

## استخدام خوارزمية أمثلية سرب العناصر في ضبط معاملات انحدار العملية الغاوصية

مصطفى مظهر رنة\*\*

رائد قراحسن \*

(الإيداع: 26 آب 2018 ، القبول: 6 شباط 2019)

### الملخص

تُعدّ تقنية انحدار العملية الغاوصية (Gaussian Process Regression) (GPR) أحد أهم تقنيات التعلم الآلي، وقد أصبحت موضوعاً شائعاً للبحث في مجال التنبؤ منذ اقتراحها ضمن نظرية التعلم الآلي العالماني Rasmussen و Williams عام 1996 بحيث تمكنت هذه التقنية بنجاح من تحليل بيانات مختلف الأنظمة والتنبؤ بسلوكها بدقة عالية، ولكن تكمن المشكلة الأساسية في هذه التقنية باختيار معاملات انحدار العملية الغاوصية بسبب غياب منهج لهيكليّة عامة لضبط قيم هذه المعاملات حتى الآن. اقترحنا في هذا البحث طريقة جديدة في اختبار ضبط قيم معاملات انحدار العملية الغاوصية بالاعتماد على خوارزمية أمثلية سرب العناصر (PSO)، وتم التحقق من جودتها عبر تطبيقها على بيانات واقعية ومولدة. كما تم مقارنة هذه الطريقة مع طريقة الاختيار التحليلية أو التجريبية (AS) باستخدام مقاييس دقة التنبؤ (MSE، RMSE، MAPE)، بهدف الوصول أخيراً لأفضل طريقة لاختيار قيم هذه المعاملات. ودلّت نتائج البحث أن الطريقة المقترحة تعطي أفضل دقة تنبؤ عند استخدامها لاختيار قيم معاملات انحدار العملية الغاوصية.

الكلمات المفتاحية: انحدار العملية الغاوصية، خوارزمية أمثلية سرب العناصر، معاملات فوقية، مقاييس دقة التنبؤ.

\* طالب دراسات عليا (دكتوراه)-قسم الإحصاء الرياضي-كلية العلوم-جامعة حلب

\*\*أستاذ مساعد-قسم الإحصاء الرياضي-كلية العلوم-جامعة حلب

## Using Particle Swarm Optimization Algorithm To Adjust The Gaussian Process Regression Parameters

Raed Kara Hasan \*

Moustafa Mazhar Rene \*\*

(Received: 26 August 2019, Accepted: 6 February 2019)

### Abstract:

Gaussian Process Regression (GPR) is one of the most important techniques of automated learning, and has become a hot research subject in prediction tasks, which proposed by Williams and Rasmussen in 1996.

GPR can successfully analyze various data sets and generate predictions with high prediction accuracy. Nevertheless, the main challenge is the selection of GPR parameters. However, there is no generally structured way, yet.

In this study, we proposed new method in investigating the capability of GPR parameters using particle swarm optimization algorithm (PSO), and their quality was verified by applying them on both artificial and realistic data. This method was also compared with the analytical or experimental selection method (AS) using the measurements of prediction error explanation (MSE, RMSE, MAPE), in order to obtain the ideal method to select these parameters.

The results showed that the proposed method gives the best predictive accuracy when used to select parameters Gaussian Process Regression.

**Keywords:** Gaussian Process regression (GPR), Particle Swarm Optimization (PSO), Hyper-Parameters, the measurements of prediction error explanation.

---

\*Postgraduate Student (PhD)–Dept. of Mathematical Statistics –Faculty of Science–University of Aleppo

\*\* Assistant Professor–Dept. of Mathematical Statistics–Faculty of Science–University of Aleppo

## 1-مقدمة: Introduction

تم تطوير العديد من النماذج لمعالجة مسائل التنبؤ، كطريقة المربعات الصغرى، وانحدار النواة، ونماذج الشبكات العصبونية، وكان آخرها انحدار العملية الغاوصية، يستخدم انحدار العملية الغاوصية (Gaussian Process Regression) أو اختصاراً (GPR) في تقنيات التعلم الآلي وهي طريقة بديلة عن آلية انحدار متجه الدعم (SVR) والتي يمكن استخدامها للانحدار. قُدمت طريقة العملية الغاوصية كأداة للانحدار (Regression) في مجال التعلم الآلي، لأول مرة من قبل العالمين Rasmussen و Williams عام 1996 حيث قاموا بوصف تحسين المعلمات في دالة التباين والتي كانت مستوحاة من استخدام العملية الغاوصية مع الشبكات العصبونية، وقد تم استخدامها في تطبيقات مختلفة مثل التنبؤ بالنفوذية الجلدية من المواد الكيميائية والتنبؤ بتركيز الأوزون في الهواء (Bishop، 2007، Williams، 2006).

تعتمد معظم خوارزميات أو طرائق اختيار معاملات GPR على الخبرة وعمليات البحث العشوائي، وبالرغم من تطور خوارزميات البحث فإنه يبقى لتطبيق هذه الخوارزميات على هذه المسألة بعض المساوئ، فهي قد تحد من جودة الأداء ودقة التنبؤ لذلك تحتاج هذه المسألة إلى مزيد من الدراسة (Bishop، 2007).

تكمن أهمية البحث في تطوير طرائق التنبؤ كونها تشكل اللبنة الأساسية لعملية معالجة البيانات واستنباط المعارف فلا يمكن أن نبني المعرفة على أساس تنبؤات لا تمثل بشكل جيد قواعد البيانات المقدمة، كما أن تطبيقات هذه الطرائق كثيرة فهي تدخل في مجالات العلوم المختلفة وتطويرها يساعد على تقدم عملية البحث العلمي.

يهدف هذا البحث بشكل أساسي إلى دراسة وتحسين أداء انحدار العملية الغاوصية عبر الوصول إلى الأسلوب الأفضل لضبط قيم معاملات هذه التقنية.

تكمن مشكلة اختيار معاملات انحدار العملية الغاوصية الأساسية بغياب منهجية عامة لضبط قيم هذه المعاملات، وذلك بسبب صعوبة دراسة المسألة تحليلياً.

تم في هذا البحث اقتراح طريقة جديدة في اختبار ضبط قيم معاملات انحدار العملية الغاوصية بالاعتماد على خوارزمية أمثلية سرب العناصر PSO (Particle Swarm Optimization)، بهدف الوصول لأفضل طريقة لاختيار قيم هذه المعاملات.

