Composante du produit nº 12-001-X
Division des méthodes d'enquêtes auprès des entreprises

Article

Estimation pour petits domaines sous un modèle à deux niveaux

par Mahmoud Torabi et J.N.K. Rao

Juin 2008

Statistique Canada   Statistics Canada

Canada

Techniques d'enquête, juin 2008
Vol. 34, N° 1, pp. 13-20
Statistique Canada, N° 12-001-X au catalogue

Estimation pour petits domaines sous un modèle à deux niveaux

Mahmoud Torabi et J.N.K. Rao [1]

Résumé

Lehtonen et Veijanen (1999) ont proposé un nouvel estimateur par la régression généralisée (GREG) assisté par modèle d'une moyenne de petit domaine sous un modèle à deux niveaux. Ils ont montré que l'estimateur proposé donne de meilleurs résultats que l'estimateur GREG habituel en ce qui concerne le biais relatif absolu moyen et l'erreur relative absolue médiane moyenne. Nous calculons l'erreur quadratique moyenne (EQM) du nouvel estimateur GREG sous le modèle à deux niveaux et nous la comparons à celle de l'estimateur fondé sur le meilleur prédicteur linéaire sans biais (BLUP). Nous présentons aussi des résultats empiriques concernant l'efficacité relative des estimateurs. Nous montrons que le nouvel estimateur GREG a de meilleures propriétés que l'estimateur GREG habituel en ce qui concerne l'EQM moyenne et l'erreur relative absolue moyenne. Nous montrons aussi que, parce qu'il emprunte de l'information aux petits domaines apparentés, l'estimateur EBLUP donne des résultats nettement meilleurs que l'estimateur GREG habituel et que le nouvel estimateur GREG. Nous fournissons les résultats de simulation sous un modèle, ainsi qu'en population finie réelle.

Mots clés : Estimateur BLUP; estimateur GREG; erreur quadratique moyenne; effets aléatoires; moyennes de petit domaine.

1. Introduction

L'Estimation pour petits domaines a suscité beaucoup d'intérêt ces dernières années à cause de la demande croissante de statistiques fiables pour des petits domaines. Les estimateurs par domaine directs classiques ne sont pas suffisamment précis, parce que, pour les petits domaines, les tailles d'échantillons sont souvent insuffisantes. Il faut donc recourir à des estimateurs indirects renforcés par l'emprunt d'information à des domaines apparentés, en particulier des estimateurs indirects fondés sur un modèle. En estimation sur petits domaines, des modèles à effets aléatoires au niveau de l'unité, dont les modèles de régression linéaire à erreurs emboîtées et les modèles à deux niveaux, sont fréquemment adoptés pour obtenir de bons estimateurs fondés sur un modèle des moyennes de petit domaine. Rao (2003) donne une revue complète de l'Estimation pour petits domaines fondée sur un modèle.

Un modèle à deux niveaux est donné par

$$y_{ij} = x'_{ij}\beta_i + e_{ij};$$
$$\beta_i = Z_i\beta + \nu_i,\ j = 1, ..., N_i; i = 1, ..., m \qquad (1)$$

où N_i est le nombre d'unités dans le i^e domaine ($i = 1, ..., m$), y_{ij} est la réponse et x_{ij} est un vecteur de dimension $p \times 1$ de covariables au niveau de l'unité attaché à la j^e unité dans le i^e domaine. En outre, Z_i est une matrice de dimensions $p \times q$ de covariables au niveau du domaine, β est un vecteur de dimension $q \times 1$ de paramètres de régression, les ν_i sont des vecteurs aléatoires indépendants de moyenne nulle et de covariance $\sum_\nu$, et les e_{ij} sont des variables aléatoires indépendantes de moyenne nulle et de variance σ_e^2 qui sont indépendantes des ν_i. Nous pouvons exprimer la moyenne $\bar{Y}_i$ du i^e domaine sous la forme

$$\bar{Y}_i \approx \mu_i = \bar{X}'_i(Z_i\beta + \nu_i),$$

en supposant que N_i est grand, où $\bar{X}_i$ est la moyenne de population connue de x_{ij} dans le i^e domaine. Nous supposons que les valeurs d'échantillon $\{(y_{ij}, x_{ij}), j = 1, ..., n_i; i = 1, ..., m\}$ obéissent au modèle (1), c'est-à-dire qu'il n'y a pas de biais de sélection d'échantillon. Pour l'échantillon, le modèle est donné par

$$y_{ij} = x'_{ij}(Z_i\beta + \nu_i) + e_{ij},\ j = 1, ..., n_i;\ i = 1, ..., m. \qquad (2)$$

En notation matricielle, (2) peut s'écrire

$$\mathrm{y}_i = X_i(Z_i\beta + \nu_i) + e_i,\ i = 1, ..., m$$

avec $\mathrm{Var}(\mathrm{y}_i) = V_i = X_i\Sigma_\nu X'_i + \sigma_e^2 I_{n_i}$, où y_i est un vecteur de dimension $n_i \times 1$ et X_i est une matrice de dimensions $n_i \times p$. Le modèle à deux niveaux (2) a été proposé pour la première fois par Moura et Holt (1999) dans le contexte de l'estimation sur petits domaines. Il intègre effectivement l'utilisation de covariables aux niveaux de l'unité et du domaine dans un seul modèle, par modélisation des pentes aléatoires β_i dans (1) en fonction de covariables au niveau du domaine Z_i.

Lehtonen et Veijanen (1999) ont proposé un nouvel estimateur par la régression généralisée (GREG) assisté par modèle d'une moyenne de petit domaine sous le modèle à deux niveaux. Ils ont montré que le nouvel estimateur GREG basé sur le modèle (1) donne de meilleurs résultats que l'estimateur GREG habituel fondé sur un modèle avec $\beta_i = Z_i\beta$ fixes. Moura et Holt (1999) ont obtenu

1. Mahmoud Torabi, Department of Pediatrics, University of Alberta, Edmonton (Alberta) T6G 2J3; J.N.K. Rao, School of Mathematics and Statistics, Carleton University, Ottawa (Ontario) K1S 5B6.

l'estimateur fondé sur le meilleur prédicteur linéaire sans biais (BLUP) de la moyenne de petit domaine μ_i et son EQM sous le modèle à deux niveaux (2); voir la section 2. Lehtonen, Särndal et Veijanen (2003) ont étudié l'effet du choix du modèle sur divers types d'estimateurs (synthétique, GREG et composite) de moyenne de petit domaine.

