Points chauds de l'analyse dialectométrique : pondération et visualisation

Autor(en): **Goebl, Hans**

Objekttyp: **Article**

Zeitschrift: **Revue de linguistique romane**

Band (Jahr): **51 (1987)**

Heft 201-202

PDF erstellt am: **26.04.2024**

Persistenter Link: https://doi.org/10.5169/seals-399808

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

POINTS CHAUDS DE L'ANALYSE DIALECTOMÉTRIQUE: *PONDÉRATION* ET *VISUALISATION* (1)

0. Remarques préliminaires.

La dialectométrie est un procédé de classification numérique qui comprend les étapes suivantes (2) :

a) choix des données empiriques à classifier
b) choix du principe métrologique permettant de mesurer les données empiriques à classifier
c) choix de l'indice de similarité
d) choix des procédés statistiques pour exploiter la matrice de similarité
e) choix de la méthode de visualisation des résultats de l'analyse dialectométrique.

Chacune de ces étapes présuppose une solide charpente *théorique* qui, elle, fournit la justification tant de l'assiette méthodologique de l'analyse taxométrique entreprise que des multiples choix méthodiques à faire. Tout dépend donc des *visées classificatoires* que le dialectométricien attribue à la classification entreprise.

Le but de cet article est de fournir des précisions méthodiques sur les étapes c) (choix de l'indice de similarité) et e) (choix de la méthode de visualisation) de la chaîne dialectométrique mentionnée ci-dessus.

---

(1) Nos travaux dialectométriques n'ont pu être réalisés que grâce à l'appui financier des organismes et institutions scientifiques suivants :
Hochschuljubiläumsstiftung der Stadt Wien (Vienne)
Dr. Adolf Schärf-Fonds zur Förderung der Wissenschaften (Vienne)
Fonds zur Förderung der wissenschaftlichen Forschung (Vienne)
Deutsche Forschungsgemeinschaft (Bonn, RFA).
Toilette stylistique : François Genton (Ratisbonne). J'adresse ici aux MM. Genton, Selberherr, Rase et Pudlatz ainsi qu'aux organismes cités ci-dessus ma plus vive reconnaissance.

(2) Pour plus de détails nous renvoyons à nos contributions dialectométriques antérieures, en l'occurrence ceux de 1981, 1982, 1983 (a-d) et 1984. Le lecteur y trouvera tant une définition exacte de la dialectométrie à l'heure actuelle qu'un historique précis de la pensée dialectométrique, ainsi que de copieuses références bibliographiques.

Vu leur rôle décisif dans la chaîne taxométrique, les deux étapes représentent de véritables « points chauds ». Il est donc nécessaire que les dialectologues désireux de se servir de l'analyse dialectométrique en prennent connaissance, d'autant plus qu'il s'agit ici de deux champs méthodiques voire méthodologiques pour lesquels la dialectologie traditionnelle dispose déjà de certains concepts pré-théoriques qui, malgré leur apparent manque de précision, pourraient néanmoins servir de point de jonction entre la dialectologie traditionnelle et la dialectométrie.

1. La taxation des données originales : de l'ALF à la matrice des données.

Nous retracerons, par la suite, la génération du corpus dialectométrique utilisé dans ses grandes lignes. Les données originales à classifier ont été tirées de l'ALF (3). En voici les coordonnées essentielles :

données dialectales brutes (taxandum) : ALF (cartes 1-1421)
objets (points d'atlas) : 70 points-ALF + 1 point artificiel (P. 999)
P. 999 : équivalent du français standard tel qu'il figure dans les en-têtes des planches de l'ALF (4)
critère discriminatoire pour l'isolement des unités taxatoires (taxats-types) : différences *lexicales* et *morphosyntaxiques*
attributs (cartes analysées ou « de travail » obtenues par la taxation des cartes ALF) : 1468 (5)

---

(3) Pour les détails de l'analyse métrologique d'un corpus géolinguistique donné cf. Goebl 1984 I, 16-73.

(4) En ce qui concerne l'utilité taxométrique de l'insertion d'un point artificiel dans un réseau d'atlas examiné cf. Goebl 1981, 351 et Goebl 1984 I, 86-87. Vu les innombrables discussions soulevées autour du problème de la « francisation » des dialectes galloromans, il semble superflu de s'attarder plus longtemps sur l'utilité de l'insertion du point 999.

(5) Il faut distinguer les cartes originales de l'ALF (plan des *données brutes*, plan du taxandum) des cartes analysées ou « de travail » (plan de la *matrice des données*). Il y aura donc opposition entre *carte d'atlas* et *carte analysée* (ou « *de travail* »). Exemple illustratif : la carte ALF 437 *échine* montre, à l'intérieur du réseau normanno-picard examiné, les 6 concepts lexicologiques suivants : ⌜*échine*⌝, ⌜*dos*⌝, ⌜*reins*⌝, ⌜*épine du dos*⌝, ⌜*épine*⌝ et ⌜*arête du dos*⌝. En appliquant le critère discriminatoire de la différenciation lexicale, l'on obtient donc, de cette carte-ALF originale, une seule « carte de travail ». La carte ALF 558 *sur le feu* en revanche peut être scindée en trois « cartes de travail » :

a) une carte *lexicale mono*nyme avec le type (taxat) ⌜*feu*⌝
b) une carte *morphosyntaxique tétra*nyme avec les types (taxats) : ⌜*sur le*⌝, ⌜*dans le*⌝, ⌜*au*⌝, ⌜*le*⌝

nombre des unités taxatoires (taxats-types) consignées dans la matrice des données : 3959

nombre de taxats/carte « de travail » 2,6969.

Pour une description des *données originales* (ALF) et de la *matrice des données* en tirée par voie taxatoire (codage) cf. Goebl 1984 I, 55-70.

La matrice des données représente, en dernière analyse, une sorte de *photographie* — prise sous un angle bien défini — des données originales. Or, nécessairement, toute *image* n'est rien d'autre qu'une *réduction* voire une *section* de la réalité originale qu'elle reproduit. C'est là un aspect primordial de toute analyse classificatoire. Qui dit *réduction* dit aussi *perte d'information* sur le plan des *données particulières,* perte compensée par un gain d'information classificatoire au niveau de l'*introspection générale.*

La matrice des données tirée de l'atlas linguistique examiné représente un schéma à double entrée groupant N objets (ou points d'atlas) et p attributs (ou cartes d'atlas, mieux : cartes « de travail ») (5). Voir la Fig. 1.

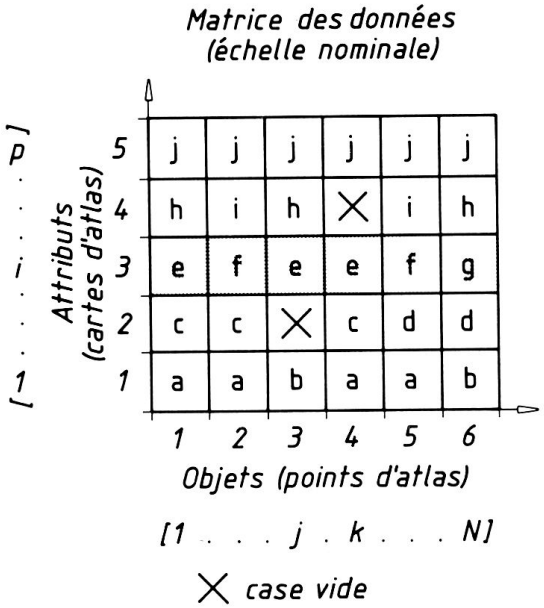


Fig. 1 :
Matrice des données constituée de 6 objets (points d'atlas) et de 5 attributs (cartes d'atlas).
En gris : l'équivalent matriciel d'une carte d'atlas.

---

c) *une carte morphosyntaxique di*nyme avec les types (taxats) ⌜*le*⌝ et ⌜*ce*⌝. D'où l'explication du fait que des 1421 cartes-*ALF* nous avons tiré, par voie taxatoire (codage), 1468 cartes *de travail.*

La Fig. 1 montre le modèle d'une matrice des données issue de la mensuration des données originales (en l'occurrence d'un atlas linguistique) sur l'échelle métrologique la plus simple, à savoir sur l'échelle *nominale* (6). Des particularités caractérologiques qui marquent un certain nombre d'objets examinés, l'échelle nominale ne retient que les différenciations oppositives (à opposition simple, binaire, ternaire, etc.) sans préciser davantage le degré tant qualitatif que quantitatif des particularités des objets mesurés. En regardant la Fig. 1 l'on constate que la carte 1 dispose de deux taxats(-types) (a et b : carte *di*nyme), la carte 2 également de deux taxats (c et d : carte *di*nyme), la carte 3 de trois taxats (e, f, g : carte *tri*nyme), la carte 4 de deux taxats (h et i : carte *di*nyme) et la carte 5 d'un seul taxat j : carte *mono*nyme). Un examen analogue du morcellement taxatoire (qui, en fait, n'est rien d'autre qu'un morcellement *onomasiologique*) de la matrice des données utilisée dans cet article aboutit à la liste suivante :

N-TYP	N-KART	%-TOT
1	594	40.46
2	351	23.91
3	184	12.53
4	111	7.56
5	78	5.31
6	41	2.79
7	33	2.25
8	20	1.36
9	18	1.23
10	10	.68
11	7	.48
12	4	.27
13	8	.54
15	2	.14
16	4	.27
17	1	.07
19	2	.14
	1468	

Nombre des taxats par carte analysée.

(6) Pour plus de détails relatifs aux échelles de mesure cf. Chandon/Pinson 1981, 33-38.

La première ligne de cette liste est à interpréter comme suit : « De 1468 cartes analysées (colonne centrale, N-KART, en bas) 594 (colonne N-KART, en haut), ou 40,46 % (= 594 sur 1468) (colonne %-TOT, en haut), ne disposent que d'un *seul* taxat-type (colonne N-TYP, en haut) et sont donc *mono*nymes. » (7)

Nous recommandons aux lecteurs de bien retenir le symbolisme utilisé dans la Fig. 1. Nous nous en servirons pour toutes les explications ultérieures.

2. Pondération et mesure de la similarité.

Il y a, grosso modo, deux manières de mesurer la *similarité* (*identité, proximité, ressemblance, affinité, parenté,* etc.) *nominale* entre deux vecteurs d'objet d'une matrice de données :

a) la mesure *non pondérée* (*isocratique*) de la similarité

b) la mesure *pondérée* (*anisocratique*) de la similarité (8).

Nécessairement, le choix entre ces deux possibilités dépendra des intentions classificatoires du dialectométricien. Sans trop nous attarder sur le sujet très épineux de la finalité de toute analyse classificatoire — cf. à ce sujet les lignes lumineuses de l'anthropologue américain Kroeber (1960) — nous nous bornerons par la suite à traiter uniquement des aspects méthodiques des deux possibilités de mesure mentionnées ci-dessus.

2.1. Mesure non pondérée (isocratique) de la similarité : calcul de l'IRI_{jk}.

Il s'agit du calcul de l'« Indice Relatif d'Identité (IRI_{jk}) » (9) (en allemand : « Relativer Identitätswert (RIW_{jk}) ») entre les vecteurs

(7) En ce qui concerne d'autres aspects tant qualitatifs que quantitatifs de la matrice des données utilisée dans cet article cf. Goebl 1984 I, 55-70.

(8) Pour plus ample information :
point de vue dialectométrique : cf. Goebl 1981, 356 ; 1983b, 12-17 ; 1984 I, 83-86 ;
point de vue taxométrique : cf. Chandon/Pinson 1981, 61-62, 64 et Bock 1974, 68-71, 72-73.

(9) Pour l'historique du nom de cet indice de similarité cf. Goebl 1981, 357, note 10.

d'objet j et k. Pour une présentation de l'IRI_{jk} cf. aussi Goebl 1976, 173 ; 1981, 357-361 ; 1983a, 363-366 et 1984 I, 75-77. Pour une explication du schéma de mesure voir la Fig. 2.

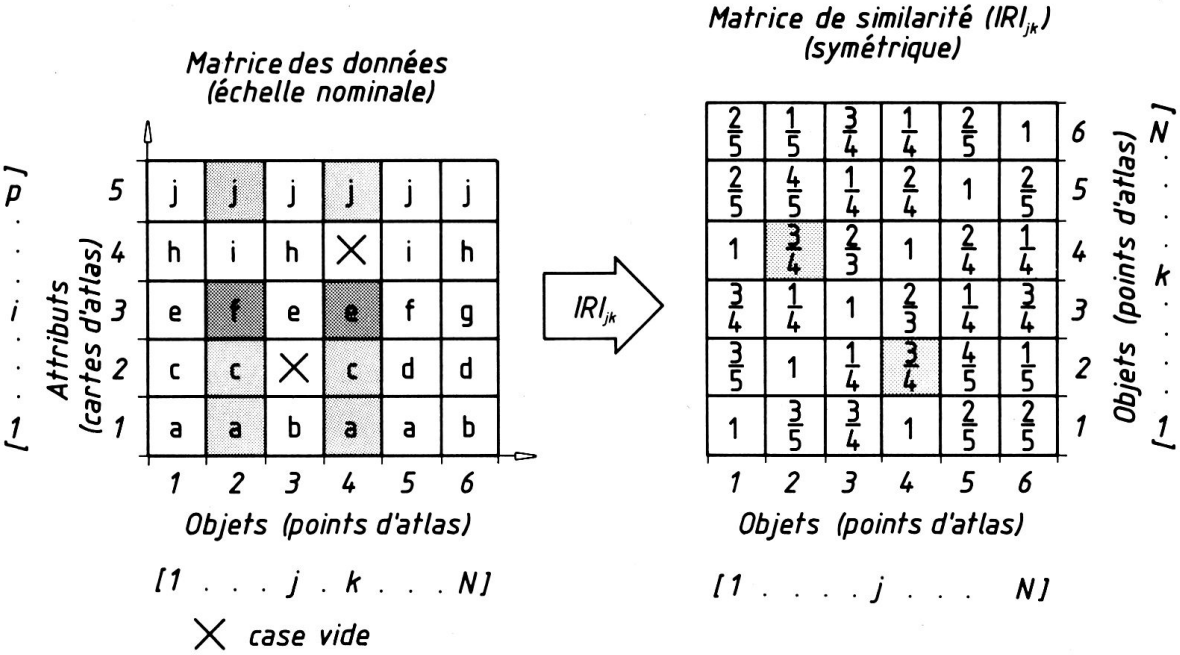


Fig. 2 :

Génération de la matrice de similarité à partir de la matrice des données par l'application de l'Indice Relatif d'Identité (IRI_{jk}).

— A gauche (matrice des données) :
- grisé clair : co-identités (COI) pour $i = 1$ (a, a), pour $i = 2$ (c, c) et pour $i = 5$ (j, j) ;
- grisé sombre : co-différence (COD) pour $i = 3$ (f, e).

— A droite (matrice de similarité) :
- en gris : $IRI_{2,4} = IRI_{4,2} = \frac{3}{4}$ (= 0,75)
  (valeurs sans mise en pourcentage).

La formule de l'IRI_{jk} repose sur le concept taxométrique bien connu de la *co-identité* (COI_{jk}) (10) et de la *co-différence* (COD_{jk}) entre deux

(10) Pour le concept de la *co-identité* (COI) cf. Chandon/Pinson 1981, 73-75. Les notions de *co-identité, co-différence* et *co-présence* (non citée dans cet article) s'expliquent par le fait que toute mesure de similarité s'effectue par *paires* de vecteurs d'attribut.

vecteurs d'objet (j : point de référence, k : point comparé) à l'endroit d'un attribut donné (i). En voici la formule :

$$IRI_{jk} = 100 \cdot \frac{\sum_{i=1}^{\tilde{p}} (COI_{jk})i}{\sum_{i=1}^{\tilde{p}} (COI_{jk})i + \sum_{i=1}^{\tilde{p}} (COD_{jk})i} \qquad (1) \cdot$$

Dans cette formule,

IRI_{jk}	est l'Indice Relatif d'Identité.
$\tilde{p}$	est le nombre des attributs présents tant dans le vecteur du point j que dans le vecteur du point k.
$(COI_{jk})i$	est une co-identité entre les points j et k à l'endroit de l'attribut i.
$(COD_{jk})i$	est une co-différence entre les points j et k à l'endroit de l'attribut i.
j	est l'indicatif du point de référence.
k	est l'indicatif du point comparé.
i	est l'indicatif d'un attribut.

Etant donné le fait que — à cause des lacunes très fréquentes dans les atlas linguistiques — la présence de *cases vides* (voir la Fig. 2) est inévitable dans la matrice des données et que, partant, les vecteurs d'objet ne disposent pas toujours de p attributs, il a fallu insérer dans la formule (1) le terme $\tilde{p}$ qui peut être ou bien inférieur ou bien égal à p.

Vérification de l'IRI_{jk} pour $j = 2$ et $k = 4$ à l'aide de la Fig. 2 :

- Remarquer que la présence d'une case vide dans $i = 4$ cause la diminution de $\tilde{p}$ de 5 à 4.
- Nombre des co-identités entre les vecteurs d'objet (i.e. entre les points) 2 et 4 :

$$\sum_{i=1}^{4} (COI_{2,4})i = 3.$$

- Nombre des co-différences entre les vecteurs d'objet (i.e. entre les points) 2 et 4 :

$$\sum_{i=1}^{4} (COD_{2,4})i = 1.$$

- Calcul de l'$IRI_{2,4}$ selon la formule (1) :

$$IRI_{2,4} = (100) \cdot \frac{3}{3+1} = (100) \cdot \frac{3}{4} = 75\ \%.$$

Les valeurs de la matrice de similarité de la Fig. 2 n'ont pas été mises en pourcentage.

La vertu isocratique de l'IRI consiste dans le fait que toutes les co-identités (COI) et toutes les co-différences (COD) prennent la valeur 1. Une mensuration anisocratique des mêmes données empiriques pourrait se faire à l'aide d'une pondération numérique adéquate de la valeur ou bien des co-identités (COI) ou bien des co-différences (COD). Ici, nous ne prenons en considération que la pondération des co-identités (COI). Il en sera question dans le paragraphe suivant.

2.2. Mesure pondérée (anisocratique) de la similarité : calcul de l'$IPI(w)_{jk}$.

Le sigle « $IPI(w)_{jk}$ » désigne l'« Indice Pondéré (avec le poids w) d'Identité » (en all. « Gewichtender Identitätswert ($GIW(w)_{jk}$ ») (11). Tant en géographie linguistique que dans d'autres branches de la linguistique, l'on rencontre fréquemment des points de vue qui cadrent fort bien avec le passage suivant tiré d'un excellent manuel de taxométrie : « The philosophy here is that agreement among rare character states is a less probable event than agreement for frequent character states and should therefore be weighted more heavily. » (Sneath/Sokal 1973, 140). Voir aussi les remarques judicieuses de H. Kuen, 1973, 336-339, 349, 350) qui représentent un panorama bien équilibré de la « philosophie pondératrice » traditionnelle en linguistique romane.

(11) Pour l'historique du nom allemand de l'$IPI(w)_{jk}$ cf. Goebl 1983b, 12-17.

En d'autres termes : plus une strate géolinguistique est grande, moins elle contribue à l'établissement de la similarité globale. Voir à ce sujet la Fig. 3.

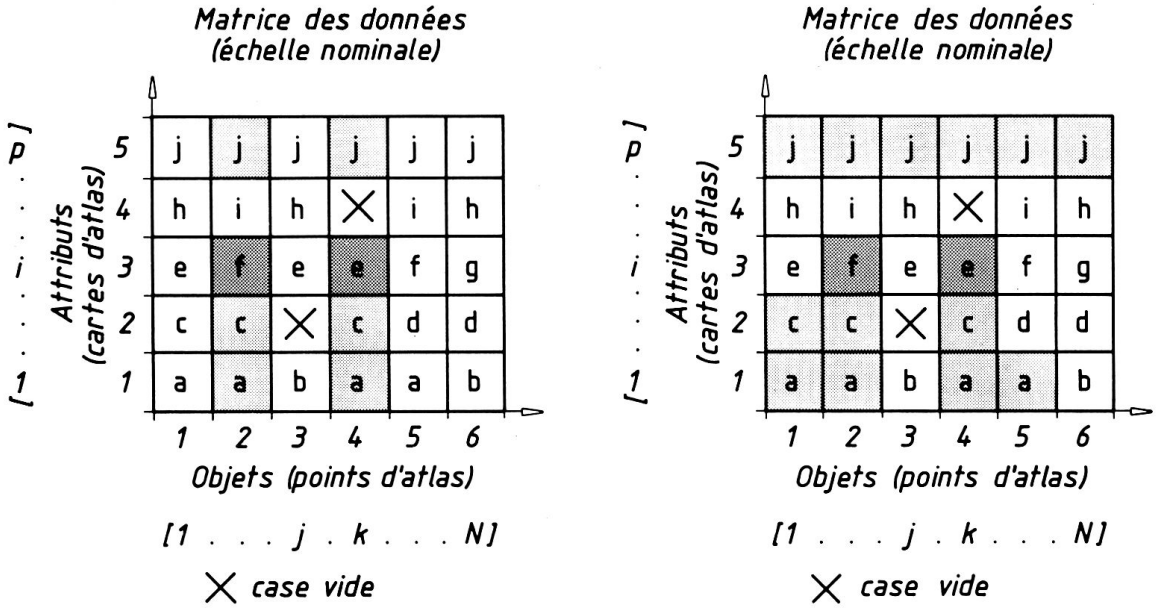


Fig. 3 :

Mesure isocratique et mesure anisocratique des données empiriques d'une matrice des données.

