Statistics and Operations Research Transactions,

vol. 44, n. 1 (2020)

SORT 44 (1) January-June 2020, 3-38 DOI: 10.2436/20.8080.02.93

Small area estimation of additive parameters under unit-level generalized linear mixed models

Tomáš Hobza[1], Yolanda Marhuenda[2] and Domingo Morales[2]

Abstract

Average incomes and poverty proportions are additive parameters obtained as averages of a given function of an income variable. As the variable income has an asymmetric distribution, it is not properly modelled via normal distributions. When dealing with this type of variable, a first option is to apply transformations that approximate normality. A second option is to use non-symmetric distributions from the exponential family. This paper proposes unit-level generalized linear mixed models for modelling asymmetric positive variables and for deriving three types of predictors of small area additive parameters, called empirical best, marginal and plug-in. The parameters of the introduced model are estimated by applying the maximum likelihood method to the Laplace approximation of the likelihood. The mean squared errors of the predictors are estimated by parametric bootstrap. The introduced methodology is applied and illustrated under unit-level gamma mixed models. Some simulation experiments are carried out to study the behaviour of the fitting algorithm, the small area predictors and the bootstrap estimator of the mean squared errors. By using data of the Spanish living condition survey of 2013, an application to the estimation of average incomes and poverty proportions in counties of the region of Valencia is given.

MSC: 62J12 Generalized linear models; 62P25 Applications to social sciences; 62D05 Sampling theory, sample surveys.

Keywords: Average income, poverty proportion, generalized linear mixed models, empirical best predictor, mean squared error, bootstrap.

1 Introduction

Many of the socioeconomic indicators published by statistical offices are additive parameters. These parameters are the sums of the transformed values that an objective variable takes in the population units and its definition depends on the selected variable and transformation. This paper deals with the small area estimation (SAE) of additive parameters, with particularizations to average incomes and poverty proportions. The problems of SAE appear when the sample sizes are small in the target population subsets,

[1] Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Trojanova 13, 120 00 Prague 2, Czech Republic. (corresponding author, email: tomas.hobza@fjfi.cvut.cz)

[2] Operations Research Center, Miguel Hernández University of Elche, Spain.

Received: September 2018
Accepted: November 2019

called small areas or domains, so that the direct estimators are not reliable. A domain direct estimator is obtained by using only the domain data. The low amount of data can be overcome by using statistical models that introduce additional information via auxiliary variables, data from other domains and variance-covariance structures. Model-based predictors of domain parameters are generally more efficient than direct estimators. See the monograph of Rao and Molina (2015) for an introduction to SAE, linear mixed models (LMM) and related issues.

Average incomes and some income-based poverty indicators are sums of transformed individual incomes. Some SAE methods based on unit-level models have been proposed in the literature for this type of parameters. Elbers, Lanjouw and Lanjouw (2003) introduced estimators based on the predictions of a fitted marginal nested error regression (NER) model. Molina and Rao (2010) proposed empirical best predictors (EBP) by employing the predictions of a NER model conditioned to the observed sample. This approach was extended to two-fold NER models by Marhuenda et al. (2017). Hobza and Morales (2013) derived predictors of means of household normalized net annual incomes under random regression coefficient models. Molina, Nandram and Rao (2015) proposed a hierarchical Bayes approach and Guadarrama, Molina and Rao (2014) compared several poverty mapping methods based on unit level models. Hobza and Morales (2016), Hobza, Morales and Santamaría (2018) derived EBPs based on unit-level logit mixed models, Tzavidis et al. (2008), Chambers, Salvati and Tzavidis (2012, 2016) introduced predictors based on M-quantile regression models. Karlberg (2014) proposed log-transformation mixed small area prediction models incorporating a logistic component for skewed data in the presence of zeroes. Dreassi, Petrucci and Rocco (2014), Fabrizi, Ferrante and Trivisano (2017) and Moura, Silva and Neves (2017) gave hierarchical Bayes procedures for skewed survey data. By using temporal and spatio-temporal area-level models, Esteban et al. (2012a, 2012b), Marhuenda, Molina and Morales (2013) and Morales, Pagliarella and Salvatore (2015) derived also model-based predictors of poverty indicators. Boubeta, Lombardía and Morales (2016, 2017) introduced empirical best predictors (EBP) of poverty proportions based on Poisson mixed models. Further references can be found in Pratesi (2016). A common feature of the above cited references is the use of predictors based on generalized linear mixed models (GLMM).

This paper extends the EBP methodology of Molina and Rao (2010) by introducing predictors of additive parameters based on unit-level GLMMs. The introduced methodology is applied to the prediction of small area average incomes and poverty indicators under unit-level gamma mixed models (GMM). The GLMMs have random effects taking into account the between-domains variability that is not explained by the auxiliary variables. The random effects are usually assumed to be normally distributed. The maximum likelihood (ML) estimation of GLMM parameters have some computational difficulties because the likelihood may involve high-dimensional integrals which cannot be evaluated analytically. For calculating the ML estimators of model parameters, this paper maximizes the Laplace approximation to the log-likelihood.

The paper introduces EBPs for estimating domain additive parameters. The proposed EBPs are based on unit-level GLMMs. Two more predictors, called plug-in and marginal, are also considered and empirically studied in a simulation experiment.

The mean squared error (MSE), also called prediction variance in the model-based approach to SAE, is a standard accuracy measure for predictors of domain parameters. Hall and Maiti (2006a,b) introduced bootstrap estimators of MSEs of predictors of functions of fixed and random effects under SAE models. As we are interested in estimating small area additive parameters, we consider the parametric bootstrap estimator of the MSE introduced by González-Manteiga et al. (2007), but adapted to GLMMs. This approach was extended by González-Manteiga et al. (2008a,b) to nested error regression models and to multivariate area-level models respectively.

In the particular case of GMMs, we carry out simulation experiments for investigating the behaviour of the fitting method, the predictors of average incomes and poverty proportions and the parametric bootstrap estimator of the MSE. We present an application to data from the Spanish living conditions survey (SLCS) of 2013 in the region of Valencia (east of Spain). The target is the estimation of 2013 average incomes and poverty proportions at county level.

The extension of the methodology of Molina and Rao (2010), where the EBPs are introduced under unit-level LMMs, to unit-level GLMMs have three main mathematical and computational difficulties: (1) under LMMs, the distribution of the unobserved part of the vector of target variables conditioned to the observed part can be calculated explicitly, but not in the case of GLMMs; (2) the likelihoods of GLMMs are high dimensional integrals, so they need more specialized fitting algorithms; (3) it can not be assumed that the shape parameters (or shape function) of GLMMs are all equal to a known common constant, so a procedure for estimating them is needed. This paper faces these three issues by studying the applicability of two unit-level GLMMs to the estimation of small area additive parameters.

The paper is organized as follows. Section 2 introduces two unit-level GLMMs. As the shape parameters of the second model are known constants multiplied by a common parameter, this model cannot be fitted by using standard software; for example by using the glmer function of lme4 library of the R programming language (R Core Team 2019). This is why Section 3 describes the employed ML-Laplace algorithm that we have programmed in R for fitting the model. Sections 4, 5, 5.1 and 5.2 present the empirical best, the marginal and the plug-in predictors of functions of model effects, additive parameters, means and poverty proportions respectively. The calculation of the EBPs uses a census file as auxiliary information. It is shown that this restriction can be avoided if the auxiliary variables are categorical. In that case, it is sufficient to have the population sizes of the domains crossed with the categories. Section 6 gives a parametric bootstrap method for estimating the MSE. Section 7 presents three simulation experiments. Simulation 1 analyses the behaviour of the fitting algorithm. Simulation 2 compares the performances of the three introduced predictors. Simulation 3 empirically studies the parametric bootstrap estimators of the MSEs. Section 8 applies the developed

methodology to unit-level data from the 2013 SLCS and takes the aggregated auxiliary information from the Spanish Labour Force Survey (SLFS). The target is the estimation of 2013 average incomes and poverty proportions at county level. Section 9 gives some conclusions. The paper contains two appendices. Appendix A gives the components of the updating equation of the ML-Laplace algorithm for the GMM. Appendix B presents some complementary tables and figures for the application to real data.

2 The unit-level generalized linear mixed models

This section introduces two unit-level GLMMs. Let D denote the number of small areas (or domains) under consideration. Both models have a set of random area effects $\{v_d : d = 1,\dots,D\}$ that are i.i.d. $N(0,1)$. In matrix notation, we have $\boldsymbol{v} = \underset{1\le d\le D}{\text{col}}(v_d) \sim N_D(\boldsymbol{0},\boldsymbol{I}_D)$, i.e.

$$f_v(\boldsymbol{v}) = (2\pi)^{-D/2}\exp\Big\{-\frac{1}{2}\boldsymbol{v}^{\top}\boldsymbol{v}\Big\}.$$

For $d = 1,\dots,D$, $j = 1,\dots,n_d$, the GLMMs assume that the conditional distribution of the target variable y_{dj}, conditioned to v_d, belongs to the exponential family, i.e. $y_{dj}|_{v_d} \sim \text{Exp}(\theta_{dj},\nu_{dj};a,b,c)$, with the probability density function (p.d.f.)

$$f(y_{dj}|v_d) = f(y_{dj}|\theta_{dj},\nu_{dj};a,b,c) = \exp\Big\{\frac{y_{dj}\theta_{dj}-b(\theta_{dj})}{a(\nu_{dj})}+c(y_{dj},\nu_{dj})\Big\}, \qquad (1)$$

where $a(\cdot) > 0$, $b(\cdot)$ and $c(\cdot)$ are known real-valued functions specifying the selected distribution and $\nu_{dj} > 0$. Further, we assume that $b(\cdot)$ is one-to-one and three times continuously differentiable with one-to-one first derivative. This is to say, we consider a nested data structure where subindexes d and j denote domain and unit (within domain) respectively and n_d is the sample size of domain d. Under (1), the expectation and variance of y_{dj}, given v_d, are

$$\mu_{dj} = E[y_{dj}|v_d] = \frac{\partial b(\theta_{dj})}{\partial\theta_{dj}} = \dot{b}(\theta_{dj}) \quad \text{var}[y_{dj}|v_d] = a(\nu_{dj})\frac{\partial^2 b(\theta_{dj})}{\partial\theta_{dj}^2} = a(\nu_{dj})\ddot{b}(\theta_{dj}).$$

Parameters μ_{dj} and ν_{dj} are called mean and shape parameters respectively. For a twice continuously differentiable and monotonous link function $g(\cdot)$ of the mean parameter, we assume that

$$\eta_{dj} = g(\mu_{dj}) = \boldsymbol{x}_{dj}^{\top}\boldsymbol{\beta}+\phi v_d, \quad d = 1,\dots,D,\ j = 1,\dots,n_d,$$

where $\phi > 0$ is a standard deviation parameter, $\boldsymbol{\beta} = \underset{1\le k\le p}{\text{col}}(\beta_k)$ is a vector of regression parameters and $\boldsymbol{x}_{dj} = \underset{1\le k\le p}{\text{col}}(x_{djk})$ is a vector of auxiliary variables which are assumed to

be constant (fixed regression design). Further, we assume that the y_{dj}'s are independent conditioned to $\boldsymbol{v}$. The sample size is $n = \sum_{d=1}^{D} n_d$ and the domain target vector is $\boldsymbol{y}_d = \underset{1\le j\le n_d}{\text{col}}(y_{dj})$. The conditional p.d.f. of $\boldsymbol{y} = \underset{1\le d\le D}{\text{col}}(\boldsymbol{y}_d)$, given $\boldsymbol{v}$, and the marginal p.d.f. of $\boldsymbol{y}$ are

$$f(\boldsymbol{y}|\boldsymbol{v}) = \prod_{d=1}^{D}\prod_{j=1}^{n_d} f(y_{dj}|v_d), \quad f(\boldsymbol{y}) = \int_{R^D} f(\boldsymbol{y}|\boldsymbol{v}) f_v(\boldsymbol{v})\, d\boldsymbol{v}.$$

Let us note that the assumption of normality of the random effects is typical for mixed models used in SAE. Sinha and Rao (2009) and Benavent and Morales (2016) carried out simulation experiments to investigate the robustness of EBLUPs of linear parameters against deviations from the hypothesis of normality under nested error regression and Fay-Herriot models respectively. They showed that EBLUPs works well when deviations are small, but their behaviour become poor when deviations are big. Similar conclusions hold also for EBP under the presented model. A specific comment concerning this issue is given in Remark 8.1 in Section 8.

An example of unit-level GLMM is the GMM, where

$$y_{dj}|v_d \sim \text{Gamma}(\nu_{dj}, \mu_{dj}/\nu_{dj}), \quad d = 1,\dots,D,\ j = 1,\dots,n_d.$$

For $y_{dj} > 0$, the conditioned p.d.f. is

$$f(y_{dj}|v_d) = \left(\frac{\nu_{dj}}{\mu_{dj}}\right)^{\nu_{dj}} \frac{y_{dj}^{\nu_{dj}-1}}{\Gamma(\nu_{dj})} \exp\left\{-\frac{\nu_{dj}}{\mu_{dj}} y_{dj}\right\} \tag{2}$$

$$= \exp\left\{\frac{y_{dj}\left(-\frac{1}{\mu_{dj}}\right) - \log\mu_{dj}}{\frac{1}{\nu_{dj}}} + \nu_{dj}\log\nu_{dj} - \log\Gamma(\nu_{dj}) + (\nu_{dj}-1)\log y_{dj}\right\}.$$

Under (2), the expectation and variance of y_{dj}, given v_d, are

$$E[y_{dj}|v_d] = \frac{\nu_{dj}}{\nu_{dj}/\mu_{dj}} = \mu_{dj}, \quad \text{var}[y_{dj}|v_d] = \frac{\nu_{dj}}{\nu_{dj}^2/\mu_{dj}^2} = \frac{\mu_{dj}^2}{\nu_{dj}}.$$

The natural parameter and the functions $a(\cdot) > 0$, $b(\cdot)$ and $c(\cdot)$ of GMMs are

$$\theta_{dj} = -\tfrac{1}{\mu_{dj}}, \qquad b(\theta_{dj}) = \log\mu_{dj} = \log\left(-\tfrac{1}{\theta_{dj}}\right) = -\log(-\theta_{dj}),$$
$$a(\nu_{dj}) = 1/\nu_{dj}, \quad c(y_{dj},\nu_{dj}) = \nu_{dj}\log\nu_{dj} - \log\Gamma(\nu_{dj}) + (\nu_{dj}-1)\log y_{dj}.$$

For the mean parameter in GMMs, we consider the link function

$$\eta_{dj} = g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \boldsymbol{x}_{dj}^{\mathrm{T}}\boldsymbol{\beta} + \phi v_d, \quad d = 1,\dots,D. \tag{3}$$

Depending on the assumptions on the shape parameters, we consider two GLMMs. Model 1 assumes that $\nu_{dj} = \nu > 0$, $d = 1,\dots,D$, $j = 1,\dots,n_d$ and ν is unknown. Model 2 assumes that $\nu_{dj} = a_{dj}\varphi$ with $a_{dj} > 0$ known and $\varphi > 0$ unknown, $d = 1,\dots,D$, $j = 1,\dots,n_d$. This is to say, Model 1 is Model 2 with $a_{dj} = 1$ and $\nu = \varphi > 0$ unknown. Under the gamma distribution (2) with the link function (3), these models are called gamma Model 1 and 2 respectively. For some distributions of the exponential family, Model 1 can be fitted with the glmer function of lme4 library of the R programming language. However, glmer cannot be applied to estimate the parameters of Model 2. Section 3 presents the ML-Laplace algorithm for fitting GLMMs, with a particularization to gamma Model 2.

Model 1 is quite popular in Gamma regression modelling. Under Model 1, the conditioned variance is $\mathrm{var}[y_{dj}|v_d] = \nu^{-1}\mu_{dj}^2$. The direct proportionality to the mean is a rigid condition that sometimes does not allow a good fit of the model to the data. This fact was observed in the application to real data and motivated the use of Model 2. Under Model 2, a good selection of a_{dj} for the conditioned variance will produce a better fit of the GMM to the data. This is illustrated in Section 8.

Alternative link functions for gamma regression models are $g(\mu_{dj}) = \mu_{dj}$ and $g(\mu_{dj}) = \log\mu_{dj}$. The link function (3) allows giving linear predictors of the natural parameter θ_{dj} and moreover it is the canonical link function for the Gamma distribution which implies some good properties of the ML estimators. That is why we investigate GMMs with the inverse link function in the simulations and we use it in the application to real data.

3 The Laplace approximation algorithm

This section describes an approximation of the loglikelihood of GLMMs and the corresponding algorithm for estimating the unknown parameters of Model 2. In what follows $\psi^{-1}(\cdot)$ denotes the inverse mapping of a one-to-one real valued function $\psi(\cdot)$. As $\mu_{dj} = g^{-1}(\eta_{dj})$ and $\theta_{dj} = (\dot{b})^{-1}(\mu_{dj})$, it holds that

$$\frac{\partial\mu_{dj}}{\partial\eta_{dj}} = \frac{1}{\frac{\partial\eta_{dj}}{\partial\mu_{dj}}} = \frac{1}{\dot{g}(\mu_{dj})}, \quad \frac{\partial\theta_{dj}}{\partial\mu_{dj}} = \frac{1}{\frac{\partial\mu_{dj}}{\partial\theta_{dj}}} = \frac{1}{\ddot{b}(\theta_{dj})}, \quad \frac{\partial\eta_{dj}}{\partial v_d} = \frac{\partial(\boldsymbol{x}_{dj}^{\intercal}\boldsymbol{\beta} + \phi v_d)}{\partial v_d} = \phi.$$

Therefore

$$\frac{\partial\mu_{dj}}{\partial v_d} = \frac{\partial\mu_{dj}}{\partial\eta_{dj}}\frac{\partial\eta_{dj}}{\partial v_d} = \frac{\phi}{\dot{g}(\mu_{dj})}, \quad \frac{\partial\dot{g}(\mu_{dj})}{\partial v_d} = \frac{\partial\dot{g}(\mu_{dj})}{\partial\mu_{dj}}\frac{\partial\mu_{dj}}{\partial v_d} = \ddot{g}(\mu_{dj})\frac{\phi}{\dot{g}(\mu_{dj})},$$

$$\frac{\partial\theta_{dj}}{\partial v_d} = \frac{\partial\theta_{dj}}{\partial\mu_{dj}}\frac{\partial\mu_{dj}}{\partial v_d} = \frac{\phi}{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})}, \quad \frac{\partial b(\theta_{dj})}{\partial v_d} = \frac{\partial b(\theta_{dj})}{\partial\theta_{dj}}\frac{\partial\theta_{dj}}{\partial v_d} = \frac{\phi\dot{b}(\theta_{dj})}{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})}.$$

The vectors $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_D$ are unconditionally independent with marginal p.d.f.

$$f(\boldsymbol{y}_d) = \int_{-\infty}^{\infty} \prod_{j=1}^{n_d} f(y_{dj}|v_d) f(v_d)\, dv_d$$

$$= \kappa_d \int_{-\infty}^{\infty} \exp\Big\{ -\frac{v_d^2}{2} + \sum_{j=1}^{n_d} \frac{y_{dj}\theta_{dj} - b(\theta_{dj})}{a(\nu_{dj})} \Big\}\, dv_d = \kappa_d \int_{-\infty}^{\infty} \exp\big\{h(v_d)\big\}\, dv_d,$$

where $\kappa_d = (2\pi)^{-1/2} \exp\Big\{ \sum_{j=1}^{n_d} c(y_{dj},\nu_{dj}) \Big\}$,

$$h(v_d) = -\frac{v_d^2}{2} + \sum_{j=1}^{n_d} \frac{y_{dj}\theta_{dj} - b(\theta_{dj})}{a(\nu_{dj})}, \tag{4}$$

$$\begin{aligned} \dot{h}(v_d) &= -v_d + \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})} \Big\{ \frac{\phi y_{dj}}{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})} - \frac{\phi \dot{b}(\theta_{dj})}{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})} \Big\}, \\ &= -v_d + \phi \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})} \big(y_{dj} - \mu_{dj}\big), \end{aligned}$$

and

$$\begin{aligned} \ddot{h}(v_d) &= -1 + \phi^2 \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})\ddot{b}^2(\theta_{dj})\dot{g}^2(\mu_{dj})} \\ &\cdot \left\{ -\frac{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})}{\dot{g}(\mu_{dj})} - (y_{dj} - \mu_{dj}) \left[\frac{\dddot{b}(\theta_{dj})}{\ddot{b}(\theta_{dj})\dot{g}(\mu_{dj})} \dot{g}(\mu_{dj}) + \ddot{b}(\theta_{dj}) \frac{\ddot{g}(\mu_{dj})}{\dot{g}(\mu_{dj})} \right] \right\} \\ &= -1 - \phi^2 \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})\ddot{b}^2(\theta_{dj})\dot{g}^2(\mu_{dj})} \left\{ \ddot{b}(\theta_{dj}) + (y_{dj} - \mu_{dj}) \left[\frac{\dddot{b}(\theta_{dj})}{\ddot{b}(\theta_{dj})} + \ddot{b}(\theta_{dj}) \frac{\ddot{g}(\mu_{dj})}{\dot{g}(\mu_{dj})} \right] \right\}. \end{aligned}$$

The Laplace algorithm looks for v_{0d} maximizing the function $h(v_d)$, i.e. such that $\dot{h}(v_{0d}) = 0$ and $\ddot{h}(v_{0d}) < 0$. The Laplace approximation to $f(\boldsymbol{y}_d)$ is

$$\begin{aligned} f(\boldsymbol{y}_d) \approx \Bigg| 1 + \phi^2 \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})\ddot{b}^2(\theta_{0dj})\dot{g}^2(\mu_{0dj})} \Big\{ \ddot{b}(\theta_{0dj}) + (y_{dj} - \mu_{0dj}) \Big[\frac{\dddot{b}(\theta_{0dj})}{\ddot{b}(\theta_{0dj})} \\ + \ddot{b}(\theta_{0dj}) \frac{\ddot{g}(\mu_{0dj})}{\dot{g}(\mu_{0dj})} \Big] \Big\} \Bigg|^{-1/2} \cdot \exp\Big\{ -\frac{v_{0d}^2}{2} + \sum_{j=1}^{n_d} \frac{y_{dj}\theta_{0dj} - b(\theta_{0dj})}{a(\nu_{dj})} \Big\} \exp\Big\{ \sum_{j=1}^{n_d} c(y_{dj},\nu_{dj}) \Big\}, \end{aligned}$$

where $\theta_{0dj} = (\dot{b})^{-1}(\mu_{0dj})$, $\mu_{0dj} = g^{-1}(\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_{0d})$. The GLMM loglikelihood is $\ell = \sum_{d=1}^{D} \ell_d$ and

$$\ell_d = \log f(\boldsymbol{y}_d) \approx \ell_{0d} = \sum_{j=1}^{n_d} c(y_{dj}, \nu_{dj}) - \frac{1}{2}\log \xi_{0d} - \frac{v_{0d}^2}{2} + \sum_{j=1}^{n_d} \frac{y_{dj}\theta_{0dj} - b(\theta_{0dj})}{a(\nu_{dj})}, \quad (5)$$

where

$$\xi_{0d} = \left| 1 + \phi^2 \sum_{j=1}^{n_d} \frac{1}{a(\nu_{dj})} \left\{ \frac{1}{\ddot{b}(\theta_{0dj})\dot{g}^2(\mu_{0dj})} + (y_{dj} - \mu_{0dj}) \left[\frac{\dddot{b}(\theta_{0dj})}{\ddot{b}^3(\theta_{0dj})\dot{g}^2(\mu_{0dj})} + \frac{\ddot{g}(\mu_{0dj})}{\ddot{b}(\theta_{0dj})\dot{g}^3(\mu_{0dj})} \right] \right\} \right|.$$

For the GMMs, we have $\theta_{dj} = -\frac{1}{\mu_{dj}}$, $g(\mu_{dj}) = \frac{1}{\mu_{dj}}$, $\dot{g}(\mu_{dj}) = -\frac{1}{\mu_{dj}^2}$, $\ddot{g}(\mu_{dj}) = \frac{2}{\mu_{dj}^3}$, $b(\theta_{dj}) = -\log(-\theta_{dj})$, $\dot{b}(\theta_{dj}) = -\frac{1}{\theta_{dj}} = \mu_{dj}$, $\ddot{b}(\theta_{dj}) = \frac{1}{\theta_{dj}^2} = \mu_{dj}^2$, $\dddot{b}(\theta_{dj}) = -\frac{2}{\theta_{dj}^3} = 2\mu_{dj}^3$,

$$h(v_d) = -\frac{v_d^2}{2} + \sum_{j=1}^{n_d} \left\{ \nu_{dj} \log(\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_d) - \nu_{dj} y_{dj} (\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_d) \right\},$$

$$\dot{h}(v_d) = -v_d + \sum_{j=1}^{n_d} \left\{ \frac{\nu_{dj}\phi}{\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_d} - \phi \nu_{dj} y_{dj} \right\} = -v_d + \phi \sum_{j=1}^{n_d} \nu_{dj}(\mu_{dj} - y_{dj}),$$

$$\ddot{h}(v_d) = -\left(1 + \phi^2 \sum_{j=1}^{n_d} \frac{\nu_{dj}}{(\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_d)^2} \right) = -\left(1 + \phi^2 \sum_{j=1}^{n_d} \nu_{dj}\mu_{dj}^2 \right).$$

In this particular case, it holds that $\ddot{h}(v_d) < 0$ for all possible values of v_d. The components of the Laplace approximation to the GMM loglikelihood are

$$\ell_{0d} = \sum_{j=1}^{n_d} \left\{ \nu_{dj} \log \nu_{dj} + (\nu_{dj} - 1)\log y_{dj} - \log \Gamma(\nu_{dj}) \right\} - \frac{1}{2}\log \xi_{0d} - \frac{v_{0d}^2}{2} + \sum_{j=1}^{n_d} \left\{ \nu_{dj} \log(\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_{0d}) - \nu_{dj} y_{dj} (\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_{0d}) \right\}, \quad (6)$$

where $\xi_{0d} = 1 + \phi^2 \sum_{j=1}^{n_d} \nu_{dj}\mu_{0dj}^2$ and $\mu_{0dj} = (\boldsymbol{x}_{dj}^{\mathsf{T}}\boldsymbol{\beta} + \phi v_{0d})^{-1}$. Under gamma Model 2, i.e. under the assumption $\nu_{dj} = a_{dj}\varphi$, Appendix A gives the partial derivatives of ℓ_{0d} with respect to the components of $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\mathsf{T}}, \phi, \varphi)^{\mathsf{T}}$. It also gives the score vector $\boldsymbol{U}_0(\boldsymbol{\theta})$ and the Hessian matrix $\boldsymbol{H}_0(\boldsymbol{\theta})$ containing the first and the second partial derivatives of ℓ_{0d} respectively.

A first Newton-Raphson algorithm maximizes $\ell_0(\boldsymbol{\theta}) = \sum_{d=1}^{D} \ell_{0d}$, with fixed $v_d = v_{0d}$, $d = 1,\ldots,D$. The updating equation is

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \boldsymbol{H}_0^{-1}(\boldsymbol{\theta}^{(k)})\boldsymbol{U}_0(\boldsymbol{\theta}^{(k)}). \tag{7}$$

For $d = 1,\ldots,D$, a second Newton-Raphson algorithm maximizes $h(v_d) = h(v_d,\boldsymbol{\theta})$, defined in (4), with $\boldsymbol{\theta} = (\boldsymbol{\beta}^\intercal, \phi, \varphi)^\intercal = \boldsymbol{\theta}_0$ fixed. The updating equation is

$$v_d^{(k+1)} = v_d^{(k)} - \frac{\dot{h}(v_d^{(k)},\boldsymbol{\theta}_0)}{\ddot{h}(v_d^{(k)},\boldsymbol{\theta}_0)}. \tag{8}$$

Algorithm. By combining the two Newton-Raphson algorithms, the ML-Laplace approximation algorithm for Model 2 is obtained. The steps are

1. Set the initial values $i = 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, $\varepsilon_3 > 0$, $\varepsilon_4 > 0$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\theta}^{(-1)} = \boldsymbol{\theta}^{(0)} + \mathbf{1}$, $v_d^{(0)} = 0$, $v_d^{(-1)} = 1$, $d = 1,\ldots,D$.
2. Until $|v_d^{(i)} - v_d^{(i-1)}| < \varepsilon_3$, $d = 1,\ldots,D$, $\|\boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}^{(i-1)}\|_2 < \varepsilon_4$, do
   (a) Apply algorithm (8) with seeds $v_d^{(i)}$, $d = 1,\ldots,D$, convergence tolerance ε_1 and $\boldsymbol{\theta} = \boldsymbol{\theta}^{(i)}$ fixed. Output: $v_d^{(i+1)}$, $d = 1,\ldots,D$.
   (b) Apply algorithm (7) with seed $\boldsymbol{\theta}^{(i)}$, convergence tolerance ε_2 and $v_{0d} = v_d^{(i+1)}$ fixed, $d = 1,\ldots,D$. Output: $\boldsymbol{\theta}^{(i+1)}$.
   (c) $i \leftarrow i+1$.
3. Output: $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(i)}$, $\hat{v}_d = v_d^{(i)}$, $d = 1,\ldots,D$.

To get some algorithm seed $\boldsymbol{\theta}^{(0)}$, we can e.g. fit Model 1 (this can be done for several distributions from the exponential family by using the glmer function of the R statistical package lme4) to obtain the estimates $\tilde{\boldsymbol{\beta}}$, $\tilde{\phi}$ and $\tilde{\nu}$ and use the seeds $\boldsymbol{\beta}^{(0)} = \tilde{\boldsymbol{\beta}}$, $\phi^{(0)} = \tilde{\phi}$ and $\varphi^{(0)} = \tilde{\nu}/\overline{a}_d$, where $\overline{a}_d = 1/n_d \sum_{j=1}^{n_d} a_{dj}$. Let us also note that the Laplace approximation algorithm gives at convergence not only estimators of the model parameters but also the mode predictors, $\hat{v}_d$, of the random effects.

4 Predictors of functions of model effects

This section considers a finite population U of N elements partitioned into D domains U_d of size N_d, $d = 1,\ldots,D$. From the population, a sample s of size n is selected with subsamples s_d of sizes n_d from domains U_d. Let $\boldsymbol{y}_d = \underset{1\le j\le N_d}{\text{col}}\,(y_{dj})$ be the random vector

containing the values of a target variable on the N_d units of domain d. Let $\boldsymbol{y}_{ds}$ be the sub-vector of $\boldsymbol{y}_d$ corresponding to the units in the sample s_d and $\boldsymbol{y}_{dr}$ the sub-vector of domain units in the non-sampled domain population $U_d - s_d$. By reordering the domain units, we can write $\boldsymbol{y}_d = (\boldsymbol{y}_{ds}^\top, \boldsymbol{y}_{dr}^\top)^\top$. We define $\boldsymbol{y}_s = \underset{1\le d\le D}{\text{col}}(\boldsymbol{y}_{ds})$ and $\boldsymbol{y}_r = \underset{1\le d\le D}{\text{col}}(\boldsymbol{y}_{dr})$ and we follow a fully model-based approach by assuming that $\boldsymbol{y} = (\boldsymbol{y}_s^\top, \boldsymbol{y}_r^\top)^\top$ follows Model 2. We are thus assuming a non-informative sampling setup.

The conditional distribution of $\boldsymbol{y}_s$, given $\boldsymbol{v}$, is

$$f(\boldsymbol{y}_s|\boldsymbol{v}) = \prod_{d=1}^{D} f(\boldsymbol{y}_{ds}|v_d),$$

where

$$f(\boldsymbol{y}_{ds}|v_d) = \exp\Big\{\sum_{j=1}^{n_d} \frac{y_{dj}\theta_{dj} - b(\theta_{dj})}{a(\nu_{dj})}\Big\} \exp\Big\{\sum_{j=1}^{n_d} c(y_{dj}, \nu_{dj})\Big\}$$

and the p.d.f. of $\boldsymbol{v}$ is

$$f(\boldsymbol{v}) = \prod_{d=1}^{D} f(v_d), \quad f(v_d) = (2\pi)^{-1/2} \exp\big\{-\frac{1}{2} v_d^2\big\}.$$

The conditional distribution of $\boldsymbol{y}_{dr}$, given $\boldsymbol{y}_{ds}$, is

$$f(\boldsymbol{y}_{dr}|\boldsymbol{y}_{ds}) = \frac{f(\boldsymbol{y}_{dr}, \boldsymbol{y}_{ds})}{f(\boldsymbol{y}_{ds})} = \frac{\int_R f(\boldsymbol{y}_{dr}, \boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}$$

$$= \frac{\int_R f(\boldsymbol{y}_{dr}|v_d) f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}. \tag{9}$$

The aim of this section is to introduce the EBP and the plug-in predictor of μ_{dj} and $\bar{\mu}_d = \frac{1}{N_d}\sum_{j=1}^{N_d} \mu_{dj}$ under Model 2. The corresponding predictors under Model 1 can be obtained in a similar way.

If $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi, \varphi)^\top$ is known, the best predictor of $\mu_{dj} = \mu_{dj}(\boldsymbol{\theta}, v_d) = g^{-1}(\boldsymbol{x}_{dj}^\top \boldsymbol{\beta} + \phi v_d)$ is $\hat{\mu}_{dj}(\boldsymbol{\theta}) = E_\theta[\mu_{dj}|\boldsymbol{y}_s]$. In this case, we have that $E_\theta[\mu_{dj}|\boldsymbol{y}_s] = E_\theta[\mu_{dj}|\boldsymbol{y}_{ds}]$ and

$$\hat{\mu}_{dj}(\boldsymbol{\theta}) = E_\theta[\mu_{dj}|\boldsymbol{y}_{ds}] = \frac{\int_R g^{-1}(\boldsymbol{x}_{dj}^\top \boldsymbol{\beta} + \phi v_d) f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d} = \frac{A_{dj}(\boldsymbol{y}_{ds}, \boldsymbol{\theta})}{B_d(\boldsymbol{y}_{ds}, \boldsymbol{\theta})},$$

where $A_{dj} = A_{dj}(\boldsymbol{y}_{ds}, \boldsymbol{\theta})$ and $B_d = B_d(\boldsymbol{y}_{ds}, \boldsymbol{\theta})$ are

$$A_{dj} = \int_R g^{-1}(\boldsymbol{x}_{dj}^\top \boldsymbol{\beta} + \phi v_d) \exp\Big\{ \sum_{i=1}^{n_d} \frac{y_{di}\theta_{di} - b(\theta_{di})}{a(\nu_{di})} \Big\} f(v_d)\, dv_d,$$
$$B_d = \int_R \exp\Big\{ \sum_{i=1}^{n_d} \frac{y_{di}\theta_{di} - b(\theta_{di})}{a(\nu_{di})} \Big\} f(v_d)\, dv_d. \tag{10}$$

The EBP of μ_{dj} is $\hat{\mu}_{dj}(\hat{\boldsymbol{\theta}})$ and can be approximated by the following Monte Carlo procedure.

1. Estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\phi}, \hat{\varphi})^\top$ and put $\hat{\nu}_{dj} = a_{dj}\hat{\varphi}$.
2. For $\ell = 1, \ldots, L$, generate $v_d^{(\ell)}$ i.i.d. $N(0,1)$ and $v_d^{(L+\ell)} = -v_d^{(\ell)}$.
3. Calculate the approximation of the EBP $\hat{\mu}_{dj} = \hat{A}_{dj}/\hat{B}_d$, where

$$\hat{A}_{dj} = \frac{1}{2L} \sum_{\ell=1}^{2L} g^{-1}(\boldsymbol{x}_{dj}\hat{\boldsymbol{\beta}} + \hat{\phi} v_d^{(\ell)}) \exp\Big\{ \sum_{i=1}^{n_d} \frac{y_{di}\hat{\theta}_{di}^{(\ell)} - b(\hat{\theta}_{di}^{(\ell)})}{a(\hat{\nu}_{di})} \Big\},$$
$$\hat{B}_d = \frac{1}{2L} \sum_{\ell=1}^{2L} \exp\Big\{ \sum_{i=1}^{n_d} \frac{y_{di}\hat{\theta}_{di}^{(\ell)} - b(\hat{\theta}_{di}^{(\ell)})}{a(\hat{\nu}_{di})} \Big\} \tag{11}$$

and $\hat{\theta}_{di}^{(\ell)} = (\dot{b})^{-1}(\hat{\mu}_{di}^{(\ell)})$ for $\hat{\mu}_{di}^{(\ell)} = g^{-1}(\boldsymbol{x}_{di}^\top \hat{\boldsymbol{\beta}} + \hat{\phi} v_d^{(\ell)})$.

The derived best predictors have minimum MSE in the class of unbiased estimators. Unfortunately, this property does not hold for EBPs which are obtained by substituting the true parameters by their estimates and therefore they are not unbiased. The EBPs are asymptotically unbiased under the assumption that the estimates of the model parameters are consistent but the domain sample sizes are usually small in SAE problems. Thus, it make sense to empirically investigate the behaviour of the plug-in predictors which are less computationally demanding. The plug-in predictor of μ_{dj} is

$$\tilde{\mu}_{dj} = g^{-1}(\boldsymbol{x}_{dj}^\top \hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d), \tag{12}$$

where $\hat{\boldsymbol{\beta}}, \hat{\phi}$ and $\hat{v}_d$ are taken from the output of the ML-Laplace approximation algorithm. The EBP and the plug-in predictor of $\bar{\mu}_d = \frac{1}{N_d}\sum_{j=1}^{N_d} \mu_{dj}$ are

$$\hat{\bar{\mu}}_d^E = \hat{\bar{\mu}}_d^E(\hat{\boldsymbol{\theta}}) = \frac{1}{N_d} \sum_{j=1}^{N_d} \hat{\mu}_{dj}, \quad \hat{\bar{\mu}}_d^P = \frac{1}{N_d} \sum_{j=1}^{N_d} \tilde{\mu}_{dj}. \tag{13}$$

5 Predictors of additive parameters

This section introduces predictors of additive parameters of small areas under Model 2. Similar predictors can be obtained under Model 1. An additive domain parameter is

$$\delta_d = \frac{1}{N_d}\sum_{j=1}^{N_d} h(y_{dj}),$$

where h is a known measurable function. If $j \in s_d$, then $E_\theta[h(y_{dj})|\boldsymbol{y}_{ds}] = h(y_{dj})$. If $j \in U_d - s_d$, then $f(y_{dj}|\boldsymbol{y}_{ds})$ is obtained from (9). Therefore, the best predictor of δ_d is

$$\hat{\delta}_d^B = \frac{1}{N_d}\sum_{j=1}^{N_d} E_\theta\big[h(y_{dj})|\boldsymbol{y}_{ds}\big] = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{dj}) + \sum_{j\in U_d - s_d} E_\theta\big[h(y_{dj})|\boldsymbol{y}_{ds}\big]\Big\},$$

where

$$E_\theta[h(y_{dj})|\boldsymbol{y}_{ds}] = \frac{\int_R\int_R h(y_{dj}) f(\boldsymbol{y}_{ds}|v_d) f(y_{dj}|v_d) f(v_d)\, dy_{dj}\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d} = \frac{A_{hdj}(\boldsymbol{y}_{ds},\boldsymbol{\theta})}{B_d(\boldsymbol{y}_{ds},\boldsymbol{\theta})},$$

$$A_{hdj} = \int_R\int_R h(y_{dj}) \exp\Big\{\sum_{i=1}^{n_d} \frac{y_{di}\theta_{di} - b(\theta_{di})}{a(\nu_{di})}\Big\} f(y_{dj}|v_d) f(v_d)\, dy_{dj}\, dv_d$$

and B_d was defined in (10). The EBP of of δ_d is

$$\hat{\delta}_d^E = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{dj}) + \sum_{j\in U_d - s_d} E_{\hat{\theta}}\big[h(y_{dj})|\boldsymbol{y}_{ds}\big]\Big\}, \quad E_{\hat{\theta}}\big[h(y_{dj})|\boldsymbol{y}_{ds}\big] = \frac{A_{hdj}(\boldsymbol{y}_{ds},\hat{\boldsymbol{\theta}})}{B_d(\boldsymbol{y}_{ds},\hat{\boldsymbol{\theta}})}.$$

If $j \in U_d - s_d$, then the numerator and denominator of $E_{\hat{\theta}}[h(y_{dj})|\boldsymbol{y}_{ds}]$ can be approximated by Monte Carlo simulation. The numerator is a bivariate integral that can be written in the form of two iterative univariate integrals. Therefore, we implement an iterative Monte Carlo algorithm which approximates the inner integral by simulating positive random numbers $y_{dj}^{(\ell_1,\ell_2)}$ from $f(y_{dj}|v_d^{(\ell_1)})$ and approximates the outer integral by simulating random numbers $v_d^{(\ell_1)}$ from $f(v_d)$. Under Model 2, the iterative Monte Carlo algorithm is

1. Estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\intercal, \hat{\phi}, \hat{\varphi})^\intercal$ and $\hat{\nu}_{dj} = a_{dj}\hat{\varphi}$.
2. For $\ell_1 = 1,\ldots,L_1$, generate $v_d^{(\ell_1)}$ i.i.d. $N(0,1)$ and $v_d^{(L_1+\ell_1)} = -v_d^{(\ell_1)}$, calculate $\hat{\mu}_{dj}^{(\ell_1)} = g^{-1}(\boldsymbol{x}_{dj}^\intercal\hat{\boldsymbol{\beta}} + \hat{\phi}v_d^{(\ell_1)})$ and $\hat{\theta}_{dj}^{(\ell_1)} = (\dot{b})^{-1}(\hat{\mu}_{dj}^{(\ell_1)})$. For $\ell_1 = 1,\ldots,2L_1$, $\ell_2 = 1,\ldots,L_2$, generate $y_{dj}^{(\ell_1,\ell_2)} \sim \mathrm{Exp}\big(\hat{\theta}_{dj}^{(\ell_1)}, \hat{\nu}_{dj}; a, b, c\big)$.

3. Approximate the EBP $\hat{E}_{\hat{\theta}}[h(y_{dj})|\boldsymbol{y}_{ds}] \approx \hat{A}_{hdj}/\hat{B}_d$, where

$$\hat{A}_{hdj} = \frac{1}{2L_1L_2}\sum_{\ell_1=1}^{2L_1}\exp\Big\{\sum_{i=1}^{n_d}\frac{y_{di}\hat{\theta}_{di}^{(\ell_1)}-b(\hat{\theta}_{di}^{(\ell_1)})}{a(\hat{\nu}_{di})}\Big\}\sum_{\ell_2=1}^{L_2}h(y_{dj}^{(\ell_1,\ell_2)}),$$

$$\hat{B}_d = \frac{1}{2L_1}\sum_{\ell_1=1}^{2L_1}\exp\Big\{\sum_{i=1}^{n_d}\frac{y_{di}\hat{\theta}_{di}^{(\ell_1)}-b(\hat{\theta}_{di}^{(\ell_1)})}{a(\hat{\nu}_{di})}\Big\}. \tag{14}$$

The plug-in predictor of δ_d is

$$\hat{\delta}_d^P = \frac{1}{N_d}\Big\{\sum_{j\in s_d}h(y_{dj}) + \sum_{j\in U_d-s_d}h(\tilde{\mu}_{dj})\Big\}, \quad \tilde{\mu}_{dj} = g^{-1}(\boldsymbol{x}_{dj}^{\mathsf{T}}\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{\nu}_d).$$

Simulation 2 shows that the plug-in predictor does not work well in some situations. For this reason we propose another predictor of the additive domain parameter δ_d. Instead of using the conditional distribution deriving the EBPs, we consider the predicted marginal distribution of y_{dj} with parameters $\hat{\nu}_{dj}$ and $\tilde{\theta}_{dj} = (\dot{b})^{-1}(\tilde{\mu}_{dj})$, where $\tilde{\mu}_d$ is the plug-in predictor of μ_{dj}. This is to say, we consider the p.d.f. $f(y_{dj}|\tilde{\theta}_{dj},\hat{\nu}_{dj};a,b,c)$ from the exponential family. Based on the marginal distribution, we define the marginal predictor of δ_d,

$$\hat{\delta}_d^M = \frac{1}{N_d}\Big(\sum_{j\in s_d}h(y_{dj}) + \sum_{j\in U_d-s_d}\hat{h}_{dj}^M\Big),$$

where

$$\hat{h}_{dj}^M \triangleq E\left[h(y_{dj})|\,\tilde{\theta}_{dj},\hat{\nu}_{dj};a,b,c\right] = \int_R h(y)f(y|\tilde{\theta}_{dj},\hat{\nu}_{dj};a,b,c)\,dy.$$

For calculating the empirical best, the plug-in and the marginal predictors, we need two files: (1) a survey file with the unit-level sample data (main file), and (2) a census file containing the values of the employed explanatory variables in all the population units (auxiliary file). However, not all the values $\boldsymbol{x}_{dj}$, $d = 1,\ldots,D$, $j = 1,\ldots,N_d$, are available in many practical cases. If in addition some of the auxiliary variables are continuous, the three introduced predictors are not applicable. An important particular case, where these predictors can be calculated under the assumed fixed regression design, is when the number of values of the vector of auxiliary variables is finite and the a_{dj}'s take a common value a_{dk} in the domain d and the covariate class k. In this situation, called "categorical setup", we only need a smaller auxiliary file containing the aggregated (domain-level) values of the explanatory variables. More concretely, the categorical setup is

$$\boldsymbol{x}_{dj} \in \{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_K\},\ a_{dj} = a_{dk} \text{ if } \boldsymbol{x}_{dj} = \boldsymbol{z}_k, \quad j = 1,\ldots,N_d,\ d = 1,\ldots,D. \tag{15}$$

Under the categorical setup (15), the EBP of δ_d is

$$\hat{\delta}_d^E = \hat{\delta}_d^E(\hat{\boldsymbol{\theta}}) = \frac{1}{N_d}\left[\sum_{j\in s_d} h(y_{dj}) + \sum_{k=1}^{K}\sum_{i=1}^{N_{dk}-n_{dk}} E_{\hat{\theta}}\big[h(y_{dki})|\boldsymbol{y}_{ds}\big]\right]$$
$$= \frac{1}{N_d}\left[\sum_{j\in s_d} h(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})E_{\hat{\theta}}\big[h(y_{dk})|\boldsymbol{y}_{ds}\big]\right], \tag{16}$$

where the size N_{dk} of $U_{dk} = \{j \in U_d : \boldsymbol{x}_{dj} = \boldsymbol{z}_k\}$ is available from external data sources (aggregated auxiliary information), n_{dk} is the size of $s_{dk} = \{j \in s_d : \boldsymbol{x}_{dj} = \boldsymbol{z}_k\}$ and y_{dki} denotes the value that the target variable takes in the ith unit of the subset $U_{dk} - s_{dk}$. Under Model 2, the expectation $E_{\hat{\theta}}\big[h(y_{dk})|\boldsymbol{y}_{ds}\big]$ is approximated (similarly to (14)) by Monte Carlo integration as follows.

1. Estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\top}, \hat{\phi}, \hat{\varphi})^{\top}$ and put $\hat{\nu}_{dk} = a_{dk}\hat{\varphi}$.
2. For $\ell_1 = 1,\ldots,L_1$ generate $v_d^{(\ell_1)}$ i.i.d. $N(0,1)$ and $v_d^{(L_1+\ell_1)} = -v_d^{(\ell_1)}$, calculate $\hat{\mu}_{dk}^{(\ell_1)} = g^{-1}(\boldsymbol{z}_k^{\top}\hat{\boldsymbol{\beta}} + \hat{\phi}v_d^{(\ell_1)})$ and $\hat{\theta}_{dk}^{(\ell_1)} = (\dot{b})^{-1}(\hat{\mu}_{dk}^{(\ell_1)})$. For $\ell_1 = 1,\ldots,2L_1$, $\ell_2 = 1,\ldots,L_2$, generate $y_{dk}^{(\ell_1,\ell_2)} \sim \mathrm{Exp}\big(\hat{\theta}_{dk}^{(\ell_1)}, \hat{\nu}_{dk}; a, b, c\big)$.
3. Calculate $E_{\hat{\theta}}\big[h(y_{dk})|\boldsymbol{y}_{ds}\big] = \hat{A}_{hdk}/\hat{B}_d$, where

$$\hat{A}_{hdk} = \frac{1}{2L_1L_2}\sum_{\ell_1=1}^{2L_1}\exp\Big\{\sum_{i=1}^{n_d}\frac{y_{di}\hat{\theta}_{di}^{(\ell_1)} - b(\hat{\theta}_{di}^{(\ell_1)})}{a(\hat{\nu}_{di})}\Big\}\sum_{\ell_2=1}^{L_2} h(y_{dk}^{(\ell_1,\ell_2)}),$$
$$\hat{B}_d = \frac{1}{2L_1}\sum_{\ell=1}^{2L_1}\exp\Big\{\sum_{i=1}^{n_d}\frac{y_{di}\hat{\theta}_{di}^{(\ell_1)} - b(\hat{\theta}_{di}^{(\ell_1)})}{a(\hat{\nu}_{di})}\Big\}. \tag{17}$$

If (15) holds, the plug-in predictor of δ_d is

$$\hat{\delta}_d^P = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})h(\tilde{\mu}_{dk})\Big\}, \quad \tilde{\mu}_{dk} = g^{-1}(\boldsymbol{z}_k^{\top}\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d),$$

and the marginal predictor of δ_d is

$$\hat{\delta}_d^M = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})\hat{h}_{dk}^M\Big\}, \quad \hat{h}_{dk}^M = \int_R h(y) f(y|\tilde{\theta}_{dk}, \hat{\nu}_{dk}; a, b, c)dy.$$

5.1 Predictors of small area means

This section introduces predictors of the small area mean

$$\overline{Y}_d = \frac{1}{N_d}\sum_{j=1}^{N_d} y_{dj},$$

which is an additive parameter with $h(y) = y$. The best predictor of y_{dj} is $\hat{y}_{dj}(\boldsymbol{\theta}) = E_\theta[y_{dj}|\boldsymbol{y}_s]$. If $j \in s_d$, then $E_\theta[y_{dj}|\boldsymbol{y}_s] = y_{dj}$. If $j \in U_d - s_d$, then $E_\theta[y_{dj}|\boldsymbol{y}_s] = E_\theta[y_{dj}|\boldsymbol{y}_{ds}]$ and

$$\hat{y}_{dj}(\boldsymbol{\theta}) = E_\theta[y_{dj}|\boldsymbol{y}_{ds}] = \frac{\int_R \int_R y_{dj} f(y_{dj}|v_d) f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dy_{dj}\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}$$

$$= \frac{\int_R \mu_{dj} f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d}{\int_R f(\boldsymbol{y}_{ds}|v_d) f(v_d)\, dv_d} = \frac{A_{dj}(\boldsymbol{y}_{ds},\boldsymbol{\theta})}{B_d(\boldsymbol{y}_{ds},\boldsymbol{\theta})} = E_\theta[\mu_{dj}|\boldsymbol{y}_{ds}] = \hat{\mu}_{dj}(\boldsymbol{\theta}),$$

where $A_{dj} = A_{dj}(\boldsymbol{y}_{ds},\boldsymbol{\theta})$ and $B_d = B_d(\boldsymbol{y}_{ds},\boldsymbol{\theta})$ are defined in (10). The EBP of y_{dj} is $\hat{y}_{dj} = \hat{y}_{dj}(\hat{\boldsymbol{\theta}}) = \hat{\mu}_{dj}(\hat{\boldsymbol{\theta}})$. Thus, the EBP of y_{dj} is $\hat{y}_{dj} = y_{dj}$ if $j \in s_d$ and $\hat{y}_{dj} = \hat{\mu}_{dj}$ if $j \in U_d - s_d$, where $\hat{\mu}_{dj}$ is the EBP of μ_{dj} given in (11).

The EBP and the plug-in and marginal predictors of $\overline{Y}_d$ are

$$\hat{\overline{Y}}_d^E = \frac{1}{N_d}\Big[\sum_{j\in s_d} y_{dj} + \sum_{j\in U_d - s_d} \hat{\mu}_{dj}\Big], \hat{\overline{Y}}_d^P = \frac{1}{N_d}\Big[\sum_{j\in s_d} y_{dj} + \sum_{j\in U_d - s_d} \tilde{\mu}_{dj}\Big],$$

$$\hat{\overline{Y}}_d^M = \frac{1}{N_d}\Big[\sum_{j\in s_d} y_{dj} + \sum_{j\in U_d - s_d} \hat{\mu}_{dj}^M\Big],$$

where $\tilde{\mu}_{dj}$ is the plug-in predictor of μ_{dj} given in (12) and

$$\hat{\mu}_{dj}^M \triangleq E\left[y_{dj}|\tilde{\theta}_{dj},\hat{\nu}_{dj};a,b,c\right] = \int_R y f(y|\tilde{\theta}_{dj},\hat{\nu}_{dj};a,b,c)\, dy = g^{-1}(\boldsymbol{x}_{dj}^\mathrm{T}\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d) = \tilde{\mu}_{dj},$$

so that $\hat{\overline{Y}}_d^M = \hat{\overline{Y}}_d^P$. Under the categorical setup (15), the EBP and the plug-in predictors of $\overline{Y}_d$ are

$$\hat{\overline{Y}}_d^E = \frac{1}{N_d}\Big[\sum_{j\in s_d} y_{dj} + \sum_{k=1}^K (N_{dk} - n_{dk})\hat{\mu}_{dk}\Big], \quad \hat{\overline{Y}}_d^P = \frac{1}{N_d}\Big[\sum_{j\in s_d} y_{dj} + \sum_{k=1}^K (N_{dk} - n_{dk})\tilde{\mu}_{dk}\Big],$$

where $\tilde{\mu}_{dk} = g^{-1}(\boldsymbol{z}_k^\mathrm{T}\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d)$, $\hat{\mu}_{dk} = \hat{A}_{dk}^z/\hat{B}_d$, $\hat{B}_d$ is defined in (10) and $\hat{A}_{dk}^z$ is the Monte-Carlo approximation of

$$A_{dk}^{z} = \int_R g^{-1}(z_k^{\intercal}\boldsymbol{\beta} + \phi v_d) \exp\Big\{ \sum_{j=1}^{n_d} \frac{y_{dj}\theta_{dj} - b(\theta_{dj})}{a(\nu_{dj})} \Big\} f(v_d)\, dv_d.$$

This is to say, $\hat{A}_{dk}^{z}$ can be calculated as

$$\hat{A}_{dk}^{z} = \frac{1}{2L} \sum_{\ell=1}^{2L} g^{-1}(z_k^{\intercal}\hat{\boldsymbol{\beta}} + \hat{\phi} v_d^{(\ell)}) \exp\Big\{ \sum_{j=1}^{n_d} \frac{y_{dj}\hat{\theta}_{dj}^{(\ell)} - b(\hat{\theta}_{dj}^{(\ell)})}{a(\nu_{dj})} \Big\},$$

where $v_d^{(\ell)}$ are i.i.d. $N(0,1)$ and $v_d^{(L+\ell)} = -v_d^{(\ell)}$, $\ell = 1,\ldots,L$, $\hat{\theta}_{dj}^{(\ell)} = (\dot{b})^{-1}(\hat{\mu}_{dj}^{(\ell)})$ and $\hat{\mu}_{dj}^{(\ell)} = g^{-1}(\boldsymbol{x}_{dj}^{\intercal}\hat{\boldsymbol{\beta}} + \hat{\phi} v_d^{(\ell)})$.

5.2 Predictors of poverty proportions

This section deals with the estimation of domain poverty proportions, which are the proportion of people in the domain whose welfare is below the poverty line. Let y_{dj} be a welfare variable (i.e. income or expenditure) for individual j from domain d and let z be the poverty line. Then, the poverty proportion is the additive parameter

$$p_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h_0(y_{dj}),$$

where $h_0(y_{dj}) = I\big(y_{dj} < z\big)$. The EBP of p_d is

$$\hat{p}_d^{E} = \frac{1}{N_d}\Big(\sum_{j\in s_d} h_0(y_{dj}) + \sum_{j\in U_d - s_d} E_{\hat{\theta}}\left[h_0(y_{dj})|\boldsymbol{y}_{ds}\right] \Big),$$

where $E_{\hat{\theta}}\left[h_0(y_{dj})|\boldsymbol{y}_{ds}\right]$ is calculated by applying (14) with $h = h_0$.

The plug-in and the marginal predictors of p_d are

$$\hat{p}_d^{P} = \frac{1}{N_d}\Big(\sum_{j\in s_d} h_0(y_{dj}) + \sum_{j\in U_d - s_d} h_0(\tilde{\mu}_{dj}) \Big),\ \ \hat{p}_d^{M} = \frac{1}{N_d}\Big(\sum_{j\in s_d} h_0(y_{dj}) + \sum_{j\in U_d - s_d} \hat{p}_{dj}^{M} \Big),$$

where $\tilde{\mu}_{dj} = g^{-1}(\boldsymbol{x}_{dj}^{\intercal}\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d)$ and

$$\begin{aligned} \hat{p}_{dj}^{M} &= E\left[I(y_{dj} < z)|\tilde{\theta}_{dj}, \hat{\nu}_{dj}; a, b, c\right] = P\Big(\mathrm{Exp}(\tilde{\theta}_{dj}, \hat{\nu}_{dj}; a, b, c) < z\Big) \\ &\triangleq F(z|\tilde{\theta}_{dj}, \hat{\nu}_{dj}; a, b, c). \end{aligned}$$

In the previous formula, $F(\cdot)$ denotes the cumulative distribution function of the corresponding distribution from the exponential family.

Under the categorical setup (15), the EBP of p_d is

$$\hat{p}_d^E = \frac{1}{N_d}\left[\sum_{j\in s_d} h_0(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})E_{\hat{\theta}}\left[h_0(y_{dk})|\boldsymbol{y}_{ds}\right]\right],$$

where $E_{\hat{\theta}}\left[h_0(y_{dk})|\boldsymbol{y}_{ds}\right]$ is calculated by applying (17) with $h = h_0$. The marginal and the plug-in predictors of p_d are

$$\hat{p}_d^M = \frac{1}{N_d}\Big[\sum_{j\in s_d} h_0(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})\hat{p}_{dk}^M\Big],$$

$$\hat{p}_d^P = \frac{1}{N_d}\Big[\sum_{j\in s_d} h_0(y_{dj}) + \sum_{k=1}^{K}(N_{dk}-n_{dk})h_0(\tilde{\mu}_{dk})\Big],$$

where $\hat{p}_{dk}^M = F(z|\tilde{\theta}_{dk},\hat{\nu}_{dk};a,b,c)$, $\tilde{\theta}_{dk} = (\dot{b})^{-1}(\tilde{\mu}_{dk})$, $\hat{\nu}_{dk} = a_{dk}\hat{\varphi}$ and $\tilde{\mu}_{dk} = g^{-1}(\boldsymbol{z}_k^\intercal\hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d)$.

6 Bootstrap estimation of the MSE

This section presents a parametric bootstrap estimator of the MSE of $\hat{\bar{\mu}}_d^E$ and $\hat{\delta}_d^E$ applicable to the categorical setup (15). Under Model 2, the algorithm steps are

1. Fit the model to the sample and calculate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\intercal, \hat{\phi}, \hat{\varphi})^\intercal$, put $\hat{\nu}_{dk} = a_{dk}\hat{\varphi}$.
2. Repeat B times ($b = 1,\ldots,B$):

   (a) The population. For $d = 1,\ldots,D$, $k = 1,\ldots,K$ generate $v_d^{*(b)}$ i.i.d. $N(0,1)$ and calculate $\mu_{dk}^{*(b)} = g^{-1}\left(\boldsymbol{z}_k^\intercal\hat{\boldsymbol{\beta}} + \hat{\phi}v_d^{*(b)}\right)$ and $\theta_{dk}^{*(b)} = (\dot{b})^{-1}(\mu_{dk}^{*(b)})$. For $j = 1,\ldots,N_d$ generate

   $$y_{dj}^{*(b)} \sim \mathrm{Exp}\left(\theta_{dk}^{*(b)},\hat{\nu}_{dk};a,b,c\right), \quad \text{where } k \text{ is such that } \boldsymbol{x}_{dj} = \boldsymbol{z}_k.$$

   Calculate the true bootstrap quantities

   $$\bar{\mu}_d^{*(b)} = \bar{\mu}_d(\hat{\boldsymbol{\theta}}, v_d^{*(b)}) = \frac{1}{N_d}\sum_{k=1}^{K} N_{dk}\mu_{dk}^{*(b)}, \quad \delta_d^{*(b)} = \frac{1}{N_d}\sum_{j=1}^{N_d} h(y_{dj}^{*(b)}).$$

   (b) The sample. The bootstrap sample has the same units as the real data sample. It is not extracted at random. The model is on the population, therefore

the source of randomness comes from the generation of the population. For each bootstrap sample, calculate $\hat{\boldsymbol{\theta}}^{*(b)}$ and the EBPs

$$\hat{\bar{\mu}}_d^{E*(b)} = \hat{\bar{\mu}}_d^E(\hat{\boldsymbol{\theta}}^{*(b)}), \quad \hat{\delta}_d^{E*(b)} = \hat{\delta}_d^E(\hat{\boldsymbol{\theta}}^{*(b)}).$$

3. Output: $mse^*(\hat{\bar{\mu}}_d^E) = \frac{1}{B}\sum_{b=1}^B \left(\hat{\bar{\mu}}_d^{E*(b)} - \bar{\mu}_d^{*(b)}\right)^2$, $mse^*(\hat{\delta}_d^E) = \frac{1}{B}\sum_{b=1}^B \left(\hat{\delta}_d^{E*(b)} - \delta_d^{*(b)}\right)^2$.

The above algorithm can be easily modified to the case of fitting a GLMM with at least one continuous auxiliary variable. For this sake, a census file is needed with the values of $\boldsymbol{x}_{dj}$ for all the units of the population. In addition, the census file must have the same unit identifier variable as the sample file. This modification is equivalent to adapting the parametric bootstrap method of González-Manteiga et al. (2007) to the current unit-level GLMMs.

7 Simulation experiments

This section presents three simulation experiments for gamma Model 2. Simulation 1 analyses the behaviour of the ML-Laplace approximation algorithm for estimating parameters. Simulation 2 compares the performances of the EBPs, the plug-in predictors and the marginal predictors. Finally, Simulation 3 empirically studies the bootstrap estimators of the MSEs.

In all the experiments data are simulated in the following way. For $d = 1,\ldots,D$ and $j = 1,\ldots,n_d$, define regressors representing four possible classes of labour status. This is to say, $(x_{dj1},x_{dj2}) = (0,0)$ for unemployed, $(x_{dj1},x_{dj2}) = (0,1)$ for employed, $(x_{dj1},x_{dj2}) = (1,0)$ for inactive and $(x_{dj1},x_{dj2}) = (1,1)$ for ≤ 15. Generate $(x_{dj1},x_{dj2}) \in \{(0,0),(0,1),(1,0),(1,1)\}$ with probabilities $p_{00} = 0.1 + \frac{(d-1)}{D-1}0.2$, $p_{01} = 0.5 - \frac{(d-1)}{D-1}0.2$, $p_{10} = 0.2$, and $p_{11} = 0.2$, respectively. For each covariate class and domain d, the constants a_{dj} (which are assumed to be known in Model 2) are generated independently from a normal distribution with mean 1.5 and standard deviation 0.2 in order to be close to the values appearing in the application to the real data. The auxiliary variables, x_{dj1}, x_{dj2}, and the shape constants, a_{dj}, are generated before starting the simulation loop, so they are constant in the three simulation experiments. The model parameters are taken as $\beta_0 = 0.8$, $\beta_1 = -0.15$, $\beta_2 = 0.2$, $\phi = 0.1$ and $\varphi = 2.5$.

Within each iteration, the three simulation algorithms generate the random effect $v_d \sim N(0,1)$, $d = 1,\ldots,D$, and the income variable $y_{dj} \sim \text{Gamma}(\nu_{dj}, \frac{\nu_{dj}}{\mu_{dj}})$, where $\nu_{dj} = a_{dj}\varphi$ and

$$\mu_{dj} = (\beta_0 + x_{dj1}\beta_1 + x_{dj2}\beta_2 + \phi v_d)^{-1}, \ d = 1,\ldots,D, \ j = 1,\ldots,n_d.$$

7.1 Simulation 1

The target of Simulation 1 is to check the behaviour of the fitting algorithm. We take $D = 30, 60, 120, 180$ and $n_d = 10, 25, 50$. The steps of simulation 1 are

1. Repeat $I = 1000$ times ($i = 1, \ldots, I$)
   1.1. Generate a sample $\{y_{dj}^{(i)} : d = 1, \ldots, D,\ j = 1 \ldots, n_d\}$ from Model 2.
   1.2. Calculate $\hat{\beta}_0^{(i)}$, $\hat{\beta}_1^{(i)}$, $\hat{\beta}_2^{(i)}$, $\hat{\phi}^{(i)}$, $\hat{\varphi}^{(i)}$.
2. Output: For $\theta \in \{\beta_0, \beta_1, \beta_2, \phi, \varphi\}$, calculate the relative bias and relative root-MSE, i.e.

$$RBIAS = \frac{1}{|\theta|}\frac{1}{I}\sum_{i=1}^{I}(\hat{\theta}^{(i)} - \theta),\ RRMSE = \frac{1}{|\theta|}\Big(\frac{1}{I}\sum_{i=1}^{I}(\hat{\theta}^{(i)} - \theta)^2\Big)^{1/2}.$$

Tables 1, 2 and 3 presents the results of the simulation experiment for $n_d = 10$, $n_d = 25$ and $n_d = 50$ respectively. The relative bias is basically negligible. The relative root-MSE decreases as D or n_d increases. Simulation 1 empirically illustrates the consistency of the implemented ML-Laplace approximation algorithm.

Table 1: *RBIAS (left) and RRMSE (right) in % for $n_d = 10$.*

	RBIAS				RRMSE			
	$D = 30$	$D = 60$	$D = 120$	$D = 180$	$D = 30$	$D = 60$	$D = 120$	$D = 180$
$\hat{\beta}_0$	0.9932	0.6967	0.7074	0.6254	6.1918	4.5492	3.3014	2.5208
$\hat{\beta}_1$	0.3361	0.5187	−0.0944	0.0275	32.9073	24.1085	16.6784	13.0500
$\hat{\beta}_2$	−0.2629	0.0504	0.0892	0.3631	25.7537	18.1448	12.4199	9.8741
$\hat{\phi}$	−11.1515	−3.9887	−0.5083	0.1928	41.0300	27.3513	18.3789	15.0615
$\hat{\varphi}$	1.5736	0.9871	0.5349	0.3370	8.5560	6.0136	4.1093	3.4281

Table 2: *RBIAS (left) and RRMSE (right) in % for $n_d = 25$.*

	RBIAS				RRMSE			
	$D = 30$	$D = 60$	$D = 120$	$D = 180$	$D = 30$	$D = 60$	$D = 120$	$D = 180$
$\hat{\beta}_0$	0.8245	0.9708	0.8585	0.9259	4.2377	3.2867	2.2593	1.9043
$\hat{\beta}_1$	−0.4093	0.3290	0.9417	0.2324	20.0357	14.1744	10.1065	8.3685
$\hat{\beta}_2$	0.4219	0.2353	−0.1183	0.0690	14.8132	11.1234	7.7572	6.3451
$\hat{\phi}$	−4.1686	−2.7112	−1.3648	−0.7450	22.0699	15.8384	10.9345	8.6654
$\hat{\varphi}$	0.6263	0.2041	0.2512	0.2021	5.1735	3.5982	2.5533	2.0827

Table 3: *RBIAS (left) and RRMSE (right) in % for $n_d = 50$.*

	RBIAS				RRMSE			
	$D=30$	$D=60$	$D=120$	$D=180$	$D=30$	$D=60$	$D=120$	$D=180$
$\hat{\beta}_0$	0.9193	1.2258	1.1286	1.2008	3.4704	2.6633	2.0673	1.8352
$\hat{\beta}_1$	0.1209	0.2595	0.0680	−0.0676	14.3711	10.0332	7.3633	5.8110
$\hat{\beta}_2$	0.5331	−0.0164	−0.0452	0.0878	11.1165	7.7034	5.3727	4.4457
$\hat{\phi}$	−5.3541	−2.2712	−1.3643	−1.1853	18.5763	12.1674	8.7474	6.9339
$\hat{\varphi}$	0.2439	0.1738	0.1221	0.0634	3.4888	2.5094	1.8352	1.4415

7.2 Simulation 2

The target of Simulation 2 is to investigate the behaviour of the EBP, the marginal and plug-in predictors of the mean $\overline{Y}_d$ and the poverty proportion p_d. Before starting the simulation loop, a first set of target variables $\{y_{dj}^{(0)} : d = 1,\ldots,D,\ j = 1,\ldots,N_d\}$ is generated and the poverty threshold z is taken as the first sample quartile of these variables.

The model is fitted by the ML-Laplace approximation algorithm. The EBP is approximated with $L_1 = L_2 = 100$ Monte Carlo iterations and the domain sizes are $N_d = 1000$, $d = 1,\ldots,D$. The steps of Simulation 2 are

1. Repeat $I = 10^4$ times ($i = 1,\ldots,I$)

   1.1. Generate the i-th population $\{y_{dj}^{(i)} : d = 1,\ldots,D,\ j = 1\ldots,N_d\}$ in the same way as the sample in Simulation 1.

   1.2. Calculate $\overline{Y}_d^{(i)} = \frac{1}{N_d}\sum_{j=1}^{N_d} y_{dj}^{(i)}$, $p_d^{(i)} = \frac{1}{N_d}\sum_{j=1}^{N_d} I(y_{dj}^{(i)} < z)$, $d = 1,\ldots,D$.

   1.3. For $d = 1,\ldots,D$, select a sample $s_d^{(i)}$ of size n_d. The indexes of the samples $s_d^{(i)}$ remain constant across the iterations. Calculate $\hat{\beta}_0^{(i)}$, $\hat{\beta}_1^{(i)}$, $\hat{\beta}_2^{(i)}$, $\hat{\phi}^{(i)}$, $\hat{\varphi}^{(i)}$.

   1.4 Calculate the predictors $\hat{\overline{Y}}_d^{E,i}$, $\hat{\overline{Y}}_d^{P,i}$, $\hat{\overline{Y}}_d^{M,i}$, $\hat{p}_d^{E,i}$, $\hat{p}_d^{P,i}$, $\hat{p}_d^{M,i}$ and the direct estimators $\hat{\overline{Y}}_d^{dir,i} = \frac{1}{n_d}\sum_{j=1}^{n_d} y_{dj}^{(i)}$, $\hat{p}_d^{dir,i} = \frac{1}{n_d}\sum_{j=1}^{n_d} I(y_{dj}^{(i)} < z)$, $d = 1,\ldots,D$.

2. For $\xi_d^{(i)} \in \{\overline{Y}_d^{(i)}, p_d^{(i)}\}$ and $\hat{\xi}_d^{(i)} \in \{\hat{\overline{Y}}_d^{dir,i}, \hat{\overline{Y}}_d^{E,i}, \hat{\overline{Y}}_d^{P,i}, \hat{\overline{Y}}_d^{M,i}, \hat{p}_d^{dir,i}, \hat{p}_d^{E,i}, \hat{p}_d^{P,i}, \hat{p}_d^{M,i}\}$, $d = 1,\ldots,D$, calculate the performance measures

$$\xi_d = \frac{1}{I}\sum_{i=1}^{I}\xi_d^{(i)},\ RE_d = \left(\frac{1}{I}\sum_{i=1}^{I}\left(\hat{\xi}_d^{(i)} - \xi_d^{(i)}\right)^2\right)^{1/2},\ RB_d = \frac{1}{|\xi_d|}\frac{1}{I}\sum_{i=1}^{I}(\hat{\xi}_d^{(i)} - \xi_d^{(i)}),$$

$$RRE_d = \frac{RE_d}{|\xi_d|},\ RB = \frac{1}{D}\sum_{d=1}^{D}|RB_d|,\ RRE = \frac{1}{D}\sum_{d=1}^{D}RRE_d.$$

Table 4 presents the results of Simulation 2 for sample sizes $n_d = 10, 25, 50, 75, 100$, which are of similar magnitude to most of the county sample sizes presented in Table 10. For estimating the domain mean, the EBP $\hat{\overline{Y}}_d^E$ has had lower *RB* but higher *RRE* than the marginal/plug-in predictor $\hat{\overline{Y}}_d^P = \hat{\overline{Y}}_d^M$. Further, the model-based predictors are more efficient than the direct estimator. For estimating the domain proportion, the plug-in estimator is not recommended and the marginal and EB predictors are both more efficient than the direct estimator. The marginal predictor has obtained slightly better results than the EBP in relative bias and root-MSE. This fact can be explained by the variability derived from approximating the EBP with $L_1 = L_2 = 100$ Monte Carlo iterations.

Table 4: *RB (left) and RRE (right) in % for* $D = 30$.

n_d	10	25	50	75	100	10	25	50	75	100
$\hat{\overline{Y}}_d^{Dir}$	4.27	2.18	1.63	1.43	0.98	17.69	10.99	7.71	6.22	5.25
$\hat{\overline{Y}}_d^E$	0.33	0.31	0.38	0.23	0.25	11.11	8.50	6.57	5.54	4.82
$\hat{\overline{Y}}_d^P$	0.79	0.53	0.46	0.24	0.25	11.09	8.45	6.54	5.46	4.75
$\hat{\overline{Y}}_d^M$	0.79	0.53	0.46	0.24	0.25	11.09	8.45	6.54	5.46	4.75
$\hat{p}_d^{Dir}$	7.53	4.21	3.03	2.49	1.78	55.10	34.32	24.03	19.33	16.37
$\hat{p}_d^E$	0.57	0.56	0.68	0.43	0.39	21.18	16.62	13.17	11.29	10.06
$\hat{p}_d^P$	98.99	97.50	95.01	92.55	90.04	101.78	100.22	97.65	95.15	92.55
$\hat{p}_d^M$	0.76	0.49	0.31	0.27	0.21	21.17	16.56	13.08	11.18	9.97

Remark 7.1 *Since it is complicated to calculate analytically the error of the Monte Carlo approximation because we approximate the integrals of numerator and denominator, we tried to study the accuracy numerically. Namely, for one choice of* $n_d = 25$ *and one selected iteration of Simulation 2 we have approximated the EBP 1000 times for each domain and different values of L. Then we calculated the standard deviation of these approximations in each domain. This quantity express the variability of the Monte Carlo approximations. The means of the obtained standard deviations over areas are presented in Table 5 for the two cases of predicting the area mean* $\overline{Y}_d$ *and the area proportion* p_d.

Table 5: *Standard deviations of 1000 MC approximations of EBP for* $n_d = 25$ *and different values of L.*

Predictor	$L = 50$	$L = 100$	$L = 200$	$L = 300$	$L = 500$
$\hat{\overline{Y}}_d^E$	0.01370	0.00960	0.00665	0.00552	0.00421
$\hat{p}_d^E$	0.00593	0.00380	0.00250	0.00202	0.00152

In the simulation experiments we have used $L = 100$ *from time reasons but from the Table 5 it follows that in practical applications a higher value of L, e.g.* $L = 300$, *could be recommended. Higher values of L than 300 increase substantially the computing time.*

As a consequence of the results of Simulation 2, we will consider only the marginal predictor in Simulation 3 and we recommend using the marginal predictor in applications to real data.

For explaining the poor behaviour of plug-in predictors $\hat{p}_d^P$ of domain proportions p_d, we recall that

$$\hat{p}_d^P = \frac{1}{N_d}\Big[\sum_{j\in s_d} I(y_{dj} < z) + \sum_{k=1}^{K}(N_{dk} - n_{dk})I(\tilde{\mu}_{dk} < z)\Big],$$

where $K = 4$ and the poverty threshold z is taken as the lower sample quartile of the generated target variables. Since $\tilde{\mu}_{dk}$ is the predictor of the expectation of the target variable distribution, then $\tilde{\mu}_{dk}$ tends to be greater than z in most of the iterations of the simulation experiment. This is to say, the probability that one of the terms $I(\tilde{\mu}_{dk} < z)$ is equal to 1 is very small. In other words, the observed value of a Bernoulli variable is a bad estimator of the probability of success. This means that the summands in the second sum will be equal to zero with high probability and therefore the auxiliary information is almost not taken into account. Moreover, only the sample s_d is used in the first sum as in the case of direct estimators, but this sum is divided by the population size N_d instead of the sample size n_d.

7.3 Simulation 3

Simulation 3 investigates the behaviour of the bootstrap MSE estimator of the marginal predictors. We take $D = 30$, $n_d = 50$, $N_d = 1000$, $I = 500$ and $B = 25, 50, 100, 200, 300$, 400. We take $E_d = (RE_d)^2$ from the output of Simulation 2. The steps of Simulation 3 are

1. Repeat $I = 500$ times ($i = 1, \ldots, I$)

   1.1. Generate the population in the same way as sample in Simulation 2.

   1.2. For $d = 1, \ldots, D$, select a sample s_d of size n_d with fixed indexes and calculate $\hat{\beta}_0^{(i)}$, $\hat{\beta}_1^{(i)}$, $\hat{\beta}_2^{(i)}$, $\hat{\phi}^{(i)}$, $\hat{\varphi}^{(i)}$.

   1.3. For each $\hat{\xi}_d^{(i)} \in \{\hat{\bar{Y}}_d^{M,i}, \hat{p}_d^{M,i}\}$, calculate $mse_d^{*(i)} = mse_d^*(\hat{\xi}_d^{(i)})$.

2. Calculate the relative performance measures

$$Rb_d = \frac{1}{|E_d|}\frac{1}{I}\sum_{i=1}^{I}(mse_d^{*(i)} - E_d),\ Re_d = \frac{1}{|E_d|}\Big(\frac{1}{I}\sum_{i=1}^{I}(mse_d^{*(i)} - E_d)^2\Big)^{1/2},$$

$$Rb = \frac{1}{D}\sum_{d=1}^{D}|Rb_d|, \quad Re = \frac{1}{D}\sum_{d=1}^{D}Re_d.$$

Table 6 summarizes the results of Simulation 3. Figure 1 presents the boxplots of the relative biases Rb_d (left) and the relative root-MSEs Re_d (right) of MSE estimators for average incomes. Figure 2 presents the same boxplots for poverty proportions.

Table 6: *Relative biases and root-MSEs (in %) of MSE estimators for average incomes (top) and poverty proportions (bottom).*

		$B=25$	$B=50$	$B=100$	$B=200$	$B=300$	$B=400$
$\overline{Y}_d$	*Rb*	5.91	5.59	6.46	6.91	6.87	6.47
	Re	32.74	24.65	19.51	16.76	15.54	14.84
p_d	*Rb*	1.77	1.51	1.62	2.06	1.57	1.77
	Re	31.10	23.20	18.09	14.83	14.23	13.29

From the figures we observe that the parametric bootstrap method slightly underestimate the MSEs of the marginal predictors and that root mean squared error of the estimates is decreasing with increasing B. On the basis of the results we recommend to use at least $B = 200$ bootstrap iterations for estimating the MSEs of the marginal predictors.

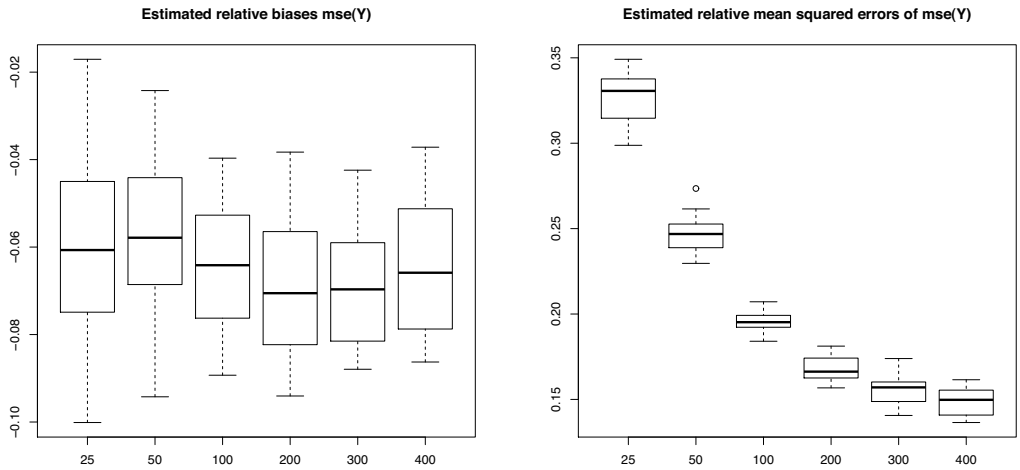

Figure 1: *Boxplots of relative biases Rb_d (left) and root-MSEs Re_d (right) of MSE estimators of average incomes.*

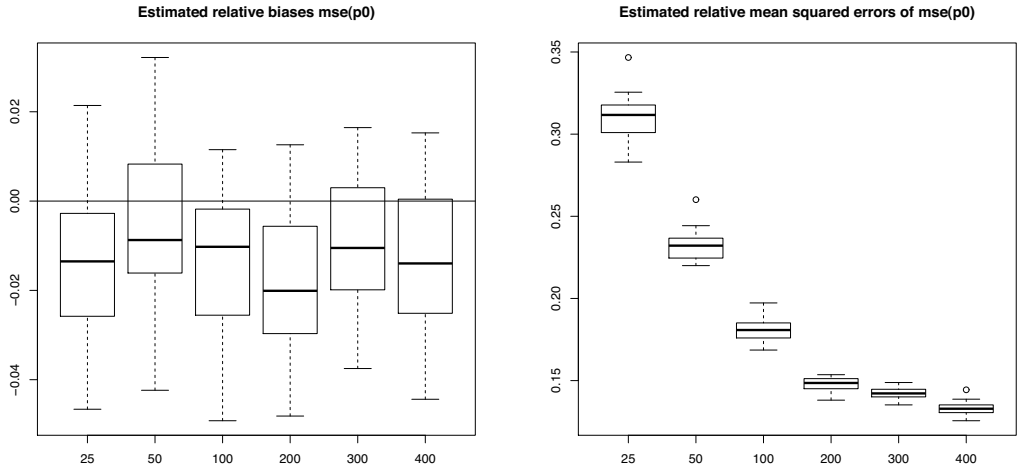

Figure 2: *Boxplots of relative biases Rb_d (left) and root-MSEs Re_d (right) of MSE estimators of poverty proportions.*

8 Application to SLCS data

This section presents an application to unit-level data from the 2013 SLCS of the region of Valencia. The SLCS is carried out in all the Spanish territory and the planned domains are the regions. The counties, within the region of Valencia, have rather small sample sizes and they are considered as small areas by the Spanish Statistical Office. We estimate average incomes and poverty proportions. The SLCS contains data from $D = 26$ Valencian counties and these counties are the domains of interest. The target variable y_{dj} is the average annual net income (in 10^4 euros) of individual j from domain d. The selected auxiliary variables are the labour status categories (employed, unemployed, inactive and below 15 years old). In addition to the SLCS data, we take auxiliary aggregated data from the SLFS, which contains survey data about the labour market. As the regional sample sizes of the SLFS are much greater than the corresponding ones of the SLCS, the sizes of counties crossed by the labour status categories are estimated from the SLFS and considered as known quantities.

