



## THÈSE DE DOCTEUR - INGÉNIEUR

Présentée  
à l'Université Pierre et Marie Curie  
PARIS - 6

Par : Jean-Pierre DELHOMME  
Pour obtenir le titre de DOCTEUR - INGÉNIEUR

Sujet de la thèse :

### APPLICATIONS DE LA THÉORIE DES VARIABLES RÉGIONALISÉES DANS LES SCIENCES DE L'EAU

(Variabilité spatiale des grandeurs hydroclimatiques et hydrogéologiques  
& précision de leur connaissance)

Soutenue le 15 Novembre 1976

devant le jury composé de :

M. JAUZEIN  
M. CASTANY  
M. LETOLLE  
M. de MARSILY  
M. MATHERON

Président

Examineurs

## REMERCIEMENTS

Je tiens, tout d'abord, à exprimer ma reconnaissance à mes camarades des Centres d'Informatique Géologique, et de Morphologie Mathématique. Il serait très long de les mentionner tous, et seuls les noms de certains apparaîtront çà ou là parmi les références bibliographiques. Mais que tous sachent qu'au cours des nombreuses discussions amicales que nous avons eues, ils m'ont beaucoup apporté.

Je sais infiniment gré à Monsieur MATHERON, Directeur du Centre de Morphologie Mathématique, d'avoir bien voulu s'intéresser à mon travail et d'avoir accepté de faire partie de mon jury. Je le remercie aussi de la possibilité qu'il m'a accordée de travailler deux années durant au sein de son Centre, où j'ai pu me familiariser avec la théorie des variables régionalisées et ses applications.

Je dois beaucoup à G. de MARSILY, Directeur du Centre d'Informatique Géologique, et rapporteur de ma thèse, qui a toujours encouragé son équipe à développer des techniques nouvelles dans le domaine des sciences de l'eau. Je lui suis également reconnaissant d'avoir relu attentivement mon manuscrit et d'avoir contribué, par ses remarques, à améliorer la clarté de l'exposé.

J'adresse mes remerciements les plus sincères à Monsieur JAUZEIN qui a bien voulu assurer la présidence de mon jury, à Monsieur LETOLLE qui a accepté d'être mon directeur de thèse à l'Université de Pierre et Marie Curie, et à Monsieur CASTANY qui m'a fait l'honneur de faire partie de mon jury.

Je dois témoigner aussi toute ma gratitude à P. DELFINER qui, au cours des années de travail commun, m'a apporté une aide importante par ses réflexions riches d'enseignement et son soutien amical.

Qu'il me soit enfin permis de remercier Madame COLLARD qui a eu la tâche ingrate de mener à bien la réalisation matérielle de ce travail.

TABLE DES MATIERES
--------------------

	page
INTRODUCTION .....	1
I - METHODES D'ESTIMATION EN USAGE EN GEOLOGIE ET DANS LES SCIENCES DE L'EAU .....	5
1 - Méthodes analytiques d'interpolation .....	5
a) Transposition du procédé manuel .....	5
b) Interpolation polynômiale .....	7
c) Fonctions splines .....	8
d) Procédés divers .....	9
2 - Méthode des moindres carrés .....	12
a) Fondements théoriques .....	12
b) Notion de "trend" .....	15
3 - Méthodes élémentaires de pondération .....	17
a) Pondération en fonction de la distance .....	17
b) Méthode des polygones d'influence .....	20
4 - Méthodes probabilistes .....	23
a) Caractérisation de la structure .....	23
α) Normales, anomalies et coefficients de corrélation .....	23
β) Fonction de corrélation .....	25
γ) Interprétation climatologique .....	27
b) Prise en compte de la structure dans l'estimation .....	28
α) Structure et erreur d'estimation .....	28
β) Interpolation optimale .....	31
γ) Moyenne pondérée optimale .....	33
δ) Applications .....	34
II - THEORIE DES VARIABLES REGIONALISEES. INFERENCE STATISTIQUE ET ESTIMATION OPTIMALE .....	39
1 - Bases conceptuelles .....	39
a) Notion de variable régionalisée .....	39
b) But de la théorie des variables régionalisées .....	41
c) Interprétation probabiliste .....	41
d) Problème de l'inférence statistique .....	42

	page
2 - Inférence statistique dans le cadre de l'hypothèse intrinsèque .....	43
a) Stationnarité d'ordre 2 .....	43
b) Hypothèse intrinsèque .....	44
c) Variogramme .....	45
α) Définition .....	45
β) Estimation du variogramme .....	45
γ) Comportement du graphe .....	47
. à l'origine .....	47
. à l'infini .....	50
δ) Ajustement à un modèle .....	53
3 - Krigeage .....	56
a) Exemple introductif .....	56
b) Cas sans dérive: krigeage simple .....	57
α) Position du problème .....	57
β) Estimation zonale .....	57
γ) Estimation ponctuelle .....	61
c) Cas avec dérive: krigeage universel .....	62
α) Modèle interprétatif .....	62
β) Calculs .....	63
d) Propriétés du krigeage .....	65
4 - Retour à l'inférence statistique: LES F.A.I.- $k$ .....	66
a) Variogramme brut et variogramme sous-jacent .....	66
b) Variogramme des résidus .....	68
c) Hypothèse intrinsèque généralisée .....	69
α) Généralisation de la notion d'accroissement .	70
β) Stationnarité des accroissements d'ordre $k$ ..	71
γ) Covariance généralisée: modèle polynômial et ajustement automatique .....	72
III - APPLICATION A L'ETUDE DES VARIATIONS SPATIALES DES GRANDEURS HYDROLOGIQUES .....	77
1 - Estimations ponctuelles et zonales .....	77
a) Vérification de la validité du modèle structural adopté .....	77
b) Emploi du krigeage en cartographie .....	79
c) Estimation optimale des valeurs moyennes .....	84
α) Cas d'un maillage .....	84
β) Cas d'un domaine quelconque .....	87
d) Adaptation de la méthode à des problèmes particuliers	93
α) Hétérogénéité de la structure .....	93
β) Transformation logarithmique .....	93
γ) Données incertaines .....	95
δ) Présence de failles .....	99

## LISTE DES FIGURES

- FIG. 1 - Méthode élémentaire de cartographie piézométrique.
- FIG. 2 - Système global et système local de coordonnées.
- FIG. 3 - Ajustement local par moindres carrés d'un polynôme du 1er degré.
- FIG. 4 - Ajustement local par moindres carrés d'un polynôme du 2ème degré.
- FIG. 5 - Pondération par l'inverse de la distance.
- FIG. 6 - Pondération par l'inverse du carré de la distance.
- FIG. 7 - Inconvénient de la pondération en fonction de la distance au point à estimer.
- FIG. 8 - Tracé des polygones et pondération de Thiessen.
- FIG. 9 - Courbes d'isocorrélation entre une station-clé et les autres stations d'un réseau.
- FIG. 10 - Report en fonction de la distance des coefficients de corrélation.
- FIG. 11 - Tracé de la fonction de corrélation spatiale pour 5 bassins versants des Etats-Unis.
- FIG. 12 - Fonction de corrélation de la pluviométrie mensuelle (Valdaï, URSS).
- FIG. 13 - Echos radar observés durant un orage (Californie, USA).
- FIG. 14 - Evolution de la profondeur de la nappe au puits n°3 de Korhogo.
- FIG. 15 - Valeurs de la transmissivité mesurées dans la nappe alluviale de la Durance.
- FIG. 16 - Comportement du variogramme à l'origine.
- FIG. 17 - Exemples de variogrammes illustrant les 4 types de comportement à l'origine.
- FIG. 18 - Variogramme non borné et variogramme à palier.
- FIG. 19 - Exemples de variogrammes présentant un palier.
- FIG. 20 - Comportements particuliers du variogramme.
- FIG. 21 - Principaux modèles de variogramme utilisés et exemple d'ajustement.
- FIG. 22 - Modélisation aux distances inférieures au pas d'échantillonnage.
- FIG. 23 - Nappe de la Crau: variogramme brut.
- FIG. 24 - Variogramme expérimental en présence d'une dérive et variogramme sous-jacent.

- FIG. 25 - Variogramme réel et variogramme des résidus.
- FIG. 26 - Réalisations de F.A.I. généralisées d'ordre 0, 1 et 2.
- FIG. 27 - Basse-Normandie (logarithmes des transmissivités): essai de krigeage.
- FIG. 28 - Bassin versant de Kadjemeur: tracé manuel des isohyètes.
- FIG. 29 - Bassin versant de Kadjemeur: carte krigée (écarts-types d'estimation en tireté).
- FIG. 30 - Bassin versant de Kadjemeur: carte obtenue par krigeage en négligeant l'effet de pépite.
- FIG. 31 - Influence de la structure sur les poids de krigeage.
- FIG. 32 - Basse-Normandie: piézométrie moyenne par maille.
- FIG. 33 - Basse-Normandie: écarts-types d'estimation des moyennes par maille.
- FIG. 34 - Bassin versant de Kadjemeur: polygones de Thiessen et pondération correspondante.
- FIG. 35 - Bassin versant de Kadjemeur: pondération optimale obtenue par krigeage.
- FIG. 36 - Bassin versant de Kadjemeur: comparaison des différentes méthodes d'estimation globale.
- FIG. 37 - Pondération de krigeage dans le cas d'une dérive linéaire.
- FIG. 38 - Pondération de krigeage dans le cas d'une dérive quadratique.
- FIG. 39 - Sables éocènes.Nord-Aquitaine: variogramme construit à partir des données de transmissivité.
- FIG. 40 - Sables éocènes.Nord-Aquitaine: histogramme des données.
- FIG. 41 - Sables éocènes.Nord-Aquitaine: variogramme construit à partir des logarithmes.
- FIG. 42 - Données sûres et données incertaines.
- FIG. 43 - Influence de l'incertitude sur le poids.
- FIG. 44 - Basse-Normandie: prise en compte des débits spécifiques dans l'estimation de la transmissivité de la nappe bathonienne.
- FIG. 45 - Failles géologiques et effet d'écran formé par un cap.
- FIG. 46 - Basse-Normandie: prise en compte de failles dans l'estimation de la cote du mur du Bathonien.
- FIG. 47 - El Aziziyah: test de la pertinence de deux failles possibles.
- FIG. 48 - El Aziziyah: problème de la cartographie automatique au voisinage d'une faille.
- FIG. 49 - Bassin versant de Kadjemeur: implantation du réseau pluviométrique.
- FIG. 50 - Méthode du point fictif: courbes d'égal gain en précision.
- FIG. 51 - Réseau piézométrique de l'Huveaune: écarts-types de krigeage.

- FIG. 52 - Ecart-types de krigeage après addition d'un piézomètre supplémentaire.
- FIG. 53 - Influence de l'incertitude affectant la donnée supplémentaire sur le gain en précision.
- FIG. 54 - Méthode des bandes tournantes.
- FIG. 55 - Bassin versant de Kadjemeur: trois simulations conditionnelles.
- FIG. 56 - Nappe de l'Huveaune: carte krigée.
- FIG. 57 - Nappe de l'Huveaune: écart-types d'estimation.
- FIG. 58 - Comparaison entre la carte krigée et deux simulations conditionnelles.
- FIG. 59 - Comparaison sur le profil AB du krigeage et de deux simulations conditionnelles.
- FIG. 60 - Nappe de Kairovan: écarts de calage.
- FIG. 61 - Nappe de Kairovan: écarts réduits.

ORIGINE DES DONNEES UTILISEES

- ASSENS, G., de MARSILY, G. de, POTIE, L. (1974):  
Basse vallée de l'Huveaune: étude des écoulements souterrains  
sur modèle mathématique.  
Ecole des Mines - Société des Eaux de Marseille.
- ASTIE, H., BOURGEOIS, M, TRUPIN, G. (1967):  
Caractéristiques hydrauliques de la nappe des sables éocènes  
de la Gironde: essai de représentation cartographique des  
transmissivités.  
BRGM, D.S.G.R. 67 A 34.
- BESBES, M. (1972):  
Ressources en eau de la plaine de Kairouan.  
D.R.E, Tunis.
- BESBES, M. (1975):  
Etude hydrogéologique de la plaine de Kairouan sur modèles  
mathématiques.  
Ecole des Mines, LFM, R75/16.
- BONNET, M., CLOUET D'ORVAL, I. (1969):  
Méthodologie d'emploi des simulateurs d'écoulement pour  
l'étude des nappes souterraines: application à la Crau.  
Géohydraulique - DGRST Action Concertée Eau 1967-1968.
- BOSSY, G. (1965):  
Documents sur la nappe alluviale de la Basse Durance et de  
la Crau (1953-1962).  
Mémoire du CERH, Montpellier.
- DASSIBAT, C., PASCAUD, P. (1975):  
Modèle mathématique des nappes du Jurassique de la plaine  
de Caen: acquisition de données hydrogéologiques.  
BRGM, 75 SGN 340 PNO.
- DEGALLIER, R., (1970):  
Communication personnelle à l'auteur.
- DEGALLIER, R., (1975):  
Interprétation des variations naturelles du niveau des nappes  
souterraines: applications aux données provenant du bassin  
versant de Korhogo, Côte d'Ivoire.  
Thèse de Doctorat es Sciences, Montpellier.

G.E.F.L.I. (1962):

Structural map of the top of Al Aziziyah formation (Libya).  
Soil and Water Resources Survey for Hydroagricultural  
Development (Western zone).

LA QUERIERE, Ph. de, ROUX, J.C. (1969):

Le bassin versant représentatif de l'Hallue.  
BRGM - PND 69/6.

ROCHE, M.A. (1968):

Ecoulement de surface, alimentation de nappe et transport  
solide des ouadis Fera, Kadjemeur et Sofoya  
ORSTOM, Fort-Lamy.

## INTRODUCTION

Sciences appliquées, les sciences de l'eau nécessitent l'acquisition d'un grand nombre de données de tous ordres: cote du niveau piézométrique dans un puits, transmissivité déterminée par un essai de pompage, cote du mur imperméable de la nappé quand celui-ci a été atteint par un forage, hauteur de précipitation mesurée dans un pluviomètre, etc.... Mais cette acquisition de données ne saurait constituer en soi un but suffisant. Pour pouvoir en tirer des enseignements, ces données numériques brutes doivent toujours subir un traitement permettant de dégager les grands traits structuraux des grandeurs étudiées.

La présentation des données sous une forme cartographique permet ainsi de donner rapidement une idée de la variation d'un paramètre dans l'espace. Cependant, comme le note le Guide des Pratiques Hydro-météorologiques de l'O.M.M., "si l'on ne dispose que d'une information peu abondante, la carte présentera, dans une certaine mesure, une interprétation subjective des données; on risque d'ignorer partiellement ou totalement les différents degrés de subjectivité qui se rapportent aux diverses parties de la carte, ce qui produira des résultats dont la précision apparente dépassera de beaucoup la précision réelle" [1]. Cette difficulté de quantifier la précision se retrouve quand on doit estimer, à partir des données, les valeurs moyennes prises par un paramètre sur chacune des mailles d'un modèle mathématique discrétisé, ou encore quand on doit estimer la lame d'eau moyenne sur un bassin versant.

Force est de constater que la plupart des méthodes utilisées couramment [2], [3] pour répondre aux besoins de la cartographie et de l'estimation de moyennes spatiales ne permettent pas d'accéder à la moindre indication sur l'incertitude du résultat; seules le permettent les méthodes prenant en compte la variabilité des phénomènes étudiés.

La théorie des variables régionalisées, due à G. MATHERON, empruntant son arsenal conceptuel et ses techniques mathématiques à la théorie des fonctions aléatoires, permet d'apporter à ces problèmes une solution reposant sur une base théorique solide. La méthode d'estimation proposée a reçu le nom de "krigeage". Ce mot, dont la consonance peut surprendre, a été forgé à partir du nom d'un géologue spécialiste des problèmes d'estimation minière, D.G. KRIGE. La méthode a vu son champ d'application, d'abord limité au seul domaine minier, s'étendre considérablement ces dernières années pour couvrir des domaines aussi divers que la bathymétrie et la foresterie. Elle s'est révélée particulièrement adaptée aux problèmes rencontrés dans le domaine de l'eau.

Après une revue des différentes méthodes d'estimation classiquement employées en hydrogéologie et en hydrométéorologie en commençant par la simple cartographie manuelle, on rappellera les traits principaux de la théorie des variables régionalisées, pour aborder enfin la présentation de quelques unes de ses applications possibles concernant les sciences de l'eau.

REFERENCES BIBLIOGRAPHIQUES DE L'INTRODUCTION

- [1] O.M.M. (1970):  
Guide des pratiques hydrométéorologiques.  
(IV - Rassemblement, traitement et publication des données).  
Publ. O.M.M., n° 168, TP 82.
- [2] UNESCO (1970):  
Les bassins représentatifs et expérimentaux: guide  
international des pratiques en matière de recherche.  
Etudes et Rapports d'Hydrologie, n°4.
- [3] UNESCO (1973):  
Groundwater studies: an international guide for research  
and practice.

CHAPITRE I  
METHODES D'ESTIMATION EN USAGE  
EN GEOLOGIE  
ET DANS LES SCIENCES DE L'EAU

1 - METHODES ANALYTIQUES D'INTERPOLATION

a) Transposition du procédé manuel

En géologie comme dans les sciences de l'eau, la cartographie manuelle est une technique très répandue. Quand on trace une carte à la main, on interpole en général visuellement entre les points expérimentaux: on fait en sorte, par exemple, que la courbe isovale 40 passe entre les données 39.7 et 40.1, un peu plus près du point de cote 40.1. Une telle interpolation visuelle est en gros linéaire.

Voulant calquer ce procédé pour bâtir une méthode de tracé automatique de cartes, certains auteurs ont pensé à déterminer d'abord des points de cotes entières, en interpolant linéairement le long des segments joignant les points expérimentaux deux à deux, et à relier ensuite les points de même cote en opérant un certain lissage des courbes obtenues [8].

Mais si l'on essaie ainsi de tenir compte de tous les couples de points expérimentaux, la méthode devient extrêmement lourde, avec des risques de conflit au moment du chaînage des courbes.

D'ailleurs, la pratique qui s'est introduite en hydrogéologie pour faciliter le tracé manuel des cartes piézométriques consiste à ne retenir que trois piézomètres à la fois. On approxime la surface à l'aide de facettes planes triangulaires reliant les points expérimentaux (Fig.1).

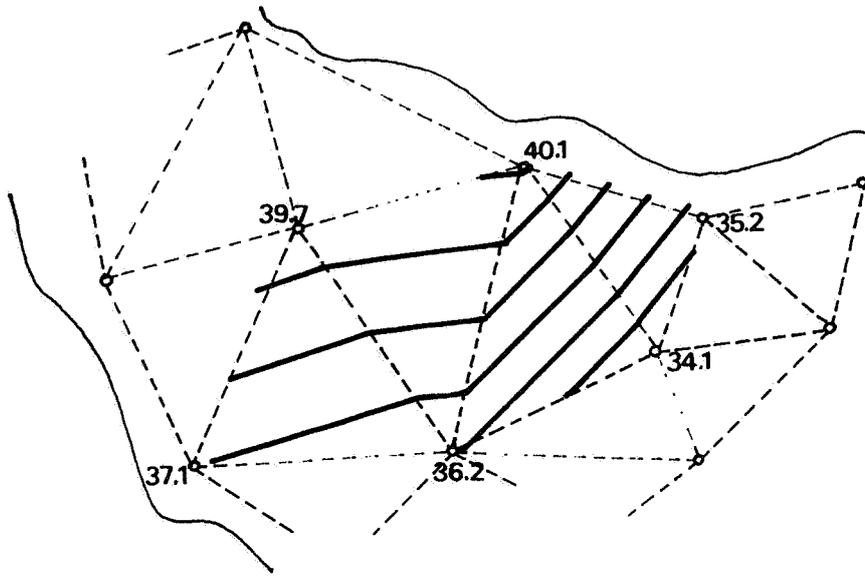


FIG. 1 - METHODE ELEMENTAIRE DE CARTOGRAPHIE PIEZOMETRIQUE

Cette méthode très répandue, encore appelée "contouring par triangulation" [2], est généralement appliquée en prenant soin de bâtir, autant que possible, des triangles à peu près équilatéraux. Un programme de calcul déterminant un tel ensemble de triangles est assez compliqué, et long à exécuter, ce qui enlève beaucoup à la simplicité apparente de la méthode quand il s'agit de l'automatiser.

Ce n'est donc pas vers une simple transposition des méthodes manuelles de cartographie que s'est tourné l'essentiel de la recherche dans ce domaine. La partie "tracé" est généralement simplifiée par le recours à une estimation préalable de la valeur du champ aux noeuds d'une grille régulière. Dans cette première phase d'interpolation proprement dite, l'utilisation des ordinateurs a permis en outre de ne pas se limiter à une simple interpolation linéaire et d'utiliser des méthodes analytiques d'interpolation plus sophistiquées.

b) Interpolation polynômiale

Pour interpoler le champ étudié à partir des valeurs mesurées aux points expérimentaux, les méthodes analytiques usuelles emploient des fonctions mathématiques comme les polynômes, plus rarement des fonctions trigonométriques. On notera le champ  $z(x)$ , les points expérimentaux  $x_i$  (cette notation condensée à une seule lettre recouvrant en fait des points de l'espace à deux dimensions, de coordonnées  $X_i$  et  $Y_i$ ), les valeurs expérimentales  $z_i = z(x_i)$ . On prend comme modèle de  $z(x)$ :

$$z(x) = \sum_{l=1}^k a_l \cdot f^l(x)$$

où les  $f^l$  sont des fonctions de base données (par exemple, les monômes  $1, X, Y, X^2, Y^2, XY$ , etc...), et les  $a_l$  des coefficients calculés en sorte que l'on ait un ajustement exact aux points expérimentaux.

Par commodité d'écriture, on notera les fonctions de base avec un indice en position supérieure,  $f^l$  désignant ainsi la  $l^{\text{ième}}$  fonction de base.

On prendra en compte un nombre de points expérimentaux identique au nombre de fonctions  $f^l$  retenu (en général  $< 10$ ), en procédant, si le nombre total de points expérimentaux est trop élevé, par voisinages glissants.

Les coefficients  $a_l$  sont obtenus en résolvant le système linéaire:

$$\sum_{l=1}^k a_l \cdot f^l(x_i) = z_i \quad i = 1, \dots, k$$

Ces méthodes sont satisfaisantes lorsque le champ étudié possède une assez bonne régularité spatiale; encore que le choix du nombre de fonctions de base  $f^l$  soit assez délicat. Des polynômes de degré élevé conduisent en effet à des cartes à l'aspect très tourmenté, avec des gradients très forts et souvent irréalistes [4]. Mais ces méthodes d'interpolation analytiques peuvent se révéler catastrophiques si le champ étudié est un peu chaotique.

Or, de nombreuses données géologiques, ou hydrogéologiques, ont une variabilité locale relativement importante qui a donc fait renoncer à la technique précédente au profit de celle des moindres carrés, où l'on n'imposera plus à la carte de respecter scrupuleusement les valeurs expérimentales.

### c) Fonctions splines

Un regain d'intérêt pour les méthodes d'interpolation polynômiale est pourtant apparu ces derniers temps par le biais des fonctions splines. Quelques auteurs ont proposé l'emploi de ces fonctions splines pour résoudre les problèmes d'interpolation qui se posent en géologie [12], géophysique [6] et en hydrométéorologie [11].

Le terme "spline" désignait, à l'origine, la languette métallique flexible utilisée par les architectes navals et les ingénieurs aéronautiques pour faciliter le tracé d'une courbe régulière passant par une série de points donnés. Ses propriétés élastiques garantissaient la continuité de la pente et de la courbure aux points imposés, les segments de courbe joignant ces points ayant, en première approximation, une allure polynômiale du 3<sup>ème</sup> degré. Cet ajustement polynômial par morceaux d'une courbe plane a été facilement adaptable au calcul numérique [9].

Par analogie avec ces courbes splines, les surfaces splines sont des surfaces polynômiales bicubiques ajustées par morceaux [3], [5]. Malheureusement, leur usage se restreint pour l'instant, pour des raisons de commodité de calcul, au cas de données disposées sur une grille rectangulaire, ou à la rigueur sur un maillage topologiquement équivalent [6]. Il faut donc attendre que la méthode soit étendue à des données irrégulièrement espacées pour qu'elle puisse être appliquée à des exemples réels. Si pour des phénomènes très continus comme la piézométrie on peut espérer de bons résultats, pour des phénomènes chaotiques on retrouvera les défauts de toute interpolation polynômiale.

d) Procédés divers

Avant d'aborder la méthode des moindres carrés, on peut encore citer, pour mémoire, deux tentatives d'application, au problème de l'interpolation, de procédés utilisés par ailleurs en hydrologie: les techniques des différences finies et des éléments finis, auxquelles on a classiquement recours dans la construction des modèles mathématiques d'écoulement souterrain.

HUTCHINSON et WALLEY [7] ont proposé l'emploi des fonctions d'interpolation utilisées en éléments finis pour la représentation de la pluviométrie. Le domaine ayant été divisé en quadrilatères convexes avec un pluviomètre à chaque sommet, ils adoptent donc comme estimation au point  $x$ , de coordonnées  $\xi$  et  $\eta$  dans le système local (cf. Fig.2), la valeur:

$$z^* = \frac{1}{4} [(1-\xi)(1-\eta)z_1 + (1+\xi)(1-\eta)z_2 + (1+\xi)(1+\eta)z_3 + (1-\xi)(1+\eta)z_4]$$

ce qui ne correspond en fait qu'à un modèle analytique particulier.

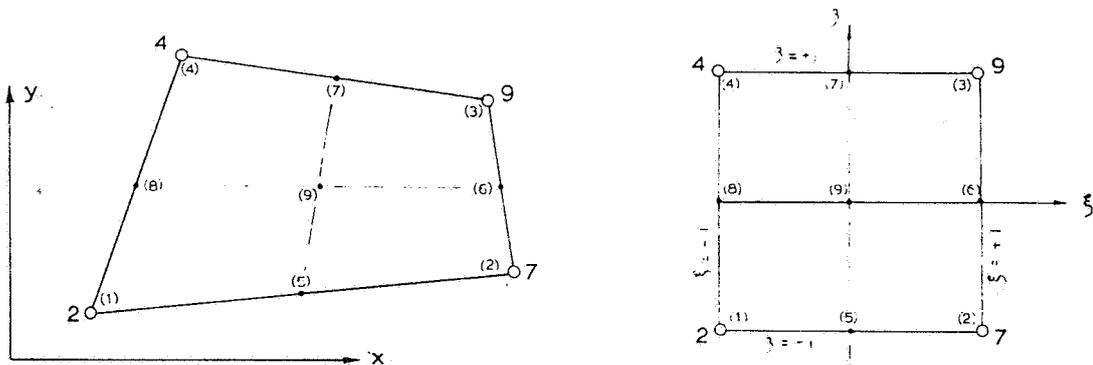


FIG. 2 - SYSTEME GLOBAL (x,y) ET SYSTEME LOCAL (xi,eta) DE COORDONNEES

NOMA et MISULIA [10] ont proposé, eux, une procédure itérative inspirée des différences finies. Après avoir affecté aux points de grille les plus proches les valeurs expérimentales et initialisé à la valeur moyenne les autres points de grille, on fait à chaque pas:

$$z(i,j) = \frac{1}{4} [z(i-1,j) + z(i+1,j) + z(i,j-1) + z(i,j+1)]$$

pour les points de grille (i,j) "non expérimentaux"; les points "expérimentaux" conservant leurs valeurs.

Cette méthode revient en fait à chercher, par la technique des différences finies, la fonction z satisfaisant l'équation de Laplace:

$$\Delta z = 0$$

avec comme conditions aux limites l'imposition des valeurs expérimentales aux points de grille les plus proches.

Ces deux méthodes n'ont jamais eu la diffusion qu'ont pu avoir l'interpolation polynômiale ou l'estimation par moindres carrés.

REFERENCES BIBLIOGRAPHIQUES - CHAP.1, § 1

- [1] AHLBERG, J.H., NILSON, E.N., WALSH, J.L. (1967):  
The theory of splines and their applications.  
Academic Press.
- [2] BATCHA, J.B., REESE, J.R. (1964):  
Surface determination and automatic contouring for mineral  
exploration, extraction and processing.  
Colo. School Mines, quart., 59, pp. 1-14
- [3] BIRKHOFF, G., GARABEDIAN, H.L. (1960):  
Smooth surface interpolation.  
Journ. Math. Physics, vol. 39, n°4, pp. 258-268
- [4] CRAIN, J.K. (1970):  
Computer interpolation and contouring of two-dimensional  
data: a review.  
Geoexploration, 8, pp. 71-86
- [5] DE BOOR, C. (1962)  
Bicubic spline interpolation  
Journ. Math. Physics, vol.41, n°3, pp. 212-218

- [6] HESSING, R.C., LEE, H.K., PIERCE, A., POWERS, E.N. (1972)  
Automatic contouring using bicubic functions.  
Geophysics, vol.37, n°4, pp. 669-674
- [7] HUTCHINSON, P., WALLEY, W.J. (1972):  
Calculation of areal rainfall using finite element techniques  
with altitudinal corrections.  
Bull. Int. Assoc. Scient. Hydrol., XVII, 3, pp. 259-272
- [8] KWAN, J.Y., RILEY, J.P., AMISIAL, R.A. (1969):  
A digital computer program to plot isohyctal maps and  
calculate volumes of precipitation.  
Symposium on the use of analog and digital computers in  
hydrology, Tucson, 1968.  
Int. Assoc. Scient. Hydrol. publ. n° 80 , pp. 240-248
- [9] MAC LAREN, M.D. (1958):  
Formulas for fitting a splined curved through a set of points.  
Appl. Math. Rep. 2, Applied Math. Staff, Boving Co., 12 p.  
cité par [12]
- [10] NOMA, A.A., MISULIA, M.G. (1959):  
Programming topographic maps for automatic terrain model  
construction.  
Surveying Mapping, 19, pp. 355-366  
cité par [4]
- [11] SHAW, E.M., LYNN, P.P. (1972):  
Areal rainfall evaluation using two surface fitting techniques.  
Bull. Int. Assoc. Scient. Hydrol., XVII, 4, pp. 419-433.
- [12] WHITTEN, E.H.T., KOELLING, M.E.V. (1973):  
Spline surface interpolation, spatial filtering and trend  
surfaces for geological mapped variables.  
Mathematical Geology, vol. 5, n°2, pp. 111-125.

## 2 - METHODE DES MOINDRES CARRES

On a vu l'échec des méthodes analytiques d'interpolation pour les grandeurs dont le comportement spatial est irrégulier, voire erratique, comme, par exemple, la pluviométrie ou la perméabilité. Quand les données montrent ainsi une importante dispersion, on ne cherche donc pas à contraindre la surface à passer par les points expérimentaux. On se contente d'imposer que les écarts aux points expérimentaux entre valeurs réelles et valeurs reconstituées par le modèle choisi soient aussi faibles que possible.

Les critères du type minimax étant d'un emploi malaisé, on a généralement recours à la méthode statistique d'ajustement par moindres carrés.

### a) Fondements théoriques

L'hypothèse sous-jacente est que le phénomène réel est régulier, - et peut être décrit par une fonction mathématique  $m(x)$  - mais qu'il s'y superpose une erreur  $\epsilon(x)$  de moyenne nulle, de variance constante et sans auto-corrélation:

$$z(x) = m(x) + \epsilon(x)$$

Prenant pour  $m(x)$  un modèle du type  $\sum_{l=1}^k a_l f^l(x)$ , on cherche à l'ajuster à l'aide de  $n$  points expérimentaux (avec évidemment  $n > k$ ), en minimisant l'expression:

$$\sum_{i=1}^n \left( \sum_{l=1}^k a_l f^l(x_i) - z_i \right)^2$$

Les coefficients  $a_l$  sont déterminés en égalant à zéro les dérivées partielles relatives à chacun d'eux, ce qui conduit au système d'équations suivant:

$$\sum_{l=1}^k a_l \cdot \left( \sum_{i=1}^n f^s(x_i) \cdot f^l(x_i) \right) = \sum_{i=1}^n f^s(x_i) \cdot z_i \quad s = 1, \dots, k$$

L'utilisateur est maître du choix des fonctions de base, de leur nombre  $k$ , et du nombre  $n$  de points expérimentaux pris en compte (s'il opère, comme c'est le cas généralement, par voisinages glissants).

En ce qui concerne les types de fonctions de base, on peut mentionner quelques essais de développements en double série de Fourier, employant des fonctions trigonométriques [3]. En dehors de ces tentatives, sont presque exclusivement utilisés des polynômes, généralement les monômes  $1, X, Y, X^2, XY, Y^2$ , etc... Parfois, les termes croisés, du type  $XY$ , sont omis pour simplifier les calculs ("abbreviated regression method" [6]), mais ce procédé n'est pas recommandé si le phénomène ne manifeste pas une orientation préférentielle nette dans la direction des axes de coordonnées. Certains auteurs [1] préconisent l'emploi de polynômes orthogonaux, c'est à dire tels que:

$$\sum_{i=1}^n f^l(x_i) \cdot f^s(x_i) = 0 \quad \forall l \neq s$$

Ces polynômes sont obtenus à partir des monômes  $1, X, Y, \dots$  par une technique d'orthogonalisation comme celle de GRAM-SCHMIDT. Les tenants de la méthode mettent en avant le fait que les coefficients  $a_l$  s'expriment ensuite simplement en fonction des  $f^l(x_i)$ ; mais le volume total des calculs semble voisin de celui correspondant à la pratique courante, avec peut-être un avantage quant à la précision en faveur des polynômes orthogonaux.

En ce qui concerne maintenant le choix des nombres de points et de fonctions de base à prendre, il est essentiellement fonction du résultat souhaité, le nombre de points étant toujours au moins de l'ordre du double du nombre de fonctions de base. Ainsi, par exemple, pour ajuster localement un plan ( $f^0=1, f^1=x, f^2=y$ ), on prend environ six points [7].

En jouant sur ces deux nombres, on décide implicitement de la part qui sera filtrée et de celle qui sera conservée et à laquelle on donnera généralement le nom de "trend" (tendance). Bien que le résultat puisse différer notablement, comme on peut le constater aux Fig. 3 et 4, selon que l'on adopte un ajustement polynômial du 1er ou du 2ème degré, on attribue généralement à cette trend une signification physique.

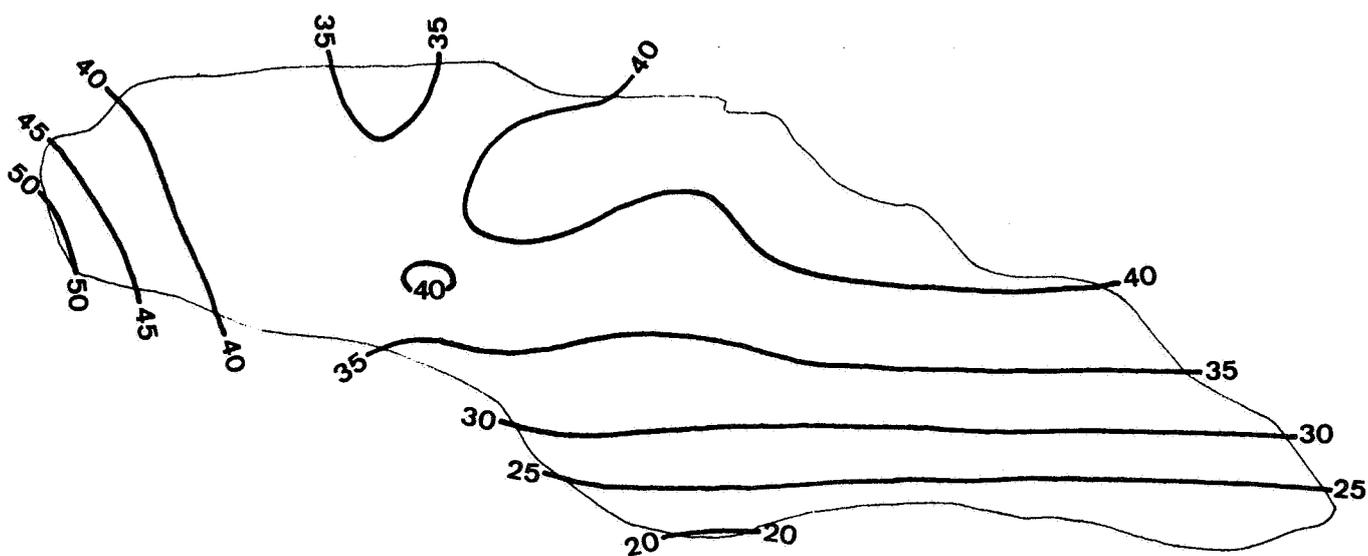


FIG. 3 - AJUSTEMENT LOCAL PAR MOINDRES CARRES D'UN POLYNÔME DU 1<sup>er</sup> DEGRE  
(Données pluviométriques sur le bassin versant de Kadjemeur)

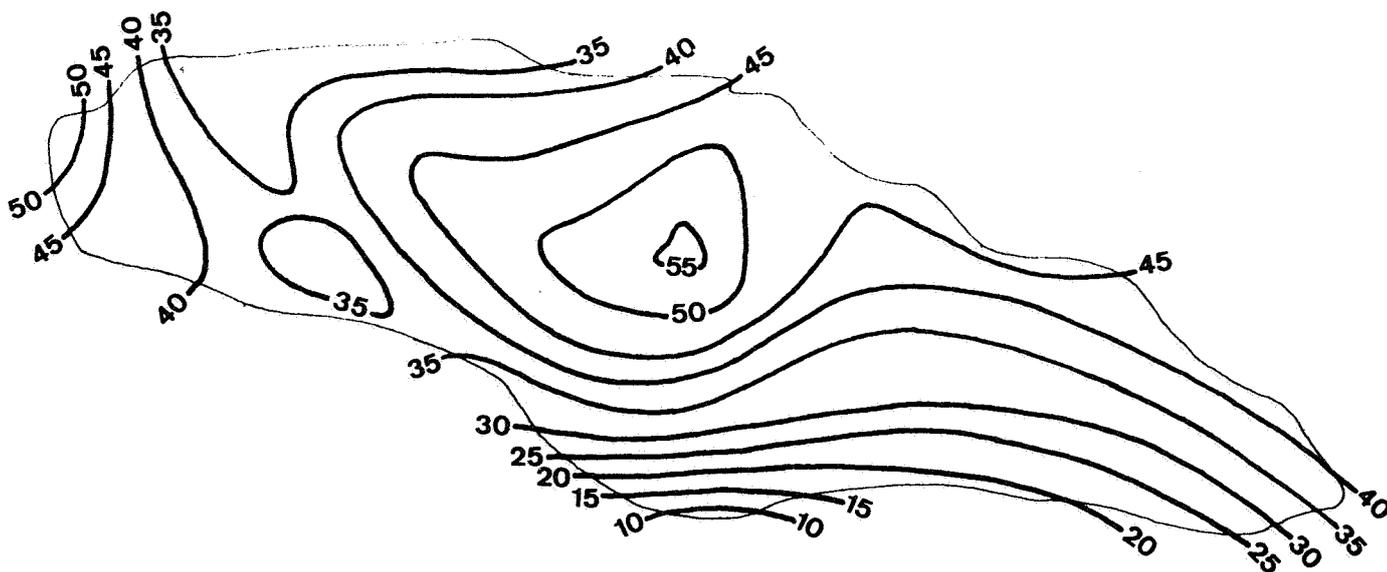


FIG. 4 - AJUSTEMENT LOCAL PAR MOINDRES CARRES D'UN POLYNÔME DU 2<sup>ème</sup> DEGRE

b) Notion de "trend"

En effet, les auteurs utilisant la technique des moindres carrés passent souvent de la formulation "mathématique" initiale:

$$\text{valeur observée} = \text{valeur d'une fonction déterministe} + \text{erreur aléatoire [8]}$$

à la formulation "physique" suivante, indiquant le but poursuivi:

$$\text{valeur observée} = \text{composante régionale} + \text{fluctuation locale}$$

(trend)                      (résidu ou anomalie)

On modélise à l'aide de fonctions mathématiques régulières la composante régionale ("car les causes hydrauliques et géologiques générales produisent toujours (en piézométrie par exemple) des effets morphologiques réguliers" [2]) et l'on impose à des résidus au statut peu clair un modèle sans auto-corrélation. Cette dichotomie entre trend et résidus est classique dans les sciences de la terre, notamment en géologie [3], [4], [5], [6]. Certains auteurs [2], s'inspirant des méthodes utilisées en géophysique, ont tenté de l'introduire en hydrogéologie pour l'étude de la piézométrie.

Mais l'assimilation entre les deux formulations présentées ci-dessus est un peu hâtive: les résidus ne constituent pas des erreurs; ils font partie intégrante du phénomène, dont ils rendent compte de fluctuations à échelle plus petite que la trend. De nombreux auteurs [3] admettent d'ailleurs que l'hypothèse d'absence d'auto-corrélation est rarement vérifiée en géologie, surtout si l'on adopte des polynômes de degré peu élevé comme modèle de "trend". Des méthodes de moindres carrés généralisées, visant à tenir compte de cette auto-corrélation, ont été parfois proposées.

Cependant, le principal reproche que l'on puisse adresser aux tenants des moindres carrés, c'est d'affirmer, comme on le lit souvent, que l'écart quadratique moyen de reconstitution aux points expérimentaux:

$$\frac{1}{n} \sum_{i=1}^n (z_i - \sum_{l=1}^k \hat{a}_l f^l(x_i))^2$$

constitue un bon critère pour juger de l'adéquation du modèle à la réalité. Le procédé des moindres carrés, minimisant les écarts entre trend et réalité aux points expérimentaux, ne garantit en rien que cet écart sera faible en dehors de ces points.

REFERENCES BIBLIOGRAPHIQUES - CHAP 1, § 2

- [1] CRAIN, I.K. (1970):  
Computer interpolation and contouring of two-dimensional data: a review.  
Geoexploration, 8, pp.71-86
- [2] GIURA, R., DE WRACHIEN, D., PONZINI, G.S. (1970):  
Procédés automatiques de calcul et de représentation des caractéristiques hydromorphologiques appliquées à l'étude de la surface piézométrique d'une nappe aquifère.  
Symposium scientifique international sur l'informatique et les problèmes de l'eau, Montpellier, 1970.
- [3] HARBAUGH, J.W., MERRIAM, D.F. (1968):  
Computer applications in stratigraphic analysis.  
John Wiley & Sons.
- [4] KRUMBEIN, W.C. (1956):  
Regional and local components in facies.  
Amer. Ass. Petrol. Geol. Bull., 40-9, pp. 2163-2194  
cité par [6]
- [5] KRUMBEIN, W.C. (1959):  
Trend surface analysis of contour-type maps with irregular control-point spacing.  
Jour. Geophysical Res., vol. 64, n°7, pp.823-824.  
cité par [3] et [6]
- [6] MILLER, R.L., KAHN, J.S. (1962):  
Statistical analysis in the geological sciences (ch. 17) some statistical approaches to mapping problems).  
John Wiley & Sons.
- [7] SHEARMAN, R.J., SALTER, P.M. (1975):  
An objective rainfall interpolation and mapping technique.  
Hydrol. Sci. Bull., XX, 3, pp. 353-363.
- [8] WATSON, G.S. (1972):  
Trend surface analysis and spatial correlation.  
The Geological Soc. of America - Special Paper 146.

3 - METHODES ELEMENTAIRES DE PONDERATION

Dans les méthodes polynômiales décrites aux deux paragraphes précédents, les coefficients  $a_1$  des différents monômes  $f^1$  pouvaient s'exprimer linéairement en fonction des données  $z_i = z(x_i)$ ; il en était donc de même pour les estimations auxquelles elles conduisaient en un point quelconque  $x_0$ :

$$\hat{z}_0 = \sum_{i=1}^k a_1 f^1(x_0) = \sum_{i=1}^n \lambda_i(x_0) z_i$$

Les poids  $\lambda_i$  dépendaient là des  $f^1(x_i)$  et  $f^1(x_0)$ , mais cette idée de pondérer les données a été exploitée bien en dehors du contexte des méthodes polynômiales

a) Pondération en fonction de la distance

La valeur  $z_0$  en un point non mesuré  $x_0$  ne devant pas être très différente des valeurs expérimentales  $z_i$  voisines, on a souvent adopté comme estimation de  $z_0$  une moyenne de ces valeurs expérimentales, calculée en donnant plus de poids aux points de mesure proches qu'au points plus éloignés:

$$\hat{z}_0 = \frac{\sum_{i=1}^n w(d_{i0}) z_i}{\sum_{i=1}^n w(d_{i0})}$$

où  $w(d)$  est la fonction de pondération,

$d_{i0}$  la distance du point estimé  $x_0$  au point expérimental  $x_i$ .

De nombreuses fonctions de pondération sont proposées dans la littérature, le choix étant plus ou moins arbitraire, question de convictions personnelles ou de tradition, et aussi fonction du résultat souhaité.

Si l'on veut disposer d'un interpolateur exact (c'est à dire que la valeur estimée soit égale à la valeur expérimentale dans le cas où le point à estimer vient à coïncider avec un point de mesure), il faut, par exemple, que la fonction de pondération satisfasse la condition:

$$w(d) \rightarrow \infty \quad \text{quand} \quad d \rightarrow 0$$

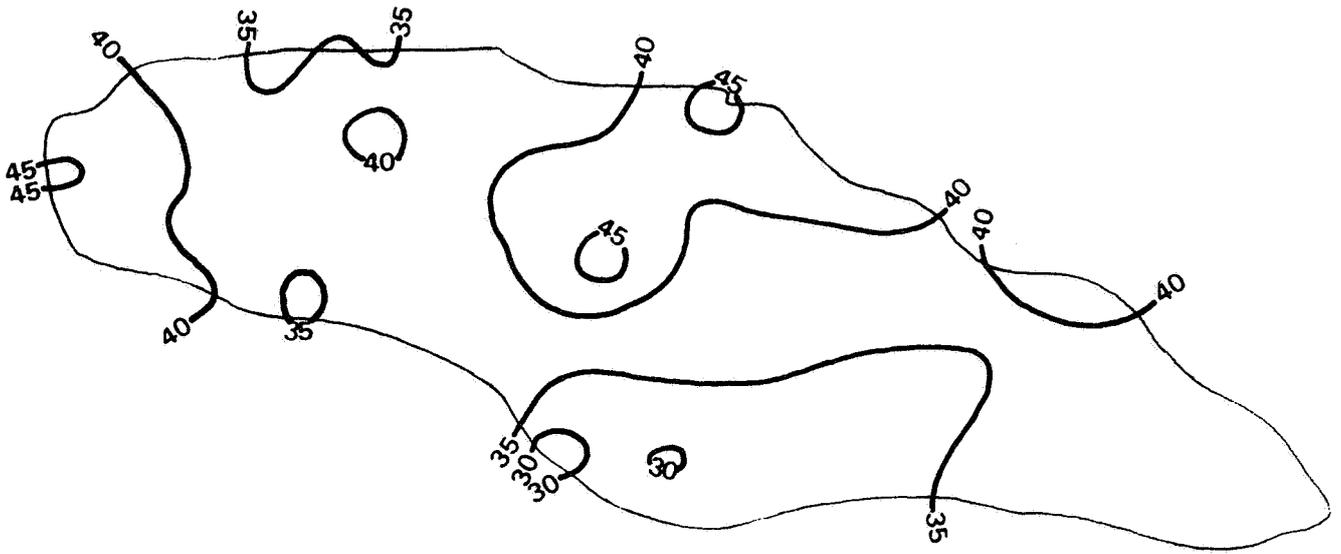


FIG. 5 - PONDERATION PAR L'INVERSE DE LA DISTANCE

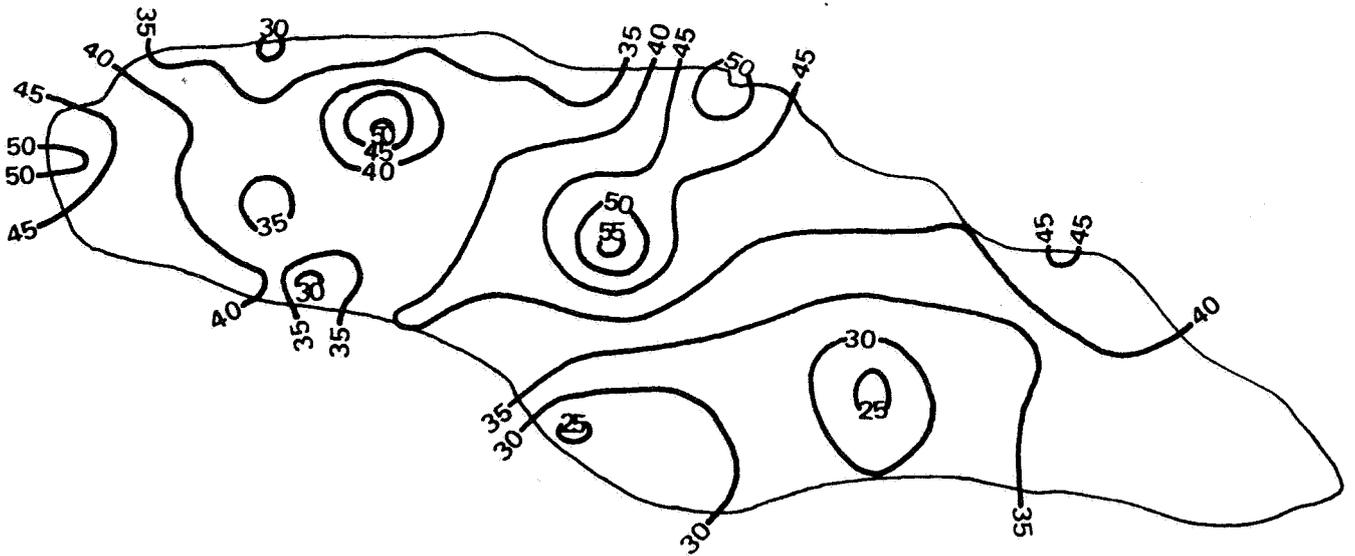


FIG. 6 - PONDERATION PAR L'INVERSE DU CARRE DE LA DISTANCE

Plus la confiance accordée aux données est faible, plus la forme de la fonction de pondération doit être plate, et plus cette confiance est forte, plus elle doit ressembler à un pic [2]. Selon la forme de cette fonction, on introduit ainsi un lissage plus ou moins important, destiné à filtrer ce que les différents auteurs appellent "bruit", "erreur" ou "fluctuations".

