

اقتراح استعمال أسلوب الشرائح في التقدير والتنبؤ بمعدلات الخصوبة في العراق للمدة 2012 - 2031

أ.م. د. عمر عبدالحسن علي* عمر فاروق خليل**

المستخلص

غالباً ما تكون الحاجة في البحوث الديموغرافية إلى أدوات أحصائية حديثة تتصف بالمرونة لمواكبة وملامعة نوع البيانات المتوفرة في العراق من حيث مرور البلد بمراحل من فترات الحروب والحصار الاقتصادي و عدم الاستقرار الأمني لفترة من الزمن. ولذا يهدف هذا البحث إلى اقتراح استعمال أسلوب الشرائح اللامعلمي كبديل عن بعض مركبات التحليل ضمن نموذج Lee-Carter الخاص بتقدير معدل الخصوبة التفصيلية كمتغير استجابة في العراق للسنوات (1977 - 2011)، ومن ثم التنبؤ بها للمدة (2012 - 2031). وتم تحقيق هذا الهدف باستعمال أسلوب لامعلمي لتجزئة القيمة المفردة باستعمال المركبات الرئيسية الدالية، ومن ثم تقدير مركبة تأثير السلاسل الزمنية، مما أفضى الى نتائج أفضل.

Keywords: Age-specific fertility rate, Non-parametric, Functional principal components.

Abstract

It is often needed in demographic research to modern statistical tools are flexible and convenient to keep up with the type of data available in Iraq in terms of the passage of the country far from periods of war and economic sanctions and instability of the security for a period of time . So, This research aims to propose the use of style nonparametric splines as a substitute for some of the compounds of analysis within the model Lee-Carter your appreciation rate for fertility detailed variable response in Iraq than the period (1977 - 2011) , and then predict for the period (2012-2031). This goal was achieved using a style nonparametric decomposition of singular value vehicles using the main deltoid, and then estimate the effect of time-series vehicles, leading to better results.

1. المقدمة وخريف البحث

1.1 المقدمة

يعد الوصول الى تحقيق اهداف الالفية الثالثة من أهم ما يرنو إليه المخطط الاقتصادي بمراقبته لنمو السكان في العراق. و لاتخفى أهمية معرفة معدل الخصوبة وأخذها كأحد المتغيرات الديموغرافية المهمة على المهتمين بهذا الشأن خصوصاً وأن العراق يعد من البلدان التي تتميز بمعدل خصوبة عالية، وإن خصوبة السكان في العراق تعد مرتفعة جداً إذا ما قورنت بالبلدان ذات المستوى العالي للخصوبة. يقدر معدل الخصوبة الكلية للمدة (1950-1955) بحدود 7.2 أنخفض الى 6.7 طفل لكل امرأة للمدة (1980-1985) ثم الى 6.0 للمدة (1985-1990) حتى بلغ 5.7 للمدة (1990-1995). ثم أنخفض معدل الخصوبة الكلي

* جامعة بغداد / كلية الإدارة والاقتصاد .

** باحث .

بحث مستل من رسالة ماجستير

مقبول للنشر بتاريخ 2014/3/31

لعام 2011 الى 4.6 طفل لكل امرأة في العراق عموماً^[1]. وتحتاج الخصوبة بهذا مستوى مرتفع الى دراسة أساليب تخفيضها لتحقيق مستوى رفاهية أعلى للأجيال القادمة.

2.1 هدف البحث

يهدف البحث الى تقدير معدل الخصوبة التفصيلي لحد عام (2011) ومن ثم التنبؤ الى حد عام (2031). بأسلوب لامعلمي مرن - الطريقة الشرائحية- كأداة أحصائية تواكب التغيرات المتفاوتة لمعدلات الخصوبة عبر الزمن ويمكن تسميتها مجازاً بأنموذج Lee-Carter اللامعلمي وهو أمر لم يتم استعماله على مستوى الوطن العربي على الأقل لحد الآن، مما يكسب البحث أهمية خاصة.

2. الجانب النظري

1.2 معدلات الخصوبة التفصيلية العمرية Age-Specific Fertility Rates (ASFR)

يتم احتساب معدل الخصوبة التفصيلي لكل فئة عمرية وذلك بقسمة عدد المواليد الحية خلال كل سنة ولكل فئة على عدد الاناث في تلك الفئة ولمنتصف السنة وفقاً على ما يأتي^{[1][3]}:

$$\text{معدل الخصوبة في العمر } x = \frac{\text{عدد المواليد الاحياء خلال السنة للنساء في عمر } x}{\text{عدد النساء في العمر } x \text{ في منتصف العام}} * 1000$$

$$f_x = \frac{B_x}{P_x^f} * 1000 \quad \dots (1)$$

2.2 معدل الخصوبة الكلي^[3] Total Fertility Rate (TFR)

هو مجموع معدلات المواليد التفصيلية حسب العمر لكل سنة من سنوات الانجاب (15-49) اي يحسب لجميع الأعمار (أحادي الأعمار) اي ان $TFR = \int_{x=15}^{49} f(x)dx$ وذلك بضرب معدل كل فئة في العدد (5) وجمع حواصل الضرب لجميع الفئات بمعنى أن: $TFR = 5 \sum_{i=1}^7 f_i$ وهو يعبر عن عدد الأطفال الذين يمكن أنجابهم خلال فترات الانجاب لـ 1000 امرأة خلال سنة معينة، وهو كذلك مقياس لربط جدول المعدلات الخاصة بالعمر وأعطاء رقم واحد بدلا من جدول كامل.

3.2 اسلوب (Lee-Carter) المعلمي

تم استعمال الأنموذج ذاته لعمل تنبؤات بمعدلات الخصوبة العمرية التفصيلية، إذ تم افتراض $f_{x,t}$ تمثل معدل الخصوبة العمرية التفصيلية (ASFR) للعمر x في السنة t ، إذ يمكن صياغة الأنموذج الرياضي له كالآتي^[15]:

$$f_{x,t} = a_x + f_t * b_x + e_{x,t} \quad \dots (2)$$

حيث ان a_x : موجه ثوابت يمثل متوسط قيم الخصوبة العمرية التفصيلية عبر الزمن. $x = 15, 16, 17 \dots 49$ or $x = 15-19, 20-24, \dots, 45-49$; $t = 1, 2, 3, \dots, T$

f_t : يمثل مؤشر الخصوبة في الزمن t ويعرف بأنه يفسر انحراف معدل الخصوبة الكلية (TFR) في الفترة t من المتوسط العام لمعدل البعيد الامد A إذ ان $A = \sum_{x=15}^{49} \hat{a}_x$ ^[15].

b_x : موجه ثوابت يمثل مركبة العمر يصف السرعة النسبية للتغير في الخصوبة عند كل فئة عمرية. $e_{x,t}$: يمثل الخطأ العشوائي في العمر x في السنة t ، مع افتراض انه يتوزع طبيعياً بمتوسط (0) وتباين (σ^2) .

والصيغة النصف اللوغاريتمية لنفس الأنموذج هي ايضا يمكن اعتبارها أنموذج كالآتي:

$$\ln(f_{x,t}) = a_x + f_t * b_x + e_{x,t} \quad \dots (3)$$

حيث ان $\ln f_{x,t}$: يمثل اللوغاريتم الطبيعي لمعدل الخصوبة العمرية التفصيلية للعمر او الفئة العمرية x في السنة t .