## 2-المواد وطرائق البحث: Materials and Methods

### 2-1- مفهوم انحدار العملية الغاوصية:

ليكن لدينا  $f = (f(x_1), f(x_2), f(x_3), \dots, f(x_d))^T$  متجه ذو  $d$  بُعد من الدوال عندئذ تسمى العملية العشوائية  $\{f(x): x \in \chi\}$  بعملية غاوص بحيث أن  $\chi$  هو فضاء المدخلات إذا كان متجه المتغيرات العشوائية  $X_1, X_2, \dots, X_d$  يتوزع وفق التوزيع الطبيعي المتعدد بمتوسط  $\mu$  و مصفوفة تباين  $K$ ، تُعرف عملية غاوص كتوزيع على الدوال  $P(f(x))$  بحيث

$$f: \chi \rightarrow \mathbb{R} \quad \text{أن } f(x) \text{ هي دالة معرفة على فضاء المدخلات } \chi \text{ كما يلي:}$$

أي أن العملية الغاوصية هي مجموعة من المتغيرات العشوائية المستمرة محدودة الأبعاد والتي كل منها يخضع للتوزيع الطبيعي وتكون جميع توزيعاتها هي توزيعات طبيعية، وتعتبر عملية غاوص (GP) من أهم تقنيات التعلم الآلي لإيجاد نماذج الانحدار غير الخطية (Rasmussen و Williams، 2006؛ Liu و زملاؤه، 2017).

لنكن لدينا  $\mu(x)$  دالة متوسط و  $k(x, x')$  دالة تباين معرفتان كما يلي:

$$\mu(x) = E[f(x)]$$

$$k(x, x') = Cov(f(x), f(x')) = E[(f(x) - \mu(x))(f(x') - \mu(x')))]$$

بحيث  $x, x' \in \chi$  عندئذ العملية الغاوصية (GP) تأخذ الشكل التالي:

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_d) \end{bmatrix} \sim N_d \left( \begin{bmatrix} \mu(x_1) \\ \vdots \\ \mu(x_d) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_d) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_d) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_d, x_1) & k(x_d, x_2) & \cdots & k(x_d, x_d) \end{bmatrix} \right) \quad (1)$$

ونرمز لذلك بالرمز:

$$P(f(x)) = \mathcal{GP}(\mu(x), k(x, x')) \quad (2)$$

نسمي الدالة  $k(x, x')$  بدالة التباين أو دالة النواة (نواة التباين) وهي دالة موجبة محدودة ولها عدة أنواع (Bishop, 2007).  
ليكن  $Y$  متغير تابع و  $X$  متغيرات عشوائية ذو  $d$  بُعد، يعطى نموذج الانحدار اللامعلمي وفق العلاقة:

$$y = f(x) + \varepsilon \quad , \varepsilon \sim N(\mathbf{0}, \sigma^2) \quad (3)$$

بحيث أن:  $f(x)$  هي دالة مجهولة أما في الانحدار المعلمي تكون معلومة، تعاني الطرائق اللامعلمية من مشكلة تعدد الأبعاد (curse of dimensionality) عندما يتم تطبيقها مع المتغيرات المتعددة (أي عندما تكون  $d$  كبيرة)، لقد تم تطوير مجموعة متنوعة من النماذج البديلة للتغلب على هذه المشكلة منها نموذج انحدار العملية الغاوسية (GPR) (Gaussian process regression model).

إن نموذج انحدار العملية الغاوسية هو نموذج لامعلمي، وهذا يعني بأنه لا يفترض شكل معين للدالة المدروسة ولكن يتم تحديد شكل العلاقة بين المدخلات والأهداف بالكامل من خلال البيانات التي قد تتضمن عدد غير محدود من الدوال، وتكون الدالة الأساسية التي تنتج البيانات مجهولة ولكن يتم توليد التنبؤات من خلال مجموعة من الدوال التي تخضع لتوزيع غاوص في فضاء الدوال، ويعتبر نموذج انحدار العملية الغاوسية من أحدث طرائق التنبؤ، وهو من نماذج بايز الاحتمالية، ففي معظم طرائق انحدار بايز يتم إيجاد معلومات مسبقة عن معلمات النموذج، وبعد ذلك يتم وضع شروط على البيانات لإعطاء معلمات النموذج اللاحق (البعدي)، حيث يمكن صياغة هذه المعلومات المسبقة بشكل توزيع احتمالي يسمى التوزيع القبلي و يحدد نموذج بايز المعلمات المجهولة للنموذج القبلي بينما يحدد نموذج عملية غاوص علاقات الدوال القبلية مباشرة بين مدخلات الاختبار ومدخلات ومخرجات التدريب (Rasmussen و Williams، 2006؛ Liu وزملاؤه، 2017).

لنفترض لدينا مجموعة من البيانات  $\{(x_i, y_i)\}_{i=1}^n$  بحيث تشير  $x_i \in \mathbb{R}^d$  إلى المدخلات والتي لها  $d$  بُعد وتشير  $y_i \in \mathbb{R}$  إلى القيم الحقيقية للناتج و  $n$  إلى عدد البيانات، عندئذ يأخذ نموذج انحدار العملية الغاوسية (GPR) الشكل التالي:

$$y_i = f(x_i) + \varepsilon_i \quad ; i = 1, \dots, n \quad , \varepsilon_i \sim N(\mathbf{0}, \sigma^2) \quad (4)$$

