



## Journal of Economics and Administrative Sciences (JEAS)



Available online at <http://jeasiq.uobaghdad.edu.iq>

### مقترح اسلوب IRWs الحصين لتقدير أنموذج الانحدار المقسم لحركة حمولة قاع نهر دجلة ونقطة التغيير لكمية تصريف المياه عند مدينة بغداد

أ.م.د. عمر عبد المحسن علي  
جامعة بغداد / كلية الإدارة والاقتصاد / قسم  
الاحصاء

الباحث / محمد أحمد عباس  
ديوان الوقف السني / دائرة التعليم الديني و  
الدراسات الاسلامية

[\(dromar72@coadec.uobaghdad.edu.iq\)](mailto:dromar72@coadec.uobaghdad.edu.iq)

[\(mohammad.stat.145@gmail.com\)](mailto:mohammad.stat.145@gmail.com)

Received:11/3/2020

Accepted :3/5/2020

Published :August / 2020

هذا العمل مرخص تحت اتفاقية المشاع الابداعي نسب المُصنّف - غير تجاري - الترخيص العمومي الدولي 4.0

[Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc-sa/4.0/)



### مستخلص البحث:

يتكون الانحدار المقسم من عدة أقسام تفصل بينها نقاط انتماء مختلفة، فتظهر حالة عدم التجانس الناشئة من عملية فصل الأقسام ضمن عينة البحث. ويهتم هذا البحث في تقدير موقع نقطة التغيير بين الأقسام وتقدير معالم الأنموذج، واقتراح طريقة تقدير حصينة ومقارنتها مع بعض الطرائق المستعملة في الانحدار الخطي المقسم. وقد تم استعمال أحد الطرائق التقليدية (طريقة Muggeo) لإيجاد مقدرات الإمكان الأعظم بالأسلوب التكراري للأنموذج ونقطة التغيير معاً، واستعمال أحد الطرائق الحصينة (طريقة IRWm) والتي تعتمد على استعمال تقنية M-estimator الحصين في أسلوب التقسيم وباستعمال دالة الوزن Tukey. وتكمن مساهمتنا في هذا البحث في اقتراح استعمال تقنية S-estimator الحصينة وباستعمال دالة الوزن Tukey، للحصول على طريقة حصينة ضد حالات انتهاك شرط التوزيع الطبيعي للأخطاء العشوائية أو تأثير القيم الشاذة، وستدعى هذه الطريقة IRWs. وقد تم تطبيق الطرائق المذكورة آنفاً على مجموعة بيانات حقيقية متعلقة بحمولة قاع نهر دجلة/ مدينة بغداد كمتغير استجابة وكمية تصريف المياه كمتغير توضيحي. وقد أظهرت نتائج المقارنة أفضلية الطريقة المقترحة.

**المصطلحات الرئيسية للبحث:** الانحدار الخطي المقسم، نقطة التغيير، المعايير الخطية، مقدر M، مقدر S، حركة حمولة القاع، هيدرولوجية المسطحات المائية.

\*البحث مستل من رسالة ماجستير

## 1- المقدمة Introduction

إن أنموذج الانحدار الخطي المقسم (Segmented linear regression model) يعد من نماذج الانحدار التي تكون فيها العلاقة بين متغير الاستجابة و المتغير التفسيري خطية بشكل جزئي، وهو من الأدوات الاحصائية التي تستعمل لنمذجة الظواهر التي تمر بمراحل انتقالية، ويمكن تمثيلها بخطين مستقيمين أو أكثر وتتصل هذه الخطوط فيما بينها بقيم غير معروفة ويشار عادة إلى هذه القيم بنقاط الربط (join point) أو نقاط التغيير (change point) أو نقاط الكسر (break points)، وهناك عدة تسميات أخرى لهذا لأنموذج منها أنموذج الخط المكسور ([10]broken line model)، وأنموذج إنحدار نقطة الربط ([16] joinpoint regression model)، وأنموذج الخطي القطعي ([17]piecewise linear model).

ويمكن استعمال هذا الانموذج عندما يفشل أنموذج الانحدار الخطي الاعتيادي في تمثيل البيانات و قد يمكن أن يساعد هذا الانموذج في تحديد أفضل شكل دالي في حالة أنموذج الانحدار غير الخطي، ويساعد هذا الانموذج على إختيار أفضل مجموعة من الخطوط أو الأقسام المتصلة للحصول على أفضل شكل لدالة غير خطية مناسبة للبيانات غير المنتظمة أو ذات المجاميع المختلفة في المتغير التوضيحي أو التي يكون فيها تغيير في نمطها، وهناك العديد من التطبيقات يمكن أن تواجه مثل هكذا بيانات مثل التطبيقات في البحوث الطبية، والبيولوجيا، والبيئة، والدراسات المالية و الدراسات الهيدرولوجية.

هناك العديد من الباحثين في الادبيات قاموا بانجاز العديد من الأعمال لتقدير نقطة التغيير و المعلمات الخاصة بأنموذج الانحدار الخطي المقسم ومهنم، ([11][12]Quandt,1958,1960) ([9]Hudson,1966), ([6]Feder,1975), ([7]Gallant and Fuller,1973) معظم هذه الطرائق على أسلوب البحث الشبكي لتقدير معاملات الانحدار ونقاط التغيير وتكون هذه التقديرات معقولة ويعد هذا النهج مكلف حسابياً، وقد اقترح الباحث ([10]Mugge,2003) طريقة تقدير ذكية لهذا الأنموذج باستعمال تقنية خطية بسيطة وهي تقنية إعادة المعايرة الخطية (Linear reparametrization technique)، إذ تسمح هذه الطريقة باستدلال جميع معلمات الأنموذج ونقطة التغيير بنفس الوقت وبطريقة فعالة من الناحية الحسابية.

