



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

1360285421

91 INPL003N

INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE

Ecole Nationale Supérieure
de Géologie de Nancy

Ecole Nationale Supérieure
des Mines de Nancy

LABORATOIRE
DE GEOMECHANIQUE

LABORATOIRE DE
MECANIQUE DES TERRAINS

----- [M] 1991 KANG, Y.S.

T H E S E

Présentée à l'I.N.P.L.
en vue de l'obtention du titre de

**DOCTEUR DE L'I.N.P.L.
EN GENIE GEOLOGIQUE ET MINIER**

par

Yong Shang KANG



**ANALYSE DES CHRONIQUES: APPLICATION
A L'EVOLUTION TEMPORELLE DE L'EXHAURE
MINIERE DANS LE BASSIN FERRIFERE LORRAIN**

Soutenu publiquement le 14 Janvier 1991 devant la Commission d'Examen

Membres du Jury :

M. Cl. CHAMBON

Président

M. J.P. LABORDE

Rapporteur

M. J.J. ROYER

Rapporteur

M. R. HOUPERT

Examineur

M. M. FEUGA

Examineur

M. S. RAMON

Examineur

M. L. DEMASSIEUX

Examineur

TABLE DES MATIERES

Avant-Propos

Résumé

Introduction..... 1

1ère Partie - ANALYSE DES CHRONIQUES OU SERIES TEMPORELLES

Chapitre 1 - Théorie des processus stochastiques et des systèmes linéaires

I	Processus stochastiques et leurs propriétés.....	3
I.1	Processus stochastique $Y(t)$	3
I.2	Processus stochastique stationnaire $Y(t)$	4
II	Systèmes linéaires et leurs propriétés.....	6
II.1	Linéarité et stationnarité.....	6
II.2	Représentation et propriétés d'un système linéaire et stationnaire.....	7
II.2.1:	Relation de convolution.....	8
II.2.2 :	Propriétés d'un système linéaire et stationnaire.....	9
III	Traitement du signal et analyse des chroniques.....	13
III.1	Définitions 13	
III.1.1 :	Filtrage.....	13
III.1.2 :	Lissage.....	14
III.1.3 :	Prévision (ou prédiction).....	14
III.1.4 :	Identification.....	14
III.2	Analyse des chroniques.....	14

Chapitre 2 - Outils descriptifs utilisés dans l'analyse des chroniques

I	Coefficient d'autocorrélation et coefficient de corrélation croisée (analyse corrélatrice).....	17
I.1	Coefficient d'autocorrélation - corrélogramme.....	17
I.2	Coefficient de corrélation croisée.....	19
II	Spectre énergétique et spectre énergétique croisé.....	20
II.1	Transformée de Fourier discrète et spectre linéaire.....	21
II.2	Spectre énergétique et spectre énergétique croisé.....	22
III	Applications des analyses corrélatrice et spectrale.....	26

Chapitre 3 - Méthodes de prévision univariées

I	Méthodes de Lissage (modèles autoprojectifs).....	30
I.1	Méthodes de lissage par moyenne mobile.....	31
I.2	Méthodes de lissage exponentiel.....	32
II	Méthodes de décomposition (modèles d'ajustement).....	35

III	Méthode Univariée de Box-Jenkins (ARIMA)	37
III.1	Modèles ARMA, ARIMA et SARIMA.....	38
III.1.1	Généralité.....	38
III.1.2	Modèles ARMA(p,q).....	39
	1) Modèle autorégressif d'ordre p : AR(p).....	39
	2) Modèle "moyenne mobile" MA(q).....	41
	3) Modèles ARMA(p,q).....	42
III.1.3	Modèles ARIMA(p,d,q) et SARIMA.....	44
III.2	Méthodologie de Box-Jenkins.....	45
III.2.1	Identification.....	45
III.2.2	Estimation.....	45
III.2.3	Validation (Diagnostic).....	46
III.2.4	Prévision.....	46
IV	Conclusion	47

Chapitre 4 - Méthodes de prévision multivariées

I	Méthode de Régression Multiple	50
I.1	Méthode linéaire classique.....	50
I.1.1	Estimation des paramètres $a_0, a_1, a_2, \dots, a_j, \dots, a_p$ (ajustement des Moindres-Carrés).....	52
I.1.2	Equation d'analyse de variance et coefficient R^2	53
I.1.3	Intervalles de confiance et tests de signification pour les coefficients.....	55
I.1.4	Prévision.....	56
I.2	Méthodes des Moindres-Carrés Généralisées.....	57
II	Méthode Multivariée de Box-Jenkins (MARIMA)	58
III	Méthode de déconvolution par projection [60]	61
IV	Méthode de déconvolution $COV_x(t) - COV_y(t)$	64
IV.1	Rappel sur la théorie des variables régionalisées.....	64
IV.1.1	Variogramme.....	64
IV.1.2	Comportement du graphe $g(t)-t$	65
IV.1.3	Ajustement à un modèle.....	67
IV.2	Méthode de déconvolution géostatistique [73].....	68
V	Conclusion	69

2ème Partie - ANALYSE DES DONNEES D'EXHAURE MINIERE DANS LE BASSIN FERRIFERE LORRAIN

Chapitre 5 - Etude bibliographique sur l'exhaure minière dans le bassin ferrifère lorrain

I	Bassin ferrifère lorrain - Généralités	72
I.1	Situation géographique - cadre géologique.....	72
I.2	Situation structurale.....	72
I.3	Lithologie et stratigraphie.....	74
I.4	Situation climatique.....	74
I.5	Réseau hydrographique.....	77

II	Travaux effectués visant à connaître le phénomène "exhaure minière"	77
II.1	Travaux de 1925 à 1957.....	78
II.2	Travaux effectués de 1958 à 1969.....	81
II.3	Travaux effectués après 1970.....	83
III	Commentaires et conclusions	90

*Chapitre 6 - Analyse des données d'exhaure minière
dans le bassin ferrifère lorrain*

I	Données et Leur Traitement Préalable	96
I.1	Acquisition et commentaires des données.....	96
I.2	Calcul des pluies efficaces.....	97
II	Relation Exhaure-Pluie Efficace, Extraction de Minerai et Réserves Souterraines	100
II.1	Pluies efficaces, réserves souterraines et exhaure.....	100
II.1.1	: Analyse des pluies efficaces.....	100
II.1.2	: Analyse d'exhaure minière.....	105
II.1.3	: Relation exhaure minière-réserves souterraines et pluie efficace.....	105
II.2	Effet de l'extraction de minerai sur l'exhaure.....	110
II.3	Simulation de l'exhaure-modèle physique.....	112
III	Modélisation Numérique	113
III.1	Mines stationnaires.....	113
III.2	Mines non stationnaires.....	114
III.3	Discussion sur les termes des modèles établis.....	127
III.3.1	: Coefficient de tempérament b ($0 \leq b \leq 1$).....	127
III.3.2	: Terme constant C , mémoire m et fonction de transfert.....	128
III.3.3	: Terme résiduel e_t	131
IV	Comparaison de Différentes Méthodes de Déconvolution	135
V	Interprétation des Résultats et Conclusion	139

Chapitre 7 - Conclusions Générale

I	Quelques remarques	144
II	Bilan des résultats	145
III	Perspectives	147
	Bibliographie	150
	Annexe I	157
	Annexe II	161
	Annexe III	165

LISTE DES FIGURES

Chapitre 1 :

Fig. 1	Processus stochastique.....	3
Fig. 2	Système	6
Fig. 3	Illustration des systèmes à contre réaction et sans contre-réaction.....	10
Fig. 4	Filtrage	12
Fig. 5	Filtrage multi-composantes	13

Chapitre 2 :

Fig. 6	Corrélogramme.....	18
Fig. 7	Signal échantillonné	22
Fig. 8	Analyse de la chronique des pluies à la station de Balgué (extrait de [56])	27
Fig. 9	Analyse des chroniques des débits de trois systèmes karstiques : Fontestorbes, Baget et Aliou (extrait de [56])	27
Fig. 10	Corrélogramme croisé et spectre croisé (extrait de [57])	28

Chapitre 3 :

Fig. 11	Lissage exponentiel linéaire.....	33
Fig. 12	Lois exponentielles	35
Fig. 13	Loi logistique.....	35
Fig. 14	Filtre linéaire.....	38
Fig. 15	Enveloppe E	41
Fig. 16	Fonctions d'autocorrélation et d'autocorrélation partielle (Extrait de [82])	43
Fig. 17	Méthode de Box-Jenkins (Extrait de [24]).....	45

Chapitre 4 :

Fig. 18	Schéma du système discret	49
Fig. 19	Ajustement des Moidres-Carrés (extrait de [18]).....	52
Fig. 20	Analyse de la variance (extrait de [18]).....	54
Fig. 21	Ajustement et prévision (extrait de [49])	56
Fig. 22	Fonctions de Walsh (extrait de [60])	62
Fig. 23	Fonction de transfert et tarissement (extrait de [60]).....	63
Fig. 24	Comportement du variogramme au voisinage de l'origine(extrait de [32]).....	66
Fig. 25	Comportement du variogramme à l'infini (extrait de [32]).....	66
Fig. 26	Principaux modèles de variogrammes (extrait de [32]).....	67

Chapitre 5 :

Fig. 27	Plateaux de Briey et de Haye	73
Fig. 28	Coupe schématique du Bajocien et du Bathonien (Bassin de Briey et Longwy) ..	75
Fig. 29	Coupe géologique de la formation ferrifère (Mine d'Auboué, Moineville, Siège du Paradis, Puits 4)	76
Fig. 30	Analyse en composantes principales de l'exhaure des mines dans le bassin de l'Orne (extrait de [13]).....	86
Fig. 31	Exemples des courbes de tarissement (extrait de [13])	87
Fig. 32	Réponse impulsionnelle et reconstitution de l'exhaure pour la mine de Valleroy de 1959 à 1964 (extrait de [13])	88
Fig. 33	Report des teneurs en sodium par siège sur la carte structurale du bassin (d'après DEMASSIEUX).....	91

Chapitre 6 :

Fig. 34	Carte des mines (extrait de [8]).....	98
Fig. 35	Coupe géologique (extrait de [8]).....	99
Fig. 36	Pluies efficaces à Tomblaine et à Frescaty	101
Fig. 37	Corrélogramme de la pluie efficace à Frescaty (1945-1985).....	102
Fig. 38	Modèle ARIMA de la pluie efficace mensuelle à Frescaty (1945-1985).....	103-104
Fig. 39	Exhaure mensuelle et la moyenne mobile centrée à 12 termes pour trois mines dans le bassin ferrifère lorrain	106
Fig. 40	Corrélogrammes de l'exhaure	107
Fig. 41	Comparaison entre la pluie efficace, la hauteur piézométrique et l'exhaure minière à Roncourt (extrait de [53])	108
Fig. 42	Comparaison entre les composantes saisonnières de la pluie efficace et de l'exhaure.....	109
Fig. 43	Exploitation et exhaure (extrait de [8])	111
Fig. 44	Simulation de l'exhaure	113
Fig. 45	Mines stationnaires	114
Fig. 46	Déconvolution à Moyeuivre ($Q_t - E_t$).....	115
Fig. 47	Mines non stationnaires	116
Fig. 48	Déconvolution à Saizerais ($Q_t - S_t, E_t$).....	117
Fig. 49	Système avec exciteur	118
Fig. 50	Extraction mensuelle cumulée et transformation de l'exhaure	118
Fig. 51	Déconvolution à Saizerais ($\frac{Q_t}{S_t} - E_t$).....	119
Fig. 52	Déconvolution à Anderny ($\frac{Q_t}{S_t^{0,3}} - E_t$).....	121
Fig. 53	Déconvolution à Saizerais ($\frac{Q_t}{MA12} - E_t$).....	122
Fig. 54	Fonction de transfert (régression multiple)	125-126
Fig. 55	Exhaure et le coefficient de tempérament b	128
Fig. 56	Fonction de transfert (mine Moyeuivre)	130
Fig. 57	Fonction de transfert.....	130
Fig. 58	Ajustement de la fonction de transfert à une courbe exponentielle (mine Moyeuivre).....	131
Fig. 59	Fonction de transfert (mine Saizerais).....	131
Fig. 60	Modèle ARIMA des résidus de régression multiple à Moyeuivre.....	133-134
Fig. 61	Corrélogramme des résidus calculés par régression dynamique (mine Saizerais)	135
Fig. 62	Fonction de transfert (déconvolution par projection).....	137-138
Fig. 63	Représentation de la mémoire graphique et du coefficient B sur la carte des mines	140
Fig. 64	Superficies d'alimentation effective et potentielle.....	141
Fig. 65	Evolution de l'exhaure dans le temps.....	142
Fig. 66	Comparaison entre la pluie efficace mensuelle (à Frescaty) et l'exhaure mensuelle totale dans le bassin (1946-1983)	142

Chapitre 7 :

Fig. 67	Exhaure mesurée et reconstituée pour la mine Saizerais (1958-1980).....	148
Fig. 68	Le comportement de l'exhaure après une période sèche.....	148

LISTE DES TABLEAUX

Chapitre 6 :

Tab. I	Classement des sièges en fonction du risque de minéralisation sodique (extrait de [26]).....	92
Tab. II	Les données d'exhaure et d'extraction de minerai	96
Tab. III	Les données météorologiques	97
Tab. IV	Statistiques des pluies efficaces	101
Tab. V	Coefficients de corrélation entre exhaure et ses facteurs influents	110
Tab. VI	Résultats de déconvolution	123
Tab. VII	Coefficients en pourcentage	124
Tab. VIII	Déconvolution par projection.....	136

AVANT-PROPOS

Le travail présenté dans ce mémoire a été réalisé au Laboratoire de Mécanique des Terrains et au Laboratoire de Géomécanique.

La direction effective de mon travail a été assurée par Monsieur CHAMBON, Professeur à l'Ecole des Mines de Nancy. Je tiens à lui exprimer ma profonde gratitude pour ses conseils avisés et les facilités qu'il m'a accordés pour mener à bien ce travail. Je le remercie sincèrement de présider le jury de thèse, mais plus encore de m'avoir en grande partie initié aux statistiques et surtout de m'en avoir donné le goût.

Monsieur HOUPERT, Professeur à l'Ecole Nationale Supérieure de Géologie, m'a accueilli dans son Laboratoire et m'a permis d'effectuer ce travail dans une atmosphère chaleureuse. Aujourd'hui, il me fait l'honneur de participer au jury. Qu'il me soit permis de lui témoigner ma respectueuse reconnaissance.

Monsieur DEMASSIEUX, Maître de Conférences à l'Ecole Nationale Supérieure de Géologie, a suivi avec attention et bienveillance mon travail. Je lui adresse mes plus vifs remerciements pour ses conseils et son soutien sans lesquels ce travail n'aurait pas pu être mené à bien.

Monsieur LABORDE, Professeur à l'Université de Nice, m'a enseigné l'Hydrologie. Ses travaux de recherche sur les débits et sur les précipitations dans la région lorraine ont été pour moi une source permanente d'enrichissement. Il me fait l'honneur d'examiner mon travail en tant que rapporteur.

Monsieur ROYER, Chargé de Recherche au Centre de Recherches Pétrographiques et Géochimiques, m'a beaucoup encouragé dans la recherche. Ses conseils et ses critiques éclairés m'ont été très précieux. Je tiens à le remercier de l'intérêt qu'il a porté à mon travail en acceptant d'en être le rapporteur.

Monsieur RAMON, Ingénieur à l'Agence de l'Eau Rhin-Meuse, m'a communiqué toutes les données nécessaires pour effectuer ce travail. Aujourd'hui, il est présent dans le jury. Qu'il trouve ici l'expression de mon entière reconnaissance.

Monsieur FEUGA, Directeur du Centre d'Etudes Supérieures des Techniques Minières, a bien voulu juger ce travail. Je lui adresse mes remerciements.

Je remercie sincèrement tous les Professeurs et Enseignants au Laboratoire de Géomécanique et au Laboratoire de Mécanique des Terrains.

Mes remerciements s'adressent également à Madame RAMELLA qui a tapé plusieurs chapitres de ce texte.

Que tous mes camarades au Laboratoire de Géomécanique et au Laboratoire de Mécanique des Terrains trouvent mes remerciements pour leur aide et leur amitié.

Que tous mes amis chinois trouvent mes remerciements pour leur aide et leur amitié.

RESUME

L'analyse des chroniques présente deux aspects complémentaires: description et prévision.

L'analyse descriptive d'une chronique est une étape préliminaire à toute modélisation et prévision. Les outils descriptifs couramment utilisés sont: le corrélogramme, le corrélogramme croisé, le spectre énergétique et le spectre énergétique croisé.

La prévision d'une chronique présente souvent un intérêt principal. Elle peut se faire soit à partir des informations du passé sur la chronique elle-même, soit à partir d'une ou plusieurs autres chroniques, d'où on distingue les méthodes de prévision univariées et les méthodes de prévision multivariées.

Les méthodes de prévision univariées sont les méthodes de lissage, de décomposition et de Box-Jenkins(ARIMA). Ces méthodes peuvent convenir dans bien des cas, et en particulier dans le court terme, lorsqu'un phénomène continue sur son élan et ne change guère, dans cet horizon de temps. En revanche, elles peuvent s'avérer mal adaptées au moyen et au long terme.

Les méthodes de prévision multivariées cherchent à saisir et à mesurer l'influence d'une ou plusieurs chroniques explicatives sur la chronique en jeu Y_t . Sans changer la généralité, on restreint notre propos au cas d'une seule chronique explicative X_t . X_t peut être considéré comme l'entrée d'un système et Y_t la sortie. La relation de convolution entre X_t et Y_t est généralement admissible. La fonction de transfert représente toutes les caractéristiques du système. On peut ainsi construire la sortie correspondante à toute entrée en convoluant celle-ci par la fonction de transfert, et pour cela, il faut connaître au préalable la fonction de transfert. Le problème de prévision se convertit alors souvent en un problème de déconvolution. Les méthodes de déconvolution classiques et couramment utilisées sont les méthodes spectrales et les méthodes du type "*problème inverse*". Ces méthodes fonctionnent bien lorsque l'entrée et la sortie vérifient parfaitement une relation de convolution. Dans la pratique, cependant, les conditions de linéarité et d'invariance ne sont jamais parfaitement remplies et de plus, les observations d'entrée et de sortie sont souvent entachées d'erreurs. La relation de convolution doit s'écrire de façon à faire intervenir un terme résiduel. L'intervention du terme résiduel provoque une grande difficulté de déconvolution, puisqu'en fait, la solution n'est pas unique. C'est pour surmonter cette difficulté que quelques nouvelles méthodes ont été proposées récemment(méthodes de projection, méthode géostatistique, etc.). Par ailleurs, il nous semble que la méthode de fonction de transfert de Box-Jenkins est très prometteuse, car elle combine

les concepts de régression multiple aux modèles ARIMA, ce qui présente théoriquement un avantage décisif de cette méthode

Comme une application de l'analyse des chroniques, l'analyse des données mensuelles d'exhaure minière dans le bassin ferrifère lorrain nous permet de distinguer deux types de mines: les mines stationnaires et les mines non stationnaires. Un modèle général reliant l'exhaure à ses principaux facteurs influents(pluie efficace et extraction cumulée de minerai) a été établi.

La non stationnarité de l'exhaure dans le temps peut être expliquée par l'extraction cumulée de minerai. L'hétérogénéité de l'exhaure dans l'espace peut être expliquée généralement par les conditions géologiques locales et les conditions structurales peuvent expliquer en partie les anomalies.

Le modèle établi peut servir à prévoir l'exhaure minière à une date future ou à reconstituer les données d'exhaure.

Par ailleurs, il faut remarquer que dans la gestion du bassin, la sécheresse pose souvent de graves problèmes à l'alimentation d'eau potable. Il est important d'étudier les comportements spécifiques des nappes souterraines ou des mines vis-à-vis de la sécheresse annuelle ou pluriannuel. Un modèle d'intervention pourra être adopté. Il s'agira de formuler des critères sur la sécheresse, d'introduire des variables indicatrices, d'étudier les comportements de l'exhaure dans la période de transition.

INTRODUCTION

L'exhaure minière dans le bassin ferrifère lorrain est un phénomène important qui présente de l'intérêt pour divers organismes: les Exploitants miniers (gérer plus rationnellement les cadences de pompage comme la politique des dépilages), le Service des Mines, l'Agence de l'Eau Rhin-Meuse (alimentation en eau potable, aménagement du bassin), *etc.*

De nombreux travaux, destinés à l'étude de ce phénomène, ont été effectués : conditions climatiques, géologiques, structurales, topographiques ; écoulement superficiel ; caractéristiques des nappes et de circulation hydraulique ; modélisation de l'exhaure et qualité des eaux exhaurées.

Relativement au problème de modélisation, CHANZY (Ingénieur des Mines) avait déjà pris l'initiative de simuler l'exhaure à partir des pluies dans les années vingt, et cette simulation consiste à postuler qu'il existe une courbe de tarissement unique pour toutes les périodes sèches. Dans les années soixante-dix, CANCELL (B.R.G.M.) a mis en cause la méthode de CHANZY et montré à l'évidence que la relation pluie-exhaure peut être correctement représentée par une relation de convolution. Mais cette dernière approche, ignorant l'hétérogénéité spatiale de l'exhaure (pour les différentes mines dans le bassin), se limite au cas d'une "exhaure-type" (stationnaire).

Cependant, bien que l'exhaure soit stationnaire pour certaines mines, elle ne l'est pas pour d'autres et cette non stationnarité semble être conditionnée par l'exploitation minière elle-même. Donc, la relation de convolution n'est pas immédiatement applicable pour toutes les mines dans le bassin.

Ce mémoire présente notre travail destiné à établir un modèle général, adapté pour toutes les mines dans le bassin ferrifère lorrain, à partir de la relation de convolution. Ce modèle général doit tenir compte, évidemment, à la fois de l'hétérogénéité spatiale et de la non stationnarité temporelle de l'exhaure.

La première partie de ce mémoire est consacrée au problème d'analyse des chroniques : dans le premier chapitre, nous rappelons les théories de processus et de systèmes ; dans le deuxième chapitre, nous présentons les outils descriptifs couramment utilisés dans l'analyse des chroniques ; dans le troisième chapitre, nous discutons les méthodes de prévision univariées ; dans le quatrième chapitre, nous abordons les méthodes de prévision multivariées (problème de déconvolution).

Le problème d'exhaure minière dans le bassin ferrifère lorrain fait l'objet de la deuxième partie de ce mémoire : dans le cinquième chapitre, nous examinons brièvement ce qu'apportent les travaux antérieurs pour l'étude de l'exhaure ; dans le sixième chapitre, nous proposons d'effectuer une analyse exhaustive sur les données mensuelles d'exhaure de toutes les mines.

Pour clore ce mémoire, dans le septième chapitre, nous formulons d'abord quelques remarques sur l'analyse des chroniques. Un bilan des résultats d'analyses des données d'exhaure est ensuite présenté. Nous proposons enfin des perspectives relatives à l'exploitation de la "*banque*" des données d'exhaure dans le but d'aménagement du bassin.

Première partie

**ANALYSE DES CHRONIQUES
OU SERIES TEMPORELLES**

Chapitre 1

S O M M A I R E

I	Processus stochastiques et leurs propriétés	3
I.1	Processus stochastique $Y(t)$	3
I.2	Processus stochastique stationnaire $Y(t)$	4
II	Systèmes linéaires et leurs propriétés	6
II.1	Linéarité et stationnarité	6
II.2	Représentation et propriétés d'un système linéaire et stationnaire.....	7
	II.2.1: Relation de convolution.....	8
	II.2.2 :Propriétés d'un système linéaire et stationnaire	9
III	Traitement du signal et analyse des chroniques	13
III.1	Définitions	13
	III.1.1 : Filtrage.....	13
	III.1.2 : Lissage.....	14
	III.1.3 : Prévission (ou prédiction).....	14
	III.1.4 : Identification	14
III.2	Analyse des chroniques	14

CHAPITRE 1

THEORIE DES PROCESSUS STOCHASTIQUES ET DES SYSTEMES LINEAIRES

Une chronique définie comme une suite de données (ponctuelles ou agrégées) datées et ordonnées dans le temps est en effet une réalisation d'un processus stochastique qui s'appelle aussi "*processus aléatoire ou fonction aléatoire*". C'est dans ce sens-là qu'une chronique peut être également appelée une suite temporelle aléatoire. Par ailleurs, une chronique ou un processus stochastique est lié intimement à la notion de systèmes.

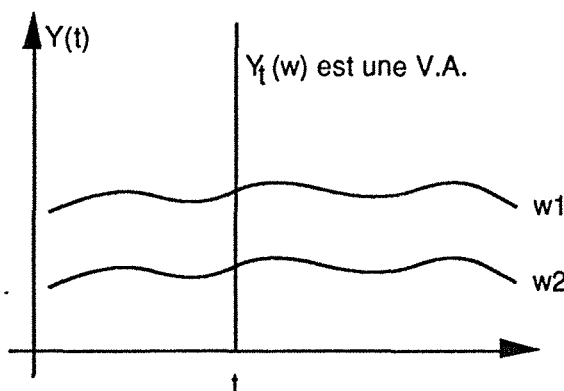
Dans ce premier chapitre, nous nous plaçons dans un contexte général. Nous rappelons d'abord très succinctement la théorie de processus stochastiques. Notre propos se limite évidemment au cas unidimensionnel. Nous désignons, dans toute la suite, un processus stochastique par $Y(t)$ et une chronique ou une variable aléatoire par Y_t . Ensuite, nous introduisons la notion de systèmes. La discussion des propriétés porte essentiellement sur les systèmes linéaires et stationnaires.

Enfin, nous précisons quelques termes bien souvent utilisés dans le traitement du signal et fixons notre objet d'étude dans l'analyse des chroniques.

I - PROCESSUS STOCHASTIQUES ET LEURS PROPRIETES

I.1 - Processus stochastique $Y(t)$

On appelle processus stochastique, une famille infinie (Y_t) et $t \in T$ de variables aléatoires Y_t , l'indice t parcourant l'ensemble T infini.



Pour t fixé, Y_t est une V.A. Si l'on fixe w , $Y_t(w)$ est une fonction mesurable de $T \rightarrow \mathbb{R}$, qui est appelée trajectoire du processus stochastique $Y(t)$ [72].

Fig. 1 - Processus stochastique.

Si de plus, la condition $E [| Y(t) |^2] < \infty$ est vérifiée, le processus stochastique est dit du second ordre. Cela assure également l'existence de l'espérance mathématique $E [| Y(t) |] < \infty$. Dans de nombreuses applications, nous nous contentons des deux premiers moments.

L'étude des processus stochastiques du second ordre porte notamment sur la caractérisation des propriétés de leurs premier et deuxième moments.

Le moment du premier ordre de $Y(t)$ est en effet son espérance mathématique :

$$m_Y(t) = E [Y(t)] \quad (1.1)$$

et celui du deuxième ordre sa matrice de variance-covariance dont un élément défini entre deux instants t_i, t_j s'écrit :

$$K_{YY}(t_i, t_j) = E [Y(t_i) Y(t_j)] \quad (1.2)$$

En physique, $K(t_i, t_j)$ possède toujours une signification énergétique. L'étude du moment de second ordre d'un processus stochastique $Y(t)$ concerne essentiellement les propriétés énergétiques de $Y(t)$ et le transfert d'énergie lorsque $Y(t)$ est impliqué dans une équation de convolution, d'où provient une propriété importante: $K_{YY}(t_i, t_j)$ est une fonction du type positive, c'est-à-dire que la relation suivante est vérifiée [4, 62, 72] :

$$\sum_i \sum_j a_i a_j K_{YY}(t_i, t_j) \geq 0 \quad (1.3)$$

... où a_i, a_j sont des coefficients réels ou complexes.

I.2 - Processus stochastique stationnaire $Y(t)$

On dit que $Y(t)$ est stationnaire au sens strict si sa loi temporelle est invariante pour toute translation dans le temps.

Cependant, cette définition paraît beaucoup trop restrictive pour être généralement admissible dans les applications. Une autre définition plus modeste est souvent applicable : c'est le processus dit stationnaire du second ordre.

Un processus $Y(t)$ est stationnaire du second ordre si les conditions suivantes sont vérifiées :

1. La moyenne de $Y(t)$ ne dépend pas de t , c'est-à-dire $E [Y(t)] = m$ [notant que $Y(t)$ est toujours remplaçable par $Y(t) - m$, nous supposons par la suite que m est identiquement nulle].

2. La covariance ne dépend que de l'intervalle de temps $\tau = t_j - t_i$, c'est-à-dire :

$$K_{YY}(t_i, t_j) = E [Y(t_i) Y(t_j)] = K_{YY}(\tau)$$

Il se peut que la stationnarité du second ordre ne vérifie pas celle au sens strict. Dans beaucoup d'applications cependant, on n'utilise que des propriétés liées aux moments des deux premiers ordres de sorte qu'il n'y a pas de distinction entre processus stationnaire au sens strict et celui du second ordre. Dans l'analyse des chroniques, on essaiera toujours de rattacher les études des chroniques à un modèle de processus stationnaire.

La covariance $K_{YY}(\tau)$ est notamment du type positif au sens (1.3). Elle vérifie par ailleurs les propriétés suivantes [72] :

$$\begin{aligned} K_{YY}(\tau) &= K_{YY}(-\tau) \text{ (fonction symétrique)} \\ K_{YY}(0) &> 0 \text{ (la variance notée } \sigma^2) \\ |K_{YY}(\tau)| &\leq K_{YY}(0) \text{ (inégalité de Schwartz)} \end{aligned} \quad (1.4)$$

Dans l'analyse des chroniques, le coefficient d'autocorrélation, désigné par $\rho_{YY}(\tau)$ est le rapport :

$$\rho_{YY}(\tau) = \frac{K_{YY}(\tau)}{K_{YY}(0)} \quad (1.5)$$

Par ailleurs, la covariance (ou corrélation) croisée entre deux processus stochastiques stationnaires du second ordre $X(t)$, $Y(t)$ est définie par :

$$K_{XY}(\tau) = E [X(t + \tau) Y(t)]$$

ou

$$\rho_{XY}(\tau) = \frac{K_{XY}(\tau)}{\sqrt{K_{XX}(0) K_{YY}(0)}} = \frac{K_{XY}(\tau)}{\sigma_X \sigma_Y} \quad (1.6)$$

et il est à noter que $\rho_{XY}(\tau) \neq \rho_{XY}(-\tau)$.

Nous rappelons ici une autre notion importante : la notion de bruit blanc $a(t)$.

Un bruit blanc est un processus stationnaire avec les caractéristiques suivantes :

$$E [a(t)] = 0 ; K_{aa}(0) = \sigma_a^2$$

$$K_{aa}(\tau) = 0 ; \tau \neq 0$$
(1.7)

Si de plus $a(t)$ suit une loi normale, à savoir...

$$a(t) \sim N(0, \sigma^2)$$

... le bruit blanc est dit Gaussien.

Dans la régression linéaire, on fait l'hypothèse que les résidus ε du modèle $Y = \alpha X + \varepsilon$ constituent un bruit blanc gaussien. Cependant, cette hypothèse est beaucoup plus restrictive dans le cas des chroniques.

II - SYSTEMES LINEAIRES ET LEURS PROPRIETES

Un système est un phénomène physique qui a été défini par Vogel comme suit [71] :

Appelons "système concret" tout assemblage d'êtres possédant une réalité physique, objets inanimés ou êtres vivants, liés de certaine façon entre eux et susceptibles d'être discernés des éléments qui ne font pas partie du "système", tout comme les éléments d'un ensemble peuvent être distingués des éléments extérieurs au moyen d'une règle d'appartenance.

Dans un système, il existe plusieurs types de variables (ou signaux) :

- les variables d'état qui, comme leur nom l'indique, caractérisent l'état du système,
- les variables d'entrée qui permettent d'influencer le système ou les variables d'état,
- les variables de sortie qui sont des fonctions des variables d'entrée et d'état du système.

II.1 - Linéarité et stationnarité

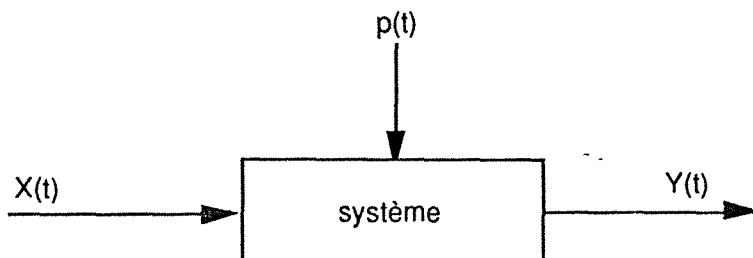


Fig. 2 - Système.

La figure 2 schématise un système dont l'entrée et la sortie sont respectivement $X(t)$, $Y(t)$. Un système peut être de type aléatoire ou déterministe. Il en est de même pour $X(t)$ et $Y(t)$. $p(t)$ est une variable d'état ou excitateur [60].

Nous n'étudions dans toute la suite que des systèmes de type déterministe, le caractère aléatoire étant donc réservé aux signaux d'entrée et de sortie.

Un système est dit linéaire s'il vérifie la propriété de superposition (additivité et homogénéité) :

$$S [a X_1(t) + b X_2(t)] = a S [X_1(t)] + b S [X_2(t)]$$

où a et b sont des coefficients et S représente la fonction du système.

Un système est dit stationnaire ou invariant dans la translation (dispositif homogène) si une translation sur la fonction d'entrée $X(t)$ se traduit par la même translation sur la fonction de sortie, à savoir :

$$\text{Si } X(t) \xrightarrow{S} Y(t) \quad \text{on a } X(t + \tau) \xrightarrow{S} Y(t + \tau)$$

Pour un système linéaire et stationnaire, la variable d'état $p(t)$ ne joue évidemment aucun rôle. Les deux termes "*système linéaire stationnaire*" et "*filtre linéaire*" sont synonymes et on les utilisera sans distinction par la suite.

Remarque 1 : La linéarité et la stationnarité ne vont pas nécessairement de pair. Il en est de même pour la non-linéarité et la non-stationnarité.

Remarque 2 : L'excitateur $p(t)$, qui modifie l'état du système, peut être le temps simplement, et on parlera d'un système non stationnaire; l'entrée elle-même et le système sera alors non linéaire, ou tout autre paramètre caractérisant l'état du système et le système sera ni linéaire, ni stationnaire [60].

II.2 - Représentation et propriétés d'un système linéaire et stationnaire

On peut souvent décrire la relation entrée-sortie d'un système à l'aide d'équations différentielles. Ceci résulte du fait que beaucoup de lois physiques s'écrivent sous forme d'équations différentielles. L'équation différentielle du type...

$$\frac{d^K Y(t)}{dt^K} + a_1 \frac{d^{K-1}(t)}{dt^{K-1}} + \dots + a_K Y(t) = X(t) \quad (1.8)$$

... possède une solution mathématique pouvant représenter des systèmes non-linéaires, des systèmes linéaires et des filtres linéaires. Elle ne définit pas en elle-même un système. Cependant, une relation de convolution peut très bien définir un système. Cette relation de convolution joue un rôle fondamental en probabilités, en géostatistique comme dans toute la physique théorique [4, 62].

II.2.1 : Relation de convolution

Pour un filtre linéaire, la relation entre l'entrée $X(t)$ et la sortie $Y(t)$ peut être exprimée par une relation de convolution définie par une intégrale du type...

$$Y(t) = \int_{-\infty}^{\infty} \phi(\tau) X(t - \tau) d\tau = \int_{-\infty}^{\infty} \phi(t - \tau) X(\tau) d\tau = \phi * X \quad (1.9)$$

... sous réserve bien entendu, d'existence de cette intégrale.

Lorsque $X(t)$, $Y(t)$ sont des processus stochastiques, la convolution définie par (1.9) est appelée "*convolution stochastique*".

Dans (1.9), la fonction $\phi(\tau)$ est baptisée réponse impulsionnelle, qui est définie comme la sortie du système tandis que l'entrée de celui-ci est une fonction de Dirac $\delta(t)$, c'est-à-dire une impulsion de masse unité qui a lieu dans un intervalle de temps extrêmement court. La fonction de Dirac $\delta(t)$ est définie telle que :

$$\int_{-\varepsilon}^{\varepsilon} \delta(t) dt = 1 \quad \varepsilon > 0 \quad (1.10)$$

La réponse impulsionnelle représente toutes les caractéristiques du système. Par ailleurs, nous pouvons définir une autre fonction, appelée "*réponse indicielle I*", par :

$$I(t) = \int_{-\infty}^t \phi(\tau) d\tau \quad (1.11)$$

... qui représente la sortie du système au moment t lorsque l'entrée de celui-ci est un échelon unité.

Dans les applications, les systèmes étudiés sont souvent causals (physiquement réalisables), c'est-à-dire que l'effet est toujours postérieur à la cause et ce qui signifie que $\phi(\tau) = 0$ pour $\tau < 0$.

Pour un système causal, la relation (1.9) peut s'écrire :

$$Y(t) = \int_0^{\infty} \phi(\tau) X(t - \tau) d\tau \quad (\text{convolution de Duhamel}) \quad (1.12)$$

C'est cette relation que nous emploierons par la suite.

Remarque : Lorsqu'un système est linéaire et non-stationnaire, la relation de convolution peut s'écrire [4, 60] :

$$Y(t) = \int_0^{\infty} \phi(\theta, \tau) X(t - \tau) d\tau$$

où θ est un paramètre qui représente l'état du système.

II.2.2 : Propriétés d'un système linéaire et stationnaire

1) Si l'entrée $X(t)$ est un processus stochastique stationnaire, la sortie $Y(t)$ l'est également.

2) Si l'entrée $X(t)$ est un bruit blanc $a(t)$, le processus $Y(t)$ (appelé processus linéaire) peut être représenté par des modèles particuliers sous forme ramassée, notés AR(p), MA(q) ou ARMA(p,q). Nous nous plaçons dans un cas discret, le modèle AR(p) s'écrit :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (1.13)$$

Le modèle MA(q) s'écrit :

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (1.14)$$

et le modèle ARMA (p,q) s'écrit :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (1.15)$$

On voit aisément que le modèle AR(p) représente un système à contre-réaction (à mémoire infinie) et il en est de même pour ARMA(p,q) (cf. fig. 3b), tandis que le modèle MA(q) représente un système sans contre-réaction (à mémoire finie ; cf. fig. 3a).

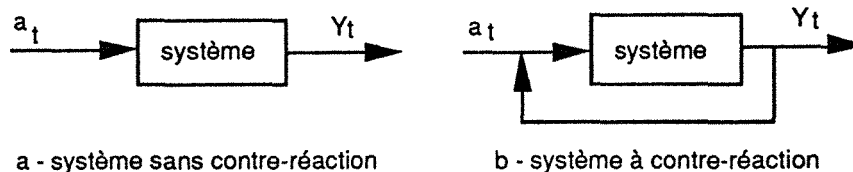


Fig. 3 - Illustration des systèmes à contre-réaction et sans contre-réaction.

3) Soit $K_{XX}(\tau)$ la covariance de $X(t)$, la covariance de $Y(t)$ désignée par $K_{YY}(\tau)$ est définie par la convolution [72,73]

$$K_{YY}(\tau) = K * \psi = \int_{-\infty}^{\infty} K_{XX}(\tau + u) \psi(u) du \quad (1.16)$$

... où $\psi(u) = \int_0^{\infty} \phi(u) \phi(u + v) dv = \phi * \hat{\phi}$ [$\phi(\tau)$ est la réponse impulsionnelle définie précédemment ; $\hat{\phi}(\tau) = \phi(-\tau)$, qui est appelée la transposée de $\phi(\tau)$].

4) Dans les applications d'identification des systèmes, la relation de convolution (1.12) doit s'écrire de façon à faire intervenir un terme résiduel [40, 60] :

$$Y(t) = \int_0^{\infty} \phi(\tau) X(t - \tau) d\tau + \varepsilon(t) \quad (1.17)$$

La théorie d'estimation linéaire de Winner propose de minimiser le terme $\varepsilon(t)$ au sens des Moindres-Carrés, c'est-à-dire :

$$\text{minimiser } \left[\varepsilon^2(t) = \left[Y(t) - \int_0^{\infty} \phi(u) X(t - u) d\tau \right]^2 \right]$$

... et ce critère aboutit à une équation appelée "équation de Winner-Hopf" [40] :

$$K_{XY}(u) = \int_{-\infty}^{\infty} \phi(\tau) K_{XX}(u - \tau) d\tau \quad u \geq 0 \quad (1.18)$$

... où $K_{XX}(u)$ est la covariance de $X(t)$ et $K_{XY}(u)$ est la covariance croisée entre $X(t)$ et $Y(t)$.

5) La convolution est une opération mathématique très liée à la transformée de Fourier que nous définissons ici [10,70].

Soit $Y(t)$ une fonction à valeurs réelles ou complexes de la variable réelle t ; on appelle transformée de Fourier (ou intégrale de Fourier) de $Y(t)$, la fonction complexe de variable réelle f :

$$Y(f) = \int_{-\infty}^{\infty} Y(t) e^{-j2\pi ft} dt \quad (-\infty < f < \infty) \quad (1.19)$$

... où f représente la fréquence.

La transformation inverse peut s'écrire :

$$Y(t) = \int_{-\infty}^{\infty} Y(f) e^{j2\pi ft} df \quad (-\infty < t < \infty) \quad (1.20)$$

Dans le domaine fréquentiel, la relation de convolution devient une relation de produit simple :

Soit $Y(W)$, $\phi(W)$ et $X(W)$ les transformées de Fourier respectives de $Y(t)$, $\phi(t)$, $X(t)$. Nous avons la relation suivante :

$$Y(W) = \phi(W) \cdot X(W) \quad (1.21)$$

... où $\phi(W)$ est appelée réponse fréquentielle ou fonction de transfert et $W = 2\pi f$.

Il en résulte que les signaux de type $e^{j2\pi ft}$ sont des fonctions propres des opérateurs filtres linéaires.

Soit $e^{j2\pi ft}$ l'entrée d'un système ; la sortie de celui-ci est $\phi(W) e^{j2\pi ft}$, $\phi(W)$ est appelé également le gain complexe du filtre.

Ainsi, la représentation de Fourier d'un signal $Y(t)$ est une décomposition de ce signal dans les bases des fonctions propres des filtres linéaires.

Remarque : On peut faire exactement le même raisonnement sur la transformée de Laplace. La transformée Z a, dans le cas discret, les mêmes propriétés que celles de Laplace dans le cas continu. La transformation Z est couramment employée en automatique notamment pour des raisons de vitesse de calcul et de précision [4, 73].

6) D'après le théorème de Winner-Khintchine, la covariance d'un processus stochastique stationnaire du second ordre $K(\tau)$ peut être décomposée en image spectrale [4,40,72] :

$$K_{YY}(\tau) = \int_{-\infty}^{\infty} \Gamma_{YY}(f) e^{j2\pi f\tau} df \quad (1.22)$$

$$\Gamma_{YY}(f) = \int_{-\infty}^{\infty} K_{YY}(\tau) e^{-j2\pi f\tau} d\tau$$

... où la fonction réelle positive $\Gamma_{YY}(f)$ est dénommée spectre énergétique.

D'une façon tout à fait analogue, le spectre énergétique croisé se définit par les transformées suivantes :

$$K_{XY}(\tau) = \int_{-\infty}^{\infty} \Gamma_{XY}(f) e^{j2\pi f\tau} df \quad (1.23)$$

$$\Gamma_{XY}(f) = \int_{-\infty}^{\infty} K_{XY}(\tau) e^{-j2\pi f\tau} d\tau$$

Il faut ajouter que la fonction $\Gamma_{XY}(f)$ est en général complexe.

III - TRAITEMENT DU SIGNAL ET ANALYSE DES CHRONIQUES

L'ensemble de techniques dans l'analyse des chroniques se réfère au traitement du signal. Dans les deux paragraphes précédents, nous avons rappelé les théories de processus stochastiques et de systèmes.

Dans ce paragraphe, nous précisons, dans un premier temps, des termes couramment utilisés dans le traitement du signal puisqu'ils seront employés dans les chapitres ultérieurs et dans un second temps, nous fixons l'objet d'étude dans l'analyse des chroniques.

III.1 - Définitions

III.1.1 Filtrage

Le filtrage est une opération qui, à un signal d'entrée X appartenant à un espace E associé à un signal de sortie Y appartenant à un sous-espace F de E (Fig. 4).

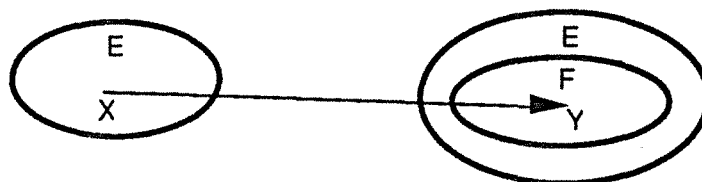


Fig. 4 - Filtrage

Il correspond à l'élimination de certaines caractéristiques du signal d'entrée X pour obtenir le signal Y de sortie. De même, le filtrage sert à l'extraction de l'information du signal bruité.

Lorsque les mesures sont échantillonnées, on parle de filtrage numérique. Le filtrage numérique est couramment utilisé dans le traitement d'image ou dans l'interprétation des diagraphies.

Un filtrage multi-composantes peut être schématisé par la figure 5. Il s'agit d'un système complexe à filtres parallèles [54].

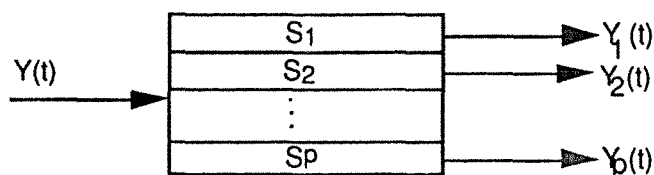


Fig 5 - Filtrage multi-composantes.

III.1.2 : Lissage

Le lissage est l'interpolation entre les instants de mesures. Ce terme est utilisé principalement lors du passage discret-analogue.

III.1.3 : Prévission (ou prédiction)

Extrapolation permettant d'estimer une grandeur physique à un instant postérieur à la dernière mesure. Il faut noter qu'en toute rigueur, la prévission ne serait possible que pour des variables représentées par des fonctions analytiques : pour de telles fonctions, en effet, la connaissance d'un petit arc de courbe détermine théoriquement la courbe entière. Mais selon toute apparence, les chroniques que l'on rencontre dans la nature ne possèdent qu'exceptionnellement une telle propriété, de sorte qu'un arc de courbe déterminé dans un intervalle donné peut coexister théoriquement avec n'importe quel autre arc de courbe dans l'intervalle voisin, ce qui signifie que l'on doit juger ou commenter les résultats de prévission avec quelques réserves.

III.1.4 : Identification

L'identification est la recherche des paramètres du modèle mathématique d'un système, signal ou bruit. Il s'agit souvent, dans la pratique, d'un problème de déconvolution qui est définie comme : soit la recherche de la fonction de transfert d'un système à partir de l'entrée et de la sortie du système ; soit le calcul de l'entrée à partir de la fonction de transfert et de la sortie du système [60, 73, 74].

III.2 - Analyse des chroniques

Dans l'analyse des chroniques existent deux aspects en fait complémentaires. Soit il s'agit de décrire l'évolution d'une chronique dans le temps, soit il s'agit de prévoir sa valeur à une date à venir [18].

L'analyse descriptive d'une chronique est une étape préliminaire à toute modélisation et prévission.

Les outils descriptifs couramment utilisés sont : le corrélogramme, le corélogramme croisé, le spectre énergétique et le spectre énergétique croisé. Ces outils seront étudiés dans le deuxième chapitre.

La prévision d'une chronique présente notamment notre principal intérêt. Elle se fait soit à partir des informations du passé sur la chronique elle-même, soit à partir d'une ou plusieurs autres chroniques, d'où on distingue deux familles de méthodes.

Les méthodes de la première famille s'appellent des méthodes univariées car elles n'incluent qu'une seule chronique. Elles peuvent convenir dans bien des cas et en particulier dans le court terme lorsqu'un phénomène continue sur son élan et ne change guère dans cet intervalle de temps. En revanche, elles peuvent s'avérer mal adaptées à moyen et à long termes. Par exemple, des changements importants de prix affecteront les ventes et probablement aussi l'impact de la publicité et de la promotion : les méthodes univariées sont alors inopérantes.

Le troisième chapitre sera consacré à la présentation des méthodes univariées : les méthodes de lissage, les méthodes de décomposition et la méthode de Box-Jenkins (dite méthode ARIMA) seront abordées.

Les méthodes de la seconde famille sont des méthodes multivariées. Elles cherchent à saisir et à mesurer l'influence d'une ou plusieurs chroniques indépendantes ou explicatives (c'est-à-dire des changements de prix et probablement aussi l'impact de la publicité et de la promotion dans l'exemple cité précédemment) sur la chronique dépendante ou expliquée (c'est-à-dire les ventes).

Dans le quatrième chapitre, nous présentons les différentes méthodes multivariées : la méthode de régression linéaire et celle de la transformée de Fourier sont les méthodes couramment utilisées ; la méthode MARIMA est une extension de la méthode ARIMA ; enfin, la méthode de déconvolution par projection sur des fonctions de base et la méthode de déconvolution géostatistique sont les méthodes récemment proposées.

En résumé, dans ce premier chapitre, nous avons rappelé brièvement les théories de processus stochastiques et de systèmes. Des outils mathématiques destinés à décrire les processus stochastiques et les systèmes ont été introduits, à savoir : la covariance (ou corrélation), la covariance croisée (ou corrélation croisée), la convolution, la transformée de Fourier, le spectre énergétique et le spectre énergétique croisé.

Par ailleurs, il a été proposé d'étudier les outils descriptifs couramment utilisés dans l'analyse des chroniques au chapitre 2, les méthodes de prévision univariées au chapitre 3 et les méthodes de prévision multivariées au chapitre 4.

Chapitre 2

S O M M A I R E

I	Coefficient d'autocorrélation et coefficient de corrélation croisée	
	(analyse corrélatrice)	17
	I.1 Coefficient d'autocorrélation - corrélogramme.....	17
	I.2 Coefficient de corrélation croisée.....	19
II	Spectre énergétique et spectre énergétique croisé.....	20
	II.1 Transformée de Fourier discrète et spectre linéaire.....	21
	II.2 Spectre énergétique et spectre énergétique croisé.....	22
III	Applications des analyses corrélatrice et spectrale	26

CHAPITRE 2

OUTILS DESCRIPTIFS UTILISES DANS L'ANALYSE DES CHRONIQUES
--

Afin de marquer l'importance de l'analyse descriptive, nous étudions dans ce chapitre entier les outils souvent utilisés. Ces outils ont été déjà introduits dans le premier chapitre. Cette étude porte notamment sur leurs propriétés et sur leurs applications.

Nous abordons dans le première paragraphe, le coefficient d'autocorrélation et le coefficient de corrélation croisée (analyse corrélatore) et dans le second paragraphe, le spectre énergétique et le spectre énergétique croisé (analyse spectrale). Enfin, pour clore ce chapitre, nous présentons dans le troisième paragraphe, un exemple d'application de ces outils dans le domaine hydrogéologique.

I - COEFFICIENT D'AUTOCORRELATION ET COEFFICIENT DE CORRELATION CROISEE (ANALYSE CORRELATOIRE)

I.1 - Coefficient d'autocorrélation - corrélogramme

Le coefficient d'autocorrélation pour un processus stochastique stationnaire $Y(t)$ est défini par (1.5) :

$$\rho_{YY}(\tau) = \frac{K_{YY}(\tau)}{K_{YY}(0)}$$

... où $K_{YY}(\tau)$ est la covariance et $K_{YY}(0)$ est la variance notée souvent par σ^2_Y .

Le graphe de la fonction $\rho_{YY}(\tau)$ est appelé corrélogramme. Il s'agit d'une fonction symétrique, c'est-à-dire que $\rho_{YY}(\tau) = \rho_{YY}(-\tau)$. C'est pour cette raison que l'on se contente toujours de la présentation pour les valeurs de τ positives, comme sur la figure 6.

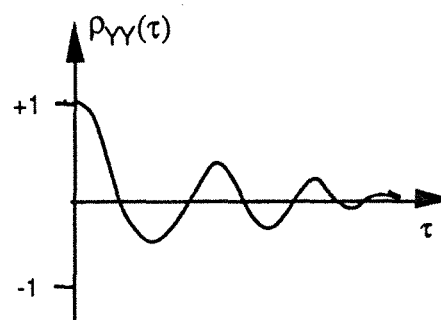


Fig. 6 - Corrélogramme.

Le corrélogramme quantifie en quelque sorte la mémoire du processus. Il traduit la dépendance entre des événements pour des intervalles de temps de plus en plus grands. Il est physiquement clair que plus τ augmente, plus la mémoire s'estompe donc que $\rho_{YY}(\tau)$ tend vers zéro.

Par ailleurs, le corrélogramme possède des propriétés qui expliquent son utilisation répandue comme un outil descriptif efficace :

$$1) \rho_{YY}(0) \equiv 1 \quad \text{et} \quad |\rho_{YY}(\tau)| \leq 1$$

2) Il a même période que le processus lui-même, par exemple pour un processus périodique de la forme :

$$Y(t) = m + A \cos \omega t + B \sin \omega t$$

... où A et B sont deux variables aléatoires indépendantes d'espérance nulle et de même variance σ^2 . Il s'agit-là d'un modèle où la composante résiduelle intervient de façon multiplicative. Son corrélogramme est [18, 19] :

$$\rho_{YY}(\tau) = \cos \omega \tau$$

Il est évident que $\rho_{YY}(\tau)$ a même période que $Y(t)$ et $|\rho_{YY}(\tau)| \leq 1$.

3) Lorsque le corrélogramme tend rapidement vers zéro, le processus représenté est stationnaire. A l'inverse, un corrélogramme qui décroît lentement indique un phénomène très structuré et donc l'existence de non-stationnarité.

4) Théoriquement, $\rho_{YY}(0) = 1$ et $\rho_{YY}(\tau) = 0$ ($\tau \neq 0$) pour un bruit blanc $a(t)$ puisque ceci caractérise une succession d'événements indépendants les uns des autres.