بحيث أن:  $f(x) \sim \mathcal{GP}(\mu(x), k(x, x'))$  و  $\mathcal{GP}(\mu(x), k(x, x'))$  هي عملية غاوص قبلية مع دالة متوسط  $\mu(x)$  ودالة تباين  $k(x, x')$  وبالتالي يعطى نموذج انحدار العملية الغاوسية وفق العلاقة:

$$y = \mathcal{GP}(\mu(x), k(x, x')) + \sigma_n^2 \delta(x, x') \quad (5)$$

بحيث أن:  $\delta(x, x')$  دالة دلتا كرونكير (Kronecker delta) و  $\delta(x, x') = 0$  عندما  $x \neq x'$  و  $\delta(x, x') = 1$  عندما  $x = x'$  و  $\sigma_n^2$  تباين الضجيج العشوائي ومن الشائع أيضاً أن نفترض  $\mu(x) = 0$  أي (دالة المتوسط للعملية الغاوسية قبلية معدومة) عندئذ يأخذ نموذج انحدار العملية الغاوسية الشكل التالي:

$$y \sim \mathcal{GP}(0, k(x, x')) + \sigma_n^2 \delta(x, x') \quad (6)$$

تم تصميم مجموعة متنوعة من دوال النواة، وسيتم في هذا البحث استخدام دالة النواة الغاوسية والموضحة وفق العلاقة الآتية:

$$k(x, x') = e^{-\frac{\|x-x'\|^2}{2\sigma^2}} \quad (7)$$

بحيث أن:  $\|x - x'\| = \sqrt{(x - x')^T(x - x')}$  تشير إلى طول الشعاع  $(x - x')$  أو نظيم الفرق بين قيمتين  $x, x'$  و  $\sigma$  معامل دالة نواة غاوص (Rasmussen و Williams، 2006).

## 2-2- الطريقة التحليلية في ضبط معاملات انحدار العملية الغاوصية:

تعتمد طريقة الاختيار التحليلي أو التجريبي (Analytic Selection) لضبط قيمة معامل الضجيج العشوائي  $\sigma_n$  ومعامل دالة النواة  $\sigma$  مباشرة من بيانات التدريب. وتتصف هذه الطريقة بسهولة حسابها وتعطي نتائج مقبولة، وتعطي قيمة المعامل  $\sigma_n$  ومعامل دالة النواة  $\sigma$  وفق الطريقة التحليلية بالعلاقات التالية:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

$$\sigma_n = \sqrt{\widehat{\sigma}_n^2}$$

و  $\widehat{\sigma}_n$  هو تقدير مستوى الضجيج لبيانات التدريب ويتم حسابه عبر الصيغة المقدمة وفق العلاقة:

$$\widehat{\sigma}_n^2 = \frac{n^{\frac{1}{5}}k}{n^{\frac{1}{5}}k - 1} \cdot \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

بحيث أن  $\hat{y}_i$  هي قيم الانحدار بواسطة طريقة الجار الأقرب و  $k$  معامل يمثل عدد نقاط الجوار في خوارزمية الجار الأقرب ويأخذ عادة إحدى القيم  $\{3, 5, 7\}$  ( Bishop، 2007؛ Rasmussen و Williams، 2006؛ Shi و Choi، 2011). تم ضبط القيمة  $k = 5$  في هذا البحث ونقترح في هذا البحث استخدام أطول مسافة إقليدية بين قيم  $x$  للاستعاضة عن مدى البيانات الذي لا يمكن حسابه في الحالة متعددة الأبعاد.

## 2-3- الطريقة المقترحة في ضبط معاملات انحدار العملية الغاوصية:

إن اختيار المعلمات الفوقية (Hyper-Parameters) في تقنية انحدار العملية الغوصية يلعب دوراً هاماً في الحصول على نموذج جيد (Shi و Choi، 2011).

قمنا في هذا البحث باقتراح استخدام خوارزمية أمثلية سرب العناصر (Particle Swarm Optimization) في اختيار ضبط قيم معاملات انحدار العملية الغاوصية وذلك بهدف الوصول لأفضل طريقة لضبط قيم هذه المعاملات.

تعتبر خوارزمية أمثلة سرب العناصر (Particle Swarm Optimization) PSO إحدى أكثر خوارزميات الذكاء الاصطناعي تطوراً وتستخدم لحل مسائل الأمثلية بشكل عام، وتحاكي السلوك الطبيعي الموجود في أسراب الطيور أو تجمعات الأسماك أثناء بحثها عن مصادر الطعام (Aich و Banerjee، 2014؛ Hu و زملاؤه، 2010).

يتم في هذه الخوارزمية أخذ مجموعة ابتدائية عشوائياً من العناصر تدعى سرب وتقوم عناصر هذا السرب بتحديث مواضعها  $P_i$  وسرعاتها  $V_i$  بالاعتماد على المعلومات التي يجمعها السرب. يؤثر كل عنصر إضافة إلى تأثير كامل السرب على توجيهه كامل السرب نحو المنطقة المثلى.

ويتم التحكم بمعدل التقارب باستخدام عدة معاملات وهي:

1- معامل القصور الذاتي  $\omega$  (Inertia Factor) يتحكم بتأثير شعاع السرعة السابق لأحد العناصر على شعاع السرعة الحالي لهذا العنصر.

2- معامل التقلص  $\psi$  (Constriction Factor) يحدد تأثير أشعة سرعة العناصر المجاورة على الموضع الحالي للعنصر ويهدف لتجنب تجمع العناصر حول نقطة واحدة من فضاء البحث لتجنب الوقوع في الحلول المثلى المحلية.

3- معامل التسارع المعرفي (Cognitive Acceleration)  $c_1$ ، ويتحكم بتأثير تجربة الأفراد عبر أفضل موقع مرّ عليه هذا العنصر.

4- معامل التسارع الاجتماعي (Social Acceleration)  $c_2$  ويتحكم بتأثير كامل السرب عبر أفضل موضع وصل إليه أحد العناصر بشكل شامل حتى الآن (Aich و Banerjee، 2014). كما أن لاختيار عدد عناصر السرب أهمية كبيرة فعند اختيار عدد قليل لعناصر السرب لن يتمكن من جمع معلومات كافية عن كامل منطقة البحث، كما أن اختيار عدد كبير لعناصر السرب سوف يؤدي إلى استهلاك وقت أطول للتقارب من منطقة الحل الأمثل، ويجب التنويه على أهمية التوزع العشوائي لعناصر السرب الابتدائي لتنتشر على كامل منطقة البحث المفترضة وذلك لتغطية كامل منطقة البحث، ويتم اختيار العدد الأعظمي للتكرارات اعتماداً على التجربة، فمن المهم الوصول إلى التقارب وإلا فإننا لن نحصل على نتائج جيدة، كما أن التكرارات التي تلي الوصول إلى الحل لا فائدة منها بسبب الوصول إلى نقطة التوازن وتكون أشعة السرعة لكل العناصر معدومة ويتم تعيين الحد الأعظمي للتكرارات كشرط للتوقف (Mullen، 2014).

