

# Appendices

## Appendix I.

**Model specification.**  $K_{space}$  was defined by a stationary Matérn covariance function over the Euclidean distance matrix  $D$  between all survey cluster locations, with spatial decay parameter  $\kappa$  and precision parameter  $\tau$  and where  $K_\nu$  is the modified Bessel function and  $\Gamma(\cdot)$  the Gamma function. This non-standard parameterisation of the Matérn covariance function corresponds to the stationary solution of the stochastic partial differential equation (SPDE) representation of the stationary Matérn family.<sup>1</sup> This facilitates the use of efficient statistical machinery for modelling with SPDEs, as described below. We set the complexity parameter  $\nu$  fixed at 1, implying a twice-differentiable process.  $K_{time}$  was defined by the covariance function corresponding to the discrete-time autoregressive stochastic process of order one (AR1). The AR1 process is typically defined over a nominal random variable:  $x_t = \rho x_{t-1} + N(0, \delta)$  where  $t$  indexes time (in our case the period) and  $\rho$  (which is constrained such that  $|\rho| < 1$ ) and  $\delta$  are parameters. On convolving the space and time correlation structures with this definition of the AR1 process, the spatial variance  $\frac{1}{\tau}$  and the temporal variance  $\delta^2$  would be non-identifiable. We therefore omit  $\delta$  from our definition above, and represent overall space-time variance via the parameter  $\frac{1}{\tau}$ .

**Model fitting.** Models were fitted by integrated nested Laplace approximations (INLA)<sup>2</sup> and a stochastic partial differential equation (SPDE) representation of the Gaussian-Markov random field (GMRF) approximation to the GP model<sup>3</sup>, using the INLA R package<sup>4</sup>. The INLA-SPDE approach makes use of the close correspondence between a GMRF defined on a sufficiently dense lattice and a GP, the efficient numerical routines enabled by representing GMRFs as SPDEs, and efficient inference over the parameters of these models using the INLA method. These approximations enable us to carry out full Bayesian inference over the model for a very large dataset, where other inference methods (such as MCMC) would be computationally prohibitive. Whilst the result is an imperfect approximation to the model posterior, this approach has been shown to have extremely high accuracy when compared with MCMC in both theoretical and real-world mapping problems<sup>56</sup>. We defined the GMRF on a lattice constructed by constrained, refined Delaunay triangulation within a convex hull no closer than ten decimal degrees (approximately 111km at the equator) from the borders of each target country. Over land this lattice was constrained to have edge length no greater than half of one decimal degree and over sea no greater than five decimal degrees. These fitted models were then used to generate 100 posterior samples each of mapped proportion with 0 years and mean years of education by random sampling from the numerical approximation to the joint posterior density of the model parameters. These estimates were combined to estimate the pixel-level (marginal) predictive mean, and also used to estimate the probability that a pixel was above or below a fixed outcome threshold.

## Appendix II.

1Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology* 2013; **7**: 39–55.

2Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations.

3Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Statist Soc B* 2011; **73**: 423–98.

4Rue H, Martino S, Lindgren F, Simpson D, Riebler A, Krainski ET. INLA: Functions Which Allow to Perform Full Bayesian Analysis of Latent Gaussian Models Using Integrated Nested Laplace Approximation. 2014.

5Schrödle B, Held L, Riebler A, Danuser J. Using integrated nested Laplace approximations for the evaluation of veterinary surveillance data from Switzerland: a case-study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2011; **60**: 261–79.

6Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations.