L'estimateur d'autocorrélation $\rho'_{YY}(\tau)$ défini dans l'intervalle $t \in \left[-\frac{T}{2}, \frac{T}{2}\right]$ s'écrit [40]

$$\rho'_{YY}(\tau) = \frac{K'_{YY}(\tau)}{K'_{YY}(0)} \quad (2.1)$$

où

$$K'_{YY}(\tau) = \begin{cases} \frac{1}{T} \int_0^{T-|\tau|} [Y(t) Y(t+|\tau|)] dt & 0 \leq |\tau| \leq T \\ 0 & |\tau| > T \end{cases} \quad (2.2)$$

$K'_{YY}(\tau)$ est l'estimateur de covariance dont l'espérance mathématique s'écrit :

$$E [K'_{YY}(\tau)] = \begin{cases} K_{YY}(\tau) \left[1 - \frac{|\tau|}{T} \right] & |\tau| \leq T \\ 0 & |\tau| > T \end{cases} \quad (2.3)$$

On constate qu'il n'est jamais possible d'obtenir la vraie image de covariance $K_{YY}(\tau)$. On ne peut que l'observer par une fenêtre $\left[1 - \frac{|\tau|}{T} \right]$. Cette constatation est très importante pour comprendre le lissage du spectre abordé dans le paragraphe suivant.

Nous revenons au cas concret d'une chronique Y_t (Y_1, Y_2, \dots, Y_n) ; le corrélogramme empirique se calcule par plusieurs formules, parmi lesquelles nous citons les formules préconisées par JENKINS et WATTS [40] :

$$\rho'_{YY}(\tau) = \frac{K'_{YY}(\tau)}{K'_{YY}(0)} \quad (2.4)$$

$$\dots \text{ où } K'_{YY}(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} (Y_t - \bar{Y}_t) (Y_{t+\tau} - \bar{Y}_t) \quad (2.5)$$

... dans lesquelles : $\tau = 0, 2, \dots, M$; M est appelé point de troncature. Le choix de M n'est pas fondé sur des concepts théoriques ; dans la littérature, il est souvent indiqué de prendre M égal à $\frac{n}{3}$ (où n est la longueur de la chronique), sans que les auteurs en fournissent la raison[56].

I.2 - Coefficient de corrélation croisée

Soit $X(t)$, $Y(t)$ deux processus stochastiques stationnaires. Le coefficient de corrélation croisée $\rho_{XY}(\tau)$ est défini par :

$$\rho_{XY}(\tau) = \frac{E [X(t + \tau) Y(t)]}{\sigma_X \sigma_Y} \quad (2.6)$$

ou

$$\rho_{YX}(\tau) = \frac{E [X(t) Y(t + \tau)]}{\sigma_X \sigma_Y} \quad (2.7)$$

Les propriétés de $\rho_{XY}(\tau)$ sont les suivantes :

1. La fonction $\rho_{XY}(\tau)$ est une fonction asymétrique, c'est-à-dire $\rho_{XY}(\tau) \neq \rho_{XY}(-\tau)$. Cependant, la relation $\rho_{XY}(\tau) = \rho_{YX}(-\tau)$ est vérifiée[40].

2. $\rho_{XY}(\tau) \leq 1$ et très souvent $\rho_{XY}(\tau) > 0$.

3. Si $\rho_{XY}(\tau) \equiv 0$ pour toute valeur τ , $X(t)$ et $Y(t)$ ne sont pas corrélés. Si de plus, $X(t)$ et $Y(t)$ suivent tous les deux une loi normale, on dit qu'ils sont indépendants.

4. Si $\rho_{XY}(0) \neq 0$ et $\rho_{XY}(\tau) = 0$ ($\tau \neq 0$), les processus $X(t)$ et $Y(t)$ sont corrélés simultanément. $\rho_{XY}(\tau)$ devient en effet un coefficient de corrélation simple défini entre deux variables.

5. Soit $X(t)$ l'entrée d'un système et $Y(t)$ la sortie, lorsque $X(t)$ est un processus purement aléatoire (bruit blanc), le coefficient de corrélation croisée $\rho_{XY}(\tau)$ correspond à la réponse impulsionnelle $\phi(\tau)$ du système. Cette propriété est particulièrement intéressante dans les applications car la recherche de la réponse impulsionnelle constitue souvent un objet principal[40,57].

II - SPECTRE ENERGETIQUE ET SPECTRE ENERGETIQUE CROISE ANALYSE SPECTRALE

La transformation de Fourier est un bel exemple d'outil mathématique créé pour les besoins de la physique. C'est en effet pour résoudre l'équation de la chaleur que Jean-Joseph-Baptiste FOURIER a découvert en 1822 les propriétés des séries qui porte maintenant son nom.

L'analyse spectrale, qui se fonde sur la transformation de Fourier, travaille dans le domaine fréquentiel.

Théoriquement, le spectre est un outil équivalent au corrélogramme. Il est parfois préférable, dans la pratique, d'utiliser le spectre et cela a pour deux raisons au moins : la première est que le spectre lui-même possède souvent une signification physique pour le phénomène étudié, et l'interprétation est donc plus facile et intuitive ; la seconde est que le spectre affine la connaissance sur le phénomène étudié par rapport à ce qu'apporte le corrélogramme.

II.1 - Transformée de Fourier discrète et spectre linéaire

Dans le premier chapitre, nous avons défini la transformée de Fourier par (1.19)...

$$Y(f) = \int_{-\infty}^{\infty} Y(t) e^{-j2\pi ft} dt \quad (-\infty < f < \infty)$$

... et la transformation inverse par (1.20) :

$$Y(t) = \int_{-\infty}^{\infty} Y(f) e^{j2\pi ft} df \quad (-\infty < t < \infty)$$

De même que la transformée de Fourier fait correspondre une fonction à une fonction, la transformée de Fourier discrète fait correspondre une suite finie à une suite finie.

Un signal périodique est caractérisé par l'information contenue dans une période.

On décompose le signal échantillonné de durée finie $Y(t) \in \left[-\frac{T}{2} \leq t \leq \frac{T}{2} \right]$ sur les racines Nième de l'unité. D'où la nécessité d'avoir :

$$N\Delta = T$$

Si Δ est la période d'échantillonnage et T la période du signal (*cf.* Fig. 7, nous supposons $N = 2n$).

La transformée de Fourier discrète se définit par [70,74,75] :

$$Y_m = \frac{1}{N} \sum_{t=-n}^{n-1} Y_t e^{-j(2\pi mt/N)}$$

$$(-n \leq m \leq n-1) \quad (2.8)$$

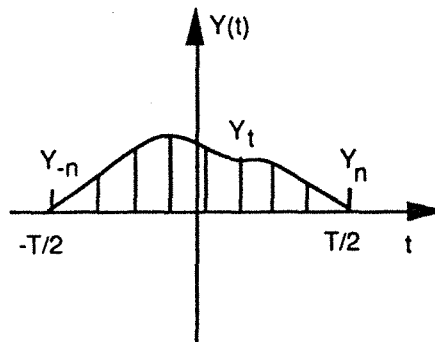


Fig. 7 - Signal échantillonné.

... et la transformation inverse par :

$$Y_t = \sum_{m=-n}^{n-1} Y_m e^{j(2\pi mt/N)}$$

$$(-n \leq t \leq n-1) \quad (2.9)$$

Le théorème de Parseval, donnant une relation entre l'énergie du signal et celle de sa transformée, s'écrit :

$$\frac{1}{N} \sum_{t=-n}^n Y_t^2 = \sum_{m=-n}^{n-1} |Y_m|^2 \quad (2.10)$$

Le graphe $|Y_m|^2 - m$ est appelé spectre linéaire ou périodogramme. Il représente évidemment la distribution de l'énergie du signal sur les fréquences harmoniques $\frac{m}{T}$ de la fréquence fondamentale $\frac{1}{T}$.

II.2 - Spectre énergétique et spectre énergétique croisé

Nous venons de voir que le spectre linéaire correspond à la transformée de Fourier de $Y(t)$. Nous cherchons maintenant une autre approche.

Tandis que Δ (période d'échantillonnage) tend vers zéro, le théorème de Parseval devient...

$$\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} Y^2(t) dt = \sum_{m=-\infty}^{\infty} |Y_m|^2 \quad (2.11)$$

... et la variance du signal $Y(t)$ ($-\infty < t < \infty$) est :

$$\begin{aligned}\sigma^2 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} Y^2(t) dt = \lim_{T \rightarrow \infty} \sum_{m=-\infty}^{\infty} |Y_m|^2 \\ &= \lim_{T \rightarrow \infty} \sum_{m=-\infty}^{\infty} T |Y_m|^2 \frac{1}{T} = \int_{-\infty}^{\infty} \Gamma_{YY}(f) df\end{aligned}\quad (2.12)$$

$\Gamma_{YY}(f) = \lim_{T \rightarrow \infty} T |Y_m|^2$ est dénommé spectre énergétique. Il est clair que le spectre énergétique possède la même signification que le spectre linéaire. Pourtant, le spectre énergétique est un outil plus efficace.

Par ailleurs, l'estimateur du spectre énergétique s'exprime par...

$$\Gamma'_{YY}(f) = T |Y_m|^2 = \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} Y(t) e^{j2\pi ft} dt \right]^2 \quad (2.13)$$

... et on voit aisément que :

$$\lim_{T \rightarrow \infty} \Gamma'_{YY}(f) = \Gamma_{YY}(f) \quad (2.14)$$

D'après le théorème de Winner-Khintchine, le spectre énergétique est la transformée de Fourier de la covariance $K_{YY}(t)$ [4, 54, 72]...

$$\begin{aligned}\Gamma_{YY}(f) &= \int_{-\infty}^{\infty} K_{YY}(\tau) e^{-j2\pi f\tau} d\tau \\ K_{YY}(\tau) &= \int_{-\infty}^{\infty} \Gamma_{YY}(f) e^{j2\pi f\tau} df\end{aligned}$$

... et le spectre énergétique croisé est la transformée de Fourier de la covariance croisée :

$$\Gamma_{XY}(f) = \int_{-\infty}^{\infty} K_{XY}(\tau) e^{-j2\pi f\tau} d\tau$$

$$\Gamma_{K_{XY}}(\tau) = \int_{-\infty}^{\infty} \Gamma_{XY}(f) e^{j2\pi f\tau} df$$

Dans ce paragraphe, la discussion porte essentiellement jusqu'à présent, sur un signal déterministe dont le spectre est univoquement déterminé (par 2.13). Pour une chronique (processus stochastique), néanmoins, le spectre $\Gamma_{YY}(f)$ n'est qu'une réalisation du processus stochastique $\Gamma_{YY}(f)$. Il en résulte que la relation (2.14) n'est plus vérifiée. Pourtant, on a une relation du type[40]:

$$\lim_{T \rightarrow \infty} E[\Gamma_{YY}(f)] = \Gamma_{YY}(f) = \int_{-\infty}^{\infty} \Gamma_{K_{YY}}(\tau) e^{-j2\pi f\tau} d\tau \quad (2.15)$$

et

$$\begin{aligned} E[\Gamma_{YY}(f)] &= \int_{-T}^T E[K_{YY}(f)] e^{-j2\pi f\tau} d\tau \\ &= \int_{-T}^T K_{YY}(\tau) \left[1 - \frac{|\tau|}{T} \right] e^{-j2\pi f\tau} d\tau \end{aligned} \quad (2.16)$$

L'expression (2.16) montre qu'il est possible, pour un processus stochastique, d'estimer la répartition moyenne de la variance suivant les fréquences à l'aide d'une fenêtre $\left[1 - \frac{|\tau|}{T} \right]$. L'effet de fenêtre est de lisser le spectre pour qu'il soit lisible et facilement interprétable. Ainsi, des fenêtres notées par la suite $W(u)$ ont été conçues. L'estimateur du spectre peut donc s'écrire de façon générale :

$$\overline{\Gamma_{YY}(f)} = \int_{-\infty}^{\infty} K_{YY}(\tau) W(\tau) e^{-j2\pi f\tau} d\tau \quad (2.17)$$

Les fenêtres souvent utilisées sont celles de Bartlett, de Tucky et de Parzen.

La fenêtre de Bartlett, par exemple, s'écrit par...

$$W(\tau) = \begin{cases} 1 - \frac{|\tau|}{M} & |\tau| \leq M \\ 0 & |\tau| > M \end{cases} \quad (2.18)$$

... où M est le point de troncature.

La largeur de bande d'une fenêtre est définie comme...

$$b = \frac{1}{\int_{-\infty}^{\infty} W^2(u) du}$$

... qui est étroitement liée à la valeur M .

Dans les applications, il s'agit de choisir une fenêtre de forme appropriée et aussi sa largeur de bande.

Nous revenons au cas d'une chronique Y_t $\{Y_1, Y_2, \dots, Y_N\}$. En supposant que $\Delta = 1$, nous avons l'estimateur du spectre [40] :

$$\overline{YY'}(f) = \sum_{\tau=-(M-1)}^M W(\tau) K_{YY}(\tau) e^{-j2\pi f\tau} \left[-\frac{1}{2} \leq f \leq \frac{1}{2} \right] \quad (2.19)$$

$\overline{YY'}(f)$ est une fonction réelle et symétrique. On se contente de calculer ses valeurs pour des fréquences positives. On aboutit à une formule du type :

$$\overline{YY'}(f) = 2 \left\{ K_{YY}(0) + 2 \sum_{\tau=1}^M W(\tau) K_{YY}(\tau) \cos 2\pi f\tau \right\} \quad 0 \leq f \leq \frac{1}{2} \quad (2.20)$$

Dans la formule ci-dessus, L est une variable continue. Mais le spectre se calcule pratiquement à des fréquences discrétisées. Nous admettons une formulation suivante...

$$\overline{YY'}(L) = 2 \left\{ K_{YY}(0) + 2 \sum_{\tau=1}^M K_{YY}(\tau) W(\tau) \cos \frac{2\pi L\tau}{F} \right\} \quad (2.21)$$

... où $L = 0, 1, 2, \dots, F$ et $F = 2 \sim 3 M$.

Par ailleurs, le spectre croisé entre deux chroniques X_t $\{X_1, X_2, \dots, X_N\}$ et Y_t $\{Y_1, Y_2, \dots, Y_N\}$ peut être calculé par...

$$\overline{\Gamma_{XY}(L)} = \sum_{\tau=-(M-1)}^M \left\{ W(\tau) K_{XY}(\tau) e^{-\frac{j2\pi L\tau}{F}} \right\} \left(-\frac{1}{2} \leq f \leq \frac{1}{2} \right) \quad (2.22)$$

... où $L = 1, 2, \dots, F$ et $F = 2 \sim 3 M$.

Le spectre croisé est exprimé à l'aide de deux fonctions : la fonction d'amplitude et la fonction de phase en raison de la non-symétrie de la corrélation entrée-sortie et sortie-entrée ; la fonction d'amplitude fournit une décomposition de la covariance totale entre entrée et sortie pour différentes fréquences et la fonction de phase donne le moyen d'estimer, pour chacune de ces fréquences, le déphasage entre entrée et sortie.

III - APPLICATIONS DES ANALYSES CORRELATOIRE ET SPECTRALE

En hydrogéologie, les travaux de MANGIN [56, 57, 58] ont montré que les outils décrits ci-dessus présentent un excellent moyen d'investigation des systèmes hydrauliques.

Les exemples traités dans ces travaux concernent trois systèmes karstiques pyrénéens : Aliou, le Baget et Fontestorbes. Le premier est très karstifié et sans réserve ; le second, moins karstifié, possède des réserves plus importantes ; le troisième est caractérisé par des réserves considérables. Les données de base sont les totaux journaliers de précipitation et les valeurs moyennes journalières de débit.

Le corrélogramme et le spectre de la pluie à la station de Balagué (fig. 8) montrent qu'elle est effectivement aléatoire.

Les corrélogrammes et les spectres pour ces trois systèmes sont illustrés sur la figure 9.

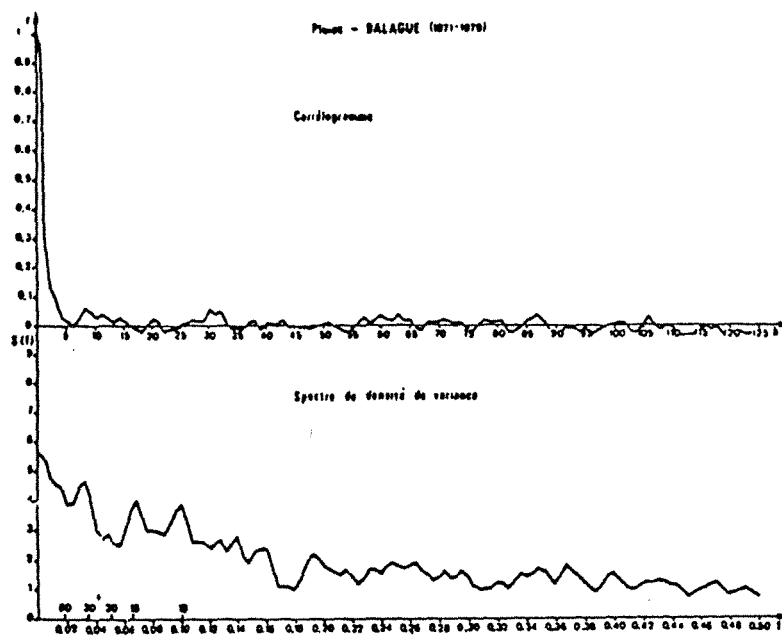


Fig. 8 - Analyse de la chronique des pluies à la station de Balagué
($n = 9$ ans ; $m = 125$ jours). (Extrait de [56]).

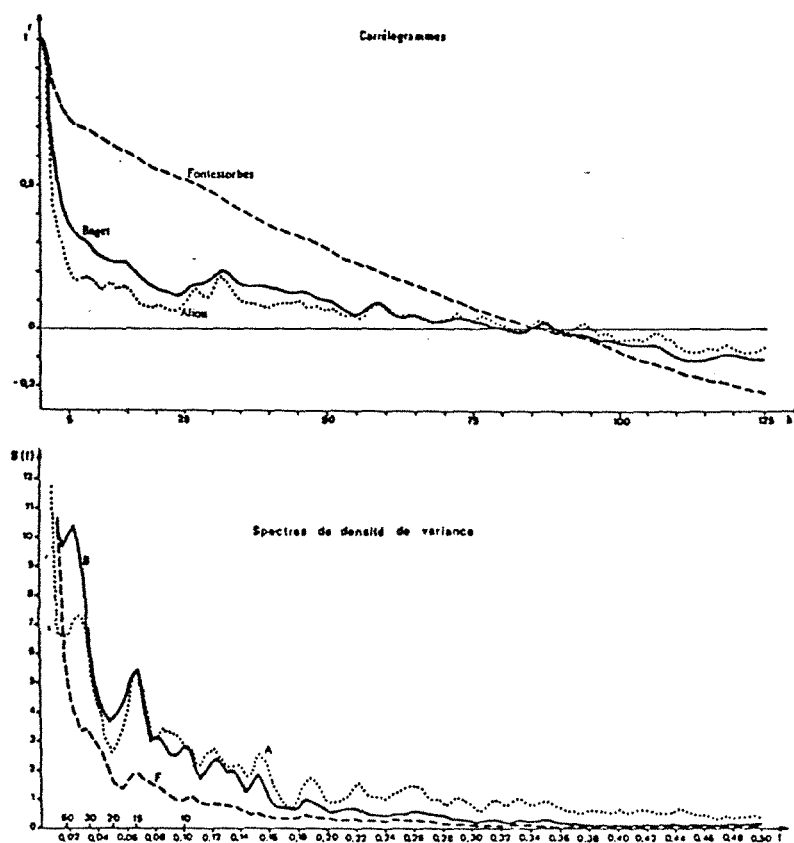


Fig. 9 - Analyse des chroniques des débits de trois systèmes karstiques : Fontestorbes, Baget et Aliou
($n = 9$ ans ; $m = 125$ jours). (Extrait de [56]).

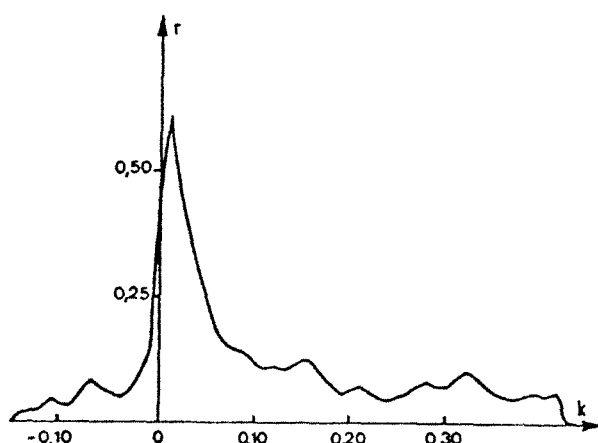
Les vitesses de décroissance des corrélogrammes pour Aliou, le Baget et Fontestorbes sont différentes. L'effet mémoire est donc différent d'un système à l'autre : il est très élevé pour Fontestorbes, très faible pour Aliou. Les spectres traduisent le même phénomène et montrent que cet effet mémoire correspond à un filtrage des hautes fréquences au profit des basses fréquences.

Cet effet mémoire est donc directement lié à l'importance des réserves, faibles pour Aliou et considérables pour Fontestorbes.

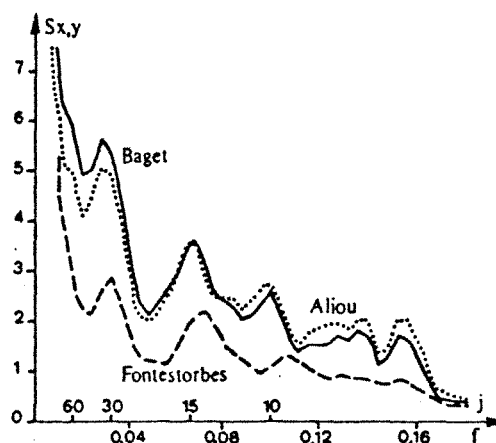
L'analyse de détail des corrélogrammes et surtout celle des spectres montrent une série de variations sur les courbes présentées aux mêmes endroits pour les trois systèmes et dont seule l'amplitude est différente. Ces variations correspondent aux pics dans le spectre de la pluie. Ceci indique que ces trois systèmes conservent, tout au moins en fréquence, l'information contenue dans les épisodes pluvieux.

En plus des analyses simples, les analyses corrélatoire et spectrale réalisées de façon croisée apportent des renseignements plus précieux sur les systèmes étudiés.

La figure 10a illustre le corrélogramme croisé pour le Baget. Il représente la réponse impulsionnelle du système car l'entrée du système (la pluie) est pratiquement assimilable à un processus purement aléatoire. La figure 10b représente les fonctions d'amplitudes correspondant aux trois systèmes. On se reportera à [57] pour plus de détails.



a - Corrélogramme croisé obtenu sur le système du Baget à partir d'une chronique s'étendant de 1968 à 1979.



b - Fonctions d'amplitudes aux trois systèmes du Baget, d'Aliou et de Fontestorbes à partir d'une chronique s'étendant de 1971 à 1979.

Fig. 10 - Corrélogramme croisé et spectre croisé (Extrait de [57]).

On verra dans le troisième chapitre, que le corrélogramme constitue un outil de base qui peut servir à identifier l'ordre q pour un processus moyen mobile $MA(q)$.

Dans le quatrième chapitre, il sera montré que le corrélogramme croisé constitue un outil de base qui peut servir à identifier un modèle $MARIMA$.

Chapitre 3

S O M M A I R E

I	Méthodes de Lissage (modèles autoprojectifs)	30
I.1	Méthodes de lissage par moyenne mobile	31
I.2	Méthodes de lissage exponentiel.....	32
II	Méthodes de décomposition (modèles d'ajustement)	35
III	Méthode Univariée de Box-Jenkins(ARIMA)	37
III.1	Modèles ARMA, ARIMA et SARIMA.....	38
III.1.1	: Généralité.....	38
III.1.2	: Modèles ARMA(p,q)	39
1)	Modèle autorégressif d'ordre p : AR(p)	39
2)	Modèle "moyenne mobile" MA(q)	41
3)	Modèles ARMA(p,q)	42
III.1.3	: Modèles ARIMA(p,d,q) et SARIMA	44
III.2	Méthodologie de Box-Jenkins	45
III.2.1	: Identification	45
III.2.2	: Estimation	45
III.2.3	: Validation (Diagnostic)	46
III.2.4	: Prévission	46
IV	Conclusion	47

CHAPITRE 3

METHODES DE PREVISION UNIVARIEES

Dans les applications d'analyse des chroniques, nous faisons appel aux modèles de prévision univariées dans les deux cas suivants...

- soit pour un phénomène étudié, une seule chronique est en notre possession,
- soit plusieurs chroniques sont enregistrées, mais notre intérêt est de faire une prévision pour une chronique quelconque en supposant qu'elle est indépendante des autres.

Supposons qu'on a les valeurs observées d'une chronique Y_t du temps $t=1$ au temps $t=T$

$$Y_t = \{Y_1, Y_2, \dots, Y_T\}$$

et le problème qui se pose ici est de prévoir la valeur de Y_t au temps $T+h$ à partir des informations sur les valeurs observées du passé. On dénote, dans la suite, la prévision par $\hat{Y}_{T(h)}$, T étant l'origine de prévision et h l'horizon de prévision.

Dans ce chapitre, trois familles de méthodes sont présentées. Les méthodes de lissage et celles de décomposition sont relativement simples et s'adaptent généralement aux chroniques dont les lois de base soient immédiatement apparentes. Pourtant, dans des situations plus complexes, la méthode de Box-Jenkins paraît beaucoup plus efficace. En effet, cette dernière méthode est très différente des méthodes de lissage et de décomposition: elle repose sur la notion du processus stationnaire et associe des processus stationnaires particuliers notés AR et d'autres notés MA et donc représente une procédure moins intuitive.

I - METHODES DE LISSAGE (MODELES AUTOPROJECTIFS)

[18, 19, 24, 81, 82]

Le lissage est une opération qui consiste à remplacer la série Y_t par une série de transformée Y_t^* :

$$Y_t^* = \sum_{j=-p1}^{p2} w_j Y_{t+j} \text{ vérifiant } \sum_{j=-p1}^{p2} w_j = 1 \quad (3.1)$$

où $p1$ et $p2$ sont des entiers non négatifs et $w_{-p1}, \dots, w_0, \dots, w_{p2}$ un ensemble de coefficients de pondération indépendants de t . Le cas particulier le plus simple d'une telle transformation est celui de la moyenne mobile symétrique d'ordre $k=2p+1$ telle que:

$$Y_t^* = \frac{1}{2p+1} \sum_{j=-p}^p Y_{t+j} \quad (3.2)$$

Le recours à une moyenne mobile symétrique dans un but de prévision n'est pas justifiée. Les moyennes mobiles non symétriques sont souvent utilisées.

I.1 - Méthodes de lissage par moyenne mobile

Si la chronique est stationnaire, la moyenne mobile simple donne la prévision...

$$\hat{Y}_T(h) = \frac{1}{N} \sum_{i=T-N+1}^T Y_i \quad (3.3)$$

... où N étant le nombre des valeurs de Y_t les plus récentes que l'on prend pour la prévision, est appelé l'ordre de la moyenne mobile. Il est évident que $\hat{Y}_T(h)$ ne dépend pas de h dans ce cas.

Dans (3.3), si l'on suppose que $N = 1$:

$\hat{Y}_T(h) = Y_T$, la prévision est la valeur la plus récente de Y_t et si $N=T$:

$\hat{Y}_T(h) = \frac{1}{T} \sum_{i=1}^T Y_i$, la prévision est la moyenne totale des valeurs $Y_1,$

Y_2, \dots, Y_T .

Si la chronique contient une tendance, il faudra faire appel à des techniques plus complexes : moyennes mobiles doubles et moyennes mobiles d'ordre supérieur.

Les moyennes mobiles doubles, désignées par MA ($M \times N$) sont les moyennes mobiles sur M points des moyennes mobiles sur N points des données originales.

Comme pour les moyennes mobiles doubles, on calcule deux fois les moyennes mobiles ; pour les moyennes mobiles d'ordre supérieur, on calcule plusieurs fois ; par exemple : trois fois pour MA ($P \times M \times N$).

I.2 - Méthodes de lissage exponentiel

L'idée est simple : on suppose que les observations influent d'autant moins la prévision qu'elles sont plus éloignées de la date T à laquelle on fait la prévision et que cette influence décroît exponentiellement. Lorsque la chronique est stationnaire, la prévision par lissage exponentiel simple s'écrit :

$$\hat{Y}_T(h) = (1 - \beta) \sum_{j=0}^{T-1} \beta^j Y_{T-j}, \quad 0 < \beta < 1 \quad (3.4)$$

On notera que la somme des coefficients de pondération...

$$(1 - \beta) (1 + \beta + \dots + \beta^{T-1}) = 1 - \beta^T$$

... est très voisine de 1 dès que T est assez grand.

L'expression (3.4) peut s'écrire...

$$\hat{Y}_T = (1 - \beta) \sum_{j=0}^{T-1} \beta^j Y_{T-j} \quad (0 < \beta < 1) \quad (3.5)$$

... car $\hat{Y}_T(h)$ est indépendant de h .

Généralement, la valeur de β est choisie indépendamment de h . On voit facilement que plus β est proche de 1, plus l'influence des observations passées remontera loin dans le temps et plus la prévision sera rigide. Au contraire, plus β est proche de 0, plus la prévision sera souple.

Il est possible de donner une autre interprétation à la formule de prévision. En effet, (3.5) peut s'écrire immédiatement :

$$\hat{Y}_T = \beta \hat{Y}_{T-1} + (1 - \beta) Y_T \quad (3.6)$$

...où \hat{Y}_T apparaît donc comme une moyenne pondérée entre la prévision faite au temps $T - 1$ et la dernière observation Y_T .

On peut encore écrire...

$$\hat{Y}_T = \hat{Y}_{T-1} + (1 - \beta) (Y_T - \hat{Y}_{T-1}) \quad (3.7)$$

... où \hat{Y}_T s'interprète comme la prévision au temps précédent \hat{Y}_{T-1} corrigée d'un terme proportionnel à la dernière erreur de prévision. Cette écriture fournit une méthode très simple de mise à jour de la prévision. Cette simplicité a été longtemps un argument décisif en faveur du lissage exponentiel simple et même avec les facilités de calculs apportées par l'informatique, c'est toujours un argument important lorsque le nombre de séries à traiter est grand.

Si de plus, la chronique contient une tendance supposée linéaire, le lissage exponentiel linéaire (Brown) conduit à une prévision dépendant de l'horizon h (fig.11)

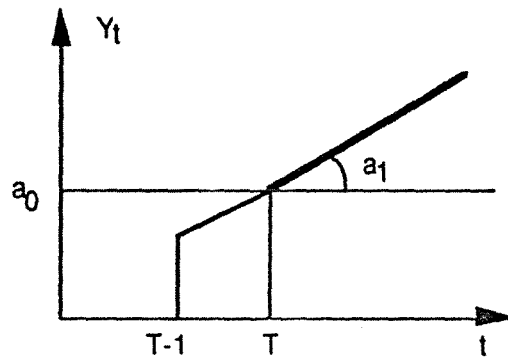


Fig.11 - Lissage exponentiel linéaire.

$$\hat{Y}_T(h) = \hat{a}_0(T) + h \hat{a}_1(T) \quad (3.8)$$

où

$$\hat{a}_0(T) = \hat{a}_0(T-1) + \hat{a}_1(T-1) + (1 - \beta^2) [Y_T - \hat{Y}_{T-1}(1)]$$

$$\hat{a}_1(T) = \hat{a}_1(T-1) + (1 - \beta^2) [Y_T - \hat{Y}_{T-1}(1)]$$

On voit que le poids de la dernière observation Y_T est égale à $(1 - \beta^2)$.

Dans le lissage exponentiel linéaire, le poids de la dernière observation est le même pour la prévision du niveau $a_0(T)$ et celle de la tendance $a_1(T)$.

Dans le lissage exponentiel double (Holt), les pondérations peuvent être différentes :

$$\begin{aligned}\hat{a}_0(T) &= \alpha [\hat{a}_0(T-1) + \hat{a}_1(T-1)] + (1 - \alpha) Y_T \\ \hat{a}_1(T) &= \beta \hat{a}_1(T-1) + (1 - \beta) [\hat{a}_0(T) - \hat{a}_0(T-1)]\end{aligned}$$

La méthode est plus flexible dans la mesure où elle fait intervenir deux constantes α et β . Mais en contrepartie, il y a deux constantes à choisir.

Si encore intervient un terme saisonnier, le lissage exponentiel triple (Winter) s'écrit :

$$\hat{Y}_T(h) = \hat{a}_0(T) + h \hat{a}_1(T) + \hat{S}(T+h-\tau) \quad (3.9)$$

... où :

$$\hat{a}_0(T) = \alpha [\hat{a}_0(T-1) + \hat{a}_1(T-1)] + (1 - \alpha) (Y_T - \hat{S}_{T-\tau})$$

$$\hat{a}_1(T) = \beta \hat{a}_1(T-1) + (1 - \beta) [\hat{a}_0(T) - \hat{a}_0(T-1)]$$

$$\hat{S}(T) = \gamma \hat{S}(T-\tau) + (1 - \gamma) [(Y_T - \hat{a}_0(T))]$$

En résumé, nous pouvons faire les remarques suivantes pour les méthodes de lissage :

* Dans le principe, le lissage exponentiel fonctionne d'une manière analogue aux moyennes mobiles et évite deux limitations majeures au moins dans l'emploi des moyennes mobiles : la première est que pour calculer une prévision par la moyenne mobile, il est nécessaire de stocker les N dernières valeurs observées. Ceci nécessite un espace considérable qui, dans bien des systèmes d'ordinateurs, est coûteux ; la seconde est que la méthode des moyennes mobiles donne un poids égal à chacune des N dernières observations et aucun poids aux observations relevées avant la période $T - N$.

* Les moyennes mobiles constituent un outil à la fois de prévision et de décomposition.

* D'une manière tout à fait analogue à celle que nous avons exposée pour les moyennes mobiles doubles et le lissage exponentiel double, on peut également développer et utiliser des formes de lissage d'ordre supérieur pour prévoir des lois de nature quadratique ou plus complexe. Les calculs nécessaires à ce type de prévision deviennent cependant très lourds et le fait qu'on ait besoin de connaître à l'avance la loi réelle contenue dans les données rend difficile l'usage de ces formes d'ordre supérieur. Aussi, bien qu'elle soit possible, l'utilisation pratique de ces méthodes est problématique. Pour cette raison et parce que l'on dispose de plusieurs autres méthodes de prévision capables de s'accommoder directement de ces lois complexes, on

emploie donc rarement dans la pratique les techniques de lissage d'un ordre supérieur à celui du lissage exponentiel double.

* Le lissage par moyennes mobiles symétriques d'ordre τ permet de filtrer la composante saisonnière de périodicité τ . C'est sur cette propriété que repose en grande partie l'intérêt pratique de cette méthode.

II - METHODES DE DECOMPOSITION (MODELES D'AJUSTEMENT)

[24, 81, 82]

Elles consistent à décomposer la chronique en quatre composantes : tendance f_t , cycle C_t , saisonnalité S_t et terme résiduel ε_t . Ces quatre composantes se combinent :

- soit de façon additive : $Y_t = f_t + C_t + S_t + \varepsilon_t$
- soit de façon multiplicative : $Y_t = f_t \cdot C_t \cdot S_t \cdot \varepsilon_t$
- soit de façon mixte, par exemple : $Y_t = (f_t + C_t) \cdot S_t + \varepsilon_t$

La composante tendancielle f_t décrit la tendance "lourde" du phénomène sur une longue période. Pour la décrire, on utilise le plus souvent une fonction polynomiale en t :

$$f_t = at + b \text{ ou } f_t = at^2 + bt + c \dots$$

Mais, on peut utiliser également les fonctions exponentielles (fig.12)

$$f_t = f_0 e^{at} \quad (a \neq 0)$$

ou la loi logistique (fig.13) :

$$f_t = \frac{a}{b + \rho^c t}$$

pour des évolutions présentant un phénomène de saturation.

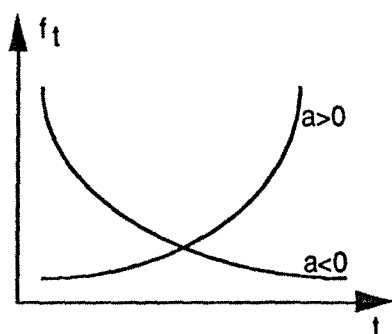


Fig.12 - Lois exponentielles.

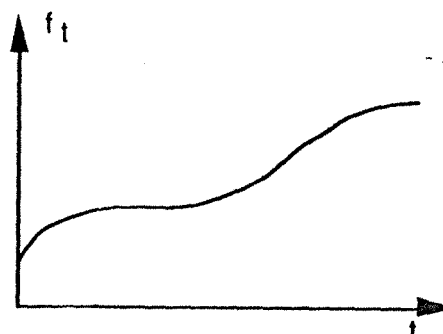


Fig. 13 - Loi logistique.

La composante saisonnière S_t est caractérisée avant tout par sa périodicité : la valeur observée au temps t est liée à celle observée au temps $(t - \tau)$, τ étant la périodicité.

La différence entre la composante cyclique et la composante saisonnière est que cette dernière représente la fluctuation saisonnière, comme par exemple, fluctuation mensuelle des pluies, tandis que la première représente la fluctuation de longue durée d'un cycle à un autre.

Le terme résiduel ϵ_t (désigné sous le nom de bruitage) est, par nature, dépourvu de toute tendance et de toute périodicité. Il est nécessaire de faire sur lui quelques hypothèses dont la plus simple est que les ϵ_t constituent un ensemble de variables aléatoires indépendantes entre elles, de moyenne nulle et de même variance : c'est le bruit blanc. Cette hypothèse, faite dans la régression linéaire, est beaucoup trop restrictive pour être généralement admissible dans le cas des chroniques. Il faudra faire appel à un modèle plus général : celui des processus aléatoires stationnaires.

La décomposition se fait en général suivant six étapes :

a) Calculer les moyennes mobiles dont l'ordre est égal à la périodicité τ , afin d'éliminer les composantes saisonnière et résiduelle dans les données originales (les moyennes mobiles se composent des composantes tendancielle f_t et cyclique C_t)

b) Séparer les moyennes mobiles des données originales (extraction ou division, selon la loi de combinaison) afin de retenir les composantes saisonnière et résiduelle (S_t, ϵ_t).

c) Isoler les facteurs saisonniers, à partir de b), en prenant la valeur moyenne de chaque facteur. Pour des données mensuelles, douze facteurs existent ; on calcule par exemple la moyenne de ($S_1, S_{13}, S_{25}, S_{37} \dots$), ce qui donne le premier facteur. Dans cette étape, on obtient la composante saisonnière (S_t).

d) Identifier la forme appropriée de la tendance (polynômes, fonctions exponentielles, etc.) et calculer ses valeurs (f_t).

e) Séparer les valeurs f_t des moyennes mobiles pour tirer la composante cyclique (C_t).

f) Séparer S_t, f_t, C_t calculé par c), d), e) respectivement, des données originales Y_t , obtenir enfin le terme résiduel (ϵ_t).

Il faut savoir que cette décomposition est un peu empirique. Dans bien des cas, la composante cyclique n'existe pas ou est ignorée et la procédure de décomposition sera simplifiée.

Enfin, il convient de souligner que la décomposition fait appel aux moyennes mobiles symétriques d'ordre τ .

Si τ est un nombre impair, on calcule immédiatement...

$$Y_t^* = \frac{1}{\tau} \sum_{j=-p}^p Y_{t+j} \quad \text{vérifiant } 2p+1=\tau$$

Si τ est un nombre pair et qui est égal, par exemple, à 12. On recourt à une procédure successive de deux moyennes arithmétiques portant respectivement sur 12 et 2 valeurs :

$$Y_t' = \frac{Y_{t-6} + Y_{t-5} + \dots + Y_t + \dots + Y_{t+5}}{12}$$

$$Y_{t+1}' = \frac{Y_{t-5} + Y_{t-4} + \dots + Y_t + \dots + Y_{t+6}}{12}$$

$$Y_t^* = \frac{Y_t' + Y_{t+1}'}{2} = \frac{0.5Y_{t-6} + Y_{t-5} + \dots + Y_t + \dots + Y_{t+5} + 0.5Y_{t+6}}{12} \quad (3.10)$$

Cette procédure permet de centrer le calcul. C'est pourquoi Y_t^* porte le nom de "moyenne mobile centrée d'ordre 12".

Pour conclure sur les méthodes de décomposition, on mentionne tout simplement une fameuse méthode qui a été développée par "The Bureau of the Census of the U.S. Department of Commerce" : c'est la méthode "Census II". Pour plus de précisions, on propose au lecteur de se reporter à l'ouvrage [81].

III - METHODE UNIVARIEE DE BOX-JENKINS (ARIMA)

[9, 19, 24, 81, 82]

La méthode de Box-Jenkins est beaucoup plus sophistiquée par rapport aux méthodes que nous venons de présenter. Ce paragraphe sera consacré à la présentation de cette méthode.

Dans un premier temps, nous étudions les différents modèles théoriques, qui sont en fait les lois sous-jacentes des processus stochastiques. Dans un deuxième temps, nous présentons les différentes étapes de la méthodologie de Box-Jenkins.

III.1 - Modèles ARMA, ARIMA et SARIMA

III.1.1 : Généralité

On introduit tout d'abord deux opérateurs utilisés dans l'exposé de ce paragraphe, ils sont :

- * Opérateur retard B, avec $BY_t = Y_{t-1}$ et $B^\tau Y_t = Y_{t-\tau}$.
- * Opérateur différence d'ordre d ou D.
 - différentiation courte (d'une période à l'autre): $(1-B)^d$ ou Δ^d .
 - différentiation longue (saisonnrière): $(1-B^\tau)^D$ ou Δ_τ^D , τ étant la périodicité.

A partir de maintenant, nous proposons de considérer le processus stochastique Y_t comme la sortie d'un filtre linéaire dont l'entrée est un bruit blanc a_t (fig.14). L'inertie du filtre s'écrit...

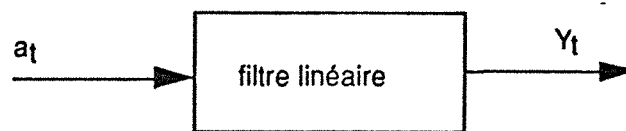


Fig. 14 - Filtre linéaire.

$$Y_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B)a_t \quad (3.11)$$

... où μ est un paramètre qui représente le niveau moyen, et la fonction $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$ un opérateur qui transforme a_t en Y_t , appelée notamment fonction de transfert du filtre linéaire (on se réfère implicitement à la théorie du système).

Pour simplifier la discussion (sans changer la généralité), on suppose que $\mu=0$ et on se ramène à une expression du type :

$$Y_t = \psi(B)a_t \quad (3.12)$$

Théoriquement, la chaîne $1, \psi_1, \psi_2, \dots$ peut être finie ou infinie. Si elle est finie ou infinie mais convergente, le filtre est stable, si bien que le processus Y_t soit stationnaire. Si la chaîne est infinie et divergente, le filtre est instable, et le processus Y_t est par conséquent non stationnaire.

La condition de stationnarité est donc assurée par la convergence de $\psi(B)$ pour $|B| \leq 1$

Sous cette condition de stationnarité, (3.12) implique que Y_t peut être représenté par la somme des valeurs antérieures pondérées de Y_t plus un bruit blanc a_t :

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j Y_{t-j} + a_t \quad (3.13)$$

ou bien $\pi(B)Y_t = a_t \quad (3.14)$

... où $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots$, et la condition d'inversibilité est que $\pi(B)$ converge pour $|B| \leq 1$

Remarque 1 : la condition de stationnarité et celle d'inversibilité sont indépendantes.

Remarque 2 : la relation entre $\pi(B)$ et $\psi(B)$ est exprimée par: $\pi(B) = \psi^{-1}(B)$ et les deux expressions (3.12 et 3.14) sont notamment équivalentes ou duales.

Un processus stochastique ainsi défini peut être représenté, convenablement, par des modèles particuliers. Si un processus est stationnaire, nous faisons appel aux modèles ARMA (p,q), tandis que s'il est non stationnaire, aux modèles ARIMA(p,d,q). Par ailleurs, si la saisonnalité se présente dans le processus, nous avons le recours aux modèles SARIMA.

III.1.2 : Modèles ARMA(p,q)

Il s'agit de décrire un processus stationnaire qui peut être représenté, soit par un modèle "autorégressif" AR(p), soit par un modèle "moyenne mobile" MA(q), soit par un modèle mixte ARMA(p,q).

1) Modèle autorégressif d'ordre p : AR(p)

Dans l'expression (3.13), si l'on suppose que seuls les premiers p coefficients de pondération ($\pi_1, \pi_2, \dots, \pi_p$) ne sont pas nuls, on aura :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad (3.15)$$

...où l'on substitue le symbole π par ϕ .

Le modèle (3.15) est dit autorégressif d'ordre p . Le caractère de régression de Y_t sur Y_{t-1} , Y_{t-2} , ..., Y_{t-p} explique la dénomination de ce modèle.

Ce modèle peut s'écrire de façon plus ramassée en s'introduisant l'opérateur B :

$$\phi_p(B)Y_t = a_t \quad (3.16)$$

...où $f_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, est l'opérateur autorégressif.

Si le polynôme $\phi_p(B)$ a des racines de module supérieur à 1, le processus Y_t est stationnaire. Pour un processus Markovien AR(1) ($Y_t = \phi Y_{t-1} + a_t$), par exemple, la condition de stationnarité s'exprime par $|\phi| < 1$.

Sous la condition de stationnarité, (3.16) peut s'écrire :

$$Y_t = \phi_p^{-1}(B)a_t \quad (3.17)$$

ce qui signifie qu'un processus autorégressif stationnaire peut être considéré comme la sortie d'un filtre linéaire dont la fonction de transfert est $\phi_p^{-1}(B)$, lorsque l'entrée du filtre est un bruit blanc a_t .

Remarque 1: un processus autorégressif est par nature inversible, puisque $\pi(B) = \phi_p(B)$ est une chaîne finie.

Remarque 2: un processus autorégressif représente un système à contre-réaction (cf. Chapitre 1).

Du modèle (3.15), nous pouvons déduire facilement l'équation de Yule-Walker:

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \dots + \phi_p \rho_{\tau-p} \quad \tau > 0 \quad (3.18)$$

$\rho_\tau, \rho_{\tau-1}, \dots, \rho_{\tau-p}$ étant les coefficients d'autocorrélation.

L'identification de l'ordre p d'un processus autorégressif $AR(p)$ fait appel au coefficient d'autocorrélation partielle défini comme le coefficient de corrélation partielle entre les valeurs à deux instants séparés de τ :

$$\gamma_{\tau} = \gamma(Y_t, Y_{t+\tau} / Y_{t+1}, \dots, Y_{t+\tau-1}) \quad (3.19)$$

Le coefficient γ_{τ} mesure ce que peut encore expliquer Y_t sur $Y_{t+\tau}$, une fois tenu compte de ce que peuvent apporter en remontant dans le temps $Y_{t+\tau-1}, Y_{t+\tau-2}, \dots, Y_{t+1}$ dans l'explication de $Y_{t+\tau}$. Ce coefficient peut être calculé à partir de l'équation de Yule-Walker (Un algorithme récursif a été donnée par DURBIN [9]). Le graphe $\gamma_{\tau} - \tau$ est appelé "*corrélogramme partiel*".

Si l'on définit l'enveloppe E par la courbe représentée sur la figure 15 on peut montrer que ρ_{τ} est une fonction amortie incluse dans l'enveloppe E , et $\gamma_{\tau} = 0$ pour $\tau > p$. C'est avec cette caractéristique de γ_{τ} qu'on peut identifier l'ordre p d'un processus autorégressif.

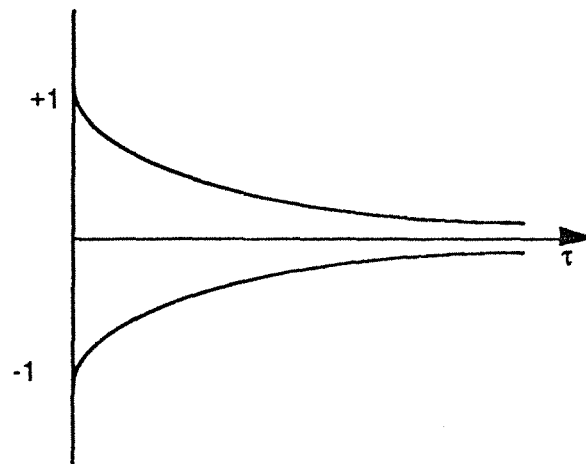


Fig.15 - Enveloppe E.

2) Modèle "moyenne mobile" $MA(q)$

Dans la formule (3.12), si l'on suppose que $\psi(B)$ est une chaîne finie, on obtient un processus qualifié de moyenne mobile d'ordre q , noté $MA(q)$:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (3.20)$$

$$\dots \text{ ou bien } Y_t = \theta_q(B)a_t \quad (3.21)$$

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ étant opérateur moyenne mobile.

Un processus moyenne mobile est par nature stationnaire, puisque la chaîne $\psi(B)$ est finie. Par ailleurs, il est inversible si les zéros de $\theta(B)$ sont de module supérieur à 1. Sous la condition d'inversibilité, (3.21) peut s'écrire:

$$\theta_q^{-1}(B)Y_t = a_t \quad (3.22)$$

On constate qu'un processus moyenne mobile peut être considéré comme la sortie d'un filtre linéaire dont la fonction de transfert est $\theta_q(B)$, lorsque l'entrée du filtre est un bruit blanc a_t .

On peut montrer que γ_τ pour un MA(q) est une fonction amortie incluse dans l'enveloppe E définie par la figure 15 et ρ_τ est une fonction telle que :

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1 \theta_{\tau-1} + \dots + \theta_q \theta_{\tau-q}}{1 + \theta_1^2 + \dots + \theta_q^2} & \tau = 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (3.23)$$

ρ_τ constitue donc un outil de base qui peut servir à identifier l'ordre q pour un processus moyenne mobile MA(q).

Remarque: Un processus moyenne mobile représente un système sans contre-réaction [cf. Chapitre 1].

3) Modèles ARMA (p,q)

En combinant les processus AR et MA, on peut définir ARMA(p,q) par :

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.24)$$

$$\text{ou bien : } \phi_p(B)Y_t = \theta_q(B)a_t \quad (3.25)$$

La condition de stationnarité est que $\phi_p(B)=1-\phi_1B-\phi_2B^2-\dots-\phi_pB^p$ a des racines de module supérieur à 1. La condition d'inversibilité est que $\theta_q(B)=1-\theta_1B-\theta_2B^2-\dots-\theta_qB^q$ a des racines de module supérieur à 1.

Sous la condition de stationnarité (3.25) peut s'écrire...

$$Y_t = \frac{\theta_q(B)}{\phi_p(B)} a_t \quad (3.26)$$

... ce qui signifie qu'un processus ARMA(p,q) peut être considéré comme la sortie d'un filtre linéaire dont la fonction de transfert est $\frac{\theta_q(B)}{\phi_p(B)}$, lorsque l'entrée du filtre est un bruit blanc a_t .

Remarque: Un processus autorégressif-moyenne mobile représente un système à contre réaction [cf. Chapitre 1].

L'étude des fonctions d'autocorrélation et d'autocorrélation partielle associée à un tel processus met en évidence un comportement moins simple que celui des modèles AR ou MA : autocorrélation est une fonction amortie incluse dans l'enveloppe E pour $\tau > q-p$, autocorrélation partielle a un comportement général inclus dans l'enveloppe E pour $\tau > p-q$.

Dans la pratique, on admette souvent que les ordres p et q sont inférieurs ou égaux à 2. Nous illustrons, dans la figure 16, les fonctions d'autocorrélation et d'autocorrélation partielle pour les modèles AR(1), AR(2), MA(1), MA(2) et ARMA(1,1).

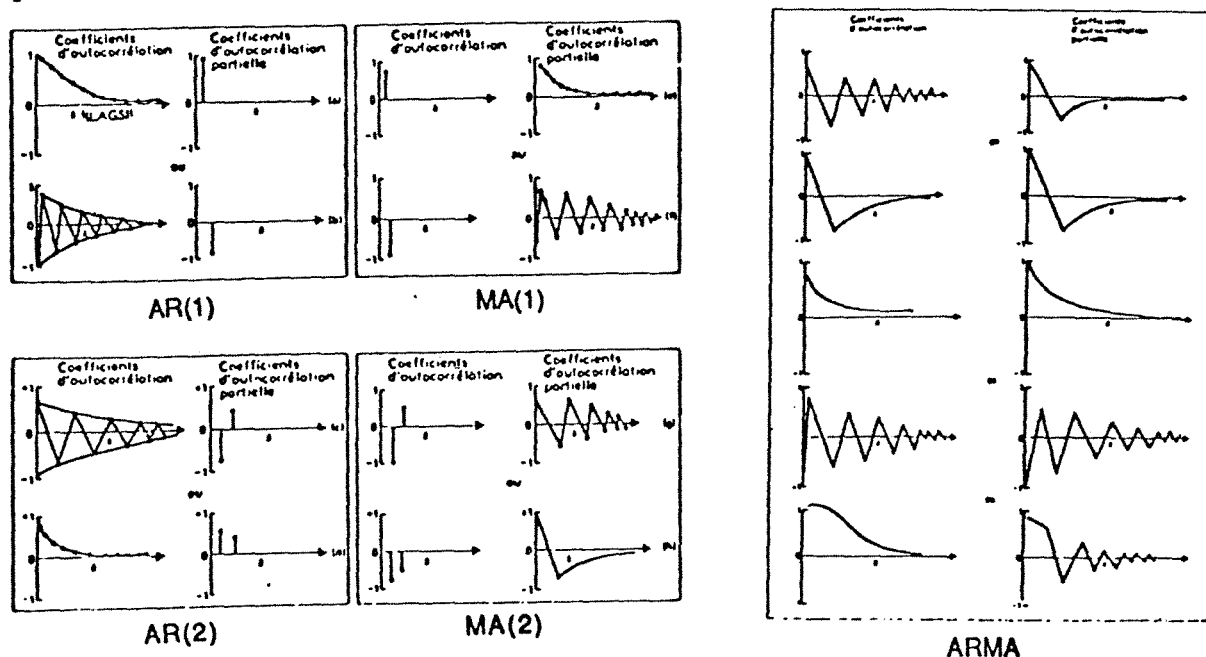


Fig.16 - Fonctions d'autocorrélation et d'autocorrélation partielle (Extrait de [82]).

