2020

Applying some hybrid models for modeling bivariate time series assuming different distributions for random error with a practical application

تطبيق بعض النماذج الهجينة (Hybrid Models) لنمذجة السلاسل الزمنية ثنائية المتغيرات بافتراض توزيعات مختلفة للخطأ العشوائي مع تطبيق عملي

أ.م. د. فراس احمد محمد / جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء drfirasmohana@gmail.com
الباحث / مؤمن عباس موسى/جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء saidmoamen@gmail.com

P - ISSN 2518 - 5764
E - ISSN 2227 - 703X

Received:18/9/2019
Accepted: 28/10/2019

مستخلص البحث:

أصبح لعملية النمذجة والتنبؤ بالسلاسل الزمنية ثنائية المتغيرات (Bivariate Time Series) مجالاً واعداً للدراسات التطبيقية في الآونة الأخيرة. ولهذا الغرض، يعتبر انموذج الانحدار الذاتي والاوساط المتحركة بمدخلات خارجيَّة المَنْشأ (ARMAX) الخطي أكثر التقنيات المستخدمة على نطاق واسع خلال السنوات القليلة الماضية في عملية النمذجة والتنبؤ بهذا النوع من البيانات. ومن اهم الافتراضات الاساسية لهذا الأنمُوذج هي الصفة الخطية وتجانس تباين الخطأ العشوائي للأنمُوذج الملائم وفي التطبيق العملي غالباً ما يتم انتهاك هذان الافتراضان لذلك تم تطبيق نماذج الانحدار الذاتي المشروطة بعدم تجانس تباين الخطأ العشوائي (ARCH &GARCH) غير الخطية مع توظيف أنموذج (GARCH) بمدخلات خارجيَّة المَنْشَأ أي أنموذج (GARCHX) من أجل التحليل والسيطرة على التغيرات الزمنية او التقلبات التي تحدث في التباين الشرطي للأنمُوذج الخطي. وبما أن مشاهدات السلاسل الزمنية نادراً ما تكون ذات مركبة خطية فقط حيث عادة ما تكون مُتَضَمّنه كلتا المركبتين معاً. لذا يعتبر الغرض الرئيسي لهذا البحث في توظيف تقنية النماذج الهجينة (Hybrid Models) طبقاً الى منهجية (Zhang) للنماذج الهجينة وذلك للدمج بين مركبة التنبوءات الخطية لأفضل أنمُوذج خطي من نماذج (ARMAX) ومركبة التنبوءات غير الخطية لأفضل أنمُوذج غير خطي من نماذج (ARCH ,GARCH & GARCHX) وبالتالي زيادة كفاءة ودقة اداء التنبؤ بالقيم المستقبلية للسلسلة الزمنية .

حيث تم بناء واقتراح توظيف عدة نماذج هجينة، ويختص هذا البحث في نمذجة وبناء النماذج الهجينة (ARMAX − GARCH) و (ARMAX − GARCHX) بافتراض ثلاث توزيعات مختلفة للخطأ العشوائي وهي التوزيع الطبيعي القياسي (Gaussian distribution)، توزيع (Student − t)، وكذلك توزيع الخطأ العام (Generalized error distribution) وهذان التوزيعان الاخيران تم تطبيقهما لغرض السيطرة على خصائص التوزيعات ذوات الذيل الثقيل (Heavy − tail) والتي تكون بقمة منحنى مرتفعة مقارنة بالتوزيع الطبيعي. واعتمد البحث هذا منهجية حديثه في تقدير معلمات الانمُوذج الهجين حيث تم تقدير معلمات النماذج الهجينة بتطبيق أسلوب مّنهجية ذاتَ المرحلتين (Two − Step Procedure) ففي المرحلة الأولى تم تقدير معلمات الأنمُوذج الخطي بأتباع ثلاث طرائق مختلفة وهي طريقة المربعات الصغرى الاعتيادية (OLS)، طريقة المربعات الصغرى التكرارية بتوظيف عامل التغاضي الاسي (RLS − EF)، وطريقة خطأ التنبؤ التكرارية (RPEM) ، وفي المرحلة الثانية تم تقدير معلمات الأنمُوذج غير الخطي باستعمال طريقة الإمكان الأعظم (MLE) وبتوظيف خوارزمية التحسين العددية (BHHH algorithm).

*البحث مستل من رسالة ماجستير

Journal of Economics and Administrative Sciences
2020; Vol. 26 , No.117
Pages: 442- 479

النماذج الهجينة المذكورة تم تطبيقها لنمذجة العلاقة بين السلسلة الزمنية للمدخلات الخارجية (x_t) والمتمثلة بمتغير سعر الصرف والسلسلة الزمنية الاصلية (y_t) والمتمثلة بمتغير معدل البطالة في الولايات المتحدة الامريكية في الفترة الزمنية من شهر يناير(2000) حتى شهر ديسمبر(2017) بواقع (216) مشاهدة، وايضاً لعملية التنبؤ خارج العينة بمعدل البطالة بالقيم الاثنتا عشر الاخيرة من عام (2018). كما تم أجراء تقييم اداء التنبؤ بين النماذج الهجينة والأنمُوذج المُنْفرد المُتَنَافِس بالاعتماد على مقاييس دوال الخسارة لدقة التنبؤ وهي متوسط مطلق الخطأ النسبي $(MAPE)$، ومقياس متوسط مطلق الخطأ (MAE)، وكذلك مقياس $(Q-LIKE)$ الحصين والمقترح توظيفهُ. وبالاعتماد على المقاييس الاحصائية فقد اظهرت نتائج البحث تفوق النماذج الهجينة حيث حَسّنة من جُودَة وكَفَاءَة الانمُوذج المُنْفرد وهذا ما يُعَزّز من قدرة وكفاءة النماذج الهجينة في القدرة على التنبؤ بالقيم المستقبلية. واتضح أن الانمُوذج $ARMAX\ (2,1,1,0)-GARCHX(1,1,1)$ الهجين عند اتباع الخطأ العشوائي لهُ توزيع الخطأ العام (GED) كان الأنمُوذج الأمثل في عملية نمذجة السلاسل الزمنية ثنائية المتغيرات قيد الدراسة والأكثر كفاءة في القدرة على التنبؤ المستقبلي بالمقارنة مع النماذج المعنية بالمنافسة وهي أنمُوذج $ARMAX\ (2,1,1,0)$ بمفرده ونظيرهُ أنمُوذج $ARMAX\ (2,1,1,0)-GARCH(1,1)$ الهجين ويعود السبب في ذلك لحصولهِ على أقل القيم لمقاييس المفاضلة. حيث تم استعمال حزم برمجية مختلفة وهي $(MATLAB(2018a), SAS\ 9.1, R\ 3.5.2\ and\ EViews\ 9)$ من قبل الباحث لتحليل البيانات قيد الدراسة .

المصطلحات الرئيسة للبحث/أنمُوذج $(ARMAX-GARCH)$، أنمُوذج $(ARMAX-GARCHX)$، التوزيع الطبيعي القياسي، توزيع $(Student's-t)$، توزع الخطأ العام (GED)، سعر الصرف، معدل البطالة.

1. المقدمة Introduction

تعتبر عملية التنبؤ بالسلاسل الزمنية واحدة من أهم اساليب التحليل الاحصائي والذي يستخدم كأساس للتخطيط في العديد من مجالات العلوم التطبيقية المختلفة. حيث يعد تحليل السلاسل الزمنية بمجالاتها المختلفة أحد الاساليب الرئيسية للبحث والتطبيق. وتعرف السلسلة الزمنية على انها عبارة عن مجموعة من المشاهدات او القياسات المرتبطة مع بعضها البعض لعدد من الظواهر التطبيقية ومنها (الاقتصادية، الاجتماعية، الهندسية، الطبية، الصناعية،...) على فترات زمنية متتابعة قد تكون متصلة او منفصلة عادة ما تكون متساوية في الطول [15]. وشهدت العقود القليلة الماضية اهتماماً متزايداً بالتطورات النظرية والتجريبية في بناء نماذج السلاسل الزمنية وتطبيقاتها المهمة في التنبؤ أذ تلعب قواعد التنبؤ دوراً مهماً في العديد من المجالات مثل الأعمال التجارية والصناعة والمنظمات الحكومية الدولية. وبصرف النظر عن عدة عوامل اقتصادية مثل سعر الفائدة، الناتج المحلي الإجمالي GDP، والتضخم وغيرها من العوامل الاقتصادية، يعد سعر الصرف Exchange rate والبطالة Unemployment أحد العوامل الرئيسية التي تحافظ على أهمية تقدم النمو الاقتصادي .

ولنمذجة هذين العاملين قدمت أدبيات السلاسل الزمنية مختلف نماذج السلاسل الزمنية الخطية، ويعد انمُوذج الانحدار الذاتي والاوساط المتحركة بمدخلات خارجيُّة $(Autoregressive\ Moving\ Average\ with\ Exogenous\ Variable)$ والذي يشار اليه اختصاراً بالرمز $(ARMAX)$ أحد الأساليب الرئيسية والمستخدمة بشكل شائع لتحليل بيانات السلاسل الزمنية ثنائية المتغيرات. ويعتبر أنمُوذج $(ARMAX)$ ثاني أكثر النماذج الخطية استخداماً بعد أنموذجي الانحدار الذاتي بمدخلات خارجية (ARX) وأنموذج المتوسط المتحرك بمدخلات خارجية (MAX) . ويعتبر أنمُوذج $(ARMAX)$ الاكثر مرونة لتميزهُ في خواصه الإحصائية في عملية النمذجة والتنبؤ حيث يجمع ما بين أنمُوذجي الانحدار الذاتي (AR) والاوساط المتحركة (MA) مع مدخلات خارجية إضافية تسمى بالمتغير الخارجي

(exogenous variable) [30]. ومن القيود الرئيسية لنماذج ARMAX هي البنية الخطية وتجانس تباين الاخطاء للأنمُوذج، وفي حالة عدم توفر هذين الافتراضين فأن ذلك يؤدي الى تَقييد تطبيق أنمُوذج ARMAX على بيانات السلاسل الزمنية. بالإضافة الى ذلك ليس لدى النماذج الخطية للسلاسل الزمنية أي إمكانية لوصف أي تقلب زمني Volatility يحدث في التباين الشرطي الفعلي في بيانات السلاسل الزمنية الحقيقة أو في بواقي الأنمُوذج الخطي.

وللتغلب على هذه المشكلة أقترح العالم (Engle,1982) أنموذج الانحدار الذاتي المشروط بعدم تجانس تباين الخطأ العشوائي (ARCH model) لغرض التحليل والسيطرة على التذبذبات الزمنية في بيانات السلاسل الزمنية ، والذي يصف التباين الشرطي للسلسلة الزمنية في الزمن الحالي t كدالة للأحجام الحقيقية للأخطاء العشوائية في الزمن السابق ε^2_{t-j} . وان أنمُوذج (ARCH) يكون ملائماً عند اتباع الخطأ العشوائي لسلسلة زمنية لأنموذج الانحدار الذاتي (AR) وبعبارة اخرى فأن أنمُوذج (ARCH) هو أنمُوذج (AR) بتباين شرطي غير متجانس (Heteroskedacity)، بالإضافة الى هذا الافتراض فأن هذا الانموذج غالباً ما يتطلب العديد من المعلمات المطلوب تقديرها لبناء الأنمُوذج وبالتالي وصف عملية التذبذبات الزمنية بشكل يتناسب مع الظاهرة قيد الدراسة. ومن أجل التَغلّب على صّعُوبات أنمُوذج (ARCH) أقترح العالم (Bollerslev,1986) أنمُوذج أكثر شمولية وهو أنموذج الانحدار الذاتي العام المشروط بعدم تجانس تباين الخطأ العشوائي (GARCH model)، والذي يكون ملائماً عند اتباع الخطأ العشوائي لسلسلة زمنية لأنمُوذج الانحدار الذاتي (ARMA) وبعبارة اخرى فأن أنمُوذج (GARCH) هو أنمُوذج (ARMA) بتباين شرطي غير متجانس [12,5].

وتؤكد العديد من الدراسات التجريبية أن النماذج غير الخطية لها أداء جيد للتنبؤ على المدى الطويل، في حين أن النماذج الخطية مناسبة للتنبؤ على المدى القصير، وكذلك نادراً ما تكون بيانات السلاسل الزمنية الحقيقية تتكون من مركب خطي فقط أو غير خطي فقط بطبيعتها وعادة ما تتضمن كلتا المركبتين معاً. ولهذا الغرض يهدف هذا البحث الى استعمال منهجية التهجين (Hybridization methodology) بين النماذج الخطية وغير الخطية لبناء أنمُوذج هجين (Hybrid model) ليُعَزّز من قدرة وكفاءة النماذج المنفردة. ولعملية بناء الأنمُوذج الهجين يتم في هذا البحث استعمال كلاً من أنمُوذج (GARCH) وأنمُوذج (GARCHX) غير الخطية في منهجية التهجين بالدمج مع أنمُوذج (ARMAX) الخطي.

بالإضافة الى ذلك، ليس هناك ما يضمن أن الأنمُوذج المحدد النهائي سوف يعطي قيم تنبؤ مثالية إذا كان هناك بعض العوامل المؤثرة مثل نوع الأنمُوذج وتوزيعات الأخطاء العشوائية المستخدمة في طريقة التهجين، وللتعامل مع هذه المشكلة تمت عملية التهجين بين نماذج التنبؤ المختلفة معاً بنماذج مختلفة وبافتراض ثلاث توزيعات مختلفة للخطأ العشوائي للأنمُوذج الهجين وهي (التوزيع الطبيعي القياسي، توزيع (Student − t)، وتوزيع الخطأ العام (GED)) وذلك لغرض السيطرة على خصائص التوزيعات ذوات الذيل الثقيل (Heavy − tail) في بواقي الأنمُوذج الخطي والتي تكون بقمة منحني مرتفعة (Leptokurtic) مقارنة بالتوزيع الطبيعي وبالتالي زيادة كفاءة النماذج الهجينة في القدرة على التنبؤ .

ومن الاسباب الرئيسية الاخرى لعملية التهجين بين النماذج المنفردة هو أن الكثير من المشاكل العالمية عادة ما تكون معقدة بطبيعتها وهذا يعني أن أي أنمُوذج مُنفرد بحد ذاته لا يكون قادر على التقاط الانماط المختلفة للسلاسل الزمنية بانتظام، وهذا يعني لا يوجد أنمُوذج سلاسل زمنية خطي أو غير خطي هو الخيار الامثل لعملية التنبؤ بالقيم المستقبلية. وبالتالي، فإن دمج النماذج الفردية المختلفة له تأثير فعال وهام جداً لزيادة فرصة السيطرة على الانماط المختلفة وتحقيق أداء أكثر كفاءة في التنبؤ. حيث قام العديد من الباحثين ومنهم (Zhao,2009)[44] و(Hickey et al.,2012) [20] وكذلك (Porshnev et al.,2016)[33] بدراسة هذا النوع من النماذج الهجينة وأظهروا تحسن أداء الأنمُوذج الهجين في دقة التنبؤ.

أن جميع الدراسات حول عملية التهجين بين أنمُوذج (ARMAX) ونماذج (GARCH) و (GARCHX) لم تتطرق الى استعمال منهجية (Zhang) للنماذج الهجينة، لذلك أحد اهم اهداف هذا البحث متمثل في عملية التهجين بين مركبة التنبوءات الخطية لأفضل أنمُوذج خطي من نماذج (ARMAX) ومركبة التنبوءات غير الخطية لأفضل أنمُوذج غير خطي من نماذج (ARCH , GARCH & GARCHX) وفقاً لهذه المنهجية . تطبيقياً، في هذه الدراسة تم اختبار عدد من النماذج الهجينة، حيث تم اختيار الأنمُوذج الأمثل استناداً إلى معايير اختيار النماذج وهي معيار معلومات اكاكي (AIC)، معيار معلومة بيز(BIC)، بالإضافة الى معيار خطأ التنبؤ النهائي (FPE). وتم استعمال أسلوب طرائق ذات المرحلتين (two − step procedure) لتقدير معلمات النماذج الهجينة (ARMAX − GARCH) و (ARMAX − GARCHX) مع ثلاثة توزيعات مختلفة إلى جانب أنمُوذج (ARMAX) المنُفرد على مجموعة البيانات الحقيقية.

وقد تم استعمال مقاييس مختلفة لدراسة دقة التنبؤ والمفاضلة بين النماذج المنُفردة والهجينة وفي ومنها متوسط مطلق الخطأ (MAE)، متوسط مطلق الخطأ النسبي (MAPE)،ومقياس (Q − LIKE) الحصين . حيث تم استخدام برامج مختلفة (MATLAB(2018a), SAS 9.1, R 3.5.2 and EViews 9) وذلك حسب توفر الحزمة البرمجية للأسلوب الاحصائي المستعمل لتحليل البيانات قيد الدراسة .

2. هدف البحث Research objective

أن الهدف الرئيسي من هذا البحث هو بناء نماذج هجينة (Hybrid Models) بافتراض توزيعات مختلفة للخطأ العشوائي والتي تجمع بين الصفات الخطية وغير الخطية للسلاسل الزمنية.
ولتحقيق هذا الهدف ينبغي ما يلي:

1- بناء أنمُوذج الانحدار الذاتي والاوساط المتحركة بمدخلات خارجية (ARMAX) واستعمال طرائق لخوارزميات مختلفة لتقدير معلمات هذا الأنمُوذج والمقارنة بين تلك الطرائق للحصول على الطريقة المثلى في عملية التنبؤ بمعدل البطالة والحصول على بواقي الأنمُوذج الامثل.

2- بناء أنمُوذج الانحدار الذاتي العام المشروط بعدم تجانس تباين الخطأ (GARCH) بتوزيعات مختلفة للخطأ العشوائي والمَبْنِيّ على أساس السلسلة الزمنية لبواقي $(\hat{\varepsilon}_t)$ الخاصة بأنمُوذج (ARMAX) الامثل.

3- بناء أنمُوذج (GARCH) بمدخلات خارجية التأثير، أي بناء أنمُوذج (GARCHX) بتوزيعات مختلفة للخطأ العشوائي والمَبْنِيّ على أساس السلسلة الزمنية للمتغير الخارجي (x_t) والسلسلة الزمنية لبواقي $(\hat{\varepsilon}_t)$ لأنمُوذج (ARMAX) الامثل.

4- فيما بعد يتم بناء النماذج الهجينة (ARMAX − GARCH) و (ARMAX − GARCHX) وبتوزيعات مختلفة للخطأ العشوائي، بالاعتماد على منهجية (Zhang) للنماذج الهجينة في السلاسل الزمنية .

5- المقارنة بين النماذج الهجينة والمنُفردة باستعمال مقاييس احصائية مختلفة لاختيار الأنمُوذج الأمثل واستعمالهُ للحصول على القيم التنبؤية النهائية لمعدل البطالة.

3. الجانب النظري: المنهجية والنماذج Methodology and models

في هذا القسم من البحث، سيتم تصنيف نوعين من نماذج السلاسل الزمنية للتنبؤ المستخدمة في طريقة التهجين:

1.3 نماذج المتوسط المشروط Conditional Mean Models

تتم عملية نمذجة معادلة المتوسط الشرطي للأنموذج الهجين باستعمال (mean equation) واحد من اهم النماذج الخطية للسلاسل الزمنية ثنائية المتغيرات وهو أنموذج (ARMAX) أحادي المدخلات - أحادي المخرجات (Single − Input − Single − Output).ويمكن أن يعبر عن أنموذج (ARMAX) رياضياً لنمذجة العلاقة بين مدخلات ومخرجات النظام باستعمال معادلة الفروق (difference equation) من خلال الصيغ الرياضية (1 و 2) الاتية [39,36]:

$$y_t = \mu + \sum_{i=1}^{n_p} \varphi_i \, y_{t-i} - \sum_{j=1}^{n_q} \theta_j \, \varepsilon_{t-j} + \sum_{\kappa=1}^{n_b} \phi_\kappa \; x_{t-n_k} + \varepsilon_t \qquad , \quad \ldots\ldots(1)$$

أو باستعمال الشكل المركب (Compact form) وكما يلي :

$$\Phi_{n_p}(L)\, y_t = \phi_{n_b}(L)\, x_{t-n_k} + \Theta_{n_q}(L)\, \varepsilon_t \qquad , \qquad \ldots\ldots(2)$$

تمثل الصيغتين اعلاه معادلة المتوسط الشرطي للنماذج الهجينة ويشار اليها اختصاراً بالرمز $ARMAX(n_p, n_q, n_b, n_k)$ ويمكن تعريف الرموز الرياضية لهذه المعادلات كالاتي :

y_t : متغير الانحدار الذاتي أو مخرجات النظام (system output).

x_t : متغير خارجي المنشأ أو مدخلات النظام (system input).

