

تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

أ.م.د. وليد عبد الله ارحيمه/الكلية التقنية الادارية- بغداد / قسم تقنيات المعلوماتية
الباحث/ كفاح اسماعيل صحن /الكلية التقنية الادارية- بغداد / قسم تقنيات المعلوماتية

تاريخ التقديم: 2017/10/24

تاريخ القبول: 2017/11/29

المستخلص

ان تأثير القيم المتطرفة في قراءات معلمات النموذج وهي متطلبات الاوكسجين الذائب BOD وكذلك الاوكسجين الذائب (DO) يمكن ان تؤدي الى قراءات خاطئة في مقدرات النموذج لتقدير نسب استهلاك واسترجاع الاوكسجين الذائب (DO) وبالتالي يؤدي الى اطلاق كميات من المياه الثقيلة الملوثة الى الانهر والتي بدورها تؤثر بشكل سلبي جدا على الحياة البيئية والثروات المائية بانواعها المختلفة. ونتيجة لهذا جاء هذا البحث في محاولة لتوظيف نماذج تقدير معلمات Streeter-Phleps وهي K_d, K_r اعادة الاوكسجين الذائب (DO) واستهلاك الاوكسجين الذائب (DO) على التوالي، بعد اكتشاف واستبعاد تأثير القيم المتطرفة باعتماد عدة خوارزميات مقترحة. اظهرت النتائج التجريبية تأثير مقدرات المعلمات ومتوسط مربعات الخطأ العائدة لها بنسب التطرف المختلفة. يمكن اعتماد طرق تقدير اخرى مثل (بييز والمقدرات الحصينة Robust Estimators) وذلك لازالة تأثير القيم المتطرفة خاصة مع نسب اعلى من (20%).

المصطلحات الرئيسية للبحث / معامل استهلاك الاوكسجين الذائب K_d ، معامل استرجاع الاوكسجين الى الماء K_r ، القيم المتطرفة، انموذج ستريتر فليبس Streeter-Phleps، المقدرات، متطلبات الاوكسجين الحيوية BOD، الاوكسجين الذائب DO





1. هدف البحث

يهدف البحث الى اكتشاف وازالة تأثير القيم المتطرفة في قيم كل من متطلبات الاوكسجين الحيوية BOD وكذلك الاوكسجين الذائب DO على مقدرات إنموذج Streeter-Phleps وهي K_d, K_r بالاضافة الى البحث عن افضل خوارزمية تقدير تقابل اقل متوسط مربعات خطأ.

2. اهمية البحث

ان وجود قراءات خاطئة لنسب الاوكسجين الذائب (DO) وكذلك متطلبات الاوكسجين الحيوية (BOD) يمكن ان يؤدي الى حدوث كوارث بيئية وذلك بسبب القرار في اطلاق دفعات مائية الى الانهر تحمل نسبة تلوث كبيرة يمكن ان تؤدي الى دمار الحياة المائية فضلاً عن الحياة البشرية.

3. مقدمة عامة General Introduction

ان قراءات نسب متطلبات الاوكسجين الحيوية (BOD) والاكسجين الذائب (DO) يجب ان تتمتع بالدقة اللازمة وذلك لما لها من اهمية بالغة في تحديد صلاحية او عدم صلاحية دفعات الماء ومن ثم تحديد القرار في اطلاقها من عدمه، ولعل وجود التطرف في بعض هذه القراءات يمكن ان يؤدي الى حدوث اخطاء تصاحب القراءات السابقة فضلاً عن كونها مؤشر لاعادة هذه القراءات تمهيدا الى الوصول الى القرار الملائم. وفي مجال البحث عن مقدرات معلمات نماذج الاوكسجين الذائب (DO) الافضل جرت العديد من البحوث منها

• في العام (1925) قام كل من (El. W. STREETER, Sanitary Engineer, and EARLE B.) بعمل بحث يتضمن دراسة التلوث والتنقية الطبيعية لنهر اوهايو حيث تضمن هذا التقرير مجموعة من الدراسات المقدمة من قبل دائرة الصحة العامة للولايات المتحدة خلال عامي 1914 و 1915 كجزء من مسح شامل للتلوث والتنقية الطبيعية لنهر أوهايو، التي أجريت تحت إشراف سورغ (W. H.) الإشارة إلى وصف تفصيلي لنهر أوهايو، حيث توصلوا الى نماذج رياضية يمكن بواسطتها معالجة عملية اعادة التهوية لنهر اوهايو والانهر التي تمتلك نفس الظروف التي تحيط بهذا النهر. [11]

اما في العام (2010) قام كل من (Gaspere Viviani و Giorgio Mannina) باجراء بحث بعنوان نمذجة جودة المياه في الانهر السريعة الزوال (التي قد تدوم ليوم واحد): تقييم الملغمة وتطوير النموذج. وتضمن البحث بناء نماذج خاصة في دراسة جودة المياه في الانهر الصغيرة بطريقة تقلل من الكلف التي ترافق البيانات الواجب توافرها في قياس جودة المياه. حيث ان المياه الصغيرة لا يمكن استثمارها كما في الانهار الكبيرة لذلك تم اللجوء لعمل نموذج واهو امتداد لنموذج ستريتر- فليس حيث يأخذ بنظر الاعتبار العمليات الفيزيائية والكيميائية والبيولوجية الأكثر صلة بنمذجة نوعية المسطحات المائية المستقبلية (أي تدهور المواد الكربونية الذائبة، أكسدة الأمونيوم، امتصاص الطحالب ونزع النتروجين، توازن الاوكسجين الذائب (DO)، بما في ذلك استفاد الاوكسجين الذائب (DO) عن طريق عمليات التحلل العظوي وغيرها واعادة توليد الاوكسجين الذائب (DO) عن طريق إعادة التأهيل الفيزيائي وإنتاج التمثيل الضوئي. [7]

في حين شهد العام (2013) قيام كل من (Husnain Haider¹, Waris Al² & Sajjad Haydar³) باجراء بحث بعنوان مراجعة نماذج الاوكسجين الذائب (DO) والمتطلبات الاوكسجين الحيوية (BOD) للانهر الكبيرة. حيث تلخص البحث في تطوير وتعديل نماذج رياضية لكل من الاوكسجين الذائب (DO) ومتطلبات الاوكسجين الحيوية (BOD)، حيث بنوا في بحثهم انه ان النماذج الرياضية والحاسوبية التي يتم اختيارها تعتمد على طبيعة الحالة الدراسية التي يتناولها الباحث. واستنتجوا انه النماذج الرياضية للاوكسجين الذائب (DO) ومتطلبات الاوكسجين الحيوية (BOD) تستخدم بشكل واسع في تطوير استراتيجيات ادارة لغرض الحفاظ على جودة المياه للانهر مما لها استخدامات مفيدة في الحياة. [6]



