A Geostatistical Framework for Combining Spatially Referenced Disease Prevalence Data from Multiple Diagnostics

Benjamin Amoah, Emanuele Giorgi and Peter J. Diggle
CHICAS, Lancaster University Medical School, Lancaster LA1 4YG, UK

Abstract

Multiple diagnostic tests are often used due to limited resources or because they provide complementary information on the epidemiology of a disease under investigation. Existing statistical methods to combine prevalence data from multiple diagnostics ignore the potential overdispersion induced by the spatial correlations in the data. To address this issue, we develop a geostatistical framework that allows for joint modelling of data from multiple diagnostics by considering two main classes of inferential problems: (1) to predict prevalence for a gold-standard diagnostic using low-cost and potentially biased alternative tests; (2) to carry out joint prediction of prevalence from multiple tests. We apply the proposed framework to two case studies: mapping *Loa loa* prevalence in Central and West Africa, using miscroscopy and a questionnaire-based test called RAPLOA; mapping *Plasmodium falciparum* malaria prevalence in the highlands of Western Kenya using polymerase chain reaction and a rapid diagnostic test. We also develop a Monte Carlo procedure based on the variogram in order to identify parsimonious geostatistical models that are compatible with the data. Our study highlights (*i*) the importance of accounting for diagnostic-specific residual spatial variation and (*ii*) the benefits accrued from joint geostatistical modelling so as to deliver more reliable and precise inferences on disease prevalence.

Keywords: disease mapping; geostatistics; malaria; neglected tropical disesaes; multiple diagnostic tests; prevalence.

1 Introduction

Disease mapping denotes a class of problems in public health where the scientific goal is to predict the spatial variation in disease risk on a scale that can range from sub-national to global (Murray et al., 2014; Liu et al., 2012). Understanding the geographical distribution of a disease is particularly

important in the decision-making process for the planning, implementation, monitoring and evaluation of control programmes (World Health Organization, 2017; Bhatt et al., 2015). In this context, model-based geostatistical methods (Diggle et al., 1998) have been especially useful in low-resource settings (Diggle and Giorgi, 2016; Gething et al., 2012; Zouré et al., 2014) where disease registries are non-existent or geographically incomplete, and monitoring of the disease burden is carried out through cross-sectional surveys and passive surveillance systems.

It is often the case that prevalence data from a geographical region of interest are obtained using different diagnostic tests for the same disease under investigation. The reasons for this are manifold. For example, when the goal of geostatistical analysis is to map disease risk on a continental or global scale by combining data from multiple surveys, dealing with the use of different diagnostic tests may be unavoidable. In other cases, gold-standard diagnostic tests are often expensive and require advanced laboratory expertise and technology which may not always be available in constrained resource settings. This requires the use of more cost-effective alternatives for disease testing in order to attain a required sample size. Different diagnostics might also provide complementary information of intrinsic scientific interests into the spatial variation of disease risk and the distribution of hotspots.

In the absence of statistical methods that allow for the joint analysis of multiple diagnostics, most studies have reported separate analyses. A shortcoming of this approach is that it ignores, and therefore fails to explain, possible correlations between prevalence of different diagnostics. Statistical inference might benefit from a joint analysis, which can yield more efficient estimation of regression parameters (Song et al., 2009) and more precise predictions of prevalence.

However, different diagnostic tests can exhibit considerable disparities in the estimates of disease prevalence for the same population, or even the same individuals. Obvious sources of such variation include differences in sensitivity and specificity. Furthermore, different diagnostics may exhibit differences in their association with the same risk factors. In a geostatistical context, there may also be differences between the spatial covariance structures of different diagnostics.

These aspects highlight the potential challenges that joint modelling of multiple diagnostics needs to take into account. In this paper, we address such issues in order to develop a general framework for geostatistical analysis and describe the application of this framework to *Loa loa* and malaria mapping in Africa.

The structure of the paper is as follows. In Section 2, we describe the two motivating applications. In Section 3, we review existing methods for combining prevalence data from different diagnostics. In Section 4, we introduce a geostatistical framework for combining data from two diagnostics and distinguish two main classes of problems that arise in this context. In Sections 5

and 6, we apply this framework to the two case studies introduced in Section 2. In Section 7, we discuss methodological extensions to more than two diagnostics.

2 Motivating applications

2.1 Loa loa mapping in Central and West Africa

Loiasis is a neglected tropical disease that has received an increased attention due to its impact on the control of a more serious infectious disease, onchocerciasis, that is endemic in large swathes of sub-Saharan Africa. Mass administration of the drug ivermectin confers protection against onchocerciasis, but individuals who are highly co-infected with *Loa loa* - the Loiasis parasite - can develop severe and occasionally fatal adverse reaction to the drug (Boussinesq et al., 1998).

Boussinesq et al. (2001) have shown that high levels in *Loa loa* prevalence within a community are strongly associated with a high parasite density. For this reason, Zouré et al. (2011) have suggested that precautionary measures should be put in place before the roll-out of mass drug administration with ivermectin in areas where prevalence of infection with *Loa loa* is greater than 20%.

In order to carry out a rapid assessment of the *Loa loa* burden in endemic areas a questionnaire instrument, named RAPLOA, was developed as a more economically feasible alternative to the standard microscopy-based microfilariae (MF) detection in blood smears (Takougang et al., 2002). To validate the RAPLOA methodology against microscopy, cross-sectional surveys using both diagnostics were carried out in four study sites in Cameroon, Nigeria, Republic of Congo and the Democratic Republic of Congo (see Wanji et al. (2012) and Additional Figure 1 in Web Appendix B).

In this study, the objective of statistical inference is to develop a calibration relationship between the two diagnostic procedures. This could then be applied to map microscopy-based MF prevalence in areas where the more economical RAPLOA questionnaire is the only feasible option.