— A gauche : mesure isocratique (non pondérée) des vecteurs d'objets 2 (j) et 4 (k) :

- grisé clair : co-identités (COI) pour $i = 1$ (a, a), pour $i = 2$ (c, c) et pour $i = 5$ (j, j) ;
- grisé sombre : co-différence (COD) pour $i = 3$ (f, e).

(Les co-identités ainsi que les co-différences équivalent à la valeur 1.)

— A droite : mesure anisocratique (pondérée) des vecteurs d'objets 2 (j) et 4 (k) :

- grisé clair : co-identités (COI) et taxats corollaires pour $i = 1$ (a, a ; a, a), pour $i = 2$ (c, c ; c) et pour $i = 5$ (j, j ; j, j, j) ;
- grisé sombre : co-différence (COD) pour $i = 3$ (f, e).

(Les valeurs des co-identités sont au-dessous de 1 et dépendent de la fréquence absolue des taxats les constituant ; les co-différences équivalent à 1. Voir la Fig. 5.)

Dans la Fig. 3, les grisés clairs et sombres marquent les parties de la matrice des données utilisées pour une mesure de similarité entre les vecteurs d'objet 2 et 4. A gauche, il s'agit d'une mesure isocratique où, nous l'avons déjà dit plus haut, toutes les co-identités et toutes les co-différences équivalent à 1. A droite par contre, seules les co-différences continuent d'avoir la valeur standard 1, alors que la valeur standard des co-identités est pondérée en fonction de l'étendue (en points d'atlas) de la strate géolinguistique à laquelle appartiennent les taxats (tokens) qui engendrent la co-identité respective.

Vérification des possibilités de pondération à l'aide de la Fig. 3 :

- Soit la carte $i = 1$ et les points 2 ($j = 2$) et 4 ($k = 4$).
- A gauche, la co-identité dépend uniquement de la concordance nominale parfaite entre le taxat a en $i = 1/j = 2$ et en $i = 1/k = 4$.
- A droite cependant, la co-identité analogue ($i = 1$, $j = 2$, $k = 4$) dépend non seulement de la concordance nominale des taxats (tokens) a en $i = 1/j = 2$ et en $i = 1/k = 4$ mais aussi de l'étendue du taxat (type) a en $i = 1$. Comme le taxat (type) a en $i = 1$ occupe 4 cases (= points d'atlas) sur 6, il représente les deux tiers de l'étendue théorique maximale. Par contre, le taxat (type) c en $i = 2$ n'occupe que 3 cases (= points d'atlas) sur 5, étant donné le fait qu'il faut retrancher une case vide de la valeur théorique maximale des cases à occuper.

Pour une pondération adéquate des co-identités en fonction de la fréquence absolue des taxats les constituant, l'on doit donc établir une *relation algorithmique* entre les termes suivants :

- fréquence absolue d'un taxat (type) susceptible de générer une co-identité,
- nombre des cases (points d'atlas) constituant un vecteur d'attribut,
- facteur pour rendre la pondération variable.

Voici donc cette relation algorithmique mise en formule :

$$(COI_{jk})i = 1 - \frac{f(TAX_{jk})i - 1}{\tilde{N}i \cdot w} \qquad (2).$$

Dans cette formule,

$(COI_{jk})i$	est une co-identité entre les points *j* et *k* à l'endroit de l'attribut *i*.
$f(TAX_{jk})i$	est la fréquence absolue du taxat (type) constituant une co-identité entre les points *j* et *k* à l'endroit de l'attribut *i*. Cette valeur peut varier entre 2 et Ñi ([12]).
Ñi	est le nombre des points d'atlas disponibles dans le vecteur d'attribut *i* (après retranchement des cases vides). La valeur Ñi peut être égale ou inférieure à N (nombre maximal des points d'atlas ; ici : N = 71).
w	est le facteur de pondération (à choisir). La valeur de w peut varier entre 1 et ∞.

Cette formule a les propriétés arithmétiques suivantes :

- Plus $f(TAX_{jk})i$ (et partant la strate générant une co-identité) est grand, plus il pèse sur la valeur standard 1 de la co-identité.
- Plus le poids w est grand, moins il pèse sur la valeur standard 1 de la co-identité. Pour $w = \infty$ la fraction

  $\frac{f(TAX_{jk})i - 1}{Ñi \cdot w}$ prend la valeur 0 : l'$IPI(w)_{jk}$ correspond alors parfaitement à l'IRI_{jk}.

- La fréquence absolue du taxat pris en considération ($f(TAX_{jk})i$) est diminuée de 1 pour éviter qu'une carte d'atlas parfaitement mononyme et sans cases vides (d'où $Ñi = f(TAX_{jk})i$) ne fasse prendre la valeur 1 à la fraction de la formule (2). Si tel était le cas, aucune carte mononyme ne pourrait contribuer à l'établissement de la similarité globale entre deux vecteurs d'objet (points d'atlas). Comme, dans ma conception classificatoire, aussi les cartes mononymes ont leur raison d'être dans un bilan taxométrique, cette diminution de 1 visant à donner leur part (fort modeste, il est vrai : voir les courbes de la Fig. 6) aussi aux cartes mononymes, semble être légitime.

(12) La valeur $f(TAX_{jk})i = 1$ n'est pas prise en considération parce qu'il est impossible d'établir une *co*-identité (et de pratiquer ainsi une mesure par *paire*) à partir d'une strate linguistique comprenant un *seul* taxat (token).

Pour la constitution arithmétique des termes $f(TAX_{jk})i$ et $\tilde{N}i$ je renvoie à la Fig. 4.

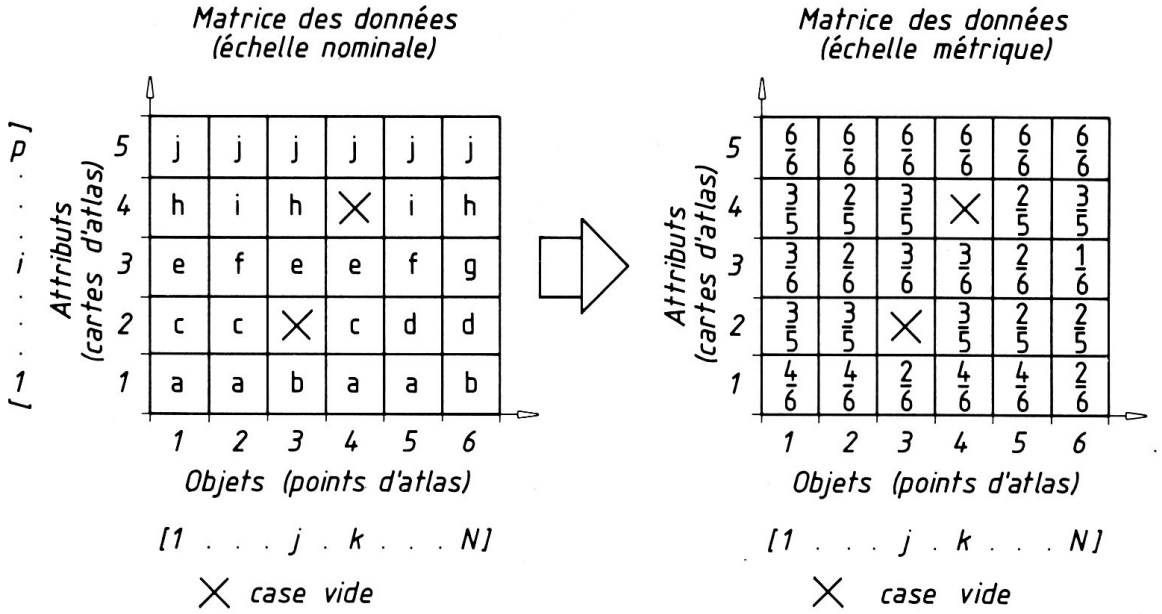


Fig. 4 :

Métrisation de la matrice des données.

La matrice des données (à gauche) correspond à celle de la Fig. 1.

Dans la Fig. 4 la matrice des données sur échelle *métrique* (à droite) contient, dans chaque case, une fraction dont le numérateur correspond à $f(TAX_{jk})i$ et le dénominateur à $\tilde{N}i$.

En combinant la formule de l'IRI_{jk} (formule 1) et la formule (2) relative à la pondération des co-identités, on peut établir la formule complète de l'$IPI(w)_{jk}$:

$$IPI(w)_{jk} = 100 \cdot \frac{\sum_{i=1}^{\tilde{p}} \left(1 - \frac{f(TAX_{jk})i - 1}{\tilde{N}i \cdot w}\right) i}{\sum_{i=1}^{\tilde{p}} \left(1 - \frac{f(TAX_{jk})i - 1}{\tilde{N}i \cdot w}\right) i + \sum_{i=1}^{\tilde{p}} (COD_{jk})i} \qquad (3).$$

Dans cette formule,

$IPI(w)_{jk}$	est l'Indice Pondéré d'Identité (avec le poids w).
$\tilde{p}$	est le nombre des attributs présents tant dans le vecteur du point *j* que dans le vecteur du point *k*.
$f(TAX_{jk})i$	est la fréquence absolue du taxat constituant une co-identité entre les points *j* et *k* à l'endroit de l'attribut *i*. Cette valeur peut varier entre 2 et Ñi ([13]).
Ñi	est le nombre des points d'atlas disponibles dans le vecteur d'attribut *i* (après retranchement des cases vides). La valeur Ñi peut être égale ou inférieure à N (nombre maximal des points d'atlas ; ici : N = 71).
$(COD_{jk})i$	est une co-différence entre les points *j* et *k* à l'endroit de l'attribut *i*.
j	est l'indicatif du point de référence.
k	est l'indicatif du point comparé.
i	est l'indicatif d'un attribut.
w	est le facteur de pondération (à choisir). La valeur de w peut varier entre 1 et ∞.

Pour l'établissement de la formule de l'$IPI(w)_{jk}$ nous nous sommes servis de propositions antérieures de Bock 1974, 68-71 ; 71-73 et d'Altmann 1977. Voir aussi la Fig. 5 qui permet de contrôler la formule (3) pour w = 1.

(13) Cf. la note 12.

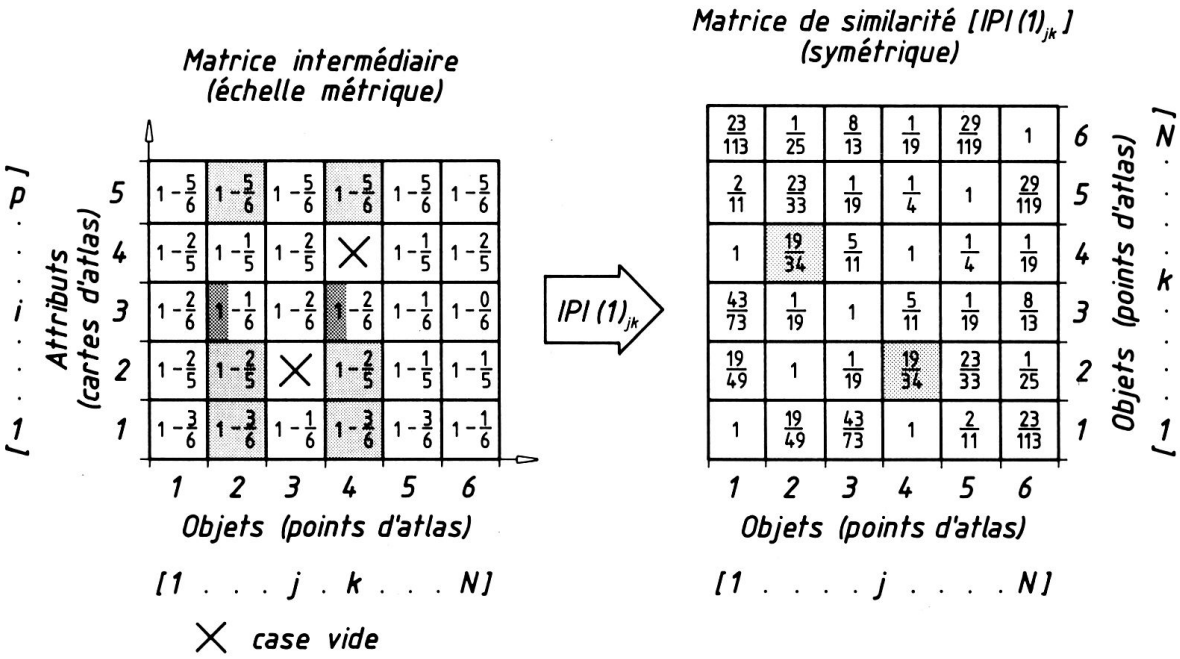


Fig. 5 :

Génération de la matrice de similarité à partir de la matrice intermédiaire par l'application de l'Indice Pondéré d'Identité [IPI(1)$_{jk}$] (avec le poids w = 1).

— A gauche (matrice intermédiaire) :

- grisé clair : co-identités pondérées [COI(1)$_{jk}$; avec le poids w = 1]

  pour $i = 1$ $(1 - \frac{3}{6})$, pour $i = 2$ $(1 - \frac{2}{5})$ et pour $i = 5$ $(1 - \frac{5}{6})$;

- grisé sombre : co-différence non pondérée (COD) pour $i = 3$ (1).

— A droite (matrice de similarité) :

- en gris : IPI(1)$_{2,4}$ = IPI(1)$_{4,2}$ = $\frac{19}{34}$

  (valeurs sans mise en pourcentage).

Vérification de l'IPI(1)$_{jk}$ pour $j = 2$ et $k = 4$ à l'aide des Fig. 3, 4 et 5 :

- Remarquer que la présence d'une case vide dans $i = 4$ cause la diminution de $\tilde{p}$ de 5 à 4 (voir les Fig. 3, 4 et 5).

- $\sum_{i=1}^{4} (1 - \frac{f(TAX_{2,4})i - 1}{Ñi \cdot 1})i = (1 - \frac{3}{6}) + (1 - \frac{2}{5}) + (1 - \frac{5}{6})$

  (voir la Fig. 5).

- $\sum_{i=1}^{4} (COD_{2,4})i = 1$

  (voir la Fig. 5).

- Calcul de l'$IPI(1)_{2,4}$ selon la formule (3) :

$$IPI(1)_{2,4} = \frac{(1 - \frac{3}{6}) + (1 - \frac{2}{5}) + (1 - \frac{5}{6})}{(1 - \frac{3}{6}) + (1 - \frac{2}{5}) + (1 - \frac{5}{6}) + 1}$$

$$= \frac{30 - 15 + 30 - 12 + 30 - 25}{30 - 15 + 30 - 12 + 30 - 25 + 30}$$

$$= \frac{90 - 52}{120 - 52}$$

$$= \frac{38}{68}$$

$$IPI(1)_{2,4} = \frac{19}{34}$$

Voir la matrice de similarité de la Fig. 5 (grisé clair) dont les valeurs n'ont pas été mises en pourcentage.

L'importance taxométrique de l'$IPI(w)_{jk}$ réside dans la possibilité de choisir librement le facteur de pondération w. Plus w est grand, plus l'$IPI(w)_{jk}$ ressemble à l'indice de similarité isocratique IRI_{jk}. Voir à cela la Fig. 6 qui réunit les tracés des fonctions pondératrices pour respectivement w = 1, w = 2 et w = 3. De l'allure des trois courbes l'on déduit aisément que pour w = 1 l'action pondératrice est maximale.

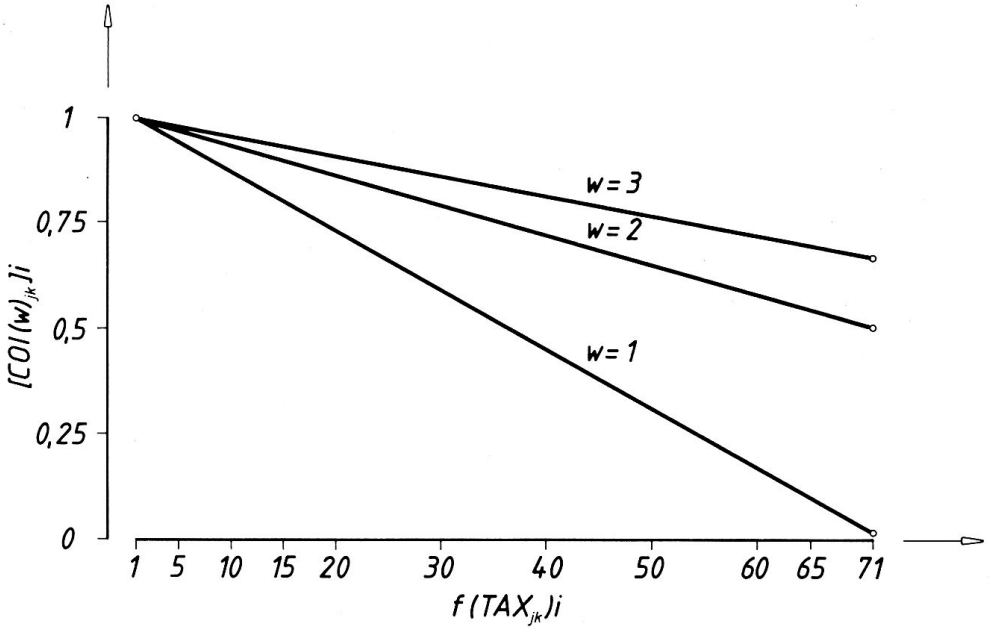


Fig. 6 :

Pondération des co-identités en fonction de la fréquence absolue des taxats les constituant et d'un poids numérique à choisir au préalable.

$[COI(w)_{jk}]i$ co-identité pondérée (par le poids w) entre les vecteurs d'objets *j* et *k* à l'endroit de l'attribut *i*.

$f(TAX_{jk})i$ fréquence absolue d'un taxat constituant une co-identité entre les vecteurs d'objets *j* et *k* à l'endroit de l'attribut *i*. $f(TAX_{jk})i$ peut varier entre 1 (avec pondération nulle) et 71 (avec pondération maximale).

Les trois courbes obéissent à la formule suivante :

$$[COI(w)_{jk}]i = 1 - \frac{f(TAX_{jk})i - 1}{w.71}$$

(pour w = 1, 2 ou 3).

Vérification de l'action pondératrice à l'aide de la Fig. 6 :

- w = 1 :

  Pour $f(TAX_{jk})i = 60$ la co-identité pondérée respective $(COI(1)_{jk})i$

est de 0,169. Pour $f(TAX_{jk})i = 20$ la valeur de $(COI(1)_{jk})i$ est de 0,732.

- $w = 3$:

  Pour $f(TAX_{jk})i = 60$ la valeur de $(COI(3)_{jk})i$ est de 0,723. Pour $f(TAX_{jk})i = 20$ la valeur de $(COI(3)_{jk})i$ est de 0,910.

Dans cet article, nous présenterons des calculs dialectométriques effectués à l'aide de $w = 1$, $w = 2$ et $w = 3$. Le choix des facteurs 1, 2 et 3 est purement arbitraire. Il eût été possible de choisir, à la place de 1, 2 et 3, tout autre nombre réel entre 1 et ∞.

3. Visualisation et mise en carte des distributions de fréquences dialectométriques.

Le pavage géométrique du fond de carte a été fait selon la méthode de Thiessen : cf. à cela Goebl 1981, 363-364 ; 1983a, 358-359 et 1984 I, 90-91. Les conventions cartographiques voire graphiques relatives tant aux cartes choroplèthes qu'aux cartes numériques respectives correspondent exactement à celles que nous avons utilisées dans nos travaux antérieurs : cf. surtout Goebl 1981, 365 sq. et 1984 I, 90 sq.

