

**Stéphanie LEDAUPHIN**

e-mail : [stephanie.ledauphin@uhb.fr](mailto:stephanie.ledauphin@uhb.fr)

e-mail :

<b>DOSSIER DE CANDIDATURE AU POSTE DE MAÎTRE DE CONFERENCE</b>
--------------------------------------------------------------------

**POSTE MCF 26**

**Mots clefs :** analyse de données fonctionnelles, alignement, B-splines, temps-intensité, multi-tableaux, performance, tests de permutations, chaînes de Markov, chaînes de Markov cachées.

---

<b>CURRICULUM VITAE.....</b>	<b>2</b>
ETAT CIVIL.....	2
SITUATION PROFESSIONNELLE.....	2
TITRES UNIVERSITAIRES ET DIPLOMES.....	2
INFORMATIQUE ET LANGUE.....	3
EXPERIENCES ET COMPETENCES.....	3
AUTRES STAGES, JOBS D'ETE ET VIE ASSOCIATIVE.....	3
<b>PUBLICATIONS.....</b>	<b>4</b>
PUBLICATIONS DANS DES REVUES INTERNATIONALES AVEC COMITÉ DE LECTURE.....	4
PUBLICATIONS DANS DES REVUES FRANCAISES AVEC COMITÉ DE LECTURE.....	4
ACTES DE CONFERENCES INTERNATIONALES AVEC COMITÉ DE LECTURE.....	4
ACTES DE CONFERENCES NATIONALES AVEC COMITÉ DE LECTURE.....	5
<b>ACTIVITÉS ADMINISTRATIVES.....</b>	<b>5</b>
<b>ACTIVITÉS D'ENSEIGNEMENT.....</b>	<b>6</b>
ATER À L'UNIVERSITÉ DE RENNES 2.....	6
VACATIONS.....	6
<b>ACTIVITÉS DE RECHERCHE.....</b>	<b>7</b>
EXPOSÉS ORAUX À DES CONFERENCES INTERNATIONALES.....	7
EXPOSÉS ORAUX À DES CONFÉRENCES NATIONALES.....	7
AUTRES CONFÉRENCES, SÉMINAIRES ET GROUPE DE TRAVAIL.....	8
POSTERS À DES CONFÉRENCES INTERNATIONALES.....	8
<b>RÉSUMÉ DES RÉSULTATS OBTENUS.....</b>	<b>9</b>
PERFORMANCE DES JUGES : ÉPREUVE DU PROFIL SENSORIEL CONVENTIONNEL.....	10
ANALYSE DE DONNÉES FONCTIONNELLES LES COURBES DE TYPE TEMPS INTENSITÉ.....	11
CHAÎNES DE MARKOV : DONNÉES SENSORIELLES AU COURS DU TEMPS.....	12
<b>PROGRAMME DE RECHERCHE.....</b>	<b>13</b>
ANALYSE DE DONNEES MARKETING, AUTRES TRAVAUX .....	13
CHAÎNES DE MARKOV CACHÉES.....	13
ANALYSE DE DONNÉES SENSORIELLES.....	13
ANALYSE DE DONNÉES FONCTIONNELLES.....	14

# CURRICULUM VITAE

## ETAT CIVIL

---

**Stéphanie LEDAUPHIN**

## SITUATION PROFESSIONNELLE

---

2007-2009     **ATER à l'Université de Rennes 2**

*Adresse Professionnelle :*

Equipe de statistique  
Université Rennes 2 Haute Bretagne  
Place du Recteur Henri Le Moal, CS 24307,  
35043 RENNES cedex

Tel : +33 (0)2-99-14-18-04 / Fax +33 (0)2-99-14-17-85 / e-mail : [stephanie.ledauphin@uhb.fr](mailto:stephanie.ledauphin@uhb.fr)

Page Web : [http://www.uhb.fr/sc\\_sociales/labstats/LEDAUPHIN/index.html](http://www.uhb.fr/sc_sociales/labstats/LEDAUPHIN/index.html)

## TITRES UNIVERSITAIRES ET DIPLOMES

---

2009           admissibilité à l'**agrégation** de mathématiques.

2008           **C.A.P.E.S.** de mathématiques.

2003-2007     **DOCTORAT en mathématiques appliquées, spécialité statistique**  
Université de NANTES, Laboratoire UMR ENITIAA<sup>1</sup>/INRA<sup>2</sup>.  
Intitulé de la thèse : « Analyse statistique d'évaluations sensorielles au cours du temps »  
Date et lieu de soutenance : le 23 mars 2007 à l'Université de NANTES  
Directeurs de thèse : Philippe Carmona et El Mostafa Qannari  
Membres du jury : Philippe Besse, Bernadette Govaerts, Georges Oppenheim,  
Denys Pommeret, Pascal Schlich  
Financement INRA et Région des Pays de La Loire.

2000-2003     Diplôme de l'école d'ingénieur de l'**ENSAI**<sup>3</sup> à RENNES.

1999-2000     **C.A.F.E.P. / C.A.P.E.S.** privé de mathématiques ;  
préparation à l'Université d'ANGERS et à l'I.U.F.M. de NANTES.

1998-1999     **Licence** de mathématiques à l'Université d'ANGERS mention Bien.