À la section 3, nous commençons par calculer l'erreur quadratique moyenne (EQM) du nouvel estimateur GREG et de l'estimateur GREG habituel (section 2) sous le modèle à deux niveaux (2), en supposant que les paramètres du modèle sont connus. Puis, nous comparons les EQM de l'estimateur GREG habituel, du nouvel estimateur GREG et de l'estimateur BLUP, et nous obtenons une expression explicite de l'accroissement de l'EQM du nouvel estimateur GREG relativement à l'EQM de l'estimateur BLUP. À la section 4, nous présentons des résultats empiriques concernant l'efficacité relative des estimateurs quand les paramètres du modèle sont estimés. Nous réalisons l'étude par simulation sous un cadre fondé sur un modèle, ainsi qu'en population finie réelle. Enfin, à la section 5, nous présentons certaines conclusions.

2. Estimateurs BLUP et GREG

Le modèle à deux niveaux (2) est un cas particulier d'un modèle linéaire mixte général avec une structure diagonale par blocs de la matrice de covariance. Par conséquent, en supposant que les paramètres du modèle sont connus, nous pouvons calculer l'estimateur BLUP de μ_i sous la forme

$$\tilde{\mu}_i^B = \bar{X}_i'(Z_i\beta + \tilde{v}_i), \tag{3}$$

où

$$\tilde{v}_i = \Sigma_v X_i' V_i^{-1}(y_i - X_i Z_i \beta),$$

et l'indice supérieur B sur $\tilde{\mu}_i$ désigne l'estimateur BLUP (Rao 2003, page 107). De même, l'estimateur BLUP d'une observation non échantillonnée j dans le i^{e} domaine peut s'écrire

$$\tilde{y}_{ij} = x_{ij}'(Z_i\beta + \tilde{v}_i). \tag{4}$$

Par ailleurs, un estimateur GREG assisté par modèle de μ_i (ou de $\bar{Y}_i$) est donné par

$$\tilde{\mu}_i^G = \frac{1}{N_i}\left[\sum_{j=1}^{N_i}\tilde{y}_{ij} + \sum_{j=1}^{n_i} w_{ij}(y_{ij} - \tilde{y}_{ij})\right],\ i = 1, \ldots, m \tag{5}$$

où $\tilde{y}_{ij}$ est le prédicteur de y_{ij} sous le modèle postulé, et w_{ij} est le poids de sondage qui est égal à N_i/n_i dans le cas de l'échantillonnage aléatoire simple (EAS) dans les domaines. Dans le présent exposé, nous nous intéressons essentiellement à l'EAS dans les domaines.

Partant de (5) avec $\tilde{y}_{ij} = x_{ij}' Z_i \beta$ pour prédicteur de y_{ij} sous le modèle (1) avec $\beta_i = Z_i\beta$ fixes, nous pouvons écrire l'estimateur GREG habituel sous la forme

$$\tilde{\mu}_i^G = \bar{y}_i + (\bar{X}_i - \bar{x}_i)' Z_i \beta, \tag{6}$$

où l'indice supérieur G sur $\tilde{\mu}_i$ signifie GREG (Särdnal, Swensson et Wretman 1992, page 225), $\bar{y}_i$ est la moyenne d'échantillon de y_{ij} dans le i^{e} domaine, et $\bar{x}_i$ est la moyenne d'échantillon de x_{ij} dans le i^{e} domaine, respectivement. En utilisant dans (5) le prédicteur (4) fondé sur le modèle à deux niveaux (1), nous obtenons un nouvel estimateur GREG de μ_i (ou de $\bar{Y}_i$) de la forme

$$\begin{aligned}\tilde{\mu}_i^{\text{LV}} &= [\bar{X}_i'(Z_i\beta + \tilde{v}_i) + (\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{v}_i))] \\ &= \bar{y}_i + (\bar{X}_i - \bar{x}_i)'(Z_i\beta + \tilde{v}_i),\end{aligned} \tag{7}$$

où l'indice supérieur LV sur $\tilde{\mu}_i$ indique qu'il a été proposé pour la première fois par Lehtonen et Veijanen (1999). Les estimateurs $\tilde{\mu}_i^B$, $\tilde{\mu}_i^G$ et $\tilde{\mu}_i^{\text{LV}}$ sont linéaires en les y_{ij} et sans biais sous le modèle à deux niveaux (1). En pratique, nous remplaçons les paramètres β, Σ_v et σ_e^2 dans (3), (6) et (7) par des estimateurs appropriés. Les estimateurs résultants sont désignés par $\hat{\mu}_i^B$, $\hat{\mu}_i^G$ et $\hat{\mu}_i^{\text{LV}}$, respectivement, où $\hat{\mu}_i^B$ est l'estimateur BLUP empirique (EBLUP). Sous l'hypothèse de normalité, $\hat{\mu}_i^B$ est le meilleur estimateur empirique (EB pour *empirical best*). L'estimateur EBLUP de $\bar{Y}_i$ est donné à la section 4.2.2. Notons que $\hat{\mu}_i^G$ et $\hat{\mu}_i^{\text{LV}}$ sont valides en tant qu'estimateurs de $\bar{Y}_i$.