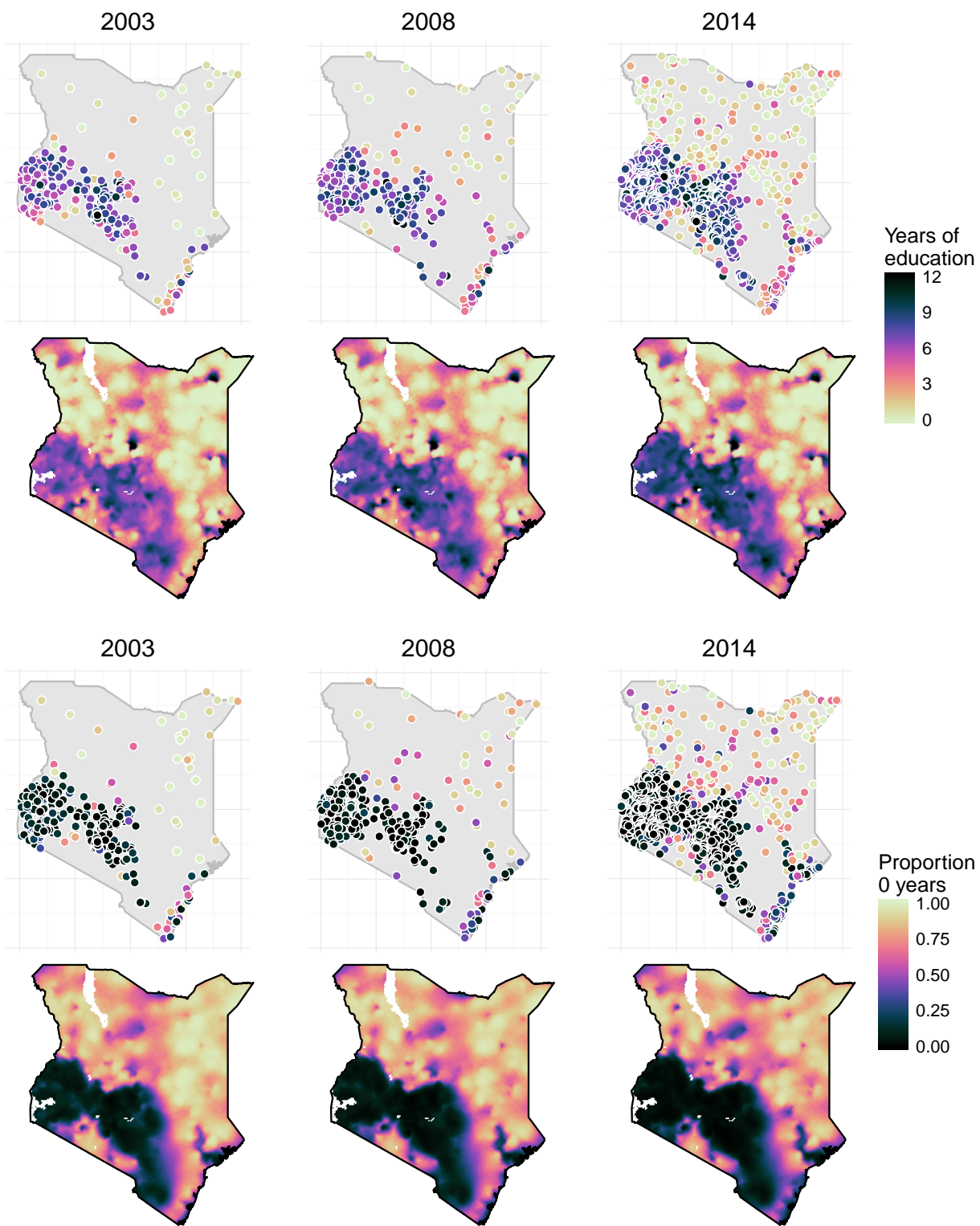


Figure 1: Data coverage and posterior mean for Kenya

