مقارنة بين معيار AIC المتين لتقدير القياس بنقطة تحطم عالية ومعيار BIC المتين للوصول لأفضل توفيق للنموذج

د. سمية محمد على

مدرس الإحصاء التطبيقى كلية التجارة جامعة الإسكندرية

د. نوال أحمد ابوزيد

مدرس الإحصاء التطبيقى كلية التجارة جامعة الإسكندرية

د. شيماء عبد الفتاح محمود

مدرس مساعد معهد العجمى للإدارة و المحاسبة

ملخص:

تقدم الدراسة الحالية مقارنة بين معيار أكايك المتين Robust AIC ومعيار Robust Bayesian (BIC) حيث يعتمد كل منهم علي مقدر MM المتين. سيتم من خلال الدراسة تقييم اداء هذه المعايير في إختيار النموذج مقارنة بالمعايير المتينة التي تم تقديمها في دراسة Saleh(2014) وهي معيار AIC المتين ومعامل التحديد المتين الذي يعتمد كل منهم علي مقدرات LTS و LMS والتي تم من خلالهما التوصل إلي إختيار النموذج الصحيح.

الكلمات المفتاحية: إختيار المتغيرات Variables Selection - معيار AIC المتين – معيار BIC المتين – مقدر MM – مقدر LTS- مقدر LMS.

١ـ مقدمة

قدم Akaike (1992) معيارمعلومات اكايك "Akaike Information Criterion" (**AIC**) بهدف اختيار النموذج الذي يمثل البيانات الفعلية بشكل صحيح (اي النموذج الذي يمكن من خلاله ان تفسر المتغيرات التفسيرية التغير الذي يحدث في متغير الإستجابة بشكل صحيح) وفي هذه الحالة نستطيع القول بأنه نموذج نحصل من خلاله علي توفيق جيد للبيانات Excellent Fit.
ولتحقيق هذا الهدف والوصول إلي ذلك النموذج هناك عدة معايير، منها معيار AIC والذي يمكن تعريفه كما يلي:

$$AIC= -2log\, L(\hat{\theta}) +2\, p \qquad (1)$$

حيث :

$\hat{\theta}$: مقدر الإمكان الأعظم للمعلمة (θ) maximum likelihood estimator

$L(\hat{\theta})$: دالة الإمكان الأعظم maximized likelihood function

P : عدد المعالم في النموذج

يكون هناك عدد من النماذج الممكنة يتم من بينها إختيار النموذج الذي تكون له أقل قيمة لمعيار AIC ، وذلك لانه في هذه الحالة يكون هذا النموذج هو أفضل نموذج تقريبي للنموذج الحقيقي. وذلك لأن معيار AIC هو تقدير للمسافة النسبية Relative Distance بين دالة الإمكان الحقيقية True Likelihood Function للبيانات ودالة الإمكان الموفقة Fitted Likelihood Function . لذلك فإن أقل قيمة للمعيار تشير إلي ان النموذج الذي تم إختياره بناء علي أقل قيمة لمعيار AIC هو النموذج الأقرب للنموذج الحقيقي.لاحقا سيتم توضيح خطوات تطبيق معيار AIC.

يمكن تعريف معيار AIC كما يلي :

$$AIC= ln\,(\frac{SSE}{n}) + 2\, p \qquad (2)$$

SSE: مجموع مربعات الخطأ

حيث : $ln\,(\frac{SSE}{n}) = -\,2\, log\, L(\hat{\theta})$

وهنا يجدر الإشارة إلي ان معيار AIC في معادلة (2) يعتمد علي المربعات الصغري (Least Square). وهذه الطريقة شديدة الحساسية للمشاهدات الشاذة وبالتالي، فإن معيار AIC الذي يعتمد علي (LS) يتأثر أيضاً بالمشاهدات الشاذة .

لذلك قام العديد من الباحثين بتقديم معيارAIC متين Robust . وفيما يلي سيتم توضيح معيار AIC الذي قدمته دراسة Ronchetti(1982)

معيار AIC المتين الذي يعتمد على مقدر M

يعتمد معيار AIC لدراسة Ronchetti(1982) علي مقدر M ، ويسمي معيار AIC المتين

(RAIC) Robust Akaike Information Criterion

يمكن تعريف RAIC كما يلي:

$$\mathrm{RAIC(p},\alpha,\rho)=\sum_{i=1}^{i=n}\rho\left(\frac{(y_i-\mu-x_i{}^T\,\hat{\beta}_m)}{\hat{\sigma}}\right)+\alpha\, p \qquad (3)$$

حيث:

x_i^T : متجه به عدد p من المتغيرات التفسيرية

β : متجه به عدد p من المعالم

$\hat{\beta}_m$: مقدر M

μ : معلمة الجزء المقطوع Intercept parameter

ρ : دالة متماثلة Symmetric Function

$\hat{\sigma}$: بعض المقدرات المتينة للمعلمة (σ) *some robust estimate of*
تشير المعادلة (*3*) إلي النموذج الذي يقلل المقدار السابق *Minimize RAIC* (أي النموذج الذي يقترب من النموذج الصحيح True Model).