2.2 Malaria mapping in the highlands of Western Kenya

Malaria continues to be a global public health challenge, especially in sub-Sharan Africa which, in 2016, accounted for about 90% of all the 445,000 estimated malaria deaths worldwide (World Health Organization, 2017). Polymerase chain reaction (PCR) and a rapid diagnostic test (RDT) are two of the most commonly used procedures for detecting *Plasmodium falciparum*, the deadliest species of the malaria parasites. PCR is highly sensitive and specific, but its use is constrained by

high costs and the need for highly trained technicians. RDT is simpler to use, cost-effective and requires minimal training, but is less sensitive than PCR (Tangpukdee et al., 2009). Recent studies have reported that PCR and RDT can lead to the identification of different malaria hotspots, i.e. areas where disease risk is estimated to be unexpectedly high (Mogeni et al., 2017). In this context, mapping of both diagnostics is of epidemiological interest since their effective use is dependent on the level of malaria transmission, with PCR being the preferred testing option in low-transmission settings (Mogeni et al., 2017).

In order to investigate this issue, a malariometric survey was conducted using both RDT and PCR in two highland districts of Western Kenya (see Additional Figure 2 in Web Appendix B); see Stevenson et al. (2015) for a descriptive analysis of this study. In this scenario, a joint model for the reported malaria counts from the two diagnostics could allow to exploit their cross-correlation and identify malaria hotspots more accurately.

3   Literature review

We formally express the format of geostatistical data from multiple diagnostics as

$$\mathcal{D} = \{(x_{ik}, n_{ik}, y_{ijk}) : j = 1, \ldots, n_{ik}; i = 1, \ldots, N; k = 1, \ldots, K\} \tag{3.1}$$

where y_{ijk} is a binary outcome taking value 1 if the j-th individual at location x_{ik} tests positive for the disease under investigation using the k-th diagnostic procedure, and 0 otherwise. We use p_{ijk} to denote the probability that an individual has a positive test outcome from the k-th diagnostic. When data are only available as aggregated counts, we replace y_{ijk} in (3.1) with $y_{ik} = \sum_{j=1}^{n_{ik}} y_{ijk}$ and p_{ijk} with p_{ik}. When all diagnostic tools are used at each location, we replace x_{ik} with x_i, although this is not a requirement in the development of our methodology.

In the remainder of this section, we review non-spatial methods for joint modelling of the p_{ijk} across multiple diagnostics and a geostatistical modelling approach proposed by Crainiceanu et al. (2008).

3.1   Non-spatial approaches

Existing non-spatial methods for the analysis of data from multiple diagnostics fall within two main classes of statistical models: generalised linear models (GLMs) and their random-effects counterpart, generalised linear mixed models (GLMMs).

Mappin et al. (2015) analysed data on *P. falciparum* prevalence from RDT and microscopy outcomes from sub-Saharan Africa, using a standard probit model

$$\Phi^{-1}(p_{i1}) = \beta_0 + \beta_1 \Phi^{-1}(p_{i2}), \tag{3.2}$$

thus assuming a linear relationship between the p_{ik} on the probit scale. Wanji et al. (2012) used a similar approach for *Loa loa* in order to study the relationship between microscopy and RAPLOA prevalence by replacing the probit link in (3.2) with the logit. This model was also used by Wu et al. (2015) to estimate the relationship between RDT, microscopy and PCR, for each pair of diagnostics. A major limitation of these approaches based on standard GLMs is that they do not account for any over-dispersion that might be induced by unmeasured risk factors.

Coffeng et al. (2013) proposes a bivariate GLMM for joint modelling of data on onchocerciasis nodule prevalence and skin MF prevalence in adult males sampled across 148 villages in 16 African countries. More specifically, the linear predictor of such model can be expressed as

$$\log\left\{\frac{p_{ijk}}{1 - p_{ijk}}\right\} = d_{ij}^\top \beta_k + Z_i + V_{ij}, \tag{3.3}$$

where the random effects terms Z_i and V_{ij} are zero-mean Gaussian variables accounting for unexplained variation between-villages and between-individuals within villages, respectively. Using this approach, Coffeng et al. (2013) estimated a strong positive correlation between nodule and MF prevalence but also reported a variation in the strength of this relationship across study sites.

3.2 The Crainiceanu, Diggle and Rowlingson model

Crainiceanu et al. (2008) proposed a bivariate geostatistical model (henceforth CDRM) to analyse data on microscopy and RAPLOA *Loa loa* prevalence (see Section 2.1). To the best of our knowledge, this is the only existing approach that attempts to model the spatial correlation between two diagnostics.

Let $k = 1$ correspond to the RAPLOA questionnaire, and $k = 2$ to microscopy. To emphasize the spatial context, we now write $p_{ij} = p_j(x_i)$; CDRM can then be expressed as

$$\begin{cases} \text{logit}\{p_1(x_i)\} = d^\top(x_i)\beta + S(x_i) \\ \text{logit}\{p_2(x_i)\} = \alpha_0 + \alpha_1 \text{logit}\{p_1(x_i)\} + Z_i, \end{cases} \tag{3.4}$$

where $\text{logit}(u) = \log\{u/[1-u]\}$, $d(x_i)$ is a vector of spatially varying covariates, $S(x_i)$ is a zero-mean

5

stationary and isotropic Gaussian process and the Z_i are zero-mean independent and identically distributed Gaussian random variables. Crainiceanu et al. (2008) also provide empirical evidence to justify the assumption of a logit-linear relationship between the two diagnostics.

A limitation of the CDRM is that it assumes proportionality on the logit scale between the residual spatial fields associated with RAPLOA and microscopy. In our re-analysis in Section 5, we use a Monte Carlo procedure to test this hypothesis.

4 Two classes of bivariate geostatistical models

We now develop two modelling strategies that address the specific objectives of the two case studies introduced in Section 2. Our focus in this section will be restricted to the case of two diagnostics (hence $K = 2$). We discuss the extension to more than two in Section 7.

