



**HAL**  
open science

# Modèles de distribution et stratégies d'échantillonnage des populations bactériennes dans un réseau de distribution publique

Armand Maul

► **To cite this version:**

Armand Maul. Modèles de distribution et stratégies d'échantillonnage des populations bactériennes dans un réseau de distribution publique. Bactériologie. Université Paul Verlaine - Metz, 1985. Français. NNT : 1985METZ004S . tel-01775525

**HAL Id: tel-01775525**

**<https://hal.univ-lorraine.fr/tel-01775525>**

Submitted on 24 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## 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](mailto: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.cfcopies.com/V2/leg/leg_droi.php)

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

**THESE**

présentée à

L'UNIVERSITÉ DE METZ

en vue d'obtenir le titre de

DOCTEUR D'ÉTAT ÈS-SCIENCES

par

**Armand MAUL****Modèles de distribution et stratégies  
d'échantillonnage des populations  
bactériennes dans un réseau  
de distribution publique**

soutenue publiquement le 9 septembre 1985 devant la Commission d'Examen

**MEMBRES DU JURY :****Président : M. J.-C. BLOCK****Examineurs : MM. C. DUBY  
A.-H. EL-SHAARAWI  
O. NASI  
M. RIZET  
P. VASSEUR**BIBLIOTHEQUE UNIVERSITAIRE  
-METZ

N° inv.	1985 008 S
Cote	S/M <sub>3</sub> 85/4
Loc	Magasin

A Corinne, Franck et Olivier

A toute ma famille

A tous mes amis

## **AVANT-PROPOS**

Ce travail, qui s'appuie sur deux disciplines : les statistiques et la microbiologie appliquée, a été réalisé au Laboratoire de Microbiologie du Centre des Sciences de l'Environnement de Metz. De son directeur, Monsieur le Professeur Jean-Claude BLOCK, qui nous a inspiré ce thème de recherche, trouve ici l'expression de ma sincère gratitude et de mon profond respect pour les conseils et les encouragements qu'il n'a cessé de me prodiguer tout au long de cette étude.

Je remercie le Docteur A.H. EL-SHAARAWI, Membre de l'Institut International de Statistiques, pour m'avoir accueilli dans son laboratoire (Canada Centre for Inland Waters) à Burlington et pour m'avoir initié à certaines méthodes statistiques au cours de son année sabbatique passée à l'Université de Metz. Il a toujours apporté une bienveillante attention à mes travaux et m'a guidé, avec beaucoup de compétence et d'amitié, dans la réalisation de cette étude.

J'adresse mes plus vifs remerciements aux autres personnalités scientifiques qui m'ont fait l'honneur de participer au jury de cette thèse :

- Madame le Professeur C. DUBY qui a bien voulu s'intéresser à ce travail et qui m'a fait profiter de ses remarques judicieuses ;
- Monsieur O. NASI, Maître-assistant, qui a chaleureusement accepté de faire la critique de ce travail et qui, par son aide constructive, a permis une approche plus rigoureuse dans l'exposé de certains problèmes statistiques ;
- Madame M. RIZET qui, par sa compétence et ses recherches antérieures m'a ainsi fait bénéficier de ses connaissances en bactériologie des eaux ;
- Madame le Professeur P. VASSEUR pour l'intérêt qu'elle a porté à mon travail et l'honneur qu'elle m'a fait en participant à ce jury.

Mes remerciements vont également au Docteur S.R. ESTERBY (Canada Centre for Inland Waters) pour sa gentillesse à l'occasion de mon séjour à Burlington où elle a mis à notre disposition un programme informatique destiné à effectuer la procédure de groupement.

Qu'il me soit permis d'associer à ces remerciements toutes les personnes ayant participé à la collecte des échantillons ainsi que M.A. DOLLARD pour sa compétence et la minutie dont elle a fait preuve au cours de sa contribution technique très qualifiée.

J'ai été sensible à l'esprit de camaraderie et d'entraide que toutes les personnes du laboratoire ont constamment manifesté à mon égard.

Ce travail a été réalisé dans le cadre d'un contrat avec  
l'Agence Financière de Bassin Rhône-Méditerranée-Corse (Lyon)  
et la Société Lyonnaise des Eaux (Paris)

## **SOMMAIRE**

INTRODUCTION .....	1
PREMIERE PARTIE - OUTILS MATHEMATQUES ET ANALYTIQUES UTILISES POUR LA MODELISATION DE LA STRUCTURE D'ABONDANCE DU PEUPEMENT BACTERIEN D'UN RESEAU DE DISTRIBUTION PUBLIQUE .....	4
<u>CHAPITRE I - MODELES MATHEMATQUES DE DISTRIBUTION D'ABONDANCE</u> .....	7
I.1. Le modèle de Poisson .....	8
I.1.1. Caractérisation de la loi de Poisson .....	8
I.1.2. Généralisation de l'indice de dispersion de Fisher .....	9
I.2. Le modèle binomial négatif .....	10
I.2.1. Caractérisation de la loi binomiale négative .....	10
I.2.2. Ajustement des données expérimentales à une distribution théorique .....	13
I.2.3. Tests de comparaison .....	16
I.2.3.1. Comparaison de plusieurs estimations du paramètre k .....	17
I.2.3.2. Comparaison de plusieurs estimations du paramètre p .....	18
I.3. Recherche d'une structure d'abondance du peuplement bactérien ..	19
I.3.1. La méthode de groupement .....	19
I.3.2. L'ordination en espace réduit .....	23
<u>CHAPITRE II - CHOIX D'UNE METHODE DE TITRAGE DES BACTERIES</u> .....	25
II.1. Comparaison expérimentale de trois méthodes de titrage : méthodes énumératives par filtration sur membrane ou par inclusion en gélose et méthode quantique en milieu liquide .	27
II.1.1. Matériel et méthodes .....	28
II.1.1.1 Echantillons d'eau .....	28
II.1.1.2 Méthode des membranes filtrantes .....	28
II.1.1.3 Méthode par inclusion en gélose .....	28
II.1.1.4 Méthode quantique en milieu liquide .....	28
II.1.1.5 Analyse statistique des quantifications bactériennes .....	29
II.1.2. Protocole expérimental .....	31
II.1.3. Résultats .....	33
II.1.4. Discussion .....	35
II.2. Comparaison expérimentale de quelques milieux de culture .....	37
II.2.1. Matériel et méthodes .....	37
II.2.2. Protocole expérimental .....	38
II.2.3. Résultats .....	39
II.2.4. Discussion .....	45
<u>CHAPITRE III - MISE EN EVIDENCE PRELIMINAIRE DE L'HETEROGENEITE DU MILIEU HYDRIQUE A DIFFERENTES ECHELLES D'OBSERVATION</u> .....	46
III.1. Matériel et méthodes .....	49
III.1.1. Collecte des échantillons d'eau .....	49
III.1.2. Numération bactérienne et analyse statistique des résultats ...	50
III.2. Protocole expérimental .....	51
III.3. Résultats .....	51
III.4. Discussion .....	60
DEUXIEME PARTIE - ETUDE EXPERIMENTALE SUR LE RESEAU DE DISTRIBUTION PUBLIQUE DE METZ .....	63
<u>CHAPITRE I - ETUDE DESCRIPTIVE DU RESEAU DE DISTRIBUTION DE LA VILLE DE METZ</u> .....	65
I.1. Caractéristiques générales du réseau et capacité de production .	66
I.2. Filières de traitement de l'eau brute .....	66
I.3. Capacité de stockage et distribution .....	68

## CHAPITRE II - STRUCTURE SPATIO-TEMPORELLE DE L'HETEROGENEITE

<u>BACTERIENNE</u> .....	69
II.1. Matériel et méthodes .....	72
II.1.1. Localisation des points de prélèvement et collecte des échantillons d'eau .....	72
II.1.2. Analyse microbiologique et physico-chimique des échantillons.	72
II.1.3. Analyse statistique des résultats .....	74
II.1.3.1 Ajustement d'un modèle de distribution d'abondance aux données expérimentales .....	74
II.1.3.2 Analyses de corrélation .....	74
II.1.3.3 Méthode de groupement .....	75
II.1.3.4 Analyse en composantes principales .....	75
II.2. Protocole expérimental .....	75
II.3. Résultats .....	77
II.3.1. Présentation des données bactériologiques .....	77
II.3.2. Configuration de l'hétérogénéité spatio-temporelle .....	81
II.3.2.1 Reproductibilité de la structure d'abondance des populations bactériennes .....	81
II.3.2.2 Zones de densité bactérienne .....	85
II.3.2.3 Modélisation de la structure d'abondance du peuplement bactérien .....	88
II.3.3. Concentration bactérienne <u>versus</u> paramètres physico-chimiques de l'eau .....	93
II.3.3.1 Temps de séjour .....	93
II.3.3.2 Chlore résiduel .....	95
II.3.3.3 Température de l'eau .....	96
II.4. Discussion .....	98
TROISIEME PARTIE - STRATEGIES D'ECHANTILLONNAGE .....	101
<u>CHAPITRE I - PROPOSITION THEORIQUE D'UNE STRATEGIE DE PRELEVEMENT POUR LE CONTROLE DE LA QUALITE BACTERIOLOGIQUE DE L'EAU DISTRIBUEE</u> .....	104
I.1. Echantillonnage pour une évaluation instantanée de la qualité bactériologique de l'eau .....	105
I.1.1. Détermination du nombre d'échantillons à prélever .....	105
I.1.2. Répartition optimale de l'effort d'échantillonnage dans les zones d'un réseau stratifié .....	108
I.2. Echantillonnage progressif à but décisionnel .....	108
<u>CHAPITRE II - APPLICATION DES PLANS D'ECHANTILLONNAGE THEORIQUES AUX DONNEES BACTERIOLOGIQUES DU RESEAU MESSIN</u> .....	111
II.1. Stratégies d'échantillonnage en fonction d'une modélisation .	112
II.1.1. Bilan instantané .....	112
II.1.2. Bilan échelonné .....	117
II.2. Discussion .....	119
CONCLUSION .....	122
REFERENCES BIBLIOGRAPHIQUES .....	126
ANNEXES .....	135

## LISTE DES FIGURES

- FIGURE 1 : Schéma du protocole expérimental pour la comparaison des titres bactériens obtenus par trois méthodes de titrage (membranes filtrantes, inclusion en gélose et numération en milieu liquide).
- FIGURE 2 : Evolution du nombre des colonies recensées sur trois milieux de culture (GN, R-2A et m-SPC) en fonction du temps d'incubation.
- FIGURE 3 : Enveloppe de quelques distributions représentatives de la variabilité observée à différentes échelles d'échantillonnage pour une valeur commune de la moyenne arbitrairement fixée à 100.
- FIGURE 4 : Représentation schématique du réseau messin (adapté d'après Société Mosellane des Eaux).
- FIGURE 5 : Représentation schématique des principales caractéristiques hydrauliques du réseau de distribution étudié.
- FIGURE 6 : Localisation et identification des points de prélèvement.
- FIGURE 7 : Distribution des fréquences cumulées croissantes des titres bactériens observés pour chaque campagne d'échantillonnage.
- FIGURE 8 : Zones de densité bactérienne observées au cours de six campagnes d'échantillonnage effectuées sur le réseau de distribution de la ville de Metz.
- FIGURE 8 a : campagne d'échantillonnage du 13 décembre
- FIGURE 8 b : campagne d'échantillonnage du 10 janvier
- FIGURE 8 c : campagne d'échantillonnage du 21 février
- FIGURE 8 d : campagne d'échantillonnage du 03 avril
- FIGURE 8 e : campagne d'échantillonnage du 15 mai
- FIGURE 8 f : campagne d'échantillonnage du 26 juin
- FIGURE 9 : Partition du système de distribution en quatre zones de densité bactérienne obtenues à partir de l'exploitation des résultats des six campagnes d'échantillonnage.
- FIGURE 10 : Partition du système de distribution en cinq zones de densité bactérienne.
- FIGURE 11 : Zones de densité bactérienne représentées dans l'espace réduit associé aux deux premiers axes principaux.
- FIGURE 12 : Représentation des six campagnes d'échantillonnage (descripteurs) dans l'espace réduit des deux premiers axes principaux.
- FIGURE 13 : Variation temporelle de la concentration bactérienne pour chaque zone de densité.
- FIGURE 14 : Evaluation du temps de séjour approximatif de l'eau dans les canalisations.

- FIGURE 15 : Régression de la moyenne géométrique de la densité bactérienne selon la température moyenne de l'eau.
- FIGURE 16 : Probabilité d'acceptation en fonction de la densité bactérienne moyenne.
- FIGURE 17 : Nombre d'échantillons à prélever en fonction de  $k$ , pour différents niveaux de la densité bactérienne  $\lambda$  et du risque  $\beta$ .
- FIGURE 18 : Probabilité d'acceptation (risque  $\beta$ ) en fonction du nombre d'échantillons analysés.
- FIGURE 19 : Illustration de la méthode d'échantillonnage progressif pour la zone 1 à partir des données observées lors des campagnes de prélèvements du 13 décembre et du 26 juin.
- FIGURE 20 : Organigramme de la procédure décisionnelle lorsque le modèle de distribution est défini progressivement à partir des données disponibles.

## ANNEXES

- FIGURE I : Résolution de l'équation de vraisemblance par la méthode de Newton.

**MODELES DE DISTRIBUTION ET STRATEGIES  
D'ECHANTILLONNAGE DES POPULATIONS BACTERIENNES  
DANS UN RESEAU DE DISTRIBUTION PUBLIQUE**

**RESUME**

L'étude quantitative des populations bactériennes contenues dans l'eau de distribution nécessite une bonne méthode d'énumération bactérienne ainsi qu'une stratégie d'échantillonnage appropriée : en l'occurrence les modalités du programme de prélèvement doivent être définies en fonction des composantes spatiale et temporelle de l'hétérogénéité bactérienne caractérisant le système de distribution considéré.

**METHODES STATISTIQUES ET ANALYTIQUES**

La loi de Poisson et la distribution binomiale négative ont souvent été utilisées pour caractériser la structure d'abondance du peuplement bactérien. Aussi, dans l'exposé des développements théoriques propres à ces deux distributions, une place prépondérante a été consacrée à l'étude des techniques d'estimation et de comparaison des paramètres qui les définissent. De plus, une étude expérimentale a été menée dans le but de pouvoir effectuer rationnellement le choix d'une technique fiable destinée à l'énumération des bactéries hétérotrophes. A cet égard, une méthode statistique générale s'inspirant du test du rapport de vraisemblance est décrite afin de permettre la comparaison des densités bactériennes moyennes issues de plusieurs séries expérimentales. Cette approche s'avère nettement plus rigoureuse que les moyens habituellement mis en oeuvre pour l'analyse statistique des données bactériologiques (e. g., l'analyse de variance), qui nécessitent le plus souvent une transformation préalable des données. Le test du rapport de vraisemblance, dont l'application requiert la connaissance de la nature exacte des lois de probabilité théoriques suivies par les variables, constitue un outil intéressant pour effectuer la comparaison des densités bactériennes. La méthode est particulièrement avantageuse lorsque les variables comparées obéissent à des lois de distribution différentes mais qui sont cependant connues. La mise en oeuvre de ce test est illustrée par la comparaison expérimentale de trois techniques de quantification des bactéries hétérotrophes à 20°C contenues dans une eau de distribution. Cette expérience a montré que si la méthode par inclusion en gélose (IG) et la méthode des mem-

branes filtrantes (MF) procurent en substance des résultats comparables, il n'en va pas de même avec la microméthode de titrage pour laquelle l'estimation de la concentration bactérienne selon la technique du nombre le plus probable (N.P.P.) peut quelquefois conduire à des estimations considérablement différentes. Les résultats font notamment apparaître une forte interaction entre la méthode utilisée et l'échantillon analysé. Par suite, trois milieux de culture, à savoir : la gélose nutritive, le milieu R-2A et le milieu m-SPC, ont été testés pour ce qui concerne le dénombrement des bactéries hétérotrophes par la méthode des membranes filtrantes après 48, 72 et 96 heures d'incubation à 20°C. Bien que le nombre des colonies apparues sur chacun des trois milieux soit approximativement le même au terme d'une incubation de 96 heures, des différences statistiquement significatives ont cependant été observées pour des temps d'incubation intermédiaires. Ce phénomène est probablement imputable aux différences relatives de la vitesse de croissance bactérienne caractérisant les milieux de culture. Finalement, la méthode des membranes filtrantes avec une période d'incubation de 72 heures sur milieu R-2A a été retenue pour tous les dénombrements bactériens effectués ultérieurement dans ce travail.

#### VARIATION SPATIALE ET TEMPORELLE DE LA DENSITE BACTERIENNE

La variabilité caractérisant les populations bactériennes à différents niveaux de l'échelle d'observation a été étudiée à l'aide d'une généralisation de l'indice de dispersion de Fisher et du paramètre  $k$  de la distribution binomiale négative. Cette démarche a permis de mettre en évidence une structure hiérarchisée de l'hétérogénéité dont l'ampleur augmente rapidement avec la taille du domaine d'échantillonnage. Six campagnes d'échantillonnage, menées selon un protocole de sélection systématique, ont été réalisées sur le réseau de distribution de la ville de Metz en vue de connaître la configuration de l'hétérogénéité spatiale et temporelle de l'abondance des bactéries hétérotrophes dans l'ensemble du réseau. Une méthode de groupement centroïde et non-hiérarchique a été utilisée afin de diviser le système de distribution en zones correspondant à des niveaux distincts de la densité bactérienne (cf. figure 9). La structure spatiale de l'hétérogénéité a montré sur ce point une grande reproductibilité. La distribution des micro-organismes à l'intérieur des zones pouvant, par ailleurs, être modélisée par une loi binomiale négative, le réseau étudié apparaît comme un système composite formé de plusieurs sous-systèmes hétérogènes. L'étroitesse de la relation de causalité existant entre la structure de la répartition spatiale des bactéries et certaines variables physico-chimiques du système est indé-

niable. Compte tenu des caractéristiques hydrodynamiques essentielles du réseau, lesquelles sont inhérentes au tracé et à la disposition des canalisations principales, il apparaît que les zones de plus forte concentration bactérienne correspondent à des teneurs en chlore résiduel plus faibles ainsi qu'à des temps de séjour prolongés de l'eau, notamment dans les réservoirs, avant d'atteindre les différentes aires de distribution. Les variations mensuelles de la concentration bactérienne globale du système de distribution ont montré une forte corrélation avec la température de l'eau distribuée. Des divergences sensibles ont cependant été observées entre les zones pour ce qui concerne l'amplitude de ces variations temporelles.

#### STRATEGIE D'ECHANTILLONNAGE POUR LE CONTROLE DE QUALITE

L'information, concernant la variabilité spatiale et temporelle, extraite des données bactériologiques disponibles, a ensuite été exploitée afin de définir les modalités d'un programme de prélèvement destiné aux contrôles de qualité ultérieurs. Si ces contrôles sont effectués dans un but décisionnel au regard d'une norme donnée (e. g., la concentration maximale acceptable de 100 bactéries/mL actuellement préconisée par les pays membres de la C.E.E.), un critère, permettant une répartition optimale des stations de prélèvement par zone, est donné lorsque l'évaluation doit présenter un caractère instantané. Le choix de ces stations est effectué : soit en fonction du risque d'échantillonnage (i. e., le risque lié à une décision erronée) accepté (cf. figure 17) soit en fonction d'un nombre total de stations prédéterminé. Dans le cas d'un bilan échelonné de la qualité de l'eau, l'échantillonnage progressif peut conduire à une réduction substantielle de l'effort de prélèvement. L'intérêt et la mise en application de ces méthodes sont illustrés à l'aide des données bactériologiques obtenues pour le réseau de distribution de la ville de Metz.

Bien que les exemples présentés dans ce travail intéressent exclusivement les bactéries hétérotrophes ainsi qu'une réglementation particulière, les modes d'échantillonnage exposés se généralisent facilement à d'autres spécifications ou encore à d'autres critères de qualité dont les mesures sont conformes au modèle binomial négatif.

## SUMMARY

The quantitative analysis for bacteria contained in drinking water requires both a good method for bacterial enumeration and an appropriate sampling strategy which has to be defined in consideration of the spatial and the temporal dispersion pattern of bacteria in the distribution system studied.

## STATISTICAL AND ANALYTICAL METHODS

The Poisson and the negative binomial probability distributions have often been used to characterize the bacterial dispersion pattern. Thus, the theoretical aspects of those two distributions are outlined with special emphasis on parameter estimations and comparisons. Further, an experiment was carried out in order to guide the choice of a valid evaluative tool for enumerating heterotrophic bacteria. In this regard, a general statistical procedure based on the likelihood ratio test is presented for the purpose of comparing estimates of mean bacterial density derived from different sets of data. This approach is much more appropriate than the conventional ways of analysing bacteriological results (e.g., analysis of variance), which usually require previous transformation of the data. The likelihood ratio test, which assumes knowledge of the exact theoretical probability distributions of the variables, offers an interesting way to compare bacterial densities, especially when the estimates to be compared follow different, but known, probability distributions. An illustrative application of the method compares three distinct titration techniques for enumerating heterotrophic bacteria in drinking water, at 20°C incubation temperature. It was shown that both the standard plate count (SPC) and the membrane filter (MF) procedures supplied substantially the same information, whereas the microplate technique using the most probable number (MPN) for bacterial enumeration could yield considerably different estimates. The results consistently indicate a strong interaction between the technique used and the sample analysed. Three different media (nutrient agar, R-2A low nutrient agar and m-SPC agar) were then evaluated for enumerating heterotrophic bacteria, using the MF technique at 48, 72 and 96 hours of incubation time at 20°C. Although the media recovered approximately the same numbers of bacteria after 96 hours of incubation, statistically significant discrepancies occurred after intermediate periods of incubation, perhaps because the relative rates of bacterial growth differed among media. Finally, it was decided to use the MF technique with R-2A medium and 72 hours incubation period for all further bacterial enumerations performed in this work.

## SPATIAL AND TEMPORAL VARIATION OF THE BACTERIAL DENSITY

The variability characterizing bacterial populations at different levels of the sampling area size was investigated by means of a generalization of Fisher's index of dispersion and also parameter  $k$  of the negative binomial distribution. This showed a hierarchical structured pattern of the heterogeneity, which increased very strongly with the size of the area sampled. Six bacteriological surveys were then performed on the drinking water distribution system of the city of Metz according to a systematic sampling design to determine the spatial and temporal distribution of heterotrophic bacteria in the whole network. A non-hierarchical nearest-centroid clustering method was used for dividing the water distribution system into zones corresponding to different levels of bacterial density (cf figure 9). The general pattern of the spatial heterogeneity showed a high degree of reproducibility. Since the frequency distributions of heterotrophic bacteria within the zones were compatible with the negative binomial distribution, the water distribution system studied may be considered as being composed of several heterogeneous subsystems. The consistency of this structured spatial dispersion pattern of bacteria in light of some physical and chemical characteristics of the system is evident. In consideration of the principal features of flow in the system relevant to the layout of water mains, the location of zones of highest bacterial concentrations have been attributed to lower levels of chlorine residuals and prolonged retention time of the water in the network, especially in the storage units, before reaching the various distribution areas. The monthly variation in the bacterial concentration of the entire system showed a marked increase which was concomitant with warmest water temperatures, nevertheless the zones were subject to noticeable discrepancies in the range of temporal variation (cf figure 13).

## SAMPLING DESIGN FOR MONITORING

Information on the spatial and temporal variability of bacteriological data was then used to develop a sampling design for use in future water quality monitoring. Under the assumption that the objective of monitoring is to determine whether or not the mean bacterial density of the water exceeds a specific standard (e. g., the 100 bacteria/mL EEC standard), a criterion is given which determines the optimal number of sampling stations allocated to each zone in case of a one-run sampling design. These stations are determined by assuming that either the risk of sampling (i. e., making the wrong decision) is prespecified (cf figure 17) or that the total number of

stations to be sampled is predetermined. If the evaluations of the bacteriological quality of the water are to be reported for specific intervals of time, then sequential sampling may result in a substantial reduction of the cost of sampling. These techniques are illustrated using the bacteriological data obtained from the drinking water distribution system of the city of Metz.

Although the examples given in this work use heterotrophic bacteria and a specific regulation, the sampling designs presented may be easily adapted to other regulations or even other bacteriological water quality data described by the negative binomial distribution.

## INTRODUCTION

De nombreux phénomènes peuvent induire de manière exogène ou endogène la présence et l'accroissement des populations bactériennes véhiculées par l'eau d'un réseau de distribution publique. La diversité et le caractère souvent imprévisible des mécanismes responsables de la propagation bactérienne, de même que le risque potentiel que représente l'eau de boisson dans la diffusion de certaines maladies infectieuses, ont conduit le législateur à fixer des normes de salubrité dont le respect est vérifié préventivement par des programmes de prélèvement organisés selon un plan déterminé.

Deux types de problèmes contribuent cependant à ébranler très sérieusement le crédit et le bien-fondé inspirés par les réglementations actuellement en vigueur et dont les modalités ont, à tout le moins, été déterminées sur des bases arbitraires. Il s'agit d'une part, de la méconnaissance relative de la signification et de l'impact épidémiologique caractérisant les micro-organismes recherchés, qui ne sont en fait que des indicateurs de contamination bactérienne eu égard aux difficultés soulevées par la mise en évidence directe des bactéries pathogènes, et d'autre part, de l'ensemble des problèmes méthodologiques et statistiques liés à l'hétérogénéité du milieu hydrique.

Outre la nécessité de disposer d'un outil analytique fiable, l'exploration du milieu hydrique et l'interprétation des résultats se heurtent à la variabilité de l'abondance bactérienne. Cette variabilité, tout en soulignant la valeur essentiellement ponctuelle de chaque mesure, entrave par là même l'extrapolation des résultats individuels au-delà d'une certaine fenêtre d'observation du domaine spatio-temporel. Aussi, il devient nécessaire d'interpréter les résultats globaux d'une campagne d'échantillonnage dans le contexte probabiliste qui les caractérise et dont les principaux aspects sont inhérents à la structure spatio-temporelle de l'hétérogénéité. La connaissance de cette structure sous-jacente s'avère par ailleurs indispensable pour la définition objective de programmes d'échantillonnage élaborés selon une exploitation rationnelle de l'information disponible. Le présent travail est entièrement tourné vers l'identification d'une hétérogénéité structurée et interprétable du peuplement bactérien d'un système de distribution. Les finalités de cette démarche sont à la fois explicatives et décisionnelles.

A cet effet, nos efforts se sont portés vers la mise en oeuvre de méthodes statistiques et de procédés d'investigations qui, à terme, doivent permettre d'arrêter les modalités d'un programme d'échantillonnage adapté au contrôle de la qualité bactériologique de l'eau d'un réseau modélisé.

La première partie de ce travail est consacrée à l'étude de quelques outils mathématiques et analytiques appropriés à l'exploration et à la description de la structure d'abondance du peuplement bactérien d'un réseau de distribution publique. Cette partie comprend trois chapitres.

Les développements théoriques d'une procédure de groupement permettant d'effectuer objectivement la stratification d'un réseau sont exposés dans le premier chapitre où nous avons également entrepris l'étude descriptive détaillée des lois de probabilité utilisées pour parachever la modélisation d'un système de distribution.

Le thème abordé dans le deuxième chapitre, qui est principalement orienté vers le choix d'une méthode de quantification bactérienne, sera l'occasion d'illustrer une méthode statistique adaptée à l'étude comparative de plusieurs séries de données microbiologiques obtenues par des techniques analytiques différentes.

Une étude préliminaire visant à apprécier l'ampleur de l'hétérogénéité du milieu hydrique, perceptible au niveau de quelques échelles d'observation privilégiées, fait l'objet du troisième chapitre.

La deuxième partie concerne la mise en oeuvre et l'analyse statistique des résultats d'un programme d'échantillonnage dont la finalité tient essentiellement dans la reconnaissance et l'identification de la configuration spatio-temporelle de l'hétérogénéité bactérienne du réseau de distribution de la ville de Metz.

Après une description sommaire du réseau messin, le deuxième chapitre de cette partie donne une représentation modélisée du réseau que nous tenterons d'interpréter au regard de quelques variables physico-chimiques jugées à la fois intéressantes et prépondérantes.

Faisant état des résultats établis antérieurement, la troisième partie contient les principes théoriques d'une stratégie d'échantillonnage destinée à optimiser le contrôle de la qualité bactériologique de l'eau distribué.

Dans un premier chapitre, nous verrons comment définir les modalités d'un programme de prélèvements, établi en fonction d'une modélisation, dans le cadre de deux modes d'échantillonnage répondant à des conceptions différentes de la procédure décisionnelle.

Dans un deuxième temps, les plans d'échantillonnage proposés et décrits de façon synthétique seront appliqués aux données bactériologiques recueillies dans le réseau de distribution de la ville de Metz, ce qui procurera une meilleure appréciation de l'incidence exercée par l'hétérogénéité du milieu hydrique sur l'effort d'échantillonnage à consentir.

## PREMIERE PARTIE

OUTILS MATHÉMATIQUES ET ANALYTIQUES UTILISÉS  
POUR LA MODÉLISATION DE LA STRUCTURE D'ABONDANCE  
DU PEUPEMENT BACTÉRIEN D'UN RÉSEAU DE DISTRIBUTION PUBLIQUE

L'étude quantitative des populations bactériennes dans les milieux aquatiques fait généralement apparaître une variabilité de la répartition des micro-organismes qui se manifeste à la fois dans le temps et dans l'espace. Cette hétérogénéité spatio-temporelle ne dépend pas uniquement des caractéristiques physico-chimiques de l'eau mais également de facteurs morphodynamiques, biogéographiques et climatiques. En particulier, la variation temporelle présente quelquefois un caractère systématique suivant un rythme journalier, saisonnier ou annuel en liaison plus ou moins étroite avec les conditions météorologiques (BELLAIR et al., 1977 ; BREITTMAYER et GAUTHIER, 1979 ; CHURCHLAND et al., 1982). De nombreux travaux consacrés à l'étude de la répartition spatiale des bactéries en milieu hydrique ont, par ailleurs, montré une tendance très marquée des micro-organismes à s'agréger. De telles associations ont été décrites à l'échelle du micron (BELL et ALBRIGHT, 1981 ; DUCKLOW et al., 1982 ; KOGURE et al., 1982 ; SEKI, 1968 ; SOROKIN, 1970 ; WILLIAMS, 1970), du centimètre (ASHBY et RHODES-ROBERTS, 1976), du mètre (PALMER et al., 1976) et même du kilomètre (EL-SHAARAWI et al., 1981).

L'existence d'une hétérogénéité microscopique sous forme d'agrégats et de complexes bactéries-particules soulève de sérieuses difficultés en microbiologie numérique car les méthodes de titrage usuelles ne distinguent pas les micro-organismes isolés des agrégats (HATTORI, 1985). Le résultat d'un titrage exprime de ce fait un nombre d'unités comprenant indistinctement des bactéries libres ou agrégées ainsi que des complexes bactéries-particules (i. e., unités formant colonie : UFC). Le nombre de bactéries estimé, dont la signification est exclusivement statistique, n'est lié à la densité bactérienne réelle du milieu étudié que par des facteurs de proportionnalité inconnus. Ces facteurs sont induits par de nombreux éléments tels que la méthode de titrage employée, l'origine et les caractéristiques de l'eau, la nature de la bactérie considérée, les phénomènes d'antagonisme et de compétition entre les différentes espèces présentes ainsi que par les interactions entre tous ces facteurs. Ainsi, les méthodes de titrage se caractérisent dans l'ensemble par un manque relatif de précision et de spécificité et n'ont, par conséquent, rien d'absolu. Dans ces conditions, il va de soi que les résultats d'un titrage bactérien n'ont qu'une valeur relative et doivent obligatoirement être accompagnés d'une description complète et précise du matériel et de la technique mise en oeuvre.

Par ailleurs, l'hétérogénéité macroscopique de la répartition des micro-organismes pose le problème de la représentativité des échantillons et entrave l'extrapolation des résultats obtenus à partir de faibles volumes d'eau lorsque la variabilité spatio-temporelle dépasse le seuil de précision des méthodes de mesure. L'étude de cette hétérogénéité dans un milieu hydrique déterminé peut se concevoir par l'utilisation de modèles mathématiques de distribution d'abondance visant à traduire la structure du peuplement bactérien de manière globale et purement formelle (HERRICKS et al., 1985 ; PIPES et CHRISTIAN, 1982).

La complémentarité des difficultés évoquées ci-dessus montre la nécessité de considérer conjointement tous les aspects propres à l'analyse quantitative en microbiologie des eaux et risque de rendre vain et infructueux tout effort réalisé unilatéralement. Ces considérations de portée très générale, et qui de ce fait s'appliquent également aux réseaux de distribution

publique, nous ont amenés, dans les chapitres suivants, à envisager successivement :

- l'étude théorique des principaux modèles mathématiques de distribution d'abondance utilisés en microbiologie appliquée,
- une expérimentation destinée à permettre le choix d'une méthode de titrage adaptée à la numération des bactéries hétérotrophes totales,
- enfin, une étude préliminaire ayant pour objectif de fournir une évaluation de la structure hiérarchique approximative de l'hétérogénéité bactérienne perceptible à différentes échelles d'observation dans un réseau de distribution publique.

## CHAPITRE I

### MODELES MATHEMATIQUES DE DISTRIBUTION D'ABONDANCE

L'ajustement d'une loi de probabilité théorique aux titres bactériens observés facilite la mise en oeuvre des méthodes statistiques pour l'exploitation et l'interprétation des résultats expérimentaux. Les paramètres de ces lois de distribution sont quelquefois utilisés pour tenter de décrire certaines caractéristiques, par exemple l'abondance ou la variabilité, de la répartition des micro-organismes dans le milieu étudié. Cette manière d'entreprendre l'étude sommaire de l'hétérogénéité des populations bactériennes a suscité de nombreux travaux visant à ajuster un modèle mathématique à une distribution empirique. Parmi les modèles les plus fréquemment utilisés, il faut citer : la loi de Poisson, la loi log-normale et la distribution binomiale négative (EL-SHAARAWI et PIPES, 1982 ; MUENZ, 1978 ; PIPES et CHRISTIAN, 1982).

La loi de Poisson, qui correspond à une répartition au hasard des bactéries dans le milieu hydrique (FISHER et al., 1922 ; GREENWOOD et YULE, 1917), est souvent postulée comme hypothèse de travail pour établir la formulation mathématique des estimateurs de la densité bactérienne associés à différentes méthodes de titrage (FONTVIEILLE et MAUL, 1985). Ce modèle, très singulier et peu réaliste, semble toutefois adapté à la description de la répartition spatiale des micro-organismes dans des volumes de liquide peu importants et bien homogénéisés (EL-SHAARAWI et al., 1981 ; FISHER et al., 1922 ; MAUL et BLOCK, 1983).

La surdispersion hiérarchisée, qui, en règle générale, est observable dans tous les milieux hydriques d'une certaine étendue, est totalement incompatible avec le modèle de Poisson. La description de celle-ci requiert des lois de distribution beaucoup moins contraignantes telles que la loi log-normale ou encore la loi binomiale négative suggérée par FISHER (1941). Ainsi, la loi log-normale (PIPES et al., 1977 ; THOMAS, 1955 ; VELZ, 1970) de même que la loi binomiale négative (EL-SHAARAWI et al., 1981 ; MUENZ, 1978 ; PIPES et al., 1977 ; THOMAS, 1955), encore plus souple que la précédente, ont souvent été utilisées pour ajuster des données bactériologiques obtenues à partir d'échantillons recueillis dans le temps ou/et l'espace et provenant d'un milieu hétérogène.

Malgré la préférence que les microbiologistes accordent plus volontiers à la distribution log-normale au détriment de la loi binomiale négative dont ils redoutent la lourdeur des calculs, l'usage de la fonction logarithme appelle quelques réserves inhérentes à la nature même de la transformation logarithmique ; celle-ci ne peut en effet s'appliquer qu'à des comptages différents de zéro. Certains auteurs ont proposé de pallier cet inconvénient en remplaçant les valeurs nulles par le nombre 1 (GREEN, 1979 ; PIPES et al., 1977), mais cet artifice n'est possible que dans certaines situations et notamment lorsque le nombre des valeurs nulles n'est pas trop élevé. Une autre difficulté réside dans le caractère continu de la distribution log-normale, ce qui par essence n'est pas très heureux étant donné la nature souvent discrète des titres bactériens estimés. En outre, le surcroît de calcul lié à l'utilisation de la loi binomiale négative ne doit plus constituer un obstacle majeur depuis l'avènement de l'informatique. Aussi, un rôle privilégié est accordé dans ce travail à la loi de Poisson ainsi qu'à la loi binomiale négative dont les aspects théoriques sont développés ci-après.

## 1.1. LE MODELE DE POISSON

### 1.1.1. Caractérisation de la loi de Poisson

Il est aisé de montrer que le nombre,  $X$ , de bactéries contenues dans des échantillons de volume unitaire prélevés successivement et de manière

indépendante dans un milieu parfaitement homogène, suit une loi de Poisson explicitée par la formule (1) :

$$P(X=r) = \frac{e^{-\lambda} \lambda^r}{r!} \tag{1}$$

dans laquelle  $\lambda$  désigne la densité bactérienne du milieu et  $r$  représente le nombre de micro-organismes contenus dans l'échantillon analysé. La distribution de Poisson est entièrement définie par le paramètre  $\lambda$  et possède de plus la particularité d'avoir une variance et une moyenne toutes deux égales à  $\lambda$ .

La fonction caractéristique de  $X$  est donnée par l'expression  $\varphi(t) = e^{\lambda(e^{it}-1)}$ .

Il découle immédiatement des propriétés des fonctions caractéristiques (VAN DER WAERDEN, 1967) que :

- . la variable aléatoire  $aX$ , où le coefficient  $a$  est un facteur multiplicatif différent de 1, n'est pas distribuée suivant une loi de Poisson,
- . la somme de plusieurs variables aléatoires indépendantes  $X_1, X_2, \dots, X_n$ , distribuées respectivement suivant une loi de Poisson de moyenne  $\lambda_1, \lambda_2, \dots, \lambda_n$ , est encore une variable aléatoire de Poisson de moyenne  $\lambda_1 + \lambda_2 + \dots + \lambda_n$ .

Ces deux propriétés auront une incidence directe sur la manière d'exploiter les résultats obtenus par les méthodes de titrage de type énumératif.

### 1.1.2. Généralisation de l'indice de dispersion de Fisher

La conformité d'une série  $r_1, r_2, \dots, r_n$  de résultats expérimentaux au modèle de Poisson peut être vérifiée au moyen de l'indice de dispersion de Fisher (FISHER et al., 1922). Cet indice s'exprime par la quantité

$$D^2 = \frac{(n-1) s^2}{\bar{r}} \tag{2}$$

où  $\bar{r}$  représente la moyenne et  $s^2$  la variance de la distribution empirique. Lorsque le modèle de Poisson est adapté à la description des données,  $D^2$  suit approximativement une loi du  $\chi^2$  avec  $n-1$  degrés de liberté. Ce résultat se démontre comme suit :

Si  $r_1, r_2, \dots, r_n$  désignent  $n$  réalisations indépendantes d'une même variable aléatoire de Poisson de moyenne inconnue  $\lambda$ , la fonction de vraisemblance associée au  $n$ -uplet  $(r_1, r_2, \dots, r_n)$  s'écrit

$$L(\lambda) = \frac{e^{-n\lambda} \lambda^T}{r_1! r_2! \dots r_n!} \tag{3}$$

avec  $T = r_1 + r_2 + \dots + r_n$ .

La variable aléatoire  $T$  est une estimation exhaustive du paramètre  $\lambda$  (VAN DER WAERDEN, 1967) ce qui signifie que la distribution conditionnelle de  $(r_1, r_2, \dots, r_n)$  par rapport à  $T$  est exempte de  $\lambda$ , permettant ainsi de tester directe-

ment l'adéquation du modèle de Poisson à partir des données expérimentales. Cette distribution conditionnelle est donnée par la formule (4) :

$$P(r_1, r_2, \dots, r_n/T) = \frac{T!}{r_1! r_2! \dots r_n!} \left(\frac{1}{n}\right)^{r_1} \dots \left(\frac{1}{n}\right)^{r_n} \quad (4)$$

qui définit une distribution multinomiale comprenant n classes équiprobables.

D'après (4) le test du  $\chi^2$  de Pearson appliqué à la distribution multinomiale donne exactement l'indice de dispersion de Fisher, à savoir :

$$\chi^2 = \sum_{i=1}^n \frac{(r_i - \bar{r})^2}{\bar{r}} \quad (5)$$

L'utilisation de l'indice de dispersion de Fisher pour tester l'homogénéité d'une série de résultats expérimentaux peut se généraliser au cas où les volumes d'eau analysés sont inégaux.

En effet, soient  $r_1, r_2, \dots, r_n$  les nombres de bactéries observés pour des échantillons dont les volumes sont respectivement égaux à  $v_1, v_2, \dots, v_n$ . La probabilité associée au n-uplet  $(r_1, r_2, \dots, r_n)$  s'exprime alors par

$$L(\lambda) = \frac{e^{-\lambda v} (\lambda v_1)^{r_1} \dots (\lambda v_n)^{r_n}}{r_1! r_2! \dots r_n!} \quad (6)$$

où  $v = v_1 + v_2 + \dots + v_n$ .

Dans ce cas, T est encore une estimation exhaustive du paramètre  $\lambda$  et la distribution conditionnelle de  $(r_1, r_2, \dots, r_n)$  par rapport à T devient

$$P(r_1, r_2, \dots, r_n/T) = \frac{T!}{r_1! r_2! \dots r_n!} \left(\frac{v_1}{v}\right)^{r_1} \dots \left(\frac{v_n}{v}\right)^{r_n} \quad (7)$$

Cette formule définit une loi multinomiale comprenant n classes associées à des probabilités inégales. Le test du  $\chi^2$  correspondant à cette situation est réalisé sur la statistique

$$D^2 = \sum_{i=1}^n (r_i - \frac{T v_i}{v})^2 / \frac{T v_i}{v} \quad (8)$$

qui constitue une généralisation de l'indice de dispersion de Fisher applicable au cas d'échantillons ayant des volumes inégaux.

## 1.2. LE MODELE BINOMIAL NEGATIF

### 1.2.1. Caractérisation de la loi binomiale négative

Toute hétérogénéité du milieu étudié de même que toute variation dans le volume des échantillons analysés se traduit inexorablement par une augmentation de la variance dont la valeur devient supérieure à la moyenne des titres bactériens observés. Le modèle de Poisson n'est alors plus approprié car la valeur de  $\lambda$  n'est pas constante pour tous les échantillons analysés.

S'il est admis, comme l'avait préconisé Fisher (1941), que la densité bactérienne moyenne  $\lambda$  d'un échantillon est distribuée suivant une loi gamma de paramètres  $k$  et  $p$ , notée  $\Gamma(k,p)$ , la densité de probabilité définissant la distribution de  $\lambda$  est donnée par la relation

$$f(x) = \frac{1}{(k-1)!} p^{-k} x^{k-1} e^{-x/p} \quad (9)$$

Dans la relation (9) le terme  $(k-1)!$  représente l'image de  $k$  par la fonction gamma, introduite par Euler au XVIII<sup>e</sup> siècle. Cette fonction, qui fournit une interpolation de la fonction factorielle, est définie pour tout nombre réel  $k$  strictement positif par

$$\Gamma(k) = \int_0^{+\infty} x^{k-1} e^{-x} dx = (k-1)!$$

La probabilité d'observer  $r$  micro-organismes dans un échantillon quelconque est obtenue par intégration à l'aide de la formule (10)

$$P(r) = \int_0^{+\infty} e^{-x} \frac{x^r}{r!} \frac{1}{(k-1)!} p^{-k} x^{k-1} e^{-x/p} dx \quad (10)$$

Le calcul de cette intégrale conduit à l'expression

$$P(r) = \frac{(k+r-1)!}{r! (k-1)!} \frac{p^r}{(1+p)^{k+r}} \quad (11)$$

qui définit la loi binomiale négative de paramètres  $k$  et  $p$ , laquelle sera parfois notée BN ( $k,p$ ) dans la suite de ce travail. Lorsque  $r$  décrit l'ensemble  $N$  des entiers naturels, les probabilités  $P(r)$ , calculées selon la formule (11), sont également les termes du développement du binôme  $(q-p)^{-k}$  avec  $q = p + 1$ , ce qui justifie l'appellation donnée à cette loi de distribution. La loi BN ( $k,p$ ) possède une moyenne,  $\lambda$ , égale au produit  $p k$  et une variance donnée par  $\sigma^2 = p (1+p)k$ . La substitution de  $p$  par  $\lambda/k$  dans l'expression de  $\sigma^2$  conduit à une autre formulation de la variance, en l'occurrence  $\sigma^2 = \lambda + \lambda^2/k$ , faisant apparaître  $\sigma^2$  sous la forme d'une somme de deux composantes, à savoir :  $\lambda$  et  $\lambda^2/k$ . La première de ces deux composantes, exprimée par la densité moyenne  $\lambda$ , correspond à une homogénéité parfaite et a déjà été mentionnée dans le modèle de Poisson alors que la seconde composante,  $\lambda^2/k$ , représente la contribution à la variance totale induite par l'hétérogénéité du milieu ou les conditions de prélèvement.

### Fonctions caractéristiques

- La fonction caractéristique associée à la distribution  $\Gamma(k,p)$  est définie par l'expression  $\varphi(t) = (1-itp)^{-k}$ .

A partir des moments d'ordre 1 et 2 obtenus par le développement en série de  $\varphi$  au voisinage de zéro, il est facile d'établir que l'espérance mathématique et la variance d'une variable aléatoire correspondant à cette loi de probabilité sont respectivement égales à  $\lambda = p k$  et  $k p^2 = \lambda^2/k$ . Il est intéressant de noter que  $\lambda^2/k$  représente la deuxième composante de la variance caractérisant la loi BN ( $k,p$ ) mentionnée ci-dessus. L'intérêt pratique de cette remarque tient au fait que, lorsque  $k$  est petit par rapport à  $\lambda$  (i. e.,  $p$  est suffisamment grand), l'expression de la variance de la loi BN ( $k,p$ ) est dominée par la composante  $\lambda^2/k$  imputable à la variabilité systématique du milieu hydrique étudié. Aussi quand le rapport  $p = \lambda/k$  est grand, l'enveloppe

de l'histogramme de la distribution BN (k,p) évolue vers la courbe de distribution de la loi  $\Gamma(k,p)$ .

- Il est aisé de montrer que la fonction caractéristique correspondant à la loi BN (k,p) est définie par  $\varphi(t) = (1 + p (1-e^{it}))^{-k}$ .

La connaissance de ces fonctions et des propriétés générales afférentes aux fonctions caractéristiques permettent d'aboutir aux propositions suivantes dont certaines auront une importance primordiale dans la suite de ce travail.

Proposition 1 :

La distribution BN (k,p) converge vers la distribution de Poisson de paramètre  $\lambda = p k$  lorsque  $k$  tend vers  $+\infty$ , (le produit  $\lambda = p k$  est supposé constant).

En effet,  $\forall t \in \mathbb{R} \quad \lim_{k \rightarrow +\infty} (1 + p (1-e^{it}))^{-k} = e^{\lambda(e^{it}-1)}$

or la fonction définie par  $\varphi(t) = e^{\lambda(e^{it}-1)}$  est continue en 0 et représente la fonction caractéristique d'une variable aléatoire de Poisson. La proposition 1 résulte immédiatement de la réciproque du théorème de continuité de Paul Lévy (METIVIER, 1972). Concrètement, ce résultat signifie que pour les grandes valeurs de  $k$  la distribution BN (k,p) est approchée de manière satisfaisante par une distribution de Poisson.

Proposition 2 :

Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes toutes distribuées suivant la même loi BN (k,p). La variable  $X = X_1 + \dots + X_n$  a pour loi de probabilité la loi BN (n k, p).

Démonstration

Si  $\varphi_{X_i}(t) = (1 + p (1 - e^{it}))^{-k}$  désigne la fonction caractéristique commune à chacune des variables  $X_i$  alors la fonction caractéristique de  $X$  est donnée par

$$\varphi_X(t) = \prod_{i=1}^n \varphi_{X_i}(t) = (1 + p (1 - e^{it}))^{-nk}.$$

Or cette dernière expression est la fonction caractéristique de la loi BN (n k, p).

Proposition 3 :

Si la variable  $X$  suit la loi  $\Gamma(k,p)$  alors la variable  $aX$  ( $a \in \mathbb{R}^{+*}$ ) suit la loi  $\Gamma(k, a p)$ .

En effet si  $\varphi_X(t) = (1 - i t p)^{-k}$  définit la fonction caractéristique de  $X$ , la fonction caractéristique de  $aX$  est déterminée par  $\varphi_{aX}(t) = \varphi_X(at) = (1-itap)^{-k}$ .

Remarque : l'énoncé de la proposition 3 ne s'applique pas au cas d'une variable  $X$  distribuée suivant une loi binomiale négative. Dans certaines

situations, notamment lorsque le rapport  $p = \lambda/k$  est suffisamment grand, il sera toutefois commode d'admettre que si  $X$  suit la loi BN ( $k, p$ ) alors  $aX$  ( $a \in \mathbb{R}^{+*}$ ) suit la loi BN ( $k, a p$ ). Cette approximation revient à négliger la composante aléatoire de la variance de  $X$  par rapport à la composante associée à la variabilité systématique.

Proposition 4 :

Si la variable aléatoire  $X$ , correspondant au nombre de bactéries contenues dans des échantillons de volume unitaire, est distribuée suivant la loi BN ( $k, p$ ) alors le nombre de micro-organismes contenus dans des échantillons de volume  $v$  suit la loi BN ( $k, v p$ ).

Démonstration

La densité bactérienne moyenne d'un échantillon de volume unitaire étant conforme à la loi  $\Gamma(k, p)$ , il résulte de la proposition 3 que les fluctuations de la concentration bactérienne moyenne dans des échantillons de volume  $v$  obéissent à la loi  $\Gamma(k, v p)$ . Or, le calcul de l'intégrale (10) effectué selon ces nouveaux paramètres conduit directement à l'expression générale définissant la loi BN ( $k, v p$ ).

Toutes ces considérations font apparaître le paramètre  $p$  d'une loi binomiale négative comme une grandeur liée au niveau de la concentration bactérienne moyenne, alors que le paramètre  $k$  constitue, de manière plus spécifique, une mesure intrinsèque de l'hétérogénéité indépendante du volume des échantillons analysés. Ainsi, les grandes valeurs de  $k$  traduisent une répartition des micro-organismes proche de l'homogénéité tandis que les petites valeurs de  $k$  indiquent au contraire une tendance à l'agrégation des populations bactériennes d'autant plus marquée que  $k$  est plus petit.

1.2.2. Ajustement des données expérimentales à une distribution théorique

.Test d'adéquation de  $\chi^2$

Le problème de l'ajustement d'une distribution empirique à une loi de probabilité théorique est mené de pair avec l'estimation des paramètres caractérisant la distribution théorique. Le procédé consiste, dans un premier temps, à formuler une hypothèse quant à la nature de la loi susceptible d'être suivie par les résultats expérimentaux, puis à procéder à l'exécution d'un test destiné à apprécier la vraisemblance de cette hypothèse. Il est loisible d'utiliser le test d'adéquation de  $\chi^2$  pour comparer une série de valeurs expérimentales à une loi parente telle que la loi binomiale négative. A cet effet, les  $n$  résultats expérimentaux  $r_1, r_2, \dots, r_n$  sont répartis en  $c$  classes obtenues en divisant l'étendue des données en  $c$  sous-intervalles disjoints. Soit  $O_i$  le nombre des observations contenues dans la classe de rang  $i$  et  $p_i$  une estimation de la probabilité pour la variable aléatoire de la loi parente d'appartenir à la même classe de rang  $i$  :  $O_i$  représente alors l'effectif expérimental de cette classe et  $np_i$  son effectif théorique. Dans ces conditions, si la distribution observée est compatible avec la loi théorique considérée, la statistique

$$\chi^2 = \sum_{i=1}^c \frac{(O_i - n p_i)^2}{n p_i} \text{ suit approximativement une loi de } \chi^2 \text{ à } \nu \text{ degrés de}$$

liberté avec  $\nu = c - \ell - 1$  où  $\ell$  est le nombre de paramètres estimés (i. e.,  $\ell = 2$  pour la loi binomiale négative). La validité de ce test suppose toutefois que les effectifs théoriques ne soient pas trop petits. Une règle de travail préconisée par SNEDECOR et COCHRAN (1957) précise que l'effectif théorique de deux classes peut descendre jusqu'à 1 à condition que toutes les autres dépassent 5. Afin de parvenir à satisfaire cette exigence, il est

quelquefois nécessaire de procéder au regroupement de plusieurs classes adjacentes. Le calcul des effectifs théoriques propres à chaque classe requiert, outre une estimation préalable des paramètres de la loi parente à tester, le calcul des probabilités  $p_i$  égales aux fréquences théoriques cumulées afférentes à chacune des classes considérées. Aussi, un programme permettant le calcul des fréquences théoriques cumulées de la loi BN (k,p) est donné en Annexe I. L'algorithme de ce programme, utilisable sur une calculatrice HP 41 C, est basé sur la relation de récurrence qui existe entre les probabilités de deux valeurs consécutives de r, calculées selon la formule (11).

.Estimation des paramètres k et p de la loi binomiale négative

Il existe plusieurs méthodes d'estimation des paramètres d'une loi binomiale négative à partir d'une série de résultats expérimentaux  $r_1, r_2, \dots, r_n$ . Si l'estimation "la plus vraisemblable" de la densité bactérienne moyenne  $\lambda$  est obtenue par la moyenne  $\bar{r}$  de la distribution empirique, plusieurs méthodes plus ou moins sophistiquées et précises ont par contre été proposées pour l'estimation des paramètres k et p (EL-SHAARAWI et al., 1981 ; FISHER, 1941 ; POOLE, 1974). Une première possibilité consiste à identifier les moments empiriques aux moments théoriques de la distribution binomiale négative. Ainsi, selon les relations mentionnées au paragraphe 1.2.1., les estimations  $\hat{k}$  et  $\hat{p}$  des paramètres k et p sont données respectivement par

$$\hat{k} = \frac{\bar{r}^2}{s^2 - \bar{r}} \text{ et } \hat{p} = \frac{\bar{r}}{\hat{k}} = \frac{s^2 - \bar{r}}{\bar{r}} \quad \text{où } s^2 \text{ désigne la variance des observations. Il}$$

faut toutefois souligner que cette manière relativement rapide et commode de procéder à l'estimation de k et de p peut quelquefois conduire à des valeurs assez différentes des estimations obtenues par la méthode du maximum de vraisemblance (FISHER, 1941), qui doit être préférée à la méthode des moments malgré la lourdeur des calculs qu'elle entraîne.

