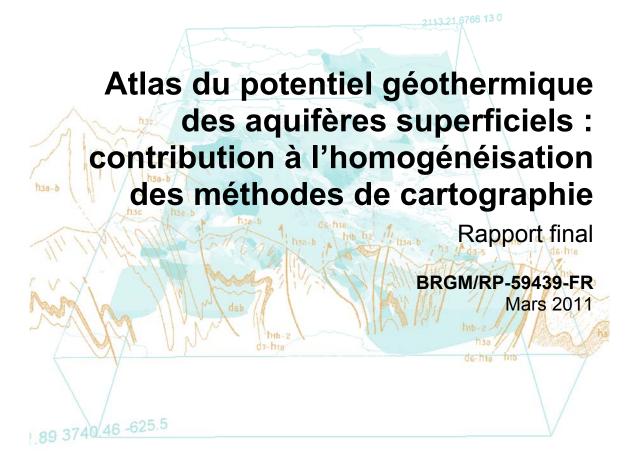


Document Public









Atlas du potentiel géothermique des aquifères superficiels : contribution à l'homogénéisation des méthodes de cartographie

Rapport final

BRGM/RP-59439-FR

Mars 2011

Étude réalisée dans le cadre des projets de Service public du BRGM 2005 10GTHB17 et de la convention ADEME – BRGM n°09 05 C0016

A. Leynet, S. Bezelgues-Courtade

Avec la collaboration de

B. Bourgine et , S. Schomburgk

Vérificateur :

Nom : JC Martin

Signatura

Approbateur:

Nom: A. Desplan

Date :

Signature :

En l'absence de signature, notamment pour les rapports diffusés en version numérique, l'original signé est disponible aux Archives du BRGM.

Le système de management de la qualité du BRGM est certifié AFAQ ISO 9001:2008.





Mots clés : Géothermie, Aquifères, Pompe à chaleur, Atlas. En bibliographie, ce rapport sera cité de la façon suivante : **Leynet A., Bézèlgues-Courtade S.,** avec la collaboration de **Bourgine B. et Schomburgk S.**(2011) - Atlas du potentiel géothermique des aquifères superficiels : contribution à l'homogénéisation des méthodes de cartographie. Rapport final BRGM/RP-59439-FR. 144p., 99 fig., 14 tabl., 4 ann. © BRGM, 2011, ce document ne peut être reproduit en totalité ou en partie sans l'autorisation expresse du BRGM.

Synthèse

Dans le cadre de la convention de partenariat n° 0905C0016, l'ADEME et le BRGM (action de service public 10GTHB17) se sont associés pour définir les principes communs à retenir pour la réalisation des atlas du potentiel géothermique régional des aquifères superficiels et assurer ainsi, autant que possible, une harmonie nationale de ces outils d'aide à la décision en cours d'élaboration dans les différentes régions de France.

En effet, dès 2003, le BRGM associé à l'ADEME et aux conseils régionaux correspondants, avait lancé la réalisation d'un outil d'aide à la décision permettant d'évaluer l'exploitabilité régionale (ou potentiel des ressources géothermique) des aquifères superficiels, dans un premier temps dans les régions Île-de-France, Lorraine, Midi-Pyrénées et Centre.

Cet outil est basé sur un atlas des ressources en eau souterraine pouvant être exploitées en géothermie très basse énergie dans les terrains superficiels. Cet atlas comprend un système d'information géographique (SIG) mis en ligne via le site INTERNET de l'ADEME et du BRGM spécifique à la géothermie www.geothermie-perspectives.fr.

Plusieurs méthodes ont été mises en œuvre, les quatre projets ayant démarré quasiment simultanément, et aucune méthode n'ayant réellement fait ses preuves, ou montré son inadaptation.

C'est pourquoi le présent projet a pour objectif, après avoir comparé les différentes méthodes suivies, de proposer une méthodologie d'évaluation cartographique du potentiel géothermique très basse énergie des aquifères superficiels.

Son champ d'application concerne l'exploitation énergétique des aquifères associées à la mise en œuvre de pompes à chaleur (PAC).

Elle aboutit à l'élaboration d'un outil d'aide à la décision répondant à un objectif de diffusion de l'information qui doit uniquement être utilisé comme tel. Il ne s'agit en effet :

- ni d'un outil de gestion des ressources en eau souterraine (dans le sens où il ne prend pas en compte toute une série de critères propres à ce type de gestion (exploitations existantes, périmètres d'influence de l'existant et du projeté, contraintes réglementaires, contraintes naturelles...) et qu'il n'est pas dynamique (évolutif dans le temps en fonction de l'évolution de l'existant);
- ni d'un outil d'évaluation de la faisabilité de projets du fait de la limitation de son échelle de validité au 1/100 000.

La méthode présentée ici se base sur une cartographie à index avec pondération de critères de potentialité géothermique. Les critères retenus sont la productivité de

l'aquifère, la température de la ressource, la profondeur d'accès à la ressource et la qualité chimique de la ressource.

Des recommandations sont émises :

- pour la classification à adopter pour chaque critère ;
- pour le choix des paramètres à prendre en compte pour cartographier ces critères (en particulier pour le critère « productivité d'aquifère »);
- pour le choix du mode de cartographie du critère « productivité d'aquifère » ;
- pour le choix du code de calcul du potentiel géothermique par aquifère à retenir.

Le critère qualité chimique de la ressource n'est pas intégré au calcul du fait de la fréquente indisponibilité de données permettant de la qualifier. Il est seulement recommandé de fournir des éléments caractéristiques (une liste de paramètres guides est fournie) des contraintes liées à la géochimie des eaux souterraines quand l'existence de ces contraintes est avérée car elles sont susceptibles de peser sur les projets d'exploitations géothermiques.

La méthode proposée reste théorique et des adaptations locales pourront être retenues en réponse à des configurations et problématiques particulières et sur avis d'un expert hydrogéologue local.

Les données disponibles permettent d'aller plus loin et de répondre à d'autres attentesdes acteurs de la géothermie. Des évolutions de cette méthode pourrontêtre envisagées, par exemple :

- production de cartes du potentiel géothermiques distinctesselon l'usage projeté (chauffage ou climatisation). Sur les cartes actuelles, le potentiel géothermique concerne les deux usages, ce qui implique une minimisation du potentiel en mode chauffage;
- représentation des contraintes (réglementaires et géologiques) locales susceptibles d'impacter la faisabilité d'un projet géothermique (ou son coût) : exploitations existantes, risques géologiques à la réalisation de forages, contraintes réglementaires...

Sommaire

1. Introd	uction	11
2. Princi	pe et méthodes de réalisation de l'atlas	13
2.1. P	RINCIPE GÉNÉRAL	13
2	.1.1.Méthodologie générale	13
	.1.2.Critères de potentialité géothermique	
2	.1.3. Cartographie du potentiel géothermique	15
	RITÈRES DE POTENTIALITÉ GÉOTHERMIQUE, SOURCES DE ONNÉES, CLASSIFICATION ET INDEXATION	19
2	2.1.Critère « Productivité d'un aquifère »	19
2	2.2.Critère température de la ressource	25
2	2.3. Critère profondeur d'accès à la ressource	25
2	2.4. Qualité de la ressource	26
2	2.5. Bilan de la classification et de l'indexation des critères	29
	VALUATION DU POTENTIEL GÈOTHERMIQUE : CHOIX DU CODE DE ALCUL MULTICRITÈRE	30
2	.3.1.Représentation : choix de l'échelle de restitution des cartes	30
3. Concl	usion	31
4. Biblio	graphie	33
Liste d	es illustrations	
Figure 1 :	Principe de l'élaboration des cartes du potentiel géothermique régional	14
Figure 2 :	Démarche de cartographie unicritère (classification et indexation)	17
Figure 3 :	Principe de l'analyse multicritère.	18
Figure 4:	Principe de calcul du variogramme.	39
Figure 5 :	Variogramme dans le cas d'un phénomène stationnaire (sans dérive) Point A	40
Figure 6 :	Variogrammes expérimentaux de 2 phénomènes ayant des portées différentes et reconnus par une maille de 100 m.	41
Figure 7:	Effet de pépite	42
Figure 8 :	Variogramme à plusieurs structures	43

Figure 9 :	Exemple de boîte à moustaches	46
Figure 10 :	Exemple de variogramme estimé (rouge) et de modèle de variogramme (bleu). Les pointillés indiquent la variance estimée, l'histogramme donne le nombre de paires de points.	49
Figure 11 :	Courbes des composantes les plus classiques (fonctions autorisées) d'un variogramme.	49
Figure 12 :	Exemple de validation croisée sous Isatis.	. 53
Figure 13 :	En haut, droite de corrélation entre Q_s et T (ρ = 0,909) ; en bas, idem avec un point négligé (ρ = 0,703).	60
Figure 14 :	Droite de corrélation entre lnQ_s et lnT (ρ = 0,820). Si l'on néglige le même point que précédemment, la droite change peu (ρ = 0,807)	61
Figure 15 :	ldem, exclus 36 points semblant alignés sur une droite (marqués en rouge), ρ = 0,783	. 62
Figure 16 :	Confrontation des Qs estimés et mesurés, en échelle logarithmique	. 63
Figure 17 :	Corrélation en Qs et T dans la Beauce.	. 64
Figure 18 :	Corrélation entre InQs et InT dans la Beauce.	. 64
Figure 19 :	Répartition des données de T et Q _s sur le territoire de la région Centre	. 70
Figure 20 :	Grès du Trias : extension et données de Qs et T disponibles	. 72
Figure 21 :	Craie du Séno-Turonien : extension et données de Q _s et T disponibles	. 73
Figure 22 :	Craie du Séno-Turonien : histogramme des InQs.	. 74
Figure 23 :	Craie du Séno-Turonien : diagramme de Henry entre InQs et une distribution Gaussienne.	. 74
Figure 24 :	Craie du Séno-Turonien : variogramme estimé de Qs.	. 75
Figure 25 :	Calcaires de Beauce : extension et données de Q _s et T disponibles	. 76
Figure 26 : I	Histogramme des débits spécifiques dans la Beauce. En bleu, partie affleurante, en vert, partie sous recouvrement	77
Figure 27 :	Histogramme des InQ _s dans la Beauce. En bleu partie affleurante, en vert, partie sous recouvrement	. 77
Figure 28 :	Calcaires de Beauce (ensemble des données) : variogramme estimé de Q_s (graphe de gauche), variogramme de lnQ_s (graphe de droite)	. 78
Figure 29 : \	Variogrammes estimés de InQ _s pour la Beauce affleurante (graphe de gauche) et la Beauce sous recouvrement (graphe de droite)	78
Figure 30 :	Répartition des points du Val de Loire (brun) et du reste de la Beauce (bleu)	. 80
Figure 31 :	Calcaires de Beauce (Val de Loire) : variogramme estimé de InQs	. 81
Figure 32 :	Beauce, sans le Val de Loire : variogramme estimé de lnQs	. 82
Figure 33 :	Sables ACF : extension et données de Q _s et T disponibles	. 83
Figure 34 :	Sables ACF : histogramme de InQ _s .	. 85
Figure 35 :	Sables ACF : diagramme de Henry pour InQ _s .	. 85
Figure 36 :	Sables ACF : variogramme estimé sur 182 données(sans doublons)	. 86

Figure 37 :	Calcaire du Jurassique : répartition des données de Qs sur le territoire de la région Centre	87
Figure 38 :	Calcaire du Jurassique : variogramme estimé de lnQs ; en bas, histogramme du nombre de paires	88
Figure 39 :	Craie du Séno-Turonien : points d'apprentissage (bleu) et de contrôle (rouge).	92
Figure 40 :	Craie du Séno-Turonien : boîtes à moustaches de InQ _s et Qs sur toute la zone	93
Figure 41 :	Craie du Séno-Turonien : boîtes à moustaches de InQ _s et Q _s (calculées sur lot d'apprentissage).	94
Figure 42 :	Craie Séno-Turonien : boîtes à moustaches de InQs et Qs (calculées sur lot de contrôle)	94
Figure 43 :	Craie du Séno-Turonien : cartes de répartition des Qs obtenues à l'issue du traitement statistique « boîte à moustache »	95
Figure 44:	Craie du Séno-Turonien : nombre de voisins pour d = 50 km	96
Figure 45 :	Craie du Séno-Turonien : calcul de Qs selon moyenne glissante avec d = 50 km.	97
Figure 46:	Craie du Séno-Turonien : calcul de Qs selon médiane glissante avec d = 50 km.	97
Figure 47 :	Craie du Séno-Turonien : nombre de voisins pour d = 40 km	98
Figure 48 :	Craie du Séno-Turonien : calcul de Qs selon moyenne glissante avec d = 40 km.	98
Figure 49 :	Craie du Séno-Turonien : calcul de Qs selon médiane glissante avec d = 40 km.	99
Figure 50 :	Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n = 338 points.	100
Figure 51 :	Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n = 160 points.	101
Figure 52 :	Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n = 80 points.	101
Figure 53 :	Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n = 40 points.	102
Figure 54:	Variogramme estimé (rouge fin) et modèle de variogramme γ_1 (rouge épais). Les nombres indiquent les quantités de doublets de valeurs	103
Figure 55 :	Modèle de variogramme γ ₂	104
Figure 56 :	Krigeage lognormal de Q _s selon γ ₁ ,voisinage unique.	105
Figure 57 :	Krigeage lognormal de Q _s , selon γ ₁ ,voisinage à 40 points	105
Figure 58 :	Krigeage lognormal de Q _s , selon γ ₁ , voisinage à 20 points	106
Figure 59 :	Krigeage lognormal de Q _s , selon γ ₁ , voisinage à 100 km.	107
Figure 60 :	Krigeage lognormal de Qs, selon γ ₁ , voisinage à 50 km	107
Figure 61:	Krigeage lognormal de Q _s selon γ ₂ , voisinage unique	108

Figure 62 :	Krigeage lognormal de Q_s , selon γ_2 , voisinage à 40 points	108
Figure 63 :	Krigeage lognormal de Q _s , selon γ ₂ , voisinage à 20 points	109
Figure 64 :	Krigeage lognormal de Q _s , selon γ ₂ , voisinage à 100 km	109
Figure 65 :	Krigeage lognormal de Qs, selon γ ₂ , voisinage à 50 km.	110
Figure 66 :	Validation croisée selon γ1.	. 111
Figure 67 :	Validation croisée selon γ2.	.111
Figure 68 :	Écart-type sur les simulations de lnQ _s .	113
Figure 69 :	Estimation de Q_s par exponentielle de la moyenne des simulations (moyenne géométrique).	. 114
Figure 70 :	Estimation de Qs par exponentielle de la médiane des simulations (≃ médiane)	. 114
Figure 71 :	Estimation de Q_s par moyenne des exponentielles des simulations (moyenne directe).	. 115
Figure 72 :	Estimation de q75(Q _s).	116
Figure 73 :	Proportion de simulations dépassant 5,556 ε−4 m²/s	116
Figure 74 :	Craie du séno-turonien : comparaison de cartes obtenues par différentes méthodes	.118
Figure 75 :	Calcaires de beauce : boîtes à moustaches de $\ln Q_s$ et Q_s sur toute la zone	121
Figure 76 :	Calcaires de Beauce - médiane mobile directe de Qs	122
Figure 77:	Calcaires de Beauce - moyenne mobile directe de Q _s .	123
Figure 78 :	Calcaires de Beauce - médiane mobile sur les logarithmes de Q _s	123
Figure 79 :	Calcaires de Beauce - moyenne mobile sur les logarithmesde Q _s	124
Figure 80 :	Calcaires de Beauce - variogramme estimé de InQs.	125
Figure 81 :	Calcaires de Beauce - variogramme estimé de InQs à très petite distance	125
Figure 82 :	Calcaires de Beauce - Modèle de variogramme Y1.	126
Figure 83 :	Calcaires de Beauce - Modèle de variogramme Y2.	127
Figure 84 :	Calcaires de Beauce - Krigeage lognormal de lnQs selon le variogramme γ_1,\dots	128
Figure 85 :	Calcaires de Beauce - Krigeage lognormal de lnQs selon le variogramme γ_2, \ldots	128
Figure 86 :	Sables du Cénomanien : boîtes à moustaches de lnQ_s et Q_s sur toute la zone	132
Figure 87 :	Sables du Cénomanien : surfaces à moins de 10 km (vert), ou à 10~20 km (jaune) des points de données	. 133
Figure 88 :	Sables du Cénomanien : moyenne mobile de Q_{s} sur voisinages de 10 km (gauche) et de 20 km (droite)	. 134
Figure 89 :	Sables du Cénomanien : médiane mobile de Q _s sur voisinages de 10 km (gauche) et de 20 km (droite)	. 134
Figure 90 :	Sables du Cénomanien - Variogramme estimé sur les 204 données	135
Figure 91 :	Sables du Cénomanien - histogramme de InQs, en rouge les points négligés	136

Figure 92 :	Sables du Cénomanien - Variogramme estimé sur 177 points	.136
Figure 93 :	Sables du Cénomanien - Variogramme estimé et modèle de variogramme $\gamma_1. \$	137
Figure 94 :	Sables du Cénomanien - Variogramme estimé et modèle de variogramme $\gamma_2. \dots$.138
Figure 95 :	Sables du Cénomanien - Krigeage lognormal selon le modèle de variogramme γ ₁	.139
Figure 96 :	Sables du Cénomanien - Krigeage lognormal selon le modèle de variogramme γ ₂	.139
Figure 97 :	Validation croisée du krigeage selon γ ₂	.140
Figure 98 :	Validation croisée variogramme γ ₂	.141
Figure 99 :	Écart-type relatif du krigeage lognormal (variogramme γ2)	142

Liste des tableaux

Tableau 1 :	Avantages et inconvénients de 4 méthodes (statistiques) de cartographiede la	00
	productivité d'aquifère.	23
Tableau 2 :	Classification et indexation des critères de potentialité géothermique	29
Tableau 3 :	Aquifères de la Région Centre retenus pour l'étude.	69
Tableau 4 :	Distribution par ensemble aquifère des populations de données de débit spécifique (Q_s) et de transmissivité (T)	71
Tableau 5 :	Sables ACF : doublons et triplet de données.	84
Tableau 6 :	Craie du Séno-Turonien : statistiques élémentaires.	93
Tableau 7 :	Comparaison des statistiques de plusieurs méthodes d'estimation	119
Tableau 8 :	Tableau des proportions d'estimations proches des valeurs réelles ;la première colonne indique la limite de proximité choisie	120
Tableau 9 :	Calcaire de Beauce : statistiques élémentaires.	121
Tableau 10	: Calcaires de Beauce - Comparaison des méthodes de cartographie	129
Tableau 11	comparaison des méthodes testées, sur un voisinage de 8 points	131
Tableau 12	: Sables du Cenomanien : statistiques élémentaires	132
Tableau 13	: Sables du Cénomanien - Comparaison des méthodes de cartographie	143
Tableau 14	: Avantages et inconvénients des méthodes de cartographiede la productivité d'aquifère	145

Liste des annexes

Annexe 1:	Traitements statistiques élémentaires de données en vue d'une cartographie géostatistique	35
Annexe 2.	Évaluation des débits spécifiquesà partir de données de transmissivité	57
Annexe 3:	Description et analyse structurale des données de la Région Centre	67
Annexe 4:	Cartographie du critère « productivité d'aquifère » : exemple de quatre méthodes de cartographie régionale appliquées à trois aquifères de Région Centre	89

1. Introduction

Pour accompagner le développement en cours de la filière géothermie très basse énergie¹ en France, le BRGM et l'ADEME se sont associés dès 2003pour élaborer un outil d'aide à la décision permettant d'évaluer l'exploitabilité régionale (ou potentiel des ressources géothermiques) des aquifères superficiels.

Cet outil est basé sur un atlas des ressources en eau souterraine pouvant être exploitées en géothermie très basse énergie Son champ d'application concerne l'exploitation énergétique des aquifères associées à la mise en œuvre de pompes à chaleur (PAC). Cet atlas comprend un système d'information géographique (SIG) mis en ligne via le site INTERNET de l'ADEME et du BRGM spécifique à la géothermie www.geothermie-perspectives.fr.

Le but est de permettre un accès du grand public, des maîtres d'œuvres (bureaux d'études, architectes) et des maîtres d'ouvrages aux cartes du potentiel géothermique afin qu'ils puissent envisager une solution géothermique lors d'un choix énergétique. Ces atlas participeront également à la définition des SRCAE² en cours et programmés suite à la loi n° 2010-788 du 12 juillet 2010, dite « Grenelle 2 ».

La réalisation de cet outil est déclinée région par région, mais la disparité des données disponibles et des contextes géologiques de chacune d'elle n'a pas toujours permis la mise en œuvre d'une méthode unique de cartographie du potentiel géothermique. Une harmonisation méthodologique est cependant proposée par le BRGM et l'ADEME. L'objectif est d'établir un socle méthodologique commun pour aboutir à la fourniture de documents de portée régionale comparables au plan national. Cette harmonisation porte sur :

- le mode d'évaluation et de représentation de la productivité des aquifères (critère essentiel à l'analyse du potentiel géothermique d'un aquifère);
- la méthode de cartographie du potentiel géothermique des aquifères (cartographie multicritères).

Document stratégique qui a vocation à définir de grandes orientations l'horizon 2020 et 2050 en matière de réduction des émissions de gaz à effet de serre, de maîtrise de la demande énergétique, de développement des énergies renouvelables, de lutte contre la pollution atmosphérique et d'adaptation au changement climatique. Il intègrera le schéma régional des énergies renouvelables et le Plan Régional de la Qualité de l'air. Les actions qui en découlent relèvent des collectivités territoriales au travers des Plans Climat Energie Territoriaux (PCET) qui devront être conformes aux orientations fixées par le SRCAE. À leur tour, les PCET seront pris en compte dans les documents d'urbanisme.

¹ Solutions énergétiques mettant en œuvre des pompes à chaleur (PAC)

² SRCAE : Schéma Régional Climat Air Énergie

Le présent rapport présente la méthodologie retenue pour l'établissement des atlas du potentiel des ressources géothermiques et les travaux qui ont permis de l'établir ainsi que des tests de différentes méthodes de cartographie de la productivité potentielle d'un aquifère à l'aide du logiciel ISATIS (de la société Géovariances)³.

Il est rappelé que les atlas s'attachent à décrire la ressource en eau souterraine et qu'ils concernent uniquement la **géothermie très basse énergie sur nappe**.

Par ailleurs, les informations fournies sont représentatives à l'échelle régionale (1/100 000). Elles revêtent donc un caractère indicatif qui doit permettre d'orienter un projet vers une solution de géothermie sur nappe ; mais les atlas ne remplacent en aucun casl'étude de faisabilité hydrogéologique (de portée locale), qui doit accompagner tout projet.

_

³ Logiciel de géostatistique

2. Principe et méthodes de réalisation de l'atlas

2.1. PRINCIPE GÉNÉRAL

2.1.1. Méthodologie générale

Le principe général de conception des atlas repose sur une méthode de cartographie à index avec pondération de critères. Cette méthode est basée sur la combinaison pondérée de cartes de paramètres (dites cartes unicritères) intervenants dans l'évaluation d'une propriété régionale (ici, le potentiel des ressources géothermales des aquifères).

Cette méthode permet la prise en compte relative de chaque critère influençant le potentiel général d'un aquifère. La principale difficulté rencontrée lors de sa mise en œuvre est la disparité régionale (tant quantitative que qualitative) des données de base utilisées pour réaliser les cartes unicritères.

La cartographie et la combinaison des cartes se fait au moyen de logiciels de traitement cartographique (SIG); les données sont traitées sous forme de grilles (ou *raster*) avec des mailles de 500 m de côté. Toutes les cartes produites sont géoréférencées.

Le schéma de la Figure 1 synthétise la méthodologie de cartographie du potentiel géothermique régional proposée.

2.1.2. Critères de potentialité géothermique

Pour rendre compte du potentiel géothermique très basse énergie des aquifères superficiels français, trois critères sont proposés :

- productivité de l'aquifère ou débit exploitable ;
- température de la ressource ;
- profondeur d'accès à la ressource.

Les deux premiers sont représentatifs du rendement de l'exploitation de l'aquifère pour un usage géothermique ; le troisième est dit « économique » car il impacte directement le coût d'investissement et le coût d'exploitation d'un projet.

Leur analyse simultanée permet de rendre compte de l'intérêt d'une opération de géothermie.

Selon les contextes régionaux, des particularités géochimiques ayant un impact sur la faisabilité de projets géothermiques peuvent exister. Elles peuvent être cartographiées, et fournies à titre informatif, mais ne sont pas intégrées au calcul multicritère car elles sont rarement disponiblesà l'échelle régionale et pour l'ensemble des aquifères.

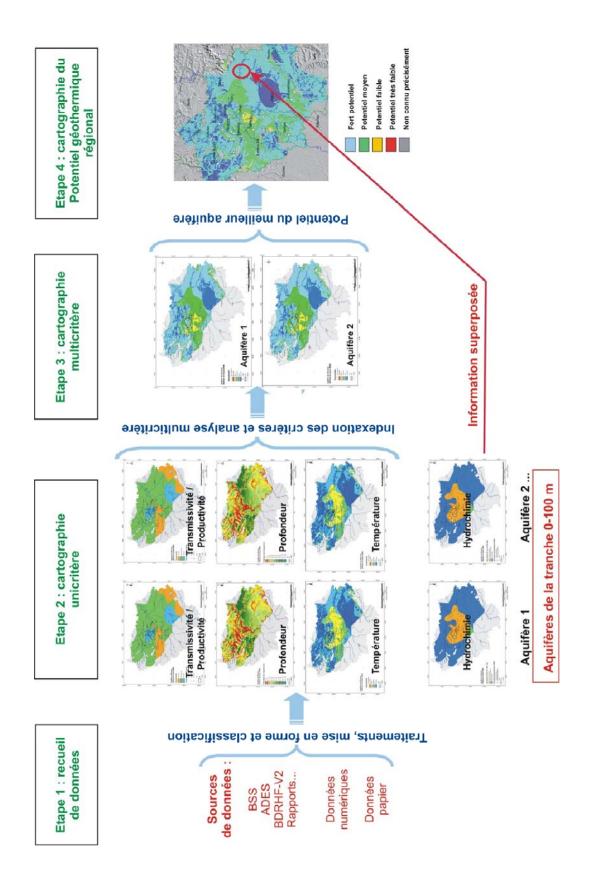


Figure 1 : Principe de l'élaboration des cartes du potentiel géothermique régional.

2.1.3. Cartographie du potentiel géothermique

L'atlas du potentiel géothermique est constitué :

- de l'ensemble des cartes unicritères de chaque aquifère superficiel d'intérêt régional ;
- d'une carte multicritère ou carte du potentiel géothermique par aquifère superficiel d'intérêt régional;
- d'une carte de synthèse générale, dite carte du potentiel géothermique régional, qui correspond à la compilation des meilleures potentialités par aquifère.

Sa réalisation repose sur quatre étapes :

- Étape 1 : recueil des données ;
- Étape 2 : cartographie unicritère ;
- Étape 3 : cartographie multicritère ;
- Étape 4 : cartographie du potentiel géothermique régional.

a) Recueil de données relatives à chaque critère

Cette étape comprend le recueil de données disponibles au format numérique dans différentes bases de données, le recueil, la numérisation et la bancarisation (mise en base de données) de données bibliographiques non numériques.

Il est important de noter que la réalisation de l'atlas repose avant tout sur le recueil de données et que la qualité des cartes produites est directement liée à la qualité et à la densité des données disponibles. La première étape de recueil des données est donc très importante ; elle permet de décider des options de cartographie à mettre en œuvre lors des étapes suivantes.

Les sources de données généralement utilisées sont :

- BSS : la Banque de données du Sous-Sol apporte des données sur la géométrie et la nature des formations géologiques, sur la piézométrie et la productivité des aquifères et sur la physico-chimie des eaux souterraines ;
- ADES : la banque nationale d'Accès aux Données sur les Eaux Souterraines qui apporte des données physico-chimiques et piézométriques complémentaires ;
- BD-LISA: la Base de Données du Référentiel Hydrogéologique Français fourni les contours des entités hydrogéologiques françaises à 3 échelles de précision (N1: nationale, N2: régionale et N3: locale);
- Carte géologique au 1/50 000 ;
- Atlas hydrogéologiques régionaux ;
- Données issues de modèles hydrodynamiques ;
- Rapports divers archivés.

b) Cartographie unicritère - classification et indexation des critères

La cartographie de chaque critère de potentialité géothermique est réalisée à partir des données recueillies précédemment et selon une grille de mailles de 500 m de côté. Elle permet d'élaborer des cartes d'indices de potentialités géothermiques comparables entre elles. Elle comprend 3 étapes :

- la classification des données : pour rendre lisibles les cartes unicritères, les données utilisées pour leur élaboration sont rangées en intervalles appelés classes (les classes à retenir sont prédéfinies dans le présent rapport ; cf. § 2.2.5.).
- l'indexation des classes : chaque classe est affectée d'un index (ou note) représentatif du potentiel géothermique associé (note faible pour le potentiel le moins bon, note élevée pour le meilleur potentiel). Les index varient de 1 à 4 de la situation la moins favorable en matière degéothermie à la situation la plus favorable ; ils sont présentés dans le Tableau 2.
- cartographie de chaque critère : des cartes dites « unicritères » sont réalisées pour rendre compte du potentiel géothermique de l'aquifère vis-à-vis de chaque critère ; la classification et l'indexation des critères y est représentée :
 - · carte de profondeur d'accès à la ressource en eau souterraine ;
 - · carte de productivité de l'aquifère (carte des débits exploitables) ;
 - · carte de température des eaux souterraines.

