نموذج معدل للإنحدار الخطي المتعدد باستخدام مرشحات تقليص الموجة الصغيرة

(دراسة تطبيقية)

د.محمد توفيق البلقيني
أستاذ الإحصاء التطبيقي والتأمين

د. فاطمة علي عبد العاطي
أستاذ الإحصاء التطبيقي

الباحث

محمد عبد المجيد بدل

الملخص

إهتم الباحث في بحثه على إيجاد نموذج كفوء لمعلمات نموذج الانحدار الخطي المتعدد وتشخيص القيم الشاذة باستخدام تحليل الموجة الصغيرة مع طريقة المربعات الصغرى الاعتيادية (OLS) وبعض الطرق الحصينة المستخدمة في معالجة القيم الشاذة وتقدير نموذج الانحدار الخطي مثل طريقة المربعات المشرذمة (المبتورة) (LTS) وطريقة M الحصين. ومن ثم كشف ومعالجة القيم الشاذة وتقدير النموذج الخطي المتعدد بالتقليص المويجي والذي يتضمن الدوال المويجية واستخدام اسلوب لتقدير مستوى قطع العتبة منها قطع العتبة الناعمة ومن ثم المقارنة بين كفاءة المقدرات المستخدمة بالطرق الاعتيادية والحصينة مع بعض مرشحات الموجة الصغيرة بالتطبيق على مرضى الذبحة القلبية من خلال اقتراح اسلوب يعمل على دمج مخرجات القيم التقديرية للطرق الحصينة واستخدامها كمدخلات في تحليل مرشح المويجة وتقدير افضل نموذج الانحدار الخطي المتعدد يبن نسبة الكولسيترول في الدم للمصابين بالذبحة القلبية من المرضى المسجلين

فـي المركـز التخصصـي للامـراض القلب بمحافظـة اربيل فـي كوردستان العــراق وحســاب بعــض المعــايير الاحصـائية منهـا (MSE , R^2, F-Cal.) للمقارنـة مع انـواع مـن الموجـة الصغيرة في نموذج الانحدار المتعدد والوصول الى افضل نموذج،وتوصل البحــث االــى ان مرشــحات تقلــيص الموجة الصغيرة تعطي أفضل النتائج فـي تقـدير نمـوذج الانحـدار الخطـي المتعــدد مقارنـــة بنمـــوذج الانحـــدار الحصـــين مـــن حيـــث المعـــايير الاحصـائية F-Cal. R^2,MSE, . وتبـين ان مرشـح الموجـة الصـغيرة هارHaar كان هو الأفضل مـن بـين انـواع ومسـتويات الموجـة الصـغيرة فـي تقـدير نمـوذج الأنحـدار الخطـي المتعدد للطريقة المقترحة وخاصة مع المقدر الحصـين LTS بالاعتمـاد علـى المعــايير (MSE, R^2, MAPE وMSAE F-cal.).

SUMMARY

The researcher was interested as for his search in finding an efficient model with parameters of multiple-linear model and in diagnosis the outlier values through the use of analyzing the wavelet with the method of ordinary least squares (OLS) and some methods used for treating the outlier values and used in the treatment of outlier values and estimate the linear model such as least trimmed squares method (LTS) and M method .Therefore, this leads to revealing and treating the Outlier values and estimating the multi-linear model with wave shrinkage that involves the wave functions and the use of a method for estimating the level of cutting threshold including soft thresholding then comparing the efficiency of the estimates with ordinary and robust methods and fortified with some filters of wavelet with the application to patients having angina pectoris through proposing a method that works for inputting the outcomes of estimated values of robust methods as for the analysis of wavelet filter and estimating the best model of multiple linear regression model between the ratio of cholesterol in blood for the patients having angina pectoris who are enrolled in Specialized Center For Cardiology in Erbil in Kurdistan-Iraq and calculating some statistical standards such as (MSE,R^2,F-Cal.) to compare types of wavelet in the multiple regression model and to reach the best model . The research Concluded that the filters wavelet give the best results as for estimating the multiple linear regression model in comparison with the robust regression model in relation to statistical standards (MSE, R^2, F- Cal) and that the filter of wavelet Haar was the best in relation to the types and levels of wavelet Connected with estimating to the multiple linear regression model especially with LTS relying upon the standards (MSE,R^2,MAPE,F-cal.,MSAE).

أولاً : المقدمة

خلال تسعينيات القرن الماضي انتشرت مقدّرات الموجة المتقلصّة (Estimators (Wavelet Shrinkage واختصاراً(WaveShrink). إن تقليص الموجة هو أسلوب لإزالة تشويش الإشارة (Signal) يعتمد على فكرة إجراء قطع عتبة (Thresholding) لمعاملات الموجة والناتجة من تطبيق التحويل المويجي، وهي مقدرّات سهلة التنفيذ من خلال خوارزميّات سريعة مما جعلها تروق للباحثين في الحالات العملية،ومع التطور المنجز في أجهزة الحواسيب من الناحية المادية ومعالجة البيانات أصبح تطور وانجاز العديد من طرائق الانحدار التحويل المويجي ممكنا عمليا ومن تلك الطرائق طرائق الموجة المتقلصّة،وإن ظهور القيم الشاذة في مجموعة من البيانات الحقيقية سيؤثر في تحليلها ومن جهة أخرى فان رفض أو حذف القيم الشاذة من البيانات ليس بالأمر المرغوب به إحصائياً [3] .

وتعْدُّ نظرية الموجة الصغيرة او المويجة (Wavelet theory) من النظريات الحديثة والمهمة ذات الإستخدامات الواسعة والمختلفة في شتى المجالات النظرية والتطبيقية ، نشأت وتطورت في نهاية القرن الماضي وبالتحديد خلال العقدين الأخيرين مع تطور علوم ونظريات الحاسوب والرياضيات ، وتعرف الموجة الصغيرة بانها عبارة عن أداة رياضية تُستخدم كتحويل (Transformation) لعمليات الإشارة processing) (signalوتقليل التشويش

ومعالجة القيم الشاذة ومن ثم تحليل المعلومات من خلال التردد (Frequency) مع الزمن (Time) متوافقاً على تحليل فورير (Fourier Tranformation) الذي يحلل المعلومات عن طريق التردد فقط مهملاً عامل الزمن. ولقد استخدمت من قبل عدة باحثين في مجالات عديدة منها علوم الرياضيات والفيزياء والهندسة منها الهندسة الفيزيائية وعلوم فيزياء الارض وتحليل اشارة الزلازل وحديثاً أستخدمت في علوم الأقتصاد وادارة الموارد المالية.وسميت بالمويجة اوالموجة الصغيرة لتموجها (waving) أعلى وأسفل المحور الأفقي (x-axis) على شكل ذبذبات (oscillates) طفيفة مقارنة بدالة الجيب (sine) التي تعتبر موجة كبيرة (big wave) [10]، ويشكل اسلوب تحويل المويجة مع قطع العتبة (Threshold) عملية معالجة للبيانات من التشويش (ومن ضمنها القيم الشاذة) وعلى هذا الأساس تم إقتراحه كطريقة يمكن استخدامها في معالجة القيم الشاذة[8].