### 2-3-1-خطوات خوارزمية PSO لاختيار معاملي GPR $(\sigma_n, \sigma)$ :

**الخطوة الأولى:** تحديد مجالات معاملات التحكم الداخلية وتم استخدام القيم القياسية ضمن الحزمة البرمجية المستخدمة وتحديد عدد عناصر السرب وتم ضبطه في هذا البحث إلى  $m = 30$  والعدد الأعظمي للتكرارات  $iter_{max} = 100$ .

**الخطوة الثانية:** من أجل التكرار الأول  $iter = 1$  يتم توليد قيم أشعة المواضع الابتدائية لعناصر السرب بشكل عشوائي ضمن مجالات منطقة البحث (بحيث يمثل كل شعاع موضع قيم معاملي GPR  $(\sigma_n, \sigma)$ ) ويتم حساب أشعة السرعة الابتدائية لكل عنصر باستخدام العلاقة التالية

$$V_{iter}^{(j)} = rang^{(j)} * rand \quad ; j = 1,2,3 \quad (10)$$

بحيث يشير  $rang^{(j)}$  إلى مجالات المعاملات  $\sigma_n, \sigma$  ويشير  $rand$  إلى عدد عشوائي ضمن المجال  $[0,1]$  ويفيد الدليل  $j$  إلى أن الحسابات سوف تتكرر ثلاثة مرات لتوليد أشعة السرعة الابتدائية لمعاملي GPR  $(\sigma_n, \sigma)$  وهكذا يتم توليد  $m$  شعاع سرعة مع المكونات الموافقة للمعاملات  $\sigma_n, \sigma$  المراد حسابها.

**الخطوة الثالثة:** ضبط  $t = 1$  و  $Pbest^t$  و  $Gbest$  تساوي شعاع الموضع الحالي.

**الخطوة الرابعة:** إذا كان  $t = m$  ننقل إلى الخطوة الثامنة وإلا نضع  $t = t + 1$  وننقل إلى الخطوة الخامسة.

**الخطوة الخامسة:** حساب مقاييس الأخطاء الثلاثة (MAPE، RMSE،MSE) المستخدمة عند الموضع  $t$  في السرب وفي حالة  $t = 1$  نحدد  $Pbest^t$  إلى شعاع الموضع الحالي وننقل إلى الخطوة السابعة وإلا ننقل إلى الخطوة السادسة.

**الخطوة السادسة:** إذا كانت قيمة مقياس الخطأ المستخدم أفضل من مقياس الخطأ للموضع  $Pbest^t$  الحالية نقوم بتحديث شعاع  $Pbest^t$  مع شعاع الموضع الحالي وإلا تبقى قيمة  $Pbest^t$  كما هي.

**الخطوة السابعة:** إذا كانت قيمة مقياس الخطأ المستخدم أفضل من مقياس الخطأ للموضع  $Gbest$  الحالية نقوم بتحديث شعاع  $Gbest$  مع شعاع الموضع الحالي وإلا تبقى قيمة  $Gbest$  كما هي وفي كلا الحالتين ننقل إلى الخطوة الرابعة.

الخطوة الثامنة: نقوم بإعادة ضبط معاملات التحكم الداخلية للتكرار الحالي رقم  $iter$  عبر العلاقات التالية

$$\omega_{iter} = \omega_{intial} + (\omega_{final} - \omega_{intial})(iter - 1)/(iter_{max} - 1)$$

$$\psi_{iter} = \psi_{intial} + (\psi_{final} - \psi_{intial})(iter - 1)/(iter_{max} - 1)$$

$$c_{1,iter} = c_{1,intial} + (c_{1,final} - c_{1,intial})(iter - 1)/(iter_{max} - 1)$$

$$c_{2,iter} = c_{2,intial} + (c_{2,final} - c_{2,intial})(iter - 1)/(iter_{max} - 1)$$

بحيث أن  $iter_{max}$  هو الحد الأعظمي للتكرارات ، كما يتم تحديث أشعة السرعة لكافة العناصر باستخدام العلاقة

$$V_{(j),iter+1}^k = \omega_{iter}(V_{(j),iter}^k) + c_{1,iter}(rand)(Pbest_{(j)}^k - P_{(j),iter}^k) + c_{2,iter}(rand)(Pbest_{(j)}^k - P_{(j),iter}^k) ; j = 1,2,3$$

بحيث أن  $V_{(j),iter}^k$  شعاع السرعة للمركبة  $j$  (مركبة المعاملات) من العنصر  $k$  في التكرار رقم  $iter$  وبحيث أن  $P_{(j),iter}^k$  شعاع الموضع للمركبة  $j$  (مركبة المعاملات) من العنصر  $k$  في التكرار رقم  $iter$  وبحيث أن  $j$  عداد يكرر الحسابات ثلاث مرات لتوليد أشعة السرعة لمعاملي GPR  $(\sigma_n, \sigma)$ .

تجدر الملاحظة أنه يمكن تقسيم العلاقة السابقة إلى مجموع ثلاثة حدود بحيث يمثل الحد الأول مركبة الذاكرة (Memorial Component) ويمثل الحد الثاني المركبة المعرفية (Cognitive Component) بينما يمثل الحد الثالث المركبة الاجتماعية (Social Component) أي أن شعاع السرعة الجديد يتأثر بالعوامل الثلاثة الذاكرة والمعرفة المكتسبة وذكاء السرب (Hu وزملاؤه، 2010).

ويتم بعدها تحديث أشعة الموضع لكل عناصر السرب باستخدام العلاقة

$$P_{(j),iter+1}^k = P_{(j),iter}^k + \psi_{iter}(V_{(j),iter+1}^k) ; j = 1,2,3 \quad (11)$$

وفي حال كان الموضع الجديد للمكونات الموافقة للمعاملين  $(\sigma_n, \sigma)$  يقع خارج منطقة البحث المحددة يتم إعادة الموضع الحالي إلى حدود منطقة البحث.

