

# Analyse d'un tableau de taux

Exploration multidimensionnelle de la circulation des agents infectieux pneumotropes chez les petits ruminants au Sénégal

E. Tillard & B. Faye

CIRAD-Elevage, Ligne Paradis, 97140 St Pierre, La réunion.

CIRAD-EMVT Campus International de Baillarguet, Montferriez sur lez, BP 5035, 34032 Montpellier cedex 1

## Résumé

On est souvent confronté en épidémiologie à l'analyse de tableaux dont les termes sont des taux. Lorsque les dénominateurs sont inégaux d'une ligne à l'autre, les centrages "colonne" classiques de l'ACP centrée ou normée ne sont pas autorisés car la moyenne d'une série de taux n'est pas une moyenne arithmétique. Pour l'analyse de base d'un tel tableau, on donne ici un exemple d'utilisation de l'analyse factorielle des correspondances décentrée (AFCD), extension de l'AFC simple, dans laquelle la marge "ligne" est remplacée par une marge extérieure imposée, contenant les effectif des dénominateurs. La structure spatio-temporelle du tableau peut ensuite être décrite par l'utilisation de méthodes déjà proposées dans une logique d'AFC (analyses interclasse, intraclasse, sur variables instrumentales, STATIS, analyse de co-inertie multiple). On montre que la structure de la programmation ADE-4 permet d'étendre les analyses K-tableaux à toutes les options d'analyse de base.

## Sommaire

Introduction.....	2
1 — Matériel et méthodes.....	3
1.1 — L'enquête sérologique.....	3
1.2 — Construction des tableaux de données.....	5
1.3 — Un double problème méthodologique.....	7
2 — Résultats.....	12
2.1 — AFCD.....	12
2.2 — Les analyses sur variables instrumentales.....	13
2.3 — Les analyses de tableaux multiples: STATIS et ACOM.....	19
3 — Discussion.....	26
3.1 — Les éléments pré-requis à la mise en oeuvre des analyses.....	26
3.2 — La méthodologie.....	26
3.3 — Discussion épidémiologique.....	29
Conclusion.....	31
Bibliographie.....	32
Annexes.....	35

## Introduction

La recherche zootechnique et vétérinaire en Afrique intertropicale a longtemps concentré ses efforts sur la seule espèce bovine. La période de sécheresse des années 1970 et les pertes importantes enregistrées dans les cheptels bovins, ont mis en évidence l'intérêt économique de développer l'élevage des petits ruminants et de mieux valoriser leurs qualités de rusticité et de renouvellement rapide. Parmi les nombreuses contraintes à l'élevage des petits ruminants identifiées au cours des 2 dernières décennies, les affections respiratoires ont fait l'objet d'une attention particulière en raison du fait qu'elles constituent une dominante pathologique majeure dans la plupart des pays de la frange sub-saharienne (Lefevre, 1987).

Sur le terrain, ces affections font malheureusement rarement l'objet d'un diagnostic précis, les agents des services de l'Élevage ayant tendance à imputer presque systématiquement aux sur-infections pasteurelliques l'entière responsabilité dans l'initiation de ces pathologies. Si certaines entités comme la peste des petits ruminants (PPR) qui sévit à l'état enzootique au Sénégal relèvent de germes à pouvoir pathogène primaire sans équivoque, les plus fréquentes correspondent au contraire à un syndrome qualifié de "complexe pneumopathique". Ce syndrome se caractérise par une étiologie multi-factorielle et séquentielle. Son apparition nécessite une infection primitive d'origine virale, chlamydiennne ou mycoplasmatique, suivie d'une sur-infection bactérienne, généralement pasteurellique, nécessaire à l'extériorisation des signes cliniques (Davies et al., 1980). La plupart des agents classiquement impliqués dans ce complexe dans les pays d'élevage intensif ont été retrouvés au Sénégal: le virus *ParaInfluenza 3* et *Oviadénovirus type 5* (Sarr et al., 1987 - Nguyen-Ba-Vy et al., 1988), *Pasteurella haemolytica* et *Pasteurella multocida* (Doutre et al., 1981).

Les études microbiologiques sur foyers de maladie, les méthodes d'enquête transversale fournissant des informations à un instant donné, ou les reproductions expérimentales ne sont pas réellement adaptées à l'étude d'un tel complexe, car elles ne permettent d'étudier correctement que des entités pathologiques d'étiologie bien définie, par des agents circulants sur un mode épizootique et déterminant une immunité de longue durée (Desoutter, 1994). C'est pourquoi, dans le cadre d'un programme de recherche sur l'élevage des petits ruminants en milieu traditionnel (programme Pathologie et Productivité des petits ruminants en milieu traditionnel - P.P.R.), conjointement mené par le CIRAD-EMVT (Centre de coopération International en Recherche Agronomique pour le Développement, département Elevage et Médecine Vétérinaire des pays Tropicaux) et l'ISRA-LNERV (Institut Sénégalais de Recherches Agricoles - Laboratoire National d'Elevage et de Recherches Vétérinaires), une enquête sérologique longitudinale de type prospectif a été menée en 1988 et 1989 chez des ovins et des caprins élevés en milieu villageois. Elle ciblait les principaux agents infectieux impliqués dans le syndrome "pneumopathie", dans 3 régions différentes du pays.

Les données issues de cette enquête ont déjà fait l'objet d'une synthèse importante, qui sera notre référence essentielle dans l'interprétation épidémiologique des résultats (Desoutter, 1994). Une analyse minutieuse des données, agent par agent, a permis d'établir des comparaisons entre zones, entre villages, entre espèces et entre classes d'âge (test du  $\chi^2$ ). Cependant, cette approche univariée apparaît trop limitative dans l'étude d'une pathologie dont l'apparition nécessite, outre l'association de plusieurs agents infectieux, la mise en jeu de facteurs environnementaux. L'objet de la présente étude consiste donc à revisiter les données sérologiques obtenues sur les ovins de la région Nord du Sénégal au moyen d'analyses multivariées, dans le but d'identifier les phénomènes d'association et de compétition dans la circulation des agents pneumotropes au niveau villageois, et leurs variations spatio-temporelles. Un accent particulier sera accordé aux spécificités méthodologiques liées à la nature des données analysées, des taux dans le cas présent, et à leur organisation dans le temps et l'espace, problème qui avait fait l'objet du rapport bibliographique.

# 1 — Matériel et méthodes.

## 1.1 — L'enquête sérologique

Une enquête longitudinale de type prospectif a été retenue. Elle permet d'appréhender en milieu réel la séquence des agents pathogènes impliqués dans les pneumopathies, de travailler sur des cinétiques d'anticorps, véritables indicateurs de la circulation des agents infectieux chez l'animal, et de détecter une éventuelle périodicité de passage, hypothèse réaliste compte tenu des variations climatiques annuelles qui caractérisent la zone. Elle sous-tend que l'infection des animaux est suffisamment fréquente et répandue pour affecter la population étudiée pendant la durée de l'enquête (Rumeau-Rouquette et al., 1993).

### 1.1.1 — Localisation et échantillonnage

L'enquête a été menée dans 3 zones différentes, la région du Delta du fleuve Sénégal (zone sahélienne), la région de Louga dans le bassin du Ferlo (zone sahélienne), et la région de Kolda (zone sud soudanienne). Ces régions ont été choisies initialement pour être représentatives de la diversité des systèmes d'élevage de ruminants dans le pays. Sur le plan logistique, elle s'est appuyée sur 2 programmes de recherche du CIRAD-EMVT qui effectuaient depuis plusieurs années des suivis en milieu villageois, "Syspro - Fleuve" dans la région du Delta du Sénégal (Tourrand, 1994), et "P.P.R." dans les régions de Louga et Kolda (Faugère et al., 1986). Ces 2 programmes ont évidemment conditionné l'échantillonnage des animaux. Dans chaque région, 10 à 20 villages ont été retenus après une enquête préalable auprès des éleveurs, destinée à s'assurer de la représentativité des systèmes de production. Ces villages, qui représentent une population de 50 à 100 troupeaux et de 1500 à 3000 têtes suivies par région, ont été répartis au hasard en 4 lots (échantillonnage en grappe), dans le cadre d'une opération de recherche destinée à évaluer la rentabilité économique de diverses prophylaxies testées en grandeur réelle. Finalement, l'enquête sérologique n'a donc pu être réalisée que sur les animaux des villages du lot témoin (25 %), seuls à ne recevoir aucun traitement (Tillard, 1991).

La présente étude sera limitée aux 2 régions situées dans la zone sahélienne du pays (figure 1), et à l'espèce ovine. En effet, les zones sahélienne et sud soudanienne sont très contrastées (climat, type génétique, systèmes de production et agents infectieux circulants différents) et par conséquent difficilement comparables. Les 2 régions de la zone sahélienne offrent une plus grande homogénéité: on y rencontre les mêmes types génétiques d'ovins (Peul Peul et Warale), et les mêmes agents infectieux pneumotropes. Elles présentent néanmoins une certaine hétérogénéité, dans leurs caractéristiques climatiques (humidité, vents), et leurs activités agropastorales. Le choix de l'espèce ovine n'est dicté que par des considérations pratiques, le nombre de caprins suivis étant généralement insuffisant pour autoriser toute interprétation à l'échelle du village.

### 1.1.2. — Protocole de terrain

L'enquête s'est déroulée sur une période de 18 mois, de novembre 1987 à juillet 1989, et inclut 2 saisons sèches fraîches, correspondant aux périodes de recrudescence habituelle des affections respiratoires, et une saison des pluies. Les mères ont été échantillonnées par classe d'âge, et leur taux de sondage par village s'est élevé à 25-30 p. cent des femelles ayant agnelé sur la période d'étude. Les jeunes animaux ont été soumis à partir de l'âge d'un mois en moyenne à des prises de sang répétées à intervalle de 45 à 60 jours, jusqu'à l'âge maximal de 15 mois.

Les animaux ne connaissent pas tous la même durée de suivi. Les dates de recrutement (premier prélèvement) se sont échelonnées sur toute la période, au rythme des naissances et beaucoup d'animaux sont sortis du suivi avant son terme, après mortalité ou exploitation décidée par l'éleveur. On observe donc une décroissance rapide de l'effectif d'animaux suivis.

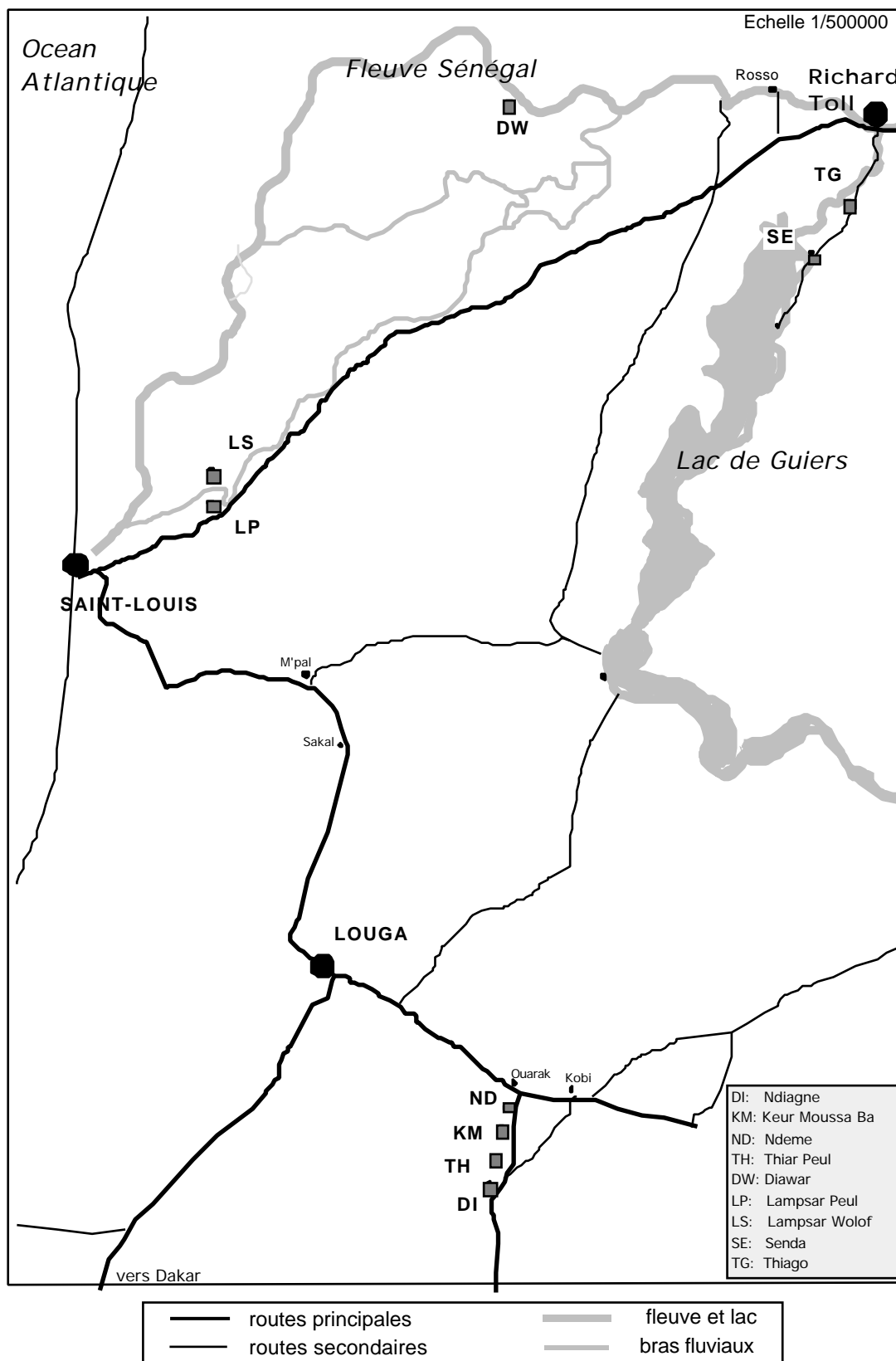


Figure 1 : Carte de la zone étudiée. Région du Delta du fleuve Sénégal et région de Louga

Au total, 45% des animaux ont été suivis pendant au moins 7 mois, et seuls 15 % d'entre eux pendant la durée complète du suivi (15 mois) (Desoutter, 1994).

L'annexe 4-5 donne un extrait des résultats bruts pour la sérologie RSV. Près de 33000 analyses sérologiques ont ainsi été réalisées sur les ovins de la zone sahélienne, dont 30 p. cent ont été utilisés dans le cadre de cette étude.

## 1.2 — Construction des tableaux de données

### 1.2.1 — Choix des villages et des agents infectieux

Le tableau I précise le nom et le code des villages enquêtés, ainsi que le nombre de troupeaux suivis et d'ovins prélevés. Trois villages ont été éliminés de l'analyse en raison d'un effectif trop faible (NO, SO et WA) et un quatrième en raison d'un nombre trop élevé de données sérologiques manquantes (TI).

Tableau I: noms et codes des villages, nombre de troupeaux et d'animaux suivis.

	Noms des villages	Codes	Nombre de troupeaux suivis	Nombres d'ovins prélevés
Louga	Ndiagne	DI	18	57
	Keur Mouso kebe	KM	7	16
	Ndeme	ND	16	37
	Thiar-Peul	TH	4	15
Delta	Diawar	DW	29	60
	Lampsar Peul	LP	4	16
	Lampsar Wolof	LS	33	64
	Ndombo	NO	5	6
	Senda	SE	1	23
	Sowanabe	SO	1	6
	Thiago	TG	10	23
	Thilene	TI	10	17
	Wadabe	WA	2	2

Parmi les 17 agents pneumotropes qui ont été recherchés systématiquement dans chaque prélèvement sanguin, 12 ont été retenus (tableau II). Ils correspondent aux agents classiquement impliqués dans le complexe pneumopathique. Ont été éliminés le virus de la peste des petits ruminants, responsable d'une entité clinique bien définie, *Pasteurella haemolytica* type 8 en raison d'un trop faible nombre de résultats positifs, et les 4 *mycoplasma* (*F38*, *capricolum*, *Mycoïdes sp mycoïdes* (LC) et *ovipneumoniae*) pour lesquels on n'enregistre que des résultats négatifs.

Tableau II: Nom et code des agents, technique sérologique utilisée et seuil de positivité.

Nom de l'agent	agent	Code	Technique sérologique	Seuil
<i>Parainfluenza III</i>	Virus	PI	Inhibition de l'hémagglutination	1/20
<i>Virus Respiratoire syncytial</i>	Virus	RS	Hémagglutination passive	1/20
<i>Adénovirus type 5</i>	Virus	AD	Séro-neutralisation	1/10
<i>Pasteurella multocida type A</i>	Bactérie	MA	Hémagglutination passive	1/20
<i>Pasteurella multocida type D</i>	Bactérie	MD	Hémagglutination passive	1/20
<i>Pasteurella haemolytica type 1</i>	Bactérie	H1	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 2</i>	Bactérie	H2	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 5</i>	Bactérie	H5	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 6</i>	Bactérie	H6	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 7</i>	Bactérie	H7	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 9</i>	Bactérie	H9	Hémagglutination passive	1/10
<i>Pasteurella haemolytica type 11</i>	Bactérie	H11	Hémagglutination passive	1/10

### 1.2.2 — Choix d'une échelle temporelle

La circulation des agents infectieux n'est pas un phénomène uniforme dans le temps. La plupart des agents impliqués dans le complexe pneumopathique présente une résistance dans le milieu extérieur ou un pouvoir pathogène pour l'animal conditionné par les caractéristiques

climatiques (températures, humidités, vents ...), que l'on sait être très fluctuantes dans l'année. De plus, les observations successives sur un même animal ne sont pas indépendantes. En effet, après une primo-infection, un animal acquiert une protection immunitaire qui le rend momentanément moins sensible à une ré-infection, et moins apte à la transmission de l'agent. La répartition saisonnière des mises bas en zone sahélienne engendre également des périodes de moindre réceptivité des animaux, protégés par les anticorps collostraux. Il est donc essentiel d'agréger les données par période, de manière à les rendre le plus indépendantes possible, et pouvoir identifier des périodes de circulation privilégiée. Sur la base de critères climatiques, 4 périodes différentes ont été fixées, qui définiront 4 tableaux différents, auxquels on affectera dans l'analyse un poids identique malgré leur durée sensiblement différente:

- de mai à juillet 1988: fin de saison sèche chaude (FS)
- de juillet à septembre 1988: saison des pluies (SP)
- de septembre 1988 à janvier 1989: saison sèche fraîche (SF)
- de janvier à mai 1989: saison sèche chaude (SC).

### 1.2.3 — Construction d'un indicateur de la circulation des agents infectieux

On ne retient que les animaux ayant fait l'objet d'un prélèvement au début et à la fin d'une période. Les 2 titres sérologiques définissent chez un même animal une cinétique d'anticorps. L'observation des cinétiques conduit à envisager 4 situations différentes décrites dans le tableau III.

Tableau III: interprétation des cinétiques d'anticorps.

Titre sérique 1	Titre sérique 2		Interprétation
0	0	$s_-$	sérologie négative - absence de circulation
0	> 0	$s_+$	sérologie positive - primo-infection
> 0	2 > 1	$s_c$	sérologie croissante - ré-infection
> 0	2 1	$s_d$	sérologie décroissante - absence de circulation

Pour caractériser la circulation des agents infectieux, on a retenu, en s'inspirant du calcul classique des taux de séroconversion (Thrusfield, 1986), pour chaque agent pathogène, l'indicateur suivant (taux de circulation), qui rapporte la somme des animaux présentant une augmentation du titre d'anticorps ( $x$ ) à la totalité des animaux prélevés en début et en fin de

période ( $r$ ):  $tc = \frac{n(s_+) + n(s_c)}{n(s_-) + n(s_+) + n(s_c) + n(s_d)} = \frac{x}{r}$ . Ce taux de circulation prend en compte à

la fois les animaux primo-infectés ( $n(s_+)$ ) et ré-infectés ( $n(s_c)$ ). Il sera calculé à l'échelle du village. Traditionnellement, en effet, l'élevage des petits ruminants est un élevage extensif. Les animaux sont parqués pour la nuit dans leurs concessions respectives. Dans la journée, les troupeaux de concession d'un même village sont rassemblés en troupeaux de jour et conduits au pâturage sous la surveillance d'un berger. Cette conduite commune aura donc tendance à homogénéiser les résultats d'un même village. Au contraire, l'éloignement relatif des villages rend peu probable la rencontre de troupeaux villageois différents.

### 1.2.4 — Standardisation de la population de référence

Dans les 2 régions, on enregistre habituellement un pic de mise bas en saison sèche fraîche, entre novembre et mars (Faugère et al., 1989 - Tourrand, 1994). Or, on constate qu'une majorité d'animaux est recrutée entre juin et novembre 1988 pour la région du Delta du fleuve (tableau IV). Ce déphasage partiel des mises bas dans le temps pour la région du Delta est dû à la survenue en octobre et novembre 1987 d'une épizootie de Fièvre de la Vallée du Rift ayant provoqué de nombreux d'avortements dans tous les villages suivis. Une grande partie des femelles a été de nouveau fécondée entre décembre 1987 et avril 1988, et a mis bas 5 mois plus tard entre avril et septembre 1988, expliquant les recrutements plus nombreux à cette époque, dans la région du Delta.