لذا يركز الباحث في البحث على عملية التقدير لمعلمات النموذج الخطي المتعدد وتشخيص القيم الشاذة باستخدام طريقة المربعات الصغرى الاعتيادية (OLS) وبعض الطرائق الحصينة المستخدمه في معالجة القيم الشاذة وتقدير النموذج الخطي مثل طريقة المربعات المشرذمة (المبتورة) (LTS) وطريقة M الحصين . ومن ثم كشف ومعالجة القيم الشاذة وتقدير النموذج الخطي المتعدد بالتقليص المويجي والذي يتضمن الدوال المويجية واستخدام اسلوب لتقدير مستوى قطع العتبة منها قطع العتبة الناعم ومن ثم المقارنة بين كفاءة المقدرات المستخدمة بالطرق الاعتيادية والحصينة مع بعض مرشحات الموجة

الصغيرة بالتطبيق على المجال الطبي .

ثانياً : مشكلة البحث

يواجه الباحث في الكثير من الدراسات التطبيقية وجود بيانات شاذة أو غير مستقرة تؤثر بشكل كبير على نتائج تحليل البيانات ومن ثم على دقة النتائج ولكون البيانات المدروسة تضمنت قيم شاذة لذلك توجب على الباحث ان يعالج هذه المشكلة قبل البدء بالتحليل لذلك أختار الباحث أسلوب تقليص الموجة الصغيرة بطريقتي قطع العتبة الناعم والانحدار الحصين،ولغرض دراسة هذا الاسلوب تم إختيار عينة من المصابين بالذبحة القلبية من خلال نسبة الكوليسترول في الدم وأجري تحليل للبيانات بطريقة الانحدار الخطي المتعدد بأسلوبي تقليص الموجة الصغيرة والانحدار الحصين لإختيار أفضل نموذج رياضي يصف بيانات الدراسة.

ثالثاً : أهمية البحث

١- التعامل مع مشكلة البيانات الشاذة في البيانات المستخدمة في تحليل الانحدار الخطي المتعدد من خلال طرق الانحدار الحصين (Robust Regression)واستخدام طريقة مقترحة تعتمدعلى التقليص المويجي(waveshrink).

٢- مقارنة أسلوب تقليص الموجة الصغيرة مع أنواع من الموجات الصغيرة مع بعض الطرائق الحصينة المستخدمة في معالجة مشكلة البيانات خلال دمج مخرجات القيم التقديرية للطرق الحصينة واستخدامها كمدخلات في تحليل مرشح المويجة وتقدير افضل نموذج الانحدار الخطي المتعدد يبن نسبة الكولسيترول في الدم للمصابين بالذبحة القلبية من المرضى المسجلين في المركز التخصصي للأمراض القلب بمحافظة اربيل في كوردستان العراق، بالإعتماد على بعض المعايير الاحصائية.

رابعاً : هدف البحث

تهدف الدراسة إلى إقتراح نموذج إنحدار خطي متعدد في المجال الطبي لبيانات المصابيين بمرض الذبحة القلبية في العراق من خلال استخدام طريقة مقترحة في التحويل تعالج مشكلة البيانات الشاذة باسلوب يعتمد على تقليص الموجة الصغيرة من خلال عدة مويجات مستخدمة مع قطع العتبة الناعم (Soft Wavelet) والصلبة (Hard Wavelet) او الوسيطة(Mid Wavelet) على انواع من الموجات الصغيرة وحساب بعض المعايير الاحصائية منها (MSE,R^2,F-cal.) ومقارنتها مع بعض الطرائق الحصينة وهي طريقة المربعات الصغرى المشرذمة (LTS) وطريقة مربعات الوسيط الصغرى (LMS) والطريقة الاعتيادية طريقة المربعات الصغرى الاعتيادية (OLS) وتطبيقها على بيانات المصابين بالذبحة القلبية في مستشفى أمراض القلب التخصصي بمحافظة اربيل في العراق ومن ثم التوصل الى أفضل نموذج أنحدار خطي متعدد يمثل الإصابة بالذبحة القلبية .

خامساً : عينة البحث

إستخدمت الدراسة بيانات لعينة من المصابين بالذبحة القلبية في مركز امراض القلب التخصصي لمحافظة أربيل لأقليم كوردستان- العراق في عام ٢٠١٤، بحجم ١٢٨ مصاب بالذبحة القلبية وأن حجم العينة يمثل إحدى شروط التحليل الرياضي لمستويات الموجة الصغيرة للأساس إثنان (2^j) بحيث ان (j) تمثل مستوى الموجة الصغيرة. وتم تقدير نموذج الإنحدار الخطي المتعدد بأسلوبين الأول الإنحدار الحصين والثاني تقليص الموجة الصغيرة.

سادساً: البرامج المستخدمة في التحليل

تمت الإستعانة بالبرنامج الإحصائي (SPSS Ver.22)، فضلاً عن استخدام برنامج ولغة (MATLAB R2014a) وبرنامج (S-PLUS Ver. 6.1) في التحليل الأحصائي ومن ثم الحصول على النتائج المطلوبة.