Le calcul de la valeur la plus vraisemblable de k, émanant d'une série  $r_1, r_2, \dots, r_n$  de n réalisations indépendantes d'une variable distribuée suivant une loi BN (k,p) de paramètres inconnus, repose sur la fonction de vraisemblance associée aux résultats expérimentaux, soit

$$P(r_1, \dots, r_n/k,p) = \prod_{i=1}^n \frac{(k+r_i - 1)! p^{r_i}}{r_i! (k-1)! (1+p)^{k+r_i}} \quad (12)$$

La méthode du maximum de vraisemblance consiste à prendre la valeur  $\hat{k}$ , qui maximise la fonction de vraisemblance, comme estimation de la vraie valeur de k. Cette estimation est par définition la "valeur la plus vraisemblable" du paramètre k.

Le maximum de la fonction de vraisemblance s'obtient en cherchant le maximum de son logarithme L (k,p), soit

$$L(k,p) = \sum_{i=1}^n [\ln (k+r_i - 1)! + r_i \ln p - \ln r_i! - \ln (k-1)! - (k+r_i) \ln (1+p)] \quad (13)$$

En remplaçant p par son estimation  $\hat{p} = \frac{\bar{r}}{k}$ , obtenue par la méthode du maximum de vraisemblance, dans la formule (13) et en dérivant par rapport à k,

il vient :

$$L'(k) = \sum_{i=1}^n \left[ \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{k+r_i-1} - \frac{r_i}{k} - \ln \left( 1 + \frac{\bar{r}}{k} \right) + \frac{(k+r_i) \bar{r}}{k(k+r)} \right] \quad (14)$$

Finalement, après simplification, la valeur cherchée est solution de l'équation

$$n \ln \left( 1 + \frac{\bar{r}}{k} \right) - \sum_{i=1}^n \sum_{t=1}^{r_i} \frac{1}{k+r_i-t} = 0 \quad (15)$$

La solution,  $\hat{k}$ , de cette équation est calculée par approximations successives suivant un procédé itératif (e. g., la méthode de Newton) à l'aide d'un calculateur électronique. Un exemple de programme de résolution de l'équation (15), écrit en langage BASIC, est proposé en Annexe II.

#### Estimation des valeurs de k et de p communes à plusieurs séries de résultats expérimentaux

Lorsque plusieurs séries de données, décrites par des lois binomiales négatives, sont disponibles, il est parfois utile (e.g., pour apprécier de manière globale le degré d'hétérogénéité ou encore pour effectuer des tests de comparaison, cf. paragraphe 1.2.3. ci-après) de procéder à l'estimation d'une valeur unique de k, ou de p, qui est alors supposée être commune à toutes ces distributions.

Soient  $(r_{i1}, r_{i2}, \dots, r_{in_i})$  les  $n_i$  observations indépendantes de la  $i^{\text{ème}}$  série,  $i = 1, 2, \dots, \ell$ , dont les valeurs obéissent à la loi BN  $(k_i, p_i)$ . En admettant l'égalité de tous les  $k_i$ , pour  $i = 1, 2, \dots, \ell$ , la fonction de vraisemblance de la valeur commune de k associée aux  $\ell$  distributions empiriques est donnée par la formule

$$P(r_{ij}, 1 \leq i \leq \ell, 1 \leq j \leq n_i/k, p_i) = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{(k+r_{ij}-1)! p_i^{r_{ij}}}{r_{ij}! (k-1)! (1+p_i)^{k+r_{ij}}} \quad (16)$$

En suivant le raisonnement déjà utilisé ci-dessus dans le cas d'une seule série expérimentale, il est aisé d'établir que la valeur la plus vraisemblable du paramètre k commun à chacune des  $\ell$  distributions est la solution de l'équation

$$\sum_{i=1}^{\ell} n_i \ln \left( 1 + \frac{\bar{r}_i}{k} \right) - \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \sum_{t=1}^{r_{ij}} \frac{1}{k+r_{ij}-t} = 0 \quad (17)$$

où  $\bar{r}_i$  représente la moyenne arithmétique des observations de la série de rang  $i$ .

Cette équation, qui constitue une généralisation de l'équation (15) peut également être résolue par la méthode de Newton. Un exemple de programme de résolution de l'équation (17) est donné en Annexe III.

Une démarche analogue peut s'appliquer à l'estimation du paramètre  $p$ , commun à plusieurs distributions binomiales négatives. En conservant les notations précédentes et en faisant l'hypothèse d'égalité des paramètres  $p_i$  pour chacune des lois BN ( $k_i, p_i$ ), la fonction de vraisemblance de la valeur commune de  $p$  devient :

$$P(r_{ij}, 1 \leq i \leq l, 1 \leq j \leq n_i/k_i, p) = \prod_{i=1}^l \prod_{j=1}^{n_i} \frac{(k_i+r_{ij}-1)! p^{r_{ij}}}{r_{ij}! (k_i-1)! (1+p)^{k_i+r_{ij}}} \quad (18)$$

D'où en annulant la dérivée, par rapport à  $p$ , du logarithme de cette expression

$$\sum_{i=1}^l \sum_{j=1}^{n_i} \left[ \frac{r_{ij}}{p} - (k_i+r_{ij}) \frac{1}{1+p} \right] = 0 \quad (19)$$

il vient immédiatement

$$\hat{p} = \frac{\sum_{i=1}^l n_i \bar{r}_i}{\sum_{i=1}^l n_i k_i} \quad (20)$$

Finalement en remplaçant  $k_i$  par  $\frac{\bar{r}_i}{\hat{p}_i}$  et  $n_i \bar{r}_i$  par  $T_i$ , il découle de (20) que

$$\hat{p} = \frac{\sum_{i=1}^l T_i}{\sum_{i=1}^l T_i / \hat{p}_i} \quad (21)$$

où  $\hat{p}_i$  représente la valeur la plus vraisemblable du paramètre  $p$  caractérisant la  $i^{\text{ème}}$  distribution.

### 1.2.3. Tests de comparaison

Soient  $l$  populations représentées par les séries expérimentales  $(r_{i1}, r_{i2}, \dots, r_{ini})$  d'effectifs  $n_i$  et dont les lois parentes sont les distributions BN ( $k_i, p_i$ ) avec  $i = 1, 2, \dots, l$ . La comparaison des  $l$  valeurs des paramètres  $k$  ou  $p$  propres à chacune de ces populations peut être effectuée au moyen du test du rapport de vraisemblance (en anglais : likelihood ratio test) (KENDALL et STUART, 1973 ; LINDGREN, 1968 ; RAO, 1973).

1.2.3.1. Comparaison de plusieurs estimations du paramètre k

Le contrôle de l'hypothèse nulle,  $H_0$ , stipulant l'égalité des paramètres  $k_i$  pour l'ensemble des populations considérées, repose sur la statistique  $\Lambda = LH_0/LH_1$ .

Le numérateur de ce rapport,  $LH_0$ , représente la fonction de vraisemblance, selon la formule (16), faisant état de l'égalité de tous les  $k_i$ , pour  $i = 1, 2, \dots, \ell$ , et dans laquelle  $k$  est remplacé par l'estimation  $\hat{k}$  obtenue en résolvant l'équation (17). Le dénominateur,  $LH_1$ , de  $\Lambda$  est formé du produit des  $\ell$  fonctions de vraisemblance associées à chacune des  $\ell$  séries expérimentales dans lesquelles  $k$  est remplacé par la solution,  $\hat{k}_i$ , de l'équation (15) se rapportant à chacune de ces séries.

Ainsi, de manière plus explicite

$$LH_0 = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{(\hat{k} + r_{ij} - 1)! \hat{p}_i^{r_{ij}}}{r_{ij}! (\hat{k} - 1)! (1 + \hat{p}_i)^{\hat{k} + r_{ij}}} \quad (22)$$

et

$$LH_1 = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{(\hat{k}_i + r_{ij} - 1)! \hat{p}_i^{r_{ij}}}{r_{ij}! (\hat{k}_i - 1)! (1 + \hat{p}_i)^{\hat{k}_i + r_{ij}}} \quad (23)$$

De (22) et (23) il vient

$$\Lambda = \frac{LH_0}{LH_1} = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{\hat{k} (\hat{k} + 1) \dots (\hat{k} + r_{ij} - 1) (1 + \hat{p}_i)^{\hat{k}_i - \hat{k}}}{\hat{k}_i (\hat{k}_i + 1) \dots (\hat{k}_i + r_{ij} - 1)} \quad (24)$$

Quand tous les  $k_i$  sont égaux, la quantité  $-2 \ln \Lambda$  suit la loi de  $\chi^2$  avec  $\ell - 1$  degrés de liberté. L'exécution du test nécessite par conséquent le calcul préalable de la fonction discriminante  $-2 \ln \Lambda$ , soit

$$-2 \ln \Lambda = 2 \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \left[ \ln \frac{\hat{k}_i (\hat{k}_i + 1) \dots (\hat{k}_i + r_{ij} - 1)}{\hat{k} (\hat{k} + 1) \dots (\hat{k} + r_{ij} - 1)} + (\hat{k} - \hat{k}_i) \ln \left( 1 + \frac{\bar{r}_i}{\hat{k}_i} \right) \right] \quad (25)$$

ou encore

$$-2 \ln \Lambda = 2 \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \sum_{t=1}^{r_{ij}} \ln (\hat{k}_i + r_{ij} - t) - 2 \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \sum_{t=1}^{r_{ij}} \ln (\hat{k} + r_{ij} - t) + 2 \sum_{i=1}^{\ell} n_i (\hat{k} - \hat{k}_i) \ln \left( 1 + \frac{\bar{r}_i}{\hat{k}_i} \right) \quad (26)$$

Cette dernière expression permet d'effectuer la comparaison de plusieurs valeurs du paramètre  $k$  en situant la statistique  $- 2 \ln \Lambda$  dans la distribution de  $\chi^2$  à  $\ell - 1$  degrés de liberté.

1.2.3.2. Comparaison de plusieurs estimations du paramètre  $p$

Le test du rapport de vraisemblance peut également être mis en oeuvre pour effectuer la comparaison des valeurs du paramètre  $p$  caractérisant plusieurs distributions binomiales négatives BN ( $k_i, p_i$ ),  $i = 1, 2, \dots, \ell$ , en utilisant les résultats de séries expérimentales issues de ces populations. En conservant la démarche ainsi que les notations du paragraphe précédent, la statistique  $\Lambda$ , sur laquelle s'appuie le test de comparaison, est formée du rapport  $LH_0/LH_1$  avec

$$LH_0 = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{(k_i + r_{ij} - 1)! \hat{p}^{r_{ij}}}{r_{ij}! (k_i - 1)! (1 + \hat{p})^{k_i + r_{ij}}} \quad (27)$$

et

$$LH_1 = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \frac{(k_i + r_{ij} - 1)! \hat{p}_i^{r_{ij}}}{r_{ij}! (k_i - 1)! (1 + \hat{p}_i)^{k_i + r_{ij}}} \quad (28)$$

Dans ces formules  $\hat{p}_i$  est égal au quotient  $\frac{\bar{r}_i}{\hat{k}_i}$  alors que  $\hat{p}$  est donné par la relation (21).

Ainsi,

$$\Lambda = \prod_{i=1}^{\ell} \prod_{j=1}^{n_i} \left( \frac{\hat{p}}{\hat{p}_i} \right)^{r_{ij}} \left( \frac{1 + \hat{p}_i}{1 + \hat{p}} \right)^{k_i + r_{ij}} \quad (29)$$

par suite,

$$- 2 \ln \Lambda = 2 \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \left[ r_{ij} \ln \frac{\hat{p}_i}{\hat{p}} + (k_i + r_{ij}) \ln \frac{1 + \hat{p}}{1 + \hat{p}_i} \right] \quad (30)$$

En remplaçant  $k_i$  par  $\frac{\bar{r}_i}{\hat{p}_i}$  et  $n_i \bar{r}_i$  par  $T_i$ , (30) devient

$$- 2 \ln \Lambda = 2 \sum_{i=1}^{\ell} T_i \left[ \ln \frac{\hat{p}_i}{\hat{p}} + \frac{1 + \hat{p}_i}{\hat{p}_i} \ln \frac{1 + \hat{p}}{1 + \hat{p}_i} \right] \quad (31)$$

Le test de comparaison des  $p_i$  s'effectue en comparant la quantité  $-2\ln \Lambda$  donnée par la relation (31) à la distribution de  $\chi^2$  avec  $\ell - 1$  degrés de liberté.

### 1.3. RECHERCHE D'UNE STRUCTURE D'ABONDANCE DU PEUPEMENT BACTERIEN

L'ajustement d'un modèle mathématique à la distribution des résultats des analyses bactériologiques permet d'appréhender l'hétérogénéité de manière formelle et quantitative mais ne donne aucune information sur la configuration et l'agencement des zones d'abondance du peuplement bactérien. C'est pourquoi cette première étape est quelquefois suivie par la recherche systématique d'une structure de discontinuité de l'environnement étudié qui se présente cependant le plus souvent sous la forme d'un continuum. Dans le cas d'une hétérogénéité structurée, l'identification de la structure d'abondance constitue une façon de connaître et de mieux comprendre le milieu tout en permettant d'envisager les campagnes d'échantillonnage ultérieures avec la perspective de tirer profit de la structure dégagée. Ainsi, l'étude de la structure géographique et temporelle de la répartition des micro-organismes dans un réseau de distribution publique doit faciliter le déroulement et la conception des investigations sur les relations entre les titres bactériens observés et différentes variables explicatives potentielles (température, temps de rétention, chlore résiduel, ...) en s'appuyant sur des bases plus solides. En particulier, si la cartographie de la concentration bactérienne fait apparaître une répartition structurée des zones d'abondance sur le réseau, il sera possible d'une part, de mieux comprendre les problèmes de contamination et d'autre part, de définir une stratégie d'échantillonnage adaptée au contrôle de la conformité de la qualité bactériologique d'une eau de distribution par rapport aux normes de salubrité en vigueur.

Dans la suite, nous envisagerons deux approches, permettant la réalisation de la stratification géographique d'un réseau de distribution à partir des résultats bactériologiques obtenus au cours d'une ou de plusieurs campagnes d'échantillonnage systématique, à savoir :

- a) une méthode de groupement basée sur la distance euclidienne entre objets (i. e., les stations de prélèvement),
- b) une méthode factorielle (i. e., l'analyse en composantes principales) destinée à visualiser les différents objets dans un espace de dimension réduite.

#### 1.3.1. La méthode de groupement

Le groupement est une opération de l'analyse multidimensionnelle qui consiste à partitionner une collection d'objets, en l'occurrence il s'agit de stations de prélèvement, en sous-collections mutuellement exclusives et collectivement exhaustives. Les points de prélèvement appartenant aux différents groupes obtenus selon un procédé objectif de classification, délimitent des parties du réseau laissant ainsi apparaître une image concrète de la structure spatiale de l'hétérogénéité bactérienne. Ces parties formées de stations contiguës ou non sont appelées zones de densité bactérienne ou plus simplement zones. Notons qu'il est recommandé de répartir les points de

prélèvement de manière systématique (i. e., points régulièrement espacés) sur toute l'étendue du réseau étudié et de conserver les mêmes points au cours des différentes campagnes d'échantillonnage.

Lorsque la méthode de groupement est effectuée sur les résultats d'une seule campagne, le caractère instantané des zones de densité bactérienne constitue une véritable photographie du réseau au moment de l'échantillonnage. Cependant, l'aspect trop ponctuel d'une telle approche de même que la mouvance des milieux hydriques montrent qu'il est préférable d'effectuer le groupement en considérant simultanément les résultats de plusieurs campagnes de prélèvements échelonnés par exemple sur une période de quelques mois. L'intégration de données collectées au cours de plusieurs campagnes réalisées dans le temps doit en effet augmenter la régularité du contour des zones en atténuant les phénomènes de contamination éphémères et accidentels. De plus, la configuration spatiale de l'hétérogénéité obtenue sur une longue période traduit des phénomènes durables dont la signification et l'interprétation sont plus aisées.

La méthode de groupement proposée est une méthode séquentielle, non-hiérarchique qui affecte chaque objet (i. e., station de prélèvement) au groupe dont le centroïde est le plus proche (ANDERSON et al., 1984). Le centroïde d'un groupe d'objets peut être imaginé comme l'objet-type du groupe, que cet objet existe effectivement dans le groupe ou qu'il ne soit qu'une création mathématique. Dans un espace multidimensionnel, le centroïde d'un groupe d'objets est le point dont les coordonnées sont les coordonnées moyennes des objets du groupe.

La méthode procède selon les étapes suivantes

#### 1) Présentation des données

Les données sont consignées dans une matrice  $X = \{x_{ij}\}$  où  $x_{ij}$  désigne le titre bactérien observé pour la  $i^{\text{ème}}$  station de prélèvement ( $i = 1, 2, \dots, n$ ) lors de la  $j^{\text{ème}}$  campagne d'échantillonnage ( $j = 1, 2, \dots, m$ ). Les différents points de prélèvement représentés par les vecteurs  $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$  pour  $i = 1, 2, \dots, n$ , sont assignés à des groupes d'objets suivant l'algorithme décrit ci-dessous.

#### 2) Transformation des données

Toutes les données sont soumises à la transformation "racine carrée" afin d'homogénéiser la variance associée à chaque résultat. Cette transformation préalable trouve sa justification dans le modèle de Poisson qui est habituellement utilisé pour décrire les titres bactériens obtenus au sein d'un faible volume d'eau. Il faut toutefois noter que la validité de ce modèle est mise en défaut par les coefficients multiplicateurs imputables aux différents niveaux de dilution utilisés pour effectuer certains titrages et parvenir à l'expression finale du titre bactérien. Ainsi, dans la matrice des données,  $x_{ij}$  est remplacé par  $x'_{ij} = \sqrt{x_{ij}}$ .

Une seconde opération destinée à équilibrer la contribution relative des différentes campagnes de prélèvements au groupement consiste à réduire les données précédentes en divisant chaque observation  $x'_{ij}$  par l'écart-type caractérisant les valeurs de la  $j^{\text{ème}}$  campagne.

Chaque élément  $x_{ij}$  de la matrice initiale est finalement remplacé

par  $x''_{ij}$  tel que

$$x''_{ij} = \frac{x'_{ij}}{s'_j}$$

$$\text{où } s'_j{}^2 = \frac{1}{n-1} \sum_{i=1}^n (x'_{ij} - \bar{x}'_{.j})^2 \quad \text{et} \quad \bar{x}'_{.j} = \frac{1}{n} \sum_{i=1}^n x'_{ij}$$

Dans la suite, les  $x''_{ij}$  seront notés plus simplement  $x_{ij}$  afin d'alléger l'écriture.

### 3) Détermination des centroïdes initiaux

Le nombre des groupes à former étant fixé égal à  $q$ , la méthode de groupement requiert la définition préalable de centroïdes initiaux servant à caractériser chaque groupe pour permettre d'engager le processus itératif. Les centroïdes désignés par  $z_{\ell} = (z_{\ell 1}, z_{\ell 2}, \dots, z_{\ell m})$ , avec  $\ell = 1, 2, \dots, q$ , sont des points de l'espace à  $m$  dimensions dont les nouvelles coordonnées sont calculées après chaque itération. Les coordonnées des  $q$  centroïdes initiaux peuvent, par exemple, être les centres des  $q$  intervalles d'égale amplitude divisant l'étendue des valeurs  $x_{ij}$  pour chaque  $j$ .

### 4) Affectation d'une station de prélèvement à un groupe

La distance euclidienne  $D_{i\ell}$  dans l'espace à  $m$  dimensions entre la station  $i$  et le centroïde du  $\ell^{\text{ème}}$  groupe est calculée selon la formule

$$D_{i\ell}^2 = \sum_{j=1}^m (x_{ij} - z_{\ell j})^2 \quad (32)$$

La station  $i$  est affectée au groupe pour lequel la distance  $D_{i\ell}$  est minimale.

### 5) Trajectoire des centroïdes

Lorsque tous les objets ont été affectés à l'un des  $q$  groupes en formation, les nouvelles coordonnées des centroïdes des groupes ainsi obtenus sont calculées comme suit

$$z_{\ell j} = \frac{1}{n_{\ell}} \sum_{x_{ij} \in w_{\ell}} x_{ij} \quad (33)$$

où  $w_{\ell}$  et  $n_{\ell}$  représentent respectivement l'ensemble des stations et le nombre des stations composant le groupe  $\ell$ . Les étapes 4 et 5 sont réitérées avec les nouveaux centroïdes dont les trajectoires tendent progressivement vers un point limite.

### 6) Test d'arrêt

Le procédé itératif est interrompu et la configuration finale est atteinte dès que la composition des groupes reste inchangée au cours de deux itérations consécutives.

. Détermination du nombre des groupes :

L'algorithme de division exposé ci-dessus permet de faire une partition du groupe initial contenant tous les objets en un ensemble de groupes dont le nombre  $q$  est compris entre 2 et  $n$ . Partant de la valeur  $q = 2$  les étapes 3, 4, 5 et 6 sont exécutées en augmentant  $q$  d'une unité à chaque nouvelle étape. L'interruption du programme est décidée dès que la fraction de la variance expliquée par le groupement atteint un niveau suffisant ou lorsque la formation d'un groupe supplémentaire n'augmente pas le rapport de la variance expliquée sur la variance totale dans des proportions jugées satisfaisantes. Ceci nous amène à décrire le mode de calcul du rapport précédent, lequel est exécuté à l'issue de l'étape 6.

. Calcul de la fraction de la variance totale expliquée par le groupement :

Soit  $C$  la matrice d'ordre  $(n \times m)$  des données centrées dans laquelle  $x_{ij} - \bar{x}_{.j}$  représente le terme de la  $i^{\text{ème}}$  ligne et de la  $j^{\text{ème}}$  colonne avec

$$\bar{x}_{.j} = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad .$$

La matrice de dispersion  $S_T$  caractérisant l'ensemble des données est définie par le produit matriciel  $S_T = C'C$ , où  $C'$  est la transposée de  $C$ . Il est facile de montrer que la trace de  $S_T$ , notée  $\text{tr}(S_T)$ , représente la variance totale des données. Ainsi,

$$\text{tr}(S_T) = \sum_{j=1}^m \sum_{i=1}^n (x_{ij} - \bar{x}_{.j})^2 \quad (34)$$

Un calcul similaire effectué au sein de chaque groupe permet d'exprimer la variabilité intragroupe.

Si  $C_\ell$  désigne la matrice d'ordre  $(n_\ell \times m)$  des données centrées appartenant au groupe  $\ell$ , le terme général de  $C_\ell$  est donné par  $x_{ij} - z_{\ell j}$ . La matrice de dispersion  $S_\ell$  correspondant à ce groupe, est alors définie par le produit  $S_\ell = C'_\ell C_\ell$ .

De même, la trace de  $S_\ell$  dont l'expression explicite est donnée par

$$\text{tr}(S_\ell) = \sum_{j=1}^m \sum_{x_{ij} \in \mathcal{G}_\ell} (x_{ij} - z_{\ell j})^2 \quad (35)$$

représente la variance intragroupe associée au groupe  $\ell$ .

Finalement, la fraction  $R$  de la variance totale, qui est expliquée par le groupement, est obtenue par la formule

$$R = \frac{\text{tr}(S_T) - \sum_{\ell=1}^q \text{tr}(S_\ell)}{\text{tr}(S_T)} \quad (36)$$

dans le cas où  $q$  groupes ont été définis.

En permettant d'apprécier l'efficacité du groupement pour les valeurs successives de q, la valeur de R, qui est généralement exprimée par un pourcentage, doit prévaloir sur l'opportunité de la formation de groupes supplémentaires. Notons enfin que la lourdeur des calculs caractérisant la méthode de groupement proposée ne doit pas constituer un obstacle majeur à la mise en oeuvre de cette technique. A titre d'exemple, la durée d'exécution du programme, rédigé en langage APL, utilisé pour le traitement des 600 données bactériologiques environ du réseau messin, n'a pas excédé quelques minutes avec le calculateur CDC 171.

### 1.3.2. L'ordination en espace réduit

Sans entrer dans les développements théoriques de l'analyse en composantes principales, dont l'étude exhaustive est présentée dans plusieurs ouvrages (KENDALL et STUART, 1968 ; LEFEBVRE, 1976 ; LEGENDRE et LEGENDRE, 1979), il convient néanmoins de rappeler très sommairement certains aspects de cette technique très puissante, parfois utilisée en tant que méthode de groupement pour la recherche d'une structure dans la distribution des données expérimentales.

Une manière d'aborder le problème multidimensionnel visant à dégager les grandes tendances de la variabilité entre objets (i. e., stations de prélèvement) observés pour un ensemble de descripteurs (i. e., campagnes de prélèvements), consisterait à vouloir représenter la dispersion des objets dans un espace multidimensionnel ayant une dimension égale au nombre de descripteurs de l'étude. Lorsque ce nombre est supérieur à trois, voire même déjà deux, les difficultés pratiques liées à la représentation d'un tel diagramme peuvent être éludées par la projection du diagramme multidimensionnel dans un espace de dimension réduite (e. g., un plan) dont les axes traduiraient au mieux l'ensemble de la variabilité de la matrice multidimensionnelle des données.

L'analyse en composantes principales, telle qu'elle a été appliquée dans ce travail, se déroule conformément à la description donnée ci-après.

- Partant de la matrice des données,  $X = \{x_{ij}\}$ , les données brutes sont soumises à la transformation "racine carrée" avant d'être centrées et réduites. Le but de cette standardisation est d'harmoniser la contribution des descripteurs au rapprochement des objets dans l'espace, et ce indépendamment de leur variance propre. Ainsi, en reprenant les notations définies à la section 1.3.1., la suite des opérations est effectuée sur les valeurs

$$x''_{ij} = \frac{x'_{ij} - \bar{x}'_{.j}}{s'_{.j}}$$

- L'étape suivante consiste à calculer la matrice de covariance des descripteurs centrés réduits qui est également la matrice de corrélation des descripteurs après la seule transformation "racine carrée".

- On procède alors à l'extraction des valeurs propres de la matrice d'association. Le nombre et la somme de ces valeurs propres sont tous deux égaux au nombre des descripteurs.

- Les vecteurs propres associés aux valeurs propres ainsi obtenues définissent les axes principaux de la matrice de corrélation. Ces axes correspondent à des fractions de plus en plus faibles de la variance totale.

- La normalisation des vecteurs propres à l'unité, qui entraîne, simultanément, une normalisation des axes-descripteurs, permet le calcul des composantes principales, formées des coordonnées des objets par rapport aux différents axes principaux.

Les axes principaux sont orthogonaux entre eux, ce qui signifie qu'ils correspondent à des directions indépendantes les unes des autres dans l'ellipsoïde de probabilité de la distribution des objets. De plus, les valeurs propres de la matrice de corrélation mesurent la quantité de variance expliquée par les axes principaux auxquels elles se rapportent de manière biunivoque.

Ces propriétés font de l'analyse en composantes principales une méthode capable de résumer, en quelques dimensions importantes, la plus grande partie de la variabilité d'une matrice d'association comprenant quelquefois un grand nombre de descripteurs et de connaître aussi la quantité de variance expliquée par ces quelques axes principaux indépendants (LEGENDRE et LEGENDRE, 1979).

- Les composantes principales permettent enfin de disposer les objets dans un espace de dimension réduite. Ainsi, par exemple, les deux premiers axes principaux forment le plan de plus grande variance pour une projection des objets qui respecte au mieux leurs distances relatives dans l'espace multidimensionnel.

En résumé, l'analyse en composantes principales, qui n'est en fait qu'une rotation dans l'espace multidimensionnel du système d'axes initial, permet :

- 1) de dégager les relations entre objets et éventuellement de procéder à leur groupement par la représentation de ces objets en espace réduit,
- 2) de connaître le rôle des descripteurs d'origine dans la formation des axes principaux,
- 3) de représenter en espace réduit les relations entre ces descripteurs.

En particulier, la contribution de chaque descripteur à la formation de l'espace réduit est évaluée au moyen du cercle de contribution équilibrée des descripteurs (LEGENDRE et LEGENDRE, 1979) ; et la relation de proportionnalité, existant entre la projection d'un axe-descripteur sur un axe principal et la corrélation de ce descripteur avec l'axe principal, permet de comparer la projection des différents axes-descripteurs sur un même axe principal.

## CHAPITRE II

### CHOIX D'UNE METHODE DE TITRAGE DES BACTERIES

Les résultats de toute étude quantitative en bactériologie des eaux sont partiellement influencés par le mode de prélèvement des échantillons, les conditions de leur conservation (BOUSQUET et DEVOUCOUX, 1982) et surtout par la méthode de quantification mise en oeuvre pour dénombrer les populations bactériennes. Selon la méthode de calcul utilisée pour l'expression de la densité bactérienne, les méthodes de titrage peuvent être classées en deux groupes, à savoir :

- les méthodes de type énumératif qui consistent à dénombrer directement les micro-organismes au microscope (microscopie en épifluorescence), ou les colonies apparues après culture sur un milieu approprié (technique des membranes filtrantes, incorporation ou étalement sur milieu gélosé).

- les méthodes de type quantique pour lesquelles l'estimation est obtenue généralement après incubation, à partir des rapports entre le nombre de réponses positives (présence d'une multiplication bactérienne) et le nombre total de milieux inoculés à un ou plusieurs niveaux de dilution. La méthode de Fisher (WYSHAK et DETRE, 1972) et surtout la méthode communément appelée méthode du nombre le plus probable (N.P.P.) se rangent dans cette deuxième catégorie. L'habitude veut qu'à une méthode de dénombrement et plus particulièrement à un mode d'ensemencement (i. e., milieu gélosé vs milieu liquide) soit associée spécifiquement une méthode de calcul des titres bactériens. Or, ce point de vue ne trouve aucun fondement étant donné que chaque mode d'ensemencement permet d'envisager plusieurs méthodes pour l'exploitation des résultats (BEYTOUT et al., 1975 ; MAUL et al., 1981 ; TIPPETT, 1932 ; WYSHAK et DETRE, 1972).

Le choix d'une technique de titrage de même que celui d'une méthode de calcul des titres bactériens est généralement guidé par des critères de justesse, de précision, de reproductibilité et de commodité d'utilisation sans toutefois oublier les méthodes de traitement ultérieur des données recueillies. A cet égard, il convient de souligner que les disparités dans les résultats, consécutives au manque de spécificité absolue caractérisant l'ensemble des méthodes de quantification, sont imputables aux propriétés sélectives des différentes méthodes. En effet, les variations des fréquences relatives des divers groupes d'organismes dans un biotope donné selon les disponibilités nutritionnelles du milieu (BERLAND et al., 1976 ; GUERRERO et al. ; 1975) sont amplifiées au niveau des résultats par la sélectivité préférentielle qu'exercent les méthodes de titrage sur les différentes espèces bactériennes. Ce phénomène se traduit par une interaction entre la méthode de titrage et l'échantillon analysé (BLOCK et al., 1982). La complexité du problème, inhérente à la pluralité des variables explicatives potentielles sur le résultat d'un titrage bactérien, explique sans doute la pléthore de travaux consacrés aux aspects qualitatifs et quantitatifs des techniques de numération bactérienne. A défaut de pouvoir parvenir à l'avènement d'une "méthode universelle", les microbiologistes ont souvent tenté d'appréhender l'incidence de certains paramètres physico-chimiques sur le "rendement" des techniques de numération. La plupart de ces études comparatives sont consacrées à la colimétrie (BLOCK et al., 1982 ; CADA, 1975 ; GREEN et LITSKY, 1980 ; MAUL et BLOCK, 1983 ; McCARTHY et al., 1961 ; THOMAS et WOODWARD, 1955 ; TOBIN et al., 1980 ; VELZ, 1970). Cependant le manque de rigueur mathématique dans l'exploitation statistique des résultats est quelquefois aggravé par une mauvaise définition du plan expérimental. Cette remarque s'applique notamment à une étude de MEANS et al., (1981) faisant état des rendements comparatifs de plusieurs milieux de culture pour l'énumération des bactéries hétérotrophes d'un réseau de distribution publique. Le manque de rigueur du protocole expérimental, en particulier pour ce qui concerne l'agencement des températures et des temps d'incubation au regard des milieux étudiés, nous amène à considérer certaines conclusions des auteurs comme étant hasardeuses et inopportunes.

Par ailleurs, la comparaison statistique des moyennes de plusieurs séries de résultats expérimentaux, obtenues par exemple pour différentes méthodes de titrage réalisées dans des conditions expérimentales déterminées, est souvent effectuée par une analyse de la variance. Cet outil remarquable, faisant partie de l'arsenal puissant des statistiques paramétriques, a souvent été utilisé à bon escient par les microbiologistes. Cependant, les principales contraintes liées à l'application de cette technique, à savoir : la normalité des données et l'homogénéité des variances associées aux différents traitements, nécessitent quelquefois une transformation préalable des données brutes dont la nature est dictée par la relation existant entre la variance et la moyenne de la population considérée (GREEN, 1979 ; SNEDECOR et COCHRAN, 1957). Cette double exigence, qu'il est parfois difficile de satisfaire par une transformation unique, s'oppose à toute utilisation abusive et systématique de l'analyse de variance dont la robustesse autorise toutefois quelques écarts entre la réalité et les fondements théoriques.

Le choix impromptu d'une méthode appropriée au titrage des bactéries hétérotrophes n'étant pas souhaitable sur les seules références bibliographiques dont nous avons connaissance, il nous est apparu nécessaire de reconsidérer le problème en mettant à profit les avantages du test du rapport de vraisemblance pour une exploitation plus rigoureuse des résultats. Ce test qui ne requiert aucune transformation des données brutes, repose exclusivement sur la loi de probabilité associée aux résultats de chacune des techniques employées pour l'estimation de la concentration bactérienne.

Une première étape sera consacrée à la comparaison expérimentale de trois techniques de quantification des bactéries hétérotrophes. En l'occurrence, il s'agit : de la méthode par inclusion en gélose (IG), actuellement considérée comme la technique de référence, de la méthode des membranes filtrantes (MF) et enfin d'une méthode de type quantique en milieu liquide basée sur le calcul du nombre le plus probable (N.P.P.) en utilisant 96 ensemencements par niveau de dilution. Cette dernière technique qui s'inspire de la microméthode proposée par PROST et HUGUES (1982) mérite considération eu égard aux résultats particulièrement intéressants qu'elle a donnés en colimétrie des eaux de rivière (BLOCK et al., 1982 ; MAUL et BLOCK, 1983).

Dans une deuxième étape, nous tenterons d'évaluer le rendement relatif, en matière de croissance bactérienne, de la gélose nutritive (GN) comparativement à deux autres milieux de culture dont la composition a été publiée récemment : le milieu R-2A (REASONER et GELDREICH, 1985) et le milieu m-SPC (TAYLOR et GELDREICH, 1979).

## 2.1. COMPARAISON EXPERIMENTALE DE TROIS METHODES DE TITRAGE :

### METHODES ENUMERATIVES PAR FILTRATION SUR MEMBRANE OU PAR INCLUSION EN GELOSE ET METHODE QUANTIQUE EN MILIEU LIQUIDE

L'objectif de cette première expérimentation est d'étudier l'amplitude des écarts entre les titres bactériens obtenus par trois méthodes de titrage distinctes afin de savoir si les résultats de celles-ci peuvent être considérés comme équivalents d'un point de vue quantitatif. La commodité de mise en oeuvre de même que la facilité de lecture et d'interprétation des résultats seront également prises en considération pour le choix d'une méthode de quantification des bactéries hétérotrophes.

### 2.1.1. Matériel et méthodes

#### 2.1.1.1. Echantillons d'eau

Des échantillons de 500 ml d'eau ont été prélevés, dans des sections du réseau de distribution publique de la ville de Metz alimentées par des unités de production différentes, au moyen de flacons en verre stérilisés à l'autoclave, et contenant 1 ml d'une solution de thiosulfate de sodium à 3%. Ces échantillons, qui ont été recueillis chez l'habitant à des robinets préalablement flambés et après avoir laissé couler l'eau pendant 30 secondes environ à la pression maximale, ont ensuite été ramenés au laboratoire à température ambiante pour être soumis à l'analyse bactériologique au plus 1 heure après le prélèvement.

#### 2.1.1.2. Méthode des membranes filtrantes

Des dilutions décimales des échantillons d'eau ont été réalisées dans de l'eau physiologique stérile contenant 0,8 % de NaCl. Des volumes de l'échantillon originel variant de 100 ml à 0,01 ml ont été filtrés sur membrane stérile d'une porosité de 0,45  $\mu$ m (Millipore HAWG 047). Pour permettre une meilleure répartition de l'aliquote sur la membrane, celle-ci a, le cas échéant, été préalablement transférée dans 9 ml d'eau physiologique. Chaque membrane a ensuite été déposée sur milieu gélosé (tryptone : 5,0 g ; extrait de levure : 2,5 g ; glucose : 1,0 g ; gélose : 15,0 g ; eau distillée : 1000 ml).

Après incubation à 20°C pendant 72 heures, toutes les colonies apparues sur la membrane ont été dénombrées. Enfin, les résultats ont été exprimés en nombre de bactéries hétérotrophes par millilitre de l'échantillon originel calculé selon le principe du maximum de vraisemblance. Ce mode de calcul consiste à diviser le nombre total des colonies recensées par le volume total de l'échantillon originel analysé toutes dilutions confondues (MAUL et al., 1981).

#### 2.1.1.3. Méthode par inclusion en gélose

A partir des dilutions obtenues conformément au paragraphe précédent, des volumes d'échantillon originel de 1 ml à 0,01 ml ont été incorporés à la gélose maintenue en surfusion. La composition du milieu de culture ainsi que les conditions d'incubation et l'expression des résultats sont en tout point restées semblables à la description donnée au paragraphe 2.1.1.2.

#### 2.1.1.4. Méthode quantique en milieu liquide

Les dilutions de la suspension bactérienne à titrer mentionnées au paragraphe 2.1.1.2. ont été réparties à l'aide de pipettes stériles dans les 96 cupules des plaques de microtitration Microtest II (Polylabo BLOCK FALCON 19013 F 3041) à raison de 0,067 ml par cupule. Auparavant, les plaques avaient été stérilisées par exposition sous rayons ultraviolets pendant 18 heures.

Une seringue de Cornwall a ensuite permis de répartir un milieu de culture liquide stérile (tryptone : 5,0 g ; extrait de levure : 2,5 g ; glucose : 1,0 g ; eau distillée : 1000 ml) sur les plaques Microtest II (0,2 ml de milieu par cupule). Puis les plaques ont été fermées au moyen d'un couvercle ou d'un film autocollant.

Après incubation à 20°C pendant 72 heures, les cupules présentant une turbidité due à la multiplication bactérienne ont été dénombrées.

Le nombre le plus probable de bactéries contenues dans un inoculum de la suspension originelle a été calculé à partir d'un triplet obtenu par la juxtaposition des comptages positifs émanant de trois niveaux de dilution successifs. Les calculs ont été effectués par une calculatrice programmable HEWLETT - PACKARD HP 41C en suivant le procédé décrit par MAUL (1982). Finalement, le titre bactérien de l'eau analysée a été exprimé par le nombre de bactéries hétérotrophes présentes dans 1 ml de l'échantillon initial.

#### 2.1.1.5. Analyse statistique des quantifications bactériennes

La comparaison statistique des résultats obtenus par les différentes méthodes de titrage bactérien a été réalisée par le test du rapport de vraisemblance (KENDALL et STUART, 1973 ; LINDGREN, 1968 ; RAO, 1973). L'intérêt principal de ce test tient dans sa généralité qui le rend applicable quelles que soient les lois de probabilité caractérisant les résultats de chacune des m méthodes de quantification à comparer. Cependant, la formulation mathématique exacte de la nature de ces lois de distribution doit, en tout état de cause, être parfaitement connue.

En notant  $x_{ij}$  le résultat correspondant à la dilution de rang j obtenu pour la méthode i, la fonction de vraisemblance associée à la densité bactérienne moyenne inconnue,  $\lambda$ , est donnée par

$$L(\lambda) = \prod_{i=1}^m P(x_{i1}, x_{i2}, \dots, x_{i\ell_i} / \lambda) \quad (37)$$

où  $P(x_{i1}, x_{i2}, \dots, x_{i\ell_i} / \lambda)$  représente la probabilité d'observer les résultats  $(x_{i1}, x_{i2}, \dots, x_{i\ell_i})$  lorsque le titre bactérien réel de l'échantillon est égal à  $\lambda$ . Si, de plus, il est admis que le nombre des colonies recensées par la méthode des membranes filtrantes et la méthode par inclusion en gélose obéit à la loi de Poisson tandis que le nombre de réponses positives à un niveau de dilution donné pour la méthode du N.P.P. répond à une loi binomiale, alors la forme explicite de  $L(\lambda)$  devient

$$L(\lambda) = L_1(\lambda) \cdot L_2(\lambda) \cdot L_3(\lambda) \quad , \text{ avec}$$

$$L_1(\lambda) = \prod_{j=1}^{\ell_1} e^{-\lambda v_{1j}} \frac{(\lambda v_{1j})^{x_{1j}}}{x_{1j}!} \quad ,$$

$$L_2(\lambda) = \prod_{j=1}^{\ell_2} e^{-\lambda v_{2j}} \frac{(\lambda v_{2j})^{x_{2j}}}{x_{2j}!} \quad , \text{ et}$$

$$L_3(\lambda) = \prod_{j=1}^{\ell_3} \binom{n_j}{p_j} e^{-\lambda v_{3j}} (\lambda v_{3j})^{p_j} (1 - e^{-\lambda v_{3j}})^{n_j - p_j} \quad (38)$$

Dans ces formules,  $\ell_i$  désigne le nombre des niveaux de dilution considérés pour la méthode i et  $v_{ij}$  le volume d'échantillon initial filtré, incorporé dans la gélose ou encore inoculé à la j<sup>ème</sup> dilution selon que, respectivement,  $i = 1, 2$  ou  $3$ , alors que  $p_j$  représente le nombre de réponses positives parmi les  $n_j$  cupulesensemencées à la dilution j pour la microméthode de titrage.

Par suite, l'estimation de la densité bactérienne moyenne,  $\hat{\lambda}_i$ , par la méthode du maximum de vraisemblance de Fisher, se calcule en maximisant la fonction de vraisemblance, définie par  $L_i(\lambda)$ , qui est associée aux résultats obtenus pour chacune des techniques de titrage considérées. La valeur  $\hat{\lambda}_i$  correspondant à ce maximum est déterminée par la résolution de l'équation obtenue en annulant la dérivée par rapport à  $\lambda$  du logarithme de  $L_i(\lambda)$ . Ce calcul est effectué distinctement pour  $L_1(\lambda)$ ,  $L_2(\lambda)$  et  $L_3(\lambda)$  mais également pour  $L(\lambda)$  en admettant l'absence d'effets liés à la technique mise en oeuvre.

Ainsi, pour  $L_1(\lambda)$ ,  $\hat{\lambda}_1$  s'obtient en résolvant l'équation

$$\sum_{j=1}^{\ell_1} x_{1j} = \lambda \sum_{j=1}^{\ell_1} v_{1j} \quad ; \quad (39)$$

$\hat{\lambda}_2$  s'obtient de manière analogue, tandis que  $\hat{\lambda}_3$  est la solution de l'équation

$$\sum_{j=1}^{\ell_3} v_{3j} (n_j - p_j) = \sum_{j=1}^{\ell_3} p_j \frac{v_{3j} e^{-\lambda v_{3j}}}{1 - e^{-\lambda v_{3j}}} \quad (40)$$

déjà donnée par COCHRAN (1950).

Sous l'hypothèse de l'absence d'un "effet méthode", l'estimation d'une valeur unique et commune,  $\hat{\lambda}$ , de la densité bactérienne conduit à l'équation

$$\frac{1}{\lambda} \left[ \sum_{j=1}^{\ell_1} x_{1j} + \sum_{j=1}^{\ell_2} x_{2j} \right] + \sum_{j=1}^{\ell_3} p_j \frac{v_{3j} e^{-\lambda v_{3j}}}{1 - e^{-\lambda v_{3j}}} =$$

$$\sum_{j=1}^{\ell_1} v_{1j} + \sum_{j=1}^{\ell_2} v_{2j} + \sum_{j=1}^{\ell_3} v_{3j} (n_j - p_j) \quad (41)$$

dont les différents termes du premier membre traduisent la quantité d'information imputable à chacune des méthodes étudiées.

Tout comme l'équation (40), la résolution de l'équation (41) nécessite l'usage d'un calculateur électronique. Un exemple de programme de résolution de ces équations est proposé en Annexe IV.

Le test du rapport de vraisemblance, destiné à comparer les résultats obtenus par les trois méthodes de titrage utilisées, est basé sur le rapport

$$\Lambda = \frac{L(\hat{\lambda})}{L_1(\hat{\lambda}_1) \cdot L_2(\hat{\lambda}_2) \cdot L_3(\hat{\lambda}_3)} \quad (42)$$

D'une manière générale, l'hypothèse nulle (i. e., équivalence des estimations données par les  $m$  méthodes comparées) est contrôlée en confrontant la statistique  $-2 \ln \Lambda$  à la distribution de  $\chi^2$  avec  $(m-1)$  degrés de liberté.

En l'occurrence la quantité

$$\begin{aligned}
 -2 \ln \Lambda = 2 \left[ & (\hat{\lambda} - \hat{\lambda}_1) \sum_{j=1}^{\ell_1} v_{1j} + \sum_{j=1}^{\ell_1} x_{1j} \ln \frac{\hat{\lambda}_1}{\hat{\lambda}} + (\hat{\lambda} - \hat{\lambda}_2) \sum_{j=1}^{\ell_2} v_{2j} \right. \\
 & + \sum_{j=1}^{\ell_2} x_{2j} \ln \frac{\hat{\lambda}_2}{\hat{\lambda}} + (\hat{\lambda} - \hat{\lambda}_3) \sum_{j=1}^3 v_{3j} (n_j - p_j) \\
 & \left. + \sum_{j=1}^3 p_j \ln \frac{1 - e^{-\hat{\lambda}_3 v_{3j}}}{1 - e^{-\hat{\lambda} v_{3j}}} \right] \quad (43)
 \end{aligned}$$

est comparée à la loi de  $\chi^2$  à 2 degrés de liberté pour un risque d'erreur de première espèce préalablement fixé.

### 2.1.2. Protocole expérimental

Des échantillons de 500 ml d'eau du robinet, numérotés de 1 à 5, ont été prélevés au cours d'une période de quinze jours environ. Ces échantillons ont été analysés conformément au protocole représenté par la figure 1.

Après homogénéisation par 10 retournements du flacon de 500 ml, un sous-échantillon de 50 ml de la suspension initiale a été recueilli dans un erlenmeyer dont 5 ml du contenu ont ensuite été introduits dans un erlenmeyer contenant 45 ml d'eau physiologique stérile. Cette suspension bactérienne à la dilution  $10^{-1}$  a permis de réaliser une dilution à  $10^{-2}$  suivant le même procédé. Tout prélèvement de liquide a été précédé par au moins deux aspirations-refoulements du liquide avec une pipette de 10 ml.

A partir de ces trois niveaux de dilution, trois titrages ont été réalisés en parallèle pour chacune des méthodes de titrage considérées : méthode des membranes filtrantes (MF), méthode par inclusion en gélose (IG) et méthode quantique en milieu liquide basée sur le calcul du nombre le plus probable (N.P.P.). Pour la première de ces trois méthodes seulement, deux essais supplémentaires ont été effectués en filtrant successivement trois sous-échantillons de 10 ml, puis trois sous-échantillons de 100 ml directement prélevés dans le flacon de 500 ml.

Afin de minimiser l'influence d'une éventuelle dérive dans le temps de la concentration bactérienne, les titrages par les différentes techniques ont été réalisés de front et la période s'écoulant entre le moment du prélèvement de l'échantillon au robinet et la fin des manipulations n'a jamais excédé 3 heures.

Après lecture et calcul des résultats comme décrit précédemment, une estimation de la densité bactérienne a été obtenue pour chaque méthode et pour chaque échantillon initial. La similarité des titres bactériens ainsi calculés à finalement été appréciée au moyen du test du rapport de vraisemblance effectué sur l'ensemble des données brutes.

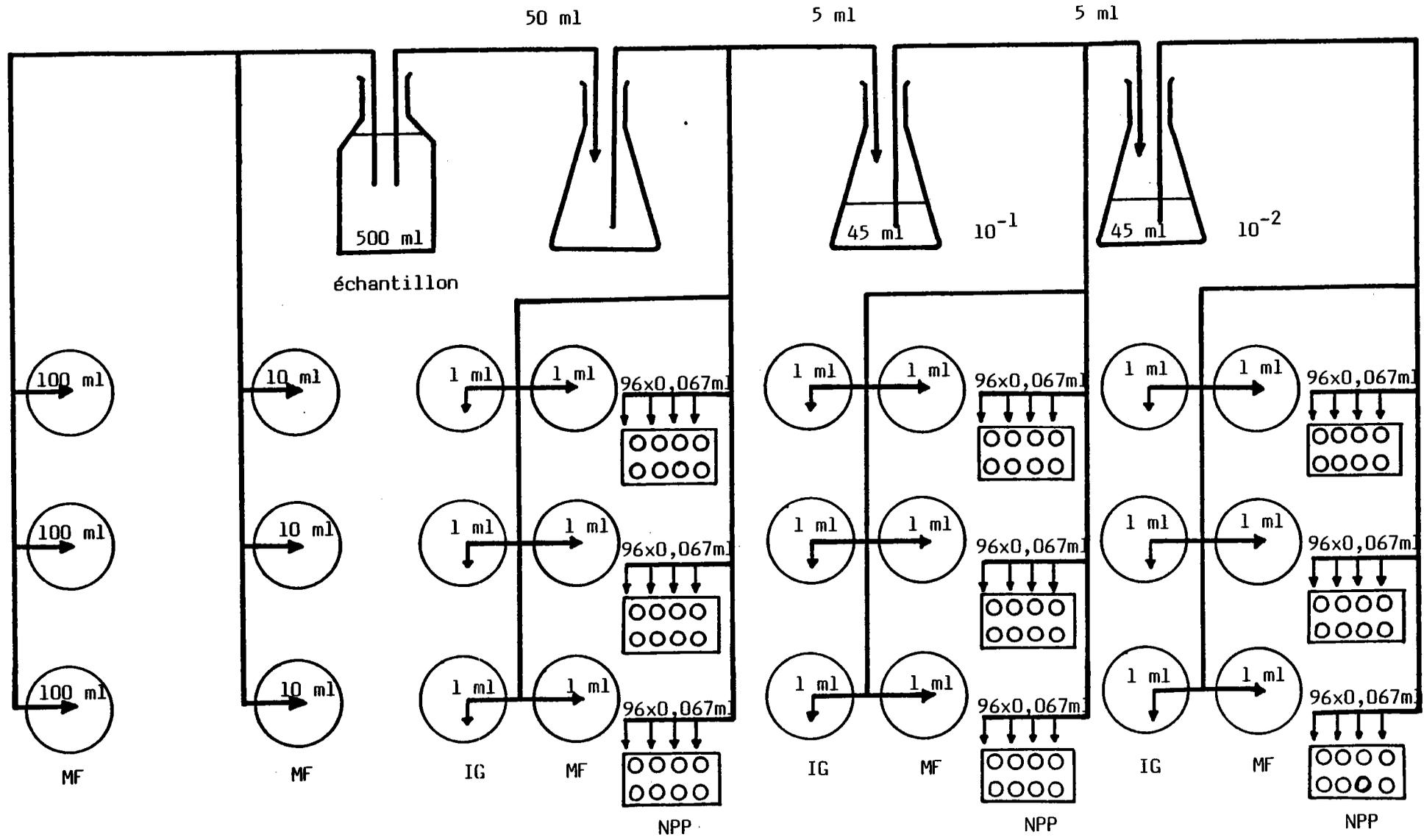


FIGURE 1 : Schéma du protocole expérimental pour la comparaison des titres bactériens obtenus par trois méthodes de titrage (membranes filtrantes, inclusion en gélose et numération en milieu liquide).

### 2.1.3. Résultats

Le tableau 1 donne, pour chaque échantillon d'eau du robinet, les résultats bruts des quantifications bactériennes réalisées par les trois techniques étudiées. Dans le cas des méthodes énumératives (i. e., MF et IG), les comptages bactériens des trois titrages effectués en parallèle sont reportés pour chaque niveau de dilution considéré. Les résultats de la méthode quantique en milieu liquide (N.P.P.) sont présentés sous la forme de triplets de réponses positives associés aux trois titrages réalisés simultanément avec ces mêmes niveaux de dilution.

Les estimations  $\hat{\lambda}_1$ ,  $\hat{\lambda}_2$  et  $\hat{\lambda}_3$  de la densité bactérienne, obtenues pour chacune des trois techniques de numération par la méthode du maximum de vraisemblance, sont consignées dans le tableau 2. Tous les titres bactériens observés sont compris entre 0 et 2367 micro-organismes par millilitre. Les tests de comparaison effectués dans le cadre de l'analyse statistique des résultats sont également résumés dans ce tableau. A cet égard, la statistique  $-2 \ln \Lambda$ , calculée d'après la formule (43) dans laquelle  $\hat{\lambda}$  a été substituée par la solution de l'équation (41) en conformité avec l'hypothèse nulle, a été comparée à la distribution de  $\chi^2$  avec 2 degrés de liberté afin de tester l'égalité des trois estimations associées à chaque échantillon. Le test s'est montré significatif ( $P < 0,05$ ) pour les échantillons n°2, 3, 4 et 5. De manière plus générale, l'indépendance des variables  $-2 \ln \Lambda$  pour chaque échantillon permet d'intégrer l'ensemble de tous les résultats en vue

d'effectuer un test unique ; en effet,  $\sum_{i=1}^5 -2 \ln \Lambda_i$  suit une loi de  $\chi^2$  avec 10 degrés de liberté. La valeur 512,31 trouvée pour  $\sum_{i=1}^5 -2 \ln \Lambda_i$  se situe

nettement au-delà de la limite supérieure du domaine d'acceptation donné par la table du khi-deux pour 10 degrés de liberté au seuil de probabilité de 0,005% ( $\chi_{0,995}^2 = 25,19$ ).

TABLEAU 1

RESULTATS OBTENUS PAR TROIS TECHNIQUES DE TITRAGE :  
MEMBRANES FILTRANTES (MF), INCLUSION EN GELOSE (IG)  
ET METHODE QUANTIQUE EN MILIEU LIQUIDE (N.P.P.)