كما يمكن التعبير عن الأنموذج (2) بالصيغة الآتية^[15]:

$$TFR_t = A + f_t + E_t \quad \dots (4)$$

حيث ان: $E_t = \sum_{t=1977}^{2011} e_{x,t}$.

وتأسيساً على ماتم ذكره في أعلاه فستكون المرحلة الاولى لتقدير موجه المعلمات $(\hat{b}_x)$ كالآتي:

$$\hat{b}_x^{(1)} = \frac{\hat{b}_x^{SVD}}{\sum_{x=1}^A \hat{b}_x^{SVD}} \quad \dots (5)$$

أذن:

$\hat{b}_x^{SVD}$: العمود الأول من المصفوفة U وهو (u_{x1}) .

$\sum_{x=1}^A \hat{b}_x^{SVD}$: مجموع قيم العمود (u_{x1}) .

أما المرحلة الأولى لتقدير موجه المعلمات (f_t) فيكون:

$$\hat{f}_t^{(1)} = \hat{f}_t^{SVD} * \sigma_1 * \sum_{x=1}^A \hat{b}_x^{SVD} \quad \dots (6)$$

حيث أن:

$\hat{f}_t^{SVD}$: العمود الأول من المصفوفة V وهو (v_{t1}) .

σ_1 : العنصر الأول من المصفوفة القطرية (Σ) .

لذلك سيكون التقدير النهائي للمعلمات كالآتي:

$$\hat{b}_x = \hat{b}_x^{(1)} * c \quad \dots (7)$$

$$\hat{f}_t = \frac{\hat{f}_t^{(1)}}{c} \quad \dots (8)$$

c : ثابت يحقق وحدانية الحل.

أما فيما يخص تقدير موجه الثوابت a_x والذي يمثل المتوسط الحسابي لمعدلات لوغاريتمات الخصوبة العمرية التفصيلية وفق الصيغة الآتية [9]:

$$\hat{a}_x = \frac{\sum_{t=1}^T \ln f_{x,t}}{T} \quad \dots (9)$$

4.2 التقدير اللامعلمي

وبإعادة كتابة الصيغة (2) ليلانم الغرض من استعمال أنموذج (L-C) يتم الحصول على الآتي [14]:

$$y_t(x_i) = g_t(x_i) + \epsilon_{t,i} \quad \dots (10)$$

أذن $y_t(x)$ يمثل لوغاريتم معدل الخصوبة للعمر x في السنة t . وان $g_t(x)$ دالة تمهيد وبخطا عند نقاط متقطعة لـ x ، ويعرف العمر بمجموعات خمسية (5-year age groups) كما مر ذكره في التحليل المعلمي في الفقرات السابقة. ويتميز التحليل اللامعلمي باستعمال إحدى أساليب التمهيد بتوافر حصانة (Robustness) كافية لتقدير (ASFR)، وسيتم عمل ذلك عن طريق توسيع منهجية (L-C) لتقليل العشوائية، وتلافي مشاكل البيانات المجمعة ضمن فترات عمرية، من جهة، ووضع نمط جديد للمركبات الرئيسية لتلافي صعوبات السنوات المطبقة من جهة أخرى. إضافة إلى ذلك استعمال أكثر من المركبة الرئيسية الأولى خلافاً لأنموذج (L-C) والذي يعتمد على المركبة الرئيسية الأولى فقط.

ويمكن تلخيص الأسلوب اللامعلمي المستعمل كالآتي [8][14]:

1. تمهيد البيانات لكل t باستعمال طريقة تمهيد لامعلمية لتقدير $g_t(x)$.

2. تجزئة المنحنيات المقدرة عن طريق مفكوك دالة أساس باستعمال الصيغة الآتية:

$$g_t(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x) \quad \dots (11)$$

أذن $\mu(x)$ تمثل مقياس الموضع لـ $g_t(x)$ ، $\{\phi_k(x)\}$ هي مجموعة دوال أساس (orthonormal) وان $e_t(x)$ يتوزع طبيعياً بوسط صفر وتباين σ^2 .

وهو ما يكافئ تحليل أنموذج (L-C) بالأسلوب المعلمي.

3. تقدير نماذج سلاسل زمنية أحادية المتغيرات لكل معاملات $(\beta_{t,k})$.

4. تنبؤ المعاملات $(\beta_{t,k})$ باستعمال نماذج السلاسل الزمنية المقدرة أعلاه.

5. استعمال معاملات التنبؤ من الخطوة (2) لاستحصا التنبؤ $g_t(x)$.

6. استعمال التباينات المقدرة لحدود الأخطاء من الخطوة (2) و (1) في احتساب حدود التقدير لغرض التنبؤات.

1.4.2 استعمال أسلوب B-Spline [11][12]

تتضمن B-Spline على متعددة حدود قطعية piecewise polynomials. وتمتاز بأنها ملائمة جداً لتكون دالة أساس للاتحاد اللامعلمي البسيط، أذن ان التوافق الخطية من الدرجة الثالثة لـ B-Spline تعطي منحنى ممهد.

أذن قام (De Boor, 1978) بوضع خوارزمية لحساب B-Spline من أي درجة من B-Spline للدرجة الدنيا. لأن الدرجة الصفري لـ B-Spline تكون ثابتة فقط على فترة واحدة بين عقدتين. وهذه

الخوارزمية تعمل في حالة المسافات بين العقد متساوية أو غير متساوية لكنها تعمل بسهولة في حالة تساوي المسافات.

لتكن $B_j(x; q)$ تعرف بأنها القيمة الناتجة عند x من الشريحة B j^{th} من الدرجة q ولشبكة معطاة من العقد متساوية المسافة (equidistant grid).

2.4.2 المركبات الرئيسية الموزونة (Weighted principal components)

نفرض بالامكان كتابة أي دالة ممهدة معدلة $\hat{g}_t^*(x)$ بشكل أساس موسع (basis expansion) كالآتي [13][14]:

$$\hat{g}_t^*(x) = \sum_{j=1}^m \gamma_{t,j} \xi_j(x) \quad \dots$$

(12)

ولتكن $A = (\gamma_{t,j})$ مصفوفة معاملات من درجة $n \times m$. الأساس الموسع أعلاه يمكن الحصول عليه طبيعياً إذا كانت حساب $\hat{g}_t(x)$ يتحقق من خلال استعمال شرائح الانحدار، والان نفرض J مصفوفة من درجة $m \times m$ وعناصرها $J = \int \xi_i(x) \xi_k(x) dx$ وبأيجاد تجزئة كولسكي (decomposition Choleski) للمصفوفة $J = U'U$. إذن

$$\varphi_k(x) = (U^{-1}z^{(k)})' \xi(x) \quad \dots (13)$$

أذا أن $z^{(k)}$ يمثل متجه الجذور المميزة الطبيعية لـ $(U^{-1})'JSJ'U^{-1}$ ، $S = (n-1)^{-1}A'W^2A$ ، حيث $W = \text{diagonal}(w_1, \dots, w_n)$ وأن $\xi(x) = [\xi_1(x), \dots, \xi_m(x)]'$. إذا Φ مصفوفة من درجة $p \times (n-1)$ مع (i, k) th قيمة $\varphi_k(x_i)$ ، و F هي مصفوفة من درجة $n \times p$ للعنصر (t, i) th المنحنيات المقدرة المعدلة $\hat{g}_t^*(x)$ إذن $\hat{\beta}_{t,k}$ هو العنصر (t, k) th من $Q = F\Phi$.

