	4.103
6	5.696	37	8.300	68	8.083	99	4.420	130	7.107	161	6.556
7	6.441	38	8.429	69	6.734	100	4.063	131	7.787	162	5.149
8	6.647	39	6.600	70	7.516	101	4.471	132	8.120	163	5.706
9	7.340	40	5.814	71	8.339	102	5.053	133	7.213	164	8.003
10	5.244	41	5.571	72	9.009	103	7.013	134	8.011	165	7.606
11	8.263	42	5.443	73	7.671	104	3.600	135	9.381	166	6.129
12	6.274	43	10.014	74	8.183	105	7.246	136	7.233	167	7.027
13	7.230	44	6.879	75	7.909	106	4.639	137	9.033	168	6.567
14	6.964	45	6.319	76	9.703	107	8.386	138	6.789	169	6.490
15	6.373	46	4.463	77	9.003	108	6.009	139	6.423	170	7.333
16	8.379	47	4.886	78	8.234	109	5.939	140	6.917	171	8.700
17	7.744	48	5.900	79	7.194	110	5.586	141	6.819	172	8.614
18	8.110	49	4.457	80	7.581	111	5.063	142	7.374	173	6.543
19	8.603	50	5.543	81	6.664	112	7.201	143	7.127	174	6.057
20	8.487	51	7.686	82	6.466	113	6.766	144	6.777	175	6.279
21	7.613	52	6.214	83	7.074	114	7.874	145	6.481	176	11.071
22	9.327	53	4.843	84	7.140	115	9.434	146	6.346	177	7.597
23	7.807	54	4.300	85	7.497	116	5.791	147	6.413	178	8.200
24	6.973	55	4.143	86	7.096	117	8.731	148	5.740	179	7.007
25	8.850	56	5.643	87	7.246	118	6.156	149	5.106	180	9.000
26	8.994	57	6.886	88	6.814	119	10.226	150	4.937	181	8.086
27	8.543	58	6.257	89	6.647	120	6.853	151	3.682	182	9.193
28	7.300	59	7.157	90	8.866	121	7.977	152	4.099		
29	6.757	60	7.543	91	6.269	122	8.930	153	3.895		
30	6.929	61	8.039	92	5.640	123	8.159	154	6.681		
31	7.543	62	7.733	93	7.583	124	7.796	155	4.409		

المصدر: جمهورية العراق، اقليم كردستان، المديرية العامة لزراعة أربيل، مركز الانواء الجوية الزراعية، نشرات شهرية منشورة (٢٠٠٨-٢٠١١).



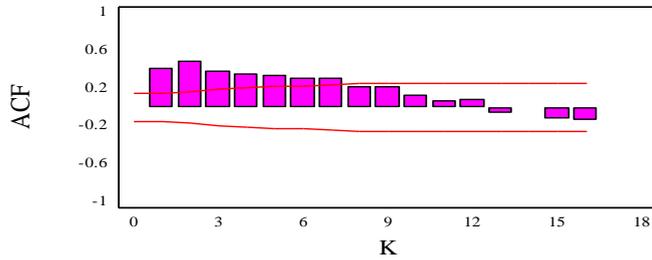
المصدر: كافة الاشكال من إعداد الباحث بالاعتماد على برنامج Statighapics

شكل (١)

السلسلة الزمنية الاصلية

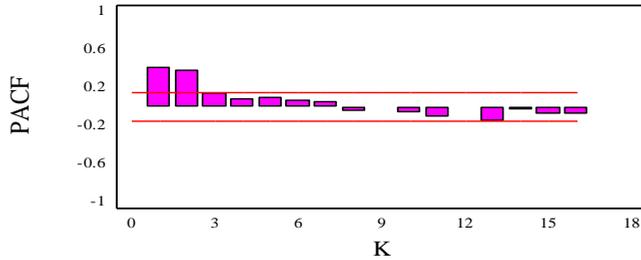
٣-٢: دراسة استقرارية السلسلة

يتم الكشف عن استقرارية السلسلة الزمنية وذلك عن طريق حساب معامل الارتباطات الذاتية والجزئية بين قيم السلسلة وقيم المتأخرة بعدد k من درجات التأخير. تعد السلسلة الزمنية مستقرة إذا هبطت معاملات الارتباطات الذاتية إلى الصفر بعد التخلف الثاني أو الثالث، بعبارة أخرى إذا هبطت هذه المعاملات إلى الصفر بسرعة دل ذلك على استقرارية السلسلة الزمنية والعكس صحيح وكما مبين في الشكل الآتي:



شكل (٢)

معاملات الارتباط الذاتي للسلسلة الأصلية



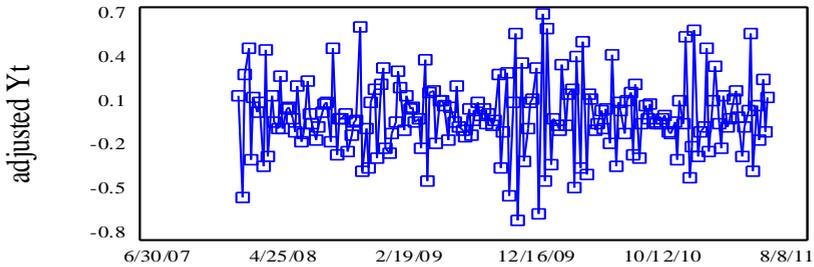
شكل (٣)

معاملات الارتباط الذاتي الجزئي للسلسلة الأصلية

من الشكل (٢) يتبين بأن هناك سبعة معاملات للارتباط الذاتي تقع خارج حدود الثقة أي إنها لا تهبط إلى الصفر بسرعة لذلك نستنتج بأن السلسلة غير مستقرة.

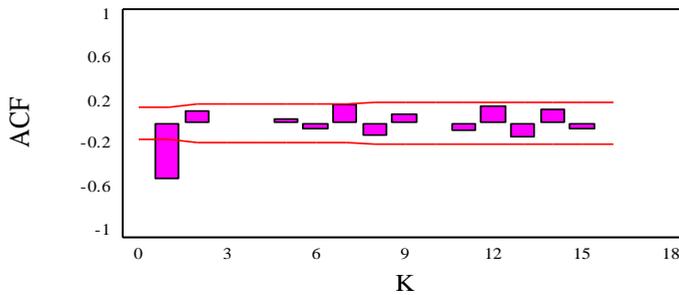
٢-٤: التخلص من عدم الاستقرار للسلسلة الزمنية

من شروط تحليل السلاسل الزمنية أن تكون السلسلة الزمنية تحت الدراسة مستقرة لذلك في حالة عدم استقرارها نحول السلسلة إلى حالة الاستقرار وذلك بإجراء تحويل معين للبيانات في حالة عدم استقرار السلسلة حول التباين أو أخذ الفروقات للسلسلة في حالة عدم استقرار السلسلة حول المتوسط. لذا تم أخذ اللوغارتم للبيانات وكذلك تم أخذ الفرق الأول للسلسلة إذ تحولت السلسلة إلى مستقرة وكما في الأشكال الآتية:



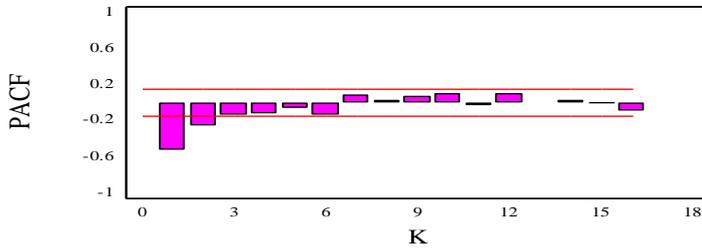
شكل (٤)

السلسلة الزمنية المعدلة



شكل (٥)

معاملات الارتباط الذاتي للسلسلة المعدلة



شكل (٦)

معاملات الارتباط الذاتي الجزئي للسلسلة المعدلة

من خلال الشكل (٥) يتبين بأن كل معاملات الارتباطات الذاتية تنزل نحو الصفر بعد فترة الإزاحة الأولى وتقع ضمن مجال الثقة وكذلك يتبين من الشكل (٦) بأن كل معاملات الارتباطات الذاتية الجزئية تنزل نحو الصفر بعد فترة الإزاحة الثانية وتقع ضمن مجال الثقة، وعليه فإن السلسلة الزمنية تعد مستقرة.