Le lissage le plus drastique est celui introduit par le procédé de la "moyenne mobile", qui consiste à assigner un poids identique à tous les points expérimentaux situés dans un certain voisinage:

$$w(d) = 1 \quad \text{si} \quad d \leq R \\ = 0 \quad \text{si} \quad d > R$$

Parmi les autres fonctions de pondération utilisées, on peut citer:

$$w(d) = \frac{1}{d} \quad \text{ou} \quad \frac{1}{d^2} \quad [2], [8] \\ = \frac{1}{d+1} \quad \text{ou} \quad \frac{1}{d^2+1} \quad [6] \\ = e^{-\alpha d} \quad [7]$$

cette dernière forme introduisant un lissage plus ou moins important selon la valeur de  $\alpha$ .

On trouvera, représentées aux figures 5 et 6, deux cartes obtenues à partir des mêmes données en adoptant une pondération dans un cas en  $\frac{1}{d}$  et dans l'autre en  $\frac{1}{d^2}$ ; on peut constater combien le choix arbitraire d'une fonction de pondération peut influencer sur l'aspect de la carte.

Mais l'arbitraire ne constitue pas le seul inconvénient de la méthode. Ne prenant en compte que les distances des différents points expérimentaux au point à estimer, elle ne parvient pas à discriminer l'information redondante.

Considérons, par exemple, la configuration de la figure 7a. Les trois points expérimentaux, situés à égale distance du point à estimer, se voient naturellement assigner des poids égaux (1/3). Mais si le point A est remplacé par deux points très proches A' et A" (Fig.7b), une méthode

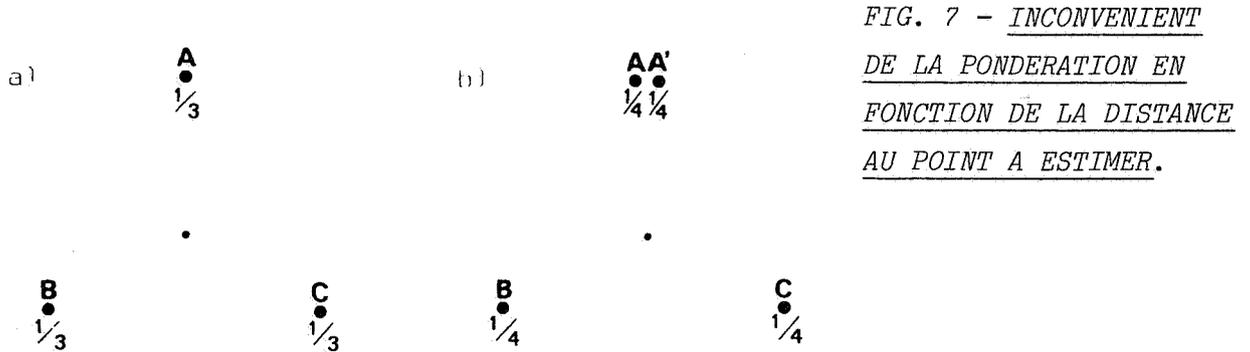


FIG. 7 - INCONVENIENT DE LA PONDERATION EN FONCTION DE LA DISTANCE AU POINT A ESTIMER.

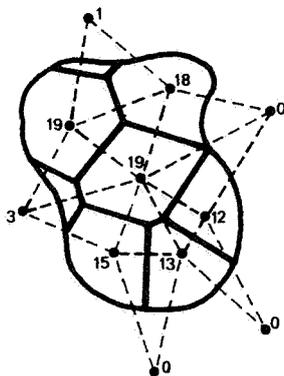
de pondération en fonction de la distance leur assignera à chacun ainsi qu'aux deux autres points B et C, un poids égal à 1/4. L'influence des deux points B et C sur l'estimation au point central viendrait ainsi à diminuer, alors que la part d'information qu'ils apportent reste identique, et que leurs poids devraient rester inchangés.

b) Méthode des polygones d'influence

Une méthode de pondération élémentaire a été également développée pour l'estimation des valeurs moyennes sur un domaine. Elle a l'avantage sur la précédente de tenir compte de la position relative des différents points expérimentaux entre eux et par rapport au domaine, et de pouvoir ainsi discriminer l'information redondante.

La méthode des polygones, proposée par THIESSEN [9], et adoptée par tous les hydrologues [1], correspond à l'hypothèse suivante: une station

est représentative de l'ensemble des points du bassin pour lesquels elle est la station la plus proche. Sa mise en oeuvre nécessite le tracé d'un réseau polygonal formé par les médiatrices des segments joignant les stations (Fig.8). On réalise ainsi une partition du bassin en n zones d'influence, de surface  $S_i$ , avec:



$$S = S_1 + S_2 + \dots + S_n$$

FIG. 8 - TRACE DES POLYONES ET PONDERATION DE THIESSEN (en %)  
(d'après [1])

Sur le plan formel, l'estimation, par la méthode de THIESEN, de la lame d'eau  $\bar{z}_S = \frac{1}{S} \int z(x)dx$  est la moyenne, pondérée par ces surfaces  $S_i$ , des hauteurs d'eau  $z_i$  mesurées aux n stations:

$$\hat{z}_S = \frac{1}{S} \sum_{i=1}^n S_i \cdot z_i$$

Dire que la station i est représentative du polygone correspondant revient, en effet, à assimiler  $z_i$  à la hauteur d'eau moyenne tombée sur le polygone, c'est à dire à  $\frac{1}{S_i} \int_{S_i} z(x)dx$ .

Parmi les inconvénients de cette méthode, le premier réside dans la relative lourdeur de sa mise en oeuvre, notamment lorsque des absences de mesures en certaines stations obligent à tracer autant de réseaux polygonaux qu'il y a eu de configurations expérimentales. Pour pallier ce dernier inconvénient, on a cherché à automatiser la détermination des poids de THIESEN, en substituant leur calcul par une méthode de Monte-Carlo [3], ou une méthode associant à une station chaque point d'une grille régulière recouvrant le bassin étudié [5], au procédé graphique traditionnel.

Le second inconvénient présenté par la méthode est qu'elle a été conçue pour être employée à grande échelle dans des régions sans influence orographique très marquée. Ceci a nécessité son adaptation pour prendre en compte les altitudes des stations, en les ramenant à la même cote par régression de la pluviométrie sur l'altitude, ou en rapportant les valeurs expérimentales à leur moyenne saisonnière ou annuelle de façon à compenser l'effet orographique [1].

Les méthodes qui viennent d'être exposées dans ce paragraphe et les précédents réalisent, en fait, des pondérations "passe-partout". Le mode de pondération n'est pas lié aux caractéristiques du phénomène étudié, mais est toujours dû à un choix plus ou moins arbitraire. Les méthodes probabilistes introduites en hydrométéorologie par l'école soviétique reposent, elles, sur une mise en évidence préalable de la variabilité spatiale des phénomènes.

REFERENCES BIBLIOGRAPHIQUES - CHAP.1, § 3

- [1] CHOW, V.T. & al. (1964):  
Handbook of applied hydrology (section 9 - Rainfall)  
Mc Graw Hill
- [2] CRAIN, I.K., BHATTACHARYYA, B.K. (1967):  
Treatment of non-equispaced two-dimensional data with a  
digital computer.  
Geoexploration, 5, pp. 173-194.
- [3] DISKIN, M.H. (1969)  
Thiessen coefficients by a Monte-Carlo procedure.  
J. Hydrol., 8, pp. 323-335.
- [4] DISKIN, M.H. (1970):  
On the computer evaluation of Thiessen weights.  
J. Hydrol., 11, pp. 69-78.
- [5] GOUBET, A., GENIN, M. (1967):  
Application de la méthode de Thiessen à l'aide de machines  
électroniques.  
1ère C.E. (Service d'Etude des Débits) - S.P.E.P.E.
- [6] NEWTON, R. (1966):  
Deriving contour maps from geological data.  
Can. J. Earth Sci., 5, pp. 165-166.
- [7] NEWTON, R. (1973):  
A statistical prediction technique for deriving contour maps  
from geophysical data.  
Mathematical Geology, vol.5, n°2, pp. 179-189
- [8] SHEARMAN, R.J., SALTER, P.M. (1975):  
An objective rainfall interpolation and mapping technique.  
Hydrol. Sci. Bull., XX, 3, pp. 353-363.
- [9] THIESSEN, A.H. (1911):  
Precipitation for large areas.  
Monthly Weather Review, vol. 39, n°7, pp. 1082-1084.

4 - METHODES PROBABILISTES

a) Caractérisation de la structure

Ce paragraphe concernant l'emploi de méthodes statistiques, ou plutôt probabilistes, ne se réfère qu'à des travaux portant sur les précipitations. L'hydrométéorologue, disposant de plusieurs réalisations du phénomène étudié (plusieurs averses, ou une série de pluies journalières,...), a toujours établi des statistiques simples ponctuelles: moyenne et variance dans le temps calculées à une station. Il a été ensuite amené naturellement à envisager l'étude des liaisons spatiales en utilisant le calcul des corrélations entre stations.

α) Normales, anomalies et coefficients de corrélation

Soient:  $z$  le champ considéré (par exemple précipitations journalières)  
 $\{x_i \mid i=1, \dots, n\}$  l'ensemble des stations de mesure  
 $z_i = z(x_i)$  la valeur du champ au point  $x_i$ .

Si une barre en position supérieure indique la moyenne temporelle, on aura en chaque station:

$$\begin{aligned} \overline{z_i} & \text{ moyenne ou "normale" au point } x_i \\ z'_i = z_i - \overline{z_i} & \text{ écart à la moyenne ou "anomalie" au point } x_i \\ s_i^2 = \overline{z_i'^2} & \text{ variance au point } x_i \\ r_{ij} = \frac{\overline{z_i' z_j'}}{s_i s_j} & \text{ coefficient de corrélation entre les points } x_i \text{ et } x_j. \end{aligned}$$

La pratique consistant à travailler sur les anomalies, éventuellement sur les pluies "standardisées" ( $z'/s$ ) [2] (ou chez certains auteurs [3] à travailler en pourcentage de la moyenne temporelle), vise à homogénéiser le champ des précipitations qui est généralement non stationnaire.

Notons que le champ des normales a les propriétés de régularité spatiale d'une "trend" (tendance en grand): les petites fluctuations sont lissées quand on prend des moyennes sur une longue période, ne laissant plus apparaître que les traits persistant à grande échelle.

A partir de  $n$  stations, on peut calculer  $C_n^2 = \frac{n(n-1)}{2}$  coefficients de corrélation différents. Ces résultats peuvent être présentés sous forme de  $n$  cartes montrant, pour chaque station, ses coefficients de corrélation avec les  $n-1$  autres stations (Fig.9). La plupart des auteurs se contentent en fait d'une seule carte se référant à une station-clé occupant une position centrale dans le réseau. On se limite alors à une vue partielle des liaisons spatiales, sur laquelle il est dangereux de raisonner, l'aspect d'une telle carte pouvant changer considérablement si l'on change de station-clé [8].

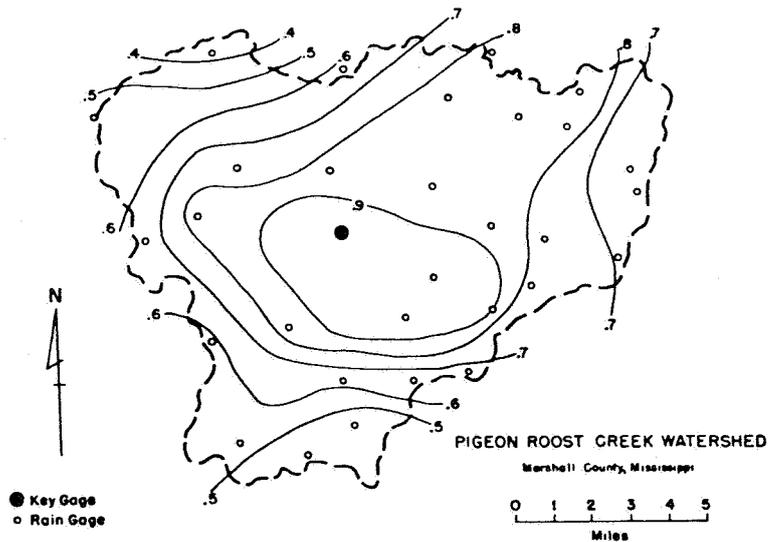


FIG. 9 - COURBES D'ISOCORRELATION ENTRE UNE STATION-CLE ET LES AUTRES STATIONS D'UN RESEAU  
(d'après [8])

β) Fonction de corrélation

Il est préférable d'adopter une vision plus globale, en utilisant la notion de fonction de corrélation introduite en hydrométéorologie par l'école soviétique (DROZDOV [5], GANDIN [7], KAGAN [9]). Sous certaines hypothèses de stationnarité du champ, on peut considérer que le coefficient de corrélation entre deux points ne dépend que du vecteur les joignant, et non de leurs emplacements.

On peut alors reporter  $r_{ij}$  en fonction de la distance  $d_{ij}$  (Fig.10), obtenant ainsi un nuage de  $C_n^2$  points illustrant la décroissance de la corrélation avec la distance.

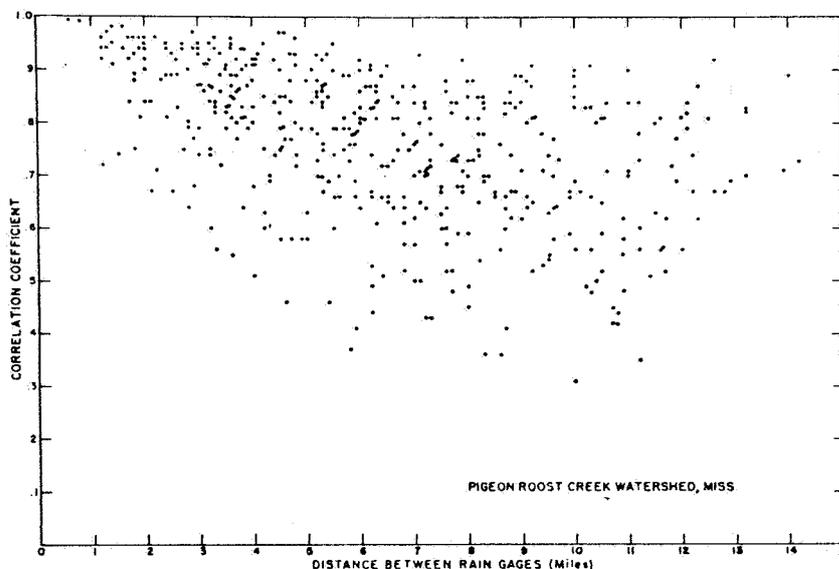


FIG. 10 - REPORT EN FONCTION DE LA DISTANCE DES COEFFICIENTS DE CORRELATION DE TOUS LES COUPLES DE PLUVIOMETRES D'UN BASSIN VERSANT (d'après [8])

Une trop grande dispersion de ce nuage pourra suggérer de classer les couples par direction, et d'étudier  $r_{ij}$  en fonction de  $d_{ij}$  et  $\theta_{ij}$  [8] et [16], en cherchant à mettre en évidence une éventuelle anisotropie.

On essaie ensuite de tracer une courbe empirique  $r = r(d)$ , en calculant, par exemple, pour différentes valeurs choisies de  $r$  la valeur moyenne de  $d$  correspondante, ou vice-versa [8] (Fig.11).

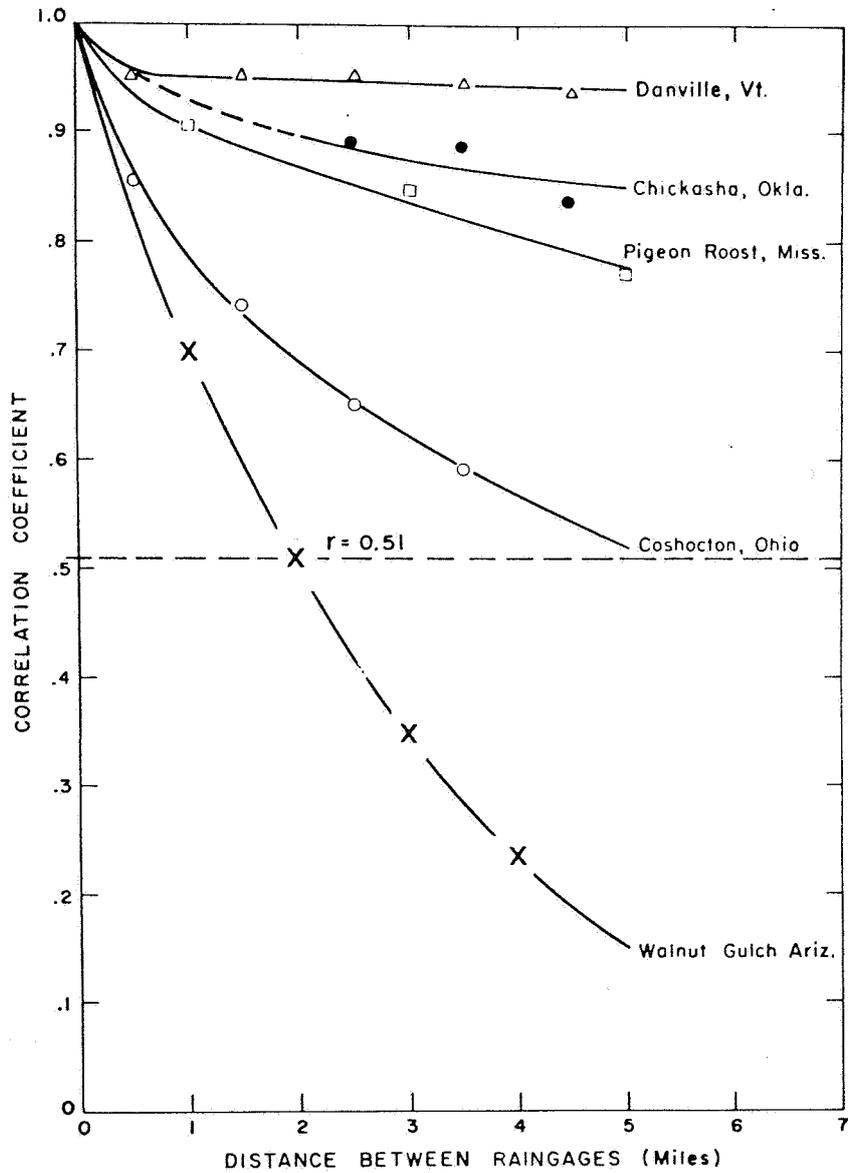


FIG. 11 - TRACE DE LA FONCTION DE CORRELATION SPATIALE POUR  
5 BASSINS VERSANTS DES ETATS-UNIS  
(d'après [8])

L'étape suivante consiste à ajuster une fonction mathématique à cette courbe empirique. Différents modèles ont été proposés: l'école soviétique a adopté une forme à décroissance exponentielle  $r(d) = e^{-\alpha d}$  [12]; RODRIGUEZ-ITURBE [14], se référant plus tard aux travaux de MATERN [11], WHITTLE [19] et YAGLOM [12], a proposé d'introduire en outre un modèle en  $r(d) = \alpha d K_1(\alpha)$ , où  $K_1$  est une fonction de Bessel modifiée de deuxième espèce. Reste la question de l'ajustement du paramètre  $\alpha$ : celui-ci est fait soit par moindres carrés en prenant l'ensemble des valeurs de coefficients de corrélation obtenues, soit en prenant comme critère de bien rendre compte des corrélations à la distance moyenne de deux points courants du bassin versant étudié [15].

γ) Interprétation climatologique

L'interprétation climatologique de ces résultats s'appuie sur le fait que, plus la décroissance de la fonction de corrélation est rapide, plus la variabilité spatiale de la pluviométrie est importante.

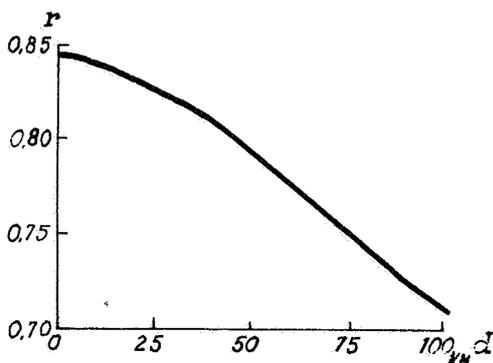


FIG. 12 - FONCTION DE CORRELATION DE LA PLUVIOMETRIE MENSUELLE (VALDAI, URSS) (d'après [9])

De nombreux auteurs, en particulier les auteurs soviétiques [9], [12], [2], ont mis en évidence le fait que, en extrapolant à la distance zéro la fonction de corrélation bâtie à partir des données, on n'obtenait pas la valeur 1, mais une valeur légèrement inférieure. Ils ont attribué cette particularité à l'existence "d'erreurs" imputables aux mesures ou à des phénomènes microclimatiques, et donc non corrélées d'une station à une autre [9]. Leur présence revient en effet à majorer les variances ponctuelles d'une quantité égale à la variance de ces "erreurs", et donc, majorant les dénominateurs, à minorer les valeurs trouvées pour les coefficients de corrélation.

L'impact des phénomènes microclimatiques sur la fonction de corrélation spatiale a été étudié par SHARON [16], [17]. Celui-ci a montré que, si les erreurs instrumentales ont pour effet une diminution de la fonction de corrélation indépendante de la distance, la présence de cellules convectives se traduit en fait par une décroissance rapide (le taux étant fonction inverse

de la taille des cellules) de cette fonction au voisinage de l'origine. Quand on dispose d'un réseau pluviométrique suffisamment dense [16], [6] ou d'un balayage radar [1] (cf. Fig.13), on peut observer, en effet, la structure cellulaire de la précipitation due à une convection à petite échelle se superposant aux phénomènes d'échelle supérieure. Celle-ci se traduit sur les isohyètes par la présence de pastilles isolées entourant des maxima locaux, d'où le nom de "spottiness" donné par SHARON [17].

L'examen de la fonction de corrélation est donc riche d'enseignement sur la structure spatiale de la pluviométrie; mais cet outil, ou d'autres outils probabilistes voisins, a permis également de donner un support mathématique aux notions d'erreur d'interpolation ou d'erreur sur le calcul des moyennes spatiales, et d'optimiser ces opérations.

b) Prise en compte de la structure dans l'estimation

α) Structure et erreur d'estimation

Les concepts de base, en ce qui concerne l'étude des réseaux pluviométriques et leur rationalisation, sont en effet l'erreur maximum admissible sur l'estimation de la pluie en un point quelconque non mesuré, et l'erreur maximum admissible sur l'estimation de la pluie moyenne sur une zone déterminée.

Le concept "d'erreur quadratique moyenne d'interpolation linéaire" est introduit en 1936 par DROZDOV [4] qui le rattache à la distance séparant les stations. Dans un travail ultérieur [5], il montre que l'erreur en question est uniquement liée à ce qu'il appelle la "fonction de structure", c'est à dire à l'évolution de la moyenne quadratique des écarts entre les valeurs prises par le phénomène en deux points, quand on fait varier la distance séparant ces deux points.

Comme on a défini  $r_{ij}$ , on définit alors:

$$b_{ij} = \overline{(z_i - z_j)^2}$$

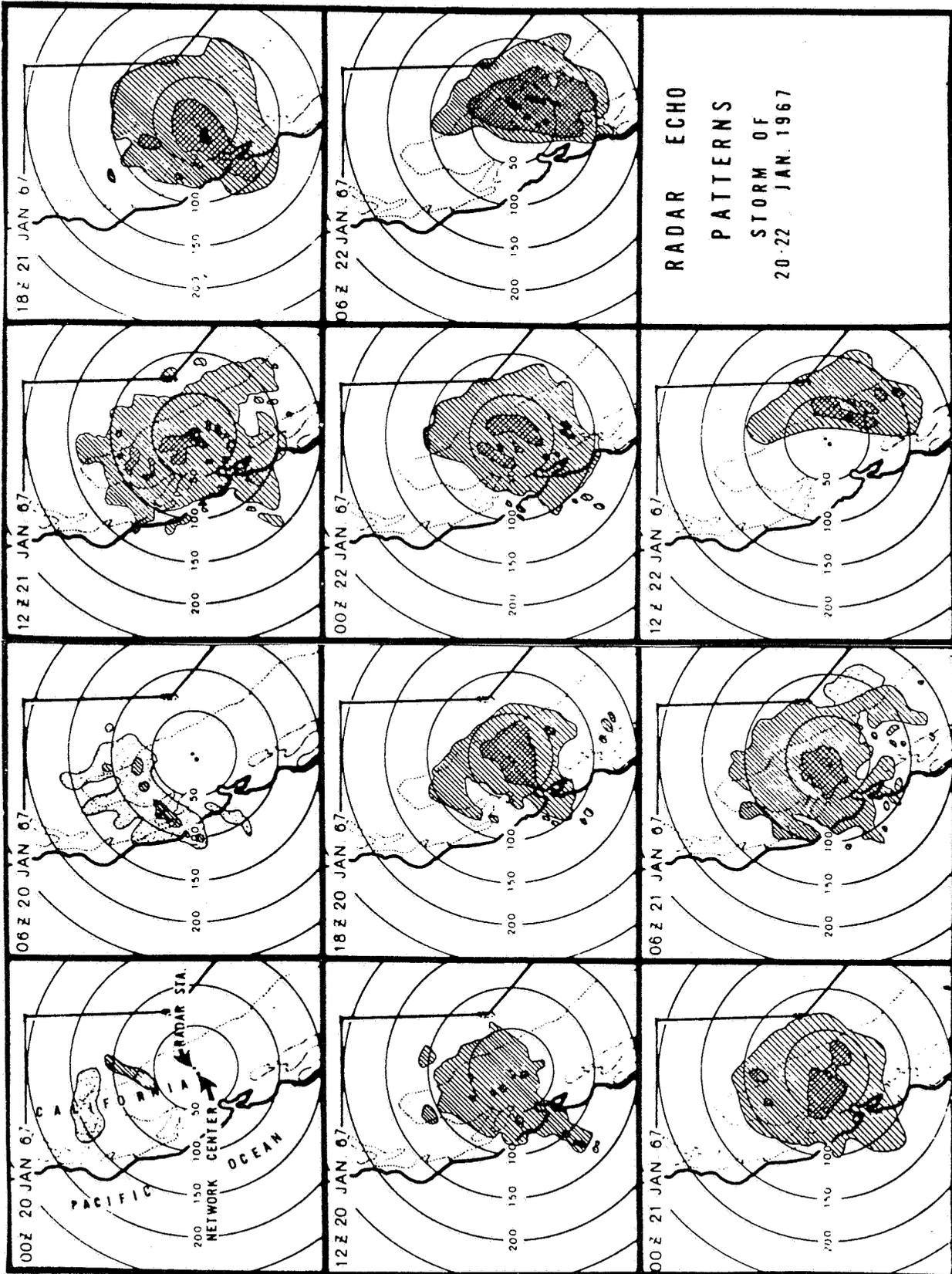


FIG. 13 - ECHOS RADAR OBSERVES DURANT UN ORAGE (CALIFORNIE, USA) - (d'après [1])

et l'on passe à la fonction de structure  $b = b(d)$  en faisant les hypothèses de stationnarité et d'isotropie nécessaires, cette fois-ci, sur les écarts.

Cette fonction de structure est reliée d'une façon simple à la fonction de corrélation. Si, par exemple, on travaille sur les pluies "standardisées" [2], on a:

$$\beta_{ij} = \left( \frac{z'_i}{s_i} - \frac{z'_j}{s_j} \right)^2 = 2(1-r_{ij})$$

ou bien encore, dans le cas stationnaire où:

$$\begin{aligned} \bar{z}_i &= \bar{z}_j = m \\ s_i^2 &= s_j^2 = s^2 \end{aligned} \quad \forall i, j$$

on a:  $b(d) = 2s^2 (1-r(d))$

Utilisant cet outil pour l'étude des erreurs d'interpolation, DROZDOV [5] montre que l'erreur quadratique moyenne correspondant à une interpolation linéaire entre deux stations est maximum au milieu du segment joignant les deux stations et vaut alors:

$$\epsilon^2 = b\left(\frac{l}{2}\right) - \frac{1}{4} b(1)$$

si  $l$  désigne la distance entre les deux stations.

KHRGIAN [10] a appliqué la même approche à l'interpolation linéaire entre trois stations. On montre alors [5] qu'au point central d'un triangle équilatéral, on a comme erreur quadratique moyenne:

$$\epsilon^2 = b\left(\frac{l}{\sqrt{3}}\right) - \frac{1}{3} b(1)$$

si  $l$  désigne là encore la distance entre les stations.

De même, en ce qui concerne l'estimation de la précipitation moyenne sur un domaine  $S$ , le résultat suivant a été établi [9], en utilisant cette fois la fonction de corrélation: à l'estimation par la moyenne arithmétique simple  $\frac{1}{n} \sum_{i=1}^n z_i$  des valeurs mesurées aux  $n$  stations correspond une

erreur quadratique moyenne égale à

$$\epsilon^2 = s^2 \cdot \left[ \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n r_{ij} - \frac{2}{n} \sum_{i=1}^n r_{iS} + r_{SS} \right]$$

où  $r_{iS}$  désigne la valeur moyenne de la corrélation entre la station  $i$  et un point décrivant le domaine  $S$ ,

et  $r_{SS}$  désigne la valeur moyenne de la corrélation entre deux points décrivant indépendamment le domaine  $S$ .

ces deux valeurs étant calculées à partir de la fonction  $r(d)$ .

De même, avec les mêmes notations, en adoptant la pondération de THIESSEN  $\sum_{i=1}^n T_i z_i$ , on a :

$$\epsilon^2 = s^2 \cdot \left[ \sum_{i=1}^n \sum_{j=1}^n T_i T_j r_{ij} - 2 \sum_{i=1}^n T_i r_{iS} + r_{SS} \right]$$

Sont ainsi quantifiées les précisions des méthodes les plus classiques d'interpolation ou d'estimation de valeur moyenne. Mais la démarche utilisée permet d'aller plus loin, et d'optimiser les modes de pondération.

### β) Interpolation optimale

En interpolation, GANDIN [7], travaillant sur les anomalies, forme, pour estimer  $z'_0$  en un point  $x_0$  situé en dehors des stations, une combinaison linéaire (ou encore une moyenne pondérée) des anomalies  $z'_i$  aux stations  $\{x_i, i=1, \dots, n\}$

$$\hat{z}'_0 = \sum_{i=1}^n p_i z'_i$$

Il cherche ensuite les poids  $p_i$  qui minimisent l'erreur quadratique moyenne :

$$\epsilon_0^2 = \overline{(\hat{z}'_0 - z'_0)^2}$$

Le calcul se conduit de la façon suivante :

$$\begin{aligned}\epsilon_0^2 &= \left( \sum_{i=1}^n p_i z'_i - z'_0 \right)^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n p_i p_j z'_i z'_j - 2 \sum_{i=1}^n p_i z'_i z'_0 + z_0'^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n p_i p_j \overline{z'_i z'_j} - 2 \sum_{i=1}^n p_i \overline{z'_i z'_0} + \overline{z_0'^2}\end{aligned}$$

Si l'on travaille, comme ci-dessus, en corrélation et avec les mêmes hypothèses, on obtient:

$$\epsilon_0^2 = s^2 \cdot \left[ \sum_{i=1}^n \sum_{j=1}^n p_i p_j r_{ij} - 2 \sum_{i=1}^n p_i r_{i0} + r_{00} \right]$$

Pour minimiser  $\epsilon_0^2$ , il suffit d'annuler les dérivées partielles par rapport aux  $p_i$ . D'où le système de n équations permettant de calculer les poids optimaux:

$$\sum_{j=1}^n p_j r_{ij} = r_{i0} \quad i = 1, \dots, n$$

La matrice  $(r_{ij})$ , étant une matrice de corrélation, est définie positive, ce qui garantit une solution unique au système.

Le second membre est déterminé à partir de la fonction de corrélation obtenue par ajustement:  $r_{i0} = r(d_{i0})$ . A noter l'importance du modèle mathématique retenu, celui-ci devant être de type positif (c'est à dire avoir une transformée de Fourier positive) pour éviter d'obtenir une valeur négative pour  $\epsilon_0^2$ .

Avec le jeu de poids optimum, l'expression donnant  $\epsilon_0^2$  se simplifie et devient:

$$\epsilon_0^2 = s^2 \left[ r_{00} - \sum_{i=1}^n p_i r_{i0} \right]$$

γ) Moyenne pondérée optimale

De même, en ce qui concerne l'estimation de la moyenne sur un domaine S, KAGAN [9] donne le système:

$$\sum_{j=1}^n p_j r_{ij} = r_{iS} \quad i=1, \dots,$$

et à l'optimum:

$$\epsilon_S^2 = s^2 \cdot r_{SS} - \sum_{i=1}^n p_i r_{iS}$$

En outre, la démarche suivie permet de tenir compte de la présence d'erreurs de mesure, si celle-ci a été décelée sur la fonction de corrélation (ou sur la fonction de structure) ajustée sur les valeurs expérimentales.

Les erreurs de mesure  $\delta_i$  sont supposées non corrélées

- aux valeurs du champ:  $\overline{\delta_i \cdot z'_j} = 0 \quad \forall i, \forall j$

- entre elles:  $\overline{\delta_i \cdot \delta_j} = 0 \quad \forall i \neq j$   
 $= C \quad \text{si } i=j$

Soit  $r_e(d)$  la fonction de corrélation expérimentale. Si  $r(d)$  est la fonction de corrélation réelle du champ, il existe alors, entre  $r_e$  et  $r$ , un rapport  $\frac{s^2}{s^2+C}$ , les variances étant surestimées de C. En extrapolant à la distance nulle, on a:

$$r_e(0) = \frac{s^2}{s^2+C}$$

et donc [9] [12]:

$$\eta = \frac{C}{s^2} = \frac{1-r_e(0)}{r_e(0)}$$

Le système donnant les poids optimaux devient alors, par exemple dans le cas de l'estimation de la moyenne sur un domaine S [12]:

$$\sum_{j=1}^n p_j r_{ij} + \eta p_i = r_{iS} \quad i=1, \dots, n$$

### δ) Applications

Cette méthodologie, basée sur l'emploi de la fonction de corrélation spatiale du champ des précipitations pour résoudre les problèmes d'estimation ponctuelle et zonale, date des années 60. Elle a eu peu d'échos en dehors de l'école qui l'avait développée.

Parmi les rares applications que l'on puisse trouver dans la littérature, on peut citer le travail de CISLEROVA et HUTCHINSON [2] sur la rationalisation du réseau pluviométrique de Zambie. Ces auteurs ont suivi l'approche de DROZDOV, utilisant comme critère l'erreur quadratique moyenne d'interpolation aux milieux des segments reliant deux à deux les stations du réseau. L'étude a porté sur les pluies annuelles préalablement "standardisées" (anomalies rapportées en chaque point aux écarts-types); normales, variances et fonction de corrélation ont été déterminées à partir des 116 stations où existaient plus de trente années de mesures. Le résultat concret de cette étude a été l'optimisation du doublement du réseau de mesure zambien. Dans leurs conclusions, les auteurs mentionnent comme une limite de la méthode employée le fait de n'avoir pas pris en compte les variations spatiales systématiques de la pluviométrie.

Bien que ne se référant pas aux mêmes auteurs, RODRIGUEZ-ITURBE [14] utilise, depuis quelques années, la fonction de corrélation spatiale pour étudier l'erreur d'estimation de la pluviométrie moyenne sur un bassin, pour différents types d'implantation des stations (échantillonnage aléatoire et aléatoire stratifié), en se limitant à l'emploi de la moyenne arithmétique simple. Il aborde en outre l'étude de la variabilité spatio-temporelle de la pluviométrie, en adjoignant à l'hypothèse de stationnarité une hypothèse de séparabilité lui permettant une décomposition en un produit de deux fonctions de corrélation, l'une spatiale, l'autre temporelle [15].

La plupart des méthodes d'estimation décrites dans ce premier chapitre utilisent en fait une pondération des observations qui peut s'écrire:

$$\hat{Z} = \sum_{i=1}^n p_i z(x_i)$$

avec  $p_i$  poids attribué au point expérimental  $x_i$ ,  
 $z(x_i)$  valeur observée au point  $x_i$ ,  
 $n$  nombre de points utilisé.

chacune de ces méthodes correspondant à un ensemble de règles permettant de déterminer les poids  $p_i$ .

La méthode développée par l'école soviétique semble, de toutes, la plus solide, car elle tient compte de la structure du phénomène étudié. Elle présente cependant encore quelques inconvénients, comme l'obligation de travailler sur les anomalies et surtout la nécessité de l'existence d'un ensemble de situations comparables. Il reste à s'affranchir de ces deux conditions. Ce sera la tâche de la théorie des variables régionalisées due à G. MATHERON.

#### REFERENCES BIBLIOGRAPHIQUES - CHAP.1, § 4

- [1] AMOROCHO, J., BRANDSTETTER, A., MORGAN, D. (1968):  
The effects of density of recording rain gauge networks on the description of precipitation patterns.  
Symposium on Geochemistry, Precipitation, Evaporation, Soil-Moisture and Hydrometry, Bern, 1967.  
Int. Assoc. Scient. Hydrol. Publ. n°78, pp. 189-202.
- [2] CISLEROVA, M., HUTCHINSON, P. (1974):  
The redesign of the raingauge network of Zambia.  
Hydrol. Sci. Bull., XIX, 4, pp. 423-434
- [3] DESI, F., CZELNAI, R., RAKOCZI, F. (1965):  
On determining the rational density of precipitation measuring networks.  
Symposium on the Design of Hydrological Networks, Pt I, Quebec 1965.  
Int. Assoc. Scient. Hydrol. Publ. n°67, pp. 127-129.
- [4] DROZDOV, O.A. (1936) cité par [13]  
A method for setting up a network of meteorological stations for a level region.  
Trudy GGO n°12 (3)

- [5] DROZDOV, O.A., SEPELEVSKIY, A.A. (1946) cité par [2] et [13]  
The theory of interpolation in a stochastic field of meteorological elements and its application to meteorological map and network rationalization problems.  
Trudy NIU GUGMS, ser.1, n°18
- [6] FELGATE, D.G., READ, D.G. (1975):  
Correlation analysis of the cellular structure of storms observed by raingauges.  
J. Hydrol., 24, pp. 201-205.
- [7] GANDIN, L.S. (1963): cité par [9]  
Ob'ektivnyj analiz meteorologicheskikh polej.  
Gidrometeoizdat, Leningrad, 1963.  
Objective analysis of meteorological fields.  
Israël Program for scientific translations, Jerusalem, 1965.
- [8] HERSHFELD, D.M. (1968)  
Rainfall input for hydrologic models.  
Symposium on Geochemistry, Precipitation, Evaporation, Soil-Moisture and Hydrometry, Bern, 1967.  
Int. Assoc. Scient. Hydrol. Publ. n°78, pp. 177-188
- [9] KAGAN, R.L. (1967):  
Nekotorye voprosy interpretacii osadkomernykh dannyykh  
(Quelques problèmes relatifs à l'interprétation des données de précipitations).  
Trudy Glavn. Geofiz. Obs. A.I. Voejkova n°208, pp. 64-75
- [10] KHRGIAN, A. Kh (1938): cité par [13]  
To what extent is temperature interpolation on a climatic map permissible ?  
Meteorologia i Hidrologia n°7.
- [11] MATERN, B. (1960):  
Spatial variation  
(Stochastic models and their application to some problems in forest survey and other sampling investigations).  
Coll. Meddelanden från Statens Skogsforskningsinstitut, Band 49, n°5.  
Almaenna Foerlaget, Stockholm.
- [12] POLISHCHUK, A.I. (1967):  
K voprosu ob osrednenii osadkomernykh dannyykh  
(Sur le problème de la moyenne des données de précipitations).  
Trudy Glavn. Geofiz. Obs. A.I. Voejkova, n°208, pp. 76-87
- [13] RODDA, J.C. & al. (1969):  
Hydrological network design - needs, problems and approaches.  
WMO Reports on WMO/IDH Projects - Report n°12, pp. 8-11
- [14] RODRIGUEZ-ITURBE, I., VANMARCKE, E.H., SCHAAKE, J.C. (1972):  
Problems of analytical methods in hydrologic data collection.  
Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resources Systems, Tucson, 1972, vol.2, pp. 451-472.



CHAPITRE II  
THEORIE DES VARIABLES REGIONALISEES  
INFERENCE STATISTIQUE  
ET  
ESTIMATION OPTIMALE

1 - BASES CONCEPTUELLES

a) Notion de variable régionalisée

Le terme "régionalisé" a été proposé par G. MATHERON [5] pour qualifier un phénomène se déployant dans l'espace (et/ou dans le temps), et y manifestant une certaine structure. Une variable caractérisant un tel phénomène est appelée "variable régionalisée" (en abrégé V.R.). En fait, presque toutes les variables descriptives du sous-sol ou de l'atmosphère, donc intéressant le cycle de l'eau, peuvent être considérées comme des V.R.

Du point de vue mathématique, une V.R. est simplement une fonction  $z(x)$  donnant la valeur au point  $x$  (de l'espace à 1, 2 ou 3 dimensions) d'une caractéristique  $z$  du phénomène naturel étudié. Mais, généralement, ces fonctions ont un comportement spatial (ou temporel) bien trop complexe pour se laisser décrire à l'aide d'expressions analytiques classiques.

L'examen de l'évolution de la cote de l'eau dans un piézomètre en nappe libre (Fig.14), ou celui des données de transmissivités recueillies dans une plaine alluviale (Fig.15), montre que l'on pourra rencontrer des phénomènes dont l'aspect structuré ne va pas sans un certain aspect erratique.

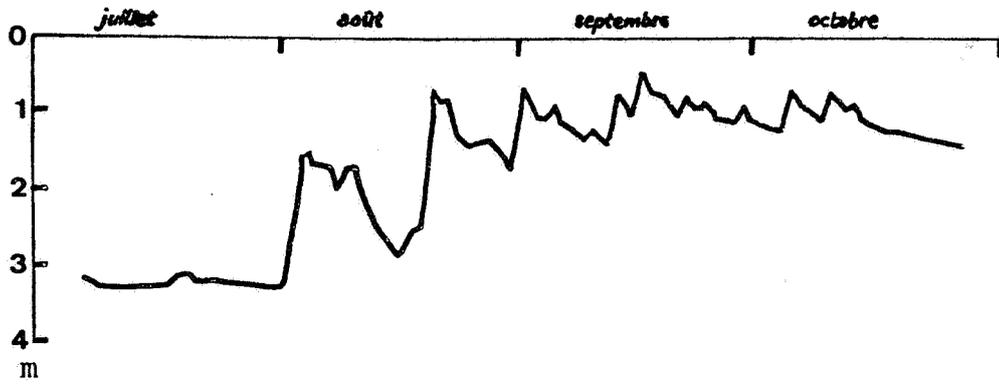


FIG. 14 - EVOLUTION DE LA PROFONDEUR DE LA NAPPE (EN METRES)  
AU PUIT N°3 DE KORHOGO



FIG. 15 - VALEURS DE LA TRANSMISSIVITE (EN  $10^{-2} m^2/s$ ) MESUREES  
DANS LA NAPPE ALLUVIALE DE LA BASSE-DURANCE

b) But de la théorie des variables régionalisées

La théorie des variables régionalisées [7] a pour but

- d'établir les bases théoriques permettant de rendre compte des caractéristiques structurales des phénomènes naturels sous une forme mathématique appropriée,
- de fournir les moyens pratiques de résoudre les divers problèmes d'estimation qui se posent à partir d'un échantillonnage fragmentaire.

Elle couvre l'ensemble des possibilités allant des surfaces (ou courbes) les plus lisses jusqu'aux plus accidentées.

c) Interprétation probabiliste

Une façon commode à la fois sur le plan conceptuel et pratique de traiter une variable régionalisée est d'utiliser la théorie probabiliste des fonctions aléatoires (en abrégé F.A.). On interprète la V.R. comme une "réalisation de fonction aléatoire", c'est à dire comme le résultat d'un tirage au sort dans un ensemble de fonctions [5], [8]. Pour préciser cette idée, on peut supposer, par exemple, que l'on a rangé dans un même groupe un ensemble d'averses analogues, autrement dit un ensemble de fonctions  $z_i(x)$  associant à chaque point  $x$  d'un bassin la hauteur de précipitation en ce point.

Sur l'ensemble  $I = \{i\}$  des évènements et le domaine  $D$  représentant le bassin, on définit la F.A.  $Z$  telle que:

$$\forall i \in I, x \in D \quad Z(x,i) = z_i(x)$$

Le tirage au sort de l'indice  $i$  de l'averse, fixant du même coup les valeurs en tous les points du bassin, donnera ainsi la fonction numérique ordinaire  $z_i(x)$ , dont on dira qu'elle est une réalisation de la F.A.  $Z$ .

On voit que le rapport entre une F.A. et une de ses réalisations est le même que celui qui existe entre une variable aléatoire ordinaire (V.A.) et le résultat numérique d'un tirage au sort selon la loi de probabilité de celle-ci.

Une F.A. est encore, si l'on veut, une variable aléatoire à une infinité de composantes correspondant chacune à un point de l'espace.

Les valeurs prises par le phénomène étudié sont donc considérées comme ayant été obtenues simultanément par un tirage au sort effectué selon la loi de probabilité à une infinité de composantes qui définit une certaine F.A.

d) Problème de l'inférence statistique:

Ce qui précède ne constitue qu'une simple interprétation de la réalité. Pour pouvoir appliquer les résultats de la théorie des F.A., il faut pouvoir reconstituer la loi de la F.A. en question, ou du moins ses premiers moments: c'est ce qu'on appelle l'inférence statistique. Ceci n'est envisageable a priori que si l'on connaît un nombre suffisant de réalisations (averses analogues par exemple). On peut estimer alors les caractéristiques de la F.A. à l'aide de moyennes sur l'ensemble des différents événements. Or, de nombreux phénomènes naturels ont une existence unique. La surface du mur imperméable d'une nappe aquifère ou son champ des transmissivités constituent des cas où la démarche adoptée semble rencontrer un obstacle insurmontable.

On peut toujours considérer le phénomène unique observé comme une réalisation donnée d'une F.A., mais on ne peut pas en général déduire les moments d'une F.A. à partir d'une seule réalisation, pas plus qu'on ne peut calculer espérance, variance ou densité de probabilité d'une V.A. à partir du résultat numérique (par exemple 33) d'une épreuve unique.

L'inférence statistique requiert alors l'introduction d'hypothèses supplémentaires sur la F.A. permettant de lever l'impossibilité, ces hypothèses devant être malgré tout suffisamment faibles pour couvrir le plus grand nombre de problèmes pratiques possible.

2 - INFERENCE STATISTIQUE DANS LE CADRE DE L'HYPOTHESE INTRINSEQUE [7]

a) Stationnarité d'ordre 2

L'hypothèse que l'on formule couramment dans la théorie des F.A. est celle de sa stationnarité.

Une F.A. est dite stationnaire si la loi de probabilité des valeurs prises par cette fonction en  $k$  points arbitraires est invariante par translation d'ensemble de ces points. Autrement dit, un phénomène est appelé stationnaire s'il a une certaine homogénéité spatiale, de nature statistique, qui fait qu'il se répète lui-même en quelque sorte dans l'espace. On admettra en outre la propriété d'ergodicité qui fait que chaque réalisation particulière suffit à rendre compte de l'ensemble des réalisations éventuelles. L'inférence statistique à partir d'une réalisation unique est alors possible, en remplaçant les moyennes, que l'on est normalement amené à faire sur l'ensemble des réalisations disponibles, par des moyennes spatiales sur l'unique réalisation.

En fait, on se limite généralement à faire l'hypothèse de la stationnarité d'ordre 2, dans laquelle on impose seulement aux deux premiers moments de la loi d'être invariants par translation, c'est à dire que:

- l'espérance mathématique soit une constante

$$E [Z(x)] = m(x) = m \quad \text{indépendante de } x$$

- la covariance entre deux points  $x$  et  $y$  ne dépende pas séparément des deux points d'appui, mais seulement du vecteur  $x-y$

$$E [Z(x)-m] [Z(y)-m] = K(x-y)$$

en particulier:  $E [(Z(x)-m)^2] = \text{var} [Z(x)] = K(0)$  indépendante de  $x$

Cette hypothèse de stationnarité d'ordre 2 a été utilisée par l'école soviétique d'hydrométéorologie. On a vu alors qu'elle pouvait être trop restrictive.

Les précipitations sont, par exemple, plus abondantes en altitude qu'en plaine. Dans le cas d'une région à relief varié, leur moyenne présente ainsi une dérive systématique (ou encore, selon la terminologie anglo-saxonne, une "trend"), et ne peut être considérée comme constante. De même, l'existence du gradient général d'écoulement dans un aquifère rend impossible une telle hypothèse.

GANDIN et l'école soviétique ont eu recours à un artifice; travaillant sur les anomalies, ils se ramenaient à un cas où l'hypothèse de stationnarité restait licite. La formulation du krigeage permettra de tenir compte de la présence de telles dérives (krigeage universel).

b) Hypothèse intrinsèque

On supposera pourtant, momentanément, que l'hypothèse d'espérance constante est satisfaite, et ce, pour les besoins de l'exposé.

