المجلة العراقية للعلوم الإحصائية

http://stats.uomosul.edu.iq

استخدام خوارزمية PSO في تقدير معلمات عملية Cox

محمد زيد حسين و مثنى صبحي سليمان

قسم الاحصاء والمعلوماتية ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق

معلومات النشر

تاريخ المقالة:
تم استلامه في 1 آب 2019
تم القبول في 26 ايلول 2019
متاح على الإنترنت في 1 حزيران 2020

الكلمات الدالة:
عملية Cox، خوارزمية PSO، مقدرالامكان الاعظم، معمل سمنت

المراسلة:
محمد زيد حسين
mohammed777zaid@gmail.com

الخلاصة

تم في هذا البحث اقتراح استخدام خوارزمية PSO لتقدير المعدل الزمني لحدوث الحوادث لعملية Cox، وتم مقارنة نتائج الطريقة الذكائية المقترحة للتقدير مع طريقة الإمكان الأعظم لتقدير المعدل الزمني للحدوث. وتضمن البحث تطبيقاً واقعياً تناول فيه فترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية بالايام من (الشركة العامة لاسمنت الشمالية/ معامل سمنت بادوش الجديد) خلال الفترة من 1/4/2018 الى 31/1/2019. كما تم تقدير المعدل الزمني لفترات تشغيل الماكنة بالطرائق المقترحة للاستخدام في البحث.

DOI: 10.33899/IQJOSS.2020.165442 , ©Authors, 2020, College of Computer Science and Mathematics, University of Mosul.
This is an open access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

1. المقدمة Introduction :

تستخدم نماذج عملية بواسون غير المتجانسة (NHPP) بشكل رئيسي للتحليل ونمذجة بيانات الفشل للأنظمة القابلة للإصلاح، وتفترض هذه النماذج أن العملية النقطية $\{N(t), t \geq 0\}$ لديها زيادات مستقلة وتتوزع توزيع بواسون، ويكون حدوث الحوادث في عملية بواسون بصورة عشوائية ورتيبة خلال فترة زمنية معينة وبنسبة حدوث ثابتة لكل وحدة زمنية يرمز لها بالرمز λ، بينما تكون النسبة التي تحدث عندها الحوادث في عملية بواسون غير المتجانسة متغيرة بتغير الزمن t، وتسمى النسبة عندئذ بالمعدل الزمني للحدوث (Rate of Occurrence) او يطلق عليها بدالة الشِدّة (Intensity Function) ويرمز لها بالرمز $\lambda(t)$. وتعد عملية كوكس (Cox Process) حالة خاصة من عملية بواسون غير المتجانسة والتي تستخدم على نطاق واسع لدراسة وتحليل بيانات الفشل (الاعطال) المتجمعة مع مرور الزمن [2].

2. عملية كوكس Cox Process:

اذا كانت العملية $\{N(t); t \geq 0\}$ تمثل عملية بواسون غير المتجانسة فان عدد الحوادث التي تحدث ضمن فترة زمنية $(0, t]$ تتبع توزيع بواسون بدالة كتلة احتمال :

$$P[N(t) = n] = \frac{[m(t)]^n e^{-m(t)}}{n!} \quad , \; n = 1, 2, \ldots \tag{1}$$

اذ ان $m(t)$ تمثل معلمة العملية (المعدل)، وهي الدالة التراكمية للمعدل الزمني للحدوث، وتعرّف بالصيغة الآتية [6]:

$$m(t) = \int_0^t \lambda(v)\, dv \quad , \quad t > 0. \tag{2}$$

اذ ان $\lambda(t)$ تمثل المعدل الزمني للحدوث، وعليه فان عملية Cox هي عملية بواسون غير متجانسة عندما يكون المعدل الزمني للحدوث معرّف وفق الصيغة الآتية [7]:

$$\lambda(t) = e^{(\alpha+\beta t)} \quad , \quad t > 0 \tag{3}$$

إذ إنّ $\alpha\,,\beta$ تمثلان معلمتا المعدل الزمني لحدوث الحوادث لعملية Cox ، ويمكن تقديرهما بعدة طرائق، وفي هذا البحث تم تقديرهما باستخدام طريقتي الإمكان الأعظم وطريقة خوارزمية PSO.