Quand le contexte local le nécessite, des cartes de qualité des eaux souterraines peuvent également être réalisées. Les données utilisées font l'objet d'une classification représentative de l'impact de la qualité des eaux souterraines sur leur exploitabilité géothermique, mais elles ne sont pas indexées.

La démarche de cartographie unicritère est présentée en Figure 2.

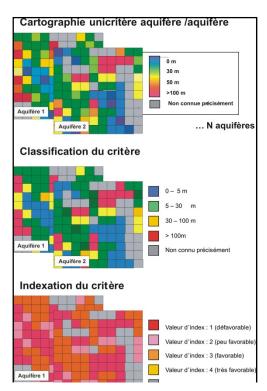


Figure 2 : Démarche de cartographie unicritère (classification et indexation).

c) Analyse multicritères : calcul du potentiel géothermique de chaque aquifère

Les cartes du potentiel géothermique de chaque aquifère sont obtenues par combinaison pondérée des cartes unicritères. Un indice de potentialité géothermique est calculé par addition pondérée des index de potentialité unicritère attribués dans chaque maille de la grille cartographique. Une carte de répartition des indices de potentialité est alors établie ; elle correspond à la carte du potentiel géothermique de l'aquifère.

La pondération des critères rend compte du fait que certains critères ont une influence plus forte que d'autres sur le potentiel géothermique de l'aquifère.

Une pondération des critères d'évaluation du potentiel géothermique des aquifères est proposée au chapitre 3.3.

Le principe du calcul multicritère est présenté en Figure 3.

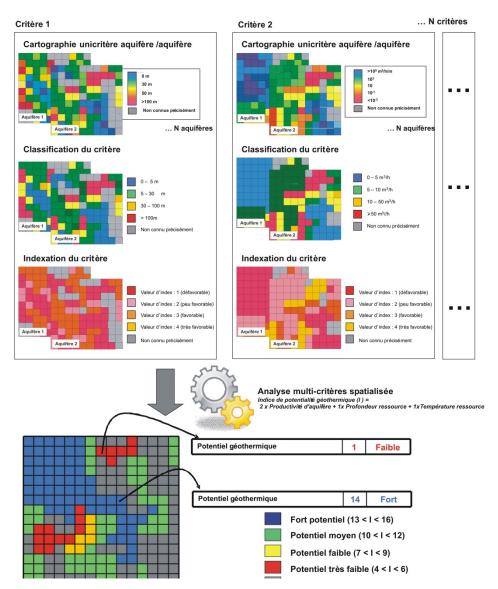


Figure 3 : Principe de l'analyse multicritère.

d) Élaboration de la carte du potentiel géothermique régional présentant les secteurs favorables à l'implantation de pompes à chaleur sur nappe.

L'analyse multicritère conduit à l'élaboration d'une carte du potentiel géothermique par aquifère.

L'obtention d'une carte unique du potentiel géothermique régional passe par la sélection, pour chaque maille de la grille cartographique, des meilleurs indices de potentialité géothermique parmi ceux calculés pour chaque aquifère.

2.2. CRITÈRES DE POTENTIALITÉ GÉOTHERMIQUE, SOURCES DE DONNÉES, CLASSIFICATION ET INDEXATION

2.2.1. Critère « Productivité d'un aquifère »

Sources de données

La productivité d'aquifère correspond à l'évaluation de la ressource en eau extractible du réservoir aquifère. Les données de base utilisées pour l'évaluer sont :

Transmissivité (notée T) et perméabilité (notée K)

Il s'agit de paramètres intrinsèques à l'aquifère (représentant sa capacité de production), très représentatifs de la ressource.

Les données de perméabilité doivent être couplées aux données d'épaisseur d'aquifère (e) pour reconstituer les données de transmissivité (T = K × e).

Les données de base sont issues des rapports de pompages d'essai (souvent archivés en BSS).

La densité de ces données intrinsèques est souvent faible, ce qui rend difficile leur exploitation pour une cartographie régionale.

Débits spécifiques (noté Q_s)

C'est le rapport du débit pompé dans un ouvrage à la hauteur du rabattement de nappe provoqué par ce pompage.

Ce paramètre est représentatif des propriétés de productivité de l'aquifère, mais il dépend également des propriétés techniques des ouvrages de captage. Il est donc moins représentatif de la ressource que la transmissivité, mais les données qui permettent de le cartographier sont la plupart du temps beaucoup plus nombreuses.

Les données de base sont l'ensemble des mesures ponctuelles de débits/rabattement disponibles en BSS. Leurs conditions de mesure ne sont pas toujours connues, ce qui rend leur précision aléatoire. De plus, elles ne sont pas forcément associées à des pompages d'essais complets, ce qui ajoute à leur imprécision.

Débits de production de forages

Ce paramètre correspond à des débits exploitables, mais est surtout représentatif des propriétés techniques des ouvrages de captage et des besoins associés aux usages de la ressource. Il est peu représentatif de la capacité de production de la ressource proprement dite.

Traitement des données de base : passage des données de base à des données de débits potentiellement exploitables

En règle générale, les partenaires des projets et futurs utilisateurs des atlas souhaitent que la productivité de l'aquifère, quelles que soient les données de bases prises en compte pour son évaluation, soit représentée sous forme de débits potentiellement exploitables.

Une carte de répartition des débits potentiellement exploitables doit donc être fournie dans les atlas régionaux.

Sa réalisation passe par le traitement des données de productivité et leur interprétation en termes de débits potentiellement exploitables. Théoriquement, différentes approches sont possibles selon les données de base prises en compte :

Données de transmissivité

Elles peuvent se traduire en débits exploitables par la recherche d'un lien du type :

$$Q_s = \alpha T(1)$$

Puis par la relation :

$$Q_{exploit} = Q_s \times s$$
 (s = rabattement).

Données de débit spécifique

Elles peuvent se traduire en débit exploitable par la relation :

$$Q_{exploit} = Q_s \times s$$

Données de débit de production de forages

Ces données sont directement utilisables pour représenter la répartition de l'exploitation d'un aquifère.

Recommandations relatives au choix des données de base

Des tests effectués à partir des données disponibles sur les aquifères de la Région Centre sont présentés en annexe 2. Ils montrent que, d'une façon générale, il n'est pas préférable d'utiliser les données intrinsèques (T, K, e) de l'aquifère car :

- même si une corrélation peut être établie entre les paramètres T et Qs, l'erreurassociée est souvent importante, du fait de la dispersion des données ;
- une cartographie des données de « productivité d'aquifère » obtenues à partir de données intrinsèques ne serait possible que dans le cas où celles-ci seraient disponibles en grande quantité et correctement réparties sur l'étendue de l'aquifère ; ce qui, dans les faits, n'est pratiquement jamais le cas.

Les données intrinsèques doivent donc être privilégiées dans les cas où :

- elles sont disponibles en grande quantité et correctement réparties sur l'étendue de l'aquifère ;
- leur répartition est évaluée en référence à d'autres paramètres (ex : carte issue d'une modélisation).

Dans les autres cas, il est recommandé d'utiliser les données de débit spécifiques après les avoir triées et avoir sélectionné les plus représentatives des aquifères.

En dernier recours, en cas d'indisponibilité de données intrinsèques et de Q_s , les données de débits exploités pourront être prises en compte si leur nombre est très important et si elles sont bien réparties régionalement.

Recommandation relative au rabattement

Les deux premières méthodes d'interprétation des données de base impliquent de définir un rabattement à prendre en compte pour le calcul du débit potentiellement exploitable.

De façon à standardiser la méthode d'évaluation de la productivité des aquifères et à éviter les disparités régionales, il est proposé :

- de limiter ce rabattement aux seuils fixés par la réglementation régionale quand ceux-ci existent ;
- de le limiter à 1/3 de l'épaisseur mouillée d'aquifère (seuil de rabattement admissible pour un aquifère généralement admis) et à un maximum de 5 m (pour les aquifères d'épaisseur mouillée supérieure à 15 m) quand aucun seuil local de rabattement n'est défini.

Recommandations pour le choix de la méthode de régionalisation du critère :

Différentes approches sont possibles pour régionaliser le paramètre productivité d'aquifère (établir des cartes régionales à partir de données ponctuelles) :

- Approche globale à dire d'expert : elle consiste à attribuer une valeur unique (ou un intervalle de valeurs unique) à un polygone représentatif d'une entité hydrogéologique sur la base de l'expertise d'un hydrogéologue local;
- Approches statistiques globales : elles consistent à attribuer une valeur unique de productivité potentielle (ou un intervalle de valeurs unique) à un polygone représentatif d'une entité hydrogéologique sur la base d'un traitement statistique de données ponctuelles :
- Approches statistiques locales (statistique mobile et géostatistiques) : elles consistent, pour une entité hydrogéologique donnée, à représenter les variations du paramètre entre les points disposant de données mesurées. Elles permettent d'attribuer une valeur à chaque maille de la carte.

Une analyse préalable des données (nombre de points de mesure, répartition, homogénéité, structure du variogramme...) telle que présentée en annexe 1 permettra d'orienter le choix de la méthode de régionalisation :

- pour que la régionalisation⁴ d'un paramètre à partir d'une approche statistique (globale ou locale) soit valable, il est préférable :
 - · que la distribution des valeurs suive une loi normale (information tirée de l'histogramme et de la droite de Henry),
 - que les valeurs du paramètre soient les plus continues possible en termes de continuité spatiale (information tirée du variogramme) ;
- pour que l'approche statistique locale soit applicable, il faut en plus que les données répondent à la condition suivante :la distance entre les points de mesure doit être inférieure à la (plus petite) portée du variogramme;
- quand la quantité et la répartition des données ne permettent pas une approche statistique (globale ou locale), la régionalisation à dire d'expert pur pourra être envisagée;
- dans le cas contraire, les approches statistiques locales doivent être privilégiées car elles rendent mieux compte de la variabilité spatiale de la productivité que les méthodes statistiques globales.

Quatre méthodes de cartographie régionale du critère « productivité d'aquifère », appliquées aux données de la Région Centre, sont présentées en annexe 1 (description des méthodes) et testées en annexe 4. Il ressort des avantages et inconvénients inhérents à chacune d'elles, résumés dans le Tableau 1.

D'autres approches existent ; elles ne sont pas à exclure, mais un compromis entre la précision possible au vu des données disponibles et le temps de traitement doit être trouvé.

Il revient toujours à l'expert hydrogéologue de choisir, en fonction de sa connaissance des systèmes hydrogéologiques :

- le mode de cartographie du critère « productivité d'aquifère » (dire d'expert ou statistique) ;
- les méthodes et paramètres de cartographie statistique (méthode statistique globale, statistique mobile, simulation statistique, nature des variables (normale ou lognormale), voisinage de calcul...).

Il est également important de toujours préciser dans les rapports la nature des données de base utilisées et de fournir une carte de localisation des points de mesure de ces données.

⁴ Construction d'une carte régionale d'un paramètre à partir de données mesurées ponctuellement

Méthode de cartographie	Avantages	Inconvénients	
Dire d'expert	Mise en œuvre très simple Approche qualitative intéressante en l'absence de données mesurées	Approche relativement subjective, non reproductible	
Statistique Mise en œuvre très simple Résultat simple à représenter		Ignore totalement les fluctuations locales ou régionales du paramètre étudié Biaisée par la répartition des mesures ⁵ Dans le cas d'une répartition hétérogène des données, peut conduire à des résultats fortement biaisés	
Statistique mobile	Prise en compte des points de mesure, d'autant plus forte qu'ils sont proches du point d'estimation Mise en œuvre relativement simple	Résultat dépendant du choix du voisinage de calcul Lissage du résultat si la distance entre les points de mesures est très grande devant la portée (estimée d'après le variogramme)	
Krigeage	Prise en compte des points de mesures selon leur distance entre eux, donc selon leur répartition spatiale Prise en compte des phénomènes régionalisés quand ils existent (structure de données dont la portée est supérieure à la maille d'échantillonnage) Résultats meilleurs que les statistiques mobile lorsque la densité d'échantillonnage est suffisante vis-àvis de la variabilité spatiale du phénomène	Résultat dépendant du choix du voisinage de calcul Besoin de calculer et modéliser le variogramme, donc de compétences particulières en géostatistiques	
Simulations géostatistiques	Idem que krigeage Fournit en tout point, non seulement une estimation (moyenne des possibilités), mais une distribution des valeurs possibles, donc une représentation de leur dispersion autour de l'estimation. Peut être utilisé pour établir une carte de probabilité de dépasser un débit seuil, ou une carte de débit dépassé avec une probabilité seuil	Idem que krigeage S'applique à une variable gaussienne Technique plus lourde à mettre en œuvre que le krigeage	

Tableau 1 : Avantages et inconvénients de 4 méthodes (statistiques) de cartographie de la productivité d'aquifère.

⁵Les données peuvent être regroupées en amas dans les zones connues pour leur productivité élevée. Si le reste de l'aquifère présente un potentiel inférieur, la moyenne des mesures ne le reflètera pas.

Mode de représentation : passage des données de débits exploitables à des données de puissances extractibles

De nombreux utilisateurs souhaiteraient disposer de cartes d'exploitabilité géothermique traduites en termes de puissances extractibles. Cela supposerait de traduire la carte de productivité d'aquifère en cartes de puissances-chaud extractibles et de puissances-froid extractibles via les relations suivantes :

- puissance chaud : P = Q_{exploit}×∆T×1,16/(1-1/COP_{chaud});
- puissance froid : $P = Q_{exploit} \times \Delta T \times 1,16/(1+1/COP_{froid})$.

Ces calculs impliquent de fixer les paramètres ΔT , COP_{chaud} et COP_{froid} à prendre en compte. Or, ces paramètres sont propres au dimensionnement d'une installation qui répond à un contexte très ponctuel (ressource disponible et besoin énergétique). Leur généralisation régionale n'aurait pas de sens ; elle conduirait à la production de cartes de potentiel géothermiques inadaptées aux opérations dont les paramètres ΔT , COP_{chaud} et COP_{froid} diffèrent de ceux retenus pour les calculs.

De telles cartes ne doivent donc pas être proposées.

Classification

La classification suivante est proposée pour le critère productivité d'aquifère:

- Q < 5 m³/h : le risque d'échec de production est élevé. La ressource est considérée comme faible ;
- 5 m³/h < Q < 10 m³/h : cette gamme de débit permet d'envisager le chauffage de logements individuels ;
- 10 m³/h < Q < 50 m³/h : cette gamme de débits permet d'envisager le chauffage de bâtiments de 1 000 à 5 000 m² ;
- $50 \text{ m}^3/\text{h} < Q < 100 \text{ m}^3/\text{h}$: cette gamme de débits permet d'envisager le chauffage de bâtiments de 5 000 à 10 000 m²⁶ ;
- Q > 100 m³/h : cette gamme de débit permet d'envisager le chauffage de grands bâtiments, voire de micro-réseaux de chaleur.

_

⁶ Valeurs données pour des bâtiments existants ; pour des bâtiments BBC ou RT2012 on considère que l'on peut multiplier la surface par 2.

2.2.2. Critère température de la ressource

Sources de données :

- Mesures ponctuelles en forage. Attention à la qualité des données; des erreurs liées au mode d'acquisition du paramètre sont possibles (équilibre entre l'eau du forage et la température atmosphérique).
- Estimation en fonction de la profondeur de l'aquifère, à partir du gradient géothermique local.
- Estimation en fonction de l'altitude du point d'eau.

Classification

Une classification de ce critère, basé sur l'évolution du degré d'efficacité d'une PAC utilisée en mode chauffage plus climatisation (PAC réversible)⁷ en fonction de la température de la ressource, est proposée :

- T < 10°C : difficultés pour le fonctionnement d'une PAC en mode chauffage ;
- 10°C ≤ T < 15°C : très favorable au fonctionnement d'une PAC en mode chauffage et climatisation et au free-cooling ;
- T > 15°C : difficulté pour le fonctionnement d'une PAC en mode climatisation.

2.2.3. Critère profondeur d'accès à la ressource

La profondeur d'accès à la ressource dans le cas d'un aquifère libre correspond à la piézométrie de la nappe, tandis qu'elle correspond à la profondeur du toit de l'aquifère considéré dans le cas d'un aquifère captif.

Ce critère rend compte de deux aspects de la potentialité géothermique d'un aquifère :

- le coût de l'opération : coûts de foration (directement lié à la profondeur des forages), coûts d'équipement (type et puissance de pompe liée à la hauteur de refoulement) et coûts de fonctionnement (consommation électrique de la pompe) :
- le risque que la surcote piézométrique provoquée par l'injection, dans le même aquifère, de l'eau prélevée, soit défavorable au projet.

Sources de données

- Données ponctuelles de piézométrie disponibles en BSS, dans ADES, et dans divers rapports ;
- Carte piézométriques existantes ;

⁷ Ce critère dépend de l'utilisation faite de la ressource géothermale ; par exemple, si l'on se contente du mode chauffage, l'aspect favorable augmentera avec la température.

- Géométrie des formations géologiques : BSS, modèles géologiques, BDRHF-V2

Classification

La classification suivante est proposée pour ce critère :

- 0 < Prof < 5 m : risque de surcote piézométrique à l'injection défavorable au projet ;
- 5 m< Prof < 15 m : forage de moindre coût permettant la réalisation économique d'un projet de PAC sur nappe pour un pavillon ;
- 15 m< Prof < 30 m : le coût du forage pèse fortement sur l'économie d'un projet de PAC sur nappe pour un pavillon ;
- 30 m
 Prof < 100 m : le coût du forage ne peut être amorti que par une consommation en surface adaptée (petit collectif ou ensemble de plusieurs pavillons);
- Prof> 100 m : le coût du forage ne peut être amorti que par une forte consommation (petit collectif ou ensemble de plusieurs pavillons).

2.2.4. Qualité de la ressource

Dans le cadre de l'évaluation du potentiel géothermique, le critère qualité chimique de la ressourcerend compte du caractère corrosif, encroûtant ou colmatant d'une eau souterraine. Il joue un rôle sur l'évaluation économique d'un projet car les typologies des matériels utilisés et de la maintenance doivent être adaptées à la chimie des eaux.

Les phénomènes mis en jeu sont complexes (réactions chimiques de précipitation, dissolution) et font intervenir une gamme importante de paramètres du milieu qui ne sont pas toujours disponibles pour tous les aquifères (une forte hétérogénéité existe d'un aquifère à l'autre et d'une région à l'autre quant à la disponibilité des données).

Les données utilisées sont les données d'analyses des eaux souterraines disponibles en BSS, dans ADES, et dans divers rapports.

Une gamme de paramètres d'évaluation des contraintes d'exploitation liées à la qualité des eaux ainsi que leur classification est définie dans le SEQ Eaux souterraines (les seuils proposés proviennent d'avis d'experts basés sur la thermodynamique des équilibres électrochimiques et/ou la bibliographie jugée comme faisant référence en ce domaine). Il est proposé de les retenir (quand les données sont disponibles) pour l'évaluation des contraintes liées à la géochimie des eaux souterraines et susceptibles de peser sur les projets d'exploitations géothermiques de très basse énergie.

-

⁸ SEQ-Eaux souterraines : grille d'évaluation de l'aptitude des eaux souterraines à différents usages et fonctions établie dans le cadre d'une étude inter-agences de l'eau et en collaboration entre Agences de l'Eau, MATE et BRGM en 2001.

Paramètres de caractérisation du pouvoir corrosif des eaux souterraines :

- CO₂ dissous
 - · [CO₂ dissous] < 50 mg/l : aptitude à satisfaire l'usage bonne ;
 - · 50 mg/l < [CO₂ dissous] < 120 mg/l : aptitude à satisfaire l'usage passable ;
 - · 120 mg/l < [CO₂ dissous] < 200 mg/l : aptitude à satisfaire l'usage mauvaise ;
 - · [CO₂ dissous] > 200 mg/l : eau brute inapte à satisfaire l'usage \rightarrow traitement ou équipement adapté nécessaire.
- O2 dissous
 - · [O₂ dissous] = 0 mg/l : aptitude à satisfaire l'usage très bonne ;
 - · 0 mg/l < $[O_2 \text{ dissous}] \le 0.1 \text{ mg/l}$: aptitude à satisfaire l'usage bonne ;
 - · 0.1 mg/l < $[O_2 \text{ dissous}] \le 4 \text{ mg/l}$: aptitude à satisfaire l'usage mauvaise ;
 - · 4 mg/l < [O₂ dissous] ≤ 8 mg/l : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.

Salinité

- · [NaCl] < 0.6 g/l : aptitude à satisfaire l'usage bonne ;
- · 0.6 g/l < [NaCl] ≤ 1.5 g/l : aptitude à satisfaire l'usage passable ;
- · 1.5 g/l < [NaCl] ≤ 3 g/l : aptitude à satisfaire l'usage mauvaise ;
- · [NaCl] > 3 g/l : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.

- Conductivité

- · C < 1300 mS/cm : aptitude à satisfaire l'usage bonne ;
- · 1300 mS/cm < C ≤3000 mS/cm: aptitude à satisfaire l'usage passable ;
- · 3000 mS/cm < C ≤ 6000 mS/cm : aptitude à satisfaire l'usage mauvaise ;
- · C > 6000 mS/cm : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.

- pH

- · pH > 9.8 : aptitude à satisfaire l'usage très bonne ;
- · 7 < pH ≤ 9.8 : aptitude à satisfaire l'usage bonne ;
- 6 ≤ pH ≤ 7 : aptitude à satisfaire l'usage passable ;
- pH < 6 : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.

- Chlorures

- · [Cl] < 150 mg/l: aptitude à satisfaire l'usage bonne;
- · 150 mg/l < [Cl⁻] < 400 mg/l : aptitude à satisfaire l'usage passable ;
- · 400 mg/l < [Cl⁻] < 1000 mg/l : aptitude à satisfaire l'usage mauvaise ;
- · [Cl⁻] > 1000 mg/l: eau brute inapte à satisfaire l'usage > traitement ou équipement adapté nécessaire.

- Sulfates

- $\cdot [SO_4^{2-}] < 250 \text{ mg/l}$: aptitude à satisfaire l'usage bonne;
- \cdot 250 mg/l < [SO₄²⁻] < 500 mg/l : aptitude à satisfaire l'usage passable ;
- · 500 mg/l < $[SO_4^2]$ ≤ 1500 mg/l : aptitude à satisfaire l'usage mauvaise ;
- \cdot [SO₄²] > 1500 mg/l : eau brute inapte à satisfaire l'usage \rightarrow traitement ou équipement adapté nécessaire.

- Ferro-bactéries

- · Absence : aptitude à satisfaire l'usage très bonne ;
- · Présence : eau brute inapte à satisfaire l'usage -> traitement ou équipement adapté nécessaire.

- Bactéries sulfato-réductrices

- · Absence : aptitude à satisfaire l'usage très bonne ;
- · sulfato-réductrices < 10 N/ml : aptitude à satisfaire l'usage bonne ;
- · 10 N/ml < sulfato-réductrices < 100 N/ml : aptitude à satisfaire l'usage mauvaise ;
- · sulfato-réductrices > 100 N/ml : eau brute inapte à satisfaire l'usage →⇒ traitement ou équipement adapté nécessaire.

- Sulfures

- · [HS⁻] < 0.1 mg/l : aptitude à satisfaire l'usage bonne ;
- · 0.1 mg/l < [HS] < 8 mg/l : aptitude à satisfaire l'usage passable ;
- · 8 mg/l < [HS⁻] < 50 mg/l : aptitude à satisfaire l'usage mauvaise ;
- · [HS⁻] > 50 mg/l : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.
- Eh (indice d'existence des ions et des précipités pour le fer)
 - · 0 mV ≤ Eh < -600 mV : aptitude à satisfaire l'usage très bonne ;
 - · -600 mV < Eh ≤ -500 mV : aptitude à satisfaire l'usage passable ;
 - · -500 mV < $[O_2$ dissous] ≤ -400 mV : aptitude à satisfaire l'usage mauvaise ;
 - · -400 mV < $[O_2$ dissous $] \le 0$ mV : eau brute inapte à satisfaire l'usage \rightarrow traitement ou équipement adapté nécessaire.

Paramètres de caractérisation du pouvoir encroûtant ou colmatant des eaux souterraines :

- pH
 - pH < 5 : aptitude à satisfaire l'usage très bonne ;
 - · pH > 5 : aptitude à satisfaire l'usage passable.
- Eh (indice d'existence des ions et des précipités pour le fer)
 - · Quelle que soit la valeur de Eh, si pH ≤ 3.5 : aptitude à satisfaire l'usage très bonne ;
 - · Si 3,5 < pH \leq 9,6 etEh > -0,166pH+1.333 : aptitude à satisfaire l'usage très bonne ;

- · Si 3,5 < pH ≤ 9,6 etEh < -0,166pH+1.333 : aptitude à satisfaire l'usage passable ;
- · pH > 9,6 et -750 mV < Eh < -250 mV : aptitude à satisfaire l'usage passable ;
- O₂ dissous
 - · [O₂ dissous] < 0,1mg/l : aptitude à satisfaire l'usage très bonne ;
 - \cdot 0,10 mg/l < [O₂ dissous] < 5,5 mg/l : aptitude à satisfaire l'usage passable ;
 - \cdot [O₂ dissous] > 5,5 mg/l: eau brute inapte à satisfaire l'usage \rightarrow traitement ou équipement adapté nécessaire.
- Ferro bactéries
 - · Absence : aptitude à satisfaire l'usage très bonne ;
 - · Présence : aptitude à satisfaire l'usage passable.
- Indice de saturation (selon la valeur du TAC)
 - · TAC <10°f et indice de Langellier < 0,2 : aptitude à satisfaire l'usage très bonne ;
 - · TAC <10°f et 0.2< indice de Langellier < 2 : aptitude à satisfaire l'usage passable ;
 - · TAC <10°f et indice de Langellier > 2 : inaptitude à satisfaire l'usage ;
 - TAC <10°f et indice de Langellier < 0.2 : aptitude à satisfaire l'usage très bonne ;
 - · TAC <10°f et 0,2 < indice de Langellier < 1 : aptitude à satisfaire l'usage passable ;
 - TAC <10°f et indice de Langellier > 1 : eau brute inapte à satisfaire l'usage → traitement ou équipement adapté nécessaire.

2.2.5. Bilan de la classification et de l'indexation des critères

La classification proposée pour les différents critères de potentialité géothermique ainsi que leur indexation est présentée dans le Tableau 2.

Critère de potentialité géothermique	Classification de critères	Indexation des classes	Potentialité
	$Q < 5 \text{ m}^3/\text{h}$	1	Défavorable
	$5 \text{ m}^3/\text{h} < Q < 10 \text{ m}^3/\text{h}$	2	Peu favorable
Productivité de l'aquifère	$10 \text{ m}^3/\text{h} < Q < 50 \text{ m}^3/\text{h}$	3	Favorable
	50 m ³ /h < Q < 100 m ³ /h	4	Très favorable
	Q > 100 m ³ /h	4	Très favorable
	T <10°C	2	Peu favorable
Température de la ressource	10°C ≤ T < 15°C	4	Très favorable
	T > 15°C	3	Favorable
	0 < Prof < 5 m	1	Défavorable
Durafa u da con dia a a barbira	5 m< Prof < 15 m	4	Très favorable
Profondeur d'accès à la	15 m< Prof < 30 m	3	Favorable
ressource	30 m< Prof < 100 m	2	Peu favorable
	Prof> 100 m	1	Défavorable

Tableau 2 : Classification et indexation des critères de potentialité géothermique

2.3. ÉVALUATION DU POTENTIEL GÉOTHERMIQUE : CHOIX DU CODE DE CALCUL MULTICRITÈRE

Après avoir procédé au recueil des données de base et réalisé des cartes unicritères pour chaque aquifère, il est procédé à une analyse simultanée de ces cartes par le biais d'un calcul automatisé multicritère, mené sous SIG. Ce calcul aboutit à une carte du potentiel géothermique pour chaque aquifère.

Les différents travaux réalisés en région montrent que des disparités régionales portant sur le code de calcul multicritère (c'est-à-dire sur la pondération des critères) existent.

Dans le but d'homogénéiser le contenu des atlas régionaux, il est proposé ici un code de calcul multicritère unique qui attribue à la productivité d'aquifère une importance de premier ordre vis-à-vis du potentiel géothermique de l'aquifère :

Potentiel géothermique = 2 × Productivité d'aquifère + 1 × Température + 1 × Profondeur d'accès

2.3.1. Représentation : choix de l'échelle de restitution des cartes

La maille de cartographie des critères et de calcul multicritères est de 500 m.

Malgré cela, et du fait des différentes approximations réalisées (approximations liées au zonage des critères et aux calculs de productivité), l'échelle de restitution des informations aux utilisateurs ne doit pas excéder le 1/100 000.

Pour limiter les zooms abusifs sur les cartes consultables via le site géothermieperspectives, un seuil de zoom limitant la visualisation des cartes du potentiel géothermique à quatre mailles (échelle du 1/100 000) doit être mis en place.