Tableau IV: évolution des effectifs d'animaux prélevés par période de recrutement (PR).

Zone	PR	01/88	02/88	04/88	06/88	07/88	09/88	11/88	01/89	03/89	05/89
Louga	01/88	48	47	44	36	21	20	17	15	14	13
	02/88		39	38	33	24	20	15	13	12	12
	04/88			6	6	4	4	2	1	1	1
	06/88				3	2	0	0	0	0	0
	07/88					0	0	0	0	0	0
	09/88						16	14	14	9	6
	11/88							12	11	9	7
	Total		48	86	88	78	51	60	60	54	45
% recrut.		100	45	7	4	0	27	20	0	0	0
Delta	01/88	22	0	21	19	17	10	7	7	7	5
	02/88		19	0	17	17	16	15	13	8	8
	04/88			8	8	7	7	6	6	5	4
	06/88				36	36	36	33	28	19	15
	07/88					37	36	34	31	19	17
	09/88						27	26	24	18	14
	11/88							36	33	29	26
	Total		22	19	29	80	114	132	157	142	105
% recrut.		100	100	28	45	32	20	23	0	0	0

Pour éliminer cette différence conjoncturelle d'âge entre les animaux des 2 régions, les données ont été standardisées par rapport à une population de référence (Rumeau-Rouquette et al., 1993). La réceptivité d'un animal aux agents infectieux a déterminé le choix des classes d'âge de cette population. Elle est conditionnée au début de son existence par la présence d'anticorps d'origine maternelle qu'il acquiert en tétant le colostrum. Ces anticorps ont presque disparu à l'âge de 3 mois (Desoutter, 1994), ce qui conduit à définir 2 classes d'âge, celle des animaux de moins de 3 mois et de plus de 3 mois. La faiblesse des effectifs par classe d'âge dans certains villages ne permet pas d'en envisager davantage. Pour chaque agent pathogène, chaque village et chaque période, on calcule un taux de circulation pour chaque classe d'âge

$$t_1 = \frac{x_1}{r_1} \text{ et } t_2 = \frac{x_2}{r_2} \text{ et un taux de circulation corrigé: } t_c = \frac{(t_1 \times r') + (t_2 \times r'')}{r' + r''} \text{ ou } r' \text{ et } r'' \text{ sont}$$

les effectifs par classe d'âge d'une population de référence pour chaque période, celle constituée par l'ensemble des animaux prélevés dans la région de Louga, qui présente une répartition normale des mises bas. Les effectifs corrigés d'animaux présentant une augmentation du titre d'anticorps sont obtenus ensuite par  $x_c = t_c \times (r_1 + r_2)$ . Un exemple de calcul est donné en annexe 5.

### 1.3 — Un double problème méthodologique

Le choix d'une méthode d'analyse doit prendre en compte la nature des données (taux) et leur organisation spatio-temporelle. Les tableaux des taux ( $t_c$ ) et des effectifs d'animaux ayant présenté une augmentation du titre ( $x_c$ ) sont présentés en annexe 5, pour les villages et les agents pathogènes retenus.

On n'a conservé dans l'analyse que les villages présentant à chaque période un effectif d'animaux prélevés suffisant (supérieur ou égal à 50 p. cent de l'effectif potentiellement utilisable), de manière à ce que la précision sur les taux ne soit pas trop disparate. Les 9 villages étudiés ne sont pas tous représentés pour chaque période.

On considère donc un plan factoriel déséquilibré à 2 facteurs (période et village) sans répétition (au plus une mesure par modalité croisée).

### 1.3.1 — L'analyse d'un tableau de taux: l'Analyse Factorielle des Correspondances Décentrée (AFCD)

On considère dans un premier temps un seul tableau  $X$  regroupant les différentes périodes, de terme général  $x_{ij}$  = nombre d'animaux présentant une augmentation du titre, pour un village  $i$  et un agent pathogène  $j$ . Pour prendre en compte les effectifs d'une population de référence, Lebart (1979) ajoutait au tableau  $X$  une colonne active supplémentaire contenant ces effectifs, qui, par son ordre de grandeur bien supérieur aux termes de  $X$ , "fixait" l'origine des axes. L'AFCD reprend ce principe tout en le formalisant de bien meilleure façon. Dans cette technique d'analyse factorielle analogue à l'AFC, la marge "ligne" du tableau  $X$  est remplacée par une marge imposée  $r_i$  (Escofier, 1981, Dolédec et al., 1995).

A chaque village  $i$  correspond un effectif d'animaux prélevés  $r_i$  (constant pour l'ensemble des agents  $j$ ). Dans cette analyse, on remplace donc le poids des lignes de l'AFC  $x_{i.} = \sum_{j=1}^p x_{ij}$

par  $r_i$  et  $x_{.j} = \sum_{i=1}^n x_{ij}$  par  $R = \sum_{i=1}^n r_i$ . Le profil ligne  $\frac{p_{ij}}{p_{i.}} = \frac{x_{ij}}{r_i} = t_{ij}$  est aussi le taux de circulation pour le village  $i$  et l'agent pathogène  $j$ , et la marge colonne  $p_{.j} = \bar{t}_j$ , le taux de circulation moyen pour l'agent pathogène  $j$  (moyenne de variables quotients). Le tableau traité a pour

terme général  $\frac{p_{ij} - p_{i.}p_{.j}}{p_{i.}p_{.j}} = \frac{\frac{x_{ij}}{R} - \frac{r_i}{R} \frac{x_{.j}}{R}}{\frac{r_i}{R} \frac{x_{.j}}{R}} = \frac{\frac{x_{ij}}{r_i} - \frac{x_{.j}}{R}}{\frac{x_{.j}}{R}} = \frac{t_{ij} - \bar{t}_j}{\bar{t}_j}$ . La pondération par  $\bar{t}_j$  tend à

augmenter le poids des agents à faible moyenne comme MD et à diminuer le poids des agents à forte moyenne (H11 et MA). Ainsi un écart de 5 points face à une moyenne de 10 (%) présentera un poids plus important qu'un écart de 5 points face à une moyenne de 50.

Pour le nuage des colonnes, l'origine des axes dans l'AFCD a pour coordonnées  $\frac{r_i}{R}$  et représente la position d'un agent pathogène fictif dont la circulation serait exactement proportionnelle à l'effectif des animaux prélevés, c'est à dire identique pour tous les villages et toutes les périodes. Un test multinomial d'ajustement ( $\chi^2$  à  $n-1$  degrés de liberté) permet de tester l'éloignement de chaque variable du tableau (agent) à cette répartition théorique (Lebart, 1979).

L'AFC du tableau  $X$  décentrée par rapport à la marge  $r_i$  est bien adaptée aux objectifs de départ. Elle permet de fixer comme point de référence l'effectif des animaux prélevés, de travailler par conséquent sur des taux et de comparer des villages entre eux et des périodes entre elles, même si les effectifs prélevés sont très différents.

### 1.3.2 — L'analyse d'un tableau de données structurées

Le tableau est structuré par un plan expérimental. Les lignes sont regroupées en  $k$  blocs (ou  $k$  sous-tableaux), chaque bloc correspondant à une période ou à un village (l'indice  $k$  est utilisé pour rapporter les éléments à un même sous-tableau). On adoptera dans cette partie les notations suivantes:  $x_{ij}^k$  = nombre d'animaux présentant une augmentation du titre, pour un village  $i$ , une période  $k$  et un agent pathogène  $j$ ,  $r_i^k$  = nombre d'animaux prélevés, pour un village  $i$  et une



période  $k$ , et  $t_{ij}^k = \frac{x_{ij}^k}{r_i^k}$  = taux de circulation, , pour un village  $i$ , une période  $k$ , et un agent pathogène  $j$ . Pour prendre en compte l'organisation spatio-temporelle des données, deux approches seront successivement envisagées, une approche non symétrique par les analyses structurées (analyse interclasse et intraclasse, analyse sur variables instrumentales) et une approche symétrique par les analyses de tableaux multiples (Chessel et al., 1993b).

### 1.3.2.1 — Les analyses structurées: décomposition de l'inertie totale du tableau

On associe au tableau  $X$  à  $n$  lignes et  $p$  colonnes, de terme  $x_{ij}^k$ , 2 variables qualitatives externes,  $P$  et  $V$  (périodes, villages), qui structurent le plan expérimental. Chacune des variables engendre 2 sous-espaces vectoriels orthogonaux entre eux, notés  $\mathbf{P}$ ,  $\mathbf{P}$  et  $\mathbf{V}$ ,  $\mathbf{V}$  sur lesquels on projette  $X$  et sur lesquels on décompose sa variance totale par des analyses interclasse et intraclasse, qui maximisent respectivement la variance des centres de gravité des classes et la variance des écarts aux centres de gravité, variance résiduelle de la précédente (Dolédec et al., 1987 - Dolédec et al., 1989). Chaque projection "extrait" du tableau initial la structure la plus caractéristique liée à un effet donné, ou son complémentaire.

L'AFC interclasse est également une AFCD et analyse un tableau à  $k$  lignes contenant les sommes des  $x_{ij}^k$  par bloc ( $x_j^k$ ), avec pour marge imposée  $\frac{R^k}{R}$ . Le tableau traité est de terme

$$\frac{p_j^k - p_{..} p_{.j}}{p_{..} p_{.j}} = \frac{\frac{x_j^k}{R^k} - \frac{x_{.j}}{R}}{\frac{x_{.j}}{R}} = \frac{\bar{t}_j^k - \bar{t}_j}{\bar{t}_j} \text{ avec } \bar{t}_j^k, \text{ taux moyen par tableau et } \bar{t}_j, \text{ taux moyen tous tableaux confondus.}$$

L'analyse ne prend par conséquent en compte que l'effet associé à  $k$  représenté par  $(\bar{t}_j^k - \bar{t}_j)$ .

L'AFC intraclasse analyse un tableau de terme 
$$\frac{p_{ij} - p_i p_{.j}}{p_i p_{.j}} = \frac{\frac{x_{ij}^k}{r_i^k} - \frac{x_{.j}}{R}}{\frac{x_{.j}}{R}} = \frac{\frac{x_{ij}^k}{r_i^k} - \frac{x_{.j}}{R^k}}{\frac{x_{.j}}{R}} =$$

$$\frac{t_{ij}^k - \bar{t}_j^k}{\bar{t}_j^k}, \text{ égal à l'écart au taux moyen par tableau rapporté au taux moyen tous tableaux confondus.}$$

L'analyse élimine ici l'effet associé à  $k$  et utilise la même pondération des colonnes que l'analyse interclasse ( $\bar{t}_j$ ). L'inertie totale du tableau  $X$  se décompose donc en une inertie

inter et une inertie intra: 
$$\frac{(t_{ij}^k - \bar{t}_j^k)}{\bar{t}_j^k} = \frac{(\bar{t}_j^k - \bar{t}_j)}{\bar{t}_j} + \frac{(t_{ij}^k - \bar{t}_j^k)}{\bar{t}_j^k} \text{ ou } I_T = I_k + I_k .$$

Pour valider le modèle associé à chaque sous espace  $\mathbf{P}$  et  $\mathbf{V}$ , on dispose d'un test de permutation aléatoire (Manly, 1991), qui consiste à établir une distribution de l'inertie du tableau par répartition au hasard les lignes dans chacun des  $k$  blocs, puis à calculer la probabilité pour que l'inertie simulée soit supérieure à l'inertie observée, avec un seuil de 5 p. cent. Des analyses univariées par régression logistique seront également utilisées pour confirmer les conclusions des tests de permutation multivariés (annexe 12).

La prise en compte des 2 variables P et V simultanément soulève une difficulté: les sous-espaces  $P$  et  $V$  ne sont pas orthogonaux entre eux du fait de la présence de cases vides (plan déséquilibré). Avec un plan expérimental orthogonal (1 valeur par modalité croisée des 2 facteurs), les  $t_{ij}^k$  forment, pour chaque variable du tableau, un plan d'analyse de variance à 2 facteurs sans répétition (Dolédec et al., 1987). On considère un effet additif des 2 facteurs P et V formant respectivement  $k1$  et  $k2$  blocs, et on analyse un tableau de terme  $\frac{(\bar{t}_{ij}^{k1} - \bar{t}_j) + (\bar{t}_{ij}^{k2} - \bar{t}_j)}{\bar{t}_j}$ , dont la variance totale se décompose ainsi (d'après Dagnélie, 1975):

$$\frac{(t_{ij}^k - \bar{t}_j)}{\bar{t}_j} = \frac{(\bar{t}_j^{k1} - \bar{t}_j) + (\bar{t}_j^{k2} - \bar{t}_j) + (t_{ij}^k - \bar{t}_j^{k1} - \bar{t}_j^{k2} + \bar{t}_j)}{\bar{t}_j} \quad \text{ou} \quad I_T = I_P + I_V + I_{res},$$

le dernier terme entre parenthèses représentant la variation résiduelle.

Avec un plan non orthogonal, on est conduit à envisager pour la réalisation des tests de permutation, par analogie avec les tests pratiqués en analyse de variance de plans factoriels déséquilibrés, 2 décompositions séquentielles symétriques de l'inertie totale du tableau (Fraile et al., 1993):  $I_T = I_P + I_{V_P} + I_{res}$  et  $I_T = I_V + I_{P_V} + I_{res}'$ . Les tests portent alors sur les sous-espaces  $I_{P_V}$  et  $I_{V_P}$  et permettent de conclure sur l'opportunité de prendre en compte un modèle incluant un effet additif des 2 facteurs.

La prise en compte d'un effet additif sera effectuée par décomposition de l'inertie totale du tableau sur des bases orthonormées définissant 2 sous-espaces,  $(P + V)$  et  $(P - V)$ , son sous-espace orthogonal, sur lequel les lignes sont centrées à la fois par village et par période.

Au total, chaque effet principal pourra être analysé par 2 projections différentes, une projection sur  $P$  (analyse interpériode) et sur  $V$  (analyse intravillage), pour l'effet période, et une projection sur  $V$  (analyse intervillage) sur  $P$  (analyse intrapériode), pour l'effet village (Dolédec et al., 1987). Un effet additif pourra être envisagé in fine si les tests de permutation sont significatifs.

### 1.3.2.2 — Les analyses de tableaux multiples: STATIS et ACOM

On ne cherche plus à décomposer la variance totale d'un tableau sur différentes variables externes, mais à décrire directement les structures induites dans les données par le plan expérimental (Thioulouse et al., 1987). On considère dans le cas présent 4 sous-tableaux distincts (un par période), définissant chacun une typologie propre des agents pathogènes. L'objectif de ces méthodes consiste à ordonner simultanément les 4 sous-tableaux, c'est à dire trouver un tableau moyen donnant une typologie moyenne des agents définie communément par les villages des 4 sous-tableaux, et analyser les écarts entre cette typologie de référence et celle induite séparément par chaque sous-tableau, avec l'aide des positions des villages des sous-tableaux.

La comparaison de  $k$  tableaux nécessite qu'ils aient des points communs, en particulier le poids des éléments de la marge commune, et les types de centrage. Dans l'AFCM de chacun des sous tableaux, le poids des lignes est imposé et le poids des colonnes ( $p_j^k$ ) est défini par les éléments du sous-tableau. Ces poids diffèrent donc pour chacun des sous-tableaux.

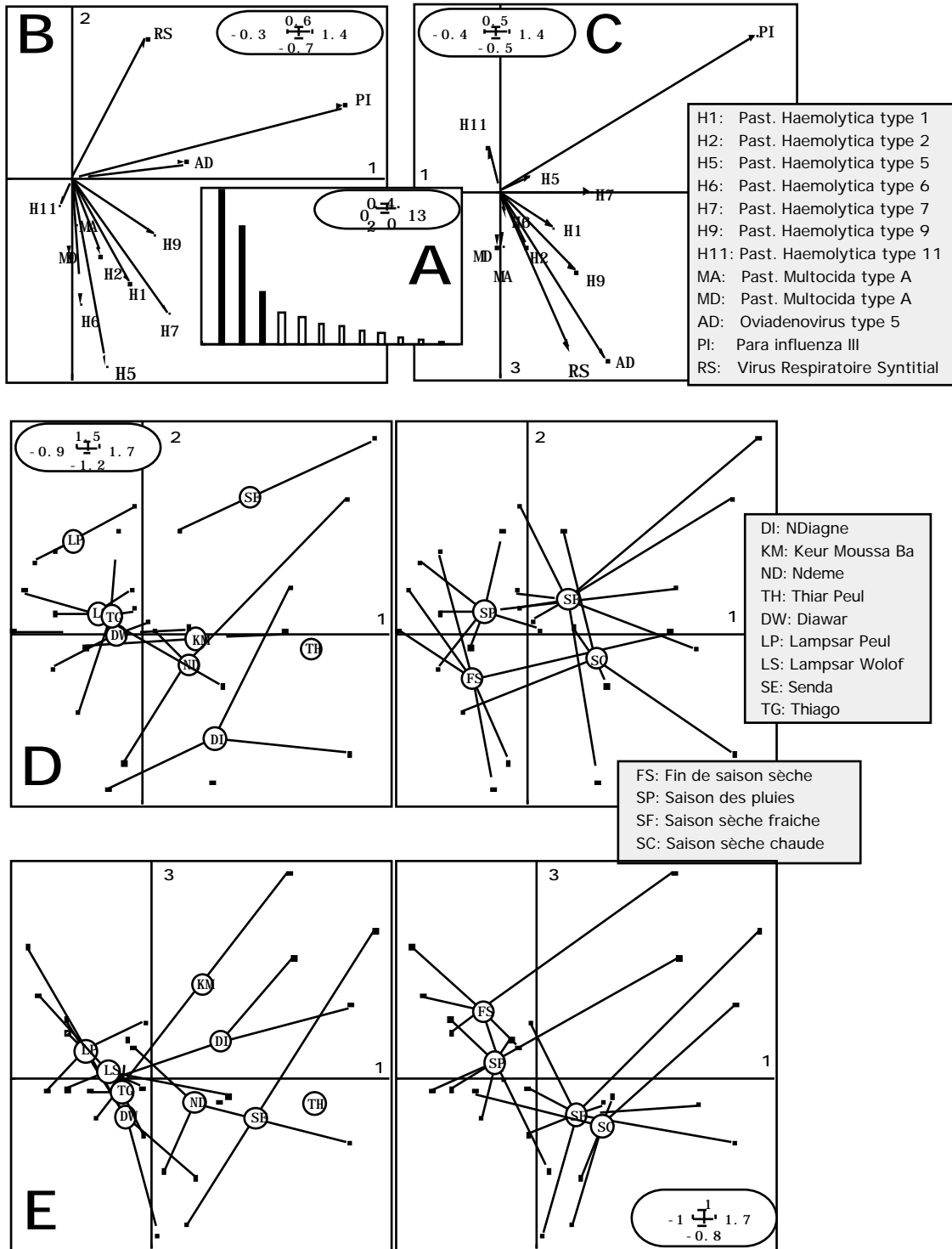


Figure 2. Analyse Factorielle des Correspondances Décentrée (AFCD)

- A: Histogramme des valeurs propres
- B: Plan factoriel 1-2 des colonnes
- C: Plan factoriel 1-3 des colonnes
- D: Plan factoriel 1-2 des lignes, regroupées par village puis par période
- E: Plan factoriel 1-3 des lignes, regroupées par village puis par période

Parmi les 3 options disponibles dans le logiciel ADE 4.0 pour la pondération et les centrages des  $k$  tableaux dans une logique d'AFC, on a retenu, par analogie avec les centrages réalisés

dans les ACP de  $k$  tableaux, l'option qui conserve les centrages de l'AFC intrapériode et conduit à l'analyse d'un triplet  $\frac{p_{ij} - p_i.p_j}{p_i.p_j}, \frac{p_i}{R^k/R}, p_j$ . Chacun des sous-tableaux est centré

sur sa moyenne et pondéré par la moyenne générale, de terme  $\frac{t_{ij}^k - \bar{t}_j^k}{\bar{t}_j}$ . La division du tableau

initial en 9 unités "village" ne sera pas envisagée du fait d'une trop grande disparité du nombre de lignes par village.

Deux méthodes présentent l'avantage de pouvoir être mise en oeuvre sur des tableaux ayant les mêmes colonnes (les agents pathogènes), la méthode STATIS (Structuration des Tableaux A Trois Indices de la Statistique) (L'Hermier des Plantes, 1976 - Lavit, 1988) et l'Analyse de CO-inertie Multiple (ACOM) (Chessel et al., 1996). Les applications de STATIS dans une logique d'AFC et de l'ACOM sont encore peu fréquentes (Chessel et al., 1995a, 1995b, 1995c), et, à notre connaissance, inexistantes sur tableau avec marge imposée. STATIS définit un tableau moyen (le compromis), décrit sa structure par ACP, et projette sur ses axes les éléments des lignes et des colonnes des tableaux initiaux. L'ACOM coordonne  $k$  analyses, trouve  $k$  systèmes d'axes de co-inertie qui maximisent chacun la covariance des coordonnées des projections des colonnes avec un système d'axes de référence (variables auxiliaires).

