

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Nationale Polytechnique
Département de génie minier



*Laboratoire Matériaux
de Génie Civil et Environnement*



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Laboratoire de Génie Minier

Thèse de Doctorat

En Génie Minier

Présentée par :

MAZARI Messaoud

**Développement de méthodes d'approche géostatistique
pour l'estimation des réserves minières appliquée au cas
de gisement de Bled El-Hadba (Tébessa)**

Soutenue le 19/05/2024 devant le jury composé de :

Président	M. YAHYAOUI Sami	Professeur	ENP
Promoteur	M. BALI Abderrahim	Professeur	ENP
Co-Promoteur	Mme. CHABOU Salima	MCA	ENP
Examineur	M. AKKAL Rezki	Professeur	ENP
Examineur	M. MAZA Mustapha	Professeur	Univ. A.M. de Bejaia
Examineur	M. BOUTRID Abdelaziz	MCA	Univ. A.L. de Khenchela
Examineur	M. MAZOUZ El-Hadi	MCA	Univ. de Oum El- Bouaghi

ENP 2024

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Nationale Polytechnique
Département de génie minier



*Laboratoire Matériaux
de Génie Civil et Environnement*



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



Laboratoire de Génie Minier

Thèse de Doctorat

En Génie Minier

Présentée par :

MAZARI Messaoud

**Développement de méthodes d'approche géostatistique
pour l'estimation des réserves minières appliquée au cas
de gisement de Bled El-Hadba (Tébessa)**

Soutenue le 19/05/2024 devant le jury composé de :

Président	M. YAHYAOUI Sami	Professeur	ENP
Promoteur	M. BALI Abderrahim	Professeur	ENP
Co-Promoteur	Mme. CHABOU Salima	MCA	ENP
Examineur	M. AKKAL Rezki	Professeur	ENP
Examineur	M. MAZA Mustapha	Professeur	Univ. A.M. de Bejaia
Examineur	M. BOUTRID Abdelaziz	MCA	Univ. A.L. de Khenchela
Examineur	M. MAZOUZ El-Hadi	MCA	Univ. de Oum El- Bouaghi

ENP 2024

ملخص

يتضمن هذا العمل تنفيذ نهج جيواحصائي يهدف إلى تحسين إجراءات تقدير المخزون لمشاريع التعدين. الطريقة المقترحة تعتمد على دراسة اختيار النماذج التغيرية الأكثر تمثيلاً للمتغيرات المستقصاة من خلال إرفاق المنحنى، التحقق المتقاطع، والكريجاج. باستخدام طريقة المربعات الدنيا، تزامنا مع معايير الدقة والفعالية، يسمح بالتحكم في التعديل وتقييم جودته. يتم تحديد اختيار النموذج المستخدم من خلال تطبيق إجراء التحقق المتقاطع (LOOCV) حيث يتم إجراء تقدير المتغيرات المدروسة بعد ذلك، باستخدام الكريجاج العادي. تم دمج مراحل النهج المقترح في برنامج حساسي، تم تطويره باستخدام Matlab، بهدف تسهيل وتأمين إجراءات التقدير.

أظهرت الدراسات السابقة على ممكن بلدة الهظبة (شرق الجزائر) وجود تباينات ملحوظة في تقدير الاحتياطيات، لا سيما في الطبقة المتوسطة، حيث قدر المكتب الوطني للبحث الجيولوجي و المنجمي (ORGM) في عام 2015 الاحتياطيات بـ 367 مليون طن في حين قدر مكتب دراسات التكنولوجيا التعدينية الألمانية (DMT) في عام 2016 الاحتياطيات بـ 464 مليون طن. لقد دفعتنا هذه الاختلافات إلى إجراء دراسة معمقة للمنهجيات المستخدمة واستخدام هذا الممكن كنموذج لتطبيق المنهجية المقترحة في هذه الأطروحة.

تطبيق النهج المقترح على ممكن بلدة الهظبة أدى إلى نتائج مرضية من حيث التنبؤات لموضع تراكيز الفوسفات وسكافة طبقة الفوسفات. حيث تظهر معايير الدقة والفعالية المتعلقة بالمتغيرات المدروسة قيماً مقبولة. يحمل هذا العمل أهمية فنية واقتصادية من خلال تعزيز إجراءات الحساب الجيواحصائي، وتحسين دقة تقدير الاحتياطيات المنجمية، مما يؤدي إلى تحسين إدارة واستغلال المواقع التعدينية.

الكلمات المفتاحية : الجيواحصائيات، الكريجاج، إرفاق المنحنى، التحقق المتقاطع (LOOCV)، الاحتياطيات المنجمية.

Abstract

This work involves implementing a geostatistical approach aimed at improving the parameter estimation procedure for a mining project. The proposed method is based on studying the selection of the most representative variographic models for the variables under investigation through a coupling of adjustment, cross-validation, and kriging. The use of the least squares method, combined with precision and efficiency criteria, allows for control of the adjustment and assessment of its quality. The choice of the model to use is determined through the application of a leave-one-out cross-validation (LOOCV) procedure. The estimation of the variables under study is then performed using ordinary kriging. The various phases of the proposed approach are integrated into a computational code developed in Matlab, with the aim of facilitating and automating the estimation procedure.

Previous studies on the Bled El Hadba deposit (Eastern Algeria) have revealed significant discrepancies in the reserve estimation, particularly in the middle layer, with the National Office of Geological and Mining Research (ORGM) in 2015 estimating 367 million tons, while Deutsche Montan Technologie (DMT) in 2016 estimated 464 million tons. This divergence motivated us to conduct an in-depth study of the methodologies employed and to use this deposit as an application example for the approach proposed in this thesis.

The application of the proposed approach to the Bled El-Hadba deposit yields satisfactory results in terms of the dispersion of concentrations and thicknesses of the mineralized layer. The precision and efficiency criteria corresponding to the variables under study exhibit acceptable values. This work holds both technical and economic significance by consolidating the geostatistical calculation procedure and improving the precision of the estimation of the mining reserves, leading to better management and exploitation of mining sites.

Keywords : Geostatistics, Kriging, adjustment, cross-validation (LOOCV), mineral reserves.

Résumé

Ce travail consiste à mettre en œuvre une approche géostatistique qui vise à améliorer la procédure d'estimation des réserves minières. La démarche proposée se base sur l'étude du choix des modèles variographiques les plus représentatifs des variables étudiées, à travers un couplage (par combinaison) ajustement-validation croisée-krigeage. L'emploi de la méthode des moindres carrés, associé à des critères de précision et d'efficacité permet de contrôler l'ajustement et d'évaluer sa qualité. Le choix du modèle à utiliser est effectué par application d'une procédure de validation croisée leave-one-out (LOOCV). L'estimation des variables étudiées est ensuite faite par krigeage ordinaire. Les différentes phases de la démarche proposée sont intégrées dans un code de calcul développé sous Matlab, dans le but de faciliter et d'automatiser la procédure d'estimation.

Des études antérieures sur le gisement de Bled El Hadba (Est Algérien) ont mis en évidence des écarts notables dans l'estimation des réserves, notamment dans la couche médiane, avec l'Office National de la Recherche Géologiques et Minière (ORGM) en 2015 estimant 367 millions de tonnes, tandis que le bureau d'études de la Technologie Minière Allemande (DMT) en 2016 estimait 464 millions de tonnes. Cette divergence nous a motivés à réaliser une étude approfondie des méthodologies employées et à utiliser ce gisement comme exemple d'application pour l'approche proposée au cours de cette thèse.

La mise en application de l'approche proposée sur le gisement de Bled El-Hadba conduit à des résultats satisfaisants en termes de dispersion des teneurs et des épaisseurs de la couche minéralisée. Les critères de précision et d'efficacité correspondant aux variables étudiées affichent des valeurs admissibles. Le présent travail porte un intérêt technique et économique à travers, la consolidation de la procédure de calcul géostatistique, et l'amélioration de la précision de l'estimation des réserves minières, conduisant à une meilleure gestion et exploitation des sites miniers.

Mots clés : Géostatistique, Krigeage, ajustement, validation croisée (LOOCV), réserves minières.

Dédicaces

A mes chers parents,

A qui je dois ce que je suis,

Qu'ils trouvent dans cette thèse, le fruit de leurs sacrifices consentis pour mon éducation, et l'expression de mon amour et de ma gratitude pour la bienveillance avec laquelle ils m'ont toujours entouré.

Que Dieu leur préserve longue vie et bonne santé.

A ma grande mère et à la mémoire de mes grands-parents.

Ma femme et ma petite Nada

A mes sœurs,

A mes frères,

A toute ma famille,

A tous mes amis.

Messaoud

Remerciements

Merci à notre bon Dieu, notre guide, notre force, notre bonheur, et la raison de notre existence. C'est lui qui nous a fait comprendre le but de cette vie, et qui nous a donné le pouvoir d'aimer les gens et d'apprécier les choses. Merci d'être là dans les moments les plus difficiles.

Je tiens à exprimer toute ma reconnaissance et ma gratitude à Madame CHABOU Salima et au Professeur BALI Abderrahim, pour leurs soutiens, leurs patiences et leurs orientations.

Une grande reconnaissance et un grand merci aux enseignants du département Génie minier, qui ont contribué à ma formation et de m'avoir donné la chance d'apprendre, et de nous faire partager leurs savoirs, sans eux je ne serai jamais arrivé à ce stade.

Je tiens à remercier :

Monsieur Sami YAHYAOUI, Professeur à l'Ecole Nationale Polytechnique d'Alger, de m'avoir fait l'honneur d'accepter la présidence du Jury.

Monsieur Rezki AKKAL, Professeur à l'Ecole Nationale Polytechnique d'Alger, qui a accepté d'examiner cette thèse et de faire partie du Jury.

Monsieur Mustapha MAZA, Professeur à l'Université Abderrahmane Mira de Béjaïa, qui a accepté d'examiner cette thèse et de faire partie du Jury.

Monsieur Abdelaziz BOUTRID, Maître de conférences à l'Université Abbes Laghrour de Khenchela, qui a accepté d'examiner cette thèse et de faire partie du Jury.

Monsieur El-Hadi MAZOUZ, Maître de conférences à l'Université Larbi Ben M'hidi de Oum El-Bouaghi, pour avoir accepté d'examiner cette thèse et de faire partie du jury.

*Mes remerciements vont à la **direction générale des mines (ministère de l'énergie et des mines)**. Je remercie infiniment tout le personnel (amis et collègues): les directeurs et sous-directeurs, les ingénieurs et les administrateurs.*

*Mes remerciements vont aussi à la **Société National du Fer et du Phosphates (FERPHOS)** pour l'accueil chaleureux dans les unités d'exploitation et de traitement à Bir-el-Ater. Je remercie infiniment tout le personnel : les directeurs de la société, les ingénieurs et les mineurs, qui sont en activité ou en retraite.*

Cette thèse n'aurait jamais pu voir le jour sans la contribution de mes amis et collègues : M. Khaled KOUIDER de l'Université HOUARI BOUMEDIEN, M.Badre Eddin CHEMALI de l'entreprise Sonatrach et à tous mes collègues du département des mines et géologie de l'Université de bejaia. Je leurs présente mes remerciements.

Pour finir, je tiens à remercier l'ensemble de mes ami(es) et collègues pour leurs sympathies et gentillesse.

Table des matières

Liste des figures

Liste des tableaux

Liste des abréviations

Introduction générale 18

Chapitre I : Classification et évaluation des réserves minières

1. Introduction..... 22

2. Revue des normes conçues pour l'estimation des réserves minières..... 22

2.1. Code de l'industrie minière Canadian (CIM)..... 23

2.2. Codes de l'industrie minière aux États-Unis..... 23

2.2.1. Code de la Société de l'exploitation minière, de la métallurgie et de l'exploration (SME Code) 23

2.2.2. Securities and Exchange Commission 7 (SEC Industry Guide 7)..... 23

2.3. Le code de l'industrie minière sud africaine (SAMREC)..... 24

2.4. Le Code de l'industrie minière Australien (JORC) 25

3. Norme internationale de classification des ressources/ réserves minières 25

3.1. Classification des ressources et des réserves minières (Code JORC)..... 26

1.1.1. Ressources Minérales 27

1.1.2. Réserves Minérales 27

4. Méthodes d'évaluation des réserves 28

4.1. Méthode des coupes transversales (sections géologiques)..... 29

4.2. Méthode des polygones 31

4.3. Méthode des triangles..... 34

4.4. Méthode de l'inverse de la distance	34
5. Conclusion	36

Chapitre II : Méthodes géostatistiques et techniques de validation

1. Introduction.....	38
2. Historique	38
3. Théorie des variables régionalisées.....	39
3.1. Variables régionalisées et fonctions aléatoires	39
3.1.1. Stationnarité.....	40
3.2. Analyse de la dépendance spatiale.....	42
3.2.1. Autocorrélation	42
3.2.2. Hypothèse intrinsèque	43
4. Le semi-variogramme (Variogramme).....	43
4.1. Propriétés du variogramme.....	44
4.1.1. Discontinuité à l'origine « Effet de pépite »	45
4.1.2. Le seuil ou palier.....	45
4.1.3. La portée « a ».....	45
4.2. Le variogramme expérimental.....	46
4.3. Principaux modèles théoriques de variogrammes isotropes.....	48
4.3.1. Modèle pépitique de palier C.....	48
4.3.2. Modèle sphérique de portée a et de palier C.....	48
4.3.3. Modèle exponentiel de paramètre a et de palier C.....	49
4.3.4. Modèle gaussien de paramètre a et de palier C.....	49
4.4. Anisotropies	50
4.4.1. Anisotropie géométrique (anisotropie « elliptique »)	51
4.4.2. Anisotropie zonale (ou stratifiée).....	52

5.	Interpolation par krigeage.....	52
5.1.	Le krigeage ordinaire.....	54
5.2.	La variance de krigeage.....	55
6.	Techniques de validation.....	56
6.1.	Validation croisée par rétention « Hold Cross-Validation »	57
6.2.	Validation croisée « K-fold ».....	57
6.3.	Validation croisée Leave-Part-Out / Leave-One-Out.....	58
7.	Conclusion	60

Chapitre III : Géologie régionale et locale du gisement de Bled El-Hadba

1.	Introduction.....	63
2.	Présentation générale du secteur minier de Djebel Onk.....	63
2.1.	Contexte géologique du secteur minier de Djebel Onk.....	64
2.2.	Stratigraphie.....	66
2.2.1.	Crétacé.....	66
2.2.2.	Paléocène.....	66
2.2.3.	Eocène.....	68
2.2.4.	Miocène.....	69
2.2.5.	Quaternaire.....	69
2.3.	Tectonique régionale.....	71
2.3.1.	Tectonique synsédimentaire.....	72
2.3.2.	Tectonique post-éocène moyen et anté-miocène.....	72
2.3.3.	Tectonique tardi-miocène, post-burdigalienne, à pliocène-pléistocène.....	72
3.	Présentation du gisement de Bled El-Hadba.....	74
3.1.	Géologie du gisement de Bled El-Hadba.....	74
3.2.	Litho-stratigraphie.....	76
3.2.1.	Paléocène.....	77

3.2.2.	Thanétien	77
3.2.3.	Miocène	77
3.2.4.	Quaternaire	78
3.3.	Tectonique du gisement.....	78
3.4.	Morphologie du corps de minerai de phosphate de Bled El-Hadba.....	78
3.4.1.	La Couche Basale.....	79
3.4.2.	La Couche Médiane	79
3.4.3.	La Couche Sommitale	79
4.	Travaux antérieurs effectués sur le gisement de Bled El Hadba	80
4.1.	Evaluation des réserves effectuée par d'ORGM.....	82
4.2.	Evaluation des réserves effectuée par DMT (2016).....	82
5.	Conclusion	83

Chapitre IV : Méthodologie et structure de l'approche employée

1.	Introduction.....	85
2.	Description de l'approche employée.....	85
2.1.	Structure de l'outil élaboré	85
2.1.1.	Variogrammes Expérimentaux et Ajustement des Modèles Théoriques	86
2.1.2.	Identification du Meilleur Modèle Ajusté via la LOOCV.....	86
2.1.3.	Établissement du Matrice de Maillage de la Zone d'Étude	86
2.1.4.	Établissement des Matrices de Krigeage.....	86
2.1.5.	Affichage des Résultats sous Forme Graphique et Tabulaire.....	86
2.2.	Organisation algorithmique pour l'approche élaborée.....	87
3.	Méthode d'ajustement	89
4.	Validation croisée.....	90
5.	Critères d'évaluation des modèles	91

6. Conclusion	92
Chapitre V : Evaluation des réserves, résultats et discussion	
1. Introduction.....	94
2. Application de l'approche géostatistique	94
2.1. Analyse statistique de données du gisement de Bled El-Hadba	95
2.1.1. Statistique descriptive et histogramme.....	95
2.2. Analyse géostatistiques du gisement de Bled El-Hadba	97
2.2.1. Estimation de la puissance de la couche minéralisé.....	97
2.2.2. Estimation des teneurs.....	101
2.3. Traitement statistique des résultats de l'estimation	105
2.3.1. Epaisseurs de la couche principale.....	106
2.3.2. Variance de krigeage des variables étudiés	107
2.1. Réserves de la couche médiane du gisement de Bled El-Hadba :	108
2.1.1. Classification des ressources/réserves de la couche médiane.....	109
3. Discussion.....	111
4. Conclusion	113
Conclusion générale	115
Références bibliographiques	117
Annexes	123

Liste des figures

Chapitre I : Classification et évaluation des réserves minières

Figure 1.1. La relation entre les résultats d'exploration, ressources et réserves minières.	27
Figure 1.2. Estimation des ressources/réserves par extrapolation des sections transversales..	29
Figure 1.3. Représentations d'un contact irrégulier entre le minerai et les parties stérile.	30
Figure 1.4. Construction d'une zone des polygones à laquelle une seule teneur contenue est appliquée (la teneur de l'échantillon contenu est étendue au polygone)	32
Figure 1.5. Ampleur des erreurs possibles lors de l'application d'une méthode des polygones à l'estimation des ressources/réserves.	33
Figure 1.6. Illustration de l'estimation par la méthode de l'inverse de la distance.	35

Chapitre II : Méthodes géostatistiques et techniques de validation

Figure 2.1. Relation entre la covariance spatiale $C(h)$ et le semi-variogramme $\gamma(h)$	42
Figure 2.2. (A) Variogramme idéalisé avec une variance de pépité nulle; (B) variogrammes observés pour les propriétés du sol avec un effet de pépité.....	45
Figure 2.3. (A) Variogrammes bornés et (B) Variogrammes non bornés	46
Figure 2.4. Variogramme expérimental du champ aléatoire dans les principales directions d'anisotropie	47
Figure 2.5. A ; B. Schéma type d'un variogramme (effet de pépité).	48
Figure 2.6. Schéma type d'un modèle de variogramme sphérique	49
Figure 2.7. Schémas type d'un modèle de variogramme : (A) Exponentiel, (B) Gaussien [39].	50
Figure 2.8. Exemple 2D d'anisotropie géométrique de rapport a_1/a_2 et l'angle θ	51
Figure 2.9. Mécanisme d'application de la méthode validation croisée K-Fold.....	58
Figure 2.10. Représentation schématique de la méthode de validation croisée Leave-Part-Out « LPOCV ».....	59
Figure 2.11. Représentation schématique de la méthode de validation croisée Leave-One-Out « LOOCV ».....	60

Chapitre III : Géologie régionale et locale du gisement de Bled El-Hadba

Figure 3.1. Carte de situation géographique de la région de Djebel Onk.....	64
Figure 3.2. Carte de situation géographique et géologique des gisements de Djebel Onk.....	65
Figure 3.3. Colonne stratigraphique de la région de Djebel Onk – Algérie orientale.....	70

Figure 3.4. Carte structurale de la région du Djebel Onk à l'intérieur du bassin de Gafsa Méthlaoui.....	71
Figure 3.5. Coupes stratigraphiques et structurales des flancs Nord et Sud du Djebel Onk. ..	73
Figure 3.6. Carte de la situation géographique du gisement de Bled El-Hadba.	74
Figure 3.7. Carte géologique du gisement de Bled El Hadba	75
Figure 3.8. Litho-stratigraphie du gisement de Bled El Hadba.....	76
Figure 3.9. Plan d'implantation des sondages des différentes campagnes de prospection.....	81

Chapitre IV : Méthodologie et structure de l'approche employée

Figure 4.1. Diagramme du processus de calcul impliquant l'ajustement des modèles, la validation croisée (LOO), et le krigeage.....	88
---	----

Chapitre V : Evaluation des réserves, résultats et discussion

Figure 5.1. Histogramme de la puissance de la couche médiane.	96
Figure 5.2. Histogramme illustrant les teneurs en P_2O_5 de la couche médiane.	97
Figure 5.3. Variogrammes de validation des épaisseurs de la couche principale	98
Figure 5.4. Modèle variographique retenu des épaisseurs de la couche principale.	99
Figure 5.5. Carte de dispersion des épaisseurs de la couche principale	100
Figure 5.6. Carte de la variance d'estimation des épaisseurs de la couche principale.	101
Figure 5.7. Variogrammes de validation des teneurs de la couche principale.....	102
Figure 5.8. Modèle variographique retenu des teneurs de la couche principale.	102
Figure 5.9. Carte de dispersion des teneurs de la couche principale.	104
Figure 5.10. Carte de la variance d'estimation des teneurs de la couche principale.....	105
Figure 5.11. Histogramme et courbe de fonction de densité cumulative des : (A) teneurs en phosphate (B) épaisseurs de la couche médiane.	106
Figure 5.12. Histogramme et graphe de fonction de densité cumulative des variances de krigeage de l'estimation (A) des épaisseurs et (B) des teneurs de la couche médiane.	107
Figure 5.13. Ressources de la couche médiane de gisement de Bled El-Hadba.....	109
Figure 5.14. Ressources et réserves en P_2O_5 de la couche médiane de gisement de Bled El-Hadba.....	110

Liste des tableaux

Chapitre V : Evaluation des réserves, résultat et discussion

Tableau 5.1 : Critères d'évaluation de la qualité d'ajustement	99
Tableau 5.2 : Critères d'évaluation de la qualité d'ajustement	103
Tableau 5.3 : Réserves prouvée et probable en P ₂ O ₅ de la couche médiane du gisement de Bled El-Hadba.	110

Annexe 4

Tableau 1. Comparaison entre le système de classification soviétique historique et les codes de déclaration Internationaux (Mining Journal, Londres, 20 août 2004 ; Stephen Henley : Russian Mineral Reporting).....	131
---	-----

Liste des sigles et abréviations

a : Portée du semi-variogramme.

BRGM : Bureau de Recherche Géologique et Minière.

C(0) : Palier.

C0 : Effet de pépite.

CIM: Code de l'industrie minière Canadian (Canadian Institute of Mining, Metallurgy and Petroleum)

Cov : Covariance.

DMT : Bureau d'étude Allemand (Deutsche Montan Technologie).

E [] : Espérance mathématique.

e : Erreur d'estimation.

Ep : Epaisseur.

EREM : Entreprise de Recherche et d'Exploitation Minière.

FERPHOS: Entreprise d'exploitation de fer et de phosphate

h : Pas du semi-variogramme ou inter-distance moyenne.

HCV: Hold Cross-Validation ;

JORC : Code de l'industrie minière Australien (Joint Ore Reserves Committee).

K-fold : K parties (groupe) ;

K. O : Krigeage Ordinaire.

LOOCV : leave-one-out cross-validation

LPOCV: Leave-Part-Out cross validation

L(λ) : Lagrangien.

m : Moyenne.

MSE : Erreur moyenne quadratique.

Mt : Millions de tonnes.

N(h) : Nombre de paires d'observations séparées par la distance h.

NSE : Critère d'efficacité de Nash-Sutcliffe

ORGM : Office National de Recherche Géologique et Minière.

P₂O₅ : Pent-oxyde de Phosphore.

RMSE : Erreur moyenne quadratique (Root Mean Square Error).

RSR : Ratio de l'erreur quadratique moyenne sur l'écart-type des observations.

SAMREC : Code de l'industrie minière sud-africain (South African Mineral Resource Committee).

SEC Industry Guide 7: Code de l'industrie minière aux États-Unis (Securities and Exchange Commission 7).

SME : Code de l'industrie minière aux États-Unis (Society for Mining, Metallurgy, and Exploration).

SONAREM : Société Nationale de Recherche et d'Exploitation Minière

SOMIPHOS : Société des Mines de Phosphate

tm : Teneur moyenne.

V : Volume.

Var : Variance.

Z (x) : Fonction aléatoire.

γ (**h**) : Semi-variogramme théorique.

γ_e (**h**) : Semi-variogramme expérimental.

θ : Angle directionnel du variogramme.

λ : Poids de krigeage.

μ : Multiplicateur de Lagrange.

σ_e^2 : Variance d'estimation.

Introduction générale

Introduction générale

La reconnaissance du sous-sol est une phase fondamentale dans la réussite d'un projet minier. Cette reconnaissance se fait à partir des observations récoltées sur le terrain, par des mesures géophysiques ou encore des forages carottés [1]. Cela augmente considérablement le coût du projet, et peut générer également des incertitudes qui sont souvent liées à la distribution des sondages. Ces observations constituent des données d'entrée des modèles géologiques, qui servent à présenter le corps minéralisé et à estimer les réserves minières. Ces modèles géologiques sont bâtis sur une large part d'interprétation du géologue à partir de différentes méthodes d'interpolation, en l'occurrence : la méthode des polygones, la méthode de l'inverse de la distance [2]. Aussi l'utilisation des méthodes d'interpolation géostatistiques qui permettent d'étudier la variabilité spatiale des différents paramètres est importante [2, 3].

L'estimation des réserves est une étape essentielle dans les études de faisabilité et la planification minière. Bien que des méthodologies avancées existent, elles peuvent ne pas être adaptées à chaque environnement géologique complexe. Plusieurs chercheurs ont proposé des modèles de prédiction des teneurs en utilisant des techniques telles que l'inverse de la distances, le krigeage et la simulation stochastique [4].

Néanmoins, le choix de la méthodologie appropriée pour l'estimation des réserves dans un projet minier dépend de divers facteurs tels que la complexité de l'environnement géologique, la disponibilité des données et le niveau de précision requis. Quoi qu'il en soit, la précision et la fiabilité des estimations des réserves sont essentielles pour déterminer la faisabilité économique d'un projet minier. Par conséquent, le choix de la méthodologie appropriée doit être soigneusement évalué.

Des études antérieures sur le gisement de Bled El Hadba ont mis en évidence des écarts notables dans les estimations des réserves. La première étude, menée par l'Office National de la Recherches Géologiques et Minière (ORGM) en 2015, estimait les réserves de la couche médiane à environ 367 millions de tonnes, tandis que la seconde, réalisée par le bureau d'études de la Technologie Minière Allemande (DMT) en 2016, évaluait les réserves à environ 464 millions de tonnes pour la même couche. Il est à noter que ces réserves ont été estimées pour des teneurs de coupure de 16 % pour l'ORGM et de 20 % pour DMT respectivement.

Cette différence a constitué le principal défi qui a motivé notre engagement dans cette étude approfondie. En présence d'une divergence significative entre ces deux évaluations, une

analyse approfondie des méthodologies employées, pour les paramètres considérés, ainsi que des variations potentielles dans les procédures d'estimation s'avèrent nécessaire.

Ce travail vise à développer une approche géostatistique qui permettra d'étudier le choix des modèles variographiques à utiliser dans la procédure d'estimation des réserves minières. Le meilleur modèle sera sélectionné à travers un couplage ajustement-validation croisée en utilisant la méthode des moindres carrés et la Validation Croisée Leave-One-Out (LOOCV). L'implémentation de cette approche consolidera la procédure de calcul géostatistique, améliorant ainsi l'estimation des réserves [5, 6]. À la suite de l'étude variographique, le couplage ajustement-validation croisée intervient systématiquement dans la sélection du modèle de variogramme correspondant à l'erreur quadratique minimale de validation. Dans ce contexte, plusieurs facteurs d'efficacité, tels que Nash–Sutcliffe, Root Mean Square Error (RMSE) et Ratio de l'erreur quadratique moyenne sur l'écart-type des observations (RSR), seront utilisés pour évaluer la qualité des modèles d'ajustement. L'intégration de LOOCV en tant que méthode de validation croisée dans la procédure d'ajustement rendra le calcul plus efficace, permettant une exploitation optimale de toutes les données du site d'étude. Cela conduira à une représentativité accrue du modèle choisi, surtout dans le cas d'un site comportant des sondages non uniformément répartis.

La présente étude revêt une importance technique et économique en fournissant une évaluation solide des réserves minières, contribuant ainsi à l'optimisation du contrôle des teneurs, à une gestion efficace du processus d'extraction et à un traitement du minerai efficace.

Le texte qui traduit l'ensemble de cette étude s'articulera en cinq (5) chapitres :

Le premier chapitre détaille les fondements essentiels de l'estimation des réserves minières, mettant en évidence les normes internationales d'une part, tout en explorant d'autre part les méthodes traditionnelles dédiées à cette estimation, en soulignant leurs limites.

Dans le deuxième chapitre, l'accent est mis sur l'efficacité d'estimation par du krigeage ordinaire, tout en examinant les méthodes de validation croisée visant à améliorer la précision des estimations géo-spatiales.

Le troisième chapitre fournit une vue d'ensemble des données d'exploration du gisement de phosphate de Bled El Hadba, jetant ainsi les notions de base pour une compréhension approfondie du site minier.

Le quatrième chapitre expose la conception de l'approche proposée pour améliorer l'estimation des paramètres des projets miniers. Il met particulièrement en évidence l'utilisation de la méthode des moindres carrés, de la validation croisée, et du krigeage, intégrées dans un code de calcul développé sous MATLAB.

Le dernier chapitre met en pratique l'approche « ajustement-validation croisée- krigeage » sur le gisement de Bled El-Hadba situé au Nord-Est de Algérie, testant ainsi la performance de la méthodologie développée pour l'estimation des réserves minières.

En fin, la conclusion résume l'ensemble des résultats obtenus, facilitant ainsi l'analyse des sources d'incertitude liées aux études antérieures sur le gisement de Bled El-Hadba. Elle propose également des recommandations et des perspectives futures dans le domaine de l'exploitation des données géo-spatiales.

Chapitre I : Classification et évaluation des réserves minières

1. Introduction

L'estimation des réserves minières est un élément fondamental de l'industrie minière. Elle consiste à évaluer la quantité, la qualité, la taille, la forme et l'emplacement des gisements des minéraux utiles, fournissant ainsi les bases essentielles pour la planification, l'exploitation et la rentabilité des opérations minières [7, 8]. Cette estimation est cruciale pour déterminer la viabilité économique d'un projet minier et répond à une série de normes et de réglementations internationales [8, 9]. Toutefois, elle comporte de nombreux défis et sources d'incertitude, allant des données géologiques limitées à l'évolution des techniques d'estimation [9].

Dans ce chapitre, nous introduisons les fondements essentiels de l'estimation des réserves minières. Nous examinerons en premier lieu les normes de classification des ressources et des réserves à l'échelle internationale, tout en soulignant l'importance de ces directives pour garantir la cohérence et la transparence dans l'industrie minière [10-13], par la suite nous aborderons les multiples sources d'incertitudes qui sont intrinsèquement liées à ces évaluations, en mettant en évidence les défis et les variables qui doivent être pris en compte pour parvenir à des estimations précises et fiables [14]. Ce chapitre projettera ainsi les bases nécessaires pour une compréhension approfondie de l'évaluation des réserves dans le domaine des ressources minérales.

2. Revue des normes conçues pour l'estimation des réserves minières

L'estimation des ressources minérales implique l'évaluation de la quantité, la qualité, les dimensions, la configuration et de la localisation des gisements [14, 15]. Pendant de nombreuses décennies, l'industrie minière a considéré l'estimation et la classification des ressources/réserves comme une simple opération nécessitant des connaissances mathématiques et géologiques de base [14, 16]. La plupart des méthodes étaient basées sur des procédures géométriques et la distribution spatiale des données. Par conséquent, l'incertitude associée aux tonnages et aux teneurs était soit ignorée, soit mal gérée, bien que divers codes miniers exigent une mesure de la confiance dans les valeurs rapportées [14]. La classification des ressources minérales en catégories définies en fonction de leurs degrés de certitude est un élément crucial de la gestion de l'industrie minière [14, 17].

Les normes de déclaration des ressources et des réserves diffèrent d'un pays à l'autre [18], et l'estimateur de ressources doit veiller à suivre les réglementations appropriées [17]. Quelques exemples d'exigences et de normes de déclaration comprennent le guide 7 de l'industrie de la Securities and Exchange Commission 7 et le code SME pour les États-Unis

[11, 19, 20], La Norme CIM pour le Canada [12, 21], le Code JORC pour l'Australie et la Nouvelle-Zélande [13, 22], et le Code SAMREC pour l'Afrique du Sud [10].

2.1. Code de l'industrie minière Canadian (CIM)

La Norme CIM (Canadian Institute of Mining, Metallurgy and Petroleum) établit des définitions claires pour les ressources et les réserves, et elle fournit des directives précises sur les procédures d'estimation de la teneur, de la taille, de la forme et de la continuité des gisements minéraux [12, 21]. Elle met également l'accent sur la présentation des résultats dans les rapports publics, encourageant la divulgation complète des informations géologiques et techniques.

L'un des aspects clés de la Norme CIM est son engagement envers la responsabilité et la durabilité dans l'industrie minière [12]. Elle encourage les entreprises à prendre en compte les facteurs environnementaux, sociaux et économiques lors de la classification des ressources et des réserves, favorisant ainsi une approche plus holistique de la gestion des actifs miniers.

2.2. Codes de l'industrie minière aux États-Unis

L'industrie minière aux États-Unis utilise généralement deux codes pour la représentation et la classification des ressources et réserves minières. Ces deux codes sont le Code SME [11], qui est élaboré par la Society for Mining, Metallurgy, and Exploration, et le SEC Industry Guide 7 (ou Guide 7 de la Securities and Exchange Commission), qui est publié par la Securities and Exchange Commission (SEC) des États-Unis [19, 20].

2.2.1. Code de la Société de l'exploitation minière, de la métallurgie et de l'exploration (SME Code)

Le Code SME [11] est un ensemble de normes et de pratiques recommandées pour l'industrie minière aux États-Unis. Il est utilisé pour l'estimation, la classification et la communication des ressources et réserves minérales. Le Code SME est généralement considéré comme un guide de référence dans le secteur minier. Il est utilisé par de nombreuses sociétés minières pour publier des informations sur leurs ressources et réserves conformément aux meilleures pratiques de l'industrie.

2.2.2. Securities and Exchange Commission 7 (SEC Industry Guide 7)

Le SEC Industry Guide 7 est un document publié par la Securities and Exchange Commission (SEC) des États-Unis [19, 20]. Il établit des directives spécifiques pour la divulgation d'informations sur les propriétés minières dans les documents déposés auprès de

la SEC par les sociétés minières cotées en bourse. Le Guide 7 définit les critères et les normes que les sociétés minières doivent suivre pour la divulgation de leurs ressources et réserves, y compris les classifications.

Les sociétés minières cotées en bourse aux États-Unis doivent se conformer aux exigences du Guide 7 de la SEC pour la divulgation de leurs propriétés minières dans les rapports aux actionnaires et les dépôts réglementaires [20]. Le Code SME est souvent utilisé en conjonction avec le Guide 7 pour s'assurer que les informations sur les ressources et réserves sont conformes aux normes internationales de l'industrie.

L'une des caractéristiques les plus importantes de ces codes est leurs insurances sur la séparation des ressources et des réserves en fonction du niveau de certitude associé à chaque catégorie [23, 24]. Les ressources sont définies comme des occurrences minérales pouvant devenir des réserves sous certaines conditions, tandis que les réserves sont des matériaux économiquement exploitables avec un haut degré de certitude. Cette distinction est cruciale pour les investisseurs et les régulateurs.

2.3. Le code de l'industrie minière sud africaine (SAMREC)

Le Code SAMREC, qui signifie "South African Mineral Resource Committee," est une norme de classification des ressources et des réserves minérales spécifiquement élaborée pour l'industrie minière en Afrique du Sud [10, 17, 24-26]. Il joue un rôle important dans la promotion de la transparence, de la responsabilité et de la fiabilité des informations relatives aux ressources minérales et aux réserves minérales dans le secteur minier sud-africain.

Le Code SAMREC [10, 17, 24, 26] encourage la séparation claire des ressources et des réserves, en mettant l'accent sur le niveau de confiance et de certitude associé à chaque catégorie. Il utilise un langage technique précis et fournit des définitions standardisées pour les termes couramment utilisés dans l'industrie. Cette normalisation facilite la compréhension et la comparaison des informations entre différentes entreprises minières. Ce code est essentiel pour garantir que les investisseurs, les régulateurs et d'autres parties prenantes disposent d'informations fiables et cohérentes pour prendre des décisions éclairées en matière d'investissement et de gestion des ressources minérales. En outre, il favorise la conformité aux normes internationales tout en tenant compte des particularités du secteur minier sud-africain.

2.4. Le Code de l'industrie minière Australien (JORC)

La norme JORC (Joint Ore Reserves Committee), est un ensemble de directives internationalement reconnues qui établit les normes pour la classification des ressources et des réserves minières [13, 17, 24]. Cette norme a été élaborée en Australie et est aujourd'hui largement utilisée dans le monde entier pour garantir des rapports précis et cohérents sur les ressources minérales et les réserves. La JORC joue un rôle essentiel dans l'industrie minière en fournissant un cadre transparent et fiable pour l'estimation des ressources et des réserves, ce qui est nécessaire pour les investisseurs, les entreprises minières et les organismes de réglementation.

Les principaux éléments de la norme JORC incluent des définitions claires pour les ressources et les réserves, des critères stricts pour l'estimation de la teneur, la taille, la forme et la continuité des gisements minéraux, et des directives sur la présentation des résultats dans les rapports publics [24]. La norme JORC met également l'accent sur la transparence, en encourageant la divulgation complète des informations géologiques et techniques, afin que les parties prenantes puissent évaluer la fiabilité des estimations.

Chaque norme de classification des ressources et réserves minérales a été élaborée pour répondre aux besoins spécifiques des régions géographiques où elles sont principalement utilisées. Bien qu'elles partagent des principes fondamentaux de rigueur et de transparence, des différences existent en ce qui concerne les catégories de ressources, les exigences de divulgation et les organismes de réglementation qui les supervisent. Les entreprises minières doivent se conformer aux normes applicables dans leur région respective pour garantir une divulgation précise et fiable des informations sur les ressources et réserves.

3. Norme internationale de classification des ressources/ réserves minières

La classification des ressources minérales vise principalement à offrir une évaluation globale de la confiance dans un projet minier, couvrant ainsi un large éventail d'acteurs, notamment les partenaires de l'industrie minière, les actionnaires et les institutions financières investissant dans le projet [18]. Bien que différents organismes gouvernementaux à travers le monde aient développé divers systèmes de classification des ressources et des réserves, la plupart partagent des objectifs similaires. Cette évaluation de la confiance joue un rôle important dans le développement des projets miniers, car il est impératif de disposer de données de ressources et de réserves suffisamment fiables pour les considérer comme des actifs.

Cependant, les lignes directrices de classification utilisent volontairement des termes génériques pour permettre leur application à une variété de types de gisements, de localisations géographiques et de méthodes d'exploitation. Elles n'imposent pas de méthodologie spécifique pour quantifier l'incertitude ou le risque, laissant plutôt une place prépondérante au jugement de l'estimateur des ressources, sous la supervision d'une personne compétente ou qualifiée. Cette flexibilité crée des défis pour établir une base de comparaison commune, car la signification des termes peut varier en fonction des circonstances et des personnes impliquées. Par conséquent, une approche émergente dans l'industrie consiste à décrire la confiance en utilisant des concepts statistiques traditionnels, en complément des critères traditionnels de classification.

Au fil des années, l'industrie minière a reconnu l'importance de ces évaluations et a établi des normes et des pratiques pour garantir des processus d'estimation et de classification robustes et reproductibles [14, 17, 24]. Cependant, avec la mondialisation croissante des entreprises minières et l'exploitation de gisements minéraux dans différents pays, il est devenu de plus en plus évident qu'il est nécessaire de développer des normes internationalement acceptées pour cette classification.

Bien que les normes régionales, telles que le CIM au Canada, le SEC Industry Guide 7 aux États-Unis et le SAMREC en Afrique du Sud, soient applicables dans leurs régions respectives, le Code JORC est souvent utilisé par les entreprises minières opérant dans divers pays [13, 14, 17, 18, 24]. Il repose sur des principes de rigueur, de transparence et de responsabilité dans la divulgation des informations sur les ressources et réserves, ce qui en fait la norme internationale la plus couramment acceptée dans l'industrie minière.

3.1. Classification des ressources et des réserves minières (Code JORC)

Le code JORC (Joint Ore Reserves Committee) propose une classification des ressources et réserves minérales qui repose sur des catégories clairement définies [14, 17], chacune reflétant un niveau de confiance et de certitude différent en ce qui concerne la présence et l'exploitation future de ressources minérales. Les réserves minérales sont classées en réserves probables et prouvées. La Figure 1.1 montre une classification générale des résultats d'exploration en fonction des niveaux de confiance dans les connaissances géologiques et des considérations techniques et économiques concernant le gisement, telle qu'établie par le Code Australasien pour la déclaration des ressources minérales identifiées et des réserves de minerai [13].