---

<sup>1</sup> Ecole Nationale d'Ingénieurs des Techniques des Industries Agricoles et Alimentaires

<sup>2</sup> Institut National de la Recherche Agronomique

<sup>3</sup> Ecole Nationale de la Statistique et de l'Analyse de l'Information

## INFORMATIQUE ET LANGUE

---

- ✓ Bureautique : Latex, Word, Excel, PowerPoint
- ✓ Logiciels Statistiques: Matlab, SAS, SPAD, R, S-Plus, SPSS
- ✓ Autres logiciels : MySQL (base de données)
- ✓ Anglais : courant (FIRST Certificate of English en 1999 et TOIC à 850 points en 2003).

## EXPERIENCES ET COMPETENCES

---

2007-2009 ATER à l'Université de Rennes 2.

Missions : **Enseignement** (2\*192h) ;  
**Communication** dans des congrès scientifiques/séminaires (nationaux et internationaux).  
Résultats et acquis : Mise en place de programmes sous R pour le traitement de données ;  
Développement de savoirs-faire pédagogiques.

2007-2009 **Colles** en classe de mathématiques spéciales à l'ICAM de Nantes.  
(Institut Catholique des Arts et Métiers)

2003-2007 Doctorat de mathématiques appliquées à l'ENITIAA à Nantes.

Missions : **Traitement statistique de données sensorielles au cours du temps** ;  
**Communications** dans des congrès scientifiques (nationaux et internationaux).  
Résultats et acquis : Elaboration d'une thèse; Mise en place de méthodes et de programmes  
pour le traitement de données ; **Publications** des résultats dans des revues nationales et  
internationales ; **Exposition de résultats** à un public spécialisé ou non (vulgarisation) ;  
**Travail en équipe** ; **Programmation** sous divers logiciels ; Développement de savoirs-faire  
pédagogiques.

2003 (7 mois) Stage de fin d'étude ENSAI à l'ENITIAA à Nantes.

Missions : Mise en place d'**indicateurs de performance** dans un panel d'experts ; **Recherche  
bibliographique** sur les courbes sensorielles.  
Résultats et acquis : Mise en place de programmes sous Matlab pour le traitement de données.

2002 (2 mois) Stage à l'Institut National de la Santé et de la Recherche Médicale à Rennes.

Missions : **Etude statistique** de l'impact des expositions professionnelles des parents sur  
l'apparition de cancers chez l'enfant.  
Acquis : Utilisation des connaissances acquises en statistiques ; **Autonomie** dans le travail de  
recherche ; **Programmation** sous SAS ; Ecriture de rapports scientifiques.

## AUTRES STAGES, JOBS D'ETE ET VIE ASSOCIATIVE

---

- 2005 Formation aux premiers secours AFPS
- 2001-2002 Administratrice de la Junior Entreprise de l'école (encadrement d'études statistiques,  
mise en place de rendez vous avec un client potentiel...).
- 2001 Stage en entreprise : assistante ingénieur qualité dans les établissements ROY  
(filiale LAPEYRE) à Saint Piat (28). Mise à jour des temps de production et enquête  
qualité auprès des ouvriers.
- 2000-2001 Présidente de l'association gala ENSAI 2001 (organisation d'une festivité, motivation  
d'une équipe et dispatch des tâches, organisation et conduite de réunions).
- 1998-2006 Soutien Scolaire en mathématiques, physique et informatique.
- 1996-2000 Animatrice de centres aérés et de camps de vacances (B.A.F.A. : Brevet d'Aptitude  
aux Fonctions d'Animateurs en 1996)
- 1985-2006 Pratique du basket en club depuis 1985 et encadrement d'équipes de 1998 à 2004.

## PUBLICATIONS

### PUBLICATIONS DANS DES REVUES INTERNATIONALES avec comité de lecture

---

- ❑ Ledauphin S., Hanafi M. et Qannari E.M.(2004) “Simplification and signification of principal components”, *Chemometrics and Intelligent Laboratory Systems*, 74, pp. 277-281.
- ❑ Ledauphin S., Vigneau E. et Causeur D.(2005) “Functional Approach for the analysis of Time Intensity curves”, *Journal of Sensory Studies*, 20, pp 285-300.
- ❑ Ledauphin S., Hanafi M. et Qannari E.M.(2006) “Assessment of the agreement among the subjects in fixed vocabulary profiling”, *Food Quality and Preference*, 17(3-4), pp 277-280. (\*)
- ❑ Ledauphin S., Vigneau E. et Qannari E.M.(2006) “A procedure for the analysis of time intensity curves”, *Food Quality and Preference*, 17(3-4), pp 290-295. (\*)
- ❑ Le S. et Ledauphin S.(2006) “You like tomato, I like tomato: Segmentation of consumers with missing values”, *Food Quality and Preference*, 17(3-4), pp 228-233.
- ❑ Ledauphin S., Pommeret D. et Qannari E.M.(2006) “A Markovian model to assess products shelf-lives”, *Food Quality and Preference*, 17(7-8), pp 598-603.
- ❑ Ledauphin S., Pommeret D. et Qannari E.M. (2008) “Hidden Markov Model to assess products shelf-lives”, *Food Quality and Preference*, 19(2), pp 156-161. (\*)

(\*) travaux adressés sur demande en cas de convocation pour audition

### PUBLICATIONS DANS DES REVUES FRANCAISES avec comité de lecture

---

- ❑ Ledauphin S., Hanafi M. et Qannari E.M.(2006) “Détermination et validation d'un tableau compromis en profil sensoriel conventionnel”, *Revue de Statistique Appliquée*, LIV (1), pp 47-60.
- ❑ Ledauphin S., Pommeret D. et Qannari E.M.(2007) “Application des chaînes de Markov pour le suivi de la dégradation de produits alimentaires”, *Journal de la Société Française de Statistique*, 148(3), pp 87-101.