4.1 Case I: Predicting prevalence for a gold-standard diagnostic

Let $S_1(x)$ and $S_2(x)$ be two independent stationary and isotropic Gaussian processes; also, let $f_1\{\cdot\}$ and $f_2\{\cdot\}$ be two functions with domain on the unit interval $[0, 1]$ and image on the real line. We propose to model data from two diagnostics, with $k = 2$ denoting the gold-standard, as

$$\begin{cases} f_1\{p_1(x_i)\} = d^\top(x_i)\beta_1 + S_1(x_i) + Z_{i1} \\ f_2\{p_2(x_i)\} = d^\top(x_i)\beta_2 + S_2(x_i) + Z_{i2} + \alpha f_1\{p_1(x_i)\}. \end{cases} \tag{4.1}$$

In our applications, we specify exponential correlation functions for $S_k(x)$, $k = 1, 2$, hence

$$\text{cov}\{S_k(x), S_k(x')\} = \sigma_k^2 \exp\{\|x - x'\|/\phi_k\},$$

where σ_k^2 is the variance of $S_k(x)$ and ϕ_k is a scale parameter regulating how fast the spatial correlation decays to zero for increasing distance. Finally, we use τ_k^2 to denote the variance of the Gaussian noise Z_{ik}.

Selection of suitable functions f_1 and f_2 can be carried out, for example, by exploring the association between the empirical prevalences of the two diagnostics in order to identify transformations that render their relationship approximately linear. Alternatively, subject matter knowledge could be used to constrain the admissible forms for f_1 and f_2; see, for example, Irvine et al. (2016) who derive a functional relationship between MF and an immuno-chromatographic test for prevalence of lymphatic filariasis by making explicit assumptions on the distribution of worms and their

reproductive rate in the general population.

The proposed model in (4.1) is more flexible than the CDRM because (i) it allows for diagnostic-specific unstructured variation Z_{ik} and, more importantly, (ii) relaxes the assumption of proportionality between the residual spatial fields of the two diagnostics through the introduction of $S_2(x)$.

4.2 Case II: Joint prediction of prevalence from two complementary diagnostics

Let $S_1(x)$ and $S_2(x)$ be two independent Gaussian processes, and Z_{ik} Gaussian noise, each having the same properties as defined in the previous section. We now introduce a third stationary and isotropic Gaussian process $T(x)$ having unit variance and exponential correlation function with scale parameter ϕ_T.

Our proposed approach for joint prediction of prevalence from two diagnostics, when both are of interest, is expressed by the following equation for the linear predictor

$$f_k\{p_{jk}(x_i)\} = d_{ij}^\top \beta_k + \nu_k[S_k(x_i) + T(x_i)] + Z_{ik}. \tag{4.2}$$

The spatial processes $S_k(x)$ and $T(x)$ accounts for unmeasured risk factors that are specific to each and common to both diagnostics, respectively. The resulting variogram for the linear predictor is

$$
\begin{aligned}
\gamma_k(u) &= \mathrm{E}\Big[\Big\{\big(\nu_k(S_k(x)+T(x))+Z_k(x)\big) - \big(\nu_k(S_k(x')+T(x'))+Z_k(x')\big)\Big\}^2\Big] \\
&= \tau_k^2 + \nu_k^2[1 - \exp(-u/\phi_T) + \sigma_k^2\{1 - \exp(-u/\phi_{S_k})\}],
\end{aligned} \tag{4.3}
$$

and the cross-variogram between the linear predictors of the two diagnostics is

$$
\begin{aligned}
\gamma_{1,2}(u) &= \mathrm{E}\Big[\Big\{\big(\nu_1(S_1(x)+T(x))+Z_k(x)\big) - \big(\nu_2(S_2(x')+T(x'))+Z_k(x')\big)\Big\}^2\Big] \\
&= 0.5\{\tau_1^2 + \tau_2^2 + \nu_1^2(1+\sigma_1^2) + \nu_2^2(1+\sigma_2^2)\} - \sigma_1\sigma_2\exp(-u/\phi_T).
\end{aligned} \tag{4.4}
$$

Given the relatively large number of parameters, fitting the model may require a pragmatic approach. In order to identify a parsimonious model for the data, we recommend an incremental modelling strategy, whereby a simpler model is used in a first analysis (e.g. by setting $S_k(x) = 0$ for all x) and more complexity is then added in response to an unsatisfactory validation check, as described below.
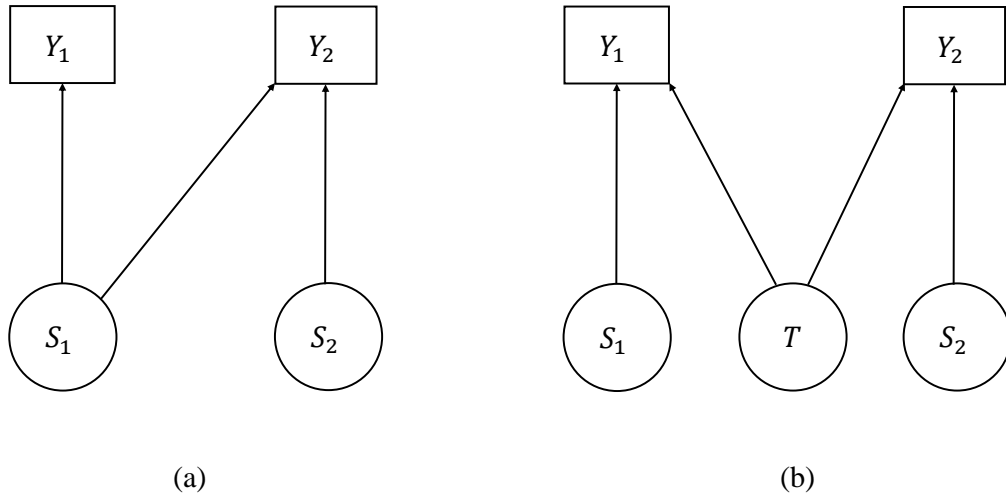
Figure 4.1: Directed acyclic graphs for the bivariate geostatistical models in (4.1) (left panel) and (4.2) (right panel). Circles and squares identify latent variables and the outcome random variables, respectively.

4.3   Comparison between the two models

Figure 4.1 gives two directed acyclic graph representations of the models in (4.1) (left panel) and (4.2) (right panel), showing their distinctive asymmetric and symmetric structures. In the first model, stochastic independence between the two diagnostics is simply recovered by setting the parameter $\alpha = 0$. If this is a scientifically relevant hypothesis, we can test it through the likelihood ratio. In the second model, independence can only be achieved if $T(x) = 0$ for all x. We do not consider this to be a credible assumption for the malaria application of Section 2.1.