3. Erreur quadratique moyenne

L'erreur quadratique moyenne (EQM) de l'estimateur GREG habituel $\tilde{\mu}_i^G$ sous le modèle à deux niveaux peut s'écrire

$$\begin{aligned}\text{EQM}(\tilde{\mu}_i^G) &= E(\tilde{\mu}_i^G - \mu_i)^2 \\ &= E[\bar{y}_i + (\bar{X}_i - \bar{x}_i)' Z_i\beta - \bar{X}_i'(Z_i\beta + v_i)]^2 \\ &= E[(\bar{x}_i - \bar{X}_i)' v_i + \bar{e}_i]^2 \\ &= (\bar{x}_i - \bar{X}_i)' \Sigma_v (\bar{x}_i - \bar{X}_i) + \frac{\sigma_e^2}{n_i},\end{aligned}$$

comme il est énoncé dans le théorème 1.

Théorème 1. *L'EQM de l'estimateur GREG* (6) *est donnée par*

$$\text{EQM}(\tilde{\mu}_i^G) = (\bar{x}_i - \bar{X}_i)' \Sigma_v (\bar{x}_i - \bar{X}_i) + \frac{\sigma_e^2}{n_i}. \tag{8}$$

En outre, nous pouvons écrire l'EQM de l'estimateur BLUP $\tilde{\mu}_i^B$ comme il suit :

$$\begin{aligned}\text{EQM}(\tilde{\mu}_i^B) &= E(\tilde{\mu}_i^B - \mu_i)^2 \\ &= E[\bar{X}_i'(\tilde{\nu}_i - \nu_i)]^2 \\ &= \bar{X}_i'\left(\Sigma_v - \Sigma_v X_i' V_i^{-1} X_i \Sigma_v\right)\bar{X}_i,\end{aligned}$$

comme il est énoncé dans le théorème 2.

Théorème 2. *L'EQM de l'estimateur BLUP* (3) *est donnée par*

$$\text{EQM}(\tilde{\mu}_i^B) = \bar{X}_i'\left(\Sigma_v - \Sigma_v X_i' V_i^{-1} X_i \Sigma_v\right)\bar{X}_i. \qquad (9)$$

Le théorème 3 donne l'EQM du nouvel estimateur GREG $\tilde{\mu}_i^{\text{LV}}$.

Théorème 3. *L'EQM du nouvel estimateur GREG* (7) *est donnée par*

$$\text{EQM}(\tilde{\mu}_i^{\text{LV}}) = \text{EQM}(\tilde{\mu}_i^B) + \left\{\bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - \bar{x}_i' \Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i}\right\}. \quad (10)$$

La preuve du théorème 3 est donnée en annexe.

Par définition, nous avons $\text{EQM}(\tilde{\mu}_i^B) \le \text{EQM}(\tilde{\mu}_i^{\text{LV}})$ et (10) donne une expression explicite de l'accroissement de l'EQM de $\tilde{\mu}_i^{\text{LV}}$ comparativement à l'EQM de l'estimateur BLUP $\tilde{\mu}_i^B$.

4. Résultats empiriques

4.1 Comparaison empirique des valeurs de l'EQM

Afin d'étudier l'efficacité du nouvel estimateur GREG, nous avons utilisé des données tirées de Moura et Holt (1999) recueillies auprès de 38 740 ménages dans les districts de recensement (petits domaines) d'un comté du Brésil. Le revenu du chef de ménage a été traité comme la variable de réponse y. Deux variables indépendantes ont été définies au niveau de l'unité, à savoir le niveau d'études du chef du ménage (échelle ordinale de 0 à 5) et le nombre de pièces dans le logement du ménage (1 à 11+). Nous avons supposé que le modèle à deux niveaux pour ces données était de la forme :

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + e_{ij}, \quad j = 1, \ldots, N_i;\ i = 1, \ldots, m \qquad (11)$$

avec

$$\beta_{i0} = \beta_0 + \nu_{i0};\ \beta_{i1} = \beta_1 + \nu_{i1};\ \beta_{i2} = \beta_2 + \nu_{i2}, \quad (12)$$

où

$$\nu = (\nu_{i0}, \nu_{i1}, \nu_{i2})' \overset{\text{iid}}{\sim} N_3(0, \Sigma_v),\ e_{ij} \overset{\text{iid}}{\sim} N(0, \sigma_e^2)$$

et x_1 et x_2 représentent, respectivement, le nombre de pièces et le niveau d'études du chef du ménage (centrées chacune autour de leur moyenne de population). Notons que le modèle (12) pour les coefficients β_i aléatoires ne contient pas de covariables Z au niveau du domaine.

Moura et Holt (1999) ont également étudié un autre modèle avec une covariable Z au niveau du domaine pour la modélisation des β_i dans (12). Pour ces données, le nombre moyen d'automobiles par ménage dans chaque domaine a été utilisé comme covariable z pour la modélisation des coefficients aléatoires β_{i1} et β_{i2} correspondant aux variables x_1 et x_2, mais non pour le terme constant aléatoire, β_{i0}. Le modèle à deux niveaux (11) avec la covariable z au niveau du domaine est donné par

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + \varepsilon_{ij}, \quad j = 1, \ldots, N_i;\ i = 1, \ldots, m \qquad (13)$$

avec

$$\beta_{i0} = \beta_0 + \nu_{i0};$$
$$\beta_{i1} = \beta_1 + \alpha_1 z_i + \nu_{i1};\ \beta_{i2} = \beta_2 + \alpha_2 z_i + \nu_{i2}. \quad (14)$$

Moura et Holt (1999) ont ajusté les modèles (11)-(12) et (13)-(14) à l'ensemble de données complet susmentionné. Nous résumons leurs résultats au tableau 1.