الخطوة التاسعة: إذا كان  $iter < iter_{max}$  فإننا نضع  $iter = iter + 1$  ونضع  $t = 1$  وننتقل إلى الخطوة الخامسة. أما في حالة  $iter = iter_{max}$  يتم اعتماد قيمة  $Gbest$  الحالية كحل أمثل أي أن مركبات الشعاع  $Gbest$  أصبحت تمثل القيم المثلى للمعاملين  $\sigma_n, \sigma$  وهي التي تجعل قيمة مقياس الخطأ المستخدم أصغر ما يمكن عند تطبيق خوارزمية GPR على بيانات التدريب.

#### 4-2- الجانب التطبيقي:

بغرض اختبار أداء الطريقة المقترحة في ضبط معاملي انحدار العملية الغاوصية  $(\sigma_n, \sigma)$  قُمنّا بتطبيق الطريقة التحليلية AS وطريقة خوارزمية سرب العناصر PSO على مجموعتي بيانات واقعية ومولدة.

#### 4-2-1 البيانات الواقعية:

تم جمع البيانات الواقعية من المشفى الوطني في مدينة حماة والتي تمثل أوزان وأطوال خمسين طفلاً مولودين حديثاً وعمر الأم وعدد أيام فترة الحمل وعدد الأطفال المنجبين للأم قبل هذا الطفل (عدد الولادات السابقة).

يهدف نموذج التنبؤ المراد بناءه إلى تقدير وزن الطفل عن طريق معرفة عمر الأم وعدد أيام فتره الحمل وعدد الأطفال المنجبين للأم قبل هذا الطفل بحيث يمثل وزن الطفل المتغير التابع  $Y$ ، بينما يمثل عمر الأم المتغير الأول المستقل  $X_1$  ويمثل عدد أيام فترة الحمل المتغير الثاني المستقل  $X_2$  وعدد الأطفال المنجبين للأم قبل هذا الطفل يمثل المتغير الثالث المستقل  $X_3$ .

## 2-4-2 البيانات المولدة:

تم توليد مجموعة البيانات المولدة بحجم 100 باستخدام دالة من النمط  $sinc$  معطاة وفق العلاقة (Hu وزملاؤه، 2010):

$$y = sinc(x) = \begin{cases} 1 & ; x = 0 \\ \frac{\sin x}{x} & ; otherwise \end{cases} \quad (12)$$

وتم إضافة ضجيج عشوائي (تشويش) وفق التوزيع الطبيعي بتوقع رياضي معدوم وانحراف معياري قدره  $0.1\sigma$ .

## 2-4-3-منهجية اختبار الطريقة المقترحة:

تعتبر الدقة معياراً لاختيار النموذج الأمثل للتنبؤ، ويُقصد بالدقة قدرة نموذج التنبؤ على إعادة إنتاج البيانات الأصل للعينة المدروسة، ومنه فإن الاختيار المناسب لمقياس دقة التنبؤ يؤثر إيجاباً في تحديد فاعلية نموذج التنبؤ المستخدم وتعمل مقاييس دقة التنبؤ القياسية بشكل عام على مفهوم الفرق بين القيم الأصلية والقيم المتوقعة أو المُتنبئ بها، وهو ما ندعوه بخطأ التنبؤ، وكلما كان مقدار الفرق قليلاً كانت دالة التنبؤ أفضل وأدق، يوجد العديد من مقاييس دقة التنبؤ، وعادة لا يتم الاعتماد على مقياس واحد في عملية ضبط نموذج التنبؤ.

تعتبر المقاييس التالية:  $MSE$ ,  $RMSE$ ,  $MAPE$  من أفضل مقاييس المقارنة بين نماذج تنبؤ مختلفة تم بناؤها باستخدام نفس مجموعة بيانات التدريب (Hyndman و Koehler، 2006).

وقد اعتمدنا في بحثنا هذا على هذه المقاييس كونها تلائم طبيعة البحث ويتم حسابهم كما يلي:

$$MSE = \text{mean}(e_i^2) \quad \text{متوسط مربعات الأخطاء} \quad \text{Mean squared error}$$

$$RMSE = \sqrt{MSE} \quad \text{جذر متوسط مربعات الأخطاء} \quad \text{Root mean squared error}$$

$$MAPE = \text{mean}(|e_i/y_i|) \quad \text{متوسط القيم المطلقة} \quad \text{Mean absolute percentage error}$$

للأخطاء النسبية

$$e_i = y_i - \hat{y}_i \quad \text{بحيث يحسب خطأ التنبؤ من العلاقة:}$$

(يُدعى خطأ التنبؤ في بعض المراجع بالبقايا residuals) (Hardle وزملاؤه، 2004).

بهدف الحكم على جودة أداء كل طريقة من طرائق التنبؤ تم حساب قيم هذه المقاييس بطريقتين: الأولى باستخدام كامل مجموعة بيانات التدريب، والثانية بتطبيق منهجية التصحيح المتقاطع k-fold Cross Validate لقياس دقة النموذج الناتج باستخدام المعاملات المحسوبة بكل طريقة بحيث تقسم هذه المنهجية البيانات إلى k قسم. ويتم اختيار أحد الأقسام بكل مرة كعينة اختبار والبقية كعينات التدريب، يتم حساب مقياس الخطأ باستخدام عينة الاختبار في كل مرة. وفي النهاية يتم حساب متوسطات قيم مقياس الخطأ التي تم الحصول عليها بكل دورة من الدورات. يتم ضبط قيمة k عادة بإحدى القيم 3 أو 5 أو 10 (Hyndman و Koehler، 2006)، وفي بحثنا هذا اعتمدنا القيمة  $k=10$ .

وبهدف تطبيق الطريقتين تم كتابة برنامج نصي باللغة البرمجية R (لغة برمجية إحصائية)، حيث تستطيع هذه اللغة القيام بالعديد من تحليلات البيانات بحيث يتم تنظيم هذه التحليلات ضمن ما يسمى بالحزم Packages مما يعني قدرة الباحثين على تطوير البرامج المختلفة الأمر الذي ساهم بانتشار استخدامها في المجالات الأكاديمية (Cotton، 2013، Matloff، 2011).



لتنفيذ هذا البحث تم الاستفادة من الحزم التالية:

الحزمة caret بهدف تطبيق خوارزمية الجار الأقرب لتقدير قيمة الضجيج وحساب مصفوفة المسافات الإقليدية لحساب معاملات انحدار العملية الغاوصية بالطريقة التحليلية، والحزمة kernlab بهدف تطبيق انحدار العملية الغاوصية والحزمة hydroPSO بهدف تطبيق خوارزمية أمثلية سرب العناصر وتم برمجة منهجية التصحيح المتقاطع ولم يتم الاعتماد على أي حزمة بهدف تطبيق أكثر من مقياس خطأ.

