

CLAIRE CHABANET

SOPHIE NICKLAUS

VINCENT BOGGIO

SYLVIE ISSANCHOU

**Modèle linéaire mixte et analyse longitudinale  
des choix alimentaires**

*Journal de la société française de statistique*, tome 143, n° 1-2 (2002),  
p. 121-129

[http://www.numdam.org/item?id=JSFS\\_2002\\_\\_143\\_1-2\\_121\\_0](http://www.numdam.org/item?id=JSFS_2002__143_1-2_121_0)

© Société française de statistique, 2002, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# MODÈLE LINÉAIRE MIXTE ET ANALYSE LONGITUDINALE DES CHOIX ALIMENTAIRES<sup>†</sup>

Claire CHABANET \*, Sophie NICKLAUS \*, Vincent BOGGIO \*\*,  
Sylvie ISSANCHOU \*

## RÉSUMÉ

Les choix alimentaires de jeunes enfants sont suivis sur un an : le nombre de plats choisis, ainsi que la composition en protéines, glucides et lipides des repas composés par les enfants sont modélisés. Des modèles linéaires avec effets aléatoires et fixes permettent de prendre en compte le caractère longitudinal des données. Des outils graphiques permettant d'examiner les résidus et les effets aléatoires sont utilisés pour construire et valider le modèle. Une simulation montre les limites de l'utilisation du graphique empirique d'autocorrélation des résidus, utilisé pour vérifier l'indépendance des erreurs ou l'adéquation du modèle de dépendance entre les erreurs. Enfin, une présentation graphique du modèle estimé est proposée.

## ABSTRACT

Food choices of young children are followed up during one year : the number of dishes, and the protein, carbohydrate and lipid contents of the meals chosen by the children are modelised. Linear models including fixed and random effects allow to take into account the longitudinal character of the data. Graphical tools are used in order to build and validate the model. A simulation shows the limits of the within group residuals autocorrelations used to check the independence between errors, or that the modelisation of the dependance is right. Finally, a graphical presentation of the model is proposed.

## 1. Données, objectifs et intérêts des modèles linéaires mixtes

Entre 1982 et 1999, les choix de 397 enfants âgés de 2 à 3 ans ont été observés dans une crèche dijonnaise, lors des repas de midi présentés sous forme de libre service. Pour chaque repas, les enfants pouvaient choisir librement entre deux entrées, un plat de viande ou poisson, deux légumes ou féculents, deux

---

<sup>†</sup> Cette étude a reçu le soutien financier de l'INRA via le réseau « comportement du consommateur » et via une AIP de la Direction Scientifique « Nutrition Humaine et Sécurité des Aliments ».

\* INRA – Unité Mixte de Recherche sur les Arômes BP86510, 21065 DIJON Cedex France; chabanet@dijon.inra.fr

\*\* Faculté de Médecine de Dijon, Laboratoire de Physiologie.

produits laitiers, et du pain. Le personnel de la crèche enregistrait les choix de chaque enfant. En moyenne, chaque enfant a participé à 110 repas. Les nombres de plats choisis par repas ont été moyennés mensuellement. Les valeurs obtenues sont présentées pour quelques enfants (Fig. 1). Les nombres de calories apportées par repas, par les protéines, les glucides et les lipides ont été estimés, puis moyennés mensuellement. Le nombre de repas entrant dans ces moyennes mensuelles vaut 10 en moyenne, il varie entre 1 et 23 selon le nombre de repas pris à la crèche dans le mois, par chaque enfant. Des variables descriptives individuelles sont disponibles, comme le sexe, le mode d'allaitement ou l'indice de masse corporelle (imc), égal au rapport du poids sur le carré de la taille; cet indice permet d'évaluer la corpulence, il est mesuré à l'entrée de l'enfant dans l'étude.

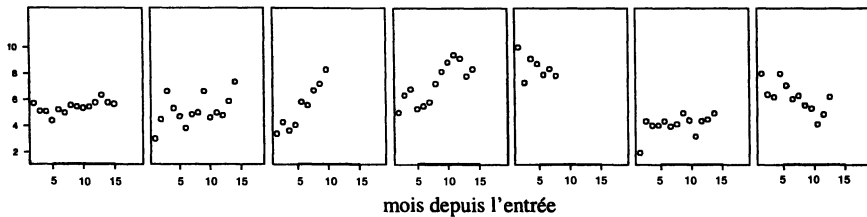


FIG. 1. – Moyennes mensuelles du nombre de plats choisis, en fonction du nombre de mois écoulés depuis l'entrée de l'enfant dans l'étude, pour sept des enfants.