Cependant, on abandonnera l'hypothèse de stationnarité de la F.A.. Il apparaît en effet que le développement des calculs d'estimation optimale n'exige pas que la F.A. elle-même, mais seulement ses accroissements, soient stationnaires d'ordre 2. On se contentera donc de faire l'hypothèse que, pour tout vecteur  $h$ , l'accroissement  $Z(x+h)-Z(x)$  possède une espérance mathématique et une variance indépendantes du point  $x$ . On supposera ainsi:

$$\begin{cases} E [Z(x+h) - Z(x)] = 0 \\ \text{var} [Z(x+h) - Z(x)] = 2\gamma(h) \end{cases}$$

Cette hypothèse, plus faible que la stationnarité d'ordre 2, a reçu le nom "d'hypothèse intrinsèque". On dira d'une F.A. satisfaisant cette hypothèse que c'est une F.A. intrinsèque. La fonction  $\gamma(h)$  s'appelle le "demi-variogramme" (conformément à une habitude établie, on emploiera par abus de langage "variogramme").

c) Variogramme

α) Définition

Le variogramme d'une fonction aléatoire intrinsèque est ainsi par définition:

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

Comme on a supposé par ailleurs que:

$$E(Z(x+h) - Z(x)) = 0$$

on peut également écrire cette expression sous la forme:

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

ce qui exprime que  $2\gamma(h)$  est l'accroissement quadratique moyen entre deux points distants de  $h$ . Le variogramme est donc comparable à la fonction de structure introduite par DROZDOV.

En supposant satisfaite l'hypothèse intrinsèque, on estimera cette fonction  $\gamma(h)$  à partir des couples de points expérimentaux disponibles sur l'unique réalisation accessible.

β) Estimation du variogramme

Quand le phénomène étudié se développe dans un espace à une dimension (par exemple en fonction du temps), ou quand le phénomène est bidimensionnel, mais a été échantillonné le long d'un profil, l'intervalle  $h$  considéré est un scalaire. Si les points expérimentaux sont alignés et régulièrement espacés, le variogramme peut être calculé, pour des valeurs de  $h$  multiples du pas, à l'aide de la formule:

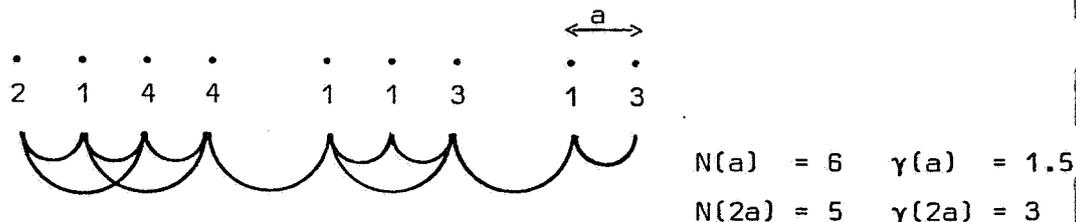
$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z(x_i+h) - z(x_i))^2$$

où  $z(x_i)$  sont les données,

$x_i$  les points pour lesquels les données sont disponibles à la fois en  $x_i$  et  $x_i+h$

et  $N(h)$  le nombre de couples de points distants de  $h$ .

Exemple:



Quand on étudie un phénomène se développant dans le plan,  $h$  est un vecteur de composantes  $(h_1, h_2)$  de même que  $x$  désigne le point de coordonnées  $(x_1, x_2)$ . Le variogramme est alors une fonction des deux variables  $h_1$  et  $h_2$ , ou, en coordonnées polaires, une fonction du module  $|h|$ , distance entre les points, et de l'angle polaire  $\theta$ .

$$\gamma(h) \equiv \gamma(h_1, h_2) \equiv \gamma(|h|, \theta)$$

Lorsque les points expérimentaux sont irrégulièrement espacés dans le plan, on procède à des regroupements par classe de distance et d'angle pour calculer les accroissements quadratiques moyens. C'est à dire que l'on considère, par exemple, tous les couples de points séparés par une distance inférieure à 1 km, comprise entre 1 et 2 km, etc..., en distinguant à chaque fois les couples orientés sensiblement N-S de ceux orientés E-O.

La comparaison des comportements du variogramme dans les différentes directions renseigne sur les anisotropies du phénomène. On essaiera toujours de se ramener au cas isotrope, soit par transformation des coordonnées, soit comme on le verra par la suite, en concentrant l'anisotropie du phénomène dans sa dérive. Le variogramme est alors une fonction du seul module  $|h| = \sqrt{h_1^2 + h_2^2}$

$$\gamma(h) \equiv \gamma(\sqrt{h_1^2 + h_2^2}) \equiv \gamma(|h|)$$

dont on étudiera le comportement du graphe.

γ) Comportement du graphe

Le graphe de cette fonction présente normalement le comportement suivant:

- il passe par l'origine, les écarts étant évidemment nuls pour  $h = 0$ ;

$$\gamma(0) = 0$$

- il croît ensuite, car les valeurs d'un phénomène différent statistiquement de plus en plus (dans le cas général) à mesure qu'augmente la distance entre les points considérés.

Cette croissance, plus ou moins rapide, caractérise ainsi d'une manière statistique la façon dont se détériore l'information apportée par une mesure ponctuelle quand on s'éloigne de ce point connu.

L'examen du comportement du graphe de  $\gamma(h)$  est donc particulièrement important; car le variogramme donne une description synthétique de la structure du phénomène étudié et permet en outre la liaison entre cette structure et la précision avec laquelle pourront être résolus les différents problèmes d'interpolation et d'estimation posés. On étudie en particulier avec soin le comportement au voisinage de l'origine (c'est à dire aux faibles distances) et à l'infini (c'est à dire aux grandes distances).

∴ Comportement à l'origine

Il traduit la continuité de la V.R. Par ordre de régularité décroissante, on distingue classiquement quatre types (Fig.16):

i) Allure parabolique: caractérise une variable extrêmement régulière (dérivable en moyenne quadratique).

ii) Allure linéaire: correspond à une variable moins régulière qu'en i) (continue en moyenne quadratique mais non dérivable).

iii) Discontinuité à l'origine (effet de pépite): signifie que la variable est très irrégulière (non continue en moyenne quadratique). Deux points distincts, mais très proches, présentent encore un écart, de variance au moins égale à l'effet de pépite C.

Le terme "effet de pépite" tire son origine des gisements d'or pépitiques, où la teneur en or varie considérablement selon que le point se trouve à l'intérieur ou à l'extérieur d'une pépite. Le terme a été conservé pour désigner cette discontinuité du variogramme qui peut être due:

- soit à une microrégionalisation d'échelle très inférieure à l'espacement des données,
- soit à des erreurs de mesure,

sans qu'il soit possible de discerner a priori la part respective de ces deux causes.

iv) Aspect plat (effet de pépite pur): correspond au cas aléatoire pur.  $Z(x)$  et  $Z(x+h)$  sont sans corrélation quelque soit leur distance  $h$  non nulle. Ceci est le cas limite de l'absence totale de structure, qui était précisément l'hypothèse faite sur les résidus dans la méthode des moindres carrés.

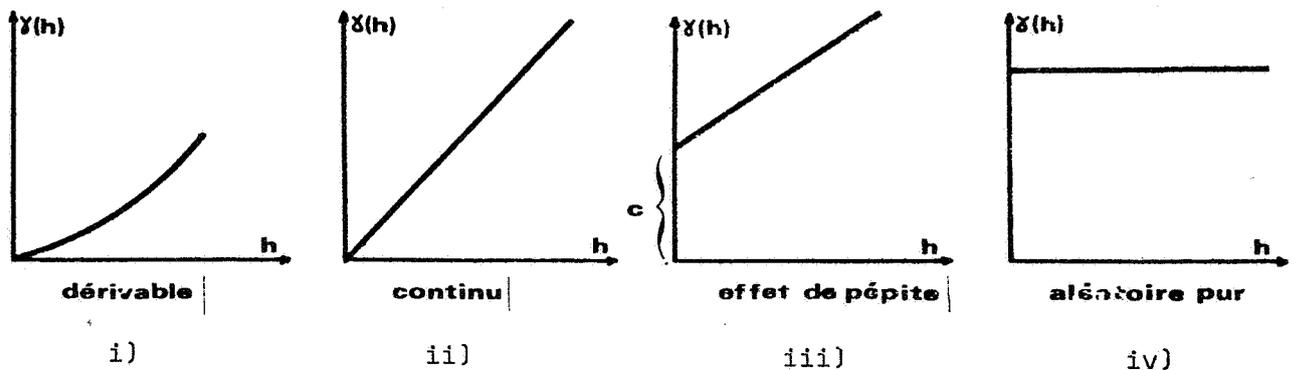
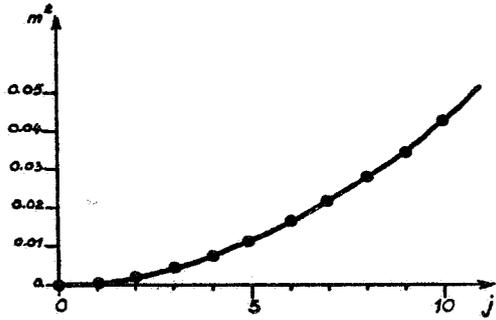


FIG. 16 - COMPORTEMENT DU VARIOGRAMME A L'ORIGINE

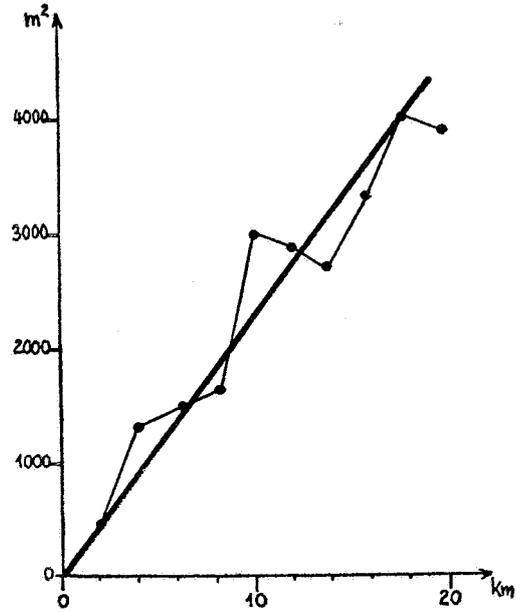
La figure 17 permet de rattacher ces quatre types théoriques à quatre exemples réels obtenus à partir de données sur:

- i) l'évolution journalière de la piézométrie dans un puits donné;
- ii) la variation spatiale de l'épaisseur d'une formation géologique;
- iii) la variation spatiale de la pluviométrie sur un bassin;
- iv) l'évolution des précipitations annuelles à une station.

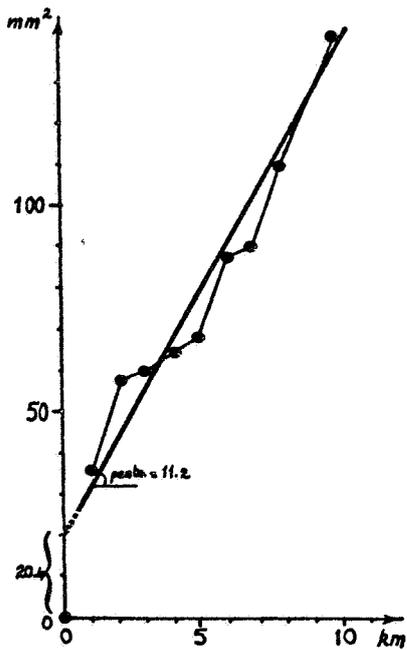
On peut constater qu'à l'extrême régularité des variations piézométriques au puits considérés s'oppose le caractère pratiquement aléatoire de la succession des précipitations annuelles.



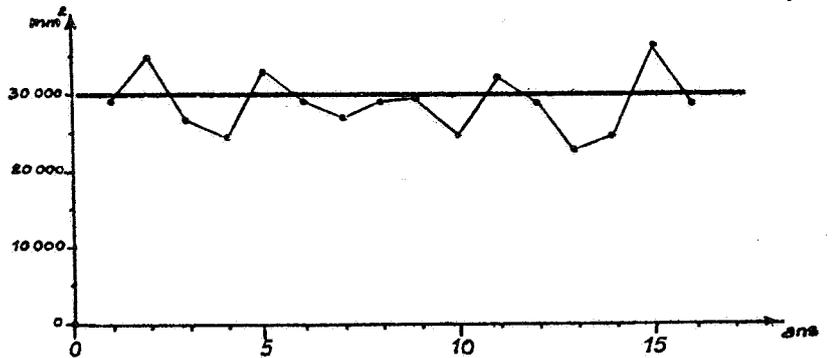
i) Série des niveaux piézométriques journaliers dans un puits (puits n°28, Korhogo, Côte d'Ivoire)



ii) épaisseur d'une formation géologique (puissance d'un réservoir pétrolier, d'après HAAS et JOUSSELIN)



iii) pluviométrie lors d'une averse (B.V. de Kadjemeur, Tchad)



iv) série des précipitations annuelles à une station (Mettmann, R.F.A.)

FIG. 17 - EXEMPLES DE VARIOGRAMMES ILLUSTRANT LES QUATRE TYPES DE COMPORTEMENT A L'ORIGINE

En ce qui concerne les variations spatiales des grandeurs hydrogéologiques ou hydrométéorologiques, si certaines grandeurs varient de façon continue, comme dans l'exemple ii) l'épaisseur d'une formation géologique, nombreuses sont celles présentant une très importante variabilité locale, comme dans l'exemple iii) la pluviométrie, ou encore comme la transmissivité. Dans l'exemple iii), la valeur de l'effet de pépité obtenue par extrapolation est de 20.4 alors que la pente est estimée à 11.2 (les distances étant exprimées en km et la pluviométrie en mm). Cela signifie que, pour une distance de 1 km, près des 2/3 de la variance de l'écart sont pris en compte par l'effet de pépité. On verra toute l'importance de ce fait en ce qui concerne l'estimation.

Outre cet examen du comportement du variogramme au voisinage de l'origine, on étudie également son comportement aux grandes distances.

\* Comportement à l'infini:

Contrairement aux fonctions de covariance (ou de corrélation) qui demeurent nécessairement bornées, le variogramme peut, lui, continuer indéfiniment à croître. Mais il peut aussi se stabiliser autour d'une valeur limite, appelée "palier". La distance à laquelle ce palier est atteint est appelée "portée"; elle renseigne sur l'extension de la zone d'influence d'un point expérimental (Fig.18). Aux distances supérieures à la portée, les corrélations sont nulles.

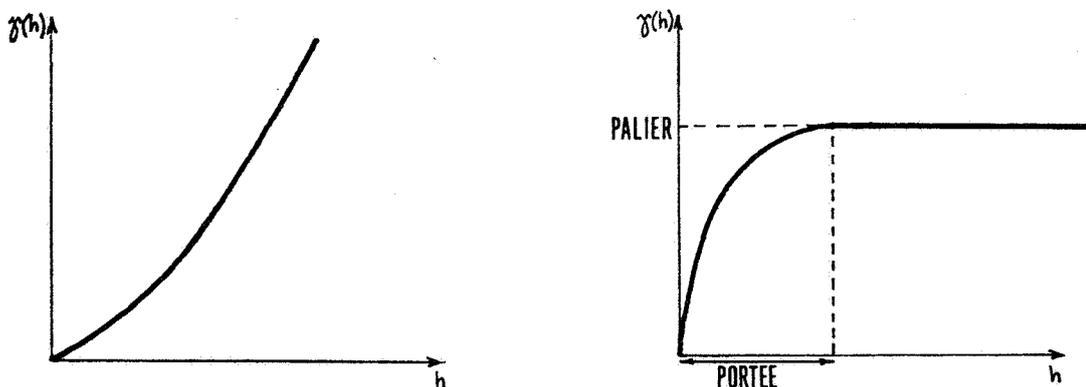


FIG 18 - VARIOGRAMME NON BORNE ET VARIOGRAMME A PALIER

La valeur du palier est égale à la variance des données. En effet, pour une valeur de  $h$  supérieure à la portée, on a :

$$\begin{aligned} \gamma(h) &= \frac{1}{2} \text{var} [Z(x+h) - Z(x)] = \frac{1}{2} (\text{var} [Z(x+h)] + \text{var} [Z(x)]) \\ &= \frac{1}{2} \cdot 2\sigma^2 = \sigma^2 \end{aligned}$$

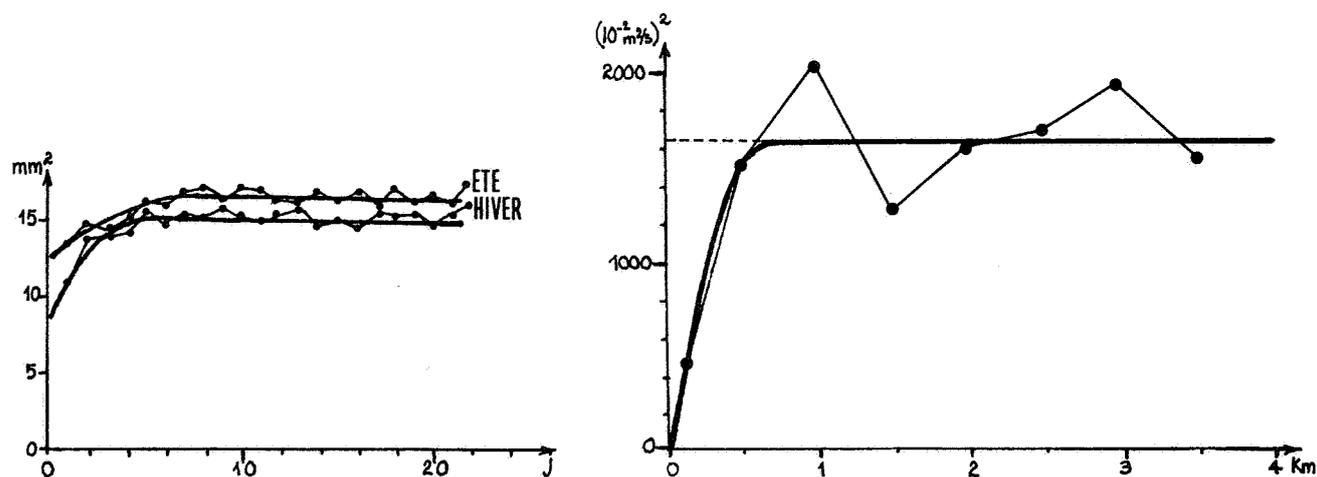
du fait de l'extinction des corrélations au-delà de la portée.

Quand le variogramme est ainsi borné, on peut également montrer qu'il est alors relié à la fonction de covariance  $K(h)$ , définie dans le cas stationnaire d'ordre 2, par :

$$\gamma(h) = K(0) - K(h)$$

La figure 19 montre deux exemples de variogrammes bornés obtenus :

- le premier, à partir de séries chronologiques de précipitations journalières à une station, en été et en hiver. On peut observer les portées et paliers correspondant à des types de précipitation différents.
- le deuxième, à partir de données de transmissivités de la nappe alluviale de la Durance. On peut estimer la portée apparente à 700 m environ, les corrélations ne s'étendant pas très loin dans cette zone où les conditions de dépôt alluvial ont été très tourmentées.



Précipitations journalières sur le bassin de l'Hallue (près d'Amiens)

Transmissivité de la nappe alluviale de la Durance

FIG. 19 - EXEMPLES DE VARIOGRAMMES PRESENTANT UN PALIER

On pourrait songer à pousser plus loin l'interprétation d'un variogramme tiré de données expérimentales, en essayant d'expliquer chacune des fluctuations apparentes du graphe. Si certains phénomènes présentent effectivement des périodicités, des structures-gigognes (superposition de variations d'échelles différentes), ou des effets de trou (existence de corrélations négatives dues à la présence de zones "riches" entourées de zones "pauvres") (Fig.20), la plupart du temps, ces fluctuations sont simplement dues à ce que les différentes valeurs de  $\gamma(h)$  calculées ne sont en fait que des estimations, obtenues parfois avec un petit nombre de couples.

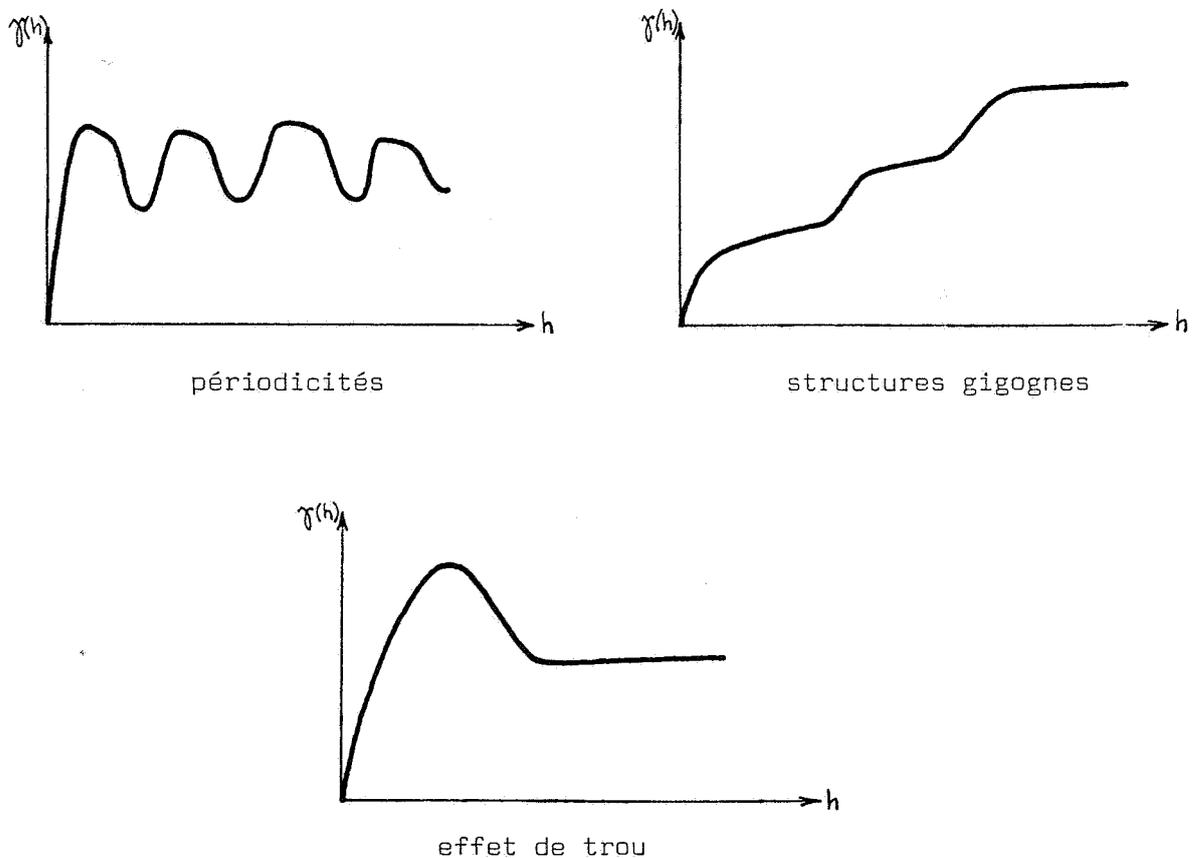


FIG. 20 - COMPORTEMENTS PARTICULIERS DU VARIOGRAMME

De là la pratique courante d'ajuster aux variogrammes expérimentaux des modèles théoriques simples, qui seuls, d'ailleurs, garantissent la cohérence mathématique nécessaire pour les calculs ultérieurs.

δ) Ajustement à un modèle

Le modèle utilisé devra garantir que les calculs de variance de combinaisons linéaires du type :

$$Y = \sum_{i=1}^n \lambda_i Z(x_i)$$

qui seront employées par la suite donnent bien toujours un résultat positif.

Pour tout ensemble de  $n$  points  $x_i$  et de  $n$  coefficients  $\lambda_i$  arbitraires vérifiant  $\sum_{i=1}^n \lambda_i = 0$  (de façon à se ramener à une combinaison linéaire d'accroissements, sur lesquels seuls on peut travailler dans le cadre de l'hypothèse intrinsèque), la fonction  $\gamma(h)$  doit être telle que :

$$\text{var}(Y) = - \sum_{i=1}^n \sum_{j=1}^n \gamma(x_i - x_j) \geq 0$$

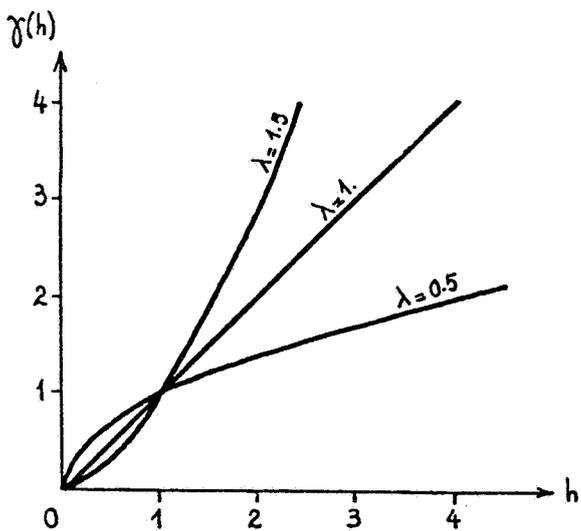
Les modèles de variogramme les plus fréquemment utilisés sont les modèles en  $h^\lambda$  et les modèles sphérique, exponentiel et gaussien, représentés Fig. 21 ci-après.

On peut également essayer d'ajuster une combinaison linéaire (à coefficients positifs) de ces différents modèles au variogramme obtenu expérimentalement.

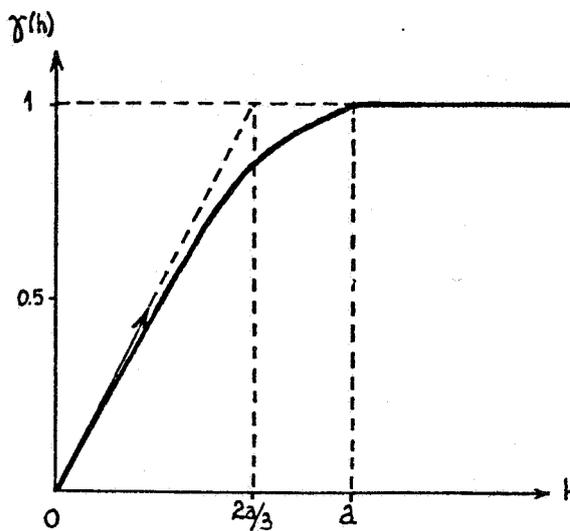
La présence d'un effet de pépite  $C$  se traduit par l'addition de la quantité  $C$  à un modèle classique  $\gamma_0$  en dehors de l'origine :

$$\gamma(h) = \begin{cases} 0 & \text{pour } h = 0 \\ \gamma_0(h) + C & \text{pour } h \neq 0 \end{cases}$$

La valeur de  $C$  est déterminée par extrapolation vers  $h=0$  sur la courbe expérimentale (cf. Fig. 17, iii), p. 49).



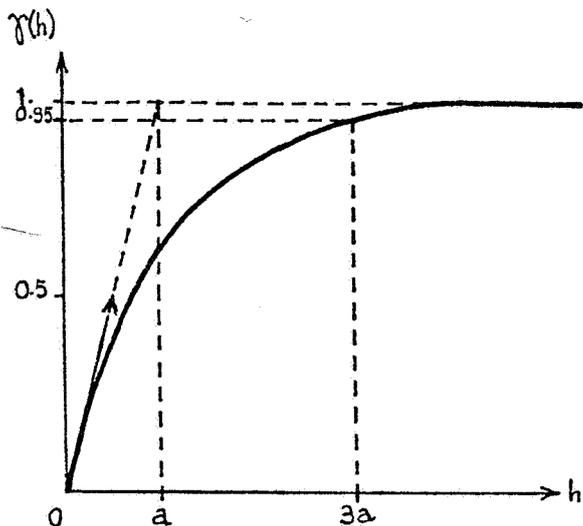
modèle en  $h^\lambda$ :  $\gamma(h) = \omega|h|^\lambda$



modèle sphérique:

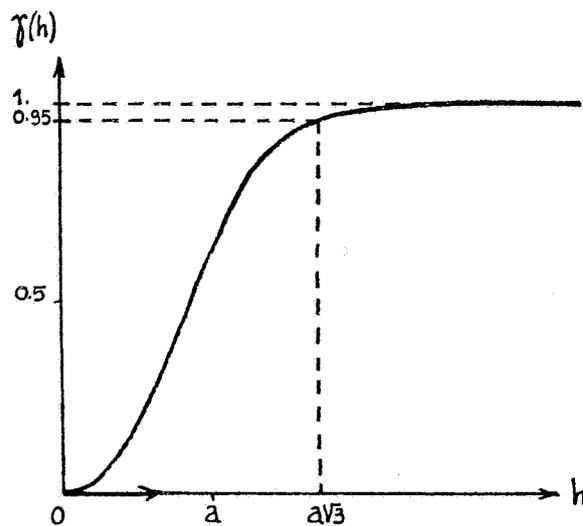
$$\gamma(h) = \omega \left[ \frac{3}{2} \frac{|h|}{a} - \frac{1}{2} \left( \frac{|h|}{a} \right)^3 \right] \quad h \leq a$$

$$\gamma(h) = \omega \quad h > a$$



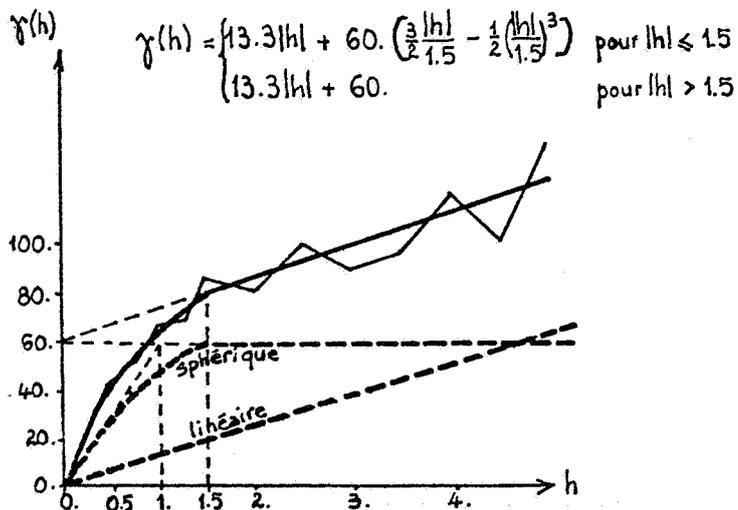
modèle exponentiel:

$$\gamma(h) = \omega \left[ 1 - e^{-\frac{|h|}{a}} \right]$$



modèle gaussien:

$$\gamma(h) = \omega \left[ 1 - e^{-\left(\frac{|h|}{a}\right)^2} \right]$$



ajustement à un modèle "linéaire + sphérique"

FIG. 21 - PRINCIPAUX MODELES DE VARIOGRAMME UTILISES ET EXEMPLE D'AJUSTEMENT

On écrit encore:  $\gamma(h) = C(1-\delta) + \gamma_0(h)$   
le symbole  $\delta$  représentant la mesure de Dirac.

Dans certains cas, assez rares, on a la certitude que le phénomène est continu et qu'il n'y a pas d'erreurs de mesure, mais on pense que le pas d'échantillonnage a été trop lâche pour mettre en évidence une croissance rapide du variogramme correspondant à une microrégionalisation. On pourrait alors remplacer l'effet de pépite  $C$  par une composante sphérique de palier  $C$  et de portée inférieure au pas d'échantillonnage  $\ell$  (Fig. 22).

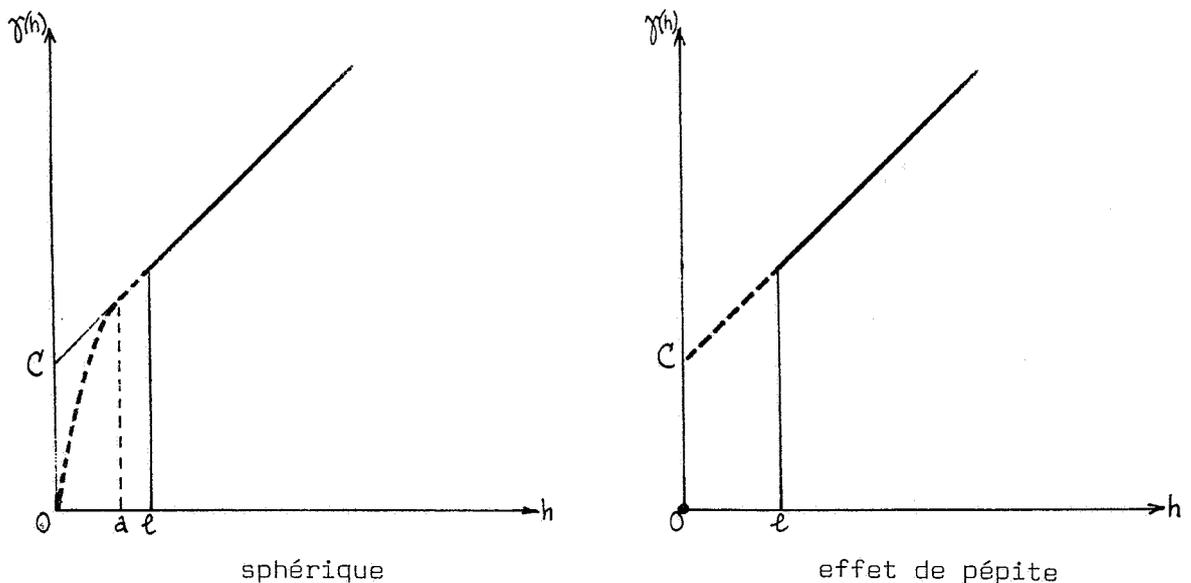


FIG. 22 - MODELISATION AUX DISTANCES INFÉRIEURES AU PAS D'ECHANTILLONNAGE

L'estimation du variogramme, l'examen du comportement de son graphe et son ajustement à un modèle ont permis de se doter d'un outil capable de caractériser la variabilité spatiale d'un phénomène. Le premier but de la théorie des variables régionalisées est ainsi atteint. Il reste à faire le lien avec les problèmes d'estimation: c'est ce que réalise la théorie du krigeage.

3 - KRIGEAGE

a) Exemple introductif

Considérant un phénomène dont on étudie les variations dans le temps (comme la température de l'eau ou son PH en un point d'une rivière), on peut se poser le problème simple suivant:

Obtient-on une estimation plus précise de sa valeur moyenne sur les intervalles de temps séparant deux heures justes, en faisant les mesures aux demies et en retenant ces valeurs, ou en faisant les mesures aux heures justes et en adoptant la demi-somme de deux mesures consécutives ?

Soient  $\sigma_A^2$  la variance de l'erreur commise en retenant la valeur de la mesure prise à la demie,  
et  $\sigma_B^2$  celle correspondant à la demi-somme des deux mesures prises aux heures justes.

On montre que  $\sigma_A^2$  et  $\sigma_B^2$  peuvent s'exprimer en fonction du variogramme. Supposons que celui-ci soit de la forme  $\gamma(h) = h^\lambda$  ( $0 \leq \lambda < 2$ ): on constate alors que  $\sigma_A^2 > \sigma_B^2$  pour  $\lambda > 1$ , et inversement, l'égalité étant réalisée pour  $\lambda = 1$ .

La question posée n'a ainsi de réponse que si l'on a caractérisé auparavant la structure, ici temporelle, du phénomène étudié.

Un variogramme en  $h^\lambda$  avec  $\lambda > 1$  correspond à un phénomène très régulier. Une mesure unique, mais bien placée au milieu de l'intervalle, est préférable à deux mesures mal placées par rapport à l'intervalle étudié.

Un variogramme en  $h^\lambda$  avec  $\lambda < 1$  correspond à un phénomène beaucoup moins régulier et, comme dans le cas aléatoire pur atteint pour  $\lambda = 0$ , le facteur déterminant est le nombre de mesures prises en compte, quelle que soit leur localisation.

Cet exemple a mis en évidence le lien étroit entre la structure du phénomène, caractérisée par le variogramme, et la précision d'une estimation. C'est autour de ce lien que va maintenant s'articuler le krigeage.

b) Cas sans dérive: krigeage simple [7]

α) Position du problème

On rappelle la situation de départ: les valeurs de la variable  $z(x)$  sont connues en  $n$  points expérimentaux  $x_1, x_2, \dots, x_n$ , et l'on désire estimer une quantité  $z_0$  qui peut être, en fait, toute fonctionnelle linéaire de la variable. On se limitera ici aux trois cas suivants qui feront l'objet d'applications au Chapitre III:

- i)  $z_0 = z(x_0)$  valeur prise par  $z$  au point  $x = x_0$
- ii)  $z_0 = \frac{1}{s} \int_{s_0} z(x) dx$  valeur moyenne de  $z(x)$  sur la maille  $s_0$  de surface  $s$ , centrée au point  $x = x_0$ .
- iii)  $z_0 = \frac{1}{S} \int_S z(x) ds$  valeur moyenne de  $z(x)$  sur le domaine  $S$ .

Ces trois objectifs correspondent aux trois besoins définis précédemment de la cartographie, de l'initialisation de modèles numériques et du calcul des lames d'eau sur un bassin versant.

De façon à simplifier l'exposé, les calculs seront développés pour le cas iii), et les résultats seront ensuite repris pour le cas i) qui ne constitue en fait que le cas limite où le domaine se réduit à un point.

β) Estimation zonale

Pour estimer  $z_0 = \frac{1}{S} \int_S z(x) dx$ , on adopte une moyenne pondérée des  $n$  données disponibles:

$$z_0^* = \sum_{i=1}^n \lambda_i z(x_i)$$

Le problème est donc de trouver les poids  $\lambda_i$  qui donneront la meilleure estimation possible.

Pour cela, on va faire appel à l'interprétation probabiliste que l'on a donnée de la V.R. étudiée, et raisonner sur la F.A. intrinsèque  $Z(x)$  dont  $z(x)$  est censée être une représentation.

On considère ainsi l'estimateur:

$$Z_0^* = \sum_{i=1}^n \lambda_i Z(x_i)$$

(On rappelle que les majuscules caractérisent, par convention dans ce texte, les grandeurs aléatoires).

On cherchera les poids  $\lambda_i$  conduisant à un estimateur:

- sans biais  $E [Z_0^* - Z_0] = 0$
- d'erreur quadratique moyenne minimale  $E [(Z_0^* - Z_0)^2]$  minimum

Comme l'on a:  $\text{var} [Z_0^* - Z_0] = E [(Z_0^* - Z_0)^2] - (E [Z_0^* - Z_0])^2$

ces conditions peuvent encore s'écrire:

$$\begin{cases} E [Z_0^* - Z_0] = 0 \\ \text{var} [Z_0^* - Z_0] \text{ minimum} \end{cases}$$

Comme l'on a conservé l'hypothèse d'espérance constante, c'est à dire:

$$E [Z(x)] = m, \quad \forall x$$

on a en particulier:  $E [Z(x_i)] = m \quad i = 1, \dots, n$

et en ce qui concerne  $E [Z_0] = E \left[ \frac{1}{S} \int_S Z(x) dx \right]$

( $Z_0$ , transposition probabiliste de  $z_0$ , est ici une intégrale stochastique)  
on a, en utilisant la permutabilité des opérations d'intégration et d'espérance mathématique:

$$E [Z_0] = \frac{1}{S} \int_S E [Z(x)] dx = \frac{1}{S} \int_S m dx = m$$

La première condition (non-biais) devient:

$$\begin{aligned} E \left[ \sum_{i=1}^n \lambda_i Z(x_i) - Z_0 \right] &= \sum_{i=1}^n \lambda_i E [Z(x_i)] - E [Z_0] \\ &= \sum_{i=1}^n \lambda_i m - m = m \left( \sum_{i=1}^n \lambda_i - 1 \right) = 0 \end{aligned}$$

Pour que cette égalité soit vérifiée quelle que soit la valeur (inconnue) de  $m$ , il faut donc que:

$$\sum_{i=1}^n \lambda_i - 1 = 0$$

Cette égalité permet d'écrire l'erreur  $Z_0^* - Z_0$  sous la forme d'une combinaison linéaire d'accroissement de la F.A.  $Z$ :

$$Z_0^* - Z_0 = \sum_{i=1}^n \lambda_i Z(x_i) - Z_0 = \sum_{i=1}^n \lambda_i (Z(x_i) - Z_0)$$

sur laquelle on pourra donc travailler dans le cadre de l'hypothèse intrinsèque.

La variance de cette erreur peut s'explicitier après quelques calculs à l'aide du variogramme:

$$\begin{aligned} \text{var} [Z_0^* - Z_0] &= \text{var} \left[ \sum_{i=1}^n \lambda_i Z(x_i) - Z_0 \right] \\ &= - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma_{ij} + 2 \sum_{i=1}^n \lambda_i \gamma_{iS} - \gamma_{SS} \end{aligned}$$

où  $\gamma_{ij} = \gamma(x_i - x_j)$

est la valeur du variogramme entre les deux points expérimentaux  $x_i$  et  $x_j$ ;

$$\gamma_{iS} = \frac{1}{S} \int_S \gamma(x_i - x) dx$$

est la valeur moyenne du variogramme entre le point  $x_i$  et un point décrivant le domaine  $S$ ;

$$\gamma_{SS} = \frac{1}{S^2} \int_S \int_S \gamma(x - x') dx dx'$$

est la valeur moyenne du variogramme entre deux points décrivant indépendamment le domaine  $S$ .

La minimisation de la forme quadratique en  $\lambda_i$ :

$$Q = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma_{ij} + 2 \sum_{i=1}^n \lambda_i \gamma_{iS} - \gamma_{SS}$$

sous la contrainte  $E = \sum_{i=1}^n \lambda_i - 1 = 0$  se fait en introduisant un multiplicateur de Lagrange  $\mu'$  et en annulant les dérivées partielles, par rapport aux différents  $\lambda_i$ , de la quantité  $Q + \mu' E$ .

Les calculs conduisent, en posant  $\mu = -\frac{\mu'}{2}$  au système linéaire suivant, appelé "système de krigeage simple":

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu = \gamma_{iS} & i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases}$$

C'est un système de  $n+1$  équations à  $n+1$  inconnues (les  $n$  poids  $\lambda_i$  et le multiplicateur de Lagrange  $\mu$ ) qui s'écrit encore matriciellement:

$$\underbrace{\begin{pmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \dots & \gamma_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix}}_{\Gamma} \times \underbrace{\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix}}_{\Lambda} = \underbrace{\begin{pmatrix} \gamma_{1S} \\ \gamma_{2S} \\ \vdots \\ \gamma_{nS} \\ 1 \end{pmatrix}}_{B}$$

en remarquant que  $\gamma_{ii} = \gamma(0) = 0$

La solution s'écrit alors:  $\Lambda = \Gamma^{-1} B$  à l'aide de la notation matricielle.

Les poids  $\lambda_i$  ainsi trouvés sont ensuite reportés dans l'expression

$$z_0^* = \sum_{i=1}^n \lambda_i z(x_i)$$

pour obtenir la valeur estimée par krigeage (ou valeur krigée).

A l'optimum, en reportant  $\sum_{j=1}^n \lambda_j \gamma_{ij} = \gamma_{iS} - \mu$  dans l'expression donnant la variance de l'erreur, on obtient:

$$\sigma_K^2 = \text{var} [Z_0^* - Z_0] = \sum_{i=1}^n \lambda_i \gamma_{iS} + \mu - \gamma_{SS}$$

$\sigma_K^2$  est la variance d'estimation par krigeage (ou variance de krigeage).

Matriciellement:  $\sigma_K^2 = \lambda \cdot B - \gamma_{SS}$  ( $\lambda$  = transposée de)

### $\gamma$ ) Estimation ponctuelle

Pour le cas de l'estimation ponctuelle,

$$\gamma_{iS} \text{ est remplacé par } \gamma_{io} = \gamma(x_i - x_o)$$

$$\gamma_{SS} \text{ est remplacé par } \gamma_{oo} = \gamma(x_o - x_o) = \gamma(0) = 0$$

Le système de krigeage s'écrit alors:

$$\left\{ \begin{array}{l} \sum_{j=1}^n \lambda_j \gamma_{ij} + \mu = \gamma_{io} \\ \sum_{j=1}^n \lambda_j = 1 \end{array} \right. \quad i=1, \dots, n$$

et la variance de krigeage:  $\sigma_K^2 = \sum_{i=1}^n \lambda_i \gamma_{io} + \mu$

Pour l'estimation ponctuelle, comme d'ailleurs pour celle de moyenne par maille, il y a autant de systèmes de krigeage que de points (ou de mailles) à estimer. Si l'on prend en compte à chaque fois la totalité des données disponibles - et on parlera alors de "voisinage unique" - on peut remarquer que seul change en fait à chaque fois le second membre. On pourra alors tirer parti de ce fait en inversant une fois pour toute la matrice  $\Gamma$ .

Pourtant, généralement, le nombre de points expérimentaux est trop important pour pouvoir procéder ainsi sans allonger par trop le temps de calcul et compromettre la précision des résultats. On reformera alors à chaque fois le système de krigeage, en ne retenant que les données (entre 10 et 20) situées au voisinage du point (ou de la maille) à estimer. On parlera alors de "voisinage glissant".

c) Cas avec dérive: krigeage universel [6], [7]

Les problèmes d'estimation posés sont ainsi résolus dans le cas où l'hypothèse intrinsèque est compatible avec la réalité du phénomène étudié. Mais il existe de nombreuses situations où l'hypothèse de moyenne constante, que nous avons formulée jusqu'ici, ne peut être raisonnablement maintenue. On a déjà cité, par exemple, le cas de la piézométrie d'une nappe ou celui de précipitations liées à l'orographie. On doit alors recourir à un modèle plus élaboré pour pouvoir prendre en compte les tendances en grand de tels phénomènes.

α) Modèle interprétatif

On supposera maintenant que l'espérance mathématique de la F.A.  $Z(x)$  (dont le phénomène est censé être une réalisation) est une fonction

$$m(x) = E[Z(x)]$$

qui varie d'une manière lente et régulière relativement à l'échelle à laquelle on travaille, et à laquelle on donnera le nom de "dérive".

Cette fonction  $m(x)$  sera suffisamment régulière pour pouvoir être représentée localement par une expression de la forme:

$$m(x) = \sum_{l=1}^k a_l f^l(x)$$

où les  $f^l$  sont des fonctions données, en pratique presque toujours des monômes, et les  $a_l$  des coefficients évidemment inconnus.

Une telle formulation englobe d'ailleurs le cas précédent où la moyenne était constante. La dérive  $m(x)$  se réduit alors à:

$$m(x) = a_1 f^1(x) = a_1$$

$f^1(x)$  étant la fonction identiquement égale à 1.

Explicitons l'expression de la dérive dans le cas d'un phénomène se développant dans le plan. Si  $X$  et  $Y$  désignent les coordonnées d'un point  $x$  du plan, on pourra avoir par exemple:

$$m(X,Y) = a_1 + a_2X + a_3Y$$

et on parlera alors de dérive linéaire ou:

$$m(X,Y) = a_1 + a_2X + a_3Y + a_4X^2 + a_5XY + a_6Y^2$$

et on parlera alors de dérive quadratique.

On conservera la seconde hypothèse qui avait été précédemment formulée, à savoir:  $\text{var} [Z(x+h)-Z(x)] = 2\gamma(h)$ . Sans préjuger de l'estimation de  $\gamma(h)$  dans le cas présent où  $E[Z(x+h)-Z(x)] \neq 0$ , problème qui sera abordé au paragraphe 4, on peut développer à nouveau les calculs qui avaient conduit au système de krigeage.

### $\beta$ ) Calculs

L'estimateur de krigeage est toujours de la forme:

$$Z_0^* = \sum_{i=1}^n \lambda_i Z(x_i)$$

mais on a maintenant:

$$E [Z(x_i)] = m(x_i) = \sum_{l=1}^k a_l f_l^1(x_i)$$

$$\begin{aligned} E [Z_0] &= \frac{1}{S} \int_S E [Z(x)] dx = \frac{1}{S} \int_S m(x) dx \\ &= \frac{1}{S} \int_S \sum_{l=1}^k a_l f_l^1(x) dx = \sum_{l=1}^k a_l \frac{1}{S} \int_S f_l^1(x) dx \end{aligned}$$

C'est à dire, en posant  $f_i^1 = f^1(x_i)$  et  $f_S^1 = \frac{1}{S} \int_S f^1(x) dx$ :

$$E [Z(x_i)] = \sum_{l=1}^k a_l f_i^1 \quad \text{et} \quad E [Z_0] = \sum_{l=1}^k a_l f_S^1$$

La condition de non-biais s'écrit donc maintenant:

$$E [Z_o^* - Z_o] = E \left[ \sum_{i=1}^n \lambda_i Z(x_i) - Z_o \right]$$

$$= \sum_{i=1}^n \lambda_i \sum_{l=1}^k a_{il} f_i^l - \sum_{l=1}^k a_{1l} f_S^l = \sum_{l=1}^k a_{1l} \left( \sum_{i=1}^n \lambda_i f_i^l - f_S^l \right) = 0$$

Cette égalité devant être réalisée, quelles que soient les valeurs inconnues (et que l'on n'a d'ailleurs pas besoin d'estimer) des  $a_{1l}$ , il faut poser:

$$\sum_{i=1}^n \lambda_i f_i^l - f_S^l = 0 \quad l = 1, \dots, k$$

En prenant  $f^1 \equiv 1$ , on retrouve comme première équation:

$$\sum_{i=1}^n \lambda_i - 1 = 0$$

qui permet encore de travailler sur des combinaisons linéaires d'accroissements.

