

**SUPPLEMENTARY MATERIALS
FOR
ESTIMATION AND PREDICTION TOTAL FERTILITY RATES ACCOUNTING FOR
UNCERTAINTY OF DATA**

BY PEIRAN LIU^{*}, AND ADRIAN E. RAFTERY[†]

SM 1. Model Specification. Here we provide a full description of the Bayesian hierarchical model, which was summarized in the main text:

$$\begin{aligned}
 \text{Level 1: } & y_{c,t,s} | f_{c,t} \sim \mathcal{N}(f_{c,t} + \delta_{c,s}, \rho_{c,s}^2), \\
 & \mathbb{E}[\delta_{c,s}] = \mathbf{x}_{c,s} \boldsymbol{\beta}, \\
 & \mathbb{E}[\rho_{c,s}] = \mathbf{x}_{c,s} \boldsymbol{\gamma}, \\
 & f_{c,t} = \frac{1}{5} [(t_{l+t} - t) f_{c,t_l} + (t - t_l) f_{c,t_{l+5}}] \text{ for } t \in [t_l, t_{l+5}]; \\
 \text{Level 2: Phase I: } & f_{c,t} = f_{c,t-5} + \varepsilon_{c,t}, \\
 \text{Phase II: } & f_{c,t} = f_{c,t-5} - g(f_{c,t-5} | \boldsymbol{\theta}_c) + \varepsilon_{c,t}, \\
 \text{Phase III: } & f_{c,t} = \mu_c + \rho_c (f_{c,t-5} - \mu_c) + \varepsilon_{c,t}, \\
 & \boldsymbol{\theta}_c = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, d_c) \\
 & \varepsilon_{c,t} \sim \mathcal{N}(0, \sigma_{c,t}^2), \\
 & g(f_{c,t} | \boldsymbol{\theta}_c) = - \frac{d_c}{1 + \exp\left(-\frac{2 \ln(9)}{\Delta_{c1}} (f_{c,t} - \sum_i \Delta_{ci} + 0.5 \Delta_{c1})\right)} \\
 & \quad + \frac{d_c}{1 + \exp\left(-\frac{2 \ln(9)}{\Delta_{c3}} (f_{c,t} - \Delta_{c4} - 0.5 \Delta_{c3})\right)}
 \end{aligned}$$

The country-specific variance, $\sigma_{c,t}$, varies according to the phase and the current fertility level, as follows:

$$\sigma_{c,t} = \begin{cases} s_\tau & \text{If } t \text{ is in Phase I.} \\ c_{1975}(t) \left(\sigma_0 + (f_{c,t} - S)(-aI_{f_{c,t} > S} + bI_{f_{c,t} < S}) \right) & \text{If } t \text{ is in Phase II.} \\ \sigma_\epsilon & \text{If } t \text{ is in Phase III.} \end{cases}$$

$$c_{1975}(t) = \begin{cases} c & t \leq 1975 \\ 1 & t > 1975 \end{cases}$$

^{*}Peiran Liu is Ph.D. Student, Department of Statistics, University of Washington, Email: prliu@uw.edu

[†]Adrian E. Raftery is a professor of Statistics and Sociology in University of Washington, Seattle, Email: raftery@u.washington.edu

The country-level parameters, $\{U_c, \rho_c, \mu_c, \gamma_{ci}, \Delta_{c4}, d_c\}$, are specified as follows:

$$\begin{aligned} \text{Level 3: } U_c & \begin{cases} = f_{c,\tau} & \tau_c \leq 1950 \\ \sim U(\min\{5.5, \max_t\{f_{c,t}\}\}, 8.8) & \tau_c < 1950 \end{cases} \\ \phi_c & = \log\left(\frac{d_c - 0.25}{2.5 - d_c}\right), \\ \phi_c & \sim \mathcal{N}(\chi, \psi^2), \\ \Delta'_{c4} & = \log\left(\frac{\Delta_{c4} - 1}{2.5 - \Delta_{c4}}\right), \\ \Delta'_{c4} & \sim \mathcal{N}(\Delta_4, \delta_4^2), \\ p_{ci} & = \frac{\Delta_{ci}}{U_c - \Delta_{c4}} \text{ for } i = 1, 2, 3, \\ p_{ci} & = \frac{\exp(\gamma_{ci})}{\sum_j \exp(\gamma_{cj})}, \\ \gamma_{ci} & \sim \mathcal{N}(\alpha_i, \delta_i^2), \\ \mu_c & \sim \mathcal{N}(\bar{\mu}, \sigma_\mu^2), \\ \rho_c & \sim \mathcal{N}(\bar{\rho}, \sigma_\rho^2); \end{aligned}$$

where τ_c is the starting year of phase II for country c .

The hyperparameters are $\{s_\tau, \sigma_0, a, b, S, c, \sigma_\epsilon, \chi, \psi, \Delta_4, \delta_4, \alpha, \delta, \bar{\mu}, \sigma_\mu, \bar{\rho}, \sigma_\rho\}$. Some of these refer to Level 2 and some to Level 3. The prior distribution of these hyperparameters is as follows:

$$\begin{aligned} \text{Level 4: } 1/s_\tau^2 & \sim \text{Gamma}(1, 0.4^2), \\ \sigma_0 & \sim U[0.01, 0.6], \\ a & \sim U[0, 0.2], \\ b & \sim U[0, 0.2], \\ S & \sim U[3.5, 6.5], \\ c & \sim U[0.8, 2], \\ \sigma_\epsilon & \sim U[0, 0.5], \\ \chi & \sim \mathcal{N}(-1.5, 0.6^2), \\ 1/\psi^2 & \sim \text{Gamma}(1, 0.6^2), \\ \Delta_4 & \sim \mathcal{N}(0.3, 1), \\ 1/\delta_i^2 & \sim \text{Gamma}(1, 1) \text{ for } i = 1, 2, 3, 4, \\ \alpha_1 & \sim \mathcal{N}(-1, 1), \\ \alpha_2 & \sim \mathcal{N}(0.5, 1), \\ \alpha_3 & \sim \mathcal{N}(1.5, 1), \\ \bar{\mu} & \sim U[0, 2.1], \\ \sigma_\mu & \sim U[0, 0.318], \\ \bar{\rho} & \sim U[0, 1], \\ \sigma_\rho & \sim U[0, 0.289] \end{aligned}$$

SM 2. MCMC Diagnostics. The model was estimated using Markov chain Monte Carlo, as described in the main text. We used 3 chains of 12000 iterations each, thinned by 10 with burn-in of 2000. We diagnosed convergence by visual inspection of trace plots, and we also used more formal convergence diagnostics.

The convergence statistics of [Gelman and Rubin \(1992\)](#) and [Brooks and Gelman \(1998\)](#) for the total fertility rates of Nigeria are shown in Table SM 1 as an example. Conventionally, these statistics indicate adequate convergence if they are close to 1, with 1.1 conventionally viewed as a threshold for acceptability. This is the case here, and indeed they are very close to 1, indicating satisfactory convergence by this criterion. This was true for all the model parameters and all the unobserved TFR values of all countries except for a very few number of countries including Djibouti, Mayotte, Saudi Arabia and South Sudan.

TABLE SM 1

Gelman-Rubin and Brooks-Gelman Diagnostics for country-specific parameters for Nigeria. The point estimate of the multivariate potential scale reduction factor is 1.01926.

	Point estimates	Upper Confidence Interval
$f_{Nigeria,1953}$	1.008	1.020
$f_{Nigeria,1958}$	1.008	1.026
$f_{Nigeria,1963}$	1.008	1.027
$f_{Nigeria,1968}$	1.011	1.041
$f_{Nigeria,1973}$	1.003	1.008
$f_{Nigeria,1978}$	1.007	1.023
$f_{Nigeria,1983}$	1.009	1.032
$f_{Nigeria,1988}$	1.005	1.018
$f_{Nigeria,1993}$	1.007	1.024
$f_{Nigeria,1998}$	1.003	1.012
$f_{Nigeria,2003}$	1.000	1.000
$f_{Nigeria,2008}$	1.002	1.006
$f_{Nigeria,2013}$	1.002	1.006

The diagnostic of [Raftery and Lewis \(1996\)](#) is designed to indicate whether the MCMC chain has been run for enough iterations to estimate posterior quantiles of interest with enough precision. We computed this diagnostic for all parameters and all unobserved TFR values for the 0.025 and 0.975 quantiles, requiring that the estimated quantile fall within 0.0125 of the true quantile in each case with probability at least 0.95. If the chain is completely independent, this yields a required number of iterations equal to 600. The largest number of iterations required was less than 5,000 for all the roughly 3,600 quantities being estimated, again indicating that the chain had been run long enough. The results for Nigerian TFR are again shown in Table SM 2 as an example, for the 0.025 quantile. In all cases, the number of iterations needed was well below the 10,000 that were actually used, indicating that the number of iterations used was adequate.

TABLE SM 2
Raftery-Lewis Diagnostics for TFR Values of Nigeria

	Burn-in	Total Iteration	Dependence Factor
$f_{Nigeria,1953}$	21	3566	5.94
$f_{Nigeria,1958}$	23	3952	6.59
$f_{Nigeria,1963}$	22	3751	6.25
$f_{Nigeria,1968}$	13	2321	3.87
$f_{Nigeria,1973}$	14	2491	4.15
$f_{Nigeria,1978}$	20	3423	5.70
$f_{Nigeria,1983}$	18	3213	5.36
$f_{Nigeria,1988}$	16	2693	4.49
$f_{Nigeria,1993}$	14	2396	3.99
$f_{Nigeria,1998}$	13	2321	3.87
$f_{Nigeria,2003}$	18	3052	5.09
$f_{Nigeria,2008}$	15	2662	4.44
$f_{Nigeria,2013}$	18	3178	5.30

SM 3. Out of sample validation Details. We are also interested in comparing the current model directly with the model proposed in (Alkema et al., 2011). In the out of sample validation settings of (Alkema et al., 2011), we should use WPP 2008 all the time, train the model with data before 1995, and make forecast for next three periods (1995 - 2000, 2000 - 2005, 2005 - 2010). All other settings are the same in Section 3 of the main paper. We summarize the out of sample validation performance with similar criterion.

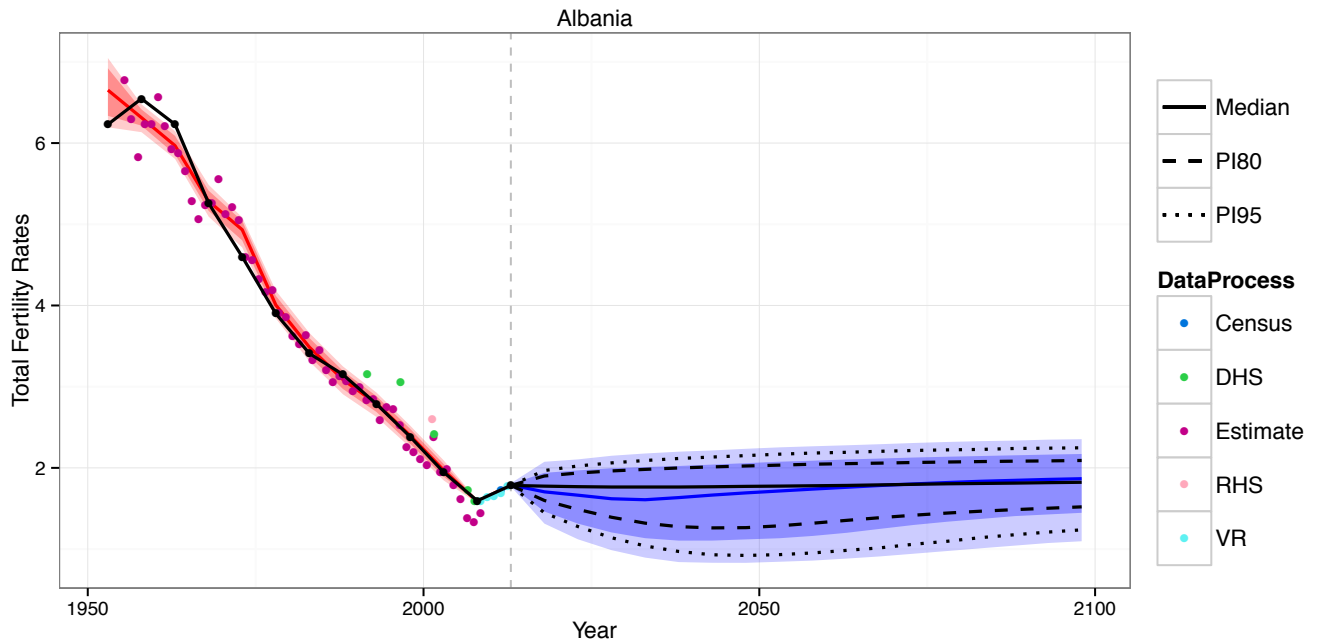
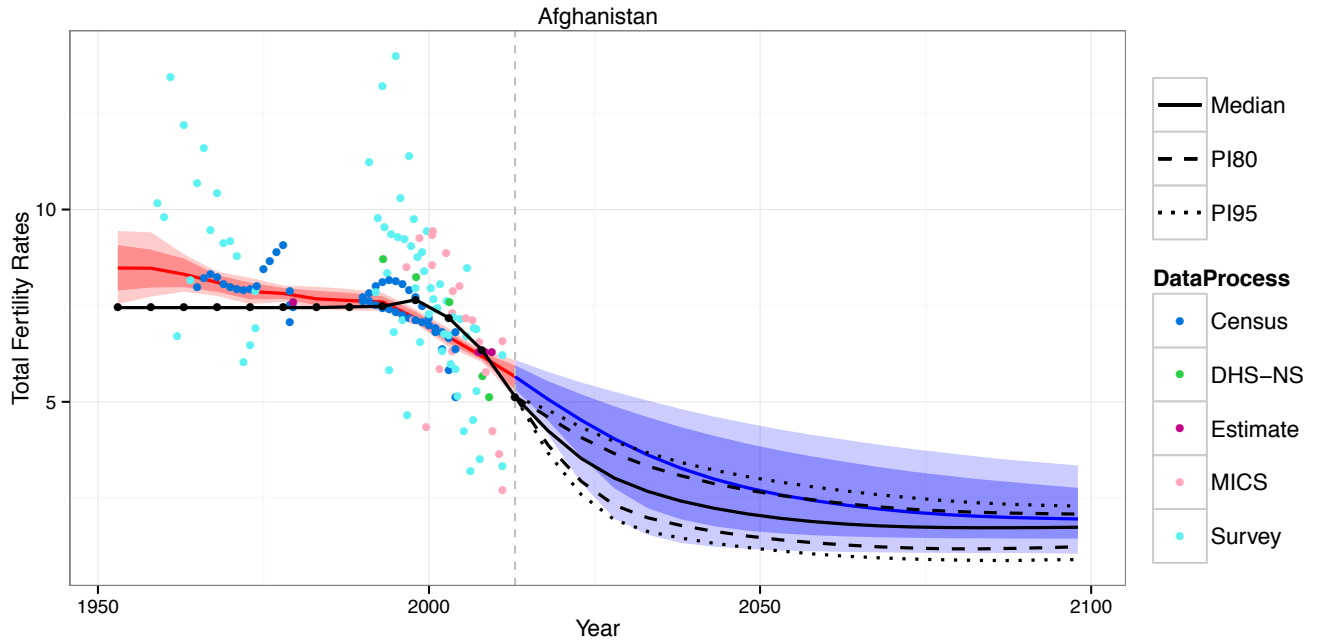
TABLE SM 3
proportion of left-out UN estimates that fall above or below their 80% and 95% projection intervals splitted by periods.

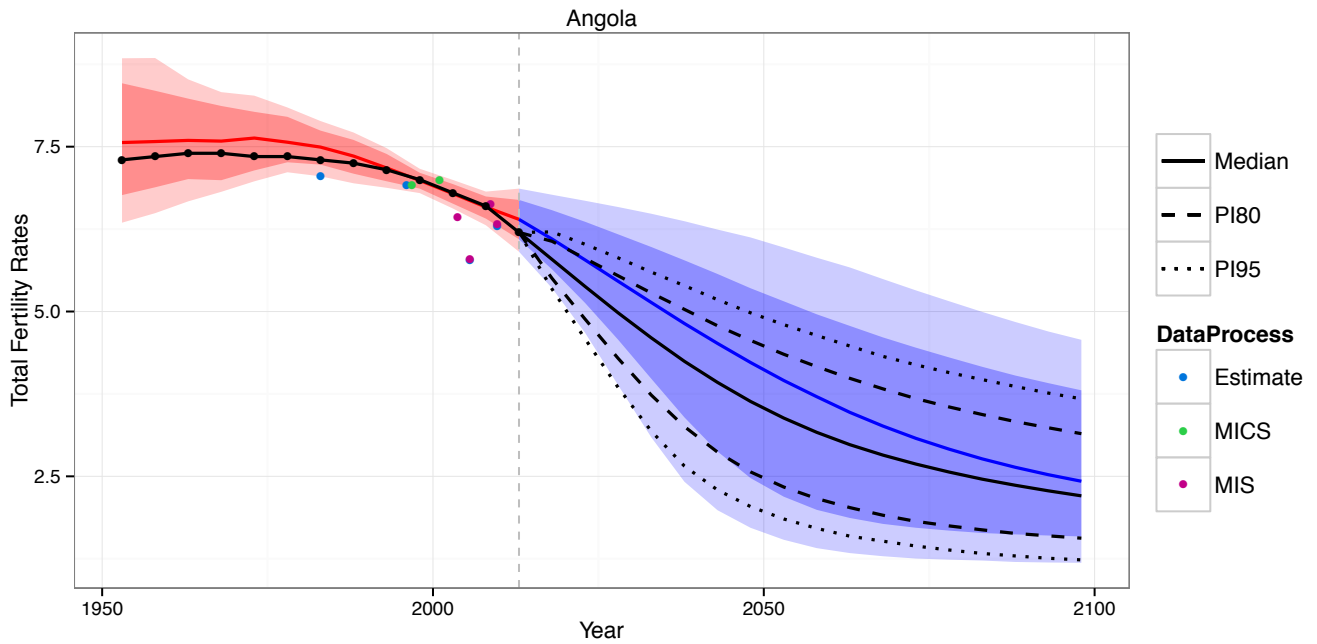
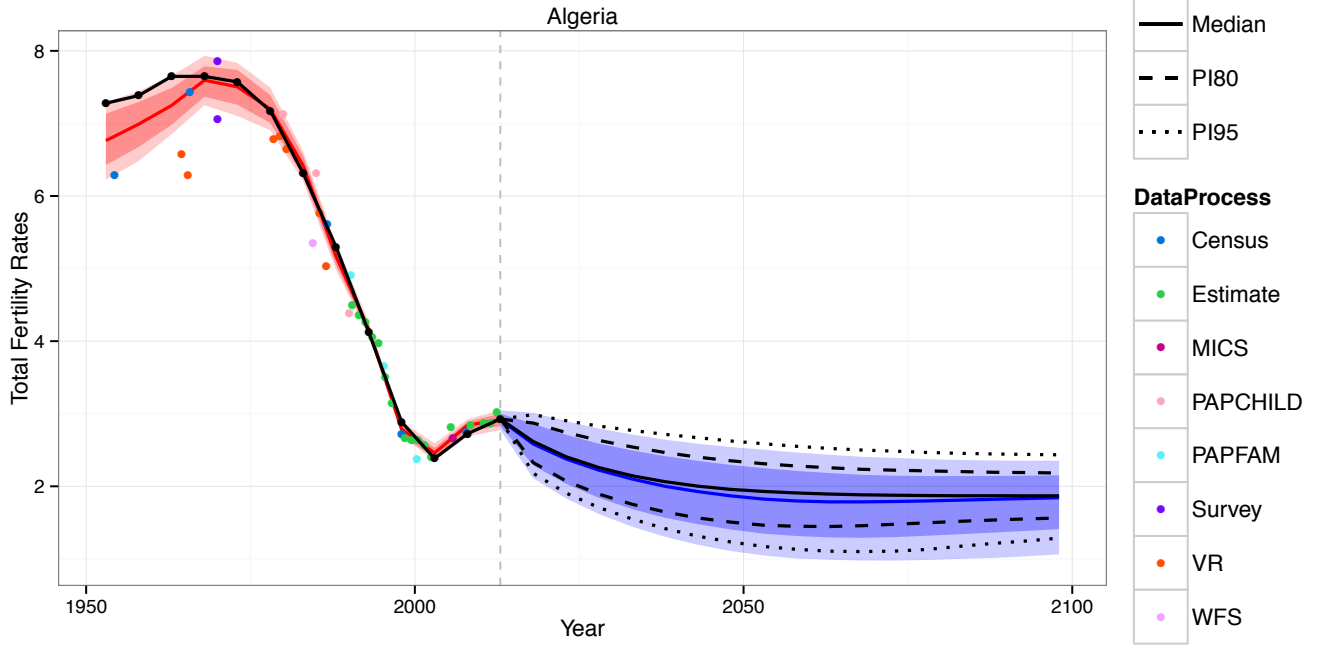
	95% PI			80% PI		
	Below	Above	Coverage	Below	Above	Coverage
1995-2000	3.0%	1.5%	94.5%	13.9%	4.5%	81.6%
2000-2005	3.5%	2.5%	94.0%	12.9%	4.0%	83.1%
2005-2010	1.0%	3.0%	96.0%	7.5%	7.5%	85.0%

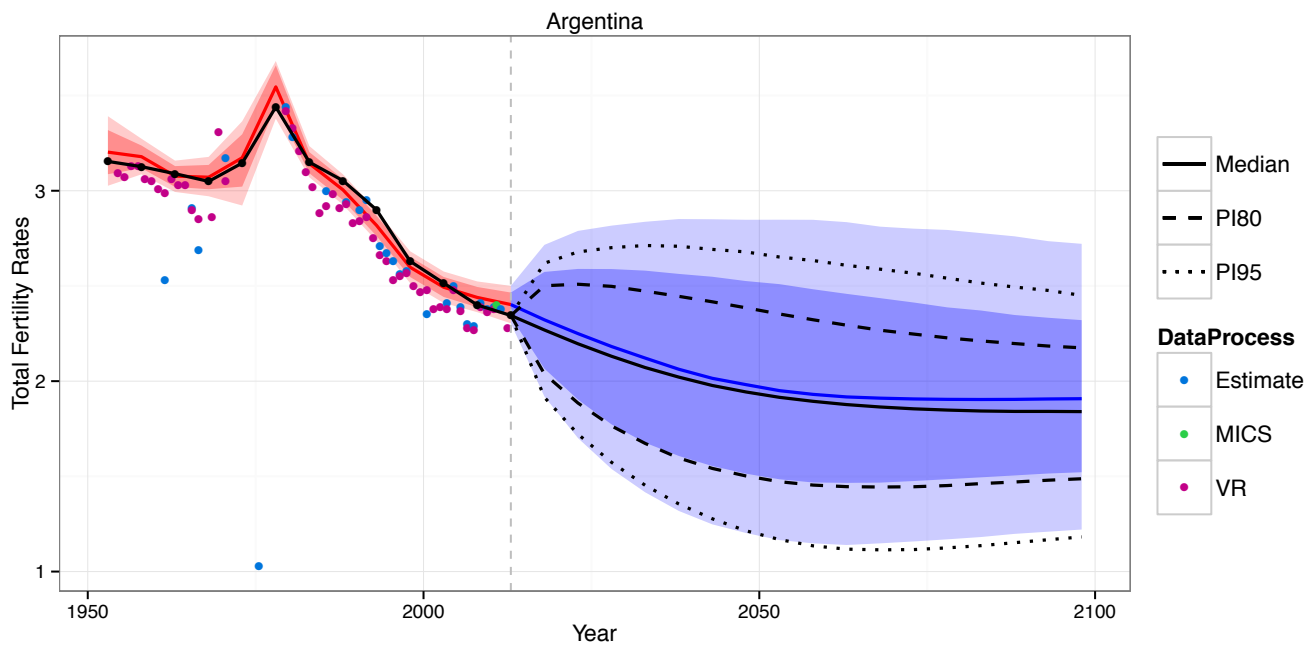
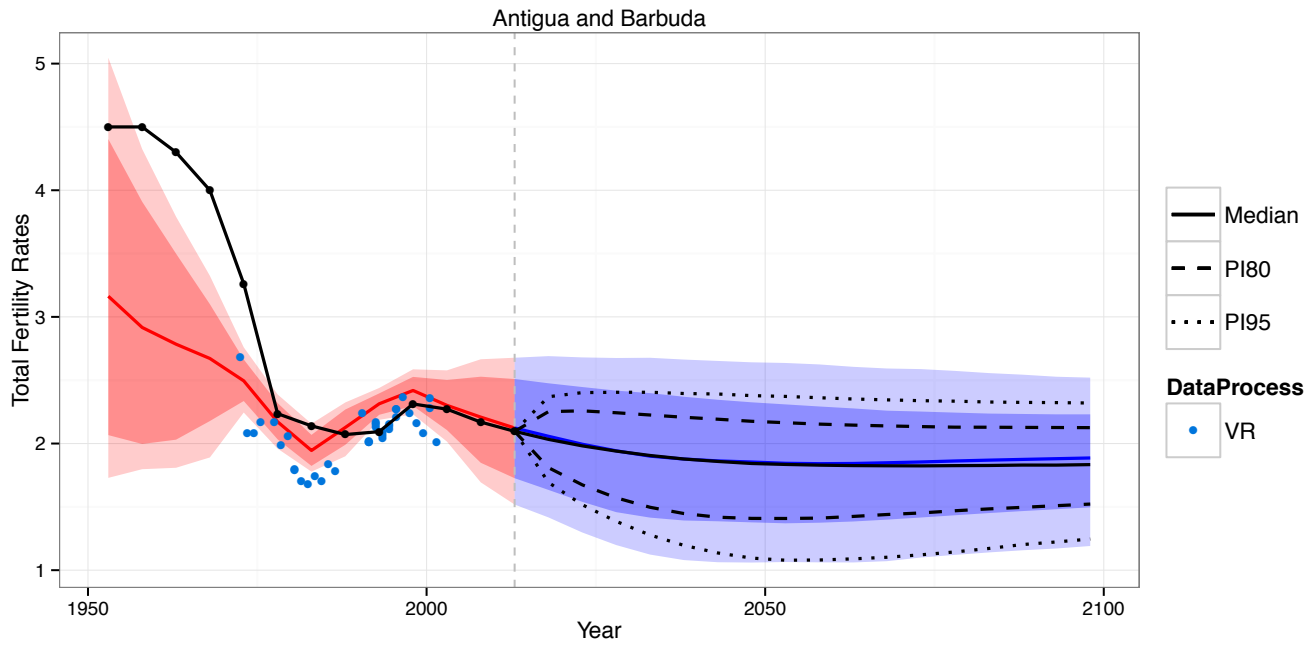
The overall coverage rates are 95.3% and 83.2% for 95% and 80% predictive intervals respectively. The coverage rate is very stable, and the 95% predictive intervals have achieved very precise coverage, which is similar to those out of sample validation in the main paper. The 80% out of sample interval is slightly wider than the nominal value. Since we just have one realization of all 201 countries and regions for out of sample validation, the coverage rate is satisfactory. We tried to compare this result in (Alkema et al., 2011), which is 91.5% and 77% for 95% and 80% predictive intervals respectively. Moreover, since in (Alkema et al., 2011), the phase III was trained without using Bayesian models, and the long term mean of phase III for all countries was set to be 2.1, which has been changed in the current model in BAYESTFR, we tried to reproduce the same out of sample validation study with hierarchical model and different long term mean. We could figure out that the coverage rate is 92% and 77.5% for 95% and 80% predictive intervals respectively.

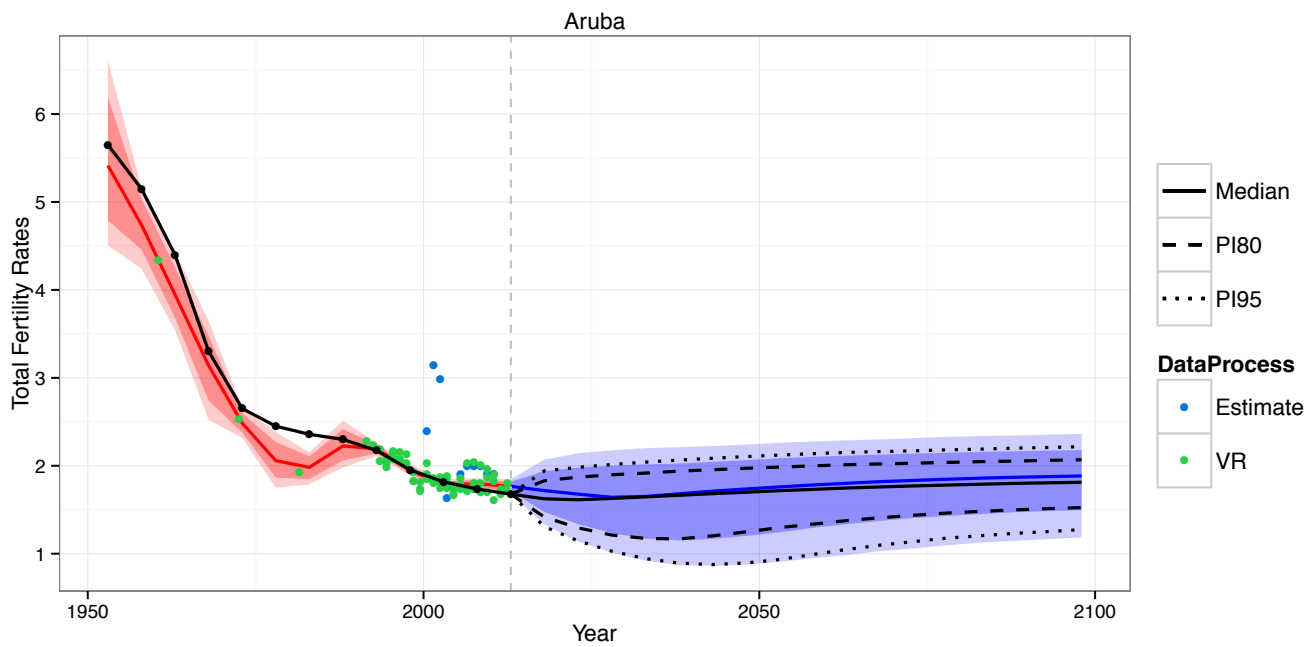
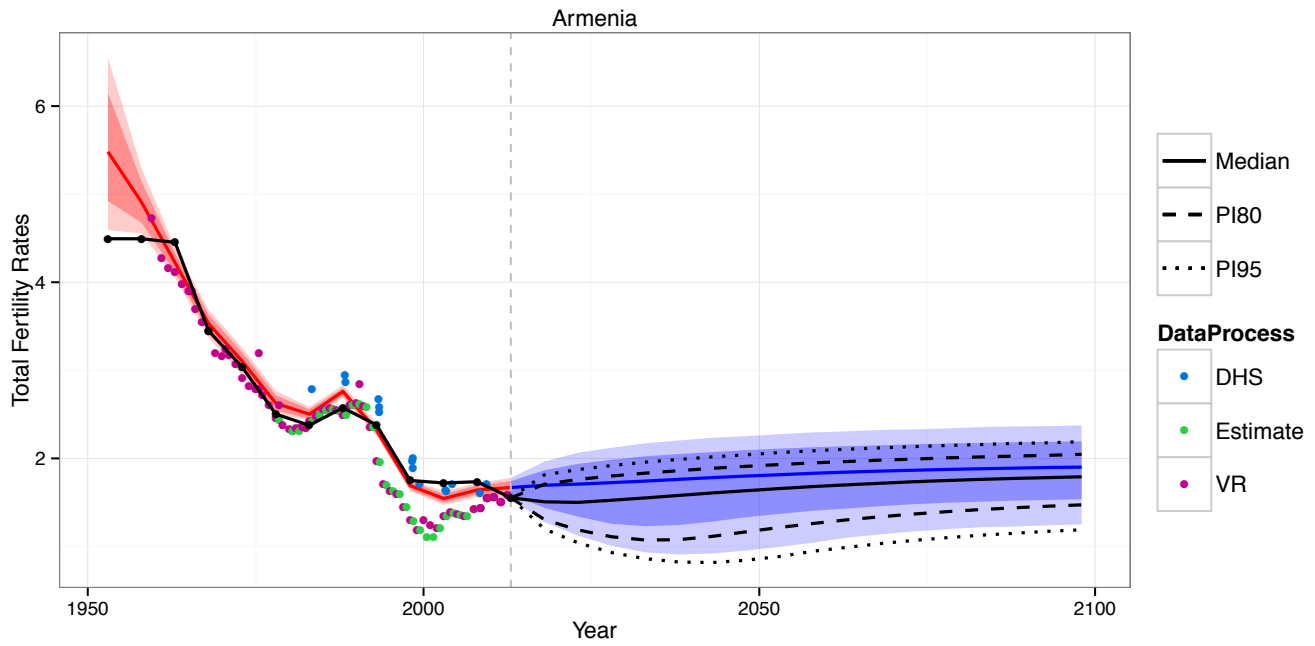
Thus, we could figure out that after we change the model into the setting of those in paper (Alkema et al., 2011), we could find that the overall coverage rate is slightly closer to the nominal rate. Thus, we could argue that our assumption that with accounting for uncertainty of past TFR estimation, the uncertainty in predictive interval is wider as we expected. If we are considering a single version of WPP package, our model could generate predictive intervals that are closer to the correct value, but the improvement is small.

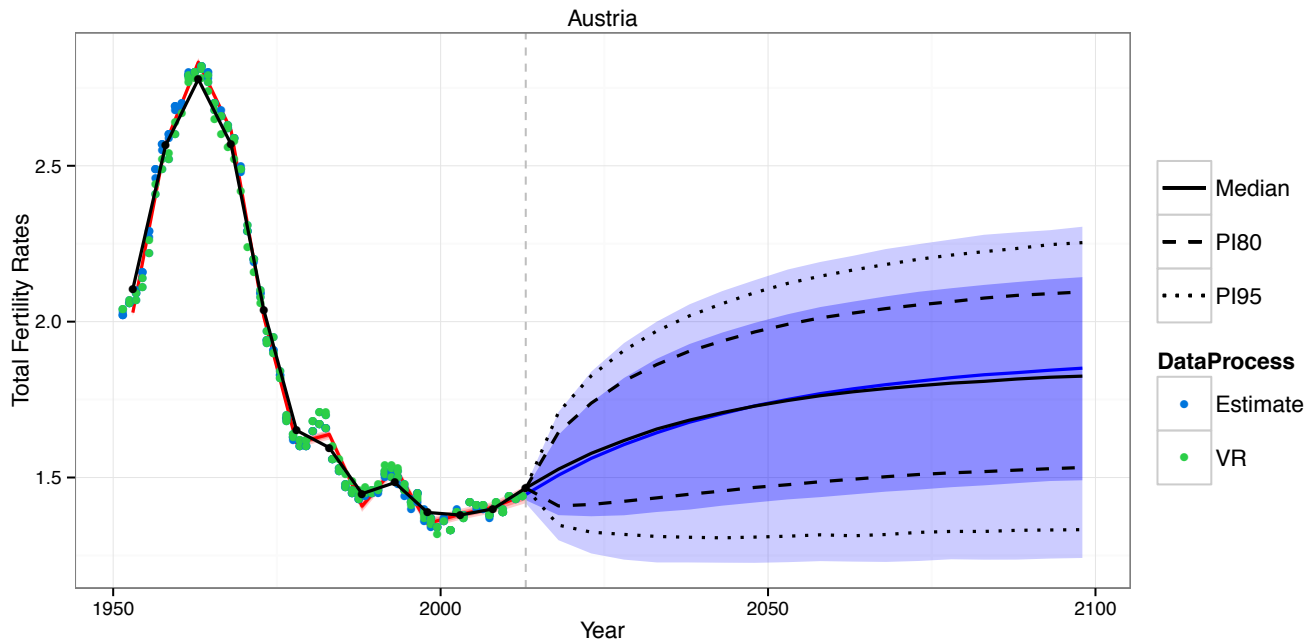
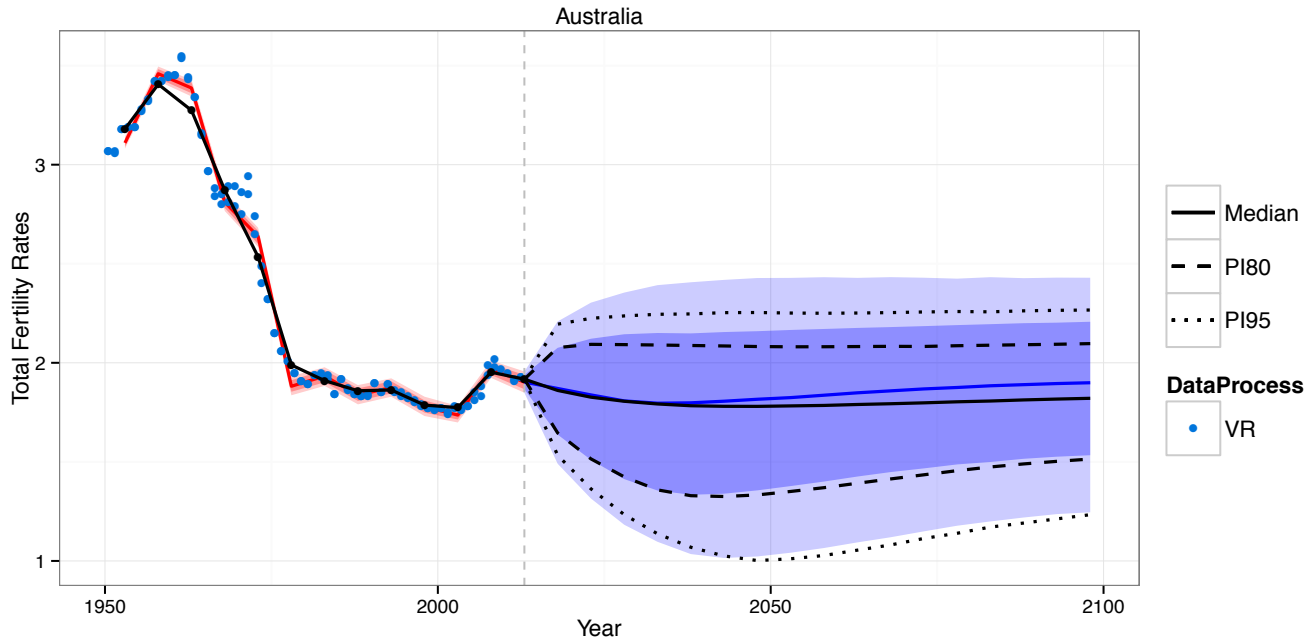
SM 4. Estimation and Projection for All Countries. We give the results of estimation of past TFR from 1950 to 2015, and probabilistic projections of future TFR from 2015 to 2100 for all countries with current populations over 100,000 except for Antigua and Barbuda and Seychelles (for about 90,000).

