Centre for Statistical and Survey Methodology

The University of Wollongong

Working Paper

16-09

Outlier Robust Small Area Estimation

R Chambers, H Chandra, N Salvati and N Tzavidis

Centre for Statistical & Survey Methodology, University of Wollongong, Wollongong NSW 2522. Phone +61 2 4221 5435, Fax +61 2 4221 4845. Email: anica@uow.edu.au

Outlier Robust Small Area Estimation

R. Chambers[1], H. Chandra[2], N. Salvati[3] and N. Tzavidis[4]

Abstract: Outliers are a well-known problem in survey estimation, and a variety of approaches have been suggested for dealing with them in this context. However, when the focus is on small area estimation using the survey data, much less is known – even though outliers within a small area sample are clearly much more influential than they are in the larger overall sample. To the best of our knowledge, Chambers and Tzavidis (2006) was the first published paper in small area estimation that explicitly addressed the issue of outlier robustness, using an approach based on fitting outlier robust M-quantile models to the survey data. Recently, Sinha and Rao (2009) have also addressed this issue from the perspective of linear mixed models. Both these approaches, however, use plug-in robust prediction. That is, they replace parameter estimates in optimal, but outlier sensitive, predictors by outlier robust versions. Unfortunately, this approach may involve an unacceptable prediction bias (but a low prediction variance) in situations where the outliers are drawn from a distribution that has a different mean to the rest of the survey data (Chambers, 1986), which then leads to the suggestion that outlier robust prediction should include an additional term that makes a correction for this bias. In this paper, we explore the extension of this idea to the small area estimation situation and we propose two different analytical mean squared error (MSE) estimators for outlier robust predictors of small area means. We use simulation based on realistic outlier contaminated data to evaluate how the extended small area estimation approach compares with the plug-in robust methods described earlier. The empirical results show that the bias-corrected predictive estimators are less biased than the projective estimators especially when there are outliers in the area effects. Moreover, in the simulation experiments we contrast the proposed MSE estimators with those generally utilized for the plug-in robust predictors. The proposed bias-robust and linearization-based MSE estimators appear to perform well when used with the robust predictors of small area means that are considered in this paper.

Key words and phrases: Linear mixed model; M-quantile model; M-estimation. Robust prediction; Bias-variance trade-off; EBLUP; Robust bias correction.

[1] *University of Wollongong Wollongong 2522, New South Wales, Australia. E-mail: ray@uow.edu.au*

[2] *Indian Agricultural Statistics Research Institute. E-mail: hchandra@iasri.res.in*

[3] *University of Pisa. E-mail: salvati@ec.unipi.it*

[4] *University of Manchester. E-mail: nikos.tzavidis@manchester.ac.uk*

1. Introduction

Outliers are a fact of life for any survey, and especially so for business surveys. As a consequence, a variety of methods have been devised to mitigate the effects of outlier values on survey estimates. Some of these, like identification and removal of these data values by ‘experienced’ data experts during survey processing, can be effective in ensuring that the resulting survey estimates are unaffected by them. However, being somewhat subjective, such methods are not amenable to scientific evaluation. As a consequence there are a number of ‘objective’ methods for survey estimation that use statistical rules to decide whether an observation is a potential outlier and to down-weight its contribution to the survey estimates if this is the case. Generally, an outlier robust estimator of this type is based on the assumption that the non-outlier sample values all follow a well-behaved working model and so it generally involves prediction of the sum (or mean) of these values under this working model. In practice, this often involves replacement of an outlying sample value by an estimate of what it should have been if in fact it had been generated under the working model. We refer to such methods as *Robust Projective* in what follows since they project sample non-outlier behaviour on to the non-sampled part of the survey population.

Robust Projective methods essentially emulate the subjective approach described earlier, and typically lead to biased estimators with lower variances than would otherwise be the case. The reason for the bias is not difficult to find – it is extremely unlikely that the non-sampled values in the target population are drawn from a distribution with the same mean as the sample non-outliers, and yet these methods are built on precisely this assumption. Chambers (1986) recognised this dilemma and coined the concept of a ‘representative outlier’, i.e. a sample outlier that is potentially drawn from a group of population outliers and hence cannot be unit-weighted in estimation. He noted that representative outliers cannot be treated on the same basis in estimation as other sample data more consistent with the working model for the population, since such values can seriously destabilise the survey estimates, and suggested addition of an outlier robust bias correction term to a Robust Projective survey estimator, e.g. one based on outlier-robust estimates of the model parameters. Welsh and Ronchetti (1998) expand on this idea, applying it more generally to estimation of the finite population distribution of a survey variable in the presence of representative outliers. A similar idea is implicit in the approach described in Chambers *et al*. (1993), where a nonparametric bias correction is suggested. In what follows, we refer to methods that allow for contributions from representative sample outliers as *Robust Predictive* since they attempt to predict the contribution of the population outliers to the population quantity of interest.

If outliers are a concern for estimation of population quantities, it is safe to say that they are even more of a concern in small area estimation, where sample sizes are considerably smaller and model-dependent estimation is the norm. It is easy to see that an outlier that destabilises a population estimate based on a large survey sample will almost certainly destroy the validity of the corresponding direct estimate for the small area from which the outlier is sourced since this estimate will be based on a much smaller sample. This problem does not disappear when the small area estimator is an indirect one, e.g. an Empirical Best Linear Unbiased Predictor (EBLUP), since the weights underpinning this estimator will still put most emphasis on data from the small area of interest, and the estimates of the model parameters underpinning the estimator will themselves be destabilised by the sample outliers. Consequently, it is of some interest to see how outlier robust survey estimation can be adapted to this situation.

Chambers and Tzavidis (2006) explicitly address this issue of outlier robustness, using an approach based on fitting outlier robust M-quantile models to the survey data. Recently, Sinha and Rao (2009) have also addressed this issue from the perspective of linear mixed models. Both these approaches, however, use plug-in robust prediction. That is, they replace parameter estimates in optimal, but outlier sensitive, predictors by outlier robust versions (a Robust Projective approach). Unfortunately, this approach may involve an unacceptable prediction bias (but a low prediction variance) in situations where the outliers are drawn from a distribution that has a different mean to the rest of the survey data.

After discussing Robust Projective estimators for small areas in Section 2, we explore the extension of Chambers (1986) Robust Predictive approach to the small area estimation situation in Section 3. In Section 4 we propose two different analytical mean squared error (MSE) estimators for outlier robust predictors of small area means. In particular, the first proposal is based on bias-robust mean squared error estimation discussed by Chambers *et al*. (2007) and represents an extension of the ideas in Royall and Cumberland (1978). We show how this approach can be useful for estimating the MSE of small area predictors based on the Sinha and Rao (2009) approach. The second MSE estimator is developed under the conditional version of the linear mixed model and it uses the first order approximations to the variances of solutions of estimating equations. This last approach can be used for estimating the MSE of a wide variety of small area 'pseudo-linear' predictors, i.e. predictors that can be written as weighted sums, where the weights are sample data dependent. Examples of such predictors are mixed model and M-quantile model-based predictors under both the Robust Projective and the Robust Predictive approaches. In Sections 5 and 6 we use model-based simulations based on realistic outlier contaminated data scenarios as well as design-based simulations to evaluate how these two different approaches compare, both in terms of estimation performance as well as in terms of MSE estimation performance. Section 7 concludes the paper with some final remarks, and a discussion of future research aimed at outlier robust small area inference.

2. Robust Projective Estimation for Small Areas

In what follows we assume that unit record data are available at small area level. For the sampled units in the population this consists of indicators of small area affiliation, values y_j of the variable of interest, values $\mathbf{x}_j$ of a $p \times 1$ vector of individual level covariates, and values $\mathbf{z}_j$ of a vector of area level covariates. For the non-sampled population units we do not know the values of y_j. However it is assumed that all areas are sampled and that we know the numbers of such units in each small area and the respective small area averages of $\mathbf{x}_j$ and $\mathbf{z}_j$. We also assume that there is a linear relationship between y_j and $\mathbf{x}_j$ and that sampling is non-informative for the small area distribution of y_j given $\mathbf{x}_j$, allowing us to use population level models with the sample data.

A popular way of using the above data in small area estimation is to assume a linear mixed model, with random effects for the small areas of interest (see Rao, 2003). Let $\mathbf{y}$, $\mathbf{X}$ and $\mathbf{Z}$ denote the population level vector and matrices defined by y_j, $\mathbf{x}_j$ and $\mathbf{z}_j$ respectively. Then

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \qquad (1)$$

where $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u)$ is a vector of mq area-specific random effects and $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$ is a vector of N

individual specific random effects. Here m is the number of small areas that make up the population and q is the dimension of $\mathbf{z}_j$. It is assumed that the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_e$ are defined in terms of a lower dimensional set of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$, which are typically referred to as the variance components of (1), while $\boldsymbol{\beta}$ is usually referred to as its fixed effect.

Let $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ denote estimates of the fixed and random effects in (1). The EBLUP of the area i mean of the y_j under (1) is then

$$\hat{\bar{y}}_i^{EBLUP} = N_i^{-1}\left\{n_i \bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T \hat{\boldsymbol{\beta}} + \bar{\mathbf{z}}_{ri}^T \hat{\mathbf{u}}\right)\right\} \tag{2}$$

where $\hat{\mathbf{u}}$ denotes the vector of the estimated area specific random effects and we use indices of s and r to denote sample and non-sample quantities respectively. Thus $\bar{y}_{si}$ is the average of the n_i sample values of y_j from area i and $\bar{\mathbf{x}}_{ri}$ and $\bar{\mathbf{z}}_{ri}$ denoting the vectors of average values of $\mathbf{x}_j$ and $\mathbf{z}_j$ respectively for the $N_i - n_i$ non-sampled units in the same area.

From a Robust Projective viewpoint, (2) can be made insensitive to sample outliers by replacing $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ by outlier robust alternatives. To motivate this approach, we initially assume the variance components $\boldsymbol{\theta}$ are known, so the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_e$ in (1) are known. Put $\mathbf{V}_s = \boldsymbol{\Sigma}_{es} + \mathbf{Z}_s \boldsymbol{\Sigma}_u \mathbf{Z}_s^T$ where $\boldsymbol{\Sigma}_{es}$ denotes the sample component of $\boldsymbol{\Sigma}_e$. Then the BLUE of the fixed effect vector $\boldsymbol{\beta}$ is

$$\tilde{\boldsymbol{\beta}} = \left\{\mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{X}_s\right\}^{-1} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{y}_s, \tag{3}$$

while the BLUP of the random effects vector $\mathbf{u}$ is

$$\tilde{\mathbf{u}} = \Sigma_u \mathbf{Z}_s^T \mathbf{V}_s^{-1}\left(\mathbf{y}_s - \mathbf{X}_s \tilde{\boldsymbol{\beta}}\right). \tag{4}$$

It is easy to see that (3) and (4) are solutions to

$$\mathbf{X}_s^T \mathbf{V}_s^{-1}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}\right) = \mathbf{0} \tag{5}$$

and

$$\boldsymbol{\Sigma}_u \mathbf{Z}_s^T \mathbf{V}_s^{-1}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}\right) - \mathbf{u} = \mathbf{0}. \tag{6}$$

A straightforward way to make the solutions to (5) and (6) robust to sample outliers is therefore to replace them by

$$\mathbf{X}_s^T \mathbf{V}_s^{-1/2} \psi\left(\mathbf{V}_s^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}\right\}\right) = \mathbf{0} \tag{7}$$

and

$$\boldsymbol{\Sigma}_u \mathbf{Z}_s^T \mathbf{V}_s^{-1/2} \psi\left(\mathbf{V}_s^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \beta\right\}\right) - \boldsymbol{\Sigma}_u^{1/2} \psi\left(\boldsymbol{\Sigma}_u^{-1/2} \mathbf{u}\right) = \mathbf{0} . \tag{8}$$

Here ψ is a bounded influence function and $\psi(\mathbf{a})$ denotes the vector defined by applying ψ to every component of $\mathbf{a}$. Unfortunately, since $\mathbf{V}_s$ is not a diagonal matrix, the solution to (8) can be numerically unstable. An alternative approach was therefore suggested by Fellner (1986), who noted that any solution to (5) and (6) was also a solution to

$$\mathbf{X}_s^T \boldsymbol{\Sigma}_{es}^{-1}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u}\right) = \mathbf{0}$$

and

$$\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u}\right) - \boldsymbol{\Sigma}_u^{-1} \mathbf{u} = \mathbf{0} .$$

He suggested that these alternative estimating equations (and hence their solutions) be made outlier robust by replacing them by

$$\mathbf{X}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \psi\left(\boldsymbol{\Sigma}_{es}^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u}\right\}\right) = \mathbf{0} \tag{9}$$

and

$$\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \psi\left(\boldsymbol{\Sigma}_{es}^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u}\right\}\right) - \boldsymbol{\Sigma}_u^{-1/2} \psi\left(\boldsymbol{\Sigma}_u^{-1/2} \mathbf{u}\right) = \mathbf{0} . \tag{10}$$

Since (9) and (10) assume the variance components $\boldsymbol{\theta}$ are known, their usefulness is somewhat limited unless outlier robust estimators of these parameters can also be defined. This is an issue investigated by Richardson and Welsh (1995). These authors propose two outlier robust variations to the maximum likelihood estimating equations for $\boldsymbol{\theta}$. One of these (ML Proposal II) leads to an estimating equation for the variance component θ_k of $\boldsymbol{\theta}$ of the form

$$\psi\left\{\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}\right)^T \mathbf{V}_s^{-1/2}\right\} \mathbf{V}_s^{-1/2}\left(\partial \mathbf{V}_s / \partial \theta_k\right) \mathbf{V}_s^{-1/2} \psi\left\{\mathbf{V}_s^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}\right)\right\} = tr\left\{\mathbf{D}_n^{\psi}\left(\partial \mathbf{V}_s / \partial \theta_k\right)\right\} \tag{11}$$

where $\partial \mathbf{V}_s / \partial \theta_k$ denotes the first order partial derivative of $\mathbf{V}_s$ with respect to the variance component θ_k

and, for $Z \sim N(0,1)$,

$$\mathbf{D}_n^{\psi} = E\left\{\psi^2(Z)\right\}\mathbf{V}_s^{-1}. \tag{12}$$

Sinha and Rao (2009) describe an approach to outlier robust estimation of $\boldsymbol{\beta}$ and $\mathbf{u}$ in (1) that builds on these results, substituting approximate solutions to both (7) and (11) into the Fellner estimating equation (10) to obtain an outlier robust estimate of the area effect $\mathbf{u}$. In particular, their approach replaces (7) by

$$\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\psi\left(\mathbf{U}_s^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}\right\}\right) = \mathbf{0} \tag{13}$$

where $\mathbf{U}_s = diag\left(\mathbf{V}_s\right)$, and replaces (12) by

$$\psi\left\{\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}\right)^T \mathbf{U}_s^{-1/2}\right\}\mathbf{U}_s^{1/2}\mathbf{V}_s^{-1}\left(\partial\mathbf{V}_s/\partial\theta_k\right)\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\psi\left\{\mathbf{U}_s^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}\right)\right\} = tr\left\{\mathbf{D}_n^{\psi}\left(\partial\mathbf{V}_s/\partial\theta_k\right)\right\}. \tag{14}$$

