

استخدام الطرق اللاخطية (MLP, ELM, SVM) في التنبؤ بالسلاسل الزمنية (دراسة حالة)

Using Nonlinear Methods (MLP, ELM, SVM) for Time Series Forecasting (Case Study)

أ.م.د. مظهر خالد عبد الحميد

جامعة تكريت - العراق

moudher@tu.edu.iq

أ.م.د. شادي اسماعيل التلباني

جامعة الأزهر - غزة

shtelbany@gmail.com

تاريخ استلام البحث 2020/ 9/29 تاريخ قبول النشر 2020/ 12 /31 تاريخ النشر 2021/3/24

المستخلص:

هدفت هذه الدراسة إلى استخدام كلاً من الشبكات العصبية الاصطناعية (ANN) و آلة متجه الدعم (SVM) والتي تعتبر نماذج غير خطية، بالإضافة الى استخدام الانحدار الذاتي والمتوسطات المتحركة التكاملية (ARIMA) والذي يمثل النموذج التقليدي، للتنبؤ بمؤشر أسعار المعادن الشهري. تم استخدام بيانات السلسلة الزمنية لمؤشر أسعار المعادن الشهري وذلك خلال الفترة الزمنية من أكتوبر 1990 حتى أكتوبر 2020 وبذلك يكون حجم العينة 361 مشاهدة. توصلت الدراسة في البداية الى تحديد النموذج الأفضل من بين نماذج كل طريقة وكانت النتائج كالتالي ARIMA(1,1,0), MLP 5-5-1, ELM 5-100-1 and SVM with (Cost(C) = 1000, Epsilon(ϵ) = 0.1, RMSE, γ)=100. وفي نهاية الدارسة تمت المقاضلة بين النماذج المقترحة باستخدام معايير دقة التنبؤ ، بالإضافة MAE, MAPE. وأشارت النتائج إلى أن نموذج ELM أفضل من نماذج ARIMA, MLP, SVM ، بالإضافة الى ان نموذج ARIMA أفضل من MLP, SVM .

الكلمات المفتاحية: آلة متجه الدعم، الشبكات العصبية، التنبؤ

Abstract:

This study aimed to use both the Artificial Neural Network (ANN) and the Support Vector Machines (SVM) which considered as a non-linear models, in addition, to use the autoregressive integrated moving average (ARIMA) which represent a traditional model to predict the monthly prices of metal index. The study used time series data for the monthly prices of metal index during the period of October 1990 to October 2020, therefore the sample size is 361 observations.

Initially, the study determined the best model for each method, results were are follow, ARIMA(1,1,0), MLP 5-5-1, ELM 5-100-1 and SVM with (Cost(C) = 1000, Epsilon(ϵ) = 0.1, γ)=100). Finally, the study differentiated between the recommended models by using the predictive accuracy measures such as RMSE, MAE, and MAPE. Results indicated that ELM model was better than ARIMA, MLP, and SVM, in addition, ARIMA model is better than MLP and SVM.

Keywords: Support Vector Machines, Neural Networks, Forecasting

1. مقدمة:

في الآونة الأخيرة ، حظي التنبؤ بالقيم المستقبلية بناءً على بيانات السلاسل الزمنية باهتمام كبير في العديد من مجالات البحث. حيث تم استخدام العديد من التقنيات لمعالجة هذه المشكلة من أجل التنبؤ بالسلوك المستقبلي لظاهرة معينة. تم استخدام المنهج التقليدي المستند إلى نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية (ARIMA) الخاصة بـ Box and Jenkins بشكل شائع لأن هذه النماذج سهلة الفهم والتفسير.

تتضمن الأساليب غير الخطية الشائعة الاستخدام في التنبؤ بالسلاسل الزمنية أسلوب شعاع الدعم الالي (SVM) وكذلك أسلوب الشبكات العصبية الاصطناعية (ANN) Neural Networks. وتعد هذه الأساليب طرقاً بديلة عن الطريقة التقليدية المعروفة بمنهجية Box and Jenkins ، ولكن على الرغم من أن النماذج غير الخطية تعطي نتائج ذات كفاءة عالية الا انه لا يجوز الحكم مسبقاً على أفضليتها في التنبؤ بالسلاسل الزمنية من النماذج التقليدية الا بعد تحليل بيانات السلسلة الزمنية محل الاهتمام. حيث تختلف طبيعة البيانات من ظاهرة لأخرى.

يتزايد استخدام SVM و ANN's بسرعة بسبب قدرتها على تشكيل منظومة غير خطية معقدة للتنبؤ بناءً على بيانات العينة. على وجه الخصوص ، في السنوات الأخيرة ، حظيت تطبيقات SVM و ANN باهتمام كبير في العديد من المجالات بسبب سعة التخزين الهائلة وقدرات التعلم والتنبؤ. تم استخدام هذه التقنيات في التنبؤ بالسلاسل الزمنية في العديد من المجالات مثل الأسواق المالية وأعمال المرافق الكهربائية والتنبؤ بالطقس والبيئة.... الخ.

هناك أنواع مختلفة من نماذج SVM، وتختلف هذه النماذج حسب اختلاف المعلمات (γ, c, ϵ) . وأيضاً هناك العديد من نماذج الشبكات العصبية المختلفة، ومنها الشبكات أمامية التغذية، وأشهرها شبكة بيرسبترون متعددة الطبقات (MLP) Multi-Layer Perceptron، وكذلك الشبكات أمامية التغذية خلفية النقل، وأحداهما شبكة إلمان المتكررة Elman Recurrent Network (ERN or ELM). ما سبق من نماذج هي ما سيتم التركيز عليها في هذا البحث.

يتمحور **الهدف الرئيسي للدراسة** حول اجراء **مفاضلة** بين الطرق اللاخطية المقترحة MLP, ELM, SVM **و الطريقة التقليدية ARIMA** لمعرفة أيهما أكثر دقة وكفاءة في التنبؤ بالسلاسل الزمنية، وللتحقق من ذلك تم التطبيق **والمفاضلة** بالاعتماد على بيانات شهرية للمؤشر العام لأسعار المعادن بالدولار الأمريكي ، ويتكون هذا المؤشر من مؤشرات أسعار النحاس والألمنيوم وخام الحديد والقصدير والنيكل والزنك والرصاص واليورانيوم، للفترة الممتدة من أكتوبر 1990 حتى أكتوبر 2020 بواقع 361 مشاهدة.

2. الشبكات العصبية الاصطناعية (ANN's) Artificial Neural Networks :

في السنوات الأخيرة ، أثارت دراسة الشبكات العصبية الاصطناعية اهتماماً كبيراً في مجالات متنوعة، ويكمن السبب الرئيسي وراء هذا الاهتمام أن الشبكات العصبية الاصطناعية تعتبر أداة مرنة، وتعد من الطرق القوية المستخدمة في التنبؤ بالسلاسل الزمنية. الشبكات العصبية (NNs) هي من فئة النماذج غير الخطية المستوحاة من عمل الخلايا العصبية البيولوجية. ولديها تطبيقات في جميع جوانب العلوم وتعتبر من النماذج التي سيكون لها تأثير كبير على حياتنا في المستقبل (Maren et al.,2014).