- أوضح Huber انه يمكن التوصل إلي معيار AIC متين من خلال إستخدام تقدير القياس للأخطاء Scale estimate of Errors الذي يمكن تعريفه كما يلي:

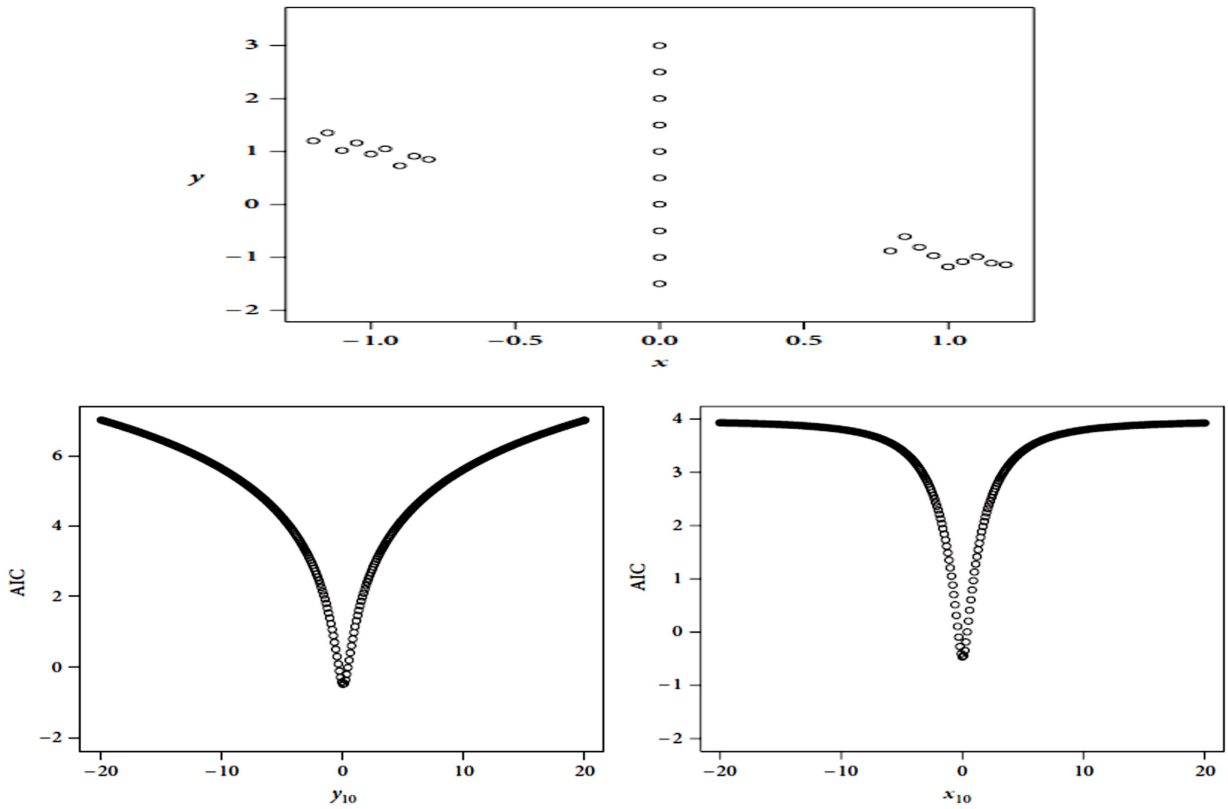
$$S=\frac{SSE}{n-p} \qquad (4)$$

يمكن إستبدال قيمة SSE في معادلة (2) بالمقدار (S) في معادلة (4) ليصبح AIC_S كما يلي:

$$AIC_S = \ln\left(\frac{(n-p)S^2}{n}\right) + 2\,p \qquad (5)$$

- وقد تم من خلال دراسات أخري تتعلق بمعيار *AIC* التوصل إلي معيار *AIC* متين آخر من هذه الدراسات :

دراسة Tharamaratnam and Claeskens(2013) حيث قدما معيار AIC يعتمد في حسابه علي مقدر S ، ودراسة Saleh(2014) حيث هدف إلي تقديم معيار AIC متين من خلال تعديل تقدير القياس modifying the estimate of scale .
حيث أنه كلما ازداد التلوث Contamination في البيانات كلما ازدادت قيمة AIC ، ويمكن توضيح ذلك من خلال شكل (١)



شكل(١): يوضح تأثير قيم كل من (x_{10}) و (y_{10})علي معيار AIC
المصدر: Saleh(2014)

يوضح الشكل السابق تأثير المشاهدات الشاذة علي AIC سواء في اتجاه y (a vertical outlier) أو اتجاه x (a leverage point). علي سبيل المثال، إذا تم إضافة قيمة ملوثة للبيانات وهي y_{10} حيث تتراوح قيم y بين (3 , 1.5-). يلاحظ من خلال ذلك أنه كلما إزداد التلوث (كلما إزدادت قيمة y_{10}) كلما إزدادت قيمة AIC . ويتضح ذلك من خلال جدول (١).