III.1.3 : Modèles ARIMA(p,d,q) et SARIMA

Un processus peut se manifester non stationnaire en moyenne. Box-Jenkins [9] ont postulé qu'il était parfois possible de le rendre stationnaire en lui appliquant un filtre différence d'ordre d. La modélisation ARMA du processus $\{\Delta^d Y_t\}$ conduit au modèle ARIMA(I comme integrated).

$$\phi_p(B)\Delta^d Y_t = \theta_q(B)a_t \quad (3.27)$$

ou bien :

$$\phi_{p+d}(B)Y_t = \theta_q(B)a_t \quad (3.28)$$

$\phi_{p+d}(B) = \phi_p(B)\Delta^d = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{p+d} B^{p+d}$ étant l'opérateur autorégressif généralisé.

Une autre cause importante de non stationnarité est que la dispersion n'est pas constante, c'est-à-dire que le processus est non stationnaire en variance. Une famille de transformation, souvent utilisée pour rendre le processus stationnaire, a été préconisée par Box et Cox [19, 24, 34] :

$$W_t = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases} \quad (3.29)$$

...où λ est un paramètre à choisir.

De plus, un processus Y_t peut être périodique ou saisonnier. La formulation générale qui tient compte de la saisonnalité est la suivante...

$$\phi_p(B)\Phi_P(B^\tau)\Delta^d \Delta_\tau^D Y_t = \theta_q(B)\Theta_Q(B^\tau)a_t \quad (3.30)$$

... où Δ^d est la différentiation non saisonnière d'ordre d; Δ_τ^D est la différentiation saisonnière d'ordre D; p, P, q, Q, d et D sont des entiers non négatifs; $\phi_p(B)$ et $\theta_q(B)$ sont des polynômes en B; $\Phi_P(B^\tau)$ et $\Theta_Q(B^\tau)$ sont des polynômes en B^τ ; τ est la périodicité.

III.2 – Méthodologie de Box-Jenkins

La méthode de Box-Jenkins peut être résumée par la figure 17. Elle implique quatre étapes : Identification, estimation, validation et prévision.

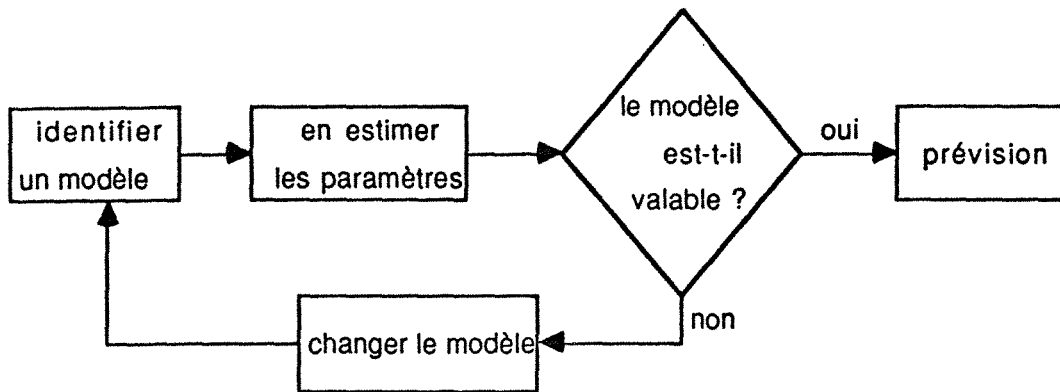


Fig.17 - Méthode de Box-Jenkins (Extrait de [24]).

III.2.1 : Identification

Cette étape consiste à déterminer les paramètres p , d , q (ainsi que P , D , Q s'il existe une composante saisonnière). La détermination de d et D est basée sur la vitesse de décroissance du corrélogramme de la chronique en jeu, et la détermination des autres paramètres repose sur le comportement théorique du corrélogramme et du corrélogramme partiel (fig.16). L'identification réside dans une comparaison de ce comportement théorique avec celui des estimateurs des coefficients d'autocorrélation et d'autocorrélation partielle de la chronique étudiée. En général, si l'autocorrélation tend exponentiellement vers zéro, on se trouve devant un modèle AR dont l'ordre dépend du nombre des coefficients d'autocorrélation partielle, significativement différents de zéro. Si l'autocorrélation partielle tend vers zéro, c'est un modèle MA : son ordre dépend du nombre des coefficients d'autocorrélation significatifs statistiquement. Lorsque ce sont les deux qui tendent exponentiellement vers zéro, nous avons un modèle mixte ARMA.

Cependant, il faut savoir que les corrélogrammes calculés n'engendrent pas toujours un choix évident pour les valeurs des ordres p , d , q , P , D , Q .

III.2.2 : Estimation

Une fois un modèle retenu expérimentalement par l'examen des corrélogrammes, les paramètres sont à estimer.

Les méthodes d'estimation recourent au principe du maximum de vraisemblance qui, moyennant certaines simplifications, s'expriment sous une version non linéaire du critère des Moindres-Carrés. Nous proposons au lecteur de consulter l'ouvrage [34, 76] pour plus de détails sur ces méthodes.

III.2.3 : Validation (Diagnostic)

La problème qui se pose ici est de savoir si le modèle choisi et ajusté dans les deux étapes précédentes peut être considéré comme valable.

Les résidus calculés à partir du modèle ajusté jouent un rôle capital dans cette étape. Nous pouvons examiner les résidus: si les résidus sont aléatoires, le modèle ajusté, ayant bien éliminé toutes les lois régissant la chronique étudiée, est valable; si non, la validité du modèle essayé doit être rejetée.

Le corrélogramme est un outil efficace dans le but de tester si les résidus sont aléatoires, car il indique les rapports existants entre les valeurs successives des résidus, et tous les coefficients d'autocorrélation doivent être proches de zéro si les résidus sont aléatoires.

Enfin, si le modèle ajusté est considéré comme valable, on peut se ramener à la prévision, si non, il ne nous reste plus qu'à reprendre les trois étapes d'identification, d'estimation et de validation pour un autre modèle, autant de fois que cela peut s'avérer nécessaire.

III.2.4 : Prévision

Une fois retenu le modèle, la prévision devient une chose simple et purement mécanique.

Supposons que l'on se trouve à l'instant T (correspondant à la dernière observation de Y_t), la prévision $\hat{Y}_T(h)$ obéit au principe de minimiser $E\{[Y_{t+h} - \hat{Y}_T(h)]\}$.

Dans le cas général du modèle ARIMA, il est montré que [9, 24] :

$$Y_{T+h} = \hat{Y}_T(h) + e_T(h) \quad (3.31)$$

$$\dots \text{où } \hat{Y}_T(h) = \sum_{j=1}^{\infty} \alpha_j(h) Y_{T+1-j} \quad \text{et} \quad e_T(h) = \sum_{i=0}^{h-1} \beta_i \varepsilon_{T+h-i} \quad \text{avec } \alpha_1(0) = \alpha_2(0) = \dots = 0; \beta_0 = 0.$$

Les coefficients $\beta_i (i=1, 2, \dots, h-1)$ s'obtiennent à partir des paramètres du modèle en égalant les coefficients des puissances identiques de B dans l'équation :

$$\phi_p(B)(1-B)^d[1+\beta_1B+\beta_2B^2+\dots]=\theta_q(B)$$

D'autre part les coefficients $\alpha_j(h)$ sont donnés par la récurrence...

$$\alpha_j(h)=\alpha_{j+h}(h-1)+\beta_{h-1}\alpha_j$$

... où les paramètres $\alpha_j=\alpha_j(1)$ sont aussi obtenus par égalisation des coefficients des mêmes puissances de B dans l'équation :

$$\theta_q(B)(1-\alpha_1B-\alpha_2B^2-\dots)=\phi(B)(1-B)^d$$

De façon générale, il est recommandé de suivre en pratique la procédure suivante :

- a) Déterminer un modèle selon la méthodologie de Box-Jenkins.
- b) Ecrire le modèle choisi en exprimant Y_{T+h} linéairement en fonction de $Y_{T+h-1}, \dots, Y_T, \dots, \varepsilon_{T+h}, \dots, \varepsilon_T, \dots$
- c) Remplacer dans le membre de droite de cette expression les valeurs futures de ε_t par 0 (0 est la valeur moyenne de ε_t), les valeurs passées de ε_t par $\varepsilon_{t-1}(1)=Y_t-\hat{Y}_{t-1}(1)$, les valeurs futures Y_{t+h} par $\hat{Y}_T(h)$ et les valeurs passées de Y_t par ses valeurs observées.
- d) Utiliser les formules de mise en jour dès qu'on dispose d'une nouvelle observation.

Le programme d'ordinateur peut fournir toutes les prévisions que nous désirons, assorties de leurs intervalles de confiance.

IV - CONCLUSION

Dans ce chapitre, trois méthodes de prévision ont été présentées. Nous formulons quelques remarques suivantes :

* Comme nous l'avons montré, ces trois méthodes sont très différentes du point de vu raisonnement :

Les méthodes de lissage et les méthodes de décomposition se limitent à certains types de lois spécifiques, qu'elles supposent toujours présentes dans les données. Le lissage exponentiel, par exemple, fait l'hypothèse d'une loi horizontale. Pourtant, dans la méthode de Box-Jenkins, il n'y a pas besoin de faire dès le départ l'hypothèse d'une loi rigide. Au lieu de cela, on commence par une loi expérimentale adaptée aux données en vue de minimiser l'erreur. Ensuite, la méthode nous fournit des informations explicites qui nous donnera la possibilité de juger si la loi adaptée à titre préliminaire est correcte pour la chronique étudiée. Si elle l'est, on peut établir directement la prévision, et si elle ne l'est pas, la méthode de Box-Jenkins donne des indications supplémentaires permettant d'identifier la loi correcte. On peut donc, expérimentalement, spécifier un modèle de prévision correspondant à une loi de rechange et isoler les informations dont on a besoin pour juger quel est le modèle le plus proche de la loi sous-jacente aux données. Ce processus d'itération permet d'apporter de proche en proche des améliorations à la méthode, pour arriver à un modèle de prévision qui assure l'optimisation de la loi de base et qui minimise l'erreur par rapport à cette loi. Par conséquent, la méthode de Box-Jenkins donne, en général, une meilleure précision par rapport aux deux autres méthodes et son application est naturellement plus coûteuse.

* Dans les méthodes de lissage, plus β est proche de 1, plus l'influence des observations passées remontera loin dans le temps et plus la prévision sera rigide. Au contraire, plus β est proche de zéro, plus la prévision sera souple. Le choix de β peut résulter de l'expérience acquise sur le phénomène en jeu, ou dans une situation nouvelle, ce choix est empirique, les valeurs les plus utilisées étant 0.7 et 0.8.

* Dans la méthode de Box-Jenkins, le choix d'un modèle n'est pas toujours univoquement déterminé, surtout pour des chroniques dont les comportements sont complexes. On peut se ramener à utiliser plusieurs modèles concurrents et rien n'indique qu'on pourra dans tous les cas obtenir un modèle de prévision satisfaisant. Il est difficile, dans un cas général, d'attacher une interprétation physique aux modèles de prévision construits à l'aide de cette méthode.

Chapitre 4

S O M M A I R E

I	Méthode de Régression Multiple	50
I.1	Méthode linéaire classique.....	50
I.1.1	: Estimation des paramètres $a_0, a_1, a_2, \dots, a_j, \dots, a_p$ (ajustement des Moindres-Carrés)	52
I.1.2	: Equation d'analyse de variance et coefficient R^2	53
I.1.3	: Intervalles de confiance et tests de signification pour les coefficients.....	55
I.1.4	: Prévision	56
I.2	Méthodes des Moindres-Carrés Généralisées	57
II	Méthode Multivariée de Box-Jenkins(MARIMA)	58
III	Méthode de déconvolution par projection [60]	61
IV	Méthode de déconvolution $COV_x(\tau) - COV_y(\tau)$	64
IV.1	Rappel sur la théorie des variables régionalisées.....	64
IV.1.1	: Variogramme.....	64
IV.1.2	: Comportement du graphe $\gamma(\tau)-\tau$	65
IV.1.3	: Ajustement à un modèle	67
IV.2	Méthode de déconvolution géostatistique [73]	68
V	Conclusion	69

CHAPITRE 4

METHODES DE PREVISION MULTIVARIEES

Les méthodes de prévision multivariées cherchent à saisir et à mesurer l'influence d'une ou plusieurs chroniques indépendantes ou explicatives sur la chronique dépendante ou expliquée. Afin de simplifier la discussion, on restreint notre propos au cas d'une seule chronique explicative, et on considère de plus que le système est discret (fig.18).

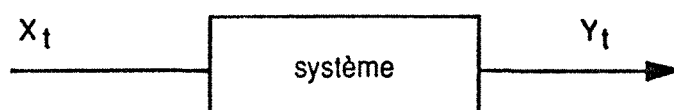


Fig. 18 - Schéma du système discret.

Sous les conditions de linéarité et d'invariance par translation dans le temps, la relation entre Y_t et X_t peut être représentée de façon discrète par la convolution [cf. Chapitre 1].

$$Y_t = \sum_{\tau=0}^m X_{t-\tau} \phi_{\tau} = \phi_0 X_t + \phi_1 X_{t-1} + \dots + \phi_m X_{t-m}$$

Cette formule sera exprimée par la suite sous une forme synthétique :

$$Y = X * \phi \tag{4.1}$$

où ϕ est, par définition, la réponse impulsionnelle. Mais dans les applications, on l'appelle souvent (par abus de langage) fonction de transfert. Elle sera notée par la suite FT. m est la mémoire du système. La FT représente toutes les caractéristiques du système. On peut ainsi construire la sortie correspondante à une entrée en convoluant celle-ci par la FT, et pour cela il faut au préalable connaître cette dernière. Donc, le problème de prévision se convertit souvent en un problème de déconvolution : c'est la recherche de la FT à partir des signaux d'entrée et de sortie.

Les méthodes de déconvolution classiques et couramment utilisées sont les méthodes spectrales (FFT, Laplace, *etc.*) et les méthodes du type problème inverse (Moindres Carrés, Moindres-Carrés avec contraintes, *etc.*) [60, 73].

Les méthodes spectrales repose sur une propriété fondamentale de la transformée de Fourier, à savoir que la convolution dans le domaine temporel devient un produit simple dans le domaine fréquentiel. On peut ainsi calculer la transformation des signaux d'entrée et de sortie du système, puis la transformée de la fonction de transfert. L'application de la transformation inverse pourra alors aboutir à une solution de FT.

Ces méthodes classiques fonctionnent bien lorsque l'entrée et la sortie vérifient parfaitement une relation de convolution. Dans la pratique, cependant, les conditions de linéarité et d'invariance ne sont jamais parfaitement remplies. De plus, les observations sur l'entrée et sur la sortie sont souvent entachées d'erreurs. La convolution doit s'écrire de façon à faire intervenir un terme résiduel [60] :

$$Y = X * \phi + \epsilon \quad (4.2)$$

L'intervention du terme résiduel provoque une grande difficulté de convolution-déconvolution, puisqu'en fait, la solution de déconvolution n'est pas unique. C'est pour surmonter cette difficulté que quelques nouvelles méthodes ont été proposées.

Dans ce chapitre, nous allons présenter quatre méthodes de déconvolution : la méthode régression multiple (méthode classique), la méthode MARIMA [9, 81, 82], la méthode de projection [60] et la méthode géostatistique [73].

I - METHODE DE REGRESSION MULTIPLE

I.1 - Méthode linéaire classique

[18, 34, 48, 67, 76]

Il s'agit de l'étude de l'influence de p variables "*explicatives*" $X_1, X_2, X_j, \dots, X_p$ sur la variable "*expliquée*" Y . On dispose pour cela de n ensembles d'observations de $(p+1)$ variables :

$$\{ y_i, x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{ip} \}$$

Nous admettrons que la relation théorique entre les variables est de la forme :

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_j x_{ij} + \dots + \alpha_p x_{ip} + \epsilon_i \quad (4.3)$$

...où ε_i représente les fluctuations aléatoires que nous supposons centrées, d'égale dispersion et indépendantes (hypothèse de Gauss-Markov) :

$$E\varepsilon_i=0, E\varepsilon_i^2=\sigma^2, E\varepsilon_i\varepsilon_{i'}=0 (i \neq i')$$

Il sera commode d'utiliser une écriture matricielle synthétique pour l'ensemble des n équations du modèle, soit sous la forme :

$$Y=\alpha_0+\alpha_1X_1+\alpha_2X_2+\dots+\alpha_jX_j+\dots+\alpha_pX_p+\varepsilon$$

$Y, X_1, X_2, \dots, X_j, \dots, X_p$ et ε sont des vecteurs colonne à n composantes.

Soit encore sous la forme matricielle :

$$Y=X\alpha+\varepsilon \quad \text{avec } E(\varepsilon)=\mathbf{0}, V(\varepsilon)=\sigma^2\mathbf{I} \quad (4.4)$$

où Y et ε sont des vecteurs colonne à n composantes et X est une matrice à n lignes et à $p+1$ colonnes :

$$X = \begin{pmatrix} 1 & x_1 & \dots & x_{1j} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2j} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nj} & \dots & x_{np} \end{pmatrix}$$

$\mathbf{0}$ est un vecteur nul à n composantes, \mathbf{I} est une matrice unité.

Plus concrètement, le modèle linéaire classique consiste à résoudre les problèmes suivants :

- estimation des coefficients $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_p$.
- liaison entre Y et les X_j (analyse de la variance).
- liaison entre Y et les X_j est-elle
 - : globalement significative ?
 - : marginalement significative ?
- précision des résultats.
- intervalle de prévision.

1.1.1 : Estimation des paramètres $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_p$ (ajustement des Moindres-Carrés)

Théoriquement, le système (4.4) admet une infinité de solutions. On appelle ajustement du modèle toute solution du système d'équations :

$$y_i = a_0 + a_1 x_{i1} + a_2 x_{i2} + \dots + a_j x_{ij} + \dots + a_p x_{ip} + e_i \quad (4.5)$$

L'ajustement des Moindres-Carrés est celui qui fournit les coefficients $a_0, a_1, a_2, \dots, a_j, \dots, a_p$ conduisant au minimum de la somme des carrés des écarts e_i définis par :

$$e_i = y_i - (a_0 + a_1 x_{i1} + a_2 x_{i2} + \dots + a_j x_{ij} + \dots + a_p x_{ip})$$

... ou sous forme matricielle, cet ajustement est défini par...

$$Y = Xa + e \quad (4.6)$$

L'ajustement des Moindres-Carrés correspond à la solution qui minimise le produit scalaire $e'e = \sum e_i^2$.

Si n est supérieur ou égal à p et si X est de plein rang (c'est-à-dire de rang $p+1$, ce qui signifie que les variables explicatives $X_1, X_2, \dots, X_j, \dots, X_p$ sont linéairement indépendantes), alors $X'X$ est une matrice d'ordre $p+1$ et de rang $p+1$, donc inversible. Le critère des Moindres-Carrés conduit à la solution unique :

$$a = (X'X)^{-1} X'Y \quad (4.7)$$

Le critère des Moindres-Carrés se prête à une interprétation géométrique claire. En effet, cette technique d'ajustement appliquée au modèle $Y = X\alpha + E$ consiste à approcher Y par sa projection orthogonale Xa sur le sous-espace engendré par les colonnes de X (fig.19).

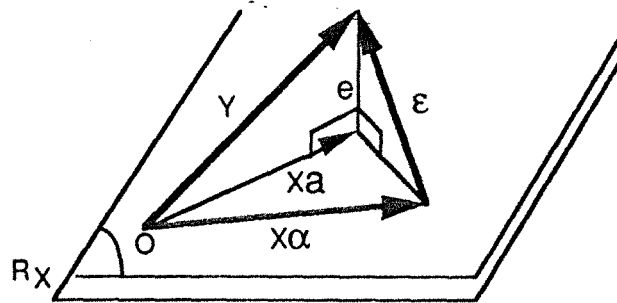


Fig. 19 - Ajustement des Moindres-Carrés (extrait de [18]).

Sous les hypothèses de Gauss-Markov, les estimateurs (4.7) fournis par la technique des Moindres-Carrés sont les meilleurs au sens suivant :

$$\begin{aligned} E(a) &= \alpha \quad (\text{estimateurs sans biais}) \\ V(a) &= \sigma^2 (XX)^{-1} \quad (\text{estimateur à variance minimale}) \end{aligned}$$

C'est en effet le théorème de Gauss-Markov.

De plus, on peut montrer que...

$$s^2 = \frac{1}{n-p} \sum e_i^2$$

.. est un estimateur sans biais de σ^2 .

1.1.2 : Equation d'analyse de variance et coefficient R^2

Le calcul précédent nécessite l'inversion de la matrice symétrique $X'X$ qui est de rang $p+1$. On peut simplifier ce calcul en centrant les variables, on écrit alors :

$$Y_c = X_c \alpha + \varepsilon \quad (4.8)$$

où α représente le vecteur des coefficients de régression, mais il s'agit maintenant du vecteur à p composantes $\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_p$ à l'exclusion de α_0 .

L'ajustement des Moindres-Carrés conduit à une solution unique qui s'écrit cette fois-ci par :

$$\begin{cases} a = (X'_c X_c)^{-1} X'_c Y_c \\ a_0 = Y - \sum_j a_j X_j \end{cases}$$

La matrice $\frac{1}{n} X'_c X_c$ (que nous notons V_{XX}) est appelée matrice des variances-covariances des variables explicatives et le vecteur $\frac{1}{n} X'_c Y_c$ (que nous notons V_{XY}) est le vecteur des covariances des variables explicatives avec la variable expliquée. On peut montrer que :

$$a = V_{XX}^{-1} V_{XY} \quad \text{et} \quad V(a) = V_{XX}^{-1} \frac{\sigma^2}{n}$$

Pour mesurer globalement le degré d'explication de la variable expliquée Y par les variables explicatives $X_1, X_2, \dots, X_j, \dots, X_p$ dans leur ensemble, on définit le coefficient de corrélation multiple par :

$$R = \frac{\|X_{ca}\|}{\|Y_c\|}$$

qui n'est autre que le cosinus de l'angle β entre les vecteurs Y_c et X_{ca} . Si il est égal à 1, la liaison est fonctionnelle. Si il est nul, il y a indépendance entre la variable Y et l'ensemble des variables $X_1, X_2, \dots, X_j, \dots, X_p$ (fig. 20).

Le carré du coefficient de corrélation multiple...

$$R^2 = \frac{\text{Cov}(Y, \hat{Y})}{\text{Var}(\hat{Y})\text{Var}(Y)} = \frac{\text{Var}(\hat{Y})}{\text{Var}(Y)} = \frac{\text{Cov}(Y, X_{ca})}{\text{Var}(Y)}$$

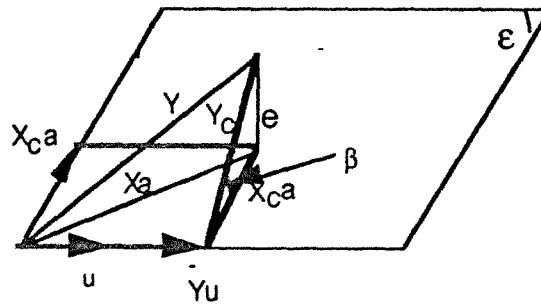


Fig. 20 - Analyse de la variance (extrait de [18]).

... prend un sens en terme de partage de la variance totale en variance expliquée et "résiduelle":

$$\text{Var}(Y) = \text{Var}(\hat{Y}) + \text{Var}(e)$$

Notons que l'on peut calculer R^2 de la façon suivante :

$$R^2 = 1 - \frac{\text{Var}(e)}{\text{Var}(Y)} = 1 - \frac{\sum e_i^2}{n\text{Var}(y)}$$

Il apparaît sur cette relation qu'en minimisant $\sum e_i^2$ on maximise R^2 . En d'autres termes, l'ajustement des Moindres-Carrés détermine la combinaison linéaire des variables ayant une corrélation maximale avec la variable expliquée Y .

On remarquera par ailleurs que l'introduction dans le modèle d'une nouvelle variable explicative quelconque ne peut que diminuer la somme des carrés des écarts, et par conséquent augmenter R^2 . Dans ces conditions la valeur prise par R ne peut être un critère absolu pour apprécier la qualité de l'ajustement.

1.1.3 : Intervalles de confiance et tests de signification pour les coefficients

On va supposer toujours vraies les hypothèses *a priori* de Gauss-Markov, et par conséquent les propriétés d'optimalité des estimateurs des Moindres-Carrés resteront toujours valides. Cependant, pour pouvoir aller plus loin dans l'induction statistique sur les coefficients du modèle, nous avons besoin de définir complètement la distribution des termes aléatoires du modèle. En effet, l'hypothèse de normalité permet de construire des tests et intervalles de confiance pour les différents estimateurs (il s'agit plus souvent d'une hypothèse conduisant à un formalisme simple plutôt qu'à une spécification de la nature du phénomène étudié) :

Ainsi la statistique...

$$T = \frac{\alpha_j - a_j}{s_j}$$

... (où s_j^2 est le produit par s^2 du j -ième terme diagonale de $(X'CX)^{-1}$) suit une loi de Student à $(n-p)$ degrés de liberté. Elle permet de calculer un intervalle de confiance pour le coefficient α_j pour un risque η :

$$a_j - \eta s_j < \alpha_j < a_j + \eta s_j$$

Elle permet également de tester la nullité d'un coefficient α_j (ce qui répond à la question "Liaison entre Y et les X_j est-elle marginalement significative ?"). Faisant l'hypothèse $\alpha_j = 0$, on calcule la quantité...

$$t_j = \frac{a_j}{s_j}$$

... que l'on compare au seuil de signification lu dans la table de Student.

Par ailleurs, la statistique...

$$\frac{n - k}{k - 1} \frac{R^2}{1 - R^2}$$

... suit une loi de Fisher-Snedecor et permet de tester globalement la significativité des coefficients des variables explicatives vraies.

1.1.4 : Prévision

Matériellement, on disposera d'une ligne supplémentaire notée X_0 dans le tableau X des variables exogènes, et on s'interrogera sur la valeur à attendre pour la variable endogène Y (fig. 21). On définit :

- la valeur à prévoir : $Y_0 = \alpha_0 + \alpha_1 X_{01} + \dots + \alpha_p X_{0p} + \varepsilon_0$;
- le prédicteur des Moindres-Carrés : $\hat{Y} = a_0 + a_1 X_{01} + \dots + a_p X_{0p}$;
- l'erreur de prévision: $Y_0 - \hat{Y}$.

On étend à la variable à prévoir Y_0 les hypothèses de Gauss-Markov et de plus, on ajoute l'hypothèse de normalité. Sous ces conditions, on constate que l'espérance de l'erreur de prévision $E(Y_0 - \hat{Y})$ est nulle et la variance $V(Y_0 - \hat{Y}) = \sigma^2 [(1 + X_0'(X'X)^{-1}X_0)]$ est minimale.

On obtient une estimation sans biais de la variance de l'erreur de prévision en remplaçant σ^2 inconnu par son estimation sans biais s^2 :

$$s_0^2 = s^2 [(1 + X_0'(X'X)^{-1}X_0)]$$

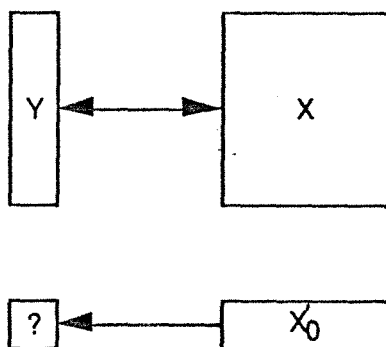


Fig. 21 - Ajustement et prévision (extrait de [49]).

La variable de Student à $(n-p)$ degrés de liberté est définie par :

$$t = \frac{Y_0 - \hat{Y}}{s_0}$$

On en déduit un intervalle de prévision au seuil η .

$$\hat{Y}_0 - t \eta s_0 < Y_0 < \hat{Y}_0 + t \eta s_0$$

I.2 - Méthodes des Moindres-Carrés Généralisées

Les formules (4.2) et (4.4) étant formellement équivalentes, la méthode de régression multiple peut être empruntée dans un but de déconvolution. Mais dans l'analyse des chroniques, il est peu naturel d'admettre l'indépendance du résidu d'une année ou d'un mois, par rapport à l'année ou au mois antérieur. On est obligé d'admettre que $E\varepsilon_i \varepsilon_j \neq 0$ ($i \neq j$), c'est-à-dire que la matrice des variances et covariances des résidus est une matrice V différente de $\sigma^2 I$ (la matrice V est dite mal conditionnée).

Dans le cas d'une autocorrélation des résidus, les estimateurs (4.7) restent sans biais, car cette propriété repose uniquement sur l'hypothèse $E\varepsilon = 0$, non supprimée. Par contre, la matrice des variances et covariances de $a...$

$$V(a) = (X'X)^{-1} X'VX(X'X)^{-1}$$

... est modifiée, puisqu'elle est différente de l'expression $\sigma^2(X'X)^{-1}$. Les estimateurs (4.7) ne sont plus les estimateurs à variance minimale.

On peut montrer que [18, 25, 34], dans la classe des estimateurs sans biais et linéaires, les estimateurs de α à variance minimale sont...

$$a^* = (X'V^{-1}X)^{-1} X'V^{-1}Y \quad (4.11)$$

... et que la matrice de variances et covariances s'écrit :

$$V(a^*) = (X'V^{-1}X)^{-1} \quad (4.12)$$

Or, la détermination de la matrice V est délicate voire impossible dans un cas général et ce problème de Moindres-Carrés généralisés est souvent traité par deux procédures d'approche :

- examen des résidus de l'ajustement des Moindres-Carrés ordinaires ; recours à des techniques de diagnostic ; recalcul des coefficients de régression en considérant la structure des résidus [76] ;

- recours à une procédure de régression dite "*ridge regression*" [5, 34, 76]. Il s'agit de calculer, pour différentes valeurs θ (θ varie en général entre 0 et 1), les estimateurs a :

$$a(\theta) = (X'X + \theta I)^{-1} X'Y \quad (4.13)$$

... et de choisir une valeur appropriée de θ (selon des critères), notée θ^* . Les estimateurs cherchés seront les estimateurs biaisés :

$$a(\theta^*) = (X'X + \theta^* I)^{-1} X'Y \quad (4.14)$$

On souligne que ces procédures d'approche représentent des secteurs les plus actifs dans le domaine de régression linéaire ou non linéaire.

II - METHODE MULTIVARIEE DE BOX-JENKINS(MARIMA)

[9, 24, 81, 82]

Les modèles MARIMA sont l'extension des modèles ARIMA. Nous utilisons les modèles multivariés MARIMA pour combiner les concepts de régression multiple aux modèles ARIMA.

Sous les notations utilisées dans la méthode univariée de Box-Jenkins [cf. Chapitre 3], la relation de convolution peut s'écrire :

$$Y_t = \phi_0 X_t + \phi_1 X_{t-1} + \dots + \phi_m X_{t-m} + \varepsilon_t$$

... ou bien :
$$Y_t = (\phi_0 + \phi_1 B + \dots + \phi_m B^m) X_t + \varepsilon_t$$

... ou encore :
$$Y_t = \phi(B)X_t + \varepsilon_t \quad (4.15)$$

Si m est assez grand, l'expression (4.15) n'est pas ramassée. C'est pour cette raison que l'on écrit :

$$y_t = \frac{W(B)}{\delta(B)} x_t - b + \varepsilon'_t$$

ou bien encore :

$$y_t = \frac{W(B)}{\delta(B)} x_t - b + \frac{\theta(B)}{\phi(B)} a_t \text{ avec } \varepsilon'_t = \frac{\theta(B)}{\phi(B)} a_t \quad (4.16)$$

x_t et y_t sont respectivement la transformée de X_t et Y_t (transformation au sens de (3.29), différentiation, désaisonnalisation).

$\frac{W(B)}{\delta(B)}$ est la fonction de transfert

$$W(B) = W_0 - W_1 B - W_2 B^2 - \dots - W_s B^s$$

$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$$

b est le décalage temporel net entre les signaux d'entrée X_t et de sortie Y_t .

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \text{ (opérateur autorégressif)}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \text{ (opérateur moyenne mobile)}$$

b, s, r, p, q sont des constantes

a_t est un bruit blanc

Comme la méthode univariée de Box-Jenkins, cette méthode implique également quatre étapes : identification, estimation, validation et prévision. Toutefois, l'étape d'identification ici est beaucoup plus complexe, car il s'agit de choisir plus de valeurs des ordres (r, s, p, q et b) permettant d'ajuster au mieux les valeurs observées de la chronique Y_t à partir des valeurs de X_t .

On décrit succinctement cette étape d'identification comme suit (on part du modèle (4.16)) :

a) Modélisation (méthode de Box-Jenkins) de la chronique d'entrée :

$$\phi_x(B)x_t = \theta_x(B)\alpha_t$$

et transformation de celle-ci en un bruit blanc α_t

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)} x_t$$

b) Afin de préserver la relation entre x_t et y_t , une transformation analogue doit se faire sur y_t , on aura donc :

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)} y_t$$

et la chronique β_t n'est pas, en général, un bruit blanc (elle l'est lorsque y_t est indépendante de x_t).

c) Calcul des coefficients d'autocorrélation $\rho_\alpha(\tau)$, $\rho_\beta(\tau)$ et des coefficients de corrélation croisée $\rho_{\alpha\beta}(\tau)$ entre α_t et β_t .

d) Estimation directe de la fonction de transfert à l'aide des coefficients de corrélation croisée :

$$\phi_\tau = \frac{\rho_{\alpha\beta}(\tau)\sigma_\beta}{\sigma_\alpha} \quad (\tau=0, 1, 2, \dots, m)$$

où σ_α et σ_β est respectivement l'écart-type de α_t et de β_t .

e) Détermination des valeurs des ordres (r,s,b) de la fonction de transfert.

f) Calcul du terme résiduel ε'_t par :

$$\varepsilon'_t = y_t - \phi_0 x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_m x_{t-m}$$

g) Modélisation (méthode de Box-Jenkins) du terme résiduel ε'_t ...

$$\phi_{\varepsilon'}(B)\varepsilon'_t = \theta_{\varepsilon'}(B)\alpha'_t \quad (\alpha'_t \text{ est un bruit blanc})$$

... afin de déterminer les valeurs des ordres (p,q).

Ayant ainsi identifié les constantes r, s, p, q et b, on dispose de programmes qui calculent les coefficients $w(B)$, $\delta(B)$, $\phi(B)$ et $\theta(B)$.

Il est à noter que, théoriquement, cette méthode représente un avantage décisif de prévision puisqu'elle calcule la fonction de transfert tout en modélisant le terme résiduel ε_t par la méthode univariée de Box-Jenkins. Cependant, son application est très complexe et réservée jusqu'à présent à des spécialistes.

III - METHODE DE DECONVOLUTION PAR PROJECTION[60]

Il s'agit de poser *a priori* que la fonction ϕ inconnue possède une forme analytique connue dépendant de quelques paramètres...

$$\phi = \sum_k a_k f_k(p_1, p_2, \dots, p_i) \quad (4.17)$$

... les a_k et p_i étant des paramètres à identifier, les f_k des fonctions connues.

Le problème revient dès lors à minimiser les résidus du système (4.2) en calculant pour cela les paramètres optimaux a_k et p_i . C'est un problème d'optimisation, qui peut être simple ou complexe suivant la nature de la norme à minimiser (L1 ou L2) et le type des fonctions de base choisie. Les fonctions f_k peuvent se présenter sous la forme de polynômes, des sommes d'exponentielles, des fonctions de Walsh, *etc.*

La méthode de déconvolution proposée par De MARSILY utilise des fonctions de Walsh comme les fonctions de base, dont le principe est présenté ci-dessous. Il faut noter que cette méthode est principalement utilisée en hydrologie. Les contraintes imposées sont liées étroitement aux problèmes hydrologiques.

La convolution peut s'écrire de façon matricielle...

$$\begin{pmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} X_0 & X_{-1} & \dots & X_{-m} \\ X_1 & X_0 & \dots & X_{1-m} \\ \vdots & \vdots & \ddots & \vdots \\ X_n & X_{n-1} & \dots & X_{n-m} \end{pmatrix} \times \begin{pmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_m \end{pmatrix}$$

... ou

$$Y - X \phi = 0 \quad (4.18)$$

On va maintenant décomposer le vecteur inconnu ϕ sur une base de vecteurs complète, celle de Walsh, par exemple (fig. 22) :

$$\phi = \lambda_0 u_0 + \lambda_1 u_1 + \dots + \lambda_m u_m$$

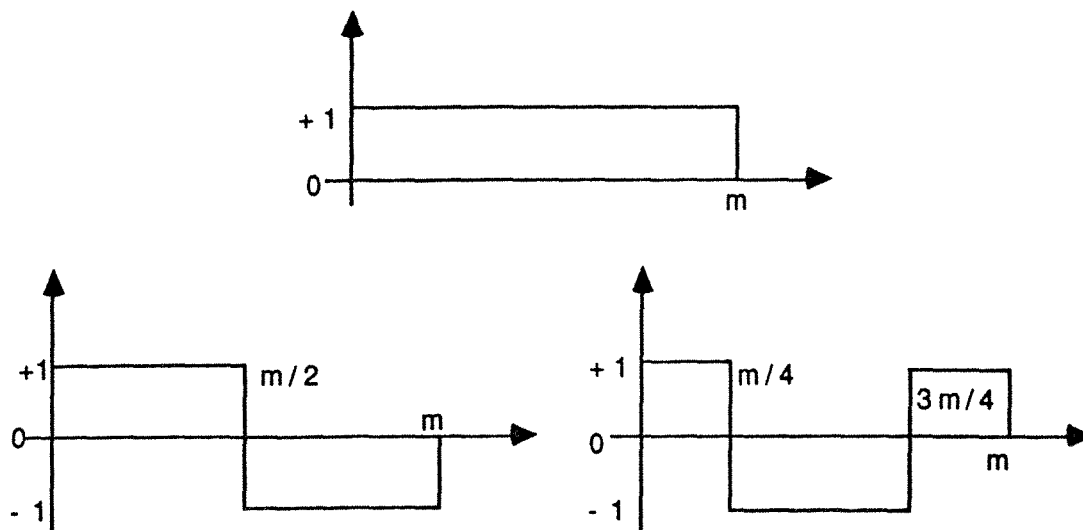


Fig. 22 - Fonctions de Walsh (extrait de [60]).

Pour chercher ϕ , il suffit de calculer ses composantes λ_j sur la base u .

Cependant, au lieu de rechercher ϕ sur une base complète, on va tronquer dès le départ, la base et n'utiliser que les $p + 1$ premiers vecteurs :

$$\text{Min}_{\lambda_j} \| Y - X \phi \|^2 \quad j = 0, 1, \dots, p$$

$$\text{Soit } \text{Min}_{\lambda_j} \| Y - \sum_{j=0}^p \lambda_j X u_j \|^2 \quad j = 0, 1, \dots, p \quad (4.19)$$

On obtient ainsi les valeurs $\lambda_0, \lambda_1, \dots, \lambda_p$.

Le résidu de ce stade est :

$$R_p = Y - \sum_{j=0}^p \lambda_j X u_j$$

Cette projection réalisée, on calcule les composantes suivantes de ϕ sur la base u jusqu'à l'ordre q en utilisant le résidu, c'est-à-dire :

$$\text{Min } \| R_p - \sum_{j=p+1}^q \lambda_j X u_j \|^2 \quad j = p + 1, \dots, q \quad (4.20)$$

On reprojette le résidu plusieurs fois sur les premiers vecteurs de la base et au fur et à mesure que l'on descend dans la base, on affine la connaissance de ϕ en tentant à chaque nouvelle projection de réduire le résidu R .

Le processus est donc par essence itératif. Du fait de cette itération, des contraintes (contrainte de positivité de la FT, contrainte sur l'aire de la FT, contrainte sur la forme de la FT, contrainte de modalité et contrainte de la loi de tarissement, *etc.*) ou des lissages de la FT cherchée peuvent être introduits pas à pas.

On voit que la base de Walsh est par construction hiérarchisée ; on peut donc retenir au début que les premiers vecteurs de base extrêmement réguliers, et éviter ainsi que n'apparaissent des instabilités ou des oscillations dans le calcul de la fonction ϕ . C'est à l'utilisateur de décider à quel stade il arrête le calcul. Ainsi une solution de compromis entre la possibilité physique et la minimisation des erreurs est obtenue. L'estimation est par conséquent biaisée.

On remarque enfin que, cette méthode permet de prendre en compte le phénomène de tarissement. Bien souvent en hydrologie, on constate que la mémoire du système est plus longue que la mémoire considérée. La partie droite de la courbe peut être assimilée à une courbe exponentielle (courbe de tarissement, fig. 23). La prise en compte du tarissement consiste essentiellement à modéliser le terme résiduel résultant du calcul de la FT.

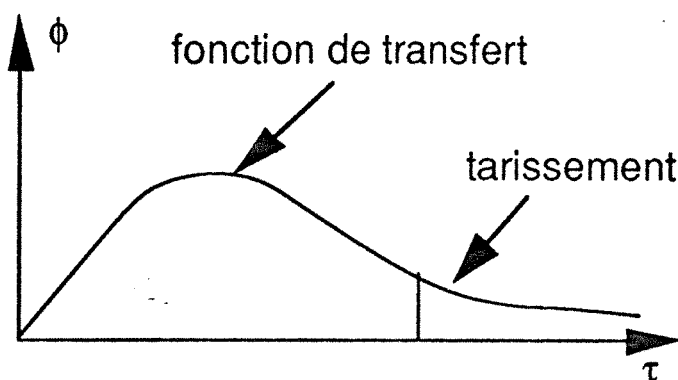


Fig. 23 - Fonction de transfert et tarissement (extrait de [60]).

IV - METHODE DE DECONVOLUTION $COV_X(\tau)$ - $COV_Y(\tau)$

IV.1 - Rappel sur la théorie des variables régionalisées

Le terme "régionalisé" a été proposé par MATHERON 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" (V.R) que l'on note $Y(t)$. Cette variable régionalisée est interprétée comme la réalisation d'une fonction aléatoire (F.A). Elle se présente sous deux aspects contradictoires (ou complémentaires) [62] :

- Un aspect structuré.
- Un aspect aléatoire (haute irrégularité, et variations imprévisibles d'un point à l'autre).

La théorie des variables régionalisées a pour but :

- d'établir les bases théoriques permettant de rendre compte des caractéristiques structurales de phénomènes naturels sous une forme mathématique appropriée : c'est l'analyse structurale (estimation du variogramme, l'examen du comportement de son graphe et son ajustement à un modèle théorique).

- de définir les moyens pratiques pour résoudre les divers problèmes d'estimation qui se posent à partir d'un échantillonnage fragmentaire : c'est la théorie du krigeage.

Sans aller trop loin dans ce sujet, on ne discute ici que l'analyse structurale. On propose au lecteur de se reporter aux ouvrages [32, 61, 62] pour plus de détails.

IV.1.1 : Variogramme

En géostatistique, on remplace l'hypothèse stationnaire d'ordre deux par une hypothèse plus faible mais à signification analogue, c'est l'hypothèse intrinsèque. Cette hypothèse simplificatrice s'énonce comme suit : pour tout vecteur τ , l'accroissement $Y(t+\tau)-Y(t)$ possède une espérance mathématique et une variance indépendantes du point t :

$$\begin{cases} E [Y(t+\tau)-Y(t)] = 0 \\ \text{Var} [Y(t+\tau)-Y(t)] = 2\gamma(\tau) \end{cases} \quad (4.21)$$

Une fonction aléatoire (F.A.) satisfaisant cette hypothèse est appelée Fonction Aléatoire Intrinsèque (F.A.I.). La fonction $\gamma(\tau)$ s'appelle "*demi-variogramme*" ou variogramme. Le variogramme joue un rôle capital dans la théorie de la F.A.I.. En effet, il décrit les propriétés structurales d'une F.A.I. de la même manière que la fonction de covariance $K(\tau)$ décrit celles d'une F.A. stationnaire d'ordre deux.

Si $Y(t)$ vérifie l'hypothèse stationnaire, elle vérifie également l'hypothèse intrinsèque et dans ce cas, on a une relation du type :

$$\gamma(\tau) = K_{YY}(0) - K_{YY}(\tau)$$

Une propriété caractéristique du variogramme est qu'il est une fonction du type positif conditionnelle : dans le cas d'une F.A. stationnaire, toutes les combinaisons linéaires finies

$\sum_i \lambda_i Y(t_i)$ ont une variance finie, et leur usage est toujours légitime ; par contre, dans le cas

d'une F.A.I., seules les combinaisons linéaires vérifiant la condition $\sum_i \lambda_i = 0$ ont une variance finie :

$$\text{Var} \left[\sum_i \lambda_i Y(t_i) \right] = - \sum_i \sum_j \lambda_i \lambda_j \gamma(t_i - t_j)$$

$$\text{avec } \sum_i \lambda_i = 0 \quad (4.22)$$

IV.1.2 : Comportement du graphe $\gamma(\tau)$ -- τ

L'examen du comportement du graphe $\gamma(\tau)$ est particulièrement important, car le variogramme donne une description synthétique de la structure du phénomène en jeu. On étudie en particulier avec soin le comportement au voisinage de l'origine et à l'infini.

Le comportement au voisinage de l'origine traduit la continuité de la V.R. Par ordre de régularité décroissante, on distingue en général quatre types illustrés par la figure 24 : a - allure parabolique, caractérisant une variable extrêmement régulière ; b - allure linéaire, correspondant à une variable moins régulière qu'en a ; c - discontinuité à l'origine (effet de pépité), signifiant que la variable est très irrégulière. Le terme "*effet de pépité*" est réservé pour désigner cette discontinuité du variogramme qui peut être due : soit à une micro-régionalisation d'échelle très inférieure à l'espacement des données, soit à des erreurs des mesures ; d - aspect plat (effet de pépité pure), correspondant au cas aléatoire pur.

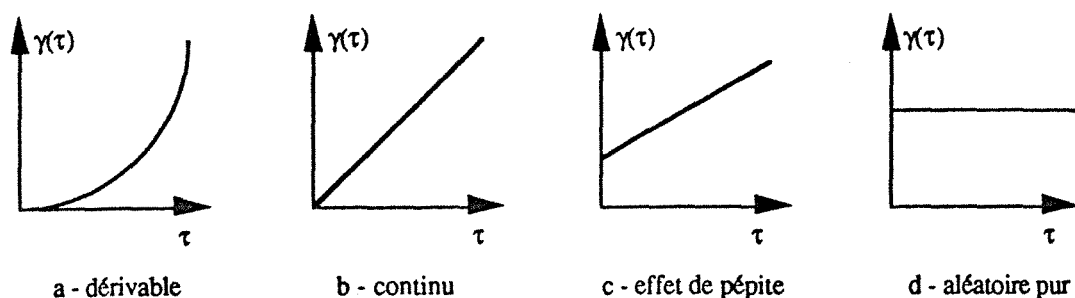


Fig. 24 - Comportement du variogramme au voisinage de l'origine (extrait de [32]).

Outre cet examen du comportement du variogramme au voisinage de l'origine, on étudie également son comportement à l'infini.

Contrairement aux fonctions de covariance qui demeurent nécessairement bornées, le variogramme peut, lui, continuer indéfiniment à croître. Bien évidemment, 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. 25). Aux distances supérieures à la portée, les corrélations sont nulles. Le variogramme à palier correspond à une V.R. stationnaire et la valeur du palier est égale à la variance des données $K(0)$.

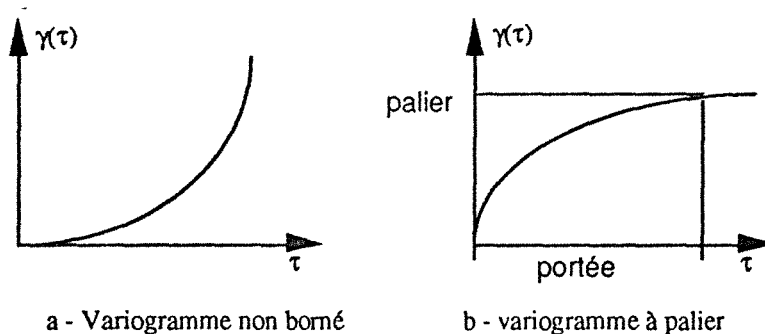


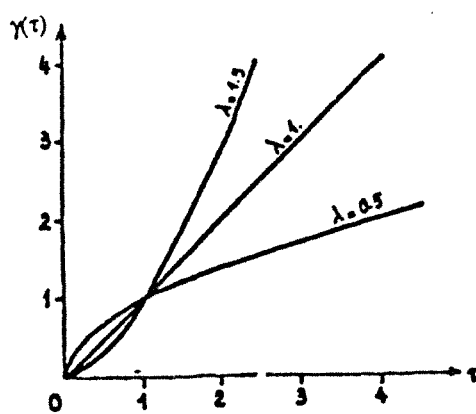
Fig. 25 - Comportement du variogramme à l'infini (extrait de [32]).

Par ailleurs, on pourrait songer à pousser plus loin l'interprétation d'un variogramme tiré des données expérimentales, en essayant d'expliquer chacune des fluctuations apparente du graphe. Si certains phénomènes présentent effectivement des périodicités, des structures gigognes ou des effets de trous (existence de corrélations négatives dues à la présence des zones "riches" entourées des zones "pauvres"). La plupart du temps, ces fluctuations sont simplement dues au fait que les différentes valeurs de $\gamma(\tau)$ calculées ne sont que des estimations, obtenues parfois avec un petit nombre de couples.

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.

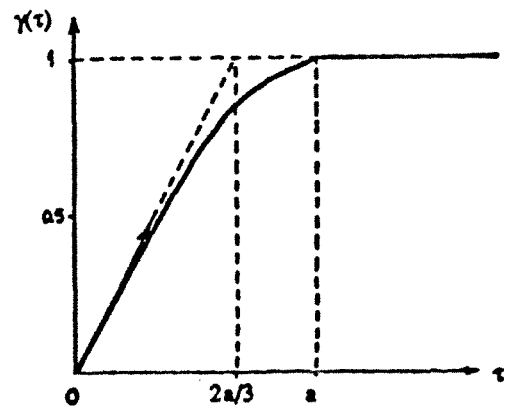
IV.1.3 : Ajustement à un modèle

On essaie d'ajuster au variogramme expérimental, un modèle théorique qui rend compte des caractéristiques observées. Ce modèle structural est une fonction simple : modèle linéaire, exponentiel, etc. Les modèles les plus fréquemment utilisés sont illustrés par la figure 26.



a - modèle en τ^λ :

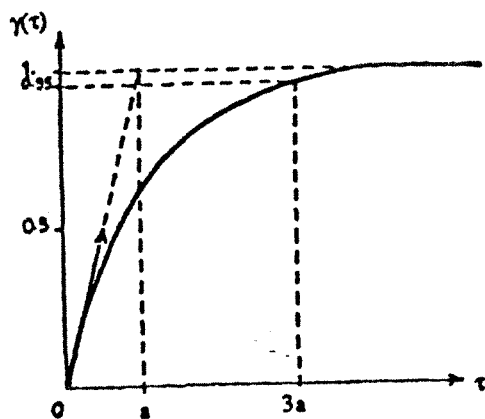
$$\gamma(\tau) = \omega |\tau|^\lambda$$



b - modèle sphérique

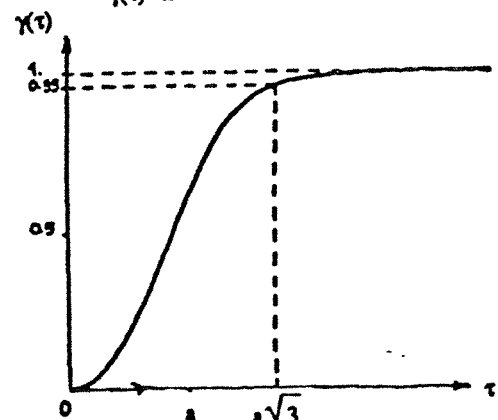
$$\gamma(\tau) = \omega \left[\frac{3}{2} \frac{|\tau|}{a} - \frac{1}{2} \left(\frac{|\tau|}{a} \right)^3 \right] \quad \tau \leq a$$

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



c - modèle exponentiel:

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



d - modèle gaussien:

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

Fig. 26 - Principaux modèles de variogrammes (Extrait de [32]).

IV.2 - Méthode de déconvolution géostatistique [73]

Cette méthode, proposée par ROYER, repose sur le calcul et la modélisation préalable des fonctions de covariance des signaux d'entrée et de sortie, puis sur la résolution de l'équation de régularisation qui relie ces deux fonctions. La résolution peut se faire numériquement par transformation de Fourier ou par méthodes quadratiques sous contraintes linéaires.

Soit $X(t)$ une fonction aléatoire stationnaire d'ordre deux, dont la moyenne est supposée identiquement nulle [cf. Chapitre 1]. La convolution de Duhamel définie par (1.12) est une convolution stochastique...

$$Y(t) = \int_0^{\infty} X(t-\tau)\phi(\tau)d\tau = \phi * X \quad (4.23)$$

... $Y(t)$ étant la régularisée de $X(t)$ à l'aide d'un opérateur ϕ appelé fonction de transfert. Il est évident que l'espérance de $Y(t)$ est également nulle.

Nous définissons ici les fonctions de covariance et de covariance croisée par :

$$\begin{aligned} K_{XX}(\tau) &= E [X(t+\tau)X(t)] \\ K_{YY}(\tau) &= E [Y(t+\tau)Y(t)] \\ K_{XY}(\tau) &= E [X(t+\tau)Y(t)] \\ K_{YX}(\tau) &= E [Y(t+\tau)X(t)] \end{aligned} \quad (4.24)$$

Il est facile de déduire une relation entre les covariances simples et les covariances croisées de X et Y , en multipliant les deux membres de (4.23) par $X(t+\tau)$ et en prenant l'espérance [72, 73] :

$$K_{YY} = \hat{\phi} * K_{YX} \quad K_{YX} = \phi * K_{XX} \quad K_{YY} = \psi * K_{XX} \quad (4.25)$$

$$\psi = \hat{\phi} * \phi \quad (4.26)$$

... où $\hat{\phi}(\tau) = \phi(-\tau)$; $\hat{\phi}(\tau)$ est appelé transposée de $\phi(\tau)$; ψ est le covariogramme transitif de $\phi(\tau)$, vérifiant l'inégalité de Schwartz [cf. Chapitre 1].

