



**HAL**  
open science

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

Armand Maul

► **To cite this version:**

Armand Maul. Définition d'une stratégie de prélèvement pour l'analyse microbiologique des eaux de surface. Sciences du Vivant [q-bio]. Université Paul Verlaine - Metz, 1982. Français. NNT : 1982METZ009S . tel-01992562

**HAL Id: tel-01992562**

**<https://hal.univ-lorraine.fr/tel-01992562>**

Submitted on 24 Jan 2019

**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>

UNIVERSITE DE METZ

BIBLIOTHEQUE UNIVERSITAIRE  
- METZ

N° inv.	19820175
Cote	S/Mz 82/9
Loc	

CENTRE DES SCIENCES  
DE L'ENVIRONNEMENT

S/MZ

THESE  
PRESENTEE A  
L'UNIVERSITE DE METZ  
EN VUE D'OBTENIR LE TITRE DE  
DOCTEUR DE 3<sup>è</sup> CYCLE  
PAR MAUL A.

---

DEFINITION D'UNE STRATEGIE DE PRELEVEMENT  
POUR L'ANALYSE MICROBIOLOGIQUE  
DES EAUX DE SURFACE

---



SOUTENUE PUBLIQUEMENT LE 12 MAI 1982  
DEVANT LA COMMISSION D'EXAMEN :

J.C. BLOCK .....Président

B. BALEUX  
D. BEYTOUT  
B. HUGUES  
G. LARBAIGT  
O. NASI

Examineurs

AVANT - PROPOS

Cette étude a été réalisée au laboratoire de microbiologie du Centre des Sciences de l'Environnement de Metz. Que son directeur, Monsieur le Professeur Jean-Claude BLOCK trouve ici l'expression de ma profonde gratitude pour m'y avoir accueilli. Qu'il soit également remercié pour les conseils et les encouragements qu'il m'a prodigués tout au long de ce travail.

Je suis très reconnaissant à Messieurs BALEUX B., BEYTOUT D., HUGUES B., LARBAIGT G. et NASI O., pour m'avoir fait l'honneur de participer au Jury de cette thèse. Je leur adresse mes plus vifs remerciements pour l'attention avec laquelle ils ont examiné ce travail.

Mes remerciements vont également à Mesdames DOLLARD M.A. et PELTRE M.C. pour la compétence et la minutie dont elles ont fait preuve au cours de leur 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.

A Corinne, Franck et Olivier

A toute ma famille

A tous mes amis

S O M M A I R E

INTRODUCTION .....	1
PREMIERE PARTIE - PROPOSITION D'UNE TECHNIQUE DE TITRAGE UTILISANT LE PRINCIPE DU NOMBRE LE PLUS PROBABLE .....	4
<u>CHAPITRE I - LA METHODE DU NOMBRE LE PLUS PROBABLE (N.P.P.) - ETUDE THEORIQUE DE L'ESTIMATEUR .....</u>	<u>6</u>
I-1 Calcul du N.P.P. ....	6
I-1-1 Hypothèses fondamentales .....	6
I-1-2 Définition et interprétation graphique du N.P.P.....	8
I-2 Etude qualitative de l'estimation par le N.P.P.....	11
I-2-1 Loi de probabilité du N.P.P.,.....	11
I-2-2 Expression de l'écart-type.....	13
I-2-3 Conditions d'application des techniques statistiques.....	17
I-2-4 Exploitation statistique des résultats expérimentaux.....	18
I-2-4-1 Calcul d'un intervalle de confiance .....	18
I-2-4-2 Analyse de variance.....	19
I-2-5 Conclusion.....	22
<u>CHAPITRE II - COMPARAISON EXPERIMENTALE DE DEUX TECHNIQUES DE TITRAGE : METHODE ENUMERATIVE SUR MILIEU GELOSE versus METHODE QUAN- TIQUE EN MILIEU LIQUIDE.....</u>	<u>24</u>
II-1 Matériel et méthodes .....	24
II-1-1 Echantillons d'eau .....	24
II-1-2 Numération des coliformes thermorésistants sur milieu gélosé.....	25
II-1-3 Numération des coliformes thermorésistants en milieu liquide .....	25
II-1-4 Analyse statistique des quantifications bactériennes.....	26
II-2 Protocole expérimental.....	27
II-3 Résultats.....	30
II-3-1 Conformité des résultats expérimentaux avec le modèle théorique.....	33
II-3-2 Comparaison des titres bactériens obtenus par les deux méthodes de titrage .....	35
II-3-3 Proportion d' <u>Escherichia coli</u> parmi les coliformes thermorésistants.....	41
II-4 Discussion .....	42



DEUXIEME PARTIE - MISE EN EVIDENCE DE L'HETEROGENEITE D'UN MILIEU HYDRIQUE:		
CONSEQUENCES SUR L'ESTIMATION DES POPULATIONS BACTERIEN-		
	NES.....	45
<u>CHAPITRE I - GENERALITES SUR LA REPARTITION SPATIALE DES INDIVIDUS D'UNE</u>		
	<u>POPULATION</u> .....	47
I-1	Différentes configurations possibles .....	47
I-2	Mise en évidence de l'hétérogénéité.....	49
<u>CHAPITRE II - ETUDE EXPERIMENTALE DE QUELQUES CAUSES D'ERREUR DANS L'ANA-</u>		
	<u>LYSE DES ECHANTILLONS</u> .....	51
II-1	Matériel et méthodes.....	51
II-2	Protocole expérimental .....	52
II-3	Résultats.....	55
II-4	Discussion.....	61
<u>CHAPITRE III - ETUDE EXPERIMENTALE DE L'HETEROGENEITE SPATIO-TEMPORELLE</u>		
	<u>D'UN MILIEU HYDRIQUE</u> .....	65
III-1	Matériel et méthodes.....	65
III-1-1	Situation des points de prélèvement et collecte des échantillons.....	65
III-1-2	Numération des coliformes thermorésistants et analyse statistique des résultats.....	66
III-2	Protocole expérimental.....	66
III-3	Résultats.....	70
III-4	Discussion.....	79
<u>CHAPITRE IV - ETUDE DE LA VARIATION HORAIRE DE LA CONCENTRATION BACTERIENNE</u>		
	<u>D'UN MILIEU HYDRIQUE</u> .....	82
IV-1	Matériel et méthodes.....	82
IV-2	Protocole expérimental.....	83
IV-3	Résultats.....	84
IV-3-1	Variation de la concentration bactérienne de la Moselle sur une période de 12 heures.....	85
IV-3-2	Variation du flux des coliformes thermorésistants dans les rejets de la station d'épuration biologique de Metz.....	91
IV-4	Discussion.....	93

TROISIEME PARTIE - PROPOSITION THEORIQUE D'UNE STRATEGIE DE PRELEVEMENT.....	96
<u>CHAPITRE I - DEFINITION D'UN PROTOCOLE ANALYTIQUE POUR UNE ESTIMATION PONC-</u>	
<u>TUELLE DE LA CONCENTRATION BACTERIENNE D'UN MILIEU HYDRIQUE.....</u>	97
I-1 Intérêt de l'échantillonnage à trois degrés.....	97
I-2 Optimisation du protocole analytique.....	98
<u>CHAPITRE II - APPLICATION DE LA METHODE D'ECHANTILLONNAGE A L'ETUDE ET AU</u>	
<u>CONTROLE DE LA CONTAMINATION BACTERIENNE D'UN COURS D'EAU.....</u>	102
II-1 Etude comparative de la contamination bactérienne chronique entre plusieurs stations de prélèvement.....	102
II-2 Les séries temporelles.....	105
II-2-1 Caractéristiques des séries.....	105
II-2-2 Analyse d'une série.....	106
CONCLUSION .....	109
REFERENCES BIBLIOGRAPHIQUES.....	113
ANNEXES.....	119

I N T R O D U C T I O N

L'analyse bactériologique des eaux s'appuie généralement sur la mise en évidence de germes indicateurs de contamination fécale. Cette recherche s'oriente fréquemment vers la quantification de Escherichia coli et des coliformes fécaux qui sont considérés traditionnellement comme de bons indicateurs de cette forme de pollution.

Cependant, l'évaluation ou la surveillance du niveau de contamination des eaux superficielles se heurte à deux difficultés majeures, en l'occurrence l'imperfection des techniques de titrage et l'hétérogénéité des milieux hydriques naturels.

- Si les titrages bactériens se caractérisent dans l'ensemble par un manque relatif de précision et de spécificité, les propriétés mathématiques de l'estimateur, en laissant préjuger de la rigueur d'une éventuelle exploitation statistique des résultats, constituent un élément déterminant dans le choix d'une méthode de titrage.

- Le défaut de représentativité des échantillons dû à l'hétérogénéité du milieu hydrique diminue la portée des estimations en s'opposant à l'extrapolation des résultats obtenus à partir d'un volume d'eau habituellement très faible. Par ailleurs, les variations temporelles de la densité bactérienne rendent sans grande signification les estimations réalisées en dehors d'une stratégie de prélèvement prenant en considération les différents aspects de l'hétérogénéité dans le temps et dans l'espace.

Ces deux problèmes bien qu'apparemment distincts ne sauraient néanmoins être traités indépendamment l'un de l'autre puisque la qualité des résultats de toute étude microbiologique dépend non seulement de la précision des mesures, mais également de la place qu'elles occupent dans le contexte spatio-temporel. Aussi, le présent travail est-il entièrement tourné vers la définition d'une méthode analytique pouvant servir de base à toute investigation en microbiologie appliquée de l'environnement.

A cet effet, nos efforts se sont portés d'une part vers l'amélioration de la technique de quantification et d'autre part vers la mise au point d'un procédé d'échantillonnage permettant d'élargir le domaine de validité des estimations.

La première partie de ce travail est consacrée à l'étude d'une technique de titrage fondée sur l'application du principe du maximum de vraisemblance. Cette partie comprend deux chapitres.

Les développements théoriques de la méthode sont exposés dans le premier chapitre où nous avons plus particulièrement souligné les propriétés mathématiques de l'estimateur.

L'étude expérimentale, qui repose principalement sur la comparaison des résultats obtenus à partir de la technique proposée avec ceux d'une méthode de référence (énumération des colonies sur milieu gélosé), fait l'objet du deuxième chapitre.

La deuxième partie concerne l'étude de l'hétérogénéité d'un milieu hydrique et ses conséquences sur l'estimation des populations bactériennes. Cette deuxième partie se subdivise en quatre chapitres.

Après quelques généralités sur la répartition spatiale des individus d'une population, l'étude expérimentale de quelques causes d'erreur dans l'analyse des échantillons est abordée dans le deuxième chapitre.

L'étude de plusieurs sources de variation entreprise simultanément dans le troisième chapitre permettra une appréhension globale de l'hétérogénéité spatio-temporelle par la relativisation de ses différentes composantes.

Dans le quatrième chapitre, il est établi, outre un bilan journalier des fluctuations de la concentration bactérienne de la Moselle, une estimation des composantes de variance caractérisant les effets de l'hétérogénéité à chaque niveau de l'échantillonnage.

Faisant état des résultats établis antérieurement, la troisième partie contient les principes théoriques d'une stratégie de prélèvement.

Dans un premier temps nous définissons un protocole analytique pour l'estimation ponctuelle de la concentration bactérienne d'un milieu hydrique, puis dans un deuxième chapitre nous tentons de mettre en relief les avantages de la méthode d'échantillonnage proposée par son application à deux situations particulières se rapportant à l'étude et au contrôle de la contamination bactérienne d'un cours d'eau.

PREMIERE PARTIE

PROPOSITION D'UNE TECHNIQUE DE TITRAGE  
UTILISANT LE PRINCIPE DU NOMBRE LE PLUS PROBABLE

Les méthodes de titrage en bactériologie peuvent être distinguées en deux groupes : les méthodes de type énumératif qui consistent à dénombrer les colonies bactériennes apparues sur un milieu gélosé et les méthodes de type quantique ou de tout ou rien pour lesquelles l'estimation est obtenue à partir des rapports entre le nombre de réponses positives (présence d'une multiplication bactérienne) et le nombre total de milieux inoculés. La méthode de FISHER (WYSHAK et DETRE, 1972) et surtout la méthode connue sous le nom de méthode du nombre le plus probable (N.P.P.) se rangent dans cette deuxième catégorie.

Largement utilisée avant les années 60, la méthode du nombre le plus probable a été quelque peu délaissée au bénéfice de la numération directe sur milieu gélosé. En effet, bien que plus séduisante par son support théorique et, de plus, offrant les avantages de la culture des bactéries en milieu liquide, la méthode du nombre le plus probable a souvent été discréditée par son manque de précision dans les conditions habituelles d'utilisation, c'est-à-dire avec un nombre relativement faible d'inoculums par niveau de dilution (3,5 ou 10 tubes par dilution) (GREEN et LITSKY, 1980; THOMAS et WOODWARD, 1955).

La numération directe des colonies sur milieu gélosé fournit un estimateur dont le calcul repose sur l'estimation du paramètre d'une loi de Poisson; elle se caractérise par une précision appréciable à condition toutefois que le nombre des colonies recensées soit suffisamment élevé. (MAUL, 1981 ; MEYNELL et MEYNELL, 1970).

Dans ce premier chapitre une microméthode de titrage utilisant 96 inoculums par dilution a été étudiée, laquelle, moyennant certaines précautions d'utilisation, fournira un outil de travail répondant aux principales exigences des techniques statistiques mises en oeuvre dans ce travail.

Dès le début du siècle, PHELPS (1908) préconise la dilution de la suspension bactérienne à analyser afin d'en estimer le titre, mais la méthode proposée est empirique et rudimentaire.

Il faut attendre 1915 (MAC GRADY, 1915) pour qu'apparaissent les premiers fondements mathématiques de l'estimation par le nombre le plus probable mais, le calcul se limite à l'estimation de la densité microbienne sans aucune indication quant à la fiabilité du résultat.

En 1917, GREENWOOD et YULE publient l'équation permettant le calcul du N.P.P. sous sa forme définitive.

C'est en 1933 que HALVORSON et ZIEGLER, après une étude minutieuse de la bibliographie, jugent les résultats de l'époque insuffisants et publient une série d'articles dans laquelle ils reconsidèrent complètement le sujet. Leurs travaux comprennent notamment l'élaboration de tables de N.P.P. et les premiers éléments du calcul d'un intervalle de confiance (HALVORSON et ZIEGLER, 1933 a et b).

Depuis, les chercheurs ont surtout concentré leurs efforts sur ces mêmes objectifs. Aussi de nombreuses tables de N.P.P. ont été publiées à ce jour.

Mais c'est à HUGUES (1981) que revient le mérite d'avoir étudié le N.P.P. en tant que variable aléatoire en précisant sa loi de probabilité ainsi que les conditions d'application de l'arsenal puissant des statistiques paramétriques.

En étudiant le problème dans le cas de plus de 10 inoculums par dilution avec les moyens actuels de calcul, il a donné une nouvelle dimension à la méthode en élargissant considérablement son champ d'application.

## CHAPITRE I : LA METHODE DU NOMBRE LE PLUS PROBABLE (N.P.P.) - ETUDE THEORIQUE DE L'ESTIMATEUR.

### I - 1 - CALCUL DU N.P.P.

L'estimation de la densité bactérienne d'une eau par la méthode du N.P.P. nécessite l'ensemencement de milieux de culture par des aliquotes provenant de dilutions successives de la suspension mère. Après incubation, des systèmes révélateurs permettent de mettre en évidence les milieux qui sont le siège d'une croissance bactérienne.

En exploitant les résultats obtenus à partir d'une ou de plusieurs dilutions successives, il est possible d'évaluer la densité bactérienne de la suspension initiale par application du principe du maximum de vraisemblance. Habituellement 3 niveaux de dilution sont pris en considération et le rapport de dilution est généralement 10.

#### I - 1 - 1 Hypothèses fondamentales.

La validité des calculs du N.P.P. développés ci-après repose sur deux hypothèses essentielles :



. Les microorganismes sont distribués au hasard dans la suspension (modèle de POISSON) ce qui signifie que chaque germe a la même probabilité de se trouver dans un inoculum donné.

La conformité des résultats expérimentaux avec cette hypothèse sera étudiée au paragraphe II-3.1.

. La présence d'un microorganisme est suffisante pour provoquer une croissance ; en d'autres termes, un résultat négatif implique l'absence de germe.

D'emblée il est nécessaire de formuler quelques réserves quant à la véracité de ces hypothèses :

a) En effet, une bactérie n'a, dans des conditions données, qu'une certaine probabilité de se développer. Cette probabilité dépend, entre autres, de la nature du germe étudié ainsi que des caractéristiques physico-chimiques des milieux de culture (ASHBY et RHODES-ROBERTS, 1976; EVANS et coll., 1981a; EVANS et coll., 1981 b). La température d'incubation et plus précisément les variations de celle-ci altèrent aussi notablement les résultats d'un titrage. Ainsi dans le cas des eaux usées chlorées et des eaux marines, GREEN et LITSKY (1980) ont montré que les résultats de la numération des coliformes fécaux par filtration sur membranes pouvaient être augmentés d'environ 50% si la période d'incubation à 44,5° C était précédée par une préincubation à 35° C pendant 5 heures. Le stress exercé sur les bactéries par ces deux types d'eau diminue vraisemblablement leur aptitude à croître sur milieu solide sélectif à des températures élevées, et l'élévation progressive de la température assure certainement une meilleure adaptation.

b) De plus, les bactéries d'un milieu hydrique forment au moins quatre populations distinctes (COSTERTON et GEESEY, 1979). Les bactéries peuvent en effet être libres, agrégées ou adsorbées sur des matières en suspension, fixées sur des supports solides immergés (bactéries épilithes et épiphytes) ou enfin intégrées dans les sédiments. Bien que la numération bactérienne dans le milieu environnant n'intéresse en général que les deux premiers types de population, la présence d'agrégats et de complexes bactéries-particules soulève de sérieuses difficultés car les méthodes de titrage précitées ne distinguent pas les microorganismes isolés des agrégats. Il en découle que le résultat d'un titrage exprime en fait le nombre d'unités formatrices de colonies (U.F.C.) comprenant des bactéries libres ainsi que des complexes bactéries-particules.

Cette dénomination doit être évidemment rapprochée du concept d'unités formatrices de plages (U.F.P.) utilisé par les virologistes.

La connaissance de la configuration de cette hétérogénéité microscopique apporterait une solution au problème du titrage bactérien. Dans ce but, SEKI (1968), SOROKIN (1970) et WILLIAMS (1970) ont étudié la proportion de bactéries adsorbées par des particules en milieu marin, ainsi que la taille des agrégats. Il purent établir que 15% à 50% des bactéries sont associées aux matières en suspension et SEKI (1971) montra que le logarithme du nombre de bactéries par agrégat était proportionnel à la masse du support organique. D'autre part, les résultats d'un titrage sont affectés par une modification de la turbidité de l'échantillon, par exemple, par ajout de quantités croissantes d'un sédiment stérile (BREITTMAYER et GAUTHIER, 1979).

La dépendance entre les résultats d'un titrage et la charge particulaire de l'eau montre que la configuration de l'hétérogénéité microscopique est étroitement liée au nombre ainsi qu'à la granulométrie des particules en suspension.

Avant de poursuivre nous faisons nôtre la considération de LURIA (1953) et reconnaissons que, le nombre de germes estimé n'a qu'une valeur statistique. Il n'est lié au nombre de microorganismes réellement présents dans le milieu à titrer que par des facteurs de proportionnalité inconnus dépendant très probablement de nombreux éléments comme la méthode de titrage, l'origine et les caractéristiques de l'eau, la nature du germe étudié, les phénomènes d'antagonisme et de compétition entre les différentes espèces présentes ainsi que des interactions entre tous ces facteurs.

Dans ces conditions, il va de soi que les résultats d'un titrage n'ont qu'une valeur relative et doivent obligatoirement être accompagnés d'une description complète et précise du matériel et de la technique utilisée.

#### I - 1-2 Définition et interprétation graphique du N.P.P.

- \* Si  $k$  désigne le nombre des niveaux de dilution pris en considération,
- . si  $p_i$  est le nombre de cupules du niveau  $i$  donnant une réponse positive
- . et si  $q_i$  est le volume d'inoculum ensemencé dans chacune des  $n_i$  cupules de la dilution de rang  $i$ ,

le  $k$ -uplet caractéristique est formé par la juxtaposition des  $k$  nombres  $p_i$  en commençant par la dilution la plus faible.

\* u étant la densité microbienne, la probabilité associée au k-uplet  $(P_1, \dots, P_i, \dots, P_k)$  est égale à :

$$(1) \quad L(u) = \prod_{i=1}^k \binom{p_i}{n_i} (1 - e^{-uq_i})^{p_i} e^{-uq_i (n_i - p_i)}$$

Cette fonction de u porte le nom de fonction de vraisemblance.

L'estimation de la densité bactérienne par la méthode du maximum de vraisemblance consiste à trouver la valeur de u correspondant au maximum de la fonction L; cette valeur qui n'est autre que le N.P.P. se détermine en annulant la dérivée du logarithme de L(u) (WYSHAK et DETRE, 1972).

Finalement, le N.P.P. est la solution de l'équation de vraisemblance (2) d'inconnue u.

$$(2) \quad \sum_{i=1}^{i=k} (n_i - p_i) q_i = \sum_{i=1}^{i=k} \frac{p_i q_i e^{-uq_i}}{1 - e^{-uq_i}}$$

Ces résultats sont établis en Annexe I.

La résolution de cette équation n'est pas simple, aussi faut-il procéder par approximations successives. Un exemple de programme de résolution par calculateur électronique est proposé en Annexe II.

La fonction de vraisemblance associée au triplet (5,2,0), dans le cas de trois dilutions décimales avec 5 tubes par dilution peut s'expliciter par :

$$L(u) = 10 e^{-0,35u} (1 - e^{-u})^5 (1 - e^{-0,1u})^2$$

Une représentation graphique en est donnée sur la figure 1.

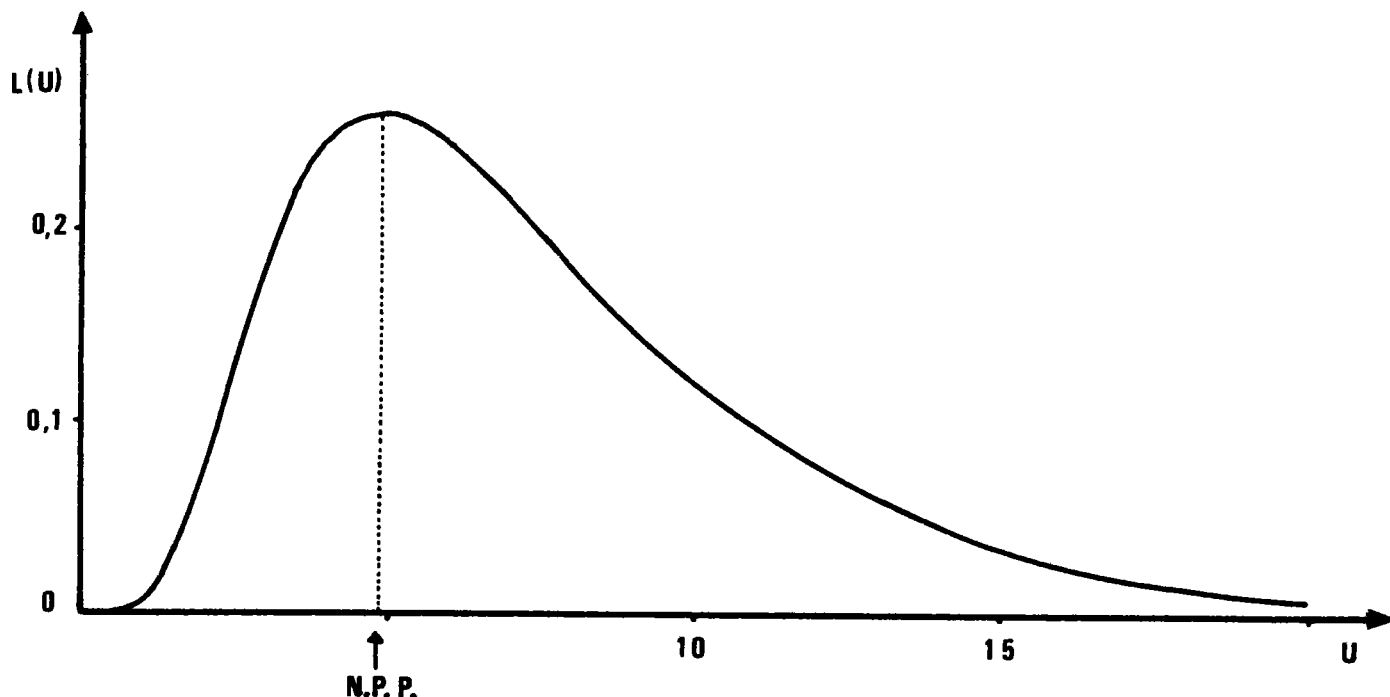


Figure 1 : Probabilité associée au triplet(5,2,0) versus densité bactérienne

Le nombre le plus probable de germes dans un inoculum de la première dilution est dans cet exemple 4,9. Cette valeur correspond au maximum de la fonction L.

Le N.P.P. par lui-même n'a pas une très grande signification si on ne dispose d'aucune indication sur sa fiabilité. Pour cette raison, il faut joindre à la valeur trouvée un intervalle dans lequel la densité réelle de la suspension a une probabilité assez grande de se trouver (en général 95% ou 99%).

D'une manière générale les limites inférieure et supérieure notées respectivement  $L_i$  et  $L_s$ , d'un intervalle de confiance au seuil  $1 - \alpha$  avec  $\alpha = \alpha_1 + \alpha_2$  sont données par les relations tirées de HUGUES (1981) et MAUL (1981).

$$\frac{\int_0^{L_i} L(u) \, du}{\int_0^{+\infty} L(u) \, du} = \alpha_1 \quad \text{et} \quad \frac{\int_0^{L_s} L(u) \, du}{\int_0^{+\infty} L(u) \, du} = 1 - \alpha_2$$

La figure 2 donne l'interprétation graphique des limites  $L_i$  et  $L_s$  avec les hypothèses de l'exemple précédent en prenant  $\alpha_1 = \alpha_2 = 0,025$ .

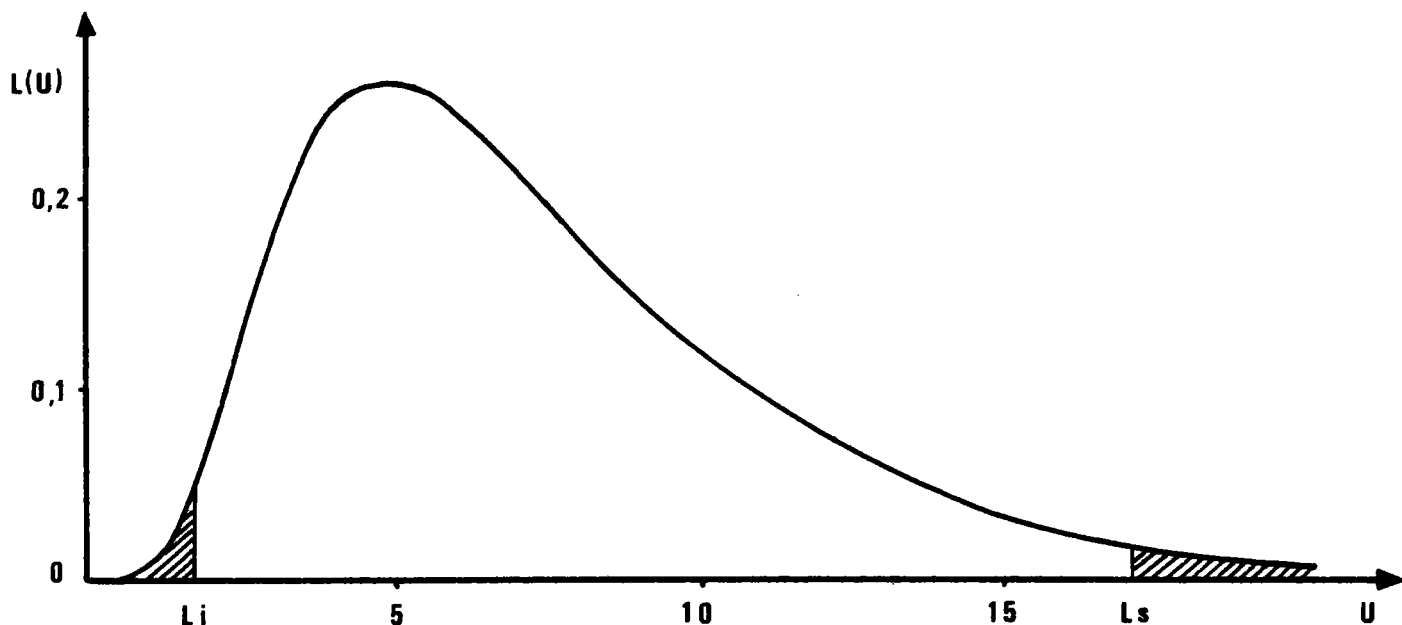


Figure 2 : Limites d'un intervalle de confiance ; interprétation graphique.

Dans ce cas particulier  $L_i$  et  $L_s$  sont telles que chacune des deux portions hachurées représente 2,5 % de l'aire totale de la partie située sous la courbe.

### I - 2. ETUDE QUALITATIVE DE L'ESTIMATION PAR LE N.P.P.

Ce paragraphe se propose d'étudier la conformité des caractéristiques de l'estimateur aux exigences des techniques statistiques mises en oeuvre dans la suite de ce travail.

#### I - 2-1 Loi de probabilité du N.P.P.

Pour une densité bactérienne  $u$  donnée et un nombre déterminé d'inoculum par dilution, chaque triplet, et par conséquent chaque indice N.P.P., a une certaine probabilité,  $L(u)$ , d'apparaître.

Il en résulte que le N.P.P. peut être considéré comme une variable aléatoire discrète, notée X, dont la distribution peut être étudiée (le nombre des valeurs possibles étant égal à  $(n_1 + 1) (n_2 + 1) (n_3 + 1)$ ). Exception étant faite du cas où la densité u attribue une probabilité appréciable à l'obtention de réponses uniformément positives à tous les niveaux de dilution, la moyenne et la variance de cette distribution peuvent être calculées par les méthodes usuelles.

La distribution de X a été étudiée pour différentes valeurs de u (0,01; 0,03 ; 0,06 ; 0,08; 0,105; 0,145; 1,45 ; 14,49 ; 145 ; 180 ; 220 ; 240 ; 260) dans le cas de trois dilutions décimales comprenant chacune 96 inoculums (HUGUES, 1981). D'emblée, le caractère dissymétrique de la distribution a pu être observé, ce qui se traduit par le signe positif du coefficient de symétrie de FISHER.

Cette particularité a suggéré l'étude de la distribution du logarithme de X dans les mêmes conditions.

Chacune des distributions ainsi obtenue a été comparée à la distribution normale ayant la même moyenne et le même écart-type par le procédé graphique des droites de HENRY.

Ces comparaisons montrent que la loi de X tend vers la loi normale quand n, le nombre de cupules inoculées à chaque niveau de dilution, augmente, mais ceci à condition que la densité u ne s'approche pas trop des valeurs extrêmes qui auraient pour effet de rendre les réponses uniformément négatives ou uniformément positives.

En effet dans ces zones critiques, le coefficient de symétrie de FISHER s'accroît progressivement et le caractère dissymétrique de la distribution s'accroît. La dissymétrie de la courbe de fréquence est notablement atténuée après transformation logarithmique de X. De plus, sous réserve que la densité bactérienne reste située en dehors des zones critiques mentionnées ci-dessus, l'approximation par la loi normale est déjà assez bonne lorsque n atteint la valeur 20.

L'asymétrie de la courbe de fréquence de la variable X est à l'origine d'une tendance à la surestimation par le N.P.P. de la densité réelle de la suspension à titrer.

L'ordre de grandeur du biais a été calculé pour différentes valeurs de n à l'aide de la formule (4) du paragraphe I.2.2.

En fait, le biais dépend également, mais dans une moindre mesure, de la densité, ce qui occasionne de légères fluctuations autour des valeurs consignées dans le tableau 1. Par souci de simplification il a été fait abstraction de ce phénomène.

Tableau 1 : Influence du nombre d'inoculums sur l'importance du biais (MAUL,1981)

nombre de tubes (cupules) utilisés pour chaque dilution	biais exprimé en % de la densité réelle
3	31
5	18
10	8
20	4
96	1

Le tableau 1 fait apparaître une diminution du biais lorsque n augmente; il devient négligeable pour n=96. Ainsi le N.P.P. est un estimateur asymptotiquement non biaisé de la densité réelle.