وفي العام (2016) قام كل من (Olga و Evhenia Uhnenko، Natalia Sorochuk و Gubareva) بإجراء بحث بعنوان نماذج التنبؤ طويل الأجل للمياه السطحية في حالة إعادة بناء الطريق السريع. حيث تضمن البحث النظر بالتفصيل في نماذج التوازن ونماذج المياه المكونة من مكون واحد والمكونات المزدوجة المستخدمة في التنبؤ طويل الأجل في إعادة إعمار الطرق السريعة، مستخدمين معادلات ثنائية وثلاثية الأبعاد. وتلخص البحث ارتفاع مستوى الحمل البشري على المسطحات المائية في إعادة بناء الطرق السريعة، التي تتجاوز في كثير من الأحيان قدرتها على التحمل، تنظيم الجريان السطحي للنهر الذي يؤدي إلى تغيرات في النظام الهيدرولوجي للمسطحات المائية واضطراب الظروف الطبيعية للنظم البيئية المائية. [10]

وايضا في العام (2017) قام كل من (Yoshifumi ، Ram K. Regmi، Binaya K. Mishra و Chitresh Saraswat و Pankaj Kumar، Kensuke Fukushi، Masago) بإجراء بحث بعنوان تقييم التلوث في نهر باكماتي في وادي كاثماندو: نموذج والتحليل القائم على اساس السيناريو لاجل التطور الحضري المستمر. وقد تضمن البحث بعرض للتحديات التي يخلقها التلوث في المياه، حيث في بحثهم هذا قاموا بتوضيح قائم على اساس تقييم التطور المستمر لمصادر المياه السطحية لوادي كاثماندو عن طريق تحليل معلمات جودة المياه والتمثلة على سبيل المثال الاوكسجين الذائب (DO) والمتطلبات الحيوية للاوكسجين (BOD). حيث تم استخدام نظام محاكاة لتقييم التلوث في مياه هذا الوادي على طول 25 كم امتدادا بين سانديريجال وجوفار. وقد تلخص البحث في انه هذا جودة المياه في هذا النهر تكون جيدة جدا في الفصول ذات الرياح الموسمية حيث يكون جريان الماء سريع على العكس من الفصل الجاف الذي يكون فيه جريان المياه ضعيف جدا. [8]

ونتيجة لذلك جاء هذا البحث في محاولة لاكتشاف وازالة تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps لتقدير نسب الاوكسجين الذائب (DO) ومتطلبات الاوكسجين الحيوي (BOD) والتي تمثل المعاملات الاساسية في قياس جودة المياه ومدى صلاحيتها للاستخدامات المختلفة، بالاضافة الى البحث عن افضل طريقة استبعاد وازالة تأثير القيم المتطرفة والتي تصاحب اقل متوسط مربعات خطأ.

4. إنموذج Streeter-Phleps:

بدأ اقتراح هذا الإنموذج من قبل الباحث (El. W. STREETER, Sanitary Engineer, and) (EARLE B. PHELPS, في العام (1925) والذي على اثره تم تقديم هذا الإنموذج على وفق الصيغة الاتية:

$$BOD(t) = BOD(0) * e^{-(k_d * t)} \quad (1)$$

$$DO(t) = DO(SAT) - \frac{k_d BOD(0) * e^{-k_d t}}{(k_r - k_d)} + \frac{k_d BOD(0) * e^{-k_r t}}{(k_r - k_d)} - (DO(SAT) - DO(0)) * e^{-k_r t} \quad (2)$$

حيث ان: BOD: تمثل متطلبات الاوكسجين الحيوية. DO: تمثل الاوكسجين الذائب في الماء. [11] K_d : هو المقدر الذي يمثل نسبة استهلاك الاوكسجين الذائب (DO) سواء بسبب التلوث او تنفس الكائنات المائية. K_r : هو المقدر الذي يمثل نسبة إعادة الاوكسجين الى المياه مرة اخرى. t : يمثل السرعة التي يتم فيها القياس لقيم كل من BOD و DO. [11]

5. مصفوفة المعلومات Information Matrix [2]

وهي تمثل الاشتقاقات الجزئية للمعادلتين رقم (1) ورقم (2) في الفقرة رقم (4) والتي يمكن كتابتها بالصيغة الاتية:

a. بالنسبة الى مصفوفة المعلومات لمتطلبات الاوكسجين الحيوية (BOD) هي كالآتي:

$$xf_1 = \begin{bmatrix} \frac{\partial f_1(t_1)}{\partial k_d} & 0 \\ \vdots & \vdots \\ \frac{\partial f_1(t_n)}{\partial k_d} & 0 \end{bmatrix} \quad (3)$$

using \hat{k}_d



حيث ان: $\frac{\partial f_2(t_1)}{\partial k_d}$: تمثل الاشتقاق الجزئي لمعادلة رقم (1) الفقرة (4) وهي كالآتي:

$$BOD(t) = BOD(0) * e^{-(k_d * t)}$$

$$\frac{\partial BOD(t)}{\partial k_d} = - BOD(0) * e^{-k_d t} * t \quad (4)$$

b. بالنسبة الى مصفوفة المعلومات للاوكسجين الذائب (DO) هي كالآتي:

$$xf_2 = \begin{bmatrix} \frac{\partial f_2(t_1)}{\partial k_d} & \frac{\partial f_2(t_1)}{\partial k_r} \\ \vdots & \vdots \\ \frac{\partial f_2(t_n)}{\partial k_d} & \frac{\partial f_2(t_n)}{\partial k_r} \end{bmatrix} \quad (5)$$

using \bar{k}_d and \bar{k}_r

حيث ان: $\frac{\partial f_2(t_1)}{\partial k_d}$: تمثل الاشتقاق الجزئي لمعادلة رقم (2) الفقرة (4) وهي كالآتي:

$$DO(t) = DO(SAT) - \frac{k_d BOD(0) * e^{-k_d t}}{(k_r - k_d)} + \frac{k_d BOD(0) * e^{-k_r t}}{(k_r - k_d)} - (DO(SAT) - DO(0)) * e^{-k_r t}$$

$$\frac{\partial DO(t)}{\partial k_d} = 0 - \frac{(k_r - k_d) * [k_d BOD(0) * e^{-k_d t} * (-t) + e^{-k_d t} * BOD(0)] - BOD(0) * k_d * e^{-k_d t} * (-1)}{(k_r - k_d)^2} + \frac{(k_r - k_d) BOD(0) * e^{-k_r t} - BOD(0) * k_d * e^{-k_r t} * (-1)}{(k_r - k_d)^2} - 0$$

وبعد عدة خطوات نحصل على المعادلة الآتية:

$$\frac{\partial DO(t)}{\partial k_d} = \frac{BOD(0) t k_r k_d e^{-k_d t}}{(k_r - k_d)^2} - \frac{BOD(0) k_r e^{-k_d t}}{(k_r - k_d)^2} - \frac{BOD(0) t k_d^2 e^{-k_d t}}{(k_r - k_d)^2} + \frac{k_r BOD(0) e^{-k_r t}}{(k_r - k_d)^2} \quad (6)$$