Since the solutions to (13) and (14) depend on the influence function ψ, we denote them by a superscript of ψ below. The Sinha and Rao (2009) Robust Projective alternative to (2) is then

$$\hat{\bar{y}}_i^{SR} = \bar{\mathbf{x}}_i^T\hat{\boldsymbol{\beta}}^{\psi} + \bar{\mathbf{z}}_i^T\hat{\mathbf{u}}^{\psi}\,. \tag{15}$$

Note that (15) estimates the area *i* mean under (1). A minor modification restricts this to the mean of the non-sampled units in area *i*, in which case (15) becomes

$$\hat{\bar{y}}_i^{REBLUP} = N_i^{-1}\left\{n_i\bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T\hat{\boldsymbol{\beta}}^{\psi} + \bar{\mathbf{z}}_{ri}^T\hat{\mathbf{u}}^{\psi}\right)\right\}. \tag{16}$$

Hereafter we call this estimator Robust EBLUP (REBLUP). An alternative methodology for outlier robust small area estimation is the M-quantile regression-based method described by Chambers and Tzavidis (2006). This is based on a linear model for the M-quantile regression of $\mathbf{y}$ on $\mathbf{X}$, i.e.

$$m_q(\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}_q \tag{17}$$

where $m_q(\mathbf{X})$ denotes the M-quantile of order *q* of the conditional distribution of y given $\mathbf{X}$. An estimate $\hat{\boldsymbol{\beta}}_q$ of $\boldsymbol{\beta}_q$ can be calculated for any value of q in the interval (0,1), and for each unit in sample we define its unique M-quantile coefficient under this fitted model as the value q_j such that $y_j = \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{q_j}$, with the sample average of these coefficients in area *i* denoted by $\bar{q}_i$. The M-quantile estimate of the mean of y_j in area i is then

$$\hat{\bar{y}}_i^{MQ} = N_i^{-1}\left\{n_i\bar{y}_{si} + \left(N_i - n_i\right)\bar{\mathbf{x}}_{ri}^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right\}. \tag{18}$$

Note that the regression M-quantile (17) model depends on the influence function ψ underpinning the M-quantile. When this function is bounded, sample outliers have limited impact on $\hat{\boldsymbol{\beta}}_q$. That is, (18) corresponds to assuming that all non-sample units in area i follow the working model (17) with $q = \bar{q}_i$, in the sense that one can write $y_j = \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i} +$ noise for all such units.

3. Robust Predictive Estimation for Small Areas

A problem with the Robust Projective approach is that it assumes all non-sampled units follow the working model, or, in what essentially amounts to the same thing, that any deviations from this model are noise and so cancel out 'on average'. Thus, under the linear mixed model (1) one can see that provided the individual errors of the non-sampled units are symmetrically distributed about zero, the REBLUP (16) of Sinha and Rao (2009) will perform well since it is based on the implicit assumption that the average of these errors over the non-sampled units in area i converges to zero. The M-quantile estimator MQ (18) is no different since it assumes that the errors $y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}$ from the area i-specific M-quantile regression model are 'noise' and hence also cancel out on average. Note that this does not mean that these non-sample units are not outliers. It is just that their behaviour is such that our best prediction of their corresponding average value is zero.

Welsh and Ronchetti (1998) consider the issue of outlier robust prediction within the context of population level survey estimation. Starting with a working linear model linking the population values of y_j and $\mathbf{x}_j$, and sample data containing representative outliers with respect to this model, they extend the approach of Chambers (1986) to robust prediction of the empirical distribution function of the population values of y_j. Their argument immediately applies to robust prediction of the empirical distribution function of the area i values of y_j, and leads to a predictor of the form

$$\hat{F}_i^{\psi\phi}(t) = N_i^{-1}\left[\sum_{j\in s_i} I(y_j \le t) + n_i^{-1}\sum_{j\in s_i}\sum_{k\in r_i} I\left(\mathbf{x}_k^T\hat{\boldsymbol{\beta}}^{\psi} + \omega_{ij}^{\psi}\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi}\right)\Big/\omega_{ij}^{\psi}\right\} \le t\right)\right]. \tag{19}$$

Here $\hat{\boldsymbol{\beta}}^{\psi}$ denotes an M-estimator of the regression parameter in the linear working model based on a bounded influence function ψ, ω_{ij}^{ψ} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}^{\psi}$ in area *i* and ϕ denotes a bounded influence function that satisfies $|\phi| \geq |\psi|$. Tzavidis *et al*. (2009) note that the robust estimator of the area *i* mean of the y_j defined by (19) is just the expected value functional defined by it, which is

$$\hat{\bar{y}}_i^{\psi\phi} = \int t d\hat{F}_i^{\psi\phi}(t) = N_i^{-1}\left[n_i \bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T \hat{\boldsymbol{\beta}}^{\psi} + n_i^{-1} \sum_{j \in s_i} \omega_{ij}^{\psi} \phi\left\{ \left(y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}^{\psi} \right) \Big/ \omega_{ij}^{\psi} \right\} \right) \right]. \quad (20)$$

These authors therefore suggest an extension to the M-quantile estimator (18) by replacing $\hat{\boldsymbol{\beta}}^{\psi}$ in (20) by $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$, which leads to a 'bias-corrected' version of (18), hereafter MQ-BC, given by

$$\hat{\bar{y}}_i^{MQ-BC} = N_i^{-1}\left[n_i \bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T \hat{\boldsymbol{\beta}}_{\bar{q}_i} + n_i^{-1} \sum_{j \in s_i} \omega_{ij}^{MQ} \phi\left\{ \left(y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}_{\bar{q}_i} \right) \Big/ \omega_{ij}^{MQ} \right\} \right) \right] \quad (21)$$

and ω_{ij}^{MQ} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}_{\bar{q}_i}$ in area *i*.

The use of the two influence functions ψ and ϕ in (21) is worthy of comment. The first, ψ, underpins $\hat{\boldsymbol{\beta}}_q$, and hence $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$. Its purpose is to ensure that sample outliers have little or no influence on the fit of the working M-quantile model. As a consequence it is bounded and down-weights these outliers. The second, ϕ, is still bounded but 'less restrictive' than ψ (since $|\phi| \geq |\psi|$) and its purpose is to define an adjustment for the bias caused by the fact that the first two terms on the right hand side of (21) treat sample outliers as self-representing. A similar argument can be used to modify the REBLUP (16). In particular, a Robust Predictive version of this estimator, hereafter REBLUP-BC, mimics the bias correction idea used in (21) and leads to

$$\hat{\bar{y}}_i^{REBLUP-BC} = \hat{\bar{y}}_i^{REBLUP} + \left(1 - n_i N_i^{-1}\right) n_i^{-1} \sum_{j \in s_i} \omega_{ij}^{\psi} \phi\left\{ \left(y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T \hat{\mathbf{u}}^{\psi} \right) \Big/ \omega_{ij}^{\psi} \right\}, \quad (22)$$

where the ω_{ij}^{ψ} are now robust estimates of the scale of the area *i* residuals $y_j - \mathbf{x}_j^T \hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T \hat{\mathbf{u}}^{\psi}$.

4. MSE Estimation for Robust Predictors

In this Section we propose two different MSE estimators for robust predictors of small area means under the Robust Projective and Robust Predictive approaches. In Section 4.1 we apply the ideas set out by Chambers *et al*. (2007) to develop a pseudo-linearization estimator of the MSE of REBLUP and REBLUP-BC. In Section 4.2 we use first order approximations to the variances of solutions of estimating equations to develop MSE estimators, under the conditional version of the linear mixed model, for the REBLUP, EBLUP and MQ predictors for small area means.

4.1 Bias-robust MSE estimation for REBLUP and REBLUP-BC

Sinha and Rao (2009) proposed a computationally intensive parametric bootstrap-based estimator for the MSE of REBLUP. An alternative MSE is the one that conditions on the realised values of the area effects (see Longford, 2007). In what follows we propose an estimator of the conditional MSE of the REBLUP and REBLUP-BC that is much less computationally demanding than the unconditional MSE estimators suggested by Sinha and Rao (2009). The proposed estimator is based on the pseudo-linearization approach to MSE estimation described by Chambers *et al*. (2007). See also Chandra and Chambers (2005, 2009) and Chandra *et al*. (2007). The MSE estimator can be used for predictors that can be expressed as weighted sums of the sample values. For this reason re-express REBLUP (16) and REBLUP-BC (22) in a pseudo-linear form, and then apply heteroskedasticity-robust prediction variance estimation methods that treat these weights (which typically depend on estimated variance components) as fixed. More precisely, under model (1) the Robust BLUP of $\bar{y}_i$ can be expressed as

$$\hat{\bar{y}}_i^{RBLUP} = \sum_{j \in s} w_{ij}^{RBLUP} y_j = \left(\mathbf{w}_{is}^{RBLUP}\right)^T \mathbf{y}_s,\ i \in 1 \ldots m \qquad (23)$$

where

$$\left(\mathbf{w}_{is}^{RBLUP}\right)^T = \frac{1}{N_i}\left\{\mathbf{1}_s^T + \left(N_i - n_i\right)\left[\bar{\mathbf{x}}_{ir}^T \mathbf{A}_s + \bar{\mathbf{z}}_{ir}^T \mathbf{B}_s \left(\mathbf{I}_s - \mathbf{X}_s \mathbf{A}_s\right)\right]\right\}.$$

Here

- $\mathbf{A}_s = \left(\mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \mathbf{W}_{1s} \mathbf{U}_s^{-1/2} \mathbf{X}_s\right)^{-1} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \mathbf{W}_{1s} \mathbf{U}_s^{-1/2}$, where $\mathbf{W}_{1s}$ is a $n \times n$ diagonal matrix of weights with j-th component $w_{1j} = \psi\left(U_j^{-1/2}\left\{y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^{\psi}\right\}\right) \Big/ U_j^{-1/2}\left\{y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^{\psi}\right\}$;

- $\mathbf{B}_s = \left(\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{W}_{2s} \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{Z}_s + \boldsymbol{\Sigma}_u^{-1/2} \mathbf{W}_{3s} \boldsymbol{\Sigma}_u^{-1/2}\right)^{-1} \left(\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{W}_{2s} \boldsymbol{\Sigma}_{es}^{-1/2}\right)$, where $\mathbf{W}_{2s}$ is a $n \times n$ diagonal matrix of weights with j-th component $w_{2j} = \psi\left(\left(\sigma_e^{\psi}\right)^{-1}\left\{y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j \tilde{\mathbf{u}}^{\psi}\right\}\right) \Big/ \left(\sigma_e^{\psi}\right)^{-1}\left\{y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j \tilde{\mathbf{u}}^{\psi}\right\}$; and
- $\mathbf{W}_{3s}$ is a $m \times m$ diagonal matrix of weights with i-th component $w_{3i} = \psi\left(\left(\sigma_u^{\psi}\right)^{-1} \tilde{u}_i^{\psi}\right) \Big/ \left(\sigma_u^{\psi}\right)^{-1} \tilde{u}_i^{\psi}$.

The Appendix provides details on the computation of such weights. Note that the REBLUP (16) can be expressed in exactly the same way, except that all quantities in the vector $\mathbf{w}_{is}^{RBLUP}$ that depend on (unknown) variance components now need a 'hat'. Given this pseudo-linear representation for the REBLUP, we develop a simple first order approximation to its MSE assuming the conditional version of the model (1), i.e. the random effects are considered as fixed. In this case we can apply the approach described by Royall and Cumberland (1978) to estimate the prediction variance of the RBLUP for $\bar{y}_i$. Let $I(j \in i)$ denote the indicator for whether unit j is in area i. Then

$$Var\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i \,\middle|\, \mathbf{X}, \mathbf{u}^{\psi}\right) = N_i^{-2}\left\{\sum_{j \in s}\left(N_i w_{ij}^{RBLUP} - I(j \in i)\right)^2 Var\left(y_j \mid \mathbf{x}_j, \mathbf{u}^{\psi}\right) + \sum_{j \in r_i} Var\left(y_j \mid \mathbf{x}_j, \mathbf{u}^{\psi}\right)\right\}, \tag{24}$$

where the first term on the right hand side above is estimated replacing $Var\left(y_j \mid \mathbf{x}_j, \mathbf{u}^{\psi}\right)$ by $\lambda_j^{-1}(y_j - \hat{\mu}_j)^2$, where $\hat{\mu}_j = \sum_{k \in s} \phi_{kj} y_k$ is an unbiased linear estimator of the conditional expected value $\mu_j = E\left(y_j \mid \mathbf{x}_j, \mathbf{u}^{\psi}\right)$ and $\lambda_j = \left\{1 - 2\phi_{jj} + \sum_{k \in s} \phi_{kj}^2\right\}$ is a scaling constant. Further details can be found in Chambers *et al*. (2007) and Salvati *et al*. (2009). The conditional prediction variance of the RBLUP is

$$\hat{V}(\hat{\bar{y}}_i^{RBLUP}) = N_i^{-2} \sum_{j \in s}\left\{a_{ij}^2 + (N_i - n_i)n^{-1}\right\} \lambda_j^{-1}(y_j - \hat{\mu}_j)^2, \tag{25}$$

where $a_{ij} = N_i w_{ij}^{RBLUP} - I(j \in i)$. Due to the well-known shrinkage effect associated with BLUPs, replacing $\hat{\mu}_j$ by the BLUP of μ_j under (1) in expression (25) can lead to biased estimation of the prediction variance under the conditional model. For this reason, Chambers *et al*. (2007) recommend that $\hat{\mu}_j$ be computed as the 'unshrunken' version of the BLUP for μ_j:

$$\hat{\mu}_j = \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^{\psi} + \mathbf{z}_j^T \mathbf{B}_s \tilde{\mathbf{u}}^{\psi} . \tag{26}$$

The conditional bias of the RBLUP under (1) is given by

$$E\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i \middle| \mathbf{X}, \mathbf{u}^{\psi}\right) = \sum_{j\in s} w_{ij}^{RBLUP} \mu_j - N_i^{-1} \sum_{j\in(r_i \cup s_i)} \mu_j ,$$

which has the simple 'plug-in' estimator

$$\hat{B}(\hat{\bar{y}}_i^{RBLUP}) = \sum_{j\in s} w_{ij}^{RBLUP} \hat{\mu}_j - N_i^{-1} \sum_{j\in(r_i \cup s_i)} \hat{\mu}_j , \tag{27}$$

with $\hat{\mu}_j$ defined by expression (26). The estimator of the conditional MSE of the RBLUP can finally be written as

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP}\right) = \hat{V}\left(\hat{\bar{y}}_i^{RBLUP}\right) + \left\{\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)\right\}^2 . \tag{28}$$

The conditional MSE of the REBLUP (16) is then estimated by replacing all unknown variance components in (28) by their estimated values. Note that: (a) $\hat{\lambda}_j = 1 + O(n^{-1})$ in this case, so that $\hat{\lambda}_j$ will be very close to one in most practical applications. This suggests that there is little to be gained by not setting $\hat{\lambda}_j \equiv 1$ when calculating the conditional prediction variance (25); (b) the square of the bias estimator (27) can be biased for the squared bias term in the MSE estimator. This bias can be corrected (see Chambers *et al.*, 2007), but a small sample size could lead to this correction becoming unstable, so we prefer use (28) since this is then a conservative estimator of the MSE of the predictor of the small area mean under model (1); (c) the heteroskedasticity-robust MSE estimator (28) ignores the extra variability associated with estimation of the variance components, and is therefore a first order approximation to the actual conditional MSE of the REBLUP. Since use of the REBLUP will typically require a large overall sample size, we expect any consequent underestimation of the conditional MSE of the REBLUP to be small.