### ACTES DE CONFERENCES INTERNATIONALES avec comité de lecture

---

- ❑ Ledauphin S., Hanafi M. et Qannari E.M. (2004) “Mesure et Test de l'accord multidimensionnel entre les sujets en profil sensoriel conventionnel”. Dans *8<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques*, pp.45-52, Rennes, France.
- ❑ Vigneau E., Ledauphin S. et Causeur D. (2004) “Approche fonctionnelle de l'analyse des signaux Temps-Intensité”. Dans *8<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques*, Rennes, France.
- ❑ Ledauphin S., Vigneau E. et Qannari E.M. (2004) “A procedure for the analysis of Time Intensity curves”. Dans *7<sup>th</sup> Sensometrics*, p. 41, UC Davis, Californie, USA.
- ❑ Ledauphin S. et Qannari E.M. (2006) “A new algorithm for the alignment of time intensity curves”. Dans *COMPSTAT2006, 17<sup>th</sup> Symposium of IASC on Computational Statistics*, p. 174, Rome, Italie.
- ❑ Ledauphin S., Pommeret D. et Qannari E.M. (2008) “Modélisation de la dégradation sensorielle des aliments par des Chaînes de Markov Cachées”. Dans *10<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques*, Louvain La Neuve, Belgique.
- ❑ Ledauphin S., Lê S. (2009) “ Typologie des consommateurs et Mesure de la loyauté/fidélité”. *En cours*.

- ❑ Ledauphin S., Hanafi M. et Qannari E.M. (2003) “La simplification et la signification des composantes principales dans le cadre d'une Analyse en Composantes Principales”. Dans *Chimiométrie 2003*, Paris, France.
- ❑ Ledauphin S., Vigneau E. et Qannari E.M. (2004) “Ajustement de courbes de type temps-intensité en évaluation sensorielle”. Dans *Chimiométrie 2004*, pp. 67-70, Paris, France.
- ❑ Ledauphin S., Vigneau E. et Qannari E.M. (2005) “L'ajustement de courbes de type temps-intensité en évaluation sensorielle en présence de répétitions”. Dans *XXXVII<sup>èmes</sup> journées de Statistiques*, Pau, France.

<b>ACTIVITÉS ADMINISTRATIVES</b>
----------------------------------

- ✓ Review 2007 : Lecture et commentaires sur une soumission dans une revue internationale (Food Quality And Preference)

## ACTIVITÉS D'ENSEIGNEMENT

### ATER à l'Université de Rennes 2

---

- ❑ Filière scientifique :
  - Master 1 Master Professionnel de Statistiques Appliquées :  
Cours et TD de sondages (30hETD) : 2007-2008 et 2008-2009  
*introduction aux sondages ; sondages aléatoires simples (avec et sans remise) ; sondage stratifié ; sondage par grappes, à plusieurs degrés ; principe de redressement*
  - Licence 1 MASS (Mathématiques Appliquées et Sciences Sociales)  
Cours, TD, TP de statistiques (30hETD) : 2007-2008 et 2008-2009  
*statistique descriptive sur une variable (vocabulaire, graphiques, concentration)  
statistique sur deux variables (contingence, test du  $\chi_2$ , régression linéaire, analyse de la variance)  
TP sur Excel et introduction à R*
- ❑ Filière Administration Economique et Sociale :
  - Licence 2 : Cours et TD de statistiques inférentielles (48hETD) : 2007-2008 et 2008-2009  
*probabilité, variables aléatoires ; échantillonnage et estimation ; comparaison et tests d'hypothèses*
  - Licence 3 : Cours et TD de sondages (48hETD) : 2007-2008 et 2008-2009  
*introduction aux sondages ; sondages aléatoires simples (sans remise) ; sondage stratifié ; sondage par grappes*
- ❑ Autres Filières :
  - Master 1 Master Professionnel Géographie et aménagement:  
TP de analyse des données (12hETD) : 2008-2009  
*statistique descriptive sur une et deux variables ; échantillonnage et estimation ; comparaison et tests d'hypothèses ; test du  $\chi_2$ , régression linéaire, analyse de la variance ; TP sur R.*
  - Licence 3  
Préparation au QCM IUFM mathématiques (36hETD) : 2007-2008  
Préparation au QCM IUFM mathématiques (24hETD) : 2008-2009