$\frac{\partial f_2(t_1)}{\partial k_r}$: تمثل الاشتقاق الجزئي لمعادلة رقم (2) الفقرة (4) وهي كالآتي:

$$DO(t) = DO(SAT) - \frac{k_d BOD(0) * e^{-k_d t}}{(k_r - k_d)} + \frac{k_d BOD(0) * e^{-k_r t}}{(k_r - k_d)} - (DO(SAT) - DO(0)) * e^{-k_r t}$$

$$\frac{\partial DO(t)}{\partial k_r} = 0 - \frac{(k_r - k_d) * [k_d BOD(0) * e^{-k_d t} * (-t) + e^{-k_d t} * BOD(0)] - BOD(0) * k_d * e^{-k_d t} * (-1)}{(k_r - k_d)^2} - \frac{(k_r - k_d) BOD(0) * e^{-k_r t} - BOD(0) * k_d * e^{-k_r t} * (-1)}{(k_r - k_d)^2}$$

وبعد عدة خطوات نحصل على المعادلة الآتية:



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phelps مع التطبيق

$$\frac{\partial DO(t)}{\partial k_r} = \frac{BOD(0)k_d e^{-k_d t}}{(k_r - k_d)^2} + \frac{(t)(k_d^2)BOD(0)e^{-k_r t}}{(k_r - k_d)^2} - \frac{(t)(k_r)(k_d)BOD(0)e^{-k_r t}}{(k_r - k_d)^2} - \frac{BOD(0)k_d e^{-k_r t}}{(k_r - k_d)^2} + (DO(SAT) - DO(0)) e^{-k_r t} * (t)$$

(7)

حيث ان: $DO(SAT)$: تمثل القيمة الابتدائية للتشبع للاوكسجين الذائب. $DO(0)$: تمثل القيمة الابتدائية الحالية. $BOD(0)$: تمثل القيمة الابتدائية لمتطلبات الاوكسجين الحيوية. [2]



6. التباين والتباين المشترك Variance and Covariance

ان الحصول على اقل محدد يمثل الوسيلة المهمة لايجاد الحجم المثالي للعينة وبالتالي يساعد الباحثين في معرفة عدد التصاميم الممكن توليدها، والمعادلة الخاصة بالتباين والتباين المشترك هي كالآتي:

$$\sum_k \hat{t} = \hat{\sigma}_1^2 (X'_1 X_1)^{-1} + \hat{\sigma}_2^2 (X'_2 X_2)^{-1} \quad (8)$$

حيث ان: $\hat{\sigma}_1^2$: يمثل تباين الدالة الاولى والتي تمثل قيم متطلبات الاوكسجين الحيوية (BOD). $\hat{\sigma}_2^2$: يمثل تباين الدالة الثانية والتي تمثل قيم الاوكسجين الذائب (DO). [7]

7. معامل [tucy] BOX PLOT

ان هذا المعامل منذ اكتشافه من قبل العالم (Tukey) في العام (1977) استخدمت هذه الطريقة بشكل متزايد في السياقات الاكاديمية، المهنية وكذلك التعليمية، مثال على ذلك (Bakker, Biehler, & Konold) في عام 2005. ان تمثيل البيانات بهذه الطريقة مفيد جدا من حيث انه يعطي انطباع اول سريع عن التغييرات في توزيعات البيانات، خاصة عبر الزمن مثال (Hoyle & Noss) في عام (2002). [12]

ان هذه الطريقة تكمن في تقسيم البيانات الى ثلاث نقاط اساسية هي (Q_1, Q_2, Q_3) والى وهذ النقاط بدورها تقسم البيانات الى اربع مجاميع متساوية، كل مجموعة تتكون من ربع البيانات المعطاة، (Q_1) يمثل الرقم الوسط بين اصغر رقم في البيانات وبين الوسيط في هذه البيانات، (Q_2) يمثل الوسيط لمجموعة البيانات المعطاة، بينما (Q_3) يمثل الوسيط بين وسيط مجموعة البيانات (Q_2) وبين اعلى قيمة في مجموعة البيانات. ولايجاد كل من Q_1, Q_2, Q_3 نتبع الصيغ الآتية:

$$Q_1 = \text{number of the location}(\text{data size} * 25\%) \quad (9)$$

$$Q_2 = \text{number of the location}(\text{data size} * 50\%) \quad (10)$$

$$Q_3 = \text{number of the location}(\text{data size} * 75\%) \quad (11)$$

$$IQR = Q_3 - Q_1 \quad (12)$$

$$\text{upper whiskers} = Q_3 - 1.5(IQR) \quad (13)$$

$$\text{lower whiskers} = Q_1 + 1.5(IQR) \quad (14)$$

حيث ان: IQR: يتمثل المدى للبيانات . *upper whiskers*: يتمثل الحد الاعلى للقيم في البيانات واي قيمة اعلى منها تسمى قيم متطرفة (Outlier). [12]

lower whiskers: يتمثل الحد الاعلى للقيم في البيانات واي قيمة اقل منها تسمى قيم متطرفة (Outlier).

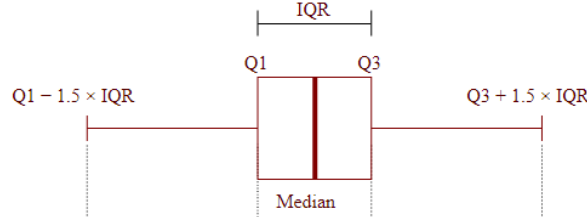
وان طريقة اكتشاف المتطرفات تكون تبعا للخطوات الآتية:

1. ترتيب القائمة بشكل تصاعدي. 2. ايجاد Q_1 وهو يمثل 25% من قيم مجموعة البيانات. 3. ايجاد Q_2 وهو يمثل 50% من قيم مجموعة البيانات. 4. ايجاد Q_3 وهو يمثل 75% من قيم مجموعة البيانات. 5. ايجاد قيم *UPPER- whisker* (U.S) المتطرفات العلوية المقبولة. 6. ايجاد قيم *LOWER- whisker* (L.S) المتطرفات السفلية المقبولة. 7. ايجاد مدى التطرف IQ.
8. استبدال القيم المتطرفة التي هي اعلى من (U.S) بقيمة (U.S). 9. استبدال القيم المتطرفة التي هي اقل من (L.S) بقيمة (L.S).

