المجلة العراقية للعلوم الاحصائية (11) 2007
ص ص [72–94]

مقارنة بين بعض المرشحات الخطية مع التطبيق

عبدالغفور جاسم* خلف صالح يوسف** سامي داود كبارة***

المستخلص :

تمت في هذا البحث دراسة بعض أساليب ترشيح السلسلة الزمنية المتمثلة بمرشح وينر من خلال نموذج الانحدار الذاتي والأوساط المتحركة والمرشح الفرقي وتم اشتقاق بعض النتائج النظرية لتمكننا من الحصول على هذه المرشحات وأهم صفاتها وخصائصها. كما تم استخدام برامج حاسوبية بلغة C^{++} لغرض إجراء عملية الترشيح وفق خوارزمية أعدت من قبلنا لهذا الغرض وبعدها تم اجراء تطبيق عملي على بيانات السلسلة الزمنية الخاصة بمرض سرطان الرئة في مدينة الموصل للفترة من 1990-2000.

Comparison between some linear filter with application

ABSTRACT

This paper deals with the comparison between some linear filter for time series represented by Weiner filter through autoregressive model, moving average model and difference filter. An application for lung cancer Time Series in Mosul city during (1990-2000) is also considered.

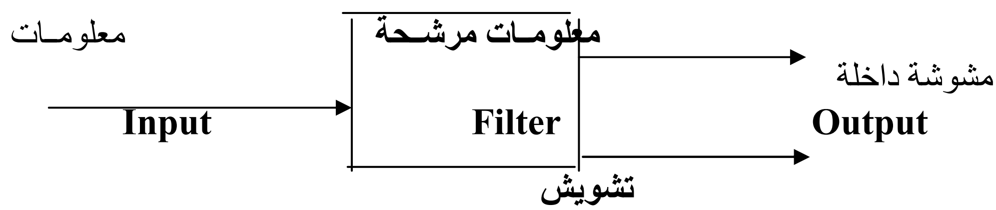
* استاذ مساعد/ كلية علوم الحاسبات والرياضيات
** مدرس مساعد/ قسم الرياضيات/ كلية العلوم/ جامعة تكريت
*** استاذ مساعد/قسم الرياضيات/ الجامعة المستنصرية

تاريخ التسلم : 25/ 6 /2006 ـــــــــــــــــــــ تاريخ القبول : 7 / 1 /2007

المقدمة :

تعد مسألة الترشيح Filtering إحد المحاور الأساسية لتحليل السلاسل الزمنية ويمكن عدّها الولادة الحقيقة لموضوع الترشيح في منتصف القرن العشرين عندما برزت أهميته في التطبيقات العسكرية. ويستخدم مصطلح الترشيح في حياتنا اليومية كثيراً كترشيح الماء من الشوائب، ترشيح الهواء الداخل الى الفم باستخدام الكمامات، ترشيح دخان السجائر وترشيح موجات الراديو من التشويش. ويتعلق هذا الموضوع بالعديد من المواضيع مثل الإحصاء والتحليل العددي والجبر الخطي وعلوم الحاسبات والهندسة الكهربائية ونظرية السيطرة.

إن الاستخدام الحديث لمصطلح الترشيح يعني فصل المشاهدات لظاهرة معينة فور ملاحظتها الى مركبتين الأولى تمثل الإشارة Signal و نهتم بمعرفة مقدارها والثانية تمثل التشويش Noise وبدوره يمثل العوامل الطارئة والمفاجئة التي تكون خارج سيطرتنا والشكل الآتي يوضح عملية الترشيح[4] .



(الشكل 1)

وقد اعتمدنا في هذا البحث على تطبيق دراسة ترشيح السلاسل الزمنية في الجانب الطبي اذ تناول البحث مرض سرطان الرئة وقد اعتمد البحث على الإحصائيات المتوفرة في مستشفى الطب الذري في مدينة الموصل لدراسة السلسلة الزمنية من سنة 1990 – 2000.

تعود الجذور الأولى لمسألة الترشيح الى نهاية العشرينات وبداية الثلاثينيات من القرن العشرين، إلا إن الاهتمام بالموضوع ازداد منذ ظهور مرشح وينر Wiener Filter(1949). الذي صممه العالم الرياضي وينر أثناء الحرب العالمية الثانية، اذ استخدم هذا المرشح في التطبيقات العسكرية والتقنية المختلفة.

وفي عام 1960 قام الباحث Kalman [2] بدراسة مرشح وينر وقدم حلا اكثر سهولة وعمومية لمسالة الترشيح وخاصة في الحالات التي يكون فيها الإشارة أو التشويش غير مستقر.

Non–stationary. وفي عام 1960 قام كل من الباحثين كالمان وبوسي بأجراء بحث مشترك احتوى على نتائج جديدة لمسالة الترشيح .

