



(١١٧) - (١٤١)

العدد الأربعون

دراسة مقارنة بين تصميم المكعب الزائدي اللاتيني وتصميم المكعب الزائدي المقطعي الأقصى في
انحدار العملية الكاوسية
راكان احمد علي عزيز

rakan.24esp18@student.uomosul.edu.iq

أ.م.د. يونس حازم الطويل

younus.altaweel@uomosul.edu.iq

قسم الرياضيات، كلية التربية للعلوم الصرفة، جامعة الموصل، الموصل، العراق

المستخلص:

أنموذج انحدار العملية الكاوسية يستخدم كتمثيل إحصائي لعدم اليقين بشأن مخرجات الأنموذج الحاسوبي. تعتمد منهجية بناء الأنموذج على استخدام دالة وسط ودالة تغاير تحقق خاصية المصفوفة الموجبة شبه المحددة. لبناء أنموذج انحدار العملية الكاوسية، نحتاج الى توليد نوعين من النقاط، نقاط التدريب ونقاط التحقق من صلاحية الأنموذج. يوجد العديد من التصاميم لتوليد هذه النقاط. تصميم المكعب الزائدي يعد من اشهر التصاميم التي تستخدم لبناء أنموذج انحدار العملية الكاوسية. ومع ذلك، فإن تصميم المكعب الزائدي في بعض الأحيان لا يغطي مجال الإدخال بشكل جيد. في هذا البحث سوف نقترح تصميم المكعب الزائدي اللاتيني المقطع الأقصى لبناء أنموذج انحدار العملية وعمل دراسة مقارنة مع تصميم المكعب الزائدي اللاتيني. تتم هذه المقارنة من خلال توليد نقاط باستخدام كلا التصميمين ومن ثم بناء أنموذج انحدار العملية الكاوسية ومقارنة التنبؤات التي نحصل عليها من أنموذج انحدار العملية الكاوسية. تتم هذه المقارنة باستخدام بعض المقاييس المقترحة التي تعتمد على الفرق بين تنبؤات نماذج انحدار العملية الكاوسية والقيم الحقيقية للأنموذج الحاسوبي. تم تطبيق هذه المنهجية على نماذج حاسوبية لتقييم أداء التصميمين، أظهرت النتائج قدرة أنموذج انحدار العملية الكاوسية على تقديم تنبؤات دقيقة ومستقرة، خاصة عند استخدام تصميم المكعب الزائدي اللاتيني المقطع الأقصى الذي أثبت كفاءة عالية. كما أكدت مقاييس التحقق أن



المكعب الزائدي اللاتيني المقطع الأقصى يقلل الخطأ ومستويات عدم اليقين مقارنة بالتصميم المكعب الزائدي اللاتيني وقد نُفذت العمليات الحسابية والبيانية كافة باستخدام برنامج كتبت بلغة البرمجة R. الكلمات المفتاحية: أنموذج انحدار العملية الكاوسية، أخطاء كوليسكي المحورية، دالة محاكاة المكبس، الانحرافات المعيارية المشروطة، تصاميم ملء الفراغ.

Comparative Study Between Latin Hypercube and Maximin Sliced Latin Hypercube Designs for Gaussian Process Regression

Rakan Ahmed Ali Aziz

rakan.24esp18@student.uomosul.edu.iq

Younus Hazim Al-Taweel

younus.altaweel@uomosul.edu.iq

Department of Mathematics, College of Education of Pure Science,
University of Mosul, Mosul, Iraq

Abstract

The Gaussian Process Regression Model (GPRM) is used as a statistical representation of the uncertainty in the computer model outputs. To construct the GPRM, we need to generate two types of design points: training points and validation points. Several designs exist for generating these points. The Latin hypercube design is the most commonly used designs for constructing GPRMs. However, the Latin hypercube design sometimes does not adequately cover the input space. In this research, we propose a Maximin sliced Latin hypercube design for constructing a GPRM and conduct a comparative study with the Latin hypercube design. This comparison is achieved by generating points using both designs, then constructing GPRMs and comparing the resulting predictions using several proposed measures that depend on the difference between the GPRM predictions and the computer model outputs. The GPRMs were applied to computer models to evaluate the performance of the two designs. The results demonstrated the ability of the GPRM to provide accurate and stable predictions, particularly when using the Maximin sliced Latin hypercube design. Validation measures also confirmed that the Maximin sliced Latin hypercube design reduces error and uncertainty levels compared to the Latin hypercube design. All



computational and graphical operations were performed using a program written in the R programming language.

Keywords: Gaussian Process Regression, Pivoted Cholesky Errors, Piston Simulation Function, Conditional Standard Deviations, Space-Filling Designs.

1. المقدمة

تعرف النماذج الحاسوبية (Computer Models) بأنها تمثيلات رياضية لنظم أو علاقات معقدة يتم تنفيذها برمجياً لمحاكاة ظواهر العالم الحقيقي. وقد شهدت هذه النماذج اعتماداً واسعاً في استكشاف الأنظمة الفيزيائية عبر مختلف المجالات العلمية والتقنية (Sacks et al., 1989)؛ ويُعزى ذلك غالباً إلى التحديات التي تفرضها التجارب الفيزيائية التقليدية، كونها باهظة التكلفة أو تستغرق زمناً طويلاً للتنفيذ. وفي سياق التجارب الحاسوبية، يتم توليد البيانات من خلال تنفيذ الأنموذج عند خيارات مدروسة ومتعددة لمتغيرات الإدخال.

تتصف النماذج الحاسوبية عادةً بكونها نماذج حتمية (Deterministic) إذ يؤدي تكرار تنفيذ الأنموذج باستخدام ذات القيم المدخلة إلى الحصول على المخرجات نفسها بشكل مطرد. ورغم طبيعتها الحتمية، تظل قيم المخرجات غير معلومة للمصمم قبل التنفيذ الفعلي عند نقاط إدخال محددة. ومن منظور بييزي (Bayesian Perspective)، يمكن التعبير عن حالة عدم اليقين (Uncertainty) المتعلقة بمخرجات الأنموذج من خلال توظيف العمليات العشوائية، مما ينتج عنه تمثيل إحصائي للأنموذج الحاسوبي يُعرف بأنموذج انحدار العملية الكاوسية (Gaussian Process Regression). وتعد العملية الكاوسية الأداة المركزية في بناء هذا الأنموذج، حيث يتم استخلاص معالمها عبر تنفيذ الأنموذج الحاسوبي على مجموعة بيانات مرجعية تُعرف ببيانات التدريب (van Oijen, 2024).

وبمجرد إتمام بناء أنموذج انحدار العملية الكاوسية، يصبح من الممكن إجراء تحليلات إحصائية متنوعة دون الحاجة إلى تنفيذ الأنموذج الحاسوبي الأصلي المكلف. ويعد التنبؤ بالمخرجات عند نقاط إدخال جديدة لم يسبق تنفيذها (Dolski, 2024). ونظراً لأن طرق مونت كارلو (Monte Carlo) التقليدية قد تكون مجهدة حسابياً في ظل عدم اليقين في المدخلات، فقد تم استخدام انموذج انحدار العملية الكاوسية في العديد من مجالات العلوم والتكنولوجيا. على سبيل المثال، في علم



البيئة، تم استخدام نموذج انحدار العملية الكاوسية من قبل كل من (Rajabi and Ketabchi, 2017)، (Domingo *et al.*, 2020)، (Donnelly *et al.*, 2022)، (Hill *et al.*, 2025). في علم الطب، تم استخدام نموذج انحدار العملية الكاوسية من قبل كل من (Oyebamiji *et al.*, 2017)، (Longobardi *et al.*, 2020)، (Sawe *et al.*, 2024). في علوم الهندسة، استخدم كل من (Pepper *et al.*, 2023) و (Servera *et al.*, 2023) أنموذج انحدار العملية الكاوسية كبديل عن النماذج الحاسوبية. اما في علم المناخ، تم استخدام أنموذج انحدار العملية الكاوسية من قبل كل من (Chaure, Behera, (Tunnell, Bowman and Carrier, 2023)، (Mansfield *et al.*, 2025)، and Bhattacharya, 2023) أنموذج انحدار العملية الكاوسية من قبل كل من (Shin and Am Seo, (Dolski, 2024)، (Paquet, 2024) و (2024). ختاماً، رغم فعالية نماذج انحدار العملية الكاوسية كتقريبات عشوائية للنماذج الحاسوبية المكلفة، إلا أن بناءها يعتمد على افتراضات وتقريبات محددة. ولضمان دقة الاستنتاجات العلمية المستمدة منها، فإنه من الضروري أن يمثل الأنموذج الإحصائي الأنموذج الحاسوبي بدقة، مما يستوجب إخضاع أنموذج انحدار العملية الكاوسية لاختبارات صارمة للتحقق من الصلاحية (Validation). تتوزع هيكلية هذا البحث على عدة محاور أساسية؛ حيث يركز القسم الثاني على الجانب النظري لبناء أنموذج انحدار العملية الكاوسية وكيفية تمثيله رياضياً باستخدام دالتي الوسط والتغاير من منظور بيز. كما يتناول القسم الفرعي (2.2) أساليب اختيار نقاط التدريب من خلال مقارنة تصميم المكعب الزائدي اللاتيني LHD بتصميم المكعب الزائدي اللاتيني المقطع الأقصى MSLHD لضمان التوزيع الأمثل للنقاط. ولتقييم دقة الأنموذج، يستعرض القسم الثالث مقاييس متقدمة للتحقق من الصلاحية تعتمد على تجزئة كوليسكي المحورية والانحرافات المعيارية المشروطة CSD. وفي القسم الرابع، تم إجراء تطبيق عملي باستخدام دالة (Dette and Pepelyshev Function) لاختبار كفاءة هذه النماذج تحت أحجام عينات مختلفة. في القسم الخامس، تم تقديم تطبيق عملي آخر باستخدام دالة محاكاة المكبس (Piston simulation function). وأخيراً، في القسم السادس نقدم الاستنتاجات التي تم التوصل إليها.