L'expression reliant les fonctions de covariance constitue la relation de régularisation.

On notera par la suite f_X et f_Y les fonctions de pondération correspondant aux covariances $K_{XX}(\tau)$ et $K_{YY}(\tau)$. Une nouvelle équation entre les fonctions de pondération de covariances et la fonction de transfert peut s'écrire :

$$f_Y = f^* f_X \quad (4.27)$$

Ainsi, la détermination de la fonction de transfert revient à résoudre le système d'équations (4.25) et (4.26), ou l'équation de convolution (4.27). Il est parfois préférable de travailler directement sur (4.27) après avoir ajusté $K_{XX}(\tau)$ et $K_{YY}(\tau)$ à un modèle théorique. Le calcul de la solution à partir de (4.25) et (4.26) ou à partir de (4.27), de préférence à (4.23), devrait être plus précis car les fonctions de covariances sont généralement moins perturbées que $X(t), Y(t)$.

La résolution numérique fait appel à la transformation de Fourier ou aux méthodes du type "*problème inverse*". Dans des cas simples où le modèle de covariance est de type exponentiel ou gaussien, il est possible de calculer la fonction de transfert sous forme analytique.

Par ailleurs, cette méthode développée dans le cadre des fonctions aléatoires stationnaires peut se généraliser aisément aux F.A. non stationnaires à l'aide de la théorie des fonctions aléatoires intrinsèques d'ordre k . En effet, il s'agit de calculer la fonction de transfert à partir des fonctions de covariances généralisées du signal d'entrée et de sortie. Le formalisme de F.A.I.k a l'avantage de filtrer la dérive (la tendance) tout en estimant la fonction de transfert. Pour plus de détails, on propose de se reporter à [73].

Il convient de noter que l'étude structurale au préalable des entrées-sorties (examen du variogramme) nous fournit de précieuses informations: zone d'influence, vérification de la stationnarité du phénomène étudié et le cas échéant découpage en morceaux localement stationnaires.

V - Conclusion

Dans ce chapitre, quatre méthodes de déconvolution ont été présentées.

La méthode de régression multiple est une méthode couramment utilisée dans de très nombreux domaines. Cette méthode peut se prêter à la déconvolution grâce à l'équivalence formelle entre la relation de régression et la relation de convolution.

La méthode MARIMA combine les concepts de régression multiple aux modèles ARIMA. En effet, elle calcule la fonction de transfert tout en considérant la structure du terme résiduel par la méthode de Box-Jenkins, ce qui représente théoriquement un avantage décisif. Cette idée de combinaison est fort intéressante.

La méthode de projection proposée par De MARSILY est une méthode principalement utilisée en hydrologie. Cette méthode a été mise en oeuvre à partir des considérations sur la stabilité des solutions physiques, et pour cela, la base de Walsh est imposée *a priori* sur la fonction de transfert. Il s'agit d'un processus itératif qui projette plusieurs fois le résidu sur les premiers vecteurs de la base. Du fait de cette itération, des contraintes ou des lissages de la FT cherchée peuvent être introduits pas à pas. La base de Walsh étant par construction hiérarchisée, on peut retenir au début que les premiers vecteurs de base extrêmement réguliers, et éviter que n'apparaissent des instabilités ou des oscillations dans la FT. Par ailleurs, cette méthode permet de prendre en compte le phénomène de tarissement, et cette démarche peut être considérée comme parallèle à la combinaison des concepts de régression multiple aux modèles ARIMA.

La méthode de déconvolution géostatistique proposée par ROYER est une méthode reposant sur le calcul et modélisation au préalable des fonctions de covariance d'entrée et de sortie du système étudié. Au lieu d'imposer une fonction de forme analytique sur la FT, cette méthode impose une étude structurale sur l'entrée et sur la sortie. Elle nous permet de retenir d'une part de précieuses informations sur les signaux, et "*filtrer*" les erreurs contenues dans les signaux ainsi que les erreurs d'estimation et rend d'autre part le calcul plus précis. Par ailleurs, cette méthode a le remarquable avantage d'être généralisé aux cas non stationnaires très souvent rencontrés dans les applications. Enfin, elle est facile à mettre en oeuvre: il suffit pour cela de disposer d'un programme de calcul de covariances et d'un programme de régression multiple sous contraintes.

Deuxième partie

**ANALYSE DES DONNES D'EXHAURE MINIERE
DANS LE BASSIN FERRIFERE LORRAIN**

Chapitre 5

S O M M A I R E

I	Bassin ferrifère lorrain - Généralités.....	72
	I.1 Situation géographique - cadre géologique.....	72
	I.2 Situation structurale.....	72
	I.3 Lithologie et stratigraphie.....	74
	I.4 Situation climatique.....	74
	I.5 Réseau hydrographique.....	77
II	Travaux effectués visant à connaître le phénomène	
	"exhaure minière".....	77
	II.1 Travaux de 1925 à 1957.....	78
	II.2 Travaux effectués de 1958 à 1969.....	81
	II.3 Travaux effectués après 1970.....	83
III	Commentaires et conclusions.....	90

CHAPITRE 5

ETUDE BIBLIOGRAPHIQUE SUR L'EXHAURE MINIERE DANS LE BASSIN FERRIFERE LORRAIN

L'exhaure minière dans le bassin ferrifère lorrain est un phénomène important. En effet, elle est en augmentation constante passant de 221 millions de m³ en 1978, à 248 millions de m³ en 1979, 266 millions de m³ en 1980 et 295 millions de m³ en 1981.

Rapporté à la production de minerai, on constate qu'il faut pomper 7 m³ d'eau par tonne de minerai produite en 1978 ; 8,2 m³/t en 1979 ; 9,6 m³/t en 1980 et 14,4 m³/t en 1981.

Les dépenses qui en résultent augmentent de façon catastrophique. En 1981, les dépenses d'exhaure (sans amortissement) s'élèvent à 49,3 millions de francs [2].

De plus, les eaux exhaurées, de bonne qualité, sont en partie exploitées pour l'alimentation en eau potable et les besoins industriels (33 millions de m³ en 1978). L'irrégularité saisonnière des débits est cependant telle qu'un stockage des eaux est nécessaire pour satisfaire les besoins des utilisateurs. C'est pour cette raison que plusieurs réservoirs souterrains ont été constitués par un noyage d'anciennes exploitations minières abandonnées. Lors du remplissage de ces réservoirs, on a constaté une minéralisation et une dégradation de la qualité des eaux qui posent un problème aux utilisateurs [39].

Les études d'exhaure minière, tant sur le plan quantitatif que sur le plan qualitatif, présentent de l'intérêt pour divers organismes : les Exploitants, le Service des Mines, l'Agence de l'Eau Rhin-Meuse, *etc.*

Depuis une quarantaine d'années, un grand nombre de travaux ont été effectués pour atteindre divers objectifs.

Dans ce chapitre, nous examinerons brièvement ce qu'apportent ces travaux à la connaissance du phénomène "exhaure". Des questions importantes seront également posées à la fin de ce chapitre.

I - BASSIN FERRIFERE LORRAIN - GENERALITES

[2, 8, 39, 42, 46, 47, 52]

I.1 - Situation géographique - cadre géologique

Le bassin ferrifère lorrain se trouve sur la bordure Est du Bassin Parisien. C'est un plateau qui se relève légèrement vers le Nord-Est. Il est formé d'une succession de couches calcaires et de marnes attaquées de façon différentielle par l'érosion. Les couches forment des auréoles successives centrées sur l'intérieur du Bassin Parisien ; les calcaires se terminent en falaises vers l'Est (fig. 27).

Le bassin ferrifère lorrain est limité à l'Ouest par la Woëvre marécageuse et se termine au Nord et à l'Est par la falaise bajocienne des plateaux de Briey et Haye.

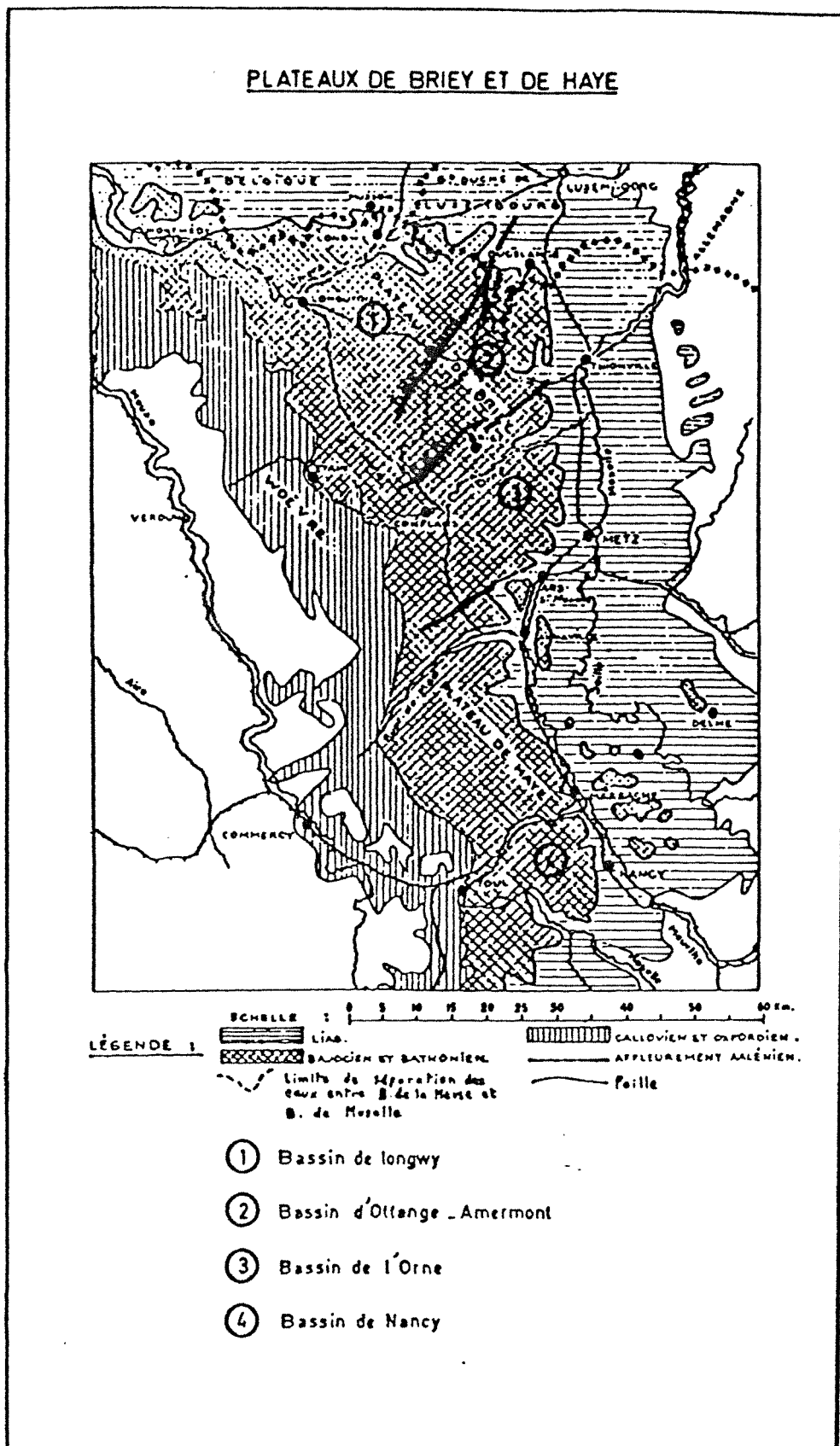
Sous ces plateaux, après 100 à 250 mètres de morts terrains, se trouve le minerai de fer. Ce dernier affleure tout le long de la falaise bajocienne sauf entre Arnaville et Marbache, zone stérile et lieu de passage de l'anticlinal principal lorrain dit aussi de "*Pont-à-Mousson*" qui divise la région minéralisée en deux : le bassin de Briey-Longwy au Nord et le bassin de Nancy au Sud.

I.2 - Situation structurale

L'ensemble du bassin ferrifère lorrain peut se diviser en quatre unités structurales :

- le bassin de Longwy s'étendant de la vallée de la Chiers jusqu'à la ligne de fractures joignant Autun-le-Tiche à Murville ;
- le bassin d'Ottange-Tucquenieux-Landres-Amermont compris entre la faille d'Autun-le-Tiche et la faille d'Avril ;
- le bassin de l'Orne allant de la faille d'Avril à la faille de Metz-Gorze ;
- le bassin de Nancy, au Sud de l'anticlinal de Pont-à-Mousson.

Fig. 27 - Plateaux de Briey et de Haye.



I.3 - Lithologie et stratigraphie

(fig. 28)

*** Formations ferrifère (15 à 65 m, 40 m en moyenne).**

La formation ferrifère, datée du Lias supérieur (Aalénien) surmonte les argiles du Toarcien. Une alternance de niveaux marneux et de calcaires ferrugineux constitue une série régressive (fig. 29).

*** Bajocien inférieur (Jurassique moyen).**

Il comprend :

- les marnes micacées (25 m) qui surmontent le toit de la formation ferrifère,
- les calcaires marneux à *Cancellophycus* (10 m),
- les calcaires du Haut-Pont (30 m).

*** Bajocien moyen (Jurassique moyen).**

- calcaires à grosses oolites et débris d'encrines, polypiers, etc. (40 à 80 m).

*** Bajocien supérieur (Jurassique moyen).**

- les Marnes de Longwy (3 à 5 m),
- l'Oolithe de Jaumont (12 à 20 m),
- les calcaires et marnes de Gravelotte (15 m),
- l'Oolithe de Vionville (0 à 15 m),
- les marnes et calcaires de Saint-Privat (10 m).

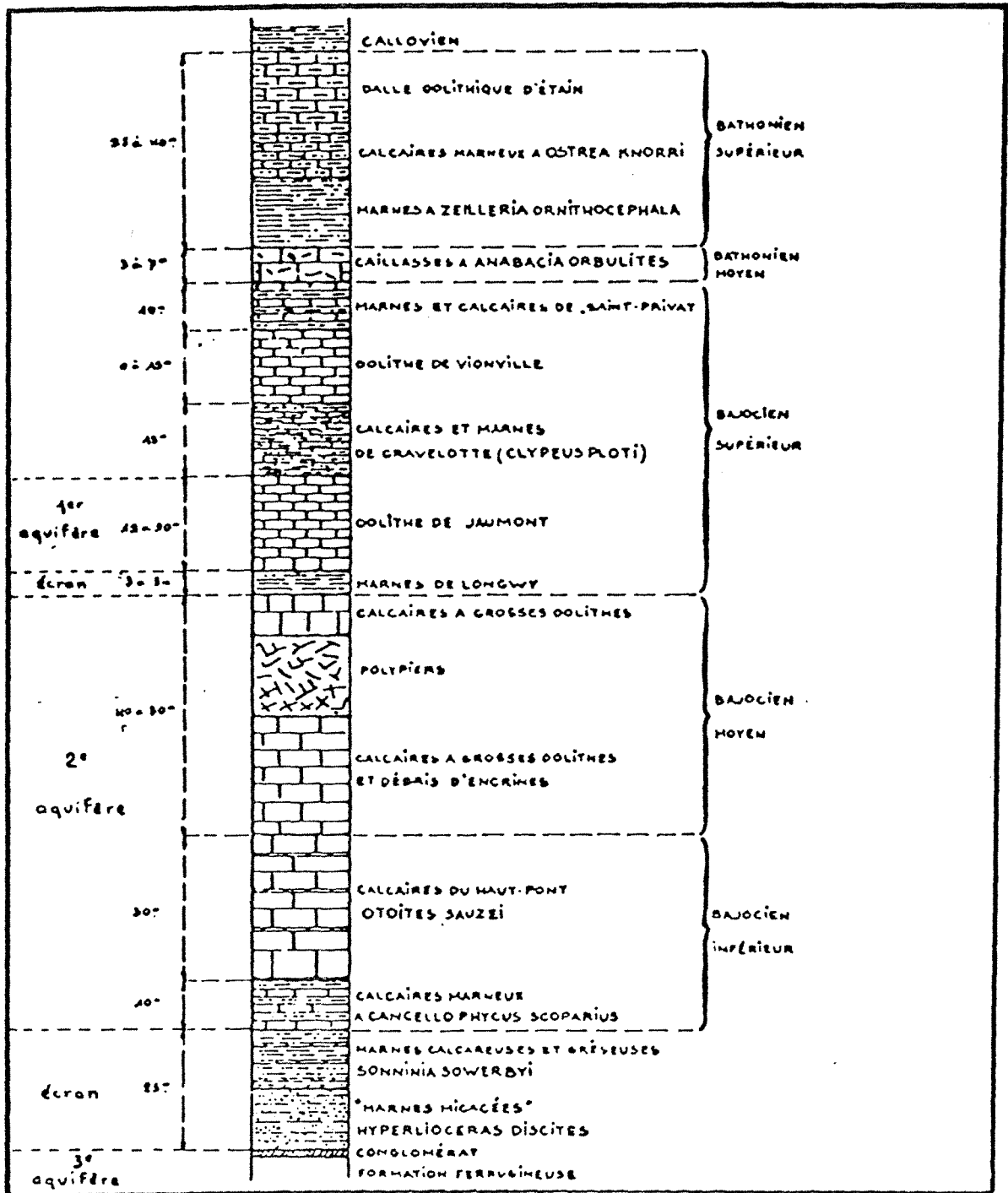
*** Bathonien moyen (Jurassique moyen).**

- les marnes à *Leilletia ornitho*,
- les calcaires marneux à *Ostra knorri*,
- la dalle oolithique d'Étain (25 à 40 m).

I.4 - Situation climatique

A 400 km de la mer vers l'Est, la Lorraine semble vouée à la continentalité. Soumis aux influences océaniques et continentales, le climat connaît une irrégularité selon que les unes ou les autres l'emportent. Mais, d'une façon générale, les températures moyennes restent modérées et les amplitudes thermiques sont moins fortes que dans les régions de climat continental type.

Fig. 28 - Coupe schématique du Bajocien et du Bathonien (Bassin de Briey et Longwy).



Bichelanne et Angeot modifié

Fig. 29 - Coupe géologique de la formation ferrifère.
(Mine d'Auboué, Moineville, Siège du Paradis, Puits 4).

Coupe (figuré)		Paragénèse
	Calcaires bajociens	
	Marnes micacées	réduite
	Calcaire ferrugineux marneux	oxydée
	Couche rouge	oxydée
	Calcaires ferrugineux et marnes	
	Calcaire gris-bleu marneux	
	Couche jaune S.2	
	Couche jaune S.3	oxydée
	Marnes gris-bleu	
	Calcaire coquillier ferrugineux	
	Couche grise	oxydée (un peu de pyrite)
	Marnes ferrugineuses	
	Marnes gris-noir	
	Couche brune	
	Marnes brunes	
	Marnes vertes oxydées	
	Couche noire	hyperréduite (pyrite, chlorite sidérose)
	Microgrès de base	

Toit de la formation ferrifère

36,5 m

Mur de la formation ferrifère

Deux saisons peuvent être différenciées en Lorraine[46, 52]:

- une saison froide, pluvieuse et peu ensoleillée, de novembre à avril, avec un minimum de températures en janvier (1,3° C) ;

- une saison chaude, ensoleillée mais peu pluvieuse, de mai à octobre où le maximum des températures est atteint en juillet (17,9° C) avec des chaleurs lourdes et fréquentes.

I.5 - Réseau hydrographique

Deux facteurs conditionnent l'écoulement des eaux de surface :

- la nature des affleurements,
- la tectonique.

Au Sud-Ouest, où affleurent des marnes imperméables, existent de nombreux petits ruisseaux drainés par l'Orne.

Au Nord-Est où les affleurements sont calcaires, l'infiltration est forte et les cours d'eau sont rares et encaissés. Ils prennent naissance soit dans la zone de transition entre marnes et calcaires, soit sur des failles (la Fensch sur la faille de Fontoy, la Crusne sur la faille de Crusne, l'Othain et le Woigot sur la faille de Mont-Bonvillers).

Les cours d'eau suivent généralement trois directions de failles, ce qui explique les coudes brusques des rivières.

L'Orne est le principal exutoire du bassin.

Dans les zones d'affleurements existent de nombreux signes de circulations karstiques (dolines, gouffres, vallées sèches).

II - TRAVAUX EFFECTUES VISANT A CONNAITRE LE PHENOMENE "EXHAURE MINIERE"

Afin de faciliter l'exposé dans ce paragraphe, nous allons diviser les travaux effectués en trois périodes :

1. les travaux de 1925 à 1957 : travaux de reconnaissance, méthode empirique de simulation de l'exhaure (proposée par CHANZY) ;
2. les travaux de 1958 à 1969 : le bilan hydrologique entre en jeu ; installation de piézomètre et approfondissement de la connaissance sur les aquifères ainsi que sur le phénomène exhaure ;
3. les travaux après 1970 : recueil et critique des données d'exhaure, analyse des données par méthodes numériques ou graphiques ; analyse de la qualité des eaux exhaurées.

II.1 - Travaux de 1925 à 1957

Depuis 1925, le bassin ferrifère lorrain est l'objet de nombreuses études.

JOLY[53] par des expériences de coloration, confirme le caractère karstique de la couverture calcaire que laisse présager l'existence de vallées sèches, de dolines, de gouffres.

CHANZY (de 1925 à 1930) est le premier qui tente de relier l'exhaure à la pluie. De ses rapports ou publications [20, 21, 22], on peut tirer des idées importantes :

- Les eaux d'exhaure d'une mine sont d'abord celles de traçage et ensuite celles de dépilage.

La première phase de l'exploitation consiste dans le traçage d'un réseau de galeries : celles-ci rencontrent des cassures innombrables, dont la direction générale est le plus souvent parallèle au fil de la mine ; mais ces cassures sont très sinueuses et d'ouverture très irrégulière : celle-ci varie très rapidement de fractions de millimètres à 20-30 cm et davantage. L'intérieur est parfois vide, parfois obstrué de blocs éboulés, parfois rempli d'un bouchon compact d'argiles. Enfin, du point de vue qui nous préoccupe, ces cassures sont parfois complètement sèches ; parfois elles suintent ; parfois elles débitent plusieurs mètres-cubes à la minute au niveau d'une simple galerie.

Au cours de l'avancement des travaux de traçage d'une mine, on produit une saignée de terrains : la venue d'eau totale augmentent d'abord assez vite pour atteindre une limite qui n'est ensuite dépassée que très lentement. D'une région à l'autre, ce débit limite varie beaucoup. Pour certaines mines, le débit des eaux de traçage s'est rapidement stabilisé à quelques centaines de litre à la minute ; pour d'autre, il a atteint environ 10 m³ à la minute, sans que l'on puisse

affirmer qu'il a ainsi atteint sa limite. Le régime de ces eaux de traçage est sensiblement constant et on ne distingue aucune loi dans leurs faibles variations.

En principe, lorsque le réseau de traçage d'une mine a découpé un nombre suffisant d'îlots, on attaque chacun de ces îlots par des chantiers laissant entre eux des piles ; ces piles sont ensuite prises à leur tour : il y a ainsi foudroyage des terrains supérieurs ; les bancs du toit tombent en gros blocs les uns après les autres et cette dislocation se propage jusqu'à la surface ; il se produit donc une quantité de fractures qui mettent la mine en communication avec les eaux supérieures : eaux souterraines et eaux de surface.

La physionomie de l'exhaure est alors radicalement bouleversée : le premier éboulement important est suivi en quelques jours d'une forte venue d'eau ; cette venue augmente à mesure que les éboulements se succèdent gagnant des couches toujours plus élevées ; si l'on cesse les travaux de défilage, on arrive ainsi en quelques mois - ou quelques années (une ou deux) à un nouveau régime.

- Origine des eaux d'exhaure : cette origine réside évidemment dans les pluies, ou plutôt dans les eaux résiduelles (eaux résiduelles = excédents = pluies - évapotranspirations [cf. annexe I]). Mais, il faut rechercher quelle est l'étendue du bassin de réception à considérer et parmi les eaux résiduelles, quelle part il faut faire à l'écoulement superficiel et quelle part à l'infiltration directe en profondeur ; enfin, si les eaux de l'écoulement superficiel sont définitivement écartées pour l'exhaure, ou si, au contraire, elles y réapparaissent par la déperdition des cours d'eau.

Relativement à l'étendue du bassin de réception, s'il ne fallait considérer que les infiltrations à l'aplomb des travaux, sur une surface sensiblement plus faible que celle des concessions exploitées, ces infiltrations suffiraient à peine à alimenter l'exhaure. Il faut donc faire appel à un bassin de réception plus vaste, mais sans aller bien loin, on trouve des eaux résiduelles en quantité largement suffisante.

Quant aux cours d'eau, il faut noter qu'on trouve de fortes exhaures dans des mines éloignées de tout cours d'eau important. De plus, lors de la sécheresse exceptionnelle de l'hiver 1920-1921, la mine d'Auboué a continué à exhaurer des cubes importants, alors que l'on était à peu près à sec et que ses infiltrations ne pouvaient certainement pas alimenter l'exhaure.

Il est donc affirmé que les eaux d'exhaure sont peu influencées par les infiltrations des cours d'eau.

Reste la possibilité de mise en jeu des réserves souterraines.

A l'époque sèche de 1920-1921, les venues de dépilage directement influencées par les pluies ont absolument disparu ; et cependant les mines ont continué à exhaurer des mètres cubes ne décroissant que très lentement, et présentant l'apport des réserves souterraines. La contribution des réserves souterraines à l'exhaure n'est pas négligeable. Les venues de traçage représentent l'écoulement dans les mines des réserves souterraines.

- Les variations de l'exhaure, dans le temps sont commandées par deux facteurs principaux : la progression des travaux qui multiplie les communications des mines avec les eaux supérieures ; le régime des pluies.

Ces deux facteurs agissent simultanément de façon très différente selon les mines considérées. Trois groupes de mines sont distingués dans le bassin de Briey :

- 1er groupe : exhaure faible et constante
- 2ème groupe : exhaure élevée et très variable en fonction des conditions climatiques ;
- 3ème groupe : exhaure en voie d'augmentation.

Cela peut signifier, en termes de séries chronologiques, qu'il existe deux types de mines vis-à-vis de leur exhaure : les mines stationnaires et les mines non stationnaires.

Pour toutes les mines, l'exhaure est influencée plus ou moins directement par les pluies. Les pluies tombant sur les calcaires engorgent les communications entre les mines et les eaux souterraines ainsi qu'entre les eaux souterraines et les eaux de surface, et mettent sous pression les orifices de débit des eaux dans les mines. Ce débit augmente avec décalage dans le temps.

A mesure que les dépilages et les foudroyages progressent, pour les mines du 3ème groupe seulement cette fois-ci, ces augmentations deviennent de plus en plus sensibles et brutales. Ce phénomène est appelé "*aggravation du tempérament de la mine*" par CHANZY.

Dans un tel cadre de connaissance, CHANZY propose une démarche qui lie l'exhaure à la pluviométrie en utilisant les eaux résiduelles. Ce résidu représente l'infiltration plus l'écoulement superficiel mais comme il ne peut les distinguer, il admet que $\frac{\text{infiltration}}{\text{écoulement superficiel}} =$ constante et ce qui permet de ne pas tenir compte de l'écoulement superficiel.

La démarche de CHANZY est mal expliquée dans ses rapports ou publications. C'est grâce aux travaux de CANCEILL présentés dans un rapport du B.R.G.M. en 1974 [13] que l'on peut comprendre les détails de cette démarche (cf. Annexe I).

En effet, cette démarche consiste à postuler que s'il existe une courbe de tarissement, elle est unique et elle doit être mise en évidence par un examen systématique de toutes les périodes sèches.

A partir de 1947, deux tendances s'affrontent :

- d'aucuns admettent les circulations karstiques comme prépondérantes et pensent qu'il faut lutter contre l'exhaure par la création de pompages dans des conduits karstiques se trouvant au-dessus des couches exploitées (gain sur la hauteur de refoulement) ;

- d'autres admettent l'existence de nappes imparfaites.

MAUBEUGE, en 1956, pense que seule la première tendance peut amener des résultats pratiques, les débits d'exhaure dus aux circulations karstiques étant jugés très supérieurs aux débits de pseudo-nappes.

Les travaux s'orientent alors vers la découverte de ruisseaux souterrains à faible profondeur. Des circulations karstiques sont mises en évidence mais aucun boyau n'est découvert. Ces recherches prouvent que les conduits karstiques sont pratiquement impossibles à localiser, sauf cas particuliers. Les moyens de prospection peuvent délimiter des zones favorables mais il est très difficile d'atteindre les conduites par forage, vu la différence de volume entre masse calcaire et conduits karstiques. Cependant, ces travaux ont toujours rencontrés des faciès de fines diaclases ou des zones très fracturées à forts débits.

II.2 - Travaux effectués de 1958 à 1969

A partir de 1958, à l'instigation du Service des Mines, le **B.R.G.M.** entreprend l'étude des ressources hydrauliques du bassin ferrifère lorrain [17, 37, 38].

Pour étudier le bilan hydrologique, le bassin ferrifère lorrain est découpé plus ou moins arbitrairement en bassins hydrologiques élémentaires appelés : bassin de Longwy, bassin de l'Orne, bassin d'Ottange-Amermont. Les conclusions suivantes ont été apportées :

- l'hypothèse de CHANZY qui rend compte de l'exhaure de la plupart des mines à l'aide d'un bassin dont la surface correspondrait au périmètre de la concession minière est vérifiée ;

- le ruissellement est peu important ;

- les courbes de tarissement de l'exhaure semblent indiquer une superposition de plusieurs régimes d'écoulement ;

- l'exhaure a atteint son maximum dans le bassin de l'Orne et le bassin d'Ottange-Amermont, tandis que l'exhaure est nettement inférieure aux prévisions dans le bassin de Longwy.

A partir de 1961, quelques piézomètres sont installés (Bazailles n° 8, 9, 10 ; Mercy-le-Haut n° 1, 2, 3, 9bis et 13 piézomètres près de Fillières). Les mesures de niveau n'ont pu être exécutées que pendant une courte période, les forages n'étant pas tubés et s'éboulant rapidement.

Cependant, il a été possible d'établir une carte piézométrique cohérente montrant une zone d'appel vers les mines de la bordure Nord du bassin de Longwy, zone d'appel ne débordant pas la superficie des concessions.

En 1968, l'A.F.B.R.M. montre que les interactions limitées entre les dépilages doivent faire considérer les réserves aquifères bajociennes comme celles d'une nappe à transmissivité moyenne et non comme celles d'un karst.

En 1969 sont installés : onze piézomètres sur la concession de Batilly, appelés R1 à R11 et trois piézomètres H119, H120 et H121 près d'Hayange.

Les piézomètres R1 à R11 surveillent simultanément jusqu'à six niveaux piézométriques différents correspondant aux couches du calcaire récifal, du calcaire de Jaumont, et de quatre niveaux des marnes de Gravelotte.

De leur étude, il se dégage que :

- les mesures des différents niveaux semblent correspondre à des surfaces piézométriques continues ;

- l'aquifère du calcaire récifal serait contenu dans des fissures peu développées (remontée très rapide de niveaux piézométriques lors des fortes pluies) ;

- tous les niveaux seraient plus ou moins en relation ;

- la partie supérieure du calcaire de Jaumont serait très karstique. En certains endroits, le calcaire de Jaumont saturé déverserait son trop-plein dans les zones dépilées par l'intermédiaire

de sa zone supérieure karstique. Un pompage effectué à l'aplomb de ces zones particulières empêcherait la zone saturée d'atteindre la partie karstifiée et l'exhaure serait réduite.

II.3 - Travaux effectués après 1970

A la demande de l'A.F.B.R.M., en 1972-1973, le B.R.G.M. effectue le recueil et la critique des valeurs des débits d'exhaure des mines de fer jusqu'en 1971 [1]. Depuis 1970, les statistiques des volumes d'exhaure sont recueillis par l'L.R.S.I.D. et font l'objet d'une publication annuelle.

L'ensemble des données mensuelles et annuelles sont stockées sur support informatique à l'Agence de l'Eau Rhin-Meuse.

CANCEILL et FEUGA en 1972 [12], font leur première étude statistique sur l'ensemble des données d'exhaure dans le bassin de l'Orne. Le principe adopté est celui d'une analyse en composantes principales. Les données mensuelles sur 25 ans de 23 mines ont été traitées. Les principaux résultats retenus sont d'une part, la mise en évidence de corrélations parfois très fortes entre mines, dont certaines se groupent autour d'une "exhaure-type", alors que d'autres conservent un comportement assez indépendant ; d'autre part, la liaison floue mais indéniable entre pluie efficace et exhaure dans l'ensemble.

LE NIR, en 1973 [53], marque une étape dans les études hydrogéologiques de l'exhaure en résumant la chronologie des travaux antérieurs et en approfondissant deux points particuliers :

- la relation exhaure-piézométrie ;
- le tarissement de l'exhaure en période sèche.

L'étude de la relation exhaure-piézométrie a pour cadre la connaissance hydrologique du bassin ferrifère lorrain. Elle a pour but d'explicitier les relations pouvant exister entre piézométrie et exhaure.

Les seuls piézomètres qui possèdent des relevés suffisamment denses sur une période assez longue sont les piézomètres de Roncourt notés R1 à R11. Par conséquent, cette étude se limite à une seule mine : la mine de Roncourt.

De cette étude, plusieurs conclusions ont été tirées parmi lesquelles on pourra citer :

- les niveaux calcaires de Jaumont et calcaires récifaux ne constituent pas deux aquifères séparés. Généralement, la hauteur piézométrique de cet aquifère unique reste au-dessous des marnes de Longwy. Il suffit donc d'étudier le réservoir constitué par les calcaires récifaux.

Le calcaire de Jaumont n'est qu'une zone de passage des eaux vers les calcaires récifaux. Ce cheminement des eaux est ralenti par les marnes de Longwy ainsi que par la présence épisodique de niveaux marno-calcaires à la base du calcaire de Jaumont qui donne lieu à la formation de petits aquifères suspendus ;

- à certaines époques, les calcaires récifaux sont saturés et le calcaire de Jaumont se sature à son tour. A ce moment, les variations très brutales de la hauteur piézométrique montrent que les vides sont peu importants dans le calcaire de Jaumont ;

- la surveillance piézométrique des calcaires récifaux constitue un très bon baromètre de l'exhaure, constatation sans application pratique immédiate ;

- l'exhaure des mines est une ressource importante d'eau de bonne qualité.

Les débits utilisables dépendent des débits des périodes d'étiage, mais, en l'absence de connaissances suffisantes sur les caractéristiques hydrauliques et sur la réalimentation de l'aquifère des calcaires récifaux ainsi que sur ses réserves. La surveillance piézométrique des calcaires récifaux est le seul moyen de surveiller l'exploitation des réserves d'eau. Pour cela, des mesures en période d'étiage sur les piézomètres R1, R2, R3 et R11 sont suffisantes.

- la surveillance simultanée de plusieurs piézomètres est nécessaire à l'élimination de valeurs aberrantes dues à des changements de comportement par colmatage ou influence de travaux miniers ;

- les réserves de l'aquifère des calcaires récifaux peut s'étudier de deux façons : par le bas (étude de l'exhaure) et par le haut (forages). Parmi les questions non résolues, on remarque :

- . le degré d'anisotropie en grand des terrains,
- . l'importance des circulations karstiques par rapport aux circulations de type nappe

- le degré d'anisotropie peut être évalué rapidement par des essais de type Lefranc sur des forages d'exploration minière ;

- l'étude des courbes de tarissement de l'exhaure devrait résoudre le second point.

L'essai d'interprétation de courbes de tarissement de l'exhaure à Roncourt (*cf.* annexe II) s'appuie sur trois hypothèses :

- deux types d'écoulement coexistent dans les calcaires : un écoulement karstique de type turbulent ; un écoulement laminaire dans les petites et moyennes fissures qui jouent le rôle d'un "volant régulateur". Les zones d'échange entre ces deux types d'écoulement sont localisées ;

- la répartition des petites et moyennes fissures est sensiblement homogène ;

- les courbes de décrue peuvent se décomposer, en représentation logarithmique, en deux droites représentant la vidange de fissures et l'écoulement des chenaux et chaque droite suit la loi de Maillet :

$$Q_t = Q_0 e^{-\alpha t}$$

Remarque: un écoulement turbulent ne suit pas la loi de Maillet, LE NIR a commis une erreur en employant le mot "turbulent".

Cet essai d'interprétation aboutit aux conclusions suivantes :

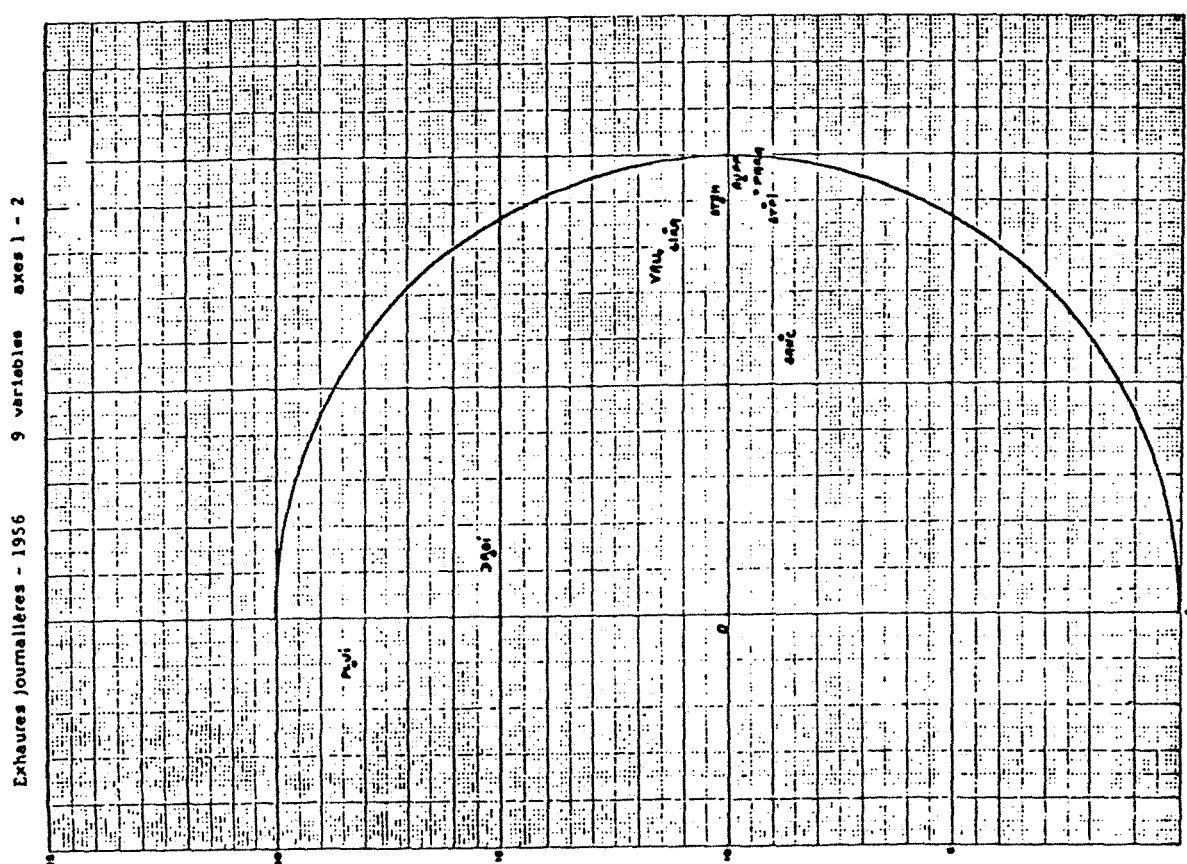
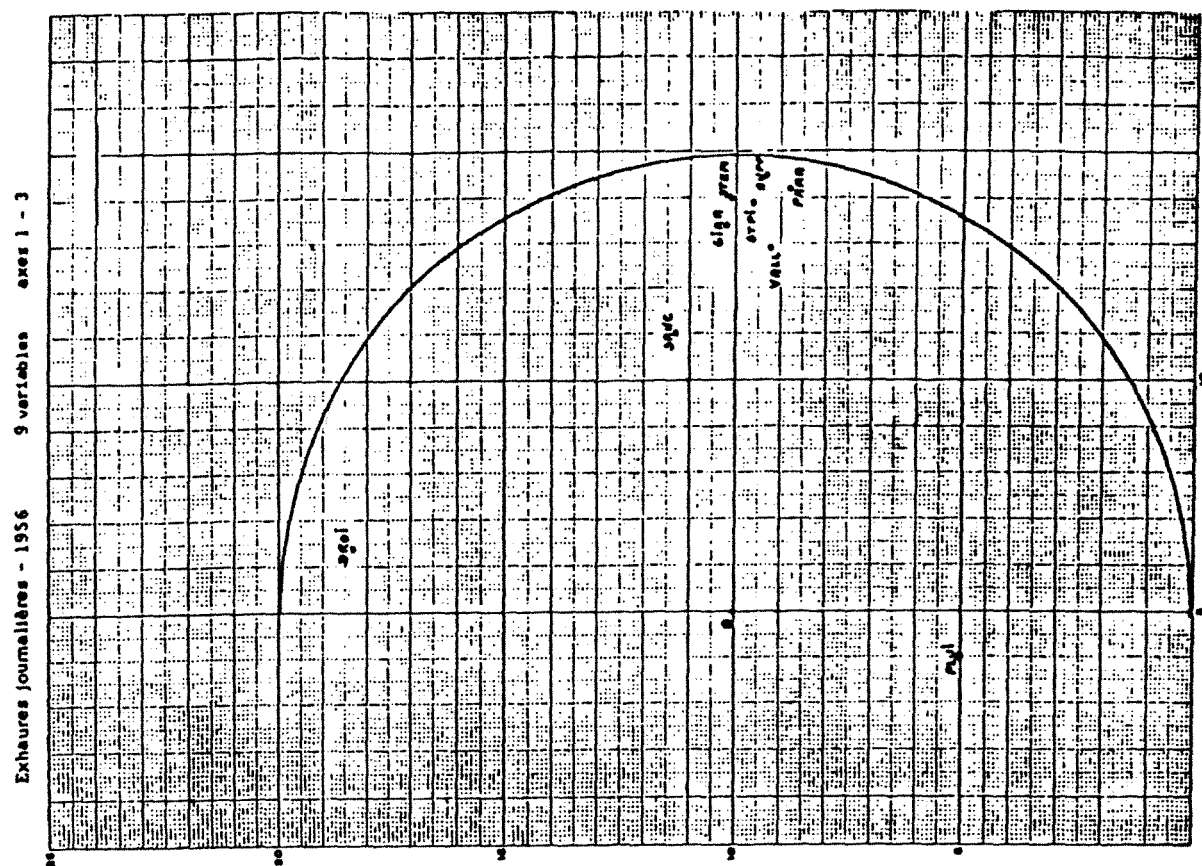
- dans les calcaires bajociens, l'écoulement karstique est un phénomène à caractère exceptionnel et local. Le volume d'eau intéressé par ce type d'eau est faible, même si ces manifestations sont spectaculaires ;

- le véritable réservoir aquifère est le réseau formé par les joints (calcaire relativement argileux) et les petites et moyennes fissures, réparti de façon homogène dans le massif à l'échelle des travaux miniers. Dans ce réseau, la loi de Darcy est applicable.

- le type de variation des niveaux piézométriques sur la zone étudiée confirme les conclusions précédentes.

CANCEILL, en 1974, a publié à nouveau un rapport [13] dans lequel deux méthodes de simulation numérique de l'exhaure (dans le bassin de l'Orne) sont étudiées : la démarche de CHANZY et la méthode de déconvolution. Il considère que le triplet Valleroy/Saint-Pierremont/Roncourt peut représenter l'ensemble des mines dans le bassin de l'Orne, car sur les graphiques d'A.C.P. (Fig. 30), on peut constater que les séries de pluies sont quasi-orthogonales aux séries d'exhaure, et les mines (dans le bassin l'Orne) sont en général très liées, à l'exception de Droitaumont. Par conséquent, la pluie efficace calculée au poste de Saint-Pierremont a été choisie comme l'entrée et l'exhaure de la mine de Valleroy comme la sortie. Ainsi, la mine de Droitaumont est considérée comme un cas exceptionnel et n'entre pas en ligne de compte.

Fig. 30 - Analyse en composantes principales de l'exhaure des mines dans le bassin de l'Orne (extrait de [13]).



Un essai d'ajustement d'une courbe de tarissement selon la démarche de CHANZY montre que les courbes ne sont pas strictement décroissantes ; la mise en coïncidence par translation horizontale de deux points de même ordonnée correspond à deux périodes sèches différentes n'entraîne pas une superposition de tronçons de courbe (fig. 31).

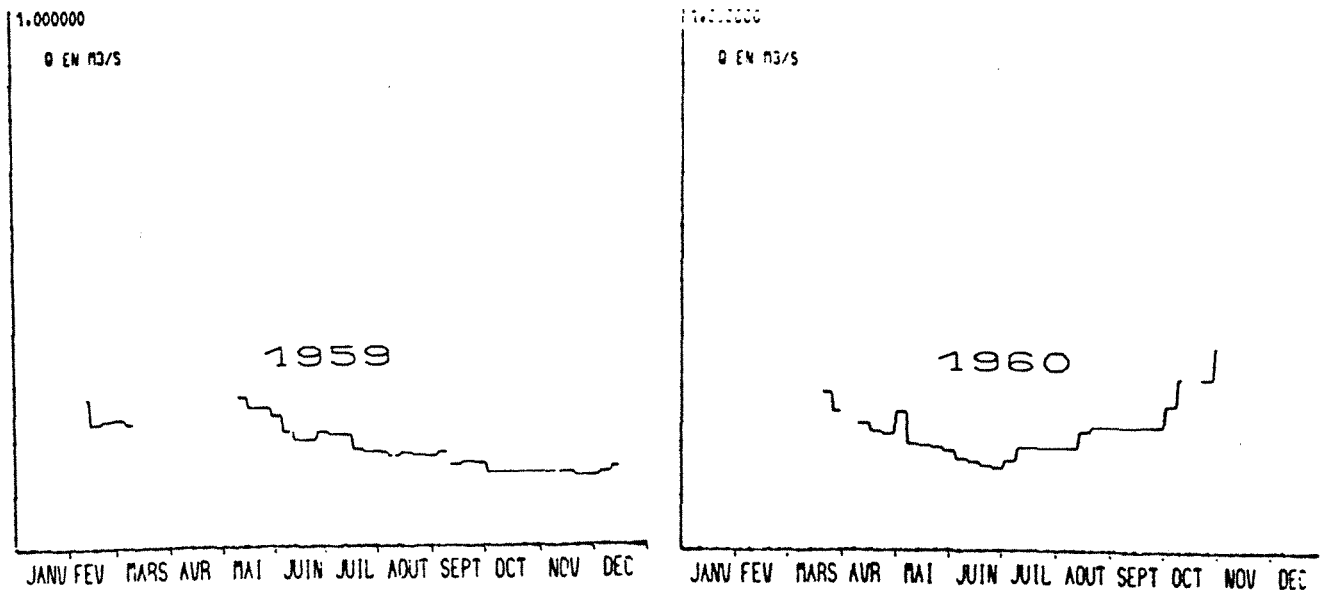


Fig. 31 - Exemples des courbes de tarissement (extrait de [13]).

La notion de courbe de tarissement ne semble donc pas correspondre à une réalité hydrologique.

Cette incompatibilité à la réalité, de la postulation de la démarche de CHANZY, conduit à un autre type de modélisation : la déconvolution.

La méthode de déconvolution utilisée est une version plus simple par rapport à la méthode mise au point au Centre d'Informatique Géologique de l'Ecole des Mines de Paris, dont la mise au point résulte d'un travail méthodologique fait au Service Géologique National en liaison avec l'équipe d'analyse numérique et traitement du signal du Groupe de Recherches Ionosphériques (G.R.I.) du C.N.R.S.

La déconvolution se réalise au niveau hebdomadaire et la réponse impulsionnelle obtenue est illustrée par la figure 32.

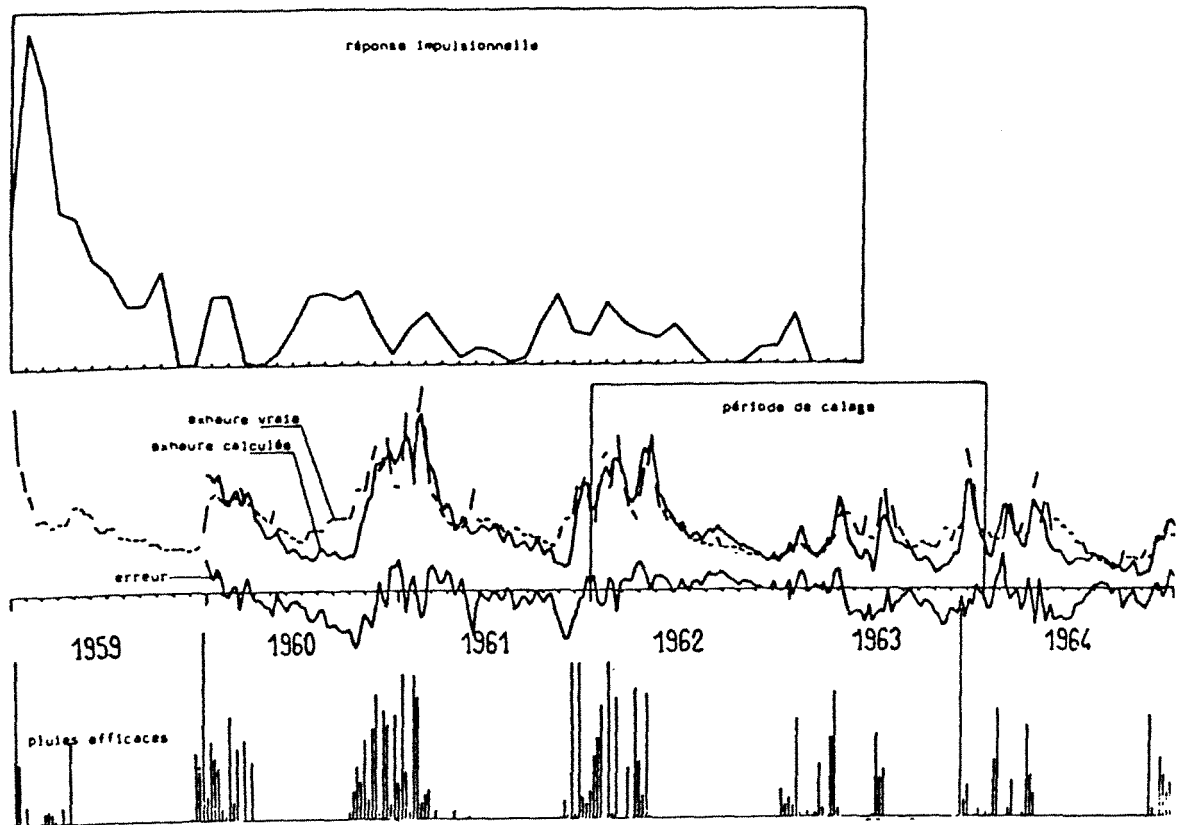


Fig. 32 - Réponse impulsionnelle et reconstitution de l'exhaure pour la mine de Valleroy de 1959 à 1964 (extrait de [13]).

La réponse impulsionnelle, dont la forme est à peu près stable au cours de différentes translations de la période de calage, peut donc être considérée comme une caractéristique du système ; sans aller jusqu'à l'interpréter comme un hydrogramme unitaire au sens propre du terme, CANCEILL pense que l'on peut cependant associer le premier pic de cette courbe à un écoulement rapide (étalé sur environ deux mois), la partie droite de la courbe pouvant représenter la composante lente de l'écoulement.

CANCEILL a conclu finalement que :

- la relation pluie-débit doit porter sur des pluies efficaces ;
- malgré tout son intérêt, la méthode de CHANZY n'est pas applicable à la représentation de la relation pluie-exhaure, principalement à cause de son manque de rigueur et d'objectivité dans l'élaboration des courbes de tarissement, mais aussi parce que le phénomène même du tarissement ne paraît pas pouvoir être isolé ;
- les méthodes dites de "déconvolution", au contraire, permettent de rendre compte correctement de la relation pluie-exhaure, en ce sens qu'elles mettent en évidence une réponse

impulsionnelle décrivant la modulation dans le temps de l'écoulement (et donnant ainsi des idées sur les différents régimes qui se superposent) ; elles permettent de reconstituer une chronique d'exhaure à partir d'une chronique de pluie avec une marge d'erreur acceptable.

Sur le plan qualitatif des eaux exhaurées des mines, les principaux travaux référés sont ceux de HERVE et ceux de DAGALLIER et DEMASSIEUX.

Les eaux exhaurées dans les mines de fer, de bonne qualité, sont en partie exploitées pour l'alimentation en eau potable et les besoins industriels. L'irrégularité saisonnière des débits est cependant telle qu'un stockage des eaux est nécessaire pour satisfaire les besoins des utilisateurs. C'est pourquoi, plusieurs réservoirs souterrains de grande dimension (0,1 à 6 Mm³) ont été constitués par un noyage d'anciennes exploitations minières abandonnées. Lors du remplissage de ces réservoirs, on a constaté que la qualité des eaux se dégrade à la suite d'une élévation de la dureté (jusqu'à 100° F) et de la teneur en sulfates (jusqu'à 1 000 mg/l de SO₄²⁻). De ce fait, ces eaux ne sont pas utilisables par l'industrie et, bien que les teneurs en sulfates obtenues ne présentent pas de danger pour les utilisateurs, il n'est pas souhaitable de les distribuer (norme conseillée : 250 mg/l).

En vue de rechercher le moyen d'améliorer la qualité des eaux stockées dans les mines, l'Agence de l'Eau Rhin-Meuse, en liaison avec le Ministère de l'Environnement, a confié l'étude de ce phénomène à l'E.N.S.G.

HERVE, dans sa thèse soutenue en 1980 [39] présente son travail réalisé dans les laboratoires de l'E.N.S.G. Deux points intéressants peuvent être tirés :

- l'oxydation à l'air humide de la pyrite, disséminée dans les roches marneuses, produit des sels sulfates solubles ; ceux-ci sont dissous à l'occasion de l'immersion des roches ;

- comme il n'est pas possible de contrôler cette réaction de sulfatation des roches, il semble qu'il faille diluer progressivement les eaux des réservoirs par des eaux d'apport peu minéralisées pour obtenir une qualité d'eau satisfaisante.

