المجلة العراقية للعلوم الاحصائية (9) 2006
ص ص [27-48]

دراسة مقارنة بين نماذج التحويل الخطية ذوات الرتب الدنيا ونموذج ARX ذي الرتب العليا باعتماد السلاسل المكانية

لميعة باقر جواد الجواد*

الملخص

يتضمن البحث مقارنة بين كفاءة النماذج الخطية المختلفة لدالة التحويل للسلاسل الزمنية أو المكانية المستقرة من الرتب الدنيا (Pss1,B.J22221,OE221,ARMAX2221,ARX11k) ونموذج ARX من رتب عليا من حيث النسبة المئوية لمطابقة النموذج لبيانات السلسلة . وقد تمت المقارنة في سلسلة مكانية – زمنية تبين الأثر الحاصل على العوامل البايولوجية للتربة نتيجة للتغير للعوامل الكيمياوية من جهة والفيزياوية من جهة ثانية والاثنين معاً من جهة ثالثة. وبشكل نموذج (مدخلات – مخرجات) بصيغة (متعدد – منفرد) مرةً وبصيغة (متعدد – متعدد) مرة أخرى . وقد تبين أن نماذج arx ذوات الرتب العليا لها نسبة تطابق عالية جداً مقارنةً بنماذج الرتب الدنيا على اختلافها جميعاً .

A comparative study between linear transfer models of low order and ARX model of high order depending on spatial series

ABSTRACT

A comparative study between linear transfer models of low order and ARX model of high order depends on spatial series.

The research includes a comparison between efficiency of the different linear transfer function f stationary time – spatial series with low order models (ARX11k , ARMAX2221 ,O.E221 , B.J22221 , pss1) and high order ARX model according to the percentage rate of coincidence of the model to the data . The comparison is made on a spatial – time series which shows the

*استاذ مساعد/ قسم الاحصاء/ كلية الادارة والاقتصاد/ جامعة بغداد

تاريخ التسلم : 1/ 8 /2005 ـــــــــــــــ تاريخ القبول : 14/ 9 /2005

effect resulted on the biological factors of the soil as a consequence of changes in the chemical factors on one hand and physical factors on the other, and the two factors all are together as an input–output model (multiple–uni) in one time and (multiple–multiple) on another time . It appears that ARX models of high order has a high percentage fit compared with the transfer models of low order .

المقدمة وهدف البحث

في التطبيقات العملية وفي بحوث المحاكاة للسلاسل الزمنية والمكانية لدالة التحويل تؤخذ عادةً النماذج ذوات الرتب الدنيا نظراً للصعوبات التي قد تعترض الباحث في تقدير معلمات النموذج واختبارها وفي اختبار صحة النموذج واختيار الأفضل من بين مجموعة من النماذج قد لا تكون بالعدد القليل . النماذج التي تطبق للسلاسل المستقرة هي على الأغلب نماذج الانحدار الذاتي ARX(r) أو نماذج الأوساط المتحركة MAX(s) أو النماذج المختلطة انحدار ذاتي – أوساط متحركة ARMAX(r,s) لقيم r=0,1 و s=0,1 . هذه النماذج قد تبدو مقنعة ومفيدة في تمثيل السلسلة وتحليلها لأغراض التنبؤ والسيطرة ولكن ما مدى مطابقة هذه النماذج للبيانات وما مدى الاعتماد عليها ؟ وهل يمكن الاكتفاء بالرتب الصغيرة تجنباً للمشاكل التي قد تنجم عن زيادة الرتبة في أي من الاتجاهين الانحدار الذاتي والمتوسطات المتحركة ؟ لغرض الإجابة عن هذه الأسئلة تم تطبيق نماذج دالة التحويل الخطية المختلفة وهي (ARX state-space , B.J , O.E , ARMAX) ذوات الرتب الدنيا على بيانات يمكن عدها سلاسل مكانية لتجربة علمية أجريت على عينات من التربة بهدف التعرف على تأثير العوامل الكيمياوية والفيزياوية (منفصلة ومجتمعة) في أنواع الأحياء المختلفة فيها لما لها من أهمية كبرى في الزراعة . وحسبت النسب المئوية للتطابق لكل نموذج . ثم طبقت نماذج ARX واعتمد معيار AIC لإيجاد أفضل نموذج وحسبت نسب التطابق لهذا النموذج بهدف المقارنة بين هذه النسب والنسب السابقة ولمعرفة أهناك فروق كبيرة بين هذين النوعين من النماذج أم أن هذه الفروق صغيرة بحيث يمكن تجاوزها والاكتفاء بالنماذج ذات الرتب الدنيا.