We start the data analysis by doing a preliminary step. We fit gamma Model 1 to the data $(y_{dj}, x_{dj1}, x_{dj2})$, $d = 1,\ldots,D$, $j = 1,\ldots,n_d$, where x_{dj1} and x_{dj2} are the dichotomic variables indicating if an individual is employed and unemployed (yes = 1, no = 0) respectively. The $K = 3$ covariate classes are $z_1 = (1,0)$, $z_2 = (0,1)$ and $z_3 = (0,0)$ for employed, unemployed and rest ($\leq$ 15 or inactive) respectively. For $d = 1,\ldots,D$, $j = 1,\ldots,N_d$, we consider the population model (Model 1)

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu, \nu/\mu_{dj}\big), \quad g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \beta_0 + \beta_1 x_{dj1} + \beta_2 x_{dj2} + \phi v_d.$$

By using the glmer function of the R statistical package lme4, we fit Model 1 by applying the ML-Laplace approximation algorithm. Table 7 (left) gives the estimates of the regression parameters and their estimated standard deviations and p-values. Table 7 (right) presents the estimates of parameters ϕ and ν and the corresponding quantile-based 95% confidence intervals calculated from 1000 parametric bootstrap samples.

Table 7: *Regression (left) and var/shape (right) parameter estimates under Model 1.*

	estimate	standard error	p-value
$\tilde{\beta}_0$	0.771	0.0310	$<$ 2E-16
$\tilde{\beta}_1$	-0.142	0.0163	$<$ 2E-16
$\tilde{\beta}_2$	0.142	0.0287	7.42E-07

	estimate	95% conf. interval
$\tilde{\phi}$	0.093	(0.0496, 0.0991)
$\tilde{\nu}$	2.858	(2.270, 2.952)

Let us note that initially we have calculated the EBP's and marginal predictors of average incomes and poverty proportions under Model 1 but the results were unsatisfactory. The predicted values had very small variability between domains and for large sample sizes n_d they did not correspond to direct estimators. So the assumption $\nu_{dj} = \nu$ for all d and j is too rigid in this case and a more general model is needed.

We take the plug-in predictors $\tilde{\mu}_{dj} = (\tilde{\beta}_0 + \tilde{\beta}_1 x_{dj1} + \tilde{\beta}_2 x_{dj2} + \tilde{\phi}\tilde{v}_d)^{-1}$ calculated under Model 1 as a preliminary step and we use them as inputs of the algorithmic procedure for fitting the more complex unit-level gamma Model 2. For $d = 1,\dots,D$, $j = 1,\dots,N_d$, we consider the population model (Model 2)

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj}, \nu_{dj}/\mu_{dj}\big),\ \nu_{dj} = a_{dj}\varphi,\ g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \beta_0 + \beta_1 x_{dj1} + \beta_2 x_{dj2} + \phi v_d.$$

For fitting Model 2 to the data, we first need the constants a_{dj}. Since they are not known in our case, we estimate them by the following algorithmic procedure.

1. For a grid of values of t in the interval (0.25, 3) and step equal to 0.01, fit the Model 2 to the data, assuming that $a_{dj} = \tilde{\mu}_{dj}^t$ is known. If $t = 2$ and a_{dj} is equal to μ_{dj}^t, then $var[y_{dj}|v_d] = 1/\varphi$, which corresponds to the homoscedastic case. Calculate the estimator $\hat{\varphi}(t)$ for each considered t.

2. For each considered t, calculate the plug-in predictors $\hat{\mu}_{dj}^{(t)}$, the raw residuals $\hat{e}_{dj}^{(t)} = y_{dj} - \hat{\mu}_{dj}^{(t)}$ and the sum of the squared residuals $r^2(t)$. Select t_* minimizing $r^2(t)$.

3. Do the inferences with Model 2 and $a_{dj} = \tilde{\mu}_{dj}^{t_*}$ known, i.e. $\nu_{dj} = \hat{\varphi}(t_*)\tilde{\mu}_{dj}^{t_*}$.

For the considered data set, the selected optimal choice of t is $t_* = 0.60$. Figure 3 presents a plot of the function $r^2(t)$ and a boxplot of the optimal shape constants $a_{dj} = \tilde{\mu}_{dj}^{t_*}$.

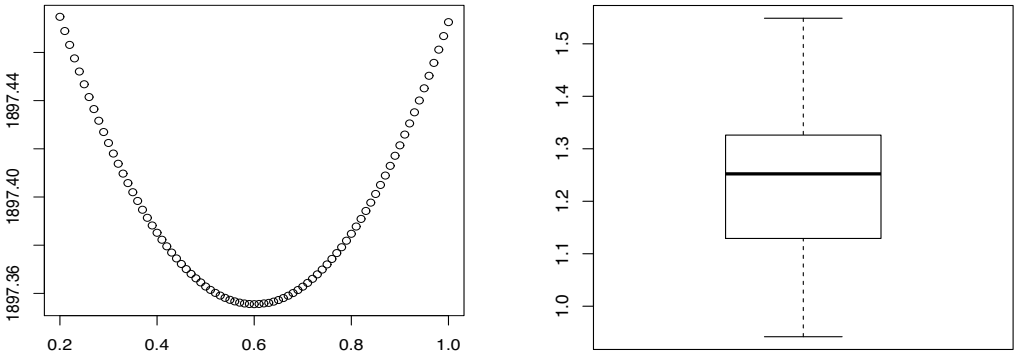

Figure 3: *Function $r^2(t)$ (left) and boxplot of a_{dj} (right).*

Table 8 gives the parameter estimates of Model 2. The signs of the regression parameter and the form of the link function indicate that employment (unemployment) has a positive (negative) effect on income.

Table 8: *Parameter estimates under Model 2.*

	estimate	standard error	p-value
$\hat{\beta}_0$	0.775	0.0132	$<$ 2E-16
$\hat{\beta}_1$	-0.141	0.0157	$<$ 2E-16
$\hat{\beta}_2$	0.140	0.0300	3.09E-06
$\hat{\phi}$	0.1113	0.0112	$<$ 2E-16
$\hat{\varphi}$	2.4646	0.0675	$<$ 2E-16

For the sake of comparisons, we also fit the unit-level log-linear normal mixed model (Model 3)

$$z_{dj} = b_0 + b_1 x_{dj1} + b_2 x_{dj2} + u_d + e_{dj}, \quad d = 1,\ldots,D,\ j = 1,\ldots,n_d,$$

where $z_{dj} = \log(y_{dj}+1)$, $u_d \sim N(0,\sigma_u^2)$, $e_{dj} \sim N(0,\sigma_e^2)$ and the random effects $u_d \sim N(0,\sigma_u^2)$ and the random errors $e_{dj} \sim N(0,\sigma_e^2)$ are mutually independent. By using the lmer function of the R statistical package lme4, we fit Model 3 by applying the REML method. The estimates of the model standard deviations are $\sigma_u = 0.0886$ and $\sigma_e = 0.3036$. Table 9 presents the estimates of the regression parameters of Model 3.

Table 9: *Parameter estimates under Model 3.*

	estimate	standard error	p-value
$\hat{b}_0$	0.803	0.0201	$<$ 2E-16
$\hat{b}_1$	0.137	0.0135	$<$ 2E-16
$\hat{b}_2$	-0.112	0.0180	5.41E-10

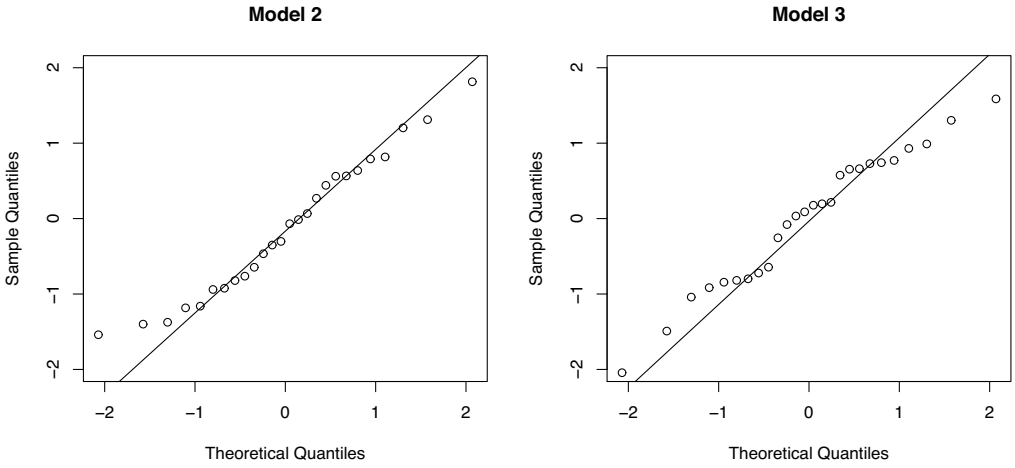

Figure 4: *Q-Q plot of random effects for models 2 (left) and 3 (right).*

In order to check the assumption that the random effects have the standard normal distribution, we take the mode predictors $\hat{v}_d$ and $\hat{u}_d$ of the random effects under Model 2 and Model 3 respectively and we plot the normal Q-Q plots in Figure 4. We do not observe significant deviations from normality. Moreover, the Kolmogorov-Smirnov test does not reject the hypothesis $H_0 : F_{\hat{v}_{1,d}} = F_{N(0,1)}$ with p-values equal to 0.763 (Model 2) and 0.925 (Model 3).

Remark 8.1 *In Figure 4 (left) there are two domains that are far from the straight line indicating normality in the bottom-left corner. To illustrate robustness of the method we have investigated what happens if we drop out all the observations of these two domains. We have fitted Model 2 without the mentioned observations and the results are presented in Table 11 of Appendix B. Since the parameter estimates are very similar to those given in Table 8, we can say that the methodology is robust with respect to small deviations from the hypothesis of normality of the random effects.*

Figure 5 presents graphs of raw residuals for Model 2 (left) and Model 3 (right). There are not significant visual differences between both models.

The sum of squares of raw residuals for models 2 and 3 are

$$r_2^2 = \sum_{d=1}^{D}\sum_{j=1}^{n_d}(y_{dj} - \hat{\mu}_{dj})^2 = 1897.35, \quad r_3^2 = \sum_{d=1}^{D}\sum_{j=1}^{n_d}(y_{dj} - (\exp(\hat{z}_{dj}) - 1))^2 = 1938.30.$$

As we observe that Model 2 has a slightly better fit to data, we do the estimation of the small area parameters (average income and poverty proportion) under Model 2. This application illustrates that it may have sense to consider more general GLMM instead of using normal mixed model for some transformation of data.

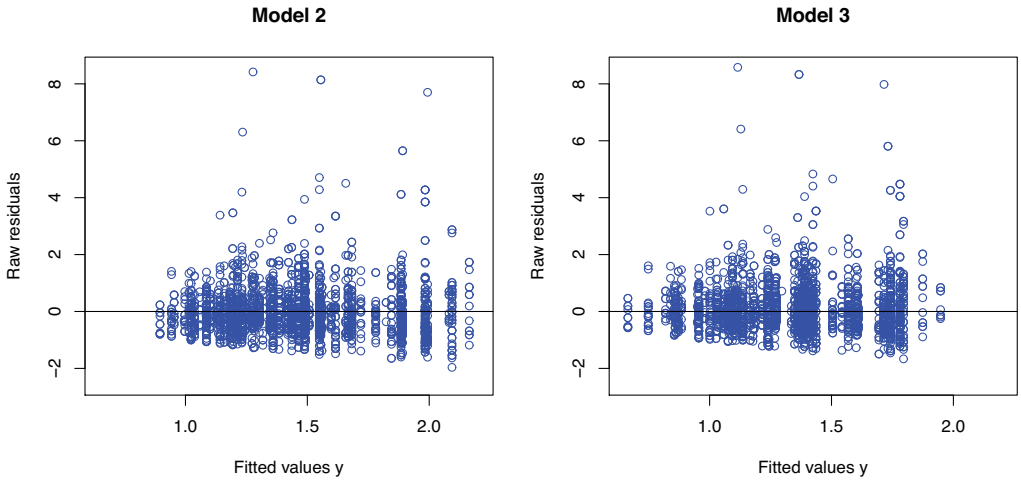

Figure 5: *Dispersion graphs of raw residuals for Model 2 (left) and Model 3 (right).*

Remark 8.2 *In order to study the importance of the random area effects in the model, we have fitted the following Gamma model without area effects (referred to as Model 0)*

$$y_{dj}|v_d \sim Gamma(\nu, \nu/\mu_{dj}), \quad d = 1,\dots,D,\ j = 1,\dots,n_d,$$

$$\eta_{dj} = g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \boldsymbol{x}_{dj}^{\top}\boldsymbol{\beta}, \tag{18}$$

and calculated the corresponding marginal predictors based on this model. The parameter estimates can be obtained by the R function glm and are presented in Table 12 in Appendix B. Unfortunately this model has a bad fit to data (the sum of squares of raw residuals $r_0^2 = 2010.56$ *in comparison with* $r_2^2 = 1897.35$ *obtained for model 2) and it does not explain the between domain variability which is not described by the auxiliary variables. Since the sizes of population classes are quite homogeneous between domains, it results in a quite over-smoothed behaviour of the predictors as can be seen from Table 13 of Appendix B. This table presents the marginal predictors under Model 2 and Model 0 and the corresponding bootstrap estimates of the MSE. The results for proportion predictions are presented also in Figure 8. From this figure one can see the smoothing effect of Model 0 and also that the estimated MSEs are for this model higher than the estimated MSEs of the direct estimators for sample sizes higher than 60.*

In order to get the marginal predictor of proportions, we need the a_{dj} values for the whole population or at least the values a_{dk} for the covariate classes $\boldsymbol{z}_k$, $k = 1,2,3$. Figures 6 and 7 were obtained by the choice

$$a_{dk} = \tilde{\mu}_{dk}^{t_*},$$

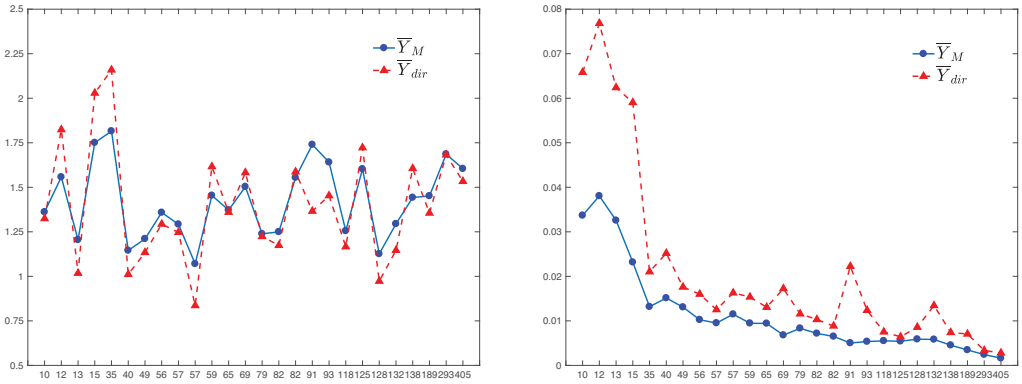

Figure 6: *Predictions of average incomes in* 10^4 *euros (left) and estimated MSEs (right).*

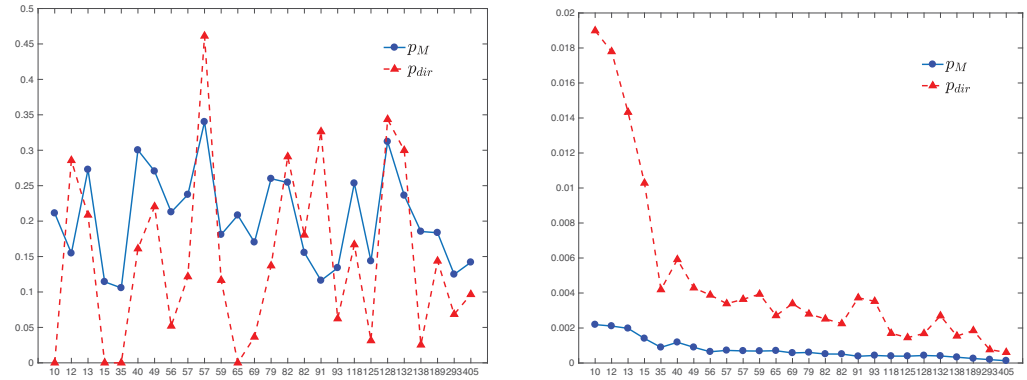

Figure 7: *Predictions of poverty proportions (left) and estimated MSEs (right).*

where $\tilde{\mu}_{dk}$ is the predictor of μ_{dk} derived under Model 1 ($\nu_{dj} = \nu$) and $t_* = 0.60$ is the optimal choice of t. This means that for every unit j of domain d and covariate class $\boldsymbol{z}_k \in \{\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3\}$, we take $a_{dj} = a_{dk}$ if $\boldsymbol{x}_{dj} = \boldsymbol{z}_k$, $d = 1, \ldots, 26$, $j = 1, \ldots, n_d$, $k = 1, 2, 3$.

Table 10 presents county codes (c), sample sizes (n_d), population sizes (N_d), marginal predictions of average incomes in 10^4 euros ($\overline{Y}_M$), marginal predictions of poverty proportions (p_M), direct estimates of average incomes in 10^4 euros ($\overline{Y}_{dir}$) and direct estimates of poverty proportions (p_{dir}). It also gives the corresponding MSE estimates (mse) based on 500 bootstrap samples generated from the fitted Model 2. As auxiliary population data, we took the SLFS data file of the region of Valencia in 2013. The poverty line for the region of Valencia was 6999.6 euros in 2013.

Figure 6 plots the marginal predictions ($\overline{Y}_M$) and the direct estimates ($\overline{Y}_{dir}$) of average incomes in 10^4 euros (left) and their corresponding model-based MSE bootstrap estimates (right). Figure 7 plots the marginal predictions (p_M) and the direct estimates (p_{dir}) of poverty proportions (left) and their corresponding model-based MSE bootstrap estimates (right). In all cases, the counties are sorted by sample size. We observe that the model-based marginal predictions have a more smooth behaviour across counties than

the direct estimates. Further, the marginal predictions and the direct estimates tend to be close when sample size increases. As expected, the marginal predictors have had lower estimated MSEs than the direct estimators.

Table 10: *Predictions and estimated MSEs for average incomes and poverty proportions.*

c	n_d	N_d	$\overline{Y}_M$	mse	p_M	mse	$\overline{Y}_{dir}$	mse	p_{dir}	mse
27	82	124270	1.2496	0.00718	0.2545	0.00051	1.1738	0.01031	0.2909	0.00252
28	57	70944	1.2924	0.00951	0.2374	0.00072	1.2471	0.01253	0.1217	0.00339
29	69	225440	1.5033	0.00680	0.1702	0.00058	1.5825	0.01725	0.0364	0.00339
30	132	227463	1.2945	0.00583	0.2362	0.00040	1.1461	0.01344	0.2998	0.00270
31	56	166774	1.3584	0.01025	0.2127	0.00064	1.2932	0.01600	0.0520	0.00388
32	293	459626	1.6860	0.00243	0.1248	0.00019	1.6821	0.00335	0.0685	0.00076
33	128	268924	1.1260	0.00589	0.3121	0.00043	0.9728	0.00858	0.3436	0.00169
34	59	292243	1.4540	0.00947	0.1809	0.00068	1.6162	0.01536	0.1166	0.00393
3	57	87560	1.0701	0.01149	0.3401	0.00069	0.8361	0.01629	0.4612	0.00364
5	91	246942	1.7397	0.00503	0.1162	0.00039	1.3659	0.02225	0.3263	0.00373
6	82	179798	1.5543	0.00651	0.1555	0.00051	1.5869	0.00885	0.1805	0.00225
7	10	26007	1.3613	0.03368	0.2112	0.00219	1.3245	0.06582	0.0000	0.01899
11	118	189865	1.2552	0.00551	0.2534	0.00040	1.1662	0.00752	0.1668	0.00170
12	15	89136	1.7504	0.02317	0.1144	0.00140	2.0290	0.05904	0.0000	0.01028
13	138	187515	1.4426	0.00454	0.1853	0.00033	1.6057	0.00738	0.0253	0.00154
14	189	370540	1.4513	0.00347	0.1836	0.00026	1.3552	0.00705	0.1438	0.00185
15	405	771129	1.6043	0.00163	0.1419	0.00013	1.5332	0.00279	0.0966	0.00061
16	93	131337	1.6408	0.00536	0.1340	0.00043	1.4531	0.01237	0.0623	0.00353
17	12	33122	1.5576	0.03806	0.1547	0.00211	1.8238	0.07683	0.2857	0.01779
18	35	54545	1.8157	0.01318	0.1057	0.00090	2.1590	0.02103	0.0000	0.00419
20	125	256553	1.6029	0.00543	0.1437	0.00039	1.7224	0.00642	0.0314	0.00145
21	49	52958	1.2107	0.01308	0.2704	0.00090	1.1340	0.01760	0.2205	0.00430
22	13	33126	1.2050	0.03257	0.2727	0.00198	1.0174	0.06239	0.2086	0.01433
23	40	70642	1.1452	0.01513	0.3002	0.00118	1.0109	0.02518	0.1611	0.00591
24	65	80434	1.3719	0.00942	0.2082	0.00070	1.3600	0.01306	0.0000	0.00271
25	79	180619	1.2386	0.00833	0.2598	0.00060	1.2238	0.01157	0.1371	0.00280

9 Concluding remarks

This paper introduces predictors of additive parameters under unit-level GLMMs. The introduced models are applicable also to continuous positive target random variables that have asymmetric distributions, like income or expenditure. In some practical cases, a GLMM can be a good alternative to the log-normal nested error model considered by Molina and Rao (2010). In the application to real data, we give a three-step procedure to determine the shape constants a_{dj} of gamma Model 2. This model has a high flexibility for fitting real data because the a_{dj}'s depend on d and j and therefore they vary within and between domains.

Among the three considered predictors, the simulations show that the empirical best and the marginal predictors have a similarly good behaviour. As the computation of the marginal predictor is less time demanding, we recommend it. Overall when reporting MSEs estimated by parametric bootstrap.

The new small area estimation methodology is applied to the SLCS data from Valencia, a region in east of Spain, in the period January-December 2013. The selected gamma Model 2 has had a slightly better fit to the data than the corresponding log-normal nested error regression model. Therefore the average incomes and poverty proportions per county are finally estimated by using the marginal predictors with its MSEs calculated by parametric bootstrap under gamma Model 2.

The simulations and the application to real data have been carried out with the programming language R. The codes are available upon request to the authors.

Acknowledgements

This work was supported by the Spanish grant ePGC2018-096840-B-I00 from Ministerio de Economía y Competitividad, the Czech grant SGS18/188/OHK4/3T/14 provided by Ministry of Education, Youth and Sports (MŠMT ČR) and from European Regional Development Fund-Project "Center of Advanced Applied Sciences"
(No. CZ.02.1.01/0.0/0.0/16_019/0000778).

References

Benavent, R. and Morales, D. (2016). Multivariate Fay-Herriot models for small area estimation. *Computational Statistics and Data Analysis*, 94, 372-390.

Boubeta, M., Lombardía, M.J. and Morales, D. (2016). Empirical best prediction under area-level Poisson mixed models. *TEST*, 25, 548-569.

Boubeta, M., Lombardía, M.J. and Morales, D. (2017). Poisson mixed models for studying the poverty in small areas. *Computational Statistics and Data Analysis*, 107, 32-47.

Chambers, R., Salvati, N. and Tzavidis, N. (2012). *M-Quantile Regression for Binary Data with Application to Small Area Estimation.* Centre for Statistical and Survey Methodology, University of Wollongong. (Working Paper 12-12, 2012, 24). Available at: http://ro.uow.edu.au/cssmwp/101.

Chambers, R., Salvati, N. and Tzavidis, N. (2016). Semiparametric small area estimation for binary outcomes with application to unemployment estimation for local authorities in the UK. *Journal of the Royal Statistical Society, Series A*, 179, 2, 453-479.

Dreassi, E., Petrucci, A. and Rocco, E. (2014). Small area estimation for semicontinuous skewed spatial data: An application to the grape wine production in Tuscany. *Biometrical Journal*, 56, 1, 141-156

Elbers, C., Lanjouw, J.O. and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71, 355-364.

Esteban, M.D., Morales, D., Pérez, A. and Santamaría, L. (2012a). Two Area-Level Time Models for Estimating Small Area Poverty Indicators. *Journal of the Indian Society of Agricultural Statistics*, 66, 1, 75-89.

Esteban, M.D., Morales, D., Pérez, A. and Santamaría, L. (2012b). Small area estimation of poverty proportions under area-level time models. *Computational Statistics and Data Analysis*, 56, 2840-2855.

Fabrizi, E., Ferrante, M.R. and Trivisano, C. (2017). Bayesian small area estimation for skewed business survey variables. *Jornal of the Royal Statistical Society, Series C*. doi.org/10.1111/rssc.12254

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D. and Santamaría, L. (2007). Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model. *Computational Statistics and Data Analysis*, 51, 2720-2733.

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D. and Santamaría, L. (2008a). Bootstrap mean squared error of small-area EBLUP. *Journal of Statistical Computation and Simulation*, 78, 443-462.

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D. and Santamaría, L. (2008b). Analytic and bootstrap approximations of prediction errors under a multivariate Fay-Herriot model. *Computational Statistics and Data Analysis*, 52, 5242-5252.

Guadarrama, M., Molina, I. and Rao, J.N.K. (2014). A comparison of small area estimation methods for poverty mapping. *Statistics in Transition new series and Survey Methodology*, 17, 41-66.

Hall, P. and Maiti, T. (2006a). Nonparametric estimation of mean-squared prediction error in nested-error regression models. *Annals of Statistics*, 34, 4, 1733-1750.

Hall, P. and Maiti, T. (2006b). On parametric bootstrap methods for small-area prediction. *Journal of the Royal Statistical Society, B*, 68, 221-238.

Hobza, T. and Morales, D. (2013). Small area estimation under random regression coefficient models. *Journal of Statistical Computation and Simulation*, 83, 11, 2160-2177.

Hobza, T. and Morales, D. (2016). Empirical Best Prediction Under Unit-Level Logit Mixed Models. *Journal of official statistics*, 32, 3, 661-692.

Hobza, T., Morales, D. and Santamaría, L. (2018). Small area estimation of poverty proportions under unit-level temporal binomial-logit mixed models. *TEST*, 27, N. 2., 270-294.

Karlberg, F. (2014). Small area estimation for skewed data in the presence of zeroes. *Statistics in Transition*, 16, 4, 541-562.

Marhuenda, Y., Molina, I. and Morales, D. (2013). Small area estimation with spatio-temporal Fay-Herriot models. *Computational Statistics and Data Analysis*, 58, 308-325.

Marhuenda, Y., Molina, I., Morales, D. and Rao, J.N.K. (2017). Poverty mapping in small areas under a two-fold nested error regression model. *Journal of the Royal Statistical Society, series A*, 180, 4, 1111-1136.

Molina, I., Nandram, B. and Rao, J.N.K. (2015). Small area estimation of general parameters with application to poverty indicators: a hierarchical Bayes approach. *The Annals of Applied Statistics*, 8, 852-885.

Molina, I. and Rao, J.N.K. (2010). Small area estimation of poverty indicators. *The Canadian Journal of Statistics*, 38, 369-385.

Morales, D., Pagliarella, M.C. and Salvatore, R. (2015). Small area estimation of poverty indicators under partitioned area-level time models. *SORT-Statistics and Operations Research Transactions*, 39, 1, 19-34.

Moura, F.A.S., Silva, D. B. N. and Neves, A.F. (2017). Small area models for skewed Brazilian business survey data. *Jornal of the Royal Statistical Society, Series A*, 180, 4, 1039-1055.

Pratesi, M. (2016). *Analysis of Poverty Data by Small Area Estimation*. John Wiley.

R Core Team (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/

Rao J. N. K. and Molina, I. (2015). *Small Area Estimation*, 2nd Edition, New York: Wiley.

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

Tzavidis, N., Salvati, N., Pratesi, M. and Chambers, R. (2008). M-quantile models with application to poverty mapping. *Statistical Methods and Applications*, 17, 3, 393-411.

A Appendix

This appendix gives the partial derivatives of the function ℓ_{0d}, defined in (6), for gamma Model 2. The first derivatives of μ_{0dj} and ξ_{0d} are

$$\frac{\partial \mu_{0dj}}{\partial \beta_r} = -x_{djr}\mu_{0dj}^2, \quad \frac{\partial \mu_{0dj}}{\partial \phi} = -v_{0d}\mu_{0dj}^2, \quad \frac{\partial \mu_{0dj}}{\partial \varphi} = 0,$$
$$\eta_{0dr} = \frac{\partial \xi_{0d}}{\partial \beta_r} = -2\phi^2 \sum_{j=1}^{n_d} a_{dj}\varphi x_{djr}\mu_{0dj}^3, \quad \eta_{0d\varphi} = \frac{\partial \xi_{0d}}{\partial \varphi} = \phi^2 \sum_{j=1}^{n_d} a_{dj}\mu_{0dj}^2,$$
$$\eta_{0d} = \frac{\partial \xi_{0d}}{\partial \phi} = \sum_{j=1}^{n_d} \{2\phi a_{dj}\varphi\mu_{0dj}^2 - 2\phi^2 a_{dj}\varphi v_{0d}\mu_{0dj}^3\}.$$

The first derivatives of ℓ_{0d} with respect to β_r, ϕ and φ are

$$\frac{\partial \ell_{0d}}{\partial \beta_r} = -\frac{1}{2}\frac{\eta_{0dr}}{\xi_{0d}} + \sum_{j=1}^{n_d} \{a_{dj}\varphi x_{djr}\mu_{0dj} - a_{dj}\varphi x_{djr}y_{dj}\},$$
$$\frac{\partial \ell_{0d}}{\partial \phi} = -\frac{1}{2}\frac{\eta_{0d}}{\xi_{0d}} + \sum_{j=1}^{n_d} \{a_{dj}\varphi v_{0d}\mu_{0dj} - a_{dj}\varphi v_{0d}y_{dj}\},$$
$$\frac{\partial \ell_{0d}}{\partial \varphi} = \sum_{j=1}^{n_d} \{a_{dj}\log a_{dj} + a_{dj}\log\varphi + a_{dj} + a_{dj}\log y_{dj} - a_{dj}\psi(a_{dj}\varphi)\} - \frac{1}{2}\frac{\eta_{0d\varphi}}{\xi_{0d}}$$
$$+ \sum_{j=1}^{n_d} \{a_{dj}\log(\boldsymbol{x}_{dj}^{\mathrm{T}}\boldsymbol{\beta} + \phi v_{0d}) - a_{dj}y_{dj}(\boldsymbol{x}_{dj}^{\mathrm{T}}\boldsymbol{\beta} + \phi v_{0d})\},$$

where $\psi(z) = \frac{d\log\Gamma(z)}{dz}$ is the digamma function. It holds that

$$\frac{\partial \eta_{0dr}}{\partial \beta_s} = \sum_{j=1}^{n_d} \{6\phi^2 a_{dj}\varphi x_{djr}x_{djs}\mu_{0dj}^4\}, \quad \frac{\partial \eta_{0dr}}{\partial \varphi} = -2\phi^2 \sum_{j=1}^{n_d} a_{dj}x_{djr}\mu_{0dj}^3,$$
$$\frac{\partial \eta_{0dr}}{\partial \phi} = \sum_{j=1}^{n_d} \{-4\phi a_{dj}\varphi x_{djr}\mu_{0dj}^3 + 6\phi^2 a_{dj}\varphi x_{djr}v_{0d}\mu_{0dj}^4\},$$
$$\frac{\partial \eta_{0d}}{\partial \beta_r} = \frac{\partial \eta_{0dr}}{\partial \phi}, \quad \frac{\partial \eta_{0d}}{\partial \varphi} = 2\sum_{j=1}^{n_d} \{\phi a_{dj}\mu_{0dj}^2 - \phi^2 a_{dj}v_{0d}\mu_{0dj}^3\},$$
$$\frac{\partial \eta_{0d}}{\partial \phi} = \sum_{j=1}^{n_d} \{2a_{dj}\varphi\mu_{0dj}^2 - 8\phi a_{dj}\varphi v_{0d}\mu_{0dj}^3 + 6\phi^2 a_{dj}\varphi v_{0d}^2\mu_{0dj}^4\},$$
$$\frac{\partial \eta_{0d\varphi}}{\partial \beta_r} = \frac{\partial \eta_{0dr}}{\partial \varphi}, \quad \frac{\partial \eta_{0d\varphi}}{\partial \phi} = \frac{\partial \eta_{0d}}{\partial \varphi}, \quad \frac{\partial \eta_{0d\varphi}}{\partial \varphi} = 0.$$

The second partial derivatives of ℓ_{0d} are

$$\frac{\partial^2 \ell_{0d}}{\partial \beta_s \partial \beta_r} = -\frac{1}{2}\frac{\frac{\partial \eta_{0dr}}{\partial \beta_s}\xi_{0d} - \eta_{0dr}\eta_{0ds}}{\xi_{0d}^2} - \sum_{j=1}^{n_d} a_{dj}\varphi x_{djr}x_{djs}\mu_{0dj}^2,$$

$$\frac{\partial^2 \ell_{0d}}{\partial \phi \partial \beta_r} = -\frac{1}{2}\frac{\frac{\partial \eta_{0dr}}{\partial \phi}\xi_{0d} - \eta_{0dr}\eta_{0d}}{\xi_{0d}^2} - \sum_{j=1}^{n_d} a_{dj}\varphi v_{0d}x_{djr}\mu_{0dj}^2,$$

$$\frac{\partial^2 \ell_{0d}}{\partial \varphi \partial \beta_r} = -\frac{1}{2}\frac{\frac{\partial \eta_{0dr}}{\partial \varphi}\xi_{0d} - \eta_{0dr}\eta_{0d\varphi}}{\xi_{0d}^2} + \sum_{j=1}^{n_d}\left\{a_{dj}x_{djr}\mu_{0dj} - a_{dj}x_{djr}y_{dj}\right\},$$

$$\frac{\partial^2 \ell_{0d}}{\partial \phi^2} = -\frac{1}{2}\frac{\frac{\partial \eta_{0d}}{\partial \phi}\xi_{0d} - \eta_{0d}^2}{\xi_{0d}^2} - \sum_{j=1}^{n_d} a_{dj}\varphi v_{0d}^2\mu_{0dj}^2,$$

$$\frac{\partial^2 \ell_{0d}}{\partial \varphi \partial \phi} = -\frac{1}{2}\frac{\frac{\partial \eta_{0d}}{\partial \varphi}\xi_{0d} - \eta_{0d}\eta_{0d\varphi}}{\xi_{0d}^2} + \sum_{j=1}^{n_d}\left\{a_{dj}v_{0d}\mu_{0dj} - a_{dj}v_{0d}y_{dj}\right\},$$

$$\frac{\partial^2 \ell_{0d}}{\partial \varphi^2} = \sum_{j=1}^{n_d}\left\{a_{dj}\varphi^{-1} - a_{dj}^2\dot{\psi}(a_{dj}\varphi)\right\} - \frac{1}{2}\frac{\frac{\partial \eta_{0d\varphi}}{\partial \varphi}\xi_{0d} - \eta_{0d\varphi}^2}{\xi_{0d}^2},$$

where $\dot{\psi}(z) = \frac{d\psi(z)}{dz}$ is the trigamma function.

For $r,s = 1,\ldots,p$ the components of the score vector and the Hessian matrix are

$$U_{0r} = \sum_{d=1}^{D}\frac{\partial \ell_{0d}}{\partial \beta_r}, \quad U_{0p+1} = \sum_{d=1}^{D}\frac{\partial \ell_{0d}}{\partial \phi}, \quad U_{0p+2} = \sum_{d=1}^{D}\frac{\partial \ell_{0d}}{\partial \varphi},$$

$$H_{0rs} = H_{0sr} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \beta_s \partial \beta_r}, \quad H_{rp+1} = H_{p+1r} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \phi \partial \beta_r},$$

$$H_{rp+2} = H_{p+2r} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \varphi \partial \beta_r}, H_{0p+1p+1} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \phi^2},$$

$$H_{0p+1p+2} = H_{0p+2p+1} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \phi \partial \varphi}, \quad H_{0p+2p+2} = \sum_{d=1}^{D}\frac{\partial^2 \ell_{0d}}{\partial \varphi^2}.$$

In matrix form, we have $\boldsymbol{U}_0 = \boldsymbol{U}_0(\boldsymbol{\theta}) = \underset{1\le r\le p+2}{\text{col}}(U_{0rs})$ and $\boldsymbol{H}_0 = \boldsymbol{H}_0(\boldsymbol{\theta}) = (H_{0rs})_{r,s=1,\ldots,p+2}$, where $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi, \varphi)^\top$.

B Appendix

This appendix presents some additional results in the form of tables and figures.

Table 11: *Parameter estimates under Model 2 without two domains.*

	estimate	standard error	p-value
$\hat{\beta}_0$	0.790	0.0137	$<$ 2E-16
$\hat{\beta}_1$	-0.142	0.0165	$<$ 2E-16
$\hat{\beta}_2$	0.134	0.0297	6.68E-06
$\hat{\phi}$	0.1081	0.0111	$<$ 2E-16
$\hat{\varphi}$	2.8269	0.0792	$<$ 2E-16

Table 12: *Parameter estimates under Model 0 without random effects.*

	estimate	standard error	p-value
$\hat{\beta}_0$	0.731	0.0130	$<$ 2E-16
$\hat{\beta}_1$	-0.158	0.0179	$<$ 2E-16
$\hat{\beta}_2$	0.151	0.0315	1.77E-06
$\hat{\nu}$	2.532		

Table 13: *Predictions and estimated MSEs for average incomes and poverty proportions under Model 2 (left) and Model 0 (right).*

c	n_d	N_d	$\overline{Y}_M$	mse	p_M	mse	$\overline{Y}_M^0$	mse	p_M^0	mse
27	82	124270	1.2496	0.00718	0.2545	0.00051	1.4581	0.06089	0.2155	0.00362
28	57	70944	1.2924	0.00951	0.2374	0.00072	1.4429	0.04642	0.2191	0.00328
29	69	225440	1.5033	0.00680	0.1702	0.00058	1.4554	0.05166	0.2178	0.00342
30	132	227463	1.2945	0.00583	0.2362	0.00040	1.4916	0.05851	0.2069	0.00304
31	56	166774	1.3584	0.01025	0.2127	0.00064	1.4776	0.06149	0.2107	0.00334
32	293	459626	1.6860	0.00243	0.1248	0.00019	1.4871	0.04135	0.2085	0.00286
33	128	268924	1.1260	0.00589	0.3121	0.00043	1.4505	0.05769	0.2176	0.00398
34	59	292243	1.4540	0.00947	0.1809	0.00068	1.4485	0.04983	0.2160	0.00341
3	57	87560	1.0701	0.01149	0.3401	0.00069	1.4739	0.05191	0.2114	0.00373
5	91	246942	1.7397	0.00503	0.1162	0.00039	1.4657	0.04944	0.2140	0.00345
6	82	179798	1.5543	0.00651	0.1555	0.00051	1.4591	0.04928	0.2155	0.00343
7	10	26007	1.3613	0.03368	0.2112	0.00219	1.4591	0.06097	0.2143	0.00382
11	118	189865	1.2552	0.00551	0.2534	0.00040	1.4683	0.05872	0.2137	0.00333
12	15	89136	1.7504	0.02317	0.1144	0.00140	1.4808	0.06417	0.2089	0.00405
13	138	187515	1.4426	0.00454	0.1853	0.00033	1.4804	0.04735	0.2097	0.00298
14	189	370540	1.4513	0.00347	0.1836	0.00026	1.4777	0.04929	0.2112	0.00297
15	405	771129	1.6043	0.00163	0.1419	0.00013	1.4792	0.03782	0.2100	0.00270
16	93	131337	1.6408	0.00536	0.1340	0.00043	1.4422	0.04166	0.2194	0.00322
17	12	33122	1.5576	0.03806	0.1547	0.00211	1.4776	0.05186	0.2100	0.00351
18	35	54545	1.8157	0.01318	0.1057	0.00090	1.4850	0.04623	0.2101	0.00329
20	125	256553	1.6029	0.00543	0.1437	0.00039	1.4736	0.04957	0.2121	0.00350
21	49	52958	1.2107	0.01308	0.2704	0.00090	1.4572	0.05194	0.2147	0.00338
22	13	33126	1.2050	0.03257	0.2727	0.00198	1.4573	0.05235	0.2167	0.00337
23	40	70642	1.1452	0.01513	0.3002	0.00118	1.4558	0.04925	0.2155	0.00342
24	65	80434	1.3719	0.00942	0.2082	0.00070	1.4588	0.05811	0.2155	0.00364
25	79	180619	1.2386	0.00833	0.2598	0.00060	1.4628	0.05700	0.2156	0.00342

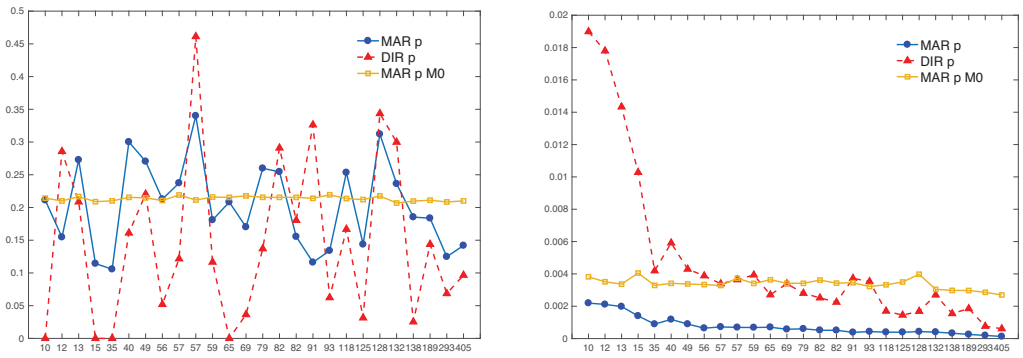

Figure 8: *Direct estimators and marginal predictors (under Model 2 and Model 0) of poverty proportions (left) and corresponding estimated MSEs (right).*

SORT 44 (1) January-June 2020, 39-66 DOI: 10.2436/20.8080.02.94

Finding archetypal patterns for binary questionnaires

Ismael Cabero[1] and Irene Epifanio[2]

Abstract

Archetypal analysis is an exploratory tool that explains a set of observations as mixtures of pure (extreme) patterns. If the patterns are actual observations of the sample, we refer to them as archetypoids. For the first time, we propose to use archetypoid analysis for binary observations. This tool can contribute to the understanding of a binary data set, as in the multivariate case. We illustrate the advantages of the proposed methodology in a simulation study and two applications, one exploring objects (rows) and the other exploring items (columns). One is related to determining student skill set profiles and the other to describing item response functions.

MSC: 62H99, 62P25, 97D60.

Keywords: Dichotomous item test, archetypal analysis, functional data analysis, item response theory, skill profile.

1 Introduction

Mining binary survey data is of utmost importance in social sciences. Many raw data from exams, opinion surveys, attitude questionnaires, etc. come in the form of a binary data matrix, i.e. examinees' responses are coded as 0/1 (1 if examinee i answers item h correctly, otherwise 0). The binary matrix can be viewed from two points of views. In the first, the interest lies in the rows, i.e. in the people, while in the second, the interest lies in the columns that contain the items or variables. In both cases, exploratory data analysis (EDA) aims to find information in data and generate ideas (Unwin, 2010). In order to be useful as a tool for EDA on data sets, the tool should be simple and easy to use, with few parameters, and reveal the salient features of the data in such a way that humans can visualize them (Friedman and Tukey, 1974).

For the first time, we propose the use of the exploratory tool Archetypoid Analysis (ADA) for this kind of data in order to understand, describe, visualize and extract information that is easily interpretable, even by non-experts. ADA is an unsupervised

[1] Departamento de Didáctica de las Matemáticas, Universidad Internacional de La Rioja, Spain. Email: ismael.cabero@unir.net

[2] Departament de Matemàtiques-IMAC-IF, Universitat Jaume I, Castelló 12071, Spain. Email: epifanio@uji.es

Received: March 2019
Accepted: February 2020

statistical learning technique (see Hastie, Tibshirani and Friedman, 2009, Chapter 14) for a complete review of unsupervised learning techniques). Its objective is to approximate sample data by a convex combination (a mixture) of k pure patterns, the archetypoids, which are extreme representative observations of the sample. Being part of the sample makes them interpretable, but also being extreme cases facilitates comprehension of the data. Humans understand the data better when the observations are shown through their extreme constituents (Davis and Love, 2010) or when features of one observation are shown as opposed to those of another (Thurau et al., 2012).

ADA was proposed by Vinué et al. (2015) for real continuous multivariate data as a derivative methodology of Archetype Analysis (AA). AA was formulated by Cutler and Breiman (1994), and like ADA, it seeks to approximate data through mixtures of archetypes, also for real continous multivariate data. However, archetypes are not actual cases, but rather a mixture of data points. Recently, Seth and Eugster (2016b) proposed a probabilistic framework of AA (PAA) to accommodate binary observations by working in the parameter space.

AA and ADA have been applied to many different fields, such as astrophysics (Chan, Mitchell and Cram, 2003), biology (D'Esposito, Palumbo and Ragozini, 2012), climate (Steinschneider and Lall, 2015; Su et al., 2017), developmental psychology (Ragozini, Palumbo and D'Esposito, 2017), e-learning (Theodosiou et al., 2013), finance (Moliner and Epifanio, 2019), genetics (Thøgersen et al., 2013), human development (Epifanio 2016; Epifanio, Ibáñez and Simó, 2020), industrial engineering (Epifanio et al., 2013; Epifanio, Ibáñez and Simó, 2018; Millán-Roures, Epifanio and Martínez, 2018; Alcacer et al., 2020), machine learning (Mørup and Hansen, 2012; Seth and Eugster, 2016a,b; Ragozini and D'Esposito, 2015; Cabero and Epifanio, 2019), market research (Li et al., 2003; Porzio, Ragozini and Vistocco, 2008; Midgley and Venaik, 2013), multidocument summarization (Canhasi and Kononenko, 2013, 2014), nanotechnology (Fernandez and Barnard, 2015), neuroscience (Tsanousa, Laskaris and Angelis, 2015; 2016) and sports (Eugster, 2012; Vinué and Epifanio, 2017, 2019).

Archetypal analysis techniques lie somewhere in between two well-known unsupervised statistical techniques: Principal Component Analysis (PCA) and Cluster Analysis (CLA). In data decomposition techniques, a data set is viewed as a linear combination of several factors to find the latent components. Different prototypical analysis tools arise depending on the constraints on the factors and how they are combined (Mørup and Hansen, 2012; Vinué, Epifanio and Alemany, 2015). The factors with the least restrictions are those produced by PCA, since they are linear combinations of variables. One of the advantages is that this helps explain the variability of the data; however, the interpretability of the factors is compromised. Instead, the greatest restrictions are found in cluster tools, such as k-means or k-medoids. Their factors are readily interpreted because they are centroids (means of groups of data) or medoids (concrete observations) in the case of k-means and k-medoids, respectively. The price that clustering tools pay for interpretability is their modeling flexibility due to the binary assignment of data to the clusters. Archetypal tools, on the other hand, enjoy higher modeling flexibility than

cluster tools but without losing the interpretability of their factors. A table summarizing the relationship between several unsupervised multivariate techniques is provided by Mørup and Hansen (2012) and Vinué et al. (2015).

AA and ADA were originally thought of for real-valued observations. The aim of this work is to extend archetypal tools to binary data. For AA, as the factors (archetypes) are a mixture of data, they would not necessarily be binary vectors, and as a consequence they would not be interpretable. In ADA though, the factors (archetypoids) are actual cases, so ADA can be applied to binary data without losing the interpretability of the factors. So, among the possible archetypal techniques (AA, PAA and ADA), we propose to use ADA for binary data.

To perform a sanity check and provide insight we analyze the solutions obtained by AA, PAA and ADA through a simulation study, where ADA shows its appropriateness versus AA or PAA for binary data sets. Furthermore, we present two real applications and compare ADA solutions with those of other established unsupervised techniques to illustrate the advantages of ADA in educational and behavioral sciences, when used as another useful tool for data mining in these fields (Slater et al., 2017). In the first application, we are interested in rows, while in the second application in columns.

The outline of the paper is as follows: In Section 2 we review AA and ADA for real-valued multivariate and functional data and PAA, besides other multivariate techniques used in the comparison. In Section 3 we introduce the analysis for binary multivariate data. In Section 4, a simulation study with binary data compares the different strategies for obtaining archetypal patterns. In Section 5, our proposal is applied to two real data sets and compared to the results of other well-known unsupervised statistical learning techniques. Section 6 contains conclusions and some ideas for future work.

The data sets and code in R (R Development Core Team, 2019) for reproducing the results for both artificial and real data are available at http://www3.uji.es/ epifanio/ RESEARCH/adaedu.rar.

2 Preliminary

2.1 AA and ADA in the real-valued multivariate case

Let $\mathbf{X}$ be an $n \times m$ real-valued matrix with n observations and m variables. Three matrices are established in AA: a) the k archetypes $\mathbf{z}_j$, which are the rows of a $k \times m$ matrix $\mathbf{Z}$; b) an $n \times k$ matrix $\alpha = (\alpha_{ij})$ with the mixture coefficients that approximate each observation $\mathbf{x}_i$ by a mixture of the archetypes ($\hat{\mathbf{x}}_i = \sum_{j=1}^{k} \alpha_{ij}\mathbf{z}_j$); and c) a $k \times n$ matrix $\beta = (\beta_{jl})$ with the mixture coefficients that characterize each archetype ($\mathbf{z}_j = \sum_{l=1}^{n} \beta_{jl}\mathbf{x}_l$). To figure out these matrices, we minimize the following residual sum of squares (RSS) with the respective constraints ($\|\cdot\|$ denotes the Euclidean norm for vectors):

$$RSS = \sum_{i=1}^{n} \|\mathbf{x}_i - \sum_{j=1}^{k} \alpha_{ij}\mathbf{z}_j\|^2 = \sum_{i=1}^{n} \|\mathbf{x}_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl}\mathbf{x}_l\|^2, \quad (1)$$

under the constraints

1) $\sum_{j=1}^{k} \alpha_{ij} = 1$ with $\alpha_{ij} \geq 0$ for $i = 1, \ldots, n$ and

2) $\sum_{l=1}^{n} \beta_{jl} = 1$ with $\beta_{jl} \geq 0$ for $j = 1, \ldots, k$.

As previously mentioned, archetypes do not necessarily match real observations. Indeed, this will only happen when one and only one β_{jl} is equal to one for each archetype, i.e. when each archetype is defined by only one observation. So, in ADA the previous constraint 2) is substituted by the following one, and as a consequence in ADA a mixed-integer problem is optimized instead of the AA continuous optimization problem:

2) $\sum_{l=1}^{n} \beta_{jl} = 1$ with $\beta_{jl} \in \{0, 1\}$ and $j = 1, \ldots, k$.

As regards the location of archetypes, they are on the boundary of the convex hull of the data if $k > 1$ (see Cutler and Breiman, 1994), although this does not necessarily happen for archetypoids (see Vinué et al., 2015). Nonetheless, the archetype is equal to the mean and to the medoid in case of the archetypoid (Kaufman and Rousseeuw, 1990), if $k = 1$.

We want to emphasize that archetypal analysis is an EDA technique based on a geometric formulation (no distribution of data is assumed). It is not an inferential statistical technique, i.e. it is not about fitting models, parameter estimation, or testing hypotheses. Nevertheless, a field to study in the future would be to view archetypal analysis as a feature extraction method (Hastie et al., 2009, Ch. 5), where the raw data are preprocessed and described by α, which can be used as inputs into any learning procedure for compositional data (Pawlowsky-Glahn, Egozcue and Tolosana-Delgado, 2015).

2.1.1 Computation of AA and ADA

The estimation of the matrices in the AA problem can be achieved by means of an alternating minimizing algorithm developed by Cutler and Breiman (1994), where the best α for given archetypes $\mathbf{Z}$ and the best archetypes $\mathbf{Z}$ for a given α are computed by turns. To solve the convex least squares problems, a penalized version of the non-negative least squares algorithm by Lawson and Hanson (1974) is used. Eugster and Leisch (2009) implemented that algorithm in the R package **archetypes**, although with some changes. Specifically, the data are standardized and the spectral norm in equation

1 is used instead of Frobenius norm for matrices. In our R implementation those changes were annulled, i.e. the data are not standardized by default and the objective function to minimize is defined by equation 1.

With respect to the estimation of the matrices in the ADA problem, it can be achieved using the algorithm developed by Vinué et al. (2015). It is composed of two steps: the BUILD step and the SWAP step. The objective of the BUILD step is to determine an initial set of archetypoids that will be upgraded during the following step. The objective of the SWAP step is to improve the primitive set by exchanging the selected instances for unselected observations and checking whether these replacements decrease the RSS. Vinué (2017) implemented that algorithm in the R package **Anthropometry** with three possible original sets in the BUILD step: $cand_{ns}$, $cand_{\alpha}$ and $cand_{\beta}$. These sets correspond to the nearest neighbor observations in Euclidean distance to the k archetypes, the cases with the maximum α value for each archetype j and the observations with the maximum β value for each archetype j, respectively. Then three possible solutions are obtained once these three sets go through the SWAP step, but only the solution with lowest RSS (often the same final set is returned from the three initializations) is chosen as the ADA solution.

One important point is the selection of k, since archetypes are not necessarily nested and neither are archetypoids. If the user has prior knowledge of the structure of the data, the value of k can be chosen based on that information. Otherwise, a simple but effective heuristic (Cutler and Breiman, 1994; Eugster and Leisch, 2009; Vinué et al., 2015; Seth and Eugster, 2016b) such as the elbow criterion can be used. With the elbow criterion, we plot the RSS for different k values and the value of k is selected as the point where the elbow is located.

2.1.2 Illustrative example

In Figure 1 a toy two-dimensional data set is used to illustrate what archetypoids mean and the differences compared with CLA and PCA, as well as to provide some intuition on what these pure and extreme patterns imply in behavioral sciences. Two numeric variables are considered from the data set personality-1.0 of the R package **neuropsychology** (Makowski, 2016), which contains personality traits data from an online questionnaire: Empathy.Agreeableness and Honesty.Humility. We apply k-means and ADA with $k = 3$, i.e. we find 3 clusters and archetypoids. We also apply PCA. Archetypoids are people with extreme values, which have clear profiles: archetypoid 1 is characterized by a very low Empathy.Agreeableness value together with a high Honesty.Humility value (1, 5.25), archetypoid 2 has the maximum values for both Empathy.Agreeableness and Honesty.Humility (7,7), while the third archetypoid has a very high Empathy.Agreeableness value together with the lowest Honesty.Humility value (6,0). Archetypoids are the purest people. The rest of the individuals are expressed as mixtures (collected in alpha coefficients) of these ideal people. For example, an individual with values of 6.25 and 0.75 for Empathy.Agreeableness and Honesty.Humility, respectively, is explained by 11% of

archetypoid 2 plus 89% of archetypoid 3.

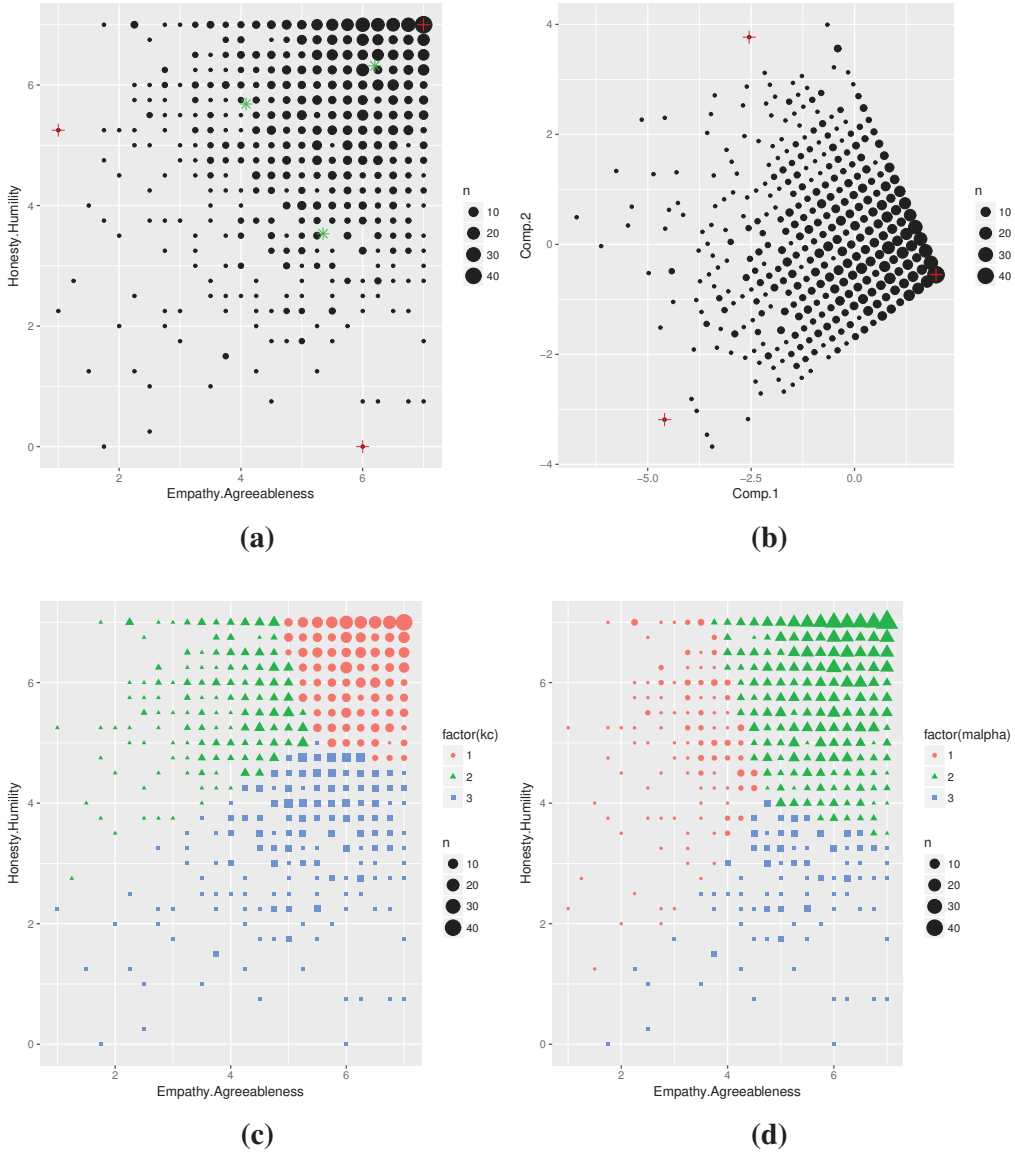

Figure 1: *(a) Plot of the toy example. The size of the points depends on their frequency. The red crosses represent the archetypoids, while the green stars represent the centroids of each cluster; (b) PC scores. Projected archetypoids are represented by red crosses; (c) k-means cluster assignments; (d) ADA assignments by the maximum alpha, i.e. assigned to the archetypoid that best explains the corresponding observation.*

This is compatible with the natural tendency of humans to represent a group of objects by its extreme elements (Davis and Love, 2010). Figure 1 d) shows the partition of the set generated by assigning the observations to the archetypoid that best explains

each individual. However, when we apply k-means to this kind of data set, without differentiated clusters, the centroids are in the middle of the data cloud. Centroid profiles are not as differentiated from each other as archetypoid profiles. This happens because centroids have to cover the set in such a way that the set is partitioned by minimizing the distance with respect to the assigned centroid (see Wu, Kamar and Horvitz, 2016) about the connection between set partitioning and clustering). On the one hand, this means that the set partition generated by k-means and ADA would be different (Figures 1 c) and d)). On the other hand, centroids are not the purest, and therefore their profiles are not as clear as those of archetypoids. For example, centroids 2 and 3 have values (4.1, 5.7) and (5.3, 3.5), which are not as intuitively interpretable as archetypoids. If we look again at the individual with values (6.25, 0.75) from the clustering point of view this individual is clearly assigned to cluster 3, with centroid (5.3, 3.5), but clustering does not say anything about the distance of this point with respect to the assigned centroid, or in which direction they are separated. In fact, (6.25, 0.75) is quite far from (5.3, 3.5). This happens because the objective of clustering is to assign the data to groups, not to explain the structure of the data more qualitatively. Finally, note that archetypoids do not coincide with the individuals with the most extreme PC scores (see Figure 1 b)).

In summary, depending on our objective, the appropriate analysis should be selected. The objective of PCA is to reduce data dimension. Although PCA returns the location of the observations in the new dimensions by PC scores, there is no guarantee that the principal components are interpretable. In other words, observations are expressed in a new base, but in general the PCA base is not easily interpretable. However, the objective of CLA is to segment the data, i.e. to make groups of data by finding modes in data. Although the modes can be easily interpretable, CLA does not return an expression about the location of each observation with respect to each mode. On the other hand, finding extreme profiles, which are easily interpretable, is not the objective of PCA or CLA, but that of AA or ADA. These techniques also return the location of the observations as a function of the extreme profiles, in fact as a mixture (a convex combination), which is more easily interpretable than a linear combination. This provides a complete overview of the data set, generally supported by visual methods, i.e. this allows data to tell us more beyond the formal modeling or hypothesis testing task.

2.2 Probabilistic archetype analysis

The idea underlying PAA is to work in a parameter space instead of the observation space, since the parameter space is often vectorial even if the sample space is not. The key is to assume that data points come from a certain distribution (from the Bernoulli distribution in the case of binary observations). Then the maximum likelihood estimates of the parameters of the distributions are seen as the parametric profiles that best describe each observation, and archetypal profiles are computed in the parameter space by maximizing the corresponding log-likelihood under the constraints for α and β. In sum-

mary, probabilistic archetypes lie in the parameter space, whereas classical archetypes lie in the observation space. Thus, archetypal profiles for binary data are the probability of a positive response. Details can be found in Seth and Eugster (2016b).

2.3 AA and ADA in the functional case

In Functional Data Analysis (FDA) each datum is a function. Therefore, the sample is a set of functions $\{x_1(t),\ldots,x_n(t)\}$ with $t \in [a,b]$, i.e. the values of the m variables in the standard multivariate context are replaced by function values with a continuous index t. We assume that these functions belong to a Hilbert space, satisfy reasonable smoothness conditions and are square-integrable functions on that interval. Simplistically, the sums are replaced by integrals in the definition of the inner product.

AA and ADA were extended to functional data by Epifanio (2016). In the functional context, functions from the data set are approximated by mixtures of archetypal functions. In functional archetype analysis (FAA), we seek k archetype functions that approximate the functional data sample by their mixtures. In other words, the objective of FAA is the same as AA, but now both archetypes ($z_j(t)$) and observations ($x_i(t)$) are functions. As a consequence, RSS is now calculated with a functional norm instead of a vector norm. We consider the L^2-norm, $\|f\|^2 =< f,f >= \int_a^b f(t)^2 dt$. The interpretation of matrices α and β is the same as in the classical multivariate case.

Analogously, FADA is also a generalization of ADA, where k functional archetypoids, which are functions of the sample, approximate the functions of the sample through the mixtures of these functional archetypoids. Again, vectors are replaced by functions and vector norms by functional norms, and the matrices are interpreted is the same way as before.

To obtain FAA and FADA in a computationally efficient way (Epifanio (2016)), functional data are represented by means of basis functions (see Ramsay and Silverman (2005) for a detailed explanation about smoothing functional data). Let B_h ($h = 1,\ldots,m$) be the basis functions and $\mathbf{b}_i$ the vector of coefficients of length m such that $x_i(t) \approx \sum_{h=1}^{m} b_i^h B_h(t)$. Then, RSS is formulated as (see Epifanio (2016) for details):

$$RSS = \sum_{i=1}^{n} \|x_i - \sum_{j=1}^{k} \alpha_{ij} z_j\|^2 = \sum_{i=1}^{n} \|x_i - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl} x_l\|^2 = \sum_{i=1}^{n} \mathbf{a}_i^{\mathsf{T}} \mathbf{W} \mathbf{a}_i, \tag{2}$$

where $\mathbf{a}_i^{\mathsf{T}} = \mathbf{b}_i^{\mathsf{T}} - \sum_{j=1}^{k} \alpha_{ij} \sum_{l=1}^{n} \beta_{jl} \mathbf{b}_l^{\mathsf{T}}$ and $\mathbf{W}$ is the order m symmetric matrix with the inner products of the pairs of basis functions $w_{m_1,m_2} = \int B_{m_1} B_{m_2}$. If the basis is orthonormal, for instance the Fourier basis, $\mathbf{W}$ is the order m identity matrix and FAA and FADA can be estimated using standard AA and ADA with the basis coefficients. If not, $\mathbf{W}$ has to be calculated previously one single time by numerical integration.

2.4 Other unsupervised learning techniques

The following well-known multivariate analysis tools for binary data are used in the comparison. We use homogeneity analysis (multiple correspondence analysis) using the R package **homals** (de Leeuw and Mair, 2009) (HOMALS). HOMALS can be considered as an equivalent to PCA in the case of categorical data. For CLA we use Partitioning Around Medoids (PAM) from the R package **cluster** (Maechler et al., 2018; Kaufman and Rousseeuw, 1990), since it returns representative objects or medoids among the observations of the data set. The pairwise dissimilarities between observations in the data set needed for PAM are computed with the *daisy* function from the R package **cluster** (Maechler et al., 2018), specifically using Gower's coefficient (Gower, 1971) for binary observations. Other popular clustering methods (Flynt and Dean, 2016) are also used in the comparison: latent class analysis (LCA) from the R package **poLCA** (Linzer and Lewis, 2011), which is a finite mixture model clustering for categorical data, and classical k-means clustering (Lloyd, 1982). It is used in the literature (Henry et al., 2015), despite not being recommended for binary data (IBM Support, 2016). For that reason, we also consider PAM, since it is a robustified version of k-means (Steinley, 2006) that can be used with distances other than Euclidean, and observations, rather than centroids, serve as the exemplars for each cluster.

3 Archetypal analysis for binary data

Let $\mathbf{X}$ be an $n \times m$ binary matrix with n observations and m variables. The idea behind archetypal analysis is that we can find a set of archetypal patterns, and that data can be expressed as a mixture of those archetypal patterns. In the case of binary data, on the one hand the archetypal patterns should also be binary data, as the population from which data come. For example, if pregnancy was one of the binary variables, it would not make sense to consider as an archetypal observation a woman who was pregnant 0.7. In other words, archetypal patterns should be binary in order to have a clear meaning and not lose their interpretability, which is the cornerstone of archetypal techniques, i.e. they should not be 'mythological', but rather something that might be observed. On the other hand, in order to describe data as mixtures, we should assume that observations exist in a vector space, i.e. that observations can be multiplied by scalars (in this case in the interval $[0, 1]$) and added together.

A solution that meets all these ideas is to apply ADA to $\mathbf{X}$, since the feasible archetypal patterns belong to the observed sample. In fact, ADA was originally created as a response to the problem in which pure non-fictitious patterns were sought (Vinué et al., 2015).

Instead, the archetypes returned by applying AA or PAA do not need to be binary, i.e. they do not need to belong to the feasible set of solutions. In fact, Seth and Eugster (2016b) binarized the archetypes obtained by AA or PAA in experiments. However,

using a continuous optimization problem to solve a problem whose feasible solutions are not continuous can fail badly (Fletcher, 2000, Ch. 13). Indeed, there is no guarantee that this approach will provide a good solution, even by examining all the feasible binary solutions in a certain neighborhood of the continuous solution.

Therefore, we propose to use ADA to handle binary observations.

4 Simulation study

We have carried out a simulation study to assess all the alternatives in a controlled scenario. The design of the experiment has been based on simulation studies that appear in Vinué et al. (2015) and Seth and Eugster (2016b). We generate $k = 6$ archetypes, ζ_i, with $m = 10$ binary variables by sampling them from a Bernoulli distribution with a probability of success $p = 0.7$, $A = [\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_6]$. Given the archetypes, we generate $n = 100$ observations as the binarized version of $x_i = \tilde{A}_i h_i + E_i$, where $\tilde{A}_i$ contains the archetypes after adding salt and pepper noise to them, h_i is a random vector sampled from a Dirichlet distribution with $\alpha = (0.8, 0.8, 0.8, 0.8, 0.8, 0.8)$, and E_i is a 10-dimensional random vector of Gaussian white noise with a mean of zero and standard deviation of 0.1. The binarized versions are obtained by replacing all values above 0.5 with 1 and others with 0. The noise density added to A is 0.05 (the default value used in MATLAB). With salt and pepper noise, a certain amount of the data is changed to either 0 or 1. To ensure that $\tilde{A}_i$'s are archetypes, we chose $\alpha = 0.8$, a value near to but less than one.

We compute PAA, AA and ADA. The archetypes returned by PAA and AA are binarized for comparison with the true ones, A. We calculated the Hamming distance (Manhattan distance between binary vectors), which is the same as the misclassification error used with binary images, between each archetypal solution and the true archetypes after permuting the columns of each archetypal solution to match the true archetypes in such a way that the least error with the city block distance is provided.

This was repeated 100 times. The first 10 times are displayed in Figure 2. The solutions returned by all the methods are quite similar to the true archetypes, i.e. the number of errors (a zero in the solution where the true value is 1, or vice versa) is very small. Nevertheless, there are differences between the methods, which are more evident in columns 5 and 6. For columns 5 and 6, the number of errors for PAA is 5 and 5, it is 4 and 2 for AA, but only 2 and 2 for ADA. Table 1 shows a the summary of the

Table 1: *Summary of misclassification errors of the archetype profiles for each method over 100 simulations.*

Method	PAA	AA	ADA
Mean (Std. dev.)	4.20 (1.86)	3.59 (1.99)	3.19 (1.88)

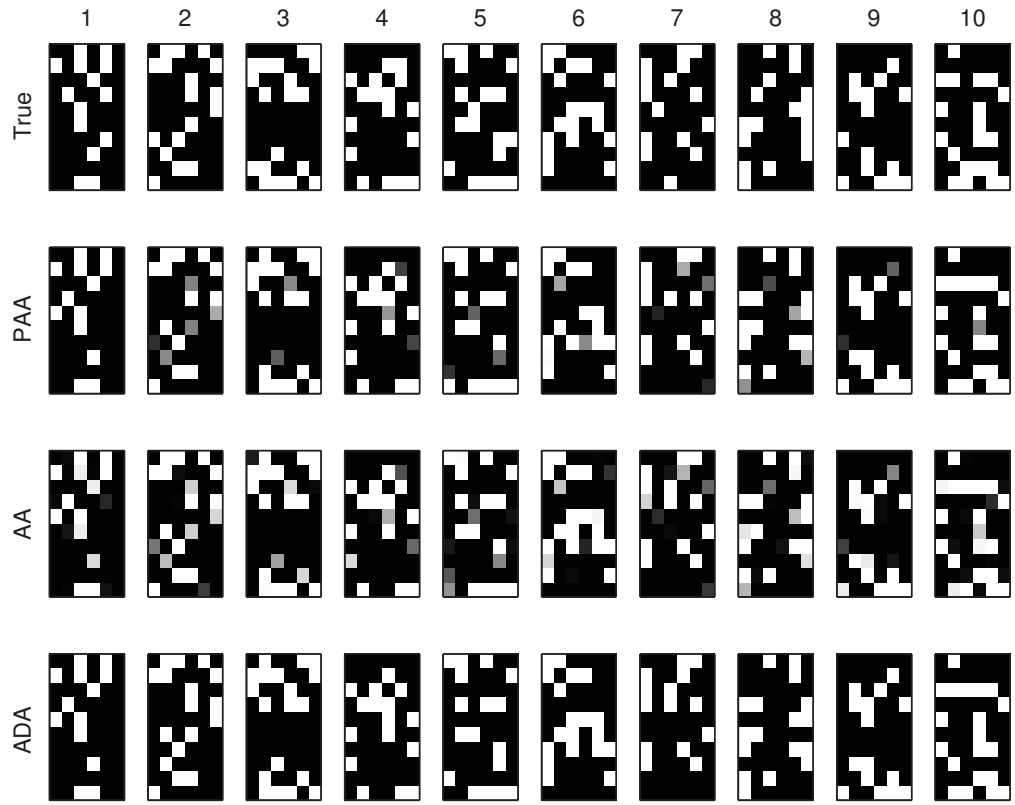

Figure 2: *Comparison between true archetypes and those returned by PAA, AA and ADA, respectively. The 10 columns represent the first 10 repetitions of the simulation. Black represents 0 and white 1.*

misclassifications. The archetypoids returned by ADA match the true archetypes better than those returned by AA or PAA, in this order, i.e. ADA provides the smallest mean misclassification error.

5 Applications

5.1 An initial mathematical skills test for first-year university students

5.1.1 Data

The first application corresponds to the first point of view of the binary matrix (analysis of the rows). We analyze the data set described by Orús and Gregori (2008), which was obtained through the application of a test on the initial mathematics skills of 690 first-year students of the College of Technology and Experimental Sciences at Jaume I University (Spain) at the beginning of the 2003-04 academic year. The test consisted

of 17 questions corresponding to 21 single items, the answers to which were coded as 0 (incorrect or unanswered) or 1 (correct). The items of the test were selected in order to ascertain some given didactic hypotheses on the didactic discontinuities between mathematics at pre-university and university levels. It is not a test designed to rank the students and return a unique score. The complete description of the questions can be seen in Orús and Gregori (2008). With ADA, we could obtain students' skill set profiles. In this way, students can be grouped by their similar mastery of skills. For instance, students showing consistently high levels of aptitude may be selected for an advanced class or students with similar difficulties could receive extra instruction as a group and also teaching strategies could also be adapted to suit their level. A classical way to group student skill set profiles is by using a clustering method, as carried out by Dean and Nugent (2013), but in terms of human interpretability, the central points returned by clustering tools do not seem as favorable as the extreme points returned by ADA. Results from different exploratory tools are compared.

5.1.2 Results and discussion

We would like to estimate the skill set profiles hidden in the data set. In other words, we would like to discover the data structure. Our intuition tells us that skill sets vary continuously across students, i.e. we do not expect there to be clearly differentiated (separate) groups of students with different abilities. Even so, CLA has been used to generate groups of students with similar skill set profiles (Chiu, Douglas and Li, 2009; Dean and Nugent, 2013). Here, we are going to consider the raw binary data and let the data speak for themselves, as ADA is a data-driven method. We compare the ADA solution with others from well-established unsupervised techniques introduced in Section 2.4 to highlight the information about the quality understanding of data provided by ADA.

For the sake of brevity and as an illustrative example, we examine the results of $k = 3$. The RSS elbow for ADA and the Bayesian Information Criterion (BIC) elbow for LCA are found at $k = 3$ (see Figure 3). According to the silhouette coefficient (a method of interpretation and validation of consistency within clusters of data, see Kaufman and Rousseeuw (1990) for details), the optimal number of clusters are $k = 2$ and $k = 3$ for PAM. However, the highest value of the silhouette coefficient is 0.22 (for $k = 2$ and $k = 3$ clusters), which means that no substantial cluster structure was found, as we predicted. We perform an h-plot (a multidimensional scaling method that is particularly suited for representing non-Euclidean dissimilarities, see Epifanio (2013) for details) on the dissimilarities used by PAM to graphically summarize the data set and to visualize the obtained clusters by PAM in two dimensions (see Figure 4). Effectively, separate clusters do not seem to exist.

This is also corroborated by Figure 5, where the students' scores from HOMALS are plotted in two dimensions. As regards the interpretation of the dimensions of HOMALS,

the loadings are displayed in Figure 6 and Table 2 shows their exact values, together with the number of correct answers. As also happens with PCA, their interpretation is not always easy and immediate. For the first dimension, all the coefficients are positive (as a measure of size), which can indicate a kind of sum score. The highest coefficients more or less correspond to the last questions of the test, which fewer students answered correctly. The second dimension compares, above all, questions 4, 5, 6a and 6b (with high positive coefficients) with 13a and 13b (with low negative values), while in the third dimension, questions 1, 3, 7, 8 and 10 (with high positive coefficients) are compared with 14a and 14b (with low negative values). However, we do know how the meaning of these contradistinctions is interpreted.

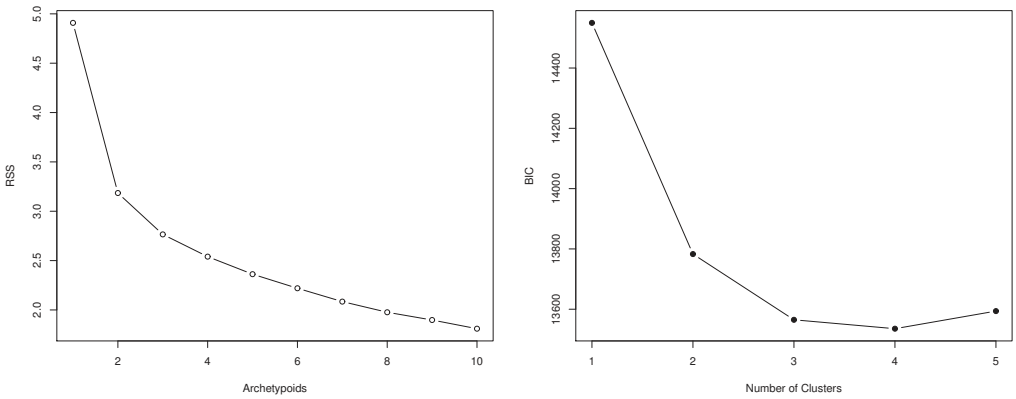

Figure 3: *Initial mathematical skills test data: Screeplot of ADA (left-hand panel); screeplot of LCA (right-hand panel).*

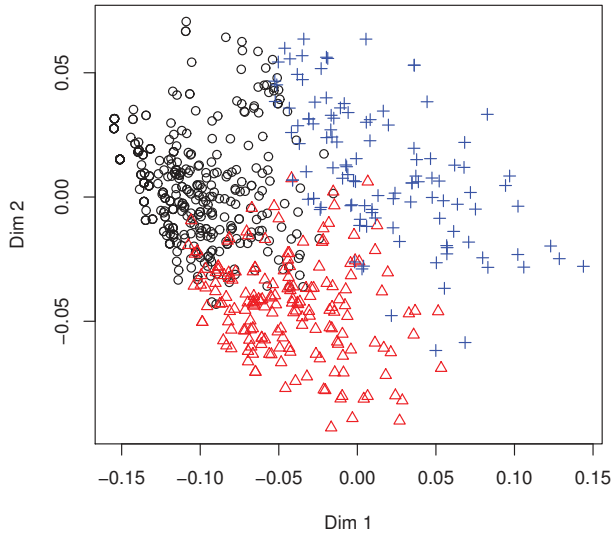

Figure 4: *H-plot of dissimilarities for the initial mathematical skills test data. We perform PAM. The black circles represent data points assigned to the first cluster, the red triangles to the second cluster and the blue crosses to the third cluster.*

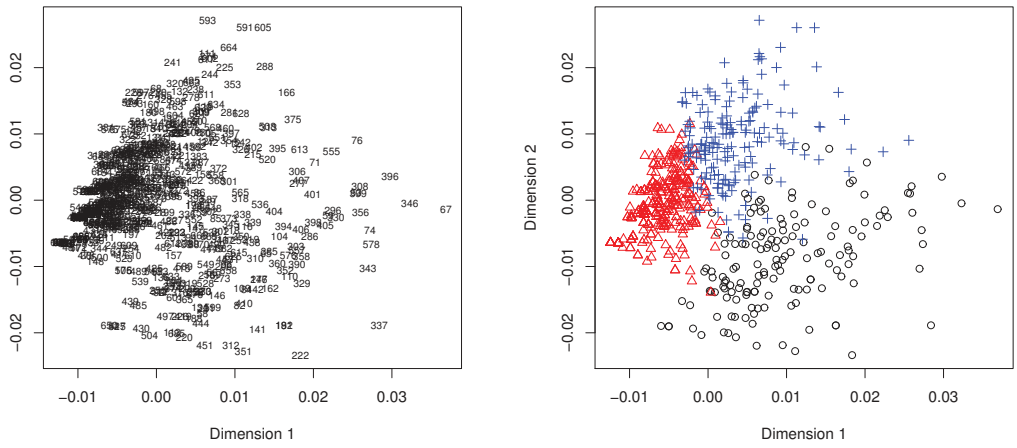

Figure 5: *HOMALS of the initial mathematical skills test data. Plot of students' scores. The numbers indicate the code of each of the 690 students (left-hand panel). We perform LCA. The black circles represent data points assigned to the first cluster, the red triangles to the second cluster and the blue crosses to the third cluster (right-hand panel).*

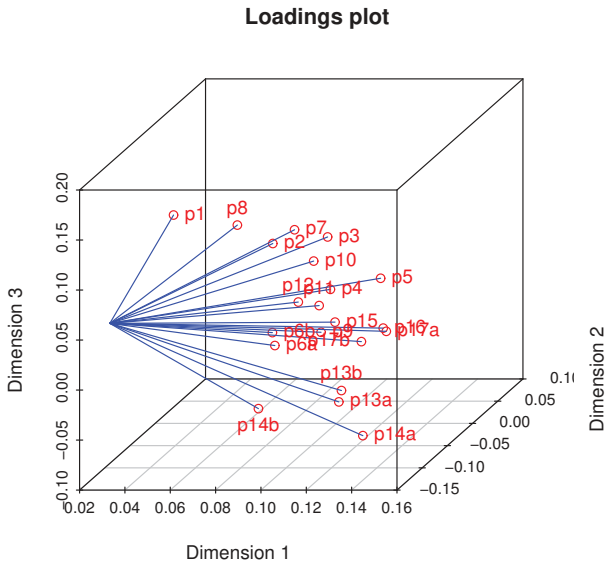

Figure 6: *HOMALS of the initial mathematical skills test data. Loadings plot.*

LCA returns the conditional item response probabilities by outcome variable for each class. Table 3 lists these probabilities for correct answer. The predicted classes for each student are shown in Figure 5, since the profiles of cluster 1 and 3 are mainly differentiated in questions 4, 5, 13a and 13b, which are the most relevant variables of dimension 2 of HOMALS.

Table 2: *Number of correct answers and loadings of the first three dimensions by HOMALS for the initial mathematical skills test data.*

Question	No. correct answers	D1	D2	D3
1	621	0.02863	−0.00232	0.10936
2	589	0.05797	0.06268	0.05200
3	301	0.09098	0.02291	0.07621
4	233	0.07597	0.09596	−0.00863
5	253	0.09922	0.09102	0.00465
6a	231	0.05414	0.08397	−0.05947
6b	105	0.05230	0.08735	−0.04804
7	270	0.07408	0.03324	0.07881
8	140	0.0601	−0.01729	0.10596
9	109	0.07749	0.07059	−0.04028
10	202	0.09541	−0.02469	0.07325
11	71	0.07485	0.07870	−0.01717
12	329	0.08006	0.01380	0.01518
13a	177	0.12934	−0.12748	−0.02185
13b	132	0.12953	−0.12312	−0.01241
14a	114	0.11951	−0.03559	−0.09643
14b	22	0.07565	−0.04544	−0.06506
15	183	0.10749	−0.03665	0.01754
16	236	0.12062	−0.00018	−0.00465
17a	47	0.12116	0.00354	−0.00952
17b	62	0.10884	0.00955	−0.02249

Table 3 also lists the profiles of the medoids, centroids of k-means and the archetypal profiles for AA, PAA and ADA. For medoids and archetypoids, the code of the corresponding observation is also displayed. To facilitate the analysis we also show the binarized profiles of AA and PAA, referred as BAA and BPAA, respectively.