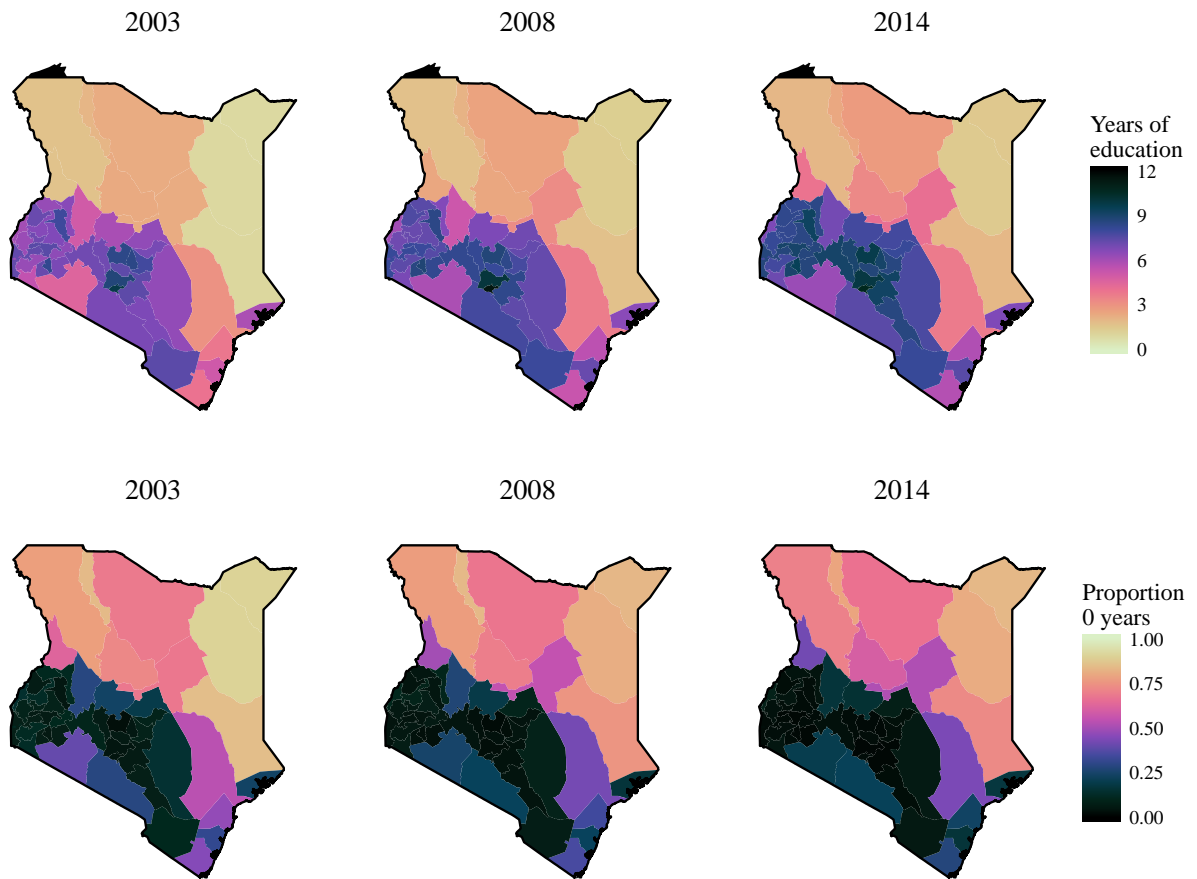
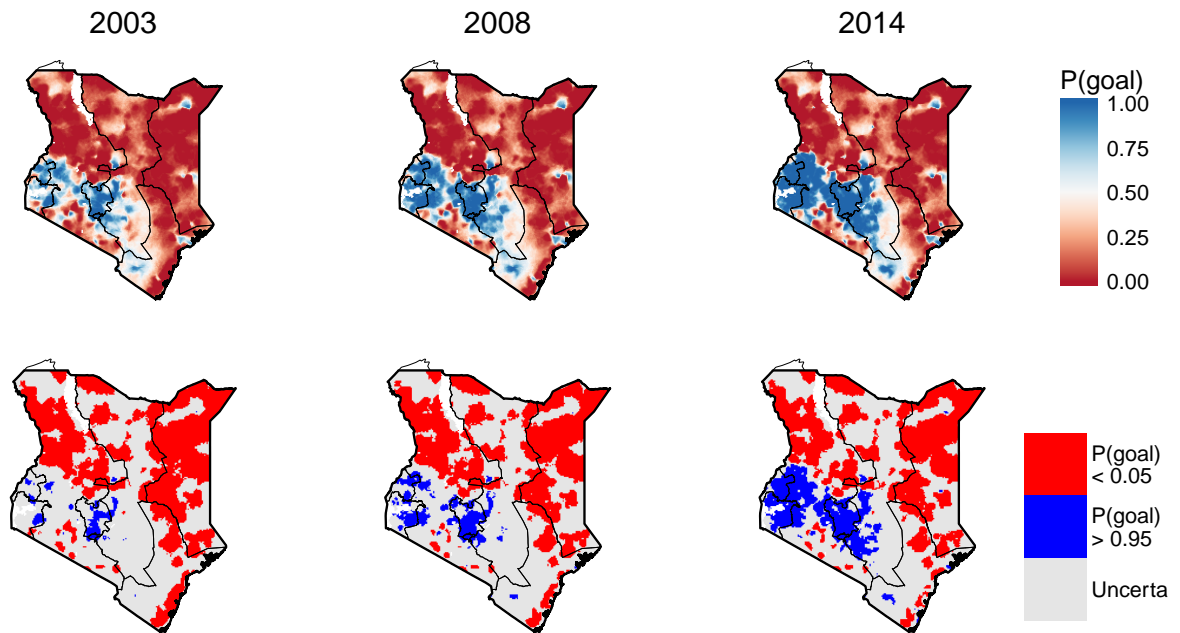