Pour l'établissement des histogrammes figurant sur les cartes choroplèthes en bas de page cf. Goebl 1981, 363 et 1984 I, 97-98.

Le but essentiel de toute démarche visualisatrice en matière de cartographie thématique est de convertir une *variabilité numérique* donnée (ici de 70 ou 71 valeurs dialectométriques différentes) en une *variabilité analogue du figuré iconique.* Il s'agit là d'un problème standard de la cartographie statistique. Pour plus ample information nous renvoyons aux ouvrages suivants, indispensables d'ailleurs pour une bonne compréhension du sujet traité : Dickinson 1973, Jenks/Caspall 1971, Jenks/Coulson 1963, Scripter 1970 et Wonka 1980 ainsi que Bertin 1977, 197-203 et Joly 1976, 199-203.

Comme la mise en carte constitue le dernier maillon de la chaîne dialectométrique (et partant taxométrique voire classificatoire), il est bon de la pratiquer d'une façon très consciencieuse et en pleine connaissance de cause. Nous présenterons par la suite trois algorithmes de visualisation capables de convertir la variation numérique (continue ou quasi-continue) en une variation iconique discrète de 6 classes (ou intervalles) représentées, dans le figuré de la carte choroplèthe, par autant de paliers de teintes hachurées (grisés).

Les trois algorithmes de *visualisation* (ou *d'intervallisation*) ont les sigles suivants :

MINMWMAX,
MEDMW,
MED.

Le lecteur désireux de suivre de près notre présentation devra toujours se reporter, pour pouvoir contrôler le cheminement de notre argumentation, aux en-têtes des cartes numériques (Fig. 28-33) Voir aussi la Fig. 7 où l'on trouvera l'en-tête de la Fig. 29 et deux flèches indiquant respectivement les seuils d'intervalle et les nombres des points d'atlas par intervalle en fonction de l'algorithme d'intervallisation MINMWMAX.

MINMWMAX	54.293	58.269	62.245	65.221	70.611	75.002	79.393
MEDMW	54.293	58.168	54.339	66.272	67.726	68.567	79.393
MED	54.293	63.005	66.623	67.609	68.035	69.580	79.393

4-3	1-3	3-8	31-14	1-0	1-1
3-3	4-4	4-4	8-8	8-8	8-8
6-6	6-6	6-6	6-6	6-6	5-5

Fig. 7 :

Détail de l'en-tête de la carte numérique de la distribution de similarité (selon $IPI(2)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

- Encadré grand : seuils numériques des intervalles de visualisation.
- Encadré petit : nombre des points d'atlas par intervalle.

La flèche renvoie aux résultats relatifs à l'algorithme d'intervallisation MINMWMAX.

Voir aussi la Fig. 29.

3.1. L'algorithme d'intervallisation MINMWMAX (à 6 intervalles).

Cet algorithme requiert les calculs suivants :

- Calcul de la *moyenne arithmétique* de la distribution de fréquence visualisée.
- Division par trois de l'*écart entre la moyenne arithmétique* et le *minimum* : il en résulte un chiffre correspondant à la *largeur des intervalles* situés au-dessous de la moyenne, à l'aide duquel on calcule aisément les limites des trois intervalles respectifs (voir les colonnes 1-3 de la Fig. 8).

- Même processus pour l'*écart* entre le *maximum* et la *moyenne arithmétique.*

Vérifier les étapes du calcul décrit ci-dessus par une consultation simultanée des Fig. 7 (voire 29) et 8.

Algorithme d'intervallisation: MINMWMAX (à 6 intervalles)				
1	2		3	4
Intervalle	Variable: $IPI(2)_{jk}$		Largeur de l'intervalle	Nombre des points d'atlas par intervalle
	de	à		
1	54,293[a)]	58,269	3,976	4 + 3 = 7
2	> 58,269	62,245	3,976	1 + 3 = 4
3	> 62,245	66,221[b)]	3,976	3 + 8 = 11
4	> 66,221	70,611	4,3906	31 + 14 = 45
5	> 70,611	75,002	4,3906	1 + 0 = 1
6	> 75,002	79,393[c)]	4,3906	1 + 1 = 2
	a) minimum (valeur discrète)			70
	b) moyenne arithmétique (valeur continue)			
	c) maximum (valeur discrète)			

Fig. 8 :

Calcul des intervalles de visualisation à l'aide de l'algorithme d'intervallisation MINMWMAX (à 6 intervalles).

Pour les données numériques voir l'en-tête de la Fig. 29.

Pour plus de détails relatifs à MINMWMAX cf. Goebl 1981, 361-363 ; 1982, 29 ; 1983a, 370 et 1984 I, 94.

3.2. L'algorithme d'intervallisation MEDMW.

Pour l'utilisation de cet algorithme il faut procéder de la manière suivante :

- Les 70 (ou 71) valeurs de la distribution de fréquence à visualiser sont rangées par ordre croissant à commencer par le minimum.

- L'on cherche, parmi les valeurs ainsi rangées par rang de taille, la valeur la plus proche de la moyenne arithmétique.
- Les valeurs entre le minimum et la valeur contiguë à la moyenne arithmétique sont réunies dans un groupe.
- Même processus pour les valeurs restantes au-dessus de la valeur contiguë à la moyenne arithmétique. L'on obtient ainsi deux groupes de valeurs rangées par rang de taille.
- Chacun des deux groupes est divisé, dans la mesure du possible, en trois classes sensiblement égales par rapport au nombre des valeurs de la distribution de fréquence à visualiser. Voir à cela la colonne 4 de la Fig. 9.

Vérifier les résultats du procédé décrit ci-dessus par une consultation simultanée des Fig. 7 (voire 29) et 9.

Algorithme d'intervallisation: MEDMW (à 6 intervalles)				
1	2		3	4
Intervalle	Variable: $IPI(2)_{jk}$		Largeur de l'intervalle	Nombre des points d'atlas par intervalle
	de	à		
1	54,293[a)]	58,168	3,875	3 + 3 = 6
2	> 58,168	64,339	6,171	4 + 4 = 8
3	> 64,339	66,272[b)]	1,933	4 + 4 = 8
4	> 66,272	67,726	1,454	8 + 8 = 16
5	> 67,726	68,567	0,841	8 + 8 = 16
6	> 68,567	79,393[c)]	10,826	8 + 8 = 16
	a) minimum (valeur discrète)			70
	b) valeur proche de la moyenne arithmétique (valeur discrète)			
	c) maximum (valeur discrète)			

Fig. 9 :

Calcul des intervalles de visualisation à l'aide de l'algorithme d'intervallisation MEDMW (à 6 intervalles).

Pour les données numériques voir l'en-tête de la Fig. 29.

3.3. L'algorithme d'intervallisation MED.

Cet algorithme rend nécessaires les calculs suivants :

- Les 70 (ou 71) valeurs de la distribution de fréquence à visualiser sont rangées par ordre croissant à commencer par le minimum.
- Par la suite, ces 70 (ou 71) valeurs sont réparties en 6 groupes sensiblement égaux. Comme la division de 70 (ou de 71) par 6 ne donne pas de résultat discret (70 : 6 = 11,6 ; 71 : 6 = 11,83) et que, parmi les 70 (ou 71) scores examinés, quelques-uns peuvent avoir des valeurs numériquement égales, l'algorithme MED doit se contenter de former des groupes de scores aussi proches de la valeur théorique (11,6 ou 11,83) que possible. Voir à cela la colonne 4 de la Fig. 10.

Vérifier les résultats de ce procédé par une consultation simultanée des Fig. 7 (voire 29) et 10.

Algorithme d'intervallisation: MED (à intervalles)				
1	2		3	4
Intervalle	Variable: $IPI(2)_{jk}$		Largeur de l'intervalle	Nombre des points d'atlas par intervalle
	de	à		
1	54,293[a)]	63,005	8,712	6 + 6 = 12
2	> 63,005	66,623	3,618	6 + 6 = 12
3	> 66,623	67,609[b)]	0,986	6 + 6 = 12
4	> 67,609	68,035	0,426	6 + 6 = 12
5	> 68,035	69,580	1,545	6 + 6 = 12
6	> 69,580	79,393[c)]	9,813	5 + 5 = 10
	a) minimum (valeur discrète)			70
	b) médiane (valeur discrète)			
	c) maximum (valeur discrète)			

Fig. 10 :

Calcul des intervalles de visualisation à l'aide de l'algorithme d'intervallisation MED (à 6 intervalles).

Pour les données numériques voir l'en-tête de la Fig. 29.

4. Notice servant à l'intelligence des Fig. 14-27 (cartes choroplèthes) et des Fig. 28-33 (cartes numériques).

Les cartes numériques (voir les Fig. 28-33) contiennent chacune un en-tête (reproduit en haut de page, au milieu) et des indications numériques relatives aux 71 points d'atlas. Pour l'intelligence des en-têtes voir la Fig. 7, pour l'intelligence des indications numériques relatives aux différents points d'atlas voir la Fig. 11.

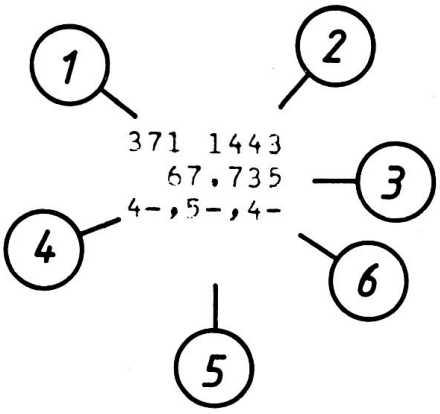


Fig. 11 :

Données numériques relatives à un point d'atlas.
Voir aussi la Fig. 29.

1 Indicatif du point d'atlas (ici : P. 371).

2 Nombre des co-présences (co-identités + co-différences) entre les vecteurs d'objets j (ici : P. 397) et k (ici : P. 371).

3 $IPI(w)_{jk}$ (ici : $IPI(2)_{397,371}$).

4 Classement par intervalle selon MINMWMAX (ici : intervalle 4 selon MINMWMAX à 6 intervalles, intervalle 4 — (= intervalle 7) selon MINMWMAX à 12 intervalles).

5 Classement par intervalle selon MEDMW (ici : intervalle 5 selon MEDMW à 6 intervalles, intervalle 5 — (= intervalle 9) selon MEDMW à 12 intervalles).

6 Classement par intervalle selon MED (ici intervalle 4 selon MED à 6 intervalles, intervalle 4 — (= intervalle 7) selon MED à 12 intervalles).

Comme les étapes classificatoires (voire taxométriques) de la pondération et de la visualisation comprennent chacune trois jalons (pondération : w = 1, 2 et 3 pour $IPI(w)_{jk}$; visualisation : MINMWMAX, MEDMW et MED) il est bon d'en étudier le rendement classificatoire par l'établissement d'un tableau à double entrée. Il s'agit là du processus heuristique classique du croisement de deux variables. Cette démarche comparative permet de bien faire ressortir les avantages et les inconvénients des deux critères classificatoires pris en considération. Voir à cela les Fig. 12 et 13.

Pondération				
	$IPI(3)_{jk}$	Fig. 16 (Fig. 30)	Fig. 19 (Fig. 30)	
	$IPI(2)_{jk}$	Fig. 15 (Fig. 29)	Fig. 18 (Fig. 29)	Fig. 20 (Fig. 29)
	$IPI(1)_{jk}$	Fig. 14 (Fig. 28)	Fig. 17 (Fig. 28)	
		MINMWMAX	MEDMW	MED
		Visualisation		

Fig. 12 :

Répartition des cartes choroplèthes (et des cartes numériques) selon les critères de la *pondération* et de la *visualisation*.

Moment taxométrique : mesure (anisocratique) de la similarité (selon $IPI(w)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

En gris : combinaisons non présentées dans cet article.

La Fig. 12 montre le traitement simultané selon les critères de la visualisation et de la pondération qu'a reçu la distribution de similarité (selon $IPI(w)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

Pondération				
	$IPI(3)_{jk}$	Fig. 23 (Fig. 33)	Fig. 26 (Fig. 33)	
	$IPI(2)_{jk}$	Fig. 22 (Fig. 32)	Fig. 25 (Fig. 32)	Fig. 27 (Fig. 32)
	$IPI(1)_{jk}$	Fig. 21 (Fig. 31)	Fig. 24 (Fig. 31)	
		MINMWMAX	*MEDMW*	*MED*
		Visualisation		

Fig. 13 :

Répartition des cartes choroplèthes (et des cartes numériques) selon les critères de la *pondération* et de la *visualisation.*
Moment taxométrique : synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(w)_{jk}$).
En gris : combinaisons non présentées dans cet article.

La Fig. 13 synthétise sept traitements différents d'un moment taxométrique bien défini, à savoir celui de la synopse des coefficients d'asymétrie de Fisher tirée de 71 distributions de similarité (selon $IPI(w)_{jk}$).

5. Calcul et mise en carte d'une distribution de similarité sous le double effet de la pondération et de la visualisation.

Voir les Fig. 14-20 (cartes choroplèthes) et les Fig. 28-30 (cartes numériques).

Les Fig. 14-20 représentent toutes des cartes de similarité dont le point de référence est le point-ALF 397 (La Trinité, Jersey). Le schéma taxométrique de la carte de similarité (relative à un point de référence donné) constitue un des piliers fondamentaux de la dialectométrie actuelle. Le but d'une carte de similarité est de faire voir la *ressemblance* (ou en d'autres termes : la *similarité,* l'*interpénétration,* la *concaténation,* la *diffusion,* l'*interaction,* etc.) du potentiel géolinguistique d'un point d'atlas préalablement choisi (i.e. du point de référence) avec celui des autres points d'atlas du réseau géolinguistique examiné. Du reste, toute carte de similarité correspond, en termes taxométriques précis, à un vecteur d'objet de la matrice de similarité : voir la Fig. 2 et

la matrice de similarité y représentée. La distribution de similarité du vecteur d'objet 1 comprend les 6 valeurs suivantes : 1, 3/5, 3/4, 1, 2/5, 2/5. Il faut en retrancher la première valeur (1) puisqu'elle correspond à la similarité du point de référence envers lui-même. Au sujet de la carte de similarité cf. Goebl 1981, 368-374 ; 1982, 31-37 et 1984 I, 100-112.

5.1. Interprétation géolinguistique des Fig. 14-20.

D'une façon générale, on peut dire que les fondements de la syntaxe iconique des Fig. 14-20 restent grosso modo inchangés. L'agencement des plages choroplèthes (représenté par l'interaction spatiale des 6 paliers de grisés hachurés) fournit à l'œil l'impression d'un massif montagneux dont les sommets se trouvent aux alentours du point de référence (plages choroplèthes dans les intervalles 6 et 5) et dont la crête principale s'étend en direction Ouest-Est tout en dégageant un « haut-plateau central » (représenté par les plages choroplèthes de l'intervalle 4) englobant les départements du Calvados, de l'Eure et de la Seine-Maritime : voir surtout les Fig. 15 et 16. On constate en outre que les polygones situés dans les départements de la Mayenne et de l'Ille-et-Vilaine d'un côté, et dans les départements du Pas-de-Calais, de la Somme et de l'Oise de l'autre, occupent presque tous des intervalles au-dessous de la moyenne arithmétique (visualisation par MINMWMAX et MEDMW) ou de la médiane (visualisation par MED). Ils constituent par là ce que l'on pourrait appeler, du point de vue géomorphologique, une *faille* et, du point de vue classificatoire, les *antipodes typologiques* du point de référence.

Dans les paragraphes suivants nous examinerons les sept cartes choroplèthes issues d'un test de similarité relatif au point de référence 397 (Fig. 14-20) en fonction d'une *variation* de la *pondération* d'abord et de la *visualisation* ensuite.

5.1.1. Pondération variable et visualisation constante.

Voir les Fig. 14-16 d'un côté et les Fig. 17-19 de l'autre.

Quant à l'effet classificatoire et partant cartographique (14) des trois degrés de pondération ($IPI(1)_{jk}$, $IPI(2)_{jk}$, $IPI(3)_{jk}$) on peut constater

(14) Soit dit en passant, la classification cartographique présentée dans cet article correspond exactement à ce que les géographes appellent la « régionalisation ». Cf. à cela Joly 1976, 223-224, Bertin 1977, 222 et le recueil d'articles fort intéressant de Sedlacek 1978.

que l'indice d'identité le plus pondéré, c'est-à-dire $IPI(1)_{jk}$, produit des agencements iconiques beaucoup plus accidentés que ne le font les indices d'identité moins pondérés. Que l'on compare, à ce sujet, la partie occidentale des Fig. 14, 15 et 16. Alors que les Fig. 15 et 16 offrent à l'œil, par le groupement compact de quelque 45 polygones dans l'intervalle 4 en position centrale, l'impression de « haut-plateau » (dont il a été question plus haut), il n'en est rien pour la Fig. 14 où les 28 polygones rangés dans l'intervalle 4 n'occupent que la Normandie occidentale (Basse-Normandie) et les zones limitrophes. L'indice $IPI(1)_{jk}$ semble donc engendrer des classifications géolinguistiques plus nuancées et moins globales, plus graduées et moins agglutinantes. On peut dire qu'il sert à dégager plutôt des *micro*-structures aux dépens des *macro*-structures. Cette impression se trouve confirmée par la comparaison des Fig. 17, 18 et 19 où, une fois de plus, la pondération plus forte engendre des implantations zonales plus accidentées (voir surtout la Fig. 17) alors qu'une pondération moins forte semble être propice à la genèse de structures choroplèthes plus homogènes (considérer surtout, sur les Fig. 18 et 19, les polygones situés au Cotentin).

Pour comprendre l'effet « pulvérisateur » d'une pondération plus intense, il est bon de se rappeler les graphiques de la Fig. 6. La droite relative au poids $w = 1$ montre que le rendement taxométrique (défini par la valeur de la co-identité pondérée) des strates géolinguistiques très étendues (de, mettons, plus de 40 à 50 points d'atlas) est de loin inférieur au rendement taxométrique des strates peu étendues. Or, tout géolinguiste sait que la plupart des strates peu étendues appartiennent à la catégorie ou bien des strates *innovatrices* ou bien des strates *conservatrices*. Les strates innovatrices sont ce que l'on appelle communément des foyers d'irradiation d'innovations linguistiques ([15]), alors que les strates conservatrices correspondent le plus souvent à des zones linguistiquement plutôt archaïques (îlots conservateurs). Or, selon le commun avis de beaucoup de linguistes, les traits innovateurs comme les trais archaïques revêtent une importance toute particulière pour la caractérisation typologique des idiomes considérés.

Le relief des Fig. 14 et 17 repose donc beaucoup plus sur les propriétés innovatrices et conservatrices du domaine dialectal examiné que celui des Fig. 15 et 18, voire 16 et 19. Soit dit en passant, il est possi-

(15) Que l'on songe au terme de « parachutage » conçu par P. Gardette (1955, 169).

ble de distinguer, sur la Fig. 14, les zones où prédominent les innovations (PP. 999, 226, 239, 249, etc.) des zones où, selon toute vraisemblance, les archaïsmes l'emportent sur les innovations linguistiques (Manche, Calvados occidental, Ille-et-Vilaine).

5.1.2. Visualisation variable et pondération constante.

Comparer les Fig. 14 et 17, les Fig. 15, 18 et 20 ainsi que les Fig. 16 et 19. Comparer en outre la Fig. 16 avec la carte 3.31 dans Goebl 1984 III, 70-71 (point de référence : P.-ALF 399).

En ce qui concerne l'effet de différents algorithmes d'intervallisation sur l'agencement de la syntaxe iconique des cartes choroplèthes, il est bon de commencer par comparer entre elles les Fig. 15, 18 et 19. Alors que l'algorithme MINMWMAX engendre un tissu polygonal peu accidenté (remarquer la vaste étendue des polygones rangés dans l'intervalle 4), les algorithmes MEDMW et MED permettent de saisir les détails du décalage spatial des relations d'affinité linguistique avec le point de référence 397 (Jersey). De par la distribution plus régulière des 6 intervalles et partant des 6 teintes hachurées à l'intérieur de la carte, l'algorithme MEDMW réussit à mieux saisir la diminution progressive de la similarité géolinguistique vis-à-vis de la dialecticité de Jersey. Remarquer l'effondrement progressif et non linéaire de l'identité typologique vers l'Est et l'importance que semble revêtir le français standard (P. 999) pour la constitution interne de la dialecticité du point de référence 397.

