



المجلة العراقية للعلوم الإحصائية

<http://stats.uomosul.edu.iq>



نمذجة وتحليل أسعار الموز في مدينة الموصل باستخدام نموذج ARFIMA "دراسة تنبؤية للسوق"

رحاب طلال احمد¹ و عمر سالم ابراهيم²

^{1,2} قسم الاحصاء والمعلوماتية ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق

الخلاصة

معلومات النشر

تاريخ المقالة:

تاريخ الاستلام: 6 كانون الاول 2023

تاريخ القبول: 19 تموز 2024

تاريخ القبول: 30 تموز 2024

متاح على الانترنت 1 كانون الاول 2024

الكلمات الدالة:

الذاكرة الطويلة، ARFIMA ،

التنبؤ، طرق تقدير معلمة التفاضل

الكسري d

المراسلة:

رحاب طلال احمد

rehab.22csp56@student.uomosul.edu.iq

تناولت هذه الدراسة استخدام نماذج ARFIMA للتنبؤ بأسعار الموز المستورد في مدينة الموصل وذلك بالاعتماد على البيانات التي تم الحصول عليها من مديره زراعه نينوى للفترة من سنة 2018 لغاية 2023 حيث استخدم في البحث عدة طرق لتقدير الذاكرة الطويلة وتحديد قيمه معلمه الفرق الكسري (d) منها طرق ذات المرحلة الواحدة والمتمثلة بطريقه Exact Maximum Likelihood (EML) المستخدمة في البحث وطرق ذات المرحلتين المتمثلة بطريقه (Geweke Porter_Hudak estimator, GPH) ، وطريقه (Rescaled_Range , R/S) ، وطريقه (Whittle_estimator) ، كما تم بناء النموذج من خلال التحقق من وجود الذاكرة الطويلة في السلسلة الزمنية من خلال عدة اختبارات ثم الانتقال الى تقدير معلمات الفرق الكسري حيث تبين تفوق طريق ذات المرحلة الواحدة ARFIMA(1,-0.06275898,0) من خلال المقارنة بمعايير BIC,MSE,RMSE,MAE واجتاز النموذج الفحوص التشخيصية والذي تم الاعتماد عليه في التنبؤ حيث كان هدف البحث التنبؤ بأسعار الموز وتوضيح خطوات بناء النموذج المناسب.

DOI: [10.33899/ijqoss.2024.185244](https://doi.org/10.33899/ijqoss.2024.185244) , ©Authors, 2024, College of Computer Science and Mathematics, University of Mosul.
This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

1. المقدمة

يمثل نموذج الذاكرة الطويلة في السلاسل الزمنية بأن تأثير الصدمات أو التغيرات على متغيرات الظاهرة طويل الأمد وسيظهر أثره في عملية التنبؤ المستقبلية وتم مشاهدة أسلوب الذاكرة الطويلة في السلاسل الزمنية لأول مرة من قبل الباحث البريطاني (Hurst, 1951). كما تطورت هذه الظاهرة فيما بعد في مجال الاقتصاد بعد أن تبين أن السلاسل الزمنية للعديد من الظواهر الاقتصادية لها ذاكرة طويلة. كان هنالك اهتمام دولي متزايد بدراسة نماذج الذاكرة الطويلة للسلاسل الزمنية، والتي تعتبر بدائل لنماذج ARIMA لأنها تحقق الاستقرار من خلال أخذ الفروق الكسرية (المعاملات التفاضلية الكسرية d) للفترة المغلقة (0.5, -0.5) ، وتمثل أهميتها في أنها تسمح بتفسير كلاً من التقلبات قصيرة الأجل للسلسلة الزمنية من خلال معلمات الانحدار الذاتي والمتوسطات المتحركة، والتقلبات طويلة الأجل من خلال معلمة التكامل الكسري.

تناولت العديد من الدراسات هذا النوع من النماذج وتطبيقها في العديد من المجالات المختلفة ومنها دراسة (MADOURI&Mohamed, 2017)، "مقارنة بين نماذج الذاكرة الطويلة ARFIMA والشبكات العصبية الاصطناعية للتنبؤ بسعر صرف الدينار الجزائري" ، تناولت هذه الدراسة مقارنة نماذج الذاكرة الطويلة ARFIMA لسعر صرف الدينار الجزائري مقابل العملات الأجنبية والهدف من هذه الدراسة هي استخدام نماذج الذاكرة الطويلة والشبكات العصبية في التنبؤ بالنموذج الافضل . كذلك دراس (Safitir. et al , 2019) "نمذجة أسعار الذهب في إندونيسيا باستخدام طريقة ARFIMA" ، يعتبر

استثمار الذهب هو الخيار الافضل في التمويل، الهدف من هذه الدراسة هو التنبؤ بأسعار الذهب في إندونيسيا حيث بيانات اسعار الذهب هي سلسلة زمنية ذات ذاكرة طويلة ARFIMA لها قيمة معلمة كسرية وتبين من هذه الدراسة كفاءة أسلوب ARFIMA. دراسة (Ben, Ayaa, 2022) (Belmukaddam, 2022) " التنبؤ بالبطالة في ظل وجود وباء كورونا في الجزائر باستخدام نماذج الذاكرة الطويلة ARFIMA خلال الفترة الزمنية 2008-2022 ، الهدف من هذه الدراسة هو تطبيق نموذج ARFIMA للتنبؤ بمعدلات البطالة الشهرية حيث تعبر مشكلة البطالة عن وجود عجز في البنى الاقتصادية وظهرت نتائج هذه الدراسة أن النموذج ARFIMA(1,0.34,10) هو النموذج الافضل واستخدم للتنبؤ ل 15 شهراً.

هدف البحث:

يهدف الى المقارنة بطرق تقدير المعلمة الكسرية (d) واختيار النموذج الافضل والتنبؤ بأسعار الموز المستورد في مدينة الموصل .

2.الذاكرة الطويلة ARFIMA:

أن النماذج الاساسية التي تسمح بتحديد الذاكرة الطويلة ،هي Autoregressive Integrated Moving Average (ARFIMA) (Hosking, 1981_Granger and Joyeux, 1980) وتعتبر امتداد لنماذج (ARFIMA) التي تأخذ معامل التفاضل الكسري d قيماً حقيقية للفترة (-0.5, 0.5)، والصيغة الرياضية لنموذج ARFIMA يمكن التعبير عنها باستخدام علاقة (Wold) بالشكل التالي:

$$\gamma_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad (1)$$

γ_t : سلسلة زمنية.

ψ_j : أوزان المتوسط المتحرك ، $\psi_j \in R, \psi_0 = 1$.

ε_t : هي عملية التشويش الابيض وبالصورة الاتية (White Noise) حيث أن:

$$\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t) = 0, \varepsilon_t \sim \text{i.i.d} (0, \sigma_\varepsilon^2)$$

وتعتبر السلسلة الزمنية ذات ذاكرة طويلة وساكنة في حال:

$$\sum_{j=0}^{\infty} |\psi_j|^2 < \infty \quad (2)$$

وعليه يمكن القول ان اي سلسلة زمنية تكون نموذج ARFIMA (p,d,q) اذا تحقق الشرط التالي:

$$\Phi_p(L)(1-L)^d \gamma_t = \Theta_q(L) \varepsilon_t \quad (3)$$

$$\begin{aligned} \Phi_p(L) &= 1 - \sum_{i=1}^p \Phi_i L^i = 1 - \Phi_1 L - \dots - \Phi_p L^p \\ \Theta_q(L) &= 1 - \sum_{j=1}^q \Theta_j L^j = 1 - \Theta_1 L - \dots - \Theta_q L^q \end{aligned} \quad (4)$$

يمثل كل من (Φ, L) ، (Θ, L) على التوالي كثيري الحدود في الجزئين AR(P), MA(q) لنموذج p,q على الترتيب. $(1-L)^d$ مؤثر الازاحة الخلفي، d معامل التفاضل الكسري .