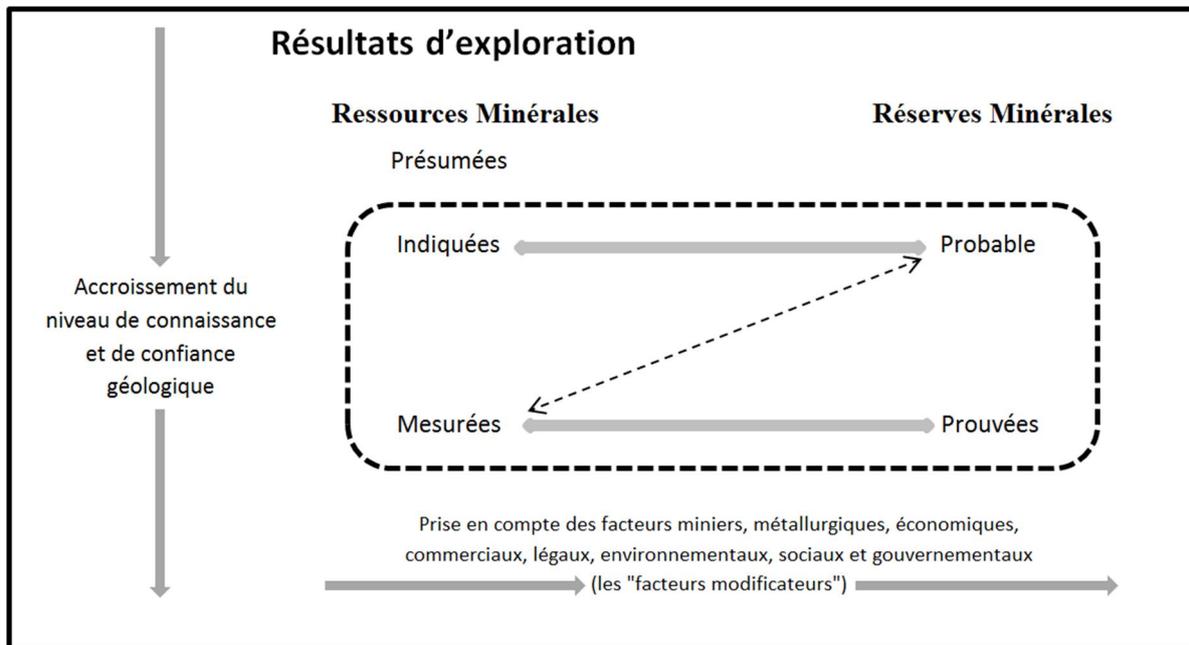


Figure 1.1. La relation entre les résultats d'exploration, ressources et réserves minières.

Modifié d'après [13].

Les principales catégories de classification selon le code JORC [13] sont les suivantes :

1.1.1. Ressources Minérales

En fonction du niveau de confiance et de la quantité d'information disponible, les ressources minérales sont subdivisées en différentes catégories :

- **Ressources Mesurées** : Les données sont de la plus haute qualité, avec une précision élevée. Elles sont basées sur des preuves solides et fiables.
- **Ressources Indiquées** : Les données sont d'une qualité légèrement inférieure à celles des ressources mesurées. Elles sont basées sur des preuves suffisantes pour soutenir une évaluation raisonnable.
- **Ressources Présumées** : Les données sont de moindre qualité et sont basées sur des preuves moins convaincantes. Il y a une incertitude plus grande par rapport aux ressources mesurées et indiquées.

1.1.2. Réserves Minérales

Contrairement aux ressources minérales, les réserves sont généralement mieux définies et confirmées par des travaux d'exploration approfondis. Ces réserves minérales représentent la quantité de minéraux économiquement viable à extraire d'un gisement existant. Elles sont

généralement prêtes pour l'extraction, et les entreprises minières planifient leur exploitation en fonction de ces réserves.

En fonction du niveau de confiance et de la quantité d'informations disponibles, les réserves minérales sont subdivisées en différentes catégories :

- **Réserves Prouvées et Probable** : Il s'agit de minerais dont l'extraction est techniquement réalisable et économiquement viable. Les réserves prouvées sont basées sur des preuves solides, tandis que les réserves probables sont basées sur des preuves moins solides.
- **Réserves Possibles** : Les réserves possibles sont celles qui pourraient être récupérées si certaines conditions étaient remplies. L'incertitude est plus grande par rapport aux réserves prouvées et probables.

Le Code JORC exige que les entreprises minières divulguent ces informations de manière transparente et qu'elles fournissent des détails sur les facteurs qui affectent l'estimation des ressources et réserves, ainsi que sur les paramètres économiques et techniques qui sous-tendent l'évaluation [13, 18]. Ces classifications fournissent une base solide pour évaluer les ressources disponibles et planifier les activités minières. Elles permettent également de prendre en compte les incertitudes inhérentes à l'extraction de minéraux.

4. Méthodes d'évaluation des réserves

Il y a deux concepts sous-tendant l'estimation des réserves : le concept d'extension, où les attributs d'un échantillon sont étendus aux blocs à estimer ; et le concept d'estimation de l'erreur, où la validité d'une méthode d'estimation est évaluée en fonction de l'erreur impliquée [21]. Les méthodes utilisées pour effectuer l'estimation sont importantes, car elles peuvent influencer la fiabilité et la précision de l'estimation. Plusieurs méthodes d'estimation ont été proposées et mises en œuvre dans la littérature. Ces méthodes sont largement classées en méthodes d'estimation géométriques et géostatistiques [21, 27], et elles sont appelées aussi techniques conventionnelles. Les techniques géométriques (par exemple, méthodes polygonales, triangulaires, et en coupes transversales) sont simples et nécessitent peu de paramètres d'entrée, et elles sont souvent appliquées aux premières étapes d'un projet minier ou pour vérifier les résultats des méthodes d'estimation sophistiquées [21].

La méthodologie d'estimation des ressources représente un processus intégral au sein de l'industrie minière, qui englobe la prédiction et la modélisation des caractéristiques physiques

d'un gisement minéral en collectant, analysant et interprétant une multitude de données provenant de diverses sources. Ces caractéristiques incluent la taille du gisement, sa forme géométrique, la continuité des zones minérales, la répartition de la teneur en minéraux utiles, ainsi que la variabilité spatiale de ces teneurs au sein du gisement. Ces informations sont prédites en se basant sur des données d'échantillonnage récoltées à travers des techniques telles que le forage, le décapage, les essais de puits et l'échantillonnage en continu. Il est essentiel de noter que, bien que ces données permettent d'établir des modèles prédictifs, les caractéristiques réelles du gisement restent partiellement connues.

4.1. Méthode des coupes transversales (sections géologiques)

La méthode des sections géologiques (ou coupes transversales) [9, 28], a longtemps été un outil essentiel dans l'estimation des réserves minérales, offrant une représentation 3D des minéraux dans une zone donnée. Elle consiste à créer des sections géologiques pour représenter la distribution des minéraux ou des minerais dans une zone donnée. Pour ce faire, des données géologiques, telles que la lithologie, la teneur en minéraux, les limites géologiques, sont collectées et utilisées pour créer des sections à travers la zone d'intérêt. Les données nécessaires proviennent souvent de sondages carottés qui fournissent des informations détaillées sur le sous-sol.

Cette méthode (Figure 1.2.) est appliquée avec le plus de succès dans le cas d'un gisement qui présente des contacts nets et relativement réguliers, comme c'est le cas pour de nombreux gisements tabulaires (veineux et en couches). Les informations sur les teneurs (par exemple, à partir de forages) sont couramment concentrées le long de sections transversales équidistantes pour produire un ensemble de données relativement systématique. Dans certaines situations souterraines, des ensembles de données plus irréguliers peuvent en résulter, par exemple, à partir de réseaux de forages.

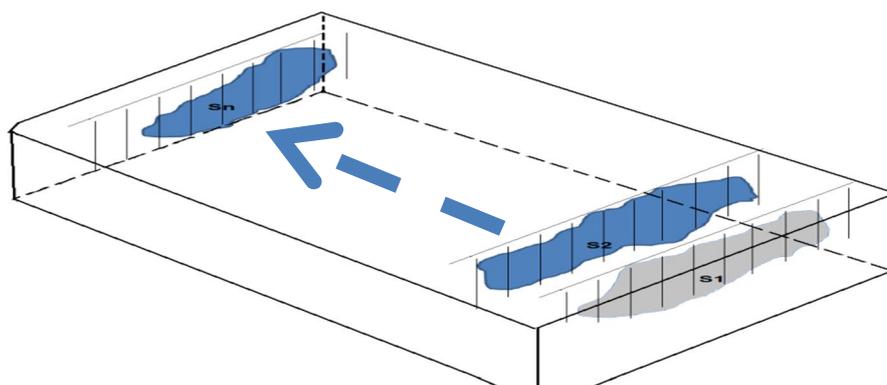


Figure 1.2. Estimation des ressources/réserves par extrapolation des sections transversales.

Une interpolation géologique des teneurs est réalisée pour chaque section transversale (ou plan), et l'interprétation sur chaque section est projetée de manière lisse sur les sections voisines. Il convient de noter que cette interpolation en deux étapes (sur et entre les sections) représente les principales hypothèses sous-jacentes à la méthode des sections transversales. Chaque section géologique représente une tranche de la zone étudiée, et ces sections sont utilisées pour estimer les volumes de minéraux ou de minerai dans chaque tranche (figure 1.2). Une teneur moyenne pondérée sur une section est généralement projetée sur un volume s'étendant à mi-chemin des sections voisines [15]. En multipliant les volumes par la teneur moyenne, il est possible d'estimer les réserves minérales dans la zone d'intérêt. Cependant, elle présente certaines limites en termes de précision, en particulier lorsque la géologie de la zone est complexe.

La Figure 1.3 démontre clairement comment l'irrégularité du contact minerai-stérile peut influencer l'estimation des réserves minières [29]. En règle générale, la teneur est surestimée car une quantité inconnue de stérile est incluse dans le minerai interprété. De plus, une quantité inconnue de minerai est perdue dans le stérile (c'est-à-dire que certaines matières à teneur de minerai se trouvent en dehors des limites d'exploitation conçues).

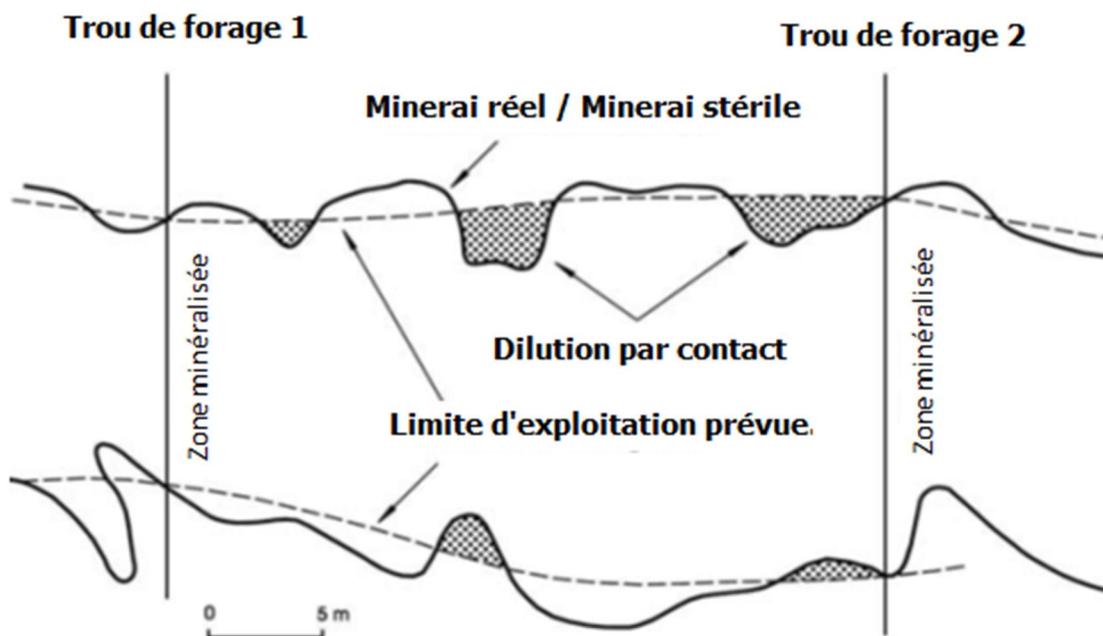


Figure 1.3. Représentations d'un contact irrégulier entre le minerai et les parties stérile modifié d'après Stone et Dunn (1996).

Les lignes pointillées lisses interprétées à l'aide d'informations provenant de forages (sur des sections), mais qui sont interpolées entre les forages. Avec de tels contacts irréguliers, la

méthode des sections doit nécessairement être incorrecte (c'est-à-dire qu'elle ne peut donner par hasard la bonne estimation de teneur pour la zone interpolée car des déchets sont inclus dans la zone interprétée de minéralisation).

L'imprécision de la méthode des sections transversales réside principalement dans le fait qu'elle ne tient pas compte des variations potentielles à l'intérieur des sections, se basant sur des données discrètes collectées aux points de sondage [9]. Cependant, cette dernière reste une approche précieuse pour les zones où les technologies plus avancées ne sont pas disponibles ou lorsque les sondages sont limités [9]. Il est essentiel de choisir la méthode d'estimation la plus adaptée en fonction des caractéristiques géologiques spécifiques de chaque site minier pour garantir des estimations de réserves précises et fiables.

4.2. Méthode des polygones

Les méthodes polygonales [29, 30], englobent diverses approches visant à estimer les volumes de polygones individuels définis de manière géométrique à partir de données limitées. Dans le cas des méthodes polygonales où les données brutes ne sont pas lissées, une seule valeur de teneur ou d'épaisseur, provenant d'un échantillon de sondage, est utilisée comme la moyenne pour un grand bloc (prisme polygonal) de minerai ou de stériles (Figure. 1.4a).

Sur une projection bidimensionnelle d'un gisement ou d'un banc montrant les emplacements des échantillons, les polygones sont définis de plusieurs manières, l'une des plus courantes étant par une série de bissectrices perpendiculaires des lignes reliant les emplacements des échantillons (voir Figure 1.4). Chaque polygone contient un emplacement d'échantillonnage unique et tous les autres points à l'intérieur du polygone sont plus proches de cet échantillon central. Des décisions arbitraires doivent être prises quant à la façon dont les prismes marginaux sont délimités sur leur bord extérieur.

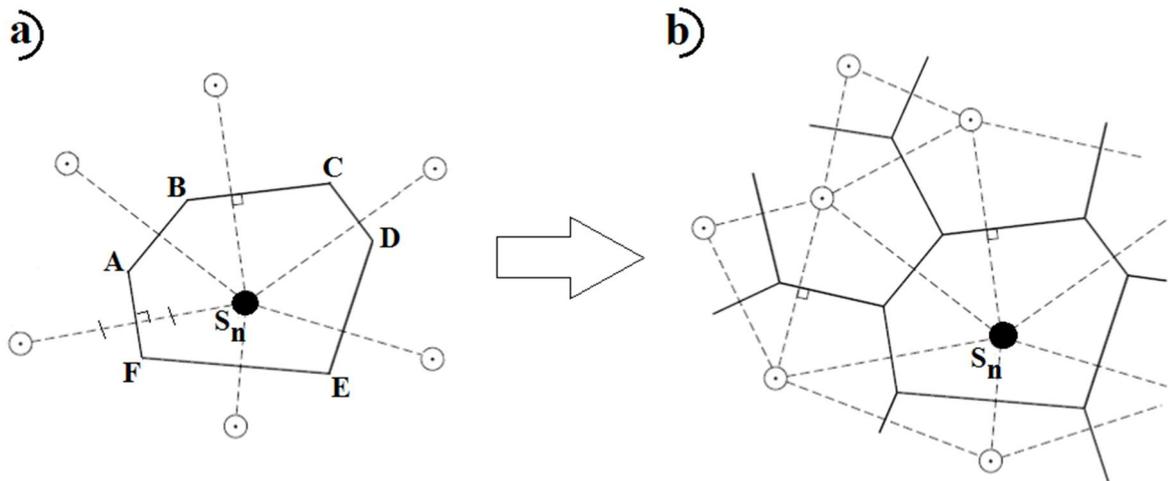


Figure 1.4. Construction d'une zone des polygones à laquelle une seule teneur contenue est appliquée (la teneur de l'échantillon contenu est étendue au polygone) Modifié d'après [30].

Les cercles représentent les points de données (sondage avec des teneurs connues) ; les lignes en pointillés relient les points de données adjacents et forment des triangles de Delaunay ; les lignes épaisses définissant un polygone, sont perpendiculaires aux lignes en pointillés et divisent les lignes en pointillés en deux segments égaux. (a) où la teneur et l'épaisseur sont extrapolées vers le polygone ABCDEF ; (b) construction d'une série de polygones représentant une zone minéralisée.

Lorsque les trous de forage sont répartis de manière aléatoire, c'est-à-dire qu'ils ne suivent pas une grille régulière, la teneur et l'épaisseur de chaque trou peuvent être attribuées à un polygone irrégulier. On suppose que ces valeurs restent constantes dans toute la zone du polygone, qui peut être défini selon deux méthodes : les bissectrices perpendiculaires et les bissectrices angulaires. Des mosaïques polygonales sont établies en construisant des bissectrices perpendiculaires aux lignes de liaison reliant les trous de forage adjacents au trou en question (figure 1.4 b).

La troisième dimension, ou "hauteur" du prisme polygonal, correspond à l'épaisseur du gisement ou du banc perpendiculaire au plan de projection. Cela crée un modèle de prismes polygonaux auxquels sont attribuées les données (teneurs) contenue. La méthode est simple, rapide et permet une désagrégation automatique des données. D'autres méthodes de détermination des prismes polygonaux individuels peuvent aboutir à ce que les données ne soient pas situées de manière optimale au sein des prismes. Par exemple, si une grille carrée est superposée sur un champ de données bidimensionnel pour un banc donné, chaque "carré"

pourrait être estimé selon la teneur de carottage la plus proche ; cette variation de la méthode polygonale est appelée procédure d'estimation du plus proche voisin.

La régularisation des teneurs (diminution de la variabilité) augmente avec le support (volume) (voir Figure 1.3). Une conséquence de cela est qu'en moyenne, l'utilisation des teneurs brutes des échantillons pour calculer les teneurs moyennes de grands volumes surestime la teneur des blocs à haute teneur (voir Figure 1.3) et, en conséquence, sous-estime la teneur des blocs à faible teneur (par exemple, un biais conditionnel, dans lequel le biais dépend de la teneur estimée). Royle (1979) discute des facteurs qui affectent l'ampleur de ce biais dans les estimations polygonales, notamment la variance des données, l'effet pépité et la portée d'influence des échantillons [31]. Ce biais est un problème particulièrement grave dans les cas où les seuils de coupure sont supérieurs à la valeur moyenne de la distribution des teneurs, car il peut entraîner des biais importants en faveur de la surestimation à la fois de la tonne de "haute" teneur et, en particulier, de la teneur moyenne au-dessus du seuil de coupure, comme illustré dans la courbe teneur-tonnage de la figure 1.5.

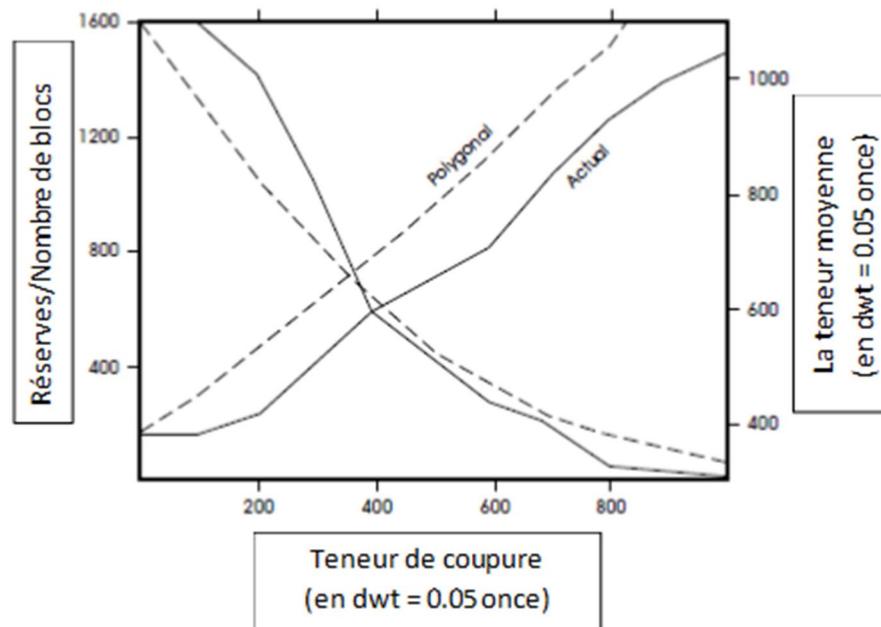


Figure 1.5. Ampleur des erreurs possibles lors de l'application d'une méthode des polygones à l'estimation des ressources/réserves (d'après Royle, 1979).

Les estimations polygonales (lignes en pointillés) sont présentées sous forme de "teneur tonnage" (le nombre de blocs remplace le tonnage) pour comparaison avec les données réelles (lignes pleines). Les courbes de tonnage s'inclinent vers le bas à droite ; les courbes de teneur moyenne s'inclinent vers le bas à gauche.

Un autre problème des méthodes des polygones est que les anisotropies spatiales de la distribution des teneurs ne sont généralement pas prises en compte [15]. De plus, les marges des gisements sont généralement connues grâce à des informations d'analyse relativement espacées, de sorte que des analyses individuelles peuvent être attribuées à des volumes extrêmement importants de roche minéralisée. De toute évidence, un tel ensemble de blocs peut être loin d'être régulier.

4.3. Méthode des triangles

Une variante triangulaire de l'approche polygonale (comme illustré dans la Figure 1.4b) présente une approche plus prudente que l'assignation de valeurs uniques à de vastes blocs. Les prismes triangulaires sont définis sur une projection bidimensionnelle, telle qu'un plan de banc, en reliant trois emplacements d'échantillonnage de manière à créer un triangle qui n'inclut aucun emplacement d'échantillonnage interne. La construction de ces triangles peut impliquer l'utilisation de la "triangulation de Delaunay"[31].

La méthode consiste à attribuer la moyenne des trois valeurs aux sommets du triangle au prisme triangulaire (bloc) en question. L'avantage principal de cette approche réside dans l'introduction d'un certain degré de lissage dans les estimations des prismes individuels. En conséquence, l'estimation de la queue de la distribution de densité de grade est plus prudente que celle obtenue avec l'approche polygonale traditionnelle. Cependant, la méthode triangulaire n'est pas exempte de défis. Premièrement, le lissage est entièrement basé sur des facteurs empiriques, ce qui signifie qu'il ne peut pas refléter la réalité de manière optimale. Deuxièmement, la pondération égale de trois échantillons est un processus arbitraire et ne garantit pas une optimisation du calcul. Troisièmement, les anisotropies, c'est-à-dire les variations directionnelles dans la distribution des teneurs, ne sont pas systématiquement prises en compte. Enfin, les unités estimées ne peuvent pas former forcément une disposition régulière de blocs, ce qui peut affecter la précision de l'estimation.

4.4. Méthode de l'inverse de la distance

Les méthodes de pondération par l'inverse de la distance représentent une approche précieuse pour estimer les propriétés des gisements minéraux, en particulier lorsqu'un réseau régulier de blocs tridimensionnels est utilisé pour modéliser la géométrie d'un gisement [15]. Cette méthode est couramment utilisée dans l'industrie minière pour estimer la teneur en minerai utile, l'épaisseur des couches minéralisées, et d'autres caractéristiques clés des gisements.

L'idée fondamentale de la méthode de l'inverse de la distance, repose sur le concept que les valeurs d'un bloc donné sont influencées de manière plus significative par les données géologiques les plus proches, en particulier en termes de distance géographique [31, 32].

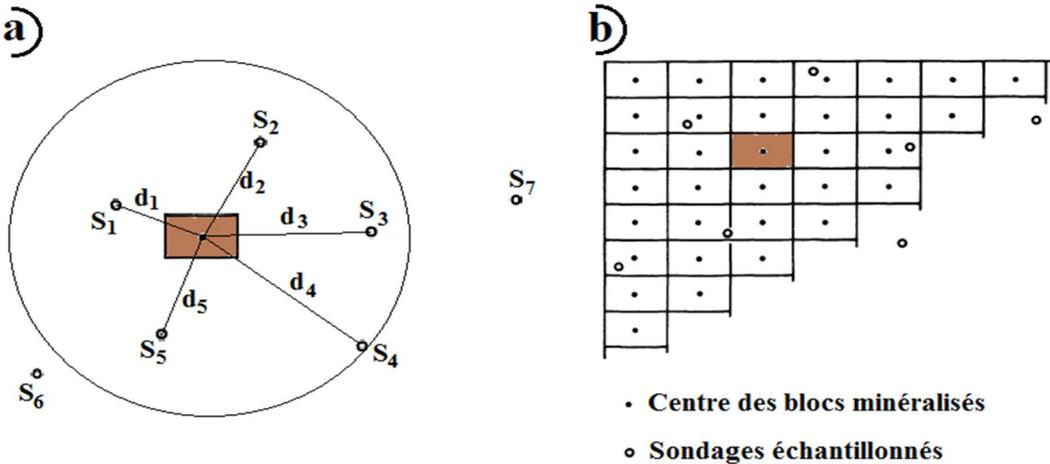


Figure 1.6. Illustration de l'estimation par la méthode de l'inverse de la distance.

a) la fenêtre mobile (circulaire) et son centre, les points de mesure (S_i) et les distances par rapport au centre (d_i), certains se trouvent à l'intérieur de la fenêtre, d'autres à l'extérieur ; b) Illustration de l'estimation de blocs, les point noires représente les centres des blocs à estimé, les petite cercle représente les emplacements des sondages.

Cela signifie que les données géologiques situées à proximité du bloc à estimer sont prises en compte de manière prépondérante, tandis que les données plus éloignées ont un impact moindre. La quantité de données prise en compte pour chaque estimation individuelle dépend de critères de distance et de la précision souhaitée (figure I.5). Ainsi, un poids est attribué à chaque donnée en fonction de sa proximité au point d'intérêt, et ces poids sont calculés en utilisant une fonction de l'inverse de la distance. Plus précisément, la formule de pondération attribue des poids plus élevés aux données proches et des poids plus faibles aux données plus éloignées. Les poids sont définis comme suit :

$$W_i = \left(\frac{1}{d_i^n} \right) / \left[\sum \left(\frac{1}{d_i^n} \right) \right] \quad \dots \dots \dots (1.1)$$

- Où :
- W_i** : les poids attribué au données de sondage selon leurs emplacement ;
- d_i** : les distance qui sépare le point à estimé des emplacement d'échantillonnage dans la zone de recherche ;
- n** : varie généralement entre 1 et 3.

Cette technique applique un facteur de pondération, basé sur une fonction de distance linéaire ou exponentielle, à chaque échantillon entourant le point central d'un bloc de minerai. Le facteur de pondération est l'inverse de la distance entre chaque échantillon et le centre du bloc, élevé à la puissance 'n'. Où 'n' varie généralement entre 1 et 3. Seuls les échantillons se trouvant dans une zone de recherche ou un volume spécifié sont pondérés de cette manière. Comme la méthode est laborieuse et répétitive, elle est généralement informatisée.

5. Conclusion

La classification des ressources/réserves minières est une étape cruciale dans l'industrie minière, car elle permet de fournir une évaluation fiable des actifs miniers, ce qui revêt une importance capitale pour les investisseurs, les partenaires miniers et les planificateurs à long terme. Plusieurs systèmes de classification ont été employés dans différents pays, sont souvent caractérisés par une terminologie volontairement vague, afin de les rendre applicables à une grande variété de types de gisements, de situations géographiques et de méthodes d'exploitation. Cette flexibilité a ses avantages, mais elle peut aussi prêter à confusion. La description de la confiance associée aux estimations est devenue de plus en plus importante, avec une tendance à utiliser une description statistique de l'incertitude en complément des critères de classification traditionnels.

L'évaluation des réserves est d'une importance capitale pour l'industrie minière. Elle constitue le fondement sur lequel reposent les décisions d'exploitation et les investissements à long terme. Cependant, cette évaluation est loin d'être une tâche simple, elle est associée des incertitudes qui peuvent influencer considérablement les résultats.

Les méthodes traditionnelles, telles que les coupes transversales, les polygones, les triangles et l'inverse de la distance, ont leurs limites en ce qui concerne la communication du degré de confiance associé aux estimations de quantités et de teneurs, comme souligné par plusieurs chercheurs [2, 15, 29, 30, 33]. En revanche, on sait que le krigeage, notamment le krigeage ordinaire (KO) fournit la meilleure estimation et sa variance associée. Il est largement utilisé pour l'estimation des ressources/réserves minières, principalement en raison de sa robustesse et de sa facilité dans l'évaluation de l'incertitude en utilisant la variance de krigeage. On sait aussi que la variance du KO est incapable de reconnaître la variabilité des données locales, un problème important lors de l'évaluation de gisements minéraux hétérogènes comportant des zones à teneur élevée et à teneur plus faible. Dans le deuxième chapitre, nous explorons des méthodes géostatistiques en combinaison avec d'autres approches avancées dans le but d'améliorer la fiabilité des estimations et d'intégrer de manière plus précise l'incertitude dans les évaluations des ressources et des réserves minières.

Chapitre II : Méthodes géostatistiques et techniques de validation

1. Introduction

Lors de l'estimation des réserves minières, deux approches peuvent être choisies : une approche des méthodes traditionnelles déterministes, telles que l'inverse de la distance, la méthode triangulaire et celle du polygone [9, 15]. Cependant, cette approche souffre de limites majeures en termes de précision. Elle néglige souvent la structure et les dimensions géographiques des données en les traitant de manière indépendante, ce qui la rend sensible aux variations locales et aux brusques changements dans la structure spatiale. En conséquence, ces méthodes déterministes génèrent fréquemment des estimations imprécises, conduisant à des sous-estimations ou des surestimations des valeurs aux endroits peu échantillonnés, avec des résultats parfois incohérents.

C'est dans ce contexte que la géostatistique [2, 15], notamment le krigeage ordinaire, se distingue. En exploitant à la fois les aspects statistiques et spatiaux des données, le krigeage ordinaire offre une solution probabiliste puissante qui prend en considération la structure spatiale des données. Il en résulte une nette amélioration de la précision des estimations géospatiales. Ce chapitre se penchera en détail sur la manière dont le krigeage ordinaire s'affirme comme une méthode précise et fiable pour la modélisation spatiale.

Après avoir examiné les avantages des approches géostatistiques en termes de précision, nous explorerons également la question de l'ajustement et de la sélection des modèles variographiques les plus performants. Nous mettrons en évidence l'utilisation des techniques de validation croisée telles que K-Fold, LOOCV, LPOCV, et d'autres [34]. Ces méthodes permettent d'évaluer la capacité des modèles à généraliser sur de nouvelles données, garantissant ainsi que les méthodes géostatistiques restent robustes, précises et capables de fournir des estimations fiables.

2. Historique

L'interpolation basée sur la dépendance spatiale des échantillons a été utilisée pour la première fois par D. G. Krige [35, 36] pour l'estimation de la teneur en or des gisements miniers en Afrique du Sud [37]. Cette innovation a vu le jour en réponse aux limites des méthodes d'interpolation statistique traditionnelles, qui étaient largement employées à l'époque. Ces méthodes classiques étaient souvent sujettes à des biais et ne tenaient pas compte de la complexité de la dépendance spatiale des données lors de l'estimation des gisements minéraux [1, 21, 38]. Le besoin de développer des procédures plus robustes et optimales pour estimer la teneur en minerai des gisements minéraux a conduit au

développement des nouvelles approches d'interpolation. Ces approches prenaient en compte les variations locales du contenu et de la teneur en minerai, ainsi que la dépendance spatiale des échantillons, dans le but d'obtenir des estimations plus précises. L'objectif était également de minimiser les coûts élevés associés à l'échantillonnage des gisements minéraux, en optimisant le placement des échantillons.

Les méthodes pratiques de D. G. Krige ont ouvert la voie à une avancée significative dans le domaine de l'estimation des variables spatialement dépendantes [2, 15, 39, 40]. Ces pratiques ont été généralisées et étendues par Matheron (1963, 1965, 1969, 1970, 1971) dans la théorie des variables régionalisées. Cette théorie repose sur le principe fondamental de prendre en compte à la fois les caractéristiques aléatoires et structurées des variables spatiales, ce qui a ouvert la voie à des outils quantitatifs permettant de les décrire et de les estimer de manière optimale et impartiale.

L'ensemble de ces procédures a été regroupé sous le terme de « géostatistique », une discipline qui a révolutionné la manière dont nous abordons l'analyse et l'estimation des variables spatialement dépendantes. [40-42] ont été parmi les premiers à appliquer le krigeage en tant que procédure d'estimation dans le génie minier.

La théorie géostatistique a continué à se développer dans les années 1970 pour inclure l'analyse de variables ayant des distributions de probabilité log-normales [2, 40, 42] ou complexes ainsi que l'estimation en présence de tendances [3, 39]. Alors que l'utilisation de la géostatistique s'était principalement concentrée sur l'industrie minière, elle est désormais largement utilisée dans la recherche sur les ressources en eau, la science du sol, et l'archéologie.

3. Théorie des variables régionalisées

La géostatistique repose sur les concepts de variables régionalisées, de fonctions aléatoires et de stationnarité. Une brève discussion théorique de ces concepts est nécessaire pour comprendre l'application pratique de la géostatistique à l'analyse de la variation des sols. Une couverture complète de la théorie des variables régionalisées et de ses applications géostatistiques est proposée par [31, 40, 42, 43].

3.1. Variables régionalisées et fonctions aléatoires

Une variable aléatoire est une mesure de données individuelles qui est susceptible de varier selon une loi de distribution de probabilité [2]. La variable aléatoire est caractérisée par

les paramètres de la distribution, tels que la moyenne et la variance de la distribution normale. Une variable régionalisée $z(x)$ est une variable aléatoire qui prend différentes valeurs z en fonction de son emplacement x à l'intérieur d'une région donnée [40]. En tant que telle, une variable régionalisée $z(x)$ peut être considérée comme une réalisation particulière d'une variable aléatoire Z pour un emplacement fixe x à l'intérieur de la région. Si toutes les valeurs de $z(x)$ sont prises en compte à tous les emplacements à l'intérieur de la région, alors la variable régionalisée $z(x)$ devient un membre d'un ensemble infini de variables aléatoires $Z(x)$ pour tous les emplacements à l'intérieur de la région. Un tel ensemble est appelé une fonction aléatoire car il associe une variable aléatoire Z à n'importe quel emplacement x [41].

3.1.1. Stationnarité

Considérons la fonction aléatoire $Z(x)$ dans un domaine D . Pour tout ensemble de k points x_1, x_2, \dots, x_k dans D , nous obtenons une variable vectorielle à k composantes $\{z(x_1), z(x_2), \dots, z(x_k)\}$, caractérisée par une distribution de probabilité à k variables :

$$F_{x_1, x_2, \dots, x_k}(Z_1, Z_2, \dots, Z_k) = \text{prob}\{Z(x_1) < Z_1, Z(x_2) < Z_2, \dots, Z(x_k) < Z_k\} \dots \dots (2.1)$$

La loi spatiale de la fonction aléatoire $Z(x)$ est définie comme l'ensemble de ces distributions de probabilité pour tout entier positif k et pour tous les choix possibles de k points.

En pratique, il est souvent suffisant de connaître les deux premiers moments de cette loi, car ils permettent de fournir une solution approximative acceptable à la plupart des problèmes courants.

- Moment d'ordre 1 :

Une fonction aléatoire $Z(x)$ est dite stationnaire du premier ordre si sa valeur attendue est la même à tous les emplacements dans toute la région d'étude.

$$E[Z(x)] = m \dots \dots \dots (2.2)$$

Où m est la moyenne des statistiques classiques, et

$$E[Z(x) - Z(x + h)] = 0 \dots \dots \dots (2.3)$$

Où h est le vecteur qui sépare entre les emplacements des échantillons.

- Moment d'ordre 2

La stationnarité du second ordre s'applique si la covariance spatiale $C(h)$ de chaque paire $Z(x)$ et $Z(x + h)$ est la même (indépendante de la position) dans toute la région d'étude et dépend de h :

$$C(h) = E[Z(x) - m][Z(x + h) - m] \dots \dots \dots (2.4)$$

À mesure que h devient plus grand, $C(h)$ diminue et la covariance spatiale décroît (Figure 2.1).

La stationnarité de $C(h)$ implique la stationnarité de la variance échantillonnée σ^2 . La covariance spatiale se rapproche de la variance échantillonnée à mesure que la distance de séparation tend vers zéro. La stationnarité du second ordre ne s'applique pas si une variance et une covariance finies ne peuvent pas être définies, comme dans le cas des phénomènes de tendance [42], et une forme plus faible de stationnarité appelée hypothèse intrinsèque doit être supposée [40].

Lorsque deux variables aléatoires $Z(x_1)$ et $Z(x_2)$ présentent des variances aux points x_1 et x_2 , leur covariance est inévitablement une fonction dépendant uniquement du vecteur h qui représente la séparation entre les points x_1 et x_2 .

La stationnarité du second ordre implique l'hypothèse intrinsèque, mais pas l'inverse. L'hypothèse intrinsèque exige que, pour tous les vecteurs de h , la variance de l'incrément $Z(x) - Z(x + h)$ soit finie et indépendante de la position à l'intérieur de la région, c'est-à-dire :

$$\begin{aligned} Var[Z(x) - Z(x + h)] &= E[Z(x) - Z(x + h)]^2 \\ &= 2\gamma(h) \dots \dots \dots (2.5) \end{aligned}$$

L'hypothèse de stationnarité intervient alors pour supposer que dans les gisements homogènes, la corrélation qui existe entre les valeurs $Z(x)$ et $Z(x+h)$ ne dépend pas des points (x) et $(x+h)$ mais du vecteur h qui les sépare.

$$C\{x_1, x_2\} = C(h) = E\{Z(x_1), Z(x_1 + h)\} - m \cdot m \dots \dots \dots (2.6)$$

La division par deux donne la semi-variance $\gamma(h)$. La semi-variance γ dépend du vecteur de séparation h . Idéalement, γ est nul lorsque $h = 0$, et augmente à mesure que h augmente (Figure2. 1).

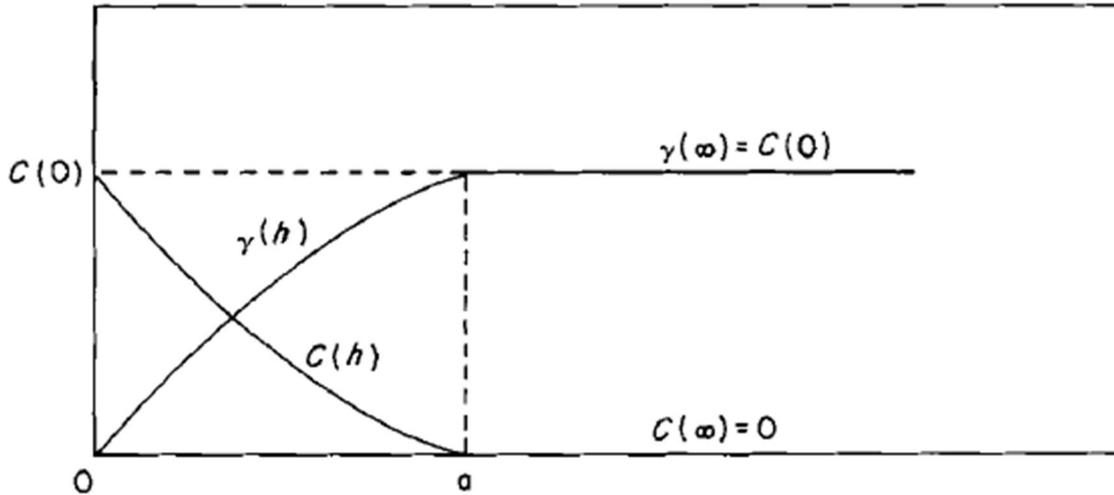


Figure 2.1. Relation entre la covariance spatiale $C(h)$ et le semi-variogramme $\gamma(h)$. [40]

3.2. Analyse de la dépendance spatiale

Les concepts de variables régionalisées et de stationnarité fournissent la base théorique pour l'analyse de la dépendance spatiale en utilisant l'autocorrélation ou les semi-variogrammes.

3.2.1. Autocorrélation

Les fonctions d'autocorrélation expriment la corrélation linéaire entre une série spatiale et la même série à un intervalle de distance supplémentaire [3]. Leur définition suppose une stationnarité d'ordre deux, auquel cas l'autocorrélation est exprimée comme suit :

$$\rho(h) = \frac{C(h)}{C(0)} \dots \dots \dots (2.7)$$

Où $\rho(h)$ représente coefficient de corrélation entre échantillons $Z(x)$ et $Z(x+h)$ à une distance de séparation h .

Un graphique des valeurs $\rho(h)$ en fonction du distance h , est appelé autocorrélation. La valeur maximale de $\rho(h)$ égale à 1 à une distance nulle ($h = 0$), et les valeurs diminuent avec l'augmentation de h . La distance (a) à laquelle $\rho(h)$ se stabilise définit la plage sur laquelle les échantillons de la variable sont spatialement dépendants.

