



www.mecsaj.com/ar

المجلة الإلكترونية الشاملة متعددة المعرفة لنشر الأبحاث العلمية والتربوية (MECSJ)

العدد التاسع (كانون الثاني) ٢٠١٩

بسم الله الرحمن الرحيم

التنبؤ بالسلاسل الزمنية باستخدام طريقة الجار الاقرب المضرب

أ.د. عبید محمود محسن الزوبعي / جامعة جيهان-السليمانية-العراق-كوردستان/ بريد اليكتروني

obaed79@yahoo.com

د.أمل السر /جامعة السودان للعلوم والتكنولوجيا -السودان/ قسم الاحصاء / بريد اليكتروني

amalalsir3@gmail.com

أ/ صفية عبدالله الفاضل خالد:كلية العلوم الرياضية والحاسوب /قسم الإحصاء /جامعة الجزيرة -السودان/ بريد الكتروني

safia318@yahoo.com

Abstract ملخص

ركزت هذه الورقة على دراسة طريقة الجار الأقرب المضرب fuzzy nearest neighbour للتنبؤ بالسلاسل الزمنية باستخدام نموذج من نوع "Abbasov-Mamedova" الذي يعمل مع السلاسل الزمنية المضطربة fuzzy time series وذلك من خلال تطبيقه على البيانات الشهرية لأسعار الغاز الطبيعي في الولايات المتحدة الأمريكية في الفترة من (1993- 2014) وتم الحصول على قيم التنبؤ لخمسة فترات في المستقبل كما تم مقارنة نتائج الطريقة مع نتائج طريقة التحليل في مجال الزمن time domain باستخدام نماذج اريما ARIMA وكذلك مع طريقة التحليل في مجال التردد frequency domain باستخدام نموذج سلسلة فوريير Fourier series model وقد أظهرت المقارنة أن نموذج طريقة الجار الأقرب المضرب ذو كفاءة عالية بناء على معياري MAE RMSE وكان وضحا انه أكثر سرعة في الحصول على النتائج مقارنة بالطرق الكلاسيكية إضافة الي سهولة التطوير مما يدل على ان الطريقة واعدة ومشجعة مستقبلا في مجال التنبؤ بالسلاسل الزمنية .

الكلمات المفتاحية :

الإستدلال التقريبي *approximate reasoning*، المنطق المضرب fuzzy logic، السلاسل الزمنية المضطربة fuzzy time series، الانظمة الخبيرة المضطربة Fuzzy expert systems .

Time series Forecasting Using Fuzzy nearest neighbour Method (FNNM)

Abstract

This paper applying Fuzzy Nearest Neighbor method (FNNM) for time series forecasting using a model of "Abbasov-Mamedova" type that works with fuzzy time series on the monthly data for natural gas price in United States from 1993 to 2014 we Obtain five periods forecasting values, the results were also compared with the results of time domain analysis method Using the ARIMA models as well as frequency domain analysis using the Fourier series model. The comparison showed the advantages of faster and highly automated model synthesis as well as improved prediction and forecasting accuracies for fuzzy nearest Neighbor method and The overall performance statistics shows that the proposed model is capable to produce good results indicating that The method is promising and encouraging for time series forecasting .

KEY WORDS: *approximate reasoning*, fuzzy logic, fuzzy time series, Fuzzy expert systems

1-1 المقدمة [2][1][4]

نظرية المجموعات المضيبة والتنبؤ THE FUZZY SETS THEORY AND FORECASTING

بنيت طريقة الجار الأقرب K-nearest neighbour (Knn) للتنبؤ بالسلاسل الزمنية على فكرة بسيطة وهي أن الحالات المتشابهة similar states التي تم ملاحظتها في الماضي من المرجح بصورة كبيرة ان تولد قيم مخرجات متشابهة similar output عن طريق النظر الي اقرب الجيران بناء على مقياس مسافة distance measure معين. الإصدار المضيبة fuzzy version من هذه الطريقة (FNN) تعمل على وزن المساهمة لمختلف الجيران في عملية التنبؤ باستخدام عضوية مضيبة من المصطلحات اللغوية linguistic terms كدالة مسافة بالنسبة للقراءة الحالية. تلعب نظرية المجموعات المضيبة والمنطق المضيب الدور الاكبر في تطبيق هذه الطريقة، حيث تعمل مرحلة التدريب على ترجمة كل السيناريوهات التي تم ملاحظتها في السلسلة الزمنية في الماضي على صورة قواعد مضيبة (تمثل النموذج) وعند الحاجة الي التنبؤ بقيمة جديدة في المستقبل يتم تضبيب المشاهدة الحالية (اخر مشاهدة) ويتم مسح شامل للقواعد واختيار القاعدة التي سيتم تفعيلها (تلك التي استوفت الشروط) ومن ثم يتم فك التضبيب للمخرج باستخدام قاعدة مركز الجذب العام . common center of gravity rule

يستطيع المحلل تعديل بعض الخيارات في طريقة الجار الأقرب مثل تغيير دالة المسافة من المسافة الاقليدية الي أي دالة مسافة اخرى كما يمكنه تعديل الاجراء المستخدم لحساب دالة العضوية كما يمكن بسهولة تعديل عدد الجيران [1]

تعتبر نظرية المجموعات المضيبة fuzzy sets theory صياغة رياضية تعمل على اتاحة الإمكانية لإزالة eliminate عدم التأكد indefiniteness والتعامل مع المعلومات غير الكاملة incomplete والمعلومات غير الدقيقة inaccurate. فالمجموعة المضيبة التي تم تطويرها بواسطة العالم لطفي زادة L.Zadeh تعتبر واحدة من التمثيلات المعروفة للرياضيات التطبيقية الحديثة كما تعتبر تمديد لمفهوم المجموعة التقليدية والتي تعرف حسب :

إذا كان $X = \{x\}$ مجموعة مكونة من عدد من العناصر وكانت $A \subset X$ فانه وعن طريق اتباع الاسلوب التقليدي يتم تعريف المجموعة A بواسطة الدالة التشخيصية $\mu_A(x)$ والتي تأخذ قيمتين فقط كما في التعريف

$$\mu_A(x) = \{if\ x \in A\ then\ \mu_A(x) = 1\ and\ if\ x \notin A\ then\ \mu_A(x) = 0\} \quad (1)$$