1. الجانب النظري

أ- النماذج الخطية لدالة التحويل للسلاسل الزمنية[2],[7]

Linear transfer function for time series models

تدعى نماذج السلاسل الزمنية التي تربط مدخلات (بصيغة سلسلة زمنية) بمخرجات (بصيغة سلسلة زمنية أخرى) نماذج التنقية filter أو نماذج المدخلات – المخرجات input - output أو نماذج التحويل transfer function.وابسط صيغة لنموذج خطي لدالة التحويل للسلاسل الزمنية هي :

$$Y(t) = \upsilon(q)\ u(t) + n(t) \qquad \dots \quad (1)$$

حيث $u(t)$هي السلسلة الداخلة و $Y(t)$ السلسلة الخارجة و $\upsilon(q) = \sum_{i=-\infty}^{\infty} \upsilon_i q^i$ هي دالة التحويل وهي متعدد حدود بعملية التراجع q و$n(t)$ هو حد الخطأ أو التشويش . وغالباً ما تكون $\upsilon_i = 0$ عندما $i < 0$ فتدعى بدالة التحويل السببية Causal ويكون النموذج مستقراً عندما $\sum_{i=-\infty}^{\infty} |\upsilon_i| < \infty$.

وقد يكون هناك تأخير زمني قبل أن تستجيب سلسلة المخرجات أو تكون حدود دالة التحويل لا نهائية لذا يفضل كتابة هذه الصيغة كما ياتي :

$$Y(t) = \frac{W_s(q)}{\delta_r(q)} q^b u(t) + n(t) \qquad \dots \quad (2)$$

حيث كل من $W_s(q)$ و $\delta_r(q)$ متعدد حدود من درجة s و r على التوالي بعملية التراجع q و b هي معلمة التأخير الزمني عن استجابة المخرجات للمدخلات . تكتب الصيغة العامة لنموذج دالة التحويل الخطي متعدد المدخلات والمخرجات كما ياتي [3],[2] :

$$A(q)\, Y(t) = \sum_{i=0}^{nu} \frac{B_i(q)}{F_i(q)}\, u_i(t - nk_i) + \frac{C(q)}{D(q)}\, e(t) \qquad \dots \quad (3)$$

حيث $u_i(t)$ هو المدخل رقم i ، $Y(t)$ هو المخرج ، A, B_i, C, D, F

متعددات حدود بعملية التراجع q ، nk_i معلمة التأخير الزمني الناتج عن استجابة المخرجات للمدخل رقم i (وهو عدد صحيح) ، nuعدد المدخلات و $e(t)$ سلسلة التشويش الأبيض .

ومن هذه الصيغة العامة يمكن الحصول على الحالات الخاصة الاتية :

1.دالة التحويل انحدار ذاتي

$$ARX : A(q)Y(t) = B(q)\,u(t-nk) + e(t) \quad \text{... (4)}$$

2. دالة التحويل المختلطة (انحدار ذاتي – اوساط متحركة)

$$ARMAX : A(q)Y(t) = B(q)\,u(t-nk) + C(q)\,e(t) \quad \text{... (5)}$$

3. دالة التحويل (Output – Error)

$$OE : Y(t) = [B(q)/F(q)]\;u(t-nk) + e(t) \quad \text{... (6)}$$

4. دالة التحويل Box – Jenkins

$$B.J : Y(t) = [B(q)/F(q)]\;u(t-nk) + [C(q)/D(q)]e(t) \quad \text{... (7)}$$

ب- نماذج فضاء الحالة State-Space Models (ss) [7],[3]

الطريقة الشائعة لوصف النماذج الخطية هي في استخدام صيغة فضاء الحالة الاتية :

$$X(t+1) = A\;X(t) + B\;u(t) \quad \text{... (8)}$$

$$Y(t) = C\;X(t) + D\;u(t) + e(t) \quad \text{... (9)}$$

إذ أن العلاقة بين المدخل $u(t)$والمخرج $Y(t)$ يتحكم بها متجه الحالة $X(t)$ ذو nx من الأبعاد (nx عدد صحيح) . وبمفهوم دالة التحويل الخطية تكون كما ياتي:

$$Y(t) = G(q)\,u(t) + n(t) \quad \text{... (10)}$$

حيث

$$G(q) = C(q\,I_{nx} - A)^{-1}\;B + D \quad \text{... (11)}$$

و I_{nx} وحدة المصفوفة من درجة nx .