كما أن خصائص نماذج ARFIMA يمكن توضيحها تبعاً للقيم المختلفة لمعامل التفاضل الكسري d (Benguesmi, 2014) عندما $d < 1/2$ والجذور كثيرة الحدود $\Theta_q(L)$ تقع خارج جذور الوحدة ، ففي هذه الحالة السلسلة γ_t قابلة للانعكاس . وعندما $d < -1/2$ كل جذور كثيرة الحدود المميزة $\Phi_p(L)$ تقع خارج جذر الوحدة ، ففي هذه الحالة السلسلة مستقرة وعندما $-1/2 < d < 0$ تكون السلسلة γ_t قابلة للانعكاس وبذاكرة قصيرة غير تامة . في حين $0 < d < 1/2$ تكون السلسلة γ_t سلسلة ساكنة بذاكرة طويلة (الاستقرار طويل الامد) ،وتكون مستمرة ايضاً حيث تتناقص دالة الارتباط الذاتي الموجبة بشكل بطيء نحو الصفر "على شكل قطع زائد" بكون عدد الفجوات k.

(3) الكشف والتحقق عن الذاكرة الطويلة باستخدام الرسوم:

أ- رسم دالة الارتباط الذاتي ("ACF")

هنالك العديد من الاشكال البيانية التي تعطي مؤشراً واضحاً على وجود الذاكرة الطويلة ،احدها هي دالة الارتباط الذاتي للسلاسل المستقرة والتي تعرف كالتالي :

$$\rho_k = \frac{\gamma_k}{\gamma_0}; k = 0, \pm 1, \pm 2, \dots \quad (5)$$

حيث أن $\hat{\rho} = r_k, k = 0, 1, \dots$ ، $\gamma_0 = \text{var}(\gamma_t)$ ، $\gamma_k = \text{Covariance}(\gamma_t, \gamma_{t-k})$ تعتبر تقدير لمعامل دالة الارتباط الذاتي

acf ويعطى بالصيغة الآتية:

$$r_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}, k = 0, 1, 2, \dots \quad (6)$$

التناقص البطيء لدالة ACF يعتبر مؤشراً على وجود ذاكرة طويلة للسلسلة .

ب - ("R/S Plot")

استخدمت احصائية R/S لاختبار وجود ذاكرة طويلة كالتالي (Hurst, 1951):

$$\ln[Q(t, k)] = \ln \left[\frac{R(t, k)}{S(t, k)} \right] \quad (7)$$

$$Y_j = \sum_{i=1}^j X_i$$

حيث أن :

$$R(t, k) = \max_{0 \leq i \leq k} \left[Y_{t+i} - Y_t - \frac{i}{k} (Y_{t+k} - Y_t) \right] - \max_{0 \leq i \leq k} \left[Y_{t+i} - Y_t - \frac{i}{k} (Y_{t+k} - Y_t) \right]$$

$$\bar{X}_{t,k} = \frac{\sum_{i=t+1}^{t+k} X_i}{k}$$

$$S(t, k) = \sqrt{\frac{\sum_{i=t+1}^{t+k} (X_i - \bar{X}_{t,k})^2}{k}}$$

R: تعبر عن المدى , S: تعبر عن الانحراف , k : تعبر عن الفجوات عبر الزمن لتطبيق هذه الاحصائية بالرسم البياني لـ (R/S) رسم $\ln[Q(t, k)]$ مقابل $\ln k$

(4) التحقق من خاصية الذاكرة الطويلة باستخدام الاختبارات الاحصائية:

❖ اختبار (Rescaled Range) R/S

استخدم (Hurst, 1951) إحصائية يتم من خلالها فحص اذا كانت السلسلة الزمنية ذات ذاكرة طويلة (ارتباطات في المدى الطويل)، أذ أن الهدف من إحصائية R/S هي حساب معامل هورست وتعرف الاحصائية R/S على أنها المدى R للمجاميع الجزئية للانحرافات لسلسلة زمنية عن المتوسط مقسوماً على انحرافها المعياري S_T .

$$R/S = Qn = \frac{1}{\sigma_y} \left[\max_{1 \leq k \leq n} \sum_{j=1}^k (y_j - \bar{y}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (y_j - \bar{y}_n) \right] \quad (9)$$

K : عدد المجاميع الجزئية بين المفردات للسلسلة y_t و متوسطها الحسابي $\bar{y}$ ويسمح تحليل R/S بحساب ما يسمى بمعامل هورست $0 < H < 1$ والذي يعرف على أنه النسبة بين لوغاريتم R/S ولوغاريتم عدد المشاهدات T :

$$H \approx \frac{\log Qt}{\log T} \quad (10)$$

وقد بين كل من (Hosking, 1981), (Lo, 1991) وجود علاقة قوية بين معامل التفاضل الكسري (d) لنماذج ARFIMA ومعامل هورست (H) حيث أن :

$$d = H - \frac{1}{2} \quad (11)$$

ومن خلال العلاقة السابقة يمكن تحديد معلمة التفاضل الكسري d وعرفة فيما اذا كانت السلسلة ذات ذاكرة طويلة ، وهذا حسب القيم المتغيرة للعامل H وكما يلي

- أ- عندما $H = \frac{1}{2}$ يكون $d=0$ في هذه الحالة لا يوجد اي ارتباط بين الظواهر للأحداث الماضية والحاضرة ,اي انه نموذج ARMA.
- ب- عندما $0 < H < \frac{1}{2}$ يكون $0 < d < \frac{1}{2}$ ، عندها يعتبر نموذج ARFIMA مستقراً بذاكرة طويلة حيث تكون الارتباطات قوية كلما اقترب H من الواحد.
- ت- عندما $0 < H < \frac{1}{2}$ يكون $-\frac{1}{2} < d < 0$ في هذه الحالة السلسلة ذات ذاكرة طويلة وفي نفس الوقت لاتتلك سلوك نماذج ARMA.

❖ اختبار Empirical Hurst Exponent

تحليل (DFA) Detrended Fluctuation Analysis هو تحليل اقترحه (Peng et al., 1994) والذي تم قياس الاعتماد طويل المدى للسلسلة الزمنية باستخدام هذا الامتداد لتحليل التقلب الطبيعي؛ بحيث تقع قيمة Exponent Hurst بين 0 و 1، ويمكن لتحليل DFA استخراج قيمة Exponent Hurst أكبر من 1 وتشير هذه القيمة إلى عدم الاستقرار (Bryce ., al et, 2012) ، حيث يقسم تحليل DFA السلسلة الزمنية تقسيم السلسلة الزمنية X للطول N إلى d فرعية من طول n, لكل تصنيف فرعي وبعد ذلك سيتم تنفيذ الخطوات التالية (Ceballos and , 2017) (Largo): إنشاء سلسلة زمنية تراكمية.

$$Y_{i,m} = \sum_{j=1}^i X_{j,m} \text{ for } i = 1, \dots, n$$

تقدير خط المربعات الصغرى.

$$\hat{Y}_m(t) = a_m * t + b_m \text{ to } \{Y_{1,m}, Y_{n,2}, \dots, Y_{n,m}\}$$

حساب الجذر التربيعي لوسط السلسلة الزمنية المشتقة.

$$F(m) = \sqrt{\frac{1}{2} \sum_{i=1}^n (Y_{i,m} - a_m(i) - b_m)^2}$$

حساب قيمة الوسط للجذر التربيعي للمتوسطات للطول n من السلسلة الزمنية.

$$F(n) = \frac{1}{d} \sum_{i=1}^n F(n)$$

❖ اختبار (Whittle estimator)

يتم تقدير معلمة التفاضل الكسري (d) باستخدام مقدر Whittle ، يستند هذا المقدر على دالة (Murad S. Taqqu and) periodogram (Teveovsky, 1998) وتتبع الصيغة التالية :

$$\varphi(\eta) = \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda, \eta)} d\lambda \quad (12)$$

اذ أن $I(\lambda)$ تمثل دالة periodogram و $f(\lambda, \eta)$ دالة الكثافة الطيفية للتكرار وأن مقدر Whittle هو قيمة η التي تعمل على تقليل الدالة φ بحيث تعتبر η في عمليات ARFIMA المعلمة H او d ، عندما تكون السلسلة الزمنية هي نموذج ARFIMA (p,d,q) فمن المفترض أن تكون η تشمل أيضاً المعاملات غير المعروفة من الانحدار الذاتي والمتوسطات المتحركة ، كما وان طريقة Whittle تختلف عن الطرق الاخرى اذ انها ليست طريقة بيانية وتعتبر من الطرق المعلمية كما وضع (Fox and Taqqu,1986) المستخدم لتقدير معلمة التفاضل الكسري d ويرمز لها $\widehat{d}_w$ في بعض الاحيان .

