Received: 13 December 2016 | Revised: 4 July 2017 | Accepted: 7 July 2017

DOI: 10.1002/bimj.201600253

RESEARCH PAPER

Biometrical Journal

Bayesian estimation of multivariate normal mixtures with covariate-dependent mixing weights, with an application in antimicrobial resistance monitoring

Stijn Jaspers[1] | Arnošt Komárek[2] | Marc Aerts[1]

[1]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, BE3590 Diepenbeek, Belgium

[2]Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, CZ-186 75 Praha 8-Karln, Czech Republic

Correspondence
Stijn Jaspers, Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, BE3590 Diepenbeek, Belgium.
Email: stijn.jaspers@uhasselt.be

Funding Information
Research Foundation Flanders, Grant no. 11E2913N; IAP Research Network P7/06

Abstract

Bacteria with a reduced susceptibility against antimicrobials pose a major threat to public health. Therefore, large programs have been set up to collect minimum inhibition concentration (MIC) values. These values can be used to monitor the distribution of the nonsusceptible isolates in the general population. Data are collected within several countries and over a number of years. In addition, the sampled bacterial isolates were not tested for susceptibility against one antimicrobial, but rather against an entire range of substances. Interest is therefore in the analysis of the joint distribution of MIC data on two or more antimicrobials, while accounting for a possible effect of covariates. In this regard, we present a Bayesian semiparametric density estimation routine, based on multivariate Gaussian mixtures. The mixing weights are allowed to depend on certain covariates, thereby allowing the user to detect certain changes over, for example, time. The new approach was applied to data collected in Europe in 2010, 2012, and 2013. We investigated the susceptibility of *Escherichia coli* isolates against ampicillin and trimethoprim, where we found that there seems to be a significant increase in the proportion of nonsusceptible isolates. In addition, a simulation study was carried out, showing the promising behavior of the proposed method in the field of antimicrobial resistance.

KEYWORDS

antimicrobial resistance, censored data, clustering, multivariate normal mixture

1 | INTRODUCTION

Infectious diseases, caused by pathogenic microorganisms or pathogens such as viruses, bacteria, parasites, and fungi, are one of the leading causes of death worldwide. The accidental discovery and isolation of penicillin by Sir Alexander Fleming changed the entire direction of approaches to treating infectious diseases and saved the lives of millions of people (Bennett & Chung, 2001; Ligon, 2004). Ever since, different antibiotics and antimicrobials have been discovered and applied widely in both human and animal populations to treat, prevent, and control infectious diseases. Nevertheless, in the past few decades, there has been a decrease in the number of antimicrobials that are effective in treating infections and antimicrobial resistance (AMR) has become one of the main public health concerns (Cohen, 2000; Spellberg et al., 2008). Not only does AMR result into an increased morbidity and mortality, it is also known to be associated with increased hospital and societal costs (Mauldin, Salgado, Hansen, Durup, & Bosso, 2010; Roberts et al., 2009).

Because of the major risk to the public health, it is very important to monitor the structure of microbial communities, which are constantly modified due to a continuous positive selection of resistant subtypes. As suggested by Cornaglia et al. (2004), one of the main reasons for reporting antimicrobial resistance data is to assess the scale of the resistance problem at the local, national, and international levels. In addition, the collected data can be used for monitoring changes in resistance rates and for detecting the emergence and spread of new resistances types.

When it comes down to the collection of data regarding antimicrobial resistance, different methods, and procedures can be followed (Caprioli, Busani, Martel, & Helmuth, 2000). Dilution experiments, discussed in Andrews (2001) and Wiegand, Hilpert, and Hancock (2008), are performed to determine the minimum inhibitory concentration (MIC) values, expressed in milligram per liter (mg/l). The MIC is defined as the lowest concentration of an antimicrobial that still inhibits the visual growth of a microorganism after overnight incubation. Dilution experiments are carried out on plates of solid growth medium, called agar, or by using microtiter plates containing liquid growth medium, referred to as broth. For example, to identify the MIC via a broth dilution method, identical doses of the bacteria are cultured in wells of liquid media containing successively lower concentrations of the antimicrobial. The MIC of the antibiotic is located between the concentrations of the last well in which no bacteria grew and the next lower dose, which allowed bacterial growth. Consider, for instance a bacterial isolate that grows at the concentrations of 1, 2, and 4 mg/l, but shows inhibition of growth at 8 mg/l. In such a situation, the traditionally reported MIC value equals 8 mg/l. Nevertheless, true inhibition of growth occurs at a certain concentration between 4 and 8 mg/l, which makes the obtained MIC value interval-censored. As such, the observed data value is an interval $\lfloor 4,\ 8 \rfloor$. It is one of the aims of this paper to take the interval-censoring, largely ignored in the analyses of the MIC data, properly into account. For the methodology developed in this paper, it does not matter whether the intervals are open, closed or half-closed. This is indicated by using the $\lfloor\ \rfloor$ notation for them.

Furthermore, due to the concept of antimicrobial resistance, the general population of microorganisms is not homogeneous. More specifically, two large subgroups in this population can be identified. On the one hand, we are dealing with wild-type susceptible isolates, that is organisms that have no acquired or mutational resistance. On the other hand, the nonwild-type population contains isolates that have developed any kind of resistance and hence show a reduced susceptibility to certain antimicrobials. This underlying population heterogeneity is also reflected in the MIC value distribution. Indeed, for a given bacterial species, the multimodal pattern of the MIC distribution can usually enable the separation of the wild-type population of microorganisms (lower MIC values) from those nonwild-type populations that show a reduced susceptibility to the antimicrobial in question (higher MIC values). In this respect, let the univariate random variable Y represent the MIC value with probability density function $f(y)$. In our context, a two-component mixture

$$f(y) = w\, f_1(y;\ \boldsymbol{\theta}_1) + (1 - w) f_2(y;\ \boldsymbol{\theta}_2) \tag{1}$$

is naturally assumed, in which f_1 and f_2 respectively represent the wild-type and nonwild-type component of the MIC distribution, possibly depending on vectors of unknown parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively. The prevalence of wild-type isolates is denoted by $w \in (0,\ 1)$, being in general unknown as well.

The wild-type susceptible population commonly shows a unimodal distribution reflecting a slight biological variability around a mode that is not altered by changing circumstances over time. Therefore, f_1 in (1) can be assumed to be of a fixed parametric form, such as a Gaussian density function (Finch, Greenwood, Whitley, & Norrby, 2010). The second component, representing the nonwild-type isolates, is often multimodal, suggesting that it is itself a mixture of different nonwild-type subpopulations that are characterized by different degrees of reduced susceptibility conferred by different mechanisms. Therefore, a simple Gaussian density function does not suffice to capture the variability in the second component and a non-parametric approach is more appealing. Different nonparametric methods exist and have been applied with success before in the field of AMR. For instance, Jaspers, Aerts, Verbeke, and Beloeil (2014b) developed a two-stage semiparametric mixture model. They assume a log-normal or gamma distribution for the first component, f_1 (wild-type subpopulation), and estimate its parameters, $\boldsymbol{\theta}_1$ using the multinomial-based method (Jaspers, Aerts, Verbeke, & Beloeil, 2014a). In a second stage, f_2 (nonwild-type subpopulation) was estimated using a censored-adjusted version of the penalized mixture approach by Schellhase and Kauermann (2012). Several drawbacks related to this two-stage approach were circumvented by the Bayesian composite link model described in Jaspers, Lambert, and Aerts (2016b) and by the back-fitting algorithm presented in Jaspers, Verbeke, Böhning, and Aerts (2016c). The latter approach approximates the unknown density f_2 by a mixture of Gaussian densities, a simple yet flexible approach to describe non-Gaussian distributions (Roeder & Wasserman, 1997; Richardson & Green, 1997).