Alors que les algorithmes MINMWMAX et MEDMW ont ceci en commun qu'ils répartissent, chacun à sa manière, les intervalles 1-3 au-dessous et les intervalles 4-6 au-dessus de la moyenne arithmétique, l'algorithme MED se passe de la moyenne arithmétique tout en distribuant les 6 intervalles d'une façon numériquement aussi juste que possible entre les valeurs extrêmes du minimum et du maximum. Il en résulte une certaine mise en relief des micro-structures de la variable cartographiée. Remarquer à titre d'illustration et en comparant les Fig. 18 et 20, que les antipodes typologiques du point de référence (rangés dans l'intervalle 1), à savoir les polygones de la Picardie occidentale, sont beaucoup plus marqués sur la Fig. 20 (MED) que sur la Fig. 18 (MEDMW). Remarquer en outre que les voisins typologiques les plus proches du point de référence (rangés dans l'intervalle 6), c'est-à-dire les polygones de la côte occidentale du Cotentin, disposent, sur la Fig. 20 (MED), de moins de polygones rangés dans l'intervalle 6 que sur la

Fig. 18 (MEDMW). D'où un certain équilibre de la syntaxe iconique entière.

Le lecteur constatera en outre que le point 999, rangé dans l'intervalle 6 sur la Fig. 18 (MEDMW), n'occupe, sur la Fig. 20 (MED), que l'intervalle 5. Cette « perte de dignité » sur le plan cartographique n'enlève d'ailleurs rien à l'importance linguistique du point 999 pour la dialecticité du point de référence 397.

Au total, on peut donc dire que l'algorithme MINMWMAX rehausse les macro-structures d'une distribution dialectométrique à visualiser, alors que les algorithmes MEDMW et MED en accentuent les micro-structures tout en fournissant des reliefs plus tourmentés. Comparer aussi les Fig. 14 et 17 d'un côté ainsi que les Fig. 16 et 19 de l'autre : partout la même transformation de macro-structures aux contours sommaires (Fig. 14, 16) en micro-structures finement accidentées (Fig. 17, 19).

5.2. Bilan cartographique différentiel des Fig. 14-20.

Le choix d'un indice pondéré d'identité ou d'un algorithme d'intervallisation devra toujours se faire en fonction d'une théorie classificatoire préconçue. Un dialectométricien désireux de classifications globales voire agglutinantes aura toujours intérêt à utiliser d'un côté un indice d'identité peu ou pas pondéré et l'algorithme MINMWMAX de l'autre, alors qu'une recherche dialectométrique visant le dégagement de micro-structures finement articulées devra recourir à une pondération plus intense et à des algorithmes d'intervallisation tels que MEDMW et MED.

6. Calcul et mise en carte de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité.

Voir les Fig. 21-27 (cartes choroplèthes) et les Fig. 31-33 (cartes numériques).

Chaque distribution de similarité dispose de plusieurs paramètres statistiques typiques comme, par exemple, le *minimum,* le *maximum,* la *moyenne arithmétique,* la *médiane,* etc., parmi lesquels compte aussi le *coefficient d'asymétrie* (dit aussi « de Fisher » ([16])). Le degré de symétrie

(16) Sir Ronald A. Fisher (1890-1962), statisticien anglais. Le coefficient d'asymétrie de Fisher s'appelle *skewness* en anglais et *Schiefe* en allemand.

voire d'asymétrie d'une distribution de similarité peut être contrôlé aisément par la considération de l'histogramme respectif : comparer à titre d'exemple les histogrammes des Fig. 14 et 17 qui, bien qu'avec des largeurs et hauteurs de colonne divergentes, montrent la même silhouette extérieure. Il va de soi que cette comparaison ne peut se faire que si les histogrammes comparés reposent sur la même distribution de fréquence. Or, les Fig. 14-20 reposent sur trois distributions de fréquence différentes. Il est donc évident que les silhouettes des trois groupes d'histogrammes en question montrent, chacune, une asymétrie typique.

Voici donc la formule du coefficient d'asymétrie de Fisher ([17]) :

$$g_j = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{n \cdot s^3_j} \qquad (4).$$

Dans cette formule,

g_j	est le coefficient d'asymétrie (de Fisher) d'une distribution de similarité relative au point de référence *j*.
j	est l'indicatif du point de référence d'une distribution de similarité.
i	est l'indicatif (allant de 1 . . . 70) des différents scores de similarité (calculés selon IRI_{jk} ou $IPI(w)_{jk}$).
n	est le total (= 70) des scores de similarité (n = N — 1 ; pour N = 71 points d'atlas).
x_i	est un des 70 scores de similarité.
$\bar{x}$	est la moyenne arithmétique d'une distribution de similarité ($\bar{x}$ = MW ; cf. le paragraphe 8).
s_j	est l'écart type (all. Standardabweichung) d'une distribution de similarité relative au point de référence *j*.

Pour plus de détails voir Goebl 1981, 394-401 ; 1982, 44-46 ; 1983b, 432-436 et 1984 I, 150-154.

Le coefficient d'asymétrie prend la valeur 0 quand la distribution de fréquence est symétrique, des valeurs positives quand la distribution

(17) Cf. aussi Gendre 1977, 38 et Bahrenberg/Giese 1975, 54.

est asymétrique vers la gauche et des valeurs négatives quand la distribution est asymétrique vers la droite. L'utilité taxométrique du coefficient d'asymétrie repose donc tant sur la taille du coefficient que sur la variation du signe opérationnel.

En ce qui concerne la valeur classificatoire et partant géolinguistique du coefficient de Fisher, on peut dire qu'il nous révèle un aspect très important de la réalité dialectale qui est resté relativement inexploré jusqu'aujourd'hui. Il s'agit d'une des propriétés les plus importantes de tout réseau interactif — et les réseaux d'atlas linguistique en font partie — à savoir celle de la *connectivité* d'un point (ou nœud) au sein d'un réseau donné. On peut affirmer sans ambages et au prix d'une simplification inévitable que plus le coefficient d'asymétrie est grand, moins le point d'atlas respectif (ou, pour être plus précis : le vecteur d'attribut du point d'atlas respectif) est « intégré » dans l'ensemble des autres points d'atlas. Ceci implique bien sûr que plus le coefficient d'asymétrie est petit, plus le point d'atlas respectif est « sociologiquement » intégré et joue donc, métaphoriquement parlant, le rôle du « médiateur » ou du « boute-en-train ». Le rôle inverse, celui du « trouble-fête », serait celui d'un point d'atlas au coefficient d'asymétrie très grand. Pour jouer le plein de cette métaphore socio-dialectologique il faut dire que l'emplacement des points d'atlas-« médiateurs » correspond le plus souvent à l'endroit où se trouvent les bourrelets-frontières (faisceaux d'isoglosses) les plus importants, alors que les points d'atlas-« trouble-fête » sont confinés dans des zones périphériques ou linguistiquement très conservatrices. Pour plus de détails cf. Goebl 1981, 394-401 et 413-414 ; 1982, 44-46 ; 1983c, 432-436 et 1984 I, 150-152.

6.1. Interprétation géolinguistique des Fig. 21-27.

Le réseau examiné englobe, outre la Normandie dialectale (y compris les Iles anglo-normandes), une bonne partie des parlers gallos, manceaux, franciens et picards. Il va de soi que chacun de ces 4 groupements dialectaux devra — sur nos cartes dialectométriques — se détacher peu ou prou de ce que l'on appelle communément la Normandie dialectale à proprement parler. Or, cette différenciation spatiale ressort fort bien des Fig. 21-27. Prenons comme point de départ la Fig. 23. Les plages choroplèthes dans les intervalles 6 et 5 occupent la Picardie occidentale, alors que le pavage constitué par les polygones dans les intervalles 1 et 2 représente, de par sa position intermédiaire entre les blocs dialectaux manceau et normand, une sorte de zone-tampon. L'alignement des polygones dans l'intervalle 1 (comprenant les

PP. 368, 358, 347 et 336) correspond en outre à la section occidentale de la fameuse *ligne Joret* (cf. Joret 1883, 113, 119 et la carte hors texte). La *ligne Joret* symbolise, comme chacun sait, les abords méridionaux des parlers normands. Elle s'inscrit donc dans ce que l'on appelle communément une zone de transition. Pour une visualisation de la *ligne Joret* d'après les données de l'ALF réalisée à l'aide de la dialectométrie interponctuelle voir Goebl 1984 III, 166-173. Pour un autre traitement dialectométrique du même domaine cf. Guiter 1978 et 1984.

La syntaxe iconique de la Fig. 23 révèle en outre, et ceci surtout par le jeu différentiel des plages choroplèthes dans les intervalles 3 et 4, que les parlers normands, eux aussi, sont bien loin de constituer une surface égale. On constate au contraire, en allant de l'Ouest vers l'Est, un petit noyau dialectal « dur », à savoir le point 394 dans le Cotentin septentrional, un large creux (polygones dans l'intervalle 2) séparant la Basse-Normandie (dans l'Ouest) de la Haute-Normandie (dans l'Est) et, au sud de la dépression choroplèthe constituée par les polygones 368, 358, 347 et 336, un petit plateau relativement homogène comprenant des polygones rangés dans les intervalles 2 et 3. Nous passons sous silence les accidents du profil dialectométrique dans la partie orientale de la Fig. 23. Il en a été question plus haut.

Le profil de la Fig. 23 montre donc un assemblage de zones typologiquement plus ou moins « dures » entre lesquelles s'inscrivent des zones de transition typologiquement peu « dures » et partant plus diffuses. Il serait intéressant de voir comment les contours de ce relief choroplèthe gagneraient en netteté en fonction d'une modification alternative des procédés de pondération et de visualisation mis en œuvre.

6.1.1. Pondération variable et visualisation constante.

Voir les Fig. 21-23 d'un côté et les Fig. 24-26 de l'autre.

La pondération plus intense a été utilisée pour la confection des Fig. 21 et 24. D'une comparaison des Fig. 21 (et 23) d'un côté et des Fig. 24 (et 26) de l'autre, il ressort clairement qu'une pondération plus forte entraîne un dénivellement moins inégal du relief choroplèthe, d'où l'impression optique d'un dégagement plus marqué de ce que nous avons appelé plus haut les « micro-structures ». En regardant la Fig. 23 l'on constate en effet (en allant d'Ouest en Est) :

- Que les parlers des Iles anglo-normandes et du Cotentin nord-occidental forment désormais un noyau dialectal doublement

articulé (intervalles 6 et 4) avec des zones de transition vers le Sud et vers l'Est.

- Que les parlers gallos et manceaux constituent également un noyau bien articulé à l'intérieur (intervalles 4, 3 et 2) et nettement différencié de son entourage.
- Que les parlers de la Seine-Maritime et de la Picardie occidentale se trouvent séparés dorénavant par un couloir bien marqué (intervalle 3 : PP. 263, 259, 258, etc.).
- Qu'un noyau francien (PP. 238 et 227) s'est détaché nettement tant des parlers picards que des parlers normands orientaux.

Dans une perspective simplificatrice, on peut donc dire que par le choix d'une pondération plus forte certains aspects amalgamateurs de la surface choroplèthe de la Fig. 23 (et 24) ont pu être affaiblis voire carrément supprimés. La même constatation a déjà été faite dans le paragraphe 5.1.1.

6.1.2. Visualisation variable et pondération constante.

Comparer les Fig. 21 et 24, les Fig. 22, 25 et 27 ainsi que les Fig. 23 et 26. Comparer en outre la Fig. 23 avec la carte 3.58 dans Goebl 1984 III, 132-133 et la Fig. 16 dans Goebl 1983c, 433.

Ici encore on peut voir que les algorithmes d'intervallisation MEDMW et MED rendent les structurations choroplèthes plus tourmentées vis-à-vis de celles obtenues à l'aide de MINMWMAX, tout en creusant çà et là des dénivellations différenciatrices. D'une comparaison des Fig. 21 et 24 il résulte clairement que les noyaux dialectaux dans la Manche, dans l'Ille-et-Vilaine et la Mayenne et dans la Seine-Maritime ressortent mieux par l'application de l'algorithme MEDMW que par celle de MINMWMAX. Même contraste entre les Fig. 22, 25 et 26 d'un côté et les Fig. 23 et 26 de l'autre. Remarquer surtout l'augmentation du nombre des polygones rangés dans l'intervalle 1 d'où une meilleure structuration du relief choroplèthe par la création d'une « douve » en position centrale entourant le Calvados occidental et la Manche.

6.2. Bilan cartographique différentiel des Fig. 21-27.

Nous ne pouvons que répéter les constatations faites au paragraphe 5.2. Le choix de la pondération et de la visualisation appropriées dépendra en premier lieu du but que le dialectométricien s'est fixé. Pour une première approche globale visant à défricher un terrain resté

jusqu'alors dialectométriquement inexploré, il est bon d'utiliser un indice plutôt isocratique et partant non pondérant d'un côté et l'algorithme d'intervallisation MINMWMAX de l'autre. Ce n'est que dans une seconde étape, lorsqu'il s'agira d'approfondir l'exploration dialectométrique des données dialectales examinées, que l'on recourra, d'une façon judicieuse, à la pondération de l'indice de similarité appliqué et à l'utilisation des algorithmes d'intervallisation MEDMW et MED.

7. Epilogue et perspectives de recherche.

La théorie de toute démarche taxométrique repose sur trois piliers conceptuels fondamentaux :

objets (en all. *Dinge*)
attributs (en all. *Eigenschaften*)
et
relations (en all. *Relationen*) ([18]).

Or, en dialectologie traditionnelle, les notions d'*objet* et d'*attribut* sont déjà bien connues. Les deux termes en question signifient, dans ce contexte, « dialecte local, parler local, locolecte, topolecte, etc. » d'un côté, et « trait dialectal, caractère dialectal, propriété dialectale ou linguistique » de l'autre. Rien cependant — ou presque — n'a été fait pour élaborer le concept de *relation,* fondamental pourtant pour toute démarche comparative voire typologique ou classificatoire. Il est donc nécessaire de faire voir aux dialectologues désireux de pratiquer une classification (ou typologie) dialectométrique, l'importance du concept de *relation* et de leur faire comprendre la nécessité de sa définition adéquate par voie mathématique.

Rien dans cet article n'aurait pu être fait sans passer par l'intermédiaire de *relations.* L'échelle nominale utilisée pour la mesure des données brutes-ALF ne fonctionne qu'à l'aide de relations choisies et définies au préalable ; les indices de similarité tant isocratique qu'anisocratique représentent des exemples types de relations algébriques bien définies. La même remarque vaut pour les algorithmes d'intervallisation et — en fin de compte et mutatis mutandis — pour le travail psychophysiologique de l'œil au moment de la considération optique des

(18) Ici, nous renvoyons expressément à l'ouvrage fondamental d'Ujomov 1965. Voir aussi notre bilan épistémologique de 1983 dédié aux notions d'*objet,* d'*attribut* et de *relation* (Goebl 1983d).

reliefs choroplèthes. Il n'est peut-être pas inutile d'ajouter que le travail (à la fois analytique que synthétique) effectué tant par l'œil que par les algorithmes dialectométriques mis en œuvre s'inscrit dans le vaste champ de ce que l'on appelle de nos jours la « reconnaissance des structures » (en all. Mustererkennung, en anglais pattern recognition) ([19]) dont la dialectométrie, sans l'ombre d'un doute, fait pleinement partie.

L'avenir de la dialectométrie dépendra donc dans une large mesure du maniement judicieux du concept de *relation* de la part des dialectologues-dialectométriciens.

8. Abréviations et glossaire (termes techniques et allemands).

Les indications suivantes sont nécessaires pour une bonne compréhension des cartes choroplèthes (Fig. 14-27) et des cartes numériques (Fig. 28-33) ainsi que de certains passages de notre texte.

Choroplethenkarte	carte choroplèthe : voir les cartes choroplèthes (Fig. 14-27).
Häufigkeiten	fréquence (absolue), nombre des points d'atlas par intervalle : voir les cartes choroplèthes (Fig. 14-27).
Häufigkeitsverteilung	distribution de fréquence : voir les cartes choroplèthes (Fig. 14-27).
$IPI(w)_{jk}$	Indice Pondéré d'Identité (pondéré avec le poids w) : cf. 2.2.
IRI_{jk}	Indice Relatif d'Identité : cf. 2.1.
MED	algorithme d'intervallisation : cf. 3.3.
MEDMW	algorithme d'intervallisation : cf. 3.2.
MINMWMAX	algorithme d'intervallisation : cf. 3.1.
MW	moyenne arithmétique (all. Mittelwert) (calculée à partir de 70 scores dans les Fig. 28-30 et à partir de 71 scores dans les Fig. 31-33) : voir les cartes numériques (Fig. 28-33).
N-KART.	nombre des attributs métrologiquement disponibles dans le vecteur d'objet du point de référence : voir les Fig. 28-30.
PUNKT : 397	point de référence (i.e. ALF-P. 397) : voir les Fig. 28-30 (cartes numériques).

(19) Cf. à titre d'exemple la synthèse excellente de Niemann 1983.

S	écart type (all. Standardabweichung) (calculé à partir de 70 scores dans les Fig. 28-30, à partir de 71 scores dans les Fig. 31-33).
SCHIEFE	coefficient d'asymétrie de Fisher : voir les Fig. 31-33.
w	poids numérique utilisé pour la pondération de l'Indice Pondéré d'Identité ($IPI(w)_{jk}$) : cf. 2.2.
6-fach, 12-fach	à 6, 12 intervalles : voir les cartes choroplèthes (Fig. 14-27).

9. Bibliographie.

ALF : Atlas linguistique de la France, éd. J. Gilliéron/E. Edmont, Paris 1902-1910.

Altmann, G. : Zur Ähnlichkeitsmessung in der Dialektologie, in : Germanistische Linguistik 3-4/77 (1977), 305-310.

Bahrenberg, G./E. Giese : Statistische Methoden und ihre Anwendung in der Geographie, Stuttgart 1975.

Bertin, J. : La graphique et le traitement graphique de l'information, Paris 1977.

Bock, H. H. : Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse), Göttingen 1974.

Chandon, J.-L./S. Pinson : Analyse typologique. Théories et applications, Paris, New York, Barcelone, Milan 1981.

Dickinson, G. C. : Statistical mapping and the presentation of statistics, Londres 1973².

Gardette, J. : Deux itinéraires des invasions linguistiques dans le domaine provençal, in : Revue de linguistique romane 19 (1955), 181-196.

Gendre, F. : L'analyse statistique univariée. Introduction à son utilisation pratique, Genève 1977.

Goebl, H. : La dialectométrie appliquée à l'ALF (Normandie), in : XIV Congresso internazionale di linguistica e filologia romanza. Atti, ed. A. Vàrvaro, Naples, Amsterdam 1976, vol. II, 165-195.

Goebl, H. : Eléments d'analyse dialectométrique (avec application à l'AIS), in : Revue de linguistique romane 45 (1981), 349-420.

Goebl, H. : Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie, in : Denkschriften der Österreichischen Akademie der Wissenschaften (Mémoires de l'Académie des Sciences de Vienne), phil.-hist. Klasse, vol. 157, Vienne 1982, 1-123.

Goebl, H. : Parquet polygonal et treillis triangulaire. Les deux versants de la dialectométrie interponctuelle, in : Revue de linguistique romane 47 (1983a), 353-412.

Goebl, H. : « Stammbaum » und « Welle ». Vergleichende Betrachtungen aus numerisch-taxonomischer Sicht, in : Zeitschrift für Sprachwissenschaft 2 (1983b), 3-44.

Goebl, H. : Messen, Matrizen und Mustererkennung. Methodisch-methodologische Bemerkungen zur Verwendung der numerischen Klassifikation in der Geolinguistik, in : Folia linguistica 17 (1983c), 403-444.

Goebl, H. : Problemi e metodi della classificazione geolinguistica, in : Linguistica e dialettologia veneta. Studi offerti a M. Cortelazzo dai colleghi stranieri, ed. G. Holtus/M. Metzeltin, Tubingue 1983d, 193-203.

Goebl, H. : Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF, Tubingue 1984, 3 vol. (1984 I, 1984 II, 1984 III).

Guiter, H. : Français central et dialectes du Nord-Ouest selon l'Atlas linguistique de la France, in : Bulletin philologique et historique, année 1976 (Paris 1978), 55-66.

