Appendix to "Integrated nested Laplace approximations for extended latent Gaussian models with application to the Naomi HIV model"

Corresponding author: Adam Howes (ath19@ic.ac.uk)

Contents

S1 Simplified Naomi model description

In this section we describe more fully the simplified Naomi model considered in the main text. In doing so we draw from the supplementary material of Jeffrey W. Eaton et al. (2021).

S1.1 Background

S1.1.1 Indexing

Let ν the most recent national household survey with HIV testing which has taken place in the country of interest. Let $x \in \{1, \ldots, n\}$ refer to a district located within the Spectrum (Stover et al. 2010) region R_x. Let $s \in \{\text{F}, \text{M}\}$ be sex, and $a \in \{\text{0-5}, \text{5-10}, \ldots, \text{75-80}, \text{80+}\}$ be five-year age bands. As short-hand, we write $a = l$ to refer to the age band with lower bound l, e.g. $a = 20$ for $a = \text{20-25}$. We index the known quantity population size $N_{x,s,a}$ by district, sex and age-band, as well as the following unknown quantities: HIV prevalence $\rho_{x,s,a} \in [0, 1]$, ART coverage $\alpha_{x,s,a} \in [0, 1]$, annual HIV incidence rate $\lambda_{x,s,a} > 0$, and proportion of HIV positive persons recently infected $\kappa_{x,s,a} \in [0, 1]$. Sometimes data are observed at an aggregate level, rather than the more granular modelled level. In this instance, we use $\{\cdot\}$ to generically refer to a aggregate set over which an observation is made, e.g. $\{a\} = \{\text{15-19}, \ldots, \text{45-49}\}$ for the adult age range 15-49. In the main text we use a more concise, but less descriptive, approach. Finally, we let $\sum_{\{x\}}$ be used as shorthand for $\sum_{x \in \{x\}}$, and likewise for s and a.

S1.1.2 Structured random effects

Let u be a generic random effect. We use use structured random effects to enable partial pooling of information across units assessed as being similar, such as neighbouring districts or adjacent age-bands. We specify the first order auto-regressive model by $u \sim \text{AR1}(\sigma, \phi)$ such that

$$\begin{aligned} u_1 &\sim \left(0, \frac{1}{1-\rho^2}\right), \\ u_i &= \rho u_{i-1} + \epsilon_t, \quad i = 2, \ldots \end{aligned}$$

where $\epsilon_i \sim \mathcal{N}(0, 1)$ is Gaussian white noise, and $|\rho| < 1$ is the lag-one correlation parameter. We use $u \sim \text{ICAR}(\sigma)$ to refer to the Besag intrinsic conditional auto-regressive model (ICAR) (Besag, York, and Mollié 1991) with full conditionals

$$u_i \,|\, u_{-i} \sim \mathcal{N}\left(\frac{\sum_{j:j\sim i} u_j}{n_{\delta i}}, \frac{\sigma^2}{n_{\delta i}}\right),$$

where u_{-i} is u with the ith unit removed, $j \sim i$ if the units are defined as adjacent, $n_{\delta i} = |\{j : j \sim i\}|$ is the total number of adjacent units, and $\sigma > 0$ is the marginal standard deviation. We follow recommendations of Freni-Sterrantino, Ventrucci, and Rue (2018) on scaling of precision matrices, disconnected adjacency graph components, and islands. For the reparameterised Besag-York-Mollie model (BYM2) (Simpson et al. 2017) we write $u \sim \text{BYM2}(\sigma, \phi)$, where u is comprised of a spatially structured ICAR component $v^\star$ with proportion $\phi \in (0, 1)$ and spatially unstructured IID component $w^\star$ with proportion $1 - \phi$, both scaled to have generalised variance equal to one, and $\sigma > 0$ is the marginal standard deviation such that

$$u = \sigma\left(\sqrt{\phi} \cdot v^\star + \sqrt{1-\phi} \cdot w^\star\right).$$

S1.1.3 Complex survey design

We assume the household survey was run according to a complex survey design where each individual $j \in U$ has non-zero probability $\pi_j \in (0, 1)$ of appearing in the sample $S \subseteq U$. Suppose we observe an outcome y_j for $j \in S$. Let $w_j = 1/\pi_j \times 1/\omega_j$ be design weights, where ω_j is a non-response factor, then the weighted mean

$$\hat{y} = \frac{\sum_{j\in S} w_j y_j}{\sum_{j\in S} w_j}.$$

The Kish effective sample size is

$$m = \frac{\left(\sum_{j \in S} w_j\right)^2}{\sum_{j \in S} w_j^2}.$$

We make these computations using the `survey` R package (Lumley 2004).