3. Conclusion

Une méthodologie d'évaluation du potentiel géothermique très basse énergie, dont le champ d'application concerne l'exploitation énergétique des aquifères superficiels associés à la mise en œuvre de pompes à chaleur (PAC) est proposée. Son champ d'application concerne l'exploitation énergétique des aquifères associée à la mise en œuvre de pompes à chaleur (PAC).

Elle aboutit à l'élaboration d'un outil d'aide à la décision répondant à un objectif de diffusion de l'information qui doit uniquement être utilisé comme tel. Il ne s'agit en effet :

- ni d'un outil de gestion des ressources en eau souterraine (dans le sens où il ne prend pas en compte toute une série de critères propres à ce type de gestion (exploitations existantes, périmètres d'influence de l'existant et du projeté, contraintes réglementaires, contraintes naturelles,...) et qu'il n'est pas dynamique : évolutif dans le temps en fonction de l'évolution de l'existant);
- ni d'un outil d'évaluation de la faisabilité de projets du fait de la limitation de son échelle de validité au 1/100 000.

La méthode présentée ici se base sur une cartographie à index avec pondération de critères de potentialité géothermique. Les critères sont la productivité de l'aquifère, la température de la ressource et la profondeur d'accès à la ressource.

Des recommandations sont émises :

- pour la classification à adopter pour chaque critère ;
- pour le choix des paramètres à prendre en compte pour cartographier ces critères (en particulier pour le critère « productivité d'aquifère »);
- pour le choix du mode de cartographie du critère « productivité d'aquifère » ;
- pour le choix du code de calcul du potentiel géothermique par aquifère à retenir.

Le critère qualité chimique de la ressource n'est pas intégré au calcul du fait de la fréquente indisponibilité de données permettant de la qualifier. Il est seulement recommandé de fournir des éléments caractéristiques (une liste de paramètres guides est fournie) des contraintes liées à la géochimie des eaux souterraines quand l'existence de ces contraintes est avérée car elles sont susceptibles de peser sur les projets d'exploitations géothermiques.

La méthode proposée reste très théorique et des adaptations locales (critères, pondération...) pourront être retenues en réponse à des configurations et problématiques régionales particulières et sur avis d'un expert hydrogéologue local.

Pour aller plus loin et répondre aux attentes récentes des acteurs de la géothermie, des évolutions de cette méthode devraient être envisagées :

- production de cartes du potentiel géothermiques distinctesselon l'usage projeté (chauffage ou climatisation). Sur les cartes actuelles, le potentiel géothermique concerne les deux usages, ce qui implique une minimisation du potentiel en mode chauffage ;
- représentation des contraintes locales susceptibles d'impacter la faisabilité d'un projet géothermique (ou son coût) : exploitations existantes, risques géologiques à la réalisation de forages, contraintes réglementaires...

32

4. Bibliographie

Système d'évaluation de la qualité des eaux souterraines. Rapport de présentation – version provisoire (janvier 2001).

- **Jaudin F., Martin J.C.** (2005) Atlas sur la géothermie très basse énergie en Région Centre Rapport intermédiaire. BRGM/RP-54340-FR, 50 p., 5 ill., 1 ann.
- **Martin J.C., Printemps J., Remaud S.** (2006) Atlas sur la géothermie très basse énergie en Région Centre Étude cartographique et statistique des débits spécifiques des forages d'eau.— Rapport intermédiaire. BRGM/RP-54542-FR, 76p., 35 ill., 2 ann.
- **Martin J.C., Bacquet I., Tourlière B.** (2006) Atlas sur la géothermie très basse énergie en Région Centre Géométrie et niveaux piézométrique des principales formations aquifères Rapport intermédiaire. BRGM/RP-54715-FR, 73p., 36 ill., 4 ann.
- **Martin J.C., Bacquet I., Tourlière B., Gomez E., Gateau C.** (2007) Atlas sur la géothermie très basse énergie en Région Centre Rapport final. BRGM/RP-55088-FR, 60 p., 25 fig., 9 tabl., 1 ann.
- **B.** Bourgine, L. Denis, T. Filhine-Tresarrieu, P. Monnot, D. Nguyen-Thé, C. Robelin (2007) Atlas du potentiel géothermique des aquifères lorrains. Rapport BRGM/RP-54987-FR, 83 p., 36 ill., 3ann.
- **Schomburgk S., Goyénèche O.***et al.* (2004) guide d'aide à la décision pour l'installation de pompes à chaleur sur nappe aquifère en région lle-de-France atlas hydrogéologique BRGM/RP-53306-FR, 94 p., 10 fig., 28 cartes, 31 pl. ht.- rapport du bureau d'études ALTO ingénierie (volume séparé)

Guide technique- Pompe à chaleur géothermique sur aquifère. Conception et mise en œuvre – collection scientifique et technique (ADEME, ARENE, BRGM).

Évaluation du potentiel géothermique du Canton de Neuchâtel. Projet CREGE novembre 2008.

Nappe du Cénomanien de Touraine : actualisation de la base de données en Indre-et-Loire. Décembre 1994. Rapport BRGM R38267.

Annexe 1

Traitements statistiques élémentaires de données en vue d'une cartographie géostatistique

La possibilité de réaliser une cartographie géostatistique d'un paramètre est contrôlée par trois aspects :

- la densité des données disponibles ;
- leur distribution spatiale;
- les relations entre les mesures disponibles en différents points (autocorrélation, représentée par le variogramme) ou avec d'autres données (corrélation, dérive externe).

Pour déterminer a priori si un lot de données est favorable ou pas, vis-à-vis de ces trois critères, à une approche géostatistique, il est recommandé, après une première étape de contrôles de base (ajustement des formats de données au logiciel de traitement statistique utilisé, élimination des doublons, analyse des erreurs de mesure et des biais d'échantillonnage), de réaliser l'analyse structurale des données.

L'objet de l'annexe 1 est de présenter les différentes techniques de traitement initial des données et quatre méthodes de cartographie régionale.

Le logiciel de traitement statistique utilisé dans la présente étude est ISATIS (GÉOVARIANCES); cependant, les rappels énoncés ici sont applicables à toute approche statistique d'un problème donné.

1. Outils d'analyse structurale d'un échantillon de données

Différents outils d'analyse des données peuvent être mis en œuvre. Ils permettent de juger a priori de l'opportunité ou pas de poursuivre la démarche de cartographie géostatistique pour une variable donnée, dans un contexte naturel donné.

Les techniques à mettre en œuvre sont les suivantes :

- Analyse de l'échantillon : taille et répartition ;
- Analyse de l'histogramme des données ;
- Analyse variographique.

1.1. Analyse de l'échantillon

L'analyse de la taille de l'échantillon de données et de la répartition géographique des données donne des indications sur la possibilité ou pas d'envisager une cartographie géostatistique : un trop faible nombre de données ou une répartition très inégale par rapport à l'extension d'un aquifère ne permettra pas d'envisager ce type de cartographie.

1.2. Analyse de la dispersion des données via l'histogramme

L'histogramme permet :

- d'identifier la présence éventuelle d'extremums correspondant à des valeurs erronées : valeurs sous le seuil de détection, valeurs anormalement fortes...;
- de vérifier s'il y a une ou plusieurs populations (nombre de modes de la distribution).
 La présence de plusieurs modes peut signaler la présence de plusieurs populations, auquel cas une subdivision en zones géographiques distinctes peut s'avérer pertinente;
- de repérer des distributions lognormale ou gaussienne. Une comparaison de Henry permettra de conclure en cela, et de suggérer des transformations du paramètre afin de rendre l'étude plus pertinente (par exemple, à une distribution lognormale sera appliquée la fonction logarithme).

L'identification d'une distribution se fait au moyen du graphique de Henry : pour tester si les valeurs mesurées suivent une distribution gaussienne, on trace la courbe portant en abscisse les quantiles de la distribution mesuréeet en ordonnée les quantiles correspondants d'une distribution normale de moyenne 0 et d'écart-type 1. Si la courbe obtenue est une droite, on peut conclure que les mesures suivent une distribution gaussienne, dont la moyenne et l'écart-type peuvent être déterminées à partir de l'équation de la droite. Pour s'assurer que la courbe de Henry est effectivement une droite, on peut également soumettre les points de la courbe à un test du χ^2 , généralement réalisé par le logiciel utilisé.

Pour tester si les valeurs mesurées suivent une distribution lognormale, on procède de même en travaillant non pas sur les mesures, mais sur leurs logarithmes. Si ceux-ci suivent une loi gaussienne, alors les mesures suivent une loi lognormale.

Dans ce dernier cas, on pourra utilement travailler sur les logarithmes des valeurs.

1.3. Analyse de la structure spatiale des données via le variogramme

1.3.1. Définition du variogramme

Pour une variable donnée, le variogramme expérimental permet de quantifier les corrélations entre valeurs mesurées en des points différents, en fonction de la distance séparant ces points.

Pour une distance h donnée, le principe de calcul consiste à identifier tous les couples de points (x, x+h) séparés de cette distance, et à calculer la demie moyenne des carrés des différences de valeur du paramètre étudié entre les 2 points de chaque couple.

Ceci donne l'expression suivante pour le variogramme $\gamma(h)$:

$$\gamma(h) = \frac{1}{2} \cdot \frac{\sum_{1}^{N(h)} \left[Z(x) - Z(x+h) \right]^{2}}{N(h)}$$

x et x+h sont deux points de mesure séparés d'une distance h

Z(x) et Z(x+h) sont les valeurs du paramètre étudié, mesurées en ces points N(h) est le nombre de couples de points séparés par la distance h

Figure 4 : Principe de calcul du variogramme.

En faisant varier la distance h, on peut mettre en évidence la façon dont les corrélations évoluent en fonction de l'interdistance entre points (fig. 2).

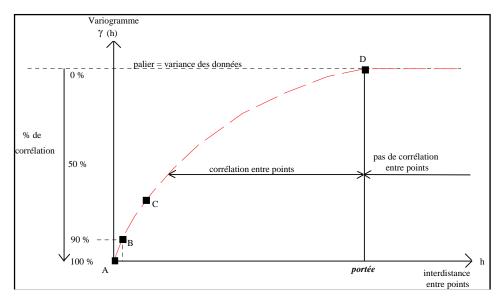


Figure 5 : Variogramme dans le cas d'un phénomène stationnaire (sans dérive) Point A.

Point A (cf. Figure 5)

Si h=0, il est évident que $\gamma(h)=0$ puisque les points x et x+0 sont confondus et ont donc la même valeur du paramètre étudié. La corrélation est maximale pour h=0. Si on classe le degré de corrélation sur une échelle variant de 0 à 100 %, elle est ici de 100 %.

Point B

En augmentant légèrement h, les points x et x+h sont voisins et ont en général de bonnes chances d'être bien corrélés. Les valeurs Z(x) et Z(x+h) sont proches et $\gamma(h)$ sera faible, mais différent de 0. Dans l'exemple de la figure ci-dessus, on suppose que les points sont encore corrélés à 90 % du maximum.

Point C

Si l'interdistance h augmente encore, les points x et x+h peuvent être encore corrélés, mais moins bien. Les différences de valeur entre Z(x) et Z(x+h) seront donc statistiquement plus fortes

Point D

Au bout d'une certaine interdistance entre points, sauf en cas de dérive, les différences Z(x)-Z(x+h) peuvent prendre toutes les valeurs possibles et la moyenne de leurs carrés se stabilise à un palier.

La distance à partir de laquelle le variogramme se stabilise est appelée "portée" du phénomène :

- 2 couples de points séparés d'une distance supérieure à la portée ne sont pas corrélés. On ne peut prévoir la valeur en un point à partir de l'autre. Les points sont corrélés à 0%;
- inversement 2 couples de points séparés d'une distance inférieure à la portée sont corrélés. Leurs valeurs sont liées entre elles.

La portée est une caractéristique physique du phénomène étudié. Par exemple cela peut être le diamètre moyen de lentilles minéralisées dans un gisement minier.

Le variogramme traduit la dégradation des corrélations entre points en fonction de l'interdistance séparant ces points.

1.3.2. Variabilité expliquée par le variogramme

Supposons que l'on reconnaisse 2 phénomènes par un échantillonnage à la maille de 100 m. Les variogrammes expérimentaux calculés pour ces 2 phénomènes sont visualisés ci-dessous.

Dans le cas A, le premier point du variogramme expérimental ne traduit que 20 % des corrélations totales. 80 % de la variabilité totale (=des corrélations) du phénomène se produit dans l'intervalle 0-100 m. La courbe expérimentale obtenue ne permet donc d'appréhender qu'une petite partie de la "variabilité" du phénomène. Il est impossible dans ce cas de prédire les teneurs locales (à une échelle de distance de 100 m ou moins) avec une bonne précision. La maille d'échantillonnage est trop grande par rapport à la portée du phénomène

Dans le cas B, le premier point du variogramme expérimental traduit 80% des corrélations totales. La maille d'échantillonnage permet de reconnaître une grande part de la variabilité totale du phénomène. Des estimations locales sont possibles avec une bonne précision.

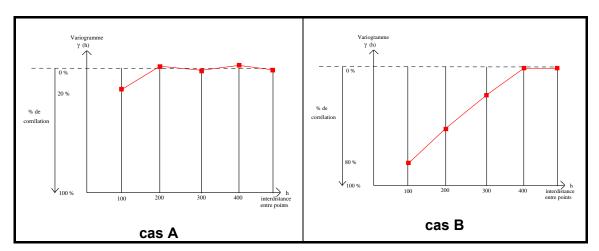


Figure 6 : Variogrammes expérimentaux de 2 phénomènes ayant des portées différentes et reconnus par une maille de 100 m.

Effet de pépite

Théoriquement on doit avoir $\gamma(0) = 0$. Toutefois, en pratique le variogramme expérimental ne semble pas toujours descendre jusqu'à l'origine. On dit alors qu'il y a un effet de pépite.

Trois explications sont possibles:

- les données sont entachées d'une erreur de mesure ou d'analyse. Dans ce cas deux mesures effectuées au même point donneront une valeur différente du paramètre étudié;
- 2. il existe une structure de courte portée, inférieure à la maille de reconnaissance, qui ne peut être mise en évidence par l'échantillonnage effectué. Ce cas peut se combiner avec le précédent et donner un effet de pépite partiel. Dans le cas de la figure 4.a, le premier point du variogramme expérimental disponible est le point A. Quel est le comportement possible entre ce point et l'origine ? Trois types de comportement sont envisageables: a) le variogramme descend jusqu'à 0 : il n'y a pas d'effet de pépite; b) le variogramme a un comportement linéaire à l'origine, avec un effet de pépite partiel d'amplitude C0; c) le variogramme a un comportement parabolique à l'origine, avec un effet de pépite partiel un peu plus fort que dans le cas précédent;
- 3. les données ne sont pas du tout corrélées dans l'espace, et on a en fait un effet de pépite pur, comme dans le cas de la fig. 4.b, où l'on s'intéresse à la pluviométrie annuelle (dans ce cas la distance est le temps).

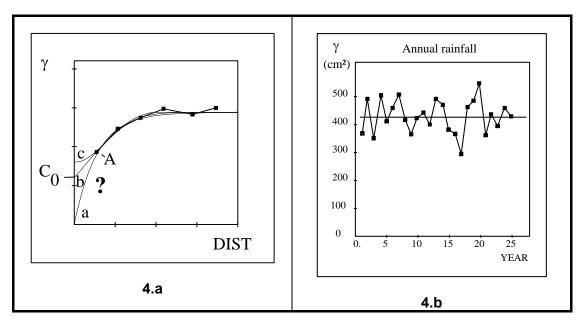


Figure 7 : Effet de pépite.

Structures gigognes

Dans certains cas le variogramme expérimental peut être la résultante de deux phénomènes ayant des portées différentes. Dans ce cas on observe une rupture de pente sur le variogramme expérimental. Chaque phénomène (structure) est caractérisé(e) par un variogramme avec un palier, la somme des paliers étant égale au palier total (fig. 5).

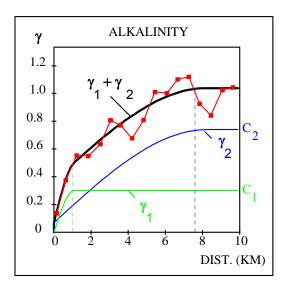


Figure 8 : Variogramme à plusieurs structures.

1.3.3. Variogramme et cartographie régionale d'une donnée

Le variogramme permet de caractériser les corrélations spatiales des valeurs mesurées :

- anisotropie dans la distribution des valeurs (comportement différent selon la direction);
- caractère très structuré ("continu") de la répartition des valeurs, ou au contraire peu ou pas structuré (répartition apparemment « aléatoire » des valeurs) ;
- présence de distances caractéristiques (appelées « portées ») en-dessous desquelles deux valeurs sont corrélées entre elles et au-delà desquelles elles sont indépendantes. Ces « portées » peuvent être liées aux caractéristiques géologiques ou hydrogéologiques de l'aquifère (exemple : diamètre de « lentilles » plus perméables dans un environnement sablo-argileux);
- présence ou non d'une dérive : tendance à l'augmentation plus ou moins systématique des valeurs dans une direction donnée. Au contraire, la stationnarité (relative « stabilité ») des valeurs montre l'absence de tendance à une augmentation systématique (ce qui n'interdit pas des fluctuations locales) ;
- dans le cas d'une estimation variographique, le variogramme devra être modélisé : un modèle de variogramme est une fonction, respectant certaines conditions mathématiques, qui approche le variogramme mesuré.

Une fois caractérisées et quantifiées les corrélations spatiales à l'aide du variogramme, il est possible de mettre en œuvre différents types de traitement :

- recherche de points aberrants ;
- interpolation d'une variable (exemple débit spécifique) sur une maille, avec estimation de l'incertitude :
- calcul de cartes de probabilité de dépasser un seuil.

Toutefois ces traitements ne sont possibles et efficaces que si le variogramme montre une structure convenable et une densité (le nombre et la répartition) des données suffisante, pour pouvoir envisager une cartographie locale avec une incertitude acceptable. Ces deux aspects sont liés, car la qualité de la densité se juge selon les portées dégagées du variogramme : il faut que les distances entre les points de mesure soient du même ordre ou plus petites qu'une portée pour que celle-ci soit effectivement utilisée dans les calculs.

Si le variogramme montre que le phénomène, à l'échelle où nous l'observons, est trop peu structuré (le paramètre étudié fluctue de façon discontinue), ou si les données sont en trop faible nombre, une cartographie locale n'a aucun sens, que cette cartographie soit élaborée par une technique géostatistique ou par une autre technique.

Ces premiers traitement et analyses sont donc fondamentaux pour permettent de juger a priori de l'opportunité ou pas de poursuivre la démarche de cartographie géostatistique pour une variable donnée, dans un contexte naturel donné.

Pour que la régionalisation d'un paramètre à partir d'une approche géostatistique soit valable, il est préférable :

- que la distribution des valeurs suive une loi normale (information tirée de l'histogramme et de la droite de Henry). Si ce n'est pas le cas, les valeurs extrêmes seront plus difficiles à restituer. La plupart des méthodes d'interpolation et de cartographie donnent des résultats optimaux lorsque les distributions sont gaussiennes et sont plus ou moins mises en difficulté lorsqu'on a affaire à des distributions très dissymétriques ;
- que les valeurs du paramètre soient les plus continues possible en termes de continuité spatiale. La continuité spatiale se lit sur le variogramme, au niveau du début de la courbe : un paramètre est spatialement continu si son variogramme démarre à zéro (en pratique, à une valeur faible). En outre, le paramètre est d'autant plus continu que son variogramme reste proche de zéro lorsque la distance augmente;
- que la distance entre les points de mesure soit inférieure à la (plus petite) portée du variogramme. Dans le cas contraire, la cartographie géostatistique ne fournira pas de meilleures estimations, en général, qu'une moyenne glissante des points sur le voisinage considéré.

2. Méthodes de cartographie régionale

Quatre méthodes de cartographies ont été appliquées et comparées sur trois aquifères de la Région Centre: une méthode globale et trois méthodes locales, dont deux basées sur l'application de méthodes géostatistiques. Rappelons qu'une méthode globale ne fournit qu'une seule valeur du paramètre sur tout le domaine étudié, tandis qu'une méthode locale fournit une valeur du paramètre en tout point (en général la valeur est fournie sur une grille régulière).

- Méthode globale : cette méthode consiste à tirer de la répartition des valeurs certaines statistiques (généralement la moyenne, la médiane et les autres quartiles), et de les exploiter pour en tirer soit une valeur unique, soit une fourchette unique attribuée à un aquifère ou à une portion d'aquifère. La méthode des boîtes à moustaches de Tukey (appliquée en Région Centre lors de la réalisation de l'atlas du potentiel géothermique) est une méthode globale.
- Méthodes locales : ces méthodes exploitent la position des données, soit par leurs distances relatives (méthodes glissantes ou du premier ordre), soit par leurs positions relatives (méthodes variographiques ou du second ordre). Ces méthodes utilisent la notion de voisinage, c'est-à-dire pour chaque point d'estimation un ensemble de points de mesures "proches". On peut définir ces voisinages par leur nombre de points (les n points les plus proches du point concerné), ou par leur rayon (les points distants de moins de r). On distingue la méthode statistique mobile et la méthode géostatistique :
 - · Statistique mobile (ou glissante)⁹: elle consiste à calculer, en tout point, une statistique (généralement médiane, moyenne arithmétique ou géométrique, quartiles) sur les valeurs connues dans le voisinage. On pourra faire varier la statistique considérée et la définition du voisinage (nombre de points et rayon).
 - Méthodes géostatistiques : il s'agit de méthodes qui s'appuient sur le variogramme des données observées pour reconstituer la répartition de ces données à l'échelle régionale. Deux méthodes sont présentées ici :
 - o Le krigeage : il consiste à proposer un modèle de variogramme (une fonction qui respecte des conditions mathématiques définies et qui approxime le variogramme mesuré) et d'utiliser celui-ci pour une estimation. Alors que les méthodes glissantes considèrent de la même manière toutes les mesures d'un même voisinage, le krigeage les différencie selon leur distance au point d'estimation et leurs distances relatives. On pourra faire varier la définition du voisinage (nombre de points et rayon) et le modèle de variogramme (en particulier les pépites, portées et paliers estimées).
 - Les simulations variographiques : diverses méthodes permettent de simuler une répartition qui respecte à la fois les valeurs connues par mesure et le modèle de variogramme proposé par l'opérateur. Leur intérêt est de fournir, en tout

_

⁹Dans la suite, on parlera de statistique (moyenne ou médiane) **directe**, pour la statistique appliquée sur un paramètre, et **logarithmique**, (ou "sur les logarithmes") pour l'exponentielle de la statistique appliquée sur les logarithmes de ce paramètre. Ainsi, la "moyenne logarithmique" est la moyenne géométrique.

point, non plus une estimation unique, mais une répartition de valeurs plus ou moins possibles; de cela on pourra tirer des approximations de la probabilité de dépasser un seuil donné, ou de se trouver entre deux bornes données, ou la valeur atteinte avec une probabilité (par exemple, la valeur dépassée dans 80 % des estimations). Cependant il ne faut pas surinterpréter les résultats; une région avec peu de mesures sera toujours mal estimée par les simulations.

On peut appliquer ces méthodes à une grille de maille donnée pour produireune représentation cartographique.

2.1. Méthode globale

La méthode statistique envisagée ici est celle des boîtes à moustaches de Tukey qui consiste à représenter graphiquement des statistiques représentatives d'un lot de valeurs.

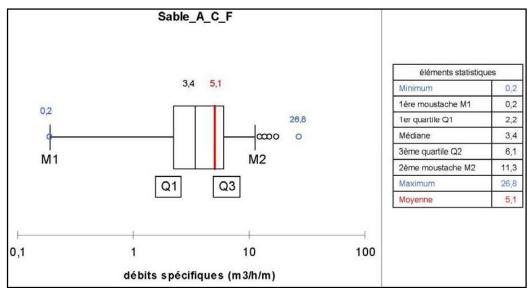


Figure 9 : Exemple de boîte à moustaches.

La boîte à moustache est constituée :

- d'un rectangle central dont les bornes correspondent aux valeurs des premier et dernier quartiles. Le rectangle central couvre donc l'étendu correspondant à l'interquartile q75-q25, qui contient la moitié des mesures;
- un trait vertical marquant, à l'intérieur du rectangle précédent, la médiane et séparant le rectangle en deux parts regroupant chacune un quart des observations;
- une croix qui représente la moyenne des valeurs ;
- des moustaches qui partent respectivement à gauche et à droite jusqu'aux valeurs q25-1,5×(q75-q25) et q75+1,5×(q75-q25). Elles couvrent donc une fois et demi l'interquartile;

- des points dépassant de part et d'autre des moustaches, appelés points aberrants, sont notés en points verts.

Ce mode de représentation permet d'associer des valeurs statistiques globales à l'entité d'où proviennent les données et de faire une cartographie sur cette base.

Plusieurs types de statistiques globales peuvent être retenus :

- Valeur moyenne : dans ce cas, la moyenne des valeurs d'une variable est associée à l'intégralité de l'entité décrite ;
- Intervalle de confiance : les bornes du rectangle de la boîte à moustache (q25 et q75) peuvent être retenues pour définir un intervalle de valeurs d'une variable à attribuer à l'intégralité de l'entité décrite. Dans ce cas, chaque portion de l'entité a une chance sur deux d'appartenir à cet intervalle;
- Seuil de confiance (ou plafond) : une valeur statistique associée à une probabilité d'occurrence peut être attribuée à l'intégralité de l'entité décrite. Par exemple : q25 est dépassé trois fois sur quatre, ouq50 est dépassé une fois sur deux, ou q75 est dépassé une fois sur quatre.

La limite de cette méthode globale réside dans le fait qu'elle remplace par une valeur unique un jeu de données caractérisé par une variabilité spatiale. Elle néglige la répartition géographique des valeurs. Par exemple, elle ne permet pas de conclure sur les valeurs aberrantes: sont-elles disséminées (ce qui suggérerait une erreur d'échantillonnage ou de mesure, ou un facteur externe localisé), ou concentrées en amas (ce qui laisserait entendre un comportement différent localisé)? Elle masque en outre les régionalisations, et il faut donc pour une meilleure prise en compte diviser la zone d'étude (la masse d'eau) en différentes zones, avec l'arbitraire ajouté par le choix des frontières, l'effet de saut de valeur à la frontière entre deux zones, et le risque d'avoir trop peu de points pour conclure sur une zone donnée.

2.2. Statistique mobile

Une méthode mobile (ou glissante) consiste à calculer, pour chaque portion de l'entité décrite (ou point de sortie), une statistique à partir les points de mesure « proches ». Il existe trois manières usuelles de définir le voisinage du calcul (c'est-à-dire les points de mesure à prendre en compte) :

- voisinage défini par un nombre *n* de points de mesures ;
- voisinage défini par l'ensemble des points de mesures situés au maximum à la distance d du point de sortie ;
- voisinage défini par un nombre maximum n de points de mesures situés à moins de la distance d.

Idéalement, la taille du voisinage (distance *d*) devrait être choisie en fonction des caractéristiques du variogramme. En effet, lorsque le variogramme présente une portée, les points situés au-delà de cette distance ne sont pas corrélés. Si l'on cherche à estimer une valeur en un point P, il n'est donc pas utile (et, de fait, pas souhaitable)

d'utiliser des données situées au-delà d'une distance (par rapport au point P) égale à la portée.

La difficulté de la méthode réside dans le choix de *d*, en particulier dans le cas d'aquifères dont certaines portions ne sont pas renseignées en données.

Enfin, il peut être intéressant de déclarer un nombre minimal *n* de points nécessaires dans le voisinage, en-deçà duquel la valeur n'est pas calculée au point de sortie. Cela permet de ne pas induire en erreur en présentant des estimations sur un nombre de mesures trop faible.

D'autre part, il est nécessaire de choisir le calcul statistique à appliquer aux points du voisinage :

- la moyenne fait sens si la densité de mesures est importante, mais est influençable par la présence de valeurs aberrantes. Elle est également sujette aux effets de voisinage (un point de mesure qui est pris en compte ou non pour un tout petit déplacement du point de sortie);
- la médiane est moins sensible aux valeurs aberrantes. Elle est par contre également sujette aux effets de voisinage ;
- un quantile peut indiquer un seuil de confiance. Deux quantiles peuvent donner un intervalle de confiance.

2.3. Krigeage

La deuxième méthode de cartographie locale testée est celle du krigeage. Le krigeage est une méthode d'interpolation géostatistique qui nécessite le calcul et la modélisation du variogramme des données.

2.3.1. Modèle de variogramme - définition

Le variogramme estimé sur des points mesurés (ou variogramme expérimental) ne présente pas toujours l'aspect idéal du variogramme théorique d'une fonction aléatoire stationnaire (la courbe expérimentale est sujette à des fluctuations statistiques, elle croît et décroît sans tendre vers la variance). La *modélisation* du variogramme consiste à choisir une fonction parmi certaines (dites fonctions *autorisées*) qui représente au mieux le variogramme expérimental. Le **modèle de variogramme** obtenu est un compromis entre le variogramme estimé, les contraintes mathématiques liées à ce type d'objet et d'éventuelles connaissances ou hypothèses sur la nature et les caractéristiques du phénomène étudié.