ε_t : متغير المتوسطات المتحركة أو الخطأ العشوائي.

n_p ، n_q ، n_b و n_k هي رتب معلمات انموذج الانحدار الذاتي، المتوسطات المتحركة، والمدخلات الخارجية، ورتبة زمن التأخير (Delay time) على التوالي [1].

φ_i ، ϕ_κ و θ_j تمثل متجه لمعلمات أنموذج الانحدار الذاتي، المتوسطات المتحركة، والمدخلات الخارجية على التوالي .

$\Phi_{n_p}(L)$: تمثل متعددة الحدود لمعلمات أنموذج الانحدار الذاتي من الرتبة n_p .

$\Theta_{n_q}(L)$: تمثل متعددة الحدود لمعلمات أنموذج المتوسط المتحرك من الرتبة n_q .

$\phi_{n_b}(L)$: تمثل متعددة الحدود لمعلمات أنموذج المتغير الخارجي من الرتبة n_b . وتعطى كل منهم بالصيغ الرياضية الاتية [8,36]:

$$\Phi_{n_p}(L) = 1 + \varphi_1 L + \varphi_2 L^2 + \cdots + \varphi_{n_p} L^{n_p} = 1 + \sum_{i=1}^{n_p} \varphi_i \, L^i \qquad , \quad \ldots(2.a)$$

$$\Theta_{n_q}(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \cdots - \theta_{n_q} L^{n_q} = 1 - \sum_{j=1}^{n_q} \theta_j \, L^j \qquad , \quad \ldots(2.b)$$

$$\phi_{n_b}(L) = \phi_1 L^{-1} + \phi_2 L^{-2} + \cdots + \phi_{n_b} L^{n_b} = \sum_{k=1}^{n_b} \phi_k \, L^k \qquad , \quad \ldots(2.c)$$

وتعتبر نماذج $AR(n_p)$، $MA(n_q)$، وأنموذج $ARMAX(n_p, n_q)$ حالات خاصه من أنموذج $ARMAX(n_p, n_q, n_b, n_k)$. اما الرمز ε_t في المعادلة رقم (1 و 2) يمثل حد الخطأ العشوائي

[1] يتم التعرف على رتبة زمن التأخير لمعادلة المتوسط الشرطي للأنموذج الهجين باستعمال دالة الارتباط التقاطعية (cross − correlation function) بين سلسلة المدخلات x_t والمخرجات y_t .

للأنمُوذج الخطي وتتم عملية نمذجة هذا الخطأ العشوائي من خلال عملية التهجين وذلك باستعمال نماذج التباين الشرطي $(\mathbf{ARCH, GARCH\ and\ GARCHX})$ غير الخطية، والأقسام التالي توضح ذلك .

2.3 نماذج التباين المشروط Conditional Variance Models

تتم عملية نمذجة البواقي $(\mathbf{Residuals})$ لأنمُوذج $\mathbf{ARMAX}(\mathbf{n_p, n_q, n_b, n_k})$ الخطي لمعادلة المتوسط والمذكورة في معادلة رقم (1)، من خلال تطبيق النماذج غير الخطية $(\mathbf{ARCH, GARCH\ and\ GARCHX})$ للتباين الشرطي والتي تمثل معادلة التباين $(\mathbf{Variance\ equation})$ للأنمُوذج الهجين، والاقسام الاتية توضح هذه النماذج :

1.2.3 انموذج الانحدار الذاتي المشروط بعدم تجانس التباين من الرتبة r

AutoRegressive Conditional Heteroskedasticity $(\mathbf{ARCH}(r))$ model

أن نماذج السلاسل الزمنية الخطية غير قادرة على نمذجة وتفسير الذبذبات الزمنية للتباين الشرطي والتي تظهر في العديد من بيانات السلسلة الزمنية الحقيقية او في بواقي الأنمُوذج الخطي، وللتعامل مع هذه المشكلة أقترح العالم $(\mathbf{Engle, 1982})$ أنموذج $(\mathbf{ARCH})$ وهو أول أنمُوذج للارتباطات الذاتية المشروط بعدم تجانس التباينات ومن النماذج الرئيسية في نمذجة التذبذبات الزمنية $(\mathbf{Volatility})$ العالية والمنخفضة. فقبل اقتراح هذا الأنمُوذج لم يكن هناك أنمُوذج دقيق متاح للتنبؤ بالتباين الشرطي، حيث ان التنبؤات للتباين المشروط في الزمن t في هذا الأنمُوذج تعتمد على مجاميع الاوزان المرجحة بقيم مربعات سلسلة البواقي للفترات السابقة $(\boldsymbol{\varepsilon}^2_{t-j})$.

ويمكن تمثيل أنمُوذج $(\mathbf{ARCH})$ رياضياً لنمذجة بواقي $\boldsymbol{\varepsilon_t}$ الأنمُوذج الخطي باستعمال معادلة التباين الشرطي $(\mathbf{Varince\ equation})$ كالاتي [12,34,17]:

$$\boldsymbol{\varepsilon_t \backslash \mathcal{F}_{t-1} \sim N(0, h_t)}$$

$$\boldsymbol{\varepsilon_t = h_t \eta_t} \quad , \quad \mathbf{t = 1, 2, \ldots n} \quad , \quad \boldsymbol{\eta_t} \sim \mathbf{IID\ N(0,1)} \quad , \quad \ldots (3)$$

تمثل المعادلة رقم (3) اعلاه نمذجة عملية $(\mathbf{ARCH})$ ويمكن تعريف الرموز الرياضية لهذه المعادلة كالاتي :

$\boldsymbol{\varepsilon_t}$: تمثل متَسَلسِلة البواقي لعملية $(\mathbf{ARCH})$.

$\boldsymbol{\mathcal{F}_{t-1}}$: دالة لمجموعة المشاهدات والاخطاء العشوائية السابقة عند الزمن $(\mathbf{t-1})$.

$\boldsymbol{\eta_t}$: متَسَلسِلة من المتغيرات العشوائية التي تكون مستقلة ومتماثله والتي تفترض أن تتبع التوزيع الطبيعي القياسي بمتوسط $(\mathbf{0})$ وتباين $(\mathbf{1})$، $\{\boldsymbol{\eta_t} \sim \mathbf{IID\ N(0,1)}\}$.

$\boldsymbol{h_t}$: يمثل التباين الشرطي للخطأ العشوائي والذي يتم فيه نمذجة الارتباطات الذاتية للتذبذبات بالاعتماد على متَسَلسِلة مربعات البواقي السابقة، ويعرف رياضياً كالاتي [15,12]:

$$\mathbf{h_t} = \alpha_o + \sum_{i=1}^{r} \alpha_i\ \boldsymbol{\varepsilon}^2_{t-i} \quad , \quad \alpha_o > \mathbf{0}\ , \quad \alpha_i \geq \mathbf{0}\ , \sum_{i=1}^{r} \alpha_i < \mathbf{1} \quad , \quad \ldots\ldots (4)$$

تمثل المعادلة رقم (4) اعلاه معادلة التذبذبات الزمنية (معادلة التباين الشرطي) لأنمُوذج ARCH من الرتبة (r) .

وتؤكد البحوث التطبيقية أن أحد عيوب أنمُوذج (ARCH) عندما تكون رتبة الأنمُوذج الامثل كبيره نوعاً ما، فأن ذلك الأمر يتطلب تقدير عدد كبير من المعلمات وهي عملية معقدة بالفعل. أي أن نماذج (ARCH) لا تحقق المستوى المطلوب من التفسير أو التنبؤ مع وجود أقل عدد ممكن من المعلمات المطلوب تقديرها [15]. وللتغلب على هذه الصعوبات اقترح العالم (Bollerslev, 1986) أنمُوذج أكثر شمولية وهو أنموذج (GARCH) [13,15].

2.2.3 أنموذج الانحدار الذاتي العام المشروط بعدم تجانس التباين من الرتبة (r, s)

Generalized Autoregressive Conditional Heteroskedasticity, GARCH(r, s)

اقترح أنمُوذج GARCH من قبل العالم (Bollerslev, 1986) وذلك من أجل التعامل مع التعقيدات الصعبة التي تعترض تطبيق أنمُوذج (ARCH) والتي من الممكن أن يكون لها تأثير كبير على دقة أداء التنبؤ. حيث يعتبر أنمُوذج ARCH(r) أكثر شمولية من أنمُوذج GARCH(r, s) ويعود السبب في ذلك الى أن التباين الشرطي لأنمُوذج GARCH(r, s) يعتمد على مجاميع الاوزان المرجحة بقيم مربعات سلسلة البواقي للفترات السابقة ε^2_{t-i} وايضاً على القيم السابقة للتباين الشرطي ذاتهُ h_{t-j} . في هذا البحث تتم عملية نمذجة السلسلة الزمنية لبواقي أنمُوذج ARMAX باستعمال معادلة التذبذبات الزمنية لأنمُوذج GARCH(r, s) والتي تُعرف رياضياً كالاتي [15,5]:

$$h_t = \alpha_o + \sum_{i=1}^{r} \alpha_i \ \varepsilon^2_{t-i} + \sum_{j=1}^{s} \vartheta_j \, h_{t-j} \qquad , \qquad \dots\dots(5)$$

تمثل المعادلة رقم (5) اعلاه معادلة التذبذبات الزمنية (معادلة التباين الشرطي) لأنمُوذج GARCH من الرتبة (r, s) . و أن α_i و ϑ_j تمثل معلمات الانمُوذج، وفقاً للقيود المفروضة على المعلمات والذي يعتبر شرطاً ضرورياً وكافياً لكي تضمن $(h_t > 0)$ تباين شرطي موجب :

$\alpha_o > 0,$

$\alpha_i \geq 0, \quad \text{for } i = 1, 2, 3, \dots r,$

$\vartheta_j \geq 0, \quad for \ j = 1, 2, 3, \dots, s .$

ويعتبر أنمُوذج GARCH(1, 1) في اغلب الدراسات والبحوث التي تناولت مشكله التقلبات في ظواهر السلاسل الزمنية هو الأنموذج الأكثر استعمالاً من غيره من النماذج برتب مختلفة لكل من s و r في نمذجة التذبذبات الزمنية [15,39,35]. ويمكن استنتاج معادلة التباين الشرطي لأنمُوذج GARCH(1, 1) من المعادلة (5) بالعلاقة الرياضية التالية :

$$h_t = \alpha_o + \alpha_1 \ \varepsilon^2_{t-1} + \vartheta_1 h_{t-1} \qquad , \qquad \dots\dots(6)$$

يتضح ان عملية نمذجة معادلة التبيان الشرطي لأنمُوذج GARCH(1, 1) اعلاه تعتمد على القيمة السابقة لمربعات بواقي الأنمُوذج ε^2_{t-1} والقيمة السابقة للتباين الشرطي ذاتهُ h_{t-1} . أما α_o ، α_1، و ϑ_1 تمثل معلمات الأنمُوذج وفقاً للقيود الاتية :

$\alpha_o \geq 0 \quad \alpha_1 > 0, \ \vartheta_1 > 0 \ \text{ and } \ \alpha_1 + \vartheta_1 < 1$

ومن خلال عملية دمج معادلة المتوسط الشرطي والمذكورة في الصيغة رقم (1) مع معادلة التباين الشرطي والمذكورة في الصيغة رقم (5) يتكون الأنمُوذج الهجين الاول $ARMAX(n_p, n_q, n_b, n_k) - GARCH(r, s)$ والذي سيتم توظيفهُ في هذا البحث في عملية التنبؤ بالقيم المستقبلية للسلسلة الزمنية .

3.2.3 انمُوذج الانحدار الذاتي العام المشروط بعدم تجانس التباين بمدخلات خارجية المنشأ $(\mathbf{GARCHX(r,s,b)})$ (مقترح توظيفهِ)

GARCH with Exogenous variable , $\mathbf{GARCHX(r,s,b)}$

أنمُوذج $\mathbf{GARCH(r,s)}$ أقترح لمعالجة عيوب أنمُوذج $\mathbf{ARCH(r)}$ وهو اول أنمُوذج استُعمل في عملية التهجين في هذا البحث، حيث تم دمج أنمُوذج **ARMAX** الخطي مع أنمُوذج **GARCH** غير الخطي وبالتالي بناء أول أنمُوذج هجين. ولتطوير عملية التهجين بشكل أكثر كفاءة للسلاسل الزمنية قيد التحليل سنقوم في هذا البحث بمُقترح توظيف نماذج هجينة أخرى بأتباع منهجية **(Zhang)** الهجينة وبتوزيعات مختلفة للخطأ العشوائي. عملية التهجين هذه مُتضَمّنه في دمج السلسلة الزمنية للمتغير ذو التأثير الخارجي (x_t) مره أخرى ولكن هذه المرة يُدمج في معادلة التباين الشرطي لأنمُوذج **(GARCH)** للتباين الشرطي ليتكون أنموذج الانحدار الذاتي العام المشروط بعدم تجانس التباين بمدخلات إضافية **(GARCH with Exogenous variable)** تتمثل بمتغير التنبؤ الخارجي للسلسة الزمنية (x_t) وبالتالي بناء أنمُوذج $\mathbf{(GARCH-X)}$ ليتكون ثاني أنموذج يسُتعمل في عملية التهجين أيضاً. ويمكن تعريف معادلة التذبذبات الزمنية للتباين الشرطي للأنمُوذج الهجين الثاني رياضياً كما يلي [2]:

$$\mathbf{h_t = \alpha_o + \alpha_1\, \varepsilon^2{}_{t-1} + \vartheta_1 h_{t-1} + ɯ_1\, x^2{}_{t-1}} \qquad , \qquad \ldots\ldots(7)$$

تمثل المعادلة أعلاه معادلة التباين الشرطي للتقلبات الزمنية لأنموذج $\mathbf{GARCH(1,1)}$ الموسع أي بأدراج تأثير السلسلة الزمنية لمتغير التنبؤ الخارجي (x_t) في معادلة التذبذبات الزمنية لأنمُوذج $\mathbf{GARCH(1,1)}$ والمذكورة في الصيغة الرياضية رقم (6) .

أما $(ɯ_1)$: تمثل معلمة متغير السلسلة الزمنية ذو التأثير الخارجي (معلمة **Exogenous variable**)، وفقاً للقيد $(ɯ_1 > 0)$. ومن خلال عملية دمج معادلة المتوسط الشرطي والمذكورة في الصيغة رقم (1) مع معادلة التباين الشرطي والمذكورة في الصيغة رقم (7) ينشأ أنمُوذج $\mathbf{ARMAX(n_p,n_q,n_b,n_k) - GARCHX(r,s,b)}$ الهجين وهو ثاني أنمُوذج هجين سيتم توظيفهُ في هذا البحث في عملية التنبؤ بالقيم المستقبلية للسلسلة الزمنية.

حيث يعتبر هذا الأنمُوذج أكثر تطوراً من الأنمُوذج الهجين الاول ويعود السبب في ذلك الى أدراج ودراسة تأثير التقلبات الزمنية للسلسلة الزمنية ذو التأثير الخارجي (x_t) على السلسة الزمنية المدخلة (y_t) مرتين: الأولى في معادلة المتوسطات الشرطية للأنمُوذج الخطي **(ARMAX)** والثانية في معادلة التباينات الشرطية للأنمُوذج غير الخطي **(GARCHX)**، وهذا بدورهِ يزيد من دقة التنبؤ المستقبلي للظاهرة قيد الدراسة.

4.2.3 التوزيعات الاحتمالية الشائعة للخطأ العشوائي عندما يتبع انمُوذج GARCH(مقترح توظيفُها)

الهدف الرئيسي الآخر من هذا البحث هو توظيف توزيعات ذات الذيل الثقيل **(Heavy − tail)** للخطأ العشوائي للنماذج الهجينة **(ARMAX-GARCH)** و **(ARMAX-GARCHX)** للحصول على تنبؤات أكثر كفاءة للنماذج الهجينة. فعادة ما تكون هناك حاجة لدراسة تأثير التوزيعات الاحتمالية على الخطأ العشوائي للنماذج غير الخطية لغرض السيطرة على التقلبات المرتبطة مع الزمن في السلاسل الزمنية. حيث تم توظيف نماذج **GARCH** في الابحاث التطبيقية بافتراض التوزيع الطبيعي. وأثبتت هذه الابحاث أن افتراض التوزيع الطبيعي غير مُجْدِي في التعامل مع خاصية الالتواء أو التفلطح والتي تظهر عادة في البواقي التقديرية للنماذج. لذلك، اقترح **Bollerslev (1987)** و **Nelson (1991)**

[2] الصيغة الرياضية لأنموذج **(GARCHX)** مذكورة في بعض المصادر (أنظر المصدر رقم (19,34 و28)) ولكن جميع المصادر لم تدرس هذا الأنمُوذج وفقاً لمنهجية **(Zhang)** الهجينة في السلاسل الزمنية.

توزيعات غير التوزيع الطبيعي للتغلب على هذه المشكلة. ومن المعادلة رقم (3) يمكن اعادة كتابة حد الخطأ العشوائي η_t للنماذج غير الخطية كالاتي[3]:

$$\eta_t = \frac{\varepsilon_t}{\sqrt{h_t}} \quad \sim N(0,1)$$

والاقسام التالية توضح دالة الكثافة الاحتمالية لهذه التوزيعات الاحتمالية [4,12,31]:

1.4.2.3 التوزيع الطبيعي القياسي Gaussian Distribution

عند تطبيق حالة التوزيع الطبيعي على الخطأ العشوائي η_t لأنموذج GARCH فأن دالة الكثافة الاحتمالية الشرطية لـ ε_t تكون على النحو التالي [12,14]:

$$f(\varepsilon_t \backslash \Omega, \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{1}{2}\frac{\varepsilon_t^2}{h_t}\right)$$

$$h_t = \sigma_t^{\,2}$$

$$\Omega = (\mu_t, \propto_o \ , \propto, \vartheta)'$$

$$\mathcal{F}_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_1\}$$

لكن مع افتراض حالة التوزيع الطبيعي للأخطاء في أنمُوذج (GARCH) وبمعنى اخر (normal − tailed) لا يمكن لمواصفات هذا الأنمُوذج ان تتكيف مع خصائص السلسلة الزمنية لبواقي الأنمُوذج حيث عادة تتوافق مع توزيعات أخرى يمكن لها ان تستوعب بعض الخصائص لعملية الخطأ العشوائي في النماذج غير الخطية. حيث وجد عدد من الباحثين في التطبيقات العملية أن توزيع البواقي التقديرية غالباً ما يكون (leptokurtic) مرتفع (أكثر تدبباً) في قمة المنحني مقارنة بقمة منحني التوزيع الطبيعي. وللتغلب على هذه المشكلة هناك اثنين من التوزيعات البديلة المقترحة والمستعملة على نطاق واسع في بحوث السلاسل الزمنية عندما يتم تطبيقها على انمُوذج GARCH تكون قادرةً على تفسير سلوكيات التوزيعات ذات الذيل الثقيل (Heavy − tail) والتي تكون بقمة منحني مختلفة عن قمة المنحني للتوزيع الطبيعي [14,4,31]، وهي :

2.4.2.3 توزيع t Student's t Distribution

مع افتراض ان عملية الخطأ العشوائي تتبع توزيع t ، والمعروف بـ (GARCH t − distribution) ويشار إليه اختصاراً $\varepsilon_t \sim t(v, h_t)$، وفي هذه الحالة تأخذ دالة الكثافة الاحتمالية الشرطية الصيغة الآتية [14,4]:

$$f(\varepsilon_t \backslash \Omega, \mathcal{F}_{t-1}) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)h_t}} \times \left[1 + \frac{\varepsilon_t^{\,2}}{(v-2)h_t}\right]^{-(v+1)/2}$$

$\Gamma(*)$: تمثل دالة Gamma.

v: تمثل درجة الحرية في توزيع (Student's − t) أو معلمة شكل التوزيع وقيمتها ضمن الفترة $(2 < v < \infty)$.