لقد استخدم الترشيح في معالجة نظام نقل البيانات الرقمية العالية وهو ما قام به Nowak عام 1968 حيث استخدم دوال ترشيح خاصة طبيعية بخصائص طوريه متميزة.

أما الباحثان Jonge&Tong [1] فقد قاما بتطبيق أسلوب خاص للترشيح اللاخطي Non-LinearFiltering في نظام الملاحة البحرية . وقد صاغا مسائل الترشيح اللاخطية لتلائم المسألة الفيزيائية، وأشتقا خوارزميات الترشيح التعاقبية لنموذج رياضي معين.

وفي عام 1976 قام كل من الباحثين Harrsision and Stevens بدراسة استخدما فيها مرشح كالمن وبتطبيق أسلوب بيز Baysian Approach لمعالجة مسالة التكهن [6] Forecasting ولاقت دراسته اهتماماً واسعا وخصوصاً في المجال الاقتصادي.

أما الباحث Priestly [8] فقد اشتق عام 1978 خوارزمية مرشح كالمن من خلال أسلوب التحليل العاملي ،كما درس العلاقة بين فضاء الحالة وتكهن بيز.

أما الدراسات الحديثة التي تناولت الترشيح بنوعيه الخطي واللاخطي ولا سيما في التسعينات، فأغلبها تتبع أسلوب فضاء الحالة State Space.

تؤكد المصادر التاريخية أن عالم الرياضيات الفرنسي Fourier أول من بحث في السلاسل الزمنية كمتسلسلة من عدد غير محدد لدوال الجيب والجيب تمام، وتقديراً لجهوده سميت سلسلة فورير Fouirer series [7].

في عام 1926 بين الباحث Yule [7] إن سلسلة فورير لايمكن أن تعطي صورة واضحة عن معالم السلسلة الزمنية لغرض التنبؤ بها أو تقدير الدورات للسلسلة الزمنية لأن طولها غير ثابت وغير محدد لذلك وضع الباحث في سنة 1927 فكرة الانحدار الذاتي (Autoregressive) أي ان السلسلة الزمنية في أي زمن هي دالة انحدار ذاتية للفترات السابقة المتعاقبة مضافاً اليها الخطأ العشوائي وقد درس Yule الانحدار الذاتي لغاية الدرجة الرابعة.

وفي عام 1931 وسع الباحث Walker [7] أسلوب Yule ودرس الانحدار الذاتي للرتبة p ويرمز له AR(P) . وفي عام 1937 وضع الباحث Stutsky [7] نموذج الأوساط المتحركة (Moving Average Model) وهو تمثيل للسلسلة الزمنية كدالة انحدار خطية للأخطاء العشوائية للفترات الزمنية السابقة والمتلاحقة ويرمز لها بالرمز MA(q).

ووضع العالم Wold [7] عام 1954 فكرة النموذج المختلط (الانحدار الذاتي- الوسط المتحرك) (Autoregressive-Moving Average Model) الذي يكتب بشكل مختصر ARMA .

وقدم الباحثان(Box&Jenkins) استنتاجات واسعة في مجال تحليل السلاسل الزمنية بصورة عامة ونماذج ARMA بصورة خاصة اذا اقترن أسمهما كمرادف لنماذج ARMA .

وأجريت دراسة دكتوراه في جامعة بغداد، اذ قامت الباحثة بثينة عبد الجادر عام 1994 بدراسة حول تحليل السلاسل الزمنية التي تحتوي على قيم مفقودة ودراسة وتشخيص النموذج وتحديد درجته وتقدير معالمه ومن ثم أجراء التنبؤات المستقبلية باستخدام مرشح كالمن.

وأجريت دراسة ماجستير عام 1996 في جامعة الموصل قام بها الباحث الحمداني حيث كون مرشح كالمن بالاعتماد على أسلوب بيز لعدد من النماذج الحركية الخطية[4].

وقام الباحث الطائي [4] عام 2002 بدراسة كفاءة بعض المرشحات باستخدام المحاكاة حيث قارن بين ثلاثة مرشحات (كالمن ، استجابة النبضة ، ألفا بيتا) وأستنتج بأن مرشح استجابة النبضة Finite –Duration Impluse Response Filter يتمتع بإمكانيات ترشيح جيدة تجعله يتفوق على المرشحين الآخرين.