The conditional MSE estimator for the REBLUP-BC (22) is obtained using the same heteroskedasticity-robust pseudo-linearization approach as outlined above for the MSE estimator for the REBLUP. The only difference from that development is that the weights w_{ij}^{RBLUP} used in (23) are now replaced by corresponding REBLUP-BC weights

$$\left(\mathbf{w}_{is}^{RBLUP-BC}\right)^T = \frac{1}{N_i}\left\{\left(1+\frac{N_i-n_i}{n_i}w_j^{\phi}\right)\mathbf{1}_s^T + \left\{\sum_{j\in r_i}\mathbf{x}_j^T - \frac{N_i-n_i}{n_i}\sum_{j\in s_i}\mathbf{x}_j^T w_j^{\phi}\right\}\mathbf{A}_s + \right.$$
$$\left. + \left\{\left(N_i-n_i\right)-\frac{N_i-n_i}{n_i}\sum_{j\in s_i}w_j^{\phi}\right\}\bar{\mathbf{z}}_i^T\mathbf{B}_s\left\{\mathbf{I}_s - \mathbf{X}_s\mathbf{A}_s\right\}\right\}, \tag{29}$$

where $w_j^{\phi} = \dfrac{\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\hat{\mathbf{u}}^{\psi}\right)\big/\omega_{ij}^{\psi}\right\}}{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\hat{\mathbf{u}}^{\psi}\right)\big/\omega_{ij}^{\psi}}$. Since the REBLUP-BC is an approximately unbiased estimator of the small area mean, the squared bias term does not impact significantly on the mean squared error estimator, and so is typically omitted.

4.2 Linearization-Based MSE estimation for small area predictors

In this Section we propose a new MSE estimator, extending the linearization approach of Street *et al*. (1988) to estimation of prediction variance for estimators based on robust estimating equations. The MSE estimator is developed on the assumption that the working model for inference is an area-specific linear model, and so the approach conditions on area effects when applied in the context of such a model. In what follows we show how this approach can be used for estimating the MSE of the REBLUP (16), the EBLUP (2) and the MQ estimator (18). The MSE estimators of REBLUP-BC and MQ-BC are reported in the Appendix. Note that when used with an estimator based on a mixed model, the proposed MSE estimator provides a second order approximation to the true MSE since it includes a term for the contribution to variability from estimation of variance components.

MSE estimation for REBLUP

Under model (1) the prediction variance of the Robust BLUP of $\bar{y}_i$ can be expressed as

$$\begin{aligned} Var\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i\right) &= Var\left\{\frac{1}{N_i}\sum_{j\in r_i}\left(\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} + \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right) - \frac{1}{N_i}\sum_{j\in r_i}y_j\right\} \\ &= \left(1-\frac{n_i}{N_i}\right)^2\bar{\mathbf{x}}_{ri}^T Var\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\bar{\mathbf{x}}_{ri} + \left(1-\frac{n_i}{N_i}\right)^2\bar{\mathbf{z}}_{ri}^T Var\left(\tilde{\mathbf{u}}^{\psi}\right)\bar{\mathbf{z}}_{ri} + \left(1-\frac{n_i}{N_i}\right)^2 Var\left(\bar{e}_{ri}\right), \end{aligned} \tag{30}$$

assuming independence between $\tilde{\boldsymbol{\beta}}^{\psi}$ and $\tilde{\mathbf{u}}^{\psi}$. It follows that we need to estimate $Var(\tilde{\boldsymbol{\beta}}^{\psi})$, $Var(\tilde{\mathbf{u}}^{\psi})$ in order to be able to calculate an estimate of the prediction variance of the RBLUP. In order to do this, put $\boldsymbol{\delta} = \left(\boldsymbol{\beta}^{\psi T}, \mathbf{u}^{\psi T}\right)^T$, so $\tilde{\boldsymbol{\delta}} = \left(\tilde{\boldsymbol{\beta}}^{\psi T}, \tilde{\mathbf{u}}^{\psi T}\right)^T$. Then, from equations (10) and (13), $\mathbf{H}(\tilde{\boldsymbol{\delta}}) = \mathbf{0}$ where

$$\mathbf{H}(\boldsymbol{\delta}) = \begin{pmatrix} \mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}) \\ \mathbf{H}_{u^\psi}(\boldsymbol{\delta}) \end{pmatrix} = \begin{pmatrix} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \psi\left(\mathbf{U}_s^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}^\psi\right\}\right) = \mathbf{0} \\ \mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \psi\left(\boldsymbol{\Sigma}_{es}^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}^\psi - \mathbf{Z}_s \mathbf{u}^\psi\right\}\right) - \boldsymbol{\Sigma}_u^{-1/2} \psi\left(\boldsymbol{\Sigma}_u^{-1/2} \mathbf{u}^\psi\right) = \mathbf{0} \end{pmatrix}.$$

Since the solutions of the equations depend on the influence function ψ, we denote them by a superscript of ψ. We can use previous results on the asymptotic variance of solutions to an estimating equation (Welsh and Richardson, 1997; Sinha and Rao, 2009) to obtain a first order approximation to $Var\left(\tilde{\boldsymbol{\delta}}\right)$ and by extension the prediction variance of the RBLUP. To do this, we note that $Var_0(\tilde{\boldsymbol{\delta}}) \approx \left\{E_0\left(\partial_{\boldsymbol{\delta}} \mathbf{H}_0\right)\right\}^{-1} Var_0\left\{\mathbf{H}(\boldsymbol{\delta}_0)\right\}\left[\left\{E_0\left(\partial_{\boldsymbol{\delta}} \mathbf{H}_0\right)\right\}^{-1}\right]^T$ where

$$\begin{aligned} Var_0\left(\mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right) &= Var_0\left[\psi\left\{U_j^{-1/2}\left(y_j - \mathbf{x}_j^T \boldsymbol{\beta}_0^\psi\right)\right\}\right] \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s \mathbf{V}_s^{-1} \mathbf{X}_s \\ Var_0\left(\mathbf{H}_{u^\psi}(\boldsymbol{\delta}_0)\right) &= E_0\left[\psi^2\left\{\left(\sigma_e^\psi\right)^{-1}\left(y_j - \mathbf{x}_j^T \boldsymbol{\beta}_0^\psi - \mathbf{z}_j^T \mathbf{u}_0^\psi\right)\right\}\right] \mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1} \mathbf{Z}_s, \end{aligned} \tag{31}$$

and

$$\begin{aligned} E_0\left\{\partial_{\beta^\psi} \mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\} &= -\mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} E_0\left[\psi'\left\{\mathbf{U}_s^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}_0^\psi\right)\right\}\right] \mathbf{U}_s^{-1/2} \mathbf{X}_s \\ E_0\left\{\partial_{u^\psi} \mathbf{H}_{u^\psi}(\boldsymbol{\theta}_0)\right\} &= -\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} E_0\left[\psi'\left\{\boldsymbol{\Sigma}_{es}^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta}_0^\psi - \mathbf{Z}_s \mathbf{u}_0^\psi\right)\right\}\right] \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{Z}_s - \boldsymbol{\Sigma}_u^{-1/2} E\left[\psi'\left\{\boldsymbol{\Sigma}_u^{-1/2} \mathbf{u}_0^\psi\right\}\right] \boldsymbol{\Sigma}_u^{-1/2}. \end{aligned} \tag{32}$$

The previous expressions lead to the estimator:

$$\begin{aligned} \widehat{Var}\left(\tilde{\boldsymbol{\beta}}^\psi\right) &= \left\{\hat{E}\left(\partial_{\beta^\psi} \mathbf{H}_0\right)\right\}^{-1} \widehat{Var}\left\{\mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\}\left[\left\{\hat{E}\left(\partial_{\beta^\psi} \mathbf{H}_0\right)\right\}^{-1}\right]^T \\ \widehat{Var}\left(\tilde{\mathbf{u}}^\psi\right) &= \left\{\hat{E}\left(\partial_{u^\psi} \mathbf{H}_0\right)\right\}^{-1} \widehat{Var}\left\{\mathbf{H}_{u^\psi}(\boldsymbol{\delta}_0)\right\}\left[\left\{\hat{E}\left(\partial_{u^\psi} \mathbf{H}_0\right)\right\}^{-1}\right]^T, \end{aligned} \tag{33}$$

where

- $\hat{E}\left\{\partial_{\beta^\psi} \mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\} = -\mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \mathbf{R} \mathbf{U}_s^{-1/2} \mathbf{X}_S$,
- $\hat{E}\left\{\partial_{u^\psi} \mathbf{H}_{u^\psi}(\boldsymbol{\delta}_0)\right\} = -\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{T} \boldsymbol{\Sigma}_{es}^{-1/2} \mathbf{Z}_s - \boldsymbol{\Sigma}_u^{-1/2} \mathbf{Q} \boldsymbol{\Sigma}_u^{-1/2}$,
- $\widehat{Var}\left\{\mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\} = (n-p)^{-1} \lambda_1 \sum_{j=1}^n \psi^2\left(r_j\right) \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s \mathbf{V}_s^{-1} \mathbf{X}_s$, and
- $\widehat{Var}\left\{\mathbf{H}_{u^\psi}(\boldsymbol{\delta}_0)\right\} = (n-p)^{-1} \lambda_2 \sum_{j=1}^n \psi^2\left(t_j\right) \mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1} \mathbf{Z}_s$.

Here, assuming use of a Huber Proposal 2 influence function, $\mathbf{R}$ is a $n \times n$ diagonal matrix with *j*-th diagonal element is 1 if $-c < r_j < c$, 0 otherwise, with $r_j = U_j^{-1/2}\left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}^\psi\right)$; the constant *c* represents the cut-off of the bounded influence function; $\mathbf{T}$ is a diagonal matrix of dimension $n \times n$ with *j*-th element diagonal element equal to 1 if $-c < t_j < c$, 0 otherwise, with $t_j = \left(\sigma_e^{\psi}\right)^{-1}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}\right)$; $\mathbf{Q}$ is a $m \times m$ diagonal matrix with *i*-th diagonal element equal to 1 if $-c < q_i < c$, 0 otherwise, with $q_i = \left(\sigma_u^{2\psi}\right)^{-1/2} u_i^{\psi}$. The values $\lambda_1 = \left\{1 + \frac{p}{n} Var\left(\psi'\left(r_j\right)\right)\left[E\left(\psi'\left(r_j\right)\right)\right]^{-2}\right\}$ and $\lambda_2 = \left\{1 + \frac{p}{n} Var\left(\psi'\left(\tilde{t}_i\right)\right)\left[E\left(\psi'\left(\tilde{t}_i\right)\right)\right]^{-2}\right\}$ are bias corrector terms (Huber, 1981).