Tableau 1
Estimations des paramètres fondées sur l'ensemble de données de Moura et Holt (1999), où σ_0^2, σ_1^2 et σ_2^2 sont les éléments diagonaux et σ_{01}, σ_{02} et σ_{12} sont les éléments non diagonaux de la matrice de covariance Σ_v

Paramètre	Covariance diagonale : modèle (14) avec z	Covariance diagonale : modèle (12) sans z	Covariance générale : modèle (12) sans z
β_0	8,442	8,688	8,456
β_1	0,451	1,321	1,223
β_2	0,744	2,636	2,596
α_1	3,779	-	-
α_2	1,659	-	-
σ_0^2	0,745	0,637	1,385
σ_1^2	0,237	0,471	0,234
σ_2^2	0,700	1,472	0,926
σ_{01}	-	-	0,354
σ_{02}	-	-	0,492
σ_{12}	-	-	0,333
σ_e^2	44,00	44,01	47,74

Les moyennes de domaine de x_1 et x_2 sont calculées d'après l'ensemble complet de données et traitées comme étant les moyennes de population $\bar{X}_{1i}$ et $\bar{X}_{2i}$. Nous avons sélectionné un sous-échantillon aléatoire de 10 % des enregistrements dans chaque petit domaine. La taille globale de l'échantillon est $n = 3\,876$ et le nombre de petits domaines est $m = 140$. En utilisant les valeurs d'échantillon de x_1, x_2 et z, ainsi que les moyennes de population $\bar{X}_{1i}$ et $\bar{X}_{2i}$, nous avons calculé $\text{EQM}(\tilde{\mu}_i^G)$, $\text{EQM}(\tilde{\mu}_i^B)$ et $\text{EQM}(\tilde{\mu}_i^{\text{LV}})$ en utilisant (8), (9) et (10), respectivement, et en traitant les estimations des paramètres de régression, Σ_v et σ_e^2 du tableau 1 pour les données complètes comme étant les valeurs réelles. Ensuite, nous avons calculé les valeurs moyennes de l'EQM sur les domaines :

$$\overline{\text{EQM}}_G = \frac{1}{m}\sum_{i=1}^m \text{EQM}(\tilde{\mu}_i^G),$$

$$\overline{\text{EQM}}_B = \frac{1}{m}\sum_{i=1}^m \text{EQM}(\tilde{\mu}_i^B)$$

et

$$\overline{\text{EQM}}_{\text{LV}} = \frac{1}{m}\sum_{i=1}^m \text{EQM}(\tilde{\mu}_i^{\text{LV}}).$$

Nous définissions l'efficacité relative de $\tilde{\mu}^B$ par rapport à $\tilde{\mu}^G$ comme étant EFF_B et l'efficacité relative de $\tilde{\mu}^{\text{LV}}$ par rapport à $\tilde{\mu}^G$ comme étant EFF_{LV}, où

$$\text{EFF}_B = \frac{\overline{\text{EQM}}_G}{\overline{\text{EQM}}_B}; \qquad \text{EFF}_{\text{LV}} = \frac{\overline{\text{EQM}}_G}{\overline{\text{EQM}}_{\text{LV}}}.$$

Nous résumons les résultats aux tableaux 2 et 3. Ces derniers révèlent que le nouvel estimateur GREG est un peu plus efficace que l'estimateur GREG habituel en ce qui concerne l'EQM moyenne : $\text{EFF}_{\text{LV}} \le 112\,\%$. Toutefois, il est considérablement moins efficace que l'estimateur BLUP sous le modèle à deux niveaux postulé. Par exemple, pour le modèle avec z et la matrice de covariance diagonale (tableau 2), $\text{EFF}_B = 292\,\%$ comparativement à $\text{EFF}_{\text{LV}} = 106\,\%$, et $\overline{\text{EQM}}_B = 0{,}62$ comparativement à $\overline{\text{EQM}}_{\text{LV}} = 1{,}72$.

Tableau 2
Comparaison des estimateurs sur petits domaines : efficacité relative (EFF) et EQM moyenne ($\overline{\text{EQM}}$) pour le cas de la matrice de covariance diagonale basé sur l'ensemble de données de Moura et Holt (1999)

Mesure de la qualité	Modèle sans z GREG	Modèle sans z Nouveau GREG	Modèle sans z BLUP	Modèle avec z GREG	Modèle avec z Nouveau GREG	Modèle avec z BLUP
EFF	100 %	112 %	306 %	100 %	106 %	292 %
$\overline{\text{EQM}}$	1,92	1,71	0,62	1,83	1,72	0,62

Tableau 3
Comparaison des estimateurs sur petits domaines : efficacité relative (EFF) et EQM moyenne ($\overline{\text{EQM}}$) pour le cas d'une matrice de covariance générale basé sur l'ensemble de données de Moura et Holt (1999)

Mesure de la qualité	GREG	Modèle sans z Nouveau GREG	BLUP
EFF	100 %	108 %	253 %
$\overline{\text{EQM}}$	2,02	1,87	0,80

4.2 Étude par simulation

4.2.1 Étude par simulation sous un cadre fondé sur un modèle

Afin d'étudier l'efficacité du nouvel estimateur GREG lorsque les paramètres du modèle sont estimés, nous avons entrepris une petite étude par simulation fondée sur les modèles à deux niveaux (11)-(12) et (13)-(14). Nous avons considéré uniquement la structure de covariance diagonale Σ_v avec les éléments diagonaux σ_0^2, σ_1^2 et σ_2^2. Nous avons de nouveau utilisé les données tirées de Moura et Holt (1999). Les estimations de $\beta_0, \beta_1, \beta_2, \alpha_1, \alpha_2, \sigma_0^2, \sigma_1^2, \sigma_2^2$ et σ_e^2 présentées au tableau 1 sont traitées comme les valeurs réelles.

Dans notre étude par simulation, nous avons pris (x_{1ij}, x_{2ij}, z_i) provenant de Moura et Holt (1999), puis nous avons généré y_{ij} en nous fondant sur les modèles (11)-(12) et (13)-(14). En utilisant les échantillons ainsi produits $(y_{ij}^{(b)}, x_{1ij}, x_{2ij}, z_i)$, $b = 1, \ldots, B = 1\,000$, nous avons calculé $\hat{\beta}^{(b)}$ par les moindres carrés généralisés pour la nouvelle méthode GREG ainsi que pour la méthode BLUP. Pour la méthode GREG, nous avons utilisé les moindres carrés ordinaires pour estimer β par $\hat{\beta}_{\text{ols}}^{(b)}$. En outre, nous avons calculé $\hat{\Sigma}_v^{(b)}$ et $\hat{\sigma}_e^{2(b)}$ par la méthode du maximum de vraisemblance restreint (REML). Pour chaque échantillon produit, nous avons calculé

$$\mu_i^{(b)} = \bar{X}'_i(Z_i\beta + \nu_i^{(b)}),\ i = 1, \ldots, m;\ b = 1, \ldots, B.$$