De toutes ces considérations, il s'ensuit que, si plusieurs triplets sont disponibles pour l'estimation de la densité bactérienne d'un échantillon, la moyenne géométrique des résultats est préférable à la moyenne arithmétique mais la différence entre ces deux moyennes diminue avec l'augmentation de n.

#### I - 2-2 Expression de l'écart-type.

La théorie du maximum de vraisemblance, et en particulier la relation existant entre la variance de l'estimateur et la quantité d'information de FISHER, permet le calcul de l'écart-type de la distribution de X (MATHER, 1965) et par suite l'écart-type, noté  $\sigma_{\log}$ , de la distribution du logarithme décimal de X (LELLOUCH et LAZAR, 1974). La valeur de  $\sigma_{\log}$  est donnée par la relation (3)

$$(3) \quad \sigma_{\log} = \frac{1}{u \log 10} \left[ \sum_{i=1}^k \frac{p_i q_i e^{-uq_i}}{(1 - e^{-uq_i})^2} \right]^{-\frac{1}{2}}$$

Les termes entrant dans la sommation de la formule (3) expriment la contribution à la quantité totale d'information de FISHER des résultats de chaque dilution et de ce fait montrent que celles-ci ne tiennent pas des rôles équivalents dans l'estimation.

Il apparaît que l'écart-type, et par conséquent la précision de l'estimation, est fonction de la densité bactérienne. Cette relation est concrétisée par le graphique de la figure 3 où  $\sigma_{\log}$  est obtenu en remplaçant les  $\pi_i$  de la formule (3) par leurs estimations, compte tenu de la densité. Ceci a pour effet d'exprimer non pas la précision réelle liée à un k-uplet particulier mais plutôt une estimation de l'écart-type théorique que permet une expérience effectuée dans ces conditions.

La courbe de la figure 3, réalisée pour trois dilutions décimales avec  $n=96$ , met en évidence une large plage de densité dans laquelle  $\sigma_{\log}$  varie relativement peu pour prendre très rapidement des valeurs considérables dans les zones situées de part et d'autre de cet intervalle où les réponses tendent à devenir respectivement uniformément négatives ou positives.



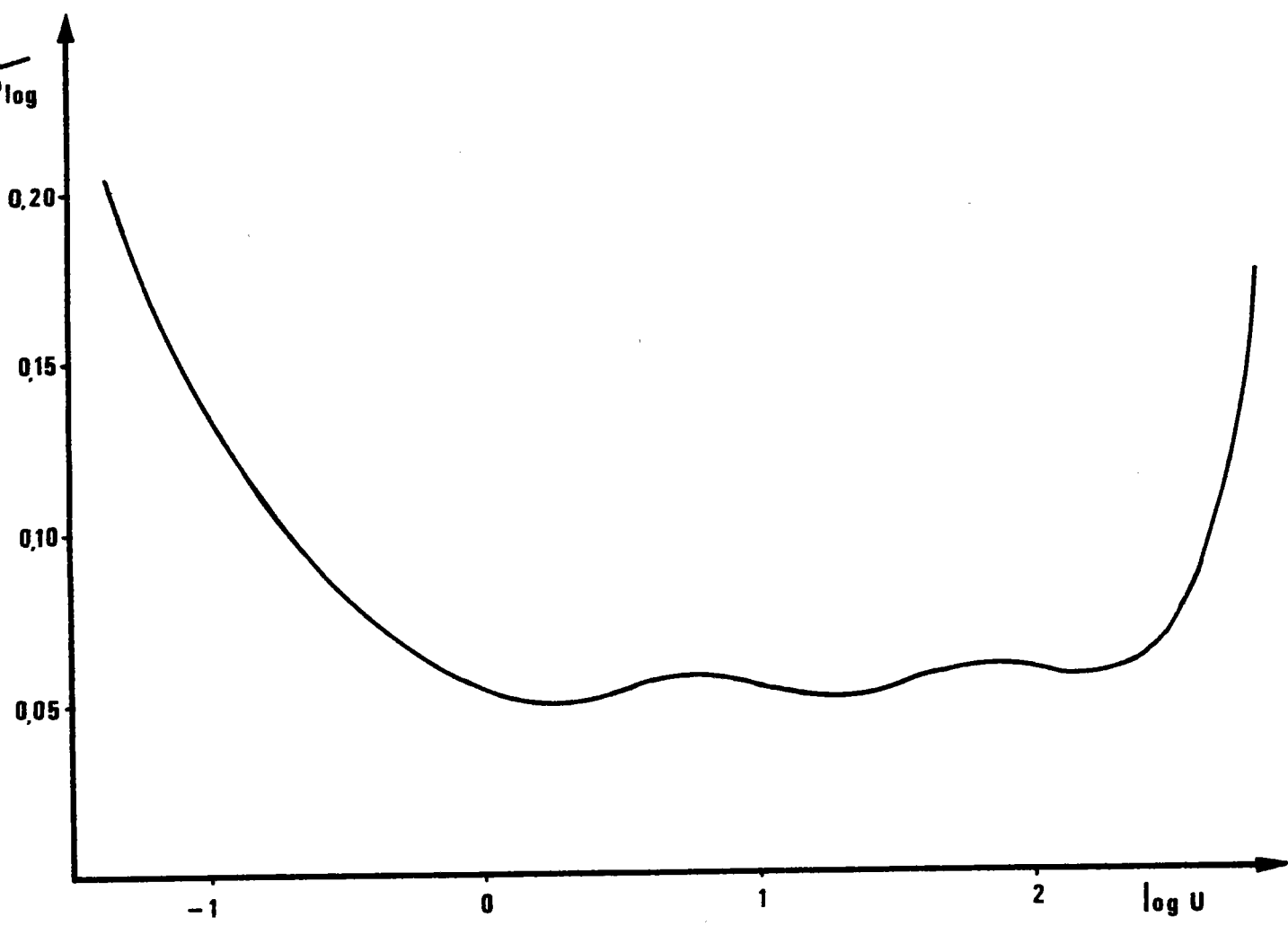


Figure 3: Ecart-type de la distribution de log X versus logarithme de la densité bactérienne dans le cas de trois dilutions décimales avec 96 inoculums par dilution.

Les fluctuations de  $\sigma_{\log}$  observables dans la partie intermédiaire trouvent leur explication dans la disparité des contributions de chaque niveau de dilution à l'estimation globale. Le maximum d'information, au sens statistique du terme (et par voie de conséquence, la meilleure précision) est atteint pour une densité moyenne de 1,6 microorganismes par inoculum (HALDANE, 1939), ce qui correspond à une proportion de 80% de réponses positives.

Ce résultat est à l'origine des minimums de  $\sigma_{\log}$  pour les densités proches de 1,6 et des deux premiers multiples décimaux de cette valeur.

Il est intéressant de noter que l'amplitude de ces fluctuations est très modérée ; les précisions, exprimées par le coefficient de variation, entre les valeurs extrêmes étant respectivement : 11,3% et 13,9%. Un nombre  $n$  plus faible aurait pour effet d'accentuer ces écarts, ainsi pour  $n=10$  le calcul donne : 35,8% et 44,8%.

La relative indépendance de  $\sigma_{\log}$  avec la densité microbienne dans l'intervalle susvisé avait déjà été mentionnée par HALVORSON et ZIEGLER (1933b), lesquels ont proposé une formule exprimant  $\sigma_{\log}$  en fonction de  $n$ , soit :

$$(4) \quad \sigma_{\log} = \frac{0,55}{\sqrt{n}}$$

La relation (4), bien que nettement moins élaborée que la formule (3), est cependant loin d'être dépourvue d'intérêt car elle met l'accent sur le lien privilégié entre la précision de l'estimation et le nombre de cupules inoculées à chaque dilution.

Le tableau 2 donne l'ordre de grandeur de la précision, exprimée par le coefficient de variation de la distribution de  $X$ , pour différentes valeurs de  $n$ .

Tableau 2 : Influence du nombre d'inoculum sur la précision de l'estimation.

Nombre d'inoculum par dilution	précision exprimée en %
5	60
10	40
20	30
96	13
200	9

Les résultats du tableau 2 montrent une amélioration très rapide de la précision pour les petites valeurs de n mais progressivement le gain de précision devient de plus en plus difficile à acquérir et risque finalement, à partir d'un certain seuil, de ne plus justifier les moyens mis en oeuvre.

I - 2-3 Conditions d'application des techniques statistiques.

Les considérations exposées aux paragraphes I-2-1 et I-2-2, nous conduisent à définir deux intervalles  $I_1$  et  $I_2$  ayant une importance primordiale dans les applications statistiques du N.P.P. (HUGUES, 1981).

Ainsi l'intervalle  $I_1$  sera formé par l'ensemble des valeurs de log N.P.P. autres que celles qui attribuent à la densité réelle une forte probabilité d'appartenir aux distributions s'écartant trop de la courbe de Gauss. Et l'intervalle  $I_2$  comprendra l'ensemble des valeurs de log N.P.P. autres que celles accordant à la densité réelle une forte probabilité d'appartenir aux distributions dont l'écart-type est lié à la densité bactérienne.

Bien qu'il soit important de souligner le caractère relativement subjectif des hypothèses servant de base de calcul des intervalles  $I_1$  et  $I_2$ , nous mentionnerons dans le tableau 3, à titre de référence, les valeurs limites de ces intervalles préconisées par HUGUES (1981). De plus, le N.P.P. étant dans une certaine mesure lié au nombre total de cupules positives, il est commode d'exprimer les bornes de  $I_1$  et  $I_2$  en pourcentage de cupules positives.

Tableau 3 : Définition des intervalles  $I_1$  et  $I_2$  (modifié d'après HUGUES, 1981)

Nombre de cupules n	Intervalle $I_1$						Intervalle $I_2$					
	Borne inférieure			Borne supérieure			Borne inférieure			Borne supérieure		
	N.P.P.	log N.P.P.	% +	N.P.P.	log N.P.P.	% +	N.P.P.	log N.P.P.	% +	N.P.P.	log N.P.P.	% +
n = 40	0,59	-0,230	17 %	92	1,964	86%	1,01	0,003	25%	157	2,197	93 %
n = 96	0,22	-0,653	7 %	158	2,200	93%	0,85	-0,069	22%	186	2,269	95%

% + : pourcentage de cupules positives.

Les chiffres de ce tableau mettent en évidence une augmentation de l'amplitude des intervalles  $I_1$  et  $I_2$  avec le nombre  $n$ . L'importance de cette propriété apparaîtra clairement ci-après.

I - 2-4 Exploitation statistique des résultats expérimentaux

I - 2-4-1 Calcul d'un intervalle de confiance

Une première application des propriétés qui précèdent est le calcul d'un intervalle de confiance, associé à un résultat expérimental, par des procédés autres que la méthode exacte décrite au paragraphe I-1-2.

La densité observée étant comprise dans l'intersection des deux intervalles  $I_1$  et  $I_2$ , les limites  $Li$  et  $Ls$  d'un intervalle de confiance peuvent être calculées par deux méthodes reposant sur la normalité de la distribution de la variable aléatoire  $\log X$  et sur le fait que l'écart-type  $\sigma_{\log}$  varie relativement peu lorsque la densité réelle reste voisine de la valeur observée.

\* Si les fluctuations de  $\sigma_{\log}$  dans l'intervalle  $I_2$  sont négligées, les logarithmes décimaux  $Li'$  et  $Ls'$  des bornes  $Li$  et  $Ls$  d'un intervalle de confiance au seuil  $1-\alpha$ , ( $\alpha = \alpha_1 + \alpha_2$ ), de la densité réelle s'obtiennent comme suit :

$$Li' = \log \text{N.P.P.} - \sigma_{\log} \cdot u_{1-\alpha_1}$$

$$Ls' = \log \text{N.P.P.} + \sigma_{\log} \cdot u_{1-\alpha_2}$$

où  $\sigma_{\log}$  qui sera désormais noté  $\sigma$  afin de simplifier l'écriture est donné par la formule (4) alors que  $u_{1-\alpha_1}$  et  $u_{1-\alpha_2}$  représentent la variable normale centrée réduite respectivement au seuil  $1-\alpha_1$  et  $1-\alpha_2$

et par suite 
$$Li = \frac{\text{N.P.P.}}{10^{\sigma \cdot u_{1-\alpha_1}}} \quad \text{et} \quad Ls = \text{N.P.P.} \times 10^{\sigma \cdot u_{1-\alpha_2}}$$

En posant  $\lambda = 10^{\sigma \cdot u_{1-\alpha/2}}$  et dans le cas particulier où  $\alpha_1 = \alpha_2 = \alpha/2$  les intervalles de confiance sont par conséquent de la forme :

$$\left] \frac{\text{N.P.P.}}{\lambda} ; \text{N.P.P.} \cdot \lambda \right[$$

Des tables donnent  $\lambda$  en fonction du rapport de dilution et du nombre  $n$  (COCHRAN, 1950).

\* S'il est tenu compte des légères fluctuations de  $\sigma$  dans l'intervalle  $I_2$ , il convient de remplacer  $\sigma$  dans les relations ci-dessus par la valeur calculée à partir de la formule (3).

Dans ce cas il n'est nullement nécessaire que les effectifs  $n_i$  soient rigoureusement égaux ce qui est commode par exemple lorsque par inadvertance certaines cupules n'ont pas étéensemencées.

Ainsi pour  $k = 3$  ;  $q_1 = 1$  ;  $q_2 = 0,1$  ;  $q_3 = 0,01$  ;  $n_1 = 96$  ;  $n_2 = 95$  ;  $n_3 = 84$  ;  $p_1 = 96$  ;  $p_2 = 38$  ;  $p_3 = 4$ , le N.P.P. est égal à 5,38, d'où  $\log \text{N.P.P.} = 0,731$  et

$$\sigma_{\log} \approx \frac{1}{5,38 \log 10} \left[ \frac{96 e^{-5,38}}{(1-e^{-5,38})^2} + \frac{0,38 e^{-0,538}}{(1-e^{-0,538})^2} + \frac{0,0004 e^{-0,0538}}{(1-e^{-0,0538})^2} \right] \frac{1}{2} = 0,059$$

Le risque étant équitablement réparti aux extrémités d'un intervalle de confiance au seuil 0,95, il vient :

$$Li' = 0,731 - 1,96 \times 0,059 = 0,615, \text{ d'où } Li = 4,12$$

$$Ls' = 0,731 + 1,96 \times 0,059 = 0,847, \text{ d'où } Ls = 7,03$$

Les valeurs  $Li$  et  $Ls$  ainsi calculées ne sont que des approximations car les hypothèses dont elles découlent sont plus ou moins en accord avec la réalité. La similitude avec les valeurs exactes émanant du calcul intégral est assez bonne à condition que le pourcentage de cupules positives ne soit ni trop faible ni trop élevé.

Avant d'aborder l'analyse de variance, qui est de loin la principale technique statistique utilisée dans ce travail, il faut noter que l'assimilation de la loi de  $\log X$  à la loi normale permet également la mise en oeuvre des tests de comparaison de deux densités (MAUL, 1981).

#### I - 2-4-2 Analyse de variance.

Sans entrer dans les développements théoriques de l'analyse de variance qui a été décrite dans de nombreux ouvrages (C.E.A., 1978; HICKS, 1973; LEGENDRE et LEGENDRE, 1979 ; OSTLE et MENSING, 1975 ; POOLE, 1974 ; SNEDECOR et COCHRAN, 1957) il convient néanmoins d'en rappeler succinctement les principes généraux.

Le support théorique de la méthode s'appuie sur quatre hypothèses fondamentales :

- . l'additivité des effets des différents traitements,
- . l'indépendance des erreurs,
- . la normalité des données,
- . l'homogénéité des variances associées aux différents traitements.

L'additivité suppose que les variables ne réagissent pas les unes avec les autres et de ce fait les résultats doivent être interprétés avec prudence en cas d'interactions significatives. Notons également que, dans le cas particulier où les effets sont multiplicatifs, une transformation logarithmique des résultats suffira à les rendre additifs.

L'indépendance des erreurs implique, quant à elle, un échantillonnage réalisé de façon aléatoire.

Si les deux premières hypothèses ne font pas, en général, l'objet de vérifications systématiques, il n'en est pas de même pour l'hypothèse de normalité ainsi que pour l'homogénéité des variances car, en toute rigueur, leur réalisation est indispensable pour que les conclusions aient un sens. En particulier l'hétéroscédasticité, qui est l'absence d'homogénéité des variances, s'oppose, lorsqu'elle est trop élevée, à toute analyse sérieuse de la variance. Les tests utilisés supportent cependant quelques écarts entre la réalité et la théorie car l'analyse de variance est relativement robuste.

Une simple transformation des données brutes suffit quelquefois à satisfaire simultanément les deux hypothèses de normalité et d'homogénéité des variances et même quelquefois l'hypothèse d'additivité comme cela vient d'être indiqué ci-dessus.

La nature de cette transformation dépend de la relation existant entre l'écart-type et la moyenne, cette relation étant fonction de la distribution spatiale des organismes à dénombrer et du principe même de la méthode de numération (GREEN, 1979).

Ainsi la transformation racine carrée est adaptée à la distribution de POISSON et devrait, par conséquent et en toute rigueur, être la seule employée pour l'analyse de variance des résultats obtenus par la numération des colonies sur milieu gélosé. Cependant les difficultés d'interprétation des résultats ainsi transformés incitent les microbiologistes à lui préférer la transformation logarithmique misant ainsi sur la robustesse de l'analyse de variance. Le pragmatisme

de cette pratique n'en est que plus renforcé par les tests d'homogénéité des variances lesquels dénotent le plus souvent une amélioration après transformation logarithmique (ASHBY et RHODES-ROBERTS, 1976 ; KAPER et coll., 1978 ; KOSINSKI et coll., 1979 ; PALMER et coll., 1976).

Pour ce qui nous concerne, et d'après les résultats émanant de ce chapitre, la transformation logarithmique du N.P.P. des données brutes réalise, du moins dans une certaine plage de densité, les conditions de validité de l'analyse de variance. Mais avant toute analyse, l'hypothèse d'égalité des variances sera malgré tout vérifiée au moyen du test de BARTLETT ou du test de COCHRAN (C.E.A., 1978) selon le protocole expérimental appliqué.

L'analyse de variance proprement dite consiste à faire une partition de la somme des carrés des écarts des observations à leur moyenne générale, en plusieurs sommes de carrés de combinaisons linéaires des observations dont chacune relate la variabilité inhérente à l'un des facteurs contrôlés étudiés ou d'une combinaison de ceux-ci (interactions).

Cette partition est réalisée de manière à rendre les diverses composantes indépendantes les unes des autres.

Une dernière somme de carrés représente l'influence de l'ensemble des facteurs non contrôlés dont il serait difficile d'établir une liste exhaustive (échantillonnage, incertitude liée à la méthode de titrage, erreurs de manipulation, etc...) Le quotient de cette somme par son nombre de degrés de liberté est une estimation de la variance résiduelle caractéristique des facteurs non contrôlés.

De même le quotient de chaque somme de carrés par son nombre de degrés de liberté donne un carré moyen qui, moyennant l'hypothèse selon laquelle le facteur contrôlé impliqué dans cette somme n'a aucune influence sur les résultats, réalise également une estimation de la variance résiduelle. Les rapports de ces carrés moyens exécutés dans un ordre préétabli et imposé par le tableau d'analyse de variance permettent, au moyen du test de SNEDECOR, d'apprécier la signification des différents facteurs contrôlés ou de leurs interactions.

Insistons sur le fait que lorsqu'un test relatif à un facteur contrôlé est non significatif il faut se garder de conclure hâtivement à l'absence de son influence, laquelle n'aura peut-être pas été mise en évidence tout simplement par le manque de puissance du test effectué.

Plus généralement, il est toujours possible de faire ressortir l'influence d'un facteur, en adaptant le nombre des essais à l'ordre de grandeur de l'impact dudit facteur sur les résultats.

Cette remarque montre que l'analyse de variance ne doit pas se limiter à la seule pratique des tests de signification; le calcul des composantes de variance permettant également d'apprécier la contribution de chaque facteur étudié à la variation totale. Ainsi l'analyse de variance apparaît comme un outil très efficace autorisant l'étude simultanée de plusieurs sources potentielles de variation en faisant ressortir leur importance relative et ceci avec un minimum d'observations.

#### I - 2-5 Conclusion

Les tests paramétriques, et notamment l'analyse de variance, exigent que les données soumises à l'analyse statistique obéissent à une loi de distribution connue. Il est également souhaitable que la variance soit connue et indépendante de la moyenne des résultats expérimentaux.

L'étude théorique de l'estimation par le N.P.P. menée tout au long de ce premier chapitre a montré que la transformation logarithmique des données brutes avait pour effet de satisfaire ces exigences, tout au moins lorsque la densité bactérienne reste comprise entre les bornes de l'intersection des intervalles  $I_1$  et  $I_2$  définis au paragraphe I-2-3.

Ces conditions sont d'autant mieux respectées que  $n$  devient plus grand. En particulier lorsque  $n = 96$  la log-normalité et la quasi-indépendance de la variance avec la densité microbienne peuvent être admises dans un intervalle de densité relativement large ce qui confère à la méthode une grande souplesse d'utilisation.

Par ailleurs, la précision de la méthode exprimée par le coefficient de variation (environ 13%) est, elle aussi, considérablement améliorée par l'augmentation de  $n$ . Le calcul montre qu'une précision comparable n'est atteinte, par la méthode de numération sur milieu gélosé, qu'à partir d'un nombre de colonies recensées environ égal à 60. En d'autres termes, si moins de 60 colonies sont dénombrées, la méthode des membranes filtrantes est théoriquement moins précise que la méthode du N.P.P. utilisant 96 inoculums par dilution.



Le niveau relativement élevé de la précision et surtout le fait que celle-ci soit, dans une certaine mesure, dépourvue de tout caractère aléatoire ont tous deux largement contribué aux motivations ayant suscité ce travail ainsi qu'à l'élaboration des résultats auxquels il a permis d'aboutir. Nous espérons par là même, contribuer à redonner à la méthode du N.P.P. une grande partie de l'attention des microbiologistes qu'elle avait perdue.

CHAPITRE II : COMPARAISON EXPERIMENTALE DE DEUX TECHNIQUES DE TITRAGE : METHODE ENUMERATIVE SUR MILIEU GELOSE versus METHODE QUANTIQUE EN MILIEU LIQUIDE.

La méthode du nombre le plus probable en milieu liquide utilisant 96 inoculums a rarement été comparée à la méthode énumérative sur milieu gélosé. A cet effet une étude a été conduite sur eau de rivière en mesurant la densité en coliformes thermorésistants (groupe a priori largement représenté par les bactéries Escherichia coli). Le plan expérimental a été conçu dans le but de répondre aux quatre questions suivantes :

. Les résultats expérimentaux sont-ils compatibles avec le modèle mathématique sur lequel s'appuient les développements théoriques exposés dans le premier chapitre ?

. Existe-t-il une interaction entre le facteur méthode et le facteur échantillon ? en d'autres termes, les écarts éventuels entre les résultats de chaque méthode sont-ils indépendants de l'échantillon analysé ?

. Les résultats des deux méthodes peuvent-ils être considérés comme équivalents ? et dans la négative, quelle est l'importance de l'écart ?

. Quelle est la proportion d'Escherichia coli parmi les coliformes thermorésistants ayant engendré des réponses positives en milieu liquide ?

II - 1 MATERIEL ET METHODES

II - 1-1 Echantillons d'eau

Des échantillons de 500 ml d'eau ont été prélevés en surface dans la Moselle à proximité de la centrale thermique de la Maxe au moyen de flacons en verre stérilisés à l'autoclave.

Ces échantillons ont été ramenés au laboratoire à température ambiante pour être soumis à l'analyse au plus 30 minutes après le prélèvement.

II - 1-2 Numération des coliformes thermorésistants sur milieu gélosé.

Ce travail étant pour une large part orienté vers la mise en évidence

de germes indicateurs de contamination fécale, nos efforts se sont concentrés plus particulièrement sur Escherichia coli.

Des dilutions décimales des échantillons d'eau ont été réalisées dans de l'eau physiologique stérile.

Un millilitre de chaque dilution ainsi obtenue a été filtré sur membrane stérile d'une porosité de 0,45  $\mu\text{m}$  (millipore HAWG 047). Pour permettre une meilleure répartition de l'aliquote sur la membrane, celui-ci a été préalablement transféré dans 9 ml d'eau physiologique.

Chaque membrane a ensuite été déposée sur milieu gélosé lactosé (I.P.64041) contenant du chlorure de triphényltétrazolium (I.P.62635) et du tergitol 7 (I.P. 62655).

Après incubation à 44° C pendant 24 heures, les colonies de coloration orangée et auréolées d'un halo jaune ont été dénombrées. Enfin les résultats ont été exprimés en nombre de coliformes thermorésistants par ml de l'échantillon originel.

Les caractéristiques macroscopiques des colonies traduisant une réduction du triphényltétrazolium et une fermentation du lactose ainsi que les conditions de culture nous ont amené à considérer ces colonies comme étant représentatives du genre Escherichieae. Cependant les caractères biochimiques des colonies n'ont pas été confirmés lors d'une subculture.

## II - 1-3 Numération des coliformes thermorésistants en milieu liquide.

Escherichia coli a été inoculé en bouillon lactosé additionné de pourpre de bromocrésol (I.P. 64041) de couleur violette. A cet effet ce milieu stérile a été préalablement réparti sur des plaques de microtitration Microtest II (Poly labo BLOCK FALCON 19013 F 3041) comportant 96 cupules (0,2ml par cupule) à l'aide d'une seringue de Cornwall. Auparavant, les plaques propres avaient été stérilisées par exposition sous rayons ultraviolets pendant 18 heures.

Chaque dilution des échantillons réalisée comme précédemment (II - 1-2) a ensuite été inoculée dans les 96 cupules des plaques Microtest II à raison de 0,05 ml par cupule à l'aide de pipettes calibrées (Poly labo 150 38).

Puis les plaques ont été fermées au moyen d'un couvercle ou d'un film autocollant.

Après incubation à 44°C pendant 24 heures, les cupules présentant une couleur jaune, due à la fermentation du lactose, ont été dénombrées.

Le nombre le plus probable de bactéries contenues dans un inoculum de la dilution la plus faible a été calculé à partir d'un triplet de résultats émanant de trois niveaux de dilution successifs. Les calculs ont été effectués par une calculatrice programmée HEWLETT-PACKARD HP 41 C (le listage du programme utilisé est donné en ANNEXE II). Finalement le titre bactérien de l'échantillon initial a été exprimé en nombre de coliformes thermorésistants par ml.

Lorsque dans un triplet de résultats, par exemple 95 - 94 - 30, une cupule a été manifestement oubliée lors de l'ensemencement à la dilution la plus concentrée, il nous est paru légitime de rectifier les chiffres bruts pour estimer la concentration bactérienne à partir du triplet 96 - 94 - 30.

Une telle omission peut en effet altérer notablement le résultat surtout lorsque la cupule oubliée avait une forte probabilité d'être positive.

Au cours d'un certain nombre d'essais un repiquage systématique de cupules positives (couleur jaune) a été pratiqué afin de confirmer ou non la présence des bactéries Escherichia coli. A cet effet une öse bouclée du contenu d'une cupule positive de bouillon lactosé a été ensemencée sur eau peptonée sans indole, par ailleurs une öse bouclée du contenu de cette même cupule a été portée dans un tube de bouillon lactosé bilié au vert brillant (test d'EIJKMAN modifié MACKENZIE).

Les deux tubes ainsi inoculés ont ensuite été placés à 44°C pendant 24 heures. La mise en évidence de l'indole sur eau peptonée par le réactif de KOVACS et une croissance bactérienne dans le tube contenant le bouillon lactosé au vert brillant ont été retenues pour caractériser Escherichia coli.

#### II - 1-4 Analyse statistique des quantifications bactériennes.

L'exploitation statistique des résultats obtenus par les titrages bactériens a été réalisée par une analyse de variance après vérification de l'homogénéité des variances par le test de BARTLETT.

Le principe de ce test consiste à calculer dans un premier temps la quantité  $\frac{M}{C}$ , où  $M = f \cdot \log 10 \cdot \left( a \log \bar{s}^2 - \sum_h \log s_h^2 \right)$

avec  $\bar{s}^2 = \frac{\sum_h s_h^2}{a}$  a étant le nombre total des estimations indépendantes  $s_h^2$

de la variance commune et f le nombre de degrés de liberté de chacune d'elles ;

et  $C = 1 + \frac{a + 1}{3af}$ .

La fonction discriminante  $\frac{M}{C}$  est ensuite comparée à la distribution du khi-deux avec (a-1) degrés de liberté.

L'absence d'influence d'un facteur contrôlé est vérifiée par le test de SNEDECOR. A cet effet le quotient du carré moyen associé au facteur étudié par un autre carré moyen imposé par le tableau d'analyse de variance est comparé à la limite supérieure du domaine d'acceptation donné par une table de la loi de SNEDECOR pour un risque de première espèce préalablement fixé.

## II - 2 PROTOCOLE EXPERIMENTAL

Des échantillons de 500 ml d'eau de rivière numérotés de 1 à 8 ont été prélevés à des jours différents sur une période totale d'environ un mois.

Le protocole expérimental de chaque manipulation est représenté par la figure 4.

Après homogénéisation par 10 retournements du flacon de 500 ml, 6 ml de son contenu ont été introduits dans un erlenmeyer contenant 54 ml d'eau physiologique stérile. Cette suspension bactérienne à la dilution  $10^{-1}$  a permis de réaliser deux autres dilutions à  $10^{-2}$  et  $10^{-3}$  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 dilutions, six titrages ont été réalisés en parallèle pour chacune des deux méthodes : méthode du nombre le plus probable en milieu liquide et numération directe sur milieu gélosé pour laquelle seules les dilutions au 1/10 et au 1/100 ont été utilisées.

Tous les ensemencements relatifs à un niveau de dilution ont été effectués avant d'entreprendre la préparation de la dilution suivante, ceci afin de minimiser l'influence d'une éventuelle dérive dans le temps de la concentration bactérienne. La période s'écoulant entre le moment du prélèvement de l'échantillon dans le milieu naturel et la fin du titrage n'a jamais excédé 2h30 minutes.

Au cours des manipulations n° 3, 5, 6 et 8, environ une centaine de cupules positives ont fait l'objet d'un test confirmatif dans le but de vérifier la présence d'Escherichia coli.

L'hypothèse de l'égalité des proportions de résultats confirmés pour chacune de ces quatre expériences a été soumise au test du khi-deux.