An estimator of the prediction variance of RBLUP can be written as:

$$\hat{V}\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i\right) = h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) \tag{34}$$

where $h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \bar{\mathbf{z}}_{ri}^T \hat{V}\left(\tilde{\mathbf{u}}^{\psi}\right)\bar{\mathbf{z}}_{ri}$ is due to the estimation of random effects, while the second term $h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \bar{\mathbf{x}}_{ri}^T \hat{V}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\bar{\mathbf{x}}_{ri}$ is due to the estimator $\tilde{\boldsymbol{\beta}}^{\psi}$. The term $h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \hat{V}\left(\bar{e}_{ri}\right)$ can be estimated from the area *i* data: $\hat{V}(\bar{e}_{ri}) = \frac{1}{(N_i - n_i)(n_i - 1)} \sum_{j \in s_i}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)^2$, or from the entire data set: $\hat{V}(\bar{e}_{ri}) = \frac{1}{(N_i - n_i)(n-1)} \sum_h \sum_{j \in s_h}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)$. Moreover, since we are working under the conditional approach, we have to add to the variance estimator (34) an estimator of the squared bias term. The result is that the estimator of the conditional MSE of the RBLUP can be written as:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP}\right) = h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) + \left\{\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)\right\}^2, \tag{35}$$

where the $\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)$ is the expression (27) developed in the previous Section. The corresponding estimator of the conditional MSE of the REBLUP (16) is obtained by adding an extra component to expression (35) due to the variability of the estimated variance components:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right) = \widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP}\right) + E\left[\left(\hat{\bar{y}}_i^{REBLUP} - \hat{\bar{y}}_i^{RBLUP}\right)^2\right]. \tag{36}$$

The last term is intractable and it is therefore necessary to approximate it. An approximation of this term is obtained by Taylor approximation following the results of Prasad and Rao (1990). Under the conditional approach

$$\hat{\bar{y}}_i^{REBLUP} - \hat{\bar{y}}_i^{RBLUP} \approx \frac{1}{N_i} \sum_{j \in r_i} \mathbf{z}_j^T \sum_{k=1}^{2}\left(\partial_{\theta_k}\mathbf{B}_s\right)\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)\left\{\hat{\theta}_k - \theta_k\right\},$$

where $\mathbf{B}_s$ is defined as in previous Section, and $\boldsymbol{\theta}=\left(\sigma_u^{2\psi},\sigma_e^{2\psi}\right)$ is the vector of the variance components. Assuming that the derivative of $\left(\tilde{\boldsymbol{\beta}}^{\psi}-\boldsymbol{\beta}^{\psi}\right)$ with respect to $\boldsymbol{\theta}$ is of lower order, the term $E\left[\left(\hat{\bar{y}}_i^{REBLUP}-\hat{\bar{y}}_i^{RBLUP}\right)^2\right]$ in (36) is then estimated by

$$h_{4i}\left(\tilde{\boldsymbol{\delta}}\right)=\left(\frac{1}{N_i}\sum_{j\in r_i}\mathbf{z}_j^T\right)\Upsilon Var(\hat{\sigma}_u^{\psi 2},\hat{\sigma}_e^{\psi 2})\left(\frac{1}{N_i}\sum_{j\in r_i}\mathbf{z}_j^T\right)^T+o\left(m^{-1}\right) \tag{37}$$

where

$$\Upsilon=\sum_{k=1}^{2}\left\{\left(\partial_{\sigma_u^{\psi 2},\sigma_e^{\psi 2}}\mathbf{B}_s\right)\left[\sum_j\sum_l\left\{\left(\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)\left(\mathbf{z}_l^T\tilde{\mathbf{u}}^{\psi}\right)+\sigma_e^{\psi 2}\mathbf{I}(j=l)\right\}\right]\left(\partial_{\sigma_u^{\psi 2},\sigma_e^{\psi 2}}\mathbf{B}_s\right)\right\}.$$

Note that $Var(\hat{\sigma}_u^{\psi 2},\hat{\sigma}_e^{\psi 2})$ in (37) is obtained using the results of the asymptotic distribution of $\left(\sigma_u^{2\psi},\sigma_e^{2\psi}\right)$ given in Sinha and Rao (2009). The MSE estimator of the REBLUP (16) then becomes:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}\left(\tilde{\boldsymbol{\delta}}\right)+h_{2i}\left(\tilde{\boldsymbol{\delta}}\right)+h_{3i}\left(\tilde{\boldsymbol{\delta}}\right)+\left\{\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)\right\}^2+h_{4i}\left(\tilde{\boldsymbol{\delta}}\right). \tag{38}$$

An estimator of $\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right)$ can be obtained by replacing all unknown variance components $\boldsymbol{\theta}$ in (38) by their estimated values $\hat{\boldsymbol{\theta}}$. This corresponds to substituting $\tilde{\boldsymbol{\delta}}=\left(\tilde{\boldsymbol{\beta}}^{\psi T},\tilde{\mathbf{u}}^{\psi T}\right)^T$ by $\hat{\boldsymbol{\delta}}=\left(\hat{\boldsymbol{\beta}}^{\psi T},\hat{\mathbf{u}}^{\psi T}\right)^T$ in the MSE approximation (38) and leads to:

$$mse\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}\left(\hat{\boldsymbol{\delta}}\right)+h_{2i}\left(\hat{\boldsymbol{\delta}}\right)+h_{3i}\left(\hat{\boldsymbol{\delta}}\right)+\left\{\hat{B}\left(\hat{\bar{y}}_i^{REBLUP}\right)\right\}^2+h_{4i}\left(\hat{\boldsymbol{\delta}}\right). \tag{39}$$

We have $E\left[h_{2i}\left(\hat{\boldsymbol{\delta}}\right)\right]=h_{2i}\left(\tilde{\boldsymbol{\delta}}\right)+o\left(m^{-1}\right)$, $E\left[h_{3i}\left(\hat{\boldsymbol{\delta}}\right)\right]=h_{3i}\left(\tilde{\boldsymbol{\delta}}\right)+o\left(m^{-1}\right)$, $E\left[h_{4i}\left(\hat{\boldsymbol{\delta}}\right)\right]=h_{4i}\left(\tilde{\boldsymbol{\delta}}\right)+o\left(m^{-1}\right)$ to the desired order of approximation. However, $h_{1i}\left(\hat{\boldsymbol{\delta}}\right)$ is not the correct estimator of $h_{1i}\left(\tilde{\boldsymbol{\delta}}\right)$ because its bias is generally of the same order as $h_{2i}\left(\hat{\boldsymbol{\delta}}\right)$, $h_{3i}\left(\hat{\boldsymbol{\delta}}\right)$, $h_{4i}\left(\hat{\boldsymbol{\delta}}\right)$. To evaluate the bias of $h_{1i}\left(\hat{\boldsymbol{\delta}}\right)$, we use a Taylor series expansion of $h_{1i}\left(\hat{\boldsymbol{\delta}}\right)$ around $\boldsymbol{\theta}=\left(\sigma_u^{\psi 2},\sigma_e^{\psi 2}\right)$:

$$\begin{aligned}h_{1i}\left(\hat{\delta}\right)&=h_{1i}\left(\tilde{\delta}\right)+\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\nabla h_{1i}\left(\delta\right)+\frac{1}{2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\nabla^2 h_{1i}\left(\delta\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\\&=h_{1i}\left(\delta\right)+\Delta_1+\Delta_2.\end{aligned}$$

If $\hat{\boldsymbol{\theta}}$ is unbiased for $\boldsymbol{\theta}$ then $E\left[\Delta_1\right]=0$. In general, if $\hat{\boldsymbol{\theta}}$ is biased, $E\left[\Delta_1\right]$ is of lower order than $E\left[\Delta_2\right]$, so

$$E\left[h_{1i}\left(\hat{\delta}\right)\right] \approx h_{1i}\left(\delta\right)+\frac{1}{2}tr\left\{\nabla^2 h_{1i}\left(\delta\right)E\left[\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\right]\right\}$$

$$=h_{1i}\left(\delta\right)+\frac{1}{2}\left(1-\frac{n_i}{N_i}\right)^2 tr\left\{\bar{\mathbf{z}}_{ri}^T\nabla^2 h_{1i}\left(\delta\right)\bar{\mathbf{z}}_i Var(\hat{\sigma}_u^{2\psi},\hat{\sigma}_e^{2\psi})\right\}+o\left(m^{-1}\right).$$

We denote the second term on the right hand side above by $h_{5i}\left(\hat{\boldsymbol{\delta}}\right)$. The estimator of the MSE of $\hat{\bar{y}}_i^{REBLUP}$ is then:

$$mse\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}\left(\hat{\boldsymbol{\delta}}\right)+h_{2i}\left(\hat{\boldsymbol{\delta}}\right)+h_{3i}\left(\hat{\boldsymbol{\delta}}\right)+\left\{\hat{B}\left(\hat{\bar{y}}_i^{REBLUP}\right)\right\}^2+h_{4i}\left(\hat{\boldsymbol{\delta}}\right)+h_{5i}\left(\hat{\boldsymbol{\delta}}\right) \tag{40}$$

and $E\left[mse\left(\hat{\bar{y}}_i^{REBLUP}\right)\right]=\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right)+o\left(m^{-1}\right)$.

MSE estimation for EBLUP

The second predictor of $\bar{y}_i$ that we consider is the well-known EBLUP based on (1). Note that EBLUP is a particular case of REBLUP when the bounded influence function is replaced by the (unbounded) identity function. Under (1) the prediction variance of the BLUP of $\bar{y}_i$ is

$$Var\left(\hat{\bar{y}}_i^{RBLUP}-\bar{y}_i\right)=\left(1-\frac{n_i}{N_i}\right)^2\bar{\mathbf{x}}_{ri}^T Var\left(\tilde{\boldsymbol{\beta}}\right)\bar{\mathbf{x}}_{ri}+\left(1-\frac{n_i}{N_i}\right)^2\bar{\mathbf{z}}_{ri}^T Var\left(\tilde{\mathbf{u}}\right)\bar{\mathbf{z}}_{ri}+\left(1-\frac{n_i}{N_i}\right)^2 Var\left(\bar{e}_{ri}\right) \tag{41}$$

assuming independence between $\tilde{\boldsymbol{\beta}}$ and $\tilde{\mathbf{u}}$. Putting $\boldsymbol{\delta}=\left(\boldsymbol{\beta}^T,\mathbf{u}^T\right)^T$, so $\tilde{\boldsymbol{\delta}}=\left(\tilde{\boldsymbol{\beta}}^T,\tilde{\mathbf{u}}^T\right)^T$ and using results on the asymptotic variance of solutions to estimating equations (Richardson and Welsh, 1997),

$$\mathbf{H}(\boldsymbol{\delta})=\begin{pmatrix}\mathbf{H}_\beta(\boldsymbol{\delta})\\ \mathbf{H}_u(\boldsymbol{\delta})\end{pmatrix}=\begin{pmatrix}\mathbf{X}_s^T\boldsymbol{\Sigma}_{es}^{-1}\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}-\mathbf{Z}_s\mathbf{u}\right)=\mathbf{0}\\ \mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1}\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}-\mathbf{Z}_s\mathbf{u}\right)-\boldsymbol{\Sigma}_u^{-1}\mathbf{u}=\mathbf{0}\end{pmatrix},$$

we obtain first order approximation to $Var\left(\tilde{\boldsymbol{\delta}}\right)$ and by extension the prediction variance of the BLUP. The starting point is $Var_0(\tilde{\boldsymbol{\delta}})\approx\left\{E_0\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right)\right\}^{-1}Var_0\left\{\mathbf{H}(\boldsymbol{\delta}_0)\right\}\left[\left\{E_0\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right)\right\}^{-1}\right]^T$, which leads to the estimators:

$$\begin{aligned}\widehat{Var}\left(\tilde{\boldsymbol{\beta}}\right)&=\left\{\hat{E}\left(\partial_\beta\mathbf{H}_0\right)\right\}^{-1}\widehat{Var}\left\{\mathbf{H}_\beta(\boldsymbol{\delta}_0)\right\}\left[\left\{\hat{E}\left(\partial_\beta\mathbf{H}_0\right)\right\}^{-1}\right]^T\\ \widehat{Var}\left(\tilde{\mathbf{u}}\right)&=\left\{\hat{E}\left(\partial_u\mathbf{H}_0\right)\right\}^{-1}\widehat{Var}\left\{\mathbf{H}_u(\boldsymbol{\delta}_0)\right\}\left[\left\{\hat{E}\left(\partial_u\mathbf{H}_0\right)\right\}^{-1}\right]^T\end{aligned} \tag{42}$$

where

- $\hat{E}\left\{\partial_\beta\mathbf{H}_\beta(\boldsymbol{\delta}_0)\right\}=-\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{X}_s$,

- $\hat{E}\left\{\partial_u \mathbf{H}_u(\boldsymbol{\delta}_0)\right\} = -\mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1} \mathbf{Z}_s - \boldsymbol{\Sigma}_u^{-1}$,
- $\widehat{Var}\left\{\mathbf{H}_\beta(\boldsymbol{\delta}_0)\right\} = \left\{\left(n-p\right)^{-1} \sum_{j=1}^{n} \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}}\right)^2\right\} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{V}_s^{-1} \mathbf{X}_s$, and
- $\widehat{Var}\left\{\mathbf{H}_u(\boldsymbol{\delta}_0)\right\} = \left\{\left(n-p\right)^{-1} \sum_{j=1}^{n} \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}} - \mathbf{z}_j^T \mathbf{u}\right)^2\right\} \mathbf{Z}_s^T \boldsymbol{\Sigma}_{es}^{-1} \boldsymbol{\Sigma}_{es}^{-1} \mathbf{Z}_s$.

An estimator of the MSE for the BLUP can therefore be written as:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{BLUP}\right) = h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) \qquad (43)$$

where $h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \bar{\mathbf{z}}_{ri}^T \hat{V}\left(\tilde{\mathbf{u}}\right) \bar{\mathbf{z}}_{ri}$ is due to the estimation of random effects, while the second term $h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \bar{\mathbf{x}}_{ri}^T \hat{V}\left(\tilde{\boldsymbol{\beta}}\right) \bar{\mathbf{x}}_{ri}$ is due to $\tilde{\boldsymbol{\beta}}$. The term $h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) = \left(1 - \frac{n_i}{N_i}\right)^2 \hat{V}\left(\bar{e}_{ri}\right)$ can be estimated just using area i data, $\hat{V}(\bar{e}_{ri}) = \frac{1}{(N_i - n_i)(n_i - 1)} \sum_{j \in s_i} \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}} - \mathbf{z}_j^T \tilde{\mathbf{u}}\right)^2$, or by using all the sample data, $\hat{V}(\bar{e}_{ri}) = \frac{1}{(N_i - n_i)(n-1)} \sum_h \sum_{j \in s_h} \left(y_j - \mathbf{x}_j^T \tilde{\boldsymbol{\beta}} - \mathbf{z}_j^T \tilde{\mathbf{u}}\right)$. Note that we have not added the squared bias estimator to (43) – as we did in the REBLUP case – because this bias is zero (see Chambers *et al*., 2007). In order to define the conditional MSE of the EBLUP, we add the term $h_{4i}\left(\tilde{\boldsymbol{\delta}}\right)$, see equation (37), to (43). In the case of the EBLUP predictor for the small area mean, $h_{4i}\left(\tilde{\boldsymbol{\delta}}\right)$ contains two differences with respect to the same expression developed for REBLUP: i) the matrix $\mathbf{B}_s = \boldsymbol{\Sigma}_u \mathbf{Z}_s^T \mathbf{V}_s^{-1}$; ii) $Var(\hat{\sigma}_u^2, \hat{\sigma}_e^2)$ is obtained using the results of the asymptotic distribution of $\left(\hat{\sigma}_u^2, \hat{\sigma}_e^2\right)$ given by Rao (2003). The MSE of the EBLUP (2) is therefore

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{EBLUP}\right) = h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{4i}\left(\tilde{\boldsymbol{\delta}}\right) \qquad (44)$$

and its estimator can be written as:

$$mse\left(\hat{\bar{y}}_i^{EBLUP}\right) = h_{1i}\left(\hat{\boldsymbol{\delta}}\right) + h_{2i}\left(\hat{\boldsymbol{\delta}}\right) + h_{3i}\left(\hat{\boldsymbol{\delta}}\right) + 2h_{4i}\left(\hat{\boldsymbol{\delta}}\right) \qquad (45)$$

since $E\left[h_{1i}\left(\hat{\delta}\right)\right] \approx h_{1i}\left(\delta\right) - h_{4i}\left(\delta\right) + o\left(m^{-1}\right)$.