Le cadre des modèles linéaires mixtes permet de décomposer la variabilité des données en une variabilité moyenne explicable par des effets fixes, une variabilité individuelle prise en compte par les effets aléatoires, et une variabilité résiduelle. L'approche choisie est exploratoire, le but de la modélisation étant la description et la quantification de la variabilité des comportements. Grâce à la modélisation, le comportement alimentaire moyen des enfants en situation de choix est caractérisé, la variabilité individuelle de ces comportements est quantifiée, les caractéristiques des enfants susceptibles d'expliquer la variabilité individuelle sont recherchées (Pinheiro, 2002).

## 2. Modélisation du nombre de plats choisis

### 2.1. stratégie de choix de modèle

Pour modéliser le nombre de choix, la stratégie de choix de modèle a été la suivante. Les variables et facteurs explicatifs sont le nombre de mois écoulés depuis l'entrée de l'enfant dans l'étude (temps), et le mois, qui permet de tenir compte d'une éventuelle variation saisonnière. Nous avons *a priori* considéré l'ordonnée à l'origine, et l'effet temps comme des effets fixes et aléatoires, en pensant qu'il existe une variabilité individuelle du niveau de consommation et de son évolution. Dans un premier temps, nous avons introduit l'effet saisonnier sous la forme d'un facteur mois à 12 niveaux à effets fixes (modèle 1). Cet effet saisonnier étant significatif, nous avons souhaité

MODÈLE LINÉAIRE MIXTE ET ANALYSE LONGITUDINALE

le modéliser, à la fois pour réduire le nombre de paramètres du modèle, et pour décrire l'effet. Nous avons modélisé les effets fixes saisonniers sous la forme d'une somme de 8 termes cosinus et sinus, de période  $T=12, 6, 4$  ou  $3$  mois (modèle 2), puis retenus les termes significatifs à 1% (modèle 3). Pour des raisons exposées en 2.2, ces effets saisonniers ont ensuite été introduits à la fois en effets fixes et en effets aléatoires (modèles 4, 5, 6). Enfin, l'introduction, en effets fixes, de caractéristiques des enfants susceptibles d'expliquer partiellement la variabilité individuelle a été envisagée, nous avons retenu un effet de l'imc qui est introduit en covariable (tous les modèles). La variance des erreurs a été supposée inversement proportionnelle au nombre de repas entrant dans le calcul de la moyenne mensuelle. Dès le modèle 1, le graphique des résidus réduits en fonction des valeurs ajustées révélait une hétérogénéité des variances. Cette hétérogénéité disparaît lorsque la variance des erreurs est supposée proportionnelle à une puissance de la moyenne, ce que nous avons choisi pour tous les modèles. Le plus complet des modèles envisagés est le suivant, tous les termes aléatoires étant supposés indépendants les uns des autres.

$$y_{ij} = \mu_i + \alpha_i \text{ temps}_{ij} + \beta \text{ imc}_i + \sum_{T=3,4,6,12} \left( \delta_{T_i} \cos\left(\frac{2\pi}{T} \text{mois}_{ij}\right) + \delta'_{T_i} \sin\left(\frac{2\pi}{T} \text{mois}_{ij}\right) \right) + \epsilon_{ij}$$

$$\mu_i \sim \mathcal{N}(\mu, \sigma_1^2), \alpha_i \sim \mathcal{N}(\alpha, \sigma_2^2), \delta_{T_i} \sim \mathcal{N}(\delta_T, \sigma_T^2), \delta'_{T_i} \sim \mathcal{N}(\delta'_T, \sigma_T'^2),$$

$$\epsilon_{ij} \sim \mathcal{N}\left(0, \frac{\sigma^2}{n \text{repas}_{ij}} \widehat{y}_{ij}^{2\gamma}\right)$$

$i = 1 \dots 397$  (enfants),  $j = 1, 2, 3, \dots$  (mois depuis l'entrée dans l'étude),  
 $\text{mois} = 1 \dots 12$  (Septembre ... Août).