سابعاً : الإطار النظري

الموجة الصغيرة(المويجة) (Wavelet) :

المويجة (Wavelet) أو ماتسمى بالموجة الصغيرة (Small Wave) هي أحد أنواع الدوال الرياضية المستخدمة لتجزئة الدالة إلى مركبات تردد مختلفة ودراسة كل مركب مع إعادة التصميم (Resolution) الملائم عند كل قياس ، وتعرف المويجة رياضياً بأنها دالة قيمة حقيقية معرفة على محور حقيقي كامل وتتذبذب صعوداً ونزولاً بشكل منتظم حول الصفر، إلى جانب آخر فهي تمثل أداة مميزة وتقنية فعالة وقوية لتمثيل وتحليل البيانات ، وسميت بالمويجة لصغرها وتميزها عن إشارة الموجة الكبيرة (Big Wave) مثل موجة دالة الجيب (Sine) وموجة الجيب تمام (Cosine) وتتميز المويجة بالخواص الآتية:

١- تعمل المويجة على تجزئة البيانات إلى مركبين يمثل المركب الأول الجزء التقريبي للبيانات (يتناسب مع معدل البيانات) وهو ما يسمى بالموجة الأب أو دالة القياس $\phi(.)$ التي تكاملها على الفترة $(-\infty,\infty)$ مساوي للواحد ، أي أن :-

$$\int_{-\infty}^{\infty} \Phi(x)dx = 1 \quad ...(1)$$

يمثل الجزء الثاني المركب التفصيلي الذي يقيس الإختلافات الموجودة في البيانات عند قياسات مختلفة وهو ما يسمى بالموجة الأم أو الدالة المويجة $\Psi(.)$ التي تكاملها على الفترة $(-\infty,\infty)$ مساوي للصفر ، أي أن :-

$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \quad ...(2)$$

بـديهياً هـذا الشـرط يضـمن أن تذبذبات المويجة يجب أن تكون متوازنة أعلى وأسفل الصفر

٢ –تكامـــــل مربـــــع دالـــــة المويجة $\Psi(.)$ يساوي الواحد فـي الفتـــــــــــــــرة $(-\infty,\infty)$:-

$$\int_{-\infty}^{\infty} \psi^2(x)dx = 1 \quad ...(3)$$

المـويجي لإشـارة مـا باسـتخدام مويجة ذات n من العزوم الزائلـة فـان هـذا سيؤدي إلـى تنـاثر تلـك الإشـارة بمعنــى أن العديـد مـن معـاملات المويجـة لأي متعـددة حدود ذات درجة n أو أقل سوف تكـون صـفرية فـي الأجـزاء الممهـدة ولا حاجـة لخزنهـا، فـي حين أن المعاملات غير الصفرية تشـير إلـى القفـزات أو الأجـزاء غيـر الممهـدة مـن الإشـارة وتكاملهـا يسـاوي صـفروكما يلـي :[7]

$$\int_{-\infty}^{\infty} x^k \psi(x)dx = 0 \quad ; k = 1,...,n \quad ...(4)$$

أنواع مرشحات الموجة الصغيرة [10]:

المويجة(هار) Wavelet Haar

المويجـة هـار تعتبـر مـن أبسـط أنــواع المويجــات المســتخدمة لأغراض التحليل ، وقد ظهـرت إلـى الوجـود مـن خـلال الدراسـة التـي قدمـها العـالم Alfred (Haar) فـي الفتـرة - 1909 1910)وبســـبب بســـاطتها وسـهولتها فهـي تعتبـر الخيـار الأفضل لـدى الراغبين فـي تعلـم ودراسة المويجات.ومن المعلوم أن هنــاك دالتــان تلعبــان دوراً أساسياً في تحليل المويجة ، دالـة المويجـــة (Wavelet Function Ψ)أو ما تسمى بدالة المويجـــة الأم ودالـــة القيـــاس (Scaling Function Φ) فبالنسـبة لدالـة المويجـة يمكـن التعبيــر عنهــا مــن خـــلال الصيغة:-

$$\Psi(u) = \begin{cases} 1, & 0 \le u < \frac{1}{2} \\ -1, & \frac{1}{2} \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

... (5)

أمـا دالـة القيـاس فيعبـر عنهـا بالشكل الآتي :-

$$\phi(u) = \begin{cases} 1, & 0 \le u \le 1 \\ 0, & \text{Otherwise} \end{cases}$$

... (6)

إن المويجـة (هـار) لـديها الخـواص الآتية:-

1- الإرتكـــــاز المرصـــــوص (Compact Support)

2- متعامدة (Orthogonal)

3- ثنائيـــــة التعامـــــد (Biorthogonal)

4- متماثلة (Symmetric)

وتـم توضـيح هـذه الخـواص فـي مرشـــح الموجـــة دوبجيـــز (Daubechies) لاحقاً .

المويجـة المتقطعـة (مـايـر) Wavelet Discrete Meyer

المويجة المتقطعة (مـاير) أو ما يسـمى (dmey) هـو مـن نتـاج العـالم (Yves Meyer) الـذي يعـود لـه الفضـل فـي تطـوير طرائـق تحليل المويجة.إن دالـة المويجـــة الأم ودالـــة القيـــاس معرفتـــان فـــي مجـــال التـــردد ، ويمكن التعبير عن دالة المويجة الأم من خلال الصيغة التالية :

$$= \begin{cases} (2\pi)^{-1/2} e^{-iu/2} \sin\left(\frac{\pi}{2} v\left(\frac{3}{2\pi}|u| - 1\right)\right) & if \ \frac{2\pi}{3} \le |u| \le \frac{4\pi}{3} \\ (2\pi)^{-1/2} e^{iu/2} \cos\left(\frac{\pi}{2} v\left(\frac{3}{4\pi}|u| - 1\right)\right) & if \ \frac{4\pi}{3} \le |u| \le \frac{8\pi}{3} \\ 0 & if \ |u| \notin \left[\frac{2\pi}{3}, \frac{8\pi}{3}\right] \end{cases}$$

... (7)

حيـــــــــــــث أن:

$$v(a) = a^4(35 - 84a + 70a^2 - 20a^2) \qquad a \in [0,1]$$

أما دالة القياس فتكون بالشكل التالي :

$$= \begin{cases} (2\pi)^{-1/2} & if \quad |u| \le \frac{2\pi}{3} \\ (2\pi)^{-1/2} \cos\left(\frac{\pi}{2} \nu\left(\frac{3}{2\pi}|u| - 1\right)\right) & if \; \frac{2\pi}{3} \le |u| \le \frac{4\pi}{3} \\ 0 & if \quad |u| > \frac{4\pi}{3} \end{cases}$$

... (8)

إن المويجة المتقطعة (ماير) لديها خاصية مهمة وهي خاصية الإرتكاز المرصوص (Compact Support) وهو يمثل تقريب جيد يقود إلى مرشحات إستجابة النبض المحدد (Finite Impulse Response) والذي من خلاله يمكن إستخدام تحويل المويجة المتقطع (DWT) ، إضافة إلى إمتلاكها نفس خواص المويجة (هار) [4].