Au cours de son travail et un peu par hasard, HERVE fait apparaître que les eaux pouvaient également se minéraliser en sodium. L'A.F.B.R.M. a confié à l'E.N.S.G. une étude complémentaire visant à préciser les causes de cette minéralisation sodique et d'évaluer le risque d'acquisition d'une telle minéralisation après fermeture et ennoyage des sièges encore en exploitation.

DAGALLIER et **DEMASSIEUX** [26], en 1986, présentent dans un rapport, leurs arguments sur l'origine du sodium et du chlore. Le report des teneurs en Na par siège sur une carte structurale du bassin (fig. 33) leur a permis de remarquer que :

- ces teneurs augmentent généralement du Nord-Est vers le Sud-Ouest et du Nord-Ouest vers le Sud-Est ; ces deux tendances recourent en fait la même observation, à savoir que dans ces deux directions, la profondeur moyenne des sièges augmentent ; autrement dit, la teneur en Na est liée à l'approfondissement des sièges et donc directement à l'aval pendage ou à l'aval écoulement général des eaux ;

- il existe néanmoins des anomalies qui s'expliquent en partie par des conditions structurales : par exemple, le siège d'Errouville est plus minéralisé que ces voisins immédiats car il est en position synclinale bloquée, par la faille d'Audun-le-Tiche.

Ils ont enfin proposé un classement des sièges en fonction du risque de minéralisation sodique (*cf.* tableau I).

III - COMMENTAIRES ET CONCLUSIONS

CHANZY a contribué remarquablement aux études du phénomène exhaure par ses observations dans les mines, ses analyses, ainsi que son initiative de relier l'exhaure à la pluviométrie. On s'inspire beaucoup de sa méthode de simulation. Cette méthode postule que s'il existe une courbe de tarissement, elle est unique, et qu'elle peut être mise en évidence par un examen systématique de toutes les périodes sèches. Notons également que CHANZY est la seule personne qui met en jeu l'influence de l'exploitation sur l'exhaure en rendant compte de l'exhaure de la plupart des mines à l'aide d'un bassin versant dont la surface correspondrait au moins au périmètre de la concession minière considérée.

L'essai d'interprétation de courbes de tarissement, réalisé par LE NIR, nous semble avoir apporté des résultats intéressants qui nous aident à mieux connaître le régime de circulation hydraulique dans les morts-terrains. Cet essai s'appuie sur trois hypothèses dont l'essentiel correspond à la postulation de CHANZY.

Fig. 33 - Report des teneurs en sodium par siège sur la carte structurale du bassin (d'après DEMASSIEUX).

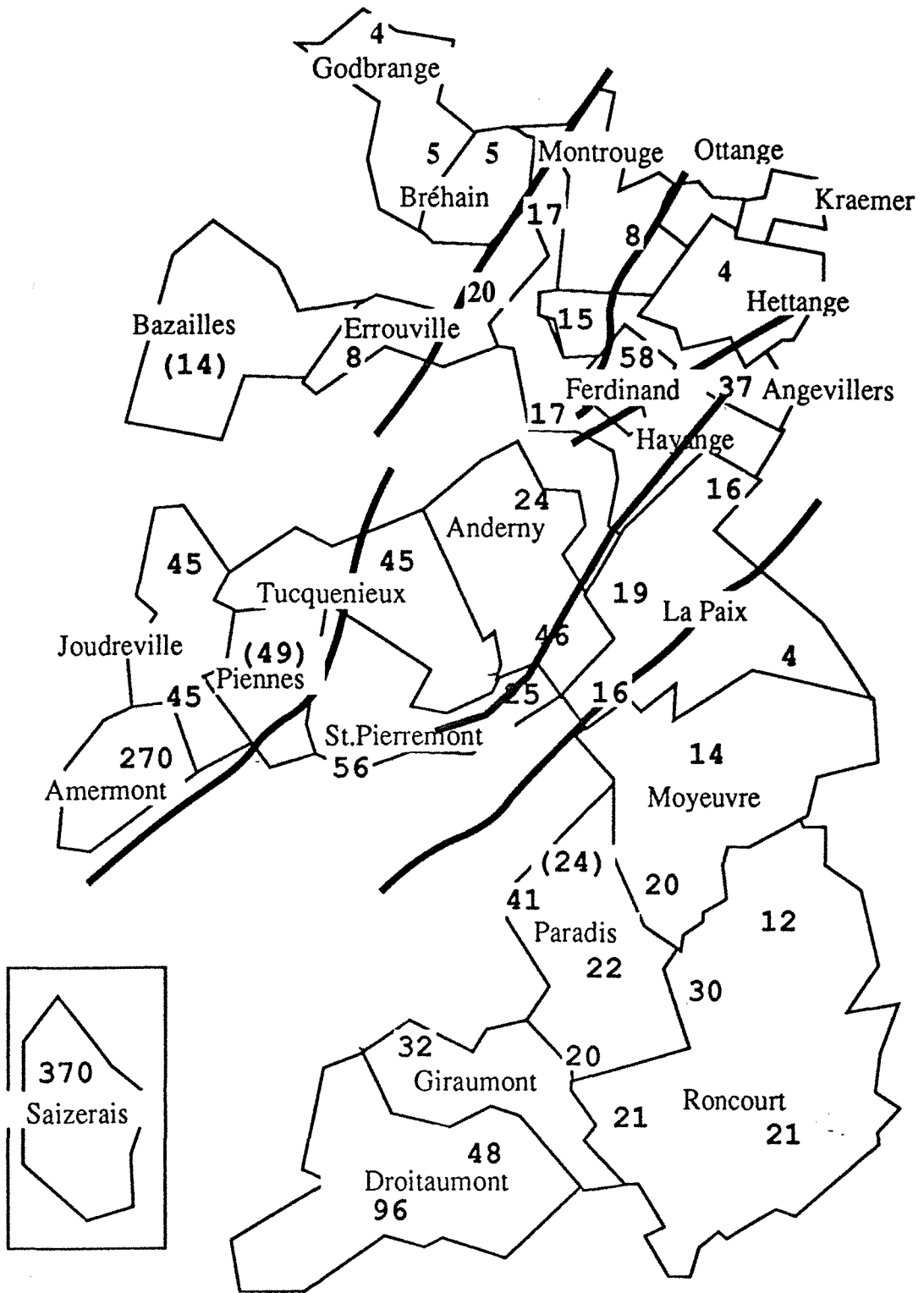


Tableau I - Classement des sièges en fonction du risque de minéralisation sodique (extrait de [26]).

1. Sièges à risques faibles ou nuls :				Teneur en Na (mg/l) observée depuis l'enneigement
Couche	Affleurement	Structure		
COBRANGE	R-V	A	a	10
TERCELET				
HAYANGE Sud	G-R	A	a	61
BREHAIN	R(?)	A	a	
ALNIETZ	R	A	s	
MONTROUGE	Toutes	A	a	
HETTANGE	G	A	a	
SERROUVILLE	R	C	a	
2. Sièges à risques modérés :				Teneur en Na (mg/l) observée depuis l'enneigement
Couche	Affleurement	Structure		
ANGEVILLERS	G-(R)	A	F	125
BURBACH	G	A	F	
HAYANGE NORD	G-R	A	s, s	84
MOYEVRE	G-R	A	f	
ORNE	G-R	A	a	
TRESSANGE	G-R	C	a	
STE ANNE	G-R	A	a	
VALLEROY	G-R	C	a	41
BAZAILLES	V	F	a	67
3. Sièges à risques certains :				
Couche	Affleurement	Structure		
ERROUVILLE	G	A	a	
FERDINAND	G-R	C	a	
LA PAIX	R	C	s, f	
SANCY	G	C	f	
CONROY	G-R	A	f	
PARADIS	G-R	A	f	
RONCOURT	G-N	A	a	
AUBOUE	G	C	f	
SANCY	G	C	f	
ANDERNY	G	C	f	
4. Sièges à forts risques (* : eau de traçage) :				Teneur en Na (mg/l) observée depuis l'enneigement
Couche	Affleurement	Structure		
TUCQUENIEUX	G-R	F	f, s	120*
ST PIERREMONT	G-R	F	f	
GERAUMONT	G-R-N	C	f	
JARNY	G-R	F	f	
DROTAUMONT	G-R	F	f	122*
PIENVE	G	F	f	
LA MOURTIERE	N	F	f	
LOUDREVILLE	N	F	a	
ANIERMONT	G	F	a	270*
SAIZERAIS	G	F-C	f, s	450

* Nature des couches exploitées : dans l'ordre de risque croissant :

- . Couche rouge ou brune (R)
- . Couche noire ou verte (N ou V)
- . Couche grise (G)

* Proximité des affleurements : dans l'ordre à risque croissant :

- . En périphérie du bassin à l'affleurement (A)
- . Au centre du bassin (C)
- . En aval pendage du bassin (F)

* Nature des accidents structuraux : dans l'ordre à risque croissant :

- . En position anticlinale ou monoclinale (a)
- . Présence de faille (f)
- . En position synclinale (s)

Nous proposons ci-dessous la classification suivante :

CANCEILL et FEUGA, en 1972, ont fait leur première étude statistique sur l'ensemble des données d'exhaure dans le bassin de l'Orne. La méthode employée est l'analyse en composantes principales. Bien que cette méthode ne puisse apporter des résultats immédiatement exploitables dans un but de prévision, elle s'avère très efficace pour tirer des informations intéressantes d'un grand tableau de données.

Suite aux travaux de CANCEILL et FEUGA, les travaux de CANCEILL en 1974 ouvrent une nouvelle voie.

La démarche de CHANZY a été d'abord mise en cause, car la notion de courbe ne semble pas correspondre à une réalité hydrologique. Ce qui conduit à un autre type de modélisation : la déconvolution.

La déconvolution s'est justifiée bien adéquate à rendre compte de la relation pluie-exhaure.

Mais cette étude se limite au cas "*exhaure-type*" (exhaure stationnaire) dans le bassin de l'Orne.

Sur le plan qualitatif, les causes de minéralisation ont été précisées par HERVE, puis par DAGALLIER et DEMASSIEUX. Un classement des sièges en fonction du risque de minéralisation sodique a été également proposé. Nous pensons que notre connaissance sur ce point est assez complète et systématique.

En conclusion, des remarques importantes peuvent se dégager des travaux effectués depuis une quarantaine d'années dans le bassin ferrifère lorrain :

- les calcaires du Bajocien forment un réservoir unique comportant une zone retardatrice des écoulements verticaux (les marnes de Longwy) et dans lequel l'écoulement karstique est un phénomène à caractère exceptionnel et local. Ce réservoir est le réseau formé par les joints, les petites et moyennes fissures, réparties de façon relativement homogène dans le massif à l'échelle des travaux miniers et la loi de Darcy y est donc applicable ;

- lors de la phase de traçage, la formation ferrugineuse reste un petit aquifère indépendant. Après dépilage par contre, la rupture de l'écran des marnes micacées met la mine en communication avec les eaux supérieures : eaux dans les calcaires du Bajocien et eaux de surface. Il y a donc, après une pluie, une augmentation de la hauteur d'eau, et le débit d'exhaure s'élève ; petit à petit, les eaux baissent et l'exhaure diminue pour reprendre successivement ses anciennes valeurs ;

- plusieurs facteurs commandent le débit d'exhaure d'une mine ; il faut d'abord prendre en compte la pluie qui alimente les aquifères souterrains. La relation pluie-exhaure peut être représentée correctement par une relation de convolution dans un cas d'"*exhaure-type*". Ensuite, l'exploitation minière peut éventuellement causer une modification du "*tempérament*" de la mine, influençant ainsi le débit d'exhaure. Le rôle régulateur des réserves souterraines, enfin, n'est pas négligeable ;

- sur le plan qualitatif, la teneur en sodium des eaux exhaurées est liée à l'approfondissement des mines et les anomalies existantes peuvent s'expliquer en partie par des conditions structurales.

Dès lors, des questions se soulèvent spontanément :

- bien que l'exhaure soit stationnaire pour certaines mines, elle ne l'est pas pour d'autres. La non stationnarité semble être provoquée par l'exploitation minière. Une fois, pour les mines stationnaires, la relation de déconvolution acceptée, comment peut-on introduire le facteur "*exploitation*" dans la convolution pour les mines non stationnaires ?

- peut-on expliquer l'hétérogénéité spatiale de l'exhaure par les conditions géologiques locales (cette hétérogénéité est négligée par CANCEILL) ?

- si la réponse à la question précédente est positive, existe-il des liaisons entre cette explication et le classement des sièges en fonction du risque de minéralisation sodique proposé par DAGALLIER et DEMASSIEUX ?

- Comment exploiter la banque des données dans un but de réaménagement du bassin ?

Ces questions s'avèrent très importantes pour l'aménagement du bassin. La recherche des réponses présente de l'intérêt à la fois méthodologique et pratique.

Nous nous proposons de tenter de répondre à ces questions dans le sixième chapitre.

Chapitre 6

S O M M A I R E

I	Données et Leur Traitement Préalable.....	96
	I.1 Acquisition et commentaires des données	96
	I.2 Calcul des pluies efficaces.....	97
II	Relation Exhaure-Pluie Efficace, Extraction de Minerai et Réserves Souterraines.....	100
	II.1 Pluies efficaces, réserves souterraines et exhaure	100
	II.1.1 : Analyse des pluies efficaces.....	100
	II.1.2 : Analyse d'exhaure minière	105
	II.1.3 : Relation exhaure minière-réserves souterraines et pluie efficace.....	105
	II.2 Effet de l'extraction de minerai sur l'exhaure	110
	II.3 Simulation de l'exhaure-modèle physique.....	112
III	Modélisation Numérique	113
	III.1 Mines stationnaires	113
	III.2 Mines non stationnaires.....	114
	III.3 Discussion sur les termes des modèles établis	127
	III.3.1 : Coefficient de tempérament β ($0 \leq \beta \leq 1$).....	127
	III.3.2 : Terme constant C, mémoire m et fonction de transfert.....	128
	III.3.3 : Terme résiduel e_t	131
IV	Comparaison de Différentes Méthodes de Déconvolution	135
V	Interprétation des Résultats et Conclusion	139

CHAPITRE 6

ANALYSE DES DONNEES D'EXHAURE MINIERE DANS LE BASSIN FERRIFERE LORRAIN

Soulignons que les questions posées dans le chapitre précédent sont liées à la non stationnarité temporelle ainsi qu'à l'hétérogénéité spatiale de l'exhaure. La non stationnarité peut être expliquée comme une conséquence de l'exploitation minière. Quant à l'hétérogénéité spatiale, elle peut être expliquée par les conditions géologiques locales.

Afin de décrire et de quantifier les variations de l'exhaure dans le temps et dans l'espace, nous proposons, dans ce chapitre, de simuler l'exhaure à partir des données météorologiques et des données d'extraction de minerai. Nous allons travailler à l'échelle mensuelle et cela, pour deux raisons: les données mensuelles sont plus complètes par rapport aux données journalières; possédant des données d'exhaure pendant une quarantaine d'années, une analyse à long terme doit se porter notamment sur les données mensuelles.

La simulation numérique est en fait un problème de déconvolution. Les méthodes utilisées sont celles présentées dans la première partie de ce mémoire, à savoir: les méthodes de décomposition, régression multiple, méthode de Box-Jenkins et la méthode de déconvolution par projection.

Dans le premier paragraphe, nous résumons les données disponibles pour cette étude. Il est toujours indispensable de faire un traitement préalable et de formuler des critiques sur les données brutes.

Dans le deuxième paragraphe, nous allons effectuer une analyse descriptive ou sémi-quantitative sur la relation exhaure-pluie efficace, extraction de minerai et réserves souterraines afin d'aboutir à un modèle physique.

La modélisation numérique présentée dans le troisième paragraphe constitue le noyau de ce chapitre important.

Les différentes méthodes de déconvolution sont également comparées dans le quatrième paragraphe. Enfin, dans le cinquième paragraphe, nous tentons d'interpréter les résultats obtenus dans un contexte géologique.

I - DONNEES ET LEUR TRAITEMENT PREALABLE

I.1 - Acquisition et commentaires des données

L'Agence Financière du Bassin Rhin-Meuse nous a communiqué les données d'exhaure et les données météorologiques.

Les données d'exhaure mensuelles (tableau II) sont complètes sauf quelques cas exceptionnels. Ces données sont "gérées" par l'A.F.B.R.M.. Un certain nombre de regroupements de mines ont été effectués par rapport aux données de base, afin de tenir compte des circulations aquifères et des regroupements "intermines" existants.

Les données météorologiques mensuelles nous semblent beaucoup moins régulières, notamment dans l'espace (tableau III). Nous n'avons que deux postes (Frescaty à Metz et Tomblaine à Nancy) où les mesures sont complètes et s'étendent de 1945 à 1985.

Les données annuelles d'extraction de minerai ont été acquises au Service des Mines (par l'intermédiaire de l'A.F.B.R.M.). Ces données sont très "disséminées". Une opération de regroupements a été faite pour que ces données soient correspondantes aux données d'exhaure. Les données dégagées de ces regroupements sont également présentées dans le tableau II. On note que les données d'extraction sont manquantes pour cinq mines (Ferdinant, Hettange, Hayange, Paradis, Roncourt).

Mine ou groupe de mines	exhaure mensuelle	extraction annuelle	Mine ou groupe de mines	exhaure mensuelle	extraction annuelle
Bazailles	67-82	51-81	Joudreville	48-86(*51)	51-84
Godbrange, Tiercelet	46-85	51-78	Amermont	46-85	51-85
Kraémer	48-82	51-77(A.P 69-71)	Piennes-Landres	66-85	51-85
Ottange	48-85	51-68	Pierremont, Mairy	46-86	51-85
Bréhain	56-85	51-80	Anderny, Sancy	49-86	51-83
Errouville, Serrouville	48-86	51-85	Hayange, Burbach	46-86	
Montrouge	48-86	51-85	Moyeuve, Joeuf	46-86	51-85
Hettange	67-86		Paradis, Valleroy	46-86(*47-49,62-67)	
Angevillers, Trassange	65-86	51-79	Roncourt, Ome	49-86	
Ferdinant	64-86		Giraumont	51-86	51-78
La Paix, Bassompierre	49-86	51-83	Droitaumont, Jamy	46-86	51-85
Tucquenieux	46-86(*52-53)	51-85	Saizerais	58-80	57-81

40-85: de 1940 à 1985 * : manque des données A.P: arrêt provisoire

Tableau II - Les données d'exhaure et d'extraction de minerai

Station	Pluviométrie mensuelle	Température mensuelle	Température maxima et minima
Frescaty(Metz)	1941-1985	1945-1985	
Tréssage Bure	1965-1985	1970-1985	
Tomblaine(Nancy)	1941-1985	1945-1985	
Sérémange-Errange	1946-1982		1959-1982
Mine Moineville	1941-1975		
Toul	1941-1983		

Tableau III - Les données météorologiques

Pour rendre l'exposé plus clair, nous présentons d'abord une carte de mines (Fig.34) et une coupe géologique (Fig.35).

I.2 - Calcul des pluies efficaces

Plusieurs travaux antérieurs [13, 20, 60] ont montré que la relation pluie-débit (exhaure) doit porter sur la pluie efficace que l'on calcule à l'aide du bilan [cf. Annexe III] :

$$P = EC + ETR + \Delta R + E \quad (6.1)$$

... où P : Pluie brute (en mm)

EC : écoulement superficiel (en mm)

ETR : évapotranspiration réelle (en mm)

ΔR : variation de RFU (réserve facilement utilisable, en mm)

E : pluie efficace au sens stricte ou infiltration(en mm). C'est la partie qui alimente les nappes souterraines

E+EC : pluie efficace au sens large(en mm)

RFU_{max} (réserve facilement utilisable maximale). Dans la région lorraine, elle est supposée égale à 40 mm. En effet, le bilan ne varie pas sensiblement avec le choix de RFU_{max}, et la valeur 40 mm paraît raisonnable pour les sols sur roches calcaires.

L'ETR (évapotranspiration réelle) est obtenue à partir d'une comparaison des grandeurs de l'évapotranspiration potentielle calculée par la formule de Thornthwaite à la pluie brute P. Trois cas peuvent se présenter :

$$1. \text{ Si } ETP < P, \text{ ETR} = ETP$$

La quantité d'eau correspondant à P-ETP est emmagasinée dans le sol jusqu'à la saturation de celui-ci . Le reste est disponible pour l'écoulement superficiel et l'infiltration vers les nappes souterraines.

Fig. 34 - Carte des mines (extrait de [8]).

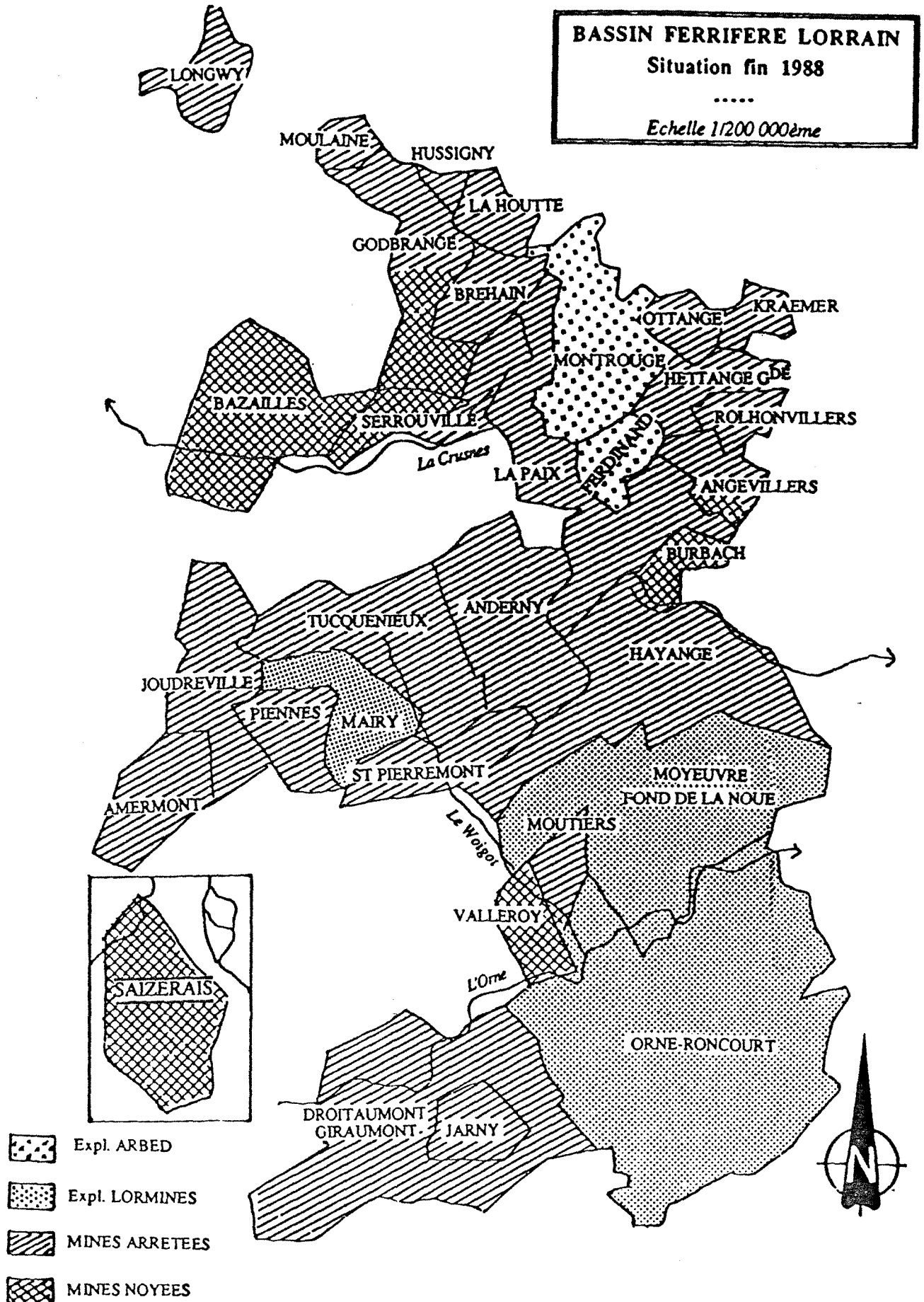
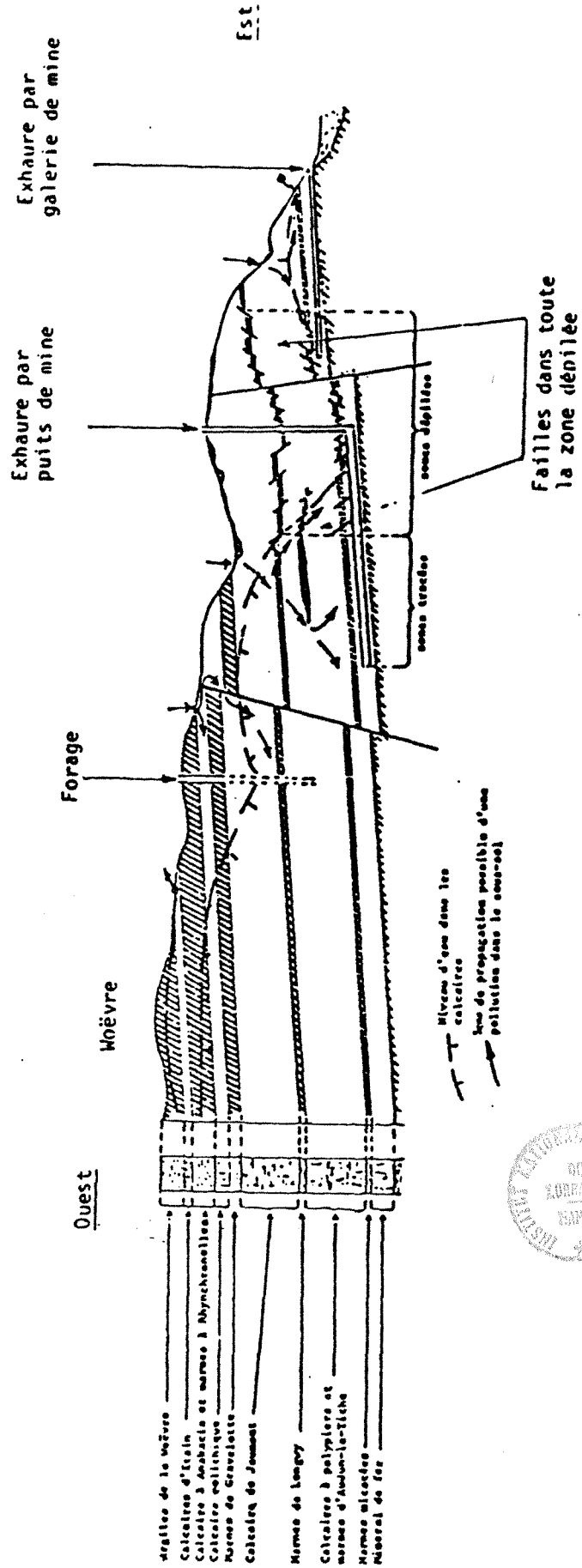


Fig. 35 - Coupe géologique (extrait de [8]).



$$2. \text{ Si } ETP = P, ETR = ETP = P$$

Les réserves en eau du sol ne varient pas.

$$3. \text{ Si } ETP > P, ETR = P + \Delta R$$

Dans le bilan (6.1), on ne peut pas distinguer la pluie efficace au sens strict E et l'écoulement superficiel, celui-ci étant perturbé par les rejets d'exhaure. On est donc obligé de calculer la pluie efficace au sens large que l'on appelle par la suite, pluie efficace tout court.

Par ailleurs, si l'on admet que l'écoulement superficiel est proportionnel à l'infiltration, on peut ne pas en tenir compte [14, 21, 22].

Les pluies efficaces mensuelles et annuelles ont été calculées à deux postes (Frescaty et Tomblaine).

II - RELATION EXHAURE-PLUIE EFFICACE, EXTRACTION DE MINERAI ET RESERVES SOUTERRAINES

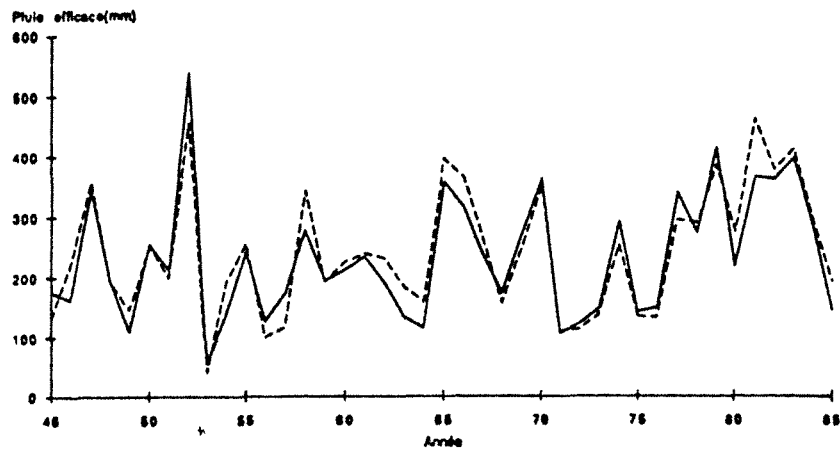
II.1 - Pluies efficaces, réserves souterraines et exhaure

II.1.1 : Analyse des pluies efficaces

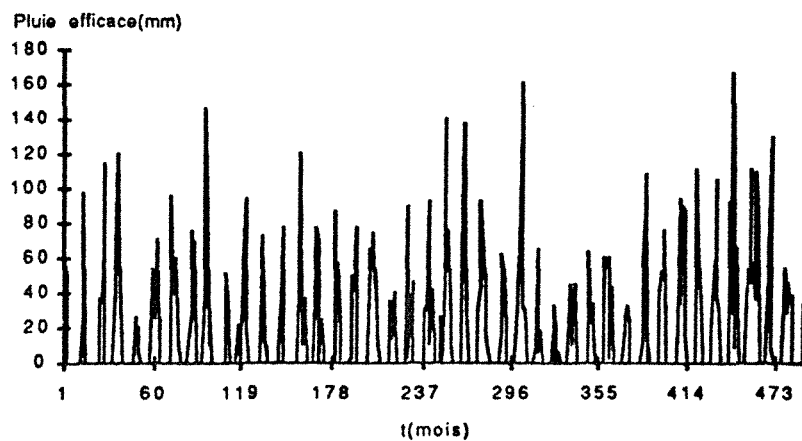
Si l'on compare les pluies efficaces annuelles aux postes Tomblaine et Frescaty (Fig. 36a), on pourra remarquer que leurs variations se rapprochent considérablement. Les statistiques (tableau IV) montrent bien que les moyennes ou écart-types des pluies efficaces à ces deux postes sont quasiment identiques. On en déduit que la faible variation météorologique dans l'espace est pratiquement négligeable. Par conséquent, on pourra utiliser les pluies efficaces à ces deux postes pour toutes les mines dans le bassin.

Le corrélogramme de la pluie efficace annuelle à Frescaty (Fig. 37a) montrent à l'évidence que celle-ci est purement aléatoire, car tous les coefficients d'autocorrélation se situent dans l'intérieur de l'intervalle de confiance à 95 %.

La pluie efficace mensuelle à Frescaty est illustrée par la figure 36b. Son corrélogramme (Fig. 37b) fait apparaître qu'elle suit une variation saisonnière de périodicité 12 mois (les coefficients d'autocorrélation à 12, 24, 36 mois sont supérieurs à leurs voisins).



a - Pluies efficaces annuelles à Tomblaine et à Frescaty (1945-1985)



b - Pluie efficace mensuelle à Frescaty (1945-1985)

Fig. 36 - Pluies efficaces à Tomblaine et à Frescaty

station	Moyenne(mensuelle)	Ecart-type(mensuelle)	Moyenne(annuelle)	Ecart-type(annuelle)
Frescaty (Metz)	20,1	31,7	240,6	105,2
Tomblaine(Nancy)	19,4	31,6	232,8	104

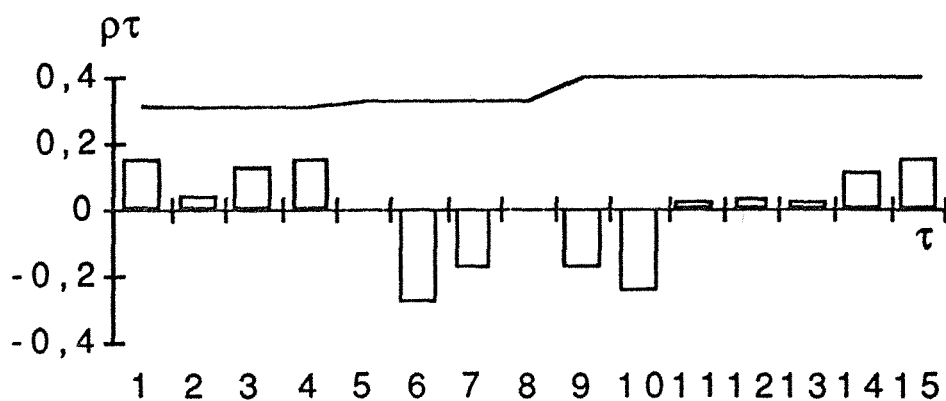
Tableau IV - Statistiques des pluies efficaces

Pour la pluie efficace à Frescaty, le recours à la méthode de Box-Jenkins nous conduit à obtenir un modèle tel que (Fig. 38) :

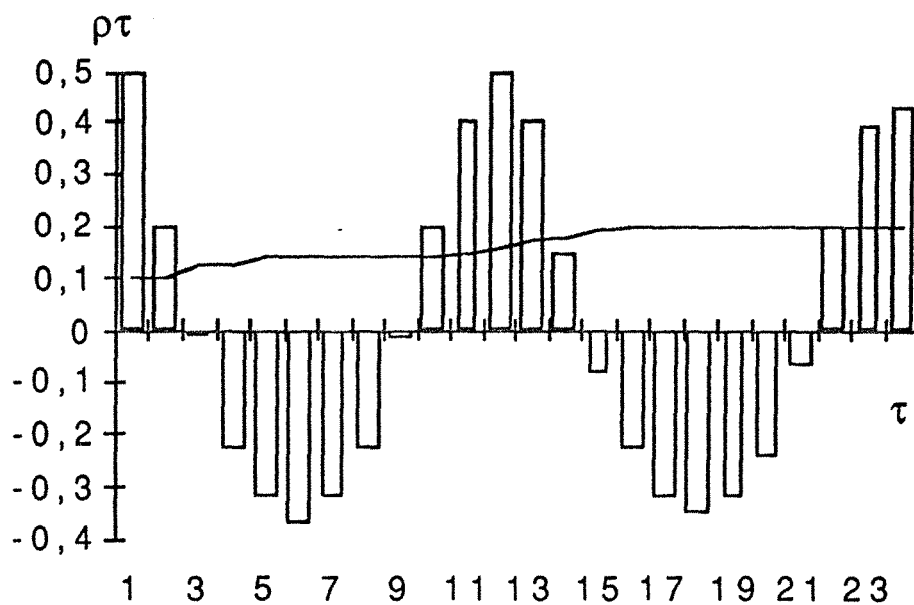
$$(1-B^{12})Y_t = (1-\theta B)(1-\Theta B^{12})a_t \quad (6.2)$$

Dans lequel $q=0,2$, $Q=0,91$.

Fig. 37 - Corrélogramme de la pluie efficace à Frescaty(1945-1985)

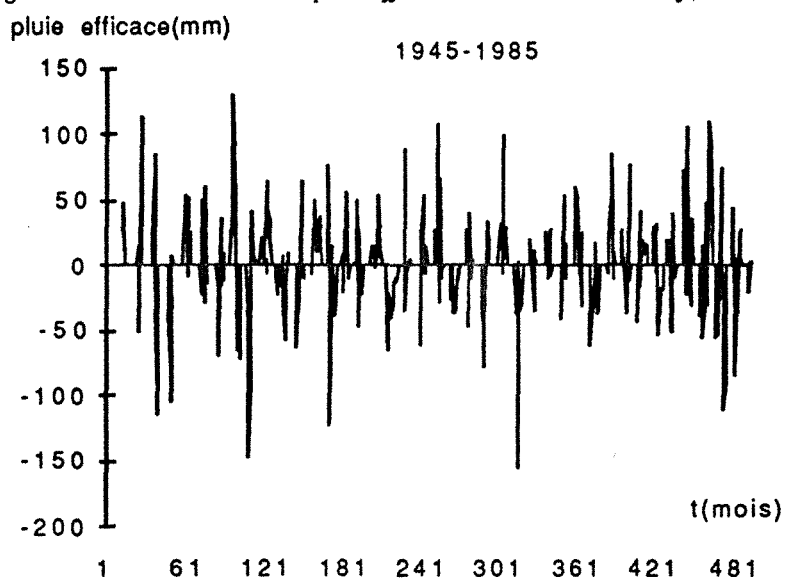


a - corrélogramme de la pluie efficace annuelle

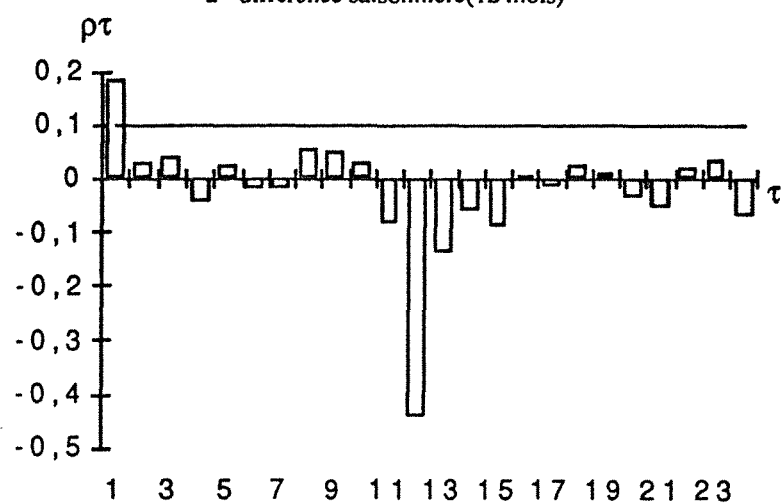


b - corrélogramme de la pluie efficace mensuelle

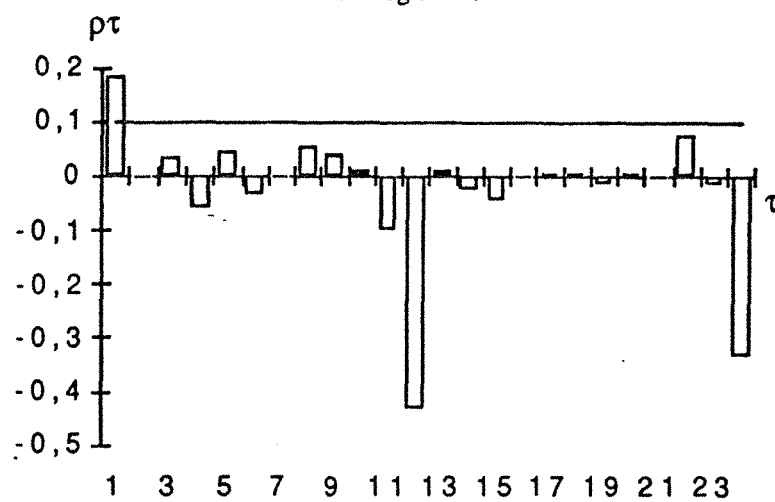
Fig. 38 - Modèle ARIMA de la pluie efficace mensuelle à Frescaty(1945-1985)



a - différence saisonnière(12 mois)



b - corrélogramme



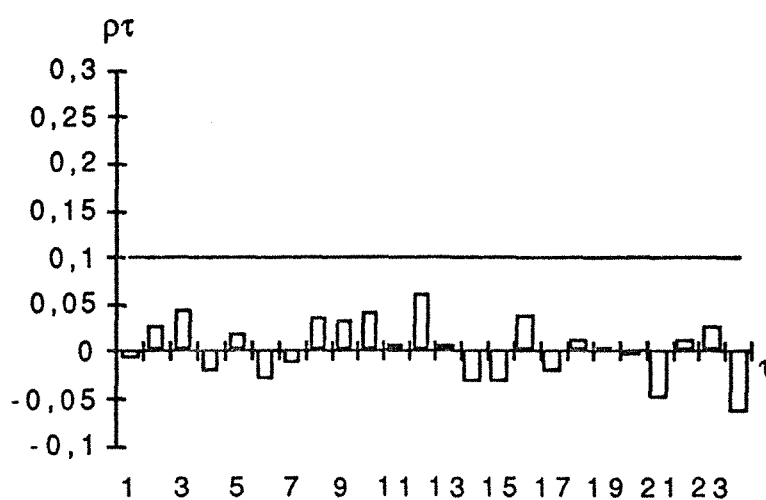
c - corrélogramme partiel

Fig. 38 - Modèle ARIMA de la pluie efficace à Frescaty(1945-1985)(fin)

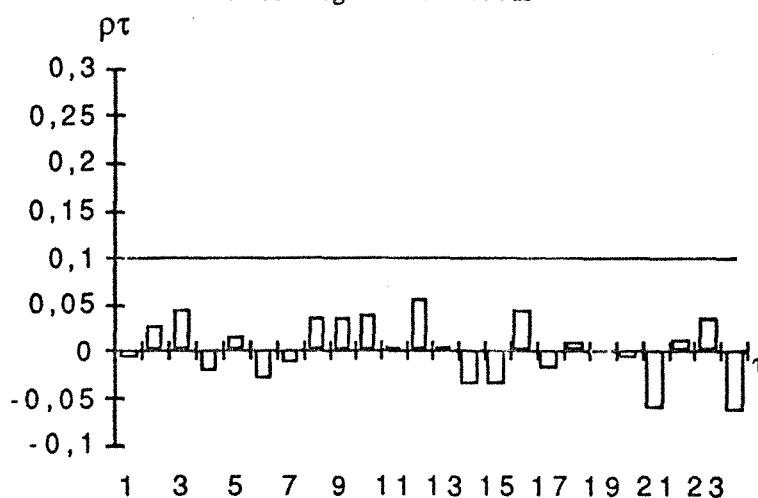
Parameter	Estimate	Std.error	T-value	P-value
MA(1)	-0,20022	0,04499	-4,44997	0,00001
SMA(12)	0,91161	0,01958	46,56819	0,00000

Model fitted to seasonal differences of order 1 with seasonal length = 12.
 Estimated white noise variance = 570,079 with 478 degrees of freedom.
 Estimated white noise standard deviation (std err) = 23,8763.
 Chi-square test statistic on first 20 residual autocorrelations = 7,85716.
 with probability of a larger value given white noise = 0,980688.

d - modèle ARIMA



e - corrélogramme des résidus



f - corrélogramme partiel des résidus

Ce modèle peut être écrit d'une façon suivante :

$$Y_t = Y_{t-12} - 0,9 a_{t-12} - 0,18 a_{t-13} + 0,2 a_{t-1} + a_t \quad (6.3)$$

Le modèle décrit par l'équation (6.3) est un modèle de prévision.

Pour la pluie efficace à Tomblaine, la même analyse a été également effectuée. Sans produire les graphiques, nous donnons immédiatement le modèle de prévision :

$$Y_t = Y_{t-12} - 0,92 a_{t-12} - 0,2 a_{t-13} + 0,22 a_{t-1} + a_t \quad (6.4)$$

On constate que le modèle (6.4) a la même structure que le modèle (6.3), seule apparaît une très légère différence entre les coefficients de pondération, ceci confirme donc que la variation météorologique dans l'espace est pratiquement négligeable.

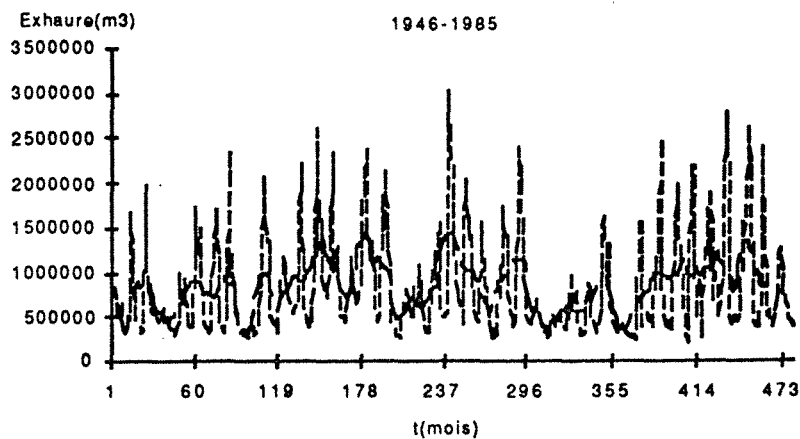
On pourra enfin remarquer qu'une chronique se comporte différemment selon l'échelle de temps et de ce fait, celle-ci doit toujours être précisée.

II.1.2 : Analyse d'exhaure minière

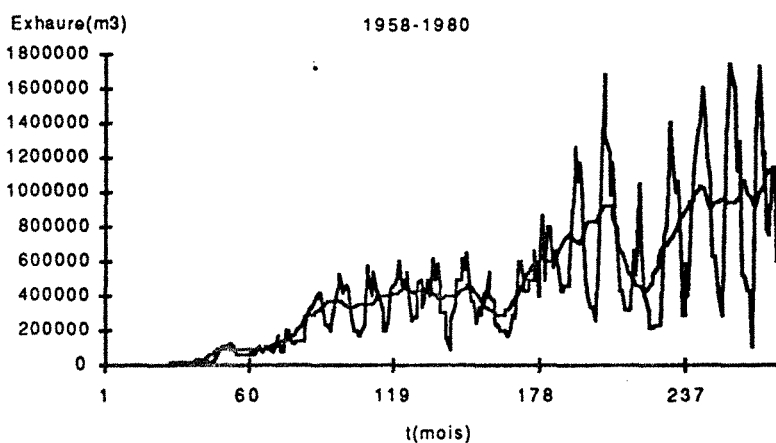
Dans le bassin ferrifère lorrain, on peut distinguer deux types de mines: les mines stationnaires et les mines non stationnaires. La figure 39 illustre ces deux types de mines. La mine Moyeuve est stationnaire et le corrélogramme de son exhaure (fig. 40a) montre que celle-ci est purement saisonnière de périodicité 12 mois. Les mines Saizerais et Anderny sont non stationnaires (fig. 39b, c) et le corrélogramme de l'exhaure à Saizerais montre que les composantes tendancielle et saisonnière se superposent (fig. 40b).

II.1.3 : Relation exhaure minière-réserves souterraines et pluie efficace

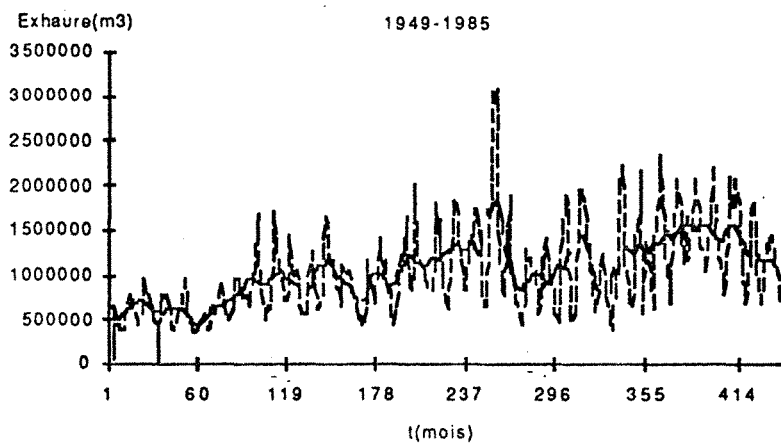
Les seuls piézomètres qui possèdent des relevés suffisamment denses sur une période assez longue sont les piézomètres de Roncourt notés R1 à R11. La figure 41 fait une comparaison entre la pluie efficace, la hauteur piézométrique et l'exhaure minière à Roncourt. Nous constatons que la hauteur piézométrique est une fluctuation saisonnière, alors que sa variation est beaucoup plus faible que la variation de la pluie efficace et que celle de l'exhaure. Par ailleurs, on peut aussi constater que l'exhaure agit presque spontanément avec la hauteur piézométrique et qu'il existe un décalage net dans le temps (1 ou 2 mois) entre la hauteur piézométrique et la pluie efficace, ceci est aussi mis en évidence par la comparaison entre les composantes saisonnières de la pluie efficace et de l'exhaure (fig. 42).



a - Moyeuve



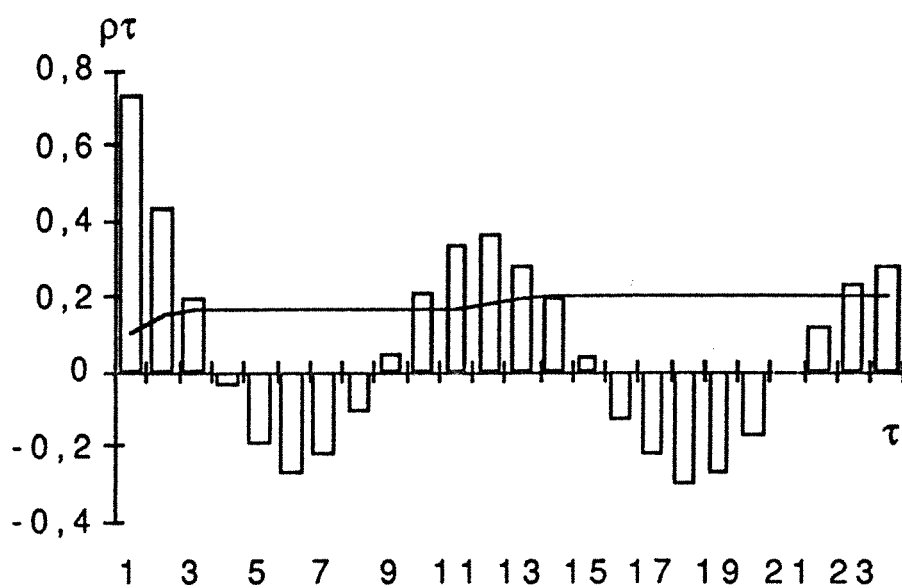
b - Saizerais



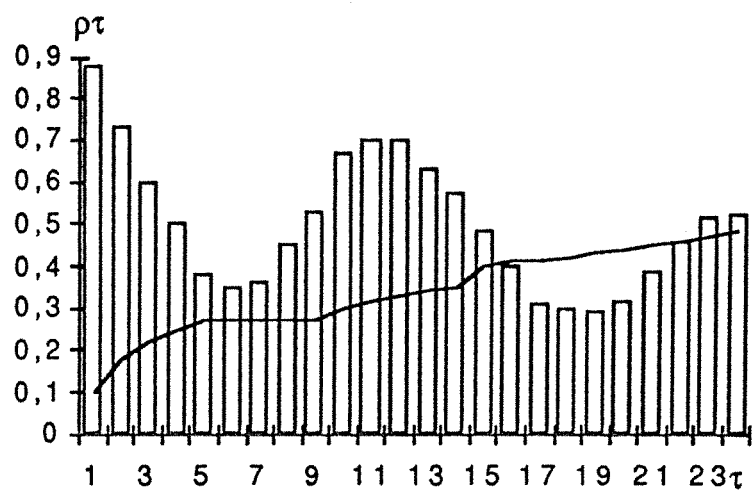
c - Anderny

Fig. 39 - Exhaure mensuelle et la moyenne mobile centrée à 12 termes pour trois mines dans le bassin ferrifère lorrain

Fig. 40 - Corrélogrammes de l'exhaure

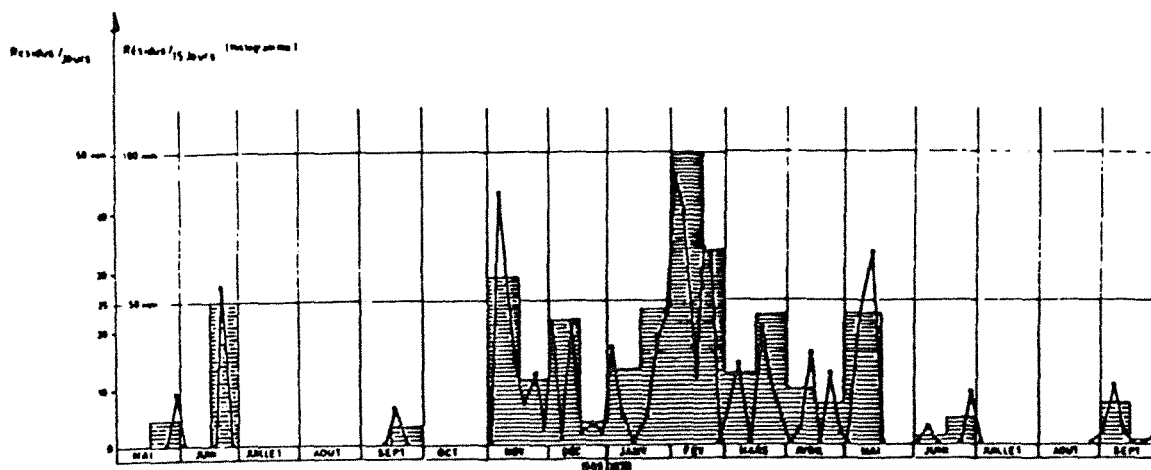


a - corrélogramme de l'exhaure à Moyeuivre(1946-1985)

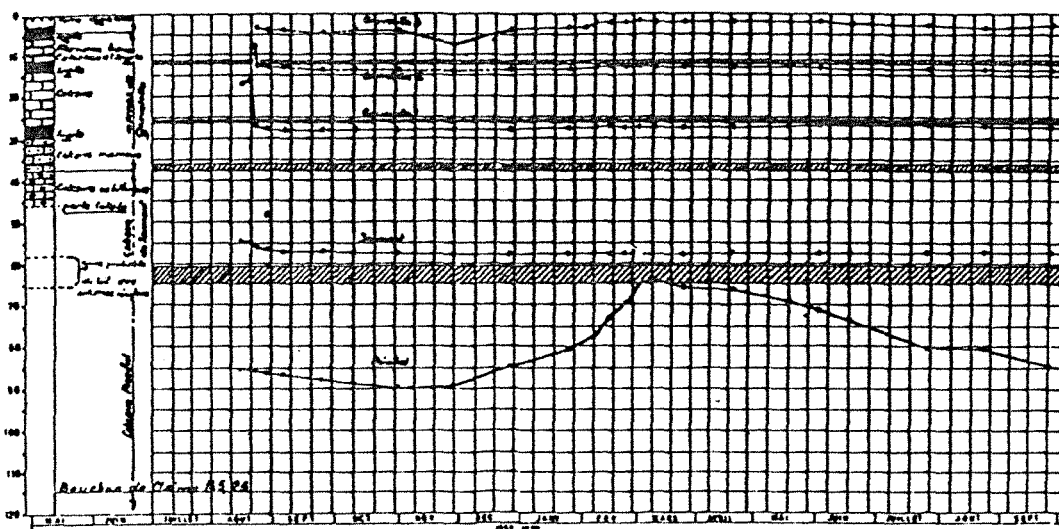


b - corrélogramme de l'exhaure à Saizerais(1958-1980)

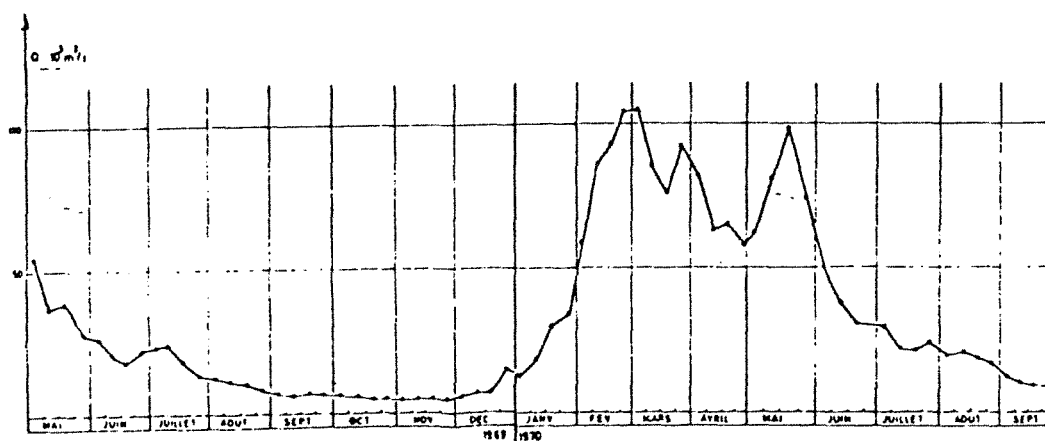
Fig. 41 - Comparaison entre la pluie efficace, la hauteur piézométrique et l'exhaure minière à Roncourt (extrait de [53])



a - pluie efficace

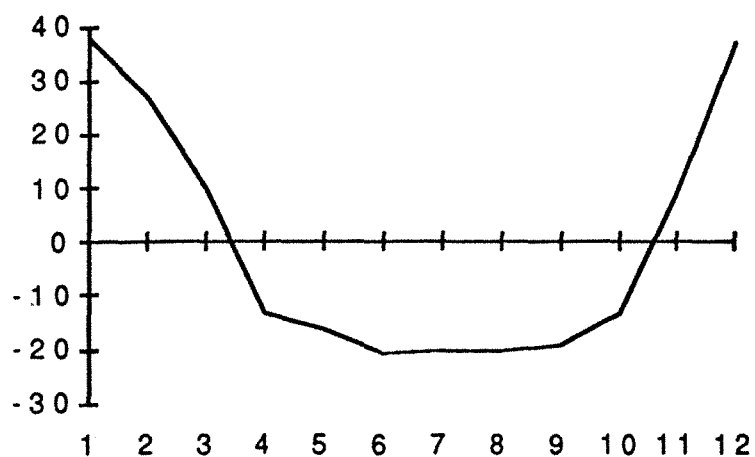


b - hauteur piézométrique

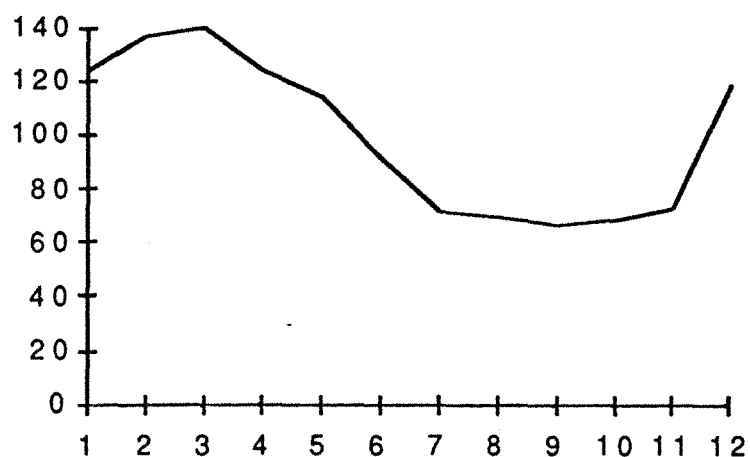


c - exhaure minière

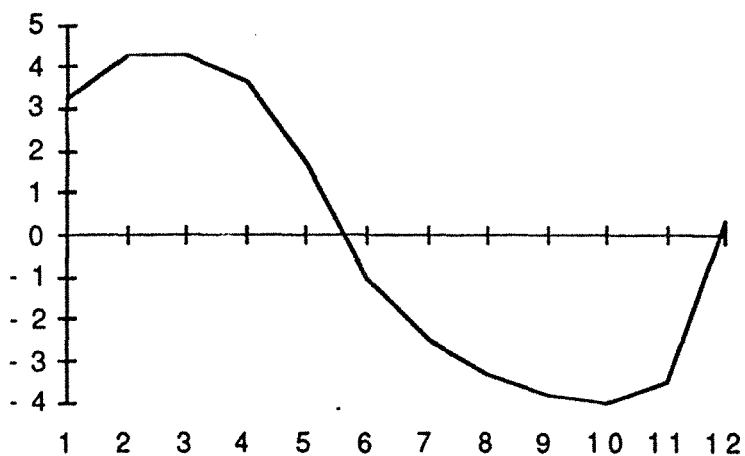
Fig. 42 - Comparaison entre les composantes saisonnières de la pluie efficace et de l'exhaure



a - pluie efficace à Frescaty(1945-1975)



b - exhaure à Saizerais(1958-1980)



c - exhaure à Moyeuivre(1951-1958)

II.2 - Effet de l'extraction de minerai sur l'exhaure

En Lorraine, l'exploitation minière s'effectue par traçage puis par dépilage de manière à obtenir un taux de "défruitement" de l'ordre de 85 % à 90 %. Si la crise de la sidérurgie a entraîné la fermeture de nombreuses exploitations, elle ne s'est pas traduite par une baisse de l'exhaure (fig. 43). Au cours des années 1971 à 1975, pour une production voisine ou légèrement supérieure à 50 millions de tonnes de minerais, les exhaures représentent en moyenne 150 millions de m³, soit 3 m³ d'eau par tonne de minerai. Au cours des années 1985 à 1987, pour une tonne de production moyenne de 12 millions de tonnes, les exhaures représentaient en moyenne 200 millions de m³, soit un peu plus de 16 m³ d'eau par tonne de minerai. Pour l'année 1987 seule, il a fallu pomper près de 21,5 m³ d'eau par tonne de minerai extrait, le tonnage totale de minerai extrait étant de l'ordre de 10,5 millions de tonnes.