	effets fixes	effets aléatoires
modèle 1	$\mu, \alpha, \beta, \text{mois}$	$\mu, \alpha$
modèle 2	$\mu, \alpha, \beta, \delta_3, \delta_4, \delta_6, \delta_{12}, \delta'_3, \delta'_4, \delta'_6, \delta'_{12}$	$\mu, \alpha$
modèle 3	$\mu, \alpha, \beta, \delta_4, \delta_{12}, \delta'_6$	$\mu, \alpha$
modèle 4	$\mu, \alpha, \beta, \delta_4, \delta_{12}, \delta'_6$	$\mu, \alpha, \delta_3, \delta_4, \delta_6, \delta_{12}, \delta'_3, \delta'_4, \delta'_6, \delta'_{12}$
modèle 5	$\mu, \alpha, \beta, \delta_4, \delta_{12}, \delta'_6$	$\mu, \alpha, \delta_6, \delta_{12}, \delta'_6, \delta'_{12}$
<b>modèle 6</b>	$\mu, \alpha, \beta, \delta_4, \delta_{12}, \delta'_6$	$\mu, \alpha, \delta_4, \delta_{12}, \delta'_6$

Les modèles envisagés résumés ci-dessus, sont des modèles linéaires mixtes, ils ont été estimés à l'aide de la fonction lme de la librairie NLME <sup>1</sup> (Pinheiro et Bates, 2000). L'estimateur utilisé est l'estimateur du maximum de vraisemblance restreint. Les tests des effets fixes ont été réalisés à l'aide des

1. version 3.3, juin 2000. (NLME est une librairie Splus).

tests de Wald, les tests des effets aléatoires à l'aide des tests du rapport de vraisemblance. Le modèle en gras est le modèle retenu.

La librairie NLME offre des outils graphiques d'aide au diagnostic qui permettent, le cas échéant, de détecter une inadéquation du modèle. La fonction «plot.lme» représente les résidus standardisés en fonction des valeurs estimées, c'est à dire en fonction des valeurs ajustées au niveau individuel  $\hat{\mu}_i + \hat{\alpha}_i \text{ temps}_{i,j} + \hat{\beta} \text{ imc}_i + \sum_{T=3,4,6,12} \left( \hat{\delta}_{T_i} \cos\left(\frac{2\pi}{T} \text{mois}_{i,j}\right) + \hat{\delta}'_{T_i} \sin\left(\frac{2\pi}{T} \text{mois}_{i,j}\right) \right)$ . Elle permet de repérer des observations extrêmes, de détecter une inadéquation de la modélisation de la variance des erreurs. Cette représentation peut être réalisée pour chacun des niveaux d'une autre variable, ce qui permet, par exemple de vérifier que la variance ne dépend pas du sexe.

La fonction «qqnorm.lme» permet de représenter les quantiles de la loi normale centrée réduite en fonction de chacun des effets aléatoires, ainsi qu'en fonction des résidus standardisés (représentation normale des effets ou des résidus). Si des points répartis autour d'une droite ne prouvent pas la normalité des erreurs ou des effets aléatoires, un écart systématique à la droite est le signe d'un écart à l'hypothèse de normalité de l'effet ou de l'erreur. Cet outil permet de détecter une inadéquation du choix de la normalité pour les erreurs ou les effets aléatoires, par exemple de détecter des distributions dissymétriques, multimodales, ou tout simplement des observations, ou des individus extrêmes. Dans le cas du nombre de choix, la normalité des ordonnées à l'origine, des pentes et des erreurs est parfaitement acceptable.

Enfin, dans notre modèle, les erreurs sont supposées indépendantes. Nous avons voulu vérifier cette hypothèse en utilisant conjointement la fonction «ACF.lme» et la fonction «plot.ACF», qui permettent de calculer et de présenter l'autocorrélation empirique des résidus standardisés.