ECHANTILLON	VOLUME D'ECHANTILLON INITIAL ANALYSE (a)	METHODE DE TITRAGE		
		MF	IG	N.P.P. (b)
1	100 ml	26 - 26 - 28		
	10 ml	2 - 3 - 2		
	1 ml	1 - 0 - 0	0 - 0 - 0	1 - 0 - 0
	0,1 ml	0 - 0 - 0	0 - 0 - 0	0 - 0 - 0
	0,01 ml	0 - 0 - 0	0 - 0 - 0	0 - 0 - 0
2	100 ml	ill. (c)		
	10 ml	205 - 220 - 192		
	1 ml	24 - 20 - 23	25 - 28 - 27	96 - 96 - 96
	0,1 ml	5 - 3 - 1	3 - 3 - 8	38 - 39 - 36
	0,01 ml	1 - 0 - 0	1 - 0 - 0	12 - 10 - 9
3	100 ml	ill.		
	10 ml	ill.		
	1 ml	ill.	ill.	96 - 96 - 96
	0,1 ml	122 - 121 - 130	148 - 145 - 130	96 - 96 - 96
	0,01 ml	14 - 15 - 12	15 - 14 - 12	33 - 56 - 61
4	100 ml	ill.		
	10 ml	ill.		
	1 ml	ill.	ill.	96 - 96 - 96
	0,1 ml	35 - 27 - 32	33 - 37 - 36	43 - 45 - 49
	0,01 ml	2 - 3 - 4		2 - 3 - 6
5	100 ml	ill.		
	10 ml	ill.		
	1 ml	ill.	ill.	96 - 96 - 96
	0,1 ml	ill.	251 - 212 - 226	96 - 96 - 96
	0,01 ml	27 - 22 - 22	23 - 22 - 23	67 - 31 - 52

(a) les volumes consignés dans cette colonne se rapportent uniquement aux méthodes MF et IG

(b) résultats exprimés sous la forme d'un triplet de réponses positives émanant de trois dilutions successives et trois répétitions

(c) illisible

ECHAN- TILLON	ESTIMATION LA PLUS VRAISEMBLABLE DE LA DENSITE BACTERIENNE MOYENNE PAR ml				- 2 ln $\Lambda$
	(hypothèse nulle) $\hat{\lambda}$	$\hat{\lambda}_1$ (MF)	$\hat{\lambda}_2$ (IG)	$\hat{\lambda}_3$ (N.P.P.)	
1	0,25	0,26	0,00	0,16	4,84
2	26,99	20,82	28,53	88,06	287,09 <sup>++</sup>
3	1297,68	1254,55	1406,06	1113,56	6,61 <sup>+</sup>
4	162,37	312,12	353,33	94,35	139,80 <sup>++</sup>
5	2000,37	2366,67	2293,94	1113,56	73,97 <sup>++</sup>

+ : valeur significative au seuil 5 %

++ : valeur significative au seuil 0,5 %

TABLEAU 2 : Comparaison des estimations de la densité bactérienne moyenne obtenues par trois techniques de titrage : membranes filtrantes (MF), inclusion en gélose (IG) et méthode quantique en milieu liquide (N.P.P.).

Par conséquent, les trois méthodes de titrage ne peuvent pas être considérées comme équivalentes pour ce qui concerne l'analyse bactériologique quantitative des échantillons d'eau potable, du moins lorsque celle-ci est réalisée dans les conditions expérimentales de la présente étude.

Les résultats des deux méthodes énumératives (i. e., MF et IG) paraissant relativement similaires, ces deux méthodes ont fait l'objet d'une comparaison statistique particulière dont le principe reste analogue à celui de l'étude précédente. Ainsi, la différence entre les comptages bactériens fut significative ( $P < 0,05$ ) pour l'échantillon n°2 seulement, et le test effectué sur l'ensemble des données recueillies s'est avéré significatif ( $p < 0,025$ ).

Les tests précédents montrent que les résultats de l'analyse statistique résumée dans le tableau 2 doivent, dans une large mesure, être attribués à la méthode quantique en milieu liquide. Les indices N.P.P. observés pour la microméthode de titrage montrent, en effet, de larges écarts comparativement aux résultats beaucoup plus concordants obtenus pour les deux autres méthodes.

#### 2.1.4. Discussion

##### a) Le test du rapport de vraisemblance

L'analyse statistique des données bactériologiques selon des procédés courants, comme par exemple l'analyse de variance, nécessite le plus souvent une transformation préalable des données brutes afin de les soumettre aux exigences spécifiées par les hypothèses des techniques statistiques mises en oeuvre. La formulation explicite d'une telle transformation, moyennant la

relation existant entre la variance et la moyenne des résultats, est quelquefois possible lorsque toutes les variables considérées obéissent à une même loi de probabilité. En ce sens, la transformation "racine carrée" est adaptée au modèle de Poisson, mais les difficultés d'interprétation des résultats ainsi transformés incitent les microbiologistes à lui préférer la transformation logarithmique (ASHBY et RHODES-ROBERTS, 1976 ; CADA, 1975 ; KAPER et al., 1978 ; KOSINSKI et al., 1979 ; PALMER et al., 1976), surtout lorsque les valeurs nulles ne sont pas trop nombreuses. Le pragmatisme de cette pratique, confortée par les tests d'homogénéité des variances lesquels dénotent le plus souvent une nette amélioration après transformation des données, ne doit cependant pas conduire à une utilisation abusive de l'analyse de variance trop souvent justifiée par sa robustesse. Toute transformation ne peut, en effet, constituer qu'un palliatif dont l'efficacité est mal connue.

De plus, les comptages bactériens effectués à des niveaux de dilution distincts et, par suite, l'utilisation de facteurs multiplicatifs destinés à exprimer la concentration bactérienne par unité de volume, sont à l'origine de discontinuités dans la relation entre la variance et la moyenne qui rendent la transformation "racine carrée" inadéquate.

La complexité du problème est encore accrue lorsque les variables de chaque groupe de données à comparer suivent des lois de probabilité différentes. Ceci se produit notamment dans le cas où les densités bactériennes à confronter sont obtenues par plusieurs méthodes de titrage et techniques de calcul des résultats (e. g., méthode par filtration sur membrane ou par inclusion en gélose *versus* méthode du N.P.P. en milieu liquide).

Le test du rapport de vraisemblance offre une alternative intéressante permettant d'éviter ces difficultés qui entravent toute comparaison rigoureuse des titres bactériens, en particulier lorsque ceux-ci obéissent à des lois de probabilité distinctes, mais dont la nature est cependant parfaitement connue. Cet outil statistique qui ne requiert aucune transformation des données en exploitant toute l'information disponible sur la base des seules distributions théoriques, mérite de retenir l'attention des microbiologistes.

#### b) Les méthodes de titrage

La microméthode de titrage en milieu liquide a donné des titres bactériens sensiblement inférieurs à ceux qui ont été observés pour les méthodes énumératives dans trois cas sur un total de cinq échantillons analysés, et l'indice N.P.P. s'est montré significativement supérieur pour une expérience seulement. Cette observation semble aller à l'encontre de certaines études qui, réalisées notamment en matière de colimétrie, relatent le plus souvent une tendance à la surestimation de l'indice N.P.P. par rapport au comptage direct des colonies effectué par exemple selon la méthode des membranes filtrantes (BLOCK et al., 1982 ; GREEN et LITSKY, 1980 ; MAUL et BLOCK, 1983 ; McCARTHY et al., 1961 ; THOMAS et WOODWARD, 1955 ; VELZ, 1970). Malgré le caractère délicat, voire inopportun d'un tel rapprochement entre la méthode quantique en milieu liquide utilisée dans ce travail et la méthode, somme toute très différente, qui est habituellement mise en oeuvre en colimétrie, il convient de rappeler que la surestimation observée dans ce dernier cas, a souvent été imputée à la nature même des milieux de culture. Le milieu liquide étant communément supposé offrir de meilleures conditions de croissance aux bactéries relativement aux milieux solides, qui, de fait, sont globalement plus sélectifs (EL-SHAARAWI et PIPES, 1982 ; THOMAS, 1955).

En tout état de cause, les disparités très importantes affectant les résultats de la microméthode, dont les écarts avec les autres techniques n'ont montré aucune structure reproductible au cours des cinq analyses d'échantillons, témoignent d'une forte interaction méthode-échantillon. Ce

phénomène trouve une explication partielle dans la sélectivité différentielle exercée par les méthodes de titrage sur les diverses espèces de micro-organismes dont les fréquences relatives peuvent varier d'un échantillon à l'autre. De plus, l'absence de système indicateur fiable pour l'identification des cupules positives entrave la lecture des plaques de microtitration et peut, par conséquent, avoir des répercussions fâcheuses sur le calcul du N.P.P. et l'expression du titre bactérien final. A cet égard, le coefficient de variation de 33 %, calculé à partir des résultats provenant des échantillons n°2, 3, 4 et 5, contraste avec les 13 % du niveau théorique du coefficient de variation de la microméthode lorsque 96 ensemencements sont effectués pour trois niveaux de dilution successifs. Cette valeur théorique qui exprime le degré de reproductibilité de la technique a été vérifiée expérimentalement en colimétrie par MAUL et BLOCK (1983). Ces considérations nous amènent à écarter la méthode quantique en milieu liquide, du moins dans sa forme actuelle, pour la numération des bactéries hétérotrophes contenues dans les eaux de distribution.

La similarité des résultats émanant des deux méthodes d'énumération directe des bactéries semble indiquer que l'anaérobiose relative et le choc thermique provoqués par l'incorporation d'un sous-échantillon d'eau dans la gélose en fusion (KLEIN et WU, 1974 ; VAN SOESTBERGER et LEE, 1969) n'ont pas exercé d'effet inhibiteur perceptible sur la reviviscence bactérienne dans les conditions expérimentales de cette étude. Cependant, le diamètre des colonies formées, qui a parfois été utilisé en tant que critère de convenance pour la comparaison expérimentale des milieux de culture (RICHARD et al., 1983), peut, par sa petitesse, quelquefois entraver le comptage des colonies, en particulier dans le cas de la méthode par inclusion en gélose. En outre, la plus grande souplesse caractérisant la filtration sur membrane, notamment pour ce qui concerne la quantité d'eau analysée, ainsi que le gain de sensibilité et de précision qui peuvent en résulter, devraient contribuer à conférer à la méthode des membranes filtrantes une place privilégiée pour le dénombrement des bactéries hétérotrophes présentes dans les eaux de distribution ; ceci au détriment de la méthode par inclusion en gélose, actuellement utilisée en qualité de technique de référence. Ces considérations nous ont amenés à retenir la méthode des membranes filtrantes pour toutes les quantifications bactériennes effectuées dans la suite de ce travail.

## 2.2. COMPARAISON EXPERIMENTALE DE QUELQUES MILIEUX DE CULTURE

Notre préférence, en matière de technique de titrage, allant à la méthode des membranes filtrantes, le but de cette deuxième expérience consiste à étudier les rendements relatifs de trois milieux de culture (i. e., la gélose nutritive (GN), le milieu R-2A et le milieu m-SPC) pour le dénombrement des bactéries hétérotrophes à 20°C. L'accent sera mis sur l'aspect cinétique de la reviviscence bactérienne caractérisant chacun des milieux étudiés afin de pouvoir fixer, sur des bases objectives, le choix d'un milieu de culture et d'une période d'incubation offrant les meilleures conditions de croissance bactérienne.

### 2.2.1. Matériel et méthodes

La collecte des échantillons d'eau a été réalisée conformément à la description donnée au paragraphe 2.1.1.1.

### a) Milieux de culture et technique de titrage des bactéries

Après filtration sur membrane des échantillons d'eau, suivant la description donnée au paragraphe 2.1.1.2., les membranes ont été déposées sur milieu gélosé.

La composition chimique des trois milieux utilisés est formulée comme suit :

- gélose nutritive (GN) (tryptone : 5,0 g ; extrait de levure : 2,5 g ; glucose : 1,0 g ; gélose : 15,0 g ; eau distillée : 1000 ml),
- milieu R-2A (extrait de levure : 0,5 g ; protéose peptone # 3 : 0,5 g ; vitamin assay casamino acids : 0,5 g ; glucose : 0,5 g ; amidon soluble : 0,5 g ;  $K_2HPO_4$  : 0,3 g ;  $MgSO_4 \cdot 7H_2O$  : 0,05 g ; pyruvate de sodium : 0,3 g ; gélose : 20,0 g ; eau distillée : 1000 ml) (REASONER et GELDREICH, 1985),
- milieu m-SPC (peptone : 20,0 g ; gélatine : 25,0 g ; glycérol : 10,0 ml ; gélose : 15,0 g ; eau distillée : 1000 ml) (TAYLOR et GELDREICH, 1979).

Toutes les colonies apparues sur la membrane ont été dénombrées après incubation à 20°C pendant 96 heures avec une lecture intermédiaire des boîtes à 48 puis 72 heures d'incubation.

### b) Exploitation statistique des résultats

La comparaison statistique des résultats, obtenus pour les différents milieux de culture après un temps d'incubation donné, a été effectuée à l'aide du test du rapport de vraisemblance exposé au paragraphe 2.1.1.5.

Les comptages bactériens étant par hypothèse régis par une loi de Poisson, le rendement relatif de deux milieux de culture a été évalué au moyen du rapport  $\hat{\rho} = \frac{r_1}{r_2}$  où  $r_1$  et  $r_2$  représentent deux variables de Poisson

indépendantes (i. e., les comptages bactériens réalisés pour un même volume d'eau filtrée) dont les moyennes respectives sont égales à  $\lambda_1$  et  $\lambda_2$ . Dans ce cas,  $\hat{\rho}$  désigne la valeur la plus vraisemblable du rapport  $\frac{\lambda_1}{\lambda_2}$  des paramètres

des deux distributions considérées. Un intervalle de confiance à 95 % de ce quotient est donné par l'ensemble des solutions de l'inéquation suivante d'inconnue  $\rho$  :

$$\frac{1 + \rho^2}{\rho} \left[ \left| r_1 - \frac{(r_1 + r_2)\rho}{1 + \rho} \right| - 0,5 \right]^2 \leq 1,96^2 (r_1 + r_2) \quad (44)$$

(EL-SHAARAWI et PIPES, 1982).

### 2.2.2. Protocole expérimental

Des échantillons de 500 ml d'eau de distribution, numérotés de 1 à 12, ont été prélevés en différents endroits du réseau de distribution de la ville de Metz.

Après homogénéisation par 10 retournements du flacon de 500 ml, des sous-échantillons de 100 ml, 10 ml et 1 ml de l'échantillon initial ont été filtrés sur membrane à raison de trois répétitions pour les sous-échantillons de 100 ml et de six répétitions pour les volumes de 10 ml et 1 ml. Par suite, trois séries de titrages ont été réalisées en parallèle en déposant équitablement les membranes sur trois milieux de culture, à savoir :

la gélose nutritive (GN), le milieu R-2A et le milieu m-SPC. Ainsi, pour chaque milieu de culture et pour chaque échantillon de 500 ml analysé, deux boîtes ont été obtenues pour les sous-échantillons de 10 ml et de 1 ml alors qu'une seule boîte était associée aux sous-échantillons de 100 ml.

Au laboratoire, toutes les manipulations ont été effectuées en moins de trois heures afin de réduire les effets indésirables, pour ce qui concerne les comparaisons entre milieux, d'une éventuelle dérive dans le temps de la concentration bactérienne.

Toutes les boîtes ont été retirées de l'étuve après 48, 72 et enfin 96 heures d'incubation dans le but de procéder au recensement des colonies bactériennes.

L'étude comparative des comptages bactériens après chaque période d'incubation considérée a été effectuée par le test du rapport de vraisemblance sur l'ensemble des données brutes se rapportant à une même quantité d'eau filtrée pour chacun des trois milieux de culture.

### 2.2.3. Résultats

Le tableau 3 donne, pour chaque échantillon d'eau et pour chaque période d'incubation considérée, les résultats des comptages bactériens obtenus pour les trois milieux de culture : gélose nutritive (GN), milieu R-2A et milieu m-SPC.

L'étude de l'évolution du nombre des colonies bactériennes apparues sur milieu gélosé en fonction du temps d'incubation, ainsi que la comparaison des rendements relatifs caractérisant les différents milieux utilisés, reposent sur le nombre total des colonies recensées pour chaque milieu de culture après une période d'incubation donnée. Ces totaux, qui ont été consignés dans la troisième colonne du tableau 4 concernent uniquement les boîtes ayant permis une lecture séquentielle des résultats après 48, 72 et 96 heures d'incubation à 20°C et portant sur des volumes d'eau filtrée équitablement répartis entre les trois milieux de culture de manière à respecter une parfaite symétrie dans le plan expérimental. Les comptages retenus pour le calcul de ces sommes ont été soulignés dans le tableau 3.

DENOMBREMENTS BACTERIENS REALISES POUR TROIS MILIEUX DE CULTURE  
(GELOSE NUTRITIVE, MILIEU R-2A ET MILIEU m-SPC)  
EN FONCTION DU TEMPS D'INCUBATION

ECHANTILLON	TEMPS D'INCUBATION	VOLUME D'EAU FILTREE								
		100 ml			10 ml			1 ml		
		GN	R-2A	m-SPC	GN	R-2A	m-SPC	GN	R-2A	m-SPC
1	48 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	80
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill.
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill.
2	48 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	47 <u>49</u>	17 <u>58</u>	<u>14</u>
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	70 <u>109</u>	62 <u>121</u>	<u>134</u>
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	167 <u>238</u>	64 <u>201</u>	<u>267</u>
3	48 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill.
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill.
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill. ill.	ill.
4	48 h	ill. nd	ill. nd	ill. nd	107 139	78 63	94 79	<u>12</u> <u>13</u>	<u>7</u> <u>6</u>	<u>10</u> <u>9</u>
	72 h	ill. nd	ill. nd	ill. nd	146 ill.	ill. ill.	ill. ill.	<u>23</u> <u>30</u>	<u>23</u> <u>27</u>	<u>62</u> <u>29</u>
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	<u>62</u> <u>71</u>	<u>62</u> <u>74</u>	<u>64</u> <u>64</u>
5	48 h	118 nd	94 nd	104 nd	<u>9</u> <u>7</u>	<u>8</u> <u>4</u>	<u>5</u> <u>4</u>	<u>0</u> <u>1</u>	<u>0</u> <u>1</u>	<u>0</u> <u>1</u>
	72 h	ill. nd	ill. nd	ill. nd	<u>53</u> <u>13</u>	<u>43</u> <u>27</u>	<u>51</u> <u>23</u>	<u>0</u> <u>2</u>	<u>3</u> <u>1</u>	<u>1</u> <u>2</u>
	96 h	ill. nd	ill. nd	ill. nd	<u>56</u> <u>13</u>	<u>49</u> <u>37</u>	<u>51</u> <u>38</u>	<u>1</u> <u>2</u>	<u>3</u> <u>1</u>	<u>1</u> <u>2</u>
6	48 h	136 nd	118 nd	142 nd	<u>11</u> <u>4</u>	<u>9</u> <u>4</u>	<u>10</u> <u>7</u>	<u>1</u> <u>1</u>	<u>0</u> <u>0</u>	<u>0</u> <u>1</u>
	72 h	186 nd	ill. nd	ill. nd	<u>18</u> <u>20</u>	<u>25</u> <u>12</u>	<u>20</u> <u>11</u>	<u>1</u> <u>1</u>	<u>0</u> <u>2</u>	<u>1</u> <u>5</u>
	96 h	ill. nd	ill. nd	ill. nd	<u>21</u>	<u>25</u> <u>18</u>	<u>20</u> <u>16</u>	<u>1</u> <u>1</u>	<u>0</u> <u>2</u>	<u>1</u> <u>5</u>

TABLEAU 3 (suite et fin)

ECHANTILLON	TEMPS D'INCUBATION	VOLUME D'EAU FILTREE								
		100 ml			10 ml			1 ml		
		GN	R-2A	m-SPC	GN	R-2A	m-SPC	GN	R-2A	m-SPC
7	48 h	37 nd	36 nd	33 nd	$\frac{5}{2}$	$\frac{3}{5}$	$\frac{3}{3}$	$\frac{1}{1}$	$\frac{0}{1}$	$\frac{1}{11}$
	72 h	39 nd	36 nd	33 nd	$\frac{6}{2}$	$\frac{6}{5}$	$\frac{3}{3}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{11}$
	96 h	39 nd	36 nd	34 nd	$\frac{7}{4}$	$\frac{6}{6}$	$\frac{3}{3}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{12}$
8	48 h	nd nd	49 nd	31 nd	$\frac{4}{1}$	$\frac{3}{4}$	$\frac{3}{2}$	$\frac{0}{0}$	$\frac{0}{1}$	$\frac{0}{1}$
	72 h	nd nd	49 nd	35 nd	$\frac{4}{2}$	$\frac{4}{4}$	$\frac{3}{2}$	$\frac{1}{0}$	$\frac{0}{1}$	$\frac{0}{1}$
	96 h	nd nd	160 nd	146 nd	$\frac{10}{2}$	$\frac{16}{4}$	$\frac{14}{2}$	$\frac{1}{0}$	$\frac{1}{1}$	$\frac{0}{1}$
9	48 h	ill. nd	ill. nd	ill. nd	$\frac{19}{18}$	$\frac{10}{19}$	$\frac{6}{19}$	$\frac{0}{1}$	$\frac{0}{2}$	$\frac{0}{1}$
	72 h	ill. nd	ill. nd	ill. nd	$\frac{25}{28}$	$\frac{25}{47}$	$\frac{20}{45}$	$\frac{0}{1}$	$\frac{1}{2}$	$\frac{1}{2}$
	96 h	ill. nd	ill. nd	ill. nd	$\frac{48}{28}$	$\frac{70}{56}$	$\frac{68}{53}$	$\frac{4}{1}$	$\frac{6}{5}$	$\frac{4}{4}$
10	48 h	101 nd	82 nd	78 nd	10 7	10 10	7 3	$\frac{1}{0}$	$\frac{1}{2}$	$\frac{1}{0}$
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{98}{99}$	$\frac{90}{99}$	$\frac{97}{99}$
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{98}{99}$	$\frac{91}{99}$	$\frac{97}{99}$
11	48 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{34}{27}$	$\frac{34}{31}$	$\frac{26}{28}$
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{41}{35}$	$\frac{43}{35}$	$\frac{41}{37}$
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{41}{52}$	$\frac{46}{52}$	$\frac{45}{45}$
12	48 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{30}{31}$	$\frac{38}{46}$	$\frac{22}{23}$
	72 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{130}{130}$	$\frac{142}{140}$	$\frac{144}{123}$
	96 h	ill. nd	ill. nd	ill. nd	ill. ill.	ill. ill.	ill. ill.	$\frac{130}{131}$	$\frac{143}{142}$	$\frac{145}{124}$

ill. : illisible  
nd : non déterminé

TABLEAU 4  
RENDEMENTS DES MILIEUX ETUDIES  
EN FONCTION DU TEMPS D'INCUBATION

TEMPS D'INCUBATION à 20°C	MILIEU DE CULTURE	NOMBRE DE COLONIES RECENSEES	RENDEMENT RELATIF *	INTERVALLE DE CONFIANCE à 95 %
48 heures	GN	389	2,05	1,61 - 2,57
	R-2A	287	1,51	1,16 - 1,94
	m-SPC	190	1,00	
72 heures	GN	855	0,91	0,80 - 1,04
	R-2A	904	0,96	0,84 - 1,09
	m-SPC	941	1,00	
96 heures	GN	1123	0,92	0,82 - 1,04
	R-2A	1192	0,98	0,87 - 1,10
	m-SPC	1217	1,00	

\* le rendement est rapporté au nombre de colonies recensées pour le milieu m-SPC

La signification des écarts entre les trois milieux de culture pour chaque temps d'incubation a été appréciée au moyen de la statistique  $- 2 \ln \Lambda$ , calculée d'après la formule (43) adaptée à la comparaison de plusieurs variables de Poisson indépendantes. Les résultats des tests de signification, effectués d'une part, pour l'ensemble des milieux étudiés, et d'autre part sur les paires de milieux comparés deux à deux, sont résumés dans le tableau 5. Les différences entre milieux, qui ont toutes été significatives ( $P < 0,005$ ) après 48 heures d'incubation, se sont progressivement atténuées avec l'augmentation du temps d'incubation. Ainsi, seule la différence entre les milieux GN et m-SPC s'est montrée significative ( $P < 0,05$ ) au terme d'une incubation de 72 heures et, par suite, aucun des écarts observés n'est resté significatif après 96 heures d'incubation.

TABLEAU 5  
 RESULTATS DU TEST DU RAPPORT DE VRAISEMBLANCE  
 (- 2 ln  $\lambda$ ) EFFECTUE SUR LES DONNEES BACTERIOLOGIQUES  
 DU TABLEAU 4

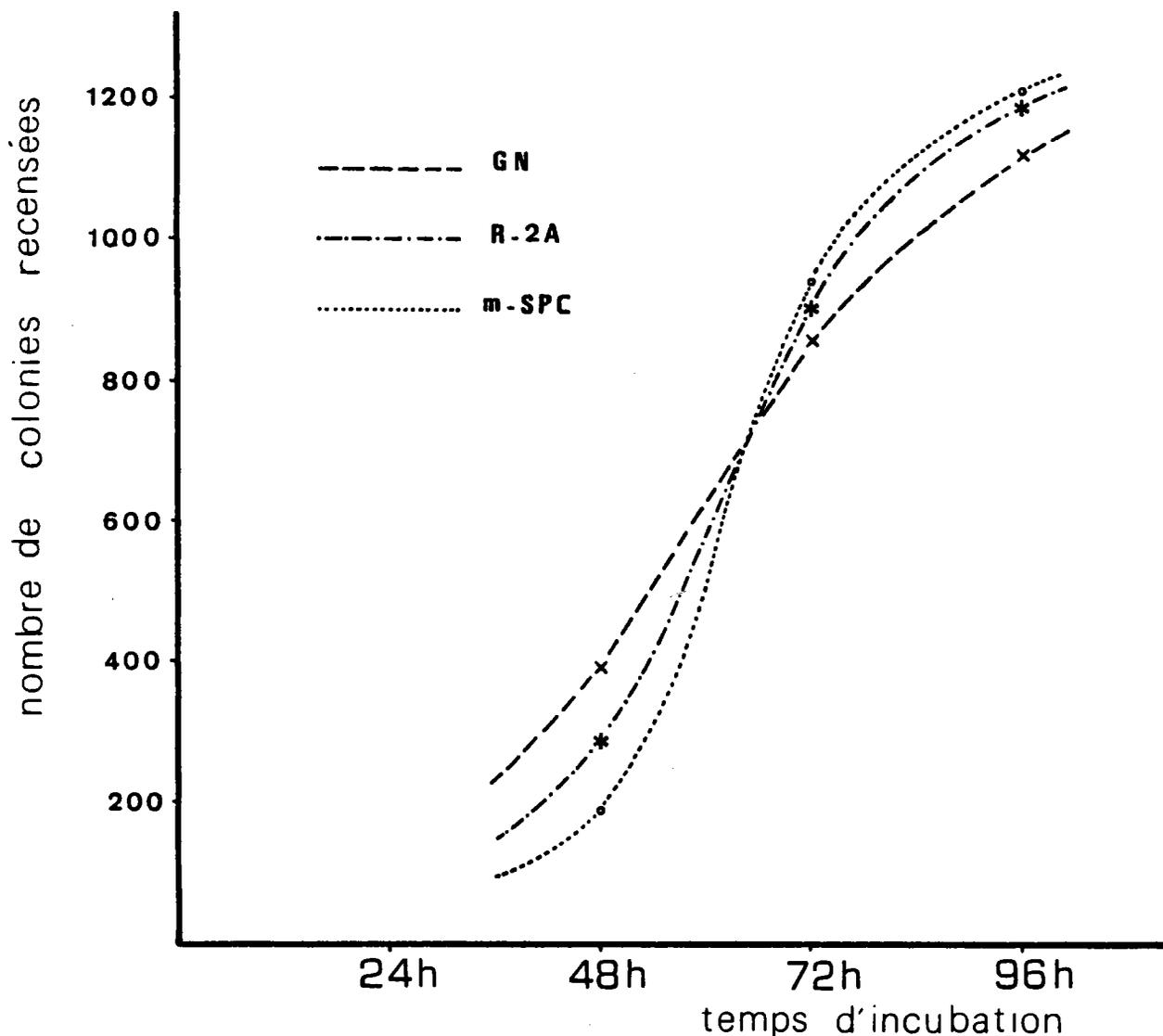
COMPARAISON	TEMPS D'INCUBATION A 20°C		
	48 heures	72 heures	96 heures
Ensemble des milieux	69,82 <sup>++</sup>	4,15	4,05
GN <u>vs</u> R-2A	15,45 <sup>++</sup>	1,37	2,06
GN <u>vs</u> m-SPC	69,81 <sup>++</sup>	4,12 <sup>+</sup>	3,78
R-2A <u>vs</u> m-SPC	19,86 <sup>++</sup>	0,74	0,26

+ : valeur significative au seuil 5 %

++ : valeur significative au seuil 0,5 %

Si la concordance de plus en plus marquée entre les dénombrements bactériens correspondant aux trois milieux avec l'augmentation du temps d'incubation ne laisse aucun doute, il est cependant intéressant de remarquer qu'après 48 heures d'incubation le rendement le plus élevé a été observé pour la gélose nutritive, suivie successivement par les milieux R-2A puis m-SPC. Les rendements relatifs des trois milieux rapportés au milieu m-SPC (Tableau 4) montrent que cet ordre a été inversé après une incubation de 96 heures. Des intervalles de confiance à 95 % de ces rapports, obtenus par la résolution de l'inéquation (44), sont reportés dans la dernière colonne du tableau 4.

Une représentation synthétique de ces phénomènes est donnée par la figure 2 qui tente de décrire l'aspect cinétique de la reviviscence des bactéries hétérotrophes à 20°C selon la nature du milieu de culture utilisé ainsi que la durée de l'incubation.



**FIGURE 2** : Evolution du nombre des colonies recensées sur trois milieux de culture (GN, R-2A et m-SPC) en fonction du temps d'incubation.

Une seconde approche à l'étude du rendement des milieux de culture en fonction du temps d'incubation peut se concevoir par l'étude comparative des dénombrements bactériens réalisés aux différents stades de l'incubation en considérant séparément chacun des trois milieux utilisés. Ainsi, le quotient du nombre total de colonies recensées pour chaque lecture intermédiaire (i. e., 48 et 72 heures d'incubation) par le nombre correspondant obtenu après 96 heures d'incubation s'élève respectivement à 0,35 et 0,76 pour la gélose nutritive, 0,24 et 0,76 pour le milieu R-2A et enfin 0,16 et 0,77 pour le milieu m-SPC. Ces rapports laissent apparaître des disparités importantes entre milieux, lorsque la période d'incubation est inférieure à 72 heures. A partir de 72 heures d'incubation, les résultats indiquent, par contre, une grande similarité pour l'ensemble des trois milieux étudiés.

#### 2.2.4. Discussion

La concordance des dénombrements bactériens pour les trois milieux de culture après 72 heures d'incubation semble, de prime abord, être en désaccord avec les travaux de MEANS et al. (1981) faisant état d'une plus grande efficacité du milieu m-SPC, comparativement à la gélose nutritive et au milieu R-2A, pour la numération des bactéries hétérotrophes présentes dans les eaux de distribution. Toutefois, le manque de rigueur expérimentale caractérisant les essais comparatifs de l'étude de MEANS et al. (1981), nous amène à formuler quelques réserves pour ce qui concerne le bien-fondé de certaines conclusions. En particulier, le rendement relatif des différents milieux testés a été évalué pour des températures et des périodes d'incubation variables selon les milieux de culture.

Notre objection se trouve confortée notamment par l'influence du temps d'incubation sur l'aspect quantitatif de la reviviscence bactérienne sur milieu gélosé. En effet, le nombre des colonies dénombrées après 72 heures d'incubation représente, pour chacun des trois milieux utilisés, environ 76 % du total des colonies recensées au terme d'une incubation de 96 heures. A cet égard, il convient de souligner que les différences entre les comptages afférents aux différents milieux semblent plus marquées par la vitesse de développement et d'apparition des colonies sur la gélose que par le résultat final observable, indépendamment du milieu considéré, lorsque l'incubation est prolongée au-delà de 72 heures.

A partir de 72 heures d'incubation les comptages bactériens sont davantage affectés par le supplément d'incubation que par la nature du milieu de culture. La prolongation de l'incubation au-delà de 72 heures ne représente cependant qu'un intérêt limité et discutable eu égard d'une part, à l'accroissement relativement modéré des comptages bactériens effectués entre 72 et 96 heures, et d'autre part à l'analogie de l'évolution des résultats pour les trois milieux étudiés à partir de 72 heures d'incubation.

Les disparités statistiquement significatives qui ont été mises en évidence entre les numérations bactériennes réalisées pour chaque milieu après une période d'incubation plus courte (i. e., 48 heures) résultent très vraisemblablement de la composition chimique des différents substrats utilisés en considération de l'équipement enzymatique des espèces bactériennes contenues dans l'eau analysée. En effet, la vitesse de croissance des colonies bactériennes dépend, dans une large mesure, de la disponibilité des éléments nutritifs apportés par le milieu de culture, compte tenu des fonctions biosynthétiques des souches présentes (BERLAND et al., 1976) et sans pour autant négliger le pouvoir de revivification des bactéries par suite de l'agression exercée par le chlore (MEANS et al., 1981).

Les résultats des deux expérimentations complémentaires réalisées successivement dans le cadre de ce chapitre nous ont amenés à retenir la méthode des membranes filtrantes pour le titrage des bactéries hétérotrophes à 20°C contenues dans les eaux de distribution. En outre, les caractéristiques cinétiques de la reviviscence bactérienne sur milieu gélosé montrent qu'il est loisible d'utiliser indifféremment chacun des trois substrats étudiés (i. e., gélose nutritive, milieu R-2A et milieu m-SPC) à condition toutefois d'observer une période d'incubation au moins égale à 72 heures.

### CHAPITRE III

MISE EN EVIDENCE PRELIMINAIRE DE  
L'HETEROGENEITE DU MILIEU HYDRIQUE A  
DIFFERENTES ECHELLES D'OBSERVATION

La présence de bactéries dans un réseau de distribution publique en quantités parfois importantes pose le problème de leur origine dont la causalité présente quatre aspects différents. Les bactéries peuvent être d'origine exogène, notamment lorsqu'elles sont :

- apportées par une contamination extérieure
- non éliminées par le traitement

ou endogène, il s'agit dans ce cas :

- de bactéries générées à partir de certaines installations de la filière de traitement
- de bactéries résultant d'une multiplication dans le réseau

### Contamination extérieure

Les bactéries peuvent être introduites dans un réseau de distribution publique, d'une part au niveau des réservoirs de stockage de l'eau qui, en particulier lorsqu'ils ne sont pas couverts, reçoivent des quantités importantes de poussières favorisant la dispersion et la persistance de certains micro-organismes fixés (SILVERMAN et al., 1983), et d'autre part lors des multiples interventions d'entretien ou de nettoyage qui apportent souvent beaucoup de matières en suspension et de bactéries.

### Bactéries non éliminées au cours du traitement

Une filière industrielle complète de traitement des eaux (préchloration - coagulation - floculation - décantation - filtration - postchloration) ne se comporte pas comme un ultrafiltre et laisse inexorablement échapper quelques micro-organismes parmi lesquels certains peuvent ne pas avoir été totalement inactivés. L'analyse d'une quantité suffisante d'eau traitée montre en effet la présence de bactéries capables de se multiplier sur milieu gélosé nutritif mais aussi de très nombreuses bactéries (jusqu'à 10 000 par ml) repérables uniquement par observation microscopique. Il est, par ailleurs, vraisemblable qu'une fraction de ces bactéries apparemment inertes puisse "récupérer" avec le temps du choc imposé par les oxydants (reviviscence). Cette perméabilité partielle des chaînes de traitement s'explique probablement par la plus grande résistance de ces micro-organismes aux désinfectants. Ainsi, une augmentation de la résistance des bactéries, provenant d'échantillons d'eau traitée, a pu être observée pour le chlore (RIDGWAY et OLSON, 1982) de même que pour les antibiotiques et certains métaux (ARMSTRONG et al., 1982 ; CALOMIRIS et al., 1984).

### Bactéries générées à partir de la chaîne de traitement

La croissance bactérienne dans les filtres à sable ou dans les filtres à charbon actif est un phénomène courant qui peut s'observer sur la plupart des filières de traitement même en présence de chlore (DOTT et TRAMPISCH, 1983 ; WILCOX et al., 1983). Le passage de l'eau, notamment lors de la remise en fonctionnement des filtres après leur lavage, entraîne de fines particules solides sur lesquelles sont fixées des microcolonies particulièrement bien protégées de l'action du chlore.

### Multiplication des bactéries dans le réseau

L'observation *in situ* des populations de micro-organismes colonisant les parois internes des canalisations, de même que l'augmentation de la concentration bactérienne de l'eau, souvent constatée dans les secteurs périphériques d'un réseau de distribution, témoignent de l'aptitude des populations bactériennes à se multiplier à l'intérieur d'un réseau. L'examen

au microscope électronique des parois internes des tubes de distribution révèle la présence d'une flore bactérienne diversifiée répartie sur des surfaces irrégulières et cratériformes (ALLEN et al., 1980 ; RIDGWAY et OLSON, 1981). Ces micro-organismes fixés sont présents en quantité variable selon l'âge et la nature du système de distribution (NAGY et al., 1983).

La multiplication des bactéries fixées ou libres dans l'eau de distribution est influencée par plusieurs paramètres comme le temps de séjour de l'eau dans le réseau, la température de l'eau et la quantité de carbone organique assimilable, non éliminé par le traitement (RIZET et al., 1984 ; van der KOOIJ et al., 1982) ou généré au cours du traitement de désinfection (van der KOOIJ et HIJNEN, 1984). Il faut en effet reconnaître qu'une très faible concentration de carbone organique assimilable ne semble pas constituer un obstacle rédhibitoire pour la croissance bactérienne, eu égard aux travaux de DOTT (1983), qui a même pu observer une multiplication bactérienne ( jusqu'à 10 000 bactéries par ml ) *in vitro* dans de l'eau bidistillée ensemencée par de l'eau de distribution.

Ces considérations générales n'apportent que des réponses partielles et qualitatives au problème de l'origine des bactéries dans un réseau de distribution qui reste un phénomène mal compris. En particulier, l'importance relative des facteurs contribuant à l'apparition et à la prolifération des bactéries dans un réseau n'est pas connue.

Par ailleurs, la diversité des sources de contamination (réparations, fuites, ...) ou encore l'incidence probable de certaines caractéristiques physico-chimiques de l'eau (température, temps de rétention, chlore résiduel, ...) sur la reviviscence ou/et la multiplication bactérienne sont à l'origine d'une hétérogénéité de la distribution des bactéries dans un réseau de distribution publique. Cette hétérogénéité qui se manifeste conjointement dans le temps et dans l'espace doit être perçue comme un continuum dans l'espace-temps où des niveaux d'hétérogénéité plus forte sont séparés par des niveaux où la structuration spatio-temporelle peut être plus faible. Cette structure hiérarchisée de la distribution des bactéries dans un réseau est imposée par quelques facteurs prédominants dont les effets encore mal connus méritent d'être étudiés. Ces facteurs en nombre limité se démarquent d'une multitude d'autres phénomènes aux effets moins spectaculaires qui, lorsqu'ils ne sont pas structurés entre eux, ne constituent qu'un bruit de fond (FRONTIER, 1982).

Ces considérations très générales soulignent la nécessité de connaître les principales caractéristiques de la configuration spatio-temporelle de l'hétérogénéité de la distribution des bactéries dans un réseau pour permettre l'interprétation des résultats de tout prélèvement sur des bases objectives. En effet, l'interprétation et la signification d'un prélèvement quelconque se heurtent à la méconnaissance de la variabilité caractérisant les différentes échelles de perception. L'étude préalable de la configuration de l'hétérogénéité apparaît ainsi comme une condition *sine qua non* à toute planification ultérieure de l'échantillonnage, si ce dernier se veut d'être efficace et utile.

L'existence d'un modèle de distribution unique, capable de décrire la répartition de l'ensemble des micro-organismes dans un milieu hydrique de grande étendue est souvent illusoire. La structure hiérarchisée de l'hétérogénéité peut, en effet, pour une échelle d'observation donnée, rendre tout à fait inadapté un modèle qui s'avère pourtant valable à un autre niveau d'observation. A cet égard, DUBY et RODOLPHE (1982) mentionnent l'existence de populations qui, ayant une répartition des individus conforme au modèle de Poisson pour une échelle assez petite, s'en écartent progressivement pour suivre des distributions binomiales négatives à de plus grandes échelles de

perception. Ces changements occasionnés par la rencontre de populations de moyennes différentes se répercutent au niveau de la formulation mathématique de ces distributions dont les paramètres peuvent alors varier en fonction des échelles retenues.

Aucune étude préalable n'ayant, à notre connaissance, fait apparaître d'échelles privilégiées dans ce qui doit être considéré comme un continuum d'échelles possibles, les niveaux d'observation retenus dans ce travail reposent principalement sur des considérations d'ordre purement pratique. L'objectif majeur de ce chapitre consiste à savoir si la variabilité statistique observée à certaines échelles est distinctement plus marquée qu'à d'autres, afin de pouvoir identifier de façon objective des niveaux dans la structure spatio-temporelle du peuplement bactérien d'un réseau de distribution publique.

Le niveau de base de la variabilité étant caractérisé par des échantillons recueillis successivement à un même robinet pendant un intervalle de temps très limité, d'autres niveaux d'hétérogénéité seront ensuite explorés, puis comparés à cette variabilité ponctuelle. Ainsi, les investigations seront progressivement étendues à la variabilité caractérisant des prélèvements effectués dans des appartements distincts d'un même immeuble au cours d'une période de 30 minutes, puis à la variabilité caractérisant des prélèvements effectués de manière systématique dans un quadrat de 650 m de côté, situé en pleine agglomération.

Cette approche, qui doit permettre une meilleure compréhension globale et intuitive de l'hétérogénéité bactérienne, constitue une étape préliminaire à la réalisation d'études plus ambitieuses devant conduire à la stratification du réseau messin.

Par ailleurs, la responsabilité des sociétés assurant la production et la distribution de l'eau potable étant, en principe, engagée seulement jusqu'au niveau du compteur volumétrique de chaque habitation, il nous est apparu opportun de conduire une étude complémentaire visant à apprécier la nature et l'ampleur des altérations éventuelles de la qualité bactériologique de l'eau en aval du compteur.

### 3.1. MATERIEL ET METHODES

#### 3.1.1. Collecte des échantillons d'eau

Les échantillons de 500 ml d'eau ont été recueillis chez l'habitant au moyen de flacons en verre stérilisés à l'autoclave et contenant 1 ml d'une solution de thiosulfate de sodium à 3 %. Les prélèvements ont été précédés par une purge des canalisations internes des différents immeubles retenus en ouvrant les robinets de prélèvement à la pression maximale jusqu'à l'obtention d'une stabilisation de la température de l'eau. La température d'équilibre ainsi obtenue, qui peut d'une certaine manière être considérée comme étant représentative de la température de l'eau circulant dans les canalisations souterraines, garantit l'achèvement de la purge. L'embout de chaque robinet, préalablement débarrassé de ses accessoires éventuels (filtres, brise-jet) a été exposé à la flamme d'une lampe à souder pendant 15 secondes environ avant de procéder au remplissage des flacons avec de l'eau recueillie à la température d'équilibre. Les échantillons ont ensuite été ramenés au laboratoire à la température ambiante pour être soumis à l'analyse bactériologique au plus 4 heures après le prélèvement.

Les habitations équipées d'installations de traitement physico-chimique de l'eau (e. g., les adoucisseurs) ont été délibérément écartées afin d'éviter les risques de contamination bactérienne pouvant éventuellement résulter d'un mauvais entretien de ces appareils.

### 3.1.2. Numération bactérienne et analyse statistique des résultats

- Après filtration sur membrane des échantillons d'eau, conformément à la description donnée au paragraphe 2.1.1.2., les membranes ont été déposées sur milieu R-2A dont la composition chimique est explicitée au paragraphe 2.2.1. Toutes les colonies apparues sur les membranes ont été dénombrées après une période d'incubation de 72 heures à la température de 20°C. Enfin, les résultats ont pu être exprimé en nombre de bactéries hétérotrophes par millilitre d'échantillon suivant le principe du maximum de vraisemblance.

- L'analyse statistique des résultats obtenus se résume à l'étude comparative de la dispersion des titres bactériens observés pour les différentes échelles de perception sélectionnées dans le continuum spatio-temporel. Dans une première étape, la distribution des résultats expérimentaux a été comparée à un modèle de référence (i. e., le modèle de Poisson) au moyen de l'indice de dispersion de Fisher ( $D^2$ ) dont la généralisation au cas d'échantillons de volumes inégaux a fait l'objet du paragraphe 1.1.2. A cet égard, il faut rappeler que si  $n$  estimations indépendantes de la concentration bactérienne d'une eau sont compatibles avec le modèle de Poisson, la statistique  $D^2$ , calculée selon la formule (8), suit approximativement une loi de  $\chi^2$  avec  $n-1$  degrés de liberté. Dans une seconde étape, les indices de dispersion de Fisher obtenus pour les différentes échelles d'observation ont permis de comparer les variabilités propres à chaque niveau de perception afin de mieux saisir la structure hiérarchisée de l'hétérogénéité.

Ces comparaisons, en l'occurrence purement empiriques, sont fondées sur la relation d'ordre des carrés moyens, obtenus en effectuant le rapport de l'indice de dispersion de Fisher par le nombre de degrés de liberté correspondant. La valeur de ces carrés moyens augmente avec le degré d'hétérogénéité des mesures.

S'il est admis que la distribution des titres bactériens observés à chaque niveau de perception est conforme au modèle binomial négatif, l'étude comparative des degrés d'hétérogénéité de plusieurs séries de résultats expérimentaux peut également s'appuyer sur le calcul des valeurs du paramètre  $k$  caractérisant chacune des échelles d'observation. La signification et l'importance de ce paramètre ont été soulignées au paragraphe 1.2.1. Le calcul de la valeur de  $k$  par la méthode du maximum de vraisemblance a été effectué pour chaque niveau d'observation en résolvant l'équation (17) à l'aide du programme informatique donné en Annexe III.

Enfin, l'étude quantitative des numérations bactériennes : compteurs *versus* robinets fait état de l'appartenance de chaque valeur observée à une distribution de Poisson dont elle constitue une estimation du paramètre, ainsi que de la proposition énoncée au paragraphe 1.1.1. relative à la somme de plusieurs variables de Poisson indépendantes. Le test de comparaison est alors effectué, moyennant la somme totale des résultats observés, soit au compteur, soit au robinet, et compte tenu du volume d'eau filtrée, en utilisant l'indice de dispersion de Fisher dont la sommation est limitée à deux termes seulement.

### 3.2. PROTOCOLE EXPERIMENTAL

#### Expérimentation réalisée en vue d'étudier l'hétérogénéité intra-muros

Des échantillons de 500 ml d'eau ont été prélevés dans des immeubles, numérotés de 1 à 5, situés dans différents quartiers de la ville de Metz. A l'intérieur de chaque immeuble deux personnes ont procédé à la collecte des échantillons pendant une période de 30 minutes environ. L'un des échantillonneurs a recueilli 5 échantillons, à raison d'un flacon toutes les 5 minutes, au robinet de purge du compteur volumétrique situé en amont des canalisations desservant l'ensemble des logements de l'immeuble. Simultanément, l'autre échantillonneur s'est rendu dans différents appartements numérotés de 1 à 5 dans chacun desquels il a prélevé successivement 2 échantillons au même robinet. Ces flacons provenant d'un même appartement ont été désignés par les lettres A et B.

#### Expérimentation réalisée à l'échelle d'un quartier de 650 m de côté

Au total, 25 échantillons de 500 ml d'eau ont été recueillis en moins de 4 heures, chez l'habitant, en des points de prélèvement répartis de façon régulière et systématique sur toute l'étendue d'un quadrat de 650 x 650 m. Ces points représentent approximativement les centres des carrés de 130 m de côté, délimités par le quadrillage du quartier étudié.

Au laboratoire, les manipulations ont été effectuées de manière identique pour chacune des deux expérimentations considérées.

Après homogénéisation par 10 retournements des flacons de 500 ml, des sous-échantillons correspondant à 10 ml, 1 ml et 0,1 ml de la suspension bactérienne originelle ont été filtrés sur membrane. La période s'écoulant entre le moment du prélèvement de l'échantillon au robinet et la fin des manipulations au laboratoire est, dans tous les cas, restée inférieure à 7 heures.

Après lecture des résultats, la dispersion des comptages bactériens observés a été appréciée par l'indice de dispersion de Fisher dont le calcul permet de tester l'applicabilité du modèle de Poisson. Finalement une approche quantitative du degré d'hétérogénéité caractérisant chaque niveau d'observation a été tentée par le calcul du paramètre  $k$  qui lui est propre.

### 3.3. RESULTATS

#### a) Etude de l'hétérogénéité

Les résultats des dénombrements bactériens effectués dans le cadre de l'expérimentation intra-muros sont consignés dans le tableau 6. Les comptages bactériens se rapportent respectivement à des volumes de 10 ml, 1 ml et 0,1 ml d'eau filtrée par échantillon de 500 ml recueilli au niveau du compteur ou dans un appartement.

TABLEAU 6

DENOMBREMENTS BACTERIENS EFFECTUES  
AU COMPTEUR AINSI QUE DANS LES APPARTEMENTS  
DE CINQ IMMEUBLES DE LA VILLE DE METZ

IMMEUBLE	POINT DE PRELEVEMENT		ECHANTILLON	VOLUME D'EAU FILTREE			
				10 ml	1 ml	0,1 ml	
1	Compteur		1	ill.	ill.	224	
			2	ill.	ill.	228	
			3	ill.	ill.	215	
			4	ill.	ill.	246	
			5	ill.	ill.	236	
	Appartement		1	A	ill.	ill.	240
				B	ill.	ill.	227
			2	A	ill.	ill.	55
				B	ill.	ill.	50
			3	A	ill.	ill.	59
B	ill.	ill.		64			
4	A	ill.	ill.	47			
	B	ill.	ill.	136			
5	A	ill.	ill.	179			
	B	ill.	ill.	158			
2	Compteur		1	ill.	178	16	
			2	ill.	170	16	
			3	ill.	163	16	
			4	ill.	159	17	
			5	ill.	126	12	
	Appartement		1	A	ill.	165	18
				B	ill.	172	16
			2	A	ill.	41	6
				B	ill.	36	5
			3	A	ill.	51	11
B	ill.	56		6			
4	A	ill.	138	18			
	B	ill.	118	12			
5	A	ill.	117	10			
	B	ill.	113	11			

ill. : illisible

TABLEAU 6 (suite 1)

IMMEUBLE	POINT DE PRELEVEMENT		ECHANTILLON	VOLUME D'EAU FILTRÉE			
				10 ml	1 ml	0,1 ml	
3	Compteur		1	ill.	190	8	
			2	ill.	184	16	
			3	ill.	181	20	
			4	ill.	171	8	
			5	ill.	161	8	
	Appartement		1	A	ill.		20
				B	ill.	ill.	17
			2	A	ill.	161	17
				B	ill.	152	6
			3	A	ill.	199	15
	B	ill.		230	18		
4	A	ill.	11	0			
	B	ill.	10	0			
5	A	ill.	193	10			
	B	ill.	224	10			
4	Compteur		1	ill.	190	25	
			2	ill.	219	22	
			3	ill.	198	22	
			4	ill.	262	26	
			5	ill.	232	25	
	Appartement		1	A	ill.	ill.	32
				B	ill.	ill.	35
			2	A	ill.	192	13
				B	ill.	167	18
			3	A	ill.	180	19
	B	ill.		228	22		
	4	A	ill.	263	27		
		B	ill.	270	25		
	5	A	ill.	226	22		
		B	ill.	203	20		

TABLEAU 6 (suite 2 et fin)

IMMEUBLE	POINT DE PRELEVEMENT	ECHANTILLON	VOLUME D'EAU FILTRÉE			
			10 ml	1 ml	0,1 ml	
5	Compteur	1	ill.	237	21	
		2	ill.	256	25	
		3	ill.	ill.	27	
		4	ill.	ill.	28	
		5	ill.	ill.	28	
	Appartement	1	A	ill.	ill.	29
			B	ill.	ill.	28
		2	A	ill.	ill.	27
			B	ill.	ill.	29
		3	A	ill.	ill.	86
			B	ill.	ill.	30
		4	A	ill.	ill.	25
			B	ill.	ill.	27
		5	A	ill.	ill.	25
			B	ill.	ill.	27

A partir de ces données brutes, la variabilité des résultats caractérisant chaque niveau d'observation a été évaluée au moyen de l'indice de dispersion de Fisher ( $D^2$ ). Toute valeur significative de  $D^2$  doit être interprétée par une surdispersion des résultats qui est incompatible avec le modèle de Poisson. Cependant, le nombre, parfois très faible, de degrés de liberté associé à certains tests, nous a suggéré de mettre à profit l'indépendance des variables  $D^2$  et la propriété d'additivité qui en découle, en vue d'effectuer un test global portant sur un ensemble de résultats plus important.

Ces tests d'adéquation sont résumés dans les tableaux 7a, 7b et 7c qui se réfèrent respectivement à la variabilité observée au compteur, la variabilité caractérisant deux prélèvements successifs effectués à un même robinet et enfin la variabilité existant entre les appartements d'un même immeuble. Les indices  $i$  ( $i = 1, \dots, 5$ ) utilisés dans ces tableaux se rapportent aux immeubles, alors que les indices  $j$  ( $j = 1, \dots, 5$ ) correspondent aux divers appartements.

TABLEAU 7a : Variabilité de la concentration bactérienne au niveau du compteur.

IMMEUBLE	DEGRES DE LIBERTE	$D^2$	DEGRES DE LIBERTE	$\sum_i D_i^2$
1	4	2,42	20	34,10 <sup>+</sup>
2	4	10,69 <sup>+</sup>		
3	4	4,46		
4	4	14,46 <sup>+</sup>		
5	4	2,06		

TABLEAU 7b : Variabilité de la concentration bactérienne entre deux échantillons prélevés successivement au même robinet.

IMMEUBLE	DEGRES DE LIBERTE	APPARTEMENT					DEGRES DE LIBERTE	$\sum_j D_j^2$	$\sum_i \sum_j D_{ij}^2$
		1	2	3	4	5			
1	1	0,36	0,24	0,20	43,28 <sup>++</sup>	1,31	5	45,40 <sup>++</sup>	90,02 <sup>++</sup> (25degrés de liberté)
2	1	0,07	0,41	0,00	2,36	0,04	5	2,88	
3	1	0,24	1,19	2,50	0,05	2,20	5	6,18	
4	1	0,13	1,03	5,79 <sup>+</sup>	0,04	1,29	5	8,29	
5	1	0,02	0,07	27,03 <sup>++</sup>	0,08	0,08	5	27,28 <sup>++</sup>	

Pour ce qui concerne l'utilisation de la statistique  $D^2$ , il faut préciser que l'approximation de la distribution de  $D^2$  par une loi de  $\chi^2$ , est davantage affectée par la valeur du paramètre de la loi de Poisson considérée que par le nombre des degrés de liberté disponibles (EL-SHAARAWI, 1983). En pratique, il est néanmoins déjà possible d'utiliser ce test pour effectuer la comparaison de deux variables de Poisson lorsque leur moyenne est supérieure à 1.