وتعد الطرائق المذكورة أعلاه من الطرائق الكلاسيكية التي تعمل بشكل جيد عندما تكون حالة البيانات طبيعية ولاحتوي على قيم شاذة أو لا يكون توزيع البواقي غير طبيعي، ولكن غالباً ماتكون للبيانات الحقيقية قيم شاذة أو توزيع غير طبيعي للبواقي والتي يمكن أن تؤثر في دقة تقدير المعلمات، لذا كان من الضروري البحث عن أسلوب بديل لمعالجة هذه المشكلة لذا فقد قدم الباحثان ([17]F. Zhang and Q. Li,2017) اقتراح المقدر المستند إلى الرتب الحصين (robust rank-based estimator) لإنحدار الخط المنحني (bent linear regression) مع نقطة تغيير غير معروفة وتكمن الفكرة الأساسية في عملية التقدير في استبدال المجموع المتبقي للمربعات في عمليات التقسيم لـ (Mugge,2003) مع دالة تشتت الرتب (rank dispersion) في الانحدارات المستندة إلى الرتب لتحقيق الرصانة ضد القيم المتطرفة والأخطاء ذات الاطراف المرتفعة، وقد قدم الباحثان ([4]O.A.Ali, M.A.Abbas,2019) مقدر حصين جديد لتقدير نقطة التغيير في أنموذج الانحدار الخطي المقسم وذلك باستعمال منهجية (M-Estimator) مع ثلاث دوال وزن معروفة (Huber و Tukey و Hampel) والتي تم دمجها مع نهج (Mugge) الكلاسيكي للحصول على مزيد من الحصانة لتقدير نقطة التغيير و معلمات الانحدار في ذات الوقت، و بين الباحثان أفضلية الطريقة الجديد باستعمال دالة الوزن (Tukey). وقد قام الباحثان ([1]Sukru A., Birdal S.,2020) بتطوير نسخة معدلة حصينة لطرائق تقدير نقطة التغيير للطرائق الكلاسيكية التي قدمها ([9] Hudson 1966) و ([10] Mugge2003)، حيث أستخدم الباحثان منهجية تعديل الإمكان الأعظم (MML) عندما يتم توزيع شروط الخطأ لأنموذج الانحدار الخطي ثنائي الاقسام (two-segmented linear regression model) بشكل مستقل ومتماثل على أنه متماثل طويل الذيل، وبذلك تكون مقدرات هذه النسخة المعدلة أكثر كفاءة وقوة من الطرائق الكلاسيكية.

وسيتم في هذا البحث إقتراح طريقة تقدير حصينة لتقدير نقطة التغيير ومعلمات الانموذج، للتغلب على مشكلة وجود بعض القيم الشاذة أو في حالة أن الخطأ العشوائي للأنموذج لايتبع التوزيع الطبيعي، وذلك باستعمال تقنية (S-estimator) الحصينة ودمجها مع أسلوب التجزئة لنهج (Mugge)، إذ تعد تقنية (S-estimator) الحصينة من مقدرات الانحدار المرتبطة بمقياس M (M-scales) والتي تعتمد على مقياس البواقي لطريقة (M-estimator) الحصينة.

## 2- الطرائق النظرية

## 1-2 وصف أنموذج الانحدار الخطي المقسم

يتم وصف العلاقة بين النتيجة والتنبؤات بنماذج قطعية مستمرة مع خطوط انحدار لها ميول مختلفة متصلة بنقاط تغيير غير معروفة، وسيقتصر اهتمامنا على أنموذج الانحدار الخطي المقسم بنقطة تغيير واحدة ومعاملات إنحدار مختلفة في أماكن مختلفة ولكنه مستمر عند تلك النقطة، ويتم نمذجتها كالآتي [8:pp.569]:

$$f(x_i) = \begin{cases} \beta_0 + \beta_1 x_i & x_1 \leq x_i \leq \tau \\ \beta_{00} + \beta_2 x_i & \tau < x_i \leq x_n \end{cases} \dots (1)$$

إذ أن:  $(\tau)$  هي نقطة التغيير غير معلومة .

$x_1$  و  $x_n$ : حدود قيم  $x_i$  المعلومة.

معلمات الإنموذج (الميل و الحد الثابت) على الجانب الأيسر والأيمن من نقطة التغيير على التوالي وهي تمثل قيماً مجهولة يجب تقديرها.

وفي حالة افتراض الأستمرارية بين الأقسام عند نقطة التغيير فسيتم فرض القيد الآتي:

$$\beta_0 + \beta_1 \tau = \beta_{00} + \beta_2 \tau, \dots (2)$$

وإذا تم افتراض أن الأنموذج مستمر فإن الصيغة رقم (1) يمكن أن تكتب بالصيغة الآتية:

$$f(x_i) = \beta_0 + \beta_1 x_i + \Delta_1(x_i - \tau)I(x_i < \tau), \dots (3)$$

إذ أن  $(\beta_0, \beta_1)$ : هي معلمات (الميل و الحد الثابت) للإنموذج قبل نقطة التغيير .

$\Delta_1$ : هي معلمة الفرق بين الميل قبل نقطة التغيير والميل بعد نقطة التغيير أي:  $\Delta_1 = \beta_2 - \beta_1$ .

و أن  $I(x_i > \tau) = 1$  إذا كانت  $x_i > \tau$  صحيحة و تساوي 0 إذا كانت خاطئة.

ويتم حساب الحد الثابت للإنحدار في القسم الذي يلي نقطة التغيير وفق الصيغة الآتية:

$$\beta_{00} = \beta_0 + \beta_1 \tau - \beta_2 \tau, \dots (4)$$

## 2-2 طرائق التقدير Estimation methods

## 1-2-2 طريقة (Muggeo) لملائمة الانحدار الخطي المقسم:

أن طريقة الإمكان الأعظم (maximum likelihood) تُعد واحدة من أهم طرائق التقدير في الإحصاء التي تستعمل في تقدير الإنحدار الخطي الأعتيادي، ويتم الحصول على تقديرات الإمكان الأعظم (ML) لجميع معلمات أنموذج الإنحدار الخطي المقسم وبما في ذلك تقدير نقطة التغيير بأسلوب تكراري مقترح من قبل الباحث ([10]Vito M. R. Muggeo, 2003)، والذي أقرح فيه تقنية بديلة لتقدير أنموذج الإنحدار المقسم، لتحويل المشكلة إلى إطار خطي، و تعد هذه الطريقة بسيطة من ناحية المفاهيم ولكنها أيضاً عامة ومناسبة للعديد من الحالات، ويتم حساب تقدير المعلمات المتعددة لأنموذج بسهولة ولكن فقط يتم طلب قيم نقاط الكسر الابتدائية [10:pp3057].

ولوضع صورة العلاقة الخطية للإنحدار المقسم بين الاستجابة  $y_i$  والمتغير التوضيحي  $x_i$  يتم استعمال الأنموذج رقم (3) والذي يعاد صاغته بالصيغة الآتية:

$$y_i = \beta_0 + \beta_1 x_i + \Delta_1(x_i - \tau)_+ \dots (5)$$

إذ أن:  $(x_i - \tau)_+ = (x_i - \tau) \times I(x_i > \tau)$

ويلاحظ هنا أن لوغاريتم دالة الامكان غير قابلة للإشتقاق عند النقطة  $x_i = \tau$ ، ويتم حل هذه المشكلة بتقنية إعادة المعايير (reparameterization) الخطي بواسطة الدالة الخطية الآتية:

$$h(x; \tau) \approx h(x; \tau^{(0)}) + (\tau - \tau^{(0)})h'(x; \tau^{(0)}) \dots (6)$$

وأن  $\tau^{(0)}$ : هي قيمة ابتدائية لنقطة الكسر ضمن مجال قيم  $x$ .

$h'(\cdot; \tau^{(0)})$ : هي المشتقة الأولى لـ  $h(\cdot)$  في  $\tau^{(0)}$ .