5. تقدير نماذج (ARFIMA)

يوجد أسلوبين من الطرق لتقدير نماذج (ARFIMA) ، حيث قسمت هذه الطرق الى طرق ذات المرحلتين بحيث يتم تقدير معلمة التفاضل الكسري (d) ومن ثم تقدير معلمات ARMA ، وطريقة التقدير بمرحلة واحدة حيث تقدر معلمة التفاضل الكسري بالتزامن مع تقدير معلمات ARMA وتعتبر طرق التقدير بمرحلة واحدة من اكثر الطرق فعالية وتتطلب هذه الطريقة ان تكون السلسلة مستقرة او تحويلها الى مستقرة ومن هذه الطرق (EML) Exact Maximum Likelihood والتي تعتبر من أكثر الطرق فعالية لتقدير معلمة الفرق الكسري (d) بالموازاة مع معلمات AR,MA لنموذج ARFIMA وشرط هذا النموذج أن تكون السلسلة مستقرة أو يتم تحويلها الى مستقرة وتسمح هذه الطريقة باستخدام كل المعلومات الطويلة الاجل والقصيرة الاجل المرتبطة بالسلسلة الزمنية .

أما طرق التقدير ذات المرحلتين نستعرض فيما يلي بعض منها :

❖ طريقة (Geweke _Porter _Hudak estimator):

هذه الطريقة اقترحها Geweke و Porter وهي عبارة عن مقدر شبه حدودي للمعلمة d بالاعتماد الى احداثيات المخطط الدوري الاول بالنسبة الى ARFIMA (p,d,q) أحادي المتغير كما ذكر في (Shang, 2020, Geweke and Hudak,1983).

$$\hat{d}_{GPH} = \frac{-\frac{1}{2} \sum_{j=1}^J [\log_{10}(\lambda_j)] - \log_{10} I(\lambda_j)}{\sum_{j=1}^J [\log_{10} \lambda_j - \log_{10}(\lambda_j)]^2}, j = 1, \dots, J \quad (13)$$

حيث أن :

$$\begin{aligned} \overline{\log_{10}(\lambda_j)} &= \sum_{j=1}^J \log_{10}(\lambda_j) \\ J &= \sqrt{n}, \quad \lambda_j = \frac{2\pi j}{n}, \quad \lambda_j \in [-\pi, \pi] \\ I(\lambda_j) &= \frac{1}{2\pi} R(0) + 2 \sum_{s=1}^{n-1} R(s) \cos s\lambda_j, \quad \text{مخطط الفترة الذي هو مقياس للتباين الذاتي} \\ R(S) &= \frac{1}{n} \sum_{t=1}^{n-s} (x_t - \bar{x})(x_{t+s} - \bar{x}), S = \pm 1, \dots, \pm(n-1) \end{aligned} \quad (14)$$

❖ طريقة (Smoothed priodogram estimation) dsprio:

أثبت كلاً من (Brockucll, and Davis, 2009) أن دالة priodogram هي مقدر غير متسق لدالة الكثافة الطيفية ، لذلك اقترح (Reisen, 1994) استخدام مقدر متسق وهو دالة ممهدة من دالة priodogram ، وبداية يجب تعريف دالة الممهد بالصيغة الاتية:

$$I(\omega) = \frac{1}{2\pi} (\hat{\gamma}X(0) + 2 \sum_{l=1}^{n-1} \hat{\gamma}X(l) \cos(l\omega)) \quad (15)$$

حيث أن:

$$\hat{\gamma}X(k) = \frac{1}{n} \sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x}), k \in \{0, 1, \dots, n-1\}$$

ويشار لهذا المقدّر بـ $\widehat{d}_{sp}$ ، ويتم الحصول على مقدّر الانحدار بواسطة (parzen lag window) مع نافذة تأخر parzen ويرمز لها $l_s(\cdot)$:

$$I_S(\omega) = \frac{1}{2} \sum_{j=-v}^v \kappa\left(\frac{j}{v}\right) \hat{\gamma}X(j) \cos(j\omega) \quad (16)$$

حيث أن $K(\cdot)$ هي تأخر نافذة Parzen وتكون بالشكل الآتي:

$$\kappa(u) = \begin{cases} 1 - 6u^2 + 6|u|^3, & \text{if } |u| \leq \frac{1}{2} \\ 2(1 - |u|)^3, & \text{if } \frac{1}{2} < |u| \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

كما يعرف بالمخطط الدوري الممهد وبمقارنة المعادلة السابقة بمعادلة خط الانحدار كما في الطريقة السابقة لنحصل على تقدير d بالكيفية نفسها من خلال طريقة المربعات الصغرى OLS التي تصغر دالة الخسارة لليواقي .

❖ طريقة (Fracdiff):

أحد الطرق المستخدمة في تقدير معلمة التفاضل الكسري d وتعتمد على استخدام مقدر ML (Maximum Likelihood Method) لإيجاد مقدرات نموذج ARFIMA (p,d,q) وتقدير التغيرات ومصفوفة الارتباط والخطأ المعياري بالإضافة إلى قيمة الاحتمالية العظمى (Likelihood) وهذه الطريقة تكون دقيقة وسريعة التقدير (Haslett and Raftery, 1989). وقد بينت التجارب أن نتائج تقدير ML غير متحيزة وفعالة في حالة استخدام بيانات محدودة الحجم بحجم العينة 100، ويتم تقدير معلمة التفاضل الكسري d بالصيغة التالية :

$$l(\alpha, \beta, d, \Phi(B), \theta(B)) = \text{constant} - \frac{1}{2} Nm \log \hat{\sigma}_\varepsilon^2 - \frac{1}{2} N |R| \quad (18)$$

حيث أن $\Phi(B), \theta(B)$ معاملات النموذج، $\hat{\sigma}_\varepsilon^2$ مقدّر $\hat{\sigma}_\varepsilon^2$ ، R مصفوفة ارتباط المتجه .

6. بناء نموذج ARFIMA

أي عملية بناء نموذج تتضمن عدد من الخطوات، لذلك سوف نقوم بعرض هذه المراحل والتي سوف نقوم بتطبيقها في الجانب التطبيقي.

• التعرف والتقدير

إن العثور على النموذج الأولي المثالي هو الخطوة الأولى في تحليل سلاسل الزمنية يعد تحديد رتبة النموذج ضرورياً لتحديد النموذج الأفضل ، أي تحديد درجات الانحدار الذاتي وعدد الفروق كما تعد مرحلة التعرف من أهم وأصعب المراحل في مجال السلاسل الزمنية بصفة خاصة وفي مجال الإحصاء بصفة عامة (Shaarawy, 2005) ، وتتضمن خطوه التعرف على النموذج عدة خطوات وهي تقدير معلمتي (p,q) من خلال التحقق من شكل دالة الارتباط الذاتي (ACF) ودالة الارتباط الذاتي الجزئي (PACF) والتي يتم من خلالها حساب عدد الارتباطات الذاتية حيث يستخدم ACF لتحديد رتبة (AR)، وتستخدم دالة PACF لتحديد رتبة MA المعنوية لتشكيل وتحديد معلمة (q).