MSE estimation for MQ

The third predictor that we consider is the MQ predictor (18) based on the M-quantile approach (Chambers and Tzavidis, 2006). For fixed q, the prediction variance of the MQ predictor is

$$Var(\hat{\bar{y}}_i^{MQ} - \bar{y}_i) = \left(1 - \frac{n_i}{N_i}\right)^2 \left\{ \bar{\mathbf{x}}_{ri}^T Var\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right) \bar{\mathbf{x}}_{ri} \right\} + \left(1 - \frac{n_i}{N_i}\right)^2 Var\left(\bar{e}_{ri}\right). \quad (46)$$

It follows that we need to estimate $Var\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)$ in order to be able to calculate an estimate of the prediction variance of this predictor. The starting point, as usual, is the first order approximation based on the estimating equations for $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$. Putting $q = \bar{q}$,

$$Var_0\left(\hat{\boldsymbol{\beta}}_q\right) \approx \left\{ E_0\left(\partial_{\beta_q} \mathbf{H}_0\right) \right\}^{-1} Var_0\left\{ \mathbf{H}\left(\boldsymbol{\beta}_{0q}\right) \right\} \left[\left\{ E_0\left(\partial_{\beta_q} \mathbf{H}_0\right) \right\}^{-1} \right]^T \quad (47)$$

with

$$\mathbf{H}(\boldsymbol{\beta}_{0q}) = \sum_{j=1}^{n} \mathbf{x}_j \psi_q(r_j) = \mathbf{X}_s^T \psi_q(r_{0q})$$

where ψ_q is a bounded influence function depending on q, $\psi_q(r_{0q})$ is the n-vector with elements $\psi_q(r_{j0q}) = \psi_q\left\{ \omega_{j0q}^{-1}\left(y_j - \mathbf{x}_j^T \boldsymbol{\beta}_{0q}\right) \right\}$ and ω_{j0q} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T \boldsymbol{\beta}_{0q}$. The $Var_0\left\{ \mathbf{H}(\beta_{0q}) \right\}$ component of expression (47) can then be written as

$$Var_0\left\{ \mathbf{H}(\beta_{0q}) \right\} = \mathbf{X}_s^T \left\{ E_0\left\{ \psi_q(r_{0q}) \psi_q^T(r_{0q}) \right\} \right\} \mathbf{X}_s,$$

because the y values are conditionally uncorrelated and $E_0\left\{ \psi_q(r_{0q}) \right\} = 0$ for each q. Assuming a Huber-type influence function, we obtain

$$E_0\left(\partial_{\beta_q} \mathbf{H}_{0q}\right) = \mathbf{X}_s^T E_0\left[2 \frac{d}{d\boldsymbol{\beta}_q} \psi_q(r_{0q}) \Big|_{\boldsymbol{\beta}_q = \boldsymbol{\beta}_{0q}} \right] = -2\mathbf{X}_s^T \mathbf{C} \mathbf{X}_s,$$

where $\mathbf{C}$ is a $n \times n$ diagonal matrix with j-th diagonal component $\omega_{j0q}^{-1} E_{0q}\left\{ qI\left(0 < r_{j0q} \le c\right) + \left(1 - q\right) I\left(-c < r_{j0q} \le 0\right) \right\}$. These expressions then lead to two types of estimators:

1. $$\widehat{Var}(\hat{\boldsymbol{\beta}}_q) = \lambda n(n-p)^{-1} \left\{ \hat{E}\left(\partial_{\beta_q} \mathbf{H}_{0q}\right) \right\}^{-1} \widehat{Var}\left\{ \mathbf{H}(\boldsymbol{\beta}_{0q}) \right\} \left[\left\{ \hat{E}\left(\partial_{\beta_q} \mathbf{H}_{0q}\right) \right\}^{-1} \right]^T \quad (48)$$

where $\widehat{Var}\left\{\mathbf{H}(\boldsymbol{\beta}_{0q})\right\}=\mathbf{X}_s^T\hat{\mathbf{F}}\mathbf{X}_s$, $\hat{\mathbf{F}}$ is a diagonal matrix of dimension $n\times n$ with *j*-th element equal to $\hat{f}_j=\hat{w}_{jq}^2\hat{r}_{jq}^2-\left(\sum_{i=1}^{n}\hat{w}_{jq}\hat{r}_{jq}\right)^2$; $\hat{E}\left(\partial_{\beta_q}\mathbf{H}_{0q}\right)=-2\mathbf{X}_s^T\hat{\mathbf{C}}\mathbf{X}_s$ where $\hat{\mathbf{C}}$ is a $n\times n$ diagonal matrix with *j*-th element $\hat{c}_j=\hat{\omega}_{jq}^{-1}\left\{qI\left(0<\hat{r}_{jq}\le c\right)+\left(1-q\right)I\left(-c<\hat{r}_{jq}\le 0\right)\right\}$. Here $\hat{w}_{jq}$ is the final weight in the iterative re-weighted least squared (IRLS) process, and $\hat{r}_{jq}=\hat{\omega}_{jq}^{-1}\left(y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_q\right)$. Note the factor $\lambda n(n-p)^{-1}$ which ensures agreement with Street *et al*. (1988) when $\mathbf{X}_s=\mathbf{1}$ and $q=0.5$. The $\lambda=\left\{1+\frac{p}{n}Var\left(\psi_q'\left(\hat{r}_{jq}\right)\right)\left[E\left(\psi_q'\left(\hat{r}_{jq}\right)\right)\right]^{-2}\right\}$ value is the bias corrector term (Huber, 1981).

2. $$\widehat{Var}(\hat{\boldsymbol{\beta}}_q)=\frac{\left(n-p\right)^{-1}\lambda\left\{\sum_{i=1}^{n}\psi_q^2\left(\hat{r}_{iq}\right)\right\}}{\left[n^{-1}\sum_{i=1}^{n}\psi_q'\left(\hat{r}_{iq}\right)\right]^2}\left(\mathbf{X}_s^T\mathbf{X}_s\right)^{-1}. \tag{49}$$

That is, the Street *et al*. (1988) estimator when $q=0.5$.

Depending on which of (48) or (49) is used, the estimator of the prediction variance of the MQ predictor when $q=\bar{q}$ can be written as:

$$\hat{V}\left(\hat{\bar{y}}_i^{MQ}\right)=\left(1-\frac{n_i}{N_i}\right)^2\left\{\bar{\mathbf{x}}_{ri}^T\widehat{Var}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\bar{\mathbf{x}}_{ri}\right\}+\left(1-\frac{n_i}{N_i}\right)^2\hat{V}\left(\bar{e}_{ri}\right) \tag{50}$$

with $\hat{V}(\bar{e}_{ri})=\frac{1}{(N_i-n_i)(n-1)}\sum_h\sum_{j\in s_h}\left(y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_h}\right)^2$. Moreover, since we are taking a conditional approach, we have to add an estimator of the squared bias based on:

$$\hat{B}\left(\hat{\bar{y}}_i^{MQ}\right)=N_i^{-1}\left\{\sum_k\sum_{j\in s_k}w_j\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_k}-\sum_{j\in i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right\}$$

where

$$w_j=\begin{cases}\frac{N_i}{n_i}+b_j & if\ j\in i\\ b_j & otherwise\end{cases}$$

and $\mathbf{b}^T=\left(\sum_{j\in r_i}\mathbf{x}_j^T-\frac{N_i-n_i}{n_i}\sum_{j\in s_i}\mathbf{x}_j^T\right)(\mathbf{X}_s^T\mathbf{W}(\bar{q}_i)\mathbf{X}_s)^{-1}\mathbf{X}_s^T\mathbf{W}(\bar{q}_i)$ is a $1\times n$ vector. The final expression for the

MSE estimator of the MQ predictor is therefore:

$$\widehat{MSE}\left(\hat{\hat{y}}_i^{MQ}\right)=\left(1-\frac{n_i}{N_i}\right)^2\left\{\bar{\mathbf{x}}_{ri}^T\widehat{Var}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\bar{\mathbf{x}}_{ri}\right\}+\left(1-\frac{n_i}{N_i}\right)^2\hat{V}\left(\bar{e}_{ri}\right)+\left\{N_i^{-1}\left\{\sum_k\sum_{j\in s_k}w_j\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_k}-\sum_{j\in i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right\}\right\}^2. \quad (51)$$

Note that (50) is a first order approximation to the asymptotic prediction variance of the MQ predictor, and so (51) could underestimate its MSE.

5. Results from Model-Based Simulation Studies

We provide model-based simulation results illustrating the comparative performances of the different outlier robust small area predictors described above. Population data are generated for $m=40$ 'small areas', with samples selected by simple random sampling without replacement within each area. Population and sample sizes are the same for all areas, and are fixed at either $N_i=100, n_i=5$ or $N_i=300, n_i=15$. Values for *X* are generated as independently and identically distributed from a lognormal distribution with a mean of 1.004077 and a standard deviation of 0.5 on the log scale. Values for *Y* are generated as $y_{ij}=100+5x_{ij}+u_i+\varepsilon_{ij}$, where the random area and individual effects are independently generated according to four scenarios:

- [0,0] – No outliers: $u\sim N(0,3)$ and $\varepsilon\sim N(0,6)$.
- [e,0] – Individual outliers only: $u\sim N(0,3)$ and $\varepsilon\sim\delta N(0,6)+(1-\delta)N(20,150)$, where δ is an independently generated Bernoulli random variable with $\Pr(\delta=1)=0.97$, i.e. the individual effects are independent draws from a mixture of two normal distributions, with 97% on average drawn from a 'well-behaved' $N(0,6)$ distribution and 3% on average drawn from an outlier $N(20,150)$ distribution.
- [0,u] – Area outliers only: $u\sim N(0,3)$ for areas 1-36, $u\sim N(9,20)$ for areas 37-40 and $\varepsilon\sim N(0,6)$, i.e. random effects for areas 1–36 are drawn from a 'well behaved' $N(0,3)$ distribution, with those for areas 37–40 drawn from an outlier $N(9,20)$ distribution. Individual effects are not outlier-contaminated.
- [e,u] – Outliers in both area and individual effects: $u\sim N(0,3)$ for areas 1-36, $u\sim N(9,20)$ for areas 37-40 and $\varepsilon\sim\delta N(0,6)+(1-\delta)N(20,150)$.

Each scenario is independently simulated 500 times. For each simulation the population values are generated according the underlying scenario model, a sample is selected in each area and the sample data are then used to compute estimates of each of the actual area means for Y.

Five different estimators are used for this purpose - the standard EBLUP, see (2), which serves as a reference; the projective M-quantile estimator MQ, see (18); the robust bias-corrected predictive MQ estimator MQ-BC, see (21); the robust projective REBLUP estimator of Sinha and Rao (2009), see (16); and its robust bias-corrected version REBLUP-BC, see (22). In all cases the 'projective' influence function ψ is a Huber Proposal 2 type with tuning constant $c = 1.345$. In contrast, the 'predictive', less restrictive, influence function ϕ used in MQ-BC and REBLUP-BC is also a Huber Proposal 2 type, but with a larger tuning constant, $c = 3$.

The performance of these estimators across the different areas and simulations is assessed by computing the median values of their area specific relative bias and relative root mean squared error, where the relative bias of an estimator $\hat{\bar{y}}_i$ for the actual mean $\bar{y}_i$ of area i is the average across simulations of the errors $\hat{\bar{y}}_i - \bar{y}_i$ divided by the corresponding average value of $\bar{y}_i$, and its relative root mean squared error is the square root of the average across simulations of the squares of these errors, again divided by the average value of $\bar{y}_i$. Table 1 sets out these median values for the different simulation scenarios and different estimators.

The relative bias results set out in Table 1 confirm our expectations regarding the behaviour of projective estimators (EBLUP, REBLUP and MQ) versus bias-corrected predictive estimators (REBLUP-BC and MQ-BC). The former are more biased than the latter as a consequence of their implicit assumption that although outlier variances may be inflated relative to non-outliers, outlier effects still have zero expectation. This increase in bias is most pronounced when there are outliers in the area effects, which is not unexpected since that is when area means are most affected by the presence of outliers in the population data. Turning to the median RRMSE results, we see that claims in the literature (e.g. Chambers and Tzavidis, 2006) about the superior outlier robustness of MQ compared with the EBLUP certainly hold true – provided the outliers are in individual effects. If there are outliers in area effects, then MQ appears to offer no extra protection compared to the EBLUP, and in fact performs worse, mainly due to its sharply increasing bias in this situation. Similarly, when we compare the EBLUP and the REBLUP we see that if outliers are associated with individual effects, then REBLUP offers better RRMSE performance than EBLUP. However, the gap between these two estimators narrows considerably when outliers are associated with area effects. In contrast, the two bias-corrected predictive estimators seem relatively robust in terms of RRMSE performance. Due to increased variability as a consequence of their bias corrections, both BC estimators are not as efficient as the projective estimators when outliers are associated with individual effects, but both also do not fail when there are outliers in the area effects.

We now turn to an examination of the performance of different methods of MSE estimation investigated in the simulations. MSE estimation for the REBLUP and REBLUP-BC is implemented via the robust MSE estimators (28) and (29) (hereafter CCT) and via the linearization-based MSE estimators (40) and (A6) (hereafter CCST), while for the MQ and the MQ-BC both (51) and (A9) (CCST) and the robust MSE estimator described in Chambers *et al*. (2007, Section 2.3 - CCT) are calculated. The bootstrap procedure proposed by Sinha and Rao (2009) for REBLUP is also investigated by using bootstrap samples of sizes 100. The MSE of the EBLUP estimator is estimated by Prasad-Rao (PR), CCT (Chambers *et al*., 2007, Section 2.3) and CCST (45) estimators.

The behaviour of the MSE estimators for each scenario and for each approach is shown in Table 2 where we report the median values of their area specific relative bias, relative root mean squared error and coverage rate for a nominal 95 per cent confidence interval. These intervals are based on 'normal theory' and are defined by the small area mean estimate plus or minus twice their corresponding estimated root mean squared error. These results show that both CCT and CCST tend to be biased low, but CCST is better in terms of coverage rate. It shows a small amount of under-coverage for all predictors. The CCST estimator is preferable to CCT for REBLUP and REBLUP-BC. It shows smaller bias and more stability. Moreover it seems that CCST is better able to handle the scenarios where outliers are present. The CCT and CCST estimators perform similarly for MQ-BC, even if CCST seems more stable. The PR estimator of MSE does well: it is very stable and shows good bias properties except in the presence of area level outliers, when it is biased downwards significantly. The bias properties of the bootstrap MSE estimator for REBLUP and REBLUP-BC are comparable with CCST, but it is much more stable.