٢-٥: اختيار النموذج الملائم

يتم تحديد النموذج الملائم لبيانات السلسلة الزمنية وذلك عن طريق دراسة سلوك مقدرات معاملات دوال الارتباطات الذاتية ACF ومعاملات دوال الارتباطات الذاتية الجزئية PACF وتحديد درجة النموذج، فقد يظهر أكثر من نموذج ملائم يناسب البيانات وهنا تظهر أهمية خبرة الباحث (Box & Jenkins, 1970).

وعند النظر إلى الأشكال (٥، ٦) لمعاملات الارتباطات الذاتية ومعاملات الارتباطات الذاتية الجزئية يتبين لنا بأن النموذج المتبع نظرياً هو $ARIMA(2,1,1)$ لأن معاملات الارتباطات الذاتية تتنازل أسياً بعد معامل الارتباط الذاتي الجزئي الثاني، وإن معاملات الارتباطات الذاتية تتنازل بشكل ترددي بعد معامل الارتباط الذاتي الأول ثم نختبر النموذج ولكن بالرجوع إلى الجانب العملي نجد أن النموذج النظري $ARIMA(2,1,1)$ غير ملائم لأن معاملات النموذج المذكور غير معنوية وكما في الجدول (٢).

ومن أجل اختيار النموذج الملائم للبيانات تم الاعتماد على أربعة معايير والتي تم ذكرها في الجانب النظري وهي:

- متوسط مربعات الأخطاء (MSE).
- معيار (AIC).
- معيار (AICc).
- معيار (BIC).

إذ تم التطبيق على ثمانية نماذج للسلاسل الزمنية. وقبل تطبيق المعايير المذكورة يجب التأكد من معنوية معاملات النماذج إذ تبين لنا أن هناك ثلاثة نماذج معلماتها معنوية لأن قيمة P.value للنماذج أقل من مستوى المعنوية ($\alpha = 0.05$) وهي ((ARIMA(1,1,0), (2,1,0), (0,1,1))، وأن بقية النماذج غير معنوية لأنه يوجد على الأقل معلمة واحدة قيمة P.value لها أكبر من مستوى المعنوية ($\alpha = 0.05$) وكما مبين في الجدول ذي الرقم (٣).

جدول (٣)

معنوية النماذج

Sig.	Sig.				Models
	MA		AR		
	ϕ_2	ϕ_1	θ_2	θ_1	
Sig. <input type="checkbox"/>			0.00000		ARIMA(1,1,0)
Sig. <input type="checkbox"/>			0.00000	0.00024	ARIMA(2,1,0)
Sig. <input type="checkbox"/>	0.00000				ARIMA(0,1,1)
Notsig. <input type="checkbox"/>	0.00000	0.18332			ARIMA(0,1,2)
Notsig. <input type="checkbox"/>	0.00000		0.11893		ARIMA(1,1,1)
Notsig. <input type="checkbox"/>	0.00001		0.38406	0.80160	ARIMA(2,1,1)
Notsig. <input type="checkbox"/>	0.37893	0.86262	0.63597		ARIMA(1,1,2)
Notsig. <input type="checkbox"/>	0.97970	0.99074	0.98410	0.99929	ARIMA(2,1,2)

المصدر: من إعداد الباحث

لذلك تتم مقارنة ثلاثة نماذج مع بعضها البعض باستخدام المعايير المذكورة أعلاه وكما في الجدول أدناه:

جدول (٤)

قيم معايير المعلومات

BIC	AICc	AIC	Mse	Models
379.92145	379.47351	379.40610	1.8875	ARIMA(1,1,0)
377.91376	377.27632	377.14072	1.78781	ARIMA(2,1,0)
372.86752	372.41958	372.35216	1.7255	ARIMA(0,1,1)

المصدر: من إعداد الباحث

بالنظر إلى الجدول ذي الرقم (٤) يتبين بأن النموذج ARIMA(0,1,1) هو أفضل نموذج لأنه يمتلك أقل قيمة للمعايير مقارنة مع قيم المعايير للنموذجين.