Les analyses séparées non coordonnées donnent 5 jeux de facteurs orientés de manière aléatoire, rendant difficiles les comparaisons des typologies 2 à 2. STATIS et l'ACOM sont des analyses de synthèse, coordonnées pour orienter les axes de manière unique, afin de permettre la comparaison des analyses des sous-tableaux entre elles (Chessel et al., 1995d), et ainsi de juger, dans le cadre de cette étude, de la stabilité de la typologie des agents dans le temps.

## 2 — Résultats

Les analyses multifactorielles, les tests de permutation et les sorties graphiques ont été réalisés avec le logiciel ADE-4. L'annexe 2 donne la liste des modules et des options utilisés. Les contributions absolues et relatives ont été systématiquement examinées. Seules les variables présentant une contribution absolue supérieure à la moyenne des contributions ont été retenues pour l'explication d'un axe. Les plans factoriels des figures ont été choisis de manière à représenter au mieux le maximum de variables, en particulier les agents viraux. Les analyses univariées (regressions logistiques) ont été réalisées avec le logiciel SPSS/PC (Norusis, 1992).

### 2.1 — AFCD

La figure 2 présente les plans factoriels 1-2 et 1-3 de l'AFCD. Le tableau V donne les résultats du  $\chi^2$  à 25 ddl, testant l'écart des distributions observées par variable à une distribution proportionnelle aux effectifs prélevés, écart matérialisé sur les plans factoriels par la longueur des vecteurs. L'annexe 6 fournit les listings de sortie (valeurs propres,  $\chi^2$  et contributions), et une représentation graphique du tableau de données traité  $P_0$ .

Tableau V : résultats du test multinomial d'ajustement pour les différentes variables.

Agent	H1	H2	H5	H6	H7	H9	H11	MA	MD	AD	VI	RS
$p(\chi^2)$	0.292	0.07	0.001	0.038	0.001	0.391	0.336	0.085	0.897	0	0	0

Les plans 1-2 et 1-3 résument la plus grande partie de l'information (70 p. cent de l'inertie totale). Six variables contribuent fortement à les expliquer, H5, H6, H7, AD, PI, et RS (tableau V). La direction des vecteurs est interprétée à l'aide des cartes factorielles des lignes. L'axe 1

est caractérisé par l'intensité de la circulation de PI (65 p. cent de l'inertie associée à l'axe 1) et supporte un effet "taille" (la majorité des agents présente des coordonnées positives sur cet axe) lié à l'opposition des villages des 2 régions, DI, KM, ND et TH pour la région de Louga et DW, LP, LS, et TG pour la région du Delta, pour une majorité de variables (fig.2-B et 2-D).

L'axe 1 oppose également les périodes FS et SP aux périodes SF et SC (fig.2-D). L'opposition des lignes sur cet axe présente donc 2 composantes intimement imbriquées, une composante spatiale, et une composante temporelle. L'axe 2 oppose la circulation de RS à la circulation des *Pasteurella* et isole 3 villages, LP et SE à forte circulation de RS et DI à forte circulation de *Pasteurella*. Le plan 1-3 est un plan "viral", qui sépare les villages au sein de chaque région et révèle une circulation précoce de AD et RS (fig.2C et 2E).

Au total, l'analyse simple aboutit à une première caractérisation du mode de circulation des agents pathogènes. Chacun des 3 virus montre une circulation relativement indépendante de celle des 2 autres, alors que les *Pasteurella* sont très corrélées entre elles. Les effets spatiaux et temporels sont évidents, mais ne peuvent être séparés de manière optimale par un centrage des lignes sur la moyenne globale. Pour cela, la décomposition de l'inertie du tableau sur les variables qui structurent le plan expérimental est indispensable.

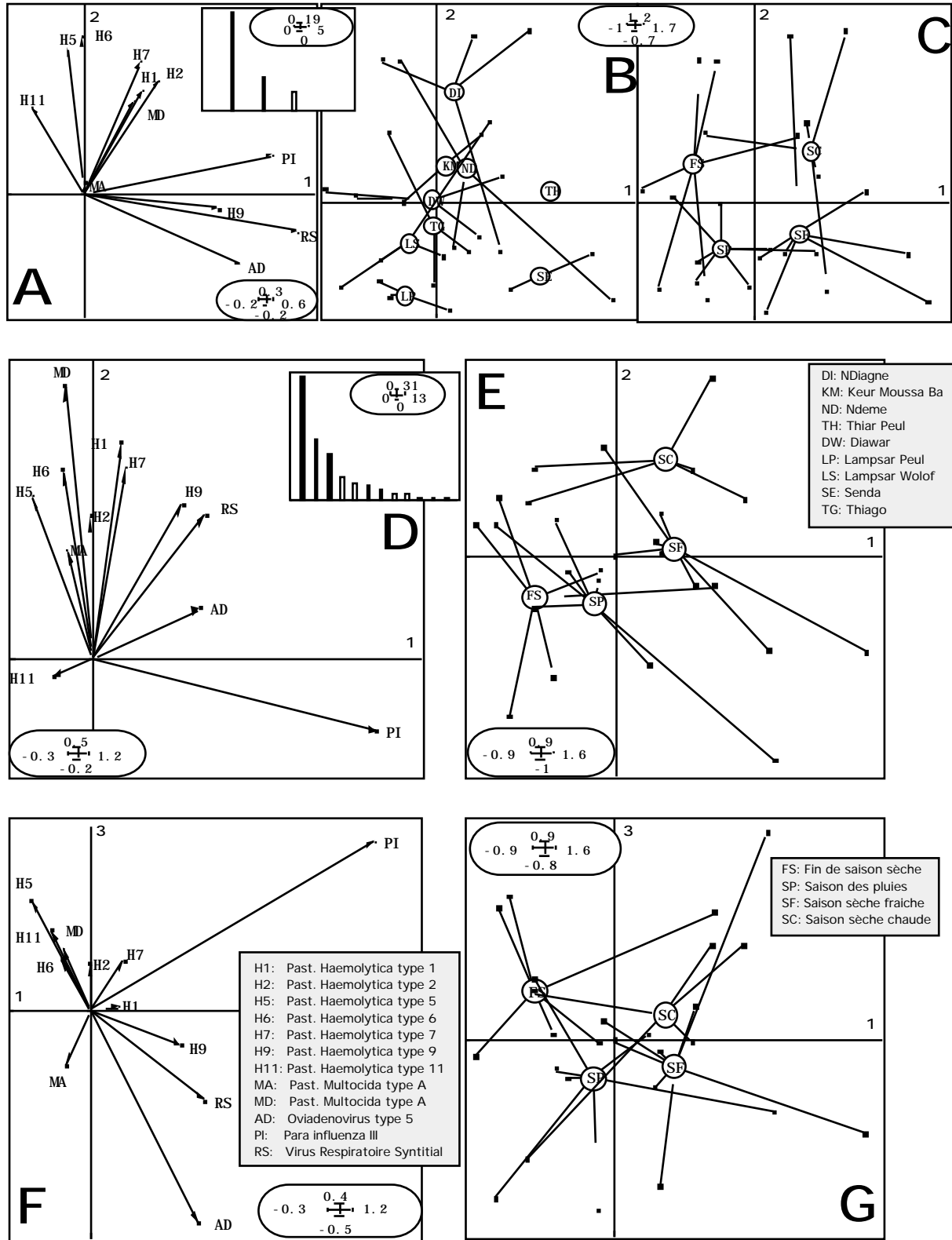
## 2.2 — Les analyses sur variables instrumentales

Les résultats sont présentés dans les figures 3 à 7. Ils sont complétés par les listings de sortie du programme et par une représentation des données centrées par village placée en annexes 7 à 11.

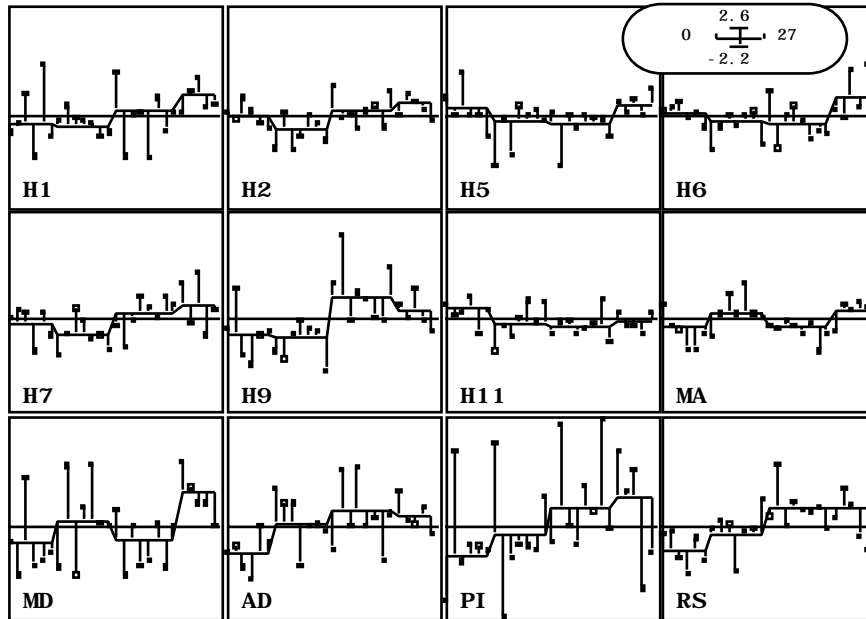
### 2.2.1 — Résultats des tests

L'inertie interpériode ne représente que 23 p. cent de l'inertie totale, mais le test de permutation effectué directement sur  $P$  est significatif au seuil de 5 p. cent. (annexe 7). Le test de l'effet période sur le tableau centré par village est très significatif ( $< 0,0001$ ). Ces résultats sont confirmés par les analyses univariées, qui montrent un facteur période significatif (niveau de signification de l'écart de déviance compris entre 0,05 et 0,0001) pour toutes les variables, à l'exception de H1 et MD. Il est donc justifié d'individualiser les périodes entre elles. L'inertie intervillage représente 38 p. cent de l'inertie totale, mais le test de permutation effectué sur  $V$  n'est pas significatif au seuil de 5 p. cent (annexe 7), ce qui ne permet pas d'envisager l'analyse intervillage sur le tableau de données brutes. L'analyse intervillage sur le tableau centré par période (interV dans intraP) analyse 46 p. cent de l'inertie intrapériode, pour laquelle un test de permutation est cette fois significatif (au seuil de 0,001). Les analyses univariées confirment ces résultats et montrent un effet village significatif pour une majorité de variables (8 sur 12), avec une probabilité supérieure à celle de l'effet période pour 3 variables, H5, MA et AD. L'individualisation des villages n'apparaît justifiée qu'une fois l'effet période pris en compte.

On envisagera successivement l'effet période, l'effet village seul, puis additionné à l'effet période.



## 2.2.2 — L'effet période (figure 3)

Figure 4. projection des données sur  $V$  et sur  $P \quad V$ 

L'analyse interpériode définit 2 axes (88 p. cent de l'inertie interpériode) et oppose sur l'axe 1 la circulation des agents viraux et sur l'axe 2 la circulation des *PasteurellaHaemolytica* (fig. 3-A). La carte B montre un échelonnement des villages sur l'axe 2. Les villages DI et LP vont participer par exemple à l'opposition qui existe entre FS-SC et SP-SF (fig. 3-C). Si l'axe "viral" parvient à séparer de manière satisfaisante FS et SP de SF et SC, une part de l'inertie interpériode maximisée sur l'axe 2 contient de l'inertie intervillage. De ce fait, cette analyse n'est pas optimale dans le cas présent car il existe une forte interaction entre les 2 variables qui définissent le plan. Il est préférable d'analyser l'effet période après avoir éliminé l'effet village, c'est à dire après avoir centré les lignes par village, par une analyse intravillage.

Cette analyse conduit à interpréter 3 axes (73 p. cent de l'inertie intravillage). Le plan 1-3 est défini par les 3 virus, peu corrélés entre eux. Les *Pasteurellas* sont regroupées en faisceau sur l'axe 2 (fig.3-D et 3-E). L'analyse intravillage fait jouer dans le cas présent un rôle clef à la saison sèche fraîche. Les positions des centres de gravité des lignes séparent 2 phases de circulation différentes (fig. 3-E), une augmentation de l'intensité de la circulation des virus jusqu'à la saison sèche fraîche, amorcée plus intensément par AD dès la saison des pluies (fig.3-G) et une augmentation de l'intensité de la circulation des *Pasteurella*, MD en particulier, à partir de la saison des pluies. L'examen des profils de coordonnées après projection sur  $V$  (analyse intravillage) et sur  $P \quad V$  (analyse inter-période du tableau centré par village) met clairement en évidence la corrélation des projections des coordonnées des 3 virus sur  $P \quad V$ , mais également l'écart entre les 2 modèles (fig. 4). La variabilité intra-village ne se résume donc pas à une variabilité interpériode, en particulier pour les 3 virus, pour lesquels la modélisation par projection sur  $P \quad V$  reste imparfaite. Ainsi, même si les taux de circulation virale décroissent rarement dans le temps (cf représentation des données centrées en annexe 7), leur progression temporelle diffère d'un village à l'autre (interaction période - village).

2.2.3 — L'effet village (fig. 5)

Le test de permutation sur *V* non significatif ne permet pas d'interpréter l'analyse intervillage. L'analyse intrapériode définit 2 axes (61 p.cent de l'inertie intrapériode) et oppose la circulation des virus PI et RS à celle des Pasteurella.

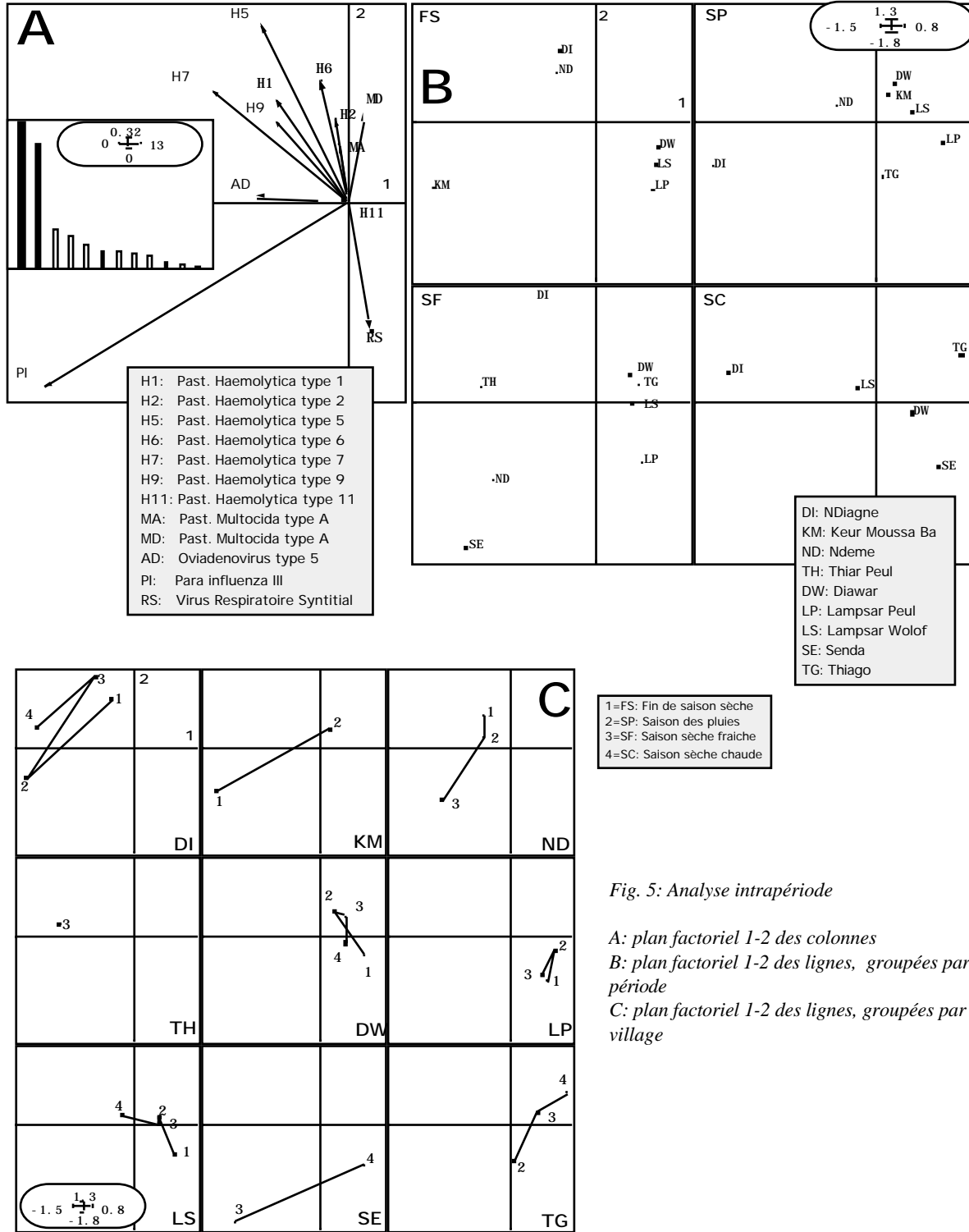


Fig. 5: Analyse intrapériode

A: plan factoriel 1-2 des colonnes  
B: plan factoriel 1-2 des lignes, groupées par période  
C: plan factoriel 1-2 des lignes, groupées par village

L'axe 3 (9 p. cent de l'inertie intrapériode) est presque uniquement défini par AD et l'axe 4 (7 p. cent de l'inertie intrapériode) par RS (annexe 7). Les 3 virus demeurent toujours peu corrélés entre eux, alors qu'au contraire, les *Pasteurella* restent très groupées sur l'axe 2.



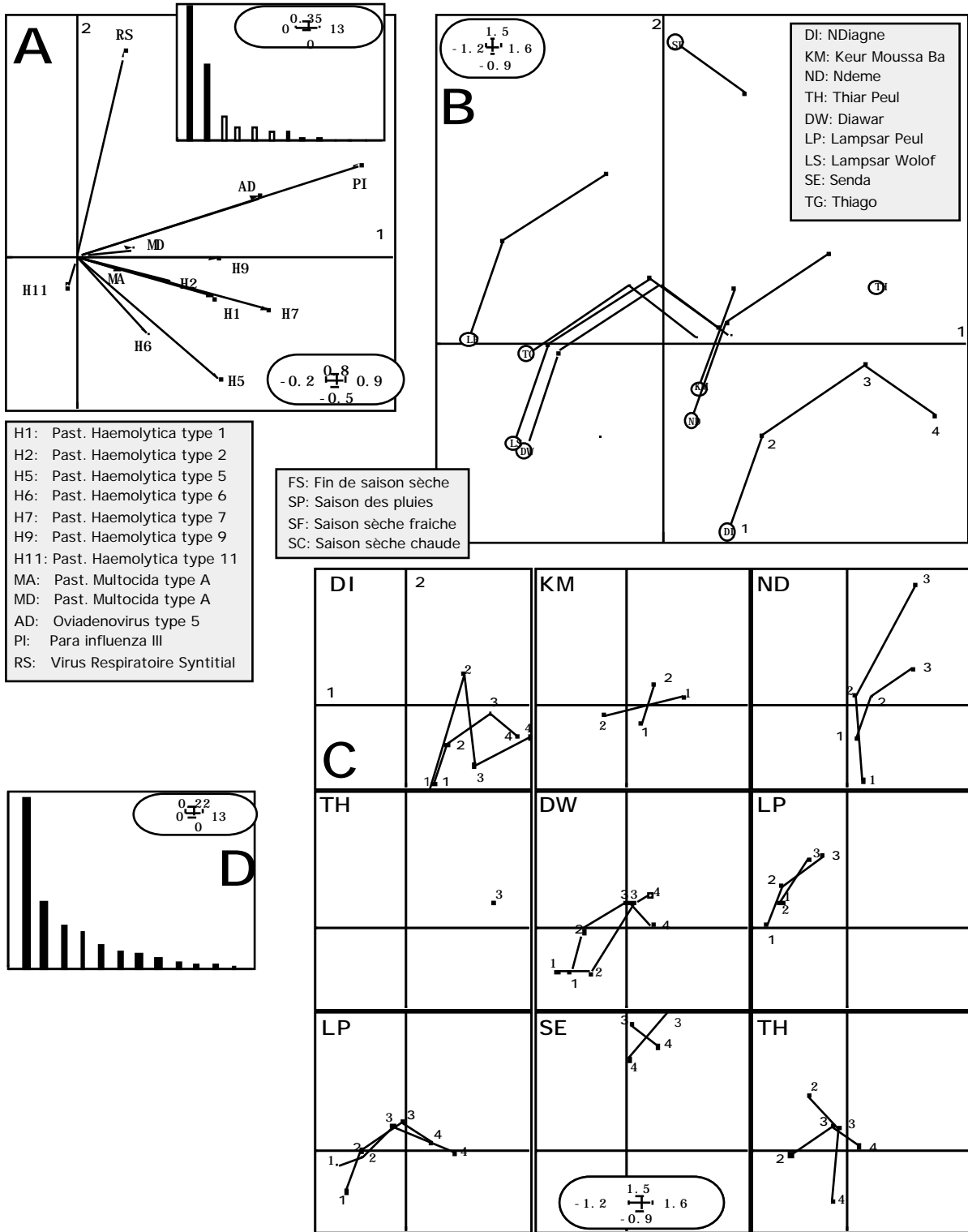


Figure 6. Analyse sur variables instrumentales. Prise en compte d'un effet additif P+V.