6. Design-Based Simulation Study

Design-based simulations complement model-based simulations for small area estimation since they allow us to evaluate the performance of small area estimation methods in the context of a real population and realistic sampling methods where we do not know the precise source of the contamination. From a practical perspective we believe that this type of simulation, by effectively fixing the differences between the small areas, constitutes a more practical and appropriate representation of the small area estimation problem from a finite population perspective. Further, it provides a good illustration of why a focus on conditional MSE is likely to be closer to the MSE of interest for people using small area methods.

The population underpinning the design-based simulation is based on a data set obtained under the Environmental Monitoring and Assessment Program (EMAP) of the U.S. Environmental Protection Agency. The background to this data set is that between 1991 and 1995, EMAP conducted a survey of lakes in the North-Eastern states of the U.S. The data collected in this survey consists of 551 measurements from a sample of 334 of the 21,026 lakes located in this area. The lakes making up this population are grouped into 113 8-digit Hydrologic Unit Codes (HUCs), of which 64 contained less than 5 observations and 27 did not have any. In our simulation, we defined HUCs as the small areas of interest, with lakes grouped within HUCs. The variable of interest is Acid Neutralizing Capacity (ANC), an indicator of the acidification risk of water bodies. A total of 1000 independent random samples of lake locations are then taken from the population of 21,026 lake locations by randomly selecting locations in the 86 HUCs that containing EMAP sampled lakes, with sample sizes in these HUCs set to the greater of five and the original EMAP sample size. Details on the data generation are in Salvati *et al.* (2008). Table 3 shows the median relative bias and the median relative root MSE of the different predictors (EBLUP, REBLUP, MQ, REBLUP-BC, MQ-BC). Similarly, Table 4 report the median relative bias, the median relative root MSE and the median coverage rate of the corresponding estimators of the MSEs of these predictors calculated from the same sample. MQ-BC and REBLUP-BC predictors work well in terms of both bias and MSE, while the EBLUP is the worst in terms of relative root MSE. The REBLUP shows a good performance in terms of RRMSE but records a big negative bias. The MQ predictor shows the worst behaviour in terms of bias and MSE.

We now turn to an examination of the performance of different methods of MSE estimation investigated in the design-based simulation. The Prasad-Rao (PR) estimator of the MSE of the EBLUP has an upward bias and larger instability than the CCST estimator for the EBLUP. This could be due to the unconditional basis of the PR estimator. The CCST estimator seems to offer the best overall results with REBLUP and REBLUP-BC, while CCT and CCST show similar performance in terms of bias and RRMSE for MQ-BC. In this simulation experiment the MSE estimation of the MQ predictor is problematic for both CCT and CCST. The bootstrap MSE estimator does not work for the REBLUP, showing big bias and instability, whereas it is a good competitor for CCT and CCST as far as REBLUP-BC is concerned. The coverage rates (for nominal 95 percent intervals) are presented in Table 4. The CCST estimation method produces intervals with median coverage close to 95 percent for EBLUP, REBLUP and REBLUP-BC. It records substantial under-coverage for MQ and MQ-BC, even if, for these estimators, it performs better than CCT. The bootstrap MSE estimator shows a degree of over-coverage for REBLUP. This occurs because the bootstrap method assumes that the linear mixed model (1) holds for the small areas, whereas this assumption is difficult to meet in many practical applications. A final comment is appropriate considering the results on the coverage rate. Chatterjee *et al*. (2008), discussing the use of bootstrap methods for constructing confidence intervals for small area parameters, argue that there is no guarantee that the asymptotic behaviour underpinning normal theory confidence intervals applies in the context of the small samples that characterize small area estimation. For this reason the authors do not recommend the use of the ‘normal theory’ to construct the prediction intervals (as we have done here).

The behaviour of the empirical true root MSE and its estimators for each area and for each approach are shown in Figures 1, 2 and 3. Examination of these results can be useful for understanding the reasons for different performances of the MSE estimators. Figure 1 shows the results for EBLUP predictor and we can note that the PR estimator does not seem to be able to capture between area differences in the design-based RMSE of the EBLUP, while the CCT MSE estimator for the EBLUP tracks the irregular profile of the area-specific empirical MSE very well. Also CCST works quite well but produces somewhat over-smoothed estimates of area-specific empirical MSE. These results confirm the poor design-based properties of the PR estimator (Longford, 2007). Figure 2 reports the results for REBLUP and REBLUP-BC predictors. For the REBLUP (top) it is evident that CCT tends to underestimate the true area-specific MSE, mainly because its squared bias component underestimates the actual squared bias of this predictor. The bootstrap MSE estimator produces over-smoothed estimates of area-specific empirical MSE, because in this simulation the assumption that linear mixed model (1) holds is violated. The CCST estimator tracks area-specific empirical MSE but it shows underestimation in a few areas. It can be seen that the CCST MSE estimator for the REBLUP-BC (bottom) has the best performance and tracks the irregular profile of the area-specific empirical MSE very well, while the bootstrap MSE estimator for the REBLUP-BC generates over-smoothed estimates of area-specific empirical MSE. Figure 3 illustrates the results for MQ (top) and MQ-BC (bottom) predictors. The MSE estimators have a similar behaviour. They track the irregular profile of the area-specific empirical MSE very well for MQ-BC, while, for MQ, the CCT and CCST underestimates the true area-specific MSE.

7. Final Remarks

In this paper we explore the extension of the Robust Predictive approach to small area estimation and we propose two different analytical mean squared error (MSE) estimators for outlier robust predictors of small area means. The first proposal is a bias-robust MSE estimator based on the 'pseudo-linearization' approach discussed by Chambers *et al.* (2007). The second method is a linearization-based MSE estimation based on first order approximations to the variances of solutions of estimating equations.

The empirical results in Sections 5 and 6 show that the bias-corrected predictive estimators (REBLUP-BC and MQ-BC) are less biased than the projective estimators (EBLUP, REBLUP and MQ) especially when there are outliers in the area effects. From the results of the simulation experiments there is evidence that the BC estimators are not as efficient as the projective estimators when outliers are associated with individual effects. This is due to increased variability as a consequence of their bias corrections. We can note also that REBLUP-BC and MQ-BC do not fail when there are outliers in the area effects. A method to compute the ‘optimal’ cut-off value c for the function ϕ and improve the efficiency of the BC estimators remains to be done. A cross-validation approach could be a possible method.

The pseudo-linearization and linearization-based MSE estimators described in Section 4 and in the Appendix can be an alternative to bootstrap MSE estimation for REBLUP and REBLUP-BC. Moreover, the CCST estimator shows a good performance also for MQ-type estimators. Overall, the CCST method performs reasonably well for the different small area predictors that we have compared in both model-based and design-based simulation experiments. We also note that the Prasad-Rao estimator of the EBLUP and the bootstrap MSE estimator of the REBLUP proposed by Sinha and Rao (2009), which work well when their model assumptions are valid, have problems, especially in terms of bias, in the presence of outliers. In the model-based simulations the CCST estimator performs quite well in all scenarios and it works better than PR and bootstrap-type MSE estimators when there outliers in the area and individual effects in terms of bias, stability and coverage rate.

Recently, the CCT estimator has been extended to estimating the MSE of M-quantile Geographically Weighted Regression small area estimators (Salvati *et. al.*, 2008) and to predictors based on nonparametric small area models (Salvati *et al.*, 2009). It could be interesting to explore whether the CCST estimator can also be used in these cases, or with nonparametric M-quantile small area estimators (Pratesi *et al.* 2008).

Finally, the CCST MSE estimator presented in this paper is developed under the conditional version of the linear mixed model, i.e. it is conditioned on area effects when applied in the context of a mixed model. However, it is possible to develop an unconditional version of the CCST MSE estimator that averages over the distribution of the random area effects under a linear mixed model, and so reduces to the Prasad-Rao MSE estimator in the case of the EBLUP. This is an avenue for further research.

Appendix

In the appendix we show how the weights $\left(\mathbf{w}_{is}^{RBLUP}\right)^T$ in expression (23) are obtained, and sketch the development of the linearization-based MSE estimators for the REBLUP-BC (22) and MQ-BC (21) predictors.

A.1 The weights $\left(\mathbf{w}_{is}^{RBLUP}\right)^T$

Under the model (1) the RBLUP is

$$\begin{aligned}\hat{y}_i^{RBLUP} &= \frac{1}{N_i}\left\{\sum\nolimits_{j\in s_i} y_j + \sum\nolimits_{j\in r_i}\left(\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} + \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)\right\} \\ &= \frac{1}{N_i}\left\{\sum\nolimits_{j\in s_i} y_j + \sum\nolimits_{j\in r_i}\left(\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}\right) + \sum\nolimits_{j\in r_i}\left(\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)\right\} \\ &= \frac{1}{N_i}\left\{\sum\nolimits_{j\in s_i} y_j + \left(\sum\nolimits_{j\in r_i}\mathbf{x}_j^T\right)\tilde{\boldsymbol{\beta}}^{\psi} + \left(\sum\nolimits_{j\in r_i}\mathbf{z}_j^T\right)\tilde{\mathbf{u}}^{\psi}\right\}\end{aligned}$$

The components $\left(\sum\nolimits_{j\in r_i}\mathbf{x}_j^T\right)$ and $\left(\sum\nolimits_{j\in r_i}\mathbf{z}_j^T\right)$ can be written as $\left(N_i - n_i\right)\bar{\mathbf{x}}_{ir}^T$, $\left(N_i - n_i\right)\bar{\mathbf{z}}_{ir}^T$, respectively, where $\bar{\mathbf{x}}_{ir}$ and $\bar{\mathbf{z}}_{ir}$ are the means for non-sampled units from area *i* for the three sets of covariates. It immediately follows that the RBLUP can be expressed as

$$\hat{y}_i^{RBLUP} = N_i^{-1}\left\{\sum\nolimits_{j\in s_i} y_j + \left(N_i - n_i\right)\bar{\mathbf{x}}_{ir}^T\tilde{\boldsymbol{\beta}}^{\psi} + \left(N_i - n_i\right)\bar{\mathbf{z}}_{ir}^T\tilde{\mathbf{u}}^{\psi}\right\}.$$

Moreover, $\tilde{\boldsymbol{\beta}}^{\psi} = \mathbf{A}_s\mathbf{y}_s$ and $\tilde{\mathbf{u}}^{\psi} = \mathbf{B}_s\left(\mathbf{I}_s - \mathbf{X}_s\mathbf{A}_s\right)$. Then the RBLUP becomes

$$\hat{y}_i^{RBLUP} = N_i^{-1}\left\{\mathbf{1}_s^T + \left(N_i - n_i\right)\left[\bar{\mathbf{x}}_{ir}^T\mathbf{A}_s + \bar{\mathbf{z}}_{ir}^T\mathbf{B}_s\left(\mathbf{I}_s - \mathbf{X}_s\mathbf{A}_s\right)\right]\right\}\mathbf{y}_s = \left(\mathbf{w}_{is}^{RBLUP}\right)^T\mathbf{y}_s$$

where the weights are defined as

$$\left(\mathbf{w}_{is}^{RBLUP}\right)^T = N_i^{-1}\left\{\mathbf{1}_s^T + \left(N_i - n_i\right)\left[\bar{\mathbf{x}}_{ir}^T\mathbf{A}_s + \bar{\mathbf{z}}_{ir}^T\mathbf{B}_s\left(\mathbf{I}_s - \mathbf{X}_s\mathbf{A}_s\right)\right]\right\}.$$

A.2 Linearization-Based MSE estimation for REBLUP-BC and MQ-BC

We develop MSE estimators for the REBLUP-BC and MQ-BC predictors by extending the linearization approach of Street *et al*. (1988) to estimation of prediction variance for estimators based on robust estimating equations.

MSE estimation of REBLUP-BC

To obtain the prediction variance of RBLUP-BC we start from the prediction error

$$\hat{\bar{y}}_i^{RBLUP-BC} - \bar{y}_i = \left(1-\frac{n_i}{N_i}\right)\left\{\left(\bar{\mathbf{x}}_{ri}^T\tilde{\boldsymbol{\beta}}^{\psi} + \bar{\mathbf{z}}_{ri}^T\tilde{\mathbf{u}}^{\psi}\right) + \frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}}\right) - \bar{y}_{ri}\right\}.$$

Then, we can write

$$\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}}\right) \approx \frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right) + \begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi} - \boldsymbol{\beta}^{\psi}\\ \tilde{\mathbf{u}}^{\psi} - \mathbf{u}^{\psi}\end{pmatrix}^T \partial_{\beta^{\psi},u^{\psi}}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\right\}$$

where

$$\begin{aligned}\partial_{\beta}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\right\} &= -\frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_j\\ \partial_{v}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\right\} &= -\frac{1}{n_j}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{z}_j\end{aligned}.$$

Hence

$$\begin{aligned}\hat{\bar{y}}_i^{RBLUP-BC} - \bar{y}_i &\approx \left(1-\frac{n_i}{N_i}\right)\left\{\left(\bar{\mathbf{x}}_{ri}^T\tilde{\boldsymbol{\beta}}^{\psi} + \bar{\mathbf{z}}_{ri}^T\tilde{\mathbf{u}}^{\psi}\right) + \frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\right.\\ &\quad\left. - \begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi} - \boldsymbol{\beta}^{\psi}\\ \tilde{\mathbf{u}}^{\psi} - \mathbf{u}^{\psi}\end{pmatrix}^T \begin{pmatrix}\frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_i\\ \frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{z}_i\end{pmatrix} - \bar{y}_{ri}\right\}\\ &= \left(1-\frac{n_i}{N_i}\right)\left\{\frac{1}{n_i}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right) + \begin{pmatrix}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_j\\ \phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{z}_j\end{pmatrix}^T\begin{pmatrix}\boldsymbol{\beta}^{\psi}\\ \mathbf{u}^{\psi}\end{pmatrix}\right.\right.\\ &\quad\left. + \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_j\\ \bar{\mathbf{z}}_{ri} - \frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)\mathbf{z}_j\end{pmatrix}^T\begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi}\\ \tilde{\mathbf{u}}^{\psi}\end{pmatrix} - \bar{y}_{ri}\right\}.\end{aligned}$$