4.4   Inference and model validation

We carry out parameter estimation for both the asymmetric and symmetric models using Monte Carlo Maximum Likelihood (MCML) (Geyer, 1991; Christensen, 2004). To carry out spatial predictions at a set of unobserved locations, we plug the MCML estimates into a Markov Chain Monte Carlo algorithm for simulation from the distribution of the random effects conditional on the data.

We summarise our predictive inferences on prevalence using the mean, standard deviation, and exceedance probabilities, i.e. the probability that the predictive distribution of prevalence exceeds a predefined threshold. Details of the derivation and approximation of the log-likelihood function are given in Web Appendix A.

For model validation we propose the following procedure. We first re-write both models in general form

$$f_k\{p_{jk}(x_i)\} = \mu_{ijk} + W_k(x_i), \tag{4.5}$$

where μ_{ijk} is the mean component expressed as a regression on the available covariates. In (4.5), if we set $W_1(x_i) = S_1(x) + Z_{i1}$ and $W_2(x_i) = S_2(x_i) + Z_{i2} + \alpha\{f_1(x_i)\}$, then (4.5) reduces to the asymmetric model (4.1); if, instead, $W_k(x_i) = \nu_k(S_k(x_i) + T(x_i)) + Z_{ik}$, we recover the symmetric model (4.2).

We define the empirical variogram of $W_k(x)$ to be

$$\hat{\gamma}_k(u) = \frac{1}{2|N(u)|} \sum_{(i,j)\in N(u)} \{\hat{W}_k(x_i) - \hat{W}_k(x_j)\}^2, \tag{4.6}$$

where $N(u) = \{(i,j) : ||x_i - x_j|| = u, i \neq j\}$ and $\hat{W}_k(x_i)$ is the mean of distribution of $W_k(x_i)$ conditioned to the data. To test whether the adopted spatial structure for $W_k(x)$ is compatible with the data, we then proceed through the following steps.

Step 0. Obtain $\hat{W}_k(x_i)$ from two separate standard geostatistical models (i.e. $W_k(x_i) = S_k(x_i) + Z_{ik}$, where $S_k(x)$, $k = 1, 2$ are independent processes) and compute the empirical variogram $\hat{\gamma}_k$, for $k = 1, 2$.

Step 1. Simulate prevalence data as in (3.1) from the adopted model for $W_k(x)$ by plugging-in the MCML estimates. Fit separate standard geostatistical models as in *Step 0* and compute the empirical variogram for the simulated dataset.

Step 2. Repeat *Step 1* a large enough number of times, say M.

Step 3. Use the resulting M empirical variograms to generate 95% confidence intervals at each of a set of pre-defined distance bins.

If the empirical variogram in *Step 0* falls fully or partly outside the 95% confidence intervals, we conclude that the model is not able to capture the spatial structure of the data satisfactorily.

5 Application I: Re-analysis of the *Loa loa* data

A total of 223 villages were sampled in the four study sites (see Web Figure 1 in Web Appendix B). The empirical prevalences from the RAPLOA and microscopy tests show a clear, approximately linear relationship when both are logit-transformed (see Web Figure 3 in Web Appendix B). Each of the two also exhibits a highly non-linear relationship with surface elevation (see Web Figure 4 in Web Appendix B), which we capture using a piecewise linear spline with knots at 750 meters and 1015 meters.

We consider the two following models.

- Model 1: a slightly modified, more flexible, version of the CDRM, given by

$$\begin{cases} \text{logit}\{p_1(x_i)\} = \mu_1(x_i) + S_1(x_i) + Z_{i1} \\ \text{logit}\{p_2(x_i)\} = \mu_2(x_i) + \alpha\text{logit}\{p_1(x_i)\} + Z_{i2} \end{cases}, \tag{5.1}$$

where

$$\begin{aligned} \mu_k(x_i) &= \beta_{k,0} + \beta_{k,1}\min\{e(x_i), e_1\} + \beta_{k,2}I(e(x_i) > e_1)\min\{e(x_i) - e_2, e_2 - e_1\} \\ &\quad + \beta_{k,3}\max\{e(x_i) - e_2, 0\}, \quad k = 1, 2, \end{aligned}$$

where $e(x)$ denotes the elevation in meters at location x, $e_1 = 750$, $e_2 = 1015$ and $I(\mathcal{P})$ is an indicator function which takes value 1 if $\mathcal{P}$ is true and 0 otherwise. In this parameterisation, $\beta_{k,1}$ is the effect of elevation on prevalence below 750 meters, $\beta_{k,2}$ its effect between 760 and 1015 meters, and $\beta_{k,3}$ its effect above 1015 meters.

- Model 2: obtained by incorporating an additional spatial process $S_2(x)$, independent of $S_1(x)$, into the linear predictor for microscopy in Model 1 to give

$$\text{logit}\{p_2(x_i)\} = \mu_2(x_i) + S_2(x_i) + \alpha\text{logit}\{p_1(x_i)\} + Z_{i2} \tag{5.2}$$

5.1 Results

Table 1 reports the MCML estimates obtained for Models 1 and 2. We observe that all parameters common to both Models 1 and Model 2 have comparable point and interval estimates, except for τ_1^2 which has a substantially narrower 95% confidence interval under Model 1 than Model 2.

As expected, both models show a significant and positive logit-linear relationship between

RAPLOA and miscroscopy. However, Model 2, which include the additional spatial process $S_2(x)$, is also able to capture spatial variation in microscopy prevalence on a scale of about 24 meters.

We use the validation procedure of Section 6.1 to test which of the two models better fits the spatial structure of the data. The results (see Web Figure 5) show a satisfactory assessment of Model 2, whereas for Model 1 the empirical variogram for microscopy partly falls outside the 95% confidence band, questioning its validity.