وبتبسيط العمليات الحسابية لتقطيع (discretize) دوال التمهيد $\hat{g}_t^*(x)$ على شبكة q من قيم المسافات المتساوية $\{x_1^*, \dots, x_q^*\}$ المعرفة على الحيز $[x_1, x_p]$. المصفوفة الناتجة G من الدرجة $n \times q$ ولتكن $G^* = WG$ حيث $W = \text{diagonal}(w_1, \dots, w_n)$. فأن طريقة تحليل القيمة المفردة (SVD) للمصفوفة G^* تعطي $G^* = \Psi \Lambda V'$ حيث $\hat{\varphi}_k(x_j^*)$ هو العنصر (j, k) th من المصفوفة Ψ . إذا كانت $B = G\Psi$ ، فأن $\hat{\beta}_{t,k}$ هي العنصر (t, k) th من المصفوفة B ، أما بالنسبة لباقي قيم $\varphi_k(x)$ بالامكان حسابها باستعمال الاستكمال الخطي (linear interpolation). ويتم تقدير نماذج السلاسل الزمنية لـ $\hat{\beta}_{t,k}$ سيكون باستعمال نماذج $(B-J)$.

5.2 التنبؤ بـ (ASFR) باستعمال الأسلوب اللامعلمي (Forecasting of ASFR by Using a Nonparametric smoothing)

الآن وبعد أن تم الحصول على $\hat{\beta}_{t,k}$ و $\hat{\varphi}_k(x)$ صار بالامكان التنبؤ بمعدلات الخصوبة (ASFR) وذلك بدمج الصيغتين (3) و (4) لينتج الصيغة الآتية [8][14]:

$$y_t(x_i) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \varphi_k(x) + e_t(x) + \epsilon_{t,i} \quad \dots (14)$$

وبأخذ التوقع لطرفي الصيغة أعلاه وبتعويض كل من $\hat{\beta}_{t,k}$ و $\hat{\varphi}_k(x)$ فضلاً عن $\hat{\mu}(x)$ في الصيغة (14) نحصل على الآتي:

$$\hat{y}_t(x_i) = \hat{\mu}(x) + \sum_{k=1}^K \hat{\beta}_{t,k} \hat{\varphi}_k(x) \quad \dots (15)$$

وللحصول على h من الخطوات التنبؤية ستكون الصيغة (15) بالشكل الآتي:

$$\hat{y}_{t,h}(x_i) = \hat{\mu}(x) + \sum_{k=1}^K \hat{\beta}_{t,k,h} \hat{\varphi}_k(x) \quad \dots (16)$$

6.2 معايير جودة النموذج Criteria of Model fitting

هناك العديد من المعايير لقياس جودة توفيق أنموذج (Lee-Carter) لمعرفة نسبة تفسير الأنموذج للانحرافات الكلية والتي تعتمد على أخطاء الأنموذج وفق الصيغة:

$$\hat{\mathcal{E}}_{x,t} = (\text{Inf}_{x,t} - \hat{a}_x + \hat{b}_x * \hat{f}_t) \quad \dots (17)$$

ويتم استعمال معامل التحديد لاحتساب القوة التفسيرية للتباين الكلي بين معدلات الخصوبة العمرية وبالصيغة الآتية:

$$R^2 = \frac{\sigma_1^2}{\sum_{t=1}^T \sigma_t^2} \quad \dots$$

(18)

اما المقاييس التي تستخدم لقياس جودة التقدير والتي تعتمد على استخراج أخطاء نموذج (L-C) فهي:

1. متوسط القيمة المطلقة للخطأ [2][1] Mean Absolute Error (MAE)

$$MAE = \frac{1}{A} \sum_{x=15}^{49} \sum_{t=1}^T |\hat{\varepsilon}_{x,t}^*| \quad \dots (19)$$

2. متوسط مربعات الخطأ Mean Square Error

$$MSE = \frac{1}{A} \sum_{x=15}^{49} \sum_{t=1}^T (\hat{\varepsilon}_{x,t}^*)^2 \quad \dots (20)$$

3. متوسط النسبة المطلقة للخطأ Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{A} \sum_{x=15}^{49} \sum_{t=1}^T \left| \frac{\hat{\varepsilon}_{x,t}}{f_{x,t}} \right| \quad \dots (21)$$

4. نسبة التباين المئوية التجميعية (Cumulative Percentage Variance: CPV)

يبين نسبة المساهمة التجميعية للمركبات الرئيسية في تفسير التباين الكلي للمركبات، واما قيمته المثلى فليس هنالك اتفاق الى حد الان عليها سوى بعض الاراء للباحثين في هذا المجال فعلى سبيل المثال هناك رأي بأن تكون النسبة المتراكمة من التباين الكلي مايعادل 70% الى 80% للمركبات الاولى من المركبات الرئيسية، وكلما ازداد كان افضل. واما صيغتها الرياضية كالآتي [16]:

$$C. P. V. = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^K \lambda_i} * 100\% \quad \dots (22)$$

3. الجانب التطبيقي

تم جمع البيانات من وزارة التخطيط - الجهاز المركزي للإحصاء حيث تم الاستناد الى بيانات التعدادات السكانية في العراق للسنوات (1977-1987-1997) بالإضافة الى المسح العنقودي متعدد المؤشرات العراق 2011 (MICS-4) وتضمنت البيانات ماهو خاص بالخصوبة وهي (عدد المواليد الحية للنساء في سن الحمل (15-49) في منتصف السنة السابقة للتعداد وعدد النساء في سن الحمل (15-49) سنة في منتصف السنة)، وتم استعمال برمجيات عديدة وتوظيفها في جانبي التحليل المعلمي واللامعلمي وكما في الملحق.

1.3 تقدير نموذج Lee – Carter المعلمي

يتم استخراج موجه تقدير المعلمات ($\hat{a}_x$) من المصفوفة (F) وفق الصيغة (9)، وتمثل قيم هذا الموجه المتوسطات الحسابية للوغاريتمات معدلات الخصوبة العمرية للفترة (1977-2011) على المستوى والبيئي (حضر، ريف، العراق).

جدول 1

القيم المقدرة للمعلمة ($\hat{a}_x$)

$\hat{a}_x$			الفئات
العراق	ريف	حضر	
-1.225	-1.1391	-1.2961	15-19
0.0317	0.043	-0.1341	20-24
0.1936	0.3304	0.1056	25-29
0.1147	0.2785	0.0279	30-34
-0.1356	0.0676	-0.2277	35-39
-0.8284	-0.5275	-0.9956	40-44
-1.9667	-1.4655	-2.2941	45-49

ويلاحظ من الجدول (1) اعلاه ان أعلى قيم للمعلمة $\hat{a}_x$ للفئة العمرية الثالثة (25-29).