جدول(١): تأثير قيم (y_{10} و x_{10}) الملوثة علي معيار AIC
دراسة Saleh(2014)

y_{10}	AIC	x_{10}	AIC
3	3.2	4.5	3.3
2.5	2.9	4	3.2
2	2.4	3.5	3.0
1.5	1.9	3	2.8
1	1.2	2.5	2.6
0.5	0.2	**0**	**−0.5**
0	**−0.5**	−2.5	2.6
−0.5	0.3	−3	2.9
−1	1.2	−3.5	3.1
−1.5	1.9	−4	3.2

قدم Saleh(2014) حالة متينة Robust Version من معيار AIC ، وذلك من خلال إستبدال AIC في معادلة (5) بمقدر متين للقياس Robust Estimator of Scale ، والذي يكون نقطة تحطمه (Breakdown٥٠%) ويعتمد علي كل من مقدر LTS و LMS.

معيار AIC المتين الذي يعتمد على مقدر LTS

يعتمد معيار **AIC المتين** الذي إقترحه Saleh(2014) علي تقدير قياس يعتمد علي مقدر متين مثل (LTS) (estimate scale from LTS)

يمكن توضيح تقدير القياس لمقدر LTS كما يلي :

$$\hat{\sigma}^2{}_{LTS} = \frac{1}{H}\sum_{i=1}^{i=H} r^2{}_{(i)}\left(\hat{\beta}_{LTS,H,N}\right) \qquad (6)$$

حيث أنه في الآونه الأخيرة تم التركيز علي طرق التقدير التي يكون لها نقطة تحطم عالية.والعديد من هذه الطرق يعتمد علي تقليل تقدير القياس أكثر من التركيز علي تقليل مجموع مربعات البواقي لذلك قام Rousseeuw(1984) بتقديم مقدر Least Median Squares (LMS) ، ولهذا المقدرنقطة تحطم عالية . وتقوم فكرته علي تقليل الوسيط لمربعات البواقي .

كما قدم أيضاً مقدر المربعات الصغري المشذبة Least Trimmed Squares (LTS) والذي يعتمد علي تقليل مجموع البواقي المربعة الصغيرة .

يتضح من خلال جدول (٢) أن معيار AIC_{LTS} (الذي يعتمد علي مقدر LTS) له أقل قيمة في ظل وجود تلوث سواء في جانب x أو y ، وذلك مقارنة بمعيار AIC المعتمد علي بعض المقدرات الاخري مثل مقدرات (M) و (S) .

جدول(٢) مقارنة بين بعض معايير AIC المتينة
دراسة Saleh(2014)

y_{10}	AIC_M	AIC_{BS}	AIC_{LTS}	x_{10}	AIC_M	AIC_{BS}	AIC_{LTS}
3	3.7	3.9	3.4	4.5	3.9	3.9	3.4
2.5	3.7	3.9	3.4	4	3.7	3.9	3.4
2	3.7	3.9	3.4	3.5	3.6	3.9	3.4
1.5	3.7	3.9	3.4	3	3.6	3.9	3.4
1	3.7	3.9	3.4	2.5	3.7	3.9	3.4
0.5	3.7	3.9	3.4	**0**	**3.5**	**3.6**	**3.2**
0	**3.6**	**3.6**	**3.2**	−2.5	3.7	3.9	3.4
−0.5	3.7	3.9	3.4	−3	3.7	3.9	3.4
−1	3.7	3.9	3.4	−3.5	3.9	3.9	3.4
−1.5	3.7	3.9	3.4	−4	4	3.9	3.4

وتعتمد **فكرة معيار AIC** في اختيار المتغيرات علي استبعاد المتغيرات التفسيرية التي يكون لها قيمة AIC كبيرة مقارنة بالمتغيرات الأخري. ويطلق علي هذه الطريقة Backward Elimination ويمكن توضيح خطوات ذلك الإجراء كما يلي:

١- إجراء الإنحدار للنموذج بكل المتغيرات predictors
٢- حذف المتغير الذي يكون له اكبر قيمة لمعيار AIC
٣- إعادة توفيق النموذج Refit the Model بدون المتغير الذي تم حذفه وتطبيق خطوة ٢ مرة أخري.