Le reste des calculs est identique, si ce n'est que la minimisation de la forme quadratique doit se faire maintenant sous les  $k$  contraintes exprimant la condition de non-biais. On introduit donc  $k$  multiplicateurs de Lagrange,  $\mu_1, \dots, \mu_k$ , et l'on aboutit au système suivant, à  $n+k$  équations et  $n+k$  inconnues:

$$\left\{ \begin{array}{l} \sum_{j=1}^n \lambda_j \gamma_{ij} + \sum_{l=1}^k \mu_l f_i^l = \gamma_{iS} \quad i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j f_j^l = f_S^l \quad l = 1, \dots, k \end{array} \right.$$

La variance du krigeage s'écrit de même:

$$\sigma_K^2 = \sum_{i=1}^n \lambda_i \gamma_{iS} + \sum_{l=1}^k \mu_l f_S^l - \gamma_{SS}$$

On passe au cas ponctuel de la même façon que précédemment. Le système de krigeage universel peut alors se mettre par exemple sous la forme matricielle suivante:

$$\begin{pmatrix} 0 & \gamma_{12} & \dots & \gamma_{1n} & 1 & f_1^2 & \dots & f_1^k \\ \gamma_{21} & 0 & \dots & \gamma_{2n} & 1 & f_2^2 & \dots & f_2^k \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & 0 & 1 & f_n^2 & \dots & f_n^k \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ f_1^2 & f_2^2 & \dots & f_n^2 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_1^k & f_2^k & \dots & f_n^k & 0 & 0 & \dots & 0 \end{pmatrix} \times \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ 1 \\ f_0^2 \\ \vdots \\ f_0^k \end{pmatrix}$$

en posant  $f^1(x_0) = f_0^1$

d) Propriétés du krigeage:

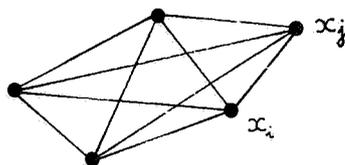
On peut énoncer maintenant quelques propriétés du krigeage valables dans les deux cas avec et sans dérive.

α) On peut remarquer que le système de krigeage prend bien en considération les éléments suivants:

- position relative du domaine (ou du point) à estimer et des points expérimentaux:



- position relative des points expérimentaux entre eux:



par l'intermédiaire des  $\gamma_{iS}$  (ou  $\gamma_{i0}$ ) et des  $\gamma_{ij}$ ,

- structure du phénomène, à travers le variogramme  $\gamma$ .

On verra, au Chapitre III, § 1, l'influence de ces trois éléments, et surtout du troisième, sur la pondération opérée par le krigeage.

β) Lorsque le point  $x_0$  coïncide avec un point expérimental  $x_i$ , on vérifie que la solution du système est  $\lambda_i = 1$ ,  $\lambda_j = 0$  pour  $j \neq i$ . On a donc bien:  $z_0 = z(x_0)$  (le krigeage est un interpolateur exact), et on a également alors  $\sigma_K^2 = 0$ .

γ) Le système de krigeage, et donc la variance de krigeage  $\sigma_K^2$ , ne font intervenir que la structure et la configuration géométrique des données et du domaine à estimer. Elle ne dépend pas des valeurs expérimentales  $z(x_i)$ . Cette propriété remarquable sera utilisée au Chapitre III, § 2, pour l'optimisation des réseaux de mesure.

#### 4 - RETOUR A L'INFERENCE STATISTIQUE: LES F.A.I.-7:

##### a) Variogramme brut et variogramme sous-jacent

On dispose donc maintenant du moyen d'estimer en tout point un phénomène présentant ou non une dérive. Mais, dans le cas où existe une dérive, un problème est resté en suspens: celui de l'estimation du variogramme  $\gamma(h)$ . En effet, la méthode proposée pour son estimation supposait en fait:  $E [Z(x+h) - Z(x)] = 0$ , c'est à dire une moyenne constante.

Que se passe-t-il si l'on essaie d'appliquer brutalement le procédé décrit alors à une variable manifestant une tendance en grand ?

Si l'on examine le variogramme expérimental (on dit encore "variogramme brut") obtenu ainsi à partir des données de piézométrie de la nappe de la Crau (Fig.23), on observe un comportement anisotrope marqué aux grandes distances, avec une croissance très rapide dans la direction NE-SO. La grande amplitude des écarts dans cette direction, par rapport aux écarts NO-SE, correspond au sens général de la nappe, qui s'écoule du NE vers le SO avec un gradient hydraulique moyen compris entre 3 et 4 m/km.

En pratique, on rencontre ainsi souvent des variogrammes bruts, dont la croissance aux grandes distances semble parabolique (ou encore plus rapide), ce qui est contraire à une propriété de  $\gamma(h)$  imposée par des

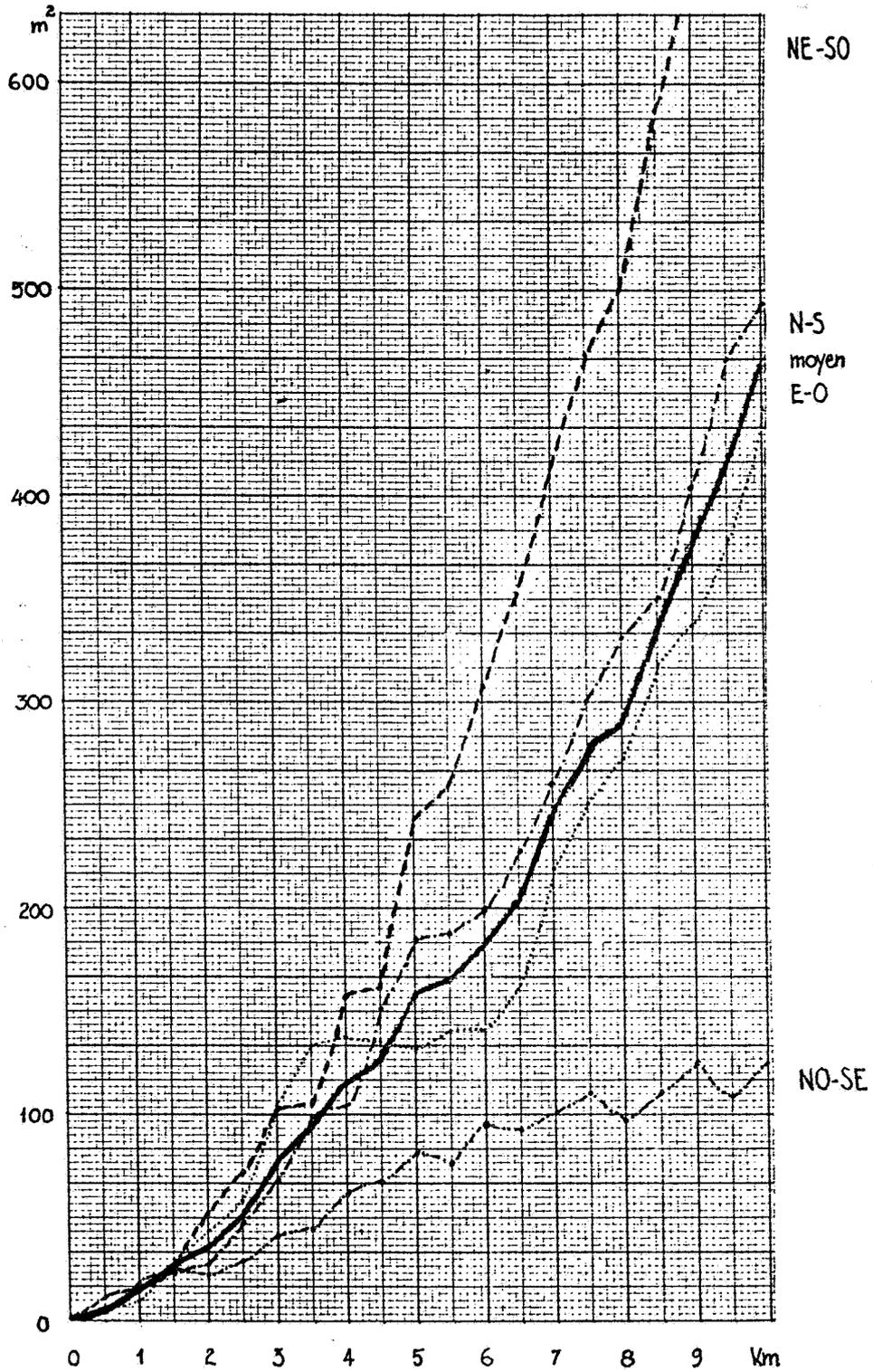


FIG. 23 - NAPPE DE LA CRAU: VARIOGRAMME BRUT

considérations de cohérence mathématique:

$$\frac{\gamma(h)}{|h|^2} \rightarrow 0 \quad \text{quand} \quad |h| \rightarrow \infty$$

En fait, ces comportements apparents viennent de ce que l'on a négligé l'existence d'une dérive.

Dans l'égalité  $2\gamma(h) = \text{var} [Z(x+h)-Z(x)]$   
 $= E [(Z(x+h)-Z(x))^2] - (E[Z(x+h)-Z(x)])^2$

le deuxième terme de la différence n'est pas nul, et en calculant le variogramme brut, on obtient en fait:

$$\begin{aligned} \gamma_{\text{exp}}(h) &= \gamma_{\text{réel}}(h) + \frac{1}{2} (E[Z(x+h)-Z(x)])^2 \\ &= \gamma_{\text{réel}}(h) + \frac{1}{2} (m(x+h)-m(x))^2 \end{aligned}$$

Dans le cas d'une dérive linéaire, il se superpose ainsi une parabole au variogramme réel.

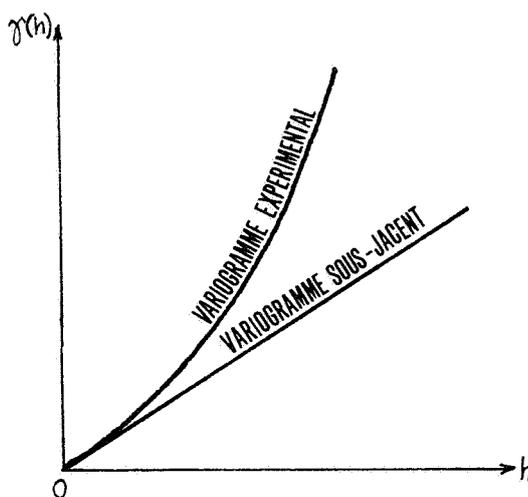


FIG. 24 - VARIOGRAMME EXPERIMENTAL EN PRESENCE  
D'UNE DERIVE ET VARIOGRAMME SOUS-JACENT

Dans certains cas simples où la dérive n'est pas apparente dans une direction, on pourra estimer directement le variogramme réel en ne considérant que des couples de points orientés selon cette direction privilégiée et, ayant ajusté celui-ci à un modèle théorique, procéder au krigeage.

Mais en dehors de ces cas simples, on ne peut accéder directement au variogramme réel (appelé variogramme "sous-jacent").

b) Variogramme des résidus

On peut songer alors à estimer, dans un premier temps, la valeur de la dérive  $m(x)$  aux différents points expérimentaux, et calculer ensuite le variogramme à partir des résidus estimés  $z(x_i) - m^*(x_i)$ .

Malheureusement, le variogramme obtenu constitue une estimation profondément biaisée du variogramme réel. Dans le cas, par exemple, d'une dérive linéaire et d'un variogramme sous-jacent linéaire, le variogramme des résidus aura une forme parabolique (Fig. 25).

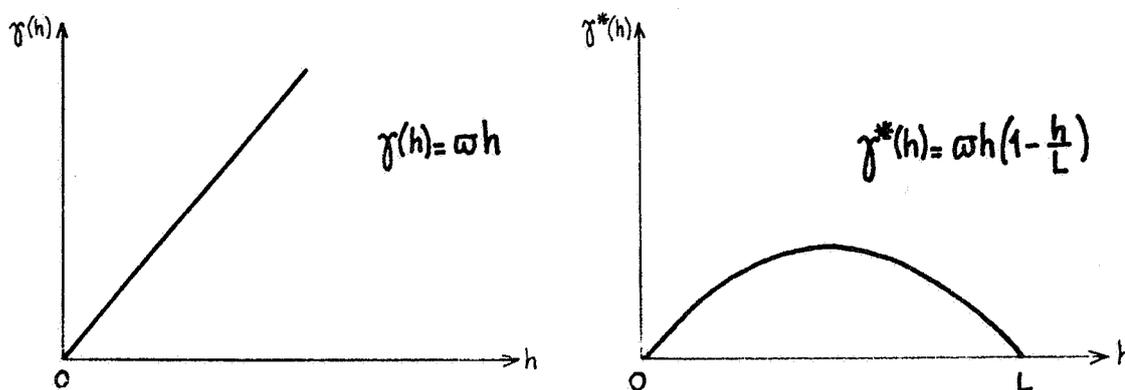


FIG. 25 - VARIOGRAMME REEL ET VARIOGRAMME DES RESIDUS

Il a été pourtant possible de mettre au point des méthodes d'identification du variogramme sous-jacent à partir de ces variogrammes des résidus [3], mais leur lourdeur, et surtout l'impossibilité de les automatiser, les ont fait pratiquement abandonner à l'heure actuelle.

c) Hypothèse intrinsèque généralisée [9], [10]:

Il est préférable en effet de repartir de la définition même du variogramme, qui est une variance d'accroissement (ou, autrement dit, de différence première).

L'intérêt des différences premières est de filtrer les constantes. Quand  $E[Z(x)]$  est une constante inconnue  $m$ , les accroissements  $Z(x+h) - Z(x)$  ont

une espérance nulle, quelle que soit cette valeur de  $m$ , et le variogramme peut alors s'estimer sans biais à partir des écarts quadratiques expérimentaux  $(z(x+h)-z(x))^2$ .

L'idée est alors, pour les phénomènes auxquels on ne peut appliquer l'hypothèse intrinsèque, de faire intervenir des différences secondes, tierces,.... qui ont la propriété de filtrer les polynômes de degré 1, 2, etc... Ce genre de procédé est d'un emploi courant dans l'étude des séries chronologiques non-stationnaires où l'on calcule des différences finies successives jusqu'à obtenir un résultat à peu près stationnaire. On généralise la méthode au cas de données réparties irrégulièrement dans le plan.

α) Généralisation de la notion d'accroissement [2]

Pour des données régulièrement espacées sur la droite, une différence première s'écrit par exemple:  $Z(x+a)-Z(x)$ . Les poids  $\lambda_1 = 1$  et  $\lambda_2 = -1$ , affectés aux données d'abscisses  $x_1 = x+a$  et  $x_2 = x$ , vérifient:

$$\sum_{i=1}^2 \lambda_i = 1 + (-1) = 0$$

Une différence seconde s'écrira:  $Z(x+a) - 2Z(x) + Z(x-a)$ . Les poids  $\lambda_1 = 1$ ,  $\lambda_2 = -2$  et  $\lambda_3 = 1$ , affectés aux données d'abscisses  $x_1 = x+a$ ,  $x_2 = x$  et  $x_3 = x-a$  vérifient:

$$\sum_{i=1}^3 \lambda_i = 1 + (-2) + 1 = 0$$
$$\sum_{i=1}^3 \lambda_i x_i = 1(x+a) + (-2)x + 1(x-a) = 0$$

Dans le plan, considérant les  $n$  points  $x_i$  de coordonnées  $(X_i, Y_i)$  et les  $n$  poids  $\lambda_i$ , on dira que  $\sum_{i=1}^n \lambda_i Z(x_i)$  est un accroissement d'ordre  $k$  de la F.A.  $Z$  si:

$$\sum_{i=1}^n \lambda_i f^1(x_i) = 0$$

pour tous les monômes  $f^1$  de degré inférieur ou égal à  $k$ , c'est à dire de la forme:  $X^p Y^q$  avec  $0 \leq p+q \leq k$ .

Ainsi • pour  $k = 0$ , on aura la condition unique:

$$\sum_{i=1}^n \lambda_i = 0$$

• pour  $k = 1$ , on aura la triple condition:

$$\sum_{i=1}^n \lambda_i = 0 ; \quad \sum_{i=1}^n \lambda_i X_i = 0 ; \quad \sum_{i=1}^n \lambda_i Y_i = 0$$

généralisant bien respectivement les conditions correspondant aux différences première et seconde sur la droite.

β) Stationnarité des accroissements d'ordre  $k$

En revenant aux notations employées pour les calculs du krigeage, la condition  $\sum_{i=1}^n \lambda_i - 1 = 0$  du krigeage simple faisait que l'erreur d'estimation

$$\sum_{i=1}^n \lambda_i Z(x_i) - Z(x_0)$$

était une combinaison linéaire d'accroissements. Seule la stationnarité de ces accroissements était nécessaire et l'on montrait que la variance de l'erreur pouvait s'exprimer en fonction du variogramme  $\gamma(h)$ .

De même, dans le krigeage universel, si on admet une dérive polynômiale de degré  $k$ , c'est à dire une dérive dont les fonctions de base  $f^1$  sont les monômes de degré inférieur ou égal à  $k$ , les conditions de non-biais:

$$\sum_{i=1}^n \lambda_i f^1(x_i) - f^1(x_0) = 0$$

qui peuvent encore s'écrire:

$$\sum_{i=0}^n \lambda_i f^1(x_i) = 0 \quad \text{en posant } \lambda_0 = -1$$

font que l'erreur d'estimation

$$\sum_{i=1}^n \lambda_i Z(x_i) - Z(x_0)$$

est un accroissement d'ordre  $k$ . L'hypothèse minimale pour l'inférence statistique est alors la stationnarité de ces accroissements d'ordre  $k$ .

Autrement dit, si  $\sum_i \lambda_i Z(x_i)$  est un accroissement d'ordre  $k$ ,  $\sum_i \lambda_i Z(x_i+h)$  doit avoir même moyenne et même variance, quel que soit le vecteur  $h$ .

C'est une généralisation de l'hypothèse intrinsèque usuelle où seuls étaient concernés les accroissements d'ordre 0.

On emploie le terme de "fonction aléatoire intrinsèque généralisée d'ordre  $k$ " (en abrégé F.A.I.- $k$ ) pour désigner une F.A. satisfaisant cette hypothèse nouvelle.

γ) Covariance généralisée: modèle polynômial et ajustement automatique

La théorie montre alors que la variance d'accroissements d'ordre  $k$  s'exprime à l'aide d'une fonction  $K(h)$  appelée covariance généralisée d'ordre  $k$ . Ainsi, pourvu que  $\sum_i \lambda_i Z(x_i)$  soit un accroissement d'ordre  $k$ , sa variance s'écrit:

$$\text{var} \left[ \sum_i \lambda_i Z(x_i) \right] = \sum_i \sum_j \lambda_i \lambda_j K(x_i - x_j)$$

Ceci étant en particulier valable pour l'erreur de krigeage, les calculs peuvent se dérouler comme précédemment, en remplaçant  $-\gamma$  par  $K$ , et tout ce qui a été dit sur les propriétés du krigeage se conserve. Pour  $k=0$ , on retrouve comme covariance généralisée d'ordre 0, au signe près, le variogramme pour lequel on avait déjà écrit:

$$\text{var} \left[ \sum_i \lambda_i Z(x_i) \right] = - \sum_i \sum_j \lambda_i \lambda_j \gamma(x_i - x_j)$$

Comme pour le variogramme, une fonction quelconque ne peut être une covariance généralisée. Elle doit satisfaire certaines conditions de cohérence mathématique assurant essentiellement que les variances d'accroissements sont toujours positives. Ici l'ensemble des covariances généralisées d'ordre  $k$  contient en particulier les polynômes de degré inférieur à  $2k+2$ .

Dans le cas du plan, les modèles polynomiaux possibles sont résumés dans le tableau suivant:

Dérive	$k$	modèles de covariance polynômiale
Constante	0	$K(h) = C\delta + A_1 h $
Linéaire	1	$K(h) = C\delta + A_1 h  + A_3 h ^3$
Quadratique	2	$K(h) = C\delta + A_1 h  + A_3 h ^3 + A_5 h ^5$

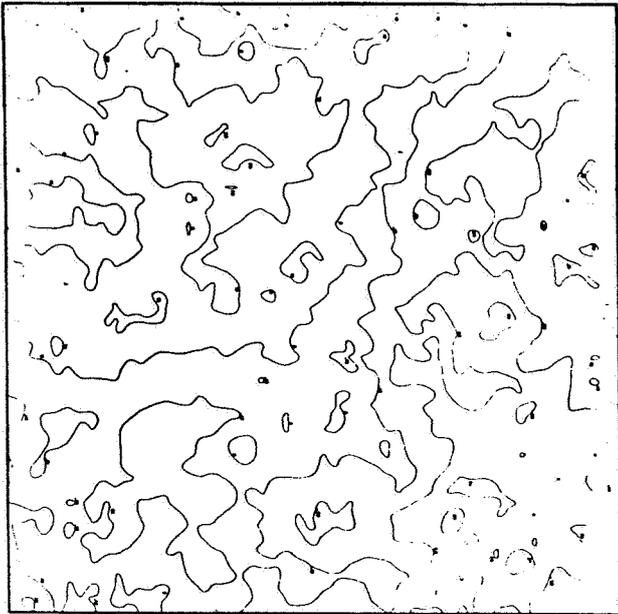
(avec  $C \geq 0$ ,  $A_1 \leq 0$ ,  $A_5 \leq 0$ ,  $A_3 \geq -\frac{10}{3} \sqrt{A_1 A_3}$ ), le Dirac  $\delta$  correspondant à l'effet de pépite.

Ces modèles sont suffisamment riches pour être adaptés à la grande majorité des applications. On pourra juger de leur aptitude à rendre compte de phénomènes très divers en se reportant à la figure 26 où sont représentées trois réalisations de F.A.I. généralisées respectivement d'ordre 0, 1 et 2.

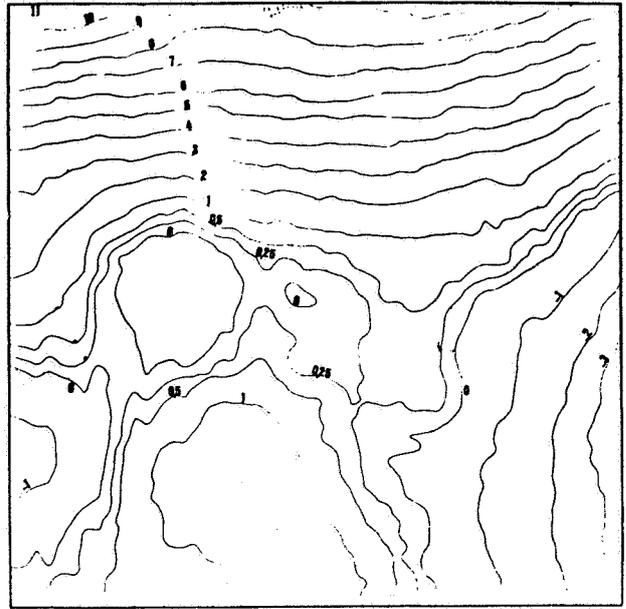
D'un point de vue pratique, du fait qu'ils dépendent linéairement de leurs paramètres, ces modèles rendent possible l'inférence statistique à l'aide d'une procédure automatique [2]. Les paramètres  $C$ ,  $A_1$ ,  $A_3$ ,  $A_5$  peuvent ainsi être obtenus par régression de variances expérimentales d'accroissements d'ordre  $k$  sur les valeurs théoriques correspondantes où ces paramètres interviennent de façon linéaire.

Après avoir brièvement rappelé la théorie des variables régionalisées, il convient maintenant, dans le dernier chapitre, d'en montrer quelques applications en hydrogéologie et hydrométéorologie.

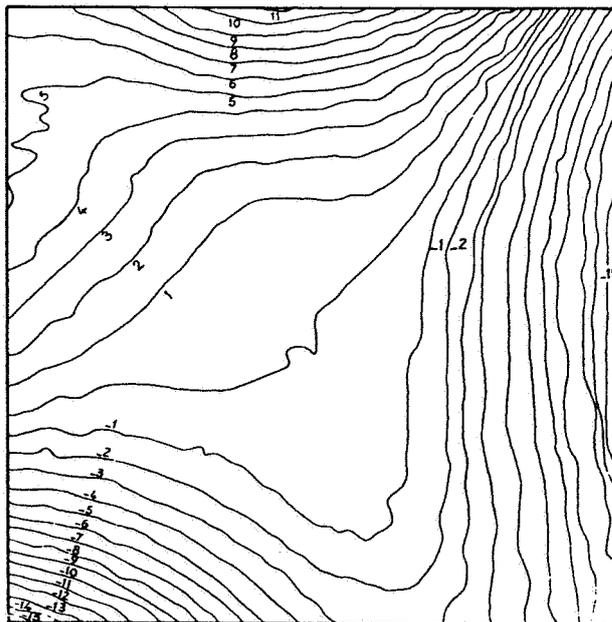
Les exemples de variogrammes expérimentaux et d'ajustement de ceux-ci à un modèle qui ont servi à illustrer la démarche théorique ne seront toutefois pas repris dans les pages qui suivent.



F.A.I. - 0  
 $K(h) = -|h|$



F.A.I. - 1  
 $K(h) = |h|^3$



F.A.I. - 2  
 $K(h) = -|h|^5$

FIG. 26 - REALISATIONS DE F.A.I. GENERALISEES D'ORDRE 0,1 ET 2  
(d'après [11])

REFERENCES BIBLIOGRAPHIQUES - CHAPITRE II

- [1] CHILES, J.P., MATHERON, G. (1975):  
Interpolation optimale et cartographie.  
Annales des Mines, Novembre 1975, pp. 1-8.
- [2] DELFINER, P. (1975):  
Linear estimation of non-stationary phenomena.  
Proceedings of NATO Adv. Study Inst., Rome, 1975
- [3] HUIJBREGTS, C. (1970):  
Le variogramme des résidus.  
Ecole des Mines, C.M.M., N-200.
- [4] MATHERON, G. (1963):  
Traité de Géostatistique appliquée.  
Mémoires du B.R.G.M. n°14, Ed. Technip.
- [5] MATHERON, G. (1965):  
Les variables régionalisées et leur estimation.  
Masson & Cie.
- [6] MATHERON, G. (1969):  
Le krigeage universel.  
Les Cahiers du C.M.M., fasc.1
- [7] MATHERON, G. (1970):  
La théorie des variables régionalisées et ses applications.  
Les Cahiers du C.M.M., fasc.5
- [8] MATHERON, G. (1970):  
Structures aléatoires et géologie mathématique.  
Revue de l'Institut International de Statistique, vol. 38,  
n°1, pp. 1-10.
- [9] MATHERON, G. (1971):  
La théorie des F.A.I.G.  
Note Géostatistique n° 117, Ecole des Mines, C.M.M., N-252
- [10] MATHERON, G. (1973):  
The intrinsic random functions, and their applications.  
Adv. in Appl. Prob., 5, pp. 439-468.
- [11] ORFEUIL, J.P. (1972):  
Simulation du Wiener-Lévy et de ses intégrales.  
Ecole des Mines, C.M.M., N-290.

CHAPITRE III  
APPLICATION  
A L'ETUDE DES VARIATIONS SPATIALES  
DES GRANDEURS HYDROLOGIQUES

1 - ESTIMATIONS PONCTUELLES ET ZONALES

Les premières applications qui ont été réalisées dans le domaine ont permis de résoudre les problèmes les plus immédiats posés par les estimations ponctuelles et zonales. Mais l'exposé commencera par ce qui, dans le déroulement actuel d'une étude, suit immédiatement la phase d'identification structurale: la vérification du modèle de structure qui a été adopté.

a) Vérification de la validité du modèle structural adopté

Après que, à partir des données expérimentales disponibles, un modèle de variogramme ait été ajusté (ou bien que le degré de dérive et les coefficients de la covariance généralisée polynômiale aient été déterminés), il y a lieu de s'assurer de la validité des paramètres structuraux retenus.

La méthode utilisée consiste à supprimer fictivement les points expérimentaux l'un après l'autre, à les kriger à partir de leurs voisins, et à comparer les erreurs d'estimation ainsi constatées aux écarts-types théoriques. On peut alors visualiser la répartition spatiale de ces erreurs et des erreurs rapportées aux écarts-types d'estimation, et en faire une analyse statistique.

Un essai de krigeage a été ainsi réalisé sur des données de transmissivité (ou plutôt leurs logarithmes) provenant de la nappe du Bathonien en Normandie [9]. Ces logarithmes de transmissivités ne présentent pas de dérive, et on a pu ajuster un modèle de variogramme linéaire avec discontinuité à l'origine:

\*\*\*\*\*  
 \* KRIGEAGE DES POINTS EXPERIMENTAUX \*  
 \*\*\*\*\*

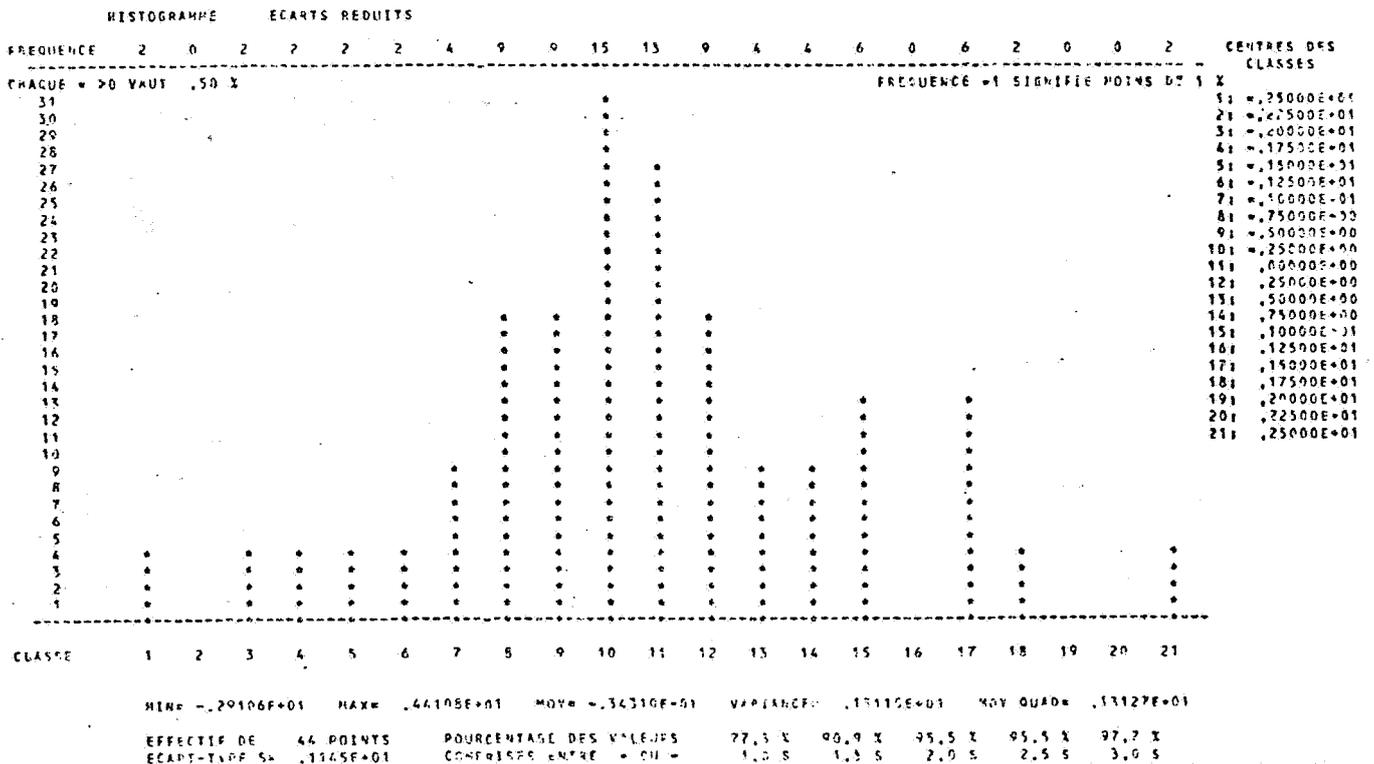
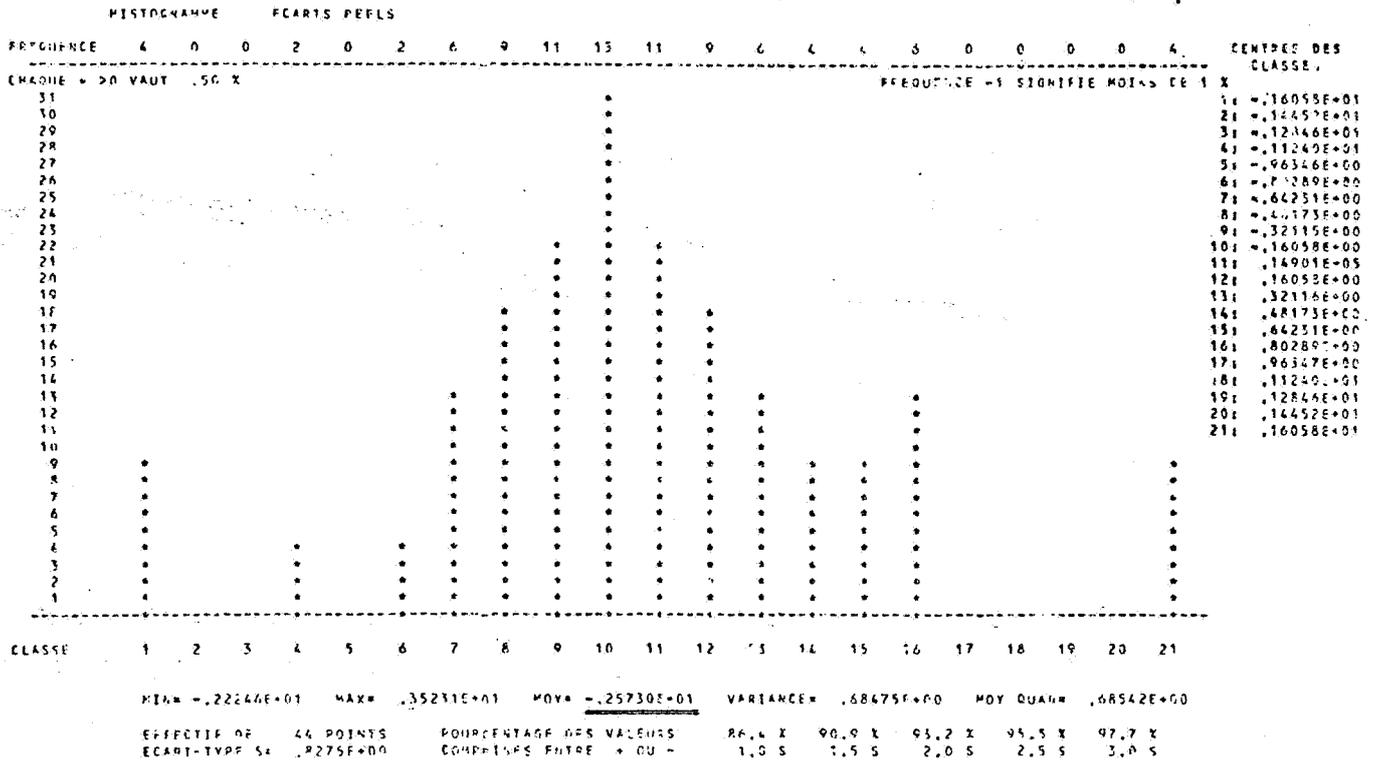


FIG. 27 - BASSE-NORMANDIE (LOGARITHME DE TRANSMISSIBILITE)  
 ESSAI DE RESSAIS

ESTIMATION ROBUSTE  
 DE L'ECART QUADRATIQUE REDUIT

$$\gamma(h) = 0.09 \delta + 0.125 h$$

les transmissivités étant exprimées en m<sup>2</sup>/s, les coordonnées en km et les logarithmes considérés étant décimaux.

Sur la figure 27, on trouvera les statistiques élémentaires et les histogrammes correspondant aux écarts réels  $z-z^*$  et aux écarts réduits  $\frac{z-z^*}{\sigma_k}$  obtenus aux 44 points expérimentaux.

On peut vérifier qu'en moyenne il n'y a bien ni sur-estimation ni sous-estimation systématique:

$$\text{erreur moyenne} = \frac{1}{n} \sum (z-z^*) = -0.026 \approx 0$$

et que les écarts-types d'estimation annoncés sont bien cohérents avec les erreurs constatées:

$$\text{erreur quadratique réduite} = \frac{1}{n} \sum \left(\frac{z-z^*}{\sigma_k}\right)^2 = 0.96 \approx 1$$

A noter enfin que 95% environ des écarts réduits sont compris entre +2 et -2, ce qui signifie encore que 95% des erreurs constatées sont inférieures en valeur absolue à deux fois l'écart-type théorique annoncé.

Ayant ainsi vérifié la validité des paramètres structuraux retenus, on pourra ensuite passer à la résolution de différents problèmes d'estimation posés: cartographie ou estimation de valeurs moyennes.

#### b) Emploi du krigeage en cartographie

Pour illustrer cette première utilisation du krigeage, on considérera l'exemple du bassin versant de l'Ouadi Kadjemur (Tchad) [2], [6], qui montre toute l'importance de la prise en compte de la variabilité des phénomènes dans les problèmes de cartographie. La figure 28 présente, sur un fond comportant le contour du bassin et les hauteurs d'eau mesurées aux 33 pluviomètres durant un orage donné (6/8/66), un tracé manuel des isohyètes correspondantes.

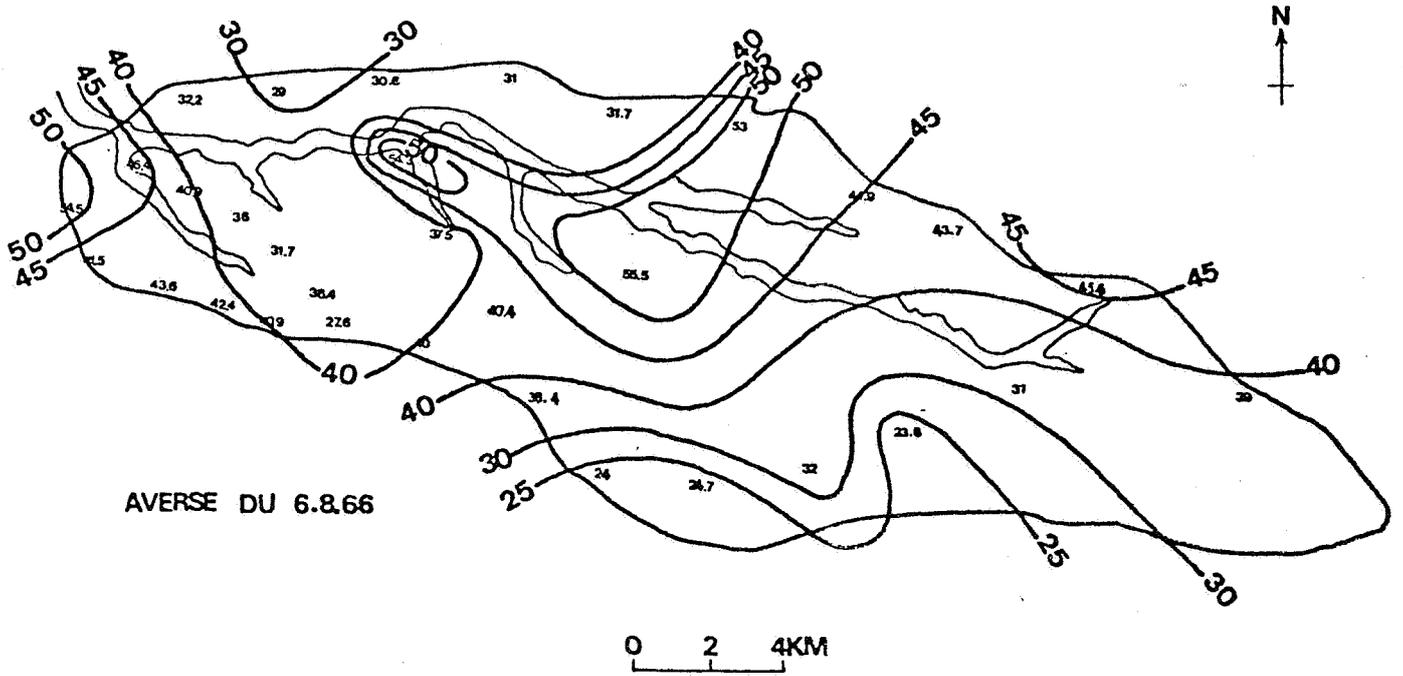


FIG. 28 - BASSIN VERSANT DE KADJEMEUR: TRACE MANUEL DES ISOHYETES (en mm.)

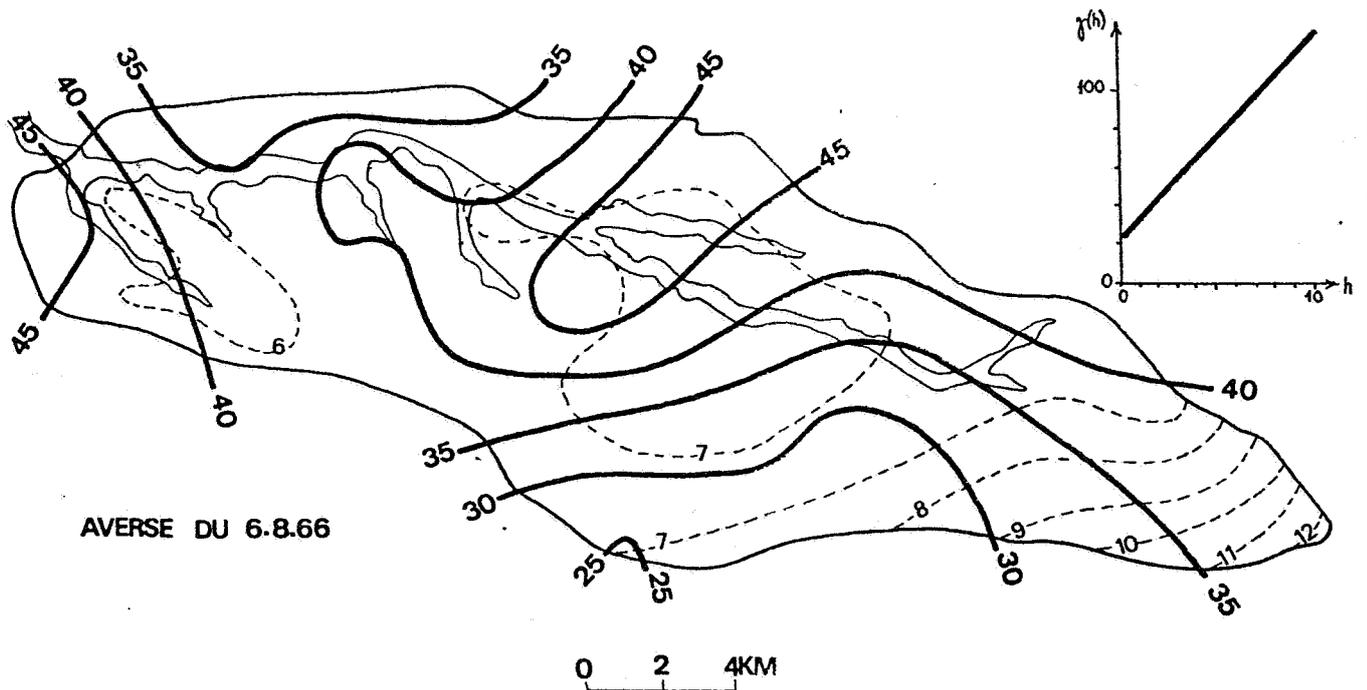


FIG. 29 - BASSIN VERSANT DE KADJEMEUR: CARTE KRIGEE  
(ECARTS-TYPES D'ESTIMATION EN TIRETE)

A partir des mêmes données, ayant estimé la pluviométrie par krigeage des valeurs ponctuelles aux noeuds d'une grille régulière de 1 km de côté, on a tracé la carte de la figure 29, que l'on désignera sous le nom de "carte krigée".

Pour ce krigeage, le bassin ne présentant pas un relief marqué et la pluviométrie n'y possédant pas de dérive systématique, on a pris comme seule fonction de base  $f^0 \equiv 1$ . Le modèle ajusté sur le variogramme expérimental (Fig. 17, p.49) était

$$\gamma(h) = \begin{cases} 0 & \text{pour } h = 0 \\ 20.4 + 11.23 h & \text{pour } h \neq 0 \text{ (en km)} \end{cases}$$

(en mm<sup>2</sup>)

La discontinuité à l'origine traduit, comme on l'a vu plus haut, une irrégularité à petite échelle. Ce phénomène a été observé depuis longtemps par les hydrométéorologues qui l'expliquent par les perturbations locales, l'instabilité de l'air au voisinage du sol et l'arrivée de la pluie sur le pluviomètre par rafales irrégulières.

Ce qui frappe en comparant les figures 28 et 29, c'est que la carte manuelle semble plus fouillée que la carte krigée. Par exemple, sur le tracé manuel, la valeur 54.3 mm, située dans le nord-Ouest du bassin, a été considérée comme un maximum local; une isohyète cotée 50 mm l'entoure et l'isohyète 45 mm fait un important détour.

Au contraire, la carte krigée ne tient pas grand compte de cette valeur 54.3 mm, et l'isohyète la plus proche est cotée 40 mm. La même chose se produit pour les données 53.0 mm et 55.5 mm, au centre-nord du bassin. On peut noter encore que l'isohyète 25 mm ne fait plus qu'une discrète apparition à l'intérieur du bassin.

La carte manuelle semble plus fidèle aux valeurs expérimentales. Mais prise entre cette fidélité aux données et la nécessité de dégager les grands traits représentatifs du phénomène, la cartographie manuelle exige en permanence des choix plus ou moins arbitraires. Ainsi, il n'a pas été tenu compte de la valeur 27.6 mm mesurée dans le sud-Ouest du bassin, alors qu'elle aurait sans doute modifié le tracé de l'isohyète 40 mm. L'intérêt n'est pas ici de faire une critique du travail du dessinateur, mais d'attirer l'attention sur la difficulté de juger de la précision d'une carte manuelle.

Le krigeage, lui, outre le fait que, par construction, il donne une estimation optimale, permet de quantifier la précision de cette estimation, à l'aide de la variance  $\sigma_k^2$  de l'erreur de krigeage. Conventionnellement, on adoptera comme intervalle de confiance pour la valeur réelle  $z(x_0)$ :

$$[z^*(x_0) - 2\sigma_k, z^*(x_0) + 2\sigma_k]$$

qui correspond à un niveau de certitude de 95% dans l'hypothèse d'une erreur gaussienne.

La pratique courante veut que l'on associe à toute carte krigée la carte d'écart-type d'estimation  $\sigma_k$  correspondante. Ici, sur la figure 29,  $\sigma_k$  a été cartographié en lignes pointillées superposées à la carte krigée. On peut voir ainsi quelle est la précision de cette dernière en tout point du bassin, la zone la plus mal connue étant la partie sud-est ( $\sigma_k \approx 11$  mm). Si on prend, par exemple, un point de grille situé au centre du bassin près de la cote 55.5 mm (plus forte valeur mesurée) et que l'on considère à la fois la valeur estimée 48.7 mm et l'écart-type d'estimation 6.25 mm, on voit que la valeur estimée est parfaitement compatible avec la donnée voisine.

L'aspect lisse de la carte krigée, qui avait surpris au départ, n'est pas dû à un lissage arbitraire, mais en fait, comme on va le voir, à la présence d'un effet de pépite.

La composante chaotique du phénomène (erreurs de mesure ou microrégionalisation) qui est modélisée par l'effet de pépite, est totalement imprévisible, car non corrélée spatialement. Aux points expérimentaux eux-mêmes, le krigeage étant un interpolateur exact restituerait la valeur mesurée. En dehors de ces points, le système de krigeage conduit à une estimation du phénomène filtré de cette composante chaotique, encore une fois imprévisible.

Si, avec les mêmes données, on avait pris un variogramme continu à l'origine (par exemple  $\gamma(h) = 15|h|$ ), le krigeage donnerait une carte (Fig.30) qui passerait scrupuleusement par les valeurs expérimentales, et même mieux que ne le faisait la carte manuelle.