[3] كذلك لم تدرس تأثير تغير التوزيع العشوائي للخطأ في أنمُوذج (GARCH) و (GARCHX) على قدرة التنبؤ للنماذج الهجينة، لذلك تعتبر أساليب مقترح توظيفها في هذا البحث.

3.4.2.3 توزيـع الخطأ العام Generalized error distribution

اقتـرح الباحـث (Nelson, 1991) أن عمليـة الخطـأ العشـوائي لأنمـوذج GARCH تتبـع التوزيـع العـام للأخطـاء (GED − GARCH) او يشـار إليـه اختصـاراً $\varepsilon_t \sim GED(v, \mu, h_t)$، والـذي يسمح بالأخـذ بعين الاعتبـار أشـكالاً مختلفة لـذيل التوزيـع ويعتبـر التوزيـع الطبيعـي القياسـي حالـة خاصـة مـن هـذا التوزيـع، وتأخـذ دالـة الكثافـة الاحتماليـة الشـرطية فـي هـذه الحالـة الصـيغة الآتيـة [31,14]:

$$f(\varepsilon_t \backslash \Omega, \mathcal{F}_{t-1}) = \frac{v\ exp\left[-\frac{1}{2}|z/\lambda|^{v}\right]}{\lambda\ \left[2^{\left(1+\frac{1}{v}\right)}\ \Gamma\left(\frac{1}{v}\right)\right]} \quad , \quad (-\infty < z < \infty)$$

و λ معرفة كالتالي :

$$\lambda = \left[2^{(-2/v)}\Gamma\left(\frac{1}{v}\right)\Gamma\left(\frac{3}{v}\right)\right]^{1/2}$$

حيـث v : تمثـل معلمـه سـماكة ذيـل التوزيـع (Tail thickness parameter)، معرفـة ضـمن الفترة $(0 < v \leq \infty)$.

3.3 مراحل بناء النماذج الهجينة (Hybrid models):

Building the Hybrid models

تـتم عمليـة بنـاء النمـاذج الهجينـة $ARMAX(n_p, n_q, n_b, n_k) - GARCH(r, s)$ و $ARMAX(n_p, n_q, n_b, n_k) - GARCHX(r, s, b)$ وفقاً للخطوات التالية :

1.3.3 المرحلـة الاولـى : تتمثـل المرحلـة الاولـى لبنـاء النمـاذج الهجينـة فـي مرحلـة التعـرف (Identification) علـى أفضـل أنمُـوذج مـن نمـاذج $ARMAX(n_p, n_q, n_b, n_k)$ وهـذا بـدورهُ معتمـد علـى اختيـار رتبـة الأنمُـوذج الامثـل، أي تحديـد المعلمـات n_p, n_q, n_b, n_k وهـي رتـب الانحـدار الـذاتي، المتوسـطات المتحركـة، رتبـة المتغيـر الخـارجي، ورتبـة زمـن التـأخير علـى التـوالي، بالاعتمـاد علـى معيـار معلومـات اكـاكي AIC، معيـار اكـاكي لخطـأ التنبـؤ النهـائي FPE، معيـار اكـاكي البيـزي BIC، وتعرف هذه المعايير رياضياً كالاتي [30,25]:

$$AIC(\beta) = \ln\left(\hat{\sigma}^2_{(\beta)}\right) + \frac{2(\beta)}{n} \qquad , \qquad \ldots\ldots (8)$$

$$FPE(\beta) = \left(\hat{\sigma}^2_{(\beta)}\right)\left(\frac{n+\beta}{n-\beta}\right) \qquad , \qquad \ldots\ldots (9)$$

$$BIC(\beta) = \ln\left(\hat{\sigma}^2_{(\beta)}\right) + \frac{\beta}{n}\ ln(n) \qquad , \qquad \ldots\ldots (10)$$

اما الرموز الرياضية للصيغ اعلاه تعرف كالاتي:

n :تمثل حجم العينة (عدد مشاهدات السلسلة الزمنية) .

β : تمثل عدد معلمات انمُوذج ARMAX المقدرة $(\beta = n_p, n_q, n_b)$.

$\hat{\sigma}^2_{(\beta)}$: يمثل تقدير تباين بواقي الانمُوذج .

2.3.3 المرحلة الثانية : الخطوة الثانية تتمثل في عملية تقدير معلمات الأنمُوذج الهجين، في هذا البحث سيتم الاعتماد على أسلوب مٌنهجية ذاتَ المرحلتين $(\text{Two} - \text{Step Procedure})$ والذي أشار اليه الباحثون $(\text{Rachev}, \text{S et all}, 2007)$ [34]، وكذلك $(\text{Tsay}, \text{R}, 2013)$ [39] في كتبهم العلمية في عملية تقدير النماذج الهجينة (Hybrid models)، ففي المرحلة الأولى يتم تقدير معلمات معادلة المتوسط لأنمُوذج (ARMAX)، في هذه البحث سيتم الاعتماد على ثلاث طرائق مختلفة لتقدير متجه المعلمات المطلوب تقديرهُ للأنمُوذج الخطي (ARMAX) وتعتمد هذه الطرائق على الخوارزميات التكرارية لتشخيص النظم الحركية الخطية وهي [4]:

1.2.3.3 طريقة المربعات الصغرى الاعتيادية (مقترح توظيُفها)

Ordinary Least Square Method, OLS

تستعمل طريقة المربعات الصغرى لتقدير متجه المعلمات غير المعلومة والمطلوب تقديره من أجل بناء أنمُوذج ARMAX وذلك من خلال جعل مجموع مربعات الأخطاء أقل ما يمكن والذي يمثل الفرق بين القيم الحقيقية $\mathbf{y(t)}$ والقيم المقدرة $\hat{\mathbf{y}}(\mathbf{t})$ [27,40] :

$$\hat{\varepsilon}(t) = y(t) - \hat{y}(t) = y(t) - \Phi^{T}(t)\hat{\beta} \quad , \quad \dots\dots (11)$$

وأن $\hat{\beta}$ يمثل مقدر المربعات الصغرى لمتجه معلمات معادلة المتوسط للأنمُوذج الهجين (أنمُوذج ARMAX) غير المعلومة β :

$$\beta = \left[\varphi_1, \varphi_2 \dots, \varphi_{n_p} \;;\; \phi_1, \phi_2 \;\dots, \phi_{n_b} \;;\; \theta_1, \theta_2 \dots, \theta_{n_p}\right]^{T} \quad , \quad \dots\dots (12)$$

ويعرف هذا المقدر $\hat{\beta}_{ols}$ رياضياً كالاتي:

$$\hat{\beta}_{ols} = (\Phi^{T}\Phi)^{-1}\Phi^{T}Y = \left[\sum_{t=1}^{n}\left(\Phi(t)\Phi^{T}(t)\right)\right]^{-1}\left(\sum_{t=1}^{n}(\Phi(t)\,y(t))\right) \quad , \quad \dots\dots (13)$$

حيث أن:

$$\Phi = [\vartheta(1), \dots, \vartheta(n)]^{T}$$

ϑ : يمثل متجه الانحدار والذي يتكون من أجزاء القياسات للمخرجات والمدخلات والاخطاء العشوائية للأنمُوذج في المعادلة رقم (1) ويعرف رياضياً كما يلي :

$$\vartheta = \left[y_{t-1}, \dots, y_{t-n_p} \,; x_{t-1}, \dots, \; x_{t-n_b} \,; \varepsilon_{t-1}, \dots, \varepsilon_{t-n_q}\right]^{T}$$

تتطلب طريقة المربعات الصغرى هذه وجود كُتْلة مُتكاملة محكمة من البيانات لعملية التقدير، ويعُدّ هذا الشي من عَيْوب استعمال طريقة OLS وذلك لأنهُ في كل لحظة زمنية يمكن أن تتوفر مشاهد ة جديدة، وبالتالي سوف ينمو حجم المعلومات حيث غالباً ما تكون هناك بيانات متاحة بشكل تسلسل زمني، وبالتالي فمن المفضل تطبيق خوارزمية تقدير تكرارية [23].

4 الطرائق الثلاث $(\text{OLS}, \text{RLS} - \text{EF}, \text{RPEM})$ المسُتعملة في تقدير معلمات معادلة المتوسط الشرطي للأنمُوذج الديناميكي ARMAX لم توظف من قبل لبناء النماذج الهجينة وفقاً لمنهجية Zhang ، لذلك تعتبر مقترح توظيفها في هذا البحث لبناء النماذج الهجينة .

2.2.3.3 طريقة المربعات الصغرى التكرارية باستعمال عامل التغاضي الاسي $(\mathbf{RLS-EF})$ (مقترح توظيفها)

Recursive Least Square with Exponential Forgetting Factor, RLS − EF

أن الصيغة التكرارية لطريقة المربعات الصغرى تعرف بطريقة المربعات الصغرى التكرارية RLS وتعتبر هذه الطريقة واحدة من أفضل الطرائق المستعملة في عملية تشخيص نظام $(\mathbf{SISO-ARMAX})$ حيث تتميز خوارزمية RLS بسرعة تقارب اسرع من غيرها من الخوارزميات ولكن تتضمن عمليات حسابية أكثر تعقيدًا من OLS [30,1] . وبما أن عملية التنبؤ عادة ما تتطلب أن يكون المُخَمّن قادراً على تتبع التغيرات في ديناميكية العملية لذلك من الضروري اعطاء أهمية نسبية أكبر للبيانات الحالية عن البيانات القديمة عند تتبع او تعقب هذه الاختلافات وهذا ما يعرف بـ بعامل التغاضي الاسي (Exponential Forgetting) عند تقدير المعلمة من البيانات حيث من المتوقع ان تكون البيانات الحديثة غنية بالمعلومات بينما البيانات التي تظل ثابتة نسبياً أو بدون تغير يمكن ان تفتقر للمعلومات، وتعتمد هذه الطريقة في ذلك على تقليل دالة الكلفة الى أقل ما يمكن، والتي تعطى بالصيغة الاتية [29,36]:

$$\mathbf{J}_t(\boldsymbol{\beta}) = \sum_{i=1}^{t} \lambda^{t-i}\, \varepsilon(\mathbf{i})^2 = \sum_{i=1}^{t} \lambda^{t-i}\, (\mathbf{y}(\mathbf{i}) - \boldsymbol{\Phi}^{\mathbf{T}}(\mathbf{i})\boldsymbol{\beta})^2 \qquad , \qquad \dots (14)$$

وأن λ يعرف بعامل التغاضي الاسي ويمثل عدداً موجباً تتراوح قيمتهُ ضمن الفترة $(0 < \lambda \leq 1)$ ، وعادة ما تكون الخيارات النموذجية لقيمة هذا العدد في الطرائق التكرارية بتوظيف عوامل التأثير تقع ما بين $(0.95 \leq \lambda < 0.99)$ [23,25] . ومن الطرق المُجْدِيَة في تحديد عامل التغاضي λ الاسي هي استعمال طريقة طول العينة التقريبي (Asymptotic sample length) والتي يشار اليها بالاختصار $(\mathbf{A}_{sl})$ وتعطى وفقاً للصيغة التالية [23]:

$$\boldsymbol{\lambda} = \mathbf{1} - \frac{\mathbf{1}}{\mathbf{A}_{sl}} \qquad \mathbf{A}_{sl} < \mathbf{1} \qquad , \qquad \dots (15)$$

وتمثل $\mathbf{A}_{sl}$ في العلاقة اعلاه عدد العينات السابقة والمعتمد عليه في تقدير المعلمات.

وتمثل الخطوات التالية منهجية عمل خوارزمية طريقة $(\mathbf{RLS-EF})$ في عملية تقدير متجه المعلمات $\boldsymbol{\beta}$ [27,40,1]:

1- تقدير خطأ التنبؤ باستعمال الصيغة الرياضية:

$$\boldsymbol{\varepsilon}(\mathbf{t}+\mathbf{1}) = \mathbf{y}(t+\mathbf{1}) - \boldsymbol{\Phi}^{\mathbf{T}}(\mathbf{t}+\mathbf{1})\hat{\boldsymbol{\beta}}(t) \qquad , \qquad \dots (16)$$

2- إيجاد تقدير أولي لعامل التغاضي الاسي $\boldsymbol{\lambda}$ وفق المعادلة (15).

3- تقدير حد الكَسْب $\mathcal{K}(\mathbf{t}+\mathbf{1})$ باستعمال المعادلة (17) التالية :

$$\mathcal{K}(\mathbf{t}+\mathbf{1}) = \mathbf{P}(\mathbf{t}+\mathbf{1})\ \boldsymbol{\Phi}(t+\mathbf{1}) = \frac{\mathbf{P}(\mathbf{t})\ \boldsymbol{\Phi}(t+\mathbf{1})}{[\boldsymbol{\lambda} + \boldsymbol{\Phi}^{\mathbf{T}}(\mathbf{t}+\mathbf{1})\mathbf{P}(\mathbf{t})\boldsymbol{\Phi}(t+\mathbf{1})]} \quad , \quad \dots (17)$$

حيث أن:

$\mathcal{K}(\mathbf{t}+\mathbf{1})$ يمثل حد الكَسْب او معامل المَنْفَعَة الربحية والذي يفسر مقدار قيمة خطا التنبؤ $\boldsymbol{\varepsilon}(\mathbf{t}+\mathbf{1})$ في قيمة المقدر.

$\mathbf{P}(\mathbf{t}+\mathbf{1})$: تمثل مصفوفة التغاير وتقدر رياضياً من خلال المعادلة رقم (18) وكما يلي :

$$\mathbf{P}(\mathbf{t}+\mathbf{1}) = \frac{\mathbf{1}}{\boldsymbol{\lambda}}[\mathbf{I} - \mathcal{K}(\mathbf{t}+\mathbf{1})\boldsymbol{\Phi}^{\mathbf{T}}(\mathbf{t}+\mathbf{1})]\mathbf{P}(\mathbf{t}) \qquad , \qquad \dots (18)$$

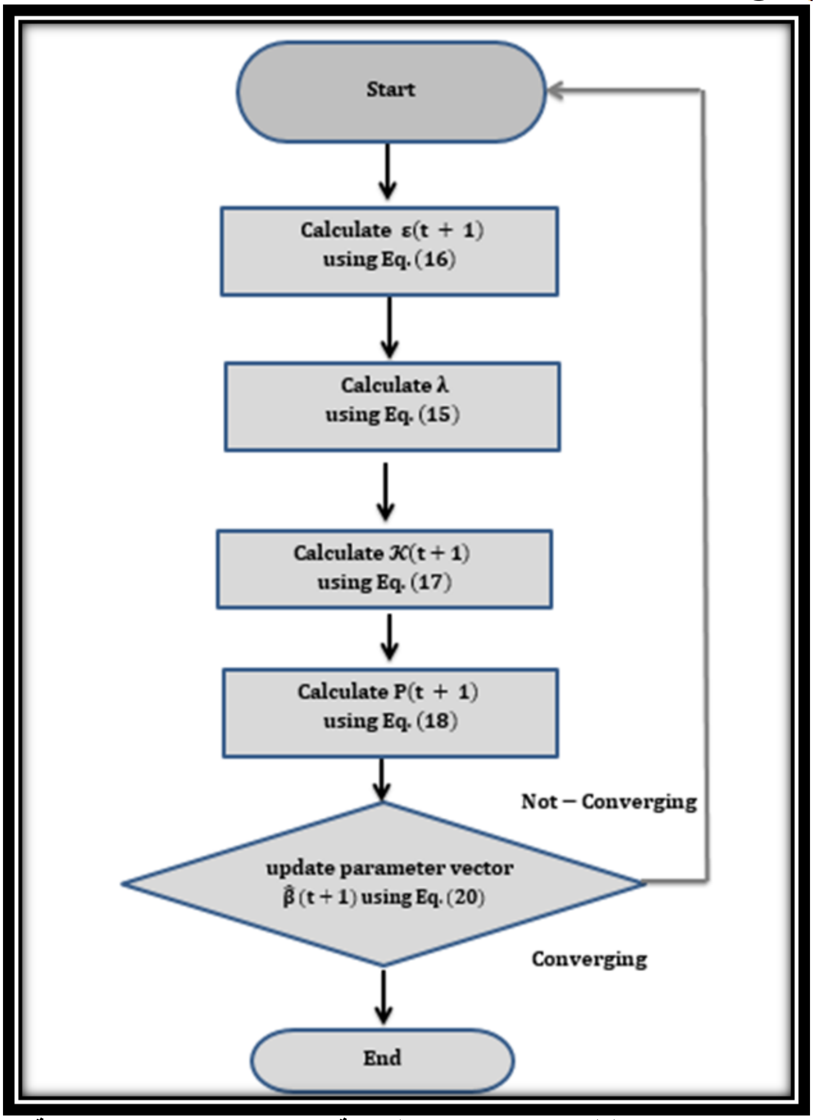
وفيها $P(t)$: تمثل مصفوفة التغاير (Covariance matrix) في الزمن الحالي t وتقدر كما يلي:

$$P(t) = [\Phi^T(t)\,\Phi(t)]^{-1} = \left[\sum_{t=1}^{n}\left(\Phi(t)\Phi^T(t)\right)\right]^{-1} \quad , \quad \dots (19)$$

4- تحديث متجه معلمات أنموذج ARMAX المقدرة وذلك من خلال المعادلة الرياضية الاتية :

$$\hat{\beta}(t+1) = \hat{\beta}(t) + \mathcal{K}(t+1)\,\varepsilon(t+1) \quad , \quad \dots (20)$$

5- الانتظار حتى انقضاء الدورة الزمنية القادمة (next time step) ومن ثم الرجوع الى الخطوة الاولى (الحصول على تقديرات متقاربة). وفيما يلي رسم مخطط يوضح خوارزمية مقدرات (RLS − EF) وفق الخطوات السابقة :



شكل (1) المخطط البياني لمقدرات طريقة المربعات الصغرى التكرارية باستعمال عامل التغاضي الاسي (RLS − EF)

3.2.3.3 طريقة خطأ التنبؤ التكرارية (مقترح توظيفُها)
Recursive Prediction Error Method, RPEM

تستعمل طريقة خطأ التنبؤ التكرارية (RPEM) لتقدير معلمات جميع تراكيب النماذج الخطية والتي تشتق من الصيغة العامة للأنموذج الخطي العام، لذلك فهي تعتبر واحدة من اهم الطرائق التكرارية التي يُستند اليها في تقدير معلمات أنموذج ARMAX، حيث تعتمد هذه الطريقة في صياغة المشكلة ومعالجتها على عامل التغاضي λ والذي يتم تضمينه في دالة الكلفة لهذه الطريقة وكما يلي [36]:

$$J_t(\beta) = \frac{1}{2}\sum_{i=1}^{t} \lambda^{t-i}\, \varepsilon(i,\beta)^T Q\, \varepsilon(i,\beta) \qquad , \qquad \dots (21)$$

أذ أن:

Q : مصفوفة موجبة معرفة (Positive definite matrix) .