As a simple summary of the profiles, we compute the percentage of correct answers for each profile. For PAM, the percentages are 9.5%, 33.3% and 57.1%; for binarized LCA, 38.1%, 9.5% and 33.3%; for binarized k-means, 38.1%, 9.5% and 42.9%; for BAA, 9.5%, 47.6% and 61.9%; for BPAA, 9.5%, 42.9% and 57.1%; and for ADA, 57.1%, 52.4% and 9.5%, respectively. Note that the median of the percentage of correct answers in the data set is 28.6% (the minimum is 0, the first quartile is 19.1%, the third quartile is 38.1%, while the maximum is 95.2%).

One profile is repeated in all the methods, a student who only answers questions 1 and 2 correctly, i.e. a student with a serious lack of competence. We therefore concentrate the analysis on the other two profiles for each method.

Table 3: *Profiles for the initial mathematical skills test data, for PAM, LCA, k-means (k-M), AA (and binarized, BAA), PAA (and binarized, BPAA) and ADA, with k = 3. The numbers in brackets for PAM and ADA indicate the code of the representative student.*

Methods	1	2	3	4	5	6a	6b	7	8	9	10	11	12	13a	13b	14a	14b	15	16	17a	17b
PAM (661)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PAM (586)	1	1	1	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
PAM (162)	1	1	1	0	1	0	0	1	0	0	1	0	1	1	1	1	0	1	1	0	0
LCA 1	0.93	0.88	0.57	0.36	0.48	0.37	0.17	0.49	0.32	0.20	0.51	0.14	0.62	1.00	0.85	0.43	0.10	0.50	0.55	0.18	0.19
LCA 2	0.88	0.79	0.28	0.17	0.15	0.24	0.07	0.30	0.14	0.03	0.15	0.02	0.32	0.05	0	0.03	0	0.13	0.14	0	0
LCA 3	0.91	0.94	0.60	0.62	0.65	0.48	0.27	0.47	0.23	0.35	0.36	0.23	0.63	0.04	0	0.20	0.03	0.31	0.53	0.09	0.17
k-M 1	0.91	0.95	0.64	0.70	0.74	0.43	0.25	0.53	0.24	0.32	0.38	0.22	0.63	0.05	0.00	0.19	0.02	0.33	0.52	0.10	0.15
k-M 2	0.88	0.78	0.26	0.13	0.11	0.27	0.09	0.27	0.13	0.05	0.14	0.02	0.32	0.06	0.01	0.03	0.01	0.12	0.15	0.01	0.01
k-M 3	0.93	0.91	0.59	0.34	0.49	0.37	0.17	0.50	0.33	0.20	0.54	0.14	0.64	1	0.88	0.46	0.10	0.52	0.58	0.18	0.19
AA 1	0.85	0.68	0.02	0	0	0.05	0	0.04	0.05	0	0.01	0	0.16	0	0	0	0	0.01	0	0	0
AA 2	0.90	1	0.87	1	1	1	0.63	0.82	0.19	0.52	0.24	0.38	0.65	0	0	0.16	0	0.07	0.43	0.09	0.15
AA 3	1	1	0.89	0.32	0.53	0.19	0.06	0.71	0.58	0.26	1	0.17	1	1	1	0.67	0.18	1	1	0.36	0.37
BAA 1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BAA 2	1	1	1	1	1	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0
BAA 3	1	1	1	0	1	0	0	1	1	0	1	0	1	1	1	1	0	1	1	0	0
PAA 1	0.86	0.72	0.13	0	0	0	0	0.12	0.07	0	0	0	0.23	0	0	0	0	0	0	0	0
PAA 2	0.90	1	0.78	1	1	1	0.61	0.73	0.27	0.40	0.38	0.31	0.66	0	0	0	0	0	0.43	0	0
PAA 3	0.99	1	0.82	0.36	0.57	0.25	0	0.66	0.44	0.27	0.86	0.12	0.85	1	1	0.73	0.15	1	1	0.32	0.42
BPAA 1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BPAA 2	1	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
BPAA 3	1	1	1	0	1	0	0	1	0	0	1	0	1	1	1	1	0	1	1	0	0
ADA (182)	1	1	1	0	0	0	0	0	1	0	1	0	1	1	1	1	0	1	1	0	1
ADA (274)	1	1	1	1	1	1	1	1	0	1	0	0	1	0	0	0	0	0	1	0	0
ADA (1)	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

In contrast with the third archetypoid, i.e. the student with very poor skills, the first and second archetypoids correspond to students with very high percentages of correct answers. In fact, the first archetypoid corresponds to the 92nd percentile of the data set, while the second archetypoid corresponds to the 88th percentile. Nevertheless, both profiles are quite different. In fact, the Hamming distance between archetypoids 1 and 2 is 13, which means that although they answered a lot of items correctly, these correctly answered items do not coincide. In other words, archetypoids 1 and 2 are somehow complementary. Both answered items 1, 2, 3, 12 and 16 correctly, which were among the most correctly answered items. Neither of them answered items 11, 14b and 17a correctly, which were among the least correctly answered items. On the one hand, the items that archetypoid 1 answered correctly, but archetypoid 2 did not are 8, 10, 13a, 13b, 14a, 15 and 17b. These items are about nonlinear systems and linear functions. On the other hand, the items that archetypoid 2 answered correctly, but

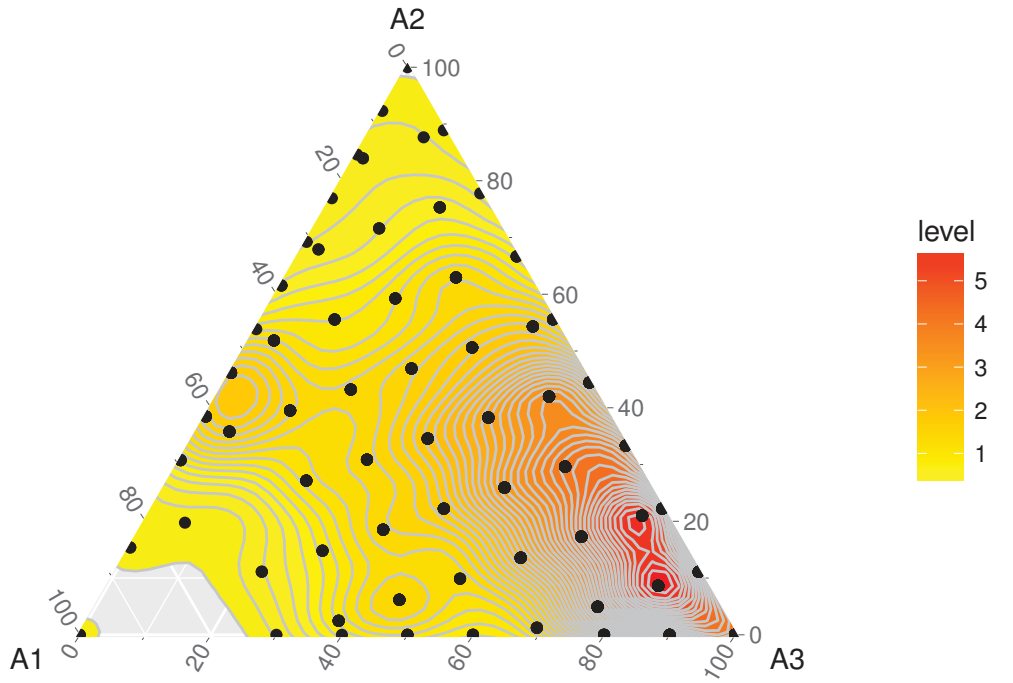

Figure 7: *Ternary plot of αs of ADA together with a plot density estimate for the initial mathematical skills test data.*

archetypoid 1 did not are 4, 5, 6a, 6b, 7 and 9. These items are about the calculation of derivatives and integrals and algebraic interpretation. The skills of these archetypoids are clear and different to each other.

We can use the alpha values for each of the students to learn about their relationship to the archetypoid profiles. The ternary plot in Figure 7 displays the alpha values that provide further insight into the data structure. Note that the majority of the data is concentrated around archetypoid 1, i.e. the one with very poor skills. If we wanted to form three groups using the alpha values, we could assign each student to the group in which their corresponding alpha is the maximum, as we did in Figure 1 (d). In this way, the number of students similar to archetypoid 1 is 113, to archetypoid 2 it is 110 and to archetypoid 3 it is 467.

The profiles of medoids 2 and 3 are not as complementary as the previous archetypoids. In fact, medoid 2 corresponds to the 56th percentile, while medoid 3 corresponds to the 92nd percentile. In this case, the percentage of correct answers for medoid 2 is not high. The Hamming distance between the two medoids is only 7. On the one hand, both answered items 1, 2, 3, 5, 7 and 12 correctly, which are the most correctly answered items. On the other hand, both failed items 6a, 6b, 8, 9, 11, 14b, 17a and 17b, many more items than in the case of ADA. The only item that medoid 2 answered correctly but

medoid 3 did not is item 4. The items that medoid 3 answered correctly but medoid 2 did not are 10, 13a, 13b, 14a, 15 and 16. It seems as if the cluster definition was guided by the number of correct answers rather than by the kind of item answered correctly. This is the reason why PAM selects medoid 2 in the middle of the data cloud. PAM, and usual clustering methods, tries to cover the set in such a way that every point is near to one medoid or one cluster center. The number of students belonging to each cluster is 398, 179 and 113, respectively. Note that the size of the cluster of students with poor skills is smaller than in the case of ADA, because some of those students are assigned to the cluster of medoid 2.

The binarized profile of LCA 1, corresponding to the 75th percentile, is similar to medoid 3, but with a lower number of correct answers (5, 7, 14a and 15), while the binarized profile of LCA 3, corresponding to the 56th percentile, is similar to medoid 2, only differentiated by two items (7 and 16). Therefore, they are even less complementary than the previous medoids. The Hamming distance between both LCA-profiles is only 5. The number of students belonging to each cluster is 155, 352 and 183, respectively. Note that the size of the cluster of students with poor skills is smaller than in the case of PAM.

The binarized profile of the first centroid of k-means, corresponding to the 75th percentile, is similar to medoid 2, only differentiated by item 16, while the binarized profile of the third centroid, corresponding to the 82nd percentile, is similar to medoid 3, but with a lower number of correct items (5, 7 and 14a). The Hamming distance between both centroids is 7. The level of complementarity between both centroids is similar to that of the medoids of PAM, but the number of correct answers of medoid 3 is higher than binarized centroid 3. The number of students belonging to each cluster is 196, 349 and 145, respectively. Note that the size of the cluster of students with poor skills is smaller than in the case of PAM, but larger than in the case of PAM for cluster 3, which in both clustering methods corresponds to the students with more correct answers.

In the clustering methods, the profiles of each cluster are not as extreme as archetypoids. Archetypoids are also more complementary, which makes it clearer to establish which kinds of features distinguish one group from another. Remember also that clustering is limited to assign each student to a group but alpha values of ADA allows to know the composition, i.e. ADA returns a richer information.

The profiles of BAA2 and BAA3 and BPAA2 and BPAA3 are quite similar to the profiles of archetypoid 2 and 1, respectively, but with slight differences. The percentiles corresponding to correctly answered items are also high, although for one of the archetypes not as high as for archetypoids. The percentiles are the 82nd and 94th for BAA2 and BAA3, and the 75th and 92nd for BPAA2 and BPAA3, respectively. Therefore, the archetypoids are more extreme than the binarized archetypes of AA and PAA. Although the profiles for BAA and BPAA are also complementary, they are not as complementary as the two archetypoids. The Hamming distance between BAA2 and BAA3 is 11, and 9 between BPAA2 and BPAA3. Archetypoids therefore manage to find more complementary profiles.

5.2 An American College Testing (ACT) Mathematics Test

5.2.1 Data

This application corresponds to the second point of view of the binary matrix (analysis of the columns). We use the same data and approach followed by Ramsay and Silverman (2002, Ch. 9) and Rossi, Wang and Ramsay (2002), although another strategy could be considered (Ramsay and Wiberg, 2017). The data used are the 0/1 (incorrect/correct) responses of 2115 males from administration of a version of the ACT Program 60-item Mathematics Test. Unlike the test introduced in Section 5.1.1, the objective of the test is to relate a student's ACT score with probability of him or her earning a college degree, i.e. to rank students. It seeks that the difficulty of questions increases as you get to higher question numbers.

Although this binary matrix does not seem curvaceous at first sight, by making the simplifying assumption that the probabilities P_{ih} (probability that examinee h gets item i right) vary in a smooth one-dimensional way across examinees, we can estimate the ability space curve that this assumption implies. Then, we can work with item response functions (IRFs) $P_i(\theta)$ as functional data (Ramsay and Silverman, 2005), where θ is the charting variable that measures out positions along the ability space curve. Or rather, we can work with the log odds-ratio functions $W_i(\theta)$, since these transformations of the item response functions have the unconstrained variation that we are used to seeing in directly observed curves. Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002) used functional PCA (FPCA) to study variations among these functions. Instead, we propose to use functional ADA (FADA), which reveals very interesting patterns that were not discovered with FPCA.

Note that in the literature, we find other terms for IRFs, such as option characteristic curves, category characteristic curves, operating characteristic curves, category response functions, item category response functions or option response functions (Mazza, Punzo and McGuire, 2014).

5.2.2 Results and discussion

As mentioned previously, we used the same data and approach followed by Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002) to estimate IRFs, $P_i(\theta)$, and their logit functions, $W_i(\theta) = \log(P_i(\theta)/(1 - P_i(\theta)))$. In particular, a penalized EM algorithm was used and functions were expanded by terms of 11 B-spline basis functions using equally spaced knots. Figure 8 displays the estimated IRFs, $\exp(W_i(\theta))/(1 + \exp(W_i(\theta)))$, and their log odds-ratio functions $W_i(\theta)$ for the 60 items. As expected, this kind of graphs with superimposed curves is largely uninformative and aesthetically unappealing (Jones and Rice, 1992).

To explore a set of curves Jones and Rice (1992) proposed the use of functions with extreme principal component scores. This could be viewed as finding the archetypoid functions. Nevertheless, the aim of PCA is not to recover extreme patterns. In fact,

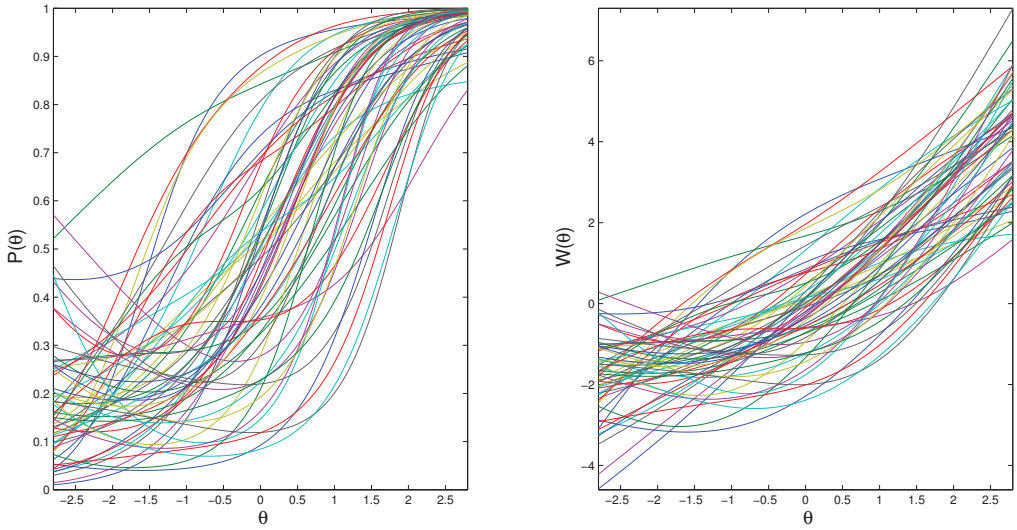

Figure 8: *Estimated IRFs (left-hand panel) and log odds-ratio functions (right-hand panel) for the ACT math exam estimated from the male data.*

curves with extreme PCA scores do not necessarily correspond to archetypal observations. This is discussed in Cutler and Breiman (1994) and shown in Epifanio, Vinué and Alemany (2013) through an example where archetypes could not be restored with PCA, even if all the components had been considered. Not only that, Stone and Cutler (1996) also showed that AA may be more appropriate than PCA when the data do not have elliptical distributions.

In order to show the advantages of ADA over PCA, we compute FPCA and FADA for $W(\theta)$, since they are unconstrained, therefore making them more appropriate for PCA application than the bounded $P_i(\theta)$. This is not a problem with FADA as it works with convex combinations. Figure 9 displays the first four PCs after a varimax rotation having been back-transformed to their probability counterparts, as performed by Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002). We base the interpretation of each PC on the detailed description carried out by Ramsay and Silverman (2002, Ch. 9).

The percentage of total variation explained by those four components is nearly 100%, while the percentage explained by each component is reported in Figure 9. The first component concentrates on the middle part of the ability range, in such a way that an item with a high (low) score in that component has a higher (lower) slope than the mean from approximately 0 to 2, i.e. it quantifies a discriminability trade-off between average students and those with rather high abilities. Analogously, the fourth component quantifies a discriminability trade-off between average examinees and those with rather low abilities. On the contrary, the second component concentrates on the upper end of the ability range. As Rossi et al. (2002) explained, the 3PL model is not well suited to modeling this type of variation. An item with a low score on this component is good at sorting out very high ability students from others of moderately high ability, whereas if

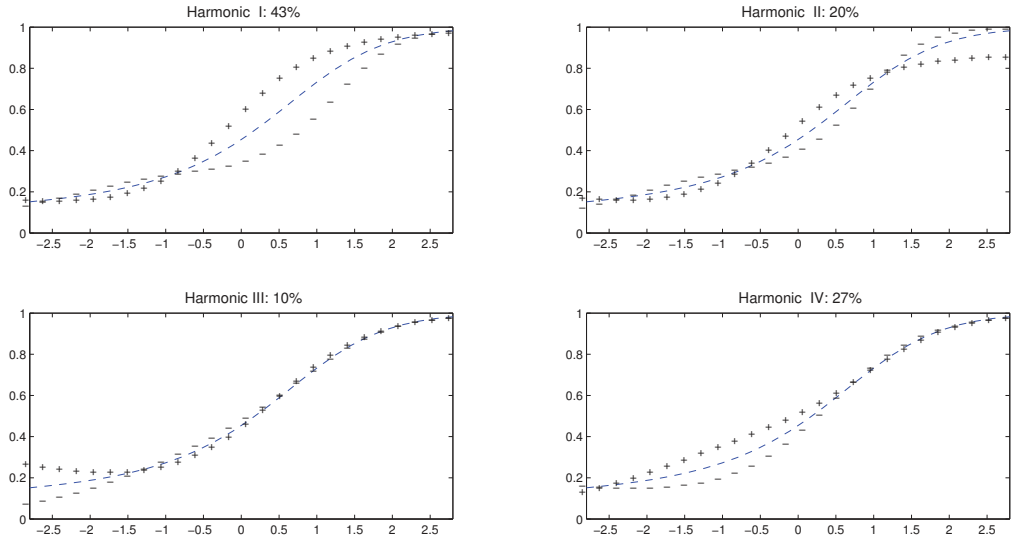

Figure 9: *The first four functional PCs in IRFs after VARIMAX rotation. Plus (negative) signs indicate the effect of adding (subtracting) a multiple of a component to the mean function. The mean function is the dashed line.*

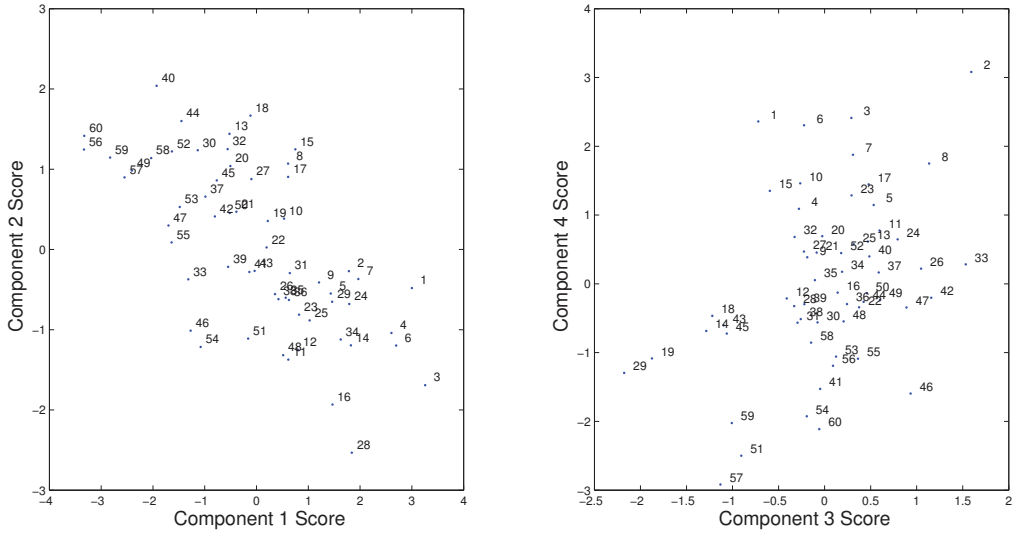

Figure 10: *Bivariate plots of principal component scores of IRFs. PC1 versus PC2 (left-hand panel); PC3 versus PC4 (right-hand panel).*

the score for this item is high, it will discriminate well among most of the population but will be found to be of approximately equal difficulty by all the very good students. Nevertheless, conclusions on the extreme part of the ability range should be made with

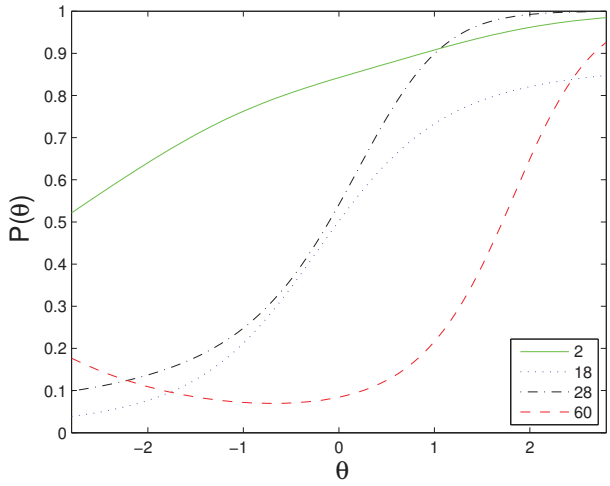

Figure 11: *ACT data: The four IRF archetypoids are items 2, 18, 28 and 60. See the legend inside the plot.*

caution, since the estimation is carried out using a relatively small numbers of students. The third component also accounts for variation in the characteristics of test items in the extreme ability range, but now in low ability ranges. PC scores for these four components can be seen in Figure 10. Note that to evaluate the 4 PC scores simultaneously and combine them to give an idea about each item, it is not easily comprehensible or human-readable.

Figure 11 displays the archetypoids for $k = 4$ explaining 97% of the variability, which is nearly as high as FPCA. The archetypoids are items 2, 18, 28 and 60. These four items describe the extreme patterns found in the sample. Item 2 has very high scores in PC 3 and PC 4, high scores in PC 1 and a score of nearly zero for PC 2. Its IRF is quite flat with a very slight slope, it seems to be a very easy item, with high probabilities of success throughout the ability range. The other archetypoids discriminate better between low and high ability students but in very different ways. Item 18 has a very high score for PC 2 and a negative score for PC 3, but nearly zero for PC 1 and PC 4. It is an item that is quite difficult even for the students in the very high ability range. The IRF of item 28 is quite similar to that of item 18 for the low ability range until θ 0, but its slope for the high ability range is higher, and the probabilities of success are higher than 0.9 for θs higher than 1. On the contrary, the probabilities of success of the IRF of item 60 are quite low as far as 1.5, which means that it is a difficult item, but the probabilities of success for the best students are high. In fact, the probabilities of success for item 60 for θ higher than 2 are higher than those of item 18. Item 28 has high score for PC 1 and low score for PC2, while it has a score of nearly zero for PC 3 and PC 4. However, item 60 has low scores for PC 1 and PC 4, a high score for PC 2 and nearly zero for PC 3. In other words, it would have been very difficult to guess the extreme representatives of the sample returned by ADA from an analysis of the scores in Figure 10.

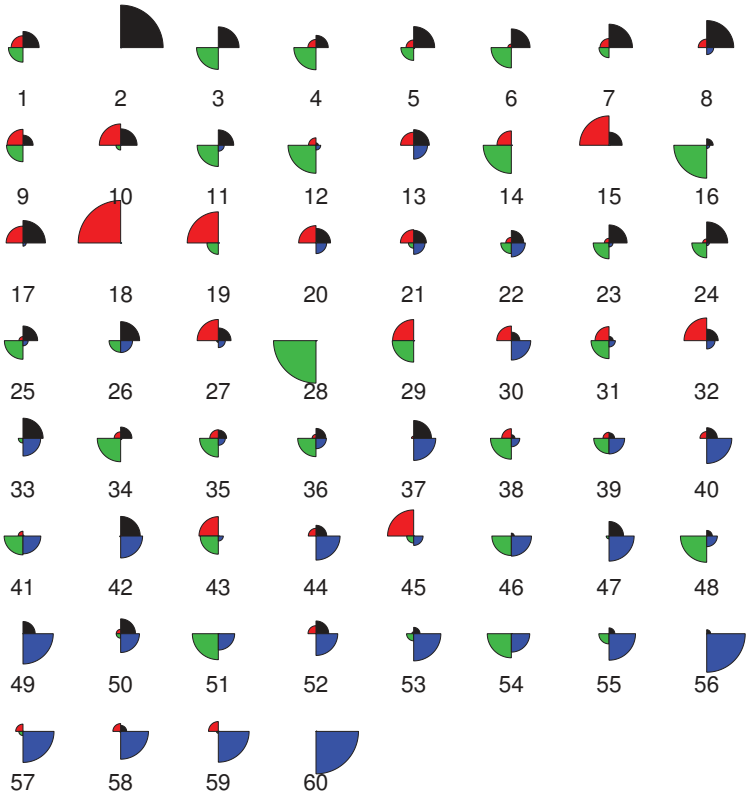

Figure 12: *Star plots of the alphas of each archetypoid for IRFs. The item number appears below each plot. The archetypoids are 2, 18, 28 and 60.*

The alpha values (from 0 to 1) tell us about the contribution of each archetypoid to each item. Remember that they add up to 1. Figure 12 shows star plots of the alpha values for each archetypoid, thus providing a complete human-readable view of the data set. The 4 alpha values in this case are represented starting on the right and going counter-clockwise around the circle. The size of each alpha is shown by the radius of the segment representing it. The items that are similar to the archetypoids can be clearly seen (for example, 7 and 8 are somehow similar to 2; 15 and 19 are somehow similar to 18; 14 and 16 are somehow similar to 28; and 56, 57 and 59 are similar to 60), as can the items that are a mixture of several archetypoids (for example, item 1 is a mixture of mainly item 2, together with items 28 and 18, to a lesser extent). Item 1 was selected by Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002) as an example of a low difficulty item, although it seems that item 2 would be a better representative of this kind of item. Item 9 was selected by Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002) as an example of a medium difficulty item, and it is mainly a mixture of items 18 and 28. Finally, item 59 was selected by Ramsay and Silverman (2002, Ch. 9) and Rossi et al. (2002) as an example of a hard item. Item 59 was mainly explained by item 60.

Results of applying FADA with kernel and parametric IRF estimates are discussed in the Supplementary Material.

6 Conclusion

We have proposed to find archetypal patterns in binary data using ADA for a better understanding of a data set. A simulation study and results provided in two applications have highlighted the benefits of ADA for binary questionnaires as an alternative that can be used instead of (or in addition to) other established methodologies.

Although, much of statistics is based on the idea that averaging over many elements of a data set is a good thing to do, in this paper we adopt a different perspective. We have selected a small number of representative observations, archetypal observations, and the data composition is explained through mixtures of those extreme observations. We have shown that this can be highly informative and is a useful tool for making a data set more "human-readable", even to non-experts.

In the first application, we have shown how ADA returns the most complementary profiles, which can be more useful in order to establish groups of students with similar mastery of skills. Furthermore, ADA returns composition information of each observation through alpha values, which is a richer information than the simple assignation to groups returned by CLA. In the second application, FADA has discovered the extreme patterns in the data, which cannot be recovered by FPCA. Furthermore, we have explained each item of the ACT math exam as a percentage of the archetypal items, which is easily understandable even for non-experts.

As regards future work, throughout the paper all variables share the same weight, but for certain situations some variables could have more weight in RSS. Another direction of future work would be to consider ADA for nominal observations, for example, by converting those variables into dummy variables, i.e. with binary codes. Furthermore, this work is limited to binary data, but questionnaires can also have Likert-type scale responses. Therefore, archetypal techniques for ordinal data would be very valuable. Another not so immediate extension, would be to consider the case of mixed data, with real valued and categorical data, together with missing data. Finally, from the computational point of view, in case of working with a very big data set, the ADA algorithm described in Section 2.1.1 could be slow. In that case, a recent alternative implemented in the R package **adamethods** (Vinue and Epifanio, 2019) for computing ADA with large data sets could be used.

Acknowledgments

This work is supported by the following grants: DPI2017-87333-R from the Spanish Ministry of Science, Innovation and Universities (AEI/FEDER, EU) and UJI-B2017-13 from Universitat Jaume I.

References

Alcacer, A., Epifanio, I., Ibáñez, M. V., Simó, A. and Ballester, A. (2020). A data-driven classification of 3D foot types by archetypal shapes based on landmarks. *PLOS ONE*, 15, 1–19.

Cabero, I. and Epifanio, I. (2019). Archetypal analysis: an alternative to clustering for unsupervised texture segmentation. *Image Analysis & Stereology*, 38, 151–160.

Canhasi, E. and Kononenko, I. (2013). Multi-document summarization via archetypal analysis of the content-graph joint model. *Knowledge and Information Systems*, 1–22.

Canhasi, E. and Kononenko, I. (2014). Weighted archetypal analysis of the multi-element graph for query-focused multi-document summarization. *Expert Systems with Applications*, 41, 535–543.

Chan, B., Mitchell, D. and Cram, L. (2003). Archetypal analysis of galaxy spectra. *Monthly Notices of the Royal Astronomical Society*, 338, 790–795.

Chiu, C.-Y., Douglas, J. A. and Li, X. (2009). Cluster analysis for cognitive diagnosis: Theory and applications. *Psychometrika*, 74, 633.

Cutler, A. and Breiman, L. (1994). Archetypal analysis. *Technometrics*, 36, 338–347.

Davis, T. and Love, B. (2010). Memory for category information is idealized through contrast with competing options. *Psychological Science*, 21, 234–242.

de Leeuw, J. and Mair, P. (2009). Gifi methods for optimal scaling in R: The package homals. *Journal of Statistical Software*, 31, 1–20.

Dean, N. and Nugent, R. (2013). Clustering student skill set profiles in a unit hypercube using mixtures of multivariate betas. *Advances in Data Analysis and Classification*, 7, 339–357.

D'Esposito, M. R., Palumbo, F. and Ragozini, G. (2012). Interval archetypes: a new tool for interval data analysis. *Statistical Analysis and Data Mining*, 5, 322–335.

Epifanio, I. (2013). H-plots for displaying nonmetric dissimilarity matrices. *Statistical Analysis and Data Mining*, 6, 136–143.

Epifanio, I. (2016). Functional archetype and archetypoid analysis. *Computational Statistics & Data Analysis*, 104, 24–34.

Epifanio, I., Ibáñez, M. V. and Simó, A. (2018). Archetypal shapes based on landmarks and extension to handle missing data. *Advances in Data Analysis and Classification*, 12, 705–735.

Epifanio, I., Ibáñez, M. V. and Simó, A. (2020). Archetypal analysis with missing data: see all samples by looking at a few based on extreme profiles. *The American Statistician*, 74, 169–183.

Epifanio, I., Vinué, G. and Alemany, S. (2013). Archetypal analysis: contributions for estimating boundary cases in multivariate accommodation problem. *Computers & Industrial Engineering*, 64, 757–765.

Eugster, M. J. and Leisch, F. (2009). From Spider-Man to Hero - Archetypal Analysis in R. *Journal of Statistical Software*, 30, 1–23.

Eugster, M. J. A. (2012). Performance profiles based on archetypal athletes. *International Journal of Performance Analysis in Sport*, 12, 166–187.

Fernandez, M. and Barnard, A. S. (2015). Identification of nanoparticle prototypes and archetypes. *ACS Nano*, 9, 11980–11992.

Fletcher, R. (2000). *Practical Methods of Optimization* (Second ed.). John Wiley & Sons.

Flynt, A. and Dean, N. (2016). A survey of popular R packages for cluster analysis. *Journal of Educational and Behavioral Statistics*, 41, 205–225.

Friedman, J. H. and Tukey, J. W. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers*, C-23, 881–890.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27, 857–871.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning. Data mining, inference and prediction*. 2nd ed., Springer-Verlag.

Henry, D., Dymnicki, A. B., Mohatt, N., Allen, J. and Kelly, J. G. (2015). Clustering methods with qualitative data: a mixed-methods approach for prevention research with small samples. *Prevention Science*, 16, 1007–1016.

Hinrich, J. L., Bardenfleth, S. E., Roge, R. E., Churchill, N. W., Madsen, K. H. and Mørup, M. (2016). Archetypal analysis for modeling multisubject fMRI data. *IEEE Journal on Selected Topics in Signal Processing*, 10, 1160–1171.

IBM Support (2016). Clustering binary data with K-Means (should be avoided). http://www-01.ibm.com/support/docview.wss?uid=swg21477401. Accessed: 2018-07-09.

Jones, M. C. and Rice, J. A. (1992). Displaying the important features of large collections of similar curves. *The American Statistician*, 46, 140–145.

Kaufman, L. and Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley.

Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems*. Prentice Hall.

Li, S., Wang, P., Louviere, J. and Carson, R. (2003). Archetypal Analysis: A New Way To Segment Markets Based On Extreme Individuals. In *ANZMAC 2003 Conference Proceedings*, pp. 1674–1679.

Linzer, D. A. and Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42, 1–29.

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28, 129–137.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. and Hornik, K. (2018). *Cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.7-1.

Makowski, D. (2016). *Package 'neuropsychology': An R Toolbox for Psychologists, Neuropsychologists and Neuroscientists*. (0.5.0).

Mazza, A., Punzo, A. and McGuire, B. (2014). KernSmoothIRT: An R package for kernel smoothing in item response theory. *Journal of Statistical Software*, 58, 1–34.

Midgley, D. and Venaik, S. (2013). Marketing strategy in MNC subsidiaries: pure versus hybrid archetypes. In *P. McDougall-Covin and T. Kiyak, Proceedings of the 55th Annual Meeting of the Academy of International Business*, pp. 215–216.

Millán-Roures, L., Epifanio, I. and Martínez, V. (2018). Detection of anomalies in water networks by functional data analysis. *Mathematical Problems in Engineering*, 2018 (Article ID 5129735), 13.

Moliner, J. and Epifanio, I. (2019). Robust multivariate and functional archetypal analysis with application to financial time series analysis. *Physica A: Statistical Mechanics and its Applications*, 519, 195–208.

Mørup, M. and Hansen, L. K. (2012). Archetypal analysis for machine learning and data mining. *Neurocomputing*, 80, 54–63.

Orús, P. and Gregori, P. (2008). *Fictitious Pupils and Implicative Analysis: a Case Study*, pp. 321–345. Berlin, Heidelberg: Springer.

Pawlowsky-Glahn, V., Egozcue, J. J. and Tolosana-Delgado, R. (2015). *Modeling and Analysis of Compositional Data*. John Wiley & Sons.

Porzio, G. C., Ragozini, G. and Vistocco, D. (2008). On the use of archetypes as benchmarks. *Applied Stochastic Models in Business and Industry*, 24, 419–437.

R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Ragozini, G. and D'Esposito, M. R. (2015). Archetypal networks. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, New York, NY, USA, pp. 807–814. ACM.

Ragozini, G., Palumbo, F. and D'Esposito, M. R. (2017). Archetypal analysis for data-driven prototype identification. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 10, 6–20.

Ramsay, J. O. and Silverman, B. W. (2002). *Applied Functional Data Analysis*. Springer.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis* (2nd ed.). Springer.

Ramsay, J. O. and Wiberg, M. (2017). A strategy for replacing sum scoring. *Journal of Educational and Behavioral Statistics*, 42, 282–307.

Rossi, N., Wang, X. and Ramsay, J. O. (2002). Nonparametric item response function estimates with the EM algorithm. *Journal of Educational and Behavioral Statistics*, 27, 291–317.

Seth, S. and Eugster, M. J. A. (2016a). Archetypal analysis for nominal observations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 849–861.

Seth, S. and Eugster, M. J. A. (2016b). Probabilistic archetypal analysis. *Machine Learning*, 102, 85–113.

Slater, S., Joksimović, S., Kovanovic, V., Baker, R. S. and Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, 42, 85–106.

Steinley, D. (2006). K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, 59, 1–34.

Steinschneider, S. and Lall, U. (2015). Daily precipitation and tropical moisture exports across the Eastern United States: An application of archetypal analysis to identify spatiotemporal structure. *Journal of Climate*, 28, 8585–8602.

Stone, E. and Cutler, A. (1996). Introduction to archetypal analysis of spatio-temporal dynamics. *Physica D: Nonlinear Phenomena*, 96, 110–131.

Su, Z., Hao, Z., Yuan, F., Chen, X. and Cao, Q. (2017). Spatiotemporal variability of extreme summer precipitation over the Yangtze river basin and the associations with climate patterns. *Water*, 9.

Theodosiou, T., Kazanidis, I., Valsamidis, S. and Kontogiannis, S. (2013). Courseware usage archetyping. In *Proceedings of the 17th Panhellenic Conference on Informatics*, PCI '13, New York, NY, USA, pp. 243–249. ACM.

Thøgersen, J. C., Mørup, M., Damkiær, S., Molin, S. and Jelsbak, L. (2013). Archetypal analysis of diverse pseudomonas aeruginosa transcriptomes reveals adaptation in cystic fibrosis airways. *BMC Bioinformatics*, 14, 279.

Thurau, C., Kersting, K., Wahabzada, M. and Bauckhage, C. (2012). Descriptive matrix factorization for sustainability: Adopting the principle of opposites. *Data Mining and Knowledge Discovery*, 24, 325–354.

Tsanousa, A., Laskaris, N. and Angelis, L. (2015). A novel single-trial methodology for studying brain response variability based on archetypal analysis. *Expert Systems with Applications*, 42, 8454–8462.

Unwin, A. (2010). Exploratory data analysis. In P. Peterson, E. Baker, and B. McGaw (Eds.), *International Encyclopedia of Education (Third Edition)*, pp. 156–161. Oxford: Elsevier.

Vinué, G. (2017). Anthropometry: An R package for analysis of anthropometric data. *Journal of Statistical Software*, 77, 1–39.

Vinué, G. and Epifanio, I. (2017). Archetypoid analysis for sports analytics. *Data Mining and Knowledge Discovery*, 31, 1643–1677.

Vinue, G. and Epifanio, I. (2019). *Adamethods: Archetypoid Algorithms and Anomaly Detection*. R package version 1.2.

Vinué, G. and Epifanio, I. (2019). Forecasting basketball players' performance using sparse functional data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 12, 534–547.

Vinué, G., Epifanio, I. and Alemany, S. (2015). Archetypoids: A new approach to define representative archetypal data. *Computational Statistics & Data Analysis*, 87, 102–115.

Wu, C., Kamar, E. and Horvitz, E. (2016). Clustering for set partitioning with a case study in ridesharing. In *IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1384–1388.

SORT 44 (1) January-June 2020, 67-98 DOI: 10.2436/20.8080.02.95

Integer constraints for enhancing interpretability in linear regression

Emilio Carrizosa[1], Alba V. Olivares-Nadal[2] and Pepa Ramírez-Cobo[3]

Abstract

One of the main challenges researchers face is to identify the most relevant features in a prediction model. As a consequence, many regularized methods seeking sparsity have flourished. Although sparse, their solutions may not be interpretable in the presence of spurious coefficients and correlated features. In this paper we aim to enhance interpretability in linear regression in presence of multicollinearity by: (i) forcing the sign of the estimated coefficients to be consistent with the sign of the correlations between predictors, and (ii) avoiding spurious coefficients so that only significant features are represented in the model. This will be addressed by modelling constraints and adding them to an optimization problem expressing some estimation procedure such as ordinary least squares or the lasso. The so-obtained constrained regression models will become Mixed Integer Quadratic Problems. The numerical experiments carried out on real and simulated datasets show that tightening the search space of some standard linear regression models by adding the constraints modelling (i) and/or (ii) help to improve the sparsity and interpretability of the solutions with competitive predictive quality.

MSC: 62J05, 90C11.

Keywords: Linear regression, Multicollinearity, Sparsity, Cardinality constraint, Mixed Integer Non Linear Programming.

1 Introduction

A plethora of real world data involve multiple features interacting between them. As a consequence, one of the most common research challenges is trying to predict a variable by making use of attributes that are deterministic or easier to access. A widely studied tool to achieve this is the linear regression model

$$\mathbf{Y} = \beta_0 + \boldsymbol{\beta}\mathbf{X} + \mathbf{a} \tag{1}$$

[1] Institute of Mathematics of the University of Seville (IMUS).

[2] The University of Chicago Booth School of Business, 5751 S. Woodlawn Ave., Chicago, Illinois 60637. Email: alba.nadal@chicagobooth.edu

[3] Department of Statistics and Operational Research, Universidad de Cádiz.

Received: January 2019
Accepted: April 2020

where $\mathbf{Y} = (y_1, \ldots, y_K)'$ contains the K realizations of the random variable to be predicted, $\mathbf{X} \in \mathbb{R}^{K \times N}$ contains the observations of the attributes $X^1, \ldots, X^N$ that influence on $\mathbf{Y}$, and $\mathbf{a} \in \mathbb{R}^K$ denotes the error term. In practice, the coefficients $\boldsymbol{\beta}$ need to be estimated and thus, the user needs to select an estimation method, which is usually derived from solving an optimization problem of the form:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\beta}} & f(\boldsymbol{\beta}) \\ \text{s.t} & \boldsymbol{\beta} \in \mathcal{B} \end{array} \tag{2}$$

where $\mathcal{B}$ denotes the feasible region.

Since the data collection technologies are improving altogether with communication systems and computers' memories and processors, the dimension of the data sets to be handled is increasing drastically. As a consequence, nowadays researchers aim to return an interpretable output which explains the main interactions between the features that conform the pile of data and the dependent variables. Usually, this is understood as a problem of choosing the most relevant features for prediction (Friedman, Hastie and Tibshirani, 2001; Cai, Tsay and Chen, 2009; Hastie, Tibshirani and Wainwright, 2015). Sparse methods will yield solutions $\boldsymbol{\beta}$ in (2) with a large number of zero coefficients, in which only the most significant features are associated with the non-zeroes (Tibshirani, 1996; Hastie et al., 2015). Although sparsity may be a desirable property for our solution, we should take into account that other characteristics need to be sought in order to obtain a more interpretable output. First, correlated variables can provide highly variable estimated coefficients that make it difficult to understand the impact of a feature on the predictive variable. Second, spurious coefficients complicate the judgement of whether a feature is truly relevant for prediction or not. We will explain these two issues with further detail in what follows and motivate why we aim to alleviate them in this paper while still seeking for a sparse solution.

It is known that ordinary least squares (OLS) provides solutions that may be highly dense. A good representative of a possibly sparse estimation method in the form (2) is the lasso (Tibshirani, 1996), which adds a ℓ_1-norm penalization term to the OLS objective:

$$\begin{array}{l} f(\boldsymbol{\beta}) = \|\mathbf{y} - \beta_0 - \boldsymbol{\beta}\mathbf{X}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1 \\ \mathcal{B} = \mathbb{R}^N. \end{array} \tag{3}$$

The lasso encompasses OLS when the penalty parameter λ equals zero, but when λ increases the solution becomes more sparse. The lasso is computationally feasible and, under certain technical conditions on the data matrix $\mathbf{X}$, it enjoys good statistical properties, see Friedman et al. (2001); Bühlmann and van de Geer (2011). However the lasso also presents certain shortcomings well-documented in the literature (for a brief review, see Bertsimas et al. (2016) and the references therein). In particular, it is known that estimation through OLS or the lasso may be quite unstable in the presence of strong

collinearity on the data (Silvey, 1969; Sengupta and Bhimasankaram, 1997; Hesterberg et al., 2008). On one hand, the presence of correlated variables may yield a high variability in the estimated coefficients, complicating thus the interpretation of the results (Farrar and Glauber, 1967; Watson and Teelucksingh, 2002; Montgomery, Peck and Vining, 2012). On the other hand, a consequence of collinearity that leads to problems for interpreting the effect of the regressors is that two variables that are highly positively (negatively) correlated may have associated estimated coefficients with different (same) signs. This problem can be illustrated through the following numerical example in Hesterberg et al. (2008). Consider the `diabetes` database (Efron and Hastie, 2003), which consists of the measures of 10 variables (age, sex, body mass index, average blood pressure and six different blood serums) on 442 patients. The top panel of Figure 1 depicts the path of solutions of the lasso for this database; that is to say, the estimates of the coefficients $\boldsymbol{\beta}$ obtained are depicted against the different values of the penalty λ. As noted by Hesterberg et al. (2008), features *tc* and *ldl* (bottom left panel), have a correlation of 0.89. However, their estimated coefficients take opposite signs, which is in contradiction with their dependence degree. Similarly, the coefficients for variables *hdl* and *tch* (bottom right panel), which show a correlation of -0.73, have the same sign when estimated by the OLS (the case $\lambda = 0$ in Figure 1). Hence it seems that the coefficients of highly correlated variables may take values that compensate each other. Finally, an additional inconsistency is that the sign of the estimated coefficient of *hdl* (squared blue line, left bottom panel) varies depending on the level of sparsity required.

The negative effects of collinearity have been differently addressed in the literature. On one hand, some authors (Chatterjee and Hadi, 2015; Montgomery et al., 2012) suggest to remove variables that are highly correlated or unimportant, often carrying out significancy tests to determine if a variable can be discarded. However, the results of these tests may be misleading in the presence of strong collinearity (Watson and Teelucksingh, 2002). In this line, the recent paper by Bertsimas and King (2015) proposes to tighten the estimation procedure (2) by adding constraints that explicitly forbid the coefficients of variables with a high pairwise correlation to be simultaneously non-zero. Nonetheless, as it will be seen in Section 3.3.1, these approaches may be detrimental if highly correlated features own a strong predictive power. On the other hand, some authors encourage highly correlated predictors to be altogether in the model. The graph-guided fused lasso (GFlasso hereafter, proposed in Kim and Xing, 2009) encourages two highly correlated variables to have similar estimated coefficients by adding the penalization $\gamma|\beta_i - sign(\rho_{ij})\beta_j|$ to the lasso objective function, where ρ_{ij} denotes the correlation between X^i and X^j. The SRIG method (Sparse Regression Incorporating Graphical Structure Among Predictors, introduced in Yu and Liu, 2016), determines the value of the coefficient β_j not only by feature X^j but also by all features X^i such that ρ_{ij} is large in absolute value. Under certain technical conditions, the SRIG is endowed with nice properties that, for instance, ensure the recovery of the original model. However, as it will be shown in Section 3.2, these conditions may not be fulfilled in real databases, yielding outputs that may not improve the performance of the

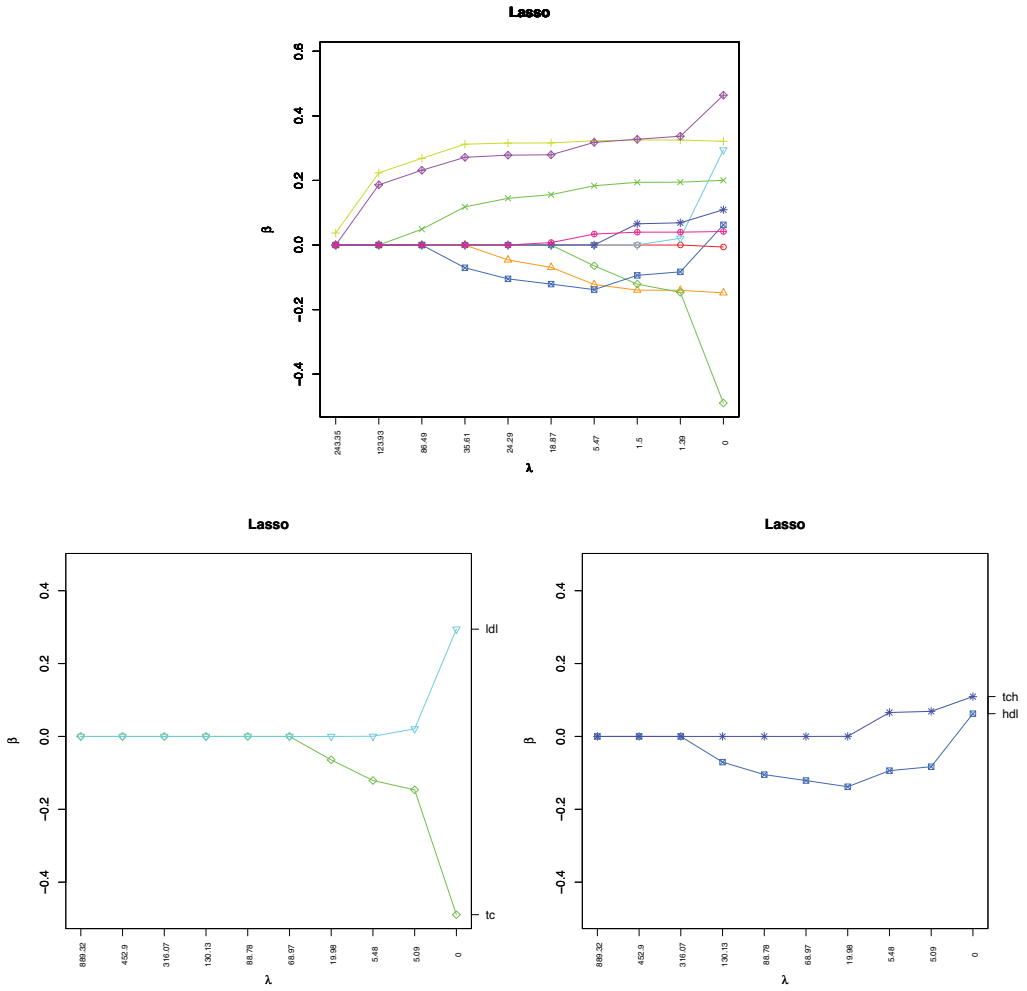

Figure 1: *Top: path of solutions of the lasso for the diabetes database (the size of the coefficients β are depicted against the values of the penalty λ). Bottom: problematic paths.*

current methodologies. Our approach is less restrictive than the previous one, since it neither encourages the removal of variables nor the presence of groups of correlated features. Instead, we propose a constraint (called *sign coherence constraint*) that aims to elude the signs' inconsistency phenomenon related to collinearity shown in Figure 1, avoiding that the coefficients of highly correlated variables compensate each other. This constraint provides more flexible models, as for some cases two highly correlated variables may appear altogether in the output, and for some other cases one feature of a highly correlated pair may be removed. This will be illustrated in Figure 5, Section 3.2. Our approach is not the first one to restrict the sign of the estimated coefficients in linear regression. For instance, Meinshausen (2013) makes use of non-negative least squares to recover the real sparsity pattern in high-dimensional data under certain conditions. Also, the LARS algorithm (Efron, Hastie, Johnstone and Tibshirani, 2004) emulates the lasso

solution by requiring the sign of the coefficients to match the sign of their correlation with the residuals. Another well-known example is the non-negative garotte (Breiman, 1995), which performs subset selection while forcing the signs of the coefficients to match the signs of the OLS estimates.

To conclude, there are further remedies to alleviate collinearity issues, consisting of harvesting more observations (Sengupta and Bhimasankaram, 1997; Montgomery et al., 2012) or applying methods that decorrelate the data (Cao, Guo and Bouman, 2010; Massy, 1965). Nevertheless, the later approaches imply the transformation of the variables and thus complicate the interpretation of the final models with respect to the original features. More recently, optimization approaches bounding the Variance Inflation Factor (VIF) and condition number of the correlation matrix have also been proposed (Tamura et al., 2019; Jou, Huang and Cho, 2014; Tamura et al., 2017).

On top of the unreliable interpretation of coefficients in presence of high correlations, the lasso also suffers a drawback, mitigated in this paper: the presence of spurious coefficients. For $\lambda > 0$ the ℓ_1-norm performs a shrinkage of the coefficients in the lasso solution that eventually attains sparsity as a side effect. However, as will be shown in the numerical section, the solutions of the lasso may be still dense for large datasets due to these spurious coefficients. In this paper we avoid this negative effect of shrinkage by defining a novel constraint (called *significance constraint*) that forces the estimated coefficients to be either zero or larger than a fixed value (to be tuned).

In summary, in this paper we model two novel constraints which will tighten the search space for $\boldsymbol{\beta}$ in Problem (2). As a result, the interpretability of the solutions is improved since (i) the signs of the coefficients are coherent with the sign of the correlations between highly or moderately correlated predictors, and (ii) the shrinkage is combatted while avoiding spurious coefficients, which may lead to the annihilation of some coefficients, thus increasing the sparsity. As will be shown in the numerical experiments, such better interpretability is obtained without damaging the predictive power of the model. When discerning the suitability of these constraints for a particular database, the user should realize that constraints modelling (i) become inactive if no highly correlated predictors are found, while constraints expressing (ii) do if all the variables have non-spurious estimated coefficients. Hence, the user might want to analyse the correlations before adding the sign coherence constraints. However, we do recommend adding the significance constraint if a regularized method is used for estimation.

The resulting optimization problems will belong to the class of Mixed Integer Quadratic Programs (MIQP), which have recently proven very suitable in different statistics problem as linear regression (Tamura et al., 2019; Bertsimas and King, 2015), time series (Carrizosa, Olivares-Nadal and Ramírez-Cobo, 2016), classification (Carrizosa, Nogales-Gómez and Morales, 2016; Carrizosa, Nogales-Gómez and Morales, 2017), or dimensionality reduction (Carrizosa and Guerrero, 2014). Indeed, Bertsimas and King (2015); Bertsimas et al. (2016) use MIQP theory to solve (in tractable way) the best

subset selection problem (Miller, 2002):

$$\begin{aligned} &f(\boldsymbol{\beta}) = \|\mathbf{y} - \beta_0 - \boldsymbol{\beta}\mathbf{X}\|_2^2 \\ &\mathcal{B} = \left\{\boldsymbol{\beta} \in \mathbb{R}^N : \|\boldsymbol{\beta}\|_0 \leq V_T\right\}, \end{aligned} \tag{4}$$

where the ℓ_0-*norm* is the cardinality function $\|\boldsymbol{\beta}\|_0 = \#(j : \beta_j \neq 0)$ and

$$\|\boldsymbol{\beta}\|_0 \leq V_T, \tag{5}$$

denotes the cardinality constraint which leads to attain the desired level of sparsity given by the value V_T. In this work, the two novel constraints (*sign coherence* and *significance* constraints) will be combined with the cardinality constraint (5), so that sparsity is also achieved in addition to a better interpretability.

The paper is structured as follows. In the next section we model the new constraints to be added to Problem (2) in order to enhance interpretability through mathematical programming. The numerical experiments are carried out in Section 3, where the estimation methods under comparison and the design of experiments are also discussed. The last section is devoted to concluding remarks and extensions.

2 Mathematical model formulation

In this paper, it is our aim to enhance the interpretability of the outputs by replacing any estimation procedure in the form (2) by:

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & f(\boldsymbol{\beta}) \\ \text{s.t} \quad & \boldsymbol{\beta} \in \mathcal{B} \cap \mathcal{S} \end{aligned} \tag{6}$$

where $\mathcal{S}$ will gather the proposed constraints. Tightening an estimation procedure by adding constraints, i.e., solving (6) instead of (2), has already been considered in the literature in order to improve the performance of linear regression estimation methods like (2), see for example Bertsimas and King (2015).

In this section we model the tightening set $\mathcal{S}$ by defining constraints that can be added to a classic (possibly sparse) linear regression estimation method (2), in order to enhance the interpretability of the outcome as well as improving its sparsity. As commented in the previous section, the first novel constraint, called *sign coherence constraint*, imposes coherence between the signs of the estimated coefficients and the signs of large pairwise correlations between predictors. The second novel constraint, the so-called *significance constraint*, allows only for truly significant features to be considered in the model. The idea is that the user should feel free to add any of these constraints, when compatible, to her selected estimation method given by (2), yielding (6).

2.1 The sign coherence constraint

The presence of correlated variables in the data is demonstrated to lead to undesired consequences, such as a high variability on the estimated coefficients and the sign inconsistencies explained in the introduction; see e.g. Bartholomew et al. (2008). As it was commented, the traditional procedure to avoid these undesired behaviour consists of removing highly correlated variables. In particular, Bertsimas and King (2015) forbids two highly correlated variables to be simultaneously non-zero. Specifically, the following pairwise correlation constraints are modelled

$$\gamma_i + \gamma_j \leq 1 \quad \forall (i,j) \in \Omega_\eta, \tag{7}$$

where $\Omega_\eta = \{(i,j) : |\rho_{ij}| \geq \eta\}$ is the set of pairs of features considered to be highly correlated, and γ_i, γ_j are defined as

$$\gamma_j \quad = \quad \begin{cases} 1 & \text{if } \beta_j \neq 0 \\ 0 & \text{if } \beta_j = 0. \end{cases} \tag{8}$$

From now on, the constraint defined by Bertsimas and King (2015) and stated as (7) will be called *correlation constraint*.

In contrast to Bertsimas and King (2015), we propose here a less restrictive approach that allows two highly correlated variables to be in the model at the same time, but forbids misleading interpretations and misrepresentative coefficients. Our aim is to avoid sign inconsistencies while allowing the model to include two correlated variables if they contribute to improve or maintain the prediction quality. Therefore, we propose to model constraints that avoid the *compensation* of coefficients for correlated variables. Under the light of the example illustrated in Figure 1, these are the requirements we aim to gather when modelling the *sign coherence* constraint:

1. The coefficients of two features that are moderately or highly positively correlated must have the same sign.
2. The coefficients of two features that are moderately or highly negatively correlated must have opposite signs.

In order to model these constraints, we introduce the following binary variables:

$$\nu_j^+ = \begin{cases} 1 & \text{if } \beta_j > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_j^- = \begin{cases} 1 & \text{if } \beta_j < 0 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, the previous requirements 1-2 can be easily formulated as constraints as:

$$\nu_i^+ + \nu_j^- \leq 1 \quad \forall (i,j) \in \Omega_\alpha^+ \tag{9}$$
$$\nu_i^- + \nu_j^+ \leq 1 \quad \forall (i,j) \in \Omega_\alpha^+ \tag{10}$$
$$\nu_i^+ + \nu_j^+ \leq 1 \quad \forall (i,j) \in \Omega_\alpha^- \tag{11}$$
$$\nu_i^- + \nu_j^- \leq 1 \quad \forall (i,j) \in \Omega_\alpha^- \tag{12}$$

where Ω_α^+ and Ω_α^- are the sets of pairs of features that are moderately or highly correlated, expressed as $\Omega_\alpha^+ = \{(i,j) : \rho_{ij} \geq \alpha\}$ and $\Omega_\alpha^- = \{(i,j) : \rho_{ij} \leq -\alpha\}$. That is to say, constraints (9)-(10) mean that, if two variables i, j are highly positively correlated (i.e. $(i,j) \in \Omega_\alpha^+$), then we do not allow one of the coefficients to be positive and the other negative. Similarly, constraints (9)-(10) imply that, if two variables i, j are highly negatively correlated (i.e. $(i,j) \in \Omega_\alpha^-$), we forbid their coefficients to be both positive or both negative.

Note that variables ν_j^+, ν_j^- are linked with γ_j, defined in Equation (8), as follows:

$$\gamma_j = \nu_j^+ + \nu_j^-,$$

and thus the cardinality constraint (5) can be also written as:

$$\sum_{j=1}^{N} (\nu_j^+ + \nu_j^-) \leq V_T.$$

In order to illustrate the impact of these constraints we compare the path of solutions depicted in Figure 1 for the lasso applied to the `diabetes` dataset, against the lasso tightened with the sign coherence constraints, as depicted in Figure 2. As it can be observed, the use of constraints (9)-(12) to tighten the feasible region of the lasso might avoid the sign of a coefficient to vary depending on the level of sparsity required, easing the interpretation of the impact of the predictors over the response variable.

2.2 The significance constraint

In this section we formulate a novel constraint that helps combatting the negative effects of shrinkage of the lasso while discarding spurious coefficients. The idea is to allow only for *significant* variables to be represented in the model and to improve the sparsity of the output. Intuitively, *large* coefficients are identified with the significance of a feature once the data are normalized. Following this reasoning we propose to establish a threshold of *significance* that a feature must be able to exceed to be allowed in the model. We model the significance constraint as follows:

$$|\beta_j| \in \{0\} \cup [\epsilon, +\infty) \qquad j = 1, \ldots, N \tag{13}$$

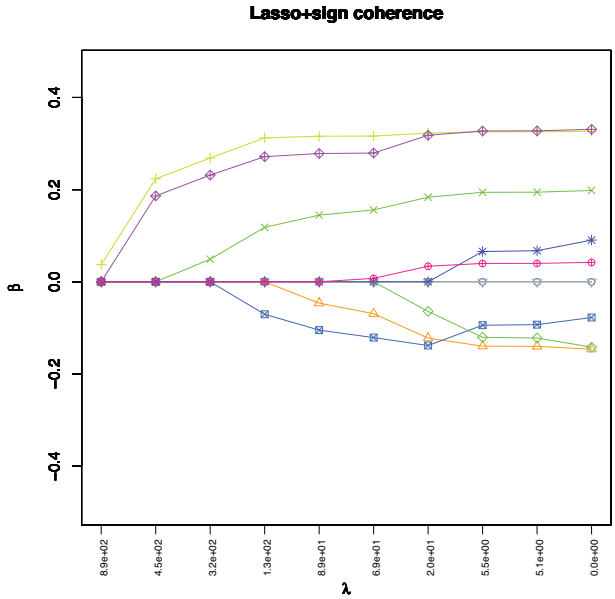

Figure 2: *Path of solutions of the tightened lasso by the sign coherence constraint for the diabetes database.*

where ϵ is called here the significance threshold, to be fixed by the user or to be tuned (see Section 3.1). In (13), a forbidden region in $(0,\epsilon)$ is defined with the aim to avoid shrinkage and to forbid spurious coefficients in the solution. This constraint was already used by Carrizosa et al. (2016) to discover potential causalities in multivariate time series. Note that using the binary variables ν_j^+ and ν_j^- defined in Section 2.1, constraints (13) can be expressed in a more manageable form via two sets of linear constraints

$$\begin{aligned} \beta_j &\geq \epsilon\nu_j^+ - \nu_j^- M \qquad \forall j = 1,\ldots,N \\ \beta_j &\leq -\epsilon\nu_j^- + \nu_j^+ M \quad \forall j = 1,\ldots,N \end{aligned} \tag{14}$$

where M is a *large* constant. This big M, often appearing when modelling problems with integer variables, is large enough so it does not exclude reasonable values of the parameters β_j (see, e.g. Camm, Raturi and Tsubakitani, 1990). In order to clarify the effect of the significance constraint (13), consider the heat map given by Figure 3. The left panel depicts the values of the estimated coefficients for the lasso, in the first column, and the lasso with the significance constraint (taking $\epsilon = 0.3$), in the second column, for the `golf2009` database (Winner, 2016). The right panel represents the values of the estimated coefficients for the OLS and its counterpart tightened with the significance constraint (taking $\epsilon = 0.05$) for the `compact` database (Torgo, 2016). Such datasets will be described with further details in Section 3. The colour represents the sign of the coefficients β_j (blue for negative, red for positive) and the intensity is related to the magnitude of such coefficients. In Figure 3 it can be observed that adding significance constraints establishes a clearer cut between zero and non-zero coefficients. Also note

that the tightened approach is not equivalent to making zero all the coefficients estimated by the lasso to be smaller in absolute value than our threshold ϵ. For instance, in the left panel β_4, estimated to be 0.143, is enlarged to $\epsilon = 0.3$, while β_3 is enlarged to -0.436 despite being estimated by a value of -0.311, which was already larger in absolute value than ϵ. Finally, it should be noted that the solutions under the significance constraint lead to an improvement of 10.58% over the out-of-sample mean squared error (MSE hereafter) of the lasso for the considered database. Moreover, on the right panel we observe that the tightened OLS shrinks coefficient $\beta_{14} = -0.024$ to zero, while enlarges coefficient $\beta_4 = 0.012$ to $\epsilon = 0.05$, despite being smaller in absolute value. Also, adding the significance constraints yields to an improvement of 1.72% over the out-of-sample MSE of the OLS.

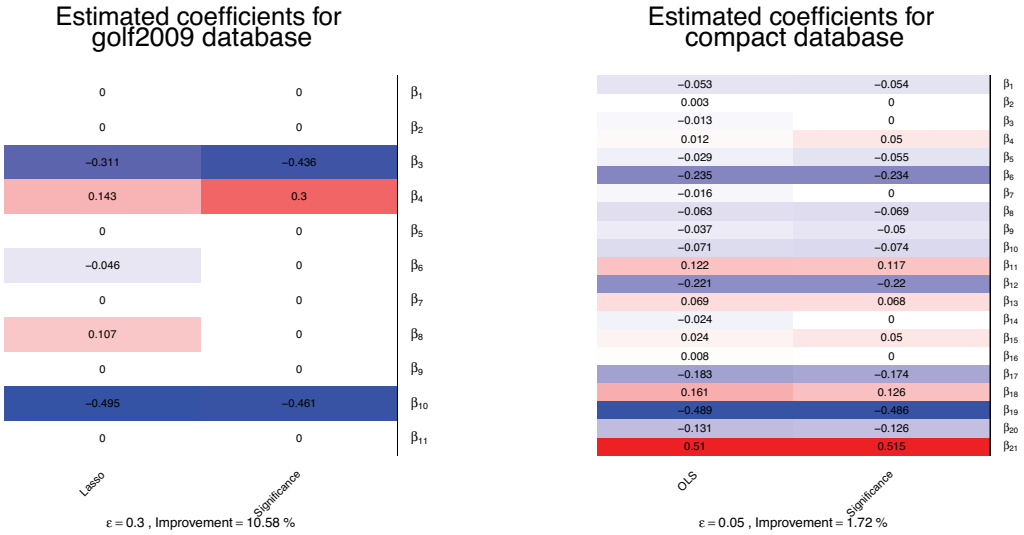

Figure 3: *Heatmaps representing the coefficients $\boldsymbol{\beta}$ estimated by the lasso (left panel) and the OLS (right panel) and their respective counterparts tightened by adding significance constraint* (13) *for the* `golf2009` *and* `compact` *datasets.*

3 Numerical illustrations

In this section we describe and undertake the numerical experiments performed to compare two benchmark estimation methods in linear regression of the form (2) against their tightened versions (6) derived from reducing the search of the coefficients $\boldsymbol{\beta} \in \mathcal{B} = \mathbb{R}^N$ to the set $\mathcal{B} \cap \mathcal{S}$, where $\mathcal{S}$ is defined through some constraints. Specifically, in the next section we outline the design of the experiments, Section 3.2 shows the results for real databases, Section 3.3.1 replicates the simulated study in Yu and Liu (2016), while in Section 3.3.2 the databases are generated following the simulations in Bertsimas and King (2015).

3.1 Design of experiments

In order to assess the impact of the novel constraints over the estimated coefficients $\boldsymbol{\beta}$ and the predictive quality of the solutions, we will analyse the differences in the performance between Problems (2) and (6). As baseline estimation methods (i.e. Problem (2)), we consider the lasso (Problem (3)) and the OLS (Problem (3) with $\lambda = 0$). Their tightened versions (Problem (6)) consider the same objective functions but reduce the search space of the coefficients $\boldsymbol{\beta}$ to the so-called tightened regions $\mathcal{B} \cap \mathcal{S}$.

In our numerical setting, we consider various tightening sets $\mathcal{S}$, whose related problems, taking form (6), are explicitly formulated in Appendix A. In order to analyse the effect of the first novel constraint proposed in this paper, the sign coherence constraint, we consider the set $\mathcal{S}_1 = \{(9)-(12)\}$. To compare our approach with the recent constraints by Bertsimas and King (2015), the correlation constraint, we will also test the performance of the set $\mathcal{S}_2 = \{(7)\}$. Both sets will be considered in the first part of Section 3.2 where the sign coherence constraint is analysed. Then, to clarify the performance of the new significance constraint, the tightening set $\mathcal{S}_3 = \{(14)\}$ will be considered in the second part of Section 3.2. Finally, we will analyse the global performance of our novel constraints when the cardinality constraint is also imposed (that is, $\mathcal{S}_4 = \{(5),(9)-(12),(14)\}$) in comparison to the tightening set of Bertsimas and King (2015), for which $\mathcal{S}_5 = \{(5),(7)\}$. In these cases we also show the predictive quality and number of non-zero coefficients for the elastic net (Enet hereafter) (Zou and Hastie, 2005), the SRIG method in Yu and Liu (2016) and the GFlasso in Kim and Xing (2009). The Enet, which trades off between lasso and ridge regression, is known to avoid erratic paths of correlated variables in the lasso (Hastie et al., 2015). In fact, for non-trivial values of the parameters, the Enet problem has a unique solution, no matter the correlations between the regressors. This shall be addressed in the last part of Section 3.2 as well as Sections 3.3.1 and 3.3.2. The Enet method was run using `R cran` package `glmnet`, and the SRIG method was run using the `R` packages recommended by the authors in Yu and Liu (2016). All the tightened procedures and the GFlasso were easily coded in the algebraic language `AMPL` (Fourer, Gay and Kernighan, 2002), but the latter was solved using `Knitro` solver. As Problems (2) and (6) are MIQPs with quadratic convex objective function and linear constraints, they were solved using `CPLEX`. For the interested reader, the code is included in Appendix D of the Supplementary Material. Even though MIQP problems may be hard to solve, the current solvers already incorporate a plethora of heuristics that turn them into highly efficient optimizers. For instance, `CPLEX` incorporates various preprocessing steps whose aim is to reduce the size of the problem and improve its formulation (Savelsbergh, 1994; Atamurk, Nemhauser and Savelsbergh, 2000). On the other hand, many other techniques and local search heuristics are implemented and implicitly run during the process (see, for instance, Danna, Rothberg and LePape, 2005; Fischetti and Lodi, 2005; Rothberg, 2007). As done in Bertsimas and King (2015), a time limit of 20 seconds was imposed to solve each MIQP for $K \leq N$, although this limit was reached

only for the largest datasets and in most cases cases the optimal solution was attained in a few seconds. For the case with $K > N$, a time limit of 40 seconds was imposed instead.

In order to make a fair comparison against existing procedures, the experiments developed here closely follow those in Bertsimas and King (2015). First, unless otherwise specified, the datasets are normalized and divided in train, test and validation sets (50%, 25% and 25% of the data, respectively). All the problems are solved in the training set, and the solution that minimizes the MSE in the test set is chosen. Two criteria are used to compare the methods, namely, the MSE and the sparsity. All the MSEs reported in this paper correspond to the values obtained in the validation sets and are normalized by dividing by the MSE of the OLS solution; that is to say, when any method attains a MSE greater than 1 their prediction power is estimated to be worse than that of the OLS, while for smaller values the accuracy has improved.

The sparsity of the solution of the unconstrained lasso increases as its regularization parameter $\lambda \in \mathbb{R}^+$ does. The critical values of λ are easily computed using any implementation of the LAR algorithm in various standard statistical packages. In particular, in this paper the lasso set of solutions was obtained by using the `lars()` function of `R-cran` package `lars` (Hastie and Efron, 2013).

For the tightened MIQPs (6), the pairwise correlation considered to generate the sets Ω_α^+ and Ω_α^- in constraints (9)-(12) is fixed to $\alpha = 0.6$. Following Bertsimas and King (2015), the maximum pairwise correlation allowed is $\eta = 0.8$; that is to say, the set Ω_η in (7) is defined here as $\Omega_\eta = \{(i,j) : |\rho_{ij}| \geq 0.8\}$.

The significance parameter ϵ in constraints (13) is tuned by chosing amongst the ten values $\{0.05, 0.06, 0.08, 0.1, 0.125, 0.15, 0.175, 0.2, 0.25, 0.3\}$ so as to minimize the MSE in the test set. The parameter V_T controlling the sparsity is chosen sequentially in $\{1, \ldots, N\}$. However, in order to restrict the search only to likely values of V_T, a stopping criterion is imposed: when no more features are added to the model (i.e., when the constraint (5) becomes inactive), no larger values of V_T are considered. On top of this, to further improve the speed of the tightened procedures, we have restricted the size of the parameters grids for large instances, as recommended in Tibshirani et al. (2005). In particular, for large simulated datasets we have required our output to have a maximum of 25% of non-zeroes over N, the number of predictors.

3.2 Real datasets

In this section we show the results obtained for some real datasets, which are easily reachable on internet and well referenced in the literature (Bertsimas and King, 2015). Further details about the data sets and their sources are displayed in Table 1. The columns provide information about the name, number of observations (K), the number of covariates (N), and data source.

Table 1: *Real data sets specifications and sources.*

	K	N	Source
`cpu`	105	6	Lichman (2016)
`yacht`	154	6	Lichman (2016)
`whitewine`	2499	11	Lichman (2016)
`redwine`	800	11	Lichman (2016)
`golf2008`	78	6	Winner (2016)
`golf2009`	73	11	Winner (2016)
`compact`	4096	21	Torgo (2016)

The median MSE and number of non-zeroes attained by the different estimation procedures are displayed in Tables 2-4, where the first column of results corresponds to the normalized MSE and number of non-zero coefficients (NZ), and each row shows the results of a real dataset. To obtain such results, the databases were randomly divided ten times in training, test and validation sets.

3.2.1 The effect of the sign coherence constraints

The sign coherence constraints described in Section 2.1 and formulated as (9)-(12) are claimed to avoid the inconsistencies shown by Figure 1. Now we analyse the effect of such constraints in the accuracy and sparsity of the obtained solutions. In addition, the results are compared with those under the correlation constraint (7) by Bertsimas and King (2015). The first three rows of Table 2 show the results for the untightened OLS, and the OLS with tightening sets $\mathcal{S}_1$ (the novel *sign coherence constraint*) and $\mathcal{S}_2$ (the *correlation constraint* of Bertsimas and King (2015)), respectively. Analogously, the remaining rows display the performance of the untightened and tightened lasso problems.

Table 2: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline methods and the approaches tightened by the correlation-based constraints.*

	Cpu		**Yacht**		**Whitewine**		**Redwine**		**Golf2008**		**Golf2009**		**Compact**	
	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ
OLS	1.000	6	1.000	6	1.000	11	1.000	11	1.000	6	1.000	11	1.000	21
$\mathcal{S}_1$	1.031	5	0.970	4	1.009	9	0.999	10	1.004	5	0.896	7.5	1.014	15
$\mathcal{S}_2$	1.042	5	0.996	5	1.007	10	1.000	10	0.998	5	0.971	8	1.016	15
lasso	1.000	5	0.959	2.5	1.000	10.5	0.997	9.5	1.000	4	0.798	9.5	1.000	20.5
$\mathcal{S}_1$	1.036	4	0.954	3.5	1.009	9	0.996	10	1	5	0.79	8	1.010	14
$\mathcal{S}_2$	1.049	4.5	0.961	4	1.008	10	0.994	8.5	0.986	4	0.966	7	1.013	14

From Table 2 it can be deduced that both constraints related to multicollinearity yield a similar performance: adding sign coherence constraints slightly improves sparsity for `yacht` and `whitewine` databases, but may also attain slightly more dense solu-

tions (golf2008 and the tightened lasso in redwine database). Regarding the predictive quality, the MSEs are quite similar in most cases. An exception is the golf2009 database, where the tightening set $\mathcal{S}_1$ improves the predictive power of the tightening set $\mathcal{S}_2$ in a 7.5% for the OLS and a 17.6% for the lasso. We conclude that, not only the sign coherence constraint improves the interpretability of the results by avoiding the inconsistencies described in Section 1, but also it does not damage the level of sparsity and predictive power. Indeed, when comparing the novel coherence constraints (9)-(12) against the correlation constraint (7), they give overall the same accuracy and sparsity. We should remark that, in addition, our constraints yield more stable results than the correlation constraint when used in the lasso model. To illustrate this, consider Figure 4, which displays the paths of solutions attained for the baseline and tightened lasso in random shuffles of the golf2009 and yacht databases. The path of solutions attained by the baseline lasso on the golf2009 database (top left panel) shows that coefficients β_{11} and β_6 grow quickly in opposite directions for small values of λ. These coefficients are inflated due to the high pairwise correlation (0.91) between these variables. This phenomenon disappears when we strictly forbid coefficients β_{11} and β_6 to be simultaneously non-zero (central left panel). However, coefficients β_3, β_4 and β_6, which are considerably less correlated ($\rho_{3,4} = 0.05$, $\rho_{3,6} = 0.77$ and $\rho_{4,6} = -0.39$), still show this behaviour. Moreover, coefficient β_{10}, which was significant even for large values of the penalty λ, suddenly disappears as λ approaches zero. On the contrary, sign coherence constraints (bottom left panel) seem to avoid the inflation of coefficients β_3, β_4, β_6 and β_{11}, also leading to smoother paths. A similar behaviour is observed for yacht database, where coefficients β_2, β_3, β_4 and β_5 explode for $\lambda = 0$ in the baseline lasso (right top panel). Since the pairwise correlations between these variables do not exceed the threshold 0.8 imposed by Bertsimas and King (2015) (indeed, the largest correlation coefficient is $\rho_{3,5} = 0.63$), they are not explicitly forbidden simultaneously in the model, thus yielding the same solution path as the lasso when adding the correlation constraint (central right panel). As sign coherence constraints are less restrictive, they allow highly correlated variables to simultaneously appear in the model. This may lead to alternative solutions that may improve the stability of the estimated parameters $\boldsymbol{\beta}$ in the presence of highly correlated variables. In fact, the bottom right panel represents a considerably stable path of solutions along λ, which clearly identify the more significant feature for prediction.

As mentioned in the introduction, our constraints differ from most of the approaches previously considered in the literature: we do no explicitly forbid two highly correlated variables in the model (as recommended in Chatterjee and Hadi (2015); Bertsimas and King (2015)), nor encourage groups of correlated variables to be altogether in or out of the model (Yu and Liu, 2016; Kim and Xing, 2009). An example of this is illustrated on Figure 5, representing the estimated $\boldsymbol{\beta}$ in the 10 shuffles of golf2009 database. The $\boldsymbol{\beta}$ have been estimated by the classic OLS (first row), the lasso (second row), and two tightened OLS approaches: adding the correlation constraint (7) of Bertsimas and King (2015) (tightening set $\mathcal{S}_2$) or adding the sign coherence constraints (9)-(12) (tightening

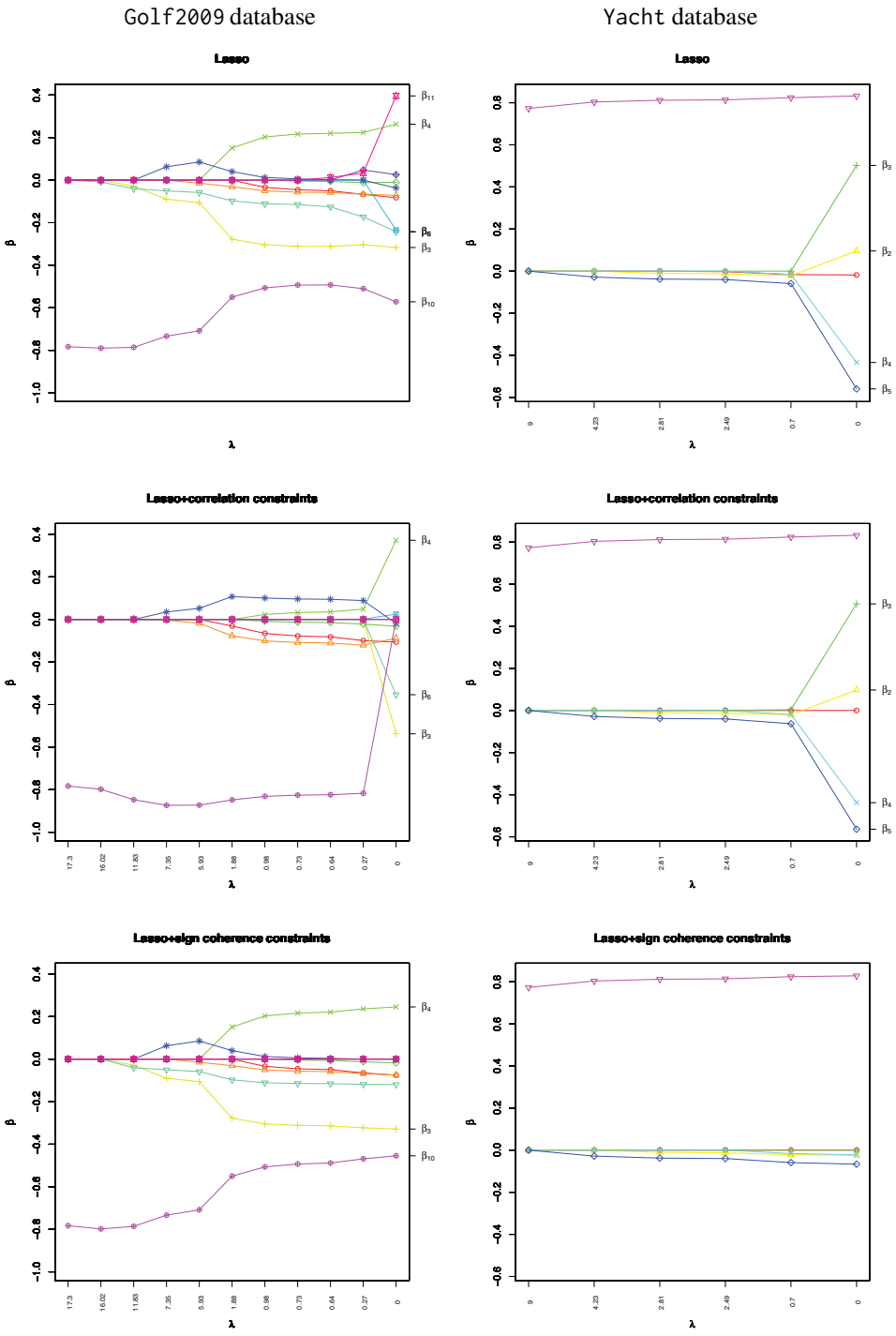

Figure 4: *Paths of solutions of the lasso (top panels), the lasso with the correlation constraint (mid panels) and sign coherence constraints (bottom panels) in shuffle 9 of the* `Golf2009` *database (left panels) and shuffle 7 of the* `Yacht` *database (right panels).*

set $\mathcal{S}_1$). In this heatmap, features X^6 and X^{10}, with correlation 0.91, do not appear simultaneously in the same shuffle when adding the correlation constraints. Nonetheless, their coefficients are non-zero simultaneously in various shuffles when considering sign coherence constraints.