- كما يمكن أيضاً إختيار النموذج عن طريق إيجاد قيم معيار (AIC) لكل النماذج الممكنة، ومن ثم يمكن إختيار النموذج الذي يكون له أقل قيمة لمعيار AIC.

هناك **معايير أخري** تستخدم لإختيار النموذج الذي يقترب من النموذج الصحيح من هذه المعايير معيار Bayesian Information Criterion (BIC) ، وفيما يلي توضيح لهذا المعيار.

معيار BIC

قدم Schwarz(1978) معيار BIC ، ويمكن تعريفه كما يلي:

$$BIC= -2 \log L(\hat{\theta}) + p \log (n) \qquad (7)$$

معيار BIC هو تقدير لدالة الإحتمال السابقة للنموذج Posterior Probability of Model . لذلك فإن أقل قيمة لهذا المعيار تشير إلي ان هذا النموذج الذي له أقل قيمة BIC ، هو النموذج المحتمل بدرجة كبيرة more likely أن يكون هو النموذج الأقرب للنموذج الصحيح.

ويستخدم معيار BIC في حالة سوء التوزيع Distributional Misspecification. ويكون في هذه الحالة أفضل من معيارAIC لإختيار النموذج.

قدمت دراسة Chang (2017) **معيار BIC متين** وذلك من خلال الإعتماد علي مقدر **MM** . حيث ان مقدر MM أكثر متانة من مقدر M وله نقطة تحطم عالية.

معيار BIC المتين الذي يعتمد على مقدر MM

يأخذ معيار BIC المتين الصيغة التالية:

$$BIC(MM) = 2 \sum_{i=1}^{n} \rho_d (r_i) + p \log(n) \qquad (8)$$

حيث:

$$r_i = \left(y_i - x_i{}^T \hat{\theta}\right) / \hat{\sigma}$$

$\hat{\theta}$: مقدر MM

تمثل ρ_d دالة وزن Tukey's biweight function

وقد قدم Chang *et al*.(2017) أيضاً معيار AIC متين يعتمد علي مقدر MM. ويمكن تعريفه كما يلي:

$$AIC(MM) = 2 \sum_{i=1}^{n} \rho_d (r_i) + \alpha p \qquad (9)$$

- هناك **معايير أخري مثل معامل التحديد** Coefficient of Determination (R^2) وقد قام Saleh(2014) بتقديم **معامل تحديد متين** يعتمد علي مقدر LTS وهو $(R^2{}_{LTS})$ وتم من خلاله التوصل إلي النموذج الأقرب للنموذج الصحيح مقارنة بمعامل التحديد التقليدي (R^2). يتم إختيار النموذج الذي له أكبر قيمة لمعامل التحديد لأن ذلك يكافي تدنية متوسط مربعات البواقي. فقد ذكر Saleh(2015) ان معامل التحديد التقليدي (R^2) الذي يعتمد علي مقدر lS يكون كما يلي:

$$R^2 = 1 - \frac{SSE}{SST} \qquad (10)$$

حيث:

$$SSE = \sum r^2{}_i$$

$$SST = \sum (y_i - \overline{y_i})^2$$

نتيجة لتأثر طريقة المربعات الصغري بالمشاهدات الشاذة ، فإن معامل التحديد الذي يعتمد علي مقدرLS يتأثر أيضاً بالمشاهدات الشاذة. ومن ثم لا يمكن الثقة بنتائج معامل التحديد التقليدي في إختيار المتغيرات. لذلك قدم Saleh(2014) معامل تحديد متين يعتمد علي مقدر LTS ويأخذ الصيغة التالية:

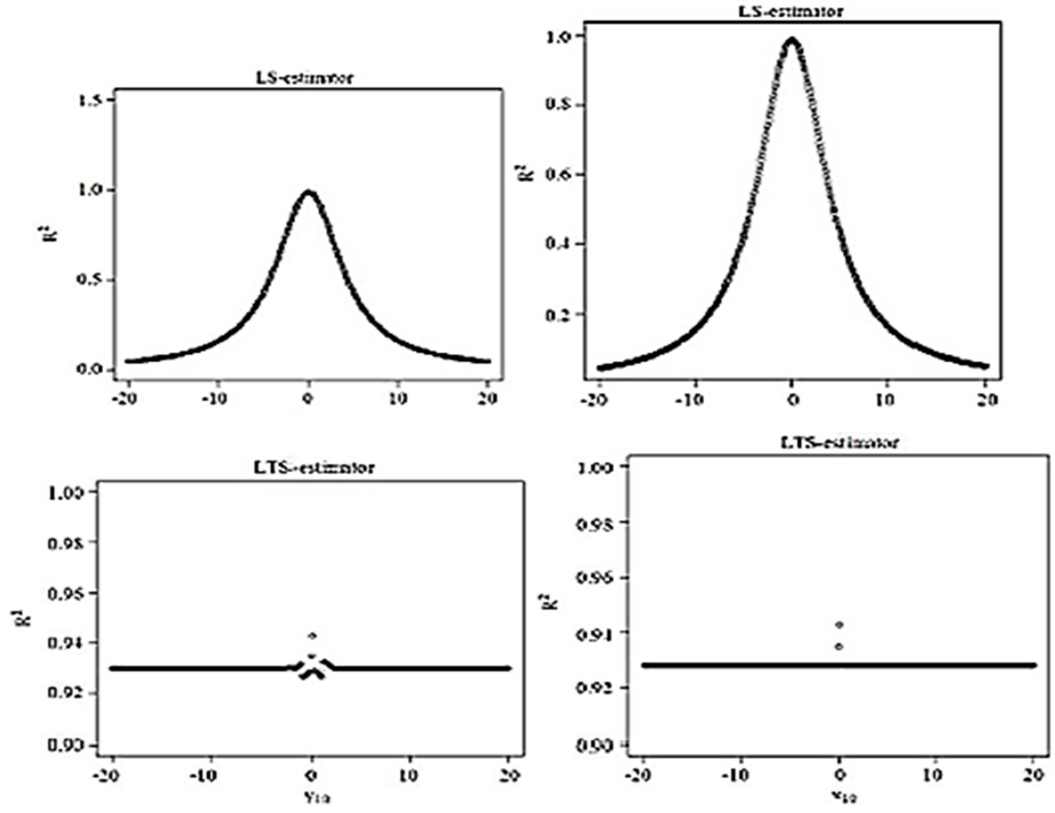
$$R^2{}_{LTS} = \frac{1 - S^2{}_1\,(y_{i-}\,\hat{\alpha} + \, x^t\,\hat{\beta}_{LTS})}{S^2{}_0\,(y_i \, - \hat{\alpha})} \qquad (11)$$

$$S^2{}_1 = \sum_{i=1}^{n} (y_{i-}\,\hat{\alpha} + \, x^t\,\hat{\beta}_{LTS})$$

$$S^2{}_0 = \sum_{i=1}^{n} \, (y_i \, - \hat{\alpha})$$

حيث $(\hat{\beta}_{LTS})$: مقدر LTS

يتضح فيما يلي تأثير التلوث في قيم كل من (x و y) علي معامل التحديد التقليدي ومعامل التحديد المتين . حيث يوضح شكل (٢) قيم (x_{10}) الملوثة والتي تتراوح بين $[-4\,,4.5]$ و قيم (y_{10}) التي تتراوح بين $[-1.5\,,3]$. ويلاحظ من خلال الشكل أن معامل التحديد التقليدي (R^2) شديد التأثر بالتلوث في إتجاه x وإتجاه y حيث أنه يقل كلما إزداد التلوث. بينما يتضح أن معامل التحديد الذي يعتمد علي مقدرLTS $(R^2{}_{LTS})$ أكثر مقاومة للتلوث من معامل التحديد التقليدي.



شكل(٢): تأثير القيم الملوثة (x_{10} , y_{10}) علي معامل التحديد التقليدي (R^2) والمتين $(R^2{}_{LTS})$

المصدر: Saleh(2014)

٢- هدف الدراسة الحالية

بناء علي ماسبق تقديمه من معايير إختيار النموذج التقليدية والمتينة ، يقترح البحث مقارنة كل من معايير (BIC) و (AIC) المتينة والتي تعتمد علي مقدر **MM** مع كل من معيار (AIC_{LTS}) و معيار (AIC_{LMS}) وذلك للإجابة علي التساؤلات التالية :

١- هل تتفق نتائج معيار BIC مع معيار AIC أم تختلف

حيث أنه في دراسة Saleh(2014) تم مقارنة معيار (AIC) التقليدي مع كل من (AIC_{LTS}) و(AIC_{LMS}) ولم يطبق معيار(BIC) التقليدي أو المتين .

٢- هل يتفق معيار AIC الذي يعتمد علي مقدر MM مع نتائج معايير (AIC_{LTS}) و(AIC_{LMS}). وذلك لأن الدراسة لم تقارن معيار(AIC) الذي يعتمد علي مقدر (MM) مع معيار AIC الذي يعتمد علي المقدرات المتينة الأخري.