يتبين من نتائج الجدول اعلاه بأن متوسطات لوغاريتمات معدلات الخصوبة دائماً تبدأ منخفضة في الفئة الاولى وسرعان ما ترتفع لتصل الى أعلى مستوياتها في الفئة الثالثة ومن ثم تبدأ بالانخفاض لتصل الى أدنى مستوياتها في الفئة الاخيرة وهي بذلك تتطابق مع نمط الخصوبة في العالم بشكل عام [5].
تجزأ المصفوفة F الى مركباتها كالآتي:

$$F_{7*35} = U_{7*7} \Sigma_{7*35} V'_{35*35}$$

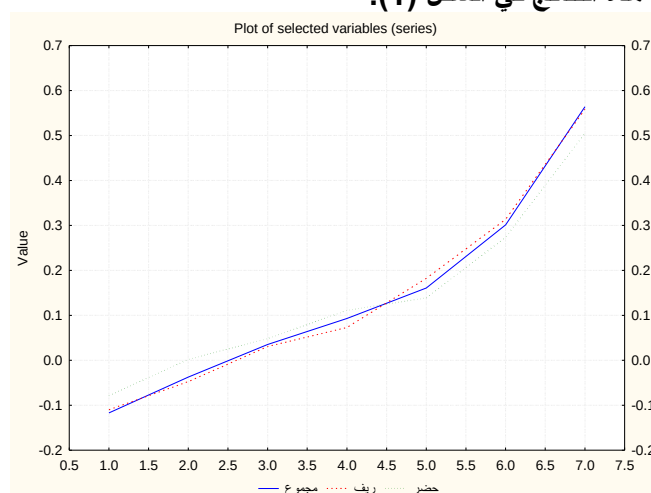
وباختيار العمود الاول من المصفوفة (U) والقيمة القطرية الاولى من المصفوفة (Σ) والعمود الاول من المصفوفة (V') نحصل على التقديرات الاولى لكل من (b_x) و (f_i) كما في المعادلتين (5) و (6) على التوالي. والنتائج في الجدول (2) وكما يأتي:

الجدول 2

القيم المقدرة للمعلمة (b_x) لـ (حضر، ريف، العراق)

$\hat{b}_x$			الفئات
العراق	ريف	حضر	
-0.117	-0.1099	-0.0786	15-19
-0.0373	-0.0475	0.0019	20-24
0.0353	0.0304	0.0485	25-29
0.093	0.073	0.1109	30-34
0.1607	0.182	0.1388	35-39
0.3012	0.3135	0.2736	40-44
0.5641	0.5585	0.5049	45-49

ويمكن توضيح هذه النتائج في الشكل (1).



شكل 1

متجهات القيم المقدرة لمعاملات ($\hat{b}_x$)

يتبين من الجدول (2) والشكل (1) تزايد رتيب في مركبة العمر ($\hat{b}_x$) ووجود فروق طفيفة بين المركبات المقدرة بالنسبة للريف والحضر.

الجدول 3

قيم العناصر القطرية للمصفوفة (Σ)

Σ		
العراق	ريف	حضر
3.8800	3.3787	3.9226
1.7025	1.3676	1.9584
0.4790	0.2194	0.4094
0.2269	0.0532	0.0691
0.0423	0.0287	0.0539
0.0213	0.0251	0.0264
0.0153	0.0075	0.0151

الجدول 4

مقياس R^2

البيئة	العراق	حضر	ريف
R^2	82.55%	79.32%	85.58%

الجدول 5

مقاييس الخطأ

MAE	MAPE	MSE	
3.3755	3.3411	0.5731	حضر
1.9741	1.9549	0.2747	ريف
2.752	2.7391	0.4546	العراق

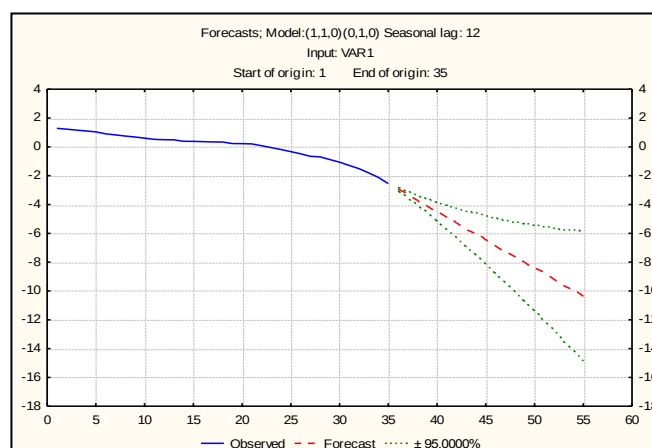
2.3 تقدير نموذج الانحدار الذاتي من الدرجة الاولى $AR(1)$

بعد اجراء عملية التشخيص للسلسلة الزمنية تبين بأن أفضل نموذج سلسلة زمنية هو نموذج الانحدار الذاتي من الدرجة الاولى $AR(1)$ بعد أخذ الفرقين الاول والثاني للسلسلة لكل من (الحضر والمجموع)، إما بالنسبة للريف فإن أفضل نموذج كان $MA(1)$ أيضاً بعد أخذ الفرقين الاول والثاني للسلسلة الاصلية للتغلب على مشكلة الاتجاه وبالاستناد على معيار تشخيص النموذج AIC.

جدول 6

القيم التنبؤية لدليل الخصوبة $(\hat{f}_t)$ في العراق

Std.Err.	Upper	Lower	Forecast	year
0.057	-2.812	-3.045	-2.928	2012
0.107	-3.107	-3.543	-3.325	2013
0.171	-3.365	-4.063	-3.714	2014
0.243	-3.610	-4.601	-4.106	2015
0.324	-3.837	-5.156	-4.497	2016
0.411	-4.050	-5.726	-4.888	2017
0.506	-4.249	-6.309	-5.279	2018
0.606	-4.436	-6.905	-5.670	2019
0.713	-4.610	-7.513	-6.061	2020
0.825	-4.773	-8.132	-6.452	2021
0.942	-4.925	-8.762	-6.844	2022
1.064	-5.067	-9.403	-7.235	2023
1.192	-5.199	-10.053	-7.626	2024
1.323	-5.321	-10.713	-8.017	2025
1.460	-5.434	-11.382	-8.408	2026
1.601	-5.538	-12.060	-8.799	2027
1.746	-5.634	-12.747	-9.190	2028
1.895	-5.721	-13.442	-9.582	2029
2.049	-5.800	-14.146	-9.973	2030
2.206	-5.871	-14.857	-10.364	2031



شكل 2

التنبؤ بدليل الخصوبة $(\hat{f}_t)$ وبمستوى ثقة 95% لمجموع العراق

يلاحظ من الجدول (6) والشكل (2)، بأن متجه الخصوبة يبدأ مستقر ومن ثم يبدأ بالانخفاض التدريجي والقيم التنبؤية توضح استمرارية الانخفاض بدليل الخصوبة مما يدل على ان العراق سيشهد انخفاض في معدلات الخصوبة في العقدين القادمين مما يسهل له عملية التحول الديموغرافي.