Figure 2: Population-weighted average by admin2 in Kenya

### GOAL: mean years of education > 6



### GOAL: < 20% of women with 0 years of education

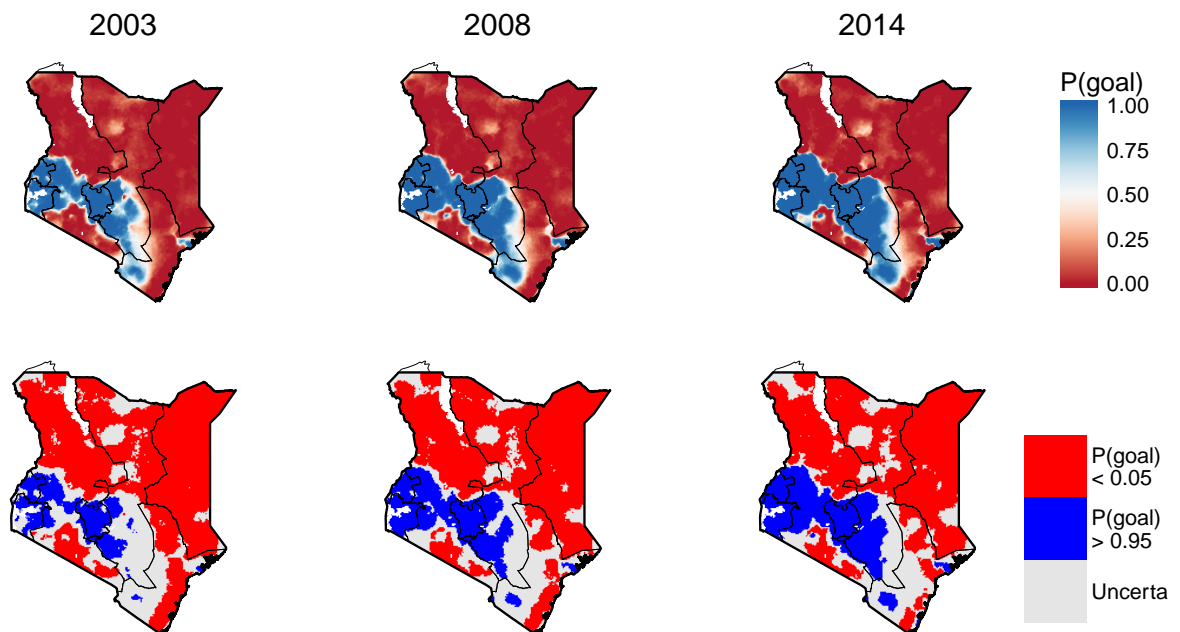


Figure 3: Probability threshold surface for Kenya relative to fixed goals. Top and bottoms rows show continuous probability and probability masked to  $\alpha < 0.05$ , respectively.