٣- الدراسة التطبيقية

سيتم التوصل إلي نتائج الدراسة من خلال التطبيق علي بيانات مشهورة تتضمن مشاهدات شاذة في جانب (x) وجانب (y) لكي تعكس مدي مقاومة المقدر ، ومن ثم يتم تقييم اداء المعايير التقليدية والمتينة. سيتم التوصل إلي نتائج الدراسة من خلال إستخدام لغة البرمجة R

٣-١بيانات الدراسة

- سيتم إستخدام بيانات **Stackloss Data** وهي بيانات تتضمن ثلاثة متغيرات تفسيرية وهي: ("x_1 " Air flow) و ("x_2 " Water Temperature) و ("x_3 " Acid Conc) والمتغير التابع (" y " Stackloss) . وتتكون البيانات من ٢١ مشاهدة يوجد بها اربعة مشاهدات شاذة في جانب (y) وهي (١ ، ٣ ، ٤ ، ٢١) ومشاهدات شاذة في جانب (x) (القيم الرافعة) وهي (١ ، ٢ ، ٣ ، ٢١). وقد تم التطبيق علي هذه البيانات في دراسة Saleh(2014).
- كذلك سيتم تطبيق المعايير المتينة علي بيانات **Wood Gravity Data** وقد تم إختيار هذه البيانات ، لانها تتضمن عدد أكبر من المتغيرات التفسيرية مقارنة ببيانات **Stackloss Data** وبالتالي يكون هناك إمكانية لتقييم اداء **المعايير المتينة** في ظل وجود العديد من المتغيرات التفسيرية. حيث تتضمن البيانات خمسة متغيرات تفسيرية (x_1, x_2, x_3, x_4, x_5) والمتغير التابع (y) .

٣-٢ نتائج الدراسة

أولاً: توضيح للنتائج التي توصلت إليها **دراسة Saleh(2014)** .

تم من خلال الدراسة تطبيق معيار AIC التقليدي الذي يعتمد علي مقدر LS ومعيار AIC المتين الذي يعتمد علي المقدرات M و LTS و LMS و S ، وذلك علي بيانات **Stackloss** وقد تم التوصل إلي النتائج التالية:

١- أقل قيمة لمعيار AIC **المتين** الذي يعتمد علي مقدرات LTS و S و LMS خاصة بالنموذج الذي يتضمن المتغيرات x_1 و x_2 وهذا النموذج الأقرب للنموذج الصحيح ، وذلك لان قيمة p.value للمتغيرين (x_1 و x_2) أقل من ٥% مما يشير إلي معنوية المتغيرين.
٢- أقل قيمة لمعيار AIC **التقليدي** خاصة بالنموذج الذي يتضمن ثلاثة متغيرات تفسيرية (x_1 و x_2 و x_3) . وأقل قيمة لمعيار AIC الذي يعتمد علي مقدر M خاصة بالنموذج الذي يتضمن المتغير (x_2) فقط .

- بالتالي إتضح من خلال نتائج دراسة Saleh(2014) انه قد تم التوصل للنموذج الذي يقترب من النموذج الصحيح عند إستخدام المقدرات المتينة LMS و LTS . بينما معيار AIC الذي يعتمد علي مقدر M لم نصل من خلاله إلي النموذج الصحيح وذلك لأن مقدر M يتأثر بالقيم الرافعة .

ثانياً: نتائج الدراسة الحالية

أ- نتائج بيانات Stackloss Data

تم إختبار النماذج الآتية :

يتضمن نموذج ١ : المتغير التابع (y) والمتغيرات (x_1 , x_2 , x_3)

نموذج ٢ : المتغير التابع (y) والمتغير المستقل (x_1)

نموذج ٣: المتغير التابع (y) والمتغير المستقل (x_2)

نموذج ٤: المتغير التابع (y) والمتغيرين (x_3)

نموذج ٥: المتغير التابع (y) والمتغيرين (x_1 , x_2)

نموذج ٦: المتغير التابع (y) والمتغيرين (x_2 , x_3)

نموذج ٧ : المتغير التابع (y) والمتغيرات (x_1 , x_3)

تتضح النتائج من خلال جدول (٣) و (٤) حيث توضح الجداول قيم معايير AIC و BIC التقليدية والمتينة التي تعتمد علي مقدر MM للنماذج السابق ذكرها.

ويمكن توضيح النتائج فيما يلي:

١- أقل قيمة لمعيار BIC **التقليدي** خاصة بالنموذج الأول الذي يتضمن المتغيرات (x_1 , x_2 , x_3) وهذا النموذج ليس هو النموذج الأقرب للنموذج الصحيح وذلك يتضح من خلال قيمة P value للمتغير (x_3) والتي تشير إلي أن المتغير غير معنوي.

٢- أقل قيمة لمعيار **BIC المتين** الذي يعتمد علي مقدر **MM** خاصة بالنموذج الخامس الذي يتضمن المتغيرين (x_1 , x_2) وهذا يتفق مع نتائج دراسة Saleh(2014) التي قامت بتطبيق معيار AIC و معيار R^2 الذي يعتمد كل منهم علي مقدر **LTS** .