جدول 7

القيم التنبؤية ASFR بطريقة (L-C) في العراق لكل امرأة ماعدا TFR

العراق					السنة الفئات
2031	2026	2021	2016	2012	
198	157	125	99	83	15-19
304	282	263	244	230	20-24
168	180	193	207	219	25-29
86	103	123	148	171	30-34
33	45	62	85	109	35-39
4	7	13	23	36	40-44
0	0	1	2	5	45-49
4.0	3.9	3.9	4.0	4.3	TFR

جدول (7) يوضح التنبؤ بعدد الولادات للمرأة العراقية حسب الفئات العمرية حيث ان العدد (83) للفئة العمرية (15-19) للعراق في سنة 2012 يمثل اعداد الولادات لكل امرأة في هذه الفئة وهكذا بالنسبة لبقية الفئات والسنوات، أما فيما يخص TFR فيمثل (4.3) عدد الولادات الحية التي تنجبها المرأة طيلة فترة حياتها الانجابية وهكذا بالنسبة لبقية قيم الجدول.

بعد أن تم تقدير كل من موجه متوسطات لوغاريتيمات معدلات الخصوبة ($\hat{a}_x$) وفق المعادلة (9) وموجه مركبة العمر ($\hat{b}_x$) وفق المعادلة (7) فضلاً عن مركبة الزمن ($\hat{f}_t$) نعوض هذه التقديرات في النموذج (3) للحصول على تقديرات معدلات الخصوبة العمرية والتي سوف تكون بشكل مصفوفة ذات ابعاد (35*7) وبعدها يتم طرح معدلات الخصوبة المقدرة من معدلات الخصوبة الحقيقية للحصول على الأخطاء التقديرية وفق المعادلة (19). اما في ما يخص القيم التنبؤية لمعدلات الخصوبة الناتجة من تقدير أنموذج Lee-Carter (7) الجدول (7) أعلاه تبين القيم التنبؤية لمعدلات الخصوبة في العراق.

الجدول 8

تقدير معدلات الخصوبة التفصيلية العمرية ASFR لكل (1000) امرأة*

معدل الخصوبة التفصيلية العمرية ASFR				الفئات
2011	1997	1987	1977	
82	45	42	86	15-19
214	214	155	263	20-24
236	224	215	335	25-29
187	217	222	294	30-34
131	168	185	242	35-39
45	89	100	142	40-44
7	31	37	58	45-49

جدول 9

تقدير معدلات الخصوبة التفصيلية العمرية بأسلوب (Lee-Carter) لكل (1000) امرأة

معدل الخصوبة التفصيلية العمرية ASFR				الفئات
2011	1997	1987	1977	
79	57	55	51	15-19
227	205	203	197	20-24
222	244	247	254	25-29
177	229	235	253	30-34
116	180	190	215	35-39
41	93	102	129	40-44
7	31	37	58	45-49

3.3 نتائج تحليل الأنموذج اللامعلمي

تم تطبيق الصيغ الرياضية الواردة في الفقرة (4.2) أعلاه لاحتساب مقدرات الأنموذج اللامعلمي المذكور انفاً في المعادلة (10)، اذ ان الجدول (10) أدناه، يبين تقديرات معلمات المتوسط العام (μ_x) اذ

يمثل قيم هذا الموجه المتوسطات الحسابية للوغاريتمات معدلات الخصوبة العمرية للمدة (1977-2011) على المستوى الوطني.

جدول 10

القيم المقدرة للمتوسط العام (μ_x) للعراق

μ_x	الفئات
10.3172	15-19
18.6080	20-24
20.0154	25-29
19.3126	30-34
17.2585	35-39
12.5465	40-44
7.1896	45-49

يتبين من الجدول (10) بأن قيم المتوسط العام تبدأ منخفضة وتستمر بالارتفاع حتى تصل الى اعلى مستوى لها في الفئة العمرية الثالثة (25-29) ومن ثم الانخفاض لتصل الى ادنى مستوياتها في الفئة العمرية (45-49) وهذا ما يظهره الشكل (3) ويظهر الشكل الاول المتوسط العام وهو يمثل بذلك منحنى الخصوبة العام.

جدول 11

القيم المقدرة للمعاملات $\varphi_k(x)$ للعراق

$\varphi_3(x)$	$\varphi_2(x)$	$\varphi_1(x)$	الفئات
0.6567	0.6464	-0.0710	15-19
-0.4609	0.5543	-0.0203	20-24
0.3286	0.4621	0.2058	25-29
0.1494	0.1446	0.3031	30-34
0.1603	0.0130	0.4284	35-39
-0.1100	-0.0918	0.5622	40-44
-0.4338	-0.1785	0.6006	45-49

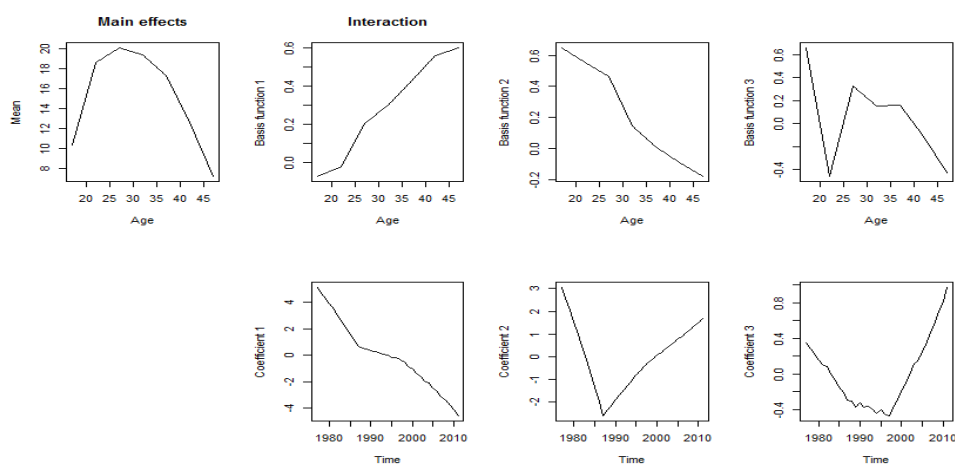
اما الجدول (11) يبين القيم المقدرة لمركبات العمر ويظهر في الجدول بأحتوانه على ثلاثة مركبات رئيسية للعمر وهذه ميزة يمتاز بها النموذج اللامعلمي على النموذج (L-C) لان الأخير يعتمد على المركبة الرئيسية الاولى في تقديره بافتراض ان هذه المركبة تحمل اكبر قدر ممكن من الانحرافات، لكن في النموذج اللامعلمي يسمح لنا بأخذ اكبر قدر من دوال الاساس (المركبات الرئيسية) للحصول على تقدير ادق لمعدلات الخصوبة وبالتالي الحصول على تنبؤات دقيقة، ويتضح في كل من الجدول (11) والشكل (3) بأن المركبة الرئيسية الاولى φ_1 هي عبارة عن دالة متزايدة رتيبة تبدأ منخفضة وتستمر بالارتفاع لتصل الى اعلى مستوياتها في الفئة العمرية (45-49) بعكس المركبة الرئيسية الثانية φ_2 اذ تمثل دالة رتيبة متناقصة وكما هو واضح، بينما المركبة الثالثة φ_3 تكون مضطربة القيم اذ تكون مرتفعة عند فئة عمرية محددة وتنخفض في الفئة التالية بشكل سريع.