المرحلة الثانية والتي تأتي بعد اختيار النموذج الأفضل للسلسلة هي عملية تقدير النموذج وعليها أن نضمن دقة تقدير النموذج حتى يتمكن من تحقيق الهدف الأساسي منه وهو التنبؤ، بعد إيجاد وتحديد رتبة النموذج ومدى قابليته للتطبيق على السلسلة الزمنية أهم طرق التقدير هي طرق ذات المرحلة الواحدة والمرحلتين والتي تم شرحها سابقاً .

• التشخيص (Diagnostic)

تعتبر هذه المرحلة من اهم المراحل المتبعة في تطوير النموذج لتوفير امكانية تطبيق معلومات النموذج للتنبؤ من خلال ملائمة معلماتها مع الفرضيات الإحصائية بحيث يعتمد تشخيص النموذج بصورة عامه على اجراء العديد من الاختبارات (Shaarawy,2005) ومنها اختبار وتحليل الاستقرار به من خلال فحص تقديرات معاملات الارتباط الذاتي التي تم الحصول عليها من مرحلة التقدير (الشعراوي,2005).

هناك عدة اختبارات لفحص جذر الوحدة والتأكد من الاستقرار وتحديد رتبة كل متغير على حدة من هذه الاختبارات اختبار ديكي-فولر (Dickey and Fuller _1979)، اختبار ديكي _ فولر الموسع (Augmented Dickey _ Fuller test)، واختبار PP (Phillip and Peron) وهو افضل وادق من اختبار ADF وخاصة عندما يكون حجم العينة صغير ومن الافضل الاعتماد على نتائج اختبار PP في حالة عدم الانسجام في نتائج الاختبارين.

الخطوة التالية هي تحليل البواقي حيث يتم تعريفها على أنها الفرق بين المشاهدات الفعلية والقيم المتوقعة وبفحص النموذج نحاول التأكد من ان لها نفس صفات التباينات العشوائية البحتة التي يعتقد انها مستقلة ولها توزيع طبيعي بمتوسط صفر وتباين ثابت والتحقق من توفر الخصائص يتم بإجراء رسم البواقي واختبار دالة الارتباط الذاتي والذاتي الجزئي للبواقي. واختبار (Ljung_Box) يقوم هذا الاختبار على اختبار الفروض $H_0: P(K) = 0$ اي البواقي مستقلة والفرضية البديلة $H_1: P(K) \neq 0$ البواقي غير مستقلة ويطبق اختبار Ljung_Box للتأكد من عشوائية السلسلة الزمنية ويتم عبر معاملات الارتباط الذاتي للبواقي لمجموعة من الازاحات كما قام كل من (Pierce and Box,1970) بتقديم الإحصائية Q لا جراء اختبار معنوية عدد من معاملات الارتباط الذاتي كمجموعة واحدة وكما موضح بالمعادلة التالية والتي تتبع توزيع χ^2 بدرجة حرة m .

$$Q_{BP} = n \sum_{k=1}^m r^2(k) \quad (19)$$

قام Ljung_Box بإجراء بعض التعديلات والتغييرات على إحصائية Q_{BP} ولنفس التوزيع χ^2 ودرجة حرية m لتكون بشكل الاتي:

$$Q_{BP} = n(n+2) \sum_{k=1}^m \frac{r^2(k)}{n-k} \quad (20)$$

حيث أن n حجم العينة ، r معامل الارتباط ، k درجات الحرية، ولهذا الاختبار عيوب ومنها تأثره بعدد معاملات الارتباط الذاتي للعينة المستخدمة والتي تتغير بتغير عدد المعاملات.

7. التنبؤ (Forecasting):

التنبؤ هو المرحلة الأخيرة من مراحل التحليل السلاسل الزمنية ولا يمكن الوصول اليها الا بعد الاختبارات الخاصة بتشخيص النموذج ، وبعد الحصول على نموذج ملائم لتمثيل البيانات يصبح النموذج جاهزا لاستخدامه للتنبؤ بالقيم المستقبلية (Tuama,2012) تستخدم في هذه المرحلة معايير او (مقاييس تقييم) التي توصلنا الى دقة النموذج وقدرته على انتاج بيانات كفؤة فيما يلي بعض من هذه المعايير :

اولا: معيار BIC (Bayesian information criterion)

قام الباحث Akaike في عام (1978 _ 1979) بتطوير وتحديث معيار AIC لنحصل على معيار BIC وصيغته الرياضية بالشكل الاتي:

$$BIC = 2 \ln \sigma_e^2 + M \ln(n) \quad (21)$$

حيث ان (M , n) عدد مشاهدات السلسلة والعدد الكلي لمعاملات النموذج على التوالي، σ_e^2 متوسط مربعات الاخطاء ، القيمة الاقل معيار BIC تمثل النموذج المناسب لبيانات السلسلة.

ثانيا: جذر متوسط مربعات الاخطاء (RMSE)

الاخطاء هي الفرق بين القيم الفعلية (الحقيقية) والقيم التقديرية ويشار اليها $\hat{y} - \gamma$ حيث γ اي تمثل القيم الحقيقية للمشاهدات $\hat{y}$ القيم المتوقعة بعد تحديد البواقي يمكننا حساب RMSE وفق الصيغة التالية:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad (22)$$

ومن مميزات هذا المعيار سهوله خصائصه الإحصائية لذلك يعتبر من اهم وافضل المعايير التي تقيس دقة التنبؤ.

ثالثا: متوسط الخطأ المطلق (Mean absolute error)

الصيغة الرياضية لهذا المقياس هي كالآتي:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right| \quad (23)$$

حيث e_t تمثل البواقي ، y_t المشاهدات الحقيقية للسلسلة ، $\hat{y}_t$ القيم المقدرة ، n عدد المشاهدات الكلي

الجانب التطبيقي

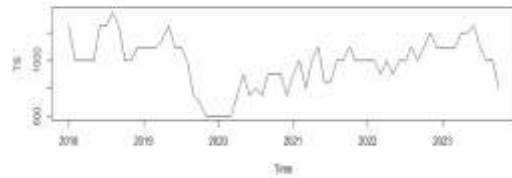
وصف بيانات السلسلة المدروسة :

تم الحصول على بيانات السلسلة الزمنية لسعر إحدى المحاصيل الزراعية من مديرية زراعة نينوى والفترة من شهر كانون الثاني 2018 لغاية كانون الاول 2023 و بشكل اسبوعي وتبلغ 288 مشاهدة سنقوم بتحليل السلسلة لبناء نموذج ARFIMA وسوف نقوم بتحليل سعر محصول الموز المستورد وإيجاد التنبؤ لـ 12 قيمة ويتم مقارنتها مع القيم الحقيقية لمعرفة دقة التنبؤ .

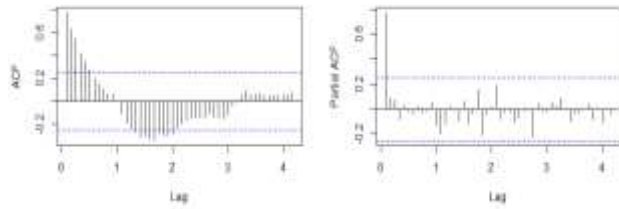
مراحل منهجية ARFIMA:

1. فحص استقرارية السلسلة:

سنقوم برسم السلسلة الزمنية لبيانات سعر محصول الموز المستورد ورسم دالة الارتباط الذاتي ودالة الارتباط الذاتي الجزئي تهدف هذه الخطوة الى فحص استقرارية السلسلة الزمنية ، وإجراء التحويلات اللازمة في حالة عدم استقرارها يوضح الشكل (1) سعر الموز المستورد في مدينة الموصل والذي اتضح منه ان السلسلة الزمنية موضع الدراسة غير مستقرة ، وللتأكد نقوم برسم كل من دالتي (ACF, PACF) ، والموضحة بالشكل (2).