TABEAU 7c : Variabilité de la concentration bactérienne entre les appartements d'un même immeuble.

IMMEUBLE	DEGRES DE LIBERTE	$D^2$	DEGRES DE LIBERTE	$\sum_i D_i^2$
1	4	393,76 <sup>++</sup>	20	1131,29 <sup>++</sup>
2	4	244,10 <sup>++</sup>		
3	4	292,75 <sup>++</sup>		
4	4	54,57 <sup>++</sup>		
5	4	46,11 <sup>++</sup>		

+ : valeur significative au seuil 5 %  
 ++ : valeur significative au seuil 0,5 %

Le tableau 8 donne les résultats des titrages réalisés par la méthode des membranes filtrantes sur les 25 échantillons d'eau recueillis à l'intérieur du quadrat de 650 m de côté. L'estimation de la concentration bactérienne exprimée en nombre de micro-organismes par ml, obtenue d'après la méthode du maximum de vraisemblance, est reportée pour chaque échantillon dans la dernière colonne du tableau 8.

TABLEAU 8

TITRES BACTERIENS OBTENUS POUR LA CAMPAGNE D'ECHANTILLONNAGE  
SYSTEMATIQUE A L'INTERIEUR DU QUADRAT DE 650 m x 650 m

ECHANTILLON	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml (a)
	10 ml	1 ml	0,1 ml	
1	22	2	0	2
2	56	5	3	6
3	96	11	3	10
4	ill.	ill.	45	450
5	126	11	3	13
6	134	13	5	14
7	ill.	22	2	22
8	ill.	34	3	34
9	29	5	2	3
10	201	19	2	20
11	ill.	38	3	37
12	9	2	0	1
13	ill.	87	10	88
14	94	5	2	9
15	100		1	10
16	30	5	2	3
17	ill.	ill.	79	790
18	ill.	69	7	69
19	4	1	4	1
20	60	5	1	6
21	ill.	28	0	25
22	42	3	1	4
23	154	12	2	15
24	98	7	3	10
25	9	2	1	1

ill. : illisible

(a) : cette estimation représente le titre bactérien le plus vraisemblable arrondi à l'entier le plus proche

L'indice de dispersion de Fisher émanant des données bactériologiques du tableau 8 s'élève à 10536,87 ; valeur qui avec 24 degrés de liberté est hautement significative ( $P < 0,005$ ). Le degré de signification de ce test interprété conjointement avec les résultats figurant dans les tableaux 7 (a, b et c) font nettement ressortir l'inadéquation des données expérimentales avec le modèle de Poisson. Compte tenu des conditions de prélèvement, il faut toutefois souligner que cette incompatibilité, qui est relativement discrète pour les échantillons recueillis successivement à un même compteur, voire à un même robinet, devient beaucoup plus marquée lorsque l'échelle d'observation est étendue à la variabilité de la concentration bactérienne entre les appartements d'un même immeuble, puis à la variabilité observée dans le quartier de 650 x 650 m.

Ces considérations s'appuient également sur l'étude comparative des carrés moyens associés aux différents niveaux d'hétérogénéité (i. e., compteurs, robinets, appartements et quartier) dont l'accroissement des valeurs respectives : 1,70 ; 3,60 ; 56,56 et 439,04 semble aller de pair avec l'extension progressive du domaine d'observation dans le continuum spatio-temporel. Par ailleurs, en considérant implicitement la distribution binomiale négative en tant que modèle de substitution lorsque le modèle de Poisson est rejeté, le calcul des coefficients  $k$  correspondant à ces mêmes niveaux d'observation a donné les valeurs respectives : 373,694 ; 116,107 ; 3,854 et 0,389. L'ordre décroissant de ces estimations traduit le même phénomène, à savoir : une dispersion de plus en plus forte des résultats.

b) Etude comparative des numérations bactériennes : compteurs versus robinets

Le tableau 9 présente, pour chaque immeuble, le nombre total des colonies recensées dans les échantillons recueillis au compteur, ainsi que dans les échantillons prélevés aux robinets des divers appartements. La dernière colonne du tableau donne le résultat du test effectué, moyennant l'indice de dispersion de Fisher ( $D^2$ ), afin de comparer les concentrations bactériennes observées pour ces deux catégories d'échantillons. Enfin, la dernière ligne du tableau 9 fait état d'un test global réalisé sur la totalité des dénombrements bactériens se rapportant à l'ensemble des immeubles considérés.

**TABEAU 9** : Comparaison des concentrations bactériennes : compteurs vs robinets.

IMMEUBLE	POINT DE PRELEVEMENT	NOMBRE TOTAL DES COLONIES RECENSEES	VOLUME D'EAU FILTRÉE (ml)	CONCENTRATION BACTERIENNE par ml	D <sup>2</sup> a)
1	compteur	1149	0,5	2298	248,07 <sup>++</sup>
	robinets	1215	1	1215	
2	compteur	873	5,5	159	98,31 <sup>++</sup>
	robinets	1120	11	102	
3	compteur	947	5,5	172	17,97 <sup>++</sup>
	robinets	1293	9	144	
4	compteur	1221	5,5	222	0,25
	robinets	1962	9	218	
5	compteur	622	2,5	249	18,56 <sup>++</sup>
	robinets	333	1	333	
TEST GLOBAL	compteurs	4812	19,5	247	174,73 <sup>++</sup>
	robinets	5923	31	191	

++ : valeur significative au seuil 0,5 %

a) : D<sup>2</sup> a été confronté à la distribution de  $\chi^2$  avec 1 degré de liberté

Les données du tableau 9 font immédiatement ressortir l'existence de disparités assez nettes entre les concentrations bactériennes observées au compteur d'eau avec celles qui ont été obtenues dans les différents appartements. Sur l'ensemble des 5 immeubles considérés, la différence s'est montrée très significative pour 4 d'entre eux, soit 3 cas présentant une concentration bactérienne plus élevée pour les échantillons prélevés au compteur contre 1 cas indiquant la situation inverse. Le test prenant en considération la totalité des résultats confirme l'existence d'une densité bactérienne globalement plus élevée pour l'eau recueillie au compteur, du moins pour ce qui concerne les échantillons analysés dans le cadre de cette expérience. Cependant, la présente étude ne permet pas de savoir si l'interaction, constatée entre le facteur point de prélèvement (i. e., compteur - robinet) et l'immeuble considéré, est de nature chronique ou s'il s'agit plus simplement de phénomènes éphémères ou aléatoires dus à la forte dispersion des résultats, observée en aval du compteur.

### 3.4. DISCUSSION

L'inadéquation du modèle de Poisson, déjà perceptible pour des échantillons recueillis successivement en un même point de prélèvement pendant un temps très limité, conduit à écarter irrémédiablement ce modèle de distribution pour l'étude quantitative des populations bactériennes à l'échelle d'un réseau de distribution publique. Notons à cet effet qu'une distribution binomiale négative théorique a été ajustée avec succès aux données bactériologiques consignées dans le tableau 8, compte tenu d'un risque de première espèce préalablement fixé à 1 %. Toutefois, ces considérations ne doivent aucunement remettre en cause l'utilité et la validité du modèle de Poisson pour des volumes de liquide peu importants et bien homogénéisés (EL-SHAARAWI et al., 1981 ; FISHER et al., 1922 ; MAUL et BLOCK, 1983).

D'une manière logique et prévisible, la variabilité des titres bactériens obtenus expérimentalement s'élève au fur et à mesure que la fenêtre d'observation du continuum spatio-temporel s'élargit. Un accroissement particulièrement net de l'hétérogénéité est cependant perceptible à l'échelle du quartier ( $k = 0,389$ ) au regard de l'hétérogénéité interappartements ( $k = 3,854$ ), laquelle, il est vrai, correspond à une période d'échantillonnage plus courte. L'hétérogénéité intra-muros se caractérise par une forte variabilité interappartements ( $k = 3,854$ ) qui contraste fortement avec la plus grande similarité ( $k = 373,694$ ) des prélèvements effectués au compteur pendant le même intervalle de temps. Il est de plus intéressant de remarquer que cet effet dispersif sur les concentrations bactériennes, imputable aux ultimes canalisations d'un immeuble, ne s'accompagne d'aucune augmentation de la densité bactérienne moyenne de l'eau entre le compteur et les robinets des divers appartements de l'immeuble. Ce dernier point devrait contribuer à dissiper le doute et la défiance des responsables de la production et de la distribution des eaux potables à l'égard des mesures de qualité bactériologique réalisées quelque peu au-delà du domaine engageant leur responsabilité, mais de ce fait beaucoup plus proches de l'intérêt même des utilisateurs.

L'enseignement essentiel de cette étude préliminaire, destinée à mieux pouvoir appréhender certains aspects liés à l'hétérogénéité du peuplement bactérien, réside somme toute dans l'ampleur de la dispersion des concentrations bactériennes, qui ont été obtenues pour des échelles d'observation encore relativement modestes du continuum spatio-temporel. Ainsi, dans le cadre de la campagne d'échantillonnage menée à l'intérieur du carré de 650 m de côté pendant une période de 4 heures environ, les résultats se sont échelonnés entre 1 et 790 bactéries par millilitre ! L'étendue de ces valeurs extrêmes permet sans doute de mieux comprendre le désarroi et la perplexité que les résultats d'analyses microbiologiques peuvent quelquefois susciter auprès des personnes ayant la charge de les interpréter, quand même ces personnes disposeraient-elles de normes guides faisant péremptoirement office de "normes guillotines".

Afin de prendre le pas sur tant de variabilité, il est primordial de placer chaque situation dans son contexte probabiliste tout en étant parfaitement averti des aspects et des conséquences liés au caractère inexorable de cette variabilité.

Les courbes représentées sur la figure 3 tentent de donner une vision plus concrète de la variabilité observée aux différentes échelles d'échantillonnage étudiées dans le réseau de la ville de Metz. Ces courbes constituent en fait les enveloppes de distributions discrètes dont la moyenne a été uniformément et arbitrairement fixée à 100. La première distribution, ainsi représentée s'identifie à la loi de Poisson, tandis que les trois suivantes font ressortir l'influence du paramètre  $k$  sur la dispersion des

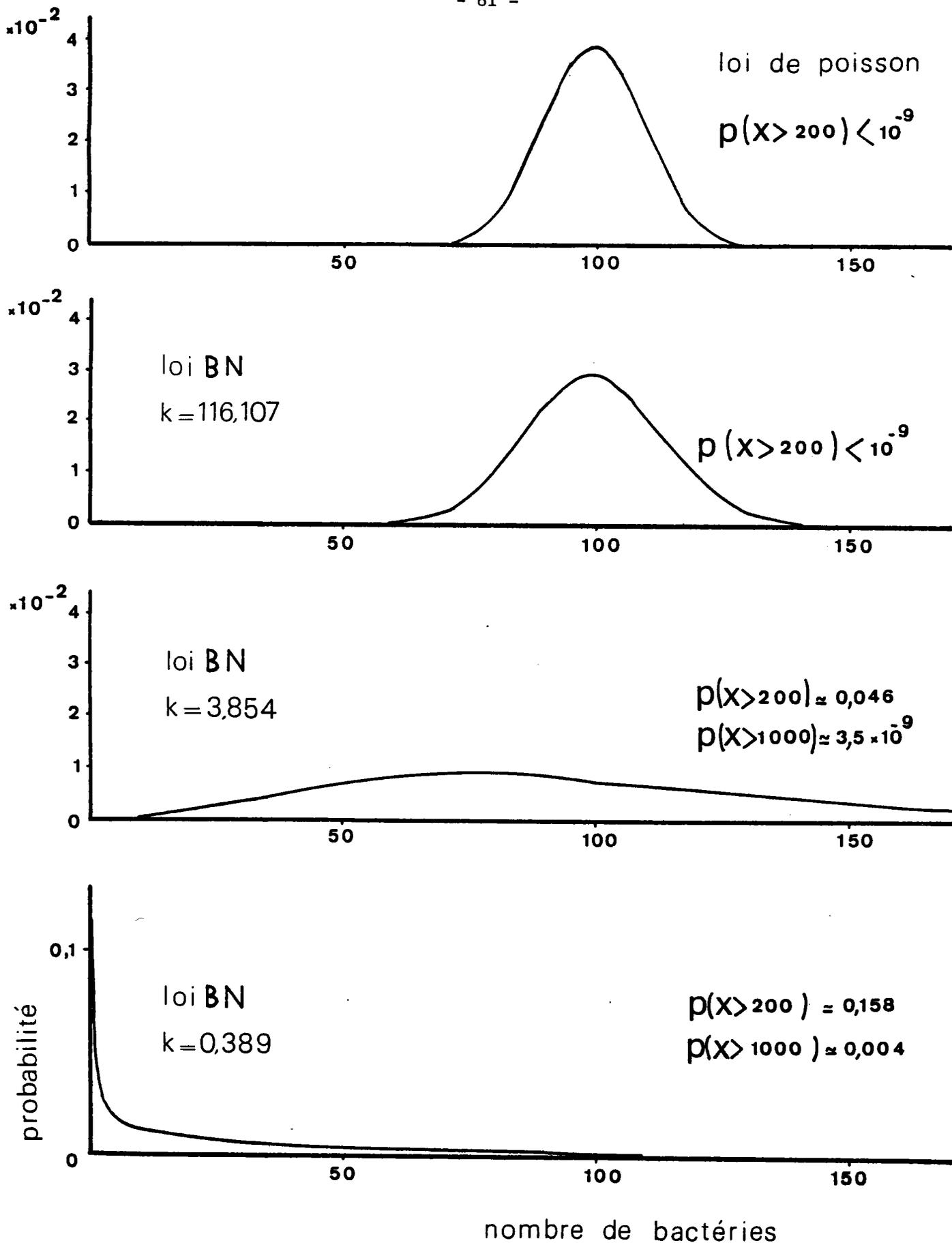


FIGURE 3 : Enveloppe de quelques distributions représentatives de la variabilité observée à différentes échelles d'échantillonnage pour une valeur commune de la moyenne arbitrairement fixée à 100.

valeurs observables en admettant implicitement la validité du modèle binomial négatif. A titre d'exemples, notons que dans ces conditions la probabilité d'observer une valeur supérieure à 200 est quasi nulle ( $P(X > 200) < 10^{-9}$ ) pour la distribution de Poisson et pour la distribution binomiale négative avec  $k = 116,107$  alors que cette probabilité avoisine 0,05 pour  $k = 3,854$  et dépasse même 0,15 lorsque  $k = 0,389$ . Pour cette dernière valeur de  $k$ , le calcul montre que même les titres bactériens supérieurs à 1000 font partie de l'éventail des résultats "observables" ( $P(X > 1000) \approx 0,004$ ).

A défaut de vouloir adopter ce point de vue probabiliste, le chercheur engagé dans l'étude d'un réseau de distribution publique s'expose au risque de travaux infructueux pouvant, en citant un mot de LEGENDRE et LEGENDRE (1979), générer plus de papier noirci que de lumière sur le problème.

**DEUXIEME PARTIE**

ETUDE EXPERIMENTALE SUR LE RESEAU  
DE DISTRIBUTION PUBLIQUE DE METZ

Les campagnes d'échantillonnage conçues dans le but de contrôler ou d'étudier la concentration bactérienne de l'eau d'un réseau de distribution publique consistent généralement à planifier la collecte d'échantillons d'eau en divers points de prélèvement sur une période déterminée. L'efficacité d'un programme de prélèvement donné est très étroitement liée à la variabilité spatiale et temporelle de l'abondance des populations bactériennes dans les milieux hydriques échantillonnés. Or, les connaissances en matière d'écologie microbienne dans les réseaux de distribution sont cependant encore très modestes. En particulier, les mécanismes qui régissent les phénomènes de contamination et de propagation bactérienne de même que les aspects quantitatifs des variations dans le temps et dans l'espace de la densité microbienne restent mal compris.

La démarche habituellement suivie par les microbiologistes pour appréhender l'hétérogénéité spatio-temporelle se résume souvent par une seule opération visant à tenter l'ajustement des données expérimentales par une loi de distribution théorique appropriée, dont la nature est jugée suffisante pour caractériser la dispersion des micro-organismes. A ce titre, il faut rappeler que la loi binomiale négative occupe une place de choix en microbiologie puisque ce modèle de distribution a souvent pu être ajusté avec succès à diverses séries de données expérimentales recueillies dans le temps ou/et dans l'espace (CHRISTIAN et PIPES, 1983 ; EL-SHAARAWI et al., 1981 ; EL-SHAARAWI et al., 1985 ; MUENZ, 1978 ; PIPES et al., 1977 ; THOMAS, 1955). Les paramètres de la loi ainsi définie, résument alors certaines caractéristiques des données et sont également pris en considération pour infléchir les modalités des programmes d'échantillonnage ultérieurs (EL-SHAARAWI et al., 1985 ; ESTERBY, 1982 ; PIPES et CHRISTIAN, 1982).

La connaissance explicite de ces lois de distribution, qui constitue avec raison une première façon d'aborder les problèmes, est toutefois insuffisante pour permettre l'identification et la description de la structure du peuplement bactérien d'un milieu hydrique. Aussi, au terme de l'étude préalable consacrée à l'observation de la structure hiérarchisée de l'hétérogénéité des populations bactériennes à petite et moyenne échelle, il convient maintenant d'entreprendre des investigations plus ambitieuses dont l'objectif sera d'explorer la structure spatiale et temporelle de l'hétérogénéité caractérisant l'ensemble d'un réseau de distribution publique. L'aboutissement de cette recherche menée à grande échelle doit permettre :

- a) de définir les modalités d'une stratégie d'échantillonnage adaptée au contrôle efficace et rationnel de la qualité bactériologique de l'eau distribuée, en tirant profit de la configuration spatio-temporelle de l'hétérogénéité mise en évidence
- b) de mieux cerner la nature et l'influence des facteurs liés à l'accroissement des concentrations bactériennes
- c) d'identifier les éventuels secteurs à problèmes qui nécessitent, à l'évidence, l'application de mesures spécifiques appropriées.

## **CHAPITRE I**

ETUDE DESCRIPTIVE DU RESEAU DE  
DISTRIBUTION DE LA VILLE DE METZ

Guidés par le souci de clarté de l'exposé, il nous est apparu utile de faire précéder l'étude expérimentale proprement dite par une présentation sommaire du contexte général ayant servi de cadre à nos investigations. Aussi, la description du réseau messin développée ci-après tentera de dégager les aspects les plus importants pour la suite de ce travail.

### 1.1. CARACTERISTIQUES GENERALES DU RESEAU ET CAPACITE DE PRODUCTION

La distribution d'eau potable de la ville de Metz, ainsi que de 26 communes voisines a été confiée à la Société Mosellane des Eaux qui assure la gestion d'un réseau comprenant environ 645 kilomètres de canalisations. La population desservie par ce réseau s'élève approximativement à 200000 habitants.

L'eau provient pour une large part d'un affluent de la Moselle, le Rupt de Mad, dont le débit, nettement insuffisant en période d'étiage, est régularisé par une retenue d'eau (lac de Madine) située à 60 kilomètres de Metz. Captée à Arnaville, à 15 kilomètres en amont de Metz, l'eau est préchlorée avant d'être définitivement traitée dans l'usine de Moulins qui représente une capacité journalière de production de l'ordre de 90 000 m<sup>3</sup> sur un potentiel de 120 000 m<sup>3</sup>/j pour l'ensemble des usines du service.

Les autres moyens de production comprennent quatre usines, situées à Hauconcourt (10 000 m<sup>3</sup>/j), Saint-Rémy (1 200 m<sup>3</sup>/j), Saint-Eloy (8 000 m<sup>3</sup>/j) et l'usine Sud (8 000 m<sup>3</sup>/j), captant toutes l'eau de la nappe alluviale de la Moselle. L'eau produite par l'usine Sud est refoulée vers l'usine de Moulins, toute proche, où elle est mélangée avec l'eau issue de cette dernière. Les sources de Gorze (3 000 m<sup>3</sup>/j en étiage, 10 000 m<sup>3</sup>/j en période de hautes eaux) et les sources de Lorry (400 m<sup>3</sup>/j en étiage, 800 m<sup>3</sup>/j en période de hautes eaux) viennent compléter des ressources largement excédentaires. La localisation de ces différentes unités de production est indiquée sur la figure 4 qui donne une représentation schématique de l'ensemble du réseau.

### 1.2. FILIERES DE TRAITEMENT DE L'EAU BRUTE

L'usine de Moulins comprend les équipements et les procédés habituellement mis en oeuvre dans la filière du traitement d'une eau de surface, soit : une régulation du niveau de l'eau à l'entrée de l'usine suivie d'une coagulation - filtration - ozonation - chloration. La coagulation a pour but de permettre la floculation des matières colloïdales par l'intermédiaire de réactifs injectés dans l'eau en amont des décanteurs. Sous l'effet de la pesanteur, le floc obtenu est maintenu dans la partie basse du décanteur, alors que l'eau clarifiée est recueillie en partie haute. Les dernières particules en suspension dans l'eau sont arrêtées par deux batteries de filtration comprenant chacune huit filtres à sable d'une hauteur de 80 cm pour une superficie unitaire de filtration de 52 m<sup>2</sup>. Le lavage se fait par courant d'eau inverse et soufflage d'air. L'eau filtrée est ensuite soumise à un traitement biologique par une ozonation dans deux tours de contact de 300 m<sup>3</sup> chacune avant de subir une chloration en vue de maintenir une concentration de 0,6 mg de chlore résiduel libre par litre d'eau traitée. Cette eau est finalement stockée dans deux bâches de stockage de 1 500 m<sup>3</sup> chacune.

Les autres unités de production sont alimentées par l'eau d'exhaure de la nappe alluviale dont le champ captant s'étend sur 60 kilomètres de longueur et 2 à 3 kilomètres de largeur. L'exhaure est assurée par des puits d'une profondeur variant de 6 à 10 mètres, répartis sur 20 kilomètres environ. Dans ces usines, le traitement de l'eau, qui est moins

uckange

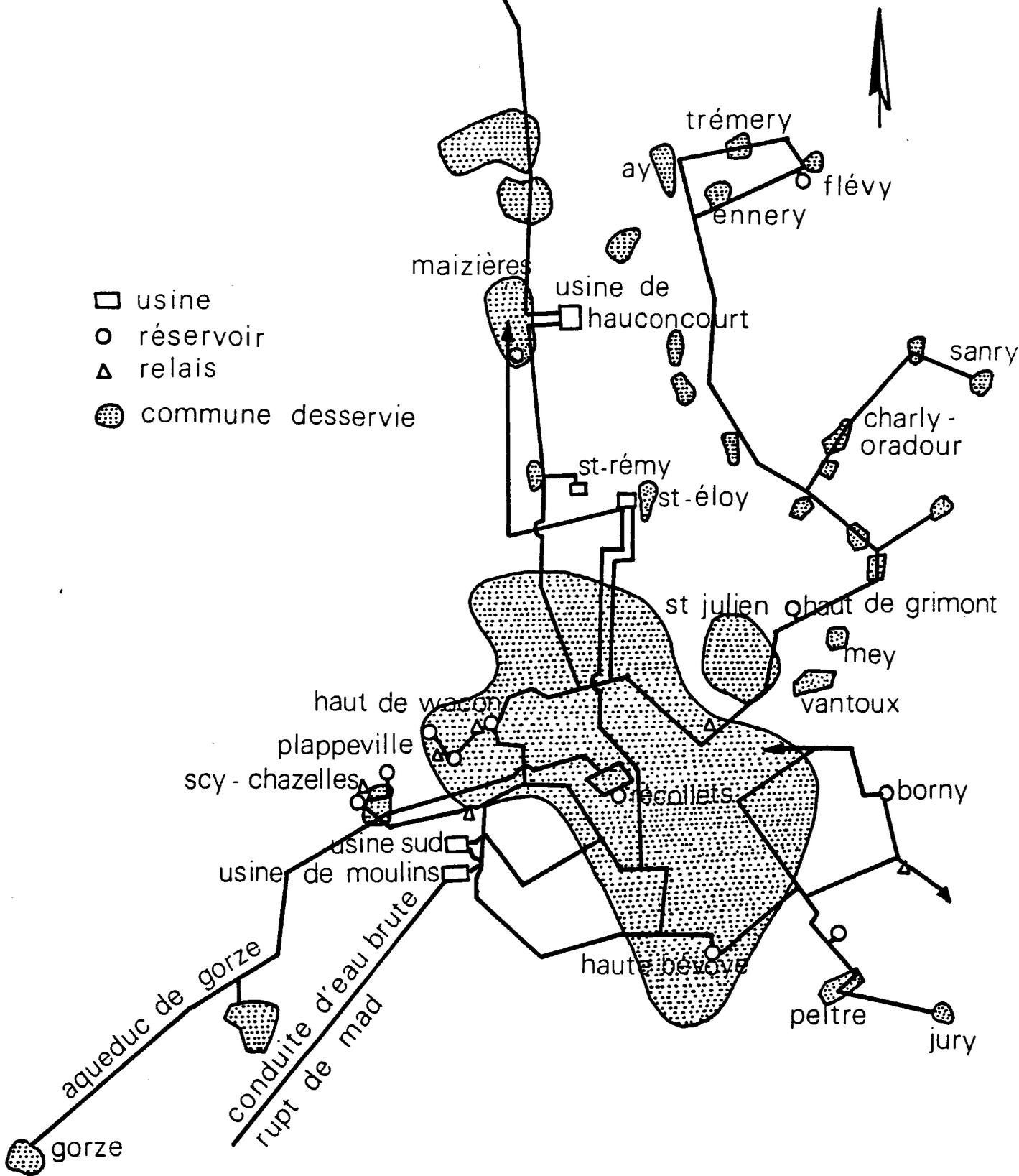


FIGURE 4 : Représentation schématique du réseau messin (adapté d'après Société Mosellane des Eaux).

sophistiqué que celui mis en oeuvre à l'usine de Moulins, se limite à une chloration suivie éventuellement d'une correction de pH ou d'une aération afin d'éliminer le gaz carbonique en excès.

### 1.3. CAPACITE DE STOCKAGE ET DISTRIBUTION

Douze réservoirs, implantés au sol ou semi-enterrés, à l'exception des réservoirs de Maizières (1 000 m<sup>3</sup>) et de Flévy (200 m<sup>3</sup>) qui sont sur tour, représentent une capacité totale de stockage de 86 000 m<sup>3</sup>.

Les cinq principaux réservoirs sont :

- HAUT DE WACON	30 000 m <sup>3</sup>	au nord-ouest
- HAUTE BEVOYE	20 000 m <sup>3</sup>	au sud-est
- BORNY	20 000 m <sup>3</sup>	à l'est
- LES RECOLLETS	5 500 m <sup>3</sup>	au centre ville
- GRIMONT	3 000 m <sup>3</sup>	au nord-est

La distribution de l'eau est assurée par un réseau de canalisations en fonte grise ou en fonte ductile dont le diamètre varie de 60 à 1100 millimètres. En dérivation sur ce réseau, les branchements d'abonnés ont été réalisés en polyéthylène haute densité, en fonte ou encore en acier galvanisé.

La quantité d'eau distribuée en moyenne est de 55 000 m<sup>3</sup> par jour, mais elle atteint 65 000 m<sup>3</sup>/j en période de forte demande. La distribution se fait principalement sur 5 étages correspondant à un niveau de pression défini par un réservoir (exceptionnellement deux).

- L'étage 1 (niveau 191 NGF) est alimenté par les sources de Gorze dont l'eau aboutit au réservoir des Récollets. Le volume d'eau distribué sur cet étage, correspondant au centre ville, est de 1 000 m<sup>3</sup>/j en moyenne.

- L'étage 2 (niveau 225 NGF) constitue la majeure partie de Metz. Cet étage, associé aux réservoirs du Haut de Wacon et de Haute Bévoye, assure une distribution journalière de 33 000 m<sup>3</sup> d'eau. A l'exception des sources de Gorze, la totalité de l'eau produite aboutit directement sur cet étage de distribution qui dessert également plusieurs communes entre Maizières et Uckange.

- L'étage 3 (niveau 242 NGF), dont le niveau de pression est défini par le réservoir de Borny, distribue quotidiennement 11 000 m<sup>3</sup> d'eau aux quartiers est de Metz (Borny, Queuleu, Vallières), ainsi qu'aux communes de Peltre, Jury, Saint-Julien (partiellement) et Vantoux.

- L'étage 4 (niveau 264 NGF) correspond à la rive droite de la Moselle. Cet étage alimente notamment les communes de Saint-Julien (l'autre partie), Mey, Charly-Oradour, Sanry-les-Vigy, Ay-sur-Moselle, Ennery et Tremery. 2 500 m<sup>3</sup>/j y sont distribués par refoulement à partir du réservoir de Grimont alimenté par un relais situé à Saint-Julien et prélevant l'eau directement sur l'étage 2.

- L'étage 5 (niveau 270 NGF) qui est dépourvu de réservoir propre est greffé sur le réservoir de Borny par l'intermédiaire d'un surpresseur. Ce petit étage (900 m<sup>3</sup>/j) alimente principalement la zone industrielle de Borny.

D'autres petits étages de distribution assurent l'approvisionnement en eau des communes de Scy-Chazelles et Plappeville sur le Mont-St-Quentin par deux chaînes de relais.

## CHAPITRE II

### STRUCTURE SPATIO-TEMPORELLE DE L'HETEROGENEITE BACTERIENNE

A notre connaissance, la seule étude d'envergure entreprise à ce jour, dans le domaine de la reconnaissance d'une structure d'abondance et l'utilisation d'informations obtenues à la suite d'un échantillonnage intensif afin de pouvoir définir ultérieurement des programmes de contrôle sur la base des données recueillies, a été réalisée aux U.S.A. par PIPES et CHRISTIAN (1982) dans neuf réseaux de distribution publique situés à proximité de Philadelphie. Cette étude, qui fait état de bactéries coliformes en considération de la spécification des normes de salubrité actuellement en vigueur aux U.S.A. (EPA, 1976), a montré l'adéquation des distributions log-normale et binomiale négative aux séries expérimentales obtenues dans le temps et dans l'espace (CHRISTIAN et PIPES, 1983). La diminution de la probabilité associée au pouvoir de détection d'une contamination de l'eau avec l'accroissement de l'hétérogénéité, amène les auteurs à mettre l'accent sur les paramètres de dispersion (i. e., écart-type géométrique et paramètre k) des modèles mathématiques retenus, afin de déterminer la fréquence des analyses à effectuer dans le cadre d'un programme de surveillance.

Dans une approche plus pragmatique, PIPES et CHRISTIAN ont ensuite tenté de mettre en évidence une structure spatiale de l'hétérogénéité bactérienne essentiellement fondée sur les caractéristiques hydrodynamiques des masses d'eau dont l'écoulement est subordonné à l'agencement des canalisations du réseau.

Cette démarche a conduit les auteurs à formuler plusieurs hypothèses concernant les mécanismes qui régissent les phénomènes de contamination et de propagation des bactéries coliformes dans les réseaux de distribution publique. En particulier, les investigateurs ont voulu savoir si les risques de contamination avaient tendance à être plus élevés à la périphérie de l'usine de traitement des eaux que dans les quartiers localisés à proximité des unités de production. Aucune des hypothèses examinées n'ayant pu être corroborée par les données recueillies, PIPES et CHRISTIAN ont été contraints d'abandonner leur projet de définition d'un programme d'échantillonnage basé sur l'exploitation mathématique d'une configuration structurée de l'hétérogénéité.

Les causes de cet insuccès, en dépit d'une problématique solide et judicieuse, tiennent vraisemblablement à deux raisons prépondérantes : il s'agit d'une part, comme PIPES et CHRISTIAN le font d'ailleurs eux-mêmes remarquer, de la taille trop modeste des réseaux étudiés, laquelle est probablement incompatible avec la formation d'un gradient stable et perceptible de l'abondance bactérienne, et d'autre part, de la "population cible", en l'occurrence les bactéries coliformes, dont la faible présence habituellement observée dans l'eau distribuée a pour effet de diminuer la puissance, la sensibilité, voire même le bien-fondé de certaines analyses statistiques.

Ces considérations nous ont amenés à reconsidérer l'ensemble du problème soulevé par les travaux de PIPES et CHRISTIAN, dans des circonstances plus favorables, à savoir : l'étude de la structure d'abondance des bactéries hétérotrophes dans un réseau de distribution s'étendant sur une grande échelle. Le choix des bactéries hétérotrophes est justifié par l'utilisation fréquente et même préconisée de ces numérations à titre de test complémentaire pour l'évaluation de la qualité bactériologique d'une eau (COLWELL et al., 1978; McFETERS et al., 1978 ; MEANS et al., 1981).

L'analyse statistique développée dans la suite de ce travail a pour objectif essentiel de donner une image concrète et précise de la variabilité spatiale et temporelle caractérisant la structure d'abondance des populations bactériennes dans un réseau de distribution publique. A cet effet, quatre types de configuration avaient initialement été formulés, pour décrire

la structure de l'hétérogénéité spatiale, selon que le réseau se présente comme :

- un unique système homogène
- un ensemble composé de plusieurs systèmes homogènes
- un unique système hétérogène
- un ensemble composé de plusieurs systèmes hétérogènes.

Par système homogène (respectivement hétérogène), nous entendons un système dans lequel le nombre de bactéries contenues dans des échantillons de même volume, est distribué suivant une loi de Poisson (respectivement une loi binomiale négative).

L'étude expérimentale, réalisée au chapitre 3 de la première partie, consacrée à l'hétérogénéité des populations bactériennes à petite et moyenne échelle, permet d'emblée d'écarter les deux premiers schémas cités qui évoquent avant tout le modèle de Poisson. Par conséquent, nous envisagerons successivement la possibilité de concevoir le réseau comme un unique système hétérogène, puis nous examinerons l'éventualité d'une configuration structurée de l'hétérogénéité spatiale faisant apparaître le réseau sous l'aspect d'un ensemble composite formé de plusieurs systèmes hétérogènes.

La mise en évidence d'une structure d'abondance du peuplement bactérien par des procédés objectifs de classification et de groupement des données en zones de densité bactérienne est largement tributaire du pouvoir de résolution de la méthode d'échantillonnage. La finesse de l'exploration de l'espace-temps est déterminée par la précision et la reproductibilité des méthodes analytiques liées au traitement des échantillons, mais aussi par la densité des observations effectuées dans le continuum spatio-temporel. La résolution spatio-temporelle est, de plus, soumise à l'interaction entre les propriétés de l'échantillonnage et celles du milieu. En effet, la collecte des échantillons, lors d'une campagne de prélèvements, étant par force effectuée d'une manière discontinue dans l'espace et pendant un certain intervalle de temps, les structures existant aux échelles inférieures restent indiscernables. En tout état de cause, il est vraisemblable qu'au-dessous d'un certain seuil d'étendue géographique, les perturbations résultant de la mouvance du milieu aquatique ne permettent plus la reconnaissance d'une structure spatiale d'abondance. Ceci contraint les investigateurs à limiter l'exploitation des données par la simple description des distributions statistiques, en deçà d'une certaine échelle d'observation. Par ailleurs, l'intervalle de temps consacré à l'échantillonnage pose le problème de la distinction des différentes directions spatio-temporelles en créant une liaison entre l'espace et le temps. Ainsi la mouvance du milieu d'une part, qui ne peut être que subie, et l'ensemble des sources de variabilité imputables aux conditions de prélèvement, de conservation et d'analyse des échantillons d'autre part, sont de nature à "brouiller" l'aspect et le caractère instantané de l'image du réseau observée à la suite d'une campagne d'échantillonnage systématique. Conscients du rôle négatif de certaines variables perturbatrices, il nous apparaît primordial d'étudier la persistance et le degré de reproductibilité des images obtenues au cours de plusieurs campagnes de prélèvements.

Enfin, la signification de la structure spatiale ainsi dégagée sera interprétée en considération de quelques caractéristiques physico-chimiques et hydrauliques du réseau. (i.e., disposition des canalisations et localisation des réservoirs, temps de séjour, chlore résiduel, température de l'eau). La variation temporelle des concentrations bactériennes fera l'objet d'une étude globale intéressant l'ensemble du système de distribution, mais également d'un examen plus minutieux rendant compte de l'évolution de la concentration bactérienne moyenne au sein de chaque zone de densité, étudiée séparément.

## 2.1. MATERIEL ET METHODES

### 2.1.1. Localisation des points de prélèvement et collecte des échantillons d'eau

Le champ d'étude de ce travail a été circonscrit à la partie du réseau de distribution messin desservie exclusivement par de l'eau en provenance de l'usine de traitement de Moulins. Les limites de la zone ainsi définie ont été déterminées, avec une marge de sécurité jugée suffisante, de manière à écarter toutes les sections du réseau susceptibles de contenir, même partiellement, de l'eau originaire des autres unités de production. La partie enclavée du système de distribution représentée par l'aire d'échantillonnage s'étend sur une superficie couvrant approximativement la moitié de l'agglomération messine. Ce domaine comprend notamment les quartiers nord-ouest (Ban-St-Martin, Plappeville), toute la partie sud de Metz (Queuleu, Magny, Sablon, Borny, Vallières, Plantières) et quelques communes avoisinantes (Scy-Chazelles, Grigy, Peltre, Jury).

Une représentation schématique des principales caractéristiques hydrauliques du réseau est donnée sur la figure 5. Cette figure indique plus particulièrement la disposition des canalisations principales et la situation des unités de stockage en précisant le temps de rétention approximatif (TR) de l'eau dans chaque réservoir de même qu'une évaluation du temps de transit moyen de l'eau entre sa sortie de l'usine de traitement de Moulins et son arrivée dans les différents réservoirs (M→R). La localisation des points de prélèvement, déterminés selon un plan d'échantillonnage systématique, est également mentionnée sur la figure 5. Ces points, qui représentent approximativement les centres des carrés de 440 m de côté obtenus par un quadrillage préalable de l'ensemble du réseau, ont été définis avec une tolérance de quelques dizaines de mètres.

Les échantillons d'eau ont été recueillis au cours d'une période de 3 heures environ en respectant scrupuleusement les principes et les consignes exposés au paragraphe 3.1.1. de la première partie. Les échantillons ont ensuite été ramenés au laboratoire à la température ambiante pour être soumis à l'analyse dans un délai n'excédant jamais 6 heures à compter du début de chaque campagne d'échantillonnage.

### 2.1.2. Analyse microbiologique et physico-chimique des échantillons

Des dilutions décimales des échantillons d'eau ont été réalisées dans de l'eau physiologique stérile contenant 0,8 % de NaCl. Des volumes variant de 10 ml à 0,01 ml de l'échantillon originel ont été filtrés sur membrane stérile d'une porosité de 0,45  $\mu\text{m}$  (Millipore HAWG 047). Afin de permettre une meilleure répartition sur la membrane de filtration, l'aliquote a, le cas échéant, été préalablement transférée dans 9 ml d'eau physiologique stérile. Les membranes ont ensuite été déposées sur milieu R-2A dont la composition chimique est donnée au paragraphe 2.2.1. de la première partie. Toutes les colonies apparues sur les membranes ont été dénombrées après une période d'incubation de 72 heures à la température de 20°C. Enfin, les résultats ont été exprimés par le nombre le plus vraisemblable de bactéries hétérotrophes contenues dans un millilitre d'échantillon. Ce nombre a été calculé suivant la méthode du maximum de vraisemblance en exploitant les résultats des comptages effectués à partir de plusieurs niveaux de dilution successifs (MAUL et al., 1981).

L'analyse physico-chimique des échantillons a été limitée à la détermination de la température de l'eau recueillie et, pour certains cas seulement, à la mesure des teneurs en chlore résiduel libre et total, effectuée par dosage volumétrique selon le test D.P.D. (RODIER, 1984).

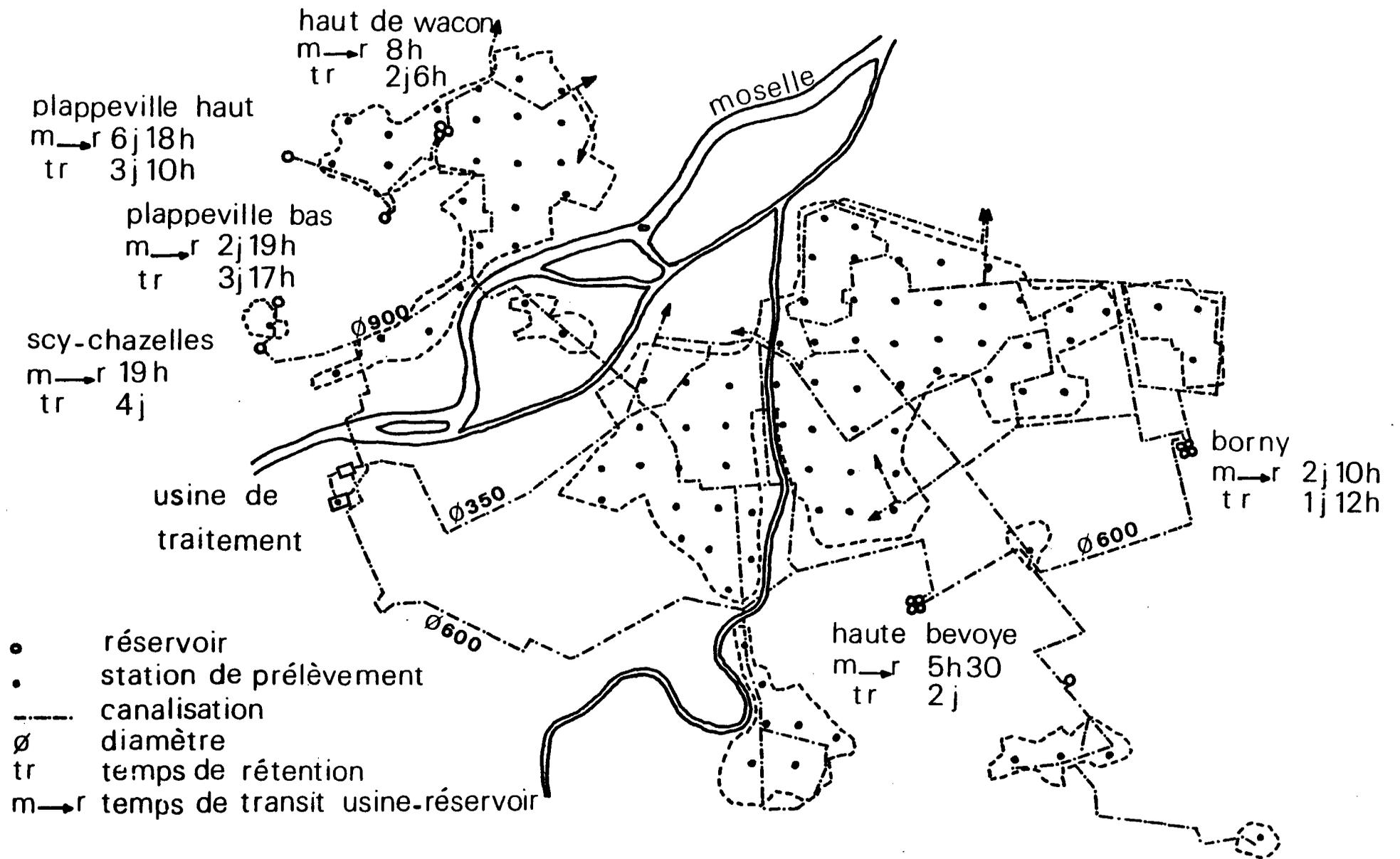


FIGURE 5 : Représentation schématique des principales caractéristiques hydrauliques du réseau de distribution étudié.

### 2.1.3. Analyse statistique des résultats

#### 2.1.3.1. Ajustement d'un modèle de distribution d'abondance aux données expérimentales

Rappelons que si le nombre des bactéries contenues dans les échantillons analysés obéit à une loi binomiale négative, la probabilité d'observer  $r$  micro-organismes dans un échantillon quelconque est donné par la formule (11) du paragraphe 1.2.1. de la première partie. Par suite, en adoptant le modèle binomial négatif comme modèle de référence, l'ajustement d'une distribution binomiale négative théorique à une série de données expérimentales indépendantes a été effectué conformément au test d'adéquation de  $\chi^2$  décrit au paragraphe 1.2.2. Le calcul des effectifs théoriques associés à chaque classe de valeurs, ainsi que les estimations du paramètre  $k$  de la loi binomiale négative par la méthode du maximum de vraisemblance ont été réalisés sur calculateur électronique au moyen des programmes développés respectivement dans les annexes I et II.

#### 2.1.3.2. Analyses de corrélation

Le coefficient de corrélation de rang de Kendall ( $\tau$ ) a été utilisé afin d'étudier a) la reproductibilité de la structure d'abondance entre les différentes campagnes de prélèvements et b) l'association entre les titres bactériens observés avec une variable physico-chimique explicative telle que le temps de séjour ou la quantité de chlore résiduel. Ce coefficient de corrélation non-paramétrique, dont le mode de calcul a été décrit par LEGENDRE et LEGENDRE (1979), est particulièrement bien adapté pour la mise en évidence d'une relation monotone entre variable explicative et variable expliquée. L'absence de liaison entre les deux variables considérées est testée à l'aide de la statistique

$$z = |\tau| \sqrt{\frac{9n(n-1)}{2(2n+5)}} - \sqrt{\frac{18}{n(n-1)(2n+5)}} \quad (45)$$

dans laquelle  $n$  désigne le nombre d'observations. Kendall (1948) a démontré que sous l'hypothèse nulle (i. e., absence de liaison monotone entre les variables) la valeur de  $z$  est distribuée suivant la loi normale centrée réduite lorsque  $n$  est supérieur à 10.

De manière plus classique, l'association entre la température moyenne de l'eau et le niveau moyen de la concentration bactérienne dans le réseau a été étudiée par une analyse de régression réalisée selon la méthode des moindres carrés. Le modèle mathématique retenu fait état d'une relation linéaire entre la moyenne géométrique de la densité bactérienne et la température moyenne de l'eau du réseau. A cet effet, chaque campagne de prélèvements a été représentée par un point affecté d'un coefficient de pondération égal au nombre des échantillons analysés au cours de ladite campagne. L'abscisse de chacun de ces points a été calculée en effectuant la moyenne arithmétique des températures relevées tandis que leur ordonnée a été obtenue par la moyenne géométrique des titres bactériens dont une estimation non biaisée est donnée par la médiane de la distribution empirique après transformation logarithmique. Après l'estimation des paramètres du modèle mathématique, la vérification de la validité de l'hypothèse affirmant l'existence d'une relation linéaire entre les deux variables étudiées a été effectuée en comparant le coefficient de corrélation de Pearson ( $r$ ) à la valeur zéro au moyen de la loi de Student (C.E.A., 1978).

### 2.1.3.3. Méthode de groupement

Une méthode de groupement séquentielle, non-hiérarchique, basée sur l'affectation de chaque point de prélèvement au groupe dont le centroïde est le plus proche a été mise en oeuvre pour effectuer une partition de l'ensemble des points de prélèvement considérés. Le programme de groupement, rédigé en langage APL, a été exécuté à l'aide d'un ordinateur CDC 171. Les parties du réseau délimitées après représentation cartographique des sous-ensembles composant la partition obtenue seront appelées "zones de densité bactérienne" ou plus simplement "zones". Les développements théoriques de la méthode de groupement ont été exposés à la section 1.3.1. de la première partie.

### 2.1.3.4. Analyse en composantes principales

L'analyse en composantes principales, qui permet la représentation de la dispersion des objets (i. e., stations de prélèvement) dans un espace de dimension réduite respectant au mieux leurs positions relatives dans l'espace multidimensionnel des données, a été utilisée en complément de la méthode de groupement précédente pour la recherche d'une structure spatiale de l'abondance bactérienne. Une description sommaire visant à dégager les principales étapes de la mise en application de cette technique, est donnée au paragraphe 1.3.2. de la première partie.

## 2.2. PROTOCOLE EXPERIMENTAL

Six campagnes de prélèvements ont été réalisées pendant la période allant de décembre 1983 à juin 1984 sur toute la partie du réseau messin représentée sur la figure 5. Compte tenu des disponibilités du laboratoire et de l'objectif fixé, les dates suivantes ont été retenues : 13 décembre, 10 janvier, 21 février, 3 avril, 15 mai et 26 juin. Au cours de chacune de ces campagnes de prélèvements, 8 échantillonneurs ont procédé simultanément à la collecte d'échantillons de 500 ml d'eau de distribution pendant une durée n'excédant pas 3 heures. Les échantillons ont invariablement été recueillis en 102 points de prélèvement, numérotés de 0 à 101, répartis sur l'ensemble du réseau. La localisation et l'identification des points de prélèvement est présentée sur la figure 6.

Au laboratoire, après homogénéisation par 10 retournements des flacons de 500 ml, des parties aliquotes représentant 10 ml, 1 ml, 0,1 ml et 0,01 ml de la suspension bactérienne originelle ont été filtrées sur membrane. La période s'écoulant entre le moment du prélèvement de l'échantillon au robinet et la fin des manipulations au laboratoire n'a jamais dépassé 6 heures. Après lecture et calcul des résultats comme décrit au paragraphe 2.1.2., la totalité des quelque 600 données bactériologiques recueillies pour l'ensemble des 6 campagnes de prélèvements a été soumise au programme de groupement.

Au cours de la journée du 26 juin seulement, les points de prélèvement désignés par les numéros 0, 8, 10, 18, 21, 27, 30, 35, 39, 45, 48, 58, 63, 68, 70, 74, 76, 83, 88, 95 et 100 (cf. Figure 6), ont fait l'objet de mesures supplémentaires destinées à déterminer la concentration en chlore résiduel libre et total de l'eau de distribution. Ces dosages, réalisés en parallèle avec l'étude microbiologique, ont été effectués immédiatement sur le lieu même de chaque prélèvement.

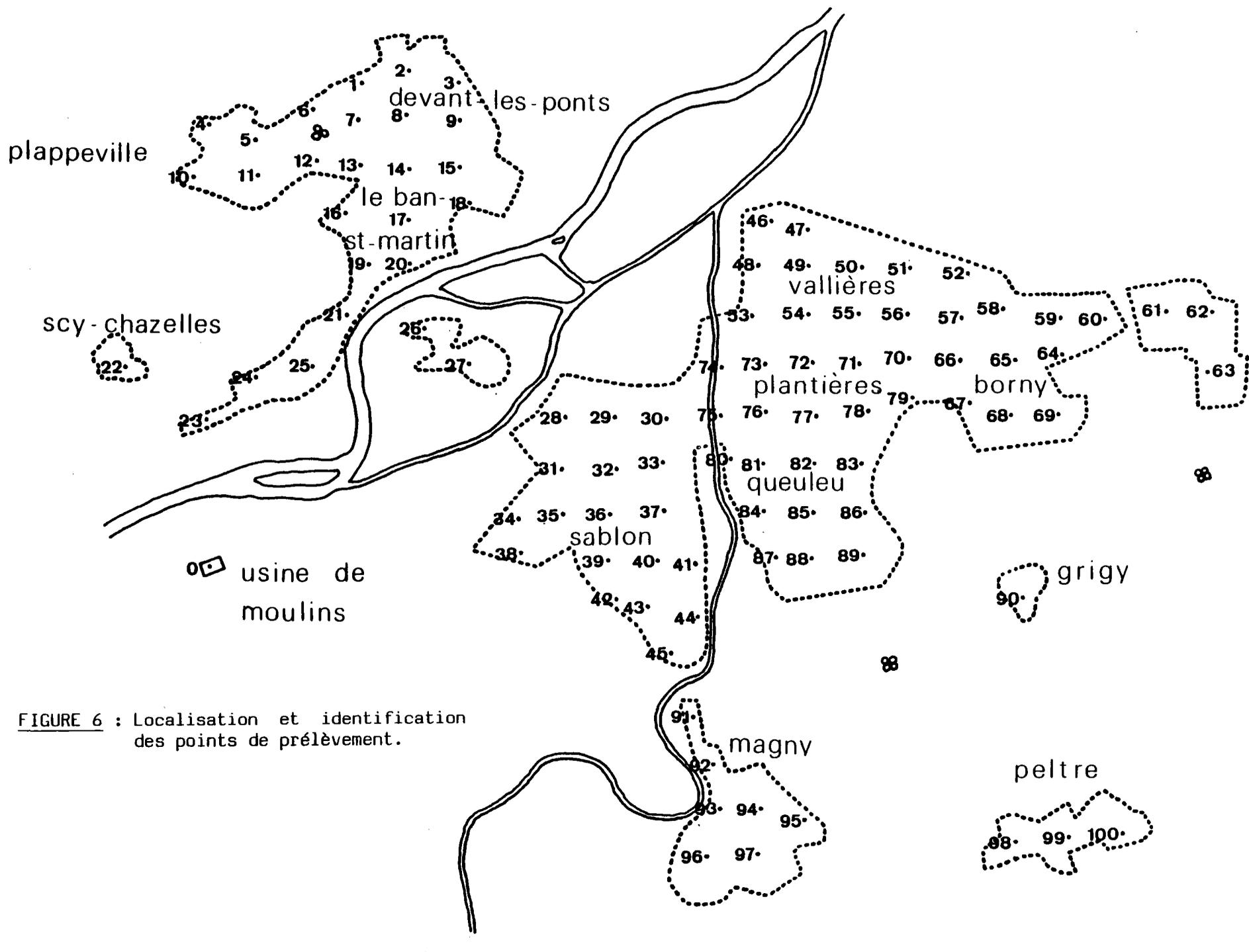


FIGURE 6 : Localisation et identification des points de prélèvement.