حيث انه شكل البياني لهذه الطريقة كالآتي: [12]



مخطط (1) : مخطط يمثل كيفية عمل BOX PLOT في تقسيم البيانات الى اربع ارباع.[12]



8. طريقة المربعات الصغرى غير الخطية في تقدير المعلمات

Nonlinear Models Ordinary, Least Squares Method for Estimation Parameter

هذه الطريقة هي وجدت على اساس فكرة مهمة جدا وهي تصغير مربعات الخطأ، حيث تم الاعتماد على هذه الطريقة في ايجاد افضل مقدر الذي تكون فيه قيمة مربعات الخطأ اقل ما يمكن، وان هذه الطريقة يمكن كتابتها بالصيغة الاتية :

$$S(\hat{\theta}) = \sum_{i=1}^n (Y_i - f_i(\zeta, \hat{\theta}))^2$$

حيث ان: $S(\hat{\theta})$ هو مجموع مربعات الخطأ MSE, Y_i : يمثل المشاهدات الحقيقية المستحصلة من الواقع

العملي في قياس. $f_i(\zeta, \hat{\theta})$: دالة غير خطية بدلالة المتغيرات ζ ومعلمات الانموذج (θ) [2,1]

9. الطريقة المقترحة Suggested Method

لضمان اكتشاف واستبعاد تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps، تم اعتماد الخوارزمية الاتية:

1. اكتشاف القيم المتطرفة باعتماد الصيغة رقم (2) Q-PLOT.

2. استبعاد القيم المتطرفة واستبدالها على وفق القيم الاتية:

a. بقيمة الوسط الحسابي. b. مساواته القيمة مقدارها الصفر. c. استبدالها باعلى واقل مدى مسموح للقيم المتطرفة. d. استبدالها بالاعتماد على القيم المجاورة لها. حيث تم في هذا البحث استخدام الطريقة C لمعالجة القيم المتطرفة.

3. تقدير K_d, K_r باستخدام باستخدام OLS.

4. ايجاد محدد التباين والتباين المشترك.

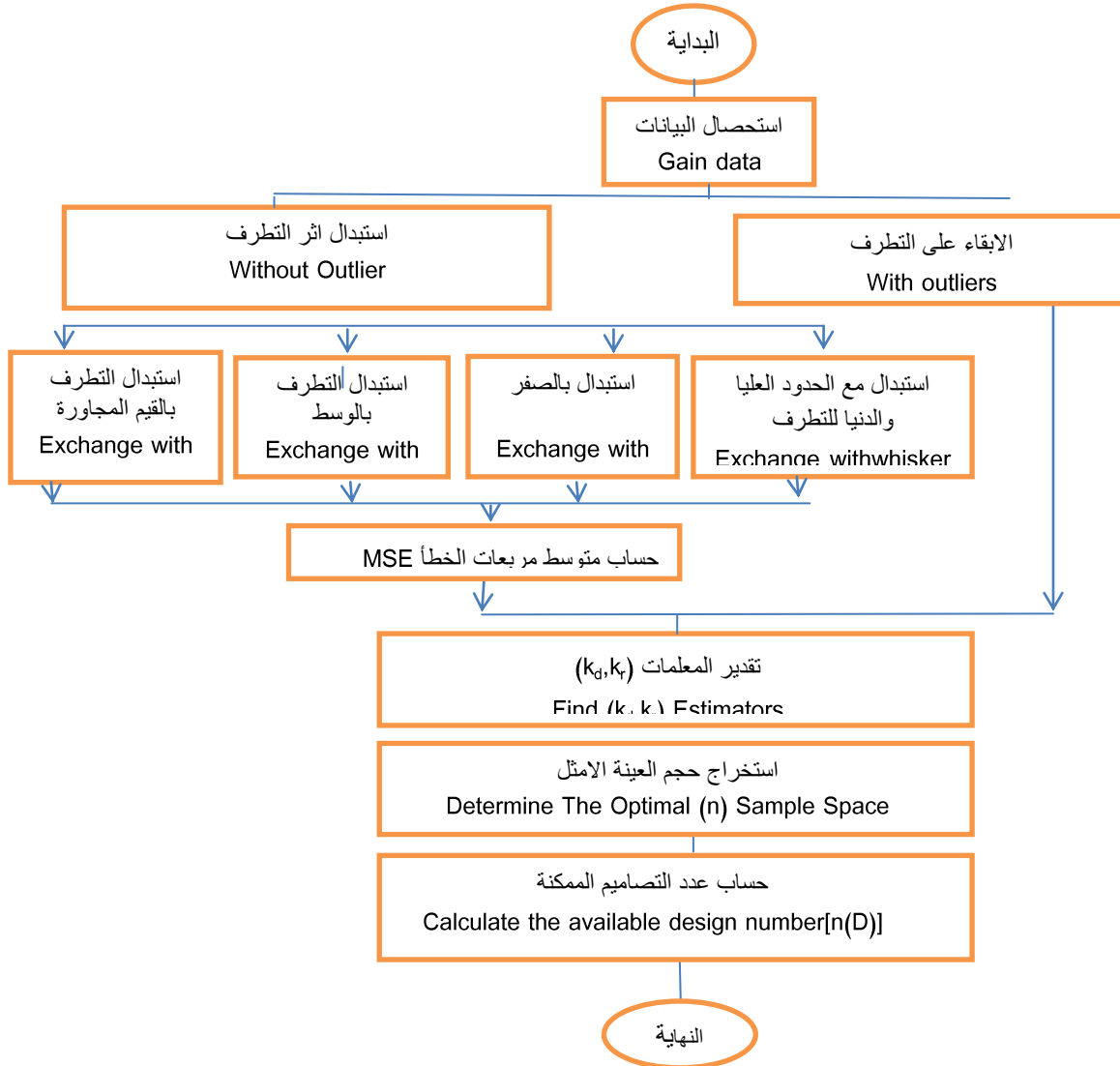
5. اخذ الحجم الذي يعادل اقل محدد للتباين والتباين المشترك يمثل الحجم المثالي للعينة (n).

المخطط الكتلي الاتي سيوضح خطوات العمل



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

مخطط رقم (2): يوضح خطوات العمل في الطريقة المقترحة. (المخطط من اعداد الباحث)



10. النتائج التجريبية Experimental Results

لغرض الحصول الى نتائج تجريبية للطريقة المقترحة تم اعتماد البيانات الخاصة والتي جمعها من احدى محطات التنقية للمياه الصرف الصحي في بغداد ولاربعة ايام متتالية، حيث تم اخذ القراءات الخاصة بكل من (t, BOD, DO) وتم تطبيق الطريقة المقترحة للحصول على المقدرات. حيث تم الحصول على قيم المعلمة (K_d) باستخدام كلا الطريقتين وبنسب تطرف مختلفة تتراوح بين (1%، 2%، 5%، 10%، 20%، 30%) وكانت النتائج كالآتي:

الجدول رقم (1) يمثل تقدير قيمة K_d مع وجود القيم المتطرفة كما هي عليه بدون اي تعامل. حيث ان: (nor): تمثل القيم المتطرفة بالحالة الاعتيادية. (mid): مضاعفة القيم باعتماد الصيغ الاتية
 $miss = miss + 2 * (miss + miss)$
(high): مضاعفة القيم باعتماد الصيغ الاتية
 $miss = miss + 6 * (miss + miss)$