وفي التطبيقات العملية فان هذه الطريقة تعتمد في تقدير المعلمات على تقنية (Gauss − Newton) والتي تستعمل تقريب متجه تدرج او انحدار المعلمة (Gradient vector)، وبالتالي، تتطلب طريقة (RPEM) بناء متجه التدرج لمخرجات الانموذج وبالنسبة الى معلماته وكما يلي [30]:

$$\psi^T(t) = \frac{\partial\, y(t)}{\partial\, \beta(t)} = \left[\frac{\partial\, y(t)}{\partial\, \beta_1(t)}\ \ \frac{\partial\, y(t)}{\partial\, \beta_2(t)} \cdots\cdots \frac{\partial\, y(t)}{\partial\, \beta_n(t)}\right] \qquad , \qquad \dots (22)$$

وعليه فأن خوارزمية خطأ التنبؤ التكرارية تكون كما يلي [30]:

1- تقدير خطأ التنبؤ وفق المعادلة الاتية:

$$\hat{\varepsilon}(t+1,\beta) = y(t+1) - \Phi^T(t+1)\hat{\beta}(t) \qquad , \qquad \dots (23)$$

2- تقدير حد الكَسْب $\mathcal{K}(t+1)$ وفقاً للعلاقة الاتية :

$$\mathcal{K}(t+1) = \frac{P(t)\,\psi(t+1)}{\left[\lambda + \psi^T(t+1)P(t)\psi(t+1)\right]} \qquad , \qquad \dots (24)$$

3- تقدير مصفوفة التغاير لخوارزمية (RPEM) وكما يلي :

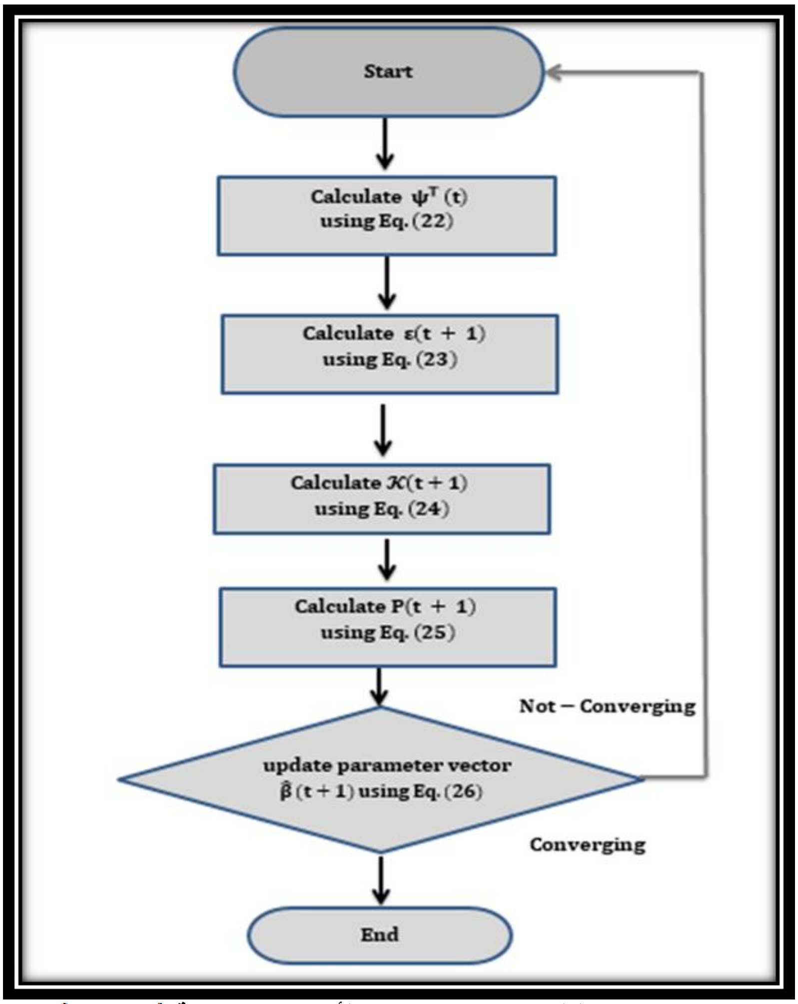
$$P(t+1) = \frac{1}{\lambda}\left[I - \mathcal{K}(t+1)\psi^T(t+1)\right]P(t) \qquad , \qquad \dots (25)$$

4- حساب متجه المعلمات وفقاً الصيغة التالية:

$$\hat{\beta}(t+1) = \hat{\beta}(t) + \mathcal{K}(t+1)\,\hat{\varepsilon}(t+1) \qquad , \qquad \dots (26)$$

وفيما يلي رسم مخطط يوضح خوارزمية مقدرات (RPEM) وفقاً للخطوات السابقة [5]:

[5] تم عمل المخططات البيانية لطرائق التقدير (RLS − EF) و (RPEM) من قبل الباحث .



شكل (2) المخطط البياني لمقدرات خطأ التنبؤ التكرارية (RPEM)

4.2.3.3 مقارنة طرائق تقدير أنموذج $\mathbf{ARMAX}(n_p, n_q, n_b, n_k)$:

تم استخدام ثلاث طرائق ولكل طريقة صفات تختلف عن الاخرى لتقدير معلمات أنموذج $\mathbf{ARMAX}(n_p, n_q, n_b, n_k)$ ولتقييم أداء هذه الطرائق المختلفة واختيار الطريقة الأمثل في عملية التقدير ومن ثم الاعتماد على بواقي هذه الطريقة لبناء النماذج غير الخطية بعد عملية التشخيص، لا بدّ وأن نعمل مقارنة لطرائق التقدير التي سبق ذكرها، ولتحقيق هذا الغرض يمكن استعمال عدة مقاييس ومن اهمها والتي أعطت اقل قيم تقديرية في الجانب العملي وهي متوسط الخطأ المطلق (MAE) وأيضاً مقياس الجذر التربيعي لمتوسط مربعات الخطأ المَعيّاري (Normalized RMSE) والتي تُعطى بالصيغ الرياضية التالية [6,38]:

A. متوسط الخطأ المطلق Mean Absolute Error (MAE)

ويمكن التعبير عن هذا المقياس بالصيغة الرياضية الاتية:

$$\mathbf{MAE}\ (\hat{y}) = \frac{1}{n}\sum_{t=1}^{n}|e_t| = \frac{1}{n}\sum_{t=1}^{n}\left|y_{actual,i} - \hat{y}_{estimated,i}\right| \quad , \quad \dots(27)$$

B. مقياس الجذر التربيعي لمتوسط مربعات الخطأ المَعيّاري (مقترح توظيفه)

Normalized Root Mean Square Error

أن مقياس الجذر التربيعي لمتوسط مربعات الخطأ المَعيّاري (Normalized RMSE) والذي يُكتب اختصاراً بـ (NRMSE) يعرف من خلال الصيغة الرياضية التالية :

$$\mathrm{NRMSE}(\hat{y}) = \frac{\mathrm{RMSE}(\hat{y})}{\bar{y}_{\mathrm{actual}}} \quad , \quad \dots\dots \ (28)$$

وان RMSE يمثل الجذر التربيعي لمتوسط مربعات الخطأ ويعُرف رياضياً كالاتي :

$$\mathrm{RMSE}(\hat{y}) = \sqrt{\mathrm{MSE}(\hat{y})} = \sqrt{\sum_{i=1}^{n} \left(y_{\mathrm{actual},i} - \hat{y}_{\mathrm{estimated},i}\right)^2 / n} \quad , \dots \ (29)$$

وتعرف الرموز الرياضية لمقاييس المقارنة أعلاه كالاتي:

$y_{\mathrm{actual},i}$: تمثل القيم الحقيقية y في الزمن i .

$\hat{y}_{\mathrm{estimated},i}$:تمثل القيمة التقديرية للأنمُوذج في الزمن i .

n : تمثل حجم العينة .

حيث تشير القيمة الادنى للمقاييس [$\mathrm{NRMSE}(\hat{y})$, $\mathrm{MAE}(\hat{y})$] الى افضل مُقدر لأنموذج ARMAX.

3.3.3 المَرحلة الثالثة: فحص مدى ملائمة الأنمُوذج

Diagnosis checks for model adequacy

بعد التعرف على الأنمُوذج وتقدير معلماته تأتي التحقق من مدى ملاءمة الانمُوذج المشخص وذلك لغرض تحّسين الانمُوذج وتطويرهُ أو الابقاء عليه كما هو. تعتبر هذه المرحلة اهم مراحل التحليل فهي المرحلة التي تحدد قبول أو استبدال الأنمُوذج بأنمُوذج اخر وتتضمن العديد من الفحوصات كالرسوم البيانية (دالة الارتباط الذاتي والذاتي الجزئي لبواقي الأنموذج الخطي) وكذلك اجراء الاختبارات الإحصائية على سلسلة بواقي أنمُوذج ARMAX وذلك لاختبار وجود تأثير ARCH لغرض تحديد ما إذا كانت نماذج GARCH قابلة للتطبيق أم لا ؟ وهذا يتم باستعمال الاختبارات الإحصائية التالية:

1.3.3.3 اختبار **Ljung – Box Test**

اقترح كل من الباحثين (Ljung & Box, 1978) إحصائية الاختبار هذه كتحسين لإحصائية (Box – Pierce)، والذي يستعمل لاختبار وجود تأثير ARCH من خلال اختبار عشوائية البواقي للسلسلة الزمنية عن طريق حساب معاملات الارتباط الذاتي للبواقي لمجموعة من الآزاحات، حيث يتميز اختبار (Ljung and Box) بأنه يأخذ في الحسبان مدى معنوية عدد من الفجوات الزمنية كمجموعة واحدة فإنها قد تبدو معنوية ويتم بعدها اختبارها بفرضية العدم والفرضية البديلة الأتية [11,24]:

$$H_0:\ \rho_1 = \rho_2 = \cdots = \rho_k = \rho_m = 0 \ , \quad \text{for } k = 1, 2, \dots, m$$
$$H_1:\ \rho_m \neq 0 \ , \quad \text{for some values of } m \quad , \quad \dots\dots \ (30)$$

وأن إحصاءه الاختبار تُعطى بالعلاقة الآتية:

$$Q_{L-B} = \left(n(n+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^{\,2}}{n-k} \right) \sim \chi^2_{(v)} \quad , \quad \dots\dots \ (31)$$

ويتم تقدير معامل الارتباط الذاتي للبواقي بالعلاقة الاتية:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n}\left(\hat{\varepsilon}_t^{\,2} - \hat{\sigma}^2\right)\left(\hat{\varepsilon}^2_{t-k} - \hat{\sigma}^2\right)}{\sum_{t=1}^{n}\left(\hat{\varepsilon}_t^{\,2} - \hat{\sigma}^2\right)^2} \quad , \quad k = 1,2,\dots m \ ,\dots (32)$$

أذ أن:

$\hat{\sigma}^2$: يمثل تباين العينة لسلسة البواقي التربيعية.

m : أكبر ازاحة للارتباطات الذاتية للبواقي ، n: تمثل حجم العينة .

k : تمثل الإزاحات الزمنية وتأخذ القيم من 1 الى m .

$\hat{\rho}_k$: تمثل مقدرات معاملات الارتباط الذاتي لسلسلة البواقي $\hat{\varepsilon}_t$ ولمربع سلسلة البواقي $\hat{\varepsilon}_t^{\,2}$.

وعند اختبار بواقي الانمُوذج وفق الصيغة (31) ويتم مقارنة القيمة الاحتمالية للاختبار مع القيم الجدولية و مستوى المعنوية α، فاذا كانت $\{p - \text{value} < 0.05\}$ فهذا يعني عدم رفض الفرضية البديلة وبمعنى وجود مشكلة في بواقي الانمُوذج الخطي أي انها لا تمثل عملية تشويش أبيض (White noise) بالإضافة الى مشكلة عدم تجانس التباين Heteroscedasticity، و غير ذلك فان الاختبار هذا يدل على ان البواقي ذوات تشويش ابيض وان الأنمُوذج المشخص ملائم .

ومن الجدير بالإشارة أن استعمال الاختبارات الإحصائية تعتبر دعم اقوى من الرسوم البيانية، ومع ذلك فأن أسلوب الاختبار والرسم البياني لا يمتاز بعملية تشخيص الأنماط غير الخطية في السلسلة الزمنية لبواقي الأنمُوذج الملائم من النماذج الخطية، لذلك اقترح اختبار $(\text{BDS} - \text{Test})$ والذي يكون قويّ وفعال للتعرف على الانماط غير الخطية في بواقي أنموذج (ARMAX) الخطي [37]، والذي سيتم استعماله تطبيقياً في هذا البحث للتحري على الأنماط غير الخطية في بواقي الأنمُوذج الخطي.

2.3.3.3 اختبار BDS $\text{BDS} - \text{Test}$

للكشَف عن البِنْيَة أو الصفة غير الخطية في البواقي التقديرية $\hat{\varepsilon}_t$ لأنمُوذج السلسة الزمنية الخطي (ARMAX) الذي تم تقديرهُ قدم اختبار $(\text{BDS} - \text{Test})$ من قبل كلاً من $(\text{Brock} - \text{Dechert} - \text{Scheinkman})$ في عام (1986) ونشر في عام (1996) كاختبار تَشْخِيصِيّ ثُبت استعمالهُ في العديد من التطبيقات للسلاسل الزمنية اللاخطية، حيث أن الصفة غير الخطية (Nonlinearity) في تركيبة التباين والتباين المشترك للأنموذج الخطي يُمكن التعرف عليها من خلال تطبيق هذا الاختبار تحت فرضيتي العدم والبديلة التالية [26,7]:

$$H_0 = \text{linearity in } \hat{\varepsilon}_t$$
$$H_0 = \text{Non} - \text{linearity in } \hat{\varepsilon}_t \qquad , \quad \dots (33)$$

وأن صيغة هذا الاختبار تعتمد على تكامل الارتباط $(\text{Correlation integral})$ وهو مقياس التردد الذي تتكرر به الأنماط الزمنية في البيانات والتي تعطى بالصيغة الرياضية الاتية :

$$\text{BDS} = \sqrt{n}\ \frac{\left(C_{1,n} - C_{m,n}{}^{m}\right)}{\hat{\sigma}_{m,n}} \ \cong \ N(0,1) \qquad , \qquad \dots (34)$$

حيث أن:

n : تمثل عدد مشاهدات السلسلة الزمنية للبواقي.

$\hat{\sigma}_{m,n}$ يمثل الانحراف المعياري لبواقي السلسة الزمنية.

$C_{m,n}{}^{m}$: يمثل تكامل الارتباط عند البعد m والذي يقدر من خلال العلاقة الرياضية الاتية :

$$C_{m,n}\ (\epsilon) = \frac{2}{(n-m+1)(n-m)} \sum_{s=1}^{n-m+1} \sum_{t=s+1}^{n-m+1} \prod_{j=0}^{m-1} \left[I_{\epsilon}\left(\varepsilon_{t+j}, \varepsilon_{s+j}\right)\right]$$

$$I_{\epsilon}(\varepsilon_t,\varepsilon_s)=\begin{cases}1 & \text{if} \quad |\varepsilon_{t+j}-\varepsilon_{s+j}|\le \epsilon \\ 0 & \text{otherwise}\end{cases}$$

وتمثـل قيمـة m البعـد الضـمني (Embedding dimension) وتكـون دائمـاً $(m \geq 2)$ وفـي التطبيقـات العملية فـأن الخيـار الأمثـل لهـذه القيمـة يكـون $(m = 4)$، أمـا الرمـز ϵ فيشـير الـى المسـافة (Distance) بـين الابعـاد للارتباطـات المتكاملـة والـذي يكـون ضـمن المـدى $\left(0.5 < \frac{\epsilon}{\hat{\sigma}_{m,n}} < 2\right)$. ويـتم رفـض فرضـية العـدم عنـد مسـتوى معنويـة α اذا كانـت قيمـة $(p - value < \alpha)$ فـرفض فرضـية العـدم يمكـن ان يكـون ناجمـاً عـن وجـود بنْيَـة ارتبـاط غيـر خطـي فـي سلسـلة البـواقي أي أن بـواقي الانمـوذج غيـر مسـتقلة ومتماثلـة التوزيـع وهـذا بـدوه يؤيـد مواصـلة التحليـل أي اجـراء التحليـل غيـر الخطـي بتطبيـق نمـاذج (GARCH) علـى بـواقي الانمُـوذج الخطـي . أما إذا تعذر رفض فرضية العدم ، فلا يمكن رفض أنمُوذج (ARMAX) الخطي.

4.3.3 المَرحلة الرابعة: بنـاء نمـاذج ARCH & GARCH غير الخطيـة :

Building a Nonlinear ARCH and GARCH models

عـادةً مـا تتمثـل عمليـة نمذجـة وبنـاء نمـاذج ARCH و GARCH للتنبـؤ باتبـاع المراحـل التقليديـة لبنـاء أي انمُـوذج مـن نمـاذج السلاسل الزمنيـة، لكـن فـي النمـاذج الهجينـة تكـون بدايـة بنـاء هـذه النمـاذج من دراسة سلسلة البواقي الناتجة من مطابقة الأنموذج الأمثل من نماذج ARMAX، وكما يلي :

1.4.3.3 تشخيص البواقي: **Diagnosis of residuals**

يـتم تشَـخيص سَلسْـلة البـواقي (Residual Series) وذلـك لاختبـار وجـود تـأثير (ARCH) للبـواقي التـي تـم الحصـول عليهـا مـن توفيـق أفضـل انمـوذج دينـاميكي مـن بـين نمـاذج (ARMAX) وذلـك لتحديـد قابليـة تطبيـق نمـاذج (GARCH)، وهـذا يـتم مـن خـلال اسـتعمال بعـض الرسـوم البيانيـة ومــن ثــم اســتعمال الاختبــارات الإحصــائية (Ljung – Box Test) و (BDS – Test) للكشــف عـن وجـود تـأثير ARCH والانمـاط غيـر الخطيـة فـي بـواقي الأنمُـوذج الخطـي [26] ، والتـي تـم التطـرق اليها في القسم (3.3.3) السابق من هذا البحث.

2.4.3.3 اختيـار رتبـة الانموذج **Model Order Determination**

تعتبـر هـذه المرحلـة مـن المراحـل الأساسـية والمهمـة جـداً فـي بنـاء انمُـوذج $GARCH(r,s)$ وأنمُـوذج $GARCHX(r,s,b)$ والتـي يـتم فيهـا اختيـار الأنمُـوذج الأمثـل والمناسـب لعمليـة التهجـين (Hybridization). وان آليـة تحديـد او التعـرف علـى رتـب كـل مـن r و s ينبغـي ان تكـون مختـارة بشـكل دقيـق جـداً، وذلـك لأن اختيـار رتبـة أدنـى أو أعلـى مـن الرتبـة الفعليـة يُـؤدى إلـى عـدم اتسـاق معلمـات الانمُـوذج، وبالتـالي فـان هـذه الجوانـب تخفـض مِـن قـدْره أداء نمـاذج GARCH فـي التنبؤ.

وأن أكثــر المعــايير المســتعملة فــي اختيــار رتبــة نمــاذج ARCH و GARCH هــي معــايير (AIC, BIC, HQC) ، معيـار AIC ومعيـار BIC تـم تعريـف كـلاً مـنهم مسـبقاً فـي القسـم السـابق مـن هـذا البحـث، أمـا معيـار HQC معيـار المعلومـات لــ (Hannan – Quinn) والـذي أقتـرح مـن قبـل (Hannan & Quinn) فــي عــام (1979) لغـرض تشـخيص رتبــة الانمُــوذج والــذي يرمــز لــه بالرمز (HQC)، ويعرف هذا المعيار رياضياً بالشكل الاتي [16,22]

$$HQC(k) = -[2\log(L) + 2k\ln(\ln(n))]/n \qquad , \qquad \ldots\ldots (35)$$

n :تمثل حجم العينة .

k : تمثــل عــدد معلمــات انمُــوذج ARCH أي $(k = r)$ أو أنمــوذج $GARCH(r,s)$ أي $(k = s + r)$المقدرة.

$\log(L)$: يمثل لوغاريتم دالة الإمكان الأعظم (Log Likelihood) .