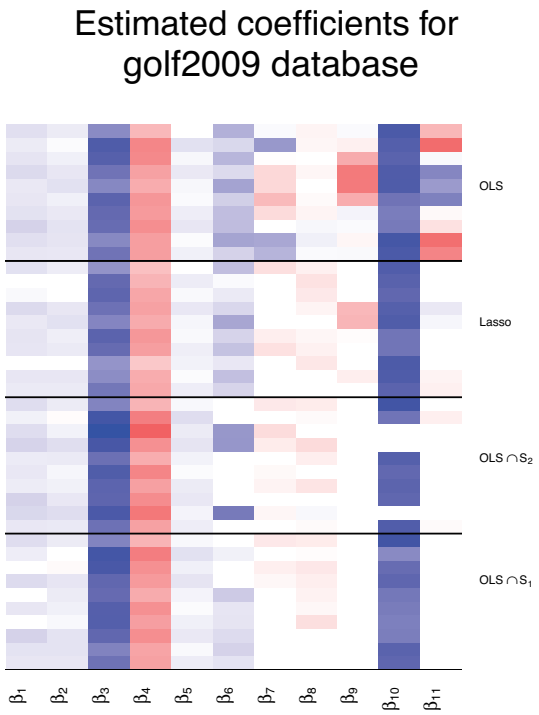

Figure 5: *Heatmap representing the coefficients β_i estimated by OLS, lasso and OLS tightened by adding correlation constraints (tightening set $\mathcal{S}_2$), or sign coherence constraints (tightening set $\mathcal{S}_1$) for* `golf2009` *database.*

The above results were obtained for fixed values of η and α, which were set to 0.6 and 0.8, respectively. These numerical examples followed the experimental design in Bertsimas and King (2015), who fixed the correlation threshold α. However, in Appendix B we explore the sensitivity of the tightening procedures to changes in η and α. We conclude that, in general, the calibration of these parameters seem to yield less sparse solutions with a similar MSE. As a consequence the results shown in our numerical experiments disregard the calibration of the correlation thresholds.

3.2.2 The effect of the significance constraint

Now we aim to study the impact of adding significance constraints (13) to the OLS and the lasso. This was briefly analysed in Section 2.2, where heat maps representing the estimated coefficients $\boldsymbol{\beta}$ were represented in Figure 3. We observed that imposing a threshold ϵ to the estimates may lead to more sparse solutions by avoiding spurious

coefficients and discarding unimportant variables. The first two rows of Table 3 display the results for the untightened OLS and its counterpart tightened with $\mathcal{S}_3$ (*significance constraint*). Analogously, the last two rows display the results for the unrestricted and tightened lasso. In this table, we observe that tightening the feasible region of the OLS and the lasso by using the set $\mathcal{S}_3$ always improves the sparsity of the output while usually attaining a competitive predictive quality. For instance, the significance constraints improves the MSE of the OLS and the lasso in a 4% and a 2.6% for `yacht` database, respectively, while reducing in 3.5 and 1.5 the number of non-zeroes. However, an exception is found in `golf2009` dataset, where the novel constraint worsens the accuracy of the OLS and the lasso, although yielding 3 more zeroes in both cases.

Table 3: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline methods and the approach tightened by the significance constraint.*

	Cpu		**Yacht**		**Whitewine**		**Redwine**		**Golf2008**		**Golf2009**		**Compact**	
	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ
OLS	1.000	6	1.000	6	1.000	11	1.000	11	1.000	6	1.000	11	1.000	21
$\mathcal{S}_3$	0.978	4.5	0.960	2.5	0.999	8	1.000	7.5	1.007	5	1.205	8	1.002	15
lasso	1.000	5	0.959	2.5	1.000	10.5	0.997	9.5	1.000	4	0.798	9.5	1.000	20.5
$\mathcal{S}_3$	0.989	4	0.934	1	0.999	8	1.001	7	0.987	3	0.936	6.5	1.003	13

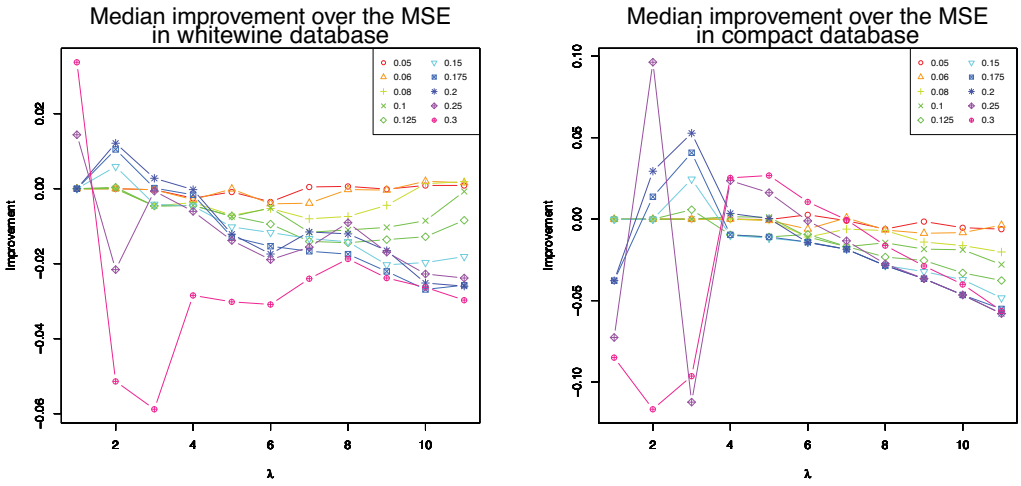

Figure 6: *Improvements over the MSE of the lasso when tightened via the significance constraint.*

Since it is not straightforward to choose a grid of thresholds ϵ to calibrate from, we will try to gain some intuition by studying the improvement of the predictive quality when adding the significance constraints for each value of ϵ considered. To this aim, Figure 6 shows the median improvement on the lasso MSE for each value of the penalty λ and each threshold in the proposed grid for `whitewine` (left panel) and `compact` (right panel) database.

From Figure 6 we observe that the behaviour for the largest ϵ $(0.3, 0.25)$ can be more effective specially for the larges values of λ. That is to say, a large ϵ may help combating the strong shrinkage of the lasso when highly sparse solutions are sought. Still large but more conservative values of the threshold ($\epsilon = 0.2, 0.175, 0.15$) seem to also improve the MSE of the lasso with large λ, also providing less extreme behaviours than the choices $\epsilon = 0.3, 0.25$. Finally, the smallest values of the threshold ($\epsilon = 0.06, 0.05$) may slightly improve the lasso with small λ and the OLS. In conclusion, there is no straightforward a priori choice for the parameter ϵ, which should be calibrated. Nevertheless, if the user is seeking a highly sparse solution (i.e., the user is choosing a high penalty λ) it seems advisable to choose larger values for ϵ in order to combat the shrinkage more effectively. On the other hand, when estimating via OLS (or lasso with small values of λ) the user might want to focus on a grid with a majority of small values of ϵ.

3.2.3 Global performance

Finally, we show in Table 4 the results when our novel constraints are jointly considered in combination to the cardinality constraint, or equivalently, when the set $\mathcal{S}_4$ (*cardinality + sign coherence + significance constraints*) is used to tighten the OLS or lasso approaches. For comparison reasons, the table also shows the results under the tightening set $\mathcal{S}_5$ (*cardinality + correlation constraints*) proposed in Bertsimas and King (2015). Table 4 also displays the predictive quality and number of non-zero coefficients for the Enet, the SRIG method in Yu and Liu (2016) and the GFlasso method in Kim and Xing (2009).

Table 4: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline estimation methods (OLS, lasso, SRIG, Enet and GFlasso), and OLS and lasso tightened with $\mathcal{S}_4$ (cardinality + novel constraints) or $\mathcal{S}_5$ (cardinality + correlation constraint), for the real datasets.*

Method	**Cpu**		**Yacht**		**Whitewine**		**Redwine**		**Golf2008**		**Golf2009**		**Compact**	
	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ
OLS	1.000	6	1.000	6	1.000	11	1.000	11	1.000	6	1.000	11	1.000	21
$\mathcal{S}_4$	0.990	4	0.934	1	1.011	6	1.002	6	0.999	3	0.963	4	1.012	9
$\mathcal{S}_5$	1.002	4	0.934	1	1.008	8	1.000	6	1.006	3	1.024	5	1.016	11
lasso	1.000	5	0.959	2.5	1.000	10.5	0.997	9.5	1.000	4	0.798	9.5	1.000	20.5
$\mathcal{S}_4$	0.993	4	0.934	1	1.013	6	0.998	6	0.982	3	0.971	4.5	1.013	9
$\mathcal{S}_5$	1.049	4	0.934	1	1.008	7.5	0.995	6	0.988	3	1.045	4	1.015	11.5
SRIG	0.988	6	0.942	2	1.000	11	0.999	11	0.983	6	1.016	11	0.999	21
Enet	0.917	5.5	0.948	2	1.000	11	0.998	11	0.994	4	0.805	9.5	1.000	21
GFlasso	0.972	6	0.959	2.5	1.000	11	0.996	11	1.000	4	0.866	9.5	1.001	20

First, we analyse the performance of the baseline estimation procedures (OLS and lasso) against their tightened counterparts. We can conclude that reducing the search space of the coefficients $\boldsymbol{\beta}$ by intersecting with either $\mathcal{S}_4$ or $\mathcal{S}_5$ always improves the sparsity of the solutions. Moreover, the predictive quality is usually similar to that of the OLS and the lasso. In particular, for `yacht` database, both tightened approaches

improve the MSE of the OLS and the lasso estimates by a 6.6% and a 2.6%, respectively. However, these tightened procedures worsens the accuracy of the lasso in the `golf2009` database.

Second, we compare the performance of the two tightening sets. From the table it can be observed that they deliver a similar accuracy-sparsity trade-off for 3 out of the 7 real datasets (`yacht, redwine, golf2008`). For the remaining databases, the approaches attain different trade-offs between sparsity and predictive quality. Indeed, the novel set $\mathcal{S}_4$ maintains the sparsity attained by the set $\mathcal{S}_5$ og Bertsimas and King (2015) on the `cpu` database, while slightly improving the MSE in a 1.2% and a 5.3% for the tightened OLS and lasso, respectively. In contrast, the proposed tightening set provides more sparse solutions, with a similar predictive quality for `whitewine` and `compact` datasets. Finally, $\mathcal{S}_4$ improves the MSE of the OLS and the lasso in a 6% and 7.1% in `golf2009` database, but it provides a slightly more dense solution for the lasso. Note that the best accuracy for the lasso in this database was attained when adding exclusively sign coherence constraints (see Table 2), although the solution provided was more dense. Third, both the SRIG and GFlasso are clearly outperformed by the tightened approaches. On top of this, for the real datasets considered here there is no guarantee that the necessary assumptions to preserve the theoretical properties of the SRIG are fulfilled. In order to compare the performance of our approach against these methods under a more favourable scenario for the later, in the next section we replicate the simulation study of Yu and Liu (2016), hence assuring the non-violation of the conditions for the SRIG method.

3.3 Simulations

In the previous section, the behaviour of the novel constraints for the case of real datasets with a small number of predictors was analysed. In this section we aim to examine the sensitivity of the tightened procedures under various settings. First, we will simulate data as in Yu and Liu (2016) to understand the behaviour of the methodology under different correlation structures and for different sizes of the training sample. Second, we aim to test the proposed methodology for larger datasets and for different correlations intensities. To do so we simulate data as in Bertsimas and King (2015).

3.3.1 Sensitivity to correlation structure and training sample size

In this section we aim to test the proposed methodology for datasets simulated following the three examples and training sizes described in Yu and Liu (2016). Ten instances have been generated for each example and for training and testing sizes of 40, 80 and 120. For both Examples 1 and 2, $\beta_i = 3$ for $i = 1,..,15$, and $\beta_i = 0$ for $i = 16,\ldots,100$. In Example 1, however, the predictors were generated as follows:

$$X^i = Z^j + 0.4\mathbf{a}_X^i, \quad Z^j \sim N(0,1), \quad \mathbf{a}_X^i \sim N(0,1)$$

for $5(j-1)+1 \leq i \leq 5j$ and $j = 1,2,3$. For $i > 15$, $X^i \sim N(0,1)$. In Example 2, the k-th vector of observations $\left(X_k^1, \ldots, X_k^N\right)$ was generated following a multivariate normal distribution with zero mean and covariance matrix $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = 0.5^{|i-j|}$. Analogously, in Example 3 the vector of observations was also drawn from a multivariate normal distribution with zero mean but with covariance matrix $\Sigma = (B + \psi I)^{-1}$, where $b_{ij} = 0$ for $i = j$, and $b_{ij} = 0.5\delta_{ij}$, $\delta_{ij} \sim Be(0.05)$ otherwise. Parameter ψ is fixed so that the condition number of Σ^{-1} equals N. The real coefficients are $\boldsymbol{\beta} = \Sigma^{-1}\Sigma_{xy}$, where Σ_{xy} is the cross-covariance vector whose elements equal 10 for the four predictors with the largest degrees and 0 otherwise. These three examples have different structures of correlation and, as it will be seen later, this may influence on the performance of the approaches making use of constraints taking into account pairwise correlations. As an illustration, Figure 8 of the Appendix C of the Supplementary Material displays the pairs of variables appearing in the correlation constraints of Bertsimas and King (2015) and our sign coherence constraints.

As done in Bertsimas and King (2015), the grid of values of the parameter λ to be tuned for the tightened MIQP with lasso objective function is logarithmically generated in the interval $(0, \lambda_{max}]$, where λ_{max} is the penalty provided by the `lars` for which only one coefficient is non-zero. Analogously to the real datasets results, Table 5 shows the median MSE and number of non-zeroes (NZ). In particular, for each example (rows) and training sizes (columns), each row shows the results obtained for an estimation method, namely the OLS, the lasso, and their tightened counterparts, which take the form of Problem (6) with tightening sets $\mathcal{S}_4$ (*cardinality + sign coherence + significance constraints*) and $\mathcal{S}_5$ (*cardinality + correlation constraints*). The last rows for each example display the results for the methods dealing with correlated variables: the Enet, the SRIG, and the GFlasso. To make it easier to discuss these results, in Figure 7 we have represented the MSE against the number of nonzeroes for the three simulated examples of Table 5. The different estimation methods have been assigned different colours, the solid items representing the approaches making use of our proposed tightening set. The diversity of training samples have been represented by unalike symbols.

In the top panel of Figure 7 we can observe that, in Example 1, the MIQPs (i.e., the approaches proposed in Bertsimas and King (2015) and in this paper) tend to attain more sparse solutions than the continuous optimization methods (i.e., the lasso, SRIG and GFlasso). As it can be observed in Figure 8 in the Appendix, the true generating model contains predictors that hold a high pairwise correlation and, therefore, are forbidden to appear simultaneously in the outputs yielded by the tightening set of Bertsimas and King (2015). As a consequence, the approaches making use of this set $\mathcal{S}_5$ are not able to recover the original graph, hence delivering more sparse solutions but with worse predictive quality than our tightening set. In constrast, the sign coherence constraints flexibility allows the simultaneous presence of all the variables in the generating model. As a consequence, $\mathcal{S}_4$ yields outputs with similar performance to that of the SRIG, that is the best approach amongst those based on continuous optimization models.

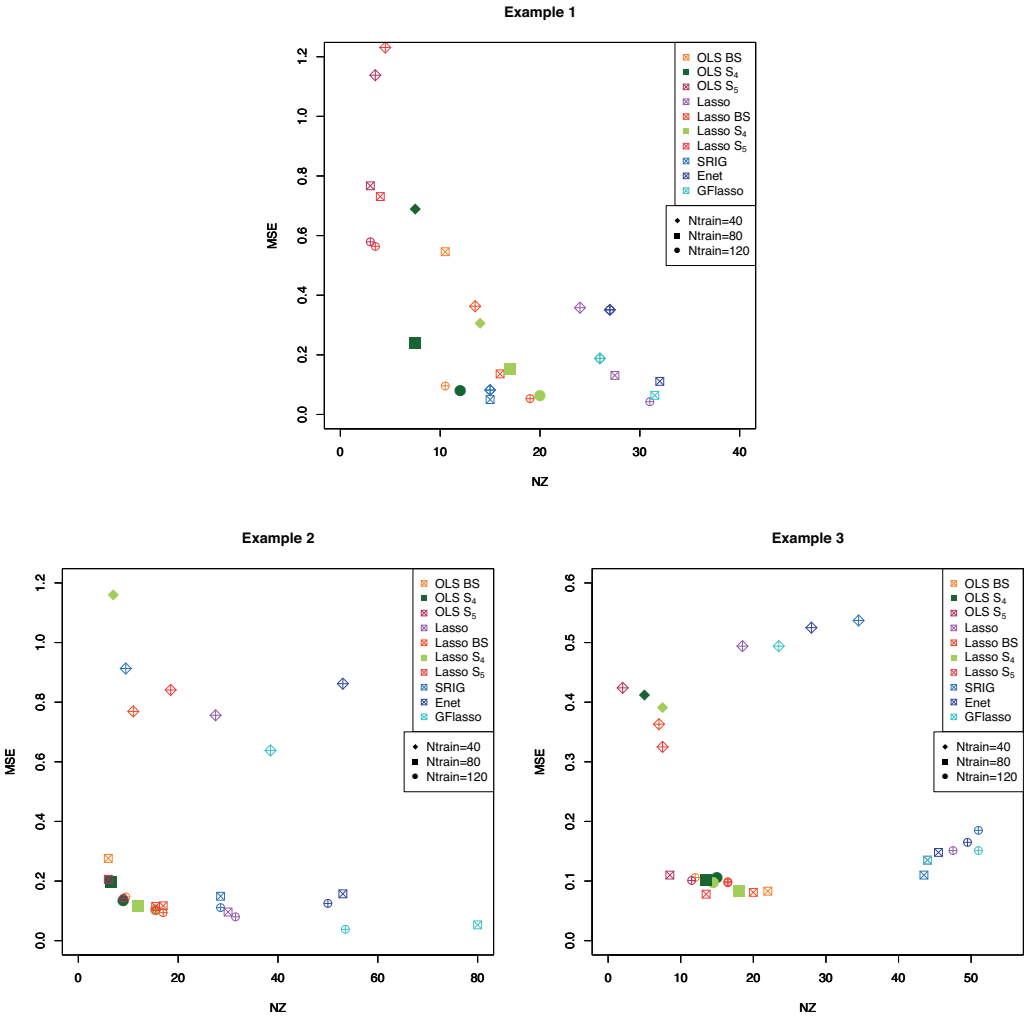

Figure 7: *Median MSE and NZ for each method in the simulated datasets of Yu and Liu (2016).*

In the bottom panels of Figure 7 we observe that the MIQP approaches are clustered together. For both Examples 2 and 3 it is evident that the methods based on integer optimization improve the sparsity of the outputs with no damage to the predictive quality. More specifically, in Example 3 all the methods based on continuous optimization conform a unique cluster with higher density and similar or slightly worse MSE than the MIQP approaches. Nonetheless, in Example 2 the SRIG and the lasso attain more sparse solutions than the rest of continuous optimization approaches, yet still yielding more non-zeroes than the tightened procedures. In this case, all the approaches attain a similar accuracy.

Table 5: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline estimation methods (OLS, lasso, SRIG, Enet and GFlasso), and OLS and lasso tightened with $\mathcal{S}_4$ (cardinality + novel constraints) or $\mathcal{S}_5$ (cardinality + correlation constraint), for the simulated datasets with $N = 100$.*

		Ntrain= 40		Ntrain= 80		Ntrain= 120	
		MSE	NZ	MSE	NZ	MSE	NZ
Example 1	OLS	1.000	100.0	1.000	100.0	1.000	100.0
	$\mathcal{S}_4$	0.689	7.5	0.240	7.5	0.081	12.0
	$\mathcal{S}_5$	1.138	3.5	0.767	3.0	0.579	3.0
	lasso	0.358	24.0	0.131	27.5	0.043	31.0
	$\mathcal{S}_4$	0.306	14.0	0.154	17.0	0.062	20.0
	$\mathcal{S}_5$	1.231	4.5	0.731	4.0	0.564	3.5
	SRIG	0.082	15.0	0.050	15.0	0.082	15.0
	Enet	0.351	27.0	0.111	32.0	0.351	27.0
	GFlasso	0.188	26.0	0.064	31.5	0.188	26.0
Example 2	OLS	1.000	100.0	1.000	100.0	1.000	100.0
	$\mathcal{S}_4$	1.380	5.0	0.197	6.5	0.133	9.0
	$\mathcal{S}_5$	1.560	4.0	0.206	6.0	0.142	9.0
	lasso	0.756	27.5	0.096	30.0	0.080	31.5
	$\mathcal{S}_4$	1.160	7.0	0.115	12.0	0.105	15.5
	$\mathcal{S}_5$	0.841	18.5	0.115	15.5	0.102	15.5
	SRIG	0.913	9.5	0.149	28.5	0.111	28.5
	Enet	0.862	53.0	0.157	53.0	0.125	50.0
	GFlasso	0.638	38.5	0.053	80.0	0.038	53.5
Example 3	OLS	1.000	100.0	1.000	100.0	1.000	100.0
	$\mathcal{S}_4$	0.412	5.0	0.101	13.5	0.106	15.0
	$\mathcal{S}_5$	0.424	2.0	0.110	8.5	0.101	11.5
	lasso	0.494	18.5	0.135	44.0	0.151	47.5
	$\mathcal{S}_4$	0.391	7.5	0.083	18.0	0.097	14.5
	$\mathcal{S}_5$	0.325	7.5	0.078	13.5	0.097	16.5
	SRIG	0.537	34.5	0.110	43.5	0.185	51.0
	Enet	0.525	28.0	0.148	45.5	0.165	49.5
	GFlasso	0.494	23.5	0.135	44.0	0.151	51.0

3.3.2 Scalability and sensitivity to correlation intensity

In this section we aim to test the proposed methodology for larger datasets with diverse correlation intensities. As the overall design of experiments, the synthetic generation of the data is similar to that of Bertsimas and King (2015). The k-th vector of observations $\left(X_k^1,\ldots,X_k^N\right)$ was generated following a multivariate normal distribution with zero mean and covariance matrix $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = \rho^{|i-j|}$. In particular, we chose $\rho = -0.9$ and $\rho = -0.5$ so as to test the performance of the constraints under highly and moderate correlations. The regression model is taken in small dimension, but with quite a number of irrelevant covariates. More precisely, the number of features was set to 500, only 10 of

which corresponding to explanatory variables, the remaining 490 being noise. For each value of ρ, 10 instances were generated as follows. The β_i were uniformly generated in the interval $(-2,2)$ for i such that i modulus $N/10 = 0$. The response was generated following (1), with $\beta_0 = 0$ and the error terms i.i.d. following a normal distribution with zero mean and variance as in Bertsimas and King (2015). As done in Section 3.3.1, the sequence of λ has been logarithmically generated. Table 6 shows the median MSE and number of non-zeroes (NZ) of the solutions for the OLS, the lasso and their tightened counterparts, as well as the SRIG, Enet and GFlasso.

Table 6: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline estimation methods (OLS, lasso, SRIG, Enet and GFlasso), and OLS and lasso tightened with $\mathcal{S}_4$ (cardinality + novel constraints) or $\mathcal{S}_5$ (cardinality + correlation constraint), for the simulated datasets with $N = 500$.*

	$\rho = -0.5$		$\rho = -0.9$	
	MSE	NZ	MSE	NZ
OLS	1.000	500	1.000	500
$\mathcal{S}_4$	0.745	6	0.732	6
$\mathcal{S}_5$	0.849	6	1.474	4
lasso	0.535	80.5	0.533	183
$\mathcal{S}_4$	0.526	10.5	0.528	11
$\mathcal{S}_5$	0.501	13	0.807	19
SRIG	0.630	34	0.792	102
Enet	0.535	50	0.534	51.5
GFlasso	0.537	70.5	0.533	54.5

Note that, for the simulated data with moderately correlated features ($\rho = -0.5$), the correlation constraints (7) and the sign coherence constraints (9)-(12) are inactive, since the highest pairwise correlation in absolute value is roughly 0.5. In this case, both sets of tightening constraints help to considerably improve the sparsity of the baseline estimation procedures. Indeed, the density of the OLS is drastically reduced by increasing in 494 the zeroes of the output, while its predictive quality is also substantially improved. However, the novel tightening set $\mathcal{S}_4$ provides an accuracy 12.2% better than the attained with the set $\mathcal{S}_5$ proposed in Bertsimas and King (2015). On the other hand, the benchmark sparse regression method, the lasso, attains a median of 80.5 non-zeroes, while its tightening counterparts produce much more sparse solutions with better MSEs. Although the predictive quality of the outputs of the tightened procedure with $\mathcal{S}_5$ is slightly better than that obtained with $\mathcal{S}_4$, this comes at the price of yielding more dense solutions. More generally, the methods based on MIQP solvers clearly outperform the approaches relying on continuous optimization techniques: the former manage to considerably reduce the sparsity of the later with a slightly better predictive quality. Amongst the later approaches, the GFlasso is outperformed by the Enet which, with a similar accuracy, provides a much more sparse solution. Nonetheless, the most sparse of these methods is the SRIG, although yielding the worst predictive quality.

Table 7: *Predictive quality (MSE) and sparsity degree (NZ) for the baseline estimation methods (OLS, lasso, SRIG, Enet and GFlasso), and OLS and lasso tightened with $\mathcal{S}_4$ (cardinality + novel constraints) or $\mathcal{S}_5$ (cardinality + correlation constraint), for the simulated datasets with $\rho = -0.9$ and $N = 50, 500, 1000$.*

	$N = 50$		$N = 500$		$N = 1000$	
	MSE	NZ	MSE	NZ	MSE	NZ
OLS	1.000	50	1.000	500	1.000	1000
$\mathcal{S}_4$	0.662	8	0.732	6	1.035	5
$\mathcal{S}_5$	0.634	8	1.474	4	1.310	5.5
lasso	0.671	24.5	0.533	183	0.555	56
$\mathcal{S}_4$	0.637	8.5	0.528	11	0.904	8
$\mathcal{S}_5$	0.595	19	0.807	19	0.406	10
SRIG	0.127	42	0.792	102	0.835	103.5
Enet	0.108	26	0.534	51.5	0.538	49.5
GFlasso	0.693	41	0.533	54.5	0.615	359

The simulated instances with highly correlated features, $\rho = -0.9$, show a similar behaviour. Nevertheless, the lasso provides significantly more dense outputs in this case, with a median of 183 non-zeroes. The tightening set $\mathcal{S}_5$ proposed in Bertsimas and King (2015) worsens its predictive quality although improving substantially its sparsity. In contrast, the novel tightening set $\mathcal{S}_4$ attains around eight more zeroes than the later while maintaining the lasso's accuracy. Regarding the OLS, both tightened estimation methods reduce drastically the density of the solutions, although the $\mathcal{S}_4$ obtains around two more non-zero coefficients than the tightening set $\mathcal{S}_5$. However, the price paid for two extra non-zeroes is shown in the accuracy of the solution: while adding cardinality, sign coherence and significance constraints improves the MSE of the OLS in a 26.8%, adding cardinality and correlation constraints instead worsens the accuracy in 47.4%. Analysing the results for the continuous optimization based methods, we observe that the best performance is yielded by the Enet and GFlasso which, with a similar MSE, significantly enhance the sparsity of the classic lasso. Nonetheless, when tightened with $\mathcal{S}_4$, the later attains outputs with comparable predictive quality and many more zero coefficients (around 40 more).

In order to analyse the scalability of our methodology we have also simulated data as in Bertsimas and King (2015) with 50 and 1000 variables, and a maximum correlation of 0.9. For the later, the MIQP to be solved would have 3001 variables and more than 2000 constraints. As the size of the problem is considerably larger, we have allowed for a time limit of 40 seconds in this case, which was also the time limit considered for high dimensional data where $N > K$. The results are collected in Table 7. As it can be observed, MIQP approaches attain solutions that are considerably more sparse than their continuous counterparts while still delivering a good predictive quality. In particular, SRIG delivers the most dense outputs from the sparse continuous methods, while the GFlasso is outperformed both in terms of accuracy and sparsity. Regarding the MIQP,

they attain different accuracy-sparsity trade-offs. Although the proposed constraints attain better accuracy when combined with the OLS objective, this is not true when the regularization penalty λ is positive. Nonetheless, in this case the tightening set $\mathcal{S}_4$ yields more sparse outputs than $\mathcal{S}_5$.

Summarizing, tightening the search space of the OLS and the lasso provides considerably more sparse solutions, although the sign coherence and significance constraints yield a better accuracy-sparsity trade-off than the correlation constraint. Indeed, the later can substantially worsen the predictive quality of the baseline methods in order to reduce the density of the outputs, while the former entails a more competitive MSE.

4 Concluding remarks

The aim of this paper is to enhance the interpretability in a regression model without worsening its predictive quality. We assume we have a baseline regression estimation procedure based on solving an optimization problem (e.g. OLS or lasso), and then the underlying optimization problems are modified by adding new constraints to those defining the search space. These constraints avoid misleading estimators that may be obtained in the presence of highly correlated variables and detect the most important features for the prediction.

In order to assess the impact of adding the two novel constraints over various estimation procedures, in our numerical experiments we consider the OLS and the lasso. The search space of $\boldsymbol{\beta}$ is reduced in these methods by using the tightening set composed by the sign coherence constraints (9)-(12) and/or the significance constraints (14), possibly in combination with the cardinality constraint (5). The first constraint forces the sign of the coefficients to be coherent with the sign of large and moderately large pairwise correlations between features, while the second avoids spurious coefficients and combats the shrinkage of regularized regression. We compare the performance of our tightening set, including all the proposed constraints, with the recent tightening set by Bertsimas and King (2015), which also defines a MIQP and includes the cardinality constraint (5) and the correlation constraint (7) which explicitly forbids two highly correlated variables to be simultaneously in the regression model. These methods are compared against other approaches also dealing with correlated variables but based on continuous optimization techniques: the Enet (Zou and Hastie, 2005), the SRIG (Yu and Liu, 2016), and the GFlasso (Kim and Xing, 2009). The results show that the novel constraints yield tractable optimization problems, solvable in short time by standard solvers, and may enhance the interpretability while often improving or maintaining the predictive quality and level of sparsity. More specifically, the MIQP approaches attain a different trade off between sparsity and predictive quality than the methods based on continuous optimization, usually yielding more sparse solution with similar or better MSE. Amongst the former methods, the novel constraints tend to improve the predictive quality of the outputs obtained with the tightening set proposed in Bertsimas and King (2015).

Although we have proposed some heuristics to further improve the speed of the tightened procedures, such as reducing the grid of parameters for large datasets, in the future we aim to develop tailored heuristics that improve the computational times of the MIQPs when they particularly model linear regression problems.

References

Atamurk, A., Nemhauser, G. and Savelsbergh, M. (2000). Conflict graphs in solving integer programming problems. *European Journal of Operational Research*, 121, 40–55.

Bartholomew, D. J., Steele, F., Moustaki, I. and Galbraith, J. (2008). *Analysis of Multivariate Social Science Data*. Chapman & Hall.

Bertsimas, D. and King, A. (2015). OR forum – An algorithmic approach to linear regression. *Operations Research*, 64, 2–16.

Bertsimas, D., King, A., Mazumder, R. (2016). Best subset selection via a modern optimization lens. *The Annals of Statistics*, 44, 813–852.

Breiman, L. (1995). Better subset regression using the nonnegative garrote. *Technometrics*, 37, 373–384.

Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data*. Springer.

Cai, A., Tsay, R. and Chen, R. (2009). Variable selection in linear regression with many predictors. *Journal of Computational and Graphical Statistics*, 18, 573–591.

Camm, J. D., Raturi, A. S. and Tsubakitani, S. (1990). Cutting big *M* down to size. *Interfaces*, 20, 61–66.

Cao, G., Guo, Y. and Bouman, C. A. (2010). High dimensional regression using the sparse matrix transform (SMT). In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 1870–1873. IEEE.

Carrizosa, E. and Guerrero, V. (2014). Biobjective sparse principal component analysis. *Journal of Multivariate Analysis*, 132, 151–159.

Carrizosa, E., Nogales-Gómez, A. and Morales, D. R. (2016). Strongly agree or strongly disagree?: Rating features in support vector machines. *Information Sciences*, 329, 256–273.

Carrizosa, E., Nogales-Gómez, A. and Morales, D. R. (2017). Clustering categories in support vector machines. *Omega*, 66, 28–37.

Carrizosa, E., Olivares-Nadal, A. V. and Ramírez-Cobo, P. (2016). A sparsity-controlled vector autoregressive model. *Biostatistics*, 18, 244–259.

Chatterjee, S. and Hadi, A. S. (2015). *Regression Analysis by Example*. John Wiley & Sons.

Danna, E., Rothberg, E. and Le Pape, C. (2005). Exploring relaxation induced neighborhoods to improve mip solutions. *Mathematical Programming*, 102, 71–91.

Efron, B. and Hastie, T. (2003). LARS software for R and Splus. https://web.stanford.edu/ hastie/ Papers/LARS/.

Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, 32, 407–499.

Farrar, D. E. and Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, 92–107.

Fischetti, M. and Lodi, A. (2005). Local branching. *Mathematical Programming*, 98, 23–47.

Fourer, R., Gay, D. and Kernighan, B. W. (2002). *The AMPL book*. Duxbury Press, Pacific Grove.

Friedman, J., Hastie, T. and Tibshirani, R. (2001). *The Elements of Statistical Learning*, Volume 1. Springer series in statistics.

Hastie, T. and Efron, B. (2013). Least Angle Regression, Lasso and Forward Stagewise. http://cran.r-project.org/web/packages/lars/lars.pdf.

Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: the Lasso and Generalizations*. CRC Press.

Hesterberg, T., Choi, N. H., Meier, L. and Fraley, C. (2008). Least angle and ℓ_1 penalized regression: A review. *Statistics Surveys*, 2, 61–93.

Jou, Y.-J., Huang, C.-C. L. and Cho, H.-J. (2014). A VIF-based optimization model to alleviate collinearity problems in multiple linear regression. *Computational Statistics*, 29, 1515–1541.

Kim, S. and Xing, E. P. (2009). Statistical estimation of correlated genome associations to a quantitative trait network. *PLoS genetics*, 5, e1000587.

Lichman, M. (2016). UCI machine learning repository. http://archive.ics.uci.edu/ml. University of California, Irvine, School of Information and Computer Sciences.

Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, 60, 234–256.

Meinshausen, N. (2013). Sign-constrained least squares estimation for high-dimensional regression. *Electronic Journal of Statistics*, 7, 1607–1631.

Miller, A. (2002). *Subset Selection in Regression* (2 ed.). Chapman & Hall/CRC.

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2012). *Introduction to Linear Regression Analysis*, Volume 821. John Wiley & Sons.

Rothberg, E. (2007). An evolutionary algorithm for polishing mixed integer programming solutions. *INFORMS Journal on Computing*, 19, 534–541.

Savelsbergh, M. (1994). Preprocessing and probing techniques for mixed integer programming problems. *ORSA Journal on Computing*, 6, 445–454.

Sengupta, D. and Bhimasankaram, P. (1997). On the roles of observations in collinearity in the linear model. *Journal of the American Statistical Association*, 92, 1024–1032.

Silvey, S. (1969). Multicollinearity and imprecise estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 539–552.

Tamura, R., Kobayashi, K., Takano, Y., Miyashiro, R., Nakata, K. and Matsui, T. (2017). Best subset selection for eliminating multicollinearity. *Journal of the Operations Research Society of Japan*, 60, 321–336.

Tamura, R., Kobayashi, K., Takano, Y., Miyashiro, R., Nakata, K. and Matsui, T. (2019). Mixed integer quadratic optimization formulations for eliminating multicollinearity based on variance inflation factor. *Journal of Global Optimization*, 73, 431–446.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 91–108.

Torgo, L. (2016). Regression data sets. http://www.dcc.fc.up.pt/ ltorgo/Regression/DataSets.html. University of Porto, Faculty of Sciences.

Watson, P. K. and Teelucksingh, S. S. (2002). *A Practical Introduction to Econometric Methods: Classical and Modern*. University of West Indies Press.

Winner, L. (2016). Miscellaneous data sets. http://www.stat.ufl.edu/ winner/datasets.html. University of Florida.

Yu, G. and Liu, Y. (2016). Sparse regression incorporating graphical structure among predictors. *Journal of the American Statistical Association*, 111, 707–720.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301–320.

A Mathematical formulation of the methods under comparison

The optimization problems that are solved in the numerical study take the form of Problem (6), whose objective function is that of the lasso or of OLS (i.e., the objective function in Problem (3), where $\lambda = 0$ in the case of the latter). The tightening sets are denoted as $\mathcal{S}_m$, $m = 1,..,5$, and were defined in Section 3.1. For the sake of comprehension, these problems will be explicitly stated now.

A.1 Lasso regression problem with tightening set $\mathcal{S}_1$

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & \|\mathbf{y} - \beta_0 - \boldsymbol{\beta}\mathbf{X}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1 \\ \text{s.t} \quad & \begin{cases} \nu_i^+ + \nu_j^- \leq 1 & \forall (i,j) \in \Omega_\alpha^+ \\ \nu_i^- + \nu_j^+ \leq 1 & \forall (i,j) \in \Omega_\alpha^+ \\ \nu_i^+ + \nu_j^+ \leq 1 & \forall (i,j) \in \Omega_\alpha^- \\ \nu_i^- + \nu_j^- \leq 1 & \forall (i,j) \in \Omega_\alpha^- \\ \nu_j^+, \nu_j^- \in \{0,1\} & \forall j = 1,\ldots,N \end{cases} \end{aligned} \tag{15}$$

A.2 Lasso regression problem with tightening set $\mathcal{S}_2$

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & \|\mathbf{y} - \beta_0 - \boldsymbol{\beta}\mathbf{X}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1 \\ \text{s.t} \quad & \begin{cases} \gamma_i + \gamma_j \leq 1 & \forall (i,j) \in \Omega_\eta \\ \gamma_j \in \{0,1\} & \forall j = 1,\ldots,N \end{cases} \end{aligned} \tag{16}$$

A.3 Lasso regression problem with tightening set $\mathcal{S}_3$

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & \|\mathbf{y} - \beta_0 - \boldsymbol{\beta}\mathbf{X}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1 \\ \text{s.t} \quad & \begin{cases} \beta_j \geq \epsilon\nu_j^+ - \nu_j^- M & \forall j = 1,\ldots,N \\ \beta_j \leq -\epsilon\nu_j^- + \nu_j^+ M & \forall j = 1,\ldots,N \\ \nu_j^+, \nu_j^- \in \{0,1\} & \forall j = 1,\ldots,N \end{cases} \end{aligned} \tag{17}$$

A.4 Lasso regression problem with tightening set $\mathcal{S}_4$

$$\begin{array}{ll}\min\limits_{\boldsymbol{\beta}} & \|\mathbf{y}-\beta_0-\boldsymbol{\beta}\mathbf{X}\|_2^2+\lambda\|\boldsymbol{\beta}\|_1 \\ \text{s.t} & \left\{\begin{array}{ll} \sum\limits_{j=1}^{N}(\nu_j^+ + \nu_j^-) \leq V_T & \\ \beta_j \geq \epsilon\nu_j^+ - \nu_j^- M & \forall j=1,\ldots,N \\ \beta_j \leq -\epsilon\nu_j^- + \nu_j^+ M & \forall j=1,\ldots,N \\ \nu_i^+ + \nu_j^- \leq 1 & \forall (i,j)\in\Omega_\alpha^+ \\ \nu_i^- + \nu_j^+ \leq 1 & \forall (i,j)\in\Omega_\alpha^+ \\ \nu_i^+ + \nu_j^+ \leq 1 & \forall (i,j)\in\Omega_\alpha^- \\ \nu_i^- + \nu_j^- \leq 1 & \forall (i,j)\in\Omega_\alpha^- \\ \nu_j^+, \nu_j^- \in \{0,1\} & \forall j=1,\ldots,N \end{array}\right. \end{array} \tag{18}$$

A.5 Lasso regression problem with tightening set $\mathcal{S}_5$

$$\begin{array}{ll}\min\limits_{\boldsymbol{\beta}} & \|\mathbf{y}-\beta_0-\boldsymbol{\beta}\mathbf{X}\|_2^2+\lambda\|\boldsymbol{\beta}\|_1 \\ \text{s.t} & \left\{\begin{array}{ll} \sum\limits_{j=1}^{N}\gamma_j \leq V_T & \\ \gamma_i + \gamma_j \leq 1 & \forall (i,j)\in\Omega_\eta \\ \gamma_j \in \{0,1\} & \forall j=1,\ldots,N \end{array}\right. \end{array} \tag{19}$$

B Calibration of correlation thresholds

In this section we calibrate the correlation thresholds η and α in the grid $\{0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0, 5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9\}$. Table 8 reports the predictive quality and the number of non-zero coefficients attained by calibrating these parameters. In comparison to results in Table 2, attained for fixed η and α, the calibrateD methods yield more dense outputs with a similar accuracy.

Table 8: *Predictive quality (MSE) and sparsity (NZ) for the approaches tightened with the correlation-based constraints when parameters η and α are calibrated.*

	Cpu		**Yacht**		**Whitewine**		**Redwine**		**Golf2008**		**Golf2009**		**Compact**	
	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ	MSE	NZ
OLS	1.000	6	1.000	6	1.000	11	1.000	11	1.000	6	1.000	11	1.000	21
$\mathcal{S}_1$	1.000	6	1.000	6	1.000	11	1.000	11	1.000	6	1.000	11	1.000	21
$\mathcal{S}_2$	1.038	5	0.987	5	1.000	10	0.999	7	1.005	3.5	0.889	8	1.016	15
lasso	1.000	5	0.959	2.5	1.000	10.5	0.997	9.5	1.000	4	0.798	9.5	1.000	20.5
$\mathcal{S}_1$	1.000	6	1.000	6	1.000	11	0.999	11	0.999	6	0.871	9.5	1.000	21
$\mathcal{S}_2$	1.042	5	0.969	5	1.000	10	1.000	8	0.998	3.5	0.812	7.5	1.015	15

C Correlated variables in simulated data with $N = 100$

In order to better understand the results of the tightened procedures displayed in Table 5 and Figure 7, in Figure 8 we have represented heatmaps that indicate whether two variables are highly correlated ($|\rho| \geq 0.8$) or moderately correlated ($|\rho| \geq 0.6$) for a random instance of each example of simulated data in Section 3.3.1. Orange colour indicates that the correlation constraint (7) is included in the tightening set $\mathcal{S}_5$ and also that sign coherence constraints (9)-(12) are added to the tightening set $\mathcal{S}_4$. Green colour stands only for the presence of sign coherence constraints (9)-(12) in $\mathcal{S}_4$. Left panels represent the correlations amongst all features, while right panels show the correlations only amongst the predictors truly appearing in the generating model (i.e. $\beta_i \neq 0$). As it can be observed, some features appearing in the generating model of Example 1 have a correlation larger than 0.8 in absolute value, and therefore they are forbidden to appear together in the output model of $\mathcal{S}_5$.

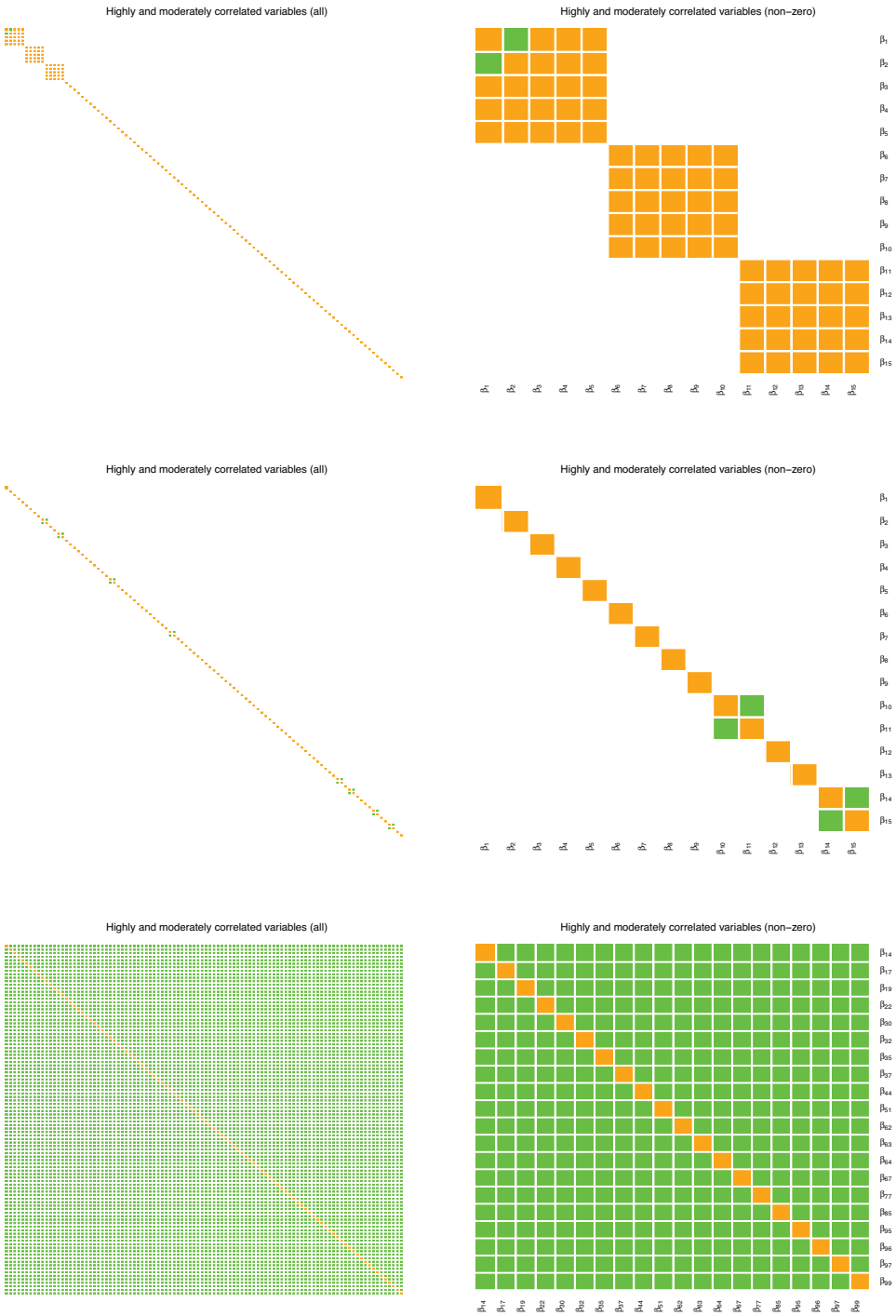

Figure 8: *Features exceeding a pairwise correlation of* 0.8 *in absolute value (orange), hence appearing in the correlation constraint* (7) *and also sign coherence constraints* (9)-(12)*, and features exceeding a pairwise correlation of* 0.6 *in absolute value (green), appearing solely in the sign coherence constraints.*

D AMPL code

```
#DATA PARAMETERS
param N;           #Number of predictors
param K;           #Number of observations
#SETS OF INDICES
set Nvar:=1..N;
set Nobs:=1..K;
#THE MATRIX OF DATA
param X {i in Nobs,j in 1..(N+1)};   # Kx(N+1) matrix of data
                                     # Includes response variable in position N+1
#PARAMETERS OF THE METHODS
param lambda;                        # Lasso penalty
param NZ;                            #Upper-bound on the total number of non-zeroes
param eps;                           #Significance threshold
param M default 100;                 #Upper bound for the coefficients beta
#SETS OF HIGHLY/MODERATELY CORRELATED VARIABLES
set conjcorrpos dimen 2;             # Positively correlated features
set conjcorrneg dimen 2;             # Negatively correlated features
#VARIABLES
var c ;                          # Intercept
var beta {j in Nvar};            # Slopes
var nupos {j in Nvar}, binary;
var nuneg {j in Nvar}, binary;
var v {j in Nvar} >=0 ;          # Auxiliar variables to express the absolute value
#OBJECTIVE FUNCTION
minimize fun:   (1/N)*sum{ i in Nobs} (X [i,p+1] -c-sum{ j in Nvar} (beta[j]*X[i,j]))^ 2
                +(1/N)*lambda*sum{ j in Nvar} v[j];
#CONSTRAINTS

#SPARSITY CONSTRAINT
subject to sparsity:   sum{ j in Nvar} (nuneg[j]+nupos[j])<=NZ;
#SIGNIFICANCE CONSTRAINTS
subject to significancepos {j in Nvar}:   beta[j]>=eps*nupos[j]-nuneg[j]*M;
subject to significanceneg {j in Nvar}:   beta[j]<=-eps*nuneg[j]+nupos[j]*M;
#SIGN COHERENCE CONSTRAINTS
subject to coherencepos1 {(j,r) in conjcorrneg}:   nupos[j]+nupos[r]<=1;
subject to coherencepos2 {(j,r) in conjcorrneg}:   nuneg[j]+nuneg[r]<=1;
subject to coherenceneg1 {(j,r) in conjcorrpos}:   nupos[j]+nuneg[r]<=1;
subject to coherenceneg2 {(j,r) in conjcorrpos}:   nuneg[j]+nupos[r]<=1;
#AUXILIAR CONSTRAINTS
subject to abs1 {j in Nvar}:      v[j]>=beta[j];
subject to abs2 {j in Nvar}:      v[j]>=-beta[j];
subject to sumnusj {j in Nvar}:   nuneg[j]+nupos[j]<=1;
```

SORT 44 (1) January-June 2020, 127-140 DOI: 10.2436/20.8080.02.97

Bartlett and Bartlett-type corrections for censored data from a Weibull distribution

Tiago M. Magalhães[1] and Diego I. Gallardo[2]

Abstract

In this paper, we obtain the Bartlett factor for the likelihood ratio statistic and the Bartlett-type correction factor for the score and gradient test in censored data from a Weibull distribution. The expressions derived are simple, we only have to define a few matrices. We conduct an extensive Monte Carlo study to evaluate the performance of the corrected tests in small sample sizes and we show how they improve the original versions. Finally, we apply the results to a real data set with a small sample size illustrating that conclusions about the regressors could be different if corrections were not applied to the three mentioned classical statistics for the hypothesis test.

MSC: Primary 62Fxx; secondary 62F12.

Keywords: Bartlett correction, censored data, Weibull distribution, chi-squared distribution, maximum likelihood estimates, type I and II censoring.

1 Introduction

Hypothesis testing is an essential step in statistical inference in order to help investigators identify and understand the effect of covariates on the response variable. Survival regression models are required when the response variable is censored, i.e., only partial information is available. Parametric survival models are often used in health economic applications (Latimer, 2013) because the survival function is fully specified (Ishak et al., 2013) and data from multiple time periods can be easily combined (Benaglia, Jackson and Sharples, 2015).

The likelihood ratio (LR), Wald, score and gradient tests are commonly used for hypothesis testing. Under the null hypothesis ($\mathcal{H}$), each test statistic is asymptotically chi-squared distributed, i.e., the four statistics are asymptotically equivalent. Since they are coupled with asymptotic properties, the chi-squared distribution may not be a good approximation to the null distribution of each statistic in small or moderate sample sizes,

[1] Department of Statistics, Institute of Exact Sciences, Federal University of Juiz de Fora, Juiz de Fora, Brazil. E-mail: tiago.magalhaes@ice.ufjf.br

[2] Departamento de Matemáticas, Facultad de Ingeniería, Universidad de Atacama, Copiapó, Chile. E-mail: diego.gallardo@uda.cl

Received: September 2019

Accepted: April 2020

then the use of these statistics become less justifiable. In practical situations, this fact can produce a type I error that should be greater (or less) than the fixed nominal value (usually 1%, 5% or 10%).

An approach to improve inferences in small/moderate samples using in the LR test is the Bartlett correction (Bartlett, 1937; Lawley, 1956). In this approach, the LR statistic is multiplied by a correction factor. Bartlett-type corrections were also developed for the score and gradient statistics, see Cordeiro and Ferrari (1991) and Vargas, Ferrari and Lemonte (2013).

Our main goal in this paper is to improve the likelihood inference in censored data from a Weibull distribution, where the scale parameter is known. Two particular models are obtained from this case: the exponential and Rayleigh distributions, but if unknown, the scale parameter may be replaced by a consistent estimate. First, we derive the Bartlett and the Bartlett-type correction for these censored data models. Next, we perform Monte Carlo simulation experiments to evaluate and compare the finite-sample performance of the improved LR, score and gradient tests with the usual LR, Wald, score and gradient tests. To the best of our knowledge, Bartlett and Bartlett-type corrections for LR, score and gradient statistics in the Weibull survival model were not specified so far. Moreover, it is the first presentation of corrections for the gradient statistic in survival models. All these results are illustrated by a comprehensive simulation study.

The paper is structured as follows. In Section 2, we describe the censored data from a Weibull distribution and discuss estimation and hypothesis testing inference on the regression parameters. The Bartlett and the Bartlett-type correction factors are derived in Section 3. Monte Carlo simulation results are presented and discussed in Section 4. An empirical application that use real data are presented and discussed in Section 5. The paper closes up with a brief discussion in Section 6.

2 Weibull distribution

A continuous random variable T is called Weibull, denoted by $\mathrm{WE}(\theta,\sigma)$, if its probability density function (pdf) is

$$f(t;\theta,\sigma)=\frac{1}{\sigma\theta^{1/\sigma}}t^{1/\sigma-1}\exp\left\{-(t/\theta)^{1/\sigma}\right\},\ t>0,$$

where $\sigma>0$ is the shape parameter and $\theta>0$ is the scale parameter. Weibull distribution is commonly used in the analysis of time-to-event or lifetime data and it is still the aim of several works, as in Jafari and Zakerzadeh (2015), Nekoukhou and Bidram (2015), Lina, Williamson and Kim (2019), Magalhães, Gallardo and Gómez (2019), for instance. Two particular models under this parametrization are obtained for $\sigma=1$ and $\sigma=1/2$, which represent the exponential and the Rayleigh models with mean θ and $\theta\sqrt{\pi}/2$, respectively. In this work, we focused on those models. However, if σ is unknown, we assume that it can be replaced by a consistent estimate. In lifetime data, censoring is very com-

mon because of time limits and other restrictions on data collection. To describe these data, we consider that, for a sample size of n, $L_1,\ldots,L_n$ are stochastically independent random variables representing the failure times and $T_1,\ldots,T_n$ are stochastically independent Weibull random variables and independent of the $L's$, denoting the censoring times. Under the right censoring, the observed information is

$$t_i = \min(T_i, L_i) \quad \text{and} \quad \delta_i = \begin{cases} 1, & T_i \leq L_i \\ 0, & T_i > L_i \end{cases}.$$

For $L_1,\ldots,L_n$ fixed, we have the type I mechanism and if $L_1,\ldots,L_n = L$, a random variable, type II censoring. Under the assumption that the censoring times $L's$ do not depend on θ (known in the literature as a non-informative censoring assumption), we have that the log-likelihood function for the two types of censoring has the form

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} \left\{ \frac{1}{\sigma\theta^{1/\sigma}} t_i^{1/\sigma-1} \right\}^{\delta_i} \exp\left\{ -(t_i/\theta)^{1/\sigma} \right\} \\ &= \left(\sigma\theta^{1/\sigma}\right)^{-r} \exp\left\{ \left(\frac{1}{\sigma} - 1\right) W_1 - \frac{1}{\theta^{1/\sigma}} W_2 \right\}, \end{aligned}$$

where $r = \sum_{i=1}^{n} \delta_i$, $W_1 = \sum_{i=1}^{n} \delta_i \log t_i$ and $W_2 = \sum_{i=1}^{n} t_i^{1/\sigma}$. The regression structure can be incorporated in the model by making $\theta_i = \exp\left(\mathbf{x}_i^\top \boldsymbol{\beta}\right)$, where $\boldsymbol{\beta}$ is a p-vector of parameters and $\mathbf{x}_i$ is a vector of regressors related to the i-th observation. Usually, the regression modelling considers the distribution of $Y_i = \log(T_i)$ instead of T_i. The distribution of Y_i is of the extreme value form with pdf given by

$$f(y_i;\mathbf{x}_i) = \frac{1}{\sigma} \exp\left\{ \frac{y_i - \mu_i}{\sigma} - \exp\left(\frac{y_i - \mu_i}{\sigma} \right) \right\}, \quad -\infty < y_i < \infty,$$

where $\mu_i = \log\theta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ is the linear predictor related to the i-th observation. The log-likelihood function derived from this regression model is given by

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[\delta_i \left(-n\log\sigma + \frac{y_i - \mu_i}{\sigma} \right) - \exp\left(\frac{y_i - \mu_i}{\sigma} \right) \right].$$

The total score function and the total Fisher information matrix for $\boldsymbol{\beta}$ are given, respectively, by

$$\mathbf{U}_{\boldsymbol{\beta}} = \sigma^{-1}\mathbf{X}^\top \mathbf{W}^{1/2}\mathbf{v} \text{ and } \mathbf{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \sigma^{-2}\mathbf{X}^\top \mathbf{W}\mathbf{X},$$

where $\mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_n)^\top$, the model matrix, assuming the $\text{rank}(\mathbf{X}) = p$, $\mathbf{W} = \text{diag}(w_1, \ldots, w_n)$, $w_i = \mathbb{E}\left[\exp\left(\frac{y_i-\mu_i}{\sigma}\right)\right]$ and $\mathbf{v} = (v_1,\ldots,v_n)^\top$, $v_i = \left\{-\delta_i + \exp\left(\frac{y_i-\mu_i}{\sigma}\right)\right\} w_i^{-1/2}$. It can be observed that the value of w_i depends on the mechanism of censoring. That means

$$w_i = 1 - \exp\left\{ -L_i^{1/\sigma} \exp(-\mu_i/\sigma) \right\} \text{ and } w_i = \frac{r}{n},$$

for types I and II censoring, respectively. The proofs are presented in Magalhães et al. (2019). The maximum likelihood estimator of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$, is the solution of $\mathbf{U}_{\boldsymbol{\beta}} = 0$. The MLE $\hat{\boldsymbol{\beta}}$ cannot be expressed in closed-form. It is typically obtained by numerically maximizing the log-likelihood function using a Newton or quasi-Newton nonlinear optimization algorithm. Under mild regularity conditions and in large samples, $\hat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \mathbf{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1})$, approximately.

Consider the p-dimensional parameter vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, where $\boldsymbol{\beta}_1$ is a q-dimensional vector and $\boldsymbol{\beta}_2$ is the remaining $p-q$ parameters. In a test of hypotheses, the interest lies in $\mathcal{H} : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$, the null hypothesis, where $\boldsymbol{\beta}_1^{(0)}$ is a known q-vector, in other words, the null hypothesis imposes q restrictions on the parameter vector. Hence, $\boldsymbol{\beta}_2$ is the vector of nuisance parameters and $\boldsymbol{\beta}_1$ is the vector of interest parameters. This partition induces the corresponding partitions

$$\mathbf{U}_{\boldsymbol{\beta}} = \left(\mathbf{U}_{\boldsymbol{\beta}_1}^\top, \mathbf{U}_{\boldsymbol{\beta}_2}^\top\right)^\top, \quad \text{with} \quad \mathbf{U}_{\boldsymbol{\beta}_1} = \sigma^{-1}\mathbf{X}_1^\top\mathbf{W}^{1/2}\mathbf{v}, \quad \mathbf{U}_{\boldsymbol{\beta}_2} = \sigma^{-1}\mathbf{X}_2^\top\mathbf{W}^{1/2}\mathbf{v},$$

$$\mathbf{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \begin{pmatrix} \mathbf{K}_{\boldsymbol{\beta}_1\boldsymbol{\beta}_1} & \mathbf{K}_{\boldsymbol{\beta}_1\boldsymbol{\beta}_2} \\ \mathbf{K}_{\boldsymbol{\beta}_2\boldsymbol{\beta}_1} & \mathbf{K}_{\boldsymbol{\beta}_2\boldsymbol{\beta}_2} \end{pmatrix} = \sigma^{-2}\begin{pmatrix} \mathbf{X}_1^\top\mathbf{W}\mathbf{X}_1 & \mathbf{X}_1^\top\mathbf{W}\mathbf{X}_2 \\ \mathbf{X}_2^\top\mathbf{W}\mathbf{X}_1 & \mathbf{X}_2^\top\mathbf{W}\mathbf{X}_2 \end{pmatrix},$$

and $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2]$, $\mathbf{X}_1$, $\mathbf{X}_2$ being $n \times q$ and $n \times (p-q)$, respectively. The LR, score and gradient statistics for testing $\mathcal{H}$ can be expressed, respectively, as

$$\begin{aligned} \text{SLR} &= 2\left[\ell\left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \sigma\right) - \ell\left(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2, \sigma\right)\right], \\ \text{SR} &= \tilde{\mathbf{v}}^\top\tilde{\mathbf{W}}^{1/2}\mathbf{X}_1\left(\tilde{\mathbf{R}}^\top\tilde{\mathbf{W}}\tilde{\mathbf{R}}\right)^{-1}\mathbf{X}_1^\top\tilde{\mathbf{W}}^{1/2}\tilde{\mathbf{v}}, \\ \text{ST} &= \sigma^{-1}\tilde{\mathbf{v}}^\top\tilde{\mathbf{W}}^{1/2}\mathbf{X}_1\left(\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)}\right), \end{aligned}$$

where $\left(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2\right)$ and $\left(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2\right)$ are the unrestricted and restricted MLEs of $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$, respectively, $\mathbf{R} = \mathbf{X}_1 - \mathbf{X}_2\mathbf{A}$, with $\mathbf{A} = \left(\mathbf{X}_2^\top\mathbf{W}\mathbf{X}_2\right)^{-1}\mathbf{X}_2^\top\mathbf{W}\mathbf{X}_1$ represents a $(p-q) \times q$ matrix whose columns are the vectors of regression coefficients obtained in the weighted normal linear regression of the columns of $\mathbf{X}_1$ on the model matrix $\mathbf{X}_2$ with $\mathbf{W}$ as a weight matrix. Here, tildes and hats indicate quantities available at the restricted and unrestricted MLEs, respectively. Under the null hypothesis $\mathcal{H}$, these three statistics have an asymptotic χ_q^2 distribution with approximation error of order n^{-1}.

3 Improved inference

As discussed in Section 2, when the sample size is not sufficiently large, the chi-squared distribution may be a poor approximation to the null distribution of the statistics. Thus, it is paramount to obtain refinements for inference based on these tests.

From the second-order asymptotic theory, three works can be mentioned: Lawley (1956), Cordeiro and Ferrari (1991) and Vargas et al. (2013). These works obtained general correction factors, respectively, for the LR, score and gradient statistics, which reduced the approximation error of the asymptotic χ^2_q distribution from n^{-1} to n^{-2}. Those correction factors are based on the derivatives of the log-likelihood function.

From the result of Lawley (1956), we derived the specific Bartlett-correction factor for LR statistic for testing $\mathcal{H}: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ in censored data from a Weibull distribution, it is given by

$$\begin{aligned}\varepsilon_p = (1/4)\sigma^{-2}\mathrm{tr}\left\{\mathbf{F}_1\dot{\mathbf{Z}}^{(2)}\right\} + (1/12)\sigma^{-6}\mathbf{1}^\top\mathbf{W}\left(2\mathbf{Z}^{(3)} + 3\dot{\mathbf{Z}}\mathbf{Z}\dot{\mathbf{Z}}\right)\mathbf{W}\mathbf{1} \\ + \sigma^{-5}\mathbf{1}^\top\mathbf{W}\left(\mathbf{Z}^{(3)} + \dot{\mathbf{Z}}\mathbf{Z}\dot{\mathbf{Z}}\right)\mathbf{W}'\mathbf{1} + \sigma^{-4}\mathbf{1}^\top\mathbf{W}'\left(\mathbf{Z}^{(3)} + \dot{\mathbf{Z}}\mathbf{Z}\dot{\mathbf{Z}}\right)\mathbf{W}'\mathbf{1},\end{aligned} \tag{1}$$

where $\mathbf{F}_1$, $\mathbf{W}'$, $\mathbf{Z}$ and $\dot{\mathbf{Z}}$ are given in the Appendix and all the algebraic manipulations are presented in the Supplementary Material, Section D.1. The three Bartlett corrected test statistics are

$$\mathrm{SLR}^*1 = \frac{SLR}{(1+c)},\ \mathrm{SLR}^*2 = \mathrm{SLR}\times\exp\{-c\} \quad \text{and} \quad \mathrm{SLR}^*3 = \mathrm{SLR}\times(1-c),$$

where $c = (\varepsilon_p - \varepsilon_{p-q})/q$, both ε_p and ε_{p-q}, can be obtained from (1). The statistic SLR*1 is the original Bartlett corrected likelihood ratio statistic. However, the others are equivalent to order $\mathcal{O}(n^{-1})$. It is noteworthy that SLR*2 assumes only positive values.

From Cordeiro and Ferrari (1991), we have written the specific Bartlett-type corrected score statistic for censored data from a Weibull distribution as

$$\mathrm{SR}^* = \mathrm{SR}\{1 - (c_R + b_R\mathrm{SR} + a_R\mathrm{SR}^2)\}, \tag{2}$$

where $a_R = A_{R3}/12q(q+2)(q+4)$, $b_R = (A_{R2} - 2A_{R3})/12q(q+2)$, $c_R = (A_{R1} - A_{R2} + A_{R3})/12q$ and, for the sake of brevity, the quantities A_{R1} to A_{R3} are presented in the Appendix.

For $\sigma = 1$, the expressions (1) and (2) reduce to exponential censored data case, derived by Cordeiro and Colosimo (1997) and Cordeiro and Colosimo (1999), respectively. For more details on the Bartlett and the Bartlett type corrections, see Cordeiro and Cribari-Neto (2014).

Finally, using the general result of Vargas et al. (2013), we obtained the specific Bartlett-type corrected gradient statistic for censored data from a Weibull distribution as

$$\mathrm{ST}^* = \mathrm{ST}\left\{1 - \left(c_T + b_T\mathrm{ST} + a_T\mathrm{ST}^2\right)\right\}, \tag{3}$$

where $a_T = A_{T3}/12q(q+2)(q+4)$, $b_T = (A_{T2} - 2A_{T3})/12q(q+2)$, $c_T = (A_{T1} - A_{T2} + A_{T3})/12q$ and the quantities A_{T1} to A_{T3} are also presented in the Appendix. For further discussion about gradient test and its Bartlett-type correction as well, see Lemonte (2016).

4 Simulation studies

In this section, we present four simulation studies to assess different aspects of our proposal. The first study is related to evaluating the type I error from the different corrected statistics under different combinations of (p,q), σ, % of censoring (C) and sample sizes. The second study is devoted to assessing the power of the corrected statistics. The third study evaluates the behaviour of the corrected statistics if the assumption of known σ is changed by the respective estimate, i.e., $\sigma = \widehat{\sigma}$, where $\widehat{\sigma}$ is some consistent estimator of σ. Finally, the fourth study assessed the performance of the corrected statistics if the scheme used to draw the censoring times is random (instead of censoring type I or type II), but considering as they were censoring type I. In all studies, we considered three values for σ: $0.5, 1$ and 3; eight combinations for (p,q): $(3,1),(3,2),(5,1),(5,2),(5,3),(7,1),(7,3)$ and $(7,5)$; 3 values for C: $10\%, 25\%$ and 50%; and 3 sample sizes: 20, 30 and 40, totaling 216 cases. We also considered $\boldsymbol{\beta} = 0_p$, i.e., a vector of zeros with dimension p. However, only the first q components of $\boldsymbol{\beta}$ were tested to be zero. For each combination of σ, (p,q), % of censoring and sample size we considered 20,000 Monte Carlo replicates. Each vector of covariates $\mathbf{x}_i$ was drawn from the multivariate standard normal distribution with dimension p. Values from the Weibull model were drawn using the inverse transformation method and right censoring type II scheme was used, i.e., the first $n \times (1 - C/100)$ times (rounded to the upper whole number) represented a failure time and the rest of units were censored at the $(1 - C/100)$-th quantile. The exception was the simulation study 4, where a right censoring scheme was used. For each sample and considering σ as known (except for simulation study 3 where such parameter was estimated from the sample) we compute the traditional statistics SLR, SR and ST and their modified version discussed in Section 3 to test $H_0 : \boldsymbol{\beta}_q = 0_q$ versus H_1 : the contrary. In all cases, we reported the percentage of times where the respective test rejected the null hypothesis giving a specified type I error. All simulations were performed using the R software (R Core Team, 2017).

4.1 Assessing the type I error

In this simulation study, we evaluate the type I error for the usual versions of SLR, SR and ST and their corrected versions discussed in Section 3 to test $H_0 : \boldsymbol{\beta}_q = 0_q$ versus H_1 : the contrary. We consider the four scenarios for (p,q), σ, % of censoring and sample sizes mentioned in the introduction of this section. We report the percentage of times where the test rejected the null hypothesis with a 5% significance. Table 1 summarizes the cases where $\sigma = 0.5$, $C = 25\%$, $n = 30$. The complete results are presented in the Supplementary Material, Section B.1. In general, the correction produces a rejection rate closer to the nominal 5% significance in the three tests. Considering the 216 involved cases, the mean of the rejection rates was 7.9%, 6.1% and 8.1% for the SLR, SR and ST tests and 5.6%, 5.5%, 5.3%, 5.6% and 5.3% for the SLR*1, SLR*2, SLR*3, SR* and ST* tests, respectively, showing a better performance in average terms to the corrected

statistics. We also compute the percentage of times where the corrected version of the test provides a rejection rate closer (in absolute value) to the nominal value. Such percentages were 99.5% for the SLR*1, SLR*2 and SLR*3, 69.0% for the SR* and 98.6% for the ST* test. Results suggest a huge improvement in the corrected version of the statistics when compared with their traditional pairs. As p and q increase, the differences in the rejection rates between the traditional statistics and the corrected ones seem to be getting larger. There are two possibles reasons, for a fixed sample size n: (a) Fixing q. As p increases, worse is the model fit and, consequently, the approximation to the null distribution of each statistic. (b) Fixing p. As q increases, there is the family-wise error rate (FWER), i.e., more restrictions in the null hypothesis make type I error larger, inflated. In the both situations, the corrected statistics seem to be less affected. Finally, we remark that the SR seems the most robust statistic among the three traditional statistics in this context.

Table 1: *Simulated rejection rates for $H_0 : \boldsymbol{\beta}_q = 0_q$, with $\sigma = 0.5$, $C = 25\%$, $n = 30$ and different values for p and q.*

p	q	SLR	SR	ST	SLR*1	SLR*2	SLR*3	SR*	ST*
3	1	7.0	6.2	7.0	5.8	5.8	5.7	5.9	5.7
	2	7.5	6.3	7.6	6.1	6.0	6.0	6.3	6.2
5	1	7.5	6.5	7.7	6.0	6.0	5.9	5.8	5.4
	2	8.3	6.5	8.4	6.1	6.0	5.9	6.0	5.9
	3	8.7	6.7	9.2	6.3	6.2	6.1	6.5	6.3
7	1	8.4	6.8	8.5	6.3	6.1	5.9	5.7	4.7
	3	10.4	7.5	10.7	6.8	6.6	6.4	6.6	6.3
	5	10.6	6.6	11.5	6.7	6.5	6.3	6.6	6.8

4.2 Assessing the power of the tests

In this simulation study, we assessed the power of the test for the usual versions of SLR, SR and ST and their corrected versions. We considered $n = 20$ and $p = 5$ in all the cases, q varying in the set $\{1,3\}$, σ in $\{0.5,1,3\}$ and C in $\{10\%,25\%,50\%\}$. To simulate the data, we further considered $\boldsymbol{\beta} = 0_p$. However, we have an interest in the hypothesis of the form $H_0 : \boldsymbol{\beta}_q = \psi \mathbf{1}_q$, where $\mathbf{1}_q$ is a vector of ones with dimension q and ψ is taken in the set $\{0.05,0.10,0.25,0.50,1.00,2.00\}$. Table 2 shows the results for $\sigma = 1$ and $C = 10\%$. The complete results are presented in the Supplementary Material, Section B.2. As expected, the power of the test is increased when ψ is increased (because the value being tested is further than the value used to simulate the data) and when q is increased. We also noted that the power of each test is greater than its corrected version for some values of ψ and is lower than its corrected version for other values of ψ. Therefore, as usual in most problems related to hypothesis tests, there is no unique most powerful test. However, the powers of the three ordinary and corrected tests seem similar.

Table 2: *Simulated rejection rates to the corrected version of the SLR, SR and ST tests for* $H_0 : \boldsymbol{\beta}_q = \psi \mathbf{1}_q$, *with* $\sigma = 1$, $C = 10\%$, $n = 20$, $p = 5$ *and different values for* q.

		ψ					
q	statistic	0.05	0.10	0.25	0.50	1.00	2.00
1	SLR	6.9	8.0	16.5	42.0	86.9	99.7
	SR	5.4	6.6	15.6	41.9	86.3	99.6
	ST	6.9	8.1	16.7	41.9	86.9	99.7
	SLR*1	7.7	8.7	17.0	41.3	86.1	99.7
	SLR*2	7.7	8.7	16.9	41.2	86.0	99.7
	SLR*3	7.6	8.7	16.9	41.2	85.9	99.7
	SR*	5.4	6.6	15.6	41.9	86.3	99.6
	ST*	6.7	7.8	17.3	44.7	88.8	98.5
3	SLR	7.5	10.0	28.9	75.9	99.3	100.0
	SR	6.4	9.8	33.0	79.1	99.4	100.0
	ST	8.1	10.8	30.1	76.5	99.3	100.0
	SLR*1	8.2	10.9	29.3	75.5	99.3	100.0
	SLR*2	8.2	10.9	29.3	75.5	99.3	100.0
	SLR*3	8.1	10.9	29.2	75.5	99.3	100.0
	SR*	6.4	9.8	33.0	79.1	99.4	100.0
	ST*	7.8	10.6	31.4	78.8	98.5	98.9

4.3 Changing the assumption of σ known

Up to this moment, we considered σ as a known value. However, in practice we also need to estimate it. An alternative is to fix $\sigma = \widehat{\sigma}_{ML}$ and apply all the discussed methodology, where $\widehat{\sigma}_{ML}$ denotes the ML estimator of σ for the complete model (i.e., with all covariates). However, as we are working in a framework with a small sample size, the bias of $\widehat{\sigma}_{ML}$ can be considerable. Previous studies performed by us suggest that the performance of the corrected statistics does not differ substantially from the traditional statistics to test $\boldsymbol{\beta}_q = 0_q$ versus $H_1 : \boldsymbol{\beta}_q \neq 0_q$. For this reason, in this simulation study, we considered fixing $\sigma = \widehat{\sigma}_J$, where $\widehat{\sigma}_J$ is the jackknife estimator for σ. A third alternative not explored by us was to fix $\sigma = \widehat{\sigma}_B$, where $\widehat{\sigma}_B$ is a bootstrap estimator for σ. However, $\widehat{\sigma}_J$ provides satisfactory results and $\widehat{\sigma}_J$ is determined in a unique form to a fixed sample, whereas $\widehat{\sigma}_B$ typically is computed based on $B >> n$ bootstrap resample, which is not unique and is more expensive in computational terms. We consider the four scenarios for (p,q), σ, % of censoring and sample sizes mentioned in the introduction of this section. We report the percentage of times where the test rejected the null hypothesis with a 5% significance. Table 3 summarizes the cases where $\sigma = 1$, $C = 25\%$, $n = 20$. The complete results are presented in the supplementary material, Section B.3. Considering the 216 involved cases, the mean of the rejection rates was 12.2%, 6.8% and 8.0% for the SLR, SR and ST tests and 8.8%, 8.5%, 8.2%, 6.3% and 5.6% for the SLR*1,

SLR*2, SLR*3, SR* and ST* tests, respectively, showing a better performance in average terms than the corrected statistics. We also compute the percentage of times where the corrected version of the test provides a rejection rate closer (in absolute value) to the nominal value. Such percentages were 100% for the SLR*1, SLR*2, SLR*3 and ST* and 73.6% for the SR*. Results suggest a huge improvement for the corrected version of the statistics when compared with their traditional counterparts.

Table 3: *Simulated rejection rates for $H_0 : \boldsymbol{\beta}_q = 0_q$, with $\sigma = 1$, $C = 25\%$, $n = 20$ and different values for p and q.*

p	q	SLR	SR	ST	SLR*1	SLR*2	SLR*3	SR*	ST*
3	1	6.9	5.6	6.9	5.6	5.5	5.4	5.2	5.2
	2	7.1	5.1	7.6	5.5	5.4	5.4	5.1	5.5
5	1	7.8	6.0	7.9	5.9	5.8	5.7	5.0	4.3
	2	8.9	6.0	9.1	6.0	5.8	5.5	5.3	5.2
	3	8.9	5.3	9.3	5.9	5.7	5.5	5.1	5.6
7	1	9.0	6.7	9.0	6.3	6.0	5.8	5.0	3.3
	3	11.5	6.6	12.0	6.7	6.4	6.0	5.6	5.0
	5	11.0	4.8	12.3	6.0	5.8	5.4	4.9	6.0

4.4 Changing the assumption of censoring type I or II

Up to this moment, all the development in this work was performed based on the assumption that the censoring scheme is either a type I or II. In this simulation study, we changed such assumption assuming that the censoring times $L_1, \ldots, L_n$ were independent random variables and independent from $T_1, \ldots, T_n$. For simplicity, we assumed that $L_i \sim WE(\lambda_i, 1)$, i.e., the exponential distribution with mean λ_i^{-1}. If the percentage of censoring times was fixed at $C\%$, we required that $\mathbb{P}(T_i > L_i) = C/100$. It was straightforward to show that such a condition was equivalent to

$$\int_0^\infty f_{T_i}(u; \theta_i, \sigma) \times e^{-\lambda_i u} du = C/100, \tag{4}$$

where $f_{T_i}(\cdot; \theta_i, \sigma)$ denotes the density function of $\mathrm{WE}(\theta_i, \sigma)$. Then, for a fixed value for θ_i, σ and C it was possible to solve numerically (4) to find λ_i. The same four scenarios for (p,q), σ, % of censoring and sample sizes mentioned in the introduction of this section were considered. We reported the percentage of times where the test rejected the null hypothesis with a 5% significance. Table 4 summarizes the cases where $\sigma = 3$, $C = 50\%$, $n = 20$. The complete results are presented in the Supplementary Material, Section B.4. Considering the 216 involved cases, the mean of the rejection rates were 6.5%, 4.8% and 6.4% for the SLR, SR and ST tests and 5.1%, 5.1%, 5.0%, 4.8% and 4.9% for the SLR*1, SLR*2, SLR*3, SR* and ST* tests, respectively, showing a better

performance in average terms to the corrected statistics. We also computed the percentage of times where the corrected version of the test provided a rejection rate closer (in absolute value) to the nominal value. Such percentages were 84.7%, 84.3% and 83.3% for the SLR*1, SLR*2 and SLR*3, respectively, 62.0% for the SR* and 73.1% for the ST* test. Furthermore, results suggest a huge improvement for the corrected version of the statistics when compared with their traditional pairs. The percentages of times where the corrected statistics are closer to the nominal value in comparison with the traditional statistics are lower than in simulation study 1, where σ is assumed as known. However, such percentages remain high and the correction is suggestible.

Table 4: *Simulated rejection rates for $H_0 : \boldsymbol{\beta}_q = 0_q$, with $\sigma = 3$, $C = 50\%$, $n = 20$ and different values for p and q.*

p	q	SLR	SR	ST	SLR*1	SLR*2	SLR*3	SR*	ST*
3	1	7.2	5.8	7.4	5.3	5.2	5.0	5.0	5.6
	2	7.5	4.6	7.7	5.5	5.4	5.2	4.7	5.9
5	1	9.0	6.9	9.1	6.2	5.8	5.5	5.3	6.0
	2	9.0	5.8	9.3	5.7	5.5	5.2	4.8	5.7
	3	10.0	4.8	10.1	6.0	5.7	5.4	4.6	5.9
7	1	10.7	8.2	10.8	6.9	6.4	5.8	5.5	5.8
	3	12.8	6.7	12.7	6.9	6.3	5.6	5.3	5.8
	5	13.3	4.5	12.3	6.7	6.2	5.6	4.8	5.6

5 Application

In this section we present a real data application related to clams in order to illustrate a case where conclusions obtained from a hypothesis test may be different if the corrections discussed in Section 3 are not considered in censored data from a Weibull distribution. In Section A of the Supplementary Material, we present a second application.

Clams data set

Bonnail et al. (2016) performed a study to assess sediment quality using the freshwater clam *Corbiculafluminea* to determine its adequacy as a biomonitoring tool in relation to theoretical risk indexes and regulatory thresholds. The clams were exposed to sediments contaminated with acid mine drainage (polymetallic acid lixiviate derived from mining activity). The study contains 27 observations with measurements, among other characteristics, of the dry weight tissue of the clams (`dry`, in gr), wet weight tissue (`wet`, in gr), shell length (`length`, in mm) and the concentration of scandium (`sc`), niobium (`nb`), beryllium (`be`) and terbium (`tb`) bioaccumulated in the soft tissue. These minerals were considered in micrograms per liter ($\mu g/L$). In this case, we focused on modelling the dry weight of such clams based on the rest of available information considering the Weibull regression model, i.e., $\texttt{dry}_i \sim \text{WE}(\theta_i, \sigma)$, where

$$\log\theta_i = \mathbf{x}_i^\top \boldsymbol{\beta} = \beta_1 \mathtt{sc}_i + \beta_2 \mathtt{nb}_i + \beta_3 \mathtt{be}_i + \beta_4 + \beta_5 \mathtt{wet}_i + \beta_6 \mathtt{length}_i,$$

with $\mathbf{x}_i^\top = (\mathtt{sc}_i, \mathtt{nb}_i, \mathtt{be}_i, 1, \mathtt{wet}_i, \mathtt{length}_i)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_6)^\top$, $i = 1, \ldots, 27$. The order of covariates was organized in order to test if minerals sc, nb and be explain the dry weight of the claims, i.e., to test $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ versus $H_1 : \beta_j \neq 0$, for at least one $1 \leq j \leq 3$. For this particular problem $p = 6$, $q = 3$ and $n = 27$, so as the sample size is small a correction might be required in traditional tests. We estimated $\widehat{\sigma}_J = 0.0317$ based on the jackknife method, which was used as known in the computation of the different statistics to test H_0. Results for traditional and corrected versions of the SLR, SR and ST tests are presented in Table 5. Note that, without correction, only the SR test does not reject H_0 considering a significance of 10%. However, all the corrected versions of the tests do not reject the null hypothesis with the same level of significance. Therefore, we cannot conclude that minerals sc, nb and be explain the dry weight of clams. Finally, to test if the WE model is suitable for this data set, we compute the quantile residuals (Dunn and Smyth, 1996). If the model was correct, these residuals would behave as a random sample from the standard normal distribution. The Kolmogorov-Smirnov test to verify such a hypothesis provides p-values of 0.206 and 0.568 to the complete and reduced model, respectively. Therefore, the assumption of WE distribution is acceptable under any usual level of significance.

Table 5: *Different statistics to test H_0 in the clams data set.*

Statistic	SLR	SR	ST	SLR*1	SLR*2	SLR*3	SR*	ST*
Observed	6.54	4.98	6.63	5.82	5.78	5.73	5.32	6.00
p-value	0.088	0.173	0.085	0.121	0.123	0.125	0.150	0.112

6 Concluding remarks

Weibull distribution is used for the analysis of time-to-event or lifetime data, with the maximum likelihood theory as the main methodology to estimate the parameters. Hypotheses regarding these parameters are tested using the likelihood, score and gradient tests. However, in small or moderate sample sizes, these procedures can not be reliable. In this paper, we derived the respectively corrected versions that improve their performance. For simplicity, we focus on the Weibull with known shape parameter (σ) to find those expressions. Nonetheless, our results show good properties for the situation when σ can be replaced by a consistent estimate based on the jackknife method. We also present an application that illustrates the usefulness of the main result of the paper. The matrices expressions for the Bartlett and Bartlett-type corrections are quite simple to be implemented in statistical software as R (R Core Team, 2017), together with the library flexsurv (Jackson, 2016), for instance. Noteworthy, we did not use bootstrap corrections because they have different natures and they add three uncertainties: the number of replications, the size of each replication and the type, parametric or nonparametric, besides being computationally costly for practitioners.