1- الجانب النظري :

المرشح Filter هو تطبيق خطي من عملية تصادفيه الى عملية تصادفيه أخرى مثل الصندوق الأسود يملك مدخلات ومخرجات بالتعاقب (لاحظ الشكل 1.1). ومن الجانب الرياضي نفرض ان $\{X_t\}$ تمثل المدخلات Input و $\{Y_t\}$ تمثل المخرجات Output ونفرض انه لدينا مجموعة من الدوال f ذات القيم المعقدة [5] Complex-valued Functions

φ تطبيق دالي معرفة كالآتي :

$$\varphi : f \longrightarrow \boldsymbol{f}$$

$$\varphi(X_t)=Y_t$$

وتكون φ مرشح خطي Linear filter إذا تحقق الشرط الآتي:

$$\forall\ (X_{t1},X_{t2})\ \in f\ \ ,\ X_{t2}\,\lambda_2 + X_{t1}\,\lambda_1\ \in f$$

بحيث ان (λ_1, λ_2) قيم معقدة فأن :

$$\varphi\,(\lambda_1 X_{t1} + \lambda_2 X_{t2}) = \varphi\,(\lambda_1 X_{t1}) + \varphi\,(\lambda_2 X_{t2})$$

وهناك نوعان من المرشحات بالاعتماد على الزمن :

أ- المرشحات ذوات الزمن المتقطع Discreate Time Filters ومنها: [5]

- المرشح ذو الإزاحة الخلفية Backward Shift Filter

$$X_t = Y_{t-1},\ t \in T$$

- المرشح الفرقي Difference Filter

$$X_t = Y_t - Y_{t-1} = \Delta Y_t\ ,\ t \in T$$

- مرشح الانحدار الذاتي Autoregressive Filter

$$X_t = \sum_{u=0}^{\infty} g_u\, Y_{t-u}$$

إذ ان $g_0=1$.

- المرشح ذو الأوساط المتحركة Moving Average Filter

$$X_t = \sum_{u=0}^{\infty} a_u Y_{t-u},\ t \in T$$

- المرشح الفيزيائي المرن Physical Realizable Filter

$$X_t = \sum_{j=-\infty}^{\infty} c_j\, X_{t-j}$$

ب - المرشحات ذوات الزمن المستمر Continous Time Filter ومنها : [5]

- المرشح الأسي Exponential Filter

$$Y_t = \int_{-\infty}^{t} e^{-B(t-s)}\ X_s\, ds \qquad t \in T$$

- المرشح الخطي العام General linear Filter

$$Y_t = \int_{-\infty}^{\infty} w(t-s) X_s ds$$

- المرشح الفيزيائي المرن Physical realizable filter

$$Y_t = \int_{-\infty}^{t} w\,(t-s)\, X_s ds$$

1.1 تمثيل مرشح الفروقات والعملية الخطية العامة:[6]

Difference Filter Representation and General Linear Operation

ان الصيغة العامة لتمثيل مرشح الفرو قات يمكن كتابتها بالشكل الآتي :

$$X_t= \Delta^d Y_t \tag{1-1}$$

حيث Y_t تمثل السلسلة الأصلية Input Series

X_t تمثل السلسلة الجديدة Output Series

وان Δ =1-B، B يمثل مؤثر الإزاحة الخلفي والذي يعرف كما ياتي:-

$$BX_t=X_{t-1},B^2X_t=X_{t-2},\ldots,B^kX_t=X_{t-k}$$

أما d فيمثل عدد الفرو قات ويأخذ القيم ...,1,2,3

عندما تكون d=1 يسمى مرشح الفرق من الدرجة الأولى وتصبح المعادلة (1-2) كما ياتي:

$$\begin{aligned}X_t&= \Delta Y_t\\&=(1-B)Y_t\\&=Y_t-BY_t\\&=Y_t-Y_{t-1}\end{aligned}$$

عندما تكون d=2 يسمى مرشح الفرق من الدرجة الثانية وتصبح المعادلة (1-2) كما ياتي:

$$\begin{aligned}X_t&= \Delta^2 Y_t\\&=(1-B)^2Y_t\\&=(1-2B+B^2)Y_t\\&=Y_t-2BY_t+B^2Y_t\\&=Y_t-2Y_{t-1}+Y_{t-2}\end{aligned}$$

و يمكن كتابة الصيغة الخطية العامة للمرشح وكما ياتي:

$$X_t = \sum_{j=o}^{n} c_j\ Y_{t-j} \qquad (2\text{-}1)$$

في حالة n=1 وباستخدام العلاقة (2-1) تكون

$$X_t = c_0 Y_t + c_1 Y_{t-1} \quad ; c_0=1, c_1= -1$$

وعندما n=2 تكون

$$X_t = \sum_{j=o}^{2} c_j\ Y_{t-j}$$

$$= c_0 Y_t + c_1 Y_{t-1} + c_2 Y_{t-2} \quad ; c_0=1, c_1= -2, c_2=1$$

وبصورة عامة يمكن حساب قيم الثوابت باستخدام نظرية ذي الحدين فالثوابت تمثل قيم التوافيق(Combinations value)

$$(x+y)^n = C_0^n x^n + C_1^n x^{n-1} + --- + C_n^n y^n$$

2.1 خصائص المرشح الفرقي (Difference Filter)