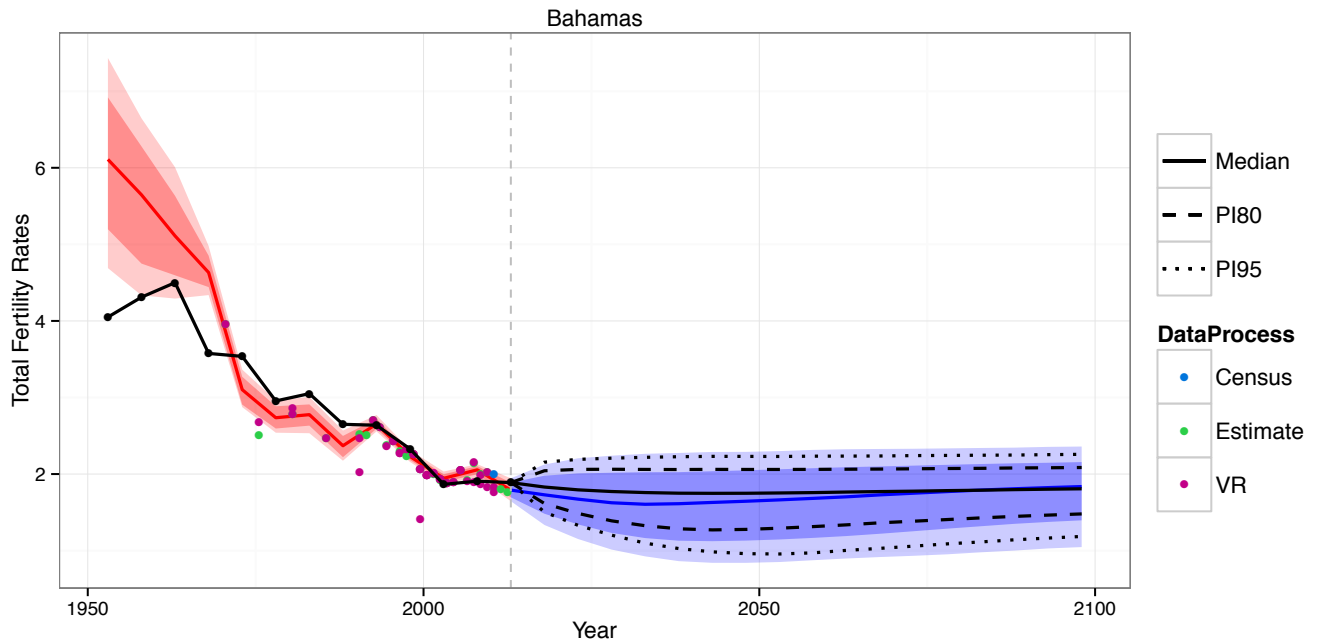
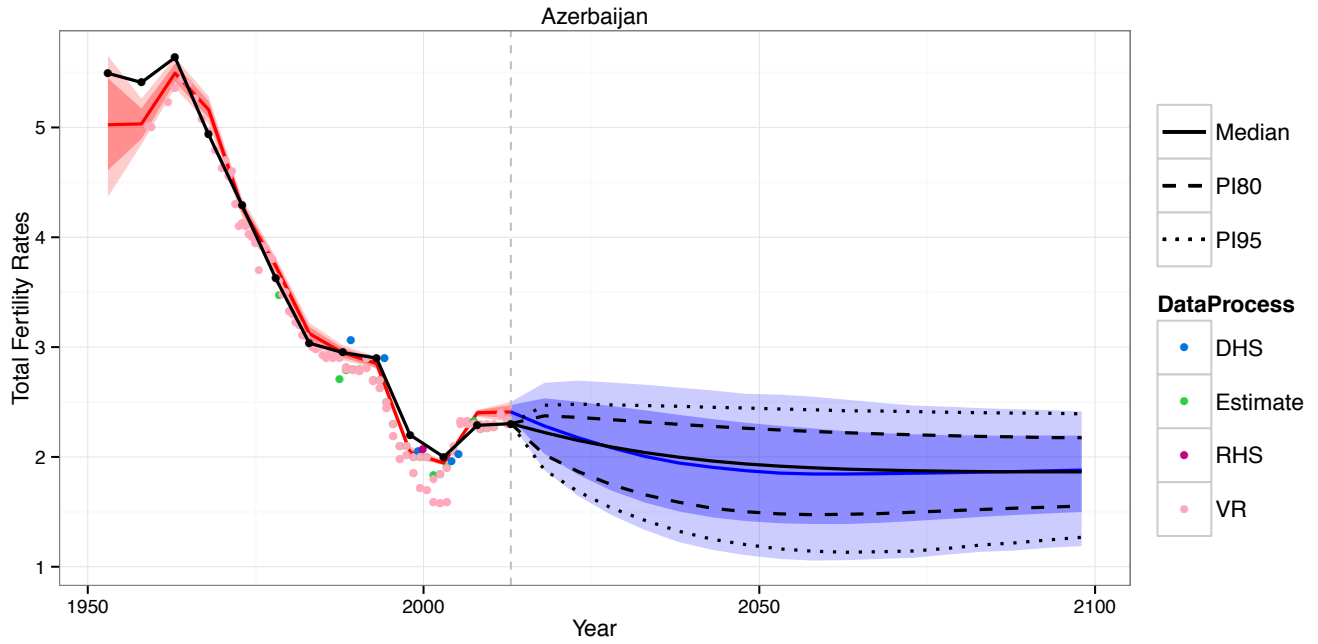


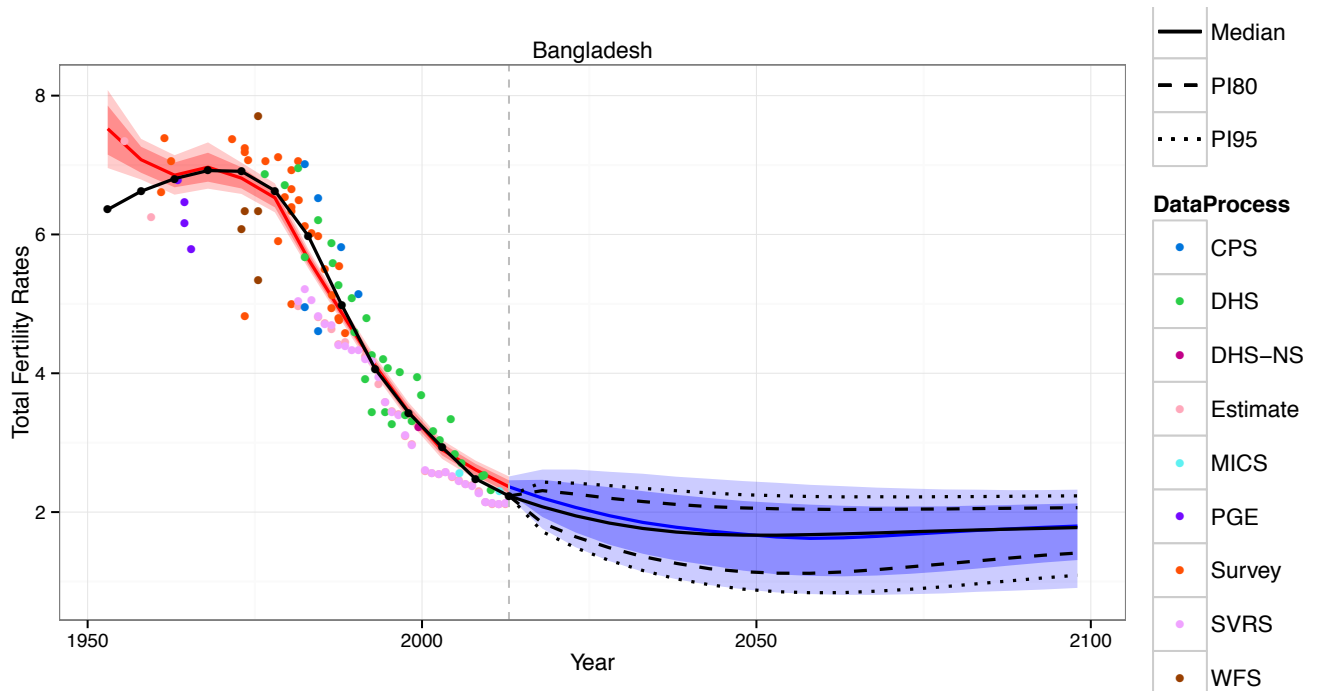
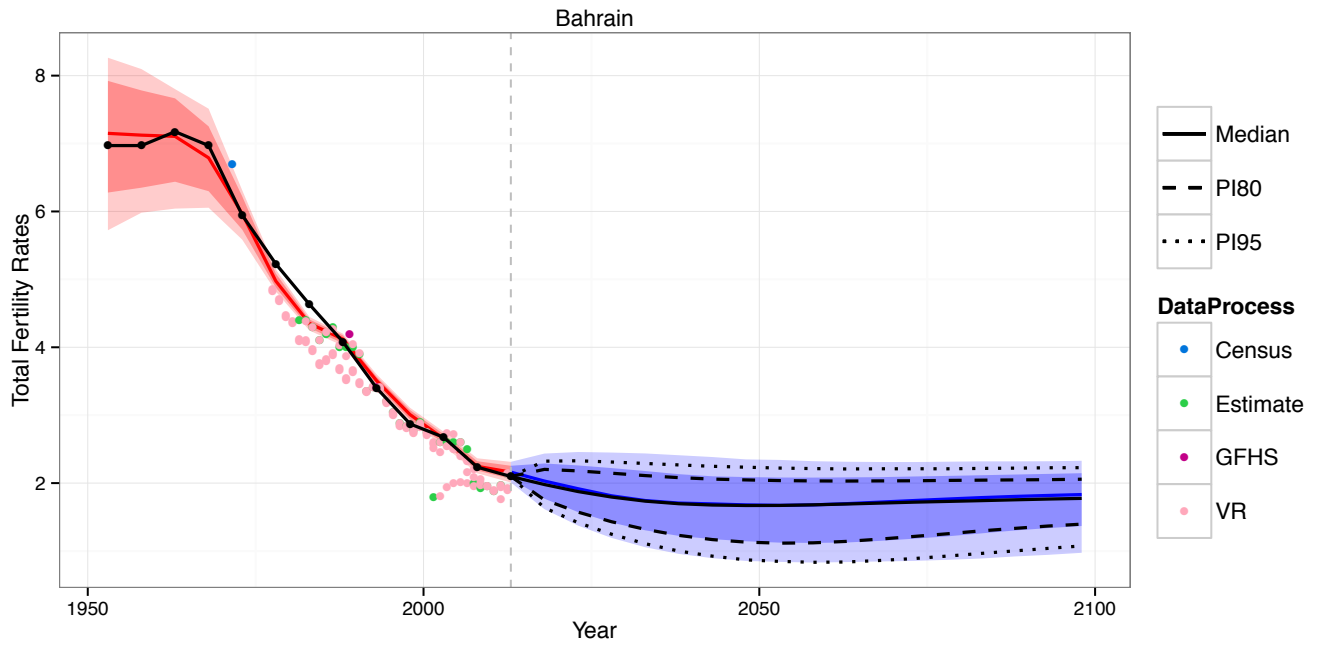


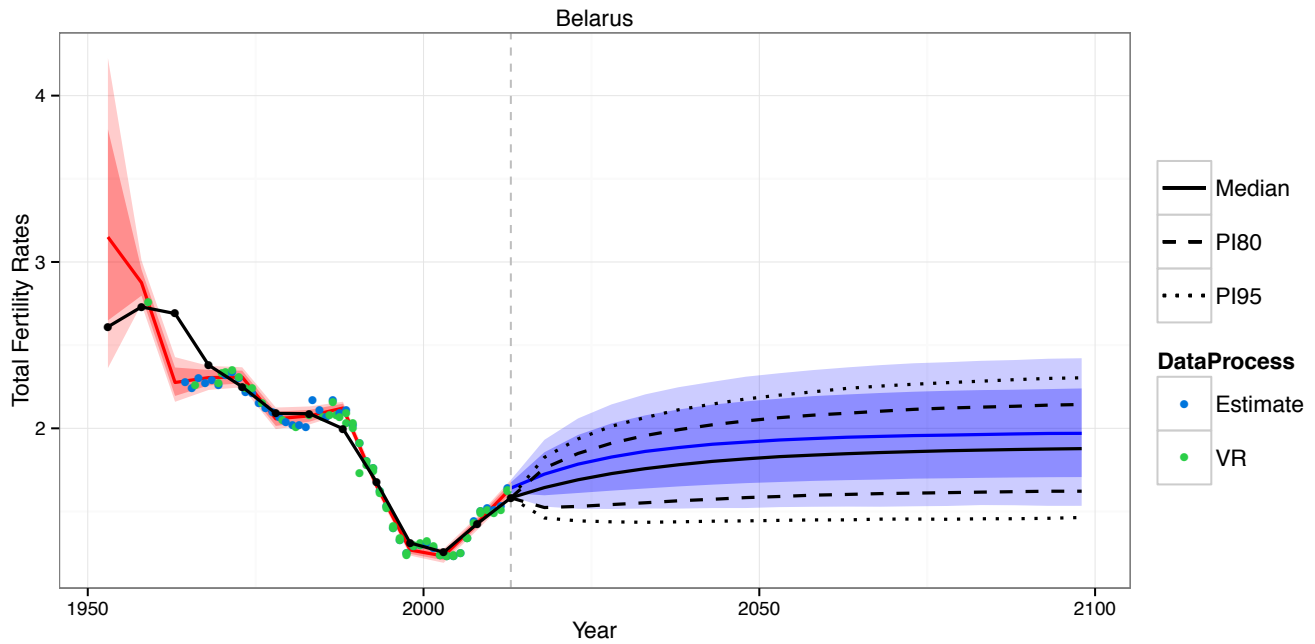
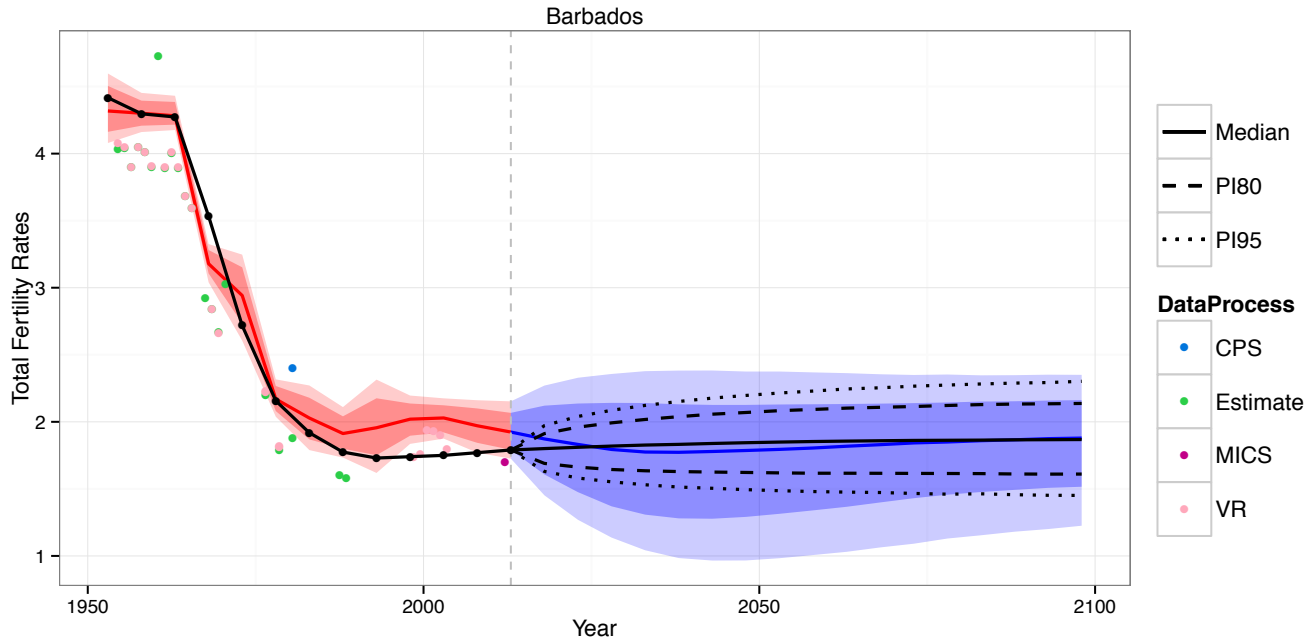


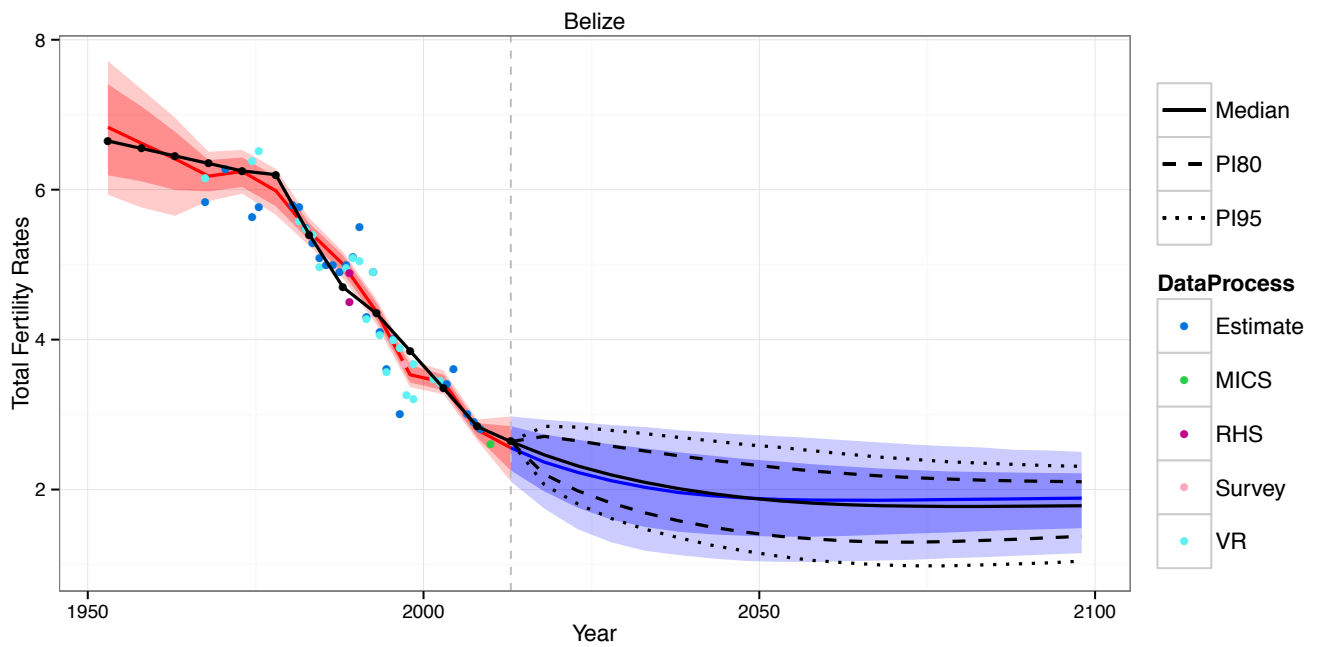
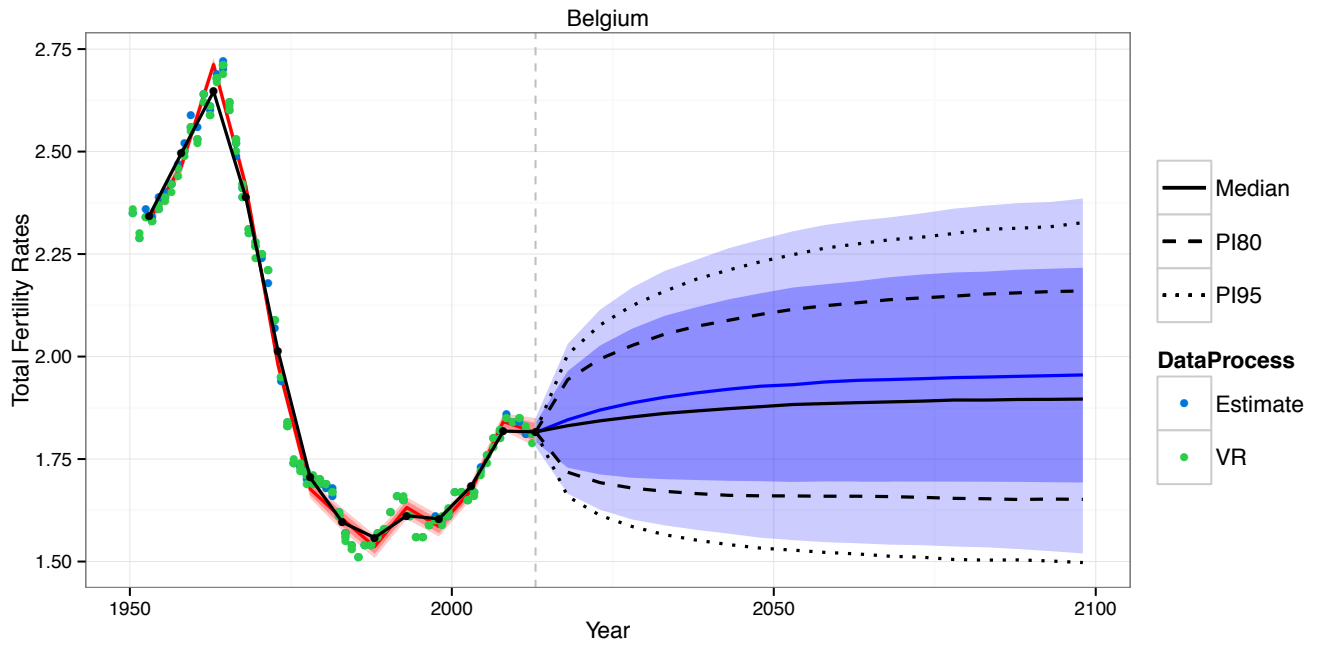


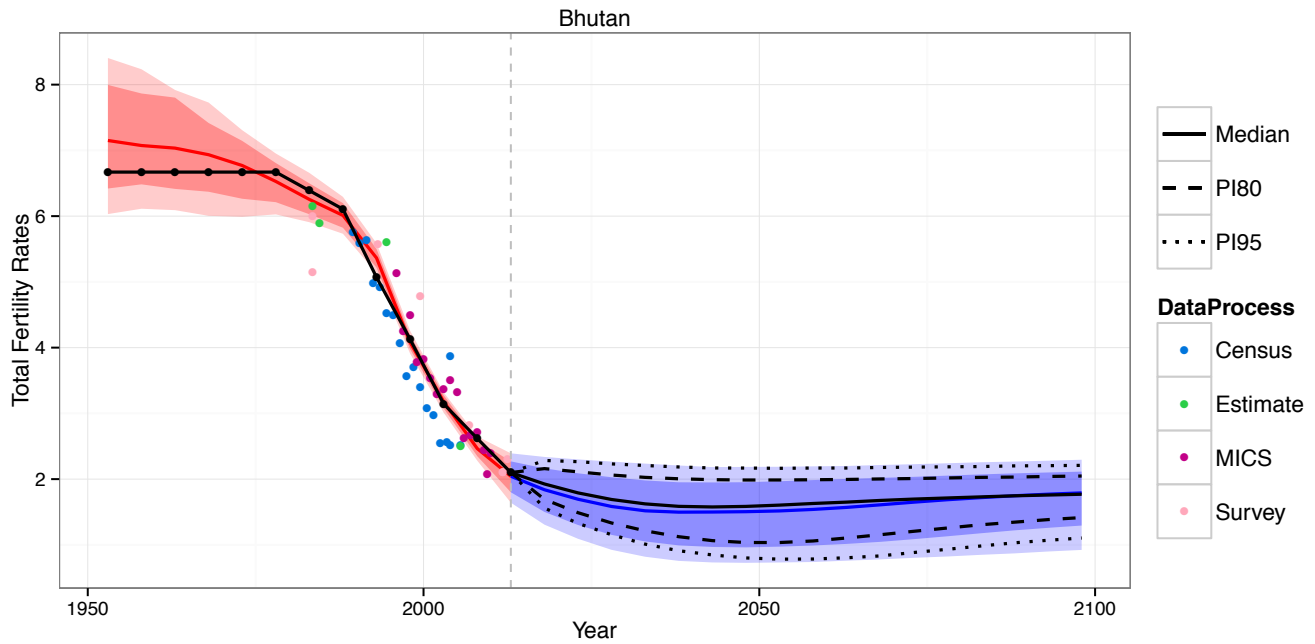
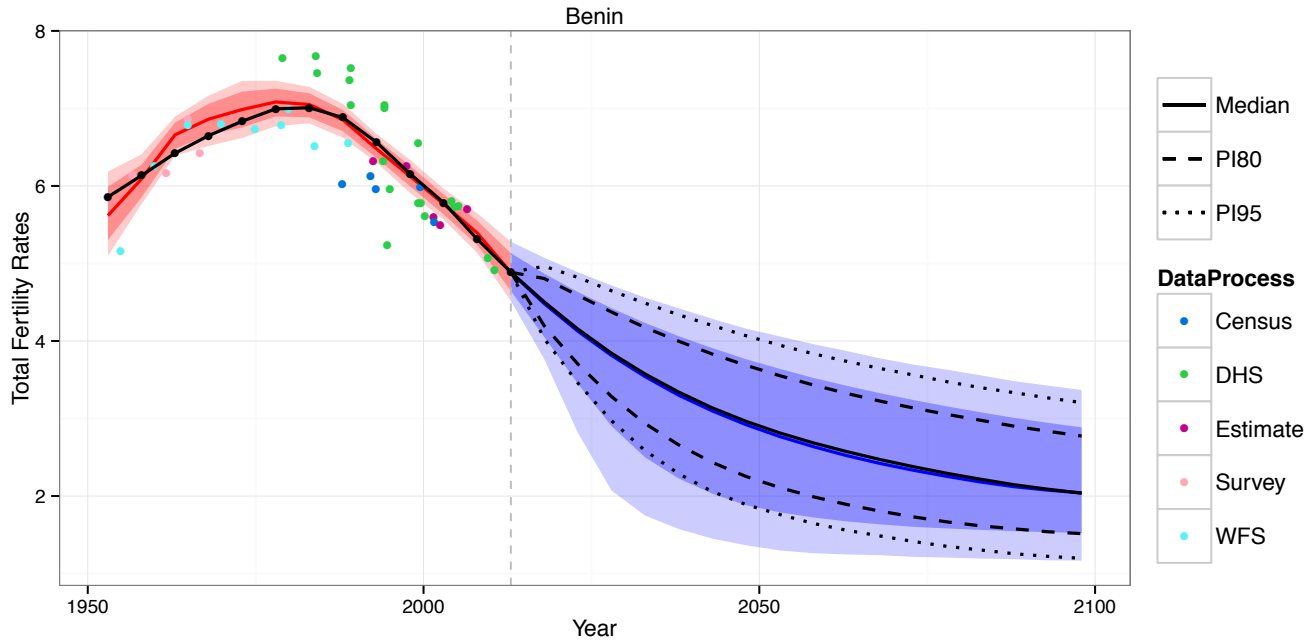


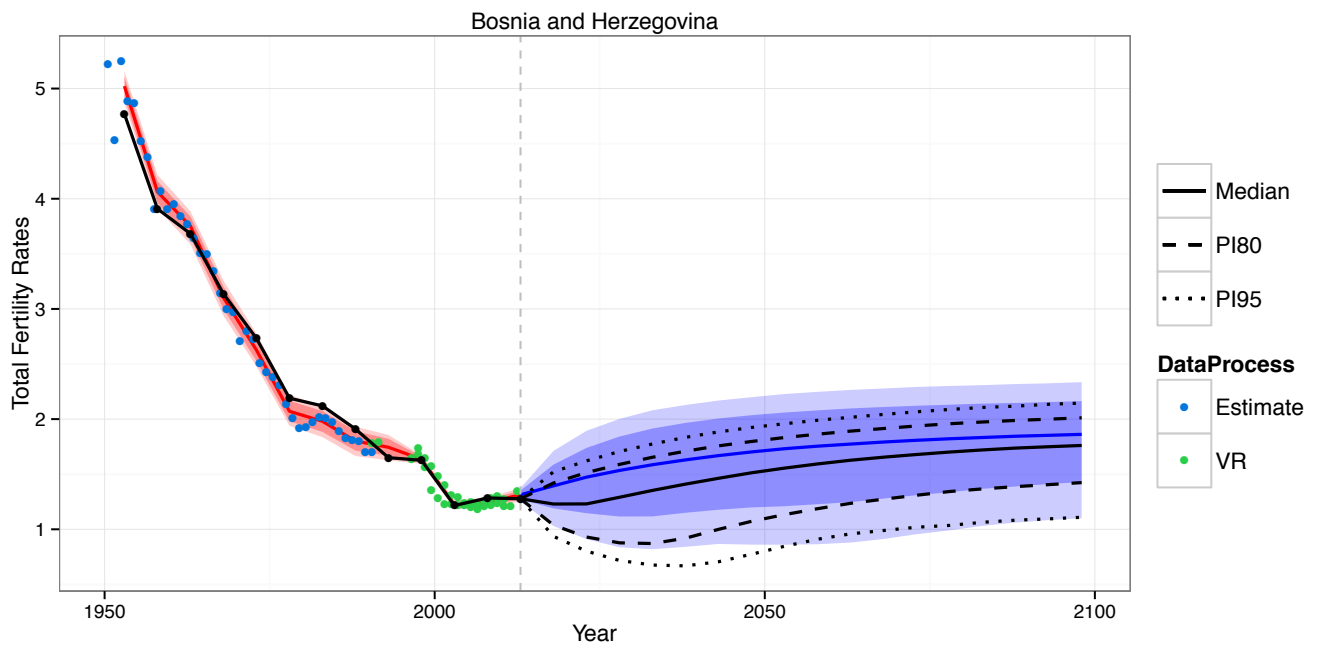
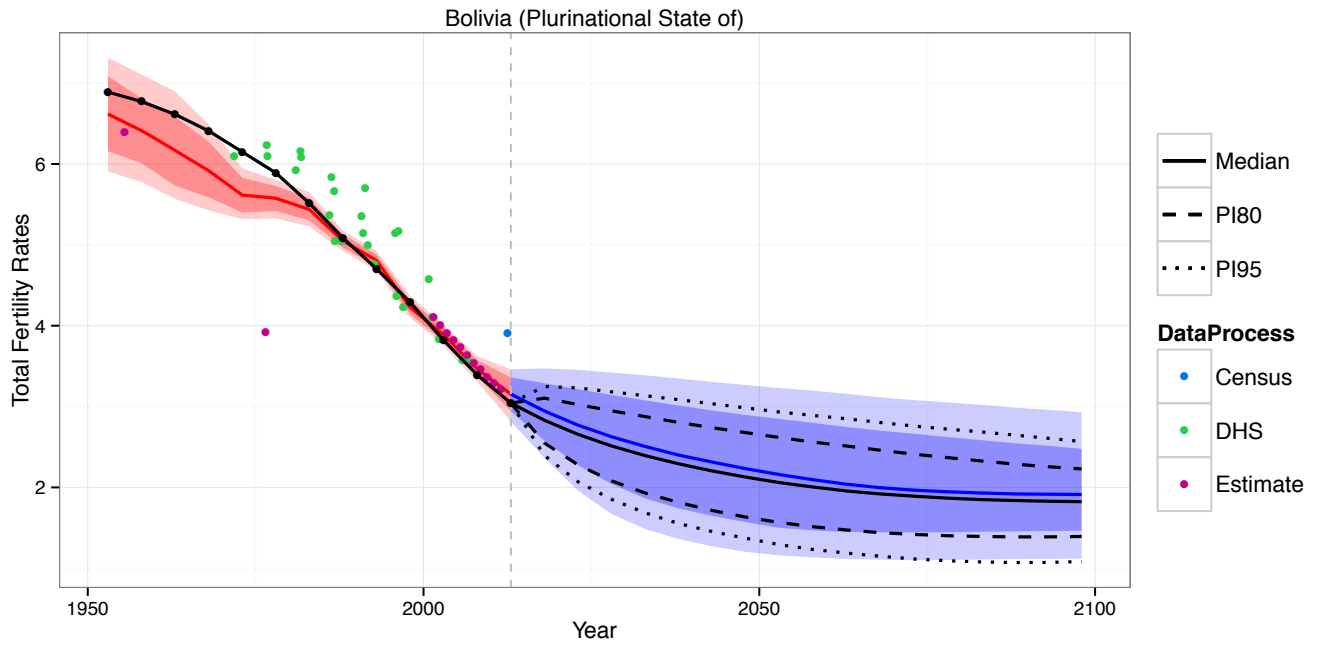


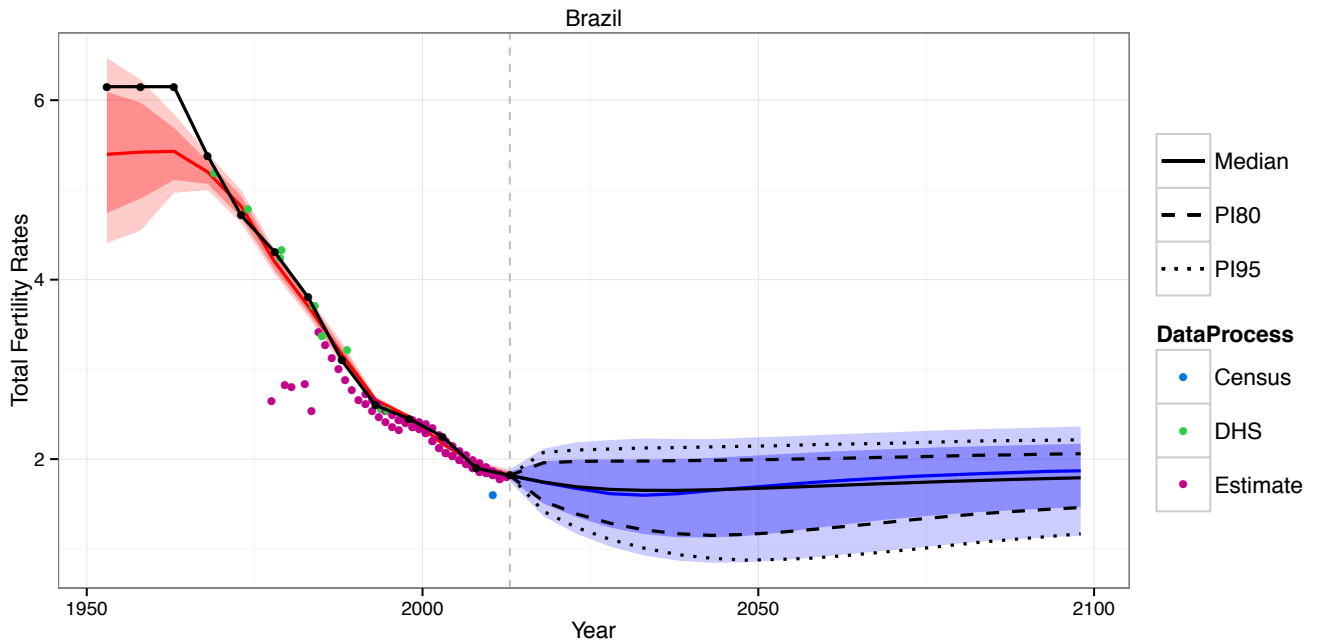
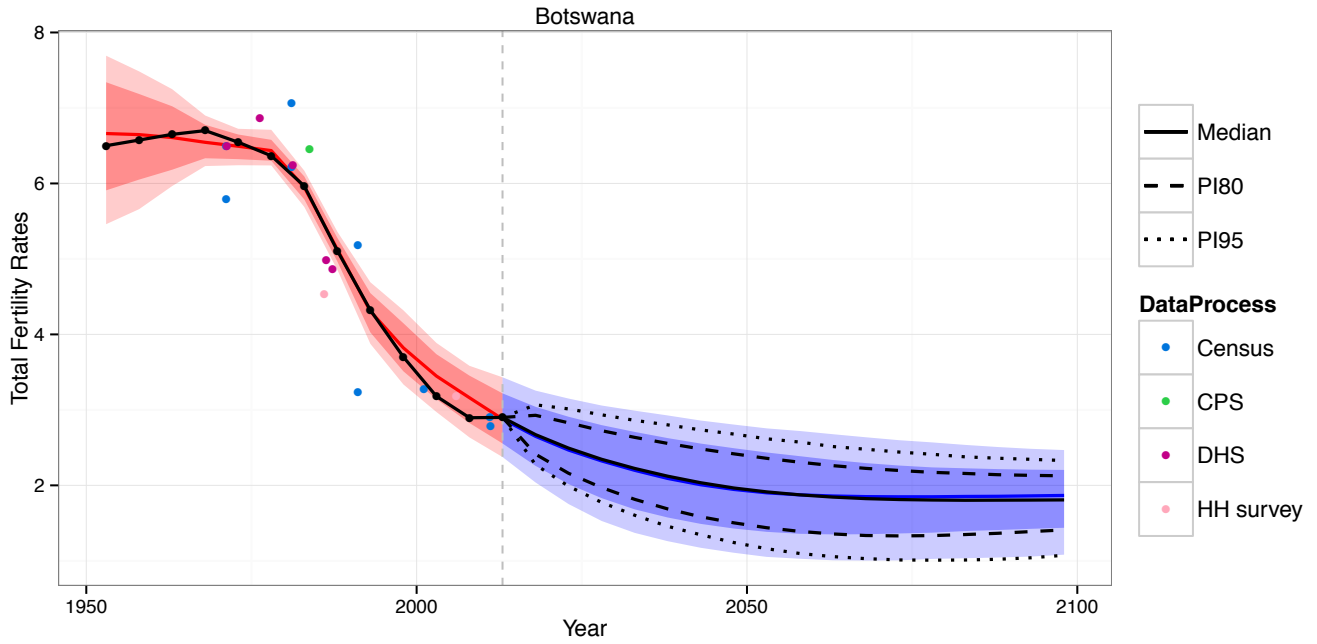


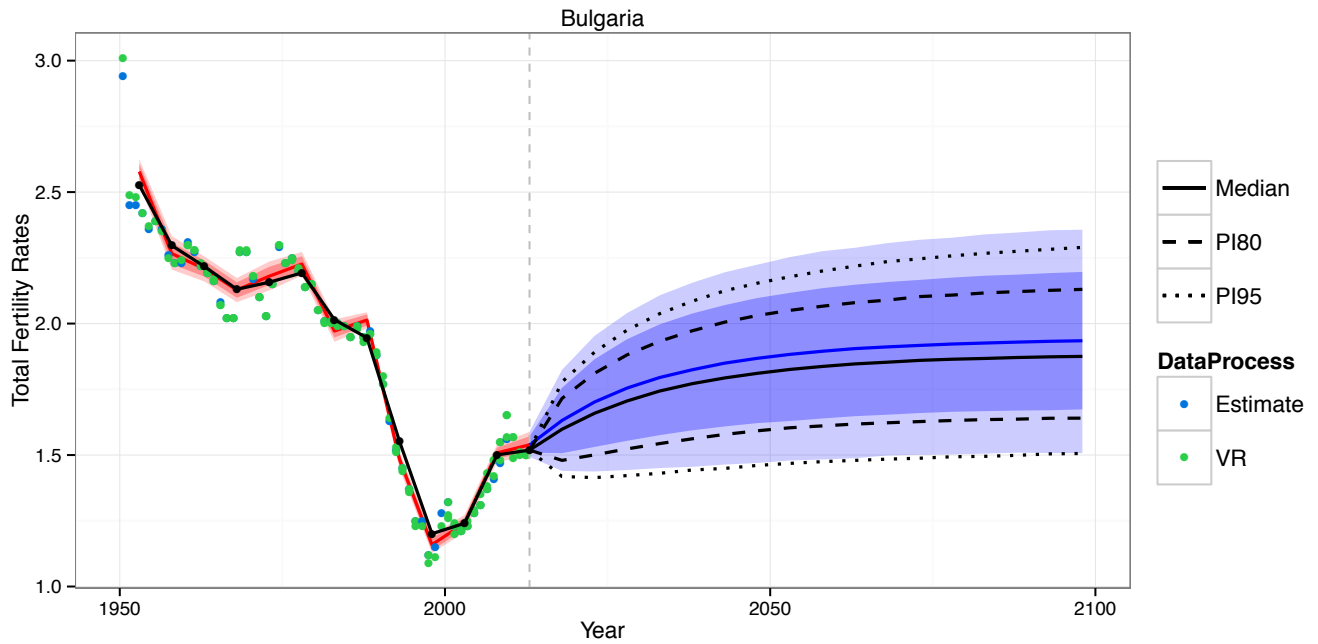
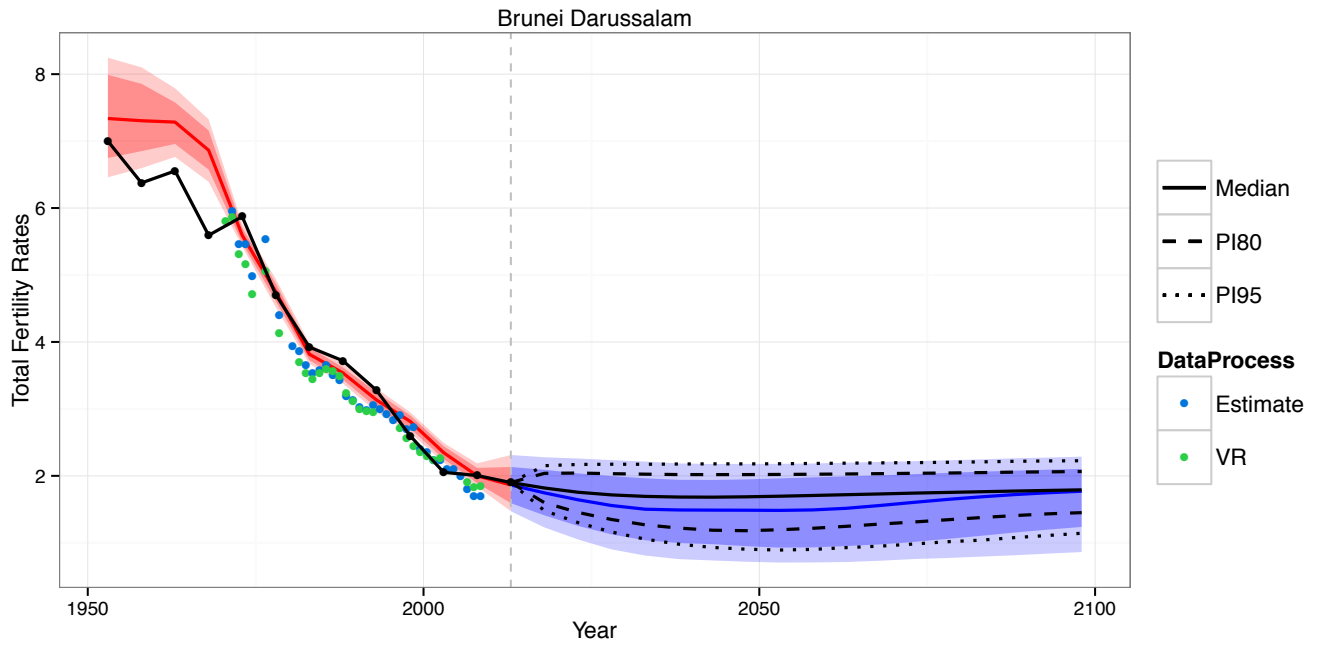