إذ إن الفكرة الأساسية للباحث (Muggeo) هي تقريب المصطلح غير الخطي  $(x - \tau)_+$  إلى الشكل الخطي من خلال (Taylor expansion) من الرتبة الأولى حول قيمة ابتدائية معروفة  $\tau^{(0)}$ ، بحيث يمكن إعادة معايرة نقطة التغيير  $\tau$  كمصطلح معامل في أنموذج خطي مستمر وتقديره مع معاملات الانحدار الأخرى كما هو الحال في الانحدار القياسي الأعتيادي [17:pp.6].

وبالتالي يمكن إعادة صياغة الحد  $(x_i - \tau)_+$  في الصيغة (5) باستعمال الدالة الخطية (6) وفق الصيغة الآتية :

$$(x_i - \tau)_+ = (x_i - \tau^{(0)})_+ + (\tau - \tau^{(0)})(-1)I(x_i > \tau^{(0)}) \quad \dots (7)$$

وبهذا يصبح النموذج (5) وفق الصيغة التالية :

$$y_i = \beta_0 + \beta_1 x_i + \Delta_1(x_i - \tau^{(0)})_+ + \gamma(-1)I(x_i > \tau^{(0)}) \quad \dots (8)$$

إذ أن :  $\gamma = \Delta_1(\tau - \tau^{(0)})$ .

وبالتالي يتم ملائمة نموذج الإنحدار الخطي المقسم بالصيغة الخطية مع المتغيرات الإضافية وفق الصيغة الآتية:

$$y_i = \beta_0 + \beta_1 x_i + \Delta_1 u_i + \gamma v_i, \quad \dots (9)$$

إذ إن  $u_i$  و  $v_i$  هي متغيرات إضافية تحسب وفق الصيغة :

$$u_i = (x_i - \tau^{(0)})_+ \quad \text{and} \quad v_i = -I(x_i > \tau^{(0)}) \quad \dots (10)$$

وأن النموذج الجديد (9) هو نموذج خطي إعتيادي يعطي تقديرات الامكان الاعظم (ML) لجميع معلمات النموذج بما في ذلك نقطة التغيير، ويمكن إيجاد هذه المقدرات وفق خوارزمية تكرارية بسيطة يمكن تلخيصها بالخطوات الآتية [10:pp.3059-3061]:

- 1- تثبيت قيمة  $\tau^{(s)}$  الابتدائية وهي قيمة تخمينية ضمن مجال قيم  $x_i$ .
- 2- يتم حساب المتغيرات الإضافية  $u_i^{(s)}$ ,  $v_i^{(s)}$  وفق الصيغة (10).
- 3- يتم ملائمة النموذج مع المتغيرات الإضافية  $u_i^{(s)}$  و  $v_i^{(s)}$  وفق الصيغة (9).
- وبذلك يتم الحصول على تقديرات الامكان الاعظم (ML) في كل دورة لجميع المعلمات.
- 4- تحديث نقطة التغيير وفق الصيغة :

$$\hat{\tau} = \frac{\hat{\gamma}}{\hat{\Delta}_1} + \tau^{(s)} \quad \dots (11)$$

إذ إن  $(\hat{\gamma}, \hat{\Delta}_1)$  هي معلمات في النموذج (9) تم تقديرها في الخطوة 3.

5- تكرر الخطوات السابقة 2 إلى 4 بعد استبدال قيمة  $\tau^{(s)}$  الابتدائية بالنقطة  $\hat{\tau}$  المحدثة في الخطوة 4 إلى s من التكرارات حتى يتم التقارب (أي  $\hat{\gamma} \approx 0$ ).

ويتم إيقاف الخوارزمية عندما يتم التقارب وذلك لأن نقطة التغيير  $\hat{\tau}$  لن تتحدث في الخطوة 4 عندما  $\hat{\gamma} \approx 0$  وستبقى كما هي، وبالتالي يتم الحصول على تقديرات الامكان الاعظم ML لجميع المعلمات في النموذج بما في ذلك نقطة التغيير  $\tau$ ، عند التقدير التقريبي  $\hat{\tau}$ ، أي  $\tau^{(s)} = \hat{\tau}$ .

## 2-2-2 طريقة مقدر M الموزون التكرارية الحصينة Robust IRWm-estimator

إن طريقة (IRWm-estimator) هي أحد الطرائق الحصينة المقترحة من قبل الباحثان ([4]O.A.Ali, M.A.Abbas,2019) والتي تعتمد على استعمال تقنية (M-estimator) الحصينة ودمجها مع نهج (Muggeo) الكلاسيكي للحصول على المزيد من الحصانة لتقدير نقطة التغيير و معلمات الإنحدار في ذات الوقت، ويعتمد نهج الحصانة في هذه الطريقة على دالة الخسارة (loss function)  $P(\cdot)$  والتي تكون قليلة التباين عند البواقي الشاذة للغاية بدلا من مجموع المربعات الصغرى الاعتيادية أي  $(\sum P(e_i) = \sum e_i^2)$  [4:pp.393-394]:

وبتصغير دالة الخسارة  $\sum P(e_i)$  يمكن الحصول على التقديرات الحصينة للمعلمات وكما موضحة في الخطوات الآتية [4:pp.394-395]:

1- تهيئة معلمات البدء  $(\hat{\beta}_0^{(s)}, \hat{\beta}_1^{(s)}, \hat{\Delta}_1^{(s)}, \hat{\gamma}^{(s)}) = (\hat{\beta}_0^{(s)}, \hat{\beta}_1^{(s)}, \hat{\Delta}_1^{(s)}, \hat{\gamma}^{(s)})$  و وضع قيمة أولية صغيرة للمعلمة  $\hat{\gamma}^{(s)}$  مثل 0.01.

2- تقدير المعلمات  $(\hat{\beta}_0^{(s+1)}, \hat{\beta}_1^{(s+1)}, \hat{\Delta}_1^{(s+1)}, \hat{\gamma}^{(s+1)})$  من خلال استعمال تقنية مقدر M الحصين للنموذج (9) في طريقة (Muggeo) بعد استعمال تقنية إعادة المعايرة الخطي على النحو الآتي:

a- إيجاد المقدر الابتدائي  $(\hat{\beta}_0^{(r)}, \hat{\beta}_1^{(r)}, \hat{\Delta}_1^{(r)}, \hat{\gamma}^{(r)}) = \hat{\theta}^{(r)}$  كأن يكون مقدر المربعات الصغرى الاعتيادي للنموذج (9) في طريقة Muggeo والمعاد كتابتها:

$$y_i = \beta_0 + \beta_1 x_i + \Delta_1 u_i^{(r)} + \gamma v_i^{(r)} + e_i, \quad \dots (12)$$

b- حساب البواقي  $e_i^{(r)} = y_i - (\beta_0 + \beta_1 x_i + \Delta_1 u_i^{(r)} + \gamma v_i^{(r)})$   
 c- إيجاد المقدرات الحصينة  $\hat{\theta}^{(r+1)} = (\hat{\beta}_0^{(r+1)}, \hat{\beta}_1^{(r+1)}, \hat{\Delta}_1^{(r+1)}, \hat{\gamma}^{(r+1)})$  باستعمال تقنية مقدر M وفق الصيغة الآتية [5:pp56]:

$$D(\hat{\theta}^{(r+1)}) = \arg \min \sum_{i=1}^n p(z_i^{(r)}), \quad \dots (13)$$

$$z_i^{(r)} = \frac{e_i^{(r)}}{\hat{\sigma}_{MAD}} \quad \text{إذ أن :}$$

$$\hat{\sigma}_{MAD} = \frac{\text{median}|e_i^{(r)} - \text{median}(e_i^{(r)})|}{0.6745} \quad \text{وأن :}$$