3. طريقة الإمكان الأعظم لتقدير معلمات عملية Cox : MLE Method to Estimate Cox Process:

تعد طريقة الإمكان الأعظم للتقدير MLE من الطرائق الاكثر استعمالاً في تقدير معلمات النماذج التصادفية لما تتميز به من خصائص جيدة، منها الثبات وخاصية المقدر غير المتحيز بأقل تباين ممكن (Minimum Variance Unbiased estimators)، ويمكن تعريف التقدير بهذه الطريقة على انه قيم المعلمات التي تجعل دالة الامكان الاعظم للمشاهدات في نهايتها العظمى. لتكن العملية $\{N(t);\, t \ge 0\}$ تمثل عملية Cox الموصوفة بالصيغة (3)، فان الدالة الاحتمالية المشتركة لأزمنة الحدوث $(t_1\,, t_2\,, \dots\,, t_n)$ بحيث $(0 < t_1 \le t_2 \le \cdots \le t_n \le t_0)$ تكون بالشكل التالي [3][5]:

$$f_n(t_1, t_2, \dots, t_n) = \prod_{i=1}^{n} \lambda(t_i) e^{-\int_0^{t_0} \lambda(u)\,\partial u} \tag{4}$$

وعليه فان الدالة التراكمية للمعدل الزمني للحدوث، والتي تمثل معلمة عملية Cox والمعرّفة في الصيغة (2) ستكون بالشكل الآتي:

$$m(t) = e^{\alpha}(e^{\beta t_0} - 1)/\beta \tag{5}$$

اما دالة الامكان لعملية Cox للفترة $[0, t_0)$ بالمعدل الزمني $\lambda(t)$ تمثل بالصيغة التالية [5]:

$$L = \prod_{i=0}^{N(t_0)} \lambda(t_i) e^{-\int_0^{t_0} \lambda(u)\,\partial u} \tag{6}$$

عندئذ تصبح الصيغة (6) على النحو الاتي:

$$L = e^{(n\alpha + \beta \sum_{i=1}^{n} t_i)} \exp\left[-e^{\alpha}\,(e^{\beta\, t_0} - 1)/\beta\right] \tag{7}$$

ومن الصيغة (7) يمكن ايجاد مقدر الامكان الاعظم للمعلمة α عندما تكون β معلومة كما يأتي:

$$\hat{\alpha} = \ln\left(\frac{n\beta}{e^{\beta t_0 - 1}}\right) \tag{8}$$

ويمكننا الاستدلال على توزيع المعلمة β من خلال التوزيع الشرطي للمتغير $\delta = \sum t_i$ ويكون مشروطاً بعدد الحوادث n، لان المشاهدات الموجودة في دالة الامكان الاعظم لعملية Cox لاتأتي الا من n و$\sum t_i$ فقط [7]. بالنسبة لمقدر الامكان الاعظم للمعلمة β فيتم ايجاد التوزيع الاحتمالي لها والذي يمثل التوزيع الشرطي للمتغير $\delta = \sum t_i$ بـ n من الحوادث، كما في الصيغة التالية:

$$f\{\delta | N(t) = n\} = \frac{n! \prod_{i=1}^{n} \lambda(t_i)}{[m(t_0)]^n} \qquad ; n = 1,2, \dots \tag{9}$$

عندما تكون المعلمة α ثابتة ومقيدة بعدد n من الحوادث فان دالة الامكان الاعظم الشرطية للمعلمة β في عملية Cox يتم تمثيلها على النحو الاتي [4]:

$$L\{\beta | N(t) = n\} = \begin{cases} \dfrac{n!\, \beta^n\, e^{\beta \sum t_i}}{(e^{\beta\, t_0 - 1})^n} & ; \quad \beta \ne 0 \\ \dfrac{n!}{t_0^n} & ; \ \beta = 0 \end{cases} \tag{10}$$

عندما $\beta = 0$ فان دالة الكثافة الاحتمالية في الفترة الزمنية $[0, t_0)$ تمثل التوزيع المنتظم (Uniform Distribution)، اما في حالة $\beta \ne 0$ عندئذ تكون دالة كثافة الاحتمال ولنفس الفترة الزمنية والمأخوذة لعينة عشوائية حجمها n من مجتمع تتوزع

توزيعاً اسياً مبتوراً (Truncated Exponential Distribution)، وعادة ماتكون هذه العملية للإحصاءات المرتبة (Order Statistics). ان دالة كثافة الاحتمال للمتغير العشوائي t في حالة وجود مشاهدة واحدة فقط للمعدل الزمني لحدوث الحوادث $\lambda(t)$ الموصوف بدالة Cox هي [7]:

$$f(t) = \begin{cases} \frac{\beta\, e^{\beta\, t}}{e^{\beta t_0 - 1}} & ;\ \beta \neq 0 \\ \frac{1}{t_0} & ;\ \beta = 0 \end{cases} \quad , 0 < t \leq t_0 \tag{11}$$

يمكن من خلال الصيغة (11) ايجاد مقدر الامكان الاعظم للمعلمة β كما في الخطوات الاتية:

$$l(\beta) = \ln L(\beta) = \ln\ [f(t)]^n \tag{12}$$

عندئذ فأن :

$$l(\beta) = \begin{cases} n \ln \beta - n \ln\left(e^{\beta\, t_0} - 1\right) + \beta \sum_{i=1}^{n} t_i + \ln(n!) & , \beta \neq 0 \\ \ln(n!) - n \ln t_0 & , \beta = 0 \end{cases} \tag{13}$$