### VACATIONS

---

- ❑ Ecole Nationale Vétérinaire de Nantes
  - TD-TP en statistiques (54h) en 1<sup>ère</sup> année (2005)  
*statistique descriptive ; échantillonnage et estimation ; comparaison et tests d'hypothèses ; test du  $\chi_2$ , régression linéaire, analyse de la variance ; TP sur Excel-Stat.*
- ❑ Ecole Nationale de la Statistique et de l'Analyse de l'Information de Bruz
  - TD en statistiques inférentielles (40h) 1<sup>ère</sup> année école d'ingénieur, option mathématiques (2004)  
*modélisation statistique ; théorie de l'estimation (exhaustivité, estimation sans biais, estimation par maximum de vraisemblance) ; théorie des tests*
  - TD en probabilités (40h) 1<sup>ère</sup> année école d'ingénieur, option mathématiques (2003)  
*Espace probabilisé, variables aléatoires ; théorie de l'intégration (mesure...) ; lois de probabilités ; espérance, convergence.*

## ACTIVITÉS DE RECHERCHE

### EXPOSÉS ORAUX À DES CONFÉRENCES INTERNATIONALES

---

**Symposium on Learning and Data Science** les 1-3 avril 2009 à l'Université de Paris Dauphine:  
Participation au "Early Career Researchers" Challenge on Marketing Data Set : Follow-Up of Purchases of a Consumer Panel  
Communication en anglais et prix du jeune chercheur attribué par le comité scientifique.

**10<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques** les 23-25 janvier 2008 à LOUVAIN LA NEUVE (Belgique) : Communication sur "Modélisation de la dégradation sensorielle des aliments par des Chaînes de Markov Cachées".

**COMPSTAT2006** les 28 août - 1 septembre 2006 à ROME (Italie) :  
Communication sur "A new algorithm for the alignment of time intensity curves".

**7<sup>th</sup> Sensometrics meeting** les 28-30 juillet 2004 à UC DAVIS (USA, Californie) :  
Communication sur "A procedure for the analysis of Time Intensity curves".

**8<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques** les 10-12 mars 2004 à RENNES :  
Communication sur la "Performance d'un jury en profil sensoriel conventionnel".

### EXPOSÉS ORAUX À DES CONFÉRENCES NATIONALES

---

**XXXVII<sup>èmes</sup> journées de Statistiques** les 6-10 juin 2005 à PAU : Communication sur "L'ajustement de courbes de type temps-intensité en évaluation sensorielle en présence de répétitions".

**Chimiométrie 2004** les 30 novembre et 1 décembre 2004 à l'INA-PG à PARIS :  
Communication sur "L'ajustement de courbes de type temps-intensité en évaluation sensorielle".

**Chimiométrie 2003** les 3-4 décembre 2003 au CNAM à PARIS :  
Communication sur "La simplification et la signification des composantes principales dans le cadre d'une Analyse en Composantes Principales".

## AUTRES CONFÉRENCES, SÉMINAIRES ET GROUPE DE TRAVAIL

---

**Groupe de Travail 2007-2008** de l'équipe de statistique de Rennes 2 sur "The Elements of Statistical Learning" (Hastie *et al.*, 2001)

Exposés sur "Basis Expansions and Regularization" (chap 5), février-mars 2008

**Séminaire de statistiques** le 30 novembre 2007 à l'INSFA :

Exposé sur "Modélisation de la dégradation sensorielle des aliments par des Chaînes de Markov Cachées et Validation par une procédure de ré-échantillonnage".

**Forum des jeunes mathématiciennes** les 30-31 janvier 2004 à l'Institut Poincaré à PARIS :

Communication sur "Les données sensorielles de type Temps Intensité" concernant une présentation succincte des données de type temps intensité.

**Doctoriales** les 19-24 septembre 2004 à LA ROCHE SUR YON :

Ecriture d'un poster récapitulatif des travaux au cours de ma thèse.

**Rencontres doctorales 2004** les 3-4 mai 2004 à PIRIAC-SUR-MER :

Présentation de "Transformation non linéaire du temps et alignement de courbes. Application à l'analyse sensorielle".

## POSTERS À DES CONFÉRENCES INTERNATIONALES

---

**IUFOST** les 17-21 septembre 2006 à NANTES :

Poster sur "Alignment of TI curves in sensory analysis: comparison of methods and validation".

**8<sup>th</sup> Sensometrics meeting** les 2-4 août 2006 à ÅS (Norvège) :

Poster sur "A Hidden Markov Model to assess products shelf life".

**6<sup>th</sup> Pangborn Sensory Science Symposium** les 7-11 août 2005 à HARROGATE International Center (UKA, Yorkshire) : Poster sur "A Markovian model to assess products shelf-lives".

## RÉSUMÉ DES RÉSULTATS OBTENUS

J'ai essentiellement travaillé pendant ma thèse sur des données issues de l'analyse sensorielle. Dans les industries agro-alimentaires ainsi que dans d'autres secteurs d'activités, l'analyse sensorielle est la clé pour répondre aux attentes des consommateurs.

Cette discipline est le plus souvent basée sur l'établissement de profils sensoriels à partir de notes attribuées par des juges entraînés selon une liste de descripteurs (variables sensorielles). Les données peuvent être présentées sous forme d'un ensemble de tableaux, chaque tableau étant associé aux notes attribuées par un juge, dont les lignes sont associées aux produits et les colonnes, aux descripteurs. Dans ce type d'étude, il importe d'étudier la performance des juges et d'en tenir compte dans l'établissement des profils sensoriels. Dans cette perspective, nous avons proposé une démarche qui permet de procurer des indicateurs de performance du jury ( $\gamma$ ) et de chacun des juges ( $\alpha_i$ ) et de tenir compte de cette performance pour une détermination d'un tableau moyen (**C**).