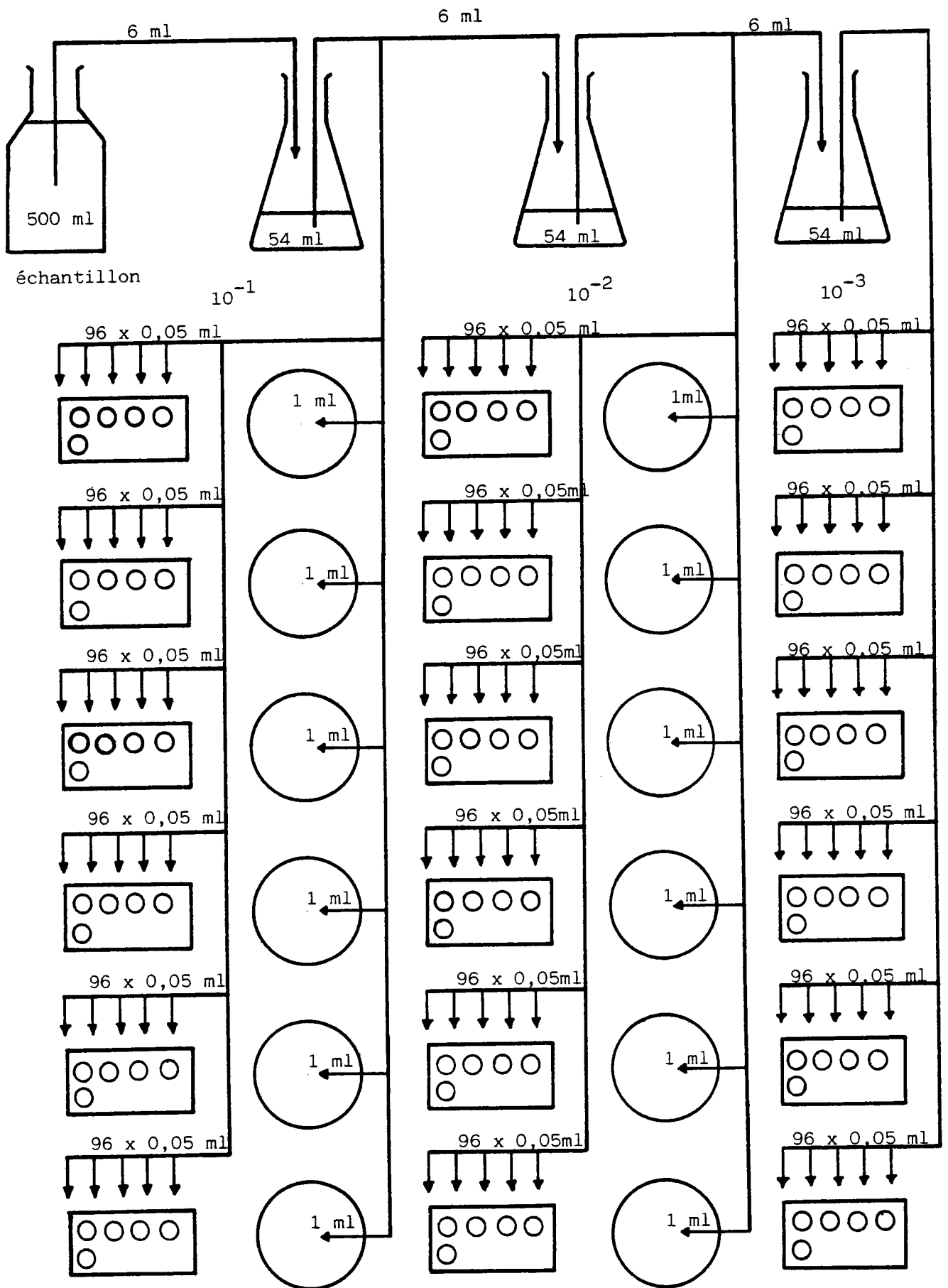


FIGURE 4 : Schéma du protocole expérimental pour la comparaison des titres bactériens obtenus par deux méthodes de titrage (numération en milieu liquide versus comptage des colonies sur milieu gélosé).

Après lecture et calcul des résultats comme décrit précédemment, six valeurs de la densité bactérienne ont été obtenues pour chaque méthode et pour chaque échantillon initial. Ces valeurs ont fait l'objet d'une transformation logarithmique avant d'être soumises à l'analyse de variance conformément au tableau n° 4.

TABLEAU 4 : Tableau d'analyse de variance.

Source de variation	Somme de carrés Q	Degrés de liberté $\nu$	Quantité estimée par $Q/\nu$
Facteur A échantillon	$Q_A = 12 \sum_i (\bar{X}_{i.} - \bar{X})^2$	7	$\sigma_R^2 + 12 \sigma_A^2$
Facteur B méthode	$Q_B = 48 \sum_j (\bar{X}_{.j} - \bar{X})^2$	1	$\sigma_R^2 + 6 \sigma_{AB}^2 + 48 \sigma_B^2$
Interaction AB	$Q_{AB} = 6 \sum_i \sum_j (\bar{X}_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X})^2$	7	$\sigma_R^2 + 6 \sigma_{AB}^2$
Résiduelle	$Q_R = \sum_i \sum_j \sum_{\alpha} (X_{ij\alpha} - \bar{X}_{ij})^2$	80	$\sigma_R^2$
Totale	$Q_T = \sum_i \sum_j \sum_{\alpha} (X_{ij\alpha} - \bar{X})^2$	95	$\sigma_T^2$

Le modèle adopté pour cette expérience est un plan factoriel à deux facteurs contrôlés avec répétitions.

Le facteur échantillon, noté A, intervient avec 8 modalités aléatoires tandis que le facteur méthode, noté B, possède deux modalités fixes. L'interaction de ces deux facteurs est notée A B.

Un résultat expérimental est noté  $X_{ij\alpha}$  lorsqu'il correspond à la modalité i du facteur contrôlé A et à la modalité j du facteur contrôlé B alors que  $\alpha$  est l'indice de l'essai considéré (dans cette expérience il y a 6 essais pour chaque combinaison i j).

La moyenne arithmétique de l'ensemble de tous les résultats expérimentaux  $X_{ij\alpha}$  est notée  $\bar{X}$ .

La moyenne arithmétique des résultats de même indice i est notée  $\bar{X}_{i.}$  et de même  $\bar{X}_{.j}$  désigne la moyenne des résultats d'indice j ; enfin la moyenne des résultats de mêmes indices i et j est notée  $\bar{X}_{ij}$ .

Par ailleurs,  $\sigma_T^2$  exprime la variance totale et  $\sigma_A^2$ ,  $\sigma_B^2$ ,  $\sigma_{AB}^2$  et  $\sigma_R^2$  désignent les composantes de variance représentatives respectivement des fluctuations des facteurs A, B, AB et de l'ensemble des facteurs non contrôlés (variance résiduelle).

## II - 3 RESULTATS

Le tableau 5 donne, pour chaque échantillon d'eau de Moselle, les résultats des titrages obtenus par les deux techniques : méthode du nombre le plus probable (N.P.P.) en milieu liquide avec 96 inoculums par dilution et méthode des membranes filtrantes (M.F.).

L'estimation de la concentration bactérienne exprimée en nombre de germes par ml d'échantillon ainsi que la moyenne géométrique des valeurs de chaque ensemble de 6 titrages réalisés dans les mêmes conditions figurent également dans ce tableau.

Les résultats des deux premières expériences proviennent de dilutions décalées d'une puissance de 10 par rapport au schéma adopté pour les autres manipulations : l'ensemencement des plaques du premier niveau de dilution ayant été effectué à partir d'eau non diluée.

Enfin les nombres soulignés correspondent aux plaques dont les cupules ont fait l'objet d'un repiquage conformément au test d'EIJKMAN.



Echantillon	Méthode	Résultat.		Concentration bactérienne par ml ( u )	log u	Moyenne géo- métrique ( / ml )	
		Triplet	N.P.P.				
1	N. P. P.	96 - 93 - 16	27,95	559	2,747	591	
		96 - 90 - 22	27,15	543	2,734		
		96 - 91 - 37	36,95	739	2,869		
		96 - 90 - 24	28,05	561	2,749		
		96 - 88 - 29	27,95	559	2,747		
		96 - 90 - 29	30,40	608	2,784		
	Membranes filtrantes	51	510	2,708	466		
		41	410	2,613			
		52	520	2,716			
		54	540	2,732			
		35	350	2,544			
		50	500	2,699			
	2	N. P. P.	96 - 94 - 30	38,10	762	2,882	876
			96 - 96 - 35	49,70	994	2,997	
96 - 94 - 35			42,10	842	2,925		
96 - 96 - 33			47,35	947	2,976		
96 - 94 - 34			41,25	825	2,916		
96 - 96 - 35			45,45	909	2,959		
Membranes Filtrantes		73	730	2,863	808		
		84	840	2,924			
		84	840	2,924			
		83	830	2,919			
		85	850	2,929			
		77	770	2,886			
3		N. P. P.	66 - 11 - 0	1,15	230	2,362	218
			61 - 9 - 0	0,99	198	2,297	
	67 - 12 - 0		1,20	240	2,380		
	64 - 6 - 0		1,02	204	2,310		
	62 - 11 - 0		1,05	210	2,322		
	64 - 13 - 0		1,13	226	2,354		
	Membranes Filtrantes	20	200	2,301	262		
		26	260	2,415			
		28	280	2,447			
		28	280	2,447			
		25	250	2,398			
		32	320	2,505			

**TABLEAU 5** : Titres bactériens obtenus par deux techniques de titrage :

N.P.P. avec 96 inoculums par dilution *versus* filtration sur membrane.

Echantillon	Méthode	Résultat		Concentration bactérienne par ml ( u )	log u	Moyenne géo- métrique ( / ml )
		Triplet	N.P.P.			
4	N. P. P.	89 - 14 - 3	2,34	468	2,670	429
		86 - 15 - 1	2,09	418	2,621	
		82 - 21 - 1	2,02	404	2,606	
		85 - 22 - 5	2,35	470	2,672	
		90 - 20 - 2	2,62	524	2,719	
		71 - 25 - 1	1,60	320	2,505	
		Membranes	33	330	2,519	
	filtrantes	38	380	2,580		
		37	370	2,568		
		44	440	2,643		
		42	420	2,623		
		49	490	2,690		
	5	N. P. P.	96 - 45 - 1	5,88	1176	3,070
92 - 41 - 3			4,08	816	2,912	
96 - 28 - 5			4,29	858	2,933	
93 - 39 - 2			4,12	824	2,916	
95 - 37 - 6			4,88	976	2,989	
95 - 36 - 5			4,70	940	2,973	
Membranes			76	760	2,881	682
Filtrantes		72	720	2,857		
		66	660	2,819		
		66	660	2,819		
		66	660	2,819		
		64	640	2,806		
6		N. P. P.	95 - 25 - 5	3,77	754	2,877
	92 - 24 - 3		3,07	614	2,788	
	93 - 28 - 2		3,39	678	2,831	
	92 - 28 - 1		3,18	636	2,803	
	91 - 29 - 2		3,14	628	2,798	
	84 - 22 - 4		2,25	450	2,653	
	Membranes		55	550	2,740	570
	Filtrantes	62	620	2,792		
		59	590	2,771		
		58	580	2,763		
		60	600	2,778		
		49	490	2,690		

TABLEAU 5 (suite)

TABLEAU 5 (suite et fin)

Echantillon	Méthode	Résultat		Concentration bactérienne par ml ( u )	log u	Moyenne géo- métrique ( / ml )	
		Triplet	N.P.P.				
7	N. P. P.	96 - 47 - 7	6,91	1382	3,141	1314	
		96 - 45 - 11	7,08	1416	3,151		
		96 - 41 - 3	5,55	1110	3,045		
		96 - 44 - 6	6,31	1262	3,101		
		96 - 43 - 7	6,27	1254	3,098		
		96 - 51 - 6	7,49	1498	3,176		
	Membranes filtrantes	106	1060	3,025	1089		
		123	1230	3,090			
		117	1170	3,068			
		92	920	2,964			
		112	1120	3,049			
	8	N. P. P.	96 - 47 - 8	7,04	1408	3,149	1491
			96 - 45 - 9	6,83	1366	3,135	
96 - 53 - 4			7,59	1518	3,181		
96 - 50 - 4			7,05	1410	3,149		
96 - 52 - 9			8,09	1618	3,209		
96 - 57 - 3			8,24	1648	3,217		
Membranes Filtrantes		114	1140	3,057	1159		
		122	1220	3,086			
		119	1190	3,076			
		110	1100	3,041			
		112	1120	3,049			
		119	1190	3,076			

II - 3-1 Conformité des résultats expérimentaux avec le modèle théorique.

La validité de l'hypothèse selon laquelle les microorganismes ou plus exactement les unités formatrices de colonies sont distribués au hasard au sein de la suspension bactérienne titrée (modèle de POISSON) peut être appréciée par le coefficient de dispersion de FISHER.

Ce coefficient calculé pour chaque groupe de 6 résultats obtenus par la méthode des membranes filtrantes sur un même échantillon d'eau est exprimé par la quantité  $D^2 = \frac{(n-1) s^2}{m}$  où  $m$  représente la moyenne et  $s^2$  la variance empirique des  $n = 6$  résultats.

Le principe du test mis en oeuvre repose sur l'égalité de la moyenne et de la variance dans le cas d'une distribution de POISSON;  $D^2$  suit alors approximativement une loi du khi-deux à  $n-1$  degrés de liberté.

Le tableau 6 donne les valeurs de  $m$ ,  $(n-1) s^2$  et  $D^2$  pour chacun des 8 essais.

**TABLEAU 6** : Coefficient de dispersion de FISHER des numérations sur milieu gélosé calculé pour chaque échantillon d'eau.

Echantillon	1	2	3	4	5	6	7	8
$m$	47,2	81,0	26,5	40,5	68,3	57,2	109,3	116,0
$(n-1) s^2$	278,8	118,0	79,5	161,5	107,3	106,8	575,3	110,0
$D^2$	5,91	1,46	3,00	3,99	1,57	1,87	5,26	0,95

Aucune des valeurs  $D^2$  trouvées n'est significative au niveau de probabilité de 95 % ( $\chi^2_{0,95} = 11,1$  avec 5 degrés de liberté).

Etant donné l'indépendance des variables  $D_i^2$  pour chaque échantillon, il est également possible d'intégrer l'ensemble de tous les résultats en vue d'effectuer un test unique. En effet,  $\sum_i D_i^2$  suit une loi du khi-deux avec 40 degrés de liberté.

La valeur 24,01 trouvée pour  $\sum_i D_i^2$  étant nettement inférieure à la limite supérieure du domaine d'acceptation au seuil de probabilité de 95% donné par la table du khi-deux pour 40 degrés de liberté ( $\chi^2_{0,95} = 55,8$ ), le modèle de POISSON semble par conséquent être respecté.

Un autre critère de compatibilité consiste à comparer, pour la méthode du N.P.P. en milieu liquide, l'écart-type empirique  $s$  à l'écart-type théorique  $\sigma_{\log}$  calculé au moyen de la formule 4.

Rappelons que  $\sigma_{\log} = 0,0561$  lorsque 96 cupules sont inoculées à chaque niveau de dilution, ce qui correspond à un coefficient de variation de 13 %. L'écart-type des résultats expérimentaux est donné par la valeur  $s = 0,0542$  à laquelle est associé le coefficient de variation 12,5 %. Le calcul des limites inférieure et supérieure d'un intervalle de confiance à 95 % conduit aux valeurs 0,0445 et 0,0694 dont les coefficients de variation respectifs sont : 10,3 % et 16,1 %.

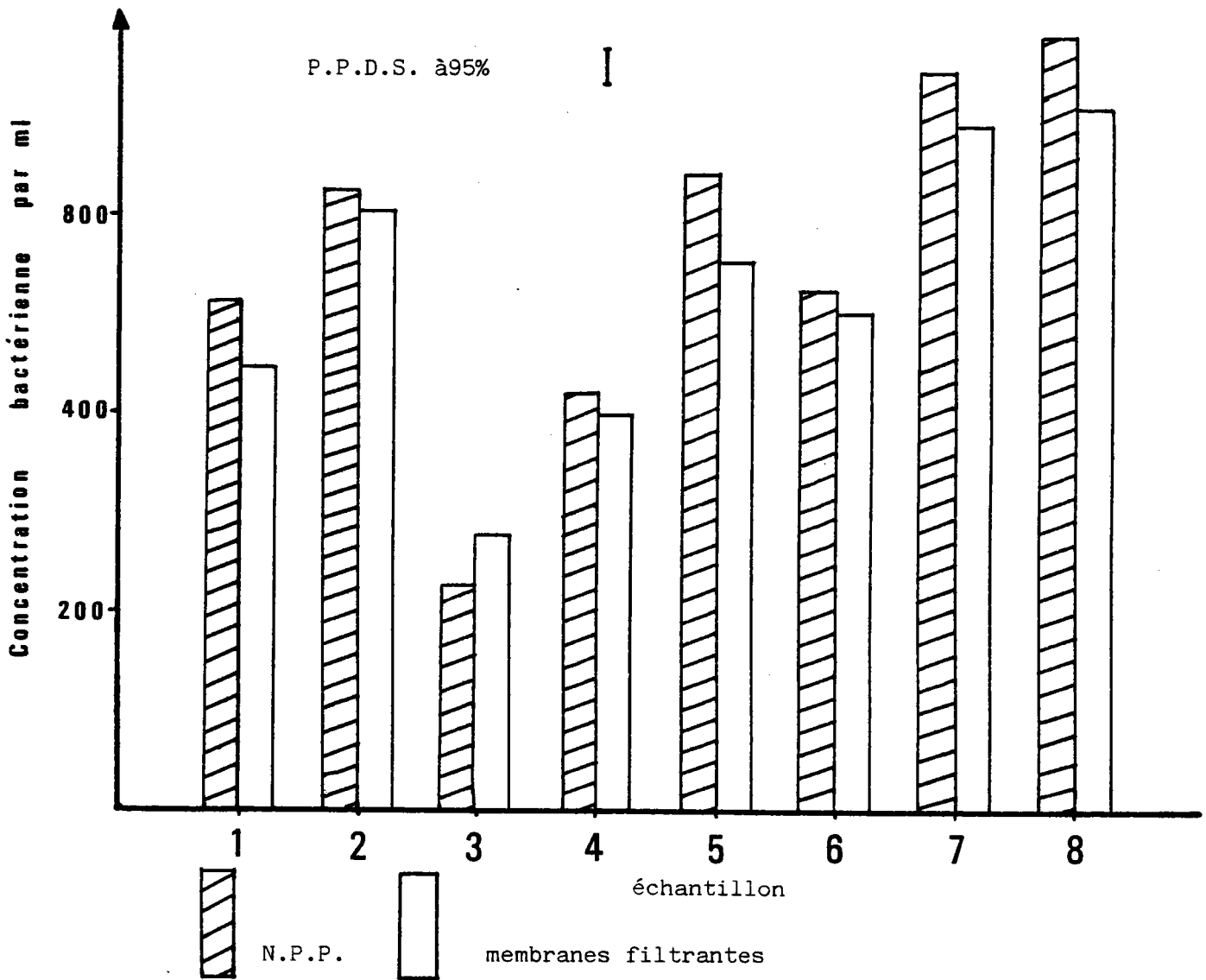
La comparaison de ces chiffres montre une très bonne similitude entre la précision théorique et celle qui est constatée expérimentalement.

Notons enfin que l'écart-type observé pour la méthode des membranes filtrantes prend la valeur 0,0487 ce qui correspond à un coefficient de variation de 11,2 %. Bien que sans véritable signification car la méthode ne possède pas de variance propre, ce dernier résultat permet cependant une comparaison aisée des précisions obtenues par les deux techniques de titrage.

## II - 3-2 Comparaison des titres bactériens obtenus par les deux méthodes de titrage.

Les données du tableau 5 sont illustrées par la figure 5 qui tente au moyen d'une représentation graphique de concrétiser les différences des résultats obtenus par les deux techniques de titrage.

**FIGURE 5:** Comparaison des moyennes géométriques des titres bactériens obtenus par deux méthodes de titrages.



L'examen de la figure 5 fait apparaître des résultats obtenus en milieu liquide plus élevés dans 7 cas sur 8, comparativement à ceux obtenus sur membranes.

Le calcul de la plus petite différence significative (P.P.D.S.) au seuil 0,95 montre que cette différence est significative pour les échantillons n° 1, 3, 5, 7 et 8.

P.P.D.S. =  $t \sqrt{\frac{2s^2}{n}}$  où n est le nombre de répétitions (ici n = 6),  $s^2$  est l'estimation de la variance résiduelle et t n'est autre que la variable de STUDENT au seuil 97,5% avec 80 degrés de liberté.

Bien que très commode ce procédé n'autorise qu'une comparaison des résultats numériques pris deux à deux. L'analyse de variance permet par contre une exploitation globale des résultats ce qui a pour effet d'augmenter la portée ainsi que la richesse des conclusions.

a) analyse de variance

Après vérification de l'homogénéité des variances au moyen du test de BARTLETT, le calcul des carrés moyens permet d'exécuter les tests de signification relatifs aux divers facteurs contrôlés.

Les estimations, notées  $s_h^2$ , des variances de chaque combinaison i j des facteurs contrôlés sont consignées dans le tableau 7.

TABLEAU 7 : Variance empirique de chaque classe de résultats.

Numéro de l'échantillon		1	2	3	4	5	6	7	8
Variance empirique $s_h^2$	N.P.P	0,00252	0,00181	0,00107	0,00551	0,00360	0,00565	0,00217	0,00117
	M.F.	0,00548	0,00071	0,00468	0,00371	0,00083	0,00133	0,00191	0,00031

Le calcul de la fonction discriminante  $\frac{M}{C}$  à partir de ces données conduit à la valeur  $\frac{M}{C} \approx 21,44$  qui est inférieure à la limite donnée par la loi du khi-deux au niveau de probabilité de 95% avec 15 degrés de liberté ( $\chi_{0,95}^2 \approx 25$ )

Le test n'étant pas significatif, l'égalité des variances a été admise.

Les carrés moyens ont été calculés à partir des quantités présentées dans le tableau 8.

Echantillon	Méthode	Répétitions						$\sum_{\alpha} x_{ij\alpha}$	$\sum_j \sum_{\alpha} x_{ij\alpha}$	$\sum_{lij} \sum_{\alpha} x_{ij\alpha}$
		1	2	3	4	5	6			
1	N.P.P.	2,747	2,734	2,869	2,749	2,747	2,784	16,630	32,642	
	M.F.	2,708	2,613	2,716	2,732	2,544	2,699	16,012		
2	N.P.P.	2,882	2,997	2,925	2,976	2,916	2,959	17,655	35,100	
	M.F.	2,863	2,924	2,924	2,919	2,929	2,886	17,445		
3	N.P.P.	2,362	2,297	2,380	2,310	2,322	2,354	14,025	28,538	
	M.F.	2,301	2,415	2,447	2,447	2,398	2,505	14,513		
4	N.P.P.	2,670	2,621	2,606	2,672	2,719	2,505	15,793	31,416	270,132
	M.F.	2,519	2,580	2,568	2,643	2,623	2,690	15,623		
5	N.P.P.	3,070	2,912	2,933	2,916	2,989	2,973	17,793	34,794	
	M.F.	2,881	2,857	2,819	2,819	2,819	2,806	17,001		
6	N.P.P.	2,877	2,788	2,831	2,803	2,798	2,653	16,750	33,284	
	M.F.	2,740	2,792	2,771	2,763	2,778	2,690	16,534		
7	N.P.P.	3,141	3,151	3,045	3,101	3,098	3,176	18,712	36,933	
	M.F.	3,025	3,090	3,068	2,964	3,049	3,025	18,221		
8	N.P.P.	3,149	3,135	3,181	3,149	3,209	3,217	19,040	37,425	
	M.F.	3,057	3,086	3,076	3,041	3,049	3,076	18,385		

**TABLEAU 8** : Résultats des titrages après transformation logarithmique et sommes nécessaires au calcul des carrés moyens.



Le tableau 9 récapitule les valeurs des carrés moyens ainsi que les quantités estimées par ceux-ci conformément au modèle d'analyse de variance adopté.

Source de Variance	Somme de Carrés Q	Degrés de liberté $\nu$	Carré moyen $q = Q/\nu$	Quantité estimée par q
Facteur A échantillon	5,05039	7	0,72148	$\sigma_R^2 + 12 \sigma_A^2$
Facteur B méthode	0,07393	1	0,07393	$\sigma_R^2 + 6 \sigma_{AB}^2 + 48 \sigma_B^2$
Interaction A B	0,09583	7	0,01369	$\sigma_R^2 + 6 \sigma_{AB}^2$
Résiduelle	0,21232	80	0,00265	$\sigma_R^2$
TOTALE	5,43247	95		

TABLEAU 9 : Tableau d'analyse de variance.

L'absence d'effet dû à un facteur contrôlé se vérifie en appliquant le test de SNEDECOR.

. Ainsi l'influence du facteur AB (interaction) se contrôle en comparant le rapport  $F_{AB} = \frac{q_{AB}}{q_R} = 5,17$  à la limite  $F_{0,95}(7;80) = 2,15$  de la loi de SNEDECOR au seuil de probabilité de 0,95 avec respectivement 7 et 80 degrés de liberté.

Puisque  $F_{AB}$  est supérieur à cette limite (en fait  $F_{AB}$  est même supérieur à  $F_{0,995}(7;80)$ ), le test est très significatif et l'existence d'une interaction entre les facteurs A et B est admise.

. L'influence éventuelle du facteur B (méthode) se contrôle en comparant le rapport  $F_B = \frac{q_B}{q_{AB}} = 5,40$  à la limite  $F_{0,95}(1;7) = 5,59$ .

Bien que légèrement inférieur à cette limite  $F_B$  est néanmoins supérieur à  $F_{0,945}(1;7)$ .

L'influence du facteur B sera admise avec un risque d'erreur inférieur à 6 %.

Le caractère conservateur de ce test peut surprendre mais son manque de puissance doit être attribué au fait que  $q_B$  est comparé non pas à la variance résiduelle mais à la variance d'interaction, ce qui a l'avantage d'augmenter considérablement la portée du résultat : la conclusion n'étant plus seulement limitée aux 8 échantillons étudiés mais il est permis de l'étendre à l'ensemble de tous les prélèvements possibles dont ils sont issus.

. Ne répondant pas à une question préalablement posée, l'étude de l'influence du facteur A (échantillon) n'a ici aucun intérêt mais il est évident que les échantillons n'avaient pas la même concentration bactérienne.

b) Importance de la surestimation des titres bactériens obtenus par la quantification en milieu liquide.

Le test de SNEDECOR effectué sur le quotient  $F_B$  ainsi que l'examen de la figure 5 montrent que les titres bactériens estimés par la méthode du N.P.P. ont tendance à être légèrement plus élevés que les estimations provenant du comptage direct des colonies apparues sur milieu gélosé.

Le logarithme décimal du rapport moyen des estimations données par chaque méthode est égal à la différence  $m_1 - m_2$  où  $m_1$  et  $m_2$  désignent la moyenne arithmétique des logarithmes de toutes les valeurs trouvées respectivement avec la méthode du N.P.P. et la méthode des membranes filtrantes.

Ainsi,  $m_1 = \frac{16,630 + \dots + 19,040}{48} = 2,842$  et  $m_2 = \frac{16,012 + \dots + 18,385}{48} = 2,786$

d'où  $m_1 - m_2 = 0,0555$  ce qui correspond à un rapport moyen de 1,136. La surestimation est donc de 14% environ.

D'autre part, un test unilatéral avec un risque de 5% bloqué à droite

réalisé sur la quantité 
$$Z = \frac{m_1 - m_2}{\sqrt{\frac{2(s^2 + 6s_{AB}^2)}{48}}} = 2,32$$

où  $s^2 + 6s_{AB}^2$  représente la variance d'interaction du tableau 9 estimée avec 7 degrés de liberté, est significatif. En effet, la quantité  $Z$  est supérieure au seuil d'acceptation égal à 1,89 donné par la table de STUDENT avec 7 degrés de liberté.

La puissance accrue de ce test unilatéral met en évidence l'écart  $m_1 - m_2$  d'une manière plus nette que lors du test effectué sur  $F_B$ .

Enfin le calcul de l'écart-type de  $m_1 - m_2$  permet d'établir que la surestimation est de  $14\% \pm 6\%$ .

II - 3-3 Proportion d'Escherichia coli parmi les coliformes thermorésistants.

Les résultats des repiquages réalisés pour un certain nombre de cupules positives des essais n° 3, 5, 6 et 8, en vue de confirmer ou non la présence des bactéries Escherichia coli, sont consignés dans le tableau 10.

Echantillon	n° 3	n° 5	n° 6	n° 8
nombre total de cupules repiquées	92	105	105	100
mise en évidence de <u>Escherichia coli</u>	85	67	82	97
<u>Escherichia coli</u> non décelé	7	38	23	3

Tableau 10: Résultat des repiquages visant à confirmer la présence de Escherichia coli.

La valeur khi-deux calculée à partir des effectifs observés et des effectifs théoriques obtenus avec l'hypothèse d'égalité des proportions de résultats confirmés est égale à 47,28, or la limite supérieure du domaine d'acceptation au seuil de probabilité 0,995 avec 3 degrés de liberté est égale à 12,8.

Le test du khi-deux est très significatif et les quatre proportions de résultats confirmant la présence d'Escherichia coli ne peuvent de ce fait être considérées comme égales.

L'échantillonnage ayant permis le calcul de ces proportions s'appelle échantillonnage à deux degrés. La première étape consiste à choisir un échantillon d'unités primaires : il s'agit des échantillons d'eau ayant été à l'origine de tests confirmatifs. Un sous-échantillon d'unités secondaires a ensuite été prélevé dans chaque unité primaire sélectionnée.

Ce sous-échantillon est formé par les cupules positives repiquées. La proportion  $\bar{p}$  de cupules contenant Escherichia coli peut être estimée par  $\bar{p} = \frac{\sum p_i}{n}$  où  $n$  est le nombre d'unités primaires et les  $p_i$  désignent les proportions de réponses confirmées calculées à partir d'un nombre identique de cupules pour chaque indice  $i$ .

Si de plus les unités primaires ayant fait l'objet de tests confirmatifs sont considérées comme étant issues d'une population infinie, alors la variance de  $\bar{p}$  s'exprime par la relation  $\sigma^2(\bar{p}) = \frac{1}{n(n-1)} \sum_{i=1}^n (p_i - \bar{p})^2$  (COCHRAN, 1977).

Avant d'effectuer les calculs il a été nécessaire d'estimer les proportions  $p_i$  à partir d'un même effectif de cupules pour chaque indice  $i$ .

Le nombre minimal étant 92, les données redondantes des essais n° 5,6 et 8 ont été éliminées par tirage au sort à l'aide d'une table de nombres au hasard.

Finalement, les calculs réalisés avec les rapports  $\frac{85}{92}$ ,  $\frac{59}{92}$ ,  $\frac{71}{92}$  et  $\frac{89}{92}$

donnent  $\bar{p} \approx 0,825$  et  $\sigma^2(\bar{p}) = 0,00555$ .

En conclusion, la proportion de cupules positives avec présence de Escherichia coli est égale à : 82,5% + 7,5 % pour le type d'eau étudié.

## II - 4 DISCUSSION

### \* Résultats expérimentaux et modèle mathématique.

Le test relatif à l'indice de dispersion de FISHER montre que le modèle de POISSON semble respecté pour les unités formatrices de colonies ce qui traduit par conséquent une bonne homogénéité au sein des petits volumes. Rappelons que tous les titrages ont été réalisés à partir d'un prélèvement de 6 ml de l'échantillon initial et que les aspirations-refoulements qui ont précédé chaque prise de liquide ont certainement contribué à assurer cette homogénéité.

Une conclusion analogue est formulée par EL-SHAARAWI et coll. (1981) tandis que d'autres auteurs (BREITTMAYER et GAUTHIER, 1979 ; KAPER et coll., 1978) mentionnent par contre de larges écarts entre les résultats expérimentaux obtenus par filtration sur membrane et le modèle théorique de POISSON.

Par ailleurs, la reproductibilité des résultats obtenus par le N.P.P. avec 96 cupules par dilution exprimée par le coefficient de variation (12,5%) est très voisine de la précision théorique égale à 13%. Ceci va dans le sens de l'affirmation précédente et semble également indiquer que les facteurs non contrôlés (erreurs lors des prélèvements de liquide, oublis éventuels de certaines cupules pendant l'ensemencement, variabilité du volume des gouttes, etc...) ont probablement un impact relativement faible sur les résultats des quantifications. La précision de 13% peut par conséquent être considérée comme un paramètre caractéristi-

que de la méthode proposée ce qui n'est pas le cas pour la méthode énumérative sur milieu gélosé où, il faut le rappeler, la précision présente un caractère aléatoire.

L'ensemble des résultats montre cependant que, sur un grand nombre de titrages, les précisions des deux méthodes sont très comparables.

Ce dernier point constitue un progrès essentiel puisque le manque de précision de la méthode quantique en milieu liquide pratiquée dans les conditions habituelles est le principal argument de ses détracteurs.

\* Comparaison des concentrations bactériennes estimées par les deux techniques de titrage.

Les estimations de la densité bactérienne données par la méthode du nombre le plus probable sont en moyenne plus élevées que les mêmes estimations obtenues par comptage des colonies apparues sur milieu gélosé. La surestimation est voisine de 14% pour l'eau de rivière utilisée dans ce travail.

A notre connaissance, tous les auteurs ayant comparé les deux méthodes (GREEN et LITSKY, 1980 ; Mac CARTHY et coll., 1961 ; VELZ, 1970 ; THOMAS et WOODWARD, 1955) ont constaté le même phénomène avec une ampleur variable selon l'origine et la nature de l'eau.

Hormis le fait que les microorganismes se développent en général plus facilement dans un milieu liquide que sur un support solide, les travaux de GREEN et LITSKY (1980) ont montré que la différence semble s'accroître pour les bactéries provenant d'une eau peu favorable à leur croissance.