بالنسبة للمجموعة المضيبة A التي يتم تعريفها من خلال دالة العضوية

$$\mu_A: X \rightarrow [0, 1] \quad (2)$$

بمعنى انه كلما كانت قيمة الدالة قريبة من 1 كلما ارتفعت درجة عضوية العنصر $x \in X$ بالنسبة للمجموعة A ورياضيا المجموعة المضيبة A تعرف كالاتي

$$A = \{(x, \mu_A(x)/x \in X\}$$

فإذا كانت $X = \{x_1, x_2, \dots, x_n\}$ مجموعة منتهية فإن المجموعة المضيبة تأخذ الشكل التالي

$$A = \{\mu_1/x_1 + \mu_2/x_2 + \dots + \mu_n/x_n\} \quad (3)$$

حيث ان μ_i/x_i عبارة عن العنصر x_i الذي ينتمي الي المجموعة المضيبة A بدرجة عضوية مقدارها μ_i والاشارة (+) تعبر عن عملية الإتحاد Union. تبرز قدرة المنطق المضيب المبني على نظرية المجموعات المضيبة على نمذجة العمليات الغير خطية nonlinear والمتعددة الأبعاد multidimensional حيث ان البنية التشخيصية للنموذج فيه تتطلب العناصر التالية: نوع type النموذج ويقصد به عدد القواعد حيث أن أساس القاعدة rule base يجب أن يستوفي كافة شروط الكفاءة من إتساق consistencies وتمام completion وإتصال continuity، العنصر الثاني هو موقع positioning المجموعات المضيبة في مجال كل متغير.

العنصر الثالث هو طرق الحوار المناسبة لتحويل القيم العادية crisp الي مضببة وتحويل القيم المضببة الي قيم عادية والعنصر الرابع يمثل مؤثر الاستدلال inference operator.

يتم تعريف بنية النموذج Model structure عن طريق عدد من الإعدادات arrangements لدوال العضوية المرتبطة بالنموذج المضبب وأبسط بنية للنموذج هي التي تستخدم دالة عضوية واحدة لكل متغير ادخال input variable ويتم حساب معامل الارتباط لمجموعة الاختبار ومن ثم يزداد تعقيد النموذج بزيادة دوال العضوية والنموذج صاحب أعلى معامل إرتباط يتم إختياره على أنه النموذج الأمثل optimal وبهذا الأسلوب يتم انشاء توليفة بين معلمات parameters النموذج ومستخلصات القواعد ، اما عملية التقدير للمعلمات فيتم على اساس طريقة المربعات الصغرى ومن هذه الناحية فإن عملية تشخيص identification النموذج المضبب قريب من التشخيص التقليدي للنموذج غير الخطي .

1-2 هدف البحث [6]

ركزت هذه الورقة على دراسة طريقة الجار الأقرب المضبب (FNNM) وتطبيقها في التنبؤ بالسلاسل الزمنية بناء على المنهجية والاجراء الذي تم تقديمه بواسطة كل من A.M. ABBASOV, M.H. MAMEDOVA (2013) ، باستخدام دالة fuzzy.ts2 من حزمة AnalyzeTS في برنامج R على بيانات أسعار الغاز الطبيعي في الولايات المتحدة (1993-2014) كما هدفت الورقة الي مقارنة نتائج الطريقة مع نتائج طريقتي التحليل في مجال الزمن ومجال التردد .

1-3 الإطار النظري

1-3-1 أساسيات التضبب fuzzy principles [2],[8]

المؤثر المنطقي و " AND " والمؤثر المنطقي أو " OR " في المنطق المضبب عبارة عن تمديد أو تعميم generalization لرصيفاتها في المنطق التقليدي classical logic ويطلق عليها تنظيم-ت (t-norms) ونظيم مرافق-ت (t-conorms) على التوالي وهي عبارة عن دوال مجالها $[0,1] \times [0,1]$ ، ومجالها المصاحب $[0,1]$. اذا كان $z = T(x, y)$ فإن x, y, z جميعا تقع داخل الفترة $[0,1]$ وكل من التنظيم-ت والنظيم المرافق-ت لها خواص التبادلية commutativity ، التوافقية monotonicity ، المحدودية boundary و الترابطية associativity.

لنضع في الاعتبار الفضاء الشامل أو المجموعة الشاملة universe of discourse والذي يرمز له بالرمز U يتم تقسيم هذا الفضاء الي عدد n من الفترات intervals كالاتي $U = \{u_1, u_2, \dots, u_n\}$ حيث ان u_i عبارة عن فترة داخل الفضاء الشامل U كما يتم تعريف المجموعة المضببة A_i للفضاء U كالاتي

$$A_i = \mu_{A_i}(u_1)/u_1 + \mu_{A_i}(u_2)/u_2 + \dots + \mu_{A_i}(u_n)/u_n \quad (4)$$

حيث أن μ_{A_i} تمثل دالة العضوية للمجموعة المضببة A_i و u_k عبارة عن عنصر من المجموعة المضببة A_i و $\mu_{A_i}(u_k)$ عبارة عن درجة العضوية degree of membership للعنصر u_k بالنسبة للمجموعة A_i .

لنفترض أن Y_t مجموعة جزئية من R و $(t = 0, 1, 2, \dots)$ حيث R الفضاء الشامل الذي تعرف عليه العضويات المضببة $\mu_i(t)$ و $i = 1, 2, \dots$ ونفرض ان $F(t)$ عبارة عن مجموعة من $\mu_i(t)$ في هذه الحالة يطلق على $F(t)$ سلسلة زمنية مضببة على Y_t أو متغير لغوي linguistic variable ويمكن عرض $\mu_i(t)$ لتمثل القيم اللغوية الممكنة ل $F(t)$ ويتم تقديمها على انها مجموعات مضببة . وكما هو الحال فإن $F(t)$ (السلسلة الزمنية المضببة) عبارة عن متغير مرتبط مع الزمن time dependent واستنادا على ما تم تقديمه بواسطة Song and Chissom اذا كانت $F(t)$ مولدة فقط بواسطة $F(t-1)$ ففي هذه الحالة يمكن تمثيل العضوية عن طريق

$$F(t-1) \rightarrow F(t) \quad (5)$$

والارتباط اعلاه يمكن تمثيله بالاتي

$$F(t) = F(t-1) \chi R(t-1, t) \quad (6)$$

حيث أن $R(t-1, t)$ تمثل العضوية المضببة بين $F(t)$ و $F(t-1)$ والرمز χ يمثل أي مؤثر an operator فيمكن ان يكون مؤثر min او مؤثر $min-max$ أو أي مؤثر رياضي arithmetic operator .