تعرف الشبكة العصبية بأنها: " نظام معالجة للمعلومات يستند إلى نماذج رياضية بسيطة لها مميزات أداء معينة بأسلوب يحاكي الشبكات العصبية الحيوية (النظام العصبي)، وذلك عن طريق معالجة ضخمة موزعة على التوازي، ومكونة من وحدات معالجة بسيطة، هذه الوحدات ما هي إلا عناصر حسابية تسمى عصبونات أو عقد أو خلايا لها خاصية عصبونية حيث تقوم بتخزين المعرفة العملية والمعلومات التجريبية لتجعلها متاحة للمستخدم وذلك عن طريق ضبط الأوزان" (Negnevitsky, 2011).

2.1 مكونات الشبكات العصبية Components of Neural Networks:

تتكون الشبكات العصبية الاصطناعية من الخلايا العصبية المستوحاة من الخلايا العصبية البيولوجية، وعادة ما يتم تنظيم الخلايا العصبية الاصطناعية في طبقات بحيث تشتمل كل شبكة على ثلاث طبقات رئيسية هي (Pantazis and Alevizakou, 2013):

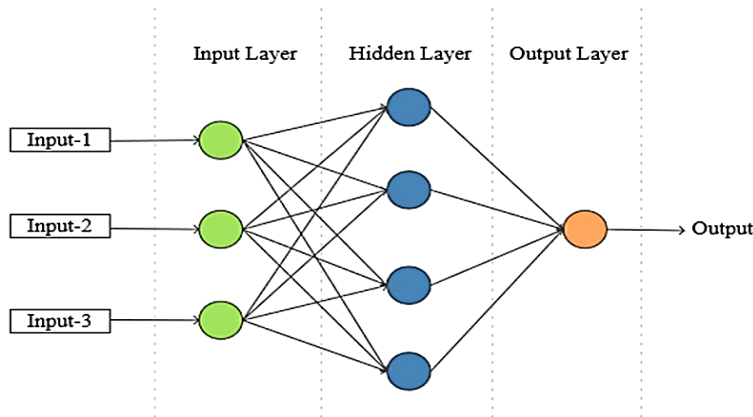
- طبقة الإدخال: تحتوي هذه الطبقة على العديد من الخلايا العصبية كمتغيرات إدخال ANN. ترتبط الخلايا العصبية في طبقة الإدخال بالخلايا العصبية للطبقات المخفية أو بالخلايا العصبية في الطبقة التالية.
- الطبقة المخفية: يمكن أن تحتوي كل طبقة مخفية على N من العصبونات مرتبطة بطرق مختلفة بالطبقات المخفية الأخرى أو مع طبقة الإخراج. يمكن أن تحصل الخلايا العصبية للطبقة المخفية على مدخلاتها من خلال طبقة الإدخال أو طبقة مخفية أخرى أو في بعض الحالات من طبقة الإخراج.
- طبقة الإخراج: يمر من خلالها متجه الإخراج، تحتوي هذه الطبقة على العديد من الخلايا العصبية. يمكن أن تحصل الخلايا العصبية لطبقة الإخراج على مدخلاتها من خلال طبقة الإدخال أو الطبقات المخفية.

2.2 أنواع الشبكات العصبية Types of Neural Networks:

هناك العديد من أنواع الشبكات العصبية المختلفة، وتصنف الشبكات العصبية الاصطناعية إلى نوعين: التصنيف بحسب عدد الطبقات، والتصنيف بحسب اتجاه تدفق المعلومات (Moseley, 2003). لكننا في هذا البحث سنتناول نوعين من هذه الشبكات .

• شبكة الـ بيرسبترون متعددة الطبقات (MLP) Multi-Layer Perceptron:

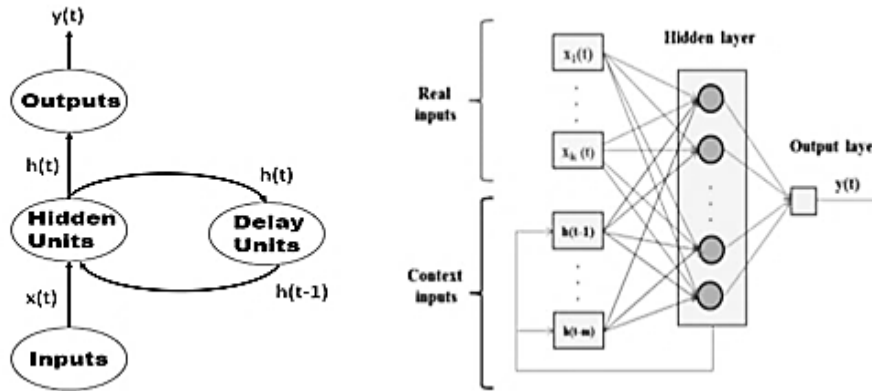
تعتبر شبكة (MLP) ضمن الشبكات أمامية التغذية التي تتحرك فيها المعلومات من طبقة إلى طبقة تالية للأمام فقط، دون السماح لها بالعودة للخلف، وهي من أشهر الشبكات استخداماً في التطبيقات المالية، بالإضافة إلى ذلك فإنّ شبكة (MLP) تعتبر خياراً جيداً إذا ما استخدمت كبديل عن نماذج التنبؤ الخطية (McNelis, 2005). وتسمى MLP أيضاً بـ Deep Learning NNs .



شكل (1): شبكة التغذية الأمامية- MLP

• شبكة إلمان المتكررة (ERN or ELM) : Elman Recurrent Network

شبكة Elman، تعتبر إحدى الشبكات أمامية التغذية خلفية النقل، وهي شبكة عصبية متكررة بسيطة شائعة قدمت أداءً متميزاً في مجموعة واسعة من التطبيقات. وتتحرك فيها المعلومات من طبقة إلى الطبقة التالية للأمام، مع السماح لها بالعودة للخلف إلى الطبقات السابقة. حيث تتكون الشبكات العصبية المتكررة Elman من طبقة إدخال، وطبقة متكررة، طبقة مخفية، وطبقة إخراج. تحتوي كل طبقة على خلية أو أكثر من الخلايا العصبية التي تنتشر المعلومات من طبقة إلى أخرى عن طريق حساب دالة غير خطية لمجموع المدخلات المرجحة. في شبكة Elman العصبية، يكون عدد الخلايا العصبية في طبقة التكرار مساوياً لعدد الخلايا العصبية في الطبقة المخفية. بالإضافة إلى ذلك، ترتبط الخلايا العصبية لطبقة التكرار بشكل كامل بجميع الخلايا العصبية في الطبقة المخفية (Lewis, 2017).