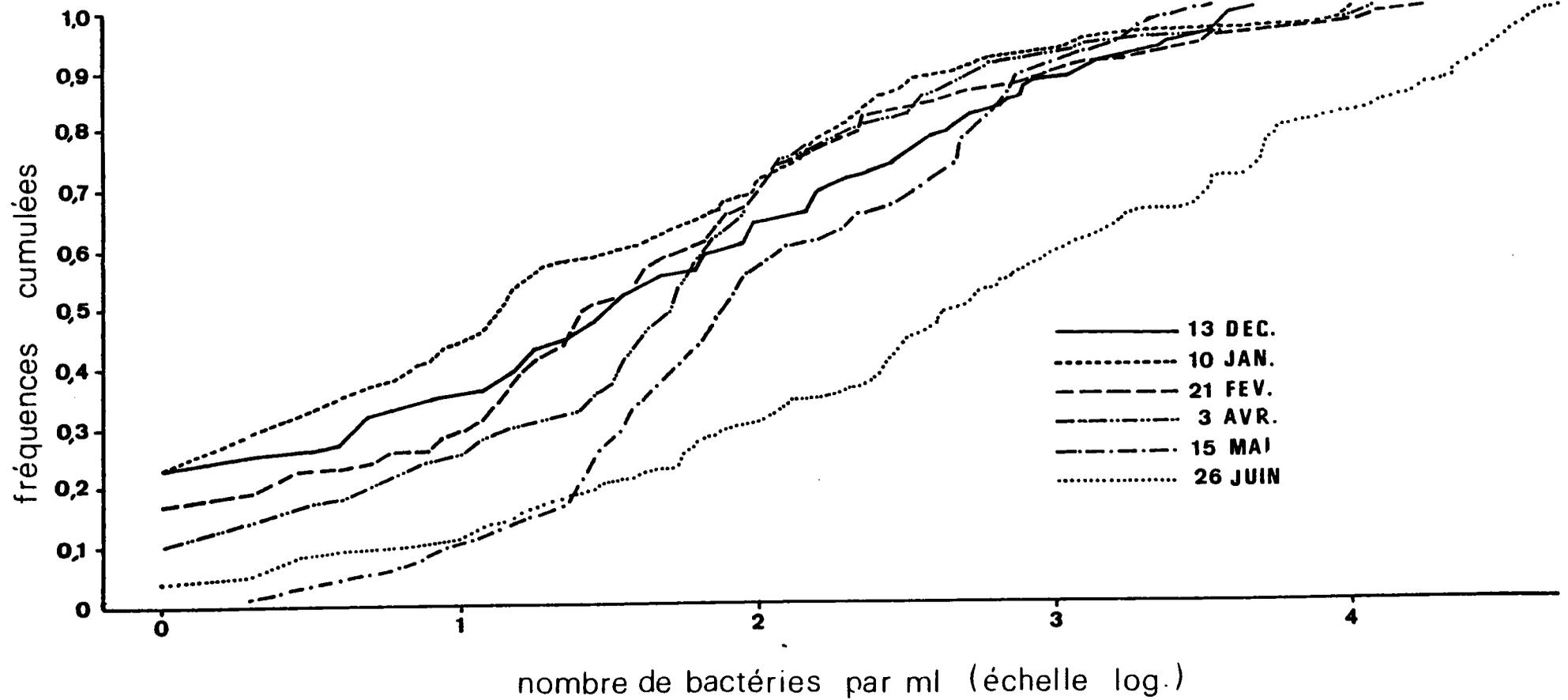
## 2.3. RESULTATS

### 2.3.1. Présentation des données bactériologiques

Les tableaux I à VI, figurant en Annexe V, donnent les résultats des six campagnes de quantification des populations bactériennes réalisées sur le réseau de distribution de la ville de Metz. Pour chaque échantillon d'eau, identifiable par le numéro du point de prélèvement auquel il se rapporte, les données consignées dans ces tableaux font état a) de la température de l'eau au moment du prélèvement, b) des résultats bruts des comptages bactériens obtenus pour trois niveaux de dilution successifs et c) d'une estimation de la concentration bactérienne par millilitre d'échantillon, calculée selon la méthode du maximum de vraisemblance.

La figure 7 tente de donner une vision globale et synthétique de l'ensemble des données bactériologiques au moyen des distributions empiriques cumulatives croissantes caractérisant les différentes campagnes d'échantillonnage. L'examen de la figure 7 fait clairement ressortir les courbes associées aux campagnes du 15 mai et surtout du 26 juin dont les profils sont nettement décalés vers le bas. Cette analyse graphique sommaire doit être interprétée par une augmentation générale du niveau de la concentration bactérienne observée au cours de ces deux dernières campagnes par rapport aux autres campagnes de prélèvements.

L'éventualité d'un réseau de distribution pouvant être décrit comme un unique système hétérogène a été testée par la tentative d'ajustement d'une loi binomiale négative à chacune des séries de données expérimentales. Le tableau 10 présente : le nombre d'échantillons analysés, la moyenne arithmétique des titres bactériens observés  $\bar{r} = \hat{\lambda}$ , les estimations  $\hat{p}$  et  $\hat{k}$  des paramètres de la loi binomiale négative théorique, ainsi que les résultats des tests d'adéquation de  $\chi^2$  destinés à évaluer la qualité de l'ajustement pour chacune des six séries expérimentales. A l'exclusion des valeurs de  $\chi^2$  associées à la première et à la dernière campagne de prélèvements (i. e., 13 décembre et 26 juin) les autres valeurs se sont toutes montrées significatives au seuil 1 %. Ce résultat indique l'inadéquation de la loi binomiale négative pour décrire d'une manière globale les données obtenues lors de quatre campagnes d'échantillonnage parmi les six études réalisées.



**FIGURE 7** : Distributions des fréquences cumulées croissantes des titres bactériens observés pour chaque campagne d'échantillonnage.

TABLEAU 10

TEST D'ADEQUATION DE  $\chi^2$  POUR L'AJUSTEMENT D'UNE LOI BINOMIALE NEGATIVE AUX SERIES EXPERIMENTALES OBTENUES POUR LES SIX CAMPAGNES DE PRELEVEMENTS EFFECTUEES SUR LE RESEAU MESSIN

DATE DE PRELEVEMENT	NOMBRE DE PRELEVEMENTS	MOYENNE $\bar{r}$	PARAMETRES DE LA LOI BINOMIALE NEGATIVE $\hat{p}$ $R$	NOMBRE x DE BACTERIES par ml	EFFECTIF THEORIQUE	EFFECTIF OBSERVE	CONTRIBUTION $\chi^2$	DEGRES DE LIBERTE	$\chi^2$
13 décembre	100	419,52	1943,66    0,21584	0 1 ≤ x ≤ 2 3 ≤ x ≤ 9 10 ≤ x ≤ 19 20 ≤ x ≤ 39 40 ≤ x ≤ 99 100 ≤ x ≤ 299 300 ≤ x ≤ 699 700 ≤ x ≤ 1499 x ≥ 1500	19,50 6,76 8,48 5,74 6,55 10,05 14,06 11,50 9,16 8,19	17 8 10 8 10 11 10 10 7 9	0,32 0,23 0,27 0,89 1,82 0,09 1,17 0,19 0,51 0,08	7	5,57
10 janvier	102	583,07	3504,02    0,16640	0 1 ≤ x ≤ 2 3 ≤ x ≤ 9 10 ≤ x ≤ 39 40 ≤ x ≤ 99 100 ≤ x ≤ 299 300 ≤ x ≤ 699 700 ≤ x ≤ 1499 x ≥ 1500	26,23 6,91 8,03 10,89 8,49 11,60 9,67 8,43 11,76	19 11 14 17 11 15 5 4 6	1,99 2,43 4,44 3,43 0,74 1,00 2,25 2,33 2,82	6	21,43 <sup>++</sup>
21 février	100	695,84	3533,64    0,19692	0 1 ≤ x ≤ 2 3 ≤ x ≤ 9 10 ≤ x ≤ 19 20 ≤ x ≤ 39 40 ≤ x ≤ 99 100 ≤ x ≤ 299 300 ≤ x ≤ 699 700 ≤ x ≤ 1499 x ≥ 1500	20,01 6,30 7,67 5,11 5,76 8,78 12,38 10,63 9,52 13,83	12 8 9 14 12 15 13 4 5 8	3,21 0,46 0,23 15,49 6,74 4,40 0,03 4,14 2,14 2,46	7	39,29 <sup>++</sup>

TABLEAU 10 (suite et fin)

DATE DE PRELEVEMENT	NOMBRE DE PRELEVEMENTS	MOYENNE $\bar{r}$	PARAMETRES DE LA LOI BINOMIALE NEGATIVE $\hat{p}$ $R$	NOMBRE x DE BACTERIES par ml	EFFECTIF THEORIQUE	EFFECTIF OBSERVE	CONTRI-BUTION $\chi^2$	DEGRES DE LIBERTE	$\chi^2$
3 avril	101	455,47	1705,75    0,26702	0 1 ≤ x ≤ 2 3 ≤ x ≤ 9 10 ≤ x ≤ 19 20 ≤ x ≤ 39 40 ≤ x ≤ 99 100 ≤ x ≤ 299 300 ≤ x ≤ 699 700 ≤ x ≤ 1499 x ≥ 1500	13,84 6,03 8,16 5,83 6,89 10,99 16,06 13,52 10,77 8,90	6 8 10 7 11 27 13 11 3 6	4,44 0,64 0,41 0,24 2,46 23,30 0,58 0,47 5,61 0,94	7	39,09 <sup>++</sup>
15 mai	98	346,33	685,59    0,50515	0 ≤ x ≤ 2 3 ≤ x ≤ 9 10 ≤ x ≤ 19 20 ≤ x ≤ 39 40 ≤ x ≤ 99 100 ≤ x ≤ 299 300 ≤ x ≤ 699 700 ≤ x ≤ 1499 x ≥ 1500	6,81 6,03 5,41 7,48 14,06 23,57 19,43 11,57 3,64	1 8 6 18 22 11 19 6 7	4,96 0,65 0,06 14,80 4,48 6,71 0,01 2,68 3,11	6	37,45 <sup>++</sup>
26 juin	101	5635,75	20214,60    0,27880	0 ≤ x ≤ 3 4 ≤ x ≤ 19 20 ≤ x ≤ 59 60 ≤ x ≤ 149 150 ≤ x ≤ 399 400 ≤ x ≤ 999 1000 ≤ x ≤ 3999 4000 ≤ x ≤ 9999 10000 ≤ x ≤ 24999 x ≥ 25000	10,14 6,07 5,87 6,43 8,88 10,58 20,48 14,98 12,01 5,56	8 8 9 9 12 13 13 12 8 9	0,45 0,62 1,67 1,03 1,10 0,55 2,73 0,59 1,34 2,13	7	12,21

++ : valeur significative au seuil 1 %

## 2.3.2. Configuration de l'hétérogénéité spatio-temporelle

### 2.3.2.1. Reproductibilité de la structure d'abondance des populations bactériennes

La procédure de groupement, mentionnée à la section 2.1.3.3. et devant conduire à la partition de l'aire d'échantillonnage en zones de densité bactérienne, a été appliquée aux données bactériologiques de chacune des six campagnes de prélèvements afin de donner une image objective et concrète de la variabilité spatio-temporelle du peuplement bactérien dans le réseau messin.

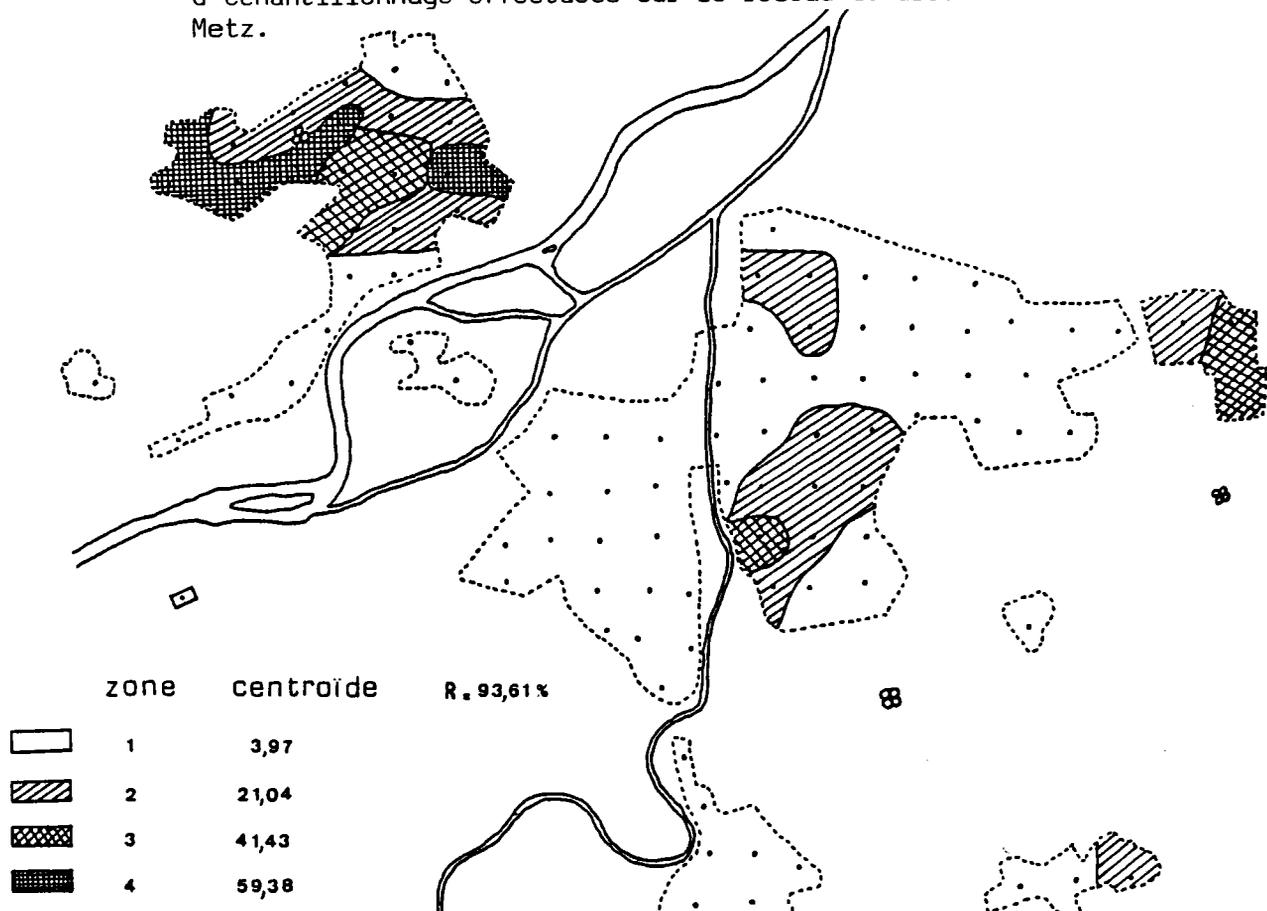
La localisation des stations de prélèvement appartenant aux différentes zones ainsi délimitées est illustrée pour chaque campagne d'échantillonnage, respectivement sur les figures 8 (a, b, c, d, e, et f). Les centroïdes caractérisant chaque zone de densité bactérienne de même que la fraction, R, de la variance totale expliquée par le groupement ont également été reportés sur ces figures.

L'étude comparative des six représentations de l'abondance bactérienne montre une grande analogie de structure compte tenu du nombre relativement faible des stations fluctuant d'une zone à l'autre au cours des diverses campagnes de prélèvements. Cette analogie se traduit notamment par une tendance très marquée pour les quartiers situés à l'est (Borny, Vallières) et au nord-ouest (Ban-St-Martin, Plappeville) de l'agglomération messine, à contenir les zones de plus forte densité bactérienne, tandis que les zones de faible concentration se trouvent invariablement à l'ouest et au sud (Sablon, Magny).

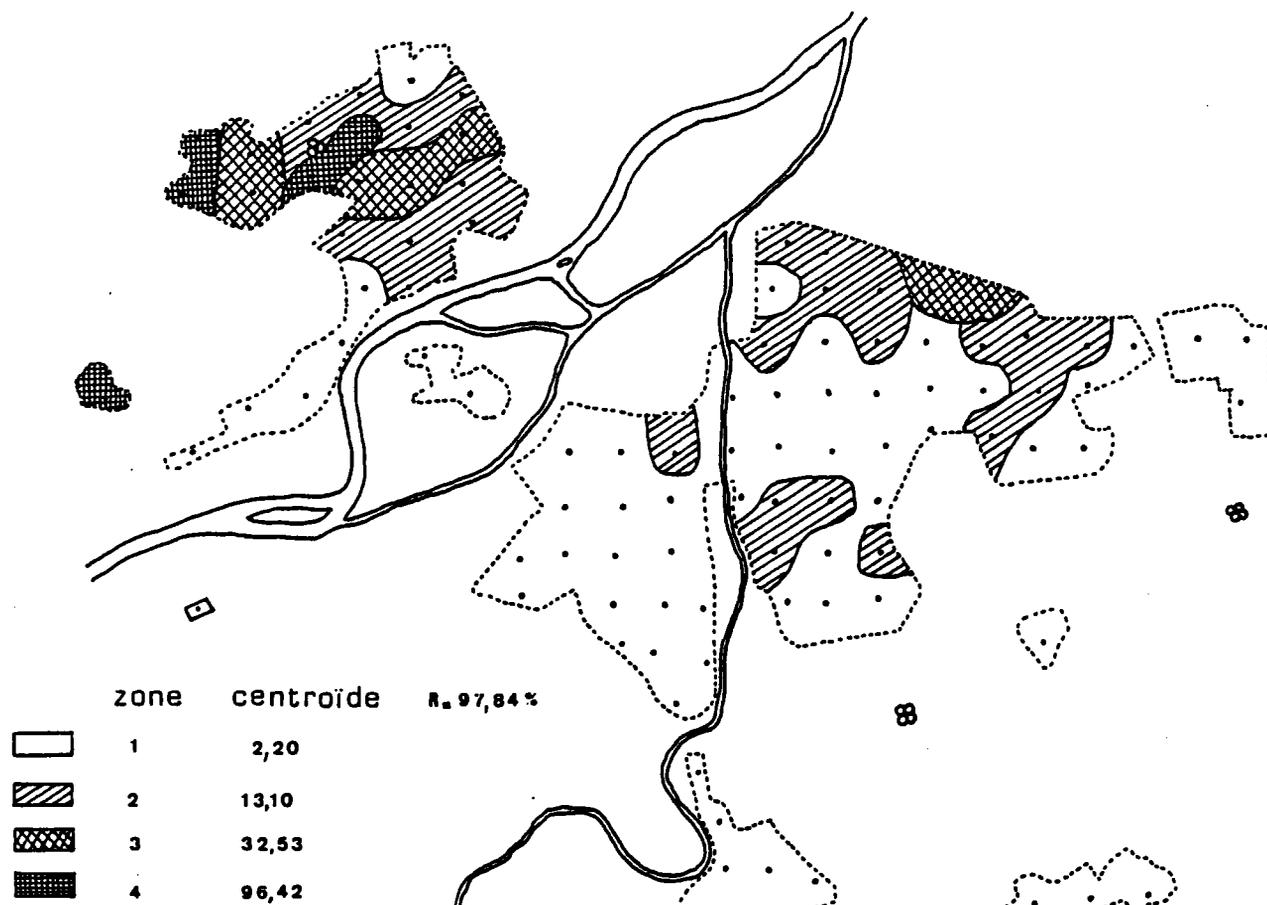
Afin d'objectiver ces observations, le degré de reproductibilité de la structure spatiale de l'hétérogénéité peut être évalué au moyen du coefficient de corrélation de rang de Kendall ( $\tau$ ) qui a été calculé entre toutes les séries expérimentales des six campagnes de prélèvements prises deux à deux. Toute valeur de  $\tau$  s'écartant significativement de zéro doit évoquer l'existence d'une liaison monotone entre les deux séries de valeurs confrontées. Cette propriété confère à  $\tau$  la qualité d'une mesure de ressemblance de la structure d'abondance du peuplement bactérien.

Les différentes valeurs de  $\tau$ , calculées pour chaque paire de séries expérimentales à l'aide d'un tableau de contingence à double entrée (LEGENDRE et LEGENDRE, 1979), sont consignées dans le tableau 11 sous la forme d'une matrice de similarité représentée par son triangle inférieur.

**FIGURE 8** : Zones de densité bactérienne observées au cours de six campagnes d'échantillonnage effectuées sur le réseau de distribution de la ville de Metz.



**FIGURE 8a** : Campagne d'échantillonnage du 13 décembre.



**FIGURE 8b** : Campagne d'échantillonnage du 10 janvier.

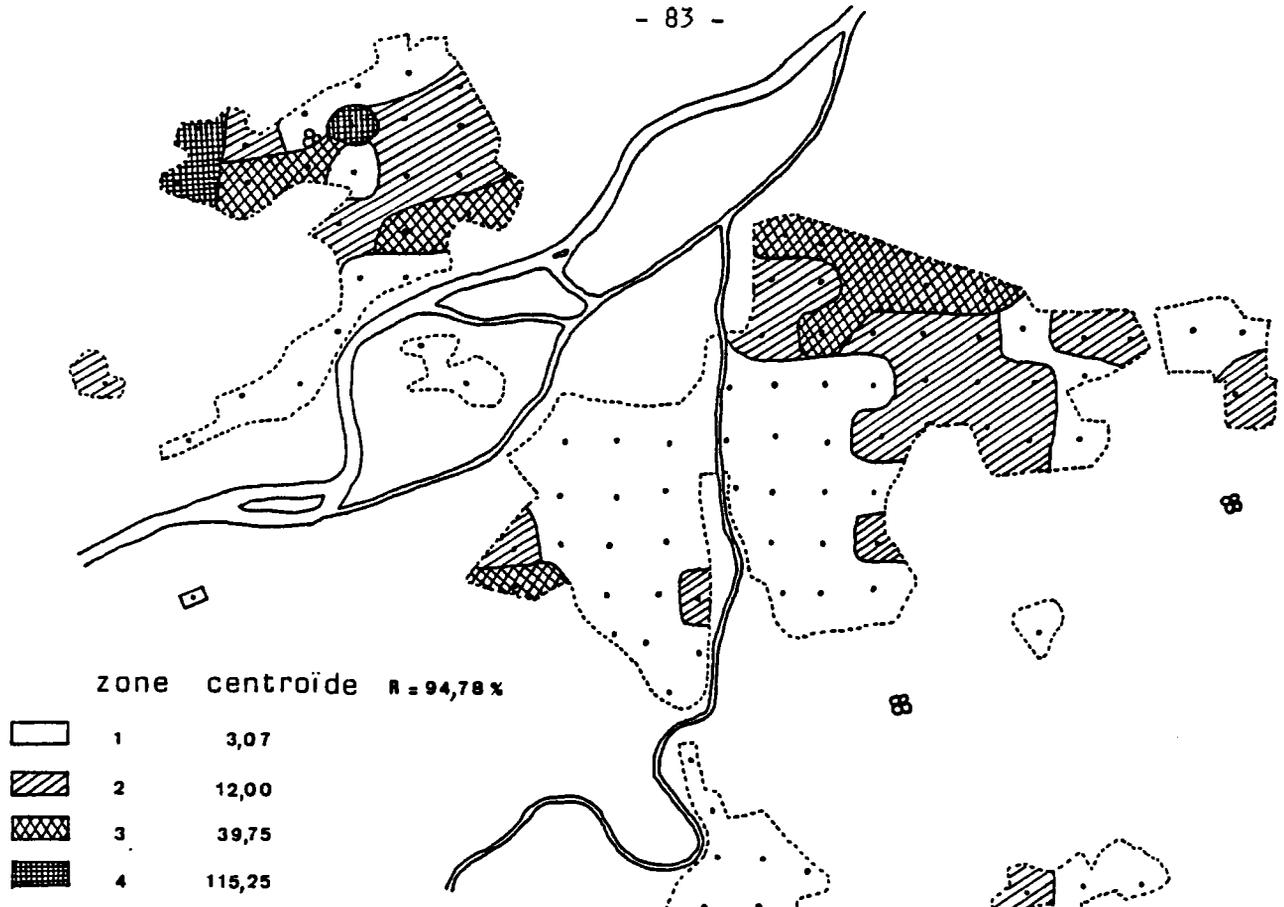


FIGURE 8c : Campagne d'échantillonnage du 21 février.

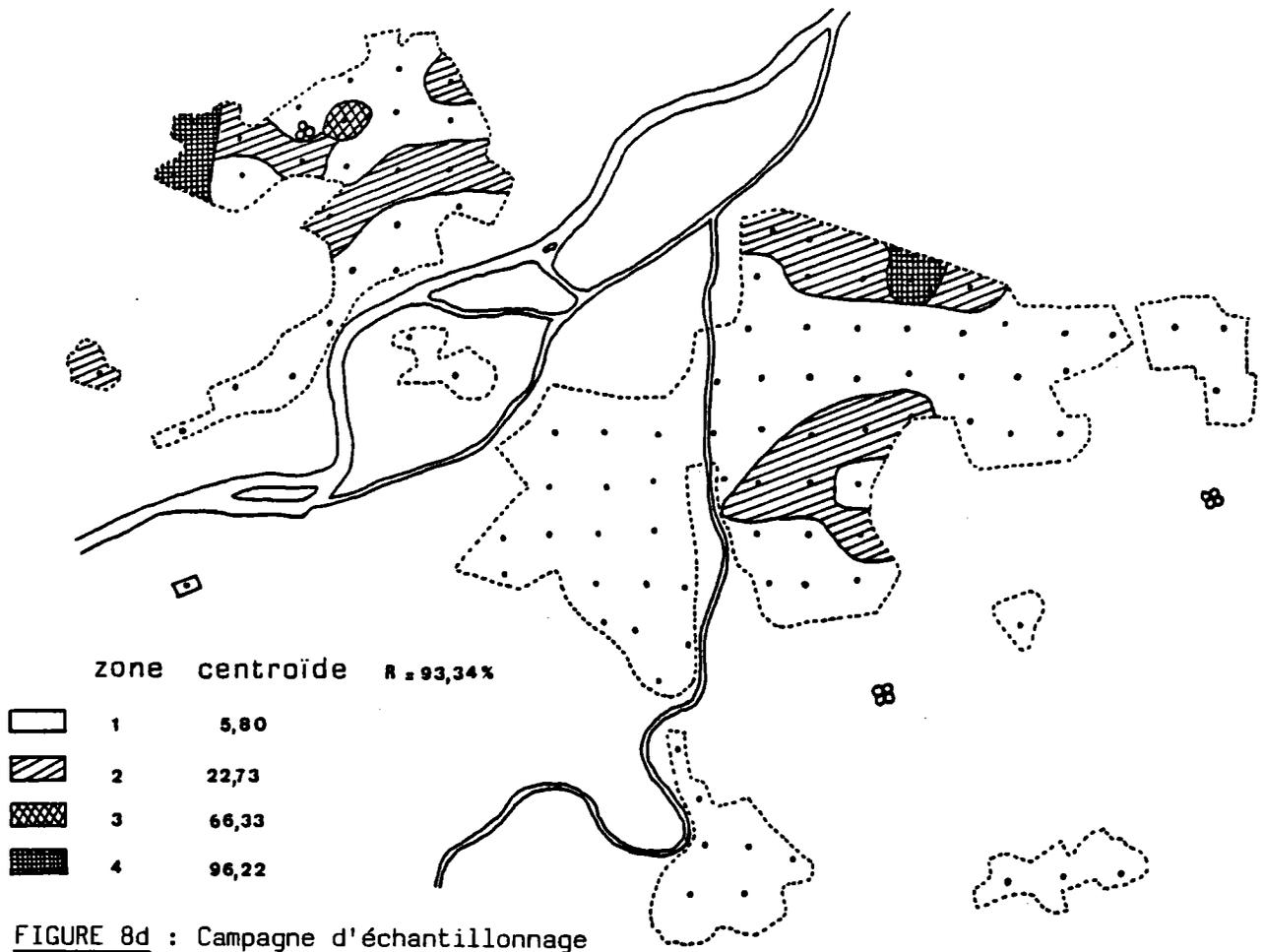


FIGURE 8d : Campagne d'échantillonnage du 3 avril.

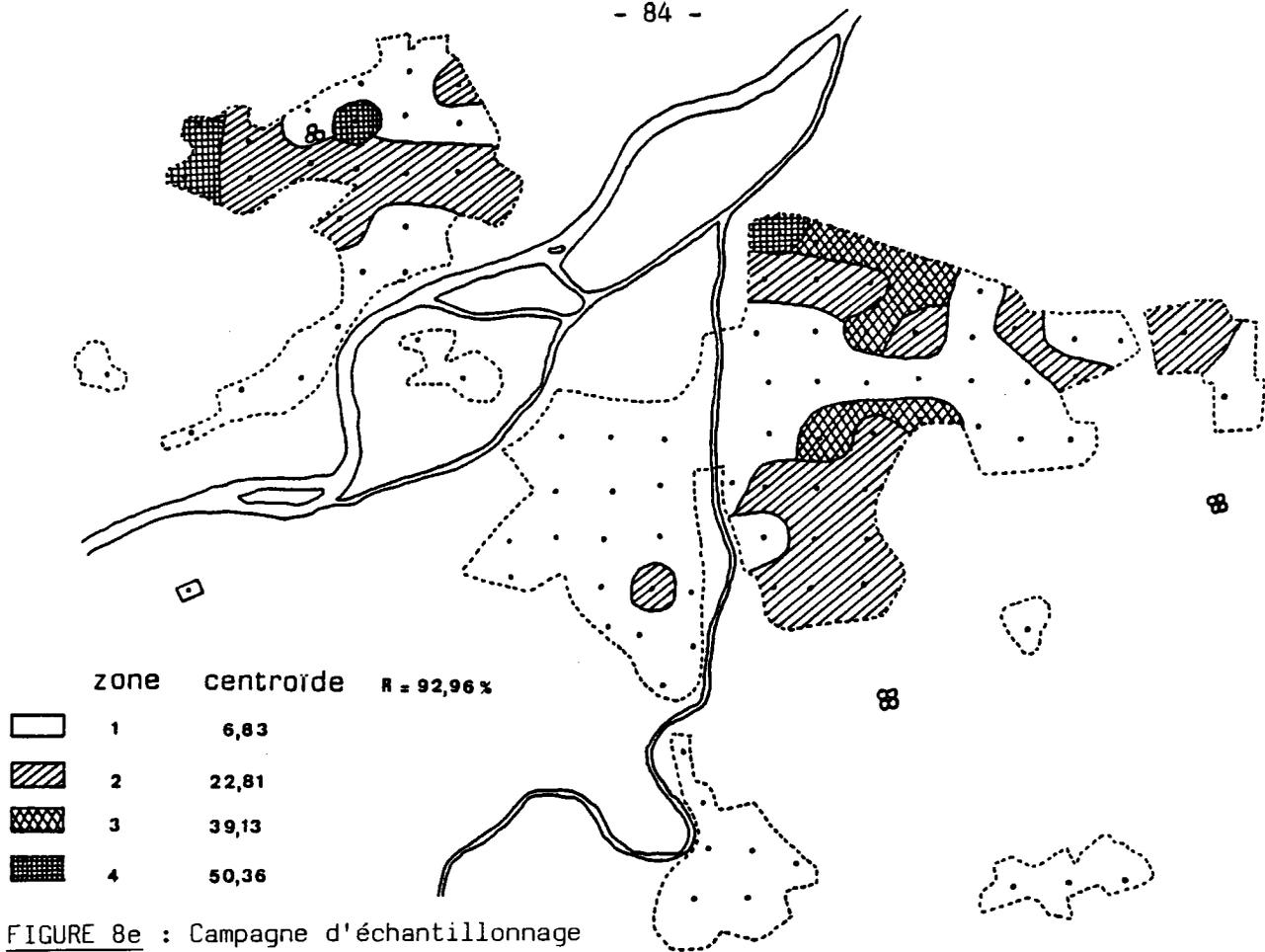


FIGURE 8e : Campagne d'échantillonnage du 15 mai.

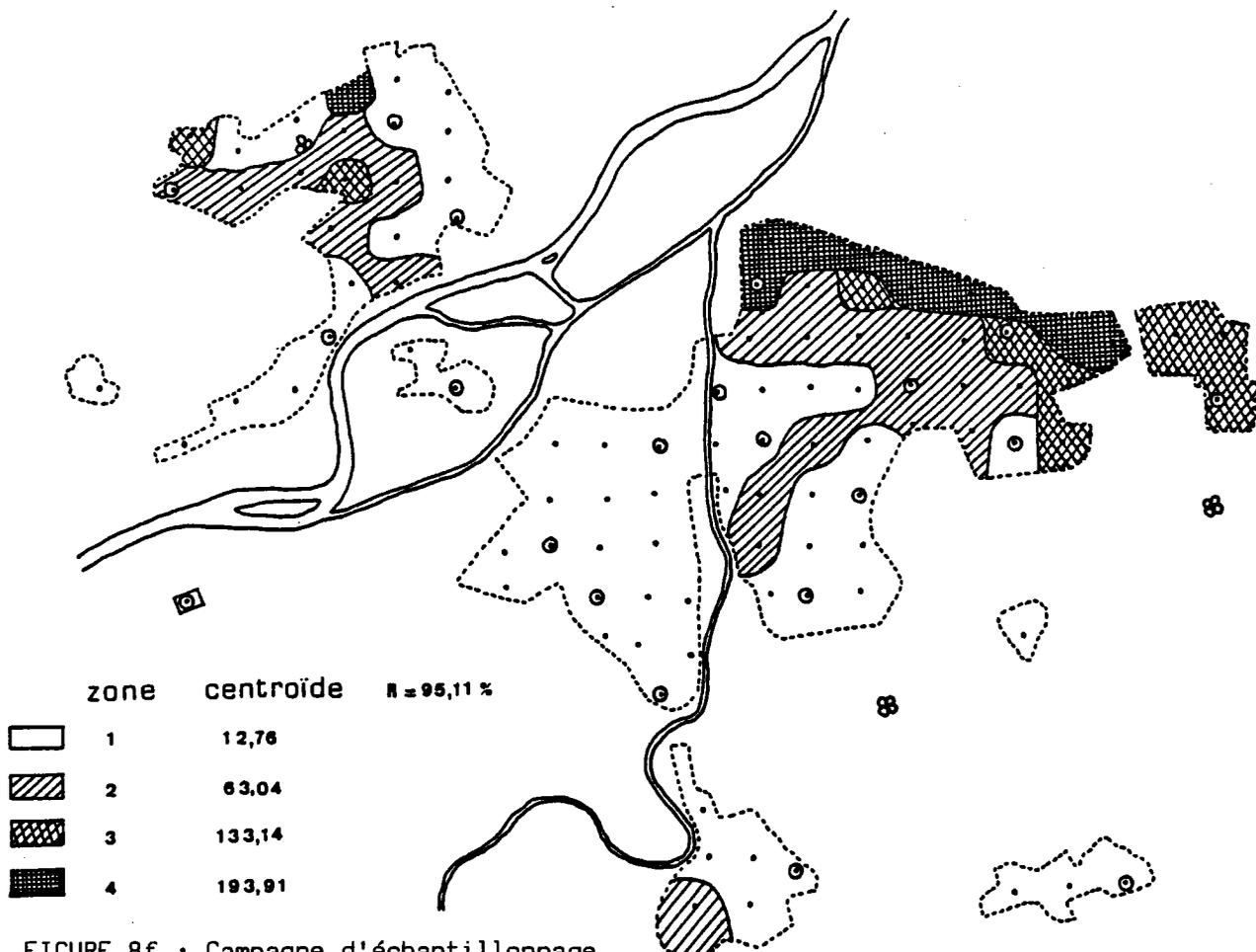


FIGURE 8f : Campagne d'échantillonnage du 26 juin.

**TABEAU 11** : Matrice de similarité pour les six campagnes d'échantillonnage, calculée à l'aide du coefficient de corrélation de rang de Kendall.

CAMPAGNES DE PRELEVEMENTS	13 décembre	10 janvier	21 février	03 avril	15 mai	26 juin
13 décembre	1,00000					
10 janvier	0,54928	1,00000				
21 février	0,49483	0,50980	1,00000			
03 avril	0,50723	0,52801	0,53684	1,00000		
15 mai	0,44917	0,37863	0,39571	0,51463	1,00000	
26 juin	0,50722	0,47804	0,47538	0,52177	0,53561	1,00000

Les tests d'indépendance, réalisés à partir de la statistique donnée par la formule (45) pour chacune des 15 valeurs de  $\tau$  ainsi obtenues, se sont tous montrés significatifs au seuil de probabilité 1 %. Ce résultat sans équivoque témoigne d'une grande stabilité de structure du peuplement bactérien dans le système de distribution étudié.

#### 2.3.2.2. Zones de densité bactérienne

La persistance de l'image, donnée par l'hétérogénéité spatiale observée au cours des six campagnes de prélèvements et mise en évidence dans la section précédente, justifie l'extension de la méthode de groupement à l'ensemble de toutes les données disponibles. Ainsi, la procédure de groupement, basée sur la prise en considération et l'analyse simultanée des six mesures de la concentration bactérienne associées à chaque point de prélèvement, a permis de diviser le réseau de distribution en quatre zones de densité bactérienne qui ont été représentées sur la figure 9. Cette partition du réseau en quatre zones explique plus de 80 % de la variabilité totale ; cependant toute tentative visant à affiner l'image de la structure spatiale du peuplement bactérien par l'augmentation du nombre des zones ne s'accompagne que d'un accroissement très modeste de la variance expliquée. A titre d'exemple, cet accroissement n'est que de 3 % environ lorsque le groupement inclut la formation d'une cinquième zone. La division du système de distribution en cinq zones de densité est illustrée par la figure 10. Finalement, le faible gain de la variance expliquée consécutif à la formation de zones supplémentaires, nous a amenés à ne retenir que les quatre zones, qui ont été délimitées sur la figure 9, pour la suite de nos investigations.

Notons, par ailleurs, que les zones représentées sur la figure 9 offrent un aspect nettement plus compact que les zones, beaucoup plus dispersées, qui ont été mises en évidence lors des différentes campagnes d'échantillonnage considérées individuellement (Figure 8). En particulier, la zone 1 qui est caractérisée par les plus faibles concentrations bactériennes, comprend 53 points de prélèvement sur un total de 102 stations échantillonnées et englobe par conséquent plus de la moitié de la superficie du système de distribution. Les zones de plus forte densité bactérienne sont, quant à elles, localisées dans les parties nord-ouest de même que les quartiers est de l'agglomération messine comme cela avait déjà été remarqué pour les campagnes d'échantillonnage prises séparément.

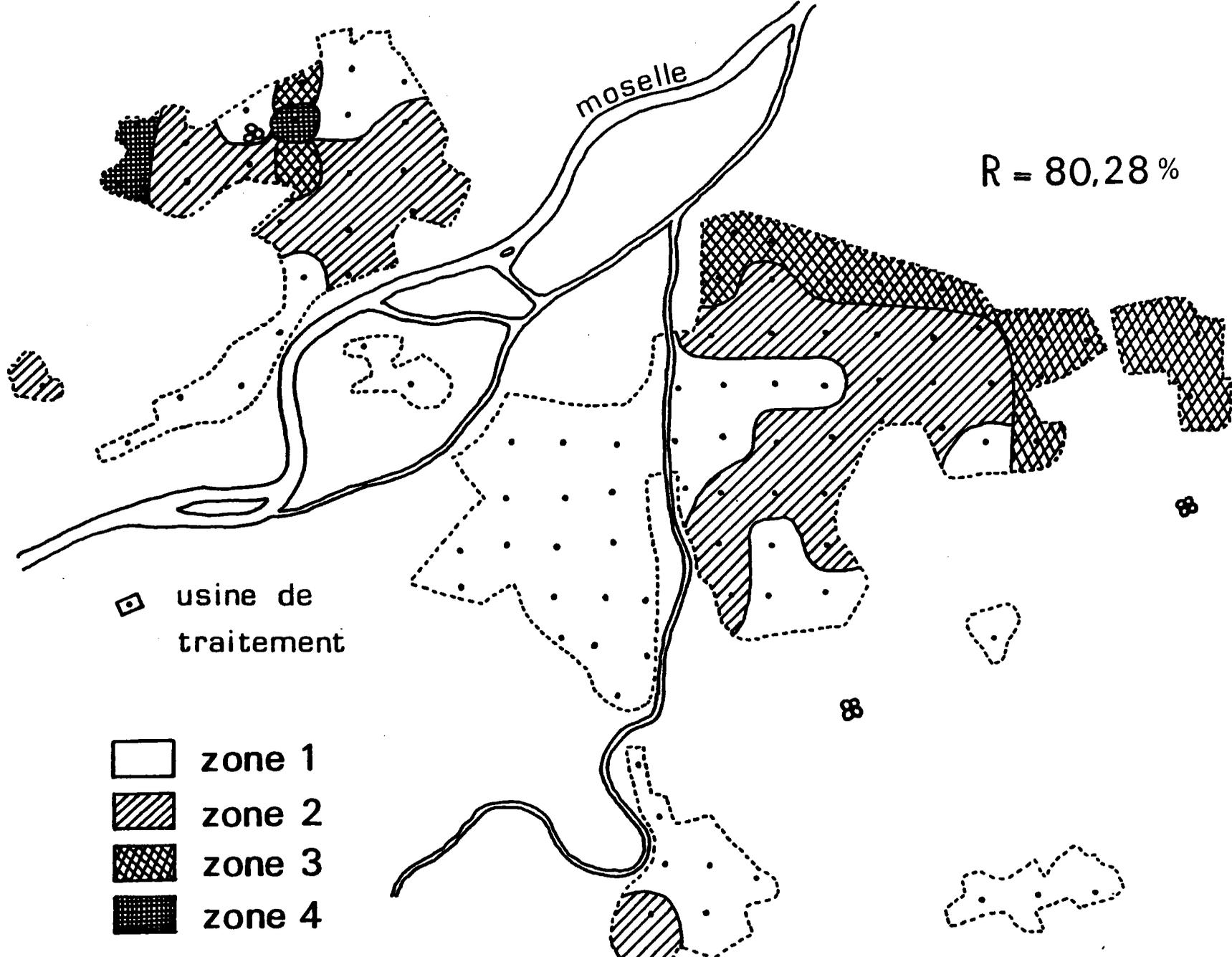


FIGURE 9 : Partition du système de distribution en quatre zones de densité bactérienne obtenues à partir de l'exploitation des résultats des six campagnes d'échantillonnage.



L'analyse en composantes principales, utilisée en tant qu'outil complémentaire à la méthode de groupement dans le but de faciliter la visualisation et la compréhension de la structure spatiale de l'hétérogénéité, a été effectuée sur l'ensemble des six séries de données bactériologiques expérimentales. Cette technique a permis de représenter les différents points de prélèvement dans l'espace réduit associé aux deux premiers axes principaux. La figure 11 donne une illustration graphique de la projection dans l'espace réduit de chaque station de prélèvement identifiable par le numéro de la zone à laquelle elle a effectivement été affectée par la méthode de groupement précédente. L'examen de la figure 11 permet aisément de reconnaître les quatre groupes de stations tels qu'ils ont déjà été définis par la méthode de groupement séquentielle. Ainsi, les zones 1 et 4 peuvent être distinguées des zones 2 et 3 sur la base du premier axe principal qui représente 63,3 % de la variabilité totale et traduit probablement le niveau de la densité bactérienne chronique caractérisant chaque point de prélèvement. Il est, par ailleurs, possible de discriminer les nuages de points composés des stations appartenant respectivement aux zones 2 et 3 au moyen du deuxième axe principal. Cet axe, qui explique 15,0 % de la variance, est vraisemblablement lié aux changements et à l'évolution de la structure d'abondance bactérienne propres à la sixième campagne de prélèvements. En effet, la représentation des six campagnes d'échantillonnage (descripteurs) dans l'espace réduit des deux premiers axes principaux, illustrée sur la figure 12, permet d'apprécier la contribution de chaque descripteur à la formation de l'espace réduit. Cet examen est facilité par la construction du cercle de contribution équilibrée des descripteurs, dont le rayon est compatible avec l'hypothèse d'une contribution équitablement partagée de chaque descripteur à la formation de tous les axes principaux. D'une manière plus précise, la contribution relative des différents axes-descripteurs, relativement à un même axe principal, est évaluée par l'étude comparative des projections des différents axes-descripteurs sur l'axe principal considéré. A cet égard, la figure 12 montre une bonne concordance de liens entre l'ensemble des six descripteurs et le premier axe principal, qu'il faut peut-être imputer à la reproductibilité de la structure d'abondance. En revanche, elle fait très nettement ressortir le rôle prépondérant de la sixième campagne d'échantillonnage quant à la signification du deuxième axe principal.

### 2.3.2.3. Modélisation de la structure d'abondance du peuplement bactérien

Le problème de l'ajustement d'une distribution binomiale négative théorique aux différentes séries expérimentales a été reconsidéré pour les données recueillies à l'intérieur de chacune des quatre zones mises en évidence par la méthode de groupement. Ainsi, pour chaque campagne de prélèvements, le tableau 12 donne : le nombre des stations échantillonnées, la moyenne arithmétique  $\bar{T} = \hat{\lambda}$  des titres bactériens observés, de même que les estimations  $\hat{p}$  et  $\hat{k}$  des paramètres de la loi binomiale négative correspondant à la série expérimentale obtenue au sein de chacune des quatre zones de densité bactérienne. Les résultats des tests d'adéquation, permettant de juger de l'efficacité de l'ajustement, sont également présentés dans le tableau 12. Cependant, le nombre trop faible des points de prélèvement appartenant aux zones 3 et 4 nous a contraints à limiter l'application de ce test aux deux premières zones seulement. Par suite, aucune des douze valeurs de  $\chi^2$  calculées ne s'est révélée significative au seuil 1 %. La distribution binomiale négative apparaît de ce fait comme un modèle mathématique adapté à la description des données bactériologiques observables à l'intérieur des zones de densité bactérienne.

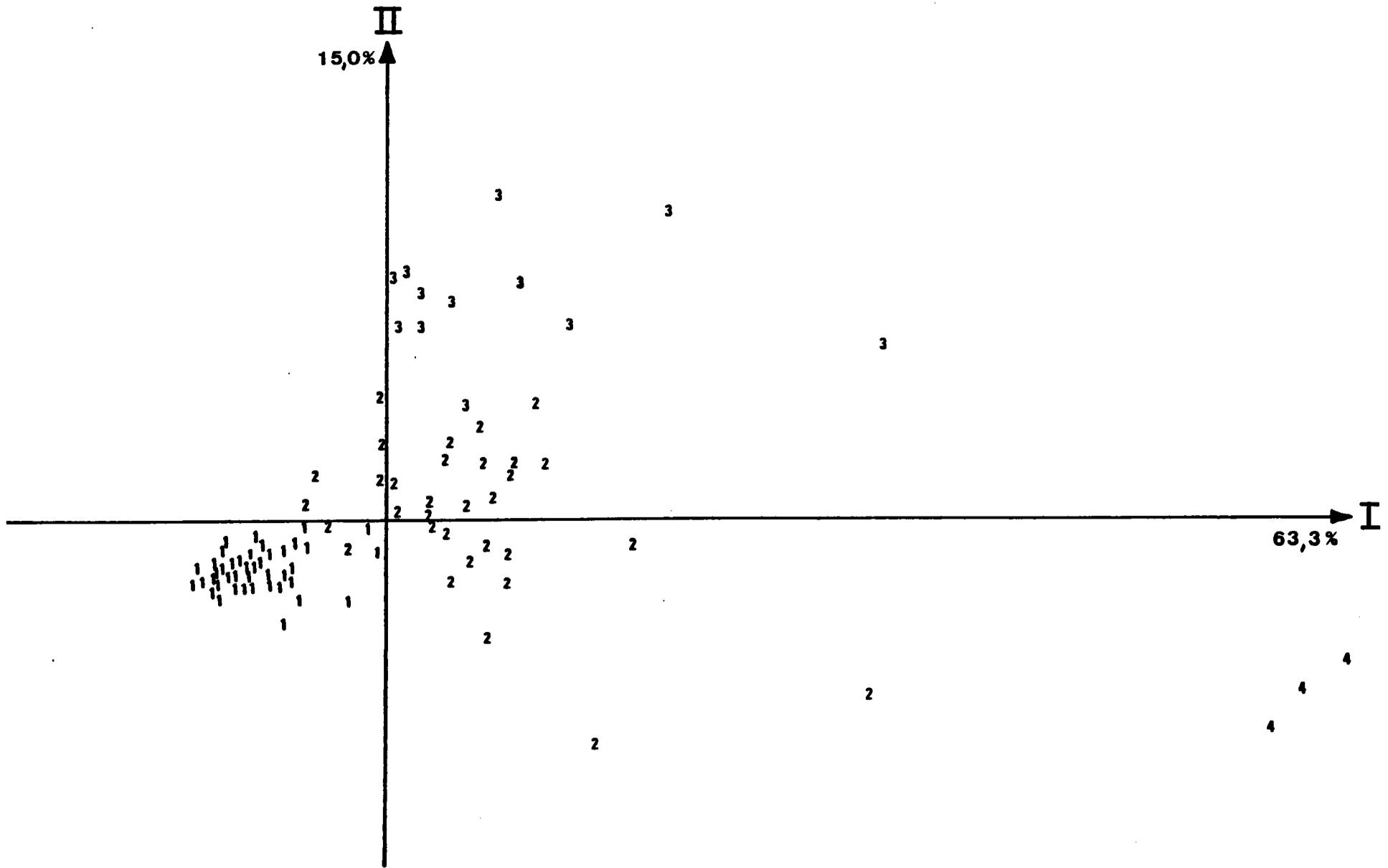


FIGURE 11 : Zones de densité bactérienne représentées dans l'espace réduit associé aux deux premiers axes principaux.

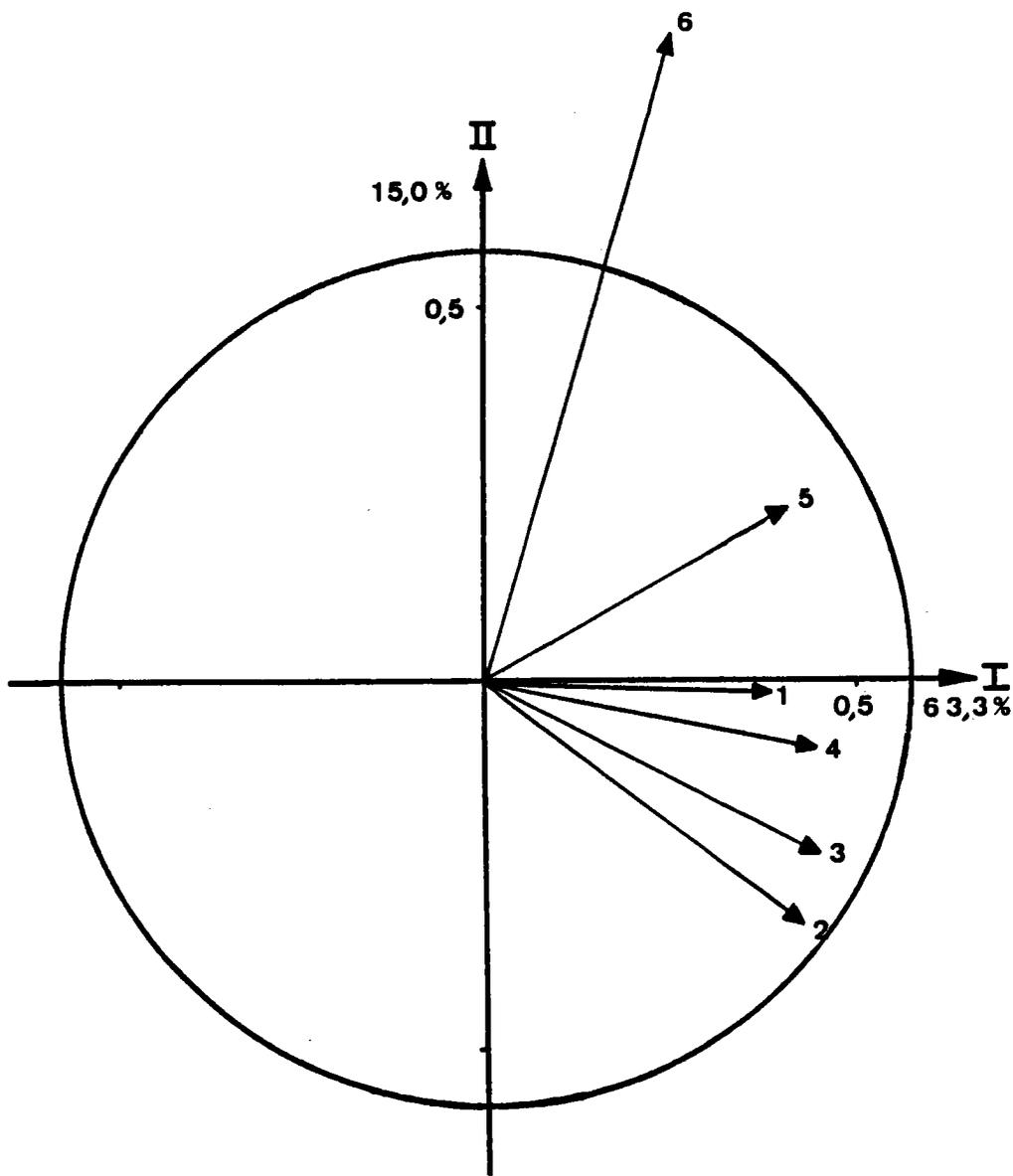


FIGURE 12 : Représentation des six campagnes d'échantillonnage (descripteurs) dans l'espace réduit des deux premiers axes principaux.

TABLEAU 12

ESTIMATION DES PARAMETRES DE LA LOI BINOMIALE NEGATIVE  
ET TEST D'ADEQUATION DE  $\chi^2$  POUR CHAQUE ZONE AU COURS  
DE CHACUNE DES SIX CAMPAGNES DE PRELEVEMENTS

DATE DE PRELEVEMENT	ZONE	NOMBRE DE PRELEVEMENTS	MOYENNE $\bar{r}$	PARAMETRES DE LA LOI BINOMIALE NEGATIVE		TEST D'ADEQUATION DE $\chi^2$	
				$\hat{p}$	$\hat{k}$	Degrés de liberté	$\chi^2$
13/12	1	52	36,81	105,54	0,34877	3 1	8,60 <sup>+</sup> 0,54
	2	31	668,23	1224,24	0,54583		
	3	14	561,64	1072,43	0,52371		
	4	3	3820,00	53,21	71,78516		
10/01	1	53	24,21	68,53	0,35324	3 1	10,53 <sup>+</sup> 2,79
	2	31	852,65	2676,59	0,31856		
	3	15	233,20	448,11	0,52041		
	4	3	9420,00	10,78	874,22991		
21/02	1	52	44,08	127,08	0,34683	3 1	4,88 4,39 <sup>+</sup>
	2	31	627,94	1547,44	0,40579		
	3	14	537,58	1206,38	0,44561		
	4	3	13433,33	487,67	27,54602		
03/04	1	52	35,92	66,95	0,53659	3 1	5,22 2,27
	2	31	284,64	308,46	0,92280		
	3	15	787,39	2266,19	0,34745		
	4	3	7833,33	1090,56	7,18284		
15/05	1	51	87,10	128,98	0,67527	3 1	11,08 <sup>+</sup> 2,22
	2	30	446,70	468,96	0,95253		
	3	14	577,72	626,41	0,92227		
	4	3	2669,67	45,20	59,05907		
26/06	1	52	197,60	420,58	0,46982	3 1	2,31 1,84
	2	31	3168,90	2387,71	1,32717		
	3	15	29119,99	3305,08	8,81069		
	4	3	7633,33	1088,53	7,01250		

+ : valeur significative au seuil 5 %

++ : valeur significative au seuil 1 %

La variation temporelle du niveau moyen de la concentration bactérienne caractérisant les différentes zones de densité est illustrée sur la figure 13 par les trajectoires des centroïdes (moyenne arithmétique de la concentration bactérienne par ml après transformation "racine carrée") de chaque zone en fonction du temps.