3.4.3.3 تقديرات الإمكان الاعظم لمعلمات أنمُوذج GARCH

MLE estimates of the GARCH model

في عملية تقدير النماذج الهجينة في هذا البحث تم الاعتماد على منهجية ذات المرحلتين (Two – Step Procedure)، المرحلة الأولى تمثل عملية تقدير النماذج الخطية والتي تم الحديث عنها في القسم (2.3.3). اما المرحلة الثانية تتمثل في عملية تقدير النماذج غير الخطية باستعمال طريقة الإمكان الأعظم. وللحصول على مقدرات الإمكان الأعظم لمتجه المعلمات غير المعلومة وليكن Θ عندما لا يتوفر حل تَحْلِيلِيّ او يصعب حل مشتقات دوال الإمكان الأعظم نظريّا، يتم استخدام خوارزميات التحسين العددي ومن الخوارزميات العددية الشائعة الاستعمال الخوارزمية التي اوجدها الباحثين (Berndt, Hall, Hall and Hausman, 1974) والتي تكتب بشكل مختصر (BHHH algorithm) وهي احدى الطرائق العددية المشتقة من طريقة (Newton – Raphson) التكرارية [3,21]. وبافتراض أن $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$ تخْضَع الى عملية أنمُوذج GARCH(1,1) واستناداً إلى افتراض التوزيع الطبيعي على عملية الخطأ العشوائي ε_t للأنمُوذج، أي أن :

$$\varepsilon_t \backslash \propto_o, \propto_1, \vartheta_1, \varepsilon_{t-1} \sim N\left(0,\ \propto_o + \propto_1 \varepsilon_{t-1}^2 + \vartheta_1 h_{t-1}\right) = N(0, h_t)$$

بالتالي فان دالة الإمكان المشتركة المشروطة بالمعلومات الماضية تكتب كما يلي:

$$L_n(\propto_o, \propto_1, \vartheta_1; (\varepsilon_t) \backslash \mathcal{F}_{t-1}) = \prod_{t=1}^{n} \left\{ \frac{1}{\sqrt{2\pi h_t}} exp\left(-\frac{{\varepsilon_t}^2}{2h_t}\right) \right\}$$

$$= (2\pi)^{-\frac{n}{2}} \prod_{t=1}^{n} \left({h_t}^{-0.5}\right) exp\left(-\frac{1}{2}\sum_{t=1}^{n} \frac{{\varepsilon_t}^2}{h_t}\right)$$

حيـــث أن h_t تتبـــع انمُـــوذج GARCH(1,1) تَبَعــاً للمعادلـــة (6) وبـــافتراض أن $\Theta = (\propto_o, \propto_1, \vartheta_1)'$ يمثـل متجـه المعلمـات غيـر المعلومـة والمـراد تقـديرها فـأن لوغـاريتم دالـة الكثافة الاحتمالية المشروطة تكون [21,41] :

$$\ell(\Theta) = -\frac{n}{2} log(2\pi) - 0.5 \sum_{t=1}^{n} log(h_t) - 0.5 \sum_{t=1}^{n} \frac{{\varepsilon_t}^2}{h_t}$$

$$= -\frac{n}{2} log(2\pi) - 0.5 \sum_{t=1}^{n} log(\propto_o + \propto_1 \varepsilon_{t-1}^2 + \vartheta_1 h_{t-1})$$

$$-0.5 \sum_{t=1}^{n} \frac{{\varepsilon_t}^2}{(\propto_o + \propto_1 \varepsilon_{t-1}^2 + \vartheta_1 h_{t-1})} \quad , \dots \dots (36)$$

بالتـالي فـأن مشـتقات الدالـة $\ell(\Theta)$ الجزئيـة بالنسـبة الـى المعلمـات $\propto_o$، $\propto_1$ و ϑ_1 علـى التـوالي تكـون كما يلي [10,41]:

$$\frac{\ell\partial(\Theta)}{\partial \propto_o} = \frac{1}{2}\sum_{t=1}^{n} \frac{{\varepsilon_t}^2}{{h_t}^2} - \frac{1}{2}\sum_{t=1}^{n} \frac{1}{h_t} \quad , \dots \dots (37)$$

$$\frac{\ell\partial(\Theta)}{\partial \propto_1} = \frac{1}{2}\sum_{t=1}^{n} \frac{{\varepsilon_t}^2 \varepsilon_{t-1}^2}{{h_t}^2} - \frac{1}{2}\sum_{t=1}^{n} \frac{\varepsilon_{t-1}^2}{h_t} \quad , \dots \dots (38)$$

$$\frac{\ell\partial(\Theta)}{\partial\,\vartheta_1}=\frac{1}{2}\sum_{t=1}^{n}\frac{\varepsilon_t^{\,2}h_{t-1}}{h_t^{\,2}}-\frac{1}{2}\sum_{t=1}^{n}\frac{h_{t-1}}{h_t}\quad,\qquad \dots\dots(39)$$

منظومة المعادلات أعلاه يمكن ان تحل بالنسبة الى متجه المعلمات المقدرة $\hat{\alpha}_0$ ، $\hat{\alpha}_1$ و $\hat{\vartheta}_1$ بالاعتماد على الطرائق العددية وباستعمال خوارزمية $(\mathbf{BHHH})$ وكما يلي [45,41,2]:

$$\Theta(i+1)=\Theta(i)+\lambda_i\;\;A(i)^{-1}\,G(i)\quad,\qquad \dots\dots(40)$$

حيث أن $\Theta(i)$ متجه المعلمات المقدرة في الدورة i .

λ_i : عدد ثابت $(\mathbf{scalar})$ يمثل طول الدورة $(\mathbf{Step-Length})$ وتكون القيمة الأولية وفي الدورة الأولى لهذا العدد تساوي الواحد الصحيح $(\lambda=1)$، ويكون متغير في كل دورة حيث يتم اختياره لتعظيم $(\mathbf{maximize})$ دالة الإمكان في الدورة الحالية.

$G(i)$ و $A(i)$ تمثل متجه ومصفوفة تقديرات معلمات أنموذج $\mathbf{GARCH}(1,1)$ على التوالي وتعُرف كما يلي :

$$G(i)=\sum_{t=1}^{n}\frac{\partial\ell_t}{\partial\Theta}=\sum_{t=1}^{n}\left[\frac{\partial\ell_t}{\alpha_0}\;\frac{\partial\ell_t}{\alpha_1}\;\frac{\partial\ell_t}{\vartheta_1}\right]\quad,\qquad \dots\dots(41)$$

$$A(i)=\sum_{t=1}^{n}\frac{\partial\ell_t}{\partial\Theta}\frac{\partial\ell_t}{\partial\Theta'}=\sum_{t=1}^{n}\left[\frac{\partial\ell_t}{\alpha_0}\;\frac{\partial\ell_t}{\alpha_1}\;\frac{\partial\ell_t}{\vartheta_1}\right]\left[\frac{\partial\ell_t}{\alpha_0}\;\frac{\partial\ell_t}{\alpha_1}\;\frac{\partial\ell_t}{\vartheta_1}\right]'\quad,\dots(42)$$

4.3 منهجية (Zhang) الهَجينة في السلاسل الزمنية
Zhang Hybrid metholodegoy for time Series

أصبح تطبيق النماذج الكمية للتنبؤ والمساعدة في اتخاذ او صنع القرارات أمرا ضَرُوريّا لا غنى عنه في التطبيقات العملية أكثر من أي وقت مضى. وعلى الرغم من تَوَافِر العديد من نماذج السلاسل الزمنية، فإن دقة التنبؤ بالسلسلة الزمنية حاليًا تعتبر أساسية لكثير من عمليات اتخاذ القرار. بالتالي، بدا البحث عن طرائق لتحسين فعالية نماذج التنبؤ، حيث توصل العديد من الباحثين في مجال التنبؤ بالسلسلة الزمنية أن اساليب التنبؤ يتحسن أدائها في النماذج الهجينة $(\mathbf{Hybrid\ Models})$. وهي نماذج مُخْتَلِطة او مُدمجة بين نموذجين او اكثر تستعمل للتنبؤ والاستِفادة من خصائص هذه الطرائق لبناء انمُوذج مشترك بينهم يجمع بين الخصائص المنُفردة للنماذج، حيث ان عملية التهجين بين نماذج السلاسل الزمنية يمكن أن يكون وسيلة فعالة للتغلب على القيود المفروضة على مركبة كل انمُوذج فضلاً عن تحسين دقة التنبؤ [42,43]. وأحد أهم الدوافع التي دعت لتوظيف النماذج الهجينة في هذا البحث هو أنهُ غالبا ما تكون السلسلة الزمنية تجمع بين عدة مشاكل، عندئذ لا يمكن أن تكون نماذج $\mathbf{ARMAX}$ ولا $\mathbf{GARCH}$ كافية في النمذجة والتنبؤ لأن الاول لا يستطيع التعامل مع العلاقات غير الخطية، بينما نماذج $\mathbf{GARCH}$ وحدها ليست قادرة على التعامل مع كل من الأنماط الخطية وغير الخطية على حد سواء بشكل جيد، هذا يعني ان النماذج المستقلة تم بناءها لمعالجة قصور واحد في السلسلة الزمنية.

ويتمثل احد أهم اهداف هذا البحث في توظيف منهجية $(\mathbf{Zhang}, \mathbf{2003})$[43] الهجينة لبناء نماذج هجينة بين أنمُوذج $\mathbf{ARMAX}(\mathbf{n_p},\mathbf{n_q},\mathbf{n_b},\mathbf{n_k})$ الخطي ونماذج $\mathbf{GARCH}(\mathbf{r},\mathbf{s})$ و $\mathbf{GARCHX}(\mathbf{r},\mathbf{s},\mathbf{b})$ غير الخطية. تفترض هذه المنهجية أن السلاسل الزمنية تكون خليط من مركبتين خطية وغير خطية ويمكن ان يعبر عنها رياضياً كما يلي [6]:

---

[6] تعتبر منهجية $(\mathbf{Zhang})$ ايضاً من الاساليب المقترح توظيفها لبناء النماذج الهجينة لهذا البحث، وذلك لان جميع المصادر التي تناولت نماذج $(\mathbf{ARMAX-GARCH})$ لم تدرس دمج اسلوب الأنظمة الديناميكية لأنمُوذج $\mathbf{ARMAX}$ مع طرائق تقديرهُ في بناء مثل هذا النوع من النماذج الهجينة بتوزيعات مختلفة للخطأ العشوائي.

$$y_t = \underbrace{L_t}_{ARMAX} + \underbrace{N_t}_{GARCH} \quad \rightarrow \quad \text{for ARMAX} - \text{GARCH} \quad \dots(43)$$

$$y_t = \underbrace{L_t}_{ARMAX} + \underbrace{N_t}_{GARCHX} \quad \rightarrow \quad \text{for ARMAX} - \text{GARCHX} \quad , \quad \dots(44)$$

حيث أن:

L_t : تمثل المركبة الخطية وتُقدر من الأنمُوذج الخطي ARMAX.

N_t : تمثل المركبة غير الخطية وتُقدر من النماذج غير الخطية GARCH و GARCHX.

تتضمن المنهجية الهجينة اعلاه الدمج بين مركبة التنبوء الخطية ومركبة التنبوء غير الخطية للحصول على التنبؤات النهائية للنماذج $(ARMAX - GARCH)$ و $(ARMAX - GARCHX)$ الهجينة ،والخطوات المتسلسلة الاتية تم الاعتماد عليها في هذا البحث للحصول على القيم التنبؤية النهائية لهذه النماذج [42,43]:

1- تتمثل المرحلة الاولى من عملية بناء النماذج الهجينة في بناء وتحديد افضل أنمُوذج من بين نماذج ARMAX الخطية للسلاسل الزمنية ثنائية المتغيرات x_t و y_t قيد الدراسة .

2- في الخطوة الثانية، يتم استخراج بواقي أنمُوذج ARMAX المُجهز في الخطوة الأولى، ولتكن e_t تمثل تقديرات بواقي الأنمُوذج في الزمن t، بالتالي :

$$e_t = y_t - \hat{L}_t \quad , \quad \dots\dots(45)$$

$\hat{L}_t$: تمثل مركبة قيم التنبوء النهائية لمعادلة المتوسط الشرطي والمذكورة في الصيغة رقم (1) والتي يتم الحصول عليها من خلال العلاقة الرياضية الاتية [39,44]:

$$\hat{L}_t(\ell) = E(y_{t+\ell}) = \sum_{i=1}^{p} \hat{\varphi}_i \, \hat{y}_t(\ell - i) - \sum_{j=1}^{q} \hat{\theta}_j \, \hat{\varepsilon}_t(\ell - j) + \sum_{\kappa=1}^{d} \hat{\Phi}_\kappa \, x_{t+\ell-k} \quad , \quad \dots(46)$$

حيث أن $(\ell = 1, 2, \dots)$ تمثل الخطوة المستقبلية للتنبوء، وكذلك :

$$\hat{y}_t(\ell - i) = y_{t+\ell-i} \quad \text{if} \quad \ell - i \le 0 \ \text{ or } \ i \ge \ell$$

$$\hat{\varepsilon}_t(\ell - j) = \begin{cases} \varepsilon_{t+\ell-j} & j \ge \ell \\ 0 & j = 0, 1, \dots, \ell - 1 \end{cases}$$

وان خطاء التنبوء (Forecast error) لمعادلة التنبوء (46) للفترة الزمنية ℓ يكون كما يلي :

$$e_t(\ell) = y_{t+\ell} - \hat{y}_t(\ell), \quad , \quad \dots\dots \ (47)$$

3- ولأهمية البواقي Residuals في تشخيص قدرة وكفاية النماذج الخطية، لذلك يتم في المرحلة الثالثة تشخيص البواقي التي تم تقديرها من انمُوذج ARMAX والذي تم تحديده في المرحلة الأولى من خلال الرسوم البيانية والاختبارات الإحصائية التي تم الحديث عنها في القسم (3.3.3)، وهنا يتم تمييز نوعين من البواقي وهي بواقي ذوات تركيبة خطية (Linear Residuals) وبواقي ذوات تركيبة غير خطية (NonLinear Residuals). إذا كانت البواقي المقدرة في الخطوة الاولى تتميز بعلاقه خطيه ففي هذه الحالة لا يمكن استعمالها لبناء الانمُوذج الهجين، بينما وجود أي نمط للبواقي غير الخطية يشير الى تقيُد استعمال انموذج ARMAX وبالتالي ادخال نماذج (GARCH, GARCHX) كنماذج هجينة .

4- أذا أكدت الاختبارات التشخيصية وجود الصفة غير الخطية في البواقي، يتم في المرحلة الرابعة بناء النماذج غير الخطية (GARCH & GARCHX) ومن ثم الحصول على مركبة قيم التنبوء النهائية $\hat{N}_t$ لمعادلة التباين الشرطي والتي يتم الحصول عليها من خلال العلاقة الرياضية الاتية [15]:

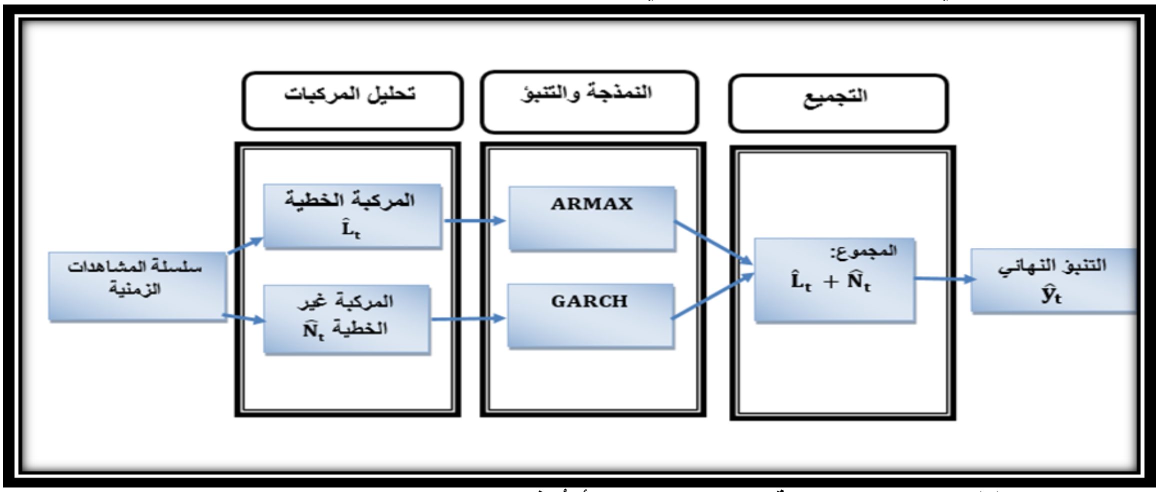
$$\tilde{N}_t(\ell) = \tilde{h}_{t+\ell\backslash t} = \tilde{\alpha}_o \sum_{i=0}^{\ell-2} (\tilde{\alpha}_1 + \tilde{\vartheta}_1)^i + (\tilde{\alpha}_1 + \tilde{\vartheta}_1)^{\ell-1}\, \tilde{h}_{t+1}\ ,\ \ell \geq 2\ ,\quad \dots(48)$$

5- ومن خلال عملية الدمج بين التنبؤات الخطية $\tilde{L}_t$ من أنموذج ARMAX الأمثل مع التنبؤات غير الخطية $\tilde{N}_t$ من نماذج GARCH و GARCHX المثلى سنحصل على القيم المستقبلية المتنبأ بها من النماذج الهجينة $\hat{y}_t(\ell)$ وعلى النحو التالي :

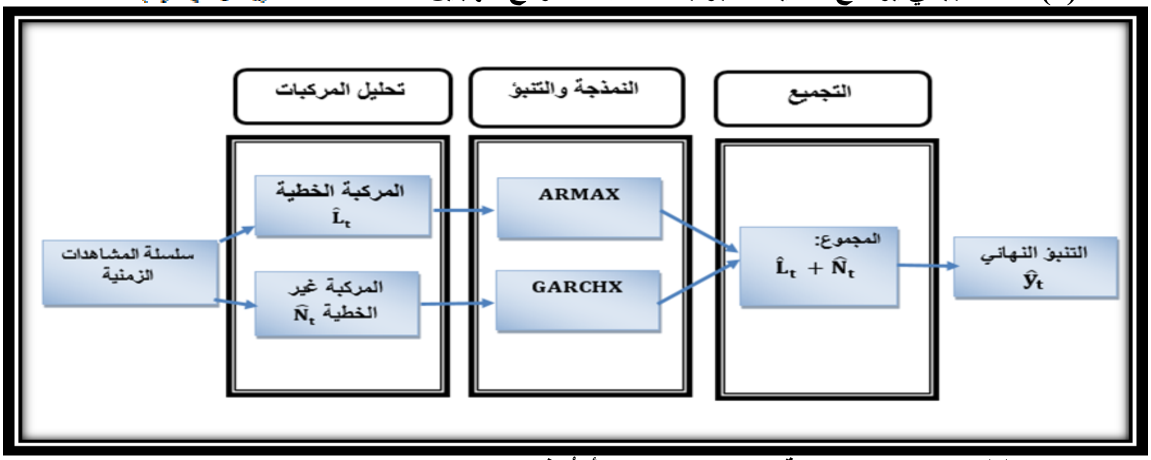
$$\hat{y}_t(\ell) = \hat{L}_t(\ell)_{ARMAX} + \tilde{N}_t(\ell)_{GARCH}\ ,\quad \rightarrow \quad ARMAX - GARCH \quad \dots(49)$$

$$\hat{y}_t(\ell) = \hat{L}_t(\ell)_{ARMAX} + \tilde{N}_t(\ell)_{GARCHX}\ ,\quad \rightarrow \quad ARMAX - GARCHX \quad \dots(50)$$

وفيما يلي خوارزمية توضح آلية عملية التنبؤ باستعمال المنهجية الهجينة للنموذجين طبقاً للخطوات السابقة، في الشكلين 3 و 4 على التوالي:[7]



شكل (3) مخطط بياني يوضح عملية التنبؤ باستعمال الأنموذج الهجين $ARMAX(n_p, n_q, n_b, n_k) - GARCH(r,s)$



شكل (4) مخطط بياني يوضح عملية التنبؤ باستعمال الأنموذج الهجين $ARMAX(n_p, n_q, n_b, n_k) - GARCHX(r,s,b)$

5.3 مقاييس دقة التنبؤ Forecast accuracy measures

[7] المخططات البيانية رقم 3 و 4 لا تمثل مراحل بناء نماذج السلاسل الزمنية الهجينة انما تمثل خوارزمية لعملية التنبؤ باستعمال المنهجية الهجينة وفقاً للصيغ رقم (49) و (50) على التوالي .