شكل (2): شبكة أمامية التغذية خلفية النقل - ELM

3. آلة المتجه الداعم Support Vector Machines

قدمت هذه التقنية من قبل (Vapnik, 1992) وهي عبارة عن أنظمة تستند إلى نظرية التعلم الاحصائية (statistical learning theory)، وتعد من خوارزميات التعلم عن طريق موجه (supervised)، وتهدف إلى إيجاد أفضل مستوى فاصل (hyperplane) يصنف البيانات بشكل صحيح قدر الامكان (قاسم، 2010). ثم تفصل الفئات بطريقة خطية عن طريق الـ hyperplane الذي ينشئ مصنفاً يقوم بتصنيف البيانات خطأً، وفي بعض الاحيان يوجد بعض المشكلات التي لا يمكن فصلها خطياً في مجال بيانات الإدخال الاصلية لذلك يقوم SVM باستخدام احد أهم المكونات التي تقوم بالتحويل من مجال إلى اخر وتدعى Kernels مختلفة سوف تسمح بعرض اساليب التعلم المختلفة. ومما سبق يتبين ان الـ SVM تحول البيانات إلى فضاء البحث feature space الذي عادة ما يكون ضخماً الأبعاد، ومن الملاحظات المهمة حول الـ SVM انها تعتمد على الخصائص الهندسية لبيانات التدريب وليس على ابعاد البيانات المدخلة (الرشيدلي، 2011). ويعتمد بناء النموذج على عدد من المعلمات مثل المستوى الفاصل (hyperplane) ومضارب لاجرانج (قاسم، 2010).

في السنوات الأخيرة، تم تقديم العديد من النماذج القائمة على آلات التعلم (Learning Machines) لتحليل السلاسل الزمنية. من بينها، الدعم آلات المتجهات (SVM) هي تقنية جديدة مصممة لحل مشاكل التصنيف

والانحدار غير الخطية في تحليل السلاسل الزمنية. تعتمد SVMs على مبدأ تقليل المخاطر الهيكلية مما يسمح بتقدير الدالة عن طريق تصغير الحد الأعلى من خطأ التعميم (Vapnik and Vapnik, 1998). نظراً لقدرتها على تحقيق أداء عالي ودقة اختبار، يمكن أن يكون SVM مفيداً بشكل خاص في التنبؤ بالسلاسل الزمنية (Okasha, 2014). إذا كان لدينا مجموعة تدريب مكونة من N من النقاط $\{X_i, Y_i\}_{i=1}^N$ مع بيانات ادخال $X_i \in \mathbb{R}^P$ ، حيث P هو العدد الإجمالي لأنماط البيانات والمخرجات وهي $Y_i \in \mathbb{R}$ ، بشكل عام، فكرة بناء الدالة التقريبية لـ SVM تتضمن خريطة البيانات x في فضاء مخصص عالي الأبعاد عبر أشكال غير خطية وتنفيذ الانحدار الخطي في الفضاء المخصص (Yang et al., 2004; Ojemakinde, 2006; Wang et al., 2009; Samsudin et al., 2010)، نفترض أن الدالة غير الخطية $y(x)$ على الشكل التالي:

$$y(x) = w^T \varphi(x) + b \quad (1)$$

عندما، w : متجه الأوزان، b : قيمة التحيز، $\varphi(x)$: متجه بيانات الادخال (Zhang et al., 2013). حيث يتم تقدير المعلمتين w و b عن طريق تقليل الدالة التالية: $Min \left(\frac{1}{2} w^T w \right)$ ، وتخضع للقيود التالية:

$$\begin{cases} y_i - w^T \varphi(x_i) - b \leq \varepsilon \\ w^T \varphi(x_i) + b - y_i \leq \varepsilon \end{cases}$$

وهذا يعطينا:

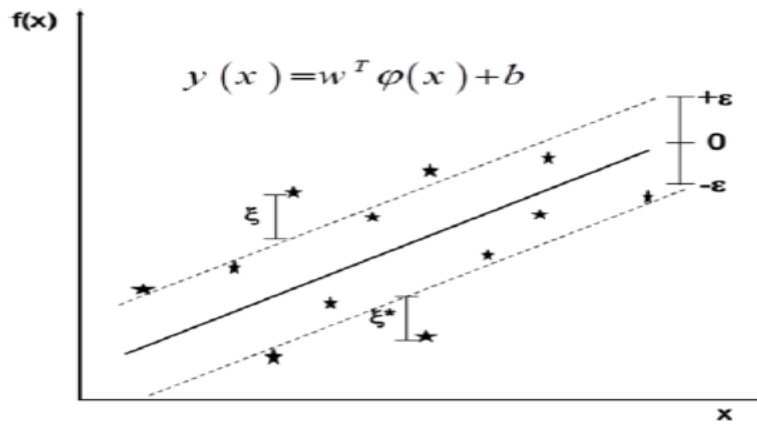
$$(2) \quad y_i (w^T x_i + b) - 1 \geq 0$$

لتقدير w و b ، المعادلة أعلاه تحول الى الدالة الاساسية أدناه من خلال ادخال متغيرات slack الموجبة ξ, ξ^* ، كالتالي (Wang et al., 2009):

$$R(w, \xi, \xi^*) = \frac{1}{2} W^T W + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (3)$$

وتخضع للقيود التالية:

$$\begin{cases} y_i - w^T x_i - b \leq \varepsilon + \xi_i \\ w^T x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$



شكل(3): Illustration of Slack Variables and ε -tube in SVM

الجزء الاول $W^T W$ (1/2) في معادلة (3) هو معيار متجه الأوزان، بينما الجزء الثاني $(C \sum_{i=1}^N (\xi_i + \xi_i^*))$ هو الجزء التجريبي ويقاس ϵ -insensitive loss function ، y_i هي القيمة المطلوبة، ويشار إلى C بالثابت المنتظم، وتحديد المبادلة بين الخطأ التجريبي والجزء المنتظم. ϵ تسمى tube size of SVM وهو مكافئ لدقة التقريب الموضوع على نقاط بيانات التدريب. يتم تقديم متغيرات ξ_i^* ، ξ_i للتعامل مع مشاكل التحسين المحتملة (Ojemakinde, 2006; Samsudin et al, 2010). لحل دالة القرار في معادلة (3) تم ادخال بعض مضاعفات

$$L = \frac{1}{2} W^T W + C \sum_{i=1}^N (\xi_i + \xi_i^*) - \sum_{i=1}^N \alpha_i (\epsilon + \xi_i - y_i + w^T x_i + b) - \sum_{i=1}^N \alpha_i^* (\epsilon + \xi_i^* + y_i - w^T x_i - b) - \sum_{i=1}^N (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad \text{s.t.} \quad \alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$$

(4)

$$\begin{aligned} \sum_{i=1}^N (\alpha_i - \alpha_i^*) &= 0 \\ \sum_{i=1}^N \alpha_i &= \sum_{i=1}^N \alpha_i^* \\ 0 \leq \alpha_i^* &\leq c, \quad i = 1, 2, \dots, N \\ 0 \leq \alpha_i &\leq c, \quad i = 1, 2, \dots, N \end{aligned}$$

يمكن التعبير عن ذلك أيضاً في الشكل:

$$y(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot (\varphi(x_i), \varphi(x)) + b \quad (5)$$

أو بشكل عام على النحو التالي:

$$y(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b \quad (6)$$

$K(x_i, x)$ هي دالة النواة (kernel function) (Mellit et al., 2012 ; Wang et al., 2012)

قيمة Kernal تساوي ناتج حاصل ضرب متجهين X_i and X_j في الفضاء المخصص $\varphi(X_i)$ and $\varphi(X_j)$. في SVM ، أكثر دوال Kernal استخداماً هي (Samsudin et al, 2010) :

- Liner: $K(X_i, X_j) = X_i^T X_j$
- Polynomial: $K(X_i, X_j) = (\gamma X_i^T X_j + r)^d, r, \gamma > 0$
- Gaussian Radial Basis Function(RBF): $K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2), \gamma > 0$
- Perceptron (Sigmoid): $K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r)$

γ, r , and d هي معاملات Kernal . يجب اختيار معلمة kernel بعناية ، لأنها تحدد ضمناً بنية الفضاء المخصص عالي الأبعاد $\varphi(x)$ وبالتالي تتحكم في تعقيد الحل النهائي (Mustaffa and Yusof, 2012) .