Acknowledgments

We gratefully acknowledge grants from FONDECYT (Chile) 11160670. We also acknowledge the editor's, the referees' and Dr. Marcio Augusto Diniz's suggestions that helped us improve this work.

References

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proccedings of the Royal Society, A*, 160, 268–282.

Benaglia, T., Jackson, C. H. and Sharples, L. D. (2015). Survival extrapolation in the presence of cause specific hazards. *Statistics in Medicine*, 34, 796–811.

Bonnail, E., Sarmiento, A. M. DelValls, T. A. Nieto, J. M. and Riba, I. (2016). Assessment of metal contamination, bioavailability, toxicity and bioaccumulation in extreme metallic environments (iberian pyrite belt) using corbicula fluminea. *Science of the Total Environment*, 544, 1031–1044.

Cordeiro, G. M. and Colosimo, E. A. (1997). Improved likelihood ratio tests for exponential censored data. *Journal of Statistical Computation and Simulation*, 56, 303–315.

Cordeiro, G. M. and Colosimo, E. A. (1999). Corrected score tests for exponential censored data. *Statistics & Probability Letters*, 44, 365–373.

Cordeiro, G. M. and Cribari-Neto, F. (2014). *An Introduction to Bartlett Correction and Bias Reduction*. Springer.

Cordeiro, G. M. and Ferrari, S. L. P. (1991). A modified score statistic having chi-squared distribution to order n^{-1}. *Biometrika*, 78, 573–582.

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5, 236–244.

Ishak, K. J., Kreif, N., Benedict, A. and Muszbek, N. (2013). Overview of parametric survival analysis for health-economic applications. *PharmacoEconomics*, 31, 663–675.

Jackson, C. H. (2016). flexsurv: A platform for parametric survival modeling in r. *Journal of Statistical Software*, 70, 1–33.

Jafari, A. A. and Zakerzadeh, H. (2015). Inference on the parameters of the weibull distribution using records. *Statistics and Operations Research Transactions*, 39, 3–18.

Latimer, N. R. (2013). Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data: inconsistencies, limitations, and a pratical guide. *Medical Decision Making*, 33, 743–754.

Lawley, D. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, 43, 295–303.

Lemonte, A. J. (2016). *The Gradient Test: Another Likelihood-Based Test*. Academic Press.

Lina, H.-M., Williamson, J. M. and Kim, H.-Y. (2019). Firth adjustment for weibull current-status survival analysis. *Communications in Statistics - Theory and Methods*. doi.org/10.1080/03610926.2019.1606241.

Magalhães, T. M., Gallardo, D. I. and Gómez, H. W. (2019). Skewness of maximum likelihood estimators in the weibull censored data. *Symmetry*, 11, 1351.

Nekoukhou, V. and Bidram, H. (2015). The exponentiated discrete weibull distribution. *Statistics and Operations Research Transactions*, 39, 127–146.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Vargas, T. M., Ferrari, S. L. P. and Lemonte, A. J. (2013). Gradient statistic: Higher-order asymptotics and bartlett-type correction. *Electronic Journal of Statistics*, 7, 43–61.

Appendix

In order to express the corrected LR, score and gradient statistics, equations (1) to (3) , it is helpful to define the quantities $w_i' = -\sigma^{-1}L_i^{1/\sigma}\exp\{-L_i^{1/\sigma}\exp(-\mu_i/\sigma)-\mu_i/\sigma\}$, $w_i'' = -\sigma^{-1}w_i'\left[L_i^{1/\sigma}\exp(-\mu_i/\sigma)-1\right]$, $f_{1i} = -\sigma^{-2}w_i - 4\sigma^{-1}w_i' - 4w_i''$, $f_{2i} = -2\sigma^{-2}w_i^2 + 6\sigma^{-2}w_i + 10\sigma^{-1}w_i' + 5w_i''$, $f_{3i} = -3\sigma^{-2}w_i^2 + 9\sigma^{-2}w_i + 14\sigma^{-1}w_i' + 6w_i''$, $i = 1,\ldots,n$, and the following matrices:

$$\begin{aligned}
&\mathbf{W}' = \mathrm{diag}(w_1',\ldots,w_n'),\ \mathbf{W}'' = \mathrm{diag}(w_1'',\ldots,w_n''),\\
&\mathbf{F}_1 = \mathrm{diag}(f_{11},\ldots,f_{1n}),\ \mathbf{F}_2 = \mathrm{diag}(f_{21},\ldots,f_{2n}),\ \mathbf{F}_3 = \mathrm{diag}(f_{31},\ldots,f_{3n}),\\
&\mathbf{Z} = \mathbf{X}\mathbf{K}_{\beta\beta}^{-1}\mathbf{X}^\top = \sigma^2\mathbf{X}\left(\mathbf{X}^\top\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^\top,\ \dot{\mathbf{Z}} = \mathrm{diagonal}\{\mathbf{Z}\},\\
&\mathbf{Z}_2 = \sigma^2\mathbf{X}_2\left(\mathbf{X}_2^\top\mathbf{W}\mathbf{X}_2\right)^{-1}\mathbf{X}_2^\top,\ \dot{\mathbf{Z}}_2 = \mathrm{diagonal}\{\mathbf{Z}_2\},\ \mathbf{Z}^{(2)} = \mathbf{Z}\odot\mathbf{Z},\ \mathbf{Z}^{(3)} = \mathbf{Z}^{(2)}\odot\mathbf{Z},
\end{aligned}$$

where $\odot$ represents a direct product and $\mathbf{1}$ is an n-dimensional vector of ones.

The remaining quantities to define an improved statistic in the score test, see equation (2), are:

$$\begin{aligned}
A_{R1} = {} & 3\sigma^{-6}\mathbf{1}^\top(\mathbf{W}+2\sigma\mathbf{W}')\dot{\mathbf{Z}}_2(\mathbf{Z}-\mathbf{Z}_2)\dot{\mathbf{Z}}_2(\mathbf{W}+2\sigma\mathbf{W}')\mathbf{1}\\
& + 6\sigma^{-6}\mathbf{1}^\top(\mathbf{W}+2\sigma\mathbf{W}')\dot{\mathbf{Z}}_2\mathbf{Z}_2\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)(2\mathbf{W}+3\sigma\mathbf{W}')\mathbf{1}\\
& + 6\sigma^{-6}\mathbf{1}^\top(3\mathbf{W}+4\sigma\mathbf{W}')\left[\mathbf{Z}_2^{(2)}\odot(\mathbf{Z}-\mathbf{Z}_2)\right](\mathbf{W}+2\sigma\mathbf{W}')\mathbf{1}\\
& - 6\sigma^{-2}\mathrm{tr}\left\{\mathbf{F}_2\dot{\mathbf{Z}}_2\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)\right\},
\end{aligned}$$

$$\begin{aligned}
A_{R2} = {} & -3\sigma^{-6}\mathbf{1}^\top(2\mathbf{W}+3\sigma\mathbf{W}')\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)\mathbf{Z}_2\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)(2\mathbf{W}+3\sigma\mathbf{W}')\mathbf{1}\\
& - 6\sigma^{-6}\mathbf{1}^\top(2\mathbf{W}+3\sigma\mathbf{W}')\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)(\mathbf{Z}-\mathbf{Z}_2)\dot{\mathbf{Z}}_2(\mathbf{W}+2\sigma\mathbf{W}')\mathbf{1}\\
& - 6\sigma^{-6}\mathbf{1}^\top(2\mathbf{W}+3\sigma\mathbf{W}')\left[\mathbf{Z}_2\odot(\mathbf{Z}-\mathbf{Z}_2)^{(2)}\right](2\mathbf{W}+3\sigma\mathbf{W}')\mathbf{1}\\
& + 3\sigma^{-2}\mathrm{tr}\left\{\mathbf{F}_3\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)^{(2)}\right\},
\end{aligned}$$

$$A_{R3} = \sigma^{-6}\mathbf{1}^\top(2\mathbf{W}+3\sigma\mathbf{W}')\left\{3\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)(\mathbf{Z}-\mathbf{Z}_2)\left(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\right)+2(\mathbf{Z}-\mathbf{Z}_2)^{(3)}\right\}(2\mathbf{W}+3\sigma\mathbf{W}')\mathbf{1}.$$

The quantities A_{T1} to A_{T3}, in the equation (3), to define an improved gradient test are, respectively:

$$
\begin{aligned}
A_{T1} = {} & 3\sigma^{-6}\mathbf{1}^{\top}\mathbf{W}\left\{(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2)(\mathbf{Z}+\mathbf{Z}_2)\dot{\mathbf{Z}}_2+\dot{\mathbf{Z}}_2(\mathbf{Z}-\mathbf{Z}_2)\dot{\mathbf{Z}}_2+2(\mathbf{Z}-\mathbf{Z}_2)\odot\mathbf{Z}_2^{(2)}\right\}\mathbf{W}\mathbf{1} \\
& +6\sigma^{-5}\mathbf{1}^{\top}\mathbf{W}\left\{(\mathbf{Z}_2+\mathbf{Z})\odot\left(\mathbf{Z}^{(2)}-\mathbf{Z}_2^{(2)}\right)+(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2)(\mathbf{Z}_2\dot{\mathbf{Z}}_2+\mathbf{Z}\dot{\mathbf{Z}})\right. \\
& \left.+2\left[\dot{\mathbf{Z}}_2(\mathbf{Z}\dot{\mathbf{Z}}-\mathbf{Z}_2\dot{\mathbf{Z}}_2)+\mathbf{Z}_2^{(2)}\odot(\mathbf{Z}-\mathbf{Z}_2)\right]\right\}\mathbf{W}'\mathbf{1} \\
& +12\sigma^{-4}\mathbf{1}^{\top}\mathbf{W}'\left\{\dot{\mathbf{Z}}\mathbf{Z}\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2\mathbf{Z}_2\dot{\mathbf{Z}}_2+\mathbf{Z}^{(3)}-\mathbf{Z}_2^{(3)}\right\}\mathbf{W}'\mathbf{1} \\
& -6\sigma^{-4}\,\mathrm{tr}\left\{\mathbf{W}(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2)\dot{\mathbf{Z}}_2+\sigma\mathbf{W}'(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_2)(\dot{\mathbf{Z}}+3\dot{\mathbf{Z}}_2)+2\sigma^2\mathbf{W}''\left(\dot{\mathbf{Z}}^{(2)}-\dot{\mathbf{Z}}_2^{(2)}\right)\right\},
\end{aligned}
$$

$$
\begin{aligned}
A_{T2} = {} & -3\sigma^{-6}\mathbf{1}^{\top}\mathbf{W}\left\{(1/4)(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})(\mathbf{Z}-\mathbf{Z}_{p-q})(3\dot{\mathbf{Z}}+\dot{\mathbf{Z}}_{p-q})\right. \\
& \left.+(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})\mathbf{Z}_{p-q}(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})+(1/2)(\mathbf{Z}-\mathbf{Z}_{p-q})^{(2)}\odot(\mathbf{Z}+3\mathbf{Z}_{p-q})\right\}\mathbf{W}\mathbf{1} \\
& -6\sigma^{-5}\mathbf{1}^{\top}\mathbf{W}\left\{(\mathbf{Z}-\mathbf{Z}_{p-q})\odot\left(\mathbf{Z}^{(2)}-\mathbf{Z}_{p-q}^{(2)}\right)+(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})(\mathbf{Z}\dot{\mathbf{Z}}-\mathbf{Z}_{p-q}\dot{\mathbf{Z}}_{p-q})\right\}\mathbf{W}'\mathbf{1} \\
& +3\sigma^{-4}\,\mathrm{tr}\left\{\mathbf{W}(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})^{(2)}+2\sigma\mathbf{W}'(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})^{(2)}\right\},
\end{aligned}
$$

$$
A_{T3} = (1/4)\sigma^{-6}\mathbf{1}^{\top}\mathbf{W}\left\{3(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})(\mathbf{Z}-\mathbf{Z}_{p-q})(\dot{\mathbf{Z}}-\dot{\mathbf{Z}}_{p-q})+2(\mathbf{Z}-\mathbf{Z}_{p-q})^{(3)}\right\}\mathbf{W}\mathbf{1}.
$$

Although the expressions for the three corrected statistics entails a great deal of algebra, the expressions only involve simple operations on diagonal matrices. Additionally, for type II censoring, i.e., $\mathbf{W}' = 0$, the expressions presented in (1) to (3) are simpler. For instance, the Bartlett-correction factor for LR statistic reduces to:

$$
\varepsilon_p = (1/4)\sigma^{-2}\mathrm{tr}\left\{\mathbf{F}_1\dot{\mathbf{Z}}^{(2)}\right\}+(1/12)\sigma^{-6}\mathbf{1}^{\top}\mathbf{W}\left(2\mathbf{Z}^{(3)}+3\dot{\mathbf{Z}}\mathbf{Z}\dot{\mathbf{Z}}\right)\mathbf{W}\mathbf{1}.
$$

SORT 44 (1) January-June 2020, 99-126 DOI: 10.2436/20.8080.02.96

Modelling count data using the logratio-normal-multinomial distribution

M Comas-Cufí*, J.A. Martín-Fernández, G. Mateu-Figueras and J. Palarea-Albaladejo

Abstract

The logratio-normal-multinomial distribution is a count data model resulting from compounding a multinomial distribution for the counts with a multivariate logratio-normal distribution for the multinomial event probabilities. However, the logratio-normal-multinomial probability mass function does not admit a closed form expression and, consequently, numerical approximation is required for parameter estimation. In this work, different estimation approaches are introduced and evaluated. We concluded that estimation based on a quasi-Monte Carlo Expectation-Maximisation algorithm provides the best overall results. Building on this, the performances of the Dirichlet-multinomial and logratio-normal-multinomial models are compared through a number of examples using simulated and real count data.

MSC: 62-07, 62F10, 62F86, 62P10, 62P25.

Keywords: Count data, Compound probability distribution, Dirichlet Multinomial, Logratio coordinates, Monte Carlo method, Simplex.

1 Introduction

A compound distribution of a random vector is the probability distribution resulting from assuming that its parameters are themselves random variables (Mosimann, 1962). This type of distribution plays an important role in mixture models (Lindsay, 1995) and Bayesian statistics, among others (Robbins, 1964, 1980). Practical applications are found in diverse areas such as genetics, microbiome studies, document classification and economics (Blei and Lafferty, 2007; Bouguila, 2008; Layton, and Siikamäki, 2009; Holmes, Harris and Quince, 2012; Silverman et al., 2018; Grantham et al., 2019).

Two classical distributions to model multivariate count data are the multinomial distribution and the multivariate Poisson distribution. Whilst in the first case the total number of counts per observation is a parameter, in the second it is not and it depends on the magnitude of the Poisson rates. In the literature, the multivariate Poisson distri-

* Department of Computer Science, Applied Mathematics and Statistics, University of Girona, Campus Montilivi (P4), E-17003 Girona. Email: mcomas@imae.udg.edu

Received: October 2018
Accepted: April 2020

butions has been compounded with Gamma distributions (Nelson, 1985) and with the multivariate log-normal distribution (Aitchison and Ho, 1989). In this article, we focus on distributions which are compounded with the multinomial distribution.

The Dirichlet-multinomial (DM) compound distribution (also called multivariate Pólya-Eggenberger distribution) is the most commonly used for modelling and analysing multivariate count data when they depend on a total number of trials and, unlike the ordinary multinomial distribution, some data overdispersion is present (Chapter 40, Johnson, Kotz and Balakrishnan, 1997). Let $\mathbf{X}$ be a random vector of counts. The DM distribution results from compounding a multinomial $\mathcal{M}(\mathbf{x};n,\boldsymbol{\pi}=\mathbf{p})$ for the measured vector of counts $\mathbf{X}$, with parameters n and $\boldsymbol{\pi}$ being the total number of trials and the vector of probabilities for the range of possible discrete outcomes respectively, and a Dirichlet $\mathcal{D}(\mathbf{p};\boldsymbol{\alpha})$ for the probabilities $\mathbf{p}$, with a parameter $\boldsymbol{\alpha}$. The probability mass function (pmf) of a DM distribution is $\mathcal{DM}(\mathbf{x};n,\boldsymbol{\alpha}) = \Pr(X=\mathbf{x};n,\boldsymbol{\alpha}) = \int_{\mathcal{S}^D} \mathcal{D}(\mathbf{p};\boldsymbol{\alpha})\,\mathcal{M}(\mathbf{x};n,\mathbf{p})\,\mathrm{d}\mathbf{p}$, where $\mathcal{S}^D$ refers to the unit simplex. The unit simplex is the sample space of random vectors $\mathbf{p}$ of length D consisting of strictly positive components adding up to one (Aitchison, 1986), i.e.

$$\mathcal{S}^D = \left\{ \mathbf{p} = (p_1,\ldots,p_D) \in \mathbb{R}^D \;\middle|\; p_k > 0 \text{ and } \sum_{k=1}^{D} p_k = 1 \right\}.$$

The closed form expression for the DM pmf is

$$\mathcal{DM}(\mathbf{x};n,\boldsymbol{\alpha}) = \frac{n!\,\Gamma\left(\sum_{k=1}^{D}\alpha_k\right)}{\Gamma\left(n+\sum_{k=1}^{D}\alpha_k\right)} \prod_{k=1}^{D} \frac{\Gamma(x_k+\alpha_k)}{x_k!\,\Gamma(\alpha_k)},$$

where $\Gamma(\cdot)$ is the well-known gamma function. The DM distribution is defined on the sample space of random count vectors. That is, the $\{n,D\}$-*simplex lattice*

$$\mathcal{S}^{(n,D)} = \left\{ \mathbf{x} = (x_1,\ldots,x_D) \;\middle|\; x_k \in \{0,1,\ldots,n\} \text{ and } \sum_{k=1}^{D} x_k = n \right\}, \tag{1}$$

consisting of random vectors of counts with components in the non-negative integer domain and sum equal to n (Scheffé, 1958). As stressed in Aitchison (1986) and Comas-Cufí, Martín-Fernández and Mateu-Figueras (2016), the Dirichlet distribution imposes a very strong independence structure: any pair of ratios between different components of $\mathbf{p}$ are assumed to be statistically independent. This heavily restricts its potential for data modelling when the analysis is based in ratios between parts, as it is the case of compositional data analysis (Comas-Cufí et al., 2016). Some generalisations of the Dirichlet have been proposed to overcome this difficulty with limited success (Connor and Mosimann, 1969; Minka, 2004; Ongaro and Migliorati, 2013).

In the 1980's John Aitchison introduced the compositional approach to model and analyse multivariate random vectors defined on the simplex (Aitchison, 1986). A number of methodological and practical contributions have been recently published in differ-

ent areas such as statistics Comas-Cufí, Martín-Fernández and Mateu-Figueras (2019), waste management (Edjabou et al., 2017), health (Chastin et al., 2015) and animal science (Palarea-Albaladejo et al., 2017).

We focus here on the logratio-normal-multinomial distribution resulting from compounding a multinomial distribution for the vector of counts $\mathbf{X}$ with a logratio-normal distribution for the corresponding vector of multinomial probabilities $\mathbf{p}$. The logratio-normal distribution was introduced in Mateu-Figueras, Pawlowsky-Glahn and Egozcue (2013) to model compositions. Also known as the normal distribution on the simplex (we denote it by $\mathcal{N}_{\mathcal{S}^D}$), it is defined using the ordinary multivariate normal probability density function (pdf) over a vector of orthonormal logratio coordinates (Egozcue et al., 2003) as follows:

$$\mathcal{N}_{\mathcal{S}^D}(\mathbf{p};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mathcal{N}(\mathbf{h};\boldsymbol{\mu},\boldsymbol{\Sigma}) = (2\pi)^{-\frac{D-1}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\mathbf{h}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{h}-\boldsymbol{\mu})\right), \quad (2)$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the usual expectation and covariance parameters, and $\mathbf{h} = (h_1,\ldots,h_{D-1})$ are orthonormal logratio coordinates of a composition $\mathbf{p}$ defined on $\mathcal{S}^D$ with respect to a predefined orthonormal basis of the simplex, see (Egozcue et al., 2003). Although the logratio-normal is a reparametrisation of the logistic-normal distribution (Aitchison and Shen, 1980), its definition avoids using the logistic transformation in order to focus on the appropriate reference measure (see Mateu-Figueras et al. (2013) for details). In our developments, we will use so-called isometric logratio (ilr) coordinates obtained from a particular choice of orthonormal basis as introduced in Egozcue et al. (2003). Namely, $\mathbf{h} = \text{ilr}(\mathbf{p})$ with elements

$$h_i = \sqrt{\frac{i}{i+1}}\ln\frac{\sqrt[i]{\prod_{j=1}^{i}p_j}}{p_{i+1}}, i = 1,\ldots,D-1. \quad (3)$$

Note that the composition $\mathbf{p}$ associated with orthonormal logratio coordinates $\mathbf{h}$ is obtained by inverse transformation $\mathbf{p} = \text{ilr}^{-1}(\mathbf{h})$. Importantly, using this particular ilr representation does not imply a lack of generality, since the results are invariant under change of orthonormal basis. This is because different orthonormal logratio coordinate systems are orthogonal rotations from one to another. In Billheimer, Guttorp and Fagan (2001) the multivariate logistic-normal-multinomial distribution was defined for modelling multinomial counts by compounding the multinomial distribution with the additive-logistic-normal distribution (Chapter 6, Aitchison (1986)). Practical applications of this model can be found in (Xia et al., 2013; Silverman et al., 2018) for microbiome data, or Hughes, Munkvold and Samita (1998) where the additive-logistic-normal was combined with the binomial distribution to model two-part compositions. In the following, we refer to the distribution obtained by composing the logratio-normal and the multinomial distribution as the logratio-normal-multinomial distribution (referred to as LNM in the following). From a probabilistic point of view, the logratio-normal-multinomial and the logistic-normal-multinomial models define the same law of prob-

ability. Nevertheless, we have decided to call it logratio-normal-multinomial in order to emphasize that, instead of using the logistic transformation, we use the orthonormal logratio coordinates together with the reference measure compatible with the algebraic-geometric structure of the simplex and with the compositional approach introduced by (Aitchison, 1986).

Markov Chain Monte Carlo (MCMC) methods have been used so far for parameter estimation with these models. For the case of multivariate logistic-normal-multinomial distribution see (Billheimer et al., 2001) and (Xia et al., 2013). Quasi-Monte Carlo methods (QMC) are well-known tools used to approximate high-dimensional integrals (Morokoff and Caflisch, 1995; Leobacher and Pillichshammer, 2014). They deviate from standard Monte Carlo in the type of sampling procedure used to approximate the integral. While classic Monte Carlo uses pseudo-random samples, QMC methods use quasi-random samples or low-discrepancy sequences. QMC methods have been successfully used in different parameter estimation scenarios (Drmota and Tichy, 1997; Pan and Thompson, 2007; Kuo et al., 2008) and have shown an improvement of the efficiency when embedded in an Expectation-Maximisation (EM) algorithm (Jank, 2005).

Building on these results, we propose more efficient tools to estimate the parameters of a LNM distribution. Their performance in modelling count data is compared with the DM distribution.

The work is organised as follows. In Section 2, some basic definitions are formally introduced. In Section 3, we derive the E and the M steps of an EM scheme for parameter estimation. We propose to combine QMC integration with the EM algorithm to estimate the parameters of the LNM distribution. Section 4 illustrates the use of DM and LNM distributions in three different examples based on simulated and real data. Lastly, Section 5 concludes with some final remarks.

All data analyses discussed in this work were conducted using the R statistical programming environment (R Development Core Team, 2015). Computer routines implementing the methods and the data sets can be obtained at https://github.com/mcomas/SORT-normal-multinomial.

2 Basic definitions and properties

The simplex $\mathcal{S}^D$ has an Euclidean vector space structure of dimension $D-1$ with its own basic operations (perturbation and powering), an inner product and a distance (so-called Aitchison distance) (Egozcue et al., 2003). According to this algebraic-geometric characterisation, compositions can be mapped onto the ordinary real space using logratio coordinates. The logratio-normal distribution is a model which is closed under the main operations in the simplex $\mathcal{S}^D$ (Mateu-Figueras et al., 2013). Also, it is a flexible distribution because it can model compositions whose components have different forms of dependence. Importantly, the density function (2) is defined with respect to what is called the Aitchison measure on the simplex, a probability measure different

from the ordinary Lebesgue measure on real space (Pawlowsky-Glahn and Egozcue, 2001; Mateu-Figueras, Pawlowsky-Glahn and Egozcue, 2011). The Aitchison measure is a natural measure on the simplex, compatible with its vector space structure and absolutely continuous with respect to the Lebesgue measure in the space of logratio coordinates (Mateu-Figueras et al., 2013).

As said above, the LNM is the distribution resulting from compounding the multinomial distribution $\mathcal{M}(\mathbf{x};n,\boldsymbol{\pi}=\mathbf{p})$ with the logratio-normal distribution $\mathcal{N}_{\mathcal{S}^D}(\mathbf{p};\boldsymbol{\mu},\boldsymbol{\Sigma})$. a random vector of counts $\mathbf{X}$ generated from a LNM distribution with parameters n, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is obtained in two steps:

- Firstly, a random composition $\mathbf{p}$ is generated using the logratio-normal distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.
- Secondly, a count random vector $\mathbf{X}$ is generated using the multinomial distribution with parameters n and $\boldsymbol{\pi}=\mathbf{p}$.

The pmf of a LNM distributions, expressed in terms of orthonormal logratio coordinates, is

$$\mathcal{LNM}(\mathbf{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) = \Pr(X=\mathbf{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) = \int_{\mathcal{S}^D} \mathcal{N}_{\mathcal{S}^D}(\mathbf{p};\boldsymbol{\mu},\boldsymbol{\Sigma})\mathcal{M}(\mathbf{x};n,\mathbf{p})\,\mathrm{d}_{\mathcal{A}}\mathbf{p} \tag{4}$$

$$= \int_{\mathbb{R}^{D-1}} \mathcal{N}(\mathbf{h};\boldsymbol{\mu},\boldsymbol{\Sigma})\,\frac{n!}{\prod_{k=1}^{D} x_k!}\,\prod_{k=1}^{D} \mathrm{ilr}_k^{-1}(\mathbf{h})^{x_k}\,\mathrm{d}\mathbf{h}, \tag{5}$$

where $\mathrm{ilr}_k^{-1}(\mathbf{h})$ stands for the k-th component of the composition $\mathbf{p}=\mathrm{ilr}^{-1}(\mathbf{h})$. Note that expression (4) is written with respect to the Aitchison measure, while expression (5) is written with respect to the Lebesgue measure in the logratio coordinate space. The LNM distribution is defined on $\mathcal{S}^{(n,D)}$ (Eq. 1).

Note that definition of the LNM distribution is similar to the definition of the DM distribution. The difference is that in the former the composition $\mathbf{p}$ is modelled by a normal distribution in terms of ilr-coordinates, instead of using a Dirichlet distribution.

Using the pmf given in (4) or (5), the following properties can be easily derived.

Property 1 *For a fixed $\boldsymbol{x}$ we have*

$$\lim_{\|\boldsymbol{\Sigma}\|\to 0} \mathcal{LNM}(\boldsymbol{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mathcal{M}\left(\boldsymbol{x};n,\boldsymbol{\pi}=\mathrm{ilr}^{-1}(\boldsymbol{\mu})\right).^{1}$$

1. $\lim_{\|\boldsymbol{\Sigma}\|\to 0}$ stands for any sequence of covariance matrices such that their highest eigenvalue goes to 0.

Proof. See Appendix A. ■

Property 2 *Let* $\boldsymbol{x} = (x_1, \ldots, x_D)$ *and* $x_1 + \cdots + x_D = n$. *If* $\lim_{n\to\infty} \frac{x_i}{n} = \pi_i$ *and* $\pi_i > 0$ *for* $1 \leq i \leq D$, *then*

$$\lim_{n\to\infty} n^{D-1} \cdot \mathcal{LNM}(\boldsymbol{x}; n, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}_{\mathcal{S}^D}(\boldsymbol{\pi}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \frac{1}{\sqrt{D}} \frac{1}{\pi_1 \ldots \pi_D}$$

Proof. See Appendix A. ■

Hence, depending on the parameters n and $\boldsymbol{\Sigma}$, the LNM distribution can be approximated by either a multinomial distribution or a logratio-normal distribution. The first property suggests that for count data sets where the random vector $\mathbf{p}$ has low variability, modelling based on either the LNM or multinomial distributions will provide similar results. The second property implies that a LNM distribution with large values of the number of trials n converges to the $\mathcal{N}_{\mathcal{S}^D}$ distribution. That is, for large n, the distribution of the random count vectors $\mathbf{X}$ on the simplex lattice $\mathcal{S}^{(n,D)}$ will be very similar to the distribution of the random vector $\mathbf{p}$ on the simplex $\mathcal{S}^D$.

3 Monte Carlo EM algorithm for logratio-normal-multinomial parameter estimation

Let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ be an independent and identically distributed sample of multivariate count data, with N denoting the sample size. To estimate the LNM parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ it is necessary to maximise the likelihood of the observed data given by

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{X}) = \prod_{i=1}^{N} \Pr(X = \mathbf{x}_i; n, \boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{6}$$

Denoting by $\mathbf{H}$ the non-observed ilr-coordinates, i.e. $\mathbf{H} = \{\mathbf{h}_1, \ldots, \mathbf{h}_N\}$, the EM algorithm (Dempster, Laird and Rubin, 1977) allows to maximise (6) by iteratively using an expected *augmented* likelihood $L(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{X}, \mathbf{H}) = \prod_{i=1}^{N} f(\mathbf{x}_i, \mathbf{h}_i; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the joint probability density function of random vectors X and H is

$$f(\mathbf{x}, \mathbf{h}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\mathbf{h}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \frac{(\sum_{k=1}^{D} x_k)!}{\prod_{k=1}^{D} x_k!} \prod_{k=1}^{D} \mathrm{ilr}_k^{-1}(\mathbf{h})^{x_k}.$$

In the E step, the expected value at the $(s+1)$-th iteration of the algorithm is calculated using the expression

$$Q(\boldsymbol{\mu},\boldsymbol{\Sigma} \mid \boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}) = \mathbb{E}_{H\mid X;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\ln L(\boldsymbol{\mu},\boldsymbol{\Sigma};\mathbf{X},\mathbf{H})\right] = \mathbb{E}_{H\mid X;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\sum_{i=1}^{N}\ln\left(f(\mathbf{x}_i,\mathbf{h}_i;\boldsymbol{\mu},\boldsymbol{\Sigma})\right)\right],$$

for a random vector H conditioned to $\mathbf{X}$ with parameters $\boldsymbol{\mu}^{(s)}$ and $\boldsymbol{\Sigma}^{(s)}$ obtained in the previous iteration. In the M step, the function $Q(\boldsymbol{\mu},\boldsymbol{\Sigma} \mid \boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)})$ is maximised with respect to the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. By expanding $Q(\boldsymbol{\mu},\boldsymbol{\Sigma} \mid \boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)})$, it holds that

$$\begin{aligned} Q(\boldsymbol{\mu},\boldsymbol{\Sigma} \mid \boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}) &= \mathbb{E}_{H\mid X;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\sum_{i=1}^{N}\ln\left(\mathcal{N}(\mathbf{h}_i;\boldsymbol{\mu},\boldsymbol{\Sigma})\,\frac{(\sum_{k=1}^{D}x_{ik})!}{\prod_{k=1}^{D}x_{ik}!}\prod_{k=1}^{D}\mathrm{ilr}_k^{-1}(\mathbf{h}_i)^{x_{ik}}\right)\right] \\ &= \sum_{i=1}^{N}\left\{\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\ln\left(\mathcal{N}(\mathbf{h}_i;\boldsymbol{\mu},\boldsymbol{\Sigma})\right)\right]\right\} + \\ &+ \sum_{i=1}^{N}\left\{\ln\left(\frac{(\sum_{k=1}^{D}x_{ik})!}{\prod_{k=1}^{D}x_{ik}!}\right) + \sum_{k=1}^{D}\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\ln\left(\mathrm{ilr}_k^{-1}(\mathbf{h}_i)^{x_{ik}}\right)\right]\right\}. \end{aligned}$$

To optimise the function Q with respect to the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, it is only necessary to optimise the terms where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are involved. That is, the term $Q^*(\boldsymbol{\mu},\boldsymbol{\Sigma}|\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}) = \sum_{i=1}^{N}\left\{\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[\ln\left(\mathcal{N}(\mathbf{h}_i;\boldsymbol{\mu},\boldsymbol{\Sigma})\right)\right]\right\}$. Using the linearity of the expectation $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}$, this term is maximised at the basic statistics $\boldsymbol{\mu}^{(s+1)} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i]$ and $\boldsymbol{\Sigma}^{(s+1)} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i^{\intercal}\mathbf{h}_i] - \hat{\boldsymbol{\mu}}^{\intercal}\hat{\boldsymbol{\mu}}$. In consequence, the critical point when applying the EM algorithm here is the calculation of the expected values $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i]$ and $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i^{\intercal}\mathbf{h}_i]$.

3.1 Quasi-Monte Carlo approximation to the E step

The expected values $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i]$ and $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[\mathbf{h}_i^{\intercal}\mathbf{h}_i]$ are calculated using Monte Carlo approximation, then turning the EM algorithm into a Monte Carlo EM algorithm (Jank, 2005; Neath, 2013). To simplify the exposition, in this subsection we denote the expected value $\mathbb{E}_{\mathbf{h}_i\mid\mathbf{x}_i;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}$ by simply $\mathbb{E}$. The vector $\mathbb{E}[\mathbf{h}] = (\mathbb{E}[h_k])_{k=1,\ldots,D-1}$ and the matrix $\mathbb{E}[\mathbf{h}^{\intercal}\mathbf{h}] = (\mathbb{E}[h_k h_r])_{k,r=1,\ldots,D-1}$ are particular cases of the general expression

$$\mathbb{E}[\varphi(\mathbf{h})] = \int_{\mathbb{R}^{D-1}}\varphi(\mathbf{h})\,f(\mathbf{h}\mid\mathbf{x};\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)})\,\mathrm{d}\mathbf{h}, \tag{7}$$

where $\varphi:\mathbb{R}^{D-1}\to\mathbb{R}$ and $f(\mathbf{h}\mid\mathbf{x};\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}) = \frac{f(\mathbf{x},\mathbf{h};\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)})}{\Pr\left(\{X=\mathbf{x}\};n,\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}\right)}$. Moreover, note that $\Pr\left(X=\mathbf{x};n,\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}\right) = \int_{\mathbb{R}^{D-1}}f(\mathbf{x},\mathbf{h};\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)})\,\mathrm{d}\mathbf{h}$. Hence, to evaluate (7), we need to approximate the integral

$$I\left(\varphi, \mathbf{x}, \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)}\right) = \int_{\mathbb{R}^{D-1}} \varphi(\mathbf{h})\, f(\mathbf{x}, \mathbf{h}; \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)})\, \mathrm{d}\mathbf{h} \tag{8}$$

for different functions φ.

In Xia et al. (2013), a MCMC method based on the Metropolis algorithm is used to estimate $\mathbb{E}[\mathbf{h}]$. Although the authors approximate the second moment, $\mathbb{E}[\mathbf{h}^\intercal\mathbf{h}]$, with the square of the first moment, $\mathbb{E}[\mathbf{h}]^\intercal \cdot \mathbb{E}[\mathbf{h}]$, we here estimate the first and the second moments, i.e. $\varphi(\mathbf{h}) = \mathbf{h}$ and $\varphi(\mathbf{h}) = \mathbf{h}^\intercal\mathbf{h}$, separately in the E step. To approximate $I\left(\varphi, \mathbf{x}, \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)}\right)$, we used Monte Carlo integration with importance sampling (Caflisch, 1998). In each E step, importance sampling is performed using a normal distribution centred at $\mathbf{m} = \mathbb{E}[\mathbf{h}]$ with covariance $\mathbf{S} = \mathbb{E}[(\mathbf{h}-\mathbf{m})^\intercal(\mathbf{h}-\mathbf{m})]$ calculated in the previous E step. The integral $I\left(\varphi, \mathbf{x}, \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)}\right)$ is approximated by

$$\begin{aligned} I\left(\varphi, \mathbf{x}, \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)}\right) &= \int_{\mathbf{h}\in\mathbb{R}^{D-1}} \frac{\varphi(\mathbf{h})\, f(\mathbf{x}, \mathbf{h}; \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)})}{\mathcal{N}(\mathbf{h}; \mathbf{m}, \mathbf{S})} \mathcal{N}(\mathbf{h}; \mathbf{m}, \mathbf{S}) \mathrm{d}\mathbf{h} \\ &= \mathbb{E}_{\mathbf{W}\sim\mathcal{N}(\mathbf{m},\mathbf{S})}\left[\frac{\varphi(\mathbf{w})\, f(\mathbf{x}, \mathbf{w}; \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)})}{\mathcal{N}(\mathbf{w}; \mathbf{m}, \mathbf{S})}\right] \\ &\approx \frac{1}{M}\sum_{r=1}^{M} \frac{\varphi(\mathbf{w}_r)\, f(\mathbf{x}, \mathbf{w}_r; \boldsymbol{\mu}^{(s)}, \boldsymbol{\Sigma}^{(s)})}{\mathcal{N}(\mathbf{w}_r; \mathbf{m}, \mathbf{S})}\,, \end{aligned} \tag{9}$$

where the values $\mathbf{w}_r$, $r = 1, \ldots, M$, are sampled from a normal distribution $\mathcal{N}(\mathbf{m}, \mathbf{S})$.

We here adopt a QMC approach which, instead of using pseudo-random normal generators, employs low-discrepancy sequences to generate the random values $\mathbf{w}_r$ (Caflisch, 1998; Wang and Fang, 2003; Leobacher and Pillichshammer, 2014). A low-discrepancy sequence is an equidistributed sample defined on a particular domain that is generated at a low computational cost (Chapter 2, Leobacher and Pillichshammer (2014)). Although different low-discrepancy sequences exist, we only considered Halton and Sobol sequences (Chapter 1, Drmota and Tichy (1997)). To choose between them we followed Morokoff and Caflisch (1995), which suggests best performance of Sobol sequences when the dimension of $\mathbf{h}$ is higher than six. Halton sequences are instead recommended for lower dimensions. QMC methods have shown to improve efficiency when combined with an EM algorithm (Jank, 2005).

Appendices B and C include a comparative of the performance of different methods in a univariate but extreme case and on a number of multidimensional cases respectively. In these scenarios, the best approximations for the first and second moments were obtained using the QMC approach. By contrast, methods based on MCMC algorithms showed the worst performance and highest computing time.

4 Examples

In this section, we consider three different contexts where N multinomial observations $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ are generated from N probability vectors $\mathbf{p}_1, \ldots, \mathbf{p}_N$. In the first scenario we consider a pure multinomial process (i.e. $\mathbf{p}_1 = \cdots = \mathbf{p}_N$). In the second scenario we consider each $\mathbf{p}_i$ in Hardy-Weinberg equilibrium (i.e. $p_{i2} = 4p_{i1}p_{i3}$, which implies a correlation of *minus one* between the logratios $\ln(\frac{p_{\cdot 1}}{p_{\cdot 2}})$ and $\ln(\frac{p_{\cdot 3}}{p_{\cdot 2}})$). The last case is a real scenario where no implicit relation between $\mathbf{p}_i$'s is assumed. In these three scenarios we compare the ability to model the sample $\mathbf{X}$ using both DM and LNM compound distributions. The main aim is to investigate how probability vectors $\mathbf{p}_i$ are modelled using the expected posterior probabilities calculated using distributions DM and LNM, i.e. $\hat{\mathbf{p}}_{i,\mathrm{DM}} = \mathbb{E}_{P|X=\mathbf{x}_i;\hat{\boldsymbol{\alpha}}}[P]$ and $\hat{\mathbf{p}}_{i,\mathrm{LNM}} = \mathrm{ilr}^{-1}\left(\mathbb{E}_{H|X=\mathbf{x}_i;\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}}}[H]\right)$ respectively.

For the LNM distribution we considered two different possibilities as starting point for the EM algorithm (SP1 and SP2 below):

- SP1: Given model parameters $\boldsymbol{\mu}_t^*$ and $\boldsymbol{\Sigma}_t^*$ evaluated at iteration t and observation $\mathbf{x}$, the maximum $\mathbf{h}^*$ of $f(\mathbf{h} \mid \mathbf{x}; \boldsymbol{\mu}_t^*, \boldsymbol{\Sigma}_t^*)$ can be easily calculated. Thus, the following iterative algorithm was defined:

  1. Set $t = 0$ and initiate $\boldsymbol{\mu}_0^*$ and $\boldsymbol{\Sigma}_0^*$ using sample mean and identity matrix in logratio coordinates.
  2. For each $\mathbf{x}_i \in \mathbf{X}$ calculate $\mathbf{h}_i^*$ maximising $f(\mathbf{h} \mid \mathbf{x}_i; \boldsymbol{\mu}_t^*, \boldsymbol{\Sigma}_t^*)$.
  3. Set $\boldsymbol{\mu}_{t+1}^* = \frac{1}{N}\sum \mathbf{h}_i^*$ and $\boldsymbol{\Sigma}_{t+1}^* = \frac{1}{N}\sum (\mathbf{h}_i^* - \boldsymbol{\mu}_{t+1}^*)^\intercal (\mathbf{h}_i^* - \boldsymbol{\mu}_{t+1}^*)$.
  4. If $\|\boldsymbol{\mu}_{t+1}^* - \boldsymbol{\mu}_t^*\|_\infty > 0.001$ go to step 2. Otherwise, stop and set $\boldsymbol{\mu}_0 = \boldsymbol{\mu}_{t+1}^*$ and $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_{t+1}^*$.

- SP2: Set $\boldsymbol{\mu}_0 = \overline{\mathrm{ilr}(\mathbf{p}_{\mathrm{DM}})}$ and $\boldsymbol{\Sigma}_0 = \mathrm{Cov}[\mathrm{ilr}(\mathbf{p}_{\mathrm{DM}})]$. That is, LNM estimation started using the final estimates of the mean vector and covariance matrix in logratio coordinates obtained from the DM compound distribution.

To estimate the parameters of the LNM distribution, we iterated the EM algorithm until the distance between two consecutive estimates was lower than a certain tolerance value $\tau = 0.001$. The expected values in (7) were approximated using 10000 iterations of a QMC simulation scheme based on Sobol sequences for the first and third example and Halton sequences for the second example.

4.1 A pure multinomial process

A sample $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ was generated, with $\mathbf{x}_i \sim \mathcal{M}(\mathbf{x}; n, \boldsymbol{\pi} = \mathbf{p})$ using a unique probability vector $\mathbf{p}$. Following Martín-Fernández et al. (2015), we designed nine different settings for $\mathbf{p}$ (see details in Appendix D). For each one we considered an increasing number of multinomial trials $n \in \{50, 100, 200, 500\}$, resulting in 9×4 different scenar-

ios. For each scenario we generated 10 different replicates of $\mathbf{X}$ with $N = 1000$ count vectors from the corresponding multinomial model. In total, we used $9 \times 4 \times 10 = 360$ samples of size 1000. For each of the replicates $\mathbf{X}$, we calculated the expected value $\hat{\mathbf{p}}_{i,\text{DM}}$ and $\hat{\mathbf{p}}_{i,\text{LNM}}$. It is reasonable that for any vector of counts $\mathbf{x}_i$, expected values $\hat{\mathbf{p}}_{i\cdot}$ will be close to $\mathbf{p}$. To evaluate how close estimations $\hat{\mathbf{p}}_{i\cdot}$ were to $\mathbf{p}$ we computed the mean of the Aitchison distances, $\text{dist}_{\mathcal{A}}$, between them; or, equivalently, the mean of the Euclidean distances, $\text{dist}_{\mathcal{E}}$, between the corresponding ilr coordinates (Egozcue et al., 2003):

$$\frac{1}{1000}\sum_{i=1}^{1000}\text{dist}_{\mathcal{A}}\left(\mathbf{p},\hat{\mathbf{p}}_{i\cdot}\right) = \frac{1}{1000}\sum_{i=1}^{1000}\text{dist}_{\mathcal{E}}\left(\text{ilr}(\mathbf{p}),\text{ilr}(\hat{\mathbf{p}}_{i\cdot})\right). \tag{10}$$

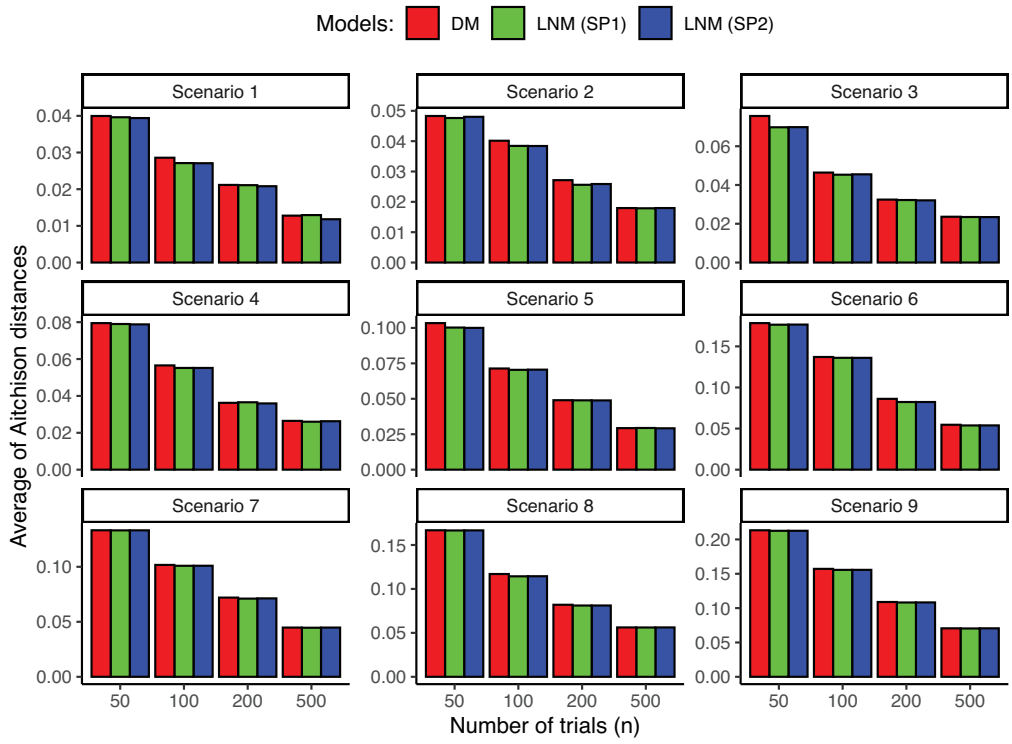

Figure 1: *Pure multinomial process: average of Aitchison distances between true* $\mathbf{p}$ *and estimated* $\hat{\mathbf{p}}_{i\cdot}$ *values using models DM (red bars), LNM with starting point SP1 (green bars) and LNM with starting point SP2 (blue bars). Nine different scenarios were set up considering four different number of trials in each one (n from 50 to 500). See Appendix D for details.*

This value was afterwards averaged across the 10 replicates. Figure 1 shows the results for the nine scenarios. As expected, DM and LNM produced similar results when modelling probabilities $\mathbf{p}$. Note that, in this example considering a multinomial setting, the estimate for $\boldsymbol{\Sigma}$ is close to the zero matrix (Property 1). Because of this, in some

scenarios after initialisations SP1 and SP2, the covariance matrix $\mathbf{\Sigma}_0$ was close to degenerate and the EM algorithm stopped in the first iteration (see Table *2*). In general, when the number of trials increases the error decreases as expected.

4.2 The Hardy-Weinberg equilibrium

In this case the variability of the unobserved vectors of probabilities $\mathbf{p}_1,\ldots,\mathbf{p}_N$ is governed by the Hardy-Weinberg equilibrium (Graffelman, and Weir, 2016). In brief, a biallelic genetic marker with alleles A and B with respective frequencies q and $(1-q)$ is in Hardy-Weinberg equilibrium if the genotype frequencies $\mathbf{x} = (f_{\mathrm{AA}}, f_{\mathrm{AB}}, f_{\mathrm{BB}})$ are given by $\mathbf{p} = (q^2, 2q(1-q), (1-q)^2)$. To obtain our sample $\mathbf{X} = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$, we generated N uniform random variables q_i, taking values between 0 and 1, in six different scenarios:

1. $q_i \sim Unif(0,1)$,
2. $q_i \sim Unif(0,0.5)$,
3. $q_i \sim Unif(0,0.25)$,
4. $q_i \sim Unif(0.25,0.5)$,
5. $\mathrm{logit}(q_i) \sim Norm(0,1)$, and
6. $\mathrm{logit}(q_i) \sim Norm(1,1)$.

The probabilities $\mathbf{p}_i$ were calculated using q_i according to the Hardy-Weinberg equilibrium. Observations $\mathbf{x}_i$ were drawn from a multinomial distribution with parameters $n \in \{50,100,200,500\}$, and $\pi = \mathbf{p}_i$ (Graffelman, 2015). After fitting DM and LNM models to sample $\mathbf{X}$, we computed the expected probability vector of probabilities $\hat{\mathbf{p}}_{i,\mathrm{DM}}$ and $\hat{\mathbf{p}}_{i,\mathrm{LNM}}$ using DM and LNM models respectively.

We compared estimations $\hat{\mathbf{p}}_{i\cdot}$ to vector $\mathbf{p}_i = (q_i^2, 2q_i(1-q_i), (1-q_i)^2)$ by using the average of Aitchison distances (10) as in the previous example. The results are displayed in Figure 2. Unlike with the pure multinomial process, there is a linear relation between the three parts of the composition. Consequently, the probabilities $\mathbf{p}_i$ could be better approximated in all cases using LNM instead of DM, with negligible differences for different starting points SP1 or SP2. Again, the error decreases with increasing number of trials as expected.

To illustrate the performance of the models in presence of variability in the probability vectors $\mathbf{p}_i$, we used ternary diagrams to graphically represent the first 50 simulated allele genotype probability vectors (Figure 3) and their genotype frequency (Figure 4).

Figure 3 (left) shows probability vectors $\{\mathbf{p}_1\ldots,\mathbf{p}_{50}\}$ satisfying the Hardy-Weinberg equilibrium which were generated from the first scenario above ($q \sim Unif(0,1)$). Note that they exactly fit a (compositional) line described by the parametric equations $\{(t^2, 2t(1-t), (1-t)^2) : 0 < t < 1\}$. Figure 3 (centre) shows that estimates $\{\hat{\mathbf{p}}_{1,\mathrm{DM}},\ldots,\hat{\mathbf{p}}_{50,\mathrm{DM}}\}$ from the DM model are far more scattered with respect to the equilibrium state than those from the LNM model $\{\hat{\mathbf{p}}_{1,\mathrm{LNM}},\ldots,\hat{\mathbf{p}}_{50,\mathrm{LNM}}\}$ (Figure 3 (right)). Taht is, while the LNM model is able to capture the variability along the compositional line the DM is

not.

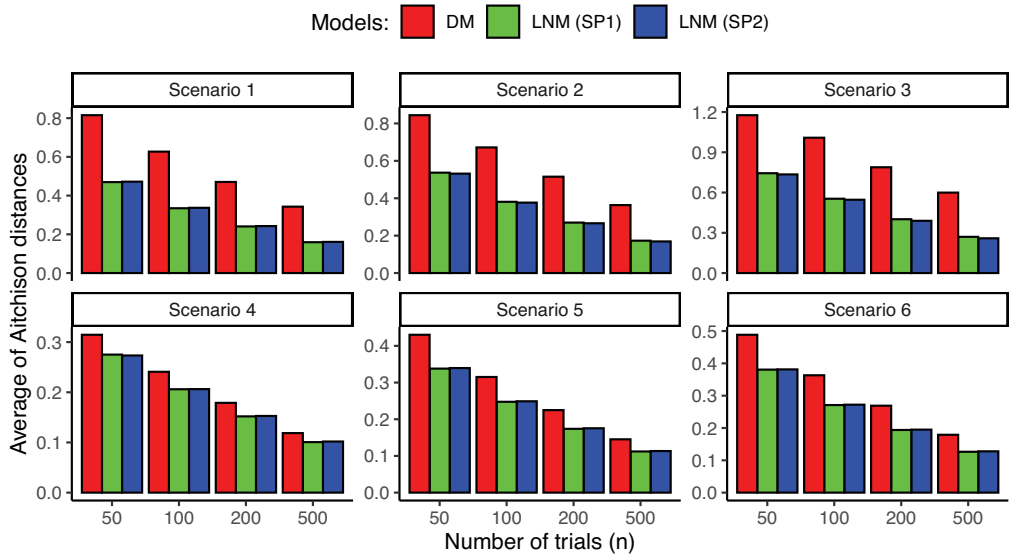

Figure 2: *Hardy-Weinberg equilibrium: average of Aitchison distances between true* $\boldsymbol{\pi}$ *and estimated* $\boldsymbol{p}_i$ *values using models DM (red bars), LNM with starting point SP1 (green bars) and LNM with starting point SP2 (blue bars). Six different scenarios were set up considering four different numbers of trials in each one (n from 50 to 500).*

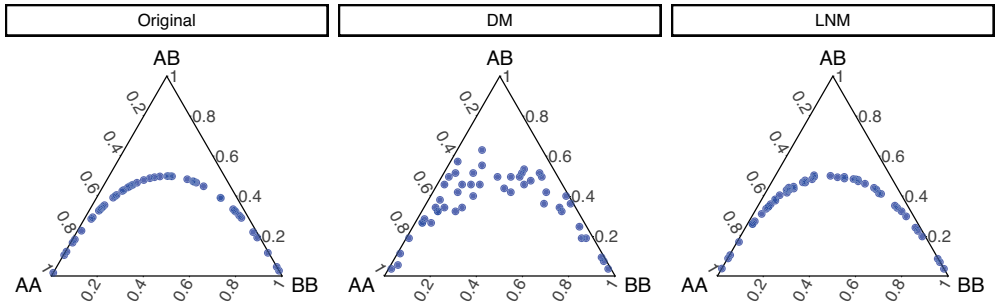

Figure 3: *Probabilities* $\boldsymbol{p}_i$ *in the Hardy-Weinberg equilibrium. Original probabilities* $\boldsymbol{p}_i$ *distributed according to Scenario 1 (left), estimates* $\hat{\boldsymbol{p}}_{i,DM}$ *using the Dirichlet-multinomial model (centre), estimates* $\hat{\boldsymbol{p}}_{i,LNM}$ *using the logratio-normal-multinomial model (right).*

This behaviour is made more evident when samples of count data are generated with both models. Figure 4 (left) shows genotype frequency data vectors $\{\mathbf{x}_1 \ldots, \mathbf{x}_{50}\}$ generated from the Hardy-Weinberg equilibrium using a multinomial distribution (Graffelman, 2015). In the centre we can see how data randomly generated from a DM model with parameter $\hat{\boldsymbol{\alpha}} = (0.642, 0.858, 0.622)$, which was estimated from sample $\mathbf{X}$, spread all over the ternary diagram. On the right-hand side, data were randomly generated using a LNM model with parameters

$$\hat{\boldsymbol{\mu}} = (-0.460, 0.343) \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 1.528 & 2.706 \\ 2.706 & 4.816 \end{pmatrix}$$

estimated from $\mathbf{X}$. In this case, the obtained sample is a fairly accurate realisation of the original data.

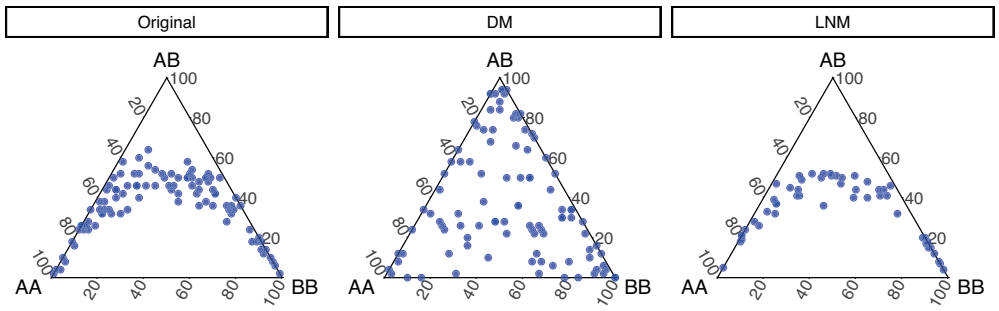

Figure 4: *Genotype frequencies $\boldsymbol{x}_i$ in the Hardy-Weinberg equilibrium. Original frequencies $\boldsymbol{x}_i$ generated in Scenario 1 (left), genotype frequencies simulated by Dirichlet-multinomial model (centre), genotype frequencies simulated by logratio-normal-multinomial model (right).*

4.3 Catalan parliamentary elections

Final vote shares by municipality from the 2015 Catalan parliamentary elections were provided by the Statistical Institute of Catalonia (http://www.idescat.cat/en/). A subsample was drawn for statistical inference based on the DM and LNM models. We aimed to evaluate by simulation the ability of an opinion poll to reflect voting intention. Table *1* shows the vote counts registered in some municipalities out of 947. We only considered votes for the six main parties obtaining seats in the parliament: Junts pel sí (jxsi), Partit socialista de Catalunya (psc), Partit popular (pp), Catalunya si que es pot (catsp), Ciutadans (cs) and Candidatura d'unitat popular (cup). Moreover, in order to have reasonable estimates for our simulations, we only considered those municipalities with at least 1000 votes registered (369 municipalities). The number of votes registered in these municipalities ranged from 1006 votes to 836687 votes.

Table 1: *Distribution of votes by municipality across political parties in 2015 Catalan parliamentary election (only some municipalities shown).*

Municipality	**jxsi**	**psc**	**pp**	**catsp**	**cs**	**cup**	**Population**
Abella de la Conca	57	11	5	3	9	18	170
Abrera	1685	1275	619	938	1771	453	12125
Àger	271	14	25	10	18	45	594
Agramunt	1940	161	185	104	252	152	5515
Aguilar de Segarra	121	0	4	5	4	18	251
Agullana	298	26	15	16	31	80	826
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Using all available data, for each municipality i we calculated the vector of probabilities $\mathbf{p}_i = \frac{1}{n'_i}\left(n'_{i,\text{jxsi}},\ldots,n'_{i,\text{cup}}\right)$, where $n'_{i,j}$ is the total number of votes to party j in municipality i and $n'_i = n'_{i,\text{jxsi}} + \cdots + n'_{i,\text{cup}}$. In this example, the vector of probabilities $\mathbf{p}_i$, $i = 1\ldots369$, was considered the gold standard. That is, the best available estimation of the vote probabilities across parties in each municipality. We created an artificial survey by selecting for each municipality a subsample consisting of n_i registered votes. More formally, we created a sample $\mathbf{X} = \{\mathbf{x}_1,\ldots,\mathbf{x}_{369}\}$ with $\mathbf{x}_i$ following a multivariate hypergeometric distribution with population size $\left(n'_{i,\text{jxsi}},\ldots,n'_{i,\text{cup}}\right)$ and sample size n_i. We considered two scenarios for n_i:

1. Proportional size: n_i as percentage of n'_i, ranging from 0.5% to 5%.
2. Constant size: n_i constant for all municipalities, with n_i ranging from 10 to 200.

In both cases, we repeated the experiment five times for each value of n_i . Given a data set $\mathbf{X}$, we compared estimates $\hat{\mathbf{p}}_{i,\text{DM}}$ and $\hat{\mathbf{p}}_{i,\text{LNM}}$ with the gold standard $\mathbf{p}_i$. Because data did not follow any particular distribution, we used three different criteria in this comparative analysis as proposed in Palarea-Albaladejo and MartÃn-Fernández (2008):

- Average of Aitchison distances: $\frac{1}{369}\sum_{i=1}^{369}\text{dist}_{\mathcal{A}}(\mathbf{p}_i,\hat{\mathbf{p}}_{i\cdot})$,
- Frobenius distance between the covariance matrix, $\mathbf{\Sigma_p} = \left(\sigma_{ij}^{\mathbf{p}}\right) \in \mathbb{R}^{5\times5}$, obtained from $\{\text{ilr}(\mathbf{p}_1),\ldots,\text{ilr}(\mathbf{p}_{369})\}$ and the covariance matrix, $\mathbf{\Sigma}_{\hat{\mathbf{p}}\cdot} = \left(\sigma_{ij}^{\hat{\mathbf{p}}\cdot}\right) \in \mathbb{R}^{5\times5}$, obtained from $\{\text{ilr}(\hat{\mathbf{p}}_{1\cdot}),\ldots,\text{ilr}(\hat{\mathbf{p}}_{369\cdot})\}$, i.e.

$$\left\|\mathbf{\Sigma_p} - \mathbf{\Sigma}_{\hat{\mathbf{p}}\cdot}\right\|_{\mathbf{F}} = \sqrt{\sum_{i=1}^{5}\sum_{j=1}^{5}(\sigma_{ij}^{\mathbf{p}} - \sigma_{ij}^{\hat{\mathbf{p}}\cdot})^2},$$

  and
- STRESS (standardised residual sum of squares) index given by

$$STRESS = \sqrt{\frac{\sum_{i=1}^{369}\sum_{j=1}^{369}\left(\text{d}_{\mathcal{A}}(\mathbf{p}_i,\mathbf{p}_j) - \text{d}_{\mathcal{A}}(\hat{\mathbf{p}}_{i\cdot},\hat{\mathbf{p}}_{j\cdot})\right)^2}{\sum_{i=1}^{369}\sum_{j=1}^{369}\text{d}_{\mathcal{A}}(\mathbf{p}_i,\mathbf{p}_j)^2}}.$$

Figure 5 shows the results for different sample sizes. The values of the three measures decreased when the size n_i increased in all cases (Fig. 5). A parallelism between the results for DM and LNM compound distributions is observed. Noticeably, the performance of DM was worse than LNM in all the cases. The alternative starting points, either SP1 or SP2, showed similar behaviour, specially using the average Aitchison dis-

tance criterium. Importantly, for a similar sample size, the results for the second scenario (constant size) were better than those for the first scenario (proportional size).

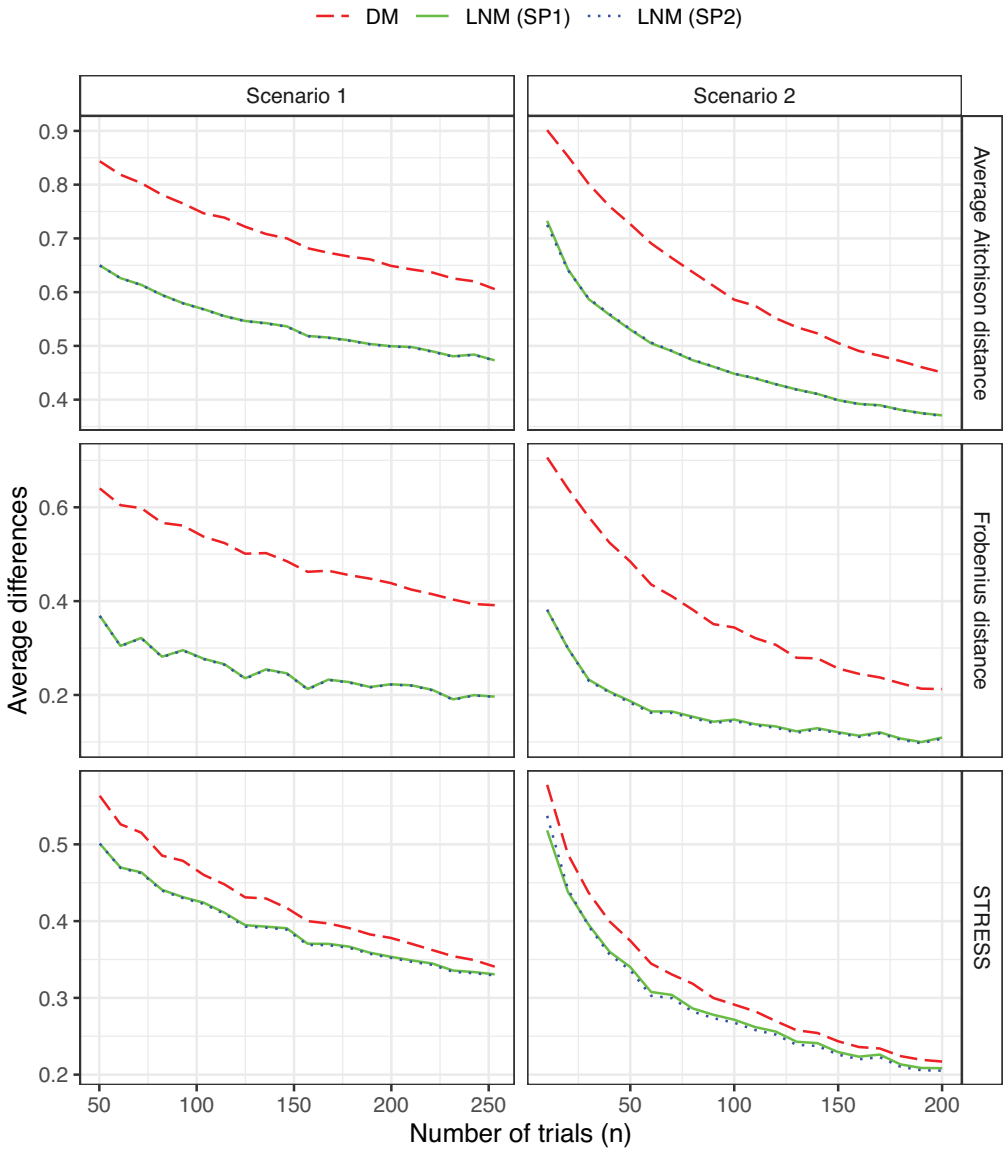

Figure 5: *Performance measures for the 2015 Catalan election example: average of Aitchison distance (top), Frobenius distance (centre), STRESS index (bottom). Estimates obtained using models DM (dashed red line), LNM (starting point SP1, dotted blue line) and LNM (starting point SP2, solid green line). Two scenarios with different sample sizes (see text for details).*

4.4 Comparison of computing times

Table 2: *Mean computing time in seconds for the three examples shown in Section 4 (95% confidence interval in parenthesis). Calculations performed on an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz. Times marked with* [1] *did not iterated trough the EM algorithm because after initialisation* Σ_0 *was singular.*

	DM	LNM (SP1)	LNM (SP2)
Pure multinomial process example			
Scenario 1 (D=9)	11.6 [4.2, 23.7]	152.6 [133.7, 172.9]	22.2 [9.6, 31.9][1]
Scenario 2 (D=12)	16.1 [3.1, 28.6]	194.7 [169.2, 223.8]	23.1 [13.6, 39.2][1]
Scenario 3 (D=15)	20.0 [10.7, 35.0]	231.3 [189.8, 310.7]	25.4 [19.1, 43.3][1]
Scenario 4 (D=16)	22.6 [6.5, 39.3]	255.9 [156.8, 334.6]	28.7 [13.6, 51.1][1]
Scenario 5 (D=20)	24.8 [11.5, 35.4]	145.3 [2.2, 310.6]	35.3 [18.8, 52.3][1]
Scenario 6 (D=25)	22.2 [12.6, 32.2]	218.2 [1.8, 535.5]	33.0 [23.7, 476.5][1]
Scenario 7 (D=30)	34.3 [9.2, 42.9]	2.4 [2.4, 2.7][1]	41.6 [10.9, 51.9][1]
Scenario 8 (D=36)	27.8 [11.2, 53.4]	3.5 [3.4, 3.6][1]	44.2 [14.9, 65.7][1]
Scenario 9 (D=50)	58.0 [11.1, 80.3]	7.0 [6.9, 9.0][1]	77.1 [45.2, 92.6][1]
Hardy-Weinberg equilibrium example			
Scenario 1 (D=3)	2.3 [1.8, 2.5]	219.2 [152.6, 286.9]	216.7 [153.8, 263.4]
Scenario 2 (D=3)	1.8 [1.6, 1.9]	878.7 [735.1, 1015.3]	565.6 [484.4, 612.3]
Scenario 3 (D=3)	2.3 [2.1, 2.6]	1319.0 [1208.2, 1548.5]	695.6 [593.3, 757.5]
Scenario 4 (D=3)	1.8 [1.6, 1.9]	66.3 [55.9, 92.7]	129.5 [89.5, 179.7]
Scenario 5 (D=3)	2.0 [1.9, 2.3]	160.7 [101.9, 208.3]	146.4 [116.1, 198.1]
Scenario 6 (D=3)	1.6 [1.5, 1.8]	327.6 [232.4, 431.0]	195.3 [140.7, 253.3]
Catalan parliamentary election example			
Scenario 1 (D=6)	1.5 [1.4, 1.7]	508.6 [430.4, 617.2]	430.2 [379.7, 510.0]
Scenario 2 (D=6)	0.8 [0.7, 1.3]	124.7 [84.3, 208.5]	186.6 [116.7, 272.2]

For the three examples above, the computing time spent on parameter estimation using LNM was higher than using DM (see Table 2). For LNM, both choices of starting points (SP1 and SP2) provided similar results, although SP1 tended to be slower. For the second example (subsection 4.2), the computing time was remarkably higher in scenarios 2 and 3 using LNM. Note that these scenarios were characterised for being the ones with the smaller probability in the first component. As expected, data dimensionality was the major factor affecting computing time.

5 Final remarks and conclusions

Count data are commonly generated in modern scientific areas such as text mining or genomic and microbiome studies based on next generation sequencing technologies. The DM distribution is a popular choice to model multivariate counts. However it may not be appropriate for complicated correlation structures because, amongst others, it imposes a negative correlation between every pair of multinomial categories. This might

not be realistic when analysing for example microbiome data (Mandal et al., 2015). Consequently, there is a need for models allowing more flexible dependence structures in multivariate count data.

The LNM compound probability distribution has been introduced in this work as a flexible model for multivariate count data. Rooted on the theoretical framework for compositional data modelling, the LNM model is fully compatible with the geometry of the simplex, the sample space where multinomial probabilities lay. Accordingly, multinomial probabilities can be conveniently mapped onto real space through logratio coordinates with respect to an orthonormal basis of the simplex for the purpose of parameter estimation. Importantly, results are invariant under changes of orthonormal basis. Parameter estimates for the LNM model cannot be computed analytically though. Different estimation approaches have been discussed and compared in this work. One based on a quasi-Monte Carlo EM algorithm is concluded to be preferable. This approach improves estimates obtained by Markov Chain Monte Carlo based on the Metropolis algorithm as used in previously works. Because inference is based on the EM algorithm, likelihood estimation can get stuck in some local maxima. Even though in the examples shown in this manuscript the global maximum is obtained, it is possible that different initialisations can be necessary to find it in particular cases.

In terms of modelling, we have shown that the LNM model produces better results than the DM. In particular, we have shown that in realistic cases LNM outperforms DM in its ability to model the underlying probabilities from the observed counts. It is important to remark that the number of parameters for a DM and a LNM grow linearly and quadratically respectively. So when the number of dimension is high, it is recommended to consider some parametrisation for the covariance matrix (Pinheiro and Bates, 1996, for an example where different restrictions on the spectral decomposition are applied to Gaussian finite mixtures see (Banfield and Raftery, 1993)).

Modelling multivariate count data using the LNM provides extra flexibility for the multinomial parameter distribution. In addition, it opens up the possibility of defining new statistical inference tools for compositional data analysis. Areas for future development include improved procedures for obtaining fast and reliable maximum likelihood estimates, e.g. along the lines of recent work by (Silverman et al., 2019). These and other questions in relation to the proposed LNM model will be addressed in future work.

References

Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman and Hall, London (UK). Reprinted in 2003 by Blackburn Press.

Aitchison, J. and Shen, S. M. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika*, 67, 261–272.

Aitchison, J. and Ho, C. H. (1989). The multivariate Poisson-Log Normal Distribution. *Biometrika*, 76, 643–653.

Banfield, J. and Raftery, A. E. (1993). Model-based Gaussian and Non-Gaussian Clustering. *Biometrics*, 49, 803–821.

Billheimer, D., Guttorp, P. and Fagan, W. F. (2001). Statistical Interpretation of Species Composition. *Journal of the American Statistical Association*, 96, 1205–1214.

Blei, D. M., and Lafferty, J. D. (2007). A Correlated Topic Model of Science. *The Annuals of Applied Statistics*, 1, 1–21.

Bouguila, N. (2008). Clustering of Count Data Using Generalized Dirichlet Multinomial Distributions. *IEEE Transactions on Knowledge and Data Engineering*, 20, 462–474.

Caflisch, R. E. (1998). Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7, 1–49.

Chastin, S. F., Palarea-Albaladejo, J., Dontje, M. L. and Skelton, D. A. (2015). Combined effects of time spent in physical activity, sedentary behaviours and sleep on obesity and cardiometabolic health markers: A novel compositional data analysis approach. *PLoS ONE*, 10, e0139984.

Comas-Cufí, M., Martín-Fernández, J. A. and Mateu-Figueras, G. (2016). Logratio methods in mixture models for compositional data sets. *SORT*, 40, 349–374.

Comas-Cufí, M., Martín-Fernández, J. A. and Mateu-Figueras, G. (2019). Merging the components of a finite mixture using posterior probabilities. *Statistical Modelling*, 19, 1–31.

Connor, R. J. and Mosimann, J. E. (1969). Concepts of independence for proportions with a generalization of the Dirichlet Distribution. *Journal of the American Statistical Association*, 64, 194–206.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1–38.

Drmota, M. and Tichy, R. F. (1997). *Sequences, Discrepancies and Applications*. Lecture Notes in Mathematics, vol. 1651. Springer, Berlin (1997).

Edjabou, M. E., Martín-Fernández, J. A., Scheutz, C. and Astrup, T. F. (2017). Statistical analysis of solid waste composition data: Arithmetic mean, standard deviation and correlation coefficients. *Waste Management*, 69, 13–23.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G. and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35, 279–300.

Graffelman, J. (2015). Exploring Diallelic Genetic Markers: The HardyWeinberg Package *Journal of Statistical Software*, 64.

Graffelman, J. and Weir, B. S. (2016). Testing for Hardy-Weinberg equilibrium at biallelic genetic markers on the X chromosome *Heredity*, 116, 558–568.

Grantham, N. S., Guan, Y., Reich, B. J., Borer, E. T., and Gross, K. (2019). MIMIX: A Bayesian Mixed-Effects Model for Microbiome Data From Designed Experiments *Journal of the American Statistical Association*, 0, 1–16.

Holmes, I., Harris, K. and Quince, C. (2012) Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics *PLOS ONE*, 7, e30126.

Hughes, G., Munkvold, G. P. and Samita, S. (1998). Application of the logistic-normal-binomial distribution to the analysis of Eutypa dieback disease incidence. *International Journal of Pest Management*, 44, 35–42.

Jank, W. (2005) Quasi-Monte Carlo sampling to improve the effciency of Monte Carlo EM *Computational Statistics & Data Anaylsis*, 48, 685–701.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions* Series in probability and statistics. John Wiley & Sons, Inc, New York (1997).

Kuo, F. Y., Dunsmuir,W. T. M., Sloan, I. H., Wand, M. P. and Womersley, R. S. (2008). Quasi-Monte Carlo for Highly Structured Generalised Response Models. *Methodology and Computing in Applied Probability*, 10, 239–275.

Layton, D. F. and Siikamäki, J. (2009). Payments for ecosystem services programs: predicting landowner enrollment and opportunity cost using a beta-binomial model *Environmental and Resource Economics*, 44, 415–439.

L'Ecuyer, P. and Lemieux, C. (2002). Recent advances in randomized quasi-Monte Carlo methods. In *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, 419–474. Kluwer Academic Publishers.

Leobacher, G. and Pillichshammer, F. (2014) *Introduction to Quasi-Monte Carlo Integration and Applications*. Compact Textbooks in Mathematics, Springer International Publishing.

Lindsay, B. G. (1995). *Mixture Models: Theory, Geometry and Applications*. Haywood, CA: Institute of Mathematical Sciences; Alexandria VA: American Statistical Association

Mandal, S., Van Treuren, W., White, R., Eggesbo, M., Knight, R. and Peddada, S. (2015) Analysis of composition of microbiomes: a novel method for studying microbial composition *Microbial Ecology in Health and Disease*, 26, 27663.

Martín-Fernández, J. A., Hron, K., Templ, M., Filzmoser, P. and Palarea-Albaladejo, J. (2015) Bayesian-multiplicative treatment of count zeros in compositional data sets *Statistical Modelling*, 15, 134–158.

Mateu-Figueras, G., Pawlowsky-Glahn, V. and Egozcue, J. J. (2011). The principle of working on coordinates. In *Compositional Data Analysis*, 29–42. John Wiley & Sons, Ltd.

Mateu-Figueras, G., Pawlowsky-Glahn, V. and Egozcue, J. J. (2013). The normal distribution in some constrained sample spaces. *SORT*, 37, 29–56.

Minka, T. P. (2004). The Dirichlet-tree distribution https://www.microsoft.com/en-us/research/publication/dirichlet-tree-distribution (last access December/2019)

Morokoff, W. J. and Caflisch, R. E. (1995). Quasi-Monte Carlo Integration. *Journal of Computational Physics*, 122, 218–230.

Mosimann, J. E. (1962). On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika*, 49, 65–82.

Neal, R. M. (2010). *MCMC Using Hamiltonian Dynamics*. Handbook of Markov Chain Monte Carlo, 54, 113–162.

Neath, R. C. (2013). On convergence Properties of the Monte Carlo EM Algorithm. In *Advances in Modern Statistical Theory and Applications: A Festschrift in honor of Morris L. Eaton*, 43–62. Institute of Mathematical Statistics.

Nelson, J. F. (1985). Multivariate Gamma-Poisson Models. *Journal of the American Statistical Association*, 80, 828–834.

Ongaro, A. and Migliorati, S. (2013). A generalization of the Dirichlet distribution. *Journal of Multivariate Analysis*, 114, 412–426.

Owen, A. B. (1995) Randomly permuted (t,m,s)-nets and (t,s)-sequences In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, 299–317. Springer-Verlag.

Palarea-Albaladejo, J. and Martín-Fernández, J. A. (2008). A modified EM alr-algorithm for replacing rounded zeros in compositional data sets. *Computers & Geosciences*, 34, 902–917.

Palarea-Albaladejo J., Rooke J. A., Nevison, I. M. and Dewhurst, R. J. (2017). Compositional mixed modeling of methane emissions and ruminal volatile fatty acids from individual cattle and multiple experiments. *Journal of Animal Science*, 95, 2467–2480.

Pan, J. and Thompson, R. (2007). Quasi-Monte Carlo estimation in generalized linear mixed models. *Computational Statistics and Data Analysis*, 51, 5765–5775.

Pawlowsky-Glahn V and Egozcue JJ (2001). Geometric approach to statistical analysis on the simplex. *Stochastic Environmental Research and Risk Assessment*, 15, 384–398.

Pinheiro, JC and Bates, DM (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing*, 6, 289–296.

R Development Core Team (2015). R: A language and environment for statistical computing. *Vienna: R Foundation for Statistical Computing*, URL http://www.r-project.org (last accessed on 23 November 2017).

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *The Annals of Mathematical Statistics*, 35, 1–20.

Robbins, H. (1980). Estimation and prediction for mixtures of the exponential distribution. *Proceedings of the National Academy of Sciences*, 77, 2382–2383.

Scheffé, H. (1958) Experiments with mixtures. *Journal of the Royal Statistical Society, series B (Methodological), 20, 344–360.*

Silverman, J. D., Durand, H. K., Bloom, R. J., Mukherjee, S. and David, L. A. (2018) Dynamic linear models guide design and analysis of microbiota studies within artificial human guts. *Microbiome, 6, 6–202.*

Silverman, J. D., Roche, K., Zachary, C. H., David, L. A. and Mukherjee, S. (2019) Bayesian Multinomial Logistic Normal Models through Marginally Latent Matrix-T Processes. *arXiv: 1903.11695.*

Wang, X. and Fang K. T. (2003). The effective dimension and quasi-Monte Carlo integration *Journal of Complexity*, 19, 101–124.

Xia, M., Chen, J., Fung K. F. and Li H. (2013). A Logistic Normal Multinomial Regression Model for Microbiome Compositional Data Analysis. *Biometrics*, 69, 1053–1063.

A Proof of properties 1 and 2

Property 1 *For a fixed* $\boldsymbol{x}$ *we have*

$$\lim_{\|\boldsymbol{\Sigma}\|\to 0} \mathcal{LNM}(\boldsymbol{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mathcal{M}\left(\boldsymbol{x};n,\mathrm{ilr}^{-1}(\boldsymbol{\mu})\right).^{2}$$

Proof. Let $\mathbf{h}$ be a real vector defined on $\mathbb{R}^{D-1}$ and let $\mathbf{x}$ be a count vector defined on $\mathcal{S}^{n,D}$. Because $\mathrm{ilr}_1^{-1}(\mathbf{h})^{x_1}\cdots\mathrm{ilr}_D^{-1}(\mathbf{h})^{x_D} \leq {x_1}^{x_1}\cdots{x_D}^{x_D}$, for any fixed $\mathbf{x}=(x_1,\ldots,x_D)$ we have:

$$\begin{aligned}\lim_{\|\boldsymbol{\Sigma}\|\to 0} \mathcal{LNM}(\mathbf{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) &= \int_{\mathbf{h}\in\mathbb{R}^{D-1}} \frac{n!}{x_1!\cdots x_D!}\mathrm{ilr}_1^{-1}(\mathbf{h})^{x_1}\cdots\mathrm{ilr}_D^{-1}(\mathbf{h})^{x_D} \lim_{\|\boldsymbol{\Sigma}\|\to 0}\mathcal{N}(\mathbf{h};\boldsymbol{\mu},\boldsymbol{\Sigma})d\mathbf{h}\\ &= \int_{\mathbf{h}\in\mathbb{R}^{D-1}} \frac{n!}{x_1!\ldots x_D!}\mathrm{ilr}_1^{-1}(\mathbf{h})^{x_1}\cdots\mathrm{ilr}_D^{-1}(\mathbf{h})^{x_D}\delta(\mathbf{h}-\boldsymbol{\mu})\,d\mathbf{h}\\ &= \frac{n!}{x_1!\ldots x_D!}\mathrm{ilr}_1^{-1}(\boldsymbol{\mu})^{x_1}\cdots\mathrm{ilr}_D^{-1}(\boldsymbol{\mu})^{x_D} = \mathcal{M}\left(\mathbf{x};n,ilr^{-1}(\boldsymbol{\mu})\right).\end{aligned}$$

■

Property 2 *Let* $\boldsymbol{x}=(x_1,\ldots,x_D)$ *and* $x_1+\cdots+x_D=n$. *If* $\lim_{n\to\infty}\frac{x_i}{n}=\pi_i$ *and* $\pi_i>0$ *for* $1\leq i\leq D$, *then*

$$\lim_{n\to\infty} n^{D-1}\cdot\mathcal{LNM}(\boldsymbol{x};n,\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mathcal{N}_{\mathcal{S}^D}(\boldsymbol{\pi};\boldsymbol{\mu},\boldsymbol{\Sigma})\frac{1}{\sqrt{D}}\frac{1}{\pi_1\ldots\pi_D}$$

Proof. Let $\boldsymbol{\pi}_n=(\pi_{n,1},\ldots,\pi_{n,D})=\frac{1}{n}(x_1,\ldots,x_D)$. We have that

$$\lim_{n\to\infty} n^{D-1}\mathcal{LNM}(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \lim_{n\to\infty}\frac{(n+D-1)!}{n!}\mathcal{LNM}(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}). \tag{11}$$

Substituting $\mathbf{x}$ by $n\boldsymbol{\pi}_n$ and using the LNM probability mass function (5), we can rewrite (11) as

$$\int_{\mathbf{h}\in\mathbb{R}^{D-1}}\mathcal{N}(\mathbf{h};\boldsymbol{\mu},\boldsymbol{\Sigma})\lim_{n\to\infty}\frac{(n+D-1)!}{n!}\frac{n!}{(n\pi_{n,1})!\ldots(n\pi_{n,D})!}\mathrm{ilr}_1^{-1}(\mathbf{h})^{n\pi_{n,1}}\ldots\mathrm{ilr}_D^{-1}(\mathbf{h})^{n\pi_{n,D}}d\mathbf{h} \tag{12}$$

Considering the change of variable given by (3), which has the Jacobian

$$d\mathbf{h} = \frac{1}{\sqrt{D}\,p_1\ldots p_D}d\mathbf{p},$$

2. $\lim_{\|\boldsymbol{\Sigma}\|\to 0}$ stands for any sequence of covariance matrices such that their highest eigenvalue goes to 0.

we can express (12) with respect to $\mathbf{p}$

$$\int_{\mathbf{p}\in\mathcal{S}^D\subset\mathbb{R}^D} \mathcal{N}(\text{ilr}(\mathbf{p});\boldsymbol{\mu},\boldsymbol{\Sigma})\frac{1}{\sqrt{D}\,p_1\dots p_D}\lim_{n\to\infty}\frac{(n+D-1)!}{(n\pi_{n,1})!\dots(n\pi_{n,D})!}{p_1}^{n\pi_{n,1}}\dots{p_D}^{n\pi_{n,D}}d\mathbf{p}, \tag{13}$$

where $d\mathbf{p}$ is measured using Lebesgue measure.