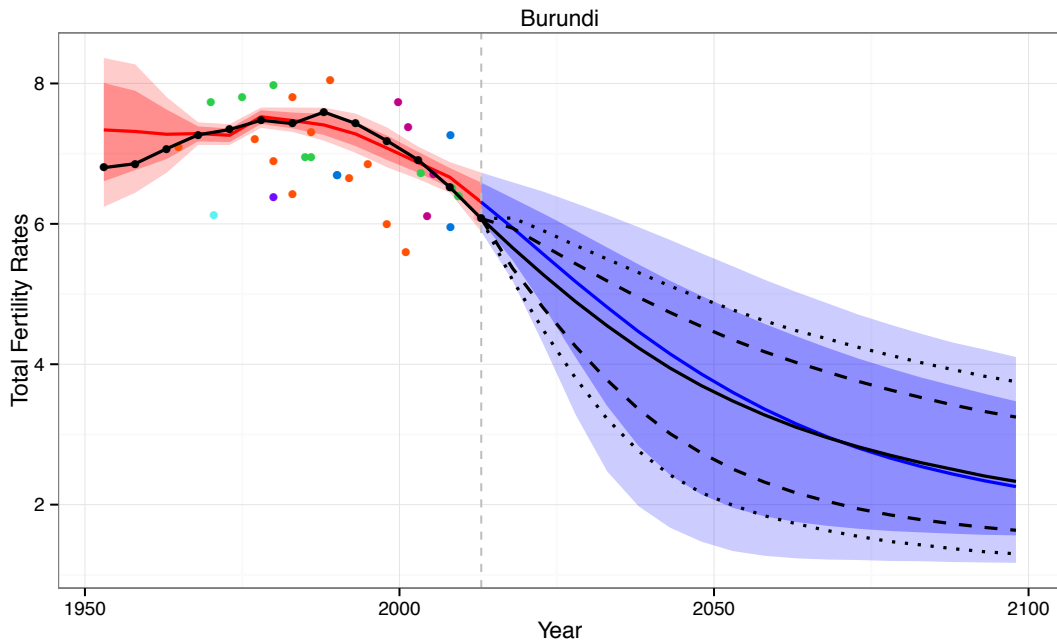
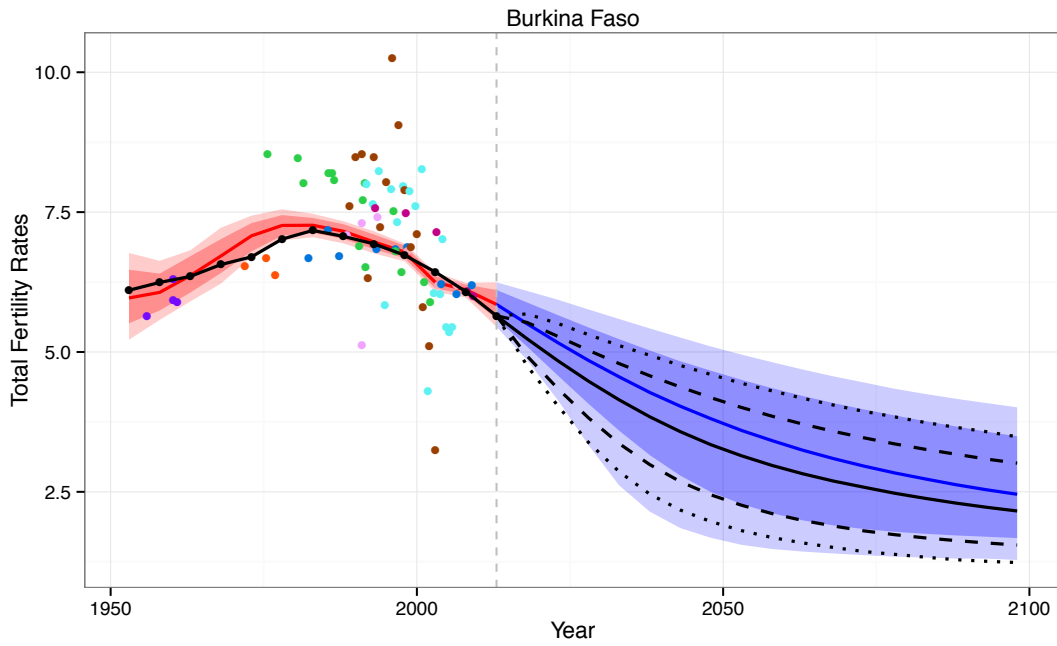


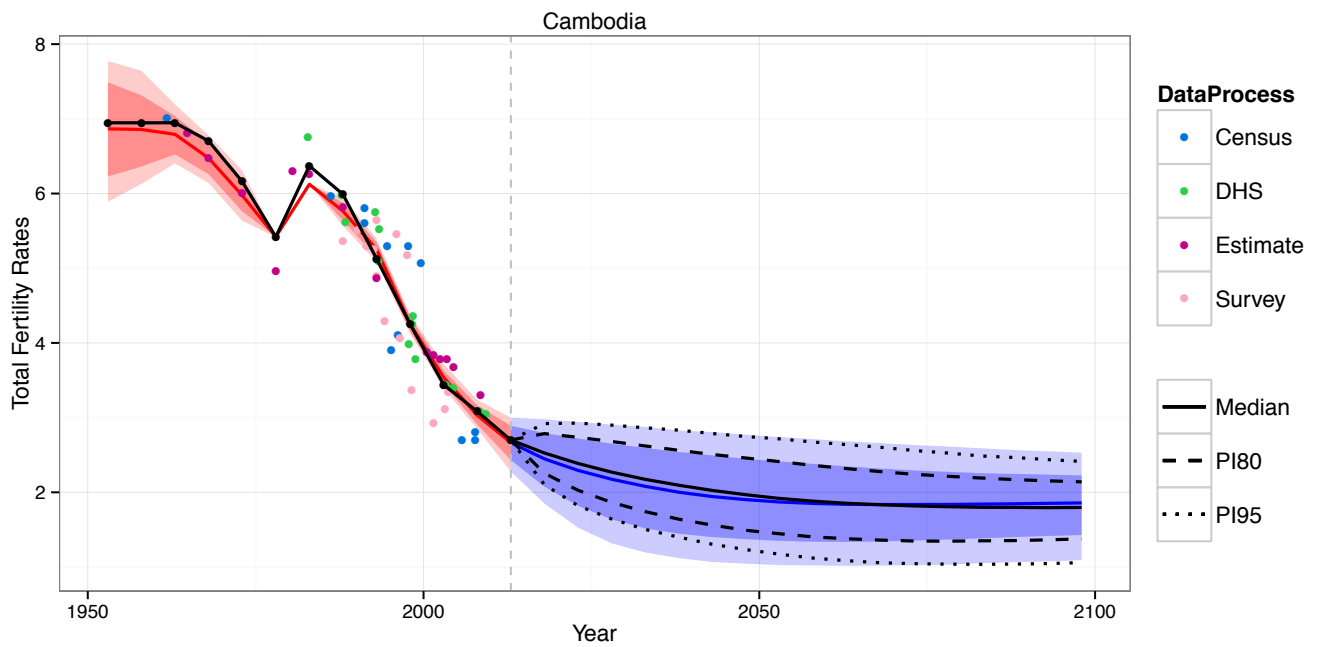
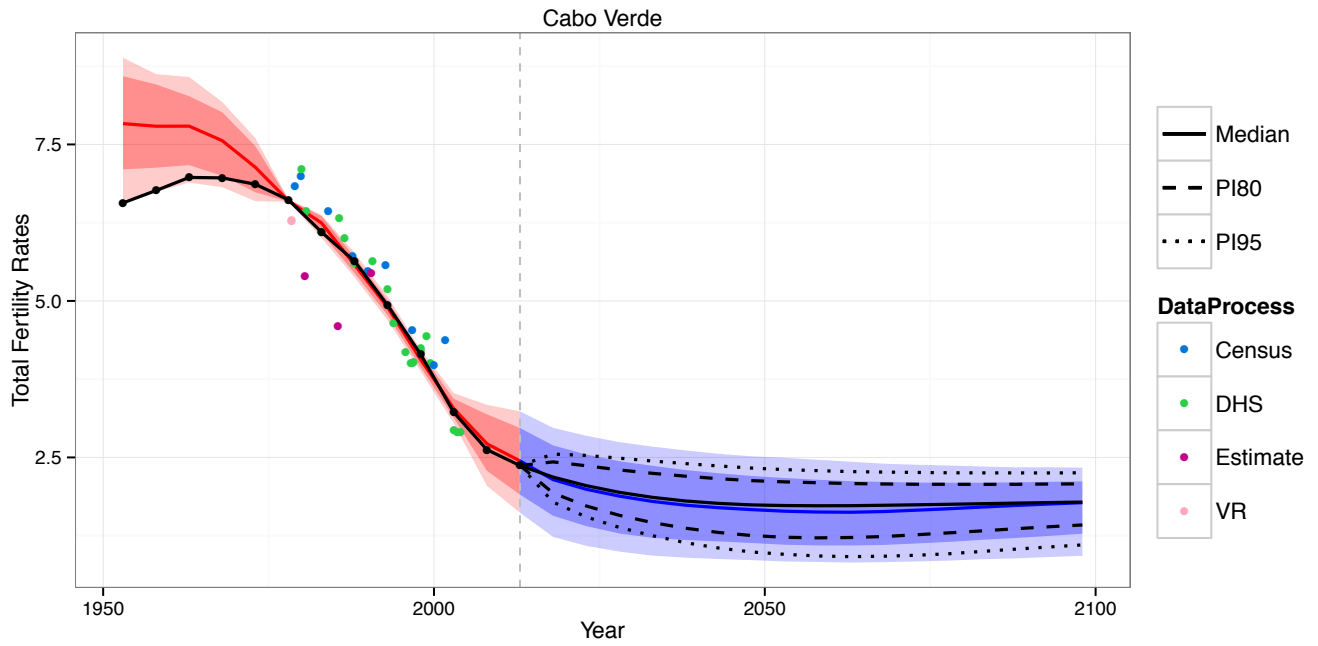


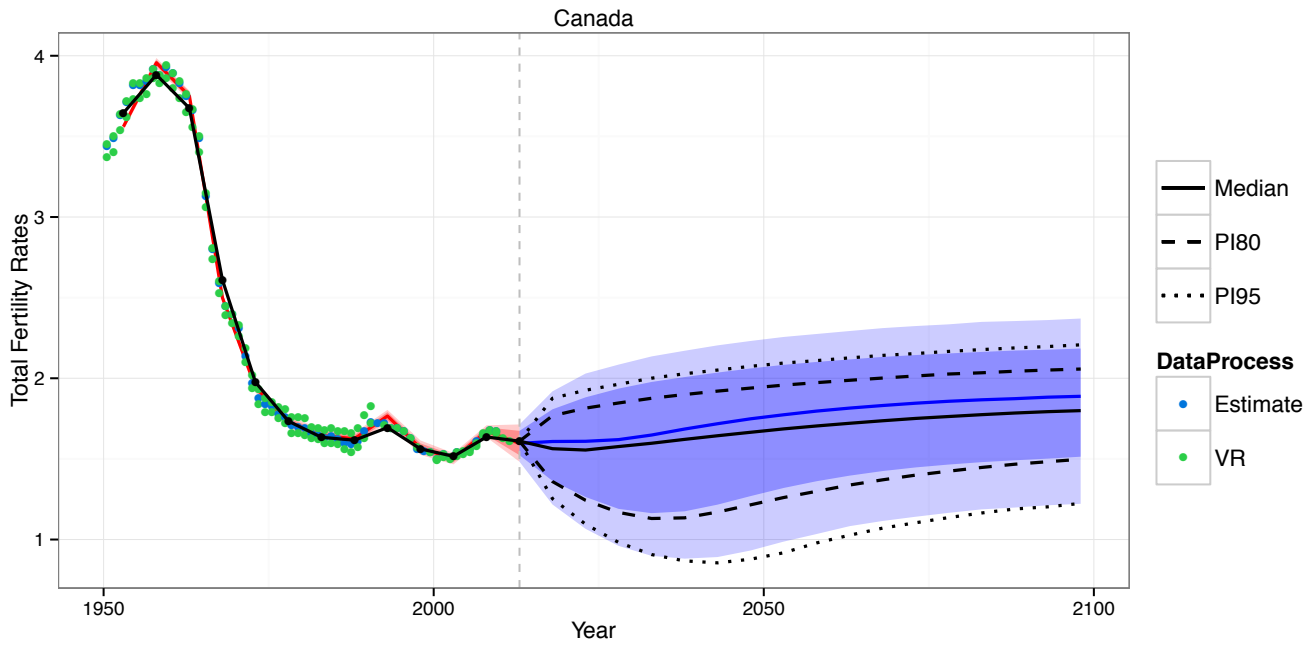
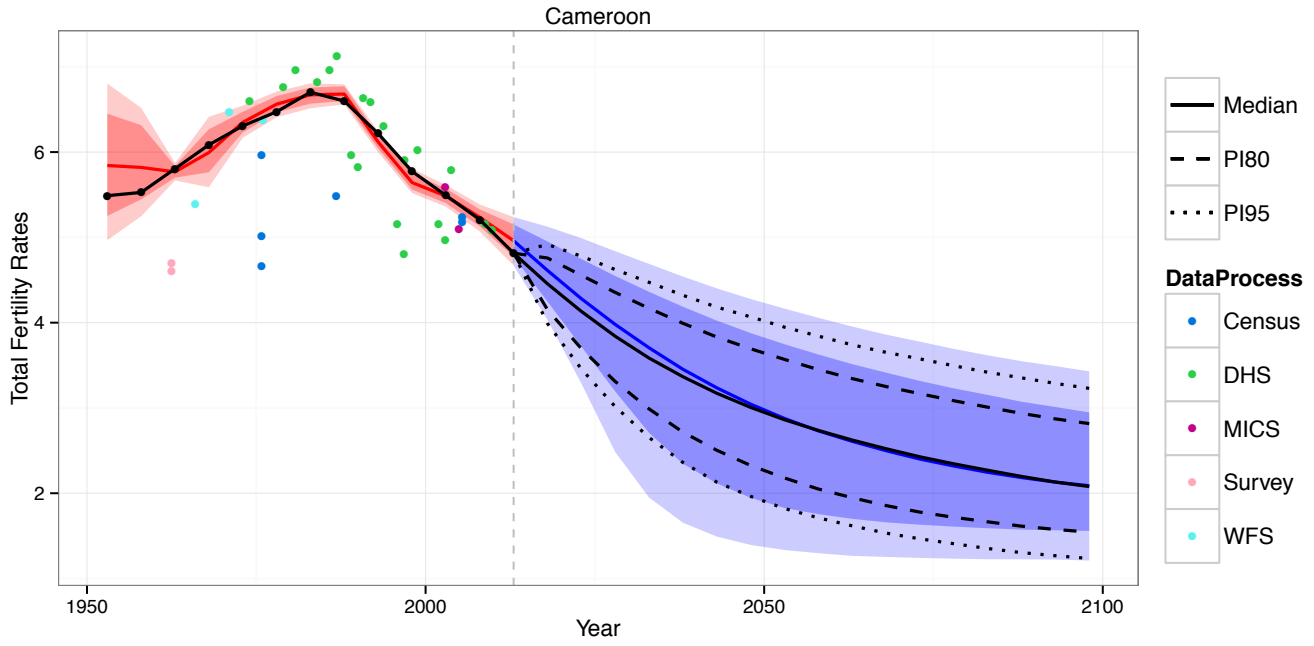


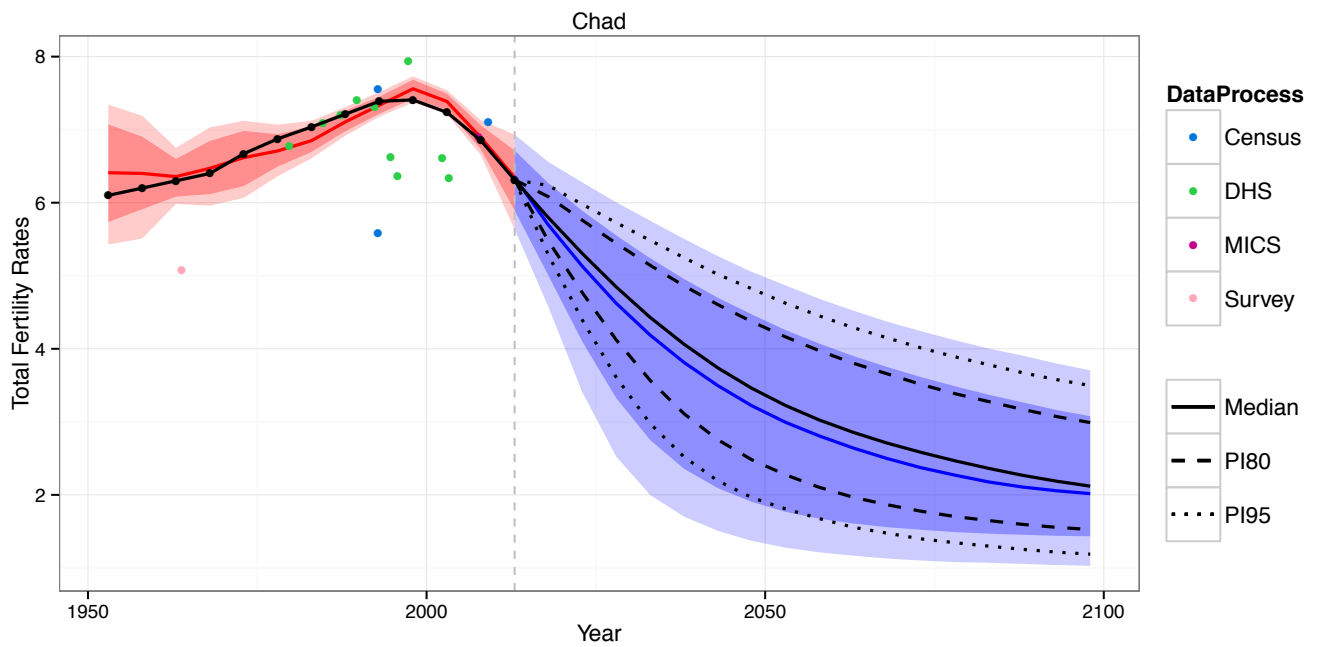
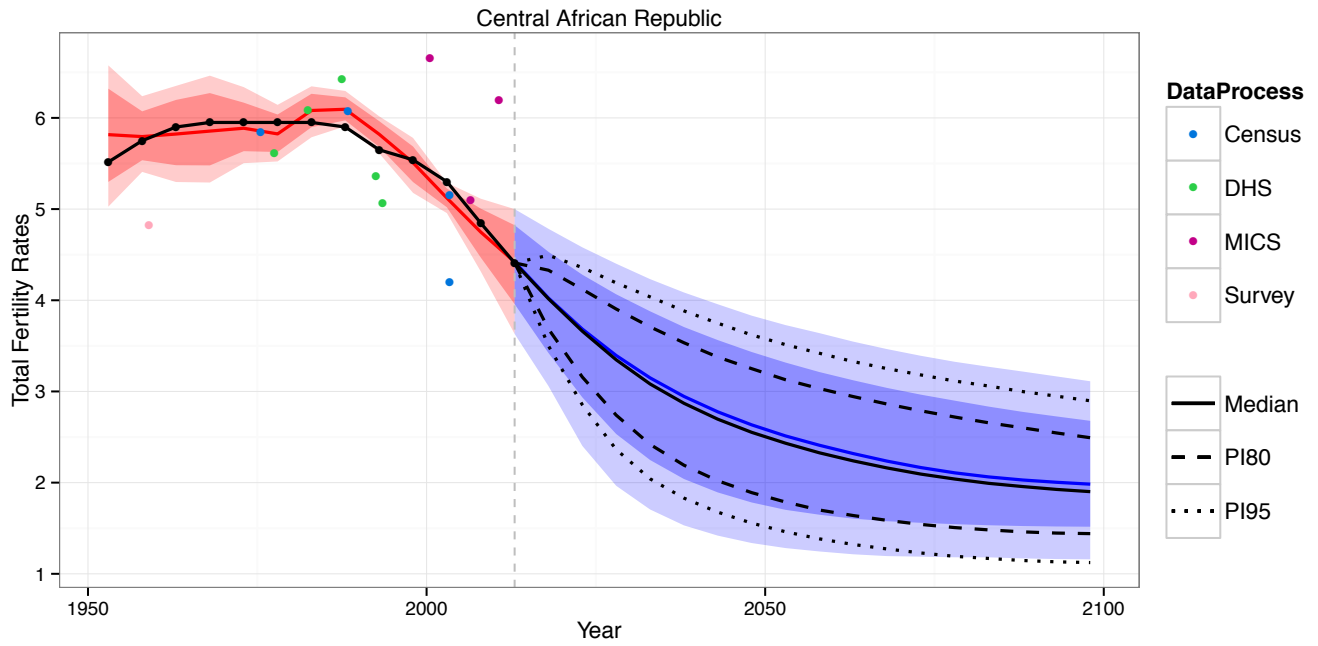


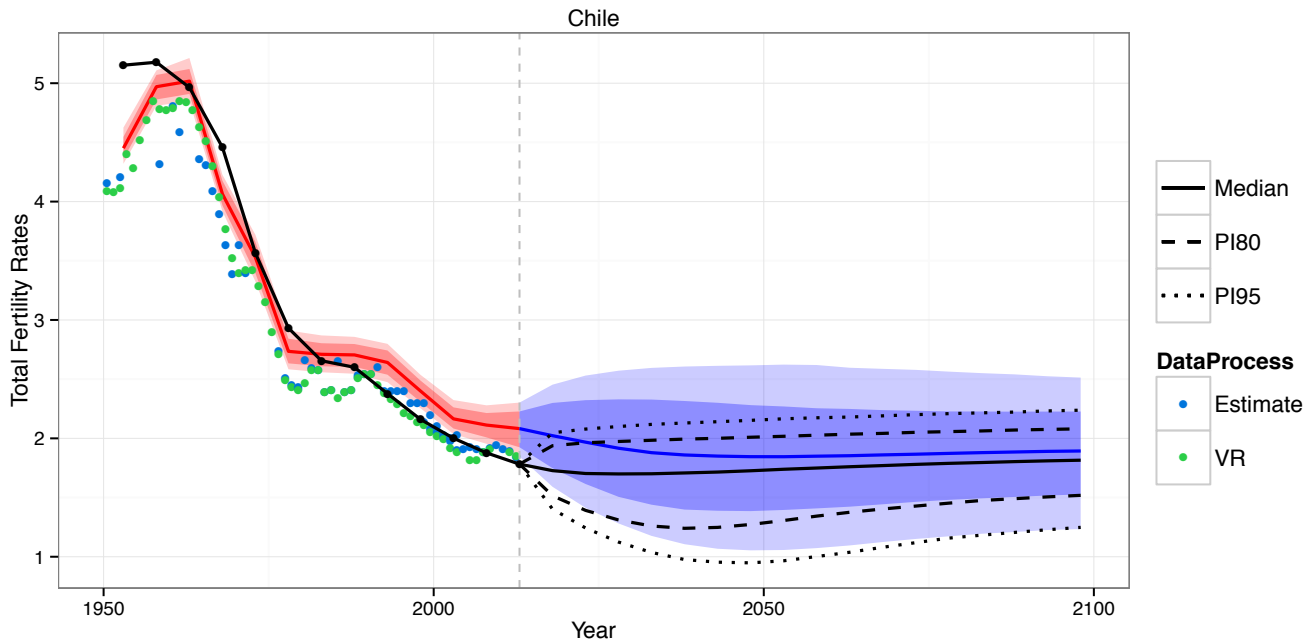
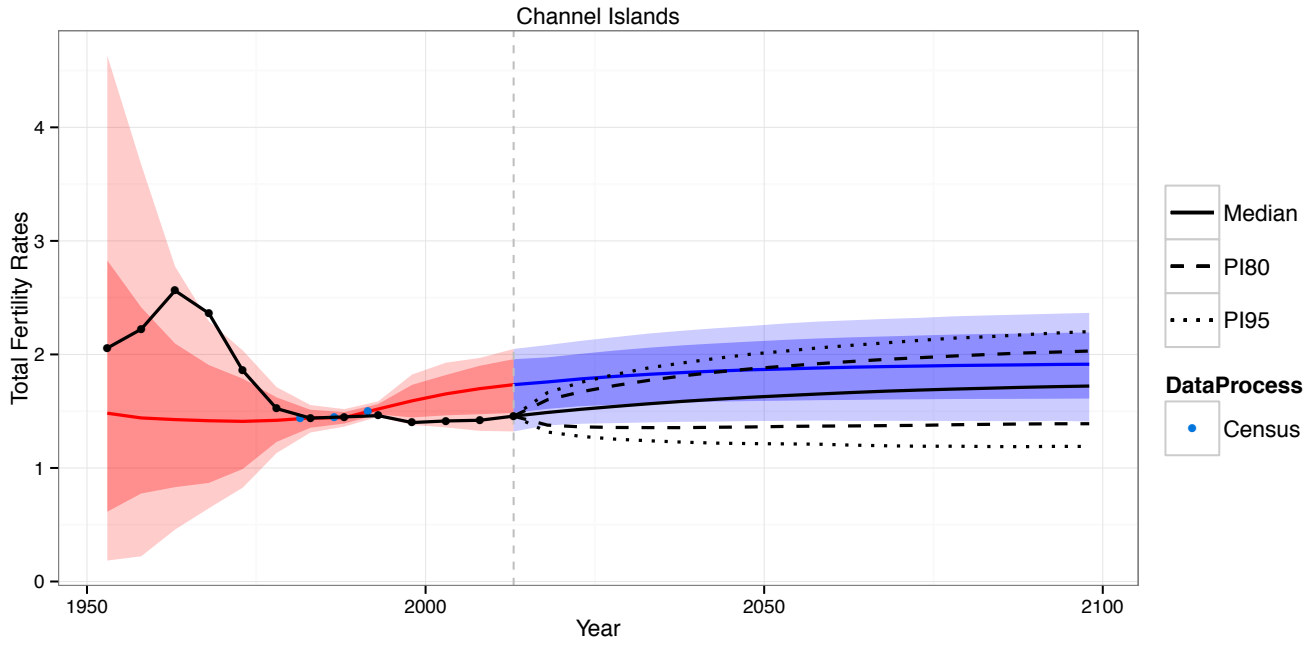


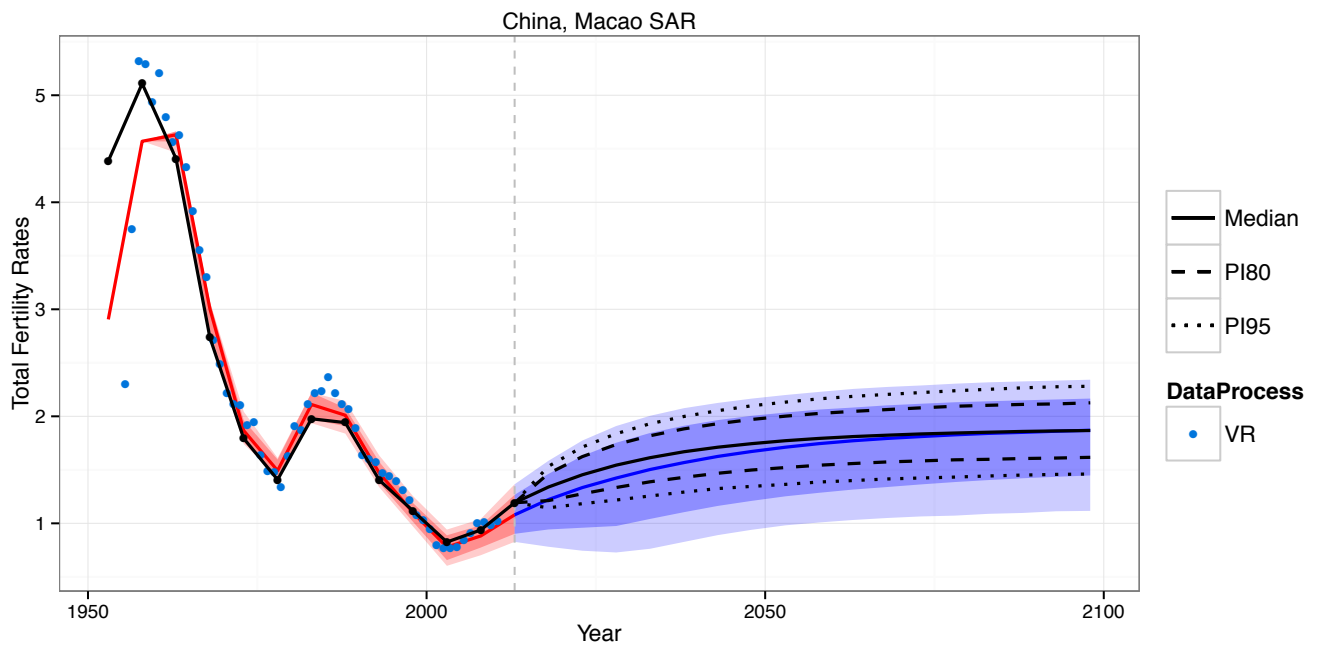
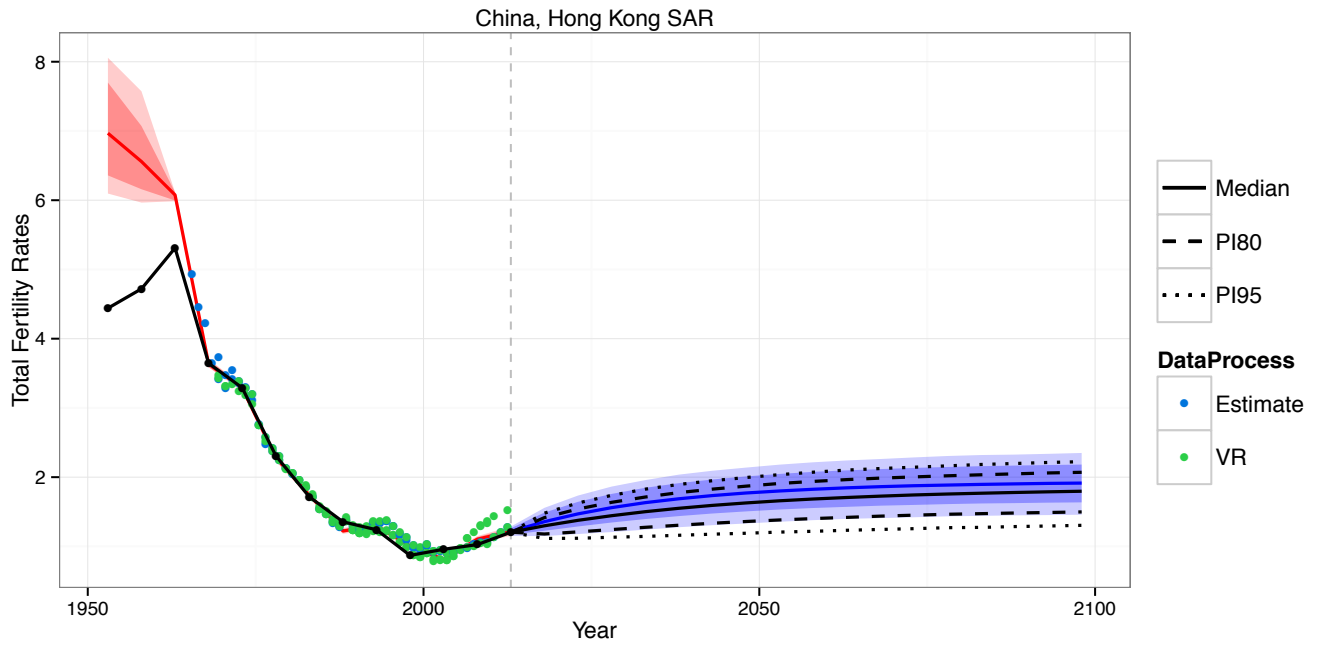


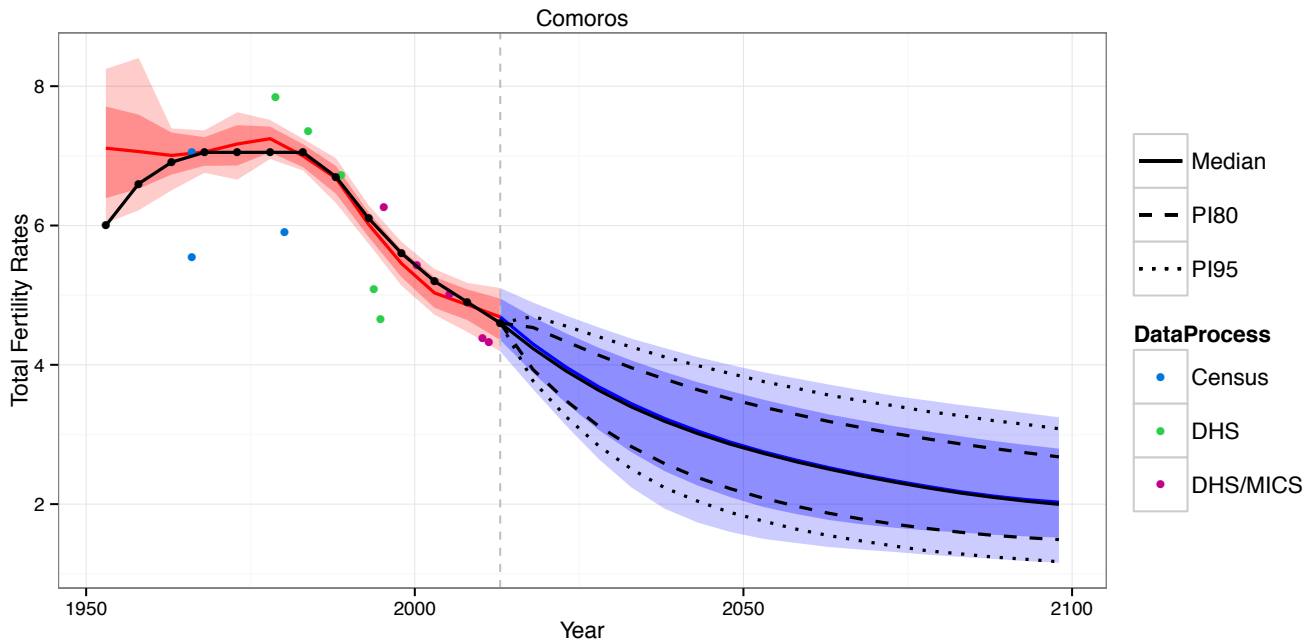
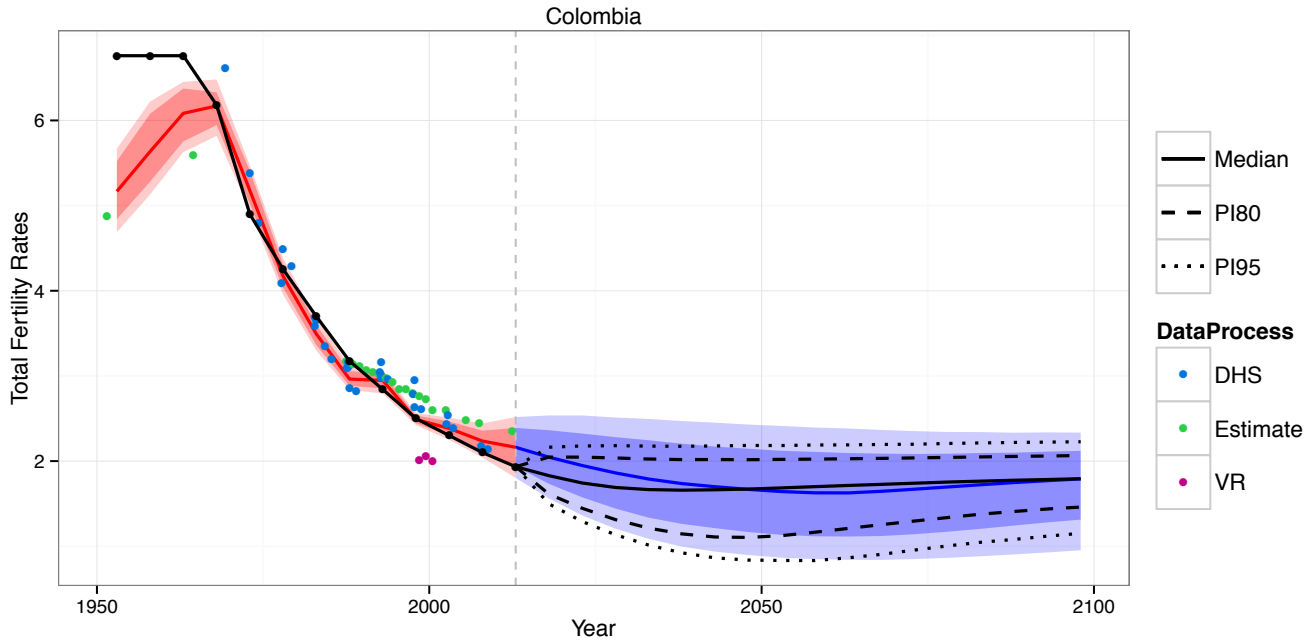


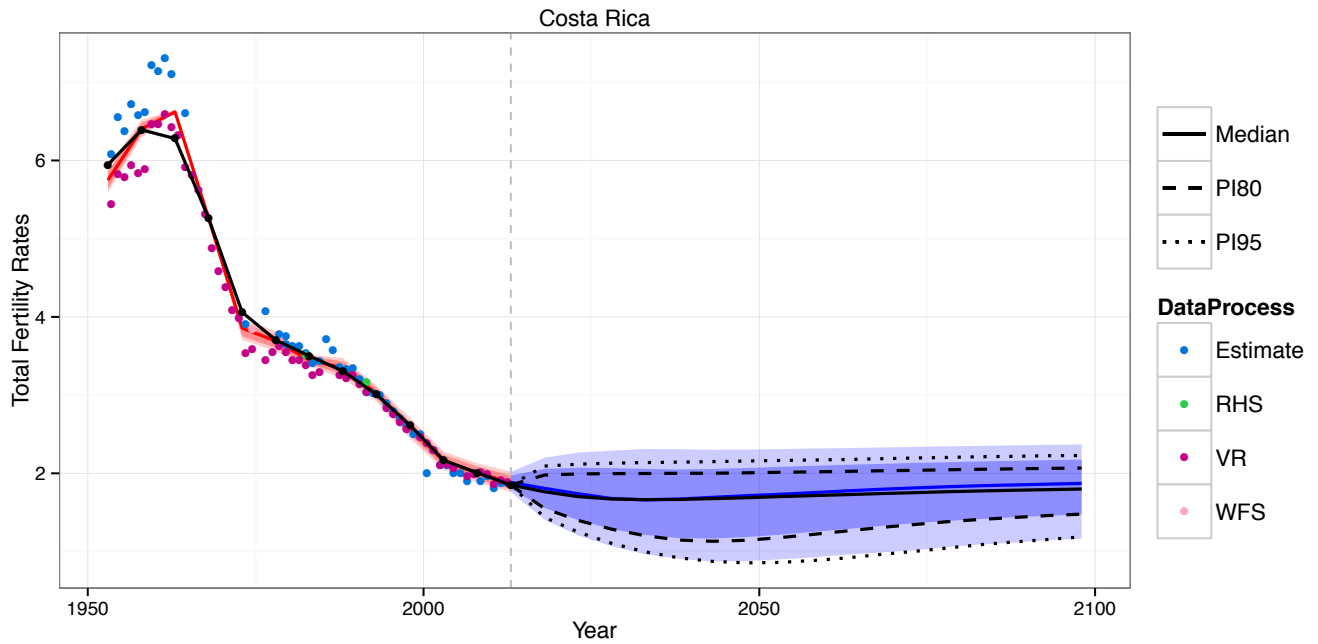
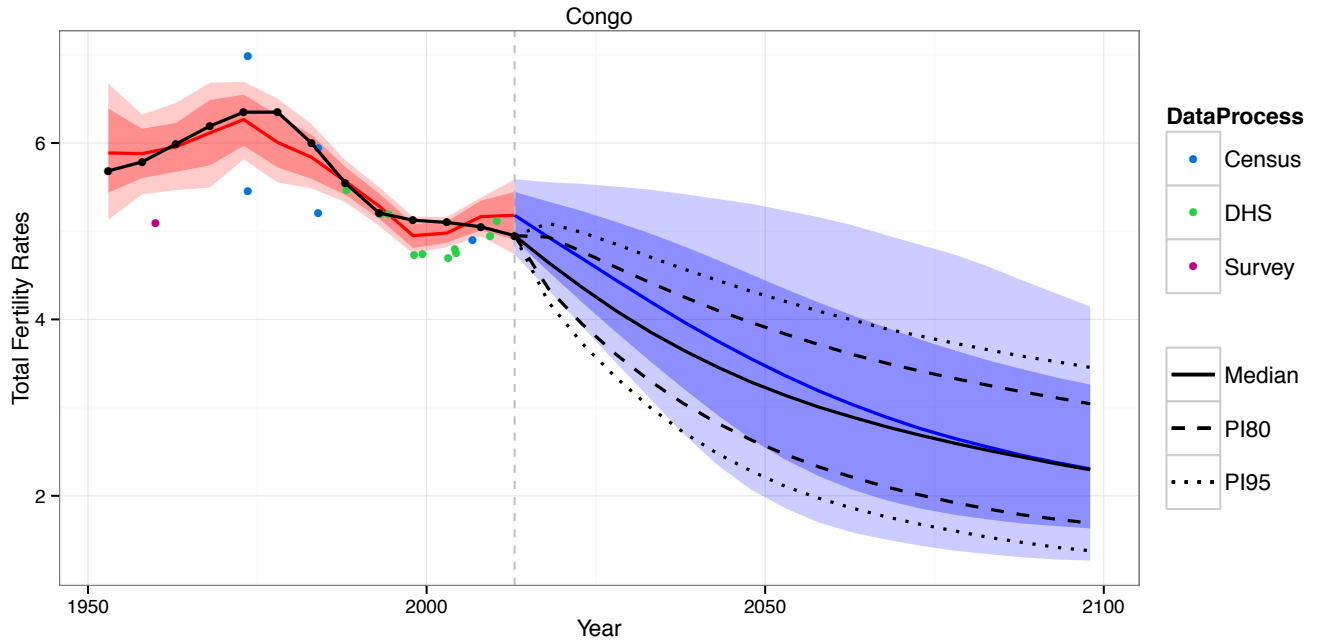