2. أنموذج انحدار العملية الكاوسية

يعرف أنموذج انحدار العملية الكاوسية بأنه عملية عشوائية تستخدم لتمثيل الانموذج الحاسوبي، إذ يتم التعامل مع الأنموذج الحاسوبي كدالة رياضية غير معلومة. وبالرغم من أن البنية الخوارزمية للأنموذج الحاسوبي قد تكون معلومة من الناحية النظرية، إلا أن طبيعته المعقدة تفرض معالجة مخرجاته كدالة غير محددة لمدخلاته. من الناحية التأصيلية، وبرؤية تستند إلى المنهج البيزي، كان لكل من (Marc C. Kennedy and O'Hagan, 2001) و (O'Hagan, 1978) السبق في توظيف العمليات الكاوسية لتوصيف سلوك الدوال الرياضية غير المعلومة. ومع حلول ثمانينيات القرن الماضي، تبلورت الفكرة الأساسية لبناء نماذج انحدار إحصائي تعتمد على العمليات الكاوسية؛ حيث قُدِّمَ الإطار العام لهذه النماذج ضمن السياق الإحصائي التقليدي (غير البيزي) بواسطة (Sacks et al., 1989)، في حين قام (Currin et al., 1991) بتطويرها وصياغتها ضمن إطار عمل بيزي متكامل.

1.2. بناء أنموذج انحدار العملية الكاوسية

نستعرض هنا الركائز الأساسية لبناء أنموذج انحدار العملية الكاوسية من منظور إحصاء بيز (van Oijen, 2024). نفترض وجود أنموذج حاسوبي مُمَثَّلٌ بالدالة $\eta(\bullet)$ ، حيث تعتمد هذه الدالة على مجموعة مدخلات يُرمز لها بـ $\mathcal{X} \subset \mathcal{X}_1 \times \dots \times \mathcal{X}_p = \mathcal{X}$ ، حيث $p = 1, 2, \dots$ وتنتج عنها مخرجات نرسم لها بـ \mathcal{R}^p ، $\mathbf{y} \in \mathcal{R}^p$. ولغرض نمذجة حالة عدم اليقين المحيطة بمخرجات هذا الأنموذج، يتم أولاً، يتم تمثيل المعلومات السابقة حول $\eta(\bullet)$ بواسطة عملية كاوسية ذات وسط $m_0(\bullet)$ ودالة تغاير $V_0(\bullet, \bullet)$. وباستخدام صيغة هرمية:

$$(1) \quad \eta(\bullet) | \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi} \sim GP(m_0(\bullet), V_0(\bullet, \bullet)),$$

إذ إن دالة الوسط $m_0(\bullet)$ تعطى بالشكل التالي:

$$(2) \quad m_0(\mathbf{x}) = h(\mathbf{x})^T \boldsymbol{\beta},$$

حيث $h(\bullet): \mathcal{X} \subset \mathcal{R}^p \mapsto \mathcal{R}^q$ هي دالة معلومة للمدخلات، حيث يمكن أن يكون بعد q مختلف عن بعد فضاء المدخلات p ، و $\boldsymbol{\beta}$ هو متجه معاملات غير معلوم. يجب اختيار الدالة $h(\bullet)$ لتشمل أي اعتقاد خبير حول شكل $\eta(\bullet)$. تُعطى دالة التغاير $V_0(\bullet, \bullet)$ بالمعادلة:

$$(3) \quad V_0(\mathbf{x}, \mathbf{x}') = \sigma^2 C(\mathbf{x}, \mathbf{x}'; \boldsymbol{\psi}),$$



حيث σ^2 يمثل معلمة قياس التباين العام غير معلومة، و $C(\bullet, \bullet; \psi)$ هي دالة ارتباط معلومة مع متجه غير معلوم من معاملات الارتباط. يجب أن تضمن دالة الارتباط المختارة، $C(\bullet, \bullet; \psi)$ ، أن تكون مصفوفة التباين لأي مجموعة من المدخلات موجبة شبه محددة.

في هذا العمل، نستخدم دالة الارتباط الكاوسية (Gaussian Correlation Function) $C(\mathbf{x}, \mathbf{x}'; \psi) = \exp\{-\sum_{k=1}^p \left(\frac{x_k - x'_k}{\psi_k}\right)^2\}$ ، إذ إن $(\psi_1, \psi_2, \dots, \psi_p)$ تسمى معاملات الارتباط. نفرض أن $\mathbf{y} = [y_1 = \eta(\mathbf{x}_1), \dots, y_n = \eta(\mathbf{x}_n)]$ مخرجات للنموذج الحاسوبي عند نقاط التصميم $\mathbf{x}_1, \dots, \mathbf{x}_n$ في فضاء الإدخال \mathcal{X} ، وهذه البيانات تمثل مجموعة بيانات التدريب. وفقاً للمعادلة (1)، يكون توزيع المخرجات توزيع طبيعي متعدد المتغيرات:

$$\mathbf{y} | \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi} \sim N_n(H\boldsymbol{\beta}, \sigma^2 A), \quad (4)$$

حيث

$$H = [h(\mathbf{x}_1), \dots, h(\mathbf{x}_n)]^T \quad (5)$$

$$A_{i,j} = C(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\psi}). \quad (6)$$

باستخدام للتوزيع الطبيعي الشرطي متعدد المتغيرات نحصل على:

$$\eta(\bullet) | \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}, \mathbf{y} \sim GP(m_0^*(\bullet), V_0^*(\bullet, \bullet)), \quad (7)$$

حيث

$$m_0^*(\mathbf{x}) = h(\mathbf{x})^T \boldsymbol{\beta} + t(\mathbf{x})^T A^{-1}(\mathbf{y} - H\boldsymbol{\beta})$$

$$V_0^*(\mathbf{x}, \mathbf{x}') = \sigma^2 [C(\mathbf{x}, \mathbf{x}'; \boldsymbol{\psi}) - t(\mathbf{x})^T A^{-1} t(\mathbf{x}')],$$

$$t(\mathbf{x}) = (C(\mathbf{x}, \mathbf{x}_1; \boldsymbol{\psi}), \dots, C(\mathbf{x}, \mathbf{x}_n; \boldsymbol{\psi}))^T$$

باستخدام توزيع اولي قليل المعلومات (Noninformative Prior) $\downarrow (\boldsymbol{\beta}, \sigma^2) \propto p(\boldsymbol{\beta}, \sigma^2)$ ، σ^{-2} وضربها مع معادلة (4) باستخدام نظرية بيز، فإن التوزيع اللاحق $\downarrow (\boldsymbol{\beta}, \sigma^2)$ هو التوزيع الطبيعي معكوس كما إذ إن

$$\boldsymbol{\beta} | \mathbf{y}, \sigma^2, \boldsymbol{\psi} \sim N(\hat{\boldsymbol{\beta}}, \sigma^2 (H^T A^{-1} H)^{-1}), \quad (8)$$

حيث

$$\hat{\boldsymbol{\beta}} = (H^T A^{-1} H)^{-1} H^T A^{-1} \mathbf{y}$$

$$\sigma^2 | \mathbf{y}, \boldsymbol{\psi} \sim \text{InvGam}\left(\frac{n-q}{2}, \frac{(n-q-2)\hat{\sigma}^2}{2}\right), \quad (9)$$

$$\hat{\sigma}^2 = \frac{\mathbf{y}^T (A^{-1} - A^{-1} H (H^T A^{-1} H)^{-1} H^T A^{-1}) \mathbf{y}}{n-q-2}$$