المويجة (دوبجيز) Daubechies wavelet

سميت هذه المويجات نسبة إلى الباحثة (Ingrid Daubechies) *والتي تعتبر الباحثة الرائدة في* موضوع المويجة ، وقد ابتكرت ما يسمى بمويجات التعامد الطبيعي(Orthonormal) ذات الإرتكاز المرصوص في عام (1988) مما جعل تحليل المويجة المتقطعة قابلة للتطبيق. وتكتب مرشحات هذه العائلة إختصاراً (DN) أو (dbL1) ، حيث أن (D) و (db) هو مختصر لإسم الباحثة (Daubechies) أما (N) فهو طول المرشح أو رتبته ، في حين أن (L1) هو عدد العزوم المتلاشية أو الزائلة لدالة المويجة ، فمثلاً (D4) و (db2) يدلان على نفس المرشح وهو المرشح من الرتبة الثانية من مرشحات هذه العائلة، ويرتبط (L1) مع (N) بالعلاقة الآتية :

L1 = N/2... (9)

بشكل عام فإن (dbN) تمثل عائلة المويجات ذات الرتبة (N) [علماً أن المويجة

هار هي أحد أفراد هذه العائلة لأن db1 هي نفس المويجة هار] . ولهذه المويجات خصائص يمكن إدراجها بالشكل الآتي :

1- الإرتكاز للموجة (dbN) هو على الفترة [0,2N-1] .

2- الموجة (dbN) لها (N) من العزوم المتلاشية أو الزائلة .

3- إن شكل الدوال للموجة (dbN) بعيدة عن التناظر .

4-يزدادالإنتضام(Regularity or Smoothness) للموجة (dbN) مع تزايد طول المرشح أو رتبته،أي (0.2075 N) والذي يمثل مؤشر الإنتظام (Regularity Index).

الموجة (كويفليتز) Coiflets Wavelet

أوجدت الباحثة (Daubechies) هذه الموجات بناءاً على طلب قدمه الباحث (Coifman) في ربيع 1989 ونسبت إليه ، حيث قام هذا الباحث بطرح فكرة الحصول على العزوم المتلاشية أو الزائلة لمرشحات التمرير الواطئ ومرشحات التمرير العال

معاً لكلتا الدالتين أي(Ψ و Φ) بدلاً أن تكون العزوم المتلاشية مقتصرة على (Ψ) وحدها . وتسمى هذه الموجات إختصاراً (Coif N) حيث (Coif) هي إختصار (Coifman) في حين أن (N) تمثل رتبة المرشح ، وهناك علاقة بين رتبة المرشح مع طوله وهي (طول المرشح =6N) كما أن عدد العزوم المتلاشية لدالة الموجة (Ψ)هي (L = 2N)، في حين أن عدد العزوم المتلاشية لدالة القياس(Φ)هي ($L_1 = 2N - 1$) .

الموجة سامليت:- Symlets Wavelet (sym)

الموجات ساملت هي الموجات المتعامده التي أقترحتهاالباحثة(Duabechies) والتي أجرت تعديلات على

عائلة (db) بزيادة التماثل مع بقاء بساطة المويجة ،وهي متماثلة ولها نفس خصائص (دوبجيز) وان رتبة دوال المويجة تتكون من (2-8) رتبة.

قطع العتبة:
[11]Thersholding

هنالك أنواع عديدة من قطع العتبة المستخدمة مع معاملات تحويل المويجة منها على سبيل المثال لا الحصر قطع العتبة الناعم (Soft Thresholding) ، قطع العتبة الصلبة (Hard Thresholding)قطع العتبة القوية أو المشتركة (Firm Thresholding) ، قطع العتبة الوسيطة (Mid Thresholding) وقطع العتبة (Garrote) غير السالبة. وفي هذا البحث تم إختيار قطع العتبة الناعمة والتي سيتم الإعتماد عليها في الجانب التطبيقي

قطع العتبة الناعم :
(Soft Thresholding)

هي التقنية القياسية لمعالجة القيم الشاذة لمعاملات المويجة W_{n} بواسطة :

$$W_{\mathrm{n}}^{(\mathrm{st})} = sign\{W_{\mathrm{n}}\}(|W_{\mathrm{n}}| - \delta)_{+}$$
...(10)

$$sign\{W_{\mathrm{n}}\} = \begin{Bmatrix} +1 & if & W_n > 0 \\ 0 & if & W_n = 0 \\ -1 & if & W_n < 0 \end{Bmatrix}$$

حيث أن :

كذلك لدينا :

$$,(|W_{\mathrm{n}}| - \delta)_{+} = \begin{bmatrix} 0 \quad if \quad (|W_{\mathrm{n}}| - \delta) \le 0 \\ (|W_{\mathrm{n}}| - \delta) \qquad otherwise \end{bmatrix}$$

إن قطع العتبة الناعمة تدفع كل المعاملات بإتجاه الصفر ، فإذا كانت معاملات المويجة أقل من مستوى قطع العتبة فهي

تذهب إلى الصفر ، لذلك فإن قطع العتبة الناعمة هو خط مستمر (Continuous Mapping). إن التقديرات ناتجة من قاعدة قطع العتبة يعتمد على الخصائص المرغوبة في التقديرات من حيث للعتبة الناعمة تحيز أقل ومتوسط مربعات خطأ كلي أقل [5].