فإذا كان من الممكن تمثيل $F(t-1)$ بواسطة A_{i-1} و يمكن تمثيل $F(t)$ بواسطة A_i اذن في هذه الحالة فإن $F(t-1) \rightarrow F(t)$ يمكن تمثيلها بواسطة $A_i \rightarrow A_{i-1}$ وبذا يمكن انشاء علاقة منطقية مضببة fuzzy logical relationship عن طريق تجميع كل الحدود في الطرف الأيمن right-hand-side للمجموعات المضببة المولدة من نفس المجموعة المضببة الي الطرف الأيسر من العضويات المنطقية المضببة . على سبيل المثال اذا كان هنالك مجموعات منطقية مضببة

$$A_i \rightarrow A_j, A_i \rightarrow A_k, A_i \rightarrow A_l, \dots$$

فإنه يمكن دمجها في مجموعة عضوية منطقية مضببة واحدة كالآتي

$$A_i \rightarrow A_j, \dots, A_k, A_l, \dots$$

تعتبر عملية تحديد طول الفترة التي تقسم السلسلة الزمنية الي عدة مجموعات مضببة من المهام الحساسة و غاية في الأهمية ذلك لان تحديد أطوال مختلفة الطول ربما يؤدي الي نتائج تنبؤات مختلفة فالطول الفعال للفترة يشترط به الا يكون كبير جدا وفي نفس الوقت ليس صغير جدا لان الفترات العريضة جدا تؤدي الي فقدان التقلبات fluctuation الموجودة في السلسلة الزمنية المضببة والفترات الصغيرة جدا تؤدي الي فقر التنقيب في السلسلة الزمنية المضببة . اذن فيجب تحديد طول الفترة المناسب بحيث يعكس أو يعمل على اظهار على الأقل نصف التقلبات في السلسلة الزمنية وبناءا على هذا المفهوم يوجد أسلوبين approaches لتحديد الطول المناسب للفترة كالآتي :

1-تحديد طول الفترة بناءا على المتوسط average-based length

2-تحديد طول الفترة بناءا على التوزيع distribution-based length .

[9] Fuzzy rules and inference والاستدلال

يمكن تعريف القاعدة المضببة fuzzy rule على انها تقرير proposition على صورة اذا فإن IF THEN والذي يحتوي على مؤثر استدلال أساسي وتقارير مضببة ويأخذ الصورة

$$((x \text{ is } A) \text{ AND } (x \text{ is } B) \text{ AND } (y \text{ is } C) \text{ AND } (z \text{ is } D)) \Rightarrow (r \text{ is } Q) \quad (7)$$

حيث ان x, y and z تمثل متغيرات الادخال و r هي متغير الاخراج و A, B and C , تمثل دوال العضوية المعرفة على x, y and z على

التوالي و Q عبارة عن دالة العضوية membership function المعرفة على r .

يطلق على الطرف الايسر الذي يمثل ادعاء القاعدة rule's premise يطلق عليه الشرط antecedent وهو يصف الي أي درجة تطبق

القاعدة بينما يطلق على الطرف الايمن أو الخلاصة conclusion يطلق عليه المترتب consequent وهي تسند دالة عضوية لكل واحد

او اكثر من متغيرات الاخراج.

أكثر الأدوات استخداما للاستدلال المنطقي المضرب ما يسمى بالاستدلال التقريبي *approximate reasoning* والذي يستخدم لاستنتاج تقارير منطقية جديدة من التقارير الموجودة مسبقا وهو يعمل على تطبيق الصيغة العامة ل *modus ponens*. الصيغة الكلاسيكية ل *modus ponens* هي *if A then B* والاصدارة المضبية تصاغ كالآتي

If X is A then Y is B, from X = A' infer that Y = B'

حيث ان *A* و *A'* عبارة عن مجموعات مضبية تنتمي الي نفس الفضاء و *B* and *B'* ايضا مجموعات مضبية على نفس الفضاء والتي ربما تختلف عن *A* و *A'*.

المتغير اللغوي linguistic variable *A* هو متغير يعبر عن قيمه بواسطة عبارات لغوية يتم تمثيلها بواسطة المجموعة المضبية، على سبيل المثال المتغير الذي يشير الي درجة الحرارة على الفترة $[0, T]$ حيث X مجال الاعداد الحقيقية هناك ثلاثة مجموعات مضبية للتعبير عن درجة الحرارة وهي منخفضة *LOW* ومتوسطة *MEDIUM* ومرتفعة *HIGH*, اختصارا *L, M, H* ولكل مجموعة من هذه المجموعات دوال عضوية μ_L, μ_M, μ_H هنا يتم التعامل مع درجة الحرارة كمتغير لغوي وقيمته اللغوية هي العبارات *LOW, MEDIUM, HIGH*

{9} Fuzzy expert systems الانظمة الخبيرة المضبية

يحتوي هذا النوع من الانظمة على تشكيلة *collection* مشفرة *encoding* من المعرفة البشرية تشمل التنبؤ *forecasting* والتصنيف

classification اضافة الي آلية استدلال *inference engine* تستخدم هذه الآلية لتقييم *evaluating* أساس القاعدة بالنسبة لمجموعة من المدخلات [7]

إذن المكونات الأساسية للنظام الخبير هي مجموعة القواعد المضبية و دوال العضوية إضافة الي آلية الاستدلال المضرب وتحتوي هذه الآلية على جدولة *scheduler* تعمل على اختيار القواعد خلال عمليات متسلسلة، وبرزت الحاجة الي هذه الجدولة لانه لابد من فحص *checked* قواعد معينة قبل اخرى بمعنى ان بعض القواعد يتم تفعيلها وجعل ما يترتب عليها *consequents* صحيحا، وفي نفس الوقت تكون هذه المترتبات هي الشروط في قواعد اخرى. لذا فإن تطبيق أمر التفعيل *firing order* للقواعد يتم بواسطة تلك الجدولة. وتحتوي الآلية ايضا على معالج قاعدة *rule processor* ليقوم بفحص الجانب الأيسر من القاعدة ليختبر قيم الصواب للشروط *conditions* ويستخدم قيم الصواب تلك لتفعيل القاعدة *fire* القاعدة ويقوم منشئ عضوية المجموعة *set membership builder* ببناء دالة عضوية تقريبية *approximate* للصيغة الناتجة من تفعيل القواعد ومن ثم يقوم مفكك التضبيب *defuzzifier* بتحويل القيم المضبية الناتجة من المترتبات *consequents* الي قيم غير مضبية.