La surestimation trouve une autre explication s'appuyant sur l'existence d'une interaction méthode - échantillon. En effet, l'amplitude des écarts observés est variable d'un échantillon à l'autre et cela d'une manière très significative. Cette variabilité de l'effet méthode selon les échantillons nous suggère l'argumentation suivante : les variations dans les fréquences relatives des différentes espèces bactériennes, qui peuvent entre autres être influencées par des facteurs climatiques (lessivage des sols par la pluie) ou les caractères physico-chimiques de l'eau, sont répercutées sur les résultats par la sélectivité préférentielle qu'exercent les méthodes de titrage sur les diverses espèces de microorganismes.

Cette hypothèse est soutenue par les résultats des repiquages qui ont décelé environ 17,5% de faux positifs. Il s'agit de coliformes thermorésistants autres que Escherichia coli, très vraisemblablement de Klebsiella dont la croissance semble peu entravée par la température de 44°C.

Ce manque de spécificité n'est pas propre aux titrages de type quantique car il est généralement admis qu'en colimétrie sur milieu gélosé 5% environ des colonies sont formées de Klebsiella. Il semble d'après ce pourcentage que le phénomène soit plus discret sur milieu solide qu'en milieu liquide dans les conditions de nos essais.

En outre le test du khi-deux effectué sur les données du tableau 10 montre de manière très significative que les proportions de réponses positives non confirmées ne peuvent pas être considérées comme étant indépendantes de l'échantillon. Il n'est pas possible d'établir de façon précise une corrélation entre la proportion de faux positifs et l'importance de la surestimation car le nombre d'expériences est trop faible. Mais il est cependant intéressant de remarquer que la plus forte proportion de non confirmés a été obtenue pour l'échantillon n° 5 donnant également le plus important écart entre les titres bactériens émanant des deux méthodes de numération et que cet échantillon a été prélevé par un jour de forte pluie. Rappelons enfin que cette étude ne permet pas dans sa conception de connaître l'importance relative de chacune des deux causes évoquées pour tenter d'expliquer la différence des résultats numériques.

DEUXIEME PARTIE

MISE EN EVIDENCE DE L'HETEROGENEITE D'UN MILIEU  
HYDRIQUE : CONSEQUENCES SUR L'ESTIMATION  
DES POPULATIONS BACTERIENNES

La portée ainsi que la qualité de toute étude microbiologique reposent sur la méthode d'échantillonnage. La stratégie de prélèvement doit être bâtie en fonction des objectifs de l'étude, des contraintes économiques et de surcroît être adaptée à l'hétérogénéité du milieu hydrique étudié.

La configuration de cette hétérogénéité se manifeste dans l'espace mais aussi dans le temps. Elle ne dépend pas uniquement des caractéristiques physico-chimiques de l'eau mais également de facteurs climatiques, biogéographiques et morphodynamiques (pente, vitesse du courant, profondeur et débit de la rivière).

De nombreux travaux consacrés à l'étude de la variabilité spatiale du plancton et des bactéries en milieu lentique et marin ont montré une tendance très marquée des microorganismes à s'agréger.

De telles associations ont été décrites à l'échelle du micron (il s'agit par exemple de l'adsorption des bactéries par des particules en suspension) (SEKI, 1968; SOROKIN, 1970 ; WILLIAMS, 1970), du centimètre (ASHBY et RHODES-ROBERTS, 1976), du mètre (PALMER et coll., 1976) et même du kilomètre (PLATT et coll., 1970).

L'augmentation de l'incertitude d'une estimation au-delà de la précision propre à la méthode de titrage provoquée par l'hétérogénéité ne va pas sans poser de problèmes lors de l'extrapolation à de grandes masses d'eau de résultats généralement obtenus à partir d'échantillons dont le volume est infiniment plus restreint.

En étudiant l'importance relative des variations spatiales et temporelles du zooplancton dans un lac philippin, LEWIS (1978) insiste sur la nécessité d'étudier simultanément les composantes temps et espace pour la compréhension globale de l'hétérogénéité. Ainsi la relativisation des diverses sources de variation fait non seulement partie intégrante des caractéristiques propres d'un milieu hydrique mais doit également servir de base de référence pour la mise au point d'une stratégie de prélèvement.

L'étude sérieuse du contrôle de la pollution d'un cours d'eau dans des conditions optimales est à ce prix !

Dans les chapitres suivants seront considérées successivement :

- des généralités sur la répartition spatiale des individus d'une population,
- les causes d'erreur dans l'analyse des échantillons,
- l'hétérogénéité spatio-temporelle d'un milieu hydrique,
- les variations journalières des concentrations bactériennes.



## CHAPITRE I : GENERALITES SUR LA REPARTITION SPATIALE DES INDIVIDUS D'UNE POPULATION

### I - 1 DIFFERENTES CONFIGURATIONS POSSIBLES.

La répartition des individus d'une population dans l'espace à deux ou trois dimensions est conditionnée par de nombreux facteurs dont l'importance relative varie selon l'espèce considérée.

Dans le cas particulier des bactéries, le rôle prépondérant de la charge particulaire de l'eau a déjà été évoqué dans la première partie de ce travail pour ce qui concerne la configuration de l'hétérogénéité microscopique de la distribution des microorganismes.

Malgré leur grande diversité, les distributions possibles des individus d'une population peuvent être distinguées en trois modèles principaux illustrés par la figure 6 où sont représentées dans l'ordre : une répartition régulière, une répartition au hasard et enfin une tendance à l'agrégation des individus dans l'espace.

Cette représentation schématique a le mérite de concrétiser et de permettre une meilleure appréhension des problèmes soulevés par l'existence d'une hétérogénéité dans la répartition des êtres vivants dans le milieu naturel. Notons qu'il peut y avoir coexistence et imbrication, à des échelles différentes, des trois modèles de la figure 6. L'extension de ces modèles à l'espace tridimensionnel est aisée, et dans le cas d'un milieu hydrique il est loisible d'identifier les points à des unités formatrices de colonies.

Si l'hétérogénéité microscopique des populations bactériennes se conçoit relativement bien, des représentations analogues avec des taches plus ou moins sombres selon l'importance de la concentration en microorganismes peuvent être imaginées pour l'illustration de l'hétérogénéité macroscopique d'un milieu hydrique.

Comme cela a déjà été pratiqué au chapitre II de la première partie, le calcul du coefficient de dispersion de FISHER à partir des résultats expérimentaux d'un échantillonnage facilite le choix d'un modèle de distribution spatiale des individus de la population étudiée.

Une exploitation statistique des résultats nécessite quelquefois la connaissance et la formulation de la loi de distribution, au sens mathématique du terme, suivie par le nombre d'individus présents dans un échantillon de taille déterminée.

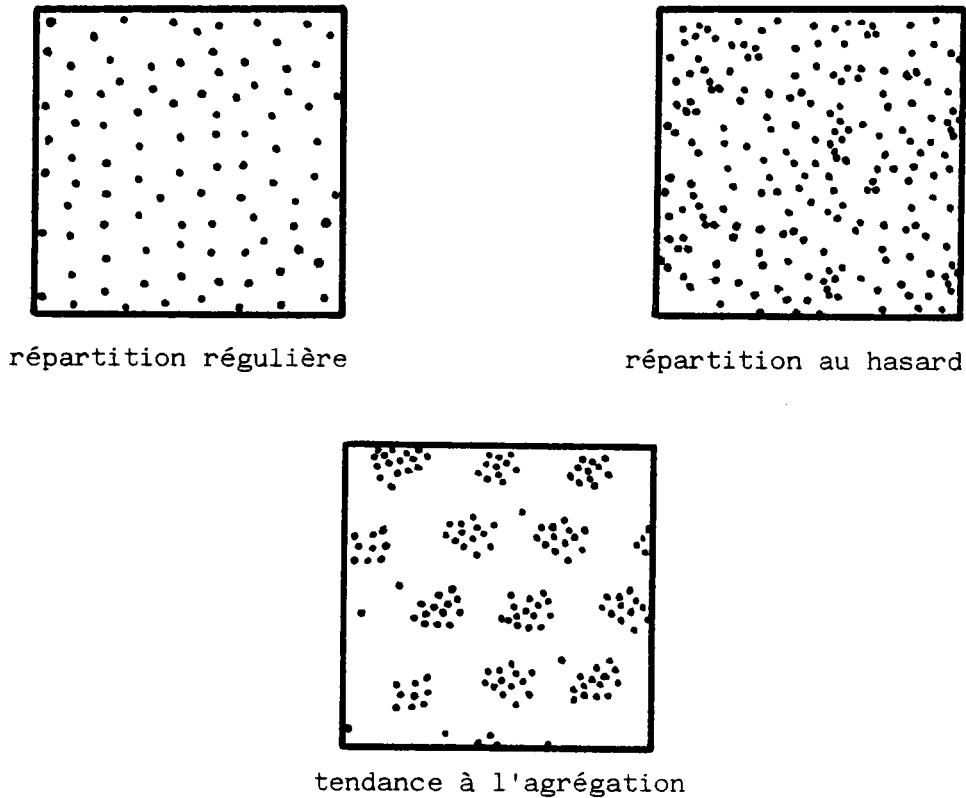


Figure 6 : Représentation schématique de trois modèles de répartition spatiale des individus d'une population (d'après POOLE, 1974).

Ainsi la loi de POISSON correspondant à la répartition au hasard des individus dans l'espace est très souvent citée comme hypothèse de base servant au calcul en microbiologie.

Le modèle de POISSON, qui constitue en fait un cas très singulier intermédiaire entre une dispersion régulière et une tendance à l'agrégation des individus dans l'espace, semble adapté à la distribution spatiale des unités formatrices de colonies à condition toutefois que les volumes de liquide soient petits et homogénéisés.

Le problème est plus délicat lorsque la densité bactérienne est variable selon le lieu de prélèvement. Cependant dans un travail remarquable ayant permis une stratification du lac ERIE en zones ayant sensiblement le même titre bactérien, EL-SHAARAWI et coll. (1981) ont montré l'intérêt de la distribution binomiale inverse pour la description de l'hétérogénéité bactérienne en limnologie.

## I - 2 MISE EN EVIDENCE DE L'HETEROGENEITE

L'imbrication possible des modèles d'hétérogénéité à des échelles différentes pose le problème du choix de la taille des échantillons pour la compréhension de la structure hiérarchique. La figure 7 tente de faire ressortir l'importance de la géométrie de l'échantillon pour la mise en évidence d'une hétérogénéité d'échelle déterminée.

En effet, un échantillonnage réalisé sur la population représentée à la figure 7 par des prélèvements de taille A ferait apparaître au travers de la variance d'essais répétés une hétérogénéité très marquée.

Par contre l'utilisation d'échantillons de taille B indiquerait une dispersion régulière des individus dans l'espace et ceci malgré une forte tendance à l'agrégation.

D'autre part, si une hétérogénéité peut être décelée par l'utilisation d'échantillons de taille adéquate, il est également possible d'entreprendre une recherche systématique de la configuration de celle-ci par un échantillonnage à plusieurs degrés dont les prélèvements aux différents niveaux sont pratiqués par des échantillons de calibres décroissants.

Toute augmentation significative du carré moyen associé à un niveau déterminé de l'échantillonnage traduit, erreurs de manipulation exclues, la présence d'une agrégation d'échelle correspondante.

C'est en appliquant ce principe que divers auteurs ont pu formuler des hypothèses concernant la configuration de l'hétérogénéité spatiale du plancton et des populations bactériennes en milieu marin ou lentique. En particulier lors d'une étude sur la variabilité spatiale de très grande échelle du phytoplancton marin côtier, PLATT et coll. (1970) ont montré que la variance interstation calculée pour 10 points de prélèvement augmentait d'abord rapidement pour ensuite se stabiliser alors que parallèlement la zone de prélèvement était progressivement élargie de 0,2 à environ 14 km<sup>2</sup>. De plus l'étude de l'autocorrélation, après

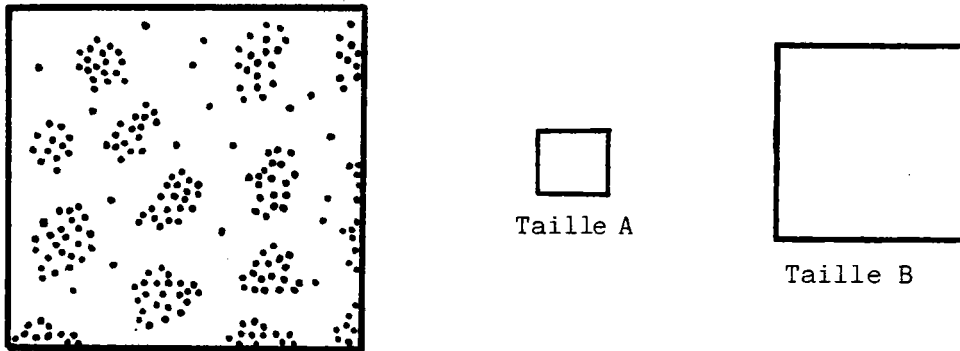


Figure 7: Mise en évidence de l'hétérogénéité et taille des échantillons (d'après POOLE, 1974).

extraction d'une tendance linéaire très nette, des valeurs numériques d'une série de 80 échantillons régulièrement prélevés le long d'une radiale de huit milles marins leur a suggéré le modèle de taches circulaires dont le diamètre moyen semble être 0,5 mille marin avec un espacement du même ordre de grandeur.

La dynamique des masses d'eau rendant une telle approche difficile en potamologie, la détermination de façon précise des paramètres géométriques de l'hétérogénéité bactérienne dans un cours d'eau ne sera pas envisagée dans le cadre de ce travail. Par contre, l'accent sera mis sur les effets de l'hétérogénéité quant à l'incertitude des estimations et cela pour des échantillons de taille fixée.

## CHAPITRE II : ETUDE EXPERIMENTALE DE QUELQUES CAUSES D'ERREUR DANS L'ANALYSE DES ECHANTILLONS.

Parmi les nombreux facteurs susceptibles d'influencer le résultat d'un titrage, quatre sources de variation (méthode, manipulateur, intra-échantillon, interéchantillon) ont plus particulièrement retenu notre attention. Aussi un plan d'expérience a été conçu de manière à pouvoir répondre aux questions suivantes :

- . La variance résiduelle des résultats numériques est-elle en accord avec la variance théorique de la méthode de titrage utilisée ?  
Cette comparaison permet notamment d'apprécier le rôle joué par l'ensemble des facteurs non contrôlés et constitue d'une certaine façon un critère de qualité quant au bon déroulement de l'expérience.
- . Les résultats sont-ils indépendants du manipulateur ?
- . Existe-t-il des différences significatives entre les densités bactériennes des sous-échantillons issus d'un même flacon (variation intraéchantillon)?
- . Enfin, les concentrations bactériennes trouvées pour plusieurs échantillons prélevés successivement au même endroit de la rivière à la surface de l'eau, peuvent-elles être considérées comme égales (variation interéchantillon) ?

L'appréciation de la contribution relative de chacun de ces quatre facteurs potentiels de variation à l'erreur globale est possible par l'étude des composantes de variance. Une telle démarche est, par ailleurs, nécessaire pour tenter d'optimiser le protocole analytique.

### II - 1 MATERIEL ET METHODES

La technique de titrage par la méthode du N.P.P. sur plaques de micro-titration ainsi que le matériel utilisé dans cette expérience ont été identiques à ceux décrits au paragraphe II-1-3 de la première partie.

L'exploitation statistique des résultats a été réalisée par analyse de la variance et la signification des effets des différents facteurs contrôlés a été vérifiée par le test de SNEDECOR.

Ne disposant que de deux degrés de liberté pour l'estimation de la variance correspondant à chaque combinaison des facteurs contrôlés, ce qui est insuffisant pour l'application du test de BARTLETT, l'homogénéité des variances a été vérifiée au moyen du test de COCHRAN. Le principe de ce test consiste à comparer la valeur numérique du rapport  $g = \frac{S^2_{\max}}{\sum_h S_h^2}$  où  $S^2_{\max}$  désigne la plus

forte des estimations  $S_h^2$  des variances des différentes classes de résultats, à la limite supérieure  $g_{1-\alpha}$  du domaine d'acceptation pour un risque d'erreur de première espèce  $\alpha$  donné. L'hypothèse d'égalité des variances est acceptée lorsque  $g$  est inférieur à cette limite.

## II - 2 PROTOCOLE EXPERIMENTAL

Trois échantillons de 500 ml d'eau ont été prélevés au même endroit de la rivière dans un intervalle de temps ne dépassant pas 3 minutes (temps nécessaire au remplissage successif des flacons).

Trois manipulateurs ont analysé chacun de ces trois flacons conformément au protocole représenté par la figure 8.

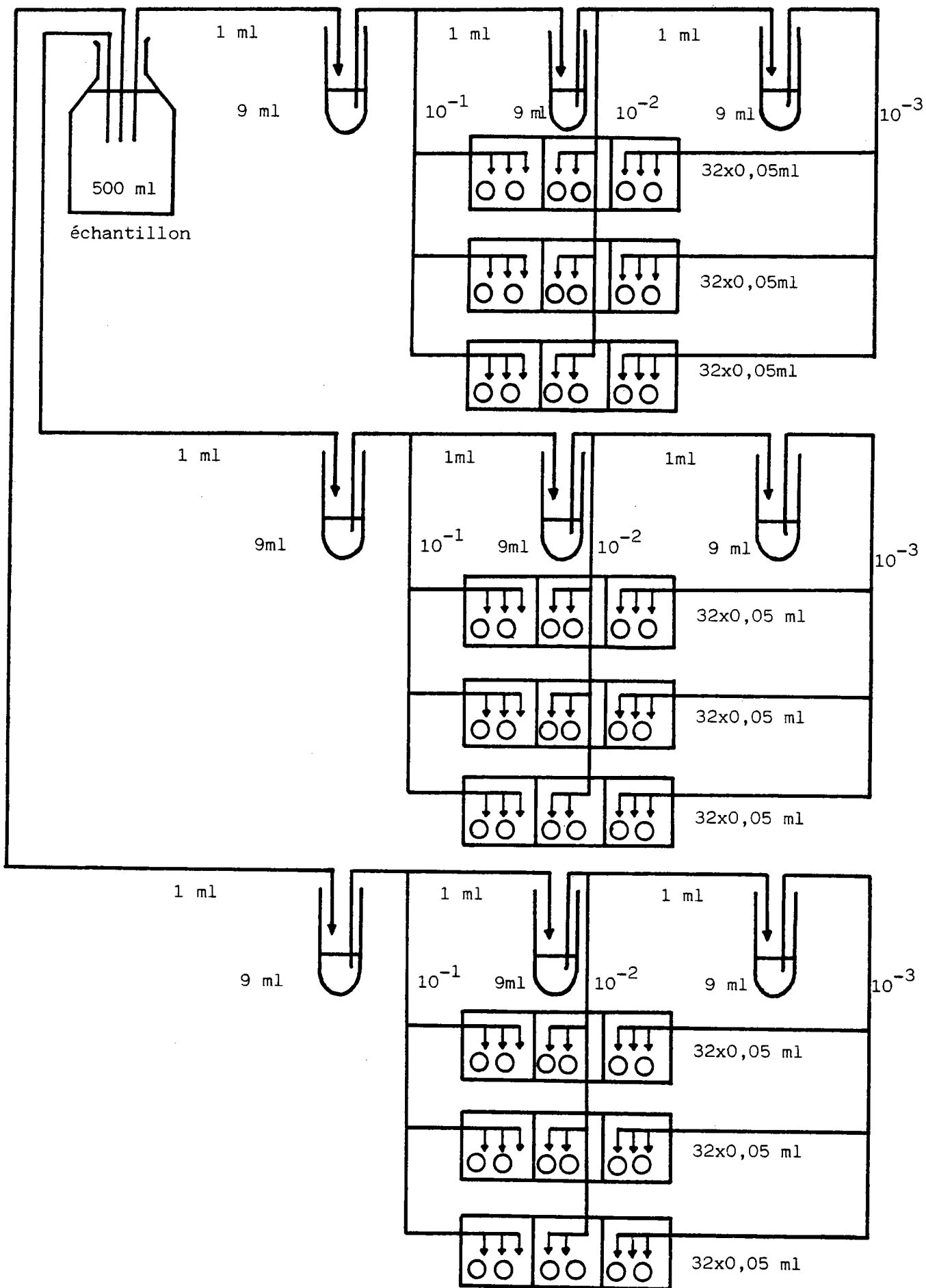
Chaque manipulateur a extrait trois sous-échantillons de 1ml de chaque échantillon de 500 ml (soit au total 9 sous-échantillons par manipulateur ). Tout prélèvement de liquide a été précédé d'une homogénéisation par 10 retournements du flacon.

Chaque sous-échantillon a été introduit dans 9 ml d'eau physiologique stérile et deux autres dilutions à  $10^{-2}$  et  $10^{-3}$  ont été préparées de manière analogue.

La série de dilutions ainsi obtenue a permis l'ensemencement de 3 plaques de microtitration à raison de 32 cupules par plaque et par dilution. Ceci nous a conduit à distinguer sur chaque plaque 3 zones, correspondant aux 3 niveaux de dilution, dont les limites ont dû être préalablement matérialisées. Tout sous-échantillon de 1 ml a ainsi été associé à un ensemble de trois plaques donnant chacune une estimation de sa concentration bactérienne par la méthode du N.P.P. avec 32 cupules par niveau de dilution.

Par ailleurs, afin de minimiser une éventuelle dérive de la densité bactérienne dans le temps et dans le but de rendre son action aléatoire, la progression suivante a été adoptée par chaque manipulateur :

Figure 8 : Schéma du protocole expérimental applicable à chaque manipulateur et pour chaque échantillon.



- dans un premier temps tous les sous-échantillons ont été prélevés (soit 9 par manipulateur) puis les dilutions ont été préparées
- ce n'est qu'après ces deux opérations que chaque manipulateur a procédé aux ensemencements à partir des 9 séries de dilution dans un ordre donné par une table de permutations de 9 nombres au hasard.

De plus, le temps s'écoulant entre le moment des prélèvements dans le milieu naturel et la fin des manipulations n'a pas dépassé 3 heures.

Après lecture et calcul des résultats comme décrit au paragraphe II-1-3 de la première partie, 27 indices N.P.P. ont été obtenus pour chaque échantillon (3 estimations de la densité bactérienne pour chacun des 3 sous-échantillons prélevés par chacun des 3 manipulateurs).

Les valeurs ainsi obtenues ont fait l'objet d'une transformation logarithmique pour être soumises à l'analyse de variance conformément au tableau 11 après vérification de l'égalité des variances par le test de COCHRAN.

Le modèle adapté à cette expérience est un plan mixte comprenant 3 facteurs contrôlés avec répétitions. Les trois facteurs contrôlés : échantillon, manipulateur et sous-échantillon, notés respectivement A, B et C interviennent chacun avec 3 modalités aléatoires. Les facteurs A et B se combinent factoriellement, par contre, chacune des modalités  $k$  de C est associée à un ensemble  $A_i B_j$  déterminé. Les notations utilisées dans le tableau 11 se soumettent aux règles données au paragraphe II - 2 de la première partie.



TABLEAU 11 : Tableau d'analyse de variance.

SOURCE DE VARIATION	SOMME DES CARRES Q	DEGRES DE LIBERTE $\nu$	QUANTITE ESTIMEE PAR Q/ $\nu$
FACTEUR A échantillon	$Q_A = 27 \sum_i (\bar{x}_{i..} - \bar{x})^2$	2	$\sigma_R^2 + 3\sigma_C^2 + 9\sigma_{AB}^2 + 27\sigma_A^2$
FACTEUR B manipulateur	$Q_B = 27 \sum_j (\bar{x}_{.j.} - \bar{x})^2$	2	$\sigma_R^2 + 3\sigma_C^2 + 9\sigma_{AB}^2 + 27\sigma_B^2$
INTERACTION AB	$Q_{AB} = 9 \sum_i \sum_j (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x})^2$	4	$\sigma_R^2 + 3\sigma_C^2 + 9\sigma_{AB}^2$
FACTEUR C sous-échantillon	$Q_C = 3 \sum_i \sum_j \sum_k (\bar{x}_{ijk} - \bar{x}_{ij.})^2$	18	$\sigma_R^2 + 3\sigma_C^2$
RESIDUELLE	$Q_R = \sum_i \sum_j \sum_k \sum_\alpha (x_{ijk\alpha} - \bar{x}_{ijk})^2$	54	$\sigma_R^2$
TOTALE	$Q_T = \sum_i \sum_j \sum_k \sum_\alpha (x_{ijk\alpha} - \bar{x})^2$	80	

II -3 RESULTATS

Le tableau 12a donne les triplets, et les indices N.P.P. associés, issus de chaque titrage ainsi que la concentration bactérienne exprimée en nombre de germes par ml de l'échantillon initial.

D'emblée il apparaît que le résultat de la première répétition du deuxième sous-échantillon prélevé dans le flacon n° 1 par le troisième manipulateur est aberrant : la concentration bactérienne estimée par le triplet 32-28-18 étant 6 à 7 fois supérieure aux deux autres essais réalisés dans les mêmes conditions. Sans qu'il soit possible de donner les raisons précises de cette anomalie, due vraisemblablement à une erreur de manipulation, le résultat douteux a été rejeté. Le remplacement de la donnée manquante par la moyenne géométrique des deux autres titrages effectués sur le même sous-échantillon a entraîné la perte d'un degré de liberté pour les variances totale et résiduelle dont les estimations respectives reposent dorénavant sur 79 et 53 degrés de liberté.

Echantillon	Manipulateur	sous échantillon	Répétitions			N.P.P.			Concentration bactérienne par ml		
			1	2	3	1	2	3			
1	1	1	31-8-0	32-11-2	32-8-2	3,08	4,87	3,36	616	974	672
		2	32-14-4	31-7-2	31-14-1	6,71	3,19	4,43	1342	639	885
		3	32-12-1	31-11-3	32-11-2	4,94	4,11	4,87	987	822	974
	2	1	31-11-0	32-9-2	31-10-1	3,59	4,27	3,57	718	854	714
		2	32-14-1	32-11-0	32-8-0	5,68	4,37	3,62	1136	874	724
		3	32-14-1	31-11-1	31-14-1	5,68	3,76	4,43	1136	752	886
	3	1	32-6-0	31-9-1	32-15-1	3,22	3,39	6,11	644	678	1222
		2	32-28-18	32-13-0	32-14-1	34,05	5,00	5,68	6810	1000	1136
		3	32-9-2	32-12-2	31-11-3	4,27	5,22	4,11	854	1044	822
2	1	1	32-10-2	31-12-6	32-9-0	4,56	4,97	3,84	912	994	768
		2	32-9-3	31-16-2	32-11-3	4,51	5,19	5,15	902	1038	1030
		3	32-4-0	32-15-4	32-15-0	2,87	7,20	5,77	574	1440	1154
	2	1	32-11-3	31-10-1	32-14-0	5,15	3,57	5,37	1030	714	1074
		2	30-6-1	32-10-2	30-13-2	2,57	4,56	3,70	514	912	740
		3	32-5-2	31-10-0	31-12-0	3,35	3,41	3,79	670	682	758
	3	1	32-10-1	30-13-2	31-6-1	4,32	3,70	2,91	864	740	582
		2	32-10-1	32-12-2	32-9-2	4,32	5,22	4,27	864	1044	854
		3	31-9-1	31-15-2	20-12-2	3,39	4,90	2,78	678	980	556
3	1	1	30-11-1	31-8-2	32-10-1	3,24	3,36	4,32	648	672	864
		2	32-9-0	32-4-3	31-12-0	3,84	3,33	3,79	768	666	758
		3	32-9-0	31-8-2	30-14-1	3,84	3,36	3,72	768	672	744
	2	1	31-7-3	31-9-2	32-9-2	3,34	3,54	4,27	668	708	854
		2	31-7-0	31-11-1	31-9-3	2,92	3,76	3,70	584	752	740
		3	28-14-1	31-9-0	30-10-1	2,91	3,24	3,09	582	648	618
	3	1	31-10-1	29-8-2	31-12-2	3,57	2,62	4,15	714	524	830
		2	31-9-1	29-9-3	29-7-3	3,39	2,84	2,61	678	568	522
		3	30-9-4	32-13-1	31-12-0	3,33	5,29	3,79	666	1058	758

TABLEAU 12 a : Résultats des titrages réalisés par la méthode du N.P.P. avec 32 cupules par dilution pour chaque combinaison des facteurs A, B et C.

L'analyse statistique des résultats transformés par la fonction logarithme décimal s'est déroulée de la manière suivante :

- . Après vérification de l'homogénéité des variances par le test de COCHRAN, l'analyse de variance a été menée en calculant dans un premier temps les carrés moyens correspondant aux divers facteurs étudiés puis en exécutant les tests de signification relatifs à ces facteurs dans l'ordre imposé par le tableau d'analyse de variance.

\* Vérification de l'homogénéité des variances.

Le calcul du rapport  $g = \frac{S_{\max}^2}{\sum_h S_h^2}$  à partir des estimations  $S_h^2$  consignées dans le tableau 12b conduit à la valeur  $g = \frac{0,04337}{0,22499} \approx 0,1928$ . Cette valeur étant inférieure à la limite supérieure du domaine d'acceptation pour un risque égal à 0,05 ( $g_{0,95} = 0,22$ ), l'égalité des variances a été admise.

\* Calcul des carrés moyens

Le carré moyen associé à chaque facteur contrôlé a été calculé à partir des sommes figurant sur le tableau 12b. Les valeurs ainsi obtenues sont résumées dans le tableau 13.

\* Tests de signification

L'absence d'effet dû à un facteur contrôlé se vérifie en appliquant le test de SNEDECOR.

- . Le rapport  $F_c = \frac{q_c}{q_R} = 0,95$  étant inférieur à 1, le test relatif

au facteur C (sous-échantillon) est non significatif. Les mesures faites ne permettent pas d'évaluer la contribution des fluctuations du facteur contrôlé C à la variance globale mais ces fluctuations seront jointes à la variance résiduelle dont une nouvelle estimation avec 71 degrés de liberté est donnée par la quantité :

$$q'_R = \frac{0,45000 + 0,14570}{53 + 18} = 0,00839$$

Echantillon	Manipulateur	sous- échantillon	1	Répétitions 2	3	Variance Empirique $S_h^2$	$\sum \frac{X_{ijkq}}{q}$	$\sum \sum \frac{X_{ijkq}}{kq}$	$\sum \sum \sum \frac{X_{ijkq}}{jkq}$	$\sum \sum \sum \frac{X_{ijkq}}{ijkq}$
1	1	1	2,790	2,989	2,827	0,01120	8,606	26,385	79,435	234,998
		2	3,128	2,805	2,947	0,02621	8,880			
		3	2,995	2,915	2,989	0,00199	8,899			
	2	1	2,856	2,931	2,854	0,00193	8,641	26,376		
		2	3,055	2,942	2,860	0,00959	8,857			
		3	3,055	2,876	2,947	0,00812	8,878			
	3	1	2,809	2,831	3,087	0,02388	8,727	26,674		
		2	3,027	3,000	3,055	0,00076	9,082			
		3	2,931	3,019	2,915	0,00314	8,865			
2	1	1	2,960	2,997	2,885	0,00326	8,842	26,805	78,808	
		2	2,955	3,016	3,013	0,00118	8,984			
		3	2,759	3,158	3,062	0,04337	8,979			
	2	1	3,013	2,854	3,031	0,00949	8,898	25,978		
		2	2,711	2,960	2,869	0,01587	8,540			
		3	2,826	2,834	2,880	0,00085	8,540			
	3	1	2,937	2,869	2,765	0,00750	8,571	26,025		
		2	2,937	3,019	2,931	0,00242	8,887			
		3	2,831	2,991	2,745	0,01559	8,567			
3	1	1	2,812	2,827	2,937	0,00466	8,576	25,748	76,755	
		2	2,885	2,823	2,880	0,00119	8,588			
		3	2,885	2,827	2,872	0,00093	8,584			
	2	1	2,825	2,850	2,931	0,00307	8,606	25,485		
		2	2,766	2,876	2,869	0,00379	8,511			
		3	2,765	2,812	2,791	0,00055	8,368			
	3	1	2,854	2,719	2,919	0,01041	8,492	25,522		
		2	2,831	2,754	2,718	0,00333	8,303			
		3	2,823	3,024	2,880	0,01073	8,727			

TABLEAU 12b : Résultats des titrages après transformation logarithmique et sommes nécessaires aux calculs.

TABLEAU 13 : Tableau d'analyse de variance.