في هذا الجزء سنُوجز مقايس مختلفة لتقييم فعالية النماذج المنفردة والهجينة والمستعملة في هذا البحث في التجربة العملية للتنبؤ. ويعتمد هذا البحث على مقاييس دقة التنبؤ لدوال الخسارة (Loss functions) الأكثر شيوعاً في التطبيقات العملية للسلاسل الزمنية وهي (RMSE, MAPE) ومقياس أخر مقترح توظيفهُ وهو مقياس (Q − LIKE) الحصين والتي تستند في حسابها الى خطأ التنبؤ (Forecast error) للنماذج والمذكورة صيغتهُ في المعادلة رقم (47). ويمكن تعريف هذه المقاييس رياضياً كالاتي [18,15]:

- متوسط مطلق الخطأ النسبي Mean Absolute Percentage Error (MAPE)

يتم استعمال هذا المقياس على نطاق واسع لمقارنة اداء طرائق التنبؤ ولغرض التخلص من الاشارة والحفاظ على جميع الأخطاء الموجبة والسالبة يتم اخذ القيم المطلقة لهذه الأخطاء ويمكن التعبير عن هذا المقياس رياضياً بالصيغة الآتية: -

$$\mathbf{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\right| * 100\% \qquad , \qquad \dots(52)$$

- متوسط مطلق الخطأ

Mean Absolute Error (MAE)

يمثل هذا المقياس متوسط مطلق انحرافات القيمة الحقيقة عن القيمة المتنبأ بها ويُعرف رياضياً كالاتي:

$$\mathbf{MAE} = \frac{1}{n}\sum_{t=1}^{n}|e_t| \qquad , \qquad \dots(53)$$

- مقياس QLIKE الحصين (مقترح توظيفه) Robust Gaussian likelihood

اقترح الباحث (Patton, 2011) مقياس يكون ذو حصانه اكثر من المقاييس الأخرى لمقارنة النماذج المتنافسة والتي تتصف بواقيها بالتقلبات الزمنية، واثبت ان هذا المقياس لا يتأثر بالصدمات الخارجية التي تطرأ على البيانات، ويعطى وفقا للصيغة الاتية [32,9]:

$$\mathbf{Q - LIKE} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{y_t}{\hat{y}_{t\backslash t-1}} - ln\left(\frac{y_t}{\hat{y}_{t\backslash t-1}}\right) - 1\right) \qquad , \qquad \dots(54)$$

في المقاييس الثلاثة اعلاه : e_t يمثل خطأ التنبؤ عند الزمن t .

y_t : القيمة الحقيقة للظاهرة.

$\hat{y}_t$: يمثل القيم التنبؤية للظاهرة.

n : يمثل عدد المشاهدات المستعلمة في التنبؤ .

4. الجانب التطبيقي

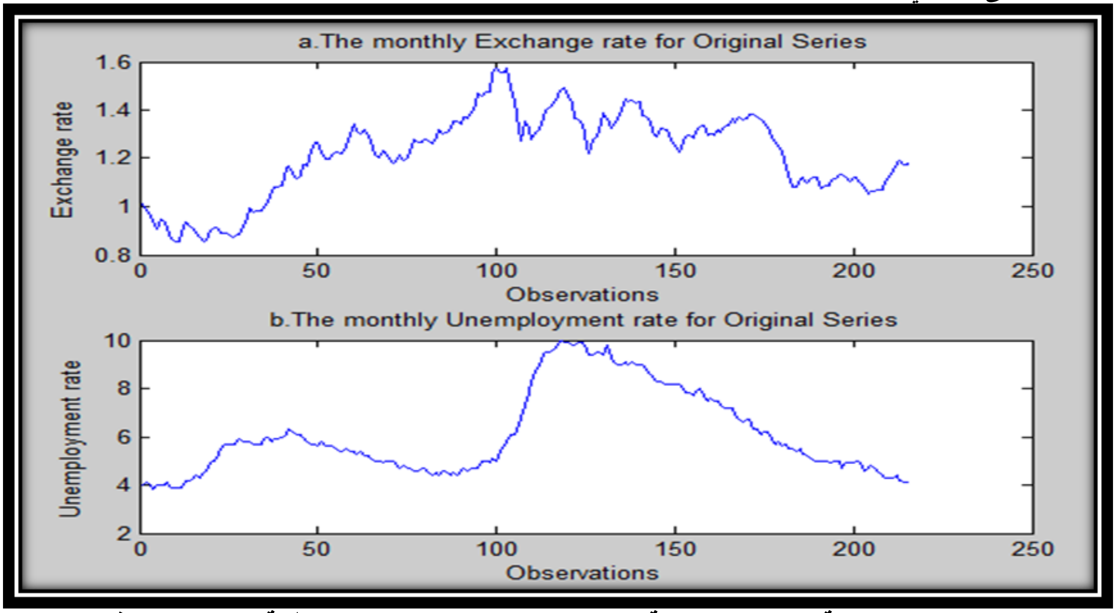
يتناول هذا القسم كافة المُجريّات العملية لتحليل مجموعات البيانات في التجربة العملية لغرض بناء النماذج الهجينة التي تم الحديث عنها في الجانب النظري من هذا البحث. حيث ركز هذا البحث بشكل أساسي على تطبيق الأساليب الهجينة مع ثلاثة نماذج توزيع مختلفة. تم بناء النماذج المنفردة والهجينة بالاعتماد على مجموعات البيانات الحقيقة وبواقي الأنمُوذج الخطي المنفرد. ثم استخدم النماذج المنفردة والهجينة للتنبؤ بالقيم المستقبلية والمفاضلة بينها وفقاً للمقاييس والاختبارات الاحصائية لتقييد دقة أداء التنبؤ.

1.4 مجموعة البيانات التطبيقية Data sets in the experiment

تمثل بيانات هذه الدراسة بيانات سلاسل زمنية ثنائية المتغيرات (bivariate time – series) لمعدل البطالة وسعر الصرف في الولايات المتحدة الأمريكية كقياسات شهرية للفترة من (كانون الثاني 2000 إلى ديسمبر 2018) حيث سيتم استخدام أول (216) مشاهدة كعينة ضابطة في تقدير وبناء النماذج الهجينة خلال الفترة الممتدة من شهر يناير (2000) حتى شهر ديسمبر (2017)، في حين تبقى المشاهدات الاثنتا عشر الأخيرة لعام (2018) كعينة متقدمة اختباريّة لفحص دقة التنبوءات المستقبلية التي سيتم الحصول عليها من الأنمُوذج المنُفرد والنماذج الهجينة [8].

2.4 النتائج التطبيقية Application results

في هذا القسم، نقوم بعملية بناء النماذج الهجينة وفقاً للمراحل التي تم التطرق اليها في الجانب النظري وبالتسلسل بالاعتماد على بيانات السلسلتين الزمنية لبناء الأنموذج الخطي ARMAX،ومن ثم توظيف بواقي هذا الأنمُوذج في عملية بناء نماذج GARCH و GARCHX غير الخطية. ويوضح الشكل $(5a)$ و $(5b)$ سلوك السلسلتين الزمنية لمعدل البطالة وسعر الصرف على التوالي :

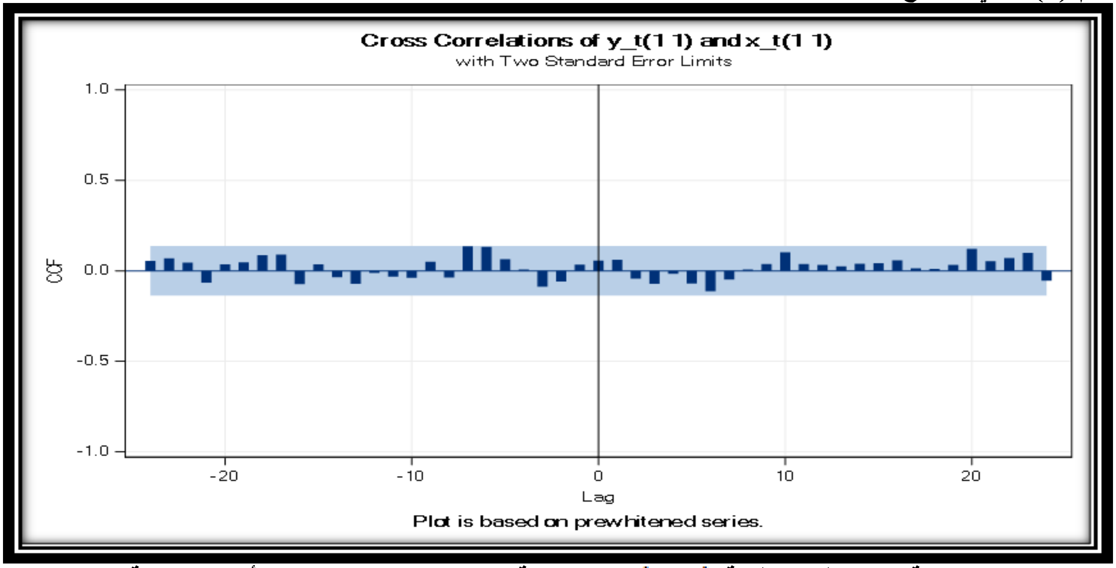


شكل (5). القياسات الشهرية للسلاسل الزمنية لـ a. سعر الصرف و b. معدل البطالة من (كانون الثاني 2000 إلى ديسمبر2017)

[8] تم عملية جمع البيانات عبر الموقع الالكتروني (fred.stlouisfed.org).

1.2.4 المرحلة الأولى : بناء أنمُوذج $\mathbf{ARMAX(n_p, n_q, n_b, n_k)}$:

بعد تهيئة سلسلتي المدخلات والمخرجات، اتضح ان السلسلتين الزمنية كانت مستقرة بعد اخذ الفرق الثاني لكل منهما، وفيما بعد تأتي واحدة من أهم مراحل بناء هذا الأنمُوذج والمتمثلة في تحديد رتبة الأنمُوذج الامثل. وللتعرف على رتبة زمن التأخير (Delay time) تم الاعتماد على دالة الارتباط التقاطعية (CCF) بين سلسلتي المدخلات (x_t) والمخرجات (y_t) والشكل البياني رقم (6) الاتي يوضح سلوك هذه الدالة :



شكل (6) : دالة الارتباط التقاطعية (CCF) بين سلسلة المدخلات المخرجات بعد أجراء عملية التبيض المسبق (Prewhitening)[9]

يتبين أن جميع معاملات الارتباط المتقاطع عند كل فجوة زمنية تقع داخل حدي فترة الثقة، بمعنى أن التغير في السلسلة الزمنية الخاصة بأسعار صرف الدولار (x_t) يؤدي الى حدوث تغير فوري في السلسلة الزمنية الخاصة بمعدل البطالة (y_t) مما يشير في هذه الحالة إلي أن رتبة زمن التأخير تساوي صفر $(n_k = 0)$.

وبعد ذلك قمنا بتوفيق أفضل أنموذج ديناميكي من خلال ملاءمة مجموعة من نماذج $\mathbf{ARMAX(n_p, n_q, n_b, n_k)}$ وبالاعتماد على مشاهدات السلاسل الزمنية المتاحة وبرتب مختلفة لاختيار الانمُوذج الأفضل للتنبؤ والذي يعطي أقل قيم للمعايير الإحصائية (AIC)، (FPE)، و(BIC) والمقدرة طبقاً للمعادلات (8)، (9)، و (10) وقد بينت النتائج أن الأنمُوذج الأمثل لتمثيل العلاقة بين السلاسل الزمنية ثنائية المتغيرات هو أنموذج :

Model	AIC	FPE	BIC
ARMAX (2, 1, 1, 0)	0.2355	1.2655	0.3292

[9] تم حساب دالة الارتباط المتقاطع (CCF) باستعمال برنامج (SAS, 9.1) بين سلسة المدخلات والمخرجات التي تمت تنقيتها أو تبيضها (Prewhitening) وذلك لتحديد أو التعرف على زمن التأخير(Delay time). ويقصد بعملية التبيض المسبق هي عملية بناء أنمُوذج ARIMA لكل سلسلة زمنية من اجل تنقية سلسلة المدخلات والمخرجات والحصول على سلسلتي البواقي المبيضة للسلسلتين الزمنية ومن ثم حساب (CCF) بينهم، ويعود السبب الرئيسي لإجراء التبييض المسبق هو للحفاظ على سلامة العلاقة الدالية بين سلسلتي الإدخال والإخراج.

نلاحظ أن رتبة الانمُوذج الديناميكي الأفضل لتمثيل العلاقة بين سلسلة المدخلات وسلسلة المخرجات هو انموذج $\mathbf{ARMAX}\ (2,1,1,0)$ وذلك لامتلاكهُ أقل قيمة للمعايير الإحصائية الثلاثة من بين عدة نماذج مختلفة تم تشخيصها [10].

2.2.4 المرحلة الثانية : مرحلة تقدير أنموذج $\mathbf{ARMAX}\ (2,1,1,0)$:

بعد الانتهاء من المرحلة الأولى والوصول الى رتبة الانمُوذج التي تم تشّخيصُها بالاعتماد على المعايير الإحصائية المذكُورة مسبقاً تكون صيغة أنموذج الانحدار الذاتي للأوساط المتحركة بمدخلات خارجية المنشأ $\mathbf{ARMAX}\ (2,1,1,0)$ المطلوب تقِدير معلماتهُ كما يلي :

$$y_t = \hat{\varphi}_1 y_{t-1} + \hat{\varphi}_2 y_{t-2} - \hat{\theta}_1 \varepsilon_{t-1} + \hat{\phi}_1 x_t + \varepsilon_t$$

ومن الانموذج أعلاه يتضح أن متجه المعلمات المطلوب تقدير معلماتهُ والذي تم توضحيهُ في المعادلة رقم (12) يكون كالاتي :

$$\beta = [\varphi_1 \;\; \varphi_2 \;\; \theta_1 \;\; \phi_1]^T$$

ومن أجل الحصول على تقديرات مُثلى تم تقدير متجه معلمات الانمُوذج أعلاه بطرائق وعوامل مختلفة وفقاً للخوارزميات المذكُورة في طرائق تقدير انمُوذج $\mathbf{ARMAX}(n_p, n_q, n_b, n_k)$ والتي تم شرحها في الجانب النظري وكما يلي :

1- يُقدر مُتجه معلمات الانمُوذج $\beta = [\varphi_1 \;\; \varphi_2 \;\; \theta_1 \;\; \phi_1]^T$ باستعمال طريقة المربعات الصغرى الاعتيادية $(\mathbf{OLS})$ تبعاً للصيغة الرياضية (13) .

2- يُقدر مُتجه معلمات الانمُوذج $\beta = [\varphi_1 \;\; \varphi_2 \;\; \theta_1 \;\; \phi_1]^T$ باستعمال طريقة المربعات التكرارية باستعمال عامل التغاضي الأسَي $(\mathbf{RLS - EF})$ تبعاً للخوارزمية المُوضحة في الشكل البياني (1).

3- وايضاً تم استعمال طريقة خطأ التنبؤ التكرارية $(\mathbf{RPEM})$ لتقدير مُتجه معلمات الانمُوذج $\beta = [\varphi_1 \;\; \varphi_2 \;\; \theta_1 \;\; \phi_1]^T$ تبعاً للخوارزمية المُوضحة في الشكل البياني (2). [11]

وبالاعتماد على هذه الطرق المختلفة فأن النتائج التقديرية التي تم الحصول عليها لمُتجه معلمات الانمُوذج $\beta = [\hat{\varphi}_1 \;\; \hat{\varphi}_2 \;\; \hat{\theta}_1 \;\; \hat{\phi}_1]^T$ والذي تم تقديرهُ تُعْطىَ من خلال الجدول (1) الاتي :

جدول (1): نتائج تقديرات معلمات انموذج $\mathbf{ARMAX}\ (2,1,1,0)$					
Error measure مقاييس المقارنة	القيم التقديرية لمعلمات انموذج $\mathbf{ARMAX}\ (2,1,1,0)$				
	$\hat{\phi}_1$	$\hat{\theta}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_1$	طريقة التقدير
$MAE = 5.312$ $NRMSE = 0.904$	0.05871	-0.04793	-0.05856	-0.06111	OLS
$MAE = 0.140$ $NRMSE = 0.027$	0.1182	-0.1341	-0.4329	-0.5391	RLS − EF
$MAE = 0.184$ $NRMSE = 0.051$	0.6893	0.7735	-0.3358	-0.5317	RPEM

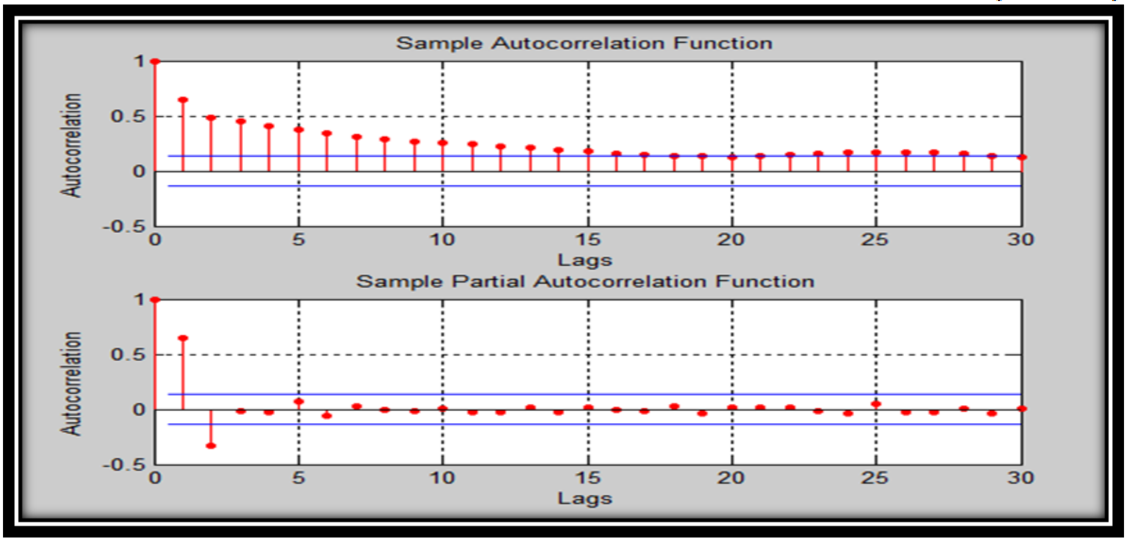
[10] تم تحديد افضل أنمُوذج من خلال كتابة برنامج باللغة البرمجية $(\mathbf{MATLAB\ (2018a)})$ من بين عدة نماذج مختلفة تم تشخيصها، حيث تم تشخيص نماذج ادنى واعلى مباشرة برتب مختلفة تم اختيارها ضمن الفترة $[0 < n_p, n_q, n_b, n_k < 10]$.

[11] الطرائق المذكُورة تم تطبيقها من خلال كتابة برنامج من قبلنا باستعمال برنامج $(\mathbf{MATLAB\ (2018a)})$ والذي يتضّمن اختبار وتشخيص الانمُوذج بالإضافة الى تحليل بواقي الانمُوذج الذي تم تقديرهُ.

نلاحظ من الجدول (1) أعلاه أن أفضل انمُوذج ديناميكي لتمثيل العلاقة بين السلسلة الزمنية للبطالة والسلسلة الزمنية لسعر الصرف هو الانمُوذج الذي يتم تقدير معلماتهِ وفقاً لطريقة المربعات الصغرى التكرارية وبتوظيف منهجّية عامل التغاضي الاسي $(RLS - EF)$ وذلك لان في هذه الطريقة كانت قيم المقاييس الإحصائية $[MAE\ (\hat{y}) = 0.140]$ وايضاً مقياس $[NRMSE(\hat{y}) = 0.0274]$ كانت أقل ما يمكن بالمقارنة مع قيم هذه المقاييس للطرق الأخرى، والمُقدرة وفقاً للمعادلتين (27) و (28) تواليا .