يتم تطبيق عمليات النظام الخبير وفق الآتي

التضبيب: Fuzzification : بمعنى ايجاد درجات عضوية للقيم اللغوية للمتغير اللغوي بناء على القيمة المدخلة بواسطة المستخدم ويتم حساب درجة الصواب degree of truth لكل شرط antecedent بواسطة تطبيق دوال العضوية لقيم الإدخال .

الاستدلال Inference: يتم في هذه الخطوة تطبيق قواعد اذا فإن The if-then rules ويتم حساب قيمة الصواب لكل قاعدة باستخدام قيمة الصواب للقيم اللغوية .

الانشاء Composition: جميع المجموعات الجزئية التي تم اسنادها بالنسبة لكل متغير اخراج يتم توليفها combined ببعضها البعض لتشكيل form مجموعة جزئية مضببة وحيدة single لجميع المخرجات الناتجة من المتغيرات اللغوية والطرق المستخدمة هي طريقة الأكبر max وطريقة المجموع sum

فك التضبيب Defuzzification: يتم هنا حساب قيمة محددة scalar من الناتج المضبب fuzzy output حيث انه في مرحلة الانشاء حصلنا على مجموعة واحدة مضببة فتقوم خطوة فك التضبيب بتجميع تلك المجموعة في قيمة واحدة غير مضببة وعادة ما يتم استخدام طريقتي المركز centroid وطريقة الأكبر maximum ، يتم في طريقة المركز حساب القيمة غير المضببة لمتغير الاخراج عن طريق ايجاد قيمة متغير مركز الجذب center of gravity لدالة العضوية للقيمة المضببة والصيغة العامة هي الوسط الحسابي كالاتي

$$\bar{X}(\text{centroid}) = \frac{\int_b^a x\mu(x)dx}{\int_b^a \mu(x)dx} \quad (8)$$

حيث ان الفترة [a, b] تمثل الفترة التي تم تجميع دالة العضوية عليها. أما طريقة الاكبر maximum method فيتم اختيار القيمة الأكبر في المجموعة الجزئية المضببة لقيمة الصواب لتمثل القيمة غير المضببة النهائية لمتغير الإدخال .

إنشاء مجموعات علاقات منطقية مضببة [7] creating logically fuzzy relationship groups

تعريف 1

نفرض ان $R(t-1, t)$ علاقة مضببة بحيث

$$F(t) = F(t-1) \circ R(t-1, t)$$

حيث \circ موثر رياضي فيقال ان $F(t)$ نتجت بواسطة $F(t-1)$ ويعبر عن هذه العلاقة بالاتي

$$F(t-1) \rightarrow F(t)$$

تعريف 2

إذا فرضنا ان

$$F(t-1) = \tilde{A}_i$$

$$F(t) = \tilde{A}_j$$

يمكن تعريف علاقة منطقية مضببة fuzzy logical relationship أو اساس قاعدة rule base كالاتي

$$\tilde{A}_i \rightarrow \tilde{A}_j$$

حيث ان \tilde{A}_i و \tilde{A}_j ارقام مضببة تسمى الطرف الايسر والطرف الايمن للعلاقة المضببة منطقيا على التوالي .

فإذا كان لدينا مجموعة من البيانات التاريخية X_t وتم تقسيم المدى الذي تقع ضمنه هذه البيانات الي فترات u_i حيث $i=1,2,\dots,m$ فإنه يكون لدينا m من الارقام المضببة fuzzy numbers يعبر عنها ب $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \dots, \tilde{A}_m$

مثلا اذا كانت الفترات $u_1 = [d_1, d_2]$, $u_2 = [d_2, d_3]$, ..., $u_{m-1} = [d_{m-1}, d_m]$, , $u_m = [d_m, d_{m+1}]$

يمكن تعريف الارقام المضببة كالاتي

$$\tilde{A}_1 = (d_0 d_1, d_2, d_3)$$

$$\tilde{A}_2 = (d_1 d_2, d_3, d_4)$$

$$\tilde{A}_3 = (d_2 d_3, d_4, d_5)$$

...

$$\tilde{A}_{m-1} = (d_{m-2} d_{m-1}, d_m, d_{m+1})$$

$$\tilde{A}_m = (d_{m-1}, d_m, d_{m+1}, d_{m+2})$$

يمكن انشاء مجموعات منطقية مضببة وفق الخطوات التالية

1- يتم تضبيب القيم التاريخية اذا كانت القيمة X_t واقعة ضمن الفترة u_j فهي تتبع الرقم المضبب \tilde{A}_j ولذلك كل القيم X_t لابد من تصنيفها الي الاعداد المضببة التابعة لها .

2- يتم توليد العلاقات المنطقية المضببة لجميع القيم التي تم تضبيبها وذلك يتم بناءا على تعريف [1] فتكون العلاقة المنطقية المضببة على الصورة

$$\tilde{A}_j \rightarrow \tilde{A}_k$$

وهي تعبر عن الاتي: اذا كانت قيمة X_{t-1} عند الزمن $t-1$ تساوي \tilde{A}_j فإن قيمة X_t عند الزمن t تساوي \tilde{A}_k

3- يتم انشاء مجموعات عناصرها علاقات منطقية مضببة fuzzy logical relationship groups عن طريق ترتيب العلاقات

المنطقية المضببة التي تتساوى فيها الارقام المضببة على الطرف الايسر وتكون هذه المجموعات على الصورة

$$\tilde{A}_j \rightarrow \tilde{A}_{k1}$$

$$\tilde{A}_j \rightarrow \tilde{A}_{k2}$$

$$\tilde{A}_j \rightarrow \tilde{A}_{k3}$$

...

$$\tilde{A}_j \rightarrow \tilde{A}_{kp}$$

ويمكن وضعها كالاتي

$$\tilde{A}_j \rightarrow \tilde{A}_{k1}, \tilde{A}_{k2}, \tilde{A}_{k3}, \dots, \tilde{A}_{kp}$$

3-1- آلية التنفيذ [3]

في هذه الورقة تم استخدام طريقة الجار الاقرب المضرب بناء على المنهجية والاجراء الذي تم تقديمه بواسطة كل من *A.M. ABBASOV, M.H. MAMEDOVA* (2013)، على بيانات أسعار الغاز الطبيعي في الولايات المتحدة حيث يستند التحليل على الاتي

أ/ اعتبار ان الديناميكية لحركة المبيعات للفترة قيد الدراسة غير معلومة

ب/ يتم التنبؤ بكل شهر داخل الفترة على اساس التغيرات في حركة الأسعار للشهور السابقة

ج/ يتم مقارنة حركة المبيعات التي تم مشاهدتها على الفترة قيد الدراسة مع نتائج تطبيق النموذج ومن ثم الحصول على خطأ النموذج ذلك باستخدام الصيغة

$$\delta(t) = \frac{V_{obsev}^t - V_{forec}^t}{N_{obsev}^t} \cdot 100\% \quad (9)$$

حيث :

V_{obsev}^t تمثل قيمة التغير المشاهد *observrd* للشهر t

V_{forec}^t تمثل قيمة التغير المتنبأ به *forecated* للشهر t

N_{obsev}^t تمثل قيمة المشاهدة للأسعار للشهر t .