بضرب المعادلتين (7) و (8) والتكامل بالنسبة ل $\boldsymbol{\beta}$ يمكننا الحصول على



$$\eta(\bullet)|y, \sigma^2, \psi \sim GP(m_1^*(\bullet), V_1^*(\bullet, \bullet)), \quad (10)$$

حيث

$$m_1(\mathbf{x}) = h(\mathbf{x})^T \hat{\beta} + t(\mathbf{x})^T A^{-1}(\mathbf{y} - H\hat{\beta}), \quad (11)$$

$$V_1^*(\mathbf{x}, \hat{\mathbf{x}}) = \sigma^2 [C(\mathbf{x}, \hat{\mathbf{x}}; \psi) - t(\mathbf{x})^T A^{-1} t(\hat{\mathbf{x}}) + (h(\mathbf{x}) - t(\mathbf{x})^T A^{-1} H)(H^T A^{-1} H)^{-1} \times (h(\hat{\mathbf{x}}) - t(\hat{\mathbf{x}})^T A^{-1} H)^T]. \quad (12)$$

ان نموذج انحدار العملية الكاوسية يتم الحصول عليه من تكامل حاصل ضرب المعادلتين (9) و

(10) بالنسبة لـ σ^2 يكون بالشكل الآتي:

$$\eta(\bullet)|y, \psi \sim Student Process(n - q, m_1(\bullet), V_1(\bullet, \bullet)), \quad (13)$$

حيث

$$V_1(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\hat{\sigma}^2}{\sigma^2} V_1^*(\mathbf{x}, \hat{\mathbf{x}}). \quad (14)$$

تمثل $V_1(\mathbf{x}, \hat{\mathbf{x}})$ مصفوفة التغاير التنبؤي، الان بأخذ $p(\psi)$ كتوزيع اولي نحصل على

$$p(\psi|y) \propto p(\psi) \iint p(y|\beta, \sigma^2, \psi) p(\beta, \sigma^2) d\beta d\sigma^2 \propto p(\psi) |A|^{-1/2} |H^T A^{-1} H|^{-1/2} (\hat{\sigma}^2)^{-(n-q)/2}, \quad (15)$$

إذ إن A و σ^2 هما دوال لـ ψ . إن تحليل بيزر بالكامل يكون بضرب دالة الكثافة في المعادلة

(13) و(15) وتكامل الناتج بالنسبة لـ ψ ، يتم تقدير متجه معلمة الارتباط ψ ، ثم استخدام التقدير

كما لو كان القيمة الحقيقية لـ ψ . أنموذج انحدار العملية الكاوسية الجديد هو نفسه في (13) مع

استخدام القيم المقدرة لـ A و $\hat{\beta}$ و $\hat{\sigma}^2$ المحسوبة باستخدام القيمة المقدرة لـ ψ (Paquet, 2024).

2.2. تصاميم بناء أنموذج انحدار العملية الكاوسية

للعلوم التربوية والنفسية وطرائق التدريس للعلوم الأساسية

إن اختيار نقاط التصميم لأنموذج انحدار العملية الكاوسية له دور مهم في الحصول على تنبؤات

دقيقة لأنموذج الحاسوبي. يجب أن تغطي نقاط التصميم كامل فضاء المدخلات وأن تنتشر بشكل

متساوٍ عبر هذا الفضاء باستخدام أقل عدد ممكن من النقاط. يُطلق على التصميم الذي يحقق هذه

الخصائص اسم تصميم ملء الفراغ (Space-filling design)، ويُطلق على نقاط التصميم

المستخدمة لبناء أنموذج انحدار العملية الكاوسية اسم نقاط التدريب، حيث يتم تنفيذ الأنموذج

الحاسوبي عند هذه النقاط تحديداً. نهدف في هذا البحث الى إجراء مقارنة تحليلية بين نوعين من فئة

تصاميم المكعب الزائدي اللاتيني لتقييم كفاءتهما في تمثيل الفضاء المدخلات، وهما:



Latin Hypercube Design (LHD)

1.2.2. تصميم المكعب الزائدي اللاتيني

تم اقتراح تصميم الـ LHD من قبل (Mckay, Beckman and Conover, 2000) ، وهو يقوم بتقسيم فترة المتغير إلى n من الفترات الفرعية ذات احتمالية متساوية، ويُنظر إليه كنوع من العينات الطباقية. يتميز تصميم الـ LHD بسهولة التنفيذ، كما أن النقاط فيه تتوزع بشكل منتظم عبر منطقة الإدخال لكل متغير. وبناءً على ذلك، يُستخدم تصميم LHD بشكل متكرر في الدراسات المتعلقة بالتجارب الحاسوبية. لنفترض أننا بحاجة إلى توليد n من النقاط x_1, x_2, \dots, x_n إذ إن $x_j = (x_{j1}, x_{j2}, \dots, x_{jp})$. افترض أن Π هي مصفوفة ذات أبعاد $n \times p$ أعمدتها عبارة عن تبديلات عشوائية مستقلة للمجموعة $\{1, 2, \dots, n\}$. افترض أيضاً أن U_{jk} هي عينة عشوائية مستقلة ومتماثلة وموزعة بانتظام على الفترة $(0, 1)$ ومستقلة عن Π . وبناءً عليه، يتم توليد تصميم LHD من خلال المعادلة التالية:

$$x_{jk} = \frac{\Pi_{jk} - U_{jk}}{n} \quad (16)$$

2.2.2. تصميم المكعب الزائدي اللاتيني المقطع الأقصى

Maximin Sliced Latin Hypercube Design (MSLHD)

يعمل تصميم الـ MSLHD عن طريق تقسيم تصميم المربع اللاتيني إلى أجزاء (مقاطع أو شرائح)، بهدف تحقيق التوزيع الأمثل للعناصر داخل كل شريحة مع الحفاظ على الخصائص والسمات الإحصائية للعينات. في البداية، يتم تشكيل تصميم LHD كلاسيكي كصفوف في مصفوفة، حيث يحتوي كل صف على مستوى واحد لكل عامل تصميم. يتبع ذلك تقسيم المصفوفة إلى عدد ثابت ومحدد مسبقاً من المقاطع. وباستخدام التوزيع العشوائي، يتم تحسين توزيع النقاط داخل كل مقطع لضمان التوزيع المتوازن والمتساوي عبر جميع المتغيرات. وأخيراً، يتم استخدام احد خوارزميات التحسين أو الطرق العددية للوصول إلى الحل الأمثل عن طريق تقليل التداخل بين الشرائح وضمان التوزيع الأمثل للعناصر داخل كل مقطع. باستخدام t من المقاطع و m من النقاط في كل عامل من العوامل الـ p ، قدم (Ba, Myers and Breneman, 2015) طريقة مباشرة ومفهومة لإنشاء تصميم SLHD بعدد تنفيذ n . تتضمن طريقة الإنشاء هذه خطوتين فقط:



1. لكل مقطع من المقاطع الـ t ، يتم إنشاء t من تصميمات LHD الصغيرة بشكل مستقل X_1, \dots, X_q بحيث يحتوي كل منها على m من النقاط في p من العوامل. لإنشاء مصفوفة $X = \bigcup_i^t X_i$ بأبعاد $n \times p$ ، يتم تجميع مستويات العوامل، التي يرمز لها بـ $1, \dots, m$.

2. استبدل كل مدخل من المدخلات الـ t في l ($l = 1, \dots, m$) في كل عمود من أعمدة X بشكل مستقل بتبديلة عشوائية λ_l للعناصر $[(1-t)t + 1, \dots, lt]$.

يمكن إنشاء إجمالي $((m!)^t (t!)^m)^p$ من تصميمات SLHD المحتملة لأي قيم معينة معطاة لـ m و t و p . ويتم تحديد أفضل SLHD على أنه الذي يحقق أقصى معيار أمثل للتصميم المحدد.

اقترح (Ba, Myers and Breneman, 2015) دالة هدف واحدة لمعيار المسافة القصوى لتصميم MSLHD. وهي كالتالي:

$$\phi_{Mm}(X) = \frac{1}{2} \left(\phi_r(X) + \frac{1}{t} \sum_{i=1}^t \phi_r(X^i) \right) \quad (17)$$

حيث

$$\phi_r(X) = \left(\frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \frac{1}{d(x_i, x_j)^r} \right)^{\frac{1}{r}} \quad (18)$$

$$d(x_i, x_j) = \left\{ \sum_{k=1}^p |x_{ik} - x_{jk}|^q \right\}^{\frac{1}{q}} \quad (19)$$

تقوم المعادلة (19) بتقليل $\phi_r(X_i)$ لكل مقطع بالإضافة إلى $\phi_r(X)$ لكل نقطة تصميم. ويُعرف الـ MSLHD بأنه التصميم الذي يقلل $\phi_{Mm}(X)$ في المعادلة (19). قدم (Ba, Myers and Breneman, 2015) تقنية عامة لإيجاد أفضل تصميمات SLHD باستخدام خوارزمية التبادل (Exchange Algorithm). افترض أن $p_0 = 0.5$ وأنا نرغب في تقليل مقياس محدد مسبقاً $\phi(X)$ لتصميم الـ SLHD عشوائياً ثم تنتقل لتحليل سلسلة من التصميمات الجديدة التي يتم إنتاجها عن طريق إحداث تغيير في التصميم السابق:

1. ضع $X \leq$ SLHD الحالي. قم بتوليد متغير عشوائي Z يتبع التوزيع المنتظم $(0,1)$.