Sur la base de ces indicateurs, nous avons proposé des tests d'hypothèses permettant de tester la significativité de la contribution de chacun des juges à la détermination du tableau moyen ainsi que la pertinence du tableau compromis obtenu. Ces tests d'hypothèses sont basés sur des procédures de permutations. Parallèlement, nous avons proposé des alternatives à ces tests basées sur des propriétés relatives à la moyenne et la variance de chacun des indicateurs lorsque toutes les permutations sont effectuées. Cette démarche a une portée générale, c'est ainsi que nous avons montré comment elle peut être aisément adaptée au cas des profils libres et la méthode de Procrustes Généralisée.

Depuis une vingtaine d'années, les courbes temps-intensité (TI) qui permettent de décrire l'évolution d'une sensation au cours de l'expérience sont de plus en plus populaires parmi les praticiens de l'analyse sensorielle. L'avantage de la description sensorielle par des courbes TI est de disposer d'une information supplémentaire sur la dynamique de la sensation dans l'objectif de mieux caractériser les produits. La difficulté majeure rencontrée en analyse sensorielle et accentuée en étude TI vient de l'appareil de mesure : les juges. En effet, il existe une grande variabilité interindividuelle qui est due à des différences physiologiques et des différences dans la manière de retranscrire l'information. Celle-ci se traduit par la présence d'une signature propre à chaque juge. Le but est donc de définir cette variabilité et de la réduire pour se focaliser sur la caractérisation des produits. Nous avons alors proposé une approche fonctionnelle basée sur les fonctions B-splines puis développé une procédure d'alignement qui permet de déterminer une transformation du temps ( $u_j$ ) propre à chaque juge  $j$  de nature à réduire les variations entre les juges. La fonction de perte associée à cette stratégie d'analyse s'inspire de celle de GPA (Generalized Procrustes Analysis) et l'algorithme de résolution consiste en un algorithme itératif qui utilise, entre autres, une adaptation de la procédure COW (Correlation Optimisation Warping) pour l'alignement de deux courbes. Enfin, une procédure de permutation permet de valider la pertinence de la démarche générale.

D'autres données sensorielles au cours du temps existent telles que le suivi de la dégradation organoleptique de produits alimentaires. Dans un premier temps, nous avons adopté une modélisation à l'aide de chaînes de Markov. Afin de tenir compte de la performance des juges, nous avons étendu la démarche précédente en considérant des chaînes de Markov cachées. Les résultats de l'analyse ont été étayés grâce à une procédure de type bootstrap et, afin de permettre une vision synthétique de l'évolution de la dégradation de plusieurs produits, nous avons effectué une AFC (Analyse Factorielle des Correspondances) sur les vecteurs d'états simulés grâce aux modèles Markoviens.

**Mots clés :** analyse de données fonctionnelles, chaînes de Markov, analyse sensorielle, temps-intensité, alignement, B-splines, performance, tableau moyen, tests de permutations.

Les données issues d'une épreuve de profil sensoriel conventionnel concernant  $n$  produits évalués par  $m$  juges entraînés pour les  $p$  descripteurs considérés, peuvent être présentées sous forme de  $m$  tableaux  $\mathbf{X}_i$  (de dimension  $n \times p$ ). L'objectif in fine est de synthétiser l'information et de caractériser les produits sur un plan organoleptique. Après une étape d'harmonisation au cours de laquelle les tableaux sont centrés et multipliés par des facteurs d'échelles de manière à avoir la même inertie, des différences entre les juges peuvent subsister ; elles sont liées à leurs performances. Afin de tenir compte de cet aspect, nous cherchons à déterminer une moyenne pondérée des différents tableaux en tenant compte de ces performances. En d'autres termes, nous souhaitons associer à un juge qui serait en désaccord avec le reste du panel un poids faible dans le calcul de la moyenne. Au-delà de cet aspect statistique consistant à pondérer les configurations des juges conduisant ainsi à une moyenne robuste, il faut souligner que les utilisateurs en évaluation sensorielle sont intéressés par la caractérisation de la performance de chacun des juges et la caractérisation de la performance globale de tout le panel. En s'appuyant sur une procédure proposée par Wakeling *et al.* (1992) dans le cadre de la méthode de Procruste généralisée (GPA) et par Kazi-Aoual *et al.* (1995) dans le cadre de la méthode STATIS, nous avons mis en place une démarche de validation de la performance des juges et du compromis, basée sur des tests de permutations.

