**Geostatistical proportional hazards model with spatially varying covariates**

A Bayesian geostatistical proportional hazards model [35] was fitted to assess the association between under-five mortality and health interventions, and to identify interventions which contribute most to mortality at the sub-national level. Let be the set of locations at which mortality data are observed, be the observed number of months lived or the censoring time for child at location ,  be the vector of socio-demographic and climatic factors and  be the coverage of intervention at location .We modeled the hazard of death by the equation, and assumed a Weibull baseline hazard i.e. where is the shape parameter, is the vector of regression coefficients with , correspond to the hazard ratio (HR).  is a cluster-specific random frailty which captures spatial correlation in mortality i.e. clusters in closer proximity are expected to have similar mortality hazard due to common exposures. We modeled  by a Gaussian process, i.e. , with an exponential correlation function of the distance between locations  and , that is [35]. The parameter gives the variance of the spatial process and is a smoothing parameter that controls the rate of correlation decay with distance. For the exponential correlation function, determines the distance at which the correlation drops to (i.e. effective range of spatial process ). Our model assumed that the relation between health interventions and mortality varied across regions by including intervention specific spatially varying coefficients, where  is the association between intervention and mortality on child mortality at global (national) level and are the varying effects at regional (sub-national) levels  with indicating the region corresponding to the location .We introduced spatial dependence among the regions via a conditional autoregressive (CAR) prior for , that is with is the variance of spatially varying disease associations, is a diagonal matrix with entries where is the number of neighbors of region ,  measures overall spatial dependence and is a proximity matrix with normalized entries that is , is 1 if region neighbors and 0 otherwise [35]. To complete Bayesian model formulation, we assumed inverse gamma priors for all spatial variances with known parameters, i.e. , a uniform prior distribution for , where  and chosen such as the effective range is within the maximum and minimum distances of the observed locations [36] and a uniform prior for where are the smallest and largest eigenvalue of  [35]. The shape parameter was assigned an exponential prior . Non-informative normal priors were adopted for the regression coefficients  for  and .

The joint posterior distribution of the model is given by

[.

Model parameters were estimated using Markov Chain Monte Carlo (MCMC) simulation [37]. We run a two chain algorithm for 250 000 iterations with an initial burn-in of 20,000 iterations. Convergence was assessed by visual inspection of trace and density plots and analytically by the Gelman and Rubin diagnostic [38].

**Bayesian variable selection**

A Bayesian stochastic search variable selection approach was used to determine the most important predictors of under-five mortality [39]. ITN coverage measures were highly correlated with more than 85%, therefore, only one (or none) ITN measure among those defining ownership and one (or none) ITN measure among those defining use [40,41] was selected. For each ITN coverage measure in the ownership group, a categorical indicator , was introduced to represent exclusion of the variable from the model (), or inclusion of one of the ITN ownership measure i.e. Prop\_ITNA (), Prop\_1ITN2 () and Prop\_1ITN (). A similar definition was adopted for the ITN use coverage measure i.e. exclusion of the variable from the model (), inclusion of Prop\_ITN5 (), Prop\_ITNS () and Prop\_ITNU (). The ITN measure with the highest probability of inclusion in each category was included in the final model. For the environmental/climatic variables except land cover, variable selection compared their linear and categorical forms and selected the one that had the highest probability of inclusion or neither of the two forms. The categorical forms were generated based on the quartiles of variables. We introduced an indicator for each environmental/climatic covariate which defines exclusion of the variable from the model (), inclusion in a categorical () or linear () form. For diseases, socio-demographic, land cover, health interventions other than ITN coverage measures, we introduced a binary indicator parameter suggesting presence () or absence () of the predictor from the model.

 has a probability mass function , where denotes the inclusion probabilities, , e.g. for ITN coverage measures so that and is the Dirac function, . We assumed a spike and slab prior for regression coefficient corresponding to the corresponding covariate, i.e. proposing a non-informative prior for in case is included in the model and an informative normal prior with variance close to zero (i.e. ) shrinking to zero if is excluded from the model. For inclusion probabilities of ITN use and ownership, a non-informative Dirichlet distribution was adopted with hyper parameters , that is, . A similar distribution was adopted for the inclusion probabilities of environmental/climatic factors. For diseases, socio-demographic, land cover, health interventions other than ITN coverage measures, a Bernoulli prior with an equal inclusion or exclusion probability was assumed for the indicator i.e. . Also, inverse Gamma priors with parameters (2.01, 1.01) were assumed for the precision hyper parameters . The predictors identified as important are those with posterior inclusion probability greater than or equal to 40% [40–42].