يتم ايجاد مشتقة اللوغاريتم لدالة الامكان الاعظم بالنسبة للمعلمة β كالاتي:

$$D_\beta\, l(\beta) = \frac{\partial\, l(\beta)}{\partial\, \beta}$$

عندئذ نستنتج الاتي:

$$D_\beta\, l(\beta) = \begin{cases} \frac{n}{\beta} - \frac{n\, t_0}{(1-e^{-\beta t_0})} + \sum_{i=1}^{n} t_i & , \beta \neq 0 \\ -\frac{1}{2} n\, t_0 + \sum_{i=1}^{n} t_i & , \beta = 0 \end{cases} \tag{14}$$

ان مقدر الامكان الاعظم للمعلمة β يمكن ايجاده من خلال حل المعادلة الاتية:

$$D_\beta\, l(\beta) = 0 \tag{15}$$

عندما $\beta \neq 0$ عندئذ نقول انه لا يوجد حل جبري للمعادلة (15)، وعليه فقد تم استخدام الطرائق العددية (Numerical Methods) للحصول على عدد من القيم المقدرة للمعلمة β. اذ تعتبر طريقة نيوتن (Newton's Method) [10]، احدى الطرائق المستخدمة لحل المعادلات التابعة لطريقة الامكان الاعظم والتي عن طريقها تم الحصول على المقدر بالصيغة الاتية:

$$\hat{\beta}_i = \hat{\beta}_{i-1} - \left[\frac{n}{\hat{\beta}_{i-1}} - \frac{n\, t_0}{\left(1-e^{-\hat{\beta}_{i-1} t_0}\right)} + \sum t_i\right]\left[-\frac{n}{\hat{\beta}^2{}_{i-1}} - \frac{n\, t_0^2 e^{-\hat{\beta}_{i-1} t_0}}{(1-e^{-\hat{\beta}_{i-1} t_0})^2}\right]^{-1} \tag{16}$$

4. خوارزمية امثلية سرب الجسيمات Particle Swarm Optimization Algorithm:

تُعّد خوارزمية (PSO) من الطرائق الذكائية المهمة في الوقت الحاضر الذي شهد تطوراً ملحوظاً في الآونة الاخيرة خصوصاً في مجال البحث العلمي والتقنيات, ويطلق عليها ايضا اسم خوارزمية (امثلية سرب الطيور)، ان هذه الطريقة تم استعمالها بشكل كبير في التطبيقات العلمية والعملية لما لها فائدة كبيرة جدا والتي تكون اهدافها الحصول على افضل مقدّر للمّعلمات باقل وقت ودقة عالية، فضلا عن تقدير مّعلمات النموذج قيد الدراسة [11]. في الفترة الاخيرة كانت هنالك أساليب كثيرة لتطوير الخوارزمية والتي تختلف من ناحية المفهوم عن الطرائق التقليدية، اذ يتم وصف تلك الأساليب بأنها طرائق حديثة وغير تقليدية للتحسين. ومعظم هذه الأساليب تستند على صفات وسلوك معينه من الحالات كالبيولوجية الجزيئية وسرب من الحشرات والنُظم العصبية. كما يستند طريقة امثلية سرب الجسيمات (PSO) على سلوك مستعمرة من الكائنات الحية، مثل سرب من الحشرات كالنحل والزنابير والنمل بنوعيه الاسود والأبيض, فضلا عن المجاميع التي تتشكل بهيئة سرب كمجموعة من الاسماك او الطيور. والسبب الذي يعود لتسمية طريقة (PSO) بهذا الاسم لأنها تعتمد في عملها على سلوك هذه المستعمرات، اما الجسيم فيشير مثلاً إلى نحلة في مستعمرة أو طير في سرب الطيور. كل فرد أو جسيم في السرب يتصرف بطريقة متوزعة باستخدام ذكائه الخاص أو الذكاء الجماعي للسرب. فعلى سبيل المثال إذا اكتشف أحد الجسيمات مساراً جيداً للغذاء، عندئذٍ ستكون بقية جسيمات السرب أيضاً قادرة على اتباع المسار الجيد على الفور حتى لو كان موقعهم بعيداً عن السرب او المجموعة. ان الاساليب المُتبعة للتحسين المستندة إلى ذكاء السرب والمستوحاة من السلوكية تُعرف بمجموعة