We now compare the predictive inferences on microscopy prevalence between the two models in order to assess whether the introduction of $S_2(x)$ makes a tangible difference. Figure 5.1 shows the point estimates for microscopy prevalence and the exceedance probabilities for a 20% prevalence threshold under Model 1 (upper panels), under Model 2 (middle panels), and the difference between the two (lower panels). Overall, the predicted spatial pattern in prevalence from the two models is similar, but with substantial local differences. The difference between the point estimates for prevalence ranges from -0.12 to 0.13, while the difference between the two exceedance probabilities ranges from -0.44 to 0.59.

5.2 Simulation Study

We carry out a simulation study in order to quantify the effects on the predictive inferences for prevalence when ignoring microscopy-specific residual spatial variation. To this end, we compare the predictive performances of Model 1 and Model 2 at 20 unobserved locations corresponding to the centroids of 20 clusters (shown as red points in Web Figure 1) that we identify using the k-means algorithm (Hartigan and Wong, 1979) and proceed as follows. We simulate 10,000 Binomial data-sets under Model 2 by setting its parameters to the estimates of Table 1 and fit both models. We then carry out predictions for microscopy prevalence over the 20 unobserved locations. We summarise the results at each of the 20 prediction locations using the 95% coverage probability (CP), the root-mean-square-error (RMSE) and the 95% predictive interval length (PIL). Table 2 shows the three metrics averaged over the 20 locations for Model 1 and Model 2. The CP of Model 1 (77.5%) is well below its nominal level of 95%. This is also reflected by a smaller PIL for Model 1, suggesting that this provides unreliably narrow 95% predictive intervals for prevalence. Finally, we note that Model 1 also has a larger RMSE than Model 2.
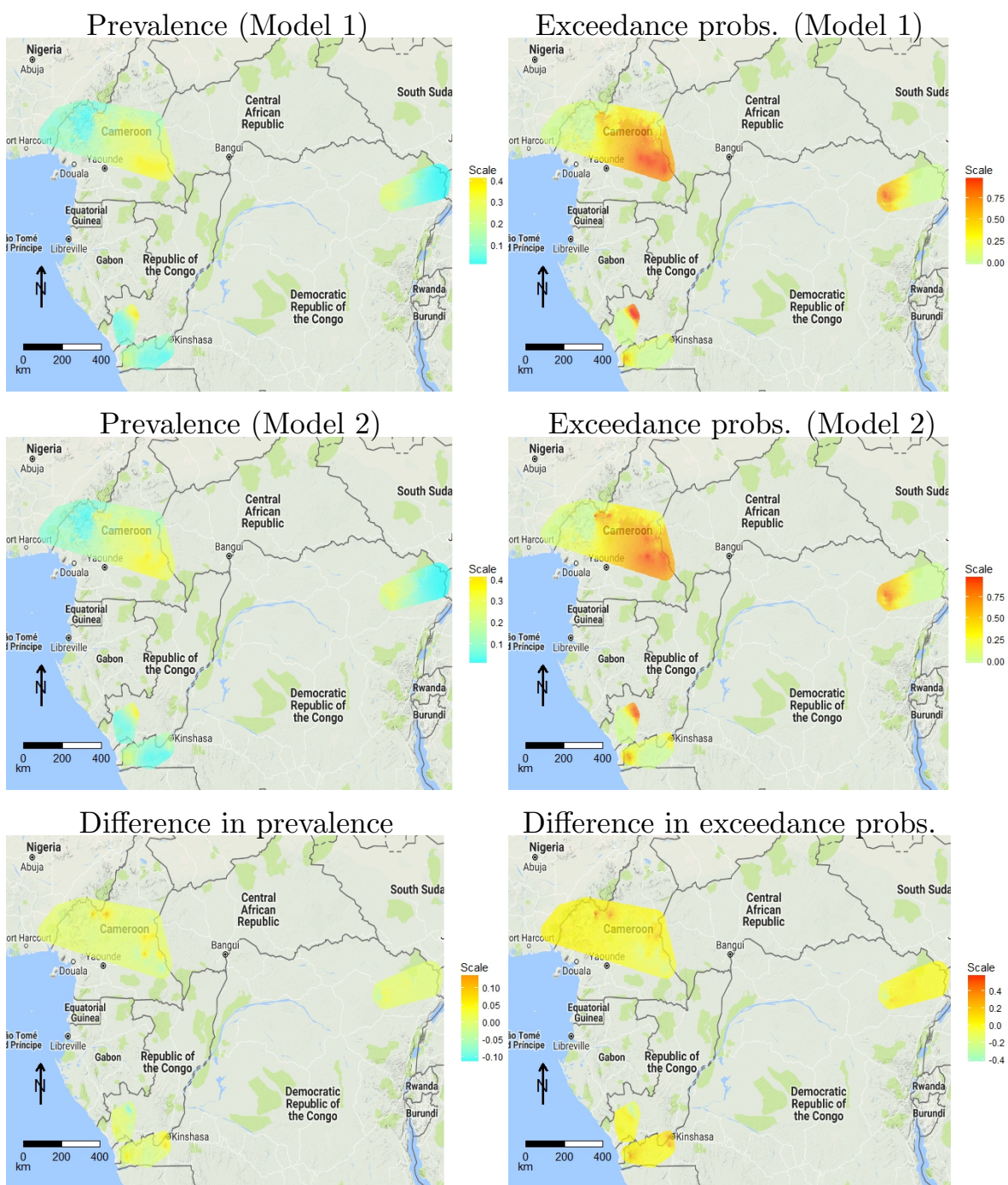
Figure 5.1: Predictive mean of *Loa loa* microfilariae prevalence (left panels) and probabilities of exceeding a 20% prevalence threshold (right panels), for Model 1 (top panels) and Model 2 (middle panels) of Section 5. The bottom panels show the differences between the predictive surfaces of Model 2 and Model 1.