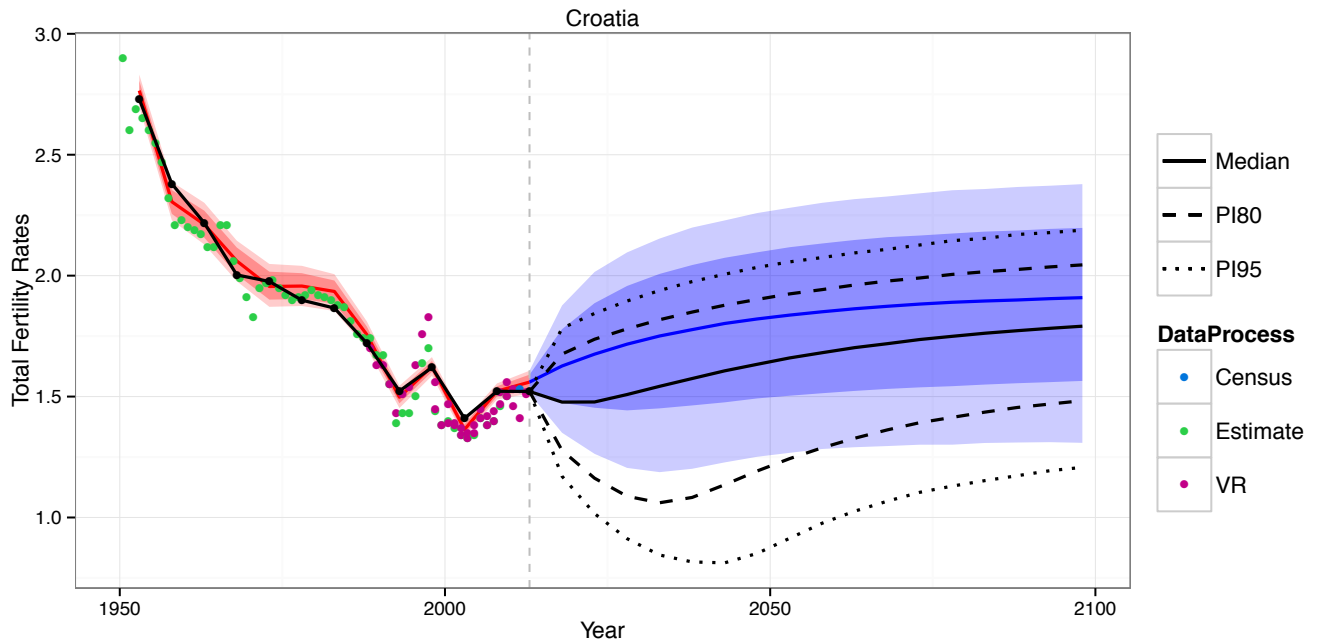
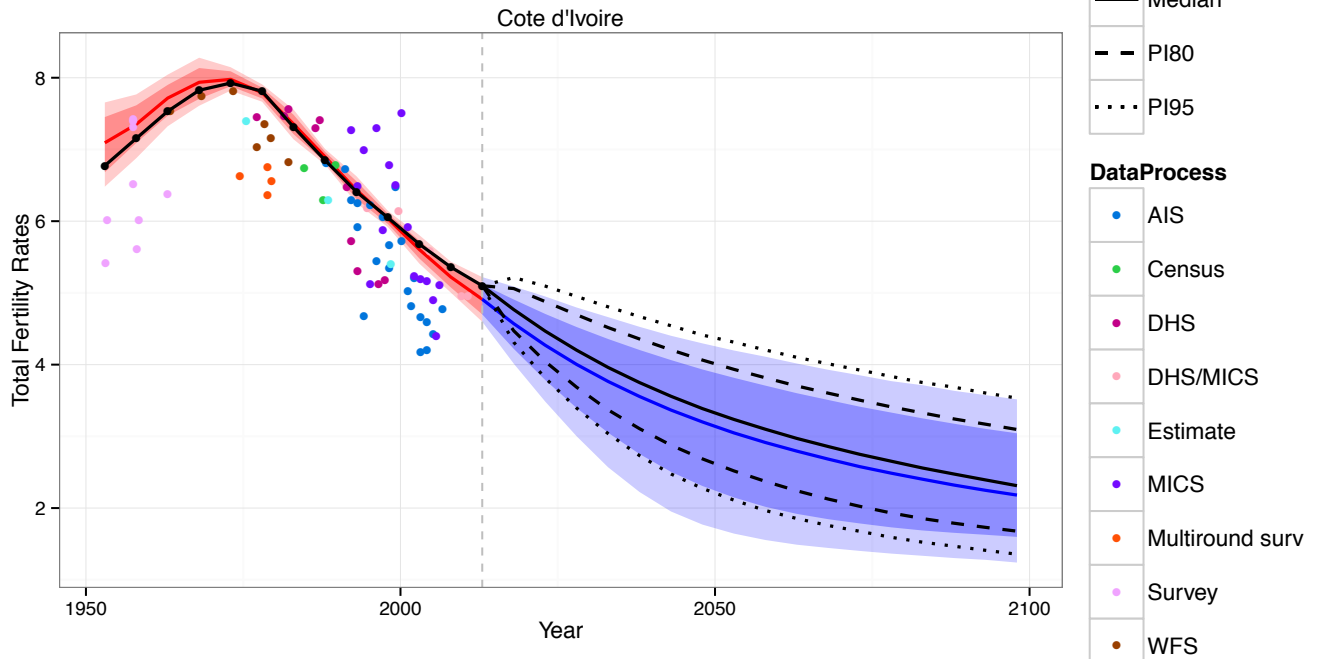


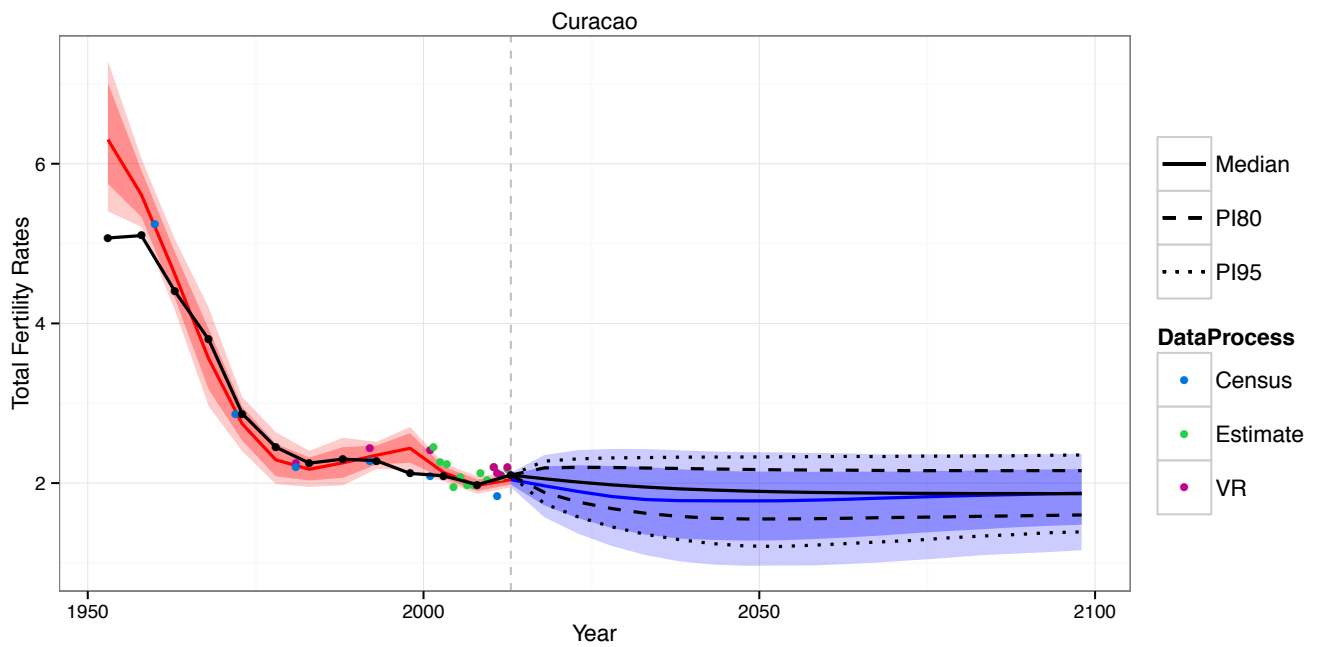
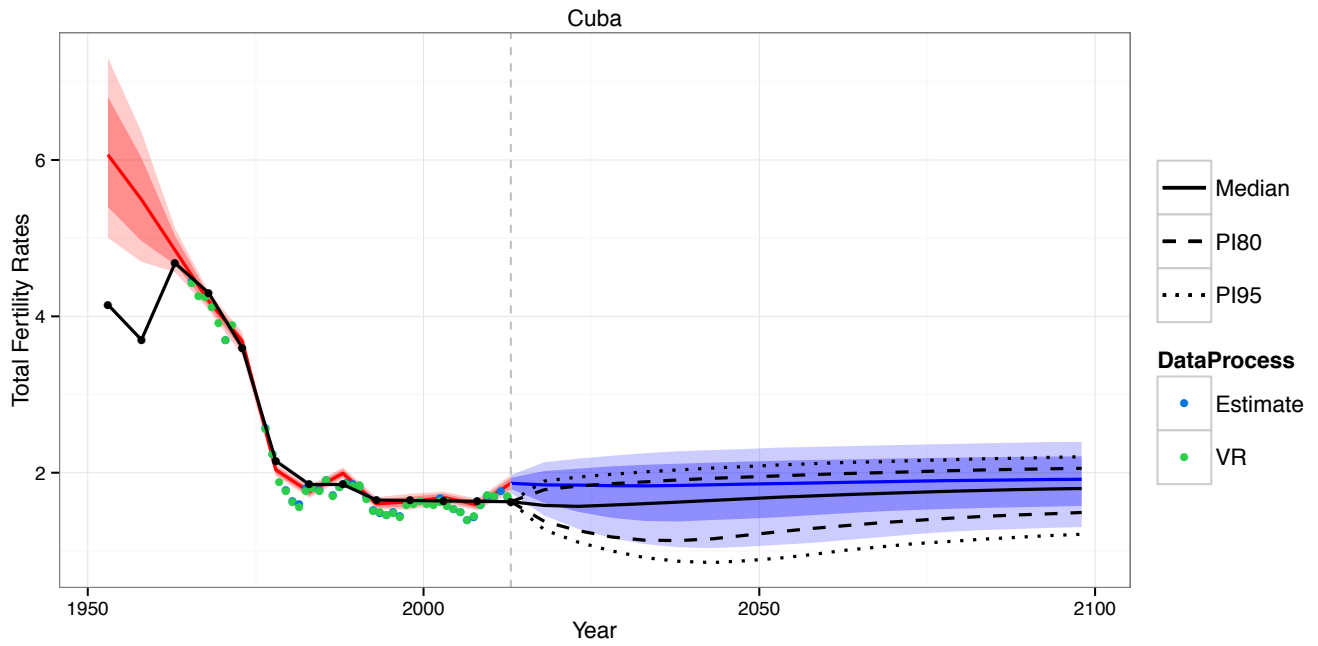


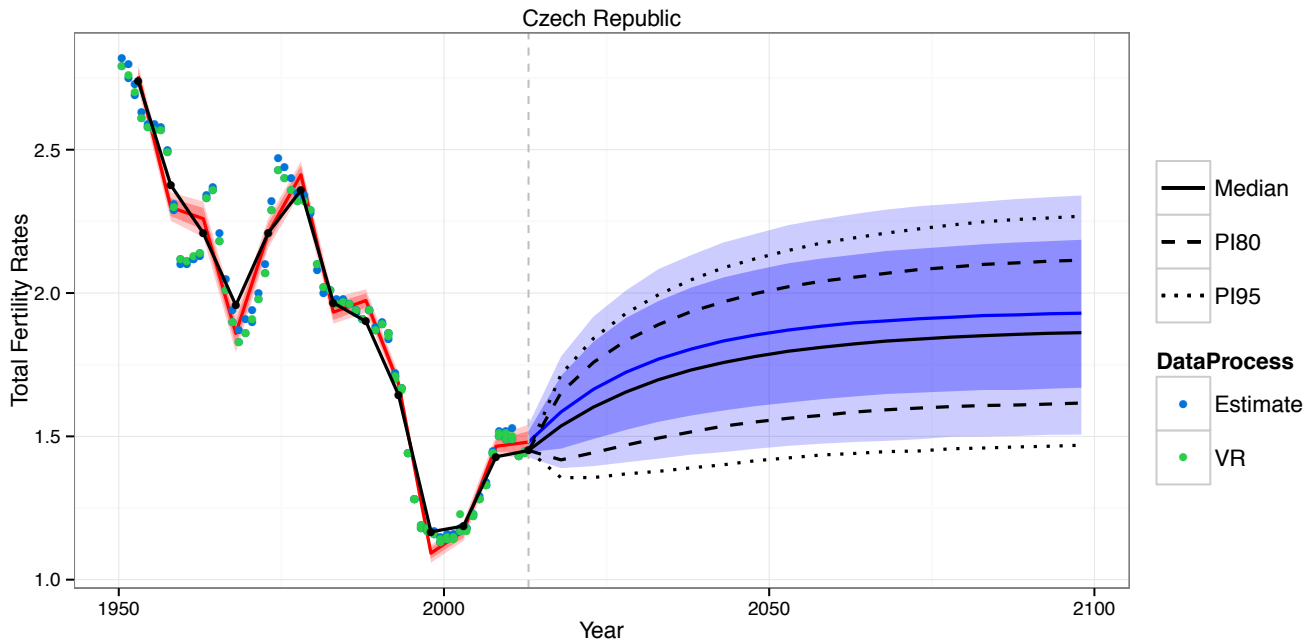
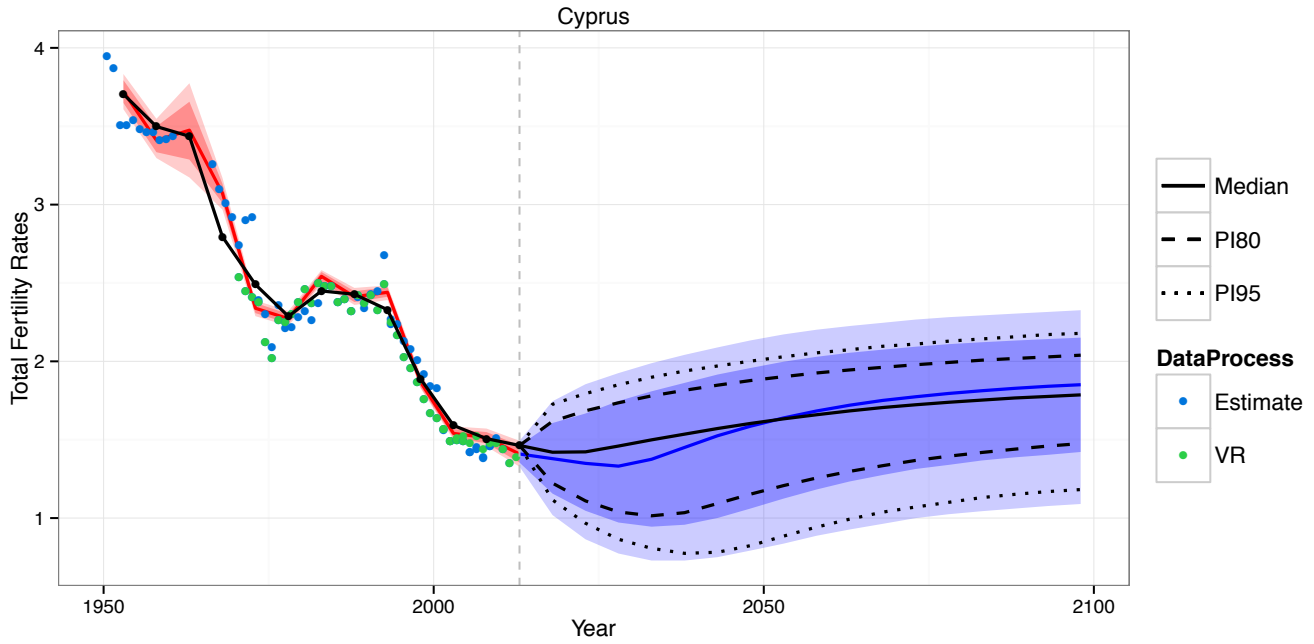


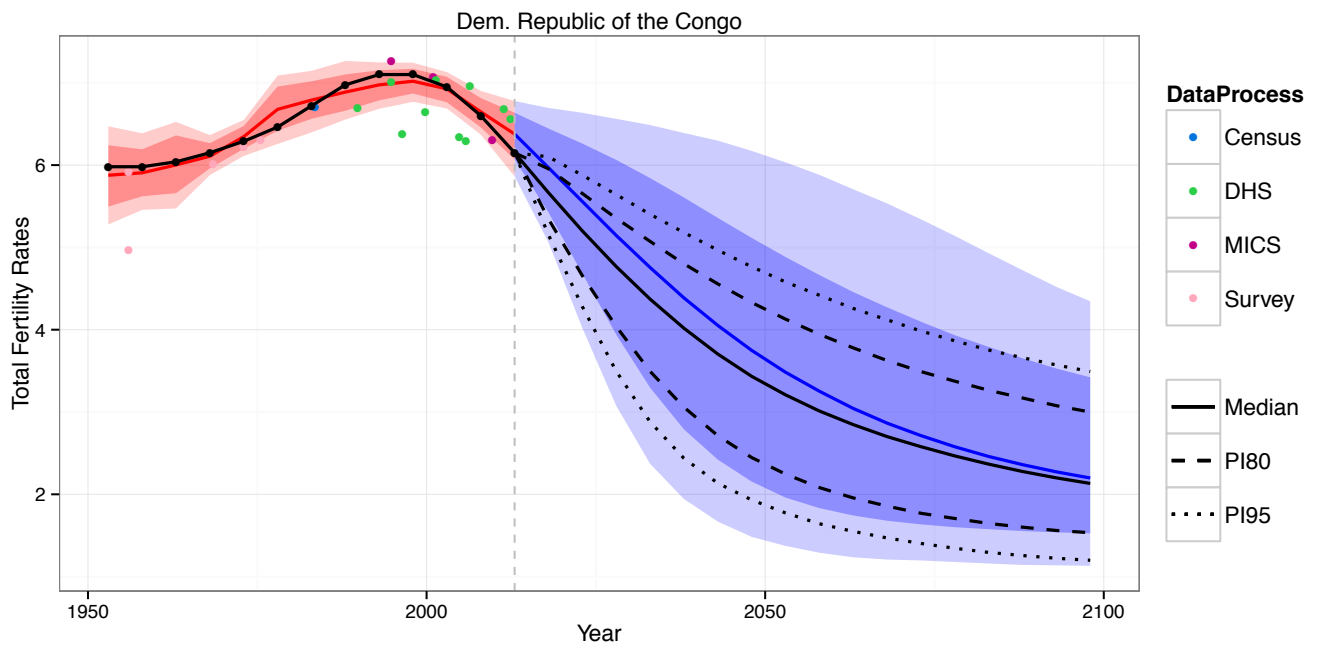
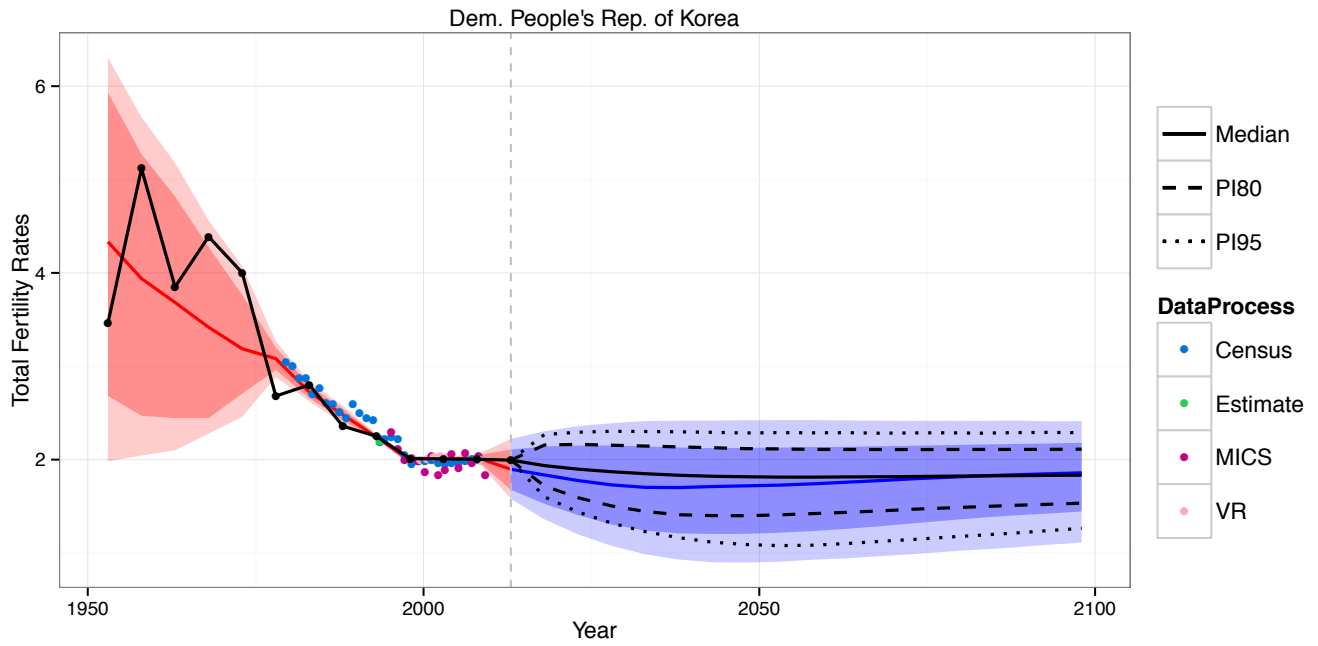


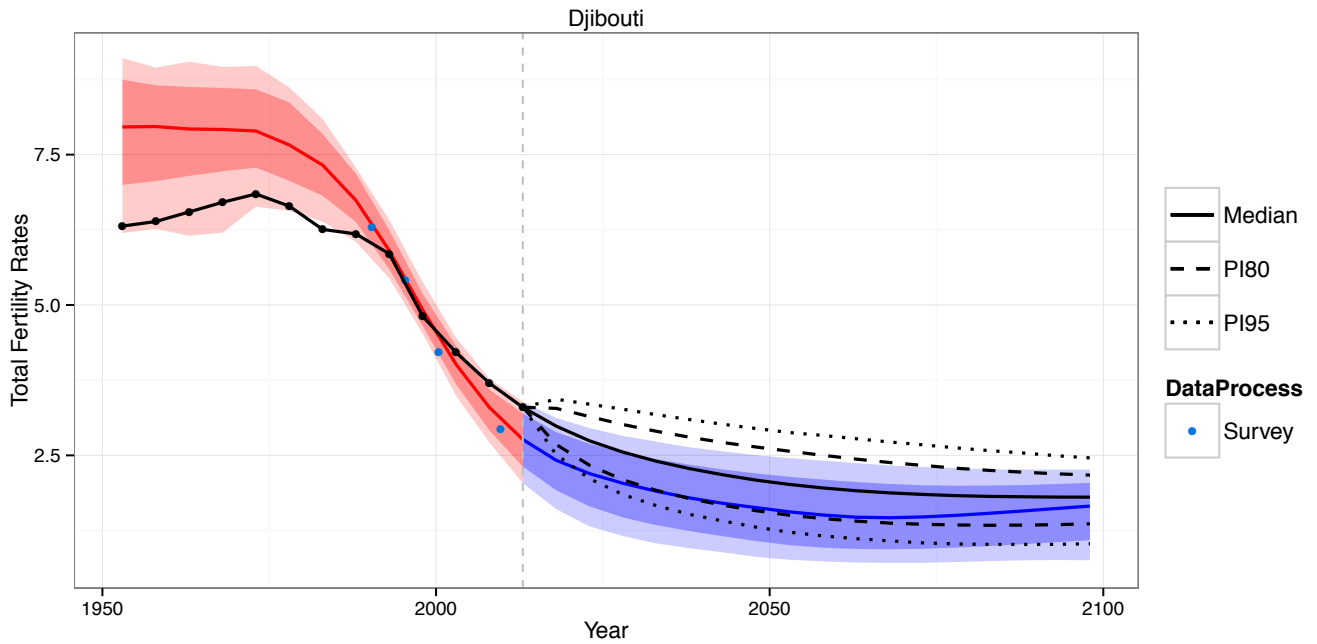
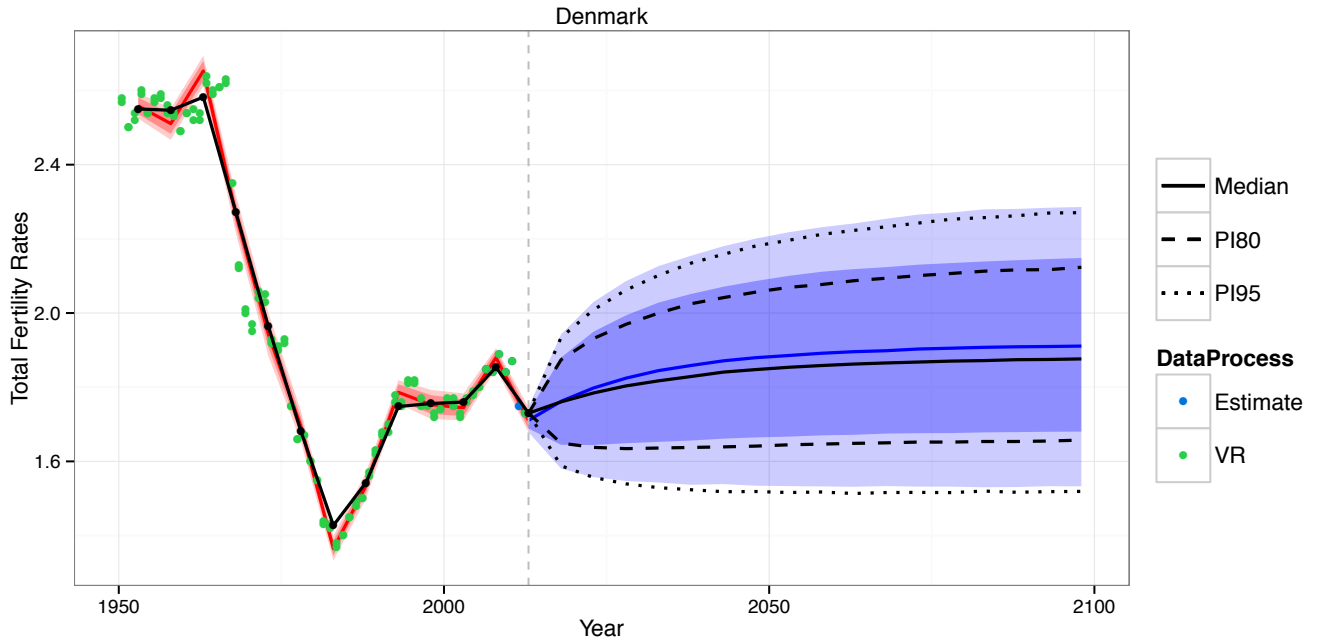


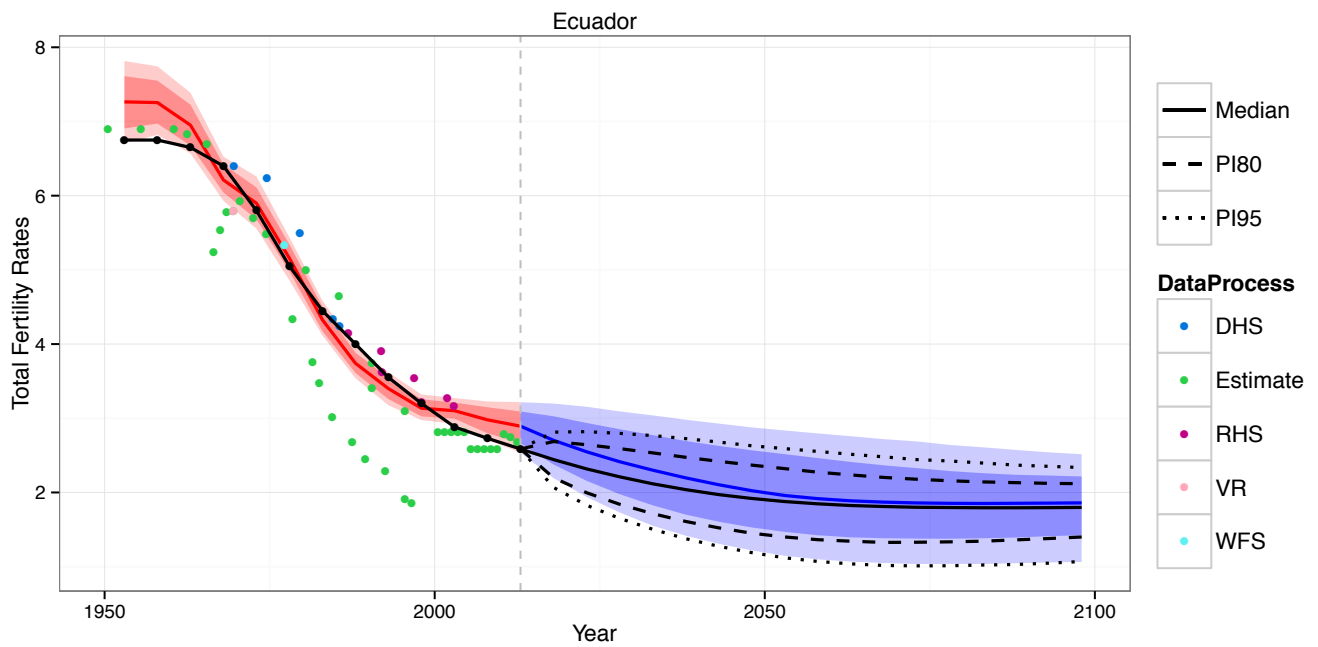
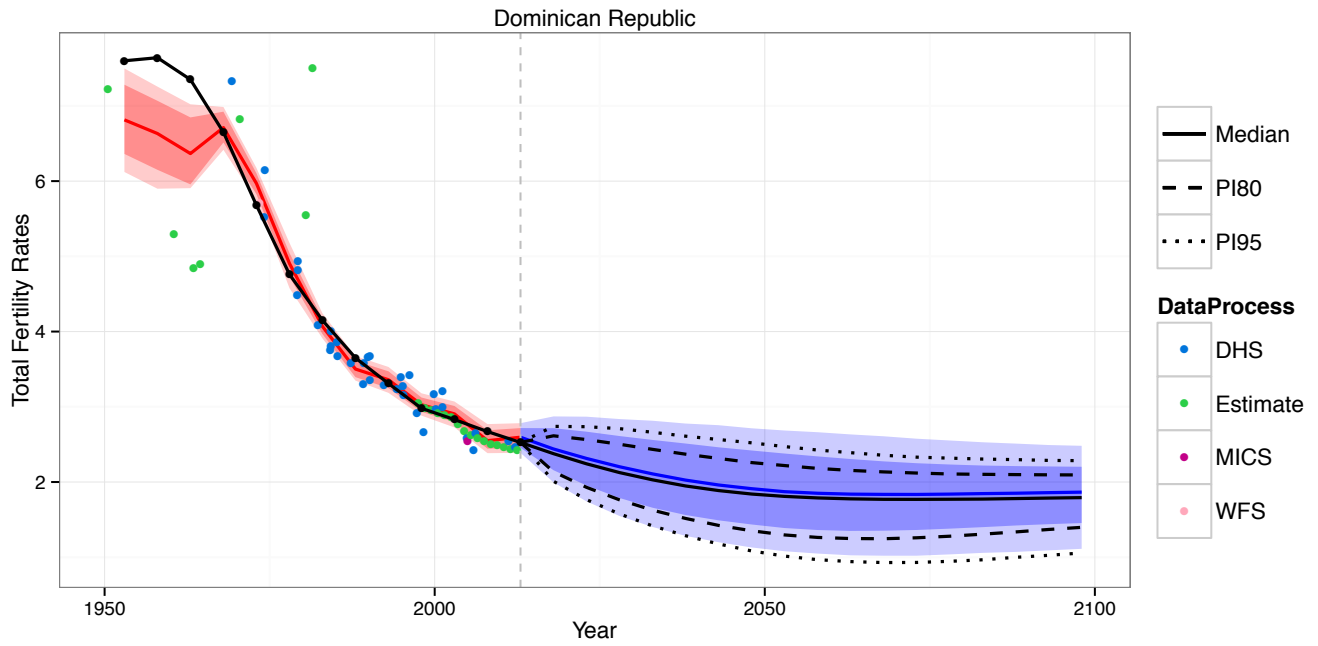


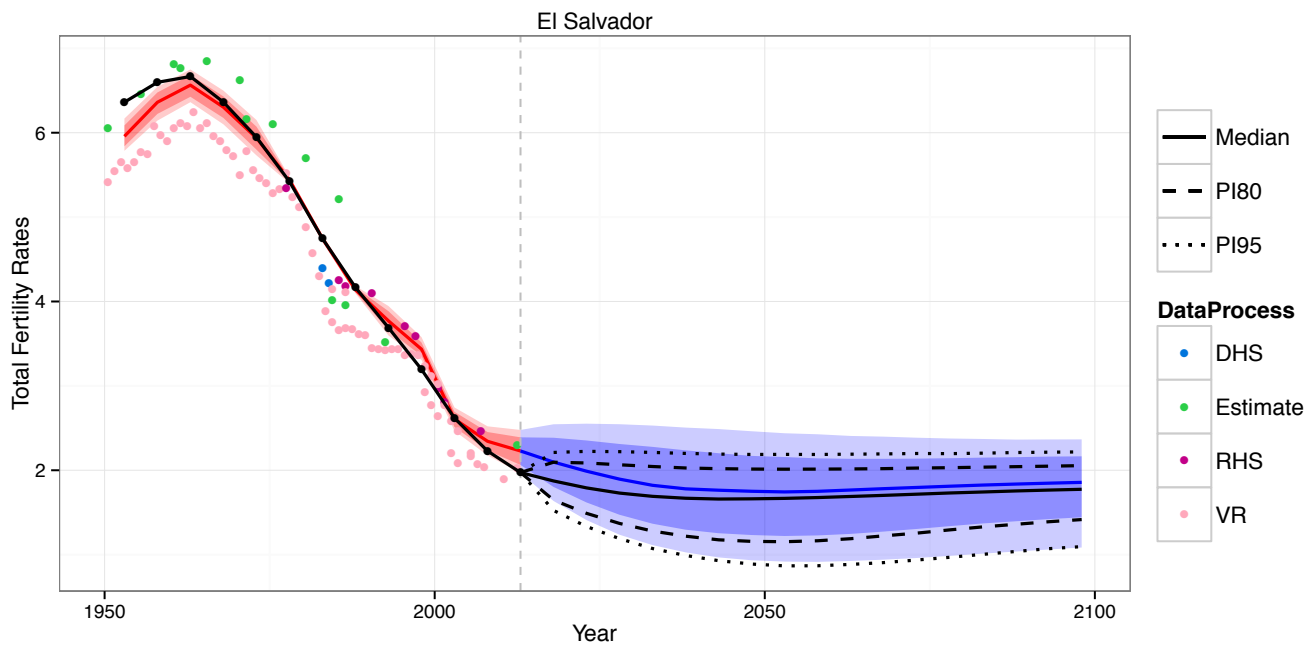
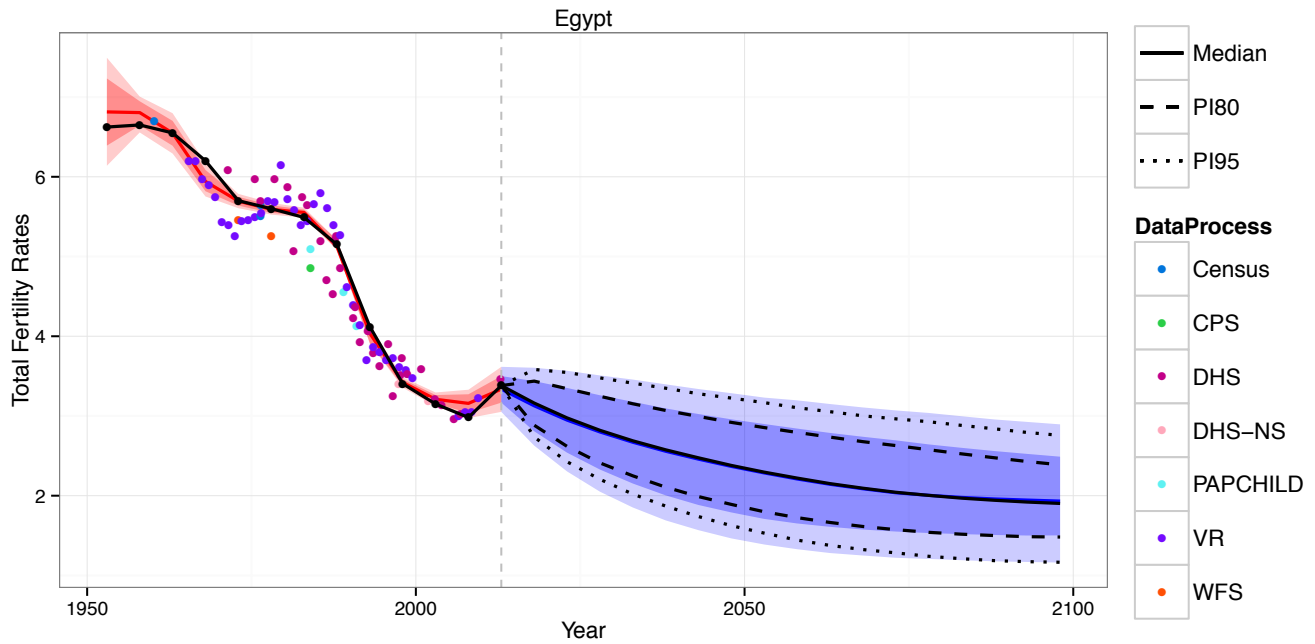


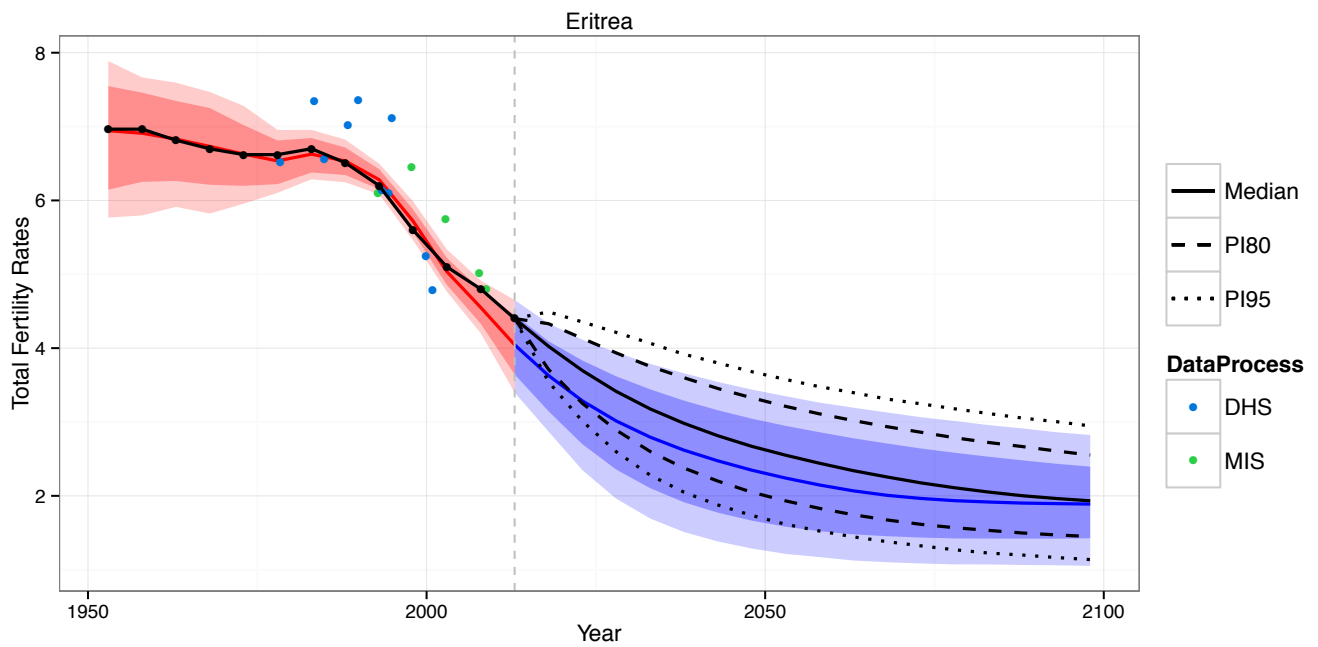
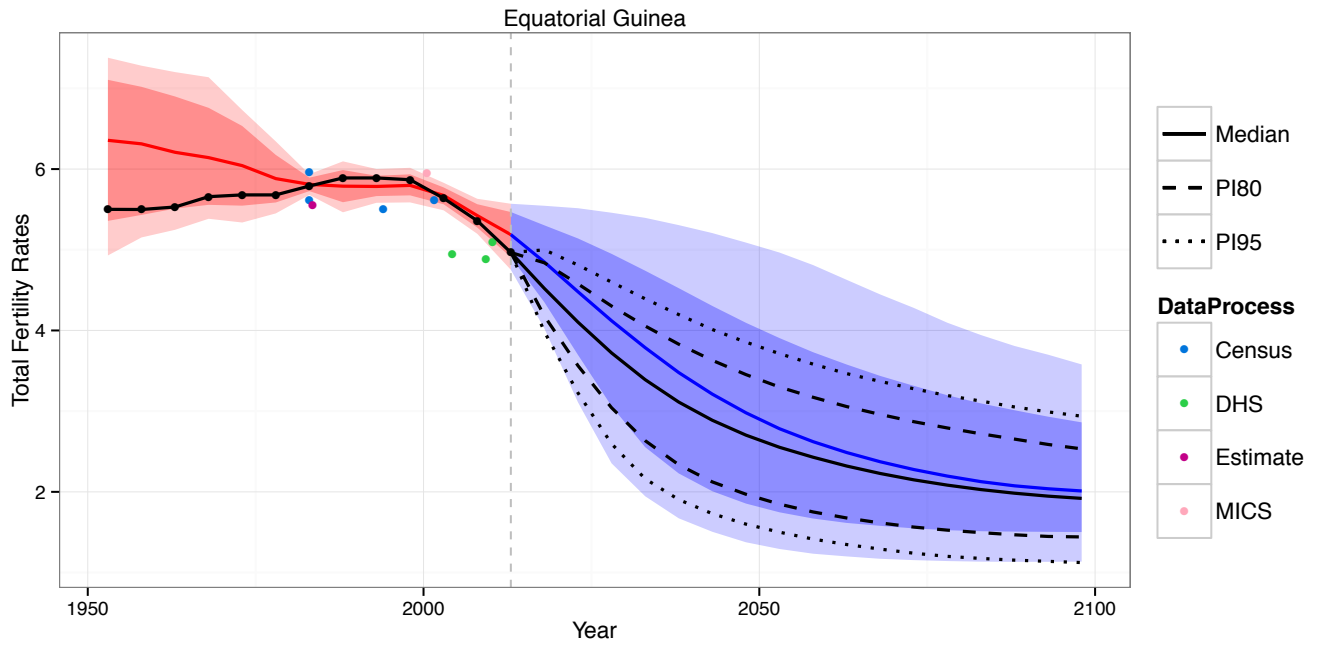