Mais si ce type de cartographie convient à un phénomène très régulier, il est irréaliste pour un phénomène comme la pluviométrie: cela

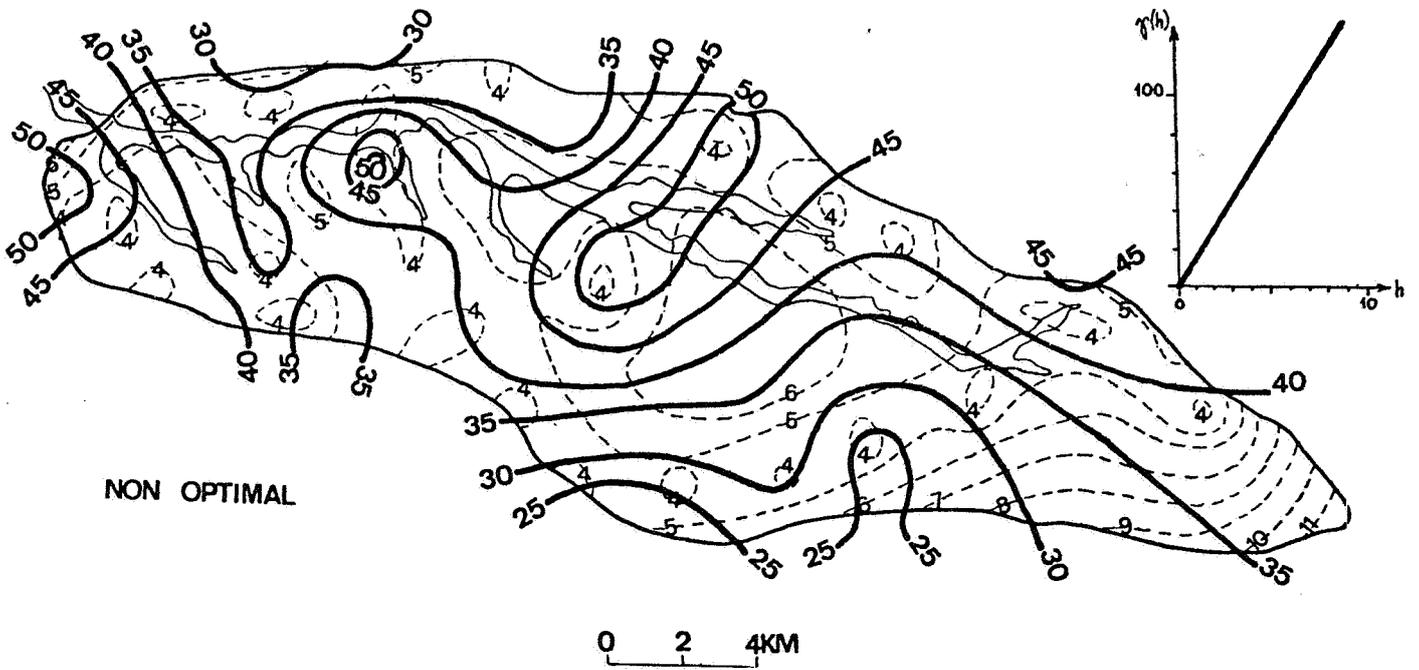
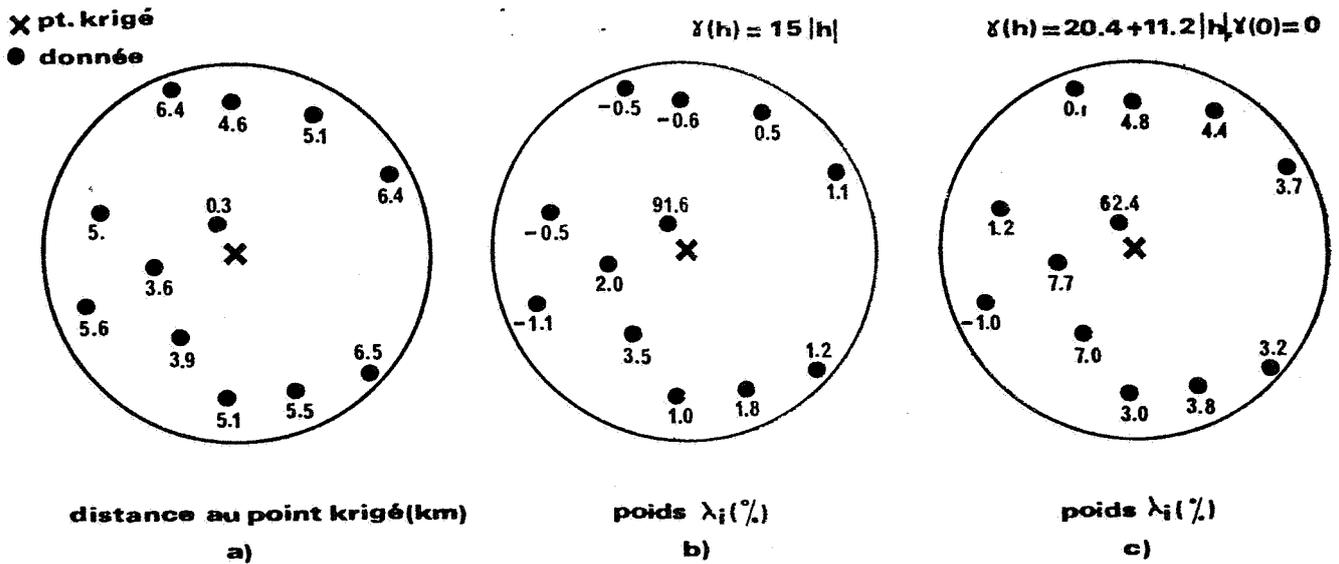


FIG. 30 - BASSIN VERSANT DE KADJEMEUR: CARTE OBTENUE PAR KRIGEAGE  
EN NEGLIGEANT L'EFFET DE PEPITE



b) Variable continue: grande influence du point le plus proche 91.6%  
 c) Variable discontinue: influence atténuée 62.4%

FIG. 31 - INFLUENCE DE LA STRUCTURE SUR LES POIDS DE KRIGEAGE

reviendrait à admettre que les épicentres d'un orage se trouvent nécessairement à la verticale d'un pluviomètre !

Pour comprendre comment la présence d'un effet de pépite a pu influencer l'aspect de la carte, notamment au centre du bassin, on peut considérer à nouveau le point de grille situé le plus près de la donnée 55.5 mm. Ce point a été estimé à partir de douze points du voisinage. La figure 31 permet de comparer les pondérations. Le poids optimum pour la donnée 55.5 mm, située à 300 m du point de grille, est seulement de 62.4%, alors qu'il aurait été de 91.6% en l'absence d'effet de pépite.

On peut constater que le krigeage réalise en quelque sorte une pondération "sur mesure" tenant compte de la variabilité du phénomène. Quand celui-ci est continu, le krigeage accorde un poids important aux valeurs expérimentales proches, parce que, précisément, la continuité signifie que deux points voisins auront des valeurs voisines. Quand le phénomène est irrégulier, ceci n'est plus vrai, et le poids accordé aux points expérimentaux proches perd de l'importance.

### c) Estimation optimale de valeurs moyennes

Envisageant maintenant l'application du krigeage aux estimations de valeurs moyennes, on montrera deux exemples, l'un concernant des moyennes par mailles, l'autre concernant une valeur moyenne sur un domaine de forme quelconque.

#### α) Cas d'un maillage

L'utilisation des modèles mathématiques s'étant répandue en hydrogéologie, on a de plus en plus besoin d'estimer des valeurs moyennes par maille. Il peut s'agir d'introduire dans le modèle les grandeurs régissant l'écoulement (transmissivités, cote des murs et toits des nappes, épaisseurs de semi-perméables,...) ou la piézométrie destinée à servir d'état de référence lors du calage.

VALEURS REELLES = VALEURS IMPRIMÉES MULTIPLIÉES PAR 100000 00

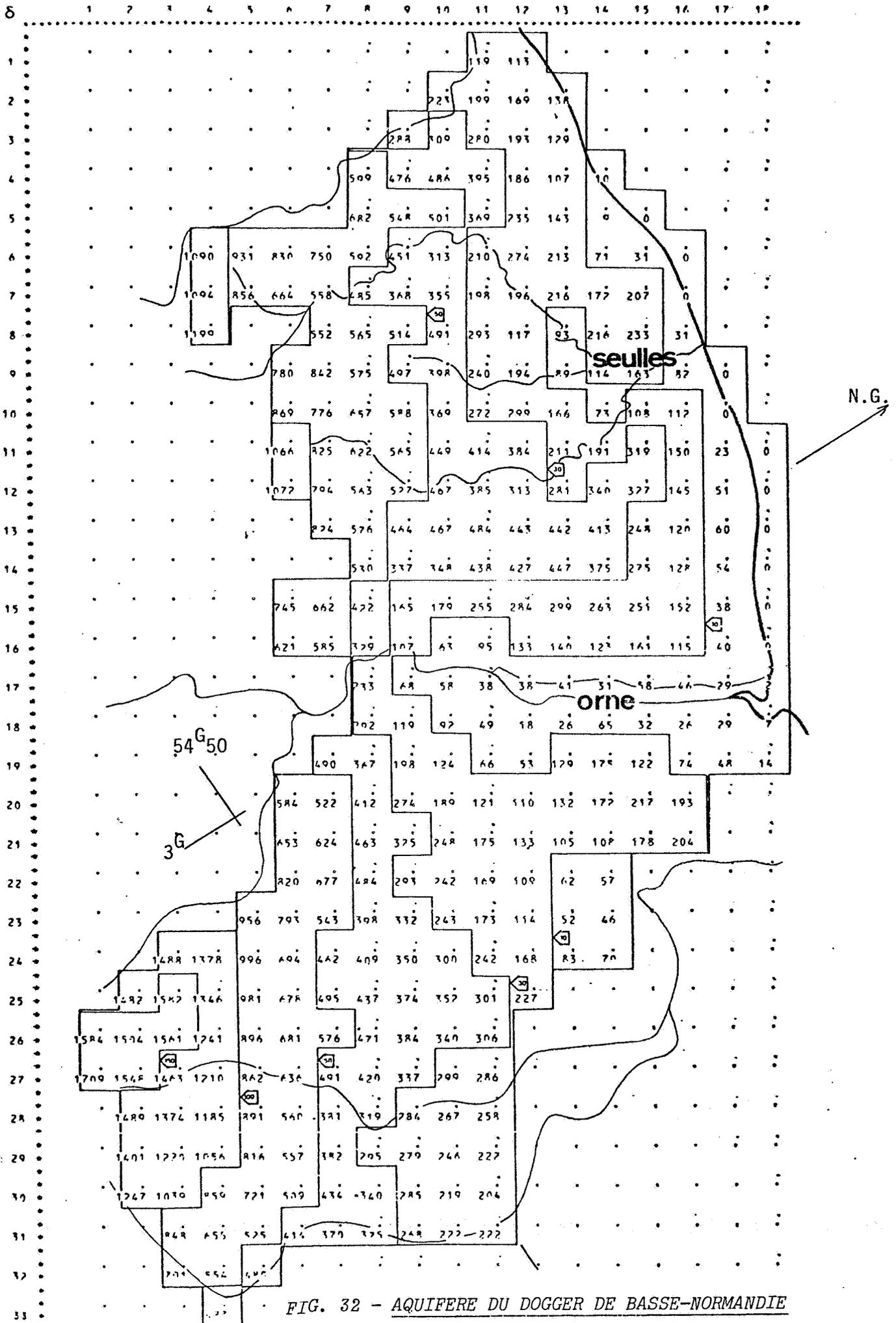


FIG. 32 - AQUIFERE DU DOGGER DE BASSE-NORMANDIE  
PIEZOMETRIE MOYENNE PAR MAILLE

VALEURS REELLES = VALEURS IMPRIMES MULTIPLIÉES PAR 100000 00

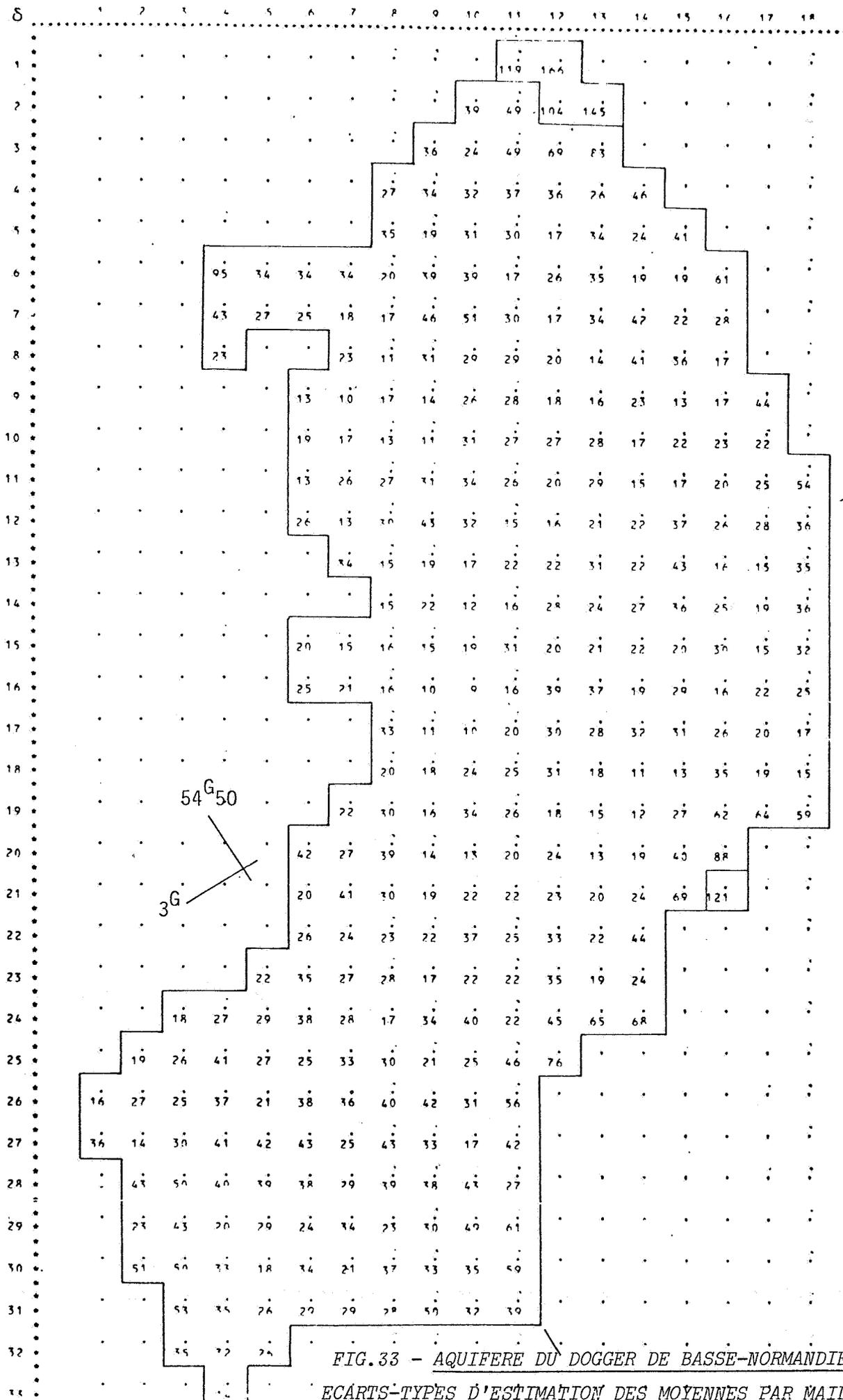


FIG. 33 - AQUIFERE DU DOGGER DE BASSE-NORMANDIE ECARTS-TYPES D'ESTIMATION DES MOYENNES PAR MAILLE

L'exemple présenté est celui de l'aquifère du Dogger en Basse-Normandie [9]. Un état piézométrique de juin 1974 a été krigé sur des mailles de 2 km de côté à partir des 208 relevés piézométriques disponibles auxquels on a ajouté quelques points correspondant à des endroits où les cours d'eau et la mer imposent le potentiel de la nappe.

Utilisant cette fois le procédé d'identification structurale automatique mis au point à partir de la théorie des F.A.I.- $k$ , on a retenu les caractéristiques structurales suivantes:

- dérive linéaire,
- covariance généralisée  $K(h) = -10|h| + 0.75|h|^3$   
(piézométrie en m, distances en km).

L'absence ici d'effet de pépite est le signe de la régularité spatiale de la piézométrie.

A noter que ces caractéristiques sont valables localement puisqu'on a travaillé par voisinages glissants.

On peut constater, sur la figure 32, que le modelé piézométrique est bien rendu, avec les entailles des vallées de l'Orne et de la Seulles. Les écarts-types de krigeage (Fig.33) sont de l'ordre de 3 mètres en général, ce qui peut sembler élevé; mais il faut se souvenir que la nappe est une nappe libre dans un réservoir calcaire entaillé par de nombreuses vallées, et que le modelé piézométrique y est assez tourmenté, avec des gradients importants.

#### β) Cas d'un domaine quelconque

Pour illustrer le calcul de quantités moyennes sur un domaine, on va reprendre l'exemple de l'Ouadi Kadjemour [6]. Les estimations par krigeage de lames d'eau moyennes ont été effectuées pour les treize épisodes pluvieux de 1966. Vu le faible nombre de points de mesure, il était difficile de procéder à l'inférence statistique du variogramme averse par averse, d'autant plus que parfois certaines données étaient manquantes. Prendre brutalement le variogramme moyen sur l'ensemble des averses eut été une



hypothèse simpliste, car ces averses diffèrent par leur intensité et leur dispersion. Une hypothèse plus raisonnable a été d'admettre que les variogrammes des épisodes pluvieux sont proportionnels:

$$\gamma_k(h) = \omega_k \bar{\gamma}(h)$$

où  $\gamma_k(h)$  est le variogramme de la  $k^{\text{ième}}$  averse,  $\bar{\gamma}(h)$  le variogramme moyen et  $\omega_k$  le coefficient de proportionnalité.

Cette relation équivaut à admettre qu'il y a conservation des corrélations spatiales sur le bassin. En notant  $s_k^2$  la variance expérimentale des hauteurs d'eau de la  $k^{\text{ième}}$  averse, et  $\bar{s}^2$  la moyenne des  $s_k^2$ , il en résulte que:

$$\omega_k = s_k^2 / \bar{s}^2$$

Le variogramme moyen est:  $\bar{\gamma}(h) = \begin{cases} 0 & \text{pour } h = 0 \\ 16.2 + 8.91 h & \text{pour } h \neq 0 \end{cases}$   
(en mm<sup>2</sup>) (en km)

Il ne reste plus ensuite qu'à poser, pour chaque averse, le système de krigeage, et à le résoudre pour obtenir les poids optimaux  $\lambda_i$  à affecter aux pluviomètres et les  $\mu_1$  nécessaires au calcul de la variance. Mais l'observation des systèmes de krigeage révèle que si l'on multiplie le variogramme par une constante  $\omega_k$ , les poids  $\lambda_i$  restent inchangés tandis que les  $\mu_1$ , ainsi que la variance d'estimation, sont multipliés par  $\omega_k$ . Avec des variogrammes proportionnels, il suffit de faire le calcul une seule fois à condition, bien sûr, que les pluviomètres relevés restent les mêmes.

La figure 35 présente les valeurs des poids de krigeage correspondant à la configuration expérimentale complète. Il a semblé intéressant de les rapprocher des poids fournis par la méthode de THIESSEN (Fig.34).

On constate une similitude entre les deux pondérations, similitude sur laquelle on reviendra en fin de paragraphe.

De façon à permettre une comparaison plus complète avec la méthode de THIESSEN, mais aussi les autres méthodes couramment employées,

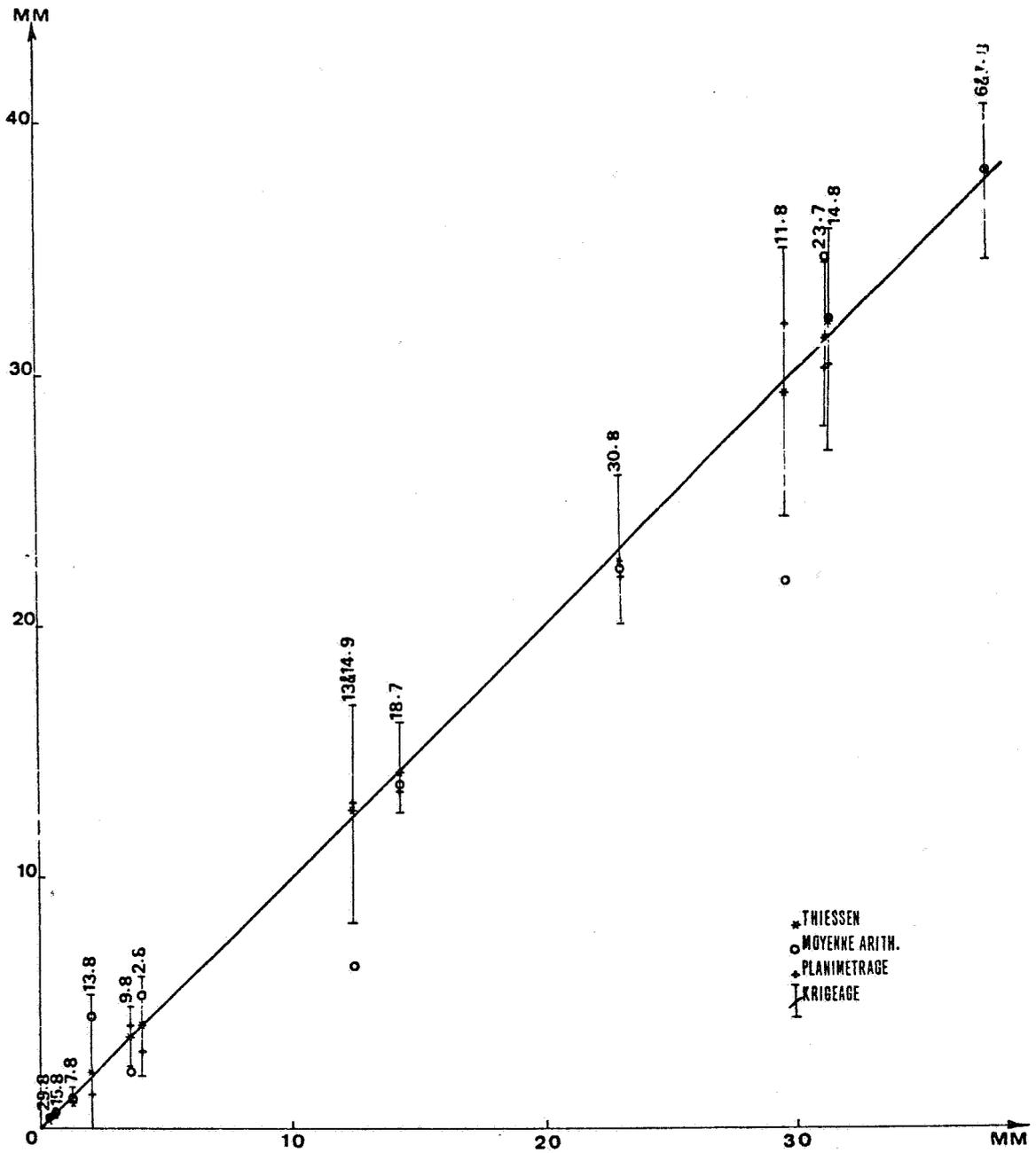


FIG. 36 - BASSIN VERSANT DE KADJEMEUR  
COMPARAISON DES DIFFERENTES METHODES D'ESTIMATION GLOBALE  
(SUR 13 EPISODES PLUVIEUX)

on a reporté (Fig.36) les évaluations des lames d'eau pour les treize averses par moyenne arithmétique, planimétrage, méthode de THIESSEN et krigeage. De part et d'autre de la bissectrice des axes, sur laquelle figure la valeur krigée, on a indiqué la fourchette à  $\pm 2$  écarts-types de krigeage.

Il ressort que trois méthodes donnent des résultats à peu près équivalents. Seule la moyenne arithmétique se singularise, en particulier sur les averses du 9 août, du 13 au 14 septembre, du 11 août et du 23 juillet, où les valeurs estimées se situent en dehors de la fourchette pourtant très large de quatre écarts-types. Avec la moyenne arithmétique, tous les pluviomètres ont même importance; le réseau étant plus fourni à l'ouest, il y a systématiquement sous-estimation ou sur-estimation de la lame d'eau selon que l'épicentre des averses se situe à l'est ou à l'ouest.

Une fois écartée la moyenne arithmétique, il semblerait donc que, du moins sur l'exemple traité, l'avantage du krigeage soit assez minime.

C'est pourtant la seule méthode autour de laquelle a pu s'articuler la comparaison, grâce à la présence d'un écart-type d'estimation indiquant avec quelle précision est donné le chiffre avancé.

Mais surtout, alors que la pondération de THIESSEN ne peut tenir compte de la présence de dérive systématique (due par exemple à l'orographie) sans recourir à des transformations sur les données, le formalisme du krigeage le permet. On a ainsi reporté sur les figures 37 et 38 les pondérations optimales qui correspondraient, avec le même variogramme, à la présence d'une dérive linéaire et d'une dérive quadratique. On peut constater que la dernière pondération diffère, elle, notablement des poids de THIESSEN.

Les applications qui viennent d'être présentées n'ont nécessité aucune adaptation du formalisme et de la méthodologie définis au Chapitre II. A côté de ces applications directes, la méthode du krigeage est suffisamment souple pour pouvoir prendre en compte des besoins particuliers comme:

- la prise en compte d'une éventuelle hétérogénéité de la structure,
- le recours à une transformation sur la variable étudiée,
- la nécessité d'utiliser des données de qualité différente,
- l'obligation de tenir compte de la présence de failles dans le domaine à cartographier.



d) Adaptations de la méthode à des problèmes particuliers

α) Hétérogénéité de la structure

Dans les sciences de l'eau, il existe peu de cas permettant d'étudier la variabilité spatiale de la structure elle-même: un variogramme unique est ainsi adopté pour tout le champ. Le modelé de la carte d'écart-types d'estimation reflète alors principalement la configuration géométrique des données, ainsi qu'éventuellement leurs qualités respectives, comme on le verra au paragraphe γ).

Quand on dispose d'une information plus dense, on peut diviser le champ étudié en carreaux, et pour chacun ajuster un variogramme différent [1]. Tenant alors compte dans l'estimation de l'hétérogénéité éventuelle du champ, on obtient des cartes d'écart-types reflétant à la fois la configuration expérimentale et la variabilité locale. Ainsi, à densité égale de points de mesure, la précision de l'estimation sera moins bonne dans une zone très tourmentée que dans une zone calme. Si l'on prend, par exemple, le cas simple où le variogramme est partout localement de la forme  $\gamma(h) = \bar{\omega} \cdot \gamma_0(h)$  (la valeur du facteur  $\bar{\omega}$  allant en croissant avec la variabilité locale), l'examen du système de krigeage montre que, à configuration géométrique identique, la variance de krigeage est proportionnelle à  $\bar{\omega}$ .

β) Transformation logarithmique

L'expérience montre que l'on a parfois intérêt à travailler sur une transformée (par exemple logarithmique) de la V.R. étudiée.

A partir des 86 données de transmissivité disponibles sur la nappe des sables éocènes de la région Nord-Aquitaine [7], on a construit, en l'absence de dérive apparente, le variogramme représenté sur la figure 39.

Son allure plate témoignait du caractère assez mal structuré de la transmissivité. L'histogramme des données (Fig.40), et la pratique courante consistant à raisonner sur les transmissivités en puissances de 10, ont suggéré de passer au logarithme.

Le variogramme construit à partir des logarithmes des 86 données, représenté sur la figure 41, peut être assimilé à un modèle linéaire avec une faible discontinuité à l'origine.

Le logarithme de la transmissivité, se révélant être mieux structuré que la transmissivité elle-même, lui sera donc préféré pour le krigeage.

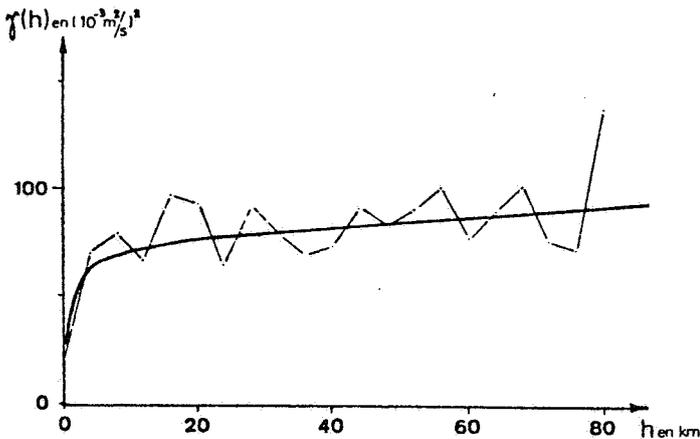


FIG. 39 - VARIOGRAMME CONSTRUIT  
A PARTIR DES DONNEES

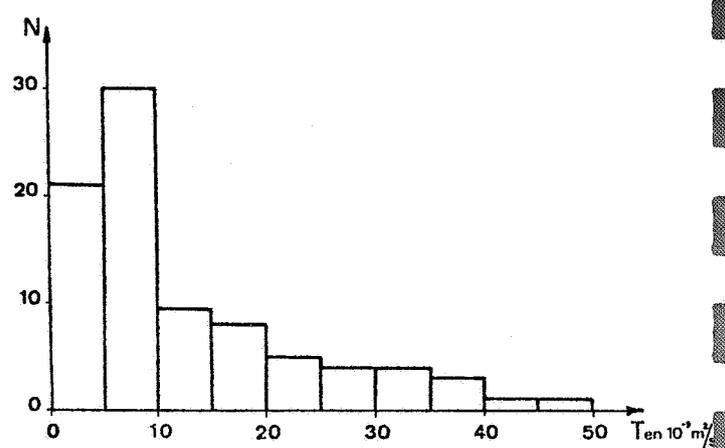


FIG. 40 - HISTOGRAMME DES DONNEES

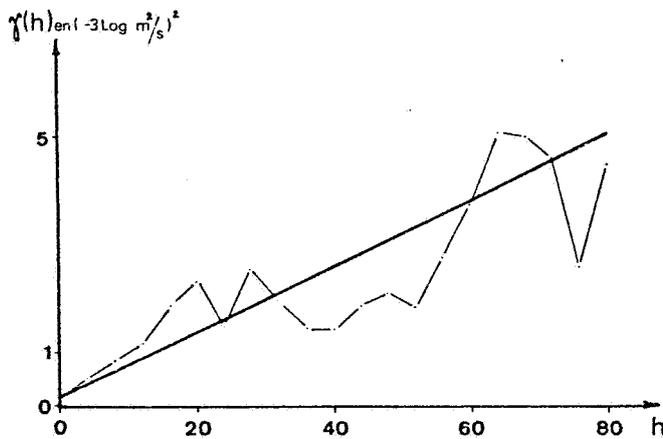


FIG. 41 - VARIOGRAMME CONSTRUIT A PARTIR  
DES LOGARITHMES DES DONNEES

TRANSMISSIVITE  
DES SABLES EOCENES  
NORD-AQUITAINS

Quand, ainsi, les données du krigeage sont constituées par des logarithmes (par exemple décimaux) de transmissivité, les calculs du krigeage donnent une estimation  $z_0^*$ , un écart-type  $\sigma_K$ , et donc un intervalle de confiance  $z_0^* \pm 2\sigma_K$  relatifs à cette variable.

Le retour à des valeurs de transmissivité en vue de l'édition des résultats conduit à une estimation médiane de  $T$  égale à  $10^{z_0^*}$  et l'assortit d'une spécification de l'intervalle de confiance sous la forme d'un coefficient multiplicateur égal à  $10^{2\sigma_K}$ :

$$10^{z_0^*} / 10^{2\sigma_k} < T < 10^{z_0^*} \cdot 10^{2\sigma_k}$$

Il faudra donc raisonner en termes multiplicatifs, et non plus additifs, quand on examine de tels documents.

Ce passage au logarithme a fait l'objet d'études depuis les débuts de la géostatistique. On a ainsi pu montrer [10] que si  $Y(x) = \log Z(x)$  est une F.A.I. gaussienne ayant un variogramme  $\gamma(h)$  sans portée (par exemple  $\gamma(h) = \bar{\omega}|h|$ ), le variogramme expérimental de  $Z(x)$  est, en espérance, de la forme  $a(1-e^{-\gamma(h)})$  et présente, par conséquent, une portée apparente. C'est sans doute un effet de cet ordre qui entraîne le caractère plat de la plupart des variogrammes expérimentaux bâtis à partir de données de transmissivité.

#### γ) Données incertaines

On peut être amené à traiter des données de différentes qualités, provenant par exemple de sources plus ou moins sûres. Certaines valeurs doivent être utilisées avec précaution sans que, pour autant, on puisse les rejeter sous peine de se priver d'une information non négligeable [8]. Une formulation adaptée du krigeage permet de prendre en compte l'ensemble des données [7].

Les incertitudes entachant certaines données peuvent provenir de ce qu'elles ont été obtenues à l'aide d'une technique de mesure rudimentaire, ou encore à l'aide d'une régression (ou de formules empiriques) permettant de les déduire d'autres variables mesurées par ailleurs. On supposera que l'incertitude sur une donnée  $z(x_i)$  peut être résumée par une variance  $\sigma_i^2$ , soit provenant d'un calcul statistique, soit même estimée grossièrement en assimilant la fourchette d'incertitude à  $\pm 2\sigma_i$ .

On supposera en outre que les erreurs  $\epsilon(x_i)$  sur les données sont:

- a) non systématiques:  $E[\epsilon(x_i)] = 0, \quad \forall i$
- b) non corrélées entre elles:  $\text{cov} [\epsilon(x_i), \epsilon(x_j)] = 0, \quad \forall i \neq j$
- c) non corrélées avec la variable étudiée:  $\text{cov} [\epsilon(x_i), Z(x)] = 0, \quad \forall i, \quad \forall x$

L'incertitude sur les données n'introduit alors qu'une légère modification dans le système de krigeage. On ne dispose plus, comme données, des vraies valeurs  $z(x_i)$ , mais seulement des sommes "vraie valeur + erreur", c'est à dire des  $z(x_i) + \epsilon(x_i)$ . L'estimateur formé pour le krigeage prend donc la forme:

$$Z_0^* = \sum_{i=1}^n \lambda_i (Z(x_i) + \epsilon(x_i))$$

Comme  $E[\epsilon(x_i)] = 0$ , les conditions de non-biais ne sont pas modifiées. La variance de l'erreur d'estimation devient:

$$\begin{aligned} \text{var} [Z_0^* - Z_0] &= \text{var} \left[ \sum_{i=1}^n \lambda_i (Z(x_i) + \epsilon(x_i)) - Z_0 \right] \\ &= \text{var} \left[ \sum_{i=1}^n \lambda_i Z(x_i) - Z_0 \right] + \text{var} \left[ \sum_{i=1}^n \lambda_i \epsilon(x_i) \right] \\ &= \text{var} \left[ \sum_{i=1}^n \lambda_i Z(x_i) - Z_0 \right] + \sum_{i=1}^n \lambda_i^2 \sigma_i^2 \end{aligned}$$

En minimisant cette expression, on retombe sur le système classique de krigeage, excepté pour les termes diagonaux où l'on a  $-\sigma_i^2$  au lieu de  $\gamma_{ii} = 0$ .

En considérant, par exemple, le cas ponctuel sans dérive, le système modifié s'écrit:

$$\begin{cases} \sum_{j=1}^n \lambda_j \gamma_{ij} - \lambda_i \sigma_i^2 + \mu = \gamma_{i0} & i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases}$$

la variance de krigeage valant toujours:

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \gamma_{i0} + \mu$$

Il peut être intéressant de visualiser les effets concrets de la présence sur une donnée  $z(x_i)$  d'une incertitude, caractérisée par  $\sigma_i^2$ . Pour ce faire, on considère l'exemple simple suivant: la variable étudiée est sans dérive et de variogramme  $\gamma(h) = |h|$ , les données sont situées aux six sommets d'un hexagone régulier de côté unité (Fig.42). Cinq de ces

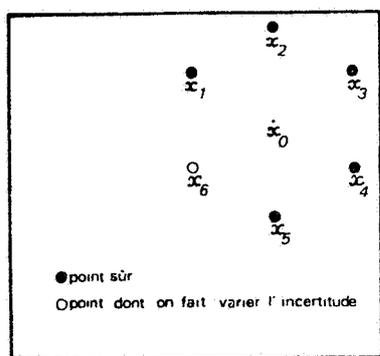


FIG. 42 - DONNEES SURES  
& DONNEE INCERTAINE

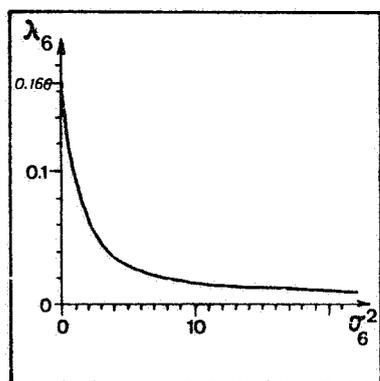


FIG. 43 - INFLUENCE DE  
L'INCERTITUDE SUR LE POIDS

données sont supposées sans erreur ( $\sigma_i^2 = 0$  pour  $i=1,2,\dots,5$ ) et l'on cherche à appréhender l'influence de  $\sigma_6^2$ .

Une façon d'y parvenir consiste à regarder le poids  $\lambda_6$  attribué à  $z(x_6)$  dans l'estimation par krigeage de la valeur au centre de l'hexagone (Fig.43) Au départ,  $\sigma_6^2 = 0$  et la sixième donnée jouant le même rôle que les cinq autres se voit attribuer un poids  $\lambda_6 = 1/6 (= 0.166)$ . A mesure que  $\sigma_6^2$  croît, le poids  $\lambda_6$  décroît et à la limite tend vers zéro quand  $\sigma_6^2$  tend vers l'infini. Une donnée totalement incertaine équivaldrait à une absence de donnée.

Cette adaptation du krigeage au cas de données de différentes qualités a été appliquée à l'étude des transmissivités de la nappe du Bathonien en Basse-Normandie [9]. Les 45

données provenant d'essais de pompage ont été complétées par 55 données de moindre qualité, obtenues à partir d'une régression linéaire entre les logarithmes des transmissivités et ceux des débits spécifiques. Les variances  $\sigma_i^2$  affectées à ces 55 dernières valeurs ont été calculées à partir de la formule classique en statistique.

L'estimation de la transmissivité a été ensuite faite par krigeage sur des mailles de 2 km de côté, une première fois en ne retenant que les 45 données sûres, puis en prenant l'ensemble des données et en tenant compte des incertitudes entachant les 55 valeurs provenant de débits spécifiques. Du fait de la prise en compte d'une information supplémentaire, même imprécise, les coefficients de l'intervalle de confiance (on a travaillé sur les logarithmes) sont inférieurs à ceux trouvés la première fois. Afin de visualiser ce resserrement des intervalles de confiance, on a fait le rapport des premiers coefficients sur les seconds et représenté, sur la figure 44, ce facteur d'amélioration de la connaissance correspondant au passage des 45 aux 100 (45+55) données. S'il reste sensiblement égal à 1 dans les régions bien reconnues par essais de pompage, il passe à des valeurs supérieures à 10 dans certaines



zones où l'appoint des débits spécifiques a été décisif.

δ) Présence de failles

Les phénomènes étudiés jusqu'ici se développaient dans l'espace sans cassures. En géologie, la présence d'une faille fait que les données situées au-delà par rapport au point  $x_0$  à estimer n'apportent pas d'information sur  $z(x_0)$ . Tous les procédés de cartographie automatique se sont heurtés à cette difficulté [11].

On est ainsi amené à introduire des lignes brisées destinées à faire écran entre les points expérimentaux devant ou non être pris en compte par le krigeage. Il peut s'agir soit d'une discontinuité réelle (par exemple, faille géologique), soit d'un effet d'écran dû aux sinuosités du contour (exemple: répartition de polluant le long d'une côte, de part et d'autre d'un cap) (Fig.45).

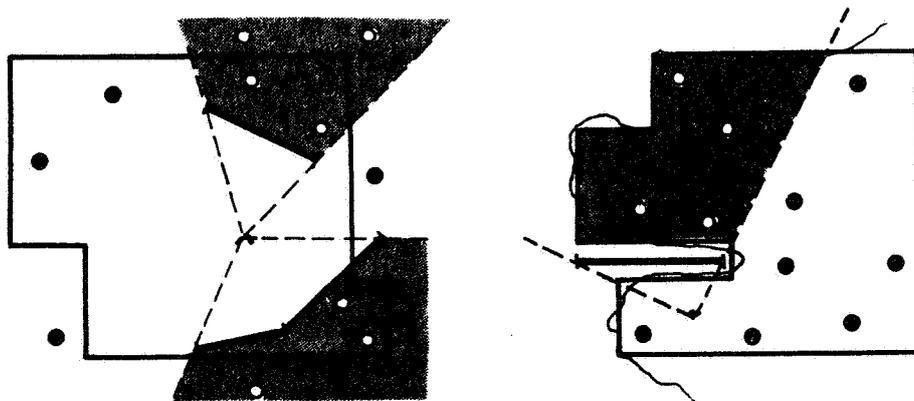


FIG. 45 - FAILLES GEOLOGIQUES ET EFFET D'ECRAN FORME PAR UN CAP

On trouvera, figure 46, un exemple d'application de cette technique à l'estimation de la cote du mur de la nappe bathonienne de Basse-Normandie.

COTE DU MUR DE LA NAPPE BATHONIENNE

MOYENNE PAR MAILLE

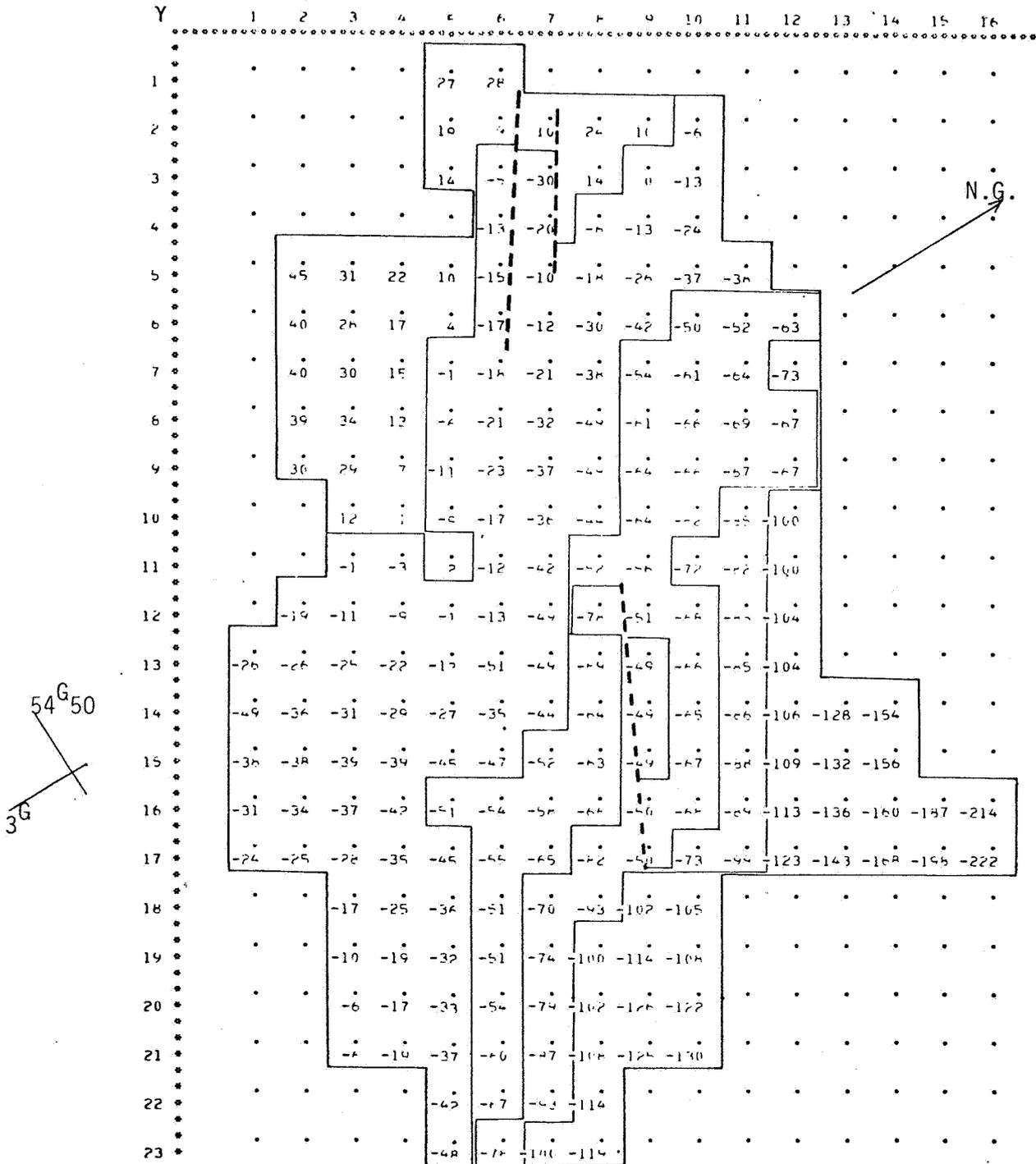


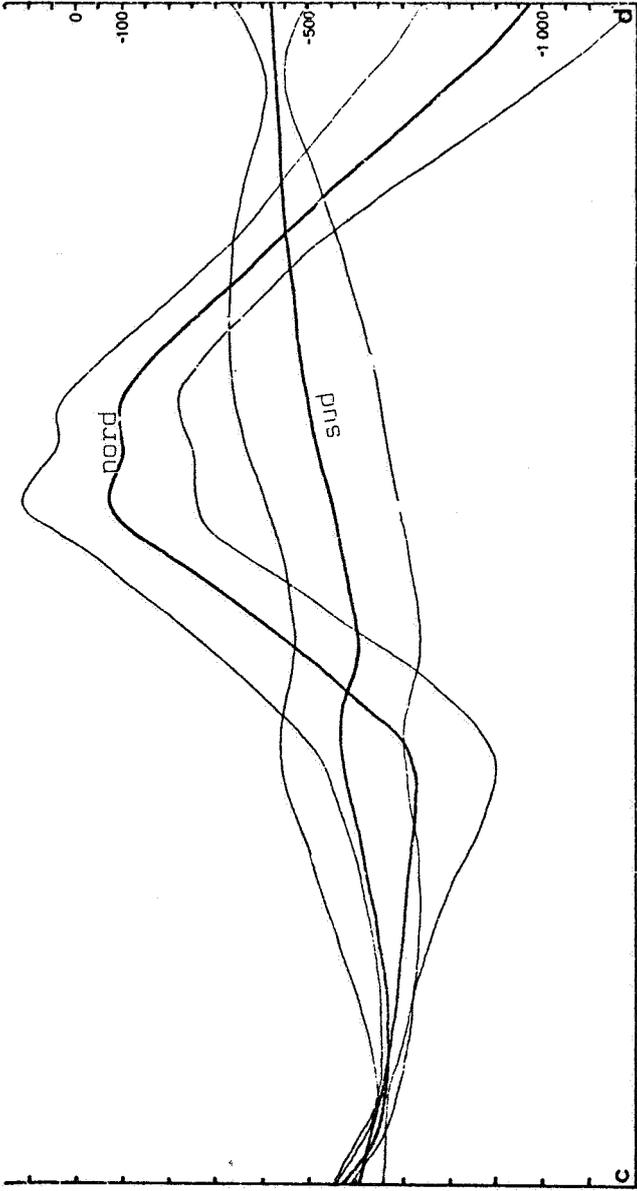
FIG. 46 - NAPPE BATHONIENNE DE BASSE-NORMANDIE  
PRISE EN COMPTE DES FAILLES DE RANVILLE ET AMBLIE  
DANS L'ESTIMATION DE LA COTE DU MUR

Le krigeage permet de tenir compte ainsi des failles connues. Il peut également aider le géologue à juger de l'existence d'une faille affectant une formation profonde, à partir des données de forage.

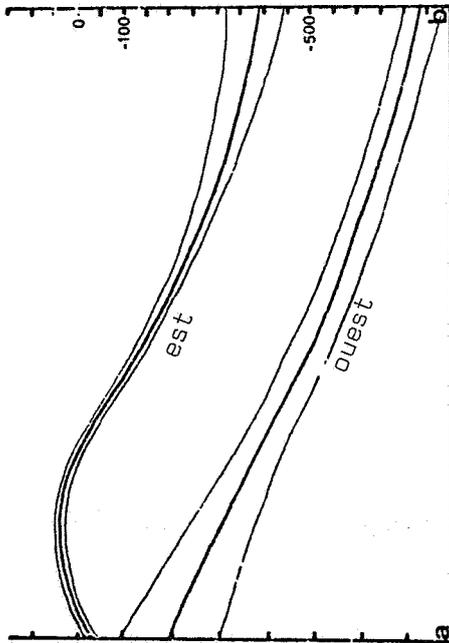
Cherchant à tester la pertinence de deux failles "possibles" dans une telle formation (réservoir calcaire d'Al Aziziyah, Libye), on a procédé, le long des lignes AB et CD (Fig. 47), à une double estimation par krigeage de la cote du toit de la formation en prenant en compte à chaque fois uniquement les forages situés d'un seul côté. On a reporté, dans les deux cas, sur un même graphique, les deux coupes obtenues ainsi de part et d'autre, assorties de leur fourchette d'incertitude à  $\pm 2\sigma$ . L'examen de la figure 47 montre que, si pour la coupe AB l'existence d'une faille est très nette, pour CD le rejet n'est pas significatif sur les 2/3 du profil, laissant ainsi planer un doute.

On peut examiner, pour terminer, les cartes krigées de la zone entourant la faille AB, obtenues respectivement en négligeant l'existence de celle-ci, puis en en tenant compte. On constate, sur la première carte, la présence d'ondulations parasites dues à l'impossibilité pour un modèle continu de franchir un ressaut aussi important (Fig. 48).

Ces différentes adaptations du krigeage ont pu montrer son aptitude à traiter toute l'étendue des problèmes d'estimation qui se posent dans les sciences de l'eau. Mais les ressources de la théorie des variables régionalisées permettent d'aborder d'autres problèmes comme l'optimisation des réseaux de mesure.