A: plan factoriel 1-2 des colonnes

B: plan factoriel 1-2 des lignes - regroupées en trajectoires temporelles par village

C: plan factoriel 1-2 des lignes - données du tableau initial (trait fin) et après projection sur P+V (trait gras)

D: histogramme des valeurs propres de l'AFCVI orthogonale

Le multifenêtrage par période de la figure 5-B permet de visualiser ce qui, dans le tableau centré par période, relève de l'effet village global, et de l'interaction période - village. On note une séparation des villages des 2 régions (SE excepté) sur l'axe 1, excepté en saison des pluies, et le rôle particulier de PI, qui constitue à la fois un élément séparateur des périodes (fig.3-D) et des villages.

A l'intérieur de chaque région, la discrimination des villages est constante, maximale en SF, et multiforme, due à RS et aux *Pasteurella* sur l'axe 2 (fig.5-B) et 4, et à AD sur l'axe 3 (les axes 3 et 4 n'ont pas été représentés). Le multifenêtrage par village met en évidence une plus grande instabilité de la circulation microbienne dans les villages de la région de Louga et à SE (fig.5-C).

#### 2.2.4 — L'effet période + village (fig. 6)

La décomposition de l'inertie totale du tableau sur des bases orthonormées est présentée dans le tableau VI (annexe 7). La part de l'inertie associée à  $V$  dans  $P + V$  ( $V / P$ ) est supérieure à celle associée à  $P$ , traduisant une plus forte variabilité entre villages qu'entre périodes. Les 2 premiers facteurs de l'analyse sur variables instrumentales représentent 73 p. cent de l'inertie projetée sur  $P + V$  et caractérisent 3 "directions", définies par RS, AD et PI, et le groupe des *Pasteurella haemolytica*.

Tableau VI. décomposition de l'inertie totale ( $PxV$ ) par projection sur différentes bases orthonormales (Chessel, 1995e)

Var	$PxV$	$P + V$	$P$	$V$	$P / V$	$V / P$	$(P + V)$	$P$	$V$	$V$	$P$
H1	0.0582	0.0325	0.0066	0.0198	0.0127	0.0259	0.0257	0.0105	0.0229		
H2	0.0740	0.0506	0.0247	0.0311	0.0195	0.0262	0.0234	0.0175	0.0260		
H5	0.1121	0.0779	0.0112	0.0701	0.0077	0.0667	0.0342	0.0072	0.0659		
H6	0.0794	0.0439	0.0182	0.0283	0.0156	0.0256	0.0355	0.0141	0.0235		
H7	0.1103	0.0791	0.0156	0.0609	0.0181	0.0635	0.0312	0.0155	0.0591		
H9	0.0540	0.0350	0.0167	0.0145	0.0206	0.0183	0.0190	0.0177	0.0172		
H11	0.0563	0.0280	0.0167	0.0136	0.0144	0.0113	0.0283	0.0125	0.0107		
MA	0.0722	0.0326	0.0160	0.0144	0.0183	0.0166	0.0396	0.0171	0.0159		
MD	0.0340	0.0166	0.0089	0.0056	0.0109	0.0077	0.0175	0.0097	0.0070		
AD	0.1182	0.0777	0.0294	0.0471	0.0306	0.0484	0.0404	0.0259	0.0470		
PI	0.3170	0.1500	0.0441	0.0909	0.0591	0.1060	0.1670	0.0483	0.1014		
RS	0.1532	0.1204	0.0711	0.0741	0.0463	0.0493	0.0328	0.0383	0.0485		
Total	1.2389	0.7443	0.2789	0.4705	0.2738	0.4654	0.4947	0.2344	0.4453		

On retrouve l'opposition entre les villages de la région de Louga et ceux de la région du Delta, induite par une intensité de circulation générale plus forte dans la première (effet taille). Dans chaque région, des villages se distinguent des autres sur l'axe 2 par l'intensité de la circulation de RS, H5 et H6. LP et SE, pour la région du Delta, et DI pour la région de Louga, se détachent ainsi des autres villages. Les 2 premiers segments de la courbe temporelle traduisent une augmentation de la circulation virale, et les 2 derniers au contraire une augmentation de la circulation des *Pasteurella haemolytica* dans les 2 phases de la saison sèche. On retrouve ici le rôle "charnière" de la saison sèche fraîche entre les circulations virale et bactérienne, déjà mis en évidence par l'analyse intravillage.

L'examen des résidus (fig.7) montre que la modélisation reste imparfaite en particulier pour 2 agents, PI et MD, pour lesquels la projection sur  $P + V$  ne caractérise pas plus de 50 p. cent de l'inertie totale du tableau (tableau VI, annexe 7).

La mise en oeuvre des analyses sur variables instrumentales permet d'adopter une démarche progressive dans la décomposition de la variabilité totale du tableau. En première approche, elle aboutit à une modélisation relativement satis-faisante des lignes du tableau.

Elle retrouve globalement les résultats des analyses des effets séparés (analyse intervillage et interpériode), tout en les dissociant au mieux. Elle identifie clairement la tendance à la croissance de la circulation des virus dans le temps et les fluctuations observées d'un village à l'autre. Mais elle ignore une partie non négligeable de l'inertie totale (40 p. cent) liée aux interactions entre villages et périodes, qu'on visualise sur la figure 6-C. La modélisation de l'inertie résiduelle par une analyse sur variables instrumentales orthogonale (AFCVI orthogonale) ne fait que confirmer les informations contenues dans le graphe des résidus (fig.7), à savoir un rôle prépondérant de AD, PI et MD dans l'origine de ces interactions.

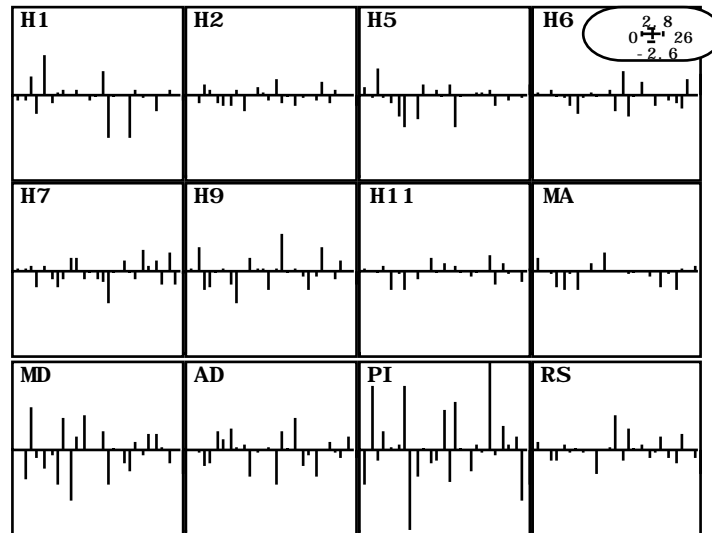


Figure 7 : résidus de la projection sur P+V

Au total, en conservant une image de la variabilité interne à l'effet étudié, l'analyse intrapériode et l'analyse intravillage renforcent la définition des positions relatives des villages et des périodes respectivement (Dolédec et al., 1987), par rapport aux analyses des effets purs (analyse interpériode, analyse intervillage, prise en compte simultanée des 2 effets). L'AFCVI aboutit à une séparation optimale des 2 effets, étroitement imbriqués entre eux. Mais la structure du tableau résiduel est faible (figure 6-D) et l'AFCVI orthogonale ne parvient pas à définir un modèle simple de l'interaction période - village, par une typologie unique des agents. Les analyses sur variables instrumentales permettent donc de mettre en évidence les interactions spatio-temporelles (fig.4 et 7), mais ne permettent pas de les analyser. Le recours aux analyses de tableaux multiples est motivé par cette insuffisance des analyses structurées.

### 2.3 — Les analyses de tableaux multiples: STATIS et ACOM

Les analyses séparées sont les AFC des 4 tableaux initiaux, munis d'une pondération commune des colonnes et d'une pondération des lignes dont la somme vaut 1 par tableau, ce qui permet de comparer les valeurs propres, bien que chaque tableau présente un nombre de lignes différent (Chessel et al., 1995d). Les plans factoriels 1-2 des analyses séparées des 4 tableaux sont présentés dans la figure 8-B et 8-C. Les structures individuelles sont très marquées avec 80 p. cent de l'inertie de chaque tableau portée par 2 axes pour 3 des 4 périodes et par 3 axes pour la saison sèche fraîche (fig 8-A). Même si on remarque des points communs comme la longueur de PI sur au moins 3 des 4 plans, l'angle que font les vecteurs AD et PI, le regroupement des *Pasteurella haemolytica* et leur direction opposée aux virus, et des différences, comme la longueur de MD, les analyses ne sont pas directement comparables deux à deux, du fait de l'orientation aléatoire des axes.

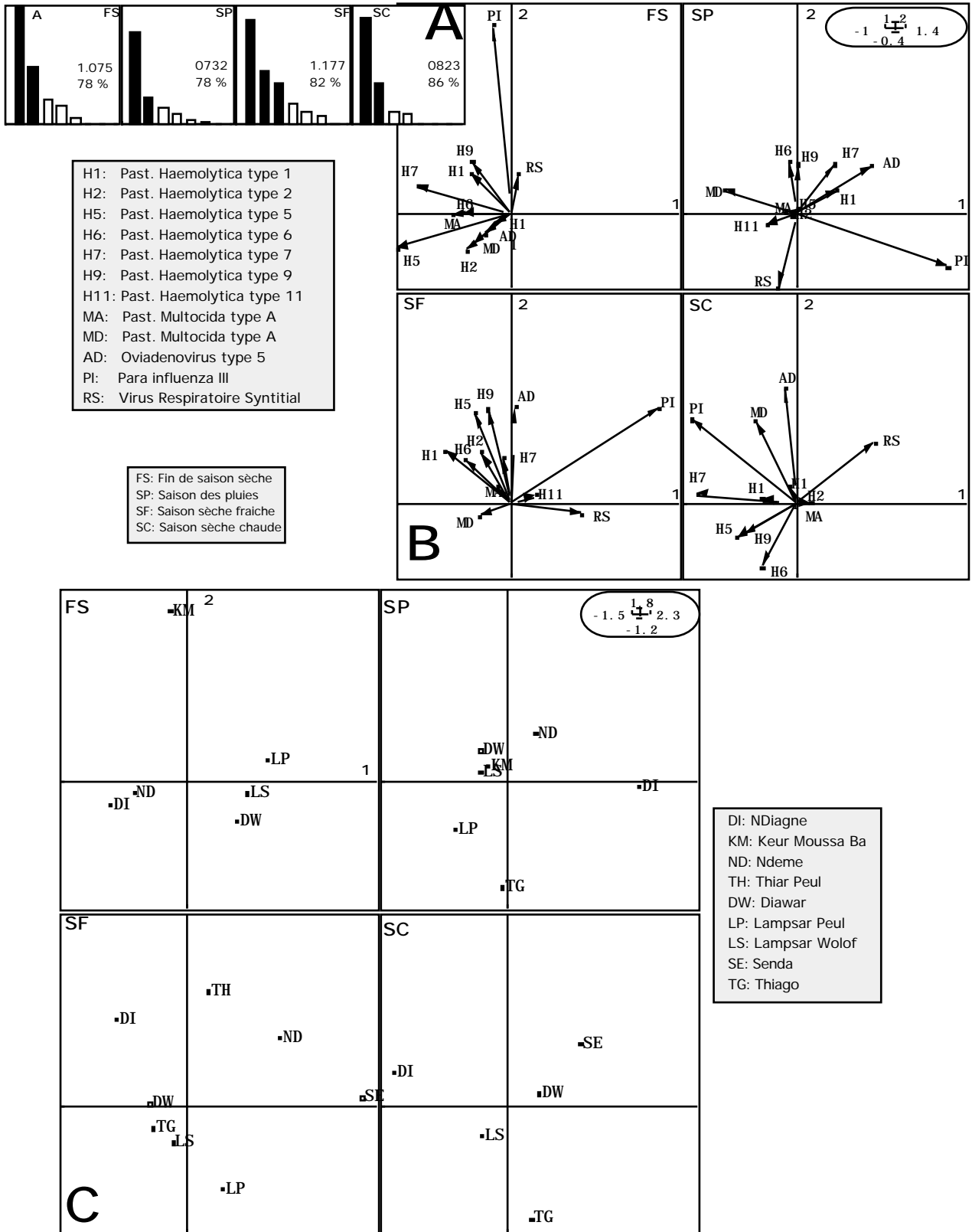


Figure 8. Analyses séparées des 4 sous-tableaux

A: valeurs propres, inertie totale et part expliquée par les premiers axes factoriels

B: plan factoriel 1-2 des colonnes des 4 sous-tableaux

C: plan factoriel 1-2 des lignes des 4 sous-tableaux

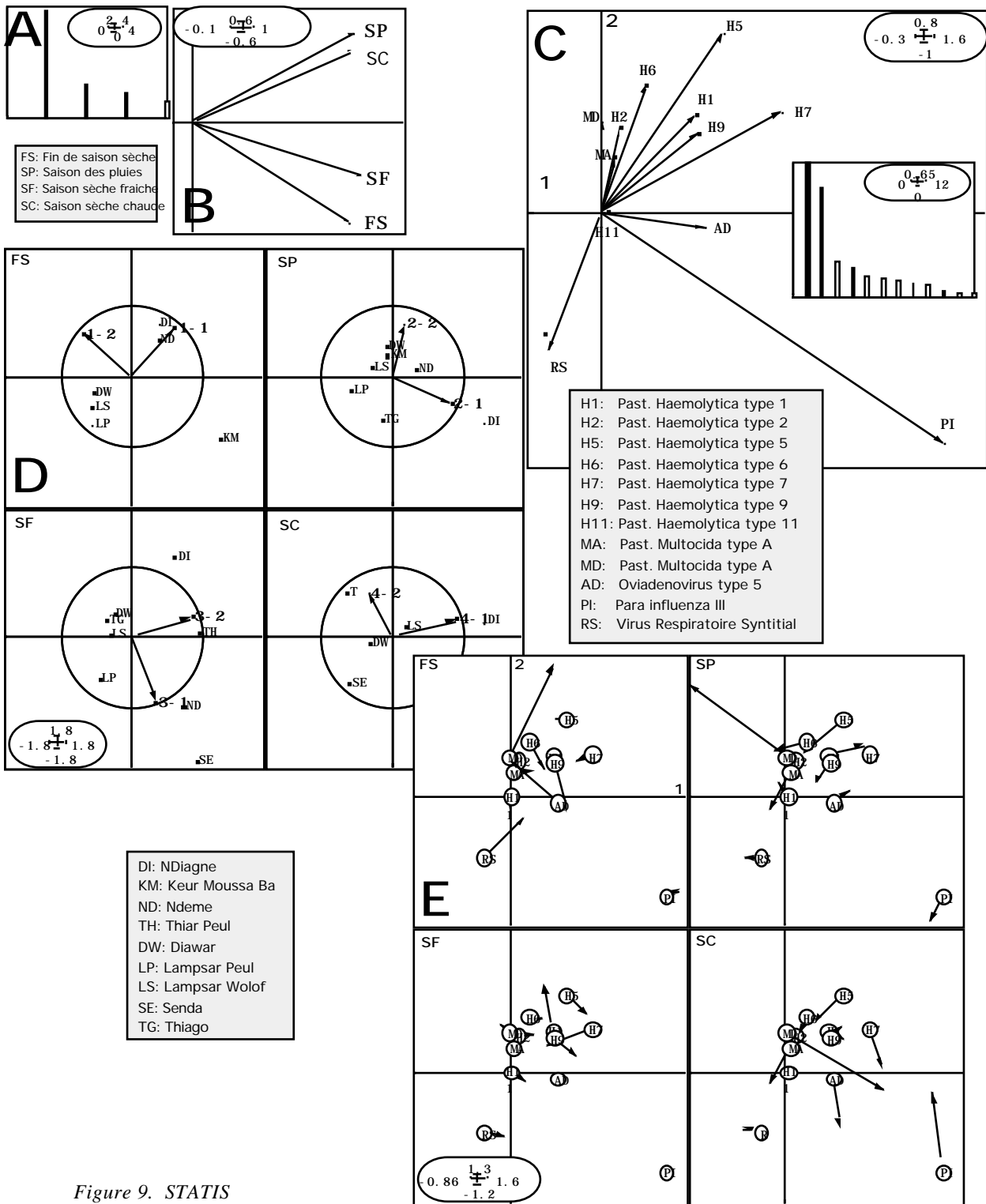


Figure 9. STATIS

STATIS - Interstructure. A: vecteurs propres de la matrice des corrélations vectorielles

B: coordonnées des tableaux sur l'axe 1

STATIS - Compromis. C: plan factoriel 1-2 des colonnes

D: projection des lignes et des axes d'inertie de chaque tableau

STATIS-aides à l'interprétation.

E: représentation des colonnes des 4 sous-tableaux dans le plan factoriel 1-2 du compromis (les étiquettes marquent les coordonnées des colonnes du compromis, les flèches les coordonnées des projections des colonnes des tableaux initiaux)

### 2.3.1 — STATIS

L'analyse de l'interstructure et du compromis de STATIS sont présentés dans la figure 9.

La diagonalisation de la matrice des coefficients RV a été retenue de manière à donner a priori à chaque tableau un poids homogène (coordonnées des tableaux sur l'axe 1 de l'interstructure) dans la constitution du compromis. Sa première valeur propre se distingue nettement des autres, justifiant la recherche d'un compromis (fig.9-A). 2 groupes de 2 périodes très corrélées entre elles s'opposent sur l'axe 2, traduisant un phénomène général d'alternance temporelle dans la circulation des agents au niveau villageois. L'analyse du compromis de STATIS aboutit à des résultats quasiment identiques à ceux de l'analyse intrapériode tant du point de vue de la structure représentée et de la typologie des agents (fig.9-C et 5-A, page 13), que de celui du positionnement des villages (fig.9-D et 5-B, page 13). La figure 9-D représente en plus la projection des axes d'inertie des 4 analyses séparées sur le plan 1-2 du compromis. Elle montre que les villages pour chaque période reproduisent avec fiabilité la structure moyenne.

STATIS permet d'analyser la stabilité de la typologie moyenne des agents. La figure 9-E représente l'erreur de reconstitution entre les coordonnées du compromis et les coordonnées des analyses séparées. Elle montre une excellente reconstitution du tableau moyen par les 4 sous-tableaux, en particulier la saison sèche fraîche (SF) et confirme le rôle clef joué par cette période, déjà mis en évidence par l'analyse intravillage (fig.3-E, page 11). Les poids, les inerties et les variances vectorielles des 4 sous-tableaux (second paragraphe, annexe 13) viennent confirmer ce constat. On note en particulier l'opposition entre la stabilité de la position des *Pasteurellahaemolytica*, et l'instabilité de la position de MD, très mal analysé par le compromis. La position de PI signe une circulation permanente de cet agent dans les villages de la région de Louga (fig.9-D), et même exclusive pour FS parmi les agents viraux, comme en témoigne le rapprochement de AD et RS de l'origine à cette période.

Au total, l'analyse intrapériode et l'analyse STATIS sur les coefficients RV aboutissent aux mêmes conclusions générales. STATIS apporte cependant 4 éléments nouveaux:

- le rôle charnière joué par SF (mis en évidence également par l'analyse intravillage)
- le rôle exclusif joué par PI en FS parmi les agents viraux
- la circulation instable de MD, que le tableau moyen interprète difficilement.
- une plus grande corrélation entre FS et SF d'une part, et SP et SC d'autre part

Cependant, STATIS ne fait pas preuve d'optimalité dans la représentation simultanée des colonnes des sous-tableaux et ne donne pas une représentation satisfaisante des trajectoires des variables (Chessel et al, 1996).

### 2.3.2 — L'ACOM

Les résultats de l'ACOM sont présentés dans la figure 10, et les aides à l'interprétation en annexe 14, ainsi qu'une représentation des données par village et par période, projetées sur les variables auxiliaires (figure 11). Les résultats des analyses séparées (fig.8-A) amènent à choisir 2 variables auxiliaires dans la construction du tableau de référence (l'analyse séparée de SF définit 3 axes), sur un maximum de 5 (dimension du plus petit des 4 sous-tableaux). La première variable auxiliaire exprime une structure stable, reproduite par chacun des tableaux (cos2 élevés). La seconde au contraire révèle un élément de structure variable en saison sèche chaude (cos2 inférieur aux autres).