وباشتقاق المعادلة (13) بالنسبة لـ  $\theta$  ومساواتها بالصفر:

$$\sum_{i=1}^n X \Psi(z_i^{(r)}) = 0 \quad \dots (14)$$

إذ أن  $p' = \Psi, X = (x_i, u_i^{(r)}, v_i^{(r)})$ :

$$w(z_i^{(r)}) = \frac{\Psi(z_i^{(r)})}{(z_i^{(r)})} \quad \text{وأن :}$$

وبالتالي فإن الصيغة (14) تصبح :

$$\sum_{i=1}^n X w(z_i^{(r)}) (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\Delta}_1 u_i^{(r)} + \hat{\gamma} v_i^{(r)})) = 0 \quad \dots (15)$$

أن المعادلة (15) تعرف بالمربعات الصغرى الموزونة (WLS) وأن الحل لهذه المعادلة يعطي تقدير  $\hat{\theta}^{r+1}$  وكما في الصيغة الآتية [5:pp.57]:

$$\hat{\theta}^{r+1} = (X'WX)^{-1}X'WY, \quad \dots (16)$$

إذ أن  $W$  هي مصفوفة قطرية من الدرجة  $n \times n$  وأن عناصر قطرها الرئيسي هو  $w(z_i^{(r)})$

d- يتم تكرار الخطوات من  $a$  إلى  $c$  حتى يتقارب المعامل المقدر [17:pp.7] أي:

$$\| \hat{\theta}^{r+1} - \hat{\theta}^r \|_{\infty} < 10^{-5} \quad \text{إذ أن } \|v\|_{\infty} = \max_j |v_j| \quad \text{لكل } v \in R^q$$

وبعد أن يتم التقارب فإن  $\hat{\theta}^{r+1} = \hat{\theta}^{s+1}$

3- تحديث نقطة التغيير  $\tau$  حسب الصيغة (11) في طريقة (Muggeo) والمعاد كتابتها :

$$\hat{\tau}^{(s+1)} = \hat{\tau}^{(s)} + \frac{\hat{\gamma}^{(s+1)}}{\hat{\Delta}_1^{(s+1)}}, \quad \dots (17)$$

4- يتم تكرار الخطوات السابقة بعد تحديث نقطة التغيير  $\tau$  حتى يتم عملية التقارب حسب معيار (Muggeo) ( $\hat{\gamma} \approx 0$ ) ، والتي عندها لا يوجد أي تحسن في تقدير نقطة التغيير، وبالتالي فإن التقدير

التقريبي  $sth$  يفترض تقدير  $\hat{\theta}^{(s+1)}$  هو تقدير (IRWm-estimator) ، وأن  $\tau^{(s)} = \hat{\tau}$ .

وهناك عدة دوال يمكن أن تستخدم في الدالة  $p$  كأوزان لحل مشكلة التصغير التي وجدت من خلال أخذ مشتقات جزئية بالنسبة لـ  $\hat{\theta}$  ومساواتها بالصفر، وسيتم استعمال دالة (Tukey) التي أثبتت كفاءتها من قبل الباحثان

: ([4]O.A.Ali, M.A.Abbas,2019)

• دالة (Tukey bisquare) [14:pp.6]:

$$p(z) = \begin{cases} \frac{z^2}{2} - \frac{z^4}{2c^2} + \frac{z^6}{6c^4}, & \text{if } |z| \leq c \\ \frac{c^2}{6}, & \text{if } |z| > c \end{cases}, \quad W(z) = \begin{cases} \left(1 - \left(\frac{z}{c}\right)^2\right)^2, & \text{if } |z| \leq c \\ 0, & \text{if } |z| > c \end{cases}$$

إذ إن :  $c = 4.685$

### 2-2-3 الطريقة المقترحة لاستعمال تقنية (S-estimator) الحصينة في الانحدار الخطي المقسم

إن مقدر S هو من مقدرات الانحدار الحصين المرتبطة بمقياس M (M-scales) والتي اقترحها الباحثان ([13]Rousseeuw and Yohai,1984)، والتي تعتمد على مقياس البواقي لطريقة مقدر M الحصين، وتستخدم الانحراف المعياري للبواقي للتغلب على نقاط الضعف (weaknesses) في الوسيط [15:pp.354].

وسيتم استعمال هذا المقدر في تقدير الانحدار الخطي المقسم باستعمال نهج التجزئة للباحث (Muggeo)، وبنفس الأسلوب المتبع في الطريقة السابقة لاستعمال مقدر M الحصين في الفقرة (2-2-2) يتم الحصول على التقدير الحصين للمعلمت عن طريق تصغير دالة الخسارة  $\sum P(e_i)$ ، ويمكن تسمية المقدر العائد بأنه مقدر S الموزون التكراري (IRWs-estimator). ويمكن تلخيص هذه الطريقة وفق الخطوات الآتية:

1- تهيئة معلمت البدء  $(\hat{\beta}_0^{(s)}, \hat{\beta}_1^{(s)}, \hat{\Delta}_1^{(s)}, \hat{\gamma}^{(s)})$  و  $\hat{\tau}^{(s)}$  وضع قيمة أولية صغيرة للمعلمة  $\hat{\gamma}^{(s)}$  مثل 0.01.