٣- أقل قيمة لمعيار **AIC المتين** الذي يعتمد علي مقدر **MM** خاصة بالنموذج الخامس الذي يتضمن المتغيرين (x_1 , x_2).

٤- تم ملاحظة أن قيمة معيار AIC الذي يعتمد علي مقدرMM للنموذج الذي تم إختياره ،أقل من قيمة معيار BIC الذي يعتمد علي مقدر MM ، ويحدث ذلك عندما يكون
(حجم العينة $\leq$ ٨).

جدول(٣): نتائج معايير AIC و BIC و R^2 التقليدية

Model	AIC	BIC	R^2
1	**114.57**	**119.79**	**0.90**
2	115.71	120.80	0.83
3	131.44	134.58	0.73
4	159.99	162.08	0.16
5	158.33	161.46	0.89
6	124.07	128.24	0.83
7	133.69	137.27	0.76
8	122.73	125.87	0.87

جدول(٤): نتائج معايير AIC و BIC المتينة التي تعتمد علي مقدر (MM)

Model	AIC(MM)	BIC(MM)
1	113.8	119.2
2	114.5	120.7
3	131.5	134.6
4	160.0	163.1
5	**112.3**	**117.5**
6	124.4	128.6
7	133.1	137.3
8	161.2	163.3

ب- نتائج بيانات Wood Gravity

تم إختيار بيانات Wood Gravity لأنها تتضمن عدد أكبر من المتغيرات التفسيرية مقارنة ببيانات Stackloss Data، وبالتالي سيتم إختبار عدد أكبر من النماذج لهذه البيانات وهو ٣١ نموذج، وتتضح النماذج فيما يلي:

يتضمن نموذج ١: المتغير التابع (y) والمتغير المستقل (x_1)

نموذج ٢: المتغير التابع (y) والمتغير المستقل (x_2)

نموذج ٣: المتغير التابع (y) والمتغيرين (x_3)

نموذج ٤: المتغير التابع (y) والمتغيرين (x_4)

نموذج ٥: المتغير التابع (y) والمتغيرين (x_5)

نموذج ٦ : المتغير التابع (y) والمتغيرات (x_1 , x_2)

نموذج ٧: المتغير التابع (y) والمتغيرات (x_1 , x_3)

نموذج ٨ : المتغير التابع (y) والمتغيرات (x_1 , x_4)

نموذج ٩: المتغير التابع (y) والمتغيرات (x_1 , x_5)

نموذج ١٠: المتغير التابع (y) والمتغيرات (x_2 , x_3)

نموذج ١١: المتغير التابع (y) والمتغيرات (x_2 , x_4)

نموذج ١٢: المتغير التابع (y) والمتغيرات (x_2 , x_5)

نموذج ١٣: المتغير التابع (y) والمتغيرات (x_3 , x_4)

نموذج ١٤: المتغير التابع (y) والمتغيرات (x_3 , x_5)

نموذج ١٥: المتغير التابع (y) والمتغيرات (x_4 , x_5)

نموذج ١٦: المتغير التابع (y) والمتغيرات (x_1 , x_2 , x_3)

نموذج ١٧: المتغير التابع (y) والمتغيرات (x_1 , x_2 , x_4)

نموذج ١٨: المتغير التابع (y) والمتغيرات (x_1 , x_2 , x_5)

نموذج ١٩: المتغير التابع (y) والمتغيرات (x_1 , x_3 , x_4)

نموذج ٢٠: المتغير التابع (y) والمتغيرات (x_1 , x_3 , x_5)

نموذج ٢١: المتغير التابع (y) والمتغيرات $(x_1 \ , \ x_4 \ , x_5)$

نموذج ٢٢: المتغير التابع (y) والمتغيرات $(x_2 \ , x_3 \ , x_4)$

نموذج ٢٣: المتغير التابع (y) والمتغيرات $(x_2 \ , x_3 \ , x_5)$

نموذج ٢٤: المتغير التابع (y) والمتغيرات $(x_2 \ , \ x_4 \ , x_5)$

نموذج ٢٥: المتغير التابع (y) والمتغيرات $(x_3 \ , x_4 \ , \ x_5)$

نموذج ٢٦: المتغير التابع (y) والمتغيرات $(x_1 \ , x_2 \ , x_3 \ , x_4)$

نموذج ٢٧: المتغير التابع (y) والمتغيرات $(x_1 \ , x_2 \ , x_3 \ , x_5)$

نموذج ٢٨: المتغيرالتابع (y) والمتغيرات $(x_1 \ , x_2 \ , x_4 \ , x_5)$

نموذج ٢٩: المتغير التابع (y) والمتغيرات $(x_1 \ , x_3 \ , x_4 \ , x_5)$

نموذج ٣٠: المتغير التابع (y) والمتغيرات $(x_2 \ , x_3 \ , x_4 \ , x_5)$

نموذج ٣١: المتغير التابع (y) والمتغيرات $(x_1 \ , x_2 \ , x_3 \ , x_4 \ , x_5)$

تتضح النتائج من خلال جدول (٥) حيث يوضح الجدول قيم معايير AIC و BIC المتينة التي تعتمد علي مقدر MM في ظل عدد أكبر من النماذج التي يتم إختبارها.