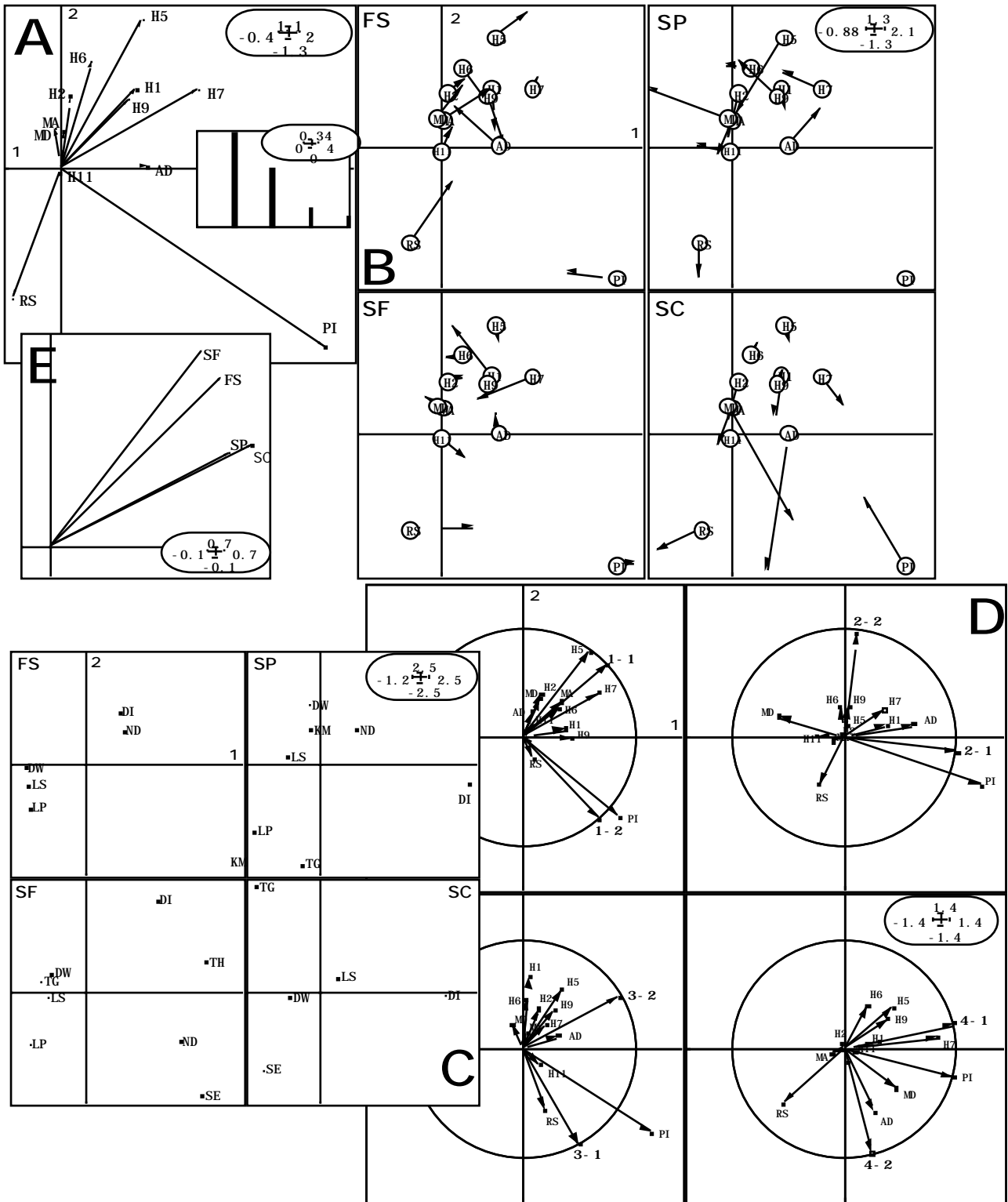


Figure 10. Analyse de CO-inertie Multiple (ACOM)

A: plan factoriel défini par les 2 premières variables auxiliaires

B: projections des colonnes des 4 tableaux dans le plan factoriel défini par les 2 premières variables auxiliaires (les étiquettes marquent les coordonnées des colonnes du tableau référence, les flèches les coordonnées des projections des colonnes des 4 tableaux)

C: représentation des 2 premiers vecteurs de co-inertie pour chaque sous-tableau (coordonnées des lignes de variance unité)

D: projection des axes d'inertie dans le plan 1-2 de co-inertie de chaque sous-tableau

E: covariances au carré entre les composantes synthétiques et les variables auxiliaires

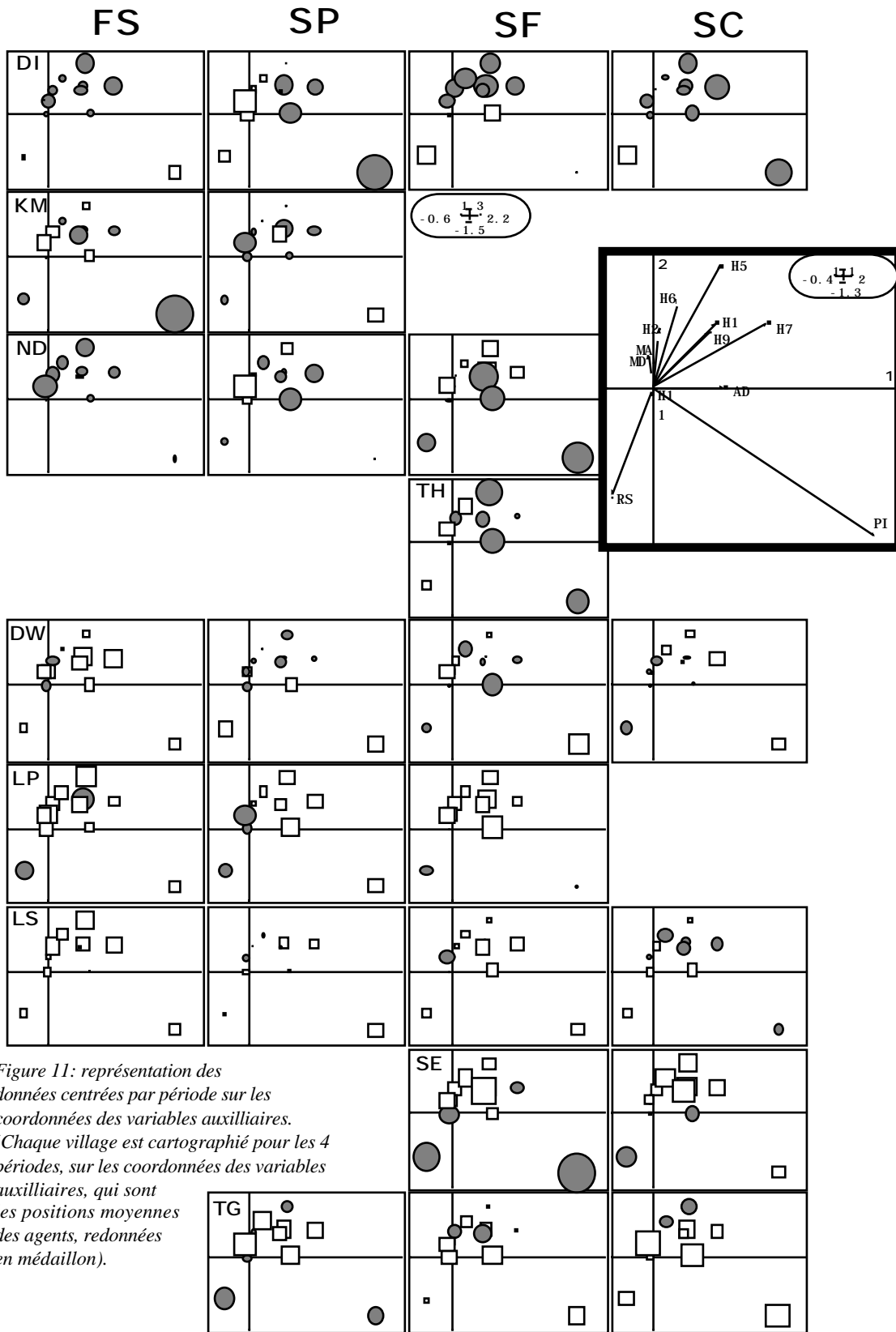


Figure 11: représentation des données centrées par période sur les coordonnées des variables auxilliaires. (Chaque village est cartographié pour les 4 périodes, sur les coordonnées des variables auxilliaires, qui sont les positions moyennes des agents, redonnées en médaillon).



Encore une fois, le plan factoriel défini par les 2 premières variables auxiliaires (non représenté) est quasiment identique au plan 1-2 de l'analyse intrapériode (à une rotation près) et au plan 1-2 du compromis de STATIS, à l'exception toutefois de MD. L'examen des pourcentages d'inertie projetée sur le plan 1-2 de co-inertie (annexe 13) montre que l'analyse, en maximisant la covariance entre les coordonnées des colonnes des sous-tableaux sur les 2 premiers axes de co-inertie et les 2 premières variables auxiliaires, a quasiment maximisé la variance interne à chaque tableau. La figure 10-D positionne les axes d'inertie des analyses séparées dans les plans 1-2 de co-inertie et confirme que l'ACOM aboutit à une bonne optimisation de l'inertie de chaque tableau. Ces éléments expliquent la ressemblance entre la carte de synthèse de l'ACOM et le plan 1-2 de l'analyse intrapériode. Leurs interprétations sont globalement identiques.

La figure 10-B représente simultanément les coordonnées normalisées des colonnes de chaque tableau (de variance unité) et les variables auxiliaires de même rang, exprimant ainsi la part de la corrélation dans la maximisation de la covariance (la longueur des flèches est inversement proportionnelle à la corrélation entre les 2 coordonnées). Contrairement à STATIS, l'ACOM fait preuve d'optimalité dans la représentation conjointe des colonnes des sous-tableaux (Chessel et al., 1996), et aboutit à une position relative des colonnes de chaque sous-tableau sur les 2 premières variables auxiliaires plus proche de la position donnée par l'analyse séparée (fig.8-B, page 17). L'ACOM fournit 4 typologies des agents différentes, comme les analyses séparées, mais cette fois ci directement comparables. Cette figure s'interprète avec la figure 10-C, qui représente les vecteurs de co-inertie (coordonnées des lignes de variance unité).

L'ACOM repère comme STATIS l'indépendance de la circulation des 3 virus, le rôle exclusif de PI en FS parmi ceux ci, la stabilité de la typologie en SF, et le comportement particulier de MD parmi le groupe des *Pasteurella* (fig.10-B) en SP et SC, et plus particulièrement l'absence de relation stable avec les autres agents. Elle apporte néanmoins des éléments nouveaux. La variabilité exprimée en saison sèche chaude par la deuxième variable auxiliaire ( $\cos^2 = 0.5268$ ) est liée à une typologie des virus différente de celle induite dans les 3 autres périodes. L'analyse met également en évidence une circulation précoce de AD et RS, dès la saison des pluies, comme l'avait déjà fait l'analyse intravillage, mais révèle en plus que cette circulation des 2 virus a lieu dans des villages différents. La circulation des *Pasteurella haemolytica* est également affinée. On constate en effet la participation de sérotypes différents selon les périodes (H5 en FS, H6 H7 et H9 en SP, H1 et H5 en SF, et H5 H6 et H7 en SC), les 2 sérotypes les mieux représentés sur les variables auxiliaires étant les types les plus fréquents (H5 et H7). Enfin, comme l'interstructure de STATIS, l'ACOM aboutit à une typologie des sous-tableaux (fig. 10-E), et met en évidence leur séparation en 2 groupes.

Tableau VII: récapitulation des principaux résultats et des analyses qui les ont mis en évidence

Résultats	AFCD	IntraV	IntraP	P+V	STATIS	ACOM
Opposition des villages des 2 régions			x	x	x	x
Discrimination des villages intra-régionale			x	x	x	x
Progression temporelle		x		x		
Rôle "charnière" de la saison sèche fraîche		x		x	x	x
Circulation des virus peu corrélée			x		x	x
Circulation très corrélée des <i>Past. Haem.</i>	x	x	x	x	x	x
Circulation précoce de RS et AD	x	x				x
Circulation particulière de MD					x	x
Circulation exclusive de PI en FS					x	x
Typologie des virus différente en SC						x
Mise en jeu de sérotypes différents de <i>P. h.</i> selon les périodes						x
Corrélation FS-SF et SP-SC					x	x

Au total, l'analyse intrapériode, STATIS et l'ACOM sont 3 méthodes équivalentes dans la représentation d'une typologie moyenne. La supériorité des analyses de tableaux multiples dans l'interprétation des interactions spatio-temporelles et des particularités typologiques propres à

chaque période est manifeste. Dans le cas présent, l'ACOM se montre plus précis dans le positionnement relatif des colonnes des sous-tableaux.

Le tableau VII résume les principaux résultats qui seront discutés dans la partie suivante.

## 3 — Discussion

### 3.1 — Les éléments pré-requis à la mise en oeuvre des analyses

#### 3.1.1 — *L'enquête sérologique*

L'enquête est réalisée en milieu expérimentalement non contrôlé, dans un réseau de troupeaux traditionnels de petits ruminants, suivis par une équipe de chercheurs pluridisciplinaire. Pour éviter les interférences ou les incompatibilités biologiques avec d'autres protocoles de recherche menés conjointement, il est nécessaire d'effectuer un sous échantillonnage aléatoire des animaux. Cela explique le nombre limité de villages échantillonnés qu'il a fallu choisir dans la présente étude parmi les villages ne recevant aucune vaccination contre les maladies étudiées. A cela s'ajoute une décroissance rapide des effectifs d'animaux prélevés, inévitable, qui oblige souvent à recenser de nouveaux animaux d'âge différent en cours d'étude (Desoutter, 1994).

Cette contrainte ne permet pas de travailler sur des cohortes d'animaux d'âge équivalent ayant un effectif suffisant, comme cela est généralement recommandé dans les études prospectives (Thrusfield, 1986), et rend nécessaire la standardisation des données. Au total, une telle enquête ne permet pas d'accéder à des données dont la précision est constante, en particulier en fin d'étude, ou dans certains villages, aux effectifs limités. C'est pourquoi l'objectif de cette étude reste essentiellement descriptif.

#### 3.1.2 — *L'indicateur de circulation*

Les taux de séroconversion classiquement décrits (Thrusfield, 1986) ne prennent généralement en compte que les animaux dont la sérologie se positive, ce qui a été également le cas dans le travail cité en référence (Desoutter, 1994). Or, un animal subit un véritable "pilonnage microbien", avec en moyenne, sur une période de 18 mois, 4 infections virales et 6 infections pasteurelliques. Il apparaît par conséquent obligatoire de prendre en compte les réinfections dans la construction d'un indicateur de circulation des agents pneumotropes au niveau villageois, si l'on veut pouvoir caractériser sa variabilité spatio-temporelle à partir d'éléments environnementaux, externes à l'animal. Néanmoins, un tel codage ne permet pas de mettre en relation les données sérologiques et les données cliniques disponibles conjointement.

#### 3.1.3 — *L'agrégation au niveau villageois*

L'analyse des taux villageois est une première étape dans l'analyse des données sérologiques. Certains sources de variabilité n'ont pu être pris en compte à ce stade, comme les différences entre espèces ou la variabilité intravillage liée aux facteurs de variation individuels, comme l'âge des animaux. Les associations virus-*Pasteurella* ne sont pas systématiques; les interactions agents - animaux sont donc mal analysées à ce niveau d'agrégation. Dans une étape ultérieure, les données individuelles seront codées selon les situations décrites dans le tableau III (page 6) et analysées en suivant la même démarche méthodologique que celle mise en oeuvre sur les taux villageois, mais cette fois ci par l'analyse des correspondances multiples (ACM).

### 3.2 — La méthodologie

#### 3.2.1 — *Choix d'une méthode d'analyse de base d'un tableau de taux*

La recherche d'une synergie méthode-donnée est un élément clef de l'analyse des données (Chessel et al., 1995a). Le choix des méthodes d'analyse multivariée est explicite: chaque prélèvement sanguin fait l'objet de 12 analyses sérologiques différentes. Une approche univariée apparaît trop limitative dans l'étude d'une pathologie dont l'apparition nécessite l'association de plusieurs agents infectieux. Face à un tableau de taux, plusieurs méthodes classiques ont été successivement envisagées, l'Analyse en Composantes Principales (ACP) des taux  $t_{ij}$ , l'Analyse factorielle des Correspondances (AFC) des numérateurs  $x_{ij}$ , et l'AFC des  $t_{ij}$ , avant d'adopter l'AFC décentrée (AFCD) sur une marge imposée.

Dans l'ACP des taux  $t_{ij}$ , on considère un tableau homogène  $T$ , de terme général  $t_{ij}$  exprimé dans une seule unité (% dans le cas présent), et pour lequel on a pas d'hypothèse de centrage à priori. Ce tableau doit être considéré comme la résultante d'un modèle évident et d'un modèle caché (Devillers et al., 1995). Le modèle évident est souvent lié à la nature des variables analysées, et constitue un effet trivial qui doit être éliminé par centrage, ce qui évite à l'analyse d'être influencée par lui. Le tableau  $T$  contient certainement un modèle évident, en particulier un effet colonne. On s'aperçoit en effet, que les taux de circulation pour certains agents pathogènes comme H11 ou MA sont toujours plus importants que pour d'autres, comme MD (annexe 5). Cet effet n'a pas grand intérêt en soi, et il est nécessaire de l'éliminer. Mais le centrage par colonne dans le cas présent n'est pas valide car la moyenne de taux ayant des dénominateurs différents n'est pas une moyenne arithmétique mais une moyenne de variables quotients (Cochran, 1977 - Scherrer, 1984). Elle s'établit par:  $\bar{t} = \frac{\sum_{i=1}^n x_{ij}}{\sum_{i=1}^n r_i}$  et non par

$\bar{t} = \frac{\sum_{i=1}^n t_{ij}}{n}$ . L'ACP des taux n'est donc pas satisfaisante d'un point de vue mathématique.

Dans l'AFC des  $x_{ij}$  et l'AFC des  $t_{ij}$ , on considère un tableau  $X$  et  $T$  à  $n$  lignes et  $p$  colonnes, de terme  $x_{ij}$  et  $t_{ij}$  respectivement, qu'on analyse par AFC. Un animal peut présenter une augmentation du titre d'anticorps vis à vis de plusieurs agents simultanément. La marge des lignes associée à l'AFC des  $x_{ij}$  est donc très différente des effectifs d'animaux prélevés (figure 12).

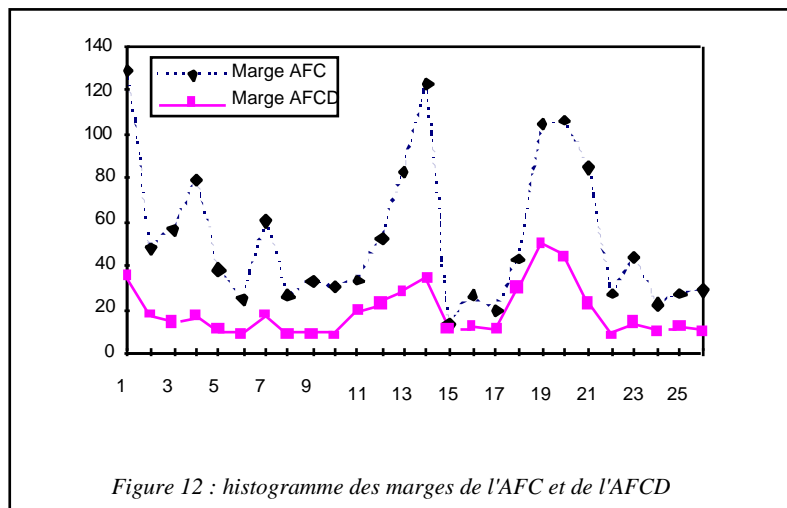


Figure 12 : histogramme des marges de l'AFC et de l'AFCD

Dans ces conditions, le profil des lignes représente pour un village la proportion de cinétiques ascendantes vis à vis d'un agent sur la totalité des cinétiques ascendantes tous agents confondus, et le profil colonne, la proportion de cinétiques ascendantes vis à vis d'un agent sur la totalité des cinétiques ascendantes tous villages confondus. L'unité élémentaire d'étude n'est

plus l'animal mais la cinétique ascendante, et on ignore ainsi les animaux ayant présenté une stabilisation ou une diminution du titre. L'analyse de ce tableau donnerait une représentation de la forme des cinétiques ascendantes par village sans nous renseigner sur le niveau de ces cinétiques par rapport à l'effectif d'animaux prélevés.