خوارزميات بدلاً من الخوارزميات الجينية، والتي تسمى بالتطور المُعتمد على الإجراءات [8]. ويفترض أن لكل جسيم في السرب له صفتان اساسيتان هما: الموقع (position)، والسرعة (velocity)، ويتحرك في مساحة او فضاء التصميم (التشكيل) الذي ينضم اليه ويحاول ان يصل الى أفضل موقع (من حيث مصدر الغذاء بالنسبة لسرب الطيور او قيمة دالة الهدف بالنسبة لمسالة رياضية معينة) تم اكتشافه. وتقوم الجسيمات بتوصيل المعلومات بالنسبة للمواقع الجيدة لبعضها البعض وتعديل او تحديث مواقعها الفردية وسرعاتها (التي تم اعتبارها سابقة) بناءاً على المعلومات الواردة حول المواقع الجيدة (التي تم اعتبارها جديدة). وكمثال على ذلك، يعتبر سلوك الطيور في منطقة معينة، على الرغم من ان كل طائر لديه ذكاء محدود في حد ذاته، فأنه يتبع القواعد البسيطة الاتية:

- يحاول عدم الاقتراب أكثر من الطيور الأخرى.
- يرسم هدفه او يشق طريق باتجاه متوسط (Average) الطيور الأخرى.
- يحاول ان يكون (متوسط الموقع) بين الطيور الأخرى دون ان يحدث فجوات كبيرة في المجموعة.

وهكذا يقوم سلوك المجموعة او السرب معتمد على مزيج من ثلاثة عوامل:

- تتماسك الجسيمات فيما بينها بحيث يحدث نوعا ما من التآزر.
- انقسام او انفصال للسرب بحيث لا يحصل تقارب جداً للجسيمات.
- اصطفاف او تشكيل المجموعة بحيث يحاذي كل جسيم مع اخر فيما بينه بحيث يقوم الجسيم بتتبع العنوان العام (المصدر) للمجموعة.

ويمكن تطوير طريقة خوارزمية (PSO) بالخطوات الاتية:

1) عندما يُحدد أحد الجسيمات او الطيور هدفاً أو طعاماً او أعظم (اكبر دالة هدف)، عندئذٍ ينقل المعلومات على الفور إلى كل الجسيمات او الطيور الأخرى.
2) جميع الجسيمات او الطيور الأخرى تنجذب إلى الهدف أو الطعام (أو الحد الأقصى لدالة الهدف) ولكن ليس مباشرة.
3) هناك مّكون من التفكير المستقل الخاص لكل طائر فضلا عن اعتماده على الذاكرة السابقة له. وبالتالي فأن النموذج يحاكي البحث العشوائي في مساحة التصميم للحصول على أقصى قيمة من دالة الهدف. على هذا النحو وبشكل تدريجي تستخدم العديد من التكرارات في هذه المحاولة، وعندها الجسيمات او الطيور تذهب إلى الهدف (أو الحد الأكبر من دالة الهدف).

1-4 آلية عمل امثلية سرب الجسيمات Particle Swarm Optimization Method (PSO)

لتكن لدينا مسألة تعظم قيمة لـ(دالة الهدف) وبقيود غير مشروطة، اذ ان الحد الاعظم Maximize لدالة الهدف $f\,(\mathbf{X})$ بمدى $X^{(l)} \leq X \leq X^{(u)}$ عندما يشير $X^{(l)}$ و $X^{(u)}$ الى الحد الادنى والحد الاعلى على التوالي بالنسبة للمتغير X، عندئذ يمكن تنفيذ هذه الخوارزمية من خلال الخطوات الاتية [11]:

1) نفترض أن عدد الجسيمات (حجم السرب) يتمثل بالرمز N، ولتقليل العدد الكلي لتقديرات الدالة اللازمة لإيجاد حل يتوجب علينا أن نفترض ايضاً حجم أصغر من عدد الجسيمات بحيث يكون صغير جدًا. أن من المحتمل الوصول إلى أطول حل او في بعض الحالات قد لا نستطيع التوصل إلى حل على الإطلاق، وعادةً ما يفترض ان يكون حجم الجزيئة من 20 إلى 30 جزيئة للسرب ككّل وسطي في حالة جعل السرب يكون مرناً.
2) توليد مجتمع اولي (ابتدائي) بصورة عشوائية لقيم X في المدى $X^{(l)}$ و $X^{(u)}$ أي $X_1 , X_2 , \dots , X_N$، ولسهولة الحساب نعتبر الجسيمات هي الموقع j والمقدرة في التكرار i تشير الى سرعتها في التكرار $X^{(i)}$ و $V^{(j)}$ على التوالي، لذلك تعتبر الجزيئات المتولدة بصورة اولية والتي تشير لقيم $X_1(0) , X_2(0), \dots , X_N(0)$ وان المتجهات التي تتمثل $X_j\,(0); (j = 1,2, \dots , N)$ تُعرف بمحاور الجسيمات وهي تكون مشابهة لعمل الكروموسومات في الخوارزميات الجينية. ان عملية تقدير دالة الهدف للقيم المقابلة للجسيمات يكون بالشكل الاتي:
$$f\,[\,X_1(0)]\ ,f\,[\,X_2\,(0)]\,, \dots\ ,f\,[\,X_N\,(0)]$$
3) ايجاد سُرع الجسيمات، اذ تتحرك جميع الجسيمات الى النقطة المثلى بسرعة معينة وفي البداية تكون جميع سُرع الجسم افتراضية وتعطى لها قيمة صفرية، اذ نضع رقم تكراري عندما $i = 1$.
4) في i من التكرارات، يتم تقدير المعّلمات والبالغ عددها 2 والتي تستخدم بواسطة جسيم نموذجي (تجريبي) والمتمثل بـ j عندها يتم الآتي:

a. اختيار افضل قيمة بصورة عامة للـ$X_j^{(i)}$ أي ان (المحاور لـ j من الجسيم في التكرار i المقابل له) تعرف بـ$P_{best\,j}$ وهي تمثل قيمة عظمى لدالة الهدف $f\,[\,X_j^{(i)}]$، اذ يتم توجيهها بواسطة الجسيم j في جميع التكرارت السابقة، وان افضل قيمة لـ $X_j^{(i)}$ الجديدة ستمثل المحاور لجميع الجسيمات الجديدة وبتكرار جديد. كذلك اختيار G_{best} والتي تعتبر قيمة عظمى لدالة الهدف $f\,[\,X_j^{(i)}]$ والتي تمثل جميع التكرارات السابقة بواسطة أي جسيم من N من الجسيمات.

b. ايجاد سرعة الجسيمj فيi من التكرار وبالصيغة الاتية:

$$V_j^{(i)} = V_j^{(i-1)} + C_1 r_1 \left[P_{best\,,j} - X_j^{(i-1)}\right] + C_2 r_2 \left[G_{best\,,j} - X_j^{(i-1)}\right] \quad ; \quad j = 1,2 \ \dots , N \qquad (17)$$

C_2, C_1 : تقومان بتكوين افراد متقاربين بحيث يعملان عمل الكروب الجماعي او يعتبران عنصران من عناصر التدريب على التوالي.

$r_2\,,r_1$: مقداران يتوزعان توزيع منتظم بأرقام عشوائية في المدى [1، 0] .

ان المعّلماتC_2, C_1 تشير الى العلاقة المهمة لذاكرة (الموقع) للجسيم نفسه الى ذاكرة (الموقع) للسرب، اذ ان القيم $C_2\,, C_1$ تكون عادة قيم مفترضة وقيمها تتمثل بالرقمين 1 و 2، لذا فان $C_1 r_1$ و $C_2 r_2$ تعتبران غير مؤكدة للجسيمات التي سوف تكون اكبر هدف والتي تستغرق نصف الوقت.

c. ايجاد موقع او محور الجسيم j في التكرار i :

$$X_j^{(i)} = X_j^{(i-1)} + V_j^{(i)} \qquad ; \ j = 1,2, \dots , N \qquad (18)$$

اذ يتم افتراض خطوة زمنية للوحدة في مدى السرعة في المعادلة (18) أعلاه لتقييم دالة الهدف المقابلة للجسيمات $f\,[\,X_1(0)]\ ,f\,[\,X_2\,(0)]\,, \dots\ ,f\,[\,X_N\,(0)]$.

d. يتم فحص التجمع او التقاء الجسيمات للحساب الحالي، اذا كانت مواقع الجسيمات متقاربة لنفس قيم المجموعة عندئذ فأن الطريقة المفترضة لها تكون متقاربة. اما اذا ظهر معيار التقارب ليس متحققاً عندئذ تُعاد الخطوة الرابعة لكن بعملية تحديث لعدد التكرارات عندما $i = i + 1$، ونحسب القيم الجديدة لأفضلP بمعنى $(P_{best,j})$ وافضل G بمعنى $(G_{best\,,j})$. وتستمر هذه العملية التكرارية حتى تتجمع كل الجسيمات مع بعضها البعض وصولاً الى افضل حل والذي يمثل الحل الامثل.

4-2 تحسين طريقة امثلية سرب الجسيمات: Improvement to the Particle Swarm Optimization Method:

لقد وجُدّ بأن سرعات الجسيمات تتراكم بسرعة أكبر من الحد الأعلى بحيث تتخطى دالة الهدف. وبالتالي يضاف مصطلح القصور الذاتي (التعطيل) θ لتقليل السرعة. عادةً قيمة θ المُفترضة تتغير خطيًا من 0.9 إلى 0.4 كعملية تكرارية متقدمة. يُفترض أن سرعة الجسيم j ، مع مصطلح القصور الذاتي θ ، تعطى بالمعادلة الاتية [11]:

$$\mathbf{V}_j(i) = \theta\, V_j^{(i-1)} + c_1 r_1 \left[\mathbf{P}_{best,j} - X_j^{(i-1)}\right] + c_2 r_2 \left[\mathbf{G}_{best} - X_j^{(i-1)}\right]; \quad j = 1,2, \dots , N \qquad (19)$$