Nous avons effectué une étude des corrélations entre exhaure annuelle-pluie efficace annuelle, extraction annuelle et extraction annuelle cumulée pour une quinzaine de mines. Les résultats rassemblés dans le tableau V nous permet de dégager de grandes lignes difficilement contestables, les coefficients négatifs étant considérés comme zéro en tenant compte du mécanisme de l'exhaure (effectivement, progression de l'exploitation ne peut jamais provoquer une diminution de l'exhaure et l'apparition des coefficients négatifs peut s'expliquer par le fait que les données d'exhaure d'une période relativement sèche, où l'exhaure diminue tandis que l'extraction continue, sont ajustées) :

mine	ρ_1	ρ_2	ρ_3	ρ	période
Bazailles	0,83	0	-0,8	0,87	67-81
Ottange	0,77	-0,64	0	0,83	51-68
Errouville	0,8	0	0,88	0,94	51-85
Angevillers	0,75	-0,65	0	0,87	65-79
Paix	0,81	0	-0,51	0,81	51-83
Tucquenieux	0,88	0	0	0,88	54-85
Joudreville	0,57	0	0	0,57	52-84
Amermont	0,4	0	-0,71	0,71	51-85
Piennes	0,77	0	-0,54	0,78	66-85
St-Pierremont	0,81	0	0,49	0,87	51-85
Anderny	0,59	0	0,78	0,86	51-83
Moyeuivre	0,78	0	-0,36	0,78	51-85
Giraumont	0,55	0,42	0	0,61	51-78
Droitaumont	0,67	0	0,85	0,91	51-85
Saizerais	0,59	0	0,89	0,9	58-80

ρ_1 : coefficient de corrélation entre exhaure et pluie efficace

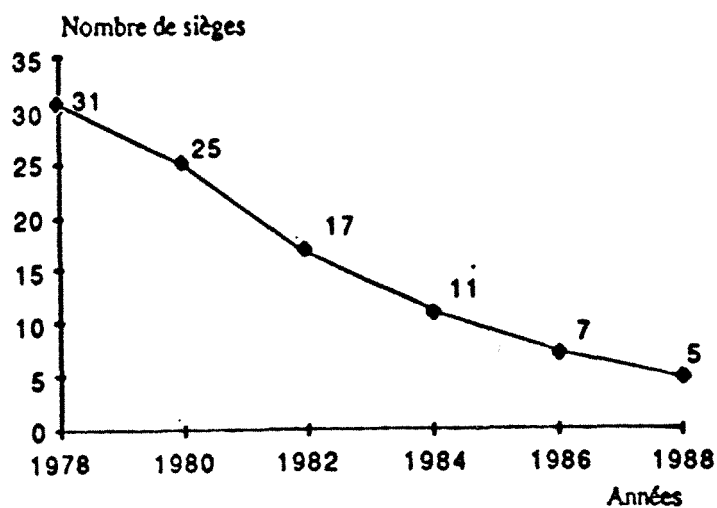
ρ_2 : coefficient de corrélation entre exhaure et extraction annuelle des minerais

ρ_3 : coefficient de corrélation entre exhaure et extraction cumulée des minerais

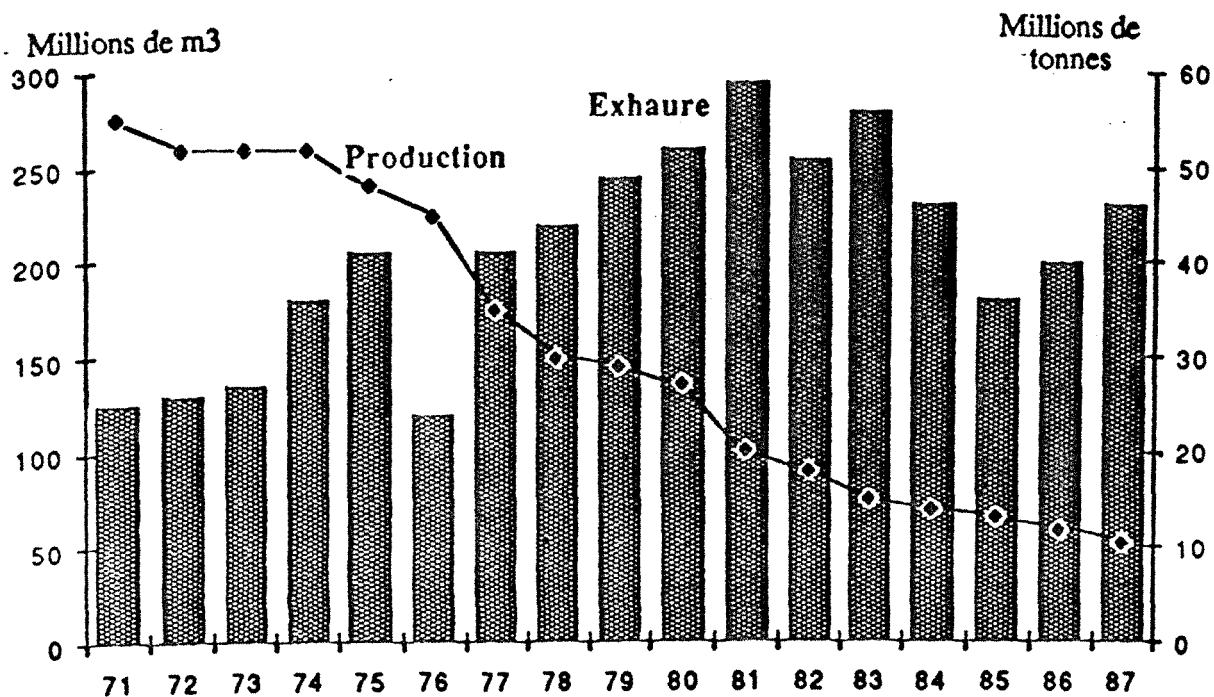
ρ : coefficient de corrélation multiple

Tableau V - Coefficients de corrélation entre exhaure et ses facteurs influents

Fig. 43 - Exploitation et exhaure (extrait de [8])



a - évolution du nombre de sièges des mines de fer de Lorraine au cours des 10 dernières années



b - évolution comparative de la production de minerai et de l'exhaure

* Pour toutes les mines, l'exhaure est bien corrélée à la pluie efficace. Les coefficients de corrélation partielle sont généralement supérieurs à 0.55, sauf pour la mine Amermont (0.4). Il convient de remarquer que la mine Amermont se trouve à une position occidentale du bassin minier et elle s'exploite à une grande profondeur (hauteur de refoulement = 210 m). Le recouvrement atteint les horizons les plus élevés dans l'échelle stratigraphique. La concession est entièrement recouverte par les marnes à Rhynchonelles imperméables [37]. Le lien entre l'exhaure et la pluie efficace est, par conséquent, moins direct.

* Pour toutes les mines l'exhaure est non corrélée à l'extraction annuelle. En effet, on peut imaginer que, dans la réalité, l'extraction annuelle provoque des variations de l'exhaure d'une manière tout à fait aléatoire (vidange des réserves dans la formation exploitée, rencontre des accidents structuraux et évacuation des eaux stockées, etc.). L'effet de l'extraction annuelle est pratiquement ou numériquement négligeable dans la modélisation.

* L'extraction annuelle cumulée, est un facteur influent de l'exhaure, notamment pour les mines non stationnaires.

On peut alors conclure que la non stationnarité de l'exhaure est effectivement provoquée par l'extraction cumulé de minerai, et la saisonnalité par la variation saisonnière de la pluie efficace.

Dans le troisième paragraphe, nous allons travailler à l'échelle mensuelle. Il nous faut convertir les données d'extraction annuelles en données mensuelles, puis en données mensuelles cumulées. Pour ce faire, nous pouvons diviser les données annuelles par 12. Vue que l'influence de l'extraction sur l'exhaure est à long terme, l'imprécision due à cette division est pratiquement négligeable.

II.3 - Simulation de l'exhaure-modèle physique

On note, par la suite, Q_t exhaure mensuelle, E_t pluie efficace mensuelle et S_t l'extraction mensuelle cumulée.

L'exhaure minière Q_t peut être assimilée à la vidange d'un récipient (fig. 44) alimenté par la pluie efficace E_t . Le niveau d'eau H_t représente les réserves souterraines (la nappe du Dogger) qui jouent un rôle régulateur entre E_t et Q_t . Ce niveau d'eau peut être considéré, en quelque sorte, comme constant, puisque sa variation est beaucoup plus faible que les variations de la pluie efficace et de l'exhaure (cf. fig. 41).

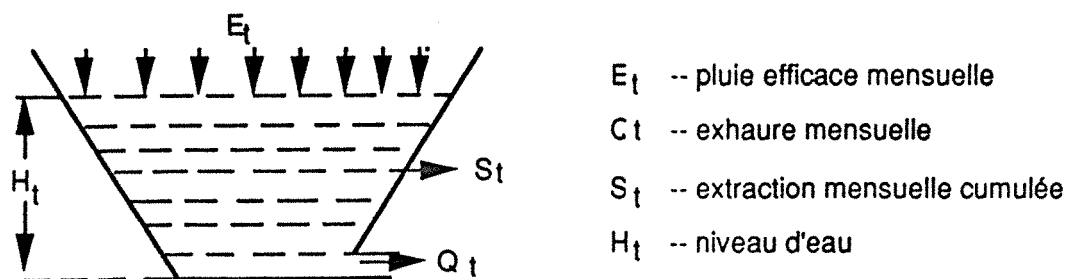


Fig. 44 - Simulation de l'exhaure

L'extraction mensuelle cumulée de minerai S_t est le moteur de la non stationnarité de l'exhaure Q_t . Son action à l'intérieur du système représenté ci-dessus se traduit par un élargissement du récipient.

Ayant ainsi défini le modèle physique, nous allons établir, dans le paragraphe suivant, un modèle mathématique qui représente la relation entre les différents facteurs illustrés par la figure 44.

III - MODELISATION NUMERIQUE

Dans ce paragraphe, nous allons effectuer la déconvolution pluie efficace-exhaure pour toutes les mines, en tenant compte de l'extraction cumulée de minerai et des réserves souterraines. Cette étude doit porter notamment sur deux cas : les mines stationnaires et les mines non stationnaires.

La méthode de régression multiple sera mise en jeu. Pour les contraintes, nous proposons que les coefficients de régression soient strictement égaux ou supérieurs à zéro (car il est dépourvu de sens physique quand il y a des portions de courbes de fonction de transfert prenant des valeurs négatives) et que les coefficients de Student soient généralement supérieurs ou égaux à 2 (test de signification).

III.1. - Mines stationnaires

L'exhaure de ces mines est stationnaire. Elle est sous la dépendance exclusive de la pluie efficace. L'extraction cumulée de minerai demeure sans influence. Le modèle est relativement simple (fig. 45)...

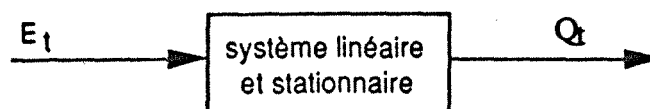


Fig. 45 - Mines stationnaires

et peut s'écrire sous la forme suivante...

$$Q_t = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \varepsilon_t \quad (6.5)$$

... où C est une constante ; $\phi_0, \phi_1, \phi_2, \dots, \phi_m$ sont les coefficients de pondération qui constituent évidemment la fonction de transfert ; m est la mémoire du système (la pluie efficace à une date antérieure à $t-m$ n'a plus d'influence sur Q_t) ; ε_t est le terme résiduel qui doit être un bruit blanc si les données sont parfaitement ajustées.

A titre d'exemple, nous présentons ici la mine Moyeuivre. Cette mine d'amont pendage est située à l'aplomb d'un plateau limité à l'Est par la vallée de la Moselle, au Sud par la Vallée de l'Orne et à l'Ouest par la vallée de Conroy. La hauteur de refoulement est faible (33 m). Les morts-terrains recouvrant les travaux miniers sont essentiellement calcaires [38].

Les résultats de déconvolution sont présentés dans la figure 46. Le coefficient de corrélation R^2 est égal à 0.68 (fig. 46a), donc l'ajustement est bon, ce qui est confirmé par le corrélogramme des résidus (fig. 46c). Mais il faut noter que les quatre premiers coefficients d'autocorrélation dépassent l'intervalle de confiance. Les résidus ne sont donc pas purement aléatoires. Le terme résiduel sera étudié à la fin de ce paragraphe.

On conclut que le modèle (6.5), sans être parfait, représente correctement la relation exhaure-pluie efficace pour la mine Moyeuivre. Cette conclusion peut s'étendre aisément à toutes les mines stationnaires.

III.2 - Mines non stationnaires

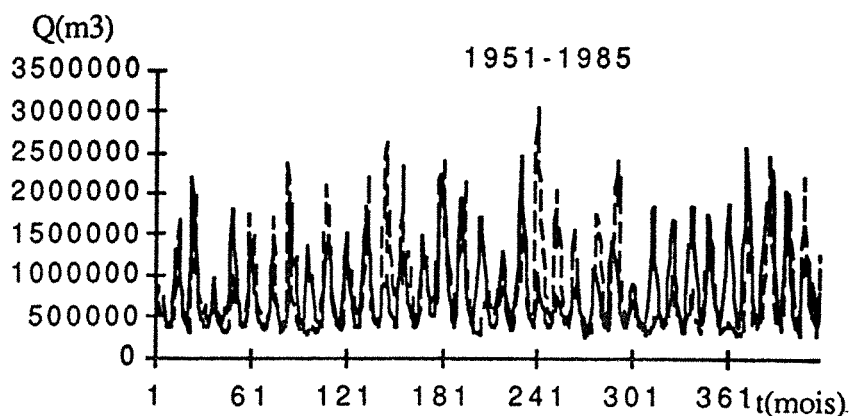
L'exhaure de ce type de mines est non stationnaire en moyenne et en variance. Elle est en augmentation avec la progression de l'exploitation minière, celle-ci pouvant être "mesurée" numériquement par le tonnage cumulé de l'extraction de minerai S_t . Ce phénomène de non stationnarité peut s'expliquer comme suit: au fur et à mesure que progresse l'exploitation minière, les communications entre la mine et les nappes des calcaires bajociens se multiplient. L'exhaure devenant de plus en plus sensible à la variation saisonnière de la pluie efficace, s'avère d'une part non stationnaire en variance (élargissement de la "bande" de variation) et d'autre part, les dimensions du récipient s'élargissant, les eaux exhaurées proviennent de plus en plus loin et par conséquent, la moyenne de l'exhaure présente une augmentation permanente.

Fig. 46 - Déconvolution à Moyeuivre ($Q_t - E_t$)

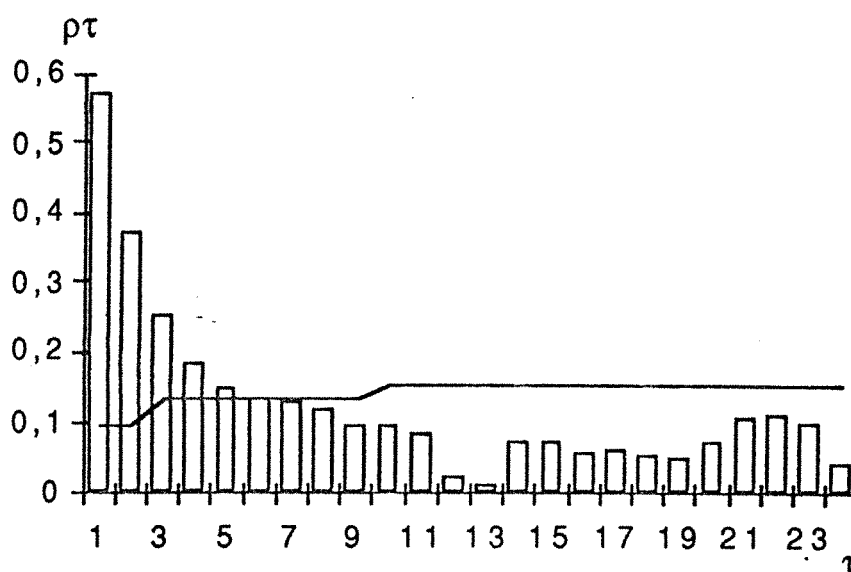
Dependent variable	coefficient	std.error	t-value	sig. level
CONSTANT	3,873143E5	2,500791E4	15,4877	0,0000
E	3358,519408	586,505161	5,7263	0,0000
1 LAGE	9215,46879	633,964659	14,5363	0,0000
2 LAGE	3736,918727	628,765192	5,9433	0,0000
3 LAGE	2371,136799	627,598511	3,7781	0,0002
4 LAGE	2120,810752	632,924417	3,3508	0,0009
5 LAGE	1861,406101	585,346528	3,1800	0,0016

R-SQ. (ADJ) = 0,6793 SE= 315013,160688 MAE= 226316,429079 Durbwat = 0,858
415 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

a - modèle de régression



b - exhaure mesurée et reconstituée



c - corrélogramme des résidus

La mine Saizerais appartient à ce type de mines. Cette mine se trouve au Sud-Ouest du bassin (en aval pendage), à l'intérieur d'une cuvette structurale faillée. L'aquifère est alimenté par la fraction infiltrée de l'impluvium s'abattant sur le bassin versant constitué par les vastes affleurements calcaires du plateau de Rosières-en-Haye. La hauteur de refoulement est environ 160 m [8].

Pour le problème de déconvolution, nous avons tenté, au début, d'ajuster un modèle schématisé par figure 47.

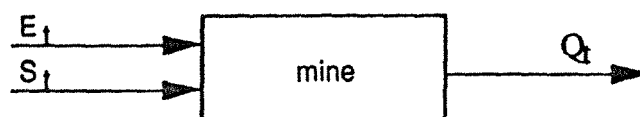


Fig. 47 - Mines non stationnaires

Nous avons supposé de plus, que l'influence de S_t était instantanée tandis que celle de E_t était en décalage. La relation s'est exprimée par :

$$Q_t = C + \alpha S_t + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \epsilon_t \quad (6.6)$$

Les résultats de déconvolution (fig. 48) nous montrent que le modèle (6.6) n'est pas adapté à la simulation de l'exhaure pour les mines non stationnaires : la figure 48b indique la différence remarquable entre Q_t mesurée et Q_t calculée ; la figure 48c montre l'existence de saisonnalité dans les résidus. Compte tenu du mécanisme de l'exhaure, ce modèle nous semble peu représentatif : S_t n'aurait pas du être traitée comme une entrée, mais plutôt comme un excitateur qui régit le tempérament de la mine (état du système). Il aurait fallu établir un modèle qui pourrait refléter cette réalité. Pour cela, nous avons conçu un autre modèle schématisé par la figure 49 dans laquelle E_t est stationnaire, Q_t est non stationnaire.

Le système n'est ni linéaire (non stationnarité en variance) ni stationnaire (augmentation de la moyenne). Par conséquent, la déconvolution n'est pas immédiatement applicable entre Q_t et E_t . Pour que la déconvolution soit possible, il faut rendre *a priori* stationnaire la sortie du système Q_t en tenant compte du moteur de la non stationnarité, c'est-à-dire l'extraction cumulée de minerai S_t .

Par des essais, nous avons vu que, une simple division $\frac{Q_t}{S_t}$ permettait de rendre l'exhaure stationnaire (fig. 50a)

Dès lors, on peut écrire :

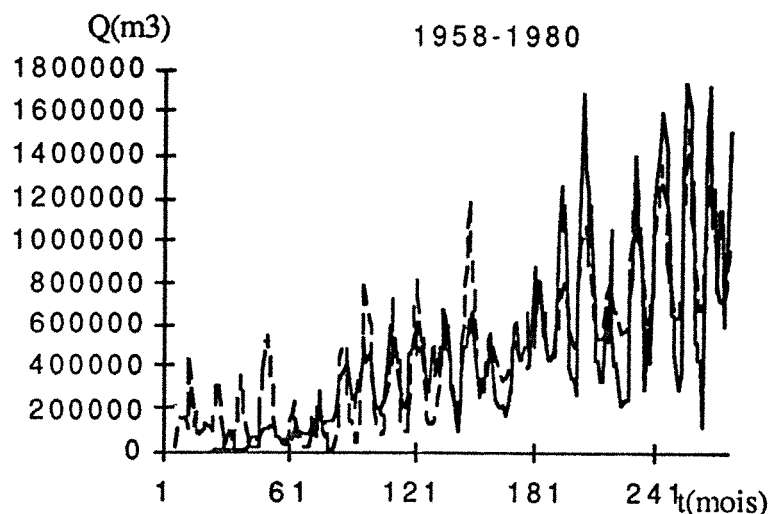
$$\frac{Q_t}{S_t} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \epsilon_t \quad (6.7)$$

Fig. 48 - Déconvolution à Saizerais ($Q_t - S_t, E_t$)

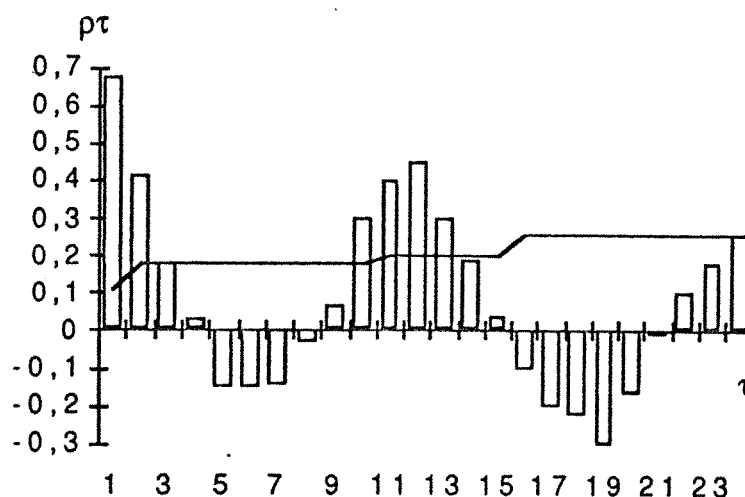
Dependent variable	coefficient	std.error	t-value	sig. level
CONSTANT	-2,055417E5	3,95041E4	-5,2030	0,0000
S	0,045486	0,002677	16,9903	0,0000
E	2510,832622	606,669591	4,1387	0,0001
1 LAGE	4012,235935	637,038303	6,2983	0,0000
2 LAGE	1411,279807	644,529969	2,1896	0,0297
3 LAGE	1231,768495	636,702145	1,9346	0,0544
4 LAGE	1394,422523	608,3613453	2,2921	0,0229

R-SQ. (ADJ) = 0,6878 SE= 219883,600666 MAE= 171866,490667 Durbwat =0,784
 216 observations fitted, forecast(s) computed for 56 missing val. of dep. var.

a - modèle de régression



b - exhaure mesurée et reconstituée



c - corrélogramme des résidus

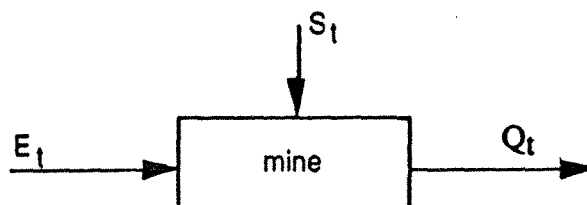
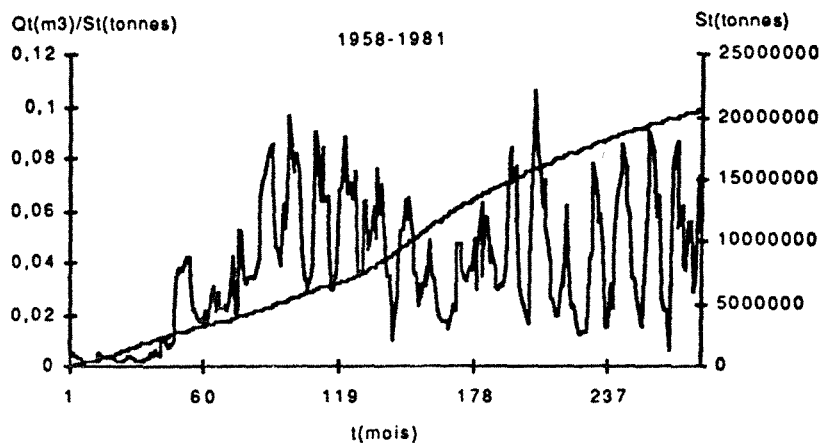
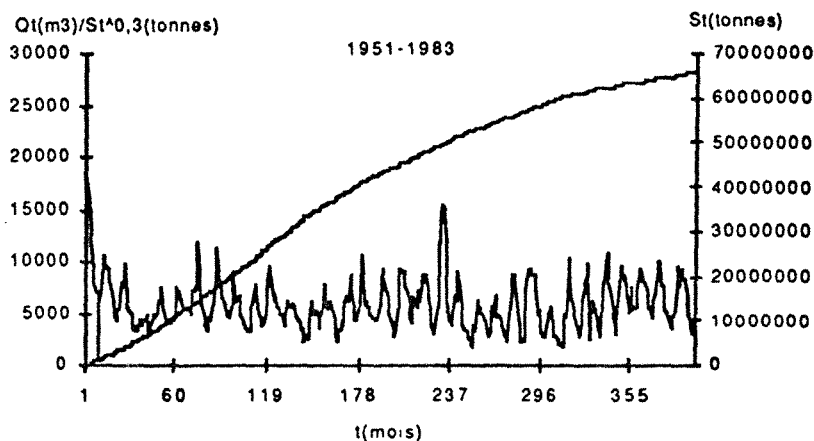


Fig. 49 - Système avec excitateur



a - Saizerais



b - Anderny

Fig. 50 - Extraction mensuelle cumulée et transformation de l'exhaure

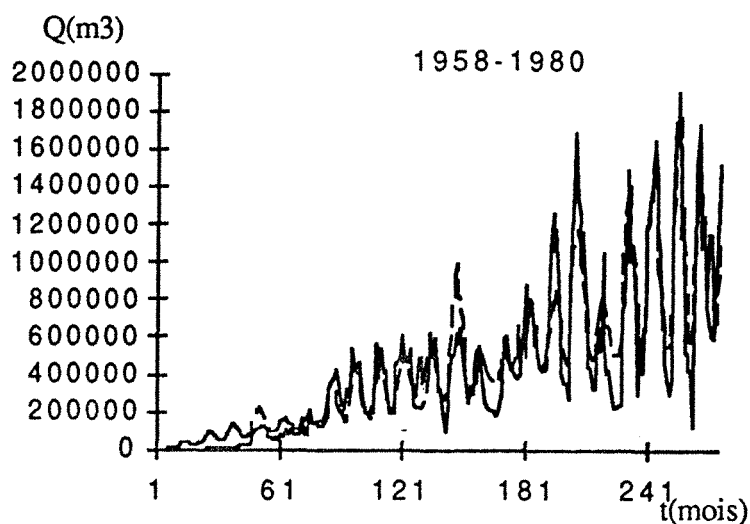
On pourra constater que l'ajustement de ce modèle présente un coefficient de corrélation multiple $R^2 = 0.57$ (fig. 51a) et que Q_t calculée se rapproche considérablement de la Q_t mesurée (fig. 51b). D'autre part, le corrélogramme des résidus (fig. 51c) montre que la saisonnalité dans Q_t a été complètement éliminée.

Fig. 51 - Déconvolution à Saizerais ($\frac{Q_t}{S_t} - E_t$)

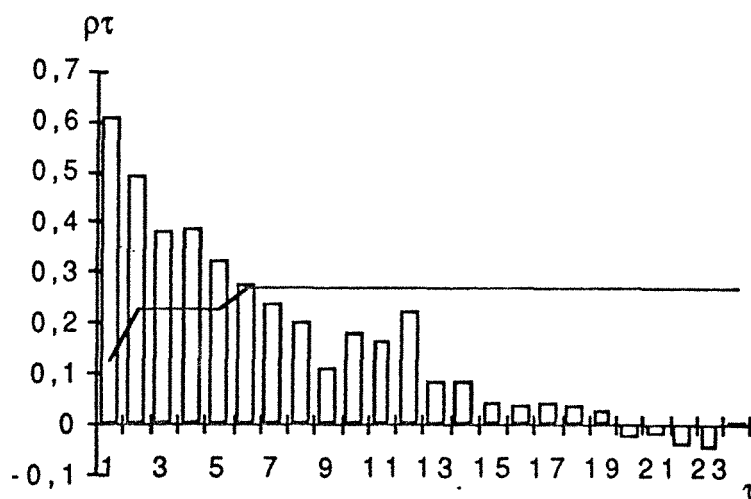
Dépendent variable	coefficient	std.error	t-value	sig. level
CONSTANT	0,028483	0,001614	17,6526	0,0000
E	0,000223	0,00004	5,5347	0,0000
1 LAG E	0,000322	0,000043	7,5082	0,0000
2 LAG E	0,000083	0,000042	1,9553	0,0519
3 LAG E	0,000101	0,000042	2,3865	0,0179
4 LAG E	0,000118	0,000043	2,7556	0,0064
5 LAG E	0,000057	0,00004	1,3958	0,1643

R-SQ. (ADJ) = 0,5716 SE= 0,014473 MAE= 0,011164 Durbwat =0,751
 216 observations fitted, forecast(s) computed for 55 missing val. of dep. var.

a - modèle de régression



b - exhaure mesurée et reconstituée



c - correlogramme des résidus

On en déduit que le modèle (6.7) est pratiquement adapté à rendre compte de la relation pluie-exhaure pour la mine Saizerais.

Pour les autres mines de ce type, il a été mis en évidence que la transformation $\frac{Q_t}{S_t^\beta}$ ($0 < \beta < 1$) peut rendre l'exhaure Q_t stationnaire. Pour la mine d'Anderny, la valeur β est égale à 0.3 (fig. 50b) et les résultats de déconvolution sont présentés dans la figure 52.

Ainsi, un modèle représentant toutes les mines de ce deuxième type peut être exprimé par la relation...

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \varepsilon_t \quad (6.8)$$

... où $0 < \beta \leq 1$. β sera appelé par la suite, "coefficient de tempérament".

Le recours à ce modèle fait apparaître une difficulté : il a été signalé dans le premier paragraphe, que les données d'extraction de minerai sont manquantes pour cinq mines. Pour surmonter cette difficulté, nous allons introduire une autre transformation du type $\frac{Q_t}{MA12}$ et :

$$MA12 = \frac{0.5Q_{t-6} + Q_{t-5} + \dots + Q_t + \dots + Q_{t+5} + 0.5Q_{t+6}}{12}$$

... où MA12 est la moyenne mobile centrée d'ordre 12.

D'après les propriétés des moyennes mobiles et de décomposition [cf. Chapitre 3], cette transformation permet de filtrer la tendance provoquée par l'extraction cumulée de minerai S_t . Le reste peut alors être expliqué par la pluie efficace E_t . On écrit :

$$\frac{Q_t}{MA12} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \varepsilon_t \quad (6.9)$$

Afin de tester ce modèle, nous l'avons appliqué pour trois mines déjà traitées par le modèle (6.8) : Saizerais, Errouville et Droitaumont.

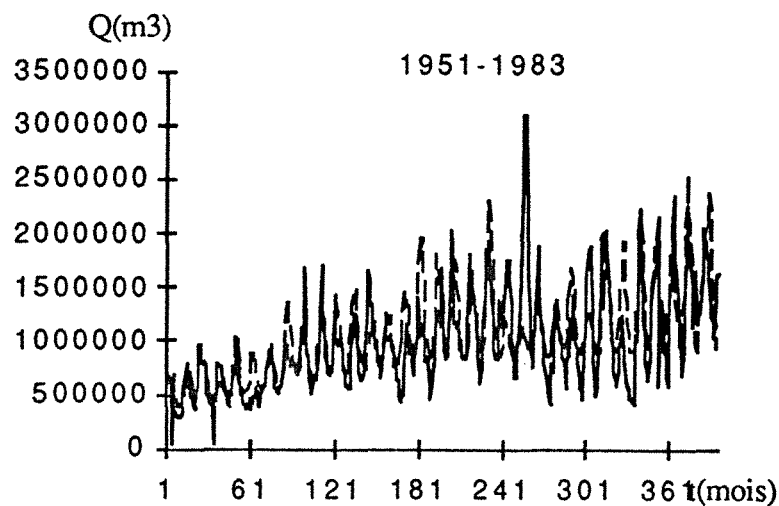
Pour la mine de Saizerais, par exemple, on pourra remarquer que le modèle (6.9) donne lieu à une bien meilleure reconstitution par rapport au modèle (6.8) (fig. 53).

Fig. 52 - Déconvolution à Anderny ($\frac{Q_t}{S_t^{0,3}} - E_t$)

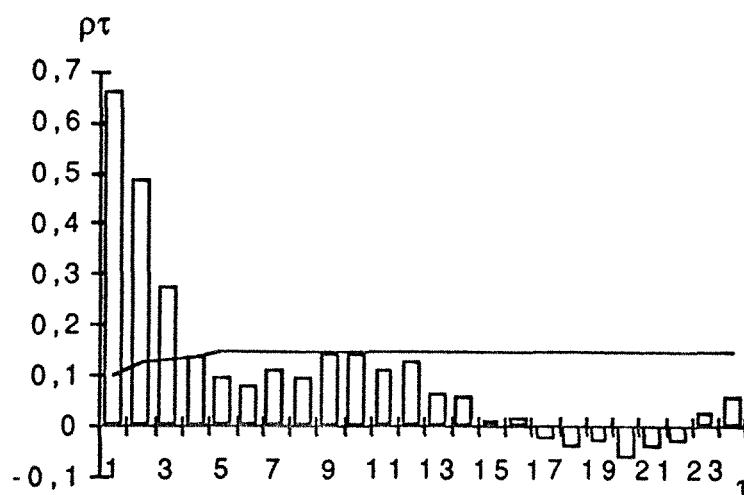
Dependent variable	coefficient	std.error	t-value	sig. level
CONSTANT	4158,740852	115,31568	36,0640	0,0000
E	9,846044	2,70067	3,6458	0,0003
1 LAGE	29,80324	2,91321	10,2304	0,0000
2 LAGE	13,340947	2,888114	4,6193	0,0000
3 LAGE	12,8522	2,881245	4,4606	0,0000
4 LAGE	12,200502	2,906851	4,1972	0,0000
5 LAGE	5,374133	2,700093	1,9904	0,0473

R-SQ. (ADJ) = 0,5774 SE= 1416,185375 MAE= 1038,5219 Durbwat =0,653
 391 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

a - modèle de régression



b - exhaure mesurée et reconstituée



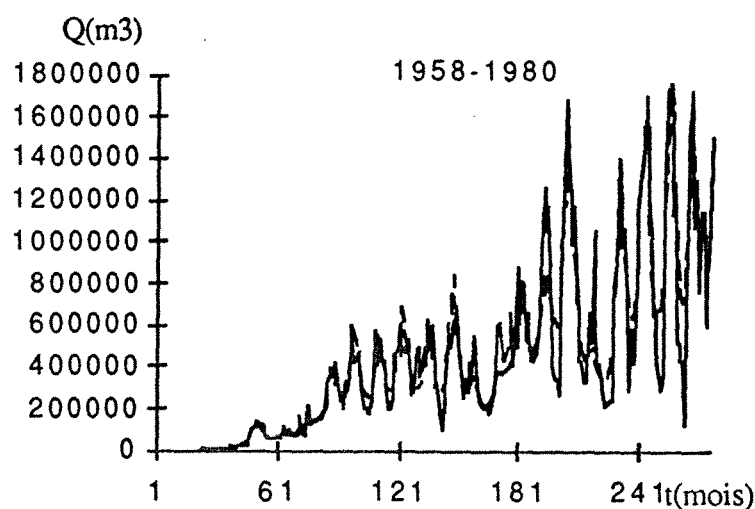
c - correlogramme des résidus

Fig. 53 - Déconvolution à Saizerais ($\frac{Q_t}{MA12} - E_t$)

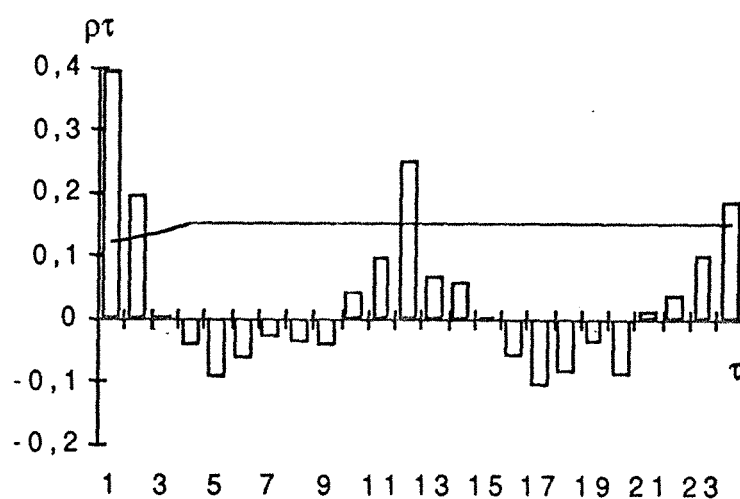
Dependent variable	coefficient	std.error	t-value	sig. level
CONSTANT	0,719301	0,028073	25,6229	0,0000
E	0,002713	0,000763	3,5579	0,0005
1 LAGE	0,006495	0,000801	8,1118	0,0000
2 LAGE	0,001882	0,00081	2,3235	0,0211
3 LAGE	0,002129	0,0008	2,6608	0,0084
4 LAGE	0,001312	0,000765	1,7154	0,0877

R-SQ. (ADJ) = 0,5494 SE= 0,276387 MAE= 0,211425 Durbwat =1,112
 216 observations fitted, forecast(s) computed for 56 missing val. of dep. var.

a - modèle de régression



b - exhaure mesurée et reconstituée



c - correlogramme des résidus

Enfin, notant que le coefficient de tempérament β peut être considéré comme zéro pour le premier type de mines, on pourra établir un modèle général pour toutes les mines dans le bassin ferrifère lorrain :

$$\frac{Q_t}{S_t \beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \dots + \phi_m E_{t-m} + \varepsilon_t$$

$$(0 \leq \beta \leq 1) \quad (6.10)$$

Par ailleurs, un modèle auxiliaire au modèle (6.10) peut être utilisé en cas de manque de données d'extraction pour les mines non stationnaires. Il peut s'écrire :

$$\frac{Q_t}{MA12} = C + \phi_0 E_t + \phi_1 E_{t-1} + \dots + \phi_m E_{t-m} + \varepsilon_t \quad (6.11)$$

Tous les résultats de déconvolution sont rassemblés dans le tableau VI.

mine	R ²	β	c	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9
Godbrange	0,59	0	1,44E5	2229	4320	638	944	504					
Kraémer	0,57	0	35908	630	1383	611	477	356	328				
Ottange	0,62	0	2,25E5	1914	5939	1888	808	858	685				
Bréchain	0,65	?	0,6915	0	0,0055	0,0047	0,0034	0,0013					
Bazailles	0,53	0	2,58E5	192	856	469	480	367	325				
Errouville	0,69	0,3	1881	10,8	22,9	10,6	9,7	7,8	5,4	4,1	5,5		
Errouville	0,5	?	0,7852	0	0,0047	0,0024	0,002	0,001					
Montrouge	0,56	0	3,09E5	1906	4922	1736	1080	1054					
Hettange	0,69	0	3,44E5	2542	5356	3156	2216	1472					
Angevillers	0,62	0	3,12E5	4537	8340	3725	3429	1930	2158	1858			
Ferdinant	0,52	?	0,71	0,0038	0,0045	0,002							
Paix	0,79	0	7,33E5	8603	14829	4808	3977	2565	1950				
Joudreville	0,40	?	0,8722	0	0,00197	0,00196	0,00089	0,00086	0,0003				
Amermont	0,29	?	0,9257	0	0	0,00084	0,00104	0,00105	0,0001	0,00064			
Piennes	0,8	0,2	3657	31	34	28	11	15	12				
Tucquenieux	0,80	0	5,09E5	4029	7107	3457	2643	1999	1317	832			
St-Pierremont	0,67	0,2	9874	113	200	73	67	38	40				
Anderny	0,58	0,3	4159	9,8	29,8	13,3	12,9	12,2	5,4				
Pier+Tuco+Ande	0,79	0,2	40754	228	471	205	170	133	91				
Hayange	0,67	0	4,66E5	2881	4473	1836	1522	1075	1020				
Moyeuivre	0,68	0	3,87E5	3359	9215	3737	2371	2120	1861				
Paradis	0,63	?	0,7688	0,002	0,0046	0,0023	0,0018						
Roncourt	0,68	0	8,48E5	10253	16030	7862	4457	4162	3475				
Giraumont	0,55	0	5,07E5	1229	2450	1346	1274	1022					
Droitaumont	0,57	0,18	46216	56	87	97	104	91	73	49	50	41	45
Droitaumont	0,55	?	0,9129	0	0,00024	0,00091	0,0012	0,00098	0,00057				
Saizerais	0,57	1	0,0285	0,00022	0,00032	0,00008	0,0001	0,00012	0,00006				
Saizerais	0,55	?	0,7193	0,0027	0,0065	0,0019	0,0021	0,0013					

R²: coefficient de corrélation multiple

β : coefficient de tempérament

? déconvolution avec transformation $\frac{Q_t}{MA12}$

Tableau VI - Résultats de déconvolution

Au vu de ce tableau, on constate que, pour les mines Amermont et Joudreville, les coefficients de corrélation multiple R² sont faibles (respectivement 0.29 et 0.40), ce qui signifie que les modèles de déconvolution représentent mal ces deux mines. En effet, ce sont des exploitations à grande profondeur et elles sont en majeure partie recouvertes par des affleu-

rements de marnes imperméables. Ces mines ne sont plus sous l'influence directe de la pluie efficace. Leur exhaure, conditionnée principalement par les réserves souterraines, présente un caractère quasiment constant si les influences "intermines" peuvent être écartées.

Afin de simplifier l'interprétation et la comparaison de différentes mines, nous présentons les coefficients de régression en pourcentage par le tableau VII :

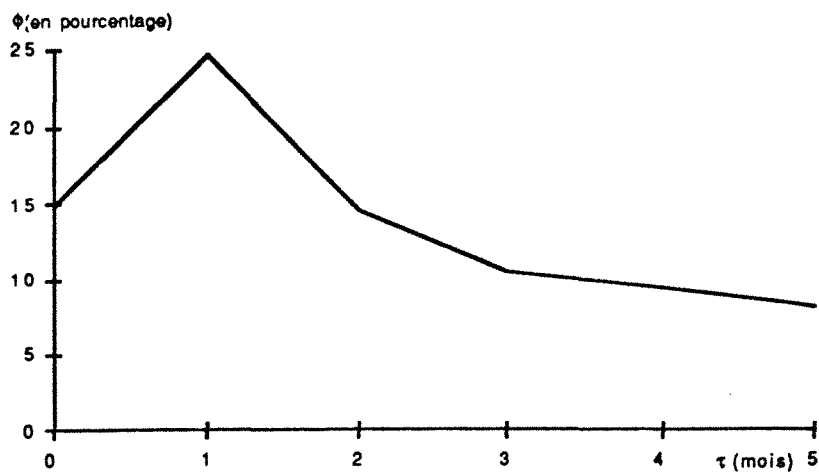
Mines	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9
Godbrange	25,8	50	7,4	10,9	5,8					
Kraemer	16,6	36,5	16,1	12,6	9,4	8,7				
Ottange	15,8	49,1	15,6	6,7	7,1	5,4				
Bréchain*	0	36,9	31,5	22,8	8,7					
Bazailles	7,1	31,8	17,4	17,9	13,6	12,1				
Errouville	14,1	29,8	13,8	12,6	10,2	7	5,3	7,2		
Errouville*	0	46,5	23,8	19,8	10					
Montrouge	17,8	46	16,2	10,1	9,8					
Hettange	17,2	36,3	21,4	15	10					
Angevillers	17,5	32,1	14,3	13,2	17,4	8,3	7,2			
Ferdinant	36,9	43,7	19,4							
Paix	23,4	40,4	13,1	10,8	7	5,3				
Joudreville*	0	32,5	32,3	16,1	14,2	4,9				
Amermont*	0	0	22,8	28,3	28,6	2,7	17,4			
Pierres	23,7	26	21,4	8,4	11,5	9,2				
Tucquenieux	18,8	33,2	16,2	12,4	9,3	6,2	3,9			
S ^l -Pierremont	21,3	37,7	13,7	12,6	7,2	7,5				
Anderny	11,8	35,7	15,9	15,5	14,6	6,5				
PTA	17,6	36,3	15,8	13,1	10,2	7				
Hayange	22,5	34,1	14,3	11,8	8,2	7,9				
Moyeuvre	14,8	24,6	14,5	10,5	9,4	8,2				
Paradis*	18,7	43	21,5	16,8						
Roncourt	22,2	34,6	17	9,6	9	7,5				
Giraumont*	16,8	33,5	18,4	17,4	14					
Droitaumont	8,1	12,6	14	15	13,1	10,5	7,1	7,2	5,9	6,5
Droitaumont*	0	6,2	23,3	30,8	25,1	14,6				
Saizerais	24,4	35,6	8,9	11,1	13,3	6,7				
Saizerais*	18,6	44,8	13,1	14,5	9					

*: déconvolution après transformation $\frac{Q_t}{MA12}$; PTA = S^l-Pierremont + Tucquenieux + Anderny

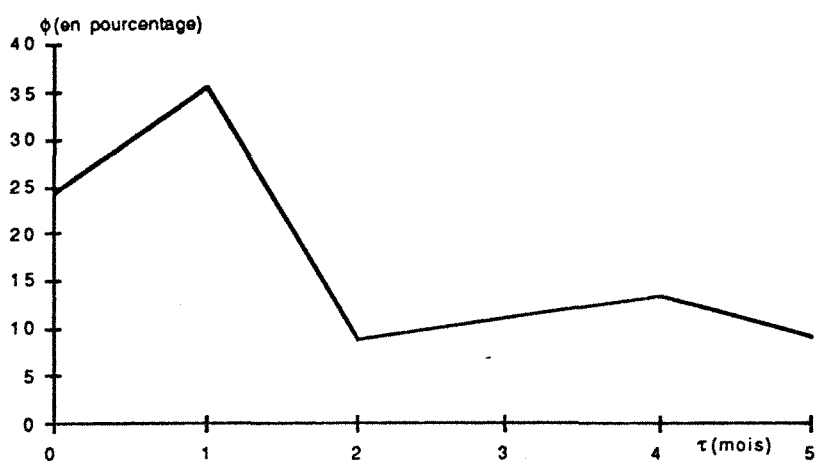
Tableau VII - Coefficients en pourcentage

Les fonctions de transfert, pour six mines réparties dans le bassin, sont schématisées en pourcentage par la figure 54. Nous constatons que ces fonctions tendent toutes vers zéro, sauf celle de la mine Amermont, pour des raisons citées ci-dessus.