3.2.4 المرحلة الثالثة : مرحلة مدى ملائمة الانمُوذج (اختبار وجود تأثير نماذج ARCH):

قبل البدء في عملية تقدير نماذج $(ARCH\ \&\ GARCH)$، تأتي مرحلة فحص مدى ملائمة دقة أنمُوذج $ARMAX\ (2,1,1,0)$ من خلال الكشف عن وجود الصفات غير الخطية في سلسلة البواقي من خلال المخططات البيانية والاختبارات الاحصائية . ويعرض الشكل (7) الاتي رسم دالتي الارتباط الذاتي (ACF) والذاتي الجزئي $(PACF)$ لمربعات سلسلة البواقي $\hat{\varepsilon}_t^{\,2}$ لأنمُوذج $ARMAX\ (2,1,1,0)$:



شكل (7) : رسم دالتي (ACF) و $(PACF)$ على التوالي المُقدرتين لمربعات بواقي $\hat{\varepsilon}_t^{\,2}$ انمُوذج $ARMAX\ (2,1,1,0)$

ومن خلال رؤية الشكل (7) نلاحظ ارتفاع معظم الارتباطات الذاتية عن حدي فترة الثقة المعنوية على أقل تقدير في الفجوات الزمنية $(k = 1 - 12)$ وهذا يعني وجود صفة الارتباط الذاتي بين البواقي وبمعنى يعني أن سلسلة الاخطاء العشوائية $\hat{\varepsilon}_t$ غير خالية من النتوءات بالتالي فأنها لا تمثل تغيرات عشوائية بحته $(Pure\ Random\ Errors)$.[12]

[12] رسم دالة الارتباط الذاتي $Autocorrelation\ function, ACF$ والذاتي الجزئي $PACF$ باستعمال برنامج MATLAB من قبل الباحث.

1.3.2.4 اختبار $\mathbf{Ljung-Box\ Test}$:

لاختبار عشوائية السلسلة الزمنية لبواقي أنمُوذج $\mathbf{ARMAX\ (2,1,1,0)}$ تم استعمال اختبار $\mathbf{(Ljung-Box)}$ لاختبار الفرضية (30) والجدول (2) الاتي يوضح نتائج هذا الاختبار والتي تم تقديرها تبعاً لصيغة الاختبار (31) حيث يبين القيمة الإحصائية المحسوبة للاختبار مع القيمة الاحتمالية وكما يأتي :

جدول (2): يبين اختبار $(Ljung-Box)$ لسلسلة البواقي ε_t H_0 : no serial correlation exists			
Test	$Q_{L-B}(Lag)$	$Q_{L-B}-Statistic$	$p-value$
Ljung – Box	$Q_{L-B}(10)$	351.37	$2.2E-16 \cong 0$
Ljung – Box	$Q_{L-B}(15)$	404.09	$2.2E-16 \cong 0$
Ljung – Box	$Q_{L-B}(20)$	428.52	$2.2E-16 \cong 0$

ويتضح من خلال الجدول اعلاه أن القيمة الاحتمالية $(p-value)$ كانت اصغر من مستوى المعنوية (0.05) وهذا يعني قبول الفرضية البديلة وبمعنى أخر وجود الارتباط المتسلسلة في بواقي الانموذج الخطي عند الإزاحات الزمنية $(k = 10,15,20)$ وبالتالي وجود حالة عدم تجانس التباين في البواقي.

2.3.2.4 اختبار BDS $\mathbf{BDS-Test}$

تم تطبيق اختبار BDS للكشف عن وجود الصفة غير الخطية في البواقي التقديرية $\hat{\varepsilon}_t$ لأنمُوذج $\mathbf{ARMAX\ (2,1,1,0)}$ أي لاختبار الفرضية الإحصائية (33) والجدول الاتي يوضح نتائج هذا الاختبار والتي تم تقديرها تبعاً لصيغة الاختبار (34) حيث يبين القيمة الإحصائية المحسوبة للاختبار مع القيمة الاحتمالية[13]:

جدول (3): يبين نتائج اختبار BDS لسلسلة البواقي ε_t $H_0 = \text{linearity in } \hat{\varepsilon}_t$			
Test	dimension	BDS – Statistic	$p-value$
BDS	$m = 2$	0.181994	0.0000
BDS	$m = 3$	0.305099	0.0000
BDS	$m = 4$	0.386709	0.0000

تُبين النتائج في الجدول رقم (3) السابق وجود الصفة غير الخطية وبأبعاد مختلفة في السلسة الزمنية لبواقي أنموذج $\mathbf{ARMAX\ (2,1,1,0)}$ حيث كانت قيمة $(p-value)$ أصغر من مستوى المعنوية المحدد (0.05) ولجميع الابعاد وهذا دليل على قبول الفرضية البديلة التي تنص على وجود الصفة غير الخطية في البواقي.

4.2.4 المرحلة الرابعة: بناء نماذج $\mathbf{(ARCH(r), GARCH(r,s)\ \&GARCHX(r,s,b))}$ المُلائِمة في عملية التهجين

يقدم هذا الجزء من البحث إجْرَاءَات التقدير للنماذج غير الخطية وذلك بالاعتماد على بواقي أنمُوذج $\mathbf{ARMAX\ (2,1,1,0)}$. حيث قمنا بتوفيق أفضل أنمُوذج غير خطي من خلال ملاءمة مجموعة من نماذج $\mathbf{(ARCH, GARCH\ \&GARCHX)}$ وبرتب مختلفة لاختيار الانمُوذج الأفضل للتنبؤ بالتذبذبات الزمنية وقد بينت النتائج أن النماذج المثلى لتمثيل بواقي الأنمُوذج الخطي هي أنموذج $\mathbf{GARCH(1,1)}$ و أنمُوذج $\mathbf{GARCHX(1,1,1)}$ وفقاً الى أقل القيم للمعايير الإحصائية (AIC)، (BIC)، و (HQC) والمقدرة طبقاً للمعادلات (8)، (9)، و (35)، والتي تم

---

[13] نتائج اختبار BDS تم التوصل اليها من خلال تطبيق برنامج (EViews 9) وذلك لان هذا البرنامج يتيح خيارات تفصيلية اكثر من بقية البرامج المستعلمة لهذا الاختبار. أما اختبار $(Ljung-Box)$ تم تطبيقهُ باستعمال برنامج (R) .

توظيفها في عملية التهجين مع أنموذج $ARMAX\ (2,1,1,0)$ بثلاث توزيعات مختلفة للخطأ العشوائي لغرض السيطرة على الصفات غير الخطية للبواقي. والجدول التالي يوضح النماذج الهجينة المقترحة بتوزيعات مختلفة (Normal, Student's t and GED Distributions) جنباً الى جنب مع المعايير الإحصائية المستعملة في المفاضلة بين هذه النماذج[14]:

الجدول (4). معايير المقارنة لاختيار أفضل أنمُوذج هجين تبعاً لتوزيعات الخطأ العشوائي المقترحة

ε_t Distribution التوزيعات المقترحة	Hybrid Models النماذج الهجينة	Criteria معايير المقارنة		
	ARMAX – GARCH	AIC	BIC	HQC
Normal Distribution	ARMAX (2, 1, 1, 0) – GARCH(1, 1)	−1.163546	−1.116667	−1.144606
Student −t Distribution	ARMAX (2, 1, 1, 0) – GARCH(1, 1)	−1.154272	−1.091767	−1.129020
Generalized Error Distribution	ARMAX (2, 1, 1, 0) – GARCH(1, 1)	−1.406024	−1.343519	−1.380772
ε_t Distribution	ARMAX – GARCHX	AIC	BIC	HQ
Normal Distribution	ARMAX (2, 1, 1, 0) – GARCHX(1, 1, 1)	−1.166272	−1.103767	−1.141020
Student −t Distribution	ARMAX (2, 1, 1, 0) – GARCHX(1, 1, 1)	−1.153658	−1.075526	−1.122092
Generalized Error Distribution	ARMAX (2, 1, 1, 0) – GARCHX(1, 1, 1)	−1.428461	−1.350330	−1.396896

يتضح من الجدول (4) السابق بعد ملائمة (Fitting) ثلاثة أنماط مختلفة للتوزيعات العشوائية لبواقي انمُوذج $ARMAX\ (2,1,1,0)$ الخطي وذلك لغرض السيطرة على التذبذبات الزمنية التي ظهرت في السلسلة الزمنية لبواقي الانمُوذج الاخير اتضح من الجدول أعلاه أن النماذج الهجينة الأمثل للسلسلتين الزمنية قيد التحليل هي أنمُوذج $[ARMAX\ (2,1,1,0) - GARCH(1,1)]$ وأنمُوذج $[ARMAX\ (2,1,1,0) - GARCHX(1,1,1)]$ المعتمدة على الخطأ العشوائي الذي يتبع توزيع الخطأ المعمم (Generalized error distribution) حيث حسْن من جَودَة النماذج الهجينة من خلال تَفوّقهِ على النماذج الهجينة الأخرى عند اتباع الخطأ العشوائي التوزيع الطبيعي وتوزيع (Student − t) بأقل القيم للمعايير الإحصائية (AIC, BIC, HQC) .

1.4.2.4 تقدير معلمات نماذج عملية التهجين $GARCH(1,1)$ و $GARCHX(1,1,1)$

في هذا البحث ولعملية التقدير تم استعمال أسلوب ذات المرحلتين لبناء النماذج الهجينة، المرحلة الأولى تم عرض نتائجها مسبقاً في القسم (2.2.4). و بعد مُلاءَمَة الانمُوذج المناسب للسلسلة الزمنية لبواقي أنمُوذج $ARMAX\ (2,1,1,0)$ يتم في هذه المرحلة والتي تمثل المرحلة الثانية لعملية التقدير، حيث يتم فيها تقدير معلمات أنمُوذج $GARCH(1,1)$ و $GARCHX(1,1,1)$ والمذكورة صيغة كل منهم في معادلة رقم (6) و (7) على التوالي عند اتباع السلسلة الزمنية للبواقي توزيع $(\varepsilon_t \sim GED)$ لبناء الانمُوذج الهجين

[14] تم تحديد افضل أنمُوذج من بين عدة نماذج مختلفة تم تشخيصها حيث تم تشخيص نماذج ادنى واعلى مباشرة برتب مختلفة تم اختيارها ضمن حدود الرتب : $[r = 1,2,3.\ \ s = 1,2, b = 1]$، ولكثرة النتائج تم ذكر فقط الانمُوذج الامثل من بين النماذج المرشحة.

(Hybrid model) باستعمال طريقة الإمكان الأعظم وبتوظيف خوارزمية التحسين العددي (BHHH) والتي تم ايضاحها في القسم (3.4.3.3) من هذا البحث. والنتائج التقديرية لمعلمات نماذج التهجين مدرجة في الجدول رقم (5) الاتي:

جدول (5): تقديرات الإمكان الاعظم لمعلمات أنمُوذج $GARCH(1,1)$ و $GARCHX(1,1,1)$ تبعاً لتوزيع GED

Models	Par.	Values	Stand. error	t − value	p-value
Variance equation for $GARCH(1,1)$ $h_t = \hat{\alpha}_o + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \hat{\vartheta}_1 h_{t-1}$	$\hat{\alpha}_o$	0.000631	0.000222	2.839891	0.0045
	$\hat{\alpha}_1$	0.609532	0.102801	5.929237	0.0000
	$\hat{\vartheta}_1$	0.271026	0.038499	7.039741	0.0000
	$\hat{v}$	4.783030	1.202647	3.977084	0.0001
Variance equation for $GARCHX(1,1,1)$ $h_t = \hat{\alpha}_o + \hat{\alpha}_1 \varepsilon_{t-1}{}^2 + \hat{\vartheta}_1 h_{t-1} + \hat{\omega}_1 x_{t-1}^2$	$\hat{\alpha}_o$	0.002267	0.001152	1.967975	0.0491
	$\hat{\alpha}_1$	0.622778	0.129305	4.816344	0.0000
	$\hat{\vartheta}_1$	0.241299	0.036489	6.613002	0.0000
	$\hat{\omega}_1$	0.002686	0.001226	2.190962	0.0285
	$\hat{v}$	4.971380	1.466313	3.390394	0.0007

الجدول (5) اعلاه يتضمن النتائج التقديرية لطريقة تقدير معلمات النماذج $GARCH(1,1)$ و $GARCHX(1,1,1)$ بالإضافة الى ($\hat{v}$) معلمة توزيع (GED) حيث يبين إحصائية اختبار t لمعنوية معلمات معادلة التباين الشرطي للنموذجين وكذلك قيم الخطأ المعياري لكل منها .ويتضح من خلالهُ ان القيمة الاحتمالية $(p - value)$ لجميع المعلمات $\beta = [\hat{\alpha}_o \ \hat{\alpha}_1 \ \hat{\vartheta}_1 \hat{\omega}_1]^T$ المُقدرة للنموذجين كانت أقل من مستوى المعنوية $(\alpha = 0.05)$ وهذا دليل على معنوية معلمات الانمُوذج.

وبعد أن تم تقدير وتشخيص معنوية الأنموذجين $(GARCH(1,1) - GED)$ و $GARCHX(1,1,1)$ يمكن كتابة صيغة النماذج الهجينة لهذا البحث كالاتي :

معادلة المتوسط الشرطي للنماذج الهجينة الأول والثاني من الجدول رقم (1) تكتب كالاتي :

$$y_t = 0.5391 y_{t-1} + 0.4329 y_{t-2} + 0.1341 \varepsilon_{t-1} + 0.1182\, x_t + \varepsilon_t\ ,$$

$$\varepsilon_t = \eta_t \sqrt{h_t} \quad , \quad \eta_t \sim GED\ (4.78)$$

ومعادلة التباين الشرطي للأنمُوذج $[ARMAX\ (2,1,1,0) - GARCH(1,1)]$ الهجين الأول تبعاً لتوزيع الخطأ المعمم (GED) من الجدول (5) السابق تكتب كالاتي:

$$h_t = 0.000061 + 0.609532\ \varepsilon_{t-1}^2 + 0.271026\ h_{t-1}$$

ومعادلة التباين الشرطي للأنموذج $[\mathbf{ARMAX}\ (2,1,1,0) - \mathbf{GARCHX}(1,1,1)]$ الهجين الثاني تبعاً لتوزيع الخطأ المعمم $(\mathbf{GED})$ من الجدول (5) السابق تكتب كالاتي:

$$\mathbf{h_t} = 0.002267 + 0.622778\, \varepsilon_{t-1}^2 + 0.241299\, \mathbf{h}_{t-1} + 0.002686\, x_{t-1}^2$$

وان معلمات النماذج غير الخطية تحقق شروط النظرية للنماذج، وعملية دمج المعادلات أعلاه يمثل تكوين للنماذج الهجينة والتي سيتم توظيفها في التنبؤ بمعدل البطالة تبعاً لمنهجية $(\mathbf{Zhang})$ للنماذج الهجينة جنباً الى جنب مع الأنموذج المنفرد . والقسم التالي يوضح ذلك:

3.4 التنبؤ باستعمال الأنموذج المُنفرد، والنماذج الهجينة تبعاً لمنهجية $(\mathbf{Zhang})$ الهجينة

بعد عدة مراحل متسلسلة لنمذجة السلاسل الزمنية ثنائية المتغيرات للظاهرة قيد الدراسة بداية من مرحلة التعرف الى مرحلة التقدير والتشخيص تأتي أهم مرحلة من مراحل تحليل السلاسل الزمنية والتي تُعد أحد أهم أهداف هذا البحث ألا وهي الحصول على القيم التنبؤية النهائية لمعدل البطالة باستعمال النماذج المُنفردة والهجينة والمقارنة بينهم باستعمال أسلوبين مختلفين وهما مقاييس دقة التنبؤ والاختبارات الإحصائية.

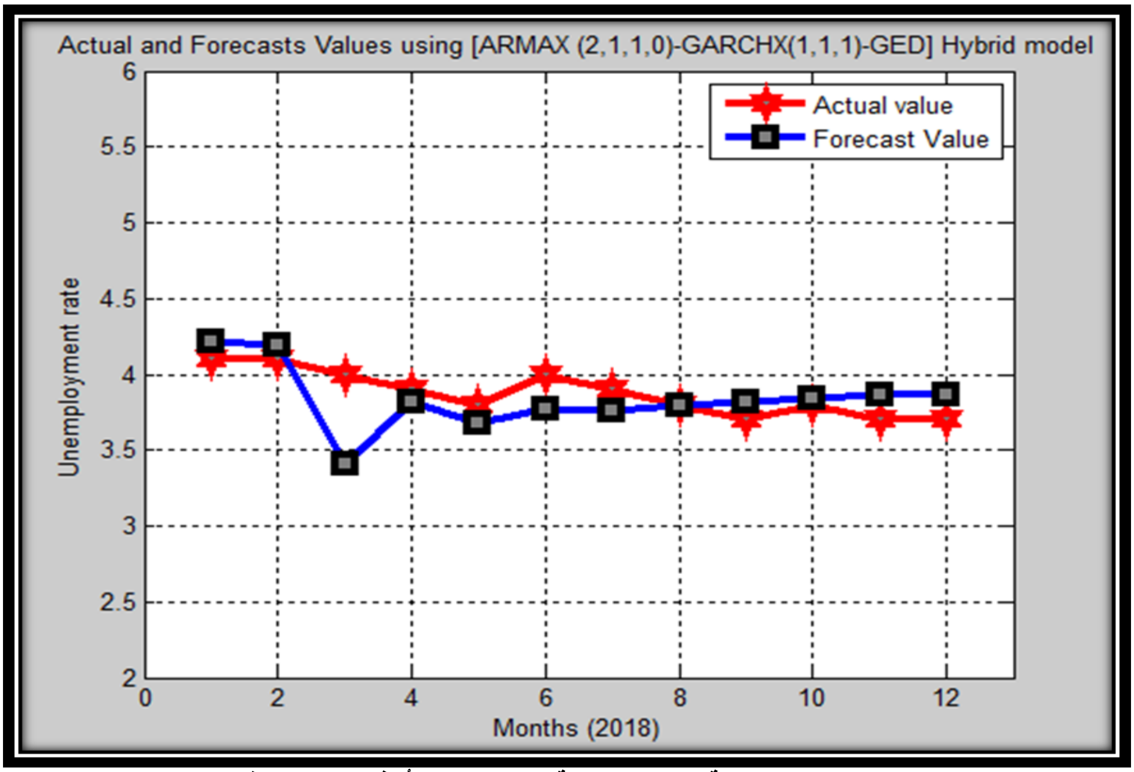
القيم الحقيقة والقيم التنبؤية التي تم الحصول عليها باستخدام أنموذج $\mathbf{ARMAX}\ (2,1,1,0)$ المُنفرد بتوظيف المعادلة رقم (46) والأنموذج $[\mathbf{ARMAX}\ (2,1,1,0) - \mathbf{GARCH}(1,1)]$ الهجين الأول والأنموذج $[\mathbf{ARMAX}\ (2,1,1,0) - \mathbf{GARCHX}(1,1,1)]$ الهجين الثاني تبعاً لتوزيع الخطأ العام $(\mathbf{GED})$ وذلك من خلال توظيف المعادلات الهجينة رقم (49) و (50) لمنهجية $(\mathbf{Zhang})$ لاثنا عشر شهراً لمعدل البطالة في عام (2018) ، هذه القيم مدرجة في الجدول رقم (6) الاتي جنباً الى جنب مع مقاييس دقة التنبؤ $(\mathbf{MAPE}, \mathbf{RMSE}, \mathbf{Q-LIKE})$ لكل أنموذج والتي تم تقدرها من خلال تطبيق الصيغ في المعادلات الرياضية (52)،(53)،و(54) على التوالي :

جدول (6): القيم الحقيقة والتنبؤية بأستعمال الانموذج المُنفرد والنماذج الهجينة مع مقاييس دقة التنبؤ لكل أنموذج

التاريخ Date	Actual value yt القيم الحقيقية	Forecast values $\hat{y}_t(\ell) = \hat{L}_t + \hat{N}_t$ ARMAX الانموذج المُنفرد	Forecast values $\hat{y}_t(\ell) = \hat{L}_t + \hat{N}_t$ ARMAX-GARCH الأنموذج الهجين الاول	Forecast values $\hat{y}_t(\ell) = \hat{L}_t + \hat{N}_t$ ARMAX-GARCHX الأنموذج الهجين الثاني
Jan/2018	4.1	4.1591	4.2162	4.2152
Feb/2018	4.1	4.1416	4.1977	4.1967
Mar/2018	4.0	3.3575	3.4148	3.4138
Apr/2018	3.9	3.7569	3.8059	3.8169
May/2018	3.8	3.6261	3.6731	3.6751
Jun/2018	4.0	3.7342	3.7716	3.7729
Jul/2018	3.9	3.7326	3.7647	3.7553
Aug/2018	3.8	3.7728	3.8040	3.7946
Sep/2018	3.7	3.8026	3.8310	3.8214
Oct/2018	3.8	3.8206	3.8474	3.8376
Nov/2018	3.7	3.8551	3.8705	3.8697
Dec/2018	3.7	3.8497	3.8649	3.8644
Forecasts Evaluation Criteria مقاييس دقة التنبؤ		MAPE = 4.16% MAE = 0.1624 Q − LIKE = 0.00197	MAPE = 4.06% MAE = 0.1585 Q − LIKE = 0.00164	MAPE = 4.006% MAE = 0.1564 Q − LIKE = 0.00160

نلاحظ من الجدول أعلاه وفقاً للحد الأدنى من مقاييس دقة التنبؤ أن النماذج الهجينة أكثر دقة وكفاءة في التنبؤ المستقبلي مقارنة بالأنموذج المنفرد. وكذلك فأن الأنموذج الهجين الثاني $[\mathbf{ARMAX}\ (2,1,1,0)-\mathbf{GARCHX}(1,1,1)]$ قد حَسّن من جُودَة الانموذج الهجين الاول $[\mathbf{ARMAX}\ (2,1,1,0)-\mathbf{GARCH}(1,1)]$ وزاد من فعّاليّة الانموذج في القدرة على التنبؤ المستقبلي، وهذا يعني ادراج السلسلة الزمنية لمتغير التنبؤ الخارجي في معادلة التباين الشرطي للأنموذج الهجين كان ذو تأثير إيجابي في قدرة التنبؤ للنماذج الهجينة[15].