1 . **الوسط الحسابي** : (Mean)

إذا فرضنا بان $\{Y_t\}$ عملية مستقرة ولها وسط يساوي صفرا $E(Y_t)=0$ نستنتج أن $\{X_t\}$ يكون وسطها الحسابي صفرا.

$$E(X_t) = E\left[\sum_{j=o}^{n} c_j Y_{t-j}\right]$$

$$= \sum_{j=o}^{n} c_j E\left[Y_{t-j}\right]$$

$$= 0$$

وهذا يعني بأن الوسط للعملية المرشحة يساوي صفرا ذا كان الوسط للعملية الأصلية يساوي صفرا.

2. **التباين**: (Variance)

$$V[X_t]=\gamma_x(0) = \sum_{j=o}^{n} c_j^2 V[Y_{t-j}]$$

$$= \sum_{j=o}^{n} c_j^2 \sigma^2 y_{t-j}$$

3. **التغاير**: (Covariance)

$$E[X_t X_{t+s}] = E[\sum_{j=o}^{\infty} c_j Y_{t-j} \sum_{k=0}^{\infty} c_k Y_{t+s-k}]$$

$$= \sum_{j=o}^{\infty}\sum_{k=0}^{\infty} c_j c_k E[Y_{t-j} Y_{t+s-k}]$$

$$= \sum_{j=o}^{\infty}\sum_{k=0}^{\infty} c_j c_k \gamma_{s-k+j}$$

وهو لا يعتمد على t ولهذا فأن X_t تكون مستقرة من الدرجة الثانية.- Second order Stationary

وبما ان $E(Y_t)=0$ فإن :

$$Cov[X_t, X_{t+k}] = \gamma_x(k) = E[X_t X_{t+k}] - E[X_t] E[X_{t+k}]$$

$$= E[X_t X_{t+k}]$$

$$= \sum_{j=o}^{\infty}\sum_{k=0}^{\infty} c_j c_k \gamma_{s-k+j}$$

4.دالة الارتباط: Correlation Function

$$\rho_k = \frac{\text{cov}(x_t, x_{t+k})}{\text{var}(x_t)} = \frac{\gamma_x(k)}{\gamma_x(0)}$$

$$= \frac{-\gamma_{y^{(K-1)}} + 2\gamma_{y^{(K)}} - \gamma_{y^{(K+1)}}}{\sum_{j=0}^{n} c_j^2 \sigma_{y_{t-j}}^2}$$

5.دالة الكثافة الطيفية: Spectral Density Function

$$f_x(\omega)=\frac{1}{2\pi}\left[\sum_{k=-\infty}^{\infty}\rho_k e^{-i\omega k}\right]$$

$$=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\left[\frac{-\gamma_y(k-1)+2\gamma_y(k)-\gamma_y(k+1)}{\sum_{j=0}^{n}c_j^2\sigma_{y_{t-j}}^2}\right]e^{-i\omega k}$$

3.1 مرشح وينر من خلال نموذج الانحدار الذاتي AR(P) : [8]

ليكن لدينا السلسلة المستقرة $\{Y_t\}$ المتمثلة بالنموذج AR(P) الذي يحقق المعادلة الفرقية الآتية:

$$y_t=a_1y_{t-1}+a_2y_{t-2}+\ldots+a_py_{t-p}+z_t \qquad (3\text{-}1)$$

وباستخدام مؤثر الإزاحة الخلفي (B) يمكن إعادة كتابة المعادلة(3-1) بما ياتي:

$$(1-a_1B-a_2B^2-\ldots-a_pB^P)y_t=z_t$$

$$y_t=(1-a_1B-\ldots-a_pB^p)^{-1}z_t$$

نفرض أن

$$G(B)=(1-a_1B-a_2B^2-\ldots-a_pB^P)^{-1}$$

$$=\sum_{u=0}^{\infty}g_uB^u,\ g_0=1$$

فاذا كانت لدينا مجموعة من المشاهدات من $\{Y_t\}$ ولتكن $\ldots, y_{t-1}, y_t$ ونرغب في تقدير x_{t+m} اذ ان m عدد صحيح ويمكن حساب القيم المرشحة بدلالة معادلة وينر الآتية :

$$\hat{x}_{t+m} = \frac{[B^{-m} G_0(B)]_+}{G_y(B)} y_t \tag{4-1}$$

$$= A(B)\, y_t$$

حيث أن:

$$G_0(B) = \frac{\sigma}{\sqrt{2\pi}}[G(B) - G^{-1}(B^{-1})]$$

$$[B^{-m} G_0(B)]_+ = \frac{\sigma}{\sqrt{2\pi}}[B^{-m}\sum_{u=0}^{\infty} g_u B^u - B^{-m} G^{-1}(B^{-1})]_+$$