All methods described above are developed with the aim of estimating a univariate MIC density with special emphasis on estimation of the prevalence of the wild-type and nonwild-type populations and also the MIC distribution in the nonwild-type population (density f_2 in the mixture 1). However, interest in the AMR community is not only in describing this density, but also

to monitor possible evolutions. More specifically, one is often interested in how the prevalence of wild-type isolates changes over time or between different countries. Indeed, in case of an increase in the proportion of nonwild-type isolates, policy makers should be alarmed in order for them to take appropriate actions. Therefore, with the aim of developing an appropriate monitoring tool, the parameter w in the mixture presented in (1), should be made dependent on certain covariates. In this regard, especially the inclusion of a time component is of major interest.

In addition, in recent monitoring programs, information is often gathered on a set of different antimicrobials and interest can go to analyzing so-called coresistance patterns. This requires the multivariate extension of the methods discussed above. In summary, we require two extensions to the existing methodology, that is we need

(E1) a procedure that is able to cope with covariate-dependent mixing weights of the mixture (1) and

(E2) a procedure that is able to estimate the joint (multivariate) MIC density of two or more antimicrobials.

Especially the extension into the multivariate setting is not straightforward for the univariate semiparametric approaches of Jaspers et al. (2014a, 2014b, 2016b, 2016c) mentioned above.

To achieve both (E1) and (E2) in this paper, we will extend a methodology of Komárek (2009), who introduced and implemented methods for Bayesian estimation of multivariate normal mixtures, allowing for the selection of the number of mixture components and also for censored data.

Section 2 elaborates on the general structure of the multivariate Gaussian mixture model with covariate-dependent mixing weights. Bayesian estimation of the model parameters is considered in Section 3, with an application to the data in Section 4. In addition to this real-life data application, the performance of the method is assessed through a small simulation study in Section 5. A short discussion will end the paper in Section 6.

2 | MODEL

Let $\mathbf{Y} = (Y_1, \ldots, Y_q)$ represent the log-transformed ($\log_2$ is traditionally used in this context) outputs from a dilution experiment in which a range of q antimicrobials was tested on a particular isolate. It has been shown on several places (Titterington, Smith, & Makov, 1985, Sec. 2.2; McLachlan & Basford, 1988; McLachlan & Peel, 2000; Fraley & Raftery, 2002) that even in a multivariate setting the Gaussian mixtures provide an appealing semi-parametric tool to model unknown distributions. This motivates us to model the unknown density f of the multivariate MIC outputs $\mathbf{Y}$ by a K component Gaussian mixture as well, namely:

$$f(\mathbf{y}) = \sum_{k=1}^{K} w_k \, \varphi_q(\mathbf{y};\ \boldsymbol{\mu}_k, \Sigma_k), \tag{2}$$

where $\varphi_q(\cdot;\ \mathbf{m},\ \mathrm{S})$ denotes a density of a q-variate Gaussian distribution with mean $\mathbf{m}$ and a covariance matrix S. The unknown parameters of the model are: the mixture weights $w_1, \ldots, w_K$ that are all assumed to be positive and sum-up to one, the mixture means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$, the mixture covariance matrices $\Sigma_1, \ldots, \Sigma_K$ that are all assumed to be positive definite. In general, also the number of mixture components K is unknown.

In a univariate setting, while using the mixtures to model the MIC distribution (e.g. Jaspers et al., 2016c), the wild-type subpopulation of microbials was characterized by the mixture component with the lowest mean value. In a multivariate setting, with K mixture means $\boldsymbol{\mu}_k = (\mu_{k,1}, \ldots, \mu_{k,q})$, $k = 1, \ldots, K$, it is however not always possible to identify a single component $g \in \{1, \ldots, K\}$ for which all marginal means $\mu_{g,j}$, $j = 1, \ldots, q$, attain their minimal value across the subpopulations. Hence, it might not always be possible to easily identify the wild-type population. Nevertheless, the multivariate setting arises in situations where different antimicrobials are tested on a population of microbials and at the same time, different subpopulations may show different levels of resistance to different antimicrobials. Hence it is perfectly possible that none of the subpopulations expresses the lowest expectations of the MIC values toward all antimicrobials tested. Consequently, none of the subpopulations might be termed as fully wild-type. On the other hand, it is not the purpose of the multivariate analysis to identify just the subpopulations of wild-type and nonwild-type isolates. The primary purpose of the multivariate analysis is to analyze the coresistance patterns. This mainly means to find their number in the population, find their proportions and to characterize the level of resistance of each pattern (subpopulation) toward different antimicrobials. All is achieved by the mixture model (2): the number of patterns corresponds to the number of mixture components K, the pattern proportions correspond to the mixture weights $w_1, \ldots, w_K$ and the resistance levels are reflected by the mixture means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$.

Model (2) allows us to tackle our aim (E2) explained in the introduction. To achieve also the aim (E1) we extend the model by allowing for dependence of the mixture weights on covariates. At this point, only saturated dependence will be considered. Suppose that the considered covariates attain jointly C possible values and $x \in \{1, \ldots, C\}$ is the value related to the particular dilution experiment. A density assumed for the MIC output vector $\mathbf{Y}$, given the covariate value x is then

$$f(\mathbf{y};\, x) = \sum_{k=1}^{K} w_k(x)\, \varphi_q(\mathbf{y};\, \boldsymbol{\mu}_k,\, \Sigma_k). \tag{3}$$

In (3) we now have C sets of unknown mixture weights, namely $\mathbf{w}(c) = (w_1(c), \ldots, w_K(c)), c = 1, \ldots, C$. Again, all the weights are positive and are constrained to sum up to one for each c. In the following, all unknown model parameters except the number of mixture components K will be referred to as the parameter vector $\boldsymbol{\theta}$. That is,

$$\boldsymbol{\theta} \equiv (\mathbf{w}(1), \ldots, \mathbf{w}(C),\, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K,\, \Sigma_1, \ldots, \Sigma_K) \equiv (\mathbb{W},\, \boldsymbol{\mu},\, \Sigma), \tag{4}$$

where $\mathbb{W}$ is an $C \times K$ matrix with the weight vectors $\mathbf{w}(1)$, ..., $\mathbf{w}(C)$ in the rows, and $\boldsymbol{\mu} = \{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K\}, \Sigma = \{\Sigma_1, \ldots, \Sigma_K\}$ are aggregate symbols for the mixture means and covariance matrices, respectively.

3 | INFERENCE ON MODEL PARAMETERS

Suppose that $\mathbf{Y}_i = (Y_{i,1}, \ldots, Y_{i,q}), i = 1, \ldots, n$, are the $\log_2$-transformed MIC values corresponding to the outcomes of the dilution experiment using n independent isolates. As explained in the introduction, the recorded MIC values are interval-censored. Hence, we only observe intervals $\lfloor L_{i,j},\, U_{i,j} \rfloor$ such that $Y_{i,j} \in \lfloor L_{i,j},\, U_{i,j} \rfloor, i = 1, \ldots, n, j = 1, \ldots, q$.

Note that if the bacterial isolate shows inhibition of growth already at the lowest tested concentration of antibiotics (let say, $\text{conc}_{\min}$), the true value of the MIC value is left-censored. The corresponding lower limit of the observed MIC interval is equal to zero and on the log-scale, $L = -\infty$, the upper limit of the interval is $U = \log_2(\text{conc}_{\min})$. Oppositely, if the bacterial isolate still grows even at the highest tested concentration of antibiotics (let say, $\text{conc}_{\max}$), the true MIC value is right-censored, the lower limit of the interval is $L = \log_2(\text{conc}_{\max})$ and the upper limit is infinity, that is $U = \infty$.