## 2.2. autocorrélation des résidus standardisés

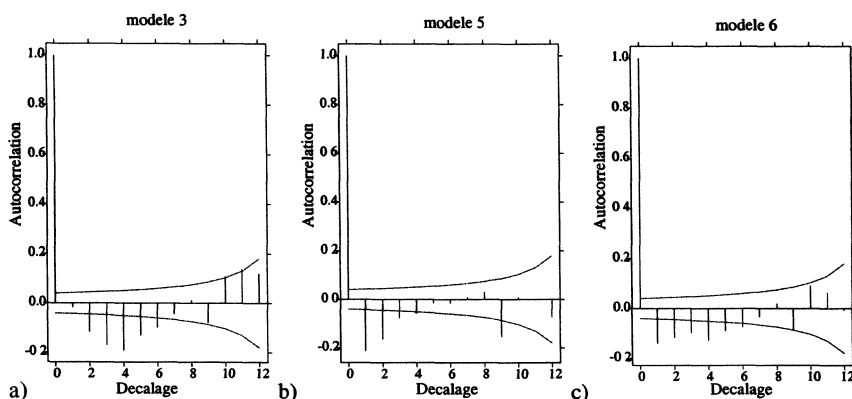


FIG. 2. – Modélisation du nombre moyen de plats choisis. Graphique empirique d'autocorrélation des résidus standardisés, dans le cas des modèles 3, 5, 6.

Le graphique d'autocorrélation des résidus standardisés du modèle 3 (Fig. 2a) révèle une structure qui nous a conduit à envisager des termes périodiques aléatoires. L'introduction de termes périodiques aléatoires modifie effectivement cette structure d'autocorrélation. Le choix des termes à introduire se pose. Une première stratégie a consisté à introduire des termes périodiques aléatoires cosinus et sinus de périodes 12, 6, 4, 3 mois (modèle 4), et à supprimer successivement les termes non significatifs, pour aboutir au modèle 5. Le modèle ainsi obtenu présente des autocorrélations négatives pour les décalages de 1 et 2 mois (Fig. 2b). Une seconde stratégie consiste à introduire en aléatoire les termes périodiques déclarés significatifs en effets fixes. L'hypothèse sous-jacente est l'existence d'un effet périodique fixe, dont l'amplitude varie aléatoirement d'un enfant à l'autre. On obtient alors des autocorrélations entre résidus significatives, négatives, ne dépendant pas du décalage (Fig. 2c). Par ailleurs, nous avons simulé un jeu de données avec 300 individus, 10 instants de mesure, une ordonnée à l'origine individuelle distribuée normalement, (moyenne 5, écart-type 1), et une pente fixe (0.05), une erreur résiduelle distribuée normalement (moyenne 0, écart-type 0.7). Le modèle composé d'une ordonnée à l'origine et d'une pente en effets fixes, et d'une ordonnée à l'origine aléatoire révèle des autocorrélations significatives, négatives, indépendantes du décalage, du même ordre de grandeur que celles que nous obtenons. Cette simulation montre les limites de l'outil : l'autocorrélation entre les résidus n'est pas égale à l'autocorrélation entre les erreurs. Si cet outil permet de détecter de fortes structures d'autocorrélation, la détection d'une faible autocorrélation des résidus ne signifie pas que les erreurs sont corrélées.

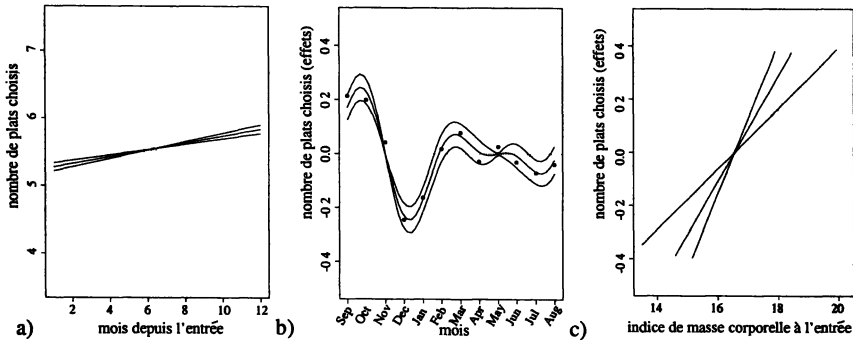


FIG. 3. - Nombre de plats choisis, effets fixes et effets aléatoires :

a) nombre moyen de plats choisis, pente et intervalle de confiance (trait plein) ; le trait pointillé vertical gauche présente la variabilité résiduelle (demi-hauteur égale à  $1.96 * \text{écart-type résiduel}$ ), celui de droite représente la variabilité individuelle, après prise en compte de l'imc. Les droites obliques en pointillé présentent la variabilité des pentes individuelles.

b) effet fixe du mois de l'année sur le nombre de plat choisis.

c) effet fixe de l'imc sur le nombre de plat choisis.

Les intervalles de confiance présentés ont une probabilité de recouvrement égale à 95 %.

Les effets fixes sont représentés en traits pleins, les effets aléatoires en pointillés.