4. معايير دقة التنبؤ :

تهدف هذه المعايير لقياس درجة دقة التنبؤ للمقارنة بين النماذج وهي :

- Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100$$

- Mean Absolute Error (MAE):

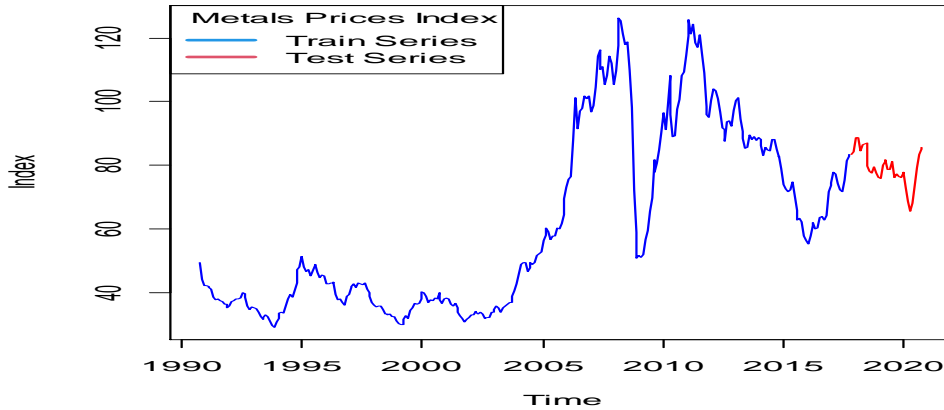
$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

• Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{(y_t - \hat{y}_t)^2}{n}}$$

5. تحليل البيانات ومناقشة النتائج:

السلسلة الزمنية المتاحة لدينا هي بيانات شهرية للمؤشر العام لأسعار المعادن بالدولار الأمريكي ، ويتكون هذا المؤشر من مؤشرات أسعار النحاس والألمنيوم وخام الحديد والقصدير والنيكل والزنك والرصاص واليورانيوم، للفترة الممتدة من شهر أكتوبر 1990 حتى شهر أكتوبر 2020 بواقع 361 مشاهدة، حيث تم تقسيم السلسلة الى جزئين: مجموعة التدريب وتتكون من 325 مشاهدة (90% من السلسلة) خلال الفترة من شهر أكتوبر 1990 حتى شهر سبتمبر 2017 ، و 10% من السلسلة خلال الفترة من أكتوبر 2017 حتى أكتوبر 2020 وتشكل 36 مشاهدة تستخدم كمجموعة اختبار للمفاضلة وتقييم أداء نماذج SVM, MLP, ELM . حيث تم استخدام **برمجية R** لتحليل بيانات الدراسة.



شكل (4): السلسلة الزمنية الشهرية لمؤشر أسعار المعادن

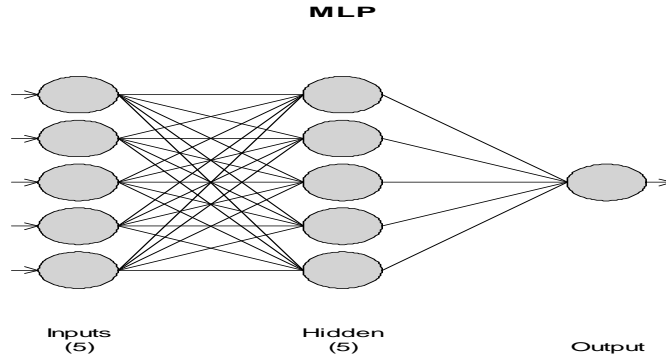
5.1 نتائج تحليل نموذجي MLP, ELM :

سيتم تحليل البيانات الشهرية للمؤشر العام لأسعار المعادن بالدولار الأمريكي باستخدام نوعين من الشبكات العصبية: الشبكات أمامية التغذية وتم استخدام واحدة منها وتعرف باسم MLP ، والشبكة أمامية التغذية خلفية النقل وتم استخدام واحدة منها وتسمى ELM . حيث سيتم الاعتماد في التحليل على نفس طريقة تقسيم البيانات الى مجموعتين: مجموعة التدريب 90% من البيانات و مجموعة الاختبار 10% من البيانات.

من خلال برمجية R والتي عملت على المفاضلة فيما بين نماذج MLP وفيما بين نماذج ELM، للوصول للنموذج الأفضل لكل منهما. حيث كان كل نموذج مكون من ثلاث طبقات: الطبقة الأولى وهي طبقة الإدخال ويقوم

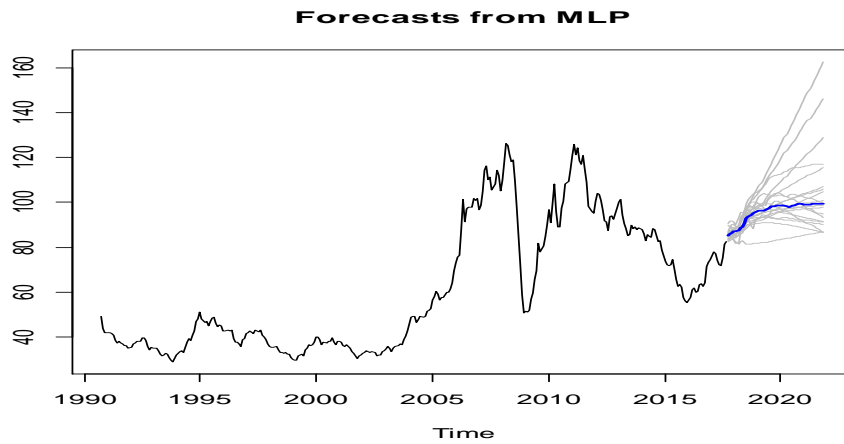
البرنامج بتحديد شكل المدخلات المناسب، الطبقة المخفية والتي تتكون من العديد من الخلايا العصبية هي الطبقة الثانية، أما الطبقة الاخيرة هي المخرجات وبها ناتج واحد.

- مخرجات نموذج MLP : الشكل التالي يوضح النموذج الأفضل الذي تم تحديده من خلال عمليات التكرار وهو نموذج MLP 5-5-1 .