طريقة المربعات الصغرى الإعتيادية لتقدير معلمات نموذج الانحدار الخطي المتعدد

Ordinary Least Squares Method (OLS)

بنيت طريقة المربعات الصغرى على أسلوب تقدير معلمات الإنحدار الخطي ولفترة طويلة من الزمن ، ولهذه الطريقة من خواص ومزايا جيدة تميزت بها مقدراتها إذا أنها تعطي تقديرات غير متحيزة وسهلة التحليل لهذا فضل (Gauss) طريقة المربعات الصغرى الإعتيادية على غيرها من الطرائق لسهولة تقدير المعلمات من البيانات ، لذا تعد طريقة (OLS) مثالية من الناحية العملية إذا ما تحققت جميع الإفتراضات الخاصة بها ، وأن الفكرة الأساسية من طريقة (OLS) هي جعل مربعات الأخطاء أقل ما يمكن والتي نحصل من خلالها على تقدير المعلمات وكما يلي :-

$$\underline{\hat{\beta}} = (x'x)^{-1} x' \underline{y} \quad \ldots(11)$$

$$\underline{Y} = X \underline{\beta} + \underline{U} \quad \ldots(12)$$

حيث أن:

$\underline{Y}$: متجه عامودي n×١ يمثل قيم مشاهدات المتغير التابع

$X_{n \times q+1}$:وتمثل مصفوفة قيم مشاهدات المتغيرات التوضيحية

$\underline{\beta}_{q+1\times1}$:متجه عامودي ١×(q+1) عبارة عن ثوابت يدعى بمتجه معالم الانحدار

U : متغير عشوائي الانحرافات عن العلاقة الخطية التامة وتأثير المتغيرات غير المحددة الأخرى.

الطرق الحصينة : Robust Method

في حالة وجود الشواذ في البيانات فإن طريقة المربعات الصغرى(OLS) لتقدير المعلمات في ظل وجود هذه الشواذ تكون غير كفوءة لانه تحدث حالة عدم تطابق بين بيانات موضوع الدراسة والفروض الأساسية الواجب توافرها في النموذج ، وبذلك تفقد الطرائق التقليدية مثل طريقة المربعات الصغرى خصائصها الجيدة لتقدير معلمات النموذج ، لذا تم إيجاد طرائق إحصائية بديلة تسمى بالطرائق الحصينة ، أما المقدرات الناتجة عن هذه الطريقة البديلة فتسمى بالمقدرات الحصينة التي تعطي مقدرات كفوءة تعادل كفاءة طريقة المربعات الصغرى الإعتيادية في حالة عدم وجود الشواذ ، فضلاً عن توصل كثير من الباحثين الى إثبات أن طريقة (OLS) غير كفؤة في حالة عدم تحقيق إحدى الإفتراضات اللازمة أو الشروط التي تعتمدها هذه الطريقة ، فإن المقدرات الحصينة لنموذج الإنحدار قريبة من كفاءة طريقة المربعات الصغرى الإعتيادية في حالة تحقق هذه الشروط وتمتاز أيضاً بانها مناسبة لفئة واسعة من التوزيعات في تقدير معلمات النموذج الخطي .وفي بحثنا هذا تم إستخدم طريقتين في تقدير الانحدار الحصين وهما [3]:

١- مقدرات -M

(M -Estimate)

من المعروف أن فكرة طريقة المربعات الصغرى (OLS) تعتمد على تصغير مجموع مربعات الخطأ أصغر ما يُمكن

أي:

$$\underset{a,b}{\mathrm{Min}} \sum_{i=1}^{n} \left[Y_i - a - bX_i\right]^2 \quad ...(15)$$

أما الأسلوب الذي تعتمده طريقة M-estimate هو تصغير المقدار الأتي:

$$\underset{a,b}{Min} \sum_{i=1}^{n} P(Y_i - a - bX_i) \quad ...(13)$$

إذ أن P دالة محدبة متماثلة SymmetricConvex Function نحصل منها على مقدرات حصينة ولتصغير المقدار أعلاه يتم أخذ المشتقة الجزئية بالنسبة للمعلمات ومساواتها للصفر وكما يأتي:

$$\left.\begin{aligned} &\sum_{i=1}^{n} \psi\,(Y_i - a - bX_i) = 0 \\ &\sum_{i=1}^{n} \psi\,(Y_i - a - bX_i)X_i = 0 \end{aligned}\right\} \quad ...(14)$$

إذ أن ψ تمثل المشتقة الجزئية للدالة p $(\psi = p')$ بالنسبة للمعلمات. وأن المعادلة (13) يُمكن حلها بعدة طرائق، منها الطرائق العددية المعروفة مثل طريقة نيوتن رافسون (Newton-Raphson Method). أو بالاعتماد على طريقة (IterativelyWeighted Least Squares المربعات الصغرى الموزونة التكرارية (IWLS) وفقاً لما يأتي:

$$\hat{\beta}_{Rob.} = \begin{bmatrix} a \\ b \end{bmatrix} = (X'WX)^{-1} X'WY$$

٢- طريقة المربعات الصغرى المشرذمة (LTS)

Least Trimmed Squares Method

إن هذه الطريقة إقترحها Rousseeuw1984 [9] وهي طريقة إحصائية لتقدير معلمات نموذج الإنحدار الخطي وتعد حصينة وبديلة للطرائق التقليدية لأنها تتميز بكفاءة إحصائية وإستقرار موضعي أفضل من طريقة مربعات الوسيط الصغرى ولنفرض بأن نموذج الإنحدار الخطي لعينة (X_i, Y_i) وللمتغير المعتمد $Y_i \in R$ ومتجه من المتغيرات المستقلة $X \in R^P$ و $\beta \in R^p$ يمكن الحصول على مقدر $\hat{\beta}_{LTS}$ بواسطة طريقة المربعات

المشـــرذمة الصـــغرى (LTS) كـــــــــــــــــالأتي:[2]

$$\hat{\beta}_{LTS} = \min_{b} \sum_{i=1}^{h} \left(e^2 i\right) \quad ...(16)$$

h : ثابت بمدى $(\frac{n}{2} < h < n)$

و أن $(e^2{}_i)$ $_{i=1,2,,,,:n}$ هي مربع الأخطاء المرتبة (الأخطاء تربع ثم ترتب) .