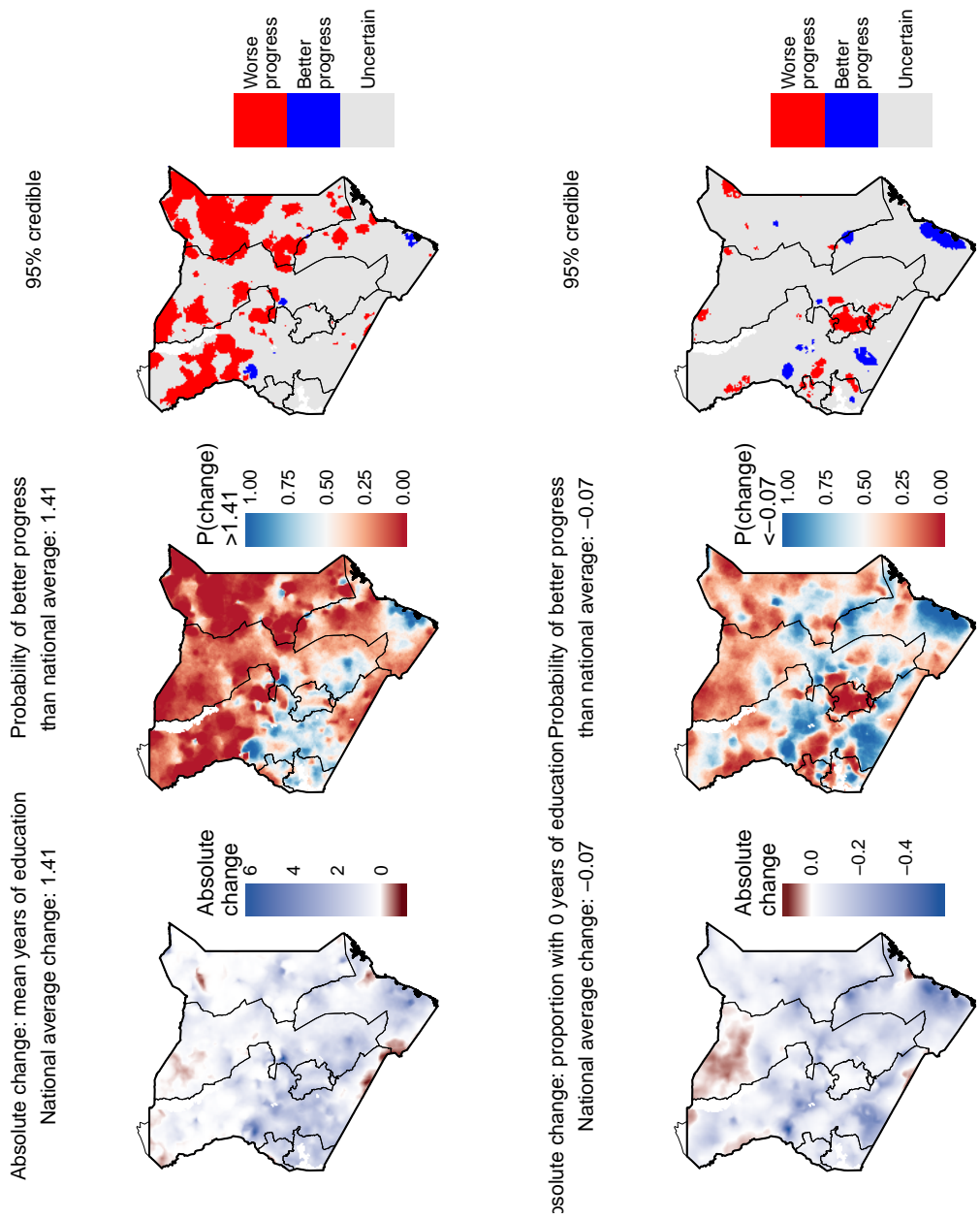


Figure 4: Probability threshold surface for Kenya relative to the absolute national population-weighted progress over the period. The top and bottom rows show progress in mean years of education and proportion of women with 0 years, respectively.