جدول (٥): نتائج معايير AIC و BIC المتينة التي تعتمد علي مقدر (MM) لبيانات Wood Gravity

Model	AIC(MM)	BIC(MM)
16	**-87.4**	**-82.4**
27	-86.6	-80.6
26	-85.7	-79.7
6	-84.5	-78.4
31	-82.4	-77.5
18	-81.9	-77.0
17	-80.8	-75.8
28	-80.0	-74.1
29	-77.2	-71.2
8	-75.7	-70.9
15	-74.9	-70.8
22	-74.8	-70.7
13	-74.5	-69.5
7	-74.5	-69.5
9	-73.5	-69.0
10	-73.0	-68.7
19	-72.9	-68.6
25	-71.7	-67.6
21	-70.9	-67.4
11	-70.6	-66.9
4	-70.4	-66.1
24	-70.1	-66.0
2	-69.7	-65.7
1	-69.4	-65.4
30	-69.3	-65.3
12	-69.0	-64.9
20	-68.9	-64.0
5	-68.0	-63.4
23	-66.4	-63.0
14	-64.4	-60.4
3	-61.6	-58.6

- يتضح من خلال جدول (٥) أن أقل قيمة لمعيار (AIC) و معيار (BIC) المتين الذي يعتمد كل منهم علي مقدر (MM) ، خاصة بالنموذج (16) الذي يتضمن المتغيرات التفسيرية

(x_1, x_2, x_3) وقد إتضح من خلال قيمة (p- value) أن هذه المتغيرات معنوية. وبالتالي فقد تم من خلال إستخدام مقدر (MM) المتين التوصل إلي إختيار النموذج الصحيح.

و قد تحقق ذلك في ظل بيانات تتضمن عدد كبير من المتغيرات التفسيرية وكذلك في ظل بيانات تتضمن عدد أقل من المتغيرات التفسيرية ، وبالتالي عدد أقل من النماذج الممكنة التي يتم إختبارها.

الخلاصة

يمكن القول أن معيار AIC ومعيار (R^2) الذي يعتمد كل منهم علي مقدرات **LTS و LMS** المتينة التي قدمتها دراسة Saleh(2014) تتفق نتائجهما (من حيث إختيار النموذج الصحيح) مع معايير AIC و BIC التي تعتمد علي مقدر **MM** (وفقاً لنتائج الدراسة الحالية)، وذلك لأن كل من LTS و LMS و مقدر MM مقدرات متينة للمشاهدات الشاذه في جانب x وجانب y. بينما لم يتم التوصل إلي النموذج الأقرب للنموذج الصحيح من خلال الإعتماد علي مقدر **M** وذلك لانه يتأثر بالمشاهدات الشاذة في جانب x . وذلك وفقاً للنتائج التي توصلت إليها دراسة Saleh(2014) لمعيار AIC الذي يعتمد علي مقدر(LMS) و مقدر(LTS) ومقدر (M).

المراجع

[1] Chang.(2017)."Essays on Robust Model Selection and Model Averaging for Linear Models".*doctoral thesis* , Australian National University.

[2] R.D.Martin,"Robust Estimation of Autoregressive Models", *in direction in time series*.pp:228-254.

[3] H.Akaike,(1992)."Information Theory and an extension of the Maximum Likelihood Principle," *In Breakthroughs in Statistics*,pp:610-624.

[4] K.Tharmaratnam and G,Claeskens,(2013). " A comparison of Robust versions of the AIC based on M-, S-, and MM –estimators,".*Statistics*. Vol.47.no.1.pp:216-235.

[5] P.J.Rousseeuw.(1984)."Least Median of Squares regression". *Journal of the American statistical Association* .vol.79.no.388.pp:871-880.

[6] P.J.Huber.(2011)." Robust Statistics". *Springer* , New York , USA.

[7]Schwarz, Gidean E.(1978)."Estimating the Dimension of a model". *Annals of Statistics*.vol.2.pp:461-464.

[8]Sh.Saleh(2014)." Robust AIC with high breakdown Scale Estimate". *Journal of Applied Mathematics*.pp:1-7.

[9]Sh.Saleh(2014)."Model Selection VIA Robust version of R-Squared Criterion". *Journal of Applied Mathematics*.vol.13.pp:414-420.