الشكل (1) السلسلة الزمنية لسعر الموز



الشكل (2) الارتباط الذاتي والارتباط الذاتي الجزئي لسعر الموز

نلاحظ من خلال رسم كل من ACF, PACF أن السلسلة غير مستقرة كما وأنه من المعروف ان هناك اختبارات لفحص استقرارية السلسلة الزمنية ، وسيتم ذلك من خلال اختبار ديكي_فولر الموسع (ADF) ، واختبار فيليبس_بيرن (PP) ويوضح الجدول هذه الاختبارات .

الجدول (1) اختبارات التحقق من استقرارية السلسلة

الاختبار الاحصائي	القيمة	P-value	حالة السلسلة
ADF	-2.0325	0.562	غير مستقرة
PP	-15.085	0.2188	غير مستقرة

تبين من معطيات الجدول اعلاه ان الدلالة الاحصائية P-value لاختبار ADF بلغت (0.562) وهي اكبر من 0.05 ، وكذلك لاختبار فيليب_بيرن بلغت (0.2188) أي ان $P\text{-value} > 0.05$ وهذا يعني ان السلسلة الزمنية غير مستقرة.

2-اختبارات الذاكرة الطويلة:

كما نعلم ان هناك العديد من الاختبارات الاحصائية والرسوم البيانية لاختبار وجود الذاكرة الطويلة في تحليل السلاسل الزمنية وفيما يلي مجموعة من هذه الاختبارات.

جدول (2) اختبار الذاكرة الطويلة

الاختبار	H
R/S	0.7484858
Empirical Hurst exponent	0.7232992
Theoretical Hurst exponent	0.5174121
Whittle	0.9472698

لنلاحظ من الجدول أن قيمة $1/2 < H < 1$ فإن السلسلة الزمنية تتميز بذاكرة طويلة وأن الارتباطات تكون قوية كلما اقترب H من الواحد

3. طرق تقدير معامل الفرق الكسري (d)

فيما يلي تقدير معامل التفاضل الكسري بطرق مباشرة وغير مباشرة حيث تعتمد على معامل هورست H و تحسب قيمة d في اختبار R/S, Whittle حسب الصيغة $d=H-1/2$ ، وفي الطرق المباشرة يتم حساب d مباشرة مثل اختبار , dSperio , Gph , Fracdiff .

الجدول (3) طرق تقدير (d)

الاختبار	قيمة d
R/S	0.2484858
Empirical Hurst exponent	0.2232992
Theoretical Hurst exponent	0.0174121
Gph	1.342017
dSperio	0.6107118
Fracdiff	0.4796818
Whittle	0.447

يتم استثناء طريقة Gph , dSperio لأن قيمة d تقع خارج الفترة (0.5 , 0) ، بينما الطرق المتبقية تقع ضمن حدود الفترة (0.5 , 0) .

4.مرحلة التعرف:

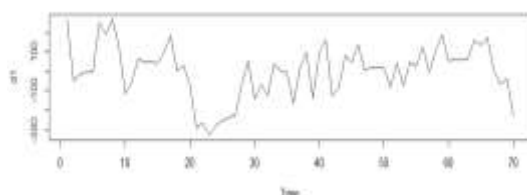
تعد هذه الخطوة مهمة وهي اولى مراحل تحليل السلاسل الزمنية حيث نتعرف من خلالها على النموذج ، أي اختيار رتب النموذج (p,d,q) ولكون السلسلة غير مستقرة سنستخدم معامل التفاضل الكسري d التي حصلنا عليه من الطرق المذكورة سابقاً من الجدول (3) يتضح لنا نتائج تقدير معامل التفاضل الكسري حيث كانت قيمة d بطريقة R/S (0.2484858) وبطريقة Empirical Hurst exponent (0.2232992) وبطريقة

Theoretical Hurst exponent (0.0174121) وبطريقة GPH (1.342017) وتبين فشل هذه الطريقة في تقدير الفرق الكسري حيث لم تقع ضمن الفترة (0,0.5) ، وكذلك فشلت بطريقة dsperio والتي تساوي (0.6107118) لاتقع ضمن الفترة ، وقيمة الفرق بطريقة Fracdiff (0.4796818) ، وبطريقة whittle (0.447) .
نقوم باختبار استقرارية السلسلة بعد اخذ الفروق الكسرية الخمسة التي نجحت باختبار الذاكرة الطويلة باستخدام اختبار PP وذلك لكونه افضل وأدق من اختبار ADF .

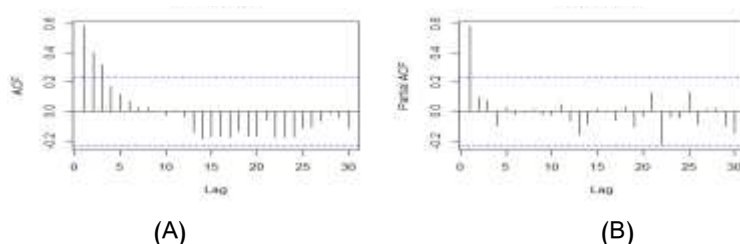
الجدول (4) نتائج اختبار PP للسلسلة بعد اجراء الفروق الكسرية للطرق السابقة

طرق التقدير	القيمة	P-value	الحالة
R/S	-31.975	0.01	مستقرة
Empirical Hurst exponent	-29.787	0.01	مستقرة
Theoretical Hurst exponent	-15.755	0.01	مستقرة
Fracdiff	-53.622	0.01	مستقرة
Whittle	-50.642	0.01	مستقرة

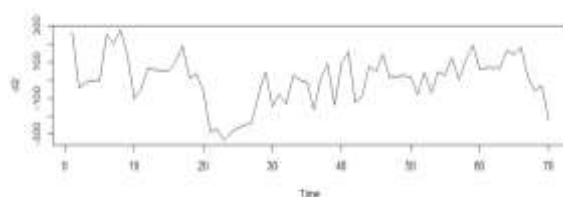
تثبت نتائج اختبارات pp المبينة في الجدول (3) أن السلسلة الزمنية مستقرة ويمكن استخدامها في بناء نماذج ARFIMA حيث أن القيمة المعنوية لاختبار الطرق الخمسة كانت اصغر من 0.05 كما سيتم رسم السلسلة الزمنية بعد اخذ الفروق الكسرية الخمسة وكذلك رسم PACF, ACF لنتحقق من تأثير الفروق الكسرية عليها ومدى تحقق الاستقرارية



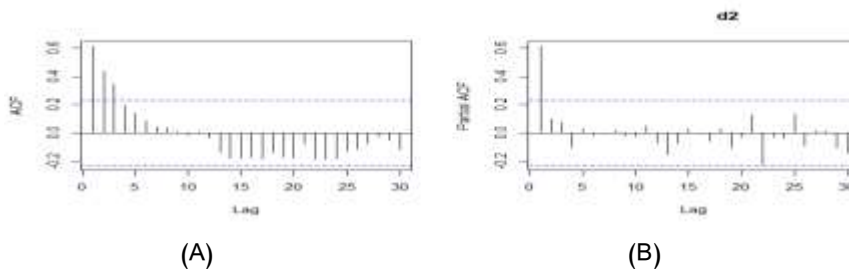
الشكل(3) السلسلة الزمنية بعد اخذ الفروق الكسري لاختبار R/S لمحصول الموز



الشكل(4) A يمثل الارتباط الذاتي و B الارتباط الذاتي الجزئي لسعر الموز بعد اخذ الفروق الكسري لاختبار R/S