S1.2 Process specification

	Model section	Latent field	Hyperparameter
S1.2.1	HIV prevalence	$36 + 3n$	9
S1.2.2	ART coverage	$36 + 3n$	9
S1.2.3	HIV incidence rate	$2 + n$	3
S1.2.4	ANC testing	$4 + 2n$	2
S1.2.5	ART attendance	n	1
	Total	$78 + 7n$	24

Table S1: The numer of latent field parameters and hyperparameters in each model section.

In this section, we describe the hyperparameter and latent field process specification for the model. Whereas in the main text process and likelihood specifications are written together, here we consider the likelihood equations separately in Section S1.3. Table S1.2 gives the number of latent field parameters and hyperparameters in each section of the model. Given a country with n districts, then the total number of unique district-age-sex combinations is $n \times 2 \times 17 = 34n$. In the case of Malawi then $n = 32$ such that the total number of latent field parameters is 491 and the total number of hyperparameters is 24.

S1.2.1 HIV prevalence

We model HIV prevalence $\rho_{x,s,a} \in [0, 1]$ on the logit scale using the linear predictor

$$\text{logit}(\rho_{x,s,a}) = \beta_0^\rho + \beta_S^{\rho,s=\text{M}} + u_a^\rho + u_a^{\rho,s=\text{M}} + u_x^\rho + u_x^{\rho,s=\text{M}} + u_x^{\rho,a<15} + \eta_{R_x,s,a}^\rho$$

where $\beta_0^\rho \sim \mathcal{N}(0, 5)$ is an intercept term, $\beta_s^{\rho,s=\text{M}} \sim \mathcal{N}(0, 5)$ is the difference in logit prevalence for men compared to women, $u_a^\rho \sim \text{AR1}(\sigma_A^\rho, \phi_A^\rho)$ are age random effects for women, $u_a^{\rho,s=\text{M}} \sim \text{AR1}(\sigma_{AS}^\rho, \phi_{AS}^\rho)$ are age random effects for the difference in logit prevalence for men compared to women age a, $u_x^\rho \sim \text{BYM2}(\sigma_X^\rho, \phi_X^\rho)$ are spatial random effects for women, $u_x^{\rho,s=\text{M}} \sim \text{BYM2}(\sigma_{XS}^\rho, \phi_{XS}^\rho)$ are spatial random effects for the difference in logit prevalence for men compared to women in district x, $u_x^{\rho,a<15} \sim \text{ICAR}(\sigma_{XA}^\rho)$ are spatial random effects for the ratio of paediatric prevalence to adult women prevalence, and $\eta_{R_x,s,a}^\rho$ are fixed offsets specifying assumed odds ratios for prevalence outside the age ranges for which data are available. We use the prior distributions $\mathcal{N}^+(0, 2.5)$ for the standard deviation terms $\{\sigma_A^\rho, \sigma_{AS}^\rho, \sigma_X^\rho, \sigma_{XS}^\rho, \sigma_{XA}^\rho\}$, $\mathcal{U}(-1, 1)$ for the AR1 correlation parameters $\{\phi_A^\rho, \phi_{AS}^\rho\}$, and Beta$(0.5, 0.5)$ for the BYM2 proportion parameters $\{\phi_X^\rho, \phi_{XS}^\rho\}$.

S1.2.2 ART coverage

We model ART coverage $\alpha_{x,s,a} \in [0, 1]$ on the logit scale using the linear predictor

$$\text{logit}(\alpha_{x,s,a}) = \beta_0^\alpha + \beta_S^{\alpha,s=\text{M}} + u_a^\alpha + u_a^{\alpha,s=\text{M}} + u_x^\alpha + u_x^{\alpha,s=\text{M}} + u_x^{\alpha,a<15} + \eta_{R_x,s,a}^\alpha$$

with terms and priors analogous to the HIV prevalence process model in Section S1.2.1 above.