6 Application II: Joint prediction of *Plasmodium falciparum* prevalence using RDT and PCR

The malaria data consist of 3,587 individuals sampled across 949 locations (see Web Figure 2). The outcomes from RDT ($k = 1$) and PCR ($k = 2$) were concordant in 92.4% of all the individuals tested for *P. falciparum*. This suggests that estimating components of residual spatial variation that are unique to each diagnostic may be difficult. For this reason our model for the data takes the following form

$$f_k(p_{jk}(x_i)) = \beta_{k,0} + \sum_{l=1}^{3} \beta_{k,l} d_{ij,l} + \nu_k T(x_i), \tag{6.1}$$

where: $d_{ij,1}$ is a binary variable taking value 1 if the j-th individual at x_i is a male and 0 otherwise; $d_{ij,2} = \min\{a_{ij}, 5\}$ and $d_{ij,3} = \max\{a_{ij} - 5, 0\}$, i.e. the effect of age, a_{ij}, is modelled as a linear spline with a knot at 5 years.

6.1 Results

Table 3 reports point estimates and 95% confidence intervals for the model parameters. Gender has a significant effect on PCR prevalence, but its effect on RDT prevalence is not significant at the conventional 5% alpha level. The effect of age is comparable between the two diagnostics, with the probability of a positive test increasing with age up to 5 years and decreasing thereafter. The estimated variance component, $\hat{\nu}_1^2 = 0.230$, associated with RDT is about three times that for PCR, $\hat{\nu}_2^2 = 0.081$. The spatial process $T(x)$, common to both diagnostics, accounts for spatial variation in malaria prevalence up to a scale of about 11.6 kilometers, beyond which the correlation falls below 0.05. The variogram-based validation procedure of Section does not show any strong evidence against the fitted model (see Web Figure 6 in Web Appendix B).

To quantify the benefit of carrying out a joint analysis for RDT and PCR, we compare the predictive inferences for prevalence that are obtained under two scenarios: (i) the fitted model in (6.1); (ii) separate fitted models that ignore the cross-correlation between the outcomes of the two diagnostic tests. Figure 6.1 shows the point predictions and standard errors for RDT and PCR prevalences for five-year-old male children under scenarios (i) (left panels) and (ii) (right panels). We observe that the point predictions for prevalence under the two models are strongly similar but the joint model in (6.1), as expected, yields smaller standard errors throughout the study area.

Having chosen (6.1) as the best model, we compare the exceedance probabilities (EPs) for a 10% threshold between RDT and PCR. Using each of the two diagnostics, we then identify malaria
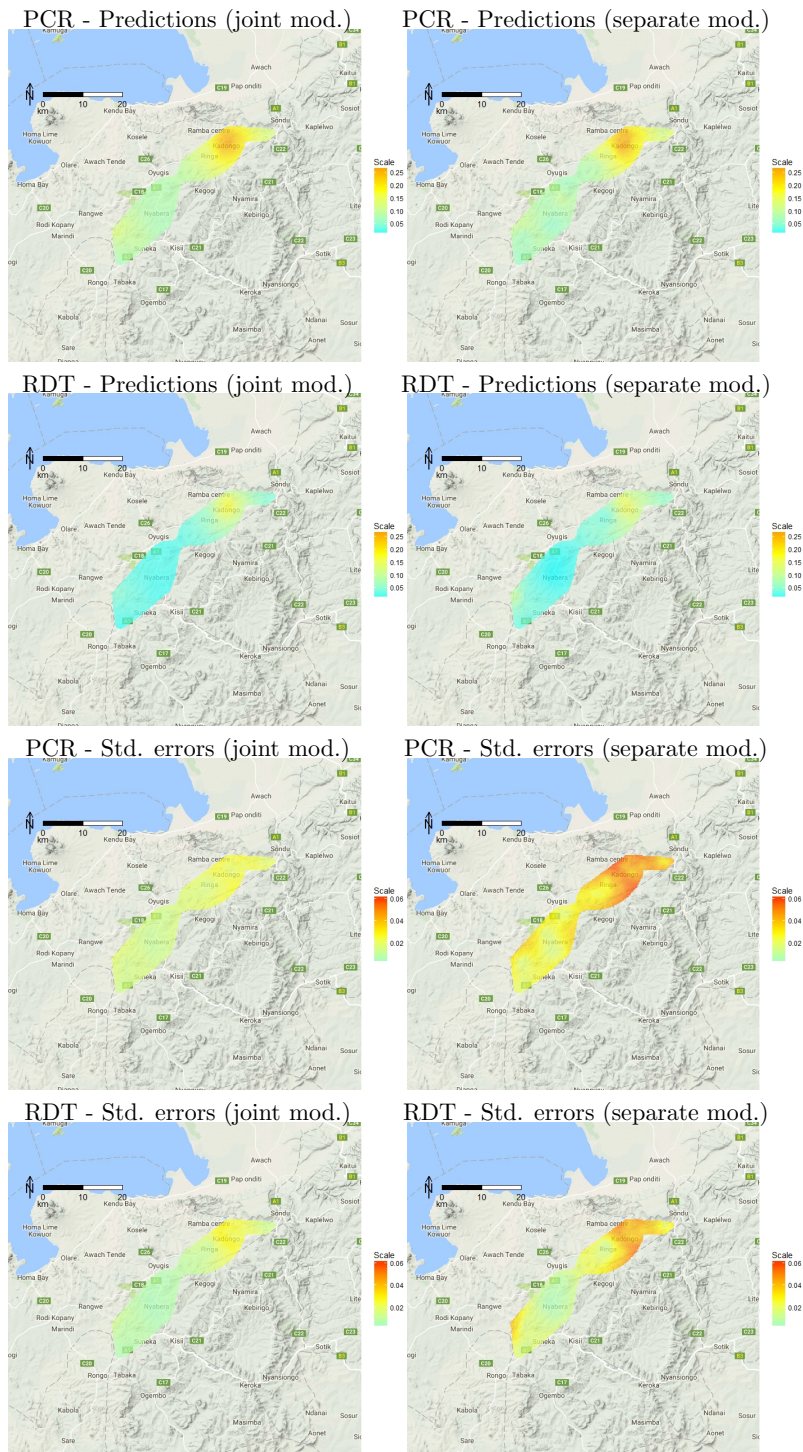
13

Figure 6.1: Point predictions (first and second rows) and standard errors (third and fourth rows) of *P. falciparum* prevalence for five-year-old children under the joint geostatistical model in (6.1) (left panels) and two separate geostatistical models (right panels) for RDT and PCR prevalence.