Guiter, H. : Confrontation de méthodes géolinguistiques en domaine normand, in : Actes du 105e Congrès national des Sociétés savantes (Caen 1980), Section de philologie et d'histoire jusqu'à 1610, Paris 1984, tome II, 19-28.

Jenks, G. F./F. C. Caspall : Error on choroplethic maps : definition, measurement, reduction, in : Annals of the Association of American Geographers 61 (1971), 217-244.

Jenks, G. F./M. R. C. Coulson : Class intervalls for statistical maps, in : Internationales Jahrbuch für Kartographie 3 (1963), 119-134.

Joly, F. : La cartographie, Paris 1976.

Joret, Ch. : Des caractères et de l'extension du patois normand, Paris 1883.

Kroeber, A. L. : On typological indices I : Ranking of languages, in : International Journal of American Linguistics 26 (1960), 171-177.

Kuen, H. : Die Stellung des Katalanischen in der romanischen Sprachfamilie, in : Studia Iberica. Festschrift für Hans Flasche, ed. K.-H. Körner/K. Rühl, Berne, Munich 1973, 331-352.

Niemann, H. : Klassifikation von Mustern, Berlin, Heidelberg, New York 1983.

Scripter, M. W. : Nested-means map classes for statistical maps, in : Annals of the Association of American Geographers 60 (1970), 385-393.

Sedlacek, P. ed. : Regionalisierungsverfahren (Wege der Forschung, vol. 195), Darmstadt 1978.

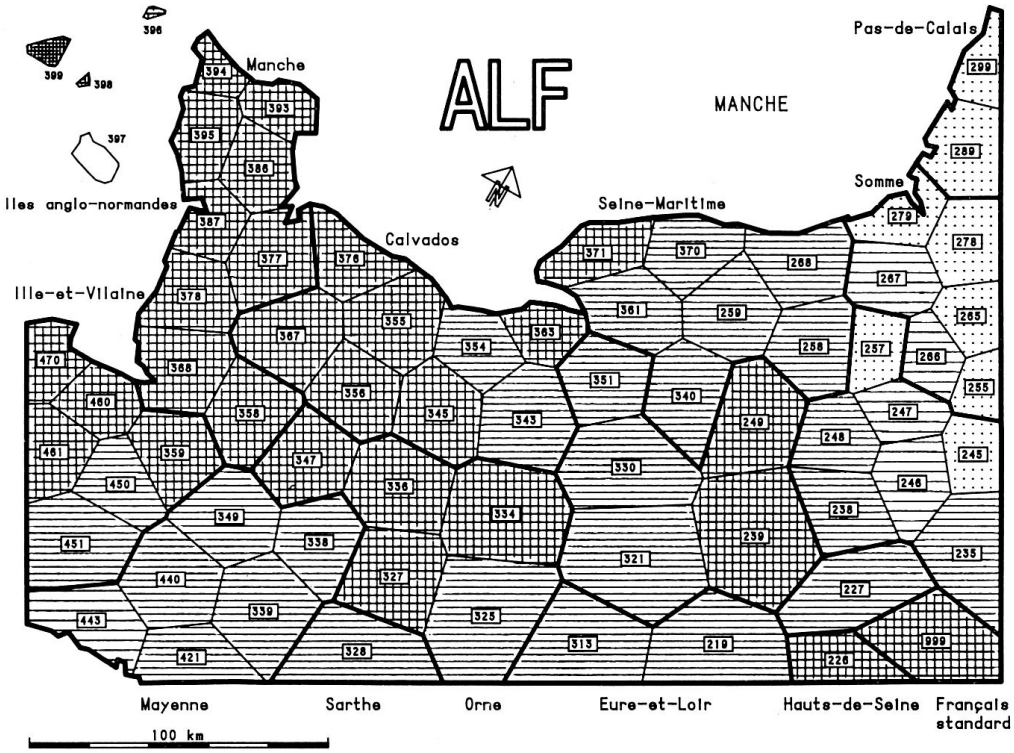
Sneath, P. H. A./R. R. Sokal : Numerical taxonomy. The principles and practice of numerical taxonomy, San Francisco 1973.

Ujomov, A. I. : Dinge, Eigenschaften und Relationen (traduit du russe), Berlin 1965.

Wonka, E. : Methoden der Wertstufenbildung und ihre Eignung für die thematische Kartographie, in : Mitteilungen der österreichischen geographischen Gesellschaft 122 (1980) 70-96.

Salzbourg. Hans GOEBL (20)

(20) Avec le concours de Siegfried Selberherr (calcul électronique), Wolf-Dieter Rase (cartographie automatique) et Hilmar Pudlatz (pavage de Thiessen).



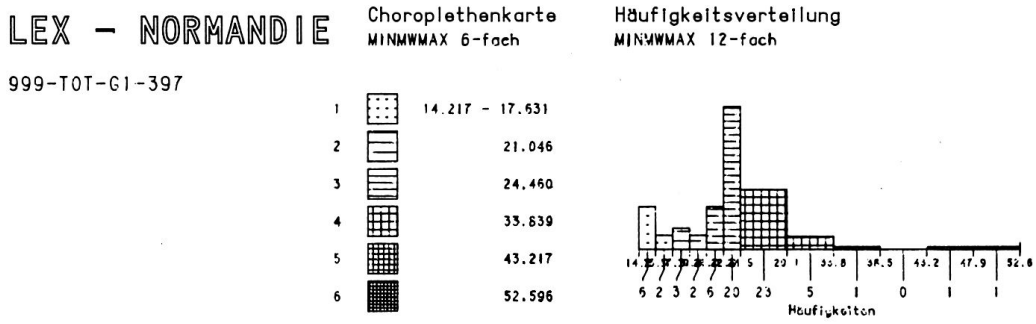
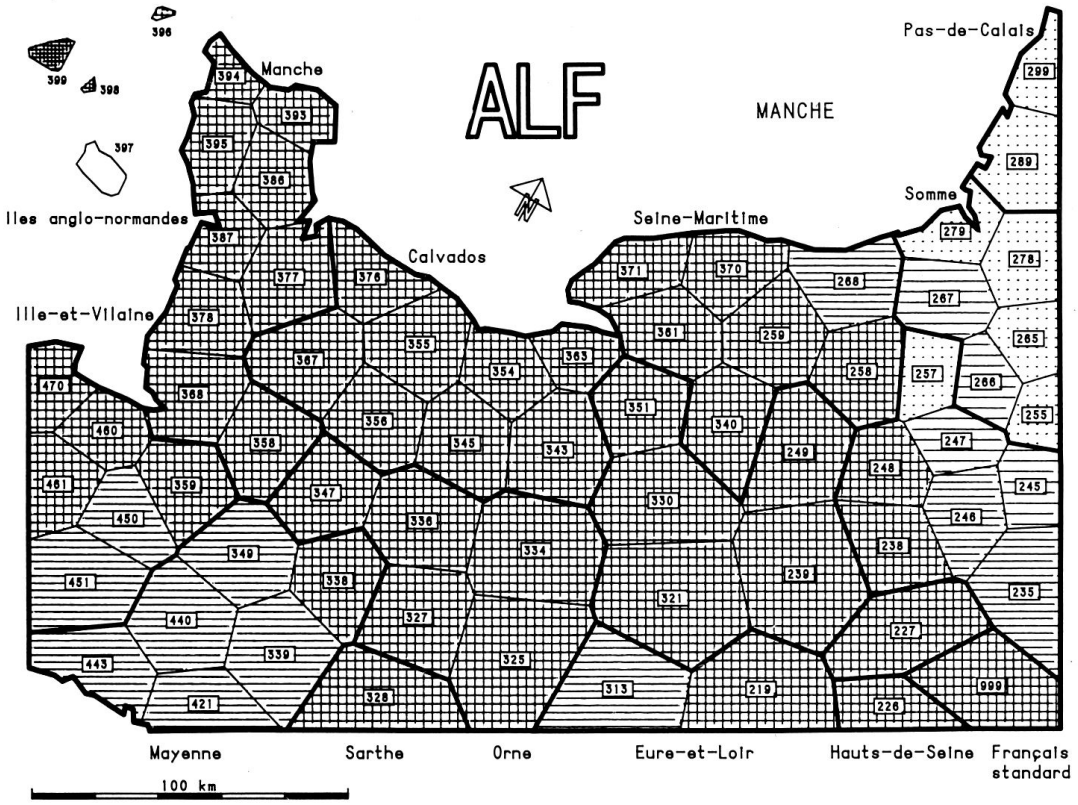


Fig. 14 :

Carte choroplèthe de la distribution de similarité (selon $IPI(1)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 28.



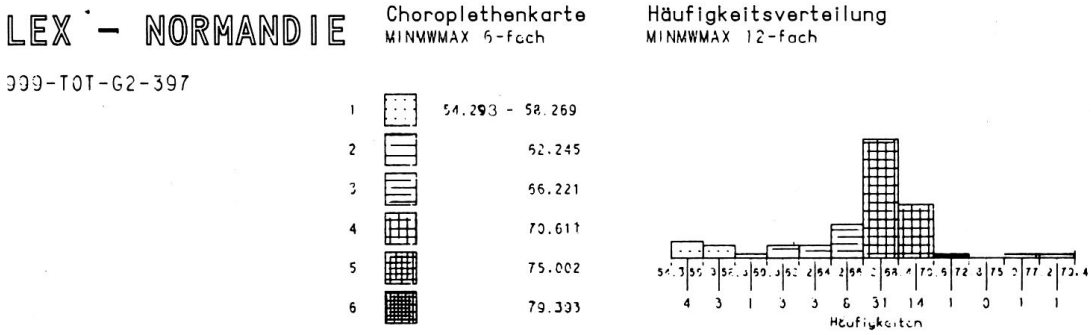
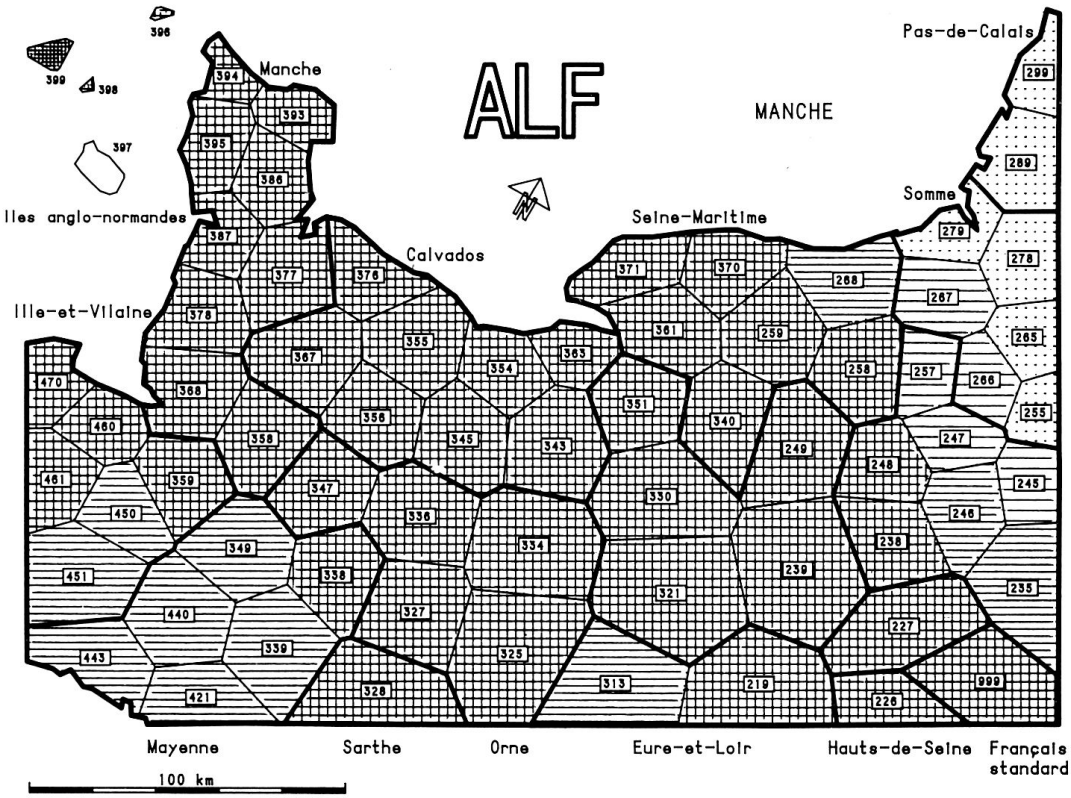


Fig. 15 :

Carte choroplèthe de la distribution de similarité (selon $IPI(2)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 29.



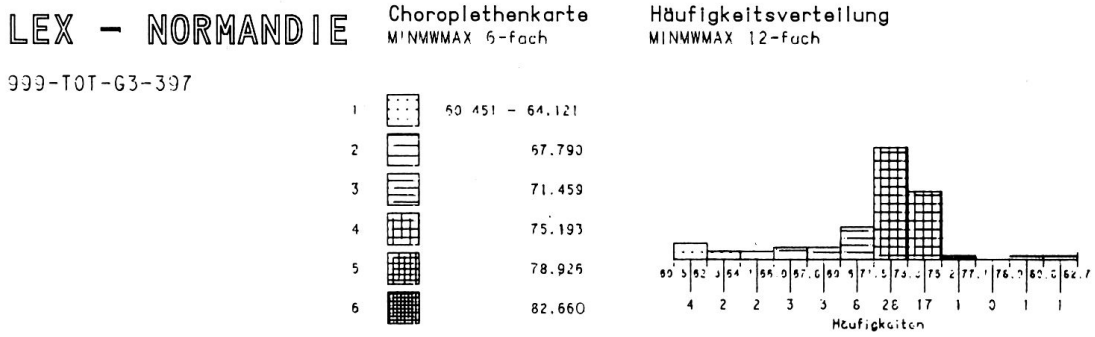


Fig. 16 :

Carte choroplèthe de la distribution de similarité (selon $IPI(3)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 30.

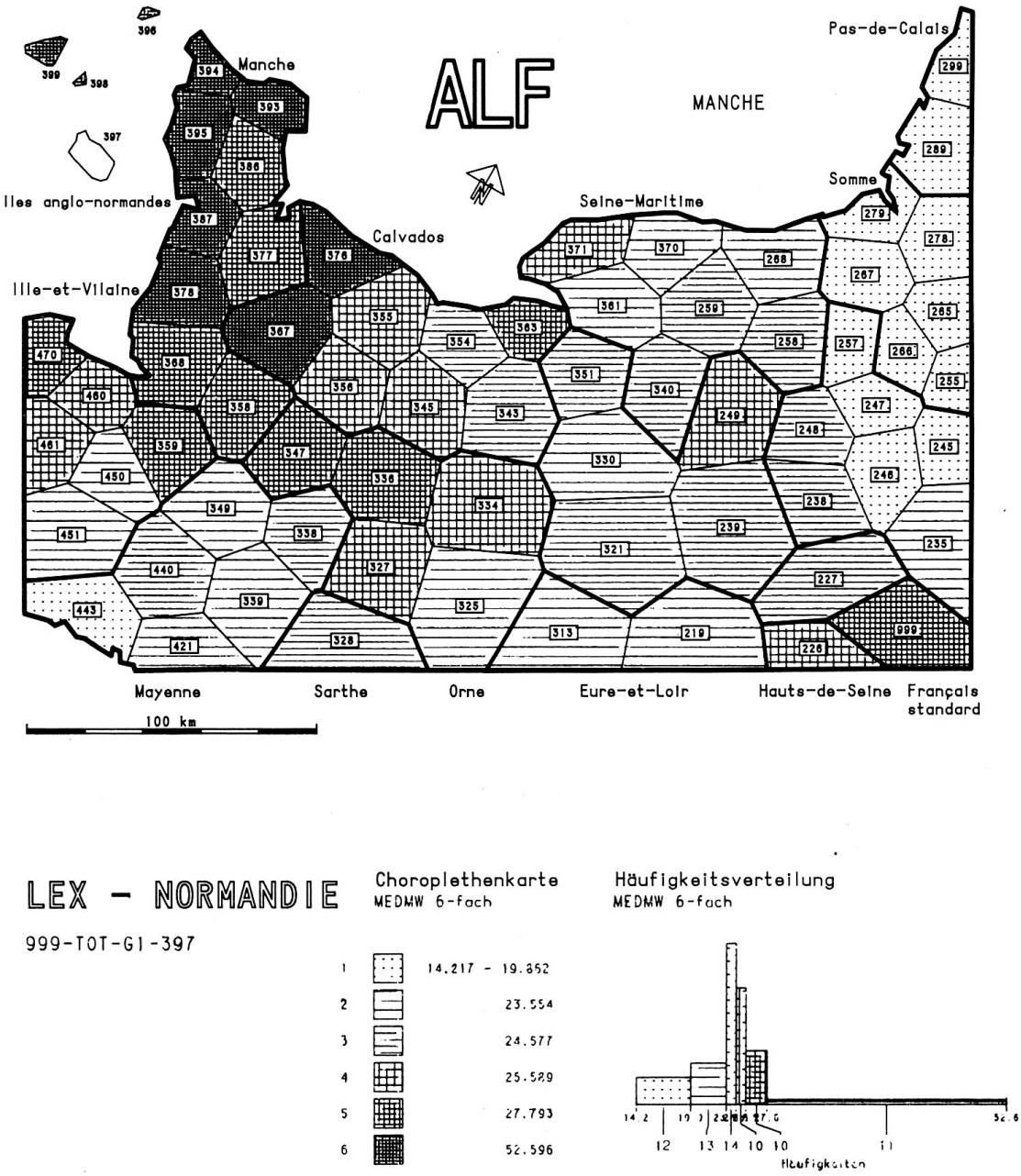
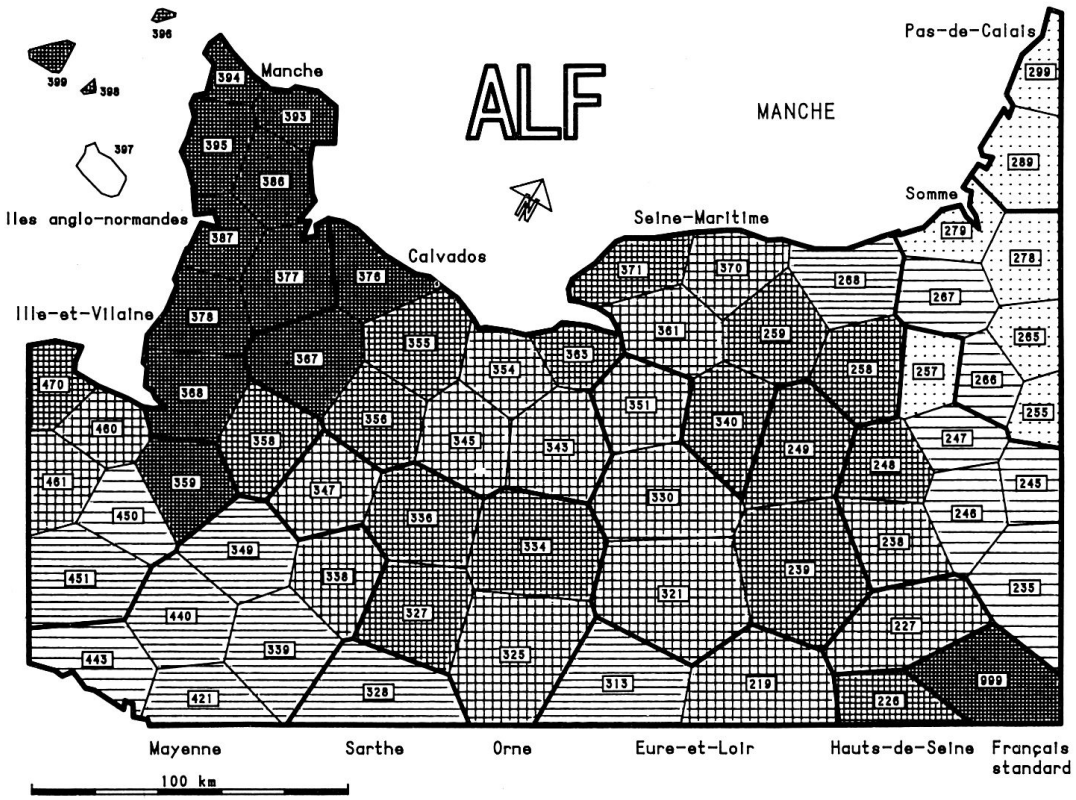


Fig. 17 :

Carte choroplèthe de la distribution de similarité (selon $IPI(1)_{jk}$) relative au point de référence 397 (La Trinité, Jersey). Algorithme d'intervallisation : MEDMW à 6 intervalles. Voir aussi la Fig. 28.