وتشتمل منهجية *A.M. ABBASOV, M.H. MAMEDOVA* (2013) على الخطوات الاساسية التالية :

1/ تعريف مجموعة الفضاء الشامل U وهي المجموعة التي تحتوي الفترة بين التغير الأقل والتغير الأكبر في مجموع المجتمع

2/ يتم تقسيم مجموعة الفضاء الشامل U الي فترات متساوية الطول تحوي قيم التغير في المبيعات

3/ وضع الوصف النوعي بالنسبة لقيم التغير في المبيعات على شكل متغير لغوي *linguistic variable*

، بمعنى تحديد القيم الخاصة بالمتغير اللغوي أو المجموعة المكونة من المجموعات المضببة $F(t)$

4/ إجراء عملية تضبيب *Fuzzifying* . لبيانات الادخال ويقصد بعملية التضبيب تحويل القيم العددية الي قيم مضببة . وهذه العملية تجعل من الممكن التعبير عن القيم العددية التي تمثل التغير في المبيعات على صورة قيم من دالة العضوية . *membership function*.

5/ يتم اختيار معلمة $W > 1$ وهذه المعلمة تخص فترة الزمن المرتبط بالشهر المعين ومن ثم حساب المصفوفة المضببة للعلاقات $P^W(T)$ ومنها حساب التنبؤ بالمبيعات للشهر التالي

6/ فك تضبيب *Defuzzifying* . القيم المتحصل عليها ذلك ويقصد بهذه العملية تحويل القيم المضببة الي قيم عددية، وفيما يلي توضيح مفصل للخطوات اعلاه كالآتي

الخطوة الاولى: يتم ايجاد حركة الأسعار خلال السنوات (1993-2014) وهي عباره عن الفرق بين الشهر السابق والشهر اللاحق بمعنى ان الفرق للشهر الحالي هو حاصل طرح حجم الاسعار للشهر الحالي من حجم الاسعار للشهر السابق . ومن اجل تعريف مجموعة الفضاء الشامل بداية لابد من ايجاد قيمة أعلى فرق وقيمة أدنى فرق داخل فترة (1993-2014) وفيما بعد للتأكد من سلاسة *smoothing* وحدود الفترة يتم اختيار عددين D_1 و D_2 وهما عبارة عن قيم موجبة، وبذا يمكن تعريف المجموعة الشاملة U حسب الصيغة

$$U = [V_{min} - D_1, V_{max} - D_2] \quad (10)$$

حيث ان V_{min} قيمة الفرق الاصغر و V_{max} قيمة الفرق الأكبر

الخطوة الثانية: يتم تقسيم المجموعة الشاملة الي بضعة فترات متساوية وهنا يتم تجزئة المجموعة الي سبعة فترات متساوية $u_1, u_2, u_3, u_4, u_5, u_6, u_7$, و لابد من إيجاد نقاط المنتصف midpoint لتلك الفترات

الخطوة الثالثة: تعريف عدد من المجموعات المضببة على المجموعة الشاملة U ، يعتبر الفرق في مستوى الأسعار متغير لغوي a linguistic variabl القيم اللغوية التالية

A_1 (نمو بمستوى قليل جدا) ، A_2 (نمو بمستوى قليل) ، A_3 (عدم وجود تغير) ، A_4 (نمو بمستوى متوسط) ، A_5 (نمو بمستوى عادي) ، A_6 (نمو مستوى عالي) ، A_7 (نمو بمستوى عالي جدا) .

كل قيمة متغير لغوي هنا تكون تابعة لمتغير مضبب بناء على قاعدة محددة يتم اسنادها الي مجموعة مضببة وتعمل هذه القاعدة على تحديد المعنى لهذا المتغير ،تستخدم الصيغة التالية لتعريف المجموعات المضببة A_1, A_2, \dots, A_7 كالآتي

$$\mu_{A_i}(u_i) = \frac{1}{1+[C.(U-u_m^i)]^2} \quad (11)$$

حيث U تمثل قيمة الفرق في حجم المبيعات ، و u_m^i تمثل نقطة المنتصف للفترة i و C عبارة عن ثابت يتم اختياره بطريقة تضمن عملية تحويل القيم الرقمية الي مضببة و $A_i = (\mu_{A_i}(u_i)/u_i)$ حيث $u_i \in U$ و $\mu_{A_i}(u_i) \in [0, 1]$.

الخطوة الرابعة: تتكون هذه الخطوة من تضبيب الفروقات التي تم حسابها بواسطة الخطوة الاولى

الخطوة الخامسة: في هذه الخطوة لابد من اختيار الأساس *basies* والذي يعبر عنه بالرمز w حيث $(1 < w < l)$ و l هي العدد الكلي للشهور ،استنادا على هذا الاساس يتم حساب ما يسمى بمصفوفة العلاقة fuzzy relationship matrix المضببة $R^w(t)$ والتي سوف تستخدم لايجاد التنبؤات، كما يتم بناء مصفوفة العملية operation matrix من الدرجة $i \times j$ ويرمز لها بالرمز $O^w(t)$ (وهنا i يمثل عدد الصفوف وهي تتطابق مع الشهور $t-2, t-3, \dots, t-w$ و j عدد الاعمدة التي تتطابق مع عدد فترات التغير).