Nous avons calculé le nouvel estimateur GREG de $\mu_i^{(b)}$ selon $\hat{\mu}_i^{\text{LV}(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, l'estimateur GREG de $\mu_i^{(b)}$ selon $\hat{\mu}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)' Z_i \hat{\beta}_{\text{ols}}^{(b)}$ et l'estimateur BLUP empirique (EBLUP) de $\mu_i^{(b)}$ selon $\hat{\mu}_i^{B(b)} = \bar{X}'_i(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, où $\hat{\nu}_i^{(b)} = \hat{\Sigma}_v^{(b)} X'_i \hat{V}_i^{-1(b)}(\mathrm{y}_i^{(b)} - X_i Z_i \hat{\beta}^{(b)})$.

Ensuite, nous avons calculé la moyenne des erreurs quadratiques moyennes ($\overline{\text{EQM}}_1$) et la moyenne des erreurs relatives absolues ($\overline{\text{ERA}}_1$)

$$\overline{\text{EQM}}_1 = \frac{1}{m}\sum_i^m \text{EQM}_{1i} \ \text{ où }\ \text{EQM}_{1i} = B^{-1}\sum_{b=1}^B (\hat{\mu}_i^{(b)} - \mu_i^{(b)})^2,$$

$$\overline{\text{ERA}}_1 = \frac{1}{m}\sum_i^m \text{ERA}_{1i} \ \text{ où }\ \text{ERA}_{1i} = B^{-1}\sum_{b=1}^B |\ \hat{\mu}_i^{(b)} - \mu_i^{(b)}|\ /\mu_i^{(b)},$$

où $\hat{\mu}_i^{(b)}$ dénote $\hat{\mu}_i^{\mathrm{LV}(b)}$, $\hat{\mu}_i^{G(b)}$ ou $\hat{\mu}_i^{B(b)}$. Nous présentons les résultats au tableau 4. Les deux modèles avec covariable z au niveau du domaine et sans covariable z produisent des valeurs un peu plus faibles de $\overline{\mathrm{EQM}}_1$ et de $\overline{\mathrm{ERA}}_1$ pour le nouvel estimateur GREG que pour l'estimateur GREG habituel. Cependant, $\overline{\mathrm{EQM}}_1$ et $\overline{\mathrm{ERA}}_1$ sont significativement plus faibles pour l'estimateur EBLUP, à cause de l'emprunt d'information aux domaines apparentés. En outre, la comparaison des tableaux 2 et 4 révèle que les valeurs de $\overline{\mathrm{EQM}}_1$ dans le tableau 4 sont un peu plus grandes que les valeurs correspondantes dans le tableau 2, à cause de l'estimation des paramètres du modèle.

Tableau 4
Comparaison des estimateurs sur petits domaines : EQM moyenne ($\overline{\mathrm{EQM}}_1$) et erreur relative absolue moyenne ($\overline{\mathrm{ERA}}_1$) sous un cadre fondé sur un modèle

Mesure de la qualité	Modèle sans z			Modèle avec z		
	GREG	Nouveau GREG	EBLUP	GREG	Nouveau GREG	EBLUP
$\overline{\mathrm{EQM}}_1$	1,93	1,73	0,67	1,84	1,73	0,73
$\overline{\mathrm{ERA}}_1$	0,14	0,13	0,08	0,13	0,12	0,08

4.2.2 Étude par simulation sous un cadre de population finie

Pour étudier les propriétés des estimateurs sous un cadre de population finie, nous avons créé une population finie synthétique d'après les données brésiliennes constituées de $n = 3\,876$ valeurs d'échantillon $(y_{ij}, x_{1ij}, x_{2ij}, z_i)$. En reproduisant les valeurs d'échantillon $(y_{ij}, x_{1ij}, x_{2ij}, z_i)$ cinq fois, nous avons obtenu le nouvel ensemble de données (y, x_1, x_2, z) de taille 19 380 que nous avons traité comme notre population réelle.

Nous avons produit 500 échantillons indépendants (B = 500), chacun de taille $n = 700$ et $n = 1\,400$, en tirant des échantillons aléatoires simples de taille $n_i = 5$ et $n_i = 10$ dans chaque domaine $i = 1, \ldots, 140$. Comme auparavant, pour chaque échantillon, nous avons calculé $\hat{\beta}^{(b)}$ pour la nouvelle méthode GREG et pour la méthode BLUP, et $\hat{\beta}_{\mathrm{ols}}^{(b)}$ pour la méthode GREG. En outre, nous avons calculé $\hat{\Sigma}_v^{(b)}$ et $\hat{\sigma}_e^{2(b)}$ par la méthode du REML. Nous avons également calculé la moyenne de population de y_{ij} pour chaque domaine i par

$$\bar{Y}_i = \sum_{j=1}^{N_i} y_{ij} / N_i, \; i = 1, \ldots, 140,$$

où N_i est la taille de la population dans le i^{e} domaine. En outre, pour chaque échantillon $b = 1, \ldots, B,$ nous avons calculé la nouvelle estimation GREG de la moyenne du i^{e} domaine selon $\hat{\bar{Y}}_i^{\mathrm{LV}(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})'(Z_i\hat{\beta}^{(b)} + \hat{v}_i^{(b)})$, l'estimation GREG selon $\hat{\bar{Y}}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})' Z_i \hat{\beta}_{\mathrm{ols}}^{(b)}$ et l'estimation EBLUP selon $\hat{\bar{Y}}_i^{B(b)} = f_i\,\bar{y}_i + (1 - f_i)[\bar{X}_i^{*\prime}(Z_i\hat{\beta}^{(b)} + \hat{v}_i^{(b)})]$, où

$$f_i = n_i / N_i, \; \bar{X}_i^* = \frac{N_i\,\bar{X}_i - n_i\,\bar{x}_i}{N_i - n_i},$$

et $\hat{v}_i^{(b)} = \hat{\Sigma}_v^{(b)} X_i' \, \hat{V}_i^{-1(b)}(\mathrm{y}_i^{(b)} - X_i\, Z_i\, \hat{\beta}^{(b)})$.