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

الجدول رقم (1): قيم متوسط مربعات الخطأ العائدة الى مقدرات نموذج Streeter-Phleps مع وجود قيم متطرفة بنسب مختلفة بدون معالجة

نسبة التطرف	طبيعة المعلومات	معلمت النموذج	المقدرات
1%	nor	k1	0.12692427525372
		bod0	1.94055792782854
		mse	4.04108665926281
	Mid	k1	-0.48581701261254
		bod0	2.505554188
		mse	28.29145871
	High	k1	-0.40355345053917
		bod0	2.154436021
		mse	193.85316569323900
2%	nor	k1	0.29881400578827
		bod0	2.019479244
		mse	4.560833849
	Mid	k1	-0.23877528012086
		bod0	2.54764760083415
		mse	55.45409352213510
	High	k1	-0.28783131503730
		bod0	2.315750353
		mse	392.03228567236700
5%	nor	k1	0.60048628904717
		bod0	2.12713471653649
		mse	4.95212968414303
	Mid	k1	0.00917620993201
		bod0	2.783958713
		mse	81.47892112
	High	k1	-0.11326769511265
		bod0	2.63320909
		mse	587.15166585443200

نسبة التطرف	طبيعة المعلومات	معلمت النموذج	المقدرات
10%	nor	k1	4.01252961171354
		bod0	4.34320500292046
		mse	4.69357174717892
	Mid	k1	92.71159855366980
		bod0	1.44015850332037
		mse	193.15488166666700
	high	k1	510.96793359511600
		bod0	2.3627E-102
		mse	1238.88021500000000
20%	nor	k1	14.29112684317270
		bod0	1.72232945495062
		mse	12.33959775499620
	Mid	k1	48.93951048723870
		bod0	29.67013756613640
		mse	597.36897753739900
	high	k1	108.31998177028700
		bod0	101.95686286773900
		mse	7826.36414315244000
30%	nor	k1	7.04764801848066
		bod0	4.17068239865444
		mse	7.71293096022233
	Mid	k1	3.59942460127557
		bod0	1555568445
		mse	2.41979313978626E+5
	High	k1	28.51283921301690
		bod0	3.00653E+18
		mse	9.03920886009559E+36

الجدول رقم (1) اظهر قيم مختلفة لمتوسط مربعات الخطأ المرافقة في تقدير معلمت انموذج Streeter-Phleps ضمن نسب مختلفة للقيم المتطرفة وهي ضمن الحدود (1%، 2%، 5%، 10%، 20%، 30%).



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

الجدول رقم (2): قيم متوسط مربعات الخطأ العائدة الى مقدرات إنموذج Streeter-Phleps بعد معالجة المتطرفة بالطريقة المقترحة

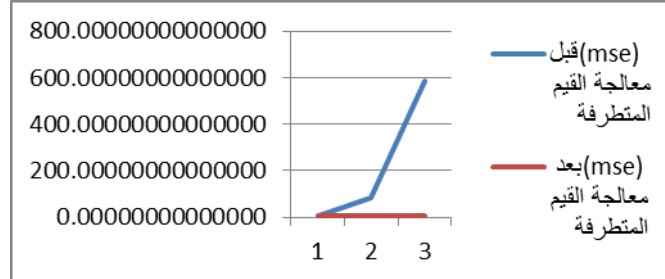
نسبة التطرف	طبيعة المعلومات	معلومات النموذج	المقدرات	نسبة التطرف	طبيعة المعلومات	معلومات النموذج	المقدرات
1%	nor	k1	0.87270214155118	10%	nor	k1	0.99607340378555
		bod0	1.99769782664432			bod0	2.02367527744261
		mse	3.23371139282077			mse	3.20738889059790
	Mid	k1	0.87270214155118		Mid	k1	0.99607340378555
		bod0	1.99769782664432			bod0	2.02367527744261
		mse	3.23371139282077			mse	3.20738889059790
	High	k1	0.87270214155118		high	k1	0.99607340378555
		bod0	1.99769782664432			bod0	2.02367527744261
		mse	3.23371139282077			mse	3.20738889059790
2%	nor	k1	0.92852632161477	20%	nor	k1	1.47069424867897
		bod0	2.00188171476882			bod0	2.27162544597655
		mse	3.24728494665915			mse	2.24413118559412
	Mid	k1	0.92852632161477		Mid	k1	1.47069424867897
		bod0	2.00188171476882			bod0	2.27162544597655
		mse	3.24728494665915			mse	2.24413118559412
	High	k1	0.92852632161477		high	k1	1.47069424867897
		bod0	2.00188171476882			bod0	2.27162544597655
		mse	3.24728494665915			mse	2.24413118559412
5%	nor	k1	1.15397411757127	30%	nor	k1	7.04764801848066
		bod0	2.03201435191609			bod0	4.17068239865444
		mse	3.17865173953290			mse	7.71293096022233
	Mid	k1	1.15397411757127		Mid	k1	3.59942460127557
		bod0	2.03201435191609			bod0	1555568445
		mse	3.17865173953290			mse	2.41979E+18
	High	k1	1.15397411757127		High	k1	28.51283921
		bod0	2.03201435191609			bod0	3.01E+18
		mse	3.17865173953290			mse	9.03921E+36

الجدول رقم (2) اظهر قيم مختلفة لمتوسط مربعات الخطأ المرافقة في تقدير معاملات إنموذج Streeter-Phleps ضمن نسب مختلفة للقيم المتطرفة وهي ضمن الحدود (1%، 2%، 5%، 10%، 20%، 30%) والتي تم معالجتها بواسطة الطريقة المقترحة.



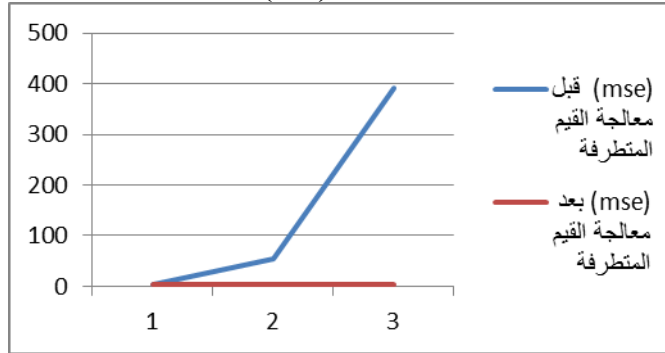
تأثير القيم المتطرفة على مقدرات انموذج Streeter-Phleps مع التطبيق

ويعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (1%) للجدولين (2و1) تم الحصول على المخطط الآتي:
مخطط رقم(3): قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (1%) للجدولين (2و1)



لقد بين المخطط رقم (3) انه بعد معالجة القيم المتطرفة ساعد ذلك على استقرار البيانات وتاهيلها للاعتماد في القياس واصدار القرار، حيث نشاهد ان قيم متوسط مربعات الخطأ (mse) والمشار إليها بالمنحنى المرسوم باللون الازرق في حالة عدم معالجة المتطرفات في حالة تزايد مستمر وقيم عالية وهذا يدل على انه هذه النتائج لا يمكن الاعتماد عليها بشكل مباشر في الانتقال الى المرحلة الثانية في المخطط الكتلي رقم (2) المذكور انفا وذلك كون القيم غير دقيقة نوعا ما، بينما نلاحظ بان قيم متوسط مربعات الخطأ (mse) والمشار إليها بالمنحنى ذي اللون الاحمر انه في حالة استقرار وينسب قليلة جدا قياسا عن الاول وبقيمة (3.91644792582314)، حيث ستكون قيم K_d المقدرة بعد اتمام المعالجة مؤهلة للمرحلة الثانية في المخطط الكتلي.