La définition du point de référence (origine) des axes est fixée par les objectifs: comparer le niveau des cinétiques ascendantes par rapport à un effectif d'animaux prélevés, entre villages et entre périodes. L'AFC classique n'est donc pas satisfaisante car elle ne répond plus aux objectifs fixés.

Plusieurs auteurs ont utilisé l'AFC sur un tableau de pourcentages (Botros, 1983 - Boukherissa, 1992), dans des cas où les colonnes ou les lignes s'excluaient mutuellement (causes de mortalité, catégories socio-professionnelles en colonne, années en ligne...). La somme des éléments d'une ligne ou d'une colonne conservait alors un sens face aux objectifs fixés (proportion de mortalité par année, par cause ...). Dans le cas présent, cette analyse cumule les défauts des 2 précédentes (absence de sens biologique dans la somme des éléments d'une ligne ou d'une colonne, et absence de sens mathématique car les dénominateurs des taux sont différents d'un village à l'autre).

Dans l'AFCD, le profil ligne est un taux, et la marge des colonnes une moyenne de variables quotients. Cette méthode est donc applicable en présence de taux ou de quotients, dont les dénominateurs sont constants entre colonnes et différents entre lignes. Avec des dénominateurs constants entre colonnes et entre lignes, l'ACP centrée par colonne, l'ACP centrée par ligne ou l'ACP doublement centrée sont également disponibles, en plus de l'AFCD, qui ajoute au centrage par colonne une pondération des variables (agent pathogène) par l'inverse de leur moyenne générale.

### 3.2.2 — Les analyses structurées

L'avantage essentiel qu'elles présentent réside dans la multiplicité des décompositions de l'inertie totale du tableau (effets principaux, effets complémentaires, effets additifs), permise par l'approche dissymétrique. En conservant dans l'inertie intravillage ou intrapériode une part de variabilité interne aux périodes ou aux villages respectivement, l'analyse intraclasse donne une représentation mieux ajustée des différences entre périodes ou entre villages, que celle de l'analyse interclasse correspondante (Dolédec et al., 1987), et permet de révéler une interaction entre effets, par comparaison des profils lignes modélisé et résiduel. Cependant, l'analyse intraclasse construit une typologie unique des variables, qui ne lui permet pas d'apporter les éléments nécessaires à l'interprétation des interactions.

### 3.2.3 — Les analyses de tableaux multiples

Le centrage utilisé rend les analyses de tableaux multiples complémentaires de l'analyse de l'effet période. L'analyse intrapériode, l'analyse du compromis de STATIS et les variables auxiliaires de l'ACOM décrivent la structure moyenne du tableau centré par période de manière identique. Ces analyses poursuivent par des méthodes différentes un objectif commun, la recherche de la partie commune aux analyses séparées, et sont donc redondantes sur ce point. Les analyses de tableaux multiples et les centrages retenus conduisent non seulement à une caractérisation de la circulation globale (en moyenne) des agents à l'intérieur de chaque période, tout comme l'analyse intrapériode, mais en plus, à une caractérisation des particularités de chaque période (écart à cette moyenne), conduisant à une description complète des interactions spatio-temporelles, grâce à l'interprétation conjointe des positions des lignes et des colonnes.

STATIS conduit à une bonne synthèse des  $n_k$  villages pour chacun des tableaux, positionnés directement par projection sur les vecteurs propres du compromis ( $z_j$ ), mais ne conduit pas à une représentation optimale des  $p_k$  variables, car les  $j$  vecteurs  $a_j^k$  ne sont pas orthogonaux entre eux (Chessel et al., 1995d). L'ACOM, par construction, aboutit à une bonne

représentation des  $n_k$  village sur les axes de co-inertie  $u_k^m$ , et par sa fonction optimisée, à une représentation satisfaisante des  $p_k$  colonnes par projection sur les axes de co-inertie ( $X_k^t D_k u_k^m$ ), sans toutefois assurer la non-corrélation des coordonnées (Chessel et al., 1996).

Aucune méthode n'est supérieure à une autre (Chessel et al., 1996) et le choix de l'une ou l'autre dépend des objectifs de l'étude. Dans le cas présent, s'il s'agit de comparer des typologies d'agents pathogènes, on retiendra l'ACOM (figure 10-B). Au contraire, s'il s'agit de comparer des typologies de villages, on se tournera soit vers l'analyse intrapériode (fig.5), soit vers STATIS (fig.9) soit vers l'ACOM (fig.10). En revenant aux objectifs de ce travail, on peut toutefois proposer un enchaînement "type" d'analyses complémentaires permettant de retrouver la majorité des résultats décrits dans le tableau VII:

- une analyse intravillage, pour la caractérisation de l'effet période et la mise en évidence des interactions spatio-temporelles,

- un centrage des données par période, suivi d'une ACOM pour la description de la structure moyenne, la caractérisation optimale des particularités typologiques propres à chaque période, et l'analyse précise des interactions spatio-temporelles.

### 3.3 — Discussion épidémiologique

#### 3.3.1 — Les virus

Dans les conditions de l'élevage traditionnel du Sénégal, les 3 virus reconnaissent une durée de vie des anticorps post-infectieux d'environ 6 mois (Desoutter, 1994), durée de vie relativement courte qui explique le mode de circulation quasi permanent, avec la survenue de très fréquentes ré-infections. On note toutefois 2 périodes préférentielles de circulation, la saison des pluies, en particulier pour AD, et la saison sèche, avec un effet particulièrement net dans les villages de la région de Louga.

PI joue en permanence un rôle prépondérant dans l'analyse. Son rythme de passage est fixé, en partie, par une action directe des facteurs climatiques. On sait en effet qu'une humidité relative faible augmente la stabilité du virus dans le milieu extérieur et qu'une vitesse moyenne des vents augmente sa dissémination (Lefèvre, 1987). Ces éléments conduisent Desoutter (1994) à proposer un modèle de l'infection par PI, que l'on retrouve avec l'analyse de l'effet additif période-village.

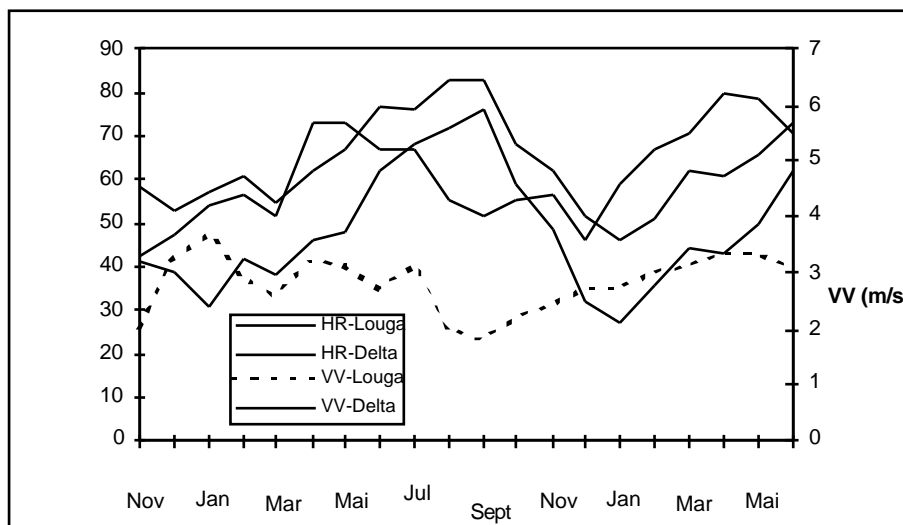


Figure 13 : Taux d'humidité relative (HR) et vitesse moyenne des vents (VV).

En SP, l'humidité relative augmente (fig.13), gênant la circulation de PI. En SF, l'humidité relative diminue, le nombre d'animaux excréteurs augmente et atteint un seuil compatible avec sa transmission efficace, favorisée par l'apparition des vents secs provenant du désert (harmattan). Enfin, en SC, la circulation est réduite par la protection immunitaire des troupeaux. La circulation plus faible dans les villages de la région du Delta suggère l'implication d'autres facteurs. Le mode de conduite des animaux spécifique de cette région, comme l'entretien en élevage de case d'un plus grand nombre d'ovins et la pratique d'une complémentation alimentaire plus intensive, a été évoqué, ainsi que l'existence de souches de PI différentes dans le Delta, atténuées par une propagation constante en zone à humidité relative plus élevée (Desoutter, 1994).

AD et RS présentent des profils moyens d'évolution temporelle très voisins. Pour ces 2 virus, les conditions climatiques (pluies, vents) pourraient jouer cette fois un rôle non plus direct mais indirect en constituant des facteurs de stress chez les animaux. La distinction en saison sèche des villages LP et SE dans la région du Delta et de DI dans la région de Louga serait liée, au moins en partie, au rôle amplificateur joué par les chèvres (nombreuses à LP et SE, absentes à DI), dans la propagation de RS (Desoutter, 1994). AD semble se transmettre d'autant plus intensément dans les villages que la circulation des autres virus faiblit ou adopte un profil minimal, exception faite de SC, où AD et RS présentent une circulation importante à SE, ce qui explique la typologie différente des virus pour cette période.

De manière générale, la circulation des 3 agents viraux montre une relative indépendance. L'explication pourrait être trouvée dans la co-existence de 2 phénomènes: (i) une alternance temporelle de la circulation, liée au niveau de protection immunitaire temporaire des animaux dans le village suite au passage d'un agent, comme le montre les figures 9-B et 10-E, mais également PI à DI, KM, et SE, RS à TG ou AD à DW; (ii) des contrastes spatiaux liés à des facteurs environnementaux comme les pratiques des éleveurs (entretien de troupeaux mixtes pour RS, exposition des animaux aux vents ...). L'alternance temporelle demeure toutefois un phénomène plus rare pour RS (figure 11), ce qui traduit la possibilité de ré-infections en présence d'anticorps, phénomène déjà décrit par Baker et al. chez les bovins (1986).

### 3.3.2 — Les pasteurelles

Les 2 espèces de *Pasteurella* sont les constituants classiques de la flore nasale normale des petits ruminants, qu'ils acquièrent dès leur plus jeune âge, par contact avec leur mère. Les anticorps colostraux pasteurelliques disparaissent très rapidement au cours des 3 premières semaines de vie (Desoutter, 1994). Sous certaines conditions de stress, certains sérotypes pourraient croître préférentiellement et se propager au niveau pulmonaire, déterminant alors une réaction immunitaire de l'hôte, visualisée par l'augmentation du titre sérologique (Confer et al., 1983). Cette source de contagion endogène pourrait expliquer la forte corrélation observée dans la circulation des *Pasteurella*. Toutefois, contrairement à des observations effectuées en pays tempéré (Thompson et al., 1977), il n'a pas été possible de noter la circulation préférentielle d'un même sérotype. Au contraire, on constate l'intervention de sérotypes différents selon les périodes, ce qui pourrait traduire non seulement l'existence de phénomènes de compétition dans les capacités invasives des bactéries au niveau des muqueuses nasales (Desoutter, 1994), mais également le rôle inhibiteur de la réaction immunitaire en retour (Frank, 1985).

La circulation des *Pasteurella* commence en saison des pluies, et atteint un maximum en saison sèche chaude, avec une saison de décalage sur la circulation virale. Cette observation traduirait le délai imposé d'une part par la colonisation nasale initiale, d'autre part, par l'attente de facteurs aptes à déclencher l'infection pulmonaire pasteurellique, comme la circulation de virus pneumotropes (Desoutter, 1994). Parmi ceux-ci, AD est le virus le plus corrélé avec les *Pasteurella*, et semblerait donc jouer un rôle majeur. RS n'est jamais associé avec les *Pasteurella*, et PI rarement (avec H7 en SC). MD montre une alternance temporelle dans une majorité de village, expliquant son instabilité dans la typologie (figure 11). Cette circulation particulière par rapport au groupe des *Pasteurella haemolytica* pourrait étayer l'hypothèse d'un pouvoir pathogène propre de MD dans l'initiation de pneumopathies chez les ovins en zone sahélienne.

## Conclusion

Pour cette étude, le choix d'une méthodologie a été orienté par 3 critères, la nature des données (variables multiples, données agrégées en taux), leur organisation temporelle (4 périodes), et un objectif essentiellement descriptif. Ces critères ont conduit à retenir comme méthode d'analyse de base une extension de l'analyse factorielle des correspondances, l'AFC décentrée, particulièrement bien adaptée à l'analyse d'un tableau de taux, que l'on a ensuite généralisée à k tableaux (analyses sur variables instrumentales, STATIS, ACOM). Les résultats retrouvent une grande partie des conclusions formulées par Desoutter (1994), sans toutefois fournir des éléments précis sur les relations virus-bactéries, que seule l'analyse des données individuelles permettra d'obtenir.

Pour la prise en compte de l'organisation spatio-temporelle, 2 groupes d'analyses ont été utilisés. Les analyses sur variables instrumentales sont faciles à mettre en oeuvre. Elles permettent de caractériser rapidement les effets principaux, leurs associations et d'identifier l'existence d'interactions entre eux, sans offrir toutefois une méthode optimale pour les analyser. Les analyses de tableaux multiples sont plus évoluées et complémentaires des premières. Elles sont efficaces dans l'analyse des particularités typologiques propres à chaque sous-tableau, et des interactions spatio-temporelles.

Le type de données analysées et les problèmes méthodologiques abordés dans cette étude sont fréquemment rencontrés en épidémiologie (analyse de taux de mortalité par cause, de taux de morbidité par maladie ...). L'intérêt majeur de ce travail réside dans la formulation de méthodes généralisables à d'autres jeux de données similaires, plus volumineux. Ainsi, la même démarche sera prochainement utilisée sur des données issues d'une enquête écopathologique réalisée en Bretagne par le laboratoire d'écopathologie de l'INRA du centre de Theix (Faye et al., 1989) en vue d'établir une typologie des pathologies dans les troupeaux laitiers, de tester sa stabilité dans le temps, et ses relations avec les pratiques d'élevage.

Les résultats pourraient être ensuite avantageusement comparés à d'autres procédures, basées sur des extensions du modèle linéaire généralisé (Generalized Estimating Equations), récemment développées par Diggle et al., (1994) pour l'analyse de mesures discrètes répétées dans le temps, ou par Norleans (1995) pour l'analyse de variables discrètes multiples.

## Bibliographie

Baker J.C., Ames T.R., Markham J.F., 1986, Sero-epizootologic study of respiratory syncytial virus in a dairy herd, *Am. J. Vet. Res.*, 47, 2, 240-245.

Benzecri J.P., 1973, L'analyse des données. II. L'analyse des correspondances, Bordas (Ed.), Paris, 620 p.

Botros F., 1983, Causes de mortalité des enfants de 0 à 1 an; leur évolution par sexe en France de 1950 à 1969, *Les Cahiers de l'Analyse des Données*, 8, 2, 233-236.

Boukherissa E., 1992, Analyses de tableaux de pourcentages calculés d'après des enquêtes: loisirs de la jeunesse et possession d'actifs dans les patrimoines, *Les Cahiers de l'Analyse des Données*, 17, 2, 159-178.

Carrol J.D., 1968, A generalization of canonical correlation analysis to three or more sets of variables, In *Proceedings of the 76th Convention of the American Psychological Association*, 3, 227-228.

Chessel D., Dolédec S., 1993a, ADE version 3.6: Hypercard ©Stacks and Programme library for the Analysis of Environmental Data. Manuel d'utilisation. 8 fascicules. URA CNRS 1451, Université Lyon 1, 750 p.

Chessel D., Mercier P., 1993b, Couplage de triplets statistiques et liaisons espèces - environnement, In *Lebreton J.D., Asselain B. (Eds.), Masson, Paris, 15-43.*

Chessel D., Thioulouse J., Hanafi M., 1995a, STATIS et ordination de plusieurs ensemble de relevés, In *Programmthèque ADE-4, Analyses multivariées et expression graphique des données environnementales, fascicule 6, Méthodes classiques d'ordination simultanée de k tableaux, Institut d'Analyse des systèmes Biologiques et Socio-économiques, Université Claude Bernard Lyon I, 33 p.*

*ftp://biom3.univ-lyon1.fr/pub/mac/ADE/ADE4/Docthem/Thema-6.2.hqx*

Chessel D., Gaertner J.C., 1995b, Analyses des correspondances et k tableaux, In *Programmthèque ADE-4, Analyses multivariées et expression graphique des données environnementales, fascicule 9, Institut d'Analyse des systèmes Biologiques et Socio-économiques, Université Claude Bernard Lyon I, 24 p.*

*ftp://biom3.univ-lyon1.fr/pub/mac/ADE/ADE4/Docthem/Thema-9.2.hqx*

Chessel D., Hanafi M., 1995c, ACOM et ordination de k ensemble de relevés, In *Programmthèque ADE-4, Analyses multivariées et expression graphique des données environnementales, fascicule 7, Co-inertie de k tableaux, Institut d'Analyse des systèmes Biologiques et Socio-économiques, Université Claude Bernard Lyon I, 20 p.*

*ftp://biom3.univ-lyon1.fr/pub/mac/ADE/ADE4/Docthem/Thema-7.3.hqx*

Chessel D., Thioulouse J., Hanafi M., 1995d, STATIS et valeur typologique des groupes faunistiques, In *Programmthèque ADE-4, Analyses multivariées et expression graphique des données environnementales, fascicule 6, Méthodes classiques d'ordination simultanée de k tableaux, Institut d'Analyse des systèmes Biologiques et Socio-économiques, Université Claude Bernard Lyon I, 29 p.*

*ftp://biom3.univ-lyon1.fr/pub/mac/ADE/ADE4/Docthem/Thema-6.1.hqx*

Chessel D., 1995e, Ordinations sous contraintes, In *Programmthèque ADE-4, Analyses multivariées et expression graphique des données environnementales, fascicule 5, Variables instrumentales, Institut d'Analyse des systèmes Biologiques et Socio-économiques, Université Claude Bernard Lyon I, 40 p.*



<ftp://biom3.univ-lyon1.fr/pub/mac/ADE/ADE4/Docthem/Thema-5.5.hqx>

Chessel D., Hanafi M., 1996, Analyses de la co-inertie de  $k$  nuages de points, *Rev. Stat. Appl.*, 44, 2, 35-60.

Cochran W.G., 1977, *Sampling techniques*, Wiley J. & sons (Ed.), New York, 428 p.

Collett D., 1991, *Modelling binary data*, Chapman & Hall (Eds.), London, 369 p.

Confer A.W., Corstvet R.E., Panciera R.J., Rummage J.A., 1983, Isolation of *P. haemolytica* and correlation with serum antibody response in clinically normal beef calves, *Vet. Microbiol.*, 8, 601-610.

Dagnelie P., 1975, *Théorie et méthodes statistiques*, Les Presses Agronomiques de Gembloux (Ed.), Gembloux, vol. 2, 463 p.

Davies D.H., Davies G.B., Price M.C., 1980, A longitudinal serological survey of respiratory virus infections in lambs, *N.Z. Vet. J.*, 2, 125-127.

Desoutter D., 1994, *Les pneumopathies des petits ruminants au Sénégal: étude séro-épidémiologique et bio-géographique*, Thèse de Doctorat d'Université, Université Paris XII Val de marne - Créteil, 189 p. et annexes.

Devillers J., Chessel D., 1995, Can the enucleated rabbit eye test be a suitable alternative for the *in vivo* eye test? A chemometrical response, *Toxicology modelling*, 1, 21-34.

Diggle P.J., Liang K.Y., Zeder S.L., 1994, *Analysis of longitudinal data*, Clarendon Press (Ed.), Oxford, 253 p.

Dolédec S., Chessel D., 1987, Rythmes saisonniers et composantes stationnelles en milieu aquatique. I. Description d'un plan d'observation complet par projection de variables, *Acta Oecologica, Oecol. Gener.*, 8, 3, 403-426.

Dolédec S., Chessel D., 1989, Rythmes saisonniers et composantes stationnelles en milieu aquatique. II. Prise en compte et éliminations d'effets dans un tableau faunistique, *Acta Oecologica, Oecol. Gener.*, 10, 3, 207-232.

Dolédec S., Chessel D., Olivier J.M., 1995, L'analyse des correspondances décentrée: application aux peuplements ichtyologiques du Haut Rhône, *Bull. Fr. Pêche Piscic.*, 336, 29-40.

Doutre M.P., Perreau P., 1981, Le portage des *Pasteurellasp.* et de *Mycoplasmaarginini* chez les moutons sains au Sénégal, *Rév. Elev. Méd. Vét. pays Trop.*, 34, 4, 365-368.

Escoufier B., 1981, Traitement des questionnaires avec non réponse, analyse des correspondances avec marge modifiée et analyse multicanonique avec contrainte, *Rapport IRISA n° 146*, 36 p.

Faugère O., Faugère B., 1986, Suivi de troupeaux et contrôle des performances individuelles des petits ruminants en milieu traditionnel africain, *Rev. Elev. Méd. Vét. pays Trop.*, 39, 1, 29-40.

Faye B., Barnouin J., Lescourret F., 1989, Objectifs principaux et stratégies de l'enquête écopathologique Bretagne sur la vache laitière, *Epidémiol. Santé Anim.*, 15, 23-31.