S1.2.3 HIV incidence rate

We model HIV incidence rate $\lambda_{x,s,a} > 0$ on the log scale using the linear predictor

$$\log(\lambda_{x,s,a}) = \beta_0^\lambda + \beta_S^{\lambda,s=\text{M}} + \log(\rho_x^{\text{15-49}}) + \log(1 - \omega \cdot \alpha_x^{\text{15-49}}) + u_x^\lambda + \eta_{R_x,s,a}^\lambda,$$

where $\beta_0^\lambda \sim \mathcal{N}(0,5)$ is an intercept term proportional to the average HIV transmission rate for untreated HIV positive adults, $\beta_S^{\lambda,s=\mathrm{M}} \sim \mathcal{N}(0,5)$ is the log incidence rate ratio for men compared to women,

$$\rho_x^{\text{15-49}} = \frac{\sum_{s\in\{\mathrm{F,M}\}}\sum_{a=15}^{45} N_{x,s,a}\cdot\rho_{x,s,a}}{\sum_{s\in\{\mathrm{F,M}\}}\sum_{a=15}^{45} N_{x,s,a}},$$

is the HIV prevalence among adults 15-49 calculated by aggregating age-specific HIV prevalences, and

$$\alpha_x^{\text{15-49}} = \frac{\sum_{s\in\{\mathrm{F,M}\}}\sum_{a=15}^{45} N_{x,s,a}\cdot\rho_{x,s,a}\cdot\alpha_{x,s,a}}{\sum_{s\in\{\mathrm{F,M}\}}\sum_{a=15}^{45} N_{x,s,a}\cdot\rho_{x,s,a}},$$

is the ART coverage among adults 15-49 calculated by aggregating age-specific ART coverages. The term ω is the average reduction in HIV transmission rate per 1% increase in population ART coverage and is fixed at $\omega = 0.7$ based on inputs to the Estimation and Projection Package (EPP) model (Jeffrey W. Eaton et al. 2019), $u_x^\lambda \sim \mathcal{N}(0,\sigma^\lambda)$ with $\sigma^\lambda \sim \mathcal{N}^+(0,1)$ are IID spatial random effects, and $\eta^\lambda_{R_x,s,a}$ specify fixed log incidence rate ratios by sex and age group calculated from Spectrum model output.

We model the proportion recently infected among HIV positive persons $\kappa_{x,s,a} \in [0,1]$ as

$$\kappa_{x,s,a} = 1 - \exp\left(-\lambda_{x,s,a}\cdot\frac{1-\rho_{x,s,a}}{\rho_{x,s,a}}\cdot(\Omega_T - \beta_T) - \beta_T\right),$$

where $\Omega_T \sim \mathcal{N}(\Omega_{T_0}, \sigma^{\Omega_T})$ is the mean duration of recent infection (MDRI), and $\beta_T \sim \mathcal{N}^+(\beta_{T_0}, \sigma^{\beta_T})$ is the false recent ratio (FRR). We use an informative prior on Ω_T based on the characteristics of the recent infection testing algorithm (RITA). For PHIA surveys this is $\Omega_{T_0} = 130$ days and $\sigma^{\Omega_T} = 6.12$ days. For PHIA surveys we assume there is no false recency, such that $\beta_{T_0} = 0.0$ and $\sigma^{\beta_T} = 0.0$.

S1.2.4 ANC testing

HIV prevalence $\rho_{x,a}^{\mathrm{ANC}}$ and ART coverage $\alpha_{x,a}^{\mathrm{ANC}}$ among pregnant women are modelled as being offset on the logit scale from the corresponding district-age indicators $\rho_{x,F,a}$ and $\alpha_{x,F,a}$ according to

$$\begin{aligned}
\mathrm{logit}(\rho_{x,a}^{\mathrm{ANC}}) &= \mathrm{logit}(\rho_{x,F,a}) + \beta^{\rho^{\mathrm{ANC}}} + u_x^{\rho^{\mathrm{ANC}}} + \eta_{R_x,a}^{\rho^{\mathrm{ANC}}},\\
\mathrm{logit}(\alpha_{x,a}^{\mathrm{ANC}}) &= \mathrm{logit}(\alpha_{x,F,a}) + \beta^{\alpha^{\mathrm{ANC}}} + u_x^{\alpha^{\mathrm{ANC}}} + \eta_{R_x,a}^{\alpha^{\mathrm{ANC}}},
\end{aligned}$$