3.1 | Bayesian model specification

Let $\mathbf{x} = (x_1, \ldots, x_n)$ denote a vector of the covariate values $x_i \in \{1, \ldots, C\}, i = 1, \ldots, n$, related to each of the isolates. Further, let $\mathcal{D}$ represent the observed response data, that is all lower and upper limits of the observed intervals. Due to the setup of the experiment, where the observed limits of the intervals correspond to prespecified values of tested concentrations of antibiotics, the censoring is noninformative and the inference on the model parameters can hence be based on the observed data likelihood. Given the assumed mixture model (3) for the true ($\log_2$-transformed) MIC values, the observed data likelihood takes the form

$$L(\boldsymbol{\theta},\, K;\, \mathcal{D},\, \mathbf{x}) = \prod_{i=1}^{n} \int_{L_{i,1}}^{U_{i,1}} \cdots \int_{L_{i,q}}^{U_{i,q}} \left\{ \sum_{k=1}^{K} w_k(x_i)\, \varphi(\mathbf{y}_i;\, \boldsymbol{\mu}_k,\, \Sigma_k) \right\} \mathrm{d}y_{i,q} \cdots \mathrm{d}y_{i,1}, \tag{5}$$

where $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,q}), i = 1, \ldots, n$.

The likelihood (5) is rather complex for a classical maximum-likelihood based inference, let alone the fact that a dimension of the corresponding parameter space is not fixed in advance due to the unknown number K of the mixture components. For the situation when the mixture weights do not depend on the covariate value, Komárek (2009) describes a Bayesian approach to the problem with the inference based on the Markov chain Monte Carlo (MCMC) methodology and provides its software implementation in a form of the contributed package of the R software (R Core Team, 2016). Here, we briefly summarize the methodology and extend it where necessary to cope with the mixture weights that vary with the value of a categorical covariate x.

To cope with a problem of an unknown dimension of the parameter space caused by an unknown number of mixture components K, we first develop a Bayesian model that assumes a fixed (known) value of K and second propose a method on how to choose an optimal value of K using a suitable information criterion. Hence, first, it is assumed that a value of K is given, the vector of unknown parameters is $\boldsymbol{\theta}$ as indicated by (4) and we describe a Bayesian model to infer on $\boldsymbol{\theta}$.

Firstly, in the mood of Bayesian data augmentation (Tanner & Wong, 1987), we augment the parameter vector in two ways.

(i) In the area of Bayesian analysis of censored data, it is useful to consider the unknown true values, in our case the true $\log_2$-transformed MIC values $Y_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, q$, as additional model parameters and we will do so here as well. In the following, let an $n \times q$ matrix $\mathbb{Y}$ represent those values.

(ii) The mixture in (3) serves, among other things, to capture heterogeneity of the studied population and its division into K subpopulations having the proportions given by the mixture weights. Another set of augmented variables used traditionally in the mixture setting are latent component allocations that reveal from which subpopulation a particular unit is sampled from. We denote them as $R_i \in \{1, \ldots, K\}$, $i = 1, \ldots, n$. In agreement with the mixture model (3), we have a priori $\mathsf{P}(R_i = k;\ x_i) = w_k(x_i)$, $k = 1, \ldots, K$. In the remainder of the paper, let $\mathbf{R} = (R_1, \ldots, R_n)$ be the vector of the component allocations for all observations in the dataset.

Consequently, the parameter vector for the Bayesian model is composed of the primary model parameters $\boldsymbol{\theta} \equiv (\mathbb{W}, \boldsymbol{\mu}, \Sigma)$, the augmented $\log_2$-transformed MIC values $\mathbb{Y}$ and the latent component allocations $\mathbf{R}$. In the following, let $p(\,)$ and $p(\,|\,)$ be generic symbols for a density and a conditional density, respectively. As dictated by the Bayesian paradigm, inference on the model parameters is based on their posterior distribution. Having performed the Bayesian data augmentation, a Bayesian hierarchical model was obtained:

$$\begin{aligned} p(\boldsymbol{\theta}, \mathbb{Y}, \mathbf{R} \mid \mathcal{D}; \mathbf{x}) = p(\boldsymbol{\mu}, \Sigma, \mathbb{W}, \mathbb{Y}, \mathbf{R} \mid \mathcal{D}; \mathbf{x}) \ \propto\ p(\mathbb{Y}, \boldsymbol{\mu}, \Sigma, \mathbb{W}, \mathbf{R}, \mathcal{D};\ \mathbf{x}) = \\ = \ p(\mathcal{D} \mid \mathbb{Y}, \boldsymbol{\mu}, \Sigma, \mathbf{R}, \mathbb{W}; \mathbf{x})\, p(\mathbb{Y} \mid \boldsymbol{\mu}, \Sigma, \mathbf{R}, \mathbb{W}; \mathbf{x}) \\ p(\boldsymbol{\mu}, \Sigma \mid \mathbf{R}, \mathbb{W}; \mathbf{x})\, p(\mathbf{R} \mid \mathbb{W}; \mathbf{x})\, p(\mathbb{W}; \mathbf{x}). \end{aligned} \tag{6}$$

The following holds or is assumed for the factors in (6). First, the term $p(\mathcal{D} \mid \mathbb{Y}, \boldsymbol{\mu}, \Sigma, \mathbf{R}, \mathbb{W}; \mathbf{x})$ corresponds to the lowest hierarchical level of the model that links the observed intervals to the augmented true values $\mathbb{Y}$. Second, the term $p(\mathbb{Y} \mid \boldsymbol{\mu}, \Sigma, \mathbf{R}, \mathbb{W}; \mathbf{x})$ corresponds to the distribution of the true values $\mathbb{Y}$ given the model parameters, that is given the assumed mixture (3), but also given the knowledge of the component allocations $\mathbf{R}$. Hence, we have

$$p(\mathbb{Y} \mid \boldsymbol{\mu}, \Sigma, \mathbf{R}, \mathbb{W}; \mathbf{x}) = p(\mathbb{Y} \mid \boldsymbol{\mu}, \Sigma, \mathbf{R}) = \prod_{i=1}^{n} p(\mathbf{Y}_i \mid \boldsymbol{\mu}, \Sigma,\ R_i) = \prod_{i=1}^{n} \varphi(\mathbf{Y}_i;\ \boldsymbol{\mu}_{R_i}, \Sigma_{R_i}). \tag{7}$$

Third, we assume in a hierarchical structure of the model, that $p(\boldsymbol{\mu}, \Sigma \mid \mathbf{R}, \mathbb{W}; \mathbf{x}) = p(\boldsymbol{\mu}, \Sigma)$, which then corresponds to the (genuine) prior distribution for the primary model parameters $\boldsymbol{\mu}$ and Σ. For the prior distribution $p(\boldsymbol{\mu}, \Sigma)$, Komárek (2009) suggests to consider two options: a semiconjugate independent Gaussian and Inverse Wishart prior or a natural-conjugate Gaussian-Inverse Wishart prior. The former option will be used in the data application here. It corresponds to assuming a priori independence between all mixture means and covariance matrices and assigning Gaussian priors to the means and Inverse Wishart priors to the covariance matrices, that is

$$p(\boldsymbol{\mu}, \Sigma) = \prod_{k=1}^{K} p(\boldsymbol{\mu}_k)\, p(\Sigma_k), \qquad \left.\begin{aligned} p(\boldsymbol{\mu}_k) &= \varphi_q\big(\boldsymbol{\mu}_k;\ \boldsymbol{\xi}_k, \mathbb{D}_k\big), \\ p(\Sigma_k) &\sim \mathcal{IW}_q(\zeta, \Xi), \end{aligned}\right\} k = 1, \ldots, K. \tag{8}$$