التقليص المويجي : Waveshrink

تــتلخص طريقــة تقــدير معلمات الإنحدار الخطي المتعدد بإستخدام تقليص المويجـة مـن خـلال معالجـة القيم الشـاذة التـي يمكـن أن تتعـرض لـه مشـاهدات المتغيـر المعتمـد بالإعتمـاد علـى التقلــيص المــويجي ومــن ثــم اســتخدام طريقــة المربعـــات الصـغرى الإعتياديـة فـي تقـدير معلمات الانحدار الخطي المتعدد لنـاتج البيانـات المعالجـة للمتغيـر المعتمــد (مــع الحفـــاظ علـــى 99.97% من طاقة البيانات) مع البيانـــات الأصـــلية للمتغيـــرات المســتقلة [8]. ولغــرض عــزل القيم الشـاذة عـن قـيم مشـاهدات المتغير المعتمد يستخدم عادةً أحد أنـواع قطـع العتبـة (مثـل قطـع العتبـة الصـلبة أو الناعمـة) مـن خــلال تقلــيص (Shrinkage) معـاملات التفصـيل والتـي يمكـن أن نحصل عليها من إعادة تغطية المشـاهدات الأصـلية وتجزئتهـا إلى مركبتين باستخدام المويجات وتمثل الأولـى مجمـوع معـاملات التفصيل (Details) ، بينما تمثل الثانيـــة معـــاملات التمهيـــد (Smooth) بالإعتمـــاد علـــى تحليـــل متعـــدد إعـــادة الحـــل (MRA)،[5] أي أن :

$$y = w^T W = \sum_{j=1}^{J_0} w_j^T W_j + v_{J_0}^T V_{J_0} \quad ...(17)$$

يـتم تقـدير مسـتوى قطـع العت

الثابتـة مـن قطـع العتبـة (Fixed from الطرائق المعروفـة ومنهـ القيم threshold) عند المستوى j = 1 فقط عند (W_1) ، ومن ثم قطع العتبة الناعمة في معالجة معاملات التحويل المتقطع للمويجة المعـاملات المتبقيـة إلـى عناصـر المتجـه W نـتمكن مـن الحصـو معاملات التحويل المتقطع للمويجة المعدلة $(Modify \;\; W)$ التي ي عادة W' والتي يمكن من خلالها إعادة تغطية مشـاهدات المتغيـر المعالج [6]، أي أن:-

$$\tilde{y} = w^T W' \quad ...(18)$$

بالإعتمـــــاد علـــــى مصفوفة المويجة W مثل

الموجـة الصــغيرة (db_N) أو الموجــــة الصــــغيرة (Coiflets). لنحصــــل علـــى قـــيم $\tilde{y}$ (مشـــاهدات المتغير المعتمد المعالجة) والتي سيتم استخدامها مع المتغير المستقل في تقدير معلمـات نمـوذج الإنحـدار الخطــي بالإعتمــاد علــى طريقـــــة المربعــــــات الصــــــــــــغرى الإعتياديـة (OLS)، أي أن :-

$$\underline{\hat{\beta}} = (x'x)^{-1} x'\tilde{y} \quad ...(19)$$

ويــــــتم تلخــــــيص الطرائـق المسـتخدمة فـي طريقـــــة المربعــــــات الصــــــــغرى (OLS) والحصــينة ثـم الاسـلوب المقتــــرح التقلـــــيص المــــويجي علـــــى دوال الموجة الصغيرة للوصول الـــى نمـــوذج الإنحـــدار الخطـي المتعدد والمقارنـة بين جميع هذه الطرق من حيث كفائتها بوجود القيم الشاذة من خلال المقاييس الاحصائية .

ثامناً : المقاييس الاحصائية

١-متوسط مربعات الخطأMSE:

هـذا المعيــار هــو احــد المعــايير الإحصـــائية المهمـــة، يستخدم للمقارنـة بين المقدرات أو النمــاذج المقـدرة. فــالنموذج الأفضل هو النموذج الذي يمتلك اقل قيمة MSE. ويمكـن حسـاب متوسط مربعات الخطـأ للنمـوذج المقدر حسب الصيغة الآتية:

$$MSE = \frac{SSE}{n-k-1} \quad ...(20)$$

٢- معامل التحديد R^2:

هـــذا المعيـــار هـــو احـــد المعايير الإحصائية التي تستخدم لمعرفة نسبة مساهمة المتغيرات التوضيحية في تفسير التغيرات الحاصلة فـي متغير الإستجابة. ويمكن حساب قيمة R^2 حسب الصيغة الآتية:

$$R^2 = \frac{SSR}{SST} \quad ...(21)$$

٣- متوســـــط الاخطـــــاء النســــــبية المطلقــــــة

MeanAbsolute PercentageError (MAPE)

وهو متوسط الاخطاء النسبية للاخطاء لمجموعة من البيانات التي اخذت من دون اشارة فهو مقياس وحيد للدقة المستخدمة بصورة عامة في الطرق الكمية للتقدير وياخذ هذا المعيار الصيغة التالية :-

$$MAPE = \frac{\sum_{i=1}^{n} |PEi|}{n} \quad ...(22)$$

$$PEi = \frac{(y_i - \hat{y}_i)}{y_i} = \frac{e_i}{y_i}$$

PEi : تمثل الخطا النسبي

٤- متوسط مجموع الاخطاء المطلقة Mean Sum (MSAE) of Absolute Error

وهو من المعايير الشائعة للمفاضلة بين طريقة المربعات الصغرى وطريقة اقل خطأ مطلق وان تصغير متوسط الاخطاء المطلقة هو بمثابة تصغير مجموع الاخطاء المطلقة ويسمى احياناً بأقل قيمة مطلقة (Least Absolute Value) ويأخذ الصيغة التالية -:

$$MSAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n} \quad ...(23)$$

تاسعاً : مجال التطبيق

إختص التطبيق بتقدير نماذج الإنحدار الخطي المتعدد على مجموعة من المتغيرات التوضيحية للمصابين بالذبحة القلبية. وإستخدمت طريقة المربعات الصغرى لتقدير نماذج الانحدار باسلوب مباشر (Enter)بادخال كل المتغيرات التوضيحية في النموذج وتطبيق اسلوب الاختيار التدريجي (Stepwise) وحذف المتغيرات التوضيحية الغير مهمة والمؤثرة في دراسة النموذج الخطي المتعدد بعد ان تم تعديل متغير الاستجابة باسلوب مرشح الموجة الصغيرة ليكون نموذج معدل بالانحدار الخطي المتعدد باسلوب الموجة الصغيرة .

وتـــم وصـــف المتغيـــرات المســتخدمة فــي الدراســة كالاتي :

أولاً: متغير الإستجابة:

Y : نســـبة الكوليســـترول للمصـــابيين بالذبحـــة القلبيـــة مجم/١٠٠ملم.