Note that for the Dirichlet distribution with parameters $\alpha_i = n\pi_{n,i}+1$, $1\le i\le D$, we have

$$\int_{\mathbf{p}\in\mathcal{S}^D\subset\mathbb{R}^D}\frac{(n+D-1)!}{(n\pi_{n,1})\dots!(n\pi_{n,D})!}{p_1}^{n\pi_{n,1}}\dots{p_D}^{n\pi_{n,D}}d\mathbf{p}=1, \tag{14}$$

and using the Stirling's approximation we have

$$\begin{aligned}\varphi(\mathbf{p}) &= \lim_{n\to\infty}\frac{(n+D-1)!}{(n\pi_{n,1})!\dots(n\pi_{n,D})!}{p_1}^{n\pi_{n,1}}\dots{p_D}^{n\pi_{n,D}}\\ &= \lim_{n\to\infty}\frac{(n+D-1)!}{(2\pi n)^{\frac{D-1}{2}}\sqrt{\pi_{n,1}\dots\pi_{n,D}}}\frac{{p_1}^{n\pi_{n,1}}\dots{p_{n,D}}^{n\pi_D}}{\pi_1^{n\pi_{n,1}}\dots\pi_D^{n\pi_{n,D}}}.\end{aligned}$$

Moreover, it can be seen that $\mathbf{p}=\boldsymbol{\pi}_n$ is a global maximum for ${p_1}^{\pi_{n,1}}\dots{p_D}^{\pi_{n,D}}$ when $p_1+\dots+p_D=1$. Moreover, because $\lim_{n\to\infty}\boldsymbol{\pi}_n=\boldsymbol{\pi}$, we have that

$$\begin{aligned}\varphi(\mathbf{p}) &= \lim_{n\to\infty}\frac{(n+D-1)!}{(2\pi n)^{\frac{D-1}{2}}\sqrt{\pi_{n,1}\dots\pi_{n,D}}}\left(\frac{{p_1}^{\pi_{n,1}}\dots{p_D}^{\pi_{n,D}}}{{\pi_{n,1}}^{\pi_{n,1}}\dots{\pi_{n,D}}^{\pi_{n,D}}}\right)^n =\\ &= \begin{cases}\infty & \text{when } \mathbf{p}=\boldsymbol{\pi}\\ 0 & \text{otherwise,}\end{cases}\end{aligned}$$

which implies, together with Equation 14, that φ is the Dirac delta function centred at $\boldsymbol{\pi}$.

Putting all together, (13) can be rewritten as

$$\begin{aligned}&\int_{\mathbf{p}\in\mathcal{S}^D\subset\mathbb{R}^D}\mathcal{N}(\text{ilr}(\mathbf{p});\boldsymbol{\mu},\boldsymbol{\Sigma})\frac{1}{\sqrt{D}\,p_1\dots p_D}\delta(\mathbf{p}-\boldsymbol{\pi})\,d\mathbf{p}=\\ &\mathcal{N}(\text{ilr}(\boldsymbol{\pi});\boldsymbol{\mu},\boldsymbol{\Sigma})\frac{1}{\sqrt{D}\,\pi_1\dots\pi_D}=\mathcal{N}_{\mathcal{S}^D}(\boldsymbol{\pi};\boldsymbol{\mu},\boldsymbol{\Sigma})\frac{1}{\sqrt{D}\,\pi_1\dots\pi_D}.\end{aligned}$$

■

Observe that we obtain the logratio-normal distribution on the simplex expressed with respect to the Lebesgue measure. The term $1/(\sqrt{D}\,\pi_1\dots\pi_D)$ relates the Aitchison measure with the Lebesgue measure (Mateu-Figueras et al., 2013).

B E-step convergence in a simple univariate but extreme case

Let $\mathbf{x} = (1,0)$ be an observed vector of counts, with mean $\mu = 0$ and standard deviation $\sigma = 1$. Consider the $\mathcal{S}^2$ basis given by $\mathcal{B} = \left\{\mathbf{b}_1 = \frac{1}{e^{\sqrt{1/2}}+e^{-\sqrt{1/2}}}\left(e^{\sqrt{1/2}}, e^{-\sqrt{1/2}}\right)\right\}$. The coordinates of an element $\mathbf{p} = (p_1, p_2)$ of $\mathcal{S}^2$ with respect to the basis $\mathcal{B}$, are

$$h = \text{ilr}(\mathbf{p}) = \sqrt{\frac{1}{2}} \ln \frac{p_1}{p_2},$$

and an element of $\mathcal{S}^2$ with respect to its coordinates is

$$\mathbf{p} = \text{ilr}^{-1}(h) = h \odot \mathbf{b}_1 = \frac{1}{e^{h\sqrt{1/2}} + e^{-h\sqrt{1/2}}}\left(e^{h\sqrt{1/2}}, e^{-h\sqrt{1/2}}\right).$$

Using Equation 5 we calculate the marginal probability

$$\Pr(\{\mathbf{X} = (1,0)\}; \mu = 0, \sigma = 1) = \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{h^2}{2}} \left(\frac{e^{h\sqrt{1/2}}}{e^{h\sqrt{1/2}} + e^{-h\sqrt{1/2}}}\right) dh \approx 0.50.$$

Using numerical integration we obtain an approximation of these expected values:

$$\mathbb{E}_{H|X;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}[h] = \int_{-\infty}^{\infty} h \frac{\left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{h^2}{2}} \left(\frac{e^{h\sqrt{1/2}}}{e^{h\sqrt{1/2}}+e^{-h\sqrt{1/2}}}\right)}{\Pr(\{\mathbf{X} = (1,0)\}; \mu = 0, \sigma = 1)} dh \approx 0.5136.$$

$$\mathbb{E}_{H|X;\boldsymbol{\mu}^{(s)},\boldsymbol{\Sigma}^{(s)}}\left[h^2\right] = \int_{-\infty}^{\infty} h^2 \frac{\left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{h^2}{2}} \left(\frac{e^{h\sqrt{1/2}}}{e^{h\sqrt{1/2}}+e^{-h\sqrt{1/2}}}\right)}{\Pr(\{\mathbf{X} = (1,0)\}; \mu = 0, \sigma = 1)} dh \approx 1.0.$$

After performing a fixed number of simulations, we compare the estimate and the variance of the error when approximating the expected values $\mathbb{E}_{h|\mathbf{x},\mu,\sigma}(h)$ and $\mathbb{E}_{h|\mathbf{x},\mu,\sigma}\left(h^2\right)$ using five different Monte Carlo approaches: MC method via importance sampling as described in Section 3.1, MC method via importance sampling and antithetic variates (Caflisch, 1998), QMC method using Halton low-discrepancy sequences, MCMC method based on the Metropolis algorithm with a standardised gaussian proposal (Xia et al., 2013), and MCMC method based on the Hamiltonian algorithm (Chapter 5, Neal (2010)). For QMC estimation, the variability was estimated using *scrambling* techniques (see Owen (1995); L'Ecuyer and Lemieux (2002) for further details). Importance sampling for MC and QMC was conducted using $m = \mathbb{E}_{h|\mathbf{x},\mu,\sigma}(h)$ and $s = 1$. MCMC methods were initiated at $h_0 = \mathbb{E}_{h|\mathbf{x},\mu,\sigma}(h)$.

Table 3: *Mean and standard deviation (in parenthesis) of 500 different approximations to* $\mathbb{E}_{h|\boldsymbol{x},\mu,\sigma}(h)$ *and* $\mathbb{E}_{h|\boldsymbol{x},\mu,\sigma}\left(h^2\right)$ *when* $\boldsymbol{x}=(1,0)$, $\mu=0$ *and* $\sigma=1$. *Computing time is shown with respect to the MC method.*

Method	First moment	Second moment	Time
Numerical approximation	0.5135884	1.0000000	
MC	0.5136272 (0.02412)	1.0008430 (0.03998)	×1.00
MC (Antithetic variates)	0.5136542 (0.00148)	1.0017754 (0.04047)	×0.96
QMC	0.5135818 (0.00071)	0.9999877 (0.00171)	×2.03
MCMC (Metropolis)	0.5133551 (0.02928)	0.9996497 (0.04716)	×6.70
MCMC (Hamiltonian)	0.5163096 (0.04394)	1.0037255 (0.07058)	×90.12

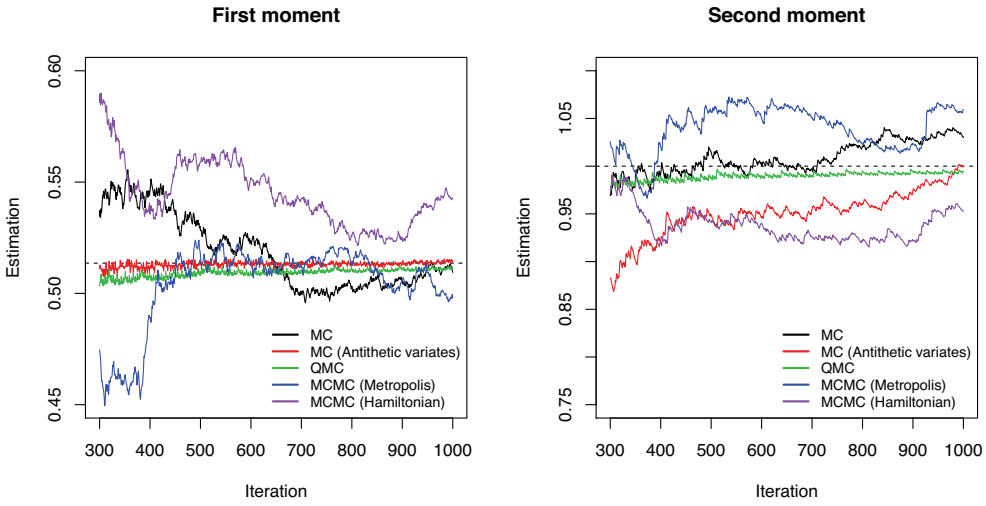

Figure 6: *Approximation of moments:* $\mathbb{E}_{h|\boldsymbol{x},\mu,\sigma}(h)$ *(left);* $\mathbb{E}_{h|\boldsymbol{x},\mu,\sigma}\left(h^2\right)$ *(right). For* $\boldsymbol{x}=(1,0)$, $\mu=0$ *and* $\sigma=1$ *using different MC techniques to approximate the E step: Monte Carlo (MC) via importance sampling (black), MC via importance sampling and antithetic variates (red), QMC using Halton sequences (green), MCMC based on the Metropolis algorithm (blue) and MCMC based on the Hamiltonian algorithm (purple). Horizontal dashed line (black) is approximation calculated by numerical integration.*

Figure 6 shows the behaviour of the methods for the first 1000 iterations in one simple approximation. The horizontal dashed line (in black) represents the expected values calculated by numerical integration. This exercise was repeated 500 times. Table 3 shows the mean and the standard deviation of the corresponding 500 approximations obtained by each procedure in the first 1000 iterations. In addition, a comparison of the computing time was conducted. The computing time was very similar for all methods, except for MCMC methods. The best results produced by the classical MC method. Regarding to the approximation of the first and second moment, QMC estimation clearly outperformes the other approaches. Remarkably, the MCMC algorithms has the worst performance. The standard deviations obtained by the Metropolis algorithm (0.044 and 0.071) were the largest and far from the standard deviations obtained by QMC. Figure 6

illustrates this behaviour. Note that the lines for the methods whose standard deviation was close to zero are close to the horizontal line representing the exact values.

C E-step convergence in multivariate cases

To evaluate the performance of the estimation procedures, we set up a simulation study parametrised by the following five parameters:

- Dimension of the random vector $\mathbf{H}$. We considered dimension $d \in \{1,5,25,125\}$.
- Multinomial sample size, $n \in \{10,100\}$.
- Location of parameter $\boldsymbol{\mu}$. Concretely, we parameterised the Aitchison norm for the mean of the multivariate normal (MVN) distribution, $\lambda \in \{0,1,2\}$.
- Variability of parameter $\boldsymbol{\Sigma}$. To this end, we parameterised the quotient between the trace and the dimension of the covariance matrix of the MVN distribution, $\nu \in \{0.5,1,2\}$.
- Agreement between count $\mathbf{x}$ and parameter $\boldsymbol{\mu}$. We considered two scenarios, a first scenario were count $\mathbf{x}$ was generated by a multinomial distribution with parameter $\boldsymbol{\pi} = \text{ilr}(\boldsymbol{\mu})$, and a second scenario were count $\mathbf{x}$ was generated by a multinomial distribution with parameter $\boldsymbol{\pi} = \text{ilr}(-\boldsymbol{\mu})$. We parameterised the two scenarios with a parameter $\xi \in \{0,1\}$ to modelate the two multinomial distributions with $\boldsymbol{\pi} = \text{ilr}^{-1}((2\xi-1)\boldsymbol{\mu})$. Parameter ξ measures the change from a situation with disagreement between $\mathbf{x}$ and $\boldsymbol{\mu}$ to a situation with agreement between them.

In each of the previous 144 scenarios we repeated the following simulation 100 times:

1. A vector $\boldsymbol{\mu} \in \mathbb{R}^d$ was uniformly generated from the d-sphere with radius λ, i.e. $\{\boldsymbol{\mu} \in \mathbb{R}^d;\ \|\boldsymbol{\mu}\|_2 = \lambda\}$.
2. A covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}_{d\times d}$ was generated as $\boldsymbol{\Sigma} = \frac{\nu}{\text{tr}(\mathbf{A})/d}\mathbf{A}$, where $\mathbf{A} \sim \text{Wishart}(d,\mathbf{I}_d)$ (ensuring that $\frac{\text{tr}(\boldsymbol{\Sigma})}{d} = \nu$).
3. A vector $\mathbf{X}$ was generated following a multinomial distribution with sample size n and probability $\boldsymbol{\pi} = \text{ilr}^{-1}((2\xi-1)\boldsymbol{\mu})$.
4. We approximated the first and second moment of the random variable $\mathbf{H}$ conditional to $\mathbf{X}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using Monte Carlo, Monte Carlo with antithetic variate, Quasi Monte Carlo and MCMC (using the Metropolis-Hastings algorithm). Approximations were conducted generating 100 random variables.

In each scenario, for the first and second moment, gold standards were obtained with 100000 replicates using standard Monte Carlo integration and MCMC. We evaluated the accuracy of each method by calculating the infinity norm of the difference between the approximation and the gold standard.

Table 4: *Results obtained after adjusting a linear model to the logarithm of the error.*

	Dependent variable: ln(error)	
	First moment	Second moment
	Effect (95% CI)	Effect (95% CI)
MC (Reference)	1	1
MC-AV	0.404 (0.378, 0.429)	0.888 (0.861, 0.915)
QMC	0.443 (0.417, 0.469)	0.612 (0.585, 0.639)
MCMC	2.182 (2.156, 2.208)	2.014 (1.987, 2.041)
d	1.020 (1.020, 1.020)	1.023 (1.023, 1.024)
n	0.994 (0.994, 0.994)	0.989 (0.989, 0.989)
λ	1.156 (1.145, 1.167)	1.532 (1.521, 1.544)
ν	1.583 (1.569, 1.598)	1.851 (1.835, 1.866)
ξ	1.056 (1.037, 1.074)	1.176 (1.157, 1.195)
R^2	0.574	0.600

To assess the results we fitted a linear model to investigate differences in logarithmic error with respect to the methods used. Results were further adjusted by the five parameters: d, n, λ, ν and ξ (Table 4). The table shows the relative effect of each parameter when estimating the error. On average, QMC produced 44.3% and 61.2% of the error to estimate the first and second moment respectively in comparison to standard MC. In contrast, MCMC methods doubled the error in both moments with respect to MC (2.182 and 2.014). The use of antithetic variables (MC-AV) provided the best results when approximating the first moment. In relation to the parameters, as it is expected, the higher the dimension the higher the error. We also observe the high impact of the MVN variability, ν, parameter in both moments (1.58 and 153 respectively). In minor measure, the same occurred for the norm, λ, and the disagreement, ξ. On the contrary, the higher the sample size, n, the lower the error.

To have a visual summary of the results, Figure 7 shows different boxplots of the parameters d, λ and ν. As seen in Table 4, this graphic illustrates how antithetic variates perform specially well in low dimensions when estimating the first moment. QMC method performs well in almost all scenarios, when estimating both moments.

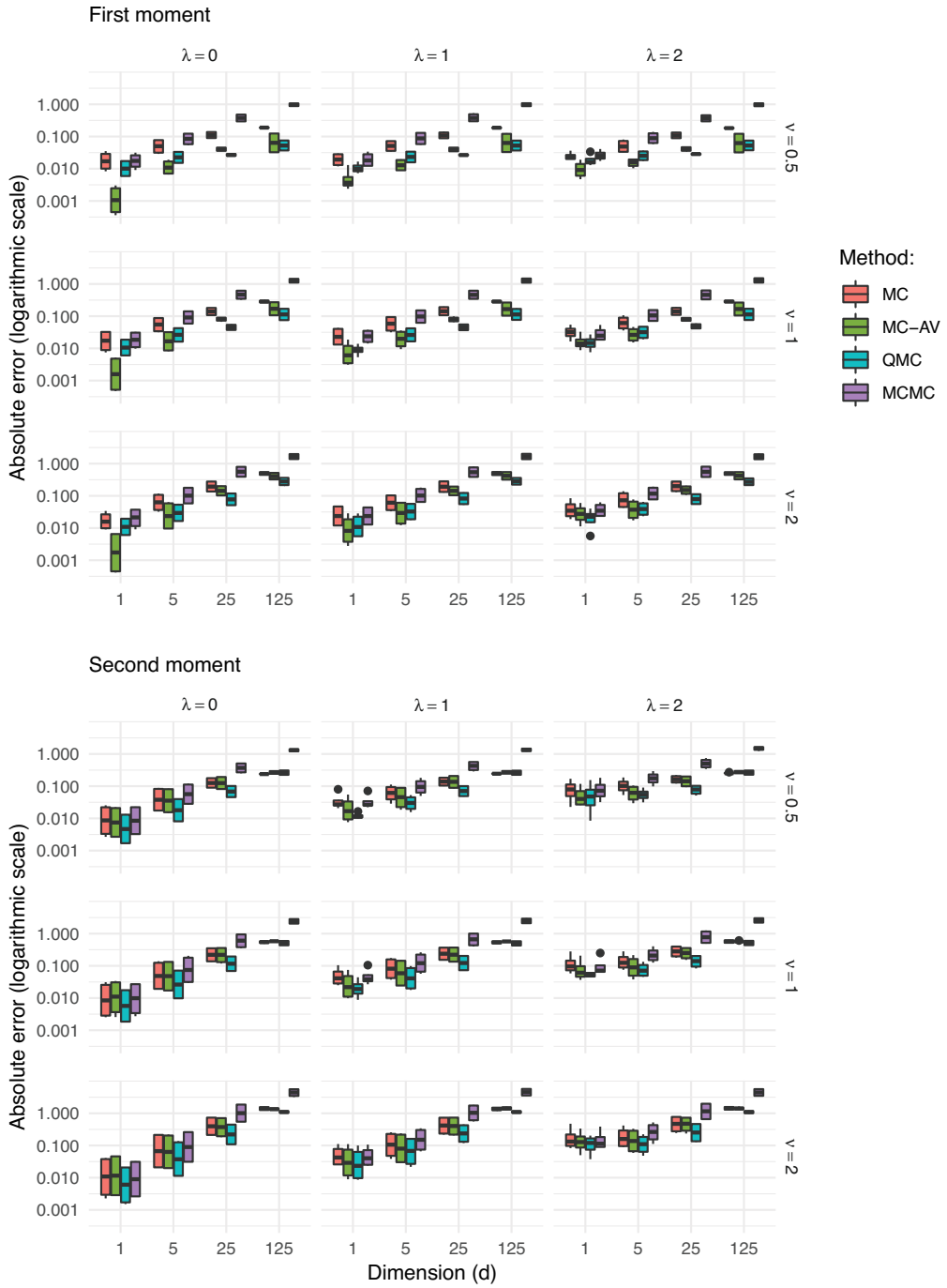

Figure 7: *Results obtained in the simulation study for the first and second moment. Results are shown for each method with respect to parameters d, λ, ν.*

D Choices for the probability vector p

Table 5: *Probability vectors* ***p*** *used in the first example. Column k shows the scenario number, column D refers to the number of components. The column on the right-hand side accounts for the initial numbers of trials.*

k	*D*	**p**
1	9	0.057 0.077 0.078 0.105 0.105 0.105 0.141 0.141 0.191
2	12	0.066 0.071 0.072 0.076 0.078 0.078 0.084 0.086 0.087 0.096 0.097 0.109
3	15	0.024 0.033 0.033 0.044 0.044 0.044 0.059 0.059 0.059 0.080 0.080 0.080 0.108 0.108 0.145
4	16	0.024 0.031 0.031 0.041 0.041 0.041 0.056 0.056 0.056 0.056 0.075 0.075 0.075 0.102 0.102 0.13
5	20	0.016 0.020 0.020 0.028 0.028 0.028 0.037 0.037 0.037 0.037 0.050 0.050 0.050 0.050 0.068 0.068 0.068 0.092 0.092 0.124
6	25	0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.02 0.02 0.02 0.02 0.02 0.02 0.02 0.02 0.02 0.02 0.05 0.05 0.05 0.05 0.50
7	30	0.021 0.023 0.023 0.025 0.025 0.026 0.028 0.028 0.028 0.028 0.030 0.030 0.031 0.031 0.031 0.033 0.034 0.034 0.034 0.034 0.037 0.038 0.038 0.038 0.042 0.042 0.042 0.047 0.047 0.052
8	36	0.019 0.019 0.020 0.020 0.021 0.021 0.021 0.021 0.022 0.022 0.023 0.023 0.024 0.024 0.024 0.024 0.024 0.026 0.026 0.027 0.027 0.028 0.028 0.029 0.029 0.030 0.032 0.032 0.033 0.033 0.037 0.037 0.038 0.043 0.043 0.050
9	50	0.02 (50 times)

SORT 44 (1) January-June 2020, 141-170 DOI: 10.2436/20.8080.02.98

Green hybrid fleets using electric vehicles: solving the heterogeneous vehicle routing problem with multiple driving ranges and loading capacities

Sara Hatami[1], Majid Eskandarpour[2], Manuel Chica[3,4], Angel A. Juan[1] and Djamila Ouelhadj[5]

Abstract

The introduction of Electric Vehicles (EVs) in modern fleets facilitates green road transportation. However, the driving ranges of EVs are limited by the duration of their batteries, which arise new operational challenges. Hybrid fleets of gas and EVs might be heterogeneous both in loading capacities as well as in driving-range capabilities, which makes the design of efficient routing plans a difficult task. In this paper, we propose a new Multi-Round Iterated Greedy (MRIG) metaheuristic to solve the Heterogeneous Vehicle Routing Problem with Multiple Driving ranges and loading capacities (HeVRPMD). MRIG uses a successive approximations method to offer the decision maker a set of alternative fleet configurations, with different distance-based costs and green levels. The numerical experiments show that MRIG is able to outperform previous works dealing with the homogeneous version of the problem, which assumes the same loading capacity for all vehicles in the fleet. The numerical experiments also confirm that the proposed MRIG approach extends previous works by solving a more realistic HeVRPMD and provides the decision-maker with fleets with higher green levels.

MSC: 90B06, 90C59, 68W20, 62P12.

Keywords: Vehicle Routing Problem, Electric Vehicles, Heterogeneous Fleet, Multiple Driving Ranges. Iterated Greedy heuristic, Successive Approximations Method

1 Introduction

Transportation is one of the main activities in modern supply chains, and accordingly, it has a significant effect on the customer level of satisfaction (Crainic, 2000). Likewise, CO_2 and greenhouse-gas emissions play an important role in producing the side

[1] IN3 – Computer Science Dept., Universitat Oberta de Catalunya, Barcelona, Spain, ajuanp@uoc.edu

[2] IÉSEG School of Management, Université de Lille, France

[3] Andalusian Research Institute DaSCI, University of Granada, Granada, Spain

[4] School of Electrical Engineering and Computing, University of Newcastle, Australia

[5] Centre for OR and Logistics, School of Math and Physics, University of Portsmouth, Portsmouth, UK

Received: June 2019
Accepted: May 2020

effects (externalities) of noise pollution, air pollution, and traffic congestion (Faulin et al., 2019). Therefore, enterprises must consider both customer satisfaction and environmental impact when planning their transportation operations. In fact, some governments are making noticeable efforts for promoting 'green' (environment-friendly) policies. One of these policies is related to shifting from Internal-Combustion-Engine Vehicles (ICEVs) to zero-emission Electric Vehicles (EVs) or, at least, to Plug-in Hybrid Electric Vehicles (PHEVs) (Mattila and Antikainen, 2011). Both ICEVs and PHEVs consume oil and produce a higher percentage of CO_2, greenhouse emissions, and other pollutant effects compared to EVs. It is then clear that a shift from a fossil fuel fleet to an electric-powered fleet is necessary to reduce pollutant emissions in cities. Also, by introducing special taxes, governments are approving policies aimed at decreasing the pollution level generated by transportation (Faulin, Lera-López and Juan, 2011). Other ways of promoting green technologies refer to offer incentives for companies to reduce carbon footprint, diminish the risk associated with the dependence on oil-based energy sources, make more affordable the acquisition of EVs, and develop alternative-energy technologies (Williams et al., 2012). Therefore, from both an environmental and energy standpoints, the use of EVs should be a first priority for the reduction of primary energy consumption.

Although EVs show many advantages regarding the use of a greener energy, this technology is currently facing some drawbacks. In particular, these vehicles make use of electronic batteries, which limit their driving-range capabilities. These batteries have long-recharge processing times and cannot be charged in classical service stations on the road (Chan et al., 2009; Wirasingha, Schofield and Emadi, 2008; Ferreira et al., 2011; Achtnicht, Bühler and Hermeling, 2012). On the contrary, the driving ranges of ICEVs and PHEVs are assumed to be unlimited as they can be easily refueled at any station along their route. With EVs becoming more prevalent among current fleets of vehicles, an efficient routing of hybrid fleets of vehicles with multiple driving-ranges is an emergent challenge in the transportation industry. Thus, the Vehicle Routing Problem with Multiple Driving ranges (VRPMD) takes into account a hybrid fleet of EVs, PHEVs, and ICEVs. However, the different vehicles of the fleet are assumed to have the same loading capacity.

In this paper, we propose a more realistic model by also considering heterogeneous fleets of vehicles in terms of loading capacity and driving ranges. The novel model is called the Heterogeneous Vehicle Routing Problem with Multiple Driving ranges (HeVRPMD). To solve this new model, we propose a Multi-Round Iterated Greedy (MRIG) metaheuristic based on a successive approximations method. The final solution for the heterogeneous case is obtained by solving a series of homogeneous cases and then combining the resulting partial solutions into a global one. In order to validate the performance of the MRIG metaheuristic, we first solve the homogeneous VRPMD and compare our results with the ones published in existing literature. Then, we extend the MRIG metaheuristic to solve the HeVRPMD. To solve this problem, we create a new set of instances based on the classical ones for the heterogeneous fleet vehicle routing

problem with unlimited driving ranges. The performance of the new metaheuristic is evaluated considering the distance-based costs and the green level of the fleet configurations. The green level of the fleet is measured by using two novel correlated green indexes. The first green index measures how green a fleet configuration is based on the fraction of 'green' vehicles in the fleet. The second green index directly computes the estimated environmental cost associated with the use of each vehicle in the fleet. The rest of the paper is organised as follows. First, we review the related literature in Section 2. Then, Section 3 provides the mathematical optimization model. The MRIG metaheuristic is described in Section 4. Section 5 presents the definition of the proposed green indexes and the computational experiments. The obtained results for the homogeneous case are analysed in Section 6, while the ones associated with the heterogeneous case are discussed in Section 7. Finally, some conclusions are provided in Section 8.

2 Literature review

The classical vehicle routing problem (VRP) and its variants have received extensive attention from practitioners and the research community (Laporte, 2009). While most VRP articles assume a fleet of homogeneous vehicles to serve the customers, in real-life it is usual to consider heterogeneous fleets of vehicles in terms of loading capacity (Koç et al., 2016; Dell'Amico et al., 2007; Dominguez et al., 2016). This section focuses on reviewing the following two streams of research: the green VRP and the VRP with a heterogeneous fleet of vehicles. The green VRP is a relatively recent variant of the VRP having the goal of reducing greenhouse gas emissions by considering alternative fuel-powered vehicle fleets, such as electricity (Erdoğan and Miller-Hooks, 2012). In addition to the reduction of the amount of gas emissions, the use of electricity imposes some restrictions, such as the limited driving-range autonomy. As a result, most of the electric VRP literature reports on alternative fuel stations to recharge the battery. In this context, Erdoğan and Miller-Hooks (2012); Schneider, Stenger and Goeke (2014) and Montoya et al. (2014) presented a green VRP considering alternative fuel stations to refill the tank or recharge the battery, while Jie et al. (2019) and Verma (2018) proposed another remedy to meet EVs' limited driving ranges . In these research papers, EVs should visit battery swapping stations to swap their batteries before their battery power runs out or their driving ranges terminate. Likewise, Lin, Zhou and Wolfson (2016) considered the vehicle load effect on the battery consumption while designing the optimal routing plan. Additionally, Keskin and Çatay (2016) included a partial recharging feature for electric VRP with time windows. This partial-recharging assumption has been included to make recharging times shorter. Also related to this research line, Felipe et al. (2014) proposed an electric VRP model which determines the amount of energy recharged and the technology used. A detailed review on the challenges of electric vehicles in logistics and transportation can be found in Juan et al. (2016).

The second stream of research addressed in this study is the heterogeneous VRP. As stated before, using a heterogeneous fleet of vehicles makes the model more realistic. In this regard, the use of hybrid fleets combining EVs, ICEVs, and PHEVs is a promising research area. Hiermann et al. (2016) and Vaz Penna et al. (2016) proposed an optimization problem combining the fleet size, mix VRP with time windows, and the use of EVs. The fleet size and mix VRP only cover conventional vehicles but they distinguish different types of vehicles according to their transportation capacity, battery size, and fixed cost. Goeke and Schneider (2015) incorporated an energy consumption model including speed, gradient, and cargo load distribution into a fleet size and mix VRP with time windows and the use of EVs. Despite the recent advances on EV-related technology and infrastructure, the development of recharging facilities throughout the road transportation networks might be only an option on the long run. Therefore, the travel range still remains as one of the main issues concerning the use of EVs in transportation activities. This issue has been addressed in Almouhanna et al. (2020), by discussing a location routing problem with constrained distance which is used EVs in the location and routing decisions. They developed a heuristic and a metaheuristic to minimize the total cost, which includes the opening cost of depots, the variable distance-based cost of vehicles, and the fixed cost of using vehicles. In Juan, Goentzel and Bektas (2014b), the authors addressed this issue by solving a VRP variant which considers a hybrid fleet of vehicles with multiple driving ranges and assume that all vehicles are homogeneous in terms of loading capacity. The use of multiple driving ranges is due to the fact that vehicles use different energy sources or even alternative battery types. These authors developed a Multi-Round Heuristic (MRH) that iteratively builds a solution for the problem. In Reyes-Rubiano et al. (2019), authors studied a VRP including homogeneous fleet of electric vehicles with a limited loading capacity and driving ranges and stochastic travel times. The authors proposed a simheuristic algorithm to design reliable routing plans in order to minimize the expected timebased cost required to complete the freight distribution plan.

A research stream related to the VRP with multiple driving ranges is the distance-constrained VRP. Few papers have studied the distance-constrained VRP. Among them, we can highlight the works of Kek, Cheu and Meng (2008), Li, Simchi-Levi and Desrochers (1992), and Laporte, Desrochers and Nobert (1984). In the context of alternative fuel-powered vehicle fleets, Erdoğan and Miller-Hooks (2012) developed a similar model to a distance-constrained VRP by considering a single driving range limitation on the tour duration. The VRP and its variants are NP-hard problems and different approaches, from exact methods to heuristics and metaheuristics, have been employed to solve them (Hokama, Miyazawa and Xavier, 2016; François et al., 2016; Andreatta et al., 2016). Soft computing based methods are very frequent when solving the VRP family of optimization problems. Authors of Brito et al. (2015) used fuzzy logic as a way of defining the constraints of the VRP optimization problem and made use of ant colony optimization as the solving strategy. Fuzzy logic was also used to define preference information of customers with respect to the satisfaction for a service time in a

multi-objective optimization problem using a solving strategy based on a genetic algorithm Ghannadpour et al. (2014). Genetic algorithms, both single- and multi-objective, have been widely used for VRPs and their variants. Karakatič and Podgorelec (2015) presented a complete review on the use of genetic algorithms for multi-depot VRPS. Recently, Pierre and Zakaria (2017) proposed a genetic algorithm with additional stochastic rules for a VRP with time windows and AbdAllah, Essam and Sarker (2017) presented an enhanced genetic algorithm that tries to increase both diversity and the capability to escape from local optima to solve a dynamic VRP in which not all customers are known in advance, but are revealed as the system progresses. The research developed in this paper extends previous work in two main directions. First, it proposes a new MRIG metaheuristic which outperforms the existing approaches for the homogeneous version of the problem – i.e., assuming that all vehicles have the same loading capacity. Secondly, it extends this MRIG metaheuristic so it can deal with the heterogeneous version of the problem too.

3 HeVRPMD optimization model

In this section, we describe the proposed HeVRPMD model, which considers a heterogeneous fleet of vehicles with respect to loading capacities as well as driving ranges. This VRP model can be seen as a combination of two distinct problems: (i) the VRP considering heterogeneous fleets of vehicles in terms of loading capacity Baldacci, Battarra and Vigo (2008); and (ii) the VRP with multiple driving ranges (VRPMD), where fleets are hybrid in terms of driving range but all vehicles are assumed to have the same loading capacity (Juan et al., 2014b). In the next subsections, we define the mathematical optimization model for the HeVRPMD. The proposed model aims to find alternative 'green' fleet configurations with minimum distance-based cost. The concept of green refers to the fact that we give priority to the use of small-size EVs over medium-size EVs, large-size ICEVs and PHEVs. In addition, as in many other vehicle routing problems, the following constraints need also to be fulfilled: (i) each route starts and ends at the depot, and it is associated with a vehicle type; (ii) each customer belongs to exactly one route; and (iii) loading capacities and driving ranges of the vehicles are never exceeded – notice that the considered vehicles are heterogeneous both in loading capacity and driving range.

A model representation of the HeVRPMD can be a directed graph $G = (N, A)$ consisting of a set N of $n+1$ nodes, $N = \{0, 1, \ldots, n\}$ and a set $A = \{(i, j) : i, j \in N, i \neq j\}$ which represents the arcs connecting pairs of nodes. Node 0 denotes the depot, where the vehicle fleets are located, and the remaining nodes represent the n customers. Each customer i has a known demand $q_i > 0$. We denote the distance-based cost associated with traveling from node i to node j by d_{ij}, with $d_{ij} = d_{ji} \geq 0$. In addition, there is a set K of k different types of vehicles, $K = \{1, 2, \ldots, k\}$. The number of vehicles for each type is assumed to be unlimited. Each vehicle of type l has a loading capacity Q^l as

well as a maximum driving range T^l. Three different decision variables are used in the model: (i) a binary decision variable x_{ij}^l, which takes the value of 1 if vehicle $l \in K$ travels from node i to j, and 0 otherwise; and (ii) two continuous decision variables u_i^l and v_i^l which represent the cumulative amount of load carried and distance traveled, respectively, by vehicle $l \in K$ when leaving customer $i \in N \backslash \{0\}$. The objective function is the minimization of total distance-based cost, subject to:

1. Satisfying all customers' demands.
2. Balancing of flows between nodes.
3. Loading capacity of vehicles.
4. Driving ranges of the vehicles.
5. Non-negativity and binary constraints.

To define the constraints we used the model provided by (Baldacci et al., 2008) to define the set of constraints related to items 1 to 3, and the model introduced by (Juan et al., 2014b) to define the set of constraints related to driving ranges of the vehicles. The objective function of the optimization model is defined in Equation 1. This function calculates the total distance-based cost of all used vehicles by adding the traveled distance by each vehicle l over all arcs $(i, j) \in A$:

$$Minimize\ z = \sum_{l \in K} \sum_{(i,j) \in A} d_{ij} x_{ij}^l \tag{1}$$

The constraints of the HeVRPMD model are defined from Equation 2 to 9. Equations 2 ensures that every customer is visited exactly once by a single vehicle:

$$\sum_{l \in K} \sum_{j \in N, i \neq j} x_{ij}^l = 1 \qquad \forall i \in N \backslash \{0\} \tag{2}$$

Equation 3 guarantees the flow conservation from and to a given customer node using a vehicle of type l. By doing so, a connection between node i and node $j \in N \backslash \{i\}$ is assured:

$$\sum_{j \in N, i \neq j} x_{ij}^l - \sum_{j \in N, i \neq j} x_{ji}^l = 0 \qquad \forall i \in N \backslash \{0\}, l \in K \tag{3}$$

Equations 4 and 5 ensure that the total load capacity of a vehicle type l on each tour does not exceed the vehicle capacity Q^l. More precisely, Equation 4 ensures that the load of the vehicle in the next node j depends on the load of the vehicle in the previous node i as well as on the demand of node j. As a result, the last node on a tour will denote the total amount of load carried by the vehicle:

$$u_i^l \leq u_j^l - q_j x_{ij}^l + Q^l (1 - x_{ij}^l) \qquad \forall l \in K, i \in N, j \in N \backslash \{0\}, i \neq j \tag{4}$$

Equation 5 ensures that load u_i^l is always greater than zero and less than the maximum capacity Q^l for a vehicle of type l:

$$0 \leq u_i^l \leq Q^l \qquad \forall l \in K, i \in N \backslash \{0\} \tag{5}$$

Constraints 6 and 7 guarantee that the total length of the route does not exceed the maximum range of vehicle l. Constraint 6 restricts the route travelled up to customer j (v_j) to be larger than the route travelled up to previous visited node i (v_i) plus the distance travelled between customer node i to customer j.

$$0 \leq v_i^l \leq v_j^l - d_{ij}x_{ij}^l + T^l(1 - x_{ij}^l) \qquad \forall l \in K, i \in N, j \in N \backslash \{0\}, i \neq j \tag{6}$$

Constraint 7 ensures that the current route travelled to be smaller than the maximum driving range of vehicle type $l \in K$ minus the route traveled between node $i \in N$ and node $j \in N$.

$$0 \leq v_i^l \leq T^l - d_{ij}x_{ij}^l \qquad \forall l \in K, \forall (i,j) \in N, i \neq j \tag{7}$$

Notice that constraints 4 to 7 in our model forbid sub-tours in the solution. In fact, similar constraints have been widely used in the VRP literature in order to eliminate sub-tours Erdoğan and Miller-Hooks (2012); Feillet (2010). Finally, Equations 8 and 9 guarantee the binary and non-negativity conditions of the decision variables:

$$x_{ij}^l \in 0,1 \qquad \forall l \in K, \forall (i,j) \in A \tag{8}$$

$$u_i^l, v_i^l \geq 0 \qquad \forall l \in K, \forall i \in N \backslash \{0\} \tag{9}$$

Even for small-scale instances of the homogeneous (simplified) version of this problem, it is not possible to obtain solutions in reasonable computing times using commercial optimization packages such as CPLEX. Therefore, in the remaining of this paper, we propose the use of a metaheuristic algorithm as the most effective way to deal with both the homogeneous and the heterogeneous versions.

4 The MRIG metaheuristic

This section describes the proposed MRIG metaheuristic to solve both the homogeneous VRPMD and its heterogeneous version. This algorithm is inspired by the successive approximations method proposed by Juan et al. (2014a) to solve the heterogeneous VRP. Accordingly, MRIG is a multi-round approach that solves the global heterogeneous VRP by dividing it into different homogeneous VRP. The main components of the algorithm are the construction of the initial solution, local improvement, and acceptance criterion. Each round of the MRIG approach consists of an optimization routing algorithm, run

inside an Iterated Greedy (IG) framework. At each round, the algorithm first selects a new subset of nodes and a new type of vehicles, thus defining a homogeneous fleet VRPMD. Then, the routing algorithm integrated into the IG framework searches for a near-optimal solution for the selected subset of nodes and vehicle type. Additionally, and in order to facilitate the generation of solutions with different fleet configurations, a penalty-based diversification mechanism is integrated within MRIG. This mechanism is applied after the construction of the initial solution and before the local improvement of the selected subset of nodes of the route. The penalty mechanism is applied for all the iterations of the algorithm and slightly modifies the initial driving range of each vehicle at random. This diversification technique has been extensively used in some metaheuristic approaches such tabu search and others. Mathematically, the penalty cost modifies the driving range of a vehicle of type l (T^l) to a new driving range value with random noise ($T^{l'}$), always with a variation below the 10% of the initial driving value of the vehicle.

The routing algorithm of the MRIG extends and enhances the popular Clarke and Wright's savings heuristic (Clarke and Wright, 1964), and the savings list of edges. One of the advantages of using this routing algorithm is that it does not require any complex parameter fine-tuning and is efficient to solve VRP, as reported in Quintero-Araujo et al. (2017); Ferone et al. (2019). Another important component of the MRIG is the heuristic. This heuristic is relatively simple, yet effective, which has obtained high-quality results in areas such as scheduling (Ruiz and Stützle, 2007), arc routing problems (González-Martín et al., 2012), and vehicle routing problems (Ruiz and Stützle, 2008; Chebbi and Chaouachi, 2015), among others. In a nutshell, it generates a sequence of promising solutions by iterating over greedy constructive heuristics using two main phases: destruction (some solution components are removed from a complete solution), and reconstruction (a greedy constructive heuristic is applied to reconstruct a complete solution). Once a candidate solution has been completed, an acceptance criterion decides whether or not the new constructed solution will replace the reference solution. Figure 1 shows a flowchart to illustrate the MRIG algorithm and its main components.

4.1 Construction of the initial candidate solution

The first step of the algorithm uses a multi-round process with a routing algorithm to obtain the initial solution. This routing algorithm is applied to each round of the process for each type of vehicle among the unused vehicles. This also means that assuming an unlimited fleet of vehicles of the same type and with the same loading capacity, a homogeneous VRPMD is solved for the nodes which are not yet served.

Thus, for example, a multi-round process will typically need three rounds to generate a global feasible solution when facing a problem with a fleet configuration composed of ICEs, PHEVs, and two types of EVs with two different battery capacities. The three rounds are associated with the unlimited range ICEs and PHEVs, medium range EVs with larger batteries, and short range EVs with smaller batteries, respectively. At each

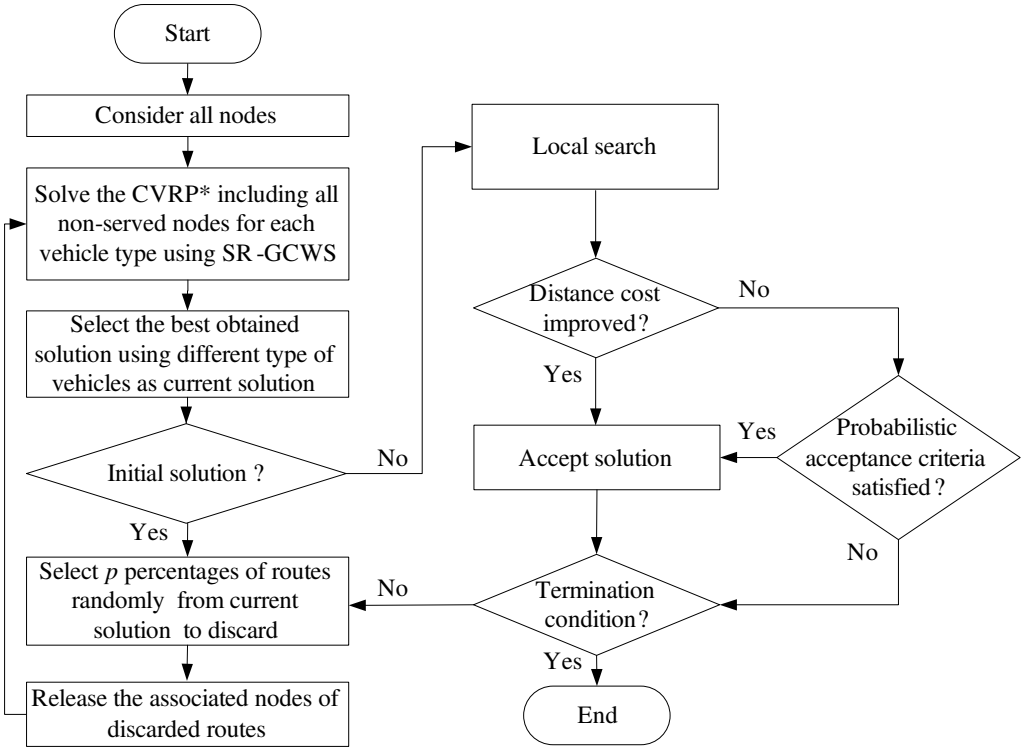

Figure 1: *Diagram of the multi-round algorithm to build the initial candidate solution.*

round of the solving method, a different homogeneous fleet capacitated VRP with a route length restrictions (CVRP*) is solved. The maximum route distance considered for each round is given by the maximum driving range of the corresponding unused vehicles. In the first round, when the problem specification includes unlimited driving-range vehicles (ICEs and PHEVs), no restriction is assumed on the route length, and the VRP is solved. In the remaining two rounds, when limited driving range vehicles are available, a CVRP* is solved considering the maximum route distance for the vehicles in this specific round (Belloso, Juan and Faulin, 2019). After solving the successive CVRP* with different types of vehicles and different driving ranges, the solution with the minimum distance-based cost is selected as the incumbent best solution. To produce the new solution, a ratio p of the routes of the current best solution are randomly selected to be discarded. The remaining $1-p$ routes are saved as partial solutions. The algorithm releases the associated nodes of the discarded routes and the same constructive process is used again to create the sub-solutions from these nodes which belong to the discarded routes. When all the nodes are served, the complete candidate solution, referred to as π_C, is built by merging the partial best sub-solutions obtained by the rounds of the algorithm. The process of the initial candidate solution construction is shown in Figure 2.

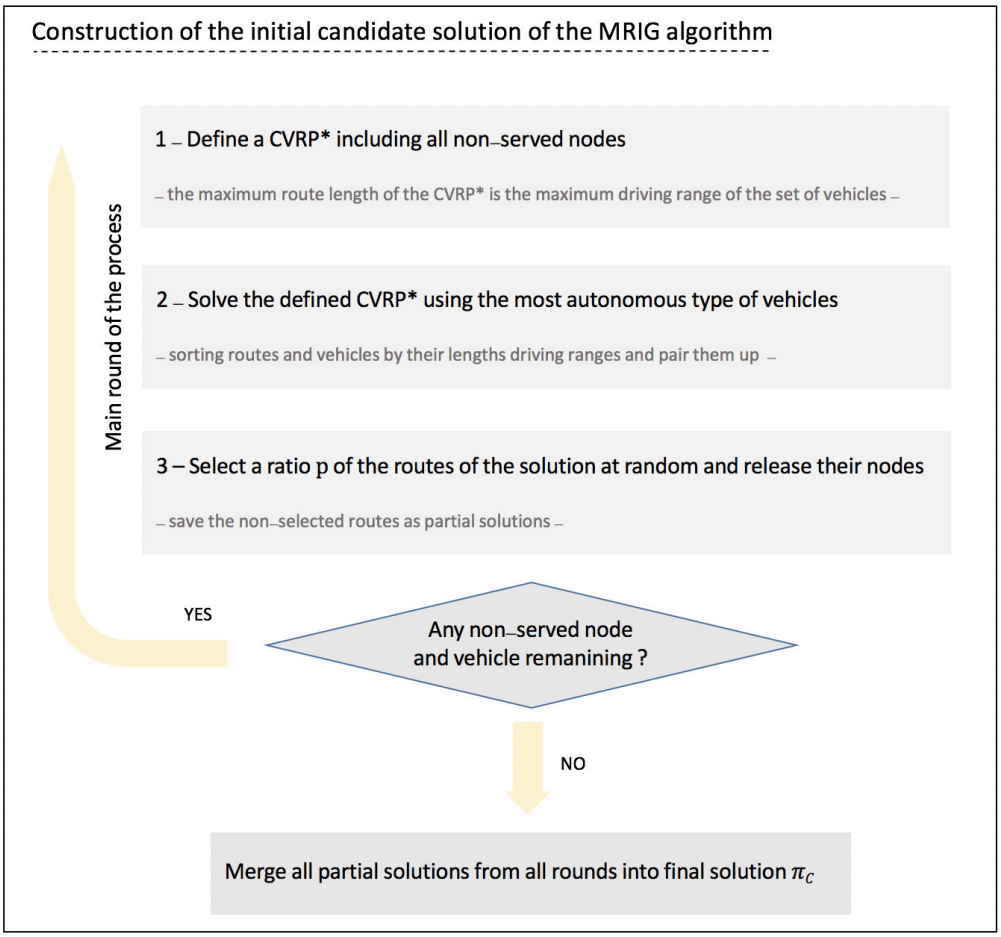

Figure 2: *Flowchart of the MRIG algorithm and its main components.*

4.2 Improvement phase

The second step of MRIG consists of two local search operators: destruction and reconstruction. The goal of this step is to improve the distance-based cost of the candidate solution, π_C. In the destruction phase, a sub-set of nodes, D, from the total n nodes is selected by using a ratio $d \in [0,1]$. This subset of nodes is removed from the routes and inserted into an archive list in the order they were selected. Note that, by following this process, there will not be any empty route (i.e., each route has at least one node). This destruction procedure explained in Algorithm 1, returns the list of removed nodes D as well as the list of routes containing the non-removed nodes. We denote τ_r to be the list of nodes assigned to route $r \in R$ of a candidate solution π_C.

During the construction phase, all the nodes of sub-set D are selected one by one according to the list order. Later, they are re-inserted into the existing routes. Among all

Algorithm 1 Destruction_$\pi_C(d)$

```
i ← 0
while i < ⌊dn⌋ do
   r ← Route randomly selected
   a ← Node randomly selected among the remaining nodes in the route r
   if |τ_r| > 1 then
      D ← Insert node a
      τ_r ← Remove node a from τ_r
      i ← i + 1
   end if
end while
return D and all τ_r, r ∈ R
```

the possible position, the chosen location for each node is the one in the route with the smallest distance-based cost. This process is repeated $\lfloor dn \rfloor$ times, until all the nodes of D are re-inserted, thus leaving D empty.

4.3 Acceptance criterion

Finally, MRIG uses an acceptance criterion that allows it to accept, from time to time, a degradation of the base solution. The criterion adds more diversity into the search and prevents the algorithm from getting stuck in a local optima. The acceptance criterion is applied once the improvement of the candidate solution, π_C, has been completed. Therefore, this step determines if the new generated solution should replace the base solution π_I even if it has a higher cost. The acceptance criterion of worse solutions is based on the probabilistic acceptance criterion of simulated annealing (Ruiz and Stützle, 2008; Yu and Lin, 2015; Wang et al., 2015). In MRIG, the acceptance criterion works as follows. Let $C_{(\pi_I)}$ denotes the distance-based cost of the current base solution π_I. The newly generated solution, π_C, is automatically accepted as the updated base solution if $C_{(\pi_C)} < C_{(\pi_I)}$. Otherwise, the solution π_C is accepted as an update of the base solution only if a certain criterion is met. This criterion relies on a probabilistic mechanism that takes into account the so-called temperature parameter and the change in the objective function value. It is defined in Equation 10, where *random* is a random number uniformly distributed between 0 and 1, and *Temp* is the temperature parameter originally proposed by Osman and Potts (1989):

$$random \leq e^{-\frac{C_{(\pi_C)} - C_{(\pi_I)}}{Temp}} \tag{10}$$

Hatami, Ruiz and Andrés-Romano (2015) simplified the latter acceptance mechanism by considering two aspects. First, they eliminated the *Temp* factor. Secondly, the probability of accepting a worse solution in the original mechanism of Equation 10 only depends on the difference between $C_{(\pi_C)}$ and $C_{(\pi_I)}$. This dependency provokes that the difference could be the same for instances with non-similar deterioration levels in terms of relative values. In order to solve this potential shortcoming, the difference, $C_{(\pi_C)} - C_{(\pi_I)}$,

is substituted by the Relative Percentage Difference (RPD) between the cost values of these two solutions. RPD value is obtained by $RPD(\pi_C, \pi_I) = \frac{C_{(\pi_C)} - C_{(\pi_I)}}{C_{(\pi_I)}} \times 100$. The improved criterion is simple and it does not need any parameter fine-tuning process. Therefore, we use it in the MRIG metaheuristic as shown in Equation 11.

$$random \leq e^{-RPD(\pi_C, \pi_I)} \tag{11}$$

5 Computational experiments

This section describes the experimental setup to evaluate the performance of MRIG. First, we describe how we evaluate the costs of the solutions (Section 5.1), as well as two new green indexes (Section 5.2). Secondly, Section 5.3 shows the benchmark instances used for the experiments.

5.1 Distance-based cost evaluation

Three types of vehicles are considered in the experiments: (i) large range ICEVs and PHEVs with no driving range limitations; (ii) medium range EVs, which have an autonomy of 200 distance units; and (iii) small range EVs, with an autonomy of 100 distance units. These three types of vehicles are denoted by L, M, and S, respectively. Accordingly, a fleet configuration for each problem instance is represented by $S/M/L$ (i.e., the number of vehicles of each type). In the computational experiments, the distance-based cost associated with the fleet configuration is computed. An analysis on how the substitution of ICEVs/PHEVs by EVs increases the distance-based cost is also provided.

5.2 Green indexes for fleet configurations

In order to compare the performance of MRIG with previous results from the literature, we have considered that one configuration is greener than another if: (i) it substitutes vehicles of type L by vehicles of type M or S, where S is always preferred over M; or (ii) vehicles of type M are substituted by vehicles of type S without increasing the number of vehicles of type L.

As mentioned before, vehicles of type S and M are EVs. It is assumed that a vehicle of type S has a lower driving range and a lower loading capacity than a vehicle of type M. A vehicle of type S can easily access high congested streets with limited parking space in many cities, and it is constrained to a lesser degree by the existence of traffic congestion or lack of parking areas than other larger-size vehicles (Juan et al., 2016). For these reasons, a vehicle of type S is considered greener than one of type M. In order to compare the green level of two different fleet configurations, we introduce two novel indexes. The first one, GI_1, is defined by Equation 12 and measures the fraction of S and M vehicles with respect to all vehicles in the fleet:

$$GI_1 = \frac{S + \omega M}{S + M + L} \tag{12}$$

where $\omega \in (0,1)$, and S, M, and L, denote the number of vehicles of each type. In the numerical experiments, we have set ω to 0.7, since the idea is that using a larger fraction of type M vehicles contributes to make the fleet greener, but not as much as using a larger fraction of vehicles of type S (considered to be the 'greenest' ones). Notice that this index will take the value of 0 whenever all the vehicles in the fleet are of type L (i.e., ICEVs or PHEVs), while it will take the value of 1 only when all the vehicles in the fleet are of type S (i.e., the greenest possible EVs). The second proposed index, GI_2, directly considers the environmental cost associated with using each type of vehicle. This index measures environmental unit cost for each fleet configuration. It is assumed that vehicles of type L have an associated cost of α monetary units. The environmental unit costs for vehicles of type M and S, which are less pollutant than type L, are set to β and γ, with $\alpha > \beta > \gamma > 0$. This index is defined by Equation 13. When applying this GI_2 index to a practical scenario, these cost values are required to be set based on additional data on the specific characteristics of each vehicle type. In our numerical experimentation we set $\alpha = 100$, $\beta = 30$, and $\gamma = 10$. These values are based on a preliminary experimentation where we tested different sets of values from the expert knowledge of a routing business collaborator. For example, in our experimental scenario, each vehicle of type L produces 10 times more environmental units than vehicles of type S; and each vehicle of type M is 3 times more pollutant than a vehicle of type S.

$$GI_2 = \gamma S + \beta M + \alpha L \tag{13}$$

These two green indexes offer alternative ways of measuring the degree of environment-friendly associated with a given fleet configuration. The idea here is that GI_1 can be used as a proxy for index GI_2 in those real-life situations in which estimating the exact values of α, β, and γ cannot be easily achieved due to the lack of accurate data. Figure 3 illustrates, for some of the solutions obtained in the numerical experiments which are discussed in the next section, the existence of a strong linear relationship between both indexes. Note that, in all the analysed instances, the determination coefficient is above 90%, which guarantees that – at least for the set of instances and numerical values considered in the experiments – GI_1 could be used to accurately estimate the value of GI_2, if necessary.

5.3 Problem instances and computing resources

We have used 33 classical VRP instances to validate our solving approach for both the VRPMD and HeVRPMD scenarios. These instances have been selected from a large set of instances available at `http://www.branchandcut.org`. The criteria we used to select these instances were to include those ones with detailed information on routes for the optimal or pseudo-optimal solution and having between 22 and 135 nodes. The charac-

teristics of these instances are also different among them (e.g., the number of nodes, the vehicle capacity, the location of the depot with respect to the clients, and their scattered or clustered topology). In addition, we need to set different loading capacities for the vehicles in the HeVRPMD scenario. To achieve this we assume that the fixed capacity in classical VRP instances, Q_0, corresponds to a vehicle of type M. Accordingly, the capacity associated with vehicles of types S and L is set to $0.8Q_0$ and $1.25Q_0$, respectively. MRIG was implemented using the Java programming language and run on an Intel® Core™i5 CPU M520 2.40GHz with 4GB RAM, and Windows 7 Pro as the operating system. The experimental results for each instance are obtained after 30 runs using different seeds for the random number generation. Our stopping criterion is the maximum CPU time, set to 300 seconds, which allows enough iterations for the metaheuristic to reach a good convergence for the majority of the instances. Finally, the p parameter of the initial solution construction is set to 0.6, while the d parameter of the destruction operator is set to 0.5. These values were obtained after a preliminary experimentation, according to the statistical learning methodology proposed in Calvet et al. (2016).

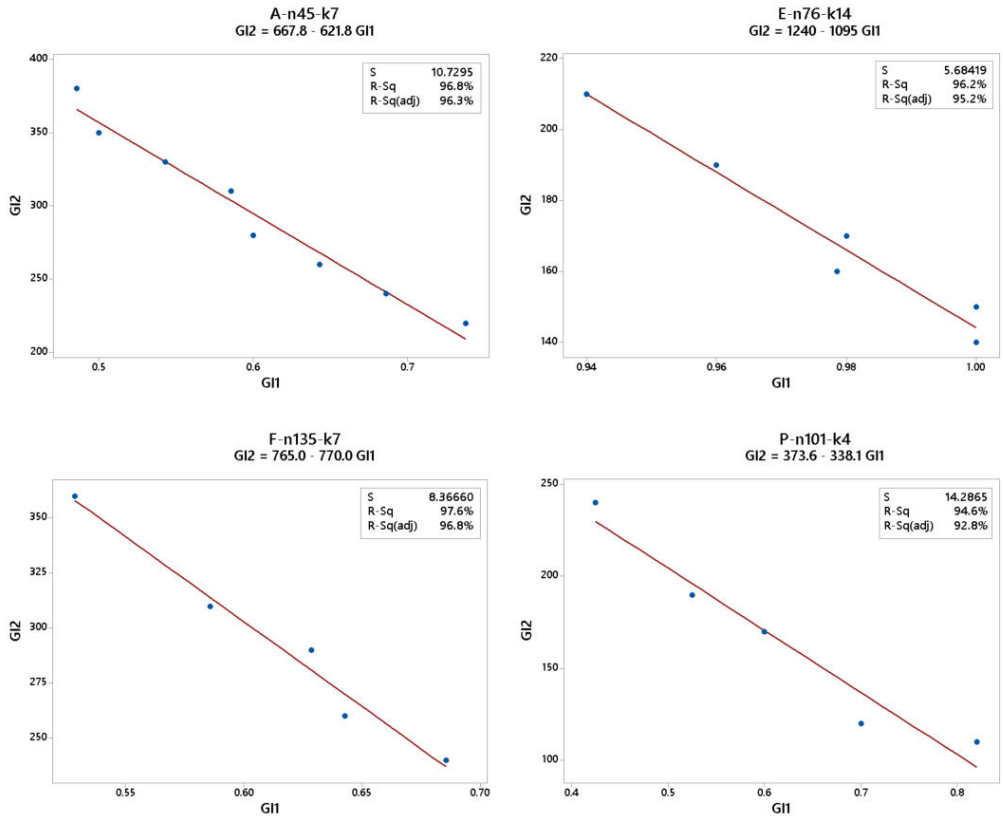

Figure 3: *Linear relationship between GI_1 and GI_2 for four problem instances.*

6 Analysis of results for the homogeneous VRPMD

To the best of our knowledge, the only algorithm considered in the literature to solve the homogeneous version of the VRPMD is the MRH one proposed in Juan et al. (2014b). The authors evaluated the performance of their MRH algorithm using 20 CVRP instances. In our paper, the number of instances tested has been increased up to 33, including the original 20 plus 13 additional ones. This allows us to directly compare the performance of our approach with the ones already published. The results are presented in Tables 1-3. These tables show the following information for each instance: its name, number of customers, vehicle capacity, and the distance-based cost of the best known solution (BKS) for the homogeneous VRP – without considering driving-range limitations – as provided in `http://www.branchandcut.org`. Also, these tables show a set of diverse feasible fleet configurations ($FleetCFG$) shown by $S/M/L$ those are found by the MRIG and MRH algorithms. For each instance, more than one feasible configuration has been found. However, only those solutions offering better values, either in distance-based cost or in green level, have been included in the tables. Notice that the "greener" configuration is considered. A configuration is greener than the other if it substitutes vehicles of type L by vehicles of type M or S (with S preferred over M), or vehicles of type M by vehicles of type S (without increasing the number of vehicles of type L). The associated distance-based cost for each fleet configuration, $DBCost$, is also included. The RPD column in each table shows the gap between the BKS and the distance-based costs provided by both algorithms. The best distance-based cost for each instance, and the associated gaps are indicated in bold. The last column of both tables shows the "diversified ratio", which is the ratio between the number of fleet configurations obtained using MRIG and MRH. Notice that, on the average, our MRIG provides 1.95 times more diversified fleet configurations than the MRH algorithm. In addition, the best distance-based cost obtained by MRIG is better than or equal to the one obtained by MRH in 80% of the instances.

Figure 4 illustrates the comparison using boxplots between three alternative solutions (A, B, and C) provided by both algorithms to show the gap differences with respect to the BKS of the problem. MRH-A and MRIG-A denote the best solution – in terms of distance-based cost – provided by each algorithm. Similarly, MRH-B and MRIG-B refer to alternative solutions that are greener than the previous ones. Finally, MRH-C and MRIG-C are solutions even greener than the ones provided by configuration B. Notice that solutions A provided by algorithms MRH and MRIG perform quite well in terms of their associated distance-based costs, since they offer very low gaps with respect to the BKS for the classical VRP – which is not considering any range constraint. As expected, the distance-based-cost gap with respect to the BKS increases as greener fleet configurations (suffixes B and C) are used. However, the average gap associated to MRIG is noticeably lower than the one associated to MRH for both B and C alternative configurations.

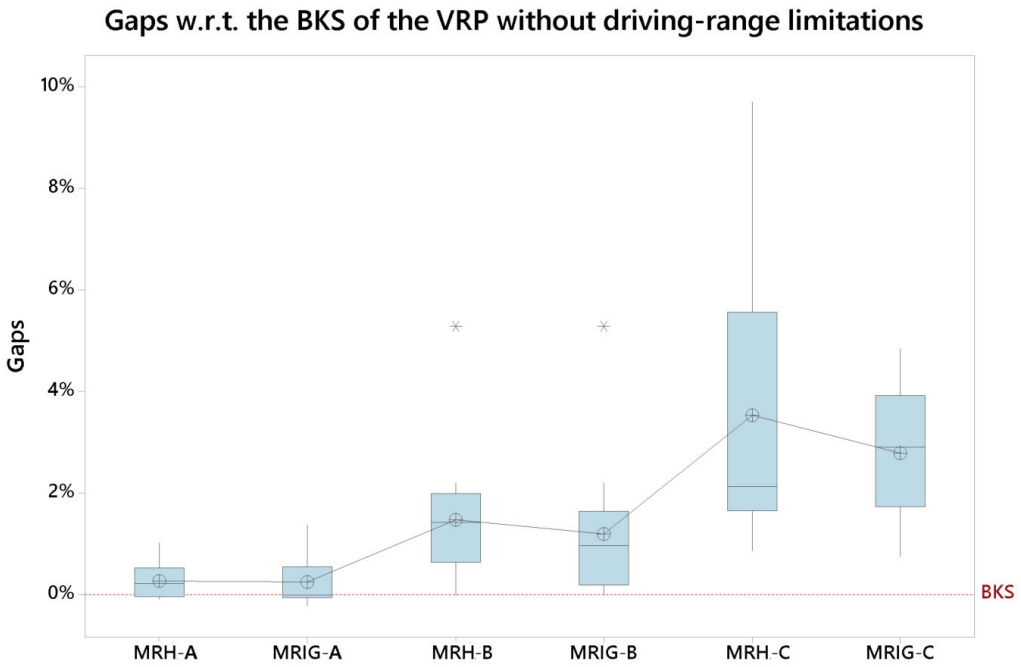

Figure 4: *Visual comparison using boxplots of three alternative solutions (A, B, and C) found by MRIG and MRH.*

Being able to select among different fleet configurations enriches the decision-making process. Thus, for example, the set of fleet configurations found by MRIG are more diversified and greener than those found by MRH in instance F-n135-k7. Moreover, MRIG can solve the instance using less vehicles of type L (ICEVs or PHEVs), since these are substituted by vehicles of types S and M (small- and medium-range EVs). Likewise, for the same fleet configuration, MRIG can find routes with a lower distance-based cost. For example, the greenest fleet configuration obtained by MRH is $3/2/2$, i.e.: it includes 3 vehicles of type S, 2 vehicles of type M, and 2 vehicles of type L. The associated distance-based cost is 1,190.07. For the same instance, MRIG was able to find a solution with an associated distance-based cost of 1,175.68 using the same fleet configuration. Furthermore, MRIG could find greener fleet configurations, such as $1/5/1$ and $2/4/1$, where the number of type L vehicles is decreased even further.

Table 1: *Experimental results for 20 classical VRP instances*

Instance name	Number of nodes	Capacity	*BKS*	*MRH*		*MRIG*		*RPD*		Diversified Ratio
				Fleet $CFG.$ $S/M/L$	*DBCost*	Fleet $CFG.$ $S/M/L$	*DBCost*	*MRH*	*MRIG*	
A-n32-k5	32	100	787.81	2/1/2	**787.08**	2/1/2	**787.08**	**0.00**	**0.00**	1
				1/3/1	829.41	1/3/1	829.41	5.38	5,38	
A-n38-k5	38	100	734.18	0/5/0	**733.95**	0/5/0	**733.95**	**0.00**	**0.00**	1.25
				1/3/1	734.18	1/3/1	734.18	0.03	0,03	
				1/4/0	735.05	1/4/0	735.05	0.15	0,15	
				3/3/0	763.13	3/3/0	755.89	3.97	2,99	
						1/5/0	**733.95**	-	-	
A-n65-k9	65	100	1181.69	1/8/0	**1183.31**	1/8/0	**1181.69**	**0.14**	**0.00**	1
				2/7/0	1191.27	2/7/0	1188.03	0.81	0.54	
				5/5/0	1276.21	5/5/0	1280.81	8.00	8.39	
				3/6/1	1238.33	3/6/1	1226.60	4.79	3.80	
				4/6/0	1253.81	4/6/0	1230.52	6.10	4.13	
				3/5/1	1297.31	3/6/0	1233.86	-	-	
A-n80-k10	80	100	1766.50	2/5/3	**1776.19**	2/5/3	**1775.75**	**0.55**	**0.52**	2.5
				1/7/2	1785.05	1/7/2	1785.04	1.05	1.05	
						2/6/2	1794.42	-	-	
						0/9/1	1994.16	-	-	
						2/8/1	2016.21	-	-	
B-n50-k7	50	100	744.78	2/5/0	**744.23**	2/5/0	**744.23**	**0.00**	**0.00**	1.67
				3/4/0	744.67	3/4/0	744.67	0.06	0.06	
				4/3/0	751.24	4/3/0	750.42	0.94	0.83	
						5/2/0	785.01	-	-	
						6/1/1	836.38	-	-	
B-n52-k7	52	100	750.08	4/2/1	**752.63**	4/2/1	**750.03**	**0.35**	**0.00**	1.5
				3/4/0	756.71	3/4/0	756.71	0.89	0.89	
						5/0/3	899.58	-	-	
B-n57-k9	57	100	1603.63	0/4/5	**1602.29**	0/4/5	**1602.29**	**0.00**	**0.00**	1.6
				0/5/4	1603.37	0/5/4	1603.37	0.07	0.07	
				0/6/3	1631.66	0/6/3	1631.85	1.83	1.84	
				1/3/5	1642.53	1/3/5	1636.34	2.51	2.13	
				1/4/4	1646.65	1/4/4	1637.44	2.77	2.19	
						1/5/3	1650.87	-	-	
						2/2/6	1694.09	-	-	
						0/7/2	1707.81	-	-	
B-n78-k10	78	100	1229.27	4/6/0	1253.10	4/6/0	**1245.64**	1.94	**1.33**	0.6
				6/5/0	1292.60	6/5/0	1288.67	5.15	4.83	
				3/7/0	1251.83	3/7/0	1246.21	1.84	1.38	
				4/5/1	**1236.33**			**0.57**	-	
				4/4/2	1252.76			-	-	
E-n22-k4	22	6000	375.28	2/2/0	**375.28**	2/2/0	**375.28**	**0.00**	**0.00**	2
				3/1/0	383.52	3/1/0	383.52	2.20	2.20	
						1/3/0	386.03	-	-	
						6/0/0	519.13	-	-	
E-n30-k3	30	4500	535.80	1/3/0	**505.01**	1/3/0	**505.01**	**0.00**	**0.00**	4
						2/1/1	579.78	-	-	
						3/0/2	597.65	-	-	
						3/1/1	633.37	-	-	

***Table 2**: Continued - Experimental results for 20 classical VRP instances*

Instance name	Number of nodes	Capacity	*BKS*	*MRH*		*MRIG*		*RPD*		Diversified Ratio
				Fleet $CFG.$ $S/M/L$	*DBCost*	Fleet $CFG.$ $S/M/L$	*DBCost*	*MRH*	*MRIG*	
E-n51-k5	51	160	524.94	3/2/0	**524.63**	3/2/0	**524.61**	**0.00**	**0.00**	3
						5/1/0	556.92	-	-	
						6/0/0	578.01	-	-	
E-n76-k10	76	140	837.36	7/3/0	**845.80**	7/3/0	**842.57**	**1.01**	**0.62**	4
				8/2/0	856.70	8/2/0	848.73	2.31	1.36	
				11/0/0	854.42	9/1/0	864.70	-	-	
						10/0/0	879.88	-	-	
E-n76-k14	76	100	1026.71	13/2/0	**1031.94**	13/2/0	1043.48	**0.51**	1.63	1.5
				14/1/0	1041.58	14/1/0	1044.28	1.45	1.71	
				13/1/0	1043.29	13/1/0	1060.05	1.61	3.25	
				15/0/0	1045.77	15/0/0	1050,79	1.86	2.35	
						12/3/0	**1038.48**		**1.15**	
						14/0/0	1075.74	-	-	
F-n135-k7	135	2210	1170.65	3/1/3	**1175.73**	3/1/3	**1168.01**	**0.66**	**0.00**	2.5
				3/2/2	1190.07	3/2/2	1175.68	1.89	0.66	
						2/3/2	1171.18	-	-	
						1/5/1	1215.14	-	-	
						2/4/1	1241.70	-	-	
M-n101-k10	101	200	819.81	8/2/0	**821.11**	8/2/0	**819.56**	**0.19**	**0.00**	3
						9/1/0	847.42	-	-	
						10/1/0	868.31	-	-	
M-n121-k7	121	200	1045.16	2/3/2	**1047.96**	2/3/2	**1044.64**	**0.32**	**0.00**	2
				1/7/0	1274.60	1/7/0	1287.52	22.01	23.25	
						3/2/3	1050.66	-	-	
						1/5/1	1129.40	-	-	
P-n50-k10	50	100	699.56	10/0/0	**700.66**	10/0/0	**700.66**	**0.16**	**0.16**	1
P-n55-k15	55	70	991.48	16/0/0	**952.02**	16/0/0	**953.18**	**0.00**	**0.12**	1
P-n70-k10	70	135	830.02	8/2/0	**834.38**	8/2/0	843.63	**0.53**	1.64	2.5
				10/0/0	841.56	10/0/0	851.39	1.39	2.57	
						6/4/0	**841.42**	-	**1.37**	
						9/1/0	844.35	-	-	
						7/3/0	842.36	-	-	
P-n76-k5	76	280	635.04	1/4/0	**638.44**	1/4/0	636.40	**0.55**	0.23	1.3
				2/3/0	647.51	2/3/0	653.07	1.97	2.85	
				4/2/0	696.63	4/2/0	666.60	9.71	4.98	
						0/5/0	**634.97**	-	**0.00**	
Average								2.08	1.92	1.95
								0.28	**0.26**	

Table 3: *Experimental results for 13 additional VRP instances*

Instance name	Number of nodes	Capacity	*BKS*	*MRIG*		RPD
				Fleet $CFG.\ S/M/L$	Cost	
A-n45-k7	45	100	1147.28	2/2/3	1146.77	0.00
				1/4/2	1154.43	0.67
				2/3/2	1155.60	0.77
				1/5/1	1191.29	3.88
				0/5/2	1174.01	2.38
				0/6/1	1230.27	7.28
				1/7/0	1463.93	27.66
				2/4/1	1186.46	3.46
A-n55-k9	55	100	1074.46	3/6/0	1074.46	0.00
				4/5/0	1092.88	1.71
				6/4/0	1150.04	7.03
A-n60-k9	60	100	1355.8	2/6/1	1357.72	0.14
A-n61-k9	61	100	1039.08	4/6/0	1040.31	0.12
				5/5/0	1045.40	0.61
				6/4/0	1057.00	1.72
				7/3/0	1091.31	5.03
E-n33-k4	33	8000	838.72	0/2/2	837.67	0.00
				0/3/1	847.37	1.16
E-n76-k7	76	220	687.60	3/4/0	690.20	0.38
				4/3/0	695.26	1.11
				5/2/0	705.97	2.67
				6/1/0	733.74	6.71
F-n45-k4	45	2010	724.57	1/2/1	723.54	0.00
				2/0/2	792.37	9.51
F-n72-k4	72	30000	248.81	4/0/0	241.97	0.00
P-n22-k8	22	3000	601.42	8/1/0	588.79	0.00
				9/0/0	647.63	9.99
P-n40-k5	40	140	461.73	5/0/0	461.73	0.00
P-n65-k10	65	130	796.67	10/0/0	797.82	0.14
P-n76-k4	76	350	598.22	0/4/0	600.55	0.39
				1/3/0	618.53	3.40
P-n101-k4	101	400	692.28	0/3/1	691.29	0.00
				0/4/0	694.67	0.49
				1/1/2	703.91	1.83
				1/2/1	700.88	1.39
				2/3/0	729.90	5.59

Regarding the 13 new instances analysed in this work, the results obtained with our MRIG approach are provided in Table 3. Notice that even considering the driving-range limitation, the distance-based cost of the best solution provided by the MRIG in 300 seconds is always similar to the classical BKS for the unconstrained problem. In other words, using an algorithm such as MRIG, it is frequently possible to find alternative solutions for the VRPMD with greener fleet configurations while, at the same time, these

solutions offer reasonably low distance-based costs – i.e., similar to the best ones that can be obtained for the classical VRP without driving-range limitations. Table 3 also shows that, in 9 out of 13 instances, the MRIG algorithm was able to generate alternative solutions with greener fleet configurations using less large vehicles. Even in these cases, the associated distance-based costs obtained by our algorithm are reasonably low. All in all, this section has shown that our algorithm is able to outperform the previous existing one for solving the homogeneous version of the VRPMD, both in terms of distance-based cost as well as in terms of green level of the solutions. Also, from a managerial perspective, the message is clear: (i) the introduction of EVs – with limited driving range – in transportation fleets does not have to cause a significant increase in distance-based costs (at least as far as an intelligent algorithm is used to optimize the associated routing problem); and (ii) among the different routing plans that such an algorithm can generate in just a few minutes, it is usually possible to choose one that offers a low distance-based cost while, at the same time, employs less contaminant vehicles. Despite the clear advantages of our approach, being a metaheuristic algorithm it cannot guarantee the optimality of the best-found solution. In addition, there is not a unique way of measuring the green level of a routing solution, since this is still a controversial concept in the scientific literature Juan et al. (2016).