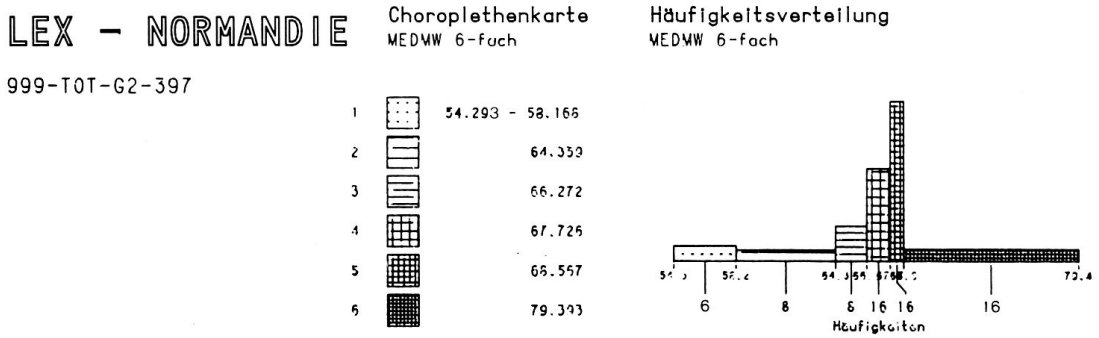
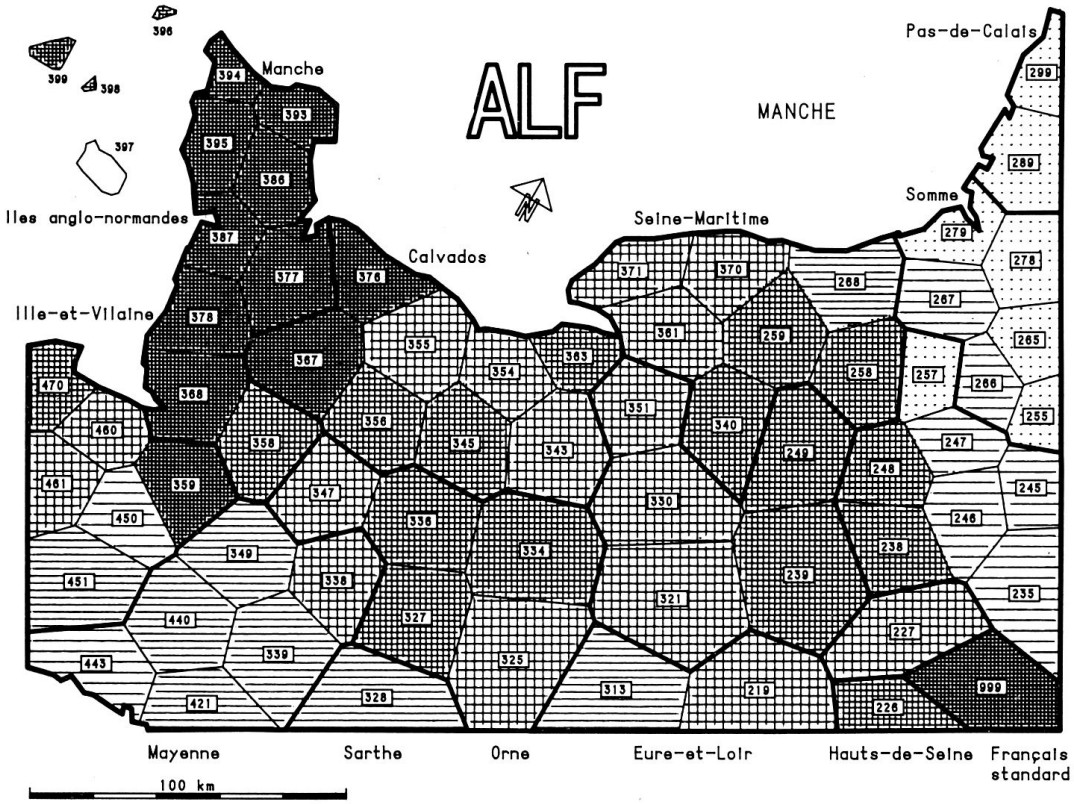


Fig. 18 :

Carte choroplèthe de la distribution de similarité (selon $IPI(2)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MEDMW à 6 intervalles.
Voir aussi la Fig. 29.



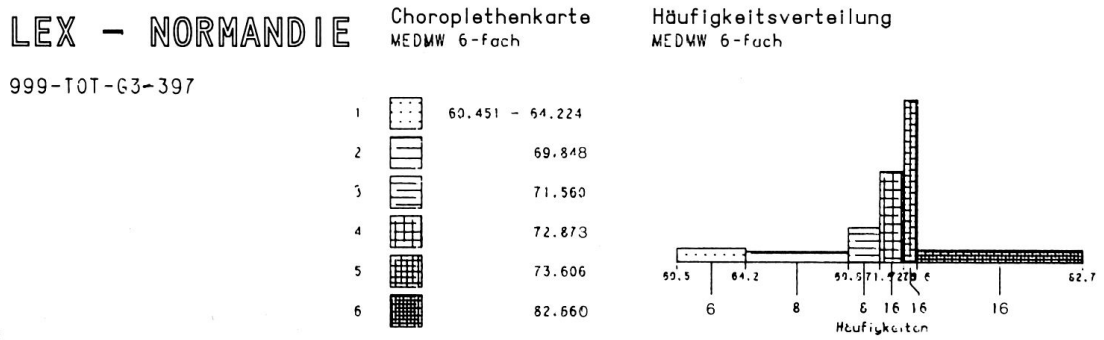
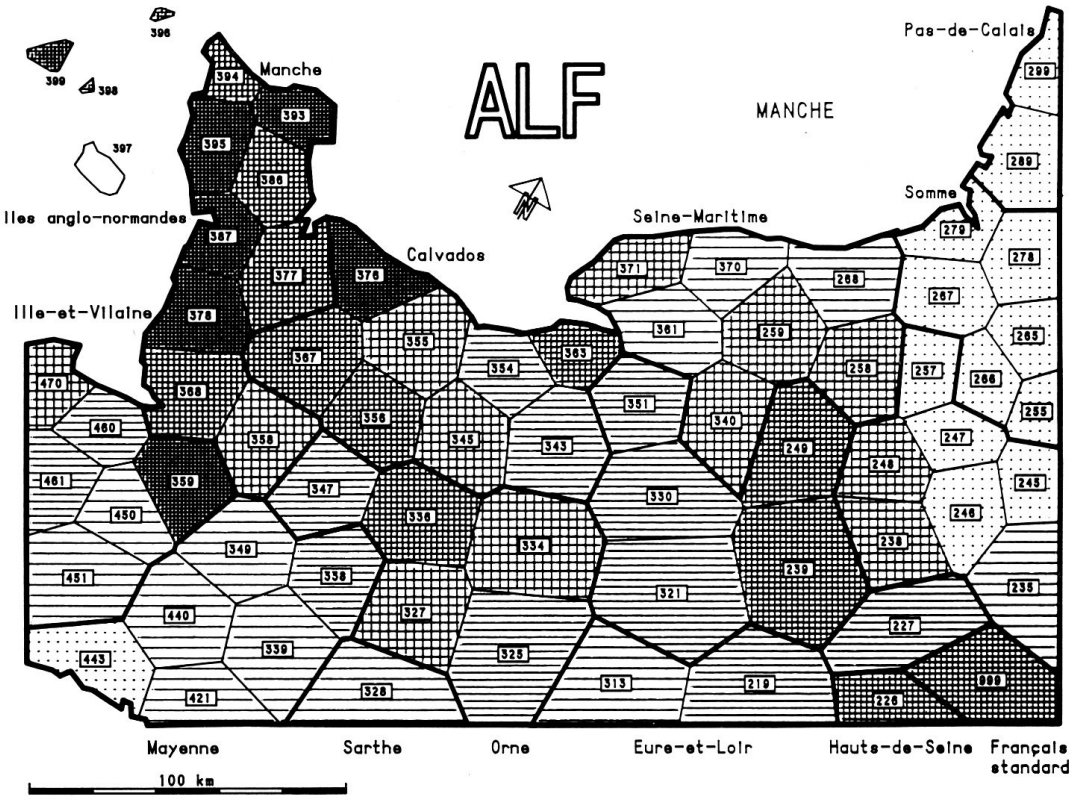


Fig. 19 :

Carte choroplèthe de la distribution de similarité (selon $IPI(3)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MEDMW à 6 intervalles.
Voir aussi la Fig. 30.



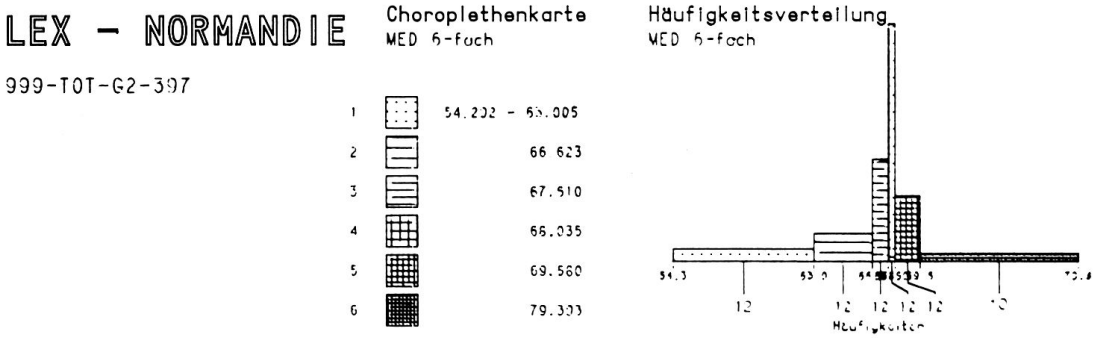
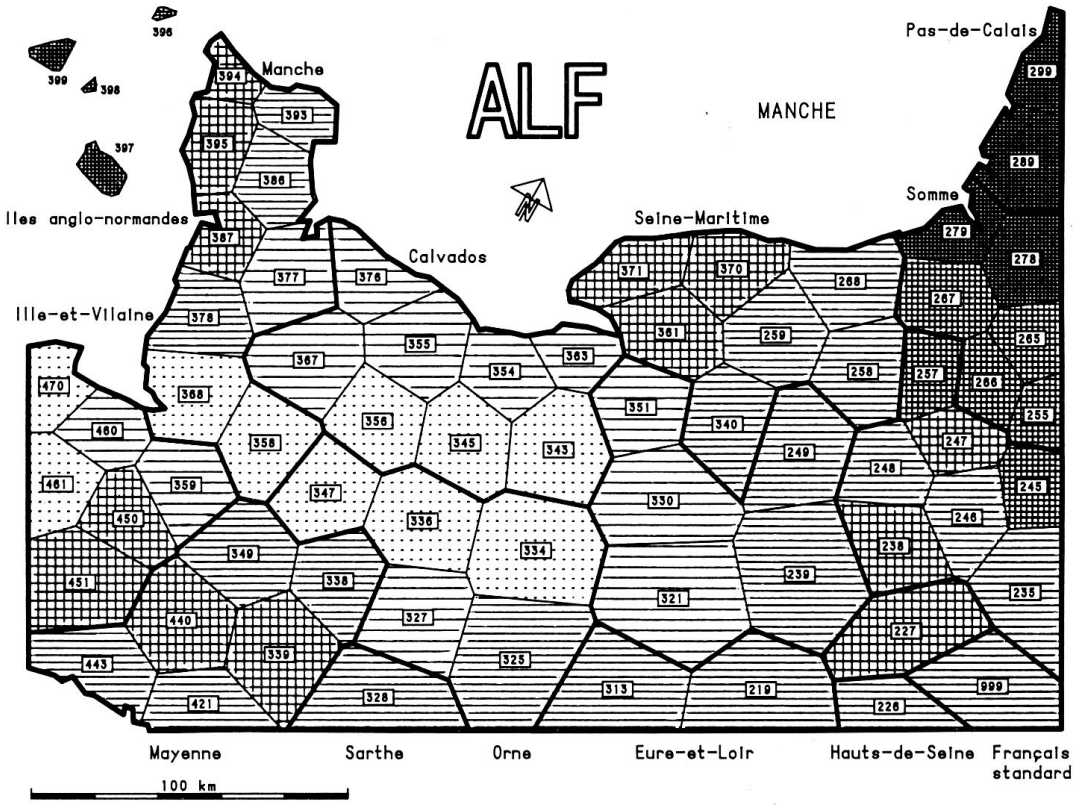


Fig. 20 :

Carte choroplèthe de la distribution de similarité (selon $IPI(2)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).
Algorithme d'intervallisation : MED à 6 intervalles.
Voir aussi la Fig. 29.



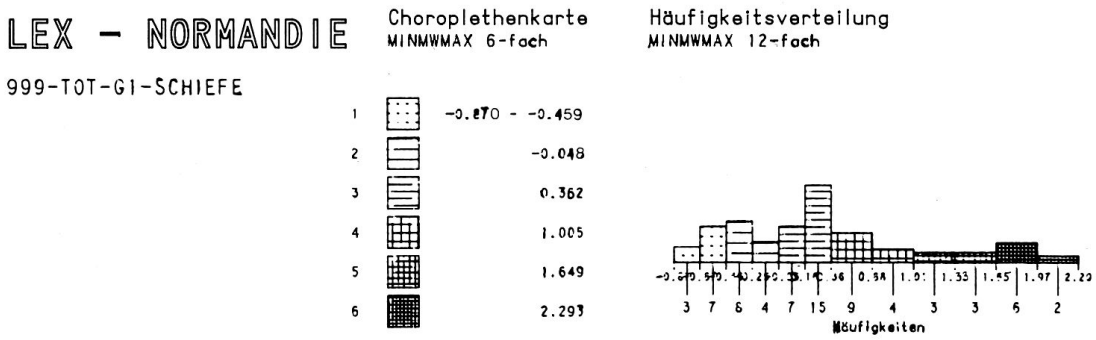
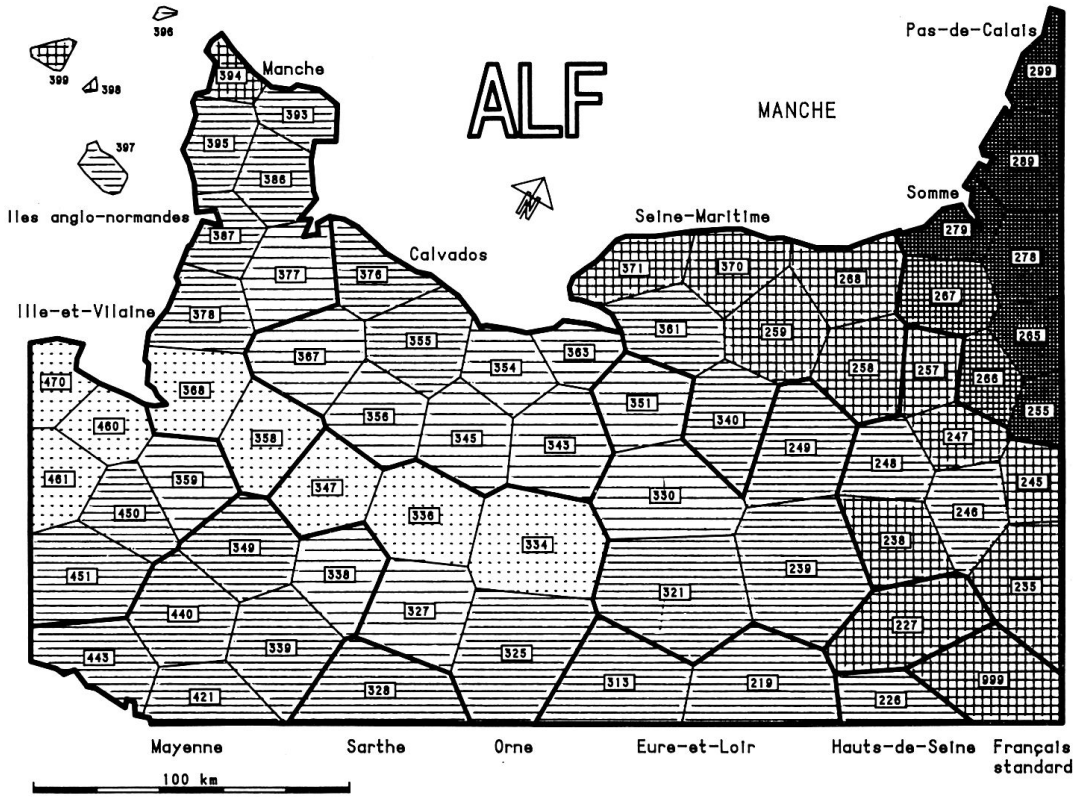


Fig. 21 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(1)_{jk}$).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 31.



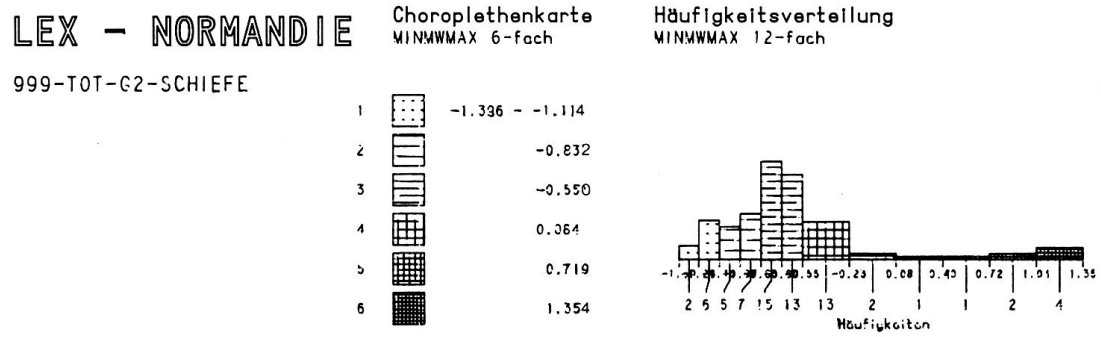


Fig. 22 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(2)_{jk}$).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 32.