La Figure 10 montre un exemple de variogramme estimé et modélisé sous Isatis. L'abscisse décrit l'interdistance entre les paires de points ; l'ordonnée s'exprime en carré de l'unité de la fonction Z étudiée. Le trait rouge fin montre le variogramme estimé, les ronds verts montrant les valeurs calculées du variogramme. Les nombres de paires de points disponibles pour le calcul sont reportés dans l'histogramme. Le trait rouge épais décrit le modèle de variogramme proposé par l'opérateur. Enfin, la droite

pointillée donne la variance du jeu de données : en théorie, variance et palier sont égaux.

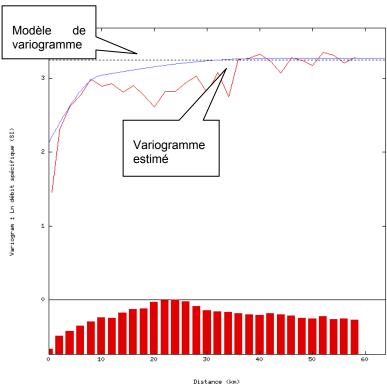


Figure 10 : Exemple de variogramme estimé (rouge) et de modèle de variogramme (bleu). Les pointillés indiquent la variance estimée, l'histogramme donne le nombre de paires de points.

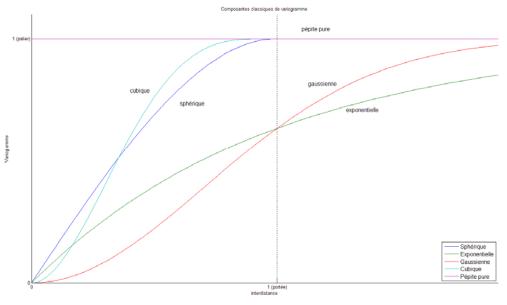


Figure 11: Courbes des composantes les plus classiques (fonctions autorisées) d'un variogramme.

2.3.2. Krigeage

Le **krigeage** est une technique d'estimation qui utilise les données et leur modèle de variogramme. Cette technique est :

- Exacte : l'estimation par krigeage en un point mesuré est la mesure en ce point. Ceci ne vaut qu'en l'absence de pépite dans le modèle de variogramme.
- Non-biaisée: l'évaluation est juste en moyenne (sous certaines hypothèses).
- Optimale dans le sens où elle minimise de façon statistique l'amplitude des erreurs (plus exactement elle minimise la variance de l'erreur d'estimation)
- Lissant : la variance des valeurs estimées est plus faible que la variance des données brutes ; le krigeage a tendance à lisser les valeurs (réduire l'aspect rugueux d'un profil, par exemple).

L'utilisation pratique du krigeage nécessite souvent, pour chaque estimation, de ne pas prendre en compte toutes les données, mais uniquement les plus proches. Ce **voisinage de krigeage** est souvent pris comme l'ensemble des points dans un disque de rayon donné autour du point d'estimation, ou comme l'ensemble d'un nombre donné de points à proximité du point d'estimation. On parle de voisinage **global** quand on prend en compte tous les points de données. En pratique, un voisinage de quelques dizaines de points est suffisant. Le krigeage est souvent présenté comme l'une des meilleures méthodes d'estimation disponibles. En effet, il prend en compte, via le variogramme, la structure du phénomène supposé.

Néanmoins ses conditions d'application (stationnarité, variogramme ne pouvant être modélisé de manière satisfaisante) et la densité de données par rapport à la portée du phénomène étudié, font qu'il n'est pas forcément plus satisfaisant qu'une estimation simple (moyenne mobile, moindres carrés...). Certains phénomènes représentés par des variogrammes peu structurés ne sont pas bien modélisables par cette méthode.

2.4. Simulations géostatistiques

La méthode par simulation consiste, non pas à chercher en chaque point une estimation unique, mais en chaque point une distribution (un histogramme) d'estimations. On simule donc un grand nombre de cas possibles, qui s'inscrivent dans un cadre défini par les valeurs mesurées, mais également, par des hypothèses sur le paramètre (sa stationnarité, sa continuité...).

La méthode testée ici consiste à :

- produire un grand nombre de simulations permettant chacune de reproduire les données d'entrée et le modèle de variogramme, donc les relations entre les valeurs en plusieurs points du paramètre étudié ;
- puis estimer une statistique (par exemple la moyenne) à partir des résultats des simulations en chaque maille de l'entité étudiée.

Il est fortement conseillé d'appliquer cette méthode à une variable gaussienne ; dans notre cas, lnQ_s .

Les simulations (200) sont construites par la méthode des bandes tournantes¹⁰. Elles respectent les valeurs de lnQ_s sur les points d'apprentissage, ainsi que le modèle de variogramme défini précédemment. Elles permettent de disposer de 200 valeurs possibles de lnQ_s pour chaque maille de calcul.

2.5. Contrôle des résultats

Pour juger de la pertinence de résultats de cartographie obtenus à partir de l'application d'une des méthodes précédentes, il est proposé de mettre en œuvre des méthodes de contrôle des résultats (0–2.6 «Comparaison des méthodes») basées sur la comparaison des valeurs mesurées aux valeurs estimées, dans les mailles où les valeurs mesurées sont disponibles.

Deux approches sont proposées ici :

- comparaison de lots de données « d'apprentissage » et « de contrôle » ;
- validation croisée.

Notons que, dans tous les cas, une simple évaluation des résultats obtenus, à « l'œil nu » par un hydrogéologue local, est recommandée. Il s'agit d'un contrôle peuprécis, mais qui permet de valider les spatialisations obtenues et d'identifier d'éventuels artefactsou effets de voisinages.

2.5.1. Contrôle par comparaison apprentissage/contrôle

Cette méthode est utilisée dans le cas où les données sont nombreuses. Son principe est le suivant : la cartographie est réalisée à partir d'un lot de données appelé « lot d'apprentissage » correspondant à ¾ des données disponibles. Le ⅓ des données restantes ou « lot de contrôle » est utilisé pour vérifier le degré de cohérence entre données estimées par la cartographie et les données mesurées. On obtient donc, sur ⅓ des points initiaux, à la fois la valeur mesurée et la valeur estimée.

Ce contrôle offre l'avantage d'être facilement programmable et rapide à exécuter, c'est pourquoi il sera effectué préférentiellement, dès que le nombre de points sera suffisant (à partir de quelques centaines). Il possède néanmoins les inconvénients suivants : il fournit ¼ d'estimations seulement, et il ne garantit pas que les lots d'apprentissage et de contrôle soient choisis de telle manière qu'ils soient équivalents (il peut donc y avoir des biais d'évaluation).

_

La méthode des bandes tournantes est l'une des plus classiques et des plus rapides. Elle consiste à tirer une grande quantité de lignes selon lesquelles sont réalisées les simulations, puis de compléter celles-ci par combinaisons linéaires.

2.5.2. Contrôle par validation croisée

Ce mode de validation est retenu dans les cas où le nombre dedonnées est insuffisant pour faire un lot d'apprentissage et un lot de contrôle¹¹.

L'expression recouvre deux applications de la même méthode : d'une part l'obtention de paires valeur mesurée – valeur estimée, et d'autre part un calcul sous Isatis et sa représentation graphique.

Cette méthode consiste, pour tout point (i) où une mesure (Pi) est connue, à estimer une valeur (Pi*) de la variable à partir de tous les autres mesures connues, en faisant donc comme si la mesure (Pi) était inconnue. On dispose alors en chaque point (i) de la vraie valeur (Pi) ainsi que de son estimation (Pi*).

En outre, Isatis offre un module de représentations graphiques associées à la validation croisée ; cependant, il n'est disponible que dans le cas d'un krigeage direct (non lognormal). Il est dans ce cas possible de calculer l'erreur d'estimation (Pi* - Pi) et l'erreur relative ((Pi* - Pi)/Pi) et, en comparant différentes méthodes d'estimation de (Pi*), d'évaluer les performances de chacune.

_

-

-

-

¹¹ Cette méthode, non présente nativement sous Isatis, est lourde à mettre en place et lente à calculer, et n'est donc employée qu'à défaut de la précédente, lorsque le nombre de mesures disponibles est trop faible (aux alentours de la centaine de points).

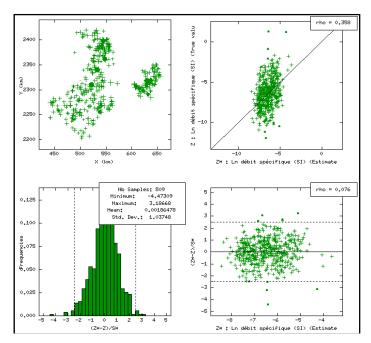


Figure 12 : Exemple de validation croisée sous Isatis.

Le résultat est présenté sous Isatis sous la forme de quatre graphiques (cf. Figure 12) :

- en haut à gauche, la carte géographique (répartition des points);
- en haut à droite, le nuage de corrélation : estimations en fonction des mesures ;
- en bas à gauche, l'histogramme des erreurs relatives réduites (rapport de l'erreur d'estimation sur son écart-type) ;
- en bas à droite, le nuage des erreurs relatives réduites par rapport aux estimations.
 Les points indiqués par des ronds sont les points anormaux, définis ici comme ceux dont l'erreur relative réduite dépasse un seuil en valeur absolue. Ce seuil est défini par l'utilisateur, et vaut par défaut 2,5.

2.6. Comparaison des méthodes

Les méthodes précédentes (0–2.5 «Contrôle des résultats») fournissent un ensemble de n points de mesure i où sont disponibles des valeurs mesurées z_i et des valeurs estimées \hat{z}_i . Les écarts entre les valeurs estimées et mesurées sont ensuite analysés pour rendre compte de la validité de la méthode de cartographie et pour comparer différentes méthodes entre elles.

Différents types d'analyses des écarts peuvent être réalisés.

2.6.1. Agrégation

Cette analyse consiste à agréger ces écarts en une valeur unique, et la méthode sera considérée d'autant plus fiable que l'agrégation est proche de zéro.

On pourra utiliser les agrégations suivantes :

- Biais : $b = \Sigma_i (\hat{z}_i z_i) \div n$.
- Erreurs absolues : $\varepsilon = \sum_{i} |\hat{z}_{i} z_{i}| \div n$.
- Erreurs quadratiques : $\eta = \sqrt{(\Sigma_i (\hat{z}_i z_i)^2 \div n)}$.

Le biais mesure l'erreur systématique de la méthode ; un biais positif (resp. négatif) souligne que la méthode a tendance à donner des estimations surestimées (resp. sous-estimées). L'erreur absolue montre la dispersion entre la mesure et l'estimation ; plus elle est faible, plus les estimations sont proches des mesures respectives. L'erreur quadratique donne un résultat similaire, en insistant sur les cas où mesure et estimations divergent fortement.

2.6.2. Marge d'erreur

Cette analyse consiste à calculer quelle proportion des estimations tombent « assez près » des mesures. On se fixe en premier lieu une marge d'erreur ("proximité") *a priori*, et on calcule la proportion d'estimations dans cette marge.

- Taux (pour une proximité p) : $j(p) = \sum_i (|\hat{z}_i - z_i| \le p) \div n$.

De cette manière, j(p) est égal à la proportion de points où l'estimation est égale à l'erreur à $\pm p$ près. Plus la marge d'erreur p sera grande, plus grande sera la proportion j(p) de points bien estimés.

2.6.3. Prédictibilité

En plus de ces statistiques habituelles, on propose d'étudier la prédictibilité de la méthode.

L'idée sous-jacente est que l'exploitant ne cherche pas forcément un fort débit, mais plus fondamentalement un débit suffisant pour l'utilisation qu'il en prévoit. En effet, l'information « le débit attendu moyen est de ... » lui sera moins utile que l'information « le débit attendu a une chance ... d'être suffisant pour tel usage ». Par rapport à un seuil d'usage, l'excédent éventuel de débit ne sera pas exploité, alors que le défaut éventuel de débit soit devra être compensé par d'autres apports énergétiques, soit conduira à l'échec du forage géothermique.

Fixons-nous un seuil de débit spécifique recherché, ou de débit exploitable. Dans la suite, afin que les grandeurs soit plus "parlantes", nous présenterons les seuils comme des valeurs de débits, tout en gardant à l'esprit qu'ils sont basés sur un rabattement arbitrairement choisi à 5 m.

Nous classons alors chaque point dans un tableau à double entrée, selon que son estimation est au-dessus ou en-dessous du seuil, et de même selon que sa valeur mesurée est au-dessus ou en-dessous du seuil.

On obtient ainsi un double classement, pour chaque point de mesure, pour chaque méthode étudiée, pour chaque seuil considéré. On peut ensuite résumer le nombre de points dans chaque configuration dans un tableau à double entrée :

«faux positifs» points supposés favorables à tort	«vrais positifs» points bien reconnus comme favorables		
«vrais négatifs» points supposés non- favorables à raison	«faux négatifs» points rejetés à tort		

Ainsi, les «faux positifs» sont les points que la méthode annonce comme favorables à une exploitation géothermique (supérieurs au seuil considéré), alors qu'en réalité ils ne le sont pas. À l'inverse les «faux négatifs» sont les points qui seraient favorables, mais que la méthode ne permet pas d'identifier. Les «vrais négatifs» sont les points rejetés à raison et les «vrais positifs» les points où la méthode trouve un débit spécifique effectivement favorable.

Pour chaque seuil, quelques paramètres sont intéressants :

- α_Q : **taux de détection**. C'est la proportion de points favorables qui sont bien reconnus par la méthode ; α_Q = *vrais positifs* ÷ (*vrais positifs* + *faux négatifs*) ;
- β_Q : **taux de confiance**. C'est la proportion de points supposés favorables qui le sont effectivement ; β_Q = *vrais positifs* ÷(*vrais positifs* + *faux positifs*) ;
- sc_Q: score. C'est les bénéfices ou coûts de chaque cas de figures. Nous poserons ici sc_Q = -1×faux positifs + 3×vrais positifs + 0×vrais négatifs + 0×faux négatifs. On suppose donc que le bénéfice de ne pas forer est nul, celui de forer et de trouver un débit suffisant de 3, et celui de forer inutilement de -1.

En d'autres termes, le taux de détection est la part des zones favorables qui est bien identifiée par la méthode, le taux de confiance est la part des zones supposées favorables par la méthode qui le sont effectivement. Ces deux valeurs permettent de s'assurer que la méthode détecte toutes les zones favorables et seulement elles. Le score est une tentative de chiffrer les bénéfices et coûts d'une exploitation à grande échelle basée sur les prédictions de la méthode. Il estime les investissements et les retombées attendus, de manière globale. Les coefficients utilisés pour le score sont choisis de façon arbitraire ; ils pourront être précisés dans le cas d'études futures.

Annexe 2

Évaluation des débits spécifiques à partir de données de transmissivité

Il est proposé ici de tester la possibilité d'établir des relations de corrélation entre les paramètres transmissivité (T) et débit spécifique (Q_s) et d'en évaluer la représentativité.

L'objectif est de vérifier la possibilité d'utiliser les données de transmissivité (représentatives des propriétés hydrauliques intrinsèques de l'aquifère) pour reconstituer des données de débits spécifique représentatives de l'aquifère et non altérées par les propriétés des ouvrages où sont réalisées les mesures de Q_s , puis pour établir des cartes de productivité (débits potentiellement exploitables) des aquifères.

1. Relations proposées

Deux modèles de corrélation possibles entre transmissivité et débit spécifique sont proposés :

- Relation directe : $T = \alpha Q_s + \beta + \varepsilon$
- Relation logarithmique : $T = \beta' Q_s^{\alpha'} e^{\varepsilon}$

Où α , β , α' , β' sont des constantes, et ε l'erreur, de moyenne nulle, qui dépend du point de mesure. La première équation propose une relation linéaire entre Q_s et T; la seconde, une relation linéaire entre lnQ_s et lnT.

2. Exemple de la craie du Séno-turonien

La craie du Séno-turonien (CST) est renseignée de 196 points de débit spécifique et transmissivité connus. C'est la zone la mieux connue en transmissivité pour la Région Centre.

2.1. Relation directe

La fonction de corrélation obtenue pour cet aquifère entre les paramètres T et Q_s est illustrée en Figure 13.

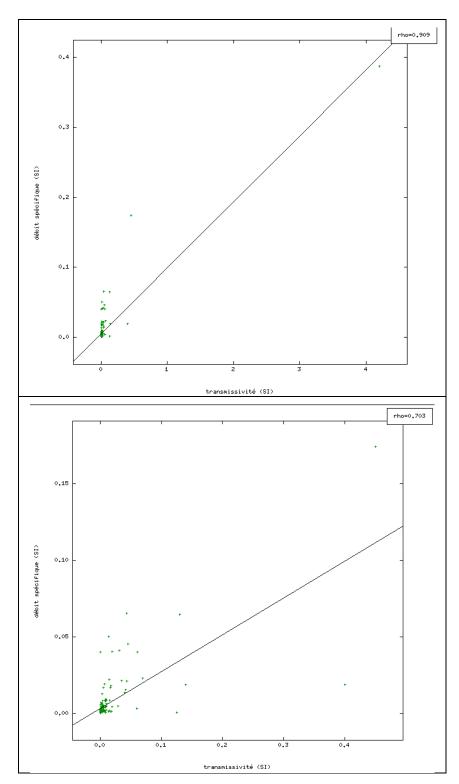


Figure 13 : En haut, droite de corrélation entre Q_s et T (ρ = 0,909) ; en bas, idem avec un point négligé (ρ = 0,703).

Le coefficient de corrélation linéaire (ρ) décrit l'alignement des valeurs de T et Q_s mesurées avec la droite de corrélation. Il y a corrélation linéaire parfaite si ρ = 1.

Les erreurs ε peuvent être estimées par la distance verticale entre chaque point et la droite. La qualité de la corrélation est estimée par lecoefficient de corrélation.

Les équations des droites de régression dans le cas de la craie du Séno-turonien sont :

- $Q_s = 0.094159 \times T 0.005299$ ($\rho = 0.909$) en prenant en compte tous les points
- $Q_s = 0.240061 \times T + 0.003328$ ($\rho = 0.703$) en écartant un point anormal

Ces corrélations sont marquées par :

- des erreurs de corrélation importantes : large dispersion verticale des données mesurées autour de la droite de régression ;
- des coefficients de corrélation sensibles à la qualité des données mesurées : le fait de négliger un point de l'échantillon de données implique une importante variation du coefficient de corrélation.

2.2. Relation logarithmique

La fonction de corrélation obtenue pour cet aquifère entre les paramètres $\ln T$ et $\ln Q_s$ est illustrée en Figure 14.

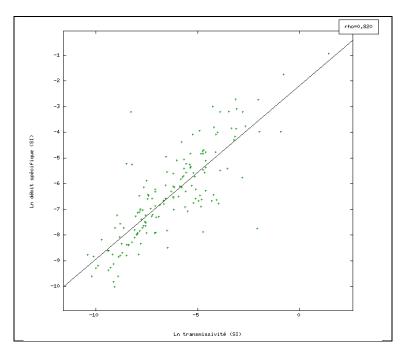


Figure 14 : Droite de corrélation entre $\ln Q_s$ et $\ln T$ (ρ = 0,820). Si l'on néglige le même point que précédemment, la droite change peu (ρ = 0,807).

Dans ce cas, l'équation de la droite de régression linéaire est :

$$ln(Q_s) = 0.675547 \times ln(T) - 2.174921 \ (\rho = 0.820)$$

La relation logarithmique semble plus robuste que la relation directe. Cependant, on peut remarquer 36 points alignés selon une droite qui n'est pas la droite de corrélation. Il peut s'agir d'un effet dû au passage en logarithme, ou cela peut être le signe que, pour certains essais, le débit spécifique a été calculé à partir de la transmissivité ou inversement.

La corrélation a donc été refaite en ne tenant pas compte de ces points (cf. Figure 15).

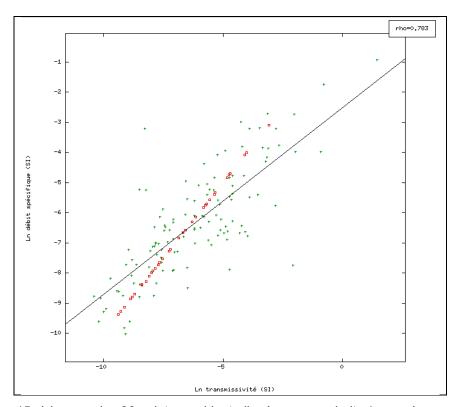


Figure 15 : Idem, exclus 36 points semblant alignés sur une droite (marqués en rouge), ρ = 0,783.

L'équation de la droite de corrélation devient :

$$ln(Q_s) = 0.619349 \times ln(T) - 2.513895 (\rho = 0.783)$$

Elle estassez notablement modifiée et se caractérise encore une fois par des erreurs de corrélation importantes : forte dispersion des données mesurées autour de la droite de corrélation.

2.3. Exemple de calcul des Q_s à partir des T

Des estimations de Q_s à partir de T disponibles sur l'aquifère de la craie du Sénoturonien ont été faites à partir de la régression logarithmique ci-dessus et comparées aux mesures de Q_s disponibles sur cet aquifère.

Le résultat de cette comparaison est illustré par la Figure 16.

Si l'estimation est globalement satisfaisante (ρ = 0,936), l'écart entre valeur estimée et mesurée peut dépasser un facteur dix, ce qui n'est pas acceptable pour des valeurs de débit spécifique.

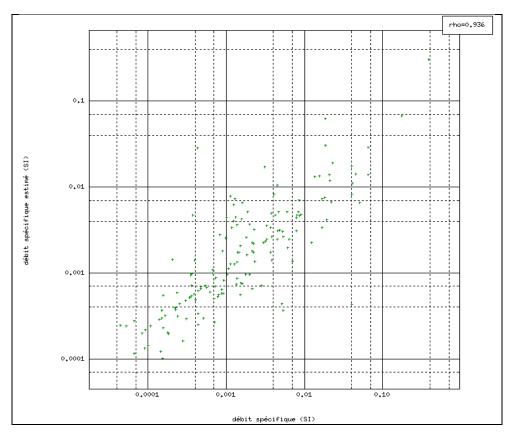


Figure 16 : Confrontation des Qs estimés et mesurés, en échelle logarithmique.

3. Exemple de la Beauce

L'aquifère des calcaires de Beauce (sous recouvrement et affleurant, hors karst) compte 99 points où à la fois le débit spécifique et la transmissivité sont connus.

La corrélation entre données de Q_s et de T établie pour cet aquifère est présentée en Figure 17 puis en Figure 18 pour les données logarithmiques.

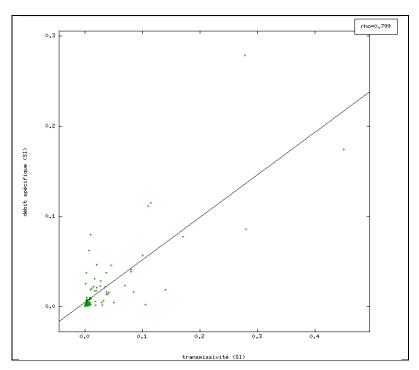


Figure 17 : Corrélation en Qs et T dans la Beauce.

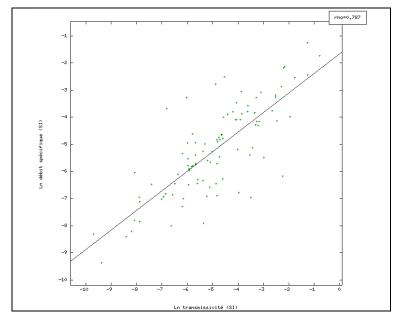


Figure 18 : Corrélation entre InQs et InT dans la Beauce.

Les équations des droites de corrélation sont respectivement :

$$Q_s = 0.471101 \times T + 0.004781 \ (\rho = 0.799)$$

$$lnQ_s = 0.721965 \times T - 1.656628 (\rho = 0.787)$$

Cet exemple met encore une fois en évidence la forte dispersion des données autour de la droite de corrélation.

4. Réflexion sur ces corrélations

Les tests effectués sur les données Q_s et T des aquifères de la nappe de Beauce et des Sables du Cénomanien en Région Centre ont montré :

- que des corrélations peuvent être établies entre les paramètres T et Q_s dans de bonnes conditions (coefficient de corrélation de l'ordre de 80 %);
- que la corrélation des logarithmes est plus stable (varie moins si l'on néglige quelques points) que celle des valeurs directes.

Cependant, ces tests montrent également :

- que les données de T sont très dispersées par rapport à celles de Q_s, ce qui implique des erreurs importantes sur les valeurs de Q_s qui seraient calculées à partir des valeurs de T;
- que le passage en logarithmes donne de meilleurs résultats, mais a tendance à lisser les écarts. Par exemple, un écart entre deux logarithmes de ±a revient à une proportion de e³ entre les valeurs initiales (multiplié ou divisé). Ainsi, deux logarithmes distants de 1 reviennent à deux valeurs initiales distantes d'un facteur 2,7 entre elles. Ce qui passe pour des approximations raisonnables peut donc cacher des amplitudes importantes, donc des erreurs d'estimation des Qs importantes.

Par conséquent, même si les mesures de transmissivités sont *a priori* plus représentatives de l'aquifère que les mesures de débits spécifiques, le fait d'en tirer des débits spécifiques calculés par une relation fixe (ne tenant pas compte d'autres paramètres, comme la géographie ou la qualité de la réalisation du forage et de la mesure) entraîne une erreur significative.

Pour conclure, la méthode d'évaluation des Q_s à partir des T via une régression linéaire ou logarithmique trouverait son intérêt uniquement dans le cas particulier d'un aquifèredont la répartition des transmissivités serait décrite de façon précise (soit à partir de nombreuses valeurs ponctuelles mesurées, soit à partir d'une carte issue d'une modélisation hydrodynamique) mais ne disposant pas de valeurs de débit spécifique. Dans le cas contraire, elle est à éviter.

Annexe 3

Description et analyse structurale des données de la Région Centre

1. Sources de données et sélection d'aquifères

Les données utilisées pour réaliser les différentes séries de test se rapportent aux paramètres transmissivité (T) et débit spécifique (Q_s) des aquifères superficiels (profondeur inférieure à 100 m) de la Région Centre. Elles sont issues du recueil de données réalisé par le SGR Centre dans le cadre de l'élaboration de l'atlas régional du potentiel géothermique des aquifères de cette région (sources : BSS, ADES). Dans la suite, ces données seront désignées comme «jeu de données principal». Une source secondaire de données a également été utilisée, uniquement pour les sables du Cénomanien (Rapport BRGM R38267, 1994).

Le Tableau 3 indique les appellations retenues dans la présente étude pour ces aquifères (ce sont celles utilisées dans l'atlas du potentiel géothermique) et établi le lien avec les entités hydrogéologiques répertoriées dans la BDLisa, qui fournit les extensions et délimitations à leur associer.

		NV3
l <u></u>	Code	
Appellation Atlas GTH	Provisoire	Libellé Provisoire Entité
	Entite	
0.11		
Sables et argiles de Sologne	F2C_2	Sables et Argiles de Sologne du Burdigalien
		Calcaires de l'Orléanais et de Pithiviers de
Calcaires de Beauce	F3C_3	l'Aquitanien
Calcalles de Deadce		Calcaires d'Etampes (Calcaires du Gâtinais)
	F3C_1	de l'Oligocène
Calcaires du Berry	F4D_2	Calcaires lacustres du Berry du Priabonien
	F6U_3	Craie du Sénonien
	F6U_X	Craie du Turonien
		Tuffeau jaune du Turonien moyen à supérieur
Craie Séno-Turonien	F6U_5	(SO région Centre)
		Craie micacée et Tuffeau blanc du Turonien
	F6U_8	moyen (SO région Centre)
	F6X_1	Craie marneuse du Turonien inférieur
Craie Cénomanien		
Craie Ceriornamen	F7C_Y	Craie du Cénomanien (Nord région Centre)
		Sables du Perche du Cénomanien supérieur
Sables ACF	F7A_3	(Ouest région Centre)
Sables ACI		Sables et grès de Vierzon, Maine, Mans,
	F7A_8	Varenne du Cénomanien inférieur à moyen
		Calcaires affleurants du Kimméridgien et de
	F10E_2	l'Oxfordien (Sud région Centre)
		Calcaires non affleurants du Kimméridgien et
Calcaire Jurrassique	F10E_1	de l'Oxfordien
	F0F_1	Marnes et Calcaires de l'Oxfordien
	F11KLMOR_1	Calcaires du Bathonien-Callovien
	F11T_1	Calcaires de l'Aalénien-Bajocien
Grès Trias	F14A_1	Grès du Trias sup. (Keuper)
Gles Illas	F15A_1	Dolomies et Grès du Trias moyen à sup.

Tableau 3 : Aquifères de la Région Centre retenus pour l'étude.

Le jeu de données principal utilisé se compose de 2 289 points de données répartis sur l'ensemble du territoire de la région Centre (Figure 19), avec comme particularités : une accumulation à proximité de la Loire en amont d'Orléans et un faible taux de couverture dans la partie sud de la région.

Les données de Q_s sont beaucoup plus nombreuses que les données de T qui par ailleurs sont absentes d'importantes plages de territoire.

Seules les valeurs de Q_s non nulles sont retenues dans la base de données source. Cela entraîne un biais d'échantillonnage et implique un risque de surestimation des Q_s , en particulier dans les secteurs étendus où existent des Q_s nuls.

Les populations de points se répartissent comme indiqué dans le Tableau 4.