where, for $\theta \in \{\rho,\alpha\}$, $\beta^{\theta^{\mathrm{ANC}}} \sim \mathcal{N}(0,5)$ are the average differences between population and ANC outcomes, $u_x^{\theta^{\mathrm{ANC}}} \sim \mathcal{N}(0,\sigma_X^{\theta^{\mathrm{ANC}}})$ are IID district random effects with $\sigma_X^{\theta^{\mathrm{ANC}}} \sim \mathcal{N}^+(0,1)$, and $\eta_{R_x,a}^{\theta^{\mathrm{ANC}}}$ for are offsets for the log fertility rate ratios for HIV positive women compared to HIV negative women and for women on ART to HIV positive women not on ART, calculated from Spectrum model outputs for region R_x.

For adult women 15-49 we model the number of ANC clients $\Psi_{x,a} > 0$ on the log scale

$$\log(\Psi_{x,a}) = \log(N_{x,\mathrm{F},a}) + \psi_{R_x,a} + \beta^\psi + u_x^\psi$$

where $N_{x,\mathrm{F},a}$ are the female population sizes, $\psi_{R_x,a}$ are fixed age-sex fertility ratios in Spectrum region R_x, β^ψ are fixed log rate ratios for the number of ANC clients relative to the predicted fertility, and $u_x^\psi \sim \mathcal{N}(0,\sigma^\psi)$ are district random effects, with $\sigma^\psi \sim \mathcal{N}^+(0,1)$.

S1.2.5 ART attendance

Let $\gamma_{x,x'} \in [0,1]$ be the probability that a person on ART residing in district x receives ART in district x'. We assume that $\gamma_{x,x'} = 0$ for $x \notin \{x, \mathrm{ne}(x)\}$ such that individuals seek treatment only in their residing district or its neighbours $\mathrm{ne}(x) = \{x' : x' \sim x\}$, where $\sim$ is an adjacency relation, and $\sum_{x'\in\{x,\mathrm{ne}(x)\}}\gamma_{x,x'} = 1$. To model $\gamma_{x,x'}$ for $x \sim x'$ we use a multinomial logistic regression model, based on the log-odds ratios

$$\tilde{\gamma}_{x,x'} = \log\left(\frac{\gamma_{x,x'}}{1-\gamma_{x,x'}}\right) = \tilde{\gamma}_0 + u_x^{\tilde{\gamma}}, \tag{1}$$

where $\tilde{\gamma}_0 = -4$ is a fixed intercept, and $u_x^{\tilde{\gamma}} \sim \mathcal{N}(0, \sigma_X^{\tilde{\gamma}})$ are district random effects with $\sigma_X^{\tilde{\gamma}} \sim \mathcal{N}^+(0, 2.5)$. Note that Equation 1 does not depend on x', such that $\gamma_{x,x'}$ is only a function of x. Choice of $\tilde{\gamma}_0 = -4$ implies a prior mean on $\gamma_{x,x'}$ of 1.8%, such that $(100 - 1.8 \times \text{ne}(x))\%$ of ART clients in district x obtain treatment in their home district, a-priori. We fix $\tilde{\gamma}_{x,x} = 0$ and recover the multinomial probabilities using the softmax

$$\gamma_{x,x'} = \frac{\exp(\tilde{\gamma}_{x,x'})}{\sum_{x^\star \in \{x, \text{ne}(x)\}} \exp(\tilde{\gamma}_{x,x^\star})}.$$

Given the total number of PLHIV on ART $A_{x,s,a} = N_{x,s,a} \cdot \rho_{x,s,a} \cdot \alpha_{x,s,a}$, the number of ART clients who reside in district x and obtain ART in district x' are $A_{x,x',s,a} = A_{x,s,a} \cdot \gamma_{x,x'}$, and the total attending ART facilities in district x' are

$$\tilde{A}_{x',s,a} = \sum_{x \in \{x', \text{ne}(x')\}} A_{x,x',s,a}.$$

S1.3 Likelihood specification

S1.3.1 Household survey data

For HIV prevalence, ART coverage and recent HIV infections, denoted by $\theta \in \{\rho, \alpha, \kappa\}$, the household survey ν furnishes weighted observations $\hat{\theta}_{\{x\},\{s\},\{a\}}$ with respective Kish effective sample sizes $m^{\hat{\theta}}_{\{x\},\{s\},\{a\}}$, and observed number of cases $y^{\hat{\theta}}_{\{x\},\{s\},\{a\}} = m^{\hat{\theta}}_{\{x\},\{s\},\{a\}} \cdot \hat{\theta}_{\{x\},\{s\},\{a\}}$. We use following three binomial working likelihoods