In (8), the prior means $\boldsymbol{\xi}_k$ and covariance matrices $\mathbb{D}_k$, $k = 1, \ldots, K$, are fixed hyperparameters. Further, $\mathcal{IW}_p(\zeta, \Xi)$ indicates an inverse-Wishart distribution with ζ degrees of freedom (fixed hyperparameter) and a scale matrix Ξ. This scale matrix is assumed to be diagonal, with let say $\gamma_1, \ldots, \gamma_q$ on a diagonal. The diagonal elements $\gamma_1, \ldots, \gamma_q$ are hyperparameters being assigned inverse-gamma priors in the additional level of hierarchy included in the model in order to allow for a weakly informative specification of the prior distribution. Fourth, the term $p\big(\mathbf{R} \mid \mathbb{W};\ \mathbf{x}\big)$ in (6) follows from the Bayesian data augmentation procedure related to the component allocations, that is

$$p(\mathbf{R} \mid \mathbb{W};\ \mathbf{x}) = \prod_{i=1}^{n} p(R_i \mid \mathbb{W};\ x_i) = \prod_{i=1}^{n} w_{R_i}(x_i). \tag{9}$$

Finally, the factor $p(\mathbb{W};\ \mathbf{x})$ corresponds to the prior distribution of the mixture weights. Here, we assume a priori independence between the sets of the mixture weights for different covariate values and assume, for each set of weights a Dirichlet prior with a common parameter δ, that is

$$p(\mathbb{W}) = \prod_{c=1}^{C} p\big(\mathbf{w}(c)\big) \propto \prod_{c=1}^{C} \prod_{k=1}^{K} \big\{w_k(c)\big\}^{\delta-1}. \tag{10}$$

To use the hierarchical model (6) in practice, one has to choose a set of prior hyperparameters. Those include: the means $\boldsymbol{\xi}_k$ and the covariance matrices $\mathbb{D}_k$, $k = 1, \ldots, K$ from the prior factor (8), the degrees of freedom ζ from the same factor, parameters of the inverse-gamma hyperpriors being used for diagonal elements of the scale matrix Ξ of the inverse-Wishart piece of (8), the Dirichlet parameter δ from (10). Guidelines are given in Komárek (2009) on how to choose those parameters in order to obtain a weakly informative prior distribution and we refer the reader therein.

3.2 | Markov chain Monte Carlo

Markov Chain Monte Carlo methods can be used to sample from the posterior distribution (6). In analogy with Komárek (2009), a block Gibbs sampler is used while updating the subsets of the parameters by sampling from their full conditional distribution. For the augmented response values, the mixture means and covariance matrices, these full conditionals do not depend on the mixture weights and sampling proceeds as described in Komárek (2009). On the other hand, an updated procedure is required to sample the latent component allocations $\mathbf{R}$ and the mixture weights. More specifically,

$$p\big(\mathbf{R}\,\big|\,\cdots\big) \propto p\big(\mathbb{Y}\,\big|\,\boldsymbol{\mu},\ \Sigma,\ \mathbf{R}\big)\, p\big(\mathbf{R}\,\big|\,\mathbb{W};\ \mathbf{x}\big) = \prod_{i=1}^{n} \varphi_q\big(\mathbf{y}_i;\ \boldsymbol{\mu}_{R_i},\ \Sigma_{R_i}\big)\, w_{R_i}(x_i).$$

That is, for each $i = 1, \ldots, n$, the new allocation R_i is sampled from the categorical counterpart of a multinomial distribution having the covariate dependent component probabilities proportional to the product $\varphi_q\big(\mathbf{y}_i;\ \boldsymbol{\mu}_k,\ \Sigma_k\big)\, w_k(x_i)$, $k = 1, \ldots, K$. In addition,

$$p\big(\mathbb{W}\,\big|\,\cdots\big) \propto p\big(\mathbf{R}\,\big|\,\mathbb{W};\ \mathbf{x}\big)\, p(\mathbb{W}) = \prod_{c=1}^{C} \left\{ \prod_{i:\ x_i=c} p\big(R_i\,\big|\,\mathbf{w}(c);\ x_i = c\big) \right\} p\big(\mathbf{w}(c)\big) \propto \prod_{c=1}^{C} \left\{ \prod_{k=1}^{K} \big\{w_k(c)\big\}^{n_{c,k}+\delta-1} \right\},$$

where $n_{c,k} = \sum_{i:\ x_i=c} \mathbb{I}(R_i = k)$ denotes the number of observations with the covariate value c ($c = 1, \ldots, C$) being currently allocated in the k-th mixture component ($k = 1, \ldots, K$). This shows that the Gibbs step for the covariate dependent mixture weights is performed by sampling independently each set of the C weights, the c-th one ($c = 1, \ldots, C$) from the Dirichlet distribution with parameters $n_{c,1} + \delta, \ldots, n_{c,K} + \delta$.

Similar to the original proposal of Komárek (2009), it is recommended to work with scaled and possibly shifted data in order to have a possibility to improve numerical properties of the MCMC procedure. It is especially desirable to rescale the outcome vector $\mathbf{Y}$ such that all its elements show variability of the same magnitude.

Finally, the selection of the number of the mixture components K in the model (3) can be performed using the penalized expected deviance (PED), where smaller values indicate a better fit (Plummer, 2008; Komárek, 2009).

4 | DATA APPLICATION

In the light of major monitoring and surveillance programs, Member States of the European Union are required to collect and evaluate data on antimicrobial resistance and food-borne outbreaks. Since 2010, data are collected on an isolate-based level (Jaspers, Ganyani, Ensoy, Faes, & Aerts, 2016a). A single bacterial isolate is tested for susceptibility against multiple antimicrobials, thereby providing the opportunity for a joint analysis. One of the bacteria of primary interest is *Escherichia coli*. Although most of these bacteria are harmless and actually are an important part of a healthy human intestinal tract, *E. coli* is one of the most frequent causes of bloodstream infections and community—and healthcare-associated urinary tract infections worldwide. A retrospective analysis of *E. coli* from urine specimens collected from patients during 1997–2007 in Switzerland showed an increasing resistance trend for ciprofloxacin, trimethoprim/sulfamethoxazole, and amoxicillin/clavulanic acid (Blaettler et al., 2009). Similarly a 30-year (1979–2009) follow-up study on *E. coli* in Sweden showed an increasing resistance trend