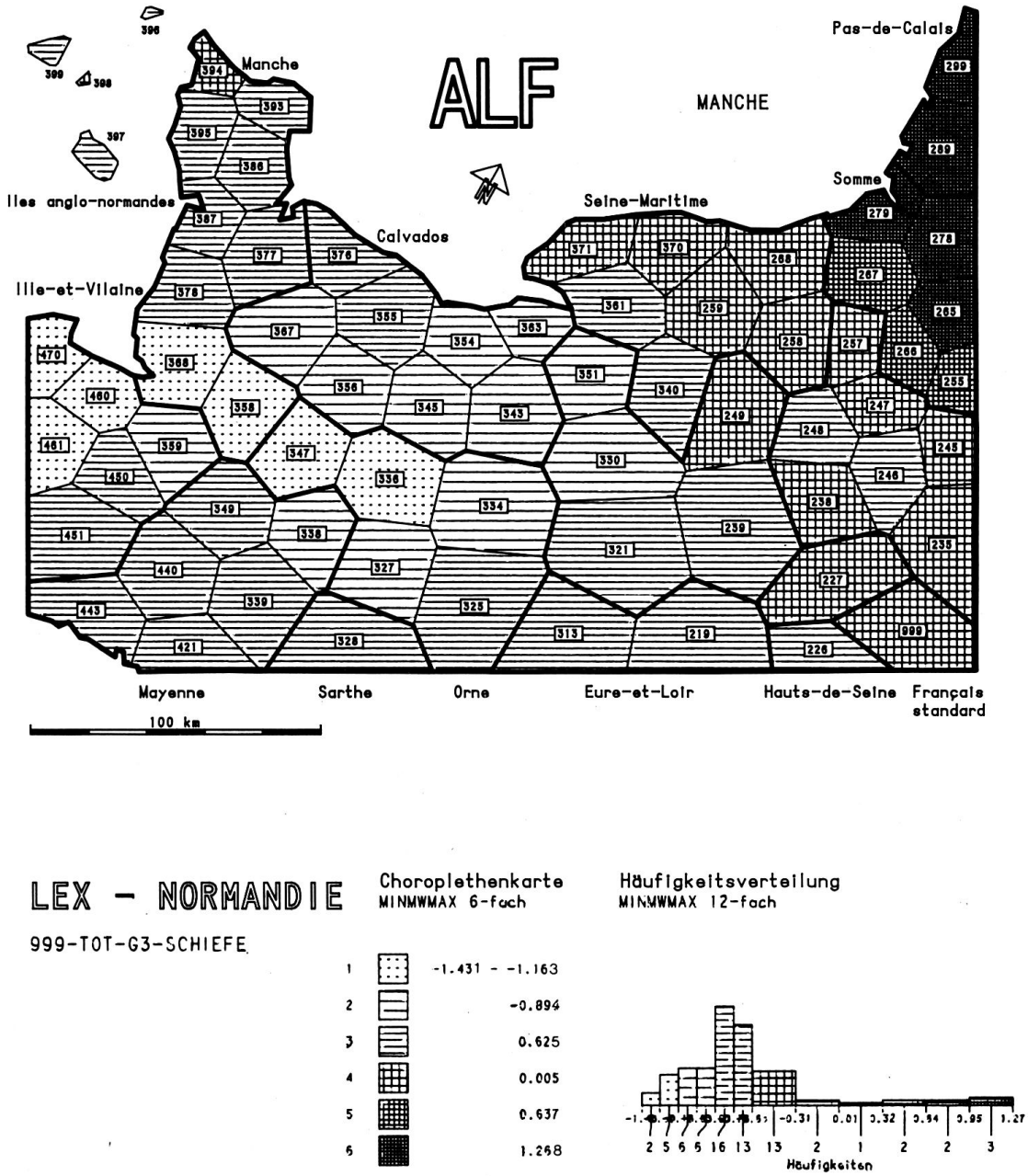
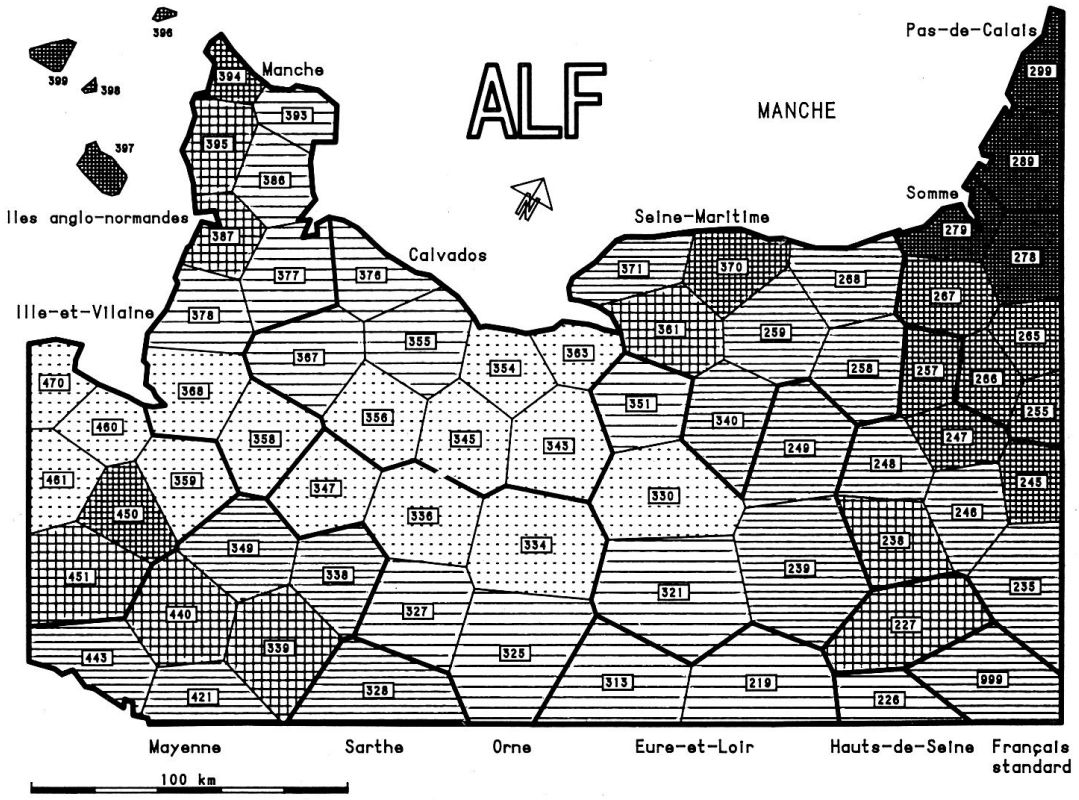


Fig. 23 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(3)_{jk}$).
Algorithme d'intervallisation : MINMWMAX à 6 intervalles.
Voir aussi la Fig. 33.



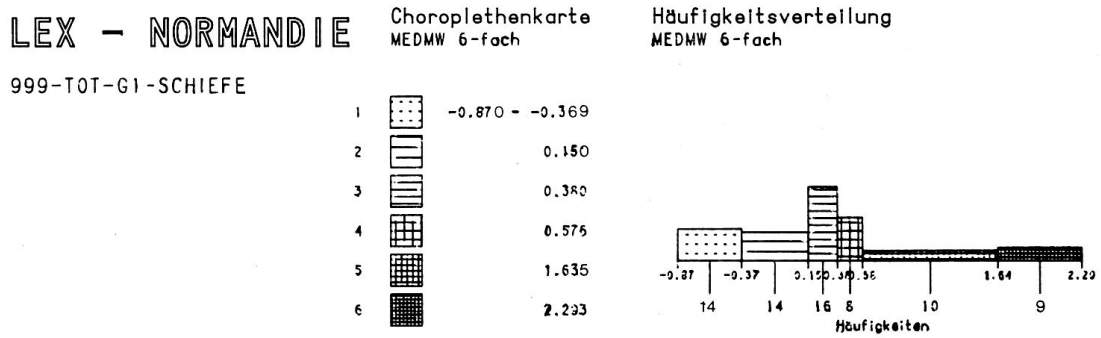
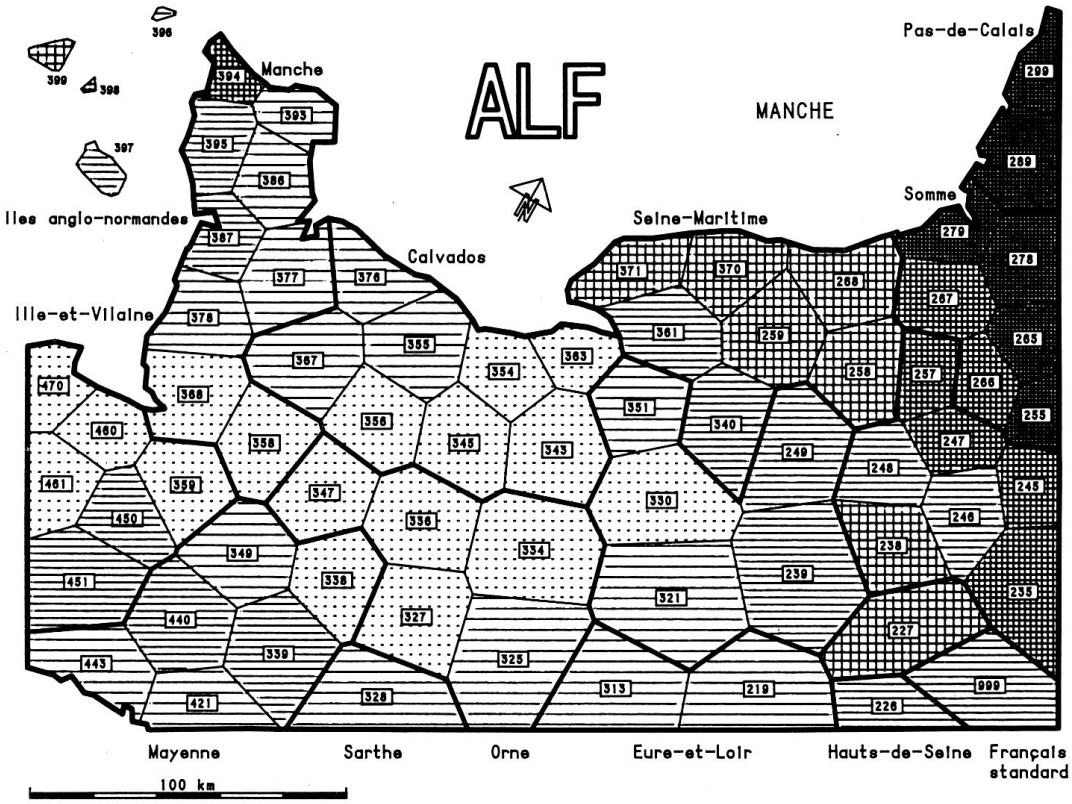


Fig. 24 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(1)_{jk}$).
Algorithme d'intervallisation : MEDMW à 6 intervalles.
Voir aussi la Fig. 31.



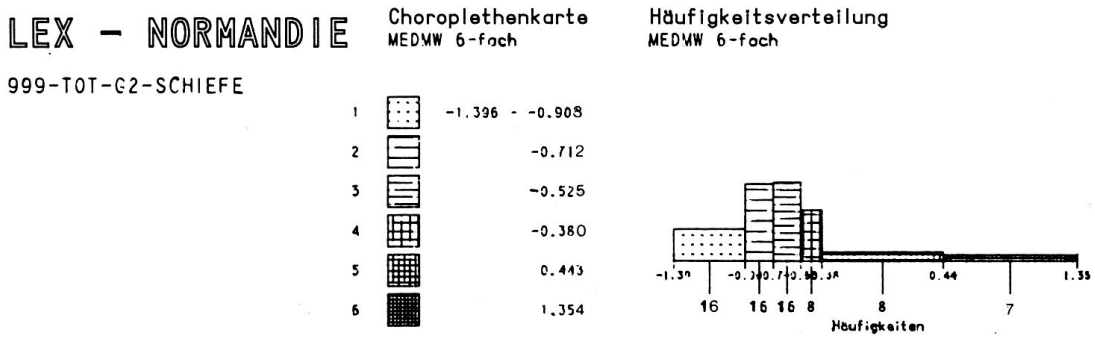
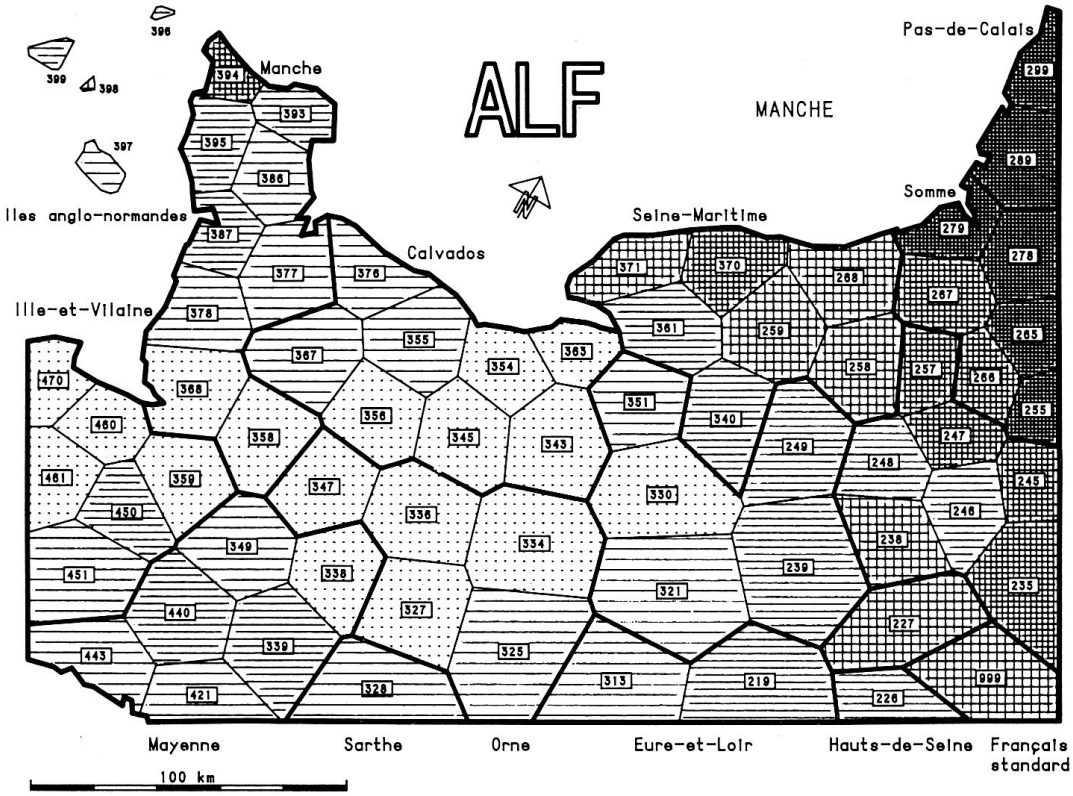


Fig. 25 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(2)_{jk}$).
Algorithme d'intervallisation : MEDMW à 6 intervalles.
Voir aussi la Fig. 32.



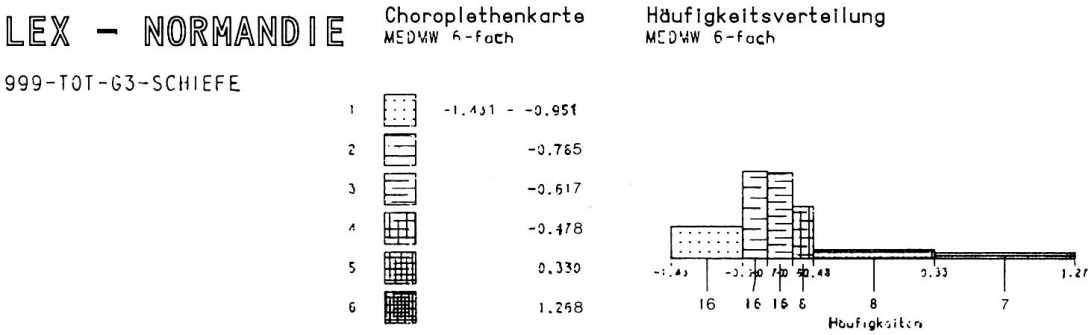


Fig. 26 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(3)_{jk}$).
Algorithme d'intervallisation : MEDMW à 6 intervalles.
Voir aussi la Fig. 33.

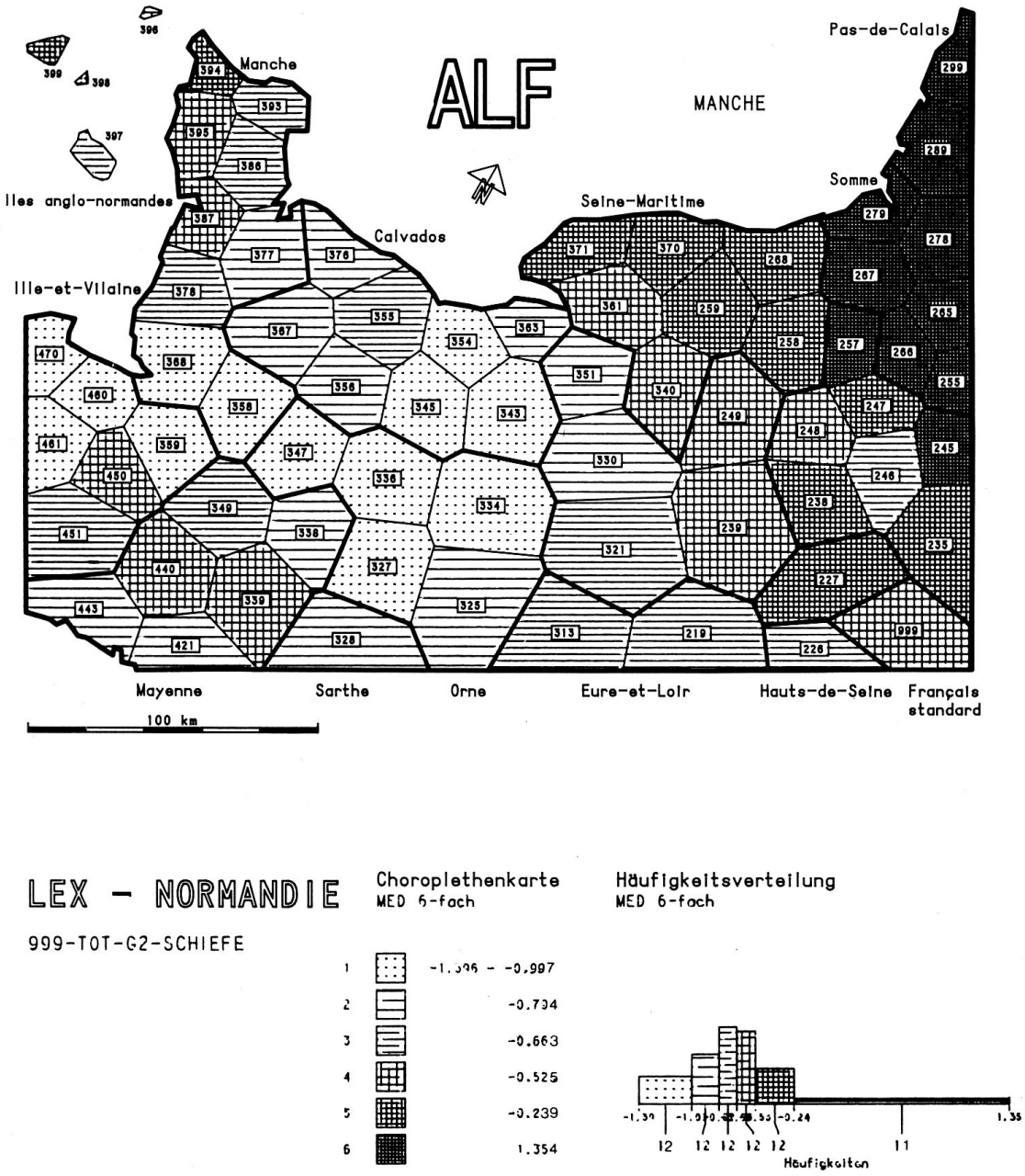


Fig. 27 :

Carte choroplèthe de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon $IPI(2)_{jk}$).
Algorithme d'intervallisation : MED à 6 intervalles.
Voir aussi la Fig. 32.

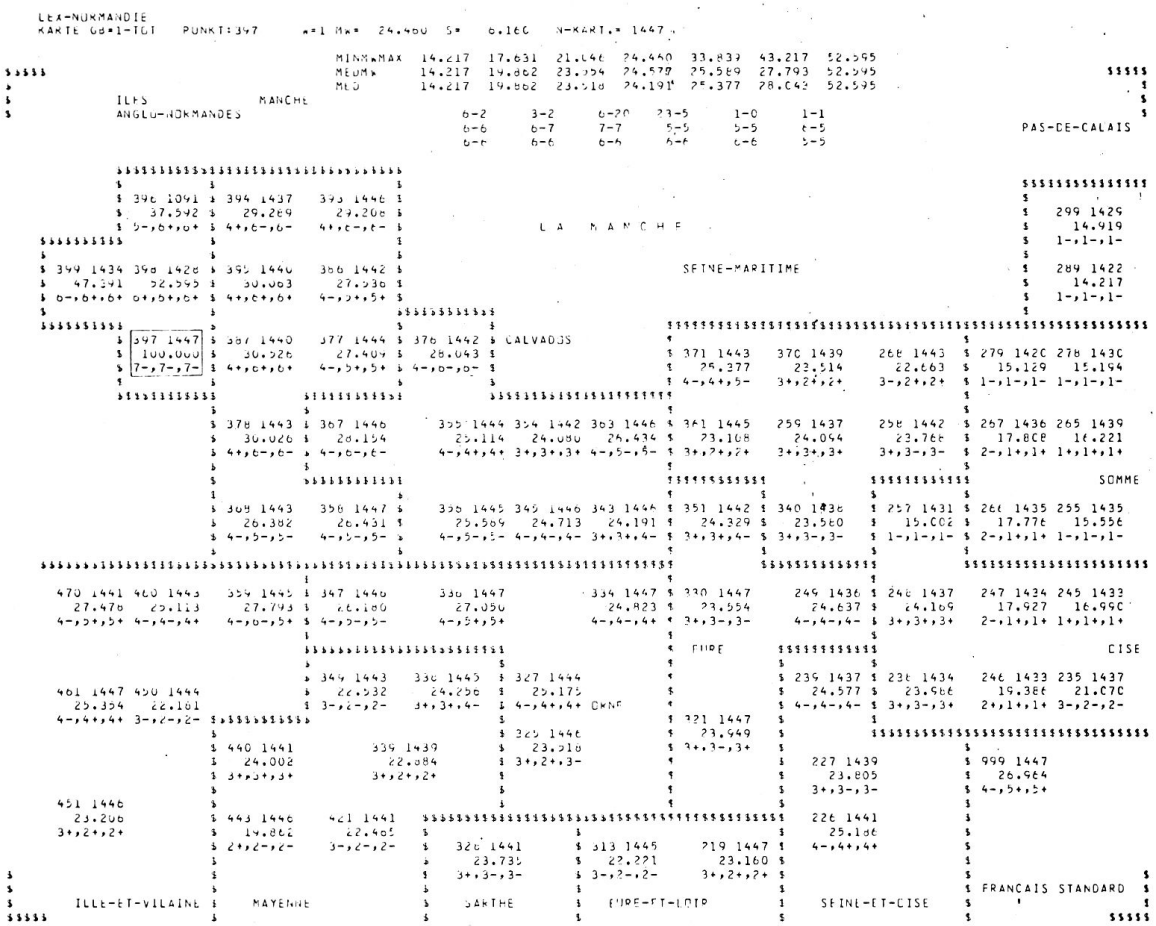


Fig. 28 :

Carte numérique de la distribution de similarité (selon IPI(1)$_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

Voir aussi les Fig. 14 et 17.