ثانياً: المتغيرات التوضيحية:

X_1 : نسـبة الشـحوم البروتينيـة العاليـــــــــــة (H.D.L) مجـم/١٠٠ملـم، X_2 : نسـبة الشـحوم البروتينيـة الواطئـة (L.D.L) مجم/١٠٠ملم، X_3 : نسبة الشحوم الثلاثية (ST.G) مجم/١٠٠ملم

X_4 : وزن المصـــاب بـــالكيلو غـرام، X_5 : عمـر المصـــاب بالسـنوات ، X_6 : مـدة الاصـابة بالذبحة القلبية بالإشهر.

١- تحليـل الانحـدار بطريقـة الانحـــدار الحصـــين :

تم تطبيق البيانـات علـى نمـوذج الانحـدار الحصـين للنمـوذج الانحـدار المتكـون مـن ثـلاث متغيـرات توضـيحية تـؤثر فـي تفسير النموذج (LDL,HDL, ST.G) باسـتخدام كـل مـن طريقتـي المربعـات الصـغرى المبتـورة LTS و طريقـة M الحصـينة بطريقـة تحليـل الانحـدار المباشـر والانحـدار التـدريجي ولخصـت النتـائج بالمقـاييس الاحصـائية كمـا فـي الجـدول(١) التـالي :

جدول (١) نتائج تحليل الانحدار الحصين

Robust		MSE	R^2	F_cal.	MAPE	MSAE
LTS	Enter	809.2848	0.9771	2210.292	0.05027	9.911382
	Stepwise	766.6609	0.61148	168.3938	0.047801	9.546725
M	Enter	806.522	0.98047	2214.637	0.04718	9.413826
	Stepwise	786.7553	0.6797	168.016	0.047116	9.413258

2-تحليـل الانحـدار باسـلوب الموجة الصغيرة

تـم تحليـل البيانـات علـى نمـوذج الانحـــدار الخطـــي المتعـــدد و

المتكـون مـن ثـلاث متغيـرات توضـيحية تـؤثر فـي تفسـير النمـوذج(LDL,HDL,ST.G) باسـتخدام الانحـدار المتعـدد الخطـي واسـلوب التقلـيص المويجي على خمسة انواع من الموجـة الصـغيرة هـي Haar,Db,Sym,Coif,dme وبعـدد مـن مسـتويات (y) الموجـة الصـغيرة ولخصـت النتائج بالمقاييس الاحصائية كما فـي الجـدول (٢) ومـن ثـم تـم اسـتخدام مقتـرح الباحـث المتضـمن اسـتخدام مخرجـات القيم التقديرية للطرق الحصينة وجعلهـا مـدخلات فـي تحليـل الموجة الصغيرة ولخصت نتائج التحليل في الجدولين (٣) و (٤) وكماهو موضح كالتالي :

جدول (٢) نتائج تحليل الإنحدارالمتعدد التدريجي بأسلوب تقليص الموجة الصغيرة على أنواع من الموجة الصغيرة

Wavelet Type	MSE	R2	F_cal.	MAPE	MSAE
Haar	476.383	0.781	147.559	0.073230	14.05470
Db2	438.671	0.778	144.584	0.070274	13.47895
Db3	487.651	0.882	144.266	0.074936	14.28492
Db4	492.663	0.773	140.970	0.075106	14.32835
Db5	499.709	0.771	139.238	0.075883	14.43427
Db6	498.222	0.772	139.682	0.075414	14.38624
Db7	506.869	0.768	137.082	0.077122	14.62218
Db8	488.806	0.776	143.052	0.074636	14.32079
Sym2	489.783	0.776	143.007	0.075036	14.31829
Sym3	487.651	0.777	144.266	0.074936	14.28492
Sym4	493.844	0.774	141.336	0.075686	14.43984
Sym5	484.884	0.777	144.086	0.073645	14.08143
Sym6	488.454	0.775	142.704	0.074525	14.27857
Sym7	486.657	0.777	143.652	0.073717	14.07008
Sym8	489.538	0.775	142.618	0.075336	14.36199
Coif1	486.053	0.778	144.770	0.074936	14.26846
Coif2	485.722	0.777	143.954	0.074821	14.26762
Coif3	490.995	0.774	141.693	0.075391	14.36923
Coif4	482.455	0.778	144.599	0.074629	14.25714

Coif5	490.657	0.774	141.488	0.075195	14.34058
Dmey	491.787	0.772	139.670	0.074819	14.24167

جدول (٣) نتـائج تحليـل الانحدار للمقدرات الحصينة بطريقـة Mالحصين بأسلوب تقليص الموجة الصغيرة

Wavelet Type	MSE	R2	F_cal.	MAPE	MSAE
Haar	3.270	0.999	32477.679	0.007747	1.435911
Db2	4.114	0.999	22821.886	0.008514	1.579911
Db3	3.975	0.998	27190.228	0.008837	1.612932
Db4	3.371	0.999	32057.241	0.007777	1.450104
Db5	4.577	0.998	23741.006	0.009277	1.716183
Db6	9.316	0.996	10355.794	0.012439	2.288767
Db7	4.324	0.998	25249.748	0.009352	1.691359
Db8	11.750	0.995	8403.010	0.014952	2.75124
Sym2	4.150	0.998	26058.342	0.008615	1.590892
Sym3	11.750	0.995	8380.514	0.01367	2.522416
Sym4	4.106	0.998	26330.776	0.008797	1.647566
Sym5	10.448	0.996	9364.397	0.014151	2.649244
Sym6	13.333	0.994	7381.312	0.015374	2.909013
Sym7	19.727	0.991	4508.691	0.018566	3.533972
Sym8	33.747	0.983	2376.531	0.024709	4.682558
Coif1	3.932	0.999	27705.067	0.008802	1.638234
Coif2	4.470	0.998	24278.683	0.009212	1.715416
Coif3	4.037	0.998	26781.192	0.00891	1.638967
Coif4	4.461	0.998	24247.977	0.009398	1.743937
Coif5	4.130	0.998	26025.079	0.009142	1.687029
Dmey	4.678	0.998	22482.868	0.009604	1.743956

جدول (٤) نتائج تحليل الانحدار للمقدرات الحصينة بطريقة LTSالحصين بأسلوب تقليص الموجة الصغيرة

