الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

Sétif 1 Université - Ferhat ABBAS
Faculté des Sciences
Département de mathématiques

Mémoire de Master

DOMAINE : **Mathématiques et Informatique**
FILIÈRE : **Mathématique**
SPÉCIALITÉ : **Mathématiques Appliquées**

Thème

Modèles géostatistiques hybrides et Machine-learning pour la prédiction des propriétés pétrophysiques des réservoirs sous Python

Présenté par :
KRAMCHA Islam

encadré par :
Pr. GRICHE Sonia

Mémoire Soutenu le 21 juin 2026, devant le jury composé de :

Présidente : Pr. AMRANI Naima	Université Sétif 1 Ferhat Abbas
Encadrante : Pr. GRICHE Sonia	Université Sétif 1 Ferhat Abbas
Examinateur : Pr. KHABER Alarbi	Université Sétif 1 Ferhat Abbas

Promotion : 2025/2026

Table des matières

1 Introduction générale et cadre méthodologique de l'étude **4**
- 1.1 Contexte général et cadre numérique . . . 4
- 1.2 Problématique . . . 4
- 1.3 Objectifs de l'étude . . . 4
- 1.4 Méthodologie générale . . . 4
- 1.5 Rôle des environnements R et Python . . . 4
- 1.6 Données utilisées et analyse exploratoire . . . 4
- 1.7 Organisation du mémoire . . . 4
- 1.8 Conclusion du chapitre . . . 4

2 Analyse exploratoire des données sous Python **5**
- 2.1 Introduction . . . 5
- 2.2 Présentation du jeu de données . . . 5
- 2.3 Nettoyage et préparation des données . . . 5
- 2.4 Analyse statistique descriptive . . . 5
- 2.5 Analyse des distributions . . . 5
- 2.6 Transformation des variables . . . 5
- 2.7 Analyse des corrélations . . . 5
- 2.8 Analyse spatiale des données . . . 5
- 2.9 Détection des valeurs aberrantes . . . 5
- 2.10 Première lecture de la continuité spatiale . . . 5
- 2.11 Synthèse générale du chapitre . . . 5
- 2.12 Conclusion . . . 5

3 Modélisation de la structure spatiale par variogramme **6**
- 3.1 Introduction . . . 6
- 3.2 Rappel sur la variable régionalisée . . . 6
- 3.3 Hypothèse de stationnarité . . . 6
- 3.4 Définition du variogramme . . . 6
- 3.5 Préparation des données sous Python . . . 6
- 3.6 Calcul du variogramme expérimental . . . 6
- 3.7 Interprétation du variogramme expérimental . . . 6
- 3.8 Types de variogrammes théoriques . . . 6
- 3.9 Implémentation des modèles sous Python . . . 6
- 3.10 Comparaison graphique des modèles . . . 6
- 3.11 Choix du modèle le plus adapté . . . 6
- 3.12 Limites du variogramme . . . 6
- 3.13 Conclusion . . . 6

4 Interpolation spatiale des propriétés pétrophysiques par krigeage **7**
- 4.1 Introduction . . . 7
- 4.2 Définition du krigeage . . . 7
- 4.3 Condition de non-biais . . . 7
- 4.4 Minimisation de la variance de l'erreur . . . 7
- 4.5 Système du krigeage ordinaire . . . 7
- 4.6 Les différents types de krigeage . . . 7

4.7 Application à la porosité (PHIE1) 7
4.8 Application à la perméabilité (logK1) 7
4.9 Mise en œuvre avancée sous Python 7
4.10 Résultats et interprétation 7
4.11 Conclusion 7

5 Modélisation des propriétés pétrophysiques par Machine Learning et approche hybride 8
5.1 Introduction 8
5.2 Définitions générales du Machine Learning 8
5.3 Intérêt du Machine Learning dans l'étude des réservoirs 8
5.4 Données et formulation du problème 8
5.5 Principaux modèles de Machine Learning utilisés 8
5.6 Implémentation sous Python 8
5.7 Interprétation des résultats des modèles 8
5.8 Limites des approches purement Machine Learning 8
5.9 Approche hybride : principe général 8
5.10 Implémentation hybride sous Python 8
5.11 Discussion générale du chapitre 8
5.12 Conclusion 8