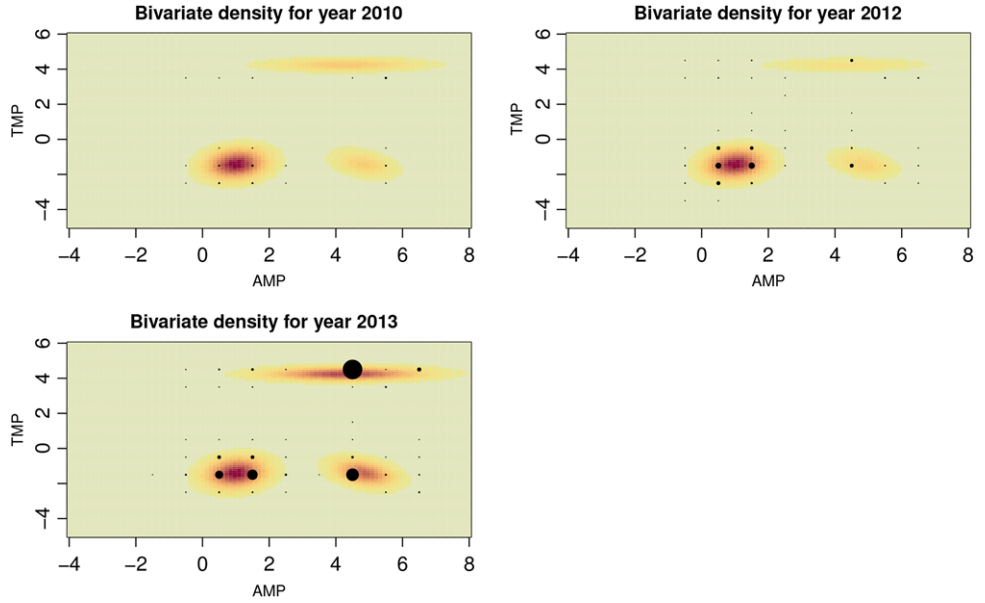

FIGURE 1 Scatterplot of the data, with the dots located at the midpoints of the censoring intervals and the size of the dots proportional to the number of observations. Overlaid are the predictive densities for the final fit with three components

for ampicillin, sulfonamide, trimethoprim, and gentamicin (Kronvall, 2010). In the following application, we focus on the joint susceptibility of *E. coli* isolates against two-third-line agents, namely ampicillin (AMP) and trimethoprim (TMP). Over the years 2010, 2012, and 2013, $n = 2\,539$ isolates were tested for susceptibility in nine European Union Member States.

Hence, in our general notation, $q = 2$ and the two components of the response vector $\mathbf{Y} = (Y_1, Y_2)$ correspond to the $\log_2$-transformed MIC values evaluating susceptibility of *E. coli* against AMP and TMP. The observed intervals $\lfloor L_{i,j}, U_{i,j} \rfloor$, $i = 1, \ldots, n$, $j = 1, 2$, arise from performing the dilution experiments with concentrations of respective antimicrobials being 0.5, 1,..., 128 mg/l. The aim of the analysis is to evaluate the coresistance patterns of *E. coli* against those two agents and moreover, to monitor changes of those patterns over the years 2010, 2012, and 2013. Hence, the covariate value x takes $C = 3$ values and $x = 1, 2, 3$ distinguishes between the experiments performed in the three respective years.

We fitted a two to four component bivariate mixture to the data. The reported results are based on 8000 iterations of 1:10 thinned MCMC obtained after a burn-in period of 1000 iterations. Convergence of the MCMC procedure was evaluated by standard tools while using the R coda package (Plummer, Best, Cowles, & Vines, 2006). Comparing the penalized expected deviance for the three models (with $K = 2, 3, 4$), a three-component mixture was found to be the most optimal. The predictive joint densities resulting from the optimal model fit can be found in Figure 1. The observed data are overlaid, with the size of the dots proportional to the number of isolates.

A first component, located at mean $\widehat{\boldsymbol{\mu}}_1 = (1.01, -1.41)$ corresponds to isolates that are susceptible to both antimicrobials of interest and could be termed as fully wild type. The second component, located at $\widehat{\boldsymbol{\mu}}_2 = (4.85, -1.41)$ has a higher mean MIC value for AMP compared to component 1. It can therefore be considered as the subpopulation of isolates that show a reduced susceptibility against AMP, but that are still susceptible against TMP. Finally, the third component shows elevated MIC values for both AMs under consideration and is located at $\widehat{\boldsymbol{\mu}}_3 = (4.29, 4.25)$.

The predictive marginal density functions can be found in Figure 2, where they are overlaid on histograms of the MIC frequencies per year. When interpreting the histograms, it must however be taken into account that they are included purely for better visualization. Note that the histograms do not provide correct density estimators due to censoring and the fact that midpoints of observed intervals and the right-censoring values, respectively, were used to produce them. From the plots, it is observed that the assumption of a fixed first component over the different years seems justified. Also the nonwild-type component is captured relatively well by the marginal densities. These separate views on the wild-type and nonwild-type components provide a nice addition to the global view that results from creating the histograms. In addition, formal inference allows for assessing whether there truly is an increase in the prevalence of one of the specific subgroups over time. Some of the discrepancies that can be observed are mainly due to different dilution ranges for different Member States. This issue will be resolved in the future as formal rules for the collection of AMR data, including the selection of ranges, have been postulated on the European Union level (2013/652/EU, 2013). Currently, a drawback of this is that the visual representation of the results can be misleading. For

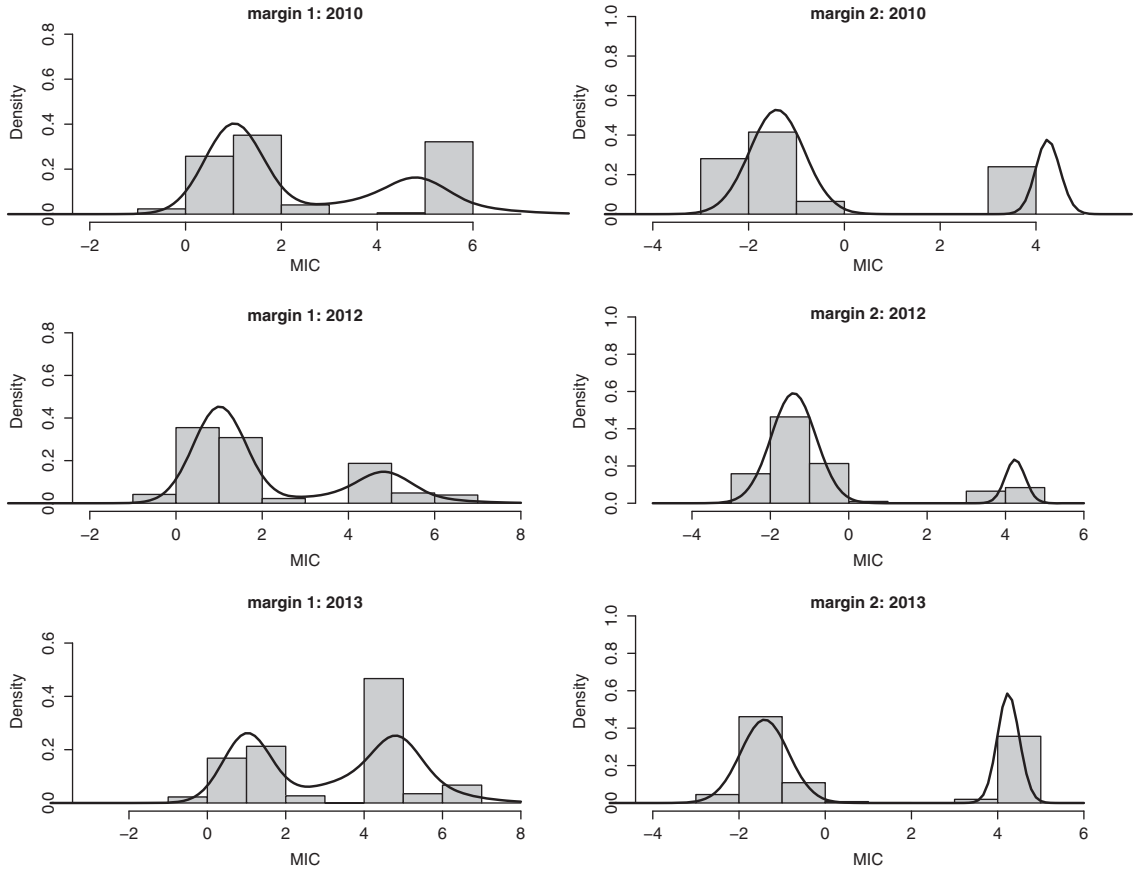

FIGURE 2 Predicted marginal densities for ampicillin (margin 1) and trimethoprim (margin 2), together with the histograms of the observed data per year