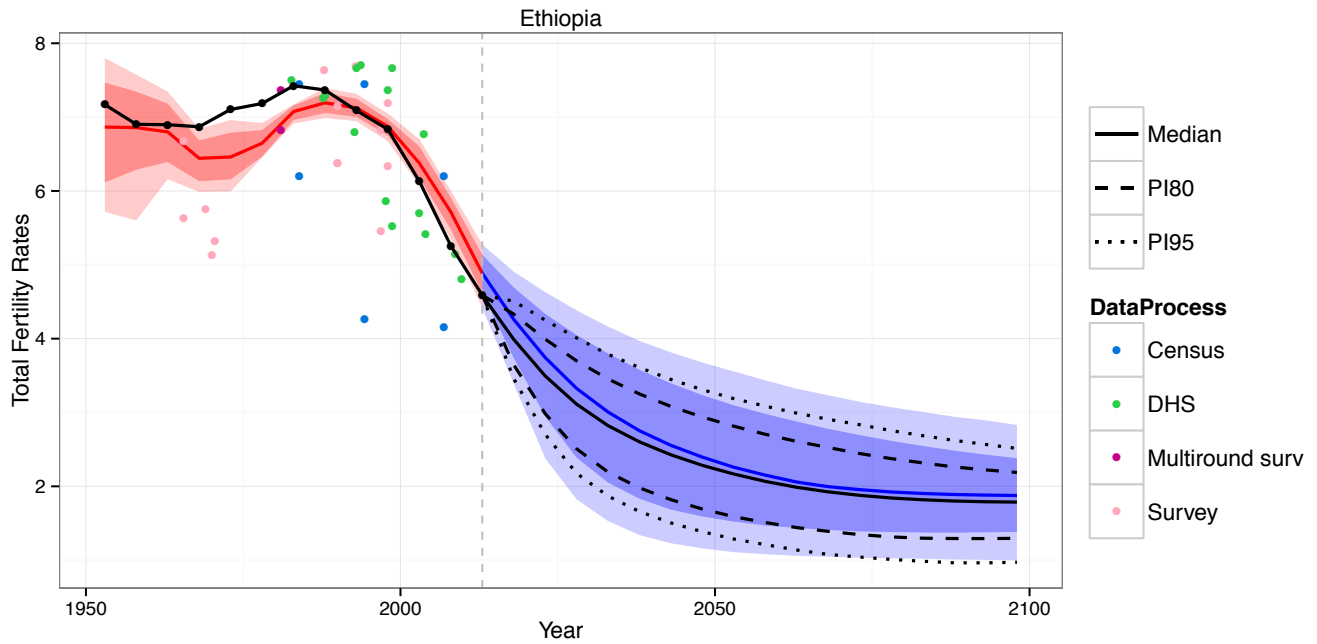
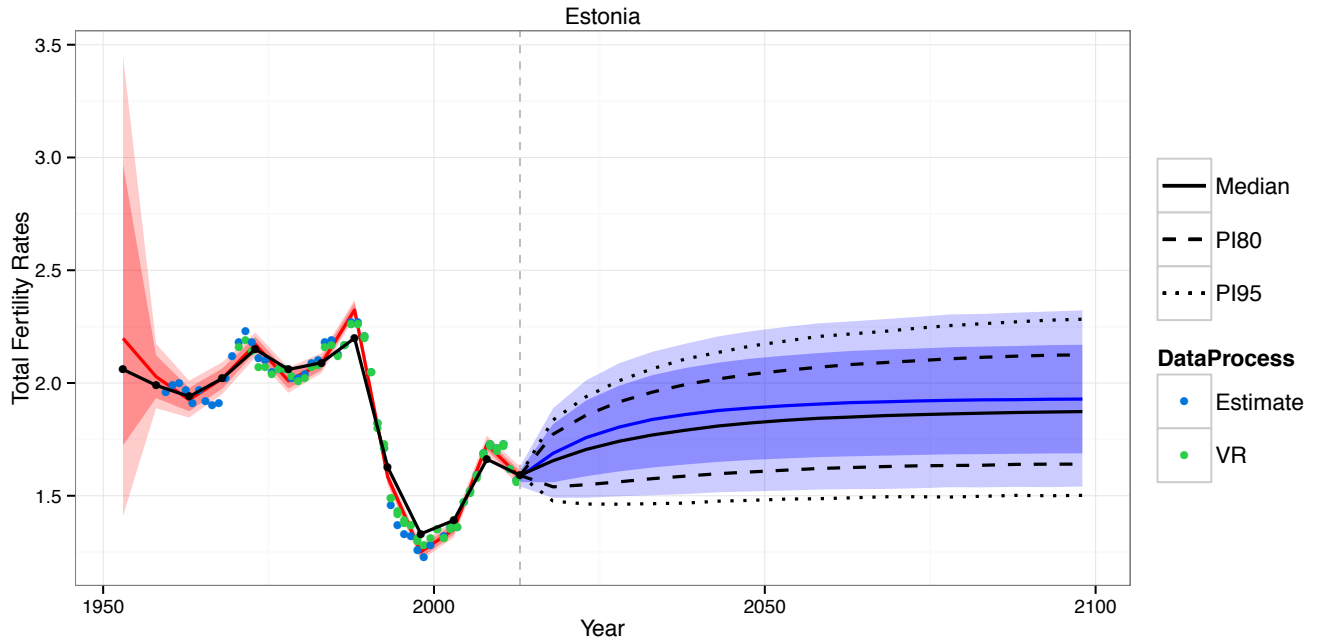


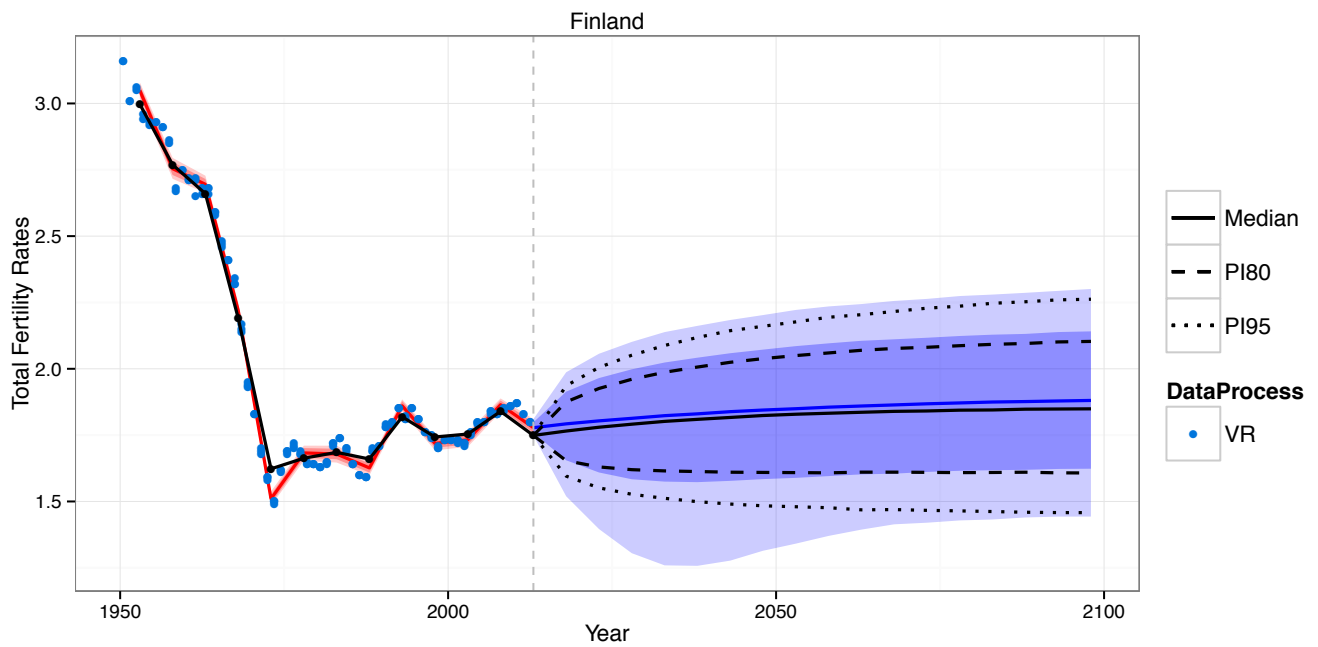
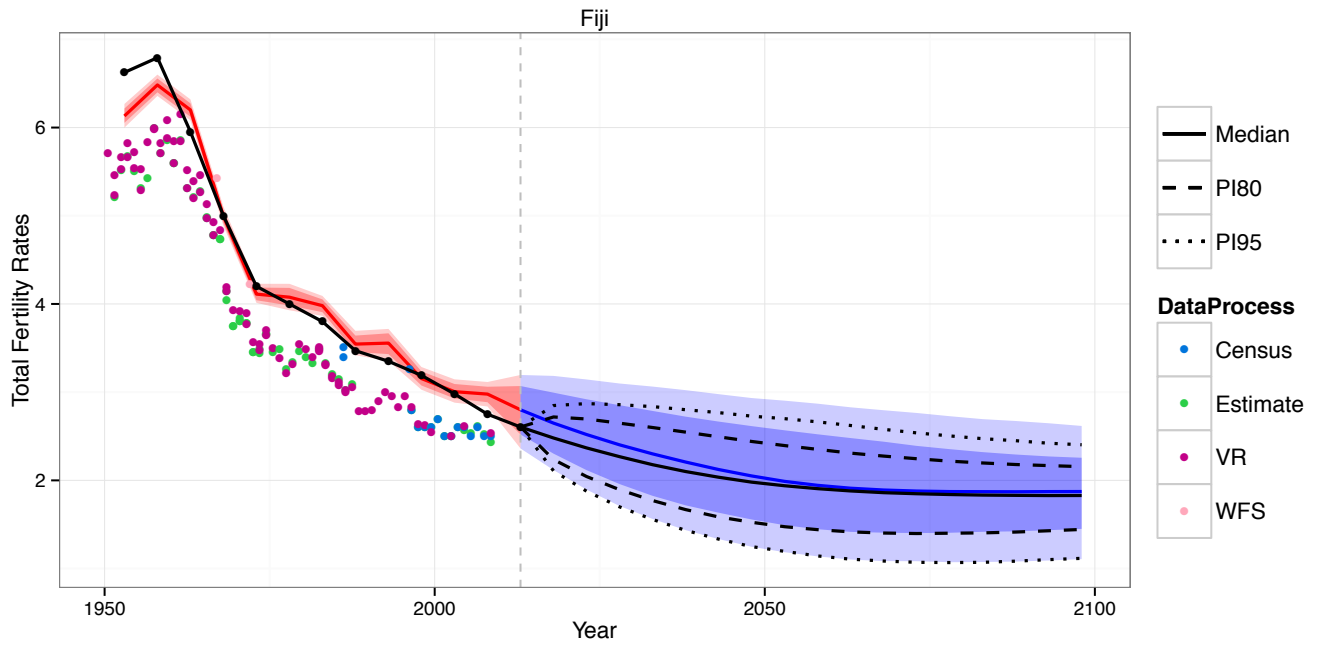


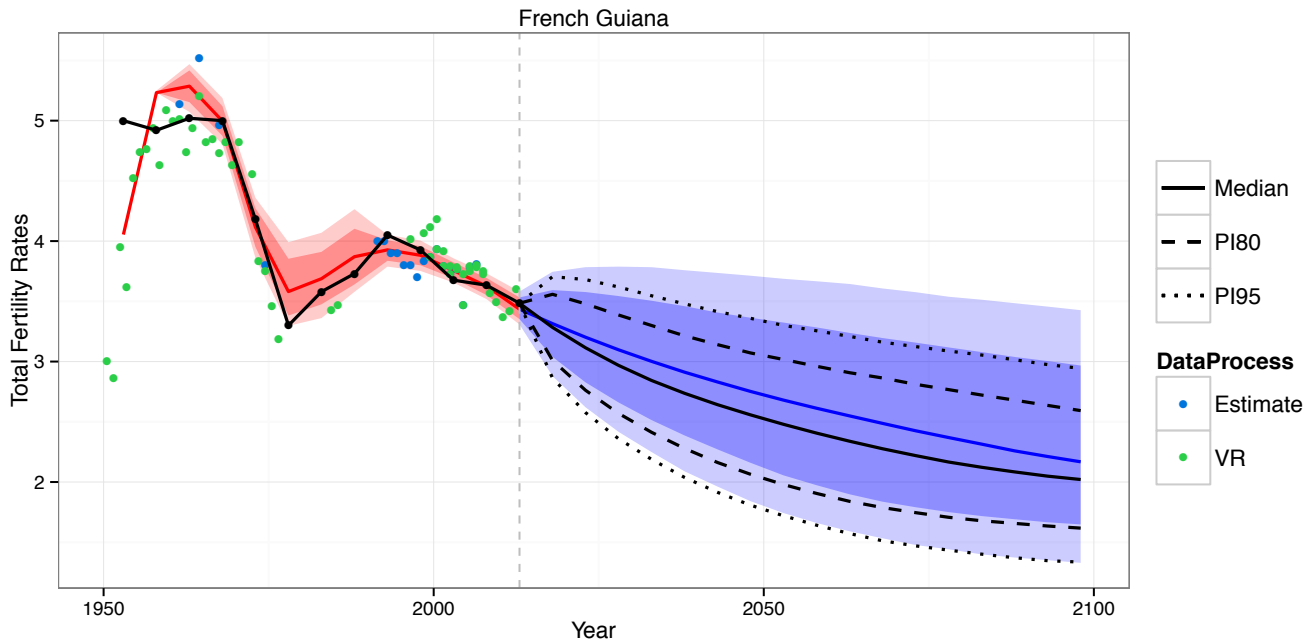
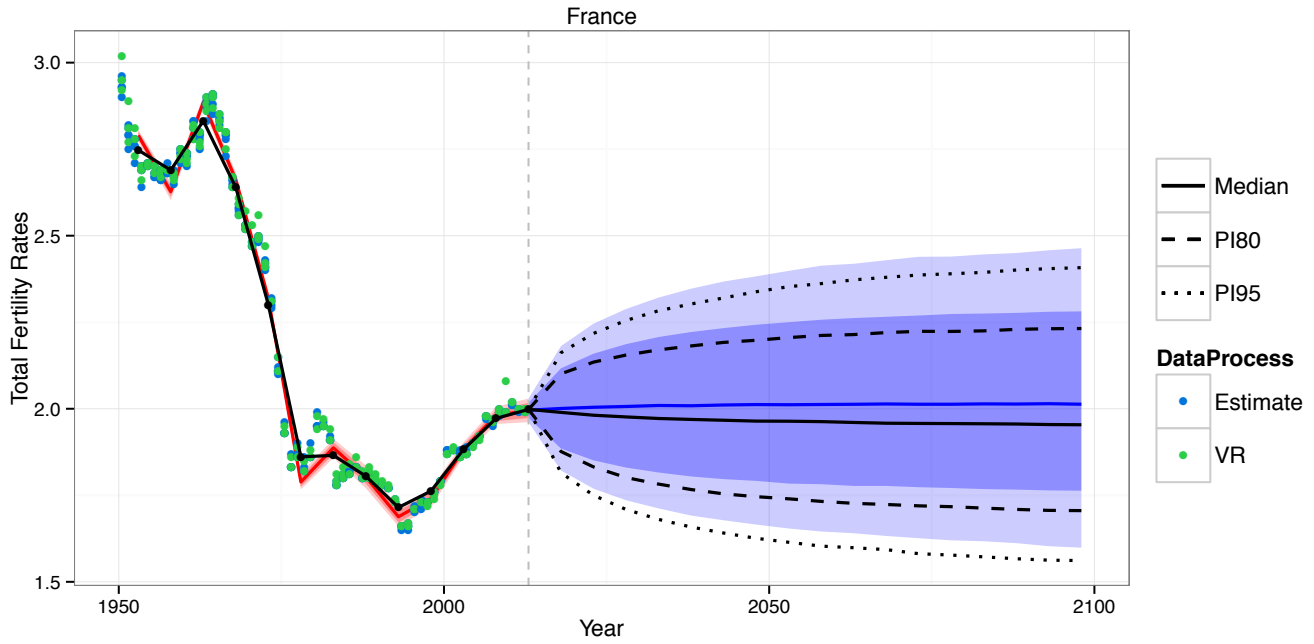


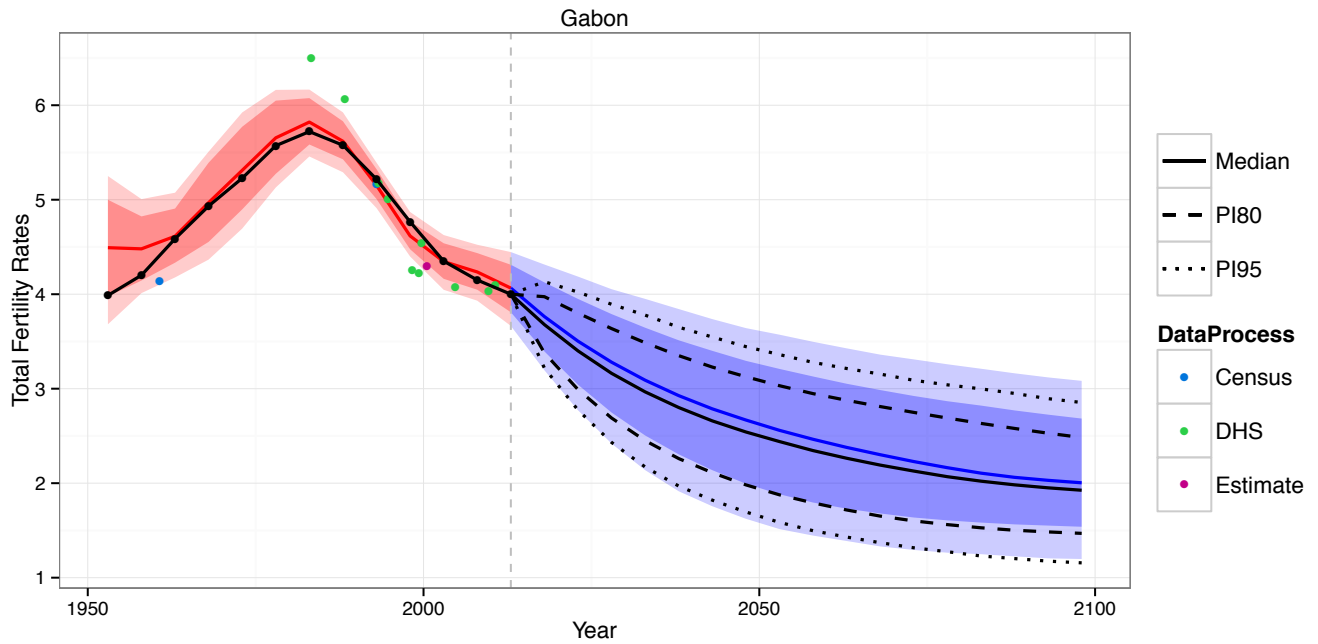
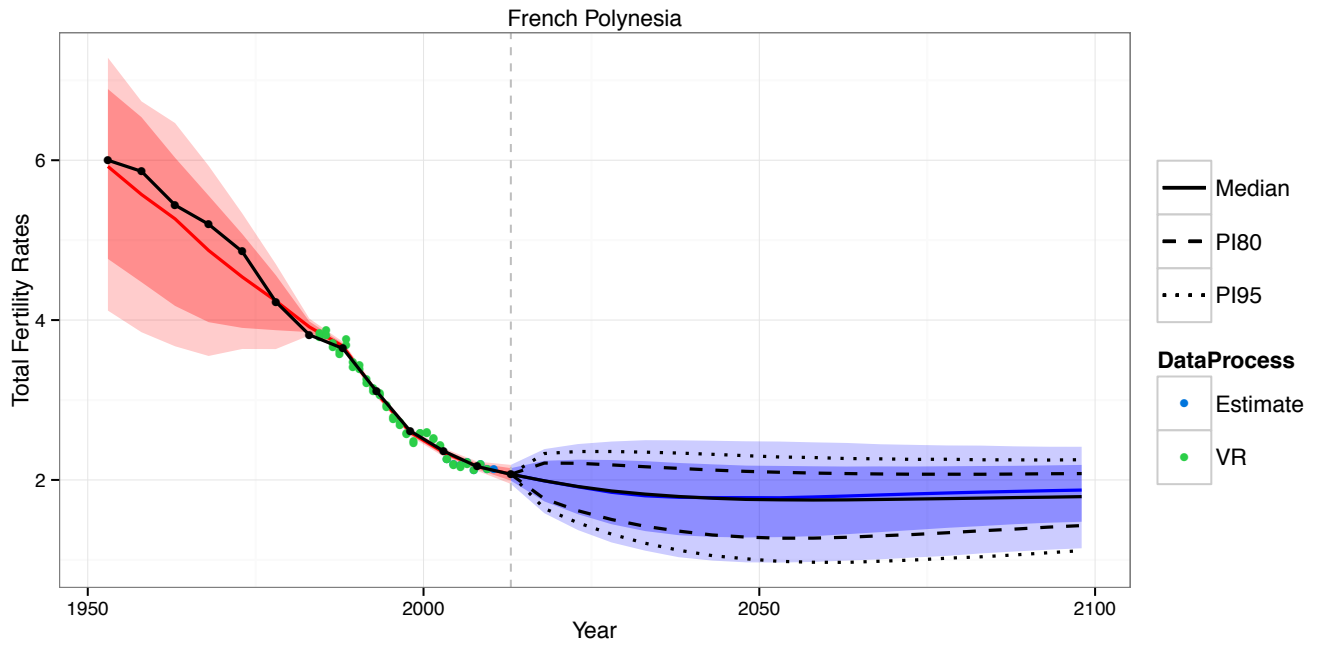


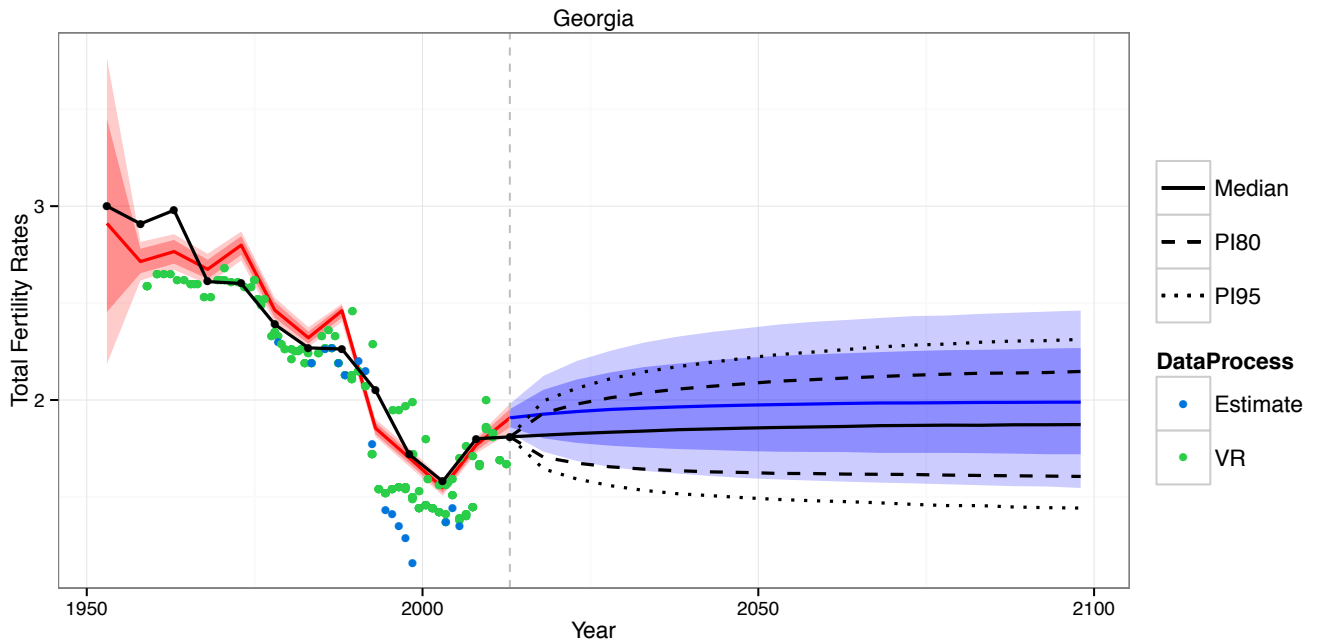
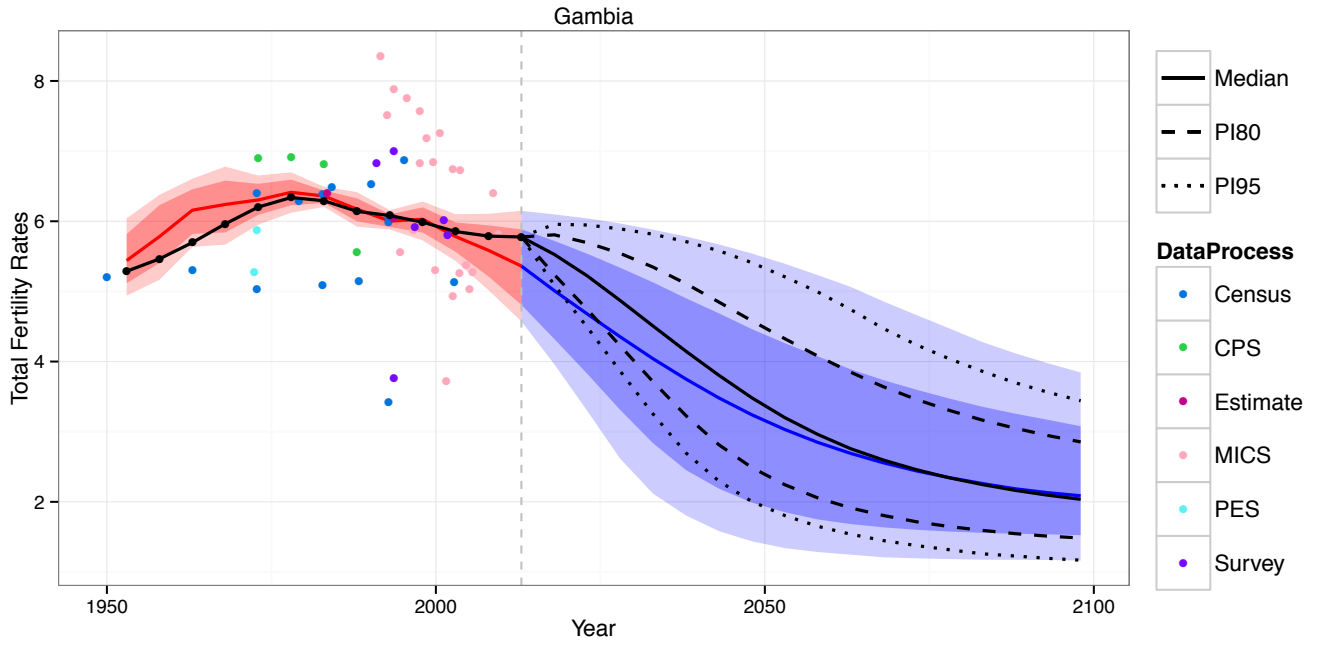


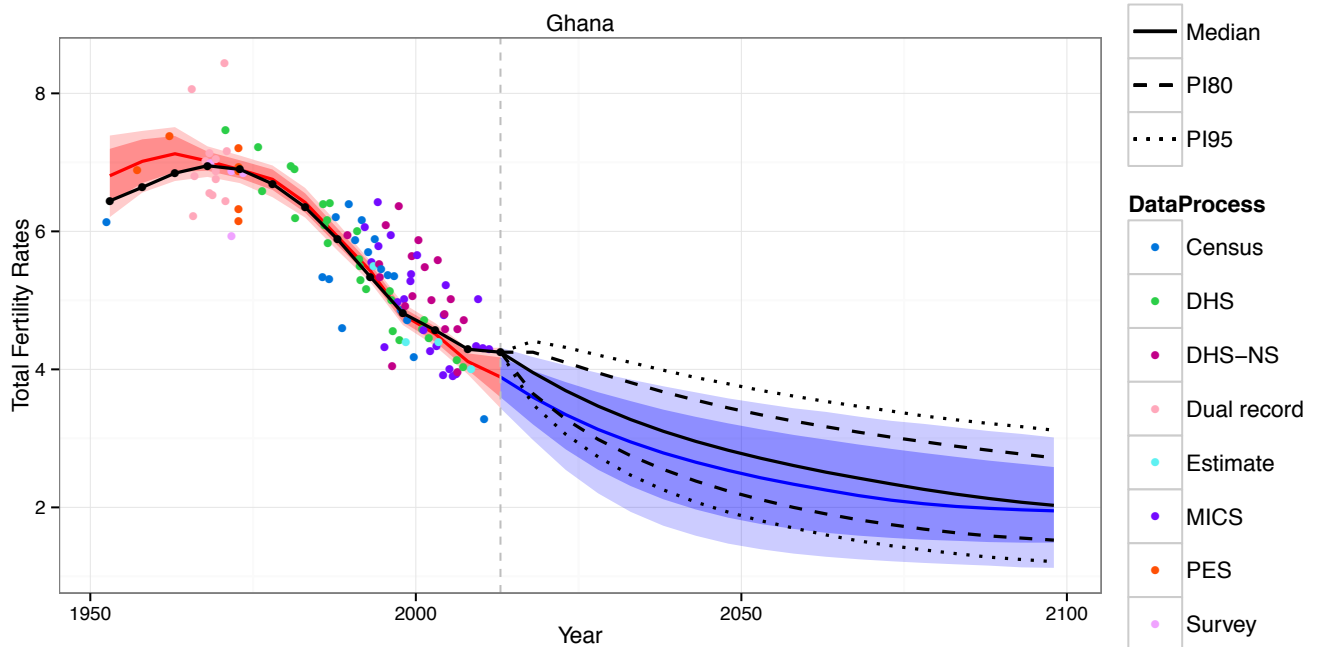
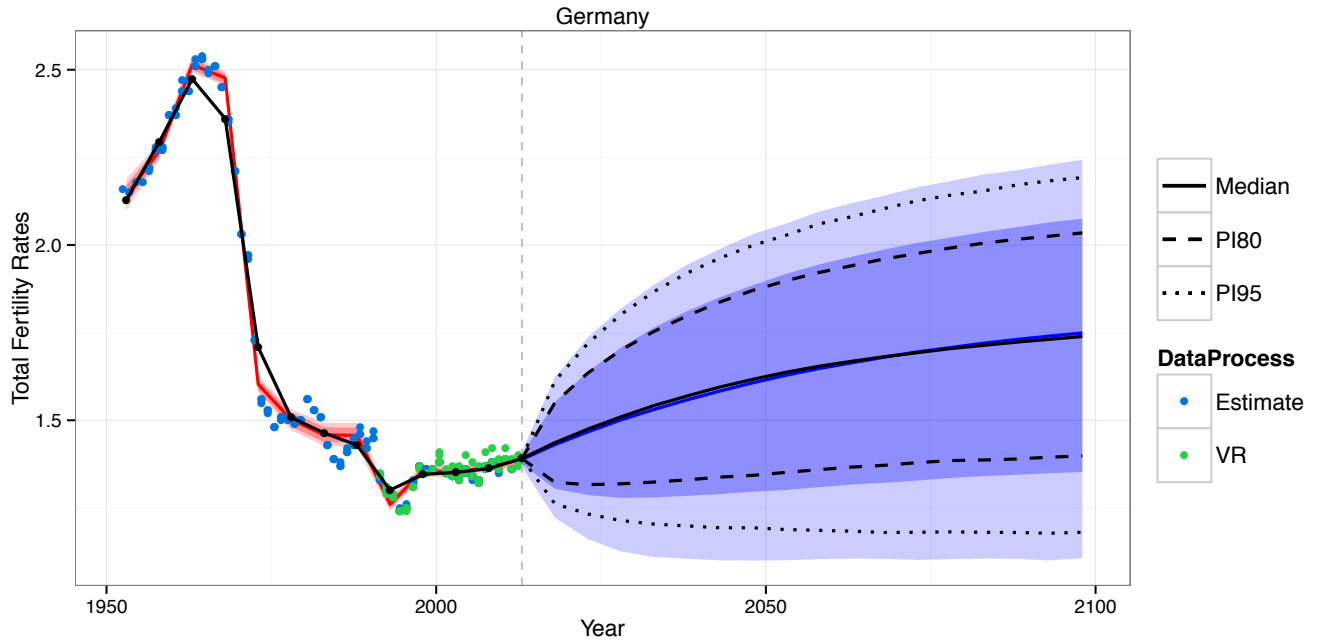


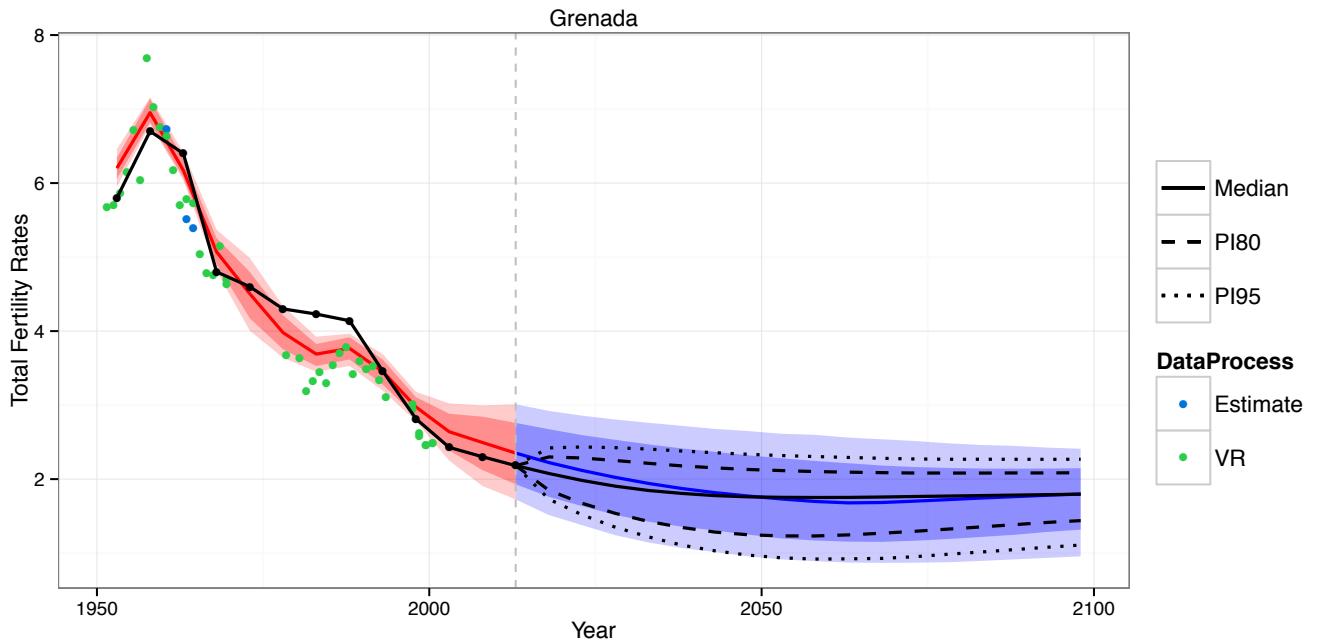
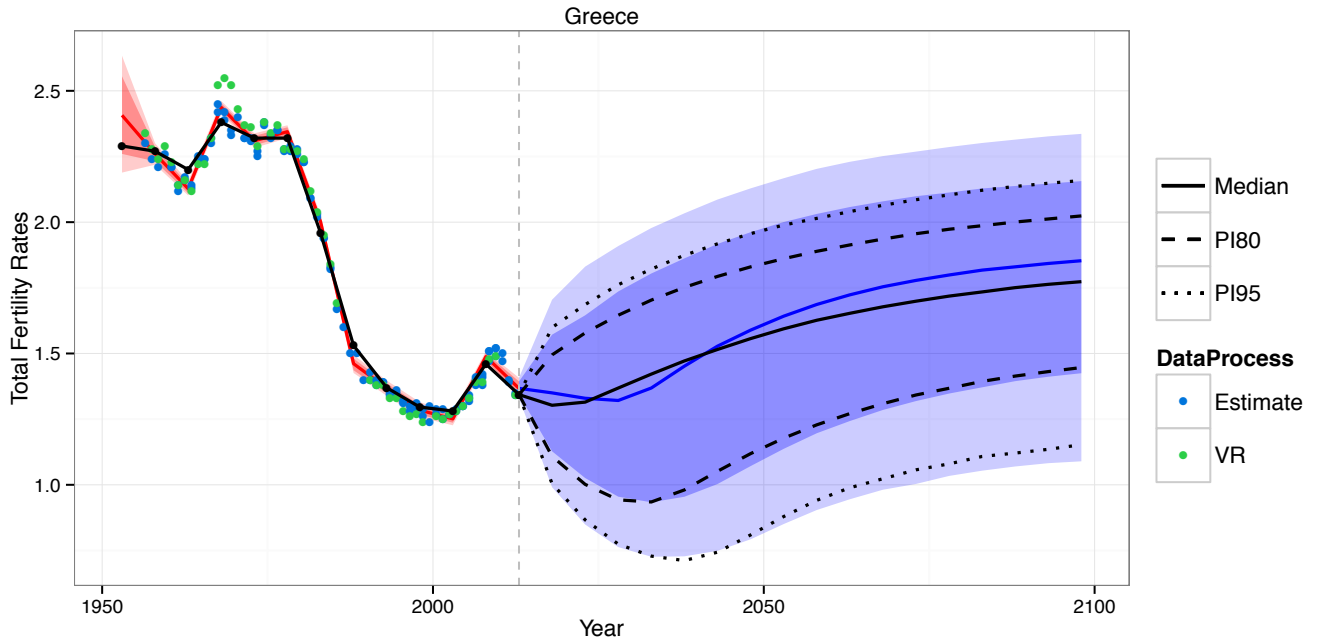