Frailé L., Escoufier Y., Raibaut A., 1993, Analyses des correspondances de données planifiées: étude de la chémotaxie de la larve infestante d'un parasite, *Biometrics*, 49, 1142-1153.

Frank G.H., 1985, Serotype-specific resistance to nasal colonization by *P. haemolytica* in cattle, *Am. J. Vet. Res.*, 46, 11, 2245-2248.

Lavit Ch., 1988, Analyse conjointe de tableaux quantitatifs, Masson (Ed.), Paris, 240 p.

Lebart L., 1979, Exemple d'analyse d'un tableau dont l'une des colonnes a un poids prédominant, *Les Cahiers de l'Analyse des Données*, 4, 4, 417-422.

Lebart L., Morineau A., Piron M., 1995, Statistique exploratoire multidimensionnelle, Dunod (Ed.), Paris, 439 p.

Lefevre P.C., 1987, Recherches sur la répartition bio-géographique de 2 virus des petits ruminants sur le continent Africain: influence des facteurs écologiques, Thèse de Doctorat es-science, Université Paris XII, 247 p.

L'Hermier Des Plantes H., 1976, Structuration des tableaux à trois indices de la statistique, théorie et application d'une méthode d'analyse conjointe, Thèse d'Université, Université du Languedoc, 97 p.

Manly B.F.J., 1991, Randomization and Monte Carlo methods in biology, Chapman & Hall (Eds.), London, 281 p.

Nguyen-Ba-Vy, Leforban Y., Gillet J.P., Thery P., 1988, Identification d'*oviadénovirus type 5* sur les chèvres du Sénégal, *Rév. Elev. Méd. Vét. pays Trop.*, 41, 1, 35-39.

Norleans M.X., 1995, Two simultaneous inference procedures for multiple responses on arbitrary scales, unpublished manuscript, Mark\_X\_Norleans%notes@sb.com.

Norusis M.J., 1992, SPSS/PC+ Advanced Statistics 5.0, SPSS Inc (Ed.), Chicago, 481 p.

Rumeau-Rouquette C., Blondel B., Kaminski M., Bréart G., 1993, Epidémiologie, méthodes et pratiques, Flammarion (Ed.), Paris, Collection Statistique et biologie et en médecine, 312 p.

Sarr J., Diop M., Cissocko S., 1987, Association entre *adénovirus* apparentés au type 2 bovin et virus de la *peste des petits ruminants* dans le complexe pneumopathique chez le mouton et la chèvre en région sahélienne, réf ISRA N<sup>o</sup>54/viro.

Scherrer B., 1984, Biostatistique, Ga'tan Morin (Ed.), Montréal, 850 p.

Thrusfield M., 1986, Veterinary epidemiology, Butterworths & co (Ed.), London, 280 p.

Thioulouse J., Chessel D., 1987, Les analyses multitableaux en écologie factorielle. I. De la typologie d'état à la typologie de fonctionnement par l'analyse triadique, *Acta oecologica - Oecol. gener.*, 8, 4, 463-480.

Thomson D.A., Fraser J., Gilmour N.J.L., 1977, Serotypes of *P. haemolytica* in ovine pasteurellosis, *Res. vet. Sci.*, 22, 130-131.

Tillard E., 1991, Evaluation technico-économique des effets de prophylaxies en milieu villageois chez les petits ruminants de la région de Kolda, Thèse de Doctorat vétérinaire, Faculté de médecine de Créteil, 172 p.

Tourrand J.F., 1994, L'élevage dans la révolution agricole du Waalo. Ruptures et continuité, thèse de Doctorat ès-Sciences, Université Paris XII - Val de Marne - Créteil, 416 p.

# Annexes

## 1 — Liste des abréviations

FS:	fin de saison sèche
SP:	saison des pluies
SF:	saison sèche fra"che
SC:	saison sèche chaude
AD:	oviadénovirus type 5
PI:	parainfluenza type III
RS:	virus respiratoire syncitial
H1:	pasteurella haemolytica type 1
H2:	pasteurella haemolytica type 2
H5:	pasteurella haemolytica type 5
H6:	pasteurella haemolytica type 6
H7:	pasteurella haemolytica type 7
H9:	pasteurella haemolytica type 9
H11:	pasteurella haemolytica type 11
MA:	pasteurella multocida type A
MD:	pasteurella multocida type D
DI:	Ndiagne
KM:	Keur Mousso Kebe
ND:	Ndeme
TH:	Thiar Peul
DW:	Diawar
LP:	Lampsar Peul
LS:	Lampsar Wolof
SE:	Senda
TG:	Thiago
AFC:	analyse factorielle des corespondances
ACP:	analyse en composantes principales
ACM:	analyse des correspondances multiples
AFCD:	analyse factorielle des corespondances décentrée
AFCVI:	analyse factorielle des correspondances sur variables instrumentales
ACOM:	analyse de co-inertie multiple
STATIS:	structuration des tableaux à trois indices de la statistique

## 2 — Liste des modules et options utilisés

Analyses	Figure	Module	Option
AFCD	2	COA	Decentred COA
Contributions colonnes	2	Ddutil	Columns/Inertia analysis
Histogramme des valeurs propres	2A	Curves	Eigenvalues
Plan factoriel colonne	2B - 2C	Scatters	Labels
Plan factoriel lignes	2D - 2E	Scatterclass	Stars
ANALYSE INTERPÉRIODE	3	DISCRIMIN	Initialize/link prep
Test de permutation	annexe 7	DISCRIMIN	Between analysis/run
Contributions colonnes	3	Ddutil	Columns/Inertia analysis
Histogramme des valeurs propres	3A	Curves	Eigenvalues
Plan factoriel colonne	3A - 3D	Scatters	Labels
	3F		
Plan factoriel lignes	3B - 3C	Scatterclass	Stars
	3E - 3G		
PROFILS DE COORDONNÉES	4	CurveModels	Numerical
Analyse intrapériode	5	DISCRIMIN	Within analysis/run
Contributions colonnes	5	Ddutil	Columns/Inertia analysis
Histogramme des valeurs propres	5A	Curves	Eigenvalues
Plan factoriel colonne	5A	Scatters	Labels
Plan factoriel lignes	5B	Scatters	Labels
Plan factoriel lignes	5C	Scatters	Trajectories
ANALYSE SUR VAR. INSTRUM.	annexe 7	PROJECTORS	2 categ var orthonormal bases
	6		PCA on instrumental variables
Contributions colonnes	6	Ddutil	Columns/Inertia analysis
Histogramme des valeurs propres	6A	Curves	Eigenvalues
Plan factoriel colonne	6A	Scatters	Labels
Plan factoriel lignes	6B	Scatters	Trajectories
Plan factoriel lignes	6C	Scatters	Trajectories et superposition
RÉSIDUS	7	Ddutil	Residuals
Analyse séparées	8	KTA - MFA	Separate analysis
Histogramme des valeurs propres	8A	Curves	Eigenvalues
Plan factoriel colonne	8B	Scatters	Labels
Plan factoriel lignes	8C	Scatters	Labels
STATIS	9	STATIS	Operator averaging
Histogramme des valeurs propres	9A - 9C	Curves	Eigenvalues
Coordonnées des tableaux	9B	Scatters	Labels
Plan factoriel colonne	9C	Scatters	Labels
Plan factoriel lignes	9D	Scatters	Labels et superposition
Plan factoriel lignes	9E	Scatters	Match 2 scatters
ACOM	10	KTA - MFA	Multiple Coinertia analysis
Histogramme des valeurs propres	10A	Curves	Eigenvalues
Plan factoriel colonne	10A	Scatters	Labels
Plan factoriel colonne	10B	Scatters	Match 2 scatters
Plan factoriel lignes	10C	Scatters	Labels
Axes d'inertie et plan de coinertie	10D	Scatters	Labels et superposition
Composantes synthétiques	10E	Scatters	





## 5 — Tableau des numérateurs standardisés et tableau des taux standardisés

Tableau des numérateurs standardisés (X<sub>ij</sub>)

Village	V	P	H1	H2	H5	H6	H7	H9	H11	MA	MD	AD	PI	RS	Effectif
DI	1	1	6,0	15,1	17,1	14,0	16,0	4,0	24,0	27,1	1,0	4,0	0,0	1,0	35
KM	2	1	1,9	2,0	2,1	3,9	3,9	2,0	7,1	5,0	0,0	0,0	8,0	1,9	11
ND	3	1	3,5	9,4	8,4	8,0	6,0	1,1	10,2	7,8	2,0	1,8	1,8	0,9	18
DW	5	1	0,0	8,3	3,6	5,6	0,0	0,0	15,5	0,9	0,0	0,0	0,0	0,0	19
LP	6	1	3,2	0,8	0,0	1,5	1,6	0,0	3,1	0,0	0,0	0,0	0,0	3,2	11
LS	7	1	1,0	3,9	1,0	4,9	1,9	2,0	15,0	10,8	1,0	1,9	0,0	0,0	30
DI	1	2	3,8	2,8	2,8	2,8	4,7	1,0	2,8	7,6	0,0	6,6	12,3	1,0	18
KM	2	2	2,1	2,1	1,5	2,1	2,1	0,0	4,7	5,3	1,1	1,5	0,0	2,1	9
ND	3	2	1,1	1,1	0,5	3,2	3,2	1,1	2,6	7,5	0,0	3,2	1,1	2,1	9
DW	5	2	2,6	5,5	5,5	5,5	3,9	2,6	13,6	10,5	1,5	1,3	0,0	0,2	23
LP	6	2	1,3	1,7	0,0	1,5	0,0	0,2	7,8	9,1	1,3	0,0	0,0	4,0	12
LS	7	2	1,8	8,7	8,3	13,6	3,3	3,3	18,8	28,4	3,3	6,7	0,0	8,3	50
TG	9	2	0,1	2,5	2,4	0,0	0,0	0,0	5,2	4,9	0,0	0,1	2,4	4,8	10
DI	1	3	5,0	9,0	5,9	7,3	5,9	3,2	5,0	7,6	0,9	1,8	3,5	1,8	14
ND	3	3	0,0	3,3	0,0	1,3	1,3	3,3	4,0	4,0	0,0	4,4	6,3	5,1	9
TH	4	3	1,1	4,2	4,4	0,0	2,2	1,9	2,7	3,3	0,0	3,9	4,2	2,2	8
DW	5	3	3,9	8,4	3,7	10,1	8,2	5,1	11,3	8,7	0,0	12,5	0,0	11,2	28
LP	6	3	0,0	1,9	0,0	1,0	1,9	1,0	3,8	2,1	0,0	0,0	2,9	4,9	11
LS	7	3	6,9	13,8	5,5	4,4	6,6	4,4	15,8	20,1	2,8	6,9	7,7	11,8	44
SE	8	3	0,0	1,2	0,4	0,0	2,8	0,0	6,4	1,2	0,0	1,2	7,6	6,4	8
TG	9	3	0,9	7,4	1,9	0,9	2,8	3,1	0,9	3,7	0,0	0,9	1,2	3,7	12
DI	1	4	4,0	7,0	8,0	7,0	11,0	3,0	10,0	8,0	2,0	6,0	11,0	2,0	17
DW	5	4	7,0	19,0	7,0	10,0	5,0	4,0	17,0	19,0	3,0	8,0	8,0	16,0	34
LP	7	4	5,0	8,0	5,0	12,0	9,0	4,0	8,0	14,0	2,0	3,0	8,0	7,0	23
SE	8	4	1,0	4,0	1,0	2,0	2,0	0,0	6,0	10,0	1,0	5,0	3,0	9,0	14
IG	9	4	1,0	4,0	4,0	5,0	2,0	1,0	4,0	6,0	0,0	0,0	0,0	2,0	10

Tableau des taux standardisés (t<sub>ij</sub>) (=numérateur / effectif, en pourcentage)

Village	V	P	H1	H2	H5	H7	H8	H9	H11	MA	MD	AD	PI	RS
DI	1	1	17	43	49	40	46	11	69	77	3	11	0	3
KM	2	1	18	18	19	36	36	18	64	46	0	0	73	18
ND	3	1	20	52	47	44	33	6	57	43	11	10	10	5
DW	5	1	0	44	19	30	0	0	81	5	0	0	0	0
LP	6	1	29	7	0	14	15	0	28	0	0	0	0	29
LS	7	1	3	13	3	16	6	7	50	36	3	6	0	0
DI	1	2	21	16	16	16	26	5	16	42	0	37	68	5
KM	2	2	24	24	17	24	24	0	53	59	12	17	0	24
ND	3	2	12	12	5	36	36	12	29	83	0	36	12	24
DW	5	2	11	24	24	24	17	11	59	45	6	6	0	1
LP	6	2	11	14	0	12	0	2	65	75	11	0	0	33
LS	7	2	4	17	17	27	7	7	38	57	7	13	0	17
TG	9	2	1	25	24	0	0	0	52	49	0	1	24	48
DI	1	3	35	65	42	52	42	23	35	54	6	13	25	13
ND	3	3	0	37	0	14	14	37	44	44	0	49	69	57
TH	4	3	14	52	56	0	28	24	34	42	0	48	52	28
DW	5	3	14	30	13	36	29	18	40	31	0	45	0	40
LP	6	3	0	17	0	9	17	9	35	19	0	0	26	45
LS	7	3	16	31	12	10	15	10	36	46	6	16	18	27
SE	8	3	0	15	5	0	35	0	80	15	0	15	95	80
TG	9	3	8	62	15	8	23	26	8	31	0	8	10	31
DI	1	4	24	41	47	41	65	18	59	47	12	35	65	12
DW	5	4	21	56	21	29	15	12	50	56	9	24	24	47
LP	7	4	22	35	22	52	39	17	35	61	9	13	35	30
SE	8	4	7	29	7	14	14	0	43	71	7	36	21	64
IG	9	4	10	40	40	50	20	10	40	60	0	0	0	20

Exemple à DI (Ndiagne) en fin de saison sèche (FS), soit entre P4 et P6 (cf annexe pour la sérologie RS).

$r = 35$  animaux, prélevés à la fois à P4 et P6

$n(s_+) = 0$  et  $n(s_c) = 1$

$$t_c = \frac{x}{r} = \frac{0+1}{35} = 0,02857, \text{ arrondi à } 0,03 \%$$

$$t_1 = \frac{x_1}{r_1} = \frac{0}{20} \text{ et } t_2 = \frac{x_2}{r_2} = \frac{1}{15}, \text{ soit } t_c = \frac{(t_1 \times r_1) + (t_2 \times r_2)}{r_1 + r_2} = \frac{(0 \times 20) + (1/15 \times 30)}{20 + 30} = 0,02941,$$

où 20 et 30 sont les effectifs des animaux de moins de 3 mois et de plus de 3 mois, respectivement, prélevés dans la région de Louga (4 villages) en FS.

## 6 — Analyse Factorielle des Correspondances Décentrées (AFCD)

Col: 1	H1	Khi2 = 28.34	DDL = 25	Proba = 0.292	Trace Cont. = 1.67
Col: 2	H2	Khi2 = 36.05	DDL = 25	Proba = 0.0706	Trace Cont. = 1.68
Col: 3	H5	Khi2 = 54.58	DDL = 25	Proba = 0.000583	Trace Cont. = 1.73
Col: 4	H6	Khi2 = 38.67	DDL = 25	Proba = 0.0397	Trace Cont. = 1.72
Col: 5	H7	Khi2 = 53.72	DDL = 25	Proba = 0.000751	Trace Cont. = 1.73
Col: 6	H9	Khi2 = 26.31	DDL = 25	Proba = 0.391	Trace Cont. = 1.67
Col: 7	H11	Khi2 = 27.41	DDL = 25	Proba = 0.336	Trace Cont. = 1.67
Col: 8	MA	Khi2 = 35.17	DDL = 25	Proba = 0.00.897	Trace Cont. = 1.68
Col: 9	MD	Khi2 = 16.58	DDL = 25	Proba = 0.897	Trace Cont. = 1.62
Col:10	AD	Khi2 = 57.54	DDL = 25	Proba = 0.00024	Trace Cont. = 1.74
Col:11	PI	Khi2 154.40	DDL = 25	Proba = 0	Trace Cont. = 1.84
Col:12	RS	Khi2 74.59	DDL = 25	Proba = 9.83E-07	Trace Cont. = 1.75

Total inertia: 1.23893

File afcd.frvp contains the eigenvalues and relative inertia for each axis

Num	Eigenval.	R.Iner.	R.Sum	Num	Eigenval	R.Iner.	R.Sum
01	+4.1508E-01	+0.3350	+0.3350	02	+3.140E-01	+0.2535	+0.5885
03	+1.4014E-01	+0.1131	+0.7016	04	+8.653E-02	+0.0698	+0.7714
05	+7.2632E-02	+0.0586	+0.8301	06	+5.531E-02	+0.0446	+0.8747
07	+4.6383E-02	+0.0374	+0.9121	08	+3.875E-02	+0.0313	+0.9434
09	+3.0313E-02	+0.0245	+0.9679	10	+1.926E-02	+0.0155	+0.9834
11	+1.3099E-02	+0.0106	+0.9940	12	+7.426E-03	+0.0060	+1.0000

All contributions are in 1/10000

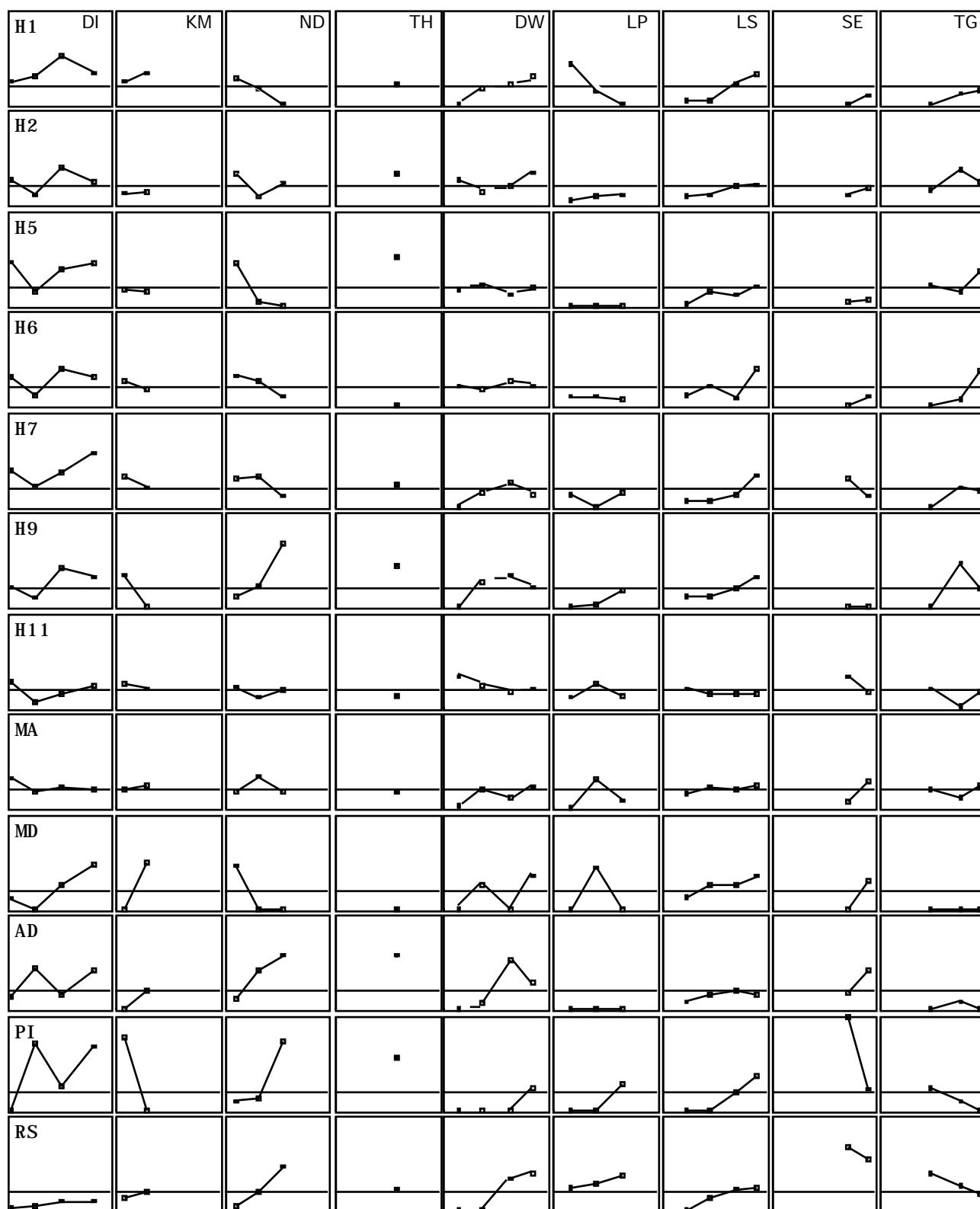
-----Absolute contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Fac 4	Fac 5
1	H1	212	557	87	16	39
2	H2	125	719	517	1182	2760
3	H5	116	2746	23	347	181
4	H6	9	1587	99	1	91
5	H7	998	1531	0	88	77
6	H9	340	127	350	397	925
7	H11	31	126	453	1909	17
8	MA	3	379	730	139	4897
9	MD	0	107	75	372	906
10	AD	1028	17	2377	2551	0
11	PI	6506	379	2251	60	100
12	RS	627	1719	3033	2932	2

-----Relative contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Fac 4	Fac 5	Remains	Weight	Cont
1	H1	1512	3007	209	23	49	5197	1318	469
2	H2	702	3050	980	1382	2708	1175	3184	597
3	H5	432	7694	29	268	11	1457	2053	904
4	H6	47	6279	175	1	83	3413	2620	640
5	H7	3757	4359	0	69	51	1761	2203	890
6	H9	2618	741	909	636	1243	3851	1053	436
7	H11	230	704	1130	293	22	4976	4626	454
8	MA	20	1649	1417	167	4925	1820	4776	582
9	MD	1	993	308	946	1933	5815	470	274
10	AD	3613	47	2819	1868	0	1650	1657	953
11	PI	8518	375	995	16	23	70	1827	2558
12	RS	1700	3526	2775	1656	1	339	2312	1236





Annexe 6. Suite. Représentation graphique du tableau de données traité.