□ Recherche du 'compromis'  $\mathbf{C}$

Nous considérons la matrice de similarité  $\mathbf{T}$  composée des indices ( $t_{ij}$ ) puis  $\lambda$  plus grande valeur propre de  $\mathbf{T}$  et  $\beta$  vecteur propre (normé) associé.  $\mathbf{C}$  est alors défini comme la moyenne pondérée des tableaux  $\mathbf{X}_i$  où le poids associé à chaque juge  $i$  est  $\beta_i$ .

$$t_{ii} = \frac{1 + \text{trace}(\mathbf{X}_i' \mathbf{X}_i)}{2}$$

□ Indicateurs de performance du jury ( $\gamma$ ) et de chacun des juges ( $\alpha_i$ )

$$\alpha_i = \text{trace}(\mathbf{X}_i' \mathbf{C}) \quad \gamma = \frac{1}{m} \sum_{i=1}^m \alpha_i.$$

□ Procédure de validation par des tests de permutations

Nous avons proposés des tests d'hypothèses pour juger si, à l'échelle globale, le compromis reflète de manière significative un point de vue global des juges. A l'échelle de chaque juge, nous posons une question similaire, à savoir : est-ce que la performance du juge  $i$  reflétée par  $\alpha_i$  traduit un accord significatif entre le juge  $i$  et le tableau compromis? En d'autres termes,  $\gamma$  et  $\alpha_i$  devraient être significativement différents de zéro et positifs.

La démarche que nous avons développée procure à l'utilisateur des outils qui lui permettent de porter un regard critique sur les données. Elle lui permet également de prendre des décisions telles que la mise à l'écart des données d'un juge non performant ou la mise en place d'un entraînement complémentaire. De plus, le compromis obtenu revêt un caractère robuste. Les tests d'hypothèses sont utiles car ils permettent de mieux étayer les conclusions. Les résultats théoriques concernant les tests de permutations procurent un moyen rapide pour effectuer les tests sans avoir à recourir à l'étude de simulations qui peut s'avérer fastidieuse.

Il faut souligner que la démarche concernant la détermination d'une moyenne pondérée peut être étendue à d'autres contextes.

Nous avons également eu recours à cette stratégie de pondération pour la détermination de courbes 'robustes'. De même, nous allons nous inspirer de la démarche poursuivie dans le cadre de la méthode GPA et notamment du problème d'optimisation, pour proposer une démarche d'analyse de courbes de type temps-intensité.

Dans les épreuves de profil sensoriel conventionnel, les notes attribuées peuvent être considérées comme des évaluations ponctuelles ne rendant donc pas compte de l'évolution des sensations. Cependant, de nombreux phénomènes intervenant durant l'évaluation d'un produit sont dynamiques, comme l'évolution de la texture durant la mastication ou la persistance d'une saveur. Il est alors possible d'enrichir les évaluations sensorielles statiques par l'acquisition d'un signal dit 'Temps-Intensité' (TI) décrivant l'évolution au cours du temps de l'intensité de la perception. De telles courbes ont tout d'abord été tracées pour étudier les évolutions de saveurs. Depuis 1978, l'intérêt de ces courbes pour l'analyse de la texture des aliments a été particulièrement souligné. Depuis une vingtaine d'années, le développement de systèmes automatisés dédiés à l'acquisition des signaux Temps-Intensité a considérablement accéléré la diffusion de cette méthodologie.

Pour l'analyse statistique des données issues d'une procédure TI, plusieurs approches ont été proposées; basées pour la plupart sur une extraction de paramètres caractéristiques des courbes. Plus récemment, de nouvelles modélisations ont visé à mieux retranscrire la continuité dans le temps du signal TI. Ces approches, dites fonctionnelles, reposent sur l'hypothèse que le signal observé est l'expression bruitée d'une courbe lisse. S'inspirant à la fois de l'approche classique d'extraction de paramètres et de l'approche fonctionnelle, la modélisation sur laquelle s'appuie notre travail consiste, entre autres, à décomposer le signal dans une base de fonctions B-splines.

Il est alors possible de déterminer une courbe moyenne des répétitions et une courbe compromis pour les juges en attribuant aux juges des poids en fonction de leur accord avec le reste du panel. Néanmoins, une particularité notable des courbes de type Temps Intensité, est la présence d'un effet juge important. En effet, il est bien connu que chaque juge a tendance à générer une courbe particulière (signature) qui ne peut pas être facilement éliminée (Van Buuren, 1992). L'objectif d'une étude TI est de pouvoir tenir compte de ces différences interindividuelles pour caractériser les produits et évaluer la performance des dégustateurs.

Notre démarche d'identification et de réduction de la variabilité interindividuelle présente une certaine similarité avec une approche proposée par Dijksterhuis *et al.* (1997). Elle permet de mettre en évidence les 'signatures' propres aux différents juges. Cette démarche consiste à déterminer un facteur  $\rho_j$  et une transformation du temps  $u_j$  associés au juge  $j$  ( $j=1, \dots, m$ ) de manière à minimiser le critère:

$$\sum_{i=1}^n \sum_{j=1}^m \|\rho_j f_{ij} \circ u_j - f_i\|^2, \text{ produit } i=1, \dots, n \text{ et juge } j=1, \dots, m.$$

Sous la contrainte :

$$\sum_{i=1}^n \sum_{j=1}^m \rho_j^2 \|f_{ij} \circ u_j\|^2 = \sum_{i=1}^n \sum_{j=1}^m \|f_{ij}\|^2$$

Il est clair que ce problème d'optimisation présente une grande similarité avec le critère qui est à la base de GPA. L'algorithme de résolution consiste en un algorithme itératif qui utilise, entre autres, une adaptation de la procédure COW (Correlation Optimisation Warping) pour l'alignement de deux courbes (Nielsen *et al.*, 1998). Il est également clair que ce problème d'optimisation reflète tout à fait l'objectif poursuivi par l'utilisateur. En effet, à travers ce problème d'optimisation, il apparaît que nous cherchons une courbe  $f_i$  représentative de l'évolution de la sensation du produit  $i$  ( $i=1, \dots, n$ ). Cette courbe 'compromis' apparaît comme une synthèse des courbes données par les différents juges. Une transformation  $u_j$ , qui est la même pour tous les produits et de ce fait caractéristique du juge  $j$  ( $j=1, \dots, m$ ), est déterminée de manière à mieux aligner les différentes courbes  $f_{ij}$ .