Source de variation	Somme de carrés Q	$\nu$	Carré moyen Q/ $\nu$	Quantité estimée par Q/ $\nu$ = q
Facteur A échantillon	0,14556	2	0,07278	$\sigma_R^2 + 3 \sigma_C^2 + 9 \sigma_{AB}^2 + 27 \sigma_A^2$
Facteur B manipulateur	0,02306	2	0,01153	$\sigma_R^2 + 3 \sigma_C^2 + 9 \sigma_{AB}^2 + 27 \sigma_B^2$
Interaction AB	0,03578	4	0,00894	$\sigma_R^2 + 3 \sigma_C^2 + 9 \sigma_{AB}^2$
Facteur C sous-échantillon	0,14570	18	0,00809	$\sigma_R^2 + 3 \sigma_C^2$
Résiduelle	0,45000	53	0,00849	$\sigma_R^2$
Totale	0,80010	79		

. D'autre part, le rapport  $F_{AB} = \frac{q_{AB}}{q_R} = 1,07$  est inférieur à la limite  $F_{0,95} (4 ; 71) \approx 2,50$  donnée par la loi de SNEDECOR ; l'existence d'une interaction entre les facteurs A et B n'est donc pas mise en évidence.

Cette interaction, qui ne trouve d'ailleurs aucune justification, joue le rôle d'un facteur non contrôlé. Ces fluctuations se conjuguent avec celles des autres facteurs non contrôlés pour aboutir à la variance résiduelle dont le quotient  $q_R'' = \frac{0,59570 + 0,03578}{71 + 4} = 0,00842$  constitue désormais une estimation avec 75 degrés de liberté.

.  $F_B = \frac{0,01153}{0,00842} = 1,37$  valeur inférieure à la limite donnée par la table

de SNEDECOR :  $F_{0,95} (2 ; 75) \approx 3,13$ .

Bien que l'influence du facteur B n'ait pas été mise en évidence, la composante de variance associée à l'effet manipulateur est estimée par  $S_B^2 = 0,00012$  ce qui correspond à un coefficient de variation logarithmique de 2,5%.

Enfin, l'influence éventuelle du facteur A (échantillon) se contrôle en comparant la fonction discriminante  $F_A = \frac{q_A}{q_R} = 8,64$  à la valeur  $F_{0,95}(2;75) = 3,13$ .

Dépassant largement cette limite, la quantité  $F_A$  est même supérieure à  $F_{0,995}(2;75) \simeq 5,70$  ce qui traduit de façon très significative une différence dans la concentration bactérienne des trois échantillons. Le coefficient de variation logarithmique calculé à partir de la composante de variance associée au facteur A vaut approximativement 11%.

Le tableau 14 donne les estimations de la concentration bactérienne de chaque échantillon ainsi qu'une estimation de la densité bactérienne globale pour la rivière.

Les calculs reposent sur l'hypothèse de normalité du logarithme de la densité moyenne et les limites de confiance sont obtenues par la loi de STUDENT à partir des carrés moyens appropriés extraits du tableau 13.

**TABLEAU 14:** Concentration bactérienne moyenne de chaque échantillon et estimation de la charge bactérienne de la rivière.

	Estimation de la concentration bactérienne par ml.	Limites d'un intervalle de confiance à 95 %		
Echantillon 1	875	806	-	950
Echantillon 2	829	764	-	900
Echantillon 3	696	642	-	756
Rivière	797	592	-	1072

## II - 4 DISCUSSION

Le résultat numérique de tout titrage bactérien est entaché d'une incertitude qui peut être considérée comme la résultante des effets de plusieurs sources de variation plus ou moins importantes et que nous avons tenté de mettre en évidence.

Nous examinerons successivement la variabilité inhérente à la méthode de titrage, l'influence du manipulateur puis les variations intra et interéchantillon.

### a) Incertitude liée à la méthode de titrage

La variance résiduelle, qui reflète en quelque sorte la précision de la méthode de numération, se traduit par un coefficient de variation égal à 21,5% alors que la précision théorique correspond approximativement à 22,7%. La similitude de ces deux valeurs confirme la conclusion déjà formulée dans la première partie. En effet, la méthode du N.P.P. se distingue par une variance propre qui ne dépend, en première approximation, que du nombre  $n$  de cupules inoculées à chaque niveau de dilution.

### b) Influence du manipulateur

Aucune différence significative entre les résultats émanant des trois manipulateurs n'a pu être mise en évidence. La faible valeur du coefficient de variation associé à ce facteur (2,5%) montre que l'incertitude sur le résultat éventuellement engendrée par l'identité du manipulateur est vraisemblablement très faible.

Au cours d'une expérience à but similaire réalisée avec la méthode de numération sur milieu gélosé, BREITTMAYER et GAUTHIER (1979) ont montré que le manipulateur peut introduire un biais relativement important (35%).

La formulation d'une conclusion s'appuyant sur ces deux expériences isolées et visant à exploiter leurs résultats fort différents n'est pas souhaitable.

Cependant il convient de souligner que la netteté des réponses pour la méthode quantique en milieu liquide enlève toute subjectivité au comptage des résultats ce qui n'est pas toujours le cas lors du recensement des colonies apparues sur milieu gélosé.

c) Variation intraéchantillon.

Le plan d'expérience n'a pas décelé de variation intraéchantillon, ce qui dénote l'absence d'hétérogénéité de petite échelle perceptible à l'intérieur des flacons. Mais le caractère ponctuel de cette expérience ne se prête guère à la généralisation de ce résultat. Des travaux réalisés sur des eaux marines (ASHBY et RHODES-ROBERTS, 1976, KAPER et coll., 1978) ainsi qu'en faciès lentique (KOSINSKI et coll., 1979; PALMER et coll., 1976) montrent la présence dans ces eaux de micropopulations bactériennes dont la taille est de l'ordre du centimètre.

A notre connaissance, ces micropopulations, dont l'origine est mal connue et dont l'observation est probablement étroitement liée à la qualité de l'homogénéisation précédant la prise des sous-échantillons, n'ont pas été mentionnées en milieu lotique.

Notre expérience ne permettant nullement d'en rejeter l'existence, le problème soulevé par cette hétérogénéité de petite échelle sera de nouveau abordé ultérieurement.

d) Variation interéchantillon

Le test relatif au facteur échantillon fait apparaître une réelle différence de la densité bactérienne entre les trois flacons et cela d'une manière très significative. Ce résultat témoigne une hétérogénéité du milieu hydrique considéré dont l'échelle n'a rien de commun avec celle des micropopulations précédemment évoquées.

Les chiffres du tableau 14 montrent que les nombreux titrages réalisés permettent une très bonne connaissance des taux bactériens au sein de chaque échantillon. Il en est tout autrement lorsque l'estimation est étendue à la concentration bactérienne moyenne de la rivière au moment des prélèvements : celle-ci est en effet estimée à 797 germes par millilitre avec des limites de confiance à 95 % respectivement égales à 592 et 1072 microorganismes par ml. La médiocrité de la précision du résultat malgré le nombre important de titrages pris en compte (au nombre de 80 !) suscite l'étonnement et semble tenir du paradoxe.

Bien que ce point n'ait pas été le but de cette manipulation, il apparaît clairement qu'un tel protocole analytique n'est absolument pas adapté à l'étude de l'état de pollution bactérienne de la rivière.



L'explication de ce phénomène, bien connu des statisticiens, se trouve dans les considérations suivantes.

Si r indices N.P.P. sont obtenus à partir de chacun des sous-échantillons prélevés dans p échantillons de l'eau à analyser à raison de q sous-échantillons par flacon, l'échantillonnage ainsi mené comporte trois degrés dont chacun apporte sa contribution à la variance totale. L'incertitude associée au résultat d'un titrage dépend dans ce cas de plusieurs sources de variation emboîtées les unes dans les autres. Ainsi les effets de l'hétérogénéité interéchantillon se superposent à la variabilité caractérisant l'hétérogénéité intraéchantillon dans laquelle viennent enfin s'imbriquer les fluctuations inhérentes à la technique de titrage.

Le calcul montre que la densité bactérienne de la rivière est alors estimée par la moyenne géométrique des résultats de tous les titrages.

Si  $\bar{x}$  désigne la moyenne arithmétique des logarithmes de ces données brutes, la variance de  $\bar{x}$ , notée  $\sigma_{\bar{x}}^2$ , est donnée par la relation (5)

$$(5) \quad \sigma_{\bar{x}}^2 = \frac{\sigma_{inter}^2}{p} + \frac{\sigma_{intra}^2}{pq} + \frac{\sigma_{N.P.P.}^2}{pqr} \quad (\text{SNEDECOR et COCHRAN, 1957})$$

où  $\sigma_{inter}^2$ ,  $\sigma_{intra}^2$  et  $\sigma_{N.P.P.}^2$  représentent respectivement les composantes de variance inter et intraéchantillon et la variance propre à la méthode de titrage.

L'examen de la formule (5) montre qu'une augmentation des nombres p, q et r provoque la diminution de  $\sigma_{\bar{x}}^2$  et par conséquent une amélioration de la précision. Si cette remarque peut sembler évidente, plus intéressant est de constater que pour un nombre total p q r d'indices N.P.P. donné,  $\sigma_{\bar{x}}^2$  est minimal lorsque q et r sont tous deux égaux à 1.

Le protocole le plus efficace consiste donc à pratiquer un titrage unique sur le plus grand nombre possible d'échantillons. Il serait par contre peu judicieux de multiplier les titrages à l'intérieur de chaque flacon. Une telle démarche ne manquerait pas de procurer une bonne connaissance de leur concentration bactérienne mais l'extrapolation des résultats à l'ensemble de la rivière risque d'aboutir à une estimation dont la précision n'est pas à la hauteur des moyens mis en oeuvre. Cette disparité est accentuée lorsque l'hétérogénéité du milieu hydrique est notable et si de surcroît le nombre d'échantillons est limité.

S'il est maintenant acquis que les répétitions doivent se faire de préférence sur les unités primaires c'est-à-dire les échantillons prélevés dans le milieu naturel, des considérations de coût posent le problème sous un angle nouveau car la solution préconisée n'est pas forcément la moins onéreuse.

Par ailleurs, la connaissance des composantes de variance caractérisant les fluctuations à chaque niveau de l'échantillonnage permet :

- d'une part, de chiffrer l'incidence du protocole analytique sur la précision des estimations,
- d'autre part d'en envisager l'optimisation compte tenu d'une formulation exacte des coûts comparés des différents plans possibles.

Une réponse explicite à ces questions pour le moment prématurées ne sera développée qu'ultérieurement lorsque des estimations plus fiables des composantes de variance seront disponibles. Aussi les expériences à venir tenteront-elles, entre autres, d'appréhender de manière plus précise l'impact que peut exercer l'hétérogénéité du milieu naturel sur les résultats.

### CHAPITRE III : ETUDE EXPERIMENTALE DE L'HETEROGENEITE SPATIO-TEMPORELLE D'UN MILIEU HYDRIQUE.

Ce chapitre aborde simultanément l'étude de l'hétérogénéité dans l'espace et dans le temps de la concentration bactérienne d'un milieu hydrique particulier en faciès lotique.

D'emblée il faut souligner la difficulté que soulève quelquefois la distinction entre les variations spatiales et temporelles de la qualité bactériologique d'une eau de rivière. Ainsi une hétérogénéité à l'échelle kilométrique est pour une large part solidaire des caractéristiques morphodynamiques du cours d'eau. Mais en adoptant par la force des choses une présentation linéaire, cette hétérogénéité peut aussi se manifester en un endroit fixé par des fluctuations dans le temps.

De plus, l'hétérogénéité spatiale peut présenter un caractère fixe et permanent imputable au site, et à une échelle moindre un aspect instantané et imprévisible lié aux turbulences de l'eau. C'est pourquoi, la possibilité d'une variation dans la concentration bactérienne selon le lieu de prélèvement et le problème des variations inter et intraéchantillon ont été envisagés dans cette étude sur une portion très limitée du cours d'eau. L'étude de l'hétérogénéité dans le temps a été limitée à la mise en évidence de fluctuations journalières et horaires des taux bactériens sans prétendre pour le moment réaliser une étude précise de ces variations. La connaissance de l'importance relative de tous ces facteurs de variation doit permettre d'une part de saisir globalement et d'une manière plus complète les différents aspects de l'hétérogénéité bactérienne du cours d'eau et d'autre part de poser clairement et en connaissance de cause des questions motivant éventuellement de nouvelles expérimentations.

#### III - 1 MATERIEL ET METHODES.

##### III - 1-1 Situation des points de prélèvement et collecte des échantillons.

Les échantillons ont été prélevés à la surface de l'eau à 6 mètres environ de la berge dans des flacons en verre de 500 ml préalablement stérilisés à l'autoclave et fixés à l'extrémité d'une canne rigide.

Trois points de prélèvement espacés de 150 m environ ont été choisis le long du canal dérivé de la Moselle alimentant la centrale thermique de la MAXE en eaux de refroidissement.

L'entrée du canal est située à 600 m environ en aval de la station d'épuration biologique de la ville de METZ.

La figure 9 donne la position des points de prélèvement désignés par les lettres A,B et C ainsi qu'une représentation schématique du site.

Les échantillons ont été ramenés au laboratoire à température ambiante pour être soumis à l'analyse immédiate.

Le temps s'écoulant entre le début de la collecte et l'arrivée des flacons au laboratoire n'a jamais dépassé 45 minutes.

### III 1-2 Numération des coliformes thermorésistants et analyse statistique des résultats.

Les titrages bactériens ont été réalisés par la méthode du N.P.P sur plaques de microtitration conformément à la description qui en a été donnée en première partie, paragraphe II-1-3.

L'exploitation statistique des titres bactériens a été effectuée par analyse de la variance après vérification de l'homogénéité des variances par le test de COCHRAN dont le principe est exposé au paragraphe II-1 de la deuxième partie. L'absence d'effet dû à un facteur contrôlé a été vérifiée par le test de SNEDECOR.

### III - 2 PROTOCOLE EXPERIMENTAL.

Deux prélèvements (9h et 15 h) ont été réalisés au cours de 5 journées choisies pendant la période du 15 mai au 15 juin 1981. Compte tenu des disponibilités du laboratoire, les dates suivantes ont été retenues :

lundi 18 mai, mardi 26 mai, vendredi 29 mai, lundi 1er juin et vendredi 5 juin.

Au cours de chaque prélèvement un échantillon de 500 ml d'eau de rivière a été recueilli à chacun des trois endroits désignés par A,B et C.

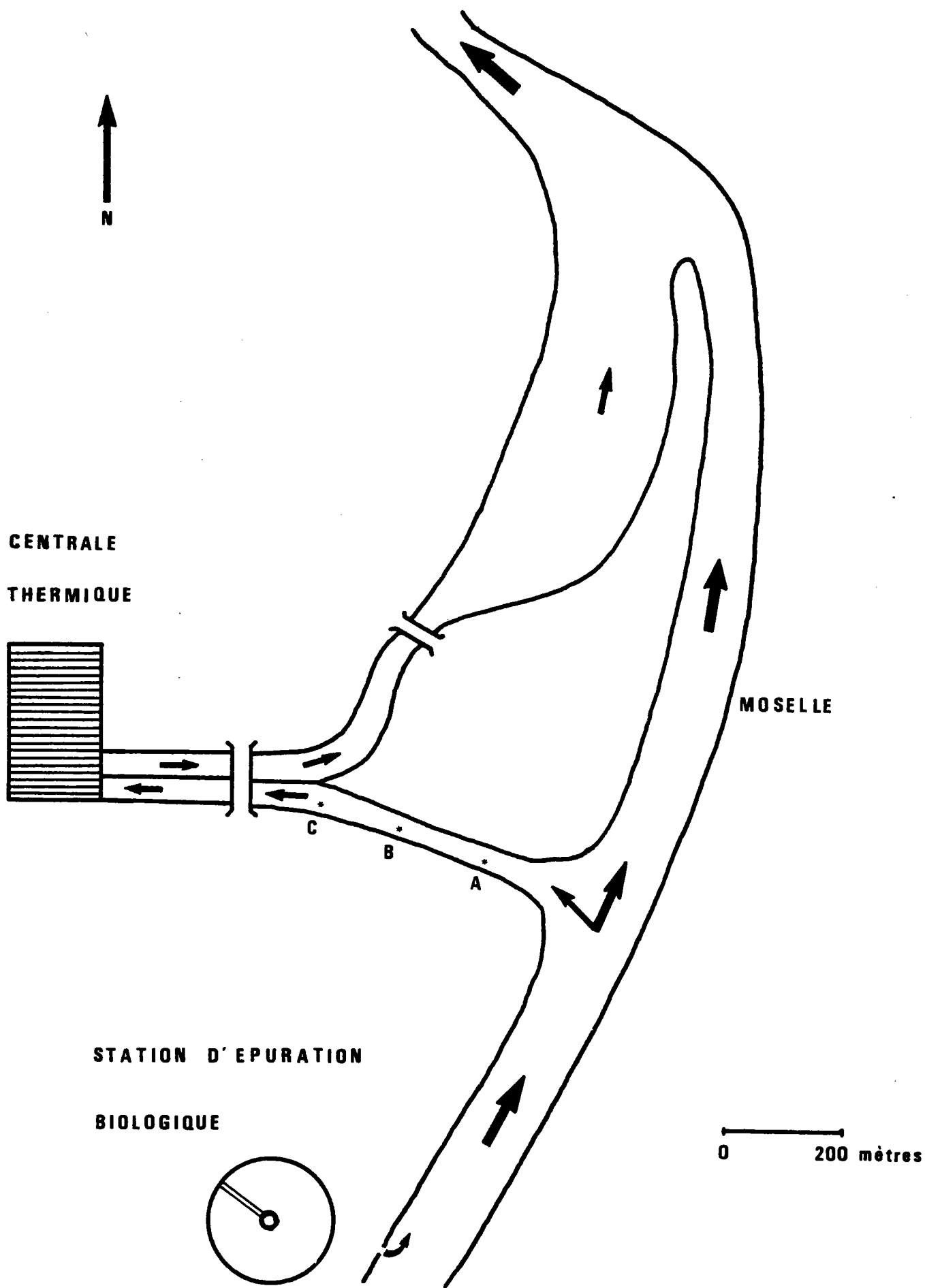
La collecte a toujours été exécutée dans l'ordre C, B puis A (en remontant le courant).

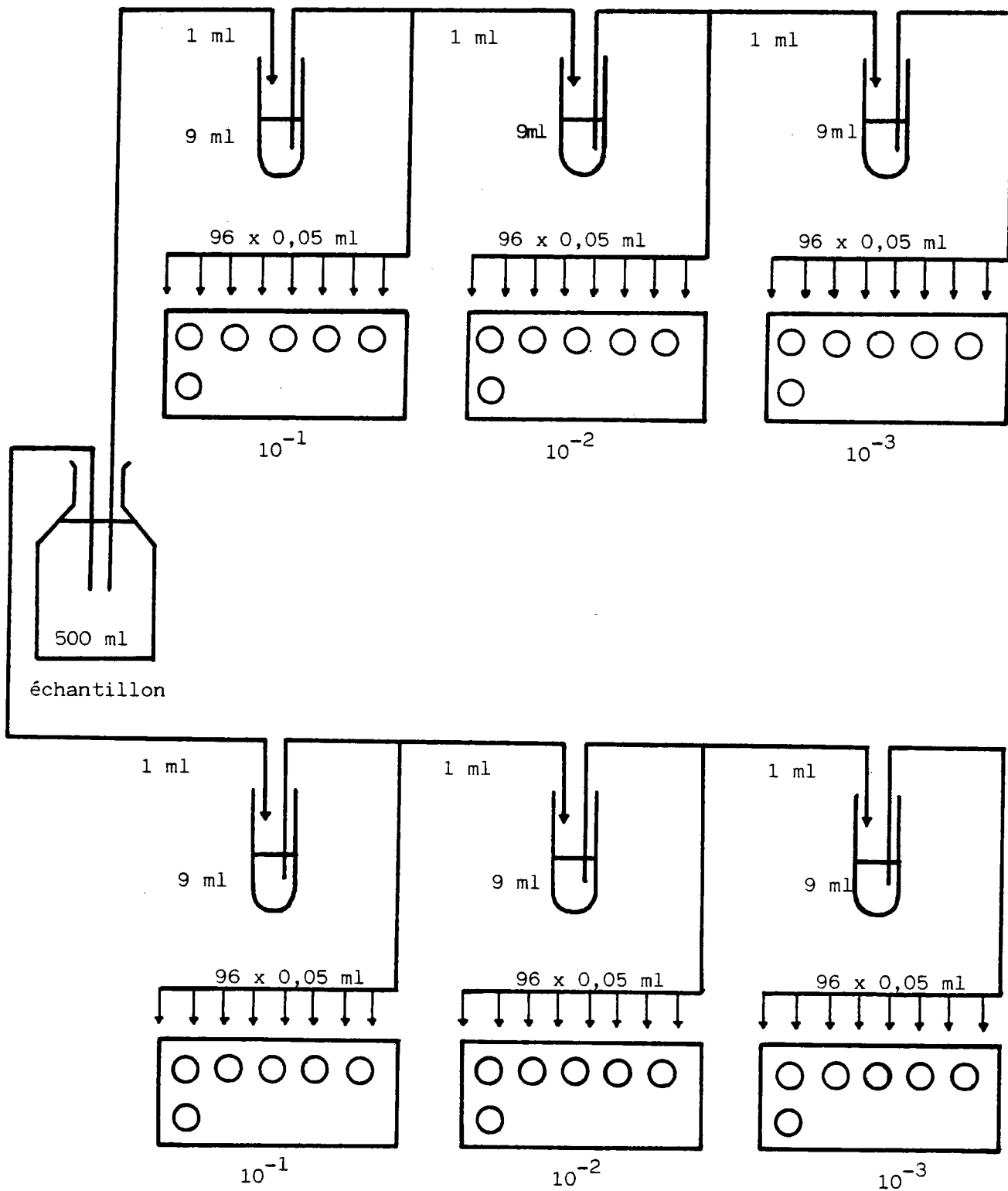
Chaque flacon a fait l'objet d'un double titrage en suivant le protocole analytique représenté par la figure 10.

A cet effet et après homogénéisation par 10 retournements du flacon de 500 ml, deux sous-échantillons de 1ml chacun ont été prélevés distinctement.

Chaque sous-échantillon a été introduit dans 9 ml d'eau physiologique stérile. Deux autres dilutions à  $10^{-2}$  et  $10^{-3}$  ont été préparées de manière analogue.

**FIGURE 9** : Représentation schématique du site et position des points de prélèvement.





**FIGURE 10** : Schéma du protocole analytique applicable pour chaque échantillon.

Chacune des deux séries de dilutions ainsi obtenues a permis la réalisation d'un titrage par la méthode du N.P.P. en milieu liquide.

Dans le but de rendre aléatoire l'action éventuelle du temps sur les concentrations bactériennes, les 3 flacons ont été traités séparément dans un ordre établi au hasard. La période s'écoulant entre le début des prélèvements dans le milieu naturel et la fin des titrages n'a jamais excédé 2h30.

Après lecture et calcul des résultats comme décrit au paragraphe II-1-3 de la première partie, deux estimations indépendantes de la densité bactérienne ont été obtenues pour chaque flacon.

L'ensemble de toutes les valeurs ainsi recueillies au cours de cette campagne de prélèvements a fait l'objet d'une transformation logarithmique avant d'être soumis à l'analyse de variance selon le modèle représenté par le tableau 15.

Il s'agit d'un plan factoriel à 3 facteurs contrôlés avec répétitions. Le facteur jour, noté A, intervient avec 5 modalités aléatoires alors que les facteurs lieu et heure, notés B et C, possèdent respectivement 3 et 2 modalités fixes. Les notations utilisées sont en accord avec les conventions d'écriture établies dans la première partie, paragraphe II-2.

En particulier, les indices  $i$ ,  $j$  et  $k$  se rapportent respectivement aux facteurs A, B et C.

**TABLEAU 15** : Tableau d'analyse de variance.

SOURCE DE VARIATION	SOMME DE CARRES Q	Degrés de liberté	QUANTITE ESTIMEE PAR Q/v
FACTEUR A jour	$Q_A = 12 \sum_i (\bar{x}_{i.} - \bar{x})^2$	4	$\sigma_R^2 + 12\sigma_A^2$
FACTEUR B lieu	$Q_B = 20 \sum_j (\bar{x}_{.j} - \bar{x})^2$	2	$\sigma_R^2 + 4\sigma_{AB}^2 + 20\sigma_B^2$
FACTEUR C heure	$Q_C = 30 \sum_k (\bar{x}_{..k} - \bar{x})^2$	1	$\sigma_R^2 + 6\sigma_{AC}^2 + 30\sigma_C^2$
INTERACTION AB	$Q_{AB} = 4 \sum_i \sum_j (\bar{x}_{ij.} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$	8	$\sigma_R^2 + 4\sigma_{AB}^2$
INTERACTION AC	$Q_{AC} = 6 \sum_i \sum_k (\bar{x}_{i.k} - \bar{x}_{i.} - \bar{x}_{..k} + \bar{x})^2$	4	$\sigma_R^2 + 6\sigma_{AC}^2$
INTERACTION BC	$Q_{BC} = 10 \sum_j \sum_k (\bar{x}_{.jk} - \bar{x}_{.j} - \bar{x}_{..k} + \bar{x})^2$	2	$\sigma_R^2 + 2\sigma_{ABC}^2 + 10\sigma_{BC}^2$
INTERACTION ABC	$Q_{ABC} = Q_T - Q_A - Q_B - Q_C - Q_{AB} - Q_{AC} - Q_{BC} - Q_R$	8	$\sigma_R^2 + 2\sigma_{ABC}^2$
RESIDUELLE	$Q_R = \sum_i \sum_j \sum_k \sum_\alpha (x_{ijk\alpha} - \bar{x}_{ijk})^2$	30	$\sigma_R^2$
TOTALE	$Q_T = \sum_i \sum_j \sum_k \sum_\alpha (x_{ijk\alpha} - \bar{x})^2$	59	

III - 3 RESULTATS.

Le tableau 16 donne les résultats des concentrations bactériennes obtenues au cours de la campagne d'échantillonnage du 18 mai au 5 juin 1981 selon le jour, l'heure et l'endroit du prélèvement. Une estimation de la variance intra-échantillon a été calculée pour chaque paire de résultats provenant d'un même échantillon.



Date	Point de prélèvement	Heure du prélèvement	Triplet	N. P. P.	Concentration bactérienne par ml (u)	log u	Variance empirique $S_h^2$
18 MAI	A	9 h	87-31- 6	2,89	578	2,762	0,000490
			87-48- 5	3,63	726	2,861	
		15 h	96-41- 3	5,55	1110	3,045	0,04090
			90-27- 1	2,87	574	2,759	
	B	9 h	88-21- 3	2,50	500	2,699	0,00039
			88-21- 8	2,67	534	2,727	
		15 h	91-32- 3	3,33	666	2,824	0,00101
			94-28- 4	3,70	740	2,869	
	C	9 h	89-32- 1	2,98	596	2,775	0,00490
			86-23- 2	2,37	474	2,676	
		15 h	91-24-5	3,02	604	2,781	0,00266
			88-24-1	2,55	510	2,708	
26 MAI	A	9 h	96-49-14	8,20	1640	3,215	0,00769
			96-43- 6	6,16	1232	3,091	
		15 h	96-51- 8	7,76	1552	3,191	0,00470
			95-48- 6	6,21	1242	3,094	
	B	9 h	96-51-15	8,74	1748	3,243	0,00205
			96-58-14	10,14	2028	3,307	
		15 h	96-59- 5	8,97	1794	3,254	0,00042
			96-59- 9	9,59	1918	3,283	
	C	9 h	96-54-10	8,64	1728	3,238	0,000002
			96-51-14	8,60	1720	3,236	
		15 h	96-58- 7	9,05	1810	3,258	0,00020
			96-54-10	8,64	1728	3,238	

TABEAU 16 : Titres bactériens de l'eau de Moselle au cours de la campagne de prélèvements du 18 mai 1981 au 5 juin 1981.

Date	Point de prélèvement	Heure du prélèvement	Triplet	N. P. P.	Concentration bactérienne par ml (u)	log u	Variance empirique $S_h^2$	
29 MAI	A	9 h	96-49-10	7,66	1532	3,185	0,00001	
			96-50- 8	7,57	1514	3,180		
		15 h	92-52- 1	4,82	2964	2,984	0,02142	
			94-16- 4	2,99	598	2,777		
		B	9 h	96-41- 5	5,76	1152	3,061	0,01051
				96-56- 3	8,04	1608	3,206	
	15 h		93-27- 3	3,39	678	2,831	0,00520	
			93-15- 1	2,68	536	2,729		
	C	9 h	96-62- 7	9,98	1996	3,300	0,00001	
			96-63- 6	10,06	2012	3,304		
		15 h	90-25- 3	2,86	572	2,757	0,00115	
			87-26-2	2,56	512	2,709		
1 JUIN	A	9 h	96-56-19	10,44	2088	3,320	0,00097	
			96-55-14	9,45	1890	3,276		
		15 h	96-33- 5	4,78	956	2,980	0,00378	
			95-42-10	5,83	1166	3,067		
	B	9 h	96-56-10	9,06	1812	3,258	0,00003	
			96-49-19	8,90	1780	3,250		
		15 h	95-51-19	8,19	1638	3,214	0,00051	
			96-48-11	7,61	1522	3,182		
	C	9 h	96-40- 9	6,06	1212	3,084	0,00002	
			95-46- 8	6,15	1230	3,090		
		15 h	95-42-5	5,35	1070	3,029	0,00684	
			95-51-9	6,99	1398	3,146		

TABLEAU 16: (suite)

Date	Point de prélèvement	Heure du prélèvement	Triplet	N. P. P.	Concentration bactérienne par ml (u)	log u	Variance empirique $S_h^2$
5 juin	A	9 h	94-32-11	4,45	890	2,949	0,00020
			92-40- 7	4,25	850	2,929	
		15 h	96-43- 1	5,60	1120	3,049	
			96-35- 9	5,39	1078	3,033	
	B	9 h	94-30- 9	4,15	830	2,919	0,00405
			95-39- 6	5,10	1020	3,009	
		15 h	95-40- 5	5,12	1024	3,010	
			95-33- 5	4,42	884	2,946	
	C	9 h	93-36- 6	4,15	830	2,919	0,00451
			95-42- 3	5,16	1032	3,014	
		15 h	96-42- 6	6,01	1202	3,080	
			91-29- 5	3,27	654	2,816	

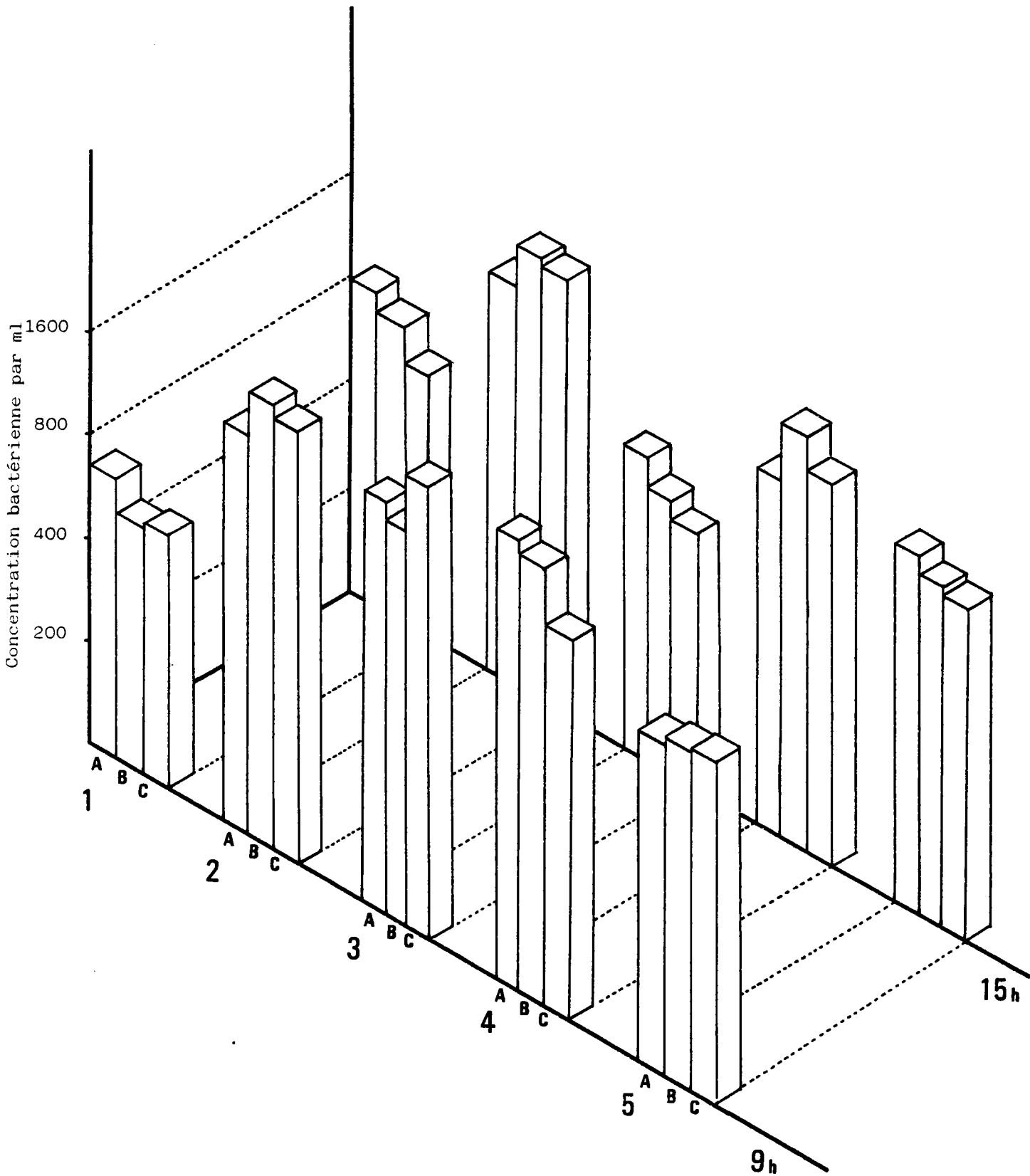
TABLEAU 16 : (suite et fin)

La figure 11 tente au moyen d'un diagramme en perspective cavalière de donner une vision globale de ces résultats.