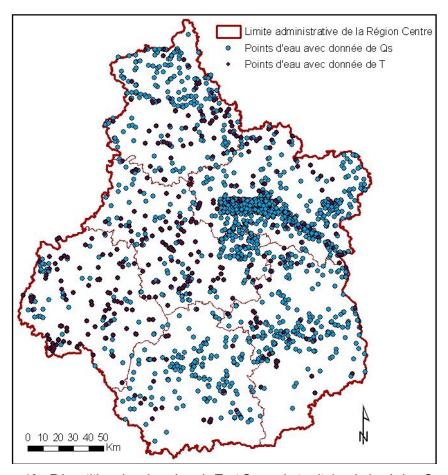


Figure 19 : Répartition des données de T et Q_s sur le territoire de la région Centre.

	Annellation Atlan CTU	Nb de valeurs / atlas		
	Appellation Atlas GTH	Q_s	T	
ss	Sables et argiles de Sologne	149	11	
	Calcaires de Beauce ¹²	954	99	
	Calcaires de Beauce affleurants	140	55	
lée	Calcaires de Beauce sous recouvrement	814	44	
Jeu de données principal	Calcaires du Berry	11	5	
	Craie Séno-Turonien	641	196	
	Craie Cénomanien	40	4	
	Sables ACF (Albien, Cénomanien,	27	27	
	Fontainebleau)	21		
	Calcaire Jurassique	271	28	
	Grès Trias	19	3	
Jeu secondaire	Sables ACF (Albien, Cénomanien, Fontainebleau)	177	153	

Tableau 4 : Distribution par ensemble aquifère des populations de données de débit spécifique (Q_s) et de transmissivité (T)

Pour être valides, les travaux qui suivent doivent être réalisés sur des jeux de données de T et Q_s suffisamment importants. Sont donc retenus les aquifères présentant le plus de données :

- Pour les données de débit spécifique : calcaire de Beauce, craie Séno-Turonienne, sables ACF (Albien, Cénomanien, Fontainebleau), calcaire jurassique ;
- Pour les données de transmissivité : aucun aquifère.

La faible quantité de données disponibles pour un aquifère et leur répartition très inégale est un critère permettant d'écarter une solution de cartographie locale – géostatistique ou autre - (exemple des grès du Trias, Figure 20, qui existent dans presque toute la région Centre, mais qui ne sont renseignés que sur sa bordure sud). Pour ces mêmes aquifères, toutes les méthodes donneront des résultats peu fiables, y compris les méthodes globales.

_

¹² Hors karsts. Onze points sont étiquetés comme karsts pour le Calcaire de Beauce, quatorze pour la Beauce sous recouvrement, deux pour la Craie du Séno-Turonien.

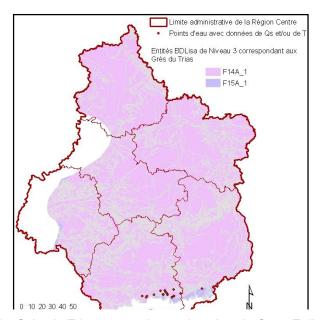


Figure 20 : Grès du Trias : extension et données de Qs et T disponibles.

2. Craie du Séno-Turonien

La craie du Séno-turonien ("CST", code BD-LISA : F6) couvre une large moitié nord de la région Centre. 643 points de mesure de Q_s (dont deux karsts : valeurs fortes) se répartissent régulièrement, excepté dans une zone au centre et au nord de l'aquifère (cf. Figure 21).

Parmi les 643 points de mesure, plusieurs se trouvent hors de l'extension de l'aquifère définie dans la BD Lisa. Nous les avons éliminés de notre sélection de données qui comprend donc :

- 509 points renseignés en débit spécifique (dont deux valeurs fortes) ;
- 160 points renseignés, en outre, en transmissivité.

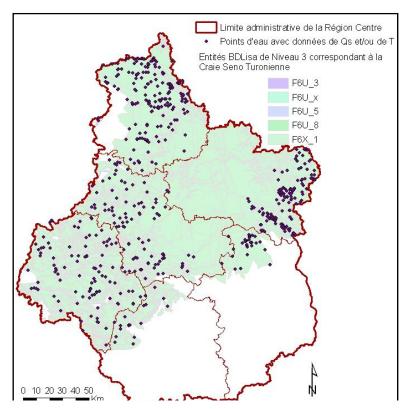


Figure 21 : Craie du Séno-Turonien : extension et données de Q_s et T disponibles.

Comme le montre l'histogramme de la , la distribution de InQ_s suit une distribution normale, malgré des valeurs fortes surreprésentées. Une comparaison sur le graphique de Henry entre InQ_s et une loi gaussienne de paramètres (-6,49 ; 1,86²) confirme l'hypothèse de normalité (cf. Figure 23), avec un risque d'erreur inférieur à 5 % (risque du premier ordre).

Il est donc pertinent de travailler, non pas en valeur directe, mais en valeurs logarithmiques.

Les deux points très forts ($lnQ_s > 0$), considérées comme des valeurs anomales, sont éliminés de notre sélection.

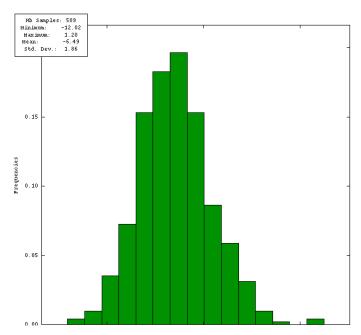


Figure 22 : Craie du Séno-Turonien : histogramme des InQs.

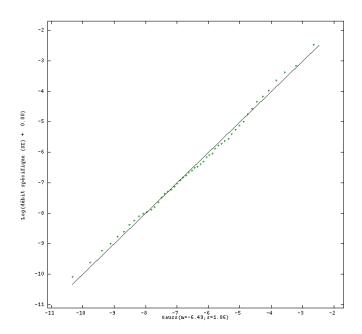


Figure 23 : Craie du Séno-Turonien : diagramme de Henry entre InQs et une distribution Gaussienne.

Le variogramme associé (cf.) montre une nette structuration, en particulier une portée d'environ 5 à 8 km qui, associée à l'importante densité de points de mesure se prête à l'application de méthodes géostatistiques à cet aquifère, sauf dans la zone vierge de points de mesure.

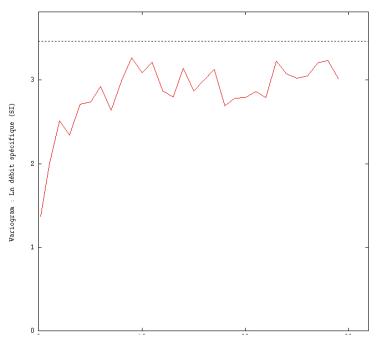


Figure 24 : Craie du Séno-Turonien : variogramme estimé de Qs.

3. Calcaires de Beauce

Cet ensemble aquifère correspond aux entités hydrogéologiques de niveau 3 F3C_3 et F3C_1 de la BD Lisa.

Les données de Q_s disponibles sont nombreuses (954 points de mesure) et bien réparties sur sa zone d'extension avec une forte concentration (environ 640 points de mesure) sur le val de Loire, au niveau d'Orléans comme indiqué en Figure 25.

Les calcaires de Beauce comprennent une partie affleurante (¼ nord-est) et une partie sous recouvrement.

Du point de vue géologique, il s'agit de calcaires intensément fissurés et localement karstifiés (avec une densité de karstification qui augmente à proximité de la Loire).

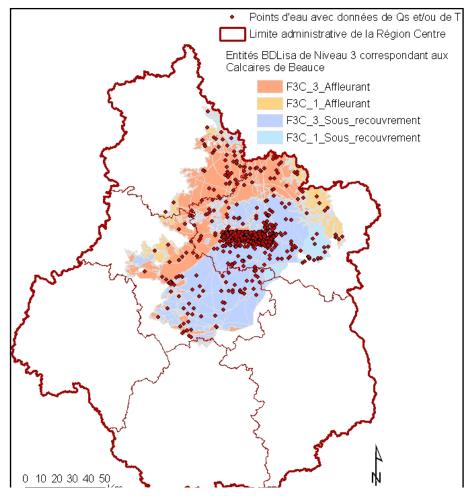


Figure 25 : Calcaires de Beauce : extension et données de Q_s et T disponibles.

3.1. Ensemble des données

Les calcaires de Beauce affleurants comptent 151 points de mesure de débit spécifique, dont 11 notés comme karsts. Les calcaires de Beauce sous recouvrement en compte 828, dont 14 karsts. Soit 954 points hors karst, 140 dans la partie affleurante (15%), et 814 dans la partie sous recouvrement (85%). Les points sont concentrés dans le cours de la Loire, le reste de la zone est couvert de manière assez homogène.

Comme le montrent les deux histogrammes (et Figure 27), la distribution des données sous forme logarithmique est plus proche d'une loi normale que la distribution des données brutes, ce qui pousse à travailler sous forme logarithmique. Toutefois, les distributions sur la partie affleurante et la partie sous recouvrement ne sont pas similaires : la première est plus aplatie que la seconde, ce qui témoigne d'une variabilité supérieure. Dans le cadre de cette étude, on peut néanmoins considérer l'ensemble comme une seule distribution à étudier.

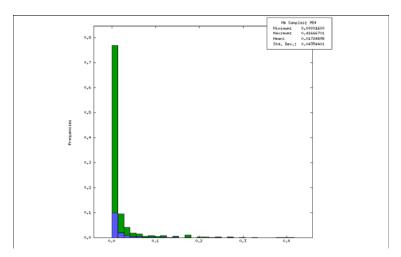


Figure 26 : Histogramme des débits spécifiques dans la Beauce. En bleu, partie affleurante, en vert, partie sous recouvrement.

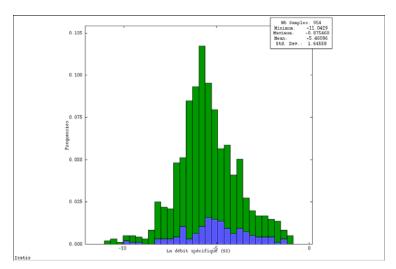


Figure 27 : Histogramme des lnQ_s dans la Beauce. En bleu partie affleurante, en vert, partie sous recouvrement.

Le variogramme de l'ensemble des données de Q_s () ne montre aucune structure dans la répartition des valeurs de Q_s . Il est même décroissant, ce qui est un non-sens mathématique et signe généralement des effets dus à la répartition hétérogène des points de mesure (amas, effets de bord, anisotropie des points de mesure...). Ce résultat confirme l'anormalité de la loi de distribution des données brutes identifiée cidessus ().

Le variogramme des données de lnQ_s () montre une certaine structure, mais un fort effet de pépite et une portée faible. Ce variogramme traduit une forte hétérogénéité dans la répartition des données qui est certainement à relier au caractère karstique des calcaires de Beauce au niveau du val de Loire.

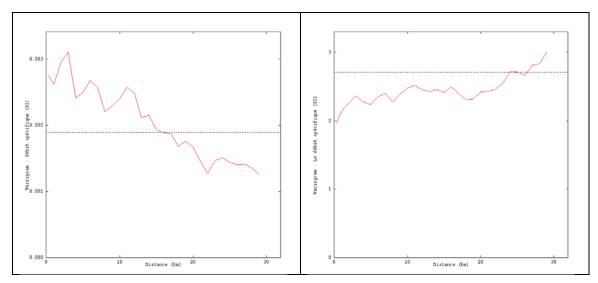


Figure 28: Calcaires de Beauce (ensemble des données) : variogramme estimé de Q_s (graphe de gauche), variogramme de lnQ_s(graphe de droite).

En distinguant les contributions des deux parties (affleurante et sous recouvrement, Figure 29), on voit que les structures apparaissant sur le variogramme global proviennent essentiellement de la partie sous recouvrement (l'essentiel des points de mesures). Le variogramme de la partie affleurante est moins bien structuré, en partie à cause du moindre nombre de points de mesures, en partie peut-être intrinsèquement (karst).

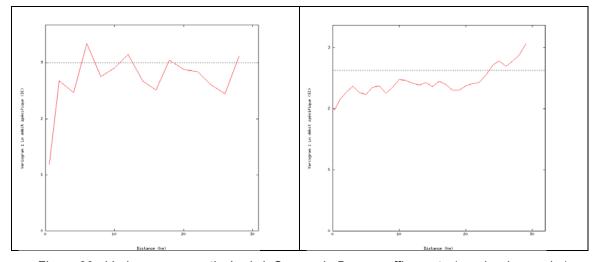


Figure 29 : Variogrammes estimés de lnQ_s pour la Beauce affleurante (graphe de gauche) et la Beauce sous recouvrement (graphe de droite).

Le variogramme estimé de InQ_s () se prête bien à une modélisation en vue d'appliquer les méthodes géostatistiques. Notons cependant que l'effet pépite élevé (au moins 65% du palier) conduira, en krigeage, à des estimations très lissées et prenant mal en compte les valeurs très fortes. De plus, on peut voir sur ce variogramme une première portée de quelques kilomètres (environ 3 km), et donc les estimations par krigeage

seront d'autant moins précises que la densité de points de mesures sera localement faible (que les estimations seront réalisées à plus de 3 km des points de mesures).

3.2. Données du Val de Loire

Le secteur du Val de Loire compte 625 points (65,5% des points de la Beauce) de mesures de Q_s répartis comme indiqué en Figure 30.

Le variogramme de InQ_s sur le Val de Loire (cf. Figure 31) a un comportement très pépitique (environ 80 %,), une portée rapidement atteinte (vers 3 km), un palier assez stable, puis une décroissance à partir de 16 km environ.

La pépite témoigne ici d'un phénomène très variable latéralement, que l'on peut rattacher aux karsts présents sur la zone. La faible portée va dans le même sens :débit spécifique guidé essentiellement par les karsts, donc très localisé. La longueur du palier montre l'absence d'autre phénomène significatif à prendre en considération. Enfin, la décroissance finale témoigne surtout d'une anisotropie de la zone, qui s'étend approximativement sur 40 km de longitude pour 13 km de latitude.

D'après ce variogramme, il ne semble pas envisageable d'évaluer le débit spécifique sur cette zone par une approche plus fine que l'approche statistique globale. Pour améliorer cela, il faudrait disposer d'une carte d'indice de présence de karsts, qui permettrait alors peut-être d'appliquer les méthodes géostatistiques.

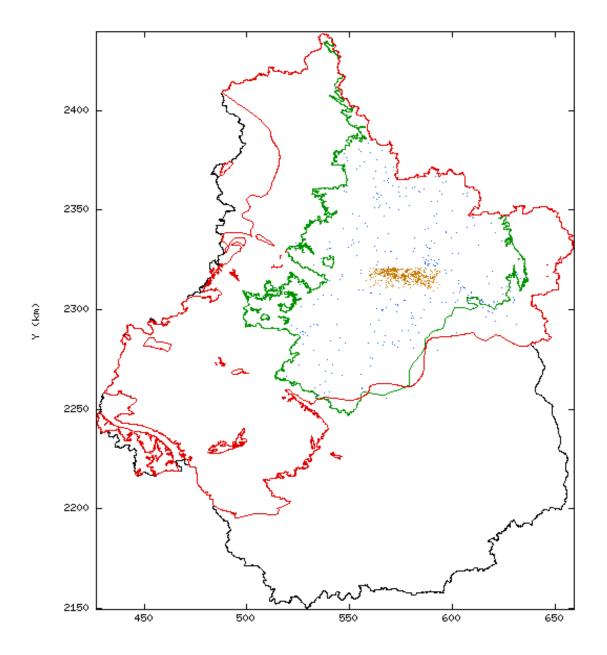


Figure 30 : Répartition des points du Val de Loire (brun) et du reste de la Beauce (bleu).

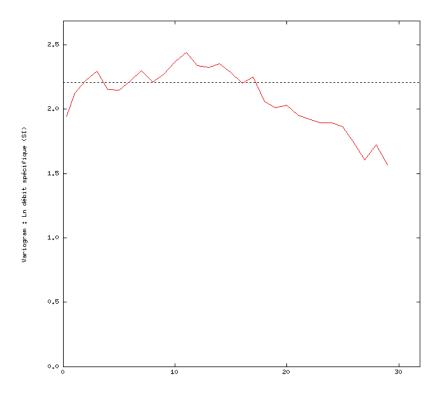


Figure 31 : Calcaires de Beauce (Val de Loire) : variogramme estimé de InQs.

3.3. Données des Calcaires de Beauce, sans le Val de Loire

Le variogramme estimé de InQ_s sur l'aquifère de la Beauce sans les données du Val de Loire () montre un comportement moins pépitique que le précédent (inférieur à 40 %) et une portée bien plus grande (au-delà de 50 km). Les pseudo-oscillations du variogramme entre 10 et 40 km peuvent provenir d'effets de voisinage, et peuvent être écartées de l'interprétation.

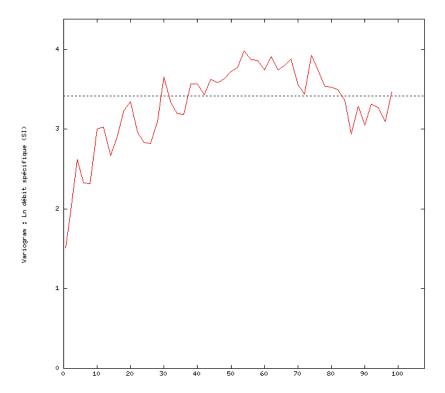


Figure 32 : Beauce, sans le Val de Loire : variogramme estimé de InQs.

En conclusion, dans cette zone moins karstique, présentant une structuration dans la répartition des débits spécifiques, les méthodes de géostatistiques pourront être appliquées.

4. Sables du Cénomanien (ACF)

Les sable de l'Albien, du Cénomanien et de Fontainebleau ("ACF", code BD-LISA : F7A) s'étendent dans la moitié ouest de la région Centre. Les points de mesure sont concentrés dans leur tiers le plus occidental.

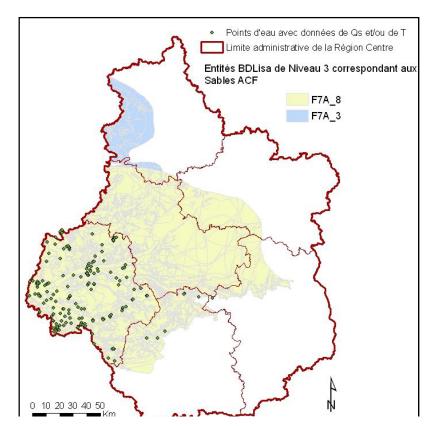


Figure 33 : Sables ACF : extension et données de Q_s et T disponibles.

Sur cet aquifère, 204 points de mesure fournissent des données de débit spécifique. Deux d'entre eux présentent des débits spécifiques très forts ($Q_s \ge 40 \text{ m}^3 \text{ h}^{-1} \text{ m}^{-1}$).

Une recherche de doublons¹³ montre 20 paires et un triplet de points de coordonnées identiques (cf. Tableau 5). À trois groupes exceptés (en gras), les valeurs de débits spécifiques sont sensiblement égales. Il est donc cohérent de garder une valeur par point.

Après élimination des doublons, il restera 182 mesures disponibles.

_

¹³ Dans la présente étude statistique, il n'est pas nécessaire de traiter ces doublons, et l'on pourrait tracer des variogrammes avec les 204 points de mesure. Cependant, les méthodes d'estimation variographique excluent la présence de doublons dans les points de mesure.

numér d'écha		Х	Υ	débit spécifi	que	
169	30	446180,00 m	2231000,00 m	0,0023806	0,0004640	
2	104	449475,00 m	2221150,00 m	0,0008640	0,0008778	
26	160	449550,00 m	2220220,00 m	0,0023370	0,0020583	
143	7	449725,00 m	2223580,00 m	0,0014167	0,0013260	
8	191	450460,00 m	2223025,00 m	0,0041510	0,0039889	
152	6	450680,00 m	2222110,00 m	0,0017583	0,0017580	
110	31	453350,00 m	2227660,00 m	0,0009806	0,0009810	
11	197	453520,00 m	2227630,00 m	0,0047640	0,0047639	
29	146	455120,00 m	2224260,00 m	0,0014960	0,0014972	
16	179	455250,00 m	2281225,00 m	0,0026060	0,0026056	
97	28	456870,00 m	2226540,00 m	0,0007944	0,0007940	
60	27	457230,00 m	2282880,00 m	0,0003361	0,0005880	
18	48	457480,00 m	2282690,00 m	0,0002220	0,0002222	
25	190	458820,00 m	2235600,00 m	0,0037970	0,0037972	
17	92	459320,00 m	2281880,00 m	0,0007420	0,0007417	
34	77	459420,00 m	2281950,00 m	0,0005030	0,0005028	
101	5	468050,00 m	2224950,00 m	0,0008611	0,0008610	
4	127	484250,00 m	2228550,00 m	0,0012060	0,0012056	
37	112	484250,00 m	2228600,00 m	0,0009860	0,0009861	
71	24	489575,00 m	2230725,00 m	0,0003972	0,0003970	
70	89 3	484800,00 m	484800,00 m	0,0003861	0,0006750 0,0006750	

Tableau 5 : Sables ACF : doublons et triplet de données.

L'analyse de l'histogramme des InQ_s (Figure 34) et de la courbe de Henry associée (Figure 35) indique que les données suivent une loi lognormale avec un risque de 8 % (risque que la loi soit normale mais non reconnue comme telle). Il est donc pertinent de travailler en valeurs logarithmiques.

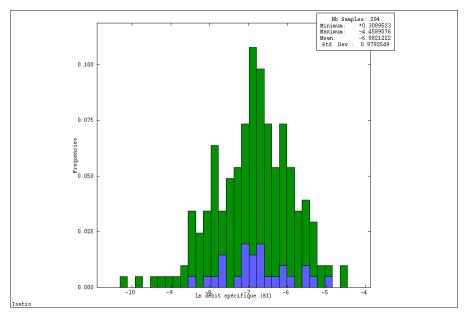


Figure 34 : Sables ACF : histogramme de InQs.

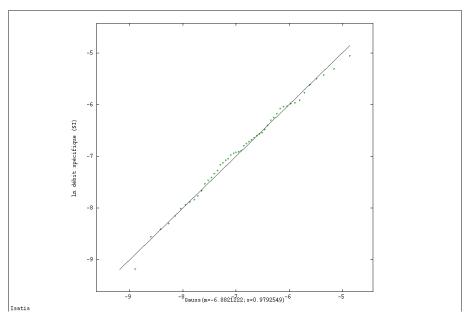


Figure 35 : Sables ACF: diagramme de Henry pour InQs.

Le variogramme des InQ_s (Figure 36) semble mettre en relief deux comportements distincts :

- Structure régionale : à une interdistance supérieure à 20 km, le variogramme oscille autour d'un palier à peu près égal à la variance (0,99);

- Structure très locale : à une interdistance inférieure à 20 km le variogramme montre une grande variabilité, oscillant autour d'une valeur de 0,75 environ

Cette partition en deux peut être le signe d'une hétérogénéité géographique des mesures. Elle nuit à la bonne modélisation du variogramme. Malgré cela, la mise en œuvre de méthodes géostatistiques reste envisageable sur cet aquifère, sauf dans la partie nord est où il n'y a pas de données.

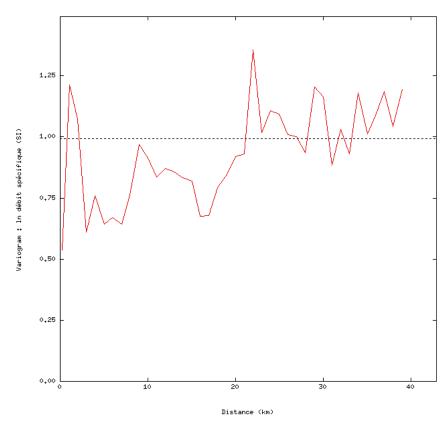


Figure 36 : Sables ACF : variogramme estimé sur 182 données(sans doublons).

5. Calcaires Jurassique

Les calcaires du Jurassique (entités $F10E_2$, $F10E_1$, $F10F_1$, $F11KLMOR_1F11T_1$ de la BD LISA) occupent l'intégralité du territoire de la Région Centre, mais les points de mesures avec données de Q_s sont concentrés sur sa partie affleurante qui correspond uniquement à l'extrême sud du territoire (cf. Figure 37). Dans ce cas, la répartition très inégale des données ne permet pas d'utiliser une approche géostatistique pour définir la répartition de propriétés physiques sur l'ensemble de l'aquifère.

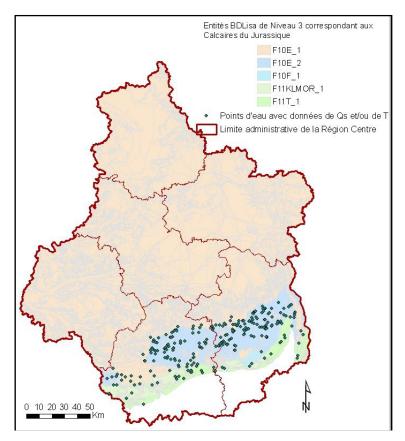


Figure 37 : Calcaire du Jurassique : répartition des données de Qs sur le territoire de la région Centre.

Le variogramme estimé de InQ_s (Figure 38) sur les calcaires du Jurassique montre un effet pépite assez faible (environ 30 %), suivi d'une structure de portée significative (environ 25 km), enfin un palier assez stable. Comme précédemment, les pseudo-oscillations de la courbe peuvent s'expliquer par effet de voisinage.

Par conséquent, Q_s est assez bien modélisable (d'après le variogramme), mais uniquement sur la marge sud de l'aquifère, où les données sont suffisamment nombreuses. Par contre, à supposer que Q_s soit stationnaire à l'échelle de l'aquifère (donc qu'on puisse parler de variogramme constant sur toute son extension), la structure mise en relief laisse à penser que les débits spécifiques évoluent assez lentement latéralement, et donc que des forages et tests de débits futurs sur la zone actuellement non couverte pourront nous apporter une information précieuse.

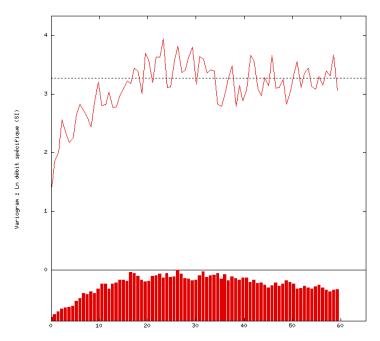


Figure 38 : Calcaire du Jurassique : variogramme estimé de lnQs ; en bas, histogramme du nombre de paires.

Annexe 4

Cartographie du critère « productivité d'aquifère » : exemple de quatre méthodes de cartographie régionale appliquées à trois aquifères de Région Centre

La cartographie géostatistique permet d'éditer des cartes prédictives d'une variable considérée. Le choix de la méthode de cartographie et de ses paramètres est à réaliser par essais, comparaisons et prise en compte de l'avis d'un expert hydrogéologue local. L'objet du présent chapitre est de tester plusieurs de ces méthodes de façon à évaluer leurs intérêts et limites respectives.

L'utilisation des données de débit spécifique sera souvent privilégiée pour établir les cartes de productivité d'aquifère. C'est donc cette variable qui a été retenue pour les tests.

Trois aquifères de la région Centre, sélectionnés pour leur représentativité de caractères hydrogéologiques classiquement rencontrés ont été pris comme exemples.

- Craie séno-turonienne : aquifère à double porosité et présentant de fortes variabilités de productivité selon son caractère « sous couverture » ou « affleurant » ;
- Calcaires de Beauce : aquifère à simple porosité (porosité de fissure) et à caractère karstique ;
- Sables cénomanien : aquifère à simple porosité (porosité d'interstices) et variogramme de données indiquant une relative homogénéité spatiale des données.

Quatre méthodes de cartographies ont été testées sur chaque aquifère (une description plus détaillée est disponible en annexe 1 (chapitre 2) :

- Statistique globale : boîtes à moustaches de Tukey (méthode développée en Région Centre lors de la réalisation de l'atlas du potentiel géothermique) :
- Statistique mobile (ou glissante) : test de plusieurs paramètres statistiques (médiane, moyenne, quartiles) et de plusieurs rayons de définition des zones de calcul;
- Krigeage;
- Simulations statistiques;
- Représentation cartographique de la probabilité 80% d'obtenir les classes de débits suivants (5, 10, 50 et 100 m³/h), calculés à partir des débits spécifiques et d'un rabattement admissible de 5 m.

Une validation des résultats de cartographie a été réalisée selon deux méthodes (cf. annexe 1 chapitre 2):

- Comparaison de lots de données « d'apprentissage » et « de contrôle » ;
- Validation croisée.

Remarque:

Dans la suite, les notations suivantes seront utilisées :

- q25 pour le 25^{ème} centile (la valeur telle que 25 % des mesures y soient inférieures)
- **2ε-3** pour la notation scientifique 2×10⁻³

1. Craie du Séno-Turonien

1.1. Méthode statistique globale

Les résultats de l'application de cette méthode à la craie du Séno-Turonien sont présentés dans le Tableau 6 et en Figure 40, Figure 41 et Figure 42.

Les données ont été scindées en deux lots (un lot « d'apprentissage » et un lot « de contrôle ») pour réaliser le contrôle des résultats. Leur répartition est présentée en Figure 39.