$$= \frac{\sigma}{\sqrt{2\pi}}[B^{-m}\sum_{u=0}^{\infty} g_u B^u]$$

كون الحد الثاني في المعادلة أعلاه سالباً

$$= \frac{\sigma}{\sqrt{2\pi}}[\sum_{u=0}^{\infty} g_u B^{u-m}]$$

$$= \frac{\sigma}{\sqrt{2\pi}}[\sum_{u=0}^{\infty} g_{u+m} B^u]$$

$$A(B) = \frac{\sum_{u=0}^{\infty} g_{u+m} B^u}{\sum_{u=0}^{\infty} g_u B^u}$$

$$\hat{x}_{t+m} = A(B)y_t$$

$$= \frac{\sum_{u=0}^{\infty} g_{u+m} B^{u}}{\sum_{u=0}^{\infty} g_{u} B^{u}} y_t \qquad (5\text{-}1)$$

$$= \frac{\sum_{u=0}^{\infty} (a_1 + a_2 B + a_3 B^2 + \ldots)^{m+u} B^u}{\sum_{u=0}^{\infty} (a_1 + a_2 B + a_3 B^2 + \ldots)^{u} B^u} y_t$$

$$= (a_1 + a_2 B + a_3 B^2 + \ldots)^{m} y_t \qquad (6\text{-}1)$$

إن المعادلة(2-6) هي المعادلة الترشيحية العامة لنموذج الانحدار الذاتي AR(p).
عندما تكون m=1 تصبح المعادلة الترشيحية (2-6) كالآتي :

$$\hat{x}_{t+1} = (a_1 + a_2 B + a_3 B^2 + \ldots)\, y_t$$

$$= a_1 y_t + a_2 y_{t-1} + a_3 y_{t-2} + \ldots$$

عندما تكون m=2 تصبح المعادلة الترشيحية (2-6) كالآتي:

$$\hat{x}_{t+2} = (a_1 + a_2 B + a_3 B^2 + \ldots)^{2} y_t$$

ان القيم المرشحة $\hat{x}_{t+m}$ يمكن الحصول عليها عندما m=1,m=2 ولكن عندما تكون قيمة mاكبر من 2 فان اشتقاق المعادلة الترشيحية يكون معقداً .

4.1 مرشح وينر من خلال نموذج الأوساط المتحركة MA (q) [1]

ليكن لدينا النموذج MA(q) والذي يحقق المعادلة الفرقية الآتية:

$$y_t = z_t + b_1 z_{t-1} + \ldots + b_q z_{t-q}$$

وباستخدام مؤثر الإزاحة الخلفي (B) تصبح المعادلة اعلاه كالآتي:

$$y_t = (1 + b_1 B + \ldots + b_q B^q) z_t$$

نفرض إن

$$G(B) = 1 + b_1 B + \ldots + b_q B^q$$

ويمكن حساب القيم المرشحة بدلالة معادلة وينر [8]

$$\hat{x}_{t+m} = \frac{[B^{-m} G_0(B)]_+}{G_y(B)} y_t \qquad (7-1)$$

ولتسهيل عملية الاشتقاق نفرض إن لدينا النموذج MA (1) فتصبح المعادلة الفرقية

$$y_t = (1 + b_1 B) z_t$$

البسط يكون موجباً في المعادلة (7-2) فقط عندما تكون m =1 فتصبح المعادلة كالآتي :

$$\hat{x}_{t+m} = \frac{b_1}{1 + b_1 B} y_t$$

$$= b_1 (1 + b_1 B)^{-1} y_t$$

$$= b_1 (1 - b_1 B + b_1^2 B^2 - b_1^3 B^3 + \ldots) y_t$$

$$= b_1 [\sum_{u=0}^{\infty} (-1)^u b_1^u y_{t-u}]$$

$$= \sum_{u=0}^{\infty} (-1)^u b_1^{u+1} y_{t-u}$$

في حالة m=1 تصبح المعادلة الترشيحية

$$\hat{x}_{t+1} = \sum_{u=0}^{\infty} (-1)^u b_1^{u+1} y_{t-u}$$

$$= b_1 y_t - b_1^2 y_{t-1} + b_1^3 y_{t-2} - \ldots$$

وان القيم المرشحة $\hat{x}_{t+m} = 0$ عندما m=2,3,… لأن البسط في المعادلة(1-7) يكون سالباً فالمقدار الموجب يساوي صفر.

وبصورة عامة تكون المعادلة الترشيحية كالآتي :

$$\hat{X}_{t+m} = \frac{[B^{-m}\{b_1 + b_2 B + b_3 B^2 + \ldots\}]_+}{(1 + b_1 B + b_2 B^2 + \ldots)} y_t$$

2- الجانب التطبيقي

تضمن الجانب التطبيقي من هذا البحث جمع البيانات المتوافرة من مستشفى الطب الذري في مدينة الموصل والخاصة بمرض سرطان الرئة.