L'estimateur EBLUP tient compte des corrections pour population finie f_i.

Nous avons calculé la moyenne des erreurs quadratiques moyennes ($\overline{\mathrm{EQM}}_2$) et la moyenne des erreurs relatives absolues ($\overline{\mathrm{ERA}}_2$) comme étant

$$\overline{\mathrm{EQM}}_2 = \frac{1}{m}\sum_i^m \mathrm{EQM}_{2i}, \quad \overline{\mathrm{ERA}}_2 = \frac{1}{m}\sum_i^m \mathrm{ERA}_{2i},$$

où

$$\mathrm{EQM}_{2i} = \frac{1}{B}\sum_{b=1}^{B}(\hat{\bar{Y}}_i^{(b)} - \bar{Y}_i)^2, \quad \mathrm{ERA}_{2i} = \frac{1}{B}\sum_{b=1}^{B}|\hat{\bar{Y}}_i^{(b)} - \bar{Y}_i| / \bar{Y}_i,$$

et $\hat{\bar{Y}}_i^{(b)}$ dénote $\hat{\bar{Y}}_i^{\mathrm{LV}(b)}$, $\hat{\bar{Y}}_i^{G(b)}$ ou $\hat{\bar{Y}}_i^{B(b)}$. Nous présentons les résultats aux tableaux 5 et 6 pour $n_i = 5$ et $n_i = 10$, respectivement. Nous envisageons les modèles avec covariable au niveau du domaine z et sans la covariable z.

Tableau 5
Comparaison des estimateurs sur petits domaines : EQM moyenne ($\overline{\mathrm{EQM}}_2$) et erreur relative absolue moyenne ($\overline{\mathrm{ERA}}_2$) sous un cadre de population finie ($n_i = 5$)

Mesure de la qualité	Modèles sans z			Modèle avec z		
	GREG	Nouveau GREG	EBLUP	GREG	Nouveau GREG	EBLUP
$\overline{\mathrm{EQM}}_2$	11,03	10,02	6,50	10,76	10,06	7,06
$\overline{\mathrm{ERA}}_2$	0,27	0,24	0,18	0,25	0,23	0,22

Le tableau 5 montre que, pour $n_i = 5$, le nouvel estimateur GREG donne de meilleurs résultats que l'estimateur GREG habituel en ce sens que $\overline{\mathrm{EQM}}_2$ et $\overline{\mathrm{ERA}}_2$ sont plus faibles. Par ailleurs, le tableau 6 révèle que, pour $n_i = 10$, l'estimateur GREG habituel a d'un peu meilleures propriétés que le nouvel estimateur GREG en ce qui concerne $\overline{\mathrm{EQM}}_2$, mais non $\overline{\mathrm{ERA}}_2$. Cependant, l'estimateur EBLUP donne des valeurs considérablement plus faibles de $\overline{\mathrm{EQM}}_2$ et $\overline{\mathrm{ERA}}_2$ que l'estimateur GREG habituel et le nouvel estimateur GREG dans les deux cas, parce qu'il emprunte de l'information aux petits domaines apparentés. Par exemple, pour le modèle sans z et $n_i = 5$, $\overline{\mathrm{EQM}}_2$ = 10,02, 11,03 et 6,50 pour le nouvel estimateur GREG, l'estimateur GREG habituel et l'estimateur EBLUP, respectivement.

Tableau 6
Comparaison des estimateurs sur petits domaines : EQM moyenne ($\overline{\text{EQM}}_2$) et erreur relative absolue moyenne ($\overline{\text{ERA}}_2$) sous un cadre de population finie ($n_i = 10$)

Mesure de la qualité	Modèle sans z			Modèle avec z		
	GREG	Nouveau GREG	EBLUP	GREG	Nouveau GREG	EBLUP
$\overline{\text{EQM}}_2$	6,53	6,77	4,73	6,75	6,96	5,24
$\overline{\text{ERA}}_2$	0,20	0,18	0,15	0,19	0,18	0,19

5. Sommaire

Dans le présent article, nous avons calculé l'erreur quadratique moyenne (EQM) de modélisation pour un nouvel estimateur GREG assisté par un modèle à deux niveaux d'une moyenne de petit domaine proposé par Lehtonen et Veijanen (1999). En outre, nous avons utilisé un ensemble de données de Moura et Holt (1999) pour démontrer empiriquement que l'estimateur BLUP est considérablement plus efficace que le nouvel estimateur GREG en ce qui concerne l'EQM sous le modèle, tandis que le nouvel estimateur GREG n'est que légèrement plus efficace que l'estimateur GREG habituel fondé sur le modèle de régression $y_i = X_i Z_i \beta + e_i, i = 1, \ldots, m.$ En outre, à l'aide d'une étude par simulation sous un cadre fondé sur un modèle, nous avons montré que le nouvel estimateur GREG a systématiquement de meilleures propriétés que l'estimateur GREG habituel en ce qui concerne l'EQM moyenne, $\overline{\text{EQM}}$, et l'erreur relative absolue moyenne, $\overline{\text{ERA}}$. Cependant, étant donné l'emprunt d'information à des petits domaines apparentés, l'estimateur EBLUP donne de nettement meilleurs résultats que les estimateurs GREG habituel et nouveau. En outre, nous avons réalisé une étude par simulation sous un cadre de population finie et montré que l'estimateur EBLUP surpasse le nouvel estimateur GREG et l'estimateur GREG habituel en ce qui concerne l' $\overline{\text{EQM}}$ et l' $\overline{\text{ERA}}$.

Remerciements

La présente étude a été financée par une subvention du Conseil de recherches en sciences naturelles et en génie du Canada. Elle est fondée sur un chapitre de la thèse de doctorat de M.Torabi rédigée sous la supervision de J.N.K. Rao. Les auteurs remercient les examinateurs et le rédacteur adjoint de leurs commentaires constructifs concernant la version originale du présent article.