ان وزن القصور الذاتي θ الأصلي الذي قدمه العالم Shi and Eberhart في عام 1999 يؤدي إلى إخماد السُرع بمرور الوقت (أو التكرارات)، مما يُمّكن السّرب من الالتقاء أكثر بدقة وفاعلية مقارنة مع خوارزمية PSO الأصلية في المعادلة (17). وتشير المعادلة (19) إلى صياغة سرعة التكّيف، مما يُحّسن من ظهورها وضبط قدرتها في حل البحث. بالاضافة الى انها توضح أن أكبر قيمة لـ θ تعزز الفحص الشامل اما أصغر قيمة فهي تعزز البحث الموقعي. وبالتالي فأن القيمة الكبيرة لـ θ تجعل الخوارزمية من استكشاف مواقع جديدة بأستمرار دون الكثير من البحث المحلي وبالتالي الفشل في العثور على الحل الامثل الصحيح. لتحقيق التوازن بين الاستكشاف الشامل والمتخصص لتسريع التقارب إلى الحل الأمثل الحقيقي، نستخدم وزن القصور الذاتي θ الذي تنخفض فيه القيمة خطيًا باستخدام عدد من التكرارات وكما في الصيغة الاتية:

$$\theta(i) = \theta_{max} - \left(\frac{\theta_{max} - \theta_{min}}{i_{max}}\right) i \qquad (20)$$

حيث ان:

θ_{max} وθ_{min}: تمثل القيم الأولية والنهائية لوزن القصور الذاتي θ على التوالي.

i_{max} : يمثل الحد الأعلى لعدد التكرارات المستخدمة في PSO.

وتم ملاحظة ان القيم $\theta_{max} = 0.9$ و $\theta_{min} = 0.4$ تستخدم بشكل واسع لدى الباحثين.

5. خطأ النسبة الاعظم Maximum Percentage Error (MPE):

يعد MPE احد المعايير المهمة من معايير قياس جودة التوفيق (Goodness of fit) ، اذ تم استخدام هذا المعيار في هذا البحث للمقارنة بين طريقتي الامكان الاعظم وطريقة PSO لتقدير معلمات عملية Cox، ويهتم هذا المعيار بالتراكيب الفردية لمجموعة البيانات [9]:

$$S_i = \sum_{j=1}^{i} x_j \quad \hat{S}_i = \sum_{j=1}^{i} \hat{x}_j \tag{21}$$

ويعرف معيار MPE بالصيغة الاتية :

$$MPE = \max_{1 \le i \le n} \left[\left| S_i - \hat{S}_i \right| / S_i \right] \tag{22}$$

6. اختبار تجانس عملية Cox: Test of the Homogeneity of Cox Process

تعد عملية Cox عملية غير متجانسة وذلك لان المعدل الزمني لحدوث الحوادث يتغير بتغير الزمن t، أي أنها تتأثر بالزمن t في سلوكها، ومن الملاحظ أن المعلمة β مقترنة بالزمن t ، وبذلك فإن عملية Cox تكون متجانسة في حالة $\beta = 0$ ، وغير متجانسة في حالة $\beta \neq 0$، ولإجراء عملية الاختبار فيما إذا كانت العملية متجانسة أم غير متجانسة فيتم اختبار الفرضية الآتية [1]:

$H_0 : \beta = 0$

$H_1 : \beta \neq 0$

وان المختبر الاحصائي المستخدم لأختبار الفرضية اعلاه هو:

$$Z = \frac{\sum_{i=1}^{n} t_i - \frac{n t_n}{2}}{\sqrt{\frac{n t_n^2}{12}}} \tag{23}$$

اذ ان:

$\sum_{i=1}^{n} t_i$: تمثل مجموع اوقات حدوث الحوادث للفترة الزمنية $[0, t_n)$

n : تمثل عدد الحوادث التي تحدث في للفترة الزمنية $[0, t_n)$

7. الجانب التطبيقي:

يعد معمل سمنت بادوش الجديد في محافظة نينوى احد اهم المعامل التابعة للشركة العامة لاسمنت الشمالية والذي يمثل مصدر مهم ورئيسي لانتاج مادة السمنت لمحافظات العراق عموماً ومحافظة نينوى خصوصاً، اذ تم اعتماد فترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية لمعمل السمنت بالايام خلال الفترة من 1/4/2018 الى 31/1/2019.

1-7 اختبار تجانس البيانات قيد الدراسة:

تم اختبار تجانس البيانات قيد الدراسة باستخدام المختبر الاحصائي في الصيغة (23)، وباستخدام البرنامج المعد لهذا الغرض باللغة البرمجية MATLAB/R2017b تم الحصول على القيمة المحسوبة ($Z = 49.4796$)، وهي اكبر من القيمة الجدولية المقابه لها (1.96) عند مستوى معنوية 0.05، وعليه تم رفض فرضية العدم وقبول البديلة اي ان العملية قيد الدراسة غير متجانسة.