La covariance et le corrélogramme montrent comment cette corrélation évolue avec la séparation, ou le décalage, h . Notez que h est un vecteur. Ces fonctions dépendent donc à la fois de sa longueur (la distance entre x et $x+h$), et de sa direction. Lorsque la covariance

dépend uniquement de la distance, on dit qu'elle est isotrope [3]. Une covariance est une fonction paire, et par l'inégalité de Schwarz, elle est bornée par sa valeur à l'origine.

$$C(\mathbf{h}) = C(-\mathbf{h})|C(\mathbf{h})| \leq C(\mathbf{0}) \dots \dots \dots \dots \dots \dots \dots \dots \dots (2.8)$$

Plus précisément, nous montrerons qu'une covariance est une fonction définie positive. Comme la fonction aléatoire a une variance finie, elle fluctue autour de la moyenne. Certains phénomènes n'ont pas ce comportement : si nous calculons la moyenne et la variance de l'échantillon sur des domaines de plus en plus grands, la moyenne de l'échantillon ne se stabilise pas, et la variance de l'échantillon augmente toujours.

En réalité, les hypothèses précédemment énoncées ne sont souvent pas satisfaites. Lorsqu'il existe une tendance marquée, on ne peut pas supposer que la moyenne reste constante [44]. Nous nous limitons à des situations où la moyenne est constante. Même dans ces cas, la covariance peut ne pas être présente. Par conséquent, il est nécessaire de prendre en compte ces considérations et d'assouplir davantage ces hypothèses.

3.2.2. Hypothèse intrinsèque

Une fonction aléatoire est dite intrinsèque si (i) l'espérance mathématique existe et ne dépend pas du point de support x , c'est-à-dire $E\{Z(x)\} = m$, et (ii) pour tout vecteur h , l'incrément $[Z(x + h) - Z(x)]$ a une variance finie qui est indépendante du point x . En d'autres termes, $E[Z(x + h) - Z(x)] = 0$ et $Var[Z(x + h) - Z(x)] = \gamma(h)$, une valeur finie qui ne dépend pas de x . La fonction $\gamma(h)$ est appelée le semi-variogramme. En abrégé, on l'appelle simplement variogramme. Il s'agit de l'outil de base pour l'interprétation structurale des phénomènes ainsi que pour l'estimation.

L'hypothèse intrinsèque est donc plus générale. Elle est beaucoup plus faible que l'hypothèse de stationnarité d'ordre 2.

4. Le semi-variogramme (Variogramme)

L'analyse structurale de la dépendance spatiale à l'aide des variogrammes peut être effectuée en utilisant des hypothèses de stationnarité moins contraignantes que celles nécessaires pour l'autocorrélation [3, 37, 44]. L'analyse des variogrammes présente l'avantage supplémentaire de définir les paramètres nécessaires pour l'estimation locale par krigeage. La statistique de variogramme $\gamma(h)$ peut être définie en termes de la variance $C(0) = \sigma^2$ et de la covariance spatiale $C(h)$ de $Z(x)$ si la stationnarité d'ordre deux s'applique, c'est-à-dire

$$\gamma(\mathbf{h}) = \sigma^2 - C(\mathbf{h}) \dots \dots \dots (2.9)$$

Les corrélations spatiales d'une fonction aléatoire intrinsèque sont caractérisées par une fonction de variogramme définie comme suit [41] :

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{var}[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})] \dots \dots \dots (2.10)$$

Étant donné qu'il a été supposé que la moyenne de $Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})$ est nulle, $\gamma(\mathbf{h})$ représente simplement la moitié de la valeur quadratique moyenne de la différence. C'est-à-dire :

$$\gamma(\mathbf{h}) = \frac{1}{2} E[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})]^2 \dots \dots \dots (2.11)$$

En notation discrète, le variogramme peut être exprimé comme :

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^n [Z(x_i) - Z(x_i + \mathbf{h})]^2 \dots \dots \dots (2.12)$$

Où : $N(\mathbf{h})$ est le nombre de paires d'observations séparées par la distance \mathbf{h} .

Ici, \mathbf{x} et $\mathbf{x} + \mathbf{h}$ font référence à des points dans l'espace tridimensionnel. Pour un angle fixe, le variogramme donne différentes valeurs à mesure que la distance augmente. Lorsque l'angle est modifié, le variogramme révèle les caractéristiques directionnelles, le cas échéant, du phénomène, telles que son anisotropie.

4.1. Propriétés du variogramme

Le taux d'augmentation du variogramme reflète le degré de dissimilarité entre les échantillons de plus en plus éloignés. Le variogramme peut augmenter indéfiniment si la variabilité du phénomène n'a pas de limite à de grandes distances [3] . En revanche, si le variogramme se stabilise à une valeur, appelée seuil (voir la figure 2.2), cela signifie qu'il existe une distance au-delà de laquelle $Z(\mathbf{x})$ et $Z(\mathbf{x}+\mathbf{h})$ ne sont pas corrélés. Cette distance est appelée « portée » (Figure 2.2). Elle donne un sens précis à la notion conventionnelle de zone d'influence d'un échantillon. Le variogramme peut révéler des structures emboîtées, c'est-à-dire des structures hiérarchiques, chacune caractérisée par sa propre portée (voir la figure 2.2.B).

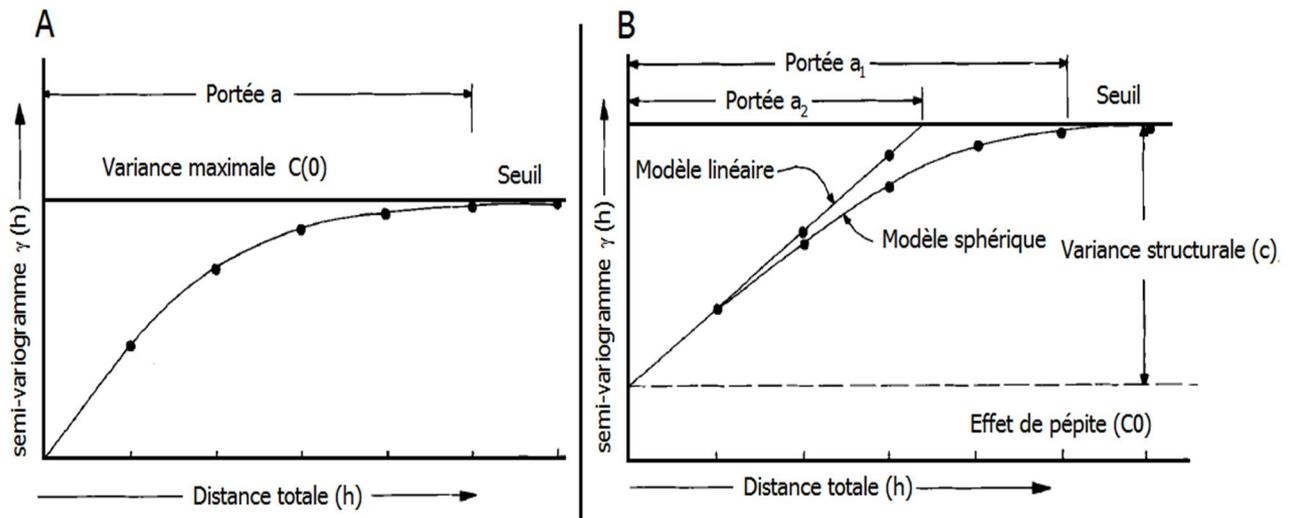


Figure 2.2. (A) Variogramme idéalisé avec une variance de pépité nulle; (B) variogrammes observés pour les propriétés du sol avec un effet de pépité. Modifié d'après [45].

4.1.1. Discontinuité à l'origine « Effet de pépité »

Lorsque h tend vers zéro, $\gamma(h)$ ne tend pas vers zéro. Cela signifie que la variable n'est même pas continue en moyenne carrée. Elle est donc très irrégulière à de courtes distances. La plupart des variables géologiques, y compris les teneurs en métaux tels que l'or, le cuivre, le plomb et le zinc, présentent ce type de comportement [40, 43, 44, 46]. Ce saut à l'origine est appelé effet pépité car il a été observé pour la première fois dans les gisements d'or en Afrique du Sud, où il était associé à la présence de pépites dans le minerai. Le terme effet pépité (voir la figure 2.2 B) est utilisé pour décrire la variabilité à courte portée, même si elle peut être due à d'autres facteurs tels que les erreurs de mesure, les erreurs de localisation, ... etc.

4.1.2. Le seuil ou palier

On appelle seuil du variogramme sa valeur limite pour de grandes valeurs de h . Ce paramètre suscite un intérêt très important dans l'analyse de la régularité des variations locales. En effet, si le seuil d'un variogramme est infini, alors le processus des variations locales n'est pas stationnaire.

4.1.3. La portée « a »

Le variogramme expérimental peut prendre de nombreuses formes, en fonction des données et de l'intervalle d'échantillonnage utilisé. Idéalement, le semi-variogramme augmente avec la distance entre les emplacements d'échantillonnage, atteignant une valeur plus ou moins constante (le seuil) à une portée a . Le seuil se rapproche de la variance d'échantillon $C(0) = \sigma^2$ pour les données stationnaires (Figure 2.3A). Les échantillons séparés

par des distances plus courtes que la portée sont spatialement liés. Ceux séparés par des distances supérieures à la portée ne sont pas spatialement liés car le variogramme est égale à σ^2 , ce qui implique une variation aléatoire. Les portées des variogrammes dépendent de l'échelle d'observation et de l'interaction spatiale des processus du sol affectant chaque propriété à l'échelle d'échantillonnage utilisée.

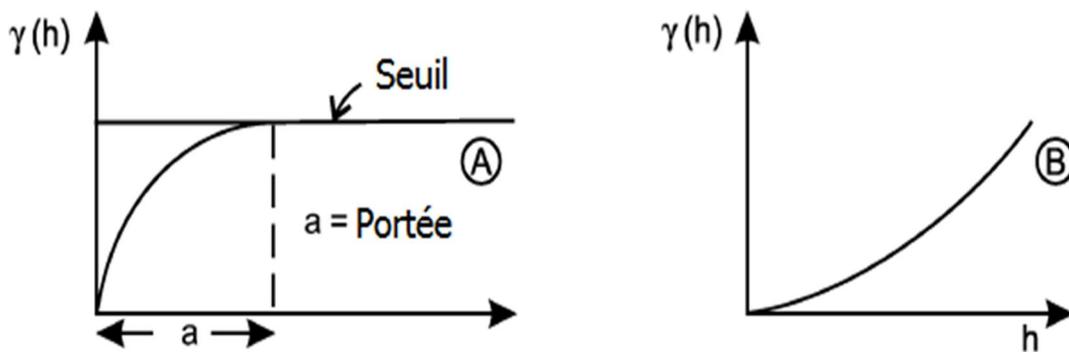


Figure 2.3. (A) Variogrammes bornés et (B) Variogrammes non bornés. Modifié d'après [44]

Les variogramme peuvent également augmenter continuellement sans montrer de portée et de seuil définis (voir Figure 2.3 B), empêchant ainsi la définition d'une variance spatiale, ce qui indique la présence d'effets de tendance et de non-stationnarité [3, 44]. Il n'est pas nécessaire que tous les variogrammes atteignent un seuil. Certains variogrammes continuent d'augmenter avec h . La plage de la portée peut varier d'une direction à l'autre, illustrant l'**anisotropie** du phénomène. De plus, pour une direction donnée, il peut y avoir plusieurs plages, particulièrement lorsque plusieurs structures emboîtées interagissent à différentes échelles de distance.

4.2. Le variogramme expérimental

On peut aussi calculer le variogramme selon certaines directions spécifiques à l'aide d'une expression particulière du variogramme :

$$\gamma_e(h, \theta) = \frac{1}{2N(h, \theta)} \sum_i^{N(h, \theta)} [(Z(x_i) - Z(x_i + h))^2] \dots \dots \dots (2.13)$$

Où : $N(h, \theta)$: nombre de paires de points séparés d'une distance h dans la direction θ .

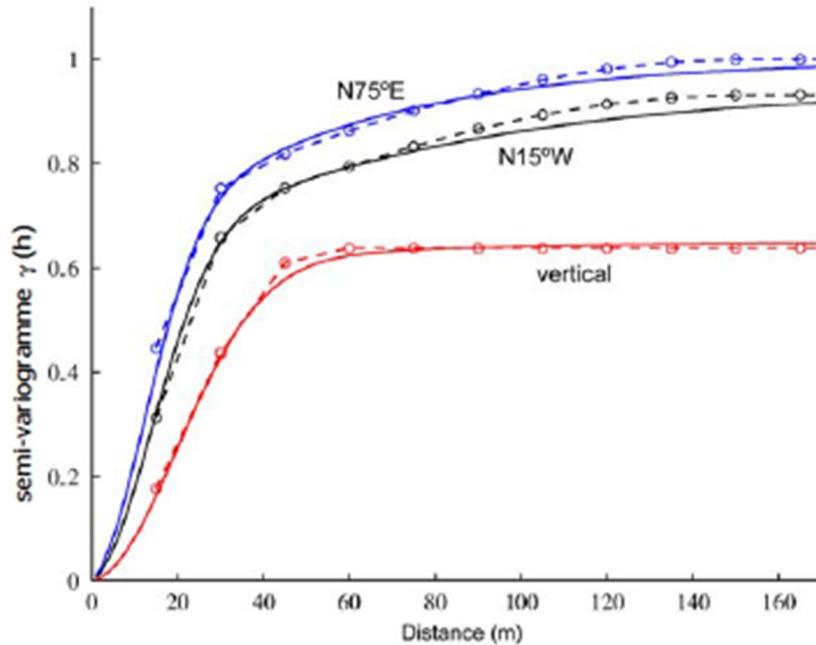


Figure 2.4. Variogramme expérimental du champ aléatoire dans les principales directions d'anisotropie. D'après [46]

En pratique, le variogramme expérimental est généralement calculé à l'aide de vecteurs h dont la longueur est inférieure à la moitié du diamètre de la région [33, 47]. Pour les paires d'échantillons avec des vecteurs h de longueur presque égale au diamètre de la région, les échantillons correspondants sont situés près de la frontière. Les classes de vecteurs h_i formées avec de telles paires ne recevront aucune contribution des échantillons situés au centre de la région et ne seront donc pas représentatives de l'ensemble des données [47]. On s'accorde une tolérance sur h et sur θ afin d'avoir suffisamment de paires de points pour chaque h et chaque θ (Figure 2.4).

Le variogramme expérimental peut présenter une forme qui suggère d'utiliser une fonction de variogramme bornée pour l'ajuster. La borne supérieure la plus basse de la fonction de variogramme est décrite comme le seuil. Lorsque le variogramme expérimental présente un seuil, il est possible de l'ajuster avec un variogramme théorique qui est en réalité une fonction de covariance $C(h)$ sur la base de la formule des variogrammes bornés.

4.3. Principaux modèles théoriques de variogrammes isotropes

4.3.1. Modèle pépitique de palier C

La situation de plus forte d'irrégularité des variations locales correspond à une absence de dépendances entre des mesures effectuées à des sites différents, ce qui se traduit par l'effet de pépité (Figure 2.5A) :

$$\text{➤ Pour tout } h \neq 0 \Rightarrow \begin{cases} C(h) = 0 \\ \gamma(h) = C(0) \end{cases} \dots\dots\dots (2.14)$$

Ce modèle traduit des discontinuités de la variable Z, dues soit à des erreurs de mesure, soit à la présence d'une microstructure [33, 39]. Inversement, la situation théorique de plus grande régularité correspond à des dépendances locales maximales (Figure 2.5B) :

$$\text{➤ Pour tout } h \neq 0 \Rightarrow \begin{cases} C(h) = C(0) \\ \gamma(h) = 0 \end{cases} \dots\dots\dots (2.15)$$

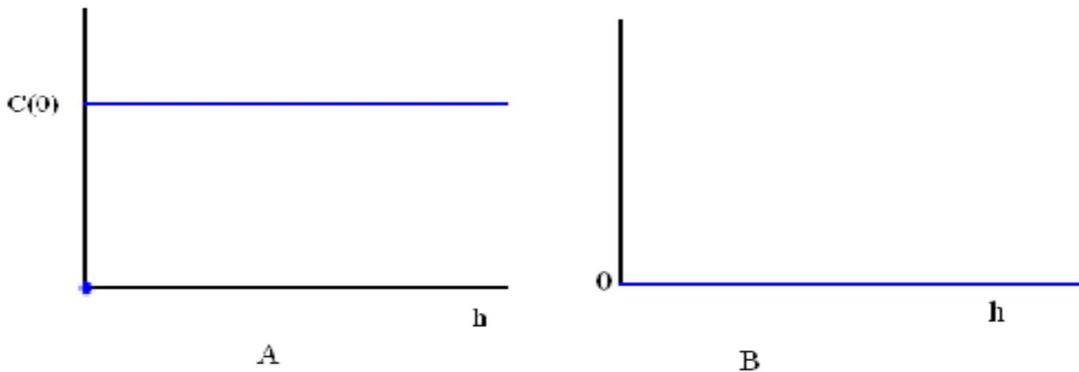


Figure 2.5. A ; B. Schéma type d'un variogramme (effet de pépité), [48].

4.3.2. Modèle sphérique de portée a et de palier C

Le modèle de variogramme sphérique est l'un des modèles les plus couramment utilisés en géostatistique pour décrire la continuité spatiale des données. Ce modèle est particulièrement adapté pour représenter des phénomènes qui présentent une portée limitée dans l'espace.

$$\gamma(h) = \begin{cases} C_0 + C \left(\frac{3h}{2a} - \frac{h^3}{2h^3} \right) & 0 \leq h \leq a \\ C_0 + C & h > a \end{cases} \dots\dots\dots (2.16)$$

L'effet de pépite peut être considéré comme une variation très locale d'un modèle sphérique, avec une portée infiniment petite [3, 39]. Cependant, il est important de noter qu'il existe une distinction fondamentale entre les deux modèles du point de vue de la nature du phénomène : le premier modèle implique une discontinuité à l'échelle d'observation, avec des valeurs changeant brusquement d'un point à l'autre, tandis que le second modèle décrit un phénomène continu mais rugueux (Figure 2.6).

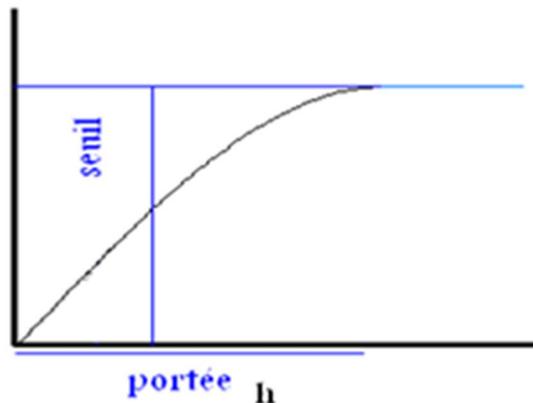


Figure 2.6. Schéma type d'un modèle de variogramme sphérique [48].

4.3.3. Modèle exponentiel de paramètre a et de palier C

Il est particulièrement adapté pour représenter des phénomènes qui décroissent progressivement en corrélation spatiale avec la distance.

$$\gamma(h) = \gamma_0 + \sigma^2 \left[1 - \text{Exp} \left(-\frac{h}{a} \right) \right] \dots \dots \dots (2.17)$$

Notons que le modèle sphérique atteint effectivement son palier pour $h=a$ alors que le modèle exponentiel ne l'atteint qu'asymptotiquement.

4.3.4. Modèle gaussien de paramètre a et de palier C

Ce modèle est souvent appliqué lorsque la distribution des données suit une forme de cloche ou de gaussienne.

$$\gamma(h) = \gamma_0 + \sigma^2 \left[1 - \text{Exp} \left(-\frac{h^2}{a^2} \right) \right] \dots \dots \dots (2.18)$$

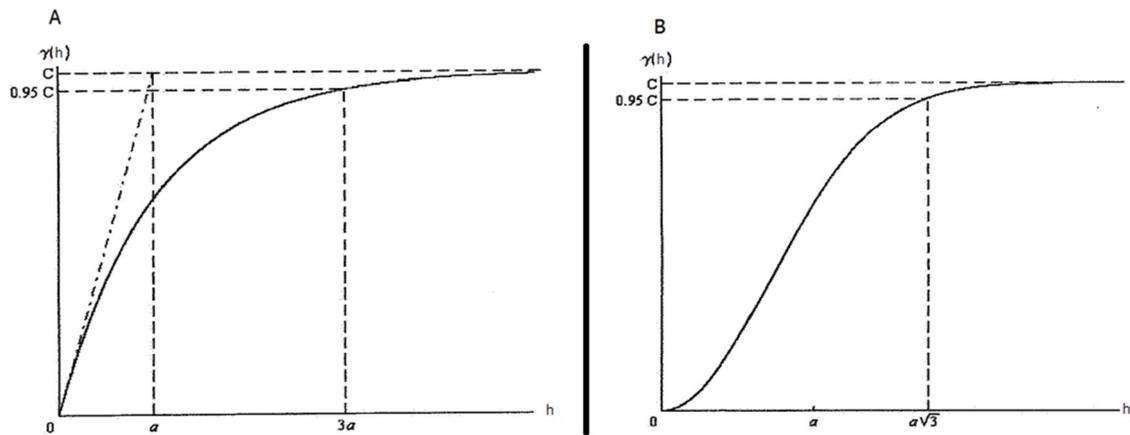


Figure 2.7. Schémas type d'un modèle de variogramme : (A) Exponentiel, (B) Gaussien [39].

Pour ces deux derniers variogrammes (Figure 2.7 A ; B), on peut prendre une portée pratique, pour laquelle le variogramme atteint 95% de la valeur de son palier [3, 39].

Il est essentiel de s'assurer de la représentativité des points expérimentaux. Pour chaque point, il est recommandé d'indiquer le nombre de paires de données utilisées pour son calcul ; un faible nombre de paires indique un point peu fiable [33]. Egalement, il est parfois utile de visualiser la nuée variographique, afin d'examiner la variabilité des données et de détecter celles qui peuvent causer une instabilité numérique dans le variogramme expérimental.

Le choix entre un modèle de variogramme avec ou sans palier, dépend du comportement des variogrammes expérimentaux dans les directions étudiées, en particulier à de grandes distances. En règle générale, la fiabilité du variogramme expérimental diminue au-delà d'une distance équivalente à la moitié du diamètre du champ dans la direction considérée [3, 39]. L'ajustement du palier éventuel est basé sur les fluctuations du variogramme et/ou la variance empirique des données.

4.4. Anisotropies

Lorsque le variogramme est calculé pour chaque paire de points dans des directions spécifiques telles que Nord-Sud ou Est-Ouest, il peut parfois révéler des variations de comportement, ce qui est appelé anisotropie [3]. Si cela ne se produit pas, le variogramme dépend simplement de la distance entre les points, et il est qualifié d'isotrope. On peut distinguer deux principaux types d'anisotropie : l'anisotropie géométrique et l'anisotropie zonale.

4.4.1. Anisotropie géométrique (anisotropie « elliptique »)

L'anisotropie géométrique se produit lorsque les variogrammes présentent la même variabilité globale, et en particulier le palier mais ont des portées différentes [3].

L'anisotropie géométrique signifie que la corrélation est plus forte dans une direction que dans les autres directions [3, 33]. Mathématiquement, si l'on trace les portées directionnelles, elles tomberaient sur le bord d'une ellipse (voir Figure 2.8), où les axes majeur et mineur de l'ellipse/ellipsoïde correspondent aux portées les plus longues et les plus courtes des variogrammes directionnels.

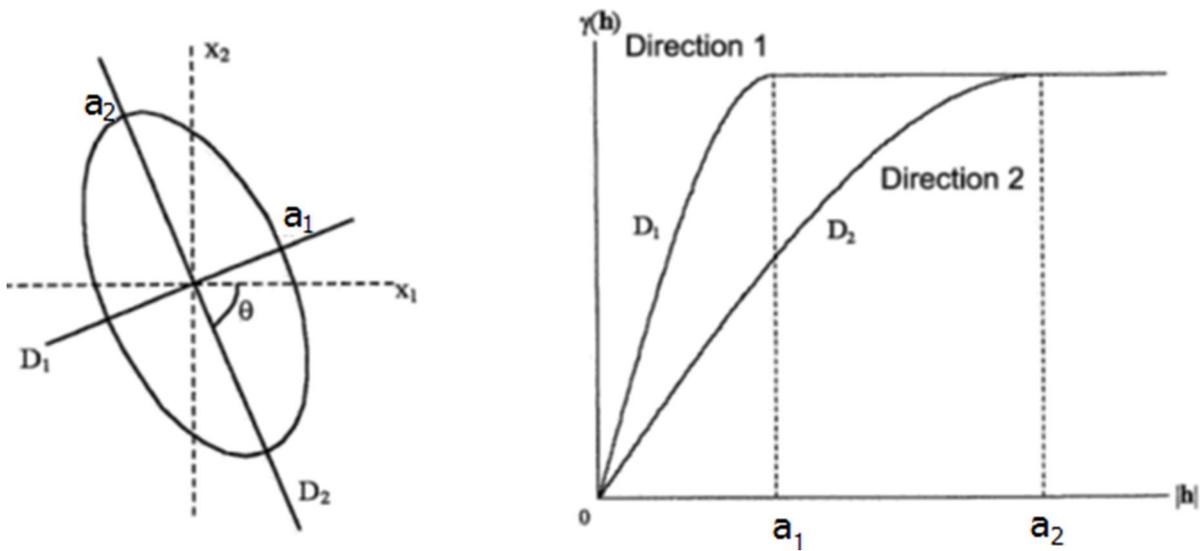


Figure 2.8. Exemple 2D d'anisotropie géométrique de rapport a_1/a_2 et l'angle θ . Modifié d'après [39]

Une manière d'identifier l'anisotropie géométrique est de calculer et de représenter graphiquement des variogrammes expérimentaux directionnels. Les différences dans les variogrammes échantillonnés calculés avec différents angles, pourraient indiquer une anisotropie. L'anisotropie géométrique peut être modélisée en modifiant le modèle de variogramme, pour un processus isotrope en transformant les coordonnées. Cette anisotropie peut souvent être corrigée en appliquant une transformation linéaire des coordonnées spatiales pour rendre la variabilité spatiale isotrope, c'est-à-dire uniforme dans toutes les directions [39, 44, 47]. Cela permet de simplifier l'analyse et la modélisation des données géo-spatiales.

Ensuite, si l'équation du variogramme dans la direction 1 est $\gamma_1(h)$, le variogramme global après correction de l'anisotropie est de la forme [44]:

$$\gamma(h) = \gamma_1 \sqrt{(x_1 - x_2)^2 - b^2(y_1 - y_2)^2} \dots \dots \dots (2.19)$$

Le meilleur estimateur est celui qui fournit l'erreur e la plus petite possible au point estimé. Pour l'ensemble des estimations réalisées, la variance de cette erreur doit être la plus petite possible.

La variance de cette erreur est la variance d'estimation σ_e^2 :

$$\mathbf{Var}(e) = \mathbf{Var}(\mathbf{Z}(x)) + \mathbf{Var}(\mathbf{Z}^*(x)) - 2\mathbf{Cov}(\mathbf{Z}(x), \mathbf{Z}^*(x)) \dots \dots \dots (2.23)$$

Substituant $\mathbf{Z}^*(x)$ par son expression, en fonction des $\mathbf{Z}(x_i)$, donnée en (équation 22), on obtient:

$$\sigma_e^2 = \mathbf{Var}(\mathbf{Z}(x)) + \sum_i \sum_j \lambda_j \mathbf{Cov}(\mathbf{Z}(x_i), \mathbf{Z}(x_j)) - 2 \sum_i \lambda_i \mathbf{Cov}(\mathbf{Z}(x_i), \mathbf{Z}(x)) \dots \dots (2.24)$$

Qui peut être réécrit en fonction du variogramme :

$$\sigma_e^2 = (\sigma^2 - \bar{\gamma}(x, x)) + \sum_i \sum_j \lambda_i \lambda_j (\sigma^2 - \gamma(x_i, x_j)) - \sum_i \lambda_i ((\sigma^2 - \bar{\gamma}(x_i, x))) \dots (2.25)$$

Puis finalement, puisqu'on a habituellement $\sum \lambda_i = 1$, par simplification on arrive à :

$$\sigma_e^2 = 2 \sum_i \lambda_i \bar{\gamma}(x_i, x) - \bar{\gamma}(x, x) - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i, x_j) \dots \dots \dots (2.26)$$

Comme on vient de voir, on peut calculer la variance d'estimation soit en utilisant le covariogramme (équation 2.24) soit le variogramme (équation 2.25)

Rappelons que pour avoir la meilleure estimation on doit avoir la plus petite variance d'estimation possible ; c'est-à-dire qu'on doit minimiser σ_e^2

$$\sigma_e^2 = \mathbf{Var}[\mathbf{Z} - \mathbf{Z}^*] = \mathbf{Var}[\mathbf{Z}] + \mathbf{Var}[\mathbf{Z}^*] - 2\mathbf{Cov}[\mathbf{Z}, \mathbf{Z}^*] \dots \dots \dots (2.27)$$

Rappelons que, ce qui nous intéresse ici est de calculer les poids λ_i qui sont les inconnues. Pour minimiser la variance d'estimation on doit calculer les dérivées partielles $\frac{d\sigma_e^2}{d\lambda_i}$ qui doivent être égale à zéro :

$$\frac{d\sigma_e^2}{d\lambda_i} = 0 \quad i = 1 \dots n \dots \dots \dots (2.28)$$

Cela représente un système d'équation à n inconnues (les n poids), qui donnera la variance la plus petite possible, mais dont la somme des poids ne sera pas nécessairement égale à 1 puisqu'il n'y a aucune contrainte dans ce sens.

Pour que l'estimateur soit sans biais (équilibré), il faut que :

$$\sum_{i=1}^n \lambda_i = 1 \quad \dots \dots \dots (2.29)$$

En effet, dans ce cas, on a

$$E[Z^*] = \sum_{i=1}^n \lambda_i \cdot E[Z_i] = \sum_{i=1}^n \lambda_i \cdot m = m \quad \dots \dots \dots (2.30)$$

Pour obtenir le meilleur estimateur linéaire sans biais, nous devons donc résoudre le système à n+1 équations mais seulement n inconnues :

$$\begin{cases} \frac{d\sigma_e^2}{d\lambda_i} = 0 & i = 1 \dots n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad \dots \dots \dots (2.31)$$

Pour équilibrer le système nous devons introduire une nouvelle inconnue sous la forme du multiplicateur de Lagrange μ et on forme le Lagrangien qui doit être minimisé pour trouver les n inconnues :

$$L(\lambda) = \sigma_e^2 + 2\mu(\sum_{i=1}^n \lambda_i - 1) \\ = \mathbf{Var}[Z_v] + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbf{Cov}[Z_i, Z_j] - 2 \sum_{i=1}^n \lambda_i \mathbf{Cov}[Z_v, Z_i] + 2\mu(\sum_{i=1}^n \lambda_i - 1) \dots (2.32)$$

Le minimum est atteint lorsque toutes les dérivées partielles par rapport à chacun des λ_i et par rapport à μ s'annulent. Ceci conduit au système de krigeage ordinaire :

5.1. Le krigeage ordinaire

La technique de krigeage ordinaire [40, 41], suppose une moyenne inconnue, mais elle exploite l'invariance dans le voisinage du point à estimer. Cette méthode repose uniquement sur l'hypothèse de stationnarité intrinsèque.

$$\left\{ \begin{array}{l} \sum_j^n \lambda_j \text{Cov}[Z_i, Z_j] + \mu = \text{Cov}[Z_i, Z_j] \quad \forall i = 1 \dots n \\ \sum_j^n \lambda_j = 1 \end{array} \right. \dots \dots \dots (2.33)$$

5.2. La variance de krigeage

La variance de Krigeage représente la variance d'estimation minimale, est calculée en insérant les équations de krigeage dans la formule générale de la variance d'estimation.

$$\sigma_k^2 = \sigma_e^2 = \text{Var}[Z_v] - \sum_{i=1}^n \lambda_i \cdot \text{Cov}[Z_v, Z_i] - \mu \dots \dots \dots (2.34)$$

Il est à signaler que La variance de krigeage est indépendante des valeurs observées, elle est uniquement liée au variogramme et à la disposition des points utilisés pour l'estimation par rapport au point (ou bloc) à estimer.

Comme la variance d'estimation s'écrit aussi en termes de variogramme, on peut aussi réécrire le système de krigeage en fonction du variogramme.

Ceci tient au fait que $C(h) = \sigma^2 - \gamma(h)$ et que $\sum \lambda_i = 1$.

$$\left\{ \begin{array}{l} \sum_i^n \lambda_i \gamma(x_i, x_j) - \mu = \bar{\gamma}(v, x_i) \quad \forall i = 1 \dots n \\ \sum_j^n \lambda_j = 1 \end{array} \right. \dots \dots \dots (2.35)$$

Et, alors

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \cdot \bar{\gamma}(v, x_i) - \bar{\gamma}(v, v) - \mu \dots \dots \dots (2.36)$$

Il est intéressant de visualiser le système de krigeage ordinaire et la variance de krigeage ordinaire sous forme matricielle :

$$K\lambda = k$$

$$\sigma_k^2 = \sigma_k^2 - \lambda \cdot k \dots \dots \dots (2.37)$$

Où :

$$K = \begin{bmatrix} \sigma^2 & Cov(Z_1, Z_1) & \cdot & Cov(Z_1, Z_1) & \mathbf{1} \\ Cov(Z_2, Z_1) & \sigma^2 & \cdot & Cov(Z_n, Z_1) & \mathbf{1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ Cov(Z_n, Z_1) & Cov(Z_n, Z_1) & \cdot & \sigma^2 & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \cdot & \mathbf{1} & \mathbf{0} \end{bmatrix} \dots\dots\dots (2.38)$$

$$k = \begin{bmatrix} Cov(Z_1, Z_v) \\ Cov(Z_2, Z_v) \\ \cdot \\ Cov(Z_n, Z_v) \\ \mathbf{1} \end{bmatrix} \text{ et } \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \lambda_n \\ \mu \end{bmatrix} \text{ et } \sigma_v^2 = \bar{C}(v, v) \dots\dots\dots (2.39)$$

Il est essentiel de fournir une estimation aussi précise que possible de la variable dans les zones non échantillonnées. Cependant, cette estimation demeure incomplète sans une évaluation de sa fiabilité, appelée variance d'estimation σ_v^2 .

Il est important de noter que cette variance ne représente pas une mesure directe de la précision de l'estimation de la valeur prédite. Toutefois, elle offre un indice de la qualité de l'estimation. Sa particularité réside dans le fait qu'elle ne dépend pas des valeurs observées aux points d'appui, mais plutôt de la configuration spatiale des données.

6. Techniques de validation

Après avoir ajusté un modèle, nous voulons souvent mesurer sa précision prédictive, que ce soit pour son propre intérêt ou dans le cadre de la comparaison ou de la sélection des modèles [49, 50]. La validation croisée est l'une des approches d'estimation de la précision prédictive hors échantillon en utilisant des ajustements à l'intérieur de l'échantillon [51, 52].

La validation croisée est une technique basée sur la division des données pour réaliser des évaluations prédictives de modèles statistiques [52, 53]. Bien que l'objectif spécifique d'une analyse statistique, tel que le test d'hypothèses ou la prédiction, puisse restreindre l'ensemble des modèles pris en considération, l'évaluation prédictive constitue une base largement applicable et objective pour la comparaison et la sélection de modèles, quelle que soit la gamme d'objectifs de modélisation [54].

En tant que technique d'évaluation prédictive, la validation croisée est extrêmement flexible en raison de la diversité des mesures prédictives (scores) avec lesquelles elle peut être utilisée [55], de la disponibilité de stratégies de division des données qui peuvent être employées pour prendre en compte la structure des données et/ou gérer les coûts computationnels et le biais d'estimation [56], et de son applicabilité étendue aux cadres d'optimisation et bayésiens. Des innovations méthodologiques récentes, souvent impliquant

des méthodes d'approximation [57], ont également amélioré l'efficacité computationnelle des algorithmes de validation croisée.

Le choix d'un modèle approprié et l'optimisation des hyper-paramètres sont le plus souvent effectués en minimisant une estimation de la performance de généralisation obtenue par validation croisée [52].

6.1. Validation croisée par rétention « Hold Cross-Validation »

Hold-out cross-validation est une méthode couramment utilisée pour évaluer les performances des modèles d'apprentissage automatique [58, 59]. Elle consiste à diviser l'ensemble de données en deux parties distinctes : l'ensemble d'entraînement (utilisé pour former le modèle) et l'ensemble de test (utilisé pour évaluer le modèle). Typiquement, une proportion des données est réservée pour l'ensemble de test, tandis que le reste est utilisé pour l'entraînement. Après l'entraînement, le modèle est évalué en comparant ses prédictions aux véritables valeurs dans l'ensemble de test. Cette méthode est simple mais peut introduire de la variabilité dans les performances du modèle, en particulier avec un petit ensemble de test.

Cette méthode, présente l'inconvénient de ne pas utiliser toutes les données disponibles [59, 60]. Si nous utilisons la moitié des données pour l'ensemble de test, alors nous n'entraînons notre modèle que sur la moitié des données, ce qui peut conduire à une mauvaise performance. D'un autre côté, si nous réservons seulement 10% des données pour l'ensemble de test, nous risquons, par pur hasard statistique, d'obtenir une mauvaise estimation de l'exactitude réelle. Nous pouvons tirer davantage de nos données tout en obtenant une estimation précise en utilisant une technique appelée validation croisée en k-folds.

6.2. Validation croisée « K-fold »

L'idée de validation croisée « K-fold » [50, 61-63] est que chaque ensemble de données sert à la fois de données d'entraînement et de données de test. Nous divisons d'abord les données en k sous-ensembles égaux. Ensuite, nous effectuons k cycles d'apprentissage, à chaque cycle, 1/k des données est réservé comme ensemble de test et l'ensemble de données restants sont utilisés comme données d'entraînement. Cette procédure est répétée jusqu'à ce que chaque sous-ensemble ait été utilisé comme ensemble de validation figure 2.9. Enfin, la performance du modèle est évaluée en calculant la moyenne des mesures de performance sur l'ensemble des sous-ensembles de validation.

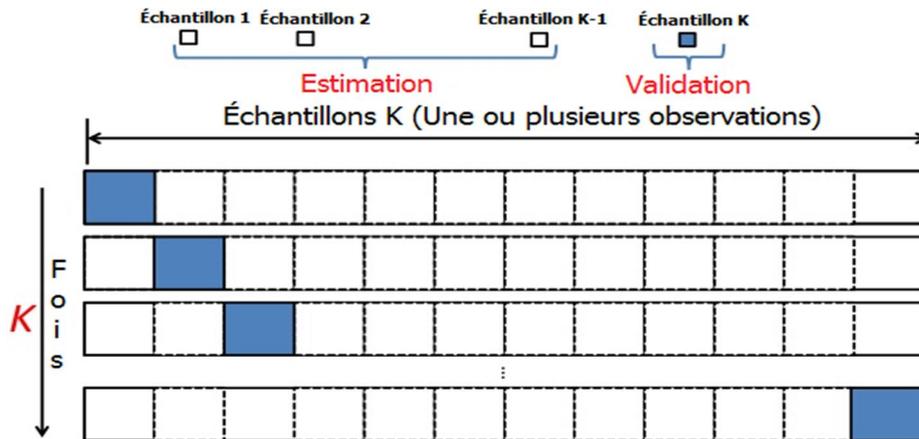


Figure 2.9. Mécanisme d'application de la méthode validation croisée K-Fold.

Cette technique permet d'estimer la performance du modèle de manière plus robuste en utilisant l'ensemble des données disponibles, tout en garantissant que chaque sous-ensemble soit utilisé comme ensemble de validation. La note moyenne de l'ensemble de test des k cycles devrait alors fournir une meilleure estimation qu'un seul cycle. Les valeurs populaires pour k sont 5 et 10, suffisantes pour donner une estimation statistiquement précise, au prix d'un temps de calcul 5 à 10 fois plus long.

Le cas extrême lorsque $k = n$, également connu sous le nom de validation croisée leave-one-out ou « LOOCV ».

6.3. Validation croisée Leave-Part-Out / Leave-One-Out

La validation croisée Leave-part-out (LPOCV) [64, 65] teste un modèle en utilisant toutes les combinaisons possibles de P échantillons sélectionnés de manière aléatoire parmi l'ensemble d'échantillons comme ensemble de test, tandis que les $(n - P)$ échantillons restants servent comme ensemble d'entraînement. Ce processus est répété $C(p, n)$ fois, où " $C(p, n)$ " représente le nombre de combinaisons pour sélectionner P échantillons parmi l'ensemble total de n échantillons (Figure 2.10).

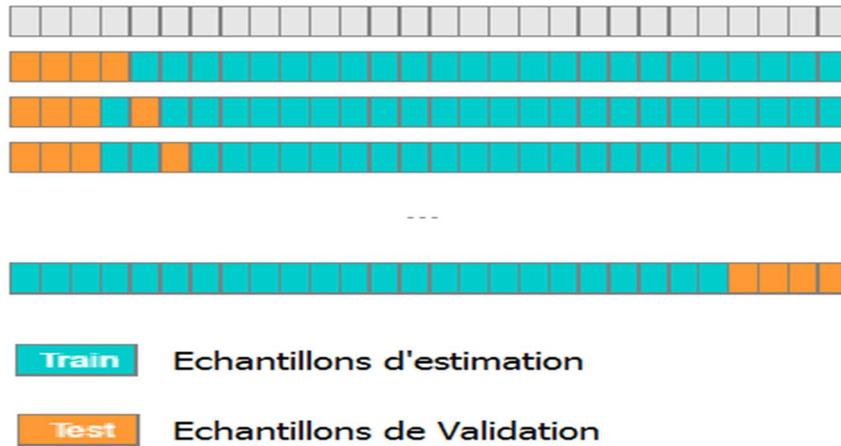


Figure 2.10. Représentation schématique de la méthode de validation croisée Leave-Part-Out « LPOCV ».