$$y^{\hat{\rho}}_{\{x\},\{s\},\{a\}} \sim \text{xBin}(m^{\hat{\rho}}_{\{x\},\{s\},\{a\}}, \rho_{\{x\},\{s\},\{a\}}), \quad \rho_{\{x\},\{s\},\{a\}} = \frac{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a} \cdot \rho_{x,s,a}}{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a}},$$

$$y^{\hat{\alpha}}_{\{x\},\{s\},\{a\}} \sim \text{xBin}(m^{\hat{\alpha}}_{\{x\},\{s\},\{a\}}, \alpha_{\{x\},\{s\},\{a\}}), \quad \alpha_{\{x\},\{s\},\{a\}} = \frac{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a} \cdot \rho_{x,s,a} \cdot \alpha_{x,s,a}}{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a} \cdot \rho_{x,s,a}}.$$

$$y^{\hat{\kappa}}_{\{x\},\{s\},\{a\}} \sim \text{xBin}(m^{\hat{\kappa}}_{\{x\},\{s\},\{a\}}, \kappa_{\{x\},\{s\},\{a\}}), \quad \kappa_{\{x\},\{s\},\{a\}} = \frac{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a} \cdot \rho_{x,s,a} \cdot \kappa_{x,s,a}}{\sum_{\{x\}} \sum_{\{s\}} \sum_{\{a\}} N_{x,s,a} \cdot \rho_{x,s,a}},$$

The generalised binomial $y \sim \text{xBin}(m, p)$ is defined for $y, m \in \mathbb{R}^+$ with $y \leq m$ such that

$$\log p(y) = \log \Gamma(m+1) - \log \Gamma(y+1) - \log \Gamma(m-y+1) + y \log p + (m-y) \log(1-p),$$

where the gamma function Γ is such that $\forall n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$.

S1.3.2 ANC testing data

We include ANC testing data for the year of the most recent survey. Let $W^{\text{ANC}}_{\{x\}}$ be the number of ANC clients, $X^{\text{ANC}}_{\{x\}}$ the number of those with ascertained status, $Y^{\text{ANC}}_{\{x\}}$ the number of those with positive status (either known or tested) and $Z^{\text{ANC}}_{\{x\}}$ the number of ANC clients already on ART prior to first ANC, such that $W^{\text{ANC}}_x \geq X^{\text{ANC}}_x \geq y^{\text{ANC}}_x \geq Z^{\text{ANC}}_x$ for all $x \in \{x\}$. When ANC testing data are only available for part of a given year, we denote $m^{\text{ANC}} \in \{1, \ldots, 12\}$ the number of months of reported data reflected in counts for that year. The observed number of HIV positive and already on ART among ANC clients is modelled by

$$Y^{\text{ANC}}_{\{x\}} \sim \text{Bin}\left(X^{\text{ANC}}_{\{x\}}, \rho^{\text{ANC}}_{\{x\},\{15,\ldots 45\}}\right),$$
$$Z^{\text{ANC}}_{\{x\}} \sim \text{Bin}\left(Y^{\text{ANC}}_{\{x\}}, \alpha^{\text{ANC}}_{\{x\},\{15,\ldots 45\}}\right),$$

where predicted prevalence and ART coverage are aggregated weighted by the predicted number of pregnant women by age $\Psi_{x,a}$

$$\rho^{\text{ANC}}_{\{x\}\{a\}} = \frac{\sum_{\{x\}} \sum_{\{a\}} \Psi_{x,a} \cdot \rho^{\text{ANC}}_{x,a}}{\sum_{\{x\}} \sum_{\{a\}} \Psi_{x,a}},$$

$$\alpha^{\text{ANC}}_{\{x\}\{a\}} = \frac{\sum_{\{x\}} \sum_{\{a\}} \Psi_{x,a} \cdot \rho^{\text{ANC}}_{x,a} \cdot \alpha^{\text{ANC}}_{x,a}}{\sum_{\{x\}} \sum_{\{a\}} \Psi_{x,a} \cdot \rho^{\text{ANC}}_{x,a}}.$$