Annexe

Calcul de $\text{EQM}(\tilde{\mu}_i^{\text{LV}})$:

$$\begin{aligned}\text{EQM}(\tilde{\mu}_i^{\text{LV}}) &= E(\tilde{\mu}_i^{\text{LV}} - \mu_i)^2 \\ &= E[\bar{X}_i'(\tilde{v}_i - v_i)]^2 + E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{v}_i)]^2 \\ &\quad + 2E[(\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{v}_i))\bar{X}_i'(\tilde{v}_i - v_i)], \quad \text{(A.1)}\end{aligned}$$

où le premier terme du deuxième membre de (A.1) est l'EQM de l'estimateur BLUP sous le modèle à deux niveaux donnée par (9). De surcroît, nous pouvons écrire

$$\begin{aligned}E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{v}_i)]^2 &= E[\bar{y}_i - \bar{x}_i' Z_i \beta]^2 \\ &\quad + E(\bar{x}_i'\tilde{v}_i)^2 - 2E[(\bar{y}_i - \bar{x}_i' Z_i\beta)(\bar{x}_i'\tilde{v}_i)], \quad \text{(A.2)}\end{aligned}$$

où

$$E[\bar{y}_i - \bar{x}_i' Z_i\beta]^2 = \text{Var}(\bar{y}_i) = \bar{x}_i'\Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i},$$

et

$$\begin{aligned}E(\bar{x}_i'\tilde{v}_i)^2 &= \text{Var}(\bar{x}_i'\tilde{v}_i) + [E(\bar{x}_i'\tilde{v}_i)]^2 \\ &= \text{Var}\left[\bar{x}_i'\Sigma_v X_i' V_i^{-1}(\mathrm{y}_i - X_i Z_i\beta)\right] \\ &\quad + \left[E\left(\bar{x}_i'\Sigma_v X_i' V_i^{-1}(\mathrm{y}_i - X_i Z_i\beta)\right)\right]^2 \\ &= \bar{x}_i'\Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i.\end{aligned}$$

En outre,

$$\begin{aligned}E[(\bar{y}_i - \bar{x}_i' Z_i\beta)(\bar{x}_i'\tilde{v}_i)] &= E\left[\bar{y}_i \bar{x}_i'\Sigma_v X_i' V_i^{-1}(\mathrm{y}_i - X_i Z_i\beta)\right] \\ &\quad - E\left[\bar{x}_i' Z_i\beta \bar{x}_i'\Sigma_v X_i' V_i^{-1}(\mathrm{y}_i - X_i Z_i\beta)\right],\end{aligned}$$

où le deuxième terme est nul. Par conséquent, nous pouvons écrire

$$\begin{aligned}E[(\bar{y}_i - \bar{x}_i' Z_i\beta)(\bar{x}_i'\tilde{v}_i)] &= E\left[\bar{y}_i \bar{x}_i'\Sigma_v X_i' V_i^{-1}\mathrm{y}_i\right] \\ &\quad - \bar{x}_i' Z_i\beta \bar{x}_i'\Sigma_v X_i' V_i^{-1} X_i Z_i\beta,\end{aligned}$$

où le premier terme peut s'écrire

$$\begin{aligned}&E\left[\frac{1'}{n_i}\mathrm{y}_i \bar{x}_i'\Sigma_v X_i' V_i^{-1}\mathrm{y}_i\right] = \\ &\qquad \frac{1'}{n_i}\left(X_i\Sigma_v\bar{x}_i + X_i Z_i\beta\bar{x}_i'\Sigma_v X_i' V_i^{-1} X_i Z_i\beta\right),\end{aligned}$$

en utilisant le lemme suivant :

LEMME 1 (Searle 1971). *Si y est un vecteur de dimension* $n \times 1$ *de moyenne* μ *et de matrice de variance-covariance* Σ *et que* b *est un vecteur de dimension* $n \times 1$, *alors* $E(\mathrm{y}\, b'\mathrm{y}) = \Sigma b + \mu b' \mu$.

Donc,

$$\begin{aligned} &E[(\bar{y}_i - \vec{x}_i' Z_i \beta)(\vec{x}_i' \tilde{\nu}_i)] \\ &\quad = \vec{x}_i' \Sigma_\nu \bar{x}_i + \vec{x}_i' Z_i \beta \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta \\ &\quad - \vec{x}_i' Z_i \beta \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta \\ &\quad = \vec{x}_i' \Sigma_\nu \bar{x}_i. \end{aligned}$$

Nous pouvons alors écrire (A.2) sous la forme :

$$\begin{aligned} E[\bar{y}_i - \vec{x}_i'(Z_i \beta + \tilde{\nu}_i)]^2 &= \vec{x}_i' \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i} \\ &+ \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{x}_i - 2 \vec{x}_i' \Sigma_\nu \bar{x}_i \\ &= \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{x}_i - \vec{x}_i' \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i}. \end{aligned} \quad \text{(A.3)}$$

Enfin, nous devons trouver le terme produit croisé de (A.1). Nous avons

$$\begin{aligned} &E[(\bar{y}_i - \vec{x}_i'(Z_i \beta + \tilde{\nu}_i)) \bar{X}_i' (\tilde{\nu}_i - \nu_i)] \\ &\quad = E[\bar{y}_i \bar{X}_i' (\tilde{\nu}_i - \nu_i)] \\ &\quad - E[\vec{x}_i'(Z_i \beta + \tilde{\nu}_i) \bar{X}_i' (\tilde{\nu}_i - \nu_i)], \quad \text{(A.4)} \end{aligned}$$

où le premier terme du deuxième membre de (A.4) peut s'écrire

$$\begin{aligned} &E[\bar{y}_i \bar{X}_i' (\tilde{\nu}_i - \nu_i)] \\ &\quad = E\left[\bar{y}_i \bar{X}_i' \left(\Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta) - \nu_i\right)\right] \\ &\quad = E\left[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} \mathrm{y}_i\right] \\ &\quad - E\left[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta\right] \\ &\quad - E[\bar{y}_i \bar{X}_i' \nu_i]. \end{aligned} \quad \text{(A.5)}$$