4.3 نتائج تقدير تأثير الزمن ($\beta_{t,k}$)

يتم الان ايجاد تقديرات تأثير $\beta_{t,k}$ الزمن للنموذج اللامعلمي بعد ان تم الحصول على المتوسط العام ومركبات العمر ويوضح الجدول (12) ثلاثة مركبات (الزمن) رئيسية حيث سيتم اعتبار هذه المركبات الثلاثة فيما بعد سلاسل زمنية احادية المتغيرات ومن الجدول (12) والشكل (3) يتبين بأن المركبات الرئيسية الثلاثة $\beta_{t,k}$ وعند عرضها كسلسلة زمنية يظهر الاضطراب الواضح فيها بالاضافة بعض القيم الشاذة التي لابد من معالجتها عند تقدير السلاسل الزمنية وأيجاد افضل النماذج المعنوية والتي تعطي دقة عالية في نتائج التقديرات.

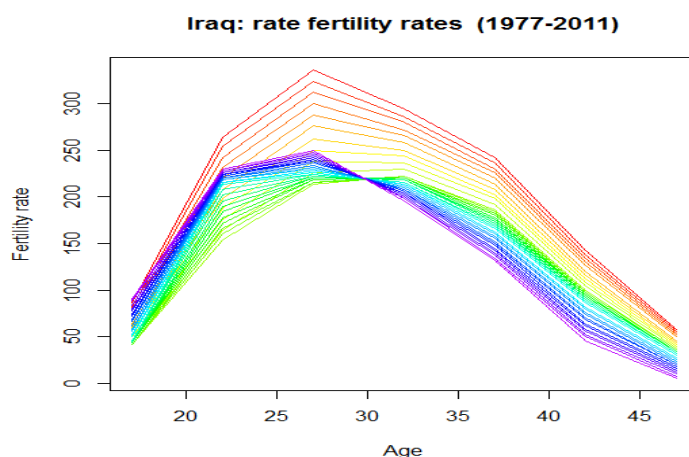
جدول (12)
القيم المقدرة للمعاملات $(\beta_{t,k})$ للعراق

$\beta_{t,3}$	$\beta_{t,2}$	$\beta_{t,1}$	السنة
0.3546	3.0427	5.0924	1977
0.2841	2.5755	4.6844	1978
0.2350	2.0563	4.2929	1979
0.1545	1.5339	3.9005	1980
0.1031	0.9850	3.4950	1981
0.0867	0.4320	2.9372	1982
-0.0004	-0.1293	2.4736	1983
-0.0651	-0.6693	2.0384	1984
-0.1462	-1.3445	1.6039	1985
-0.2034	-1.9932	1.1515	1986
-0.2896	-2.6132	0.6613	1987
-0.3032	-2.3626	0.5517	1988
-0.3698	-2.1759	0.4904	1989
-0.3223	-1.8749	0.3253	1990
-0.3723	-1.6821	0.2897	1991
-0.3671	-1.4458	0.2082	1992
-0.4033	-1.2579	0.0532	1993
-0.4402	-1.0490	0.0328	1994
-0.3998	-0.7944	-0.1535	1995
-0.4514	-0.6071	-0.2004	1996
-0.4695	-0.4172	-0.3324	1997
-0.3842	-0.2317	-0.5361	1998
-0.2857	-0.1171	-0.8438	1999
-0.2005	0.0639	-1.0568	2000
-0.0977	0.1665	-1.3734	2001
-0.0109	0.3442	-1.5975	2002
0.1185	0.4958	-1.9422	2003
0.1419	0.6164	-2.1052	2004
0.2547	0.7854	-2.4757	2005
0.3342	0.8813	-2.7281	2006
0.4660	1.0779	-3.1091	2007
0.5558	1.1646	-3.3755	2008
0.7005	1.3649	-3.7900	2009
0.8222	1.5000	-4.1000	2010
0.9708	1.6789	-4.5625	2011



شكل 3
متجهات القيم المقدرة للمعاملات $(\beta_{t,k}, \mu_x, \varphi_k)$

يتضح من الجدول الشكل (3) الاختلاف الكبير والواضح بين مركبات الزمن الرئيسية الثلاثة فضلاً عن مركبات العمر وهذا بدوره يتطلب معالجة كل مركبة على حده فالمركبة الاولى تعاني من مشكلة الاتجاه والتي سيتم معالجتها بأخذ الفروق للسلسلة الاصلية حتى تصبح مستقرة أما بالنسبة للمركبتين الثانية والثالثة كلتيهما تعانيان من عدم الاستقرارية ايضاً لذا يتوجب معالجة هذه المشكلة.



شكل 4

معدلات الخصوبة التفصيلية العمرية (ASFR) للمدة (2011-1977)

جدول 13

مقاييس الخطأ

Averages across ages:			
ME	MSE	MPE	MAPE
-0.00002	0.1109	0.0041	0.02114
Averages across years:			
IE	ISE	IPE	IAPE
-0.00088	2.66253	0.0822	0.5179

الان وبعد ان تم الحصول على تقديرات $(\beta_{t,k}, \mu_x, \phi_k)$ للأنموذج اللامعلمي بالأسلوب الشرائحي، وبتعويضها في الأنموذج (15) للحصول على معدلات الخصوبة التفصيلية العمرية كما في الجدول الاتي:

جدول 14

تقدير معدلات الخصوبة التفصيلية العمرية (ASFR) بالأسلوب اللامعلمي لكل (1000) امرأة للعراق

معدل الخصوبة التفصيلية العمرية ASFR				الفئات
1977	1987	1997	2011	
61	42	49	62	15-19
100	86	93	96	20-24
113	94	98	101	25-29
107	95	95	92	30-34
98	87	85	77	35-39
75	66	62	49	40-44
48	41	36	19	45-49

5.3 نتائج أنموذج الاحدار الذاتي AR(P)

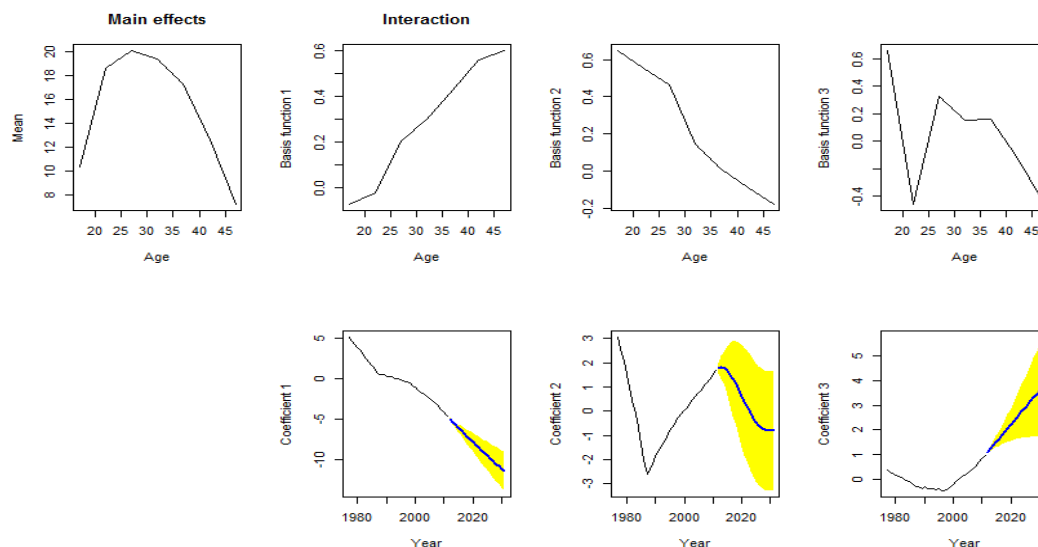
الان وبعد ان تم الحصول على تقديرات $(\beta_{t,k}, \mu_x, \phi_k)$ تأتي المرحلة قبل الاخيرة وهي التنبؤ بمركبة الزمن المقدرة $\beta_{t,k}$ والتي تتطلب ايجاد افضل أنموذج سلاسل زمنية ليتم توظيفه في مرحلة التنبؤ بدليل الخصوبة $(\hat{\beta}_{t,k})$ ، وبعد اجراء عملية التشخيص للسلسلة الزمنية تبين بأن أفضل أنموذج سلسلة زمنية هو أنموذج الاحدار الذاتي من الدرجة الثانية AR(2) للمركبة الزمنية الاولى $\hat{\beta}_{t,1}$ بعد أخذ الفرقين الاول للسلسلة إما بالنسبة للمركبة الثانية $\hat{\beta}_{t,2}$ فإن افضل أنموذج كان AR(2) أيضاً، اما المركبة الثالثة $\hat{\beta}_{t,3}$ فإن افضل أنموذج سلسلة زمنية هو AR(1) وبلاستناد على معيار تشخيص الأنموذج AIC وفيما يلي جدول (15) يبين تقديرات معلمات النماذج .