يتم ايضا انشاء مصفوفة المعيار criteria matrix ويرمز لها بالرمز $K(t)$ من الدرجة $1 \times j$ عبارة عن مصفوفة من صف واحد تابعة للفرق المضبب بالنسبة للشهر $t-1$. على سبيل المثال اذا اخذنا الاساس $w=7$ يمكن تعريف مصفوفة العملية $O^7(t)$ من الدرجة 6×7 والتي تعبر عن الفروقات المضببة لمبيعات الشهور $(t-2, t-3, t-4, t-5, t-6, t-7)$ كما يمكن تعريف مصفوفة المعيار $K(T)$ من الرتبة 1×7 (مصفوفة الفروق المضببة للشهر $t-1$)، ولهذا بالنسبة للاساس $w=7$ يتم استخدام بيانات الشهور الثمانية الماضية (لابد من توفرها لايجاد الفرق للشهر $(t-7)$) و يتم حساب مصفوفة العلاقة $R(t)$ كما في الصيغة

$$R(t)_{[i,j]} = O^w(t)_{[i,j]} \cap K(T)_{[1,j]} \quad (12)$$

أو بصورة اخرى

$$R(t) = O^w(t) \otimes K(T) = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1j} \\ R_{21} & R_{22} & \dots & R_{2j} \\ \dots & \dots & \dots & \dots \\ R_{i1} & R_{i2} & \dots & R_{ij} \end{bmatrix} \quad (13)$$

حيث $O^w(t)$ مصفوفة العملية operation matrix و $K(T)$ مصفوفة المجموعات المضببة و \otimes عبارة عن عملية $\min(\cap)$. وفيما بعد يتم تعريف قيمة التنبؤ $F(t)$ للشهر t على الصيغة المضببة التالية

$$F(t) = [\text{Max}(R_{11}, R_{21}, \dots, R_{i1}) \text{Max}(R_{12}, R_{22}, \dots, R_{i2}) \dots \text{Max}(R_{1j}, R_{2j}, \dots, R_{ij})] \quad (17)$$

وفي هذه الحالة $1 \leq n \leq 6, 1 \leq \tau \leq 7$

الخطوة السادسة : في هذه الخطوة يتم فك تضبيب defuzzify نتائج الخطوة الخامسة باستخدام الصيغة التالية formula

$$V(t) = \frac{\sum_{i=1}^7 \mu_t(u_i) \cdot u_m^i}{\sum_{i=1}^7 \mu_t(u_i)} \quad (14)$$

حيث $\mu_t(u_i)$ القيمة المحسوبة لدالة العضوية للشهر t و u_m^i تمثل نقطة المنتصف للفترة u_m

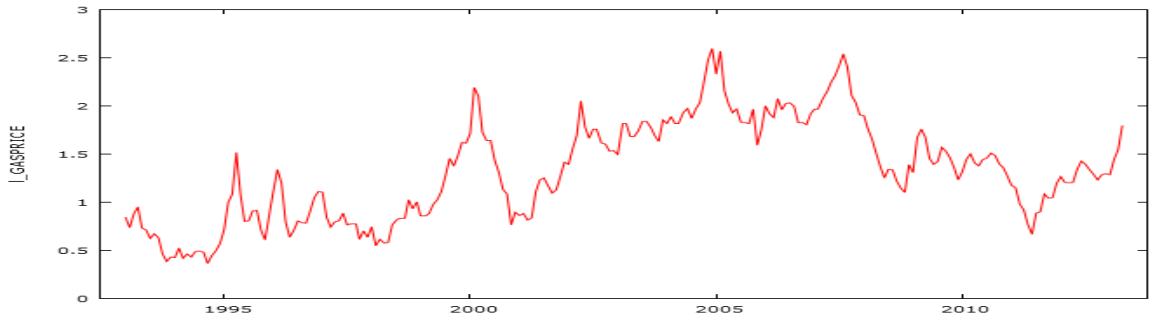
4-1 الجانب التطبيقي [5]

تم في الجزء التالي من هذه الورقة التطبيق على بيانات اسعار الغاز الطبيعي في الولايات المتحدة الامريكية وهي بيانات شهرية اخذت في الفترة من (1993-2014) ونتائج تحليل طريقة الجار الاقرب تم الحصول عليها عن طريق تطبيق دالة $fuzzy.ts2$ من حزمة $AnalyzeTS$ [6] بواسطة برنامج R . وقد تم ايضا استخدام برنامج R للحصول على نموذج $ARIMA$ ونموذج سلسلة فوريير حيث تم استخدام البيانات الشهرية لبناء تلك النماذج.

تظهر في الشكل (١) السلسلة الزمنية بعد اجراء التحويلة اللوغرتمية لها من اجل الحصول على سکون البيانات يمثل رسم بياني للمجموعات المضطربة وهي سبعة مجموعات

بعد ملاحظة عدم سکون السلسلة الزمنية تم اجراء التحويلة اللوغرتمية للحصول على السكون كما يظهر في الشكل (1)

شكل (١) رسم للسلسلة الزمنية $L-GASPRICE$ بعد اجراء التحويلة اللوغرتمية على البيانات الاصلية (من اعداد الباحث بواسطة برنامج $GRET$)



وباستخدام طريقة تحليل السلسلة الزمنية في مجال الزمن $time domain$ تم تطبيق نموذج $ARIMA(1,0,1)$ وذلك للحصول على تنبؤات بأسعار الغاز لستة فترات مستقبلية وكان النموذج المقدر على الصورة

$$X_t = 4.2793 + 0.9026X_{t-1} - 0.0906 \epsilon_{t-1} \quad (15)$$

كما تم ايضا تطبيق تحليل السلسلة الزمنية في مجال التردد $Frequency domain$ وذلك من اجل الحصول على تنبؤات بأسعار الغاز لستة فترات وكان النموذج المقدر لتحليل سلسلة فوريير الذي تم تطبيقه على السلسلة على الصيغة

$$\log X_t = 3.3718 + 0.0073t + 0.094 \cos \omega t - 0.059 \sin \omega t \quad (16)$$

وكان نموذج طريقة الجار الاقرب على صورة $Abbasov-Manedova$ model of 7 fuzzy set with $w=7$ and $c=0$

وفيما يلي اجزاء من نتائج التحليل لبيانات السلسلة الزمنية لبيانات أسعار الغاز الطبيعي في الولايات المتحدة في الفترة (1993-2014) باستخدام طريقة الجار الاقرب المضرب وكما يتضح من النتائج ان النموذج الذي تم تطبيقه نموذج من نوع



type

[1] "Abbasov-Manedova model"

حيث يوضح جدول (table1) التالي والمكون من اربعة اعمدة حيث تمثل *set* اسم المجموعات المضطربة وتمثل *low* و *up* الحدود العليا والدنيا للمجموعات المضطربة كما تعبر *mid* عن قيمة المنتصف *midpoint* التابعة لكل مجموعة مضطربة . (من اعداد الباحث باستخدام برنامج R)

table1

	set	dow	up	mid
1	u1	-2.1323639	-1.5380051	-1.83518450
2	u2	-1.5380051	-0.9436462	-1.24082562
3	u3	-0.9436462	-0.3492873	-0.64646673
4	u4	-0.3492873	0.2450716	-0.05210785
5	u5	0.2450716	0.8394305	0.54225104
6	u6	0.8394305	1.4337894	1.13660992
7	u7	1.4337894	2.0281482	1.73096881