La LPOCV peut fournir une estimation d'erreur extrêmement précise, mais peut rapidement devenir épuisante pour de grands ensembles de données. Le nombre d'itérations de test qu'un modèle doit traverser en utilisant la LPOCV peut être calculé en utilisant une combinaison mathématique de $C(p, n)$. Nous pouvons voir, par exemple, qu'une exécution de LPOCV avec un jeu de données de 10 points et 3 points de test nécessiterait 120 itérations.

C'est pourquoi et pour des raisons de calcul, on a tendance à fixer la valeur de P à 1. Cette valeur nous amène à utiliser la méthode de validation croisée LOO (Leave-One-Out) [65].

La Validation Croisée Leave-One-Out (LOOCV) [34, 50, 53, 66] est le cas extrême de la validation croisée k -fold, où k est égal au nombre total d'observations (n). Un classificateur est formé n fois en utilisant $n-1$ observations, et la seule observation restante est utilisée pour les tests de validation (Figure 2.11).

La méthode de Validation Croisée Leave-One-Out (LOOCV) est basée sur la division de l'ensemble de données étudié en deux parties : une paire d'observations pour la validation, c'est-à-dire (x_1, y_1) , et le reste des échantillons $(x_2, y_2), \dots, (x_n, y_n)$ comme échantillons d'apprentissage. Il est important de noter que la paire (x_1, y_1) n'est pas utilisée pour ajuster le modèle. La valeur x_1 et la fonction d'estimation \hat{f}_1 sont utilisées pour trouver la valeur de \hat{y}_1 . Ensuite, l'erreur moyenne quadratique (MSE) du test pour cette paire d'observations est calculée comme suit :

$$MSE_1 = (y_1 - \hat{y}_1)^2 \dots \dots \dots (2.40)$$

Ce processus est répété pour chaque paire (x_i, y_i) , $i = 1, \dots, n$.

L'erreur de test est calculée pour chaque i :

$$MSE_i = (y_i - \hat{y}_i)^2 \dots \dots \dots (2.41)$$

L'erreur de test pour la méthode LOOCV est la moyenne des erreurs :

$$CV = \frac{1}{n} \sum_{i=1}^n MSE_i \dots \dots \dots (2.42)$$

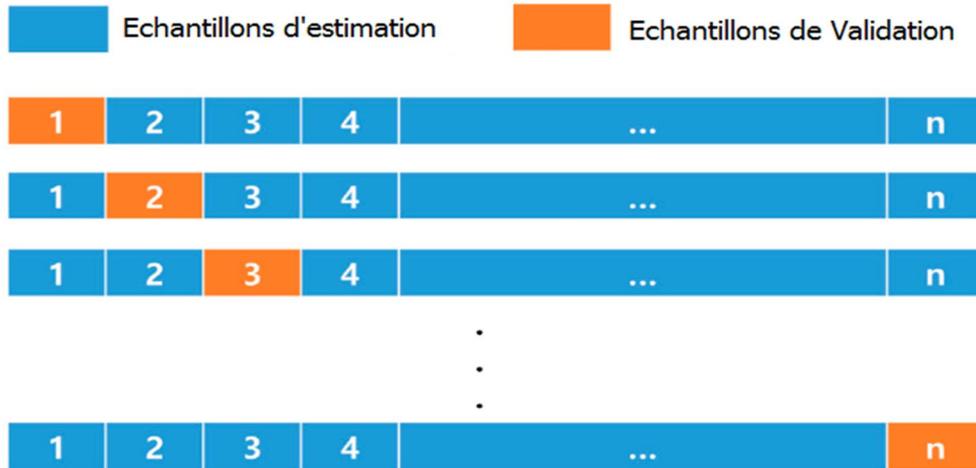


Figure 2.11. Représentation schématique de la méthode de validation croisée Leave-One-Out « LOOCV ».

Le LOOCV est souvent utilisé pour évaluer la performance d'un modèle lorsqu'il y a un nombre limité d'observations dans l'ensemble de données, car il maximise l'utilisation des données disponibles. Cependant, il peut être coûteux en termes de temps de calcul, car il nécessite de former le modèle autant de fois qu'il y a d'observations dans l'ensemble de données.

7. Conclusion

La géostatistique, en particulier le krigeage ordinaire, représente une approche puissante et hautement fiable pour la modélisation spatiale et l'estimation des réserves. En combinant les informations statistiques et géographiques des données, elle offre une solution puissante pour analyser et prédire les valeurs géospatiales inconnues. Cette approche repose sur la création de modèles variogramme minutieusement ajustés, qui capturent les corrélations spatiales et permettent d'obtenir des estimations précises. L'ajustement adéquat de ces modèles est essentiel pour garantir des résultats fiables en prenant en compte la structure spatiale des données et en minimisant les erreurs d'estimation. De plus, l'utilisation de

techniques de validation appropriées, comme K-Fold, LOOCV, LPOCV et d'autres, joue un rôle clé dans l'évaluation de la capacité des modèles géostatistiques à généraliser sur de nouvelles données, assurant ainsi que ces modèles demeurent robustes et précis, même dans des conditions changeantes.

L'application de la géostatistique, notamment dans des secteurs tels que l'exploration minière, contribue considérablement à réduire l'incertitude entourant les estimations des réserves, renforçant ainsi la qualité de la prise de décision. En conséquence, cette approche améliore la gestion des ressources géologiques essentielles pour de nombreuses industries qui dépendent de données spatiales précises pour leurs activités. Elle offre un potentiel inestimable en termes de réduction des risques et d'optimisation des ressources, soulignant ainsi son importance cruciale dans un large éventail de contextes.

Chapitre III : Géologie régionale et locale du gisement de Bled El-Hadba

1. Introduction

Ce présent chapitre a pour but de fournir une vue d'ensemble de données d'exploration du gisement de phosphate de Bled El Hadba. Cette démarche vise à établir les fondements essentiels permettant d'acquérir une compréhension plus approfondie des particularités de ce gisement, notamment ses caractéristiques géologiques [67, 68]. En d'autres termes, c'est pour établir un cadre initial qui servira de point de départ d'une exploration plus approfondie de ce site minier. Cette approche aidera à contextualiser les informations ultérieures et à faciliter une analyse plus complète de ce gisement.

2. Présentation générale du secteur minier de Djebel Onk

Sur le plan administratif, la région de Djebel Onk est rattachée à la wilaya de Tébessa, située à environ 100 kilomètres au sud et à seulement 20 kilomètres de la frontière algéro-tunisienne (Figure 3.1). Cette zone se trouve dans la région Nord-Est de l'Algérie, dans la partie orientale des montagnes de Nememcha, qui sont les derniers maillons de l'Atlas Saharien. Cette zone représente une frontière géographique naturelle entre les hauts plateaux de Constantine et la région saharienne. Le massif de Djebel Onk se caractérise par une structure calcaire s'étendant sur une distance d'environ 20 kilomètres, atteignant son point le plus élevé à 1198 m d'altitude au sommet du Djebel Tarfaya. Cette région est parcourue par de multiples vallées creusées par les oueds (Figure 3.2), qui prennent leur source sur le flanc nord de la zone et convergent vers le sud.

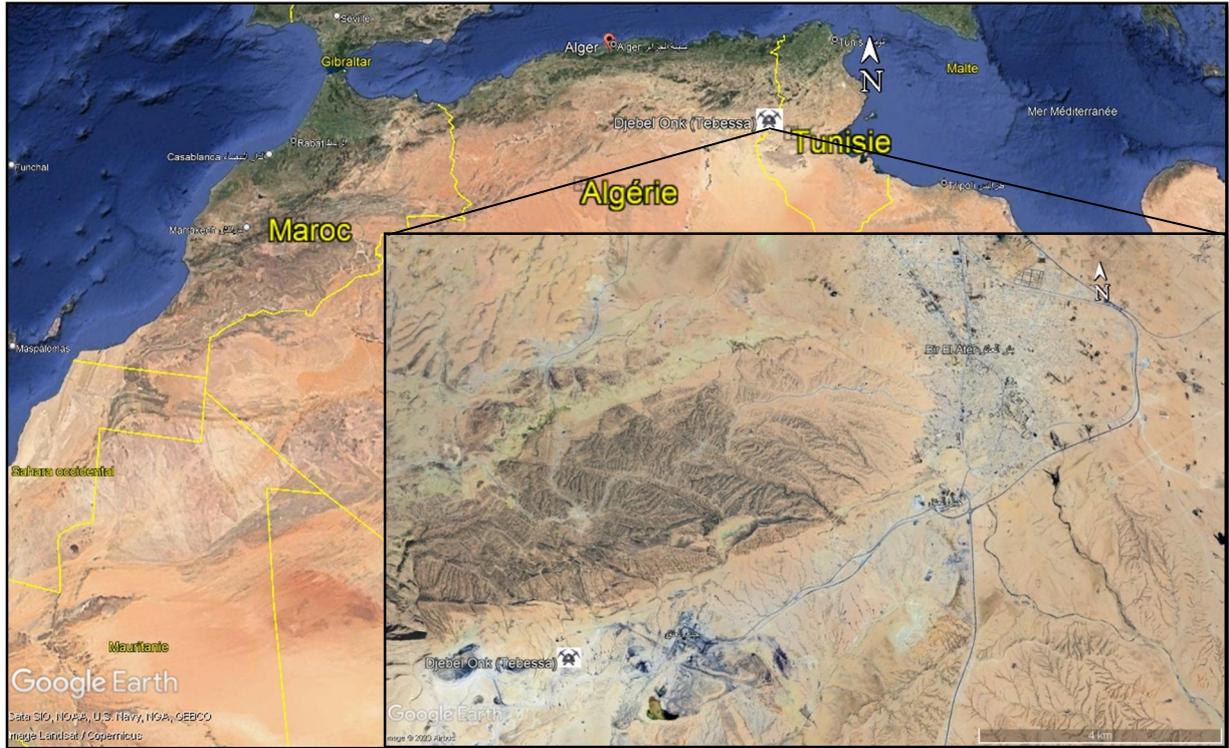
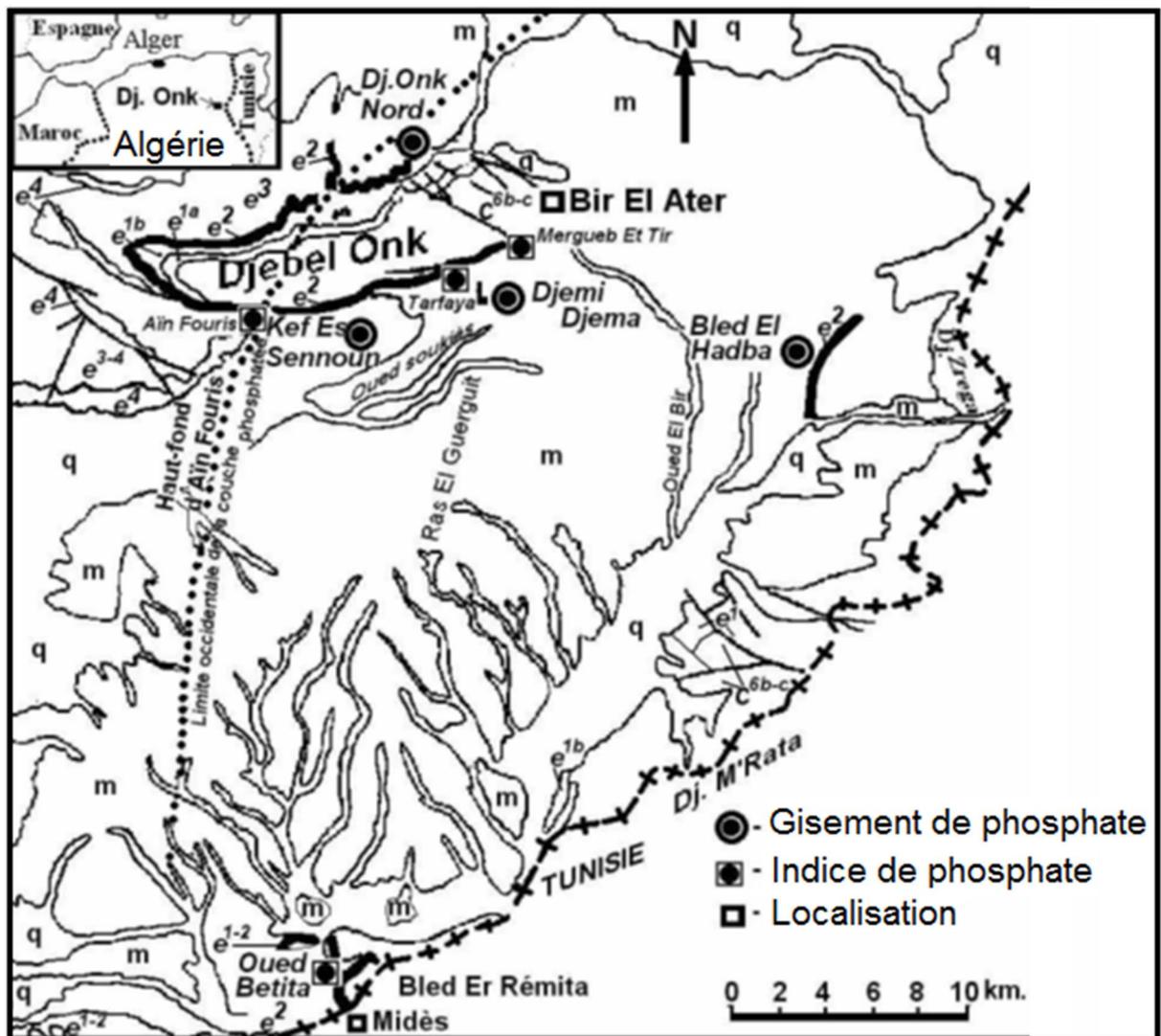


Figure 3.1. Carte de situation géographique de la région de Djebel Onk.

2.1. Contexte géologique du secteur minier de Djebel Onk

Les gisements de phosphates dans la région de Djebel Onk remontent à l'âge Thanétien supérieur. La couche phosphatée présente une épaisseur d'environ 30 mètres. Ces terrains affichent une structure géologique régionale caractérisée par une série d'anticlinaux et de synclinaux fortement asymétriques, avec des failles marquées sur leurs flancs [69]. Cette structure se caractérise par des axes orientés en direction du N80°E au sein du massif du Djebel Onk (Figure 3.2).



Légende:

q Quaternaire : sables, limons, alluvions

m Miocène : sables, argiles

EOCENE

e⁴ Lutétien : argiles bariolées, formation à gypse (formation Souar)

e³⁻⁴ Lutétien à Yprésien : calcaires, marnes, calcaire à nodules de silex.

e³ Yprésien : marnes, calcaire, à silex, quelques niveaux phosphatés (formation mélaoui S.S)

e¹ Thanétien : marnes, calcaires, couche principale de phosphate (formation Chouabine)

e¹⁻² Thanétien à Montien : calcaires, marnes, couches de phosphate inférieures

—+— Stratification horizontale

—+— Stratification à pendage faible à fort

— Contour géologique

⋯ Limite Occidentale de la couche principale de phosphate

e^{1b} Montien : calcaire, lumachelles, marnes (formation Tselja)

e^{1a} Danien : marnes (formation El-Haria)

CRETACE SUPERIEUR

e^{6b-c} Maestrichtien : calcaire compactes (formation Abiod)

Figure 3.2. Carte de situation géographique et géologique des gisements de Djebel Onk. [70]

2.2. Stratigraphie

La région de Djebel Onk a été explorée par de nombreux chercheurs au fil des décennies. Parmi les travaux les plus importants, nous pouvons citer les études menées par des chercheurs tels que Dussert en 1924, Flandrin en 1948, Visse en 1951, Ranchin en 1963, Villain en 1979, et Chabou-Mostefai en 1987.

La séquence sédimentaire exposée dans cette zone est remarquablement diverse, couvrant une succession stratigraphique allant du Crétacé supérieur (Maestrichtien) à l'Éocène moyen (Lutétien). Cette série sédimentaire marine présente une épaisseur considérable d'environ 500 mètres [69-71]. Il convient de noter que cette succession stratigraphique est surmontée de manière discordante par une série continentale de sables-argileuse de Miocène, suivie par des dépôts quaternaires qui remplissent les structures synclinales et conservent les caractéristiques des séquences antérieures à la période miocène (Figure 3.2).

2.2.1. Crétacé

Au cours du Crétacé, il a été observé que le cœur de l'anticlinal du Djebel Onk est formé de couches sédimentaires les plus anciennes. Sur le plan tectonique, ces formations présentent une complexité remarquable, caractérisée par une déformation significative et des irrégularités géologiques prononcées [67]. De plus, seulement les dépôts du Maestrichtien affleurent à la surface.

- **Maestrichtien** : Ces dépôts sédimentaires de Maestrichtien se manifestent sous la forme de calcaires blancs massifs, intercalés de couches marneuses, au cœur de la structure anticlinale de Djebel Onk. Le sommet de cette séquence, est caractérisé par une surface ferrugineuse et tarudée.

2.2.2. Paléocène

Il est représenté par des sédiments marins qui atteignent d'assez grandes épaisseurs (350 m) [72]. Les variations lithologiques et les restes des organismes permettent d'en faire les subdivisions stratigraphiques suivantes :

- a) **Danien** : Il est constitué par un puissant complexe marneux où on note la séparation entre le Danien supérieur et le Danien inférieur :
 - **Danien supérieur** : il est représenté par des calcaires et des intercalations de marnes assez épaisses, avec des caractéristiques spécifiques; calcaires dur de couleur crème ou

blanc s'intercalant avec des marnes tendres argileuses schisteuses, dont l'épaisseur varie d'une dizaine de centimètres à un mètre.

- **Danien inférieur** : il est représenté par des argiles marneuses schisteuses d'une couleur gris foncé ou brun vert avec des intercalations des marnes très dures en bancs irréguliers. L'épaisseur du Danien inférieur est de 30 à 40 m.
- b) **Montien** : Cet étage est caractérisé par la présence de puissantes couches calcaires, où l'on retrouve des spécimens tels que l'*Ostrea Canaliculita* Sow et l'*Ostrea Multicostata* Desh, appartenant à la classe des Pélécy-podes. Ces couches calcaires sont ponctuées d'intercalations de marnes [67]. Ces caractéristiques géologiques spécifiques sont visibles dans plusieurs zones de la région, notamment à Djebel Onk, Djebel Darmoun, Djemi-Djema, Oued Betita et Bled El Hadba. En termes d'épaisseur, cet étage s'étendant sur environ 60 mètres.
- c) **Thanétien** : L'horizon géologique du Thanétien revêt une importance particulière en tant que couche porteuse de la minéralisation dans la région. Cette strate est clairement visible le long du flanc de l'anticlinal de Djebel Onk, s'étalant sur une épaisseur de 72 mètres. Il convient de noter que le Thanétien se subdivise naturellement en deux parties distinctes, chacune présentant ses propres caractéristiques géologiques [67, 69] :
 - **Thanétien inférieur** : se caractérisé par la présence de marnes de couleur gris foncé à noire, qui présentent une texture schistifiée. Dans la partie inférieure de cette couche, on peut observer des conglomérats ainsi que de minces couches phosphatées. De plus, deux niveaux de marnes à faune caractéristique. À mesure que l'on s'approche du sommet de cette strate, des intercalations phosphatées deviennent de plus en plus fréquentes, atteignant parfois jusqu'à 2 mètres d'épaisseur. Ces intercalations sont surmontées par des couches de calcaires et de marnes contenant de gros gastéropodes. L'épaisseur totale de cette strate varie généralement de 30 à 40 mètres.
 - **Le Thanétien supérieur** : il débute par un niveau dolomitique qui contient des gastéropodes. Just en-dessous de ce niveau, on trouve une couche phosphatée, dont l'épaisseur moyenne atteint environ 30 mètres à Djebel Onk et Bled El Hadba, mais qui décroît progressivement au nord jusqu'à disparaître. Du côté ouest et sud de ce gisement, on observe généralement une strate lumachéllique. Il est important de noter que c'est principalement cette couche qui fait l'objet d'exploitation à Kef Essennoun. Cette description détaillée des caractéristiques du Thanétien supérieur offre un aperçu

précieuses des formations géologiques de cette région et de leur potentiel en termes de ressources minérales.

2.2.3. Eocène

- a) **Yprésien** : Il repose directement sur les dépôts du Thanétien, s'étend sur une vaste superficie dans la région de Djebel Onk. Son épaisseur moyenne est d'environ 32 mètres, et il se compose de plusieurs couches distinctes :
- Une couche de calcaire à silex, avec une strate phosphatée intercalée d'une épaisseur d'environ 3 mètres.
 - Une couche s'étalant sur environ 4,5 mètres, composée principalement de calcaire marneux, avec des couches de phosphates situées principalement dans la partie supérieure de cette couche.
 - Une couche de calcaire renfermant des nodules de silex noir. Ces nodules présentent une patine ferrugineuse, allant du noir au brun rougeâtre. L'épaisseur totale de cette couche atteint environ 26 mètres.
- b) **Lutétien** : Cet étage recouvre de manière concordante les roches Yprésiennes. Ces formations géologiques sont largement répandues dans la région de Djebel Darmoun, où elles atteignent une épaisseur d'environ 270 mètres. En analysant cette strate du bas vers le haut, on peut distinguer :
- **Lutétien inférieur** : Se caractérise par des changements significatifs de faciès géologiques. Cette couche présente les éléments suivants lorsqu'on l'analyse du bas vers le haut :
 - Les calcaires laissent place à des marnes blanches, crayeuses, renfermant des géodes de quartz et des nodules de silex. Cette strate a une épaisseur d'environ 30 mètres.
 - La présence de calcaires massifs, de teinte blanchâtre, associés à des nodules de quartz, couvrant une épaisseur d'environ 15 mètres.
 - Cette couche géologique comprend du gypse, avec des bancs de calcaires intercalés, s'étalant sur une épaisseur d'environ 150 mètres.
 - **Lutétien supérieur** : Présente un faciès évaporitique, composé essentiellement de gypse, d'argile verdâtre, d'argile phosphatée verte et de bancs de calcaire. Lorsqu'on examine cette couche géologique de bas en haut, on peut identifier les caractéristiques suivantes :

- À la base, on trouve une strate d'argile verte phosphatée intercalée avec des bancs de calcaire, qui a une épaisseur d'environ 10 mètres.
- Vers la partie supérieure de cette séquence, on rencontre des argiles vertes avec des bancs de gypse intercalés, s'étendant sur une épaisseur totale d'environ 65 mètres.

2.2.4. Miocène

Caractérisé par la présence d'un complexe de roches terrigènes comprenant de conglomérat, d'argiles, de sable et de schiste. Cette période géologique est subdivisée en trois unités distinctes :

- a) **Miocène inférieur** : Se caractérise par la présence de conglomérats, de sables à gros grains, ainsi que de sables blancs à grains moyens, avec de fines couches d'argiles silteuses intercalées. Cette unité géologique a une épaisseur totale d'environ 200 mètres.
- b) **Miocène moyen** : Principalement composé d'argiles de couleur marron, parfois schisteuses, avec des intercalations de sables à grains fins et moyens. L'épaisseur totale de cette unité géologique est d'environ 100 mètres.
- c) **Miocène supérieur** : Principalement composé de roches sablo-argileuses, de grès avec des bancs de graviers, et il contient également de gros blocs arrondis.

2.2.5. Quaternaire

Les formations quaternaires s'étendent sur de vastes surfaces dans la région. Elles se composent principalement d'éboulis de pente, de dépôts sableux, de dépôts éoliens, de gravières, de blocs, ainsi que de dépôts alluvionnaires et fluviaux.

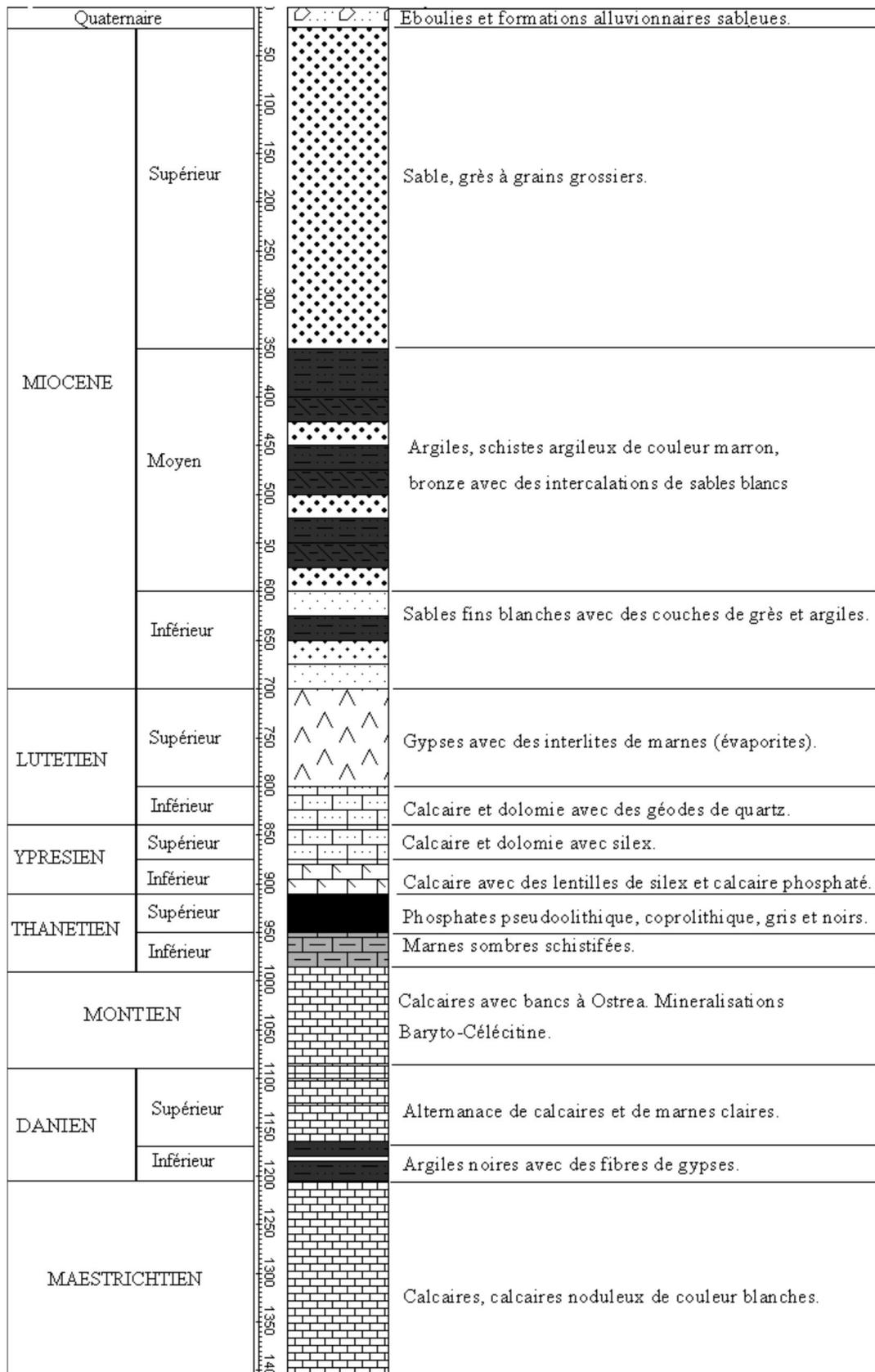


Figure 3.3. Colonne stratigraphique de la région de Djebel Onk – Algérie orientale. [73]

2.3. Tectonique régionale

Les structures géologiques de la région de Djebel Onk sont généralement affectées par des failles sur leurs flancs, et elles suivent principalement une orientation SW-NE. De plus, elles sont décrochées par des accidents transverses orientés en général N120° à N140° E.

La structure géologique dominante dans la région de Bir-El-Ater est l'anticlinal de Djebel Onk, situé au cœur du Crétacé supérieur. Il s'étend sur environ une vingtaine de kilomètres selon une orientation N70°E et possède une largeur d'environ 3 km. Cet anticlinal se caractérise par une forte asymétrie et a été décrit par Visse (1951) comme une flexure anticlinal formée après le Pliocène. Cette structure anticlinale est fortement asymétrique, avec un flanc Nord affichant un pendage faible de moins de 10 degrés, qui supporte le gisement de Djebel Onk Nord. En revanche, son flanc Sud est très raide, parfois vertical, et présente même des zones renversées (Figure 3.4).

De plus, ces anticlinaux des Djebels Onk, Djemi-Djema et Oued Betita se situent à la périphérie de la flexure Sud-Atlasique (Figure 3.5). Cette flexure est le résultat de l'interaction entre la zone mobile de l'Atlas et la plateforme saharienne stable. Plus précisément, ces anticlinaux sont situés dans la branche Nord de la flexure, avec une orientation E-W (Figure 3.4).

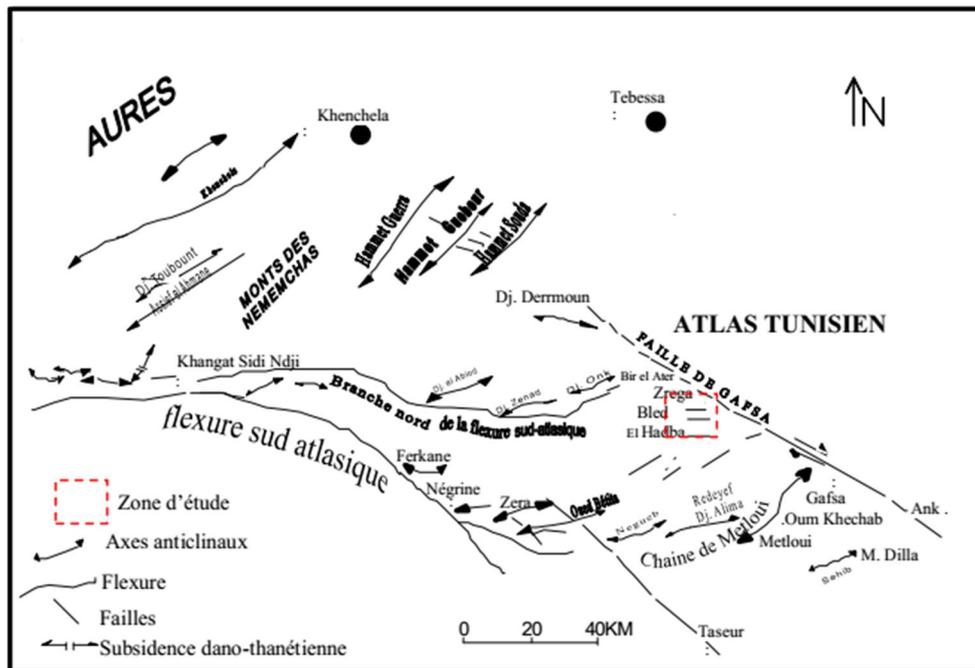


Figure 3.4. Carte structurale de la région du Djebel Onk à l'intérieur du bassin de Gafsa Mélaoui. [74]

Trois phases tectoniques d'ampleurs très inégales ont marqué de leur empreinte la région de Djebel Onk. A savoir :

2.3.1. Tectonique synsédimentaire

Elle est caractérisée par une compression dirigée du NW-SE, survenue pendant la période de sédimentation du Paléogène. Cette compression a eu pour conséquence la formation d'ondulations et de failles au sein des dépôts sédimentaires en cours de formation. Cette activité tectonique précoce a entraîné des interruptions dans le processus de sédimentation pendant la période du Paléocène à l'Éocène. De plus, elle a contribué à la création de structures géographiques telles que des hauts-fonds et des sillons qui ont eu un impact significatif sur le dépôt des phosphates. Il est possible que ces déformations précoces soient liées à la montée diapirique des couches d'évaporites datant du Trias, notamment au Paléocène-Eocène.

2.3.2. Tectonique post-éocène moyen et anté-miocène

Elle se caractérise par plusieurs événements significatifs. Au cours de cette période, la série géologique émergee au Lutétien supérieur, marquant une phase importante de plissement dans la région des Aurès. En parallèle, des déformations distensives transverses ont également émergé, avec une direction principale variant de N120° à N150°. Ces déformations ont initialement commencé à déstructurer les couches géologiques du Crétacé supérieur à l'Éocène en créant de petits horsts et grabens alignés dans une direction approximative de N170°E. Ces plissements et déformations sont en grande partie antérieurs au dépôt des sables miocènes. C'est au cours de cette phase tectonique que l'anticlinal de Djebel Onk a fait son apparition [70].

2.3.3. Tectonique tardi-miocène, post-burdigalienne, à pliocène-pléistocène

Elle a marqué une phase majeure de plissement dans l'Atlas tunisien. Au cours de cette période, les plissements ont affecté les couches géologiques du Crétacé supérieur, du Paléocène, de l'Éocène et du Miocène, créant d'importantes structures en antiformes et en synformes. Ces plis se sont superposés aux structures antérieures, formant des horsts et des grabens dans ces structures préexistantes. Vers le sud et sur le flanc opposé, ces plissements ont donné naissance à d'énormes structures en forme de méga-plis dans les régions de Djebel Onk et Djemi Djema. C'est cette tectogenèse qui a contribué à façonner la structure géologique actuelle de la région du Djebel Onk (Figure 3.5).

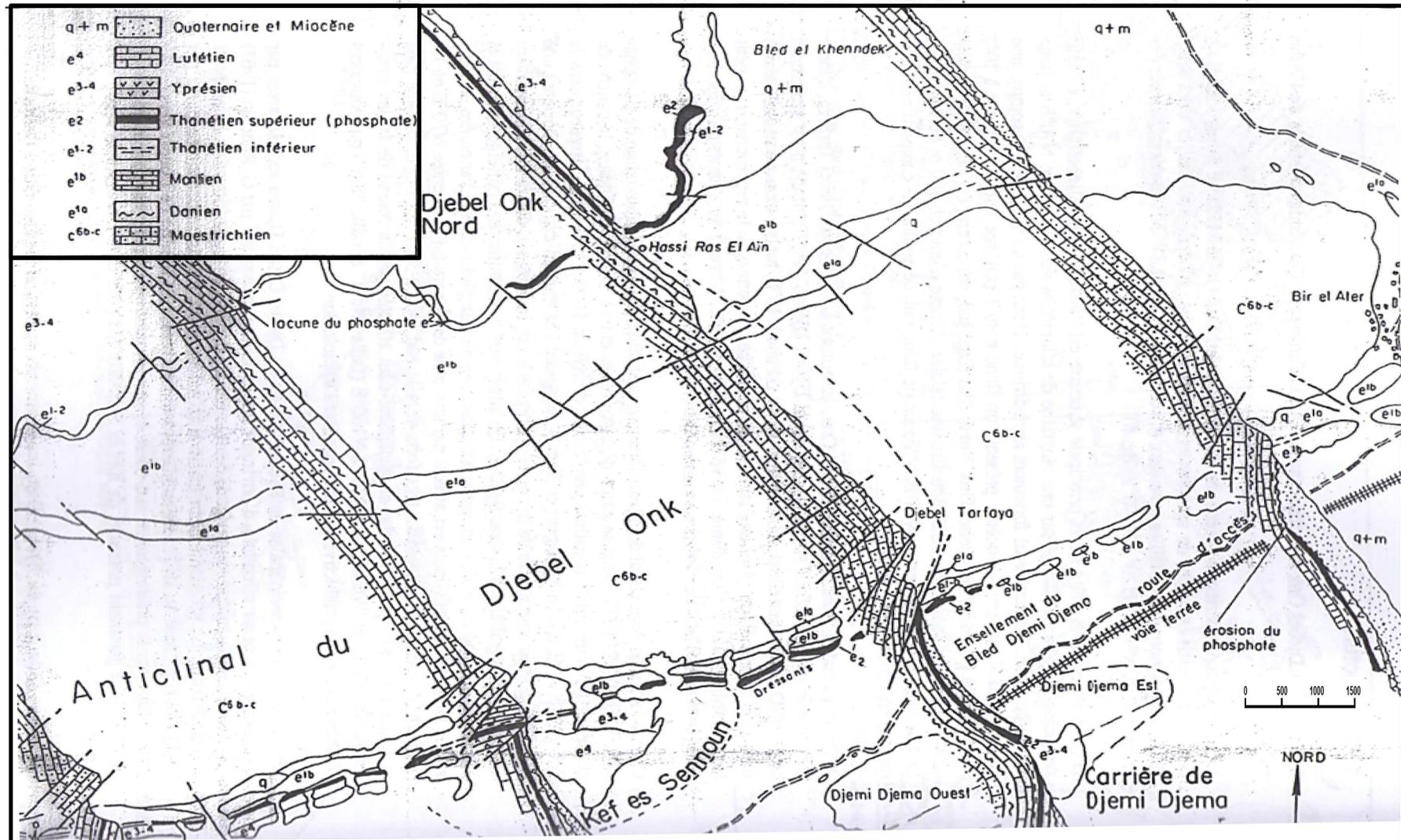


Figure 3.5. Coupes stratigraphiques et structurales des flancs Nord et Sud du Djebel Onk. [67]

3. Présentation du gisement de Bled El-Hadba

Le gisement de Bled El Hadba est situé à 14 Km au Sud est de la ville de Bir El Ater et à 6 Km à l'Ouest de la frontière Algéro-tunisienne. A vol d'oiseau, le gisement est distant de 16 Km de la carrière de Djebel Onk (Figure 3.6).

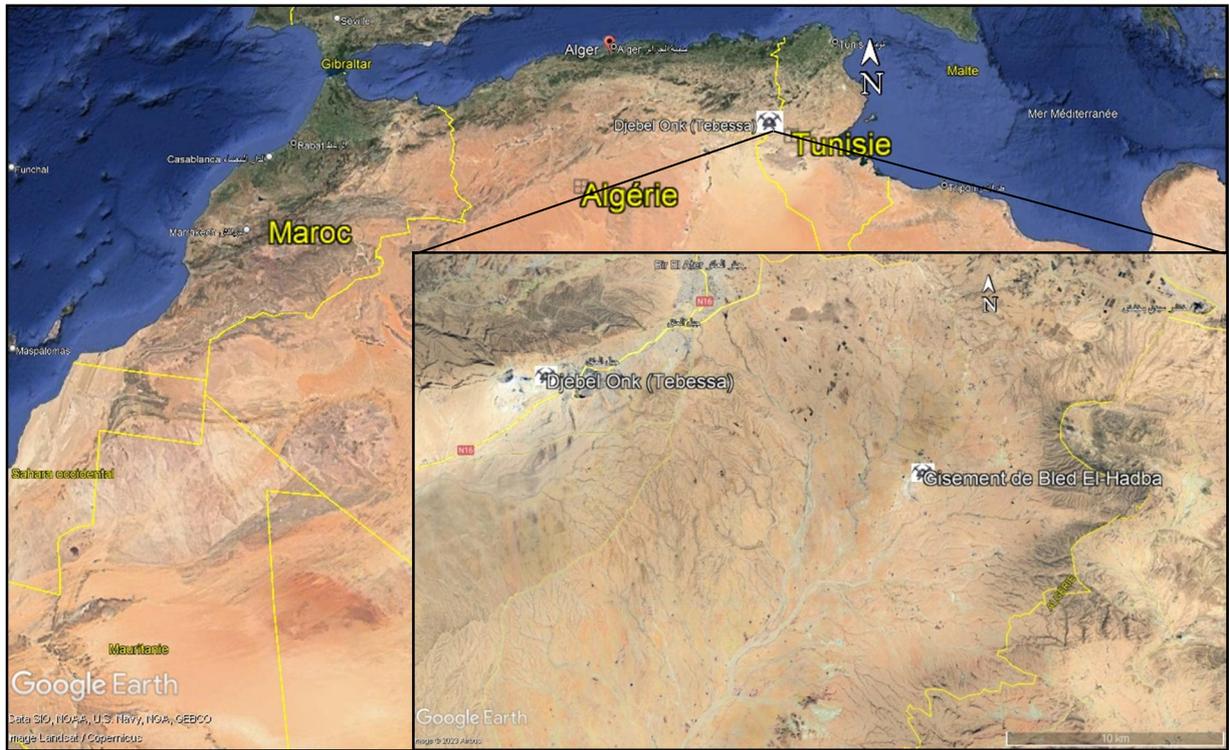


Figure 3.6. Carte de la situation géographique du gisement de Bled El-Hadba.

3.1. Géologie du gisement de Bled El-Hadba

Les phosphates situés à Bled El Hadba ont été identifiés depuis longtemps, remontant à 1888 (Thomas). Dès le début du XXe siècle, des efforts de reconnaissance ont été entrepris à travers la création de puits et de galeries pour explorer ce gisement. La géologie de la zone phosphatée est décrite par Dussert (1924). Quelques données concernant la pétrographie des phosphates sont fournies par Cayeux (1941).

En 1962, Ranchin a repris les données fournies par Dussert et Cayeux, ce qui a abouti à la mise en évidence de la présence de quatre couches distinctes de phosphate dans ce gisement. Au fil du temps, Bled El Hadba a été soumis à plusieurs campagnes d'exploration et de prospection (Figure 3.9), ce qui a contribué à une meilleure compréhension de ses caractéristiques et de son potentiel minier, estimé à environ 1,2 milliard de tonnes.

La géologie du gisement est relativement simple, les couches phosphatées et leurs encaissants plongent d'une façon monoclinale et sous une pente douce de 6 à 10° vers l'Ouest et le Nord-Ouest, ce pendage devient plus accentué en se dirigeant vers le sud du gisement (Annexe N°1 : a, b, c, d, f).