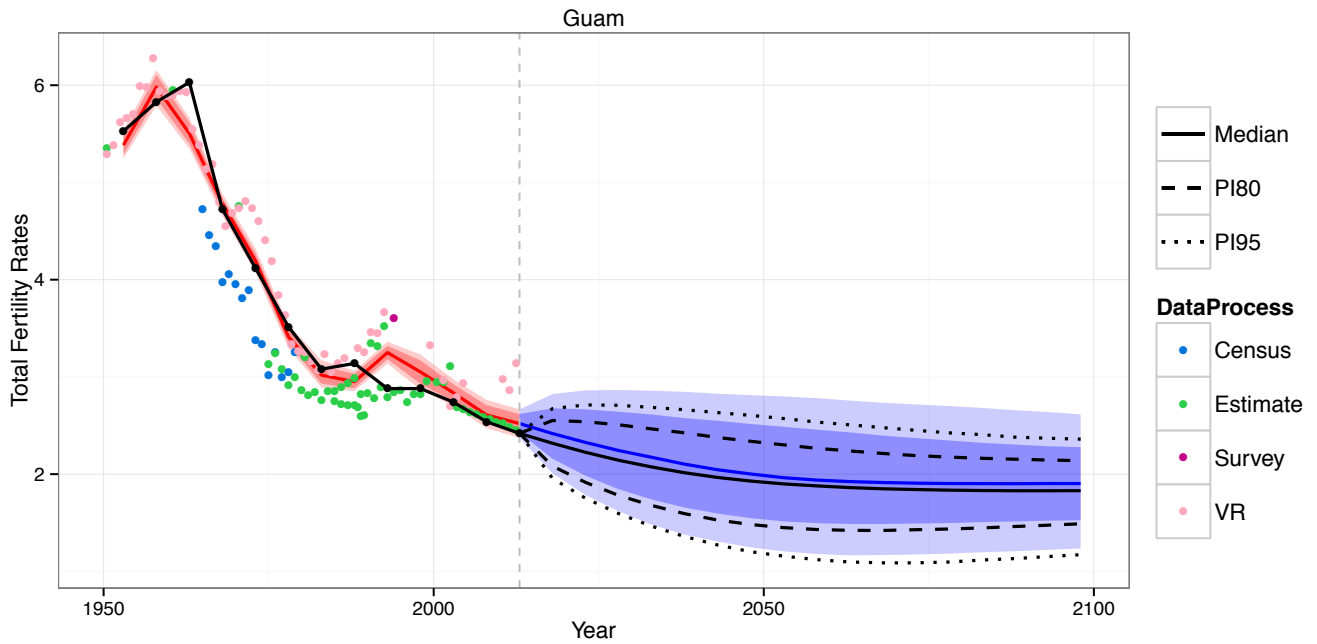
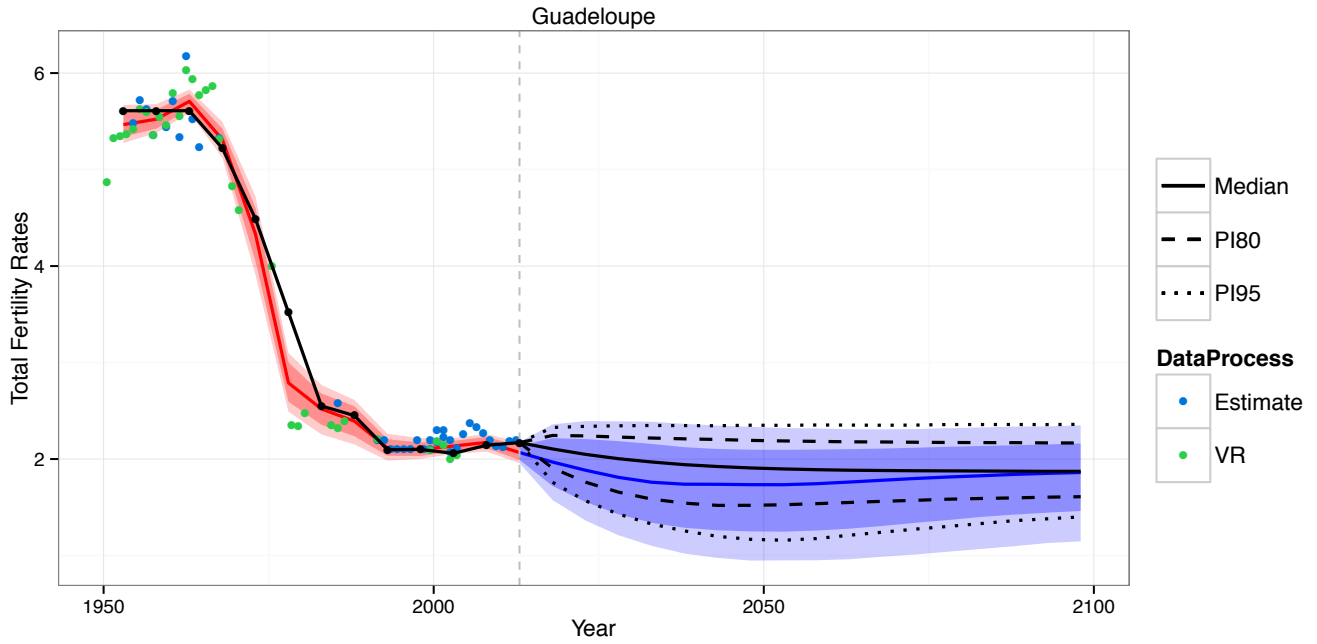


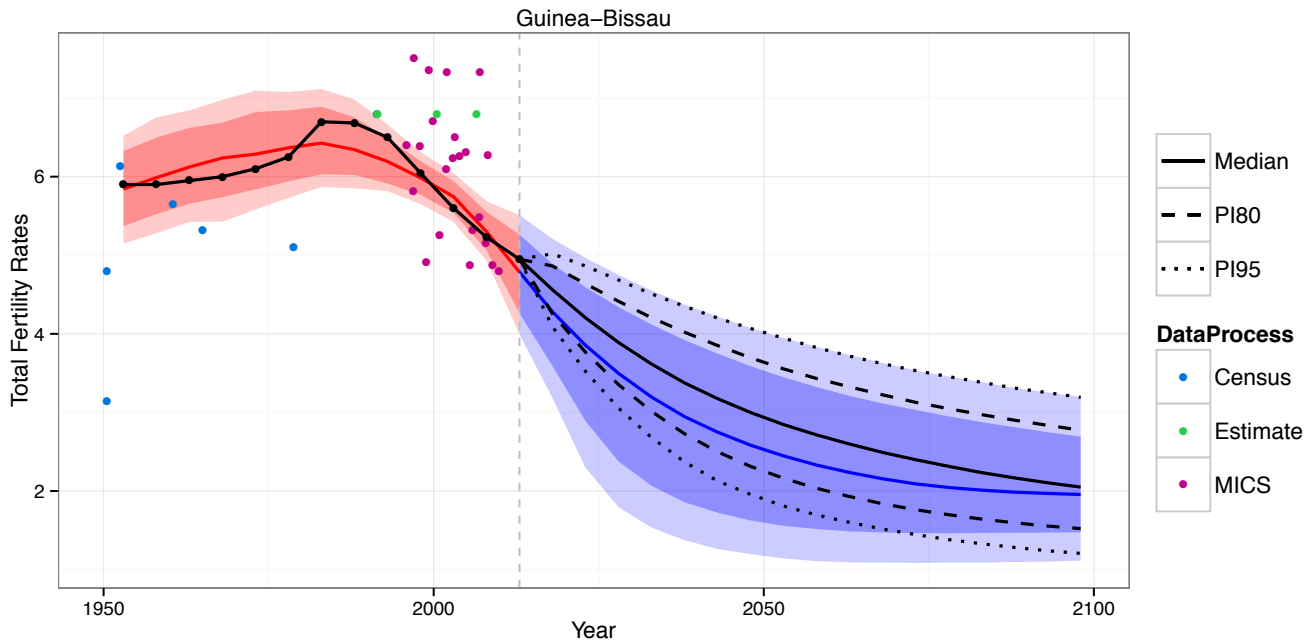
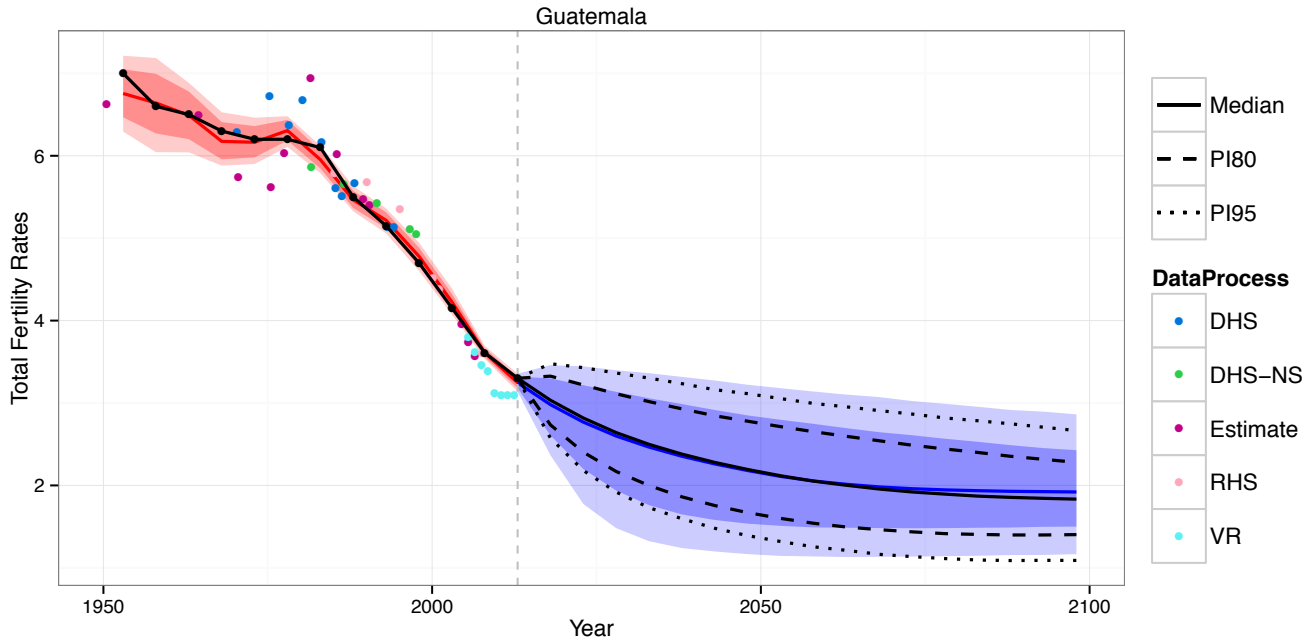


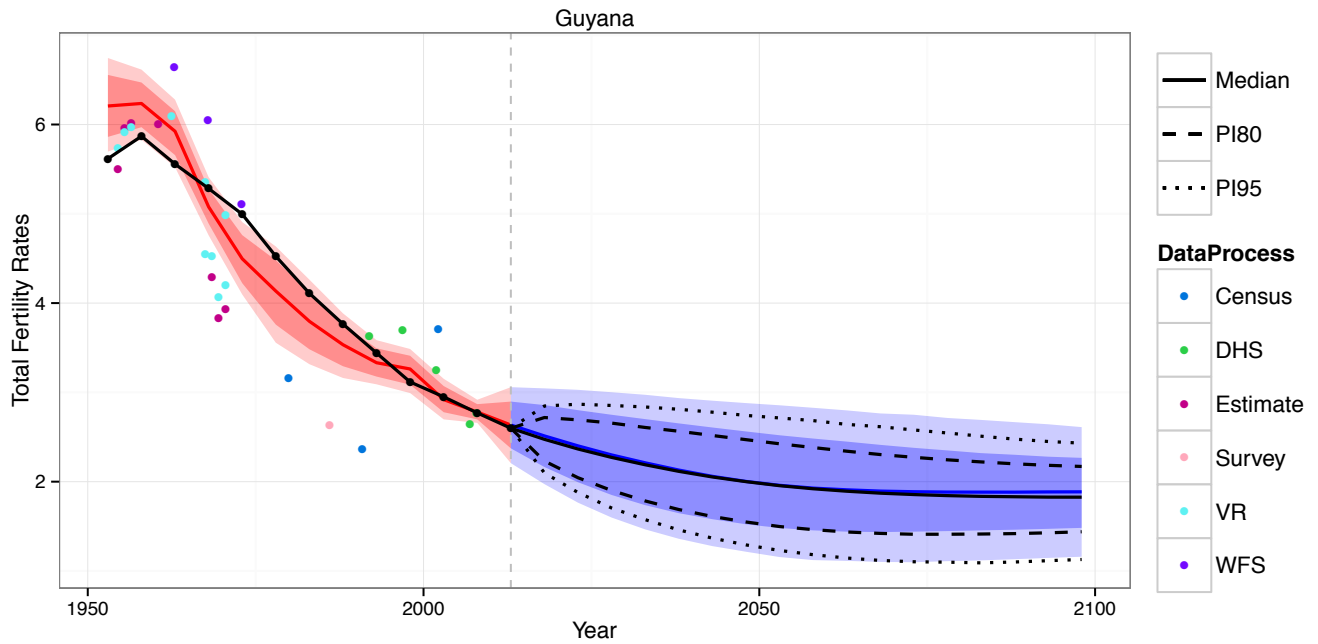
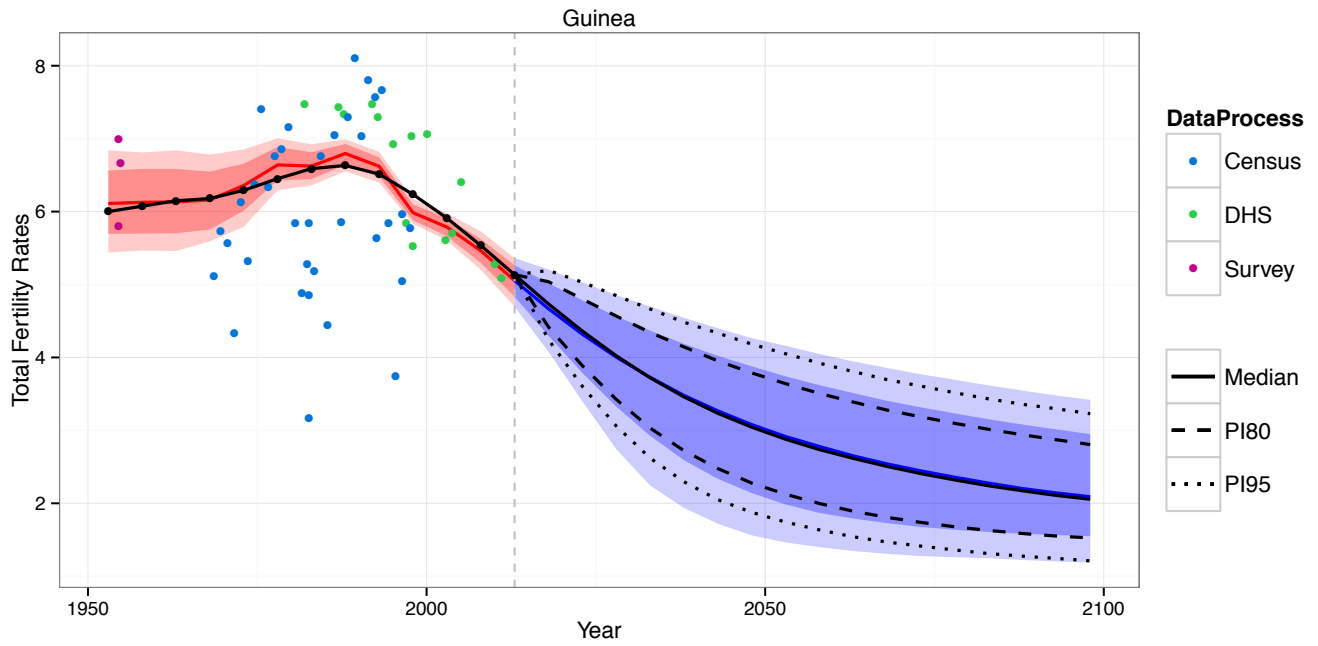


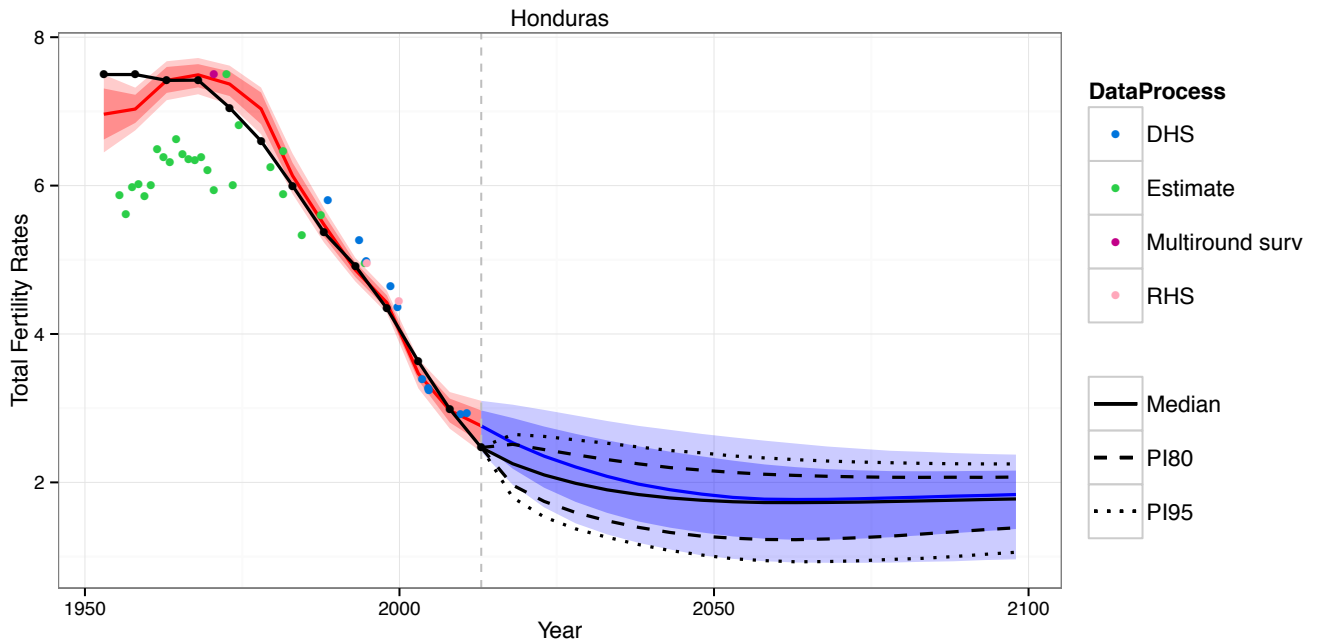
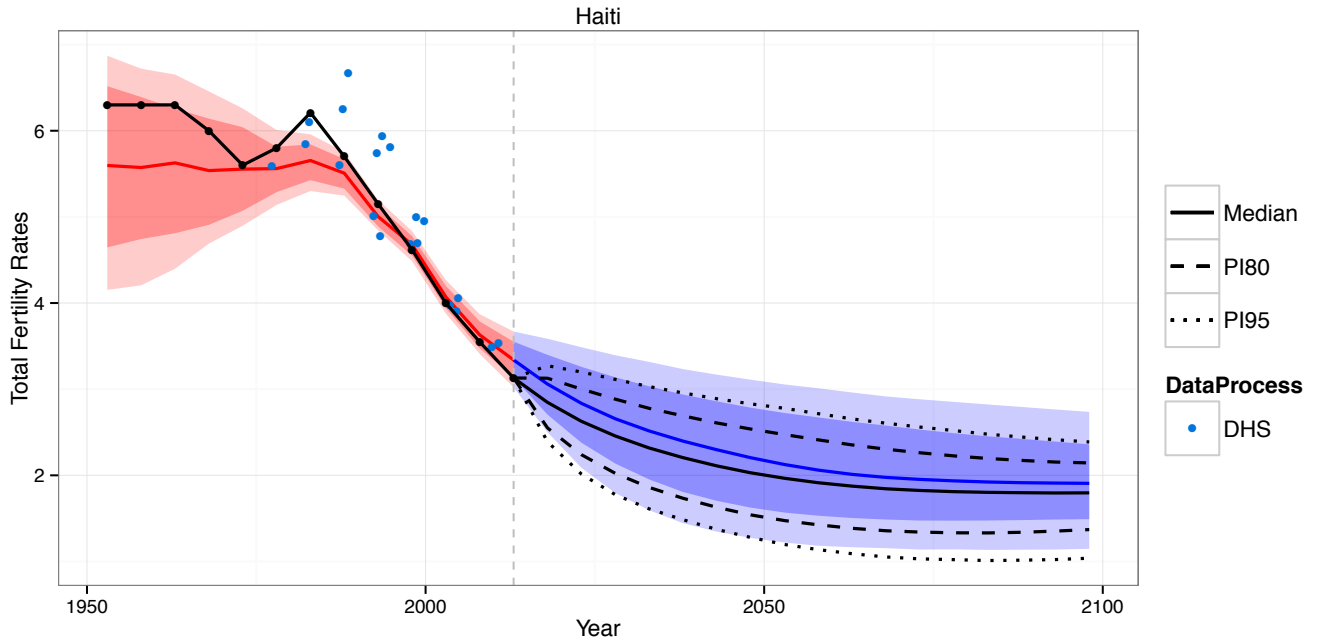


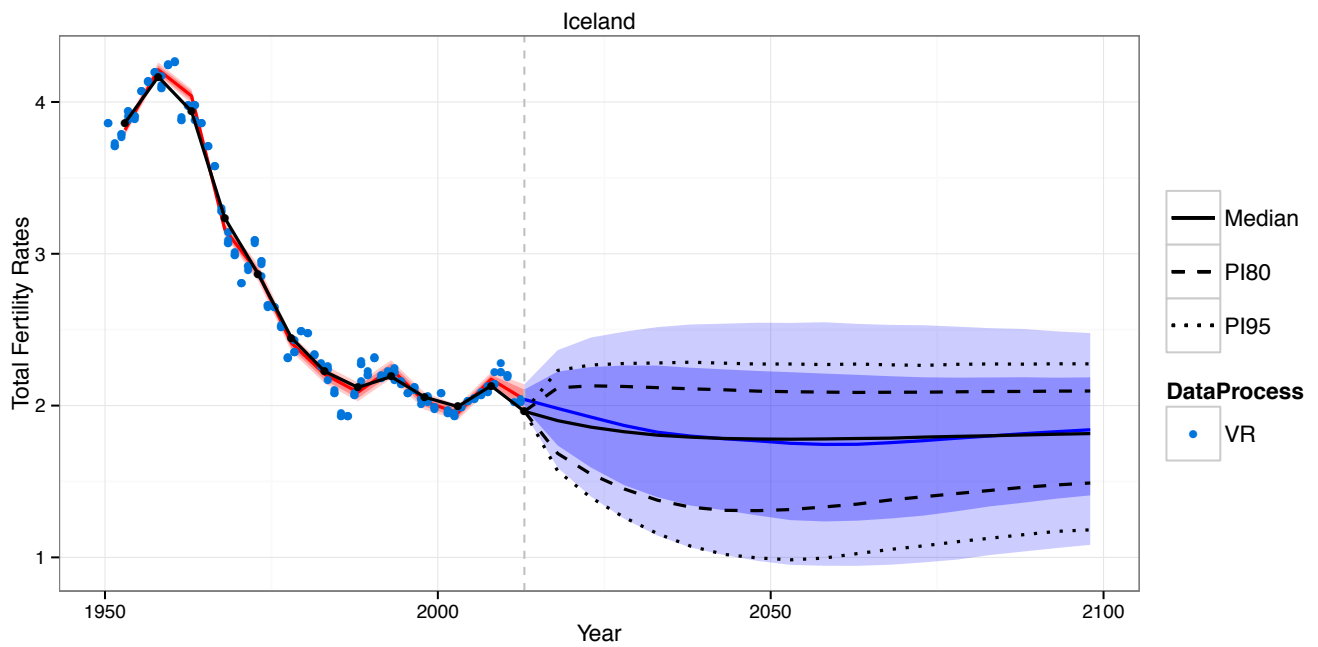
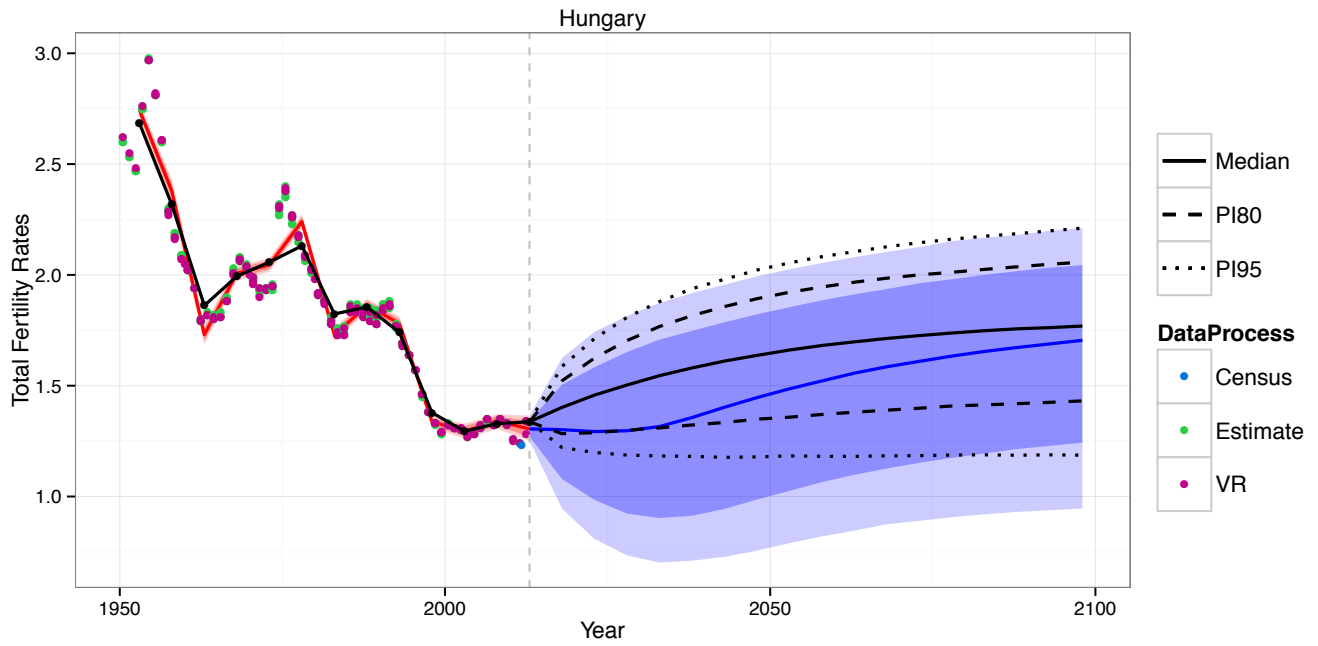


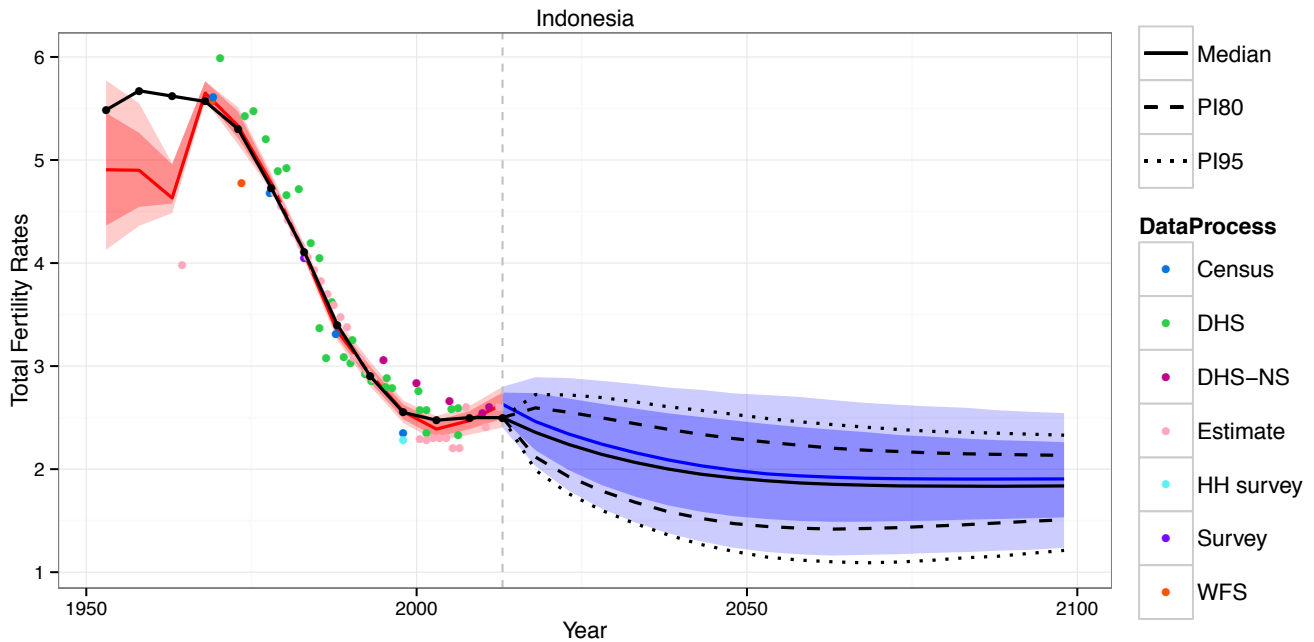
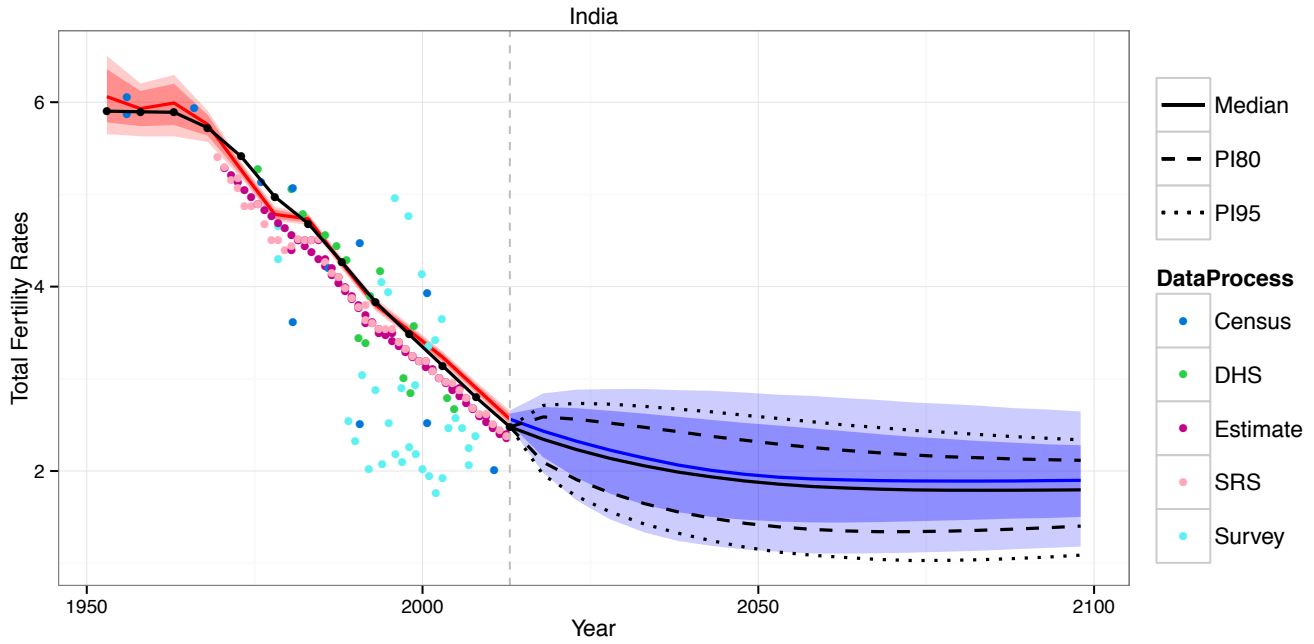


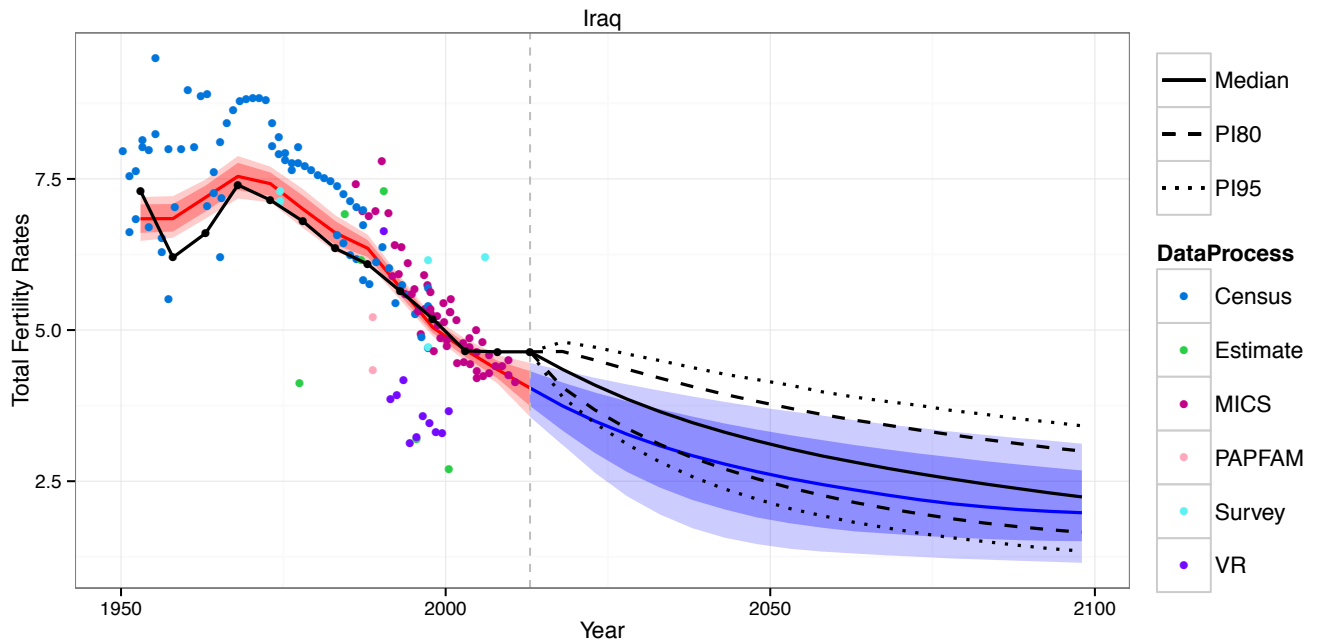
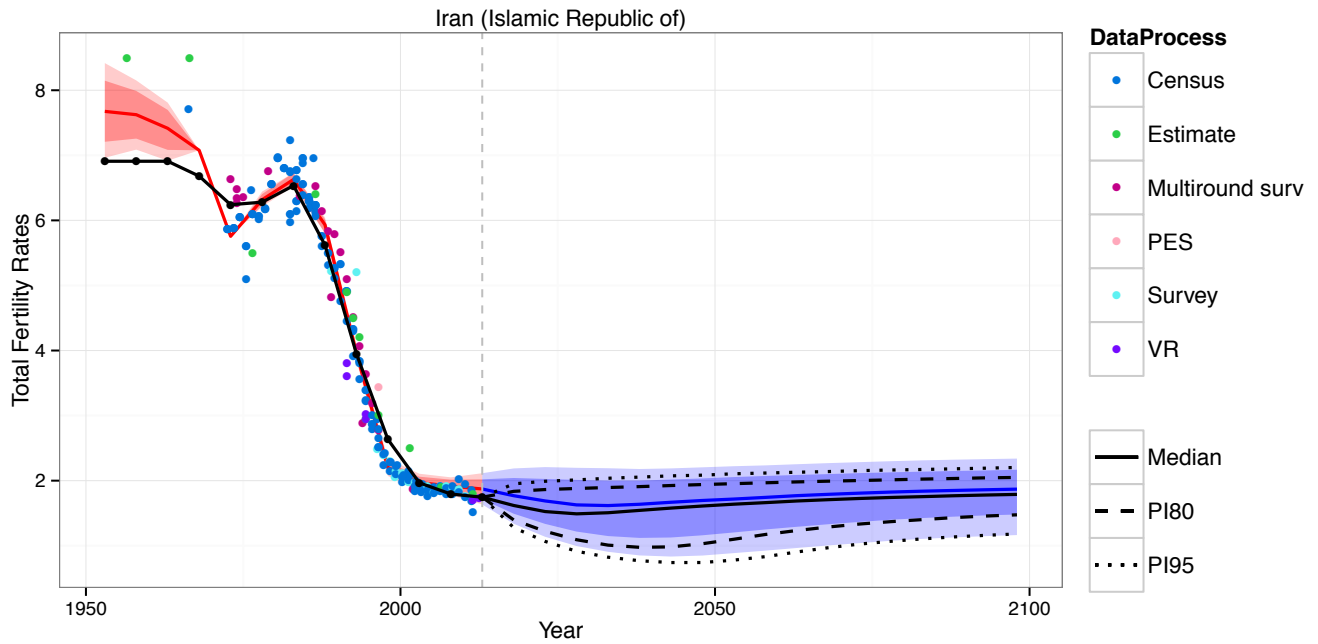


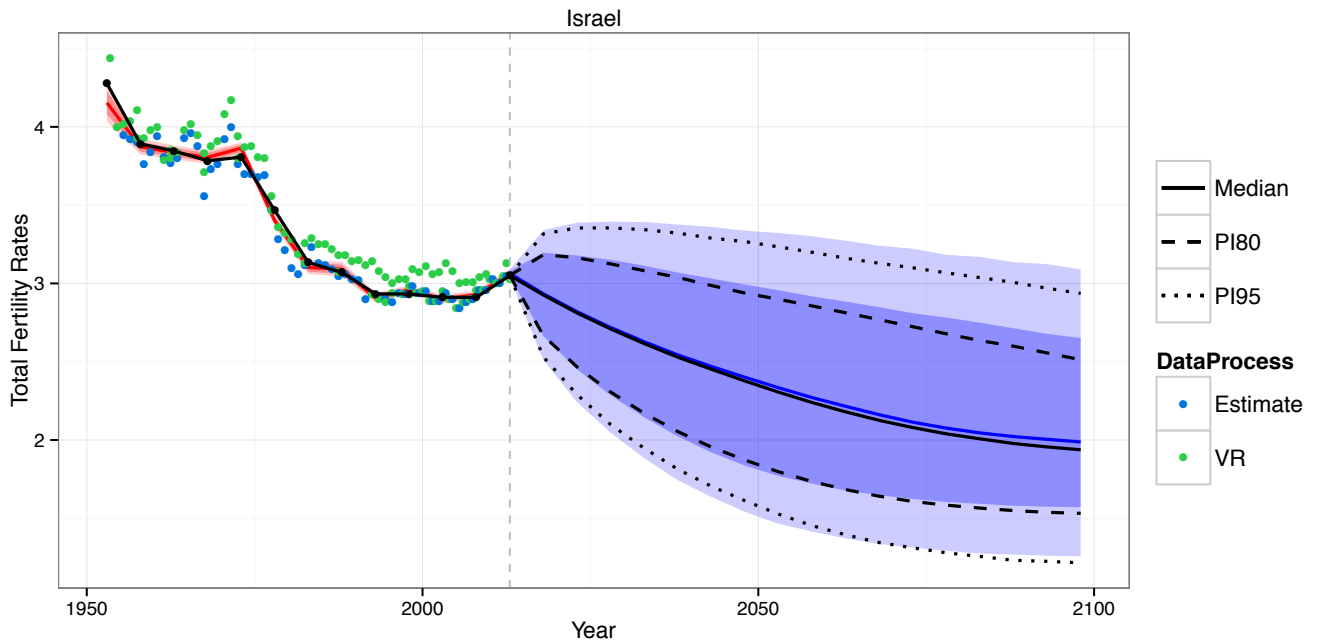
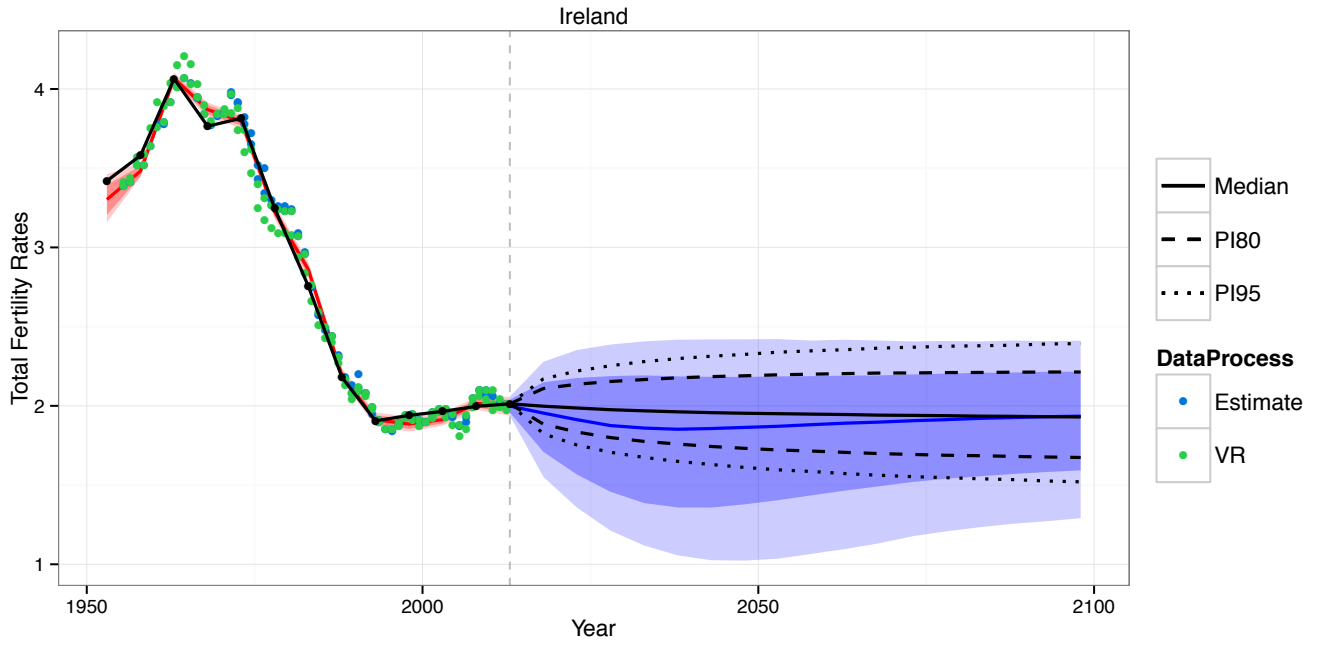