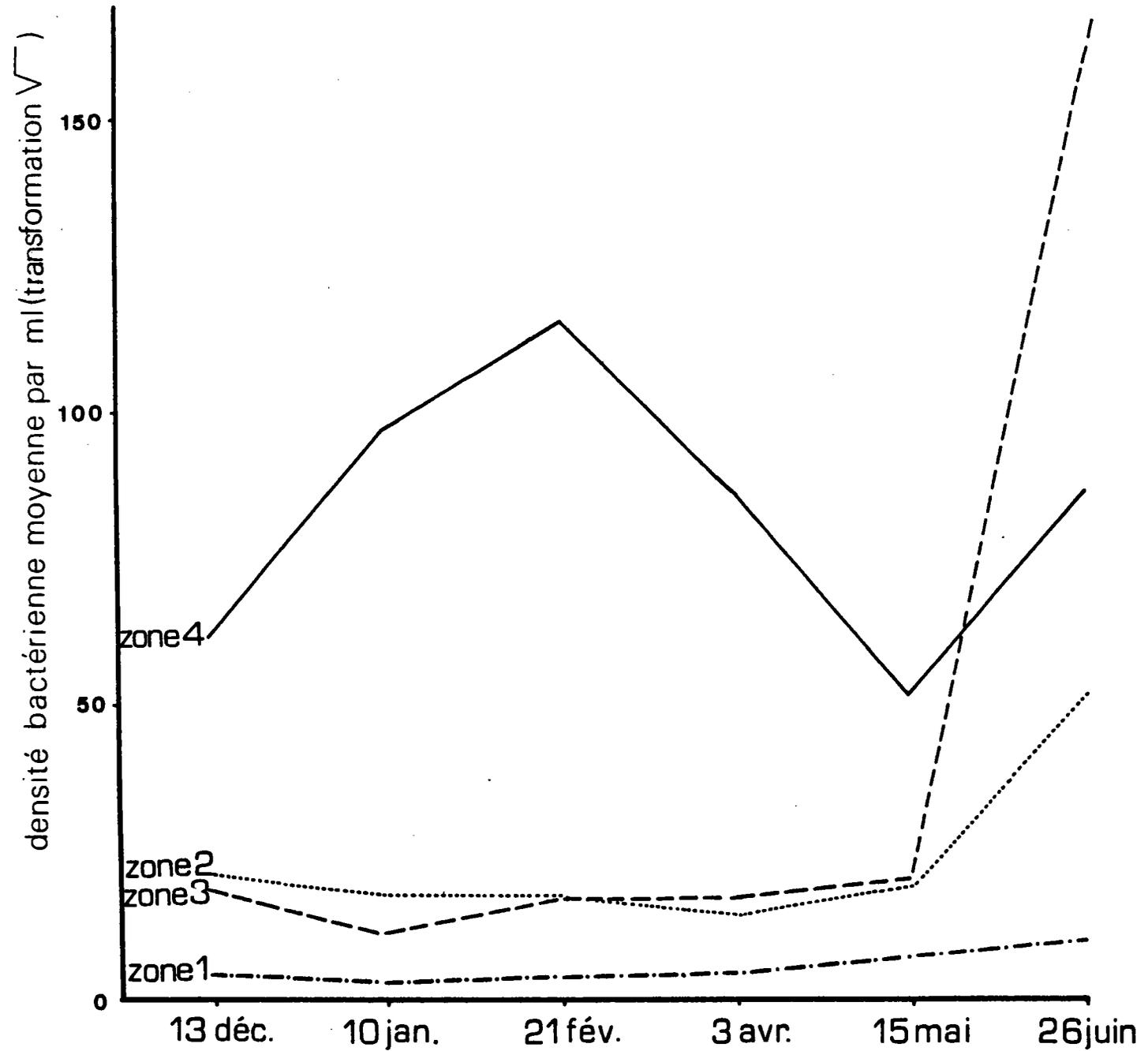


FIGURE 13 : Variation temporelle de la concentration bactérienne pour chaque zone de densité.

Le graphique fait clairement ressortir l'augmentation générale de la concentration bactérienne qui s'est manifestée pour l'ensemble du réseau au cours de la dernière campagne de prélèvements. Il est, de plus, intéressant de remarquer le manque de cohérence et de synchronisme dans l'évolution des trajectoires des centroïdes, ce qui est révélateur de l'existence d'une interaction entre le "facteur zone" et le temps. En d'autres termes, les diverses zones ne sont pas touchées de manière uniforme par les variations de la concentration bactérienne, et, par conséquent, elles semblent plutôt se comporter indépendamment les unes des autres. Ce manque de rigidité absolue dans l'interdépendance des niveaux de concentration bactérienne est surtout vérifié pour les zones 2 et 3 qui se différencient mutuellement lors de la dernière campagne d'échantillonnage après avoir suivi une évolution très similaire. Ce phénomène avait déjà été signalé à la section précédente.

### 2.3.3. Concentration bactérienne versus paramètres physico-chimiques de l'eau

#### 2.3.3.1. Temps de séjour

La dépendance éventuelle entre la concentration bactérienne et le temps de séjour de l'eau dans les canalisations, considéré ici en tant que variable explicative, a été exprimée par le coefficient de corrélation de rang de Kendall ( $\tau$ ). Pour chaque campagne d'échantillonnage,  $\tau$  a été calculé entre la série des temps de transit moyens de l'eau depuis l'usine de traitement de Moulins jusqu'aux différentes stations de prélèvements et la série correspondante des concentrations bactériennes observées. A cet effet, les temps de séjour approximatifs, établis sur des considérations hydrodynamiques liées à l'infrastructure du réseau, ont été répartis en cinq classes de niveau conformément aux indications reportées sur la figure 14.

Les six valeurs de  $\tau$  ainsi obtenues (13 décembre :  $\tau \approx 0,52$  ; 10 janvier :  $\tau \approx 0,49$  ; 21 février :  $\tau \approx 0,51$  ; 3 avril :  $\tau \approx 0,57$  ; 15 mai :  $\tau \approx 0,51$  et 26 juin :  $\tau \approx 0,59$ ) se sont révélées significatives avec un risque d'erreur de première espèce fixé à 1 %. Ces résultats montrent l'existence d'une relation monotone croissante entre le temps de séjour et la densité bactérienne observée.

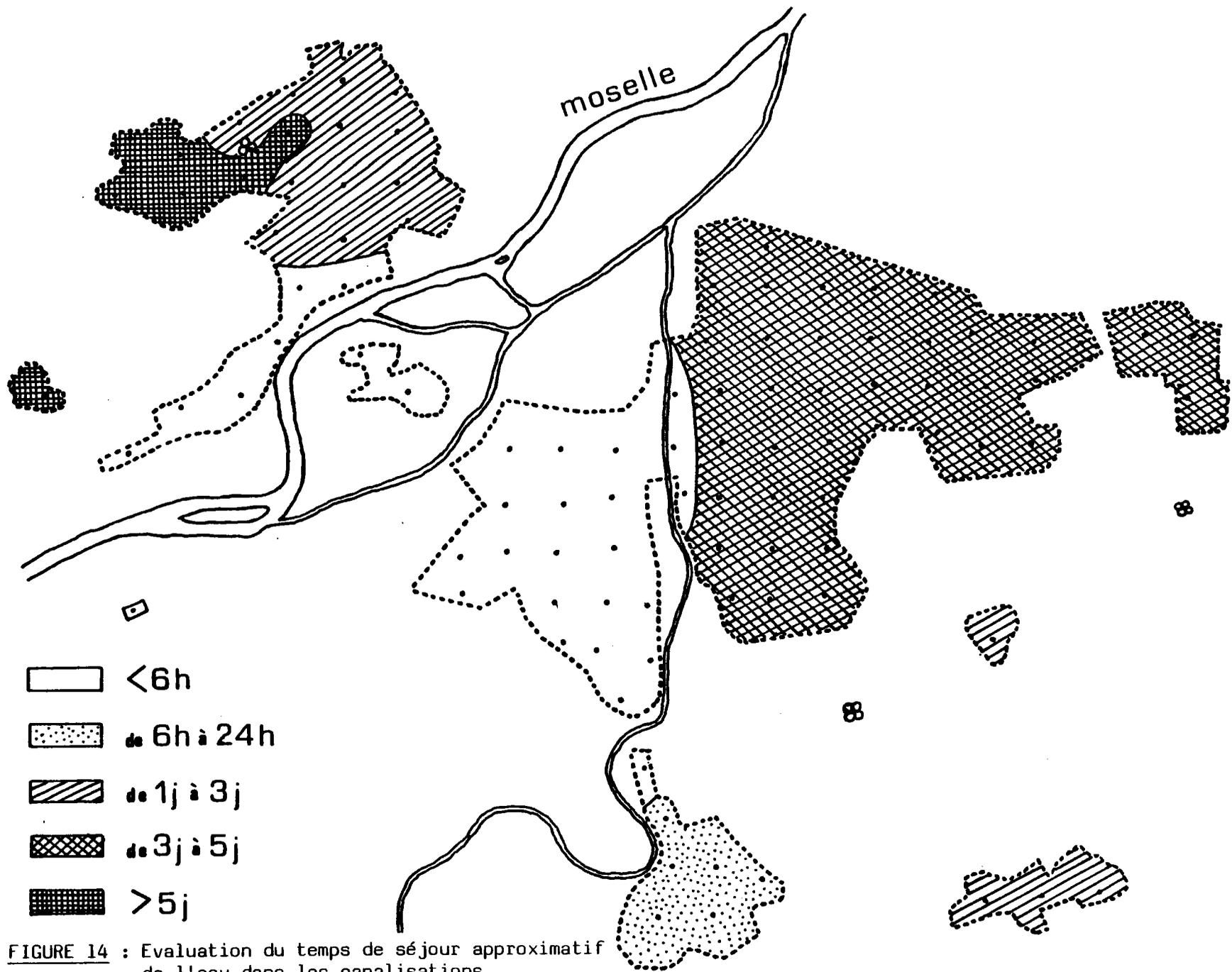


FIGURE 14 : Evaluation du temps de séjour approximatif de l'eau dans les canalisations.

### 2.3.3.2. Chlore résiduel

Les stations de prélèvement ayant fait l'objet d'un dosage volumétrique du chlore résiduel libre et total contenu dans l'eau de distribution lors de la sixième campagne d'échantillonnage, ont été encerclées sur la figure 8f. Les résultats de ces mesures ainsi que les concentrations bactériennes correspondantes sont consignés dans le tableau 13 pour chacun des 21 points de prélèvement considérés.

TABLEAU 13 : Quantités de chlore résiduel libre et total *versus* concentration bactérienne au cours de la campagne de prélèvements du 26 juin.

POINT DE PRELEVEMENT	CHLORE RESIDUEL LIBRE (mg/l)	CHLORE RESIDUEL TOTAL (mg/l)	CONCENTRATION BACTERIENNE par ml
0	0,50	0,60	1
8	0,00	0,10	173
10	0,00	0,00	5600
18	0,00	0,08	627
21	0,10	0,22	12
27	0,12	0,25	12
30	0,00	0,12	120
35	0,05	0,20	29
39	0,00	0,15	23
45	0,05	0,15	0
48	0,00	0,00	42000
58	0,00	0,05	11100
63	0,00	0,10	14200
68	0,00	0,10	245
70	0,00	0,05	4400
74	0,00	0,20	1200
76	0,00	0,10	491
83	0,00	0,05	818
88	0,00	0,10	200
95	0,00	0,00	20
100	0,00	0,00	130

Les résidus de chlore libre et total, dont les concentrations observées s'échelonnent respectivement de 0 à 0,50 mg/l et de 0 à 0,60 mg/l, diminuent rapidement avec l'accroissement de la distance par rapport à l'usine de traitement. Ainsi, aucune trace de chlore résiduel libre n'a pu être décelée dans les secteurs périphériques du réseau associés aux zones de plus forte densité bactérienne. De plus, les valeurs de  $\tau$  calculées entre la série des concentrations bactériennes et les niveaux correspondant de chlore résiduel libre et total (i. e.,  $\tau \approx -0,62$  et respectivement,  $\tau \approx -0,54$ ) se sont montrées significatives au seuil 1 %. Ce résultat permet d'objectiver l'existence d'une corrélation négative entre les quantités de chlore résiduel contenues dans l'eau de distribution et la concentration bactérienne.

### 2.3.3.3. Température de l'eau

Le choix d'un modèle log-linéaire pour l'étude quantitative d'une liaison éventuelle entre la température moyenne de l'eau et l'abondance bactérienne a été guidé par la relation de proportionnalité apparente observée entre l'écart-type et la moyenne arithmétique des résultats de chacune des six séries de données expérimentales. L'existence d'une régression linéaire avec variance résiduelle constante implique le modèle mathématique suivant :

$$\log m_g = a \theta + b + \epsilon$$

où  $\epsilon$  désigne une variable aléatoire normalement distribuée, de moyenne nulle et dont la variance, appelée variance résiduelle, est supposée indépendante de  $\theta$ . La figure 15 donne la disposition des points correspondant aux différentes campagnes d'échantillonnage. Les abscisses représentent la température moyenne de l'eau, soit, dans l'ordre chronologique : 7,8° C ; 8,0° C ; 7,3° C ; 8,8° C ; 12,2° C et 16,7° C, tandis que les ordonnées respectives de ces points sont égales à la moyenne géométrique de la densité bactérienne, à savoir : 33 ; 14 ; 26 ; 53 ; 80 et 504, exprimée sur une échelle logarithmique.

L'estimation des paramètres du modèle de régression adopté conduit à l'équation

$$\log m_g = 0,140 \theta + 0,316 \quad (46)$$

pour la droite de régression de la moyenne géométrique de la concentration bactérienne ( $m_g$ ) en fonction de la température ( $\theta$ ).

L'étroitesse de la relation linéaire entre  $\theta$  et  $\log m_g$  a été évaluée par le coefficient de corrélation linéaire de Pearson ( $r = 0,947$ ) des deux distributions. Le test d'indépendance mentionné dans le paragraphe 2.1.3.2., permet de vérifier que le coefficient de corrélation linéaire de Pearson des deux variables est différent de zéro avec un risque d'erreur égal à 1 %. De plus, la valeur du coefficient de détermination ( $r^2 = 0,897$ ) montre que la fraction de la variance de  $m_g$ , expliquée par  $\theta$ , représente près de 90 % de la variabilité totale de  $m_g$ . Enfin, le modèle mathématique décrit par la formule (46) fait état d'une régression exponentielle de la moyenne géométrique des concentrations bactériennes selon la température de l'eau. Dans les limites du domaine exploré, ce modèle implique un accroissement de 38 % environ de la moyenne géométrique  $m_g$  chaque fois que la température de l'eau s'élève de 1° C.

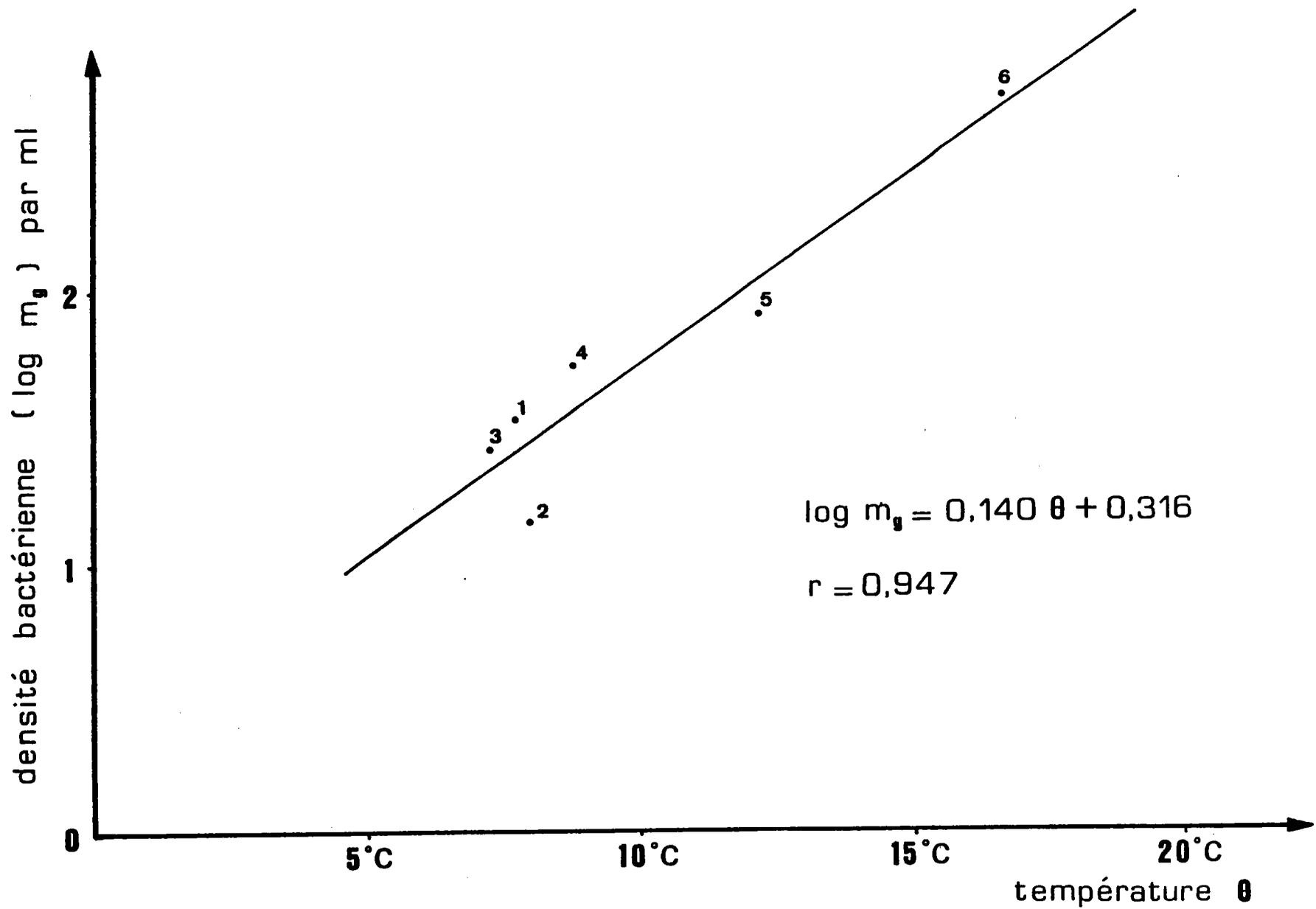


FIGURE 15 : Régression de la moyenne géométrique de la densité bactérienne selon la température moyenne de l'eau.

## 2.4. DISCUSSION

Dans ce chapitre, l'accent a été mis sur l'intérêt que présente l'utilisation de la loi binomiale négative pour la description des données bactériologiques obtenues dans les systèmes de distribution publique. Les tests d'adéquation résumés dans le tableau 10 montrent cependant clairement que les distributions empiriques observées au cours des six campagnes d'échantillonnage ne sont pas toutes conformes au modèle binomial négatif. Ce résultat semble, de prime abord, être en contradiction avec les conclusions formulées dans d'autres travaux (CHRISTIAN et PIPES, 1983 ; EL-SHAARAWI et al., 1985 ; PIPES et CHRISTIAN, 1982) faisant état de l'adéquation de la distribution binomiale négative avec les séries de données bactériologiques étudiées. L'explication de ce désaccord est vraisemblablement liée à la grande taille du réseau échantillonné dans le cadre de la présente étude, mais aussi au caractère systématique du plan d'échantillonnage lui-même qui a permis de couvrir uniformément l'ensemble de toutes les sections du réseau messin. En effet, l'examen minutieux des effectifs observés dans les diverses classes d'abondance du tableau 10 laisse entrevoir la nature multimodale de la distribution empirique des titres bactériens dont l'origine serait imputable aux différents paliers matérialisés par les zones de densité bactérienne. Cette hypothèse est corroborée par les tests d'adéquation effectués à l'intérieur de chaque zone de densité bactérienne. Ces tests font nettement ressortir la prééminence du modèle binomial négatif pour ajuster les données bactériologiques recueillies au sein d'une même zone du système de distribution. Par conséquent, il devient loisible de considérer le réseau de distribution de la ville de Metz comme un système composite formé de plusieurs sous-systèmes hétérogènes caractérisés par différents niveaux d'abondance bactérienne. A cette image très contrastée de l'hétérogénéité spatiale, il faut adjoindre une variation temporelle de la concentration bactérienne globale particulièrement prononcée, qui n'affecte cependant pas nécessairement les zones de manière uniforme et équitable.

La partition d'un réseau de distribution en zones de densité bactérienne témoigne d'une configuration structurée et persistante de l'hétérogénéité spatiale. Par suite, la structure d'abondance du peuplement bactérien est probablement, en partie, induite par les effets conjugués de quelques variables physiques et chimiques du système, telles que la nature et l'ancienneté des canalisations (NAGY et al., 1983), les caractéristiques hydrodynamiques des masses d'eau, les quantités de chlore résiduel ainsi qu'une multitude d'autres facteurs dont les effets contribuent à la formation d'une structure hiérarchisée et écologiquement stable (MEANS et al., 1981). Il serait alors intéressant de savoir si la disposition et la configuration des zones observées dans le réseau messin sont compatibles avec les aspects généraux de l'écoulement des masses d'eau à l'intérieur du réseau ; cette étude hydrodynamique préalable est basée sur la localisation des réservoirs ainsi que le tracé des canalisations principales. L'observation d'une analogie entre les deux structures (i. e., abondance bactérienne *vs* agencement des canalisations) permettrait, de fait, une meilleure compréhension des mécanismes régissant la répartition spatiale des bactéries dans un réseau. Plus généralement, une approche similaire, s'appuyant exclusivement sur des considérations liées à l'hydraulique, pourrait, par suite, avantageusement s'appliquer à d'autres études afin de circonscrire a priori les secteurs présentant un risque élevé de contamination bactérienne.

L'infrastructure et la configuration générale du réseau messin, dont une représentation schématique est illustrée sur la figure 5, permettent de connaître les aspects prépondérants de l'écoulement de l'eau dans les canalisations les plus importantes. Cependant, il va de soi que des informations fiables ne pourront pas être obtenues ainsi, au-delà d'un certain niveau de ramification où le sens des courants d'eau devient progressivement

plus incertain, voire même imprévisible. L'étude comparative de la figure 5 avec la figure 9, qui donne une image concrète de la structure spatiale de l'abondance bactérienne, a permis de faire les observations suivantes :

- (i) : la totalité de l'aire d'échantillonnage correspondant à la zone 1, excepté le secteur localisé à l'extrême nord (Devant-les-Ponts) ainsi que quelques communes avoisinantes (Grigy, Peltre, Jury), forme un ensemble compact qui regroupe des stations de prélèvement desservies directement, c'est-à-dire sans stockage intermédiaire, par l'usine de traitement de Moulins. La partie enclavée de la zone 1 située à Devant-les-Ponts doit peut-être son origine à la présence occasionnelle en cet endroit d'un mélange avec de l'eau faiblement contaminée en provenance des usines de Saint-Eloy ou Hauconcourt, et ce malgré les précautions prises pour délimiter le terrain d'études. Ces unités de production localisées plus au nord sont indiquées sur la figure 4. Une autre explication tient probablement au tracé de la canalisation desservant ce secteur (voir figure 5). Cette canalisation, alimentée par le réservoir du Haut de Wacon, se dirige tout d'abord vers le nord-est, avant de former un coude pour ensuite reprendre la direction sud. Il s'ensuit que Le Ban-St-Martin se trouve de fait situé en aval de Devant-les-Ponts ;
- (ii) : les autres zones de densité bactérienne sont, quant à elles, localisées relativement plus loin de l'usine de Moulins et concernent plutôt les parties périphériques de l'agglomération messine (i. e., Plappeville, Le Ban-St-Martin, Vallières, Borny et Queuleu, partiellement). L'eau distribuée dans ces zones a préalablement été stockée dans un réservoir au moins, avant d'atteindre l'aire de distribution ; ceci est notamment le cas pour l'eau distribuée dans les zones 3 et 4, qui a successivement été stockée dans deux réservoirs en série, à savoir : Haute Bevoye-Borny et Haut de Wacon-Plappeville.

Ces remarques montrent clairement que les densités bactériennes élevées sont plutôt observées à la périphérie du réseau, loin de l'usine de traitement, qu'à proximité de celle-ci. Ces régions de forte concentration bactérienne sont associées à des temps de séjour prolongés et coïncident, par ailleurs, avec les secteurs où la quantité de chlore résiduel est la plus faible. Aussi, la structure d'abondance du peuplement bactérien, qui a été mise en évidence dans le cadre de cette étude, est très vraisemblablement largement conditionnée par ces deux variables explicatives dont les modalités sont sans aucun doute empreintes d'une certaine stabilité géographique. Par exemple, l'augmentation du temps de séjour de l'eau dans les canalisations avec l'éloignement par rapport à l'usine de production, s'accompagne d'une diminution de la quantité de chlore résiduel. Malheureusement, la concomitance de ces deux variables, qui est induite par cette relation monotone, ne permet pas de séparer leurs effets propres par la formulation d'un modèle mathématique valable. Cependant, les effets antagonistes de ces deux grandeurs suggèrent des mesures visant à faire progresser le front de chlore résiduel le plus loin possible de l'usine de traitement afin d'éviter les concentrations bactériennes trop importantes à la périphérie du réseau. C'est ainsi que les résultats de notre étude sur le réseau messin ont amené les responsables de la Société Mosellane des Eaux à envisager prochainement l'installation de plusieurs postes de chloration secondaires destinés à réintroduire du chlore en amont des sections les plus contaminées et dans lesquelles celui-ci commençait à faire défaut. A ces moyens supplémentaires il conviendrait néanmoins d'associer un programme de surveillance et d'entretien plus assidu des réservoirs de stockage, qui détiennent selon nous un rôle majeur dans l'accroissement de la densité bactérienne.

L'influence d'une autre variable explicative, en l'occurrence : la température de l'eau, est exprimée de manière globale par le modèle mathématique donné par la formule (46). La validité de ce modèle laisse entrevoir la possibilité de variations cycliques obéissant à des rythmes saisonniers (BECKER, 1978 ; EL-SHAARAWI et al., 1985). Cette variabilité temporelle dont l'ampleur est considérable, n'affecte cependant pas les différentes zones de densité de manière uniforme. L'interaction observée ne pouvant pas être attribuée à des différences de température de l'eau entre les zones, il est vraisemblable que celle-ci soit le fait d'autres variables non identifiées. En tout état de cause, ce phénomène ne fait que s'ajouter à la complexité de fonctionnement du réacteur biologique que constitue un réseau de distribution publique.

Les considérations précédentes, confortées par la configuration et l'étendue géographique des zones de densité bactérienne observées sur le réseau messin semblent indiquer l'inopportunité d'une recherche systématique réalisée en vue de parvenir à la stratification d'un réseau de distribution lorsque la taille de celui-ci est trop modeste. Ainsi, la petitesse des réseaux étudiés par PIPES et CHRISTIAN (1982) constitue peut-être la cause principale de leur échec. Cependant, le choix des bactéries coliformes, dont l'abondance est normalement très faible dans les eaux de distribution, n'est sans doute pas étranger à cet insuccès.

Finalement, la reconnaissance d'une configuration structurée et cohérente de l'hétérogénéité spatiale des populations bactériennes dans le réseau de distribution de la ville de Metz doit être attribuée :

- (i) : à la taille du réseau étudié, qui est suffisante pour permettre la formation d'un gradient stable et perceptible de l'abondance bactérienne ;
- (ii) : à la population cible, à savoir : les bactéries hétérotrophes dont l'abondance relativement élevée autorise la mise en oeuvre de méthodes statistiques efficaces ; et
- (iii) : au mode d'échantillonnage adopté, en l'occurrence un plan d'échantillonnage systématique réitéré de manière à inclure invariablement toutes les parties du réseau.

La structure de l'hétérogénéité bactérienne ainsi mise en évidence permet ultérieurement de concevoir ou d'améliorer les programmes d'échantillonnage effectués pour le contrôle de la qualité bactériologique de l'eau distribuée.

**TROISIEME PARTIE**

**STRATEGIES D'ECHANTILLONNAGE**



La qualité bactériologique de l'eau d'un réseau de distribution publique est habituellement évaluée à l'aide de prélèvements réalisés périodiquement selon un programme d'échantillonnage déterminé. Les résultats de ces prélèvements sont confrontés à des niveaux guides dont la finalité est de garantir la qualité bactériologique de l'eau distribuée. Le manque de données épidémiologiques fiables et la connaissance très imparfaite de l'incidence pathologique sur la population, imputable à la présence des micro-organismes habituellement ou occasionnellement présents dans l'eau "potable", ont contraint les autorités responsables en santé publique à fixer les normes de salubrité sur des bases peu objectives et plutôt arbitraires. Il en est de même pour les modalités des programmes d'échantillonnage dont certains aspects comme la localisation des points de prélèvement ne sont pas précisés alors que d'autres paramètres importants s'appuient sur des considérations dénuées de tout fondement mathématique rigoureux. Ceci est en particulier le cas pour le nombre des prélèvements à effectuer qui, dans certaines législations et notamment la réglementation actuellement appliquée aux Etats-Unis, doit augmenter avec les effectifs de la population desservie selon une relation essentiellement empirique (EPA, 1976).

Les normes de salubrité actuellement en vigueur font généralement état d'un groupe de bactéries indicatrices de contamination fécale (e. g., les bactéries coliformes) dont l'absence est interprétée comme une garantie contre la présence possible dans l'eau de bactéries pathogènes d'origine intestinale. La formulation des normes guides présente généralement deux aspects selon qu'elles fixent a) une valeur maximale de la moyenne arithmétique des titres bactériens observés pour l'ensemble des échantillons analysés pendant une période donnée ou/et b) une proportion maximale d'échantillons dont la concentration bactérienne peut dépasser un seuil déterminé au cours de la même période (EPA, 1976 ; HEALTH and WELFARE CANADA, 1978). Afin de compléter la panoplie des critères de salubrité communément utilisés, des niveaux guides portant sur la numération totale à 20° C ont quelquefois été proposés à titre de tests complémentaires (COLWELL et al., 1978 ; McFETERS et al. ; 1978 ; MEANS et al., 1981). La fixation d'une concentration maximale acceptable pour les bactéries hétérotrophes a été préconisée par GELDREICH (1973). Ainsi, le seuil de 500 bactéries par millilitre a été proposé aux Etats-Unis pour la densité bactérienne moyenne (EPA, 1976) alors que les pays membres de la Communauté Economique Européenne (C.E.E.) ont proposé de fixer ce seuil à 100 bactéries par millilitre à 20° C. En tout état de cause, il apparaît que trop souvent ces normes sont le résultat de compromis plus ou moins arbitraires tentant de concilier des contingences financières et des exigences sanitaires mal définies.

L'élaboration de tout programme d'échantillonnage destiné au contrôle de la qualité bactériologique d'une eau doit obligatoirement spécifier certaines modalités comme : le nombre des prélèvements à effectuer ainsi que leur distribution dans le temps et dans l'espace. Dans la plupart des cas, la décision de conformité ou de violation de la réglementation en vigueur est basée sur l'analyse d'un échantillon, ce qui implique inévitablement le risque d'aboutir à une conclusion erronée. Ce risque présente dans des conditions d'échantillonnage déterminées deux aspects antagonistes. D'une part, l'analyse bactériologique peut conduire à la décision de non-conformité de l'eau avec la réglementation alors qu'en réalité la densité bactérienne moyenne est inférieure à la concentration maximale acceptée. Le risque, noté  $\alpha$ , décrit par cette situation est appelé risque de première espèce ou encore "risque fournisseur". D'autre part, l'observation de la conformité de l'eau avec une spécification donnée alors qu'en réalité la densité bactérienne moyenne dépasse le seuil maximal toléré est associée au risque  $\beta$ , appelé risque de deuxième espèce ou de manière plus expressive "risque consommateur".

La finalité même du contrôle bactériologique de l'eau en réseau

doit amener les autorités responsables à concevoir des plans d'échantillonnage dont les modalités assurent le maintien du risque  $\beta$  au-dessous d'un seuil prédéterminé compte tenu des fluctuations possibles de certains paramètres déterminants.

Ces considérations montrent que les résultats bactériologiques doivent toujours être interprétés avec la souplesse et la prudence inhérentes au contexte probabiliste qui les accompagne. A cet égard, il faut savoir que la probabilité d'observer la conformité d'un échantillon quelconque avec une spécification particulière de la réglementation dépend (i) de la densité bactérienne moyenne réelle ( $\lambda$ ) de l'eau distribuée, (ii) du nombre  $n$  d'échantillons analysés, et (iii) de la variance caractérisant l'ensemble des observations possibles (ESTERBY, 1982 ; PIPES et CHRISTIAN, 1982). Cette variance, qui résulte à la fois des fluctuations réelles de la concentration bactérienne du milieu et de l'incertitude sur la valeur même du résultat d'analyse, est exprimée par les paramètres de dispersion des lois de distribution suivies par les titres bactériens. Ces remarques montrent pourquoi il n'est pas souhaitable de vouloir établir une correspondance entre le nombre d'échantillons à prélever et la taille de la population desservie comme cela est actuellement mentionné dans la réglementation applicable aux Etats-Unis (EPA, 1976). Mieux vaut en effet, pour des raisons de rigueur scientifique, bâtir un plan d'échantillonnage en fonction de la nature et de la formulation explicite des modèles de distribution bactérienne caractérisant le milieu étudié (EL-SHAARAWI et al., 1985). Notons toutefois qu'il est vraisemblable que l'effectif de la population desservie puisse, d'une certaine manière, être corrélé avec le degré d'hétérogénéité d'un système de distribution ; mais dans ce cas il ne s'agit que d'un facteur indirect permettant d'arrêter les modalités du plan d'échantillonnage.

Les développements théoriques exposés dans le premier chapitre de cette partie ont pour but de montrer comment il est possible de définir un programme d'échantillonnage efficace pour le contrôle de la qualité bactériologique de l'eau distribuée. Ce programme doit être établi à partir de données obtenues antérieurement et permettant de préciser les modèles de distribution des populations bactériennes du réseau considéré. Deux approches seront étudiées successivement, selon que l'échantillonnage est réalisé au cours d'une seule et même campagne de prélèvements (bilan instantané) ou dans le cadre d'un échantillonnage progressif à but décisionnel (bilan échelonné).

Dans le premier cas, les différents aspects caractérisant la structure de l'hétérogénéité bactérienne seront exploités afin de : (i) déterminer la localisation et le nombre des prélèvements à effectuer au cours d'une même campagne pour une valeur donnée du risque  $\beta$ , et (ii) pouvoir procéder à la répartition optimale de l'effort d'échantillonnage dans les zones de densité bactérienne lorsque le nombre total des stations à échantillonner a préalablement été fixé.

Le deuxième mode de prélèvement consiste à fragmenter l'effort d'échantillonnage en planifiant la collecte des données de manière discontinue et séquentielle. L'information est alors interprétée au fur et à mesure de son obtention et l'expérience est arrêtée dès qu'une décision, au regard de la normalisation, peut être prise moyennant des risques d'erreurs fixés.

Dans le deuxième chapitre, ces plans d'échantillonnage théoriques seront appliqués aux données bactériologiques recueillies au cours des six campagnes de prélèvements réalisées sur le réseau de la ville de Metz.

## CHAPITRE I

PROPOSITION THEORIQUE D'UNE STRATEGIE DE  
PRELEVEMENT POUR LE CONTROLE DE LA  
QUALITE BACTERIOLOGIQUE DE L'EAU DISTRIBUEE

## 1.1. ECHANTILLONNAGE POUR UNE EVALUATION INSTANTANEE DE LA QUALITE BACTERIOLOGIQUE DE L'EAU

### 1.1.1. Détermination du nombre d'échantillons à prélever

Dans toute cette section la distribution des micro-organismes dans le système considéré est supposée obéir à une loi binomiale négative de paramètres  $k$  et  $p$ . La moyenne de la distribution BN ( $k,p$ ) est égale à  $\lambda = p k$ .

D'après la proposition 2 énoncée au paragraphe 1.2.1. de la première partie, la loi de probabilité suivie par la somme de  $n$  variables aléatoires indépendantes, obéissant toutes à une même loi binomiale négative de paramètres  $k$  et  $p$ , est la loi binomiale négative définie par les paramètres  $n$  et  $p$ . En d'autres termes, la loi de distribution associée à la somme  $(\sum_{i=1}^n r_i)$ , des réalisations  $r_1, r_2, \dots, r_n$  de  $n$  variables indépendantes de la loi BN ( $k,p$ ), est définie par l'expression

$$P\left(\sum_{i=1}^n r_i = r\right) = \frac{(nk + r - 1)!}{r! (nk - 1)!} \frac{p^r}{(1+p)^{nk+r}} \quad (47)$$

Cette propriété fondamentale est à la base de tous les développements exposés ci-après.

La capacité d'un plan d'échantillonnage à mettre en évidence les transgressions de la réglementation peut être appréciée par le calcul de la probabilité d'acceptation (PA) de la conformité de l'eau avec une norme de référence. Cette probabilité est une fonction des paramètres de la distribution suivie par les échantillons, du nombre  $n$  des analyses effectuées et de la spécification en vigueur. Dans le cas où la densité bactérienne réelle de l'eau dépasse la norme guide, PA s'identifie au risque de deuxième espèce  $\beta$  alors que la différence  $1-PA$  correspond au risque  $\alpha$  lorsque la qualité bactériologique réelle de l'eau vérifie la réglementation.

D'une manière plus formelle, si l'évaluation de la qualité bactériologique de l'eau repose sur la densité bactérienne moyenne  $\bar{r}$  des échantillons analysés et si en outre, le critère de qualité retenu spécifie une concentration bactérienne moyenne devant rester inférieure à  $\mu$ , la condition  $\bar{r} \leq \mu$  peut être exprimée de façon équivalente par l'inégalité

$$\sum_{i=1}^n r_i \leq n \mu. \text{ Or d'après la formule (47), il résulte immédiatement que la}$$

probabilité PA est obtenue par l'expression

$$PA = \sum_{r=0}^{\mu n} \frac{(nk + r - 1)!}{r! (nk - 1)!} \frac{p^r}{(1+p)^{nk+r}} \quad (48)$$

(EL-SHAARAWI et al., 1985)

Une simplification appréciable de la formule (48), est apportée par l'utilisation du théorème central limite dont l'énoncé stipule en substance que la distribution de la somme des  $n$  titres bactériens,

$\sum_{i=1}^n r_i$ , est approximativement normale lorsque le nombre  $n$  des échantillons considérés est suffisamment élevé.

Ainsi, la probabilité PA est avantageusement approchée par :

$$PA = \Phi \left( \frac{\mu - \lambda}{\sqrt{\lambda/n} \sqrt{1+\lambda/k}} \right) \quad (49)$$

où  $\Phi(z)$  représente l'aire de la partie du plan située sous la courbe de la distribution normale centrée réduite entre  $-\infty$  et  $z$  tandis que  $\lambda = p k$  désigne la moyenne de la loi binomiale négative de paramètres  $k$  et  $p$ .

La figure 16 donne la représentation graphique de la probabilité PA (calculée selon les formules (48) ou (49), pour plusieurs valeurs du paramètre  $k$  et du nombre d'échantillons  $n$ , en fonction de la densité bactérienne moyenne réelle de l'eau ( $\lambda$ ). Cette figure, qui a été conçue pour une norme guide  $\mu = 100$  bactéries par millilitre, donne également une interprétation graphique des risques  $\alpha$  et  $\beta$ .

A partir de la relation (49), il est facile de montrer que dans l'éventualité d'une concentration bactérienne moyenne  $\lambda$  supérieure à  $\mu$ , le nombre d'échantillons à prélever, permettant de déceler ce phénomène avec un risque d'erreur de deuxième espèce  $\beta$  préalablement fixé, doit satisfaire l'inégalité

$$n \geq \frac{\lambda z_0^2 (1+\lambda/k)}{(\mu-\lambda)^2} \quad (50)$$

dans laquelle  $z_0$  est la valeur limite de la variable normale centrée réduite associée au seuil de probabilité  $\beta$ .

L'examen de la formule (50) montre que le nombre minimal d'échantillons à recueillir et la valeur de  $k$  varient en sens inverse. Le paramètre  $k$  qui traduit le degré d'hétérogénéité d'un système peut dès lors, et à ce titre, être considéré comme une mesure de l'entropie du système étudié (cf. proposition 4, § 1.2.1., première partie). Les grandes valeurs de  $k$  reflètent des situations de plus en plus proches du modèle de Poisson (cf. proposition 1, § 1.2.1., première partie) et sont à l'origine de conditions d'échantillonnage devenant progressivement plus avantageuses : les conditions idéales de prélèvement étant réalisées au sein d'un milieu parfaitement homogène modélisé par la distribution de Poisson. Les résultats émanant de l'étude théorique de ce cas limite, peu réaliste au-delà d'une certaine masse d'eau, sont obtenus en faisant tendre  $k$  vers l'infini dans les formules précédentes.

Réciproquement, le risque  $\beta$ , inhérent aux circonstances et aux modalités d'un programme de prélèvement donné, est calculé à l'aide de la relation (48) ou de son approximation conformément à la formule (49). Il s'ensuit, moyennant cette dernière expression de  $\beta$ , que si  $n(k_i, \lambda, \beta)$  désigne le nombre d'échantillons à recueillir dans un système régi par la loi BN ( $k_i, p_i$ ), avec  $\lambda = p_i k_i$ , l'effort d'échantillonnage relatif à consentir dans deux systèmes distincts doit vérifier la formule

$$n(k_2, \lambda, \beta) = n(k_1, \lambda, \beta) \cdot \frac{k_1 (\lambda+k_2)}{k_2 (\lambda+k_1)} \quad (51)$$

afin de maintenir le même niveau du risque  $\beta$  dans les deux systèmes considérés. L'intérêt essentiel de la relation (51) est de permettre la comparaison de plusieurs systèmes caractérisés par des valeurs de  $k$  différentes (EL-SHAARAWI et al., 1985).

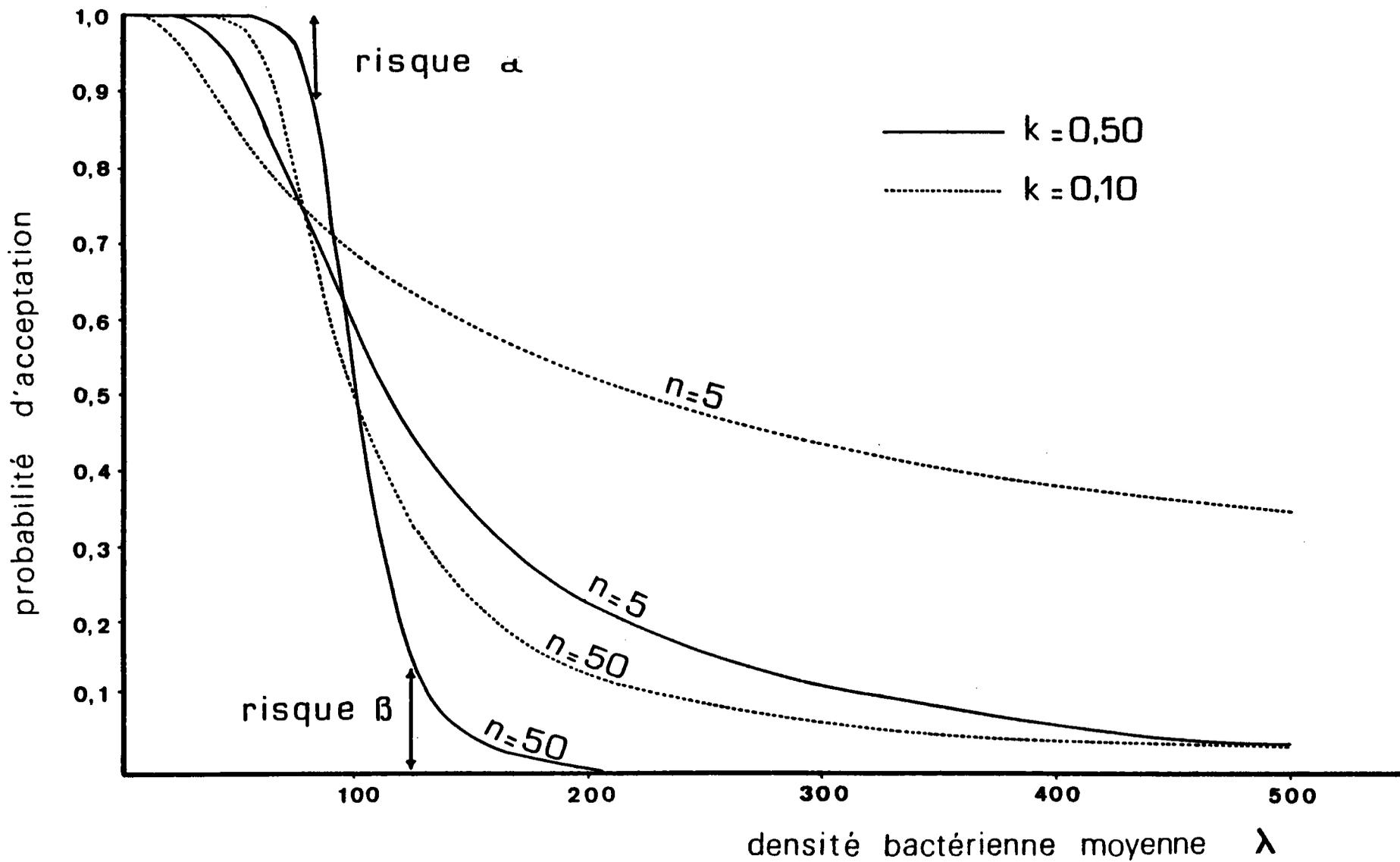


FIGURE 16 : Probabilité d'acceptation en fonction de la densité bactérienne moyenne.

### 1.1.2. Répartition optimale de l'effort d'échantillonnage dans les zones d'un réseau stratifié

Plus généralement, lorsqu'un réseau de distribution publique peut faire l'objet d'une partition en  $\ell$  zones pouvant chacune être décrite par une loi binomiale négative de paramètre  $k$  connu, les formules précédentes autorisent le calcul du nombre d'échantillons à recueillir dans chaque zone de densité bactérienne.

Ainsi, lorsque le niveau du risque  $\beta$  accepté est imposé pour chacune de ces zones, le nombre  $n_i$  d'échantillons à prélever dans la  $i^{\text{ème}}$  zone est calculé directement d'après l'inégalité (50).

Il peut cependant arriver que des contraintes administratives ou budgétaires par exemple dictent le nombre total  $N$  des échantillons à prélever pour l'ensemble du réseau. La planification des études ultérieures consiste dans ce cas à répartir les  $N$  échantillons dans les différentes zones de manière à contrôler uniformément le risque de deuxième espèce sur toute l'étendue du réseau. Il est alors facile d'établir au moyen de la formule (51) que la répartition optimale des  $N$  échantillons est donnée conformément à l'égalité

$$n_i = N \frac{1 + \lambda/k_i}{\sum_{i=1}^{\ell} (1 + \lambda/k_i)} \tag{52}$$

où  $k_i$  représente la valeur du paramètre  $k$  de la loi binomiale négative caractérisant la zone  $i$ .

Finalement, le risque  $\beta$  associé à un tel mode d'échantillonnage est exprimé par

$$\beta = \Phi \left( \frac{\mu - \lambda}{\sqrt{\lambda/N} \sqrt{\sum_{i=1}^{\ell} (1 + \lambda/k_i)}} \right) \tag{53}$$

Ce risque peut également être calculé à l'aide de la formule (48) ou (49) pour l'une quelconque des  $\ell$  zones considérées.

### 1.2. ECHANTILLONNAGE PROGRESSIF A BUT DECISIONNEL

L'échantillonnage progressif constitue un autre type de prélèvement conçu en fonction d'une modélisation (DUBY et RODOLPHE, 1982). Dans ce mode d'échantillonnage particulier, la taille de l'échantillon n'est pas déterminée à l'avance puisqu'elle dépendra du modèle utilisé mais aussi des résultats des observations faites graduellement au cours de l'échantillonnage. La méthode peut être utilisée pour les contrôles de conformité de la qualité bactériologique d'une eau par rapport à une norme de salubrité lorsque, d'une part, le système étudié est modélisé par une loi binomiale négative, et d'autre part, le programme de surveillance fait état de bilans périodiques intéressants des intervalles de temps donnés (e. g., bilans mensuels).

Sans entrer dans le détail des aspects techniques de ces plans que l'on pourra trouver dans le livre de WALD (1947), initiateur de ces méthodes,

les principales étapes de l'échantillonnage progressif à but décisionnel se déroulent de la façon suivante :

Soit à comparer la moyenne ( $\lambda$ ) de la distribution des titres bactériens à deux valeurs théoriques  $\lambda_0$  et  $\lambda_1$ , avec  $\lambda_0 < \lambda_1$ . Les valeurs de  $\lambda_0$  et  $\lambda_1$  pourront, par exemple, être respectivement égales à 100 et 200.

La méthode permet de choisir entre deux hypothèses :

$$\begin{aligned} H_0 &: \lambda \leq \lambda_0 \\ H_1 &: \lambda \geq \lambda_1 \end{aligned}$$

à partir des résultats de  $n$  observations  $r_1, r_2, \dots, r_n$  obtenues au hasard et de manière séquentielle. La valeur de  $n$  n'est pas fixée, cependant l'effectif des échantillons est limité au nombre d'observations juste suffisant pour prendre une décision et non plus déduit des risques d'erreurs admis, bien que ceux-ci doivent toujours être fixés à l'avance.

A l'étape  $n$ , la décision est basée sur la valeur du rapport de vraisemblance

$$R_n = \frac{\prod_{i=1}^n P_1(r_i)}{\prod_{i=1}^n P_0(r_i)} \quad (54)$$

où les probabilités  $P_0(r_i)$  et  $P_1(r_i)$  sont définies par les lois de distribution des titres bactériens correspondant à l'hypothèse  $\lambda = \lambda_0$ , respectivement  $\lambda = \lambda_1$ .

WALD a montré que :

- pour  $R_n < \frac{\beta}{1-\alpha}$  le risque d'erreur est inférieur à  $\beta$  si l'hypothèse  $H_0$  ( $\lambda \leq \lambda_0$ ) est retenue alors qu'en réalité c'est l'hypothèse  $H_1$  ( $\lambda \geq \lambda_1$ ) qui est exacte ;
- pour  $R_n > \frac{1-\beta}{\alpha}$  le risque d'erreur est inférieur à  $\alpha$  si l'on retient l'hypothèse  $H_1$  alors qu'en réalité c'est l'hypothèse  $H_0$  qui est vérifiée.

Le rapport  $R_n$ , qui est calculé après chaque étape de l'échantillonnage, est comparé aux limites

$\frac{\beta}{1-\alpha}$  et  $\frac{1-\beta}{\alpha}$  avant de procéder éventuellement à l'analyse d'échantillons supplémentaires. De manière plus explicite :

- si  $R_n < \frac{\beta}{1-\alpha}$  l'hypothèse  $H_0$  ( $\lambda \leq \lambda_0$ ) est acceptée ;
- si  $R_n > \frac{1-\beta}{\alpha}$  l'hypothèse  $H_1$  ( $\lambda \geq \lambda_1$ ) est acceptée ;
- enfin si  $\frac{\beta}{1-\alpha} < R_n < \frac{1-\beta}{\alpha}$ , l'échantillonnage doit être poursuivi jusqu'à l'obtention du premier rapport  $R$  dont la valeur est extérieure à ces limites.

Dans la pratique, l'écriture de la double inégalité ci-dessus est mise sous la forme

$$A(n) < f_n(r_1, r_2, \dots, r_n) < B(n)$$

où  $f_n$  désigne une fonction des  $n$  premières observations. Les représentations graphiques des fonctions définies par  $A(n)$  et  $B(n)$  constituent les courbes d'arrêt de l'échantillonnage. Après chaque nouvelle observation, le point représentant  $f_n$  en fonction de  $n$  est porté sur le graphique. L'échantillonnage est poursuivi aussi longtemps que ce point reste situé entre les deux courbes. Dès que la trajectoire du point ainsi obtenue franchit la courbe d'arrêt supérieure, l'hypothèse  $H_1$  est acceptée ; à l'inverse, si celle-ci atteint la courbe d'arrêt inférieure c'est  $H_0$  qui est retenue. Il faut cependant noter que le nombre d'observations, nécessaire à l'aboutissement du test progressif, peut, dans certains cas, être relativement important ; aussi, il est quelquefois commode de fixer préalablement une taille maximale de l'échantillon à ne pas dépasser.

Lorsque le paramètre  $k$  de la loi binomiale négative servant de modèle pour la distribution des micro-organismes dans le système étudié est connu, il découle immédiatement des formules (11) et (54) que

$$\ln R_n = \ln \left[ \frac{\lambda_1 (k+\lambda_0)}{\lambda_0 (k+\lambda_1)} \right] \cdot \sum_{i=1}^n r_i + n k \ln \frac{k + \lambda_0}{k + \lambda_1} \quad (55)$$

Moyennant cette expression de  $\ln R_n$  et en remplaçant  $\sum_{i=1}^n r_i$  par  $T_n$  dans la double inégalité  $\ln \frac{\beta}{1-\alpha} < \ln R_n < \ln \frac{1-\beta}{\alpha}$ , il vient :

$$A(n) < T_n < B(n)$$

$$\text{avec } A(n) = \frac{\ln(\beta/1-\alpha)}{\ln \left[ \frac{\lambda_1 (k+\lambda_0)}{\lambda_0 (k+\lambda_1)} \right]} - \frac{n k \ln(k+\lambda_0/k+\lambda_1)}{\ln \left[ \frac{\lambda_1 (k+\lambda_0)}{\lambda_0 (k+\lambda_1)} \right]} \quad (56)$$

et

$$B(n) = \frac{\ln(1-\beta/\alpha)}{\ln \left[ \frac{\lambda_1 (k+\lambda_0)}{\lambda_0 (k+\lambda_1)} \right]} - \frac{n k \ln(k+\lambda_0/k+\lambda_1)}{\ln \left[ \frac{\lambda_1 (k+\lambda_0)}{\lambda_0 (k+\lambda_1)} \right]} \quad (57)$$

En résumé, l'échantillonnage progressif est interrompu et une décision peut être prise dès que l'effectif cumulé  $T_n$  des résultats obtenus lors des  $n$  premières mesures de l'expérience se trouve à l'extérieur de l'intervalle défini par les nombres  $A(n)$  et  $B(n)$ .