2. إذا كان $z \leq p_0$ ، اختر مقطعاً في X بشكل عشوائي، ثم قم بتبديل عنصرين تم اختيارهما عشوائياً داخل عمود اختيار عشوائياً. أشر إلى الـ SLHD الجديد بـ X_{try} وانتقل إلى الخطوة 4.

3. في حالة كان $z > p_0$ ، اختر عموداً من X بشكل عشوائي وقم بتبديل عنصرين عشوائياً من مجموعة مختارة (Λ_l ($l \in \{1, \dots\}$)). انتقل إلى الخطوة 4 بعد تسمية الـ SLHD الناتج بـ X_{try} .

4. استبدل القيمة الحالية X بـ X_{try} إذا كانت X_{try} أصغر من X ، وإذا لم يكن كذلك، استبدل X بـ X_{try} باحتمالية $\Lambda = \exp\left\{\frac{X_{try}-X}{t}\right\}$ ، حيث t هو معلمة محددة مسبقاً.

5. حتى تتحقق شروط معينة للتقارب، كرر الخطوات من (1) إلى (4).

6. يتم تحديد التصميم المثالي على أنه تصميم المكعب اللاتيني المقطع الأقصى MSLHD الذي تجد الخوارزمية أنه يمتلك أقل قيمة لـ $\phi(X)$.

عند حساب مصفوفة التباين، إذا كانت المسافة الدنيا بين نقطتين صغيرة جداً، فإن دالة الارتباط تعطي قيمة قريبة من الواحد، وهذا يجعل الصفوف أو الأعمدة في مصفوفة التباين متشابهة جداً، مما يجعل المصفوفة غير محددة إيجابياً بشكل جيد. إن تعظيم المسافة الدنيا بين نقاط التدريب يقلل من وجود نقاط متقاربة جداً ويضمن تغطية أفضل للمجال، والذي يؤدي إلى تحسين أداء نموذج انحدار العملية الكاوسية عبر تقليل الارتباط بين البيانات، وهذا الإجراء يحد من عدم استقرار مصفوفة التباين ويقود بالتالي إلى تقديرات أكثر دقة.

إن معيار تعظيم المسافة الدنيا يعمل على توزيع نقاط العينة بشكل متساوٍ قدر الإمكان، مما يقلل من المساحات (الفجوات) التي لا تحتوي على معلومات دقيقة. وهذا بدوره يساهم في تقليل خطأ التكامل العددي من خلال توزيع نقاط العينة لتعظيم الحد الأدنى للمسافة بينها، مما يضمن تغطية شاملة لنطاق التكامل، وبالتالي تقليل الحد الأقصى للخطأ المحتمل، خاصة في الأبعاد العالية.

3. التحقق من صلاحية نموذج انحدار العملية الكاوسية

يتم إنشاء نموذج انحدار العملية الكاوسية بناءً على افتراضات معينة للتباين والارتباط. ونتيجة لذلك، يجب استخدام المقاييس الإحصائية لتقييم أداء نموذج انحدار العملية الكاوسية والتحقق من مدى واقعية الافتراضات التي اسند إليها البناء الرياضي للنموذج. في هذا القسم، نقدم بعض



المقاييس الإحصائية المعتمدة للتحقق من صلاحية أنموذج انحدار العملية الكاوسية وضمان كفاءته في التمثيل والتنبؤ، والتي سنعتمد لاحقاً للمقارنة بين كفاءة تصميم LHD وتصميم MSLHD.

1.3. مقياس Q^2 و RMSE

إن مقياس جذر متوسط مربعات الخطأ RMSE يعطى بواسطة:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i^* - m_1(\mathbf{x}_i))^2} \quad (20)$$

حيث كلما كانت قيمة $RMSE \rightarrow 0$ كلما كانت التنبؤات أفضل. أما مقياس Q^2 فيعطى بواسطة:

$$Q^2 = 1 - \frac{\sum_{i=1}^m (y_i^* - m_1(\mathbf{x}_i))^2}{\sum_{i=1}^m \left(y_i^* - \frac{1}{m} \sum_{i=1}^m y_i^* \right)^2} \quad (21)$$

إذ إنّه كلما كانت قيمة $Q^2 \rightarrow 1$ كلما كانت التنبؤات أفضل.

2.3. تجزئة مصفوفة التباين

أن التنبؤ في انحدار العملية الكاوسية يعتمد على مصفوفة التباين، لنفترض أنّ \mathbf{G} هي مصفوفة الانحراف المعياري بإذ إنّ $V_1(\mathbf{x}, \hat{\mathbf{x}}) = \mathbf{G}\mathbf{G}^T$. وبناءً عليه، فإن الأخطاء المحولة الناتجة:

$$K_{\mathbf{G}}(\mathbf{y}^*) = \mathbf{G}^{-1}(\mathbf{y}^* - m_1(\mathbf{x})) \quad (22)$$

ستتصف بقيم مستقلة وغير مرتبطة إحصائياً. وبموجب افتراض التوزيع الطبيعي، يتبع كل خطأ توزيع t بدرجات حرية $(n - q)$ على المقياس $K_{\mathbf{G}}(\mathbf{y}^*)$ ، مما يتيح لنا فحص القيم المتطرفة سواء كانت كبيرة أو صغيرة جداً. ولتسهيل التعامل مع المصفوفة اكيده الايجابية (Positive Definite)، يمكن تحليلها إلى حاصل ضرب مصفوفة الجذر التربيعي في منقولها عبر استخدام خوارزمية تجزئة كوليسكي المحورية (Bastos and O'Hagan, 2009).

Pivoted Cholesky

3.3. تجزئة كوليسكي المحوري

Decomposition

اقترح (Bastos and O'Hagan, 2009) استخدام تجزئة كوليسكي المحورية لتوليد أخطاء غير مترابطة، كما هو موضح في المعادلة (20). ويمكن الحصول على هذه التجزئة من خلال إعادة ترتيب نقاط التحقق من الصلاحية بحيث يُقابل الخطأ الأول أكبر تباين تنبؤي، بينما يُقابل الخطأ الثاني أكبر تباين تنبؤي مشروطاً بالخطأ الأول، وتستمر هذه العملية على النحو نفسه لبقية الأخطاء. وعليه، يمكن حساب أخطاء كوليسكي المحورية وفق العلاقة:

$$K_{PC} = \mathbf{G}_{PC}^{-1}(\mathbf{y}^* - m_1(\mathbf{x})) \quad (23)$$

إذ إن $\mathbf{G}_{PC} = \mathbf{P}\mathbf{R}$ و \mathbf{P} تمثلان مصفوفة التبديل. وبعد ذلك، من خلال تمثيل أخطاء كوليسكي المحورية مقابل الدليل، أي الترتيب المحوري، يتوقع أن تتمركز الأخطاء حول 0 مع تباين ثابت.

ويعد انموذج انحدار العملية الكاوسية عالي الثقة أو منخفض الثقة إذا أظهر الرسم عدداً كبيراً من الأخطاء الكبيرة أو عدداً كبيراً من الأخطاء الصغيرة، إذ يمكن أن تشير كلتا الحالتين أيضاً إلى وجود عدم استقرار في الأنموذج. فضلاً عن ذلك، يمكن ربط تجزئة كوليسكي المحورية ببنية الارتباط؛ إذ تشير الأخطاء الكبيرة أو الصغيرة الواقعة في الجهة اليسرى من الرسم إلى عملية غير مستقرة أو إلى تقدير ضعيف للتباين التنبؤي. في المقابل، فإن ظهور أخطاء متطرفة (كبيرة أو صغيرة) في الجهة اليمنى من الرسم قد يدل على بنية ارتباط غير ملائمة أو على تقدير غير دقيق لمعاملات الارتباط.