جدول (table2) عبارة عن مشاهدات السلسلة الذي يتكون من ثلاثة اعمدة *point* هي زمن المشاهدة والعمود الثاني *ts* يمثل السلسلة الاصلية والعمود *diff.ts* يمثل سلسلة الفروق التي اخذت من السلسلة الاصلية (من اعداد الباحث باستخدام برنامج R)

table2

	point	ts	diff.ts
1	1993 Nov	3.8712010	NA
2	1993 Dec	3.5553481	-0.315852949
3	1994 Jan	3.9512437	0.395895657
4	1994 Feb	4.1271344	0.175890666
5	1994 Mar	3.5263605	-0.600773860
6	1994 Apr	3.4965076	-0.029852963
7	1994 May	2.8903718	-0.606135804



www.mecsaj.com/ar

- 8 1994 Jun 3.0910425 0.200670695
- 9 1994 Jul 2.9444390 -0.146603474
- 10 1994 Aug 1.9459101 -0.998528830
- 11 1994 Sep 0.6931472 -1.252762968
- 12 1994 Oct 1.6094379 0.916290732
- 13 1994 Nov 1.3862944 -0.223143551
- 14 1994 Dec 2.3978953 1.011600912
- 15 1995 Jan 1.0986123 -1.299282984
- 16 1995 Feb 1.9459101 0.847297860

جدول (table3) يمثل الفروق المضطربة للسلسلة الأصلية (من اعداد الباحث باستخدام برنامج R)

table3

[1] NA

[2] "A[1993 Dec]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[3] "A[1994 Jan]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[4] "A[1994 Feb]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[5] "A[1994 Mar]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[6] "A[1994 Apr]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[7] "A[1994 May]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[8] "A[1994 Jun]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[9] "A[1994 Jul]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[10] "A[1994 Aug]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[11] "A[1994 Sep]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[12] "A[1994 Oct]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

جدول (table4) التالي هو استقراء السلسلة ويتكون من ثلاثة اعمدة *point* تمثل زمن الاستقراء و *inter-polate* يعبر عن استقراء السلسلة و *diff-interpolate* عبارة عن فروق السلسلة المحسوب من عمود استقراء السلسلة , (من اعداد الباحث باستخدام برنامج R)

table4



www.mecsaj.com/ar

point interpolate diff.interpolate

1	1994 Jul	3.03893461	-0.05210785
2	1994 Aug	2.89233113	-0.05210785
3	1994 Sep	1.89380230	-0.05210785
4	1994 Oct	0.64103933	-0.05210785
5	1994 Nov	1.55733006	-0.05210785
6	1994 Dec	1.33418651	-0.05210785
7	1995 Jan	2.34578742	-0.05210785
8	1995 Feb	1.04650444	-0.05210785
9	1995 Mar	1.89380230	-0.05210785
10	1995 Apr	1.55733006	-0.05210785
11	1995 May	2.14511673	-0.05210785
12	1995 Jun	2.25047725	-0.05210785

جدول (table5) يمثل سلسلة التنبؤ وتتكون من ثلاثة اعمدة *point* وتمثل زمن التنبؤ و *diff.forecast* هو الفروق المتحصل عليها من سلسلة التنبؤ (من اعداد الباحث باستخدام برنامج R)

table5

	point forecast	diff.forecast
1	2014 Mar 4.984845	-0.05210785
2	2014 Apr 4.932737	-0.05210785
3	2014 May 4.880629	-0.05210785
4	2014 Jun 4.828521	-0.05210785
5	2014 Jul 4.776413	-0.05210785

جدول (6) تضبيب الفروق للسلسلة الزمنية المتنبأ بها (من اعداد الباحث باستخدام برنامج R)

table6

- [1] "A[2014 Mar]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"
- [2] "A[2014 Apr]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"
- [3] "A[2014 May]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

www.mecsaj.com/ar

[4] "A[2014 Jun]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

[5] "A[2014 Jul]={(1/u1),(1/u2),(1/u3),(1/u4),(1/u5),(1/u6),(1/u7)}"

واخيرا في نهاية النتائج نجد الدقة *accuracy* والتي تمثل مقاييس كفاءة النموذج التي تم حسابها بواسطة الاجراء

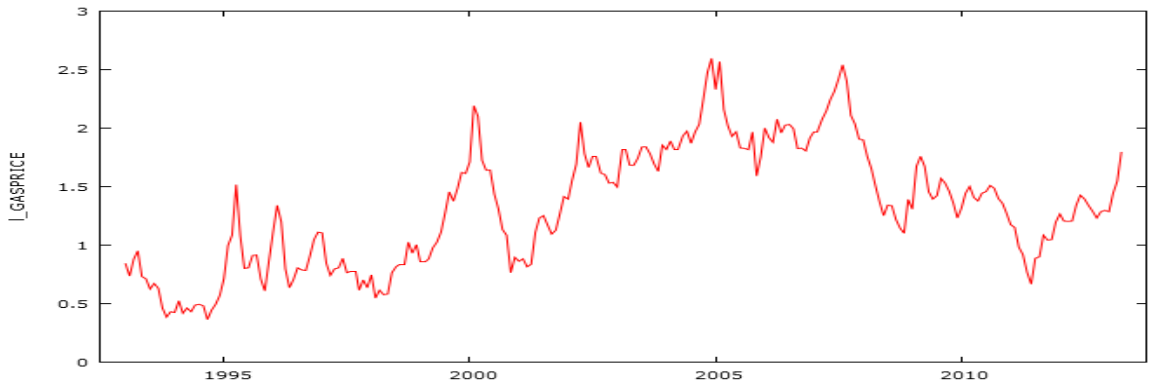
(من اعداد الباحث باستخدام برنامج R)