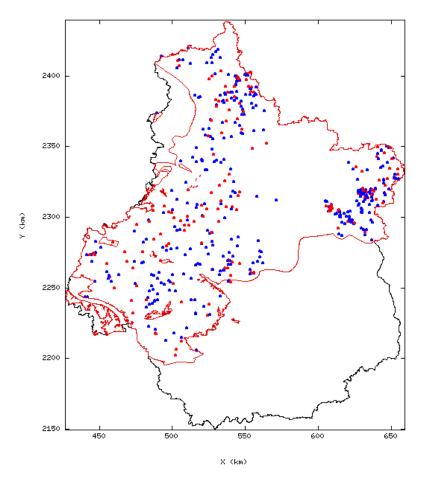


Figure 39 : Craie du Séno-Turonien : points d'apprentissage (bleu) et de contrôle (rouge).

	Variable	Effectif	Min	Max	Moy	Écart- type	q25	d20	q75
Lot ppren sage	Q_s	338	6ε−6	0,1736	7,076€−3	1,776€−2	3,860€−4	1,453€−3	4,861€−3
Lot d'appren tissage	InQ _s	338	-12,02	-1,751	-6,499	1,813	-7,860	-6,534	-5,327
de rôle	Q_s	169	1,7€−5	0,3864	8,778€−3	3,336€−2	4,170€−4	1,263€−3	3,660€−3
Lot de contrôle	InQ _s	169	-10,98	-0,9509	-6,549	1,780	-7,782	-6,674	-5,610

Tableau 6 : Craie du Séno-Turonien : statistiques élémentaires.

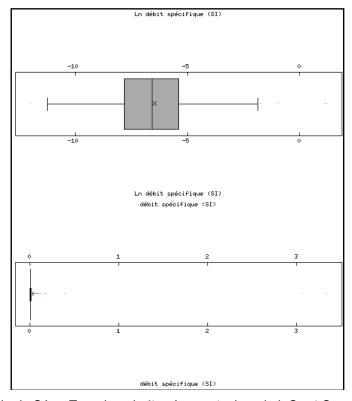


Figure 40 : Craie du Séno-Turonien : boîtes à moustaches de InQ_s et Qs sur toute la zone.

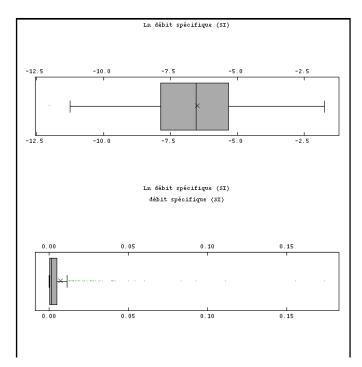


Figure 41 : Craie du Séno-Turonien : boîtes à moustaches de lnQ_s et Q_s (calculées sur lot d'apprentissage).

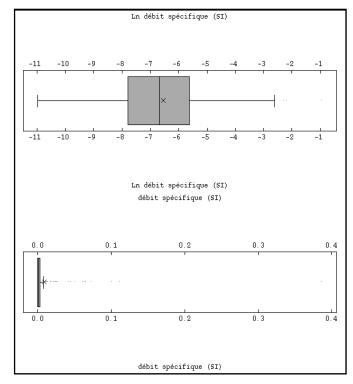


Figure 42 : Craie Séno-Turonien : boîtes à moustaches de InQs et Qs (calculées sur lot de contrôle).

La comparaison des boîtes calculées sur le lot d'apprentissage et sur le lot de contrôle ne met pas en évidence de différence flagrante, excepté un point très fort (>0,3) présent dans le jeu de contrôle.

Les données du lot d'apprentissage sont en moyenne (et en médiane) légèrement plus fortes que les données du lot de contrôle ; l'estimation cartographique présentée en Figure 43 risque donc d'être surévaluée. Cependant, l'obtention de deux lots ayant exactement les mêmes statistiques n'est pas recherchée, car l'intérêt est justement de voir comment les méthodes d'estimations prendront en compte ce biais systématique.

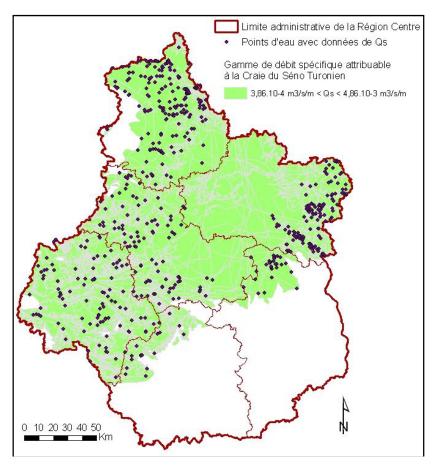


Figure 43 : Craie du Séno-Turonien : cartes de répartition des Qs obtenues à l'issue du traitement statistique « boîte à moustache ».

1.2. Méthode statistique mobile

Pour illustrer cette méthode, nous l'avons appliquée à une grille de maille 1 km, définie sur l'étendue de l'aquifère de la Craie du Séno-Turonien. La variable retenue est InQ_s . Le calcul statistique porte surla médiane et la moyenne.

Les cartes résultats et les cartes du nombre de points pris en compte dans le voisinage considéré sont fournies à chaque test.

1.2.1. Voisinage par distance

Quatre situations ont été testées :

- Test 1 : d = 50 km ; au moins 20 points de mesure ; calcul statistique = moyenne glissante ;
- Test 2 : d = 50 km ; au moins 20 points de mesure ; calcul statistique = médiane glissante ;
- Test 3 : d = 40 km ; au moins 20 points de mesure ; calcul statistique = moyenne glissante ;
- Test 4 : d = 40 km ; au moins 20 points de mesure ; calcul statistique = médiane glissante.

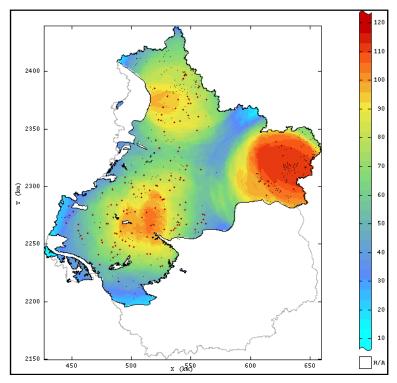
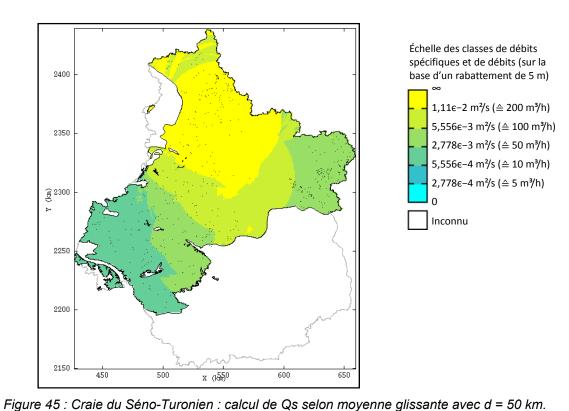


Figure 44 : Craie du Séno-Turonien : nombre de voisins pour d = 50 km.



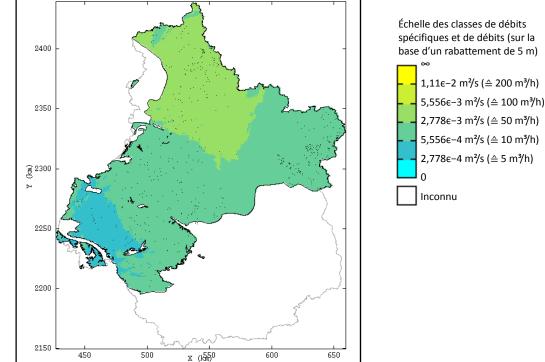
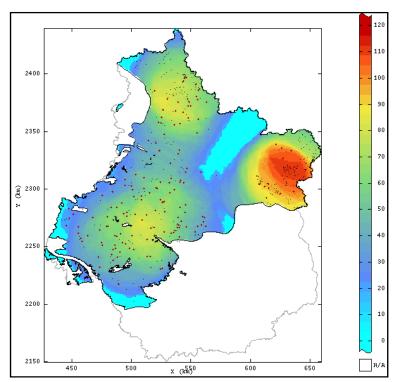


Figure 46 : Craie du Séno-Turonien : calcul de Qs selon médiane glissante avec d= 50 km.



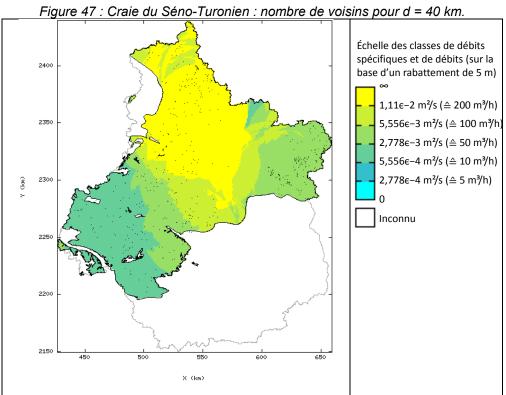


Figure 48 : Craie du Séno-Turonien : calcul de Qs selon moyenne glissante avec d = 40 km.

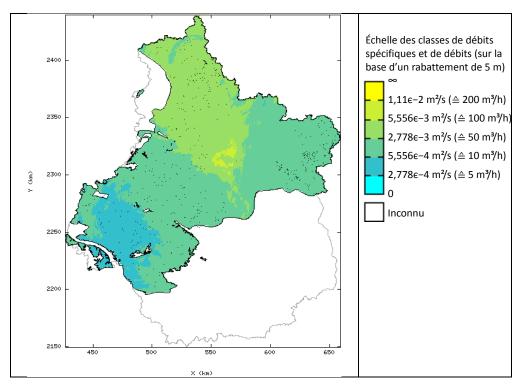


Figure 49 : Craie du Séno-Turonien : calcul de Qs selon médiane glissante avec d = 40 km.

Les résultats obtenus sont cohérents sur la répartition générale de la variable :

- accroissement général du débit spécifique dans l'orientation sud-ouest-nord-est ;
- zonation des débits spécifiques en trois grands ensembles :
 - · valeurs faibles, entre $\exp(-8)$ et $\exp(-7)$, au sud-ouest en Touraine, croissant selon SO-NE;
 - · valeurs moyennes, entre exp(-7) et exp(-6), au centre et à l'est, croissant vers l'Orléanais ;
 - · valeurs fortes, entre exp(-6) et exp(-5), dans l'Orléanais et au nord, croissant vers l'Orléanais et l'extrême pointe nord de la région.

La moyenne glissante et la médiane glissante donnent des résultats similaires pour InQ_s ; ce qui s'explique par le caractère gaussien de cette variable. Cela serait faux pour une variable non gaussienne, comme Q_s .

Le variogramme des InQ_s de la Craie du Séno-Turonien indique une portée de l'ordre de 10 km. Le voisinage à retenir pour la cartographie géostatistique devrait être de cet ordre. Or, les résultats ci-dessus montrent que la densité des données (lacune de données dans la partie nord-est) ne permet pas, pour un voisinage inférieur à 50 km, de couvrir l'intégralité de l'extension de l'aquifère. Dans ce cas, et pour que la cartographie géostatistique conserve une pertinence, il est recommandé de conserver un voisinage cohérent avec le variogramme quitte à ce que la cartographie ne couvre par l'extension complète de l'aquifère.

En conclusion, la méthode glissante sur un grand voisinage peut donner des indications sur les zonations et les tendances. Néanmoins, elle ne permet pas d'obtenir une carte à la fois cohérente (présence d'artefacts de voisinages), complète (zones non calculées) et précises (voisinage de 50 km au minimum).

1.2.2. Voisinage par nombre de points

Une autre possibilité consiste à prendre comme voisinage les n points de mesure les plus proches de chaque point de sortie.

Quatre situations ont été testées :

- Test 1 : n = 338 points de mesure (= totalité du lot d'apprentissage) ; calcul statistique = médiane glissante ;
- Test 2 : n = 160 points de mesure ; calcul statistique = médiane glissante ;
- Test 3 : n = 80 points de mesure ; calcul statistique = médiane glissante ;
- Test 4 : n = 40 points de mesure ; calcul statistique = médiane glissante.

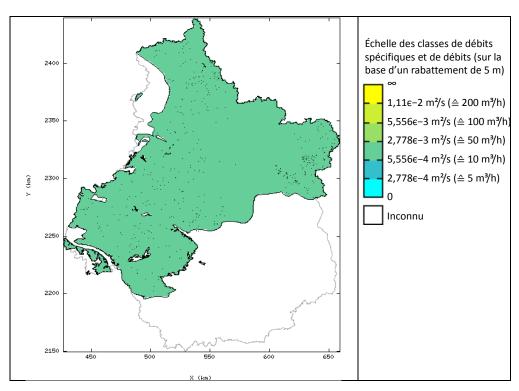


Figure 50 : Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n =338 points.

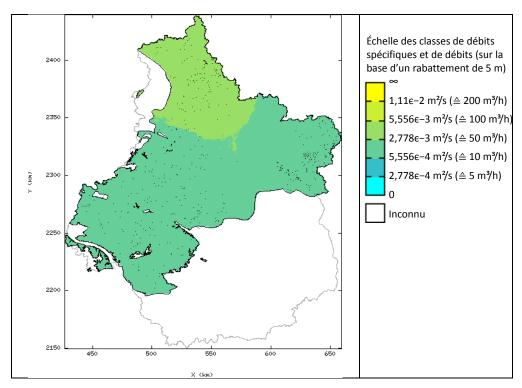


Figure 51 : Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n=160 points.

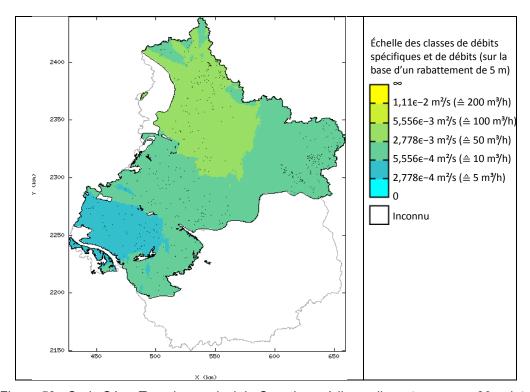


Figure 52 : Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n= 80 points.

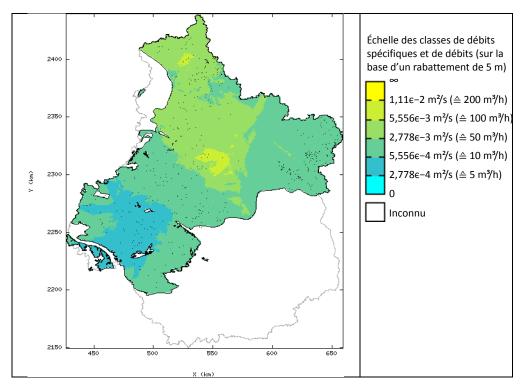


Figure 53 : Craie Séno-Turonien : calcul de Qs selon médiane glissante avec n=40 points.

Les cartes obtenues permettent de distinguer des tendances et des zonations. Le résultat est fortement sensible au nombre n.

1.3. Méthode du krigeage

La variable prise en compte pour ce test est InQ_s ,.

Deux modèles de variogramme sont proposés : γ 1 et γ 2.

1.3.1. Modèle de variogramme : γ1

Le variogramme est modélisé par trois structures :

$$y_1$$
 = Pépite(2,13) + Sphérique(10 km; 0,78) + Sphérique(40 km; 0,36)

L'effet de pépite est fort (65 %). La première portée partielle de 10 km semble découler de l'accumulation de points espacés de moins d'une dizaine de kilomètres. La seconde portée partielle de 40 km est très grande par rapport aux dispositions géographiques des points.

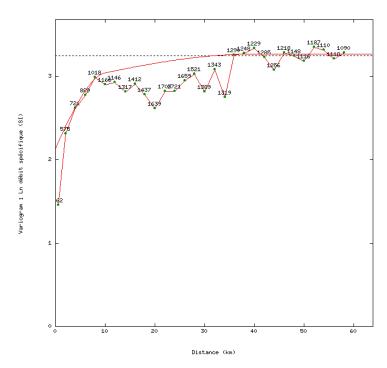


Figure 54 : Variogramme estimé (rouge fin) et modèle de variogramme γ₁ (rouge épais). Les nombres indiquent les quantités de doublets de valeurs.

1.3.2. Modèle de variogrammeγ2

On remarque que cette seconde option donne un variogramme sensiblement équivalent à la première. Le modèle de variogramme ici choisi s'attache plus à décrire la structure à faible distance.

$$y_2$$
 = Pépite(1,13) + Exponentielle(5 700 m; 1,6) + Sphérique(45 000 m; 0,52)

Le choix entre ces deux modèles dépend surtout de l'effet pépite, et donc de l'importance que l'on accorde à ces 62 couples de points distants entre eux de moins d'un kilomètre. Dans toute la suite, on utilisera y_1 , sauf mention contraire.

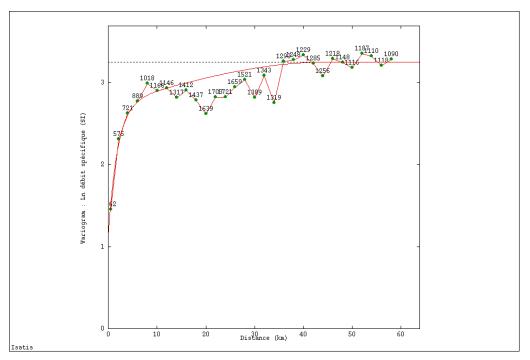


Figure 55: Modèle de variogramme γ_2 .

1.3.3. Krigeage

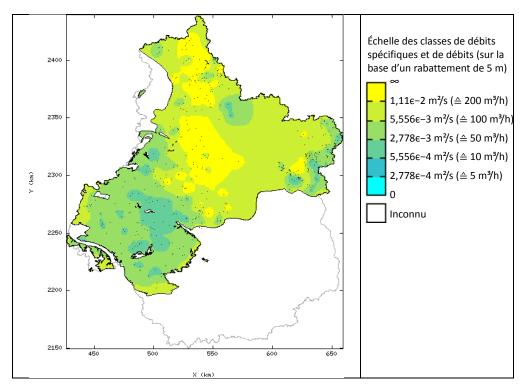


Figure 56 : Krigeage lognormal de Q_s selon γ_1 , voisinage unique.

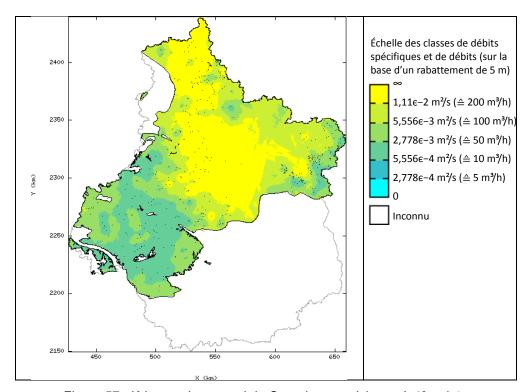


Figure 57 : Krigeage lognormal de Q_{s} , selon γ_1 , voisinage à 40 points.

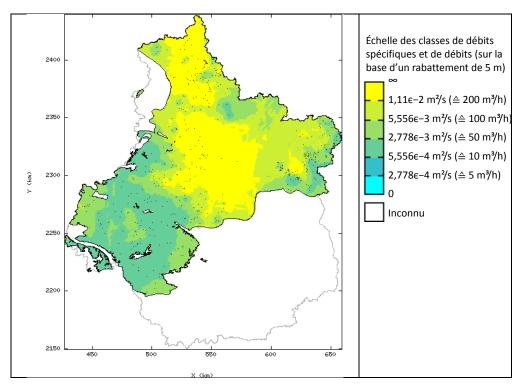


Figure 58 : Krigeage lognormal de Q_s , selon γ_1 , voisinage à 20 points.

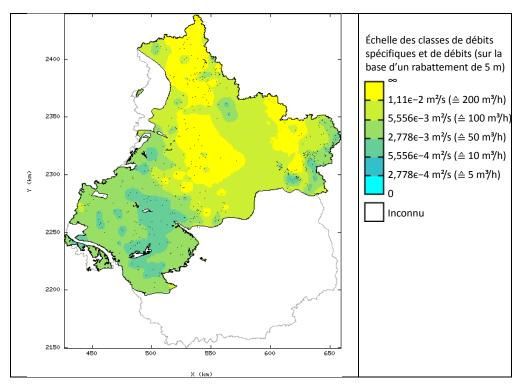


Figure 59 : Krigeage lognormal de Q_s , selon γ_1 , voisinage à 100 km.

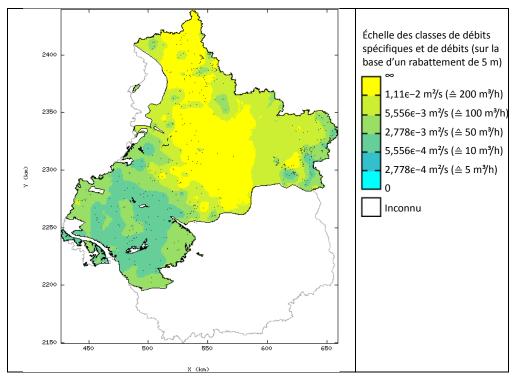


Figure 60 : Krigeage lognormal de Qs, selon γ_1 , voisinage à 50 km.

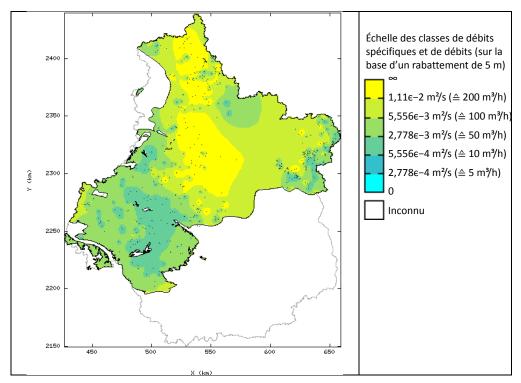


Figure 61 : Krigeage lognormal de Q_s selon γ_2 , voisinage unique.

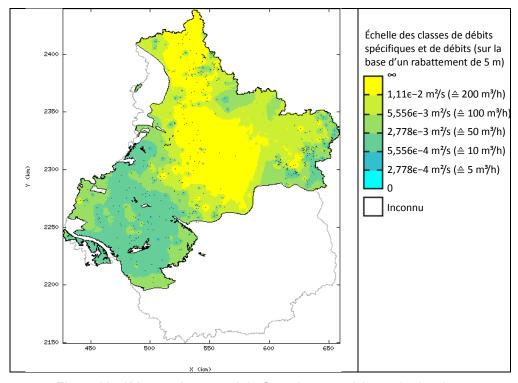


Figure 62 : Krigeage lognormal de $Q_{\text{s}},$ selon $\gamma_2,$ voisinage à 40 points.

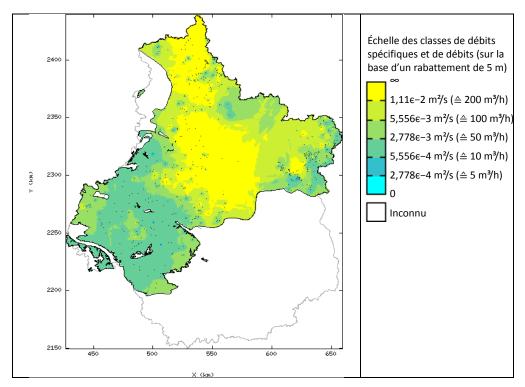


Figure 63 : Krigeage lognormal de Q_s , selon γ_2 , voisinage à 20 points.

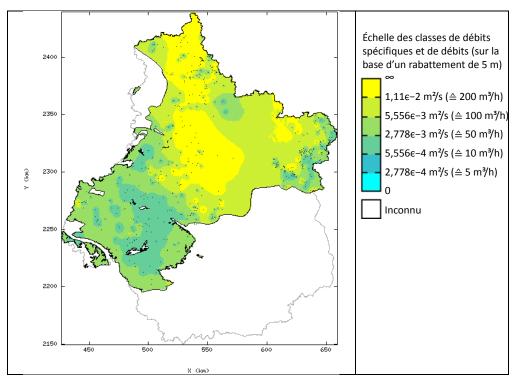


Figure 64 : Krigeage lognormal de Q_s , selon γ_2 , voisinage à 100 km.

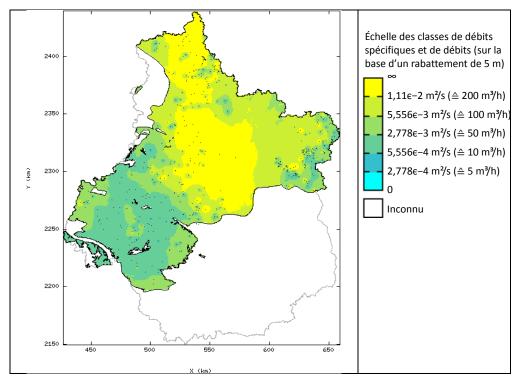


Figure 65 : Krigeage lognormal de Qs, selon γ₂, voisinage à 50 km.

Les estimations par krigeage donnent des cartes plus hétérogènes que les méthodes mobiles, même à voisinage grand (voire unique). Même si ce fait seul ne permet pas de conclure sur la qualité des estimations, elles semblent permettre des régionalisations plus puissantes. La réduction des tailles de voisinage donne des cartes de plus en plus "granuleuses", mais rien ne permet de dire si ces variations représentent des phénomènes réels ou des dispersions aléatoires des mesures. La comparaison entre les résultats selon les deux modèles de variogrammes montre des différences locales et légères, qui ne remettent pas en cause l'allure générale des cartes.

1.3.4. Contrôle des résultats par validation croisée

Isatis permet la représentation graphique d'une validation croisée, mais uniquement dans le cas d'une estimation par krigeage direct, c'est-à-dire par krigeage de lnQ_s (cf. p 52). Cette quadruple représentation graphique permet de détecter éventuellement des points ou zones aberrants.

La Figure 66 présente les résultats de la validation croisée sur tous les points de la Craie du Séno-Turonien, en utilisant le variogramme γ_1 et un voisinage à 50 km, et la Figure 67 en utilisant le variogramme γ_2 .

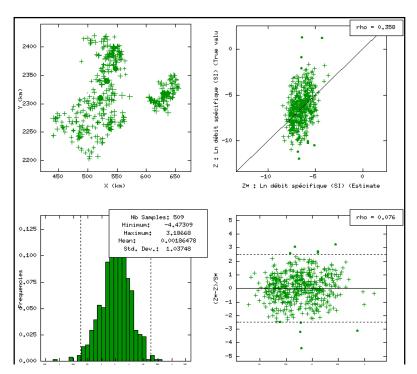


Figure 66 : Validation croisée selon γ1.

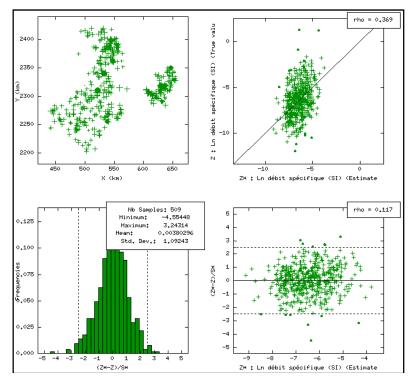


Figure 67 : Validation croisée selon γ2.

Ces deux validations croisées donnent des résultats similaires, ce qui laisse à penser que le choix précis du variogramme influe peu sur les résultats. Les validations croisées semblent acceptables (peu de valeurs anormales). Par contre, les estimations ont tendance à fortement lisser les valeurs, ce dont témoigne le graphique croisé (haut à gauche), où les points sont concentrés dans une ellipse de grand axe presque vertical. Notons que les autres méthodes provoquent un phénomène similaire, et généralement plus marqué : les méthodes globales donnent un résultat unique, donc "lissé à l'extrême", et les méthodes mobiles ont une forte tendance à lisser les résultats, et donc à mal représenter les valeurs très fortes, comme le montrent les cartes ci-dessus et comme le montreront les comparaisons entre les méthodes, ci-dessous.

1.4. Simulations géostatistiques

Les simulations suivantes sont réalisées sur InQ_s sur voisinage unique, à partir du variogramme γ_1 et des mesures sur la CST sauf deux (très fortes), sous la forme de 1 000 simulations utilisant chacune 200 bandes tournantes.

Nous obtenons donc 1 000 cartes simulées de lnQ_s , ou dit autrement 1 000 valeurs possibles de lnQ_s en chaque point d'estimation. Pour les interpréter, on peut construire la carte agrégée par exemple des moyennes géométriques de Q_s (c'est-à-dire exponentielle des moyennes de lnQ_s), des médianes ou de guantiles donnés.

On peut également à partir des simulations de lnQ_s obtenir 1 000 simulations de Q_s , par simple passage à l'exponentielle. Néanmoins, ce passage, pour simple qu'il semble, demande à être validé par des considérations théoriques.

Commençons par observer les écarts-type en chaque point des simulations.

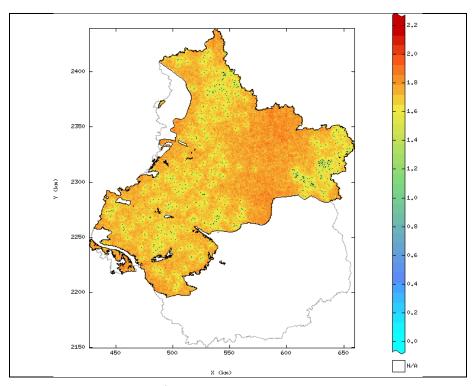


Figure 68 : Écart-type sur les simulations de InQ_s.