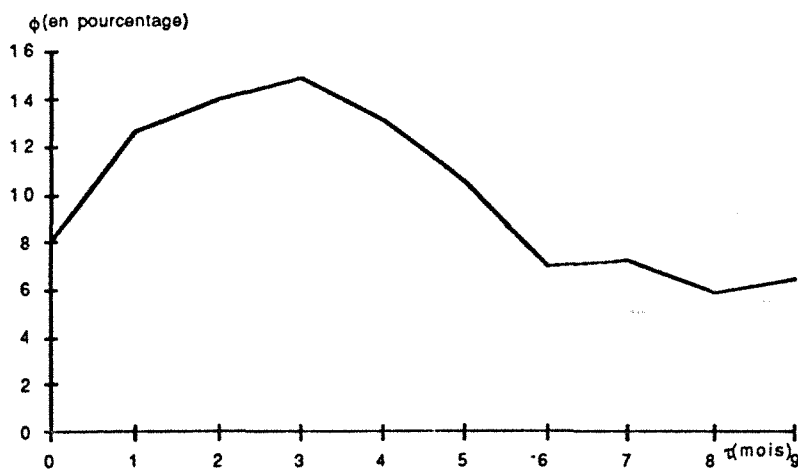
Fig. 54 - Fonction de transfert (régression multiple)



a - Moyeuve

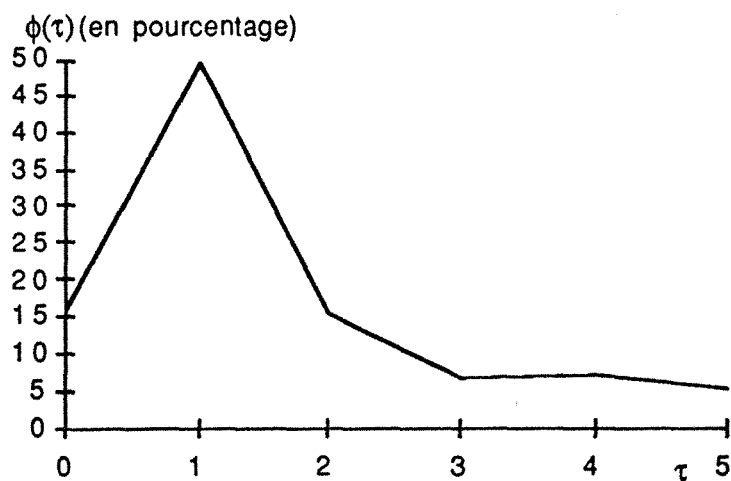


b - Saizerais

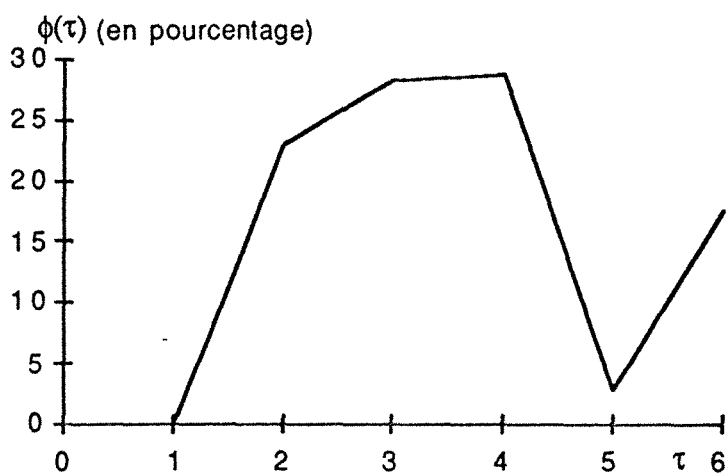


c - Droitaumont

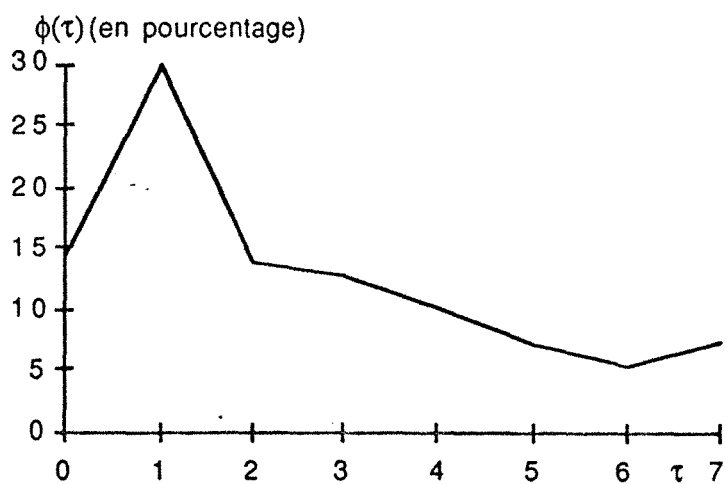
Fig. 54 - Fonction de transfert (régression multiple) (fin)



d - Ottange



e - Amermont



f - Errouville

III.3 - Discussion sur les termes des modèles établis

Reprenons le modèle général (6.10) :

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \epsilon_t \quad (0 \leq \beta \leq 1)$$

Nous allons analyser, en tenant compte du mécanisme d'exhaure et de l'hydrogéologie, les différents termes de ce modèle.

III.3.1 : Coefficient de tempérament β ($0 \leq \beta \leq 1$)

Pour toutes les mines du premier type, β est égal à zéro. Le tempérament d'une mine ne change pas avec la progression de l'exploitation minière dans le temps. L'exhaure est sous dépendance exclusive de la pluie efficace.

Pour les mines du deuxième type, le coefficient β varie d'une mine à l'autre ($0 < \beta \leq 1$). Il représente la réaction d'une mine vis-à-vis de l'exploitation minière. Ce coefficient pourrait être assimilé, en quelque sorte, au coefficient d'élasticité (module d'Young) de la mécanique des matériaux.

En tant que coefficient représentant un sens physique, sa valeur β ne peut pas être quelconque. L'étude de toutes les mines a effectivement montré que β prend une valeur dans l'intervalle $[0,1]$. Les cas $\beta < 0$ et $\beta > 1$ ne peuvent jamais être rencontrés. L'explication est la suivante (cf. fig. 55) :

Si $\beta < 0$, le modèle (6.10) peut s'écrire...

$$Q_t = \frac{C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \epsilon_t}{S_t^{|\beta|}}$$

... et cette expression implique que Q_t diminue lorsque S_t augmente.

Si $\beta > 1$, le modèle (6.10) conduit à obtenir des valeurs Q_t qui croissent indéfiniment avec t .

Or, dans la pratique, Q_t ne peut ni croître indéfiniment dans le temps, ni diminuer lorsque S_t augmente. Il est donc certain que le coefficient de tempérément β se définit strictement dans l'intervalle $[0,1]$.

Les principaux facteurs qui conditionnent le coefficient de tempérément β sont les suivants :

- conditions géologiques, structurales et hydrogéologiques au fond des mines ;
- politique des dépilages et taux de "défruitement" ;
- l'épaisseur de la couche ferrugineuse exploitée.

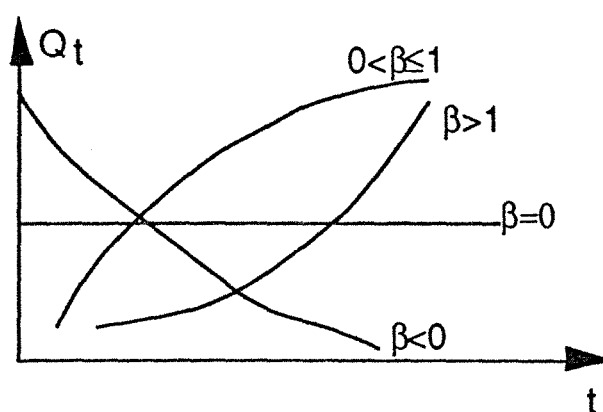


Fig. 55 - Exhaure et le coefficient de tempérément β

III.3.2 : Terme constant C , mémoire m et fonction de transfert

Dans les applications de la convolution pour la prévision des étiages dans des bassins hydrologiques, ce terme peut être considéré comme un parasite, dont la prise en compte améliore la qualité de la reconstitution, mais présente souvent un handicap pour la simulation des bas débits [79].

En l'absence de pluie efficace, on obtient après un certain temps :

- si la constante est négative: un débit négatif, ce qui n'est pas acceptable ;
- si la constante est positive: un débit minimal qui se maintient même en l'absence prolongée de pluie efficace.

Il est possible de remédier à ce défaut en imposant un terme constant nul au prix d'une petite perte de qualité de la reconstitution.

Cependant, l'exhaure minière entraîne la vidange d'une nappe importante (la nappe du Dogger). L'effet des réserves souterraines n'est donc pas négligeable en réalité.

La mémoire m signifie que la pluie efficace E_t à une date antérieure à $t-m$ n'a pas d'influence sur Q_t . Elle est liée notamment :

- à la topographie et aux conditions hydrologiques de surface ;
- à la profondeur des mines ;
- aux caractéristiques hydrogéologiques des morts-terrains.

Souvent en hydrologie, la mémoire réelle du système est plus longue que la mémoire estimée selon différents critères. Outre la première contrainte (les coefficients de régression soient strictement égaux ou supérieurs à zéro), si l'on admet également la deuxième (les coefficients de Student soient généralement égaux ou supérieur à deux), comme on a fait jusqu'à présent, la mémoire obtenue peut être appelée "*mémoire significative*". Mais il faut remarquer que l'on peut laisser de côté la deuxième contrainte pour les problèmes de déconvolution, car notre intérêt est d'obtenir une meilleure reconstitution ou prévision, et non plus de tester si une variable explicative influence significativement la variable expliquée.

On va négliger maintenant la deuxième contrainte pour la mine Moyeuivre. La fonction de transfert calculée (le calcul a été effectué en utilisant un programme de déconvolution réalisé sur Macintosh par Monsieur DEMASSIEUX) avec un modèle comportant une constante C s'annule vers 9 mois, tandis que la fonction calculée avec un modèle sans constante C résiste même à 12 mois (fig. 56), ce qui signifie un gain important de précision lorsque une constante C rentre en ligne de compte.

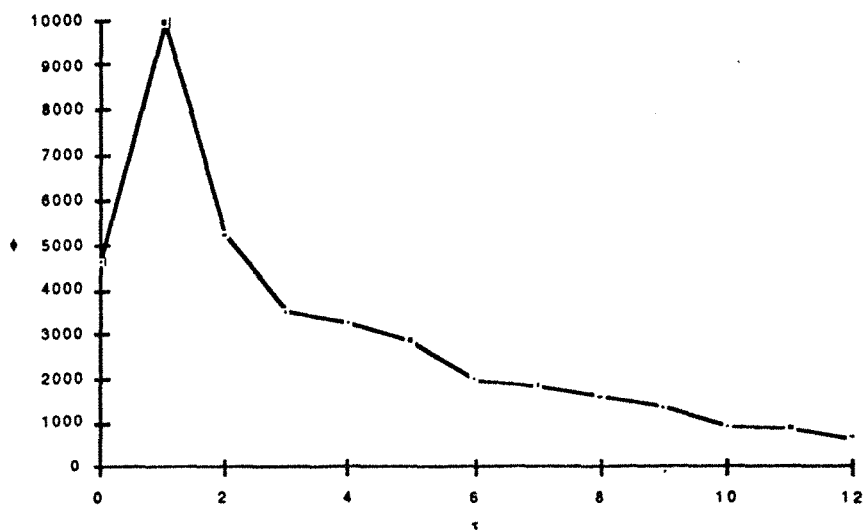
La fonction de transfert, étant une exponentielle à partir d'un moment T_0 [60,79], est souvent interprétée comme un hydrogramme unitaire par des hydrologues (fig. 57) :

$$\phi_{\tau+\Delta\tau} = e^{-\alpha\Delta\tau}\phi_{\tau} \quad (\alpha \text{ est le coefficient de tarissement}) \quad (6.12)$$

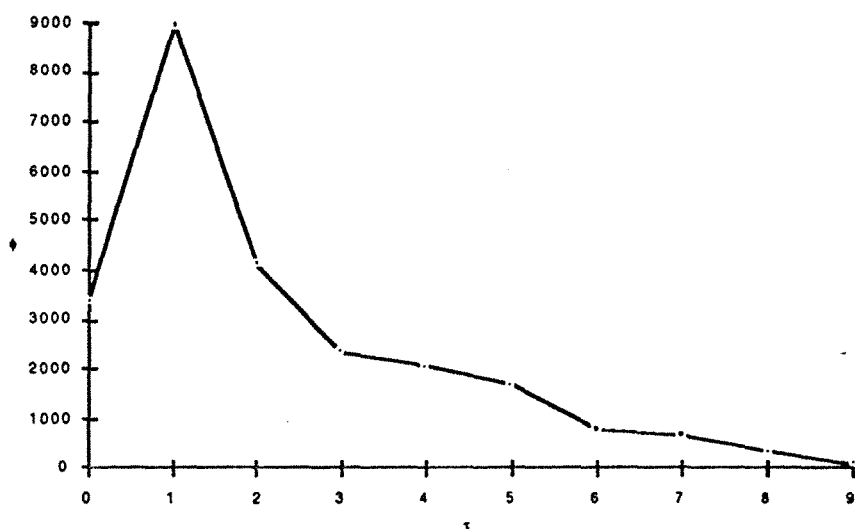
Pour la mine Moyeuivre, T_0 est d'un mois. Un ajustement sur les données $\phi_1, \phi_2, \dots, \phi_9$ donne lieu à un modèle (fig. 58)...

$$\ln(\phi_{\tau}) = 9,52 - 0,49\tau \quad (\rho = -0,97) \quad (6.13)$$

... et le coefficient de tarissement est égal à $0,49 \text{ mois}^{-1}$.



a - Sans terme constante



b - Avec terme constante

Fig. 56 - Fonction de transfert (mine Moyeuve)

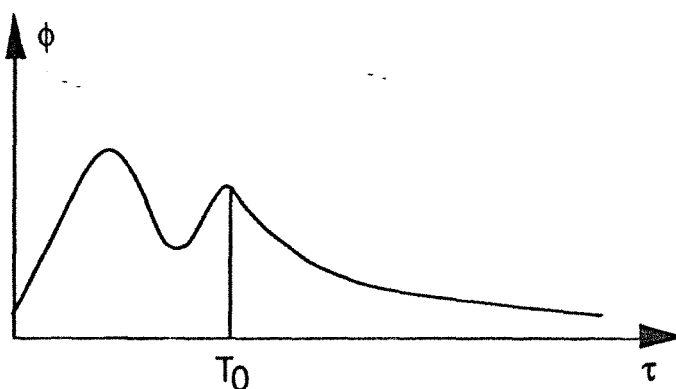


Fig. 57 - Fonction de transfert

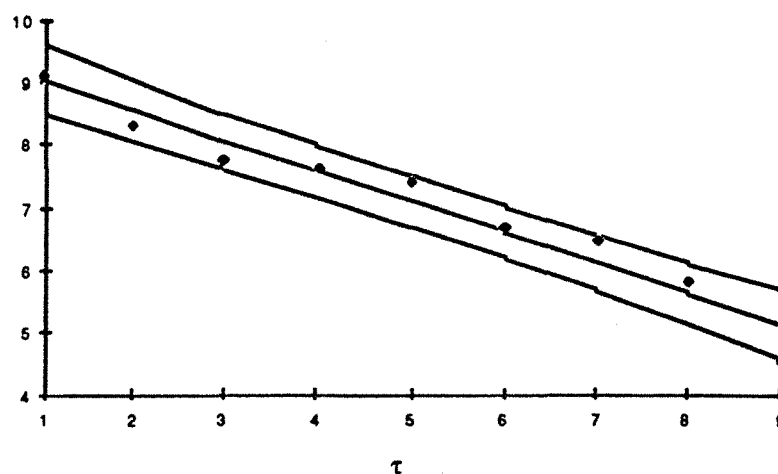


Fig. 58 - Ajustement de la fonction de transfert à une courbe exponentielle (mine Moyeuve)

Pourtant, pour la mine Saizerais, la fonction de transfert semble moins stable (fig. 59). Cela peut s'expliquer par la superposition de plusieurs régimes de circulation.

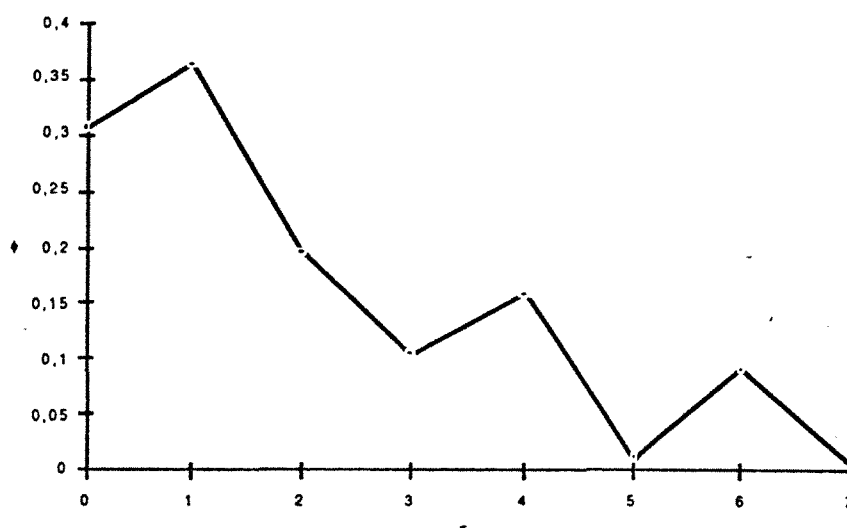


Fig. 59 - Fonction de transfert (mine Saizerais)

III.3.3 : Terme résiduel ε_t

Les résidus du modèle (6.10) ne constituent pas un bruit blanc. Pour la mine Moyeuve, par exemple, le corrélogramme des résidus (cf. fig. 46c) montre à l'évidence qu'il existe une certaine d'inertie, qui peut être expliquée par plusieurs facteurs :

- la pluie efficace n'est pas parfaitement estimée: sa sous-estimation d'un mois entraîne celle de l'exhaure des m mois suivants (résidu positif et autocorrélé). De plus, bien que la répartition des pluies brutes au cours d'un mois influence l'exhaure, le fait de calculer la pluie efficace à l'échelle mensuelle ignore ce phénomène[79] ;

- les observations d'exhaure sont entachées d'erreurs ;

- les hypothèses de linéarité et d'invariance par translation dans le temps ne sont pas parfaitement vérifiées.

L'inertie stockée dans les résidus n'est pas précisément explicable. Dans le but de modifier le modèle de prévision, et sans savoir ce qui se passe dans la "boîte noire", on fait appel à la méthode de Box-Jenkins.

Pour la mine Moyeuivre, les résidus constituent un processus autorégressif d'ordre 1 (processus Markovien fig. 60), on écrit alors :

$$(1 - 0.57 B) \varepsilon_t = a_t$$

... ou bien
$$\varepsilon_t = 0.57 \varepsilon_{t-1} + a_t \quad (6.14)$$

a_t étant un bruit blanc. Cette approche conduit à un modèle corrigé tel que :

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \dots + \phi_m E_{t-m} + 0.57 \varepsilon_{t-1} + a_t \quad (6.15)$$

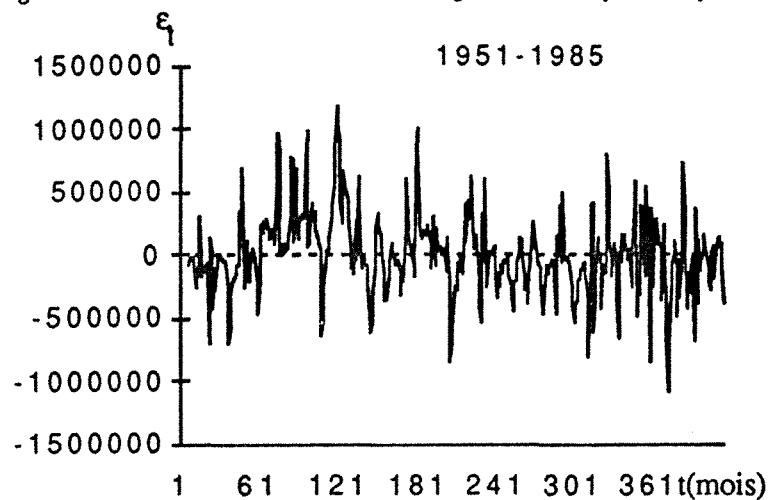
Pour les autres mines, on constate que le modèle des résidus est souvent du type AR(1), et parfois AR(2). Le modèle corrigé peut être exprimé symboliquement par :

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + f(\varepsilon_t) + a_t \quad (6.16)$$

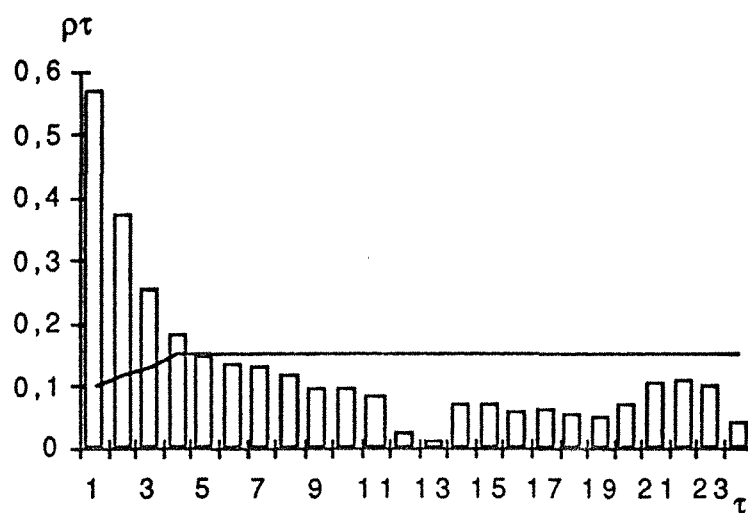
Pour le modèle auxiliaire (6.11), la même démarche peut être effectuée et on aura :

$$\frac{Q_t}{MA12} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + f(\varepsilon_t) + a_t \quad (6.17)$$

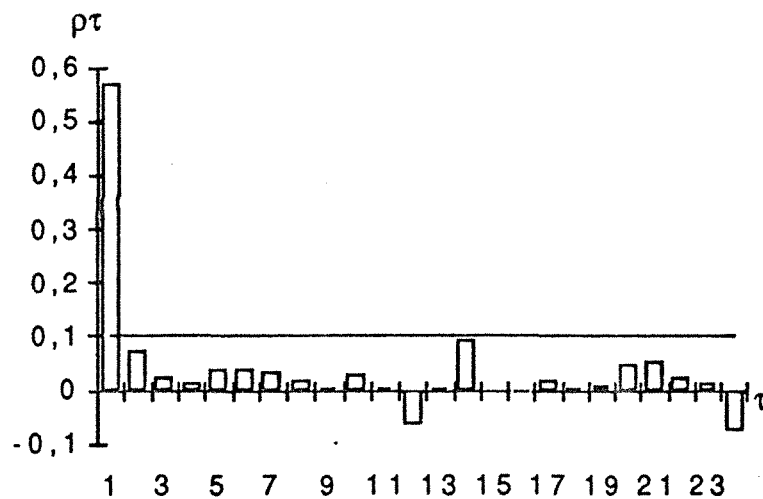
Fig. 60 - Modèle ARIMA des résidus de régression multiple à Moyeuvre



a - résidus de régression



b - correlogramme



c - correlogramme partiel

Fig. 60 - Modèle ARIMA des résidus de régression multiple à Moyeuvre (fin)

Parameter	Estimate	Std.error	T-value	P-value
AR(1)	0,57046	0,04044	14,10475	0,00000

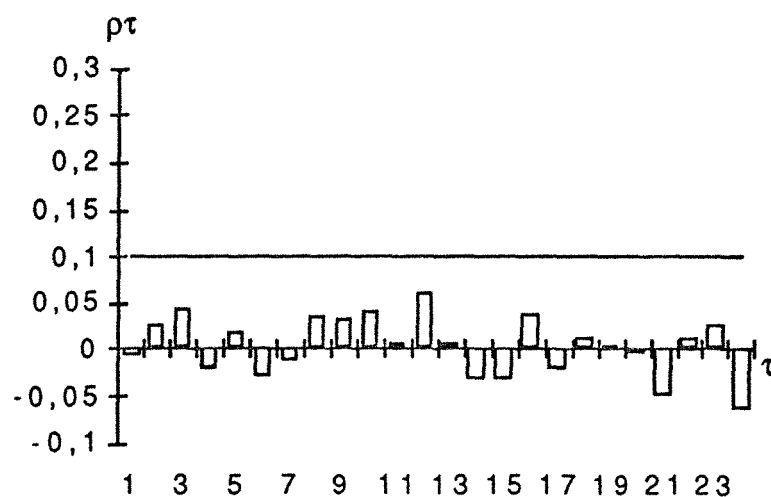
Estimated white noise variance = 6,60058E10 with 414 degrees of freedom.

Estimated white noise standard deviation (std err) = 256916.

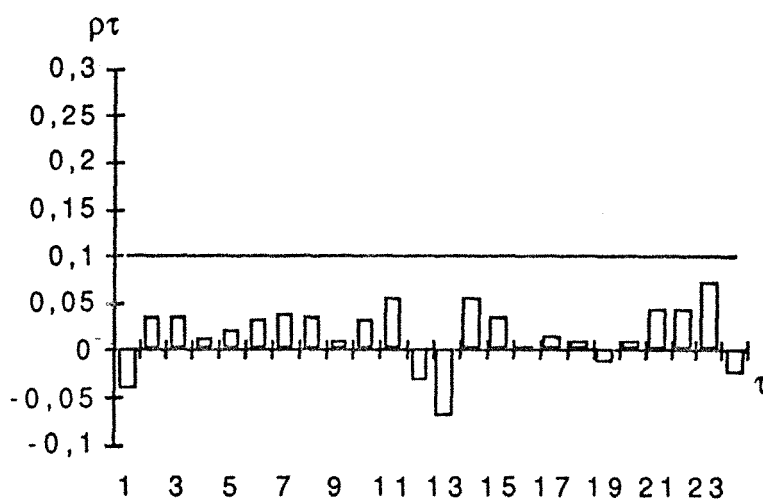
Chi-square test statistic on first 20 residual autocorrelations = 10,0104.

with probability of a larger value given white noise = 0,952689.

d - modèle ARIMA



e - corrélogramme des résidus



f - corrélogramme partiel des résidus

En conclusion, nous pouvons faire les remarques suivantes :

* Dans la pratique, on se contente de la prévision du modèle général (6.10) et de son auxiliaire (6.11). S'il en est besoin, on pourra faire appel aux modèles (6.16) et (6.17), mais leur application est certainement plus complexe et plus coûteuse.

* Notre démarche de modification repose essentiellement sur le principe des méthodes MARIMA qui combinent les concepts de régression multiple à la méthode de Box-Jenkins.

* Un système expert, appelé régression dynamique, recherchant une fonction de transfert de façon à faire intervenir le terme résiduel dans le calcul, fournit évidemment une prévision plus précise par rapport à la méthode de régression multiple. Pour la mine Saizerais, cette méthode conduit à des résidus purement aléatoires (fig. 61) et le modèle s'écrit:

$$\begin{aligned} \text{Log} \frac{Q_t}{S_t} = & -3,84 + 0,0052E_t + 0,0059E_{t-1} + 0,0017E_{t-2} + 0,0029E_{t-4} \\ & + 0,0014E_{t-5} + 0,938a_{t-1} + a_t \end{aligned}$$

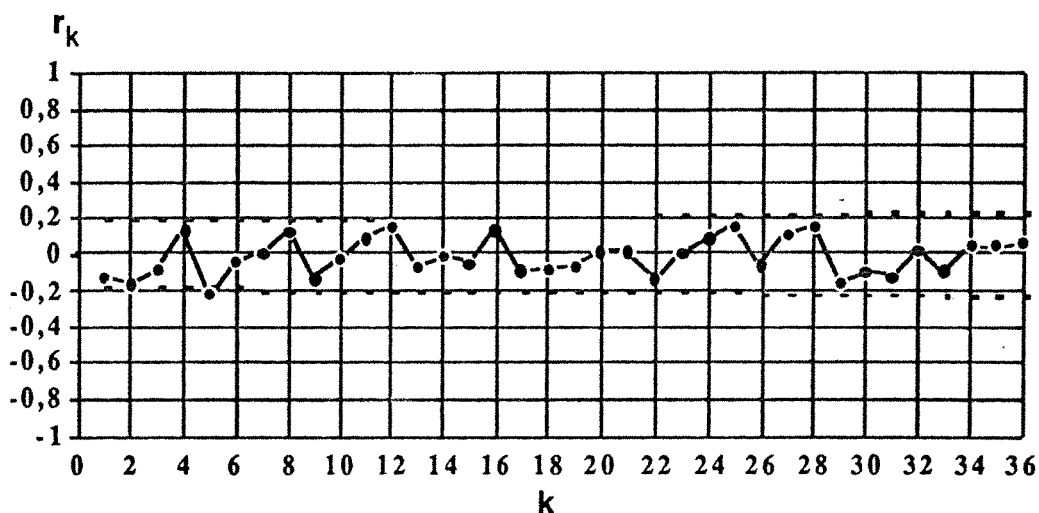


Fig. 61 - Corrélogramme des résidus calculés par régression dynamique (mine Saizerais)

IV - COMPARAISON DE DIFFERENTES METHODES DE DECONVOLUTION

Les différentes méthodes de déconvolution ont fait l'objet du quatrième chapitre. Il convient de souligner que :

La méthode de projection [60] a été mise au point à partir des considérations sur la stabilité physique des solutions. La recherche de la fonction de transfert se fait par projection

successive sur une base de vecteurs particuliers hiérarchisés (base de Walsh). Du fait de cette succession, des contraintes ou des lissages de la fonction de transfert cherchée peuvent être introduits pas à pas. La stabilité de cette fonction de transfert est ainsi garantie et il est toujours possible de proposer une fonction de transfert qui réalise un compromis entre plausibilité physique et norme d'ajustement.

Cette méthode est principalement utilisée en hydrologie: fonction de transfert d'un bassin ou d'une nappe, mouvement de l'eau en zone non saturée, interprétation des diagraphies, *etc.*

D'autre part, la méthode de déconvolution géostatistique [70] repose sur le calcul et modélisation préalable des fonctions de covariance des signaux d'entrée et de sortie, puis sur la résolution de l'équation de régularisation qui relie ces deux covariances. La résolution peut se faire numériquement par la transformation de Fourier ou par méthodes quadratiques sous contraintes linéaires. Les avantages de cette méthode sont :

- Les fonctions de covariance fournissent de précieuses informations sur la structure des données ;

- Ajustement à un modèle de covariance filtre des erreurs d'estimation ;

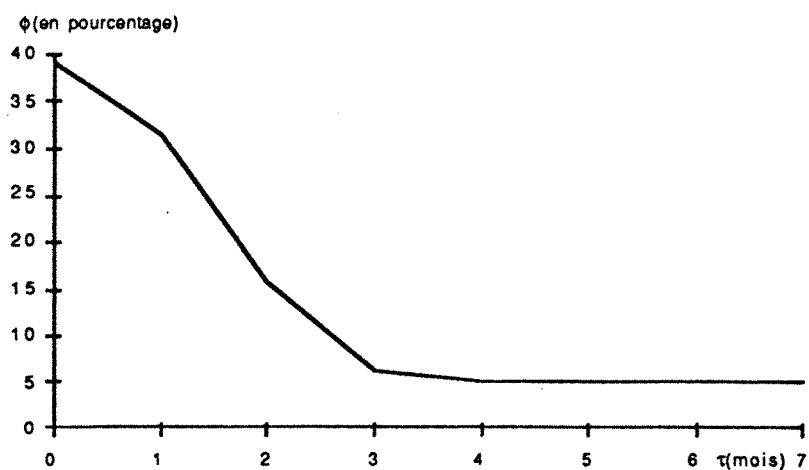
- Dans le cas où le modèle de covariance est de type exponentiel ou gaussien, il est possible de calculer la fonction de transfert sous forme analytique.

La méthode de régression multiple, qui appartient aux méthodes classiques, a été largement utilisée dans la modélisation décrite au paragraphe précédent. Afin de comparer cette méthode classique avec la méthode de projection, nous avons appliqué celle-ci à quatre mines : Saizerais, Moyeuve, Errouville et Droitaumont. Les résultats des déconvolutions sont résumés dans le tableau VIII. Les fonctions de transfert sont illustrées par la figure 62a, b, c, d et les prévisions (pour la mine Saizerais) de ces deux méthodes par la figure 62e.

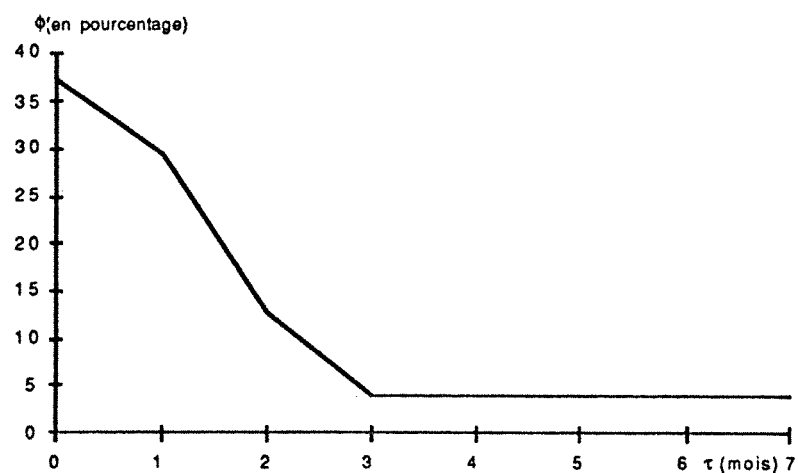
mine	β	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	ϕ_9	ϕ_{10}	ϕ_{11}	ϕ_{12}
Saizerais	1	0,00055	0,00043	0,00019	0,00006	0,00006	0,00006	0,00006	0,00006					
Moyeuve	0	9986	8081	4004	1569	1304	1304	1304	1304					
Errouville	0,3	19,87	19,87	17,05	11,42	8,44	8,11	7,94	7,94	7,94	7,94	7,94	8,96	12,04
Droitaumont	0,18	226,72	226,72	222,72	214,71	208,14	200,34	189,89	182,1	177,87	174,55	172,89	179,14	197,88

Tableau VIII - Déconvolution par projection

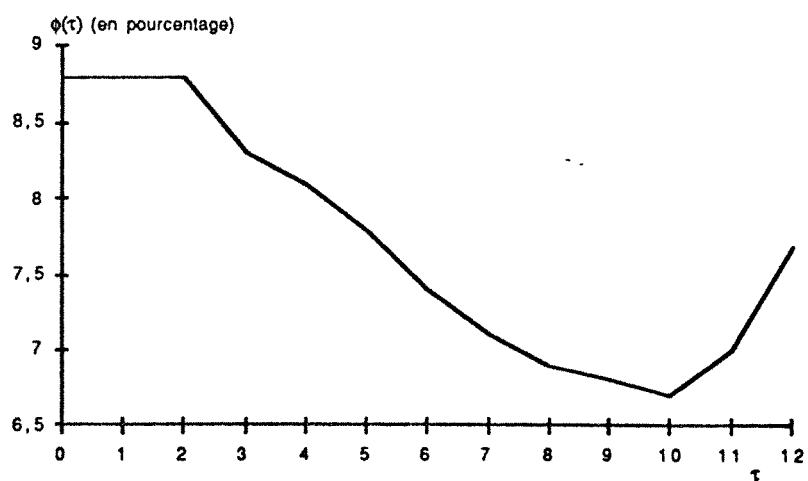
Fig. 62 - Fonction de transfert (déconvolution par projection)



a - Moyeuvre

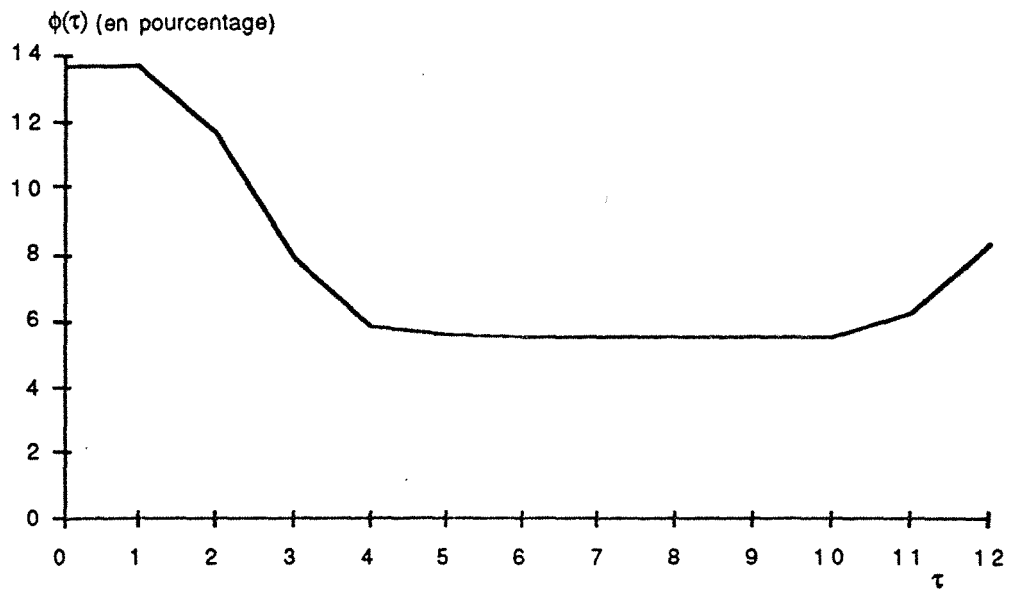


b - Saizerais

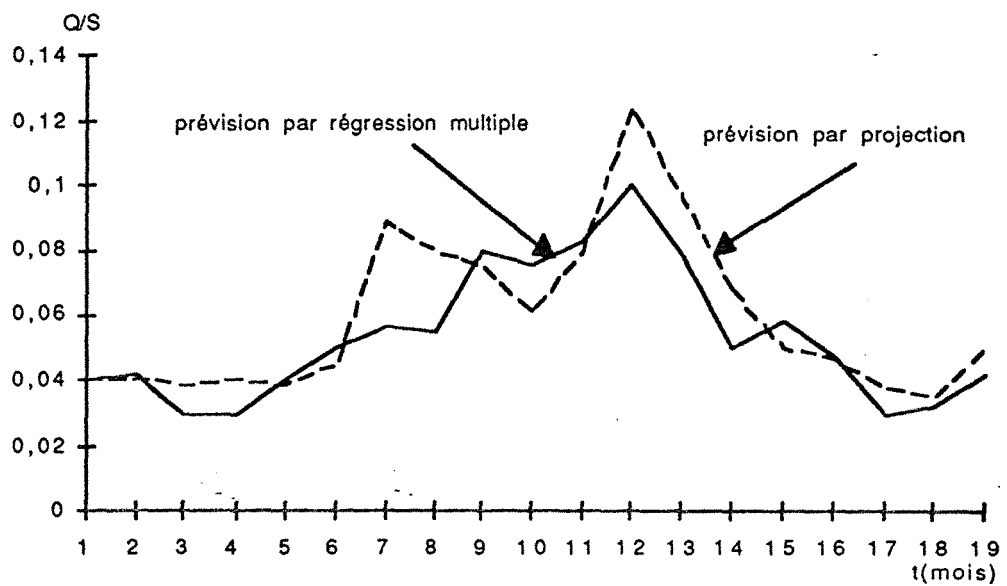


c - Droitaumont

Fig. 62 - Fonction de transfert (déconvolution par projection) (fin)



d - Errouville

e - prévision de l'exhaure à Saizerais
(Juin 1982 - Décembre 1983)

On peut conclure que :

* Du point de vue prévision, les résultats obtenus par les deux méthodes sont proches et ces deux méthodes se valident mutuellement. Pourtant lorsque l'on regarde les fonctions de transfert, nous avons une préférence pour la méthode de régression multiple car elle aboutit à des fonctions qui tendent vers zéro. La méthode de régression est par conséquent celle qui est la mieux adaptée à notre problème.

* Par rapport aux autres méthodes de déconvolution, la méthode de régression est facile à mettre en oeuvre. Les programmes de calcul fournissent les résidus et leur examen constitue un diagnostic du modèle adopté et aboutit à une éventuelle modification de celui-ci.

* La méthode de déconvolution par projection a remporté un grand succès dans les applications en hydrologie, mais elle ne s'adapte pas très bien à notre cas d'étude et cela est dû à deux raisons :

. cette méthode (ou le programme utilisé) ne permet pas de tenir compte du rôle des réserves souterraines ;

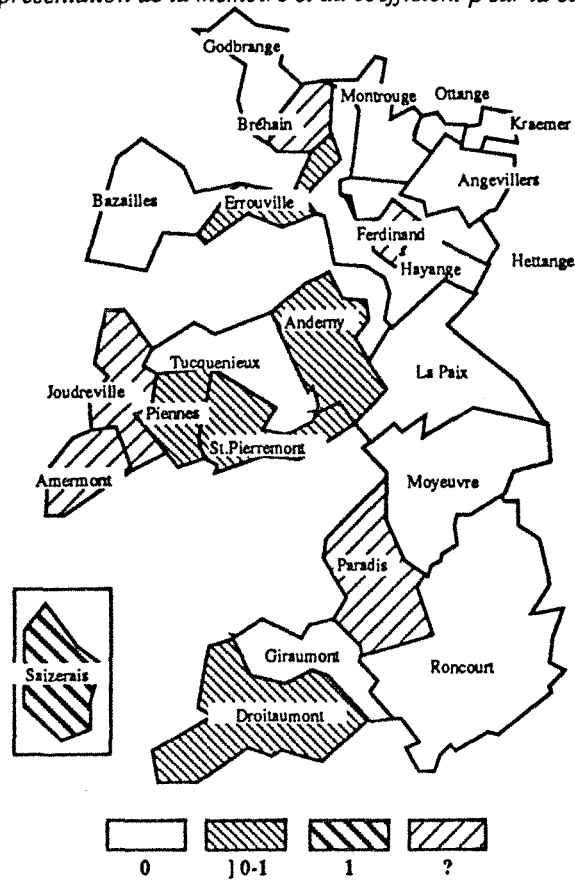
. nous travaillons à l'échelle mensuelle et la mémoire du système est courte, le problème de stabilité des solutions physiques n'est pas primordiale. La méthode de projection semble être trop contraignante vis-à-vis de la forme supposée de la fonction de transfert.

V - INTERPRETATION DES RESULTATS ET CONCLUSION

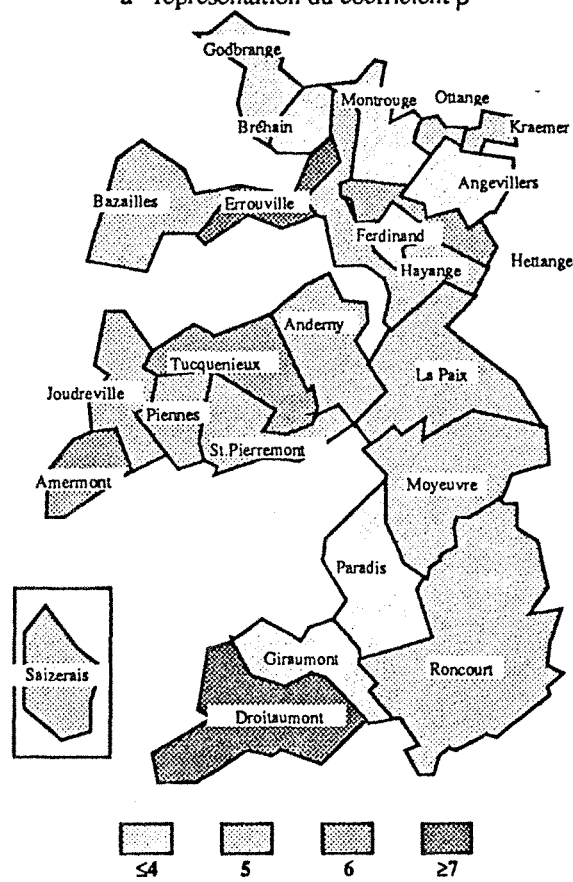
Les paramètres caractérisant une mine vis-à-vis de son exhaure sont la mémoire m , la fonction de transfert ϕ ainsi que le coefficient de tempérament β . Nous avons déjà illustré la fonction de transfert pour six mines par la figure 54. Nous représentons maintenant la mémoire significative m et le coefficient β sur la carte des mines (fig. 63), ce qui nous permet de constater que :

* L'exhaure des mines d'aval pendage a vraisemblablement plus de chances d'être non stationnaire que celles d'amont pendage (nous rappelons que le pendage se dirige du Nord- Est vers le Sud-Ouest ; cf. Fig. 35). La mémoire significative m est en général de 4 ou 5 mois et elle est relativement longue pour des mines d'aval pendage (10 pour Droitaumont, 6 pour Tucquenieux, 7 pour Errouville, etc.).

Fig. 63 - Représentation de la mémoire et du coefficient β sur la carte des mines



a - représentation du coefficient β



b - représentation de mémoire m

Ces constatations peuvent s'expliquer, entre autres, par les conditions topographiques et géologiques : les mines à l'Est ou au Nord (amont pendage) se situent très fréquemment à l'aplomb de plateaux isolés. La superficie d'alimentation potentielle est alors limitée et la profondeur d'exploitation est faible. La "*superficie d'alimentation effective*" atteint vite à la "*superficie d'alimentation potentielle*" (Fig.64). Par conséquent, l'exhaure en début de phase d'exploitation atteint un niveau autour duquel elle varie en fonction de la pluie efficace (cf. fig. 65a). Par contre, les mines à l'Ouest (aval pendage), s'exploitent à une grande profondeur et drainent sans doute une quantité d'eau d'origine plus lointaine. La "*superficie d'alimentation potentielle*" est importante et l'exhaure est en augmentation permanente, mais n'atteint pas la valeur limite jusqu'à la fermeture des mines. Si l'on reprend l'exploitation, on atteindra cette valeur limite un jour et au-delà, l'exhaure deviendra alors stationnaire (cf. fig.65b). On cite un phénomène comparable en géophysique [15] : à l'échelle de temps de quelques secondes, la glace a un comportement très proche de l'élasticité parfaite, si l'on l'observe sur quelques heures ou quelques jours, on s'aperçoit qu'elle flue lentement en présence d'une contrainte déviatorique. (La figure 65 schématise l'évolution de l'exhaure dans le temps : lors du traçage, la formation ferrugineuse demeure un petit aquifère indépendant; au cours de l'avancement des travaux, on produit une saignée de terrains, l'exhaure augmente d'abord très vite, pour atteindre une limite, qui n'est ensuite dépassée que très lentement. Lors de l'exploitation, l'exhaure se comporte différemment pour les différents types de mines).

* Les conditions structurales jouent un rôle également important. En effet, elles expliquent en partie les anomalies.

La mine Ferdinand et la mine Errouville se situent au Nord-Est du bassin, position favorable pour une exhaure stationnaire, leur exhaure est néanmoins non stationnaire, et la mémoire significative d'Errouville est longue (7 mois). La mine Ferdinand est en effet en position d'une cuvette coincée par deux failles, une structure assez semblable à celle de Saizerais. Quant à la Mine Errouville, elle est en position synclinale, bloquée par la faille d'Autun-le-Tiche et s'exploite à une grande profondeur.

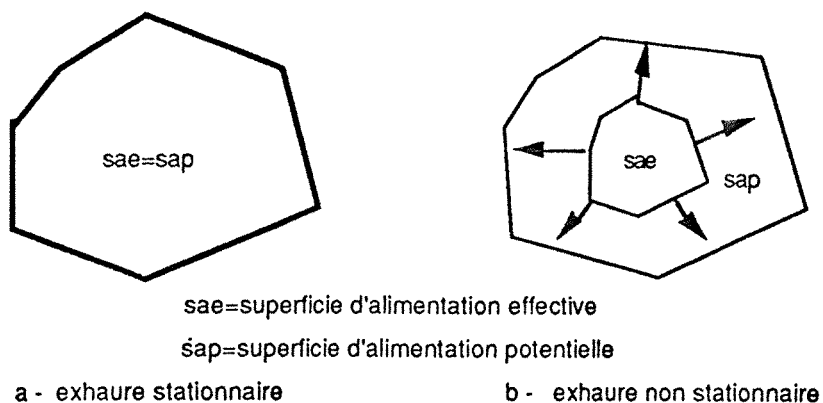


Fig. 64 - Superficies d'alimentation effective et potentielle

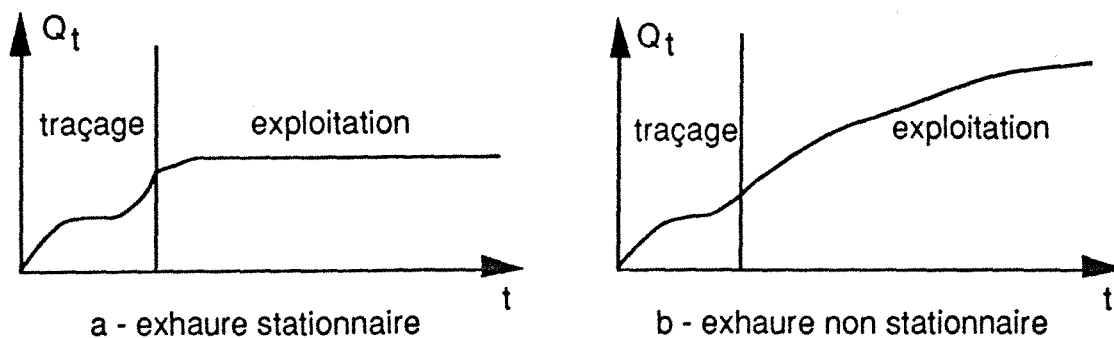


Fig. 65 - Evolution de l'exhaure dans le temps

* Il convient de noter que les analyses ci-dessus coïncident partiellement avec les analyses géochimiques des eaux exhaurées (la minéralisation sodique augmentent d'amont pendage en aval pendage assez systématiquement dans chaque domaine structural). Ce n'est pas un hasard : la quantité et la qualité des eaux exhaurées sont conditionnées toutes les deux par la circulation hydrodynamique souterraine.

* Dans l'ensemble, l'exhaure minière du bassin ferrifère lorrain présente une tendance d'augmentation en moyenne et en variance, tandis que la pluie efficace est quasiment stationnaire (fig. 66).

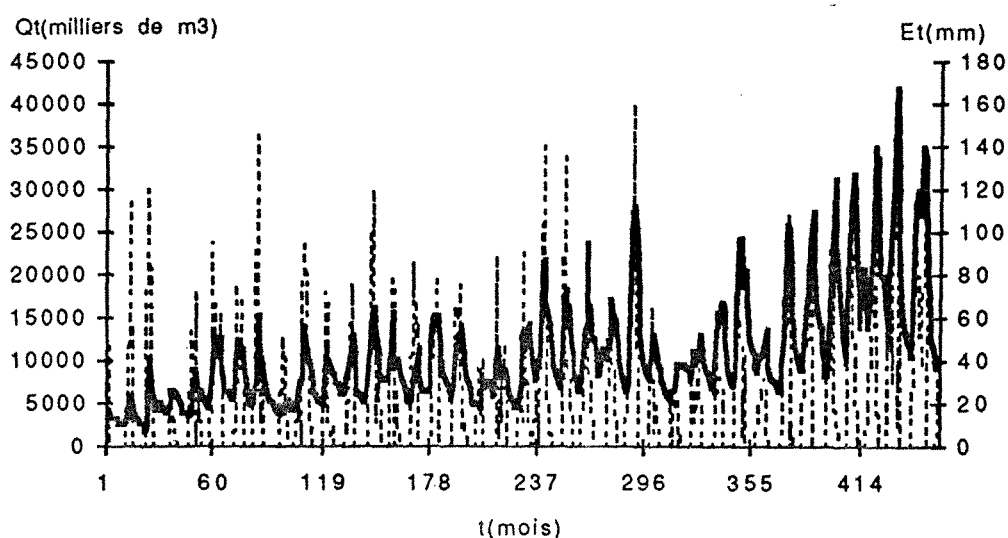


Fig. 66 - Comparaison entre la pluie efficace mensuelle (à Frescaty) et l'exhaure mensuelle totale dans le bassin (1946-1983)

En conclusion de ce chapitre, nous avons d'abord présenté les données et nous avons mis en évidence l'influence de l'extraction cumulée sur l'exhaure minière. Ensuite, nous avons insisté sur la modélisation de l'exhaure en tenant compte du rôle de l'extraction cumulée des minerais et de celui des réserves souterraines. Un modèle général a été établi. Nous avons aussi fait une comparaison entre deux méthodes de déconvolution (méthode régression multiple et méthode de déconvolution par projection). Enfin, l'interprétation des résultats dans un contexte géologique nous a permis de mieux connaître et d'améliorer la compréhension du phénomène exhaure.

Chapitre 7

S O M M A I R E

I	Quelques remarques.....	144
II	Bilan des résultats.....	145
III	Perspectives.....	147

CHAPITRE 7

CONCLUSION GENERALE

La première partie de ce mémoire a été consacrée au problème d'analyse des chroniques, et la seconde partie au problème d'exhaure minière dans le bassin ferrifère lorrain.

Dans cette conclusion générale, nous allons présenter tout d'abord quelques remarques sur le problème d'analyse des chroniques, et ensuite, un bilan des résultats apportés par l'analyse des données d'exhaure minière.

L'analyse des données d'exhaure minière a permis de décrire et de quantifier le phénomène exhaure dans l'espace et dans le temps. Les résultats apportés s'avèrent très intéressants dans le but d'aménagement du bassin.

Nous allons enfin proposer des perspectives relatives à l'exploitation de la "banque" des données d'exhaure pour la gestion de la future ressource lorsque l'ennoyage des mines sera en voie de généralisation.

I - QUELQUES REMARQUES

SUR LE PROBLEME D'ANALYSE DES CHRONIQUES

* Une chronique se comporte différemment selon l'échelle de temps, en raison notamment de "l'effet mémoire" du phénomène étudié. Dans l'analyse des chroniques, il faut donc toujours préciser l'échelle de temps à laquelle nous travaillons.

Pour un phénomène étudié, existe toujours une échelle de temps mieux adaptée que l'on note f_m . D'autre part, on note f_t l'échelle de temps à laquelle on travaille. Si $f_t > f_m$, on risque d'introduire une perte de l'information contenue dans le signal. Par contre, si $f_t < f_m$, il n'y a plus guère de gain en précision. Il convient de noter que cette échelle de temps f_m peut être déterminée par analyse spectrale (fréquence de Naquist).

* Une chronique peut être non stationnaire en moyenne et/ou en variance. Une chronique non stationnaire se comporte différemment sous représentation locale et sous représentation globale. L'étude d'une chronique consiste notamment à la mise en évidence de ses comportements à court, moyen et long termes.