If the tuning constant used in the BC term is large, $\phi' \approx 1$ and we can write

$$\hat{\bar{y}}_i^{RBLUP-BC} - \bar{y}_i = \left(1-\frac{n_i}{N_i}\right)\left\{\frac{1}{n_i}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right) + \begin{pmatrix}\mathbf{x}_j\\ \mathbf{z}_j\end{pmatrix}^T\begin{pmatrix}\boldsymbol{\beta}^{\psi}\\ \mathbf{u}^{\psi}\end{pmatrix}\right\} + \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\\ \bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\end{pmatrix}^T\begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi}\\ \tilde{\mathbf{u}}^{\psi}\end{pmatrix} - \bar{y}_{ri}\right\}. \quad \text{(A1)}$$

The covariance between the first and second terms on the right hand side should be of a lower order of magnitude than either of their variances, so we can write

$$Var(\hat{\bar{y}}_i^{RBLUP-BC} - \bar{y}_i) \approx \left(1-\frac{n_i}{N_i}\right)^2 \left\{ Var\left(\frac{1}{n_i}\sum_{j\in s_i}\left\{ \omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}} \right) + \begin{pmatrix}\mathbf{x}_j \\ \mathbf{z}_j\end{pmatrix}^T \begin{pmatrix}\boldsymbol{\beta}^{\psi} \\ \mathbf{u}^{\psi}\end{pmatrix} \right\} \right) \right.$$
$$\left. + \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si} \\ \bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\end{pmatrix}^T Var\begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi} \\ \tilde{\mathbf{u}}^{\psi}\end{pmatrix} \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si} \\ \bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\end{pmatrix} + Var\left(\bar{e}_{ri}\right) \right\}$$
$$= \left(1-\frac{n_i}{N_i}\right)^2 \left\{ \frac{1}{n_i^2}\sum_{j\in s_i} E\left\{ \omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}} \right)\right\}^2 + \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si} \\ \bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\end{pmatrix}^T Var\begin{pmatrix}\tilde{\boldsymbol{\beta}}^{\psi} \\ \tilde{\mathbf{u}}^{\psi}\end{pmatrix} \begin{pmatrix}\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si} \\ \bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\end{pmatrix} + Var\left(\bar{e}_{ri}\right) \right\}.$$

The corresponding estimator of the MSE of RBLUP-BC is therefore

$$MSE\left(\hat{\bar{y}}_i^{RBLUP-BC}\right) = h_{1i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{6i}\left(\tilde{\boldsymbol{\delta}}\right)$$
$$= \left(1-\frac{n_i}{N_i}\right)^2 \left\{ \left(\bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\right)^T \hat{V}\left(\tilde{\mathbf{u}}^{\psi}\right)\left(\bar{\mathbf{z}}_{ri} - \bar{\mathbf{z}}_{si}\right) + \left(\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right)^T \hat{V}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\left(\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right) \right. \quad \text{(A2)}$$
$$\left. + \hat{V}\left(\bar{e}_{ri}\right) + \frac{1}{n_i(n_i - p)}\sum_{j\in s_i}\left\{ \omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}} \right)\right\}^2 \right\}$$

with $\tilde{\boldsymbol{\delta}} = \left(\tilde{\boldsymbol{\beta}}^{\psi T}, \tilde{\mathbf{u}}^{\psi T}\right)^T$. For a random intercepts model, this become

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP-BC}\right) = h_{2i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{3i}\left(\tilde{\boldsymbol{\delta}}\right) + h_{6i}\left(\tilde{\boldsymbol{\delta}}\right)$$
$$= \left(1-\frac{n_i}{N_i}\right)^2 \left\{ \begin{array}{c} \left(\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right)^T \hat{V}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\left(\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right) + \hat{V}\left(\bar{e}_{ri}\right) \\ + \frac{1}{n_i(n_i - p)}\sum_{j\in s_i}\left\{ \omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}} \right)\right\}^2 \end{array} \right\}. \quad \text{(A3)}$$

Expressions for $\hat{V}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)$, $\hat{V}\left(\tilde{\mathbf{u}}^{\psi}\right)$ and $\hat{V}\left(\bar{e}_{ri}\right)$ are set out in the RBLUP development in Section 4.2. The conditional MSE of REBLUP-BC is obtained by adding an extra term to (A2) and (A3) due to the uncertainty of the estimated variance components. The approach already used for RBLUP can be used to approximate $\hat{\bar{y}}_i^{REBLUP-BC} - \hat{\bar{y}}_i^{RBLUP-BC}$ by

$$\sum_{k=1}^{2}\left\{ \frac{1}{N_i}\sum_{j\in r_i}\mathbf{z}_j^T\left(\partial_{\theta_k}\mathbf{B}_s\right)\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi}\right) + \frac{N_i - n_i}{N_i n_i}\sum_{j\in s_i}\partial_{\theta_k}\left(\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}^{\psi} - \mathbf{z}_j^T\mathbf{B}_s\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)}{\omega_{ij}^{\psi}} \right)\right)\right\}\left\{\hat{\theta}_k - \theta_k\right\}$$

where $\mathbf{B}_s$ is defined in Section 4.1 and $\boldsymbol{\theta}$ is the vector of the variance components. Hence,

$$h_{4i}\left(\boldsymbol{\delta}\right)=E\left[\left(\hat{\bar{y}}_i^{REBLUP-BC}-\hat{\bar{y}}_i^{RBLUP-BC}\right)^2\right]$$
$$\approx\left(1-\frac{n_i}{N_i}\right)^2\mathbf{D}_i^TVar\left\{\sum_{k=1}^{2}\left(\partial_{\theta_k}\mathbf{B}_s\right)\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)\left\{\hat{\theta}_k-\theta_k\right\}\right\}\mathbf{D}_i+o\left(m^{-1}\right) \qquad \text{(A4)}$$

where $D_i=\bar{\mathbf{z}}_{ri}-\frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}^{\psi}-\mathbf{z}_j^T\mathbf{B}_s\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)}{\omega_{ij}^{\psi}}\right)\mathbf{z}_j$. Note that $\mathbf{D}_i=\mathbf{0}$ when ϕ is the identity function (i.e. the Chambers and Dunstan (1992) version of BC) and the model is a random intercepts one. Finally, we estimate the variance-covariance matrix of the variance components $E\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_g-\theta_g\right)$ using results already obtained for the RBLUP, and hence calculate an estimate of the prediction variance of the REBLUP-BC. The MSE of $\hat{\bar{y}}_i^{REBLUP-BC}$ for the random intercept model can be written as

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP-BC}\right)=h_{2i}\left(\tilde{\boldsymbol{\delta}}\right)+h_{3i}\left(\tilde{\boldsymbol{\delta}}\right)+h_{4i}\left(\tilde{\boldsymbol{\delta}}\right)+h_{6i}\left(\tilde{\boldsymbol{\delta}}\right). \qquad \text{(A5)}$$

An estimator of (A5) is:

$$mse\left(\hat{\bar{y}}_i^{RBLUP-BC}\right)=h_{2i}\left(\hat{\boldsymbol{\delta}}\right)+h_{3i}\left(\hat{\boldsymbol{\delta}}\right)+h_{4i}\left(\hat{\boldsymbol{\delta}}\right)+h_{6i}\left(\hat{\boldsymbol{\delta}}\right) \qquad \text{(A6)}$$

with $E\left[mse\left(\hat{\bar{y}}_i^{REBLUP-BC}\right)\right]=\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP-BC}\right)+o\left(m^{-1}\right)$.

MSE estimation of MQ-BC

As in the Section 4.2 for the MQ case we assume that the $\bar{q}_i$ values are known. The estimation or prediction error for MQ-BC is

$$\hat{\bar{y}}_i^{MQ-BC}-\bar{y}_i=N_i^{-1}\sum_{j\in r_i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}+\frac{N_i-n_i}{N_in_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)-N_i^{-1}\sum_{j\in r_i}y_j.$$

We can write

$$\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\approx\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)+\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}-\boldsymbol{\beta}_{\bar{q}_i}\right)^T\partial_{\beta_{\bar{q}_i}}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_i-\mathbf{x}_i^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}$$

where

$$\partial_{\beta_{\bar{q}_i}}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}=-\frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_i$$

and hence

$$\hat{\bar{y}}_i^{MQ-BC} - \bar{y}_i \approx \left(1-\frac{n_i}{N_i}\right)\left\{\frac{1}{n_i}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)+\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}\right\}\right.$$
$$\left.+\left\{\bar{\mathbf{x}}_{ri} - \frac{1}{n_i}\sum_{j\in s_i}\phi'\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\mathbf{x}_j\right\}^T \hat{\boldsymbol{\beta}}_{\bar{q}_i} - \bar{y}_{ri}\right\}.$$

If the tuning constant used in the BC term is large, $\phi' \approx 1$ and we can write

$$\hat{\bar{y}}_i^{MQ-BC} - \bar{y}_i \approx \left(1-\frac{n_i}{N_i}\right)\left(\frac{1}{n_i}\sum_{j\in si}\left\{\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i} + \omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\} + \left\{\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right\}^T \hat{\boldsymbol{\beta}}_{\bar{q}_i} - \bar{y}_{ri}\right). \quad \text{(A7)}$$

The covariance between the first and second terms on the right hand side should be of a lower order of magnitude than either of their variances, so

$$Var\left(\hat{\bar{y}}_i^{MQ-BC} - \bar{y}_i\right) = \left(1-\frac{n_i}{N_i}\right)^2\left[\begin{array}{c}\left\{\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right\}^T Var\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\left\{\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right\} + Var\left(\bar{e}_{ri}\right)\\ +\frac{1}{n_i^2}\sum_{j\in s_i}E\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2\end{array}\right]. \quad \text{(A8)}$$

The corresponding estimator of the MSE of MQ BC, when $q = \hat{\bar{q}}$, is therefore:

$$\widehat{MSE}(\hat{\bar{y}}_i^{MQ-BC}) = \left(1-\frac{n_i}{N_i}\right)^2\left[\begin{array}{c}\left\{\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right\}^T \hat{V}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\left\{\bar{\mathbf{x}}_{ri} - \bar{\mathbf{x}}_{si}\right\} + \hat{V}\left(\bar{e}_{ri}\right)\\ +\frac{1}{n_i^2}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2\end{array}\right] \quad \text{(A9)}$$

where $\hat{V}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_j}\right)$ is obtained as in the case of MQ estimator and $\hat{V}(\bar{e}_{ri}) = \frac{1}{(N_i - n_i)(n-1)}\sum_h \sum_{j\in s_h}\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_h}\right)^2$.

The expression $\frac{1}{n_i^2}\sum_{j\in s_i}E\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2$ is estimated by $\frac{1}{n_i^2}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2$ because a standard Taylor series argument implies

$$E\left\{\phi\left(\frac{y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2 \approx E\left\{\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2 + \left(\omega_{ij}^{\psi}\right)^{-2}\mathbf{x}_j^T Var\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\mathbf{x}_j.$$

So a more appropriate estimator of $\frac{1}{n_i^2}\sum_{j\in s_i}E\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j - \mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2$ would seem to be

$$\frac{1}{n_i^2}\sum_{j\in s_i}\left(\omega_{ij}^{\psi}\right)^2\left[\left\{\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2-\frac{\mathbf{x}_j^T\hat{V}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\mathbf{x}_j}{\left(\omega_{ij}^{\psi}\right)^2}\right].$$

In any case, a conservative estimator of $\frac{1}{n_i^2}\sum_{j\in s_i}E\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2$ is $\frac{1}{n_i^2}\sum_{j\in s_i}\left(\omega_{ij}^{\psi}\right)^2\left\{\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2$.

Acknowledgments. Work partially supported by the SAMPLE project (SSH-CT-2007-217565; FP7-SSH-2007-1) and by ARC Linkage Grant LP0776810. The authors are grateful to Space-Time Aquatic Resources Modeling and Analysis Program (STARMAP) for data availability. The views expressed here are solely those of the authors.

REFERENCES

Chambers, R.L. (1986). Outlier robust finite population estimation. *Journal of the American Statistical Association* **81**, 1063-1069.

Chambers, R.L., Dorfman, A.H. and Wehrly, T.E. (1993). Bias robust estimation in finite populations using nonparametric calibration. *Journal of the American Statistical Association* **88**, 268-277.

Chandra, H. and Chambers, R. (2005). Comparing EBLUP and C-EBLUP for Small Area Estimation. *Statistics in Transition*, **7**, 637-648.

Chambers, R. and Tzavidis, N. (2006). M-quantile models for small area estimation. *Biometrika* **93**, 255-268.

Chambers, R., Chandra, H. and Tzavidis, N. (2007). On Bias-Robust Mean Squared Error Estimation for Linear Predictors for Domains. Working Papers, 09-08, Centre for Statistical and Survey Methodology, The University of Wollongong, Australia. (Available from: http://cssm.uow.edu.au/publications).

Chandra H., Salvati N. and Chambers R. (2007). Small Area Estimation for Spatially Correlated Populations. A Comparison of Direct and Indirect Model-Based Methods. *Statistics in Transition*, **8**, 887-906.

Chandra, H. and Chambers, R. (2009). Multipurpose Weighting for Small Area Estimation. *Journal of Official Statistics*, **25**, 3, 379-395.

Chatterjee , S. , Lahiri , P. & Huilin, L. (2008). Parametric Bootstrap Approximation to the Distribution of EBLUP and Related Prediction Intervals in Linear Mixed Models. To appear in the *Annals of Statistics*.

Huber, P. (1981) *Robust Statistics*. J. Wiley, New York.

Longford, N.T. (2007). On Standard Errors of Model-Based Small-Area Estimators. *Survey Methodology*, **33**, 69-79.

Prasad, N.G.N. and Rao, J.N.K. (1990). The Estimation of the Mean Squared Error of Small Area Estimators. *Journal of the American Statistical Association*, **85**, 163-171.

Pratesi, M., Ranalli, M.G. and Salvati, N. (2008). Semiparametric M-quantile regression for estimating the proportion of acidic lakes in 8-digit HUCs of the Northeastern US. *Environmetrics*, **19**, 687-701

Rao, J.N.K. (2003). *Small Area Estimation*. Wiley, New York.

Richardson, A.M. and Welsh, A.H. (1995). Robust restricted maximum likelihood in mixed linear models. *Biometrics* **51**, 1429-1439.