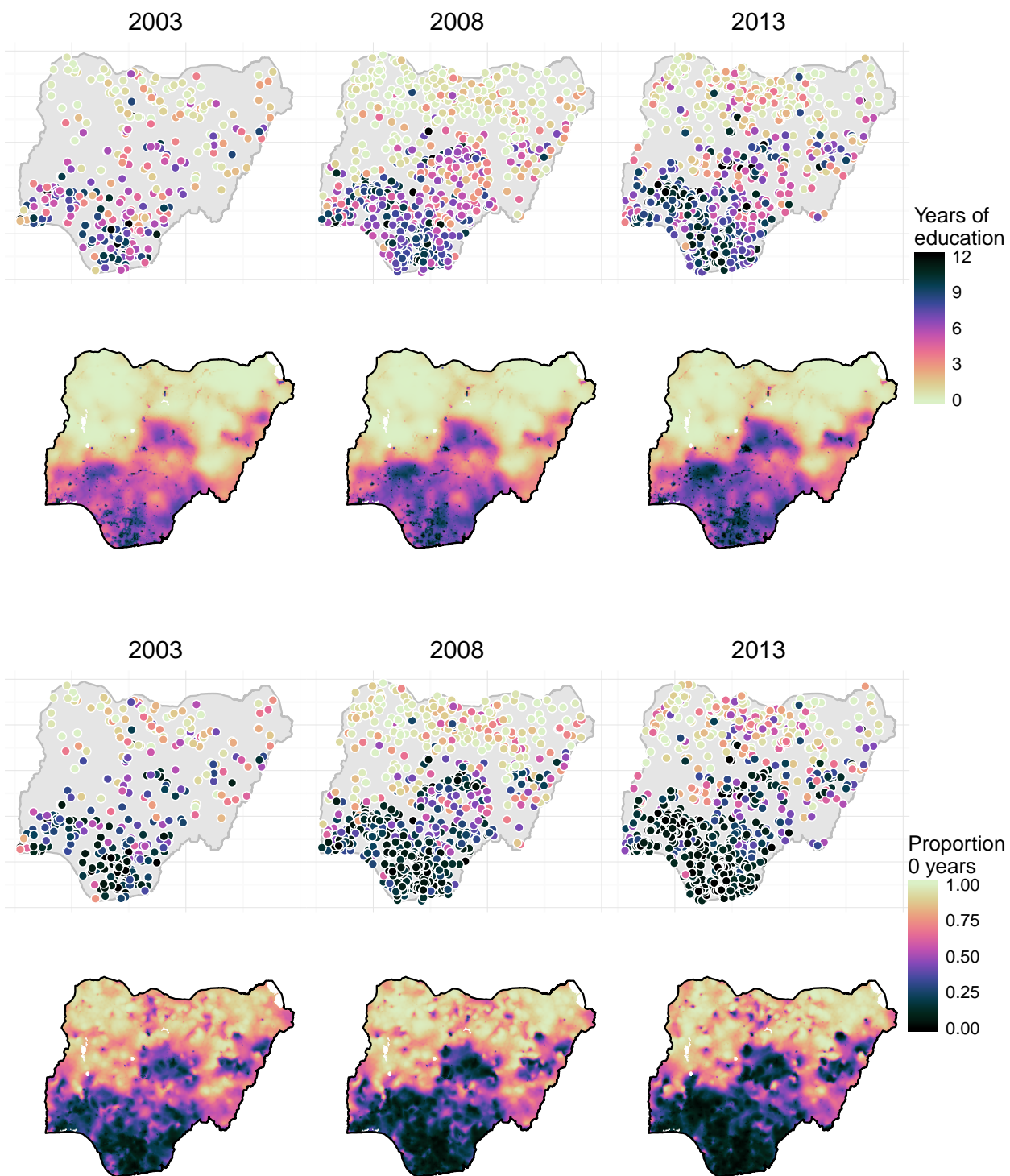


Figure 5: Data coverage and posterior mean for Nigeria

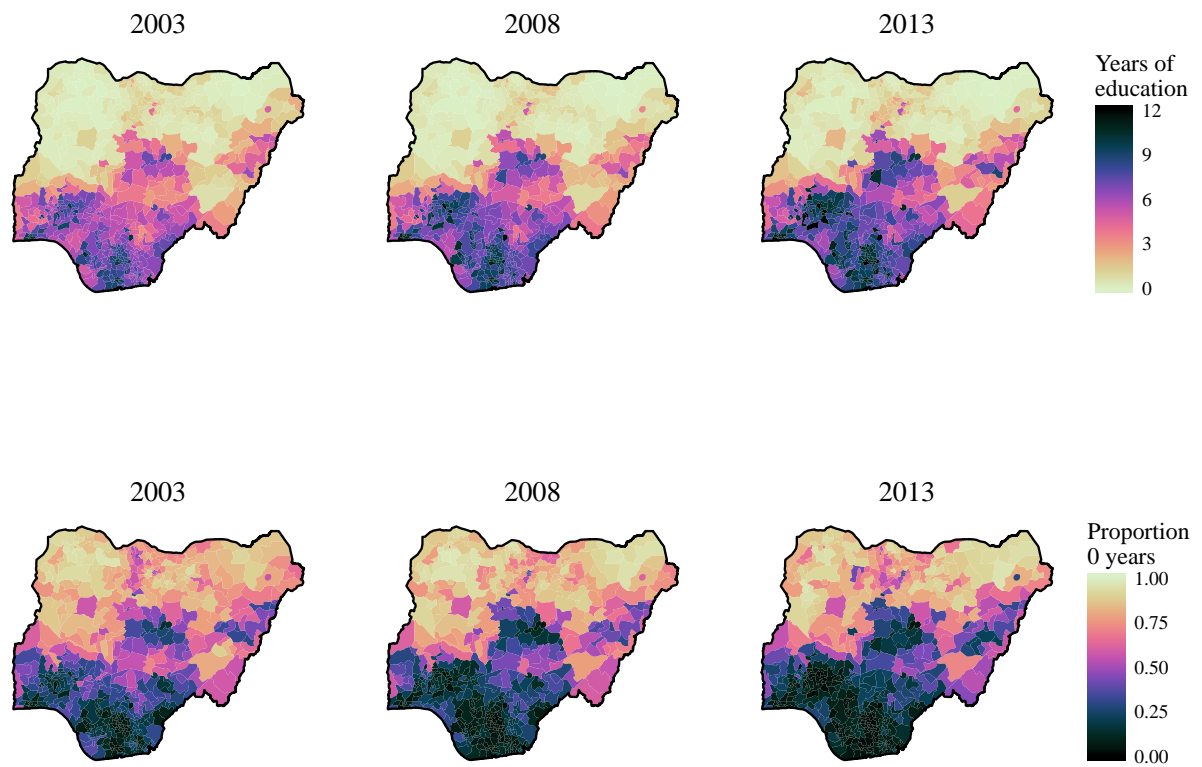
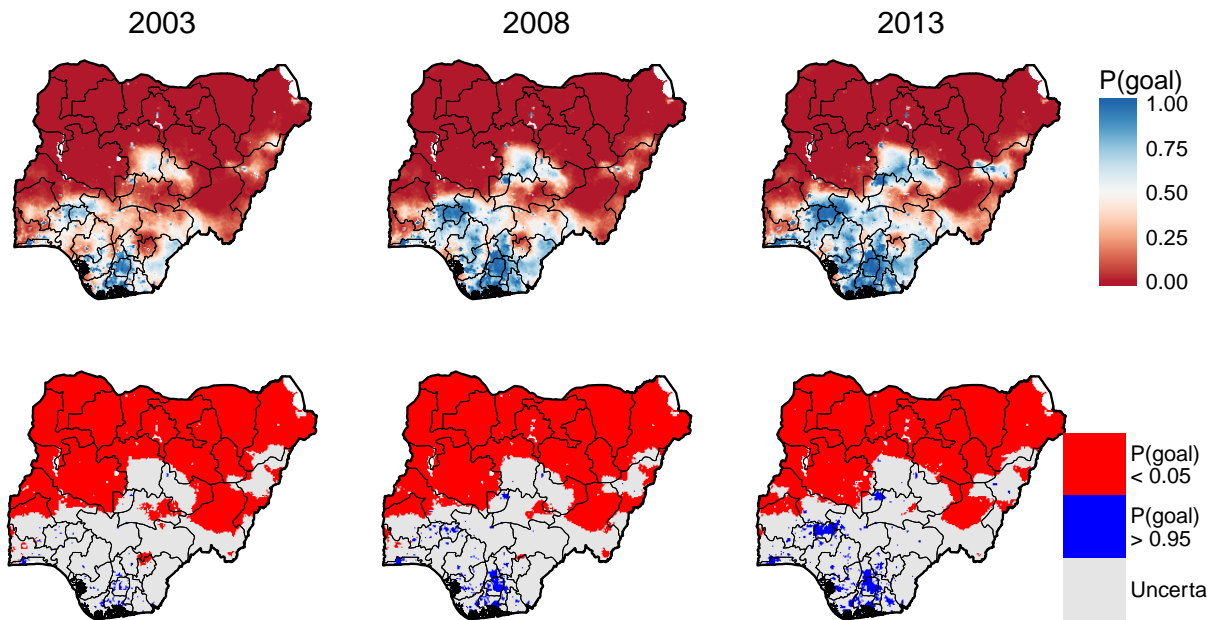