14

hotspots, as the sets of locations such that their EP is at least 90%. Figure 7 of Web Appendix B shows that PCR identifies a considerably larger hotspot in the north east of the study area than does RDT, and a smaller hotspot in the south west that is undetected by RDT. These results are consistent with the main findings of Mogeni et al. (2017).

7 Conclusions and extensions

We have developed a flexible geostatistical framework to model reported disease counts from multiple diagnostics and have distinguished two main classes of problems: (1) prediction of prevalence as defined by a gold-standard diagnostic using data obtained from a more feasible low-cost, but potentially biased, alternative; (2) joint prediction of prevalences as defined by two diagnostic tests. As the burden of disease declines in endemic regions, the use of multiple transmission metrics and diagnostics becomes necessary in order to better inform and adapt control strategies. It is thus important to develop suitable methods of inference that allow the borrowing of strength of information across multiple diagnostics. As our study has shown, the main benefit of this approach is a reduction in the uncertainty associated with the predictive inferences on disease risk.

Our application to Loiasis mapping has shown the importance of acknowledging the existence of residual spatial variation specific to each diagnostic test. Through a simulation study, we have also shown that ignoring this source of extra-Binomial variation can lead to unreliably narrow prediction intervals for prevalence, with actual coverages falling well below their nominal level.

The second application on malaria mapping has highlighted the benefits of a joint analysis of data from two diagnostic tests when both are of scientific interest. A joint model can yield estimates of prevalence with smaller standard errors than estimates obtained from two separate geostatistical models.

Although we have only considered the case of two diagnostic tests throughout the paper, our methodology can be easily extended to more than two. However, the nature of the extension will be dependent on the specific context and scientific goal. For example, a natural extension of the models of Section 4.1 would be to use multiple biased diagnostic tools (for $k = 1, \ldots, K - 1$) to better predict a gold-standard ($k = K$). In this case, the cross-correlation between the outcomes of the biased diagnostic tests could be modelled using the symmetric structure of the model in Section 4.2, while preserving an asymmetric form for the linear predictor of the gold-standard. Formally,

this is expressed as

$$
\begin{cases}
f_k\{p_{jk}(x_i)\} = d_{ijk}^\top \beta_k + \nu_k \left[S_k(x_i) + T(x_i) \right] + Z_{ik}, \quad k = 1, \ldots, K-1 \\
f_K\{p_{jK}(x_i)\} = d_{ijK}^\top \beta_K + S_K(x_i) + Z_{iK} + \sum_{k=1}^{K-1} \alpha_k f_k\{p_j(x_i)\}
\end{cases}
. \tag{7.1}
$$

However, we would be wary of attempting to fit this, or other comparably complex models without an initial exploratory analysis that might help to understand the extent of the cross-correlations between the outcomes of different diagnostics, with a view to reducing the dimensionality of the model.

Acknowledgements

Supplementary Materials

Web Appendix A, referenced in Section 6.1; and Web Appendix B, referenced in Sections 2, 5, and 6 are available with this paper at the Biometrics website on Wiley Online Library.

References

Bhatt, S., Weiss, D., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C., Henry, A., Eckhoff, P., et al. (2015). The effect of malaria control on plasmodium falciparum in africa between 2000 and 2015. *Nature* **526,** 207–211.

Boussinesq, M., Gardon, J., Gardon-Wendel, N., Kamgno, J., Ngoumou, P., and Chippaux, J.-P. (1998). Three probable cases of loa loa encephalopathy following ivermectin treatment for onchocerciasis. *The American journal of tropical medicine and hygiene* **58,** 461–469.

Boussinesq, M., Gardon, J., Kamgno, J., Pion, S., Gardon-Wendel, N., and Chippaux, J.-P. (2001). Relationships between the prevalence and intensity of loa loa infection in the central province of cameroon. *Annals of Tropical Medicine & Parasitology* **95,** 495–507.

Christensen, O. F. (2004). Monte carlo maximum likelihood in model-based geostatistics. *Journal of computational and graphical statistics* **13,** 702–718.

Coffeng, L. E., Pion, S. D., O'Hanlon, S., Cousens, S., Abiose, A. O., Fischer, P. U., Remme, J. H., Dadzie, K. Y., Murdoch, M. E., De Vlas, S. J., et al. (2013). Onchocerciasis: the pre-control association between prevalence of palpable nodules and skin microfilariae. *PLoS neglected tropical diseases* **7,** e2168.

Crainiceanu, C. M., Diggle, P. J., and Rowlingson, B. (2008). Bivariate binomial spatial modeling of loa loa prevalence in tropical africa. *Journal of the American Statistical Association* **103,** 21–37.

Diggle, P. J. and Giorgi, E. (2016). Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* **111,** 1096–1120.

Diggle, P. J., Tawn, J., and Moyeed, R. (1998). Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47,** 299–350.

Gething, P. W., Elyazar, I. R., Moyes, C. L., Smith, D. L., Battle, K. E., Guerra, C. A., Patil, A. P., Tatem, A. J., Howes, R. E., Myers, M. F., et al. (2012). A long neglected world malaria map: Plasmodium vivax endemicity in 2010. *PLoS neglected tropical diseases* **6,** e1814.

Geyer, C. J. (1991). Markov chain monte carlo maximum likelihood.

Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28,** 100–108.

Irvine, M. A., Njenga, S. M., Gunawardena, S., Njeri Wamae, C., Cano, J., Brooker, S. J., and Deirdre Hollingsworth, T. (2016). Understanding the relationship between prevalence of microfilariae and antigenaemia using a model of lymphatic filariasis infection. *Transactions of The Royal Society of Tropical Medicine and Hygiene* **110,** 118–124.

Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., et al. (2012). Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet* **379,** 2151–2161.

Mappin, B., Cameron, E., Dalrymple, U., Weiss, D. J., Bisanzio, D., Bhatt, S., and Gething, P. W. (2015). Standardizing plasmodium falciparum infection prevalence measured via microscopy versus rapid diagnostic test. *Malaria journal* **14,** 460.