ومن هذا يتبين لنا انه بعد التعامل مع المتطرفات سوف نكون قد تخلصنا من الاخطاء الناتجة عن القياس لقيم كل من (متطلبات الاوكسجين الحيوية BOD والاكسجين الذائب DO).
ويعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (2%) للجدولين (2و1) تم الحصول على المخطط الآتي:
مخطط رقم(4): قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (2%) للجدولين (2و1)

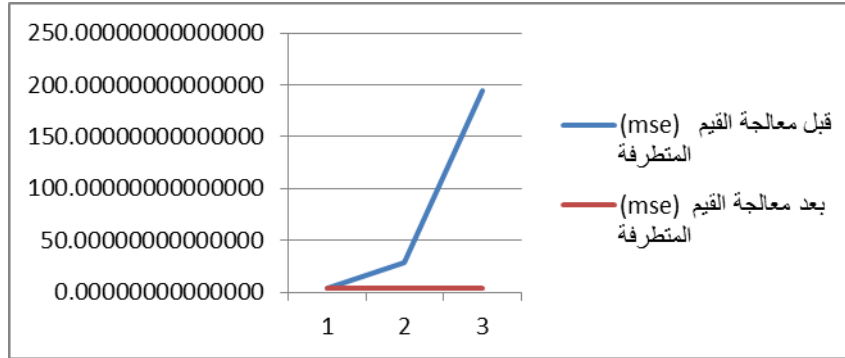


نلاحظ من خلال هذا المخطط اننا حصلنا على نتائج مستقرة كما هو الحال عندما كانت نسبة المتطرفات هو 1%، ولكن هنا كان قيمة متوسط مربعات الخطأ بعد معالجة المتطرفات هي مستقرة على القيمة (4.33308473799828)، وهذا الشيء يثبت بانه لو ازلنا المتطرفات يساعد في استقرار البيانات وجعلها مهينة بشكل جيد بغية ايجاد المقدرات المعنية في موضوع البحث، حيث انه كلما كانت القيم مستقرة كلما كانت نتائج التقدير اقرب للموثوقية والاطمئنان. وهذا يدل على انه الباحث في الجانب الامن في الانتقال الى المرحلة التالية.

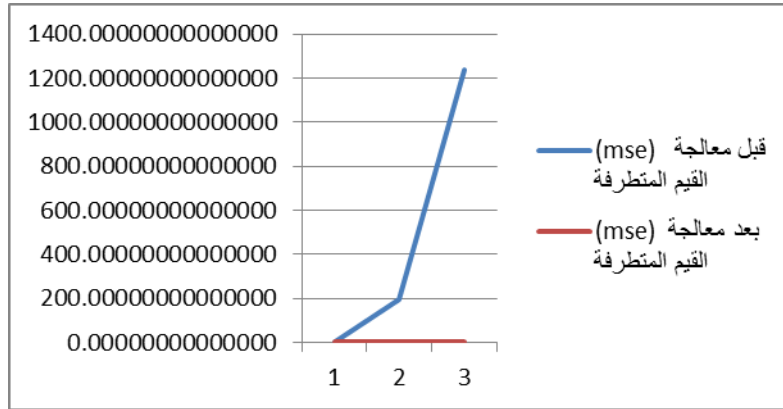


تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

وبعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (5%) للجدولين (2و1) تم الحصول على المخطط الآتي:
مخطط رقم(5): قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (5%) للجدولين (2و1)



ان هذا المخطط يوضح لنا بانه بعد معالجة القيم المتطرفة في البيانات يساعد ذلك على استقرارها وبالتالي تكون مؤهلة للمرحلة الثانية في اجراءات موضوع البحث.
وبعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (10%) للجدولين (2و1) تم الحصول على المخطط الآتي:
المخطط رقم (6): قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (10%) للجدولين (2و1)

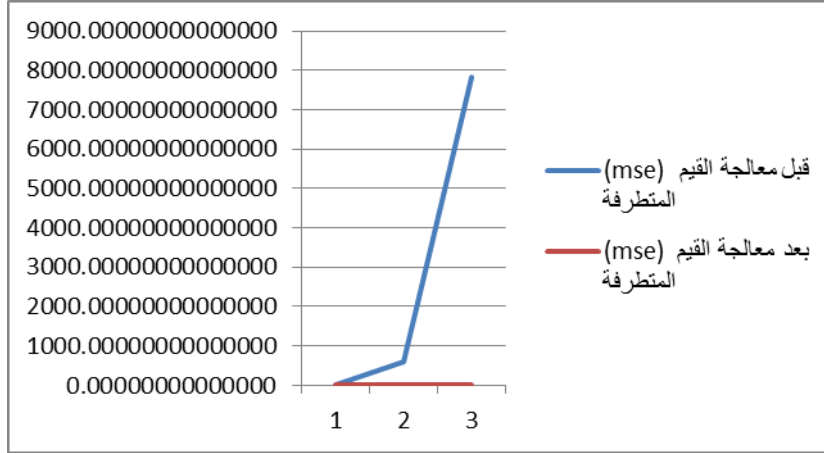


من هذا المخطط نلاحظ ايضا انه استقرار البيانات يكون بعد ازالة القيم المتطرفة التي هي نتائج اخطاء في القياسات بسبب الاجهزة المستخدمة للقياس وعدم توفر الظروف المناسبة للقياس في بعض الاحيان.
وبعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (20%) للجدولين (2و1) تم الحصول على المخطط الآتي:



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

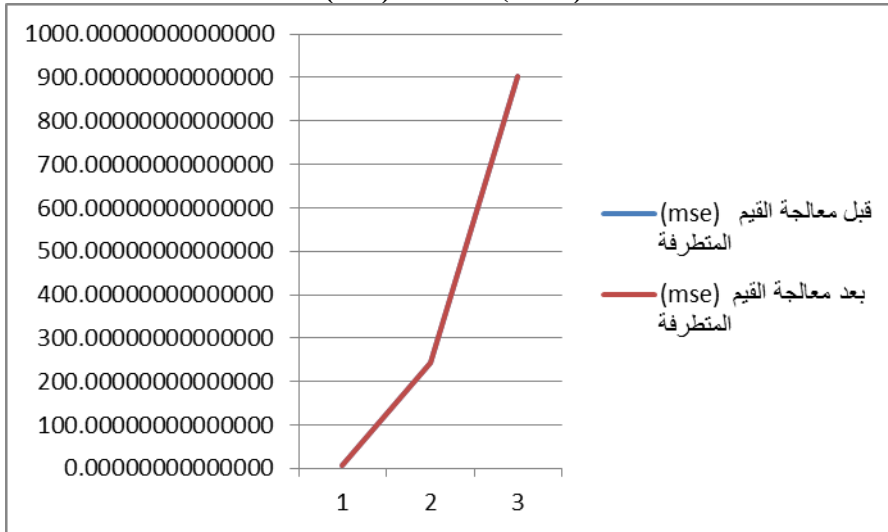
مخطط رقم (7): قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (20%) للجدولين (2و1)



ايضا في هذه الحالة نلاحظ في هذا المخطط انه الاستقرار يكون بعد ازالة المتطرفات من البيانات بسبب الاخطاء المتوقعة التي تم ذكرها في السابق.

ويعمل مخطط لقيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (30%) للجدولين (2و1) تم الحصول على الاتي:

مخطط رقم (8) : قيم متوسط مربعات الخطأ MSE قبل وبعد معالجة القيم المتطرفة ضمن نسبة التطرف (30%) للجدولين (2و1)



عند ملاحظة هذا المخطط يتضح لنا بانه القيم المقاسة التي تحتوي على متطرفات بنسبة 30% يدل على انها قيم غير صحيحة ويجب الابعاز الى مسؤول المختبر باعادة النظر في الاجهزة المستخدمة في قياس القيم المراد ايجادها اي يجب اجراء الصيانه على هذه الاجهزة او استبدالها. وعند مقابلة الباحث مع مسؤولي المختبر تبين بان الاجهزة تتعطل عن العمل بسبب الغازات التي تنبعث من مياه الصرف الصحي تسبب في حدوث ثقوب في انابيب الاجهزة وبالتالي تسبب عطلها وعدم تمكنها من اعطاء القياسات المضبوطة. وتمثل هذه المشكلة احد المشاكل المهمة التي يعاني منها الكادر في المشروع.



11. إيجاد الحجم الأمثل Determine the Optimal Size of Sample

لايجاد الحجم الأمثل للعينة سوف يتطلب إيجاد محدد المصفوفة عند سرعة عمل مختلفة للفحوصات والتي تعمل بها الجهة تم أخذ البيانات منها، وعند أحجام عينة مختلفة وهي (2,3,4,5,6) وكانت النتائج كالآتي :
Exchange The Outliers by Whiskers (11-1) عندما يتم استبدال المتطرفات الحدود العليا والدنيا للمتطرفات

جدول رقم (3): محددات مصفوفة التباين والتباين المشترك لكل حجم عينة عند سرعة مختلفة بعد معالجة القيم المتطرفة

المحدد	حجم العينة	السرعة
33.28172531	2	0.55
14.88288325	3	
8.443288528	4	
5.462676115	5	
3.843578013	6	
44.44737327	2	0.59
19.84669342	3	
11.23645547	4	
7.251145332	5	
5.086285505	6	
60.62661249	2	0.63
27.04234117	3	
15.28784621	4	
9.847194258	5	
6.891778382	6	
84.24631663	2	0.67
37.54774582	3	
21.20324603	4	
13.63807756	5	
9.528603331	6	
119.0154797	2	0.71
53.01078596	3	
29.90914316	4	
19.21638278	5	
13.40796973	6	
169.6610068	2	0.75
75.53752133	3	
42.59430141	4	
27.34629676	5	
19.06343004	6	

لقد أوضح هذا الجدول قيم المحددات التي ترافق كل حجم عينة وسرعة معينة والذي تبين انه يرتفع من سرعة الى اخرى.

(11-2) عندما يتم الإبقاء المتطرفات الحدود العليا والدنيا للمتطرفات **With outliers** كانت النتائج كالآتي:



تأثير القيم المتطرفة على مقدرات إنموذج Streeter-Phleps مع التطبيق

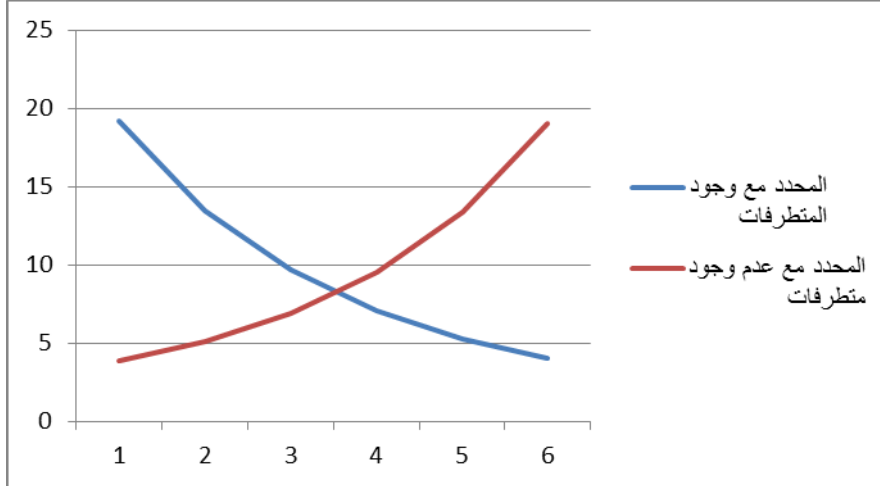
جدول رقم (4): محددات مصفوفة التباين والتباين المشترك لكل حجم عينة عند سرع مختلفة في حالة ابقاء القيم المتطرفة دون معالجة

المحدد	حجم العينة	السرعة
170.6351485	2	0.55
75.96479962	3	
42.8301775	4	
27.49358097	5	
19.16259027	6	
119.5431486	2	0.59
53.25062391	3	
30.04824026	4	
19.30885125	5	
13.47510908	6	
85.46618265	2	0.63
38.09513765	3	
21.5152719	4	
13.84116261	5	
9.672510647	6	
62.10635878	2	0.67
27.70539573	3	
15.66505867	4	
10.09210266	5	
7.064817907	6	
45.96708906	2	0.71
20.52789402	3	
11.62417575	4	
7.503026153	5	
5.264376989	6	
34.73939949	2	0.75
15.53706827	3	
8.816252346	4	
5.705474688	5	
4.015669541	6	

لقد اوضح هذا الجدول قيم المحددات التي ترافق كل حجم عينة وسرعة معينة والذي تبين انه يقل كلما زادت السرعة التي يتم عندها اخذ عينة.