Figure 6: Population-weighted average by admin2 in Nigeria

### GOAL: mean years of education > 6



### GOAL: < 20% of women with 0 years of education

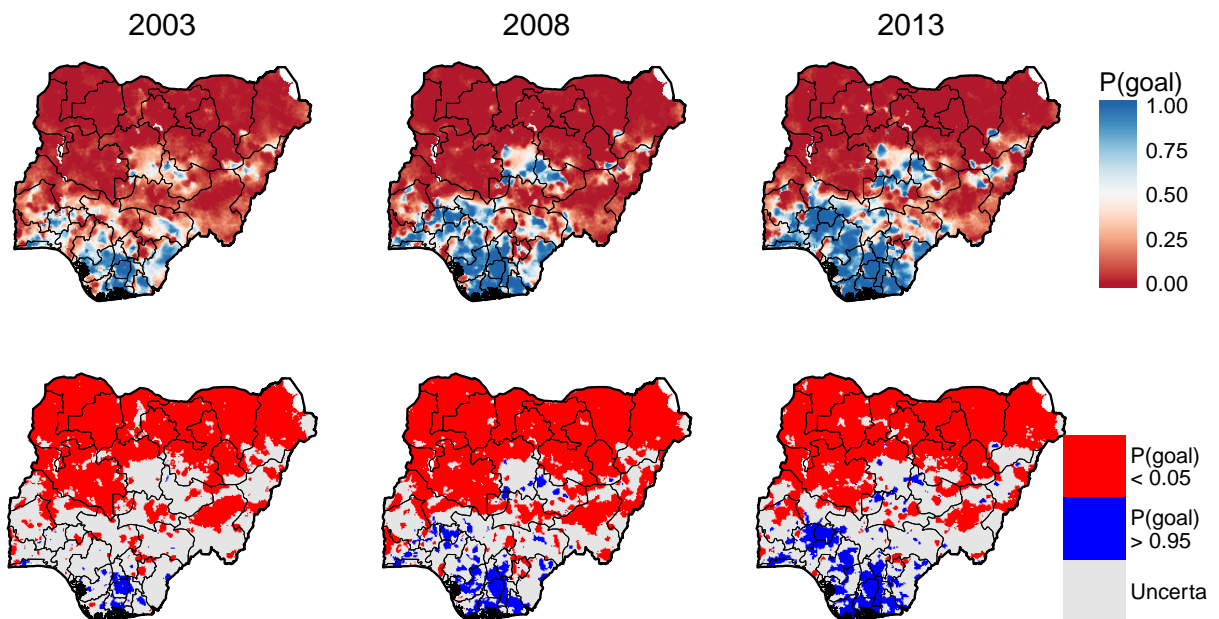


Figure 7: Probability threshold surface for Nigeria relative to fixed goals. Top and bottoms rows show continuous probability and probability masked to  $\alpha < 0.05$ , respectively.