accuracy

ME MAE MPE MAPE MSE RMSE U

[1,] 0.06 0.248 -Inf Inf 0.202 0.449 1.008949

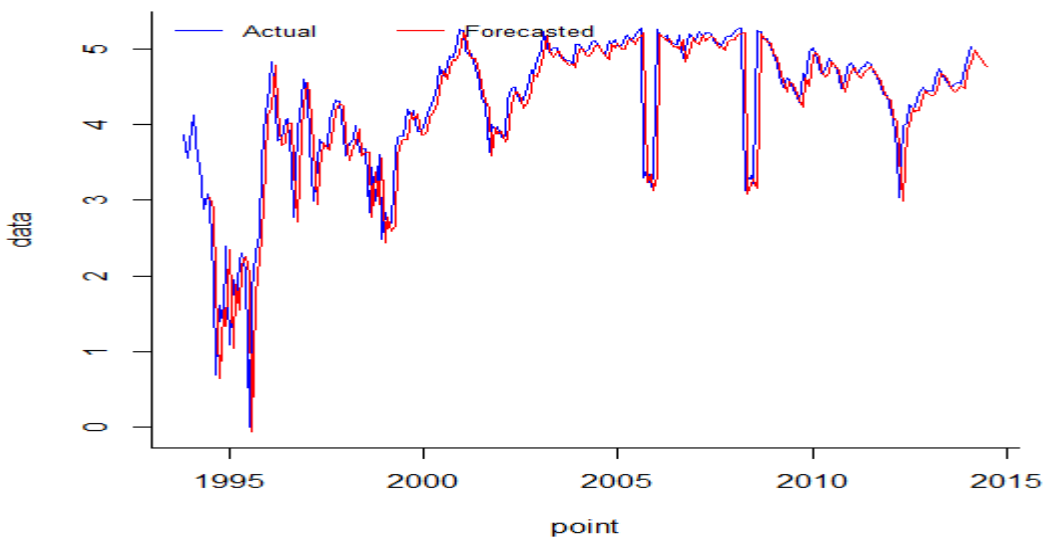
شكل (1) رسم للسلسلة الزمنية L-GASPRICE بعد اجراء التحويلة اللوغرتمية على البيانات الاصلية



(من اعداد الباحث باستخدام برنامج R)

شكل [2] يوضح القيم الفعلية والمنتبأ بها لبيانات أسعار الغاز الطبيعي للسلسلة بواسطة نموذج من نوع Abbasov-Mamedova

Actual series vs forecated series by Abbasov-Mamedova model of 7 fuzzy set with $w = 7$ and $C = 0$



من اعداد الباحث باستخدام برنامج R

5-1 النتائج والتوصيات conclusion

بعد تطبيق منهجية Abbasov-Mamedova ضمن تحليل السلاسل الزمنية المضطربة تم الحصول على النموذج وهو عبارة عن القواعد المضطربة التي من خلال تشغيلها تم الحصول على تنبؤات 5 فترات لبيانات مبيعات الغاز الطبيعي وقد اظهرت الطريقة كفاءة عالية مقارنة بطريقتي تحليل السلاسل الزمنية في مجال الزمن وفي مجال التردد حسب جدول (1) التالي الذي يوضح معايير الدقة لكل طريقة.

ونخلص أخيرا الى ان طريقة الجار الأقرب المضطرب تمتاز بسرعتها العالية في ايجاد النموذج مقارنة بالطرق الكلاسيكية إضافة الى سهولة تطوير النموذج بإدخال قيمة جديدة في السلسلة الزمنية ومن ثم تطبيق الدالة للحصول على تنبؤات لأي فترات في المستقبل مما يدل على ان الطريقة واعدة ومشجعة مستقبلا في مجال التنبؤ بالسلاسل الزمنية.

جدول (1) ملخص لمعايير الاداء لطريقة الجار الأقرب مقارنة بطريقتي مجال الزمن ومجال التردد

معايير الدقة	RMSE	MAE	MASE
طريقة مجال التردد	0.7652809	0.5989277	0.8774864
طريقة مجال الزمن	0.4293453	0.2432592	0.9996888
طريقة الجار الأقرب المضطرب	0.449	0.248	-

فهرسة Bibliography

1/ Burtschy .B,Albeanu.G,Popentiu.Fl –Time series forecasting in software

reliability by fuzzy nearest 16neighbor method-Available at

anale-mi.spiruharet.ro/upload/full_2009_s_a3.pdf

2/ Debasish Sena* and Naresh Kumar Nagwani- A time-series forecasting-based prediction model to estimate groundwater levels in India available at

<http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=00113891&AN=118356587&h=d43Am4S4UrdeY%2BNqD3kZsVsqDcIHfdaq1J3b%2BydQhk%2BVCWqS7len%2Bnc%2BGFHLtDa0cxr%2FkZPnreuCWO3onH0ZeQ%3D%3D&crl=c>

3/A.M. ABBASOV, M.H. MAMEDOVA,(2003)- Application of fuzzy time series to population forecasting

available at:

<http://papers.cumincad.org/data/works/att/50b1.content.pdf>

4-Flores, Juan J., José Ortiz, José R.Cedeñ. González, Carlos Lara, and Rodrigo López Farias. 2016. “FNN a Fuzzy Version of the Nearest Neighbor Time Series Forecasting Technique.” 2015 IEEE International Autumn Meeting on Power, Electronics and Computing, ROPEC 2015.



www.mecsjs.com/ar

available at

<http://ieeexplore.ieee.org/iel7/7389728/7395064/07395125.pdf>

5/ Tran Thi Ngoc Han, Doan Hai Nghi, Mai Thi Hong Diem, Nguyen Thi Diem My, Hong Viet Minh,

Vo Van Tai, Pham Minh Truc(Package 'AnalyzeTS') November 24, 2016 available at

<https://cran.r-project.org/web/packages/AnalyzeTS/AnalyzeTS.pdf>

6/ Min-Yuan Cheng, Nhat-Duc Hoang " Groutability Estimation of Grouting Processes with Microfine Cements Using an Evolutionary Instance-Based Learning Approach" available at

<https://ascelibrary.org/doi/full/10.1061/%28ASCE%29CP.1943-5487.0000370>

7- Okan Duru,shigeru yoshida "comparative analysis of statistical and fuzzy integrater time series and judgmental forecasting :an empirical study of forecasting dry bulk shipping index "-available at

http://www.ymf.or.jp/wp-content/uploads/58_4.pdf

8- Siler, W; Buckley, J. J., *Fuzzy expert systems and fuzzy reasoning*, John Wiley & Sons, Inc., ISBN 0-471-38859-9,2005.

9- Kantrowitz, M., Horstkotte, E., Joslyn, J.,"FAQ: Fuzzy Logic and Fuzzy Expert Systems 1/1", <http://www.faqs.org/faqs/fuzzy-logic/part1>, March 1997