Mogeni, P., Williams, T. N., Omedo, I., Kimani, D., Ngoi, J. M., Mwacharo, J., Morter, R., Nyundo, C., Wambua, J., Nyangweso, G., et al. (2017). Detecting malaria hotspots: A comparison of rapid diagnostic test, microscopy, and polymerase chain reaction. *The Journal of infectious diseases* **216,** 1091–1098.

Murray, C. J., Ortblad, K. F., Guinovart, C., Lim, S. S., Wolock, T. M., Roberts, D. A., Dansereau, E. A., Graetz, N., Barber, R. M., Brown, J. C., et al. (2014). Global, regional, and national incidence and mortality for hiv, tuberculosis, and malaria during 1990–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet* **384,** 1005–1070.

Song, P. X.-K., Li, M., and Yuan, Y. (2009). Joint regression analysis of correlated data using gaussian copulas. *Biometrics* **65,** 60–68.

Stevenson, J. C., Stresman, G. H., Baidjoe, A., Okoth, A., Oriango, R., Owaga, C., Marube, E., Bousema, T., Cox, J., and Drakeley, C. (2015). Use of different transmission metrics to describe malaria epidemiology in the highlands of western kenya. *Malaria journal* **14,** 418.

Takougang, I., Meremikwu, M., Wandji, S., Yenshu, E. V., Aripko, B., Lamlenn, S. B., Eka, B. L., Enyong, P., Meli, J., Kale, O., et al. (2002). Rapid assessment method for prevalence and intensity of loa loa infection. *Bulletin of the World Health Organization* **80,** 852–858.

Tangpukdee, N., Duangdee, C., Wilairatana, P., and Krudsood, S. (2009). Malaria diagnosis: a brief review. *The Korean journal of parasitology* **47,** 93.

Wanji, S., Akotshi, D. O., Mutro, M. N., Tepage, F., Ukety, T. O., Diggle, P. J., and Remme, J. H. (2012). Validation of the rapid assessment procedure for loiasis (raploa) in the democratic republic of congo. *Parasites & vectors* **5,** 25.

World Health Organization (2017). World malaria report 2017.

Wu, L., van den Hoogen, L. L., Slater, H., Walker, P. G., Ghani, A. C., Drakeley, C. J., and Okell, L. C. (2015). Comparison of diagnostics for the detection of asymptomatic plasmodium falciparum infections to inform control and elimination strategies. *Nature* **528,** S86.

Zouré, H. G., Noma, M., Tekle, A. H., Amazigo, U. V., Diggle, P. J., Giorgi, E., and Remme, J. H. (2014). The geographic distribution of onchocerciasis in the 20 participating countries of the african programme for onchocerciasis control:(2) pre-control endemicity levels and estimated number infected. *Parasites & vectors* **7,** 326.

Zouré, H. G. M., Wanji, S., Noma, M., Amazigo, U. V., Diggle, P. J., Tekle, A. H., and Remme, J. H. F. (2011). The geographic distribution of loa loa in africa: Results of large-scale implementation of the rapid assessment procedure for loiasis (raploa). *PLOS Neglected Tropical Diseases* **5,** 1–11.

Table 1: Monte Carlo maximum likelihood estimates and associated 95% confidence intervals for the fitted Model 1 and Model 2 to the *Loa loa* data; see Section 5 for more details.

Parameter	Model 1	Model 2
$\beta_{1,0}$	-0.791	-0.763
	(-1.984, 0.402)	(-1.963, 0.437)
$\beta_{1,1} \times 10^3$	0.515	0.588
	(-0.977, 2.008)	(-0.922, 2.098)
$\beta_{1,2} \times 10^3$	-3.529	-3.412
	(-7.314, 0.255)	(-7.155 , 0.331)
$\beta_{1,3} \times 10^3$	-0.110	-0.059
	(-1.531, 1.312)	(-1.501 , 1.382)
$\beta_{2,0}$	-1.762	-1.736
	(-2.075, -1.449)	(-2.244, -1.229)
$\beta_{2,1} \times 10^3$	0.208	0.126
	(-0.386, 0.802)	(-0.799, 1.050)
$\beta_{2,2} \times 10^3$	-0.223	-0.039
	(-2.023, 1.576)	(-2.944 , 2.865)
$\beta_{2,3} \times 10^3$	-0.591	-0.612
	(-1.666, 0.485)	(-2.429, 1.205)
σ_1^2	1.581	1.617
	(0.669, 3.738)	(0.679, 3.851)
σ_2^2	—	0.216
		(0.111, 0.419)
ϕ_1	182.037	187.388
	(64.657, 512.512)	(65.171, 538.807)
ϕ_2	—	23.686
		(6.150, 91.220)
τ_1^2	0.205	0.324
	(0.081, 0.521)	(0.052, 6.229)
τ_2^2	0.324	0.104
	(0.055, 5.873)	(0.018, 5.797)
α	1.005	1.017
	(0.902, 1.107)	(0.939, 1.095)

Table 2: Results of the simulation study including the 95% coverage probability (CP), the root-mean-square-error (RMSE), the 95% predictive interval length (PIL) averaged over the 20 unobserved locations. For more details, see the main text in Section 5.2.

	CP	RMSE	PIL
Model 1	0.770	4.948	0.140
Model 2	0.943	3.932	0.185

Table 3: Monte Carlo maximum likelihood estimates and associated 95% confidence intervals for the model in (6.1) fitted to the malaria data.

Parameter	RDT ($k = 1$)	PCR ($k = 2$)
$\beta_{k,0}$	-6.186	-4.373
	(-7.234, -5.138)	(-17.008, 8.261)
$\beta_{k,1}$	-0.003	0.251
	(-0.415, 0.395)	(0.009, 0.494)
$\beta_{k,2}$	0.261	0.220
	(0.070, 0.453)	(0.095, 0.344)
$\beta_{k,3}$	-0.059	-0.020
	(-0.081, -0.037)	(-0.028, -0.012)
ν_k^2	0.230	0.081
	(0.145, 0.364)	(0.052, 0.126)
ϕ_T	11.581	
	(10.618, 12.63)	