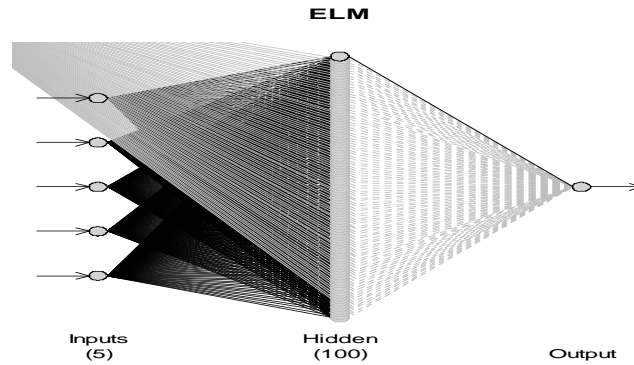
شكل (5): نموذج MLP 5-5-1

من خلال النموذج الأفضل MLP 5-5-1 تم التنبؤ بالقيم المستقبلية (مجموعة الاختبار) ويتضح ذلك من خلال الشكل التالي. وسيتم لاحقاً مفاضلة هذا النموذج مع النماذج الأخرى من حيث دقة التنبؤ.



شكل (6): التنبؤ باستخدام نموذج MLP 5-5-1

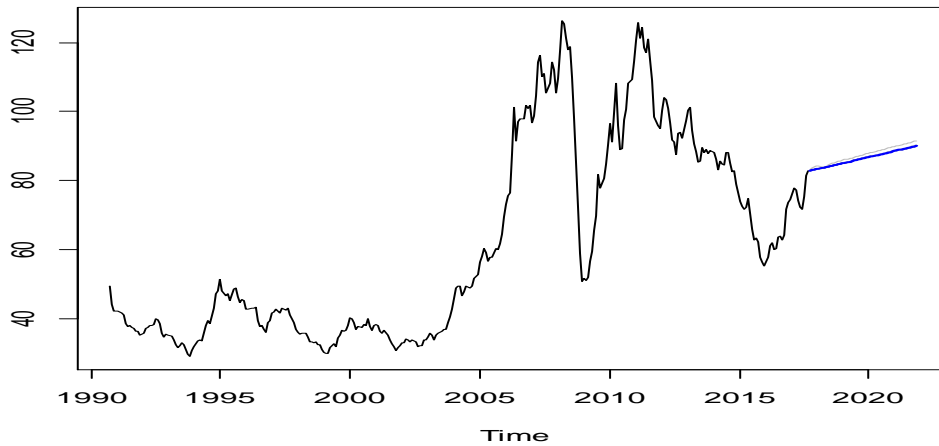
- مخرجات نموذج ELM : من خلال عمليات التكرار تم التوصل الى ان النموذج الأفضل الذي تم تحديده هو نموذج ELM 5-100-1 ويتضح ذلك من الشكل التالي.



شكل (7): التنبؤ باستخدام نموذج ELM 5-100-1

الشكل التالي يوضح التنبؤ بالقيم المستقبلية (مجموعة الاختبار) من خلال النموذج الأفضل ELM 5-100-1 وسيتم **مفاضلة** هذا النموذج مع النماذج الأخرى من حيث دقة التنبؤ لاحقاً.

Forecasts from ELM



شكل (8): التنبؤ باستخدام نموذج ELM 5-100-1

5.2 نتائج تحليل نموذج SVM :

هناك أنواع مختلفة من نماذج SVM، وتختلف هذه النماذج حسب اختلاف المعلمات (γ ، C ، ϵ) التي تم اختيارها بشكل تجريبي لعدد محدود من القيم، **والمفاضلة** بين هذه النماذج حيث يحتوي النموذج الأفضل على أصغر خطأ تدريب وأقل عدد من متجهات الدعم.

قيمة ϵ المستخدمة في التحليل لها قيمة افتراضية ثابتة هي 0.1 (Dimitriadou et al., 2008 ; Garg, 2012). ويقترح كلا من (Meyer, 2018 ; Suganyadevi and Babulal, 2014) أن يكون المدى المناسب لـ C من 1 إلى 1000 ، ومن أجل الحصول على تحليل أكثر شمولاً قمنا بتوسيع الحد الأقصى للمدى ليصبح 10000 . بينما المدى المناسب لـ γ من 0.0001 إلى 100 (Okasha, 2014). الجدول التالي يوضح عدة نماذج SVM مختلفة باختلاف معالماتها، حيث تمت **المفاضلة** بين هذه النماذج من خلال معيارين أساسيين وهما خطأ التدريب وعدد متجهات الدعم. حيث كانت قيم المعلمات المستخدمة كالتالي:

C: 1, 10, 100, 1000, 10000

γ : 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 10, 30, 50, 70, 100

جدول (1): نماذج SVM لمؤشر أسعار المعادن

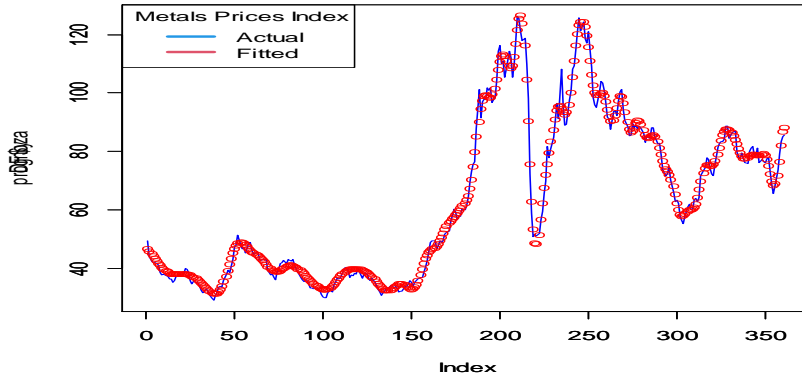
Model	C	γ	Number of Support Vectors	Training Error	Model	C	γ	Number of Support Vectors	Training Error
M1	1	0.0001	300	855.420090	M8	1	0.5	239	173.878496
M2		0.0005	286	688.145774	M9		1	228	148.636349
M3		0.001	287	589.905211	M10		10	159	65.147006
M4		0.005	280	459.904905	M11		30	115	43.666341
M5		0.01	281	438.084253	M12		50	112	36.465828
M6		0.05	276	401.016705	M13		70	102	27.069152
M7		0.1	265	338.236153	M14		100	101	18.575818
M15	10	0.0001	287	589.804309	M22	10	0.5	233	151.646288
M16		0.0005	280	460.201191	M23		1	214	157.776148
M17		0.001	281	440.478947	M24		10	138	46.452316
M18		0.005	282	416.143116	M25		30	112	39.302589
M19		0.01	281	418.968175	M26		50	102	20.260336
M20		0.05	262	334.496193	M27		70	98	14.677885
M21		0.1	244	244.775817	M28		100	97	13.016402
M29	100	0.0001	281	440.491873	M36	100	0.5	215	151.433438
M30		0.0005	282	414.062541	M37		1	212	154.759723
M31		0.001	282	413.620532	M38		10	120	43.297815
M32		0.005	277	427.123881	M39		30	101	27.125486
M33		0.01	277	421.447840	M40		50	99	14.068326
M34		0.05	249	245.836908	M41		70	97	13.331464
M35		0.1	239	212.410599	M42		100	95	11.352560
M43	1000	0.0001	282	412.845481	M50	1000	0.5	212	155.191077
M44		0.0005	284	413.579120	M51		1	203	138.616708
M45		0.001	279	421.247872	M52		10	114	44.185863
M46		0.005	275	422.647653	M53		30	102	18.173453
M47		0.01	264	352.864893	M54		50	100	13.973946
M48		0.05	244	221.600583	M55		70	94	12.144676
M49		0.1	237	204.668194	M56		100	90	9.1954390
M57	10000	0.0001	283	413.115947	M64	10000	0.5	216	151.295728
M58		0.0005	278	430.008691	M65		1	189	131.231292
M59		0.001	277	436.191571	M66		10	114	47.578424
M60		0.005	259	339.833640	M67		30	106	15.568179
M61		0.01	246	256.743872	M68		50	102	13.236034
M62		0.05	241	218.016140	M69		70	95	11.210499
M63		0.1	242	179.231176	M70		100	91	9.961322