باستخدام البرنامج الإحصائي الجاهز Statistica (1999) تم تطبيق الأفكار التي طرحت في الفصل الثاني، اذ تمت ملاءمة نموذج انحدار ذاتي من الرتبة الثالثة AR(3) للسلسلة ونموذج أوساط متحركة من الرتبة السادسة MA(6) وكذلك تم إيجاد مرشح الفروقات Difference Filter.

وتم إيجاد القيم الترشيحية لكل مرشح من خلال المعادلات الترشيحية وباستخدام برامج تم إعدادها من قبلنا وبلغة C^{++} مرفقة لاحقاً.

وأخيراً تم فحص كفاءة تلك المرشحات الثلاث من خلال معيار معدل مربع الخطأ Mean Square Error ومعيار معدل القيم المطلقة لنسب الأخطاء Mean Absolute Percentage Error.

ملاءمة نموذج الانحدار الذاتي AR(P) :

باستخدام البرنامج الإحصائي الجاهز Statistica (1999) تم تقدير معالم النموذج المشخص بطريقة الامكان الاعظم المضبوط والجدول الاتي يمثل مقدار الرتبة (p) للنموذج والانحراف المعياري (SD) لبواقي النموذج والمعيارين (BIC) و (MSE).

الجدول (1): ملاءمة نموذج AR(P) لسلسلة المصابين بمرض سرطان الرئة للفترة من 1990−2000 والمسجلين في مستشفى الطب الذري في مدينة الموصل

P	SD	MSE	BIC
1	5.569	14.363	458.226
2	3.702	14.212	355.306
3	3.605	13.443	353.178
4	3.604	13.548	357.987
5	3.583	13.451	361.326
6	3.572	13.500	365.397
7	3.573	13.602	370.353
8	3.573	13.712	375.235
9	3.535	13.487	377.938

ومن ملاحظة الجدول أعلاه نجد إن النموذج الملائم للسلسلة هو نموذج الانحدار الذاتي من الرتبة الثالثة AR(3) حيث كانت قيم MSE,BIC هي اقل القيم مقارنة مع الرتب الأخرى و الصيغة الرياضية له هي :-

$$y_t = 0.7188 \quad y_{t-1} - 0.0507 \quad y_{t-2} + 0.24909 \quad y_{t-3} + z_t$$

S.E 0.0856 0.1066 0.8612

S.D= 3.605

MSE=13.443

BIC=353.178

ملاءمة نموذج الأوساط المتحركة MA(q) :

بالطريقة المشار اليها بالفقرة السابقة تم تقدير معالم النموذج كما موضح في الجدول الاتي :

الجدول (2): ملاءمة نموذج الأوساط المتحركة MA(q) لسلسلة المصابين بمرض سرطان الرئة للفترة من1990–2000 والمسجلين في مستشفى الطب الذري في مدينة الموصل.

q	SD	MSE	BIC
1	4.144	17.175	380.200
2	3.980	15.959	374.422
3	3.733	14.153	362. 390
4	3.687	13.915	363.998
5	3.571	13.151	360.441
6	3.496	12.709	359.719
7	3.466	12.591	362.326
8	3.466	12.692	373.087
9	3.460	12.748	371.632

ومن ملاحظة الجدول الذي في أعلاه نجدان أفضل نموذج يلائم السلسلة هو من الرتبة السادسة MA(6) وصيغته الرياضية هي :-

$$y_t = z_t - 0.6735\, z_{t-1} - 0.3742\, z_{t-2} - 0.4972\, z_{t-3} - 0.3352\, z_{t-4} - 0.2977\, z_{t-5} - 0.2602\, z_{t-6}$$

Std.Error 0.0908 0.1032 0.1054 0.1026
0.0856 0.1034

S.D=3.496
MSE=12.709
BIC=359.719

ومن مقارنة الجدولين (1) و (2) نلاحظ تفوق نموذج الانحدار الذاتي AR(3) على نموذج الاوساط المتحركة بأقل BIC .

وبما إن بحثنا مقارنة بين مرشحات الانحدار الذاتي والاوساط المتحركة سوف نستخدم النموذجين للترشيح ومن المقارنة نختار النموذج الافضل الذي يلائم السلسلة .