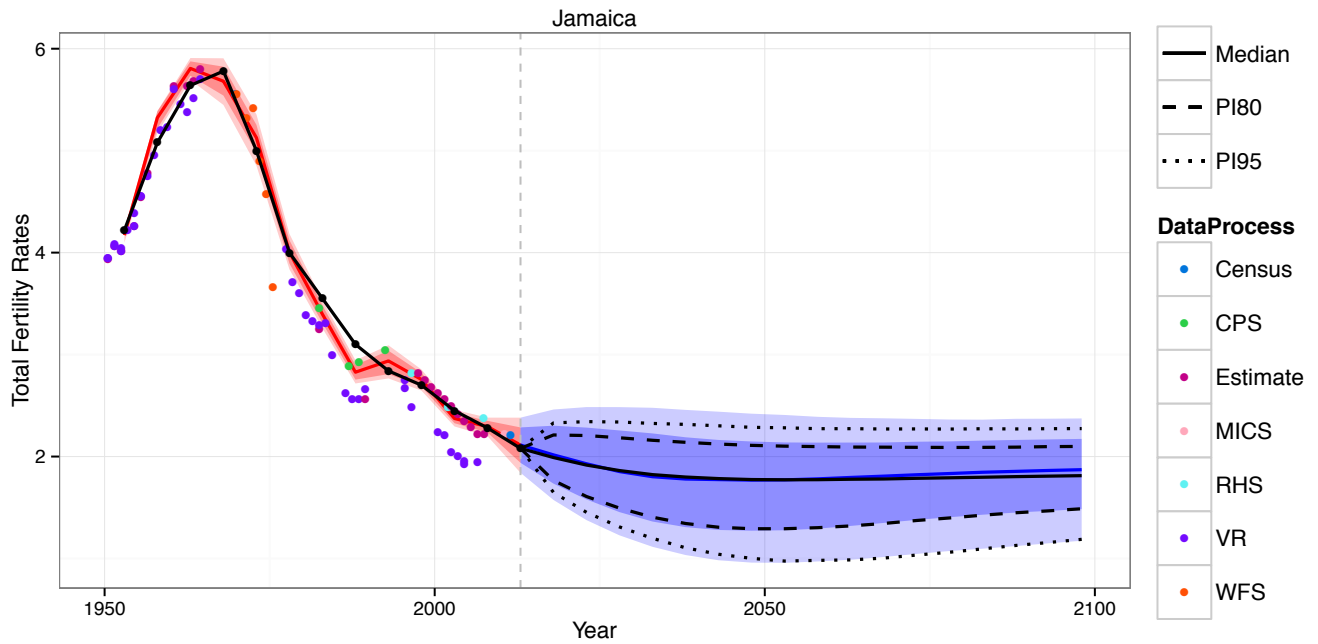
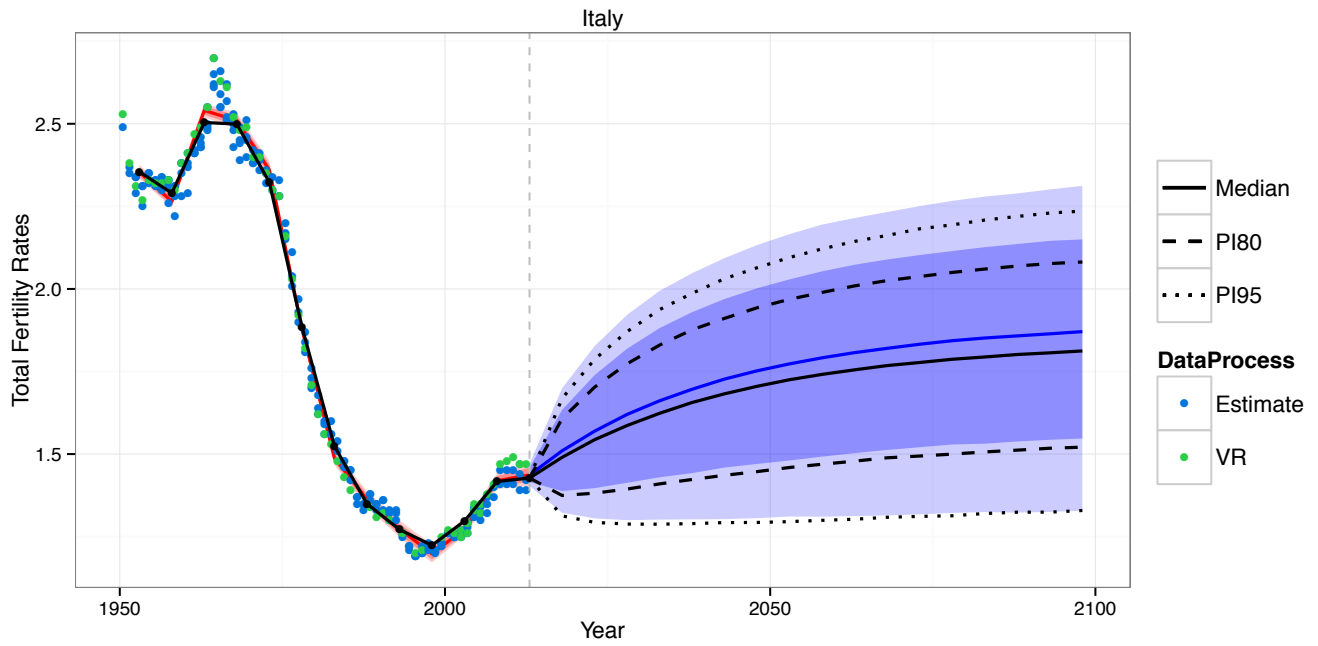


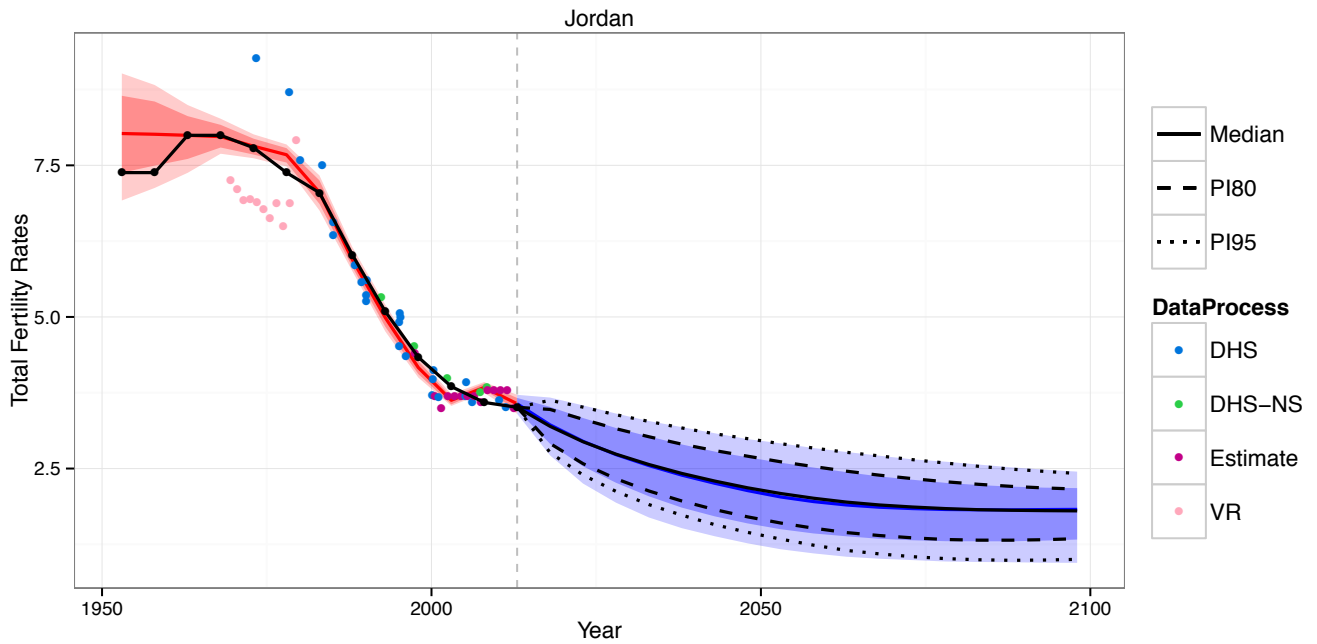
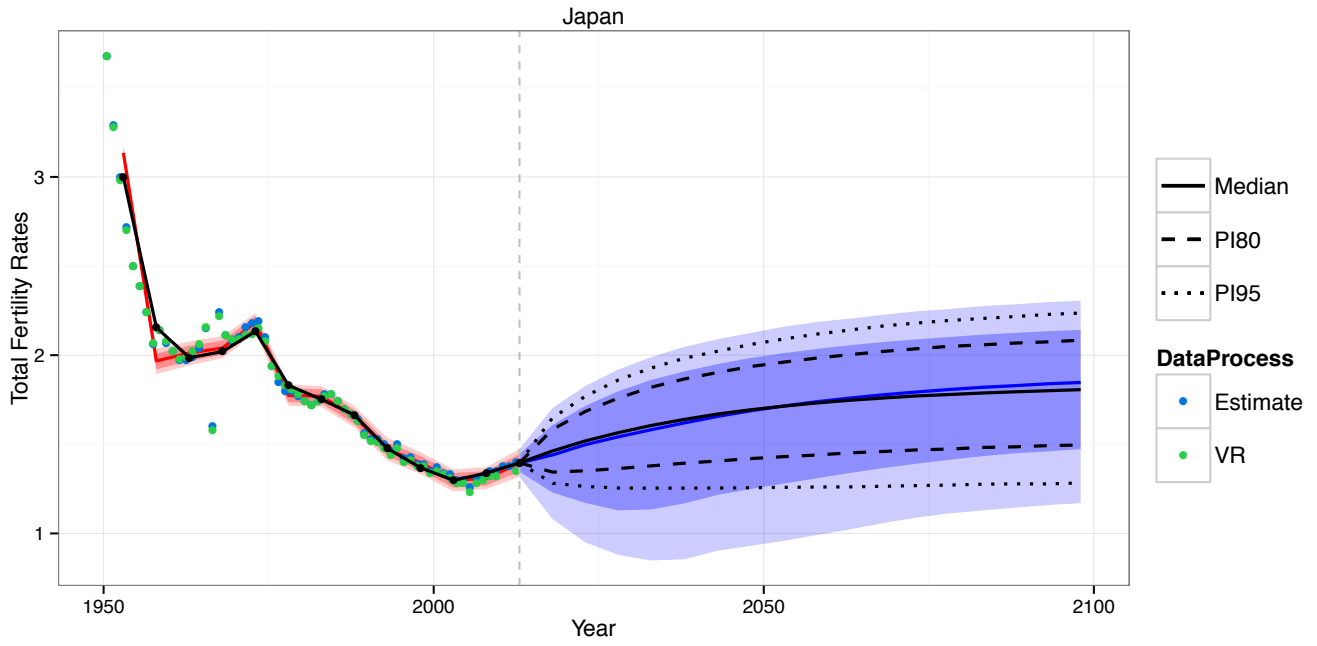


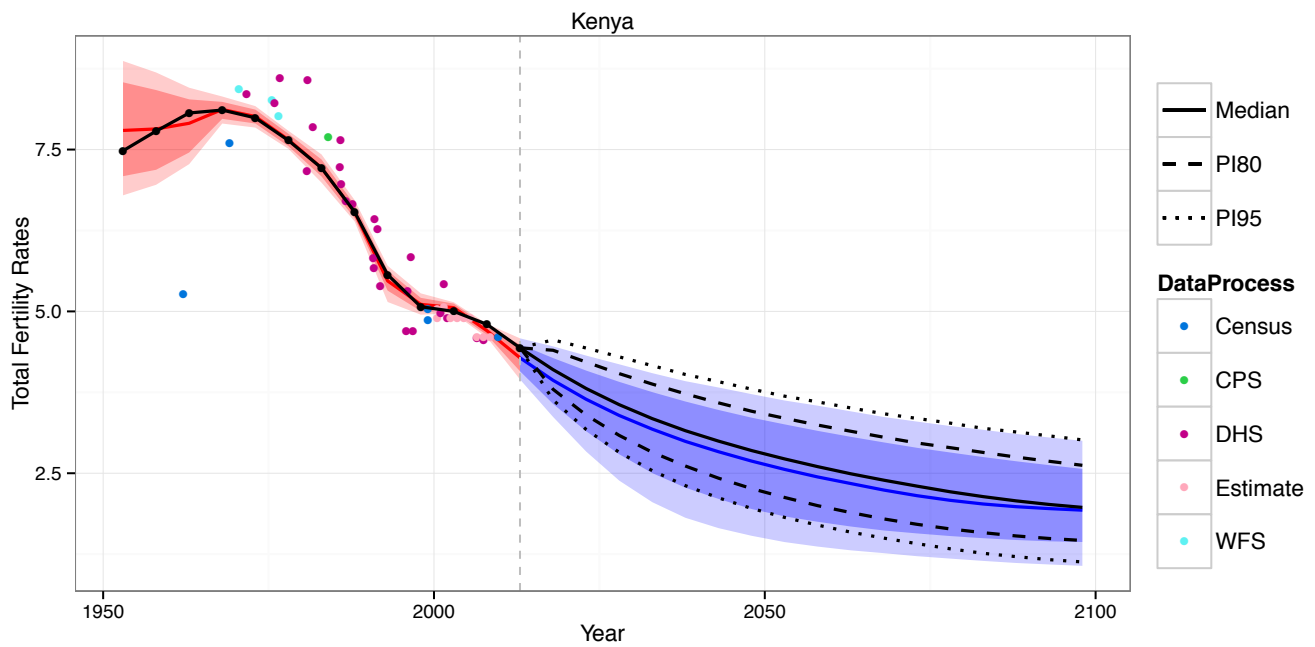
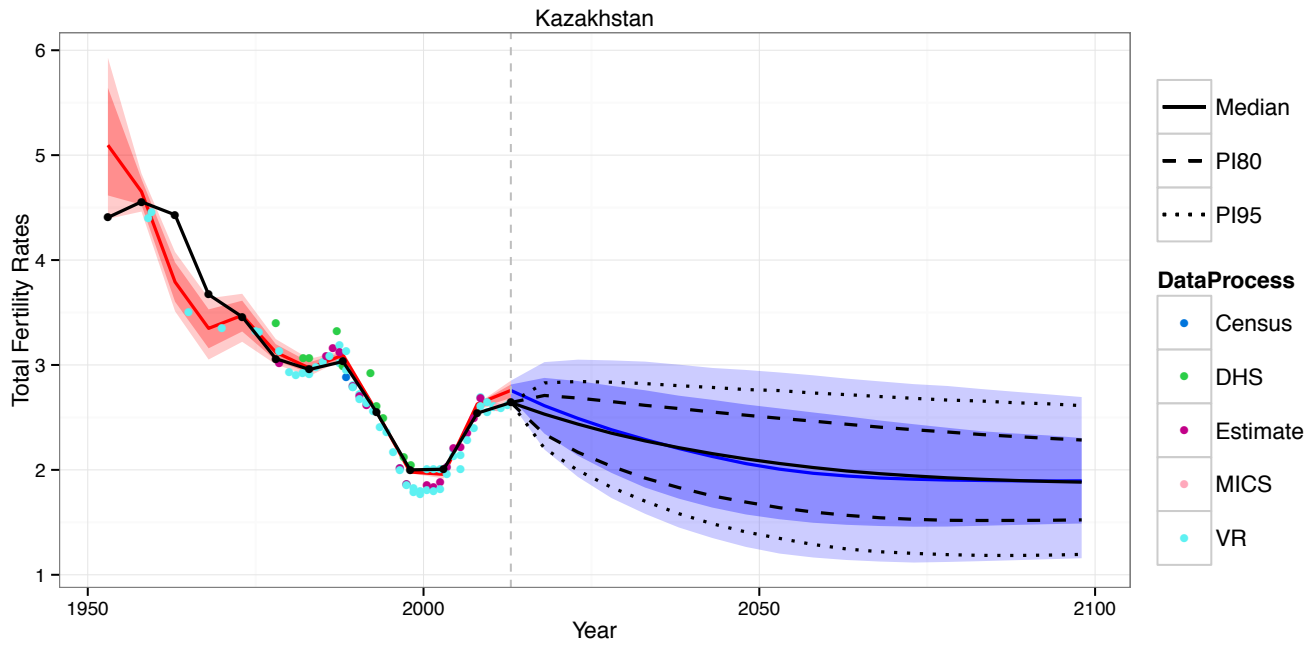


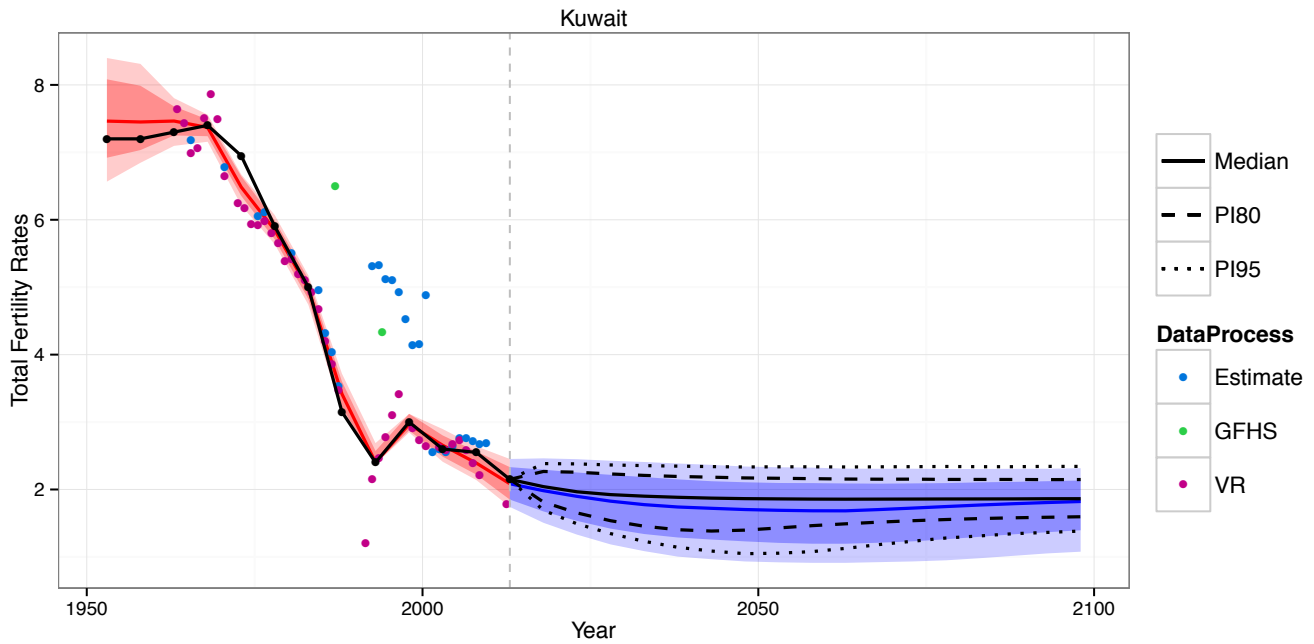
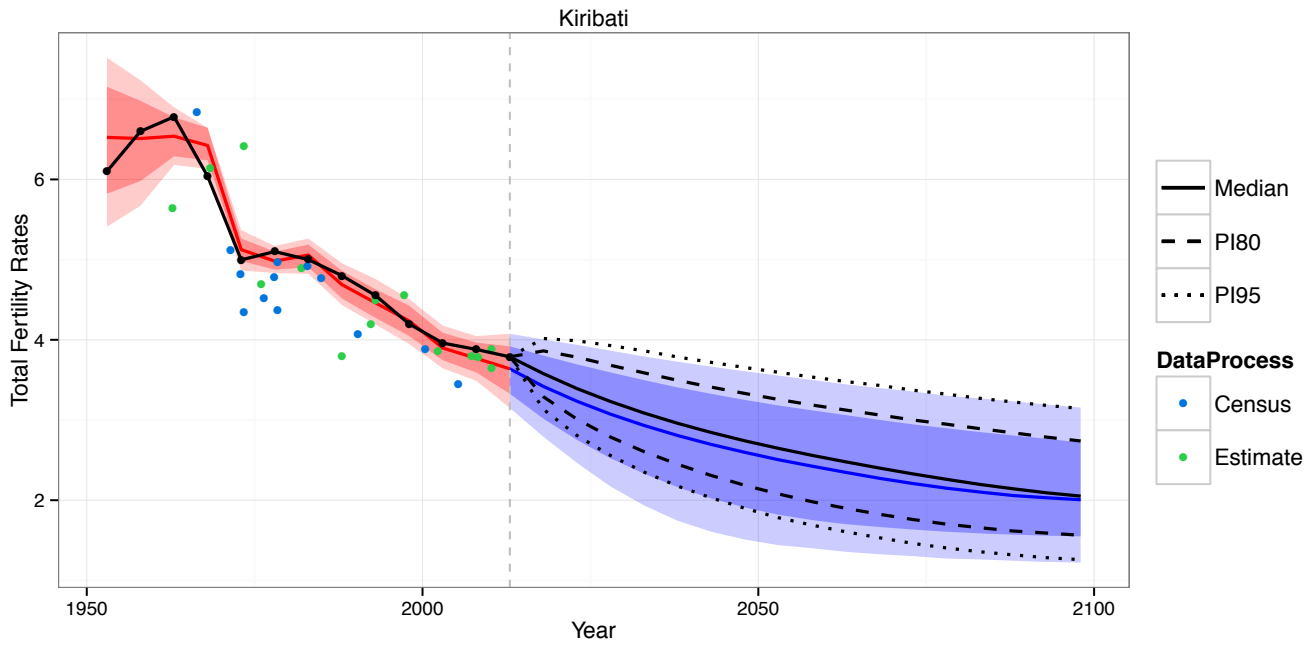


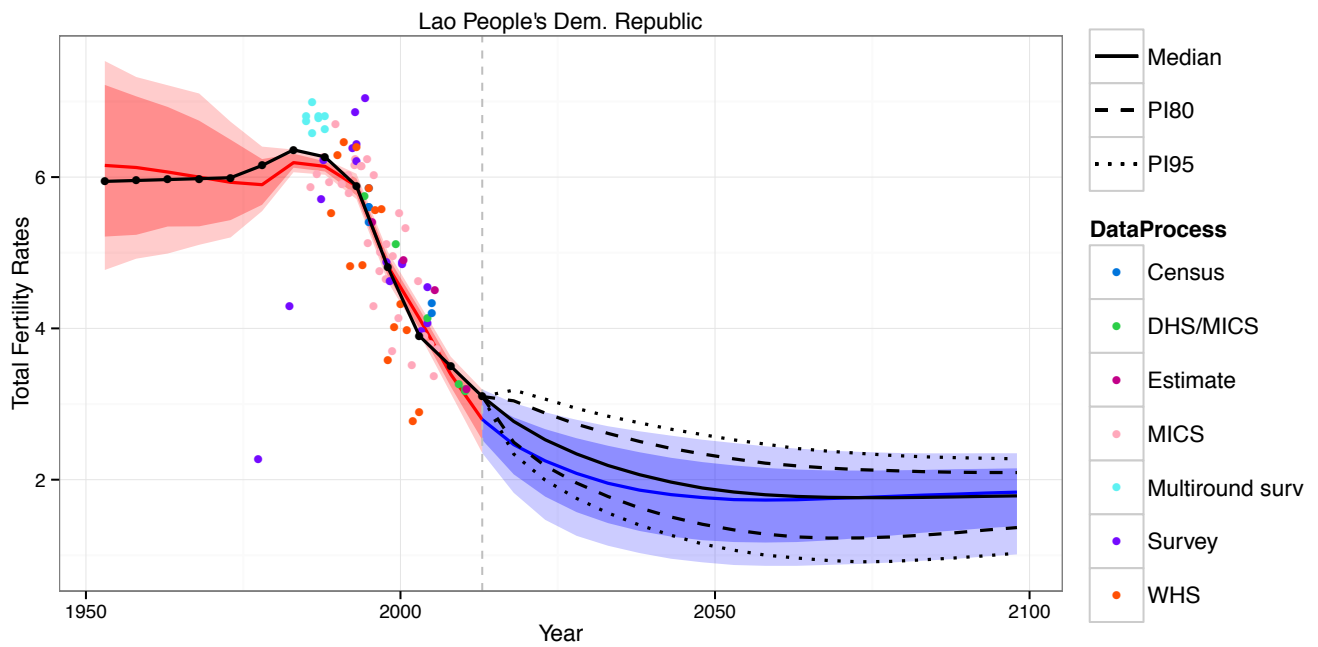
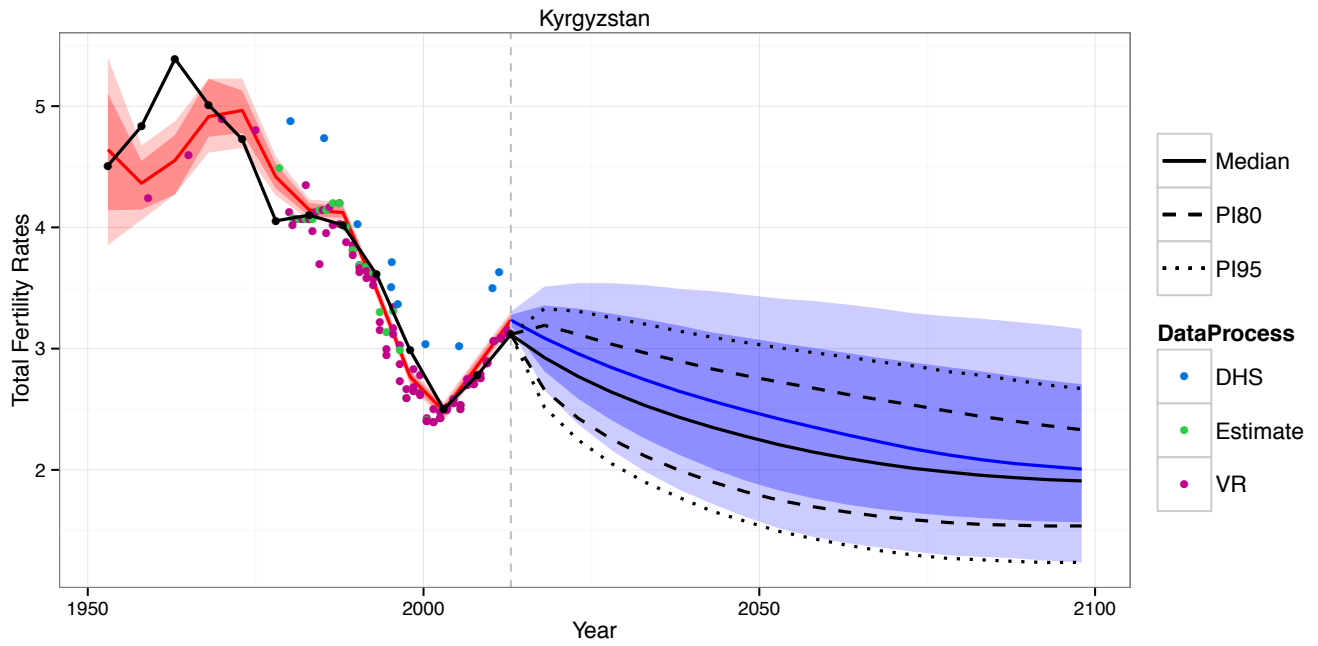


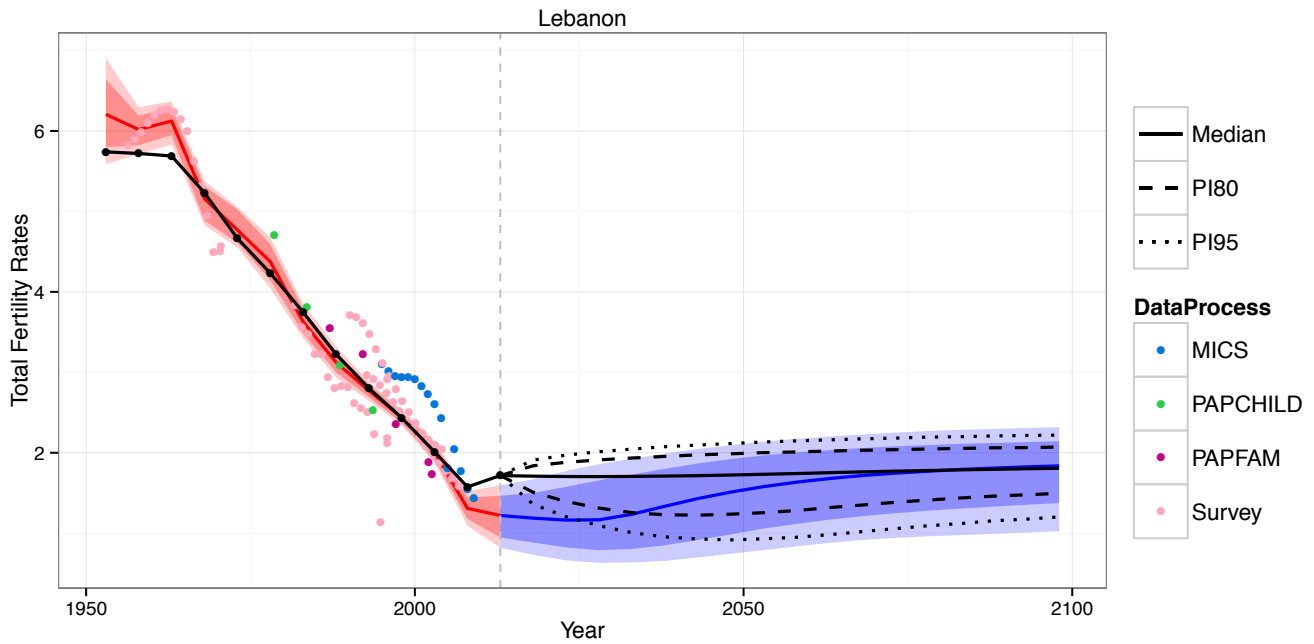
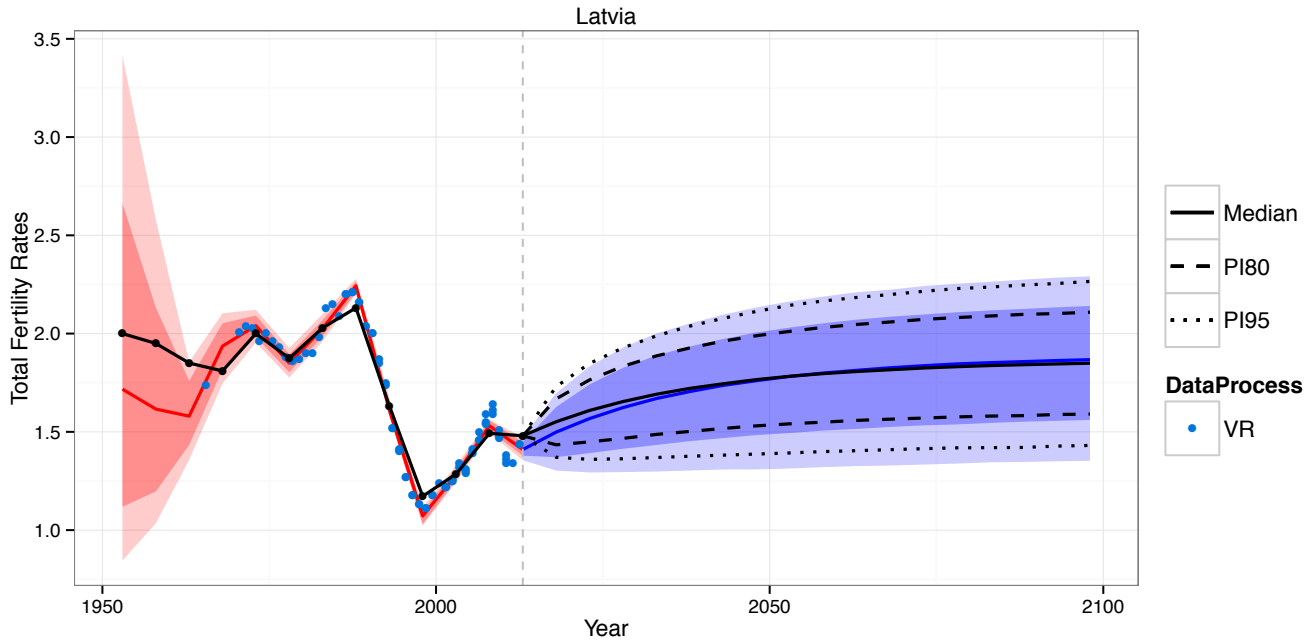


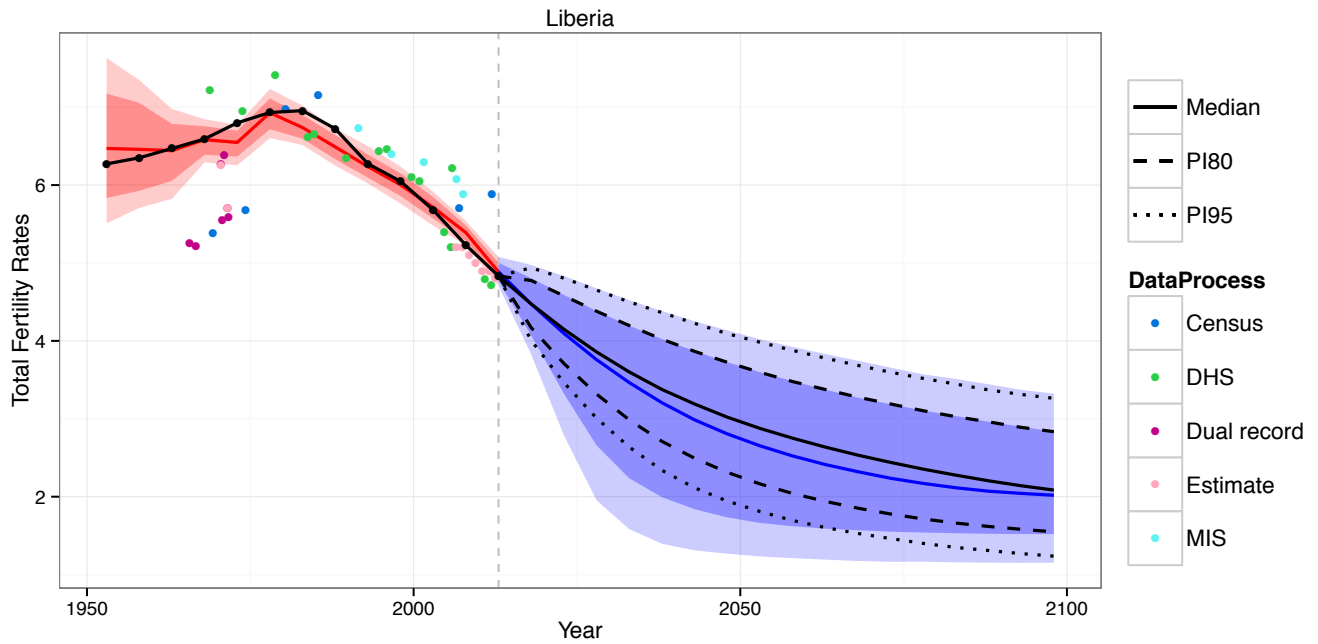
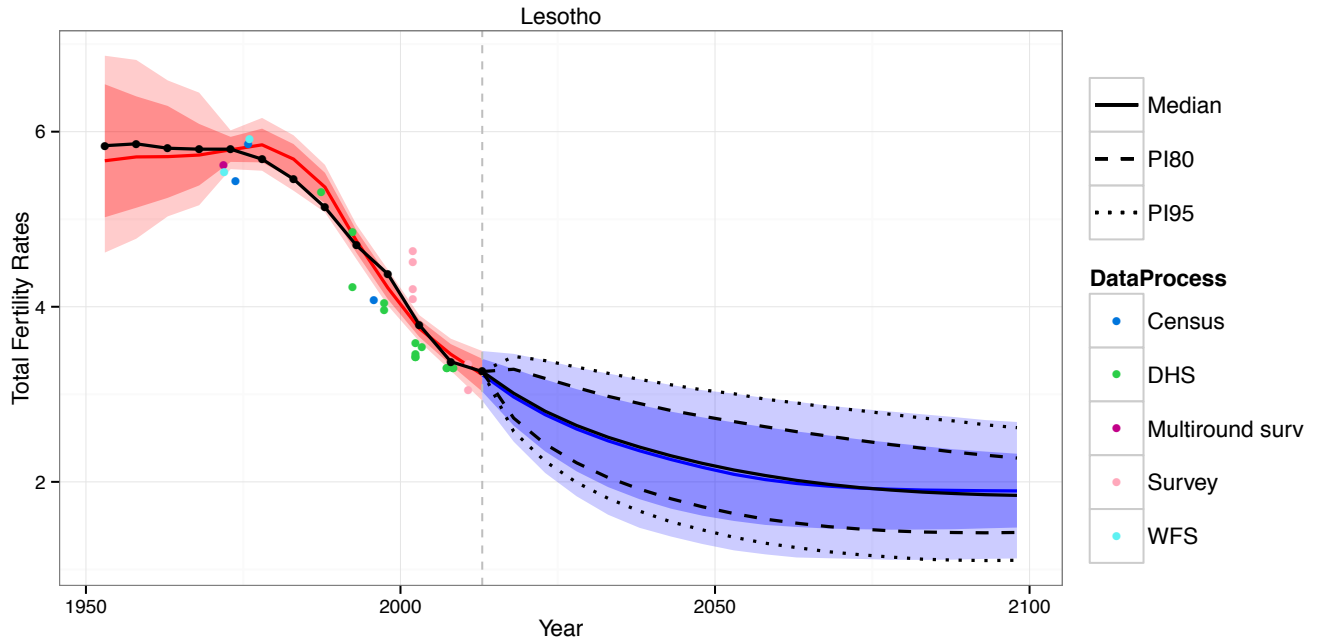


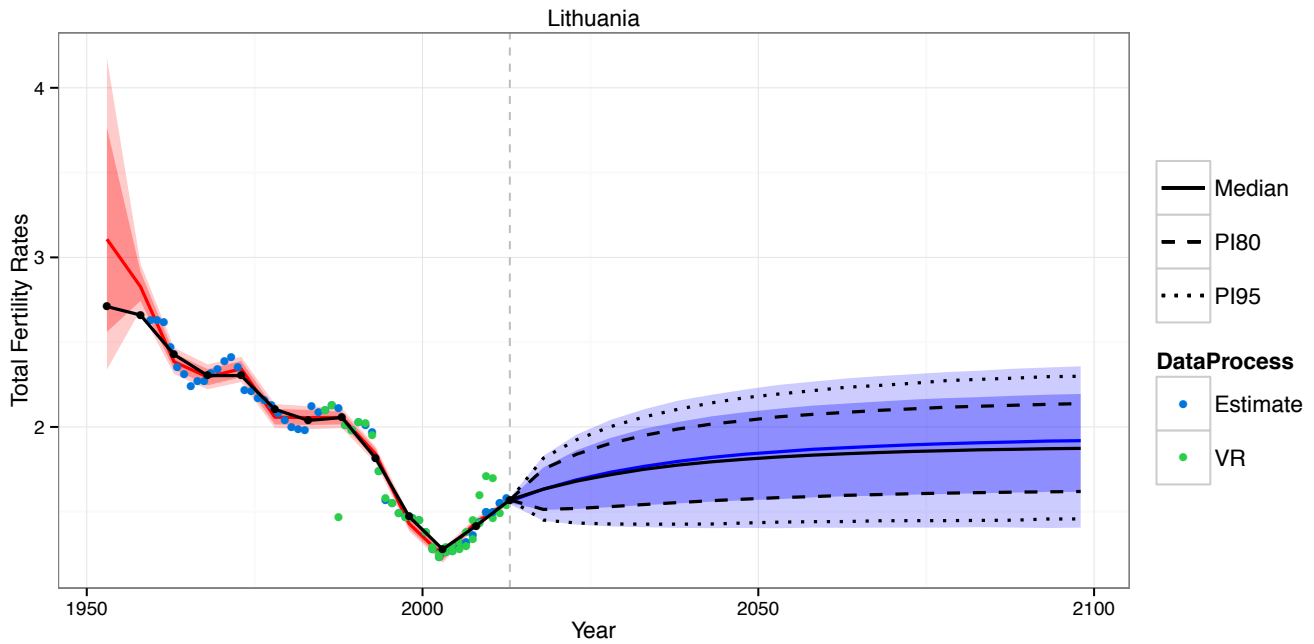
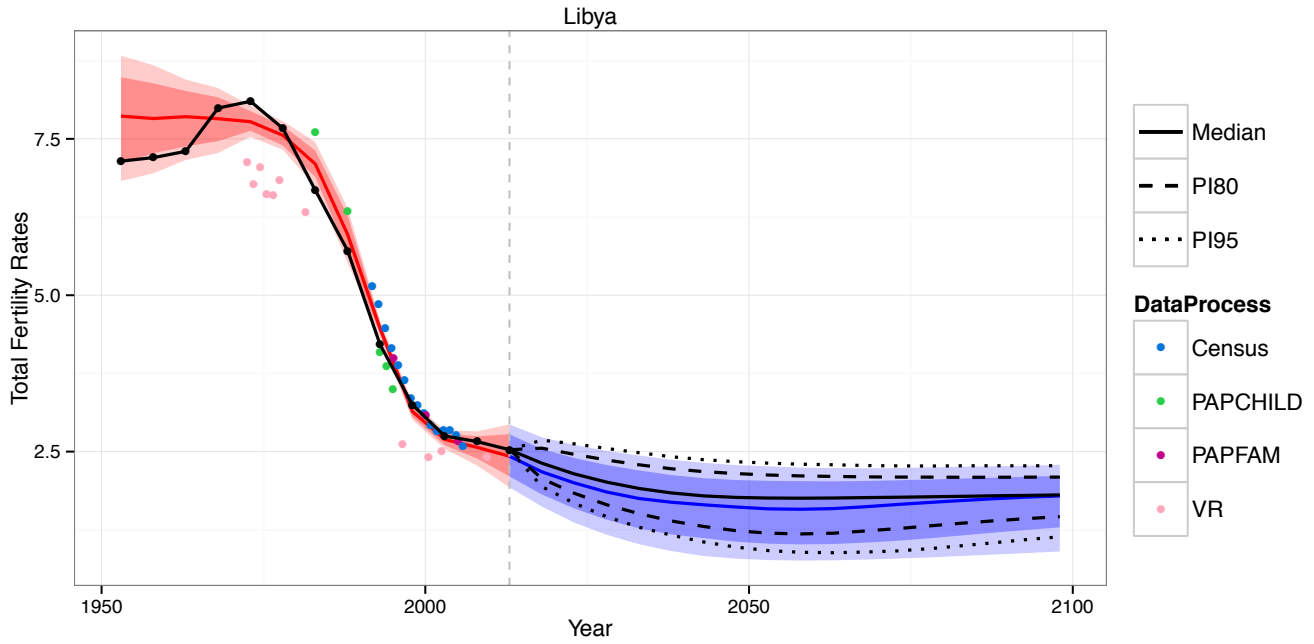