### 3-النتائج ومناقشتها:

سنقوم الآن بعرض النتائج التطبيقية على كل من مجموعتي البيانات:

بعد التطبيق العملي ظهرت لدينا النتائج الموضحة بالجدول والأشكال كما يلي:

الجدول رقم (1) قيم المعاملات المستخدمة ضمن مجموعة البيانات الواقعية

أسلوب حساب المعاملات	قيم المعاملات	
	$\sigma_n$	$\sigma$
الطريقة التحليلية AS	0.2238204	118.7886
الطريقة المقترحة PSO	0.06710821	0.22939607

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

بعد تطبيق الطريقة التحليلية والطريقة المقترحة وفق المعاملات الموضحة بالجدول رقم (1) على مجموعة بيانات الأطفال الواقعية وحساب قيم مقاييس الأخطاء لكل طريقة ظهرت لدينا النتائج التالية:

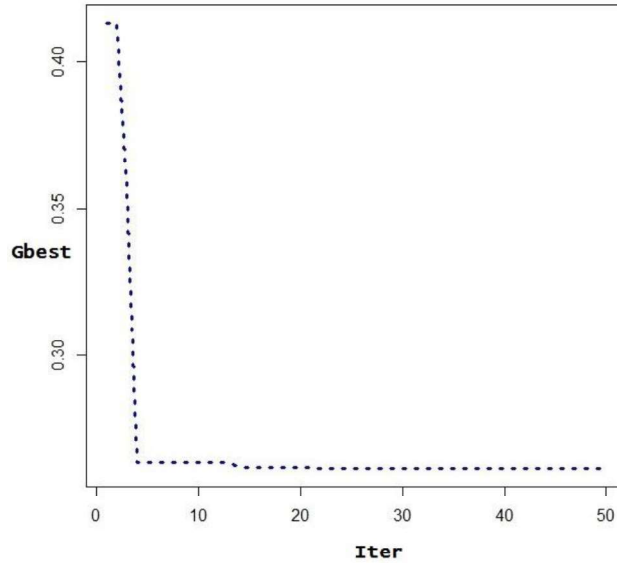
الجدول رقم (2): نتائج تطبيق الطريقة التحليلية والطريقة المقترحة على مجموعة البيانات الواقعية

	الطريقة التحليلية AS			الطريقة المقترحة PSO		
	MSE	RMSE	MAPE	MSE	RMSE	MAPE
10-CV التصحيح المتقاطع	0.641398 1	0.764291 8	0.326710 2	0.1821091	0.4242735	0.151673
All Data كامل البيانات	0.484555 6	0.696100 3	0.270241 1	0.0868622	0.2947241	0.0875836

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

يُظهر الجدول رقم (2) من اليسار إلى اليمين اسم الطريقة المستخدمة لتقدير قيم المعاملات وقيم مقاييس الأخطاء على مجموعة البيانات الواقعية، حيث أن قيم مقاييس الأخطاء المحسوبة بطريقة التصحيح المتقاطع بعشرة حقول (وهي متوسط قيم

الأخطاء في كل تكرار) في السطر الأول، وقيم مقاييس الأخطاء المحسوبة باستخدام كامل مجموعة بيانات الأطفال في السطر الثاني.



الشكل رقم (1): معدل تقارب خوارزمية أمثلية سرب العناصر PSO المقترحة من الحل في حالة البيانات الواقعية  
 كما يُظهر الشكل رقم (1) معدل تقارب خوارزمية أمثلية سرب العناصر PSO المقترحة لضبط معاملي *GPR* نحو أفضل قيمة ضمن 50 تكرار. وتم تمثيل تقارب هذه الخوارزمية بالخط باللون الأزرق المنقط، بحيث يمثل المحور الأفقي دليل التكرارات ويمثل المحور الشاقولي أفضل ملائمة ضمن التكرار. وتم الرسم باستخدام لغة R باستخدام التعليمتين *plot* و *lines*.  
 نلاحظ من الجدول رقم (2) بأنه كان للطريقة المقترحة قيماً أصغر لمقاييس الأخطاء الثلاثة (*MAPE*، *RMSE*، *MSE*) وذلك سواءً باستخدام كامل مجموعة البيانات الواقعية لحساب الأخطاء أو باستخدام منهجية التصحيح المنقطع.  
 كما نلاحظ أيضاً من خلال الشكل رقم (1) فعالية خوارزمية أمثلية سرب العناصر PSO المقترحة في تقدير معاملي *GPR* في حالة البيانات الواقعية بسبب قدرتها على التقارب والوصول للحل الأمثل بأقل عدد من التكرارات حيث تقاربت من أفضل ملائمة بالخطوة 13 بحيث اعتمدنا لقياس سرعة تقارب الخوارزمية من الحل على زمن تنفيذ الخوارزمية بالثواني إضافة إلى رقم تكرار التقارب (التكرار الذي تم الوصول فيه لأفضل قيم للمعاملات) ونذكر هنا بأنه تم ضبط الحد الأعظمي للتكرارات لهذه الخوارزمية إلى 50 تكرار، ولا تتطلب الطريقة التحليلية أي تكرار.  
 وبإعادة الحسابات بالنسبة لمجموعة البيانات المولدة ظهرت لدينا النتائج التالية:

الجدول رقم (3) قيم المعاملات المستخدمة ضمن مجموعة البيانات المولدة

أسلوب حساب المعاملات	قيم المعاملات	
	$\sigma_n$	$\sigma$
الطريقة التحليلية AS	2	5.860907
الطريقة المقترحة PSO	0.08873033	3.36093609