مجلة العلوم الأساسية
للعلوم التربوية والنفسية وطرائق التدريس للعلوم الأساسية

4.3. الانحرافات المعيارية المشروطة (CSD) Conditional Standard Deviations

نقترح في هذا القسم إجراء تعديل على مقياس كوليسكي المحوري، وذلك عبر رسم أخطاء كوليسكي المحورية K_{PC} مقابل الانحرافات المعيارية المشروطة CSD في الحالة الاعتيادية، يتم رسم الأخطاء بناءً على دليلها التسلسلي، مما يجعلها تظهر منتشرة بشكل متساوٍ على المحور الأفقي X في رسم K_{PC} مقابل الترتيب المحوري؛ وبذلك يعبر المحور X عن مقياس رتبي فقط يتعلق بترتيب النقاط. بدلاً من ذلك، نوصي برسم K_{PC} مقابل الانحرافات المعيارية المشروطة، وهو ما يمنحنا مقياس نسبة على المحور X . يتيح هذا المقياس إجراء مقارنات دقيقة بين النقاط، مما يجعله أكثر



إفادة في تحليل البيانات؛ حيث يمكننا، على سبيل المثال، تمييز مدخلات التدريب القريبة من نقاط التحقق أو البعيدة عنها. لتكن $\mathbf{x}_1^*, \dots, \mathbf{x}_m^*$ هي مدخلات اختبار، حيث يتم ترتيب هذه المدخلات بناءً على البيانات الشرطية لمخرجات اختبار. ويمكن الحصول على الانحرافات المعيارية المشروطة باتباع الخطوات التالية:

• احسب $v_1 = V(\eta(\mathbf{x}_1^*) | \mathbf{y}), \dots, v_m = V(\eta(\mathbf{x}_m^*) | \mathbf{y})$ ، ثم ضع $\mathbf{x}_{(1)}^* = \mathbf{x}_{i_1}^*$ إذ $i_1 = \arg \max_i v_i$

• احسب $v_{i|(1)} = V(\eta(\mathbf{x}_i^*) | \mathbf{y}, \eta(\mathbf{x}_{(1)}^*))$ ، $i = 1, \dots, m$ و $(i \neq i_1)$ ، ثم ضع $\mathbf{x}_{(2)}^* = \mathbf{x}_{i_2}^*$ إذ $i_2 = \arg \max_i v_{i|(1)}$

• احسب $v_{i|(1),(2)} = V(\eta(\mathbf{x}_i^*) | \mathbf{y}, \eta(\mathbf{x}_{(1)}^*), \eta(\mathbf{x}_{(2)}^*))$ ، $(i \neq i_1 \neq i_2)$ ، ثم ضع $\mathbf{x}_{(3)}^* = \mathbf{x}_{i_3}^*$ إذ $i_3 = \arg \max_i v_{i|(1),(2)}$

وبناءً على ما سبق، يتم رسم قيم K_{PC} مقابل قيم CSD الممثلة في المجموعة \mathbf{S} :

$$\mathbf{S} = \left\{ \begin{array}{l} S_1 = \sqrt{v_{i_1}}, S_2 = \sqrt{v_{i_2|(1)}}, S_3 = \sqrt{v_{i_3|(1),(2)}}, \dots, \\ S_m = \sqrt{v_{i_m|(1),(2), \dots, (m-1)}} \end{array} \right\}$$

إذ $v_{i_1} = V(\eta(\mathbf{x}_1^*) | \mathbf{y})$ ، $v_{i_2|(1)} = V(\eta(\mathbf{x}_i^*) | \mathbf{y}, \eta(\mathbf{x}_{(1)}^*))$ ، \dots
 $v_{i_m|(1),(2), \dots, (m-1)} = V(\eta(\mathbf{x}_m^*) | \mathbf{y}, \eta(\mathbf{x}_{(1)}^*), \dots, \eta(\mathbf{x}_{(m-1)}^*))$

ولغرض توحيد المقاييس، يمكن تحويل الانحرافات المعيارية المشروطة إلى قيم معيارية باستخدام الصيغة التالية:

$$CSD = \frac{4 \times \mathbf{S}}{R} \quad (24)$$

إذ إن $Y = (\mathbf{y}, \mathbf{y}^*)$ و $R = \max(Y) - \min(Y)$. يشير هذا المقياس بشكل مباشر إلى عرض فترات الثقة بنسبة 95% بالنسبة للمجال الكلي للبيانات.

4 مثال دالة Dette and Pepelyshev (DP) Dette and Pepelyshev Function

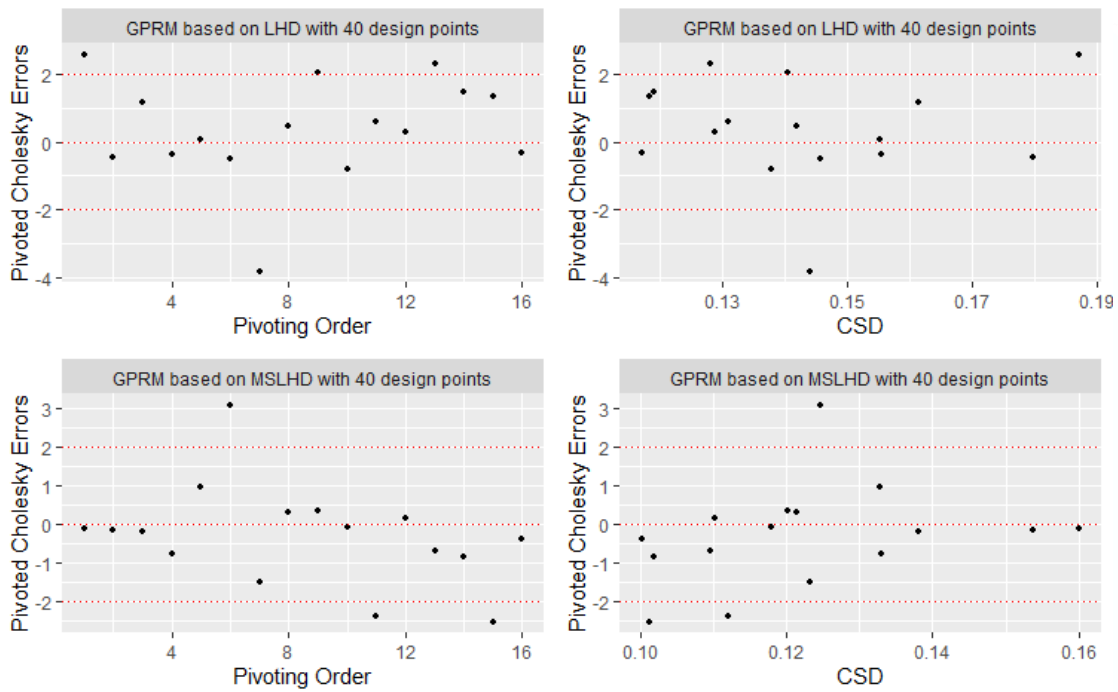


دالة DP هي نموذج رياضي ذو ثمانية أبعاد تم استخدامها من قبل (Dette and Pepelyshev, 2010). مخرجات هذه الدالة تعطى بواسطة:

$$f(\mathbf{x}) = 4(x_1 - 2 + 8x_2 - 8x_2^2)^2 + (3 - 4x_2)^2 + 16\sqrt{x_3 + 1} + (3x_3 - 1)^2 \sum_{i=4}^8 i \ln(1 + \sum_{i=3}^i x_j) \quad (25)$$

إذ إن \mathbf{x} هو متجه متغيرات الإدخال التي تقع قيمها ضمن الفترة $[0, 1]$ ، وإن $f(\mathbf{x})$ متغير الاستجابة.

تم بناء أنموذج انحدار العملية الكاوسية، حيث تم تحديد حجم مدخلات التدريب ليكون بمقدار خمسة أضعاف أبعاد الأنموذج الحاسوبي $5p$ ، أي تم توليد 40 نقطة تدريب باستخدام LHD و MSLHD. تم الحصول على مخرجات دالة DP عند نقاط التدريب هذه. تم بناء أنموذج انحدار العملية الكاوسية باستخدام دالة وسط خطية وفقاً للمعادلة (2) ودالة الارتباط الكاوسية. كما تم الحصول على تنبؤات لدالة DP باستخدام أنموذج انحدار العملية الكاوسية من خلال المعادلة (11) ومعادلة (14). تم توليد نقاط تحقق عددها $2p = 16$ وتم حساب مخرجات دالة DP عند هذه النقاط. للتحقق من صلاحية أنموذج انحدار العملية الكاوسية، تم حساب أخطاء كوليسكي المحورية، المعادلة (23)، والانحرافات المعيارية المشروطة، المعادلة (24)، الشكل (1) يوضح أخطاء كوليسكي المحورية مقابل ترتيبها المحوري (Pivoting Order) وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة.



الشكل (1): أخطاء كوليسكي المحورية مقابل ترتيبها المحوري وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة لدالة DP بالاعتماد على نقاط تدريب $5p$.