والجدول ذو الرقم (٥) يبين القيمة التقديرية للمعلمة في نموذج ARIMA(0,1,1) وكما يأتي:

جدول (٥)

تقدير معلمة نموذج النموذج ARIMA(0,1,1)

Parameter	Estimate	Stnd. Error	t
MA(1)	0.6828	0.0554	12.3261

المصدر: من إعداد الباحث

٦-٢: اختبار مدى ملاءمة النموذج

بالاعتماد على المعايير يعد نموذج $ARIMA(0,1,1)$ أفضل نموذج ملائم، بعد ذلك يجب أن نتحقق من مدى ملاءمة النموذج من خلال استخدام اختبارات مدى الملاءمة (الفحوص التشخيصية)، ويتم ذلك باختبار معاملات كل من الارتباط الذاتي والارتباط الذاتي الجزئي وذلك باحساب إحصاءات الاختبار وكما يأتي:

جدول (٦)

إحصاءات الاختبارات الخاصة

بمعاملات الارتباط الذاتي والذاتي الجزئي

إحصاءات الاختبارات الخاصة بمعاملات الارتباط الذاتي			النموذج
A	Q_{L-B}	Q_{B-P}	
13.54666503	14.36944723	0.1022236	ARIMA(0,1,1)
إحصاءات الاختبارات الخاصة بمعاملات الارتباط الذاتي الجزئي			
B	AQ_1	$Q_{mon.}$	
12.88237632	12.15364569	0.02930535	

المصدر: من إعداد الباحث

وبمقارنة القيم المحتسبة في الجدول ذي الرقم (٦) مع القيم الجدولية والمذكورة في الجدول ذي الرقم (٧) وبمستوى معنوية $(\alpha = 0.05)$ نلاحظ أن الفروقات غير معنوية وأن معاملات الارتباطات الذاتية والجزئية عشوائية وعلية فإن النموذج المحدد يعد ملائماً ويمكن استخدامه للتنبؤ للسلسلة الزمنية قيد البحث.

جدول (٧)

القيم الجدولية لإحصاءات الاختبار

A,B	χ^2 value	d.f=m-(p+q)	Model
± 0.587	25	15	ARIMA(0,1,1)

المصدر: من إعداد الباحث

٧-٢: التنبؤ المستقبلي وحدود الثقة

تم استخدام النموذج ARIMA(0,1,1) في عملية التنبؤ لأنه اجتاز اختبارات جودة أو ملائمة النماذج المشخصة وتم التنبؤ لتسعة أسابيع وكما يأتي:

جدول (٨)

القيم التنبؤية وحدود الثقة للنموذج ARIMA(0,1,1)

Period	Zt	L.L	U.L
7/5/2011	8.39203	5.67454	12.4109
7/12/2011	8.39203	5.56652	12.6517
7/19/2011	8.39203	5.46526	12.8862
7/26/2011	8.39203	5.36989	13.115
8/2/2011	8.39203	5.27971	13.339
8/9/2011	8.39203	5.19415	13.5588
8/16/2011	8.39203	5.11272	13.7747
8/23/2011	8.39203	5.03504	13.9872
8/30/2011	8.39203	4.96074	14.1967

المصدر: من إعداد الباحث

بالنظر إلى الجدول أعلاه يتبين لنا أن التنبؤ باستخدام النموذج ARIMA(0,1,1) يعطي القيم نفسها لنفسها للفترات المختلفة أي نستطيع التنبؤ بقيمة واحدة وذلك بسبب طبيعة النموذج مع اختلاف فترات الثقة للتنبؤ.

٣: الاستنتاجات والتوصيات

١-٢: الاستنتاجات

تم التوصل إلى استنتاجين وهما كما يأتي:

(١) عند استخدام نموذج $ARIMA(0,1,1)$ فإنه يعطي قيم متشابهة لفترات التنبؤ المختلفة مع اختلاف فترات الثقة، أي نستطيع التنبؤ بقيمة واحدة، وذلك بسبب طبيعة النموذج. لذا نعلم على فترات الثقة للتنبؤ بالسلسلة الزمنية وكما مبين في الجدول ذي الرقم (٨).