7 Analysis of results for the heterogeneous HeVRPMD

The proposed algorithm is not only able to outperform the state-of-the-art results for the homogeneous VRPMD, but it can also solve the realistic heterogeneous version of the problem (i.e., HeVRPMD). Tables 4 to 6 show the results obtained by the algorithm when solving the HeVRPMD problem. These tables show the instance name and loading capacity in the homogeneous case Q_0 (first column), the best known solution (BKS) for the classical VRP without driving range limitations (second column), and loading capacities VS-VM-VL for small, medium, and large vehicles, respectively (third column). As stated in Section 5.3, heterogeneous instances were generated by considering $VS = 0.8Q_0$, $VM = Q_0$, and $VL = 1.25Q_0$. The MRIG algorithm provides, for each heterogeneous instance, a set of solutions. The solution with the minimum distance-based cost (*DBCost*) found by the MRIG is shown in bold and therefore, its corresponding *RPD* value is 0. Apart from the solution with the minimum *DBCost*, two sets of three solutions each are shown ($SetGI_1$ and $SetGI_2$). These three different fleet configurations are shown based on the minimum, medium, and maximum number of large vehicles, noted by L_s, L_m, and L_l, respectively. Green indexes GI_1 and GI_2 are shown for all the solutions (sixth and seventh columns for those of $SetGI_1$ and, 11-th and 12-th columns for those of $SetGI_2$). Therefore, Tables 4 to 6 report between six or seven solutions for each instance, depending on the case the minimum *DBCost* solution also belongs to the set of "green" solutions. At the end of Table 6, the average *RPD* values are also shown. About 42% of the solutions provided by MRIG reach the maximum green level

Table 4: *Experimental results for the HeVRPMD.*

Instance name (Q_0)	BKS Cost	VS-VM-VL	$SetGI_1$					$SetGI_2$				
			Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*	Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*
A-n32-k5(100)	787.81	80-100-125	**1/0/3**	**687.58**	0.25	310	**0.00**					
			0/0/4^{L_s}	730.19	0.00	400	6.20	0/0/5^{L_s}	1151.57	0.00	500	67.48
			1/1/2^{L_m}	695.44	0.43	240	1.14	1/1/3^{L_m}	726.91	0.34	340	5.72
			4/2/1^{L_l}	830.40	0.77	200	20.77	2/2/1^{L_l}	738.58	0.68	180	7.42
A-n38-k5(100)	734.18	80-100-125	**0/0/4**^{L_s}	**644.25**	0.00	400	**0.00**	0/1/5^{L_s}	1235.46	0.12	530	91.77
			2/1/3^{L_m}	672.16	0.45	350	4.33	2/0/3^{L_m}	676.67	0.40	320	5.03
			6/3/0^{L_l}	903.39	0.90	150	40.22	3/3/0^{L_l}	780.15	0.85	120	21.09
A-n45-k7(100)	1147.28	80-100-125	**1/0/5**	**990.18**	0.17	510	**0.00**					
			0/0/6^{L_s}	1079.11	0.00	600	8.98	1/1/7^{L_s}	1663.92	0.19	740	68.04
			3/1/5^{L_m}	1167.23	0.41	560	17.88	2/2/4^{L_m}	1100.33	0.43	480	11.12
			5/7/0^{L_l}	1675.26	0.83	260	69.19	1/7/0^{L_l}	1459.94	0.74	220	47.44
A-n55-k9(100)	1074.46	80-100-125	**0/0/7**^{L_s}	**942.84**	0.00	700	**0.00**	1/0/8^{L_s}	1326.54	0.11	810	40.70
			4/0/6^{L_m}	1014.94	0.40	640	7.65	1/3/4^{L_m}	968.22	0.39	500	2.69
			8/4/0^{L_l}	1272.33	0.90	200	34.95	8/4/0^{L_l}	1272.33	0.90	200	34.95
A-n60-k9(100)	1355.8	80-100-125	**0/1/6**	**1153.56**	0.10	630	**0.00**					
			0/0/7^{L_s}	1164.19	0.00	700	0.92	0/1/8^{L_s}	1697.11	0.08	830	47.12
			2/2/5^{L_m}	1184.15	0.38	580	2.65	1/2/5^{L_m}	1175.03	0.30	570	1.86
			5/6/1^{L_l}	1422.37	0.77	330	23.30	2/6/1^{L_l}	1334.07	0.69	300	15.65
A-n61-k9(100)	1039.08	80-100-125	**1/0/7**	**909.60**	0.13	710	**0.00**					
			0/0/8^{L_s}	935.72	0.00	800	2.87	2/0/8^{L_s}	1323.97	0.20	820	45.55
			4/1/5^{L_m}	965.43	0.47	570	6.14	4/2/4^{L_m}	988.82	0.54	500	8.71
			11/3/0^{L_l}	1311.24	0.94	200	44.15	9/3/0^{L_l}	1230.46	0.93	180	35.27
A-n65-k9(100)	1181.69	80-100-125	**1/0/7**	**1051.34**	0.13	710	**0.00**					
			0/0/8^{L_s}	1068.09	0.00	800	1.59	1/1/8^{L_s}	1604.55	0.17	840	52.62
			4/0/5^{L_m}	1142.82	0.44	540	8.70	1/4/4^{L_m}	1081.21	0.42	530	2.84
			8/5/0^{L_l}	1464.39	0.88	230	39.29	7/5/0^{L_l}	1398.35	0.88	220	33.01
A-n80-k10(100	1766.5	80-100-125	**0/1/7**	**1511.77**	0.09	730	**0.00**					
			0/0/8^{L_s}	1516.03	0.00	800	0.28	0/0/10^{L_s}	2368.14	0.00	1000	56.65
			4/0/7^{L_m}	1674.59	0.36	740	10.77	1/2/6^{L_m}	1533.72	0.27	670	1.45
			3/8/1^{L_l}	1948.24	0.72	370	28.87	1/8/1^{L_l}	1857.37	0.66	350	22.86
B-n50-k7(100)	744.78	80-100-125	**2/0/4**	**615.88**	0.33	420	**0.00**					
			0/0/5^{L_s}	619.18	0.00	500	0.54	1/0/6^{L_s}	952.27	0.14	610	54.62
			0/4/2^{L_m}	643.47	0.47	320	4.48	1/2/3^{L_m}	623.59	0.40	370	1.25
			9/2/0^{L_l}	950.75	0.95	150	54.37	7/2/0^{L_l}	848.22	0.93	130	37.73
B-n52-k7(100)	750.08	80-100-125	**0/0/5**^{L_s}	**650.00**	0.00	500	**0.00**	1/0/6^{L_s}	948.90	0.14	610	45.98
			3/0/4^{L_m}	686.33	0.43	430	5.59	2/2/3^{L_m}	662.08	0.49	380	1.86
			5/4/0^{L_l}	802.93	0.87	170	23.53	3/4/0^{L_l}	773.18	0.83	150	18.95
B-n57-k9(100)	1603.63	80-100-125	**0/0/7**^{L_s}	**1317.27**	0.00	700	**0.00**	0/0/9^{L_s}	2183.29	0.00	900	65.74
			2/1/6^{L_m}	1337.82	0.30	650	1.56	1/1/6^{L_m}	1324.36	0.21	640	0.54
			1/9/2^{L_l}	2013.19	0.61	480	52.83	0/6/2^{L_l}	1565.40	0.53	380	18,84
B-n78-k10(100)	1229.27	80-100-125	**0/0/8**^{L_s}	**1048.11**	0.00	800	**0.00**	0/1/9^{L_s}	1718.67	0.07	930	63.98
			4/1/6^{L_m}	1124.33	0.43	670	7.27	2/2/5^{L_m}	1066.13	0.38	580	1.72
			6/6/0^{L_l}	1327.24	0.85	240	26.63	6/6/0^{L_l}	1327.24	0.85	240	26.63
E-n22-k4(6000)	375.28	4800-6000-7500	**1/1/2**	**369.19**	0.43	240	**0.00**					
			0/0/3^{L_s}	377.68	0.00	300	2.30	1/1/3^{L_s}	439.38	0.34	340	19.01
			2/0/2^{L_m}	376.90	0.50	220	2.09	1/3/1^{L_m}	373.37	0.62	200	1.13
			6/0/0^{L_l}	533.15	1.00	60	44.41	6/0/0^{L_l}	533.15	1.00	60	44.41

Table 5: *Continued - Experimental results for the HeVRPMD.*

Instance name (Q_0)	BKS Cost	VS-VM-VL	$SetGI_1$					$SetGI_2$				
			Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*	Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*
E-n30-k3 (4500)	535.8	3600-4500-5625	**0/1/2**	**477.66**	0.23	230	**0.00**					
			0/0/3^{L_s}	507.64	0.00	300	6.28	1/0/4^{L_s}	922.54	0.20	410	93.14
			1/1/2^{L_m}	480.54	0.43	240	0.60	2/1/2^{L_m}	480.54	0.54	250	0.60
			3/3/0^{L_l}	507.17	0.85	120	6.18	0/3/0^{L_l}	550.91	0.70	90	15.33
E-n33-k4(8000)	838.72	6400-8000-10000	**0/0/3^{L_s}**	**704.32**	0.00	300	**0.00**	1/0/4^{L_s}	930.78	0.20	410	32.15
			1/2/3^{L_m}	856.47	0.40	370	21.60	**0/0/3^{L_m}**	**704.32**	0.00	300	**0.00**
			2/3/1^{L_l}	998.74	0.68	210	41.80	0/3/1^{L_l}	829.35	0.53	190	17.75
E-n51-k5(160)	524.94	128-160-200	**0/0/4^{L_s}**	**497.74**	0.00	400	**0.00**	1/1/4^{L_s}	656.74	0.28	440	31.94
			3/0/3^{L_m}	522.59	0.50	330	4.99	2/1/2^{L_m}	520.60	0.54	250	4.59
			7/0/0^{L_l}	598.69	1.00	70	20.28	7/0/0^{L_l}	598.69	1.00	70	20.28
E-n76-k7(220)	687.60	176-220-275	**1/0/5**	**648.07**	0.17	510	**0.00**					
			0/0/6^{L_s}	1083.78	0.00	600	67.23	0/0/6^{L_s}	1083.78	0.00	600	67.23
			4/0/4^{L_m}	695.36	0.50	440	7.30	2/1/3^{L_m}	667.59	0.40	350	3.01
			9/0/0^{L_l}	791.22	1.00	90	22.09	9/0/0^{L_l}	791.22	1.00	90	22.09
E-n76-k10(140)	837.36	112-140-175	**0/0/8^{L_s}**	**757.81**	0.00	800	**0.00**	1/1/8^{L_s}	903.53	0.17	840	19.23
			5/0/5^{L_m}	824.65	0.50	550	8.82	1/6/3^{L_m}	819.66	0.52	490	8.16
			18/0/0^{L_l}	1235.73	1.00	180	63.07	13/0/0^{L_l}	986.11	1.00	130	30.13
E-n76-k14(100)	1026.71	80-100-125	**1/0/11**	**902.69**	0.08	1110	**0.00**					
			0/0/12^{L_l}	930.71	0.00	1200	3.10	18/0/0^{L_s}	1251.03	1.00	180	38.59
			8/0/8^{L_m}	990.48	0.50	880	9.73	4/5/5^{L_m}	988.25	0.54	690	9.48
			19/0/0^{L_s}	1238.74	1.00	190	37.23	1/0/12^{L_l}	946.84	0.08	1210	4.89
F-n45-k4 (2010)	724.57	1608-2010-2512	**0/0/3^{L_s}**	**690.89**	0.00	300	**0.00**	2/0/4^{L_s}	1126.57	0.33	420	63.06
			2/0/3^{L_m}	714.90	0.40	320	3.48	0/3/2^{L_m}	810.13	0.42	290	17.26
			4/2/1^{L_l}	744.86	0.77	200	7.81	3/1/1^{L_l}	860.73	0.74	160	24.58
F-n72-k4(30000)	248.81	24000-30000-37500	**0/3/1**	**237.53**	0.53	190	**0.00**					
			1/0/3^{L_s}	242.82	0.25	310	2.23	3/0/3^{L_s}	269.56	0.50	330	13.48
			2/0/3^{L_m}	250.35	0.40	320	5.40	**0/3/1^{L_m}**	**237.53**	0.53	190	**0.00**
			5/0/0^{L_l}	270.60	1.00	50	13.92	5/0/0^{L_l}	270.60	1.00	50	13.92
F-n135-k7(2210)	1170.65	1768-2210-2762	**1/0/5**	**1053.07**	0.17	510	**0.00**					
			0/0/6^{L_s}	1070.75	0.00	600	1.68	2/0/7^{L_s}	1838.11	0.22	720	74.55
			4/0/5^{L_m}	1206.49	0.44	540	14.57	1/2/4^{L_m}	1071.17	0.34	470	1.72
			10/2/1^{L_l}	1455.24	0.83	260	38.19	4/3/1^{L_l}	1217.92	0.76	230	15.65
M-n101-k10(200)	819.81	60-200-250	**1/0/7**	**757.61**	0.13	710	**0.00**					
			0/0/8^{L_s}	787.39	0.00	800	3.93	0/0/8^{L_s}	787.39	0.00	800	3.93
			4/2/5^{L_m}	861.15	0.49	600	13.67	4/1/4^{L_m}	790.40	0.52	470	4.33
			14/1/0^{L_l}	1122.67	0.98	170	48.19	11/1/0^{L_l}	1014.53	0.98	140	33.91
M-n121-k7(200)	1045.16	60-200-250	**1/0/5**	**953.34**	0.17	510	**0.00**					
			0/0/6^{L_s}	965.17	0.00	600	1.24	3/0/7^{L_s}	1680.06	0.30	730	76.23
			4/0/6^{L_m}	1537.13	0.40	640	61.24	1/2/4^{L_m}	967.07	0.34	470	1.44
			4/7/0^{L_l}	1393.23	0.81	250	46.14	1/7/0^{L_l}	1289.11	0.74	220	35.22
P-n22-k8(3000)	601.42	2400-3000-3750	**1/1/5**	**507.16**	0.24	540	**0.00**					
			0/0/7^{L_s}	524.97	0.00	700	3.51	0/0/7^{L_s}	524.97	0.00	700	3.51
			3/2/4^{L_m}	536.91	0.49	490	5.87	2/3/3^{L_m}	541.46	0.51	410	6.76
			14/1/0^{L_l}	876.21	0.98	170	72.77	10/1/0^{L_l}	783.15	0.97	130	54.42

Table 6: *Continued - Experimental results for the HeVRPMD.*

Instance name (Q_0)	BKS Cost	VS-VM-VL	$SetGI_1$					$SetGI_2$				
			Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*	Fleet CFG. S/M/L	*DBCost*	GI_1	GI_2	*RPD*
P-n40-k5(140)	461.73	112-140-175	**0/1/3**	**431.67**	0.18	330	**0.00**					
			$0/0/4^{L_s}$	432.23	0.00	400	0.13	$2/0/4^{L_s}$	584.80	0.33	420	35.47
			$3/0/3^{L_m}$	457.78	0.50	330	6.05	$4/0/2^{L_m}$	463.83	0.67	240	7.45
			$6/0/0^{L_l}$	514.97	1.00	60	19.30	$6/0/0^{L_l}$	514.97	1.00	60	19.30
P-n50-k10(100)	699.56	80-100-125	$\mathbf{0/0/8^{L_s}}$	**607.39**	0.00	800	**0.00**	$0/1/8^{L_s}$	658.36	0.08	830	8.39
			$5/0/5^{L_m}$	669.00	0.50	550	10.14	$0/6/3^{L_m}$	657.15	0.47	480	8.19
			$13/0/0^{L_l}$	805.71	1.00	130	32.65	$13/0/0^{L_l}$	805.71	1.00	130	32.65
P-n55-k15(70)	991.48	56-70-87	$\mathbf{0/0/13^{L_s}}$	**824.21**	0.00	1300	**0.00**	$0/1/13^{L_s}$	883.51	0.05	1330	7.20
			$8/0/8^{L_m}$	915.58	0.50	880	11.09	$3/8/4^{L_m}$	919.94	0.57	670	11.62
			$20/0/0^{L_l}$	1126.70	1.00	200	36.70	$20/0/0^{L_l}$	1126.70	1.00	200	36.70
P-n65-k10(130)	796.67	104-130-162	$\mathbf{0/0/8^{L_s}}$	**726.51**	0.00	800	**0.00**	$3/0/8^{L_s}$	831.83	0.27	830	14.50
			$5/0/5^{L_m}$	779.95	0.50	550	7.36	$0/6/3^{L_m}$	766.30	0.47	480	5.48
			$13/0/0^{L_l}$	931.96	1.00	130	28.28	$13/0/0^{L_l}$	931.96	1.00	130	28.28
P-n70-k10(135)	830.02	108-135-196	$\mathbf{0/0/8^{L_s}}$	**760.93**	0.00	800	**0.00**	$1/1/8^{L_s}$	916.60	0.17	840	20.46
			$5/0/5^{L_m}$	821.68	0.50	550	7.98	$1/6/3^{L_m}$	812.82	0.52	490	6.82
			$13/0/0^{L_l}$	969.13	1.00	130	27.36	$13/0/0^{L_l}$	969.13	1.00	130	27.36
P-n76-k4(350)	598.22	280-350-437	**1/1/2**	**594.64**	0.43	240	**0.00**					
			$0/0/4^{L_s}$	695.78	0.00	400	17.01	$2/1/4^{L_s}$	935.17	0.39	450	57.27
			$2/0/2^{L_m}$	606.86	0.50	220	2.06	$0/2/2^{L_m}$	597.13	0.35	260	0.42
			$8/0/0^{L_l}$	744.71	1.00	80	25.24	$8/0/0^{L_l}$	744.71	1.00	80	25.24
P-n76-k5(280)	635.04	224-280-350	$\mathbf{0/0/4^{L_s}}$	**601.29**	0.00	400	**0.00**	$2/0/5^{L_s}$	974.60	0.29	520	62.09
			$3/0/3^{L_m}$	632.18	0.50	160	5.14	$1/3/2^{L_m}$	646.81	0.52	300	7.57
			$8/0/0^{L_l}$	767.63	1.00	80	27.66	$8/0/0^{L_l}$	767.63	1.00	80	27.66
P-n101-k4(400)	692.28	320-400-500	$\mathbf{0/0/3^{L_s}}$	**679.68**	0.00	300	**0.00**	$3/1/4^{L_s}$	983.43	0.46	460	44.69
			$2/0/2^{L_m}$	711.54	0.50	220	4.69	$2/2/2^{L_m}$	690.92	0.57	280	1.65
			$11/0/0^{L_l}$	1004.26	1.00	110	47.76	$11/0/0^{L_l}$	1004.26	1.00	110	47.76
Average							3.94^{L_s}					45.04^{L_s}
							8.85^{L_m}					4.62^{L_m}
							35.31^{L_l}					27.19^{L_l}

regarding index GI_1. Moreover, MRIG is able to provide routes with a lower cost than the ones in the homogeneous case. This is due to an increase in the loading capacity of vehicles of type *L*. For a randomly selected subset of 20 instances, Figure 5 shows a scatter-plot of distance-based cost (*DBCost*) versus green level (as measured by the GI_1 index). The left point (circle) in each panel represents the best solution in terms of distance-based cost, while the right point (square) represents an alternative 'greener' solution (i.e., one with a higher value of GI_1). Notice that it is usually possible to choose a greener solution without having to assume a noticeable increase in the distance-based cost (instances E-n76-k10, F-n72-k4, P-n101-k4, or P-n40-k5 are good examples). Just in a few cases (e.g., instance M-n121-k7), a noticeable increase in the green level might also require paying a much higher distance-based cost.

Finally, Figure 6 illustrates eight alternative fleet configurations for the A-n45-k7 instance. The auxiliary index $GI_2i = 1/GI_2$ has been employed instead of GI_2 to fa-

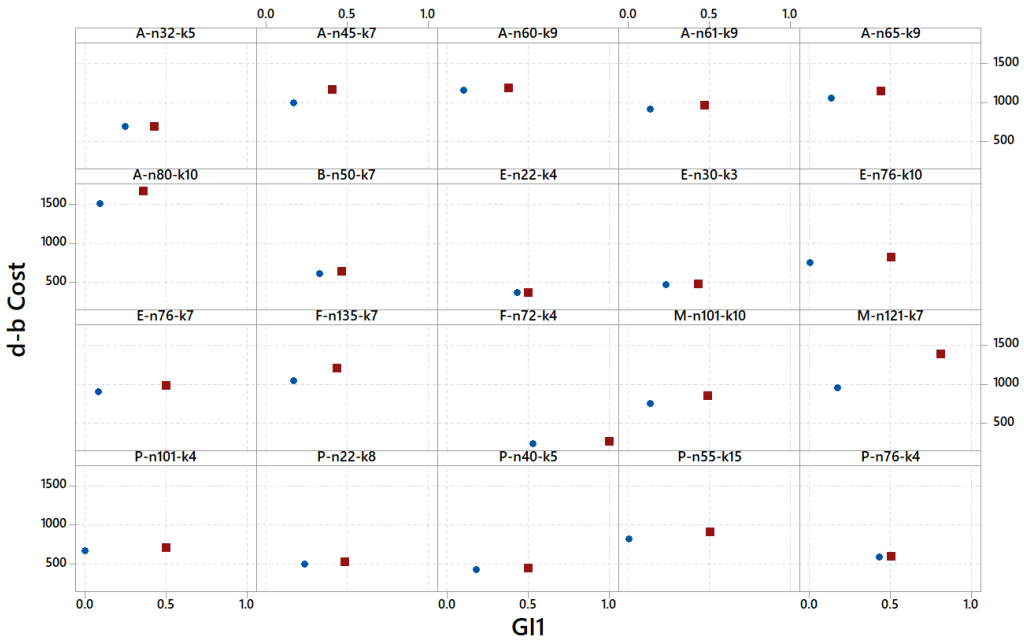

Figure 5: *A comparison of best DBCost solutions (circles) and 'greener' alternatives (squares).*

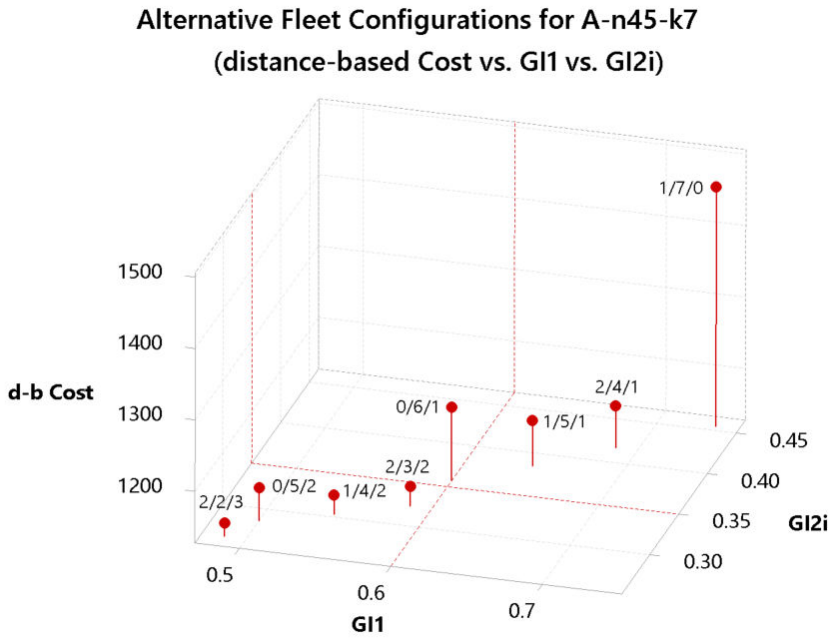

Figure 6: *Cost and green indexes values of eight alternative fleet configurations for instance A-n45-k7.*

cilitate the readability of the figure (i.e., higher values of both GI_1 and $GI_2 i$ represent greener configurations). In this figure, we can see that the best configuration, in terms of distance-based cost, is $2/2/3$. However, the green level of the other solutions is higher. Therefore, the decision-maker could prefer alternative configurations such as $2/3/2$ and $2/4/1$. The latter two solutions offer greener configurations without a significant increase in distance-based cost. This illustrative plot shows how decision-makers can benefit from being able to choose from alternative fleet configurations with different green levels.

8 Conclusion and future work

The paper has introduced a realistic version of the vehicle routing problem in which hybrid fleets of gas and electric vehicles are considered. The introduction of electric vehicles in the model offers clear benefits in terms of making transportation more environment-friendly. However, due to the limited driving-range capabilities of electric batteries, the use of these vehicles also imposes new challenges that need to be solved. We have proposed a novel metaheuristic, MRIG, to solve both the homogeneous and heterogeneous version of the vehicle routing problem with multiple driving ranges. Our MRIG algorithm has outperformed the state-of-the-art results in the case of the homogeneous version. In addition, our approach has been able to solve the heterogeneous version too. This version considers hybrid fleets of vehicles with both different driving ranges and loading capacities.

The metaheuristic is designed to generate a set of fleet configurations with different green levels including small- and medium-driving range electric vehicles instead of gas-fueled vehicles. This solution can be used by decision-makers to help them choose the fleet configuration that fits best with their needs among a set of provided solutions. In some cases, greener configurations might have a higher distance-based cost, but the extensive experiments carried out in this paper showed that it is frequently possible to choose a greener configuration without a significant increase in the distance-based cost. We highlight here some research directions to extend this work, based on considering more realistic variants thus, following a common trend in logistics Solos, Tassopoulos and Beligiannis (2016); Solano-Charris, Prins and Santos (2015). First, we aim to design a multi-objective optimization problem and method Martin et al. (2009) to optimize both the distance-based cost and green level of the fleet configuration. Additionally, we will consider the development of a stochastic model and the corresponding solving approach in the presence of uncertainty (e.g., random demands or random traveling times).

Acknowledgements

This work is jointly supported by the Spanish Ministry of Science (RED2018-102642-T), the Andalusian Government, the National Agency for Research Funding AEI, and ERDF (EU) under grants EXASOCO (PGC2018-101216-B-I00), AIMAR (A-TIC-284-UGR18), SIMARK (PY18-4475) and the Ramón y Cajal program (RYC-2016-19800).

References

AbdAllah, A. M. F., D. L. Essam, and R. A. Sarker (2017). On solving periodic re-optimization dynamic vehicle routing problems. *Applied Soft Computing*, 55, 1–12.

Achtnicht, M., G. Bühler, and C. Hermeling (2012). The impact of fuel availability on demand for alternative-fuel vehicles. *Transportation Research Part D: Transport and Environment*, 17, 262–269.

Almouhanna, A., C. L. Quintero-Araujo, J. Panadero, A. A. Juan, B. Khosravi, and D. Ouelhadj (2020). The location routing problem using electric vehicles with constrained distance. *Computers & Operations Research*, 115, 104864.

Andreatta, G., M. Casula, C. De Francesco, and L. De Giovanni (2016). A branch-and-price based heuristic for the stochastic vehicle routing problem with hard time windows. *Electronic Notes in Discrete Mathematics*, 52, 325–332.

Baldacci, R., M. Battarra, and D. Vigo (2008). Routing a heterogeneous fleet of vehicles. In *The Vehicle Routing Problem: Latest Advances and New Challenges*, Chapter 10, pp. 3–27. Springer.

Belloso, J., A. A. Juan, and J. Faulin (2019). An iterative biased-randomized heuristic for the fleet size and mix vehicle-routing problem with backhauls. *International Transactions in Operational Research*, 26, 289–301.

Brito, J., F. J. Martínez, J. A. Moreno, and J. L. Verdegay (2015). An ACO hybrid metaheuristic for close–open vehicle routing problems with time windows and fuzzy constraints. *Applied Soft Computing*, 32, 154–163.

Calvet, L., A. A. Juan, C. Serrat, and J. Ries (2016). A statistical learning based approach for parameter fine-tuning of metaheuristics. *SORT-Statistics and Operations Research Transactions*, 40, 201–224.

Chan, C. C., Y. S. Wong, A. Bouscayrol, and K. Chen (2009). Powering Sustainable Mobility: Roadmaps of Electric, Hybrid, and Fuel Cell Vehicles. *Proceedings of the IEEE*, 97, 603–607.

Chebbi, O. and J. Chaouachi (2015). Multi-objective Iterated Greedy Variable Neighborhood Search Algorithm For Solving a full-load automated guided vehicle routing problem with battery constraints. *Electronic Notes in Discrete Mathematics*, 47, 165–172.

Clarke, G. and J. W. Wright (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*, 12, 568–581.

Crainic, T. G. (2000). Service network design in freight transportation. *European Journal of Operational Research*, 122, 272–288.

Dell'Amico, M., M. Monaci, C. Pagani, and D. Vigo (2007). Heuristic approaches for the fleet size and mix vehicle routing problem with time windows. *Transportation Science*, 41, 516–526.

Dominguez, O., A. A. Juan, I. A. de la Nuez, and D. Ouelhadj (2016). An ils-biased randomization algorithm for the two-dimensional loading hfvrp with sequential loading and items rotation. *Journal of the Operational Research Society*, 67, 37–53.

Erdoğan, S. and E. Miller-Hooks (2012). A Green Vehicle Routing Problem. *Transportation Research Part E: Logistics and Transportation Review*, 48, 100–114.

Faulin, J., S. E. Grasman, A. A. Juan, and P. Hirsch (2019). Sustainable transportation: concepts and current practices. In *Sustainable Transportation and Smart Logistics*, pp. 3–23. Elsevier.

Faulin, J., F. Lera-López, and A. A. Juan (2011). Optimizing routes with safety and environmental criteria in transportation management in spain: a case study. *International Journal of Information Systems and Supply Chain Management*, 4, 38–59.

Feillet, D. (2010). A tutorial on column generation and branch-and-price for vehicle routing problems. *4OR - A Quarterly Journal of Operations Research*, 8, 407–424.

Felipe, N., M. T. Ortuño, G. Righini, and G. Tirado (2014). A heuristic approach for the green vehicle routing problem with multiple technologies and partial recharges. *Transportation Research Part E: Logistics and Transportation Review*, 71, 111–128.

Ferone, D., A. Gruler, P. Festa, and A. A. Juan (2019). Enhancing and extending the classical grasp framework with biased randomisation and simulation. *Journal of the Operational Research Society*, 70, 1362–1375.

Ferreira, J., P. Pereira, P. Filipe, and J. Afonso (2011). Recommender system for drivers of electric vehicles. *ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology*, 5, 244–248.

François, V., Y. Arda, Y. Crama, and G. Laporte (2016). Large neighborhood search for multi-trip vehicle routing. *European Journal of Operational Research*, 255, 422–441.

Ghannadpour, S. F., S. Noori, R. Tavakkoli-Moghaddam, and K. Ghoseiri (2014). A multi-objective dynamic vehicle routing problem with fuzzy time windows: Model, solution and application. *Applied Soft Computing*, 14, 504–527.

Goeke, D. and M. Schneider (2015). Routing a mixed fleet of electric and conventional vehicles. *European Journal of Operational Research*, 245, 81–99.

González-Martín, S., A. A. Juan, D. Riera, Q. Castellà, R. Muñoz, and A. Pérez (2012). Development and assessment of the sharp and randsharp algorithms for the arc routing problem. *AI Communications*, 25, 173–189.

Hatami, S., R. Ruiz, and C. Andrés-Romano (2015). Heuristics and metaheuristics for the distributed assembly permutation flowshop scheduling problem with sequence dependent setup times. *International Journal of Production Economics*, 169, 76–88.

Hiermann, G., J. Puchinger, S. Ropke, and R. F. Hartl (2016). The Electric Fleet Size and Mix Vehicle Routing Problem with Time Windows and Recharging Stations. *European Journal of Operational Research*, 252, 995–1018.

Hokama, P., F. K. Miyazawa, and E. C. Xavier (2016). A branch-and-cut approach for the vehicle routing problem with loading constraints. *Expert Systems with Applications*, 47, 1–13.

Jie, W., J. Yang, M. Zhang, and Y. Huang (2019). The two-echelon capacitated electric vehicle routing problem with battery swapping stations: Formulation and efficient methodology. *European Journal of Operational Research*, 272, 879–904.

Juan, A. A., J. Faulin, J. C. Cruz, B. B. Barrios, and E. Martinez (2014a). A successive approximations method for the heterogeneous vehicle routing problem: analysing different fleet configurations. *European Journal of Industrial Engineering*, 8, 762.

Juan, A. A., J. Goentzel, and T. Bektaş (2014b). Routing fleets with multiple driving ranges: Is it possible to use greener fleet configurations? *Applied Soft Computing Journal*, 21, 84–94.

Juan, A. A., C. A. Mendez, J. Faulin, J. De Armas, and S. E. Grasman (2016). Electric vehicles in logistics and transportation: A survey on emerging environmental, strategic, and operational challenges. *Energies*, 9, 1–21.

Karakatič, S. and V. Podgorelec (2015). A survey of genetic algorithms for solving multi depot vehicle routing problem. *Applied Soft Computing*, 27, 519–532.

Kek, A. G. H., R. L. Cheu, and Q. Meng (2008). Distance-constrained capacitated vehicle routing problems with flexible assignment of start and end depots. *Mathematical and Computer Modelling*, 47, 140–152.

Keskin, M. and B. Çatay (2016). Partial recharge strategies for the electric vehicle routing problem with time windows. *Transportation Research Part C: Emerging Technologies*, 65, 111–127.

Koç, C. a., T. Bektaş, O. Jabali, and G. Laporte (2016). Thirty years of heterogeneous vehicle routing. *European Journal of Operational Research*, 249, 1–21.

Laporte, G. (2009). Fifty Years of Vehicle Routing. *Transportation Science*, 43, 408–416.

Laporte, G., M. Desrochers, and Y. Nobert (1984). Two exact algorithms for the distance-constrained vehicle routing problem. *Networks*, 14, 161–172.

Li, C.-L., D. Simchi-Levi, and M. Desrochers (1992). On the distance constrained vehicle routing problem. *Operations research*, 40, 790–799.

Lin, J., W. Zhou, and O. Wolfson (2016). Electric Vehicle Routing Problem. *Transportation Research Procedia*, 12, 508–521.

Martin, D., R. del Toro, R. Haber, and J. Dorronsoro (2009). Optimal tuning of a networked linear controller using a multi-objective genetic algorithm and its application to one complex electromechanical process. *International Journal of Innovative Computing, Information and Control*, 5, 3405–3414.

Mattila, T. and R. Antikainen (2011). Backcasting sustainable freight transport systems for Europe in 2050. *Energy Policy*, 39, 1241–1248.

Montoya, A., C. Guéret, J. E. Mendoza, and J. G. Villegas (2014). A multi-space sampling heuristic for the green vehicle routing problem. *Transportation Research Part C: Emerging Technologies*, 70, 113–128.

Osman, I. and C. Potts (1989). Simulated annealing for permutation flow-shop scheduling. *Omega*, 17, 551–557.

Pierre, D. M. and N. Zakaria (2017). Stochastic partially optimized cyclic shift crossover for multi-objective genetic algorithms for the vehicle routing problem with time-windows. *Applied Soft Computing*, 52, 863–876.

Quintero-Araujo, C. L., J. P. Caballero-Villalobos, A. A. Juan, and J. R. Montoya-Torres (2017). A biased-randomized metaheuristic for the capacitated location routing problem. *International Transactions in Operational Research*, 24, 1079–1098.

Reyes-Rubiano, L., D. Ferone, A. A. Juan, and J. Faulin (2019). A simheuristic for routing electric vehicles with limited driving ranges and stochastic travel times. *SORT-Statistics and Operations Research Transactions*, 43, 3–24.

Ruiz, R. and T. Stützle (2007). A simple and effective iterated greedy algorithm for the permutation flow-shop scheduling problem. *European Journal of Operational Research*, 177, 2033–2049.

Ruiz, R. and T. Stützle (2008). An Iterated Greedy heuristic for the sequence dependent setup times flow-shop problem with makespan and weighted tardiness objectives. *European Journal of Operational Research*, 187, 1143–1159.

Schneider, M., A. Stenger, and D. Goeke (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science*, 48, 500–520.

Solano-Charris, E., C. Prins, and A. C. Santos (2015). Local search based metaheuristics for the robust vehicle routing problem with discrete scenarios. *Applied Soft Computing*, 32, 518–531.

Solos, I. P., I. X. Tassopoulos, and G. N. Beligiannis (2016). Optimizing shift scheduling for tank trucks using an effective stochastic variable neighbourhood approach. *International Journal of Artificial Intelligence*, 14, 1–26.

Vaz Penna, P. H., H. M. Afsar, C. Prins, and C. Prodhon (2016). A Hybrid Iterative Local Search Algorithm for The Electric Fleet Size and Mix Vehicle Routing Problem with Time Windows and Recharging Stations. *IFAC-PapersOnLine*, 49, 955–960.

Verma, A. (2018). Electric vehicle routing problem with time windows, recharging stations and battery swapping stations. *EURO Journal on Transportation and Logistics*, 7, 415–451.

Wang, C., D. Mu, F. Zhao, and J. W. Sutherland (2015). A parallel simulated annealing method for the vehicle routing problem with simultaneous pickup-delivery and time windows. *Computers & Industrial Engineering*, 83, 111–122.

Williams, J. H., A. Debenedictis, R. Ghanadan, A. Mahone, J. Moore, W. R. M. Iii, S. Price, and M. S. Torn (2012). 2050 : The Pivotal Role of Electricity. *Science*, 335, 53–60.

Wirasingha, S. G., N. Schofield, and A. Emadi (2008). Plug-in hybrid electric vehicle developments in the US: Trends, barriers, and economic feasibility. In *2008 IEEE Vehicle Power and Propulsion Conference*, pp. 1–8.

Yu, V. F. and S.-Y. Lin (2015). A simulated annealing heuristic for the open location-routing problem. *Computers & Operations Research*, 62, 184–196.

SORT 44 (1) January-June 2020, 171-200 DOI: 10.2436/20.8080.02.99

Bayesian structured antedependence model proposals for longitudinal data

Edwin Castillo-Carreno[1], Edilberto Cepeda-Cuervo[1]
and Vicente Núñez-Antón[2]

Abstract

An important problem in Statistics is the study of longitudinal data taking into account the effect of other explanatory variables, such as treatments and time and, simultaneously, the incorporation into the model of the time dependence between observations on the same individual. The latter is specially relevant in the case of nonstationary correlations, and nonconstant variances for the different time point at which measurements are taken. Antedependence models constitute a well known commonly used set of models that can accommodate this behaviour. These covariance models can include too many parameters and estimation can be a complicated optimization problem requiring the use of complex algorithms and programming. In this paper, a new Bayesian approach to analyse longitudinal data within the context of antedependence models is proposed. This innovative approach takes into account the possibility of having nonstationary correlations and variances, and proposes a robust and computationally efficient estimation method for this type of data. We consider the joint modelling of the mean and covariance structures for the general antedependence model, estimating their parameters in a longitudinal data context. Our Bayesian approach is based on a generalization of the Gibbs sampling and Metropolis-Hastings by blocks algorithm, properly adapted to the antedependence models longitudinal data settings. Finally, we illustrate the proposed methodology by analysing several examples where antedependence models have been shown to be useful: the small mice, the speech recognition and the race data sets.

MSC: 62F15, 62J05, 62P10.

Keywords: Antedependence models, Bayesian methods, Gibbs sampling, Mean-covariance modelling, Nonstationary correlation.

1 Introduction

Continuous longitudinal data consist of repeated measurements on the same subject over time. These measurements are typically correlated and there have been several propos-

[1] Department of Statistics, Faculty of Sciences, Universidad Nacional de Colombia, Carrera 30, 45-03, Building 476, Bogotá D.C. (Colombia), ecastillo@unal.edu.co, ecepedac@unal.edu.co.

[2] (Corresponding author) Department of Applied Economics III (Econometrics and Statistics), Faculty of Economics and Business, University of the Basque Country UPV/EHU. Avda. Lehendakari Aguirre 83, 48015 Bilbao. (Spain). vicente.nunezanton@ehu.eus.

Received: July 2019
Accepted: May 2020

als in the literature to handle stationary or nonstationary correlations and variances, as well as balanced or unbalanced longitudinal data (Diggle et al., 2002; Weiss, 2005; Verbeke and Molenberghs, 2000; Fitzmaurice et al., 2009). A general fixed effects regression model for longitudinal data can be defined by assuming that the response variable $\mathbf{Y}_i$ can be explained with the model given by:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i,\ i = 1,\ldots,m, \tag{1}$$

where $\mathbf{Y}_i = (Y_{i1},\ldots,Y_{in_i})^\top$ is the $n_i \times 1$ vector of responses for subject i, X_i is the $n_i \times q$ design matrix of rank q, which includes the covariates for the i-th subject; $\boldsymbol{\epsilon}_i$ is the vector of errors, assumed to follow a multivariate normal distribution with mean $\mathbf{0}$, and a given variance-covariance matrix so that $\text{Var}(\mathbf{Y}_i) = \boldsymbol{\Sigma}_i(\boldsymbol{\theta}) = \sigma^2 V_{0i}$, whereas $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)^\top$ and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_q)^\top$ are k and q-dimensional vectors of unknown parameters for the variance-covariance and mean model, respectively. Here, n_i represents the number of observations available for the i-th subject. If the number of observations available for each subject is the same (i.e., $n_i = n$, $\forall i$), we have a balanced data set. However, observations are, in general, not equally spaced. In addition, m represents the number of individuals in the study, and $N = \sum_{i=1}^{m} n_i$ represents the total number of observations. Fitting for the mean and covariance structure can be carried out by using maximum likelihood estimation methods with numerical maximization, such as the Newton Raphson or the EM algorithms (Ware, 1985). The model's assumptions include independence of responses from different subjects, multivariate normality of responses, and either no missing data or, at worst, ignorably missing responses (Laird, 1988).

The approach of fitting a regression model for longitudinal data by means of specifying the variance-covariance structure includes the possibility of having several different structures, which can be stationary or nonstationary in terms of correlation between observations along time, and homogeneous or heterogeneous in terms of variance as a function of the time at which observations are taken. Among the most commonly used covariance structures featuring homogeneous variances and stationary correlations are the compound symmetry (CS), autoregressive structures of order p (AR(p)), autoregressive with moving average structures of order p and q (ARMA(p,q)) models (Weiss, 2005). Models for nonstationary correlations and heterogeneous variances include the heterogeneous versions of the previous models, which are not always the best-fitting models for this type of settings. Therefore, mode general models, such as the integrated autoregressive with moving average model (ARIMA) or generalizations of the autoregressive models, such as the unstructured and structured versions of the antedependence models of order s (AD(s) or SAD(s)) need to be implemented (Núñez-Antón and Zimmerman, 2001; Zimmerman and Núñez-Antón, 2010). Estimation in any of these variance-covariance parametric models is commonly carried out by restricted maximum likelihood methods, together with the use of recursive algorithms. Some computational software packages such as, for example, SAS© or SPSS©, include the possibility of specifying some particular variance-covariance parametric choices to estimate

this type of models. Estimation in higher order AD or SAD models usually requires the use of specific numerical algorithms that need to be directly programmed in specific computing languages or available software statistical packages.

Bayesian estimation proposals for longitudinal data settings where no specific variance-covariance structure is considered in the model specification were previously developed (Brown, Kenward and Bassett, 2001), where a Wishart prior distribution was assumed for the covariance structure. In addition, the joint estimation for the mean structure and some simple covariance structures under the assumption of prior normal distribution for the mean parameters, and inverse gamma distributions for the variance in the proposed model were previously proposed (Hui and Berger, 1983). Moreover, the advantages of this proposal for fitting a growth curve model in post-menopause female bone calcium loss were also illustrated. The first proposal introduced Bayesian longitudinal models by taking into account regression structures in both the mean and the variance-covariance matrix of normal observations (Cepeda-Cuervo, 2001). This approach was based on the modelling proposal that used the Cholesky's matrix decomposition (Macchiavelli and Arnold, 1994; Pourahmadi, 1999). More specifically, by assuming normal prior distributions for the mean and variance regression structures parameters, a Bayesian methodology was introduced to fit the proposed models building the kernel transition functions from observational working variables (Cepeda-Cuervo, 2001). Results and some of the extensions of this work have also been presented by several authors (Cepeda and Gamerman, 2004; Cepeda-Cuervo and Núñez-Antón, 2007; Cepeda-Cuervo and Núñez-Antón, 2009; Cepeda-Cuervo, 2011), where, in addition, observational units are allowed to be correlated. These proposals included a detailed description of the optimization algorithms, as well as simulation and case studies that allowed for the comparison of the Bayesian and classic proposals for the analysis of this type of data. A Bayesian version of first-order multivariate antedependence model has also been developed (Jiang et al., 2015). Finally, it is interesting to briefly mention Bayesian AD models within the framework of Bayesian hierarchical mixed linear models, where, in general, authors have assumed that the errors are independent and identically distributed (i.i.d.), and also conjugate prior distributions for the parameters in the proposed models (Congdon, 2020; Gelman et al., 2014a; Gill, 2014). More specifically, by following the proposals in Congdon, 2020, some possible extensions of Bayesian hierarchical mixed linear models can be considered, so that allowing for autocorrelated errors is possible. That is, there exists the possibility of assuming that the covariance matrix of the random errors, or that of the random effects, follows an AD model (Fahrmeir, Kneib and Lang, 2013). Our proposals would allow researchers to develop these models and to extend non-Bayesian previous methods for hierarchical mixed linear models with AD structures, such as, for example, the ones in Jaffrézic and Pletcher (2000), Jaffrézic et al. (2002) and Yang and Tempelman (2012), to a Bayesian context.

In this paper we propose a Bayesian method for the joint estimation of the mean and covariance parameters in the regression longitudinal models settings under the normality assumption, and also allowing for the specification of several different variance-

covariance structures. Our proposals start by considering variance-covariance models with stationary correlations and homogeneous variances, as is the case in the CS, AR(1) and ARMA(1,1) models, so that they are then generalized to consider nonstationary correlations and heterogeneous variances, such as is the case in the structured antedependence model of order one, or SAD(1) model. That is, we extend the previous proposal (Cepeda-Cuervo, 2001) to consider parametric more parsimonious variance-covariance models that have been shown to be more useful in longitudinal data settings than those of the unstructured AD model previously considered therein. For each one of the variance-covariance structures considered here, we provide a detailed description of the estimation algorithm constructed for each specific case, including the Gibbs sampling and the Metropolis-Hasting by blocks algorithm used under each of the assumed covariance structures. In order to be able to assess the behaviour of the estimation proposed algorithms, for the specific cases of CS, AR(1) and ARMA(1,1) variance-covariance structures, a real data set analysis for the *Small Mice* balanced data set (Izenman and Williams, 1989; Weiss, 2005) is carried out. As for the SAD(1) variance-covariance structures and given that, as previously mentioned (Zimmerman and Núñez-Antón, 2010), the proposed variance-covariance model depends on the specific data sets and on their underlying structure, we compare two specific structured models based on the analysis of the *Speech Recognition* data set (Tyler et al., 1988; Núñez-Antón and Woodworth, 1994; Zimmerman, Núñez-Antón and El Barmi, 1998), and also on the analysis of the *100-Km Race* data set, kindly provided by Ian Jollife of the University of Kent (Zimmerman et al., 1998).

The paper is organized as follows. In Section 2, we introduce and describe the basic characteristics of the variance-covariance models we consider. In Section 3 we include the Bayesian longitudinal model proposals, as well as the posterior distributions and kernel transition functions for each of the models considered, which include the proposed algorithms and prior distribution assumptions required for each of them. In Section 4 we introduce and describe the data sets to be analysed, as well as the main objectives of the data set analyses. In Section 5, we analyse the different data sets under the Bayesian proposals included, describe the results and compare them with those obtained with previous classic approaches. We also present a sensitivity analysis for the estimates obtained under our proposals. Finally, in Section 6, we provide some general conclusions and final practical recommendations.

2 Some covariance structures

As already mentioned by several authors (Weiss, 2005; Núñez-Antón and Zimmerman, 2001), some of the clear advantages of parametric modelling approaches for the variance-covariance matrix in longitudinal data settings are the following: (a) they help to optimize the obtention of estimates for the parameters in the mean structure; (b) they allow to obtain the most appropriate estimates for the standard errors for the estimators

of the parameters included in the mean structure (i.e., $\boldsymbol{\beta}$); (c) in most cases, they provide a feasible and effective solution when estimating models in data sets with missing data or when times at which measurements are taken are not the same for all of the individuals in the study; and (d) estimates are still valid even for the cases where the number of observations on each individual is relatively large when compared to the number of individuals in the study.

Specific variance-covariance structures to be introduced in this paper consider that all of the variances and covariances within a given individual are functions of a vector of parameters with a small or moderate number of elements, which, as in equation (1), will be denoted by $\boldsymbol{\theta}$. That is, the covariance model $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ defines a family of possible variance-covariance matrices depending on the k-order vector of parameters $\boldsymbol{\theta}$. Parameter estimation for the covariance structure is usually carried out by maximum likelihood or restricted maximum likelihood methods (Diggle et al., 2002). One of the main challenges and problems when modelling covariance structures within longitudinal data settings is to be able to select the so-called "best-fitting" or "most appropriate" covariance structure for the specific data set under study. Most researchers agree that, in order to do so, a combination of graphical methods, exploratory descriptive analysis, as well as profile plots tools provides the necessary and required information to be able to narrow down the possible covariance structure choices to the ones that can be considered as optimal choices for the data set under study (Verbeke and Molenberghs, 2000; Fitzmaurice et al., 2009). Our suggestion for proposing or considering "reasonable" covariance structures for a specific data set, which we have followed in Section 5, can be summarized in the following items (Zimmerman and Núñez-Antón, 2010):

- Compute the means for the different time points and build a profile plot, using the `matplot` function in R, for the observations in your data set. The behaviour of the means along time will provide the user clear ideas about the type of mean function that needs to be used for the mean model. In addition, the profile plot also provides information about the possible behaviour of the variance for the different time points in the data set. Compute the variances for the different time points, as well as the correlation matrix for the corresponding data set, so that stationarity in variance and correlation can be better assessed.
- Build the corresponding ordinary scatterplot graph - OSM, using the `splom` function in the **lattice** package in R, semivariogram, or PRISM (Zimmerman, 2000), to better assess the correlation structure in your data set. These graphs are built for a saturated mean model, which considers a mean parameter for each time point. Inspection of these graphs will provide the user with a clear idea about the different covariance models that could be considered for the data set under study. The user is now able to propose a set of suitable covariance structures for the data set under study, and the best fitting covariance model will be selected on the basis of some specific goodness-of-fit criteria.

- Finally, the user can test or assess for possible mean model reductions by fitting alternative mean models and selecting the best fitting one on the basis of specific tests or goodness-of-fit criteria. Later in this section, specially for antedependence models, we describe different formulations for the covariance models considered here. For easiness of comprehension and understanding of these formulations, we recommend the use of variance-covariance formulations such as the ones in equations (4), (7), (8) and (9), which are the ones we use in the applications in Section 5, as well as in the Bayesian proposals in Section 4.

We now introduce the different variance-covariance structures, $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$, that we include in our methodological Bayesian proposals. It is worth mentioning that, besides the covariance structures introduced here, there are additional structures that interested readers may wish to read about (Weiss, 2005; Núñez-Antón and Zimmerman, 2001).

The simplest variance-covariance structure, besides the obvious independence structure which is not of real interest within these settings, corresponds to the so-called compound symmetry (CS), equicovariance or equicorrelation model, which is defined by assuming that homogeneous or constant variances in time and equal correlations between different measurements on the same subject. That is, $\text{Var}(Y_{ij}) = \sigma^2,\ j = 1,\ldots,n_i$, and $\text{Corr}(Y_{ij},Y_{il}) = \rho,\ j \neq l$. There is a heterogenous version of the CS model, CSH, where variances are allowed to change over time (Núñez-Antón and Zimmerman, 2001).

The first order autoregressive structure, AR(1), includes two covariance parameters, σ^2 and ρ, with $\text{Var}(Y_{ij}) = \sigma^2,\ j = 1,\ldots,n_i$, and ρ is the correlation parameter such that $\text{Corr}(Y_{ij},Y_{il}) = \rho^{|t_{ij}-t_{il}|},\ j \neq l$. This type of serial correlation differs from the CS model correlation because in the autoregressive model of order one, the correlation decreases as a power function of time. As can be easily seen, the AR(1) model assumes homogeneous variances and stationary correlations. That is, variances are constant over time and correlations between observations taken at equally spaced time points are also constant. There is, however, a heterogeneous version of the AR(1) model, ARH(1), where variances are allowed to change with time (Núñez-Antón and Zimmerman, 2001). Differences between the AR(1) and CS models are very difficult to assess from the exploratory analysis or the individuals' profile plots, specially when there are only few observations available per subject (Fitzmaurice et al., 2009).

The autoregressive with moving average model or order (1,1), ARMA (1,1), represents a generalization of the previous two models, CS and AR(1). In this model, the correlation between consecutive observations of the same observational unit is given by:

$$\text{Corr}(Y_{ij},Y_{il}) = \begin{cases} \phi & |j-l| = 1 \\ \phi\rho^{|t_{ij}-t_{il}|-1} & |j-l| > 1, \end{cases} \tag{2}$$

with $\text{Var}(Y_{ij}) = \sigma^2,\ j = 1\ldots,n_i$, and where ϕ, $0 < \phi < 1$, is the correlation between consecutive observations of the same observational unit, ρ, $0 < \rho < 1$, is an additional parameter, which allows the correlation to feature an exponential decreasing behaviour.

As can be seen the ARMA(1,1) model reduces to the previous models: to the CS model if $\rho = 1$, and to the AR(1) model if $\phi = \rho$; and a moving average model of order 1, MA(1), if $\rho = 0$ (Weiss, 2005). In addition, the ARMA(1,1) model also assumes homogeneous variances and stationary correlations.

The concept of antedependence was originally introduced in 1962 (Gabriel, 1962), and the antedependence models within the longitudinal data settings first defined in 2010 (Zimmerman and Núñez-Antón, 2010). Let $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})$ be the vector of measurements taken on the i-th subject, which is assumed to follow a multivariate normal distribution. The antedependence longitudinal model of order s, AD(s) with an autoregressive specification (Zimmerman and Núñez-Antón, 2010), is defined as:

$$\begin{aligned} Y_{i1} &= \mathbf{x}_{i1}^{\mathsf{T}}\boldsymbol{\beta} + \epsilon_{i1} \\ Y_{ij} &= \mathbf{x}_{ij}^{\mathsf{T}}\boldsymbol{\beta} + \sum_{k=1}^{s^*} \phi_{j,j-k}(Y_{i,j-k} - \mathbf{x}_{i,j-k}^{\mathsf{T}}\boldsymbol{\beta}) + \epsilon_{ij} \qquad j = 2, \ldots, n_i, \end{aligned} \tag{3}$$

where $\mathbf{x}_{ij}$ be a q-vector of covariates associated to Y_{ij}, $s^* = \min(s, j-1)$, the ϵ_{ij}'s are independent $N(0, \sigma_j^2)$ random variables, and σ_j^2 and $\phi_{j,j-k}$ are unstructured parameters. In this model, each variable is regressed on the previous s^* predecessors in the ordered list and, in addition, it is also allowed that autoregressive coefficients vary with time (i.e., that they depend upon j). In this sense, AD models are nostationary in both variance and correlation, because variances may vary with time and correlations between equidistant observations in time are not necessarily assumed to be constant. Specific elements of the variance-covariance matrix $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ in this model can be recursively obtained by using the well known Yule-Walker equations approach, so that, if an AD(1) model with a covariance specification is assumed (Zimmerman and Núñez-Antón, 2010), $\mathrm{Var}(\boldsymbol{Y}_i) = \boldsymbol{\Sigma}_i(\boldsymbol{\theta})$ can be specified as:

$$\begin{aligned} [\boldsymbol{\Sigma}_i(\boldsymbol{\theta})]_{kk} &= \sigma_k^2, \;\; k = 1, \ldots, n_i \\ [\boldsymbol{\Sigma}_i(\boldsymbol{\theta})]_{kl} &= \sigma_k \sigma_l \prod_{j=k}^{j=l-1} \rho_j, \;\; k < l, \;\; k, l = 1, \ldots, n_i \\ [\boldsymbol{\Sigma}_i(\boldsymbol{\theta})]_{kl} &= \boldsymbol{\Sigma}_i(\boldsymbol{\theta})_{lk}, \;\; k > l, \;\; k, l = 1, \ldots, n_i, \end{aligned} \tag{4}$$

with $\rho_j = \rho_{j,j+1}$. Antedependence models of order s can be not so parsimonious mainly because the vector of variance-covariance parameters $\boldsymbol{\theta}$ has, for $n_i = n, \forall i$, $(s+1)(2n - s)/2$ parameters (Zimmerman and Núñez-Antón, 1997). In addition, as the autoregressive coefficients and the variances of the ϵ_{ij}'s in (3) depend on the time at which measurements are taken, variances in this model are heterogeneous and correlations are nonstationary. That is, variances are allowed to change over time and correlations between observations taken at equally spaced time points are not constant and, thus, are allowed to vary. The same holds for the AD(1) model in (4).

Zimmerman and Núñez-Antón (1997) originally proposed the structured antedependence (SAD) models in 1997. Their proposed models specify that the correlation parameters are determined by a Box-Cox power function and the variances for each time point are determined by a polynomial function of not so many parameters, were able to model nonstationary correlations and variances. Moreover, Núñez-Antón and Woodworth (1994) and Zimmerman and Núñez-Antón (1997) have previously defined the specific commonly used functions for the parameters in model (3), for a general structured antedependence model or order s, SAD(s), with $n_i = n$, as:

$$\phi_{j,j-k} = \phi_k^{f(t_{ij},\lambda_k)-f(t_{i,j-k},\lambda_k)},\ j = s+1,\ldots,n;\ k = 1,\ldots,s \tag{5}$$

$$\sigma_j^2 = \sigma^2 G(t_{ij},\boldsymbol{\psi}),\ \ j = 1,\ldots,n, \tag{6}$$

or equivalently, for (4), with:

$$\rho_{j,j-k} = \rho_k^{f(t_{ij},\lambda_k)-f(t_{i,j-k},\lambda_k)},\ j = s+1,\ldots,n;\ k = 1,\ldots,s \tag{7}$$

$$\sigma_j^2 = \sigma^2 G(t_{ij},\boldsymbol{\psi}),\ \ j = 1,\ldots,n, \tag{8}$$

where

$$f(t_{ij},\lambda_k) = \begin{cases} (t_{ij}^{\lambda_k}-1)/\lambda_k, & \text{if } \lambda_k \neq 0 \\ \log(t_{ij}), & \text{if } \lambda_k = 0, \end{cases} \tag{9}$$

with $\phi_k > 0, 0 < \rho_k < 1, \forall k$, $\sigma_j^2 > 0, \forall j$, and $\{\boldsymbol{\psi} : G(t_{ij},\boldsymbol{\psi}) > 0\}$, in such a way that the variance-covariance matrix for the i-th subject, $\boldsymbol{\Sigma}_i(\boldsymbol{\theta})$, is positive definite. Here, as will be seen in the applications in Section 5, $G(t_{ij},\boldsymbol{\psi})$ is usually assumed to be a positive power or step function of time. In addition, given that the SAD models are special cases of the AD models, variances in these models are heterogeneous and correlations are also nonstationary. Equation (9) represents a Box-Cox power law. Moreover, in the SAD(1) model settings, and if measurement times are equally spaced, then the lag-one correlations (and, as a matter of fact, all same-lag correlations) are a monotone function of t: they increase if $\lambda < 1$ and decrease if $\lambda > 1$. For $\lambda = 1$, same-lag correlations remain constant and, in addition, they coincide with those of the AR(1) model. That is, the Box-Cox power law can be seen as a transformation to the time scale that effects a nonlinear deformation upon the time axis, such that correlations between measurements equidistant in the deformed scale remain constant. There are some specific special cases of the SAD(1) model that are worth mentioning:

1. **Type 1 - SAD model**: We assume an SAD(1) model as in (4), such that (8) holds with $G(t_{ij},\boldsymbol{\psi}) \equiv 1$, with $\rho_j = \rho_{j,j+1} = \rho^{f(t_{i,j+1},\lambda)-f(t_{i,j},\lambda)}$, and $f(t,\lambda)$ given by equation (9). This model assumes homogeneous variances and nonstationary correlations.
2. **Type 2 - SAD model**: We assume an SAD(1) model as in (4), such that (8) holds, with

$$G(t,\psi) = \begin{cases} 1, & \text{if } t = 1 \\ \psi, & \text{otherwise,} \end{cases} \tag{10}$$

and ρ_j as in the previous model. This model assumes a specific case of heterogeneous variances and nonstationary correlations.

3. **Type 3 - SAD model**: We propose an SAD(1) model as in (4), with ρ_j defined as in the previous models and

$$h(\sigma_j^2) = \psi_0 + \psi_1 t_{ij} + \cdots + \psi_r t_{ij}^r, \tag{11}$$

where h is an appropriately chosen link function so that the σ_j^2 variances are positive. Some authors (Zimmerman et al., 1998) have previously proposed h to be the identity link function, whereas we propose, without loss of generality, to use the logarithmic link function instead. Our proposal is more general in the sense that it does not require any additional constraints on the parameters for the variances to be positive. Moreover, this model assumes heterogeneous variances and nonstationary correlations.

3 Bayesian longitudinal model methodological proposals

Let $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{in_i})^\top$, represent the times at which observations on the i-th subject were taken, and Y_{ij} represent the observation taken on subject i at time t_{ij}, $j = 1, \ldots, n_i$. Let $\mathbf{x}_{ij}$ be a q-vector of covariates associated to Y_{ij}, so that $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{in_i})^\top$ is the $n_i \times q$ design matrix of rank q. In this way, we have that model (1) holds. Thus, if $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_m)^\top$ denotes the vector of measurements for all of the m individuals in the study, having a design matrix $\mathbf{X} = (\mathbf{X}_1^\top, \mathbf{X}_2^\top, \ldots, \mathbf{X}_m^\top)^\top$, containing the values for the covariates for all individuals, we have that:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{12}$$

where $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m)^\top$ is a vector of random errors associated to the corresponding component in the responses vector $\boldsymbol{Y}$, so that the $\boldsymbol{\epsilon}_i$'s are assumed to be independent from each other, $\boldsymbol{\epsilon} \sim MVN$ with mean $\mathbf{0}$ and block diagonal variance-covariance matrices, so that $\text{Var}(\mathbf{Y}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})$ will be a block diagonal matrix with diagonal components $\boldsymbol{\Sigma}_1(\boldsymbol{\theta}), \ldots, \boldsymbol{\Sigma}_m(\boldsymbol{\theta})$.

3.1 Prior parameter distributions

In order to provide the required details for our proposed Bayesian longitudinal method, prior distribution should be assumed for the mean and for the variance-covariance regression structure parameters (Gelman, 2006). For the mean regression parameters, we assume a q-multivariate normal distribution, so that $p(\boldsymbol{\beta}) \sim N(\mathbf{b}_0, \mathbf{B}_0)$. As for the

variance-covariance parameters, we assume a prior distribution $p(\boldsymbol{\theta})$ that will depend on the assumed covariance structure. More specifically:

1. For the CS and AR(1) models, if we let $\varphi = 1/\sigma^2$, the variance-covariance vector parameter in these models is $\boldsymbol{\theta} = (\varphi, \rho)^{\intercal}$, so that its assumed prior distribution is $p(\boldsymbol{\theta}) = p(\varphi)p(\rho)$, where:

$$p(\varphi) \equiv \text{Gamma}\left(\frac{g_0}{2}, \frac{g_0\sigma_0^2}{2}\right) \tag{13}$$

$$p(\rho) \equiv \text{Beta}(a,b), \tag{14}$$

   where g_0, σ_0^2, a and b are assumed to be known hyperparameter values (Gelman, 2006).

2. In the ARMA(1,1) structure, given that $0 < \phi < 1$ and $0 < \rho < 1$, the parameter vector is $\boldsymbol{\theta} = (\varphi, \rho, \phi)^{\intercal}$, so that its assumed prior distribution is $p(\boldsymbol{\theta}) = p(\varphi)p(\rho)p(\phi)$, where $p(\phi) \equiv \text{Beta}(a_1, b_1)$, $p(\rho) \equiv \text{Beta}(a_2, b_2)$, and, for $\varphi = 1/\sigma^2$, we have the same prior distributional assumption as in (13).

3. For the structured antedependence models, we assume the following independent prior distributions:

   (a) **Type 1 - SAD model**: In this model, assumed prior distributions for σ^2 and ρ, are as above. That is, for $\varphi = 1/\sigma^2$, we have $p(\varphi) \equiv \text{Gamma}\left(\frac{g_0}{2}, \frac{g_0\sigma_0^2}{2}\right)$ and, for ρ, we have $p(\rho) \equiv \text{Beta}(a,b)$. For λ, we assume a uniform prior distribution, so that $p(\lambda) \equiv U(-a,a)$.

   (b) **Type 2 - SAD model**: For this model, the same prior distributions as above are assumed for $\varphi = 1/\sigma^2$, ρ and λ. For ψ, if we let $\psi = \exp(\eta)$, we then assume that the prior distribution for η is such that $p(\eta) \equiv N(0, \nu^2)$.

   (c) **Type 3 - SAD model**: In this model, the same prior distributions as above are assumed for ρ and λ. For $\boldsymbol{\psi} = (\psi_0, \psi_1, \ldots, \psi_r)^{\intercal}$, we assume a multivariate prior normal distribution, so that $p(\boldsymbol{\psi}) \equiv MVN(\boldsymbol{\psi}_0, \mathbf{K}_0)$.

A final comment related to the aforementioned assumed prior distributions for the different covariance models: we believe that it is relevant to mention that, given that we do not really have prior information related to the parameters in the models, which will be the ones that will to be estimated in the applications in Section 5, we have decided to assume vague prior distributions so we do not include any prior unknown information that can generate unjustified and unnecessary changes in the posterior distributions that will be used for inferential purposes in Section 5. However, if we have prior information available for the mean regression parameters, it can be easily incorporated in the model, by assuming appropriate values for $\mathbf{b}_0$ and $\mathbf{B}_0$ in $p(\boldsymbol{\beta}) \sim N(\mathbf{b}_0, \mathbf{B}_0)$. With regard to the

variance parameters, our recommendation is that the prior information is specified as follows: (i) the gamma prior distribution for $\varphi = 1/\sigma^2$ can be specified by the mean g_0 and the variance σ_0^2 from the prior information for the parameter σ^2 available in the specific application; (ii) the parameters a and b in the beta prior distribution for ρ can be specified from their prior mean and variance for ρ available in the specific application; (iii) the parameters a and b in a more general uniform prior distribution for λ (i.e., $U(a,b)$), can be also specified from the prior information for λ available in the specific application; (iv) the prior distributions for η and ψ, or $\boldsymbol{\psi}$, in the Type 2 and Type 3 SAD models, respectively, can be directly specified from their corresponding prior means and variances information for these parameters available in the specific application.

3.2 Posterior conditional distributions and estimation proposals

Under the model assumptions, apart from a constant term, the likelihood function is given by:

$$\mathbb{L}(\boldsymbol{\beta},\boldsymbol{\theta}|\mathbf{Y}) \propto \prod_{i=1}^{m} |\boldsymbol{\Sigma}_i(\boldsymbol{\theta})|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left[(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^{\intercal}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})\right]\right\}, \tag{15}$$

where $\boldsymbol{\theta} = (\sigma^2,\rho)^{\intercal}$ in the CS and AR(1) models, $\boldsymbol{\theta} = (\sigma^2,\rho,\phi)^{\intercal}$ in the ARMA(1,1) model, and $\boldsymbol{\theta} = (\sigma^2,\boldsymbol{\rho},\lambda,\boldsymbol{\psi})$ in the SAD models. Thus, the posterior parameter distribution is given by $p(\boldsymbol{\beta},\boldsymbol{\theta}|\mathbf{Y}) \propto \mathbb{L}(\boldsymbol{\beta},\boldsymbol{\theta}|\mathbf{Y})p(\boldsymbol{\beta})p(\boldsymbol{\theta})$. Moreover, given that, under the assumed prior distribution for $\boldsymbol{\beta}$ we have that:

$$p(\boldsymbol{\beta}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\mathbf{b}_0)^{\intercal}\mathbf{B}_0^{-1}(\boldsymbol{\beta}-\mathbf{b}_0)\right\}, \tag{16}$$

the posterior conditional distribution of $\boldsymbol{\beta}$ will be $p(\boldsymbol{\beta}|\boldsymbol{\theta},\mathbf{Y}) \equiv N(\mathbf{b}^*,\mathbf{B}^*)$, where (Cepeda-Cuervo, 2001; Cepeda and Gamerman, 2004):

$$\mathbf{b}^* = \mathbf{B}^*(\mathbf{B}_0^{-1}\mathbf{b}_0 + \mathbf{X}^{\intercal}\boldsymbol{\Sigma}^{-1}\mathbf{Y}) \tag{17}$$

$$\mathbf{B}^* = (\mathbf{B}_0^{-1} + \mathbf{X}^{\intercal}\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1} \tag{18}$$

Samples of $\boldsymbol{\beta}$ are are taken from the conditional posterior distribution $p(\boldsymbol{\beta}|\boldsymbol{\theta},\mathbf{Y}) \equiv N(\mathbf{b}^*,\mathbf{B}^*)$, and accepted with probability one (Gamerman and Lopes, 2006; Gelman et al., 2014b).

3.3 Posterior conditional distributions for σ^2 and ρ in the CS and AR(1) models

Taking into account that, in the in CS and AR(1) models, the variance-covariance matrix can be written as $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \frac{1}{\varphi}\boldsymbol{C}(\rho)$, with $\varphi = 1/\sigma^2$, samples of φ and ρ are obtained from their conditional posterior distributions. More specifically, samples of φ are obtained

from its posterior distribution, given by:

$$p(\varphi) \propto \varphi^{\left(\frac{m+g_0}{2}-1\right)} \exp\left(-\varphi\frac{g_0\sigma_0^2+R}{2}\right), \tag{19}$$

where $R = (\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^{\intercal}\mathbf{C}^{-1}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$. That is, values for φ can be obtained from the conditional gamma posterior distribution Gamma $\left(\frac{m+g_0}{2}, \frac{g_0\sigma_0^2+R}{2}\right)$. For the parameter ρ, and given that its posterior distribution $p(\rho|\boldsymbol{\beta},\varphi)$ is analytically intractable for these covariance models, we propose that samples generating the posterior distribution for ρ be obtained, using the MCMC algorithm (Gamerman and Lopes, 2006; Gelman et al., 2014b), from the kernel transition function:

$$q(\rho^{(*)}|\rho^{(k)}) = \begin{cases} \rho^{(*)} \sim U(0, 2\rho^{(k)}) & \rho^{(k)} \leq 0.5 \\ \rho^{(*)} \sim U(2\rho^{(k)}-1, 1) & \rho^{(k)} > 0.5 \end{cases} \tag{20}$$

3.4 Posterior conditional distributions for σ^2, ρ and ϕ in the ARMA(1,1) model

We assume a longitudinal model as in (12) with variance-covariance structure similar to (2). As in Section 3.3, samples of $\varphi = 1/\sigma^2$ are obtained from the posterior distribution described before; that is, from the corresponding gamma posterior distribution Gamma $\left(\frac{m+g_0}{2}, \frac{g_0\sigma_0^2+R}{2}\right)$. However, for the parameters ρ and ϕ, and given that their posterior distributions are analytically intractable for the ARMA(1,1) model, we propose that samples generating the posterior distribution for ρ be obtained as before, from (20), and samples for the posterior distribution of ϕ be obtained, using the MCMC algorithm (Gamerman and Lopes, 2006; Gelman et al., 2014b), from the kernel transition function:

$$q(\phi^{(*)}|\phi^{(k)}) = \begin{cases} \phi^{(*)} \sim U(0, 2\phi^{(k)}) & \phi^{(k)} \leq 0.5 \\ \phi^{(*)} \sim U(2\phi^{(k)}-1, 1) & \phi^{(k)} > 0.5 \end{cases} \tag{21}$$

3.5 Posterior conditional distributions for σ^2, ρ, λ and ψ in the SAD(1) models

As the number of parameters in the variance-covariance matrix for the structured antedependence models of order one, SAD(1), depend on the specific selected $G(\boldsymbol{\psi},t)$ function in (8) for the type 1 and 2 models, or (11) for the type 3 model, we have to propose specific Bayesian estimation modelling approaches for each of them, which will depend on the type of SAD model being considered. Based on the types of SAD models described in Section 3.1 above, we describe the different distributions and estimation algorithms for each of them in what follows.

1. **Type 1 - SAD model**: To estimate σ^2, we use the proposal in Section 3.3, so that samples of $\varphi = 1/\sigma^2$ are obtained from the posterior distribution described above. As for the ρ parameter, we propose that samples generating their posterior distribution be obtained, using MCMC and a transition kernel such as the one in (20). Given that we have assumed a $U(-a,a)$ uniform prior distribution for the parameter λ, samples from its posterior conditional distribution are obtained by using an MCMC algorithm (Gamerman and Lopes, 2006; Gelman et al., 2014b), assuming that $\lambda^{(*)} = a(2\nu^{(*)} - 1)$, where $\nu^{(*)}$ is obtained from a kernel transition function similar to those previously defined in equations (20) and (21).

2. **Type 2 - SAD model**: Our Bayesian proposal to estimate σ^2, ρ and λ is similar to the one described for the type 1 SAD model. As for the parameter ψ, we let $\psi = \exp(\eta)$ and, in addition, assume that the prior distribution for η is such that $p(\eta) \sim N(0,\nu^2)$. In this way, the complete conditional posterior distribution is not known, so that samples for the posterior distribution of ψ can be obtained, using the MCMC algorithm (Gamerman and Lopes, 2006; Gelman et al., 2014b), from the kernel transition function:

$$\begin{aligned} q(\psi^{(*)}, \psi^{(k)}) &= \psi^{(k)} + N(0,\nu^2) \\ \psi^{(k)} &= \exp(\eta^{(k)}) \end{aligned} \tag{22}$$

3. **Type 3 - SAD model**: Given that the posterior conditional distribution for $\boldsymbol{\psi}$, $p(\boldsymbol{\psi}|\boldsymbol{\beta},\lambda,\rho,\mathbf{Y})$, is analytically intractable, we propose a kernel transition function given by the observational model obtained from $\tilde{Y}_j = \frac{1}{m-1}\sum_{i=1}^{m}(Y_{ij} - \bar{Y}_j)^2$, where $\bar{Y}_j = \frac{1}{m}\sum_{i=1}^{m} Y_{ij}$, and by assuming, without loss of generality, that $n_i = n$, and that the working observational model

$$\widetilde{w}_j = \log(\tilde{Y}_j) = \psi_0 + \psi_1 X_{1j} + \psi_2 X_{2j} + \varepsilon_j \tag{23}$$

follows a normal distribution, where $\varepsilon_j \in N(0,\sigma^2)$, with σ^2 known, and such that $\tilde{\mathbf{X}}_j = (1, X_{1j}, X_{2j})$ and $\tilde{\mathbf{X}} = (\tilde{\mathbf{X}}_1^\intercal, \ldots, \tilde{\mathbf{X}}_n^\intercal)^\intercal$. Thus, the kernel transition function for $q(\boldsymbol{\psi})$ is obtained from the combination of the normal prior distribution and the observational model in (23). That is,

$$q(\boldsymbol{\psi}|\mathbf{Y}) \equiv N(\boldsymbol{\mu}_\psi, \mathbf{K}_\psi), \tag{24}$$

where $\boldsymbol{\mu}_\psi = \mathbf{K}_\psi(\mathbf{K}_0^{-1}\boldsymbol{\psi}_0 + \tilde{\mathbf{X}}^\intercal\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{W}})$, with $\tilde{\boldsymbol{\Sigma}} = \text{diag}(\sigma^2)$, $\tilde{\mathbf{W}} = (\widetilde{w}_1, \ldots, \widetilde{w}_n)^\intercal$ and $\mathbf{K}_\psi = (\mathbf{K}_0^{-1} + \tilde{\mathbf{X}}^\intercal\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{X}})^{-1}$ (Gelman et al., 2014b).

As a final comment to the posterior inferences related to all of the models described in Section 3, and given that the proposed Bayesian inference is based on the posterior distribution of the parameters and, in addition, given that, for the models considered

here, the posterior distributions do not have a closed form expression, inferences are based on the simulation of the posterior distributions obtained by applying the MCMC algorithm (Gamerman and Lopes, 2006; Gelman et al., 2014b). A first approach to simulate samples of the posterior distribution may be to apply the Gibbs sampler algorithm, but this is possible only if all conditional posterior distributions are known. If some or all of the conditional posterior distributions are not known, as is the case in the longitudinal Bayesian models proposed here, kernel transition functions should be built in order to be able obtain samples of the unknown conditional distributions using the Metropolis Hastings algorithm. Therefore, a Metropolis-Hastings-within the Gibbs algorithm is used to be able to draw samples of the posterior distributions, from which the posterior inferences can be straightforwardly obtained. For example, in CS and AR(1) models, samples of the mean regression parameters $\boldsymbol{\beta}$ are proposed from a normal distribution with mean and variance given by equations (17) and (18), respectively, and samples of $\varphi = 1/\sigma^2$ are obtained from the gamma distribution given in equation (19), where in both cases the Gibbs sampler algorithm is used. Moreover, samples of ρ are proposed from the kernel transition function in equation (20), by applying the Metropolis Hastings algorithm. Therefore, samples of the target posterior distributions are obtained by applying an iterative algorithm, so that posterior inferences on the parameters in the models can be obtained. A proper construction of the kernel transition function is very important to improve the convergence of the chains and to be able to obtain better posterior inferences.

4 Data

4.1 Small Mice Data

The Small Mice data set (Izenman and Williams, 1989) was used to illustrate the proposals along the lines of spectral models for the analysis of longitudinal data. The study analysed more than 600 mice at birth, 7 days after birth (onset of growth), 14 days (when eyes open and consumption of solid food begins), 21 days (end of maternal influence for food) and 42 days (when most mice reach sexual maturity). Of these 600 observations a particular group of 35 male mice was divided into 4 groups. The Small Mice data constitutes a balanced set of longitudinal data with the weights in milligrams of 14 mice which make up groups 3 and 4 of the original study (Izenman and Williams, 1989). These weights were taken in the days corresponding to $t = 2,5,8,11,14,17,20$ after birth by the same person using the same measurement scale. The objective is to find a parsimonious model describing in the best possible way how weight is related to the time at which measurements were taken, and weight on previous times.

4.2 Speech Recognition Data

This set of data comes from the audiological study presented in 1988 (Tyler et al., 1988). The general study was performed with five different types of implants, three single channel implants and two multichannel implants. The implants were surgically implanted five to six weeks before connecting electrically to an external voice processor. The data includes the scores obtained when performing a speech recognition test on patients with multichannel cochlear implants. These patients were divided into two groups depending on the type of implant received (namely A and B): 21 subjects received implant A and 21 subjects received implant B. The individuals in the study were bilaterally deaf, therefore the base values of the test were all equal to zero. Measurements were taken at 4 time points: 1,9, 18 and 30 months after having received the implant. In the study there was a variation in the actual follow-up times, so these times were not exact. In addition, some subjects did not show up in one or more of their programmed follow-ups, so some data were missing (there were eight missing observations at month 18 and twenty missing observations at month 30). It was assumed observations were missing at random (Zimmerman and Núñez-Antón, 2010). The interest of studying these data focuses on describing the audiological performance of the individuals who receive each type of implant and how their performance depends on the time elapsed since implantation, as well as on the type of implant. More specifically, the goal is to assess how the average means of the types of implants are compared to each other, and, secondarily, whether the audiologic performance of a subject tends to be more consistent over time (Zimmerman and Núñez-Antón, 2010).

4.3 100-Km Race Data

These data set was kindly provided by Ian Jollife of the University of Kent, and originally analysed in 1998 (Zimmerman et al., 1998). The data correspond to each of the partial times in minutes for each of the 80 competitors in each of the 10-kilometer sections of a 100-km race in the United Kingdom in 1984. In addition to the partial times, the data features the age of 76 of the 80 competitors. Some descriptive graphs and exploratory analyses of this data have been previously reported (Everitt, 1994a; Everitt, 1994b). The objective is to find a parsimonious model describing in the best possible way how competitor's performance on each 10-km section is related to the section number $(j = 1, 2, \ldots, 10)$, and performance on previous sections.

5 Applications

In this section we illustrate the usefulness of the proposed Bayesian methodology with the statistical analysis of the three data sets that were briefly introduced in Section 4. Longitudinal models with compound symmetry, CS, autoregressive of order one, AR(1), and autoregressive with moving average, ARMA(1,1), models for the variance-

covariance structure were fitted to the first data set, and structured antedependence models of order one, SAD(1), were fitted to the last two data sets. Unless indicated otherwise, in all of the analyses reported in this section, parameter estimates were obtained from 20000 iterations, after a burn-in of 10000 samples. As specific information on the parameters prior distributions is not available, a $N(\mathbf{b}_0, \mathbf{B}_0)$ distribution was assumed, where independence between the individual distributions for each one of the parameters was assumed, with $\mathbf{b}_0 = (0, \ldots, 0)^{\top}$ and variances for each one of the distributions being equal, so that $\mathbf{B}_0 = \text{diag}(10^k)$, where $k = 5$. In addition, Beta$(1,1)$ prior distributions were assumed for the correlation parameters ρ and ϕ, a Gamma $\left(\frac{g_0}{2}, \frac{g_0\sigma_0^2}{2}\right) \equiv$ Gamma$(10^{-k}, 10^{-k})$, $k = 1, 2, \ldots$ prior distribution was assumed for the variance parameter $\varphi = 1/\sigma^2$, a $U(-1,1)$ uniform prior distribution was assumed for the time-scale transforming parameter λ in equation (9), and a $N(0, \nu^2)$, with $\nu^2 = 1$, distribution was assumed for η, in the Type 2 - SAD model, with $\varphi = \exp(\eta)$. Given that the posterior estimates of φ may change significantly for different values of k, a sensitivity analysis was performed concluding that, for our specific applications, $k = 8$ is the appropriate value minimizing this effect.

5.1 Small Mice Data

From the correlation matrix reported in Table A.1 in the Supplementary Material, it is worth mentioning that there exists a high correlation between consecutive observations or lag-one correlations, with the smallest correlation being the one corresponding to the weights taken between days 5 and 8, and the remaining ones featuring similar values. Moreover, the values for the correlations outside the super diagonal are smaller, but not negligible at all. More specifically, the lag-one correlations range from 0.77 to 0.96, with correlations not being exactly equal, but quite similar to one another, except for the 0.77 value, which is smaller than the others. Thus, it seems reasonable to consider the initial hypothesis that the lag-one correlations are approximately equal. If we move to the lag-two and lag-three correlations, they seem to be quite similar for their first two values, and then their values suddenly increase for the later values. A close analysis of this matrix seems to suggest that there may be two groups of observations, the early ones, corresponding to times 2, 5, 8 and 11, and the late ones, corresponding to times 14, 17 and 20. The former feature a pattern of high lag-one correlations, intermediate values for lag-two correlations, and low lag-three correlations, whereas the latter, if we consider observations at times 11, 14, 17 and 20, all feature high correlations. Based on the above, variance-covariance models such as the CS, AR(1) and ARMA(1,1), as described in Section 2, should be considered for this specific data set. From the correlation matrix reported in Table A.1 (see Supplementary Material), it is worth mentioning that there exists a high correlation between consecutive observations or lag-one correlations, with the smallest correlation being the one corresponding to the weights taken between days 5 and 8, and the remaining ones featuring similar values. Moreover, the values for

the correlations outside the super diagonal are smaller, but not negligible at all. More specifically, the lag-one correlations range from 0.77 to 0.96, with correlations not being exactly equal, but quite similar to one another, except for the 0.77 value, which is smaller than the others. Thus, it seems reasonable to consider the initial hypothesis that the lag-one correlations are approximately equal. If we move to the lag-two and lag-three correlations, they seem to be quite similar for their first two values, and then their values suddenly increase for the later values. A close analysis of this matrix seems to suggest that there may be two groups of observations, the early ones, corresponding to times 2, 5, 8 and 11, and the late ones, corresponding to times 14, 17 and 20. The former feature a pattern of high lag-one correlations, intermediate values for lag-two correlations, and low lag-three correlations, whereas the latter, if we consider observations at times 11, 14, 17 and 20, all feature high correlations. Based on the above, variance-covariance models such as the CS, AR(1) and ARMA(1,1), as described in Section 2, should be considered for this specific data set. Figure A.1 in the Supplementary Material displays the profiles for the different mice in the data set, where we can see that there is an increasing trend for their weights. Given the increasing structure featured by the Small Mice Data (Weiss, 2005), we assume a longitudinal model with the following mean regression structure:

$$Y_{ij} = \beta_0 + \beta_1 \text{Day} + \beta_2 \text{Day}^2 + \epsilon_{ij}, \tag{25}$$

with CS, AR(1) and ARMA(1,1) variance-covariance structures. Parameter estimates are compared to those obtained by applying restricted maximum likelihood methods and reported in Weiss (2005). Tables 1, 2 and 3 include the parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations, and including median values, as well as estimates obtained by restricted maximum Likelihood methods (REML), using the SPSS statistical software package, for the CS, AR(1) and ARMA(1,1) variance-covariance structures, respectively. To implement and obtain the estimates under the Bayesian proposal we have used OpenBugs (Spiegelhalter et al., 2003), together with R (R Core Team, 2013). Based on the estimated parameter values for the different variance-covariance models fitted to the data, we can conclude that estimates and standard deviations under the Bayesian proposals and those obtained by REML are quite similar, which can be used as evidence supporting the fact that the proposed method is behaving as expected and its results are stable under the prior distributional assumptions. However, we should be careful about these conclusions in the sense that this is a very simple, well behaved and balanced data set, and the considered variance-covariance models are, in terms of complexity, very simple and very parsimonious models. More complex variance-covariance models, such as the AD or SAD models, cannot be fitted in most statistical packages and, thus, specific programming is required to fit these models. Selection of the model that best fits the data will be assessed by using the well-known and commonly used Akaike Information Criterion (AIC) (Akaike, 1974), the Bayesian Information Criterion (BIC) (Schwarz, 1978) and

the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). Smaller values of AIC, BIC or DIC indicate better fitting models.