2- تقدير المعلمت  $(\hat{\beta}_0^{(s+1)}, \hat{\beta}_1^{(s+1)}, \hat{\Delta}_1^{(s+1)}, \hat{\gamma}^{(s+1)})$  من خلال استعمال تقنية مقدر S الحصين للنموذج (9) في طريقة (Muggeo) بعد استعمال تقنية إعادة المعايرة الخطي على النحو الآتي:

a- إيجاد المقدر الابتدائي  $(\hat{\beta}_0^{(r)}, \hat{\beta}_1^{(r)}, \hat{\Delta}_1^{(r)}, \hat{\gamma}^{(r)})$  كأن يكون مقدر المربعات الصغرى الاعتيادي للنموذج (9) في طريقة Muggeo والمعاد كتابتها بالصيغة:

$$y_i = \beta_0 + \beta_1 x_i + \Delta_1 u_i^{(r)} + \gamma v_i^{(r)} + e_i \quad \dots (18)$$

b- حساب البواقي  $e_i^{(r)} = y_i - (\beta_0 + \beta_1 x_i + \Delta_1 u_i^{(r)} + \gamma v_i^{(r)})$ .

c- إيجاد تقدير المعلمت  $(\hat{\beta}_0^{(r+1)}, \hat{\beta}_1^{(r+1)}, \hat{\Delta}_1^{(r+1)}, \hat{\gamma}^{(r+1)})$  كمقدر حصين باستعمال تقنية مقدر S وفق الصيغة الآتية [15:pp. 354]:

$$D(\hat{\theta}^{(r+1)}) = \arg \min \sum_{i=1}^n p(z_i^{(r)}), \quad \dots (19)$$

$$z_i^{(r)} = \frac{e_i^{(r)}}{\hat{\sigma}_{si}^{(r)}} \quad \text{إذ أن:}$$

ويتم حساب  $\hat{\sigma}_{si}^{(r)}$  وفق الصيغة:

$$\hat{\sigma}_{si}^{(r)} = \begin{cases} \frac{\text{median}|e_i^{(r)} - \text{median}(e_i^{(r)})|}{0.6745}, & \text{if } r = 1 \\ \sqrt{\frac{1}{nk} \sum_{i=1}^n w_i^{(r-1)} e_i^2{}^{(r)}}, & \text{if } r > 1 \end{cases} \quad \dots (20)$$

وإن:  $k = 0.199$

ويتم حساب قيمة الوزن  $w_i^{(r)}$  وباستعمال دالة (Tukey) حسب الصيغة الآتية:

$$w_i^{(r)} = \begin{cases} \left(1 - \left(\frac{z_i^{(r)}}{1.547}\right)^2\right)^2, & \text{if } |z_i^{(r)}| \leq 1.547 \\ 0, & \text{if } |z_i^{(r)}| > 1.547 \end{cases}, \quad r = 1 \quad \dots (21)$$

$$\frac{p(z_i^{(r)})}{z_i^2{}^{(r)}}, \quad r > 1$$

وبعد اشتقاق الصيغة (19) ومساواتها بالصفر وبنفس الأسلوب المتبع في الطريقة السابقة في الفقرة (2-2-2) السابقة يتم إيجاد مقدر S الحصين بطريقة المربعات الصغرى الموزونة (WLS) باستعمال الوزن  $w_i^{(r)}$  حسب الصيغة الآتية [15:pp. 355]:

$$\hat{\theta}^{r+1} = (X'WX)^{-1}X'WY, \quad \dots (22)$$

إذ أن  $W$  هي مصفوفة قطرية من الدرجة  $n \times n$  وأن عناصر قطرها الرئيسي هو  $w_i^{(r)}$

$d$  - يتم تكرار الخطوات من  $a$  إلى  $c$  حتى يتقارب المعامل المقدر أي [17:pp.7]:

$$v \in R^q \text{ لكل } \|v\|_{\infty} = \max_j |v_j| \text{ إذ أن } \|\hat{\theta}^{r+1} - \hat{\theta}^r\|_{\infty} < 10^{-5}$$

وبعد أن يتم التقارب فإن  $\hat{\theta}^{s+1} = \hat{\theta}^{r+1}$

3- تحديث نقطة التغيير  $\tau$  حسب المعادلة (11) في طريقة (Muggeo) والمعاد كتابتها :

$$\hat{\tau}^{(s+1)} = \hat{\tau}^{(s)} + \frac{\hat{\gamma}^{(s+1)}}{\hat{\Delta}_1^{(s+1)}}, \quad \dots (23)$$

وتكرر العمليات السابقة بعد تحديث نقطة التغيير  $\tau$  حتى يتم عملية التقارب وفقاً لمعيار (Muggeo) ( $\hat{\gamma} \approx 0$ ) [10:pp.3061] ، والتي عندها لا يوجد أي تحسن في تقدير نقطة التغيير، وبالتالي فإن التقدير

التقريبي  $\hat{\tau}^{(s+1)}$  هو تقدير (IRWs-estimator)، وأن  $\hat{\tau}^{(s)} = \hat{\tau}$ .

3-2 الخطأ المعياري لنقطة التغيير The standard error for the change point

يمكن استخلاص الاستدلالات على نقطة التغيير  $\hat{\tau}$  عن طريق أساليب التمهيد أو الامكان أو طريقة Wald، وللتمكن من استعمال إحصائيات Wald البسيطة على وجه الخصوص يمكن الحصول على الخطأ المعياري لنقطة التغيير  $\hat{\tau}$  باستعمال التقريب الخطي لنسبة متغيرين عشوائيين (أي طريقة دلتا) وبالتالي فإن صيغة التباين لنقطة التغيير  $\hat{\tau}$  تعطى وفق الصيغة الآتية [17:pp.8]:

$$Var(\hat{\tau}) = \frac{Var(\hat{\gamma}) + Var(\hat{\Delta}_1) \left(\frac{\hat{\gamma}}{\hat{\Delta}_1}\right)^2 + 2 \left(\frac{\hat{\gamma}}{\hat{\Delta}_1}\right) Cov(\hat{\gamma}, \hat{\Delta}_1)}{\hat{\Delta}_1^2}, \quad \dots (24)$$

إذ إن:  $var(.)$  و  $cov(., .)$  تعني التباين و التباين المشترك على التوالي .

وبالتالي فإن الخطأ المعياري لنقطة التغيير ( $SE(\hat{\tau})$ ) في الإنحدار الخطي المقسم يكون وفق الصيغة الآتية :

$$SE(\hat{\tau}) = \sqrt{Var(\hat{\tau})}, \quad \dots (25)$$

وعندما تتقارب الخوارزمية (converges) في أسلوب (Muggeo) فإنه من المتوقع أن تكون  $\hat{\gamma}$  صفر تقريباً وبالتالي فإن الخطأ المعياري لنقطة التغيير لهذه الخوارزميات ببساطة تكون وفق الصيغة [10:pp.3061] :