Les écarts-types des simulations de InQ_s (Figure 68) évoluent entre 0 et 2,2. Ils sont proches de zéro autour des points de mesures, car ceux-ci forcent les simulations à avoir des valeurs compatibles avec les mesures, donc très proches entre elles ; les écarts-types autour des points de mesure seront d'autant plus faibles que le modèle de variogramme choisi aura un effet pépite faible et montera lentement (grande interdépendance à faibles interdistances). Lorsqu'on s'éloigne des points de mesure, les écarts-type se rapprochent d'un plafond d'environ 2,2, ce qui est du même ordre de grandeur que l'écart-type des mesures, 1,8 ici¹⁴. L'aspect granuleux des cartes d'écart-type de InQ_s et de Q_s ci-dessous montre que le nombre de simulations n'est pas suffisant pour des évaluations précises, même si à notre niveau il reste suffisant pour une première approche.

Il convient de ne pas sur-interpréter ces résultats. Ils traduisent la dispersion des simulations, et donc donnent une approche des imprécisions sur les estimations, cependant ils supposent un cadre théorique (phénomène continu et stationnaire, modèle de variogramme), qui n'est pas forcément respecté, et donc peuvent sous-estimer cette dispersion.

_

¹⁴ À notre connaissance, la théorie ne donne pas de valeur pour l'écart-type de krigeage au loin des points de mesure.

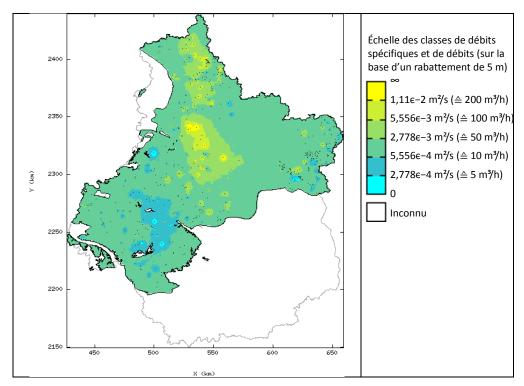


Figure 69 : Estimation de Q_s par exponentielle de la moyenne des simulations

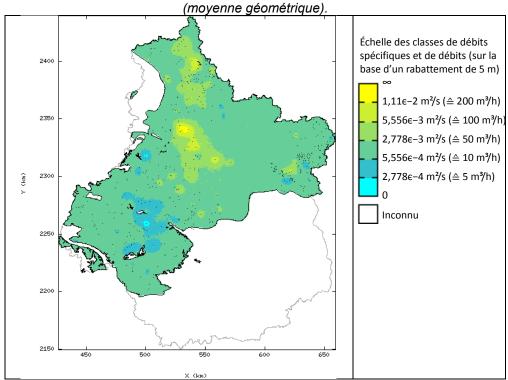


Figure 70 : Estimation de Qs par exponentielle de la médiane des simulations (~médiane).

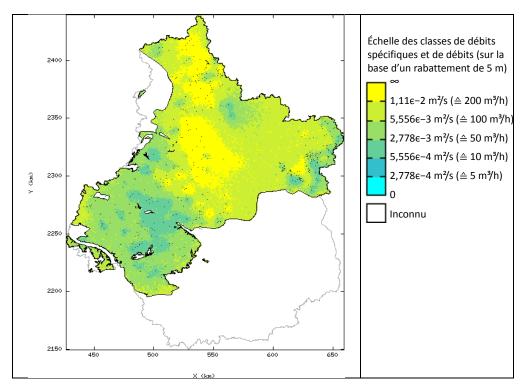
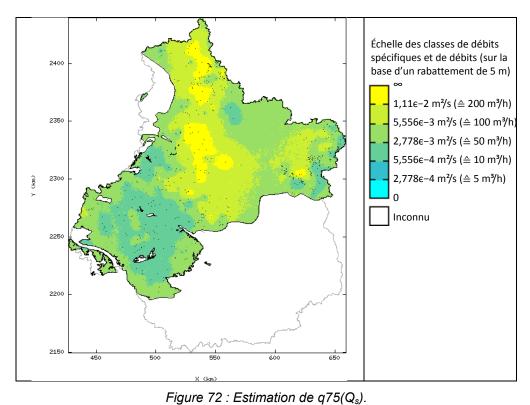


Figure 71 : Estimation de Q_s par moyenne des exponentielles des simulations (moyenne directe).

On peut calculer à partir des simulations, en tout point, leur moyenne géométrique (Figure 69), leur médiane (Figure 70), et leur moyenne directe (Figure 71). Cette dernière carte est proche du résultat du krigeage lognormal (, p. 105), ce qui est prédit par la théorie. Néanmoins, quelques développements théoriques doivent être menés avant de pouvoir confirmer cette approche.

Enfin, les deux figures suivantes montrent des exemples de cartes supplémentaires apportées par les simulations. La Figure 72 montre en tout point le quantile 75 % des simulations, c'est-à-dire la valeur atteinte ou dépassée par ¼ des simulations. La Figure 73 montre en tout point la proportion de simulations dépassant 5,556 ε –4 m²/s (soit 10 m³/h sous nos hypothèses), et donc la probabilité qu'une simulation dépasse ce seuil. Ces deux figures montrent ainsi comment obtenir mieux qu'une estimation ponctuelle.

Encore une fois, il convient de ne pas sur-interpréter les résultats : dans les zones où manque la donnée, les traitements sur les simulations ont tendance à ne pas régionaliser le phénomène : la moyenne des simulations tendra vers la moyenne des mesures, par conséquent les résultats seront inexploitables localement.



Tigure 12 . Estimation de q15(Qs).

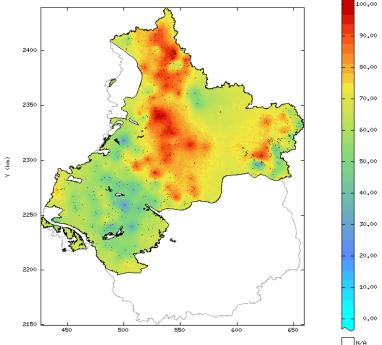


Figure 73 : Proportion de simulations dépassant 5,556 ϵ -4 m²/s.

1.5. Comparaison des différentes méthodes

1.5.1. Comparaison des cartes

Sont comparées ici (Figure 74) les résultats cartographiques obtenus par méthodes :

- statistiques mobiles (moyenne et médiane glissante) ;
- krigeage variographique;
- simulation variographique.

Par rapport aux estimations par krigeage, il apparaît que les méthodes mobiles ont tendance à sous-estimer les débits spécifiques, sauf pour la moyenne directe mobile.

La Figure 74 compare quatre approches d'estimation de Q_s sur la Craie du sénoturonien. Cette comparaison doit se faire en gardant à l'esprit que les conditions et les paramètres d'estimations influent fortement sur le résultat. On constate que, comme il s'agit d'estimer la valeur espérée de Q_s , la médiane glissante donne un résultat différent des autres méthodes, et fortement sous-estimé. Les cartes produites par les trois autres méthodes présentent des différences et également des effets de voisinages pour la moyenne directe glissante (arcs de cercles visibles), des effets de palier pour le krigeage (bulles d'une dizaine de kilomètres autour des points de données), et une granulosité pour les simulations (nécessité d'un très grand nombre de simulations).

Enfin, ces cartes peuvent donner la fausse impression que l'estimation peut être bien contrainte. Or le variogramme montre un quasi-palier assez faible : au-delà d'une dizaine de kilomètres, la valeur attendue est quasiment indépendante des valeurs déjà mesurées. Il s'agit donc de ne pas oublier le caractère peu prédictible de Q_s à des distances myriamétriques des points de données.

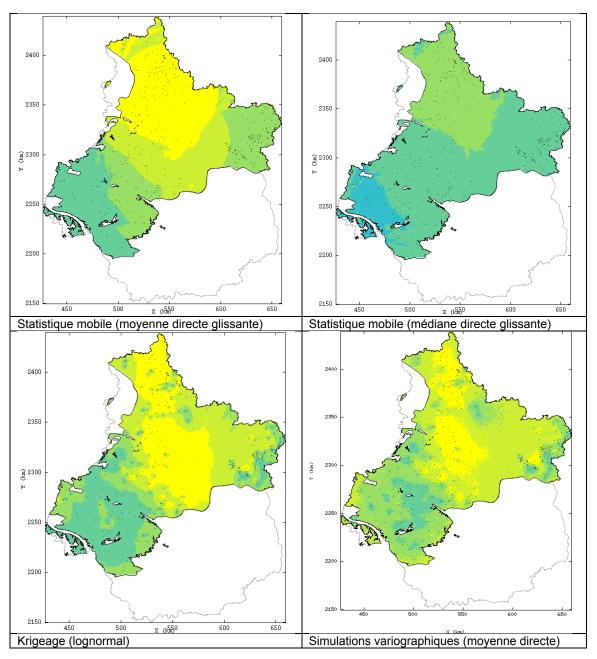


Figure 74 : Craie du séno-turonien : comparaison de cartes obtenues par différentes méthodes.

1.5.2. Comparaison des résultats de validation des cartographies

Les tableaux suivants donnent les résultats de validation des méthodes d'estimation utilisées ci-dessus. Le contrôle par comparaison apprentissage/contrôle (voir p. 51) a été utilisé; il fournit des couples valeur mesurée / valeur estimée. Ces couples sont ensuite agrégés pour pouvoir comparer les méthodes entre elles (p. 53).

Le tableau 7 compare les agrégations usuelles : moyenne, écart-type, biais, erreurs absolue et quadratique.

méthode	krigeage lognormal (modèle γ ₁)			krigeage lognormal (modèle γ₂)				simulati on*		diane ite sur Q _s	moyenn e globale			
voisinage	unique	40 points	20 points	100 km	50 km	unique	40 points	20 points	100 km	50 km	unique	6 points	20 points	unique
moyenne	8,20 ^{€-3}	8,34 ^{€-3}	8,91 ^{€-3}	8,40 ^{€-3}	8,34 ^{€−3}	7,76 ^{€-3}	7,97 ^{ε−3}	8,22 ^{€-3}	7,87 ^{€-3}	7,85 ^{ε−3}	2,00 ^{€-3}	4,94 ^{€-3}	2,56 ^{€-3}	7,08 ^{€-3}
écart-type	5,01 ^{ε−3}	5,82 ^{ε−3}	7,87 ^{ε−3}	5,27 ^{ε−3}	5,82 ^{ε−3}	5,73 ^{ε−3}	7,04 ^{ε−3}	7,82 ^{ε−3}	5,87 ^{ε−3}	6,35 ^{€-3}	1,30 ^{€-3}	8,43 ^{ε−3}	2,62 ^{€−3}	0
biais	1,67 ^{ε−3}	1,81 ^{€-3}	2,38 ^{ε−3}	1,87 ^{ε−3}	1,81 ^{ε−3}	1,23 ^{€-3}	1,44 ^{€-3}	1,69 ^{€-3}	1,34 ^{ε−3}	1,32 ^{€-3}	−4,53 ^{ε-3}	−1,59 ^{ε−}	−3,97 ^{ε−3}	5,45 ^{€-4}
erreurs absolues	8,92 ^{€-3}	8,95 ^{€-3}	9,46 ^{€-3}	8,98 ^{€−3}	8,95 ^{€-3}	8,49 ^{€-3}	8,71 ^{€-3}	8,85 ^{€-3}	8,51 ^{€-3}	8,54 ^{€-3}	5,98 ^{€-3}	7,33 ^{€-3}	6,15 ^{€-3}	8,80 ⁶⁻³
erreurs quadratiques	1,58 ^{ε−2}	1,58 ^{€−2}	1,66 ^{€-2}	1,57 ^{€-2}	1,58 ^{€−2}	1,58 ^{€-2}	1,62 ^{€-2}	1,65 ^{€-2}	1,57 ^{€-2}	1,59 ^{€-2}	1,66 ^{€-2}	1,68 ^{€-2}	1,65 ^{€-2}	1,63 ^{€-2}

Tableau 7 : Comparaison des statistiques de plusieurs méthodes d'estimation.

*) exponentielle de la valeur moyenne des simulations en chaque point, variogramme γ1

Toutes les méthodes expriment un biais.

Celui-ci est le plus faible pour la moyenne globale. De fait, le biais de la moyenne globale devrait être nul si les lots d'apprentissage et de contrôle étaient similaires (avait même moyenne); le biais sur la moyenne globale montre que cela n'est pas le cas. Par contre, on aurait pu imaginer que les méthodes géostatistiques échappent à cet effet d'altérité entre les deux lots. Ce n'est pas le cas ici. Cela attire l'attention sur les limites des méthodes d'estimations; nous avons déjà insisté sur le fait que les estimations ponctuelles pouvaient être très différentes des mesures "réelles". On voit ici que des estimations sur un groupe de points peuvent également être biaisées, bien que de manière moindre.

Le modèle de variogramme γ_2 donne des résultats légèrement meilleurs que γ_1 en erreurs absolues. La prise en compte des corrélations à faible distance (<10 km) présente donc un intérêt.

Les erreurs absolues et les erreurs quadratiques ne mettent pas en lumière de méthode préférable.

Le tableau 8 compare les proportions d'estimations "assez proches" des mesures, c'est-à-dire dans une marge d'erreur acceptable. Cette marge est indiquée dans la première colonne.

métho de	krigeag	krigeage lognormal (modèle γ₁)			krigeage lognormal (modèle γ₂)				simulati on*	médian glissan Q _s	-	moyen ne global e		
marge	unique	40 poi	20 poi nts	100 k m	50 km	unique	40 poi	20 poi nts	100 k m	50 km	unique	6 point	20 poi nts	unique
à 0,001	5,36%	7,74%	7,74%	5,36%	7,74%	7,14%	10,12 %	12,50 %	7,14%	10,12 %	38,69%	33,33 %	40,48 %	4,17%
à 0,003	22,02 %	28,57 %	32,14 %	22,62 %	28,57 %	33,93 %	40,48 %	42,26 %	37,50 %	38,69 %	76,19%	67,26 %	70,83 %	9,52%
à 0,005	49,40 %	49,40 %	50,00 %	50,00 %	49,40 %	56,55 %	55,95 %	57,74 %	56,55 %	56,55 %	83,33%	77,38 %	83,33 %	26,79%
à 0,01	77,98 %	76,79 %	75,00 %	76,79 %	76,79 %	76,79 %	76,79 %	76,79 %	76,19 %	77,38 %	88,69%	83,93 %	88,10 %	89,88%
à 0,06	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	98,81 %	97,02%	97,62 %	97,02 %	98,21%

Tableau 8 : Tableau des proportions d'estimations proches des valeurs réelles ; la première colonne indique la limite de proximité choisie.

On voit que les méthodes glissantes employées ici donnent de bons résultats. Ainsi, pour une médiane glissante sur un voisinage à 20 points, 41 % des estimations sont justes à 0,001 près. À l'inverse, seules 5 à 13 % des estimations par krigeage sont justes à 0,001 près. Ce fait semble paradoxal : les différents krigeages donnent des estimations moins concentrées autour des mesures que les méthodes glissantes ou globales testées.

Au-delà d'une certaine marge (cf. dernière ligne du tableau), la médiane glissante donne des résultats inférieurs au krigeage. L'explication est un peu complexe et nous ne la détaillerons pas ici. Elle est basée sur le fait que la distribution est lognormale, sur un voisinage la médiane est souvent un indicateur proche de la valeur la plus probable que la moyenne¹⁵. En contrepartie, lorsque la valeur mesurée est forte (dans la queue de la distribution lognormale), la médiane glissante donne une valeur beaucoup plus éloignée de la mesure que les autres méthodes.

En définitive, la médiane glissante peut être préférée à la moyenne glissante lors de la recherche de valeurs que l'on sait faibles (d'après l'histogramme de distribution). Par contre, pour établir une cartographie, nous préconisons les méthodes recherchant une moyenne locale (moyenne glissante ou krigeage lognormal). Les méthodes de simulations offrent l'avantage de fournir les deux angles d'approche, et pourront être préférées si le temps de travail et les données le permettent.

_

¹⁵ Pour une distribution lognormale de paramètres μ et σ , la médiane e^{μ} est plus proche du mode (valeur la plus fréquente) $e^{\mu-\sigma^2}$ que l'espérance (valeur moyenne) $e^{\mu+\sigma^2/2}$.

2. Calcaires de Beauce

Les données des calcaires de Beauce ont été traitées selon le contrôle par validation croisée.

2.1. Méthode statistique globale

Les résultats de l'application de cette méthode aux calcaires de Beauce sont présentés dans le Tableau 6 et en Figure 75.

Variable	Effectif	Min	Max	Moy	Écart- type	q25	d20	d75
Qs	954	1,600€−5	4,167ε−1	1,725€−2	4,354€−2	1,587€−3	3,663€−3	1,175€−2
InQs	954	-11,04	-0,8755	-5,461	1,646	-6,446	-5,609	-4,444

Tableau 9 : Calcaire de Beauce : statistiques élémentaires.

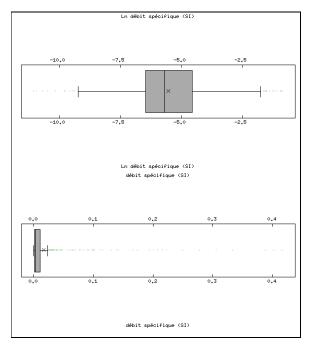


Figure 75 : Calcaires de beauce : boîtes à moustaches de lnQ_s et Q_s sur toute la zone.

2.2. Méthode statistique mobile

Pour illustrer cette méthode, nous l'avons appliquée à une grille de maille 1 km, définie sur l'étendue de l'aquifère des calcaires de Beauce. Les variables retenues sont Q_s et InQ_s . Le voisinage retenu pour ce test est un voisinage de distance avec d=10 km. De ce fait, certaines zones ne seront pas cartographiées (zones sans donnée à moins de 10 km). Cette approche permet de tester l'effet de petits voisinages sur les méthodes, d'autant mieux que des points de mesure forment un amas le long de la Loire.

Les cartes suivantes illustrent le comportement de Q_s d'après les méthodes mobiles : médiane, moyenne, moyenne géométrique (moyenne sur les logarithmes) et médiane géométrique (sur les logarithmes).

Comme dans le cas de la craie du Séno-turonien :

- les résultats obtenus sont cohérents sur la répartition générale de la variable ;
- la moyenne glissante et la médiane glissante donnent des résultats similaires pour lnQ_s ; ce qui s'explique par le caractère gaussien de cette variable. Cela est faux pour la variable non gaussienne qu'est Q_s .

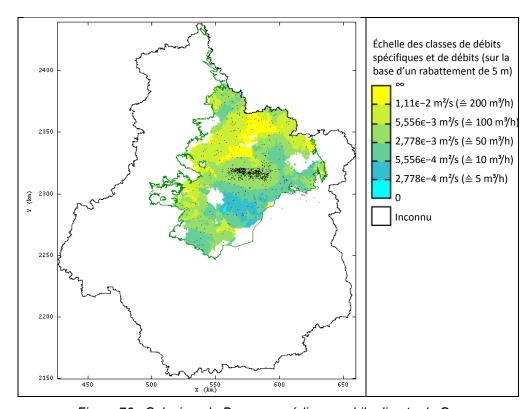
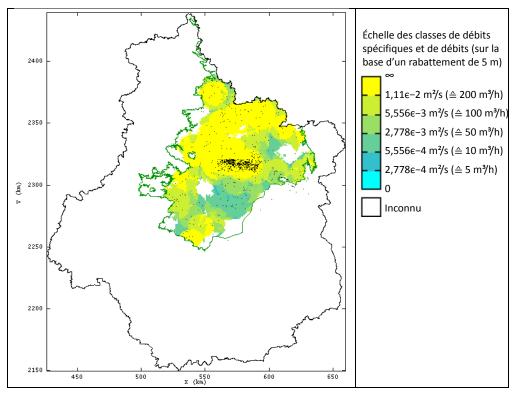


Figure 76 : Calcaires de Beauce - médiane mobile directe de Qs.



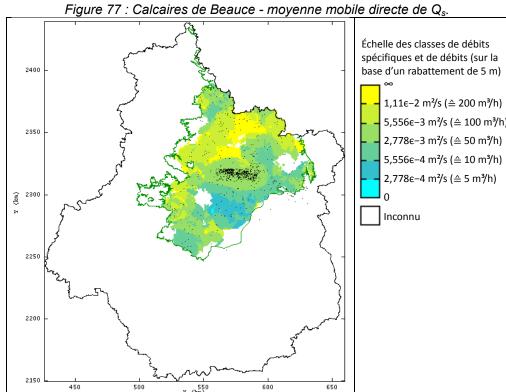


Figure 78 : Calcaires de Beauce - médiane mobile sur les logarithmes de Q_s.

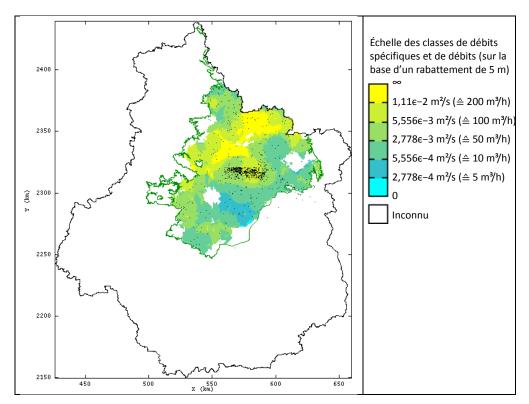


Figure 79 : Calcaires de Beauce - moyenne mobile sur les logarithmesde Q_s.

2.3. Méthode du krigeage

La variable prise en compte pour ce test est InQs.

2.3.1. Variogramme

Le variogramme estimé de *InQ*_s sur les calcaires de Beauce est très peu structuré :

- très important effet de pépite (environ 2 sur une variance de 2,7, soit environ ¾). Le phénomène est très peu spatialisé, et donc les méthodes géostatistiques auront en général tendance à fournir des estimations fortement lissées. Les éventuelles zones de très fort débit, si elles sont de taille réduite, n'apparaîtront sans doute pas.
- structure de faible portée (sur les deux premiers kilomètres). Elle semble n'être créée que par le premier point du variogramme. Un regard sur le variogramme à petites distances montre néanmoins que cette structure existe.
- comportement plat entre 2 et 24 km.
- comportement linéaire croissant au-delà. Il trahit une dérive, pouvant être due à la répartition des points dans l'espace.

En résumé, on observe une structure très peu marquée sur les quelques premiers kilomètres, et un fort effet pépite. La modélisation du variogramme ne permettra donc pas de rendre une carte fine du débit spécifique.

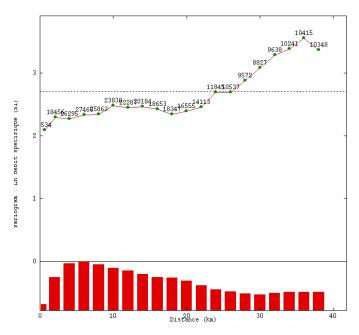


Figure 80 : Calcaires de Beauce - variogramme estimé de InQs.

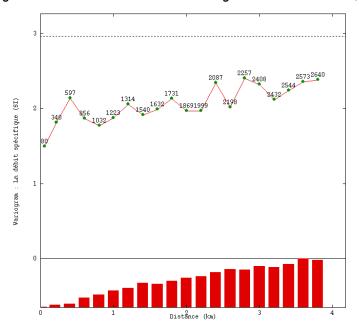


Figure 81 : Calcaires de Beauce - variogramme estimé de InQs à très petite distance.

2.3.1. Variogrammey₁

On propose un premier modèle du variogramme γ_1 décrit par la relation :

```
y_1 = Pépite(1,49) + Sphérique(1310 m; 0,39) + Linéaire(20 640 m; 0,53)
```

L'effet de pépite modélise le comportement faiblement auto-corrélé du phénomène, le sphérique montre une structure à très courte portée (kilométrique), la dérive linéaire suit le comportement à interdistance moyenne.

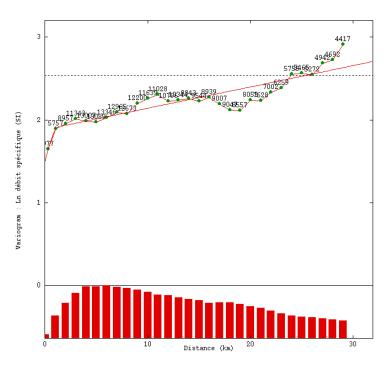


Figure 82 : Calcaires de Beauce - Modèle de variogramme Y1.

2.3.2. Variogrammey₂

Un deuxième modèle de variogramme, sans dérive, c'est-à-dire avec stationnarité du variogramme à grandes distances ("palier"), est proposé ; il est décrit par la relation :

$$y_2$$
 = Pépite(1,73) + Sphérique(1 100 m; 0,36) + Cubique(35 000 m; 0,53)

Ce variogramme et le précédent ont des valeurs sensiblement égales pour les premiers kilomètres (≤ 10 km). Néanmoins, celui-ci possède maintenant un palier, égal à la variance des données (quelques points aberrants ont été négligés, pour un meilleur résultat).

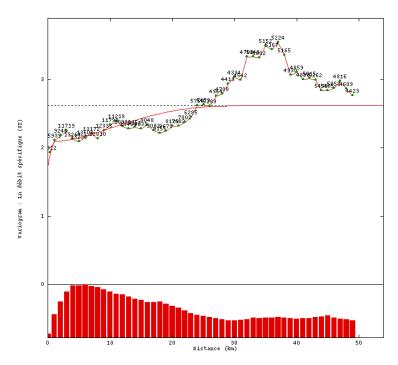


Figure 83 : Calcaires de Beauce - Modèle de variogramme Y2.

2.3.3. Krigeage

La carte de la Figure 84 présente le résultat du krigeage lognormal selon γ_1 , à voisinage unique, sur une grille de 1 km. Le jeu de données contient quelques doublons, ils sont négligés.

La carte du krigeage lognormal de InQ_s selon γ_2 (Figure 85) montre quelques différences avec celle de γ_1 . En particulier, l'influence de la présence ou de l'absence de dérive est bien visible le long de la bordure sud-est de la zone, où peu de données sont disponibles.

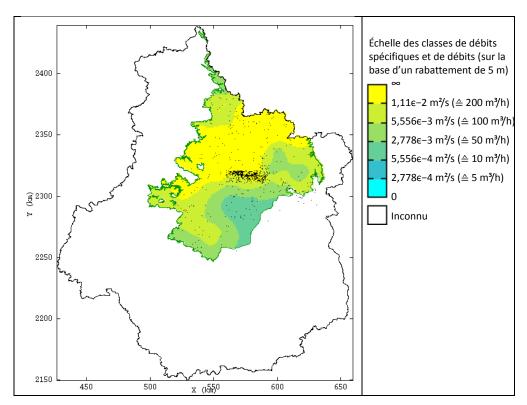


Figure 84 : Calcaires de Beauce - Krigeage lognormal de lnQs selon le variogramme γ₁.

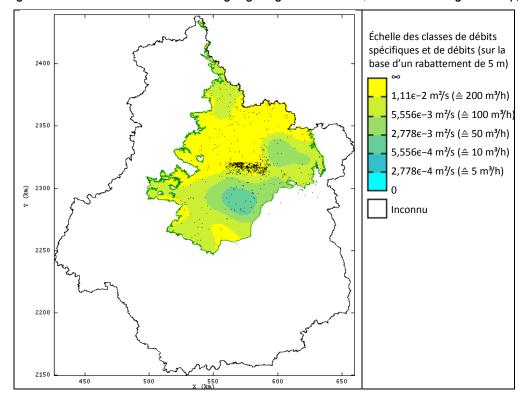


Figure 85 : Calcaires de Beauce - Krigeage lognormal de lnQs selon le variogramme γ₂.

2.4. Comparaison des différentes méthodes

Le tableau suivant donne les résultats de validation des méthodes d'estimation utilisées ci-dessus auxquelles s'ajoute la méthode de la moyenne globale. Les comparaisons sont effectuées au moyen des comparaisons d'agrégation et de prédictibilité.