1.2 **الترشيح** Filtering

2.1.2: مرشح وينر من خلال النموذج AR (P) :

بعد إن تمت ملاءمة نموذج الانحدار الذاتي AR(3) وتتمثل معادلته بالآتي:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + z_t$$

$$a_1 = 0.71988 \quad a_2 = -0.0507 \quad a_3 = 0.24909$$

والآن يمكننا البدء بالعملية الترشيحية من المعادلة الآتية :- (لاحظ [2])

$$\widehat{x}_{t+m} = (a_1 + a_2 B + a_3 B^2 + \ldots)^m y_t$$

حيث m عدد صحيح يمثل عدد خطوات الترشيح .

عندما m=1 تصبح المعادلة الترشيحية للنموذج كالآتي :-

$$\widehat{x}_{t+1} = 0.71988\, y_t - 0.0507\, y_{t-1} + 0.2490\, y_{t-2}$$

وعندما m=2 فأن المعادلة الترشيحية تصبح كالآتي :-

$$\widehat{x}_{t+2} = (a_1 + a_2 B + a_3 B^2)^2 y_t$$

$$= a_1^2 y_t + 2a_1 a_2 y_{t-1} + (a_2^2 + 2a_1 a_3) y_{t-2} + 2a_2 a_3 y_{t-3} + a_3^2 y_{t-4}$$

$$= 0.5181 y_t - 0.0729 y_{t-1} + 0.7194 y_{t-2} - 0.0252 y_{t-3} + 0.0620 y_{t-4}$$

وباستخدام البرنامج الحاسوبي الذي تم إعداده من قبلنا بلغة C^{++} حصلنا على القيم المرشحة .والشكلان (3-12)،(3-13) يمثلان مقارنة بين السلسلة الأصلية و السلسلة المرشحة من خلال النموذج AR(3) عندما m=1 ،m=2 على التوالي .

ويمثل الشكل (3) رسم السلسلة المرشحة من خلال النموذج أعلاه .

2.1.2 مرشح وينر من خلال النموذج MA (q) : [2]

نلاحظ بأن نموذج الأوساط المتحركة هو MA(6) .

عندما m=1 تكون المعادلة الترشيحية كالآتي :

$$\hat{x}_{t+1} = [\frac{b_1 + b_2 B + b_3 B^2 + b_4 B^3 + b_5 B^4 + b_6 B^5}{1 + b_1 B + b_2 B^2 + b_3 B^3 + b_4 B^4 + b_5 B^5 + b_6 B^6}] y_t$$

ولإيجاد مفكوك معكوس المقام (inverse) نستخدم النتيجة الاتية [1] :

$$(1 + b_1 B + b_2 B^2 + b_3 B^3 + b_4 B^4 + b_5 B^5 + b_6 B^6)^{-1} = \sum_{u=0}^{\infty} (-1)^u (b_1 B + b_2 B^2 + b_3 B^3 + b_4 B^4 + b^5 B^5 + b_6 B^6)^u$$

$$\cong 1 - (b_1 B \ + b_2 B^2 + b_3 B^3 + b_4 B^4 + b_5 B^5 + b_6 B^6)$$

$$\hat{x}_{t+1} = [b_1 (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$+ [b_2 B (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$+ [b_3 B^2 (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$+ [b_4 B^3 (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$+ [b_5 B^4 (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$+ [b_6 B^5 (1 - b_1 B - b_2 B^2 - b_3 B^3 - b_4 B^4 - b_5 B^5 - b_6 B^6)] y_t$$

$$= b_1 y_t - (b_1^2 - b_2) y_{t-1} - (2b_1 b_2 - b_3) y_{t-2} - (2b_1 b_3 + b_2^2 - b_4) y_{t-3} - (2b_1 b_4 + 2b_2 b_3 - b_5) y_{t-4}$$
$$- (2b_1 b_5 + 2b_2 b_4 + b_3^2 - b_6) y_{t-5} - (2b_1 b_6 + 2b_2 b_5 + 2b_3 b_4) y_{t-6} - (2b_2 b_6 + 2b_3 b_5 + b_4^2) y_{t-7}$$
$$- (2b_3 b_6 + 2b_4 b_5) y_{t-8} - (2b_4 b_6 + b_5^2) y_{t-9} - (2b_5 b_6) y_{t-10} - b_6^2 y_{t-11}$$

وبتعويض القيم العددية للمعلمات $b_1,b_2,b_3,\ldots,b_6$ تصبح المعادلة الترشيحية كالآتي:

$$\hat{X}_{t+1} = -0.6735y_t - 0.8278y_{t-1} - 1.0012y_{t-2} - 1.4490y_{t-3} - 1.1213y_{t-4} - 1.967y_{t-5}$$
$$-0.9064y_{t-6} - 0.6030y_{t-7} - 0.4582y_{t-8} - 0.2630y_{t-9} - 0.1549y_{t-10} - 0.0686y_{t-11}$$

وباستخدام البرنامج الذي تم إعداده بلغة C^{++} حصلنا على القيم المرشحة .
ويمثل الشكل (4) رسم السلسلة المرشحة من خلال النموذج أعلاه ونلاحظ إن النموذج لا يصلح للترشيح .