Néanmoins la tectonique est assez compliquée, qui se caractérise par des décrochements nets, horizontaux et des failles de différentes directions. La puissance des couches de phosphates du gisement est importante. Elle est de l'ordre de 41 m. Une grande partie du gisement est à une profondeur de 0 à 60 m.

La superposition de trois sous-couches (sommitale, principale et basale) dans le faisceau total des phosphates a été mise en évidence grâce aux coupes, révélant ainsi la variation en termes d'épaisseur et d'extension latérale de ces sous-couches dans le gisement.

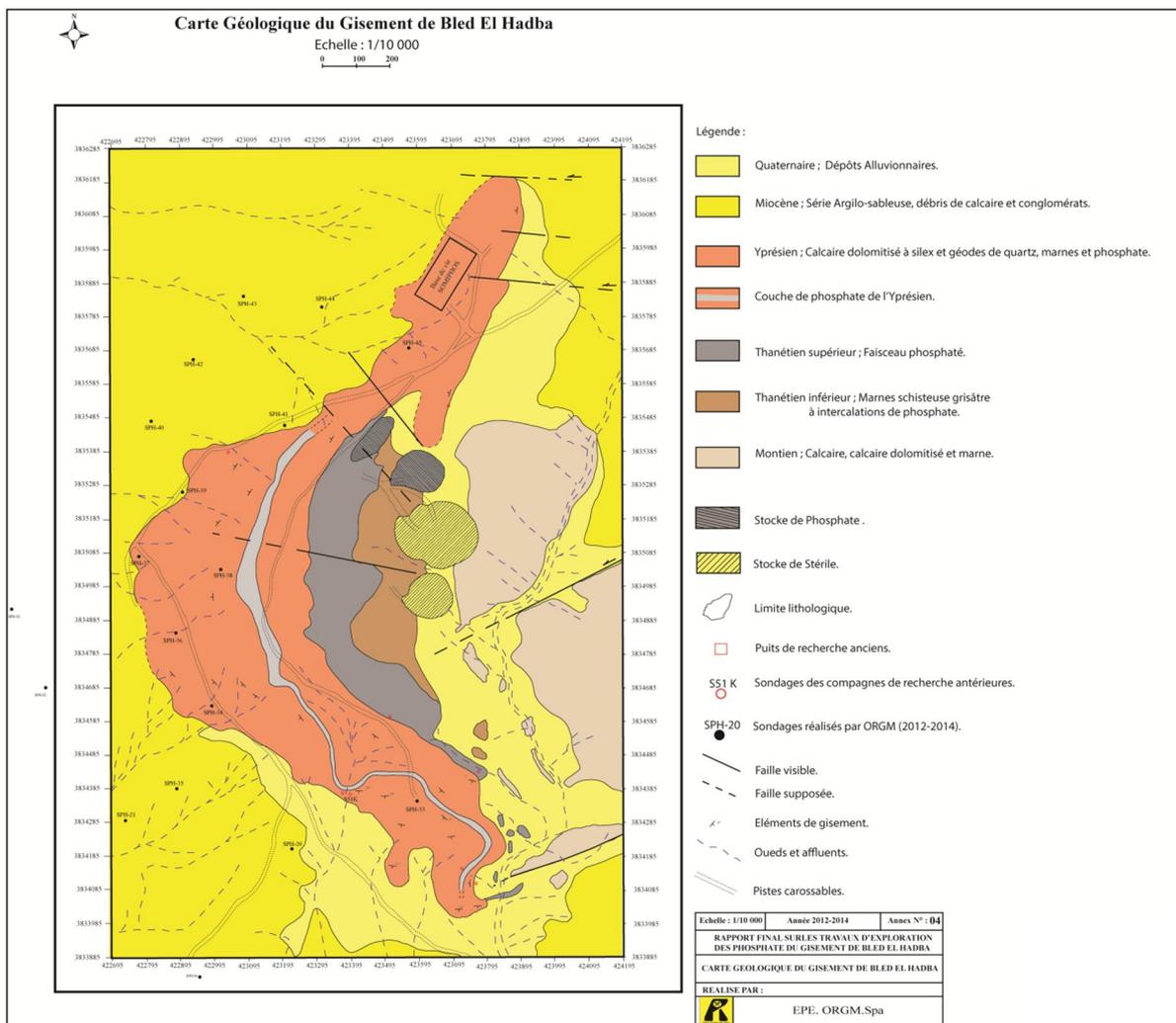


Figure 3.7. Carte géologique du gisement de Bled El Hadba. [75]

3.2. Litho-stratigraphie

La séquence litho-stratigraphique dans cette région présente, dans ses grandes lignes, des similitudes avec celles des gisements de Djemi-Djema et Kef Essenoun. Cependant, il convient de noter d'importantes variations lithologiques à l'intérieur du faisceau phosphaté [67, 75]. Ces variations se manifestent sous la forme d'une bande d'affleurement orientée NE-SW, qui s'enfonce vers l'ouest, sous les formations calcaires à silex de l'Yprésien, tout en étant recouverte par les sables du Miocène (Figure 3.7).

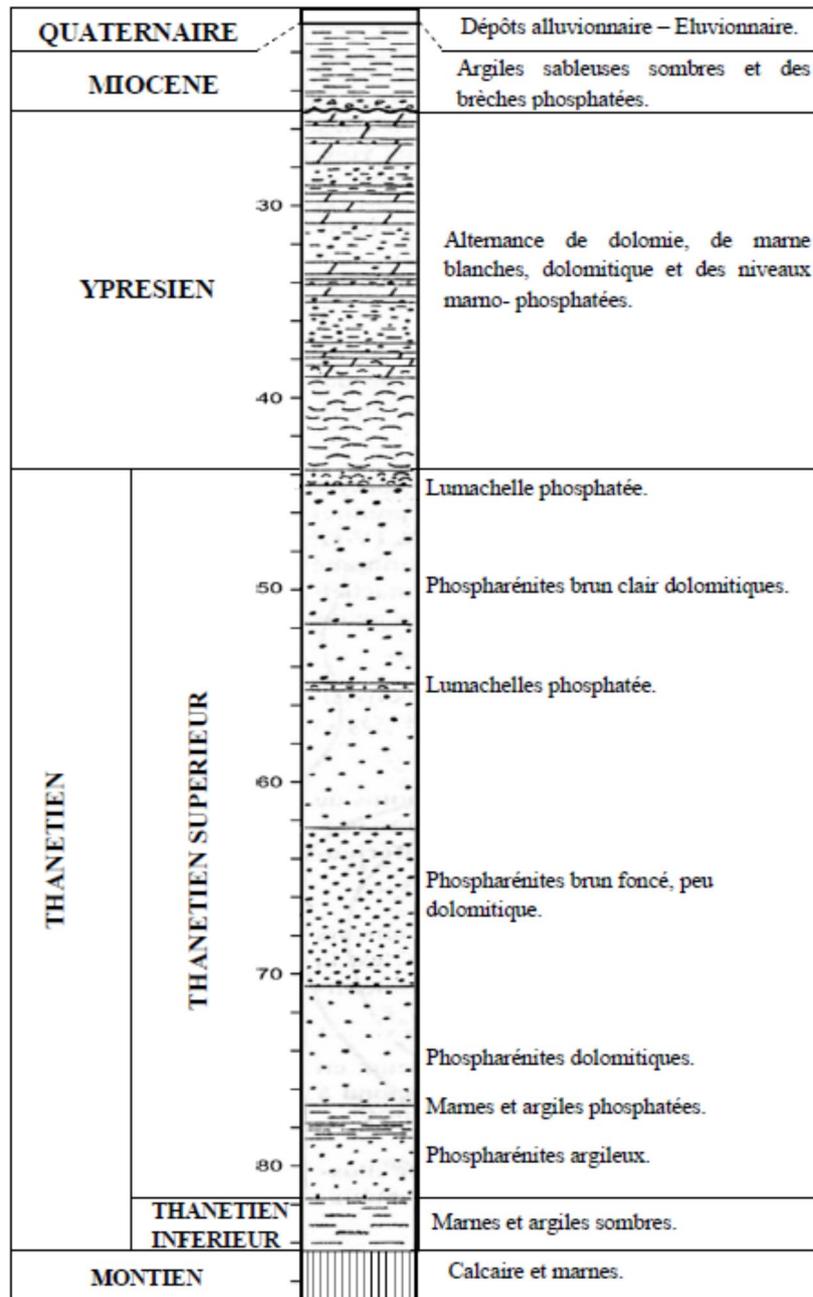


Figure 3.8. Litho-stratigraphie du gisement de Bled El Hadba.[73]

3.2.1. Paléocène

- a) **Montien** : Se trouve en affleurement au sud-est du gisement. À la base de cette couche, on observe des calcaires silicifiés contenant des lumachelles bioclastiques ainsi que d'autres organismes fossiles. Ces calcaires sont surmontés par une série d'alternances comprenant de l'argile lumachéllique, des argiles calcaireuses et des marnes dolomitiques. Ensuite, on trouve un banc de dolomie sur lequel reposent des calcaires gris clairs, légèrement enrichis en phosphate. L'épaisseur de cette strate peut atteindre jusqu'à 100 mètres.

3.2.2. Thanétien

- b) **Thanétien inférieur** : À la base, on trouve des marnes noires, au-dessus desquelles se trouvent des calcaires légèrement lumachelles, intercalées de petites couches de phosphates de couleur marron et grise [75]. L'épaisseur de cette séquence varie généralement de 0.80m dans le sondage SPH-40 à 5.90m dans le sondage SPH-21 (Annexe N° 2a).
- c) **Thanétien supérieur** : Cet étage constitue la couche phosphatée, se distinguant par la présence de grains phosphatés de type coprolithique et pseudo-lithique, à ciment carbonaté. On y observe également la présence d'oxyde de fer, de débris osseux et de dents. Cette couche phosphatée est subdivisée en trois sous-couches distinctes, à savoir la couche basale, la couche principale et la couche sommitale [67]. L'épaisseur totale de cette couche phosphatée est d'environ 30 mètres [75].
- d) **Yprésien** : Au-dessus de la série du Thanétien, on trouve une série calcaro-dolomitique contenant des formations de silex. Cette série, qui correspond à l'Yprésien, est caractérisée par une composition carbonatée, marneuse, dolomitique et calcaire. Elle présente une légère teneur en phosphate et peut atteindre une épaisseur allant jusqu'à 56 m (S-5H) [67, 75].

Au sein de la série yprésienne, on peut observer, au niveau de sa section médiane, de fines couches de phosphate dolomitique (Annexe N° 3), d'une épaisseur variant entre 1,7 et 2,3 mètres. Ces couches de phosphate sont intercalées entre des strates de marnes dolomitiques et de calcaires.

3.2.3. Miocène

Cette série recouvre la série précédente et présente une épaisseur qui augmente progressivement en direction de l'ouest et du nord-ouest. Il est constitué essentiellement par une puissante série de sables quartzeux, de couleur jaunâtre, meuble, intercalant des couches

de marnes argileuses de couleur jaunâtre à verdâtre. Le Miocène débute généralement par des conglomérats de base, formés par des éléments polygéniques, anguleux à sub-anguleux (quartz, calcaire, marnes) et à ciment argilo-gréseux. La puissance de cette série augmente progressivement d'est en ouest, et dans les limites de ce gisement, l'épaisseur maximale atteint 301.00 mètres.

3.2.4. Quaternaire

Les dépôts du Quaternaire couvrent principalement la majeure partie du gisement de Bled El-Hadba. Ils se composent principalement de dépôts éoliens et alluvionnaires.

3.3. Tectonique du gisement

La région de Bled El Hadba se trouve à l'ouest de la structure en forme d'anticlinal du Djebel Zerga. Cette structure anticlinal est formée pendant le Crétacé supérieur et constitue une extension vers le nord de la structure de Djebel M'rata en Tunisie. En comparaison avec cette extension anticlinale principale, les dépôts phosphatés de Bled El Hadba sont situés du côté ouest [67].

La structure géologique de l'ensemble du gisement est relativement simple, car la série stratigraphique du Paléocène, de l'Éocène et du Miocène adopte une disposition monoclinale avec un pendage qui varie entre 5 et 15°. Cette inclinaison est orientée vers le nord-ouest en suivant le flanc ouest du Djebel Zerga. Cependant, la zone de Bled El Hadba est affectée par une série de failles et décrochements horizontaux orientés au NW-SE (voir Figure 3.6). Bien que ces décrochements soient visibles sur les cartes géologiques, ils n'entraînent pas de modifications majeures dans la configuration géométrique de la couche de phosphate [67]. De plus, il est à noter la présence de failles secondaires avec de petits déplacements, variant entre 5 et 20 mètres, orientées WNW - ESE.

Les coupes géologiques réalisées dans le gisement de Bled El Hadba (voir annexes 2 : a, b, c, d, e) révèlent que la couche de phosphate s'étend en formant une bande d'affleurement orientée NE - SW. Cette couche phosphatée s'enfonce progressivement vers l'ouest, sous les couches calcaires de l'Yprésien et les sables du Miocène. L'angle d'inclinaison de cette couche varie entre 5 et 15 degrés.

3.4. Morphologie du corps de minerai de phosphate de Bled El-Hadba

Les diverses strates du minerai dans le faisceau du Thanétien supérieur, ont été classées en trois couches phosphatées à l'aide de l'analyse macroscopique et microscopique, ainsi que

des résultats de l'analyse chimique [75]. Afin de mieux visualiser la distribution du minerai des cartes d'isopaques, ont été élaborées (voir les annexes 3 : b, c, d). Ces couches sont intersectées par l'ensemble des forages et sont présentées de bas en haut comme suit :

3.4.1. La Couche Basale

Présente une épaisseur variable, avec des mesures allant de 2.40 mètres (SPH-40) à 18.00 mètres (S52K) (Annexe N°2b). Cette couche est principalement constituée de phosphates à grains fins et de grains de taille moyenne, associés à un ciment calcaro-dolomitique. Elle présente également des intercalations de lumachelles phosphatées, avec des zones contenant des géodes et des cristaux de quartz.

3.4.2. La Couche Médiane

Cette couche, la principale, se distingue par son épaisseur significative, atteignant jusqu'à 29.75 mètres dans le sondage SPH-18, et par la qualité élevée du minerai, avec des teneurs dépassant les 27% en P_2O_5 et un faible taux de MgO, mesuré à moins de 4% dans le sondage SPH-22. Globalement, cette couche présente une épaisseur considérable dans la partie ouest du bassin (Annexe N°2c). Cependant, elle est caractérisée par des phosphates de haute qualité à travers tout le gisement, affichant un taux de P_2O_5 supérieur à 20% et un taux de MgO inférieur à 4%. Les phosphates se présentent le plus souvent sous forme de grains fins, intégrés dans une matrice argilo-marneuse, avec une texture gréseuse, tendre et friable, présentant des teintes allant du beige au grisâtre, voire au gris foncé avec des nuances de bleu. De plus, cette couche contient des intercalations de lumachelle phosphatée associée à un ciment calcaro-dolomitique ou de phosphate lumachellique, d'une épaisseur variant de 2.50 mètres (SPH-38) à 6.00 mètres (SPH-11), mais avec une qualité de phosphate inférieure, affichant un taux de P_2O_5 inférieur à 20% et un taux de MgO supérieur à 4.5%.

3.4.3. La Couche Sommitale

Affiche une épaisseur variant de 2.30 mètres (SPH-11) à 18.00 mètres (SH-27). Dans la partie nord-ouest du gisement, cette couche a subi une érosion totale, comme observé aux abords du sondage SPH-32 (Annexe N°2d). Sur le plan lithologique, elle présente des similitudes marquées avec la couche basale, composée de phosphates lumachelliques à grains fins et de taille moyenne, associés à un ciment calcaro-dolomitique. Ces phosphates se caractérisent par leur texture dure et compacte, avec des teintes allant du grisâtre au beige, voire au gris foncé, et comprennent des intercalations de lumachelles phosphatées, ainsi que des géodes de quartz.

4. Travaux antérieurs effectués sur le gisement de Bled El Hadba

Dans le cadre des travaux d'exploration des phosphates du gisement de Bled El Hadba qui ont pour objectif, la précision et l'augmentation des réserves de minerais de phosphates vers les catégories supérieurs, ainsi que la recherches de nouvelles ressources, L'ORGM a réalisé 45 sondages carottés, noté SPH01-SPH02...SPH45, couvrant une superficie de 660.43 ha, totalisant une longueur de 8221.40 m. Les sondages les plus profondes sont le SPH02 et SPH18 (336 m), le sondage le plus court est le SPH45 (66 m).

Sur le plan, les sondages sont projetés en fonction de la maille d'exploration appropriée pour chaque catégorie des réserves, en incluant les ouvrages des campagnes antérieurs dans l'ensemble des blocs.

- ✓ Pour le bloc « B », la maille est de (200 x 250 m) ;
- ✓ Le bloc « C1 » : (400 x 500 m) ;
- ✓ Le bloc « C2 » : > 500 m.

Tous les sondages (Figure 3.9), ont été implantés verticalement, vu le pendage doux de la couche de minerai ; de 5 à 10° vers l'ouest.

Les sondages des campagnes de recherches antérieurs qui sont pris en considération sont réparés comme suite :

- SONAREM.1978: S27K, S51K, S52K, S53K, S54K, de maille de 400x900 m.
- EREM.1987: S01H, S02H, S03H, S04H, S05H, de maille de 400x600 m.
- BRGM.1993:D20.
- EREM/FERPHOS.1998: BH02, BH04, BH06, BH8, BH10.

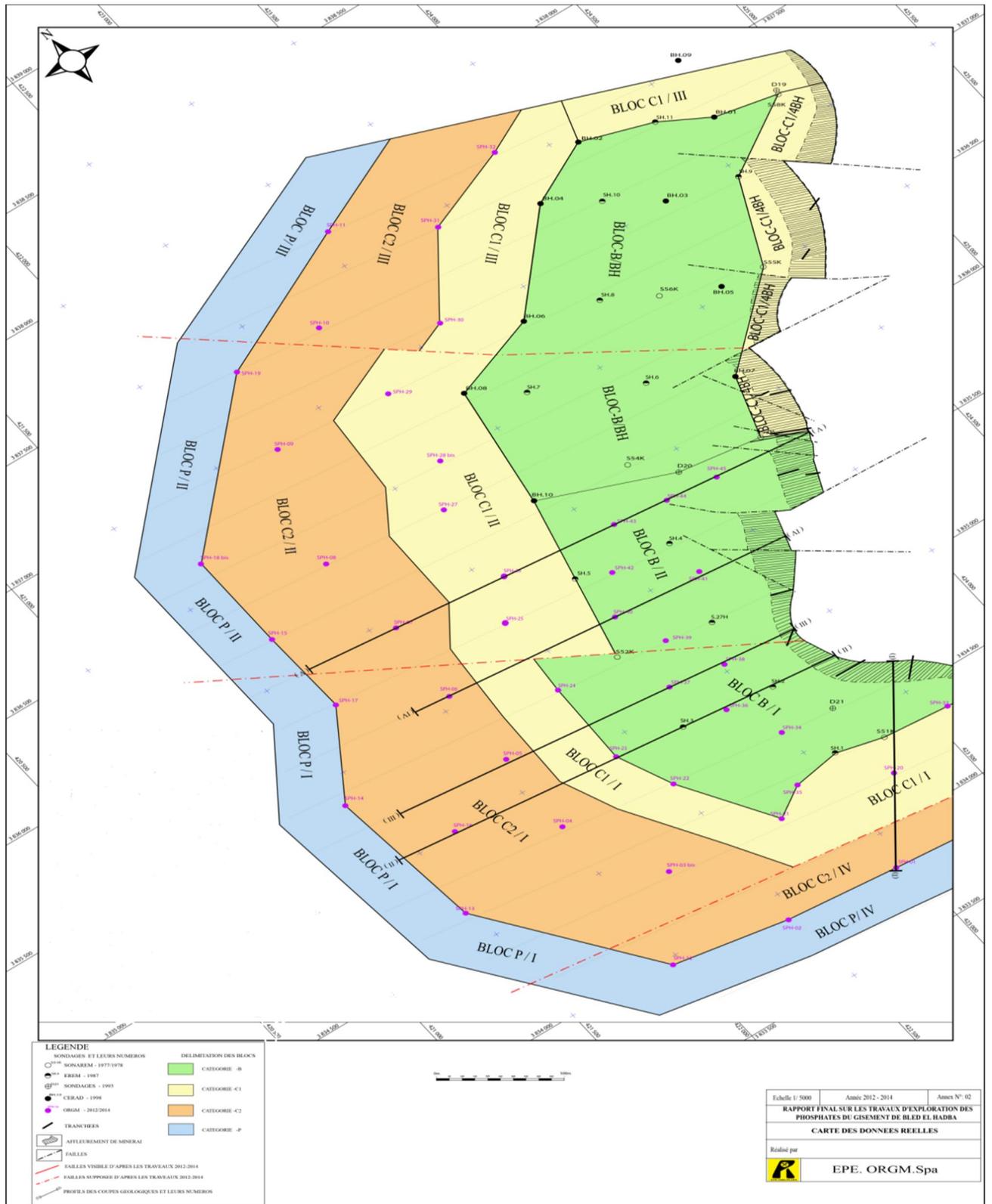


Figure 3.9. Plan d'implantation des sondages des différentes campagnes de prospection.

En se fondant sur les données recueillies lors des sondages, l'entreprise ORGM et le bureau d'étude DMT, ont procédé à l'estimation des réserves minières du gisement de Bled el Hadba. Cette estimation a été réalisée à l'aide de différentes méthodes, notamment celle des

blocs géologiques employée par l'ORGM en 2015, ainsi que la méthode de l'inverse de la distance mise en œuvre par DMT en 2016. Ces approches ont été utilisées pour évaluer les quantités potentielles de ressources minérales présentes dans le gisement, offrant ainsi des perspectives variées sur la répartition et l'importance des réserves minières.

4.1. Evaluation des réserves effectuée par d'ORGM

Le programme de recherche (2012/2014) de phosphate du gisement de Bled El Hadba, entrepris par l'ORGM, a pour objectif la précision et l'augmentation des réserves de minerais de phosphates vers les catégories supérieures, ainsi que la recherche de nouvelles ressources.

Dans cette perspective, l'ORGM a réalisé 45 sondages notés (SPH-01 à SPH-45) , couvrants une superficie de 660.43 ha. Les réserves géologiques globales du gisement, toutes catégories confondues sont évaluées à environ 889.17 Mt. avec une teneur moyenne en P_2O_5 de 21.05 %. Tandis que la couche médiane (principale) recèle 367.40 Mt [75].

Les cinq évaluations historiques (SONAREM, EREM, BRGM, FERPHOS et ORGM) des ressources ont été préparées d'après des niveaux d'exploration différents. Cependant, toutes les déclarations de ressources historiques reposaient sur des estimations de ressources par bloc « coupe géologique » qui est fonction des épaisseurs de la minéralisation, des teneurs moyennes interceptées par les trous de forage et des zones d'influence correspondantes. En combinaison avec une densité moyenne globale, un tonnage total par bloc avec des teneurs moyennes a été calculé. La classification des ressources se conformait aux lignes directrices soviétiques basées principalement sur le maillage des forages.

4.2. Evaluation des réserves effectuée par DMT (2016)

La déclaration des ressources de DMT pour BEH repose sur les mêmes données de base que celles dont disposait le BRGM en 1992, tandis que le nombre de forage a augmenté de 32 à 61 trous de forage.

La déclaration des ressources de DMT correspond aux teneurs moyennes est estimé de l'ordre de 24.1%, le tonnage a augmenté passant de 367 à 463 Mt pour la couche principale (médiane) [76], et ce pour le même niveau d'exploration en comparaison avec l'estimation de l'ORGM 2015. Contrairement aux estimations historiques des ressources, la déclaration des ressources de DMT ne se fonde pas sur une estimation par bloc, mais sur l'interprétation et la modélisation géologiques en 3D des couches de phosphorite. En outre, l'estimation de DMT

se fonde sur des analyses statistiques et une interpolation des teneurs et d'épaisseurs par des méthodes géostatistiques.

5. Conclusion

Ce chapitre offre une exploration approfondie du secteur minier de la région de Djebel Onk. En examinant la géologie complexe de la région, qui comprend une diversité de formations du Crétacé supérieur à l'Éocène moyen, l'accent est mis sur les dépôts phosphatés d'une épaisseur notable. Notre étude s'est ensuite concentrée sur le gisement de Bled El-Hadba, caractérisé par une géologie simple, avec des couches phosphatées inclinées vers l'ouest, le gisement est également influencé par des failles et des décrochements horizontaux, ajoutant ainsi une dimension complexe à sa structure.

Les études antérieures sur le gisement de Bled El Hadba ont révélé des divergences significatives dans les résultats d'estimation des réserves. En effet, en se basant sur les mêmes données d'exploration, la première étude réalisée par l'ORGM en 2015 a donné une estimation d'environ 367 millions de tonnes, tandis que la deuxième étude, conduite par DMT en 2016 sur la même couche (principale), a abouti à une estimation d'environ 463 millions de tonnes. Cette différence a constitué le principal défi qui a motivé notre engagement dans cette étude approfondie.

Face à une divergence substantielle entre les deux évaluations, une analyse approfondie est indispensable. Il est essentiel d'examiner les méthodologies utilisées, les paramètres pris en compte, ainsi que l'éventuelle variation dans les procédures d'estimation. Le chapitre à venir expose la structure d'une approche intégrant l'évaluation géostatistique des réserves minières à une procédure de validation croisée Leave-One-Out. Cette démarche vise à sélectionner le modèle variographique optimal, qui joue un rôle important dans la phase de krigeage et, par conséquent, exerce une influence directe sur les résultats d'estimation.

Chapitre IV : Méthodologie et structure de l'approche employée

1. Introduction

Ce chapitre est consacré à la présentation de l'approche géostatistique, visant à améliorer la procédure d'estimation des paramètres des projets miniers. Cette méthodologie repose sur une démarche approfondie, qui se concentre sur la sélection des modèles variographiques les plus représentatifs des variables d'intérêt, en utilisant une approche de couplage qui intègre l'ajustement des modèles, la validation croisée, et le krigeage. Au cœur de cette méthodologie, la méthode des moindres carrés est employée en conjonction avec des critères de précision et d'efficacité, permettant ainsi un suivi minutieux de l'adéquation des modèles et une évaluation de leur qualité. Le choix du modèle optimal est déterminé grâce à une procédure de validation croisée leave-one-out (LOOCV), tandis que l'estimation des variables d'intérêt est effectuée à l'aide du krigeage ordinaire. Pour simplifier et automatiser l'ensemble du processus d'estimation, les différentes étapes de cette approche ont été intégrées dans un code de calcul développé sous MATLAB, offrant ainsi une solution plus accessible et efficace.

2. Description de l'approche employée

Dans le but de faciliter et d'améliorer l'estimation des réserves minières, une approche géostatistique est employée. Cette approche comporte plusieurs étapes décrivant le processus d'interpolation par krigeage, associé à un couplage ajustement-validation croisée [52, 53].

Un code de calcul sous MATLAB est mis en œuvre afin d'automatiser le traitement et d'améliorer la performance de la procédure de calcul des réserves. Cette initiative vise non seulement à accroître l'efficacité de la procédure, mais également à renforcer sa robustesse et sa fiabilité, en minimisant les erreurs potentielles associées au choix manuel des paramètres des variogrammes ou à l'utilisation de logiciels qui traitent directement les données issues de l'ajustement automatique. Le code de calcul élaboré, constitue ainsi un outil essentiel pour garantir une gestion plus précise des données, permettant une meilleure estimation des ressources minérales et contribuant à l'optimisation globale des opérations l'estimation.

2.1. Structure de l'outil élaboré

Le code de calcul est développé sous forme de boucles imbriquées, ce qui permet de gérer de manière systématique et structurée les différentes étapes du processus géostatistique. Ces boucles permettent une manipulation fluide et efficace des complexités inhérentes aux données géologiques et géostatistiques. Les observations recueillies lors des campagnes d'exploration jouent un rôle important en tant que données d'entrée pour le modèle élaboré. Ces données, sont exploitées de manière itérative par le code, garantissant une analyse

approfondie et une estimation précise des réserves minérales. Ce mode de développement du code assure une intégration harmonisée des différentes phases du calcul géostatistique, contribuant ainsi à la qualité et à la fiabilité des résultats obtenus.

Les étapes du processus de calcul sont organisées de la manière suivante :

2.1.1. Variogrammes Expérimentaux et Ajustement des Modèles Théoriques

Cette phase débute par le calcul des variogrammes expérimentaux en utilisant les données de sondages fournies. Ces variogrammes expérimentaux sont ensuite soumis à un ajustement itératif des modèles théoriques par la méthode des moindres carrés. Cette étape garantit une représentation précise de la variabilité spatiale des paramètres géologiques.

2.1.2. Identification du Meilleur Modèle Ajusté via la LOOCV

Pour déterminer le modèle de variogramme optimal, la méthode de validation croisée leave-one-out (LOOCV) est employée. Cette technique implique le calcul du variogramme expérimental à partir de (n-1) échantillons, puis la comparaison des variables estimées avec l'échantillon restant. Cette procédure est répétée n fois, permettant ainsi d'identifier le modèle de variogramme offrant la meilleure précision et adéquation aux données.

2.1.3. Établissement du Matrice de Maillage de la Zone d'Étude

La zone d'étude est divisée en sections discrètes, et une matrice de maillage est créée en fonction du pas d'espace sélectionné. Cette étape prépare le terrain pour l'estimation des variables dans des emplacements spécifiques de la zone d'étude.

2.1.4. Établissement des Matrices de Krigeage

Des matrices de krigeage sont élaborées pour chaque point du maillage, facilitant l'estimation précise des variables et de leurs variances d'estimation associées. Cette approche offre une représentation spatiale détaillée des paramètres géologiques dans la zone d'étude.

2.1.5. Affichage des Résultats sous Forme Graphique et Tabulaire

Les résultats d'estimation sont présentés de manière visuelle à travers des graphiques, mettant en évidence la distribution spatiale des variables étudiées. Des représentations tabulaires complètent cette visualisation, offrant une compréhension plus complète des caractéristiques géostatistiques de la zone d'étude.

2.2. Organisation algorithmique pour l'approche élaborée

Le diagramme algorithmique dans la figure 4.1 englobe différentes phases de calcul. Commenant par le calcul des variogrammes expérimentaux à partir des données de sondages, l'ajustement itératif des modèles théoriques par la méthode des moindres carrés, et l'identification du meilleur modèle de variogramme ajusté grâce à la méthode de validation croisée (LOOCV). Enfin, les paramètres du variogramme sélectionné sont utilisés pour l'estimation par krigeage (figure 4.1).

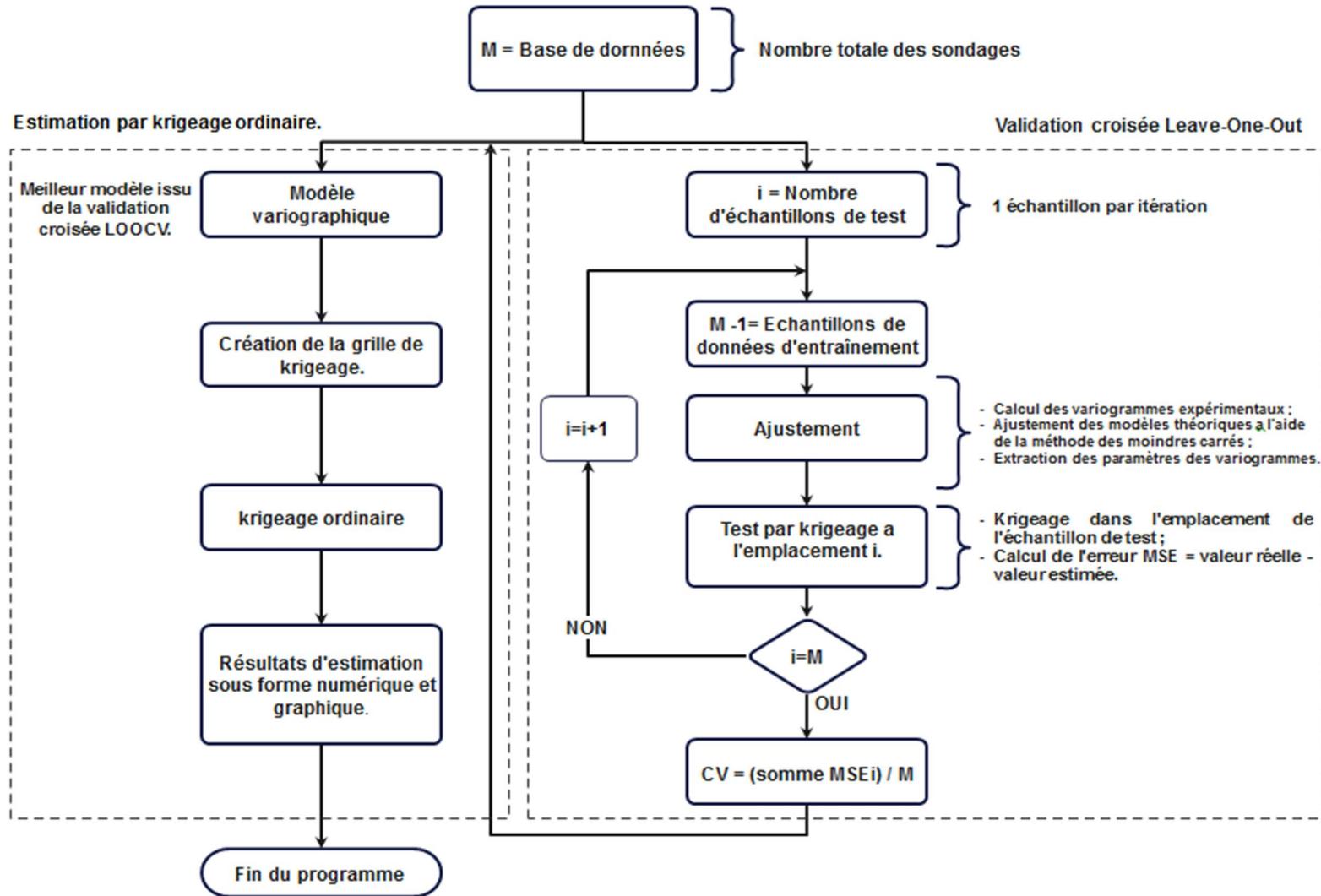


Figure 4.1. Diagramme du processus de calcul impliquant l'ajustement des modèles, la validation croisée (LOO), et le krigeage.

3. Méthode d'ajustement

La méthode des moindres carrés est couramment utilisée dans divers domaines pour ajuster un modèle mathématique aux données expérimentales [77, 78]. Cette approche repose sur la minimisation des carrés des écarts entre les observations réelles et les valeurs prédites par le modèle. Elle constitue une technique fondamentale dans l'optimisation de paramètres afin d'obtenir une représentation mathématique qui capture au mieux le comportement des données. En géostatistique et d'autres disciplines scientifiques, l'application des moindres carrés est essentielle pour estimer les paramètres des modèles variographiques, contribuant ainsi à une modélisation plus précise de la variabilité spatiale des phénomènes étudiés.

Dans le cas des moindres carrés non-linéaires, il s'agit de minimiser la fonction

$$g(x) = \frac{1}{2} r(x)' r(x) = \frac{1}{2} \sum_{i=1}^m r_i(x)^2 \quad \dots \dots \dots (4.1)$$

Avec:

$$r_i(x) = y_i - f(t_i, x) \quad i = 1, \dots, m$$

Où : $f(t_i, x)$ est une fonction non-linéaire (le modèle) avec t_i les variables indépendantes et $x \in \mathbb{R}_n$ le vecteur de paramètres à estimer, $y_i =$ les valeurs observées.

La description du modèle quadratique pour la minimisation de l'équation (4.1) passe par les les dérivées premières et secondes de $g(x)$.

La dérivée première s'écrit :

$$\nabla g(x) = \sum_{i=1}^m r_i(x) \cdot \nabla r_i(x) = r(x)' r(x) \quad \dots \dots \dots (4.2)$$

Où :

$$\nabla r(x) = \begin{bmatrix} \frac{\partial r_1(x)}{\partial x_1} & \dots & \frac{\partial r_1(x)}{\partial x_n} \\ \vdots & \dots & \vdots \\ \frac{\partial r_m(x)}{\partial x_1} & \dots & \frac{\partial r_m(x)}{\partial x_n} \end{bmatrix} \dots \dots \dots (4.3)$$

Cette matrice est appelée Jacobienne.

Le vecteur : $\nabla r_i(x) = \begin{bmatrix} \frac{\partial r_i(x)}{\partial x_1} \\ \vdots \\ \frac{\partial r_i(x)}{\partial x_n} \end{bmatrix}$ correspond à la ligne i de la matrice Jacobienne.

La dérivée seconde s'écrit :

$$\begin{aligned} \nabla^2 g(x) &= \sum_{i=1}^m (\nabla r_i(x) \cdot \nabla r_i(x)' + r_i(x) \cdot \nabla^2 r_i(x)) \\ &= \nabla \mathbf{r}(x)' \nabla \mathbf{r}(x) + S(x) \quad \dots \dots \dots (4.4) \end{aligned}$$

Où :

$$S(x) = \sum_{i=1}^m r_i(x) \cdot \nabla^2 r_i(x) \dots \dots \dots (4.5)$$

La méthode Gauss-Newton utilise une approximation de la matrice des dérivées secondes (4.4) en omettant le terme $S(x)$. Comme $S(x)$ est composé d'une somme de termes $r_i(x) \nabla^2 r_i(x)$, cette simplification se justifie dans une situation où les résidus $r_i(x)$ sont petits.

4. Validation croisée

La taille de l'ensemble de données joue un rôle important dans le choix de la méthode de validation croisée (voir chapitre 2, section 6). La méthode LOOCV peut être préférable pour des ensembles de données plus petits, tandis que K-fold peut être plus appropriée pour des ensembles de données plus importants, offrant un équilibre entre la robustesse de l'évaluation et le coût des calculs. LPO peut être une option intermédiaire lorsque l'on souhaite explorer différentes tailles de sous-ensembles exclus.

Pour des ensembles de données de taille modérée, K-fold Cross-Validation peut constituer un compromis raisonnable. Cette méthode divise l'ensemble de données en K parties (plis) et effectue l'évaluation K fois, en utilisant chaque pli comme ensemble de test une fois. K-fold est moins coûteuse en termes de calculs par rapport à LOO, tout en fournissant une estimation robuste de la performance du modèle.

Enfin, Leave-p-Out (LPO) Cross-Validation, peut être envisagée lorsque l'on souhaite une certaine flexibilité dans le choix du nombre spécifique de points à exclure pour chaque validation. Cette personnalisation offre un compromis entre les méthodes Leave-One-Out (LOO) et K-fold Cross-Validation. Toutefois, il est important de noter que cette flexibilité

s'accompagne d'un coût en termes de calcul. En effet, le nombre élevé d'itérations nécessaire pour évaluer différentes configurations peut rendre la méthode LPO plus exigeante en ressources de calcul, nécessitant ainsi des infrastructures matérielles performantes pour assurer une exécution efficace. Ainsi, bien que LPO offre une adaptabilité précieuse, il est crucial de prendre en compte ses exigences en termes de coût de calcul lors de sa sélection pour l'évaluation de modèles.

En considérant le cas de la présente étude qui implique un ensemble de données comprenant 61 échantillons, la méthode de validation croisée Leave-One-Out (LOOCV) se distingue comme un choix optimal. Avec un ensemble de données de cette taille, LOOCV offre une évaluation robuste du modèle, éliminant un seul point de données à la fois pour chaque itération de validation. Cette approche garantit une utilisation optimale des données, réduisant ainsi le risque de biais lors de l'évaluation du modèle. Bien que LOOCV puisse être plus coûteuse en termes de calculs, ce facteur devient moins contraignant avec des ensembles de données de taille modeste comme celui-ci, rendant la méthode LOOCV préférable pour assurer une performance fiable dans notre cas d'étude. Ainsi, en tenant compte de la taille de l'ensemble de données, de la robustesse de l'évaluation et des considérations de coût de calcul, LOOCV se positionne comme le choix privilégié pour la validation croisée dans notre cas spécifique.

5. Critères d'évaluation des modèles

Le modèle utilisé a été choisi en fonction de la valeur minimale de MSE (Mean-Square error). La concordance existante entre les valeurs issues des variogrammes expérimentaux et théoriques est évaluée à l'aide des indicateurs d'efficacité (Nash-Sutcliffe NSE) et de précision (RMSE, RSR) [79-81].

Le critère de Nash-Sutcliffe (NSE) est calculé comme suit :

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \dots \dots \dots (4.6)$$

Un moyen simple pour avoir la précision des modèles consiste à calculer le RMSE (root-mean-square error) [79, 80, 82]. Le RMSE représente les différences entre les valeurs prédites par un modèle et les valeurs réellement observées.

Le RMSE est calculé comme suit :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \dots \dots \dots (4.7)$$

Le RSR (Root mean square error-observations standard deviation ratio) est calculé par le rapport entre le RMSE et l'écart type des données mesurées [79, 80, 83]:

$$RSR = \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \dots \dots \dots (4.8)$$

Le RSR varie de la valeur optimale de 0, qui indique zéro RMSE ou bien une variation résiduelle et donc une simulation de modèle parfaite, à une grande valeur positive. Plus le RSR est bas, plus le RMSE est faible indiquant une meilleure correspondance entre les données observées et simulées.