S1.3.3 Number receiving ART

Let $\dot{A}_{\{x\},\{s\},\{a\}}$ be data for the number receiving ART

$$\dot{A}_{\{x\},\{s\},\{a\}} = \sum_{\{x\}}\sum_{\{s\}}\sum_{\{a\}}\sum_{x\sim x',x=x'} \dot{A}_{x',x,s,a}.$$

We model the unobserved numbers of ART clients travelling from x' to x as

$$\dot{A}_{x',x,s,a} \sim \text{Bin}(N_{x',s,a}, \pi_{x',x,s,a})$$

where $\pi_{x',x,s,a} = \rho_{x',s,a} \cdot \alpha_{x',s,a} \cdot \gamma_{x',x,s,a}$. This likelihood is approximated using a normal for the sum of binomials by

$$\dot{A}_{\{x\},\{s\},\{a\}} \sim \mathcal{N}(\tilde{A}_{\{x\},\{s\},\{a\}}, \sigma^{\tilde{A}}_{\{x\},\{s\},\{a\}})$$

where the mean is

$$\tilde{A}_{\{x\},\{s\}\{a\}} = \sum_{\{x\}}\sum_{\{s\}}\sum_{\{a\}}\sum_{x\sim x',x=x'} N_{x',s,a} \cdot \pi_{x',x,s,a},$$

and the variance is

$$\left(\sigma^{\tilde{A}}_{\{x\},\{s\}\{a\}}\right)^2 = \sum_{\{x\}}\sum_{\{s\}}\sum_{\{a\}}\sum_{x\sim x',x=x'} N_{x',s,a} \cdot \pi_{x',x,s,a} \cdot (1 - \pi_{x',x,s,a}).$$

S1.4 Identifiability constaints

If data are missing, some parameters are fixed to default values to help with identifiability. In particular:

1. If survey data on ART coverage by age and sex are not available then we set $u^{\alpha}_{a} = 0$ and $u^{\alpha}_{a,s=\text{M}} = 0$ and use the average age-sex pattern of ART coverage from the Spectrum offset $\eta^{\alpha}_{R_x,s,a}$.
2. If no ART data (survey or ART programme) are available but data on ART coverage among ANC clients are available, the level of ART coverage is not identifiable, but spatial variation is identifiable. In this instance, overall ART coverage is determined by the Spectrum offset, and only area random effects are estimated such that $\text{logit}\,(\alpha_{x,s,a}) = u^{\alpha}_{x} + \eta^{\alpha}_{R_x,s,a}$.
3. If survey data on recent HIV infection are not included in the model, then $\beta^{\lambda}_{0} = \beta^{\lambda,s=\text{M}}_{S} = u^{\lambda}_{x} = 0$. The sex ratio for HIV incidence is determined by the sex incidence rate ratio from Spectrum in the same years and the incidence rate in all districts is modelled assuming the same average HIV transmission rate for untreated adults, but varies according to district estimates of HIV prevalence and ART coverage.

For the Malawi example considered in the main text, all required data are available.

S2 C++ `TMB` user template

This section contains C++ for the negative log posterior of the simplified Naomi model we describe in Section S1. For ease of understanding, Table S2 provides correspondence between the mathematical notation used and the variable names used, for all hyperparameters and latent field parameters. For further reference on the `TMB` software see `https://kaskr.github.io/adcomp/_book`.

S3 MCMC convergence and suitability

We assessed MCMC convergence and suitability using a range of graphical and numerical tests. These included the potential scale reduction factor $\hat{R}$, bulk and tail effective sample size (ESS), autocorrelation decay plots, univariate traceplots, pairs density plots, and NUTS specific divergent transition and energy assessments.

For the time being, this analysis is available from `athowes.github.io/elgm-inf/mcmc-convergence`. Once the MCMC results are finalised, the analysis will be moved to this appendix and expanded upon. The following draft text and references may be useful in that expanded write-up.

- Improved $\hat{R}$ statistic from Vehtari et al. (2021). Recommended only to use the sample if the value is less than 1.05.
- ESS and autocorrelation related to efficiency
- Traceplots helpful for diagnosis of problems
- Pairs density plots helpful for understanding relationships between parameters, including possible identifiability issues
- Energy plot from Betancourt (2017)

References

Besag, Julian, Jeremy York, and Annie Mollié. 1991. "Bayesian image restoration, with two applications in spatial statistics." *Annals of the Institute of Statistical Mathematics* 43 (1): 1–20.