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

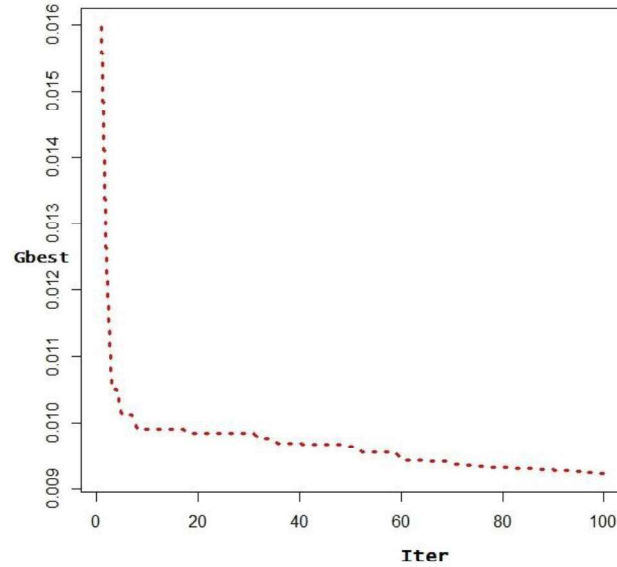
بعد تطبيق الطريقة التحليلية والطريقة المقترحة وفق المعاملات الموضحة بالجدول رقم (3) على مجموعة البيانات المولدة وحساب قيم مقاييس الأخطاء لكل طريقة ظهرت لدينا النتائج التالية:

الجدول رقم (4): نتائج تطبيق الطريقة التحليلية والطريقة المقترحة على مجموعة البيانات المولدة

	الطريقة التحليلية AS			الطريقة المقترحة PSO		
	MSE	RMSE	MAPE	MSE	RMSE	MAPE
10-CV التصحيح المتقاطع	0.011423 1	0.106879 3	2.089634	0.0092290	0.0925018	1.191489
All Data كامل البيانات	0.010411 2	0.102035 6	0.821003 5	0.0073840	0.0859307	1.073861

المصدر: من إعداد الباحث اعتماداً على مخرجات اللغة البرمجية R

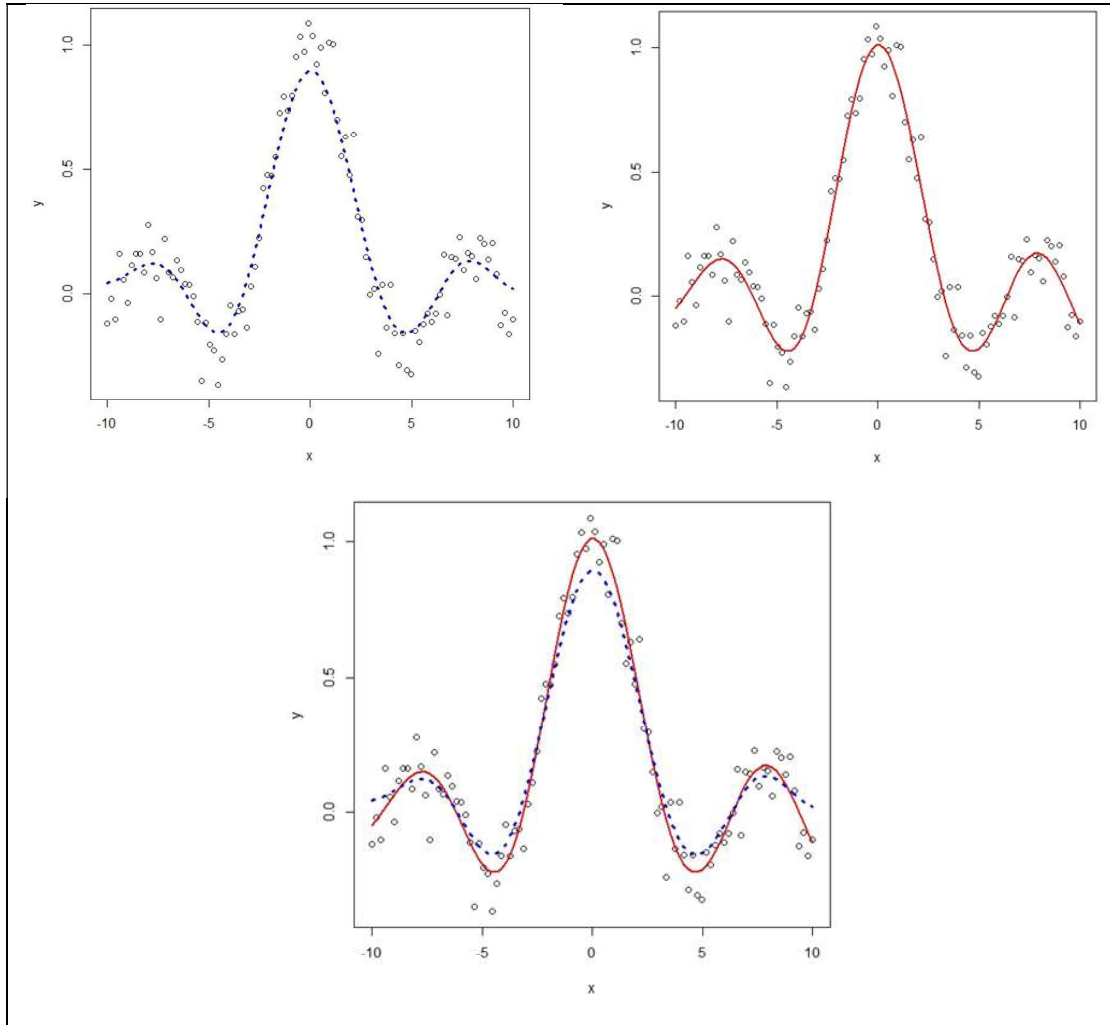
يُظهر الجدول رقم (4) من اليسار إلى اليمين اسم الطريقة المستخدمة لتقدير قيم المعاملات وقيم مقاييس الأخطاء على مجموعة البيانات المولدة، حيث أن قيم مقاييس الأخطاء المحسوبة بطريقة التصحيح المتقاطع بعشرة حقول (وهي متوسط قيم الأخطاء في كل تكرار) في السطر الأول، وقيم مقاييس الأخطاء المحسوبة باستخدام كامل مجموعة البيانات المولدة في السطر الثاني.