Une procédure de permutation permet de valider la pertinence de la démarche générale. Une fois l'effet juge réduit, la caractérisation des produits devient plus pertinente.

L'analyse sensorielle est utilisée pour le développement et l'amélioration des produits et également pour le suivi de la qualité des produits. Cela permet, entre autres, d'évaluer la date limite de consommation et l'évolution de la qualité des produits en fonction du temps. Dans beaucoup de situations, le vieillissement se manifeste par une dégradation des produits sur le plan organoleptique. C'est pourquoi il est préconisé de procéder à des évaluations sensorielles des produits à différents intervalles de temps. Nous nous sommes intéressés à l'évaluation sensorielle effectuée par un jury entraîné qui est appelé à s'exprimer sur la qualité d'un ou plusieurs produits à intervalles de temps réguliers. L'évaluation consiste à attribuer aux produits des états exprimant un niveau de fraîcheur allant de 'très frais' à 'impropre à la consommation'.

Comme nous l'avons souligné ci-dessus, le protocole expérimental qui est à la base de cette étude est utilisé pour la détermination de la date limite de consommation. Des analyses statistiques usuellement proposées dans ce cadre sont basées sur des modèles de survie. Cependant, leur objectif est relativement restreint en ce sens qu'ils visent à déterminer la date à partir de laquelle les produits deviennent impropres à la consommation. L'approche discutée ici permet d'avoir une vision dynamique de l'évolution des produits.

Nous considérons les différentes modalités (par exemple, de 'très frais' à 'très dégradé') comme étant des états d'une chaîne de Markov homogène et, pour chaque produit, nous ajustons un modèle markovien de premier ordre. L'accent est mis sur les tests d'hypothèses mis en oeuvre pour juger de l'adéquation du modèle et la comparaison des produits. Afin d'avoir une vision globale de l'évolution des produits sur la base d'une représentation graphique, nous proposons d'effectuer une analyse factorielle des correspondances (AFC) sur un tableau de données obtenu à partir des modèles associés aux différents produits.

Il est à souligner que la modélisation à l'aide d'une chaîne de Markov suppose que les juges sont cohérents dans leur évaluation, en ce sens qu'ils attribuent au produit considéré une suite d'états compatibles avec le sens de la dégradation. En pratique, ceci n'est pas garanti. Afin de pallier ce problème, nous proposons d'utiliser des chaînes de Markov cachées. Ceci permet d'éviter des prétraitements fastidieux sur les données.

## BIBLIOGRAPHIE

---

- Van Buuren (1992). Analysing time intensity responses in sensory evaluation. *Food Technology*, 46(2), pp. 101-114.
- Dijksterhuis et Eilers (1997). Modelling time-intensity curves using prototype curves. *Food Quality and Preference*, 8(2), pp. 131-140.
- Kazi-Aoual, Hitier, Sabatier et Lebreton (1995). Refined approximations to permutation tests for multivariate inference. *Computational Statistics and Data Analysis*, 20(6), pp. 643-656.
- Nielsen, Carstensen et Smedsgaard (1998). Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping. *Journal of Chromatography A*, 805, pp. 17-35.
- Wakeling, Raats et MacFie (1992). A comparison of consensus tests for generalised procrustes analysis. *Journal of Sensory Studies*, 7, pp. 91-96.

## PROGRAMME DE RECHERCHE

### ANALYSE DE DONNEES MARKETING, AUTRES TRAVAUX

---

J'ai travaillé dernièrement sur un jeu de données marketing sur le suivi d'un panel de consommateurs (tableau 3D dont une dimension est le temps), l'objectif étant de définir des groupes de consommateurs. Face aux problèmes de pouvoir d'achat qu'ils peuvent rencontrer, les consommateurs changent leur manière d'acheter: certains se dirigent vers les marques premiers prix d'autres restent fidèles à leurs marques habituelles. Ainsi, la définition de typologies de consommateurs concernant leur fidélité à une marque au cours du temps permettrait la mise en évidence de l'impact de campagnes marketing.

Ce travail a été réalisé en collaboration avec Sébastien Lê. En introduisant une notion de 'profil de consommation', nous avons défini une mesure de la loyauté/fidélité envers une marque. Nous avons ainsi pu mettre en avant quatre groupes de consommateurs et visualiser la position des consommateurs vis-à-vis des différentes marques présentes sur le marché grâce à une Analyse Factorielle des Correspondances (AFC).

Dans un second temps, nous nous sommes attachés à caractériser la régularité de cette fidélité au cours du temps en utilisant les résultats d'une Analyse Factorielle Multiple (AFM). En effet, l'inertie-intra associée à chaque individu diminue avec la proximité des profils au cours du temps.

Enfin, nous avons cherché le lien éventuel entre les marques et les typologies considérées.