A titre d'information, notons que des formules analogues se rapportant au modèle de Poisson peuvent être calculées soit directement, soit à partir des relations (56) et (57) en faisant tendre  $k$  vers l'infini (cf. proposition 1, § 1.2.1., première partie). Celles-ci deviennent alors respectivement :

$$A(n) = \frac{\ln(\beta/1-\alpha)}{\ln \lambda_1/\lambda_0} - \frac{(\lambda_0-\lambda_1)n}{\ln \lambda_1/\lambda_0} \quad (58)$$

et

$$B(n) = \frac{\ln(1-\beta/\alpha)}{\ln \lambda_1/\lambda_0} - \frac{(\lambda_0-\lambda_1)n}{\ln \lambda_1/\lambda_0} \quad (59)$$

## CHAPITRE II

APPLICATION DES PLANS D'ECHANTILLONNAGE

THEORIQUES AUX DONNEES BACTERIOLOGIQUES

DU RESEAU MESSIN

Les développements théoriques exposés dans le chapitre précédent seront maintenant illustrés par les données obtenues au cours de six campagnes d'échantillonnage systématique effectuées sur le réseau messin dans le but de parvenir à une modélisation de la structure d'abondance des populations bactériennes. Il faut rappeler que ces investigations ont permis de stratifier le réseau de la ville de Metz en le faisant apparaître comme un ensemble composite formé de quatre systèmes hétérogènes (i. e., modélisés par une loi binomiale négative). La structure ainsi dégagée sera mise à profit pour définir et optimiser les modalités des programmes d'échantillonnage ultérieurs sur la base de l'information disponible.

## 2.1. STRATEGIES D'ECHANTILLONNAGE EN FONCTION D'UNE MODELISATION

### 2.1.1. Bilan instantané

Les considérations théoriques de la section 1.1. de cette troisième partie ont surtout fait ressortir le rôle tout à fait prépondérant du paramètre  $k$  associé au modèle binomial négatif observé. La valeur de ce paramètre, qui constitue de fait une mesure de l'entropie du système considéré, détermine l'importance de l'effort d'échantillonnage à consentir compte tenu du risque d'erreur de deuxième espèce  $\beta$  accepté. Dans cette perspective, il devient particulièrement intéressant de vouloir étudier le degré d'homogénéité des valeurs de  $k$  obtenues au sein d'une même zone de densité bactérienne à l'occasion de plusieurs campagnes de prélèvements. A l'origine de cette démarche, il convient de citer une étude statistique, réalisée rétrospectivement par EL-SHAARAWI et al. (1985), portant sur les données bactériologiques obtenues lors des programmes routiniers de surveillance menés sur le réseau de distribution de la ville de Nancy au cours des années 1981 et 1982. Les conclusions de cette étude font notamment état du caractère invariable et intrinsèque du paramètre  $k$  correspondant aux distributions binomiales négatives ajustées avec succès aux différents bilans mensuels.

Pour chaque zone de densité bactérienne mise en évidence dans le réseau messin (cf. figure 9), le tableau 14 donne les valeurs  $\hat{k}_i$  (cf. équation (15)) correspondant à chacune des six campagnes d'échantillonnage, ainsi que l'estimation  $\hat{k}_c$  de la valeur de  $k$  commune aux six séries expérimentales. La valeur  $\hat{k}_c$ , calculée selon l'équation (17), repose sur l'hypothèse nulle faisant état de l'égalité des  $k_i$  pour  $i$  allant de 1 à 6. Les résultats du test d'homogénéité des  $k_i$  sont présentés dans la dernière colonne du tableau à l'exception de la zone 4 qui comprend un nombre d'échantillons jugé trop faible. Le test de comparaison a été effectué conformément à la formule (26) et la procédure décrite au paragraphe 1.2.3.1. de la première partie.

La démarche exposée ci-après a pour but d'associer une valeur du paramètre  $k$  à chaque zone de densité bactérienne, à partir de l'information recueillie au cours des campagnes de prélèvements réalisées antérieurement. Bien que le principe de la méthode proposée n'ait rien d'absolu, le critère utilisé constitue cependant une base de travail satisfaisante. En effet, l'estimation du paramètre  $k$ , retenue à l'issue de cette méthode, est obtenue conformément à l'intention de situer sa valeur à un niveau minimal ; toute valeur réelle de  $k$ , supérieure à cette estimation, se traduit par conséquent par une surestimation du nombre des échantillons à prélever, ce qui, a fortiori, contribue à assurer le maintien du risque de deuxième espèce au-dessous du seuil fixé à l'avance.

TABLEAU 14

ESTIMATION DE LA VALEUR DE  $k$  POUR CHAQUE ZONE  
ET CHAQUE CAMPAGNE D'ECHANTILLONNAGE, ESTIMATION DE  $k_c$   
POUR CHAQUE ZONE ET RESULTATS DU TEST  
D'HOMOGENEITE DES VALEURS  $k_i$

ZONE	CAMPAGNE DE PRELEVEMENTS	$\hat{k}_i$	$\hat{k}_c$	SIGNIFICATION DU TEST DE COMPARAISON DES $k_i$
1	13 décembre 10 janvier 21 février 03 avril 15 mai 26 juin	0,34877 0,35324 0,34683 0,53659 0,67527 0,46982	0,43927	$P \approx 0,100$
2	13 décembre 10 janvier 21 février 03 avril 15 mai 26 juin	0,54583 0,31856 0,40579 0,92280 0,95253 1,32717	0,65823	$P < 0,001$
3	13 décembre 10 janvier 21 février 03 avril 15 mai 26 juin	0,52371 0,52041 0,44561 0,34745 0,92227 8,81069	0,84629	$P < 0,001$
4	13 décembre 10 janvier 21 février 03 avril 15 mai 26 juin	71,78516 874,22991 27,54602 7,18284 59,05907 7,01250	41,54600	

L'hypothèse nulle (i.e., égalité des  $k_i$ ) a été retenue pour la zone 1 uniquement, ce qui a permis de lui associer la valeur 0,43927. Cependant, lorsque l'hypothèse d'égalité des paramètres  $k_i$  est rejetée pour un ensemble de  $l$  séries expérimentales, il devient intéressant de chercher un sous-ensemble de  $m$  ( $m < l$ ) séries expérimentales ajustées par  $m$  lois binomiales négatives possédant un paramètre  $k$  commun. C'est ainsi qu'une valeur commune de  $k$  a pu être assignée à la zone 2 de même qu'à la zone 3 conformément à l'algorithme suivant :

- (i) les  $l$  valeurs  $\hat{k}_i$  sont rangées par ordre croissant ;
- (ii) le test de comparaison est effectué, avec un risque d'erreur prédéterminé, sur les distributions empiriques correspondant aux deux premières valeurs de  $k$  ainsi ordonnées ;
- (iii) si l'égalité de ces deux valeurs est acceptée, le test est étendu aux trois premières séries expérimentales ;
- (iv) les termes de la suite ordonnée des  $\hat{k}_i$  sont ainsi pris en considération un à un et l'itération est poursuivie tant que le test aboutit à un résultat non significatif ;

(v) dès qu'une différence significative entre les valeurs  $\hat{k}_i$  comparées est observée, le processus est interrompu et une estimation de  $k_c$  est calculée en résolvant l'équation (17) appliquée aux séries expérimentales considérées à ce stade, à l'exception de la dernière.

Enfin, la plus petite valeur  $\hat{k}_i$  observée (i.e., 7,01250) a été choisie pour caractériser la zone 4.

Lorsque chaque zone est représentée par une valeur de  $k$  spécifique, la formule (50) est utilisée pour déterminer le nombre des stations d'échantillonnage qu'il faudrait attribuer à chacune de ces zones. Ce nombre est, par ailleurs, calculé en fonction d'un risque  $\beta$ , d'une norme  $\mu$  et d'une valeur  $\lambda$  prédéterminés. A titre d'exemple et pour des raisons de commodité, la norme  $\mu$  a été fixée, dans la suite de cet exposé, à 100 bactéries hétérotrophes par millilitre, en accord avec la réglementation préconisée au sein des pays membres de la C.E.E.. La figure 17 donne une illustration graphique de la variation du nombre minimal d'échantillons à recueillir ( $n$ ) en fonction de la valeur du paramètre  $k$ , et cela pour différents niveaux de  $\lambda$  et du risque  $\beta$ .

Si, dans le cadre d'une approche sensiblement différente, le nombre total des échantillons à prélever est contingenté, la répartition optimale d'un nombre  $N$  d'échantillons est déterminée à l'aide de la formule (52). Cette répartition est compatible avec une distribution équilibrée du risque  $\beta$  dans l'ensemble des zones composant le système de distribution. Le risque  $\beta$  commun à toutes les zones peut alors être calculé selon la formule (53) ou encore être obtenu par simple lecture graphique au moyen des courbes représentées sur la figure 18. Ces courbes, qui correspondent à différentes valeurs de  $k$  alors que  $\lambda$  a préalablement été fixé à 200, traduisent les variations du risque  $\beta$  en fonction de  $n$ .

Le tableau 15 présente les résultats des calculs précédents qui ont été effectués pour les quatre zones de densité bactérienne délimitées sur la figure 9.

TABLEAU 15  
REPARTITION OPTIMALE DES STATIONS D'ECHANTILLONNAGE  
ENTRE LES DIFFERENTES ZONES DU RESEAU DE DISTRIBUTION PUBLIQUE  
DE LA VILLE DE METZ

ZONE	$k_c$	NOMBRE DES STATIONS D'ECHANTILLONNAGE		
		$n$	$\tilde{n}$	$\hat{n}$
1	0,43927	52	24,70	35,96
2	0,50514	31	21,48	31,28
3	0,51945	15	20,89	30,42
4	7,01250	3	1,59	2,33

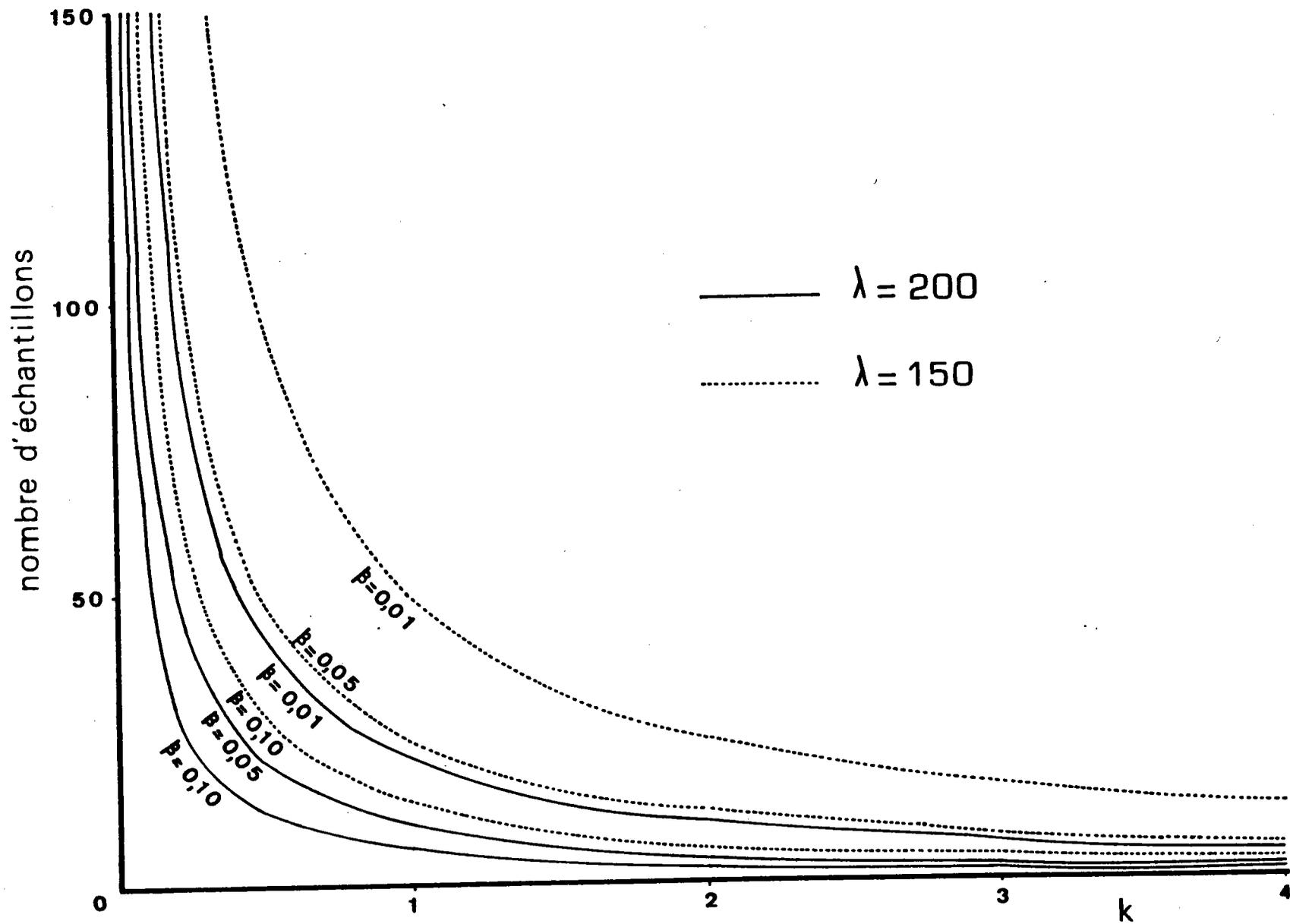


FIGURE 17 : Nombre d'échantillons à prélever en fonction de  $k$ , pour différents niveaux de la densité bactérienne  $\lambda$  et du risque  $\beta$ .

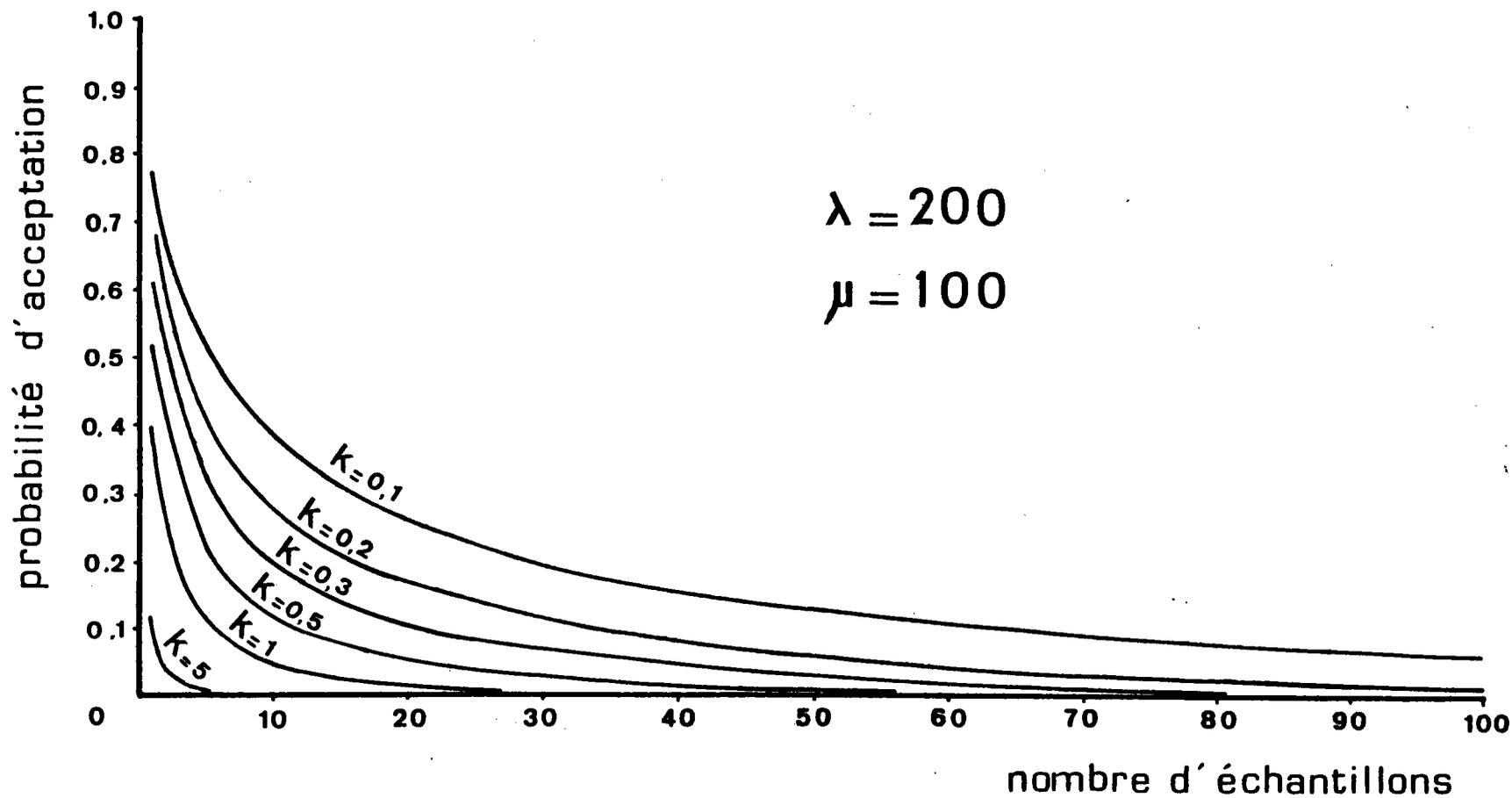


FIGURE 18 : Probabilité d'acceptation (risque  $\beta$ ) en fonction du nombre d'échantillons analysés.

En prenant pour valeur de référence  $\lambda = 200$ , la colonne introduite par le symbole  $\tilde{n}$  donne l'effectif des échantillons à prélever dans chaque zone pour un risque  $\beta$  égal à 0,05. La colonne surmontée par  $\hat{n}$  présente la répartition optimale d'un lot composé de 100 échantillons ; dans ce cas, le risque  $\beta$  d'aboutir à une conclusion inexacte est sensiblement égal à 0,0235. Conformément à la méthode suggérée par EL-SHAARAWI (1982), il est possible d'évaluer l'efficacité du plan d'échantillonnage systématique mis en oeuvre pour le réseau messin en comparant les valeurs de  $\tilde{n}$  ou  $\hat{n}$  avec les nombres d'échantillons effectivement recueillis durant les six campagnes de prélèvements réalisés en 1983 et 1984. Ces nombres figurent dans le tableau 15 sous la lettre n.

Les résultats reportés dans le tableau 15 montrent: (i) que 70 stations au moins doivent être échantillonnées dans le réseau étudié afin de pouvoir déceler un manque de conformité de la qualité de l'eau au regard de la réglementation avec une probabilité de 0,95, lorsque la densité bactérienne moyenne  $\lambda$  est égale à 200 et, (ii) que le plan d'échantillonnage optimisé, correspondant à un risque  $\beta = 0,05$ , se traduit par une réduction du nombre des stations de prélèvement dans les deux premières zones et une légère augmentation du nombre des stations dans la zone 3.

### 2.1.2. Bilan échelonné

Bien que les expériences menées à Metz en 1983 et 1984 n'aient pas été effectuées selon un programme d'échantillonnage progressif, les résultats observés pour la zone 1, qui est caractérisée par une estimation de  $k_C$  égale à 0,43927, ont été utilisés afin d'illustrer la méthode par simulation. La figure 19 montre le profil des courbes obtenues d'après les formules (56) et (57) dans le cas où les risques  $\alpha$  et  $\beta$  ont tous deux été fixés à 0,05. Ces courbes, qui correspondent à plusieurs systèmes de formulation des hypothèses  $H_0$  et  $H_1$ , permettent de distinguer des régions du plan associées à différents niveaux de la concentration bactérienne moyenne. La figure 19 donne également la trajectoire des points de coordonnées  $(n, \log T_n)$  où  $T_n$  représente l'effectif cumulé des titres bactériens tirés un à un et de manière aléatoire dans l'ensemble des résultats observés pour la zone 1 au cours de la première et de la sixième campagne de prélèvements. L'échantillonnage fictif ainsi défini a été poursuivi jusqu'à l'obtention d'un point situé à l'extérieur de la région d'incertitude comprise entre deux courbes adjacées. Ainsi, la conformité de la qualité bactériologique de l'eau, eu égard à la réglementation appliquée, a pu être statuée à partir du 11ème échantillon pour ce qui concerne les résultats provenant de la campagne du 13 décembre alors que les valeurs de l'expérimentation du 26 juin ont conduit à reconnaître le dépassement de la norme guide à compter du 28ème échantillon pris en considération. A titre d'exemple, la même procédure mise en oeuvre pour les données recueillies lors des autres campagnes de prélèvements, à savoir : le 10 janvier, le 21 février, le 3 avril et le 15 mai, a permis de conclure au respect de la réglementation (i.e., 100 bactéries par millilitre) respectivement à partir de la 12ème, 22ème, 14ème et 14ème étape.

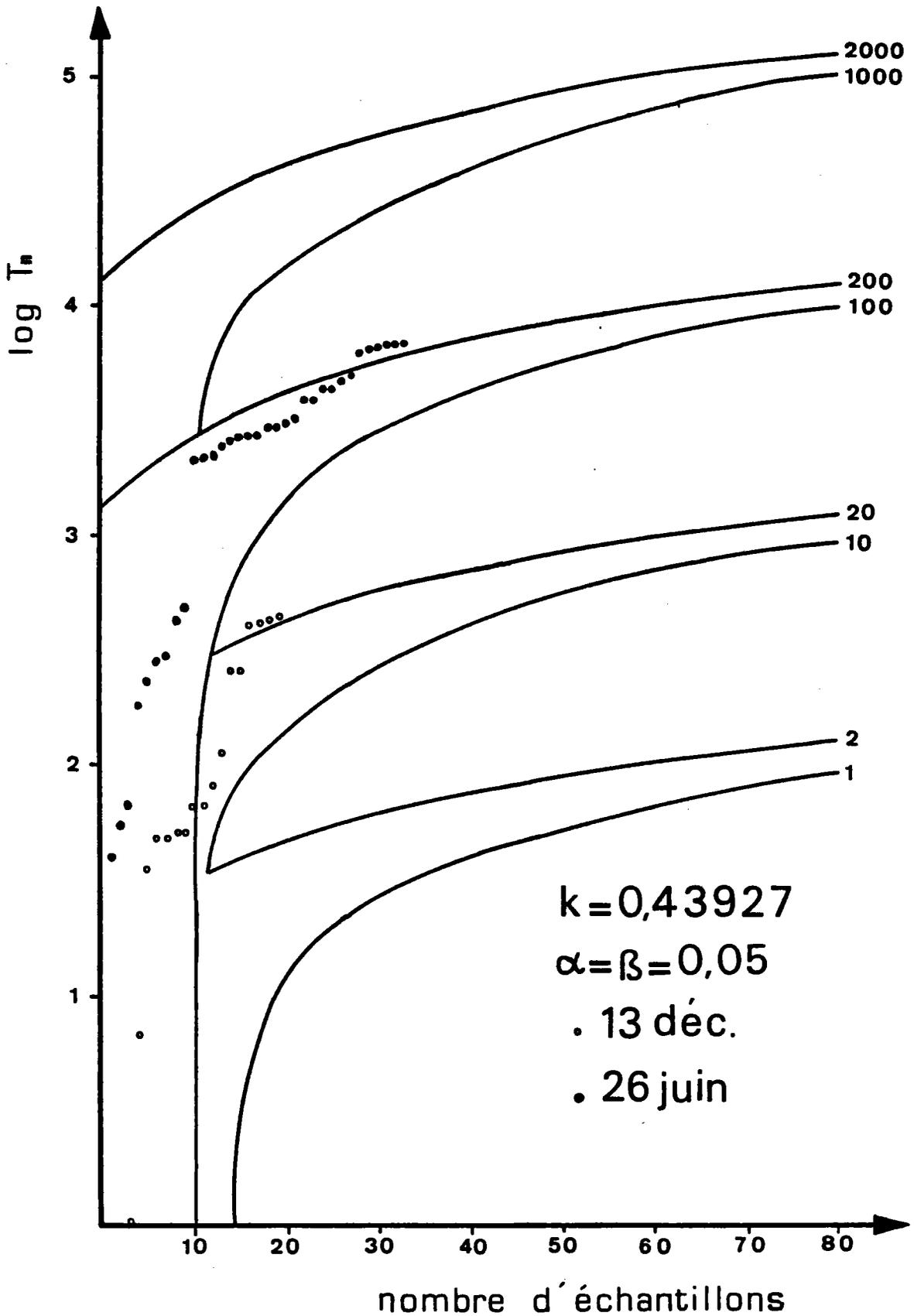


FIGURE 19 : Illustration de la méthode d'échantillonnage progressif pour la zone 1 à partir des données observées lors des campagnes de prélèvements du 13 décembre et du 26 juin.

## 2.2. DISCUSSION

Les deux modes de prélèvement présentés à la section 2.1 correspondent à des conceptions très différentes de l'échantillonnage. Ces deux approches, qui répondent à des objectifs particuliers, ont clairement fait ressortir le rôle et l'incidence de l'hétérogénéité du peuplement bactérien sur les modalités des programmes de surveillance. Dans les sous-systèmes composant le réseau étudié, cette hétérogénéité a été exprimée par le paramètre  $k$  dont la signification peut, d'une certaine manière, être rapprochée du concept d'entropie des physiciens, soit une mesure du degré de désorganisation des systèmes considérés, ou encore de la notion d'entropie utilisée dans la théorie mathématique de l'information.

La constance relative de  $k$ , observée pour la zone 1, traduit une plus grande stabilité de la distribution spatiale des micro-organismes à l'intérieur de cette zone comparativement aux autres parties du réseau où des variations significatives de  $k$  ont été mises en évidence au cours des six campagnes de prélèvements. Cette observation, essentiellement quantitative, trouve une interprétation plausible dans les faibles distorsions affectant l'aire géographique de la zone 1 au fil des expérimentations, et ce, en particulier au regard des zones 2 et 3 dont les contours ont fait l'objet de déformations plus importantes. Les disparités consécutives au manque de cohérence éventuel des valeurs de  $k$  caractérisant une zone donnée, devraient normalement s'accompagner d'un réajustement continu et approprié de l'effort d'échantillonnage en vue de maintenir le risque d'erreur  $\beta$  à un niveau constant pour tous les contrôles instantanés de la qualité bactériologique de l'eau. Il semble toutefois préférable, compte tenu de certaines contraintes administratives ou plus simplement d'ordre pratique, de fixer le nombre d'échantillons à prélever de manière à contrôler le risque maximal au-dessous d'un seuil prédéterminé. Une telle attitude ne doit cependant pas entraîner une fixation irrévocable des moyens mis en oeuvre, puisque l'échantillonnage est avant tout un processus évolutif au cours duquel le modèle peut être enrichi après chaque étape d'échantillonnage.

La définition d'un programme de prélèvement rigoureux et efficace destiné au contrôle de la qualité bactériologique de l'eau d'un réseau doit obligatoirement être précédée par une recherche méthodique et systématique de la structure d'abondance des populations bactériennes. Cette recherche devra notamment spécifier si le réseau étudié peut être considéré comme un unique système hétérogène, ou bien comme une structure plus complexe composée de plusieurs sous-systèmes hétérogènes. Ce dernier cas semble à l'évidence beaucoup plus probable pour les réseaux au-delà d'une certaine importance. La stratification d'un réseau permet, par ailleurs, de circonscrire les secteurs posant des problèmes, ce qui facilite également la mise en oeuvre de moyens appropriés pour remédier le cas échéant à des résultats chroniquement douteux. Sur ce point, notons par exemple que, pour ce qui concerne le réseau de distribution de la ville de Metz, la densité bactérienne moyenne maximale de 100 bactéries hétérotrophes par millilitre a été dépassée lors de toutes les campagnes de prélèvements dans les zones 2, 3 et 4 et même dans la zone 1 au cours de la dernière campagne d'échantillonnage.

L'échantillonnage progressif offre une alternative intéressante d'échantillonnage, en fonction d'une modélisation, permettant d'effectuer le contrôle de la qualité bactériologique d'une eau en réseau. En plus de la valeur du paramètre  $k$ , le procédé prend en considération les résultats enregistrés au fur et à mesure de leur obtention après chaque étape de l'échantillonnage. Cette approche permet de se soustraire aux exigences parfois assez contraignantes des programmes d'échantillonnage utilisant un nombre fixé à l'avance et souvent important d'échantillons à recueillir au

cours d'une seule campagne de prélèvements. De plus, l'échantillonnage progressif peut, dans certaines circonstances, s'avérer nettement plus efficace que la méthode plus classique décrite dans le cadre d'un bilan instantané. Ceci se produit en particulier lorsque la densité bactérienne moyenne réelle de l'eau est suffisamment éloignée au-dessous ou au-dessus de la valeur de référence spécifiée par la norme guide appliquée. Les différences sont alors rapidement mises en évidence et l'effort d'échantillonnage concédé est réduit à un niveau minimal. Il peut toutefois arriver que le point de coordonnées  $(n, T_n)$  se maintienne durablement entre les limites  $A(n)$  et  $B(n)$  définies par les formules (56) et (57). Cette éventualité, qui fait obstacle à l'interruption du processus en retardant son issue dichotomique malgré un nombre parfois élevé d'observations, doit évoquer une densité bactérienne moyenne de l'eau analysée comprise entre les valeurs  $\lambda_0$  et  $\lambda_1$  associées aux limites précédentes. En outre, il faut souligner que les deux types d'erreur, à savoir : le risque fournisseur ( $\alpha$ ) et le risque consommateur ( $\beta$ ) sont pris en considération dans l'échantillonnage progressif. Un avantage supplémentaire de ce mode de prélèvement tient sans doute dans l'interprétation graphique illustrée sur la figure 19, et la possibilité de classer l'eau analysée dans un éventail de catégories correspondant à différents niveaux de qualité bactériologique.

L'ordinateur permet en fait d'étendre le champ d'utilisation de l'échantillonnage progressif à tout système homogène ou hétérogène, même lorsque la valeur de l'entropie du système (i.e., le paramètre  $k$ ) n'est pas supposée connue a priori. Cette perspective a le mérite d'assouplir quelque peu les hypothèses relativement fortes, nécessaires à la construction de ce type de plans. Il est en effet possible de calculer de manière itérative une nouvelle estimation de  $k$  après chaque étape de l'échantillonnage. Mais dans ce cas, la perte d'un degré de liberté, consécutive à l'identification du modèle, entraîne l'obligation de comparer  $T_n$  à  $A(n-1)$  et  $B(n-1)$  au lieu de  $A(n)$  et  $B(n)$ . De plus, au cours de chaque nouvelle itération, les formules permettant le calcul des bornes de l'intervalle d'incertitude doivent être choisies en fonction du modèle retenu. Ainsi, les relations (56) et (57) seront utilisées s'il y a "surdispersion" des résultats tandis que les formules (58) et (59) seront appliquées à des séries expérimentales compatibles avec le modèle de Poisson : c'est-à-dire si, par exemple,  $T_n = 0$  ou, plus généralement, lorsque  $k$  est grand par rapport à la moyenne  $T_n/n$ . L'algorithme de cette procédure, qui se prête éventuellement à la visualisation ainsi qu'au suivi des résultats sur écran, est illustré à la figure 20.

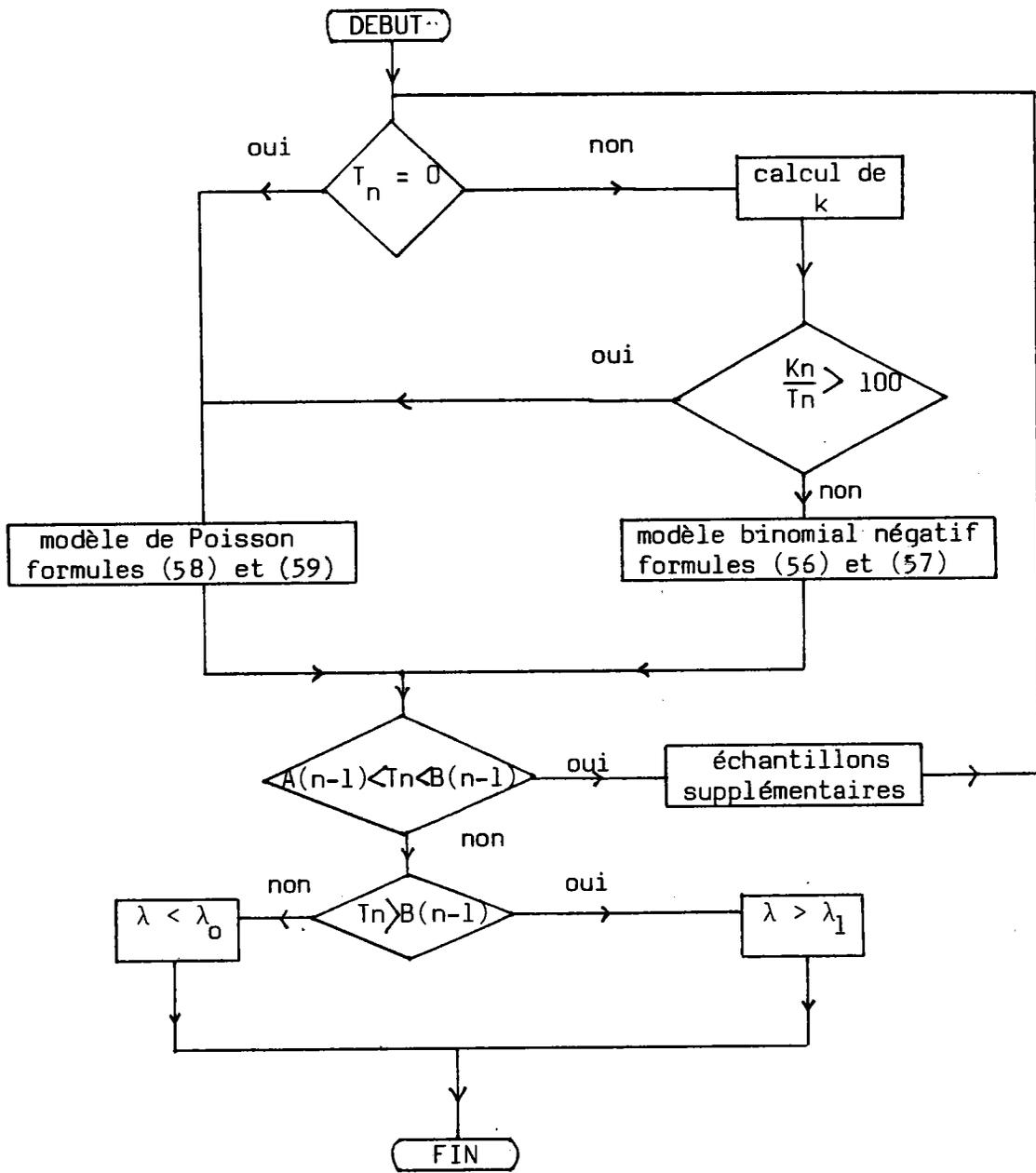


FIGURE 20 : Organigramme de la procédure décisionnelle lorsque le modèle de distribution est défini progressivement à partir des données disponibles.

Les méthodes d'échantillonnage présentées dans ce travail ont été explicitées dans le cadre d'une réglementation particulière intéressant les bactéries hétérotrophes à 20°C. Les concepts utilisés se généralisent cependant aisément à d'autres spécifications, voire même à d'autres critères de qualité faisant état de données pouvant être décrites par la distribution binomiale négative.

## **CONCLUSION**

L'étude des modèles de distribution de même que la définition d'une stratégie d'échantillonnage des populations bactériennes dans un réseau de distribution publique nous ont amenés à mettre en pratique un certain nombre de méthodes statistiques afin de tenter :

- d'une part, de promouvoir une technique de titrage adaptée à la quantification des bactéries hétérotrophes,

- d'autre part, d'appréhender la configuration spatio-temporelle de l'hétérogénéité de la concentration bactérienne dans un système de distribution. La description de la structure d'abondance des micro-organismes par un modèle mathématique se trouve justifiée par ses finalités à la fois explicatives et décisionnelles. La structure observée permet en effet de mieux comprendre les aspects déterministes liés à certaines variables physico-chimiques prépondérantes mais elle doit également servir de fondement à la formulation rationnelle des modalités de tout programme de contrôle et de surveillance de la qualité de l'eau distribuée.

\* La comparaison expérimentale de trois méthodes de titrage (méthodes énumératives par filtration sur membrane ou par inclusion en gélose et méthode quantique en milieu liquide avec 96 ensemencements par dilution) a montré :

- l'existence d'une forte interaction méthode-échantillon. Ce phénomène, qui s'est surtout manifesté entre la méthode quantique et les deux méthodes énumératives doit être imputé au manque de spécificité absolue des méthodes de titrage, ainsi qu'aux difficultés de lecture consécutives à l'absence de système indicateur fiable pour la microméthode,

- une valeur du coefficient de variation empirique (33 %) nettement supérieure à la reproductibilité théorique (13 %) de la méthode quantique en milieu liquide,

- une assez bonne concordance des résultats obtenus par les deux techniques énumératives. Cependant, nous avons préféré retenir la méthode des membranes filtrantes au détriment de la méthode par inclusion en gélose pour des raisons de commodité d'utilisation et de sensibilité.

\* L'étude comparative des rendements relatifs de plusieurs milieux de culture a surtout fait ressortir l'aspect cinétique de la reviviscence bactérienne sur milieu gélosé. Des différences significatives de rendement ont été observées après 48 heures d'incubation, mais celles-ci s'atténuent progressivement et les trois milieux testés (i. e., gélose nutritive, milieux R-2A et m-SPC) donnent des résultats assez comparables au terme d'une incubation de 96 heures. Cette évolution, qui s'effectue de manière similaire pour l'ensemble des trois milieux étudiés à partir de 72 heures d'incubation, se caractérise toutefois par un net ralentissement de l'augmentation relative du nombre des colonies recensées entre 72 et 96 heures.

Les comparaisons précédentes ont toutes été effectuées par le test du rapport de vraisemblance. Cet outil statistique se distingue des autres techniques habituellement utilisées pour les comparaisons de moyennes en microbiologie, en exploitant toute l'information disponible sur la base des seules distributions théoriques des données non transformées.

\* Les investigations préliminaires menées afin d'étudier l'hétérogénéité du milieu hydrique à différentes échelles d'observation ont permis, moyennant une généralisation de l'indice de dispersion de Fisher, de

mettre en évidence une structure hiérarchisée des niveaux d'hétérogénéité. Cette structuration du peuplement bactérien est caractérisée par une variabilité dont l'ampleur augmente rapidement avec l'élargissement de la fenêtre d'observation du continuum spatio-temporel ; et le modèle de Poisson doit être irrémédiablement écarté à l'échelle d'un réseau de distribution publique. De plus, les ultimes canalisations d'un immeuble exercent sur les concentrations bactériennes un effet dispersif qui ne s'accompagne cependant d'aucune augmentation de la densité bactérienne moyenne intra-muros.

\* La recherche intentionnelle et systématique de la configuration spatio-temporelle de l'hétérogénéité bactérienne du réseau messin a permis d'aboutir aux conclusions suivantes :

- les données expérimentales recueillies au cours d'une même campagne d'échantillonnage ne sont généralement pas distribuées suivant une loi binomiale négative unique. Par conséquent, l'hypothèse d'un réseau formé d'un seul système hétérogène a dû être rejetée,

- la méthode de groupement dont les résultats ont été corroborés par l'ordination en espace réduit des stations de prélèvement a conduit à la réalisation d'une partition du réseau en plusieurs zones pouvant être individuellement modélisées par la distribution binomiale négative. Le réseau messin apparaît de ce fait comme un ensemble composite formé de plusieurs sous-systèmes hétérogènes associés à des niveaux d'abondance bactérienne propres,

- la configuration structurée et persistante de l'hétérogénéité spatiale est compatible avec les caractéristiques hydrauliques principales qui découlent de l'infrastructure du réseau ; en particulier, les zones de forte densité bactérienne sont plutôt localisées à la périphérie du système de distribution, loin de l'usine de traitement.

\* En ce qui concerne l'étude des liaisons existant entre l'abondance bactérienne et certaines variables physico-chimiques, nous avons pu montrer que la concentration bactérienne était en corrélation étroite avec

- le temps de séjour de l'eau dans les canalisations, ce qui contribue à objectiver la remarque précédente,

- les concentrations de chlore résiduel libre et total,

- la température de l'eau, conformément à un modèle log-linéaire qui, dans les limites du domaine exploré, implique un accroissement global de la population bactérienne de 38 % environ chaque fois que la température de l'eau s'élève de 1°C. Cette liaison, qui traduit indirectement une grande partie de la variation temporelle, laisse entrevoir des variations cycliques obéissant à des rythmes saisonniers dont l'ampleur considérable ne semble cependant pas affecter les zones de manière identique. Celle-ci paraissent évoluer relativement indépendamment les unes des autres.

\* Enfin, nous avons montré comment définir les modalités d'un programme d'échantillonnage visant à contrôler la conformité, au regard d'une spécification donnée, de la qualité bactériologique de l'eau distribuée dans un réseau modélisé. Une étude préliminaire devra donc, le cas échéant, être réalisée dans le but de stratifier le réseau considéré en le définissant comme un ensemble formé d'un ou plusieurs sous-systèmes hétérogènes. Indépendamment de l'issue de cette modélisation, la stratégie d'échantillonnage peut se concevoir de deux manières différentes selon que l'on souhaite établir un bilan instantané ou bien un bilan échelonné. Dans le premier cas, la taille de

l'échantillon à recueillir dans chaque zone est calculée en fonction du risque d'erreur accepté ou encore de manière à répartir équitablement ce risque entre les différentes zones. A l'opposé, le nombre des mesures à effectuer n'est pas déterminé à l'avance lorsque l'échantillonnage est progressif. Quel que soit le mode de prélèvement adopté, l'effort d'échantillonnage à consentir est largement tributaire de la valeur du paramètre  $k$  exprimant l'hétérogénéité de chacun des sous-systèmes considérés. A cet égard, nous avons montré que 70 stations au moins doivent être échantillonnées dans le réseau messin afin de pouvoir déceler, lors d'un bilan instantané, des violations de la réglementation européenne avec une probabilité de 0,95 lorsque la densité bactérienne moyenne réelle est égale à 200 bactéries par millilitre. De conception légèrement plus sophistiquée mais d'application également moins contraignante, l'échantillonnage progressif peut, dans certaines circonstances, se montrer nettement plus performant que le mode de prélèvement attaché au bilan instantané.

Au terme de nos investigations théoriques menées conjointement avec une étude sur le terrain, nous pouvons affirmer que les stratégies de prélèvement préconisées permettent d'effectuer d'une manière objective et rationnelle le contrôle de la qualité bactériologique d'une eau en fonction d'une modélisation préalable du système de distribution.

**REFERENCES BIBLIOGRAPHIQUES**

- ALLEN M.J., TAYLOR R.H. and GELDREICH E.E. (1980)  
The occurrence of microorganisms in water main encrustations.  
J. Am. Wat. Wks Ass. 72, 614-625.
- ANDERSON J.E., EL-SHAARAWI A.H., ESTERBY S.R. and UNNY T.E. (1984)  
Dissolved oxygen concentrations in Lake Erie (U.S.A. - CANADA),  
1. Study of spatial and temporal variability using cluster and regression  
analysis.  
J. Hydrol. 72, 209-229.
- ARMSTRONG J.L., CALOMIRIS J.J. and SEIDLER R.J. (1982)  
Selection of antibiotic - resistant standard plate count bacteria during water  
treatment.  
Appl. Environ. Microbiol. 44, 308-316.
- ASHBY R.E. and RHODES-ROBERTS M.E. (1976)  
The use of analysis of variance to examine the variations between samples of  
marine bacterial populations.  
J. Appl. Bacteriol. 41, 439-451.
- BECKER R.J. (1978)  
Implementation of the coliform standard by a community water utility, p.  
169-180.  
In Evaluation of the microbiology standards for drinking water (Hendricks C.W.  
ed.).  
U.S. Environmental Protection Agency, Washington, D.C.
- BELL C.R. and ALBRIGHT L.J. (1981)  
Attached and free-floating bacteria in the Fraser River estuary, British  
Columbia, Canada.  
Mar. Ecol. Prog. Ser. 6, 317-327.
- BELLAIR J.T., PARR-SMITH G.A. and WALLIS I.G. (1977)  
Significance of diurnal variations in fecal coliform die-off rates in the  
design of ocean outfalls.  
J. WPCF 49, 2022-2030.
- BERLAND B.R., BONIN D.J., DURBEC J.P. and MAESTRINI S.Y. (1976)  
Bactéries hétérotrophes aérobies prélevées devant le delta du Rhône.  
II. Détermination des exigences nutritionnelles. Comparaison avec des souches  
provenant d'autres biotopes.  
Hydrobiologia 49, 123-128.
- BEYTOUT D., LAVERAN H. et REYNAUX M.P. (1975)  
Méthode pratique d'évaluation numérique applicable aux techniques  
miniaturisées de titrage en plages.  
Ann. Biol. Clin. 33, 379-384.
- BLOCK J.C., MAUL A et DOLLARD M.A. (1982)  
Estimation du nombre le plus probable (N.P.P.) des coliformes en rivière en  
réalisant 96 inoculums par niveau de dilution.  
Sciences de l'Eau 1, 387-399.
- BREITMAYER J.P. et GAUTHIER M.J. (1979)  
Dénombrement des bactéries en milieu marin : facteurs de variation et  
d'incertitude.  
Ann. Microbiol. (Inst. Pasteur) 130 A, 245-256.

BOUSQUET G. et DEVOUCOUX J. (1982)

Prévision et interprétation des résultats en bactériologie des eaux : facteurs de variabilité et stratégies.

J. Fr. Hydrol. 13, 157-169.

CADA R.L. (1975)

Proficiency test specimens for water bacteriology.

Appl. Microbiol. 29, 255-259.

CALOMIRIS J.J., ARMSTRONG J.L. and SEIDLER R.J. (1984)

Association of metal tolerance with multiple antibiotic resistance of bacteria isolated from drinking water.

Appl. Environ. Microbiol. 47, 1238-1242.

CHRISTIAN R.R. and PIPES W.O. (1983)

Frequency distribution of coliforms in water distribution systems.

Appl. Environ. Microbiol. 45, 603-609.

CHURCHLAND L.M., KAN G. and AGES A. (1982)

Variation in fecal pollution indicators through tidal cycles in the Fraser River estuary.

Can. J. Microbiol. 28, 239-247.

COCHRAN W.G. (1950)

Estimation of bacterial densities by means of the "most probable number".

Biometrics 6, 105-116.

COLWELL R.R., AUSTIN B. and WAN L. (1978)

Public health considerations of the microbiology of "potable" water. p. 65-75.

In Evaluation of the microbiology standards for drinking water (Hendricks C.W. ed.). U.S. Environmental Protection Agency, Washington, D.C.

COMMISSARIAT A L'ENERGIE ATOMIQUE (1978)

Statistique appliquée à l'exploitation des mesures.

Tomes 1 et 2, Masson édit., Paris.

DOTT W. (1983)

Qualitative und quantitative Bestimmung von Bakterienpopulationen aus aquatischen Biotopen. 6. Mitteilung : Wiederverkeimung im Trinkwasser.

Zbl. Bakt. Hyg., I. Abt. Orig. B 178, 263-279.

DOTT W. und TRAMPISCH H.J. (1983)

Qualitative und quantitative Bestimmung von Bakterienpopulationen aus aquatischen Biotopen. 5. Mitteilung : Vergleichende Untersuchungen an zwei Schnellsandfiltern.

Zbl. Bakt. Hyg., I. Abt. Orig. B 177, 141-155.

DUBY C. et RODOLPHE F. (1982)

Echantillonnage en fonction d'une modélisation. p. 437-455.

Dans Stratégies d'échantillonnage en écologie. Masson édit., Paris.

DUCKLOW H.W., KIRCHMAN D.L. and ROWE G.T. (1982)

Production and vertical flux of attached bacteria in the Hudson River plume of the New York Bight as studied with floating sediment traps.

Appl. Environ. Microbiol. 43, 769-776.

EL-SHAARAWI A.H. (1983)

Goodness of fit statistics for generalized power series distributions.

Communication : 44th Session of the International Statistical Institute, Madrid.

EL-SHAARAWI A.H. (1982)

Sampling strategy for estimating bacterial density in large lakes.  
J. Fr. Hydrol. 13, 171-187.

EL-SHAARAWI A.H., BLOCK J.C. and MAUL A. (1985)

The use of historical data for estimating the number of samples required for monitoring drinking water.  
Sci. Tot. Environ. 42, 289-295.

EL-SHAARAWI A.H., ESTERBY S.R. and DUTKA B.J. (1981)

Bacterial density in water determined by poisson or negative binomial distributions.  
Appl. Environ. Microbiol. 41, 107-116.

EL-SHAARAWI A.H. and PIPES W.O. (1982)

Enumeration and statistical inferences. p. 43-66. In Bacterial indicators of pollution. CRC Press, Boca Raton, Fl.

ENVIRONMENTAL PROTECTION AGENCY (1976)

National interim primary drinking water regulations, EPA - 570/9-76-003.  
U.S. Environmental Protection Agency, Washington, D.C.

ESTERBY S.R. (1982)

Fitting probability distributions to bacteriological data. Considerations for surveys and for water quality guidelines.  
J. Fr. Hydrol. 13, 189-203.

FISHER R.A. (1941)

The negative binomial distribution  
Ann. Eugen. 11, 182-187.

FISHER R.A., THORNTON H.G. and Mc KENZIE W.A. (1922)

The accuracy of the plating method of estimating the density of bacterial populations.  
Ann. Appl. Biol. 9, 325-359.

FONTVIEILLE D. et MAUL A. (1985)

Dénombrement et mesures d'activité des bactéries. Dans Bactériologie des eaux douces, aspects écologiques et sanitaires, coordinateur Martin G., collection Le Point, vol. 2, Lavoisier édit., Paris, à paraître.

FRONTIER S. (1982)

Stratégies d'échantillonnage en écologie.  
Masson édit., Paris.

GELDREICH E.E. (1973)

Is the total count necessary ? Section VII. 1 - VII. 2.  
Proceedings AWWA Water Quality Technology Conference, Cincinnati, Ohio.

GREEN R.H. (1979)

Sampling design and statistical methods for environmental biologists.  
Wiley and Sons Publishers, New York.

GREEN B.L. and LITSKY W. (1980)

Evaluation of membrane filter methods for enumeration of faecal coliforms from marine water.  
Marine Environ. Res. 3, 267-276.

GREENWOOD M. and YULE G.U. (1917)

On the statistical interpretation of some bacteriological methods employed in water analysis.

J. Hyg. 16, 36-54.

GUERRERO R., RODA F., ABELLA C. and TORRELLA F. (1975)

Optimal growth temperatures and media parameters of bacterial communities from lakes of different trophic states.

Verh. Internat. Verein. Limnol. 19, 2620-2626.

HATTORI T. (1985)

Kinetics of colony formation of bacteria : an approach to the basis of the plate count method.

Rep. Inst. Agr. Res. 34, 1-36.

HEALTH and WELFARE CANADA (1978)

Guidelines for canadian drinking water quality.

Canadian Government Publishing Centre, Supply and Services Canada, Hull, Quebec.

HERRICKS E.E., SCHAEFFER D.J. and KAPSNER J.C. (1985)

Complying with NPDES permit limits : when is a violation a violation ?

J. WPCF 57, 108-115.

KAPER J.B., MILLS A.L. and COLWELL R.R. (1978)

Evaluation of the accuracy and precision of enumerating aerobic heterotrophs in water samples by the spread plate method.

Appl. Environ. Microbiol. 35, 756-761.

KENDALL M.G. (1948)

Rank correlation methods.

Charles Griffin and Co., London.

KENDALL M.G. and STUART A. (1968)

Design and analysis, and time-series. p. 285-291. In The advanced theory of statistics, vol. 3, 2<sup>nd</sup> ed.

Hafner Publ. Co., New York.

KLEIN D.A. and WU S. (1974)

Stress : a factor to be considered in heterotrophic microorganism enumeration from aquatic environments.

Appl. Microbiol. 27, 429-435.

KOGURE K., SIMIDU U. and TAGA N. (1982)

Bacterial attachment to phytoplankton in sea water.

J. Exp. Mar. Biol. Ecol. 56, 197-204.

KOSINSKI R.J., SINGLETON F.L. and FOSTER B.G. (1979)

Sampling culturable heterotrophs from microcosms : a statistical analysis.

Appl. Environ. Microbiol. 38, 906-910.

LEFEBVRE J. (1976)

Introduction aux analyses statistiques multidimensionnelles.

Masson édit., Paris.

LEGENDRE L. et LEGENDRE P. (1979)

Ecologie numérique.

Tomes 1 et 2, Masson édit., Paris.

LINDGREN B.W. (1968)

Statistical theory, 2<sup>nd</sup> ed.  
Collier - Macmillan, London.

MAUL A. (1982)

Définition d'une stratégie de prélèvement pour l'analyse microbiologique des eaux de surface.

Thèse de Doctorat de 3<sup>e</sup> cycle, Université de Metz.

MAUL A. et BLOCK J.C. (1983)

Microplate fecal coliform method to monitor stream water pollution.  
Appl. Environ. Microbiol. 46, 1032-1037.

MAUL A., DOLLARD M.A. et BLOCK J.C. (1981)

Application du principe du maximum de vraisemblance (N.P.P.) à la numération bactérienne sur milieu gélosé.

J. Fr. Hydrol. 12, 245-254.

MCCARTHY J.A., DELANEY J.E. and GRASSO R.J. (1961)

Measuring coliforms in water.  
Water Sew. Works 108, 238-245.

McFETERS G.A., SHILLINGER J.E. and STUART D.G. (1978)

Alternative indicators of water contamination and some physiological characteristics of heterotrophic bacteria in water. p. 37-48.

In Evaluation of the microbiology standards for drinking water

(Hendricks C.W. ed.). U.S. Environmental Protection Agency, Washington, D.C.

MEANS E.G., HANAMI L., RIDGWAY H.F. and OLSON B.H. (1981)

Evaluating mediums and plating techniques for enumerating bacteria in water distribution systems.

J. Am. Wat. Works Ass. 73, 585-590.

METIVIER M. (1972)

Notions fondamentales de la théorie des probabilités, 2<sup>e</sup> éd.

Dunod édit., Paris.

MUENZ L.R. (1978)

Some statistical considerations in water quality control. p. 49-56.

In Evaluation of the microbiology standards for drinking water

(Hendricks C.W. ed.). U.S. Environmental Protection Agency, Washington, D.C.

NAGY L.A., SHOVLIN M.G. and OLSON B.H. (1983)

Factors influencing the occurrence of bacteria, filamentous fungi and yeasts on the surface of drinking water pipes.

Communication au 3<sup>e</sup> Symposium International "Microbial Ecology"

Michigan State University, Abstract E 4.

PALMER F.E., METHOT R.D. and STALEY J.T. (1976)

Patchiness in the distribution of planktonic heterotrophic bacteria in lakes.

Appl. Environ. Microbiol. 31, 1003-1005.