وللزيادة في وصف حد الخطأ $n(t)$ تكون الصيغة لنموذج فضاء الحالة كما ياتي :

$$X(t+1) = A\ X(t) + B\ u(t) + K\ e(t) \quad \ldots \ (12)$$

$$Y(t) = C\ X(t) + D\ u(t) + e(t) \quad \ldots \ (13)$$

فيكون

$$Y(t) = G(q)\ u(t) + H(q)\ e(t) \quad \ldots \ (14)$$

حيث $G(q)$ كما في المعادلة رقم (11) أعلاه و

$$H(q) = C(q\ I_{nx} - A)^{-1}\ K + I_{ny} \quad \ldots \ (15)$$

حيث ny يمثل أبعاد $Y(t)$ و $e(t)$.

ويكفي لحساب نموذج فضاء الحالة الخطي معرفة رتبة النموذج

ج- دالة الارتباط المتقاطعة لنماذج دوال التحويل (CCF) [7]

The Cross-Correlation function for transfer function models

دالة الارتباط المتقاطعة بين المدخلات $X(t)$ والمخرجات $Y(t)$ في النموذج البسيط (المعادلة رقم 1) واللتان يفترض ان تكونا مستقرة بالاشترك (jointly stationary) أي إن كلا منهما مستقرة ($COV(X(t),Y(s))$ والتغاير بينهما هو دالة الى t-s فقط وتكون دالة الارتباط المتقاطعة ρ_{xy} كما ياتي:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y} \quad \ldots(16)$$

$\gamma_{xy}(k)$ هي دالة التغاير بين $Y(t), X(t)$ ، σ_y^2, σ_x^2 هما تباين السلسلتين $Y(t), X(t)$ على التوالي .

يمكن الربط بين دالة الارتباط المتقاطعة $\rho_{xy}(k)$ ودالة التحويل كما ياتي :

$$\rho_{xy}(k) = \frac{\sigma_x}{\sigma_y}[\upsilon_0 \rho_x(k) + \upsilon_1 \rho_x(k-1) + \cdots\cdots] \quad \ldots(17)$$

فإذا كانت المدخلات هي سلسلة التشويش الأبيض والتي تكون فيها $\rho_x(k) = 0$ لكل $k \neq 0$ فان :

$$\upsilon_k = \frac{\sigma_y}{\sigma_x} \rho_{xy}(k)$$

...(18)

وهذا يعني إن دالة الاستجابة النبضية υ_k تتناسب طردياً مع دالة الارتباط المتقاطعة $\rho_{xy}(k)$.

لكي يتم تقدير معلمات النموذج بشكل دقيق تجرى التحويلات اللازمة على كل من المدخلات والمخرجات لتحويلهما إلى سلسلتين مستقرتين ، كما يفضل معرفة النموذج المناسب لكلا السلسلتين وكتابتهما بصيغة تركيبية خطية من سلسلة التشويش الأبيض وكما ياتي :

$$\upsilon_k = \frac{\sigma_\beta}{\sigma_\alpha}\rho_{\alpha\beta}(k) \qquad \text{...(19)}$$

حيث

$$\alpha(t) = \frac{\phi_x(q)}{\theta_x(q)}X(t) \quad ، \quad \beta(t) = \frac{\phi_y(q)}{\theta_y(q)}Y(t) \text{ ، } \beta(t) = \upsilon(q)\alpha(t) + \varepsilon(t) \quad \text{...(20)}$$

د- النسبة المئوية لأفضل تطابق [3] Percentage of Best Fit

تحسب النسب المئوية لأفضل تطابق بالصيغة الاتية :-

$$Fit = \frac{1 - norm\left(Y - \hat{Y}\right)}{norm\left(Y - \hat{Y}\right)} \times 100$$

حيث Y هي المخرجات ، $\hat{Y}$ القيمة التقديرية أما بطريقة المحاكاة أو بطريقة اقل خطأ للتنبؤ .

وكلما كانت النسبة اكبر وفي الاتجاه الموجب كان التطابق بين النموذج والبيانات أفضل .

2. الجانب التطبيقي

جرت محاولة علمية لمعرفة التأثير في العوامل البايولوجية للتربة الحاصل نتيجة لتغير العوامل الكيمياوية والفيزياوية جنبا إلى جنب فضلا عن تغير المكان والزمان من موقع الى اخر ومن زمن الى اخر . فقد تم جمع بيانات من أعماق مختلفة من الأرض وبارتفاعات متغيرة تبعاً لمستوى سطح البحر وبمواقع مختلفة تبعا لقربها أو بعدها عن النهر في شهر كانون الأول . وقد أعيدت التجربة في شهر نيسان إذ سحبت عينات ترابية من أعماق ومواقع مناظرة لسابقتها . وتمت قراءة المتغيرات ذوات العلاقة للمواقع الأفقية المختلفة