6. Conclusion

La combinaison judicieuse de la méthode des moindres carrés, les critères de précision et du choix de la validation croisée LOOCV a jeté les bases d'une approche fiable et précise pour l'interpolation spatiale par krigeage.

La méthode des moindres carrés a été utilisé dans le but de garantir un ajustement précis des modèles variographiques. L'utilisation des critères de précision et d'efficacité des modèles, tels que le critère de Nash-Sutcliffe, a permis d'évaluer la qualité d'ajustement et d'orienter le choix vers des modèles variographiques offrant les meilleurs résultats.

La mise en œuvre de la validation croisée LOOCV a apporté une dimension supplémentaire à la sélection des modèles variographiques en évaluant leur performance sur l'ensemble des données. Cette approche, bien qu'il soit coûteux en termes de calculs, s'est avérée important pour garantir la robustesse des modèles retenus, surtout dans un contexte où la taille des données peut être modeste. Cette méthodologie orientée vers l'optimisation des ajustements et d'avoir une précision maximale dans l'estimation par krigeage ordinaire des réserves du gisement, contribuant ainsi à la prise de décision éclairée dans la gestion et l'exploitation du site.

Chapitre V : Evaluation des réserves, résultats et discussion

1. Introduction

Après avoir mis en œuvre l'approche géostatistique, il est essentiel de l'appliquer à un cas réel afin de tester la performance de l'approche suivie. Le gisement de Bled El-Hadba, situé au Nord-Est de l'Algérie, a été choisi pour étudier la variabilité spatiale des teneurs en P_2O_5 et des épaisseurs de la couche principale du gisement, qui correspond à la couche intermédiaire.

Les données géochimiques, provenant de 61 forages effectués par l'ORGM lors des campagnes d'exploration de ce gisement, représentent une ressource principale pour cette étude. Cependant, la répartition irrégulière de ces échantillons (sondages) présente des défis significatifs. Les distances entre les points d'échantillonnage varient considérablement, atteignant parfois jusqu'à 500 mètres dans certaines zones. Cette hétérogénéité dans la distribution spatiale des échantillons soulève des questions quant à la représentativité des données, nécessitant une approche minutieuse pour garantir une interprétation précise.

2. Application de l'approche géostatistique

L'application de l'approche ajustement-validation-krigeage sur le gisement de Bled El Hadba représente une étape essentielle dans notre démarche. Cette méthodologie englobe un ensemble intégré de processus visant à optimiser l'estimation des réserves minières. Tout d'abord, le volet d'ajustement s'attache à choisir le modèle variographique le plus adéquat, influant directement sur la qualité de l'estimation. Cette étape est réalisée grâce à une combinaison de méthodes des moindres carrés et de validation croisée Leave-One-Out (LOOCV), garantissant une adéquation optimale du modèle aux données géochimiques du gisement.

Une fois l'étape de validation (LOOCV) des modèles variographique terminée, l'étape suivante consiste à déterminer le support géométrique sur lequel le krigeage sera effectué. La précision voulue est due non seulement à la qualité du matériel à utiliser pour la modélisation, mais aussi au nombre de points à estimer. Cela nous oblige à prendre en compte le pas (inter-distance) de la grille à choisir. Dans notre cas, nous avons considéré une grille régulière horizontale en deux dimensions (2D) qui s'étend sur toute la superficie du gisement de Bled El-Hadba. La maille unitaire de cette grille a pour dimensions de $5 \times 5 \text{ m}^2$, la grille ainsi construite comprend au total 581×821 points.

2.1. Analyse statistique de données du gisement de Bled El-Hadba

Avant d'entreprendre une étude géostatistique, une phase préliminaire de l'analyse statistique est essentielle. Cette étape vise à comprendre la distribution, la variabilité et les relations entre les données géospatiales, fournissant ainsi une base solide pour l'application de méthodes plus avancées. L'analyse statistique initiale permet d'explorer les caractéristiques fondamentales des variables d'intérêt, d'identifier les tendances générales et de détecter d'éventuelles anomalies. Ces informations préliminaires sont cruciales pour orienter le choix des méthodes géostatistiques les plus appropriées, garantissant ainsi la pertinence et la fiabilité des résultats obtenus dans le cadre de l'étude géostatistique ultérieure.

Dans le cadre de cette analyse, nous examinons les épaisseurs des couches minéralisées et les teneurs équivalentes en P_2O_5 en tant que paramètres régionalisés, dans le but d'évaluer les réserves de phosphate dans le gisement de Bled El Hadba.

2.1.1. Statistique descriptive et histogramme

De nombreuses variables dans les ensembles de données des sciences de la terre, telles que les teneurs et les épaisseurs des couches minéralisées, présentent des distributions qui s'éloignent de la normalité. Il est fréquent d'observer de nombreuses valeurs relativement faibles et quelques-unes très élevées. La distribution normale s'avère souvent inappropriée comme modèle pour ce type de distribution asymétrique, et une distribution log-normale peut parfois constituer une alternative pertinente. Une variable aléatoire (X) suit une loi log-normale lorsque son logarithme suit une loi normale. De la même manière que dans le tracé de probabilité normale, les fréquences cumulées formeront une ligne droite si les valeurs données suivent une distribution log-normale.

Dans ce qui suit, nous présentons une étude statistique et variographique approfondie de la couche médiane, qui, du point de vue de l'exploitation, présente des caractéristiques plus favorables en termes de teneur et d'épaisseur.

a) Épaisseurs de la couche médiane

L'analyse statistique des données de terrain recueillies lors de la phase d'exploration indique que la couche médiane présente une meilleure qualité en termes de P_2O_5 , du point de vue de l'exploitation, par rapport aux autres couches phosphatées. L'épaisseur maximale de cette couche peut atteindre 29.75 m, avec une moyenne de 18.51 m.

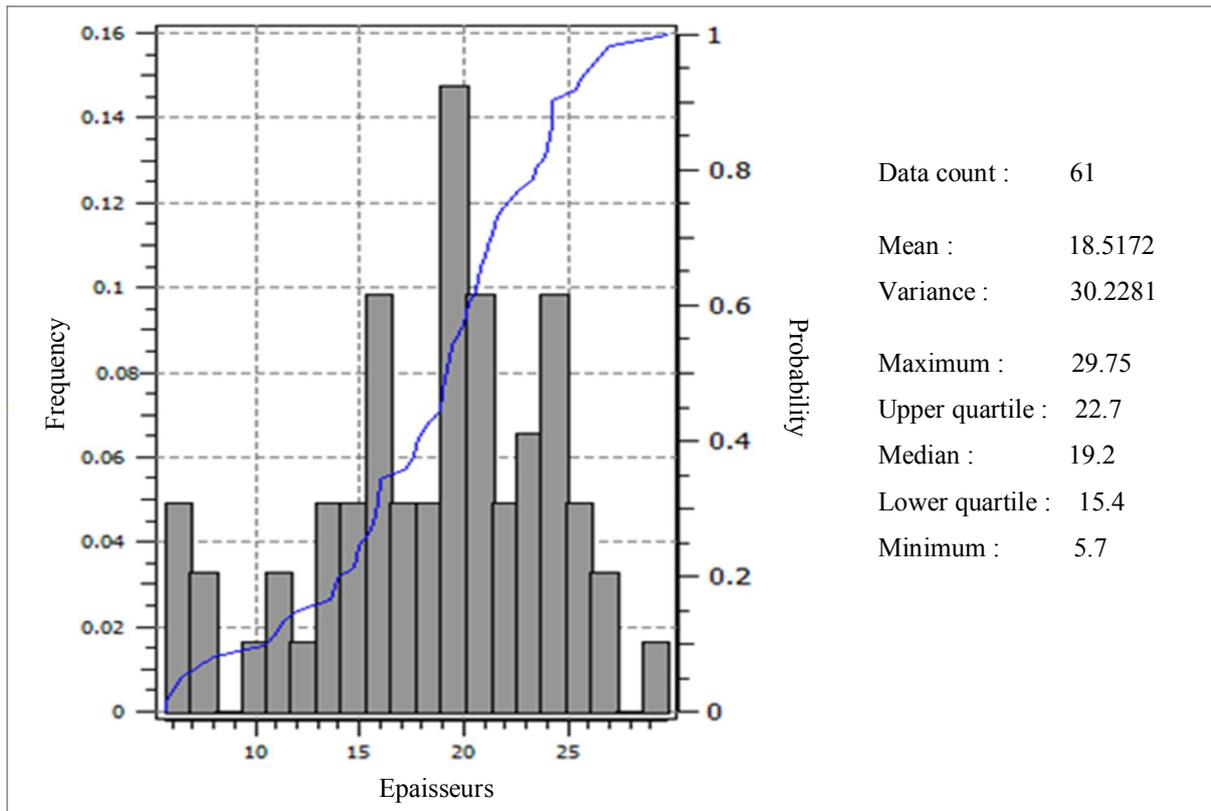


Figure 5.1. Histogramme de la puissance de la couche médiane.

Environ trois quarts des sondages de la couche médiane présente des épaisseurs supérieures à 22.7 m, démontrant une variance de 30.22 qui témoigne de l'hétérogénéité de la distribution de cette couche phosphatée (Figure 5.1). L'épaisseur minimale de cette couche est de 5.7 m.

b) Teneurs de de la couche médiane

La concentration moyenne en P_2O_5 des carottes de forage s'élève à 24.67 %, avec une valeur maximale pouvant atteindre 27.9 % et une valeur minimale de 18.68 % (Figure 5.2). La variance, évaluée à 4.04, témoigne de l'homogénéité de la distribution des concentrations dans cette couche.

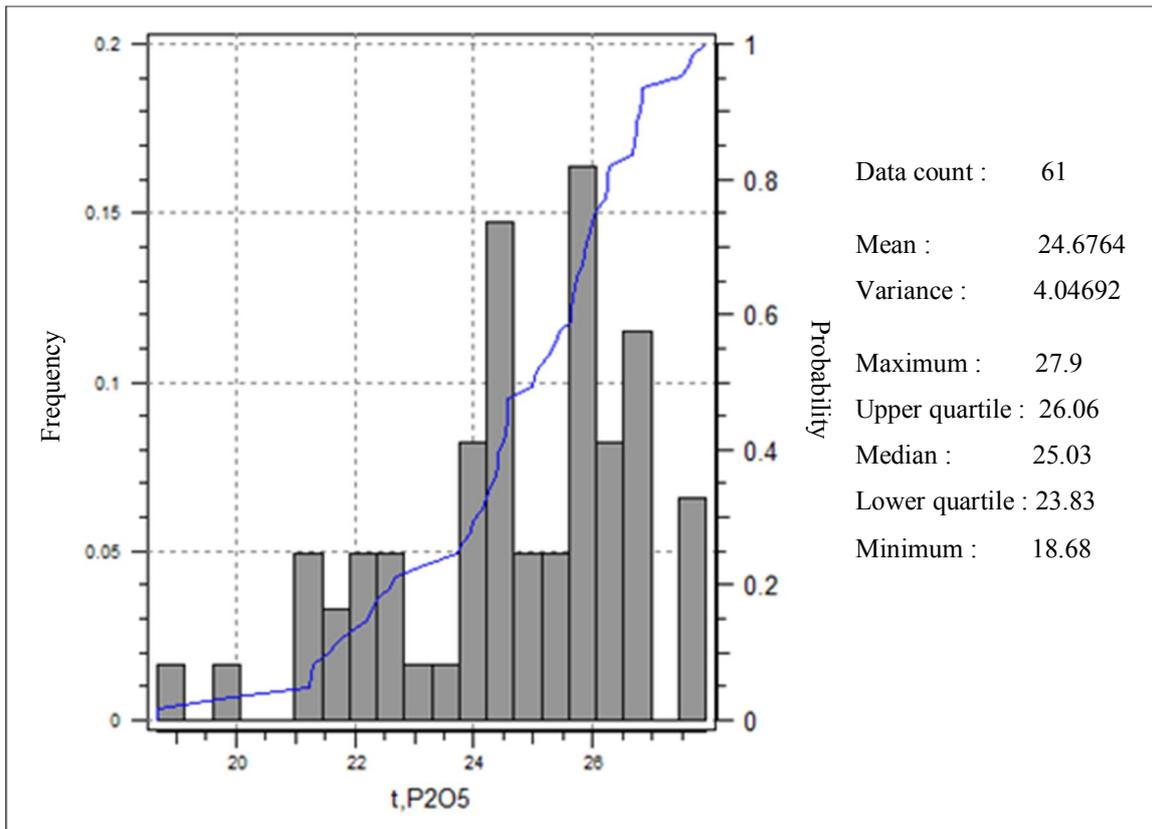


Figure 5.2. Histogramme illustrant les teneurs en P₂O₅ de la couche médiane.

2.2. Analyse géostatistiques du gisement de Bled El-Hadba

La mise en application de l'approche géostatistique sur le gisement de Bled-El-Hadba nous a permis d'estimer les réserves en phosphate, sur la base des données géochimiques disponibles. Nous présentons ci-après les résultats de calcul des teneurs et épaisseurs de la zone d'étude, sous forme de modèles variographiques et de cartes d'estimation.

2.2.1. Estimation de la puissance de la couche minéralisé

Les variogrammes théoriques et expérimentaux issus de la procédure de validation LOOCV sont présentés dans la Figure suivante.

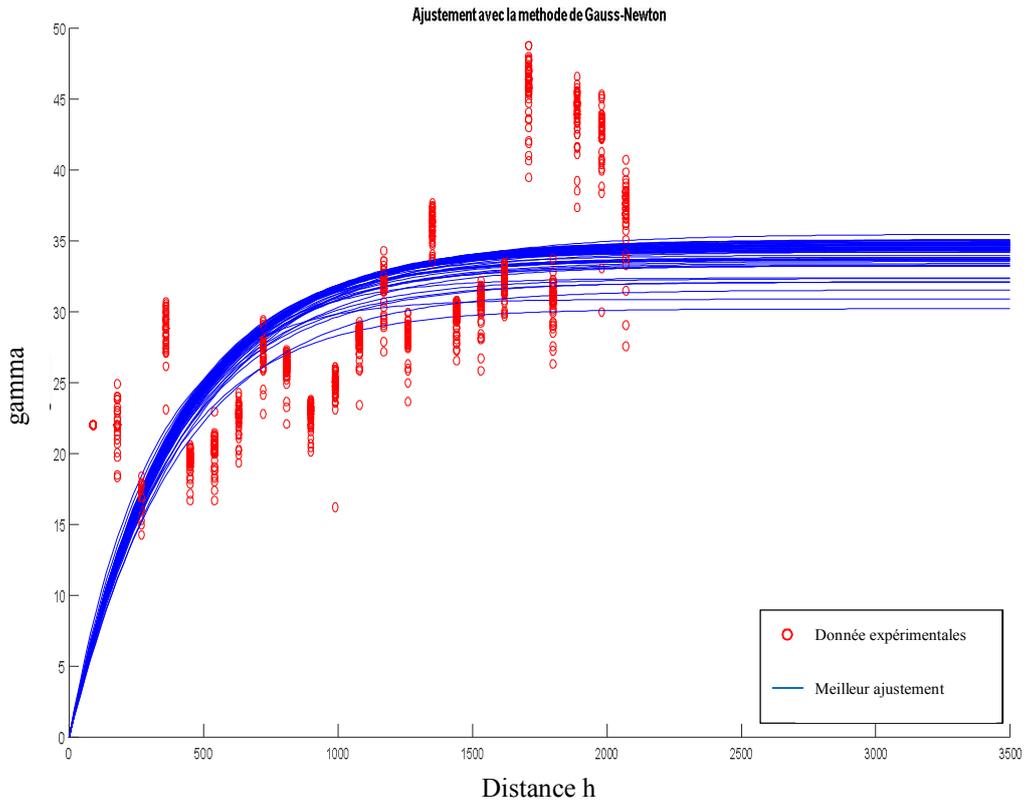


Figure 5.3. Variogrammes de validation des épaisseurs de la couche principale

Les variogrammes expérimentaux ont mis en évidence l'influence de certains échantillons dans le calcul des variances (l'écart peut atteindre 15). L'ajustement des variogrammes théoriques aux données expérimentales montre une bonne adéquation suivant le modèle exponentiel employé (Figure 5.3).

Après avoir calculé les variogrammes et appliqué la procédure de validation, il est question de choisir le modèle le mieux représentatif (Figure 5.4).

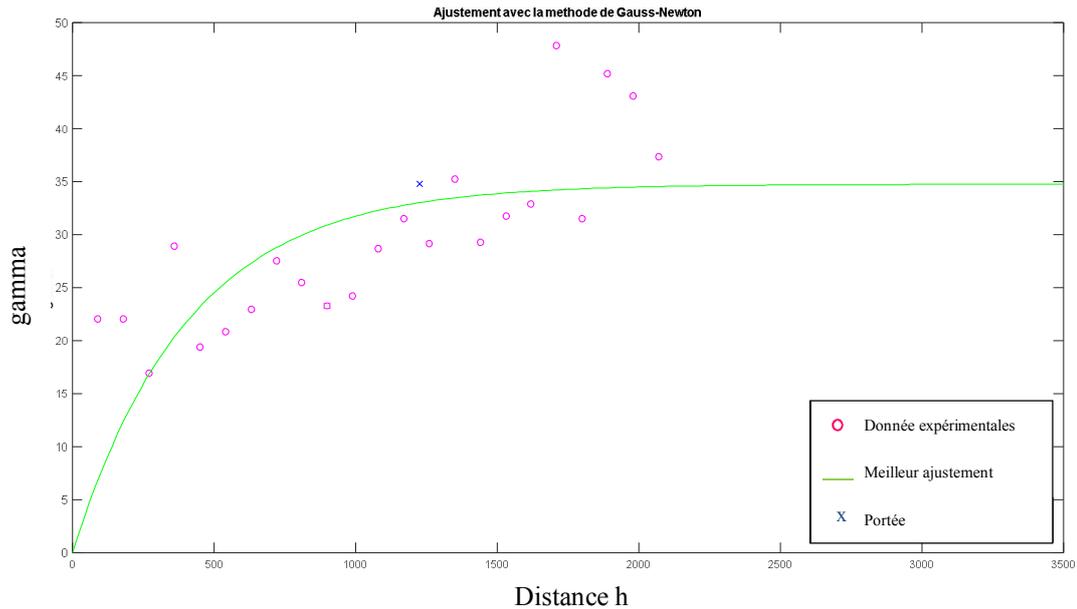


Figure 5.4. Modèle variographique retenu des épaisseurs de la couche principale.

Ce variogramme atteint un palier de $C(0) = 35.6677$ pour une portée maximale de $a = 1380$ m. le modèle choisi montre une continuité près de l'origine illustrée par un effet de pépité nul ($C_0 = 0$).

L'adéquation entre modèles théoriques et expérimentaux est quantifiée par des critères d'évaluation de la qualité d'ajustement (MSE, RMSE, Nash–Sutcliffe, RSR). Les résultats sont présentés dans le tableau ci-dessous.

Tableau 5.1 : Critères d'évaluation de la qualité d'ajustement

		MSE	RMSE	Nash–Sutcliffe	RSR
Modèles théoriques	MIN	0.036	6.226	0.318	0.514
	MAX	182.335	8.851	0.694	0.845
	MOY	23.037	8.205	0.570	0.646
Modèle retenu	LOOCV	0.036	6.599	0.572	0.642

La concordance entre modèles théoriques et expérimentaux est évaluée par les critères RMSE, Nash–Sutcliffe et RSR qui affichent des valeurs admissibles (Tableau 1).

L'erreur moyenne quadratique (MSE) est déterminante dans le choix du meilleur modèle de validation. Par conséquent, le modèle choisi affiche une MSE minimale de 0.036.

Le modèle théorique choisi est utilisé dans une procédure de Krigeage Ordinaire, sous une grille régulière de 5 x 5 m, afin d'évaluer la dispersion des épaisseurs dans la zone d'étude (Figure 5.5).

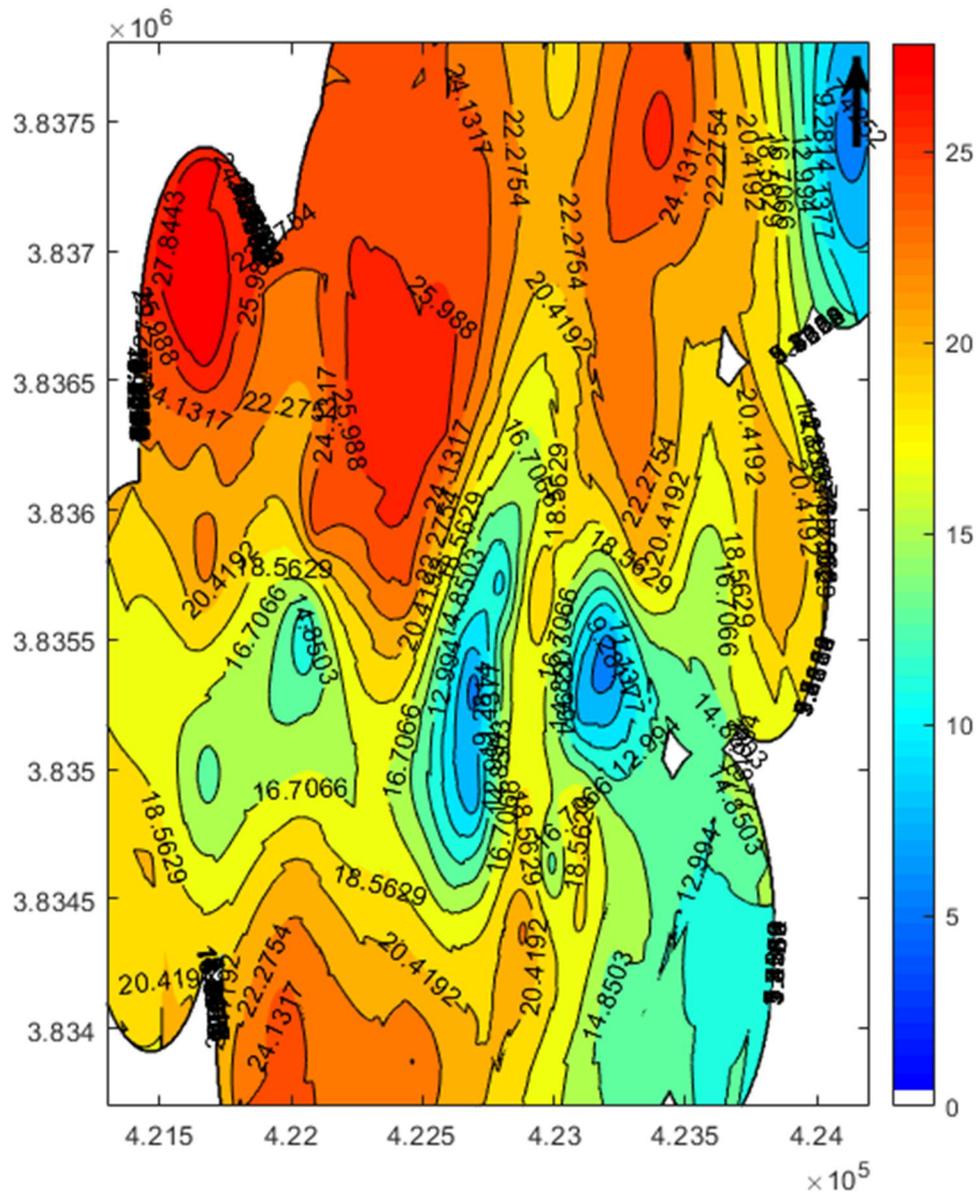


Figure 5.5. Carte de dispersion des épaisseurs de la couche principale

Cette carte a mis en évidence la distribution des épaisseurs dans le gisement, qui varient entre 8 et 25 mètres. Afin d'évaluer la qualité des résultats d'estimation, il est nécessaire de calculer la variance d'estimation (Figure 5.6).

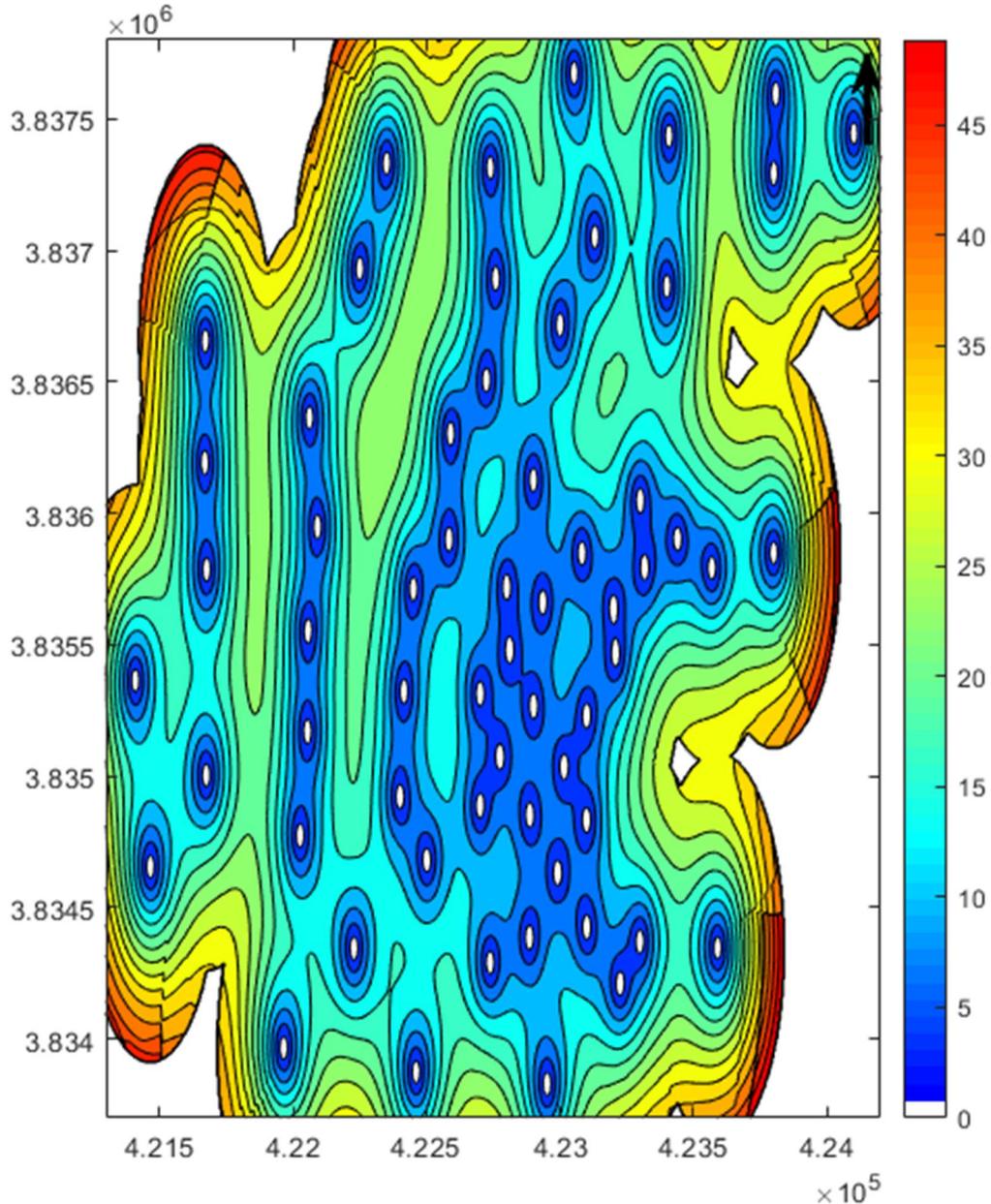


Figure 5.6. Carte de la variance d'estimation des épaisseurs de la couche principale.

La zone échantillonnée affiche une qualité d'estimation admissible (variance < 35.66) dont la meilleure qualité est constatée dans la zone centrale du gisement (variance < 12), caractérisée par une densité d'échantillonnage plus importante. Une relation de proportionnalité est établie entre la variance d'estimation et le nombre d'échantillons.

2.2.2. Estimation des teneurs

L'application de l'approche géostatistique dans l'étude de la variabilité spatiale des teneurs a donné des résultats similaires à ceux des épaisseurs, correspondant aux : variogrammes de validation (Figure 5.7), modèle variographique retenu (Figure 5.8), carte de

dispersion des teneurs (Figure 5.9), carte de la variance d'estimation des teneurs (Figure 5.10).

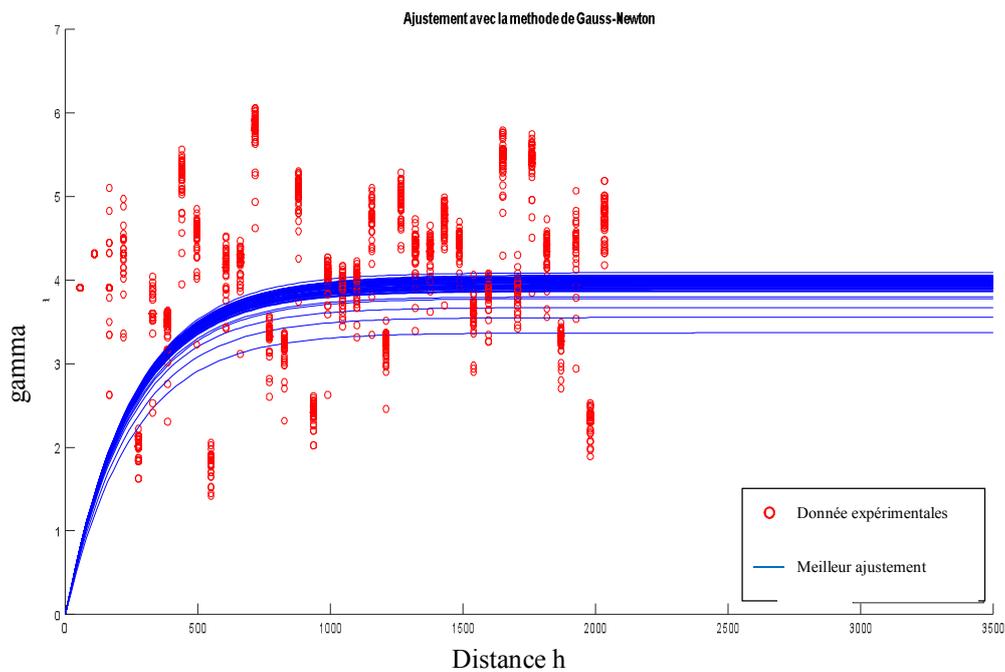


Figure 5.7. Variogrammes de validation des teneurs de la couche principale

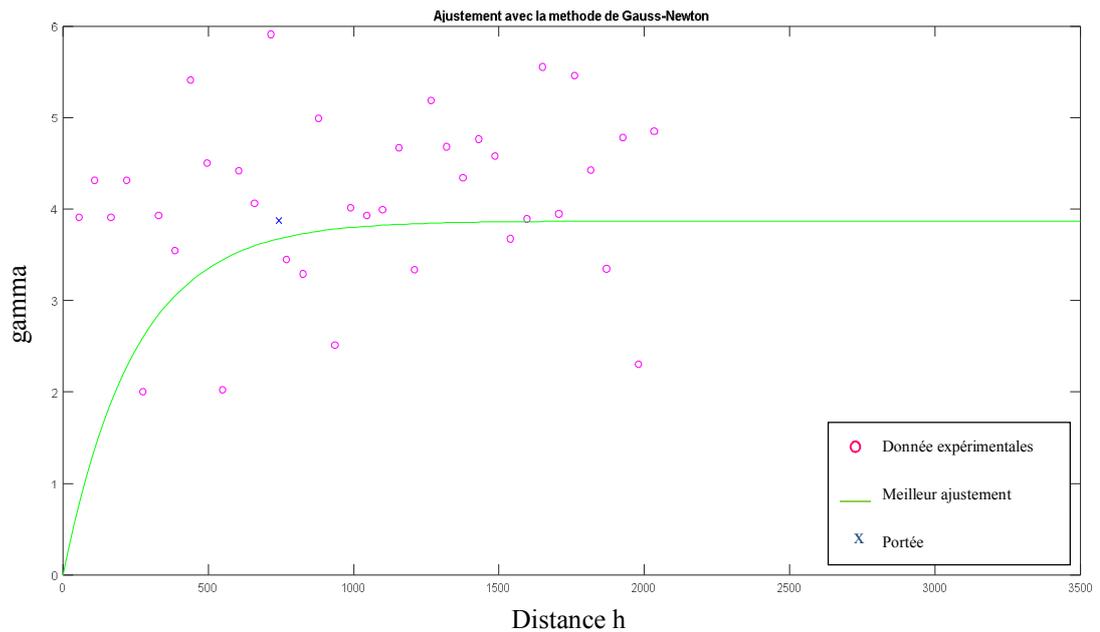


Figure 5.8. Modèle variographique retenu des teneurs de la couche principale.

La valeur de la portée correspondant au variogramme des teneurs retenu a considérablement diminué (742 m) comparativement à celle du variogramme des épaisseurs. Le palier de ce variogramme atteint $C_0=3.87$.

Les critères d'évaluation de la qualité d'ajustement relatif aux teneurs aboutissent à des valeurs admissibles (tableau 5.2).

Tableau 5.2: Critères d'évaluation de la qualité d'ajustement

		MSE	RMSE	Nash–Sutcliffe	RSR
Modèles théoriques	MIN	1.54×10^{-7}	34.791	0.084	0.883
	MAX	214.554	35.883	0.281	1.076
	MOY	23.053	34.900	0.199	0.990
Modèle retenu	LOOCV	1.54×10^{-7}	34.953	0.248	0.989

Dans un premier temps, nous élaborons une cartographie exhaustive représentant la distribution spatiale des teneurs en P_2O_5 en utilisant la grille générée par la méthode du krigeage ordinaire (Figure 5.9).

La figure 5.10, illustre la variance sur la grille résultant du processus de krigeage. Cette représentation nous offre un moyen d'évaluer la qualité des informations utilisées dans la procédure de calcul, en identifiant les zones où la variance est élevée, notamment autour des limites. La variance issue du processus de krigeage est exploitée par la suite dans la classification des ressources/réserves, conformément à diverses catégories établies par la norme JORC [13, 26, 84]. Ce paramètre de variance joue un rôle important dans l'évaluation de la fiabilité des données géostatistiques, permettant ainsi de catégoriser de manière plus précise les ressources et réserves minérales en fonction de leur niveau d'incertitude associé.

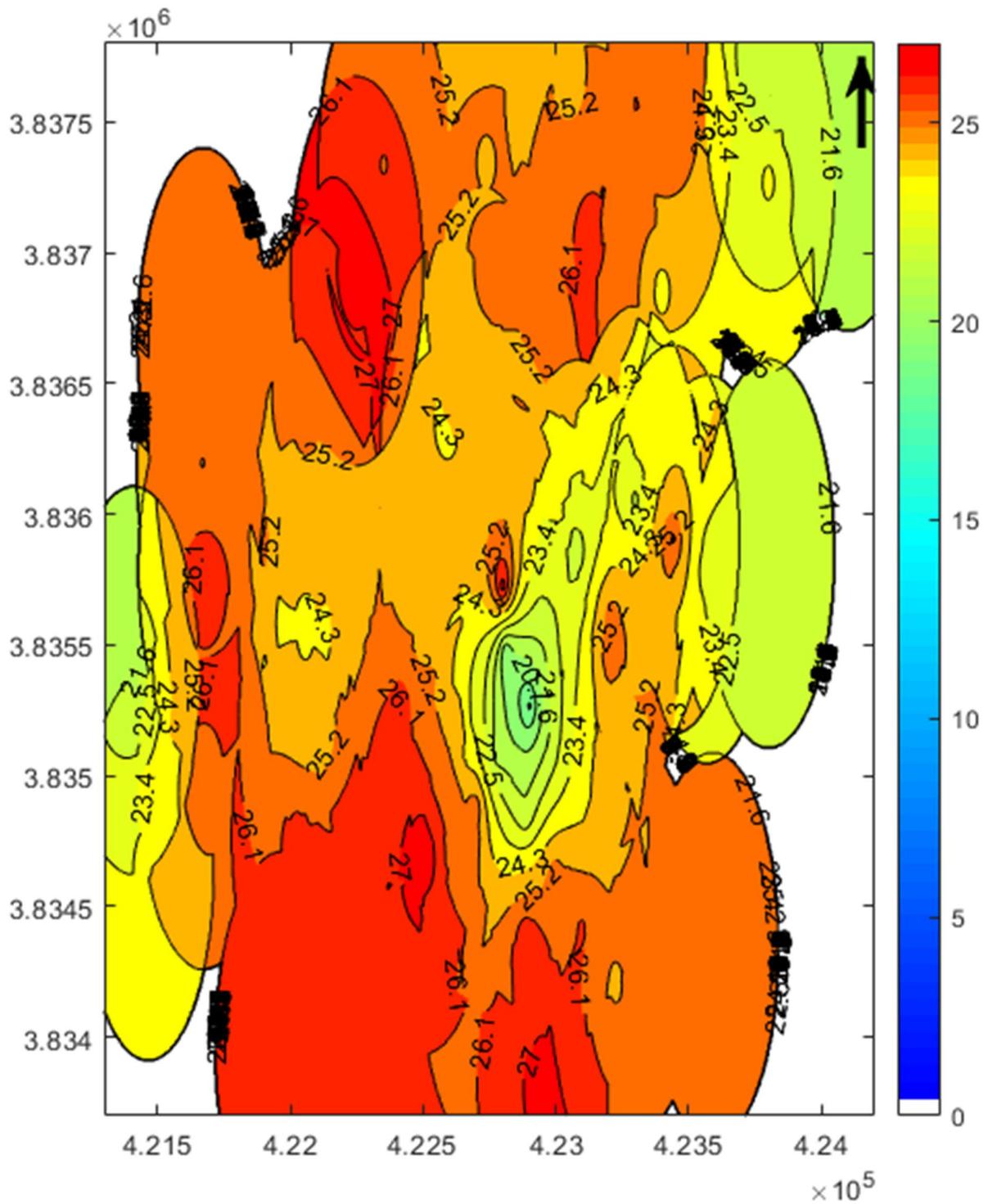


Figure 5.9. Carte de dispersion des teneurs de la couche principale.

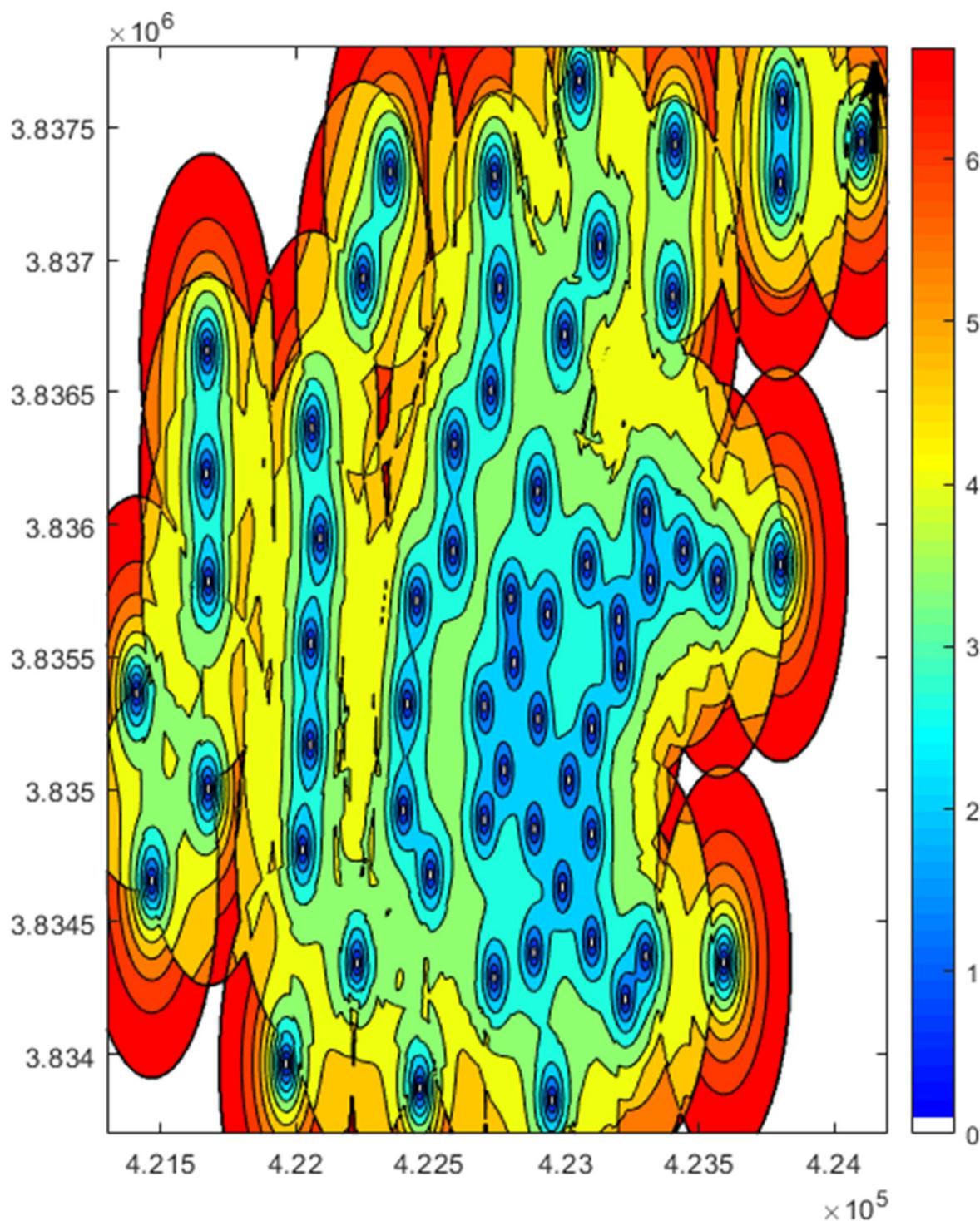


Figure 5.10. Carte de la variance d'estimation des teneurs de la couche principale.