Il présente pour chaque méthode les statistiques principales, les biais, erreurs absolues, erreurs quadratiques, ainsi que les taux de détection (α), taux de confiance (β), scores (sc), tels que décrits en page 54. Pour ces trois derniers, le nombre dans la première colonne indique le débit potentiel d'exploitation en m³/h (calculé à partir de Q_s , sur la base d'un rabattement de 5 m) posé comme seuil.

méthode	médiane	moyenn	krigeagel	krigeagel	médiane	moyenne	moyenn
	mobile	e mobile	ognorma	ognorma	mobile	mobile	е
	directe	directe	lγ ₁	lγ ₂	logarithmiqu	logarithmiqu	globale
				-	е	е	
moyenne	4,58∈−3	1,84∈−2	6,14∈−2	1,49∈−2	4,58∈−3	5,25∈−3	1,75∈−2
écart-type	3,92∈−3	8,60€−3	6,10∈−2	9,98∈−3	3,92∈−3	3,08∈−3	0
biais	-1,31∈−2	7,38€−4	4,38∈−2	-2,74∈−3	-1,31∈−2	-1,24∈−2	-1,21∈−4
erreurs	1,61∈−2	2,21∈−2	5,70∈−2	1,96∈−2	1,61∈−2	1,60∈−2	2,19∈−2
absolues							
erreurs	4,63∈−2	4,41∈−2	8,37∈−2	4,41∈−2	4,63∈−2	4,58∈−2	4,44∈−2
quadratiq							
ues							
α5	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
α10	98,81%	99,88%	99,88%	100,00%	98,81%	99,16%	100,00%
α50	92,07%	98,38%	98,56%	98,92%	92,07%	91,89%	100,00%
α100	10,74%	94,49%	93,39%	95,87%	10,74%	52,07%	100,00%
α200	4,74%	85,78%	85,34%	76,72%	4,74%	4,74%	100,00%
α500	0,79%	7,94%	85,71%	13,49%	0,79%	0,00%	0,00%
β5	96,17%	96,17%	96,17%	96,17%	96,17%	96,17%	96,17%
β10	91,90%	91,58%	91,58%	91,58%	91,90%	92,13%	91,58%
β50	64,28%	61,76%	62,59%	62,32%	64,28%	64,72%	60,66%
β100	44,32%	41,18%	41,09%	41,78%	44,32%	49,09%	39,67%
β200	42,31%	26,53%	26,98%	32,01%	42,31%	37,93%	25,36%
β500	14,29%	27,78%	15,63%	22,37%	14,29%	0,00%	_
sc5	2 605	2 605	2 605	2 605	2 605	2 605	2 605
sc10	2 411	2 434	2 434	2 437	2 411	2 422	2 437
sc50	1 249	1 300	1 314	1 315	1 249	1 252	1 305
sc100	68	539	531	559	68	371	537
sc200	18	46	58	156	18	15	13
sc500	-3	4	-259	-8	-3	-3	0

Tableau 10 : Calcaires de Beauce - Comparaison des méthodes de cartographie.

Certains comportements s'expliquent simplement :

- Les taux de détection de la moyenne globale sont forcément de 100 % tant que le seuil est inférieur à cette moyenne (car toutes les estimations sont égales à la moyenne, donc toujours "vrais positifs" ou "faux positifs"). Quand le seuil est supérieur à la moyenne des mesures, il n'y a pas de "vrai positif", donc le taux de reconnaissance est de 0 %.
- Les médianes mobiles appliquées aux variables Q_s et InQ_s donnent les mêmes résultats : en effet les calculs sont à peu près identiques, car le point médian ne change pas lors du passage au logarithme.

D'autre part, on note les comparaisons suivantes entre les méthodes :

- Si l'on se fie aux biais et aux erreurs absolues et quadratiques, le krigeage selon y_1 semble la moins bonne méthode, puisque ces valeurs y sont les plus fortes ;
- Pour des seuils faibles (jusqu'à un débit de 50 m³/h), les méthodes géostatistiques donnent des détections et des confiances similaires et à peine meilleures que celles de la moyenne globale;
- Par contre, pour des seuils de 100 ou 200 m³/h, le krigeage donne de meilleurs résultats en termes de détections et confiances. La moyenne mobile appliquée à la variable Q_s est également une bonne méthode. Remarquons le très bon résultat de la méthode basée sur V_2 pour un très fort seuil (200 m³/h).
- Pour un seuil de 500 m³/h (qui n'est pas recherché dans le cadre de notre étude), toutes les méthodes sont mauvaises. Pourtant ce seuil est dépassé pour 126 points de données sur les 915, il n'est donc pas rare. Notons cependant que le calcul de scores est très dépendant des coefficients appliqués.

En résumé, les méthodes variographiques montrent une bonne reconnaissance pour des seuils élevés et une confiance significative. Elles permettent mieux de faire ressortir des événements rares (détectent mieux les points à débits spécifiques forts) que les autres méthodes. Dans la configuration présente, la moyenne mobile appliquée à la variable Q_s sur un voisinage de 10 km donne également de bons résultats, et peut aisément être implémentée et automatisée.

Enfin, quelques points sur la méthodologie doivent être soulignés :

- une part des défaillances des méthodes peut provenir des mesures elles-mêmes. Par exemple, si un point est surévalué par une méthode, cela peut provenir d'un mauvais forage ou d'une mesure erronée; paradoxalement, dans ce cas, les qualités d'une méthode seront dégradées si elle donne un résultat conforme à la réalité physique, c'est-à-dire si elle tente d'approximer le débit spécifique réel, et non le débit spécifique mesuré avec une erreur inévitable. Nous ne voyons pas de moyen de tenir compte de cet effet dans les comparaisons.

À titre illustratif, nous nous proposons de réaliser la même comparaison, cette fois en considérant un voisinage donné par nombre de points : soit donc un voisinage à 8 points.

méthode	médiane	moyenne	krigeagelognormaly ₁	krigeagelognormaly ₂	médiane	moyenne	moyenne
	mobile	mobile			mobile	, mobile	globale
	directe	directe			logarithmique	logarithmique	
moyenne	7,47∈−3	1,77∈−2	1,39∈−2	1,49∈−2	7,47∈−3	6,25∈−3	1,76∈−2
écart-type	9,58€−3	1,86∈−2	1,00∈−2	9,98∈−3	9,58∈−3	5,77∈−3	_
biais	−1,02∈−2	4,64∈−5	−3,78∈−3	-2,74∈-3	−1,02∈−2	−1,14∈−2	2,46∈−10
erreurs	1,69∈−2	2,23∈−2	1,90∈−2	1,96∈−2	1,69∈−2	1,62∈−2	2,19∈−2
absolues							
erreurs	4,60∈−2	4,64∈−2	4,42∈−2	4,41∈−2	4,60∈−2	4,56∈−2	4,44∈−2
quadratiques							
α5	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
α10	98,93%	99,88%	100,00%	100,00%	98,93%	98,93%	100,00%
α50	82,52%	94,77%	98,92%	98,92%	82,52%	82,16%	100,00%
α100	51,79%	79,89%	93,66%	95,87%	51,79%	50,41%	100,00%
α200	26,72%	59,91%	65,52%	76,72%	26,72%	22,84%	100,00%
α500	6,35%	30,95%	14,29%	13,49%	6,35%	2,38%	0,00%
β5	96,17%	96,17%	96,17%	96,17%	96,17%	96,17%	96,17%
β10	92,01%	91,58%	91,58%	91,58%	92,01%	92,01%	91,58%
β50	66,76%	64,54%	62,53%	62,32%	66,76%	67,56%	60,66%
β100	52,37%	44,07%	42,77%	41,78%	52,37%	52,89%	39,67%
β200	40,00%	31,74%	33,19%	32,01%	40,00%	40,15%	25,36%
β500	19,51%	18,84%	25,00%	22,37%	19,51%	30,00%	_
sc5	2 605	2 605	2 605	2 605	2 605	2 605	2 605
sc10	2 415	2 434	2 437	2 437	2 415	2 415	2 437
sc50	1 146	1 289	1 318	1 315	1 146	1 149	1 305
sc100	393	502	565	559	393	386	537
sc200	93	118	150	156	93	80	13
sc500	-9	-51	0	-8	-9	2	0

Tableau 11 : comparaison des méthodes testées, sur un voisinage de 8 points.

L'intérêt d'un voisinage en nombre de points est qu'il permet de cartographier toute la zone demandée; de plus, il se resserre là où l'information est abondante (à la différence d'un voisinage en rayon, qui intégrera toutes les données à proximité). Il est donc pertinent si l'on veut établir une cartographie complète de la zone, à des fins de première estimation.

On voit un léger avantage des méthodes de krigeage par rapport aux autres. Cependant les différences sont minimes, et on pourrait donc imaginer réaliser des cartes préliminaires utilisant une méthode glissante à nombre de points fixés (par exemple moyenne mobile à 8 points), qui serait facile à implémenter en "pressebouton".

3. Sables du Cénomanien

3.1. Méthode statistique globale

Variable	Effectif	Min	Max	Moy	Écart- type	q25	d20	d75
Qs	204	3,333€−5	1,158€−2	1,583€−3	1,645€−3	5,597€−4	1,037ε−3	2,088€−3
InQs	204	-10,31	-4,459	-6,882	0,9793	-7,488	-6,871	-6,172

Tableau 12 : Sables du Cenomanien : statistiques élémentaires.

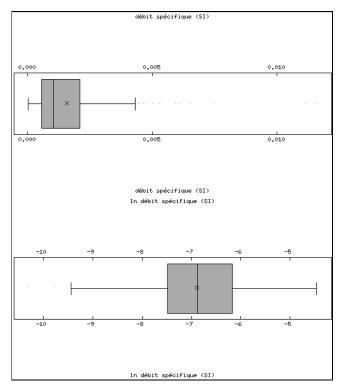
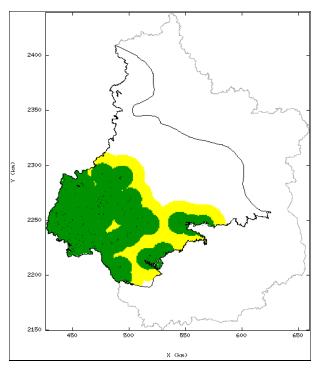


Figure 86 : Sables du Cénomanien : boîtes à moustaches de $\ln Q_s$ et Q_s sur toute la zone.



3.2. Méthode statistique mobile

Figure 87 : Sables du Cénomanien : surfaces à moins de 10 km (vert), ou à 10~20 km (jaune) des points de données.

Une première difficulté surgit de la forte hétérogénéité des points de mesures disponibles. Environ la moitié de l'aquifère n'a aucune donnée dans un voisinage de 10 km (figure 87). De plus, l'anisotropie des données –toutes dans la partie sud-ouest ne permet pas d'évaluer la stationnarité de Q_s sur la zone sans données ; rien ne certifie que les conclusions que nous pourrions tirer de la partie sud-ouest (par exemple, un variogramme) restent valables au nord-est 16 .

De cela, une cartographie en méthode mobile nous semble compromise. Nous allons néanmoins la réaliser. Les cartes suivantes sont calculées sur 182 points de données (son exclus : les doublons, deux points aberrants).

¹⁶ Ceci est d'autant plus vrai que cette absence de données correspond dans les faits à un approfondissement de l'aquifère en direction du nord-est, donc éventuellement à un changement de ses paramètres

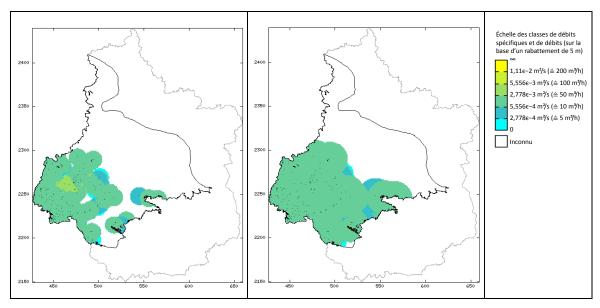


Figure 88 : Sables du Cénomanien : moyenne mobile de Q_s sur voisinages de 10 km (gauche) et de 20 km (droite).

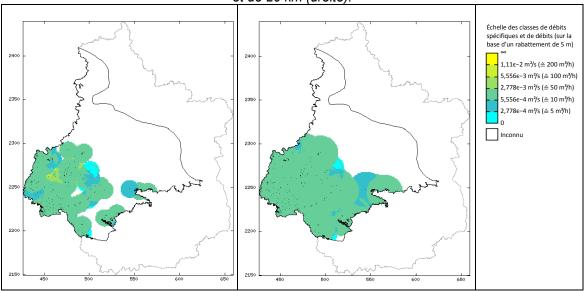


Figure 89 : Sables du Cénomanien : médiane mobile de Q_s sur voisinages de 10 km (gauche) et de 20 km (droite).

La figure 88 (moyenne mobile) et la figure 89 (médiane mobile) montrent que les estimations par méthode mobiles sont peu adaptées au problème présent. On note de très nombreux artefacts de voisinages (arcs de cercles). De fait, pour un voisinage de 20 km, les cartes sont presque exclusivement de la couleur de la classe dominante (10~50 m³/h), les marges prenant la couleur de la valeur la plus proche.

En conséquence, les méthodes mobiles sont à rejeter pour cette configuration de données.

3.3. Méthode du krigeage

3.3.1. Variogramme

Un premier variogramme de InQ_s est calculé sur les 204 points (Figure 90). Il apparaît peu structuré. Des essais montrent que cela n'est pas le fait de valeurs aberrantes.

Le variogramme montre une très forte variation aux faibles distances. Cela suggère la présence de points très proches ayant des débits spécifiques très différents. Effectivement, des doublons sont présents parmi les données.

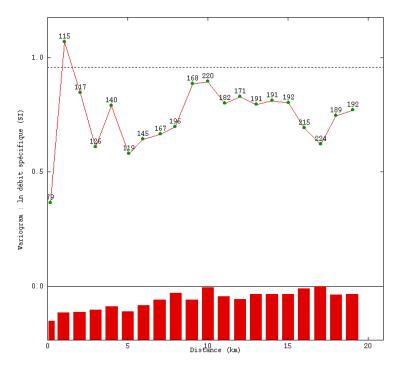


Figure 90 : Sables du Cénomanien - Variogramme estimé sur les 204 données.

Un deuxième variogramme est calculé avec 177 points (Figure 92): sont exclus (Figure 91) des doublons et des points particuliers (à l'aide du nuage variographique).

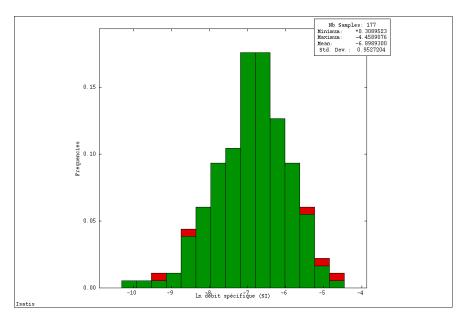


Figure 91 : Sables du Cénomanien - histogramme de InQs, en rouge les points négligés.

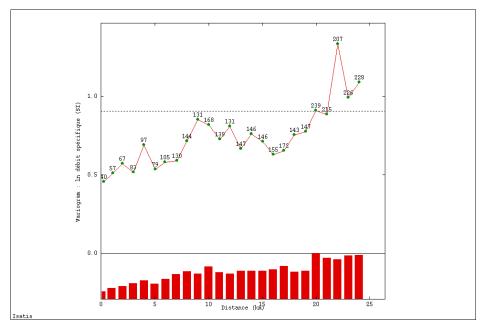


Figure 92 : Sables du Cénomanien - Variogramme estimé sur 177 points.

Au vu des très faibles structurations de ces variogrammes estimés, il ne nous paraît pas faisable de proposer des cartes d'estimations fiables basées sur les méthodes géostatistiques (mobiles ou variographiques). Cependant, nous allons illustrer la méthode sur quelques approches.

3.3.2. Modèle de variogramme y_1

Le modèle proposé pour le variogrammey₁ (Figure 93) est décrit par la relation :

 γ_1 = Pépite(0,36) + Sphérique(13 km; 0,44)

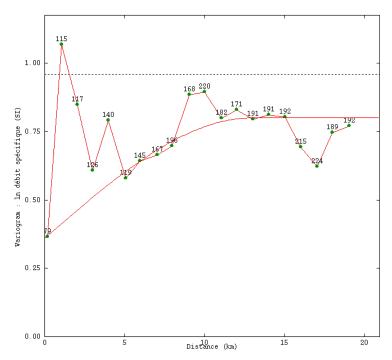


Figure 93 : Sables du Cénomanien - Variogramme estimé et modèle de variogramme γ₁.

3.3.3. Modèle de variogramme y₂

On propose le modèle suivant (Figure 94) :

 y_2 = Pépite(0,46) + Sphérique(0,22 ;21 km) + Gaussienne(0,24 ;21 km)

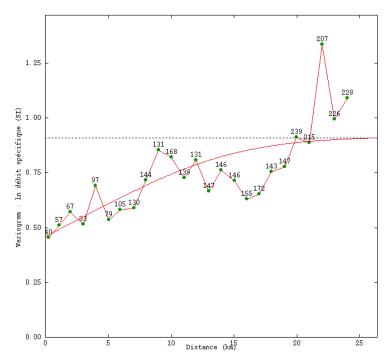


Figure 94 : Sables du Cénomanien - Variogramme estimé et modèle de variogramme γ₂.

3.3.4. Discussion au sujet des variogrammes

Les deux modèles de variogrammes présentés suivent les variogrammes estimés de manière très approximative. De plus, les effets pépites sont très forts, respectivement de 45 % et 50 % du palier total. Enfin, les deux modèles se distinguent par leurs portées, de 13 km et 21 km respectivement.

Il ressort de ces constats que les estimations ne pourront être menées qu'avec une grande incertitude et sur de petites zones (à moins d'une dizaine de kilomètres des points de mesure).

Cela conforte notre conclusion de ne pas établir de carte à l'aide des méthodes statistiques mobiles, car ces cartes auraient été trop fausses.

Néanmoins, nous réalisons une courte étude sur le krigeage et la comparaison entre les méthodes ci-dessous.

3.3.5. Krigeage

La figure 95 présente le résultat du krigeage lognormal selon γ_1 en voisinage unique. La carte du krigeage lognormal de lnQ_s selon γ_2 est présentée en figure 96.

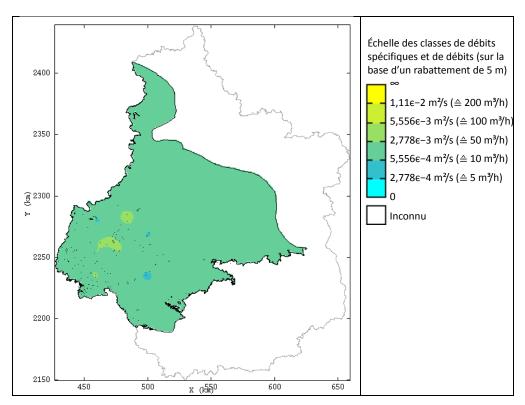


Figure 95 : Sables du Cénomanien - Krigeage lognormal selon le modèle de variogramme γ₁.

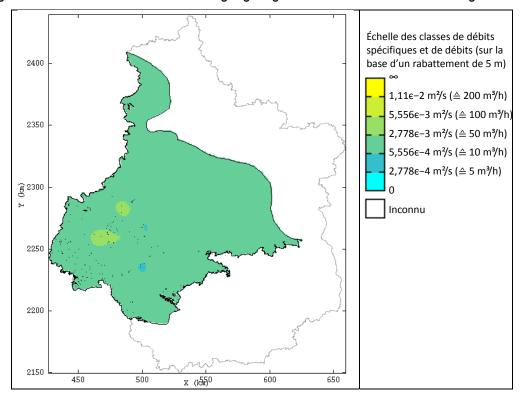


Figure 96 : Sables du Cénomanien - Krigeage lognormal selon le modèle de variogramme γ_2 .

Ces cartes montrent qu'outre les problèmes de position des points de mesures déjà exposés, les variations de Q_s sont très faibles et sans régionalisation bien apparente.

Même si ces cartes semblent plus complètes que celles par méthodes mobiles (Figure 88 et Figure 89), elles sont à prendre avec précaution, car le krigeage peut introduire une précision "artificielle", qui n'est pas forcément gage de justesse.

3.3.6. Évaluation des krigeages par validation croisée

Validation croisée du krigeage selon y₁

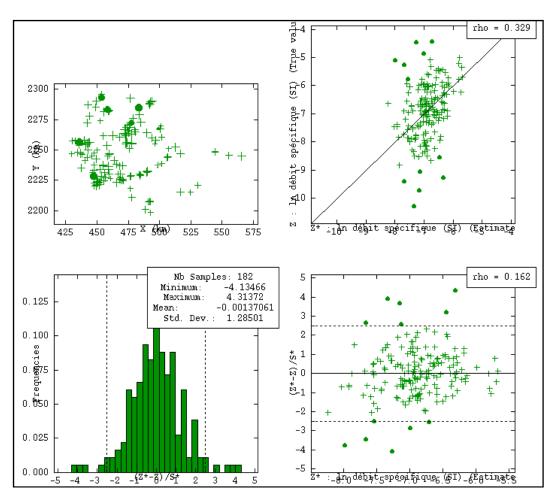


Figure 97 : Validation croisée du krigeage selon γ₂.

Validation croisée du krigeage selon y₂

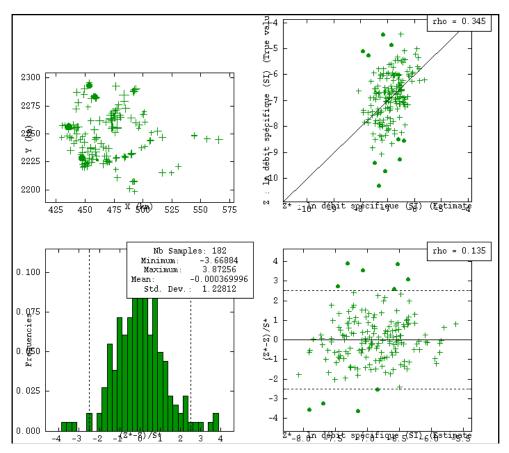


Figure 98 : Validation croisée variogramme γ₂.

Les deux validations croisées montrent un fort lissage des valeurs (graphique haut droite, aplatissement des points selon un grand axe vertical).

De plus, on note un assez grand nombre de points aberrants (selon la définition de la validation croisée : erreur commise dépassant en valeur absolue 2,5 fois l'erreur-type). Ces points correspondent autant à des sur-estimations qu'à des sous-estimations (signe de l'erreur relative), mais ils se trouvent dans l'ouest de la zone de travail, ce qui signe une plus grande variabilité de lnQ_s dans cette zone.

En outre, les erreurs relatives ne laissent pas apparaître de biais d'évaluation majeur, tout au plus une meilleure estimation des points à lnQ_s élevés (graphique bas droite, dispersion autour de l'horizontale centrale à peu près constante).

La moyenne des erreurs relatives est plus proche de zéro, et leur écart-type est plus faible (voir graphiques en bas à gauche), dans le second cas (variogramme γ_2). Ce variogramme est donc préférable, même si la différence ne semble pas très grande.

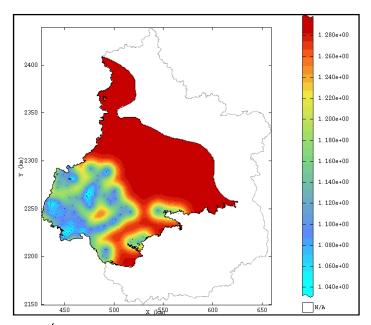


Figure 99 : Écart-type relatif du krigeage lognormal (variogramme γ2).

L'écart-type relatif de krigeage est également calculé lors des krigeages (Figure 99 pour le krigeage selon γ_2). Il propose une mesure de la précision du krigeage, sous l'hypothèse, forte, de stationnarité. Cette hypothèse n'est pas validée ici, néanmoins l'écart-type relatif permet de visualiser les zones où le krigeage est plus précis ou plus lâche. On note en particulier toute la moitié nord-ouest, où cet écart-type est maximal : au-delà d'une distance égale à la portée, mesurée depuis les points de mesure, le krigeage donne une valeur égale à la moyenne des points, et donc il estime l'écart-type sur son estimation comme égal à l'écart-type des mesures.

Cette carte permet donc un jugement préliminaire sur la qualité des estimations par krigeage, sans pourtant être forcément fiable, à cause de la non-stationnarité de Q_s . On retrouve à peu près les contours des zones à 10 km des points de données (Figure 87), ce qui est en accord avec les portées estimées.

3.4. Comparaison des différentes méthodes

Bien que la cartographie de l'aquifère par les méthodes géostatistiques ne puisse être menée à bien à cause de la répartition et du comportement des données, nous pouvons comparer par une validation croisée les différentes méthodes, afin de voir leur comportement dans un cas nettement défavorable.

Le tableau suivant donne les résultats de validation des méthodes d'estimation envisagées.

Il présente pour chaque méthode les statistiques principales, les biais, erreurs absolues, erreurs quadratiques, ainsi que les taux de détection (α) , taux de confiance (β) , scores (sc), tels que décrits en page 55. Pour ces trois derniers, le nombre dans la

première colonne indique le débit en m³/h (sur la base d'un rabattement de 5 m) posé comme seuil (5, 10, 50).

Pour les 163 points sélectionnés, le débit spécifique a une moyenne de 1,53€−3.

méthode	moyenne glissante	médiane glissante	lognormal γ1	lognormal γ2	lognormal γ1	lognormal γ2	moyenne globale
voisinage	10 km	10 km	10 km	10 km	unique	unique	_
moyenne	1,55€−3	1,28ε−3	1,50ε−3	1,50ε−3	1,48ε−3	1,50ε−3	1,53€−3
écart-	6,40€−4	6,33€−4	8,83€−4	8,08€−4	7,64ε−4	7,58ε−4	3,48€−18
type							
biais	2,14€−5	-2,46€-4	-3,03€−5	-3,36€−5	-4,65€−5	-2,62€−5	0
erreurs	9,88€−4	8,97€−4	9,35ε−4	9,29ε−4	9,21ε−4	9,34ε−4	1,05€−3
absolues							
α5	100%	98,63%	99,32%	100%	100%	100%	100%
α10	98,39%	95,16%	93,55%	94,35%	95,97%	96,77%	100%
α50	4,55%	4,55%	22,73%	22,73%	22,73%	22,73%	0%
β5	89,57%	89,44%	89,51%	89,57%	89,57%	89,57%	89,57%
β10	76,25%	77,12%	77,85%	76,97%	76,28%	75,95%	76,07%
β50	16,67%	100%	45,45%	41,67%	62,50%	50%	_
sc5	421	415	418	421	421	421	421
sc10	328	319	315	316	320	322	333
sc50	-2	3	9	8	12	10	0

Tableau 13 : Sables du Cénomanien - Comparaison des méthodes de cartographie.

Le biais et l'erreur absolue ne permettent pas de conclure, ces valeurs étant proches pour toutes les méthodes.

Le taux de détection passe pour la moyenne globale de 100 % à 0 %; c'est logique, puisque la moyenne globale pose toutes les estimations égales entre elles; soit le seuil est choisi inférieur (100 % des estimations sont alors supposées favorables), soit supérieur (0 %).

Les méthodes variographiques fournissent d'aussi bonnes détections que les méthodes glissantes. Pour les seuils de débit forts ($\alpha50$), leurs résultats sont même meilleurs. Cependant la détection est globalement mauvaise pour les débits forts, puisque $\alpha50$ ne dépasse guère un cinquième.

Le taux de confiance est homogène entre les méthodes pour les débits faibles et moyens (β 5, β 10). Par contre, la comparaison des confiances pour des débits forts (β 50) montre la faiblesse de la méthode de la moyenne glissante par rapport aux autres méthodes. La bonne confiance de la méthode de la médiane glissante (100 %) est à mettre en rapport avec son faible taux de détection α 50 : elle détecte un seul point au-dessus des 50 m³s⁻¹, sur les 22 existants. Les méthodes variographiques en détectent 5, mais en contrepartie voient 3 à 7 points «favorables», selon la méthode, alors qu'ils sont sous le seuil.

Dans ce cas-ci, les méthodes variographiques permettent donc globalement de mieux identifier les zones intéressantes pour la prospection géothermique.

4. Conclusions

Les avantages et inconvénients des différentes méthodes testées ci-dessus sont synthétisés dans le Tableau 14.

Les méthodes statistiques mobiles et géostatistiques offrent des résultats comparables et permettent de tenir compte de la variabilité spatiale des débits spécifiques Q_s .

Elles doivent être utilisées systématiquementquand les données le permettent.

Méthode de cartographie	Avantages	Inconvénients
Dire d'expert	Mise en œuvre très simple Approche qualitative intéressante en l'absence de données mesurées	Approche relativement subjective, non reproductible
Statistique globale	Mise en œuvre très simple Résultat simple à représenter	Ignore totalement les fluctuations locales ou régionales du paramètre étudié Dans le cas d'une répartition hétérogène des données, peut conduire à des résultats fortement biaisés
Statistique mobile	Prise en compte des points de mesure, d'autant plus forte qu'ils sont proches du point d'estimation Mise en œuvre relativement simple	Résultat dépendant du choix du voisinage de calcul Lissage du résultat si la distance entre les points de mesures est très grande devant la portée (estimée d'après le variogramme)
Krigeage	Prise en compte des points de mesures selon leur distances entre eux, et donc selon leur répartition spatiale Prise en compte des phénomènes régionalisés quand ils existent (structure de données dont la portée est supérieure à la maille d'échantillonnage) Résultats meilleurs que les statistiques mobile lorsque la densité d'échantillonnage est suffisante vis-à-vis de la variabilité spatiale du phénomène	Résultat dépendant du choix du voisinage de calcul Besoin de calculer et modéliser le variogramme, donc de compétences particulières en géostatistiques
Simulations géostatistiques	Idem que krigeage Fournit en tout point, non seulement une estimation (moyenne des possibilités), mais une distribution des valeurs possibles, donc une représentation de leur dispersion autour de l'estimation. Peut être utilisé pour établir une carte de probabilité de dépasser un débit seuil, ou une carte de débit dépassé avec une probabilité seuil	Idem que krigeage S'applique à une variable gaussienne technique nettement plus lourde à mettre en œuvre que le krigeage(temps de mise en œuvre)

Tableau 14 : Avantages et inconvénients des méthodes de cartographie de la productivité d'aquifère.