من خلال جدول (2)، تم التوصل الى ان نموذج M56 الذي يحتوي على $\gamma=100$, $C=1000$ هو النموذج الأفضل حيث يمتلك أقل خطأ تدريب وأقل عدد من متجهات الدعم. وبالتالي سيتم استخدامه في التنبؤ بالقيم المستقبلية (مجموعة الاختبار). وسيتم المفاضلة مع النماذج الأخرى من حيث دقة التنبؤ.

جدول (2): تحديد نموذج SVM الأفضل

Model	C	γ	N. Support Vectors	Training Error
M14	1	100	101	18.575818
M28	10	100	97	13.016402
M42	100	100	95	11.352560
M56	1000	100	90	9.195439
M70	10000	100	91	9.961322

الشكل البياني التالي يوضح القيم الفعلية مقابل القيم المقدرة لمجموعة التدريب وكذلك لمجموعة الاختبار ومدى تقاربهما من خلال استخدام نموذج SVM الأفضل الذي تم تحديده في الخطوة السابقة.


 شكل (9): القيم الفعلية والقيم المتنبأ بها باستخدام نموذج SVM عندما $\gamma=100$, $C=1000$

6. المفاضلة بين نماذج MLP, ELM, SVM :

في البداية، تم تحليل البيانات باستخدام النموذج التقليدي ARIMA واختيار النموذج $ARIMA(1,1,0)$ من بين عدة نماذج وإضافته مع النماذج المقترحة MLP, ELM, SVM لمعرفة أي من هذه النماذج لديه الأفضل من حيث دقة التنبؤ بالقيم المستقبلية لمؤشر أسعار المعادن، وذلك بالاعتماد على مقاييس دقة التنبؤ MAE, RMSE, MAPE، وكانت النتائج كما هي موضحة في الجدول التالي:

جدول (3): مقاييس دقة التنبؤ للمفاضلة بين النماذج

Model	MAE	RMSE	MAPE(%)
ARIMA (1,1,0)	7.46898	9.07991	9.90
MLP 5-5-1	15.50511	17.7467	20.30
ELM 5-100-1	7.265267	8.77185	9.60
SVM	18.63877	21.0316	23.50

يتضح من الجدول السابق، عند إجراء المفاضلة بين النماذج اللاخطية المقترحة MLP, ELM, SVM والنموذج الخطي $ARIMA(1,1,0)$ ، تبين أن النموذج الأفضل هو النموذج ELM 5-100-1 وهو أكثر دقة من

MAE, RMSE, ARIMA(1,1,0), SVM, MLP 5-5-1 وذلك لامتلاكه أقل القيم لمعايير دقة التنبؤ. وهذا يدعونا للقول أنه لا يجب الحكم مسبقاً على أن النماذج الحديثة تتفوق دائماً على النماذج التقليدية، والسبب في ذلك يعود لطبيعة البيانات محل الاهتمام.

للتأكيد على ما سبق، سنستخدم الكفاءة النسبية (R.E) Relative Efficiency بالاعتماد على معيار RMSE والذي كلما كانت قيمته أقل كلما كان النموذج أكثر دقة وكفاءة، وذلك للنماذج السابقة المبينة في الجدول السابق، ويمكن حسابها بالصيغة التالية: $RE(i, j) = \frac{RMSE(i)}{RMSE(j)}$ التي اقترحها كلا من (White and Safi, 2016) كالتالي:

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of ELM } 5-100-1}{RMSE \text{ of ARIMA } 1,1,0} = \frac{8.77185}{9.07991} = 0.9661$$

قيمة RMSE لنموذج ELM 5-100-1 تشكل 96.6% من قيمة RMSE لنموذج ARIMA(1,1,0) ، أي أن نموذج ELM 5-100-1 أكثر دقة وكفاءة في التنبؤ بـ 1.03 مرة من نموذج ARIMA(1,1,0).

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of ELM } 5-100-1}{RMSE \text{ of SVM}} = \frac{8.77185}{21.0316} = 0.4171$$

قيمة RMSE لنموذج ELM 5-100-1 تشكل 41.7% من قيمة RMSE لنموذج SVM ، أي أن نموذج ELM 5-100-1 أكثر دقة وكفاءة في التنبؤ بـ 2.40 مرة من نموذج SVM .

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of ELM } 5-100-1}{RMSE \text{ of MLP } 5-5-1} = \frac{8.77185}{17.7467} = 0.49432$$

قيمة RMSE لنموذج ELM 5-100-1 تشكل 49.4% من قيمة RMSE لنموذج MLP 5-5-1 ، أي أن نموذج ELM 5-100-1 أكثر دقة وكفاءة في التنبؤ بـ 2.02 مرة من نموذج MLP 5-5-1.

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of MLP } 5-5-1}{RMSE \text{ of SVM}} = \frac{17.7467}{21.0316} = 0.8438$$

قيمة RMSE لنموذج MLP 5-5-1 تشكل 84.4% من قيمة RMSE لنموذج SVM ، أي أن نموذج MLP 5-5-1 أكثر دقة وكفاءة في التنبؤ بـ 1.19 مرة من نموذج SVM .

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of ARIMA}(1,1,0)}{RMSE \text{ of SVM}} = \frac{9.07991}{21.0316} = 0.4317$$

قيمة RMSE لنموذج ARIMA(1,1,0) تشكل 43.1% من قيمة RMSE لنموذج SVM ، أي أن نموذج ARIMA(1,1,0) أكثر دقة وكفاءة في التنبؤ بـ 2.32 مرة من نموذج SVM .

$$\bullet \text{Relative efficiency (R.E.)} = \frac{RMSE \text{ of ARIMA}(1,1,0)}{RMSE \text{ of MLP } 5-5-1} = \frac{9.07991}{17.7467} = 0.5116$$

قيمة RMSE لنموذج ARIMA(1,1,0) تشكل 51.1% من قيمة RMSE لنموذج MLP 5-5-1 ، أي أن نموذج ARIMA(1,1,0) أكثر دقة وكفاءة في التنبؤ بـ 1.95 مرة من نموذج MLP 5-5-1.