(٢) إن المنطقة قيد البحث (ناحية بحركة- أربيل) تصلح لوضع التوربينات الهوائية لتوليد الطاقة الكهربائية، لأن سرعة الرياح في المنطقة كمعدل تساوي 6.92279 متر/ثانية أسبوعياً (Martin, 2003).

٢-٢: التوصيات

يوصي الباحث بأخذ مناطق أخرى في محافظة أربيل وباقي محافظات القطر للتطبيق والتنبؤ بالسلاسل الزمنية المناخية مثل الرطوبة ودرجات الحرارة... الخ، وأخذ حالات أخرى مثل وجود الموسمية في البيانات وكذلك وجود تأثيرات دورية في السلاسل الزمنية.

قائمة المصادر

المصادر العربية

أولاً: الدوريات والمجلات

١. الناصر، جمعة (٢٠٠٩)، "بعض الاختبارات المعدلة لملائمة النماذج للسلسلة الزمنية المناخية في العراق"، المؤتمر الإحصائي العربي الثاني، سرت- الجماهيرية العربية الليبية الشعبية الاشتراكية العظمى، ٢-٤ تشرين الثاني.
[http://www.aitrs.org/Portals/PCBS/Documents/sasc res/res7.pdf](http://www.aitrs.org/Portals/PCBS/Documents/sasc%20res/res7.pdf)

ثانياً: الكتب

١. البري، عدنان ماجد (٢٠٠٢)، طرق التنبؤ الإحصائي، الجزء الأول، كلية العلوم، جامعة الملك سعود، المملكة العربية السعودية.
٢. الوردي، عدنان هاشم (١٩٨٢)، أساليب التنبؤ الإحصائي - طرق وتطبيقات، وزارة التعليم العالي والبحث العلمي، كلية الإدارة والاقتصاد، جامعة البصرة، العراق.

نشرات وإحصائيات:

١. جمهورية العراق، إقليم كردستان، المديرية العامة لزراعة أربيل، مركز الأنواء الجوية الزراعية، نشرات شهرية منشورة، ٢٠٠٨-٢٠١١.

المصادر الاجنبية

1. Akaike, H. (1973), "Information theory and extension of the maximum likelihood principle", In: B. N. petrov and F. Csaki, eds, 2nd International Symposium on Information Theory, Academia Kiado, Budapest, pp.267-281.
2. Box G, Jenkins G. (1970), "Time Series Analysis Forecasting and Control", San Francisco: Holden-Day.
3. Box, G. E. P. & Pierce, David A. (1970), "Distribution of Residual Autocorrelations in Autoregressive", Integrated

- Moving Average Time Series Models), JASA, Vol.65, No.332, PP. (1509-1526).□
4. Davies, Neville & Newbold, Paul (1979), "**Some Power Studies of Portmanteau Test of Time Series Model Specification**", Biometrika, Vol. 66, No.1, PP. (153-155).□
 5. Ljung, G. M. and Box, E. P. (1978), "**On a Measure of Lack of Fit in Time Series Models**", Biometrika, Vol. 65, No. 2, PP. (297-303).□
 6. Makridakis, S. & Wheelwright, S.C. & McGee, V.E. (1983), "**Forecasting Methods and Application**", 2nd ed., John Wiley, New York.
 7. Martin, O.L. Hansen (2003), Aerodynamics of wind turbines, Cromwell press, UK.□
 8. Monti, Anna Clara (1994), "**A Proposal for a Residual Autocorrelation Test in Linear Methods**", Biometrika, Vol. 81, No. 4, PP. (776-780).□
 9. Schwarz, G. (1978), "**Estimating the dimension of a model**", Annuals of Statistics 6, pp.461-464.
 10. Wei, William W.S. (1990), "**Time Series Analysis**, Addison, Wesley Publishing Company.
-