Le premier terme de (A.5) est

$$\begin{aligned} &E\left[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} \mathrm{y}_i\right] \\ &\quad = E\left[\frac{1'}{n_i} \mathrm{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} \mathrm{y}_i\right] \\ &\quad = \frac{1'}{n_i} X_i \left(\Sigma_\nu \bar{X}_i + Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta\right), \quad \text{(A.6)} \end{aligned}$$

le deuxième terme de (A.5) peut s'écrire

$$\begin{aligned} &E\left[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta\right] \\ &\quad = \vec{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta, \quad \text{(A.7)} \end{aligned}$$

et le troisième terme peut être exprimé par

$$\begin{aligned} E[\bar{y}_i \bar{X}_i' \nu_i] &= E[(\vec{x}_i' Z_i \beta + \vec{x}_i' \nu_i + e_i) \bar{X}_i' \nu_i] \\ &= E[\vec{x}_i' \nu_i \bar{X}_i' \nu_i] = \vec{x}_i' \Sigma_\nu \bar{X}_i. \end{aligned} \quad \text{(A.8)}$$

Par conséquent, en substituant (A.6), (A.7) et (A.8) dans (A.5), nous avons

$$E[\bar{y}_i \bar{X}_i' (\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.9)}$$

Nous nous penchons maintenant sur le deuxième terme de (A.4). Nous avons

$$\begin{aligned} &E[\vec{x}_i'(Z_i \beta + \tilde{\nu}_i) \bar{X}_i' (\tilde{\nu}_i - \nu_i)] \\ &\quad = E[\vec{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] + E[\vec{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] \\ &\quad - E[\vec{x}_i' Z_i \beta \bar{X}_i' \nu_i] - E[\vec{x}_i' \tilde{\nu}_i \bar{X}_i' \nu_i]. \quad \text{(A.10)} \end{aligned}$$

Alors, nous obtenons l'expression qui suit pour les quatre termes du deuxième membre de (A.10) :

$$\begin{aligned} &E[\vec{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] \\ &\quad = E\left[\vec{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta)\right] = 0, \quad \text{(A.11)} \end{aligned}$$

$$\begin{aligned} E[\vec{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] &= \vec{x}_i' \Sigma_{\tilde{\nu}} \bar{X}_i + \vec{x}_i' E(\tilde{\nu}_i) \bar{X}_i' E(\tilde{\nu}_i) \\ &= \vec{x}_i' \operatorname{Var}\left[\Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta)\right] \bar{X}_i \\ &+ \vec{x}_i' E\left(\Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta)\right) \\ &\times \bar{X}_i' E\left(\Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta)\right) \\ &= \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{X}_i, \end{aligned} \quad \text{(A.12)}$$

$$E[\vec{x}_i' Z_i \beta \bar{X}_i' \nu_i] = 0 \quad \text{(A.13)}$$

et

$$\begin{aligned} &E[\vec{x}_i' \tilde{\nu}_i \bar{X}_i' \nu_i] \\ &\quad = E\left[\vec{x}_i' \Sigma_\nu X_i' V_i^{-1} (\mathrm{y}_i - X_i Z_i \beta) \bar{X}_i' \nu_i\right] \\ &\quad = E\left[\vec{x}_i' \Sigma_\nu X_i' V_i^{-1} \mathrm{y}_i \bar{X}_i' \nu_i\right] \\ &\quad - E\left[\vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta \bar{X}_i' \nu_i\right] \\ &\quad = E\left[\vec{x}_i' \Sigma_\nu X_i' V_i^{-1} (X_i Z_i \beta + X_i \nu_i + e_i) \bar{X}_i' \nu_i\right] \\ &\quad = E\left[\vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \nu_i \bar{X}_i' \nu_i\right] \\ &\quad = \vec{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{X}_i. \end{aligned} \quad \text{(A.14)}$$

Par conséquent, en introduisant (A.11) à (A.14) par substitution dans (A.10), nous obtenons

$$E[\vec{x}_i'(Z_i \beta + \tilde{\nu}_i) \bar{X}_i' (\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.15)}$$

Donc, il découle de (A.4), (A.9) et (A.15) que

$$E[(\bar{y}_i - \vec{x}_i'(Z_i \beta + \tilde{\nu}_i)) \bar{X}_i' (\tilde{\nu}_i - \nu_i)] = 0. \quad \text{(A.16)}$$

Il découle maintenant de (A.1), (A.3) et (A.16) que

$$\text{EQM}(\tilde{\mu}_i^{\text{LV}}) = \text{EQM}(\tilde{\mu}_i^{B}) + \left\{ \bar{x}_i' \Sigma_v X_i' V_i^{-1} X_i \Sigma_v \bar{x}_i - \bar{x}_i' \Sigma_v \bar{x}_i + \frac{\sigma_e^2}{n_i} \right\},$$

comme il est énoncé au théorème 3.

Bibliographie

Lehtonen, R., et Veijanen, A. (1999). Domain estimation with logistic generalized regression and related estimators. *IASS Satellite Conference on Small Area Estimation*, Riga : Latvian Council of Science, 121- 128.

Lehtonen, R., Särndal, C.-E. et Veijanen, A. (2003). L'effet du choix d'un modèle dans l'estimation par domaine, dont les petits domaines. *Techniques d'enquête*, 29, 37-49.

Moura, F.A.S., et Holt, D. (1999). Production d'estimations régionales à partir de modèles multiniveau. *Techniques d'enquête*, 25, 81-89.

Rao, J.N.K. (2003). *Small Area Estimation*. Hoboken : New York : John Wiley & Sons, Inc.

Särndal, C.-E., Swensson, B. et Wretman, J.H. (1992). *Model Assisted Survey Sampling*. New York : Springer-Verlag.

Searle, S.R. (1971). *Linear Models*. New York : John Wiley & Sons, Inc.