Salvati, N., Tzavidis, N., Pratesi, M. and Chambers, R. (2008). Small Area Estimation Via M-quantile Geographically Weighted Regression. Paper submitted to *Test* and available from the authors upon request.

Salvati, N., Chandra, H., Ranalli, M.G. and Chambers, R. (2009). Small Area Estimation Using a Nonparametric Model Based Direct Estimator. Paper submitted to *Computational Statistics and Data Analysis* and available from the authors upon request.

Sinha, S.K. and Rao, J.N.K. (2009). Robust small area estimation. *The Canadian Journal of Statistics*, 37, 381-399.

Tzavidis, N., Marchetti, S. and Chambers, R. (2008). Robust prediction of small area means and distributions. Forthcoming in *Austalian & New Zealand Journal of Statistics.*

Welsh, A.H. and Richardson, A.M. (1997). Approaches to the Robust Estimation of Mixed Models. G. S. Maddala and C. R. Rao, eds., *Handbook of Statistics,* Vol. 15, Chapter 13.

Welsh, A.H. and Ronchetti, E. (1998). Bias-calibrated estimation from sample surveys containing outliers. Journal of the Royal Statistical Society, Series B, 60, 413-428.

Table 1 Model-based simulation results: performances of predictors of small area means.

Estimator		No outliers		Individual outliers	Area outliers	Both types
Scenario/Areas	[0,0]	[0,u]/ 1-36	[e,0]	[e,u]/ 1-36	[0,u]/ 37-40	[e,u]/ 37-40
Median values of Relative Bias (expressed as a percentage)						
EBLUP	0.019	0.097	-0.019	0.166	-0.536	-1.592
REBLUP	0.027	0.108	-0.391	-0.296	-0.468	-0.998
MQ	0.020	0.088	-0.428	-0.323	-0.942	-0.988
REBLUP-BC	0.022	0.026	-0.286	-0.276	0.020	-0.318
MQ-BC	0.022	0.030	-0.276	-0.262	-0.068	-0.297
Median values of Relative Root MSE (expressed as a percentage)						
EBLUP	0.805	0.854	1.215	1.369	0.966	2.389
REBLUP	0.822	0.842	1.008	0.985	1.019	1.436
MQ	0.824	0.833	1.030	1.008	1.464	1.570
REBLUP-BC	0.913	0.918	1.232	1.240	0.859	1.270
MQ-BC	0.913	0.915	1.238	1.256	0.931	1.486

Table 2 Performance of Root MSE estimators in model-based simulation experiments.

Estimator	MSE Estimator	No outliers		Individual		Area	Both
Scenario/Areas		[0,0]	[0,u]/ 1-36	[e,0]	[e,u]/ 1-36	[0,u]/ 37-40	[e,u]/ 37-40
		Median values of Relative Bias (expressed as a percentage)					
EBLUP	Prasad-Rao	-0.34	3.82	1.74	11.32	-17.31	-40.86
	CCT	3.61	1.55	31.24	5.95	2.15	-3.05
	CCST	5.64	4.78	33.95	8.52	77.26	8.28
REBLUP	CCT	-17.71	-20.24	-15.76	-19.51	-34.79	-36.63
	CCST	-2.01	-5.31	-8.46	-7.91	-3.58	-22.51
	Bootstrap	-1.19	7.38	-4.42	11.37	-19.42	-31.44
MQ	CCT	-2.98	-12.56	-16.29	-24.02	6.69	177.42
	CCST	0.11	-7.77	-8.21	-14.10	8.95	163.38
REBLUP-BC	CCT	-10.56	-11.88	-12.46	-12.57	-10.54	-18.37
	CCST	12.98	12.19	7.79	7.90	13.63	4.67
	Bootstrap	-0.21	-0.52	-6.76	-4.90	-1.25	-12.96
MQ-BC	CCT	-6.35	-7.19	3.48	1.87	3.92	5.96
	CCST	-7.18	-7.42	-11.38	-11.42	3.21	-9.20
		Median values of Relative Root MSE (expressed as a percentage)					
EBLUP	Prasad-Rao	6.24	7.20	18.57	22.28	17.90	43.19
	CCT	31.51	31.25	76.20	61.57	28.37	51.30
	CCST	26.65	15.20	66.72	29.28	88.30	39.97
REBLUP	CCT	29.52	28.67	30.82	29.00	28.58	38.70
	CCST	27.86	20.89	28.47	20.25	22.87	29.24
	Bootstrap	10.27	10.67	34.92	16.61	14.62	33.04
MQ	CCT	61.94	59.88	61.50	59.67	43.76	205.30
	CCST	54.77	50.63	49.14	45.34	40.58	189.92
REBLUP-BC	CCT	33.64	33.21	45.20	45.48	33.56	47.18
	CCST	38.14	37.65	51.03	50.34	37.63	53.71
	Bootstrap	10.12	10.20	15.27	14.53	10.60	18.35
MQ-BC	CCT	36.68	36.19	65.37	65.70	38.33	64.26
	CCST	33.93	33.55	44.81	44.65	35.30	50.55
		Median values of Coverage Rate (expressed as a percentage)					
EBLUP	Prasad-Rao	95	96	95	96	90	73
	CCT	90	90	92	93	93	91
	CCST	94	96	96	96	98	97
REBLUP	CCT	87	86	86	86	86	78
	CCST	92	92	90	92	94	86
	Bootstrap	94	96	92	96	93	83
MQ	CCT	72	69	67	66	90	91
	CCST	79	79	81	80	92	92
REBLUP-BC	CCT	86	85	86	86	85	82
	CCST	91	91	91	91	92	88
	Bootstrap	95	95	93	94	95	91
MQ-BC	CCT	86	86	82	82	87	86
	CCST	87	87	86	86	88	88

Table 3 Median values of the relative bias (RB) and relative root mean squared error (RRMSE) generated by estimators in design-based simulation. All values are expressed as percentages and are over the regions of interest.

Estimator	RB(%)	RRMSE(%)
EBLUP	10.79	35.18
REBLUP	-13.08	30.59
MQ	-22.98	35.07
REBLUP-BC	-4.13	31.94
MQ-BC	-6.17	31.57

Table 4 Performance of Root MSE estimators in design-based simulation: median values of the percentage relative bias, relative root MSE and coverage rate (for nominal 95 per cent interval). Intervals are defined by the small area mean estimate plus or minus twice their corresponding estimated root mean squared error.

Estimator\MSE estimator	Prasad-Rao	CCT	CCST	Bootstrap
Median values of Relative Bias (expressed as a percentage)				
EBLUP	6.37	1.79	5.85	
REBLUP		-23.06	3.59	32.12
MQ		-31.59	-24.48	
REBLUP-BC		-14.58	3.51	0.48
MQ-BC		-6.40	-11.01	
Median values of Relative Root MSE (expressed as a percentage)				
EBLUP	30.61	30.67	28.16	
REBLUP		45.79	43.72	61.95
MQ		62.19	55.88	
REBLUP-BC		39.78	43.13	39.81
MQ-BC		45.53	38.38	
Median values of Coverage Rate (expressed as a percentage)				
EBLUP	96	94	94	
REBLUP		82	91	99
MQ		70	81	
REBLUP-BC		82	92	95
MQ-BC		78	83	

Figure 1 Area specific values of true RMSE (solid line) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the PR estimator are indicated by △, those for the CCT estimator are indicated by ◯, and those for the CCST estimator are indicated by +. Plots show results for the EBLUP predictor.

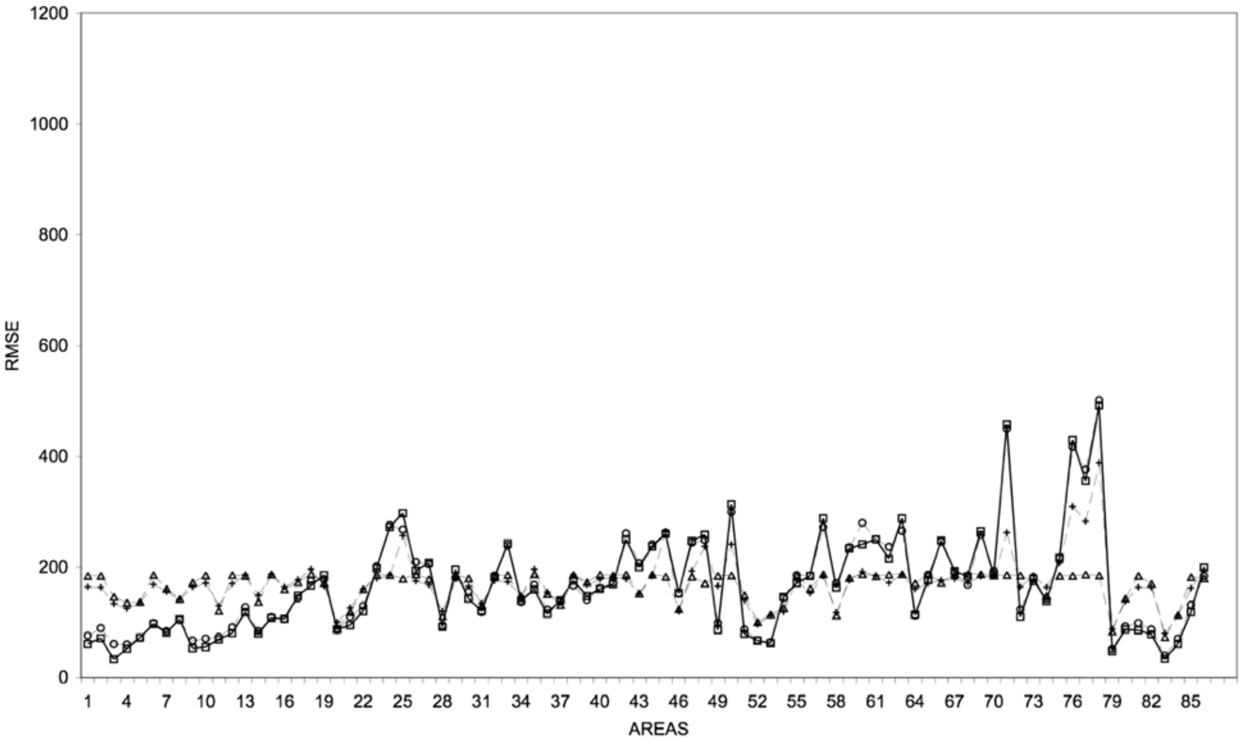

Figure 2 Area specific values of true RMSE (solid line) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the CCT estimator are indicated by ○, those for the CCST estimator are indicated by +, while those for the MSE bootstrap estimator are indicated by ◇. Plots show results REBLUP (top) and REBLUP-BC (bottom) predictors.

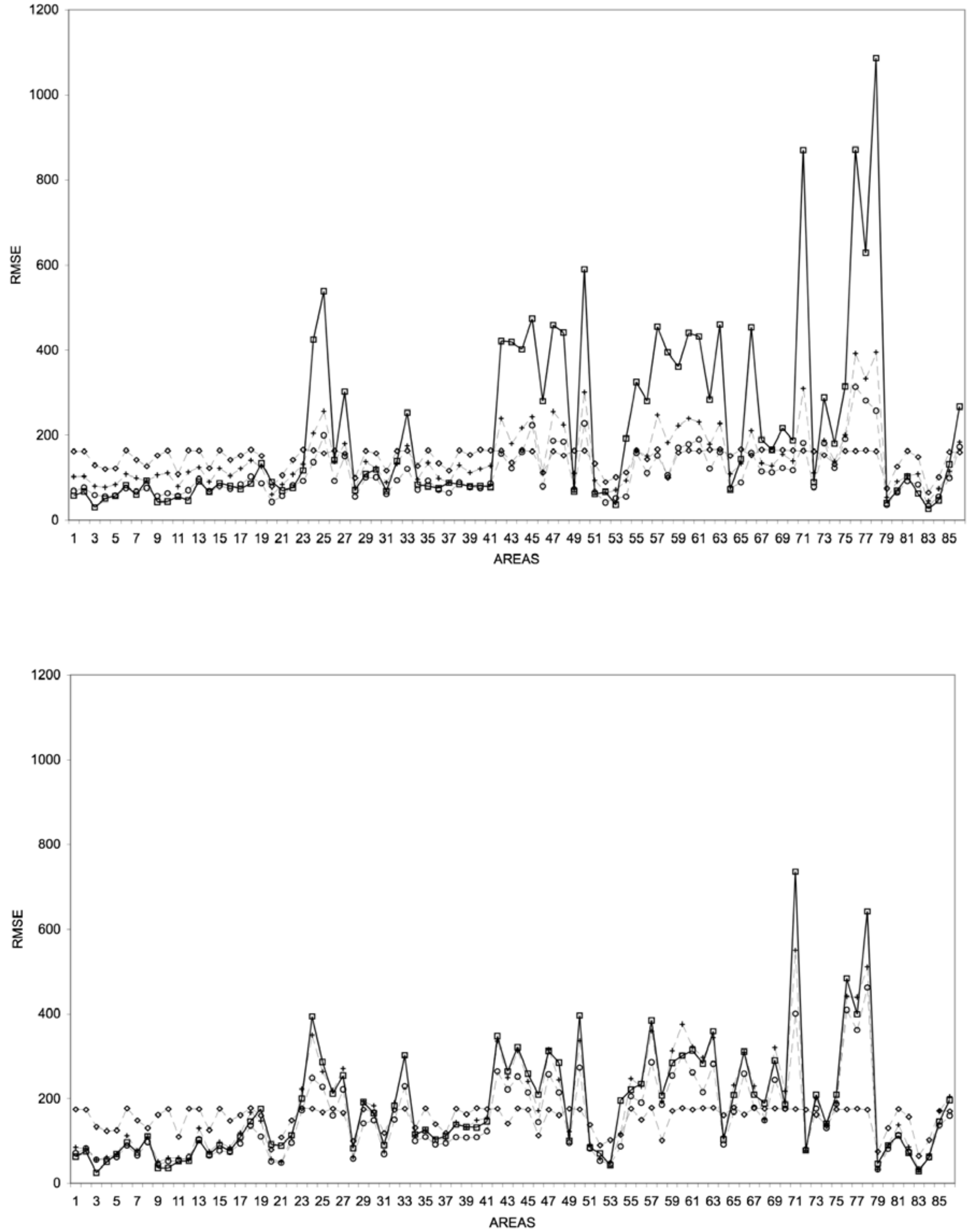

Figure 3 Area specific values of true RMSE (solid line) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the CCT estimator are indicated by ○, while those for the CCST estimator are indicated by +. Plots show results MQ (top) and MQ-BC (bottom) predictors.