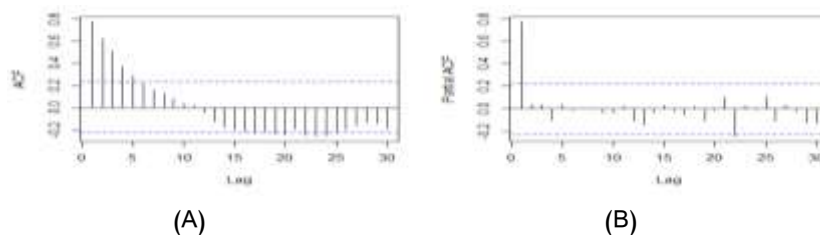
الشكل(5) السلسلة الزمنية بعد اخذ الفروق الكسري لاختبار Empirical Hurst exponent لمحصول الموز



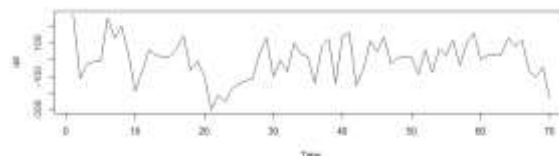
الشكل (6) A يمثل الارتباط الذاتي و B الارتباط الذاتي الجزئي لسعر الموز بعد اخذ الفرق الكسري لاختبار Empirical Hurst exponent



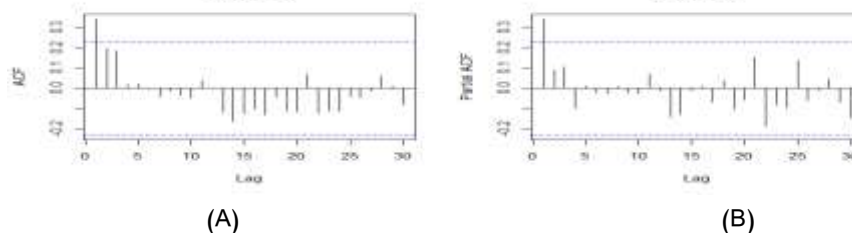
الشكل (7) السلسلة الزمنية بعد اخذ الفرق الكسري لاختبار Theoretical Hurst exponent لمحمول الموز



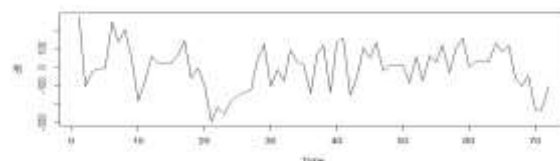
الشكل (8) A يمثل الارتباط الذاتي و B الارتباط الذاتي الجزئي لسعر الموز بعد اخذ الفرق الكسري لاختبار Theoretical Hurst exponent



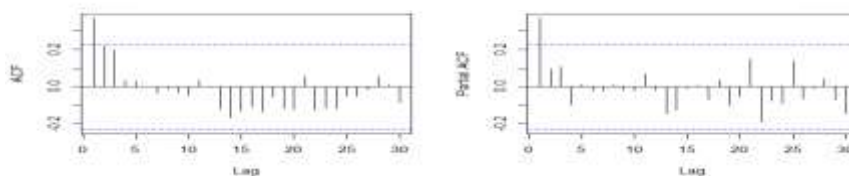
الشكل (9) السلسلة الزمنية بعد اخذ الفرق الكسري لاختبار Fracdiff لمحمول الموز



الشكل (10) A يمثل الارتباط الذاتي و B الارتباط الذاتي الجزئي لسعر الموز بعد اخذ الفرق الكسري لاختبار Fracdiff



الشكل (11) السلسلة الزمنية بعد اخذ الفرق الكسري لاختبار Whittle لمحمول الموز



(A)

(B)

الشكل (12) A يمثل الارتباط الذاتي و B الارتباط الذاتي الجزئي لسعر الموز بعد اخذ الفرق الكسري لاختبار Whittle

5. التشخيص والتقدير:

الهدف من هذه الخطوة هو تحديد واحد أو أكثر من نماذج ARFIMA من خلال معرفة رتبة $AR(P)$ ورتبة $MA(q)$ حيث سنستخدم معيار (BIC) لمقارنة النماذج ولتقدير الطرق المستخدمة للفروقات الكسرية.

الجدول (5) تقدير نماذج ARFIMA بطريقة المرحلتين

قيمة d	النموذج	BIC	1θ	θ_2	θ_1	2θ
R/S $d = 0.2484858$	(0,d,0)	684.5347	0	0	0	0
	(1,d,0)	682.7829	0.8881	0	0	0
	(0,d,1)	686.3361	0	0	-0.2317	0
	(1,d,1)	686.9793	0.9025	0	-0.0660	0
	(2,d,0)	686.996	0.9634	-0.0558	0	0
	(2,d,1)	690.4313	0.2434	0.5959	-0.7332	0
	(0,d,2)	690.4946	0	0	-0.2466	-0.0391
	(1,d,2)	694.6892	0.89855	-0.0370	0.03953	0
	(2,d,2)	697.1149	0.300925	0.52699	-0.60817	0.06012
	(0,d,0)	684.8847	0	0	0	0
Empirical Hurst exponent $d = 0.2232992$	(1,d,0)	682.6296	0.85729	0	0	0
	(0,d,1)	686.3078	0	0	-0.27479	0
	(1,d,1)	686.82	0.87245	0	-0.06776	0
	(2,d,0)	686.819	0.9485	-0.06546	0	0
	(2,d,1)	690.576	0.1950	0.5938	-0.7558	0
	(0,d,2)	690.4464	0	0	-0.2927	-0.0563
	(1,d,2)	691.0854	0.8717	0	-0.0603	0.0151
	(2,d,2)	694.7816	0.242692	0.5370	-0.6467	0.0549
Theoretical Hurst exponent $d = 0.0174121$	(0,d,0)	691.4151	0	0	0	0
	(1,d,0)	681.2809	0.85315	0	0	0
	(0,d,1)	686.8192	0	0	0.3804	0
	(1,d,1)	685.4	0.87564	0	-0.0998	0
	(2,d,0)	685.3848	1.0079	-0.1125	0	0
	(2,d,1)	689.2045	0.2374	0.5537	-0.7179	0
	(0,d,2)	690.4923	0	0	-0.4031	-0.1033
	(1,d,2)	689.6586	0.8752	0	-0.09599	0.00789
	(2,d,2)	693.4508	0.2647	0.5224	-0.6602	0.0278
	(0,d,0)	684.5347	0	0	0	0

0	0	0	0	683.9711	(0,d,0)	Fracdiff d =0.4796818
0	0	0	0.267334	683.343	(1,d,0)	
0	-0.17326	0	0	686.7703	(0,d,1)	
0	0.10061	0	0.849300	688.1882	(1,d,1)	
0	0	0.07204	0.757320	688.1884	(2,d,0)	
0	0.36433	-0.21914	1.115947	691.9314	(2,d,1)	
-0.01881	-0.18175	0	0	691.0024	(0,d,2)	
0.071529	0.102481	0	0.8688244	694.9398	(1,d,2)	
0.145399	-0.50702	0.51848	0.26349	696.0568	(2,d,2)	
0	0	0	0	683.7226	(0,d,0)	Whittle d= 0.447
0	0	0	0.270140	683.489	(1,d,0)	
0	-0.17427	0	0	686.7097	(0,d,1)	
0	0.09708	0	0.850729	688.0461	(1,d,1)	
0	0	0.069864	0.76193	688.0466	(2,d,0)	
0	0.36063	0.21949-	1.11698	691.7841	(2,d,1)	
-0.01905	-0.182854	0	0	690.938	(0,d,2)	
0.070923	0.098695	0	0.869811	694.8997	(1,d,2)	
0.142767	-0.51034	0.519013	0.26486	695.7113	(2,d,2)	