2-7 تقدير معلمات عملية Cox لفترات تشغيل طاحونة المواد الاولية:

لغرض تقييم اداء طريقة PSO الذكائية في تقدير معلمات عملية Cox، تم تقدير معلمات العملية قيد الدراسة باستخدام الطريقة المقترحة PSO للتقدير ومقارنتها مع الطريقة التقليدية والمتمثلة بطريقة الامكان الاعظم. اذ تم تقدير معلمات عملية Cox لفترات تشغيل طاحونة المواد الاولية لمعمل سمنت بادوش الجديد، وتم اعداد برامج بأستخدام اللغة البرمجية MATLAB/R2017b لكتابة الخوارزميات الخاصة بكل طريقة وكما في الجدول التالي:

جدول (1): مقدرات معلمات عملية Cox لفترات التشغيل بالأيام لطاحونة المواد الاولية.

Methods	Parameters Estimation	
	$\hat{\alpha}$	$\hat{\beta}$
MLE	-1.0181	0.0018
PSO	0.5424	0.0165

الجدول (1) يوضح مقدرات معلمات عملية Cox لفترات التشغيل بين توقفين متتالين بالأيام لطاحونة المواد الاولية باستخدام طرائق التقدير المقترحة للاستخدام في البحث، علماً انه تم اجراء عدة تجارب في تقدير معلمات العملية قيد الدراسة باستخدام طريقة PSO من خلال اعطاء قيم اولية (ابتدائية)، وتم الاستنتاج بان افضل قيم مقدرة للمَعلمتين $\hat{\alpha}$ و $\hat{\beta}$ تم الحصول عليهما بالاعتماد على القيم الاتية:

- الاوزان الابتدائية C2,C1 ثم مساواتها بالواحد: C1 = C2 =1
- سرعة الجسيم الذي ينطلق نحو الهدف هي: V= 0.8
- عدد التكرارات لتنفيذ خطوات البرنامج هي: Iteration = 25
- فترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية بالايام خلال الفترة الزمنية الممتدة من 1/4/2018 الى 31/1/2019 والتي تمثل 53 تشغيل / اليوم هي:

t = [3 8 2 4 1 1 2 3 1 1 1 1 3 2 3 1 1 1 2 3 5 6 5 2 1 1 4 1 4 3 1 3 1 1 7 2 5 1 2 1 1 3 3 1 6 1 2 3 3 1 3 2 1]

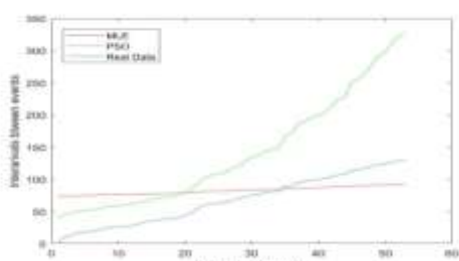
8. مناقشة النتائج Discussion of Results:

لغرض المقارنة بين الطرائق المستخدمة لتقدير معلمات عملية Cox، تم استخدام معيار خطأ النسبة الاعظم MPE حسب الصيغة (22)، وباستخدام البرنامج المعد لهذا الغرض باللغة البرمجية MATLAB\R2017b تم الحصول على العدد المتوقع لفترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية لمعمل سمنت بادوش الجديد خلال الفترة الزمنية قيد الدراسة، وتم حساب معيار MPE بين القيم الحقيقية والقيم المقدرة للمعدل الزمني لتوقف المعمل، وكما في الجدول التالي:

جدول (2): قيم MPE للطرائق المستخدمة لتقدير معلمات عملية Cox .

Methods	MPE
MLE	0.9972
PSO	0.7547

يلاحظ من الجدول (2) ان قيمة خطأ النسبة الاعظم MPE لمقدرات طريقة PSO اقل من قيمة خطأ النسبة الاعظم لطريقة الامكان الاعظم في التقدير، ممايدل على كفاءة الطريقة الذكائية في تقدير معلمات عملية Cox. والشكل التالي يمثل دالة عملية Cox المقدرة باستخدام طرائق التقدير التقليدية والذكائية المستخدمة في البحث مقارنةً مع القيم التراكمية الحقيقية التي تمثل فترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية لمعمل سمنت بادوش الجديد:



الشكل(1): الدوال المقدرة للعدد المتراكم لفترات التشغيل المتتالية بين توقفين متتالين لطاحونة المواد الاولية لمعمل سمنت بادوش الجديد باستخدام طرائق التقدير.

باستخدام طرائق التقدير لفترات التشغيل بين توقفين بالايام لبيانات البحث في الشكل (1)، تم ملاحظة ان طريقة PSO كانت الاقرب الى البيانات الحقيقية، مما يدل على كفاءة هذه الطريقة في التقدير مقارنة مع طريقة الامكان الاعظم لبيانات البحث.