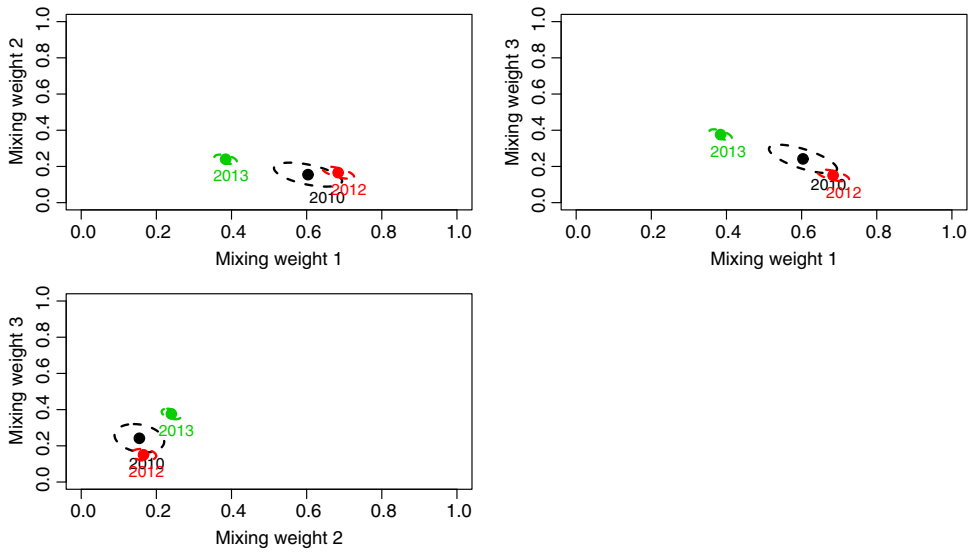

FIGURE 3 Credible ellipses for the mixing weights based on the final fitted model

example, the apparent misplacement of the estimated density of margin 2 in the year 2010 is merely the effect of having a large amount of right-censored observations (i.e. MIC≥4), for which the true values are located to the right of the final bar in the histogram. Nevertheless, their exact locations are unknown.

On Figure 3, we have plotted the posterior means of the mixing weights accompanied by their 95% credible ellipses. While the first component receives a relatively large weight in 2010 (0.6036) and 2012 (0.6838), there is a huge drop in 2013 (0.3843).

Moreover, in 2013, there seems to be a shift towards the third component, which now receives a weight of 0.3759 compared to 0.2416 in 2010 and 0.1504 in 2012. It is also observed that the constructed credible ellipses for 2010 and 2012 are overlapping, while the one for 2013 is clearly separated from the others. We could therefore conclude that for 2013, the proportion of bacteria that have a reduced susceptibility against both ampicillin and trimethoprim is significantly higher compared to 2010 and 2012.

5 | SIMULATION STUDY

In order to assess the performance of the introduced method, we performed a small simulation study. The basic idea in this study is to analyze an extreme example of evolving resistance strains. More specifically, we considered four years in which we followed the susceptibility of a microorganism tested against two antimicrobials, denoted by AM1 and AM2. In the first year, we assume that the isolates show no resistance against any of the antimicrobials tested. In the second year, half of the isolates have developed resistance traits against AM2, while still no resistance is observed against AM1. Hence, in this second year, 50% of the isolates are contained in the wild-type sub-population. This proportion still holds for the third year, in which half of the isolates that were already resistant to AM2 have also developed resistance traits against AM1. Finally, in the last year under consideration, the remaining 25% have also developed resistance traits against AM1. In summary, data are sampled from a three-component mixture of bivariate normal densities, with $x \in \{1, 2, 3, 4\}$ corresponding to the year and represented by

$$f(\mathbf{y};\ x) = w_1(x)\mathcal{N}_2\left(\begin{bmatrix}0.5\\-2\end{bmatrix}, \begin{bmatrix}0.5 & 0.2\\0.2 & 0.5\end{bmatrix}\right) + w_2(x)\mathcal{N}_2\left(\begin{bmatrix}0.5\\1\end{bmatrix}, \begin{bmatrix}0.5 & 0.2\\0.2 & 0.5\end{bmatrix}\right) + w_3(x)\mathcal{N}_2\left(\begin{bmatrix}3.2\\1\end{bmatrix}, \begin{bmatrix}0.8 & 0.5\\0.5 & 0.8\end{bmatrix}\right), \tag{11}$$

where

$$\boldsymbol{w}(x) = \begin{bmatrix} w_1(1) & w_2(1) & w_3(1)\\ w_1(2) & w_2(2) & w_3(2)\\ w_1(3) & w_2(3) & w_3(3)\\ w_1(4) & w_2(4) & w_3(4)\end{bmatrix} = \begin{bmatrix} 0.98 & 0.01 & 0.01\\ 0.5 & 0.49 & 0.01\\ 0.5 & 0.25 & 0.25\\ 0.5 & 0.01 & 0.49\end{bmatrix}.$$

We considered 1000 samples with a sample size of 2000, thereby mimicking the size of most real-life data applications. From these 2000 data points, 500 were sampled in year 1, 520 in year 2, 480 in year 3 and again 500 in year 4. For each simulation run, the simulated data were censored in order to resemble real-life datasets as closely as possible. For both antimicrobials, right-censoring was established at two dilution steps below the maximum, while the remaining observations were rounded to the next integer, thereby becoming interval-censored.

When estimating a parameter of interest, denoted by $\boldsymbol{\theta}$, we can have a look at the mean squared error (MSE) to assess the performance of the applied method. Ideally, we would compare different methods based on this summary measure. However, in absence of existing alternatives, we will only regard the MSE values for the current method. Attention is paid to the estimation of the mixing weights, the component means and their variance-covariance matrices. Their posterior means based on appropriately relabeled MCMC sample are considered as their estimators. A summary of the calculated MSE values is presented in Table 1. It can be seen there that all measures are extremely small, indicating a promising behavior of the developed method.

In addition to analyzing the joint distribution of the antimicrobials, one can also be interested in the marginal density functions. Therefore, Figure 4 shows the estimated marginal densities for the individual runs, together with the true marginals. Again here, the estimated densities provide an excellent approximation to the truth. It is also observed that a change in the distribution of the nonwild-type component can be captured by setting one or more of the mixing weights to zero. Therefore, the developed method is also appropriate when the nonwild-type distribution is not constant over time (as might be the case in real-life applications).