لوحظ من خلال الجدول (4) عند الفرق الكسري $d=0.2484858$ وبالمفاضلة بين النماذج حسب معيار (BIC) تبين ان النموذج $ARFIMA(1,d,0)$ هو افضل نموذج بطريقة R/S وذلك لكونه يمثل اقل قيمة معيار (BIC) حيث كانت القيمة (682.7829) ، وعند الفرق الكسري $d=0.2232992$ تبين ان النموذج $ARFIMA(1,d,0)$ هو افضل بطريقة Empirical Hurst exponent والذي يمثل اقل قيمة معيار (BIC) حيث كانت القيمة (682.6296) ، وعند الفرق الكسري $d=0.0174121$ تبين ان النموذج $ARFIMA(1,d,0)$ هو افضل بطريقة Theoretical Hurst exponent والذي يمثل اقل قيمة معيار (BIC) حيث كانت القيمة (681.2809) ، وعند الفرق الكسري $d=0.4796818$ حيث تبين ان النموذج $ARFIMA(1,d,0)$ بطريقة Fracdiff هو افضل نموذج حيث يمثل اقل قيمة لمعيار (BIC) وعند الفرق الكسري $d=0.447$ حيث تبين ان النموذج $ARFIMA(1,d,0)$ بطريقة Whittle هو افضل نموذج حيث يمثل اقل قيمة لمعيار (BIC) .

سنقوم الآن باستخدام طريقة التقدير ذات المرحلة الواحدة وذلك من خلال تحديد الفرق الكسري (d) وتقدير معاملات النموذج $AR(p)$ ، $MA(q)$ كما هو موضح بالجدول (5)

(6) تقدير نماذج $ARFIMA$ بطريقة المرحلة الواحدة

تقدير d	النموذج	BIC	1θ	θ_2	θ_1	2θ
0.48351315	(0,d,0)	683.3659	0	0	0	0
-0.06275898	(1,d,0)	680.7693	0.852632	0	0	0
0.4410549	(0,d,1)	686.6982	0	0	-0.39169	0
-0.168688	(1,d,1)	684.1877	0.87656	0	-0.10451	0
-0.2052839	(2,d,0)	683.9041	1.01909	-0.12086	0	0
-0.1352635	(2,d,1)	688.2598	0.24919	0.24919	-0.70983	0
0.4179445	(0,d,2)	690.8766	0	0	-0.41591	-0.10167
-0.1617120	(1,d,2)	688.4897	0.87613	0	0.0095	0.00702
-0.10887092	(2,d,2)	692.6679	0.27485	0.514911	-0.65542	0.02631

يلاحظ من معطيات الجدول السابق وعند المقاضلة بين النماذج حسب معيار (BIC) تبين ان النموذج (0, -0.06275898, 1) هو افضل نموذج وذلك لكونه يمثل اقل قيمة معيار (BIC) حيث كانت القيمة هي (680.7693) .

(7) مقارنة نماذج ARFIMA بطريقة المرحلتين والمرحلة الواحدة

MAE	RMSE	MSE	2θ	θ1	θ2	1θ	BIC	النموذج
87.63704	109.5444	11999.98	0	0	0	0.8881	682.7829	(1, 0.2484858,0)
87.48044	109.2756	11941.16	0	0	0	0.85729	682.6296	(1, 0.2232992,0)
85.50975	106.9305	11434.13	0	0	0	0.85315	681.2809	(1, 0.0174121,0)
88.09449	111.2834	12383.99	0	0	0	0.267334	683.343	(1, 0.4796818,0)
88.05866	111.1203	12347.71	0	0	0	0.270140	683.489	(1,0.447,0)
84.8198	106.228	11284.38	0	0	0	0.852653	680.7693	(1, - 0.06275898,0)

بمقارنة نماذج تقدير الذاكرة الطويلة ذات المرحلتين مع الذاكرة الطويلة ذات المرحلة الواحدة يتبين ان افضل نموذج هو (0, -0.06275898, 1) بطريقة التقدير ذات المرحلة الواحدة حيث تمتلك اقل معيار BIC وكذلك اقل قيمة لل MSE ,RMSE,MAB .

جدول (8) تقدير معلمات النموذج (0, -0.06275898, 1)

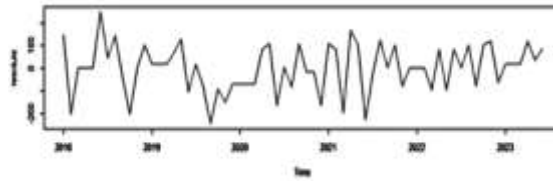
p.value	التقدير	المعلمات	النماذج
0.0000	0.852653	θ1	AR (1)

كما أن الصيغة الرياضية للنموذج (1, -0.06275898, 0) ARFIMA تكون كالتالي :

$$(1 - 0.852653L)(1 - L)^{-0.06275898}Y_t = \varepsilon_t$$

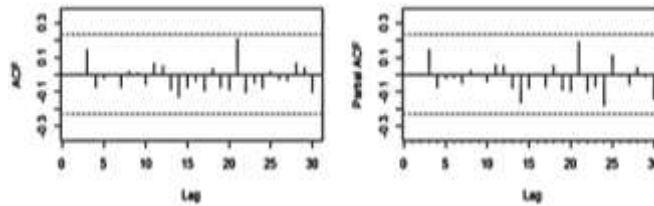
• تحليل البواقي ARFIMA(1, -0.06275898, 0):

يعتبر تحليل البواقي أهم مرحلة لمعرفة مدى ملائمة النموذج ARFIMA(1,0.0000458,0) لاستخدامه في التنبؤ ومن خلال الشكل (16) لرسم البواقي نلاحظ أنها تتأرجح بشكل ثابت حول الصفر .



الشكل (13) منحني البواقي لبيانات السلسلة المدروسة

من الصعوبة الحكم على شكل البواقي لذلك نستعين برسم دالة ACF ودالة PACF .



(A)

(B)

الشكل (14) دالة الارتباط الذاتي للبواقي B دالة الارتباط الذاتي الجزئي للبواقي

ومن الشكل السابق نجد أن جميع معاملات الارتباط الذاتي والارتباط الذاتي الجزئي للبواقي تقع ضمن فترة حدود الثقة ، وهذا يعني عدم ارتباط بواقي هذا النموذج مع بعضها (الخطاء الناتجة عشوائية) مما يدل على سكون البواقي .

اختبار $Ljung_Box$ أدناه جدول يبين نتائج الاختبار لبواقي النموذج $(1, -0.06275898, 0)$:جدول (9) اختبار $Ljung_Box$

النموذج	Q*	p-value
$(1, -0.06275898, 0)$	2.8209	0.9853

نلاحظ أن القيمة المعنوية $p\text{-value} > 0.05$ ، وهذا يعني أن البواقي مستقلة وأن النموذج قد أجتاز التشخيص .