Après transformation logarithmique des titres bactériens, l'étude statistique par analyse de variance des valeurs ainsi transformées a été menée de la manière suivante :

- dans un premier temps l'homogénéité des variances a été vérifiée par le test de COCHRAN puis le calcul des carrés moyens correspondant aux divers facteurs étudiés a été entrepris ; enfin les tests de signification se rapportant à ces facteurs contrôlés ont été exécutés dans un ordre imposé par le tableau d'analyse de variance.

**FIGURE 11** : Diagramme en perspective des titres bactériens obtenus au cours des 5 journées d'étude (numérotées chronologiquement de 1 à 5) selon l'heure et l'endroit du prélèvement.



\* Vérification de l'homogénéité des variances

La valeur de la fonction discriminante  $g = \frac{S_{\max}^2}{\sum_h S_h^2}$  calculée à partir des

estimations  $S_h^2$  consignées dans le tableau 16 conduit au rapport  $g = \frac{0,04090}{0,16609} \approx 0,2462$

Cette valeur étant inférieure à la limite  $g_{0,95} = 0,2929$  du domaine d'acceptation pour un risque de première espèce égal à 0,05, l'égalité des variances a été admise.

\* Analyse de variance

Le tableau 17 résume les valeurs et les quantités estimées par les carrés moyens correspondant aux différentes sources de variation ainsi que les résultats des tests de signification. A cet effet, les interactions puis les effets des facteurs isolés ont été étudiés successivement en comparant la valeur numérique de la fonction discriminante appropriée à la limite supérieure du domaine d'acceptation tirée de la table de SNEDECOR.

. L'examen du tableau 17 montre qu'une interaction de 2<sup>e</sup> ordre des facteurs A,B et C (respectivement jour, lieu et heure) n'est pas mise en évidence de façon significative ( $F_{ABC} < F_{0,95} (8;30)$  mais  $F_{ABC} > F_{0,90} (8;30)$ ). Cette interaction doit être attribuée aux variations interéchantillons et traduit par conséquent une hétérogénéité spatiale macroscopique de grande échelle dans la concentration bactérienne du cours d'eau. Le manque de puissance du test utilisé n'a pas permis de mettre en relief ce phénomène relativement discret (la composante de variance associée à l'interaction ABC vaut en effet 0,00277 et correspond à un coefficient de variation logarithmique égal à 12,1%).

Cette hétérogénéité qui a été mise en évidence au chapitre II ne peut cependant pas être ignorée.

La valeur inférieure à l'unité du rapport F calculé pour le facteur BC indique l'absence d'interaction heure-lieu.

Par contre l'interaction AC (jour-heure) est mise en évidence de façon significative ce qui dénote des fluctuations horaires variables d'un jour à l'autre. Par ailleurs, aucune différence significative entre les charges bactériennes obtenues le matin et l'après-midi n'ayant pu être mise en évidence (absence de l'effet C), l'effet heure se manifeste surtout par son interaction avec le jour.

**TABEAU 17** : Analyse de variance de l'expérience factorielle à 3 facteurs contrôlés réalisée sur les titres bactériens de l'eau de la Moselle obtenus au cours de la campagne de prélèvements du 18 mai au 5 juin 1981.

Source de variation	Somme de carrés $Q$	Degrés de liberté $\nu$	Carré moyen $q = Q / \nu$	Quantité estimée par $q$	F calculé	Seuil d'acceptation.	Signification
Facteur A jour	1,36988	4	0,34247	$\sigma_R^2 + 12\sigma_A^2$	61,82	$F_{0,995}(4;30) = 4,62$	**
Facteur B lieu	0,01320	2	0,00660	$\sigma_R^2 + 4\sigma_{AB}^2 + 20\sigma_B^2$	0,51		N.S.
Facteur C heure	0,09923	1	0,09923	$\sigma_R^2 + 6\sigma_{AC}^2 + 30\sigma_C^2$	0,86		N.S.
Interaction AB	0,10410	8	0,01301	$\sigma_R^2 + 4\sigma_{AB}^2$	2,35	$F_{0,95}(8;30) = 2,27$	*
Interaction AC	0,45974	4	0,11493	$\sigma_R^2 + 6\sigma_{AC}^2$	20,75	$F_{0,995}(4;30) = 4,62$	**
Interaction BC	0,00837	2	0,00418	$\sigma_R^2 + 2\sigma_{ABC}^2 + 10\sigma_{BC}^2$	0,38		N.S.
Interaction ABC	0,08867	8	0,01108	$\sigma_R^2 + 2\sigma_{ABC}^2$	2,00	$F_{0,95}(8;30) = 2,27$	N.S.
Résiduelle	0,16609	30	0,00554	$\sigma_R^2$			
Totale	2,30927	59					

N.S. : test non significatif

\* : significatif au seuil de 5%

\*\* : significatif au seuil de 0,5%

Le test relatif au facteur B (lieu) étant négatif, il n'apparaît pas d'hétérogénéité fixe et persistante propre à chacun des trois points de prélèvement. L'interprétation de l'interaction AB (jour-lieu), significative au seuil de 5% est plus délicate. Aucune explication tangible d'un tel phénomène ne pouvant être proposée, l'erreur de première espèce ne doit pas être écartée.

Enfin, l'existence de variations journalières en ce qui concerne l'état de pollution du cours d'eau ne fait aucun doute. Le rapport des moyennes journalières extrêmes est sensiblement égal à 3. Notons que des taux bactériens singulièrement faibles ont été observés à plusieurs reprises pendant les semaines qui suivirent cette campagne de prélèvements. Ainsi des expériences réalisées dans les mêmes conditions ont donné des moyennes journalières voisines de 65 germes par ml, ce qui porte le rapport précédent à 25.

. Dans le but d'améliorer la puissance des tests de signification, les fluctuations émanant de certains facteurs n'ayant pas été mis en évidence ont été jointes à celles des facteurs auxquelles elles peuvent s'ajouter d'après le plan d'expérience. Ainsi les sommes des carrés correspondant aux facteurs B, C et BC ont été additionnées respectivement à celles des facteurs AB, AC et ABC. Les degrés de liberté associés à ces groupements sont obtenus par la somme des degrés de liberté de leurs composants. Le tableau 18 résume l'analyse de variance ainsi simplifiée.

Le test de SNEDECOR effectué sur le rapport  $F_{B+AB}$  (l'effet isolé du facteur B (lieu) n'étant pas distingué de son interaction avec A (jour) ) n'est pas significatif. Ce résultat montre que si une hétérogénéité spatiale fixe et de caractère permanent existe dans la section étudiée du cours d'eau, celle-ci est en tout état de cause trop discrète pour que les trois points de prélèvement ne puissent être considérés comme équivalents à l'occasion d'expériences ultérieures.

Les autres résultats présentés sur le tableau 18 confirment ceux qui ont été établis lors de l'analyse de variance complète.

. De plus, le double titrage réalisé sur chaque flacon de 500 ml permet de contrôler l'existence d'une variation intraéchantillon.

La variance résiduelle  $\sigma^2_R$  mentionnée dans les tableaux 17 et 18 est en effet la somme de la variance propre à la méthode de titrage, notée  $\sigma^2$ ,  
N.P.P.

**TABLEAU 18** : Tableau simplifié d'analyse de variance de l'expérience factorielle à 3 facteurs contrôlés.

Source de variation	Somme de carrés Q	Degrés de liberté v	Carré moyen q = Q/v	Quantité estimée par q	F calculé	Seuil d'acceptation.	Signification
Facteur A JOUR	1,36988	4	0,34247	$\sigma_R^2 + 12 \sigma_A^2$	61,82	$F_{0,995} (4;30)$ = 4,62	**
Facteur B + Interaction AB	0,11730	10	0,01173	$\sigma_R^2 + 4(\sigma_{AB}^2 + \sigma_B^2)$	2,12	$F_{0,95} (10;30)$ = 2,16	N.S.
Facteur C + Interaction AC	0,55896	5	0,11179	$\sigma_R^2 + 6(\sigma_{AC}^2 + \sigma_C^2)$	20,19	$F_{0,995} (5 ;30)$ = 4,23	**
Interaction BC + Interaction ABC	0,09704	10	0,00970	$\sigma_R^2 + 2(\sigma_{ABC}^2 + \sigma_{BC}^2)$	1,75	$F_{0,95} (10;30)$ = 2,16	N.S.
Résiduelle	0,16609	30	0,00554	$\sigma_R^2$			
Totale	2,30927	59					

N.S. : test non significatif

\*\* : significatif au seuil de 0,5 %



C'est le cas notamment d'une estimation unique réitérée journallement sur une rivière dont la densité bactérienne est soumise à des variations cycliques se manifestant sur une durée de 24 heures.

Le choix irréfléchi de l'heure des prélèvements peut conduire à des estimations fortement biaisées et dénuées de toute valeur. Cette situation nécessite une connaissance plus profonde de la variabilité journalière de la concentration bactérienne du cours d'eau. Et ce n'est qu'après avoir étudié la structure de la périodicité, par des expérimentations préliminaires semblables à celle que nous avons menée au chapitre IV de la 2ème partie, qu'il devient généralement possible d'arrêter les modalités de l'échantillonnage systématique (COCHRAN, 1977). Le moment du jour se prêtant au mieux à l'échantillonnage devra autant que possible se distinguer par une certaine stabilité et correspondre à un niveau moyen de la pollution.

Une série peut présenter une évolution déterministe, une variabilité systématique et des fluctuations aléatoires. La figure 15 illustre à l'aide d'un exemple fictif ces trois composantes éventuelles d'une série de données.

La composante déterministe interprétée comme la tendance peut constituer un phénomène d'intérêt écologique.

La composante systématique (oscillation) correspond à une variabilité périodique à signification écologique dont l'étude est entravée par des fluctuations irrégulières et imprévisibles qualifiées de bruit aléatoire.

## II - 2-2 Analyse d'une série.

Sans aborder l'étude exhaustive des méthodes d'analyse des séries dont l'exposé, donné par LEGENDRE et LEGENDRE (1979), dépasse nettement le cadre de ce travail, il convient néanmoins d'en dégager le cheminement.

Lorsque la nature purement aléatoire de la série a été écartée, l'étude de la variabilité systématique ne peut être entreprise qu'après l'extraction d'une tendance éventuelle (composante de basse fréquence). La mise en évidence de la composante systématique (par l'autocorrélation, l'analyse harmonique ou l'analyse spectrale), sur la série stationnaire ainsi obtenue, dépend essentiellement de l'amplitude des oscillations stationnaires par rapport à l'intensité du bruit aléatoire (composante de haute fréquence).

La variabilité systématique doit être suffisamment importante pour ne pas se confondre avec le bruit aléatoire.

et de la composante de variance, notée  $\sigma^2_{\text{intra}}$ , caractérisant les fluctuations à l'intérieur des échantillons.

Il s'ensuit qu'une éventuelle hétérogénéité intraéchantillon peut être mise en évidence par le test de SNEDECOR appliqué au rapport

$$F_{\text{intra}} = \frac{\sigma_R^2}{\sigma_{\text{N.P.P.}}^2};$$

en remplaçant  $\sigma_{\text{N.P.P.}}^2$  par l'estimation qui en a été donnée au paragraphe II-3-1 de la première partie avec 40 degrés de liberté, il vient :

$$F_{\text{intra}} = \frac{0,00554}{0,00294} \approx 1,88.$$

Cette valeur étant supérieure à la limite donnée par la table de SNEDECOR au niveau de probabilité 0,95 ( $F_{0,95}(30; 40) = 1,74$ ), il apparaît que la concentration bactérienne n'est pas uniforme à l'intérieur des flacons de 500 ml. La composante de variance  $\sigma^2_{\text{intra}}$  est estimée par  $S^2_{\text{intra}} = 0,00260$  ce qui correspond à un coefficient de variation logarithmique de 11,7 %.

### III - 4 DISCUSSION.

Dans cette expérimentation la mise en évidence ainsi que l'étude succincte de divers aspects de l'hétérogénéité dans l'espace et dans le temps d'un milieu hydrique particulier ont été menées simultanément.

L'absence d'hétérogénéité fixe et permanente perceptible nous amène à considérer les trois points de prélèvement comme étant équivalents. Ce résultat n'est pas surprenant compte tenu qu'aucun apport d'eau ne se produit le long de la section considérée du cours d'eau et que celle-ci est trop courte pour que le phénomène d'autoépuration puisse être évoqué.

Les autres manifestations de l'hétérogénéité, tant spatiale que temporelle, sont à l'origine de fluctuations dans les résultats et peuvent de ce fait constituer une source d'incertitude.

Le graphique à secteurs circulaires représenté par la figure 12 tente de donner une vision globale de l'importance relative de plusieurs facteurs de variation. Les secteurs angulaires sont proportionnels aux composantes de variance caractérisant les variations propres à la technique de titrage, les variations intra et interéchantillon ainsi que les fluctuations horaires et journalières. Les estimations de ces composantes ont été calculées à partir des données des tableaux 17 et 18 et sont exprimées en pourcentage de leur somme totale.

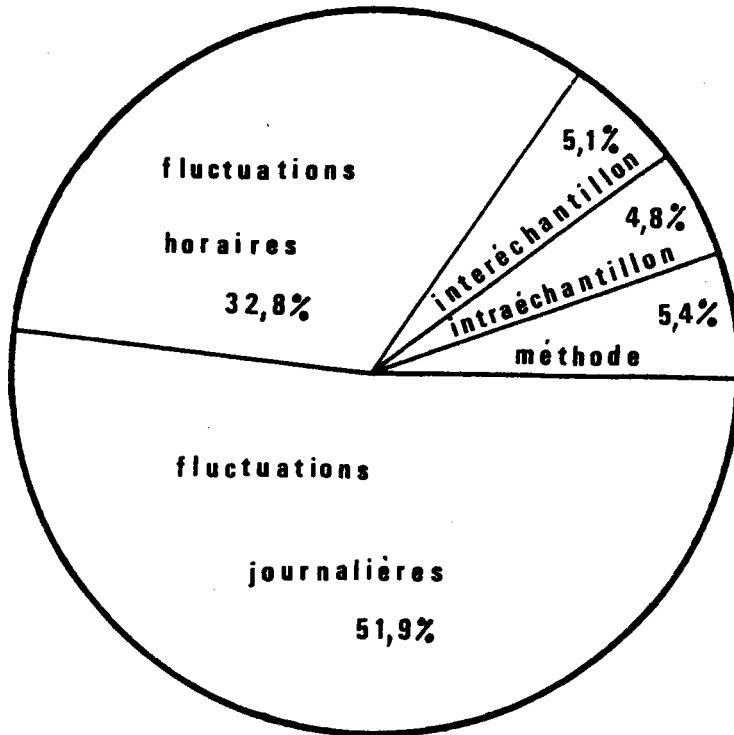


Figure 12 : Comparaison de quelques sources de variation.

Il importe de souligner le caractère très approximatif d'une telle comparaison car les valeurs retenues ne doivent en aucun cas être considérées comme des constantes caractéristiques du milieu hydrique étudié. Certaines d'entre elles sont très probablement sujettes à des distorsions très importantes (c'est le cas notamment des variations horaires et journalières). Par ailleurs, la composante relative à l'effet heure n'a qu'une portée très restreinte puisqu'elle ne s'applique qu'aux deux moments de la journée retenus pour l'expérience.

Malgré ces réserves il apparaît nettement que les fluctuations dans le temps sont plus marquées que les trois autres causes de variation lesquelles ont des effets comparables.

La précision relativement bonne (13%) exprimée par la composante de variance propre à la technique de titrage, notée  $\sigma_{N.P.P.}^2$ , de même que son caractère invariant et incoercible ont grandement facilité l'exploitation statistique des résultats dont les conclusions viennent d'être énoncées. Ces deux caractéristiques constituent l'intérêt essentiel de la méthode du nombre le plus probable avec 96 inoculums par dilution.

Rappelons que l'hétéroscédasticité des variances est incompatible avec toute analyse sérieuse de la variance. De plus, une précision insuffisante de la méthode de titrage s'oppose à la mise en évidence des autres sources de variation en diminuant la puissance des tests de signification. A titre d'exemple, lorsque 5 tubes seulement sont inoculés à chaque niveau de dilution, la variance  $\sigma_{N.P.P.}^2$  (0,0605) est supérieure à la somme de toutes les composantes de variance représentées sur la figure 12 (elle représenterait alors 54 % de la variance totale). Cette dernière remarque montre que les études microbiologiques réalisées dans ces conditions sont vouées à un manque certain de finesse et de prétention.

La méthode de titrage utilisée dans ce travail, avec un protocole analytique approprié, permet d'envisager des études structurées et conçues dans le but de répondre à des objectifs plus ambitieux.

#### CHAPITRE IV : ETUDE DE LA VARIATION HORAIRE DE LA CONCENTRATION BACTERIENNE D'UN MILIEU HYDRIQUE.

Après l'étude générale du chapitre précédent dont nous retiendrons plus particulièrement la prépondérance des variations temporelles, une étude plus fine visant à apprécier l'ampleur des fluctuations de la concentration bactérienne de l'eau, au cours d'une même journée, a été entreprise sur le même site. Afin d'évaluer l'impact de la station d'épuration biologique située en amont, les résultats obtenus ont été rapprochés de la charge bactérienne des effluents de la station.

L'expérience a tenté également de montrer qu'il est possible de prétendre à un niveau appréciable de précision et de justesse dans les estimations en adoptant un protocole analytique approprié dont les modalités seront précisées ultérieurement. Les développements théoriques reposant pour une large part sur les composantes de variance  $\sigma_{N.P.P.}^2$ ,  $\sigma_{intra}^2$  et  $\sigma_{inter}^2$ , le plan d'expérience a été conçu de manière à permettre une nouvelle estimation de ces composantes (composantes de la variance analytique).

#### IV - 1 MATERIEL ET METHODES

Les échantillons d'eau ont été recueillis en surface, à 6 mètres environ de la berge, à l'endroit représenté sur la figure 9 et désigné par la lettre C.

Le matériel utilisé pour les prélèvements a été décrit au paragraphe III-1 de cette deuxième partie.

Les titrages bactériens ont été réalisés par la méthode du N.P.P. sur plaques de microtitration conformément à la description donnée dans la première partie au paragraphe II-1-3.

L'application du test de COCHRAN, décrit au paragraphe II-1 de la présente partie, a permis la vérification de l'homoscédasticité des variances avant d'entreprendre le traitement statistique des résultats par une analyse de la variance.

Enfin les intervalles de confiance à 90 % des composantes de variance  $\sigma_{intra}^2$  et  $\sigma_{inter}^2$  ont été déterminés d'après la méthode proposée par MORIGUTI (SNEDECOR et COCHRAN, 1957) en faisant l'hypothèse de normalité.

IV - 2      PROTOCOLE EXPERIMENTAL

Trois échantillons de 500 ml d'eau ont été prélevés au même endroit de la rivière dans un intervalle de temps ne dépassant jamais 3 minutes (temps nécessaire au remplissage successif des trois flacons). Cette opération commencée à 8 heures le 16 juillet 1981 a été réitérée toutes les 2 heures jusqu'à 20 heures.

Chacun des 21 échantillons ainsi recueillis au cours de cette journée a été analysé comme suit.

Deux sous-échantillons de 10 ml chacun ont été prélevés distinctement dans le flacon de 500 ml. Une homogénéisation de l'échantillon initial par 10 retournements a précédé chaque prélèvement de liquide. Un millilitre de chaque sous-échantillon a ensuite été introduit dans 9 ml d'eau physiologique stérile puis une suspension à la dilution  $10^{-2}$  a été préparée en suivant le même procédé. Chacune des deux séries de dilutions ainsi obtenues a permis la réalisation d'un titrage par la méthode du N.P.P. en milieu liquide : la première plaque ayant étéensemencée par l'aliquote d'eau non diluée.

La période s'écoulant entre la collecte des 3 échantillons et la fin des titrages n'a jamais dépassé 2 heures.

Par ailleurs, dans le but d'évaluer l'importance de l'apport bactérien émanant des rejets de la station d'épuration biologique située en amont du point C, un prélèvement de 500 ml effectué sur les effluents de la station a précédé de 10 minutes environ chaque échantillonnage dans le milieu naturel. Trois plaques de microtitration ont étéensemencées par les dilutions  $10^{-2}$ ,  $10^{-3}$  et  $10^{-4}$  préparées à partir de 1 ml de l'échantillon initial.

Après lecture et calcul des résultats comme décrit au paragraphe II-1-3 de la première partie, deux estimations indépendantes de la densité bactérienne ont été obtenues pour chaque flacon.

La signification des différences entre les concentrations bactériennes de l'eau de rivière pour chaque prélèvement a été testée par une analyse de la variance des logarithmes décimaux des titres bactériens estimés.

Auparavant, l'égalité des variances a été vérifiée par le test de COCHRAN.

Le facteur heure, noté A, intervient avec 7 modalités fixes alors que les facteurs échantillon et sous-échantillon, notés B et C, interviennent respectivement avec 3 et 2 modalités aléatoires.

Chacune des modalités k de C est associée à une modalité définie j du facteur B et chacune des modalités j est elle-même associée à une modalité définie i du facteur A.

Le modèle adapté à cette expérience est par conséquent un plan pyramidal à 3 facteurs contrôlés sans répétitions.

Les notations utilisées dans le tableau 19, représentant le plan d'expérience, respectent les conventions d'écriture utilisées au paragraphe II-2 de la première partie.

**TABLEAU 19** : Tableau d'analyse de variance.

SOURCE DE VARIATION	SOMME DE CARRES Q	DEGRES DE LIBERTE	QUANTITE ESTIMEE PAR Q/ν
FACTEUR A Heure	$Q_A = 6 \sum_i (\bar{x}_{i..} - \bar{x})^2$	6	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2 + 2\sigma_{inter}^2 + 6\sigma_A^2$
FACTEUR B échantillon	$Q_B = 2 \sum_i \sum_j (\bar{x}_{ij.} - \bar{x}_{i..})^2$	14	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2 + 2\sigma_{inter}^2$
FACTEUR C sous-échantillon	$Q_C = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij.})^2$	21	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2$
TOTALE	$Q_T = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x})^2$	41	

IV - 3 RESULTATS.

Les résultats sont présentés en deux parties :

- la première met en relief les fluctuations de la concentration bactérienne de la rivière sur une période de 12 heures et donne une estimation des composantes de la variance analytique.
- La deuxième partie montre les variations du flux bactérien des rejets de la station d'épuration au cours de la même période.

IV - 3-1 Variation de la concentration bactérienne de la Moselle sur une période de 12 heures.

Trois flacons de 500 ml d'eau de la Moselle ont été prélevés toutes les 2 heures sur une période de 12 heures. Les valeurs des titres bactériens correspondant à chaque échantillon sont rapportées dans le tableau 20.

Les logarithmes de ces valeurs ont été soumis à l'analyse de variance après vérification de l'homogénéité des variances par le test de COCHRAN.

\* Vérification de l'homoscédasticité des variances.

La valeur de la fonction discriminante  $g = \frac{S_{\max}^2}{\sum_h S_h^2}$  calculée à partir des estimations  $S_h^2$  consignées dans le tableau 20 est égale à  $g = \frac{0,04898}{0,11846} \approx 0,4135$

La limite supérieure du domaine d'acceptation pour un risque de première espèce fixé à 0,01 valant approximativement 0,4661, l'égalité des variances a été admise.

\* Analyse de variance.

Les valeurs des carrés moyens correspondant aux différents facteurs contrôlés, les quantités estimées par celles-ci ainsi que les résultats des tests de signification sont résumés dans le tableau 21.

L'examen de ce tableau montre que les effets des facteurs C (sous-échantillon) et B (échantillon) sont significatifs au seuil de 5%. Ces résultats traduisent d'une part la présence d'une hétérogénéité à l'intérieur des flacons de 500 ml et d'autre part des différences entre les concentrations bactériennes des échantillons recueillis successivement au même moment de la journée.

Les données du tableau 21 ont rendu possible l'estimation des trois composantes de la variance analytique, c'est à dire  $\sigma_{N.P.P.}^2$ ,  $\sigma_{\text{intra}}^2$  et  $\sigma_{\text{inter}}^2$ . Ces estimations accompagnées d'un intervalle de confiance à 90 % (calculé selon la méthode proposée par MORIGUTI en ce qui concerne  $\sigma_{\text{intra}}^2$  et  $\sigma_{\text{inter}}^2$ ) sont présentées dans le tableau 22.

De plus, le test se rapportant au facteur A (heure des prélèvements) met en évidence une différence très significative entre les sept valeurs de la concentration bactérienne de la rivière observées au cours de l'expérience.



heure	n° de l'échantillon	Triplet	N.P.P.	Concentration bactérienne par ml (u)	log u	Variance empirique $S_h^2$	Moyenne géométrique (/ ml)
8	1	96-59-9	9,59	192	2,283	0,02290	346
		96-76-14	15,70	314	2,497		
	2	96-79-16	17,49	350	2,544	0,00218	
		96-81-23	20,38	408	2,610		
	3	96-81-21	19,83	397	2,598	0,00583	
		96-87-25	25,40	508	2,706		
10	1	96-85-30	25,44	509	2,707	0,00289	425
		96-84-20	21,39	428	2,631		
	2	96-79-21	18,75	375	2,574	0,00432	
		96-74-15	15,13	303	2,481		
	3	96-91-17	26,12	522	2,718	0,00186	
		96-85-22	22,72	454	2,657		
12	1	96-82-21	20,42	408	2,611	0,00080	374
		96-85-21	22,39	448	2,651		
	2	96-83-16	19,60	392	2,593	0,00320	
		96-79-11	16,29	326	2,513		
	3	96-74-12	14,49	290	2,462	0,01008	
		96-80-24	20,08	402	2,604		
14	1	96-96-45	64,66	1293	3,112	0,00061	986
		96-96-48	70,22	1404	3,147		
	2	96-93-26	33,35	667	2,824	0,00396	
		96-93-37	40,91	818	2,913		
	3	96-95-37	47,56	951	2,978	0,00005	
		96-95-38	48,66	973	2,988		
16	1	96-92-19	28,16	563	2,751	0,00151	588
		96-95-18	32,02	640	2,806		
	2	96-89-24	26,96	539	2,732	0,00020	
		96-87-26	25,77	515	2,712		
	3	96-85-21	22,39	448	2,651	0,04898	
		96-93-43	46,02	920	2,964		

**TABLEAU 20** : Titres bactériens de l'eau de la Moselle obtenus entre 8h et 20h.

Heure	n° de l'échantillon.	Triplet	N.P.P.	Concentration bactérienne par ml (u)	log u	Variance empirique $S_h^2$	Moyenne géométrique (/ ml)
18	1	96-96-73	142,89	2858	3,456	0,00344	2675
		96-96-79	173,11	3462	3,539		
	2	96-96-72	138,63	2773	3,443	0,00034	
		96-96-70	130,63	2613	3,417		
	3	96-96-67	119,72	2394	3,379	0,00120	
		96-96-63	106,83	2137	3,330		
20	1	96-85-17	21,14	423	2,626	0,00378	366
		96-76-21	17,30	346	2,539		
	2	96-78-24	18,99	380	2,580	0,00016	
		96-78-21	18,24	365	2,562		
	3	96-78-18	17,51	350	2,544	0,00014	
		96-77-17	16,81	336	2,527		

TABLEAU 20 : Titres bactériens de l'eau de la Moselle obtenus entre 8 h et 20h.  
(Suite et fin)

**TABEAU 21** : Analyse de variance du plan pyramidal à 3 facteurs contrôlés réalisée sur les titres bactériens de l'eau de Moselle obtenus au cours du bilan journalier.

Source de variation	Somme de carrés Q	degrés de liberté $\nu$	Carré moyen $q = Q/\nu$	Quantité estimée par q	F calculé	Seuil d'acceptation	signification.
Facteur A heure	3,81436	6	0,63573	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2 + 2\sigma_{inter}^2 + 6\sigma_A^2$	41,77	$F_{0,995} (6;14) = 5,26$	* *
Facteur B échantillon	0,21309	14	0,01522	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2 + 2\sigma_{inter}^2$	2,70	$F_{0,95} (14;21) = 2;20$	*
Facteur C sous-échantillon	0,11846	21	0,00564	$\sigma_{N.P.P.}^2 + \sigma_{intra}^2$	1,92	$F_{0,95} (21;40) = 1,83$	* (1)
Totale	4,14591	41					

\* : test significatif au seuil de 5%

\*\* : test significatif au seuil de 0,5%

(1) Le carré moyen associé au facteur C est comparé à l'estimation de  $\sigma_{N.P.P.}^2$  donnée dans la première partie avec 40 degrés de liberté.

TABLEAU 22 : Estimation des composantes de variance

Composante de variance	Valeur estimée	Intervalle de confiance à 90%
$\sigma^2$ N.P.P.	0,00294 (12,5%)	0,00211 (10,6%) - 0,00444 (15,4%)
$\sigma^2$ intra	0,00270 (12 %)	0,00019 (3,2%) - 0,00734 (19,9%)
$\sigma^2$ inter	0,00479 (16 %)	0,00098 (7,2%) - 0,01344 (27,1%)

Les nombres entre parenthèses représentent les coefficients de variation logarithmique associés aux différentes estimations.

Une représentation graphique des fluctuations de la qualité bactériologique de l'eau pendant la journée du 16 juillet 1981 est donnée par la figure 13.