.9 الاستنتاجات والتوصيات Conclusions and Recommendations:

تم في هذا البحث التوصل الى بعض الاستنتاجات والتوصيات التي يمكن ادراجها بالنقاط الاتية:

a) من خلال استخدام اختبار التجانس لعملية Cox لفترات تشغيل طاحونة المواد الاولية لمعمل سمنت بادوش الجديد، تم ملاحظة ان العملية غير متجانسة.

b) تم الاستنتاج بأن تقدير معلمات عملية Cox بالطريقة الذكائية والمتمثلة بطريقة (PSO) هي افضل من الطريقة التقليدية والمتمثلة بطريقة الامكان الاعظم للتقدير، وذلك لأنها اعطت اقل قيمة لمعيار خطأ النسبة الاعظم (MPE) وبأقل وقت وجهد.

c) ان قيم المعلمات $\hat{\alpha}$ و $\hat{\beta}$ المقدرة في طريقة PSO تم الحصول عليهما في البرنامج المعد لهذا الغرض عند التكرار (25) وهي تمثل بذلك سرعة كبيرة واختصاراً للوقت مقارنة مع طريقة الامكان الاعظم التي تمثل وقتاً اطول.

مما تقدم يمكن إعطاء التوصيات الاتية:

a) نوصي بأعتماد خوارزمية PSO في تقدير المعدل الزمني لعملية Cox وذلك لكفاءتها وسرعتها في التقدير.

b) نوصي بأستخدام خوارزميات اخرى كطريقة ذكائية لتقدير معلمات عملية Cox.

c) نوصي باستخدام مقدرات خوارزمية PSO لتقدير معلمات عمليات تصادفية اخرى وذلك لكفاءتها في التقدير.

References

1- Jaafar, Ayat Sadiq, (2016) “Bayesian and traditional methods for estimating the parameters of some heterogeneous Poisson models with practical application - a comparative research”, an unpublished Master of Science thesis in statistics, College of Administration and Economics, University of Baghdad.

2- Al-Khayyat, Basil Younes, and Suleiman, Muthanna Sobhi, (2007), “Statistical Analysis of the Heterogeneous Poisson Process with Application”, Future Research Journal, Issue (17), pp. (133-153), Iraq.

3- Noman, Inam Abdel Rahman (2012), “Designing acceptance sampling plans for the General Company for Electronic Industries using the general exponential distribution,” PhD thesis in statistics, College of Administration and Economics, University of Baghdad.

4- Naima, Ali Bandar, (2009), “Comparison of ML&WLS estimates for some models of heterogeneous Poisson processes”, unpublished Master of Science thesis in Statistics, College of Administration and Economics, University of Baghdad.

5- Basawa, I. V., and Prakasa Rao, B. L. S. (1980),"Statistical Inference for Stochastic Processes". Academic Process, London.

6- Burr, D. (1994),"On inconsistency of Breslow's estimator as an estimator of the hazard rate in the Cox model". Biometrics 50, 1142-1145.

7- Cox, D. R., and Lewis, P. A. (1966),"Statistical Analysis of Series of Events". Chapman and Hall, London, United Kingdom.

8- Erkanli, A., Merrick, J. R. and Soyer, R., (2002), "Parametric and semi-parametric Bayesian models for accelerated life tests". Journal of Computational and Graphical Statistics, Institute of Mathematical Statistics, Vol. 11, No. 2, PP. 289–305. USA.

9- Leemis, L. M. (1991), "Non parametric estimation of the cumulative intensity function for a non-homogeneous Poisson process". Management Science, 37(7), 886-900.

10- Silvey, S. D. (1975),"Statistical Inference". Chapman and Hall, London.

11- Soyer, R. and Tarimcilar, M.M.,(2005)," Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach", Department of Management Science, The George Washington University, Washington, DOI: 10.1287/ mnsc.1070.0776 . Source: DBLP, Monroe Hall 403,Vol. 40, pp. 1-25,USA.

Use the PSO algorithm to estimate the Cox process parameters

Muhammad Zaid Hussain & Muthanna Subhi Suleiman
Department of Statistics and Informatics, College of Computer science and Mathematics,

University of Mosul, Mosul, Iraq

Abstract : In this paper it was proposed to use the PSO algorithm to estimate the time interval of the Cox process. The results of the proposed method of estimation were compared with the maximum likelihood estimator of estimating the rate of occurrence. The research included a realistic application in which the successive operating periods between two consecutive stops of the raw material mill in per day (from the General Company for Northern Cement / Cement Badush new plant) during the period from 1/4/2018 to 31/1/2019. The average time for machine operating periods was estimated by the methods proposed for use in the research.

Keywords: Cox process, PSO algorithm, maximum likelihood estimator, Cement Plant.