$$SE(\hat{\tau}) = SE(\hat{\gamma}) / |\hat{\Delta}_1| \quad \dots (26)$$

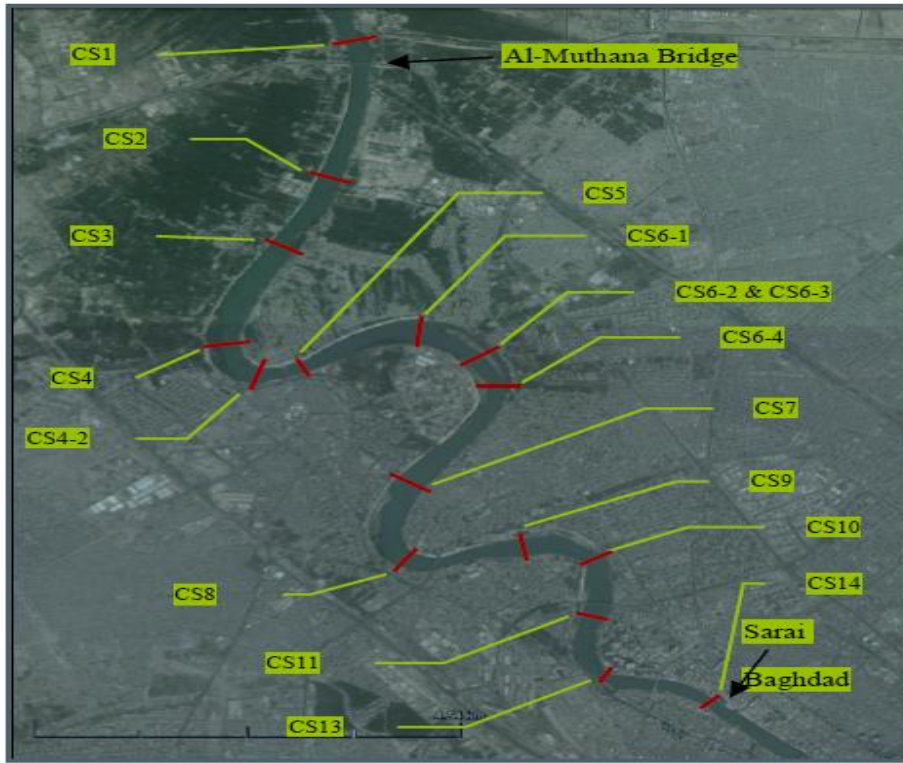
### 3- الجانب التطبيقي The Application

#### 1-3 وصف البيانات الحقيقية Real Data Description

تعد حركة الرواسب (Sediment transport) في الأنهار عنصراً هاماً في موضوع هيدرولوجية الأنهار ويجب قياسه في محطة هيدرومترية، ويجب إجراء القياسات من أجل تعليق وتصريف حمولة القاع في الأنهار ذات الأنظمة الطبيعية. وإن تأثير مشكلة الرواسب على تنمية الموارد المائية هو الدافع الرئيسي لقياس تصريف الرواسب في بعض المواقع على امتداد النهر. وتعد البيانات التي يتم الحصول عليها من القياسات مهمة لفهم عملية التقلبات وظروف التعرية أو الترسبات في مجرى النهر [2:pp.7-8].

وقد تم الاستناد على مجموعة من البيانات الحقيقية ذات عينة بحجم 8 مشاهدات للتطبيق العملي والتي تخص البيانات المحلية لحمولة قاع النهر و التصريف لنهر دجلة في مدينة بغداد، وتم الاعتماد على هذه البيانات من البحث المنشورة من قبل (Sven Knutsson, 2017) حيث قام الباحثون بحساب حمولة القاع و التصريف لنهر دجلة على 16 موقع على طول الجزء الشمالي من نهر دجلة في مدينة بغداد وكما موضحة في الشكل (1)، وقد بين الباحثون أن هناك بعض المواقع قد أضررت نقاط تغيير في انحدار التصريف [3:pp.67].

وسيتم تطبيق بيانات الموقع CS6-1 من نهر دجلة وهي من المواقع التي ظهر فيها نقطة التغيير حيث يحتوي هذا الموقع على 8 مشاهدات وكما موضحة في الجدول (1) وأن الشكل (2) يوضح شكل إنتشار البيانات.



الشكل (1): يوضح موقع العينات لحمولة قاع النهر عبر مقاطع على طول الجزء الشمالي من نهر دجلة في مدينة بغداد [3:p p.60]

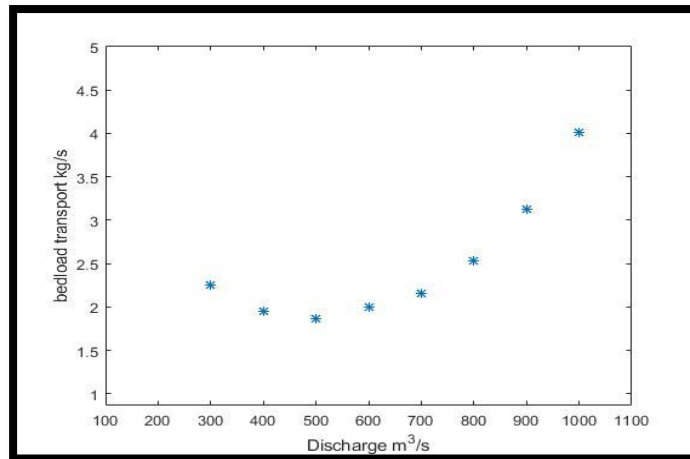
الجدول (1): مجموعة بيانات حمولة القاع و التصريف للموقع CS6-1 لنهر دجلة في مدينة بغداد حسب صيغة (van Rijn 1984) لحساب حمولة القاع<sup>1</sup>

X	300	400	500	600	700	800	900	1000
Y	2.251	1.954	1.866	2.003	2.157	2.529	3.122	4.011

المصدر رقم [3] وتم الحصول عليها من الدكتور (عمار عادل علي) // هندسة موارد المائية.

<sup>1</sup> :  $kg s^{-1}$  مقاس بـ bed load تمثل حمولة القاع Y و  $m^3 s^{-1}$  مقاس بـ Discharge تمثل التصريف X :