Wavelet Type	MSE	R^2	F_cal.	MAPE	MSAE
Haar	3.262	0.999	31192.417	0.007683	1.42587
Db2	4.126	0.998	22178.250	0.008518	1.58036
Db3	3.983	0.998	25744.363	0.008864	1.61681
Db4	3.891	0.998	26643.521	0.008605	1.58369
Db5	4.569	0.998	23199.089	0.009239	1.71223
Db6	3.847	0.998	27119.587	0.009243	1.71221
Db7	4.289	0.998	24239.726	0.009304	1.68600
Db8	3.910	0.998	26809.594	0.008362	1.53436
Sym2	4.142	0.998	25021.728	0.008622	1.59397
Sym3	3.981	0.998	25744.363	0.008864	1.61680
Sym4	4.043	0.998	25449.953	0.008721	1.64119
Sym5	3.658	0.999	28110.989	0.008427	1.55420
Sym6	4.160	0.998	24825.100	0.008739	1.65469
Sym7	4.268	0.998	24336.570	0.009121	1.68633
Sym8	4.645	0.998	22701.702	0.009392	1.73547
Coif1	3.870	0.998	26565.474	0.008656	1.60781
Coif2	4.468	0.998	23303.652	0.009209	1.71643
Coif3	4.045	0.998	25762.787	0.008885	1.63521
Coif4	4.389	0.998	23222.183	0.009322	1.73289
Coif5	4.049	0.998	25087.976	0.009072	1.67253
Dmey	4.676	0.998	21551.187	0.009578	1.74376

عاشراً : الإستنتاجات والتوصيات

١-الإستنتاجات:

توصل البحث الى جملة من الإستنتاجات المبنية على هذه النتائج. فيما يلي اهم هذه الإستنتاجات :

١- اعطت مرشحات تقليص الموجة الصغيرة أفضل النتائج في تقدير نموذج الانحدار الخطي المتعدد مقارنة بنموذج الانحدار الحصين من حيث المعايير الاحصائيةً R^2,MSE, .F-Cal.

٢- ساهمت الطريقة المقترحة لمرشح الموجة الصغيرة مع المقدرات الحصينة على الحصول على أفضل نموذج للأنحدار الخطي المتعدد واعطت أفضل النتائج من خلال مقارنة المعايير الاحصائية مع تقديرات نموذج الموجة الصغيرة والتقدير الحصين .

٣- أن مرشح الموجة الصغيرة هارHaar هو الأفضل في تقدير نموذج الأنحدار الخطي المتعدد للطريقة المقترحة وخاصة مع المقدر الحصين LTS بالاعتماد على المعايير MSE, R^2 ,MAPE, F-cal. و MSAE

٤- ان طريقة تحليل الموجة الصغيرة ساعد على معالجة وتقليل القيم الشاذة في البيانات مثل طريقتي المقدرات الحصينة بحيث اصبحت القيم الشاذة قيمتين فقط بعد ما كانت ستة قيم شاذة.

٢-التوصيات :

بناءاً على النتائج والإستنتاجات التي تم التوصل اليها خرج االبحث بالتوصيات التالية :

١.اعتماد مرشح الموجة الصغيرة هار بالطريقة المقترحة في تقدير نموذج الانحدار الخطي المتعدد ومعالجة القيم الشاذة في البيانات.

٢.تطبيق مرشحات الموجة الصغيرة فضلاً عن الطريقة المقترحة في تقدير نماذج السلسلة الزمنية ومقارنتها مع الشبكات العصبية الاصطناعية .

٣.إجراء بحوث حول استخدام الطريقة المقترحة في هذا البحث مع لوحات السيطرة النوعية ومقارنة نتائجها .

٤.استخدام الموجة الصغيرة وتطبيقها على قطع العتبة

الوسط والخشن في نموذج الانحدار اللاخطي والنموذج اللامعلمي في الانحدار .

إحدى عشر : المصادر

اولاً : المصادر العربية

١. الزبيدي، طه حسين علي(٢٠٠٩)"استخدام الموجة الصغيرة المتقطعة في تحليل السلسلة الزمنية AR(1) ومقارنته مع مرشحات اخرى" اطروحة دكتوراه في الاحصاء ،كلية الادارة والاقتصاد،جامعة بغداد، العراق.

٢. القهوجي، وصفي طاهر صالح (2000)"بعض طرق التقديرات الحصينة في تحسين الصور الرقمية "اطروحة دكتوراه فلسفة في الاحصاء (غير منشورة)، كلية الادارة والاقتصاد،جامعة بغداد، العراق.

٣. عبدالله،به ري خان (٢٠٠٥)"مقارنة بين المقدارت الاعتيادية وبعض المقدرات الحصينة لمعلمات الانموذج الانحدار الخطي المتعدد بتطبيق في مجال الانواء الجوية" رسالة ماجستير في الاحصاء، الادارة والاقتصاد،جامعة صلاح الدين، العراق.

ثانياً - المصادر الاجنبية :

4.Burrus C. S. , Gopinath , R. A. & Guo H. (1998) "Introduction to wavelets and wavelet transforms", Printice - Hall , Inc,Upper Saddle River, N.J. , USA ,P.1, PP.76-97 , PP.196-213.

5.CascioL.L.(2007)"Wavelet analysis and denoising:New tools for economists" , Queen Marry Press , university of London ,Paper work.

6.Ergen,B.&Tatar,y.&Culcur,H.(2010)",Time-Freguency analysis of Phonocardiongram SignalUsingWavelet Transform(aComparativeStudy),Dep.of Computer Enginerering ,Firat Unviersity , Istanbul,Tukey .

7.Fugal D. L. (2009) "Conceptual wavelets in digital signal processing",Space and Signal technical publishing , San Diego , USA .

8.Merry.R.J.E,(2005)" Wavelet Theory and Applications Eindhoven, June 7 .

9.Rousseeuw, P.J.& Van Aelst,S.(1999)"Positiv Breakdown Robust Methods in Computer Vision" . John Wiley and Sons, New York, U.S.A.

10.Ramazan,Gangay,Faruk Selguk andBrandonWhitcher(2002),An Introduction to Wavelet and other Filtering Methods in Finance and Economics

11. Xuli Zong , Andrew F. Laine , Edward A. Geiser and David C. Wilson , (1996) , De-Noising and Contrast Enhancement via Wavelet Shrinkage and nonlinear adaptive gain.