The results in this simulation study derived from a relatively balanced data setting (i.e. almost equal amount of data points in each year). Nevertheless, also imbalanced data settings lead to similar results. Indeed, another simulation study was performed on a smaller scale. More specifically, 100 data sets were simulated from mixture (11), where 200 data points were sampled in year 1, 1250 in year 2 and 275 in both year 3 and year 4. The results from this simulation study are provided as supplementary information on the journal web page (link). It can be observed there that, even with this smaller simulation size, the values of the performance measures are very similar to those observed in Table 1. The reason for this is that only the estimates for the mixing weights depend on the specific years, while the estimates for the mean and variance-covariance matrix of the different

TABLE 1 Bias, variance, and mean squared error (MSE) values regarding the estimated mixture weights, the component means and their variance/covariance matrices

	Mixing weights ($\times 10^{-5}$)				Means ($\times 10^{-3}$)				Variance-covariance ($\times 10^{-3}$)			
Component		Bias	Var	MSE		Bias	Var	MSE		Bias	Var	MSE
1	$w_1(1)$	−381.05	0.85	2.30	μ_{11}	1.24	0.48	0.48	$\Sigma_{1,11}$	3.24	0.63	0.64
	$w_1(2)$	202.42	5.10	5.51	μ_{12}	6.12	0.58	0.61	$\Sigma_{1,12}$	2.54	0.39	0.39
	$w_1(3)$	224.98	6.04	6.55					$\Sigma_{1,22}$	8.74	0.75	0.83
	$w_1(4)$	122.68	6.87	7.02								
2	$w_2(1)$	188.78	0.55	0.90	μ_{21}	1.56	1.96	1.96	$\Sigma_{2,11}$	5.02	2.37	2.39
	$w_2(2)$	−399.20	5.42	7.01	μ_{22}	8.33	1.94	2.01	$\Sigma_{2,12}$	−1.24	1.21	1.21
	$w_2(3)$	−126.14	6.80	6.96					$\Sigma_{2,22}$	−9.67	2.64	2.73
	$w_2(4)$	242.52	1.00	1.59								
3	$w_3(1)$	192.28	0.51	0.88	μ_{31}	5.07	3.37	3.39	$\Sigma_{3,11}$	−5.32	8.00	8.03
	$w_3(2)$	196.78	0.89	1.28	μ_{32}	4.64	3.21	3.23	$\Sigma_{3,12}$	−7.40	5.07	5.12
	$w_3(3)$	−98.83	8.01	8.10					$\Sigma_{3,22}$	1.64	7.78	7.79
	$w_3(4)$	−365.20	7.74	9.07								

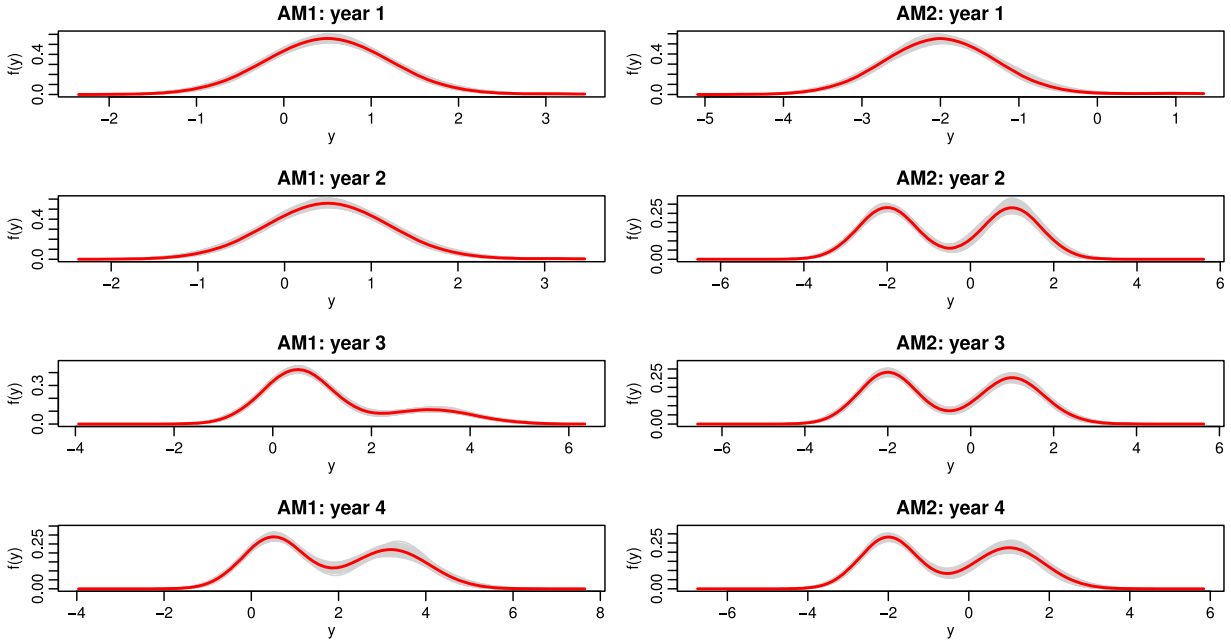

FIGURE 4 Marginal density estimates obtained after simulating 1000 datasets from mixture (11). The individual results from each run are shown in gray, while the dashed curve indicates the true underlying marginal density

components are based on the pooled data. Therefore, the only requirement is that a sufficient amount of information is present in each of the years, but an imbalance in this available information is not an issue as such. In the supplementary information, the marginal density estimates are plotted as well, given also an image that is nearly identical to the results in Figure 4. In conclusion, the proposed method is quite robust to imbalanced data scenarios. Finally, it is worthwhile to note that an increase in the amount of overlap between wild-type and nonwild-type subgroups might have a negative effect on the quality of the estimate. Indeed, in case of highly overlapping subpopulations, problems related to identifiability might arise, making it harder to correctly estimate the component-specific parameters. These issues might for example arise in case a subgroup of the susceptible population is starting to develop resistance to specific antimicrobials. In this initial stage of development, the MIC values might still be too close together and the model will not be able to identify the additional component. This issue of overlapping subpopulations rather a general problem to the estimation of mixture models and it is hence not specifically related to the methodology in this paper. Similar remarks were made for the univariate setting (Jaspers et al., 2014a, 2014b).

6 | DISCUSSION

Antimicrobial resistance is one of the major public health burdens of the last decades and it is therefore important to monitor the distribution of isolates that show a reduced susceptibility to one or more antimicrobials of interest. In this paper, we presented an extension to the Bayesian estimation procedure for multivariate normal mixtures, introduced by Komárek (2009). The extension concerned the inclusion of covariate-varying mixing weights. On the other hand, the parameters of the component densities were assumed to be independent of any covariates. The developed method was applied to a real-life data application, in which we investigated the joint susceptibility of *Escherichia coli* isolates against the antimicrobials ampicillin and trimethoprim. Based on observed data from three years, a possible decrease in susceptible isolates was detected, while the number of isolates nonsusceptible to both antimicrobials under investigation had increased in the final year. These observations might trigger policy makers to further identify possible causes of the observed shift over time and to undertake plausible actions regarding this phenomenon. In addition, a simulation study was performed to assess the performance of the proposed method, both in terms of estimating the parameters of the component densities as well as estimating the marginal density functions. In both cases, promising results were found and we therefore believe that the introduced procedure can be a valuable tool for monitoring antimicrobial resistance.

The choice for a saturated model was based on the analysis of real-life data applications. Indeed, when interest is in comparing between few different years or different countries, the choice follows naturally. On the other hand, when an evolution over time is of prior interest, one might feel more for a (non-)parametric effect of time. Nevertheless, at this point AMR data are only collected on an isolate-based level for a small amount of years and fitting any trend would suffer from the lack of data to justify it. Of course, updating the model to include such a time trend is possible and is part of our further research. In addition, the parameters of the component densities are considered to be independent of any covariate. For the wild-type first component, this assumption is very tenable. Indeed, it is well-documented that the wild-type component commonly shows a unimodal distribution reflecting a slight biological variability around a mode that is not altered by changing circumstances over time (Finch et al., 2010; Lee & Whitmore, 1999; Turnidge, Kahlmeter, & Kronvall, 2006). For the nonwild-type part of the MIC distribution, less information is present and there could well changes in their distribution over time. The current method deals with these shifts by not fixing the number of components beforehand, but allowing for a selection of the number of component densities based on the penalized expected deviance. When a certain component is not present for a certain moment in time or for a certain country, the assigned weight will reduce to zero. This could also be observed in the simulation study, where it was shown that weights that were almost zero were also estimated with great precision. As an alternative, one could also consider to allow the means and variance-covariance matrices of the distinct components to depend on certain covariates. Although possible in theory, the implementation of such an extension is less trivial and might suffer from identifiability issues. In summary, we are confident that the proposed method provides a very nice monitoring tool, for which a further extension regarding the identification of (non-)parametric trends are a nice topic of future work.