Table 1: *Parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML methods for the CS variance-covariance structure for the Small Mice Data.*

Parameter	Mean	Median	REML-estimates
β_0	65.200 (30.158)	64.715	65.745 (29.459)
β_1	70.776 (4.469)	70.766	70.328 (4.406)
β_2	-1.351 (0.198)	-1.350	-1.349 (0.195)
σ^2	9889.113 (2663.290)	9348.963	9671.253 (2621.890)
ρ	0.603 (0.095)	0.606	0.626 (0.107)

Table 2: *Parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML methods for the AR(1) variance-covariance structure for the Small Mice Data.*

Parameter	Mean	Median	REML-estimates
β_0	73.843 (28.129)	73.518	74.083 (28.186)
β_1	68.613 (4.089)	68.629	68.588 (3.994)
β_2	-1.252 (0.173)	-1.254	-1.251 (0.169)
σ^2	8622.617 (2663.290)	8130.642	8796.697 (2488)
ρ	0.856 (0.037)	0.857	0.874 (0.038)

Table 3: *Parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML methods for the ARMA(1,1) variance-covariance structure for the Small Mice Data.*

Parameter	Mean	Median	REML-estimates
β_0	77.556 (28.099)	77.5126	77.418 (28.494)
β_1	67.687 (4.534)	67.735	67.620 (4.317)
β_2	-1.208 (0.194)	-1.210	-1.204 (0.183)
σ^2	8169.547 (2293.082)	7684.294	8796.697 (2488)
ρ	0.792 (0.0539)	0.797	0.832 (0.055)
ϕ	0.842 (0.034)	0.845	0.8732 (0.035)

Table 4: *Goodness-of-fit AIC, BIC and DIC values for the CS, Bayesian CS-BCS, AR(1), Bayesian AR(1)-BAR(1), ARMA(1,1) and Bayesian ARMA(1,1)-BARMA(1,1) variance-covariance structures for the Small Mice Data.*

Model	AIC	BIC	DIC
CS	1109.3	1110.6	–
BCS	1124.3	1137.2	1122.9
AR(1)	1039.3	1040.5	–
BAR(1)	1054.7	1067.6	1053.3
ARMA(1,1)	1038.9	1040.8	–
BARMA(1,1)	1054.6	1070.1	1052.7

Alternative more recent model selection criteria include the Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010). Given that the main objective here is the proposal of a Bayesian methodology and its comparison with previous maximum likelihood estimation methods, we have used the AIC and BIC criteria to be able to assess and compare the performance of our models with those previously fitted. In any case, and given that the DIC and WAIC are standard model evaluation tools and considered more appropriate criteria for model selection purposes within the Bayesian framework (Watanabe, 2010; Choi, Jang and Alemi, 2018), we have also provided the DIC values for some of the models fitted here. Moreover, we believe that model selection, within the Bayesian framework, and given the common use of the DIC criterion and the well known advantages of the WAIC criterion, should be proposed together with the use of both model selection criteria (Piironen and Vehtari, 2017; Vehtari, Gelman and Gabry, 2017). However, in our view and given that our main objective is to compare and assess the behaviour of our Bayesian proposals with previous non-Bayesian approaches, we consider that the reported AIC and BIC criteria are appropriate within this context, specially taking into account that we have assumed vague prior distributions. Table 4 includes the AIC, BIC and DIC values for the CS, Bayesian CS-BCS, AR(1), Bayesian AR(1)-BAR(1), ARMA(1,1) and Bayesian ARMA(1,1) variance-covariance structures for the Small Mice Data. Based on these values and keeping in mind that these are simple models than can be fitted by using REML methods in SPSS or other alternative statistical packages, the best fitting model based on AIC is the ARMA(1,1) model, with the AR(1) model being a close competitor. If we use BIC, the best fitting model is the AR(1), with the ARMA(1,1) model being also a close competitor. The same conclusion is reached if DIC is used as a model selection criterion, with the BARMA(1,1) model being the best fitting one, and the BAR(1) following quite closely, a fact that is also supported is we use the AIC or BIC criteria for the Bayesian model proposals. In summary, for the Small Mice data, the best fitting models are the autoregressive model of order one and the autoregressive model with moving average, ARMA(1,1), model.

A final remark on the basis of a comment raised by an anonymous reviewer is that practitioners may consider using the logarithm of the weight as a response variable instead. They should be aware that when we have higher variance sample values for the different time points it may be convenient to use this transformation so that these variance values may be more parsimoniously modelled. However, given that our main objective was to compare our results to those in previous analyses (see, e.g., Weiss, 2005), and also to assess if the proposed methodology was able to model specific variance and correlation behaviours, such as the ones in the small mice data, we decided to use weight as the response variable for the analyses reported here.

5.2 Speech Recognition Data

Previous analyses (Zimmerman et al., 1998) reported that the likelihood-ratio test for the equality of the within-group covariance matrices indicated that it was reasonable to

pool them ($p = 0.35$). From the pooled correlation matrix reported in Table A.2 in the Supplementary Material, it is worth mentioning that correlations are positive and quite large, that correlations between test scores at times t and $t+k$ seem to decrease monotonically as k increases, and that correlations between test scores at adjacent measurement times increase over time. This latter statement is somehow consistent with a prior belief that subjects may "learn" over time, as with the result that responses equidistant in time become more highly correlated as the study progressed, which is a clear sign of nonstationary correlation structures, such as the one modelled by SAD-type models and the proposed time-transforming scale in equation (9). In addition, variances seem to be homogeneous at all times points except for the first one. Based on the above, variance-covariance models such as the SAD, as described in Section 2, should be considered for this specific data set. Figure A.2 in the Supplementary Material displays the profiles for the different individuals for each type of implant. As Zimmerman and Núñez-Antón (2010) have already mentioned in previously presented exploratory analyses, these plots suggest that there is an increasing trend for the mean audiologic performance, at least for the initial months, and that audiologic performance seems to stabilize for the later months, which provides some empirical evidence for the consistency of audiologic performance over time. These plots also suggest that variances seem to increase slightly from the first to the second measurement, but remain constant thereafter. Several previous different models were fitted (Zimmerman and Núñez-Antón, 2010), such as, for example, homogeneous and heterogeneous versions of the CS and AR(1) models, but finally concluded that the best fitting models models for this data are the structured antedependence model of order one or SAD(1) models. Given the increasing mean featured by this data (Núñez-Antón and Woodworth, 1994), in order to be able to compare our results when fitting the Type 1 - SAD model, we initially propose the mean regression structure:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \epsilon_{ij}, t_{i1} = 1,\ t_{i2} = 9,\ t_{i3} = 18,\ t_{i4} = 30 \tag{26}$$

with the Type 1 - SAD model variance-covariance structure described in Section 2. Table 5 includes the parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations, and including median values, as well as estimates obtained by restricted maximum Likelihood methods (REML), previously reported by Núñez-Antón and Woodworth (1994), where standard deviations for the variance-covariance parameters were not provided. Initial values for the regression parameters were assumed so that $\boldsymbol{\beta}_0 = (20, 1, 0)^{\intercal}$. In addition, initial values for the Bayesian estimation were assumed so that $\rho_0 = 0.50$, $\lambda_0 = 0.50$ and $\sigma_0^2 = 100$. The acceptance rates for ρ and λ were equal to 39% and 32%, respectively. The corresponding goodness-of-fit information criteria values for this model were AIC=1322.752, BIC=1341.351, and DIC=1320.714. Núñez-Antón and Woodworth (1994) did not report these values in earlier analyses, and their computation was not straightforward unless specific programs to fit the proposed model are implemented. This issue is clearly out of the scope

of this paper. Based on the estimated parameter values reported in Table A.2 in the Supplementary Material, we can conclude that estimates and standard deviations under the Bayesian proposal and those obtained by REML are quite similar, which can be used as evidence supporting the fact that the proposed method is behaving as expected and its results are stable under the prior distributional assumptions. In addition, fitting of this not so parameterized and parsimonious model by REML methods requires a more specific and complex programming and maximization than the ones proposed in this paper. As an illustration of fitting the Type 2 - SAD model and in order to be able to compare the results obtained with our proposed methodology, we fitted the same model previously proposed (Zimmerman et al., 1998), with mean regression structure given by:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \beta_3 z_i + \beta_4 z_i t_{ij} + \beta_5 z_i t^2 + \epsilon_{ij}, \tag{27}$$

and also with $t_{i1} = 1$, $t_{i2} = 9$, $t_{i3} = 18$, $t_{i4} = 30$, and $z_i = 1$ if the i-th individual received implant type A, and $z_i = 0$, otherwise. As for the variance-covariance structure, we assume a Type 2 - SAD model given by (4) and (10). Table 6 includes the parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations, as well as median values. Values previously obtained by restricted maximum likelihood methods (REML) were not reported (Zimmerman et al., 1998). Initial values for the regression parameters were assumed so that $\boldsymbol{\beta}_0 = (20,1,0,8,1,0)^{\top}$. In addition, initial values for the Bayesian estimation were assumed so that $\rho_0 = 0.50$, $\lambda_0 = 0.50$, $\psi_0 = 1$ and $\sigma_0^2 = 100$. Table 7 includes the parameter mean values under the Bayesian proposal for the variance-covariance parameters, together with their standard deviations, including median values, for the Type 2 - SAD variance-covariance structure, as well as those obtained by restricted maximum Likelihood methods (REML) (Zimmerman et al., 1998), where standard deviations for the variance-covariance parameters were not provided.

Table 5: *Parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML methods for the Type 1 - SAD variance-covariance structure for the Speech Recognition Data.*

Parameter	Mean	Median	REML-estimates
β_0	22.330 (4.294)	22.375	22.850 (4.260)
β_1	2.537 (0.313)	2.540	2.520 (0.340)
β_2	-0.048 (0.008)	-0.048	-0.04695 (0.009)
σ^2	602.028 (112.636)	585.562	587.15
ρ	0.933 (0.025)	0.9395	0.940
λ	0.297 (0.144)	0.300	0.300

The acceptance rates for ρ, λ and ψ were equal to 35%, 32% and 34%, respectively. The corresponding goodness-of-fit information criteria values for this model were AIC=1287.364, BIC=1318.362, and DIC=1283.554. These values were not originally reported in previous analyses (Zimmerman et al., 1998), and their computation is not

straightforward unless specific programs to fit the proposed model are implemented. This issue is clearly out of the scope of this paper. Based on the estimated parameter values reported in Table 7, we can conclude that estimates under the Bayesian proposal and those obtained by REML are comparable, except for parameter σ^2, which can be used as evidence supporting the fact that the proposed method is behaving as expected and its results are stable under the prior distributional assumptions. In addition, fitting of this not so parameterized and parsimonious model by REML methods requires a more specific and complex programming and maximization than the ones proposed in this paper. Moreover, estimates reported in Table 6 for the mean regression structure do not support the conclusions previously reported (Zimmerman et al., 1998) with regard to the significance of the parameter β_5 in (27). Given the robustness of the proposed methodology, the above differences could question the appropriateness of estimates obtained by REML methods.

Table 6: *Parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations within parentheses, including median values, for the Type 2 - SAD variance-covariance structure for the Speech Recognition Data.*

Parameter	**Mean**	**Median**
β_0	13.827 (3.959)	13.837
β_1	2.249 (0.386)	2.246
β_2	-0.044 (0.010)	-0.044
β_3	14.719 (5.663)	14.634
β_4	0.395 (0.551)	0.398
β_5	-0.009 (0.015)	-0.009

Table 7: *Parameter estimated mean values for the variance-covariance structure under the Bayesian proposal, together with their respective standard deviations in parentheses, including median values, for the Type 2 - SAD variance-covariance structure for the Speech Recognition Data.*

Parameter	σ^2	ρ	λ	ψ
Mean	334.046	0.928	0.323	1.773
Standard Deviation	71.858	0.0381	0.192	0.316
Median	325.052	0.936	0.330	1.746
REML-estimates	388.7	0.935	0.240	1.615

Table 8: *Parameter estimated mean values under the Bayesian proposal for the Type 3 - SAD variance-covariance structure, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML-methods for the 100-Km Race Data.*

Parameter	**Mean**	**Median**	**REML-estimates**
β_0	44.585 (1.632)	44.573	43.428
β_1	-2.410 (2.102)	-2.421	1.354
β_2	1.327 (0.752)	1.326	0.253
β_3	-0.097 (0.072)	-0.097	-0.017

5.3 100-Km Race Data

From the correlation matrix and sample variance values reported in Table A.3 in the Supplementary Material, it can be observed that variances tend to increase as the race progresses, that the correlations among split times are positive and quite large, that the correlations between the split time for a fixed 10-Km section and split times for successive sections tend to decrease monotonically, and that the correlations between split times for adjacent sections are smaller in the later sections of the race than in the earlier sections. In addition, variances seem to increase as the race progresses, with the exception of the seventh section of the race. Based on the above, variance-covariance models such as the SAD, as described in Section 2, should be considered for this specific data set. Figure A.3 in the Supplementary Material displays the profiles for the individuals in the data set, where we can see that there is an increasing trend for the times as the race progresses. In addition, variances for the different sections also seem to increase monotonically. Based on the above, some authors (Zimmerman et al., 1998) have previously suggested the fitting of an SAD model of order one, as well as a cubic in time mean regression model, so that:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \beta_3 t_{ij}^3 + \epsilon_{ij},\ i = 1,\ldots,80;\ j = 1,\ldots,10 \tag{28}$$

As for the variance-covariance structure, we assume a Type 3 - SAD model with variances given by $\sigma_j^2 = \exp(\psi_0 + \psi_1 t_{ij} + \psi_2 t_{ij}^2),\ \ j = 1,\ldots,10,$, and covariance structure given by (4). In the model proposal for the Type 3 - SAD model, there is a slight difference with that in previous analyses (Zimmerman et al., 1998), where the proposed variance function was, instead, $\sigma_j^2 = \sigma^2(1 + \psi_1 t_{ij} + \psi_2 t_{ij}^2),\ \ j = 1,\ldots,10$. Therefore, parameter estimates are not directly comparable. Our variance model variation was necessary for the Bayesian proposal in this paper. In the data analysis reported here, parameter estimates were obtained from 15000 iterations, after a burn-in of 5000 samples. Initial values for the regression parameters were assumed so that $\boldsymbol{\beta}_0 = (20,1,0,0)^{\intercal}$. In addition, initial values for the Bayesian estimation were assumed so that $\rho_0 = 0.50$, $\lambda_0 = 0.50$, $\boldsymbol{\psi}_0 = (1,1,1)^{\intercal}$ and $\mathbf{K}_0 = \text{diag}(k_0, k_0, k_0)$, with $k_0 = 0.1384$. Table 8 includes the regression parameter estimated mean values under the Bayesian proposal, together with their respective standard deviations, and including median values, as well as estimates obtained by restricted maximum Likelihood methods (REML) (Zimmerman et al., 1998), where standard deviations for the variance-covariance parameters were not provided. It is worth mentioning that, even though there are differences between the REML estimates and those obtained by the Bayesian proposal, the values previously reported (Zimmerman et al., 1998) for the regression parameters are all within the 95% credibility intervals listed here: $\text{CI}(0.95)_{\beta_1} = (-11.77, 6.95)$, $\text{CI}(0.95)_{\beta_2} = (-4.90, 7.547)$, and $\text{CI}(0.95)_{\beta_3} = (-0.5433, 0.3513)$, which were generated with the obtained estimated values under the Bayesian proposal. Table 9 includes the estimated values for the variance-covariance parameters under the Bayesian proposal, together with their respective standard deviations, and including median values, as well as estimates obtained by

restricted maximum Likelihood methods (REML), when available (Zimmerman et al., 1998), where standard deviations for the variance-covariance parameters were not provided. In any case and in order to be able to compare the estimated variances at each split time, we also include their REML-estimates for the variance parameters: $\hat{\sigma}^2 = 16.952$, $\hat{\psi}_1 = 0.590$, and $\hat{\psi}_2 = 0.450$.

Table 9: *Parameter estimated mean values for the variance-covariance parameters under the Bayesian proposal for the Type 3 - SAD variance-covariance structure, together with their respective standard deviations within parentheses, including median values, and parameter estimates under REML-methods, when available, for the 100-Km Race Data.*

Parameter	Mean	Median	REML-estimates
ρ	0.918 (0.031)	0.924	0.929
λ	1.680 (0.261)	1.684	1.600
ψ_0	2.771 (0.308)	2.767	–
ψ_1	0.677 (2.128)	0.683	–
ψ_2	-0.034 (0.021)	-0.034	–

The acceptance rates for ρ, λ and $\boldsymbol{\psi}$, the latter resulting from the working variable in equation (23), were equal to 34%, 37% and 84%, respectively. The corresponding goodness-of-fit information criteria values for this model were AIC = 1401.88, BIC = 1425.31, and DIC = 1403.078. These values were not reported in previous analyses (Zimmerman et al., 1998), and their computation is not straightforward unless specific programs to fit the proposed model are implemented. This issue is clearly out of the scope of this paper. In order to better assess the behaviour of the estimated split time variances obtained under the Bayesian proposal, we have computed the estimated variances under our proposal and under the REML method proposal (Zimmerman et al., 1998) and report this information, as well as the estimated sample variance values obtained from the data, in Table 10. In our opinion, it is clear that the Bayesian and REML estimated values for each of the sections in the race differ from each other, as well as from the reported sample values. However, it is worth mentioning that the observed increase for the estimated variances under the Bayesian proposal is smaller than that obtained under the REML methods.

Based on the estimated parameter values reported in Tables 8, 9 and 10, we can conclude that estimates under the Bayesian proposal and those obtained by REML are not exactly similar, but show a similar behaviour, which can be used as evidence supporting the fact that the proposed method is behaving as expected and its results are quite stable even under very general prior distributional assumptions. In addition, fitting of this not so parameterized and parsimonious model by REML methods requires a more specific and complex programming and maximization than the ones proposed in this paper. Given the robustness of the proposed methodology, the above differences could question the appropriateness of estimates obtained by REML methods. In addition, to be able to better compare mean split times estimated values under the Bayesian proposal with the corresponding fitted values that can be obtained from the estimates previously

reported (Zimmerman et al., 1998), Figure A.4 in the Supplementary Material shows the residuals obtained for the 100-Km Race Data obtained under the Bayesian and classic REML methods for the Type 3 - SAD variance-covariance structure. As can be seen from this figure, there are no significant differences between the residuals resulting from the model estimation by classic REML estimation and Bayesian estimation methods. Moreover, the residual sum of squares computed on the model estimated by restricted maximum likelihood methods is RSS $= 83455.21$, whereas the corresponding one for the proposed Bayesian method is RSS $= 94598.3$. As an additional way of comparing the behaviour of the residuals for each section of the race, Figures A.5 and A.6 in the Supplementary Material include the corresponding boxplots for the residuals resulting from the REML and Bayesian method proposals. Conclusions that can be obtained from the information provided in these figures suggest that residuals for the different sections of the race obtained by the two methods do not significantly differ from each other, which supports the claim that results obtained by the REML classic methodology can be well approximated by means of a simpler and more flexible Bayesian method, such as the one included in this manuscript.

Table 10: *Estimated sample variances, and parameter estimated variances under the Bayesian proposal for the Type 3 - SAD variance-covariance structure, and REML-methods for the 100-Km Race Data.*

Parameter	Sample values	REML-estimates	Bayesian estimates
σ_1^2	26.89	34.58	31.01
σ_2^2	34.78	67.48	54.60
σ_3^2	49.01	115.61	90.08
σ_4^2	58.89	179.013	139.28
σ_5^2	91.41	257.67	201.82
σ_6^2	149.90	351.58	274.05
σ_7^2	107.85	460.75	348.73
σ_8^2	152.22	585.18	415.86
σ_9^2	144.99	724.86	464.73
σ_{10}^2	167.21	879.80	486.70

5.4 Sensitivity analysis

In this section we study the behaviour of the Bayesian estimate for the variance σ^2 under different values of the hyperparameters used in the assumed prior distribution for $\varphi = 1/\sigma^2$, which, as already mentioned in previous sections, was assumed to be a Gamma(k,k) distribution, with $k = 10^{-5}$. In this case, we illustrate this behaviour in the analysis of the three different models (i.e., CS, AR(1) and ARMA(1,1)) fitted to the Small Mice Data (SMD) and the two SAD (i.e., Type 1 and Type 2) models fitted to the Speech Recognition Data (SRD). Changes in the estimated values for σ^2 are observed

for different values of k in the gamma distribution, such that the assumed values for the hyperparameter k for this analysis are $k = 1 \times 10^{-3}$, $k = 1 \times 10^{-5}$, $k = 1 \times 10^{-8}$ and $k = 1 \times 10^{-10}$. Table 11 includes the average variance estimated value of the chains by means of a Gibbs sample of the resulted conditional posterior distribution, together with their corresponding standard deviations in parentheses, for different values of the hyperparameter k in the prior Gamma(k,k) distribution assumed for $\varphi = 1/\sigma^2$. Fitted models correspond to the SC, AR(1) and ARMA(1,1) models for the Small Mice Data (SMD) and to the Type 1 and Type 2 - SAD models for the Speech Recognition Data (SRD). From the information reported in Table 11, we can conclude that, as the hyperparameter in the assumed gamma distribution becomes smaller, the standard deviation and estimated values obtained under the Bayesian proposal approach those obtained by the REML estimating method. In addition, and given that variance estimates and their standard deviations obtained for values of $k = 1 \times 10^{-8}$ and $k = 1 \times 10^{-10}$ are quite similar, we can conclude that once the value of k in the prior distribution is equal to 1×10^{-8}, changes in the means of the corresponding chains are negligible, and this was the main reason for the use of this specific hyperparameter value in the prior distribution assumed for the analysis of the three data sets in Section 5.

Table 11: *Estimated variances, together with their corresponding standard deviations within parentheses, for different values of the hyperparameter k in the prior* Gamma(k,k) *distribution assumed for* $\varphi = 1/\sigma^2$. *Fitted models correspond to the CS, AR(1) and ARMA(1,1) models for the Small Mice Data (SMD) and to the Type 1 and Type 2 - SAD models for the Speech Recognition Data (SRD).*

k	1×10^{-3}	1×10^{-5}	1×10^{-8}	1×10^{-10}
SMD-CS	12302.35 (4103.29)	10672.29 (3504.29)	9889.11 (2663.29)	9890.02 (2662.79)
SMD-AR(1)	10025.26 (3234.567)	9989.30 (3012.23)	8622.617 (2663.29)	8621.85 (2661.95)
SMD-ARMA(1.1)	10054.53 (2997.72)	9867.56 (2900.32)	8169.55 (2488.56)	8169.00 (2488.32)
SRD-Type 1 SAD	768.34 (172.3452)	727.02 (156.14)	602.03 (112.64)	602.24 (112.45)
SRD-Type 2 SAD	380.65 (83.579)	372.99 (78.93)	334.05 (71.86)	334.05 (71.81)

6 Conclusions and final recommendations

We have proposed alternative Bayesian longitudinal models for fitting compound symmetry, autoregressive or order one, autoregressive with moving averages, as well as unstructured and structured antedependence models for nonstationary in variance and/or correlation longitudinal data settings. Very flexible distributional prior assumptions were proposed, and the specific methods to obtain the conditional posterior distribution were

described. The usefulness of the proposed method was illustrated with the analysis of the Small Mice Data, the Speech Recognition Data and the 100-Km Race Data, and results were compared to those obtained by restricted maximum likelihood methods. Results suggested that the proposed methods behave well under general conditions, and estimated values are in line with with those obtained by classic methods. However, classic methods require specific programming, whereas the proposed Bayesian methods can be easily adjusted to the data sets under study by using very flexible and easy programming, as well as general available software, such as R and OpenBugs. Future work includes extending these proposals to more complex unstructured and structured antedependence higher order models. Finally, we would like to mention that, even though the proposed Bayesian methodology has been shown to have a fast convergence and reasonable acceptance rates, our future research in the area includes the study of acceptance rates improvements in terms of making the proposed methodology more efficient, providing at the same time recommendations useful for researchers in the area. In practice, acceptance rates are known to improve with the adequate selection of initial values from the information available in the data, as well as from the appropriate parameter selection for the prior distributions. They do so by making use of a thorough analysis of the available prior information, such as, for example, variables rank or proposed models motivation and/or parameterization. For example, if we let $\varphi' = \log(\varphi)$ and we assume a normal prior distribution for φ' instead in the CS and AR(1) models. Convergence rates and acceptance rates can also be improved by applying alternative Monte Carlo resampling methods, such as the reduced-rejection-rate method (Baldassi, 2017). This parametrization of φ can be also important for the aforementioned problem of sensitivity of the posterior variance (i.e., precision) estimates to the gamma prior distributions Gamma$(10^{-k}, 10^{-k})$ assumption, for $k = 1,2,3,4,\ldots$. Thus, when this prior distribution is assumed and no prior information on φ is available, a sensitivity analysis, like the one described in Section 5.4, should always be included in any statistical data analysis, so that the sensitivity of φ to the smallest changes in the value of k is minimized.

Acknowledgments

This work was supported by the Department of Statistics, Faculty of Sciences of the Universidad Nacional de Colombia, Ministerio de Economía y Competitividad, Agencia Estatal de Investigación (AEI), Fondo Europeo de Desarrollo Regional (FEDER), the Department of Education of the Basque Government (UPV/EHU Econometrics Research Group), and Universidad del País Vasco UPV/EHU under research grants MTM2013-40941-P (AEI/FEDER, UE), MTM2016-74931-P (AEI/FEDER, UE), IT-642-13, IT1359-19 and UFI11/03.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Baldassi, C. (2017). A method to reduce the rejection rate in Monte Carlo Markov chains. *Journal of Statistical Mechanics: Theory and Experiment*, 2017, 033301.

Brown, P.J., Kenward, M.G. and Bassett, E.E. (2001). Bayesian discrimination with longitudinal data. *Biostatistics*, 2, 417–432.

Cepeda-Cuervo, E. (2001). Modelagem da Variabilidade em Modelos Lineares Generalizados. Unpublished Math Ph.D. Thesis. Mathematics Institute, Universidade Federal do Río de Janeiro, Brazil.

Cepeda-Cuervo, E. (2011). Generalized spatio-temporal models. *SORT*, 35, 165–178.

Cepeda, E.C. and Gamerman, D. (2004). Bayesian modeling of joint regressions for the mean and covariance matrix. *Biometrical Journal*, 46, 430–440.

Cepeda-Cuervo, E. and Núñez-Antón, V. (2007). Bayesian joint modelling of the mean and covariance structures for normal longitudinal data. *SORT*, 31, 181–200.

Cepeda-Cuervo, E. and Núñez-Antón, V. (2009). Bayesian modelling of the mean and covariance matrix in normal nonlinear models. *Journal of Statistical Computation and Simulation*, 79, 837–853.

Choi, H., Jang, E. and Alemi, A.A. (2018). WAIC, but Why? Generative ensembles for robust anomaly detection. arXiv preprint arXiv:1810.01392.

Congdon, P. (2020). *Bayesian Hierarchical Models With Applications Using R*. Second edition. CRC Press, London.

Diggle, P.J., Heagerty, P., Liang, K.-Y. and Zeger, S. (2002). *Analysis of Longitudinal Data (2nd edition)*. Oxford University Press, Oxford.

Everitt, B. (1994a). Exploring multivariate data graphically: A brief review with examples. *Journal of Applied Statistics*, 21, 63–94.

Everitt, B. (1994b). *A Handbook of Statistical Analyses Using S-Plus*. Chapman and Hall, London.

Fahrmeir, L., Kneib, T. and Lang, S. (2013). Bayesian multilevel models. In: Scott, M.A., Simonoff, J.S. and Marx. B.D. (eds). *The SAGE Handbook of Multilevel Modeling*. SAGE, London, 53–71.

Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2009). *Longitudinal Data Analysis*. CRC Press, New York.

Gabriel, K.R. (1962). Ante-dependence analysis of an ordered set of variables. *Annals of Mathematical Statistics*, 33, 201–212.

Gamerman, D. and Lopes, H.F. (2006). *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference (2nd edition)*. CRC Press, New York.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, 515–533.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., (2014a). *Hierarchical Linear Models in Bayesian Data Analysis*. Chapman and Hall/CRC Press, Boca Raton, FL.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2014b). *Bayesian Data Analysis, vol. 2*. Chapman and Hall/CRC Press, Boca Raton, FL.

Gill, J. (2014). *Bayesian Hierarchical Models in Bayesian Methods: A Social and Behavioral Sciences Approach*. CRC Press, Boca Raton, FL.

Hui, S.L. and Berger, J.O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association*, 78, 753–760.

Izenman, A.J. and Williams, J.S. (1989). A class of linear spectral models and analyses for the study of longitudinal data. *Biometrics*, 45, 831–849.

Jaffrézic, F. and Pletcher, S.D. (2000). Statistical Models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics*, 156, 913–922.

Jaffrézic, F., White, I.M.S., Thompson. R. and Visscher, P.M. (2002). Contrasting models for lactation curve analysis. *Journal of Dairy Science*, 85, 968–975.

Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z. and Liu, J.-F. (2015). Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity*, 115, 29–36.

Laird, N.M. (1988). Missing data in longitudinal studies. *Statistics in Medicine*, 7, 305–315.

Macchiavelli, R.E. and Arnold, S.F. (1994). Variable-order antedependence models. *Communications in Statistics - Theory and Methods*, 23, 2683–2699.

Núñez-Antón, V. and Woodworth, G.G. (1994). Analysis of longitudinal data with unequally spaced observations and time-dependent correlated errors. *Biometrics*, 50, 445–456.

Núñez-Antón, V. and Zimmerman, D.L. (2001). Modelización de datos longitudinales con estructuras de covarianza no estacionarias: Modelos de coeficientes aleatorios frente a modelos alternativos. *Qüestiió*, 25, 225–262.

Piironen, J. and Vehtari, A. (2017). Comparison of Bayesian predictive methods for model selection. *Statistics and Computing*, 27, 711–735.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika*, 86, 677–690.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Schwarz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society - Series B*, 64, 583–616.

Spiegelhalter, D.J., Thomas, A., Best, N. and Lunn, D. (2003). *WinBUGS User Manual*. MRC Biostatistics Unit, Cambridge, UK. www.mrc-bsu.cam.ac.uk/bugs.

Tyler, R.S., Abbas, P., Tye-Murray, N., Gantz, B.J., Knutson, J.F., McCabe, B.F., Lansing, C., Brown, C., Woodworth, G., Hinrichs, J. and Kuk, F. (1988). Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. *The Laryngoscope*, 98, 1100–1106.

Vehtari, A., Gelman, A. and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out-cross-validation and WAIC. *Statistics and Computing*, 27, 1413–1432.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer, New York.

Ware, J.H. (1985). Linear models for the analysis of longitudinal studies. *The American Statistician*, 39, 95–101.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.

Weiss, R.E. (2005). *Modeling Longitudinal Data*. Springer, New York.

Yang, W. and Tempelman, R.J. (2012). A Bayesian Antedependence Model for Whole Genome Prediction. *Genetics*, 190, 1491–1501.

Zimmerman, D.L. (2000). Viewing the correlation structure of longitudinal data through a PRISM. *The American Statistician*, 54, 310–318.

Zimmerman, D.L. and Núñez-Antón, V. (1997). Modelling longitudinal and spatially correlated data: Antedependence models for longitudinal data. In: Gregoire, T.G., Brillinger, D.R., Diggle, P.J., Russell-Cohen, E., Warren, W.G. and Wolfinger, E. (eds). *Modelling Longitudinal and Spatially Correlated Data. Methods, Applications, and Future Directions*. Springer-Verlag, New York. Lecture Notes in Statistics No. 122, 63–76.

Zimmerman, D.L and Núñez-Antón, V. (2010). *Antedependence Models for Longitudinal Data*. CRC Press, New York.

Zimmerman, D.L., Núñez-Antón, V. and El Barmi, H. (1998). Computational aspects of likelihood-based estimation of first-order antedependence models. *Journal of Statistical Computation and Simulation*, 60, 67–84.

SORT 44 (1) January-June 2020, 201-220 DOI: 10.2436/20.8080.02.100

On interpretations of tests and effect sizes in regression models with a compositional predictor

Germà Coenders[1] and Vera Pawlowsky-Glahn[2]

Abstract

Compositional data analysis is concerned with the relative importance of positive variables, expressed through their log-ratios. The literature has proposed a range of manners to compute log-ratios, some of whose interrelationships have never been reported when used as explanatory variables in regression models. This article shows their similarities and differences in interpretation based on the notion that one log-ratio has to be interpreted keeping all others constant. The article shows that centred, additive, pivot, balance and pairwise log-ratios lead to simple reparametrizations of the same model which can be combined to provide useful tests and comparable effect size estimates.

MSC: 62J05, 62H15.

Keywords: Compositional regression models, CoDa, composition as explanatory, centred log-ratios, pivot coordinates, pairwise log-ratios, additive log-ratios, effect size.

1 Introduction

Compositional Data (CoDa) can be defined as positive vectors containing information about the relative importance of parts of a whole, not necessarily with a constant sum.

The CoDa tradition started with Aitchison's seminal work in Aitchison (1982), and Aitchison (1986) e.g. on chemical and geological compositions. There, only the proportion of each part or component is of interest, since absolute amounts are in general either not available or irrelevant, as they only inform about the size of the chemical or soil sample. In the last three decades, CoDa have provided a standardized toolbox for statistical analyses where the research questions concern the relative importance of magnitudes, in both hard sciences and social sciences (Coenders and Ferrer-Rosell, 2020). The term *compositional analysis* (Barceló-Vidal and Martín-Fernández, 2016) has even been coined to stress the fact that what is ultimately compositional is not the data, which

[1] Dept. of Economics, University of Girona, Faculty of Economics and Business, Campus Montilivi, 17003 Girona, Spain. E-mail: germa.coenders@udg.edu

[2] Dept. of Computer Sciences, Applied Mathematics and Statistics, University of Girona, Polytechnic IV, Campus Montilivi, 17003 Girona, Spain. E-mail: vera.pawlowsky@udg.edu

Received: July 2019

Accepted: May 2020

may even not be parts of any whole, but the research objectives, research questions or hypotheses focusing on relative importance rather than absolute values. Along similar lines, CoDa have also been defined as "arrays of strictly positive numbers for which ratios between them are considered to be relevant" without any further requirement (Egozcue and Pawlowsky-Glahn, 2019). Examples of applications to data which do not represent parts of any whole can be found in Ortells et al. (2016) and Linares-Mustarós, Coenders and Vives-Mestres, 2018. Accessible handbooks have contributed to extending the use of CoDa (Buccianti, Mateu-Figueras and Pawlowsky-Glahn, 2006; Boogaart and Tolosana-Delgado, 2013; Filzmoser, Hron and Templ, 2018; Greenacre, 2018; Pawlowsky-Glahn and Buccianti, 2011; Pawlowsky-Glahn, Egozcue and Tolosana-Delgado, 2015), as has dedicated user-friendly software (Boogaart and Tolosana-Delgado, 2013; Filzmoser et al., 2018; Greenacre, 2018; Palarea-Albaladejo and Martín-Fernández, 2015; Thió-Henestrosa and Martín-Fernández, 2005), although in many cases standard software can be used after transforming the data.

In compositional research problems, most of the basic statistical analysis tools are flawed unless they are re-expressed by means of logarithms of ratios as proposed in the so-called log-ratio CoDa methodology.

The appeal of log-ratios is that once they are computed, standard statistical methods can be used in many cases, as long as the relative character of the information is taken into account when interpreting the results. Since one component can only increase *in relative terms* if some other(s) decrease, the effects of components as explanatory in a regression model cannot be interpreted in isolation. The effect of increasing one component in relative terms unavoidably depends on which other components are reduced in its stead. We emphasise the phrase "in relative terms" according to a compositional research focus, because it could be the case that all components increase in absolute terms.

In this article we stress the importance of the notion that in ordinary-least-squares multiple regression models interpretation of a predictor is always subject to keeping all other predictors constant. In log-ratio terms, the effect of increasing one log-ratio is understood while keeping all other log-ratios constant. The fact that the same log-ratio can have different effects and interpretations depending on the manner in which the remaining log-ratios in the regression model are constructed is frequently overlooked by applied researchers. This notion may also make interpretation of log-ratios as explanatory variables differ from other statistical analyses, which is also often overlooked.

Many ways of constructing and interpreting log-ratios have been suggested in the literature, which often lead to the same predictions, residuals and goodness of fit of the model. Given this circumstance, it is difficult to provide arguments to choose among them, since "there seems to be little to distinguish between forms of comparable goodness of fit. Much discussion has turned on attempts to provide interpretations for the parameters" (Aitchison (1986), p285). In a sense, the alternative log-ratios do not lead to different models, but to different reparametrizations of one and the same model. Each reparametrization aims at one particular manner of interpreting the results. The aim of

this article is to review some of the most common alternative parametrizations and highlight their implications regarding parameter interpretation, "keeping all other log-ratios in the model constant". To the best of our knowledge, some coincidences and similarities between the interpretations of these parametrizations are reported for the first time in this article, which will hopefully help researchers find their way in the crossroads of the many methods proposed in the literature. Some of the parametrizations involve rerunning the model more than once in order to shed additional light on the meaning of parameters.

The five particular parametrizations chosen in this article are aimed at easing interpretation and all have comparable and readily interpretable effect sizes, which are obscured in compositional analyses with more complicated alternatives, whose use prevents effect size interpretation from being a common practice in the applied literature (Müller et al., 2018).

The article starts with the first parametrizations, chronologically speaking (additive and centred log-ratios) and continues with some more recently proposed alternatives. Both statistical tests and effect sizes are interpreted and compared. An illustration using one of Aitchison's classic data sets follows. The last section concludes.

2 Basic form of the regression model with an explanatory composition. Additive log-ratios

Consider a *composition* $\mathbf{x}$, i.e. a vector in the positive orthant of D-dimensional real space carrying information about the relative importance of its components. For ease of formulation and illustration, in this article we consider $D = 4$ components closed to one without loss of generality:

$$\mathbf{x} = (x_1, x_2, x_3, x_4) \in R^4_+, \text{ with } x_j > 0, j = 1,2,3,4, \sum_{j=1}^{4} x_j = 1. \tag{1}$$

The most common CoDa approach is to represent $\mathbf{x}$ in terms of logarithms of ratios among its components (Aitchison, 1986; Egozcue et al., 2003). Log-ratios may, for instance, be computed among all possible pairs of components in the so-called *pairwise log-ratios* (Aitchison, 1986; Greenacre, 2019). In this article we follow Müller et al. (2018) in computing logarithms to base 2, which make for a simple interpretation. A unit increase in the logarithm to base 2 corresponds to a twofold increase in the original magnitude.

$$\log_2\left(\frac{x_j}{x_k}\right) = \log_2\left(x_j\right) - \log_2\left(x_k\right), \text{ with } j < k, k = 2,3,4, \; j = 1,2,3. \tag{2}$$

A particularly interesting case of pairwise log-ratios is that of *additive log-ratios* (Aitchison, 1982), in which only $D-1$ pairwise log-ratios are computed with a common component in the denominator, for instance the last. This yields an invertible log-ratio covariance matrix and additive log-ratios can thus be directly used as predictors in an ordinary-least-squares regression model:

$$\log_2\left(\frac{x_j}{x_4}\right), \text{ with } j = 1,2,3. \tag{3}$$

The most useful and general expression of a log-ratio is the log-contrast (Aitchison, 1983; Aitchison and Bacon-Shone, 1984):

$$\sum_{j=1}^{4} \alpha_j \log_2(x_j), \quad \text{with } \sum_{j=1}^{4} \alpha_j = 0. \tag{4}$$

Log-contrasts led to the first formalization of a regression with a compositional explanatory variable (Aitchison and Bacon-Shone, 1984). The regression problem can be understood as obtaining the log-contrast which is maximally correlated with the dependent variable:

$$\begin{aligned} &y = \alpha_0 + \alpha_1 \log_2(x_1) + \alpha_2 \log_2(x_2) + \alpha_3 \log_2(x_3) + \alpha_4 \log_2(x_4) + \varepsilon, \\ &\text{with } \sum_{j=1}^{4} \alpha_j = 0, \end{aligned} \tag{5}$$

where ϵ follows the usual assumptions in the linear regression model. The zero-sum constraint of the coefficients in Eq. (5) reflects the fact that a component can only increase its relative importance if one or more of the others decrease. Geometrically, the zero sum constraint implies that the vector $[\alpha_1, \alpha_2, \alpha_3, \alpha_4]^T$ is orthogonal to the unit vector $[1,1,1,1]^T$ as required for a composition, which is key to the scale invariance property in CoDa (Egozcue and Pawlowsky-Glahn, 2019): multiplying the individual compositions by arbitrary positive constants will not modify the regression results. The constraint can be handled by running the regression model by ordinary least squares on the additive log-ratios (Aitchison and Bacon-Shone, 1984):

$$y = \beta_0 + \beta_1 \log_2\left(\frac{x_1}{x_4}\right) + \beta_2 \log_2\left(\frac{x_2}{x_4}\right) + \beta_3 \log_2\left(\frac{x_3}{x_4}\right) + \varepsilon. \tag{6}$$

All formulations presented in this article lead to the same coefficients when re-expressed according to Eq. (5). In the additive log-ratio case the re-expression and the fulfilment of the constraint are trivial:

$$y = \beta_0 + \beta_1 \log_2(x_1) + \beta_2 \log_2(x_2) + \beta_3 \log_2(x_3) + (-\beta_1 - \beta_2 - \beta_3) \log_2(x_4) + \varepsilon. \tag{7}$$

The α_4 coefficient corresponding to $\log_2(x_4)$ can also be obtained by rerunning the model with a different denominator in the additive log-ratio transformation, for instance as the coefficient of $\log_2(x_4/x_3)$ in a model in which all log-ratios have x_3 in the denominator.

The interpretation is as follows. The expected value of the dependent variable increases when increasing the relative importance of components with positive α coefficients in Eq. (5), especially those with the largest coefficients in absolute value, at the expense of reducing that of components with negative α coefficients (especially those with the largest coefficients in absolute value). In compositional regression models, the interpretation can never be that of "increasing one component while keeping all other components constant" because this statement is, in relative terms, nonsensical.

Overall tests of all $D-1$ effects in Eq. (6) simultaneously (e.g., joint F tests in linear regression) are invariant for all approaches described in this article. Since the composition is multivariate by nature, the joint test is normally the one with the greatest interest. Rejecting the null hypothesis means that the composition as a whole has an effect on the dependent variable.

Having said this, researchers sometimes like to test other more specific hypotheses. For this purpose, the proper interpretation of the coefficients of each log-ratio is crucial. Interpretation does change among the alternative approaches discussed in this article. In this section we interpret the coefficients of additive log-ratios. As in any multiple regression model by ordinary least squares, the effect of one log-ratio and its test is understood as the expected change in y for a one-unit change of the log-ratio when the other log-ratios are held constant (Pindyck and Rubinfeld, 1976).

Accordingly, when using logarithms to base 2, β_1 is the effect of doubling the ratio between x_1 and x_4 *while keeping all other log-ratios constant*. Keeping the second log-ratio constant means that x_2 can only vary by the same factor as x_4. Keeping the third log-ratio constant means that x_3 can only vary by the same factor as x_4. The interpretation of β_1 is thus the change in the dependent variable expected value when the ratio between x_1 and each of components 2 to D doubles. It is also the change in the dependent variable expected value when the ratio between x_1 and the geometric mean of all other components doubles, with the restriction that components 2 to D vary by a common factor. All effect sizes are hence readily interpretable and comparable: the interpretation of β_j is the change in the dependent variable expected value when the ratio between x_j and each and every of the components $x_1,...,x_{j-1},x_{j+1},...,x_D$ doubles. If we consider the fact that in relative terms one component can only increase if other components decrease, statistically testing the β_j parameter means testing if increasing the x_j component at the expense of reducing all other components by a common factor has any impact on the dependent variable.

Table 1 shows an example of a fictitious population with $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2$, and $\beta_3 = 3$ as in Eq. (6). It can be noted that, as compared to case 1, case 2 doubles the ratio of x_1 over each and every of the remaining components. As compared to case 1, case 2 increases the first log-ratio by one unit while keeping the remaining log-ratios constant.

As a result, as compared to case 1, case 2 increases the expected value E(y) by $\beta_1 = 1$. The interested reader may compare cases 3 and 4 with case 1 to arrive at β_2 and β_3. The comparison between case 5 and case 1 leads to α_4.

Regression effects can also be interpreted in terms of what in CoDa is known as the *perturbation* operator, which can be explained in brief as the product of two compositions, component-wise. Increases in log-ratios correspond to perturbations when expressed with respect to the original compositions. Therefore, Table 1 can also be interpreted in terms of perturbations. Cases 2 to 4 in the table correspond to the perturbation of the $[x_1, x_2, x_3, x_4]$ composition when increasing each log-ratio by one unit. For instance, increasing $\log_2\left(\frac{x_1}{x_4}\right)$ by one unit while keeping all other log-ratios constant is equivalent to perturbing the original composition with $[0.4, 0.2, 0.2, 0.2]$. The product of $[0.4, 0.2, 0.2, 0.2]$ and $[0.25, 0.25, 0.25, 0.25]$ yields $[0.1, 0.05, 0.05, 0.05]$ which is closed back to a unit sum as $[0.4, 0.2, 0.2, 0.2]$. The *inverse log-ratio transformation* is the manner in which log-ratios can be expressed back as the original composition. It should be clear that these perturbations are nothing other than the inverse log-ratio transformations of vectors $[1, 0, 0]$, $[0, 1, 0]$ and $[0, 0, 1]$. Indeed, using logarithms to base 2, the inverse additive log-ratio transformation of $[1, 0, 0]$ is $[2^1, 2^0, 2^0, 2^0]$ which, after closing to unit sum, equals $[0.4, 0.2, 0.2, 0.2]$.

Table 1: *Fictitious population with $\beta_0 = 0, \beta_1 = 1, \beta_2 = 2, \beta_3 = 3$. Additive log-ratios.*

Case	x_1	x_2	x_3	x_4	$\log_2\left(\frac{x_1}{x_4}\right)$	$\log_2\left(\frac{x_2}{x_4}\right)$	$\log_2\left(\frac{x_3}{x_4}\right)$	E(y)
1	0.250	0.250	0.250	0.250	0	0	0	0
2	0.400	0.200	0.200	0.200	1	0	0	1
3	0.200	0.400	0.200	0.200	0	1	0	2
4	0.200	0.200	0.400	0.200	0	0	1	3
5	0.200	0.200	0.200	0.400	-1	-1	-1	-6

It must be noted that even if the construction of the log-ratio $\log_2(x_j/x_4)$ suggests increasing x_j in relative terms to only x_4, this does not correspond to its interpretation when the composition is explanatory, because control of the other log-ratios is a key issue.

3 Regression model with explanatory centred log-ratios

Log-ratios are often computed between each component and the geometric mean of all components including itself, in the so-called *centred log-ratio* (Aitchison, 1983):

$$\log_2\left(\frac{x_j}{\sqrt[4]{x_1 x_2 x_3 x_4}}\right), \text{ with } j = 1, 2, 3, 4. \tag{8}$$

In order to prevent perfect collinearity, one centred log-ratio must be dropped from the regression equation. This is by no means a nuisance, as often argued, but is the key to the proper parameter interpretation, as shown below. Without loss of generality, if we leave out the last centred log-ratio, the model formulation is:

$$y = \beta_0 + \beta_1 \log_2 \left(\frac{x_1}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right) + \beta_2 \log_2 \left(\frac{x_2}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right) + \beta_3 \log_2 \left(\frac{x_3}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right) + \varepsilon. \tag{9}$$

By expressing Eq. (9) as a the log-contrast in Eq. (5) we obtain:

$$y = \beta_0 + \left(\beta_1 - \frac{1}{4} \sum_{j=1}^{3} \beta_j \right) \log_2 (x_1) + \left(\beta_2 - \frac{1}{4} \sum_{j=1}^{3} \beta_j \right) \log_2 (x_2) + \left(\beta_3 - \frac{1}{4} \sum_{j=1}^{3} \beta_j \right) \log_2 (x_3) + \left(-\frac{1}{4} \sum_{j=1}^{3} \beta_j \right) \log_2 (x_4) + \varepsilon. \tag{10}$$

Univariate tests referring to each particular log-ratio are interpreted as follows. Since all four centred log-ratios in Eq. (8) add-up to zero, increasing a given centred log-ratio while keeping the remaining two log-ratios in the equation constant means increasing the given centred log-ratio while decreasing the omitted centred log-ratio by the same amount. Individual coefficients and their tests thus show the existence of significant trade-offs between pairs of components. A positive significant β_j coefficient means that increasing component x_j at the expense of reducing component x_4 has a significant positive effect on the dependent variable. If we use the logarithm to base 2, β_j is interpreted as the expected change in the dependent variable when the ratio between x_j and x_4 increases fourfold. In this manner effect sizes in the model are once more readily interpretable and comparable. Table 2 has an example of a fictitious population with $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2$, and $\beta_3 = -1$ as in Eq. (9). For instance, as compared to case 1, case 2 increases the first log-ratio by one unit while keeping the remaining log-ratios constant and shows a fourfold increase in the ratio between x_1 and x_4. Compared to case 1, the ratio between x_1 and any of the remaining components (x_2 and x_3) is doubled, while the ratio between x_4 and any of the remaining components (x_2 and x_3) is halved. It must be noted that the omitted log-ratio $\log_2 \left(\frac{x_4}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right)$ is equal to -1.

Table 2: *Fictitious population with $\beta_0 = 0, \beta_1 = 1, \beta_2 = 2, \beta_3 = -1$. Centred log-ratios, where the one with x_4 in the numerator has been dropped.*

Case	x_1	x_2	x_3	x_4	$\log_2 \left(\frac{x_1}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right)$	$\log_2 \left(\frac{x_2}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right)$	$\log_2 \left(\frac{x_3}{\sqrt[4]{x_1 x_2 x_3 x_4}} \right)$	E(y)
1	0.250	0.250	0.250	0.250	0	0	0	0
2	0.444	0.222	0.222	0.111	1	0	0	1
3	0.222	0.444	0.222	0.111	0	1	0	2
4	0.222	0.222	0.444	0.111	0	0	1	-1

In order to get tests and estimates for all possible pairwise trade-offs, the model can be rerun D times by dropping each time a different centred log-ratio.

It must be noted once more that even if the construction of the log-ratio $\log_2\left(\frac{x_j}{\sqrt[4]{x_1x_2x_3x_4}}\right)$ suggests increasing x_j in relative terms to all components, this does not correspond to its interpretation when the composition is explanatory, because control of the other log-ratios is a key issue.

4 Regression model with explanatory pivot coordinates

Egozcue et al. (2003) were the first to advocate for an orthonormal basis to compute the log-ratio transformation. The advantages of this approach in many statistical analyses can be found in Pawlowsky-Glahn et al. (2015). This recommendation was translated to models with explanatory compositions by Tolosana-Delgado and Boogaart (2011) in the form of *balance coordinates* (Egozcue and Pawlowsky-Glahn, 2005). In short, balance coordinates are *scaled log-ratios* of the geometric means of two groups of components, chosen in such a way that the basis is orthonormal.

Within balance coordinates, one particular form (Egozcue et al., 2003; Fišerová and Hron, 2011; Hron, Filzmoser and Thompson, 2012) which later became known as *pivot coordinates* (Filzmoser et al., 2018), makes it possible to interpret the effect of increasing one component at the expense of decreasing all others by a common factor and has gained widespread acceptance, partly due to the unawareness that the original formulation as additive log-ratios by Aitchison and Bacon-Shone (1984) is interpreted in the same manner up to a scaling constant when used as explanatory (Coenders, 2019).

In order to provide an easily interpretable and comparable measure of effect size, Müller et al. (2018) wisely changed the requirement of orthonormality of the basis to mere orthogonality by removing scaling constants from pivot coordinates, unaware that this resulted in the same estimates and test statistics as the additive log-ratio representation. This approach was first referred to as *orthogonal coordinates for compositional regression* (Müller et al., 2018). Henceforth we refer to them as *simplified pivots*.

The first coordinate under the simplified pivot approach is the log-ratio of the first component over the geometric mean of all other components, the second is the log-ratio of the second component over the geometric mean of components 3 to D, the third is the log-ratio of the third component over the geometric mean of components 4 to D, and so forth. Constructed as just described, the following log-ratios make it possible to interpret the first log-ratio, which is the one to be called *pivot*, as the effect of increasing the first component while reducing all others by a common factor (Hron et al., 2012; Müller et al., 2018):

$$y = \beta_0 + \beta_1 \log_2\left(\frac{x_1}{\sqrt[3]{x_2x_3x_4}}\right) + \beta_2 \log_2\left(\frac{x_2}{\sqrt[2]{x_3x_4}}\right) + \beta_3 \log_2\left(\frac{x_3}{x_4}\right) + \varepsilon. \qquad (11)$$

The model can be rerun D times by permuting the components so that each time one different component plays the role of the first, which is in the numerator of the first log-ratio. The order of all other components is irrelevant. Each run provides one of the α coefficients in Eq. (5).

Table 3 shows an example of a fictitious population with $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2$, and $\beta_3 = 3$ as in Eq. (11). The reader will note that keeping the second and third log-ratios constant forces all components in the denominator of the first log-ratio to change by a common factor. Thus, as compared to case 1, case 2 doubles the ratio of x_1 over each and every of the remaining components, exactly as in Table 1. Also as compared to case 1, case 2 increases the first log-ratio by one unit while keeping the remaining log-ratios constant. Cases 3 and 4 and coefficients β_2 and β_3 are usually not interpreted in the pivot coordinate case.

Table 3: *Fictitious population with $\beta_0 = 0, \beta_1 = 1, \beta_2 = 2, \beta_3 = 3$. Simplified pivots.*

Case	x_1	x_2	x_3	x_4	$\log_2\left(\frac{x_1}{\sqrt[3]{x_2 x_3 x_4}}\right)$	$\log_2\left(\frac{x_2}{\sqrt[2]{x_3 x_4}}\right)$	$\log_2\left(\frac{x_3}{x_4}\right)$	E(y)
1	0.250	0.250	0.250	0.250	0	0	0	0
2	0.400	0.200	0.200	0.200	1	0	0	1
3	0.240	0.380	0.190	0.190	0	1	0	2
4	0.243	0.243	0.343	0.172	0	0	1	3

Since only the β_1 coefficient is interpreted in each of the D model runs, sometimes researchers compile a table including only these, which can give the misleading impression that there is only one regression model with D log-ratios while there actually are D regression models, each with $D-1$ log-ratios. We do not discuss further the estimates and tests and their interpretation because they are identical to the additive log-ratio case, albeit in the simplified pivot case, interpretation is more intuitive in accordance with the way in which the log-ratio is constructed.

5 Regression model with other explanatory orthogonal coordinates

Besides pivot coordinates, any balance coordinates can be re-expressed as orthogonal coordinates for compositional regression (Müller et al., 2018) by just dropping the scaling constants. They are thus just the logarithm of the geometric means of two groups of components, one in the numerator and one in the denominator taking care that the basis is orthogonal. As ordinary balance coordinates, they can be formed from a *sequential binary partition* of the components (Egozcue and Pawlowsky-Glahn, 2005). There are potentially many ways in which components can be partitioned, and the choice can be tailored to the research objectives. We provide only an example: we firstly partition the whole composition into the group of components x_1 and x_2 on the one hand and the group x_3 and x_4 on the other, we secondly partition the first group into its two single components, and thirdly we do likewise with the second group, according to the rows of

the following sign matrix, in which positive signs indicate the numerator of the log-ratio and negative signs the denominator:

$$\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & 0 & 0 \\ 0 & 0 & +1 & -1 \end{array}. \tag{12}$$

We get the following reparametrization:

$$y = \beta_0 + \beta_1 \log_2 \left(\frac{\sqrt[2]{x_1 x_2}}{\sqrt[2]{x_3 x_4}} \right) + \beta_2 \log_2 \left(\frac{x_1}{x_2} \right) + \beta_3 \log_2 \left(\frac{x_3}{x_4} \right) + \varepsilon. \tag{13}$$

Since it is feasible to compute potentially many sets of orthogonal coordinates by partitioning the components in different ways, the interpretation has to be tailored to the particular log-ratios. If we consider what it means to keep the second and third log-ratios constant while interpreting the first one, the estimates and tests of β_1 have to be interpreted as the effect of increasing x_1 and x_2 by a common factor and reducing x_3 and x_4 by a common factor in such a way that the ratio of the geometric means of the first pair over the second doubles (assuming we use the logarithm to base 2). A sequential binary partition chosen by the researcher as in Eq. (12) makes it possible to test the effect of jointly increasing *any subset of components* by a common factor while decreasing *any other subset of components* by a common factor. If we consider what it means to keep the first and third log-ratio constant while interpreting the second log-ratio, the estimates and tests of β_2 have to be interpreted as the effect of doubling the x_1 to x_2 ratio without modifying the relative importance of x_3 to x_4, nor the relative importance of x_3 and x_4 to x_1 and x_2 in geometric mean terms, in the same way as in the centred log-ratio case. The actual estimate is half of that obtained with the centred log-ratio and the test result is identical. The reader will note that, by coincidence, the formulation of the last log-ratio in Eq. (13) coincides with the additive log-ratio case in Eq. (6), but not its interpretation.

Table 4 has an example of a fictitious population with $\beta_0 = 0$, $\beta_1 = 1$, $\beta_2 = 2$, and $\beta_3 = -2$ as in Eq. (13). For instance, as compared to case 1, case 2 increases the first log-ratio by one unit while keeping the remaining log-ratios constant and shows a twofold increase in the ratio between x_1 and x_2 on the one hand and x_3 and x_4 on the other.

Summing up, orthogonal coordinates for compositional regression have the attractive property that effects can always be interpreted as increasing the components in the numerator by a common factor while decreasing those in the denominator by a common factor in such a way that the ratio has a twofold increase. The perturbation $[0.333, 0.333, 0.167.0.167]$ in the second row of Table 4 associated to the $\log_2 \left(\frac{\sqrt[2]{x_1 x_2}}{\sqrt[2]{x_3 x_4}} \right)$ log-ratio is a good example. In the orthogonal coordinate case, interpretation is intuitive in accordance with the way in which the log-ratio is constructed.

Table 4: *Fictitious population with $\beta_0 = 0, \beta_1 = 1, \beta_2 = 2, \beta_3 = -2$. Orthogonal coordinates for compositional regression.*

Case	x_1	x_2	x_3	x_4	$\log_2\left(\frac{\sqrt[2]{x_1 x_2}}{\sqrt[2]{x_3 x_4}}\right)$	$\log_2\left(\frac{x_1}{x_2}\right)$	$\log_2\left(\frac{x_3}{x_4}\right)$	E(y)
1	0.250	0.250	0.250	0.250	0	0	0	0
2	0.333	0.333	0.167	0.167	1	0	0	1
3	0.343	0.172	0.243	0.243	0	1	0	2
4	0.243	0.243	0.343	0.172	0	0	1	−2

Of course, the original balance coordinates which include scaling constants are equivalent up to a change in scale, which preserves the statistical test results but makes effect sizes less readily comparable.

6 Regression model with explanatory pairwise log-ratios

Greenacre (2019) suggested a general approach to selecting $D-1$ pairwise log-ratios, which, when introduced as explanatory, provide yet another flexible way of testing hypotheses that can be tailored to the research objectives. It boils down to taking care that each component participates in at least one log-ratio and that exactly $D-1$ log-ratios are computed. This results in an acyclic connected graph in which the D components act as nodes and the $D-1$ log-ratios as edges (Greenacre, 2019). Once more, this makes for a very high number of possible reparametrizations. As in the section above, we present just one example. The reader will note that the formulation of the first log-ratio coincides with the additive log-ratio case in Eq. (6), but not its interpretation.

$$y = \beta_0 + \beta_1 \log_2\left(\frac{x_1}{x_4}\right) + \beta_2 \log_2\left(\frac{x_2}{x_1}\right) + \beta_3 \log_2\left(\frac{x_3}{x_4}\right) + \varepsilon. \tag{14}$$

The log-ratios in Eq. (14) correspond to the graph in Figure 1.

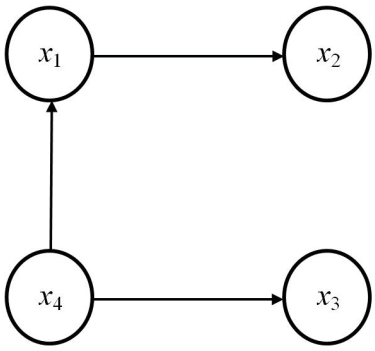

Figure 1: *Acyclic connected graph representing the pairwise log-ratios. Arrows point from the denominator to the numerator.*

When interpreting each log-ratio the researcher will have to be well aware of what the model is controlling for. If we use the logarithm to base 2, β_1 is interpreted as the effect of doubling the ratio of x_1 over x_4. However, keeping the second log-ratio constant means that x_2 can only increase by the same factor as x_1, and keeping the third log-ratio constant means that x_3 can only decrease by the same factor as x_4. The graph in Figure *1* also shows x_2 to be connected to x_1, and x_3 to x_4. Thus the estimates and tests of β_1 have to be interpreted as the effect of multiplying x_1 and x_2 by a common factor and x_3 and x_4 by a common factor in such a way that the ratio of the first pair over the second doubles. The interpretation is thus not the same as in the additive log-ratio case, even if the formulation of the log-ratio is the same as in Eq. (6).

Rearranging the components in the pairwise log-ratios makes it possible to test the effect of jointly increasing any subset of components by a common factor while decreasing *all the remaining components by a common factor*. The remaining log-ratios and the acyclic connected graph inform the researcher of which other components are linked to the numerator and which to the denominator of the log-ratio which is being interpreted, for which purpose great care has to be exerted. When used as explanatory, pairwise log-ratios are thus more closely related to orthogonal coordinates for compositional regression than previously thought, although less flexible, because orthogonal coordinates make it possible to leave certain components out of both the denominator and the numerator for interpretation.

Because of the way in which the pairwise log-ratios are computed in this particular example, the reader can apply the reasoning above to find out that the interpretation of β_2 and β_3 is the same as in the additive log-ratio case, and also corresponds to two particular simplified pivots. For instance, when interpreting β_2, keeping the first and third log-ratios constant implies that x_4 and x_3 vary by the same factor as x_1, respectively, while no component varies by the same factor as x_2. The graph in Figure *1* also shows x_4 and x_3 to be connected to x_1. β_2 thus refers to doubling the ratio of x_2 over all other components assuming that they decrease by a common factor.

Table 5 shows an example of a fictitious population with $\beta_0 = 0$, $\beta_1 = 2$, $\beta_2 = 1$, and $\beta_3 = -1$ as in Eq. (14), which illustrates the interpretations above.

Table 5: *Fictitious population with $\beta_0 = 0, \beta_1 = 2, \beta_2 = 1, \beta_3 = -1$. Pairwise log-ratios.*

Case	x_1	x_2	x_3	x_4	$\log_2\left(\frac{x_1}{x_4}\right)$	$\log_2\left(\frac{x_2}{x_1}\right)$	$\log_2\left(\frac{x_3}{x_4}\right)$	E(y)
1	0.250	0.250	0.250	0.250	0	0	0	0
2	0.333	0.333	0.167	0.167	1	0	0	2
3	0.200	0.400	0.200	0.200	0	1	0	1
4	0.200	0.200	0.400	0.200	0	0	1	−1

It must be noted again that even if the construction of the log-ratio $\log_2(x_j/x_k)$ suggests increasing x_j in relative terms to only x_k, this does not correspond to its interpretation when the composition is explanatory, because control of the other log-ratios is a key issue.

It must be reminded that cases 2 to 4 in all Tables 1 to 5 can also be interpreted in terms of the perturbation of the $[x_1, x_2, x_3, x_4]$ composition when increasing each log-ratio by one unit while keeping all other log-ratios constant. In all cases, the perturbations are obtained as the inverse transformations of vectors $[1,0,0]$, $[0,1,0]$ and $[0,0,1]$ according to the given log-ratio transformation.

7 Illustration

As an illustration we use one of the original simulated data sets provided by Aitchison (1986), called *Bayesite*, which is freely available in the R library compositions (Boogaart and Tolosana-Delgado, 2013).

In the development of bayesite, a new fibreboard, experiments were conducted to obtain some insight into the nature of the relationship of its permeability (measured in microdarcies) to the mix of its four ingredients ($n = 21$):

- Short fibres (x_1).
- Medium fibres (x_2).
- Long fibres (x_3).
- Binder (x_4)

All model parametrizations have an intercept term equal to 317.302, a residual standard error equal to 46.61 on 17 degrees of freedom, multiple R-squared equal to 0.419, adjusted R-squared equal to 0.316, and a significant joint F statistic (4.078 on 3 and 17 degrees of freedom, p-value = 0.024), telling that the mix as a whole has an impact on permeability.

Table 6 shows the coefficients. Those in italics are either redundant or not needed for interpretation. The most correlated log-contrast with permeability is also the same for all parametrizations:

$$19.414\log_2(x_1) + 27.406\log_2(x_2) - 25.953\log_2(x_3) - 20.866\log_2(x_4)\,. \qquad (15)$$

This means that permeability increases together with increases of x_1 and x_2 coupled with decreases in x_3 and x_4, in terms of relative importance, x_2 and x_3 having a greater impact than x_1 and x_4.

The additive log-ratio results tell that increasing the relative importance of x_2 at the expense of reducing all other components by a common factor in such a way that the ratio of x_2 over any other component doubles, leads to an expected increase of 27.406 microdarcies in permeability, which is statistically significant at the 0.05 level. The results also tell that relatively increasing x_3 at the expense of reducing all other components by a common factor in such a way that the ratio of x_3 over any other component doubles, leads to a significant expected decrease of 25.953 microdarcies.

Table 6: *Estimates and tests in four alternative reparametrizations. Redundant or not needed effects in italics.*

Additive log-ratios:	Estimate	Std. Error	t-value	p-value
Denominator x_4				
Numerator x_1	19.414	11.560	1.679	0.111
Numerator x_2	27.406	11.560	2.371	0.030
Numerator x_3	−25.953	11.560	−2.245	0.038
Denominator x_3				
Numerator x_1	*19.414*	*11.560*	*1.679*	*0.111*
Numerator x_2	*27.406*	*11.560*	*2.371*	*0.030*
Numerator x_4	−20.866	16.229	−1.286	0.216
Centred log-ratios:				
Log-ratio with x_4 numerator omitted				
Numerator x_1	40.281	23.929	1.683	0.111
Numerator x_2	48.272	23.929	2.017	0.060
Numerator x_3	−5.087	23.929	−0.213	0.834
Log-ratio with x_3 numerator omitted				
Numerator x_1	45.368	17.694	2.564	0.020
Numerator x_2	53.360	17.694	3.016	0.008
Numerator x_4	*5.087*	*23.929*	*0.213*	*0.834*
Log-ratio with x_2 numerator omitted				
Numerator x_1	−7.991	17.694	−0.452	0.657
Numerator x_3	*−53.360*	*17.694*	*−3.016*	*0.008*
Numerator x_4	*−48.272*	*23.929*	*−2.017*	*0.060*
Log-ratio with x_1 numerator omitted				
Numerator x_2	*7.991*	*17.694*	*0.452*	*0.657*
Numerator x_3	*−45.368*	*17.694*	*−2.564*	*0.020*
Numerator x_4	*−40.281*	*23.929*	*−1.683*	*0.111*
Simplified pivots:				
x_1 in the first place				
Pivot	19.414	11.560	1.679	0.111
Second log-ratio	*33.877*	*11.541*	*2.935*	*0.009*
Third log-ratio	*−2.544*	*11.964*	*−0.213*	*0.834*
x_2 in the first place				
Pivot	27.406	11.560	2.371	0.030
Second log-ratio	*28.550*	*11.541*	*2.474*	*0.024*
Third log-ratio	*−2.544*	*11.964*	*−0.213*	*0.834*
x_3 in the first place				
Pivot	−25.953	11.560	−2.245	0.038
Second log-ratio	*10.763*	*11.541*	*0.933*	*0.364*
Third log-ratio	*24.136*	*11.964*	*2.017*	*0.060*
x_4 in the first place				
Pivot	−20.866	16.229	−1.286	0.216
Second log-ratio	*12.459*	*10.216*	*1.220*	*0.239*
Third log-ratio	*26.680*	*8.847*	*3.016*	*0.008*
Other orthogonal coordinates (example)				
$\log_2\left(\sqrt{x_1x_2}/\sqrt{x_3x_4}\right)$	46.820	14.880	3.147	0.006
$\log_2(x_1/x_2)$	−3.996	8.847	−0.452	0.657
$\log_2(x_3/x_4)$	−2.544	11.964	−0.213	0.834
Pairwise log-ratios (example)				
$\log_2(x_1/x_4)$	46.820	14.880	3.147	0.006
$\log_2(x_2/x_1)$	27.406	11.560	2.371	0.030
$\log_2(x_3/x_4)$	−25.953	11.560	−2.245	0.038

The centred log-ratio formulation shows that increasing x_1 at the expense of reducing x_3 and increasing x_2 at the expense of reducing x_3 both lead to a significant increase in permeability. Doubling x_1 at the expense of halving x_3 (i.e., multiplying their ratio by four) leads to a 45.368 microdarcy increase in expected permeability, while multiplying the ratio between x_2 and x_3 by four leads to a 53.360 increase in expected permeability.

The results with simplified pivots are identical numerically and interpreted in the same way as those with additive log-ratios.

In this particular example, the second and third orthogonal coordinates are trade-offs between pairs of components and are thus related to the results of a centred log-ratio (estimates are halved but test statistics are identical). For instance, the test statistic for the coordinate $\log_2(x_1/x_2)$ is equivalent to the x_1 statistic in the centred log-ratio formulation with x_2 omitted.

Also in this particular example, the second and third pairwise log-ratios provide the same result as in the additive log-ratio case. Researchers need to carefully tailor interpretation to the particular log-ratios chosen, especially in the pairwise case. For instance, keeping the first and third pairwise log-ratios constant while increasing the second implies increasing the ratio of x_2 over all other components by the same factor.

Finally, the results of the first log-ratio both in the particular pairwise log-ratio example and the particular orthogonal coordinate example we have chosen, show that the effect of multiplying x_1 and x_2 by a common factor and x_3 and x_4 by another common factor in such a way that the ratio of the geometric mean of the first pair over the second doubles is significant, and amounts to 46.820, in terms of expected permeability in microdarcies.

8 Discussion

One attractive feature of CoDa is that once the raw composition has been transformed into log-ratios, classical statistical techniques for unbounded data can, in many cases, be applied in the usual way, and even with standard software. Log-ratio transformations thus constitute the easy way out in compositional problems. This includes models in which the composition is the explanatory variable. The applied researcher can concentrate his or her efforts in interpreting the results taking the compositional nature of the data and the research questions into account: what does increase at the expense of decreasing what? Along these lines, some quick and useful highlights to recap the article are:

- All alternatives considered in this article are reparametrizations of the same model. In a sense, none can be worse or better than any other as long as the parametrization provides answers to the researcher's questions and, above all, is interpreted correctly. If the researcher wants to interpret the results from more than one perspective or to test more than one type of hypotheses, he or she can use more than one parametrization.

- When used as explanatory variables, additive log-ratios are not interpreted as increasing a component at the expense of reducing the last component, as their formulation suggests, but as increasing a component at the expense of reducing *all other components*.
- When used as explanatory variables, centred log-ratios are not interpreted as increasing a component at the expense of reducing all other components, as their formulation suggests, but as increasing a component at the expense of reducing *the component whose log-ratio is omitted*.
- When used as explanatory variables, simplified pivot coordinates are equivalent to additive log-ratios.
- Orthogonal coordinates can be tailored to testing particular hypotheses of interest related to increasing any subset of components at the expense of reducing any other subset. Moreover, the interpretation of the regression coefficients is intuitive following the formulation of the corresponding log-ratios.
- When used as explanatory variables, pairwise log-ratios are not interpreted as increasing a component at the expense of reducing another component, as their formulation suggests, but as a tailored tool to interpret the effect of increasing a subset of components at the expense of reducing all the remaining components. Proper interpretation requires exerting great care.
- It often pays to embed theoretical knowledge or research questions into (possibly more than one) parametrizations of the model.
- As in any multiple regression, the full formulation of the model has much to tell about the log-ratio whose effect is being interpreted.
- Using logarithms to base 2 and removing scaling constants enhances interpretability and provides comparable effect size estimates.
- Expressing the effects of the log-ratios as the effects of the corresponding perturbations may help clarify their interpretation under all approaches, and even more so when tailoring orthogonal coordinates and pairwise log-ratios to the research objectives. The effect of the first log-ratio in the regression equation is that of perturbing the composition with the corresponding inverse log-ratio transformation of vector $[1,0,0,0,...,0]$, the second log-ratio refers to perturbing the composition with the inverse of vector $[0,1,0,0,...,0]$, and so on.

Great care must be taken if using the test results for simplifying the model (Pawlowsky-Glahn et al., 2015). The significance of one log-ratio depends both on the components present in the analysis and on the remaining $D-2$ log-ratios, which jointly frame the interpretation as the significant effect of relatively increasing what and how at the expense of relatively decreasing what and how. If we put it otherwise, dropping log-ratios changes the interpretation and estimates of whatever is left in the model. For instance, if we drop the second and third log-ratios in the orthogonal coordinate case in Eq. (13), then the coefficient of the first log ratio loses its original sharpness and shifts its interpretation into merely increasing the ratio of the product x_1x_2 over the product x_3x_4,

without knowing if increases and decreases are by a common factor. This is so because the perturbation can no longer be computed from the inverse transformation. If we put it yet otherwise, all methods are interpreted with respect to the given set of components in $\mathbf{x}$. For instance, if we drop x_1 in the additive log-ratio case in Eq. (6), then the interpretation of the coefficient of $\log_2(x_2/x_4)$ is the outcome of increasing x_2 while decreasing only x_3 and x_4 by a common factor. Besides that, different parametrizations will unavoidably suggest different simplifications, and the simplified models will no longer be equivalent. In the bayesite example, the pivot and additive log-ratio approaches suggest that x_1 may be dropped from the composition whereas the centred log-ratio approach reveals a significant trade-off between x_1 and x_3. The centred log-ratio approach suggests to drop x_4 while in the pairwise log-ratio approach all the coefficients are significant. The easy way out is not to simplify the model at all. Of course, if the research is carried out for predictive or exploratory purposes rather than for theory building or theory testing, then simplifying the model can be the wise path to follow (see below) and parameter interpretation may not be essential.

We have not dealt with the diagnostic tools used in linear regression models with a compositional predictor because they are the same as in the general linear regression case, according to the distributional assumptions for the ϵ disturbance term, for instance the normal distribution. Any of the parametrizations can be obtained from any other parametrization by linear transformations. Since the regression model is affine equivariant, this implies that all parametrizations lead to the same goodness of fit, residuals, predicted values, and even leverage values and Cook's distances (Filzmoser et al., 2018). Conversely, the statistical distribution of the compositional variables plays no specific role. For this reason, orthogonality or isometry do not constitute requirements for using compositions as explanatory variables.

Having said this, orthogonal isometric log-ratios, among which balance coordinates constitute a common example, have very desirable properties in other compositional analyses, and can be blindly applied with virtually any statistical method. In the explanatory role, any orthogonal coordinates, isometric or not, also have the attractive property that effects can always be interpreted as increasing the components in the numerator by a common factor while decreasing those in the denominator by a common factor. Both advantages have no doubt contributed to their widespread use.

The extension from a linear model to a generalized linear model is straightforward (Coenders, Martín-Fernández and Ferrer-Rosell, 2017). For instance, if the dependent variable is a count, a *Poisson regression* can be specified, or if the dependent variable is ordinal or binary, an ordered or a binary *logit model* can be specified. Interpretation would then refer to the log expected count, to the logit, or to the appropriate expression in each case, taking the link function of the generalized linear model into account.

Adding non-compositional predictors in the same model can also be done in a straightforward manner (Coenders et al., 2017) and nested models can be used to assess the predictive power of the compositional versus non-compositional predictors. The results of the non-compositional predictors are invariant under any of the parametrizations of

the composition presented in this article. The interpretation of the compositional predictors is the same as outlined in this article "keeping the non-compositional predictors constant". A very interesting particular case is including the total as predictor, which Coenders et al. (2017) recommend doing when the composition does not have a constant sum.

This article is by no means comprehensive. We have purposely selected only the simplest parametrizations with comparable effect sizes and leading to the same predictions. There are other ways to introduce a composition as explanatory in a regression model. A first group of methods (stability-based model selection, stepwise selection of the pairwise log-ratios with the highest explanatory power, spike-and-slab lasso regression modelling, principal balances, selection of the balance coordinate with highest explanatory power, and compositional principal component analysis, among others) always simplify the model, each in its own way, and thus lead to different predictions, do not control for all possible components or all possible log-ratios, and modify the interpretation. The interested reader may resort to the original sources (Combettes and Müller, 2019; Greenacre, 2019; Lin et al., 2014; Louzada, Shimizu and Suzuki, 2019; Martín-Fernández et al., 2018; Quinn and Erb, 2020; Rivera-Pinto et al., 2018; Solans et al., 2019). These data-driven approaches are especially useful when the number of components is very large, sometimes even larger than the sample size, when the model is built with predictive purposes, or when theory is weak and the researcher prefers to embrace a data mining perspective. A second group of methods does not imply simplifying the model. Among them we highlight interpreting the effects of balance coordinates, which up to a scaling constant are equivalent to those of orthogonal coordinates for compositional regression (Pawlowsky-Glahn et al., 2015), comparing predictions with different composition values (Dumuid et al., 2019), converting estimates into a gradient (Tolosana-Delgado and Boogaart, 2011), and converting estimates into elasticities (Morais, Thomas-Agnan and Simioni, 2018).

Having said this, we hope that by focusing on the most simple alternatives and on the comparative interpretation of their effect sizes and tests, we make it easier for researchers to draw fruitful, precise and clear conclusions about the influence of a composition on a dependent variable. Especially, a focus on effect sizes is as of now lacking in most applications, with few exceptions.

Acknowledgements

The article was supported by the Spanish Ministry of Science, Innovation and Universities/FEDER (grant RTI2018-095518-B-C21), the Spanish Ministry of Health (grant CIBER CB06/02/1002) and the Catalan Government (grant 2017SGR656). Acknowledgements are due to Josep Antoni Martín Fernández, Michael Greenacre and Juan José Egozcue for their comments on previous drafts, to Karel Hron, Michael Greenacre, John Bacon-Shone, and Raimon Tolosana-Delgado for fruitful discussions at the 8th Interna-

tional Workshop on Compositional Data Analysis, and to Berta Ferrer-Rosell and Esther Martínez Garcia, for keeping asking "is that effect large or small?" in all articles we co-authored.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44, 139–177.

Aitchison, J. (1983). Principal component analysis of compositional data. *Biometrika*, 70, 57–65.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability*. Chapman and Hall, London.

Aitchison, J. and Bacon-Shone, J. (1984). Log contrast models for experiments with mixtures. *Biometrika*, 71, 323–330.

Barceló-Vidal, C. and Martín-Fernández, J.A. (2016). The mathematics of compositional analysis. *Austrian Journal of Statistics*, 45, 57–71.

Boogaart, K.G. Van den and Tolosana-Delgado, R. (2013). *Analyzing Compositional Data with R*. Springer, Berlin.

Buccianti, A., Mateu-Figueras, G. and Pawlowsky-Glahn, V. (2006). *Compositional Data Analysis in the Geosciences: from Theory to Practice*. Geological Society, London.

Coenders, G. (2019). Compositional explanatory variables: which are the differences between alr and pivot coordinates? *8th International Workshop on Compositional Data Analysis CoDaWork 2019*. Terrassa, Spain, 3-8 June 2019. https://doi.org/10.13140/RG.2.2.22987.44325

Coenders, G. and Ferrer-Rosell, B. (2020). Compositional data analysis in tourism. Review and future directions. *Tourism Analysis*, 25, 153–168.

Coenders, G., Martín-Fernández, J.A. and Ferrer-Rosell, B. (2017). When relative and absolute information matter. Compositional predictor with a total in generalized linear models. *Statistical Modelling*, 17, 494–512.

Combettes, P.L. and Müller, C.L. (2019). Regression models for compositional data: General log-contrast formulations, proximal optimization, and microbiome data applications. *arXiv*,1903.01050.

Dumuid, D., Pedišić, Ž., Stanford, T.E., Martín-Fernández, J.A., Hron, K., Maher, C.A., Lewis, L.K. and Olds, T. (2019). The compositional isotemporal substitution model: a method for estimating changes in a health outcome for reallocation of time between sleep, physical activity and sedentary behaviour. *Statistical Methods in Medical Research*, 28, 846-857.

Egozcue, J.J. and Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology*, 37, 795–828.

Egozcue J.J. and Pawlowsky-Glahn, V. (2019). Compositional data: the sample space and its structure. *TEST*, 28, 599–638.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G. and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35, 279–300.

Filzmoser, P., Hron, K. and Templ, M. (2018). *Applied Compositional Data Analysis with Worked Examples in R*. Springer, New York.

Fišerová, E. and Hron, K. (2011). On interpretation of orthonormal coordinates for compositional data. *Mathematical Geosciences*, 43, 455–468.

Greenacre, M. (2018). *Compositional Data Analysis in Practice*. Chapman and Hall/CRC press, New York.

Greenacre, M. (2019). Variable selection in compositional data analysis using pairwise logratios. *Mathematical Geosciences*, 51, 649–682.

Hron, K., Filzmoser, P. and Thompson, K. (2012). Linear regression with compositional explanatory variables. *Journal of Applied Statistics*, 39, 1115–1128.

Lin, W., Shi, P., Feng, R. and Li, H. (2014). Variable selection in regression with compositional covariates. *Biometrika*, 101, 785–797.

Linares-Mustarós, S., Coenders, G. and Vives-Mestres, M. (2018). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting*, 40, 1–10.

Louzada, F., Shimizu, T.K.O. and Suzuki, A.K. (2019). The Spike-and-Slab Lasso regression modeling with compositional covariates: An application on Brazilian children malnutrition data. *Statistical Methods in Medical Research*. https://doi.org/10.1177/0962280219863817

Martín-Fernández, J.A., Pawlowsky-Glahn, V., Egozcue, J.J. and Tolosona-Delgado, R. (2018). Advances in principal balances for compositional data. *Mathematical Geosciences*, 50, 273–298.

Morais, J., Thomas-Agnan, C. and Simioni, M. (2018). Interpretation of explanatory variables impacts in compositional regression models. *Austrian Journal of Statistics*, 47, 1–25.

Müller, I., Hron, K., Fišerová, E., Šmahaj, J., Cakirpaloglu, P. and Vančáková, J. (2018). Interpretation of compositional regression with application to time budget analysis. *Austrian Journal of Statistics*, 47, 3–19.

Ortells, R., Egozcue, J.J., Ortego, M.I. and Garola, A. (2016). Relationship between popularity of key words in the Google browser and the evolution of worldwide financial indices. In: Martín-Fernández, J.A. and Thió-Henestrosa, S. (eds), *Compositional Data Analysis. Springer Proceedings in Mathematics & Statistics*, Vol. 187. Springer, Cham.

Palarea-Albaladejo, J. and Martín-Fernández, J.A. (2015). zCompositions—R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, 143, 85–96.

Pawlowsky-Glahn, V. and Buccianti, A. (2011). *Compositional Data Analysis. Theory and Applications*. Wiley, New York.

Pawlowsky-Glahn, V., Egozcue, J.J. and Tolosana-Delgado, R. (2015). *Modelling and Analysis of Compositional Data*. Wiley, Chichester.

Pindyck, R.S. and Rubinfeld, D.L. (1976). *Econometric Models and Economic Forecasts*. MacGraw-Hill, New York.

Quinn, T. and Erb, I. (2020). Interpretable log contrasts for the classification of health biomarkers: a new approach to balance selection. *mSystems*, 5, e00230–19.

Rivera-Pinto, J., Egozcue, J.J., Pawlowsky-Glahn, V., Paredes, R., Noguera-Julian, M. and Calle, L. (2018). Balances: a new perspective for microbiome analysis. *mSystems*, 3, e00053–18.

Solans, M., Coenders, G., Marcos-Gragera, R., Castelló, A., Gràcia-Lavedan, E., Benavente, Y., Moreno, V., Pérez-Gómez, B., Amiano, P., Fernández-Villa, T., Guevara, M., Gómez-Acebo, I., Fernández-Tardón, G., Vanaclocha-Espi, M., Chirlaque, M.D., Capelo, R., Barrios, R., Aragonés, N., Molinuevo, A., Vitelli-Storelli, F., Castilla, J., Dierssen-Sotos, T., Castaño-Vinyals, G., Kogevinas, M., Pollán, M. and Saez, M. (2019). Compositional analysis of dietary patterns. *Statistical Methods in Medical Research*, 28(9), 2834–2847.

Thió-Henestrosa, S. and Martín-Fernández, J.A. (2005). Dealing with compositional data: The freeware CoDaPack. *Mathematical Geology*, 37, 773–793.

Tolosana-Delgado, R. and Boogaart, K.G. Van den (2011). Linear models with compositions in R. In: Pawlowsky-Glahn, V. and Buccianti, A. (eds), *Compositional Data Analysis. Theory and Applications*. Wiley, New York.