والشكل البياني التالي يوضح القيم الحقيقية والقيم المتنبأ باستعمال الأنموذج الهجين الثاني $[\mathbf{ARMAX}\ (2,1,1,0)-\mathbf{GARCHX}(1,1,1)]-\mathbf{GED}$ باثنتي عشرة قيمة مستقبلية لمتغير البطالة من يناير حتى ديسمبر (2018) :



شكل (8) القيم الحقيقة والقيم التنبؤية باستعمال الانموذج الهجين الثاني $[\mathbf{ARMAX}\ (2,1,1,0)-\mathbf{GARCHX}(1,1,1)-\mathbf{GED}]$ باثنتي عشرة قيمة مستقبلية لمتغير البطالة.

[15] النتائج والرسوم البيانية تم التوصل اليها من خلال كتابة برنامج بلغة ($\mathbf{MATLAB}\ (2018a)$) من قبل الباحث.

5. الاستنتاجات والتوصيات:

1.5 الاستنتاجات Conclusions

باستعمال نتائج البيانات الحقيقية للسلاسل الزمنية ثنائية المتغيرات قد تم التوصل إلى جملة من الاستنتاجات فيما يلي نصّها:

1- بعد مطابقة عدة نماذج لمشاهدات السلاسل الزمنية (y_t) و (x_t) المتاحة، واعتماداً على أقل القيم لمعايير التقييم $(\mathbf{AIC})$، $(\mathbf{BIC})$ و $(\mathbf{FPE})$ للمفاضّلة بين النماذج المقترحة، فقد اتضح أن انمُوذج $\mathbf{ARMAX}\ (2,1,1,0)$ هو الأنمُوذج الأمثل والملائم لتحليل بيانات السلاسل الزمنية الشهرية، وذلك من خلال تفوّقهُ على غيره من النماذج المقترحة في اجتياز جميع الفحوص والاختبارات التشخيصية بنجاح.

2- بينت نتائج مرحلة تقدير أنمُوذج $\mathbf{ARMAX}\ (2,1,1,0)$ أن طريقة $(\mathbf{RLS-EF})$ كانت الطريقة المُثلى في عملية تقدير مُتجه معلمات الأنمُوذج بالمقارنة مع طرائق التقدير $[\mathbf{OLS}, \mathbf{RPEM}]$ وتبين ذلك من خلال إقتِناء أقل القيم لمقاييس المقارنة $[\mathbf{NRMSE}(\hat{y}), \mathbf{MAE}(\hat{y})]$.

3- عند تحليل السلسلة الزمنية لبواقي الأنمُوذج الخطي ومن خلال رسم دالة $(\mathbf{ACF})$ و $(\mathbf{PACF})$ كما في الشكل البياني (7) لمربعات البواقي $\hat{\varepsilon}_t^{\,2}$، بالإضافة الى تطبيق الاختبارات التشخيصية $[\mathbf{Ljung-Box}, \mathbf{BDS-Test}]$ كما في الجدول (2) و (3) توالياً، إتّضَح من ذلك ان السلسلة الزمنية للبواقي التقديرية تمتاز بوجود تأثير $(\mathbf{ARCH})$ وايضاً بالصفة غير الخطية.

4- في عملية التهجين بينت النتائج أن الانمُوذج الأفضل الذي تخضع لهُ سلسلة البواقي في عملية التهجين من بين عدة نماذج غير خطية مُقتَرَحه وبرتب مختلفة هو أنمُوذج $\mathbf{GARCH}(1,1)$ تبعاً لتوزيع الخطأ العام $(\hat{\varepsilon}_t \sim \mathbf{GED})$ وفقاً لأقل القيم لمعايير اختيار رتبة الأنمُوذج $(\mathbf{AIC}, \mathbf{BIC}, \mathbf{HQC})$ وباقل عدد ممكن من المعلمات.

5- أنْتَج دمج السلسلة الزمنية (x_t) مره أخرى ولكن في معادلة التباين الشرطية، وبناء أنمُوذج $(\mathbf{GARCHX}(1,1,1))$ بثلاث توزيعات للخطأ العشوائي تأثير إيجابي في عملية التنبؤ، واتضح ذلك من خلال نتائج الانمُوذج الهجين الثاني $[\mathbf{ARMAX}\ (2,1,1,0) - \mathbf{GARCHX}(1,1,1)]$ تبعاً لتوزيع $(\mathbf{GED})$، حيث حسّن من جَودَة الأنمُوذج الهجين الأول والأنمُوذج المنفرد ايضاً من خلال تفوّقهِ على هذه النماذج باقتناء أقل القيم لمقاييس دقة التنبؤ $[\mathbf{MAE}, \mathbf{MAPE}, \mathbf{Q-LIKE}]$.

2.5 التوصيات Recommendations

في ضوء النتائج السابقة والتي تعد ذات أهمية كبيرة للظاهرة قيد الدراسة، ومن خلال ما تقدم في هذا البحث يمكن التوصية بما يلي:

1- نوصي الباحثين بزيادة الاهتمام باستعمال النماذج الهجينة للتنبؤ بالسلاسل الزمنية ثنائية المتغيرات التي تتضمن بياناتها تقلبات زمنية، وذلك لقدرة هذه النماذج على سد النقص الذي تعاني منهُ النماذج المنفردة.

2- نوصي بأجراء عملية التهجين ما بين النماذج الخطية وغير الخطية، ونقترح منها أنموذج هجين يجمع ما بين أنموذج $(\mathbf{ARMAX})$ الخطي وأنموذج الشبكات العصبية غير الخطي $(\mathbf{ANN})$،وذلك لأنَّها تتميز باحتفاظها بأحداث الماضي مما يؤدي إلى زيادة دقَّة التنبؤ.

3- نوصي بدراسة النماذج الهجينة من خلال الدمج ما بين أنمُوذج الانحدار الذاتي للأوساط المتحركة الخطي بمدخلات لمتغيرات خارجية متعددة وبمُخْرَجٍ واحد $(\mathbf{MISO-ARMAX})$ أو أنمُوذج متعدد المدخلات والمخرجات $(\mathbf{MIMO-ARMAX})$ مع النماذج غير الخطية $(\mathbf{GARCH})$ أو $(\mathbf{ANN})$.

4- اقتــراح اجــراء دراســة مُعَمَّقــه لنمــاذج السلاســل الزمنيــة الهجينــة الكســرية مثــل أنمُــوذج $(AFRIMAX-GARCH)$ وكــــذلك للنمــــاذج الهجينــــة الموســــمية مثــــل أنمــــوذج $(SARMAX-GARCH)$.

5- نوصــي بتطــوير عمليــة التهجين مــن نمــوذجين الــى ثــلاث نمــاذج ونقتــرح بدراســة أنمــوذج يجمــع مــا بين $(ARMAX-GARCH-ANN)$.

المصادر References

1. Aldemġr, A and Hapoğlu, H.(2015)." Comparison of ARMAX model identification results based on least squares method". IJMTER, Vol. (02), No.(10), PP(27-35).
2. Arnerić ,H. Babić ,Z. and Škrabić,B.(2007)." Maximization of the likelihood function in financial time series models". Quantitative Economics and Finance. pp(1-12).
3. Berndt, E.; Hall, B.; Hall, R.; Hausman, J. (1974). "Estimation and inference in nonlinear structural models" . Annals of Economic and Social Measurement. 3 (4): 653–665
4. Bollerslev, T. (1987). "A Conditional heteroscedastic time series model for speculative prices and rates of return". Review of Economics and Statistics, 69, 542-547.
5. Bollerslev, Tim.(1986). "Generalized Autoregressive Conditional Heteroscedasticity,". Journal of Econometrics, Vol. (31), No (3), pp (307-327).
6. Botchkarev,A.(2019). "A new typology design of performance metrics to measure errors in machine learning regression algorithms". Interdisciplinary Journal of Information, Knowledge and Management. Volume (14) .pp. (045-076).
7. Brock,.W. A., Dechert,.W., Scheinkman,.J., and LeBaron,.B. (1996)."A test for independence based on the correlation dimension". Economic Reviews, Vol(15),No.(3).pp(197–235).
8. Brooks, C. (2008) ."Introductory econometrics for finance". 2nd edition. Cambridge University Press.
9. Charles, A., and Darné, O. (2018). "The accuracy of asymmetric GARCH model estimation". International Economics, JEL Classification: C22.
10. Edward, N., (2011). "Modeling and forecasting using time series GARCH models: an application of tanzania inflation rate data". Master thesis of Science (Mathematical Modeling) of the university of Dares Salaam, Morocco.
11. Enders, W. (2014). "Applied econometric time series". 4th Edition. John Wiley & Sons Inc, New York.
12. Engle, R. (1982)."Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation". Econometrica ,Vol.(50),No(4),pp (987-1007).

13. Engle, R. (2001). "GARCH 101: The use of ARCH/GARCH models in applied econometrics". Journal of Economic Perspective, Vol. (15), No. (4), pp.(157-168).

14. Feng, L., & Shi, Y. (2017). "A simulation study on the distributions of disturbances in the GARCH model". Cogent Economics and Finance, Vol.(5), No.(1), pp (1-19).

15. Franses, P. H., van Dijk, D. J. C., and Opschoor, A. (2014). "Time series models for business and economic forecasting", 2nd Edition. Cambridge University Press.

16. Greene, William H. (2012). "Econometric analysis", 7th Edition. Upper Saddle River, NJ: Prentice-Hall.

17. Guidolin, M., Pedio, M. (2018) ."Essentials of time series for financial applications". London: Elsevier Inc.

18. Hamoud.,M,Y. Gomaa,.A,A. and Firas,.A, M.(2019)."Time Series Analysis". Part I. AL-Daad, Baghdad, Iraq.

19. Han, H., and Kristensen, D. (2014). "Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates". Journal of Business and Economic Statistics , Vol.(32),No.(3), pp(416–429).

20. Hickey, E., Loomis, D. G., & Mohammadi, H. (2012). "Forecasting hourly electricity prices using ARMAX–GARCH models: An application to MISO hubs". Energy Economics, Vol (34), No(1), pp(307–315).

21. Hurn, S. (2009)." Likelihood methods in financial econometrics". Economic Research Southern Africa, University of Stellenbosch.

22. IHS Global Inc. (2016). "EViews 9 user's guide II". California: IHS Global Inc.

23. Kanjilal, P.P.,(1995)."Adaptive Prediction and Predictive Control", Peter Peregrines Ltd. London.

24. Ljung, Greta M. and Box E.P. (1978), "On a measure of lack of fit in time series models". Biometrika, Vol. (65), No. (2), pp (297-303).

25. Ljung, L., (1999)."System identification theory for user",2nd ed. Prentice Hall Upper Saddle River N.J. London UK.

26. Modarres, R., & Ouarda, T. B. M. J. (2013). "Modeling rainfall–runoff relationship using multivariate GARCH model". Journal of Hydrology, Vol. (499), pp (1–18). New Zealand.

27. Morata,A and Loira,L (2017). "Yeast - Industrial applications". Rijeka, Croatia.

28. Nana,.G, Korna,R., and Sayera,E (2013)."GARCH-extended models: theoretical properties and applications". American Economic Association, JEL Classification: C32, C50.

29. Navrátil,P and Ivanka,J.(2014)."Recursive estimation algorithms in Matlab & simulink development Environment". Wseas Transactions On Computers .Vol. (13).pp(691-702).

30. Nelles, O. (2001). "Nonlinear System Identification from Classical Approaches to Neural Networks and Fuzzy Models". Springer, New York.

31. Nelson,D.B. ,.(1991)."Conditional heteroskedasticity in asset returns: A new Approach", Econometrica ,Vol (59), pp(347-370).

32. Patton, A., (2011). "Volatility forecast comparison using imperfect volatility proxies", Journal of Econometrics, Vol.(160),No.(1), pp(246-256).

33. Porshnev, A.,Valeria ,L. and Ilya,R,(2016). "Could emotional markers in twitter posts add information to the stock market ARMAX-GARCH model". Higher School of Economics Research Paper No. WP BRP 54/FE.

34. Rachev, S., Mittnik, S., Fabozzi, F., Focardi, S., and Jasic, T. (2007). "Financial econometrics: From basics to advanced modeling techniques". John Wiley & Sons, Inc. New York.

35. Sadorsky, P. (2006). "Modeling and forecasting petroleum futures volatility" .Energy Economics, Vol.(28), pp(467-488).

36. Soderstrom, T. and Stoica, P. (2001) ."System Identification" Prentice-Hall International, Hemel Hempstead, U.K.

37. Stephen J. G., and Robert A.M (2011)."Nonlinear dynamical systems analysis for the behavioral sciences using real data". Taylor and Francis Group, LLC.

38. Taebi, A., and Mansy, H. (2017). "Time-Frequency distribution of seismocardio graphic signals". A Comparative Study. Bioengineering, Vol.(4), No(32), pp(1-22).

39. Tsay, Ruey S.(2013). "An introduction to analysis of financial data with R". John Wiley & Sons .Hoboken.

40. Wellstead, P. E., & Zarrop, M. B. (1991). "Self-tuning systems: Control and signal processing". Chichester , UK: John Wiley & Sons.

41. Xekalaki, E. and Degiannakis, S. (2010). "ARCH models for financial applications", John Wiley & Sons Ltd., New York.

42. Yaziz, S.R. Azizan, N. A. Zakaria, R. and Ahmad, M.H.,(2013). "The performance of hybrid ARIMA-GARCH modeling in forecasting gold price ". International Congress on Modelling and Simulation, Adelaide, Australia.

43. Zhang, G.P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model". Neurocomputing. Comput., Vol.(50), pp(159-175).

44. Zhao, J. H., Dong, Z. Y., & Zhao, M. L. (2009). "A statistical model for flood forecasting". Australasian Journal of Water Resources, Vol.(13), No.(1), pp(43–52). Australia.

Applying some hybrid models for modeling bivariate time series assuming different distributions for random error with a practical application

Firas Ahmmed Mohammed [a] and Moamen Abbas Mousa [b]

[a] Ass.Prof. Department of statistics, College of Management and Economics, Baghdad University, Baghdad, Iraq. (drfirasmohana@gmail.com)

[b] PhD student. Department of Statistics, College of Management and Economics, Baghdad University, Baghdad, Iraq. (saidmoamen@gmail.com)

Abstract

Bivariate time series modeling and forecasting have become a promising field of applied studies in recent times. For this purpose, the Linear Autoregressive Moving Average with exogenous variable ARMAX model is the most widely used technique over the past few years in modeling and forecasting this type of data. The most important assumptions of this model are linearity and homogenous for random error variance of the appropriate model. In practice, these two assumptions are often violated, so the Generalized Autoregressive Conditional Heteroscedasticity (ARCH) and (GARCH) with exogenous variable (GARCHX) are applied to analyze and capture the volatility that occurs in the conditional variance of a linear model. Since time series observations rarely have linear or nonlinear components in nature or usually included together. Therefore, the main purpose of this paper is to employ the hybrid model technique according to Zhang methodology for hybrid models to combine the linear forecasts of the best linear model of ARMAX models and the nonlinear forecasts of the best nonlinear models of (ARCH, GARCH & GARCHX) models and thus increase the efficiency and accuracy of performance forecasting future values of the time series.

This paper is concerned with the modeling and building of the hybrid models (ARMAX-GARCH) and (ARMAX-GARCHX), assuming three different random error distributions: Gaussian distribution, Student-t distribution, as well as the general error distribution and the last two distributions were applied for the purpose of capturing the characteristics of heavy tail distributions which have a Leptokurtic characteristic compared to the normal distribution. This research adopted a modern methodology in estimating the parameters of the hybrid model namely the (two-step procedure) methodology. In the first stage, the parameters of the linear model were estimated using three different methods: The Ordinary Least Squares method (OLS), the Recursive Least Square Method with Exponential Forgetting Factor (RLS-EF), and the Recursive Prediction Error Method (RPEM). In the second stage, the parameters of the nonlinear model were estimated using the MLE method and employing the numerical improvement algorithm (BHHH algorithm).

The hybrid models have been applied for modeling the relationship between the exogenous time series (x_t) represented by the exchange rate and the endogenous time series (y_t) represented by the unemployment rate in the USA for the period from (January 2000 to December 2017 i.e. 216 observations), and also the out-of-sample forecasts of unemployment rate in the last twelve values of (2018). The forecasting performance of the hybrid models and the competing individual model was also evaluated using the loss function accuracy measures (MAPE), (MAE), and the robust (Q-LIKE). Based on statistical measurements, the results showed the hybrid models improved the accuracy and efficiency of the single model. ($\mathbf{ARMAX\ (2,1,1,0) - GARCHX(1,1,1)}$) hybrid model error whose conditional variance follow a GED distribution is the optimal model in modeling the bivariate time series data under study and more efficient in the forecasting process compared with the individual $\mathbf{ARMAX\ (2,1,1,0)}$ model and the hybrid $\mathbf{ARMAX\ (2,1,1,0) - GARCH\ (1,1)}$ model. This is due to having the lowest values for accuracy measures. Different software packages (MATLAB (2018a), SAS 9.1, R 3.5.2 and EViews 9) were used to analyze the data under consideration.

Keywords: Bivariate time series ARMAX-GARCH, ARMAX-GARCHX, Normal distribution, Student-t distribution, General Error distribution (GED), Unemployment rate, Exchange rate.