Profils CD (avec intervalles de confiance)



Profils AB (avec intervalles de confiance)

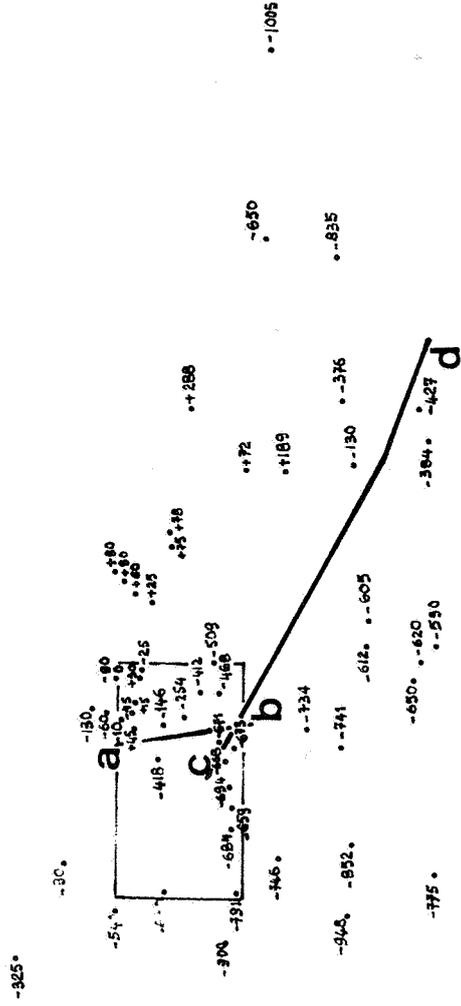


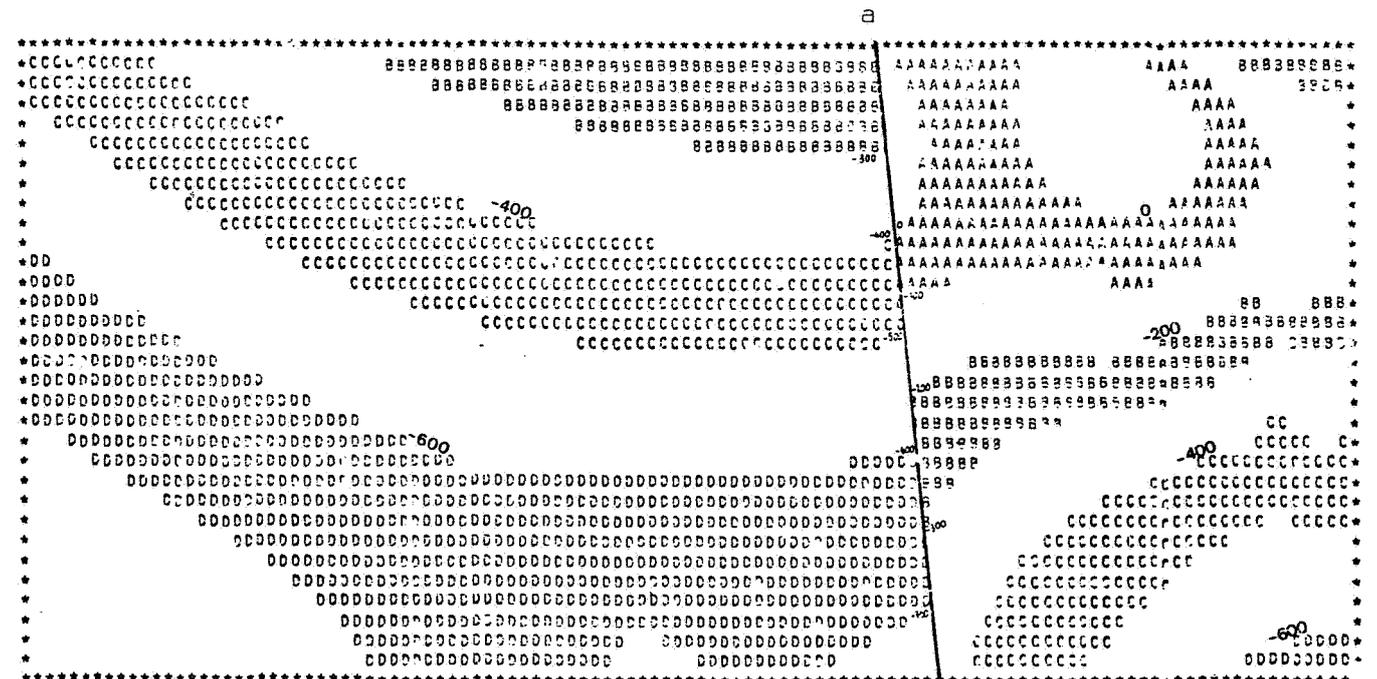
FIG. 47 - FORMATION GEOLOGIQUE D'EL AZIZIYAH  
TEST DE LA PERTINENCE  
DE DEUX FAILLES "POSSIBLES"

Cotes du toit de la formation et emplacement des failles

• -1070



Carte du toit de la formation tracée en négligeant l'existence de la faille



Carte du toit de la formation tracée en tenant compte de la faille

FIG. 48 - FORMATION GEOLOGIQUE D'AL AZIZIYAH  
 PROBLEME DE LA CARTOGRAPHIE AUTOMATIQUE  
 AU VOISINAGE D'UNE FAILLE

REFERENCES BIBLIOGRAPHIQUES - CHAP. III, § 1

- [1] CHILES, J.P., CHAUVET, P. (1975):  
Le krigeage: méthode de cartographie des fonds marins.  
Revue Hydrographique Internationale, vol. 52, n°1, pp. 29-45
- [2] DELFINER, P., DELHOMME, J.P. (1973):  
Optimum interpolation by kriging.  
Display and analysis of spatial data, pp. 96-114.  
(Proc. of NATO Adv. Study Inst., Nottingham, 1973)  
John Wiley & Sons.
- [3] DELFINER, P., DELHOMME, J.P., CHILES, J.P. (1976):  
BLUEPACK: notice d'utilisation.  
Ecole des Mines, C.M.M.-C.I.G.
- [4] DELHOMME, J.P. (1970):  
Présentation d'une méthode objective d'interpolation pour  
la construction de cartes isopiézométriques.  
Groupe d'Etude des Bassins Versants Souterrains, Douai, 1970.
- [5] DELHOMME, J.P. (1971):  
Traitement géostatistique des données piézométriques: le  
krigeage en hydrogéologie.  
Ecole des Mines, L.H.M., R71/14.
- [6] DELHOMME, J.P., DELFINER, P. (1973):  
Application du krigeage à l'optimisation d'une campagne  
pluviométrique en zone aride.  
Colloque UNESCO-OMM-AIHS sur l'élaboration des projets  
d'utilisation des ressources en eau sans données suffisantes,  
Madrid, 1973.  
Actes du Colloque, tome I, pp. 191-210.
- [7] DELHOMME, J.P. (1974):  
La cartographie d'une grandeur physique à partir de données  
de différentes qualités.  
Congrès d'Hydrogéologie, Montpellier, 1974.  
Mémoires de l'Assoc. Int. des Hydrogéologues, tome X, 1,  
pp. 185-194.
- [8] GILLILAND, J.A. (1972):  
Principles of groundwater data acquisition.  
Water Resour. Res., vol.8, n°1, pp. 182-187.
- [9] LEVASSOR, A., DELHOMME, J.P., TALBOT, A. (1976):  
Etude du bassin de l'Orne.  
Ecole des Mines, L.H.M.
- [10] MATHERON, G. (1974):  
Effet proportionnel et lognormalité ou le retour du serpent  
de mer.  
Note géostatistique n° 124.  
Ecole des Mines, C.M.M., N-374.
- [11] WALTERS, R.F. (1969):  
Contouring by machine: a user's guide.  
Amer. Ass. Petrol. Geol. Bull., vol.53, n°11, pp. 2324-2340.

## 2 - OPTIMISATION DU RENFORCEMENT D'UN RESEAU DE MESURES

De même que les divers problèmes d'estimation n'ont pu être résolus qu'après avoir été convenablement formulés, la question de l'optimisation d'un réseau de mesure n'a de sens que si le but poursuivi a été clairement défini.

Dans le cas d'une reconnaissance hydroclimatique en vue d'un aménagement hydraulique, l'hydrologue, ayant pour objectif l'étude de la relation pluie-débit, s'intéresse en premier lieu aux hauteurs d'eau moyennes tombées, journallement par exemple, sur le bassin. La variance d'estimation par krigeage a permis de donner la fourchette d'incertitude avec laquelle cette quantité peut être calculée.

Dans le cas d'une étude de la répartition spatiale des transmissivités ou d'une reconnaissance du modelé piézométrique, l'hydrogéologue s'intéresse aux moyennes par mailles (ou aux valeurs ponctuelles) de la grandeur étudiée. Ici encore, pour chaque maille, ou en chaque point, les variances de krigeage ont permis de définir la confiance devant être accordée à l'estimation.

### a) Méthode du point fictif

La variance de krigeage constitue donc tout naturellement l'indicateur de précision nécessaire pour:

- décider de l'opportunité de renforcer le réseau,
- déterminer l'emplacement optimal d'un éventuel point de mesure supplémentaire.

On mettra à profit le fait que la variance de krigeage ne fait intervenir que la structure du phénomène et la configuration géométrique des données  $\{x_i, i=1, \dots, n\}$  et du domaine à estimer ( $x_0, s_0$  ou  $S$ ), sans faire intervenir les valeurs  $z(x_i)$  elles-mêmes. Si on considère un point supplémentaire  $x_{n+1}$ , il est ainsi possible de bâtir le système de krigeage correspondant aux  $n+1$  points  $x_1, \dots, x_n, x_{n+1}$ , et d'en déduire la variance d'estimation relative à cette nouvelle configuration expérimentale, sans avoir à faire d'hypothèse

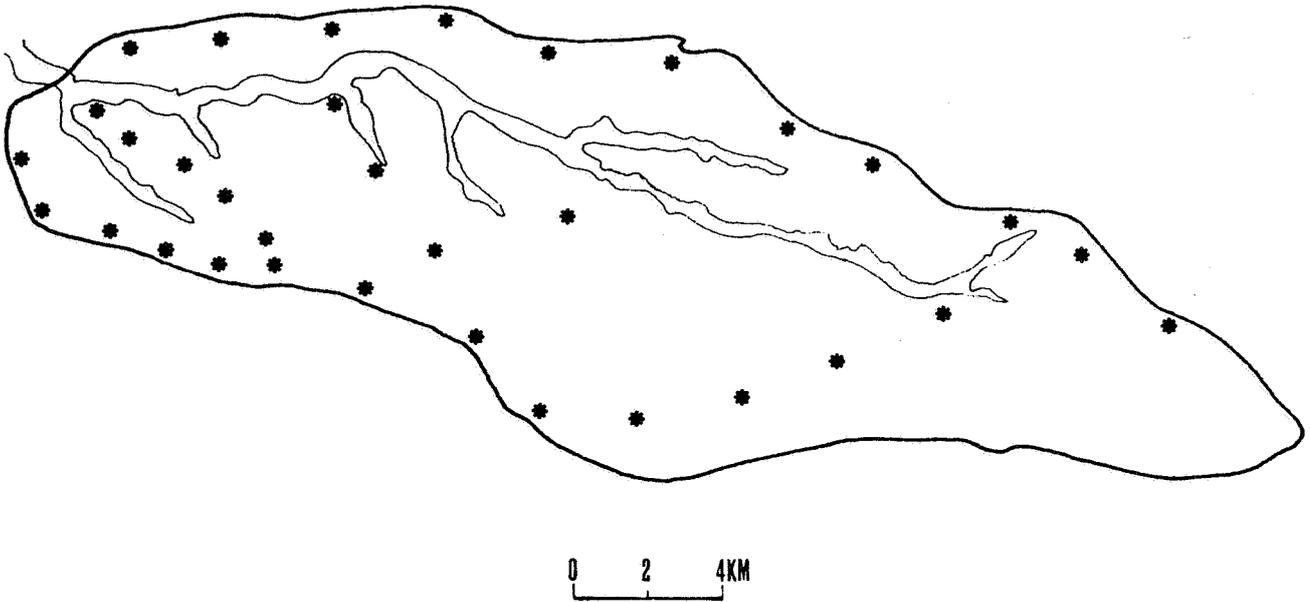


FIG. 49 - BASSIN VERSANT DE KADJEMEUR  
IMPLANTATION DU RESEAU PLUVIOMETRIQUE

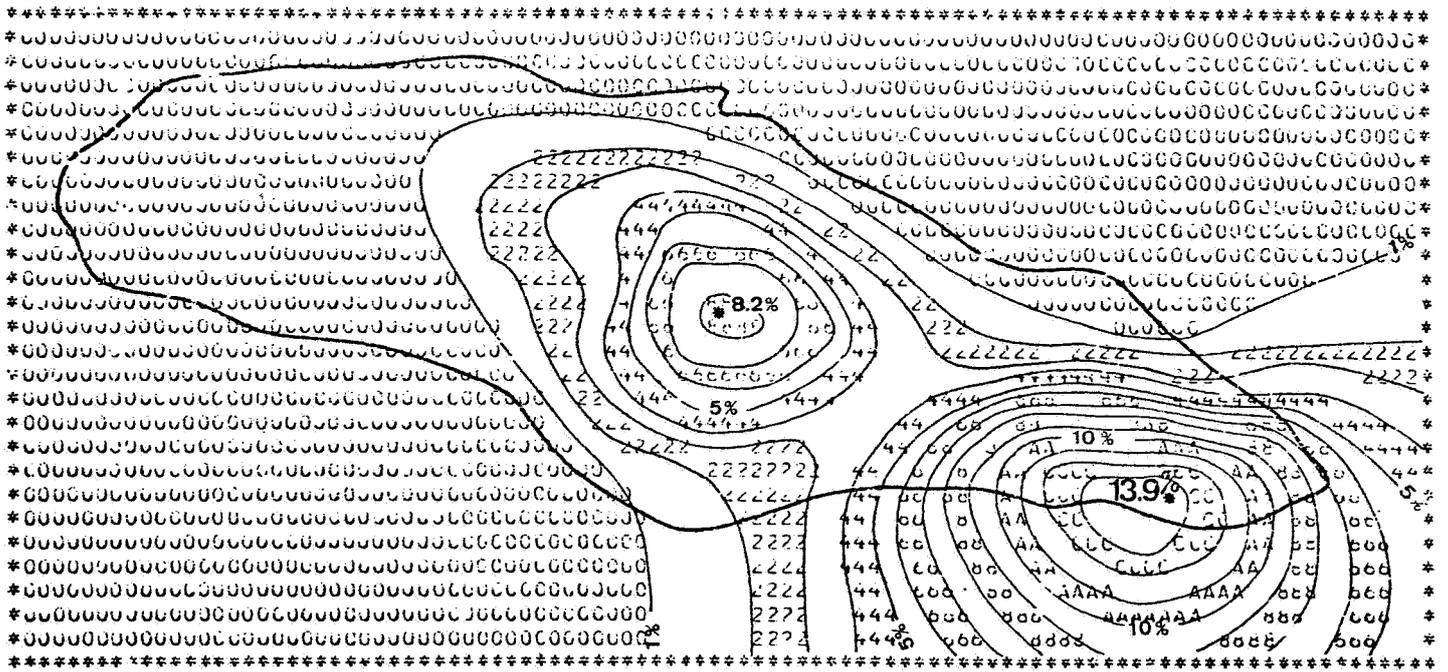


FIG. 50 - METHODE DU POINT FICTIF: COURBES D'EGAL GAIN EN PRECISION

a priori sur  $z(x_{n+1})$ . On donne à cette méthode le nom de "méthode du point fictif" [1]. On examinera successivement comment elle peut être utilisée dans les deux cas mentionnés plus haut.

b) Objectif global

Dans le cas où l'objectif est constitué par l'estimation d'une quantité moyenne sur un domaine S, on cherchera à déterminer quel est le point x où l'implantation d'un  $n+1^{\text{ème}}$  point de mesure aux côtés des n déjà existants conduit à la variance d'estimation la plus faible. Si  $\sigma_{k_0}^2$  désigne la variance d'estimation de départ avec les n points expérimentaux, et  $\sigma_k^2(x)$  (toujours  $< \sigma_{k_0}^2$ ) celle qui correspond à la configuration  $\{x_1, \dots, x_n, x\}$ , le gain en précision peut être défini comme:

$$G(x) = \frac{\sigma_{k_0}^2 - \sigma_k^2(x)}{\sigma_{k_0}^2}$$

Si le choix ne se pose que parmi un certain nombre de points présélectionnés selon un autre critère (accès facile, relevé aisé, etc...), on plantera fictivement un point de mesure en chacun d'eux, et on déterminera les gains en précision correspondants. On procédera de même le long d'un cheminement si l'on a décidé a priori de retenir une ligne caractéristique du terrain (voie d'accès, accident de terrain, etc...).

Quand, au contraire, l'on ne possède aucun a priori, on fera parcourir au point fictif x une portion de plan englobant le domaine S et ses abords immédiats, et on tracera une carte d'isogains en précision sur l'estimation de la valeur moyenne sur le domaine S. C'est la solution qui a été adoptée pour l'exemple de l'Ouadi Kadjemeur (Fig.50) [1].

Si l'on n'avait tenu aucun compte de l'existence de corrélations spatiales, le gain correspondant à un  $34^{\text{e}}$  pluviomètre aurait été de  $1/34 \approx 3\%$  et identique pour tout emplacement à l'intérieur du bassin.

L'examen comparé des figures 49 et 50 montre que la méthode utilisée est riche d'enseignements. Si le gain reste inférieur à 1% pour la partie ouest du bassin, qui comporte déjà de nombreux pluviomètres, le gain maximum dépasse 13% au lieu des 3% par une analyse sommaire. Et, contrairement

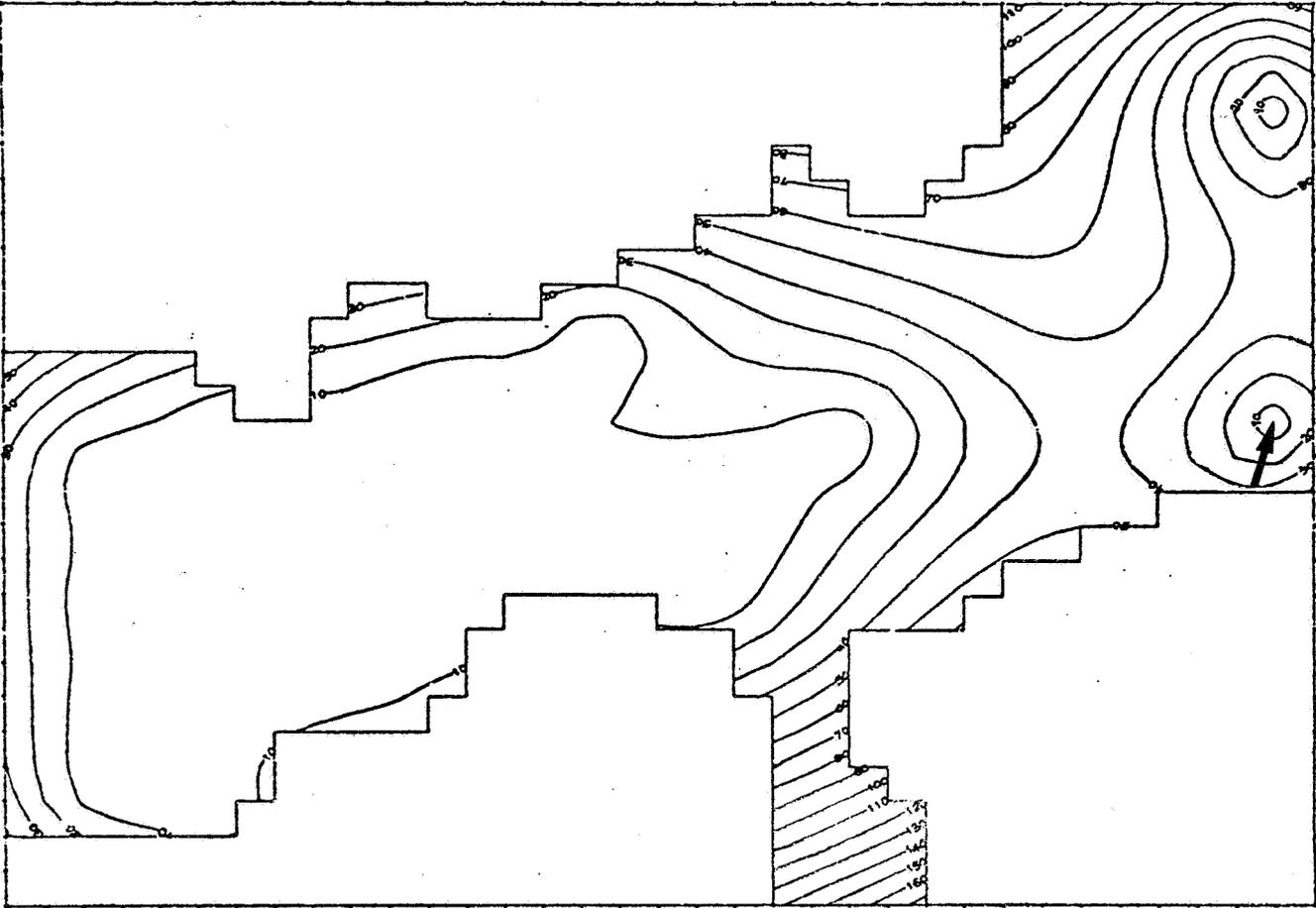


FIG. 52 - ECARTS-TYPES DE KRIGEAGE APRES ADDITION  
D'UN PIEZOMETRE SUPPLEMENTAIRE

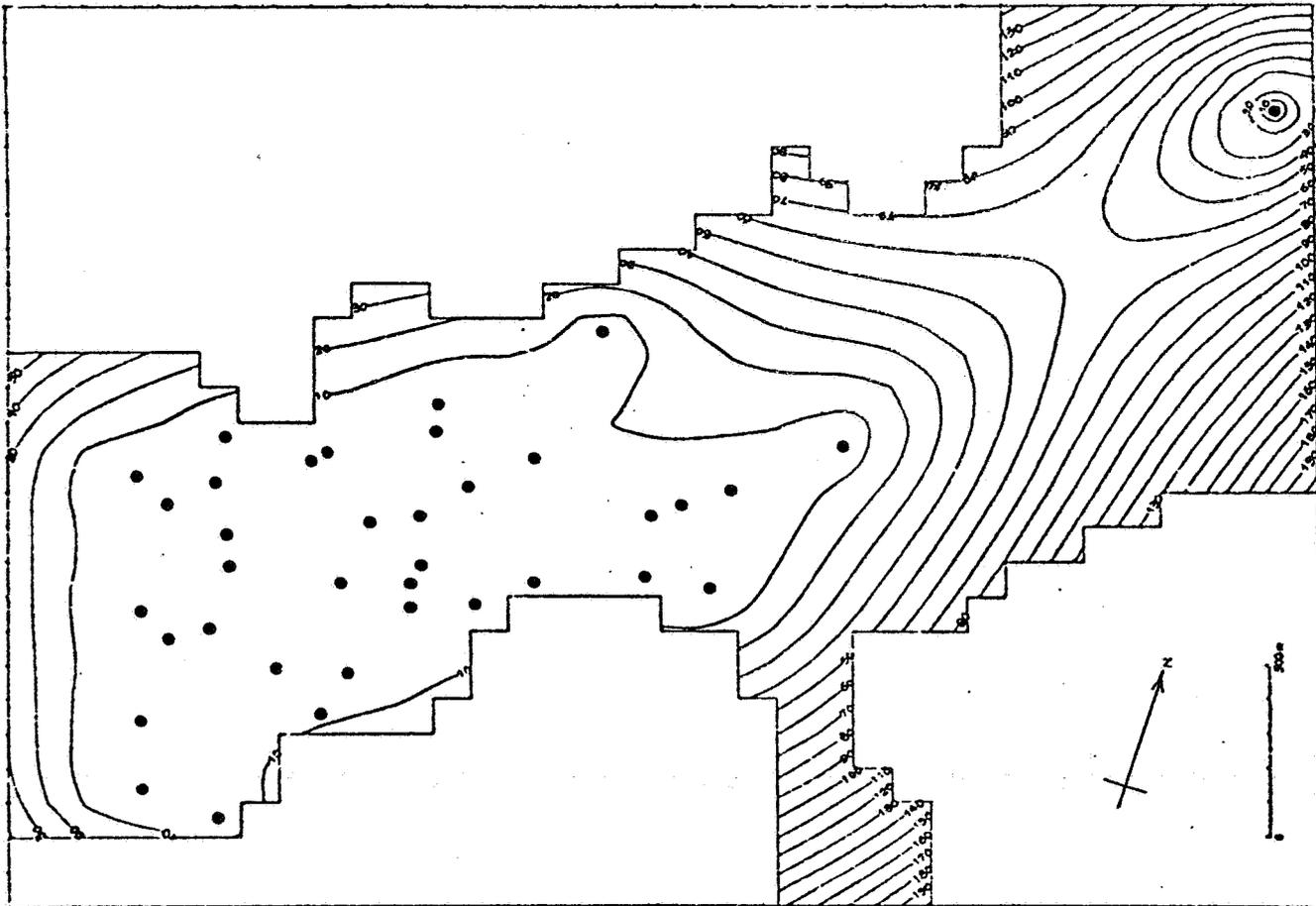


FIG. 51 - RESEAU PIEZOMETRIQUE DE L'HUVEAUNE:  
ECARTS-TYPES DE KRIGEAGE (en cm)

à ce que l'on aurait pu penser, le meilleur endroit pour implanter un pluviomètre supplémentaire n'est pas la partie centrale du bassin, mais sa partie sud-est. A la limite, un point de mesure extérieur au bassin, mais dans la région sud-est, est plus informatif qu'un point central. La raison en est que la frange sud-est n'est pas reconnue du tout, alors que la région centrale est, elle, bien cernée.

Si l'on revient aux poids de krigeage ou à ceux de THIESSEN (Fig. 34 et 35, p.88), on constate que ces résultats vont dans le sens d'un "soulagement" des pluviomètres de poids les plus élevés. On tend vers une égalisation, en quelque sorte, de la contribution des différentes stations, ce qui satisfait le sens physique de l'hydrologue.

#### c) Objectif local

Dans le cas où l'objectif est constitué par une connaissance locale (ponctuelle, comme en cartographie, ou encore par maille), le problème se pose d'une manière légèrement différente, en ce sens qu'il y a maintenant autant de pondérations et de variances de krigeage qu'il y a de noeuds ou de mailles sur la grille utilisée. L'implantation d'un point de mesure supplémentaire aura une influence sur les variances de krigeage de tous les points (ou mailles) qui feront intervenir ce point dans leur pondération du krigeage. Cette influence, allant en s'affaiblissant avec la distance, ne se fera pourtant sentir en pratique que dans un certain voisinage.

On prendra l'exemple du réseau piézométrique de la nappe alluviale de l'Huveaune, à Marseille. On a superposé sur la Fig.51 les emplacements des différents points de mesure existants à la carte d'écart-types d'estimation par krigeage. On suppose l'existence d'un piézomètre supplémentaire dans la partie nord de la nappe, où il n'y avait qu'un seul piézomètre isolé. En comparant la carte d'écart-type de krigeage alors obtenue (Fig.52) à celle qui correspondait à la configuration expérimentale réelle, on peut juger de l'influence de ce piézomètre fictif: dans toute la zone l'entourant immédiatement, les écart-types, primitivement de presque 2 m, diminuent considérablement; l'influence reste malgré tout assez localisée, avec, en particulier, une baisse très faible dans la corne sud-est de l'aquifère.

La démarche à suivre pour renforcer un réseau est donc ici d'implanter des points de mesure supplémentaires dans toutes les zones où existent, sur la carte d'écart-type de krigeage, des maxima locaux très marqués, de façon à les faire disparaître.

Ce problème s'étend aux cas de données de différentes qualités, soit parce que l'on peut envisager deux types d'appareil de mesure, soit parce qu'on peut employer, par exemple, deux techniques, l'une directe, l'autre indirecte, d'acquisition de la donnée pour le point expérimental supplémentaire.

d) Influence de la qualité de la station supplémentaire

Pour illustrer cette question, on a repris ici l'exemple des données situées aux sommets d'un hexagone [3]. Cinq points de mesure (supposés sans erreur) étant en place, on a étudié l'information apportée par le sixième point, affecté d'une incertitude de variance  $\sigma_6^2$ , en regardant la réduction relative de la variance de krigeage due à sa présence. Cette fois-ci on se propose d'estimer  $z(x)$  en tout point  $x$  (et non plus seulement au centre de l'hexagone). En utilisant seulement les cinq premières données, on aurait une variance de krigeage  $\sigma_{k_5}^2$ ; en utilisant aussi la sixième, on aura  $\sigma_{k_6}^2$  ( $< \sigma_{k_5}^2$ ). Sur la figure 53, on a cartographié la réduction relative de variance (ou gain en précision):

$$g = \frac{\sigma_{k_5}^2 - \sigma_{k_6}^2}{\sigma_{k_5}^2}$$

pour différentes valeurs de  $\sigma_6^2$ .

Quand  $\sigma_6^2 = 0$  (cas d'une donnée normale), le gain maximum est de 100% en  $x_6$  (puisque alors  $\sigma_{k_6}^2 = 0$ , d'après une propriété du krigeage signalée au Chap. II, § 3d), ce gain allant en décroissant quand le point krigé s'éloigne de  $x_6$ .

Pour des valeurs de plus en plus élevées de  $\sigma_6^2$ , la zone d'influence de la sixième donnée va en se rétrécissant, tendant à la limite à disparaître quand  $\sigma_6^2$  devient infinie. Une donnée totalement incertaine n'apporterait aucune information.

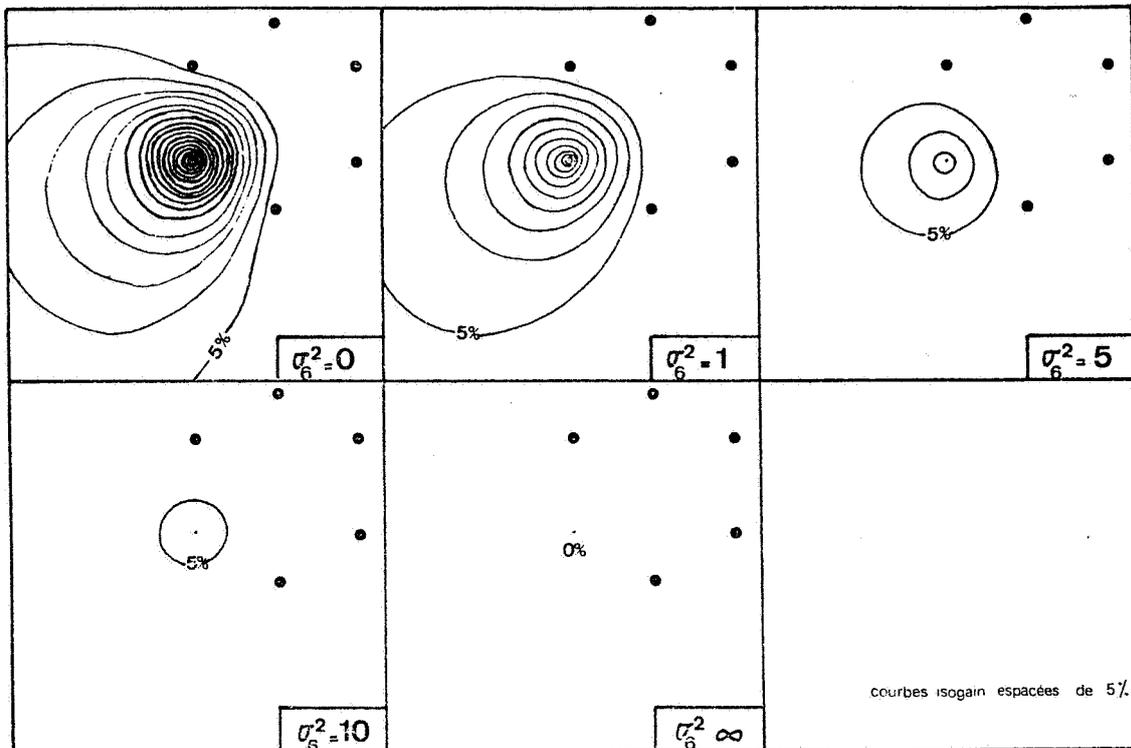


FIG. 53 - INFLUENCE DE L'INCERTITUDE AFFECTANT LA DONNEE SUPPLEMENTAIRE  
SUR LE GAIN EN PRECISION

e) Recherche d'un optimum économique

Le krigeage permet donc de quantifier le gain d'information qu'apporterait un éventuel point de mesure supplémentaire, et ce, en prenant en compte la qualité de la mesure en question, qualité d'ailleurs souvent liée au coût. Or, comme l'ont souligné GANDIN et KAGAN [4], qui avaient déjà adopté la variance de l'erreur d'interpolation comme critère de calcul, le problème de l'optimisation d'un réseau ne saurait être séparé de son aspect économique.

Si on associe au coût d'implantation d'une station nouvelle le gain en précision correspondant qui est déterminé par le krigeage, on peut faire un calcul d'optimum économique et orienter la poursuite d'une campagne de mesures. Par exemple, en ce qui concerne la transmissivité, on pourra décider de faire soit un essai de pompage pour avoir une bonne mesure, soit plusieurs simples déterminations du débit spécifique beaucoup moins coûteuses, mais aussi moins précises, et pour chaque éventualité on pourra déterminer les emplacements optimaux [9].

On a vu jusqu'à présent que le krigeage a des applications multiples en hydrogéologie. Ces possibilités ont été encore enrichies par un développement ultérieur de la théorie des variables régionalisées: la méthode des simulations conditionnelles.

REFERENCES BIBLIOGRAPHIQUES - CHAP. III, § 2

- [1] DELHOMME, J.P., DELFINER, P. (1973):  
Application du krigeage à l'optimisation d'une campagne pluviométrique en zone aride.  
Colloque UNESCO-OMM-AIHS sur l'élaboration des projets d'utilisation des ressources en eau sans données suffisantes, Madrid, 1973.  
Actes du Colloque, tome I, pp. 191-210
- [2] DELHOMME, J.P., LEVASSOR, A. (à paraître):  
Reconnaissance optimale du champ des transmissivités.
- [3] DELHOMME, J.P. (1974):  
La cartographie d'une grandeur physique à partir de données de différentes qualités.  
Congrès d'Hydrogéologie, Montpellier, 1974.  
Mémoires de l'Assoc. Int. des Hydrogéologues, tome X, 1  
pp. 185-194.
- [4] GANDIN, L.S., KAGAN, R.L. (1967):  
The economic approach to the planning of a network of meteorological stations.  
Soviet Hydrology: selected papers n°6  
Traduit de: Trudy GGO, n° 208, pp. 120-131

### 3 - SIMULATIONS CONDITIONNELLES

On a vu que, appliqué à la cartographie, le krigeage donne en chaque point de la carte la meilleure estimation linéaire possible et l'écart-type de l'erreur associée, à partir desquels on peut bâtir un intervalle de confiance pour la valeur réelle (qui demeure inconnue). Il peut être intéressant, en premier lieu, de visualiser la latitude qui est ainsi laissée à la surface, et de voir les différentes allures que pourraient avoir en réalité les courbes isovales.

Mais surtout, la carte krigée est plus lisse que la "carte réelle" parce qu'un estimateur, même optimal, ne peut restituer des détails qui n'ont pas été observés. L'intérêt peut ainsi résider également dans la visualisation de la variabilité du phénomène.

Cette idée de chercher à produire des valeurs ayant la même variabilité que le phénomène réel sous-tend l'analyse classique de sensibilité, ou "error analysis", utilisée dans les modèles numériques. A partir des histogrammes expérimentaux, par exemple de transmissivité, on affecte, de façon aléatoire, des valeurs équiprobables à tous les noeuds du maillage, et ce, autant de fois qu'on veut constituer de jeux de transmissivité. Mais cette méthode ignore ainsi complètement la structure spatiale de la variable. Le modèle des fonctions aléatoires, que nous avons jusqu'ici utilisé, semble pouvoir conduire à de meilleurs résultats.

#### a) Principe

On a, en effet, considéré le phénomène étudié comme une réalisation particulière d'une F.A. Ceci constituait alors un moyen de poser le problème dans un cadre probabiliste et de pouvoir ainsi utiliser des outils comme l'espérance mathématique, la variance, le variogramme ou les covariances généralisées.

L'idée avec les "simulations conditionnelles" (le terme "simulation" doit être pris ici dans son acception statistique, et n'a aucun lien avec les techniques de simulation utilisées pour prévoir le comportement des aquifères) est d'exhiber d'autres réalisations de cette même F.A.

Par construction, ces autres réalisations auront même variabilité spatiale que le phénomène réel (même variogramme ou même covariance généralisée). Mais, en outre, on leur imposera de conserver aux points expérimentaux les valeurs mesurées, de façon à assurer une bonne coïncidence avec ce qui est connu du phénomène.

Une simulation conditionnelle est donc définie comme une surface qui:

- a la même variabilité que le phénomène étudié,
- passe par les points expérimentaux.

On peut, en outre, dans le cas sans dérive, imposer que l'histogramme des valeurs simulées reflète la même loi de dispersion que l'histogramme des données (lognormal par exemple si l'on travaille sur des transmissivités)[5].

Une telle simulation n'est évidemment pas la réalité, mais seulement une version de celle-ci. Les diverses simulations conditionnelles présentent des variantes possibles de la carte réelle (inconnue).

Si l'on construit un grand nombre de ces variantes, la moyenne de leurs valeurs en un point redonnera la valeur estimée par krigeage, et leur variance la variance d'estimation correspondante.

La méthode est basée sur le principe suivant. Partant de la décomposition triviale:

$$\begin{aligned} \text{valeur réelle} &= \text{valeur krigée} + \text{erreur de krigeage} \\ z(x) &= z^*(x) + (z(x) - z^*(x)) \end{aligned}$$

où l'erreur reste inconnue puisque la valeur réelle  $z(x)$  n'est pas disponible en dehors d'un point expérimental, on simulera ces erreurs  $(z(x) - z^*(x))$  en les empruntant à une autre réalisation, que l'on connaîtra, elle, de façon exhaustive, pour l'avoir engendrée.

On est en effet capable d'obtenir une réalisation  $s(x)$  d'une F.A.  $S(x)$  ayant même variogramme ou même covariance généralisée que  $Z(x)$ , en utilisant la technique des bandes tournantes [4], [6], [8]. Cette technique est celle qui a été d'ailleurs employée pour engendrer les réalisations de F.A.I. d'ordre 0, 1 et 2, montrées à la figure 26 (p.74).

b) Méthode des bandes tournantes

Elle consiste à sommer un grand nombre de réalisations indépendantes définies sur des droites balayant le plan, se ramenant ainsi à des simulations à une dimension, beaucoup plus simples à obtenir.

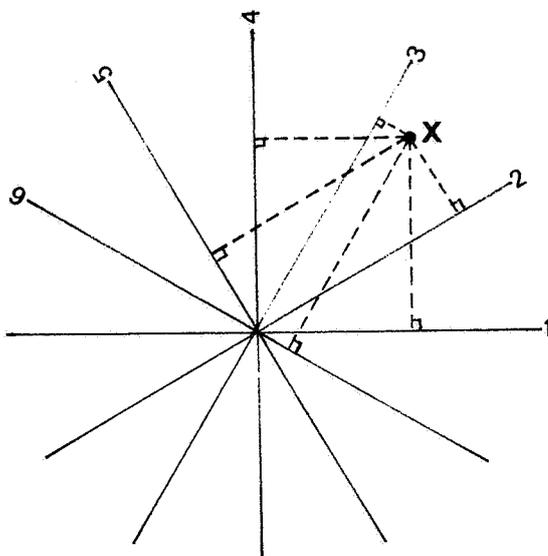


FIG. 54 - METHODE DES BANDES TOURNANTES  
(schéma de principe)

Pour simuler la valeur en tout point  $x$  du plan, on ajoute les valeurs prises, aux projections de  $x$  sur les différentes droites, par les simulations à une dimension correspondantes. On montre par exemple que, en simulant sur les droites des réalisations d'une F.A. à une dimension ayant un variogramme en  $h^\lambda$ , on obtient également un schéma en  $h^\lambda$  dans le plan [4]. Cette méthode peut de même être utilisée pour une F.A. ayant une covariance généralisée polynômiale dans le plan [6], en simulant sur les droites des réalisations d'une F.A. à une dimension admettant elle-même une covariance généralisée polynômiale.

c) Aperçu de la théorie

Sur une telle simulation  $s(x)$ , il est possible, avec la même configuration de points expérimentaux  $x_i$  que dans la réalité, et les valeurs  $s(x_i)$  simulées en ces points, de calculer au point  $x$ , avec les mêmes poids  $\lambda_i$  que pour  $z^*(x)$ , la valeur krigée  $s^*(x)$ . Dans l'égalité:

$$s(x) = s^*(x) + (s(x) - s^*(x))$$

l'erreur  $(s(x) - s^*(x))$  est, cette fois, connue exactement.

Considérant la somme:  $z_s(x) = z^*(x) + (s(x) - s^*(x))$

on peut montrer que  $z_s(x)$  constitue bien une simulation conditionnelle au sens défini au § a).

Le fait qu'elle coïncide avec  $z(x)$  en tout point expérimental  $x_i$  provient directement de ce que le krigeage est un interpolateur exact. On a:

$$z^*(x_i) = z(x_i)$$

$$s^*(x_i) = s(x_i)$$

et finalement:  $z_s(x_i) = z(x_i)$

Le fait qu'elle ait également même covariance généralisée a été démontré par ailleurs [2] en faisant appel à d'autres propriétés de l'estimateur de krigeage, et sera admis ici.

Le rapport des simulations conditionnelles au krigeage, signalé au § a), se démontre comme suit.

En un point  $x$  quelconque,  $z^*(x)$  étant fixé, on considère:

$$Z_s(x) = z^*(x) + [S(x) - S^*(x)]$$

on a  $E [Z_s(x)] = z^*(x) + E[S(x) - S^*(x)]$

le krigeage étant sans biais:

$$E [S(x) - S^*(x)] = 0$$

et l'on a:  $E [Z_s(x)] = z^*(x)$

de même:  $\text{var} [Z_s(x)] = E [(Z_s(x) - z^*(x))^2] = E [(S(x) - S^*(x))^2]$   
 $= \sigma_k^2$

#### d) Exemples

On examinera maintenant deux applications de cette méthode dans le domaine des sciences de l'eau. Le premier exemple est constitué par trois simulations conditionnelles d'un orage sur le bassin de l'Ouadi Kadjemeur déjà cité. Sur les cartes de la figure 55, l'isohyète 40 mm a été figuré d'un trait plus épais pour faciliter la lecture.

Alors que la carte krigée était, on l'a vu, entièrement lisse, ces cartes présentent un aspect plutôt bosselé, qui s'accorde bien avec ce à quoi on pouvait s'attendre au sujet de la variabilité spatiale d'un orage. La structure cellulaire des précipitations est traduite d'une manière assez comparable à celle qui a pu être observée ailleurs dans la réalité grâce aux échos radar (Fig.13, p.29).

On peut noter également que les trois variantes diffèrent notablement (de 10 à 60 mm) dans la partie est du bassin qui a été peu reconnue et où les écarts-types de krigeage étaient importants.

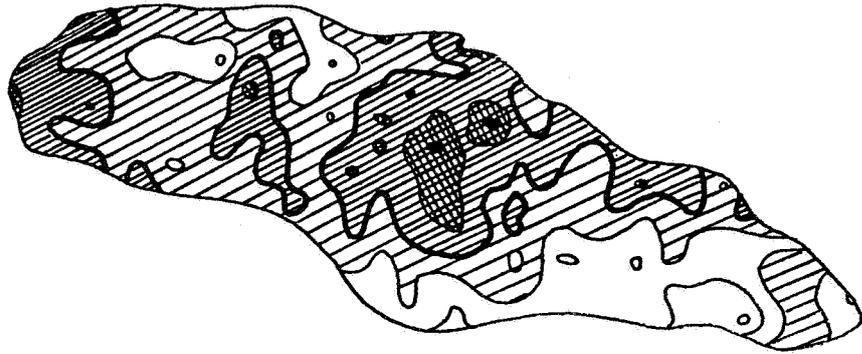
Le deuxième exemple, obtenu à partir de données piézométriques, montre encore mieux le rapport des simulations conditionnelles au krigeage. On reprendra le cas de la nappe alluviale de l'Huveaune, qui, bien que réel, se présente avec la clarté d'un cas d'école.

On a repris, à côté de la carte krigée (Fig.56), la carte d'écart-type correspondante (Fig.57). L'aquifère est, on l'a vu, connu par de nombreux piézomètres à l'aval (d'où une zone où les écarts-types sont inférieurs à 10 cm), mais par un seul piézomètre à l'amont. On peut constater que la précision, excellente à l'aval, décroît rapidement à mesure que croît la distance aux points expérimentaux. Les figures 58 et 59

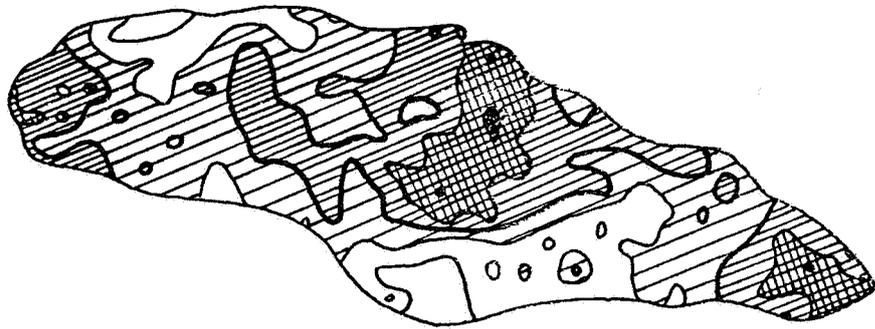
comparent l'estimation par krigeage à deux simulations conditionnelles. Celles-ci tombent toujours à l'intérieur de la fourchette à  $\pm 2\sigma$  autour du krigeage. Quand l'écart-type d'estimation  $\sigma$  est faible, toutes les simulations conditionnelles sont approximativement confondues avec le résultat du krigeage (cf. moitié aval). Elles passent toutes par les points expérimentaux (cf. le piézomètre isolé de l'amont). Mais, quand  $\sigma$  est grand, les simulations peuvent différer de façon importante, en demeurant cependant compatibles avec les données disponibles (cf. moitié amont).

Ce type d'outil permet ainsi de rechercher parmi les modèles possibles celui qui semble le mieux explicable à partir du contexte hydrogéologique. Il permet également de mieux visualiser toute la latitude qui est laissée lors du calage d'un modèle.