6 Analyse de l'incertitude et simulation géostatistique 9
6.1 Introduction 9
6.2 Notion d'incertitude en géostatistique 9
6.3 Limites du krigeage 9
6.4 Principe de la simulation géostatistique 9
6.5 Simulation gaussienne séquentielle (SGS) 9
6.6 Validation des modèles 9
6.7 Résultats des simulations pour PHI1 9
6.8 Résultats des simulations pour K1 9
6.9 Implémentation sous Python 9
6.10 Analyse probabiliste (P10, P50, P90) 9
6.11 Cartes d'incertitude 9
6.12 Analyse des risques 9
6.13 Recommandations 9
6.14 Discussion 9
6.15 Conclusion 9

7 Discussion générale et perspectives 10
7.1 Introduction 10
7.2 Synthèse des principaux résultats 10
7.3 Apports de la géostatistique 10
7.4 Apports du Machine Learning 10
7.5 Intérêt de l'approche hybride 10
7.6 Apport de la simulation géostatistique 10
7.7 Limites du travail 10
7.8 Perspectives de recherche 10
7.9 Conclusion du chapitre 10

7.10 Importance des environnements R et Python dans la modélisation géostatistique et l'apprentissage automatique 10

Conclusion générale **11**

Résumé 12

Abstract 13

1 Introduction générale et cadre méthodologique de l'étude

1.1 Contexte général et cadre numérique

1.2 Problématique

1.3 Objectifs de l'étude

1.4 Méthodologie générale

1.5 Rôle des environnements R et Python

1.6 Données utilisées et analyse exploratoire

1.7 Organisation du mémoire

1.8 Conclusion du chapitre

2 Analyse exploratoire des données sous Python

2.1 Introduction

2.2 Présentation du jeu de données

2.3 Nettoyage et préparation des données

2.4 Analyse statistique descriptive

2.5 Analyse des distributions

2.6 Transformation des variables

2.7 Analyse des corrélations

2.8 Analyse spatiale des données

2.9 Détection des valeurs aberrantes

2.10 Première lecture de la continuité spatiale

2.11 Synthèse générale du chapitre

2.12 Conclusion

3 Modélisation de la structure spatiale par variogramme

3.1 Introduction

3.2 Rappel sur la variable régionalisée

3.3 Hypothèse de stationnarité

3.4 Définition du variogramme

3.5 Préparation des données sous Python

3.6 Calcul du variogramme expérimental

3.7 Interprétation du variogramme expérimental

3.8 Types de variogrammes théoriques

3.9 Implémentation des modèles sous Python

3.10 Comparaison graphique des modèles

3.11 Choix du modèle le plus adapté

3.12 Limites du variogramme

3.13 Conclusion

4 Interpolation spatiale des propriétés pétrophysiques par krigeage

4.1 Introduction

4.2 Définition du krigeage

4.3 Condition de non-biais

4.4 Minimisation de la variance de l'erreur

4.5 Système du krigeage ordinaire

4.6 Les différents types de krigeage

4.7 Application à la porosité (PHIE1)

4.8 Application à la perméabilité (logK1)

4.9 Mise en œuvre avancée sous Python

4.10 Résultats et interprétation

4.11 Conclusion

5 Modélisation des propriétés pétrophysiques par Machine Learning et approche hybride

5.1 Introduction

5.2 Définitions générales du Machine Learning

5.3 Intérêt du Machine Learning dans l'étude des réservoirs

5.4 Données et formulation du problème

5.5 Principaux modèles de Machine Learning utilisés

5.6 Implémentation sous Python

5.7 Interprétation des résultats des modèles

5.8 Limites des approches purement Machine Learning

5.9 Approche hybride : principe général

5.10 Implémentation hybride sous Python

5.11 Discussion générale du chapitre

5.12 Conclusion

6 Analyse de l’incertitude et simulation géostatistique

6.1 Introduction

6.2 Notion d’incertitude en géostatistique

6.3 Limites du krigeage

6.4 Principe de la simulation géostatistique

6.5 Simulation gaussienne séquentielle (SGS)

6.6 Validation des modèles

6.7 Résultats des simulations pour PHI1

6.8 Résultats des simulations pour K1

6.9 Implémentation sous Python

6.10 Analyse probabiliste (P10, P50, P90)

6.11 Cartes d’incertitude

6.12 Analyse des risques

6.13 Recommandations

6.14 Discussion

6.15 Conclusion