Betancourt, Michael. 2017. "A Conceptual Introduction to Hamiltonian Monte Carlo." *arXiv Preprint arXiv:1701.02434*.

Eaton, Jeffrey W, Tim Brown, Robert Puckett, Robert Glaubius, Kennedy Mutai, Le Bao, Joshua A Salomon, John Stover, Mary Mahy, and Timothy B Hallett. 2019. "The Estimation and Projection Package Age-Sex Model and the r-Hybrid Model: New Tools for Estimating HIV Incidence Trends in Sub-Saharan Africa." *AIDS (London, England)* 33 (Suppl 3): S235.

Eaton, Jeffrey W., Laura Dwyer-Lindgren, Steve Gutreuter, Megan O'Driscoll, Oliver Stevens, Sumali Bajaj, Rob Ashton, et al. 2021. "Naomi: A New Modelling Tool for Estimating HIV Epidemic Indicators at the District Level in Sub-Saharan Africa." *Journal of the International AIDS Society* 24 (S5): e25788. https://doi.org/https://doi.org/10.1002/jia2.25788.

Freni-Sterrantino, Anna, Massimo Ventrucci, and Håvard Rue. 2018. "A Note on Intrinsic Conditional Autoregressive Models for Disconnected Graphs." *Spatial and Spatio-Temporal Epidemiology* 26: 25–34.

Lumley, Thomas. 2004. "Analysis of Complex Survey Samples." *Journal of Statistical Software* 9: 1–19.

Simpson, Daniel, Håvard Rue, Andrea Riebler, Thiago G Martins, and Sigrunn H Sørbye. 2017. "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors." *Statistical Science* 32 (1): 1–28.

Stover, J, P Johnson, T Hallett, M Marston, R Becquet, and IM Timaeus. 2010. "The Spectrum Projection Package: Improvements in Estimating Incidence by Age and Sex, Mother-to-Child Transmission, HIV Progression in Children and Double Orphans." *Sexually Transmitted Infections* 86 (Suppl 2): ii16–21.

Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. "Rank-Normalization, Folding, and Localization: An Improved r for Assessing Convergence of MCMC (with Discussion)." *Bayesian Analysis* 16 (2): 667–718.