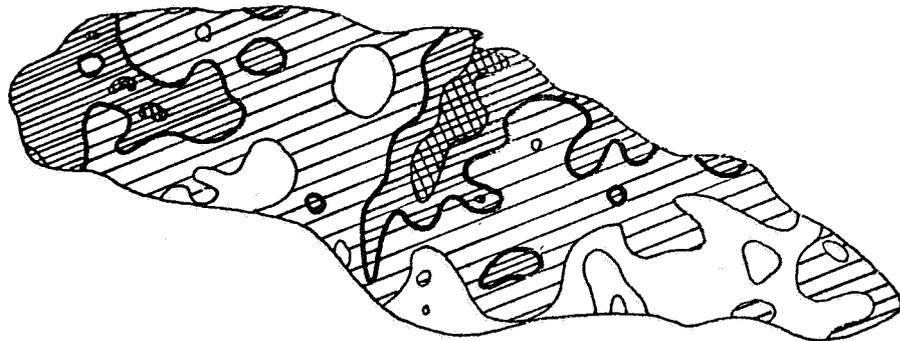
Ce dernier problème est au centre de la pratique actuelle en hydrogéologie mathématique, et il convient, avant de conclure, d'examiner comment les résultats précédents permettent de l'éclairer d'un jour nouveau.



variante n°1

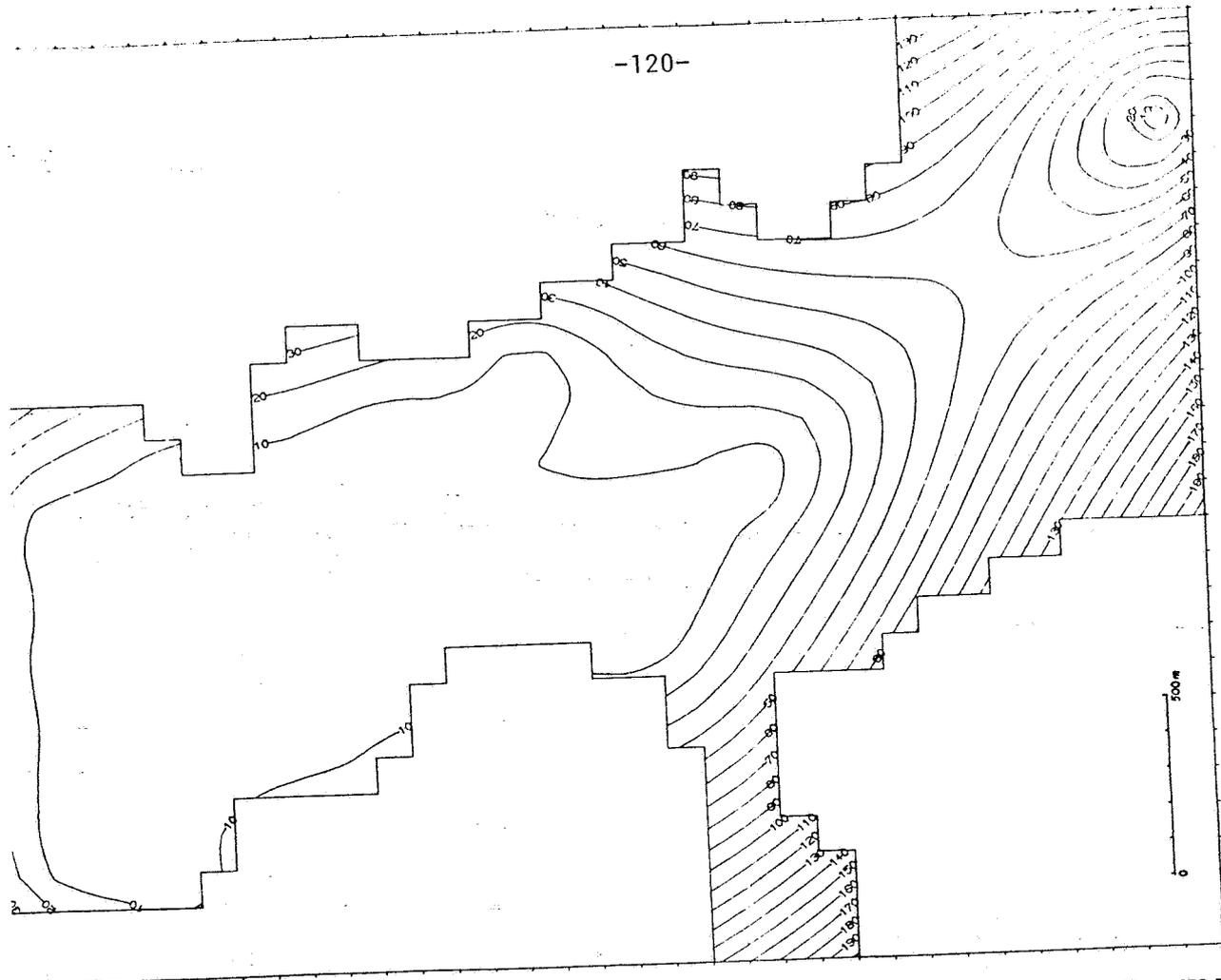


variante n°2



variante n°3

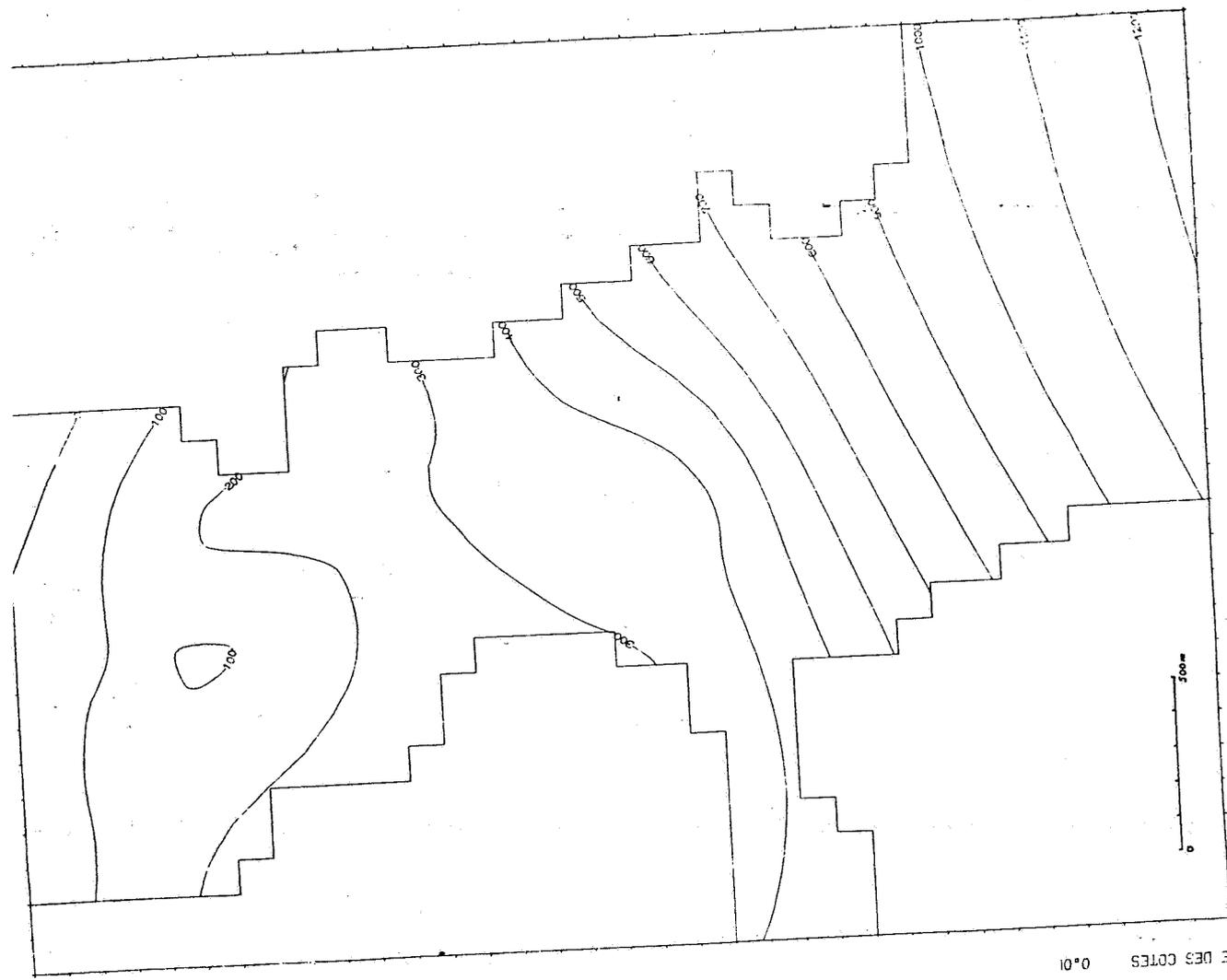
FIG. 55 - BASSIN-VERSANT DE KADJEMEUR  
TROIS SIMULATIONS CONDITIONNELLES  
DE L'AVERSE DU 6/8/66



-120-

LIMITES DES COTES 0.01  
 PIÈCE-HUVEAUNE ECART-TYPE

ECARTS-TYPES D'ESTIMATION (en cm)



LIMITES DES COTES 0.01  
 PIÈCE-HUVEAUNE KRIGEAGE

ET

FIG. 56 & 57 - NAPPE DE L'HUVEAUNE: CARTE KRIGEE

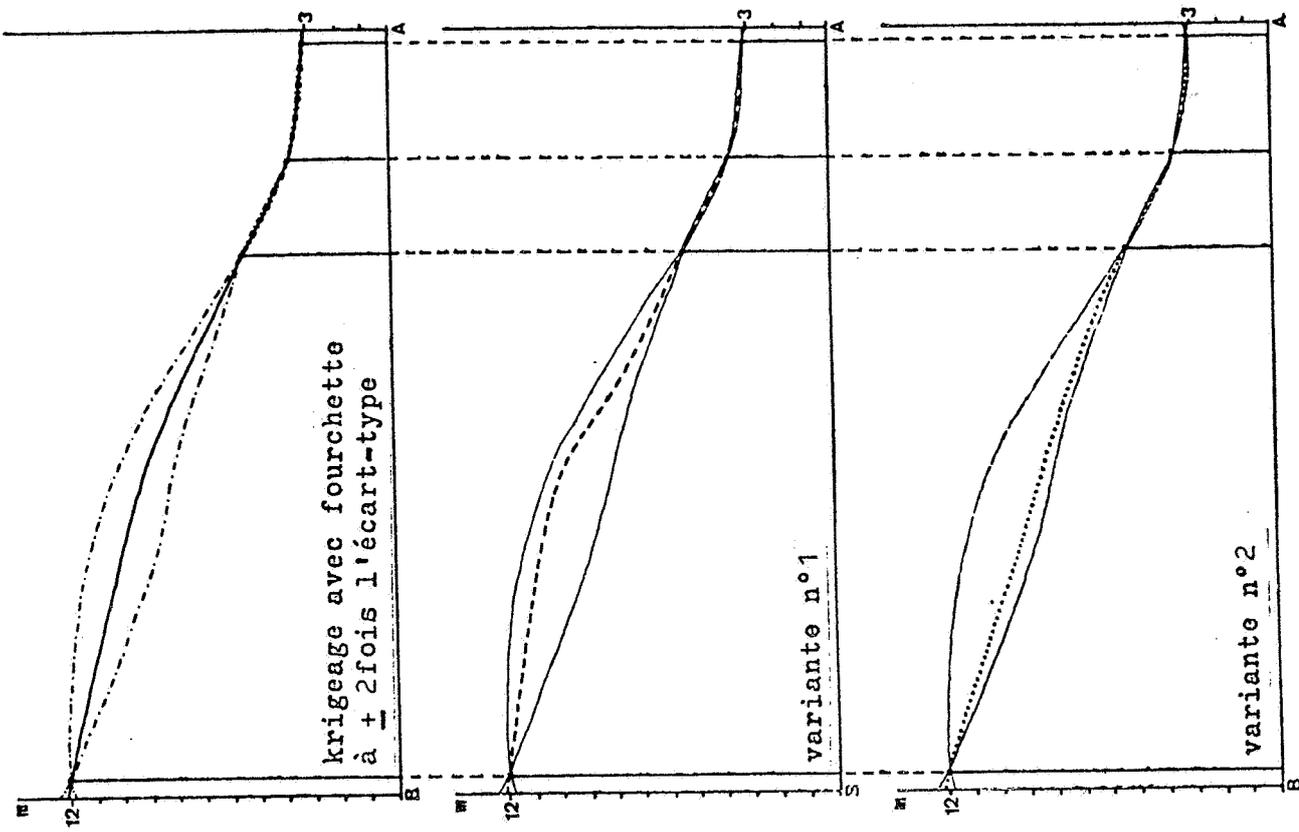


FIG. 59 - COMPARAISON SUR LE PROFIL AB DU KRIGEAGE ET DE DEUX SIMULATIONS CONDITIONNELLES

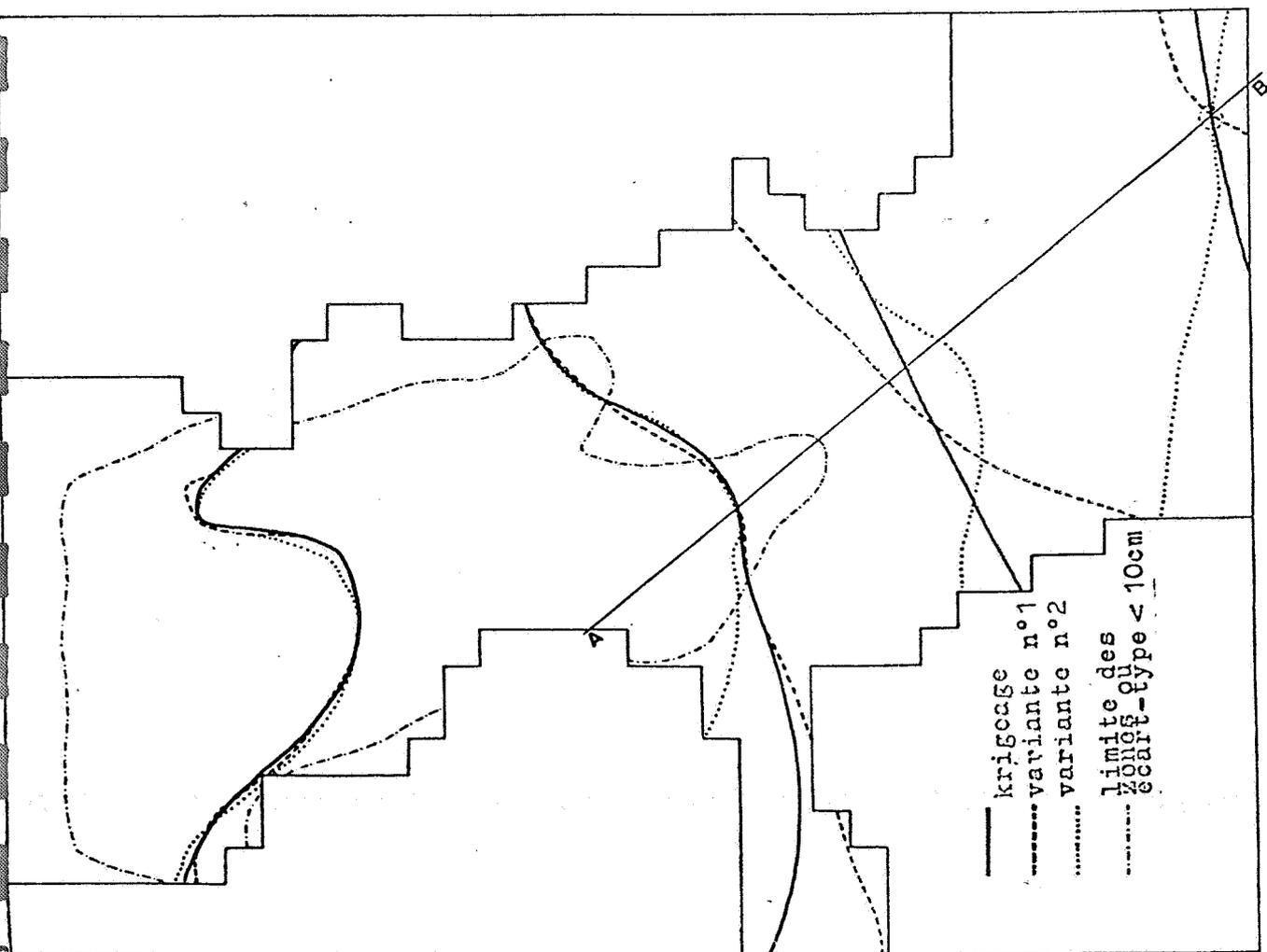


FIG. 58 - COMPARAISON ENTRE LA CARTE KRIGEE ET DEUX SIMULATIONS CONDITIONNELLES

#### 4 - IDENTIFICATION DES SYSTEMES HYDROGEOLOGIQUES

Le calage d'un modèle mathématique de nappe souterraine consiste à essayer de reproduire des états piézométriques observés, en ajustant différents paramètres (transmissivités, débit d'infiltration, etc...). Cet ajustement se fait généralement par essais et erreurs, en comparant les cartes piézométriques tracées à partir des résultats des calculs et la carte de référence.

Le krigeage permet d'établir, à partir des données de terrain, un document qui pourra être comparé directement aux résultats du modèle; il permet en outre de définir un critère d'adéquation entre le modèle et la réalité.

Le procédé classique de superposition des deux jeux d'isopièzes est trop subjectif pour conduire à un critère formulable de façon mathématique. De plus, il est à noter que dans les modèles, la cote piézométrique attribuée à une maille est représentative plus d'une valeur moyenne sur la maille que d'une valeur ponctuelle au centre de celle-ci. On a alors intérêt à estimer directement par krigeage les valeurs moyennes par maille, cette estimation étant toujours assortie de la fourniture d'un écart-type d'estimation. Disposant ainsi d'un indicateur de la précision de la connaissance piézométrique sur les différentes mailles, on peut utiliser cette information pour bâtir un critère objectif d'adéquation du calage [1].

L'écart de calage, qui peut être défini comme la différence entre la piézométrie calculée sur le modèle et celle obtenue par krigeage, doit être compatible avec la valeur de la fourchette d'incertitude correspondante. Si l'on adopte la fourchette conventionnelle à  $\pm 2\sigma$ , on jugera le calage correct dans les zones où l'écart de calage réduit, c'est-à-dire rapporté à l'écart-type d'estimation, est en valeur absolue inférieur à 2. On rejoint là une intuition qui veut que l'on affine le calage dans les zones bien reconnues, et que l'on puisse se contenter d'un calage approximatif dans les zones mal reconnues.

On peut encore exprimer le critère adopté sous la forme:

$$\max_i \left| \frac{z_i^* - z_i^c}{\sigma_{k_i}} \right| \leq 2$$

l'indice  $i$  décrivant l'ensemble des différentes mailles, et  $z_i^c$  représentant la valeur obtenue sur le modèle à la maille  $i$ .

On trouvera, figures 60 et 61 la répartition des écarts de calage et des écarts réduits pour un même état du calage de la nappe de Kairouan. On remarquera que le passage à l'écart réduit a permis de mieux focaliser les zones où il subsiste un problème. Le calage peut ainsi être considéré comme satisfaisant sur les limites ouest et dans la partie sud de la nappe, où pourtant on trouve des écarts de calage supérieurs parfois à 5 mètres: la piézométrie  $y$  est en fait assez mal connue. Mais, par contre, on constate que, dans l'état du calage considéré, la bordure est de la nappe pose encore un problème (conditions aux limites ou relation avec d'autres nappes).

L'effort doit être poursuivi dans cette voie de l'identification des systèmes hydrogéologiques, où le recours aux méthodes probabilistes devrait permettre de sortir d'une relative stagnation, en donnant une impulsion nouvelle.



ETUDE HYDROGEOLOGIQUE  
DE LA PLAINE DE KAIROUAN  
SUR MODELES MATHÉMATIQUES

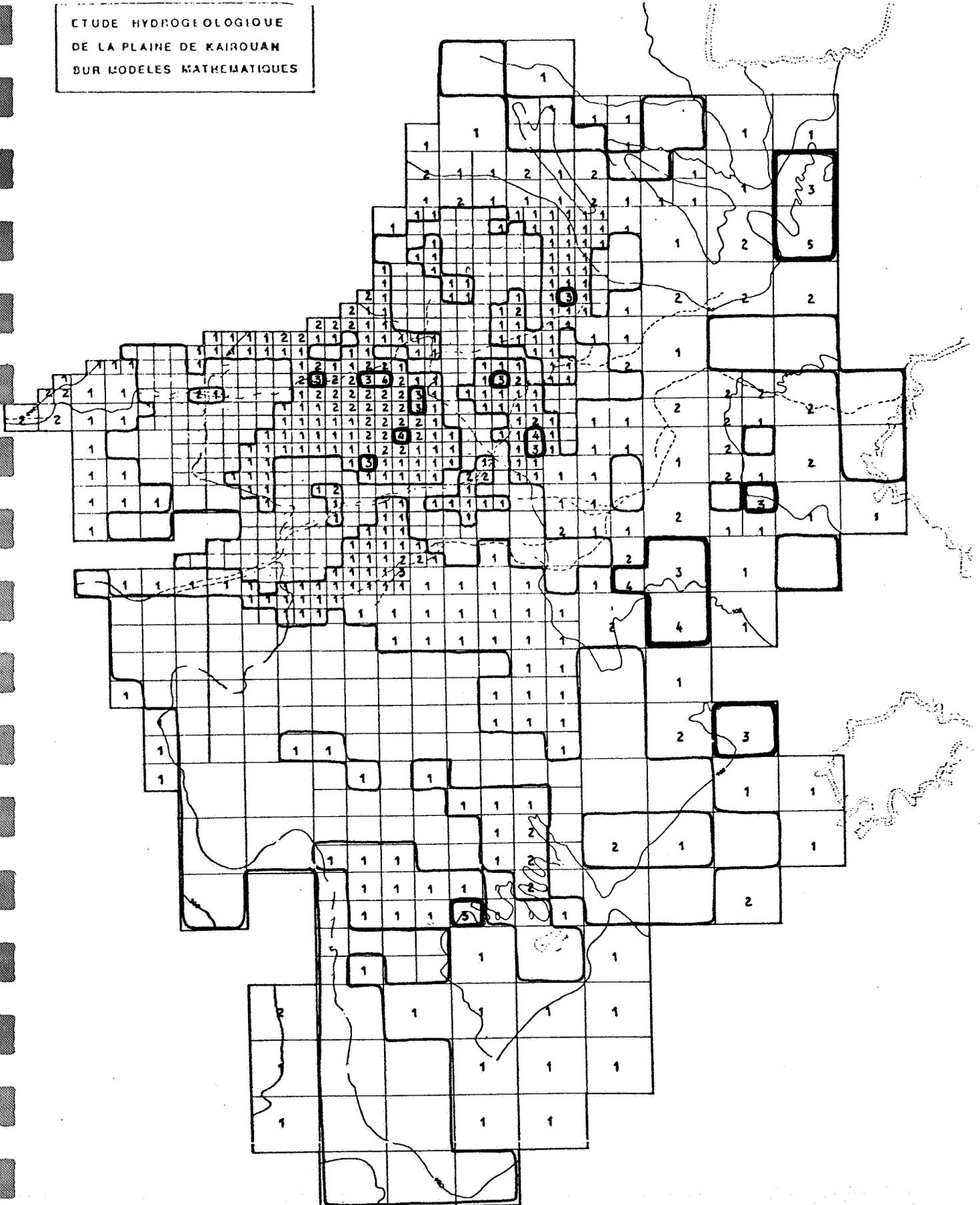


FIG. 61 - NAPPE DE KAIROUAN: ECARTS REDUITS

$$= \frac{\text{ECARTS DE CALAGE}}{\text{ECARTS-TYPES DE KRIGEAGE}}$$

REFERENCES BIBLIOGRAPHIQUES - CHAP. III, § 3 et 4

- [1] BESBES, M., DELHOMME, J.P., de MARSILY, G. (1975):  
Précision de la connaissance de la piézométrie d'un  
aquifère: conséquence sur le calage d'un modèle.  
Congrès International d'Hydrogéologie, Porto-Alegre, 1975  
Ecole des Mines, L.H.M., RD/75/23 bis
- [2] DELFINER, P. (1975):  
Linear estimation of non-stationary phenomena.  
Proc. of NATO Adv. Study Inst., Rome, 1975
- [3] DELFINER, P., DELHOMME, J.P. (1973):  
Optimum interpolation by kriging.  
Display and analysis of spatial data, pp. 96-114.  
(Proc. of NATO Adv. Study Inst., Northham, 1973).  
John Wiley & Sons.
- [4] GUIBAL, D. (1972):  
Simulation de schémas intrinsèques.  
Ecole des Mines, C.M.M., N-291
- [5] JOURNEL, A.G. (1974):  
Geostatistics for conditional simulation of orebodies.  
Economic Geology, vol. 69, n°5, pp. 673-687
- [6] MATHERON, G. (1972):  
The turning bands: a method for simulating random functions  
in  $R^n$ .  
Ecole des Mines, C.M.M., N-303.
- [7] MATHERON, G. (1973):  
The intrinsic random functions and their applications.  
Adv. in Appl. Prob., 5, pp. 439-468.
- [8] ORFEUIL, J.P. (1972):  
Simulation du Wiener-Lévy et de ses intégrales.  
Ecole des Mines, C.M.M., N-290.

CONCLUSION

L'acquisition de données ponctuelles ne constituait pas un but en soi, mais était destinée à permettre d'étendre cette connaissance fragmentaire à l'ensemble du domaine. A nouveau, la résolution du problème d'estimation qui était ainsi posé ne saurait être le terme du travail, mais ne constitue qu'une étape vers la compréhension du système hydrologique ou hydrogéologique étudié.

On a vu qu'en ce qui concerne les écoulements souterrains, on débouchait par le biais du calage des modèles numériques sur un problème d'identification, et non plus seulement d'estimation. A partir de la connaissance que l'on a de la forme des équations régissant les écoulements et des quelques fenêtres ponctuelles ouvertes sur le système physique, il s'agit de parvenir à simuler correctement le comportement de ce système pour offrir au gestionnaire des eaux un modèle sur lequel il pourra raisonner. Mais les précipitations et la piézométrie ne sont pas connues de façon exhaustive; un procédé quelconque d'identification du champ des transmissivités à partir de ces données (calage manuel d'un modèle ou procédure automatisée) ne peut conduire à un résultat sans incertitude. La quantification de la confiance qu'on peut lui accorder pourrait demeurer une question académique si toute incertitude sur la transmissivité ne se traduisait immédiatement par une incertitude sur les débits qui pourront être tirés de l'aquifère, et ne se répercutait ainsi sur les prévisions du gestionnaire.

Alors qu'après les premiers succès des modèles mathématiques, l'hydrogéologie semblait le règne du déterminisme, la nécessité se fait jour, actuellement, du recours à une approche probabiliste pour sortir de

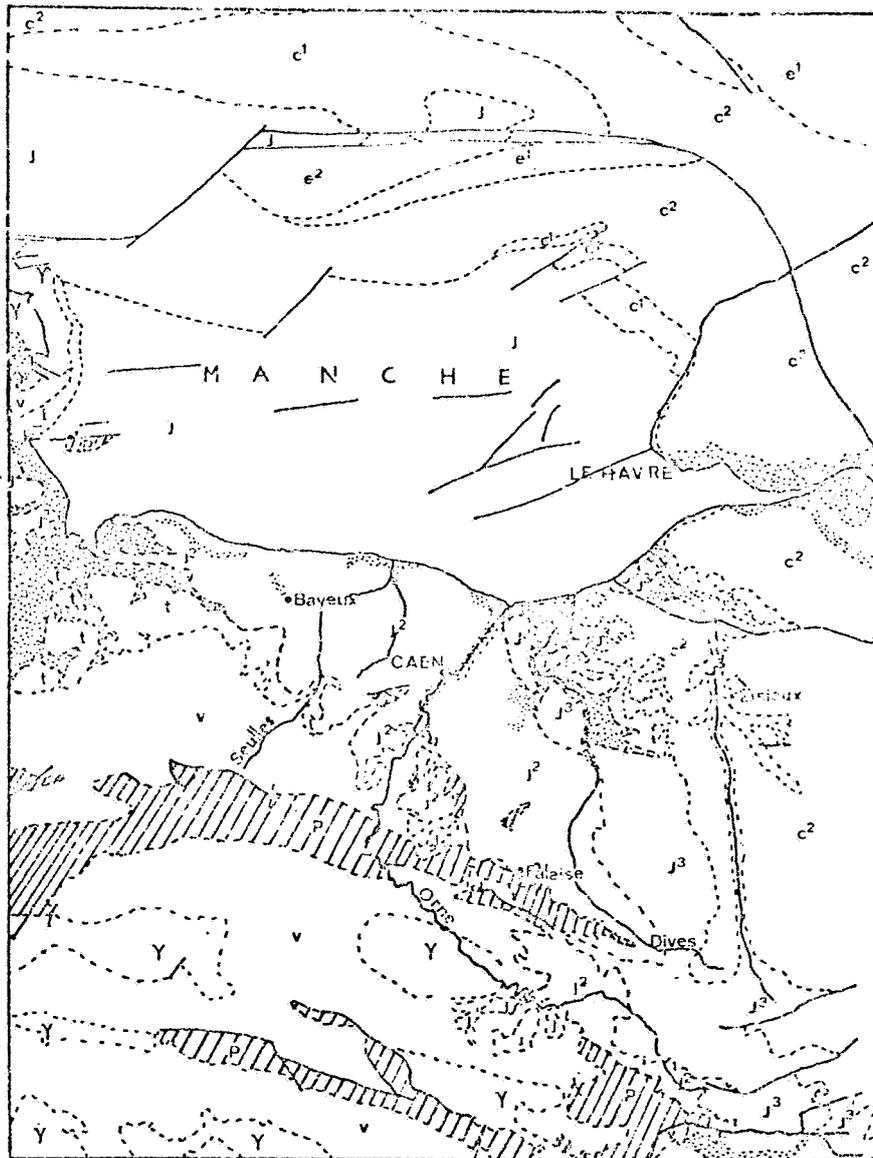


A N N E X E

FIG. A-1

CADRE DE L'ETUDE

échelle: 1/1000000



Légende

- |                |                               |   |                    |
|----------------|-------------------------------|---|--------------------|
| a              | alluvions                     | t | permo-trias, trias |
| e <sup>2</sup> | éocène moyen et supérieur     | P | primaire           |
| e <sup>1</sup> | éocène inférieur              | v | précambrien        |
| c <sup>2</sup> | crétacé supérieur             | Y | granite            |
| c <sup>1</sup> | crétacé inférieur             |   |                    |
| J <sup>3</sup> | jurassique supérieur          |   |                    |
| J <sup>2</sup> | jurassique inférieur et moyen |   |                    |
| J              | jurassique indifférencié      |   |                    |

d'après carte géologique de la France - BRG

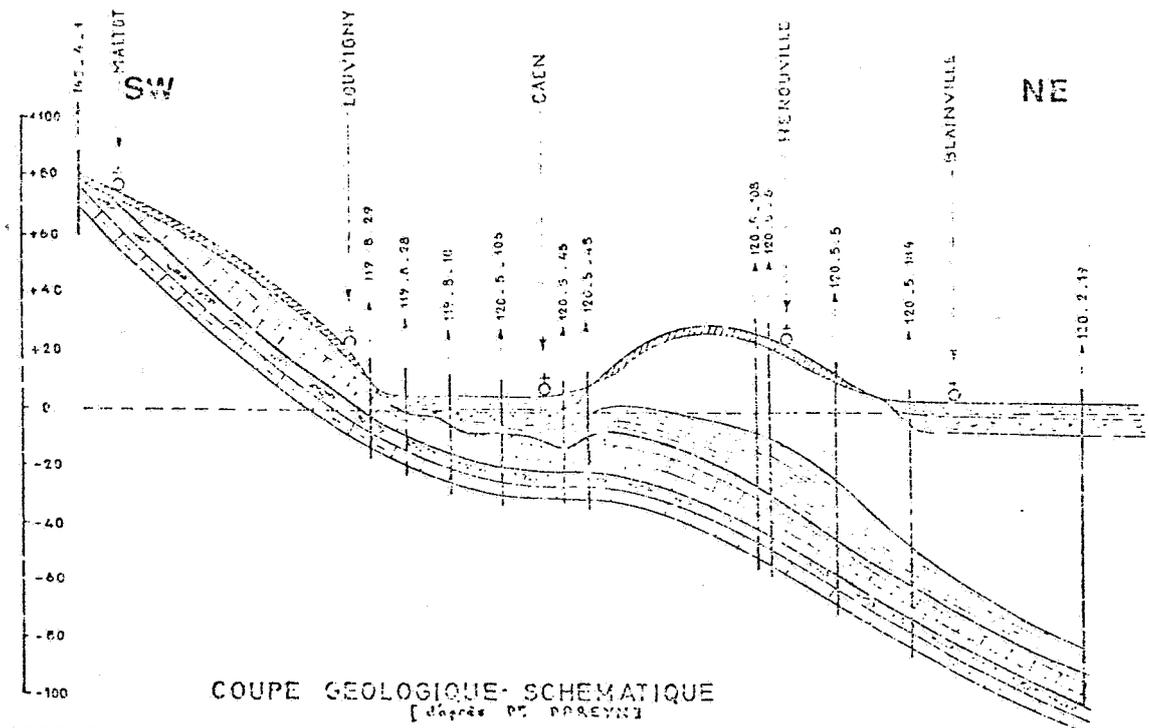
# UN EXEMPLE COMPLET D'APPLICATION HYDROGEOLOGIQUE

## L'ETUDE DES RESERVOIRS AQUIFERES DU DOGGER EN BASSE-NORMANDIE

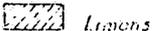
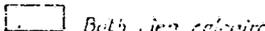
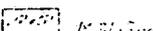
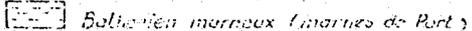
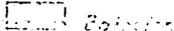
### CADRE DE L'ETUDE

Les problèmes d'alimentation en eau du bassin de l'Orne, et plus particulièrement ceux de la ville de Caen, ont rendu nécessaire la construction d'un modèle mathématique des aquifères de la région. Les terrains jurassiques des plaines de Caen, Falaise et Argentan constituent, en effet, d'importantes réserves d'eau souterraine qui apportent l'appoint indispensable aux ressources de surface.

Une coupe SO-NE passant par Caen permet de voir les relations pouvant exister entre les différentes unités structurales.



LEGENDE :

- |   |   |
|---|---|
|  Limons                              |  Alluvions |
|  Bathonien calcaire                  |  Mâtière  |
|  Bathonien marneux (marines de Port) |  Lias     |
|  Dogger                              |   |

ECHELLE : Longueurs

Fig 7

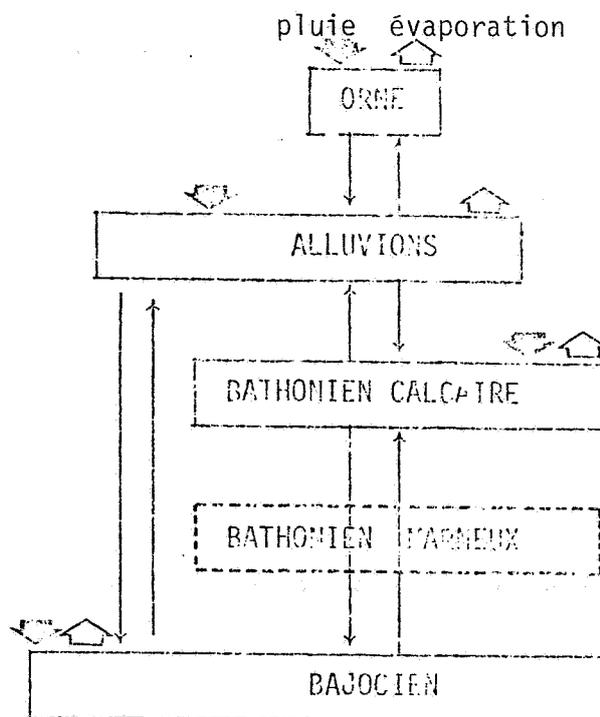
La succession la plus typique:

- alluvions des vallées,
- calcaire bathonien,
- argiles et marnes de Port-en Bessin,
- calcaire bajocien,
- substratum: argiles à poisson du Toarcien

est localement altérée:

- soit par disparition d'un ou plusieurs des termes de la série: ainsi l'aquifère bajocien existe seul dans la partie occidentale au-delà de la limite d'affleurement des argiles de Port-en-Bessin; de même, la partie sud de l'aquifère bathonien est entaillée par les bancs de grès armoricain affleurant dans le synclinal de la Brèche au Diable;
- soit au contraire par convergence des faciès stratigraphiques de termes contigus: toute la haute vallée de la Dives et du Laison, au sud-est de Caen, est constituée d'un empilement, sous un faciès essentiellement calcaire, de Bathonien sur le Bajocien; la structure y est, d'un point de vue hydraulique, de type monocouche.

Les transferts d'eau ont été schématisés de la façon suivante:



Le niveau de Bathonien marneux disparaissant dans la partie sud de la zone étudiée.

Les caractéristiques du modèle qui a été construit peuvent se résumer ainsi:

Couche \ Taille de maille	Nombre de mailles par couche			
	Micro 0.25 km <sup>2</sup>	Petite 1 km <sup>2</sup>	Moyenne 4 km <sup>2</sup>	Grande 16 km <sup>2</sup>
0 Rivières	65	65	18	0
1 Alluvions	92	40	15	2
2 Bathonien (+ Bajocien.)	256	252	195	10
3 Bajocien seul	27	282	165	14

La surface cumulée des différentes couches représentant le milieu poreux atteint 4.880 km<sup>2</sup>. Le modèle comporte au total 1351 mailles.

Afin de synthétiser l'ensemble de l'information hydrogéologique rassemblée en vue de sa construction, on a décidé d'appliquer la technique du krigeage à l'étude des variations spatiales des grandeurs suivantes:

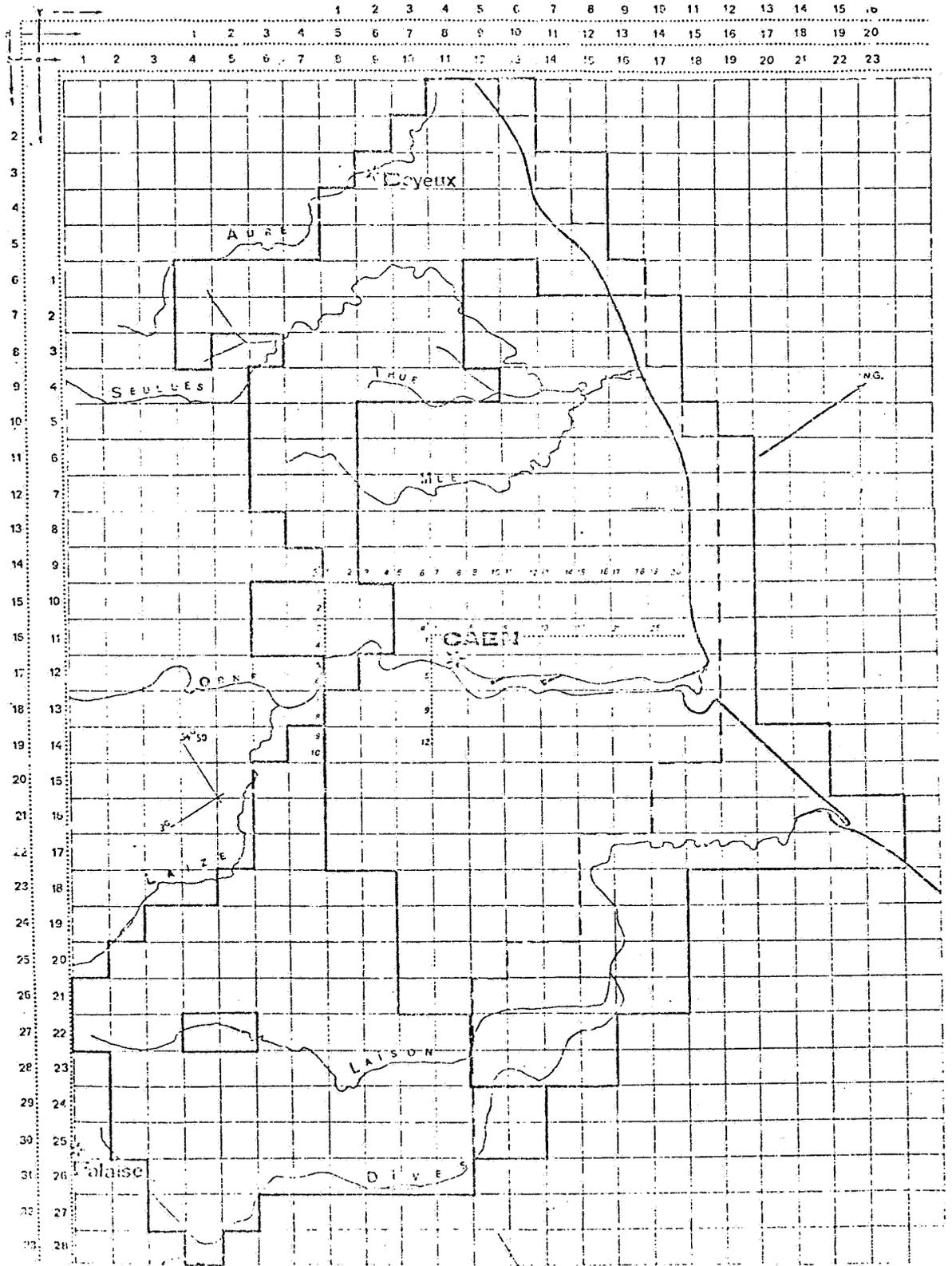
- transmissivités des deux aquifères bathonien et bajocien,
- piézométrie de l'ensemble du Dogger,
- cotes des murs des deux aquifères, et épaisseur des marnes de Port-en-Bessin.

La figure A-3 regroupe les différents maillages utilisés.

### ETUDE DES TRANSMISSIVITES

Pour chacun des deux aquifères on a estimé par krigeage les valeurs moyennes de la transmissivité sur des mailles de 2 x 2 km (correspondant à la taille des mailles moyennes du modèle).

FIG. A-3 - DIFFERENTES GRILLES UTILISEES  
POUR LE KRIGEAGE



a) Données disponibles:

Une analyse rapide de l'ensemble des données de transmissivité et de débit spécifique existant sur les aquifères du Dogger a montré:

- le caractère sensiblement lognormal des variables T et Q/S, suggérant de recourir à une transformation logarithmique sur ces variables;
- l'existence d'un coefficient de corrélation de 0,90 entre leurs logarithmes (calculé à partir des 59 couples disponibles).

Il a donc été décidé, pour chacun des deux aquifères étudiés - Bathonien et Bajocien - d'utiliser les données de débit spécifique pour étendre le jeu de transmissivités disponible à partir des seuls essais de pompage, en procédant à une régression linéaire de  $\log T$  en  $\log Q/S$ .

• Les figures A-4 et A-6 montrent la position des différents puits et forages pour chacune des deux nappes, les valeurs de transmissivités n'étant indiquées qu'en cas d'existence d'un essai de pompage.

• Sur les figures A-5 et A-7, où sont reportés graphiquement les couples  $(\log_{10} T, \log_{10} \frac{Q}{S})$ , apparaît nettement l'existence d'une liaison linéaire entre ces deux variables, en ce qui concerne le Bathonien que le Bajocien.

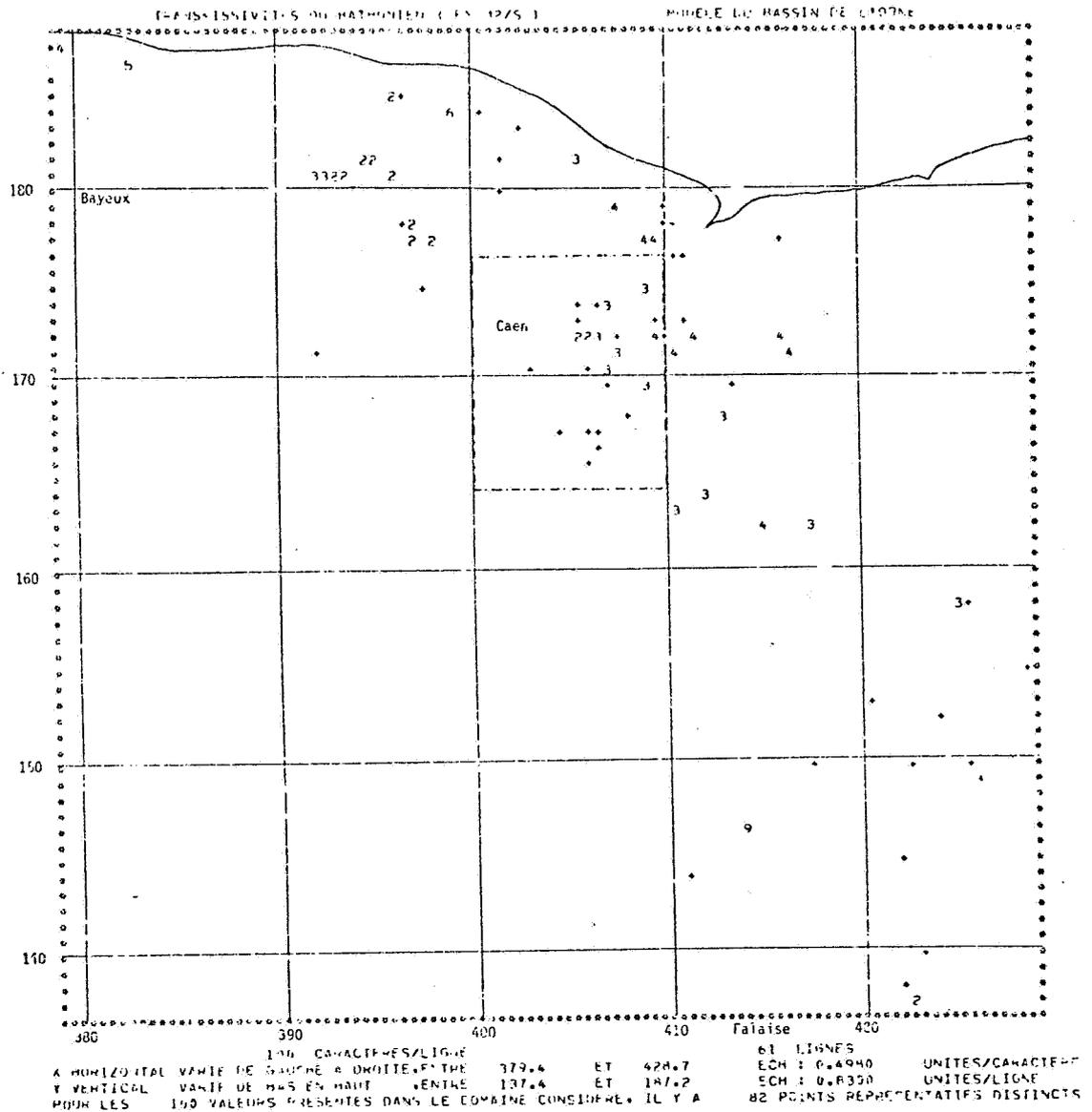
Les coefficients a et b des équations de régression:  $\log_{10} T = a \log_{10} Q/S + b$  ont été calculés par moindres carrés, et les droites de régression tracées sur les figures A-5 et A-7. Pour l'aquifère du Bathonien, à partir des 45 couples disponibles, on a obtenu:

$$\log_{10} T = 0,898 \log_{10} Q/S - 0,209$$

En ce qui concerne le Bajocien, malgré le faible nombre de couples (10), le calcul a néanmoins été fait et a donné:

$$\log_{10} T = 1,30 \log_{10} Q/S + 0,563$$

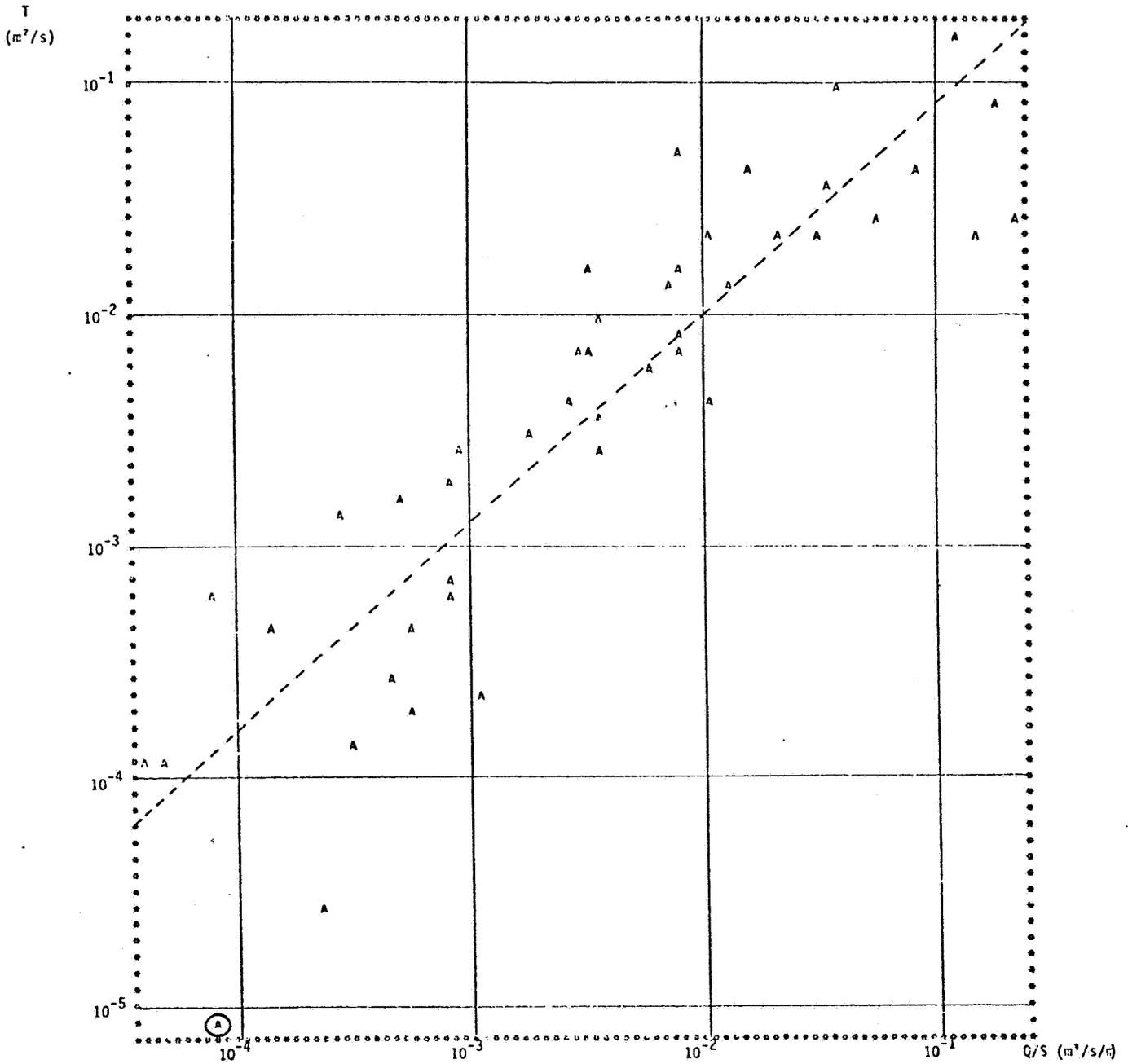
FIG. A-4 - BATHONIEN: POSITION DES FORAGES ET VALEURS DE TRANSMISSIVITES OBTENUES PAR ESSAIS DE POMPAGE



- Le signe + est réservé aux forages où aucun essai de pompage n'a permis de déterminer la transmissivité.
- Les valeurs numériques sont affectées aux forages où la transmissivité est connue par essais de pompage.

- 1 - représente les valeurs de T supérieures à  $10^{-1}$  m<sup>2</sup>/s
- 2 - représente les valeurs de T comprises entre  $10^{-1}$  et  $10^{-2}$  m<sup>2</sup>/s
- 3 - représente les valeurs de T comprises entre  $10^{-2}$  et  $10^{-3}$  m<sup>2</sup>/s
- 4 - représente les valeurs de T comprises entre  $10^{-3}$  et  $10^{-4}$  m<sup>2</sup>/s
- 5 - représente les valeurs de T comprises entre  $10^{-4}$  et  $10^{-5}$  m<sup>2</sup>/s
- 6 - représente les valeurs de T comprises entre  $10^{-5}$  et  $10^{-6}$  m<sup>2</sup>/s
- 7 - représente les valeurs de T comprises entre  $10^{-6}$  et  $10^{-7}$  m<sup>2</sup>/s
- 8 - représente les valeurs de T comprises entre  $10^{-7}$  et  $10^{-8}$  m<sup>2</sup>/s
- 9 - représente les valeurs de T inférieures à  $10^{-9}$  m<sup>2</sup>/s.