* Les outils descriptifs présentés dans le deuxième chapitre sont très efficaces et utiles dans le but de diagnostic des modèles. En effet, les analyses autocorrélatrice et spectrale sur le terme résiduel d'un modèle apportent souvent des informations importantes sur la structure du résidu, ce qui peut suggérer une éventuelle modification du modèle.

* Pour un phénomène étudié, le choix d'une méthode de prévision appropriée est plutôt heuristique que mathématiquement optimal, car les qualités d'une méthode peuvent varier selon le type des données traitées, le choix du découpage du temps (échelle de temps), et même selon la longueur des signaux. Il est donc nécessaire de s'appuyer sur des comparaisons empiriques de performances et d'analyse de comportements mis en évidence par des simulations.

* La description d'une chronique est une étape préliminaire à toute modélisation et prévision. La prévision présente souvent l'intérêt principal de la simulation numérique et de plus, elle n'est pas une fin en soi mais une aide à la prise de décision par les éclairages qu'elle fournit.

* Le principe d'analyse des chroniques est basé sur les théories du système et du traitement du signal. Le système étudié est considéré comme une "boîte noire", et il est difficile de saisir des informations sur ce qui se passe à l'intérieur du système.

II - BILAN DES RESULTATS

* L'évolution de l'exhaure dans le temps peut être décrite de la manière suivante (cf. Fig. 65, page 142) :

Lors du traçage, on produit une saignée de terrains ; l'exhaure augmente d'abord très vite, pour atteindre une limite qui n'est ensuite dépassée que très lentement. Le régime de ces eaux de traçage est pratiquement constant et on ne distingue aucune loi dans leurs faibles variations.

Lors de l'exploitation, pour des mines, l'exhaure augmente au début pour atteindre un niveau et par la suite, elle varie toujours autour de ce niveau selon la variation saisonnière de la pluie. Par contre, pour les autres mines, l'exhaure augmente constamment sans atteindre la limite jusqu'à la fermeture des sièges. D'où deux types de mines peuvent être distingués : les mines stationnaires et les mines non-stationnaires. Si l'on reprend l'exploitation, l'exhaure non stationnaire devra atteindre sa valeur limite un jour, et au-delà, elle sera stationnaire.

* La non stationnarité de l'exhaure dans le temps est provoquée par l'extraction cumulée de minerai : au fur et à mesure que progresse l'exploitation minière, les communications entre la mine et les nappes de calcaires bajociens se multiplient et les eaux exhaurées proviennent de plus en plus loin ("*la superficie d'alimentation effective*" s'élargit). Il en résulte l'augmentation continue de l'exhaure.

* L'hétérogénéité de l'exhaure dans l'espace peut être généralement expliquée par les conditions géologiques locales: les mines d'aval pendage ont très vraisemblablement plus de chances d'être non stationnaires par rapport aux mines d'amont pendage, car elles possèdent une immense "*superficie d'alimentation potentielle*". Par ailleurs, les conditions structurales peuvent expliquer en partie les anomalies. Il faut noter que ces conclusions coïncident partiellement avec les analyses géochimiques des eaux exhaurées. Ce n'est pas un hasard, en effet, la qualité et la quantité des eaux sont conditionnées, toutes les deux, par la circulation hydrodynamique souterraine.

* L'exhaure minière peut être assimilée à la vidange d'un récipient (*cf.* Fig. 44, page 113) alimentée par la pluie efficace. Les réserves souterraines (hauteur d'eau h_t dans le récipient) jouent un rôle régulateur entre l'exhaure et la pluie efficace. L'extraction cumulée de minerai, joue un rôle d'élargir les dimensions du récipient de sorte que l'exhaure augmente dans le temps.

* Le modèle général adapté pour toutes les mines peut être mis sous la forme :

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \varepsilon_t$$

$$(0 \leq \beta \leq 1)$$

et un modèle auxiliaire, utilisé pour les mines non stationnaires en cas de manque de données d'extraction, peut être exprimé par :

$$\frac{Q_t}{MA12} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + \varepsilon_t$$

Dans les deux modèles ci-dessus, β est appelé "*coefficient de tempérament*"; $\phi_0, \phi_1, \phi_2 \dots \phi_m$ sont les coefficients de pondération constituant la "*fonction de transfert*"; m est la "*mémoire*" du système étudié et ε_t le "*terme résiduel*". De plus, nous pouvons raffiner le modèle général et son auxiliaire en "*blanchissant*" le terme résiduel à l'aide de la méthode de Box-Jenkins, nous aurons :

$$\frac{Q_t}{S_t^\beta} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + f(\epsilon_t) + a_t$$

et

$$\frac{Q_t}{MA12} = C + \phi_0 E_t + \phi_1 E_{t-1} + \phi_2 E_{t-2} + \dots + \phi_m E_{t-m} + f(\epsilon_t) + a_t$$

... où $f(\epsilon_t)$ est une fonction calculée par la méthode de Box-Jenkins et a_t un bruit blanc.

* Le coefficient de tempérament β représente la réaction d'une mine vis-à-vis de l'exploitation minière et pourrait être assimilé, en quelque sorte, au coefficient d'élasticité de la mécanique des matériaux. En tant que coefficient représentant un sens physique, sa valeur ne peut pas être quelconque. En effet, ce coefficient est défini dans l'intervalle $[0, 1]$, pour les mines stationnaires $\beta=0$, pour les mines non stationnaires $0 < \beta \leq 1$.

III - PERSPECTIVES

Avec les modèles établis, il est possible de prévoir l'exhaure minière à une date donnée. Par exemple, dans 5 ans, si l'on exploite chaque année 3 millions de tonnes de minerai, on pourra calculer le tonnage d'eau qu'il faut pomper par mois ou par an. Cette prévision est très importante, surtout pour les mines non stationnaires, car elle nous aide à gérer plus rationnellement les pompages et à calculer la rentabilité d'une mine en tenant compte des dépenses d'exhaure.

Dans le bassin ferrifère lorrain, la plupart des mines ont été fermées à cause de la crise de la sidérurgie. Par conséquent, une prévision dans un temps futur ne présente plus un grand intérêt pratique. Mais les mines fermées et ennoyées constituent des réservoirs souterrains qui pourront être exploitées comme des ressources en eau potable. Le problème d'aménagement du bassin est actuellement primordial et l'extrême sécheresse de l'été 1989 a accentué la nécessité de l'exploitation de la "banque" des données d'exhaure minière.

On sait que les hydrogéologues chargés de la gestion du bassin s'appuient souvent sur des lois de distribution (Gumbel, Galton, Montana, *etc.*) et sur des paramètres statistiques (Moyenne, Ecart-type, Gradex saisonniers, *etc.*). Dans le cas des précipitations ajustées à des lois de distribution de Gumbel ou de Montana, les paramètres de ces distributions évoluent avec

la durée de la mesure de la précipitation, et souvent selon des lois "simples", en tout cas, que l'on peut exprimer de façon analytique. On peut espérer que l'exhaure aura un comportement analogue, puisqu'elle dépend des précipitations [cf. Consistance de l'Etude, E.N.S.G. et A.F.B.R.M.].

Pour les mines stationnaires, l'ajustement des lois statistiques peut se faire immédiatement avec les données connues. Par contre, pour les mines non stationnaires, il est indispensable de "filtrer" d'abord l'effet de l'extraction cumulée sur l'exhaure. Pour ce faire, on peut reconstituer ce qu'aurait été l'exhaure, si les exploitations avaient été dans l'état actuel : en effet, à partir de leur fermeture, c'est bien leur état actuel qui conditionnera leur fonctionnement hydraulique. La reconstitution des données fait appel aux modèles que nous avons pu élaborer. La figure 67 illustre un exemple de reconstitution des données.

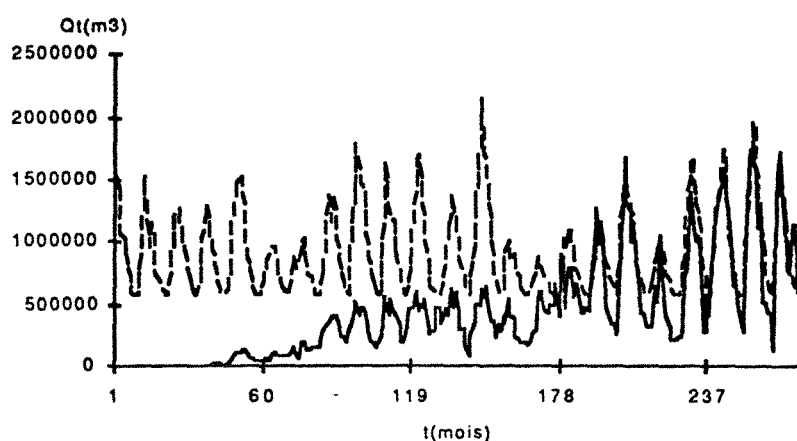


Fig. 67 - Exhaure mesurée et reconstituée pour la mine Saizerais (1958-1980).

D'autre part, dans la gestion du bassin, la sécheresse pose souvent de graves problèmes (quantité et qualité des eaux pour utilisations industrielle et domestique). Il nous semble très important d'étudier les comportements spécifiques des nappes souterraines ou des mines vis-à-vis de la sécheresse annuelle ou pluriannuelle. Cette étude pourra fournir une prévision plus précise qui servira aux exploitants d'eau souterraine qui peuvent, soit être assurés d'une production normale, soit, en cas de risque de "pénurie", mettre en oeuvre des solutions d'appoint ou améliorer des dispositifs d'exploitation.

Possédant de suffisamment de données, dans le bassin ferrifère lorrain, cette étude est possible. Les années 1949 et 1976 étant les années sèches, nous avons des données d'exhaure, pour la moitié des mines en 1949 et pour toutes les mines en 1976. C'est une richesse à exploiter !

Nous pensons que cette étude pourra être menée selon deux approches : approche déterministe et approche statistique. Pour l'approche déterministe, nous pouvons reprendre la méthode de CHANZY [cf. Annexe I] et effectuer une éventuelle modification, car pendant la période sèche, c'est bien le phénomène de tarissement qui est un facteur dominant. En ce qui concerne à l'approche statistique, nous pouvons adopter un modèle d'intervention (cette méthode a été utilisée par exemple dans le but d'étudier l'influence de l'embargo pétrolier sur l'économie des Etats-Unis). Il s'agira de formuler des critères sur la sécheresse, d'introduire des variables indicatrices, d'étudier les comportements de l'exhaure dans la période de transition (cf. Fig. 68).

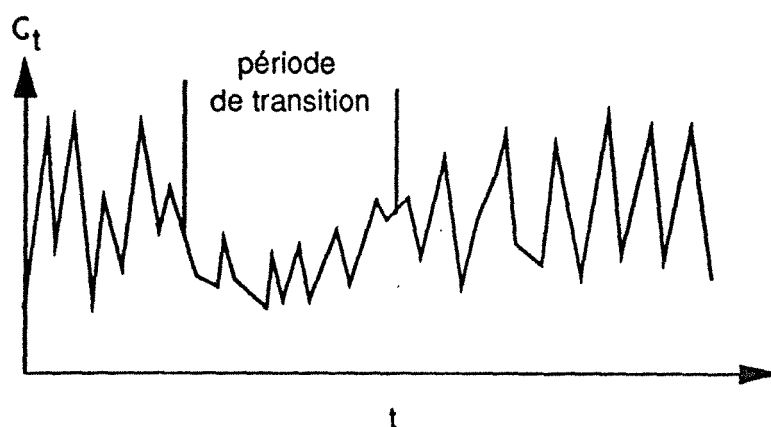


Fig. 68 - Le comportement de l'exhaure après une période sèche.

BIBLIOGRAPHIE

- [1] A.F.B.R.M. (1980). "*Ressources en Eau d'Exhaure du Bassin Ferrifère Lorrain - Exploitation du Fichier Eaux d'Exhaure des Mines de Fer*".
- [2] A.F.B.R.M. (1982). "*Schéma de Restructuration des Exhaures des Mines de Fer - Actualisation au 1er Novembre*".
- [3] AZENCOTT R., DACUNHA-CASTELL D. (1984). "*Séries d'Observation Irrégulières - Modélisation et Prévision*". Masson, Paris, 198p.
- [4] BLANC-LAPIERRE A., PICINBONO B. (1981). "*Fonctions Aléatoires*". Masson, Paris, 419p.
- [5] BELSLEY D.A., KUH E., WELSCH R.E. (1980). "*Regression Diagnostics - Identifying Influential Data and Sources of Collinearity*". John Wiley & Sons, New-York, 292p.
- [6] BENZECRI J.P. & Collaborateurs (1976). "*L'Analyse des Données*". Tome I (La Taxonomie), 2ème Edition, Dunod, Paris, 631p.
- [7] BOULY S. (1984). "*Rapport Relatif à la Détermination des Paramètres de Protection de la Mine de Saizerais*" , E.N.S.G., Nancy.
- [8] BOULY S. (1988). "*Protection des Eaux Souterraines du Bassin Ferrifère Lorrain - Propositions pour une Protection Acceptable des Ressources en Eaux*". A.F.B.R.M. et E.N.S.G., Nancy, 65p.
- [9] BOX G.P., JENKINS G.M. (1976). "*Time Series Analysis - Forecasting and Control*". Hodey-day, San Francisco, Calif., 575p.
- [10] BRACEWELL R. (1989). "*L'Analyse de Fourier*". Edition Française de Scientific Américan, Mensuel n° 142, Paris, 74-85.
- [11] B.R.G.M. (1967). "*Chronique d'Hydrogéologie*". N° 10, Mars 1967, Orléans.
- [12] CANCEILL M., FEUGA B. (1972). "*Bassin Ferrifère Lorrain - Analyse Statistique des Données d'Exhaure Minière (Bassin de l'Orne)*". B.R.G.M., 72 SGN 421 NES, Orléans.

- [13] CANCEILL M. (1974). "*Bassin Ferrifère Lorrain - Relation entre Pluie et Exhaure des Mines de Fer (Bassin de l'Orne)*". B.R.G.M. et A.F.B.R.M.74 SGN 252 AME, Orléans.
- [14] CANCEILL M., FORKASIEWICZ J., THIERY D. (1977). "*La Prévision du Régime Naturel des Nappes d'Eau Souterraine Appliquée à leur Gestion*". B.R.G.M. 77 SGN 662 HYD, Orléans.
- [15] CARA M. (1989). "*Géophysique*". Dunod, Paris, 185p.
- [16] CASTANY G. (1963). "*Introduction à l'Etude des Courbes de Tarissement*". Annales de Spéléologie, 18ème fasc. 4, 24-30.
- [17] CASTANY G., GUILLAUME M., JEANDEL G., DE MAUTORT J. (1960). "*Inventaire des Ressources Hydrauliques du Bassin Lorrain - Le Bassin de Longwy*". B.R.G.M. A.1641, Orléans.
- [18] CHAMBON C. "*Analyse Statistique et Prévision*". 2ème Partie, Cours à l'Ecole des Mines de Nancy, 107p.
- [19] CHAMBON C. "*La Prévision dans le Temps*". Cours à l'Ecole des Mines de Nancy.
- [20] CHANZY M. (1926). "*Influence des Eaux Superficielles sur les Eaux d'Exhaure dans le Bassin de Briey*". Nancy, 21p.
- [21] CHANZY M. (1926). "*Contribution à l'Etude de l'Exhaure dans le Bassin de Briey (Action des Pluies)*". Ann. des Mines, Dunod, Paris, 29p.
- [22] CHANZY M., BICHELONNE J. (1930). "*Etat Actuel de l'Exhaure des Mines du Bassin de Briey*". Ann. des Mines, Dunod, Paris
- [23] COHN B.P., ROBINSON J.E., MERRIAM D.F. (1980). "*Forecasting Lake Levels*". Sciences de la Terre, n°14, Juillet 1980, Nancy, 121-130.
- [24] COUTROT B., DROESBEKE F. (1984). "*Les Méthodes de Prévision*", Que Sais-je ? Presse Universitaire de France, Paris, 127p.
- [25] COURTOT S. "*Analyse des Chroniques*". Cours à l'Ecole des Mines de Nancy.

- [26] DAGALIER R., DEMASSIEUX L. (1986). "*Origine des Teneurs en Sodium des Eaux des Mines de Fer en Cours d'Exploitation ou après Ennoyage*". E.N.S.G. et A.F.B.R.M., Nancy, 10p.
- [27] DARRICAU-BEUCHER H. (1981). "*Approche Géostatistique du Passage des Données de Terrain aux Paramètres des Modèles en Hydrogéologie*". Thèse de Docteur-Ingénieur, Ecole des Mines de Paris, Paris, 118p.
- [28] DAVID M. (1977). "*Geostatistical Ore Reserve Estimation*". Developments in Geomatics 2, Elsevier Scientific Publishing Company, Amsterdam, 364p.
- [29] DAVIS J.C., SAMPSON R.T. (1973). "*Statistics and Data Analysis in Geology*". John Wiley & Sons, New-York, 550p.
- [30] DEGALIER R. (1972). "*Un Modèle de Simulation des Ecoulements Souterrains - Le Modèle SIMERO*". Bull. B.R.G.M. (2), 3-1972, Orléans, 42p.
- [31] DELFINER P., MATHERON G. (1980). "*Les Fonctions Aléatoires Intrinsèques d'Ordre k*". Fontainebleau, France, 35p.
- [32] DELOHME J.P. (1976). "*Applications de la Théorie des Variables Régionalisées dans les Sciences de l'Eau*". Thèse de Docteur-Ingénieur, Université de Paris VI, Paris, 130p.
- [33] DEMASSIEUX L. "*Eléments d'Hydrodynamique Souterraine*", Cours à l'E.N.S.G., Nancy, 202p.
- [34] DRAPER N.R., SMITH H. (1981). "*Applied Regression Analysis*". Second Edition, John Wiley & Sons, New-York, 709p.
- [35] FERREIRA H.G. (1980). "*Traitement des Données de Prospection Géochimique Basé sur l'Analyse Structurale*". Sciences de la Terre, n°14 Juillet, 143-159.
- [36] GASCUEL-ODOUX C. (1984). "*Application de la Géostatistique à l'Etude de la Variabilité Spatiale des Propriétés Hydriques du Sol*". Thèse Docteur-Ingénieur, Ecole des Mines de Paris et I.N.R.A. Rennes, 235p.
- [37] GUILLAUME M., DE MAUTORT J., JEANDEL G. (1960). "*Etude des réserves en Eaux Souterraines des Horizons Aquifères de l'Aalénien-Bajocien (Exhaure des Mines de Fer du Bassin d'Ottange-Amermont)*". D.S.G.R., 64 A2 B.R.G.M., Orléans.

- [38] GUILLAUME M., DE MAUTORT J., JEANDEL G. (1960). "*Etude des Réserves Souterraines des Horizons Aquifères de l'Aalénien-Bajocien (Exhaure des Mines de Fer du Bassin de l'Orne)*". D.S.G.R., 63 A39 B.R.G.M., Orléans.
- [39] HERVE D. (1980). "*Etude de l'Acquisition d'une Teneur en Sulfates par les Eaux Stockées dans les Mines de Fer de Lorraine*". Thèse de Docteur-Ingénieur, I.N.P.L., Nancy, 80p.
- [40] JENKINS G.M, WATTS D.G. (1968). "*Spectral Analysis and Its Applications*". Hoday, San Francisco, Calif., 525p.
- [41] JOURNEL A. (1975). "*Analyse Structurale - Guide Pratique de Géostatistique Minière*". Introduction de Géostatistique Minière, Cours à l'Ecole des Mines de Nancy, 72-122.
- [42] KANG Y.S. (1986). "*Analyse des Données de 1983 à 1985 sur la Qualité des Eaux dans la Mine Ennoyée de Saizerais*". Rapport D.E.A., I.N.P.L., Nancy, 73p.
- [43] KRIGE D.G. (1981). "*Lognormal de Wjjsian Geostatistics for Ore Evaluation*". South African Institute of Mining and Metallurgy, Johannesburg, Second edition, 50p.
- [44] KRUMBEIN W.C., GRAYBILL F.A. (1965). "*An Introduction to Statistical Models in Geology*". Mc Graw-Hill, New-York, 475p.
- [45] LABORDE J.P. (1982). "*Eléments d'Hydrologie de Surface. Critique et Analyse Statistique des Données Hydrologiques*". I.N.P.L., Nancy, 330p.
- [46] LABORDE J.P. (1982). "*Etude de la Répartition Spatiale des Paramètres Statistiques de la Distribution des Pluies Journalières sur le Bassin Rhin-Meuse*", I.N.P.L., Nancy, 37p.
- [47] LABORDE J.P. (1984). "*Analyse des données et Cartographie automatique en Hydrologie - Elements d'Hydrologie Lorraine*", Thèse Docteur ès Sciences, I.N.P.L., Nancy, 484p.
- [48] LEBART L., MORINEAU A., TABARD N. (1977). "*Techniques de la Description Statistique - Méthodes et Logiciels pour l'Analyse des Grands Tableaux*". Dunod, Paris, 351p.

- [49] LEBART L., MORINEAU A., FENELON J.P. (1979). "*Traitement des Données Statistiques - Méthodes et Programmes*". Dunod, Paris, 513p.
- [50] LEBART L., MORINEAU A., WARWICK K.M. (1984). "*Multivariate Descriptive Statistical Analysis - Correspondance Analysis and Related Technics for Large Matrices*". John Wiley & Sons, New-York, 231p.
- [51] LEDOUX E. (1980). "*Modélisation Intégrée des Ecoulements de Surface et des Ecoulements Souterrains sur un Bassin Hydrologique*". Thèse Doctorale, E.N.S.M.P. et Université de Paris VI, Paris, 230p.
- [52] LEMPREUR R. (1988). "*Analyse Spatio-temporelle des Pluies sur la Site de Nancy*", Thèse Doctorale, I.N.P.L, Nancy, 169p.
- [53] LE NIR M. (1973). "*Exhaure des Mines de Fer du Bassin Lorrain - Synthèse des Etudes Réalisées à la Date 30 Septembre 1973*". B.R.G.M. 73 GSN 320 NES, Orléans, 23p.
- [54] MA Y.Z. (1987). "*Filtrage Géostatistique des Images Numériques*". Thèse Doctorale, I.N.P.L., Nancy, 143p
- [55] MALLET J.L. (1988). "*Analyse des Données*". Tomes 1-2, I.N.P.L, Nancy, 780p.
- [56] MANGIN A. (1981). "*Utilisation des Analyses Corrélatoire et Spectrale dans l'Approche des Systèmes Hydrologiques*". C.R. Acad. Sc. Paris, T. 293, Série 2, 401-404.
- [57] MANGIN A. (1981). "*Apports des Analyses Corrélatoires et Spectrales Croisées dans la Connaissance des Systèmes Hydrologiques*". C.R. Acad. Sc. Paris, t.293, Série 2, 1011-1014.
- [58] MANGIN A. (1984). "*Incidence du Choix du Pas de Discrétisation d'un Processus sur la Simulation à Partir de Modèles*". Journal of Geology, 67 (1984) 25-43, Elsevier Science Publishers, B.V., Amsterdam.
- [59] MARSILY G. de(1981). "*Hydrogéologie Quantitative*", Masson, Paris, 215p.
- [60] MARSILY G. de (1978). "*De l'Identification des Systèmes Hydrogéologiques*". Tomes I-II, Thèse de Doctorat d'Etat, Université de Paris VI.

- [61] MATHERON G. (1962). "*Traité de Géostatistique Appliquée*". Tomes I-II, Edition Technip, Paris.
- [62] MATHERON G. (1970). "*La Théorie des Variables Régionalisées et Ses Applications*". Fascicule 5, Fontainebleau, France, 212p.
- [63] MATHERON G. (1978). "*Le Krigeage Disjonctif et le Paramétrage Local des Réserves*". Fontainebleau, France, 20p.
- [64] POINTET T. (1975). "*Essai de Caractérisation du Coefficient de Tarissement des Nappes*". Thèse Doctorale, I.N.P.L. et Université de Nancy I, 115p.
- [65] RADIX J.C. (1984). "*Filtrage et Lissage Statistiques Optimaux Linéaires*". Cépadus Edition, Toulouse, France, 350p.
- [66] RAMON S. (1975). "*Note sur le Débit de la Mine de Saizerais (Mthe-et Melle)*", 75/DPR/2., Metz.
- [67] RAO C.R. (1973). "*Linear Statistical Inference and Its Applications*". John Wiley & Sons, New-York, 625p.
- [68] RENBU J.M. (1981). "*An Introduction to Geostatistical Methods of Mineral Evaluation*". Johannesburg.
- [69] RIVOIRARD J. (1984). "*Une Méthode d'Estimation du Récupérable Local Multivariable*". Fontainebleau, France.
- [70] RODDIER F. (1985). "*Distributions et Transformation de Fourier*". Mc Graw-Hill, New-York.
- [71] ROUSSEUW P.J., LEROY A.M. (1987). "*Robust Regression & Outlier Detection*". John Wiley & Sons, New-York, 329p.
- [72] ROYER J.J. (1982). "*Cours de Probabilités*". I.N.P.L. Tome II : Processus Stochastiques, I.N.P.L., Nancy, 185p.
- [73] ROYER J.J. (1988): "*Analyse Multivariable et Filtrage des Données Régionalisées*". Thèse Docteur ès Sciences, I.N.P.L., Nancy, 240p.

- [74] SALMAN W.P., SOLOTAREFF M.S. (1982). *"Le Filtrage Numérique"*. Editions Eyrolles, Paris, 174p.
- [75] SCHWARTZ M., SHAW L. (1975). *"Signal Processing - Discrete Spectral Analysis, Detection and Estimation"*. Mc Graw-Hill, New-York.
- [76] SEBER G.A.F., WILD C.J. (1989). *"Nonlinear Regression "*. John Wiley & Sons, New-York.
- [77] STEARNS S.D. (1975). *"Digital Signal Analysis"*. (Traduction en Chinois par GAO S.Q. & Collaborateurs).
- [78] TAUPIN D. (1988). *"Probabilities Data Reduction and Error Analysis in the Physical Sciences"*. Les Editions de Physique Paris, 164-203.
- [79] THIERY D. (1978). *"Application à Quatre Bassins Hydrologiques des Méthodes de Prédiction des Etiages par Déconvolution"*. B.R.G.M. 78 SGN 621 HYD, Orléans, 203p.
- [80] WEAVER H.J. (1983). *"Applications of Discrete and Continuous Fourier Analysis"*. John Wiley & Sons, New-York.
- [81] WHEELWRIGHT S.C., MAKRIDAKIS S. (1978). *"Forecasting"*. Hoden-day, San Francisco, Calif.
- [82] WHEELWRIGHT S.C., MAKRIDAKIS S. (1983). *"Méthodes de Prédiction pour la Gestion"*. Traduit de l'Américain par De MARICOURT R.

LISTE DES ANNEXES

- ANNEXE I METHODE DE SIMULATION PROPOSEE PAR
 CHANZY**
- ANNEXE II ESSAI D'INTERPRETATION DES COURBES
 DE TARISSEMENT DE L'EXHAURE DE
 RONCOURT**
- ANNEXE III BILAN HYDROLOGIQUE**

ANNEXE I

METHODE DE SIMULATION PROPOSEE PAR CHANZY

(d'après CANCEILL [13])

La méthode de CHANZY, dont le point clé est la construction de la courbe de tarissement, peut être décomposée en cinq étapes :

* On considère pas les pluies brutes mais les pluies corrigées des quantités absorbées par l'évapotranspiration par la végétation. Ces pluies corrigées sont appelées "*eaux résiduelles*" ou "*résidus*" (Ce sont en fait les pluies efficaces). Le ruissellement est supposé proportionnel à l'infiltration, ce qui permet de ne pas en tenir compte.

En application de ce principe, CHANZY calcule les résidus par période de cinq jours, en considérant que "les résidus de cinq jours consécutifs se totalisent pour donner lieu à une évaporation égale au sixième de l'évaporation du mois". Les valeurs d'évaporation sont donc considérées comme constantes à l'intérieur de chaque mois. Les chiffres proviennent des mesures d'évaporation relevées jusqu'en 1910 par le service de la navigation de Meurthe-et-Moselle à Parroy, sur bac d'évaporation, et sont tout à fait empiriquement corrigés pour tenir compte de la végétation ; CHANZY ne donne que peu de détails sur ces corrections, mais propose *ex-abrupto* une série de valeurs dans son article de 1926, valeurs légèrement corrigées en 1930.

* La vidange est théoriquement assimilée à celle d'un récipient où l'eau est à un niveau h , et dont le débit est Q (fig. A1). Q et h sont alors deux fonctions du temps t satisfaisant aux équations :

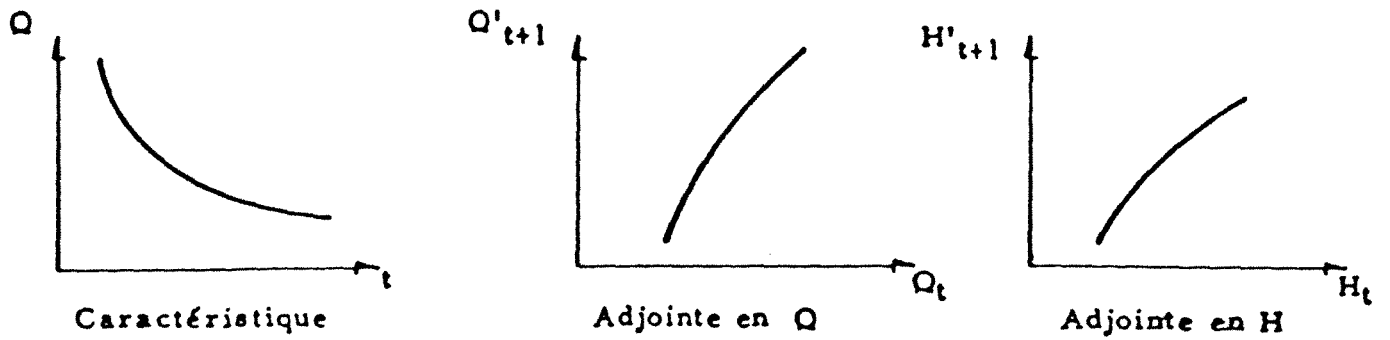
$$Q = k \sqrt{h}$$

$$Q dt = -S dh$$

... où S est la section (fonction de la cote h) horizontale de réservoir et k est une constante.

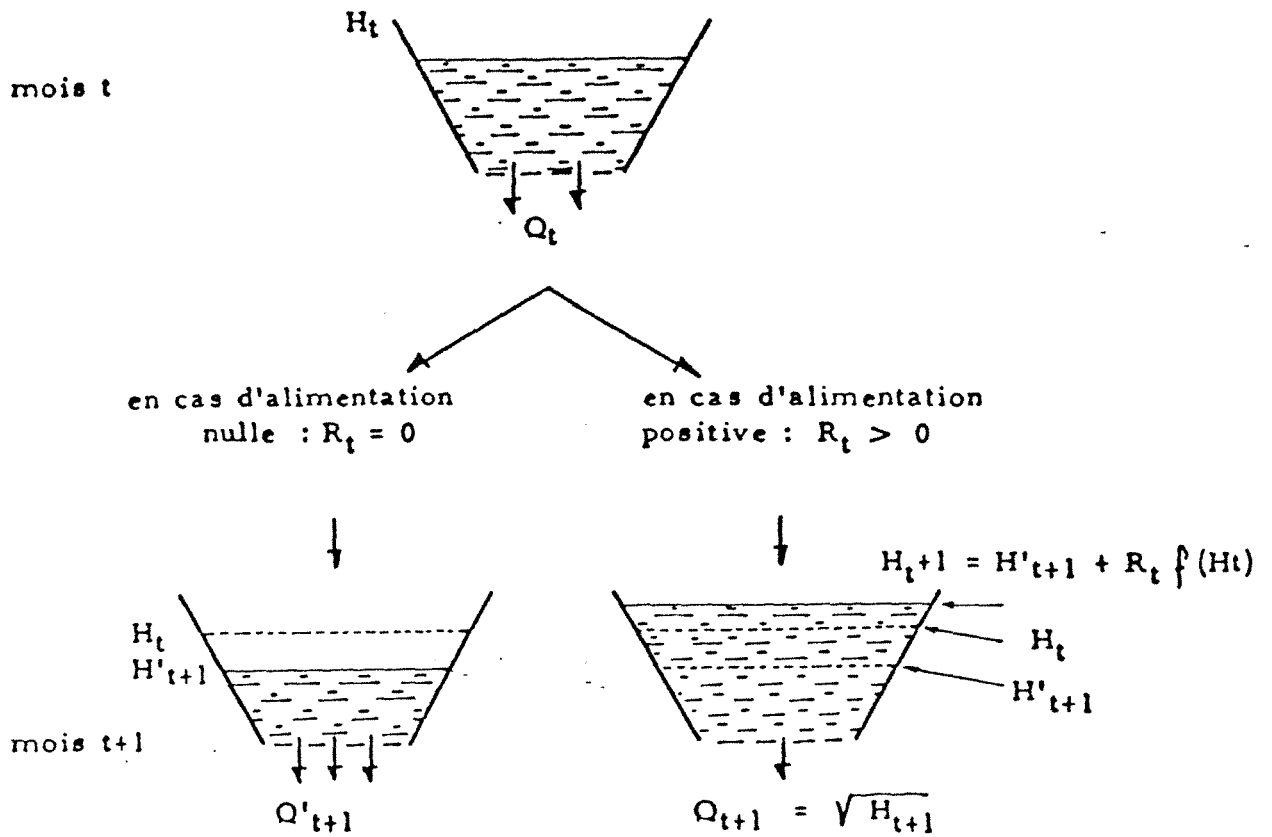
La courbe de vidange ou de tarissement donnant Q en fonction de t est appelée "*caractéristique de la mine*".

Fig. A1 - Représentation schématique de la méthode de CHANZY (Extrait de [13]).



$$Q = K \sqrt{h}$$

$$H = Q^2 = K^2 h$$



* Si l'on admet qu'un résidu produit son action sur l'exhaure en moins de cinq jours, il est possible de construire empiriquement la caractéristique en "juxtaposant les portions du graphique réel correspondant aux périodes sèches écourtées de cinq jours à leur début".

Dans toute la suite, on ne travaille plus au pas de temps de cinq jours, mais au pas de temps mensuel.

* De la caractéristique CHANZY déduit l'adjointe, courbe qui donne Q'_{t+1} , valeur que prendrait le débit d'exhaure en période sèche à l'instant $t+1$ en fonction de sa valeur Q_t à l'instant t . CHANZY définit une nouvelle quantité H par $H=Q^2$ et trace l'adjointe donnant H'_{t+1} en fonction de H_t .

* Entre le temps t et $t+1$ (qui sont, pour CHANZY, des mois), il se produit un accroissement $\Delta H=H_{t+1}-H_t$, accroissement qui dépend de l'alimentation R_t (résidu) et de la valeur H_t :

$$\Delta H=R_t f(H_t)$$

Comme $H_{t+1}=H'_{t+1}+\Delta H$, on obtient H_{t+1} en lisant H'_{t+1} en fonction de H_t sur l'adjointe et en calculant ΔH à partir de la formule $\Delta H=R_t f(H_t)$. Dans son papier de 1926, CHANZY propose une fonction linéaire pour $f(H_t)$:

$$f(H_t)=AH_t+B$$

Il suggère, en 1930, de modifier cette expression en $f(H_t)=C\lambda$.

... où λ est défini théoriquement par $\lambda=\frac{1}{Q} \frac{dH}{dt}$ et pratiquement calculé par $\lambda=\frac{H_t-H'_{t-1}}{\sqrt{H_t}}$.

Les constantes A , B , C sont déterminées par tâtonnements dans chaque cas particulier.

Dans le papier de 1952, enfin, la constante C est interprétée comme la superficie du bassin versant.

Tout ce processus, donc conduit à calculer, d'une part à l'aide de l'adjointe, d'autre part à l'aide des résidus et de la fonction $f(H_t)$, les valeurs H_{t+1} de mois en mois successivement, et en déduire les valeurs de l'exhaure Q_{t+1} par $Q_{t+1}=\sqrt{H_{t+1}}$.

Le fond de la méthode, en fait, revient à ajuster deux relations f et g : l'une, $H'_{t+1}=g(H_t)$ et l'autre, $f(H_t)$, définit la variation de niveau produite par un résidu unité, c'est-à-dire l'alimentation du réservoir. La relation $H'_{t+1}=g(H_t)$ correspond au phénomène de vidange.

L'exhaure est alors donnée par la formule (où les H_t sont remplacés par les Q_t^2) :

$$Q_{t+1} = \sqrt{g(Q_t^2) + R_t f(Q_t^2)}$$

... où Q_t est le débit, R_t est le résidu à l'instant t .

ANNEXE II

ESSAI D'INTERPRETATION DES COURBES DE TARISSEMENT DE L'EXHAURE DE RONCOURT

(d'après LE NIR [53])

Cette interprétation s'appuie sur trois hypothèses :

- Deux types d'écoulement coexistent dans les calcaires : un écoulement franchement karstique de type turbulent ; un écoulement laminaire, dans les petites et moyennes fissures, qui jouent le rôle d'un "*volant régulateur*". Les zones d'échange entre ces deux types d'écoulement sont localisées.

- La répartition des petites et moyennes fissures est sensiblement homogène.

- Les courbes de décrue peuvent se décomposer, en représentation semi-logarithmique, en deux droites représentant la vidange de fissures et l'écoulement des chenaux et chaque droite suit la loi de MAILLET :

$$Q=Q_0e^{-\alpha t}$$

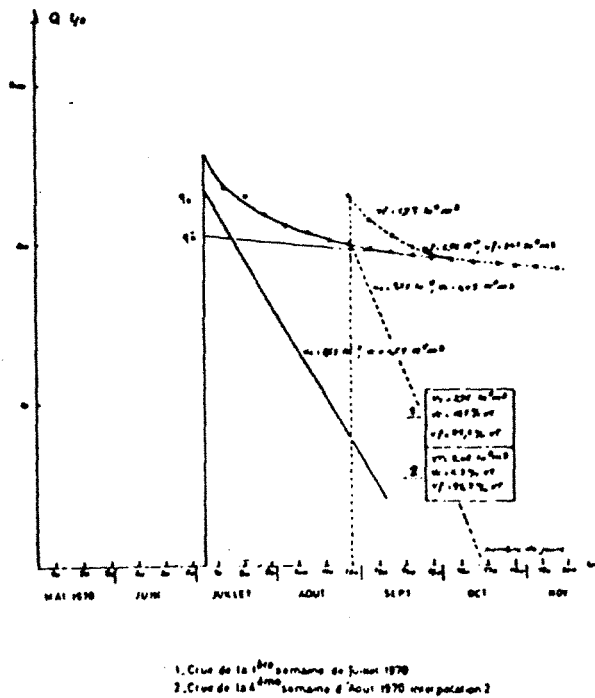
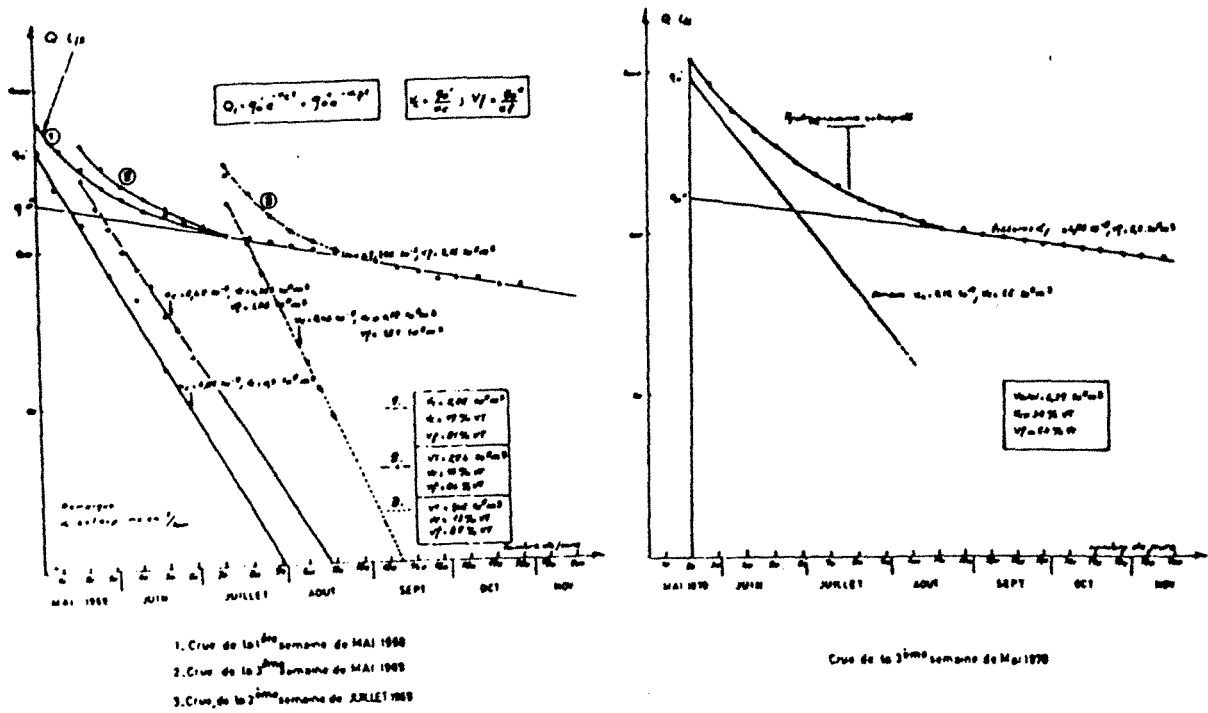
Les débits retenus sont les moyennes hebdomadaires entre mai 1969 et novembre 1970. Les tarissements étudiés sont ceux des crues débutant en : première semaine de mai 1969 ; troisième semaine de mai 1969 ; deuxième semaine de juillet 1969 ; troisième semaine de mai 1970 ; première semaine de juillet 1970 ; quatrième semaine de juillet 1970.

Les courbes de tarissement sont obtenues par extrapolation entre le début de la courbe de crue et la zone des plus faibles débits observables s'y rattachant.

La décomposition de ces courbes se présente dans la figure A2. La formule suivante est obtenue :

$$Q_t=q'_0e^{-\alpha ct}+q''_0e^{-\alpha ft}$$

Fig. A2 - Courbes de tarissement (Extrait de [53]).



... où α_c est le coefficient de tarissement correspondant à la vidange des chenaux (régime turbulent) ; α_f est le coefficient de tarissement correspondant à la vidange des fissures (régime laminaire).

Le volume initial disponible est :

$$V_c = \frac{q'0}{\alpha_c} \quad \text{pour les chenaux}$$

$$V_f = \frac{q''0}{\alpha_f} \quad \text{pour les fissures}$$

$$V_t = \frac{q'0}{\alpha_c} + \frac{q''0}{\alpha_f} \quad \text{est le volume total}$$

Les résultats sont rassemblés dans le tableau A1 et les conclusions dégagées sont les suivantes :

- La proportion du volume d'eau s'écoulant en régime laminaire est très supérieur au volume d'eau s'écoulant en régime turbulent.

- Plus la crue est petite, plus petite est la proportion d'eau s'écoulant en régime turbulent.

- La loi de décroissance du pourcentage du volume du régime laminaire " V_l ", par rapport au volume total " V_t " en fonction de l'accroissement des débits maximums de crue, est logarithmique et de la forme :

$$Q \text{ maximum de la crue} = A \log\left(\frac{V_l}{V_t}\%\right) + B$$

Pour l'exhaure de Roncourt(fig. A3) :

$$Q_{\text{maximum de la crue}} = -0,028 \log\left(\frac{V_l}{V_t}\%\right) + 2$$

- Pour les fortes crues (période pluvieuse de février à mai 1970), le volume d'eau circulant en régime laminaire représente encore 65 % et 70 % du volume d'eau contenu dans les calcaires au passage du maximum de la crue. Cette proportion augmente très vite quand les crues diminuent.

Tableau A1 - Tableau des résultats (Extrait de [53]).

DATE DU MAXI- MUM DE LA CRUE	DEBIT MAXIMUM DE LA CRUE l/s	af 10^{-3}	ac 10^{-2}	$V_T \cdot 10^6 \text{ m}^3$	$\frac{Vf}{V_T} \%$	$\frac{Vc}{V_T} \%$
5 - 3 - 70	1200	4,07	4,45	5,27	65	35
26 - 2 - 70	1200	7,94	5,90	4,40	70	30
17 - 5 - 70	1200	4,85	5,12	4,39	64	36
3 - 5 - 69	625	7,86	6,89	2,46	81	19
18 - 5 - 69	450	7,86	6,42	2,26	84	16
3 - 7 - 70	347	3,92	6,55	2,76	89,5	10,5
12 - 7 - 69	310	7,86	8,45	1,46	87	13
23 - 8 - 70	196	3,92	9,32	2,08	95,7	4,3
20 - 12 - 69	196	4,07	12,9	2,36	97,7	2,3

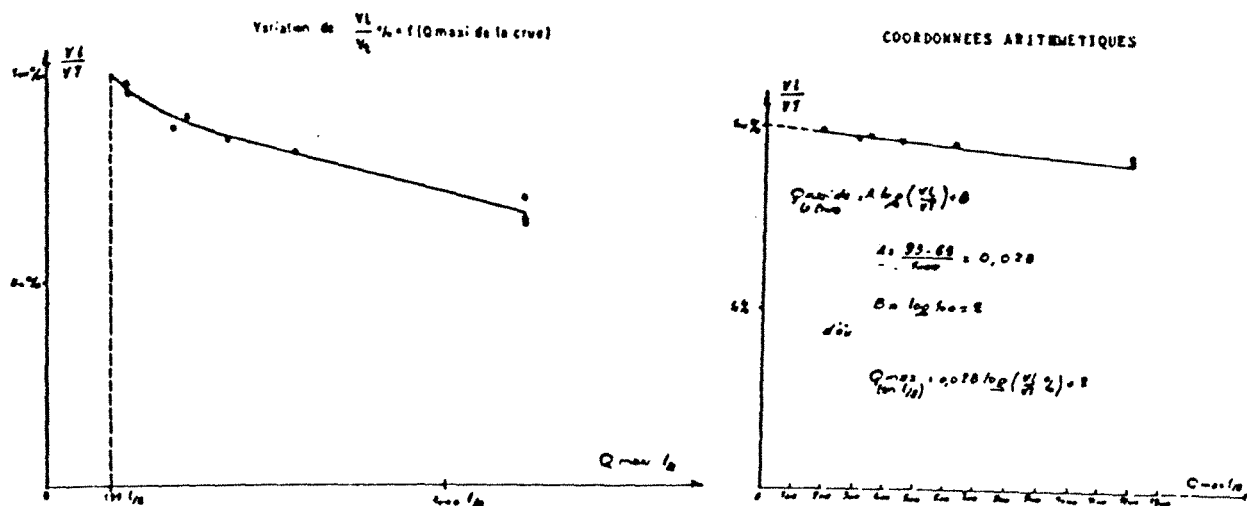
où :

af = coefficient de tarissement correspondant à la vidange des fissures (régime laminaire).

ac = coefficient de tarissement correspondant à la vidange des chenaux (régime turbulent).

V_T = volume total initial

Fig. A3 - Interprétation des courbes de tarissement (Extrait de [53])



ANNEXE III

BILAN HYDROLOGIQUE

Dans un bassin superficiel, le bilan hydrologique s'écrit :

$$P = EC + ETR + \Delta R + E$$

- où
- P : pluie brute en mm
 - EC : écoulement superficiel (en mm)
 - ETR : Evapotranspiration réelle (en mm)
 - ΔR : variation de RFU (réserve facilement utilisable, en mm)
 - E (= P - EC - ETR - ΔR) : pluie efficace au sens strict, en mm
 - E + EC (= P - ETR) : pluie efficace au sens large, en mm

La pluie efficace au sens stricte est appelée aussi excédent : c'est la portion qui infiltre et alimente les aquifères.

Le calcul du bilan consiste d'abord à estimer ETR à partir de ETP (évapotranspiration potentielle) et ETP à son tour s'estime par plusieurs méthodes parmi lesquelles deux sont citées ci-après :

I - Calcul de l'évapotranspiration potentielle mensuelle

1. Formule de Turc

$$\text{Si } H_r \geq 50 \% : \text{ETP} = (I_g + 50) \frac{0.4 t}{t + 15}$$

$$\text{Si } H_r < 50 \% : \text{ETP} = (I_g + 50) \frac{0.4 t}{t + 15} \left[1 + \frac{50 - H_r}{70} \right]$$

Pour les deux cas ci-dessus, on a :

$$I_g = I_g A (0.18 + 0.62 \frac{h}{H}) : \text{la radiation globale moyenne d'origine solaire (cal/cm}^2\text{/jour) .}$$

où $H_r = \frac{h}{H}$: humidité relative de l'air (en %) pendant le mois

t : température moyenne du mois (en °C)

h : durée d'insolation durant le mois considéré (en heures par mois)

H : durée astronomique du mois considéré (en heures par mois)

I_{gA} : énergie de radiation qui atteindrait le sol en l'absence d'atmosphère (cal/cm²/jour)

H et I_{gA} sont fonction de la latitude et sont données par un tableau.

2 . Formule de Thornthwaite

Le calcul de l'évapotranspiration potentielle mensuelle n'est possible que si la température moyenne mensuelle est inférieure ou égale à 38° C. Deux cas sont alors à considérer :

$$* \text{ Si } t < 26.5^\circ \text{ C : ETP} = 16 (10 t / D)^\alpha \text{ K}$$

ETP : évapotranspiration potentielle mensuelle corrigée (en mm)

$$I = \sum_{I}^{12} i : \text{ indice thermique annuel}$$

$$i = (t / 5)^{1.5} : \text{ indice thermique mensuel}$$

t : température moyenne mensuelle (en ° C)

$$\alpha = (1.6 / 100) I + 0.6$$

k : coefficient de correction mensuel dépendant de la latitude

Les valeurs K sont données dans le tableau ci-après :

Mois	J	F	M	A	M	J	J	A	S	O	N	D
K	0,73	0,78	1,02	1,15	1,32	1,33	1,33	1,24	1,05	0,91	0,75	0,70

Tableau A2 - Les valeurs k.

* Si $26.5^\circ \text{ C} \leq t \leq 38^\circ$: ETP mensuelle non corrigée est donnée directement par une table. L'évapotranspiration potentielle mensuelle corrigée est alors calculée par application du coefficient de correction mensuelle.

II - Calcul de l'évapotranspiration réelle mensuelle

La comparaison des valeurs de l'ETP mensuelle et de la pluie brute mensuelle (P) permet de calculer l'ETR mensuelle.

Plusieurs cas peuvent se présenter :

* Si $ETP < P$, $ETR = ETP$

La quantité d'eau correspondant à $P - ETP$ est emmagasinée dans le sol jusqu'à saturation de celui-ci. Le reste est disponible pour l'écoulement superficiel et l'infiltration vers les nappes souterraines.

* Si $ETP = P$, $ETR = P +$ tout ou partie de la réserve en eau du sol jusqu'à l'épuisement de celle-ci.

Lorsque les réserves en eau du sol sont épuisées et que $ETP > ETR$, $ETP - ETR =$ "water-deficiency".

* Si $ETP = P$, $ETR = ETP = P$

Les réserves en eau du sol restent les mêmes que celles du mois précédent.

La valeur de ETR annuelle est la somme des douze valeurs ETR mensuelles. La valeur de ETR journalière est la division de l'ETR mensuelle par le nombre de jours du mois considéré.



ABSTRACT

Time Series Analysis : Application to the Temporal Evolution of Mining Exhaust in the Lorrain's Ferriferous Basin

Time series analysis presents two complementary aspects: description and prevision.

The descriptive analysis of a time series is a preliminary step to the modelisation and prevision. The descriptive tools are: autocorrelation coefficients, cross-autocorrelation coefficients, power spectrum and cross-power spectrum.

The prevision of a time series presents generally the principal goal. It can be done either from the informations of the time series itself in the past, or from one or several other time series. In consequence, we can distinguish the univariate prevision methods and the multivariate prevision methods.

The univariate prevision methods are: smoothing methods, decomposition methods and Box-Jenkins method. These methods can be convenient for some simple cases, particularly for short term analysis when a phenomenon continue on an horizon. But they have some disadvantages when used for medium and long terms analysis.

The multivariate prevision methods consist of discovering and mesuring the influence of independent variables on dependent variable Y_t . Without change the generality, we restraint our proposition for a case of an only independent variable X_t . X_t can be considered as the input of a linear system and Y_t the output. The relation of convolution between X_t and Y_t is generally admissible. The transfert function presents all the characteristics of the system. We can construct the output corresponding to an input by convolving the input by the transfert function of the system. Hence it's necessary to recognize the transfert function before any simulation and the problem of prevision can be converted to a probleme of deconvolution.

The classical and widely used methods of deconvolution are the spectral methods and regression methods (inverse problem). These methods work well if the input and output verify perfectly a relation of convolution. But in prattice, the conditions of linearity and stationnarity are never perfectly verified and the observations are always containinated of errors. Thus the relation of convolution should be written in a manner of including a residual term. The intervention of the residual term excites a difficulty of deconvolution because in fact the solution is not unitary. To overcome this difficulty, some new methods have been proposed recently (method of projection and geostatistical method). It appears to us that the transfert function method of Box-Jenkins is very promising because it combines the concepts of regression to models ARIMA, presenting un decisive theoretical avantage of this method.

As an application of analysis of time series, the analysis of mensal data of mining exhaust in the lorrain's ferriferous basin has permitted to distinguish two types of mines: the stationnary mines and non stationnary mines. A general model relating the exhaust to its principal influential factors (effective rainfall and cumulated extraction of mineral) has been established.

The non stationnarity of exhaust in the time can be explained by the cumulated extraction of mineral. The heterogeneity of exhaust in the space can be explained generally by the local geological conditions and the structual conditions can explain partially the anomaly.

The established model can be used to predict the mining exhaust in a future moment or to reconstitute the exhaust data.

Otherwise, in the gestion of basin, the drought poses frequently a grave problem to the alimentation of potable water. It's important to study the special compartments of ground aquifert or of mines facing to an annual or multiannual drought. This study will furnish a more precise prevision. An intervention model can be adopted. It consists of formulating some criteria on the drought, introducing some indicating variables and studying the compartments of mining exhaust during the period of transition.