جدول 15
قديرات معلمات النماذج المعنوية للعراق

تقدير المعلمة					الأنموذج	مركبات الزمن
AIC	MSE	AR(1)	AR(2)	drift		
-50.3	0.010	0.3452 S.E=0.15	0.4541 S.E=0.15	-0.3164 SE=0.07	AR(2,1,0) with drift	$\beta_{t,1}$
-19.45	0.023	1.8737 S.E=0.06	-0.9092 S.E=0.06	--	AR(2,0,0)	$\beta_{t,2}$
-110.98	0.002	-0.6484 S.E=0.12	--	--	AR(1,2,0)	$\beta_{t,3}$

جدول 16
القيم التنبؤية لدليل الخصوبة ($\hat{\beta}_{t,k}$) للعراق

$\hat{\beta}_{t,3}$			$\hat{\beta}_{t,2}$			$\hat{\beta}_{t,1}$			السنة
Upper	Lower	Forecast	Upper	Lower	Forecast	Upper	Lower	Forecast	
1.16242	1.04142	1.10192	2.00162	1.5622	1.78191	-4.7807	-5.0721	-4.9264	2012
1.3461	1.14266	1.24438	2.27895	1.34569	1.81232	-5.0814	-5.5698	-5.3256	2013
1.54337	1.21562	1.37949	2.51347	1.03774	1.77561	-5.321	-6.0633	-5.6921	2014
1.74749	1.29125	1.51937	2.69396	0.66439	1.67918	-5.566	-6.5608	-6.0634	2015
1.95928	1.35304	1.65616	2.81525	0.2485	1.53188	-5.7913	-7.0517	-6.4215	2016
2.17763	1.41227	1.79495	2.87667	-0.1896	1.34356	-6.0145	-7.5399	-6.7772	2017
2.40197	1.46291	1.93244	2.88106	-0.6318	1.12463	-6.2304	-8.0218	-7.1261	2018
2.6321	1.50946	2.07078	2.83407	-1.0628	0.88565	-6.4443	-8.4988	-7.4716	2019
2.86747	1.54967	2.20857	2.74356	-1.4697	0.63691	-6.6557	-8.9698	-7.8127	2020
3.10797	1.58545	2.34671	2.61902	-1.8427	0.38815	-6.8664	-9.4354	-8.1509	2021
3.35324	1.61601	2.48463	2.471	-2.1746	0.1482	-7.0766	-9.8955	-8.4861	2022
3.60316	1.64222	2.62269	2.31066	-2.4611	-0.0752	-7.2873	-10.35	-8.8188	2023
3.85748	1.66383	2.76065	2.14914	-2.7005	-0.2757	-7.4986	-10.8	-9.1494	2024
4.11609	1.68128	2.89868	1.99711	-2.8934	-0.4482	-7.711	-11.245	-9.4781	2025
4.3788	1.69455	3.03667	1.86419	-3.0423	-0.5891	-7.9248	-11.686	-9.8052	2026
4.64551	1.70387	3.17469	1.75842	-3.1509	-0.6963	-8.14	-12.122	-10.131	2027
4.91607	1.70929	3.31268	1.6858	-3.2238	-0.769	-8.3568	-12.554	-10.455	2028
5.19041	1.71098	3.45069	1.6499	-3.2655	-0.8078	-8.5753	-12.982	-10.779	2029
5.46839	1.709	3.58869	1.65165	-3.2805	-0.8144	-8.7956	-13.407	-11.101	2030
5.74994	1.70346	3.7267	1.68944	-3.2724	-0.7915	-9.0177	-13.828	-11.423	2031



شكل 5

تنبؤات متجهات القيم المقدرة للمعلمات ($\beta_{t,k}, \mu_x, \phi_k$)
يتبين من الشكل (5) بأن القيم التنبؤية لمركبات الزمن تتجه نحو الانخفاض على المستوى المتوسط للمركبتين الأولى والثانية وهذا بدوره يؤثر على القيم التنبؤية لمعدلات الخصوبة التفصيلية (ASFR) وهذا ما سيوضحه جدول (17).

جدول 17

القيم التنبؤية لمعدلات الخصوبة التفصيلية (ASFR) في العراق لكل 1000 امرأة
للاسلوب اللامعلمي للمدة (2031-2012)

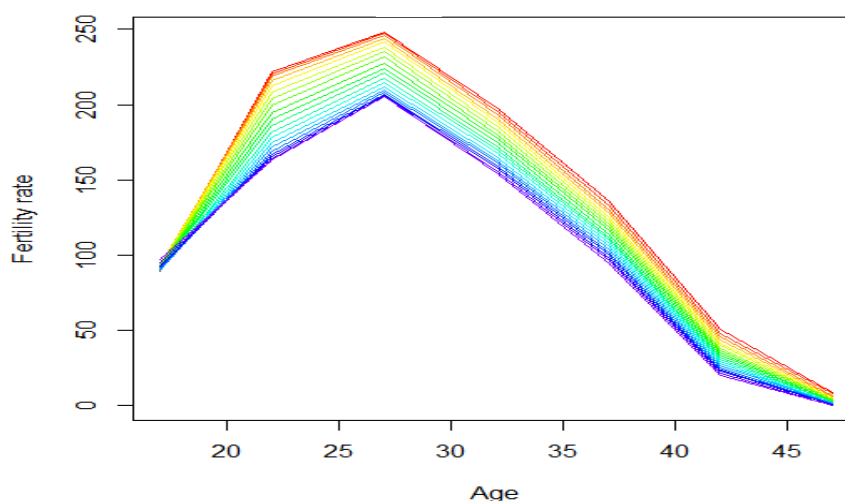
السنة	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
15-19	89	91	92	93	93	93	93	92	92	91
20-24	222	221	219	216	212	209	204	200	195	190
25-29	248	248	246	244	242	238	235	231	228	224
30-34	198	196	194	191	189	186	183	181	178	175
35-39	136	133	131	128	126	123	121	119	117	114
40-44	50	48	46	44	42	40	38	36	35	33
45-49	9	8	7	6	5	5	4	4	3	3
TFR	4.8	4.7	4.7	4.6	4.5	4.5	4.4	4.3	4.2	4.2