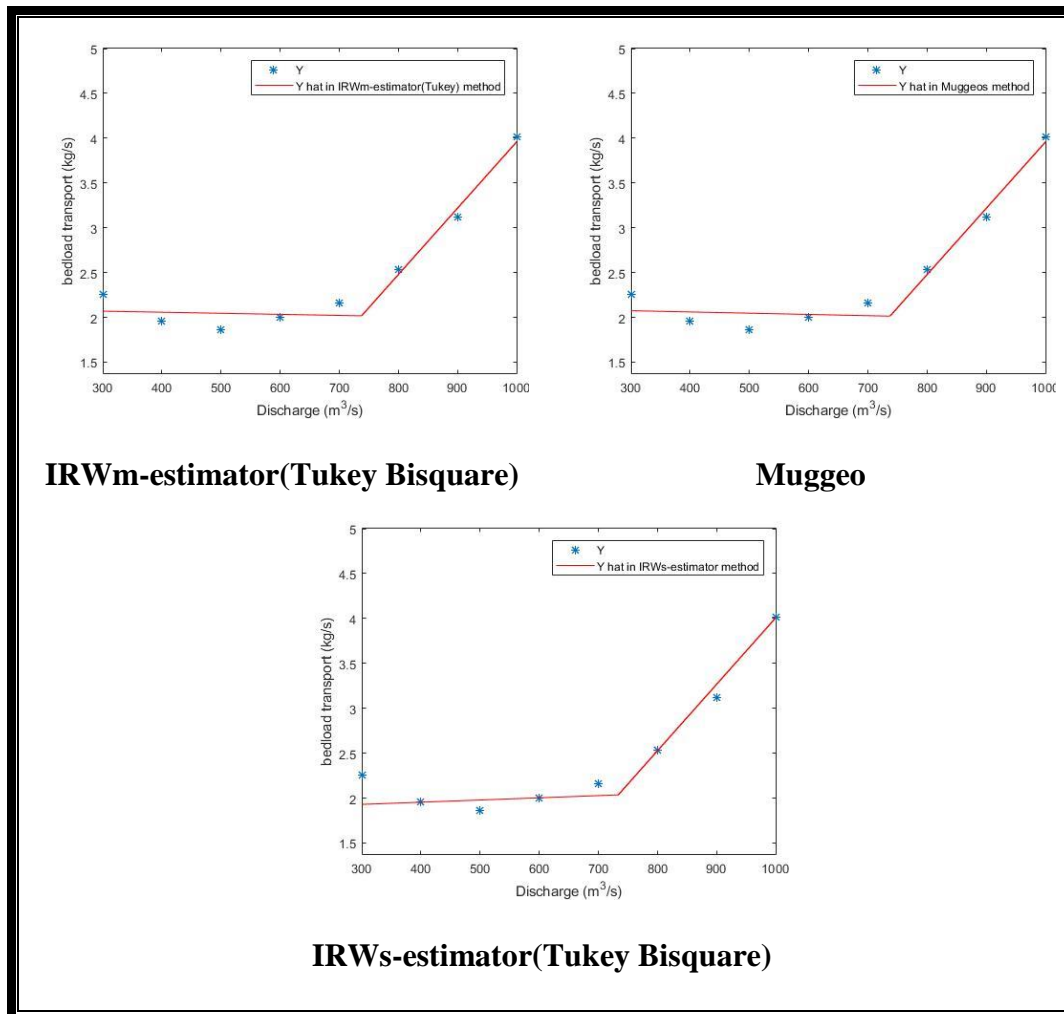
الشكل (2): يوضح أنتشار البيانات لحمولة القاع و التصريف للموقع CS6-1 لنهر دجلة في مدينة بغداد والمحسوبة حسب صيغة (van Rijn 1984) لحساب حمولة القاع

### 2-3 مناقشة النتائج Results Discussion

بعد أن تم الحصول على بيانات المواقع التي ظهر فيها نقطة التغيير في التصريف من الباحث الدكتور في هندسة الموارد المائية (عمار عادل علي) أحد الباحثين في البحث المنشور الذي تم اعتماد بياناته للتطبيق العملي، سيتم تطبيقها على طرائق تقدير الانحدار الخطي المقسم المذكورة في الجانب النظري في الفقرة (2)، وأن الشكل رقم (3) يوضح القيم التنبؤية لكل طريقة، وقد تم حساب الخطأ المعياري لنقطة التغيير ( $S.E.(\hat{\tau})$ ) ومتوسط مربعات الخطأ (MSE) للأنموذج لكل طريقة لأجراء المقارنة بين الطرائق وقد تم الحصول على النتائج في الجدول رقم (2) الآتي:

الجدول (2): يمثل نتائج تطبيق طرائق تقدير الانحدار الخطي المقسم لبيانات حمولة القاع والتصريف للموقع CS6-1 لنهر دجلة في مدينة بغداد

Parameters of the segmented linear regression model	Methods		
	Muggeo	IRWm-estimator	IRWs-estimator
		Tukey Bisquare	Tukey Bisquare
$\hat{\tau}$	737.0557	737.4091	733.4267
$\hat{\beta}_0$	2.1157	2.1049	1.8560
$\hat{\beta}_1$	-0.00014	-0.00012	0.00025
$\hat{\beta}_{00}$	-3.4483	-3.4475	-3.3990
$\hat{\beta}_2$	0.00741	0.00741	0.00741
S.E. ( $\hat{\tau}$ )	34.151	33.203	0.124
MSE	0.02743	0.02522	7.6698e-8



الشكل (3): يوضح القيم التنبؤية لطرائق تقدير الإنحدار المقسم لبيانات الموقع CS6-1 لنهر دجلة في مدينة بغداد

من خلال نتائج الجدول (2) يلاحظ إن الطريقة التي حصلت على أقل قيمة لمتوسط مربعات الخطأ (MSE) والخطأ المعياري لنقطة التغيير ( $\hat{\tau}$ ) S.E. هي طريقة (IRWs-estimator) المقترحة باستعمال دالة (Tukey bisquare) وتليها طريقة (IRWm-estimator) باستعمال دالة (Tukey bisquare).

#### 4- الاستنتاجات والتوصيات Conclusions and Recommendations

##### 1-4 الإستنتاجات Conclusions

بعد التوصل إلى النتائج العددية لتطبيق مجموعة البيانات الحقيقية تم أستنتاج ما يلي:  
1- إن نتائج طريقة (IRWs-estimator) المقترحة كطريقة حصينة وباستعمال دالة (Tukey) أفضل من نتائج طريقة الامكان الأعظم ML لـ (Muggeo) ومن طريقة (IRWm-estimator) الحصينة ويمكن ملاحظة ذلك من خلال نتائج متوسط مربعات الخطأ (MSE) للأنموذج و الخطأ المعياري لنقطة التغيير ( $\hat{\tau}$ ) S.E.

2- يمكن الإعتماد على طريقة (IRWs-estimator) المقترحة وباستعمال دالة (Tukey) لإيجاد نقطة التغيير و معلمات أنموذج الانحدار الخطي المقسم كطريقة حصينة لمعالجة المشاكل التي قادت تواجهنا في الأنموذج مثل انتهاك بعض الفروض الاساسية كأن يكون توزيع الخطأ العشوائي غير طبيعي أو وجود قيم متطرفة تؤثر على كفاءة التقدير.

وبالتالي يمكن القول إن الطريقة الحصينة المقترحة (IRWs-estimator) وباستعمال دالة (Tukey) هي أفضل طريقة حصينة لأنموذج الإنحدار الخطي المقسم مقارنة مع بقية الطرائق المستعملة في الجانب النظري ويمكن الإعتماد عليها في تقدير نقطة التغيير و معلمات الأنموذج.

**2-4 التوصيات Recommendations**

إستناداً إلى الإستنتاجات التي تم التوصل إليها من النتائج العددية للتطبيق لابد من ذكر بعض التوصيات التي بالإمكان الأخذ بها:

- يمكن استعمال طريقة (IRWs-estimator) المقترحة لإيجاد نقطة التغيير أو التحويل في كمية التصريف لمياه النهر لتحديد أنظمة التدفق الضرورية للصيانة الذاتية لقنوات طبقة الحصى الغرينية في قاع النهر.