2.3. Traitement statistique des résultats de l'estimation

Après avoir réalisé les estimations par krigeage ordinaire des variables régionalisées, notamment les teneurs en phosphate (P_2O_5) de la couche principale et sa puissance, une phase importante d'analyse statistique s'impose. Cette analyse vise à catégoriser les résultats en fonction de leur niveau de confiance et de précision obtenus lors de l'étape d'estimation. Cette

section détaillera le processus d'analyse statistique, offrant ainsi une compréhension approfondie de la qualité des données estimées.

2.3.1. Epaisseurs de la couche principale

Dans ce qui suit, nous exposons l'histogramme relatif à l'épaisseur de la couche principale de phosphate, accompagné d'une courbe représentant la fonction de densité cumulative (CDF) (figure 5.11).

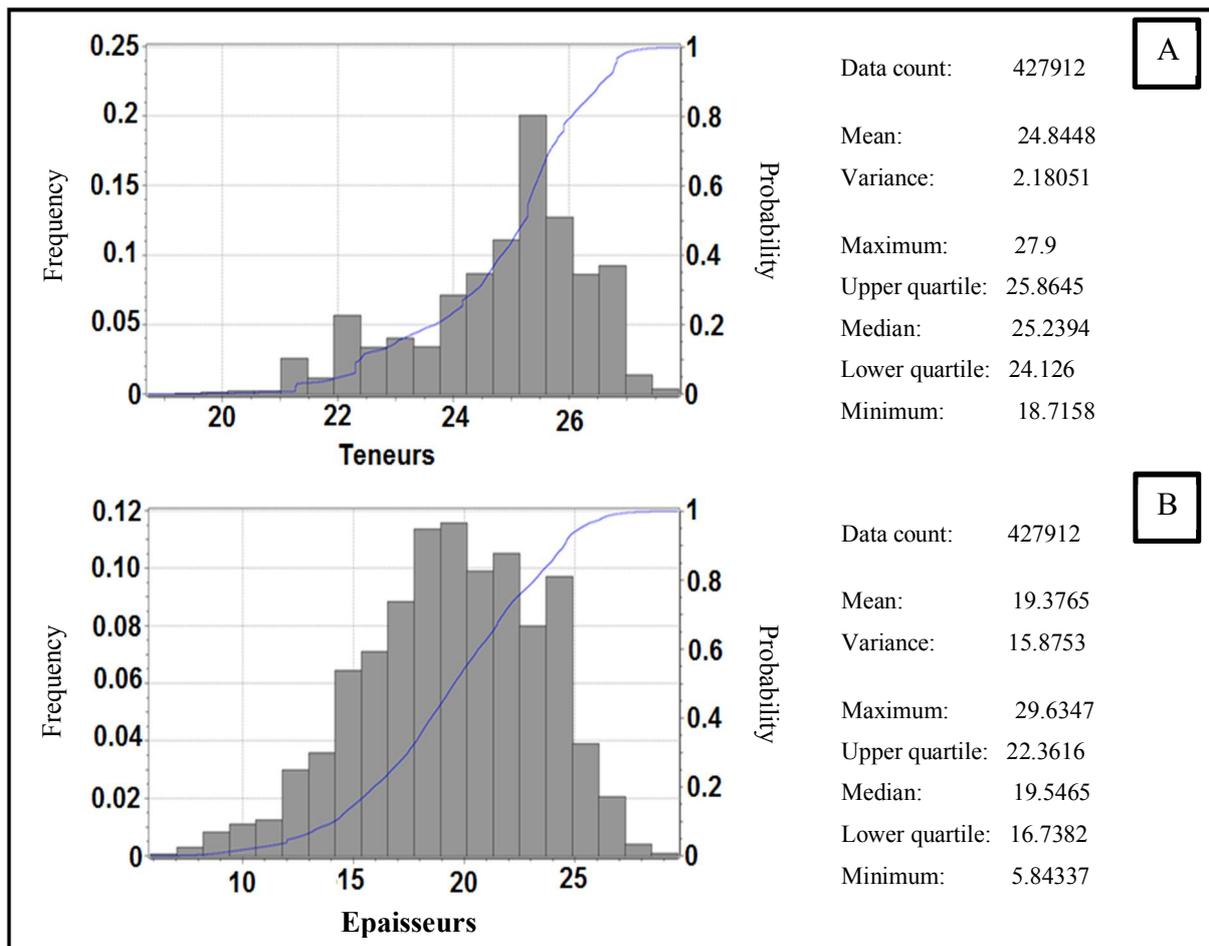


Figure 5.11. Histogramme et courbe de fonction de densité cumulative des : (A) teneurs en phosphate (B) épaisseurs de la couche médiane.

L'analyse de ce graphique met en évidence une puissance moyenne de 19.37 m et une médiane de 19.54 m, accompagnées d'une moyenne des teneurs de 24.84% et d'une médiane de 25.86%, illustrant une remarquable proximité entre ces deux mesures centrales. Cela indique une stabilité dans la mesure de la tendance centrale des données. La convergence étroite entre la médiane et la moyenne de la distribution des épaisseurs et des teneurs de la couche principale de phosphate au niveau du gisement de Bled El-Hadba suggère que ces

valeurs sont réparties de manière équilibrée de part et d'autre de la médiane, signalant ainsi une symétrie dans la distribution. Cependant, la variance des teneurs de la couche étudiée est de 2,18, tandis que celle de la puissance est nettement plus élevée, atteignant 15,87 (figure 5.11). Ces résultats peuvent être expliqués par la distribution géométrique de la couche phosphatée, influencée par les failles existantes au niveau de gisement de Bled El-Hadba.

2.3.2. Variance de krigeage des variables étudiées

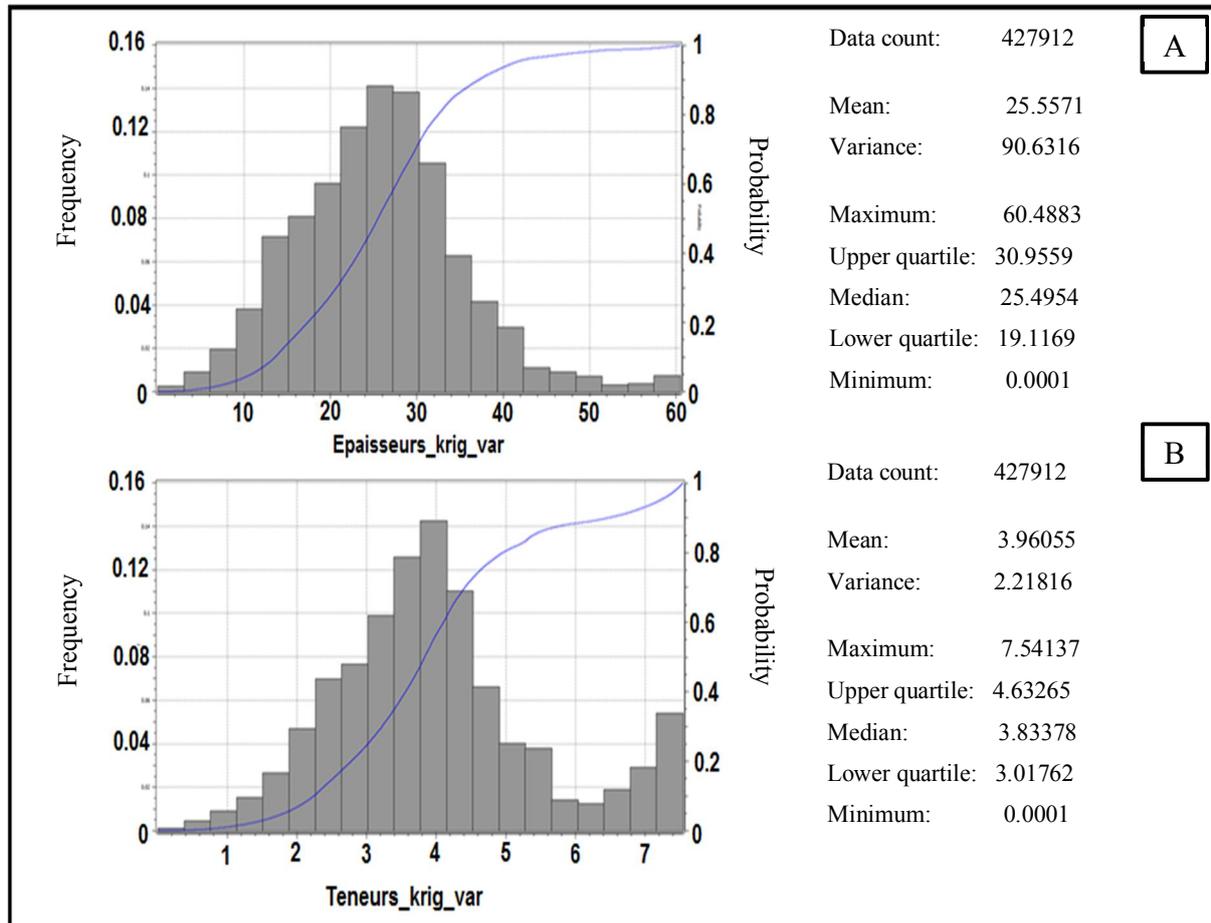


Figure 5.12. Histogramme et graphe de fonction de densité cumulative des variances de krigeage de l'estimation (A) des épaisseurs et (B) des teneurs de la couche médiane.

La progression ultérieure des courbes de la fonction de densité cumulative vers une forme gaussienne indique une concentration qui se stabilise autour de la moyenne, manifestant une dispersion relativement équilibrée des valeurs de part et d'autre. Cependant, les variances les plus élevées sont observées au niveau de la distribution des épaisseurs des couches minéralisées, atteignant un maximum de 60,4, tandis que les variances de la distribution des teneurs atteignent un pic à 7,54 (figure 5.12). Ces résultats s'expliquent par l'estimation

extrapolée des épaisseurs et des teneurs, avec une large plage d'estimation dans la distribution de l'épaisseur de la couche minéralisée.

Sur la base de ces résultats, il devient impératif de procéder à une classification en fonction du degré de certitude de ces estimations, afin d'apporter une clarification approfondie et une meilleure compréhension des implications de ces données.

2.1. Réserves de la couche médiane du gisement de Bled El-Hadba :

DMT (2016) [76] a évalué les réserves en P_2O_5 (prouvées + probables) pour cette couche à environ 464 millions de tonnes, avec une teneur de coupure fixée à 20%. Cette estimation présente une différence significative par rapport à celle de l'ORGM 2015[75], qui a évalué la même couche à environ 367 millions de tonnes pour des teneurs de coupure de 16%.

Selon la littérature, il n'existe pas de classification standardisée des réserves, mais diverses études se fondent sur la variance du krigeage. Généralement, on attribue les valeurs entre 30% et 40% de la variance maximale du krigeage à la catégorie prouvée, tandis que la catégorie probable englobe jusqu'à 90% de cette variance maximale. Dans le cadre de cette étude, les limites des classes pour les différentes catégories de réserves sont définies en utilisant de 40% et 90% du palier de variogramme des épaisseurs (35.6677). Ces pourcentages seront utilisés respectivement pour présenter les catégories des ressources mesurée/indiqué et des réserves prouvées/probables.

Au cours de cette étude, les ressources et réserves en phosphate sont déterminées à partir d'une analyse exploratoire, en utilisant les données issues de la grille de krigeage. La figure 5.13 illustre une courbe de densité cumulative des ressources en phosphate.

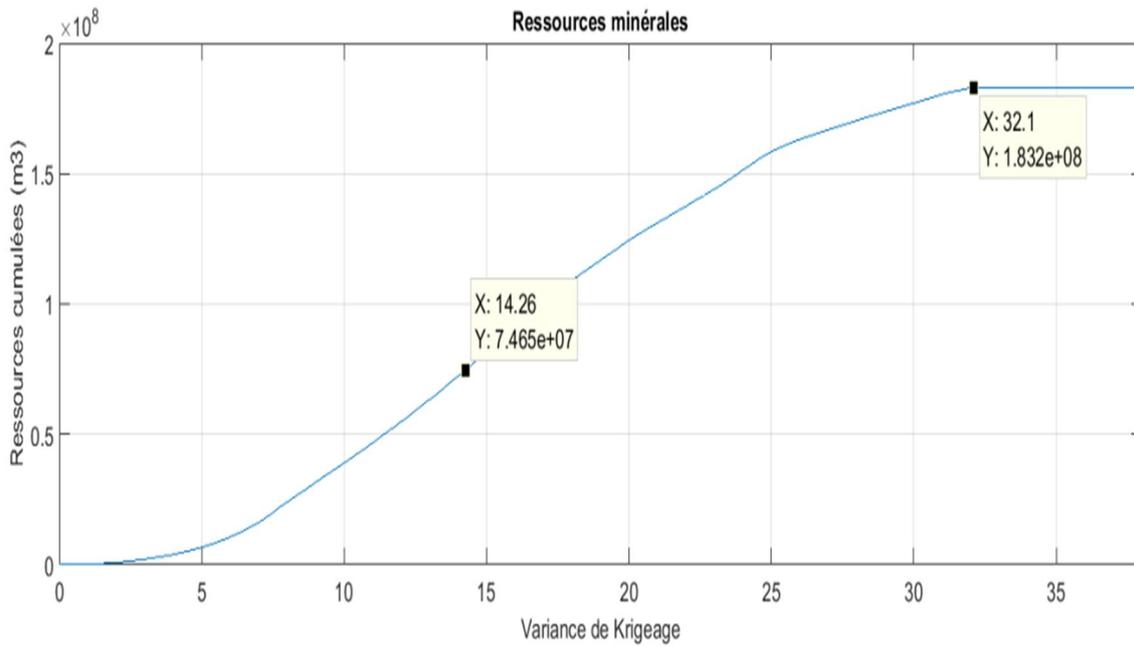


Figure 5.13. Ressources de la couche médiane de gisement de Bled El-Hadba.

Le volume total de la couche, qui englobe les volumes mesuré et indiqué (figure 6 en annexe 5), est évalué à 183,2 millions de mètres cubes, dont 74,65 millions de mètres cubes représentent les ressources mesurées (figure 5 en annexe 5), comme le montre la figure 5.13. Cette estimation équivaut à 403,04 millions de tonnes au total, dont 164,23 millions de tonnes sont classées comme ressources mesurées (figure 5.13).

2.1.1. Classification des ressources/réserves de la couche médiane

La classification des ressources peut être effectuée en fonction de différents critères tels que la distance entre les forages, la variance de krigeage et les modèles d'incertitude fournis par l'utilisation de simulations géostatistiques [84]. A la différence des l'étude antérieures (ORGM 2015, DMT 2016) [75, 76], qui ont utilisé la distance entre les forages comme source de classification. Cependant, les résultats présentés au cours de cette étude utilisent la variance de krigeage comme facteur limitatif pour évaluer l'incertitude d'estimation d'une part, et d'autre part, classifient les ressources/réserves de cette couche de phosphate.

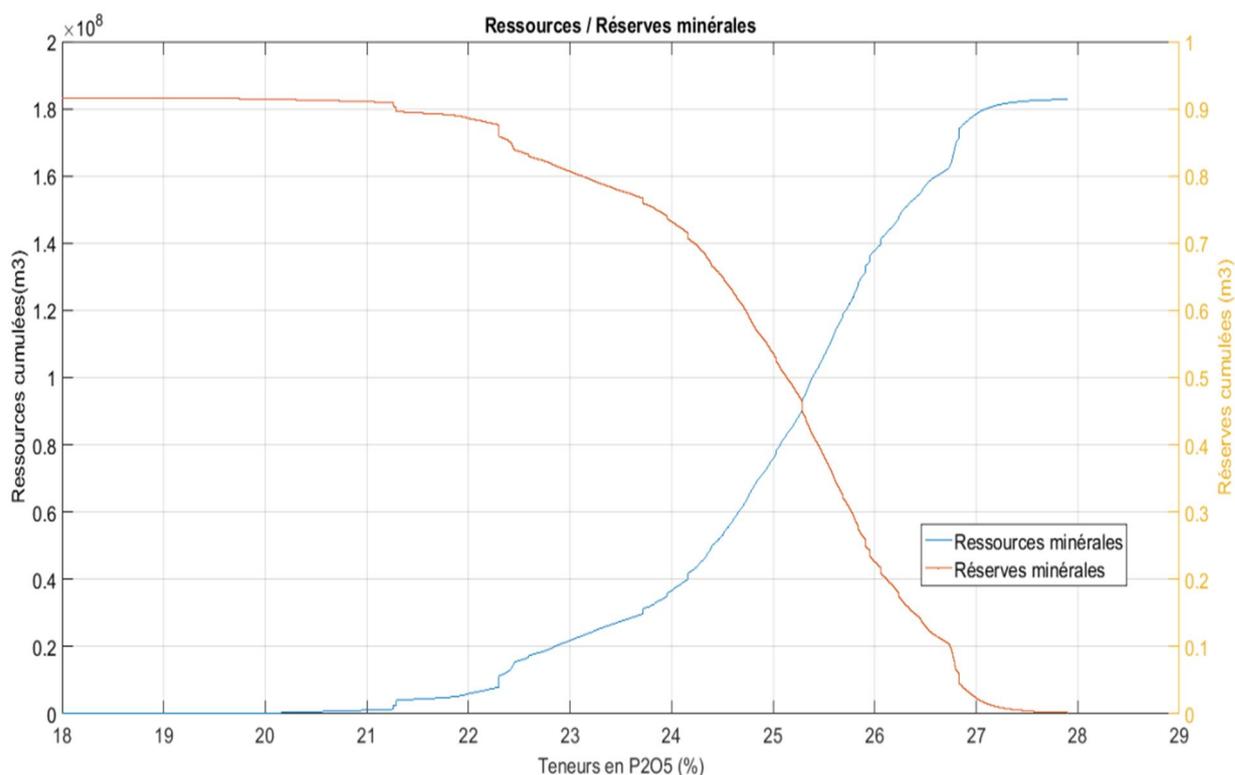


Figure 5.14. Ressources et réserves en P_2O_5 de la couche médiane de gisement de Bled El-Hadba.

Le graphique présenté dans la figure 5.14 synthétise les réserves en phosphate de la couche principale du gisement de Bled El-Hadba. Les réserves sont détaillées dans le tableau 3, en précisant leurs quantités en fonction des teneurs de coupure.

Tableau 5.3 : Réserves prouvées et probables en P_2O_5 de la couche médiane du gisement de Bled El-Hadba.

Teneur de coupure de P_2O_5 (%)	Réserves (million de m^3)	Masse volumique (t/m^3)	Réserves (million de tonnes)	P_2O_5 (%)
19	183.1	2.2	402.82	24.85
20	182.9		402.38	24.85
21	182.1		400.62	24.87
22	177.2		389.4	25.02
23	161.4		355.08	25.31
24	146.4		322.08	25.52
25	107		235.4	25.87

En adoptant une teneur de coupure de 20%, les réserves sont estimées à environ 402,38 millions de tonnes, créant ainsi une mine géante avec une perspective d'exploitation à long terme. Cette évaluation encourage les entreprises minières à s'engager dans l'exploitation de ce gisement de phosphate, stimulant par conséquent les investissements dans des projets de transformation dans la région du Nord-Est algérien. Cette abondance de ressources promet non seulement une durée prolongée d'activités minières, mais également offre des opportunités stratégiques pour le développement économique régional grâce à la transformation des phosphates.

3. Discussion

L'approche géostatistique proposé vise à optimiser l'estimation des réserves en examinant attentivement le choix du modèle variographique. Cette méthodologie est concrétisée à travers un code de calcul développé sous MATLAB, intégrant les différentes phases d'ajustement, de validation, et de krigeage. Lors du krigeage, les modèles sont méticuleusement choisis grâce à une procédure de validation croisée (LOOCV), renforçant ainsi la fiabilité des résultats obtenus par le krigeage ordinaire (comme indique dans les Figures 5. (5, 6, 9 et 10)). L'approche de validation croisée Leave-One-Out (LOOCV) génère une multitude de variogrammes (voir la Figure 5.3 et 5.7) en fonction du nombre d'échantillons utilisés, permettant une analyse approfondie de la variabilité spatiale des épaisseurs et des teneurs de la couche minéralisée.

La méthode des moindres carrés a été appliquée lors de l'ajustement des modèles théoriques aux variogrammes expérimentaux. L'évaluation de la qualité de cet ajustement a été réalisée à l'aide de plusieurs critères tels que Nash–Sutcliffe, RMSE, et RSR. Ces critères, utilisés pendant l'ajustement, permettent d'évaluer la qualité de l'ajustement ainsi que l'erreur d'estimation. Le critère NSE fournit des indications sur la qualité de l'ajustement, tandis que le RMSE et le RSR quantifient l'erreur d'estimation. Une combinaison de ces critères offre une évaluation globale de la qualité du modèle [85]. Cela suggère que le critère NSE peut être non seulement un indicateur de la qualité des modèles utilisés dans cette étude, mais également un outil pour améliorer la qualité dans l'industrie minière.

Les résultats des tableaux 5.1 et 5.2 indiquent que les modèles utilisés dans cette étude démontrent une fiabilité dans la modélisation des variogrammes d'épaisseur et des teneurs en phosphate. Les valeurs médianes de NSE s'élèvent à 0,570 pour les épaisseurs et à 0,199 pour les teneurs, avec des valeurs de RSR respectives de 0,646 et 0,990. Le choix du meilleur

variogramme est fondé sur la minimisation de l'erreur quadratique des variogrammes d'épaisseur et de teneur.

Les critères d'efficacité et de précision affichent également des valeurs acceptables, indiquant une qualité d'ajustement satisfaisante des données expérimentales aux modèles théoriques. Par conséquent, la validation croisée leave-one-out peut être utilisée pour estimer et comparer les performances des modèles. Ceci indique que les critères utilisés pour évaluer la qualité d'ajustement des données expérimentales aux modèles théoriques sont satisfaisants. L'erreur moyenne quadratique (MSE) des variogrammes d'épaisseur et de teneur est de 0,036 et $1,54 \times 10^{-7}$, respectivement, ce qui indique que les modèles s'ajustent bien aux données (Figure 5.4 et 5.8). De plus, les valeurs de NSE (efficacité Nash-Sutcliffe) et RMSE (erreur quadratique moyenne) de 0,572 et 6,599 respectivement indiquent que les modèles fournissent des estimations raisonnables des données observées.

L'utilisation de la validation croisée leave-one-out (LOOCV) dans cette étude est appropriée car l'ensemble de données utilisé est relativement petit, avec seulement 61 échantillons. LOOCV est une méthode puissante pour la sélection de modèles, car elle est indépendante du modèle sous-jacent et fournit une estimation sans biais de la performance du modèle [86]. Cependant, elle peut être coûteuse en termes de calculs pour des grands ensembles de données en raison de la nécessité de refaire le modèle pour chaque échantillon de manière répétée. Selon [87], LOOCV est privilégiée pour un nombre limité de données, et dans ce cas, elle offre une meilleure solution avec un coût temporel minimal.

Les résultats obtenus à partir de la phase de krigeage ordinaire sont considérés comme représentatifs et sont soutenus par l'approche géostatistique et la procédure de validation croisée LOOCV (Figure 5. (3, 4, 7 et 8)). Le recours à ces méthodes renforce la confiance en la précision et la fiabilité des résultats obtenus.

Cette validation peut être considérée sous deux perspectives. La première, utilisée dans la communauté informatique, vise à améliorer la qualité du modèle ajusté. La deuxième perspective est pratique, où le modèle validé peut être utilisé pour prendre de meilleures décisions en matière de gestion des ressources et d'exploitation. En ayant une estimation plus précise des réserves minérales, les entreprises minières peuvent optimiser leurs plans d'extraction, réduire les déchets et minimiser les coûts. Cela peut entraîner des améliorations significatives de la rentabilité et de la durabilité des opérations minières.

De plus, l'utilisation de la géostatistique et des méthodes de validation croisée peut accroître la fiabilité et la crédibilité des estimations de ressources, ce qui est important pour les investisseurs, les organismes de réglementation et les autres parties prenantes impliquées dans l'industrie minière. Dans l'ensemble, la validation des modèles de variogrammes à l'aide de la méthode LOOCV offre des avantages techniques et pratiques, en en faisant un outil important pour l'industrie minière.

En fin, la comparaison des résultats des différentes études menées sur le gisement de Bled El-Hadba a mis en évidence une divergence significative. D'une part, l'ORGM a estimé les réserves en utilisant la méthode des blocs géologiques, adoptant ainsi une approche prudente envers le potentiel du gisement, ce qui aurait pu conduire à une sous-estimation de celui-ci. De plus, la classification des catégories de réserves reposait sur la distance entre les sondages, alors que les normes internationales, notamment la norme JORC, exigeaient une classification en fonction de la précision de l'estimation effectuée et du niveau d'exploration du site. D'autre part, DMT a utilisé des méthodes de modélisation géostatistiques, fournissant ainsi des résultats plus fiables. Cependant, une critique de cette approche résidait dans l'utilisation de modèles en trois dimensions, malgré l'existence d'une anisotropie géométrique notable entre l'étendue et la puissance du gisement, un aspect qui n'a pas été pris en compte lors de l'estimation du gisement. Cette lacune aurait pu conduire à d'éventuelles surestimations du gisement de Bled El-Hadba.

4. Conclusion

L'approche géostatistique développée dans cette étude a démontré son efficacité pour améliorer l'estimation des réserves minérales. L'utilisation des méthodes des moindres carrés en combinaison avec LOOCV dans la sélection du modèle de variogrammes dans les phases d'estimation permet une représentation plus précise de la variabilité spatiale des paramètres étudiés. L'évaluation des facteurs d'efficacité confirme la qualité des modèles d'ajustement obtenus, bien que la sensibilité des moindres carrés aux données initiales souligne la nécessité de prudence durant son utilisation. L'applicabilité de la méthode LOOCV peut être limitée à des sites avec un nombre réduit de points de données. Le développement d'un code informatique offre un outil pratique pour la gestion des sites miniers. Dans l'ensemble, cette étude contribue à l'avancement des techniques géostatistiques dans l'industrie minière.

Les réserves du gisement, estimées à environ 402,38 millions de tonnes pour une teneur de coupure de 20%, créent une mine géante avec des perspectives d'exploitation à long terme. Cette évaluation positive incite les entreprises minières à exploiter activement le gisement, stimulant les investissements dans des projets de transformation dans la région du Nord-Est algérien.

Conclusion générale

Conclusion générale

Le présent travail a eu comme objectif de développer une approche géostatistique, permettant d'étudier le choix des modèles variographiques à utiliser dans la procédure d'estimation des réserves minières. En effet, le meilleur modèle est choisi à travers un couplage ajustement-validation croisée, et ce en employant la méthode des moindres carrées et LOOCV.

La mise en œuvre de cette approche nous a permis de consolider la procédure de calcul géostatistique et d'améliorer par conséquent l'estimation des réserves. A l'issue de l'étude variographique, le couplage ajustement-validation croisée intervient systématiquement dans le choix du modèle de variogramme ayant les caractéristiques qui correspondent à l'erreur quadratique minimale de validation.

Dans ce contexte, plusieurs facteurs d'efficacité sont utilisés afin d'évaluer la qualité des modèles d'ajustement en l'occurrence : Nash–Sutcliffe, RMSE, RSR. Ces derniers affichent des valeurs admissibles dans le cas de la mine de Bled El-Hadba, ce qui montre l'efficacité de la méthode d'ajustement mise en œuvre.

Associer LOOCV comme méthode de validation croisée à la procédure d'ajustement rend la procédure de calcul plus efficace, permettant ainsi d'exploiter toutes les données du site d'étude. Cela conduit à une meilleure représentativité du modèle choisi, notamment dans le cas d'un site comportant des sondages non uniformément répartis.

En outre, il est à signaler que l'emploi de la méthode des moindres carrées dans la phase d'ajustement est une technique simple, puissante et efficace, pour trouver des solutions analytiques des problèmes non linéaires. Néanmoins, elle présente une sensibilité envers les données initiales, ce qui peut provoquer des divergences. C'est pourquoi, le choix des données initiales doit se faire avec précaution.

L'approche géostatistique est mise en œuvre à travers le développement d'un code de calcul sous MATLAB, permettant de disposer d'un outil d'aide à la décision plus précis, avec lequel les décideurs peuvent se baser, pour améliorer la gestion et l'exploitation des sites miniers.

Il est à noter que le choix de la méthode de validation croisée dépend du nombre et de la distribution spatiale des échantillons. Lorsqu'il s'agit d'un échantillonnage limité, la méthode LOOCV est recommandée, car elle implique l'élimination séquentielle d'un échantillon pour

la vérification. Cette méthode garantit que la continuité spatiale des paramètres étudiés n'est pas affectée en utilisant $(n-1)$ échantillons pour le calcul du variogramme expérimental.

Nous suggérons une généralisation de la méthode d'évaluation de la précision en utilisant la validation croisée Leave-One-Out (LOOCV), une technique de validation des modèles théoriques largement utilisée dans des domaines d'estimation tels que la géostatistique. Cette méthode est particulièrement adaptée aux petits ensembles de données, comme on en trouve fréquemment dans l'industrie minière.

À l'avenir, il est important d'explorer des méthodes offrant des évaluations plus robustes tout en tenant compte de la simplicité d'application et du coût computationnel minimal. En recherchant de telles méthodes, il est possible de trouver un équilibre entre précision et efficacité computationnelle. Cependant, il est essentiel de choisir la méthode de validation en fonction des caractéristiques spécifiques du problème et des données disponibles. Différentes méthodes peuvent être plus adaptées à différentes situations, et une réflexion approfondie doit être menée pour s'assurer que la méthode choisie est en phase avec les exigences et les limites de l'étude.

Les résultats de cette étude indiquent que les réserves de phosphate dans la couche médiane, avec une teneur de coupure de 20%, sont estimées à environ 402,38 millions de tonnes. Cette évaluation favorable présente le gisement comme une mine géante avec des perspectives d'exploitation à long terme, incitant les entreprises minières à investir dans l'exploitation de ce gisement et à développer des projets de transformation dans la région du Nord-Est algérien. Cette abondance de ressources offre non seulement une durée prolongée d'activités minières mais ouvre également des opportunités stratégiques pour le développement économique régional grâce à la transformation des phosphates.

Références bibliographiques

Références bibliographiques

1. Marjoribanks, R., *Geological methods in mineral exploration and mining*. 2010: Springer Science & Business Media.
2. Isaaks, E.-H. and R.-M. Srivastava, *Applied geostatistics*. , ed. i. Oxford University Press. 1989.
3. Chiles, J.-P. and P. Delfiner, *Geostatistics: modeling spatial uncertainty*. Vol. 713. 2012: John Wiley & Sons.
4. Li, L. and H. Lan, *Integration of Spatial Probability and Size in Slope-Unit-Based Landslide Susceptibility Assessment: A Case Study*. *Int J Environ Res Public Health*, 2020. **17**(21).
5. Afeni, T.B., V.O. Akeju, and A.E. Aladejare, *A comparative study of geometric and geostatistical methods for qualitative reserve estimation of limestone deposit*. *Geoscience Frontiers*, 2021. **12**(1): p. 243-253.
6. Zerzour, O., et al., *Geostatistics-Based Method for Irregular Mineral Resource Estimation, in Ouenza Iron Mine, Northeastern Algeria*. *Geotechnical and Geological Engineering*, 2021. **39**(5): p. 3337-3346.
7. Kaplan, U.E., Y. Dagasan, and E. Topal, *Mineral grade estimation using gradient boosting regression trees*. *International Journal of Mining, Reclamation and Environment*, 2021. **35**(10): p. 728-742.
8. Kaplan, U.E. and E. Topal, *A new ore grade estimation using combine machine learning algorithms*. *Minerals*, 2020. **10**(10): p. 847.
9. Sinclair, A.J. and G.H. Blackwell, *Applied mineral inventory estimation*. 2006: Cambridge University Press.
10. SAMREC, *The South African Code for the reporting of exploration results, mineral resources and mineral reserves (THE SAMREC CODE)*. 2016, Group of the SSC Committee under the joint auspices of the Southern African Institute of Mining and Metallurgy (SAIMM) and the Geological Society of South Africa (GSSA).
11. SME, *The SME guide for reporting exploration results, mineral resources, and mineral reserves*. 2014, the Society for Mining, Metallurgy, and Exploration.
12. CIM, *CIM Definition Standards for Mineral Resources & Mineral Reserves*. 2014, Canadian Institute of Mining, Metallurgy and Petroleum: Canada.
13. JORC, *Australasian Code for Reporting of Exploration Results, Mineral Resources and Ore Reserves (The JORC Code) [online]*. 2012.
14. Dominy, S.C., M.A. Noppé, and A.E. Annels, *Errors and uncertainty in mineral resource and ore reserve estimation: The importance of getting it right*. *Exploration and Mining Geology*, 2002. **11**(1-4): p. 77-98.
15. Alastair, J.S. and H.B. Garston, *Applied Mineral Inventory Estimation*. 2004: cambridge university press.
16. Armitage, M. and I. Robinson, *Sampling, grade control and ore reserve estimation at Renco gold mine, Zimbabwe: past, present and future*. Geological Society, London, Special Publications, 1992. **63**(1): p. 121-123.
17. Emery, X., J.M. Ortiz, and J.J. Rodriguez, *Quantifying uncertainty in mineral resources by use of classification schemes and conditional simulations*. *Mathematical geology*, 2006. **38**: p. 445-464.
18. Rendu, J.-M., *International aspects of resource and reserve reporting standards*. AUSIMM website, 2001.
19. Rendu, J.-M., *Reporting mineral resources and mineral reserves in the United States of America*. Technical and Regulatory Issues, 2006.

20. Ellis, T.R., *Reporting standards-The USA experience*. CIM Bulletin-Canadian Mining and Metallurgical Bulletin, 2003. **96**(1067): p. 37-47.
21. Rossi, M.E. and C.V. Deutsch, *Mineral resource estimation*. 2014, London: Springer Science & Business Media.
22. Stephenson, P.R., *The JORC code*. Applied Earth Science, 2001. **110**(3): p. 121-125.
23. Glacken, I.M., D.V. Snowden, and A.C. Edwards, *Mineral resource estimation*. Mineral resource and ore reserve estimation-the Aus. Inst. Mining and Metallurgy guide to good practice, 2001: p. 189-198.
24. Noppé, M., *Communicating confidence in mineral resources and mineral reserves*. Journal of the Southern African Institute of Mining and Metallurgy, 2014. **114**(3): p. 213-213.
25. Camisani-Calzolari, F., D. Krige, and J. Dixon, *The South African code for reporting of mineral resources and mineral reserves and the geostatistical implications involved*. Proc. 6th Int. Geostat. Cong., Cape Town, South Africa,, 2000.
26. Lomborg, K. and S. Rupprecht, *The 2016 SAMREC code*. Journal of the Southern African Institute of Mining and Metallurgy, 2017. **117**(12): p. 1095-1100.
27. Hartman, H.L. and J.M. Mutmanský, *Introductory mining engineering*. Second edition ed. 2002: John Wiley & Sons.
28. Abzalov, M., *Applied mining geology*. Vol. 12. 2016: Springer.
29. Stone, J.G. and P.G. Dunn, *Ore reserve estimates in the real world*. 1996: Society of Economic Geologists Littleton.
30. Annel, A.E., *Mineral deposit evaluation: A practical approach*. First edition ed. 1991, Chapman & Hall, 2-6 Boundary Row, London SE1 8HN: British Library Cataloguing in Publication Data.
31. Royle, A.G., *Plane projections of tabular ore bodies for evaluation purposes*. Vol. 88: A87-A91. 1979: Transactions IMM.
32. Setianto, A. and T. Triandini, *Comparison of kriging and inverse distance weighted (IDW) interpolation methods in lineament extraction and analysis*. Journal of Applied Geology, 2013. **5**(1).
33. Arnaud, M. and X. Emery, *Estimation et interpolation de données spatiales. Méthodes déterministes et méthodes géostatistiques*. 1999.
34. Vehtari, A., et al., *Limitations of "Limitations of Bayesian Leave-one-out Cross-Validation for Model Selection"*. Computational Brain & Behavior, 2018. **2**(1): p. 22-27.
35. Krige, D.G., *A statistical approach to some basic mine valuation problems on the Witwatersrand*. Journal of the Southern African Institute of Mining and Metallurgy, 1951. **52**(6): p. 119-139.
36. Krige, D.G., *On the departure of ore value distributions from the lognormal model in South African gold mines*. Journal of the Southern African Institute of Mining and Metallurgy, 1960. **61**(4): p. 231-244.
37. Trangmar, B.B., R.S. Yost, and G. Uehara, *Application of geostatistics to spatial studies of soil properties*. Advances in agronomy, 1986. **38**: p. 45-94.
38. Pyrcz, M.J. and C.V. Deutsch, *Geostatistical Reservoir Modeling*. 2014, Oxford University Press.
39. Xavier, E., *Géostatistique linéaire*. 2001, France: Ecole des mines de Paris.
40. Journel, A.G. and C.J. Huijbregts, *Mining Geostatistics*. 1978: Academic Press Limited, London.
41. Matheron, G., *Kriging or polynomial interpolation procedures*. CIMM Transactions, 1967. **70**(1): p. 240-244.

42. David, M., *Geostatistical Ore Reserve Estimation*. First edition ed. Developments in Geomathematics 2. 1977, Amsterdam: Elsevier Scientific Publishing Company.
43. Clark, I., *Practical geostatistics*. Vol. 3. 1979: Applied Science Publishers London.
44. Sarma, D.D., *Geostatistics with applications in earth sciences*. Second Edition ed. 2009, India: Springer Science & Business Media.
45. Wilding, L.P. and L. Drees, *Spatial variability and pedology*, in *Developments in Soil Science*. 1983, Elsevier. p. 83-116.
46. Emery, X., *Change-of-support models and computer programs for direct block-support simulation*. *Computers & geosciences*, 2008. **35**(10): p. 2047-2056.
47. Wackernagel, H., *Multivariate geostatistics: an introduction with applications*. 2003: Springer Science & Business Media.
48. Causeur, D., *Modélisation géostatistique*. Laboratoire de Mathématiques Appliquées. Pôle d'Enseignement Supérieur et de Recherche Agronomique de Rennes. 26p, 2000.
49. Vehtari, A. and J. Lampinen, *Bayesian model assessment and comparison using cross-validation predictive densities*. *Neural computation*, 2002. **14**(10): p. 2439-2468.
50. Vehtari, A. and J. Ojanen, *A survey of Bayesian predictive methods for model assessment, selection and comparison*. *Statistics Surveys*, 2012. **6**(none).
51. Stone, M., *Cross-validation: A review*. *Statistics: A Journal of Theoretical and Applied Statistics*, 1978. **9**(1): p. 127-139.
52. Stone, M., *Cross-validatory choice and assessment of statistical predictions*. *Journal of the royal statistical society: Series B (Methodological)*, 1974. **36**(2): p. 111-133.
53. Arlot, S. and A. Celisse, *A survey of cross-validation procedures for model selection*. *Statistics Surveys*, 2010. **Vol. 4** p. 40–79.
54. Yates, L.A., et al., *Cross validation for model selection: a review with examples from ecology*. *Ecological Monographs*, 2022. **93**(1): p. e1557.
55. Gneiting, T. and A.E. Raftery, *Strictly proper scoring rules, prediction, and estimation*. *Journal of the American statistical Association*, 2007. **102**(477): p. 359-378.
56. Arlot, S., *V-fold cross-validation improved: V-fold penalization*. arXiv preprint arXiv:0802.0566, 2008.
57. Vehtari, A., A. Gelman, and J. Gabry, *Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC*. *Statistics and computing*, 2017. **27**: p. 1413-1432.
58. Yadav, S. and S. Shukla. *Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification*. in *2016 IEEE 6th International conference on advanced computing (IACC)*. 2016. IEEE.
59. Blum, A., A. Kalai, and J. Langford. *Beating the hold-out: Bounds for k-fold and progressive cross-validation*. in *Proceedings of the twelfth annual conference on Computational learning theory*. 1999.
60. Stuart, R. and N. Peter, *Artificial Intelligence A Modern Approach Third Edition*. 2010, PEARSON.
61. Russell, S.J. and P. Norvig, *Artificial intelligence a modern approach*. 2010: London.
62. Berrar, D., *Cross-Validation*. *Encyclopedia of Bioinformatics and Computational Biology*, 2019. **Vol. 1**: p. 542-545.
63. Marcot, B.G. and A.M. Hanea, *What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?* *Computational Statistics*, 2020. **36**(3): p. 2009-2031.
64. Liu, S., *Leave- p -Out Cross-Validation Test for Uncertain Verhulst-Pearl Model With Imprecise Observations*. *IEEE Access*, 2019. **7**: p. 131705-131709.