7 Discussion générale et perspectives

7.1 Introduction

7.2 Synthèse des principaux résultats

7.3 Apports de la géostatistique

7.4 Apports du Machine Learning

7.5 Intérêt de l'approche hybride

7.6 Apport de la simulation géostatistique

7.7 Limites du travail

7.8 Perspectives de recherche

7.9 Conclusion du chapitre

7.10 Importance des environnements R et Python dans la modélisation géostatistique et l'apprentissage automatique

Conclusion générale

Résumé

Ce mémoire développe une méthodologie intégrée pour la prédiction des propriétés pétrophysiques (perméabilité K1 et porosité PHIE1) des réservoirs pétroliers en combinant géostatistique et apprentissage automatique (Machine Learning) sous Python. À partir de données spatiales réelles (fichier TPgroupe2 . x1s), une analyse exploratoire et une transformation logarithmique (logK1) sont réalisées. La structure spatiale est modélisée par variogramme (modèle sphérique), puis le krigeage ordinaire est appliqué pour produire des cartes d'estimation et d'incertitude.

Les modèles de Machine Learning (Random Forest, XGBoost) améliorent les prédictions non linéaires mais ignorent la dépendance spatiale. Une approche hybride est proposée : elle combine la tendance de Random Forest avec le krigeage des résidus, améliorant à la fois la précision et la cohérence spatiale. Enfin, une simulation gaussienne séquentielle (SGS) génère de multiples réalisations, permettant de cartographier les quantiles P10, P50, P90 et l'écart-type (spread), quantifiant rigoureusement l'incertitude spatiale pour l'aide à la décision.

Abstract

This thesis presents an integrated methodology for predicting petrophysical properties (permeability K1 and porosity PHIE1) of oil reservoirs by combining geostatistics and machine learning using Python. Based on real spatial data (file TPgroupe2. xls), an exploratory analysis and a logarithmic transformation (log K1) are performed. Spatial continuity is modeled using an experimental variogram (spherical model), followed by ordinary kriging to generate estimation maps with uncertainty.

Machine learning models (Random Forest, XGBoost) capture nonlinear relationships but ignore spatial dependence. A hybrid approach combining the Random Forest trend with kriging of residuals is proposed, improving both predictive accuracy and spatial consistency. Sequential Gaussian Simulation (SGS) generates multiple equiprobable realizations, yielding P10 (pessimistic), P50 (median), P90 (optimistic) maps and a spread (P90 - P10) map, providing rigorous quantification of spatial uncertainty for decision support in reservoir engineering.

ملخص

تقدم هذه المذكرة منهجية متكاملة للتنبؤ بالخصائص الصخرية للمكامن النفطية (النفاذية والمسامية) من خلال دمج تقنيات الإحصاء الجغرافي وتعلم الآلة باستخدام بايثون. تم الاعتماد على بيانات مكانية حقيقية مستمدة من ملف البيانات. بدأت الدراسة بتحليل استكشافي شامل للبيانات، وتطبيع المتغيرات عبر التحويل اللوغاريتمي للنفاذية. تم نمذجة الاستمرارية المكانية باستخدام الرسم البياني التجريبي للتباين المكاني واختيار النموذج النظري الأكثر ملاءمة (الكروي، الأسي، الغاوسي)، ثم تطبيق طرق الكريجين لإنشاء خرائط الاستيفاء المكاني مصحوبة بخرائط لقياس عدم اليقين.

لمعالجة العلاقات غير الخطية والمعقدة بين المتغيرات، تم تطوير نماذج تعلم الآلة (الغابات العشوائية وإكس جي بوست) وأظهرت تفوقاً على الانحدار الخطي الكلاسيكي. لكن لتجاوز محدودية تعلم الآلة في إغفال البعد المكاني، تم اقتراح النموذج الهجين الذي يجمع بين اتجاه نموذج الغابات العشوائية وإعادة كريجينة البواقي، مما أدى إلى تحسين الدقة التنبؤية والاتساق المكاني معاً.

تم أخيراً إجراء محاكاة غاوسية متسلسلة لتوليد مئة تحقيق مكاني، واستخراج خرائط الاحتمالات للمئين العاشر والخمسين والتسعين (السيناريو المتشائم والمتوسط والمتفائل) وخرائط الانتشار لتقييم عدم اليقين المكاني ودعم اتخاذ القرار. أثبتت النتائج أن النموذج الهجين والمحاكاة الجيولوجية يقدمان أداة قوية لتحسين توصيف المكامن وتحديد المناطق الواعدة والمناطق عالية المخاطرة.