تابع جدول 17

السنة	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
15-19	90	90	90	90	90	91	92	93	95	97
20-24	186	182	178	175	172	169	167	165	164	163
25-29	221	218	215	212	210	208	207	206	205	205
30-34	173	170	168	166	164	162	160	158	156	155
35-39	112	110	108	106	104	102	100	99	97	95
40-44	32	30	29	28	26	25	24	22	21	20
45-49	2	2	2	1	1	1	1	0	0	0
TFR	4.1	4.0	3.9	3.9	3.8	3.8	3.8	3.7	3.7	3.7

يلاحظ من الجدول (17) بأن القيم التنبؤية لمعدلات الخصوبة التفصيلية (ASFR) متسقة مع القيم المقدرة لهذه المعدلات للمدة (2011-1977) فضلاً عن الانخفاض البطيء لهذه المعدلات والموضح في قيم معدل الخصوبة الكلي (TFR).

Iraq: rate fertility rates (2012-2031)



شكل 6

معدلات الخصوبة التفصيلية العمرية (ASFR) التنبؤية للمدة (2031-2012)

جدول 18

نسبة التباين المئوية التجميعية CPV

العراق	المركبة الرئيسية PC	أنموذج (L-C)	الانموذج اللامعلمي
PC1	40.0	73.2	
PC2	66.6	95.5	
PC3	80.7	100	
PC4	90.3	100	
PC5	94.5	100	
PC6	97.5	100	
Total Percentage%	100.0%	100%	

يلاحظ من الجدول (18) بأن النسبة التباين للمركبة الرئيسية الأولى لأمودج (L-C) كانت 40% بينما بلغت هذه النسبة للمركبة الرئيسية الأولى للأمودج اللامعلمي 73.2% ويدل ذلك على تفوق الأمودج اللامعلمي على أمودج (L-C).

4. الاستنتاجات

كانت الطريقة اللامعلمية في التقدير والتنبؤ بمعدلات الخصوبة أفضل و أجدر بالاستعمال من طريقة L-C المعلمية من خلال نتائج الجانب التطبيقي، وكما يأتي:

1. أثبتت الطريقة اللامعلمية أفضليتها عن نمودج (L-C) وهذا ما افضاه الجدول (6) والجدول (13) حيث كانت قيمة متوسط مربعات الخطأ لنمودج (L-C) مساوية (0.4546) بينما كانت قيمته باستعمال الاسلوب اللامعلمي (0.1109).

2. برزت أفضلية الاسلوب اللامعلمي كما بين ذلك نتائج الجدول (18)، حيث كانت قيمة نسبة التباين المنوية التجميعية CPV للمركبة الرئيسية الأولى للطريقة اللامعلمية (73.2%) وهذا يعني بأن هذه المركبة تحمل اكبر نسبة من انحرافات بينما سجلت نسبة التباين المنوية لنمودج (L-C) للمركبة الرئيسية الأولى (40.0%) في حالة نمودج L-C المعلمي.

3. كانت نتائج الطريقة اللامعلمية أكثر واقعية ومنطقية بالنسبة لقيم معدلات الخصوبة التفصيلية العمرية المتنبئ بها ويبدو ذلك واضحاً من الجدول (17) وهو ما يتوافق مع المعدلات الحقيقية المدرجة في الجدول (8) اي انها متسقة مع المعدلات الاصلية

المصادر العربية

1. العبيدي، عمر فاروق خليل، (2014)؛ "تقدير وتنبؤ معدلات الخصوبة في العراق باستعمال طريقة لامعلمية حصينة مع مقارنتها بأسلوب L-C وبعض الطرائق المباشرة وغير المباشرة"، رسالة ماجستير غير منشورة في الاحصاء، كلية الادارة والاقتصاد، جامعة بغداد، جمهورية العراق.
2. علي، عمر عبدالمحسن ؛ و خليل، عمر فاروق؛ "تقدير معدلات الخصوبة في العراق باستعمال نمودج Lee-Carter والتنبؤ للفترة 2012 – 2031"، المؤتمر العلمي الدولي الرابع لاتحاد الاحصائيين العرب؛ المحور الاول: التحولات الديموغرافية في البلدان العربية، (الصفحة 214 – 232).
3. خواجة، خالد زهدي، "إحصاءات الخصوبة"، المعهد العربي للتدريب والبحوث الاحصائية.
4. الجهاز المركزي للاحصاء، وزارة التخطيط، (1978)، "نتائج التعداد العام للسكان لعام (1977)"، بغداد.
5. الجهاز المركزي للاحصاء، وزارة التخطيط، (1988)، "نتائج التعداد العام للسكان لعام (1987)"، بغداد.
6. الجهاز المركزي للاحصاء، وزارة التخطيط، (1998)، "نتائج التعداد العام للسكان لعام (1997)"، بغداد.
7. الجهاز المركزي للحصاء، وزارة التخطيط، (2012)، مديرية الاحصاء الاجتماعي والتربوي، "نتائج المسح العنقودي متعدد المؤشرات مراقبة اوضاع النساء والاطفال _ العراق 2011".

المصادر الأجنبية

8. Andreozzi. L.; Blaconá .M.T and Arnes.N; (2011); "The Lee – Carter Method for Estimating and Forecasting Mortality: An Application for Argentina"; School of Statistics Faculty of Economics and Statistics, National University of Rosario, Argentina.
9. D' Amato, V., Piscopo, G., Russolillo, M.; (2011); "The mortality of the Italian population: Smoothing techniques on the Lee--Carter model ". Annals of Applied Statistics, Vol. 5, No. 2A, pp. 705-724.
10. De Boor, C.; (1978); "A Practical Guide to Splines". Applied Mathematical Sciences, Springer-Verlag, New York, Vol. 27.
11. Eilers, Paul H. C., Marx, B. D.; (1996); "Flexible Smoothing with B-splines and Penalties". Statistical Science, Vol. 11, No. 2, pp. 89-102.
12. He, X., Ng, P.; (1999); "COBS: qualitatively constrained smoothing via linear programming". Computational Statistics 14, pp.315-337.
13. Hyndman, R. J. Ullah, M. S. (2007); "Robust forecasting of mortality and fertility rates: a functional data approach". Computational Statistics & Data Analysis, Vol. 51, pp. 4942-4956.
14. Lee, R. D.; (1993); "Modeling and forecasting the time series of U.S. fertility: age distribution, range, and ultimate level"; International Journal of Forecasting, Vol. 9, pp. 187-202.
15. Nikitović, V. (2011); "Functional Data Analysis in Forecasting Serbian Fertility". Stanovništvo Journal, Vol. 49, No. 2, pp. 73-89.
16. Tim, Neil H.; (2002); "Applied Multivariate Analysis". Springer-Verlag, New York.

الملحق خلاصة استعمال البرمجيات لادخال ومعالجة وتحليل البيانات