5.التنبؤ: التنبؤ بالنموذج $(1, -0.06275898, 0)$ ARFIMA

بعد اختيار $(1, -0.06275898, 0)$ ARFIMA كأفضل نموذج في التشخيص وأفضل نموذج ممثل للبيانات قمنا باستخدامه للتنبؤ وسيكون التنبؤ ل 12 قيمة ومن ثم مقارنتها مع القيم الحقيقية وكما مبين في الجدول أدناه

الجدول (10) القيم المتنبأ بها ل 12 قيمة

القيمة المتنبأ بها $(\hat{y}_t)$	القيمة الحقيقية (y_t)	التسلسل
994	1100	1
990	1150	2
984	1250	3
980	1100	4
1038	1100	5
1024	1100	6
1012	1200	7
1002	1200	8
1058	1400	9
1082	1500	10
1112	1500	11
1152	1500	12

نلاحظ من الجدول (10) أن القيم المتنبأ بها اعطت قيم قريبة من القيم الحقيقي

الاستنتاجات والتوصيات:

1. من خلال النتائج في الجانب التطبيقي والتطبيق على البيانات الخاصة بأسعار الموز المستورد في مدينة الموصل وبشكل اسبوعي لبناء افضل نموذج ARFIMA بالتنبؤ استنتجنا أن السلسلة الزمنية لأسعار الموز غير مستقرة وتم الكشف عن ذلك من خلال رسم الاتجاه العام لكل من دالتي (ACF, PACF) وكذلك الاختبارات الخاصة بفحص استقراريه السلسلة (PP,ADF).
2. يلاحظ من النتائج أيضا أن السلسلة الزمنية تمتلك ذاكرة طويلة وتم ذلك بأجراء عدة اختبارات (R/S, THE Empirical hurst, Whittle, Fracdiff, dsprio, Gph, exponent)، ومن خلال طرق تقدير الفرق الكسري لنموذج ARFIMA ظهر فشل (Gph, dsprio) حيث كانت قيمة $d > 0.5$ وتساوي (1.342017, 0.6107118) على التوالي ونلاحظ بنفس الوقت نجاح بقية الاختبارات وطرق التقدير باثبات خاصية الذاكرة الطويلة .
3. من خلال أخذ الفروق الكسرية وأدخالها على السلسلة الزمنية الأصلية وبعد اجراء اختبار (pp) للتحقق من الاستقراريه نلاحظ استقرار السلسلة حيث كانت قيمه p_value اصغر من 0.05 ، ومن خلال المقارنة بين الطرق والنماذج ظهر لنا أن افضل نموذج هو (-1, 0.06275898, 0) بطريقة التقدير بمرحلة واحدة وذلك لامتلاكه اقل قيمه معيار BIC والذي يثبت تفوق النموذج على النماذج الاخرى وكذلك اقل قيمه MSE, RMSE, MAE. ومن ذلك يمكن استنتاج إمكانية استخدام النموذج بالتنبؤ بالقيم المستقبلية .
4. أن جودة النتائج التي تم الحصول عليها من خلال استخدام الاختبارات وطرق التقدير والتنبؤ تشير الى إمكانية استخدامها في تحليل البيانات والحصول على تنبؤات دقيقة وخاصة في مجال الزراعة وأسعار المنتجات الزراعية .
5. نوصي باستخدام طرق اخرى للتقدير منها الطرق الشبه معلمية والطرق اللامعلمية والتطبيق على أنواع مختلفة من البيانات الزراعية والصناعية والتجارية .

References

1. Akaike, H.J.B., A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. 1979. 66(2): p. 237-242.
2. Belmukaddam, Mustafa, and Ben, Ayaa, Wafaa, predicting unemployment in light of the presence of the Corona epidemic in Algeria using long memory models (ARFIMA) during the time period from 2008 to 2020. 2022. 8(2)
3. Benguesmi, T., Using seasonal time series models to forecast electric energy sales – a case study of the National Electricity and Gas Company. 2014, Mohamed Khider Biskra University.
4. Berri, Adnan. Methods of statistical forecasting (2002), Part One, King Saud University.
5. Box, G.E.P. and D.A. Pierce, Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. Journal of the American Statistical Association, 1970. 65(332): p. 1509-1526.
6. Brockwell, P.J. and R.A. Davis, Time series: theory and methods. 2009: Springer science & business media.
7. Bryce, R. M. and Sprague, K. B. (2012). "Revisiting detrended fluctuation analysis", Scientific Reports 315(2)
8. Ceballos, R.F. and F.F.J.a.p.a. Largo, On the estimation of the Hurst exponent using adjusted rescaled range analysis, detrended fluctuation analysis and variance time plot: A case of exponential distribution. 2018.
9. Dickey, D. A., and Fuller, W. A. (1979). "Distribution of the estimators for Autoregressive Time Series With a unit Root", Journal of the American Statistical Association, N. 74: pp .427-431
10. Dickey, D.A. and W.A.J.J.o.t.A.s.a. Fuller, Distribution of the estimators for autoregressive time series with a unit root. 1979. 74(366a): p. 427-431.
11. Fox, R. and M.S. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. The Annals of Statistics, 1986. 14(2): p. 517-532.
12. Geweke, J. and S.J.J.o.t.s.a. Porter-Hudak, The estimation and application of long memory time series models. 1983. 4(4): p. 221-238. Journal of time series analysis
13. Granger, C.W. and R.J.J.o.t.s.a. Joyeux, An introduction to long-memory time series models and fractional differencing. 1980. 1(1): p. 15-29
14. Haslett, J. and A.E.J.J.o.t.R.S.S.S.C. Raftery, Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. 1989. 38(1): p. 1-21. Journal of the Royal Statistical Society: Series C

15. Hosking, J.R.M., Fractional Differencing. *Biometrika*, 1981. 68(1): p. 165-176.
16. Hurst, H.E.J.T.o.t.A.s.o.c.e., Long-term storage capacity of reservoirs. 1951. 116(1): p. 770-799. *Transactions of the American society of civil engineers*.
17. Lo, A., Long-term Memory in Stock Market Prices, *Econometrica*. 1991. *Journal of the Econometric Society*
18. MacKinnon, J.J.L.-r.e.r., Critical values for cointegration tests. 1991. 13.
19. MADOURI, H. and M.J.e.-B.R. MKIDICHE, A Comparative Study of ARFIMA and Artificial Neural Networks to Forecast Exchange Rate of Dinar Algerian. 2017. 17(1): p. 159-171.
20. Peng, C., S. Buldyrev, and M.J.P.R. Simons, Nature and fractals. 1994. *Physics Rev.* 168,.
21. Reisen, V.A.J.J.o.T.S.A., Estimation of the fractional difference parameter in the ARIMA (p, d, q) model using the smoothed periodogram. 1994. 15(3): p. 335-350. *J Journal of Time Series Analysis*
22. Safitri, D. and D. Ispriyanti. Gold price modeling in Indonesia using ARFIMA method. in *Journal of Physics: Conference Series*. 2019. IOP Publishing.
23. Shaarawy, Samir. Introduction to modern time series analysis. 2005. King Abdulaziz University.
24. Shang, H.L.J.J.o.T.S.E., A comparison of Hurst exponent estimators in long-range dependent curve time series. 2020. 12(1): p. 20190009, *Journal of Time Series Econometrics*
25. Taqqu, M.S. and V. Teverovsky, On estimating the intensity of long-range dependence in finite and infinite variance time series. A practical guide to heavy tails: statistical techniques and applications, 1998. 177: p. 218.
26. Tuama, S.A.J.A.-A.U.j.o.E. and A. Sciences, Using Analysis of Time Series to Forecast numbers of The Patients Malign

Modeling and analyzing banana prices in the city of Mosul using the ARFIMA model “Predictive Market Study”

Rehab Talal Ahmed¹ & Omar Salim Ibrahim

Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul , Iraq

Abstract: This study examined the use of ARFIMA models to forecast imported banana prices in the city of Mosul, based on data obtained from the Directorate of Agriculture in Nineveh for the period from 2018 to 2023. Several methods were used to estimate long memory and determine the fractional differencing parameter (d), including single-stage methods such as the maximum likelihood (EML) method used in this research, and two-stage methods such as the Geweke-Porter-Hudak (GPH) estimator, the dsprio (Smoothed periodogram estimation) method, the Fracdiff method, the Rescaled Range (R/S) method, and the Whittle estimator. The model was built by verifying the presence of long memory in the time series through several tests, and then estimating the fractional differencing parameters. The single-stage ARFIMA (1,-0.06275898,0) model outperformed the other methods based on criteria such as BIC, MSE, RMSE, and MAE. The model passed diagnostic tests and was used for forecasting banana prices, with the aim of clarifying the steps for constructing an appropriate model.

Keywords: Long memory, ARFIMA , Forecasting , Methods for estimating the fractional differential parameter d.