يظهر الرسم العلوي الخاص بتصميم LHD أن معظم أخطاء كوليسكي المحورية تقع ضمن الحدود $[-2, 2]$ ، إلا أن هناك نقاطاً واضحة تجاوزت هذه الحدود، مما يشير إلى وجود تعارض في بعض مناطق فضاء الإدخال. كما نلاحظ أن قيم الانحرافات المعيارية المشروطة (CSD) تتركز تقريباً في الفترة $(0.11, 0.19)$ ، مما يعكس عدم يقين أكبر في التقدير. في المقابل، يظهر الرسم السفلي الخاص بتصميم MSLHD تحسناً ملحوظاً، حيث وقعت غالبية أخطاء كوليسكي المحورية داخل



الحدود المطلوبة، مع وجود نقاطاً خارج الحدود وهي قريبة من الحد الاعلى للفترة $[-2,2]$. علاوة على ذلك، نلاحظ ان قيم الانحرافات المعيارية المشروطة (CSD) تقع في الفترة $(0.10,1.16)$ وأن قيمها أقل من تلك التي حصلنا عليها باستخدام تصميم LHD.

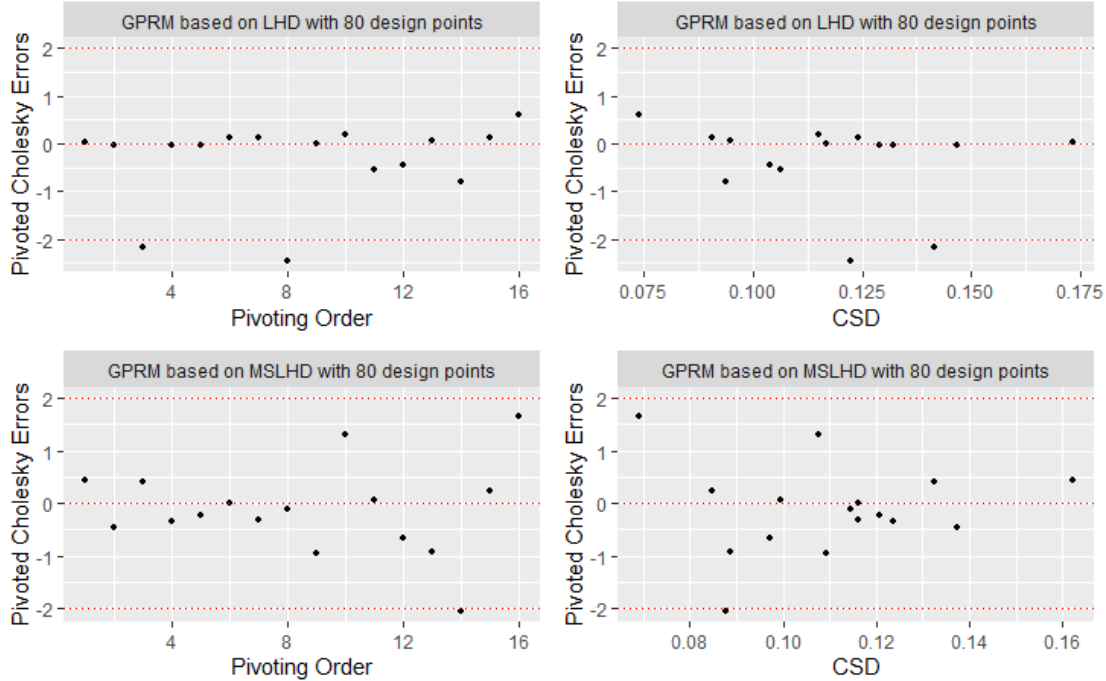
تم أيضاً حساب قيم RMSE و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج $[-2,2]$ لكافة نماذج انحدار العملية الكاوسية كما هو موضح في الجدول (1).

جدول (1): قيم RMSE و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لأنموذج انحدار العملية الكاوسية لدالة DP بالاعتماد على نقاط تدريب $5p$.

نوع التصميم	RMSE	Q^2	متوسط CSD	نسبة الأخطاء خارج $[-2, 2]$
LHD	4.925	0.705	0.220	25%
MSLHD	2.813	0.881	0.188	18%

نلاحظ أن قيمة RMSE لأنموذج انحدار العملية الكاوسية المعتمد على MSLHD أصغر من تلك التي تم الحصول عليها عند استخدام LHD. بالإضافة الى ذلك، فإن قيمة Q^2 ، ومتوسط CSD، ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لأنموذج انحدار العملية الكاوسية باستخدام MSLHD كانت أفضل مقارنةً بتلك المنحصّل عليها عند استخدام LHD.

تمت زيادة عدد نقاط التدريب لتصبح $10p$ ، حيث تم توليد 80 نقطة تدريب باستخدام كل من تصميم LHD وتصميم MSLHD. بعد ذلك تم الحصول على مخرجات دالة DP عند هذه النقاط. يوضّح الشكل (2) أخطاء كولييسكي المحورية مقابل ترتيبها المحوري (Pivoting Order)، وكذلك أخطاء كولييسكي المحورية مقابل الانحرافات المعيارية المشروطة. وذلك باستخدام 80 نقطة تدريب.



الشكل (2): أخطاء كوليسكي المحورية مقابل ترتيبها المحوري وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة لدالة DP بالاعتماد على نقاط تدريب $10p$.

نلاحظ أن أخطاء كوليسكي المحورية لتصميم LHD قد تحسنت إلا أن بعض النقاط لا تزال خارج الحدود للفترة $[-2, 2]$ أو قريبة منها، مع قيم الانحرافات المعيارية المشروطة (CSD) التي تتراوح بين 0.07 و 0.175. أما بالنسبة لتصميم MSLHD، فتظهر النتائج تظهر كفاءة عالية جداً؛ حيث تقع جميع أخطاء كوليسكي تقريباً داخل الحدود المسموحة وبقيم أصغر، مع تمركز قيم الانحرافات المعيارية المشروطة (CSD) في نطاق أكثر ضيقاً.

تم أيضاً حساب قيم RMSE و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لكافة نماذج انحدار العملية الكاوسية كما هو موضح في الجدول (2).



جدول (2): قيم RMSE و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لأنموذج انحدار العملية الكاوسية لدالة DP بالاعتماد على نقاط تدريب $10p$.

نوع التصميم	RMSE	Q^2	متوسط CSD	نسبة الأخطاء خارج $[-2, 2]$
LHD	1.835	0.955	0.149	12.5%
MSLHD	1.007	0.985	0.159	0.6%

نلاحظ تحسناً واضحاً في قيم RMSE و Q^2 ومتوسط CSD ونسبة الأخطاء خارج المجال $[-2, 2]$. ومع ذلك، فإن هذه المؤشرات لأنموذج انحدار العملية الكاوسية المعتمد على MSLHD لاتزال أفضل مقارنةً بتلك المتحصّل عليها عند استخدام LHD.

Piston simulation function

5. مثال دالة محاكاة المكبس

استخدم (Fuchs and Kenett, 1998) دالة محاكاة رياضية لنمذجة الأداء الديناميكي للمكبس داخل الأسطوانة. تعبر الاستجابة في هذه الدالة عن المدة الزمنية اللازمة لإتمام دورة كاملة للمكبس مقاسة بالثواني، وتمثل رياضياً كما يلي:

$$\eta(x) = 2\pi \sqrt{\frac{M}{k + S^2 \frac{P_0 V_0 T_a}{T_0 V^2}}} \quad (26)$$

إذ $V = \frac{S}{2k} \left(\sqrt{A^2 + 4k \frac{P_0 V_0}{T_0} T_a} - A \right)$ و $A = P_0 S + 19 \cdot 62M - \frac{kV_0}{S}$ ، ان متغيرات الإدخال التي تؤثر على أداء المكبس هي كما يلي:

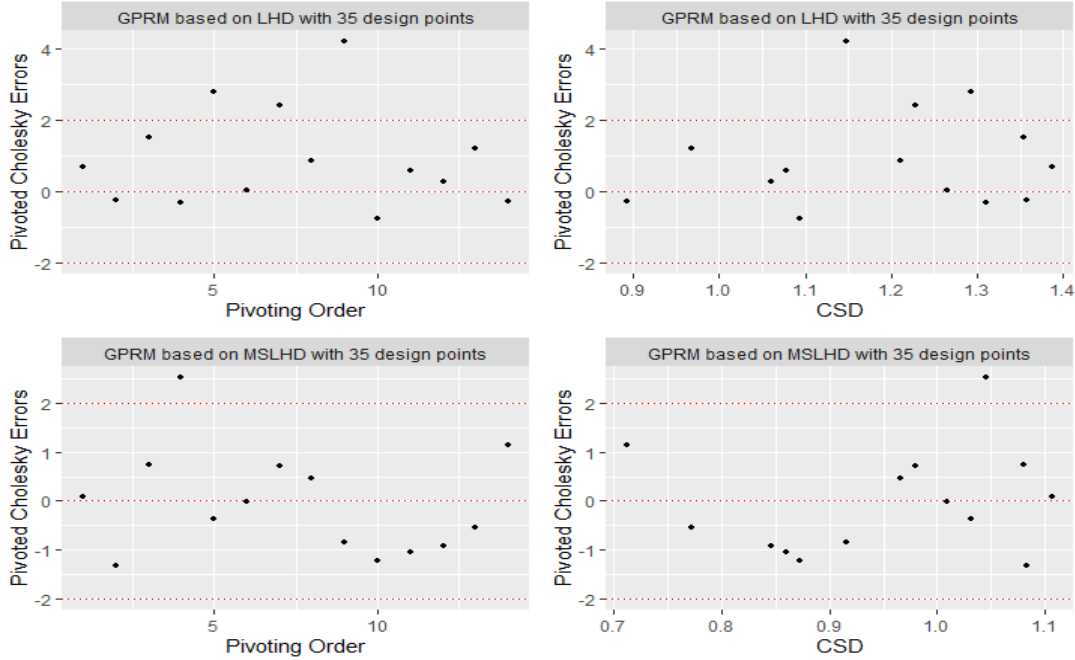
- وزن المكبس $M \in [30, 60]$ ووحدة قياسه كيلو غرام (kg).
- مساحة سطح المكبس $S \in [0 \cdot 005, 0 \cdot 020]$ ووحدة قياسه متر مربع (m^2).
- حجم الغاز الابتدائي $V_0 \in [0 \cdot 002, 0 \cdot 010]$ ووحدة قياسه متر مكعب (m^3).
- معامل الزنبرك $k \in [1000, 5000]$ ووحدة قياسه نيوتن لكل متر (N/m).



- الضغط الجوي $P_0 \in [90000, 110000]$ ووحدة قياسه نيوتن لكل متر مربع (N/m^2) .
- درجة حرارة المحيطة $T_a \in [290, 296]$ ووحدة قياسه الكلفن (K) .
- درجة حرارة تعبئة الغاز $T_0 \in [340, 360]$ ووحدة قياسه الكلفن (K) .

تم توليد $5p = 35$ نقطة تدريب باستخدام تصميم LHD و MSLHD. تم الحصول على مخرجات دالة محاكاة المكبس عند نقاط التدريب. تم بناء نموذج انحدار العملية الكاوسية باستخدام دالة وسط خطية وفقاً للمعادلة (2) ودالة الارتباط الكاوسية. كما تم الحصول على تنبؤات لدالة محاكاة المكبس باستخدام نموذج انحدار العملية الكاوسية معادلة (11) ومعادلة (14).

للتحقق من صلاحية نموذج انحدار العملية الكاوسية، تم توليد نقاط تحقق عددها $2p = 14$ وتم حساب مخرجات دالة محاكاة المكبس عند هذه النقاط. بعد ذلك، تم حساب أخطاء كوليسكي المحورية، المعادلة (23)، والانحرافات المعيارية المشروطة، المعادلة (24). الشكل (3) يوضح أخطاء كوليسكي المحورية مقابل ترتيبها المحوري واخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة.



الشكل (3): أخطاء كوليسكي المحورية مقابل ترتيبها المحوري وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة لدالة محاكاة المكبس بالاعتماد على نقاط تدريب $5p$.

يظهر الرسم العلوي الخاص بتصميم LHD أن معظم أخطاء كوليسكي المحورية تقع ضمن الحدود $[-2, 2]$ ، إلا أن هناك نقاطاً واضحة تجاوزت هذه الحدود، مما يشير إلى وجود تعارض في بعض مناطق فضاء الإدخال. كما نلاحظ أن قيم الانحرافات المعيارية المشروطة (CSD) تتركز تقريباً في الفترة $(0.9, 1.4)$ ، مما يعكس عدم يقين أكبر في التقدير. في المقابل، يظهر الرسم السفلي الخاص بتصميم MSLHD تحسناً ملحوظاً، حيث وقعت غالبية أخطاء كوليسكي المحورية داخل الحدود المطلوبة، مع وجود نقطة واحدة فقط خارج الحدود وهي قريبة من الحد الأعلى للفترة $[-2, 2]$ ، وهو أمر مقبول إحصائياً. نلاحظ أيضاً أن قيم الانحرافات المعيارية المشروطة (CSD) تقع في الفترة $(0.7, 1.1)$ وأن قيمها أقل من تلك التي حصلنا عليها باستخدام تصميم LHD.

تم أيضاً حساب قيم $RMSE$ و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج $[-2, 2]$ لنماذج انحدار العملية الكاوسية كما هو موضح في الجدول (3).

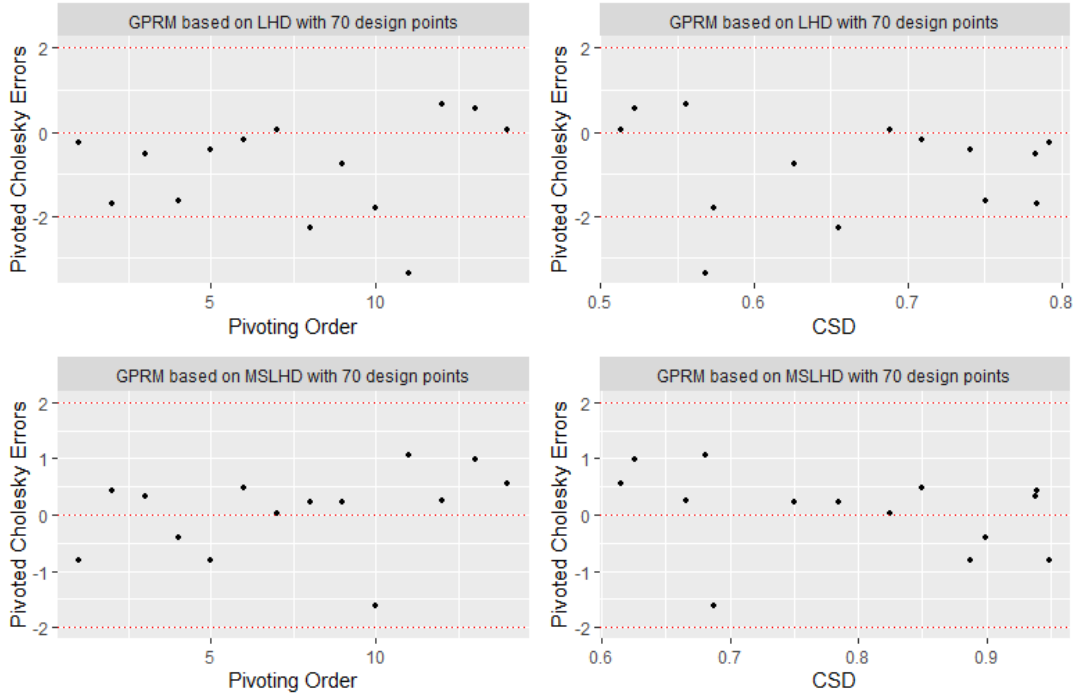


جدول (3): قيم Q^2 و $RMSE$ ومتوسط CSD ونسبة الاخطاء خارج $[-2,2]$ لأنموذج انحدار العملية الكاوسية لدالة DP بالاعتماد على نقاط تدريب $5p$.

نوع التصميم	RMSE	Q^2	متوسط CSD	نسبة الأخطاء خارج $[-2, 2]$
LHD	0.030	0.949	0.115	21%
MSLHD	0.003	0.989	0.096	7%

نلاحظ أن قيمة $RMSE$ لأنموذج انحدار العملية الكاوسية المعتمد على MSLHD أصغر من تلك التي تم الحصول عليها عند استخدام LHD. بالإضافة الى ذلك، فإن قيمة Q^2 ، ومتوسط CSD ، ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لأنموذج انحدار العملية الكاوسية باستخدام MSLHD كانت أفضل بتلك المتحصّل عليها عند استخدام LHD.

تم زيادة عدد نقاط التدريب لتكون $10p$ حيث تم توليد 70 نقطة تدريب بواسطة كل من تصميم LHD وتصميم MSLHD، بعد ذلك تم الحصول على مخرجات دالة محاكاة المكبس. بناءً على هذه النقاط السبعين من بيانات التدريب، تم تقدير معاملات الارتباط بطريقة (MLE) معادلة (15). للتحقق من أنموذج انحدار العملية الكاوسية الجديد، نعيد خطوات التحقق السابقة حيث تم توليد نقاط تحقق بعدد $2p = 14$ ، وتم حساب مخرجات دالة محاكاة المكبس، تم أيضا حساب أخطاء كوليسكي المحورية، المعادلة (23)، والانحرافات المعيارية المشروطة، المعادلة (24). الشكل (4) يوضح أخطاء كوليسكي المحورية مقابل ترتيبها المحوري (Pivoting Order) وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة باستخدام 70 نقطة تدريب.