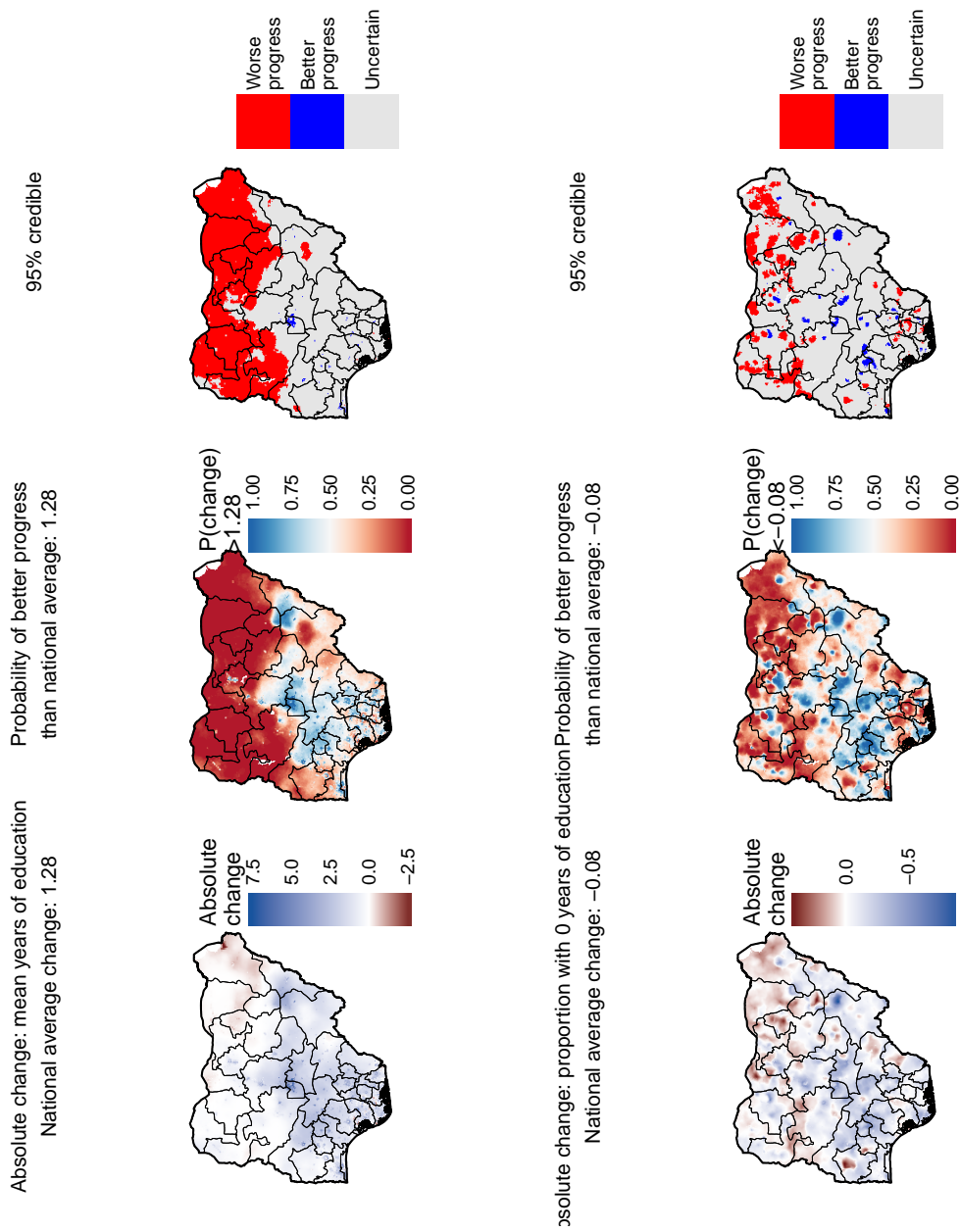


Figure 8: Probability threshold surface for Nigeria relative to the absolute national population-weighted progress over the period. The top and bottom rows show progress in mean years of education and proportion of women with 0 years, respectively.