الشكل رقم (2): معدل تقارب خوارزمية أمثلية سرب العناصر PSO المقترحة من الحل في حالة البيانات المولدة

باستخدام الدالة Sinc مع إضافة تشويش  $\rho=0.1$

كما يظهر الشكل رقم (2) معدل تقارب خوارزمية أمثلية سرب العناصر PSO المقترحة في حالة البيانات المولدة باستخدام الدالة Sinc مع إضافة تشويش  $\rho=0.1$  نحو أفضل قيمة ضمن 100 تكرار. بحيث تم تمثيل تقارب خوارزمية PSO بالخط باللون الأحمر المنقط و يمثل المحور الأفقي دليل التكرارات ويمثل المحور الشاقولي أفضل ملائمة ضمن التكرار. ويُظهر الشكل رقم (3) نموذجي انحدار GPR باستخدام المعاملين  $(\sigma_n, \sigma)$  الناتجين من خلال خوارزمية أمثلية سرب العناصر على مجموعة البيانات المولدة باستخدام الدالة Sinc مع إضافة تشويش  $\rho = 0.1$  على اليمين ومن خلال الطريقة التحليلية على اليسار بحيث تم تمثيل خط انحدار العملية الغاوصية GPR باستخدام المعاملات المحسوبة بالطريقة التحليلية باللون الأزرق المنقط بالإضافة إلى الخط باللون الأحمر المتصل والذي يمثل خط انحدار العملية الغاوصية GPR باستخدام المعاملات المحسوبة بخوارزمية أمثلية سرب العناصر PSO المقترحة.



الشكل رقم (3): نمذجي انحدار *GPR* باستخدام المعاملين  $(\sigma_n, \sigma)$  الناتجين من خلال خوارزمية أمثلية سرب العناصر و الطريقة التحليلية على مجموعة البيانات المولدة باستخدام الدالة *Sinc* مع إضافة تشويش  $\rho = 0.1$

4-الاستنتاجات والتوصيات:

#### 4-1-الاستنتاجات: Conclusions

من خلال ما ذكر سابقاً وما سُجل من نتائج نورد ما يلي:

1-أظهرت الدراسة التي أجريناها على مجموعتي البيانات الواقعية والمولدة تفوق الطريقة المقترحة والمتمثلة باستخدام خوارزمية أمثلية سرب العناصر PSO في ضبط معاملات انحدار العملية الغاوصية *GPR* على الطريقة التحليلية التجريبية AS وذلك من خلال تحقيقها لأصغر قيمة من قيم مقاييس الأخطاء الثلاثة (MSE، MAPE، RMSE) سواء باستخدام كامل مجموعتي البيانات الواقعية والمولدة لحساب الأخطاء أو باستخدام منهجية التصحيح المتقاطع.

2-من خلال نتائج هذه الدراسة تبين فعالية خوارزمية أمثلية سرب العناصر PSO المقترحة في ضبط معاملي *GPR* في حالة البيانات الواقعية والمولدة بسبب قدرتها على التقارب والوصول للحل الأمثل.

3- أيضاً تم التوصل إلى أن خوارزمية أمثلية سرب العناصر PSO تعطي أفضل دقة تتبؤ وأفضل نتائج وأسرع تقارب من الحل وذلك لتكرار عدد الأفضلية بالاعتماد على أصغر قيمة من قيم مقاييس الأخطاء المستخدمة وبسبب قدرة منحني الانحدار المثل لها على ملائمة وتمثيل البيانات بشكل أفضل من الطريقة التحليلية.

#### 4-2-التوصيات: Recommendations

1- مما سبق نوصي باستخدام خوارزمية PSO المقترحة في ضبط معاملي GPR كونها تعطي أفضل النتائج، كما نوصي ببرمجة هذه الطريقة باستخدام البرمجة المتوازية Parallel Programming بحيث يتم تنفيذ العديد من العمليات في وقت واحد والتي تعتمد على مبدأ أنه يمكن في كثير من الأحيان تقسيم المشاكل الكبيرة إلى مشاكل أصغر حجماً ليتم حلها بشكلٍ متوازٍ في الوقت نفسه لتقليل الزمن اللازم لتنفيذها، ونوصي باعتماد شرط توقف إضافي عند عدم وجود تحسن في قيمة الملائمة بعد عدد محدد من التكرارات للاستفادة من خاصية سرعة التقارب في خوارزمية PSO .

2- نوصي الباحثين في المجال الصحي أو الطبي والأطباء العاملون في المشافي أو مديريات الصحة باستخدام خوارزمية أمثلية سرب العناصر PSO المقترحة في ضبط معاملي GPR لأنها أعطت أفضل النتائج وتطبيقها في المجال الصحي أو الطبي.

3- إجراء دراسات مستقبلية حول ضبط معاملي GPR باستخدام طرائق أخرى غير الطريقة المقترحة في هذا البحث.

## 5-References:

- 1- **Aich, U., and Banerjee, S., (2014).** Modeling of EDM Responses by Support Vector Machine Regression with Parameters Selected by Particle Swarm Optimization, Applied Mathematical Modelling, 38 2800–2818.
- 2- **Bishop, C. M., (2007).**Pattern Recognition and Machine Learning. Springer.
- 3- **Cotton, R., (2013).** Learning R, O'Reilly Media, Inc., United States of America, 377.
- 4- **Hardle, W., Muller, M., Sperlich, S.,and Werwatz A., (2004)** .Nonparametric and Semiparametric Models, Springer, Berlin, 301.
- 5- **Hu, Z.,Min, W, Huang, X., (2010).**Parameter Selection of Support Vector Regression Based on Particle Swarm Optimization, International Conference on Granular Computing, 5, 251 – 256.
- 6- **Hyndman, R., J., Koehler, A. B., (2006).** Another Look at Measures of Forecast Accuracy, International Journal of Forecasting, 22, 679–688.
- 7- **Liu, Y., Keller, Y., Song, PH., Bond, J., Jiang, G., (2017).**Prediction of concrete corrosion in sewers with hybrid Gaussian processes regression model. RSC Advances, 7, 30894–30903.
- 8- **Matloff, N., (2011).**The Art of R Programming, Malloy Incorporated, United States of America, 373.
- 9- **Mullen, K., (2014).** Continuous Global Optimization in R, Journal of Statistical Software, 60, Issue 6.
- 10- **Rasmussen, C. E., Williams C. K. I., (2006).**Gaussian Processes for Machine Learning. MIT, Press.
- 11- **Shi, J. Q., Choi, T., (2011).**Gaussian Process Regression Analysis for Functional Data. Chapman & Hall/CRC, London, rst edition.