Le logarithme décimal de la densité bactérienne moyenne de la rivière à un moment donné ainsi que les limites d'un intervalle de confiance à 95% de cette estimation ont été calculés par la formule :

$$\bar{X}_{i..} \pm \sqrt{\frac{q_B}{pq}} \cdot t_{0,975}$$

où  $q_B$  désigne le carré moyen relatif au facteur échantillon (B)

( $q_B = 0,01522$  d'après le tableau 21) ;

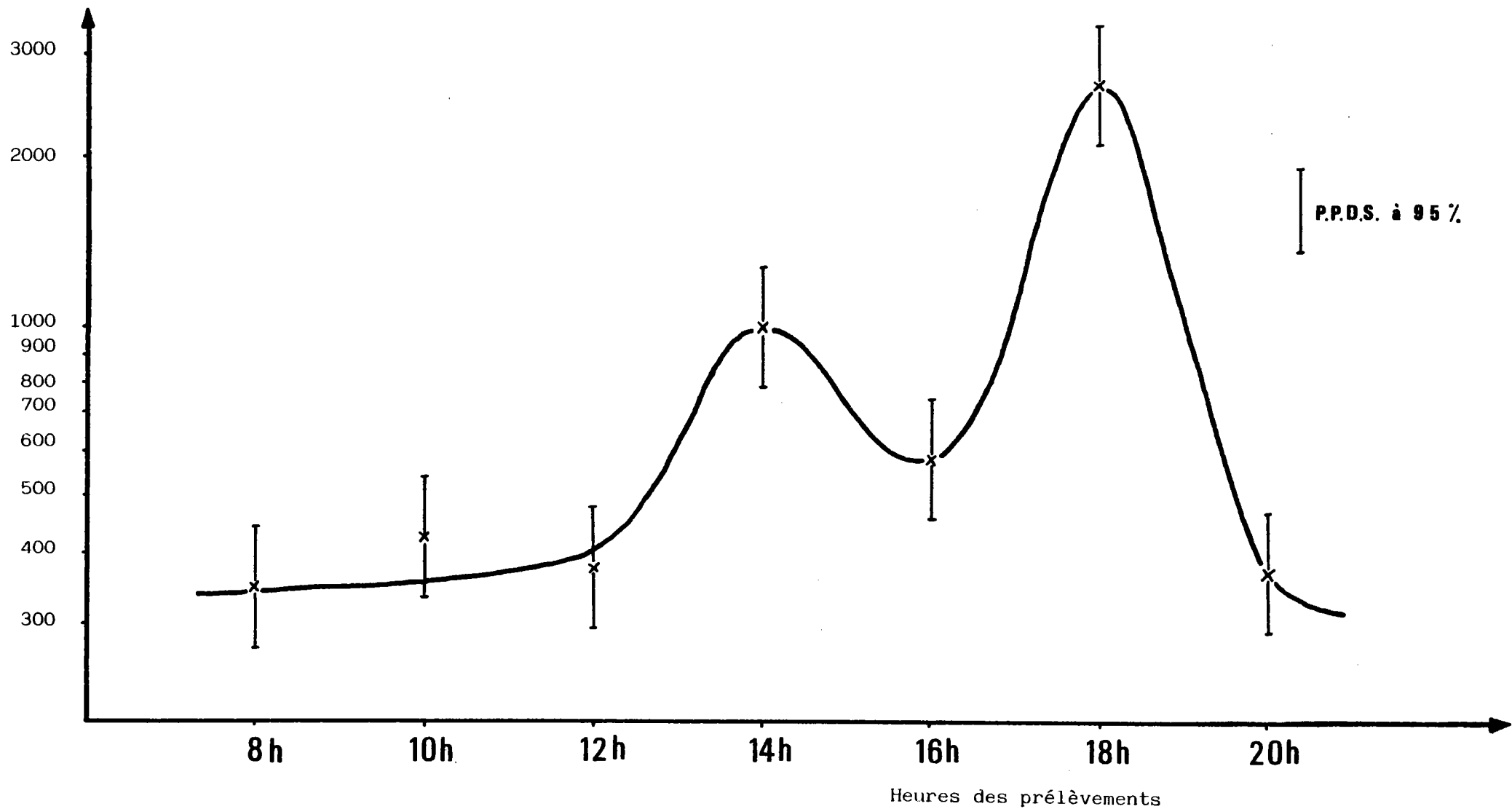
$\bar{X}_{i..}$  est la moyenne des  $pq$  estimations disponibles pour une même heure ( $p$  échantillons et  $q$  titrages par échantillon), ce produit est ici égal à 6 ;

enfin  $t_{0,975}$  représente la variable de Student au seuil de probabilité 0,975 déterminée avec 14 degrés de liberté.

En conservant les mêmes notations, la plus petite différence significative (P.P.D.S. au seuil de probabilité 0,95) entre deux moyennes particulières est obtenue par la

relation :

$$P.P.D.S = \sqrt{\frac{2q_B}{pq}} \cdot t_{0,975}$$



**FIGURE 13** : Variation des concentrations bactériennes de l'eau de Moselle enregistrées au cours de la journée du 16 juillet 1981.

La figure 13 fait apparaître une concentration bactérienne relativement constante entre 8h et 12h au cours de la journée d'expérimentation : les titres bactériens observés à 8h, 10h et 12h ne diffèrent pas de façon significative. Deux pointes de contamination bactérienne ont été mises en évidence à 14 h puis à 18h. Si la première est plutôt discrète, le deuxième pic est par contre remarquable par son ampleur. A 20h la rivière retrouvait un état de pollution comparable à celui de la matinée.

#### IV - 3-2 Variation du flux des coliformes thermorésistants dans les rejets de la station d'épuration biologique de METZ.

L'enregistrement en continu du débit de l'effluent de la station d'épuration biologique permet l'étude précise de la variation du débit instantané exprimé en litres/seconde.

La variation de ce débit pour le jour de l'expérimentation est représentée sur la figure 14.

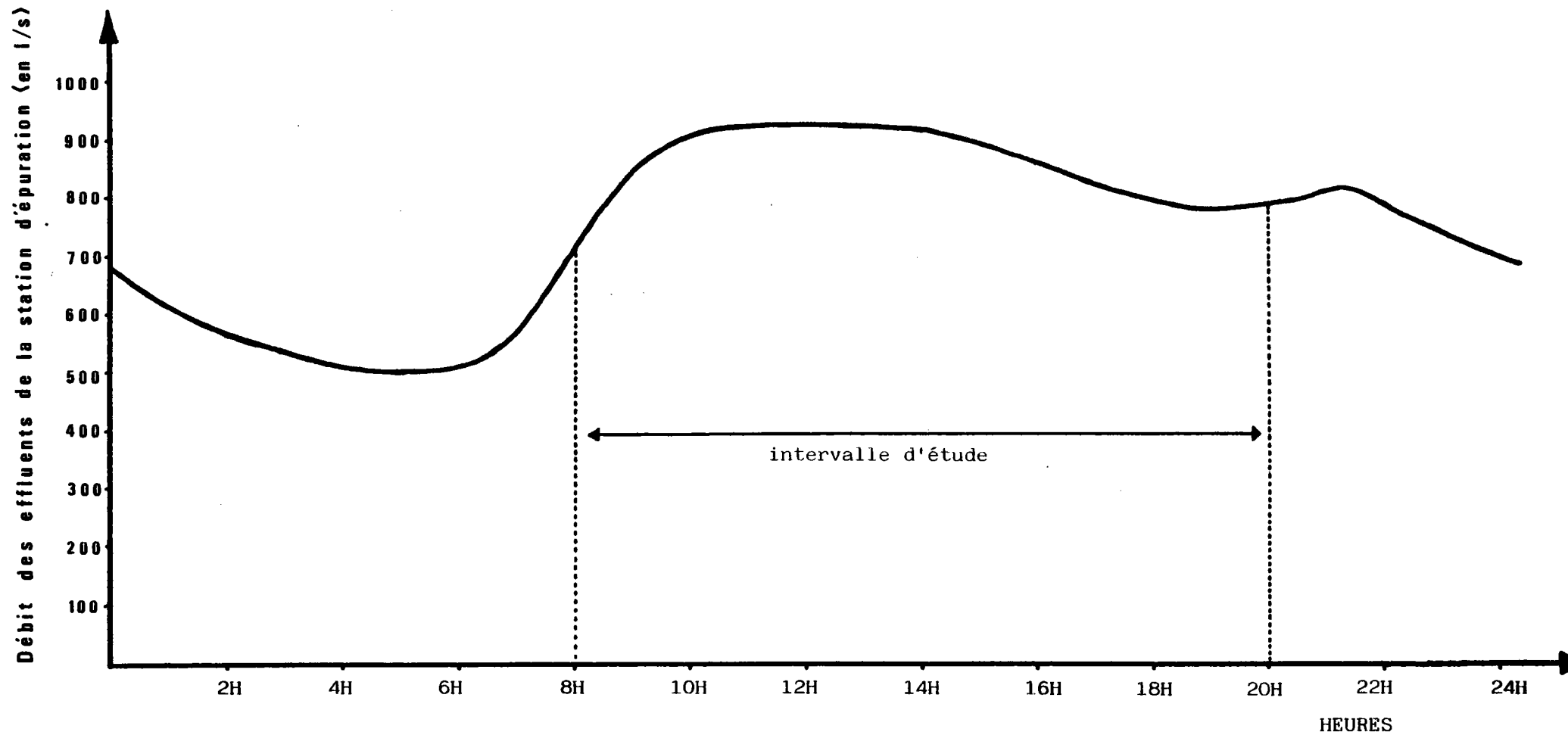
Etant à son niveau le plus bas de 500l/s vers 5h le matin, le débit augmente progressivement pour atteindre sa valeur maximale de 900l/s vers 10h. Après une phase en plateau se prolongeant jusqu'à 14h environ, le débit amorce une diminution très lente interrompue toutefois par une légère reprise vers 21h avant de retrouver le niveau minimal de 500 l/s vers 5h au jour J+1.

Les résultats des analyses bactériologiques effectuées sur les eaux traitées de la station sont présentés dans le tableau 23.

Etant donné que les estimations de la concentration bactérienne ne s'appuient que sur un titrage unique réalisé à partir d'1 ml d'effluent, les intervalles de confiance mentionnés dans le tableau 23 ne s'appliquent qu'à l'échantillon initial de 1 ml.

Le protocole analytique ne permettant pas d'intégrer l'hétérogénéité du milieu des eaux traitées dans ces estimations, toute extrapolation de ces résultats à des volumes plus importants de liquide est impossible.

Malgré cette restriction, qui s'oppose à toute comparaison statistique rigoureuse, les concentrations en coliformes thermorésistants des rejets de la station d'épuration biologique dans le milieu naturel ne semblent pas présenter des différences très marquées pendant la durée de l'expérience. Le flux de pollution bactérienne de l'effluent, calculé à partir des débits instantanés et des résultats des analyses bactériologiques, varie approximativement entre  $0,6 \times 10^{10}$  et  $2,2 \times 10^{10}$  germes par seconde au cours de l'expérimentation.



**FIGURE 14** : Variation du débit des effluents de la station d'épuration biologique de la ville de METZ (16 juillet 1981).

**TABLEAU 23** : Concentration bactérienne des effluents de la station d'épuration biologique de la ville de METZ pendant la journée du 16 juillet 1981.

heure du prélèvement	Triplet associé aux dilutions $10^{-2}$ - $10^{-3}$ - $10^{-4}$	N.P.P.	Concentration bactérienne par ml	limites d'un interval- le de confiance à 95 %	
8h	95 - 60 - 20	10,05	20100	16040	25180
10h	96 - 59 - 11	9,90	19800	15550	25210
12h	96 - 46 - 11	7,25	14500	11230	18730
14h	95 - 46 - 8	6,15	12300	9560	15830
16h	96 - 37 - 7	5,44	10880	8360	14150
18h	95 - 42 - 8	5,63	11260	8740	14510

#### IV - 4 DISCUSSION

Par cette expérience, nous avons voulu :

- a) étudier les variations de la concentration bactérienne de la rivière pendant une période de 12h.
- b) apprécier la part que tient dans ces variations le flux de pollution émanant des rejets de la station d'épuration par voie biologique située en amont.
- c) reconsidérer le problème de l'hétérogénéité du milieu hydrique par une nouvelle estimation des composantes de la variance analytique.

##### a) Fluctuations de la densité bactérienne dans le milieu naturel.

La plus petite différence significative (P.P.D.S.) entre deux estimations indépendantes de la densité bactérienne de la rivière (chaque estimation étant la moyenne de tous les titres bactériens obtenus selon un protocole analytique donné) dépend de la variance analytique, du nombre de degrés de liberté de son estimation



et du risque d'erreur de première espèce accepté.

Pour un risque d'erreur fixé à 5% la P.P.D.S. vaut 0,15279. Cette différence entre deux logarithmes correspond en fait à un écart de 42% entre la valeur la plus forte et la valeur la plus faible.

Le niveau relativement élevé du pouvoir séparateur de la méthode analytique utilisée, qui peut s'exprimer par la P.P.D.S., a permis une bonne connaissance des fluctuations horaires de la densité bactérienne du milieu étudié. Des variations à la fois très subites et très importantes ont été mises en évidence. Les concentrations extrêmes observées à 8h (346 germes/ml) et à 18h (2675 germes/ml) se situent dans un rapport de 1 à 8 environ.

L'ampleur et la soudaineté des pics de contamination bactérienne montrent combien il est illusoire de prétendre obtenir des résultats représentatifs de l'état de pollution chronique d'une rivière à partir de prélèvements ponctuels lorsque les fluctuations de la densité bactérienne, au cours du temps sont ignorées.

b) Fluctuations de la densité bactérienne des rejets de la station d'épuration biologique.

Les analyses bactériologiques de même que le débit des eaux usées traitées rejetées par la station d'épuration biologique se distinguent par une certaine stabilité (variation d'un facteur 2).

Des résultats sensiblement différents ont été obtenus par ROLLAND (1981) qui relate des variations de la concentration en coliformes thermorésistants dans l'effluent de la station d'épuration biologique de NANCY allant de 400 germes/ml à 10h jusqu'à 21200 germes/ml à 18h pendant la même journée (variation d'un facteur 50).

Les faibles variations ainsi que l'importance relativement limitée du flux de contamination fécale représenté par les rejets de la station de METZ, par rapport au flux bactérien de la rivière, ne sauraient, en tout état de cause, expliquer les pics de contamination observés dans le milieu naturel à 600 m en aval des installations.

c) Les composantes de la variance analytique.

Le double titrage réalisé pour chaque échantillon de 500 ml a été pratiqué dans le seul but de permettre l'estimation des composantes de la variance analytique,  $\sigma^2_{intra}$  et  $\sigma^2_{inter}$ , dues à l'hétérogénéité intra et interéchantillon.

Les estimations de ces composantes sont très comparables aux valeurs obtenues au cours des expériences antérieures. En particulier le coefficient de variation (12%) exprimant la variabilité entre les sous-échantillons de 10 ml prélevés à l'intérieur d'un même flacon n'est pas significativement différent de la valeur obtenue au chapitre III (11,7%) pour des sous-échantillons de 1 ml.

La composante interéchantillon (16%), mise en évidence avec un risque inférieur à 5%, montre que des échantillons prélevés successivement au même endroit de la rivière ne peuvent pas être considérés comme équivalents.

Ce résultat soulève une nouvelle fois le problème de l'extrapolation des titres bactériens provenant d'un échantillon unique.

Malgré la concordance des estimations, faites à plusieurs reprises dans ce travail, des composantes de la variance analytique, rien n'autorise à les considérer comme des constantes caractéristiques du milieu hydrique étudié.

L'inhérence de ces composantes à la configuration de l'hétérogénéité leur confère très probablement une certaine variabilité pour un milieu donné et a fortiori lorsque les milieux hydriques sont différents.

TROISIEME PARTIE

PROPOSITION THEORIQUE D'UNE STRATEGIE  
DE PRELEVEMENT

CHAPITRE I :       DEFINITION D'UN PROTOCOLE ANALYTIQUE POUR UNE ESTIMATION PONCTUELLE  
DE LA CONCENTRATION BACTERIENNE D'UN MILIEU HYDRIQUE.

I - 1       INTERET DE L'ECHANTILLONNAGE A TROIS DEGRES.

L'étendue des milieux hydriques naturels contient les microbiologistes à estimer la concentration bactérienne de ces eaux à partir de titrages réalisés sur des échantillons dont la taille est comparativement beaucoup plus faible. Le titrage est quelquefois effectué sur un seul sous-échantillon issu d'un échantillon unique prélevé dans le milieu naturel étudié. Ce sous-échantillon, dont le volume ne dépasse généralement pas quelques millilitres, est alors supposé être représentatif d'une quantité d'eau considérable (plusieurs centaines de mètres cubes ou davantage). Malheureusement l'hétérogénéité des milieux n'autorise pas l'extrapolation du titre bactérien obtenu et ceci quelle que soit la précision de la technique de numération utilisée. Cette restriction s'oppose à la comparaison rigoureuse de plusieurs prélèvements et, de ce fait, rend délicate toute étude sérieuse de la qualité bactériologique d'une eau. Le problème se pose de manière analogue lorsque plusieurs sous-échantillons, prélevés successivement à l'intérieur d'un même flacon, font chacun l'objet d'un titrage distinct (échantillonnage à deux degrés). En effet, quel que soit le nombre de ces titrages, la moyenne des résultats obtenus ne peut s'appliquer qu'à l'échantillon initial et son extrapolation se heurte cette fois à l'hétérogénéité interéchantillon. Ces quelques considérations montrent que pour augmenter la portée des estimations il est nécessaire d'y intégrer les effets de l'hétérogénéité du milieu hydrique.

Dans ce but nous préconisons un échantillonnage à trois degrés comprenant plusieurs échantillons (unités primaires) dans lesquels seront prélevés des sous-échantillons (unités secondaires), le troisième degré étant constitué par le titrage bactérien de ces unités secondaires.

Les variations observées à chaque niveau de l'échantillonnage se traduisent par une incertitude sur le résultat qui tient compte à la fois de l'hétérogénéité du milieu et de la reproductibilité de la méthode de numération.

Bien qu'il ne soit pas possible de définir avec précision les limites de l'espace hydrique pris en considération, la pluralité des échantillons recueillis dans un volume d'eau relativement important contribue à donner plus de poids à l'estimation ponctuelle de la concentration bactérienne d'une rivière.

Rappelons que si l'échantillonnage comporte  $r$  répétitions (indices N.P.P.) pour chacun des  $q$  sous-échantillons prélevés dans  $p$  échantillons, la densité bactérienne est estimée par la moyenne géométrique des  $p q r$  résultats obtenus.

La variance,  $\sigma_{\bar{x}}^2$ , de la moyenne arithmétique,  $\bar{x}$ , des logarithmes des titres bactériens est donnée par la formule (5):

$$(5) \sigma_{\bar{x}}^2 = \frac{\sigma_{\text{inter}}^2}{p} + \frac{\sigma_{\text{intra}}^2}{pq} + \frac{\sigma_{\text{N.P.P.}}^2}{pqr} \quad (\text{SNEDECOR et COCHRAN, 1957})$$

où  $\sigma_{\text{inter}}^2$ ,  $\sigma_{\text{intra}}^2$  et  $\sigma_{\text{N.P.P.}}^2$  désignent les composantes de la variance analytique correspondant à chacun des trois niveaux de l'échantillonnage.

L'examen de la formule (5) montre qu'une diminution de la valeur de  $\sigma_{\bar{x}}^2$ , et par suite une précision accrue dans l'estimation du titre bactérien moyen, peut être obtenue

a) en réduisant la valeur des composantes de variance : une amélioration de la technique de titrage entraînerait la diminution de  $\sigma_{\text{N.P.P.}}^2$ , et une meilleure homogénéisation des échantillons aurait pour effet d'atténuer l'importance de  $\sigma_{\text{intra}}^2$ . La variance  $\sigma_{\text{inter}}^2$ , directement liée à l'hétérogénéité du milieu, présente par contre un caractère incoercible pour des conditions de prélèvement données.

b) en augmentant les nombres  $p$ ,  $q$  et  $r$  avec toutefois une préférence pour les unités primaires (échantillons) ; ainsi lorsque le produit  $p q r$  est fixé,  $\sigma_{\bar{x}}^2$  est minimal quand  $q$  et  $r$  sont tous deux égaux à l'unité.

En supposant constantes les valeurs des composantes de variance, il devient possible de comparer l'efficacité relative de plusieurs protocoles analytiques fictifs. Le tableau 24 tente de mettre en relief l'incidence du choix d'une combinaison des nombres  $p$ ,  $q$  et  $r$  sur la qualité de l'estimation.

Les calculs, pour l'estimation  $S_{\bar{x}}^2$  de la variance  $\sigma_{\bar{x}}^2$ , ont été effectués en remplaçant dans la formule (5) les composantes  $\sigma_{\text{inter}}^2$ ,  $\sigma_{\text{intra}}^2$  et  $\sigma_{\text{N.P.P.}}^2$  par les estimations obtenues au chapitre IV de la deuxième partie.

Les trois premières lignes du tableau 24 font ressortir l'avantage de faire figurer les répétitions le plus possible en amont du plan expérimental pour un nombre total de titrages donné ( $p q r = 4$  dans le cas étudié).

**TABLEAU 24** : Comparaison de quelques protocoles analytiques fictifs.

nombre d'échantillons p	nombre de sous-échantillons q	nombre de titrages r	Variance $S_x^2$	Coefficient de variation en%	(a) Intervalle de confiance à 95%
4	1	1	0,00261	11,8	79,4 - 125,9
2	2	1	0,00381	14,3	75,7 - 132,1
1	1	4	0,00823	21,1	66,4 - 150,6
1	1	1	0,01043	23,8	63,1 - 158,6
1	1	(b) ∞	0,00749	20,1	67,7 - 147,8
(c) 1	1	1	0,00294	12,5	78,3 - 127,7

(a) : le titre bactérien moyen est supposé être égal à 100.

(b) : le symbole "∞" indique qu'aucune incertitude n'est imputable à la technique de titrage.

(c) : cette estimation repose sur l'hypothèse d'un milieu parfaitement homogène.

Les conséquences de l'imperfection de la méthode de titrage sur l'incertitude du résultat peuvent s'apprécier par la confrontation des estimations données à la quatrième et la cinquième ligne du tableau.

Le faible gain de précision procuré par une méthode de titrage supposée parfaite (symbolisée par ∞) laisse entrevoir les limites d'une recherche inconsidérée visant à améliorer indéfiniment la précision des techniques de numération lorsque parallèlement l'hétérogénéité du milieu est négligée. Cette hétérogénéité des milieux hydriques naturels induit, en effet, sur les résultats une incertitude inexorable et cela indépendamment de la reproductibilité de la méthode de titrage utilisée.

Le dernier cas de figure mentionné dans le tableau fait état d'un titrage unique avec l'hypothèse d'une parfaite homogénéité du milieu. Il est intéressant de noter que pour atteindre la précision indiquée (12,5%), il est en fait nécessaire d'effectuer quatre titrages sur quatre échantillons distincts (première ligne du tableau).

L'optimisation du protocole analytique, c'est-à-dire le choix de la meilleure combinaison des nombres p, q et r, nécessite une formulation exacte des coûts comparés des différents plans expérimentaux possibles. En effet, si pour des raisons purement mathématiques les répétitions doivent de préférence être le privilège des seules unités primaires (échantillons), la solution préconisée n'est pas forcément la plus économique.

Une valeur approximative du coût (C) de l'analyse est donnée par la formule suivante

$$C = C_1 p + C_2 pq + C_3 p q r$$

dans laquelle les coefficients  $C_i$  représentent le coût d'une unité supplémentaire au niveau correspondant de l'échantillonnage et cela indépendamment des autres niveaux.

En considérant simultanément le coût (C) et la variance ( $S_x^2$ ), le plan expérimental le plus efficace est celui qui minimise le produit  $S_x^2.C$

$$(6) \quad S_x^2.C = \left( \frac{S_{inter}^2}{p} + \frac{S_{intra}^2}{pq} + \frac{S_{N.P.P.}^2}{pqr} \right) (C_1 p + C_2 pq + C_3 p q r)$$

$S_{inter}^2$ ,  $S_{intra}^2$  et  $S_{N.P.P.}^2$  désignent respectivement les estimations des composantes de variance  $\sigma_{inter}^2$ ,  $\sigma_{intra}^2$  et  $\sigma_{N.P.P.}^2$ .

Les solutions de ce problème sont données par les relations (7) et (8)

$$(7) \quad r = \sqrt{\frac{C_2 \cdot S_{N.P.P.}^2}{C_3 \cdot S_{intra}^2}} \quad \text{et} \quad (8) \quad q = \sqrt{\frac{C_1 \cdot S_{intra}^2}{C_2 \cdot S_{inter}^2}}$$

alors que p est obtenu en résolvant l'équation (6) soit en fonction du coût soit en fonction de la variance, selon que le coût ou la variance est imposé (SNEDECOR et COCHRAN, 1957).

En introduisant les valeurs numériques des composantes de variance, obtenues au chapitre IV de la deuxième partie, dans les relations (7) et (8) il vient :

$$r = \sqrt{\frac{0,00294 C_2}{0,00270 C_3}} = 1,04 \sqrt{\frac{C_2}{C_3}} \quad \text{et} \quad q = \sqrt{\frac{0,00270 C_1}{0,00479 C_2}} = 0,75 \sqrt{\frac{C_1}{C_2}}$$

Le choix de r et de q se portera sur les nombres entiers approchant au mieux les valeurs ainsi calculées.

Le coût d'un sous-échantillon ( $C_2$ ) n'étant pas supérieur à celui d'un titrage ( $C_3$ ), il ne serait pas économique de faire plus d'un titrage par sous-échantillon. De plus, sans pour autant donner avec précision les coûts  $C_1$  et  $C_2$ , le rapport  $\frac{C_1}{C_2}$  ne saurait justifier une valeur de q supérieure à l'unité.

Ainsi, des considérations d'ordre économique montrent qu'il n'y a pas lieu de modifier la combinaison qui, de prime abord, s'avérait la plus efficace, à savoir la collecte de plusieurs échantillons de 500 ml faisant chacun l'objet d'un titrage unique.

Cette conclusion, établie sur des hypothèses particulières notamment en ce qui concerne l'incertitude sur le résultat provenant des différents niveaux de l'échantillonnage, devrait cependant jouir d'une certaine généralité si l'on en juge par l'ordre de grandeur relatif des composantes de variance observées à plusieurs reprises dans ce travail.

Reste à déterminer le nombre  $p$  des échantillons qu'il convient de recueillir dans le milieu naturel pour permettre une estimation ponctuelle dans le temps et dans l'espace de la concentration bactérienne d'un cours d'eau. Il va de soi que  $p$  dépend non seulement de la précision souhaitée pour le résultat mais également de contraintes économiques. Si le nombre trois constitue, à juste titre, un minimum, il ne nous paraît cependant pas nécessaire de prélever plus de six échantillons compte tenu des exigences habituelles des études microbiologiques.

Lorsque aucune information fiable n'est disponible sur la variance propre du protocole analytique adopté, l'estimation ne doit être fondée que sur les seules données expérimentales. Ainsi en faisant comme à l'accoutumée l'hypothèse de normalité, le logarithme de la densité bactérienne conformément au protocole optimal susvisé est estimé par :

$$\bar{X} \pm \frac{S}{\sqrt{p}} \cdot t_{1-\alpha/2}$$

$p$  : nombre d'échantillons analysés

$\bar{X}$  : moyenne arithmétique des logarithmes des  $p$  titres bactériens obtenus

$S$  : écart-type empirique

$t_{1-\alpha/2}$  : variable de STUDENT au niveau de probabilité  $1-\alpha$  avec  $p-1$  degrés de liberté.





## CHAPITRE II : APPLICATION DE LA METHODE D'ECHANTILLONNAGE A L'ETUDE ET AU CONTROLE DE LA CONTAMINATION BACTERIENNE D'UN COURS D'EAU.

En mettant l'accent sur la portée accrue des estimations obtenues à partir d'un échantillonnage à trois degrés, l'intérêt d'une telle démarche a été mis en relief dans le premier chapitre.

Toute estimation ponctuelle étant réalisée sur la base d'un protocole analytique optimisé s'appuyant sur la pluralité des échantillons recueillis, nous tenterons dans ce dernier chapitre de faire ressortir les avantages de la méthode proposée pour ce qui concerne les études microbiologiques de plus grande envergure et de caractère moins restrictif.

Nous envisagerons dans un premier temps l'étude comparative de l'état de pollution d'un cours d'eau en plusieurs endroits différents puis l'étude des séries temporelles de données obtenues pour une même station de prélèvement. Ces deux cas de figure particuliers ne sont nullement exhaustifs, d'autres modèles pouvant être imaginés afin d'apporter une réponse à des questions dont la formulation ne saurait cependant comporter la moindre équivoque. La clarté, de même que la pertinence des questions motivant toute étude microbiologique, constituent une condition sine qua non à l'élaboration d'un plan d'expérience adapté au problème à résoudre.

Parmi les autres modèles possibles nous citerons sans les traiter les études d'impact. Les microbiologistes peuvent être amenés, par exemple, à juger du degré de toxicité d'un effluent en comparant les densités des populations bactériennes observées en amont et en aval de l'émissaire. Le principe général de la méthodologie permettant cette évaluation, compte tenu de l'hétérogénéité du milieu, est donné par GREEN (1979). Le plan d'expérience préconisé par GREEN trouve avantageusement son application en microbiologie en tirant bénéfice de notre procédé d'échantillonnage.

### II -1 ETUDE COMPARATIVE DE LA CONTAMINATION BACTERIENNE CHRONIQUE ENTRE PLUSIEURS STATIONS DE PRELEVEMENT.

La localisation des sources de pollution ainsi que les caractéristiques hydrographiques d'un cours d'eau sont à l'origine d'une hétérogénéité spatiale fixe et persistante de la concentration bactérienne laquelle est estompée par le phénomène d'autoépuration.

L'étude globale d'une section déterminée et suffisamment longue d'un cours d'eau passe par la connaissance de ce gradient de contamination bactérienne, et cela indépendamment des variations cycliques évoquées au chapitre précédent. Nous proposons à cet effet un plan d'expérience mettant à profit le protocole analytique établi au chapitre I de cette troisième partie.

La comparaison de l'état de pollution chronique entre 1 stations de prélèvement requiert plusieurs expérimentations réparties au hasard sur une période donnée. Chacune des t expérimentations ainsi programmées devra prévoir la réalisation d'une estimation ponctuelle pour chaque station de prélèvement. Ces estimations seront effectuées simultanément selon le protocole analytique précité.

Le plan d'expérience adapté à cette étude est un plan factoriel comprenant deux facteurs contrôlés avec répétitions. Le tableau 25 donne une représentation de ce modèle d'analyse de variance en respectant les conventions d'écriture qui ont été faites au paragraphe II-2 de la première partie.

Le facteur contrôlé A (lieu de prélèvement) intervient avec 1 modalités fixes tandis que le facteur B (expérimentation) possède t modalités aléatoires. De plus, r résultats indépendants (r échantillons faisant l'objet d'un titrage unique) sont disponibles pour chaque combinaison  $A_i B_j$  de ces modalités, c'est-à-dire pour chaque estimation ponctuelle.

Le modèle d'analyse de variance adopté permet de vérifier séparément l'influence de chacun des facteurs contrôlés A et B. En outre, l'interaction éventuelle de ces deux facteurs peut être distinguée de l'effet des facteurs non contrôlés dont la variance, notée  $\sigma_R^2$ , correspond à la variance analytique.

**TABLEAU 25** : Tableau d'analyse de variance du plan factoriel à deux facteurs permettant l'étude comparative de la contamination bactérienne entre 1 stations de prélèvement.

Source de variation	Somme de carrés Q	Degrés de liberté $\nu$	Quantité estimée par $q/\nu$
Facteur A lieu	$Q_A = rt \sum_i (\bar{x}_{i.} - \bar{x})^2$	1 - 1	$\sigma_R^2 + r \sigma_{AB}^2 + rt \sigma_A^2$
Facteur B expérimentation	$Q_B = rl \sum_j (\bar{x}_{.j} - \bar{x})^2$	t - 1	$\sigma_R^2 + rl \sigma_B^2$
Interaction AB	$Q_{AB} = r \sum_j \sum_i (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$	(1-1) (t-1)	$\sigma_R^2 + r \sigma_{AB}^2$
Résiduelle	$Q_R = \sum_i \sum_j \sum_{\alpha} (x_{ij\alpha} - \bar{x}_{ij})^2$	1 t(r-1)	$\sigma_R^2$
Totale	$Q_T = \sum_i \sum_j \sum_{\alpha} (x_{ij\alpha} - \bar{x})^2$	ltr - 1	

La mise en évidence de l'interaction AB s'explique par la présence de différences significatives parmi les écarts observés d'une expérimentation à l'autre entre les 1 stations de prélèvement. Cette possibilité, qui exclut un comportement simplement additif des facteurs A et B, ne saurait, selon toute vraisemblance, être écartée.

Le risque de première espèce étant préalablement fixé, l'influence du facteur A (lieu du prélèvement) est alors vérifiée en comparant la valeur du rapport  $F_A = \frac{q_A}{q_{AB}} = 1 + \frac{rt \sigma_A^2}{\sigma_R^2 + r \sigma_{AB}^2}$  à la limite supérieure du domaine d'acceptation donné par la loi de SNEDECOR avec (1-1) et (1-1) (t-1) degrés de liberté.

Le risque  $\beta$  de conclure à tort à l'absence d'effet du facteur A (risque de deuxième espèce) est une fonction décroissante des nombres de degrés de liberté (1-1) et (1-1) (t-1) ainsi que de l'expression  $1 + \frac{rt \sigma_A^2}{\sigma_R^2 + r \sigma_{AB}^2}$  (C.E.A., 1978).

Le test apparaît donc d'autant plus conservateur que  $\sigma_R^2 + r \sigma_{AB}^2$  est plus grand devant  $rt \sigma_A^2$ .