وبعمل مخطط للنتائج المستحصلة في كل من الجدولين (4 و3) بغية المقارنة نحصل على الاتي:
مخطط رقم (8): قيم المحددات لمصفوفة التباين والتباين المشترك في حالة معالجة وعدم معالجة القيم المتطرفة



نلاحظ من خلال هذا الرسم انه في عادل عدم وجود المتطرفات يبين لنا ما هو قريب من الواقع انه كلما كانت السرعة التي عندها يقوم مسؤول المختبر اخذ العينة كبيرة كلما كانت قيمة المحدد كبيرة وبالتالي نسبة الخطأ تكون كبيرة، والعكس بالعكس عندما تكون السرعة قليلة كلما كان المحدد اقل وبالتالي نسبة الخطأ تكون قليلة، اما بالنسب للمعالجة مع وجود المتطرفات فانه يؤدي الى نتيجة بالعكس وهي خلاف المنطق حيث انه كلما كانت السرعة كبيرة كان المحدد صغير وهذا يعني انه نسبة الخطأ تكون قليلة وهذا الشيء غير قريب من الواقع، لانه السرعة البطيئة تعطي دقة اكبر في معرفة مدى التلوث في المياه.

12) الاستنتاجات والتوصيات Suggestions and Conclusions

بعد اجراء عمليات التقدير وحساب متوسط مربعات الخطأ العائدة لها ظهرت لدينا عدد من الاستنتاجات والتوصيات منها :

12-1) الاستنتاجات Conclusions

1. ان وجود قيم متطرفة يمكن ان يؤدي الى حدوث اخطاء كبيرة تصاحب عمليات التقدير وبالتالي ظهور متوسط مربعات خطأ عالي المصاحب لها.
2. ان نسب القيم المتطرفة يؤدي الى زيادة في قيم متوسط مربعات الخطأ بالتوافق مع زيادة نسب القيم المتطرفة.
3. ان معالجة القيم المتطرفة يمكن ان يدي الى تقليل من قيم متوسط مربعات الخطأ المصاحبة لعمليات التقدير.
4. ان عدم معالجة القيم المتطرفة يمكن ان يؤدي الى حدوث اخطاء في عمليات التقدير لمعلممات الانموذج وهذا بدوره يؤدي الى حدوث اخطاء في قرار اطلاق الدفغات المانية ممايؤدي الى دمار الحياة المائية.
5. وجود القيم المتطرفة يؤدي الى حديث اخطاء في تصاميم الحجم الامثل للمعاينة وبالتالي حدوث اخطاء في قراءات اخذ العينة ضمن السرعة الفعلية التي يعمل بها المشروع .

12-2) التوصيات Suggestions

1. اعتماد نماذج اخرى لغرض تمثيل نسب الاوكسجين الذائب (DO) ومقارنتها مع انموذج Streeter-Phleps يمكن ان يؤدي الى الوصول الى مقدرات افضل وبالتالي الحصول على متوسط مربعات خطأ اقل.
2. اعتماد طرق (QQPLOT) في اكتشاف القيم المتطرفة ومقارنتها مع الانموذج المقترح لغرض مقارنة النتائج.
3. اعتماد طرق اخرى لتقدير معلممات الانموذج مثل طريقة ببيز والتي تفترض وجود توزيع اولي لمعلممات الانموذج يمكن توظيفه للحصول على مقدرات افضل ومتوسط مربعات خطأ اقل.



المراجع References

1. النقاش، ا. ع. (1997). تخطيط شبكات توزيع كهرباء الاحياء السكنية " اطروحة دكتوراه، كلية الادارة والاقتصاد، الجامعة المستنصرية.
2. عباس، ش. (2010). "تقدير دالة البقاء لتوزيع ويبيل ذو المعلمات الثلاثة لمرض سرطان الدماغ في البيانات الكاملة." رسالة ماجستير، كلية التربية ابن الهيثم / قسم الرياضيات، جامعة بغداد للعام 2010.
3. Atkinson, A., A. Donev, et al. (2007). Optimum experimental designs, with SAS, Oxford University Press.
4. Azar, A. T. (2010). "Adaptive neuro-fuzzy systems." Fuzzy Systems. Vienna, Austria: IN-TECH 10: 7220.
5. Casman, E. (1986). A robust version of classical D-optimal design applied to dissolved-oxygen sag-curve calibration. Integrated design of hydrological networks. Proceedings of the Budapest symposium.
6. Haider, H. (2013). A Review of Dissolved Oxygen and Biochemical Oxygen Demand Models for Large Rivers.
7. Mannina, G. and G. Viviani (2010). "Water quality modelling for ephemeral rivers: Model development and parameter assessment." Journal of Hydrology 393(3): 186-196.
8. Mishra, B. K., R. K. Regmi, et al. (2017). "Assessment of Bagmati river pollution in Kathmandu Valley: Scenario-based modeling and analysis for sustainable urban development." Sustainability of Water Quality and Ecology.
9. Montgomery, D. C. (2001). "Design and Analysis of Experiments, John Wiley & Sons." New York: 64-65.
10. Sorochuk, N., E. Uhnenko, et al. (2016). "Models of Long Term Forecasting of Surface Waters State in Reconstruction of Highway." Procedia Engineering 134: 139-145.
11. Streeter, H. W., Phelps, E.B., (1925). "A study of the pollution and natural purification."
12. Tukey, J. W. ((1977)). "Exploratory data analysis." . . Schematic summaries. In J. W. Tukey (Ed.), . Reading: Allison-Wesley.: (pp. 27e56).
13. Waterhouse, T. H. (2005). "Optimal experimental design for nonlinear and generalised linear models."



The Effect of Extreme Values on Streeter-Phleps Model Parameter Estimators With Application

Abstract

The extremes effects in parameters readings which are BOD (Biological Oxygen Demands) and DO(Dissolved Oxygen) can caused error estimating of the model's parameters which used to determine the ratio of de oxygenation and re oxygenation of the dissolved oxygen(DO),then that will caused launch big amounts of the sewage pollution water to the rivers and it's turn is effect in negative form on the ecosystem life and the different types of the water wealth.

As result of what mention before this research came to employees Streeter-Phleps model parameters estimation which are (K_d, K_r) the de oxygenation and re oxygenation ratios on respectively, after detect and putting aside the extremes effects by follow some suggested algorithms.

The experimental results shows the effects of parameter estimators and the mean square errors(MSE) that belong to it by the different ratio of extremes.

We can follow some other estimating methods such as (bias and robust estimators) to remove the effect of the extreme values specially with the ratio more than (20%).

Keywords: K_d (deoxygenation ratio), k_r (reoxygenation ratio), Extreme Values, Streeter-Phleps model ,Estimator ,BOD (biological oxygen demans), DO(dissolved oxygen).