The original mixAK package was extended to also incorporate the methodology introduced in this paper. A sample script to perform the analysis shown in this paper is available as supplementary material at [appropriate link will be added at final stage of review].

ACKNOWLEDGMENT

The research of the first author was supported by the Research Foundation Flanders (FWO), grant 11E2913N. Support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) is gratefully acknowledged by all authors. The authors also thank EFSA for providing the data, which can also be retrieved from the EU summary reports published at the EFSA website. Finally, the first author is grateful for the hospitality during his visit to Charles University in Prague and the second author for the hospitality during his visit to Hasselt University.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

Stijn Jaspers http://orcid.org/0000-0003-0201-2067

REFERENCES

2013/652/EU (2013). Commission Implementing Decision of 12 November 2013 on the monitoring and reporting of antimicrobial resistance in zoonotic and commensal bacteria (notified under document C(2013) 7145). Text with EEA relevance. Technical Report, The European Commission.

Andrews, J. (2001). Determination of minimum inhibition concentrations. *Journal of Antimicrobial Chemotherapy*, *48*, S1.5–S1.16.

Bennett, J. W., & Chung, K. T. (2001). Alexander Fleming and the discovery of penicillin. *Advances in Applied Microbiology*, *49*, 163–184.

Blaettler, L., Mertz, D., Frei, R., Elzi, L., Widmer, A., Battegay, M., … Flückiger, U. (2009). Secular trend and risk factors for antimicrobial resistance in *Escherichia coli* isolates in Switzerland 1997–2007. *Infection*, *37*, 534–539.

Caprioli, A., Busani, L., Martel, J. L., & Helmuth, R. (2000). Monitoring of antibiotic resistance in bacteria of animal origin: Epidemiological and microbiological methodologies. *International Journal of Antimicrobial Agents*, *14*, 295–301.

Cohen, M. L. (2000). Changing patterns of infectious disease. *Nature*, *406*, 762–767.

Cornaglia, G., Hryniewicz, W., Jarlier, V., Kahlmeter, G., Mittermayer, H., Stratchounski, L., … ESCMID Study Group for Antimicrobial Resistance Surveillance,. (2004). European recommendations for antimicrobial resistance surveillance. *Clinical Microbiology and Infection*, *10*, 349–383.

Finch, R. G., Greenwood, D., Whitley, R. J., & Norrby, S. R. (2010). *Antibiotic and Chemotherapy*. Saunders, CA: Elsevier.

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, *97*, 611–631.

Jaspers, S., Aerts, M., Verbeke, G., & Beloeil, P. (2014a). Estimation of the wild-type minimum inhibitory concentration value distribution. *Statistics in Medicine*, *33*, 289–303.

Jaspers, S., Aerts, M., Verbeke, G., & Beloeil, P. (2014b). A new semi-parametric mixture model for interval censored data, with applications in the field of antimicrobial resistance. *Computational Statistics and Data Analysis*, *71*, 30–42.

Jaspers, S., Ganyani, T., Ensoy, C., Faes, C., & Aerts, M. (2016a). Development and application of statistical methodology for analysis of the phenomenon of multi-drug resistance in the EU: Demonstration of analytical approaches using antimicrobial resistance isolate-based data. *EFSA Supporting Publication*, *13*, 1–54.

Jaspers, S., Lambert, P., & Aerts, M. (2016b). A Bayesian approach to the semi-parametric estimation of a MIC distribution. *Annals of Applied Statistics*, *10*, 906–924.

Jaspers, S., Verbeke, G., Böhning, D., & Aerts, M. (2016c). Application of the Vertex Exchange Method to estimate a semi-parametric mixture model for the MIC density of *Escherichia coli* isolates tested for susceptibility against ampicillin. *Biostatistics*, *17*, 94–107.

Komárek, A. (2009). A new R package for Bayesian estimation of multivariate normal mixures allowing for selection of the number of components and interval-censored data. *Computational Statistics and Data Analysis*, *53*, 3932–3947.

Kronvall, G. (2010). Antimicrobial resistance 1979–2009 at Karolinska Hospital, Sweden: Normalized resistance interpretation during a 30-year follow-up on *Staphylococcus aureus* and *Escherichia coli* resistance development. *APMIS*, *118*, 621–639.

Lee, M.-L. T., & Whitmore, G. A. (1999). Statistical inference for serial dilution assay data. *Biometrics*, *55*, 1215–1220.

Ligon, B. L. (2004). Sir Alexander Fleming: Scottish researcher who discovered penicillin. *Seminars in Pediatric Infectious Diseases*, *15*(1), 58–64.

Mauldin, P. D., Salgado, C. D., Hansen, I. S., Durup, D. T., & Bosso, J. A. (2010). Attributable hospital cost and length of stay associated with health care-associated infections caused by antibiotic-resistant gram-negative bacteria. *Antimicrobial Agents Chemotherapy*, *54*, 109–115.

McLachlan, G. J., & Basford, K. E. (1988). *Mixture Models: Inference and Applications to Clustering*. New York, NY: Marcel Dekker, Inc.

McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. New York, NY: Wiley & Sons.

Plummer, M. (2008). Penalized loss functions for Bayesian model comparison. *Biostatistics*, *9*, 523–539.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*, 7–11. http://CRAN.R-project.org/doc/Rnews/.

R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, AT. https://www.R-project.org/.

Richardson, S., & Green, P. (1997). On Bayesian analysis of mixtures with unknown number of components (with discussion). *Journal of the Royal Statistical Society, Series B*, *59*, 731–792.

Roberts, R. R., Hota, B., Ahmad, I., Scott II, R. D., Foster, S. D., Abbasi, F., … Weinstein, R. A. (2009). Hospital and societal costs of antimicrobial-resistant infections in a Chicago teaching hospital: Implications for antibiotic stewardship. *Clinical Infectious Diseases*, *49*, 1175–1184.

Roeder, K., & Wasserman, L. (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association*, *92*, 894–902.

Schellhase, C., & Kauermann, G. (2012). Density estimation and comparison with a penalized mixture approach. *Computational Statistics*, *27*, 757–777.

Spellberg, B., Guidos, R., Gilbert, D., Bradley, J., Boucher, H. W., Scheld, W. M., … Infectious Diseases Society of America. (2008). The epidemic of antibiotic-resistant infections: A call to action for the medical community from the Infectious Diseases Society of America. *Clinical Infectious Diseases*, *46*, 155–164.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation (with Discussion). *Journal of the American Statistical Association*, *82*, 528–550.

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. New York, NY: Wiley & Sons.

Turnidge, J., Kahlmeter, G., & Kronvall, G. (2006). Statistical characterisation of bacterial wild-type MIC value distributions and the determination of epidemiological cut-off values. *Clinical Microbiology and Infection*, *12*, 418–425.

Wiegand, I., Hilpert, K., & Hancock, R. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature Protocols*, *3*, 163–175.

SUPPORTING INFORMATION

Additional Supporting Information including source code to reproduce the results may be found online in the supporting information tab for this article.

How to cite this article: Jaspers S, Komárek A, Aerts M. Bayesian estimation of multivariate normal mixtures with covariate-dependent mixing weights, with an application in antimicrobial resistance monitoring. *Biometrical Journal*. 2018;60:7–19. https://doi.org/10.1002/bimj.201600253