Les 12 variables (agents pathogènes) en ligne sont mesurées à quatre périodes, dans 9 villages. Les valeurs représentent les écarts au taux moyen par agent pathogène, toutes périodes confondues (centrage par colonne), rapportés au taux moyen par agent pathogène. Le calcul du taux constitue un centrage des données du tableau par rapport à la marge imposée.



## 8 — Analyse interperiode

Discrimin: Between analysis: Run 26/06/96 15/0  
 total inertia: 1.238928  
 File afcd-disc.bevp contains the eigenvalues and relative inertia for each axis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+1.8150E-01	+0.6507	+0.6507	02	+6.2940E-02	+0.2257	+0.8764
03	+3.4470E-02	+0.1236	+1.0000	04	+0.0000E+00	+0.0000	+1.0000

DDUtil: Columns: inertia analysis 26/06/96 15/03  
 All contributions are in 1/10000  
 -----Absolute contributions-----

Num	Var	Fac 1	Fac 2
1	H1	169	547
2	H2	651	1583
3	H5	18	1722
4	H6	0	2529
5	H7	259	1501
6	H9	700	11
7	H11	442	1338
8	MA	0	19
9	MD	44	154
10	AD	1486	345
11	PI	2375	103
12	RS	3850	144

-----Relative contributions-----

Num	Var	Fac 1	Fac 2	Remains	Weight	Cont.
1	H1	4691	5259	48	1318	234
2	H2	4830	4071	1098	3184	877
3	H5	306	9693	0	2053	400
4	H6	2	8726	1271	2620	654
5	H7	3026	6075	897	2203	557
6	H9	7597	43	2359	1053	600
7	H11	4803	5033	162	4626	600
8	MA	11	75	9912	4776	574
9	MD	904	1099	7995	470	318
10	AD	9191	740	67	1657	1052
11	PI	9777	147	75	1827	1580
12	RS	9829	127	42	2312	2548

## 9 — Analyse intravillage

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT & DC

Discrimin: Within Analysis: Run 04/07/96 16/49

total inertia: 1.238928

File afcd-disc-vil.whvp contains the eigenvalues and relative inertia for each axis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+3.0260E-01	+0.3938	+0.3938	02	+1.4994E-01	+0.1951	+0.5889
03	+1.1096E-01	+0.1444	+0.7333	04	+5.7040E-02	+0.0742	+0.8075
05	+4.2960E-02	+0.0559	+0.8634	06	+3.5853E-02	+0.0467	+0.9101
07	+2.3061E-02	+0.0300	+0.9401	08	+1.7661E-02	+0.0230	+0.9631
09	+1.4495E-02	+0.0189	+0.9820	10	+6.9140E-03	+0.0090	+0.9910
11	+4.4527E-03	+0.0058	+0.9968	12	+2.4965E-03	+0.0032	+1.0000

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT & DC

DDUtil: Columns: inertia analysis 04/07/96 16/46

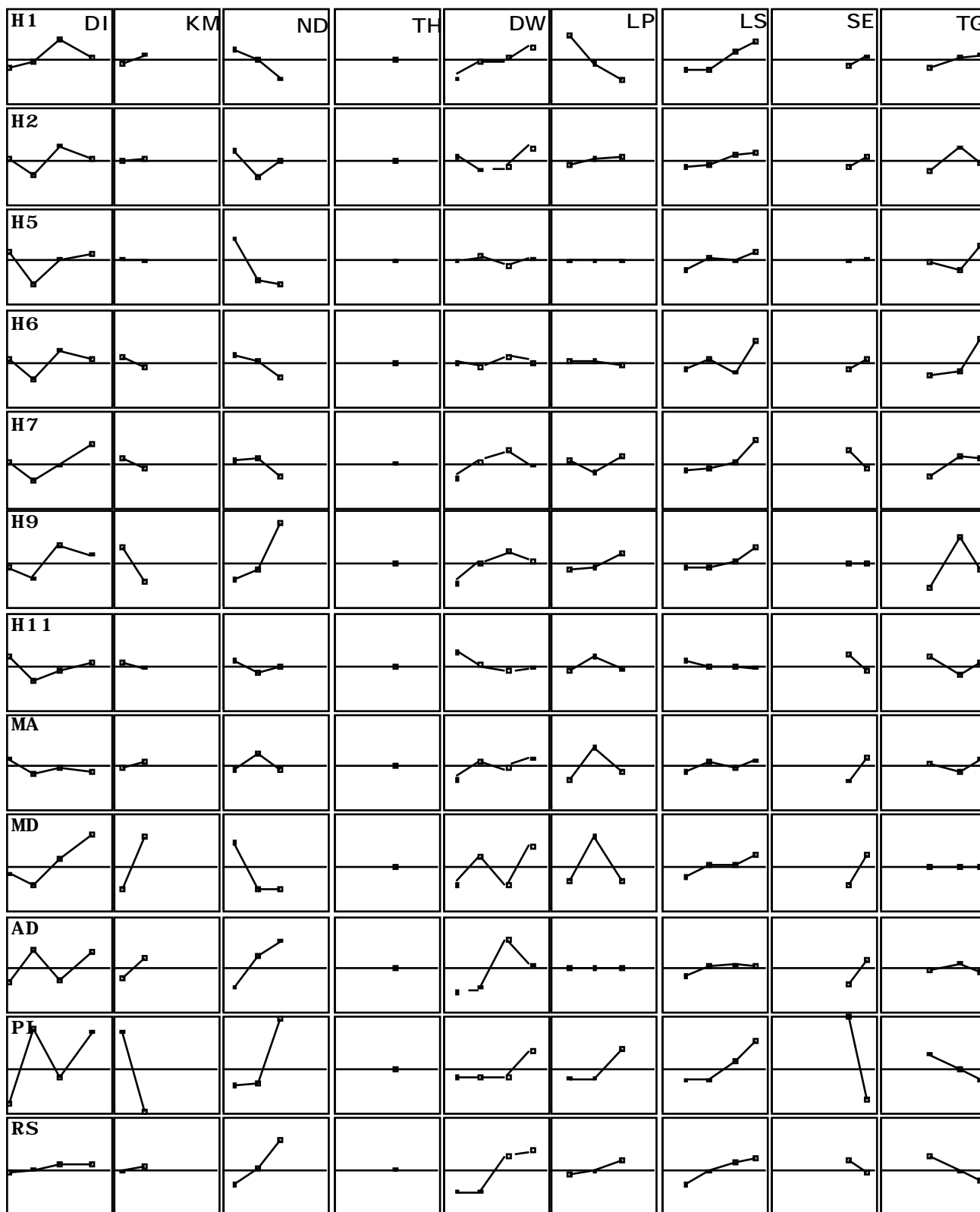
All contributions are in 1/10000

-----Absolute contributions-----

Num	Var	Fac 1	Fac 2	Fac 3
1	H1	49	1152	0
2	H2	0	1219	292
3	H5	311	1013	1029
4	H6	90	1748	281
5	H7	112	1502	219
6	H9	383	466	54
7	H11	298	30	1245
8	MA	122	1056	624
9	MD	15	656	75
10	AD	848	78	3183
11	PI	6445	183	2184
12	RS	1320	891	808

-----Relative contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Remains	Weight	Cont.
1	H1	388	4503	2	5105	1318	499
2	H2	0	4260	757	4981	3184	558
3	H5	2251	3623	2722	1403	2053	545
4	H6	534	5128	610	3726	2620	665
5	H7	691	4565	493	4249	2203	642
6	H9	2934	1768	153	5142	1053	514
7	H11	2117	106	3239	4536	4626	555
8	MA	643	2738	1197	5420	4776	752
9	MD	162	3463	293	6081	470	369
10	AD	3616	166	4972	1244	1657	924
11	PI	8626	121	1072	179	1827	2942
12	RS	5051	1689	1133	2125	2312	1029



Annexe 9. Suite. Représentation graphique du tableau de données centrées par village

Les 12 variables (agents pathogènes) en ligne sont mesurées à quatre périodes, dans 9 villages. Les valeurs représentent les écarts au taux moyen par village, rapportés au taux moyen tous villages confondus.

**10 — Analyse intrapériode**

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT &amp; DC

Discrimin: Within Analysis: Run 27/06/96 11/53

total inertia: 1.238928

File afcd-disc.whvp contains the eigenvalues and relative inertia for each axis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+3.1743E-01	+0.3307	+0.3307	02	+2.6600E-01	+0.2771	+0.6077
03	+8.7257E-02	+0.0909	+0.6986	04	+6.9911E-02	+0.0728	+0.7714
05	+5.3241E-02	+0.0555	+0.8269	06	+4.0337E-02	+0.0420	+0.8689
07	+3.7360E-02	+0.0389	+0.9078	08	+3.3158E-02	+0.0345	+0.9424
09	+2.5921E-02	+0.0270	+0.9694	10	+1.2834E-02	+0.0134	+0.9827
11	+9.9894E-03	+0.0104	+0.9931	12	+6.5786E-03	+0.0069	+1.0000

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT &amp; DC

DDUtil: Columns: inertia analysis 27/06/96 11/54

All contributions are in 1/10000

-----Absolute contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Fac 4
1	H1	263	468	0	323
2	H2	22	750	243	719
3	H5	611	2208	383	967
4	H6	73	1301	41	130
5	H7	160	929	123	0
6	H9	218	235	130	179
7	H11	2	0	229	2503
8	MA	19	532	506	1443
9	MD	4	127	112	79
10	AD	527	1	6893	24
11	PI	6598	2147	560	0
12	RS	48	1296	774	3628

-----Relative contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Fac 4	Remains	Weight	Cont.
1	H1	1622	2411	0	437	5528	1318	537
2	H2	145	4028	428	1015	4382	3184	516
3	H5	1922	5823	331	670	1251	2053	1050
4	H6	380	5660	59	148	3750	2620	637
5	H7	5389	2609	113	0	1886	2203	986
6	H9	1864	1681	306	335	5812	1053	388
7	H11	21	2	505	4426	5043	4626	411
8	MA	111	2519	786	1795	4786	4776	585
9	MD	59	1343	390	220	7985	470	262
10	AD	1884	5	6773	19	1316	1657	925
11	PI	7673	2092	179	0	54	1827	2843
12	RS	187	4201	823	3091	1696	2312	854

**11 — Analyse sur variables instrumentales: projection sur P+V**

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT & DC  
 Projectors: PCA on Instrumental Variables28/06/96 11/12  
 Inertia: 7.4427e-01

File afcd-vi.ivvp contains the eigenvalues and relative inertia for each axis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+3.4615E-01	+0.4651	+0.4651	02	+1.9565E-01	+0.2629	+0.7280
03	+6.6838E-02	+0.0898	+0.8178	04	+3.9415E-02	+0.0530	+0.8707
05	+3.7348E-02	+0.0502	+0.9209	06	+2.9076E-02	+0.0391	+0.9600
07	+2.0273E-02	+0.0272	+0.9872	08	+4.4600E-03	+0.0060	+0.9932
09	+4.0132E-03	+0.0054	+0.9986	10	+9.8240E-04	+0.0013	+0.9999
11	+6.3685E-05	+0.0001	+1.0000				

ADE-4 \* Metrowerks CodeWarrior C \* CNRS-Lyon1 \* JT & DC  
 DDUtil: Columns: inertia analysis 28/06/96 11/24  
 All contributions are in 1/10000

-----Absolute contributions-----

Num	Var	Fac 1	Fac 2	Fac 3
1	H1	579	126	38
2	H2	662	114	3316
3	H5	973	1681	219
4	H6	318	830	7
5	H7	1877	342	53
6	H9	488	0	1286
7	H11	10	253	3
8	MA	231	52	2259
9	MD	34	2	790
10	AD	1282	337	280
11	PI	3415	867	1351
12	RS	125	5391	392

-----Relative contributions-----

Num	Var	Fac 1	Fac 2	Fac 3	Remains	Weight	Cont.
1	H1	6178	759	79	2982	1318	436
2	H2	4530	442	4378	649	3184	680
3	H5	4329	4226	188	1255	2053	1046
4	H6	2513	3703	11	3771	2620	589
5	H7	8221	846	45	886	2203	1062
6	H9	4825	1	2454	2719	1053	470
7	H11	127	1773	9	8088	4626	376
8	MA	2450	314	4626	2607	4776	438
9	MD	722	24	3186	6066	470	222
10	AD	5713	849	241	3196	1657	1044
11	PI	7879	1130	601	387	1827	2015
12	RS	360	8762	218	659	2312	1617

## 12 – Résultats des analyses variable par variable des effets «période» et «village» par regression logistique

Variable	H1		H2		H5		H6	
	Déviante	p-value	Déviante	p-value	Déviante	p-value	Déviante	p-value
(C)	373,737		609,838		481,277		553,844	
(C)+P	370,200	0,3123	590,674	0,0003	472,166	0,0278	539,190	0,0021
(C)+V	362,428	0,1833	588,440	0,0062	432,042	0,0000	533,697	0,0098
(C+P)+V	355,827	0,0725	573,616	0,0295	425,592	0,0000	521,563	0,0242
(C+P+V)+I	337,734	0,2026	558,071	0,3420	404,141	0,0906	495,657	0,0266
<b>Modèle retenu</b>	<b>Logit = C</b>		<b>Logit = C+P+V</b>		<b>Logit=C+V+P</b>		<b>Logit=C+P+V+I</b>	

Variable	H7		H9		H11		MA	
	Déviante	p-value	Déviante	p-value	Déviante	p-value	Déviante	p-value
(C)	494,871		321,216		666,068		665,503	
(C)+P	481,264	0,0035	310,466	0,0132	651,740	0,0025	655,991	0,0232
(C)+V	455,759	0,0000	311,224	0,2656	653,383	0,1232	646,773	0,0164
(C+P)+V	438,370	0,0000	297,524	0,1138	641,311	0,2362	635,552	0,0088
(C+P+V)+I	413,816	0,0392	285,197	0,5800	616,872	0,0405	596,353	0,0003
<b>Modèle retenu</b>	<b>Logit=C+P+V+I</b>		<b>Logit=C+P</b>		<b>Logit = C+P</b>		<b>Logit = C+V+P+I</b>	

Variable	MD		AD		PI		RS	
	Déviante	p-value	Déviante	p-value	Déviante	p-value	Déviante	p-value
(C)			433,266		449,042		500,132	
(C)+P	172,675	0,1801	413,837	0,0002	421,269	0,0000	451,758	0,0000
(C)+V	167,786	0,8589	399,975	0,0001	398,578	0,0000	464,445	0,0000
(C+P)+V	168,695	0,8062	379,874	0,0000	352,273	0,0000	421,742	0,0002
(C+P+V)+I	163,253	0,5458	361,612	0,1951	266,493	0,0000	405,194	0,2811
<b>Modèle retenu</b>	<b>Logit = C</b>		<b>Logit = C+V+P</b>		<b>Logit = C+P+V+I</b>		<b>Logit = C+P+V</b>	



**13 — STATIS sur RV**

----- Correlation matrix -----  
 [ 1] 1000  
 [ 2] 327 1000  
 [ 3] 541 484 1000  
 [ 4] 412 515 398 1000  
 -----

File afcd-ktab-rv.oa+EV contains the eigenvalues of the interstructure analysis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+2.3404E+00	+0.5851	+0.5851	02	+7.2784E-01	+0.1820	+0.7670
03	+5.5216E-01	+0.1380	+0.9051	04	+3.7964E-01	+0.0949	+1.0000

File afcd-ktab-rv.oa+vp contains the eigenvalues of the compromise analysis

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+6.4401E-01	+0.3377	+0.3377	02	+5.1761E-01	+0.2714	+0.6091
03	+1.7296E-01	+0.0907	+0.6998	04	+1.3726E-01	+0.0720	+0.7718
05	+1.0484E-01	+0.0550	+0.8268	06	+8.0836E-02	+0.0424	+0.8692
07	+7.3039E-02	+0.0383	+0.9075	08	+6.6603E-02	+0.0349	+0.9424
09	+5.2121E-02	+0.0273	+0.9697	10	+2.5712E-02	+0.0135	+0.9832
11	+1.9470E-02	+0.0102	+0.9934	12	+1.2571E-02	+0.0066	+1.0000

Typological value indices

Rows = row number of each table

Weights = Weights of operators in the compromise

NS norm2 = Square norm (Hilbert-Schmidt) of operators

Cos2 = Square cosine between operator and approximated compromise

Num	Rows	Weights	NS norm2	Cos2
1	6	0.4547	0.4151	0.549
2	7	0.4972	0.2198	0.463
3	8	0.5223	0.3741	0.564
4	9	0.4951	0.3062	0.514

## 14 — Analyse de Co-inertie Multiple (ACOM)

-> Total inertia from separate analysis

vp level 1=3.3009e-01

vp level 1=2.0517e-01

vp level 1=6.3225e-02

vp level 1=3.5964e-02

Table n°: 1 \* Rows: 6 \* Col: 12 \* Rank: 5  
\* Weight (x1000): 213 \* RV (x1000): 767

N	Iner	Iner+	Var	Var+	Sca. Pro.2	Cos2
1	5.636e-01	5.636e-01	4.321e-01	4.321e-01	3.014e-01	6.976e-01
2	2.716e-01	8.352e-01	4.002e-01	8.323e-01	3.068e-01	7.668e-01
3	1.233e-01	9.586e-01	1.179e-01	9.502e-01	5.663e-02	4.805e-01
4	8.705e-02	1.046e+00	9.428e-02	1.044e+00	5.190e-02	5.505e-01

Table n°: 2 \* Rows: 7 \* Col: 12 \* Rank: 6  
\* Weight (x1000): 313 \* RV (x1000): 796

N	Iner	Iner+	Var	Var+	Sca. Pro.2	Cos2
1	4.411e-01	4.411e-01	4.325e-01	4.325e-01	3.287e-01	7.599e-01
2	1.261e-01	5.672e-01	1.254e-01	5.579e-01	9.323e-02	7.432e-01
3	7.854e-02	6.458e-01	8.177e-02	6.397e-01	7.311e-02	8.941e-01
4	4.895e-02	6.947e-01	4.935e-02	6.891e-01	3.027e-02	6.133e-01

Table n°: 3 \* Rows: 8 \* Col: 12 \* Rank: 7  
\* Weight (x1000): 194 \* RV (x1000): 838

N	Iner	Iner+	Var	Var+	Sca. Pro.2	Cos2
1	4.999e-01	4.999e-01	3.133e-01	3.133e-01	2.369e-01	7.562e-01
2	2.584e-01	7.583e-01	4.391e-01	7.524e-01	4.099e-01	9.335e-01
3	2.021e-01	9.604e-01	2.015e-01	9.539e-01	1.242e-01	6.163e-01
4	1.025e-01	1.063e+00	1.016e-01	1.056e+00	5.979e-02	5.882e-01

Table n°: 4 \* Rows: 5 \* Col: 12 \* Rank: 4  
\* Weight (x1000): 278 \* RV (x1000): 669

N	Iner	Iner+	Var	Var+	Sca. Pro.2	Cos2
1	5.094e-01	5.094e-01	4.918e-01	4.918e-01	4.188e-01	8.517e-01
2	2.005e-01	7.099e-01	2.088e-01	7.006e-01	1.100e-01	5.268e-01
3	6.361e-02	7.735e-01	7.121e-02	7.718e-01	1.448e-02	2.034e-01
4	4.964e-02	8.232e-01	5.134e-02	8.232e-01	1.350e-02	2.629e-01

Iner: inertie projetée sur l'axe N d'inertie  
 Iner+: inertie précédente cumulée  
 Var: inertie projetée sur l'axe N de co-inertie  
 Var+: inertie précédente cumulé  
 Sca. Pro 2: carré de covariance entre les coordonnées et les variables auxiliaires de même rang  
 Cos2: carré de corrélation entre les coordonnées et les variables auxiliaires de même rang