7. الخلاصة:

تم تحليل بيانات السلسلة الزمنية لمؤشر أسعار المعادن الشهري خلال الفترة الزمنية من أكتوبر 1990 حتى أكتوبر 2020، واستخدام عدة طرق لاختبية هي MLP, ELM, SVM بالإضافة الى نموذج خطي هو ARIMA، حيث تمت المفاضلة بين هذه النماذج لمعرفة أكثرها دقة في التنبؤ المستقبلي. وقد أشارت نتائج هذه الدراسة الى أن نموذج ELM 5-100-1 أكثر دقة وكفاءة في التنبؤ من ARIMA(1,1,0) و SVM و MLP 5-5-1. والنتيجة الأهم هي أن النموذج التقليدي الخطي ARIMA(1,1,0) كان افضل وأكثر دقة وكفاءة في التنبؤ من نموذجين لاختبيين هما SVM و MLP 5-5-1. وهذا يشير الى ان النماذج التقليدية قد تكون أفضل من النماذج الحديثة في التنبؤ، وذلك حسب طبيعة بيانات الظاهرة المدروسة. ويجب عدم تحديد الأفضلية مسبقاً للنموذج الحديث على النموذج التقليدي.

وبناء على نتائج هذه الدراسة، نوصي باستخدام الطرق الهجينة بين الخطية واللاخطية بدلا من استخدام كل نموذج على حده، والتي من الممكن ان يكون لها نتائج جيدة مقارنة مع النماذج الاخرى المفردة. مثل: ARIMA-MLP, ARIMA-ELM, ARIMA-RBF, ARIMA-SVM, ARIMA-GARCH

المراجع:

1. الرشيدلي، كرم (2010) "تصميم وتنفيذ نظام كشف التطفل باستخدام التقنيات الذكائية"، رسالة ماجستير، قسم هندسة البرمجيات، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.
2. قاسم، عمر (2010)، "تطبيق التقنيات الذكائية في المعلومات الحياتية"، رسالة دكتوراه، قسم الرياضيات، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.
3. Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., & Weingessel, A. (2008). Misc functions of the Department of Statistics (e1071), TU Wien. R package, 1, 5-24.
4. Garg, A. (2012). Forecasting exchange rates using machine learning models with time-varying volatility. Master Thesis. Linköping University.
5. Lewis, N. D. (2017). Neural networks for time series forecasting with R: An intuitive step by step blueprint for beginners. CreateSpace Independent Publishing Platform.
6. Maren, A. J., Harston, C. T., Pap, R. M. (2014). Handbook of neural computing applications. Academic Press.
7. McNelis, D. (2005). "Neural Networks in Finance", Elsevier Inc, California, USA.
8. Mellit, A.; Pavan, A.; Benghane, M. (2013). Least squares support vector machine for short-term prediction of meteorological time series. Theoretical & Applied Climatology, 111(1), 297-307.
9. Meyer, D. (2018). Support Vector Machines—the Interface to libsvm in package e1071. FH Technikum Wien, Austria. David.Meyer@R-Project.org.
10. Moseley, N. (2003). "Modeling Economic Time Series Using A focused Time Lagged Feed Forward Neural Network", Proceeding of Student Research Day, CSIS, Pace University.
11. Mustafa, Z., & Yusof, Y. (2012). A hybridization of enhanced artificial bee colony-least squares support vector machines for price forecasting. Journal of Computer Science, 8(10), 1680-1690.
12. Negnevitsky, M. (2011). "Artificial Intelligence", A Guide to Intelligent Systems 3rd ed, University of Tasmania, Addison-Wesley Longman Publishing Co, Inc. Boston, MA, USA.
13. Ojemakinde, B. T. (2006). Support Vector Regression for Non-Stationary Time Series. Master Thesis. University of Tennessee – Knoxville.
14. Okasha, M. K. (2014). Using support vector machines in financial time series forecasting. International Journal of Statistics and Applications, 4(1), 28-39.
15. Pantazis, G., & Alevizakou, E. G. (2013). The use of artificial neural networks in predicting vertical displacements of structures. International Journal of Applied Science and Technology, 3(5).
16. Samsudin, R., Shabri, A., & Saad, P. (2010). A comparison of time series forecasting using support vector machine and artificial neural network model. Journal of applied sciences, 10(11), 950-958.
17. Suganyadevi, M. V., & Babulal, C. K. (2014). Support vector regression model for the prediction of loadability margin of a power system. Applied Soft Computing, 24, 304-315.

18. Vapnik, V. N., & Vapnik, V.(1998). Statistical learning theory, (Vol. 1). New York: Wiley.
19. Wang, J., Li, L., Niu, D. & Tan, Z. (2012). An annual load forecasting model based on support vector regression with differential evolution algorithm. Applied Energy, 94, 65-70.
20. Wang, W. C., Chau, K. W., Cheng, C. T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of hydrology, 374(3-4), 294-306.
21. White, A. K. and Safi, S. K. (2016). The efficiency of artificial neural networks for forecasting in the presence of autocorrelated disturbances. International Journal of Statistics and Probability, 5(2):51.
22. Yang, H., Huang, K., Chan, L., King, I., & Lyu, M. R. (2004). Outliers treatment in support vector regression for financial time series prediction. In International Conference on Neural Information Processing (pp. 1260-1265). Springer, Berlin, Heidelberg.
23. Zhang, L., Zhou, W. D., Chang, P. C., Yang, J. W., & Li, F. Z. (2013). Iterated time series prediction with multiple support vector regression models. Journal Neurocomputing, 99, 411-422.