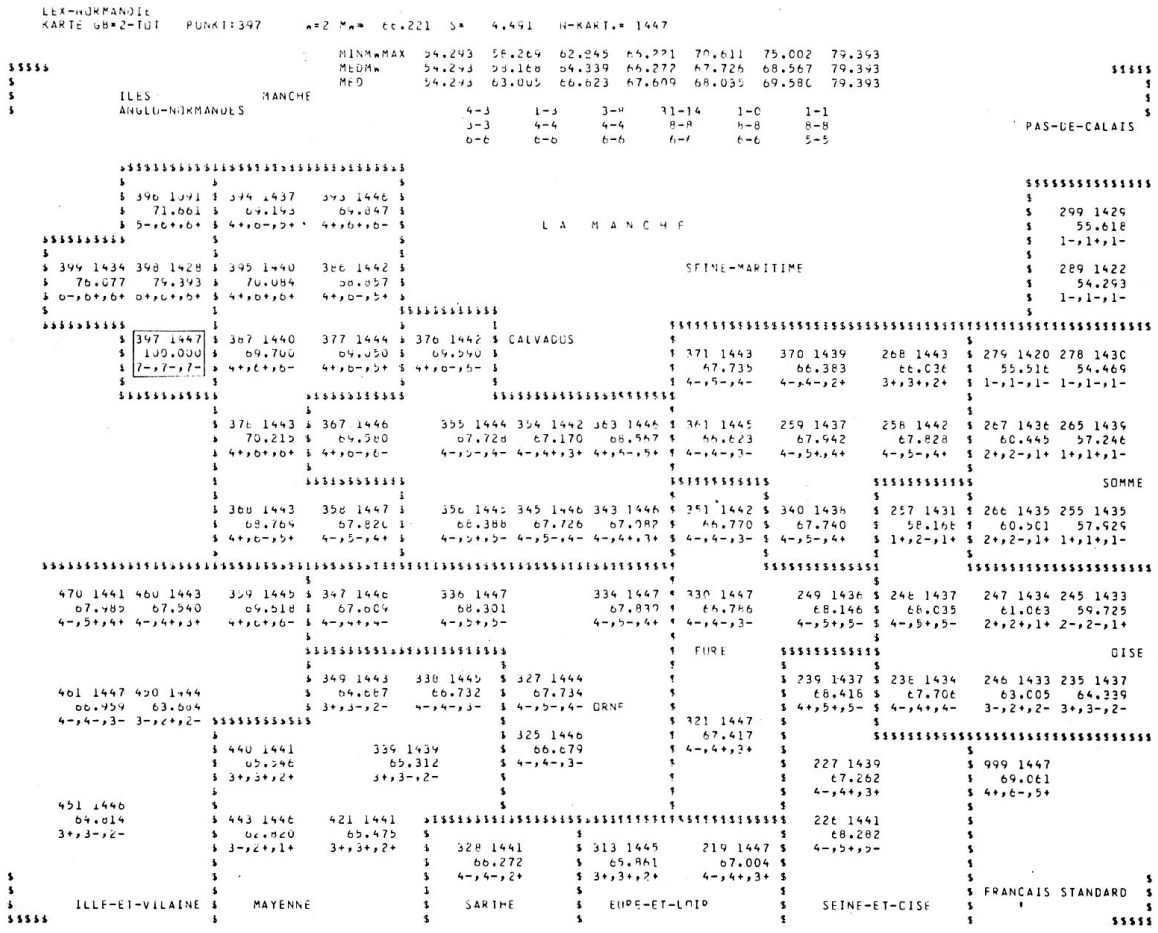


Fig. 29 :

Carte numérique de la distribution de similarité (selon $IPI(2)_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

Voir aussi les Fig. 15, 18 et 20.

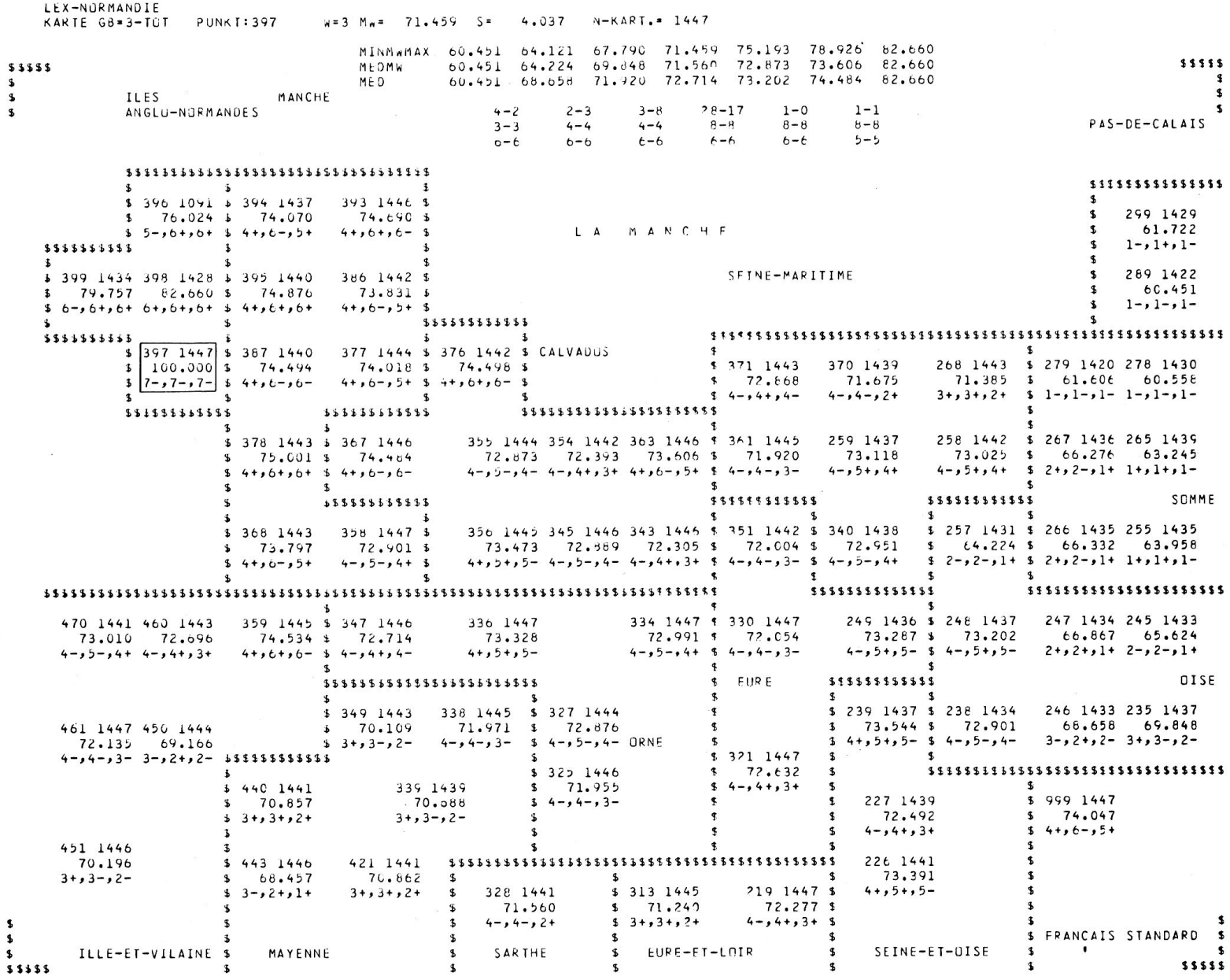


Fig. 30 :

Carte numérique de la distribution de similarité (selon IPI(3)$_{jk}$) relative au point de référence 397 (La Trinité, Jersey).

Voir aussi les Fig. 16 et 19.

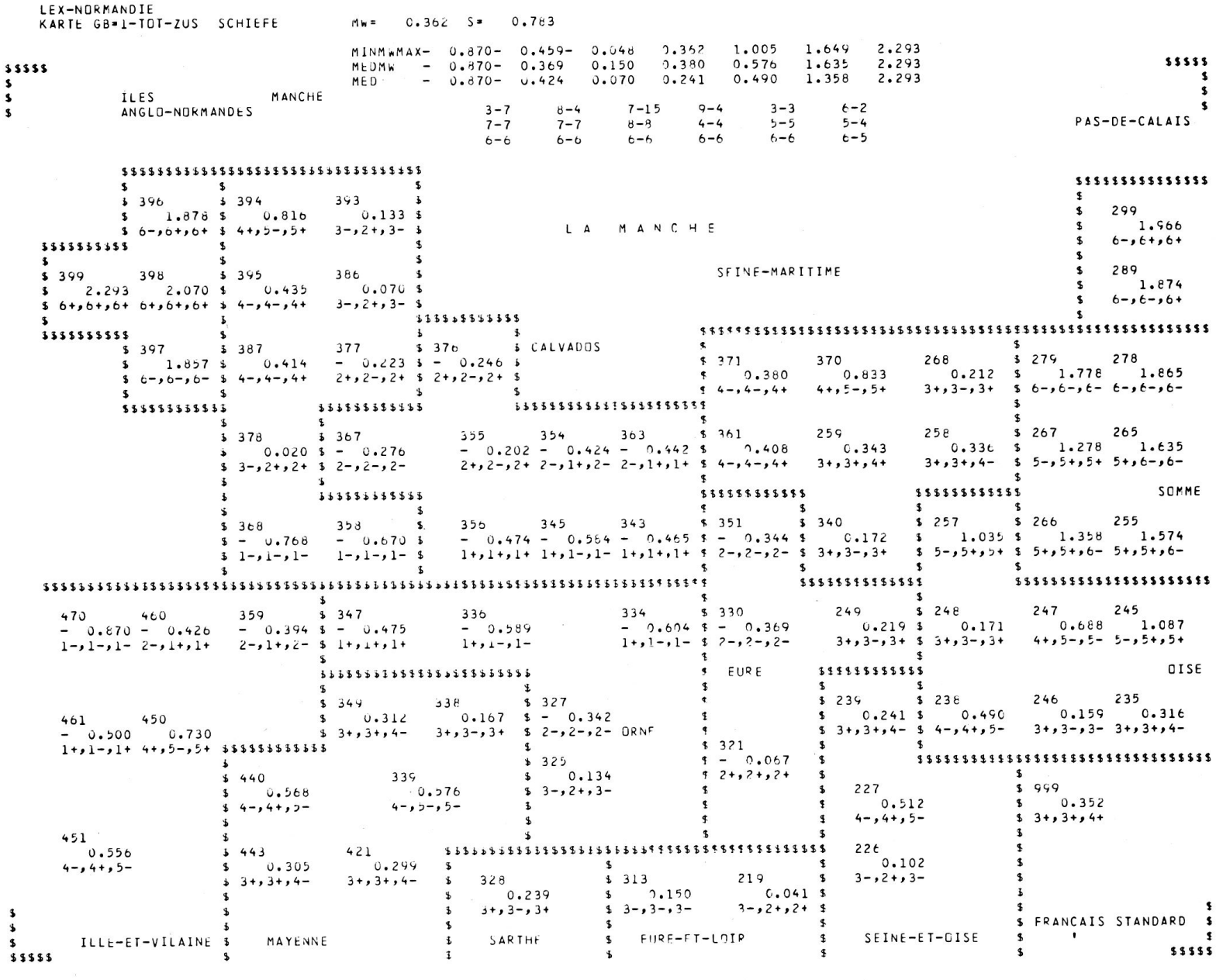


Fig. 31 : Carte numérique de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon IPI(1)$_{jk}$).

Voir aussi les Fig. 21 et 24.

LEX-NORMANDIE
KARTE GB=2-TOT-ZUS SCHIEFE Mw=- 0.550 S= 0.592

MINMWMAX- 1.396- 1.114- 0.832- 0.550 0.084 0.719 1.354
MEDMW - 1.396- 0.903- 0.712- 0.525- 0.360 0.443 1.354
MED - 1.396- 0.997- 0.794- 0.663- 0.525- 0.239 1.354

ILES ANGLO-NORMANDES — MANCHE

2-6 5-7 15-13 13-2 1-1 2-4
8-8 8-8 8-8 4-4 4-4 4-3
6-6 6-6 6-6 6-6 6-6 6-5

PAS-DE-CALAIS

396 - 0.582 3+,3+,4+
394 - 0.377 4-,5-,5+
393 - 0.715 3-,2+,3+

L A M A N C H E

299 1.083 6+,6+,6+

399 - 0.434 4-,4+,5-
398 - 0.324 4-,5-,5+
395 - 0.634 3+,3-,4-
386 - 0.766 3-,2+,3+

SEINE-MARITIME

289 1.040 6+,6-,6+

397 - 0.707 3-,3-,3+
387 - 0.602 3+,3+,4+
377 - 0.837 2+,2-,2+
376 - 0.795 3-,2-,2+

CALVADOS

371 - 0.515 4-,4-,5-
370 - 0.380 4-,5-,5+
268 - 0.492 4-,4-,5-
279 1.085 6+,6+,6+
278 1.354 6+,6+,6+

376 - 0.775 3-,2+,3-
367 - 0.896 2+,2-,2-
355 - 0.794 3-,2+,3-
354 - 1.003 2-,1+,1+
363 - 0.951 2+,1+,2-

361 - 0.592 3+,3+,4+
259 - 0.496 4-,4-,5-
258 - 0.466 4-,4+,5-
267 0.351 5-,5+,6-
265 0.940 6-,6-,6+

SOMME

368 - 1.212 1+,1-,1-
358 - 1.198 1+,1-,1-
356 - 0.943 2+,1+,2-
345 - 0.997 2-,1+,2-
343 - 0.999 2-,1+,1+

351 - 0.872 2+,2-,2+
340 - 0.589 3+,3+,4+
257 - 0.053 4+,5+,6-
266 0.443 5+,6-,6-
255 0.750 6-,6-,6-

470 - 1.396 1-,1-,1-
460 - 1.131 1+,1-,1+
359 - 1.099 2-,1+,1+
347 - 1.194 1+,1-,1-
336 - 1.190 1+,1-,1-
334 - 1.120 1+,1-,1+

330 - 0.906 2+,1+,2-
249 - 0.554 3+,3+,4+
248 - 0.632 3+,3-,4-
247 - 0.239 4-,5+,6-
245 0.050 4+,5+,6-

FURE

OISE

461 - 1.260 1-,1-,1-
450 - 0.589 3+,3+,4+
349 - 0.788 3-,2+,3-
338 - 0.903 2+,2-,2-
327 - 1.033 2-,1+,1+
ORNE
239 - 0.615 3+,3-,4-
238 - 0.429 4-,4+,5+
246 0.663 3+,3-,4-
235 - 0.379 4-,5-,5+

321 - 0.775 3-,2+,3-

440 - 0.614 3+,3+,4-
339 - 0.609 3+,3+,4-
325 - 0.808 3-,2-,2+

227 - 0.403 4-,4+,5+
999 - 0.525 4-,4-,5-

451 - 0.695 3-,3-,3+
443 - 0.797 3-,2-,2+
421 - 0.794 3-,2+,3-
328 - 0.818 3-,2-,2+
313 - 0.794 3-,2+,3-
219 - 0.712 3-,3-,3+
226 - 0.684 3+,3-,3+

ILLE-ET-VILAINE — MAYENNE — SARTHE — EURE-ET-LOIR — SEINE-ET-OISE — FRANCAIS STANDARD

Fig 32.

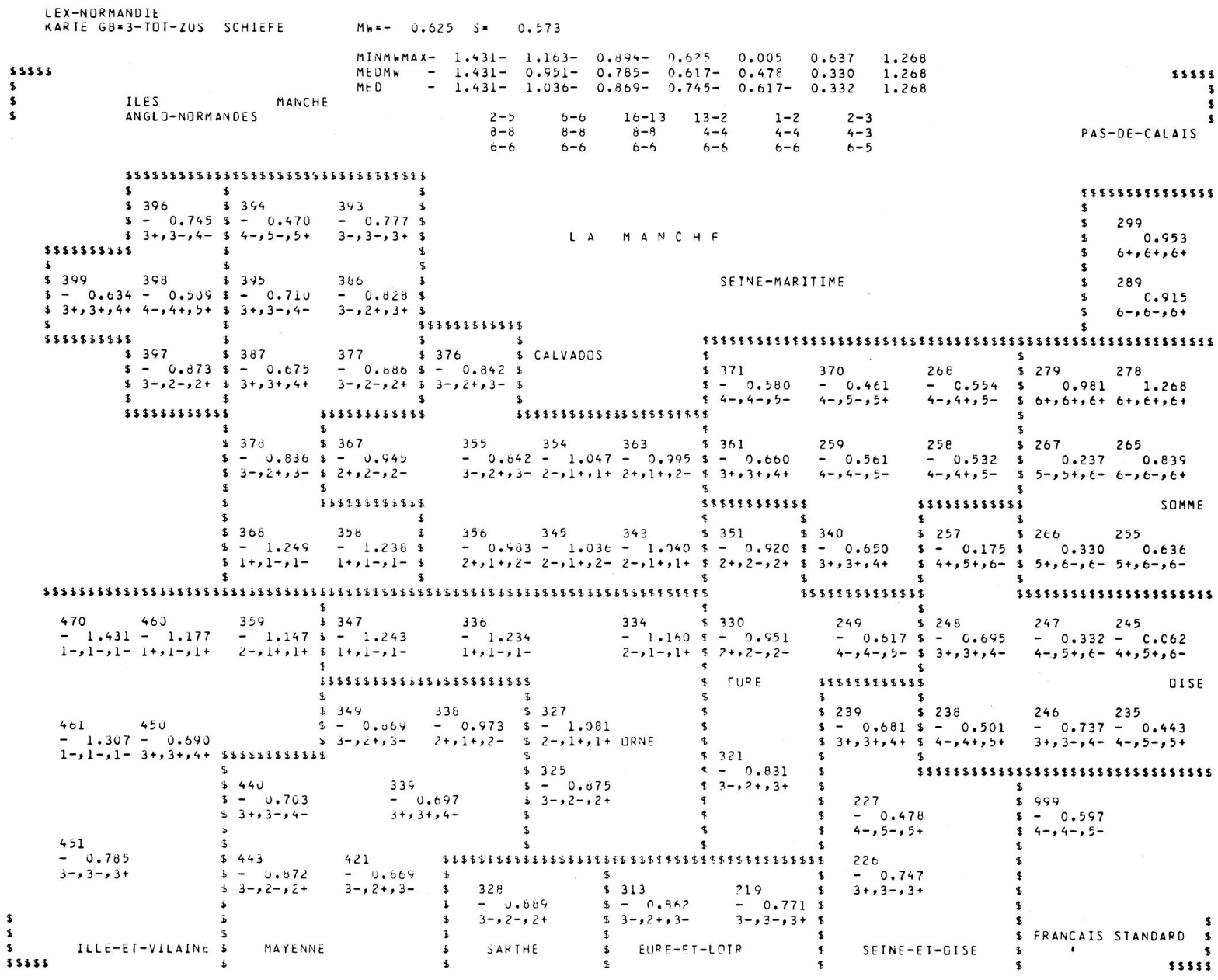


Fig. 33 :

Carte numérique de la synopse des coefficients d'asymétrie (de Fisher) de 71 distributions de similarité (selon IPI(3)$_{jk}$).