FIG. A-5 - AQUIFERE DU BATHONIEN: REGRESSION LINEAIRE DE LOG<sub>10</sub>(T)  
EN LOG<sub>10</sub> (Q/S)



49 CARACTERES/LIGNE  
 A HORIZONTAL VARIE DE GAUCHE A DROITE, ENTRE -4.409 ET -0.6540 ECH : 0.4261E-01 UNITES/CARACTERE  
 Y VERTICAL VARIE DE BAS EN HAUT, ENTRE -5.079 ET -0.7959 ECH : 0.7139E-01 UNITES/LIGNE  
 POUR LES 45 VALEURS PRESENTES DANS LE DOMAINE CONSIDERE, IL Y A 45 POINTS REPRESENTATIFS DISTINCTS

ON OBTIENT  $\Delta \text{LOG}_{10}(T) = 0.698E 00 \Delta \text{LOG}_{10}(Q/S) - 0.209E 00$  A PARTIR DES 45 POINTS DISPONIBLES VAR. RES. = 0.205E 00

FIG. A-6 - BAJOCIEN: POSITION DES FORAGES ET VALEURS DES TRANSMISSIVITES  
OBTENUES PAR ESSAIS DE POMPAGE

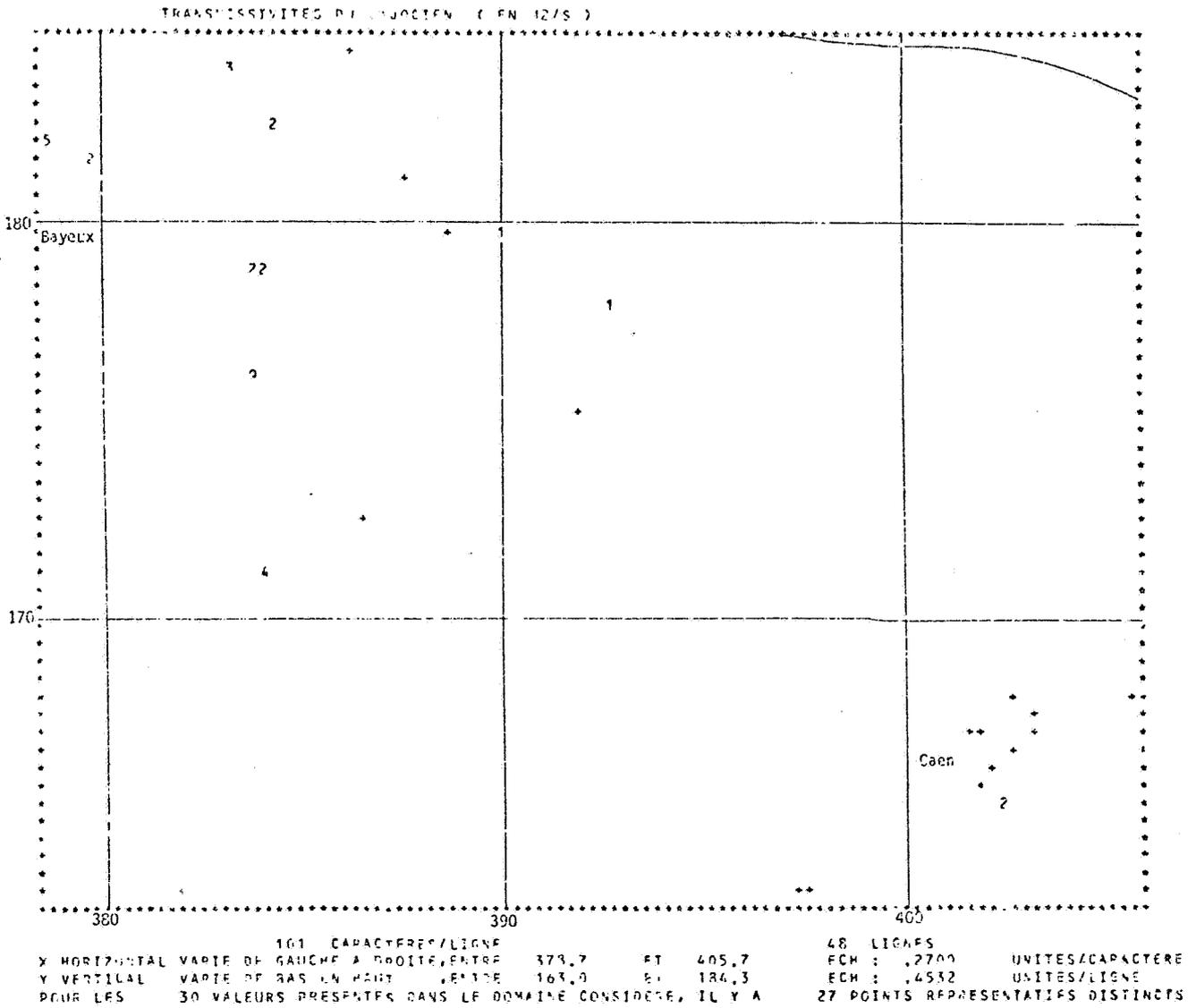
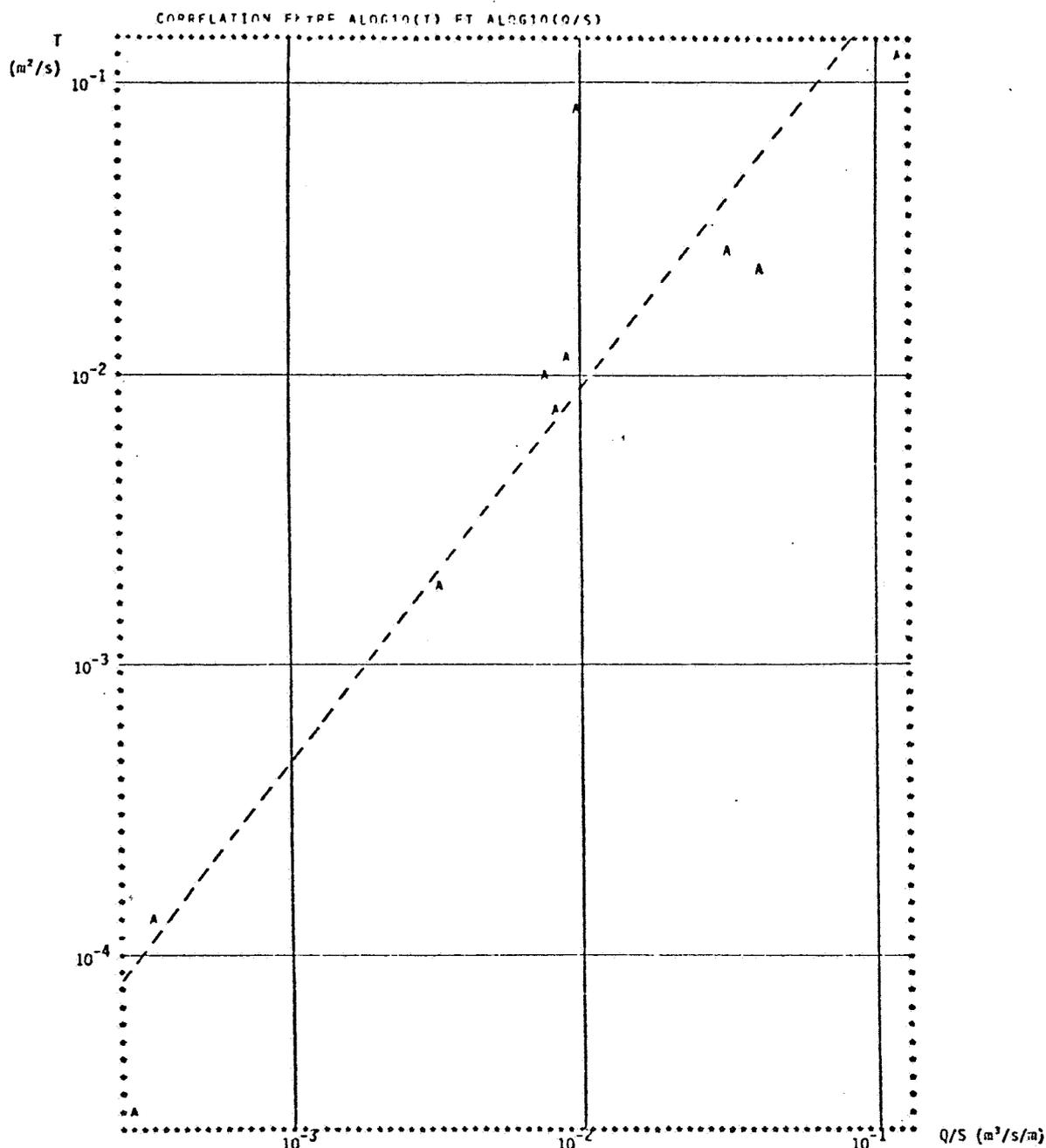


FIG. A-7 - AQUIFERE DU BAJOCIEN: REGRESSION LINEAIRE DU  $\text{LOG}_{10}(T)$   
EN  $\text{LOG}_{10}(Q/S)$



<p>73 CARACTERES/LIGNE X HORIZONTAL VARIE DE GAUCHE A DROITE, ENTRE -3.540 ET -.9234 Y VERTICAL VARIE DE BAS EN HAUT, ENTRE -4.556 ET -.9079 POUR LES 10 VALEURS PRESENTES DANS LE DOMAINE CONSIDERE, IL Y A</p>	<p>61 LIGNES FCW 1 .5634E-01 UNITES/CARACTERE- FCW 1 .6091E-01 UNITES/LIGNE 10 POINTS REPRESENTATIFS DISTINCTS</p>
--	--

ON OBTIENT  $\text{ALOG}_{10}(T) = .130E+01 \cdot \text{ALOG}_{10}(Q/S) + .563E+00$  A PARTIR DES 10 POINTS DISPONIBLES VAR.RES. = .184E+00

• La liaison linéaire entre  $\log_{10} T$  et  $\log_{10} Q/s$ , bien que nette, n'en est pas pour autant une liaison exacte (fonctionnelle), la qualité de cette liaison étant d'ailleurs mesurée par la variance résiduelle:

$$\sigma_{\epsilon}^2 = \frac{1}{n-2} \sum_{i=1}^n [\log_{10} T_i - (a \log_{10} (Q/s)_i + b)]^2$$

la sommation courant sur les n couples disponibles.

En estimant  $\log_{10} T_j$  à partir de  $\log_{10} (Q/s)_j$ , pour les m-n puits où seule est disponible une valeur de débit spécifique, on commet donc inévitablement une erreur de prédiction dont la variance peut être calculée comme:

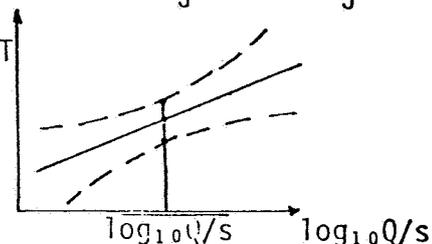
$$\sigma_j^2 = \sigma_{\epsilon}^2 \cdot \left[ 1 + \frac{1}{n} + \frac{[\log_{10} (Q/s)_j - \overline{\log_{10} Q/s}]^2}{\sum_{i=1}^n [\log_{10} (Q/s)_i - \overline{\log_{10} Q/s}]^2} \right] \text{ pour } j = n+1, \dots, m$$

où  $\overline{\log_{10} Q/s}$  désigne la valeur moyenne  $\frac{1}{n} \sum_{i=1}^n \log_{10} (Q/s)_i$ .

La précision de la reconstitution peut être approximativement fournie par l'intervalle à 95%:

$$(a \log_{10} (Q/s)_j + b) - 2 \sigma_j \leq (\log_{10} T_j) \text{ vrai} \leq (a \log_{10} (Q/s)_j + b) + 2 \sigma_j$$

La formule donnant  $\sigma_j^2$  montre que l'amplitude de cet intervalle croît quand  $\log_{10} (Q/s)_j$  s'éloigne de  $\overline{\log_{10} Q/s}$ .



En repassant des logarithmes aux valeurs de transmissivités, cet intervalle de confiance se met sous la forme d'un facteur d'incertitude égal à  $10^{2\sigma_j}$  affectant la valeur  $T_j$  estimée.

Les valeurs obtenues par essais de pompage sont, pour les besoins de la cause et faute de pouvoir apprécier l'incertitude les affectant, considérées connues exactement, c'est à dire à un facteur 1 près.

On trouvera dans les pages suivantes les différentes valeurs de transmissivités accompagnées de leur facteur d'incertitude pour les deux aquifères étudiés.

A ces données proprement dites ont été ajoutées deux valeurs de transmissivité très faibles figurant la fenêtre de la Brèche-au-Diable sur le Laison pour le Bathonien, et celle de la Haute Seulle pour le Bajocien.

FORAGES CAPTANT LE BATHONIEN

CODE	X	Y	TRASSA.	FACT.
96.5.001	379.400	187.100	0.20E-03	1.0
96.5.002	379.500	187.200	0.56E-03	1.0
96.5.006	372.900	186.400	0.20E-04	1.0
96.7.003	396.300	184.500	0.18E-00	8.9
96.7.005	396.700	184.800	0.17E-01	8.3
96.7.058	396.400	184.400	0.76E-01	1.0
96.8.002	400.900	183.700	0.31E-03	8.5
96.8.019	399.400	184.100	0.83E-05	1.0
19.3.168	394.700	181.100	0.16E-00	1.0
19.3.169	396.200	180.700	0.64E-02	1.0
19.3.170	396.400	180.600	0.24E-01	1.0
19.3.171	394.900	181.000	0.45E-01	1.0
19.3.172	395.200	181.300	0.97E-01	1.0
19.3.173	393.600	180.900	0.39E-01	1.0
19.3.179	392.300	180.500	0.69E-02	1.0
19.3.180	392.600	180.900	0.31E-02	1.0
19.3.181	393.100	180.400	0.15E-01	1.0
19.3.182	393.400	180.700	0.15E-01	1.0
19.4.007	401.800	180.100	0.10E-01	8.3
19.4.079	401.600	181.100	0.22E-03	8.6
19.4.097	402.800	183.100	0.38E-03	8.5
19.4.139	396.600	177.800	0.71E-02	8.2
19.4.140	397.500	177.500	0.22E-01	1.0
19.4.142	397.500	178.300	0.40E-01	1.0
19.4.145	398.300	177.100	0.27E-01	1.0
19.4.149	397.900	174.600	0.26E-03	8.5
19.7.045	392.300	171.400	0.35E-02	8.2
19.8.066	403.200	170.700	0.47E-02	8.2
20.1.002	407.400	173.800	0.73E-02	1.0
20.1.011	405.700	181.600	0.42E-02	1.0
20.1.012	407.900	178.700	0.19E-03	1.0
20.1.015	406.000	173.800	0.24E-02	8.2
20.1.080	409.300	177.000	0.42E-03	1.0
20.1.083	409.200	174.800	0.14E-02	1.0
20.1.156	406.700	173.800	0.86E-03	8.3
20.1.158	410.000	177.600	0.14E-03	1.0
20.2.002	410.200	178.600	0.50E-03	8.4
20.2.003	410.700	177.800	0.20E-01	8.4
20.2.006	416.100	177.000	0.41E-02	8.2
20.2.020	410.300	177.900	0.10E-01	8.3
20.2.071	410.600	176.600	0.27E-02	8.2
20.2.076	411.200	176.500	0.91E-03	8.3
20.5.002	406.800	167.200	0.60E-01	8.6
20.5.003	408.300	167.900	0.87E-02	8.3
20.5.005	406.800	171.900	0.97E-02	1.0
20.5.006	406.100	170.900	0.14E-01	1.0
20.5.007	406.400	165.800	0.46E-02	8.2
20.5.107	406.500	171.000	0.99E-02	8.3
20.5.108	406.200	170.800	0.87E-02	8.3
20.5.109	406.200	172.200	0.13E-01	1.0
20.5.110	406.000	172.400	0.50E-01	1.0
20.5.113	406.500	166.900	0.23E-00	9.0
20.5.114	406.800	166.400	0.15E-01	8.3
20.5.116	406.600	167.600	0.25E-01	1.0
20.5.117	406.700	167.400	0.22E-01	1.0
20.5.133	407.000	166.700	0.44E-01	8.5

AGES CARBANT LE BATHONIEN

(suite)

20.5.135	405.500	167.200	0.14E-00	8.9
20.5.141	408.200	167.700	0.74E-02	8.2
20.5.143	407.200	169.500	0.56E-00	9.4
20.5.151	405.500	166.800	0.41E-03	8.4
20.5.153	405.700	166.200	0.18E-02	8.2
20.5.184	408.000	172.600	0.21E-02	8.2
20.5.226	414.900	167.600	0.17E-01	8.3
20.5.254	409.700	172.700	0.41E-03	8.4
20.5.255	409.600	172.400	0.38E-02	8.2
20.5.321	405.900	172.900	0.46E-01	8.5
20.5.322	405.000	173.100	0.57E-02	8.2
20.5.384	407.700	171.800	0.78E-02	1.0
20.5.387	409.500	169.700	0.19E-02	1.0
20.5.427	409.200	171.900	0.11E-03	1.0
20.5.441	407.400	170.700	0.69E-02	1.0
20.6.100	413.300	168.300	0.15E-02	1.0
20.6.145	413.600	169.800	0.48E-03	8.4
20.6.146	412.500	163.800	0.44E-02	1.0
20.6.154	416.300	172.100	0.22E-03	1.0
20.6.173	410.300	172.400	0.71E-05	10.1
20.6.174	410.700	163.200	0.28E-02	1.0
20.6.175	411.100	172.900	0.26E-02	8.2
20.6.177	411.800	172.500	0.42E-03	1.0
20.6.179	410.900	171.600	0.11E-03	1.0
20.7.061	416.800	171.500	0.75E-03	1.0
46.2.128	415.200	162.100	0.56E-03	1.0
46.3.097	417.700	161.900	0.25E-02	1.0
46.3.114	420.600	153.400	0.26E-02	8.2
46.4.001	425.700	158.100	0.14E-02	8.3
46.4.002	425.300	158.100	0.38E-02	1.0
46.4.003	428.700	154.800	0.13E-03	8.7
46.6.057	411.100	144.100	0.29E-02	8.2
46.7.047	417.800	149.700	0.42E-01	8.5
46.8.001	425.500	149.900	0.70E-02	8.2
46.8.002	425.700	149.500	0.33E-02	8.2
46.8.003	422.400	144.700	0.61E-02	8.2
46.8.004	424.300	152.600	0.13E-01	8.3
46.8.011	422.600	149.700	0.46E-03	8.4
46.8.032	425.400	149.000	0.20E-01	8.4
76.4.009	422.500	137.400	0.23E-01	1.0
76.4.010	423.000	140.000	0.11E-01	8.3
76.4.011	423.000	139.700	0.56E-02	8.2
76.4.012	422.400	138.100	0.33E-02	8.2

FORAGES CANTANT LE BAJOCIEN

CODE	Y	Y	TRANSM.	FACT.
96.6.001	356.300	184.300	.47E-02	7.9
19.1.046	323.200	183.800	.81E-02	1.0
19.1.161	327.600	181.200	.89E-05	14.5
19.1.234	383.700	178.800	.22E-01	1.0
19.1.237	378.700	181.900	.28E-04	1.0
19.1.239	383.300	184.000	.19E-02	1.0
19.1.250	384.400	182.700	.10E-01	1.0
19.1.251	379.900	181.800	.26E-01	1.0
19.2.100	388.700	179.800	.11E-01	8.0
19.2.210	384.100	178.900	.11E-01	1.0
19.2.213	390.100	179.600	.20E+00	1.0
19.3.175	392.800	178.100	.12E+00	1.0
19.3.176	391.900	175.200	.43E-02	8.0
19.6.077	386.600	172.600	.16E-02	8.2
19.6.105	384.200	171.200	.14E-03	1.0
19.8.001	402.500	165.100	.76E-01	1.0
19.8.037	402.600	166.800	.17E-01	8.1
19.8.042	402.300	166.300	.48E-03	8.8
19.8.043	401.900	165.500	.58E-02	7.9
19.8.047	401.600	167.000	.37E-02	8.0
19.8.050	402.000	167.000	.16E-01	8.1
19.8.120	401.900	165.500	.65E-02	7.9
19.8.196	402.800	167.900	.43E-01	8.5
20.5.106	403.200	167.200	.88E-03	8.4
20.5.149	405.700	168.000	.72E-04	10.7
20.5.274	403.300	167.700	.94E-02	8.0
45.4.005	397.300	163.200	.10E-02	8.3
45.4.069	397.300	163.000	.54E-03	8.7
45.4.070	397.600	163.200	.13E-02	8.2

b) Méthode (Rappel):

Les données provenant de la régression doivent être, comme on l'a vu, utilisées avec précaution, mais elles ne peuvent être rejetées sous peine de se priver d'une information non négligeable.

On a donc eu recours à une formulation adaptée du krigeage, autorisant la prise en compte de données de différentes qualités, pour estimer les valeurs de transmissivités à afficher sur les mailles du modèle en construction.

Le krigeage prend en compte:

- les positions relatives des points expérimentaux entre eux et par rapport à la maille à estimer,
- la structure du phénomène étudié: présence éventuelle de tendances à grande échelle (dérive), et surtout variabilité, caractérisée par le variogramme (ou la covariance généralisée).

Il fournit, à l'aide d'une pondération des valeurs expérimentales du voisinage, l'estimation la plus précise possible ("optimale"), sans sur- ou (sous-) évaluation systématique ("sans biais"), ainsi que la précision de cette estimation, sous la forme d'une variance d'estimation.

Dans le cas présent, la pondération optimale des données que réalise le krigeage prend en compte, en outre, les qualités respectives des dites données. Il faut également noter que les calculs, s'effectuant sur les logarithmes des transmissivités, donnent un estimateur de krigeage  $Z_K$ , un écart-type d'estimation  $\sigma_K$  et donc un intervalle de confiance  $Z_K \pm 2 \sigma_K$  relatifs à cette variable transformée.

Le retour aux valeurs de transmissivité elles-mêmes conduit à un estimateur médian de  $T$  égal à  $10^{Z_K}$  et l'assortit d'une spécification de l'intervalle de confiance sous forme d'un coefficient multiplicateur égal à  $10^{2\sigma_K}$ .

$$10^{Z_K} / 10^{2\sigma_K} < T < 10^{Z_K} \cdot 10^{2\sigma_K}$$

c) Résultats:

α) Reconnaissance structurale: La première phase du travail, préalable au krigeage proprement dit, est constituée par l'étude de la structure du phénomène.

Le modèle de variogramme retenu pour le logarithme de la transmissivité a été, pour le Bathonien comme pour le Bajocien, un modèle linéaire avec discontinuité à l'origine.

Dans les deux cas, il a été conclu également à une absence de dérive systématique.

- En ce qui concerne le Bathonien, la reconnaissance structurale a été faite à partir des seules 45 données provenant d'essais de pompage, de façon à éviter de sur-estimer l'effet de pépite du fait de la présence de données susceptibles d'être entachées d'erreur.

Les transmissivités étant exprimées en  $m^2/s$  et les coordonnées des forages en km., les coefficients de  $\gamma(h)$  ont été déterminés comme suit:

effet de pépite:  $C = 0.09$  pente:  $A1 = 0.125$

- En ce qui concerne le Bajocien, le nombre des données d'essais de pompage n'étant que de 10, la reconnaissance a dû porter sur l'ensemble des 29 données.

Les coefficients trouvés ont alors été:

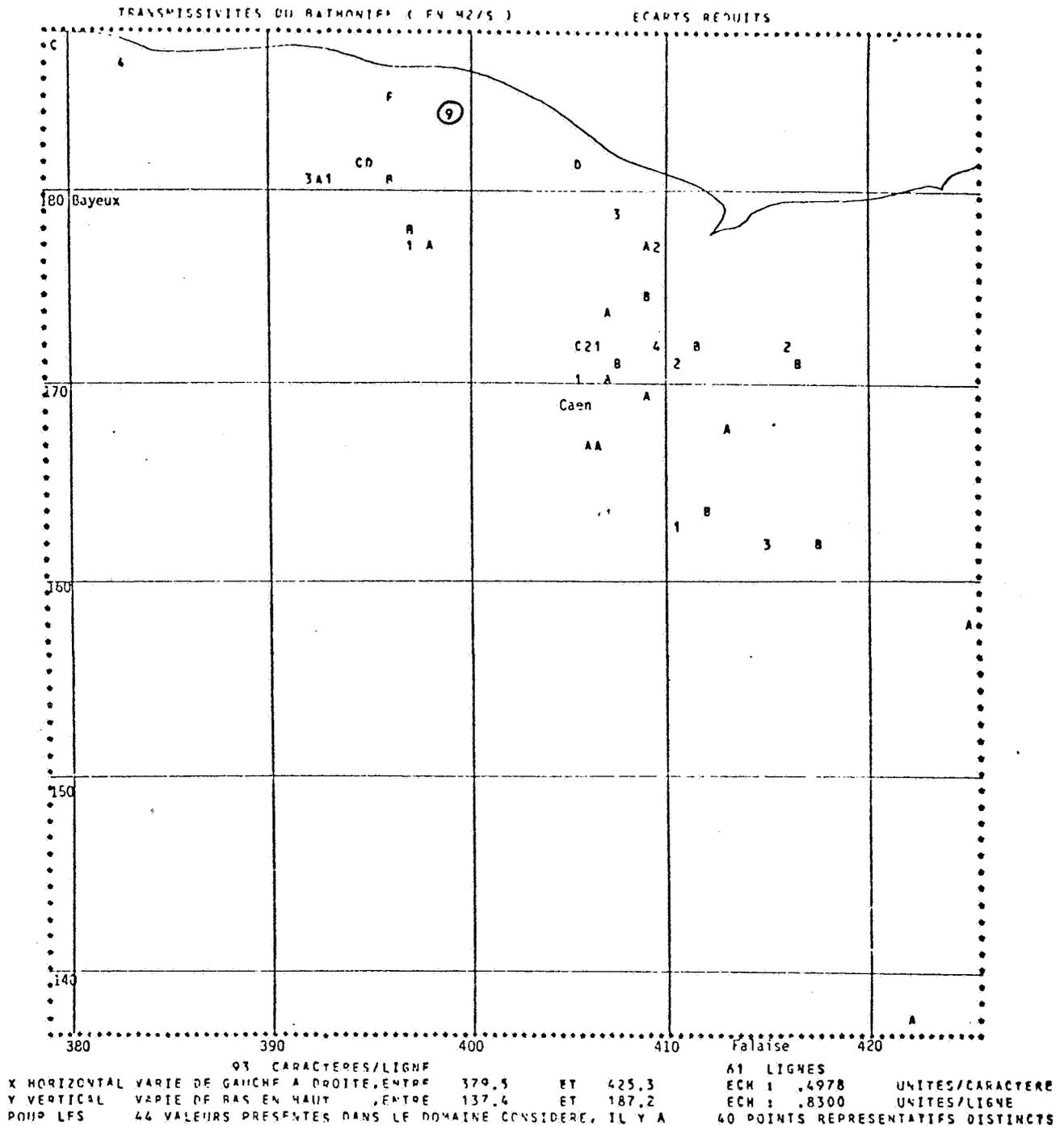
effet de pépite:  $C = 0.56$  pente:  $A1 = 0.037$

La valeur plus forte de la discontinuité  $C$  est à relier à la présence de données incertaines dans le jeu. Quant à la pente  $A1$  plus faible, elle semblerait indiquer une moindre variabilité de la transmissivité dans l'aquifère inférieur.

- Un essai de krigeage a été, en outre, réalisé à titre d'exemple sur le Bathonien, afin de vérifier la validité des paramètres calculés. La méthode consiste à supprimer fictivement les points expérimentaux l'un après l'autre, à les kriger à l'aide de leurs voisins, et à comparer les erreurs d'estimation expérimentales ainsi obtenues aux écarts-types d'estimation théoriques.

Ceci a été fait sur l'ensemble des forages où existait un essai de pompage, et les écarts réduits  $\frac{Z_{réel} - Z_K}{\sigma_K}$  ont été reportés sur la figure A-8:

FIG. A-8 - LOGARITHMES DES TRANSMISSIVITES DU BATHONIEN  
ECARTS REDUITS



1	REPRESENTE LES VALEURS COMPRISES ENTRE	.0000	ET	.5000
2	REPRESENTE LES VALEURS COMPRISES ENTRE	.5000	ET	1.0000
3	REPRESENTE LES VALEURS COMPRISES ENTRE	1.0000	ET	1.5000
4	REPRESENTE LES VALEURS COMPRISES ENTRE	1.5000	ET	2.0000
5	REPRESENTE LES VALEURS COMPRISES ENTRE	2.0000	ET	2.5000
6	REPRESENTE LES VALEURS COMPRISES ENTRE	2.5000	ET	3.0000
7	REPRESENTE LES VALEURS COMPRISES ENTRE	3.0000	ET	3.5000
8	REPRESENTE LES VALEURS COMPRISES ENTRE	3.5000	ET	4.0000
9	REPRESENTE LES VALEURS COMPRISES ENTRE	4.0000	ET	4.5000
A	REPRESENTE LES VALEURS SUPERIEURES A	4.5000		
A	REPRESENTE LES VALEURS COMPRISES ENTRE	-4.5000	ET	.0000
R	REPRESENTE LES VALEURS COMPRISES ENTRE	-1.0000	ET	-.5000
C	REPRESENTE LES VALEURS COMPRISES ENTRE	-1.5000	ET	-1.0000
D	REPRESENTE LES VALEURS COMPRISES ENTRE	-2.0000	ET	-1.5000
E	REPRESENTE LES VALEURS COMPRISES ENTRE	-2.5000	ET	-2.0000
F	REPRESENTE LES VALEURS COMPRISES ENTRE	-3.0000	ET	-2.5000
G	REPRESENTE LES VALEURS COMPRISES ENTRE	-3.5000	ET	-3.0000
H	REPRESENTE LES VALEURS COMPRISES ENTRE	-4.0000	ET	-3.5000
I	REPRESENTE LES VALEURS COMPRISES ENTRE	-4.5000	ET	-4.0000

On a pu vérifier (cf. Fig. 27, p. 78 dans le corps du texte\*) qu'en moyenne il n'y avait bien ni sous- ni sur-estimation systématique:

$$\text{erreur moyenne} = \frac{1}{n} \sum (Z - Z_K) = -0.026 \approx 0$$

et que les écarts-types d'estimation annoncés étaient bien cohérents avec les erreurs constatées:

$$\text{erreur quadratique réduite moyenne} = \frac{1}{n} \sum \left( \frac{Z - Z_K}{\sigma_K} \right)^2 = 0.96 \approx 1$$

Les écarts réduits restent partout inférieurs à 2, sauf dans la région de Bernières s/Mer où la transmissivité de  $0,83 \cdot 10^{-5} \text{ m}^2/\text{s}$  obtenue par essai de pompage au forage 96.8.19 n'apparaît pas cohérente avec l'information obtenue au cours des essais du voisinage (écart réduit > 4).

Une valeur de débit spécifique de l'ordre de  $10^{-4} \text{ m}^3/\text{s}/\text{m}$ , semblable à celle des forages avoisinants, et la position à l'écart en bas et à gauche du point correspondant sur le graphique  $T \leftrightarrow Q/s$  (Fig. A-5) semblent plaider en faveur d'un réexamen de cette valeur.

Comme on peut le voir, le krigeage peut être un bon outil de vérification de la cohérence des données.

### β) Krigeage:

#### • Aquifère du Bathonien:

- Une première estimation de la transmissivité, à partir des seules données provenant d'essais de pompage (plus 1 point figurant la fenêtre) a donné un modèle assez pauvre, avec seulement trois grandes zones: transmissivités fortes dans le Nord-Ouest, assez faibles dans le Nord-Est et très faibles dans tout le Sud-Est (du fait de la grande importance que prend le point figurant la fenêtre, rajouté dans une zone très pauvre en essais de pompage).

---

\*Un certain nombre de planches figurant déjà dans le corps du texte n'ont pas été reprises dans la présente annexe.

Les coefficients multiplicateurs, indiquant les bornes de l'intervalle de confiance à 95%, décroissent jusqu'à 2. dans les zones bien reconnues, mais dépassent 100. sur les limites où la transmissivité n'est pratiquement pas connue.

- L'estimation a été faite à nouveau en prenant, cette fois, l'ensemble des données (y compris le point figurant la fenêtre), et en tenant compte des incertitudes entachant les valeurs. Les traits majeurs de la répartition spatiale de la transmissivité semblent mieux se dégager: on obtient une meilleure focalisation des zones à forte transmissivité ( $> 10^{-2} \text{ m}^2/\text{s}$ ) autour de Caen et de la Mue. A noter aussi l'emprise moindre de la zone à très faible transmissivité ( $< 5.10^{-4} \text{ m}^2/\text{s}$ ) entourant la fenêtre de la Brèche au Diable, et la mise en évidence de zones à faible transmissivité, orientées sensiblement NW-SE, et correspondant sans doute à l'influence de l'accident de Ranville et à son prolongement (cf. Fig. A-9).

Du fait de la prise en compte d'une information supplémentaire, même si elle est entachée d'incertitude, les coefficients de l'intervalle de confiance (cf. Fig. A-10) sont inférieurs à ceux trouvés précédemment, la baisse observée étant évidemment la plus forte dans les zones peu reconnues par essais de pompage.

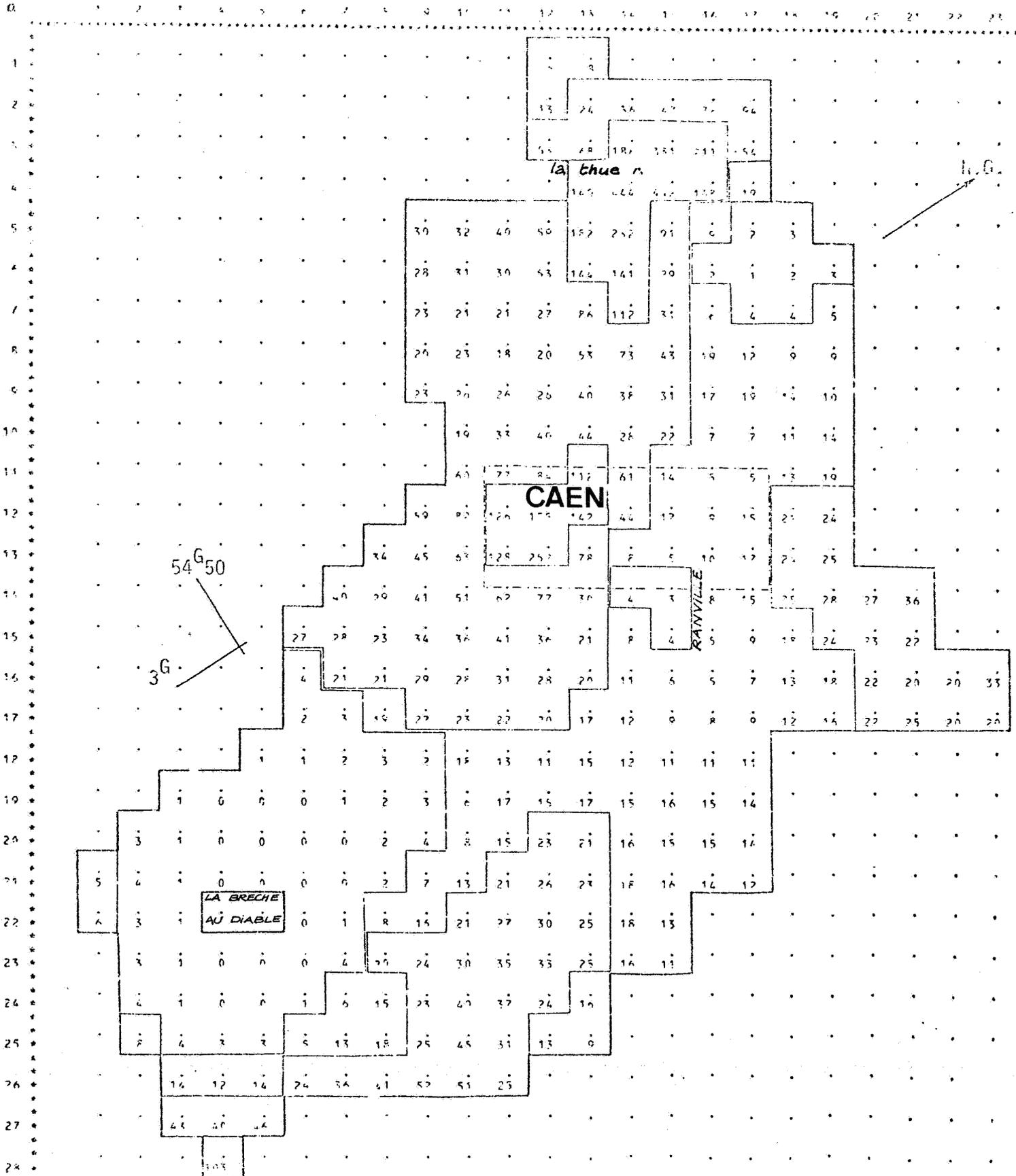
Afin de visualiser cette baisse, on a calculé et reporté, pour chaque maille, le facteur d'amélioration de la connaissance de la transmissivité correspondant au passage des seuls essais de pompage à l'ensemble englobant les valeurs obtenues par régression. Si ce facteur reste sensiblement égal à 1. dans la région caennaise bien connue à partir des seuls essais de pompage, il passe à des valeurs supérieures à 10. du côté de la Dives et de la Haute-Mue, où l'appoint des débits spécifiques a été décisif (cf. Fig. 44, p. 98 dans le corps du texte).

- Enfin, l'importance de la fenêtre de la Brèche au Diable sur le modèle des transmissivités est apparue telle que l'on a repris l'estimation en supprimant le point la représentant: on peut constater sur la Fig. A-11 que tout le tiers Sud-Est en est affecté.

Si la carte obtenue en tenant compte de la présence de la fenêtre semble plus satisfaisante (zone à faible transmissivité allongée dans le sens des barres de grès armoricain entourant la fenêtre), il faut bien admettre que l'on a ajouté, dans cette région, un point informant. Seul le modèle permettra de vérifier l'hypothèse faite.

FIG. A-9 - TRANSMISSIVITES DU BATHONIEN  
 ESTIMATEUR MEDIAN (en  $10^{-4}m^2/s$ ) OBTENU PAR KRIGEAGE  
 AVEC L'ENSEMBLE DES DONNEES

VALUES REALES & VALUES ESTIMES MULTICRONS PAR LISTED DATA







• Aquifère du Bajocien:

- Le faible nombre de données n'a pas justifié une étude aussi poussée que pour le Bathonien.

A noter seulement que les intervalles de confiance sont du même ordre de grandeur que ceux obtenus pour le Bathonien, malgré le nombre moindre de données; et ce, du fait de la plus grande régularité du phénomène, du moins en apparence.

Les zones bien connues sont centrées sur les régions de Caen et Bayeux. Les fortes transmissivités apparaissant dans les basses vallées de la Thucet de la Mue doivent être considérées avec précaution du fait de la grande amplitude des intervalles de confiance correspondants (extrapolation en direction de la mer de deux fortes transmissivités obtenues dans la région par essais de pompage)

Le travail, qui a été réalisé sur les transmissivités, a permis de tirer parti au maximum de l'information disponible à partir des essais de pompage et des calculs de débits spécifiques.

Des zones d'ombre subsistent, là où les intervalles de confiance restent très grands: la connaissance de la transmissivité ne pourra y progresser qu'à travers son rapport avec le modèle piézométrique, lors du calage du modèle d'écoulement souterrain.

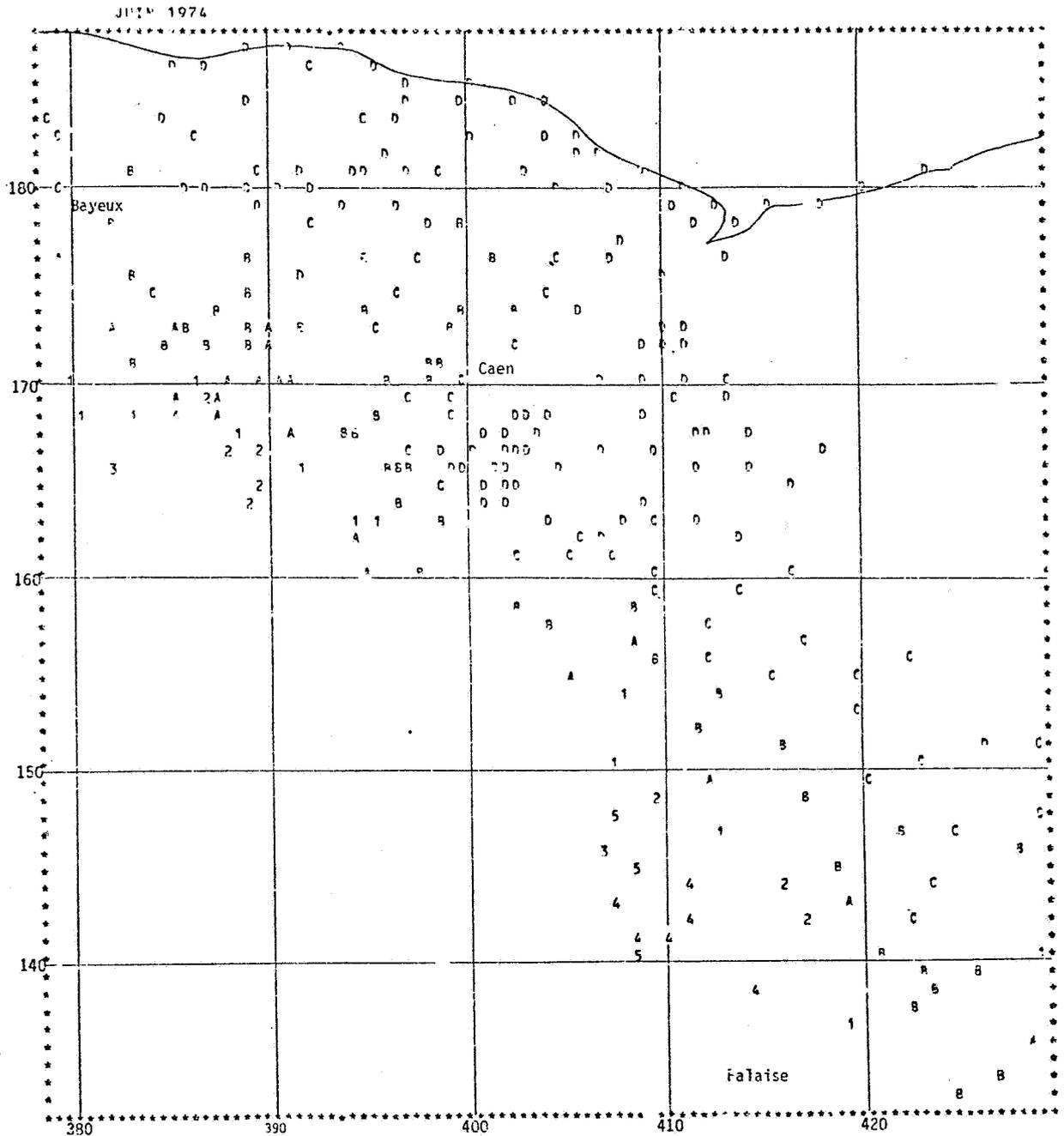
ETUDE DE LA PIEZOMETRIE

Afin de disposer d'états piézométriques de référence pour le calage du modèle, on a estimé par krigeage la piézométrie moyenne sur des mailles de 2 km x 2 km aux 3 dates suivantes :

- Juin 1974 où l'on dispose de 208 relevés piézométriques
- Octobre 1974 où l'on dispose de 201 relevés piézométriques
- Avril 1974 où l'on ne dispose que de 176 relevés, avec une absence de données du côté de la Dives.

La piézométrie a dû être étudiée, les deux nappes confondues, en l'absence de mesures séparées dans les zones où existe un bicouche.

FIG. A-12 - PIEZOMETRIE DU DOGGER  
 DONNEES DISPONIBLES POUR JUIN 1974



94 CARACTERES/LIGNE		41 LIGNES	
X HORIZONTAL VARIE DE GAUCHE A DROITE, ENTRE	378,7 ET 428,7	FCH 1, 5376	UNITES/CARACTERE
Y VERTICAL VARIE DE BAS EN HAUT, ENTRE	133,9 ET 187,0	FCH 1, R917	UNITES/LIGNE
POUR LES	254 VALEURS PRESENTES DANS LE DOMAINE CONSIDERE, IL Y A	251 POINTS REPRESENTATIFS DISTINCTS	
1 REPRESENTE LES VALEURS COMPRISES ENTRE	80,00	ET	100,00
2 REPRESENTE LES VALEURS COMPRISES ENTRE	100,00	ET	120,00
3 REPRESENTE LES VALEURS COMPRISES ENTRE	120,00	ET	140,00
4 REPRESENTE LES VALEURS COMPRISES ENTRE	140,00	ET	160,00
5 REPRESENTE LES VALEURS COMPRISES ENTRE	160,00	ET	180,00
A REPRESENTE LES VALEURS COMPRISES ENTRE	60,00	ET	80,00
B REPRESENTE LES VALEURS COMPRISES ENTRE	40,00	ET	60,00
C REPRESENTE LES VALEURS COMPRISES ENTRE	20,00	ET	40,00
D REPRESENTE LES VALEURS COMPRISES ENTRE	0,0000	ET	20,00

Cette piézométrie reflète, en fait, les parties libres des 2 nappes. Aux relevés piézométriques disponibles, on a rajouté quelques points correspondant à des endroits où les cours d'eaux (27) ou la mer (19) imposent le potentiel de la nappe, ce qui porte le nombre de données à 254 par exemple pour Juin 1974 (cf. Fig. A-12).

Les caractéristiques structurales qui ont été déterminées à l'aide de la procédure automatique issue de la théorie des FAI-k et utilisées pour le krigeage sont:

- dérive linéaire,
- covariance généralisée sans effet de pépite (phénomène continu)

$$K(h) = - 10. |h| + 0.75 |h|^3$$

On peut constater sur la figure 32 (p. 85 dans le corps du texte) que les modèles piézométriques sont bien rendus par le krigeage.

Les écarts types d'estimation sont de l'ordre de 3 mètres en général, souvent inférieurs, avec quelques zones où leurs valeurs élevées coïncident avec une absence totale de données (Vallée de la Dive en Avril 1975)

Il est à noter dès à présent que le fait que les écarts-types d'estimation de la piézométrie sont toujours supérieurs au mètre indique que l'on devra envisager une précision de calage de l'ordre métrique et non une précision centimétrique qui serait sans rapport avec le degré réel de connaissance du phénomène : on est en effet en présence d'une nappe libre en pays calcaire au modèle très tourmenté, et l'échantillonnage piézométrique est relativement lâche.

ETUDE DE LA GEOMETRIE DES AQUIFERES

Ont été estimées enfin les valeurs moyennes sur des mailles de 2 x 2 km de:

- l'épaisseur des marnes de Port en Bessin,
- la cote du mur du Bathonien,
- la cote du mur du Bajocien.

a) Epaisseur des marnes de Port en Bessin:

Cette épaisseur était disponible en 36 forages

Les caractéristiques structurales mises en évidence sont les suivantes:

- dérive linéaire,
- $K(h) = - 38 \cdot |h|$  (h en km)

Les résultats du krigeage montrent le biseautage des marnes vers le Sud-Est (cf. Fig. A-13).

Les écarts-types assez forts (> 10 m) dans cette zone indiquent une incertitude sur la position du biseau de l'ordre à peu près d'une maille (2 km).

b) Mur du Bathonien:

On n'a retenu en principe que les forages ayant effectivement atteint le mur.

Aux 97 données disponibles, on a été amené à rajouter 2 indications imprécises, fournies par 2 forages n'ayant pas atteint le mur, mais situés dans des zones mal reconnues. On a ainsi introduit dans les données la cote  $-99 \pm 50$  au forage 120.8.001 à Mery Corbon sur le Laison (de façon à indiquer le pendage vers le N.E. au sud de la grille) et la cote  $-89 \pm 10$  au forage 120.2.006 à Gonnevillie en Auge.

On a, par ailleurs, tenu compte dans le krigeage de l'existence de failles sensiblement orientées dans la direction américaine:

- une seule à Ranville (faute de données permettant de tenir compte d'un éventuel forst);
- un jeu de deux failles entourant Amolie.



Les paramètres structuraux obtenus sont:

- dérive linéaire,
- $K(h) = -23|h| + 4.4|h|^3$

Les résultats (Fig. 46, p. 100 dans le corps du texte) concordent assez bien avec une carte tracée à la main par les géologues du SGR (BRGM) figurant dans la documentation recueillie.

A noter que la cote du mur n'est réellement bien connue (écarts-types d'estimation < 10 m) que dans les régions de Caen et de la Seulles, de la Thue et de la Mue.

c) Mur du Bajocien:

67 forages sont disponibles dont un bon nombre dans la partie libre de la nappe bajocienne.

Une carte tracée à la main par un géologue ferait intervenir la similitude avec les couches supérieures.

Ne travaillant ici qu'avec les seules cotes obtenues aux forages ayant atteint le Bajocien, on a dû introduire les données fictives suivantes:

- $23 \pm 10$  à Creully (119.2.213)
- $110 \pm 30$  à Bures sur Dives (120.7.088)
- $130 \pm 30$  à Gonneville en Auge (120.2.006)
- $130 \pm 50$  à Mery-Corbon (120.8.001)

pour tenir compte de l'information de type "géologue".

Le manque de données n'a permis en outre de ne prendre en considération qu'une seule faille, celle de Ranville.





Les paramètres structuraux obtenus par l'indentification structurale automatique (Fig. A-14) sont:

- dérive linéaire,
- $K(h) = - 100 \cdot |h|$

Les résultats (Fig. A-15) obtenus par krigeage concordent assez bien, compte-tenu des remarques précédentes, avec un tracé manuel disponible dans la documentation BRGM.