C'est précisément en réduisant la quantité  $\sigma_R^2 + r \sigma_{AB}^2$  que la méthode analytique proposée apporte une amélioration appréciable. En effet, la précision accrue de la technique de titrage étudiée dans ce travail contribue à diminuer la valeur de  $\sigma_R^2$  alors que parallèlement la pluralité des échantillons recueillis, en autorisant l'extrapolation des estimations, a pour effet de situer  $\sigma_{AB}^2$  à un juste niveau. Des titrages répétés sur un échantillon unique, tout en donnant à  $\sigma_R^2$  une valeur plus faible provoquée par la perte de la composante  $\sigma_{inter}^2$ , entraîneraient par contre le report inexorable de celle-ci sur  $\sigma_{AB}^2$ . La quantité  $\sigma_R^2 + r \sigma_{AB}^2$  s'en trouverait finalement augmentée et la variance  $\sigma_{AB}^2$  perdrait par là même une partie de sa signification.

Notons, par ailleurs, qu'une amélioration de la puissance du test susvisé peut être obtenue par une augmentation du nombre de degrés de liberté et plus particulièrement du nombre t d'expérimentations. D'une manière générale le choix des paramètres t et r (nombre d'échantillons) dépend d'une part de la connaissance a priori des composantes de variance  $\sigma_{AB}^2$  et  $\sigma_R^2$  et d'autre part de l'importance des écarts que l'on souhaite pouvoir mettre en évidence entre les stations compte tenu de l'investissement prévu.

## II - 2 LES SERIES TEMPORELLES

### II - 2-1 Caractéristiques des séries

Une série temporelle est une suite d'observations ordonnées dans le temps obéissant à des éléments déterministes et systématiques ainsi qu'à des éléments aléatoires.

Les méthodes d'analyse numérique des séries ont pour objectif l'extraction des composantes déterministes et systématiques en faisant cas du contexte probabiliste lié à l'existence de phénomènes aléatoires.

L'intérêt suscité par l'étude des séries en microbiologie tient non seulement à l'importance des rythmes endogènes aux écosystèmes mais également des rythmes biogéographiques en rapport avec l'activité humaine.

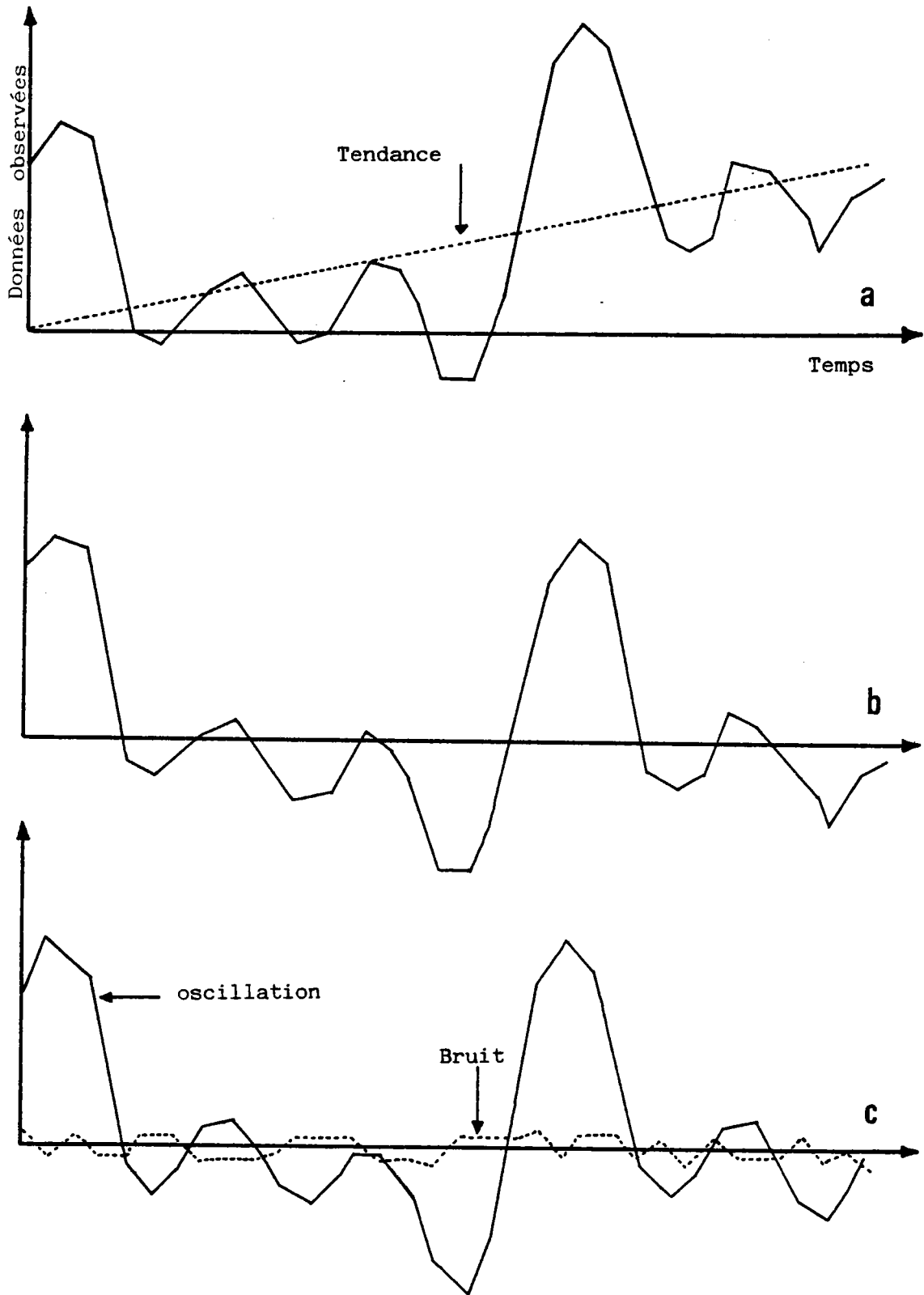
L'analyse des séries peut par conséquent apporter au microbiologiste une meilleure compréhension des phénomènes qui le préoccupent.

Une série peut, par exemple, être conçue dans le but d'évaluer les fluctuations horaires de la concentration bactérienne au cours d'une même journée (voir chapitre IV de la deuxième partie) ou à plus long terme de permettre l'étude des variations journalières, mensuelles ou saisonnières de la pollution d'origine bactérienne d'une rivière.

Afin de respecter les conditions imposées par l'analyse des données et dans le souci de traduire au mieux la variabilité temporelle des paramètres observés, l'échantillonnage devra être mené de façon rigoureuse et sensée.

Ainsi le choix du point de prélèvement, dans la mesure du possible, devra tenir compte des caractéristiques hydrographiques du cours d'eau afin d'être représentatif de la section principale du courant (VELZ, 1970). Ceci exclut les prélèvements effectués trop près de la berge surtout lorsque l'eau y est peu profonde et stagnante.

Par ailleurs, les méthodes d'analyse des séries exigent un échantillonnage systématique c'est-à-dire dont les prélèvements sont séparés par des intervalles de temps égaux. Cette forme d'échantillonnage présente un inconvénient majeur lorsque le paramètre observé présente une variation périodique dont la période ou l'un de ses multiples coïncide avec l'intervalle de temps compris entre deux estimations ponctuelles consécutives.



**FIGURE 15** : Série fictive de données (a), dont est extraite une tendance linéaire (b) et qui est constituée (c) d'une oscillation systématique, à laquelle se superpose un "bruit" aléatoire. (D'après LEGENDRE et LEGENDRE, 1979).

Dans le cas particulier des études microbiologiques le bruit est surtout imputable au manque général de précision des méthodes de numération sans oublier la nature aléatoire de la reproductibilité des méthodes de type énumératif. Les effets de l'hétérogénéité du milieu ne font, par ailleurs, qu'accentuer l'incertitude des résultats.

La méthode analytique proposée au chapitre I de la présente partie, qui, il faut le rappeler, consiste à titrer plusieurs échantillons par la méthode du N.P.P. avec 96 inoculums par dilution selon un protocole optimisé, a pour effet d'atténuer le phénomène de bruit pour les raisons suivantes :

- . la technique de titrage utilisée se distingue par une bonne précision intrinsèque (13%) laquelle est de surcroît indépendante de la densité bactérienne.
- . en tenant compte des effets de l'hétérogénéité du milieu par l'analyse simultanée de plusieurs échantillons, les estimations reflètent, de toute évidence, plus étroitement l'état de pollution réel du cours d'eau. En d'autres termes, la possibilité d'extrapoler les résultats donne plus de poids aux estimations ponctuelles.

Si de plus, l'homogénéité des variances des résultats associés à une même estimation ponctuelle a pu être vérifiée au moyen d'un test approprié (par exemple le test de BARTLETT ou le test de COCHRAN), une estimation de la variance analytique commune peut être calculée à partir de l'ensemble des observations. Ce procédé, qui a déjà été utilisé au chapitre IV de la 2ème partie, présente l'avantage de permettre le calcul d'un intervalle de confiance plus resserré, par suite du nombre accru de degrés de liberté, et dont l'amplitude sur une échelle logarithmique est identique pour chaque estimation ponctuelle.

La stabilisation ainsi que l'atténuation de la composante aléatoire qui découlent des remarques précédentes, permettent d'envisager l'étude des séries microbiologiques sur des bases plus prometteuses en favorisant la mise en exergue des oscillations systématiques.

C O N C L U S I O N

Avant de pouvoir établir les fondements d'une stratégie de prélèvement dans un milieu hydrique nous avons été amenés :

- d'une part, à étudier une méthode de titrage adaptée aux techniques statistiques mises en oeuvre dans ce travail,
- d'autre part, à relativiser les différents aspects de l'hétérogénéité spatio-temporelle de la concentration bactérienne d'un cours d'eau. La connaissance globale et précise des caractéristiques de cette hétérogénéité étant nécessaire à la mise au point d'un protocole analytique optimal ainsi qu'à la formulation des objectifs et des modalités de toute étude microbiologique.

\* De l'étude théorique de la méthode de titrage nous avons pu dégager les propriétés suivantes :

- le logarithme de l'indice N.P.P. est normalement distribué,
- sa variance, qui est surtout fonction du nombre d'inoculums par dilution, peut être considérée comme indépendante de la densité bactérienne.

La validité de ces deux propositions suppose toutefois l'appartenance de la concentration bactérienne à un intervalle dont l'amplitude croît avec le nombre d'inoculums par dilution. Ainsi, lorsque 96 cupules sont inoculées à 3 niveaux de dilution consécutifs, les bornes inférieure et supérieure de cet intervalle sont associées respectivement à un pourcentage de réponses positives compris entre 22% et 93%.

Il s'ensuit que les statistiques paramétriques et plus particulièrement l'analyse de variance peuvent être appliquées aux résultats dans une large plage de densité, après une transformation logarithmique préalable des titres bactériens.

\* L'étude expérimentale de la méthode de titrage a montré :

- une concordance parfaite de la précision de la méthode, calculée à partir de la variance empirique, avec la précision théorique.  
Les coefficients de variation observés avec 96 puis 32 inoculums par dilution sont respectivement de 12,5 % et 21,5 %.
- que l'erreur introduite par le manipulateur, et plus généralement l'ensemble des facteurs non contrôlés impliqués dans la réalisation des titrages, exercent sur les résultats un effet négligeable.



\* Lors de la comparaison expérimentale des deux techniques de titrage (méthode quantique en milieu liquide avec 96 inoculums par dilution et méthode énumérative sur milieu gélosé), nous avons pu établir que :

- le modèle de POISSON est adapté à la description de la répartition des unités formatrices de colonies dans les petits volumes d'eau homogénéisés.
- la précision propre à la microméthode est équivalente à celle qu'on peut attendre, en moyenne, d'un titrage réalisé sur milieu gélosé.
- en milieu liquide, la proportion de réponses positives imputables à la croissance de Escherichia coli s'élève à 82,5% environ. Mais cette estimation, qui ne constitue en fait qu'une moyenne, est variable selon l'échantillon considéré.
- les titres bactériens donnés par la méthode étudiée ont tendance à être légèrement supérieurs aux résultats obtenus par la méthode des membranes filtrantes. La surestimation, voisine de 14 %, n'est pas indépendante de l'échantillon.
- l'existence d'une interaction très significative entre chacun des deux paramètres précédents et l'échantillon analysé nous a amené à considérer la sélectivité préférentielle des méthodes de titrage et plus précisément le manque de spécificité de la microméthode comme étant la cause essentielle de la surestimation observée en milieu liquide.

Le haut niveau de précision de même que les propriétés mathématiques de l'indice N.P.P. ont notablement favorisé la réalisation d'études visant à procurer une meilleure appréhension des caractéristiques de l'hétérogénéité spatio-temporelle d'un milieu hydrique.

\* En ce qui concerne l'étude de cette hétérogénéité, les expérimentations réalisées sur la Moselle ont permis de mettre en évidence :

- de fortes variations journalières et horaires de la concentration bactérienne, les dernières étant remarquables tant par leur soudaineté que par leur ampleur,
- des différences significatives entre les titres bactériens provenant de sous-échantillons distincts issus d'un même échantillon de 500 ml. Le coefficient de variation caractérisant cette hétérogénéité intraéchantillon a été évalué à 12 % environ,
- des variations de la densité bactérienne entre les échantillons recueillis simultanément en un même endroit de la rivière. Cette hétérogénéité interéchantillon se traduit par un coefficient de variation estimé à 16 %.

Indépendamment de la précision de la technique de titrage, l'hétérogénéité du milieu hydrique s'oppose à l'extrapolation des résultats provenant d'un échantillon unique. Aussi, dans le but d'augmenter la portée des estimations ponctuelles en y intégrant les effets de l'hétérogénéité du milieu, nous préconisons un échantillonnage à trois degrés basé sur la pluralité des échantillons analysés. Compte tenu de nos résultats expérimentaux, le protocole analytique optimal consiste à réaliser un titrage unique par échantillon ; le nombre de ceux-ci étant lié à la précision recherchée. Bien que les composantes de la variance analytique soient probablement sujettes à des distorsions résultant de la diversité des configurations possibles de l'hétérogénéité, nous sommes en droit de penser que la démarche analytique proposée devrait bénéficier d'une certaine généralité en s'appliquant à d'autres milieux hydriques.

Pour ce qui concerne la surveillance de la contamination bactérienne, l'absence d'hétérogénéité fixe, perceptible entre des points situés le long d'une portion limitée du cours d'eau ne recevant pas d'émissaire, nous suggère, eu égard au rôle prépondérant des variations temporelles, un échantillonnage méthodique et intensif en un nombre de points aussi restreint que possible.

Au terme de cette étude menée conjointement au laboratoire et dans le milieu naturel, nous pouvons affirmer que la méthode analytique proposée (échantillonnage à trois degrés selon un protocole optimisé et méthode de titrage utilisant 96 inoculums par dilution) constitue un outil particulièrement intéressant pour les investigations microbiologiques en potamologie.

REFERENCES BIBLIOGRAPHIQUES

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.

BREITMAYER J.P. et GAUTHIER M.J. (1979)

Dénombrement des bactéries en milieu marin :  
facteurs de variation et d'incertitude.

Ann. Microbiol. (Inst. Pasteur) 130A, 245-256.

COCHRAN W.G. (1950)

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

Biometrics 6, 105 - 109.

COCHRAN W.G. (1977)

Sampling techniques (3<sup>e</sup> Edition)

Wiley and Sons Publishers, New York.

COMMISSARIAT A L'ENERGIE ATOMIQUE (1978)

Statistique appliquée à l'exploitation des mesures.

Tome 1 et 2 Masson édit., Paris.

COSTERTON J.W. and GEESEY G.G. (1979)

Which populations of aquatic bacteria should we enumerate ?

In Costerton J.W. and Colwell R.R., édit.,

Native aquatic bacteria : enumeration, activity, and ecology, ASTM STP 695,  
7 - 18.

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.

EVANS T.M., WAARVICK C.E., SEIDLER R.J. and LE CHEVALLIER M.W. (1981a)

Failure of the most-probable number technique to detect coliforms in drinking water and raw water supplies.

Appl. Environ. Microbiol. 41, 130-138.

EVANS T.M., LE CHEVALLIER M.W., WAARVICK C.E. and SEIDLER R.J. (1981b)  
Coliform species recovered from untreated surface water and drinking water  
by the membrane filter, standard, and modified most-probable-number techniques.  
Appl. Environ. Microbiol. 41, 657-663.

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. (1917)

Cité par HALVORSON H.O. et ZIEGLER N.R. (1933a).

HALDANE J.B.S. (1939)

Sampling errors in the determination of bacterial or virus density by the di-  
lution method.

J. Hyg. 39, 289-293.

HALVORSON H.O. and ZIEGLER N.R. (1933a)

Application of statistics to problems in bacteriology.

I - A means of determining bacterial population by the dilution method.

J. Bacteriol. 25, 101-121.

HALVORSON H.O. and ZIEGLER N.R. (1933b)

Application of statistics to problems in bacteriology.

III- A consideration of the accuracy of dilution data obtained by using se-  
veral dilutions.

J. Bacteriol. 26, 559-567.

HICKS C.R. (1973)

Fundamental concepts in the design of experiments (2e Edition)

Holt, Rinehart and Winston Publishers, New York.

HUGUES B. (1981)

Nouvelle utilisation de la méthode du nombre le plus probable en virologie. Application à la mise en évidence et à la quantification des virus dans le milieu hydrique.

Thèse de Doctorat d'Etat es Sciences , Université de Metz.

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.

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.

LEGENBRE L. et LEGENBRE P. (1979)

Ecologie numérique. Tomes 1 et 2

Masson, Paris.

LELLOUCH J. et LAZAR P. (1974)

Méthodes statistiques en expérimentation biologique.

Flammarion Médecine-Sciences, Paris.

LEWIS W.M. (1978)

Comparison of temporal and spatial variation in the zooplankton of a lake by means of variance components.

Ecology, 59, 666-671.

LURIA H., (1953)

Cité par DOUGHERTY R.M. (1964)

Animal virus titration techniques, in Techniques in experimental virology.

HARRIS R.C. éd., Academic Press Publishers, Londres, 169-223.

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

Cité par VELZ C.J. (1970).

Mc CRADY M.H. (1915)

Cité par HALVORSON H.O. et ZIEGLER N.R. (1933a).

MATHER K. (1965)

Analyse statistique en biologie.

Gauthier-Villars, Paris.

MAUL A. (1981)

Etude des méthodes de quantification en microbiologie.

D.E.A. Ecotoxicologie, Université de Metz.

MEYNELL G.G. and MEYNELL E. (1970)

Theory and practice in experimental bacteriology (2 e Edition)

University Press Publishers, Cambridge.

OSTLE B. and MENSING R.W. (1975)

Statistics in research (3e Edition)

Iowa State University Press, Ames.

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.

PHELPS E.B. (1908)

Cité par HALVORSON H.O. et ZIEGLER N.R. (1933a).

PLATT T., DICKIE L.M. and TRITES R.W. (1970)

Spatial heterogeneity of phytoplankton in a near-shore environment.

J. Fish. Res. Board Can. 27, 1453-1473.

POOLE R.W. (1974)

An introduction to quantitative ecology.

Mc Graw-Hill Publishers, New-York.

ROLLAND D. (1981)

Flux de virus entériques dans les eaux usées brutes et traitées d'une station d'épuration biologique.

Thèse de 3è cycle d'Ecologie, Université de Metz.

SEKI H. (1968)

Some characteristics of marine microorganisms in the pacific central water.

Records of Oceanographic Works in Japan 9, 247-255.

SEKI H. (1971)

Microbial clumps in sea water in the euphotic zone of Saanich Inlet (British Columbia).

Marine Biology 9, 4-8.

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

Statistical methods (6e Edition)

Iowa State University Press, Ames, Iowa.

SOROKIN, Yu.I. (1970)

On the degree of aggregation of marine bacterio-plankton.

Doklady Akademii Nauk S.S.S.R. 192, 905-907.

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

Estimation of coliform density by the membrane filter and the fermentation tube methods.

J. Publ. Health. 45, 1431-1437.

VELZ J.C. (1970)

Applied stream sanitation.

Wiley Interscience Publishers, New York.

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. (U.K.) 50, 859-870.

WYSHAK G. and DETRE K. (1972)

Estimating the number of organisms in quantal assays.

Appl. Microbiol. 23, 784-790.



**A N N E X E S**

ANNEXE I

LA FONCTION DE VRAISEMBLANCE

La probabilité, pour un microorganisme donné de se trouver dans un inoculum de volume q, s'exprime par le rapport  $\frac{q}{V}$  où V représente le volume de la suspension à titrer. La probabilité pour qu'il en soit absent est par conséquent égale à  $1 - \frac{q}{V}$ .

Si b représente le nombre total de microorganismes contenus dans l'échantillon, la probabilité pour qu'un inoculum soit stérile est :  $(1 - \frac{q}{V})^b = p$  (probabilité d'événements indépendants).

Le quotient  $\frac{q}{V}$  étant généralement petit,  $1 - \frac{q}{V}$  peut être approché par  $e^{-\frac{q}{V}}$  ce qui donne :

$$p = (1 - \frac{q}{V})^b \approx e^{-\frac{bq}{V}}$$

Comme  $\frac{b}{V}$  désigne la densité, u, de microorganismes par unité de volume, l'expression de p se simplifie, et finalement, la probabilité pour qu'un inoculum soit stérile est égale à  $p = e^{-uq}$ .

Considérons n prélèvements de même volume q ; la variable aléatoire, notée X, associée au nombre d'inoculums stériles, suit la loi binomiale de paramètres (n,p).

D'où  $P(X=s) = \binom{n}{s} p^s (1-p)^{n-s}$

Comme  $p = e^{-uq}$ ,  $P(X=s) = \binom{n}{s} e^{-uqs} (1 - e^{-uq})^{n-s}$

La probabilité d'obtenir respectivement  $s_1, s_2, \dots, s_i, \dots, s_k$  cupules stériles dans k niveaux de dilution successifs est donnée par :

$$L(u) = \prod_{i=1}^k \binom{n_i}{s_i} e^{-uq_i s_i} (1 - e^{-uq_i})^{n_i - s_i}$$

L(u) s'obtient en faisant le produit des probabilités d'événements indépendants.

$n_i$  : nombre de cupules inoculées à la  $i^{i\text{ème}}$  dilution.

$q_i$  : volume d'inoculum ensemencé dans chacune des  $n_i$  cupules.

Si  $p_i = 1 - e^{-uq_i}$  est le nombre de réponses positives au niveau de dilution i,

$$L(u) = \prod_{i=1}^k \binom{n_i}{p_i} e^{-uq_i (n_i - p_i)} (1 - e^{-uq_i})^{p_i} \quad (1)$$

## L'EQUATION DE VRAISEMBLANCE

L'expression  $L(u)$  et son logarithme montrent toutes les deux un maximum pour la même valeur de  $u$  qui est par définition le N.P.P.

$$\text{Log } L(u) = K - u \sum_{i=1}^k q_i (n_i - p_i) + \sum_{i=1}^k p_i \text{Log} (1 - e^{-uq_i})$$

$K$  est une constante.

$$\frac{d \text{Log } L}{du} = - \sum_{i=1}^k q_i (n_i - p_i) + \sum_{i=1}^k p_i \frac{q_i e^{-uq_i}}{1 - e^{-uq_i}}$$

Le N.P.P. s'obtient en annulant cette dérivée, et finalement le N.P.P. est la valeur de  $u$  solution de l'équation :

$$\sum_{i=1}^k q_i (n_i - p_i) = \sum_{i=1}^k \frac{p_i q_i e^{-uq_i}}{1 - e^{-uq_i}} \quad 2)$$

ANNEXE II

RESOLUTION DE L'EQUATION DE VRAISEMBLANCE

L'équation  $\sum_{i=1}^k q_i (n_i - p_i) = \sum_{i=1}^k \frac{p_i q_i e^{-uq_i}}{1 - e^{-uq_i}}$  est de la forme  $f(u) - K = 0$ .

En dehors des cas particuliers où toutes les réponses sont uniformément positives ou négatives (les N.P.P. associés à ces combinaisons sont alors respectivement :  $+\infty$  et 0), la résolution est possible par la méthode de NEWTON dont le principe est illustré par la figure 16.

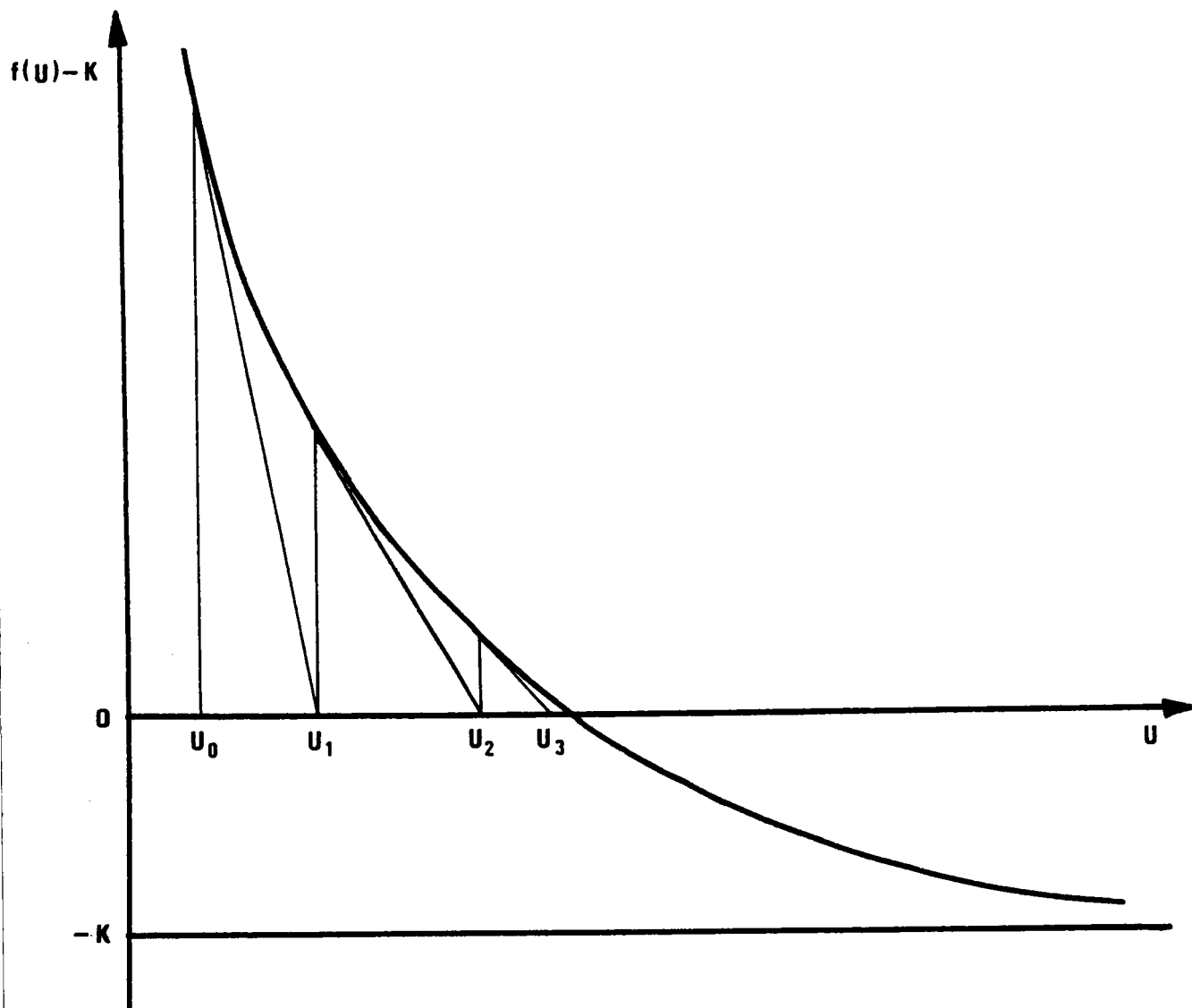


Figure 16 : Résolution de l'équation de vraisemblance par la méthode de NEWTON.

Soit  $u_1$  l'abscisse du point d'intersection avec l'axe  $Ou$  de la tangente à la courbe au point d'abscisse  $u_0$ .

La réitération de ce procédé donne une suite de valeurs  $u_0, u_1, \dots, u_n, \dots$  convergente vers la solution de l'équation, à condition toutefois que la valeur initiale  $u_0$  soit convenablement choisie.

Si la suite ne converge pas, il suffit de changer la valeur  $u_0$ .

PROGRAMME DE RESOLUTION DE L'EQUATION DE  
VRAISEMBLANCE DANS LE CAS DE TROIS DILUTIONS DECIMALES  
EFFECTUE POUR LA CALCULATRICE HEWLETT-PACKARD 41 C

01 LBL "NPP"	36 "INITIALE"	71 +	106 0.0001
02 "N1"	37 PROMPT	72 RCL 02	107 *
03 PROMPT	38 STO 04	73 RCL 07	108 1
04 STO 11	39 LBL 01	74 *	109 RCL 07
05 "P 1"	40 CHS	75 0.01	110 -
06 PROMPT	41 E $\nearrow$ X	76 *	111 X $\nearrow$ 2
07 STO 00	42 STO 05	77 1	112 /
08 "N2"	43 RCL 04	78 RCL 07	113 -
09 PROMPT	44 0.1	79 -	114 STO 09
10 STO 12	45 *	80 /	115 RCL 03
11 "P2"	46 CHS	81 +	116 RCL 08
12 PROMPT	47 E $\nearrow$ X	82 STO 08	117 -
13 STO 01	48 STO 06	83 RCL 00	118 RCL 04
14 "N3"	49 RCL 04	84 RCL 05	119 RCL 09
15 PROMPT	50 0.01	85 *	120 *
16 STO 13	51 *	86 1	121 +
17 "P3"	52 CHS	87 RCL 05	122 RCL 09
18 PROMPT	53 E $\nearrow$ X	88 -	123 /
19 STO 02	54 STO 07	89 X $\nearrow$ 2	124 STO 10
20 RCL 11	55 RCL 00	90 /	125 RCL 04
21 RCL 00	56 RCL 05	91 CHS	126 -
22 -	57 *	92 RCL 01	127 ABS
23 RCL 12	58 1	93 RCL 06	128 0.01
24 RCL 01	59 RCL 05	94 *	129 X > Y ?
25 -	60 -	95 0.01	130 GTO 02
26 0.1	61 /	96 *	131 RCL 10
27 *	62 RCL 01	97 1	132 STO 04
28 +	63 RCL 06	98 RCL 06	133 GTO 01
29 RCL 13	64 *	99 -	134 LBL 02
30 RCL 02	65 0.1	100 X $\nearrow$ 2	135 RCL 10
31 -	66 *	101 /	
32 0.01	67 1	102 -	
33 *	68 RCL 06	103 RCL 02	
34 +	69 -	104 RCL 07	
35 STO 03	70 /	105 *	

MODE D'UTILISATION

Le chargement du programme ayant été effectué directement à partir du listage ou bien par carte magnétique, le nombre le plus probable associé par exemple au triplet  $\left(\frac{96}{96}; \frac{38}{95}; \frac{4}{84}\right)$  est calculé de la manière suivante.

(les chiffres figurant au dénominateur représentent le nombre de cupules inoculées à chaque niveau de dilution).

<u>Séquence de touches</u>		<u>Affichage</u>
<span style="border: 1px solid black; padding: 2px;">XEQ</span>	<span style="border: 1px solid black; padding: 2px;">ALPHA</span> N.P.P. <span style="border: 1px solid black; padding: 2px;">ALPHA</span>	N1
96	<span style="border: 1px solid black; padding: 2px;">R/S</span>	P1
96	<span style="border: 1px solid black; padding: 2px;">R/S</span>	N2
95	<span style="border: 1px solid black; padding: 2px;">R/S</span>	P2
38	<span style="border: 1px solid black; padding: 2px;">R/S</span>	N3
84	<span style="border: 1px solid black; padding: 2px;">R/S</span>	P3
4	<span style="border: 1px solid black; padding: 2px;">R/S</span>	INITIALE
1	<span style="border: 1px solid black; padding: 2px;">R/S</span>	demande d'une valeur initiale permettant d'engager le processus itératif.

après quelques secondes d'attente, le résultat est affiché sur l'écran : 5,38.

L'affichage de la mention "OUT OF RANGE" suggère le choix d'une autre valeur initiale, la précédente n'ayant pas engendré une suite convergente.