**3-4 الدراسات المستقبلية Future Studies**

1- ضرورة الإهتمام بتقدير المشاهدات المتطرفة في البيانات بدلاً من حذفها أثناء العمل بالتحليل الإحصائي والعمل على إيجاد طرائق حصينة لذلك.

2- يمكن إجراء دراسات مستقبلية لموضوع الإنحدار المقسم كالاعتماد على دوال هدف وأوزان أخرى في الطريقة المقترحة غير التي أعتمدت في هذه الدراسة، أو عمل دراسات عن طرائق حصينة أخرى في الإنحدار المقسم.

**شكر وتقدير**

ننقدم بجزيل الشكر والامتنان لكل من السادة المحررين والمقيمين العلميين في المجلة لجهودهم في اغناء البحث بملاحظاتهم القيمة. كما ان الشكر موصول الى الدكتور (عمار عادل علي) لتزويدنا ببيانات البحث عن هيدرولوجية مياه نهر دجلة في مدينة بغداد.

**Acknowledgments**

We introduce our thanks to the editors and the reviewers of (JEAS) journal for their efforts in enriching the research with their valuable comments. Thanks also to Dr. Ammar Adel Ali for providing us with data on the hydrology of the Tigris River at Baghdad city.

**المصادر References**

1. Acitas,S., Senoglu,B.,(2020). **Robust change point estimation in two-phase linear regression models: An application to metabolic pathway data**, Journal of Computational and Applied Mathematics, Vol.363, pp. 337–349.
2. Ali A. A .(2016) "**Three Dimensional Hydro- Morphological of Tigris River**". A thesis submitted to Lulea University of Technology. Sweden .
3. Ali, A.A., Al-Ansari N.A., Al-Suhail, Q., and Knutsson S., (2017). **Spatial Measurement of Bedload Transport in Tigris River**. Journal of Earth Sciences and Geotechnical Engineering, vol.7, no. 4, 55-75.
4. Ali,O.A., Abbas,M.A.,(2019),**New Robust Estimator Of Change Point In Segmented Regression Model For Bed-Load Of Rivers**, Journal of Mechanics of Continua and Mathematical Sciences, Vol.14, No.6, pp.384-402.
5. Almetwally,E.M., Alongy,H.M.,(2018), **Comparison Between M-estimation, S-estimation, And MM Estimation Methods of Robust Estimation with Application and Simulation**, International Journal of Mathematical Archive, **55-63** .
6. Feder,P.I.,(1975),**The log likelihood Ratio In Segmented Regression**, Annals of Statistics, pp. 84-97.
7. Gallant,A. R., Fulle, W.A. ,(1973), **Fitting Segmented Polynomial Regression Models Whose Join Points have to be Estimated**, Journal of the American Statistical Association, Vol. 68, No. 341, pp. 144-147.
8. Greene, M.E., Rolfson, O.,Garellick, G.,Gordon, M. and Nemes, S., (2014), **Improved statistical analysis of pre- and post-treatment patient-reported outcome measures (PROMs) the applicability of piecewise linear regression splines**, Quality of Life Research, Vol.24,Issue3,pp.567–573.

9. Hudson, D.J., (1966), **Fitting Segmented Curves Whose Join Points Have to be Estimated**, Journal of the American Statistical Association, 61:316, 1097-1129.
10. Muggeo, V. M.R., (2003), **Estimating regression models with unknown breakpoints**, Statistics in Medicine 22, 3055–3071.
11. Quandt, R.E., (1958), **The estimation of the parameters of a linear regression system obeying two separate regimes**. J. Amer. Statist. Assoc, 53, 873–880.
12. Quandt, R.E., (1960), **Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes**, J. Amer. Statist. Assoc, 55:290, 324-330.
13. Rousseeuw, P. J., Yohai, V. J., (1984). **Robust Regression by Mean of SEstimators**, Robust and Nonlinear Time Series Analysis, New York, ,256-274, doi: 10.1007/978-1-4615-7821-5-15.
14. Salini, S., Cerioli, A., Laurini, F. and Riani, M., (2015), **Reliable Robust Regression Diagnostics**, International Statistical Review, 0, 0, 1–29.
15. Susanti, Y., Pratiwi, H., Sulistijowati, H.S. and Liana, T., (2014). **M estimation, S estimation, and MM Estimation in Robust Regression**. Int. J. Pure Appl. Math. Vol. 91 No. 3, pp. 349-360.
16. Yu, B., Barrett, M. J., Kim, H.-J. and Feuer, E. J. (2007). **Estimating joinpoints in continuous time scale for multiple change-point models**. Computational Statistics & Data Analysis 51, 2420–2427.
17. Zhang, F., Li, Q., (2017), **Robust bent line regression**, Journal of Statistical Planning and Inference.

## Proposing Robust IRWs Technique to Estimate Segmented Regression Model for the Bed load Transport of Tigris River with Change Point of Water Discharge Amount at Baghdad City

**Mohammed Ahmed Abbas**  
Researcher. Sunni Endowment  
Diwan, Department of Religious  
Education and Islamic Studies,  
Baghdad, Iraq  
([mohammad.stat.145@gmail.com](mailto:mohammad.stat.145@gmail.com))

**Omar Abdulmohsin Ali**  
Assist. Prof. Department of  
Statistics, College of Management  
and Economics, Baghdad  
University, Baghdad, Iraq  
([dromar72@coadec.uobaghdad.edu.iq](mailto:dromar72@coadec.uobaghdad.edu.iq))

Received:11/3/2020

Accepted :3/5/2020

Published :August / 2020



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

### Abstract

Segmented regression consists of several sections separated by different points of membership, showing the heterogeneity arising from the process of separating the segments within the research sample. This research is concerned with estimating the location of the change point between segments and estimating model parameters, and proposing a robust estimation method and compare it with some other methods that used in the segmented regression. One of the traditional methods (Muggeo method) has been used to find the maximum likelihood estimator in an iterative approach for the model and the change point as well. Moreover, robust estimation method (IRW method) has used which depends on the use of the robust M-estimator technique in segmentation idea and using the Tukey weight function. The research's contribution lies in the suggestion to use the S-estimator technique and using the Tukey weight function, to obtain a robust method against cases of violation of the normal distribution condition for random errors or the effect of outliers, and this method will be called IRWs. The mentioned methods have been applied to a real data set related to the bed-load of Tigris River/ Baghdad city as a response variable and the amount of water discharge as an explanatory variable. The results of the comparison showed the superiority of the proposed method.

**Keywords:** segmented linear regression, change point, reparameterization, M-estimator, S-estimator, bed-load transport, hydrology of water bodies.