الملاحق

Oct-90	49.27	May-98	34.51	Dec-05	69.37	Jul-13	85.67
Nov-90	44.03	Jun-98	33.14	Jan-06	73.13	Aug-13	89.64
Dec-90	42.07	Jul-98	33.1	Feb-06	75.58	Sep-13	88.15
Jan-91	42.11	Aug-98	32.85	Mar-06	76.35	Oct-13	89.07
Feb-91	41.98	Sep-98	33.29	Apr-06	87.8	Nov-13	87.82
Mar-91	41.7	Oct-98	32.28	May-06	101.27	Dec-13	88.71
Apr-91	40.96	Nov-98	32.26	Jun-06	91.46	Jan-14	88.08
May-91	38.4	Dec-98	31.06	Jul-06	97.17	Feb-14	86.16
Jun-91	37.58	Jan-99	30.1	Aug-06	98.01	Mar-14	82.98
Jul-91	38.02	Feb-99	29.93	Sep-06	97.79	Apr-14	85.48
Aug-91	37.32	Mar-99	29.87	Oct-06	101.7	May-14	84.85
Sep-91	37.02	Apr-99	31.53	Nov-06	101.07	Jun-14	84.36
Oct-91	36.3	May-99	32.51	Dec-06	101.77	Jul-14	88.18
Nov-91	36.29	Jun-99	31.73	Jan-07	96.83	Aug-14	88.03
Dec-91	35.19	Jul-99	34.25	Feb-07	98.69	Sep-14	85.07
Jan-92	35.59	Aug-99	35.02	Mar-07	104.59	Oct-14	82.62
Feb-92	37.12	Sep-99	36.64	Apr-07	114.38	Nov-14	82.86
Mar-92	37.41	Oct-99	36.31	May-07	116.38	Dec-14	78.83
Apr-92	38	Nov-99	36.56	Jun-07	110.02	Jan-15	73.85
May-92	37.99	Dec-99	37.86	Jul-07	111.06	Feb-15	72.37
Jun-92	38.08	Jan-00	40.05	Aug-07	105.49	Mar-15	71.78
Jul-92	39.76	Feb-00	40	Sep-07	108.18	Apr-15	72.14
Aug-92	39.66	Mar-00	38.79	Oct-07	114.42	May-15	74.64
Sep-92	38.52	Apr-00	36.8	Nov-07	112.14	Jun-15	70.31
Oct-92	35.75	May-00	37.67	Dec-07	105.31	Jul-15	65.75
Nov-92	34.52	Jun-00	37.29	Jan-08	109.55	Aug-15	62.65
Dec-92	35.44	Jul-00	38.18	Feb-08	117.66	Sep-15	63.4
Jan-93	35.25	Aug-00	38.05	Mar-08	126.26	Oct-15	62.22
Feb-93	35	Sep-00	39.85	Apr-08	125.45	Nov-15	57.83
Mar-93	33.88	Oct-00	37.68	May-08	121.31	Dec-15	56.31
Apr-93	32.34	Nov-00	36.62	Jun-08	117.98	Jan-16	55.21
May-93	31.6	Dec-00	38.01	Jul-08	118.91	Feb-16	57.68
Jun-93	32.17	Jan-01	38.3	Aug-08	109.9	Mar-16	61.19
Jul-93	32.8	Feb-01	37.87	Sep-08	98.02	Apr-16	62
Aug-93	32.32	Mar-01	36.38	Oct-08	72.29	May-16	59.98
Sep-93	30.98	Apr-01	35.82	Nov-08	58.71	Jun-16	60.26
Oct-93	29.68	May-01	36.63	Dec-08	50.63	Jul-16	63.49
Nov-93	29.12	Jun-01	35.1	Jan-09	51.69	Aug-16	63.78
Dec-93	30.6	Jul-01	33.69	Feb-09	51.01	Sep-16	62.83

Jan-94	31.91	Aug-01	32.67	Mar-09	51.82	Oct-16	64.13
Feb-94	33.48	Sep-01	31.8	Apr-09	56.74	Nov-16	71.53
Mar-94	33.78	Oct-01	30.65	May-09	59.43	Dec-16	73.54
Apr-94	33.38	Nov-01	31.62	Jun-09	65.57	Jan-17	74.54
May-94	35.64	Dec-01	32.08	Jul-09	69.67	Feb-17	77.94
Jun-94	37.83	Jan-02	32.82	Aug-09	81.63	Mar-17	77.34
Jul-94	39.43	Feb-02	33.04	Sep-09	77.94	Apr-17	74.04
Aug-94	38.52	Mar-02	33.98	Oct-09	80.53	May-17	72.24
Sep-94	40.75	Apr-02	33.63	Nov-09	84.7	Jun-17	71.71
Oct-94	42.85	May-02	33.17	Dec-09	90.29	Jul-17	75.39
Nov-94	47.11	Jun-02	33.7	Jan-10	96.58	Aug-17	81.56
Dec-94	48.11	Jul-02	33.28	Feb-10	91.11	Sep-17	82.68
Jan-95	51.23	Aug-02	31.83	Mar-10	99.42	Oct-17	83.39
Feb-95	48.01	Sep-02	31.95	Apr-10	108.15	Nov-17	83.46
Mar-95	46.51	Oct-02	32.17	May-10	96.02	Dec-17	84.11
Apr-95	47.03	Nov-02	33.66	Jun-10	89.09	Jan-18	88.57
May-95	45.19	Dec-02	33.78	Jul-10	89.2	Feb-18	88.61
Jun-95	46.81	Jan-03	34.72	Aug-10	97.59	Mar-18	84.47
Jul-95	48.51	Feb-03	35.71	Sep-10	100.64	Apr-18	86.45
Aug-95	48.8	Mar-03	35.09	Oct-10	108.27	May-18	86.79
Sep-95	46.27	Apr-03	33.76	Nov-10	109.43	Jun-18	86.98
Oct-95	44.58	May-03	35.12	Dec-10	114.52	Jul-18	79.77
Nov-95	45.5	Jun-03	35.69	Jan-11	120.81	Aug-18	78.12
Dec-95	45.1	Jul-03	36.21	Feb-11	125.82	Sep-18	77.44
Jan-96	42.8	Aug-03	36.9	Mar-11	121.37	Oct-18	79.35
Feb-96	42.63	Sep-03	36.8	Apr-11	124.29	Nov-18	77.5
Mar-96	43.03	Oct-03	38.83	May-11	118.49	Dec-18	76.07
Apr-96	42.9	Nov-03	40.4	Jun-11	116.98	Jan-19	75.77
May-96	43.23	Dec-03	42.73	Jul-11	120.99	Feb-19	80.12
Jun-96	39.22	Jan-04	45.87	Aug-11	114.77	Mar-19	81.1
Jul-96	37.76	Feb-04	48.75	Sep-11	109.08	Apr-19	81.6
Aug-96	37.91	Mar-04	49.26	Oct-11	98.39	May-19	78.65
Sep-96	36.94	Apr-04	49.42	Nov-11	95.84	Jun-19	78.53
Oct-96	36.01	May-04	46.47	Dec-11	95.08	Jul-19	81.03
Nov-96	38.81	Jun-04	47.8	Jan-12	100.5	Aug-19	76.05
Dec-96	39.43	Jul-04	49.34	Feb-12	104.01	Sep-19	77.32
Jan-97	41.51	Aug-04	48.79	Mar-12	103.54	Oct-19	76.46
Feb-97	41.81	Sep-04	49.28	Apr-12	100.95	Nov-19	76.22
Mar-97	42.75	Oct-04	51.65	May-12	96.63	Dec-19	77.47
Apr-97	41.43	Nov-04	52.1	Jun-12	91.63	Jan-20	77.7
May-97	43.02	Dec-04	52.66	Jul-12	91.19	Feb-20	73.02

Jun-97	42.72	Jan-05	56.4	Aug-12	87.65	Mar-20	68.71
Jul-97	42.42	Feb-05	57.94	Sep-12	93.59	Apr-20	65.55
Aug-97	43.06	Mar-05	60.22	Oct-12	94	May-20	68.01
Sep-97	41.01	Apr-05	59.05	Nov-12	92.42	Jun-20	73.68
Oct-97	39.95	May-05	56.65	Dec-12	97.42	Jul-20	79.09
Nov-97	38.84	Jun-05	57.62	Jan-13	100.29	Aug-20	83.48
Dec-97	36.96	Jul-05	57.77	Feb-13	101.34	Sep-20	85.12
Jan-98	35.94	Aug-05	60.14	Mar-13	94.55	Oct-20	85.49
Feb-98	35.42	Sep-05	60.04	Apr-13	90.73	-	-
Mar-98	35.58	Oct-05	61.68	May-13	88.32	-	-
Apr-98	35.73	Nov-05	64.36	Jun-13	85.41	-	-

Source: <https://www.indexmundi.com/commodities/>