65. Celisse, A. and T. Mary-Huard, *Exact Cross-Validation for k NN : application to passive and active learning in classification*. Journal de la Société Française de Statistique, 2011. **152**(3): p. 83-97.
66. James, G., et al., *An introduction to statistical learning*. Vol. 112. 2013: Springer.
67. Prian, J.P. and P. Cortial, *Etude de développement du gisement de phosphate de Djebel Onk (Algérie). Rapport d'expertise géologiques 1993*: B.R.G.M., Algérie. 288.
68. Kechiched, R., *Typologie géochimique et géostatistique des minerais de phosphates du gisement de Bled El Hadba – Djebel Onk (Algérie Orientale)*. 2011, Université Badji Mokhtar – Annaba. p. 175p.
69. Chabou, M.S., *Etude de la série stratigraphique tertiaire du Djebel Onk (Algérie). Leur situation dans le cadre des phosphates algériens*. 1987, Fac. Sci. Alger. p. 136.
70. Mezghache, H. and A. Hani, *Typologie chimique des phosphates du gisement de Djemi-Djema-bassin de Djebel Onk (Algérie orientale)*. Géologie Méditerranéenne, 2000. **27**(1): p. 95-106.
71. Chabou, M.S., *Caractérisation des phosphates du Djebel Onk*. 1998, Département Génie Minier-Ecole Nationale Polytechnique, Alger, Algérie.
72. Chabou, M.S. and R. Flicoteaux, *The Paleocene-Eocene phosphate deposits of Djebel Onk, Algeria*. Phosphate deposits of the world, 1989. **2**: p. pp 289-293.
73. Cielensky, S. and N. Benchernine, *Travaux de prospection et d'évaluation des phosphates dans la région de Bir El Ater*. Rapport interne. 1987, Algérie: EREM. 80p.
74. Aissaoui, D., *Les structures liées à l'accident sud-atlasique entre Biskra et le Djebel Mannadra, Algérie. Evolution géométrique et cinématique*. 1984, Strasbourg. p. 138p.
75. O.R.G.M., *Rapport Final sur les Resultats des Travaux D'exploration des Phosphates du Gisement de Bled el Hadba 2012–2014*. ORGM: Algeria Boumerdes, . 2015, Algeria
76. DMT-Consulting, G., *Etude de Faisabilité de Bancable du gisement de Bled El-Hadba*. N°14/SOMIPHOS/DMT Consulting GmbH ed. 2016. 1321.
77. Audibert, T., *La méthode de Newton et ses variantes pour l'optimisation*. 2013.
78. Cornillon, P.-A. and E. Matzner-Løber, *Moindres carrés généralisés. Régression: Théorie et applications*, 2007: p. 179-196.
79. Moriasi, D.N., et al., *Model evaluation guidelines for systematic quantification of accuracy in watershed simulations*. American Society of Agricultural and Biological Engineers, 2007. **50**(3): p. 885–900.
80. Zhong, X. and U. Dutta, *Engaging Nash-Sutcliffe efficiency and model efficiency factor indicators in selecting and validating effective light rail system operation and maintenance cost models*. Journal of Traffic and Transportation Engineering, 2015. **3**: p. 255-265.
81. Robert, E.C. and E.W. William, *Do Nash values have value? Discussion and alternate proposals*. Hydrological Processes, 2008. **22**: p. 2723–2725.
82. Chai, T. and R.R. Draxler, *Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature*. Geoscientific Model Development, 2014. **7**: p. 1247–1250.
83. Prasad, M.B.K., et al., *Predicting dissolved oxygen in the Chesapeake Bay: applications and implications*. Aquatic Sciences, 2011. **73**(3): p. 437-451.
84. Afzal, P., et al., *Mineral Resource Classification Using Geostatistical and Fractal Simulation in the Masjed Daghi Cu–Mo Porphyry Deposit, NW Iran*. Minerals, 2023. **13**(3): p. 370.
85. Zeybek, M., *Nash-sutcliffe efficiency approach for quality improvement*. Journal of Applied Mathematics and Computation, 2018. **2**(11): p. 496-503.

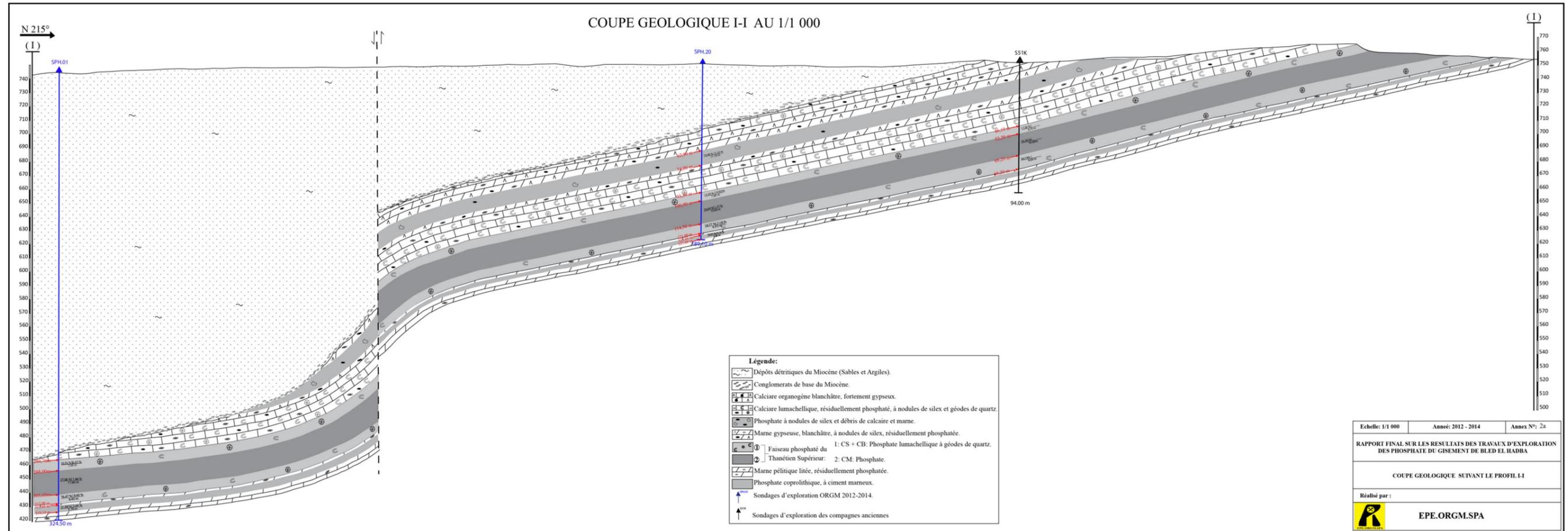
86. Magnusson, M., et al. *Leave-one-out cross-validation for Bayesian model comparison in large data*. in *International conference on artificial intelligence and statistics*. 2020. PMLR.
87. Gronau, Q.F. and E.-J. Wagenmakers, *Limitations of Bayesian leave-one-out cross-validation for model selection*. *Computational brain & behavior*, 2019. **2**: p. 1-11.

Annexes

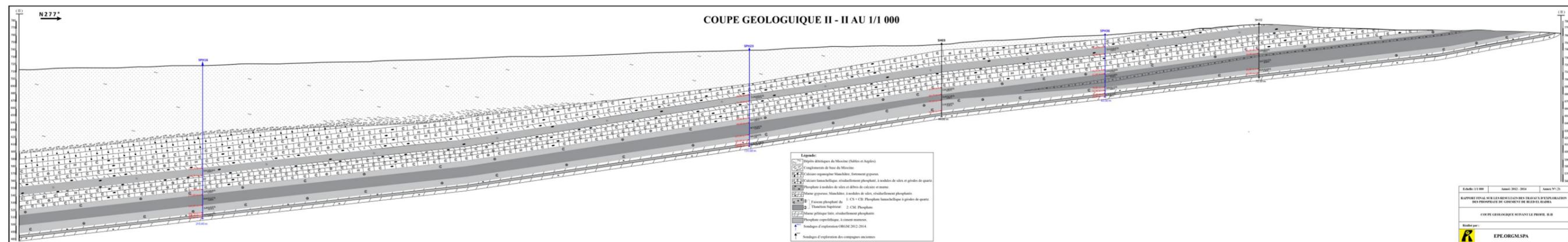
Annexes :

Annexes 1 : Coupe géologique des différents profils dans la carte de l'annexe 1.

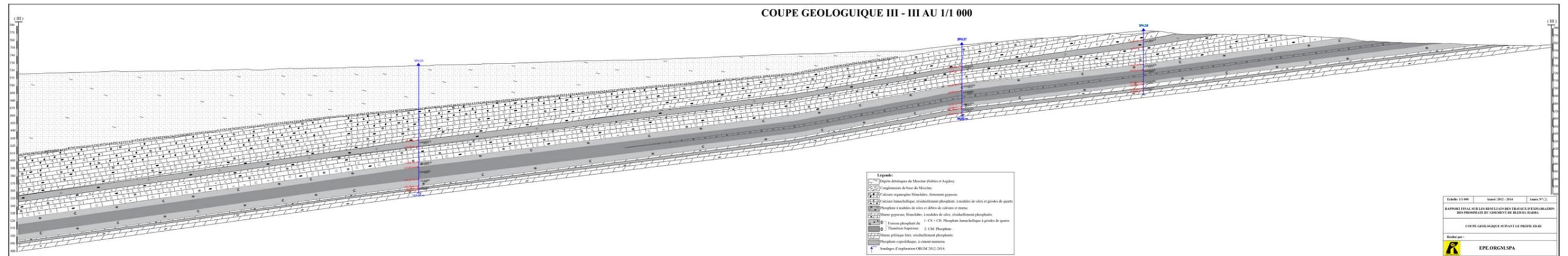
a) Coupe géologique suivant le profil I-I, Echelle 1/1000.



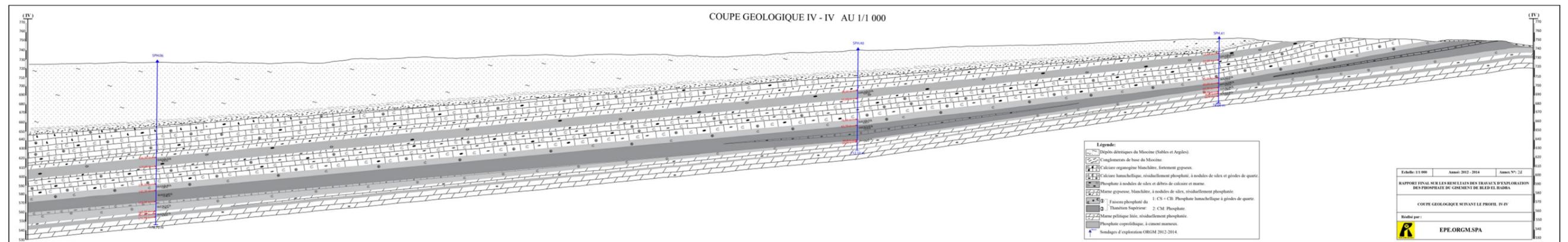
b) Coupe géologique suivant le profil II-II, Echelle 1/1000.



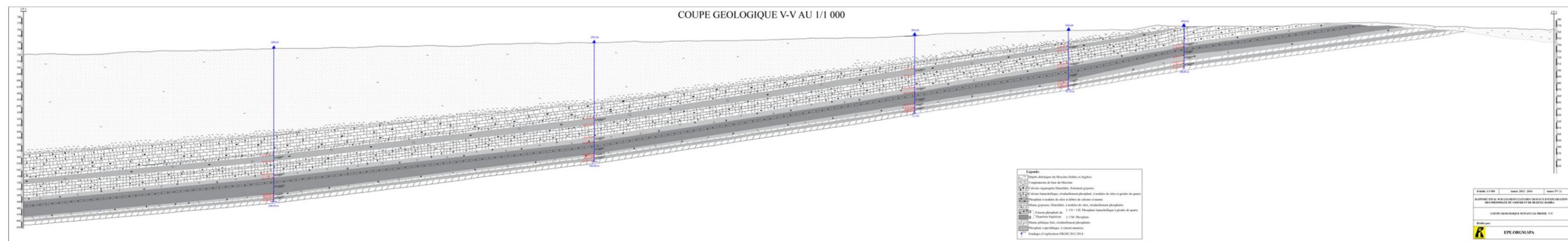
c) Coupe géologique suivant le profil III-III, Echelle 1/1000.



d) Coupe géologique suivant le profil IV-IV, Echelle 1/1000.

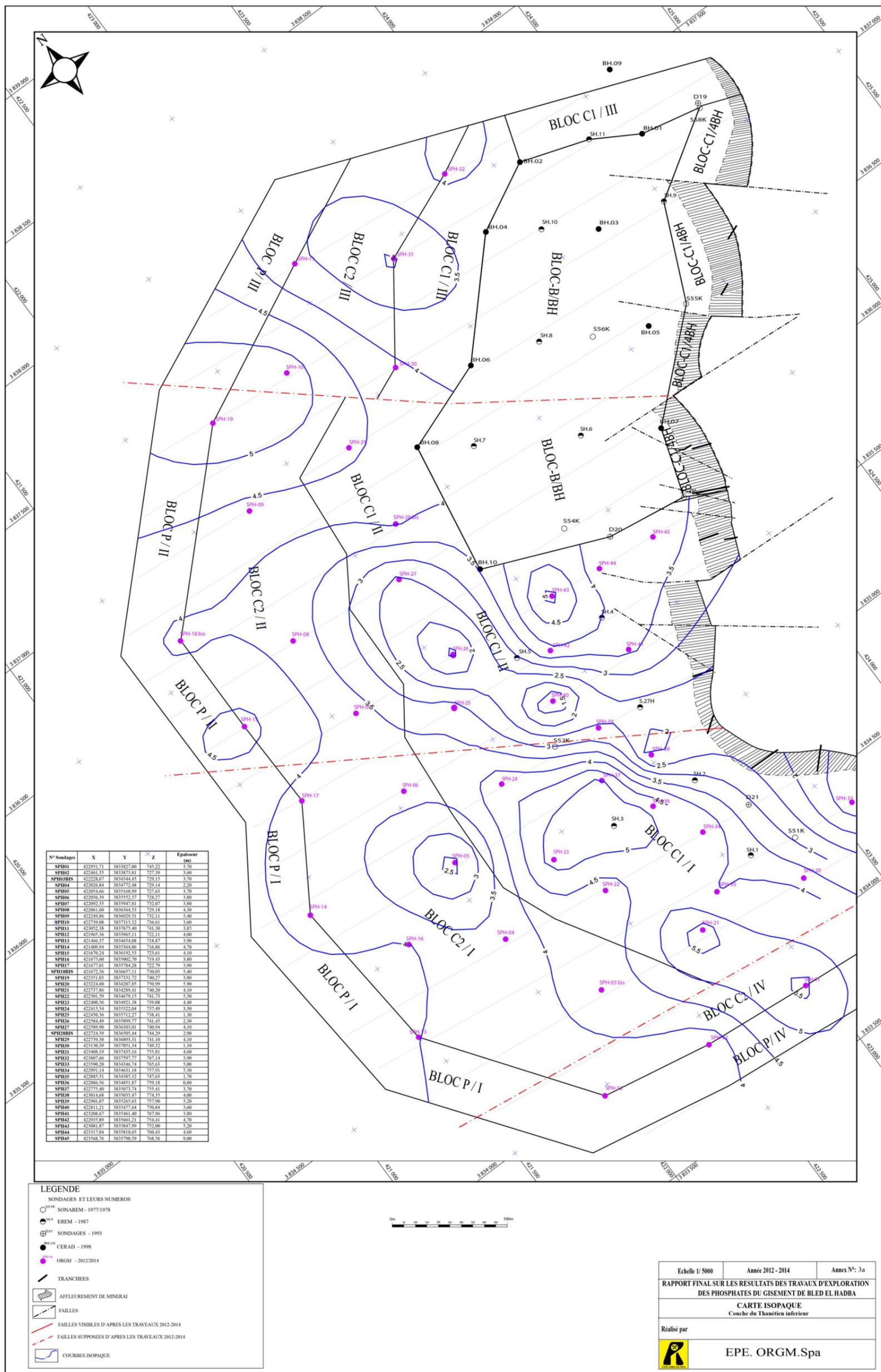


e) Coupe géologique suivant le profil V-V, Echelle 1/1000.

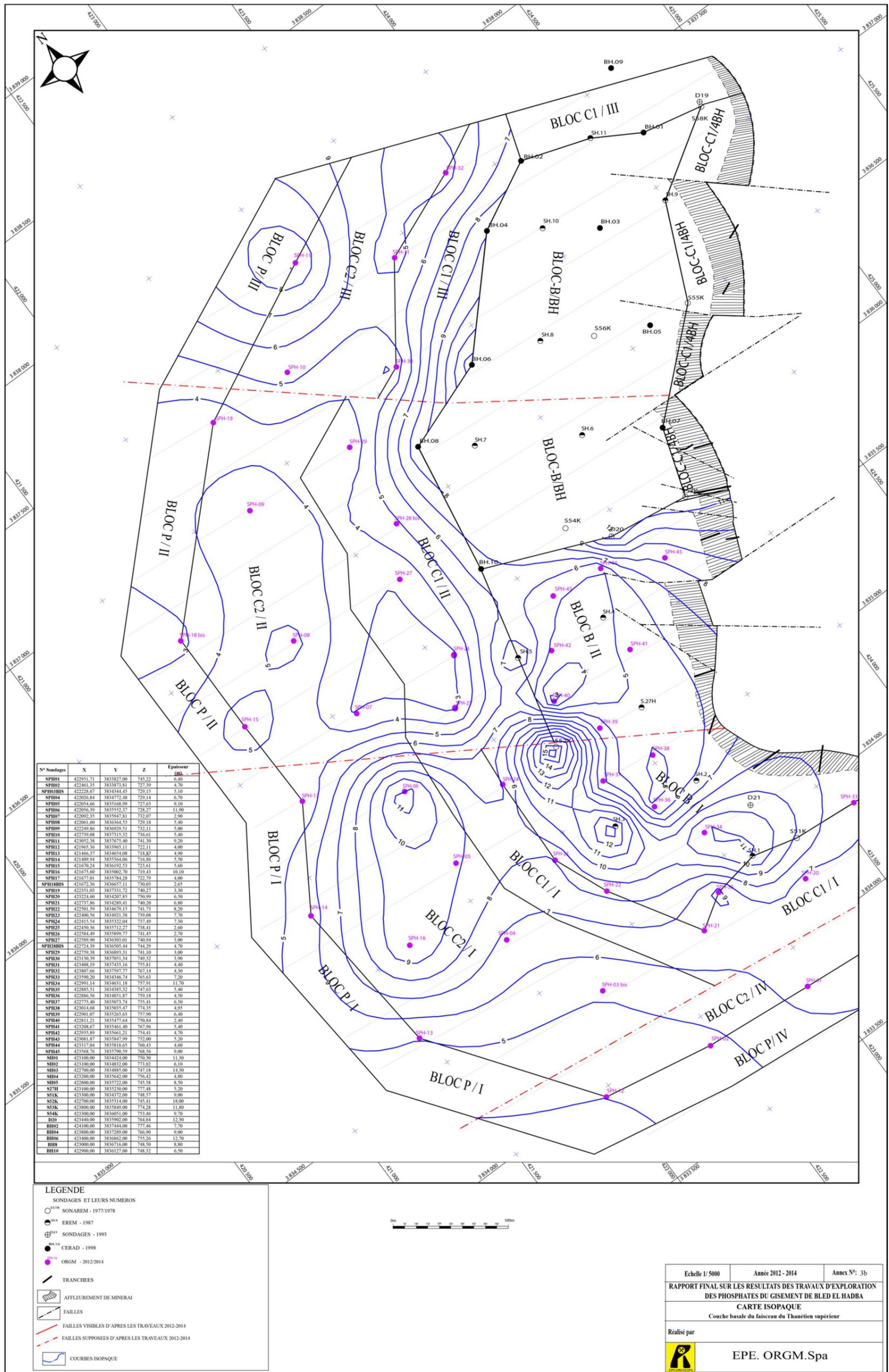


Annexes 2 :

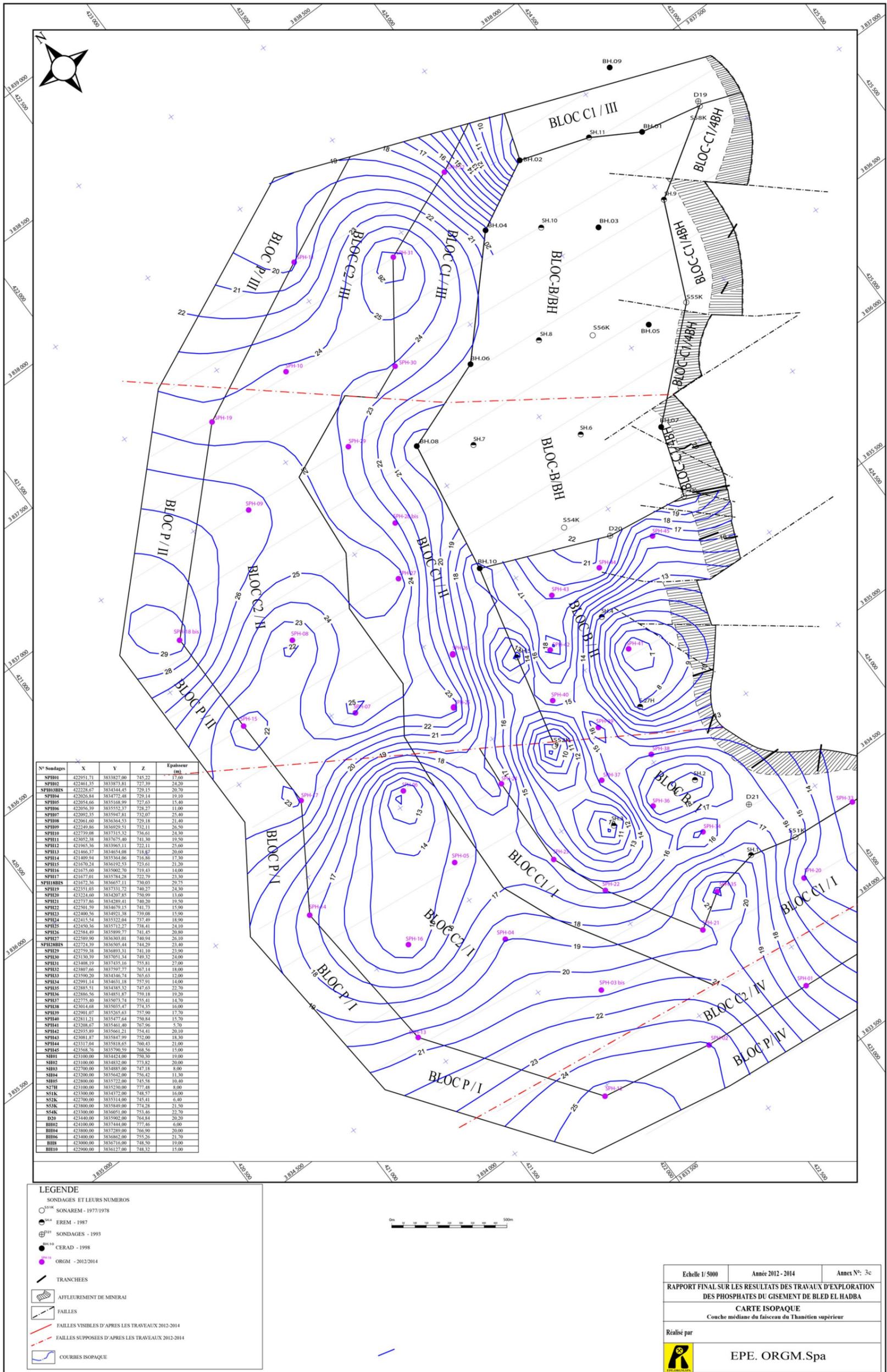
a) Carte isopaque de la couche du Thanétien inférieur, Echelle 1/5000.



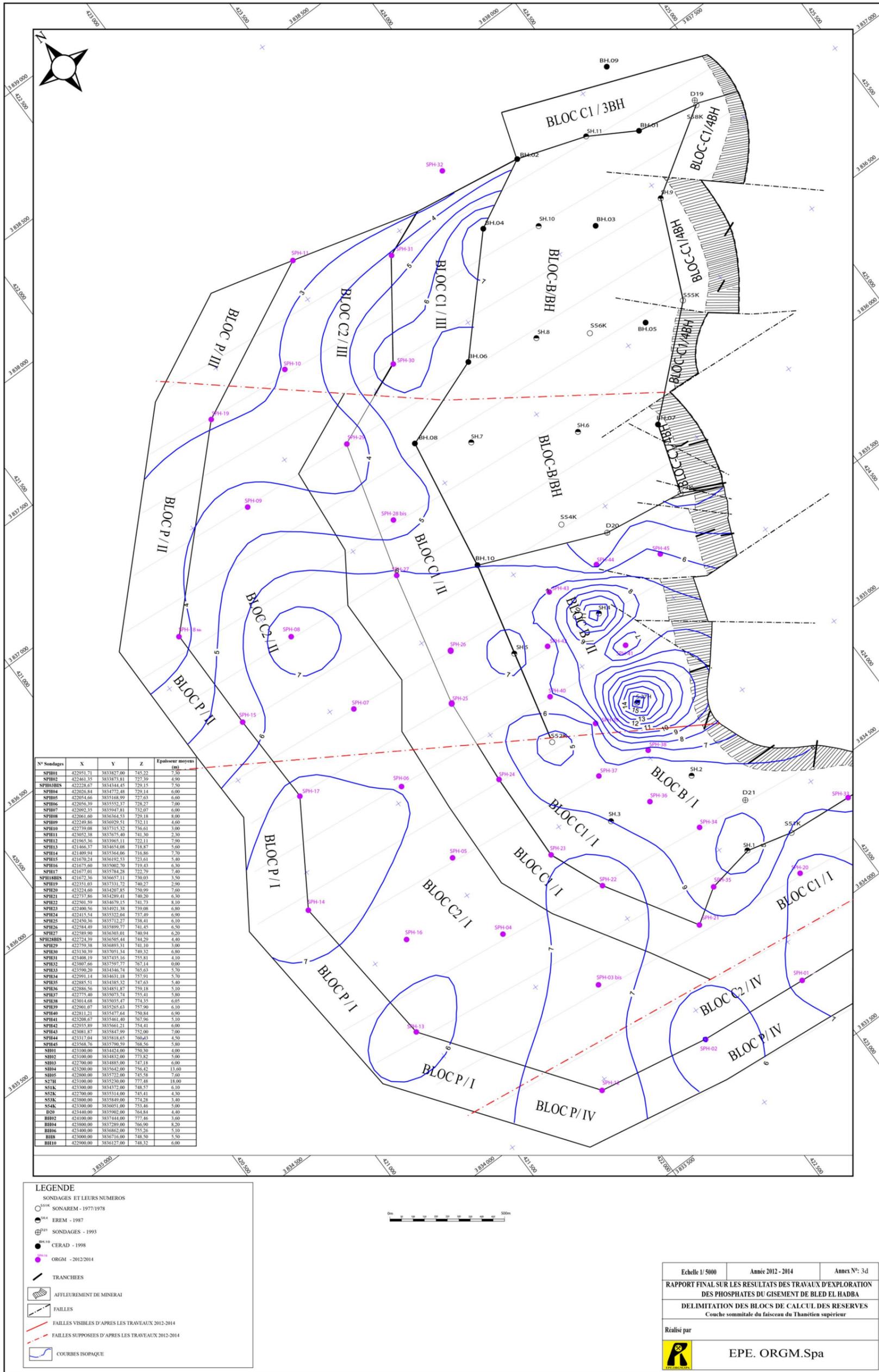
b) Carte isopaque de la couche sommitale, Echelle 1/5000.



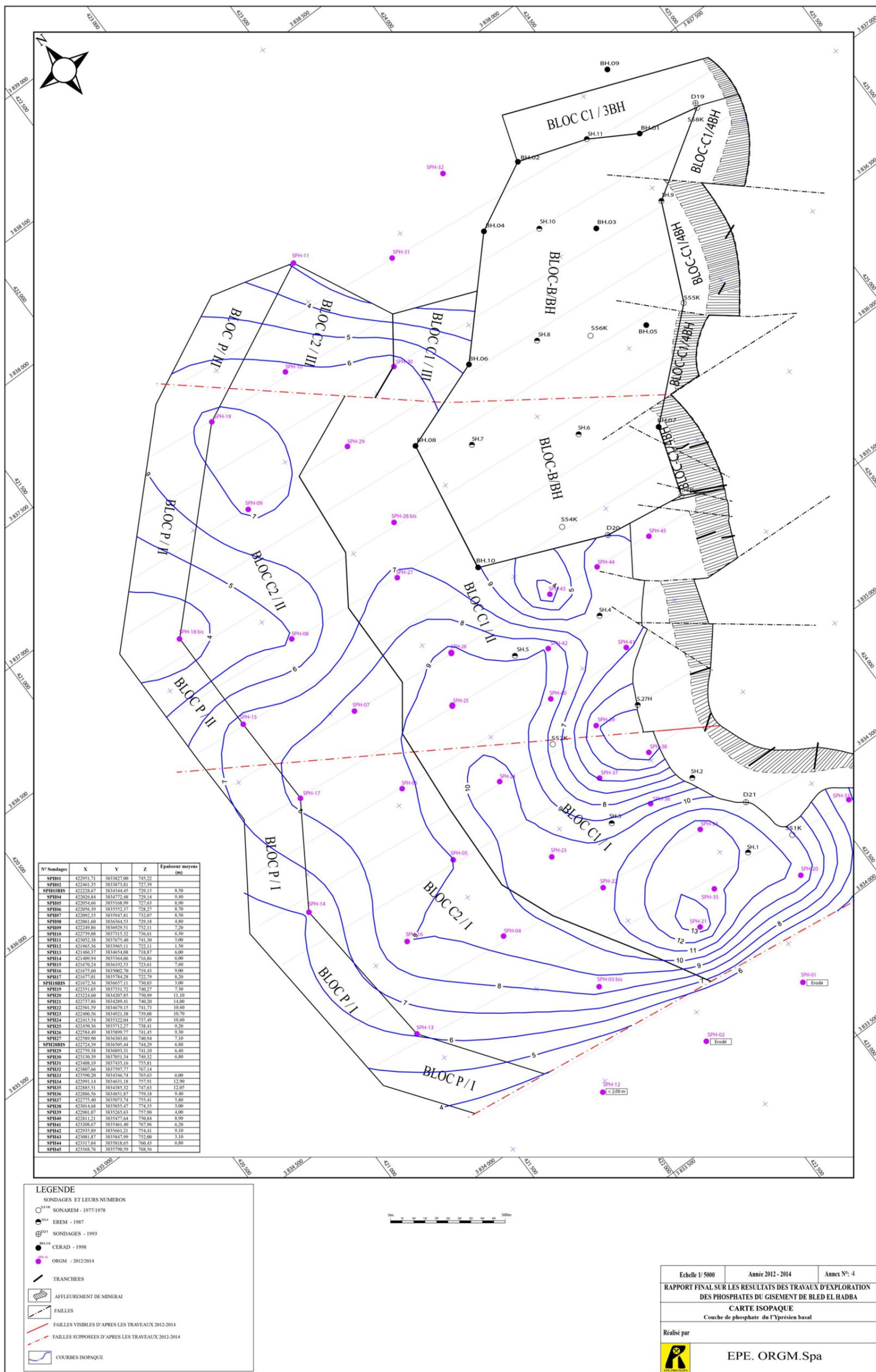
c) Carte isopaque de la couche médiane, Echelle 1/5000.



d) Carte isopaque de la couche basale, Echelle 1/5000.



Annexes 3 : Carte isopaque de la couche du l'yprésien « basal », Echelle 1/5000.



Annexe 4 :

Tableau 1. Comparaison entre le système de classification soviétique historique et les codes de déclaration Internationaux (Mining Journal, Londres, 20 août 2004 ; Stephen Henley : Russian Mineral Reporting)

Russe	Codes de déclaration internationaux
A, B	Reserve prouvé/ ressource mesuré
C1	Reserve probable/prouvé Ressource indiqué
C2	Reserve probable Ressource présumé/indiqué
P1	Ressource présumé
P2	Recherche des ressources minérales (Source : UN Framework Classification for Reserves/Resources, code334)
P3	Aucun équivalent

Selon le CRIRSCO (Committee for Mineral Reserves International Reporting Standards, c'est-à-dire le Comité pour les normes internationales pour les rapports sur les réserves et ressources minérales), les catégories de ressources A + B + C1 + C2 du code NAEN (2011) peuvent être évaluées comme équivalentes aux ressources M + I. Le code NAEN publié en 2011 établit une conversion entre les ressources/réserves classées selon le code soviétique et le code JORC.

Annexe 5 : Carte de la variance de Krigeage représente la zone d'exploration de catégorie prouvée (a) accompagnée des deux cartes : (b) d'estimation de la puissance de minéralisation et, (c) de dispersion des teneurs, du gisement de Bled El-Hadba.

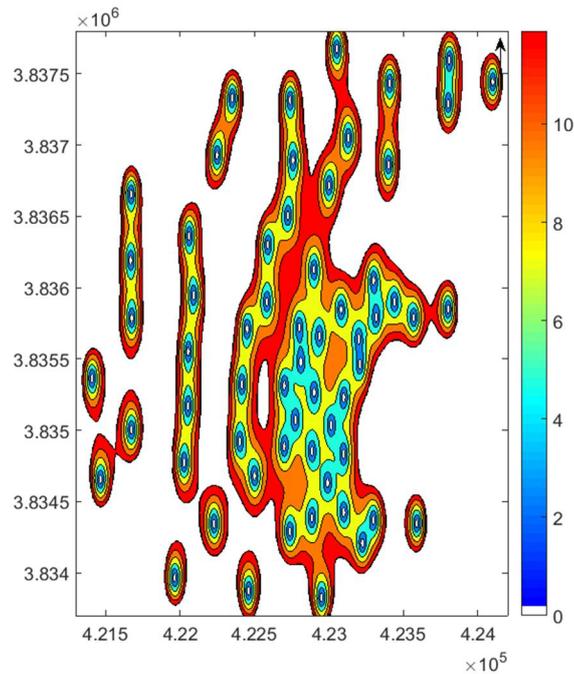


Figure 5. (a) Carte de la variance de Krigeage représente la zone d'exploration de catégorie prouvée.

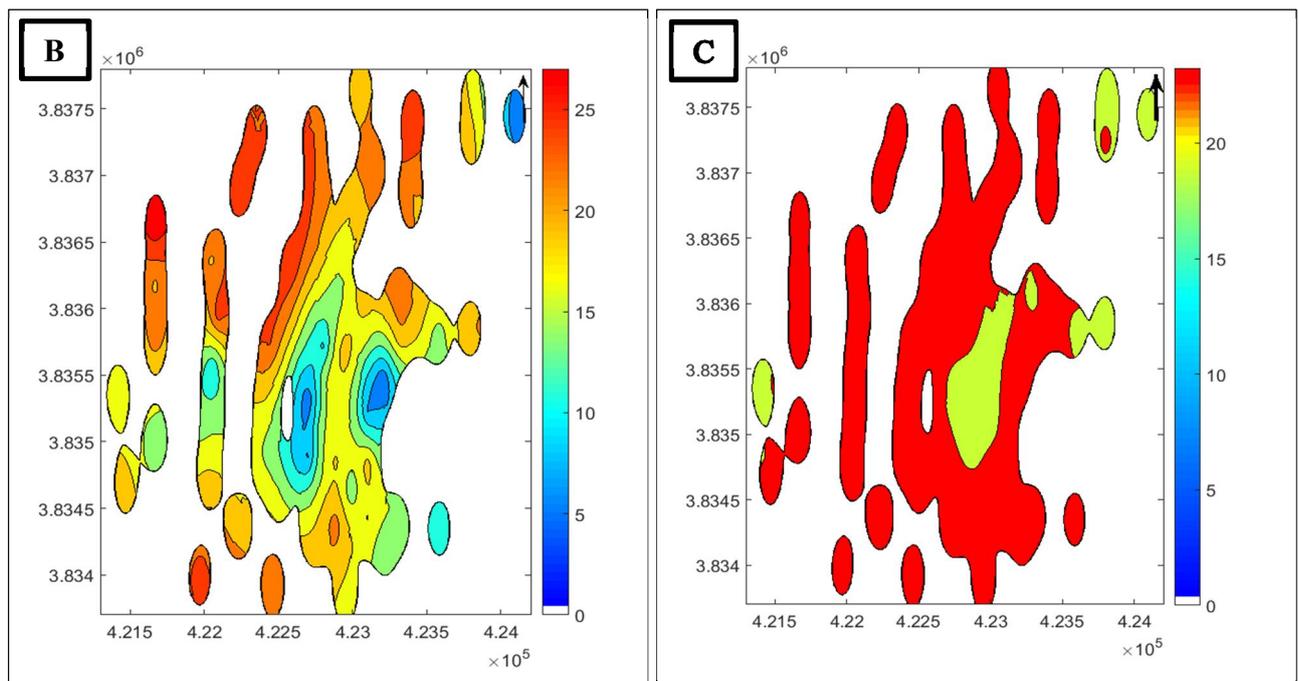


Figure 5. Carte de la zone d'exploration de catégorie prouvée, représente l'estimation de : (b) la puissance de minéralisation et (c) de dispersion des teneurs.

Annexe 5 : Carte de la variance de Krigeage représente la zone d'exploration de catégorie Probable (a) accompagnée des deux cartes : (b) d'estimation de la puissance de minéralisation et, (c) de dispersion des teneurs, du gisement de Bled El-Hadba.

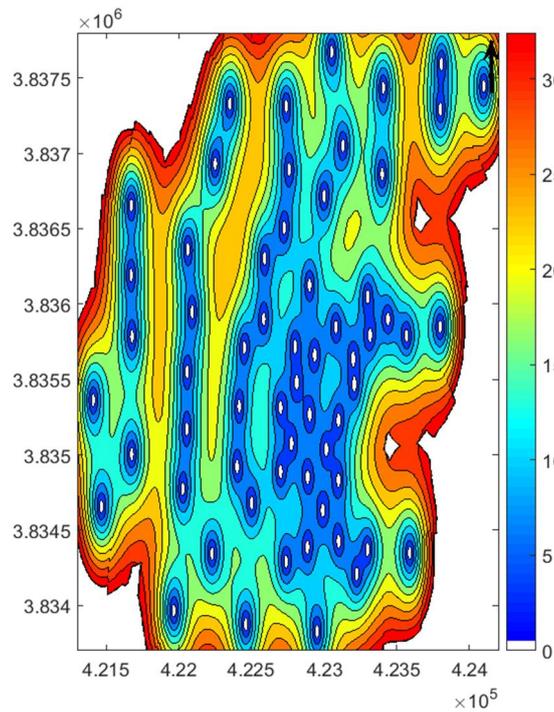


Figure 6. (a) Carte de la variance de Krigeage représente la zone d'exploration de catégorie probable.

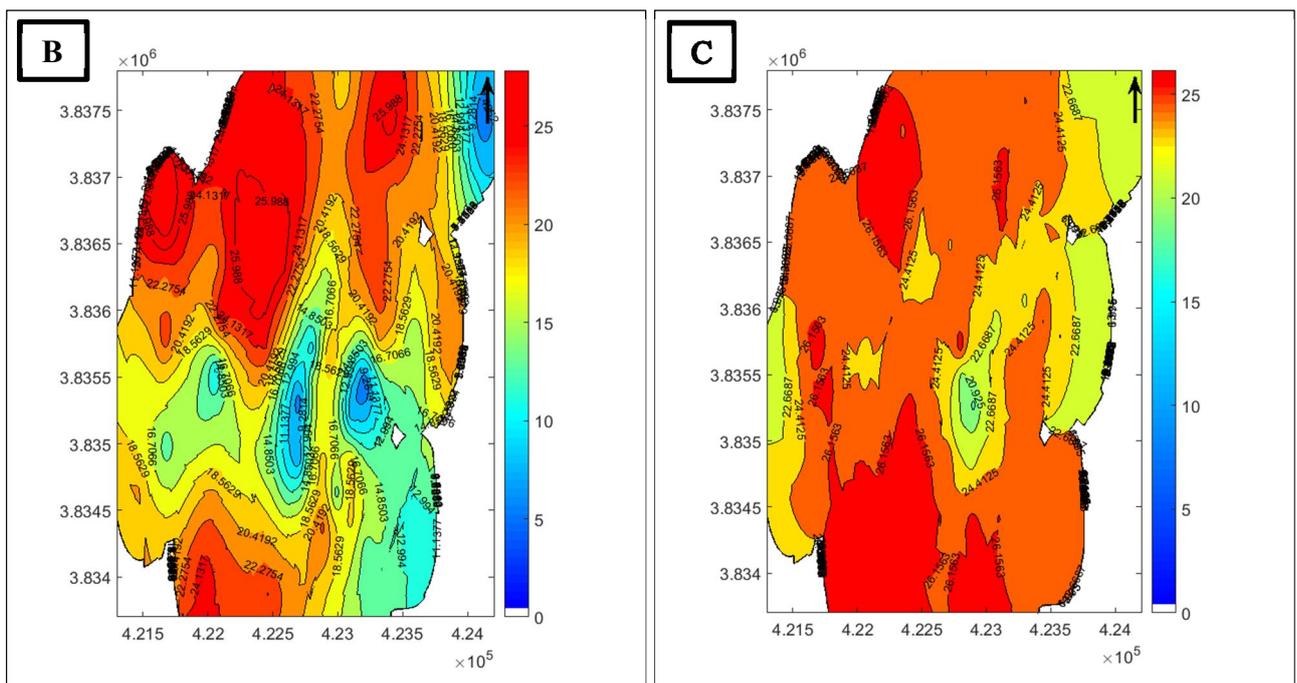


Figure 6. Carte de la zone d'exploration de catégorie probable, représente l'estimation de : (b) la puissance de minéralisation et (c) de dispersion des teneurs.