Ce résultat nous amène à considérer le modèle 6, comme satisfaisant, malgré la structure d'autocorrélation observée. Enfin, un test robuste des effets fixes, réalisé avec SAS <sup>2</sup> (proc mixed, option empirical) sur le modèle 6, confirme la significativité des effets fixes retenus.

Les effets fixes et aléatoires estimés sont présentés graphiquement (Fig. 3). La variabilité individuelle du nombre de plats choisis est très grande, elle est partiellement expliquée par l'imc de l'enfant (la prise en compte de l'imc diminue de 3% l'écart-type individuel  $\widehat{\sigma}_1$ ). De même, la variabilité des pentes est importante. Les effets mois et imc sont présentés à la même échelle, ce qui permet de comparer l'amplitude des effets; l'effet de l'imc est grand comparé à l'effet mois.

### 3. Modélisation des pourcentages de protéines, glucides et lipides

Nous avons envisagé de modéliser directement les proportions, ou de les transformer par la transformation « centered log-ratio »  $\log \frac{p_i}{(p_1 p_2 p_3)^{1/3}}$ , où  $p_1, p_2, p_3$  désignent les proportions de protéines, glucides et lipides, et  $p_i$  désigne l'une des 3 proportions. Cette transformation a été proposée pour analyser des données de composition, et interpréter correctement la matrice de corrélations entre les différentes composantes (Aitchison, 1986). Pour chacune des composantes, nous avons considéré l'ordonnée à l'origine fixe et aléatoire, la pente fixe et aléatoire, un effet fixe de l'imc, des effets saisonniers fixes non interprétés en raison de la variation saisonnière des menus proposés (modèles 1 et 3 ci-dessous).

La modélisation simultanée des trois proportions a été réalisée en créant un pseudo-facteur « composante », à trois niveaux, un niveau par variable (glucides, lipides, protéines) (Thiébaud et al., 2002). Ce pseudo-facteur a été introduit dans le modèle sous forme d'effet principal, pour tenir compte du fait que la proportion moyenne est différente d'une composante à l'autre. Tous les effets fixes, et tous les effets aléatoires ont été introduits sous forme d'interaction avec le pseudo-facteur : si une variable explicative a pour effet d'augmenter l'une des proportions, cela doit se traduire par la diminution d'au moins une des deux autres proportions, de telle façon que la somme des trois proportions estimées soit égale à un. La variance résiduelle est supposée dépendante de la composante, et inversement proportionnelle au nombre de repas. Le plus complet des modèles envisagés s'écrit :

$$y_{ijk} = \mu_{ik} + \alpha_{ik} \text{temps}_{ij} + \beta_k \text{imc}_i + \gamma_{1k} \text{pain}_i + \gamma_{2k} \text{fromage}_i + \gamma_{3k} \text{viande}_i + \sum_{T=3,4,6,12} \left( \delta_{T_k} \cos\left(\frac{2\pi}{T} \text{mois}_{ij}\right) + \delta'_{T_k} \sin\left(\frac{2\pi}{T} \text{mois}_{ij}\right) \right) + \epsilon_{ijk}$$

$$\mu_{ik} \sim \mathcal{N}(\mu_k, \sigma_{1k}^2), \alpha_{ik} \sim \mathcal{N}(\alpha_k, \sigma_{2k}^2), \epsilon_{ijk} \sim \mathcal{N}\left(0, \frac{\sigma_k^2}{n \text{repas}_{ij}}\right)$$

MODÈLE LINÉAIRE MIXTE ET ANALYSE LONGITUDINALE

$i = 1 \dots 397$  (enfants),  $j = 1, 2, 3, \dots$  (mois depuis l'entrée dans l'étude),  
 $k = 1, 2, 3$  composantes (1 : protéines, 2 : glucides, 3 : lipides),  
 mois = 1 ... 12 (Sept ... Août),  $cor(\mu_{ik}, \mu_{ik'}) = \rho_{kk'}$  et  $cor(\alpha_{ik}, \alpha_{ik'}) = \rho'_{kk'}$ .  
 Tous les autres termes aléatoires sont supposés indépendants les uns des autres.