PIPES W.O. and CHRISTIAN R.R. (1982)

Sampling frequency. Microbiological drinking water regulations - final report.

U.S. Environmental Protection Agency, EPA R-805-637/9-82-001, Washington, D.C.

PIPES W.O., WARD P. and AHN S.H. (1977)

Frequency distributions for coliform bacteria in water.

J. Am. Wat. Works Ass. 69, 644-647.

POOLE R.W. (1974)

An introduction to quantitative ecology.

McGraw - Hill Publishers, New York.

PROST H. et HUGUES B. (1982)

Dénombrement des microorganismes par la technique du nombre le plus probable (indice N.P.P.) : emploi statistique de cet indice.

Ann. Falsif. Expert. Chim. 75, 185-207.

RAO C.R. (1973)

Linear statistical inference and its applications, 2 nd ed.

Wiley and Sons Publishers, New York.

REASONER D.J. and GELDREICH E.E. (1985)

A new medium for the enumeration and subculture of bacteria from potable water.

Appl. Environ. Microbiol. 49, 1-7.

RICHARD J., ANDERSEN H.M. and GRATADOUX J.J. (1983)

Rapid method for selecting appropriate solid media for the enumeration of aerobic micro-organisms.

J. Appl. Bacteriol. 54, 329-334.

RIZET M., BLANC J., COGNET M., DERONS M., DORIZON N. et FILIPPI G. (1984)

Qualité de l'eau dans les réseaux de distribution, son évolution biologique et chimique.

J. Fr. Hydrol. 15, 207-221.

RIDGWAY H.F. and OLSON B.H. (1981)

Scanning electron microscope evidence for bacterial colonization of a drinking water distribution system.

Appl. Environ. Microbiol. 41, 274-287.

RIDGWAY H.F. and OLSON B.H. (1982)

Chlorine resistance patterns of bacteria from two drinking water distribution systems.

Appl. Environ. Microbiol. 44, 972-987.

RODIER J. (1984)

L'analyse de l'eau, 7e éd. p. 534-536.

Dunod édit., Paris.

SEKI H. (1968)

Some characteristics of marine microorganisms in the pacific central water.

Rec. Oceanogr. Works Jpn. 9, 247-255.

SILVERMAN G.S., NAGY L.A. and OLSON B.H. (1983)

Variations in particulate matter, algae, and bacteria in an uncovered, finished - drinking - water reservoir.

J. Am. Wat. Works Ass. 75, 191-195.

SNEDECOR G.W. and COCHRAN W.G. (1957)

Statistical methods, 6 th ed.

Iowa State University Press, Ames.

SOROKIN Y.I. (1970)

On the degree of aggregation of marine bacterio-plankton.  
Dokl. Akad. Nauk. SSSR 192, 905-907.

TAYLOR R. and GELDREICH E.E. (1979)

Standard plate count methodology. A new membrane filter procedure for bacterial counts in potable water and swimming pool samples.  
J. Am. Wat. Works Ass. 71, 402-405.

THOMAS H.A., Jr. (1955)

Statistical analysis of coliform data.  
Sew. Ind. Wastes 27, 212-228.

THOMAS H.A., Jr. and WOODWARD R.L. (1955)

Estimation of coliform density by the membrane filter and the fermentation tube methods.  
J. Publ. Health 45, 1431-1437.

TIPPETT L.H.C. (1932)

A modified method of counting particles.  
Proc. Roy. Soc. (London) 137, 434-446.

TOBIN R.S., LOMAX P. and KUSHNER D.J. (1980)

Comparison of nine brands of membrane filter and the most-probable-number methods for total coliform enumeration in sewage-contaminated drinking water.  
Appl. Environ. Microbiol. 40, 186-191.

VAN DER KOOIJ D. and HIJNEN W.A.M. (1984)

Substrate utilization by an oxalate-consuming *Spirillum* species in relation to its growth in ozonated water.  
Appl. Environ. Microbiol. 47, 551-559.

VAN DER KOOIJ D., ORANJE J.P. and HIJNEN W.A.M. (1982)

Growth of *Pseudomonas aeruginosa* in tap water in relation to utilization of substrates at concentrations of a few micrograms per liter.  
Appl. Environ. Microbiol. 44, 1086-1095.

VAN DER WAERDEN (1967)

Statistique mathématique.  
Dunod édit., Paris.

VAN SOESTBERGER A.A. and LEE C.H. (1969)

Pour plates or streak plates ?  
Appl. Microbiol. 18, 1092-1098.

VELZ C.J. (1970)

Applied stream sanitation.  
Wiley Interscience Publishers, Inc., New York.

WALD A. (1947)

Sequential analysis.  
Wiley and Sons Publishers, New York.

WILCOX D.P., CHANG E., DICKSON K.L. and JOHANSSON K.R. (1983)  
Microbial growth associated with granular activated carbon in a pilot water  
treatment facility.  
Appl. Environ. Microbiol. 46, 406-416.

WILLIAMS P.J. (1970)  
Heterotrophic utilization of dissolved organic compounds in the sea.  
I. Size distribution of population and relationship between respiration and  
incorporation of growth substrates.  
J. Marine Biol. Assoc. 50, 859-870.

WYSHAK G. and DETRE K. (1972)  
Estimating the number of organisms in quantal assays.  
Appl. Microbiol. 23, 784-790.

## ANNEXES

**ANNEXE I**

**CALCUL DES FREQUENCES THEORIQUES  
CUMULEES DE LA DISTRIBUTION BN (k,p)**

Le test d'adéquation de  $\chi^2$  pour une loi binomiale négative nécessite le calcul de probabilités telles que  $P(N_0 < X \leq N)$  où  $N_0$  et  $N$  désignent des nombres entiers naturels alors que  $X$  représente une variable aléatoire distribuée suivant la loi BN  $(k,p)$ . Ces probabilités sont obtenues à partir des fréquences cumulées de la loi considérée, en l'occurrence

$$P(X \leq N) = \sum_{i=0}^N P(i).$$

Partant de  $P(0) = \frac{1}{(1+p)^k}$ , le calcul de  $P(r) = \frac{(k+r-1)!}{r! (k-1)!} \frac{p^r}{(1+p)^{k+r}}$  est effectué de manière itérative au moyen de la formule de récurrence liant les probabilités de deux valeurs consécutives de  $r$ , à savoir :

$$P(r+1) = P(r) \frac{(k+r) p}{(r+1) (1+p)}$$

PROGRAMME DE CALCUL DE LA PROBABILITE P (X ≤ N)

EFFECTUE POUR LA CALCULATRICE HEWLETT-PACKARD 41 C

01 LBL "CUM"	21 STO 04	41 *
02 "K"	22 RCL 02	42 RCL 03
03 PROMPT	23 1	43 *
04 STO 00	24 -	44 STO 03
05 "P"	25 RCL 04	45 RCL 10
06 PROMPT	26 X > Y ?	46 +
07 STO 01	27 GTO 02	47 STO 10
08 "N"	28 RCL 04	48 RCL 04
09 PROMPT	29 1	49 1
10 STO 02	30 +	50 +
11 RCL 01	31 RCL 01	51 GTO 01
12 1	32 1	52 LBL 02
13 +	33 +	53 RCL 10
14 RCL 00	34 *	
15 Y ↗ X	35 1 / X	
16 1 / X	36 RCL 01	
17 STO 03	37 *	
18 STO 10	38 RCL 00	
19 0	39 RCL 04	
20 LBL 01	40 +	

### MODE D'UTILISATION

Le chargement du programme ayant été effectué directement à partir du listage ou bien par carte magnétique, la probabilité  $P(X \leq N)$  associée par exemple à la loi BN (0,21584 ; 1943,66) est calculée de la manière suivante pour  $N=19$ .

<u>Séquence de touches</u>			<u>Affichage</u>	
<input type="text" value="XEQ"/>	<input type="text" value="ALPHA"/>	CUM	<input type="text" value="ALPHA"/>	K
0,21584	<input type="text" value="R/S"/>			P
1943,66	<input type="text" value="R/S"/>			N
19	<input type="text" value="R/S"/>			

après quelques secondes d'attente, le résultat est affiché sur l'écran :  
0,405.

N.B. : Un programme de calcul de la probabilité  $P(N)$  est obtenu en supprimant les lignes 18, 45, 46, 47 et en remplaçant les lignes 01 et 53 respectivement par "LBL 'PRO'" et "RCL 03" dans le programme précédent.

## ANNEXE II

ESTIMATION DU PARAMETRE K PAR LA  
METHODE DU MAXIMUM DE VRAISEMLANCE

La valeur la plus vraisemblable du paramètre  $k$ , estimée à partir d'une série de  $n$  résultats expérimentaux  $r_1, r_2, \dots, r_n$  issus d'une population dont la loi parente est la loi BN ( $k,p$ ), est la valeur  $\hat{k}$  donnée par la solution de l'équation

$$n \ln \left( 1 + \frac{\bar{r}}{k} \right) - \sum_{i=1}^n \sum_{t=1}^{r_i} \frac{1}{k + r_i - t} = 0$$

Dans le cas où cette équation, de la forme  $d(k) = 0$ , admet une solution, sa résolution est possible par la méthode de Newton. Le principe de cette méthode est illustré par la figure I qui donne également le profil général de la courbe représentative de la fonction définie par  $d(k)$ .

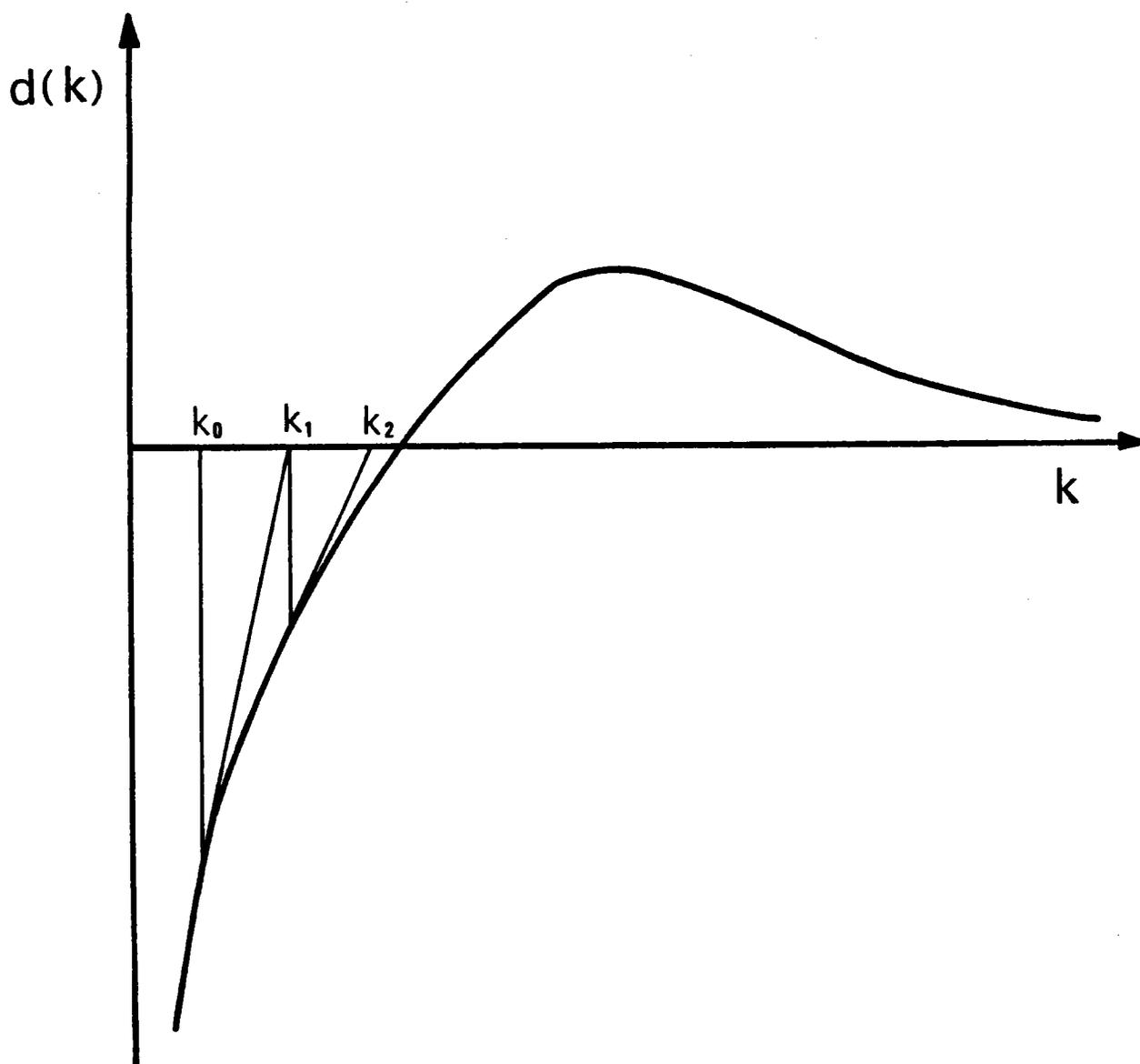


FIGURE I : Résolution de l'équation de vraisemblance par la méthode de Newton.

Le choix d'une valeur initiale  $k_0$  permet d'engager le processus itératif ; la valeur  $k_1 = k_0 - \frac{d(k_0)}{d'(k_0)}$  est alors l'abscisse du point d'intersection avec l'axe  $O k$  de la tangente à la courbe au point d'abscisse  $k_0$ . La réitération de ce procédé engendre une suite de nombres  $k_0, k_1, \dots, k_n, \dots$ , définie par la formule de récurrence  $k_{n+1} = k_n - \frac{d(k_n)}{d'(k_n)}$ . Lorsque la valeur initiale  $k_0$  est convenablement choisie, cette suite converge vers la solution de l'équation de vraisemblance.

PROGRAMME DE RESOLUTION DE L'EQUATION  
DE VRAISEMBLANCE, ASSOCIEE AU PARAMETRE K,  
ECRIT EN LANGAGE BASIC MICROSOFT

Le calcul des sommes  $S_{r_i} = \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{k+r_i-1}$  et  $S_{r_i}^2 = \frac{1}{k^2} + \frac{1}{(k+1)^2} + \dots + \frac{1}{(k+r_i-1)^2}$ , apparaissant, pour toute valeur expérimentale  $r_i$  différente de zéro, dans les expressions de  $d(k)$  et de  $d'(k)$ , peut occasionner une perte de temps non négligeable lorsque les valeurs  $r_i$  sont relativement grandes. Afin de pallier cet inconvénient il est souhaitable d'utiliser une approximation quand  $r_i$  dépasse un seuil  $R$  prédéterminé. Ainsi, pour toute valeur  $r_i$  supérieure à  $R$ ,  $S_{r_i}$  se calcule en ajoutant à la quantité fixe  $SR = \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{k+R-1}$ , la somme  $\frac{1}{k+R} + \frac{1}{k+R+1} + \dots + \frac{1}{k+r_i-1}$  dont une valeur approchée est donnée par  $\ln \frac{k+r_i-0,5}{k+R-0,5}$ . De manière analogue, la quantité  $S_{r_i}^2$  est obtenue en ajoutant à  $SR^2 = \frac{1}{k^2} + \frac{1}{(k+1)^2} + \dots + \frac{1}{(k+R-1)^2}$ , la somme  $\frac{1}{(k+R)^2} + \frac{1}{(k+R+1)^2} + \dots + \frac{1}{(k+r_i-1)^2}$  qu'il est loisible de remplacer par  $\frac{1}{k+R-0,5} - \frac{1}{k+r_i-0,5}$ . L'exécution du programme dont le listage est donné ci-après, nécessite, outre le choix de  $R$ , l'introduction du nombre total,  $n$ , des données bactériologiques qui devront être consignées dans une instruction DATA.

```
10 INPUT "NOMBRE DES DONNEES BACTERIOLOGIQUES" ; N
20 DIM R(N)
30 INPUT "VALEUR DE R" ; R
40 FOR J=1 TO N
50 READ R(J)
60 T=T+R(J)
70 NEXT J
80 K= 0.1
90 SR= 0
100 SR2= 0
110 REM CALCUL DE SR ET DE SR2
120 FOR I=1 TO R
130 SR= SR+1 / (I+K-1)
140 SR2= SR2+1 / ((I+K-1) * (I+K-1))
150 NEXT I
160 REM CALCUL DE LA SOMME DES SRI ET DES SRI2
170 SRI= 0
180 SRI2= 0
190 FOR I=1 TO N
200 IF R(I)= 0 THEN 290
210 IF R(I) ≤ R THEN 220 ELSE 270
220 FOR M= 1 TO R(I)
230 SRI= SRI+1 / (M+K-1)
240 SRI2= SRI2+1 / ((M+K-1) * (M+K-1))
250 NEXT M
260 GOTO 290
270 SRI= SRI+SR+LOG ((K+R(I)-0.5) / (K+R-0.5))
280 SRI2= SRI2+SR2+1 / (K+R-0.5)-1 / (K+R(I)-0.5)
290 NEXT I
300 D= N * LOG (1+T / (N * K))-SRI
310 IF D < - 100 THEN 320 ELSE 340
320 K= K+0.1
330 GOTO 90
340 DD= -T/ (K * (K+T/N))+SRI2
350 K1= K-D/DD
360 IF ABS (K1-K) < 10-5 THEN PRINT K1 : GOTO 400
370 IF T / (N * K1) < 10-1 THEN PRINT "K INFINI" : GOTO 400
380 K= K1
390 GOTO 90
400 END
410 DATA
```

**ANNEXE III**

**ESTIMATION DU PARAMETRE K COMMUN A  
PLUSIEURS DISTRIBUTIONS BINOMIALES NEGATIVES**

L'estimation par la méthode du maximum de vraisemblance du paramètre  $k$  commun à plusieurs distributions binomiales négatives BN ( $k, p_i$ ),  $i=1, \dots, \ell$ , représentées par les séries expérimentales ( $r_{i1}, r_{i2}, \dots, r_{in_i}$ ) conduit à la résolution de l'équation

$$\sum_{i=1}^{\ell} n_i \ln \left( 1 + \frac{\bar{r}_i}{k} \right) - \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \sum_{t=1}^{r_{ij}} \frac{1}{k+r_{ij}-t} = 0$$

La solution  $\hat{k}$  de cette équation de la forme  $\sum_{i=1}^{\ell} d_i(k) = 0$  peut aussi être

calculée par la méthode itérative de Newton.

Partant d'une valeur initiale  $k_0$ , le procédé itératif décrit en Annexe II engendre une suite récurrente définie par la formule

$$k_{n+1} = k_n - \frac{\sum_{i=1}^{\ell} d_i(k_n)}{\sum_{i=1}^{\ell} d'_i(k_n)}$$

et dont la limite finie, si elle existe, est égale au nombre cherché.

#### PROGRAMME DE CALCUL DU PARAMETRE K COMMUN A PLUSIEURS

#### POPULATIONS ECRIT EN LANGAGE BASIC MICROSOFT

Le programme développé ci-dessous s'inspire du programme présenté en Annexe II, dont il reprend les techniques de calcul ainsi que la plupart des notations. Pour l'exécution de ce programme, il est nécessaire d'introduire successivement : la valeur de  $R$ , la valeur initiale de  $k$ , le nombre  $\ell$  de séries expérimentales, ainsi que l'effectif de chacune de ces séries dont les valeurs sont consignées dans un tableau comprenant  $\ell$  lignes.

```

10 INPUT "VALEUR DE R" ; R
20 INPUT "VALEUR INITIALE DE K" ; K
30 INPUT "NOMBRE DE SERIES EXPERIMENTALES" ; L
40 DIM D(L)
50 DIM DD(L)
60 DIM T(L)
70 DIM N(L)
80 REM INTRODUCTION DE L'EFFECTIF DES SERIES EXPERIMENTALES
90 FOR I=1 TO L
100 INPUT N(I)
110 NEXT I
120 REM RECHERCHE DU PLUS GRAND EFFECTIF
130 MAX= 0
140 FOR J=1 TO L
150 IF MAX <= N(J) THEN 170
160 MAX=N(J)
170 NEXT J
180 DIM R(L, MAX)
190 REM INTRODUCTION DES DONNEES
200 FOR X=1 TO L

```

```
210 FOR J=1 TO N(X)
220 INPUT R(X,J)
230 NEXT J
240 NEXT X
250 FOR X=1 TO L
260 FOR J=1 TO N(X)
270 T(X)= T(X)+R(X,J)
280 NEXT J
290 NEXT X
300 SR=0
310 SR2=0
320 REM CALCUL DE SR ET DE SR2
330 FOR I=1 TO R
340 SR= SR+1 / (I+K-1)
350 SR2= SR2+1 / ((I+K-1) * (I+K-1))
360 NEXT I
370 REM CALCUL DE LA SOMME DES SRI ET DES SRI2
380 FOR X=1 TO L
390 SRI=0
400 SRI2=0
410 FOR I=1 TO N(X)
420 IF R(X,I)=0 THEN 510
430 IF R(X,I) ≤ R THEN 440 ELSE 490
440 FOR M=1 TO R(X,I)
450 SRI = SRI + 1/(M+K-1)
460 SRI2 = SRI2 + 1/((M+K-1) * (M+K-1))
470 NEXT M
480 GOTO 510
490 SRI= SRI+SR+LOG ((K+R (X,I)-0.5) / (K+R-0.5))
500 SRI2= SRI2+SR2+1 / (K+R-0.5)-1 / (K+R(X,I)-0.5)
510 NEXT I
520 D(X)= N(X) * LOG(1+T(X) / (N(X) * K))-SRI
530 DD(X)= -T / (K * (K+T(X) / N(X)))+SRI2
540 NEXT X
550 DS=0
560 DDS=0
570 FOR J=1 TO L
580 DS= DS+D(J)
590 NEXT J
600 IF DS < -100 THEN 610 ELSE 630
610 K=K+0.1
620 GOTO 300
630 FOR J=1 TO L
640 DDS= DDS+DD(J)
650 NEXT J
660 REM CALCUL DE LA NOUVELLE VALEUR DE K
670 K1= K-DS/DDS
680 IF ABS (K1-K) < 10 ↑ (-5) THEN PRINT K1 : GOTO 720
690 IF K1 > 200 THEN PRINT "K INFINI" : GOTO 720
700 K=K1
710 GOTO 300
720 END
```

#### **ANNEXE IV**

ESTIMATION DE LA DENSITE BACTERIENNE PAR L'ANALYSE  
STATISTIQUE DE L'ENSEMBLE DES RESULTATS OBTENUS PAR  
DES METHODES DE TITRAGE DIFFERENTES

Une estimation de la densité bactérienne d'un échantillon par l'exploitation statistique conjuguée des résultats obtenus par deux méthodes de titrage énumératives, en l'occurrence la méthode des membranes filtrantes et la méthode par inclusion dans la gélose, ainsi qu'une méthode quantique, à savoir : la technique du nombre le plus probable, est donnée par la solution de l'équation suivante

$$\frac{1}{\lambda} \left[ \sum_{j=1}^{\ell_1} x_{1j} + \sum_{j=1}^{\ell_2} x_{2j} \right] + \sum_{j=1}^{\ell_3} p_j \frac{v_{3j} e^{-\lambda v_{3j}}}{1 - e^{-\lambda v_{3j}}} =$$

$$\sum_{j=1}^{\ell_1} v_{1j} + \sum_{j=1}^{\ell_2} v_{2j} + \sum_{j=1}^{\ell_3} v_{3j} (n_j - p_j)$$

Dans cette égalité  $x_{ij}$  désigne le nombre de colonies recensées au niveau de dilution  $j$  pour la méthode des membranes filtrantes ou par inclusion dans la gélose selon que l'indice  $i$  vaut respectivement 1 ou 2 ;  $v_{ij}$  représente le volume de l'échantillon originel filtré, incorporé dans la gélose, ou encoreensemencé dans chaque cupule (i. e.,  $i = 1, 2$  ou  $3$ , respectivement) au niveau de dilution  $j$ . Enfin  $p_j$  correspond au nombre de réponses positives sur un nombre total  $n_j$  de cupulesensemencées à la dilution de rang  $j$ .

Il est facile d'établir que la fonction définie par

$$f(\lambda) = \frac{A}{\lambda} + \sum_{j=1}^{\ell_3} p_j \frac{v_{3j} e^{-\lambda v_{3j}}}{1 - e^{-\lambda v_{3j}}} - B,$$

où  $A$  et  $B$  désignent respectivement le terme entre crochets et le second membre de l'équation de vraisemblance, est strictement décroissante sur  $]0; +\infty[$ . Cette propriété suggère l'utilisation de la méthode de Newton pour la résolution de l'équation  $f(\lambda) = 0$ . Par ailleurs, il est intéressant de

remarquer que si  $A = 0$  et  $B = \sum_{j=1}^{\ell_3} v_{3j} (n_j - p_j)$ , l'équation de vraisemblance

donnée ci-dessus se résume à l'équation, mentionnée par COCHRAN (1950), permettant le calcul du nombre le plus probable (N.P.P.) associé à la technique de même nom. Un programme de résolution de l'équation ainsi simplifiée, établi par MAUL (1982), a largement inspiré le programme dont le listage est développé ci-après.

PROGRAMME DE CALCUL DE LA DENSITE BACTERIENNE LA PLUS

VRAISEMBLABLE CONÇU POUR LA CALCULATRICE HEWLETT-PACKARD 41C

01 LBL "PMV"	46 /	91 *
02 "A"	47 RCL 01	92 1
03 PROMPT	48 0.00666	93 RCL 06
04 STO 11	49 *	94 -
05 "B"	50 RCL 06	95 X ↗ 2
06 PROMPT	51 *	96 /
07 STO 00	52 1	97 -
08 "P1"	53 RCL 06	98 RCL 13
09 PROMPT	54 -	99 RCL 07
10 STO 12	55 /	100 *
11 "P2"	56 +	101 0.00066
12 PROMPT	57 RCL 13	102 X ↗ 2
13 STO 01	58 0.00066	103 *
14 "P3"	59 *	104 1
15 PROMPT	60 RCL 07	105 RCL 07
16 STO 13	61 *	106 -
17 "INITIALE"	62 1	107 X ↗ 2
18 PROMPT	63 RCL 07	108 /
19 STO 04	64 -	109 -
20 LBL 01	65 /	110 RCL 11
21 CHS	66 +	111 RCL 04
22 0.06666	67 RCL 11	112 X ↗ 2
23 *	68 RCL 04	113 /
24 E ↗ X	69 /	114 -
25 STO 05	70 +	115 STO 09
26 RCL 04	71 RCL 00	116 RCL 08
27 0.00666	72 -	117 RCL 09
28 *	73 STO 08	118 /
29 CHS	74 RCL 12	119 CHS
30 E ↗ X	75 RCL 05	120 RCL 04
31 STO 06	76 *	121 +
32 RCL 04	77 0.06666	122 STO 10
33 0.00066	78 X ↗ 2	123 RCL 04
34 *	79 *	124 -
35 CHS	80 1	125 ABS
36 E ↗ X	81 RCL 05	126 0.01
37 STO 07	82 -	127 X > Y ?
38 RCL 12	83 X ↗ 2	128 GTO 02
39 0.06666	84 /	129 RCL 10
40 *	85 CHS	130 STO 04
41 RCL 05	86 RCL 01	131 GTO 01
42 *	87 RCL 06	132 LBL 02
43 1	88 *	133 RCL 10
44 RCL 05	89 0.00666	
45 -	90 X ↗ 2	

### MODE D'UTILISATION

L'exécution du programme nécessite l'introduction des nombres A, B, p<sub>1</sub>, p<sub>2</sub> et p<sub>3</sub> dans cet ordre. A titre d'exemple, le calcul de  $\hat{\lambda}$  correspondant à l'échantillon n° 2 du tableau se déroule comme suit :

<u>Séquence de touches</u>			<u>Affichage</u>
<input type="text" value="XEQ"/>	<input type="text" value="ALPHA"/>	PMV <input type="text" value="ALPHA"/>	A
789	<input type="text" value="R/S"/>		B
38	<input type="text" value="R/S"/>		P1
288	<input type="text" value="R/S"/>		P2
113	<input type="text" value="R/S"/>		P3
31	<input type="text" value="R/S"/>		

après quelques secondes d'attente, le résultat 26,99 est affiché sur l'écran.

**ANNEXE V**

RESULTATS DES CAMPAGNES DE QUANTIFICATION DES  
POPULATIONS BACTERIENNES EFFECTUEES SUR LE RESEAU  
DE DISTRIBUTION DE LA VILLE DE METZ

TABLEAU I

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 13 DECEMBRE 1983.

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
0					nd
1	9 °C	ill.	ill.	82	820
2	6 °C	ill.	32	1	30
3	8 °C	ill.	24	3	25
4	7 °C	ill.	ill.	332	3320
5	7 °C	ill.	ill.	35	350
6	8 °C	ill.	231	22	230
7	8 °C	ill.	ill.	355	3550
8	5 °C	ill.	ill.	48	480
9	9 °C	ill.	ill.	51	510
10	6 °C	ill.	ill.	459	4590
11	8 °C	ill.	ill.	324	3240
12	7 °C	ill.	ill.	371	3710
13	9 °C	ill.	ill.	215	2150
14	7,5°C	ill.	ill.	174	1740
15	8 °C	ill.	ill.	286	2860
16	8 °C	ill.	ill.	112	1120
17	8 °C	ill.	ill.	31	310
18	7 °C	ill.	ill.	28	280
19	7,5°C	0	0	0	0
20	8,5°C	30	2	0	3
21	7 °C	ill.	159	10	154
22	5 °C	ill.	42	1	39
23	9,5°C	10	1	0	1
24	7 °C	1	0	0	0
25	6 °C	99	5	0	9
26	6,5°C	4	1	0	0

TABLEAU I (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
54	8 °C	ill.	257	16	248
55	7 °C	ill.	27	1	25
56	9 °C	133	16	1	14
57	8 °C	ill.	85	15	91
58	8 °C	ill.	93	10	94
59	8 °C	ill.	28	3	28
60	7 °C	138	14	3	14
61	11 °C	ill.	448	40	444
62	11 °C	ill.	ill.	124	1240
63	9 °C	ill.	ill.	235	2350
64	9 °C	ill.	151	17	153
65	8 °C	ill.	62	8	64
66	7 °C	ill.	67	5	65
67	10 °C	52	5	0	5
68	7 °C	51	8	0	5
69	8 °C	ill.	67	9	69
70	7 °C	ill.	46	7	48
71	7 °C	89	6	0	9
72	8 °C	54	4	1	5
73	8 °C	ill.	30	4	31
74	6,5°C	2	0	0	0
75	8 °C	7	1	0	1
76	8,5°C	57	2	0	5
77	7 °C	ill.	ill.	65	650
78	7,5°C	ill.	ill.	61	610
79	nd	ill.	29	6	32
80	7,5°C	ill.	40	6	42

TABLEAU I (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
81	nd	ill.	ill.	74	740
82	8,5°C	ill.	ill.	75	750
83	nd	ill.	ill.	75	750
84	9 °C	ill.	ill.	138	1380
85	nd	ill.	179	18	179
86	8,5°C	ill.	93	8	92
87	7,5°C	ill.	ill.	32	320
88	6 °C	11	0	0	1
89	8 °C	228	19	2	22
90	8,5°C	ill.	147	13	145
91	8 °C	52	3	0	5
92	8 °C	7	0	0	1
93	8 °C	ill.	ill.	15	150
94	8 °C	5	0	0	0
95	8 °C	14	3	1	2
96	8,5°C	ill.	17	2	17
97	7 °C	ill.	19	1	18
98	8 °C	ill.	10	3	12
99	8 °C	ill.	14	4	16
100	8 °C	ill.	192	20	193
101	8 °C	ill.	35	4	35

ill. : illisible

nd : non déterminée

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

TABLEAU I (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
81	nd	ill.	ill.	74	740
82	8,5°C	ill.	ill.	75	750
83	nd	ill.	ill.	75	750
84	9 °C	ill.	ill.	138	1380
85	nd	ill.	179	18	179
86	8,5°C	ill.	93	8	92
87	7,5°C	ill.	ill.	32	320
88	6 °C	11	0	0	1
89	8 °C	228	19	2	22
90	8,5°C	ill.	147	13	145
91	8 °C	52	3	0	5
92	8 °C	7	0	0	1
93	8 °C	ill.	ill.	15	150
94	8 °C	5	0	0	0
95	8 °C	14	3	1	2
96	8,5°C	ill.	17	2	17
97	7 °C	ill.	19	1	18
98	8 °C	ill.	10	3	12
99	8 °C	ill.	14	4	16
100	8 °C	ill.	192	20	193
101	8 °C	ill.	35	4	35

ill. : illisible

nd : non déterminée

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

TABLEAU II

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 10 JANVIER 1984.

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
0	nd	0	0	0	0
1	8,5°C	ill.	139	10	135
2	6 °C	ill.	44	2	42
3	8 °C	ill.	151	13	149
4	7 °C	ill.	ill.	1000	10000
5	7 °C	ill.	ill.	79	790
6	7,5°C	ill.	ill.	24	240
7	8 °C	ill.	ill.	938	9380
8	6 °C	ill.	ill.	32	320
9	13 °C	ill.	ill.	242	2420
10	6 °C	ill.	ill.	888	8880
11	9 °C	ill.	ill.	136	1360
12	10 °C	ill.	ill.	828	8280
13	9 °C	ill.	ill.	55	550
14	8 °C	ill.	ill.	55	550
15	7,5°C	ill.	171	21	175
16	8 °C	ill.	197	11	189
17	9,5°C	ill.	ill.	29	290
18	8 °C	ill.	ill.	25	250
19	8 °C	132	12	0	13
20	10 °C	ill.	ill.	31	310
21	6 °C	1	0	0	0
22	5 °C	ill.	ill.	1000	10000
23	9 °C	25	3	0	3
24	7 °C	1	0	0	0
25	8 °C	1	0	0	0
26	7 °C	1	0	0	0

TABLEAU II (suite 1)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTRÉE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
27	7,5°C	1	0	0	0
28	8 °C	34	2	0	3
29	7,5°C	3	0	0	0
30	11 °C	ill.	164	12	160
31	6 °C	0	0	0	0
32	7 °C	0	0	0	0
33	9 °C	45	5	1	5
34	7,5°C	9	0	0	1
35	7,5°C	16	1	0	2
36	7 °C	0	0	0	0
37	8 °C	0	0	0	0
38	7 °C	13	2	0	1
39	7 °C	124	7	1	12
40	8 °C	17	2	0	2
41	8 °C	24	3	0	2
42	8 °C	ill.	14	0	13
43	8 °C	1	0	0	0
44	nd	12	0	0	1
45	7 °C	0	0	0	0
46	10 °C	ill.	96	12	98
47	10 °C	ill.	64	8	65
48	7 °C	ill.	20	0	18
49	6 °C	ill.	ill.	16	160
50	9 °C	ill.	59	4	57
51	7 °C	ill.	ill.	112	1120
52	7 °C	ill.	ill.	110	1100
53	7 °C	ill.	114	10	113

TABLEAU II (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
54	7 °C	71	5	0	7
55	8 °C	ill.	92	15	97
56	9 °C	271	29	3	27
57	10 °C	ill.	99	10	99
58	8 °C	ill.	138	19	143
59	9 °C	ill.	220	23	221
60	8 °C	127	11	2	13
61	9 °C	141	15	2	14
62	10 °C	ill.	38	4	38
63	10 °C	ill.	16	2	16
64	8 °C	ill.	48	4	47
65	9 °C	ill.	ill.	43	430
66	9 °C	ill.	9	0	8
67	8,5°C	ill.	84	14	89
68.	8 °C	47	5	0	5
69	9 °C	55	6	2	6
70	9 °C	ill.	15	1	15
71	8 °C	ill.	10	0	9
72	8 °C	26	2	0	3
73	8,5°C	ill.	17	0	15
74	8 °C	0	0	0	0
75	8,5°C	0	0	0	0
76	8,5°C	6	0	0	1
77	7,5°C	ill.	33	2	32
78	7 °C	61	11	2	7
79	9,5°C	17	4	0	2
80	8,5°C	84	11	0	9

TABLEAU II (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
81	8 °C	ill.	ill.	22	220
82	8 °C	ill.	ill.	22	220
83	8 °C	21	3	0	2
84	8 °C	ill.	73	7	73
85	7,5°C	ill.	50	6	51
86	nd	ill.	74	7	74
87	8 °C	5	0	0	0
88	8 °C	36	3	0	4
89	9 °C	26	2	1	3
90	7 °C	189	17	3	19
91	8 °C	2	0	0	0
92	9 °C	1	0	0	0
93	8 °C	3	0	0	0
94	8 °C	13	0	0	1
95	7 °C	17	2	0	2
96	8,5°C	35	4	1	4
97	7 °C	136	14	1	14
98	7 °C	98	10	0	10
99	9 °C	ill.	18	1	17
100	7,5°C	117	13	1	12
101	7 °C	ill.	134	14	135

ill. : illisible

nd : non déterminé

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

TABLEAU III

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 21 FEVRIER 1984

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE b)			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
1	6 °C	41	5	0	41
2	6 °C	10	0	0	9
3	7 °C	98	8	1	96
4	6 °C	ill.	ill.	120	12000
5	5,5°C	ill.	38	1	355
6	5,5°C	15	1	0	14
7	8 °C	ill.	ill.	175	17500
8	7 °C	58	5	0	57
9	12 °C	ill.	50	3	482
10	6 °C	ill.	ill.	108	10800
11	7 °C	ill.	95	6	918
12	8,5°C	ill.	ill.	31	3100
13	7 °C	36	4	0	36
14	6 °C	113	12	0	113
15	5,5°C	225	20	0	221
16	6,5°C	184	20	3	186
17	7 °C	ill.	106	8	1036
18	6 °C	ill.	ill.	71	7100
19	7 °C	0	0	0	0
20	8,5°C	5	2	0	6
21	6 °C	19	2	0	19
22	5 °C	107	11	3	109

ill. : illisible

nd : non déterminé

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

b) : la concentration bactérienne des échantillons prélevés aux points n°1 à 22 a été déterminée en filtrant 1 ml, 0,1 ml et 0,01 ml de la suspension originelle alors que des volumes de 10 ml, 1 ml et 0,1 ml ont été utilisés dans tous les autres cas

TABLEAU III (suite 1)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE <sup>b)</sup>			CONCENTRATION BACTERIENNE par ml <sup>a)</sup>
		10 ml	1 ml	0,1 ml	
0	5 °C	1	0	0	0
23	8 °C	6	1	0	1
24	6,5°C	ill.	12	3	14
25	6,5°C	49	1	0	5
26	6 °C	21	4	0	2
27	6 °C	23	5	0	3
28	7 °C	ill.	16	3	17
29	8,5°C	208	23	3	21
30	11 °C	ill.	36	7	39
31	14,5°C	ill.	23	4	25
32	6 °C	ill.	14	2	15
33	7,5°C	0	0	0	0
34	8 °C	ill.	86	9	86
35	6 °C	13	4	0	2
36	6 °C	1	0	0	0
37	7,5°C	4	1	0	0
38	8 °C	ill.	ill.	126	1260
39	7 °C	2	0	0	0
40	5 °C	0	0	0	0
41	7 °C	ill.	148	23	155
42	7 °C	29	3	0	3
43	7 °C	6	2	1	1
44	7 °C	5	2	0	1
45	5 °C	11	3	0	1
46	14 °C	ill.	ill.	91	910
47	9 °C	ill.	ill.	309	3090
48	8 °C	ill.	104	15	108

TABLEAU III (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTRÉE b)			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
49	6 °C	ill.	61	7	62
50					nd
51	8 °C	ill.	ill.	194	1940
52	7 °C	ill.	ill.	73	730
53	9 °C	ill.	140	14	140
54	10 °C	ill.	ill.	398	3980
55	6 °C	ill.	ill.	41	410
56	12 °C	ill.	69	11	73
57	8 °C	ill.	215	25	218
58	10 °C	ill.	19	8	12
59	8 °C	ill.	67	8	68
60	7 °C	ill.	ill.	22	220
61	9 °C	99		2	10
62	12 °C	ill.	23	6	26
63	8 °C	ill.	ill.	30	300
64	7 °C	88	8	0	9
65	7 °C	ill.	78	6	76
66	8 °C	ill.	70	4	67
67	7 °C	ill.	144	16	145
68	7 °C	ill.	68	7	68
69	7 °C	ill.	39	3	38
70	9 °C	ill.	101	13	104
71	8 °C	ill.	26	0	24
72	13 °C	131	9	0	13
73	8 °C	34	0	0	3
74	6,5°C	1	0	0	0
75	6,5°C	9	0	0	1

TABLEAU III (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTRÉE b)			CONCENTRATION BACTERIENNE par ml a)
		10 ml	1 ml	0,1 ml	
76	7 °C	4	0	0	0
77	6 °C	ill.	40	4	40
78	6 °C	ill.	97	7	95
79	6 °C	ill.	ill.	22	220
80	7 °C	165	8	0	16
81	9 °C	173	13	2	17
82	8 °C	139	21	2	15
83	6,5°C	56	8	1	6
84	7 °C	114	12	0	11
85	6,5°C	ill.	35	5	36
86	6 °C	ill.	122	16	125
87	5,5°C	ill.	23	3	24
88	6,5°C	ill.	43	4	43
89	7 °C	ill.	25	0	23
90	6 °C	ill.	22	5	25
91	7 °C	4	0	0	0
92	6 °C	13	2	0	1
93	7 °C	1	0	0	0
94	7 °C	0	0	0	0
95	6,5°C	36	1	0	3
96	6 °C	ill.	16	1	15
97	6 °C	125	15	1	13
98	6 °C	ill.	89	9	89
99	7 °C	ill.	24	5	26
100	6 °C	ill.	47	5	47
101					nd

TABLEAU IV

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 3 AVRIL 1984

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
0	7 °C	1	0	0	1
1	10 °C	18	2	0	18
2	7 °C	42	4	0	41
3	9 °C	ill.	23	1	218
4	8 °C	ill.	ill.	116	11600
5	8 °C	ill.	45	4	445
6	8 °C	32	6	0	34
7	8,5°C	ill.	ill.	44	4400
8	8 °C	44	7	1	47
9	8,5°C	89	6	0	86
10	12 °C	ill.	ill.	75	7500
11	7,5°C	ill.	14	2	145
12	9,5°C	ill.	58	6	582
13	9 °C	113	13	1	114
14	7 °C	ill.	49	4	482
15	7,5°C	ill.	21	4	227
16	8,5°C	ill.	30	4	309
17	8,5°C	109	11	1	109
18	8 °C	78	8	1	78
19	8,5°C	5	1	0	5
20	8,5°C	29	1	0	27
21	8,5°C	12	1	0	12
22	7 °C	ill.	199	11	1909
23	9 °C	38	8	1	42
24	8,5°C	2	1	0	3
25	nd	2	0	0	2
26	9 °C	8	1	0	8

TABLEAU IV (suite 1)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
27	8 °C	2	0	0	2
28	8 °C	0	0	0	0
29	9,5°C	26	3	0	26
30	11,5°C	40	6	0	41
31	8 °C	ill.	15	0	136
32	8 °C	0	0	0	0
33	9,5°C	1	0	0	1
34	8,5°C	41	5	0	41
35	10 °C	19	1	0	18
36	9 °C	0	0	0	0
37	9,5°C	0	0	0	0
38	9 °C	2	0	0	2
39	8,5°C	3	0	0	3
40	8 °C	11	3	0	13
41	8,5°C	2	0	0	2
42	8,5°C	31	0	0	28
43	9 °C	0	0	0	0
44	8,5°C	4	0	0	4
45	8 °C	1	0	0	1
46	10 °C	ill.	106	9	1045
47	8 °C	91	10	1	92
48	8 °C	33	5	1	35
49	9 °C	ill.	115	11	1145
50	11 °C	ill.	81	9	818
51	8,5°C	ill.	ill.	89	8900
52	8 °C	ill.	42	0	382
53	11 °C	54	3	0	51

TABLEAU IV (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
54	12 °C	143	17	0	144
55	9 °C	74	5	0	71
56	9 °C	80	11	3	85
57	9 °C	63	9	2	67
58	8 °C	67	2	0	62
59	11 °C	72	7	0	71
60	11 °C	97	13	1	100
61	8 °C	12	0	0	11
62	9 °C	47	4	0	46
63	10,5°C	23	1	0	22
64	9 °C	37	5	0	38
65	9 °C	62	12	2	68
66	8 °C	51	7	1	53
67	9 °C	30	5	0	32
68	11 °C	11	1	0	11
69	9 °C	115	15	2	119
70	9 °C	33	4	0	33
71	9 °C	64	0	0	58
72	12 °C	71	3	0	67
73	9 °C	3	0	0	3
74	8,5°C	5	0	0	5
75	8,5°C	4	1	0	5
76	12 °C	186	10	2	178
77	10 °C	ill.	48	2	455
78	7 °C	ill.	34	2	327
79	9 °C	ill.	37	2	355
80	9 °C	55	6	1	56

TABLEAU IV (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
81	8 °C	ill.	38	3	373
82	9 °C	ill.	31	4	318
83	9 °C	57	3	0	54
84	8,5°C	55	4	0	53
85	9 °C	56	3	0	53
86	8 °C	ill.	58	7	591
87	8 °C	91	7	0	88
88	9 °C	101	7	0	97
89	8,5°C	114	7	0	109
90	8,5°C	5	0	0	5
91	8,5°C	0	0	0	0
92	8 °C	31	5	0	32
93	9 °C	1	0	0	1
94	9 °C	5	0	0	5
95	8,5°C	9	2	0	10
96	8 °C	ill.	19	1	182
97	8 °C	105	7	1	102
98	8 °C	67	1	1	62
99	9 °C	68	1	0	62
100	8,5°C	37	1	0	34
101					nd

ill. : illisible

nd : non déterminé

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

TABLEAU V

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 15 MAI 1984

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
0	11 °C	5	0	0	5
1	11 °C	ill.	20	2	200
2	10,5°C	44	5	0	44
3	11,5°C	ill.	45	7	473
4	11 °C	ill.	ill.	220	2200
5	11 °C	ill.	73	8	736
6	12 °C	10	2	0	11
7	12 °C	ill.	261	26	2609
8	11 °C	38	2	0	36
9	13 °C	ill.	17	3	182
10	11,5°C	ill.	ill.	32	3200
11	11 °C	ill.	40	2	382
12	14 °C	ill.	72	6	709
13	12 °C	ill.	47	5	473
14	12 °C	ill.	32	4	327
15	11,5°C	ill.	33	3	327
16	11 °C	ill.	53	7	545
17	12,5°C	92	10	2	94
18	11 °C				nd
19	12,5°C	8	2	0	9
20	11,5°C	56	6	0	56
21	12 °C	9	1	0	9
22	10,5°C	23	5	1	26
23	11 °C	25	2	0	24
24	12,5°C	ill.	19	2	191
25	12 °C	60	7	0	60
26	11 °C	24	4	0	25

TABLEAU V (suite 1)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
27	11,5°C	36	4	0	36
28	11 °C	81	9	0	81
29	11,5°C	2	1	0	3
30	13,5°C	37	4	0	37
31	13 °C	13	2	0	14
32	11,5°C	20	3	0	21
33	12,5°C	41	5	0	41
34	12 °C	ill.	11	0	100
35	11,5°C	59	7	1	60
36	12,5°C	28	3	0	28
37	14 °C	10	2	0	11
38	12 °C	7	1	0	7
39	12 °C	27	3	0	27
40	12 °C	ill.	66	7	664
41	12 °C	24	3	0	24
42	12,5°C	45	2	0	42
43	12,5°C	3	0	0	3
44	12 °C	38	4	0	38
45	12,5°C	17	3	1	19
46	15 °C	ill.	ill.	22	2200
47	13 °C	ill.	103	9	1018
48	13 °C	ill.	90	6	873
49	12 °C	ill.	45	6	464
50	14 °C	ill.	28	3	282
51	15 °C	ill.	155	11	1509
52	11 °C	36	0	0	32
53	14 °C	7	1	0	7

TABLEAU V (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
54	15 °C	74	5	1	72
55	13 °C	ill.	116	9	1136
56	nd	ill.	60	10	636
57	12 °C	ill.	21	1	200
58	15 °C	ill.	26	4	273
59	12 °C	83	5	0	79
60	12 °C	101	9	0	99
61	11 °C	ill.	56	5	555
62					nd
63	12 °C	71	5	0	68
64	12 °C	ill.	57	3	545
65	11 °C	73	8	0	73
66	12,5°C	116	13	0	116
67	12 °C	37	3	0	36
68	14 °C	36	0	0	32
69	12 °C	ill.	17	0	155
70	12 °C	69	8	0	69
71	12 °C	23	5	0	25
72	15,5°C	28	2	0	27
73	11 °C	15	3	0	16
74					nd
75					nd
76	13,5°C	50	8	2	54
77	14 °C	ill.	ill.	18	1800
78	11 °C	ill.	44	8	473
79	13 °C	ill.	158	32	1727
80	12 °C	88	7	1	86

TABLEAU V (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
81	12 °C	ill.	67	7	673
82	13 °C	ill.	44	4	436
83	15 °C	ill.	35	6	373
84	12 °C	74	5	0	71
85	12 °C	ill.	53	7	545
86	11,5°C	ill.	69	6	682
87	12,5°C	ill.	64	13	700
88	12 °C	ill.	27	7	309
89	12,5°C	ill.	59	6	591
90	11 °C	119	6	2	114
91	12,5°C	5	1	0	5
92	12,5°C	2	0	0	2
93	14 °C	16	0	0	14
94	12 °C	38	10	2	45
95	12 °C	84	6	1	82
96	12 °C	107	13	2	110
97	12 °C	32	4	0	32
98	12 °C	84	8	0	83
99	11,5°C	47	6	1	49
100	12 °C	25	6	0	28
101	11,5°C	48	7	0	50

ill. : illisible

nd : non déterminé

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche

TABLEAU VI

RESULTATS DES ANALYSES BACTERIOLOGIQUES EFFECTUEES  
LORS DE LA CAMPAGNE DE PRELEVEMENTS DU 26 JUIN 1984

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
0	18,5°C	1	0	0	1
1	17,5°C	ill.	ill.	285	28500
2	14,5°C	42	5	0	42
3	17 °C	ill.	30	3	300
4	17 °C	ill.	ill.	121	12100
5	15,5°C	ill.	22	3	227
6	16 °C	100	7	0	96
7	17 °C	ill.	ill.	52	5200
8	17 °C	ill.	16	3	173
9	17 °C	ill.	31	3	309
10	15 °C	ill.	ill.	56	5600
11	15 °C	ill.	ill.	50	5000
12	15 °C	ill.	ill.	24	2400
13	16,5°C	ill.	ill.	261	26100
14	17,5°C	ill.	ill.	86	8600
15	17 °C	ill.	26	2	255
16	15,5°C	ill.	ill.	49	4900
17	17,5°C	ill.	ill.	12	1200
18	16 °C	ill.	66	3	627
19	17,5°C	7	1	0	7
20	14,5°C	ill.	151	14	1500
21	17,5°C	13	0	0	12
22	16 °C	ill.	53	4	518
23	15 °C	9	2	0	10
24	18 °C	18	1	0	17
25	18 °C	2	0	0	2
26	17 °C	14	2	0	14

TABLEAU VI (suite 1)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTRÉE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
27	nd	12	1	0	12
28	17 °C	0	0	0	0
29	17 °C	3	0	0	3
30	18 °C	121	12	0	120
31	17,5°C	55	6	0	55
32	17 °C	0	0	0	0
33	16,5°C	ill.	36	1	336
34	16 °C	ill.	42	4	418
35	16,5°C	28	3	1	29
36	17,5°C	41	2	0	39
37	17 °C	26	3	1	27
38	17,5°C	3	0	0	3
39	17,5°C	25	1	0	23
40	18,5°C	108	7	0	104
41	15 °C	ill.	36	5	373
42	17 °C	118	10	0	115
43	18 °C	4	0	0	4
44	15 °C	ill.	29	4	300
45	18 °C	0	0	0	0
46	16 °C	ill.	ill.	394	39400
47	17 °C	ill.	ill.	309	30900
48	16,5°C	ill.	ill.	420	42000
49	16 °C	ill.	ill.	33	3300
50	16 °C	ill.	ill.	221	22100
51	17 °C	ill.	ill.	347	34700
52	16 °C	ill.	ill.	501	50100
53	18,5°C	ill.	ill.	31	3100

TABLEAU VI (suite 2)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTRÉE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
54	16,5°C	ill.	ill.	30	3000
55	16 °C	ill.	ill.	45	4500
56	16,5°C	ill.	ill.	29	2900
57	18 °C	ill.	ill.	77	7700
58	18 °C	ill.	ill.	111	11100
59	15,5°C	ill.	ill.	375	37500
60	16,5°C	ill.	ill.	329	32900
61	15 °C	ill.	ill.	172	17200
62	17,5°C	ill.	ill.	161	16100
63	15,5°C	ill.	ill.	142	14200
64	16 °C	ill.	ill.	234	23400
65	15 °C	ill.	175	12	1700
66	17 °C	ill.	ill.	50	5000
67	17 °C	ill.	ill.	32	3200
68	18,5°C	ill.	27	0	245
69	17 °C	ill.	ill.	217	21700
70	17,5°C	ill.	ill.	44	4400
71	17 °C	ill.	26	3	264
72	21 °C	ill.	78	12	818
73	15,5°C	ill.	26	4	273
74	17 °C	ill.	119	13	1200
75	17 °C	53	6	1	54
76	17 °C	ill.	46	8	491
77	17 °C	ill.	ill.	54	5400
78	16 °C	ill.	101	8	991
79	17 °C	ill.	327	34	3282
80	17 °C	ill.	43	3	418

TABLEAU VI (suite 3 et fin)

POINT DE PRELEVEMENT	TEMPERATURE DE L'EAU	VOLUME D'EAU FILTREE			CONCENTRATION BACTERIENNE par ml a)
		1 ml	0,1 ml	0,01 ml	
81	17,5°C	ill.	ill.	45	4500
82	17 °C	ill.	95	6	918
83	16 °C	ill.	84	6	818
84	17 °C	ill.	ill.	48	4800
85	17 °C	ill.	41	3	400
86	17 °C	ill.	72	7	718
87	17 °C	ill.	133	18	1373
88	17 °C	ill.	21	1	200
89	15,5°C	ill.	66	4	636
90	16 °C	ill.	6	0	55
91					nd
92	17 °C	73	14	0	78
93	17 °C	3	0	0	3
94	17 °C	18	1	0	17
95	16,5°C	18	4	0	20
96	17 °C	ill.	166	11	1609
97	16,5°C	ill.	56	3	536
98	16,5°C	66	3	0	62
99	16,5°C	71	7	1	71
100	17 °C	132	12	0	130
101	14,5°C	62	5	0	60

ill. : illisible

nd : non déterminé

a) : titre bactérien (numération totale à 20°C) le plus vraisemblable arrondi à l'entier le plus proche