Ce travail devrait mener à une publication post symposium et probablement à un autre projet de recherche.

J'ai pu, par ailleurs, pendant ma thèse réaliser d'autres travaux :

- simplification et signification des composantes principales: Nous avons montré que la procédure générale de détermination d'un tableau compromis peut s'adapter à d'autres contextes tels que l'analyse en composantes principales. Cela permet, en particulier, de tester la significativité des axes factoriels et la significativité de la contribution des variables à la détermination des composantes principales.
- segmentations de consommateurs
- prévision à partir de spectres Infra-rouge.

Je reste toujours ouverte à l'étude d'autres données que celles issues de l'agroalimentaire et continue de m'intéresser à d'autres problèmes.

### CHAÎNES DE MARKOV CACHÉES

---

Je poursuis ma collaboration avec Denys Pommeret sur la modélisation par des chaînes de Markov cachées en utilisant une démarche de ré-échantillonnage de type bootstrap pour valider ou relativiser les résultats obtenus. J'ai exposé ces résultats aux 10<sup>èmes</sup> journées Européennes Agro-industrie et méthodes statistiques en 2008. Je compte ensuite mettre en place des tests basés sur le ré-échantillonnage.

### ANALYSE DE DONNÉES SENSORIELLES

---

Pour l'alignement des courbes, nous avons mis en évidence des transformations non linéaires du temps associées aux juges. Ces transformations pourraient présenter un intérêt pour le praticien. En effet, ce dernier pourrait être intéressé par définir une typologie des juges en

fonction du comportement reflété par ces transformations et expliquer cette typologie par des facteurs physiologiques, par exemple.

Nous avons étudié le cas des courbes TI associées à une sensation. Il est bien évident, que pour avoir une caractérisation plus complète, il conviendrait de mettre en place des expériences de même type pour appréhender plusieurs sensations. La pratique la plus courante consiste à étudier les sensations les unes après les autres à l'aide d'une procédure de type TI. Il est clair qu'une telle démarche est très coûteuse et demande beaucoup de temps et beaucoup de disponibilité de la part des juges. C'est pourquoi, des méthodes alternatives ont été développées. Par exemple, la méthode DATI (Dual Attribute Time Intensity) consiste à évaluer deux sensations en même temps: une sensation suivant l'axe horizontal et l'autre suivant l'axe vertical (Duizer *et al.* 1997). L'entraînement pour ce genre d'étude est plus fastidieux et l'utilisation de deux axes simultanément ajoute une difficulté à l'expérience. Cependant, les résultats semblent satisfaisants. Pour leurs études, Duizer *et al.* (1997) n'ont pas trouvé de différences significatives entre les deux méthodes: SATI (Single Attribute Time Intensity) et DATI. En tout état de cause, ce type d'expériences conduit à des données encore plus complexes que celles que nous avons considérées. De ce point de vue, cela mériterait une étude approfondie.

Récemment, il a été proposé une nouvelle méthode d'acquisition de données sensorielles dans le temps qui permet de prendre en compte plus de descripteurs à la fois : la méthode de la Dominance Temporelle des Sensations (Pineau, 2004). Ce protocole expérimental est intéressant car il permet la caractérisation sensorielle des produits sur la base de plusieurs descripteurs dans un temps restreint et avec un coût raisonnable. Cela consiste à interroger les sujets sur plusieurs attributs. Au cours de l'expérimentation les juges doivent indiquer quel attribut est dominant et quelle note d'intensité attribuer à la sensation dominante. Le calcul du nombre de fois (parmi les différents juges et les différentes répétitions) où chaque descripteur est dominant pour un produit donné permet de tracer la série des fréquences de citations appelée courbe DTS. La collection des différentes courbes DTS d'un même produit donne une image de la temporalité du profil sensoriel du produit considéré. Cette méthode nécessite beaucoup de concentration pour les dégustateurs, une vitesse de réaction très grande et des capacités à noter tous les descripteurs. De plus, le problème de variabilité interindividuelle se pose toujours. En effet, les différentes sensations pourraient être perçues par le juge à des moments différents du fait de facteurs physiologiques ou autres inhérents aux juges. Par conséquent, il serait intéressant d'explorer comment adapter notre méthodologie de réduction de la variabilité des juges à ce type d'épreuves.

## ANALYSE DE DONNÉES FONCTIONNELLES

Je souhaiterais poursuivre mes recherches sur l'alignement de courbes en optimisant l'algorithme développé et m'intéresser à l'application des différentes modélisations à d'autres données que celles issues de l'analyse sensorielle.

## BIBLIOGRAPHIE

- Duizer, Bloom et Findlay (1997). Dual attribute time-intensity sensory evaluation: a new method for temporal measurement of sensory perceptions. *Food Quality and Preference*, 8(4), pp. 261-269.
- Pineau (2004). Méthodes d'acquisition, de codage et d'analyse de profils sensoriels temporels. *Dans 8èmes Journées Agro-Industries et Méthodes Statistiques*, pp. 87-93, Rennes, France.
- Pineau, Pessina, Cordelle, Imbert, Rogeaux et Schlich (2004). Temporal dominance of sensations: comparison with conventional profiling and time-intensity. *Dans 7th Sensometrics Meeting*, p. 40, Davis, Californie.