Variable name	Mathematical notation	Type	Dimension	Domain
`logit_phi_rho_x`	$\text{logit}(\phi^{\rho}_{X})$	Hyper	1	$\mathbb{R}$
`log_sigma_rho_x`	$\log(\sigma^{\rho}_{X})$	Hyper	1	$\mathbb{R}$
`logit_phi_rho_xs`	$\text{logit}(\phi^{\rho}_{XS})$	Hyper	1	$\mathbb{R}$
`log_sigma_rho_xs`	$\log(\sigma^{\rho}_{XS})$	Hyper	1	$\mathbb{R}$
`logit_phi_rho_a`	$\text{logit}(\phi^{\rho}_{A})$	Hyper	1	$\mathbb{R}$
`log_sigma_rho_a`	$\log(\sigma^{\rho}_{A})$	Hyper	1	$\mathbb{R}$
`logit_phi_rho_as`	$\text{logit}(\phi^{\rho}_{AS})$	Hyper	1	$\mathbb{R}$
`log_sigma_rho_as`	$\log(\sigma^{\rho}_{AS})$	Hyper	1	$\mathbb{R}$
`log_sigma_rho_xa`	$\log(\sigma^{\rho}_{XA})$	Hyper	1	$\mathbb{R}$
`logit_phi_alpha_x`	$\text{logit}(\phi^{\alpha}_{X})$	Hyper	1	$\mathbb{R}$
`log_sigma_alpha_x`	$\log(\sigma^{\alpha}_{X})$	Hyper	1	$\mathbb{R}$
`logit_phi_alpha_xs`	$\text{logit}(\phi^{\alpha}_{XS})$	Hyper	1	$\mathbb{R}$
`log_sigma_alpha_xs`	$\log(\sigma^{\alpha}_{XS})$	Hyper	1	$\mathbb{R}$
`logit_phi_alpha_a`	$\text{logit}(\phi^{\alpha}_{A})$	Hyper	1	$\mathbb{R}$
`log_sigma_alpha_a`	$\log(\sigma^{\alpha}_{A})$	Hyper	1	$\mathbb{R}$
`logit_phi_alpha_as`	$\text{logit}(\phi^{\alpha}_{AS})$	Hyper	1	$\mathbb{R}$
`log_sigma_alpha_as`	$\log(\sigma^{\alpha}_{AS})$	Hyper	1	$\mathbb{R}$
`log_sigma_alpha_xa`	$\log(\sigma^{\alpha}_{XA})$	Hyper	1	$\mathbb{R}$
`OmegaT_raw`	Ω_T	Hyper	1	$\mathbb{R}$
`log_betaT`	$\log(\beta_T)$	Hyper	1	$\mathbb{R}$
`log_sigma_lambda_x`	$\log(\sigma^{\lambda})$	Hyper	1	$\mathbb{R}$
`log_sigma_ancrho_x`	$\log(\sigma^{\rho^{\text{ANC}}}_{X})$	Hyper	1	$\mathbb{R}$
`log_sigma_ancalpha_x`	$\log(\sigma^{\alpha^{\text{ANC}}}_{X})$	Hyper	1	$\mathbb{R}$
`log_sigma_or_gamma`	$\log(\sigma^{\tilde{\gamma}}_{X})$	Hyper	1	$\mathbb{R}$
`beta_rho`	$(\beta^{\rho}_{0}, \beta^{\rho,s=\text{M}}_{s})$	Latent	2	$\mathbb{R}^2$
`beta_alpha`	$(\beta^{\alpha}_{0}, \beta^{\alpha,s=\text{M}}_{S})$	Latent	2	$\mathbb{R}^2$
`beta_lambda`	$(\beta^{\lambda}_{0}, \beta^{\lambda,s=\text{M}}_{S})$	Latent	2	$\mathbb{R}^2$
`beta_anc_rho`	$\beta^{\rho^{\text{ANC}}}$	Latent	1	$\mathbb{R}$
`beta_anc_alpha`	$\beta^{\alpha^{\text{ANC}}}$	Latent	1	$\mathbb{R}$
`u_rho_x`	u^{ρ}_{x}	Latent	n	$\mathbb{R}^n$
`us_rho_x`		Latent	n	$\mathbb{R}^n$
`u_rho_xs`	$u^{\rho,s=\text{M}}_{x}$	Latent	n	$\mathbb{R}^n$
`us_rho_xs`		Latent	n	$\mathbb{R}^n$
`u_rho_a`	u^{ρ}_{a}	Latent	10	$\mathbb{R}^{10}$
`u_rho_as`	$u^{\rho,s=\text{M}}_{a}$	Latent	10	$\mathbb{R}^{10}$
`u_rho_xa`	$u^{\rho,a<15}_{x}$	Latent	n	$\mathbb{R}^n$
`u_alpha_x`	u^{α}_{x}	Latent	n	$\mathbb{R}^n$
`us_alpha_x`		Latent	n	$\mathbb{R}^n$
`u_alpha_xs`	$u^{\alpha,s=\text{M}}_{x}$	Latent	n	$\mathbb{R}^n$
`us_alpha_xs`		Latent	n	$\mathbb{R}^n$
`u_alpha_a`	u^{α}_{a}	Latent	13	$\mathbb{R}^{13}$
`u_alpha_as`	$u^{\alpha,s=\text{M}}_{a}$	Latent	10	$\mathbb{R}^{10}$
`u_alpha_xa`	$u^{\alpha,a<15}_{x}$	Latent	n	$\mathbb{R}^n$
`ui_lambda_x`	u^{λ}_{x}	Latent	n	$\mathbb{R}^n$
`ui_anc_rho_x`	$u^{\rho^{\text{ANC}}}_{x}$	Latent	n	$\mathbb{R}^n$
`ui_anc_alpha_x`	$u^{\alpha^{\text{ANC}}}_{x}$	Latent	n	$\mathbb{R}^n$
`log_or_gamma`	$u^{\tilde{\gamma}}_{x}$	Latent	n	$\mathbb{R}^n$

Table S2: Correspondence between mathematical notation and variable names used in our `TMB` code. The total number of hyperparameters is 24, and the total number of latent field parameters is $51 + 14n$.