الشكل (4): أخطاء كوليسكي المحورية مقابل ترتيبها المحوري وأخطاء كوليسكي المحورية مقابل الانحرافات المعيارية المشروطة لدالة محاكاة المكبس بالاعتماد على نقاط تدريب $10p$.

نلاحظ أن أخطاء كوليسكي المحورية لتصميم LHD قد تحسنت إلا أن بعض النقاط لا تزال خارج الحدود للفترة $[-2, 2]$ أو قريبة منها، مع قيم الانحرافات المعيارية المشروطة (CSD) التي تتراوح بين 0.5 و 0.8. أما بالنسبة لتصميم MSLHD، فتظهر النتائج تظهر كفاءة عالية جداً؛ حيث تقع جميع أخطاء كوليسكي تقريباً داخل الحدود وقيم أصغر. يؤكد ذلك أن زيادة حجم العينة يؤدي إلى إنتاج أنموذج انحدار العملية الكاوسية صالح إحصائياً وأكثر موثوقية في التنبؤ.

تم أيضاً حساب قيم $RMSE$ و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج $[-2, 2]$ لكافة نماذج انحدار العملية الكاوسية كما هو موضح في الجدول (4).

جدول (4): قيم $RMSE$ و Q^2 ومتوسط CSD ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$ لأنموذج انحدار العملية الكاوسية لدالة DP بالاعتماد على نقاط تدريب $10p$.



نوع التصميم	RMSE	Q^2	متوسط CSD	نسبة الأخطاء خارج [-2, 2]
LHD	0.044	0.955	0.073	14%
MSLHD	0.003	0.989	0.077	0%

نلاحظ تحسناً في قيم RMSE ومتوسط CSD ونسبة الأخطاء الواقعة خارج المجال $[-2, 2]$. إلا أن هذه القيم لأنموذج انحدار العملية الكاوسية المعتمد على MSLHD لاتزال أفضل مقارنةً بتلك المتحصّل عليها عند استخدام LHD.

6. الاستنتاجات

أثبتت الدراسة أن اختيار تصميم نقاط التدريب يلعب دوراً محورياً في دقة وكفاءة أنموذج انحدار العملية الكاوسية GPRM، حيث حقق تصميم المكعب الزائدي اللاتيني المقطع الأقصى MSLHD تفوقاً جوهرياً على تصميم LHD التقليدي من حيث استقرار التنبؤات وتقاربها مع المخرجات الحقيقية، خاصة عند زيادة حجم عينة التدريب إلى 70 و 80 نقطة. كما أظهرت مقاييس التحقق من الصلاحية، المتمثلة في أخطاء كوليسكي المحورية والانحرافات المعيارية المشروطة (CSD) وبعض المقاييس الأخرى، أن تصميم MSLHD يضمن تغطية أفضل لفضاء المدخلات ويقلل من مستويات عدم اليقين والتعارض في البيانات، مما يجعل أنموذج انحدار العملية الكاوسية تمثيلاً إحصائياً صالحاً وموثوقاً للنماذج الحاسوبية المعقدة.

شكر وتقدير
للعلوم التربوية والنفسية وطرائق التدريس للعلوم الأساسية

الباحثون يتقدمون بالشكر والامتنان الى قسم الرياضيات في كلية التربية للعلوم الصرفة/جامعة الموصل على الجهود المبذولة والتسهيلات المقدمة والتي ساعدت على تحسين واكمال جودة هذا البحث.



References

المصادر

1. Ba, S., Myers, W.R. and Brenneman, W.A. (2015) "Optimal sliced Latin hypercube designs," *Technometrics*, 57(4), pp. 479–487.
2. Bastos, L.S. and O'Hagan, A. (2009) "Diagnostics for Gaussian process emulators," *Technometrics*, 51(4), pp. 425–438.
3. Chaure, A., Behera, A.K. and Bhattacharya, S. (2023) "Gaussian Process Regression for Climate Modeling: Potentials, Limitations, and Advances in Emulation Techniques."
4. Currin, C. et al. (1991) "Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments," *Journal of the American Statistical Association*, 86(416), pp. 953–963.
5. Dette, H. and Pepelyshev, A. (2010) "Generalized Latin hypercube design for computer experiments," *Technometrics*, 52(4), pp. 421–429.
6. Dolski, T. (2024) Gaussian Process Emulation: Theory and Application to Coupled Physics. PhD Thesis.
7. Domingo, D. et al. (2020) "Using ice cores and Gaussian process emulation to recover changes in the Greenland ice sheet during the last interglacial," *Journal of Geophysical Research: Earth Surface*, 125(5), p. e2019JF005237.
8. Donnelly, J. et al. (2022) "Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model," *Water Research*, 225, p. 119100.
9. Fuchs, C. and Kenett, R.S. (1998) *Multivariate quality control: theory and applications*. Chapman and Hall/CRC.
10. Hill, T. et al. (2025) "Computationally efficient subglacial drainage modelling using Gaussian process emulators: GlaDS-GP v1. 0," *Geoscientific Model Development*, 18(13), pp. 4045–4074.
11. Kennedy, Marc C. and O'Hagan, A. (2001) "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 63(3), pp. 425–464. Available at: <https://doi.org/10.1111/1467-9868.00294>.
12. Kennedy, Marc C and O'Hagan, A. (2001) "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), pp. 425–464.
13. Longobardi, S. et al. (2020) "Predicting left ventricular contractile function via Gaussian process emulation in aortic-banded rats," *Philosophical Transactions of the Royal Society A*, 378(2173), p. 20190334.



14. Mansfield, L.A. et al. (2025) "Using a Gaussian Process Emulator to approximate the climate response patterns to greenhouse gas and aerosol forcings," EGU sphere, 2025, pp. 1–27.
15. McKay, M.D., Beckman, R.J. and Conover, W.J. (2000) "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code," Technometrics, 42(1), pp. 55–61. Available at: <https://doi.org/10.1080/00401706.2000.10485979>.
16. O'Hagan, A. (1978) "Curve fitting and optimal design for prediction," Journal of the Royal Statistical Society: Series B (Methodological), 40(1), pp. 1–24.
17. van Oijen, M. (2024) "Gaussian Processes and Model Emulation," Bayesian Compendium. Springer, pp. 105–117.
18. Oyebamiji, O.K. et al. (2017) "Gaussian process emulation of an individual-based model simulation of microbial communities," Journal of Computational Science, 22, pp. 69–84.
19. Paquet, J.-F. (2024) "Applications of emulation and Bayesian methods in heavy-ion physics," Journal of Physics G: Nuclear and Particle Physics, 51(10), p. 103001.
20. Pepper, N. et al. (2023) "A probabilistic model for aircraft in climb using monotonic functional Gaussian process emulators," Proceedings of the Royal Society A, 479(2271), p. 20220607.
21. Rajabi, M.M. and Ketabchi, H. (2017) "Uncertainty-based simulation-optimization using Gaussian process emulation: application to coastal groundwater management," Journal of hydrology, 555, pp. 518–534.
22. Sacks, J. et al. (1989) "Design and analysis of computer experiments," Statistical science, 4(4), pp. 409–423.

23. Sawe, S.J. et al. (2024) "Gaussian process emulation to improve efficiency of computationally intensive multidisease models: a practical tutorial with adaptable R code," BMC Medical Research Methodology, 24(1), p. 26.
24. Servera, J.V. et al. (2023) "Multifidelity Gaussian process emulation for atmospheric radiative transfer models," IEEE Transactions on Geoscience and Remote Sensing, 61, pp. 1–10.
25. Shin, Y. and Am Seo, Y. (2024) "Development of an Emulator for Radiation Physics Simulator Using the Stochastic Variational Gaussian Process Model," international journal on advanced science, Engineering & Information Technology, 14(2).
26. Tunnell, M., Bowman, N. and Carrier, E. (2023) "Fast Gaussian Process Emulation of Mars Global Climate Model," Earth and Space Science, 10(9), p. e2022EA002743.

JOBS



مجلة العلوم الأساسية
Journal of Basic Science



Print -ISSN 2306-5249
Online-ISSN 2791-3279

العدد الأربعون
٢٠٢٦ م / ١٤٤٧ هـ



مجلة العلوم الأساسية
للعلوم التربوية والنفسية وطرائق التدريس للعلوم الأساسية