Les termes saisonniers conservés sont les termes significatifs à 1 %, dans le modèle 1 ci-dessous :  $\delta'_{12_1}, \delta_{6_1}, \delta'_{6_1}, \delta_{12_2}, \delta'_{4_2}, \delta_{12_3}, \delta'_{12_3}, \delta_{6_3}, \delta'_{6_3}$ . Les mêmes termes sont retenus, à partir du modèle 3.

modèle 1	$\mu_{ik}, \alpha_{ik}, \beta_k,$ termes saisonniers	sans transformation
<b>modèle 2</b>	$\mu_{ik}, \alpha_{ik}, \beta_k, \gamma_{1k}, \gamma_{2k}, \gamma_{3k},$ termes saisonniers	sans transformation
modèle 3	$\mu_{ik}, \alpha_{ik}, \beta_k,$ termes saisonniers	avec transformation « centered log-ratio »
modèle 4	$\mu_{ik}, \alpha_{ik}, \beta_k, \gamma_{1k}, \gamma_{2k}, \gamma_{3k},$ termes saisonniers	avec transformation « centered log-ratio »

Pour les modèles 1 et 3, les effets enfant aléatoires estimés présentaient des distributions fortement dissymétriques (Fig. 4a, modèle 1). Sachant qu'il existe une grande variabilité du niveau de consommation en fromage, que 7% des enfants n'en consomment pas du tout (Fig. 5c), nous avons pensé que le niveau de consommation en fromage pouvait influencer fortement la proportion de lipides consommés. De même, les niveaux individuels de consommation en viande et pain sont susceptibles d'influencer respectivement les proportions de protéines et de glucides.

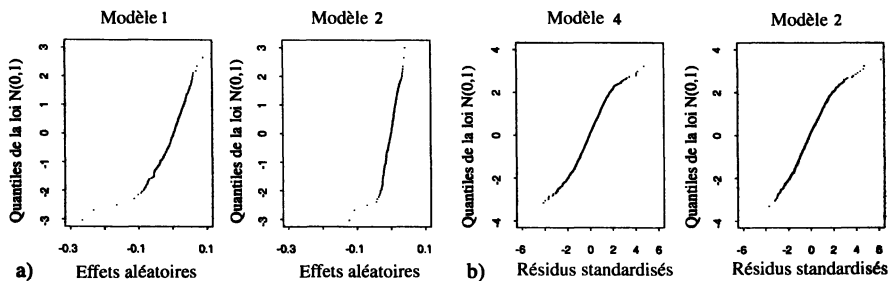


FIG. 4. – Modélisation de la proportion de lipides : a) quantiles de la loi normale  $N(0,1)$  en fonction des ordonnées à l'origine aléatoires, dans le cas des modèles 1 et 2; b) quantiles de la loi  $N(0,1)$  en fonction des résidus standardisés, dans le cas des modèles 4 et 2.

Nous avons donc estimé, pour chaque enfant, son niveau de consommation en fromage, en viande ou poisson, et en pain, en calculant le nombre moyen de choix par repas sur l'ensemble du temps de séjour à la crèche. Ces trois



variables sont introduites en effets fixes, dans les modèles 2 et 4. L'objectif est ici de rendre plus acceptable l'hypothèse de normalité des effets individuels. Il s'agit en même temps, de « réduire la variabilité individuelle » en l'expliquant partiellement par des caractéristiques individuelles introduites en covariables. La prise en compte des niveaux de consommation réduit les écarts-types individuels de 37 à 52 % selon les composantes. De même, nous avons recherché si des variables comme le sexe, le mode d'allaitement de l'enfant à sa naissance, etc ... pouvaient expliquer les pourcentages de protéines, glucides, lipides. Une fois pris en compte les niveaux individuels de consommation, la normalité des effets individuels est acceptable (Fig. 4a, modèle 2).

Par ailleurs, l'hypothèse de normalité des résidus est plus acceptable pour le modèle 2 que pour le modèle 4 (Fig. 4b). Nous avons donc choisi le modèle 2 basé sur la modélisation directe des proportions. La représentation normale des effets enfant aléatoires a permis de détecter quatre ou cinq enfants ayant des comportements extrêmes (consommation de glucides particulièrement élevée, ou consommation de lipides ou de protéines particulièrement faible). Le modèle a été estimé de nouveau en omettant ces quelques enfants pour ne pas surestimer la variabilité individuelle.