3.1.2 : **المرشح الفرقي :Difference Filter**

من اجل تطبيق ما تطرقنا إليه في الجانب النظري نأخذ السلسلة الزمنية y_t وبأستخدام الفرق الأول d=1 نحصل على المرشح الفرقي من الدرجة الأولى (First Difference Filter) وتكون معادلته كالاتي :

$X_t=Y_t-Y_{t-1}$

الشكل (5) يمثل رسم السلسلة المرشحة مع السلسلة الاصلية عندما d=1 ونلاحظ منه إن السلسة كأنما تملك وسطاً ثابتاً ، واننا فقدنا مشاهدة واحدة فقط اذ إن السلسة المرشحة تحوي على 131 قيمة مرشحة ويستخدم الفرق الأول لازالة الاتجاه العام.
إما عندما d=2 فأننا نحصل على مرشح فرقي من الدرجة الثانية (Second Difference Filter) وتكون معادلته كالاتي :

$X_t=Y_t-2Y_{t-1}+Y_{t-2}$

والشكل (6) يمثل رسم السلسة المرشحة مع السلسة الاصلية عندما d=2 فأننا نفقد مشاهدتين وبهذا تحوي السلسلة المرشحة على 130 قيمة مرشحة .
واغلب السلاسل الزمنية نرشحها باستخدام الفرق الأول فقط لانه يكفي لجعل السلسة مستقرة في الوسط وبهذا يكون المرشح عندما d=1 أفضل منه عندما d=2 ويتبين ذلك واضحاً من خلال المعايير الاحصائية الموجودة في الجدول رقم (3) ومن ملاحظة الاشكال المكورة انفاً والمعايير الاحصائية في الجدول الاتي نلاحظ تفوق مرشح الانحدار الذاتي على المرشحين الآخرين .

الجدول (3): يوضح المعايير الإحصائية للمرشحات الثلاثة

المرشح	AR(3)		MA(6)	Difference	
	m=1	m=2	m=1	d=1	D=2
MSE	1.846	10.441	Big value	52.2	79.9
MAPE	0.238	0.757	12.087	1.357	1.543

الاستنتاجات Conclusions

إن اهم الاستنتاجات التي تم التوصل اليها من خلال هذا البحث هي كألاتي :

1- تمت ملاءمة نماذج الانحدار الذاتي والأوساط المتحركة للسلسلة الزمنية لمرض سرطان الرئة باستخدام طريقة بوكس – جينكنز وبالاعتماد على المعيارين BIC,MSE وجدنا أن النماذج الأفضل هي AR(3) ، MA(6) .

2- من خلال المرشحات الخطية وهي مرشح وينر لنموذج الانحدار الذاتي AR(3) ومرشح وينر لنموذج الأوساط المتحركة MA(6) فضلا عن المرشح الفرقي Difference Filter تم توليد سلسلة مرشحة جديدة منها.

وأظهرت نتائج التطبيق العملي تفوق مرشح وينر من خلال النموذج AR(3) على بقية المرشحين اذ أعطى أقل قيمة للمعيارين MAPE،MSE.

3- تشابه نتائج معيار MSE مع معيار MAPE في تحديد المرشح الافضل في المعيارين أعطيّا اقل قيمة في المرشح نفسه .

المصادر *References*

[1]. الجابري ، عبد الرحمن محمد مرشد (2000) " دراسة لترشيح بعض السلاسل الزمنية مع تطبيقات "اطروحة دكتوراه، كلية علوم الحاسبات والرياضيات-جامعة الموصل.

[2]. الجميلي ، خلف صالح يوسف (2005) " دراسة بعض المرشحات الخطية مع تطبيق "رسالة ماجستير، كلية العلوم – الجامعة المستنصرية .

[3]. الحمداني، مهند سعد الله داؤد (1996) " مرشح كالمن لبعض النماذج الحركية مع المحاكاة " رسالة ماجستير، كلية الإدارة والاقتصاد-جامعة الموصل.

[4]. الطائي، طلال فاضل (2002) " دراسة كفاءة بعض المرشحات مع المحاكاة " رسالة ماجستير، كلية علوم الحاسبات والرياضيات-جامعة الموصل.

[5]. Edward J-Wegman (1998) “ Time Series Analysis: Theory, Data analysis and Computation ” .

[6]. Harrison, P.J and Stevens, F,C (1976), “ Baysian forecasting ”, J. Ray stat. Soc. 38,205-248.

[7]. Makridakis and Whellwright (1976), “ Forecasting: Methods and Application ”, John wily & sons, inc, second edition U.S.A.

[8]. Priestly, M. B(1981), “ Spectral Analysis and time series ”, London : Academic Press.