Les niveaux de consommation ont des effets déterminants sur les compositions. Ils expliquent partiellement la variabilité individuelle. En moyenne, la proportion de lipides est la plus élevée; les proportions de glucides et de lipides augmentent linéairement entre 2 et 3 ans, alors que la proportion de protéines diminue. Mais les proportions sont extrêmement variables d'un enfant à l'autre, de même que les pentes. Les déterminants principaux de la composition en glucides, lipides et protéines, sont respectivement les niveaux individuels de consommation en pain, fromage, et viande. Les enfants dont l'indice de masse corporelle est plus élevé que la moyenne consomment légèrement plus de glucides que les autres. Enfin, la variabilité résiduelle, c'est à dire la variabilité d'un mois à l'autre, pour un même enfant, est très élevée. Les résultats obtenus pour la proportion de lipides sont présentés graphiquement (Fig. 5).

#### 4. Conclusion

La construction d'un modèle requiert le choix des effets fixes, des effets aléatoires, de la structure de corrélation éventuelle entre effets aléatoires, d'un modèle pour la variance des erreurs, outre le choix des lois des termes aléatoires. Les outils de diagnostic, basés sur l'examen des effets aléatoires et des résidus sont indispensables à la construction du modèle. Nous avons montré que si les autocorrélations empiriques des résidus peuvent révéler des autocorrélations entre erreurs qu'il est nécessaire de modéliser, le fait de détecter une faible autocorrélation n'est pas nécessairement le reflet d'un modèle inadéquat. Nous présentons les modèles estimés sous forme graphique, en représentant simultanément les effets fixes, la variabilité associée aux effets aléatoires et la variabilité résiduelle.

L'intérêt de la modélisation réside dans la description du comportement alimentaire d'enfants dans une situation d'observation réelle, et surtout dans

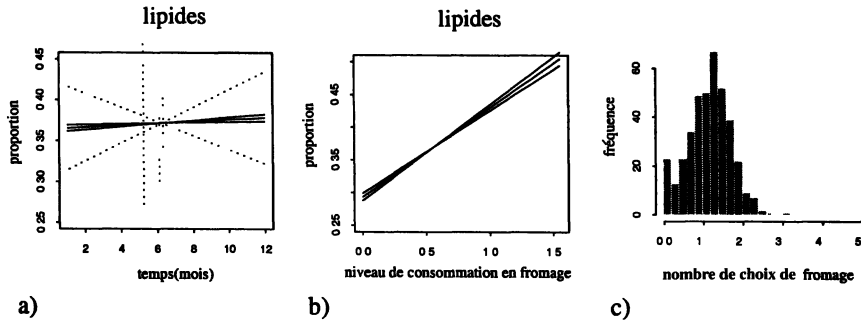


FIG. 5. – a) proportion moyenne de lipides, pentes et intervalles de confiance à 95 % sont représentés en traits pleins. Les 3 barres verticales en pointillé présentent, de gauche à droite, la variabilité résiduelle, la variabilité individuelle de l'ordonnée à l'origine, la variabilité individuelle de l'ordonnée à l'origine restant après prise en compte des niveaux de consommation en pain, fromage et viande (intervalles de confiance à 95 %). Les droites obliques en pointillé présentent la variabilité individuelle des pentes (intervalles de confiance à 95 %).  
 b) effets fixes et intervalles de confiance à 95 % du niveau de consommation en fromage (nombre de prises par repas) sur la proportion de lipides.  
 c) distribution des nombres de choix de fromage par repas.

l'observation de l'évolution de ce comportement. L'imc est rarement montré comme étant un facteur d'influence des niveaux de consommation dans les études expérimentales, ce qui renforce l'intérêt de ces données. Concernant la proportion de nutriments consommés, cette approche permet d'avancer des niveaux moyens mais souligne surtout l'importance de la variabilité individuelle, rarement présentée par ailleurs.

## RÉFÉRENCES

- AITCHISON J. (1986), *The Statistical Analysis of Compositional Data*, Chapman and Hall, London.
- PINHEIRO J.C., BATES D.M. (2000), *Mixed-effects Models in S and S-PLUS*, Springer-Verlag, New-York.
- PINHEIRO J.C. (2002), Model building using covariates in nonlinear mixed-effects models, *Journal de la Société Française de Statistique*, 143, 1-2, p. 79-101.
- THIÉBAUT R, JACQMIN-GADDA H., CHÈNE G., LEPORT C., COMMENGES D. (2002), Bivariate linear mixed models using SAS proc MIXED, *Comput. Methods Programs Biomed.*, Nov, 69 (3), 249.