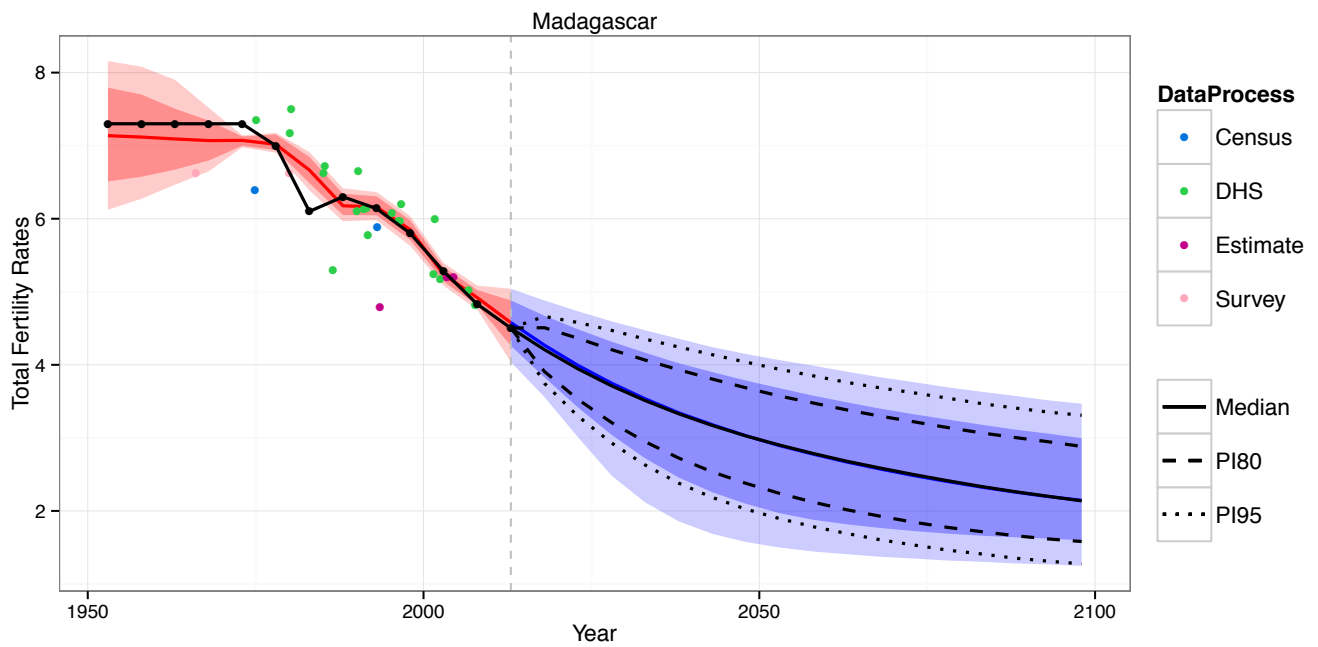
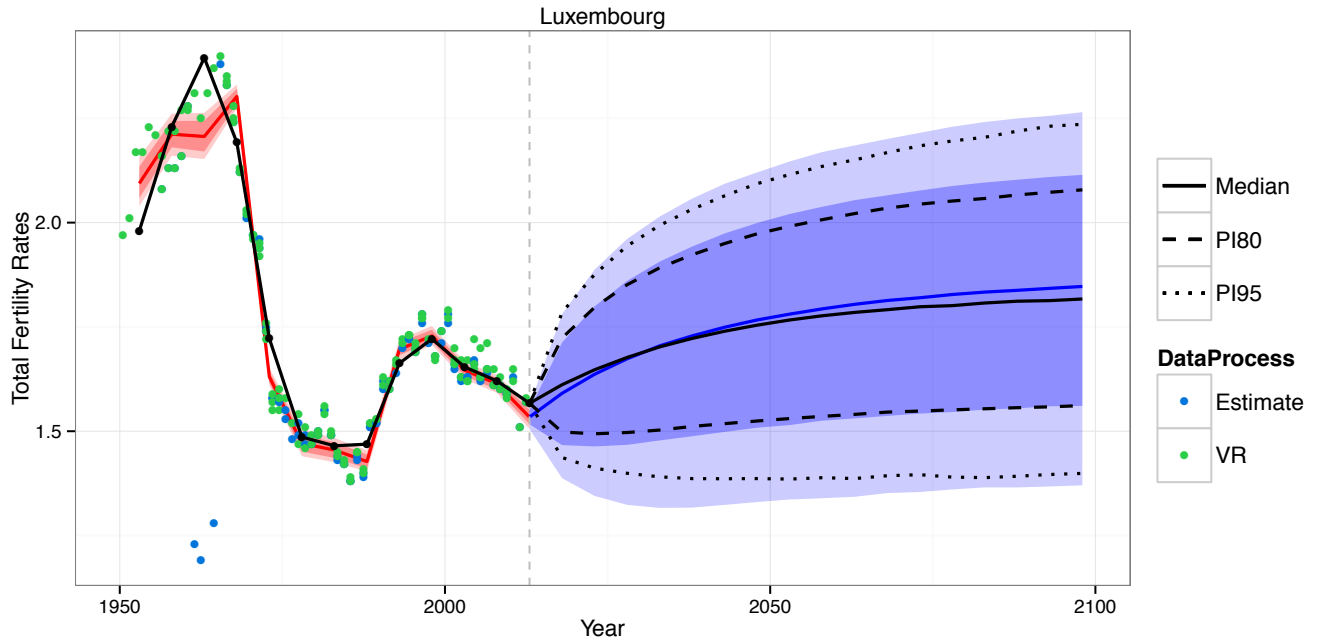


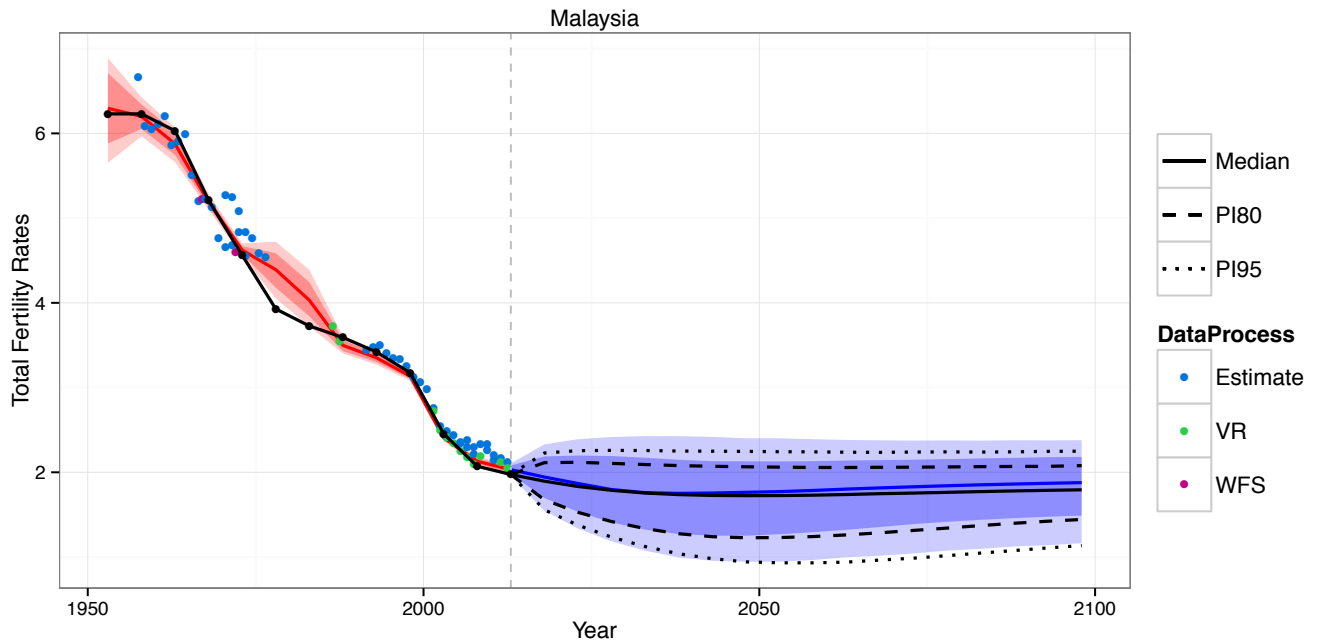
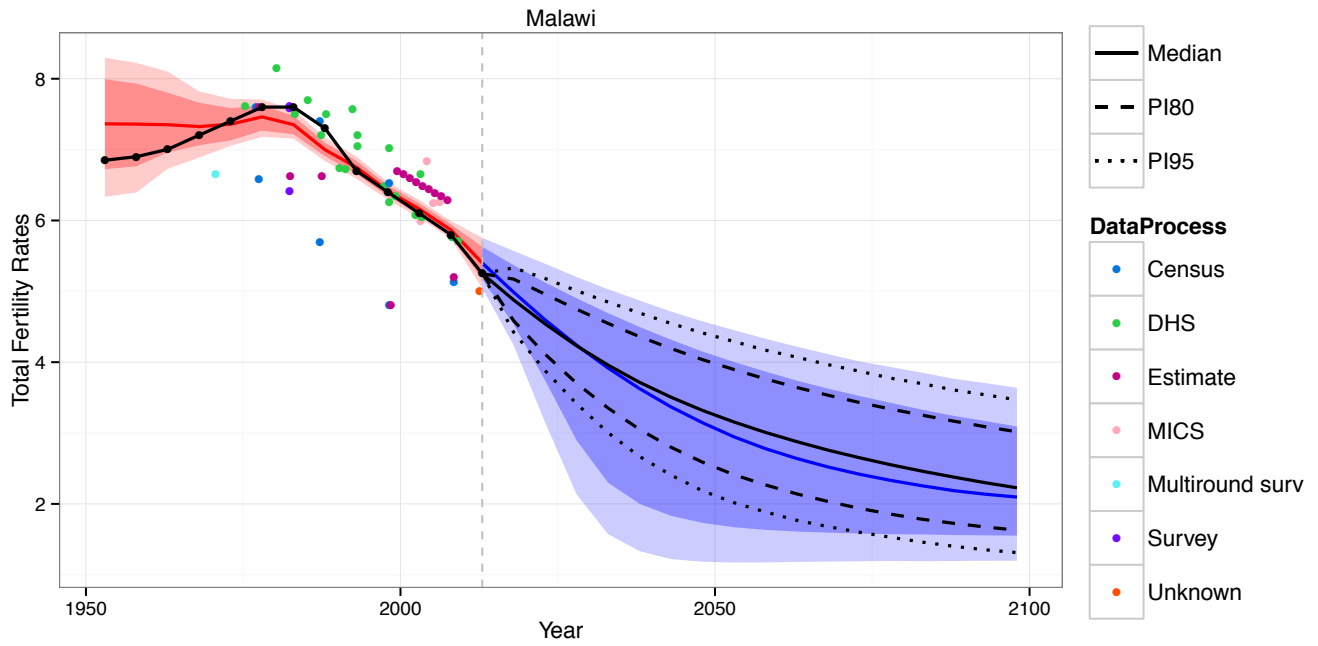


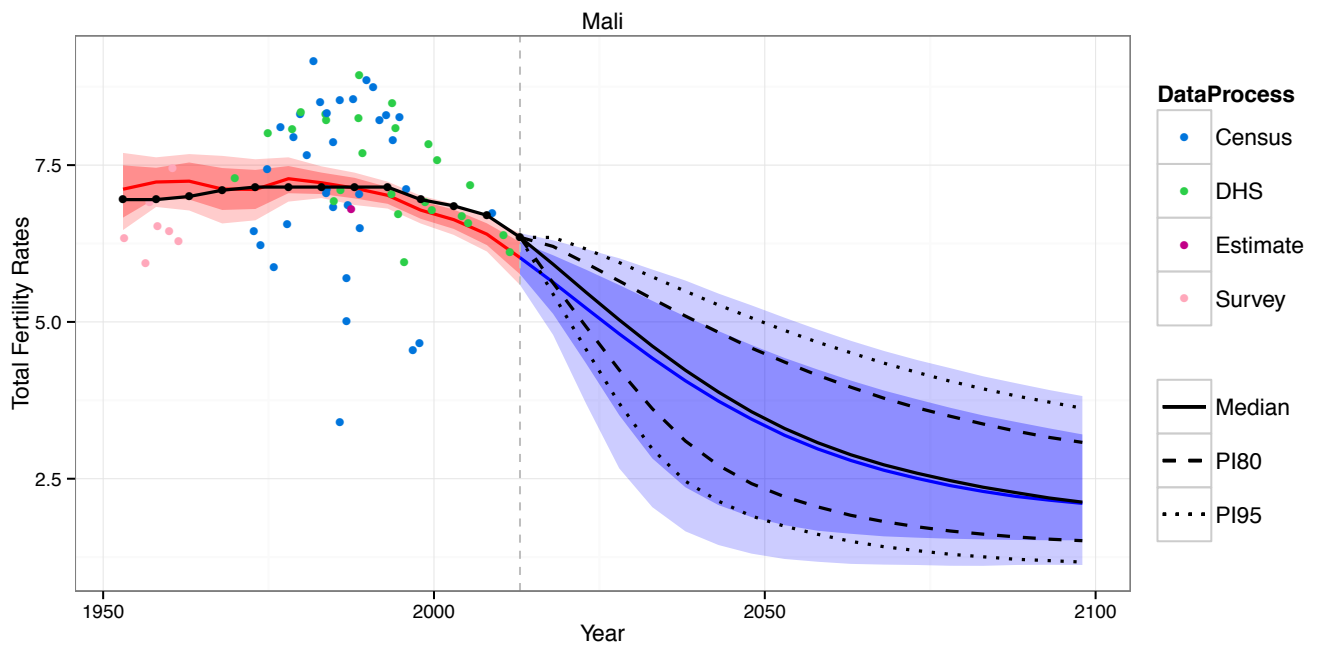
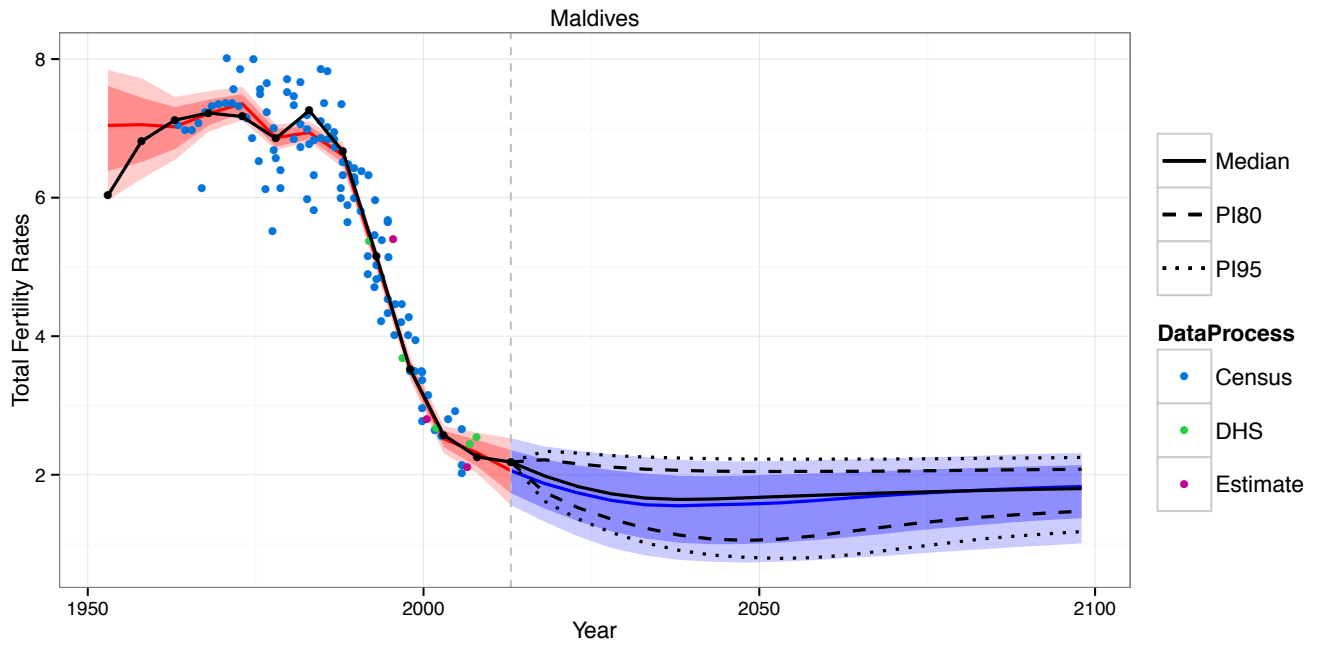


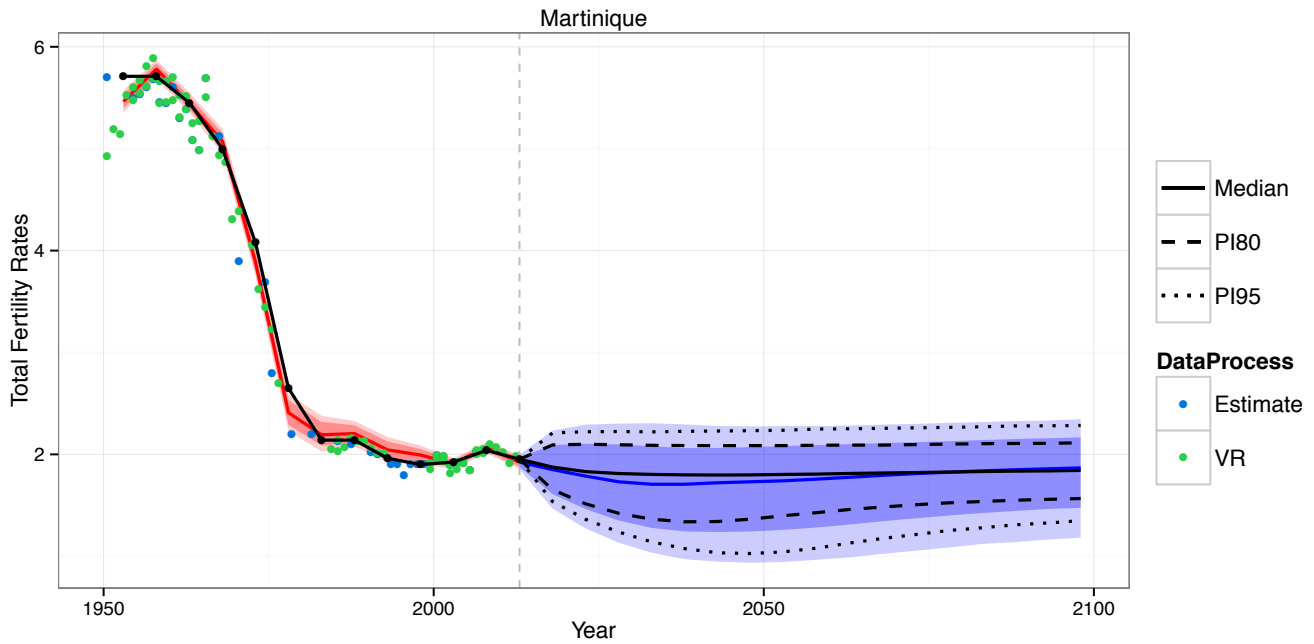
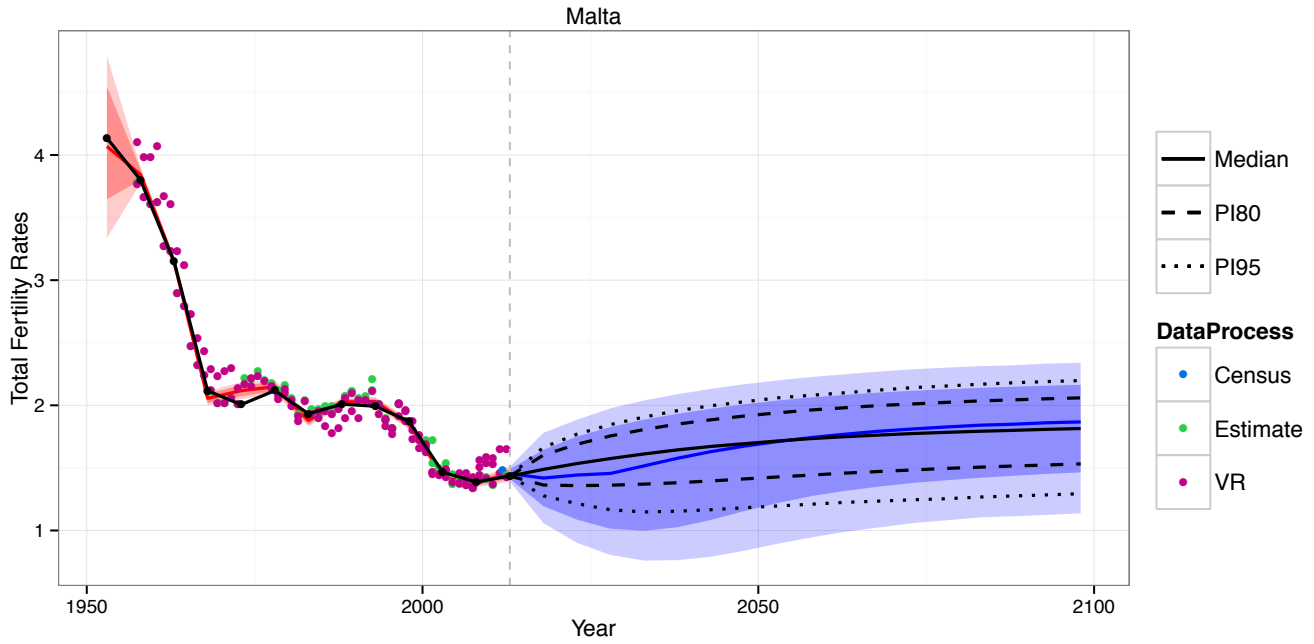


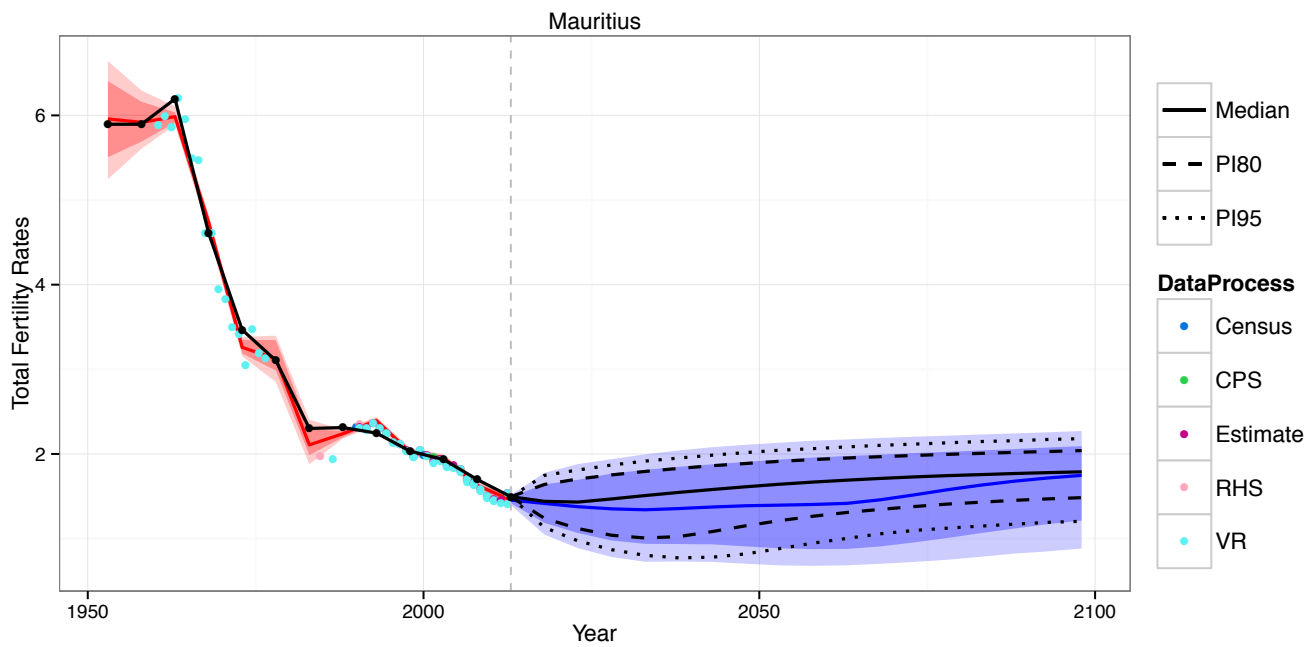
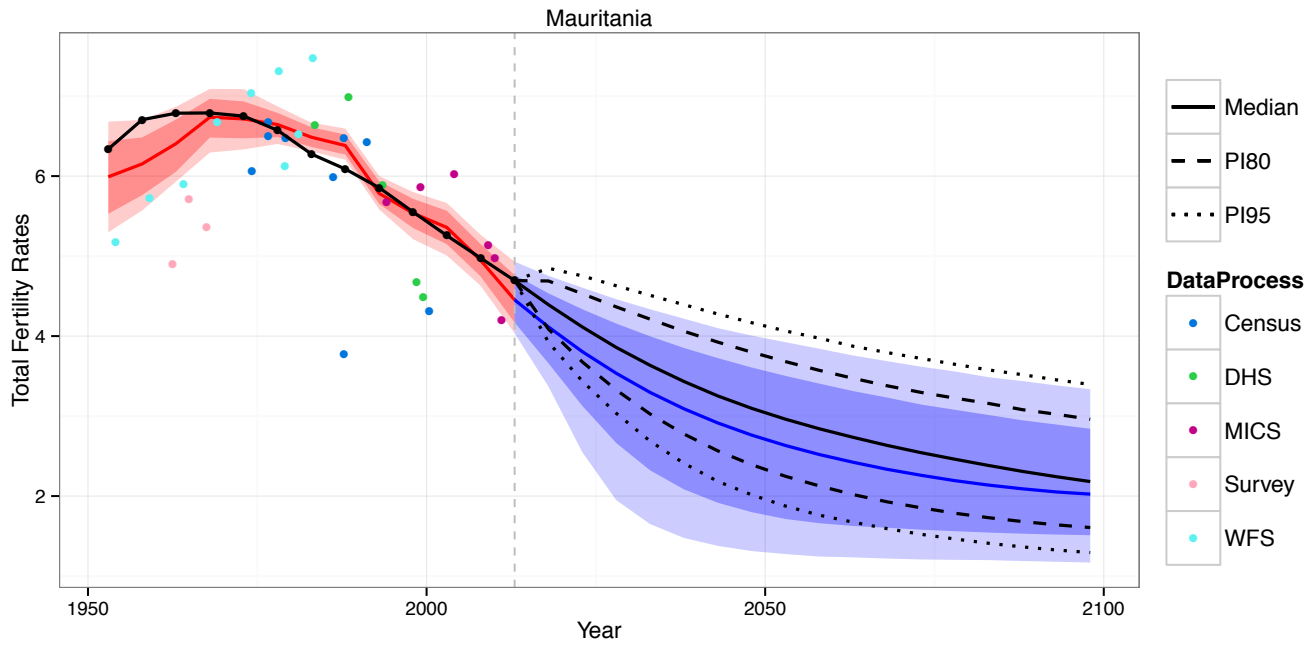


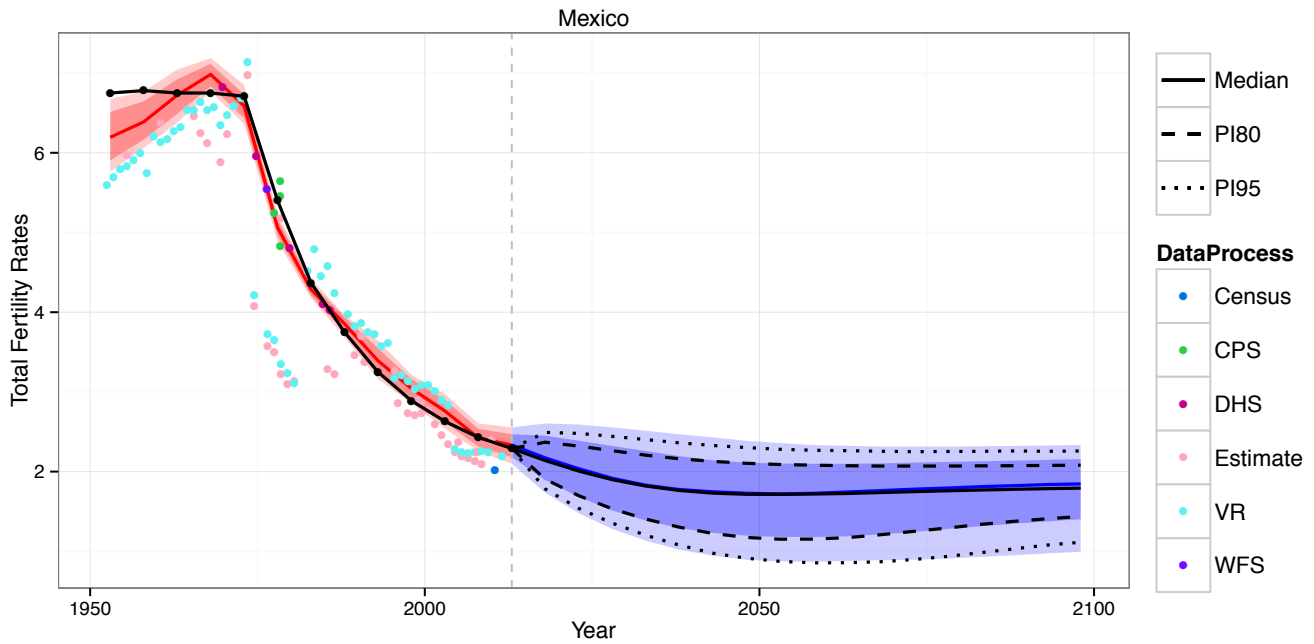
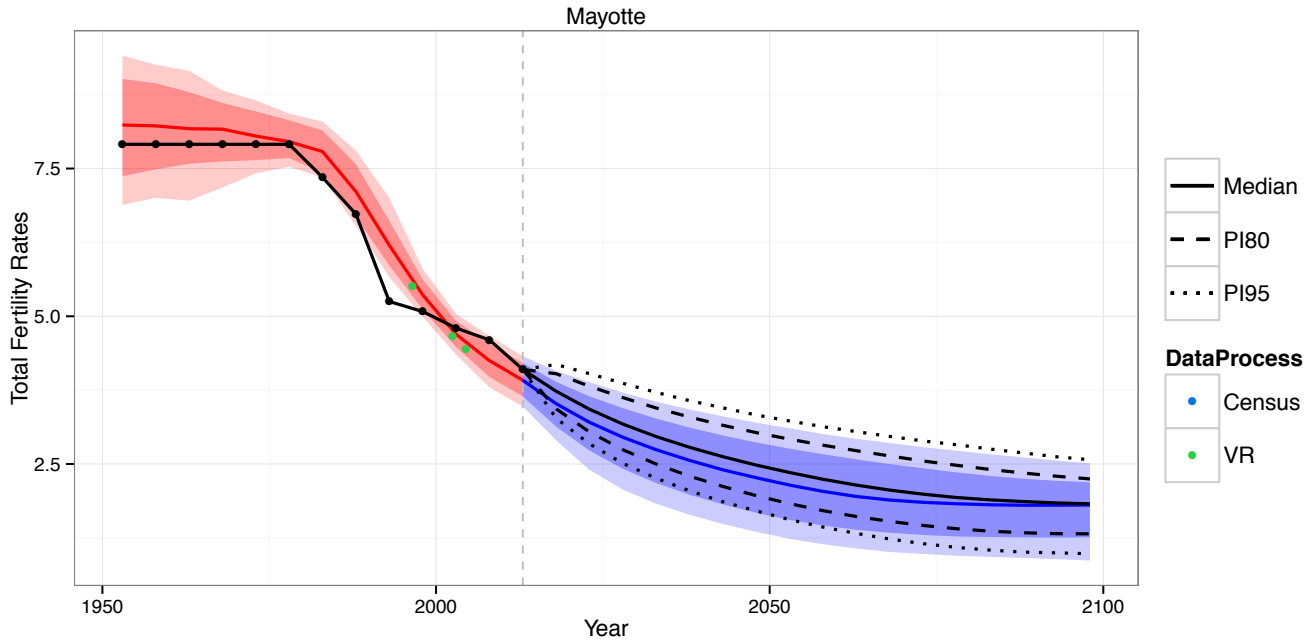


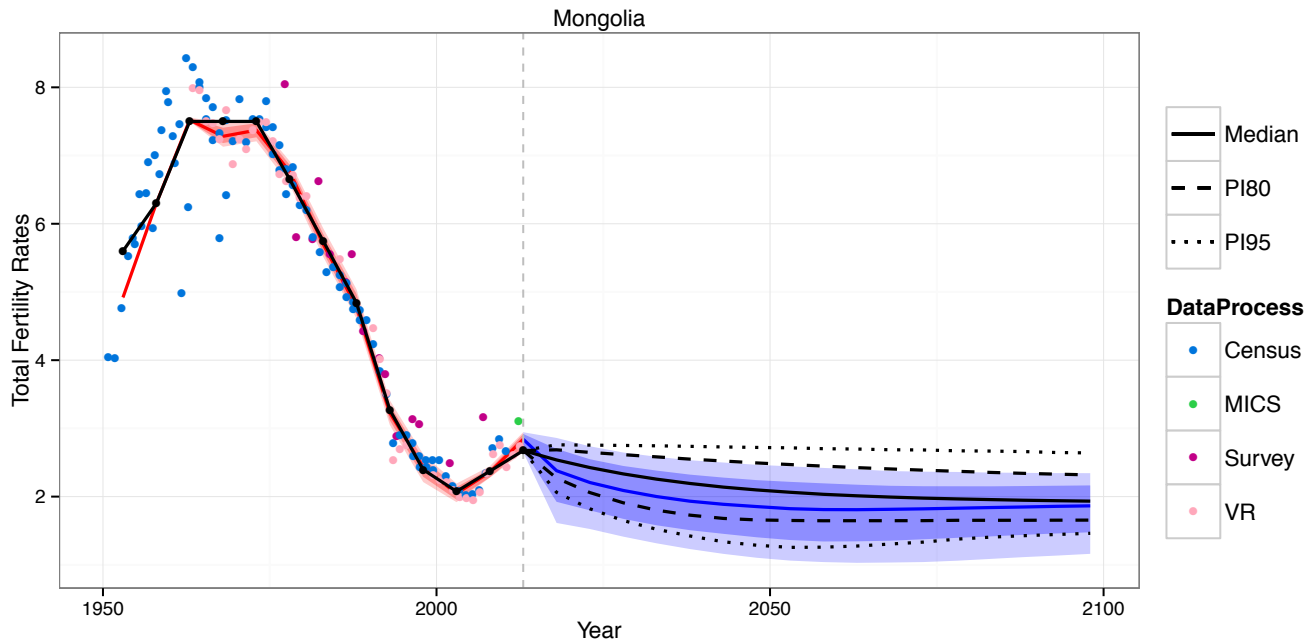
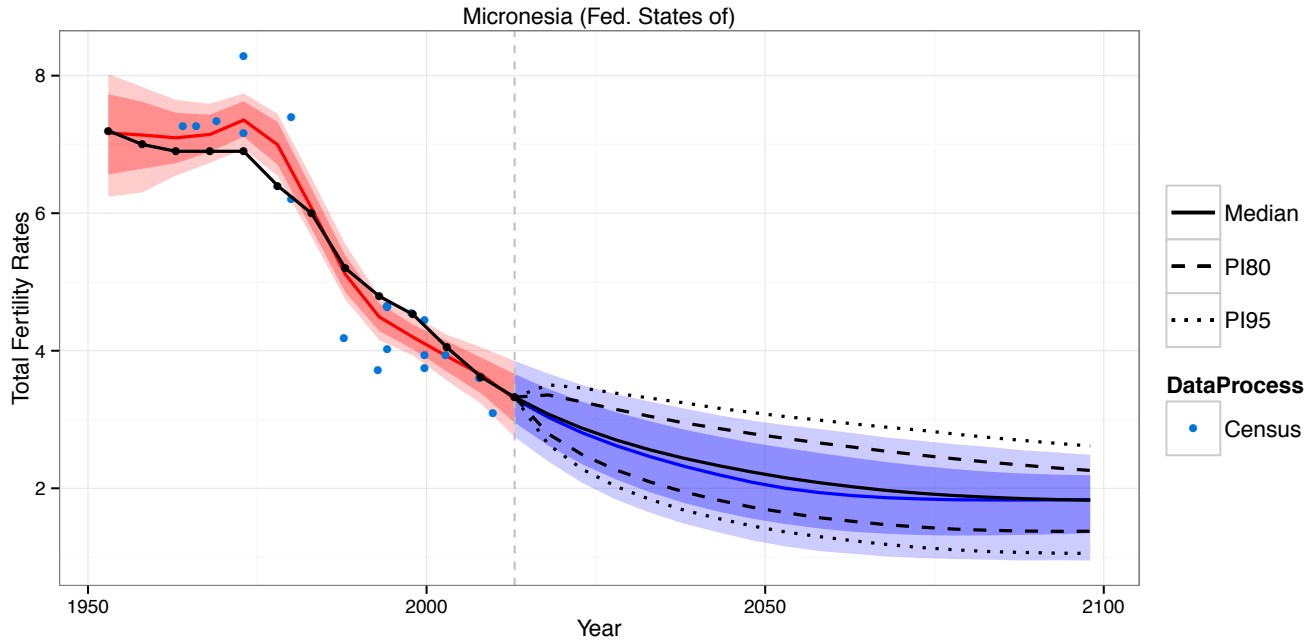


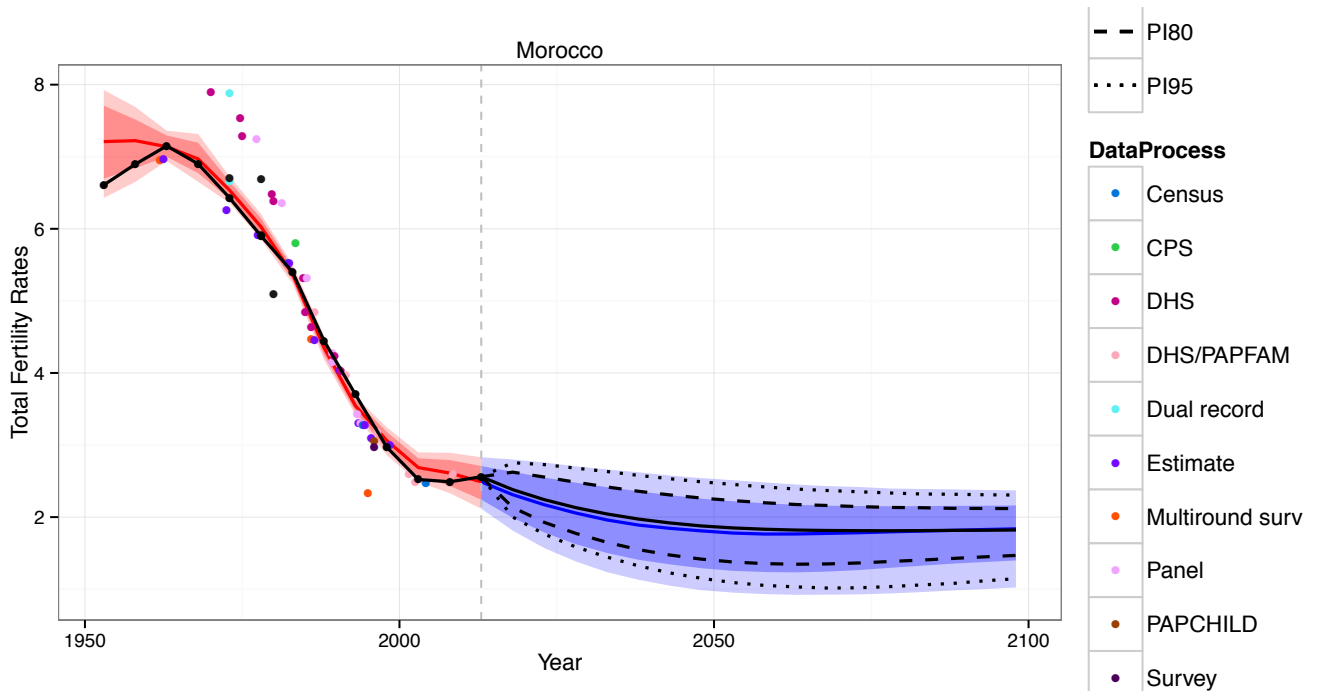
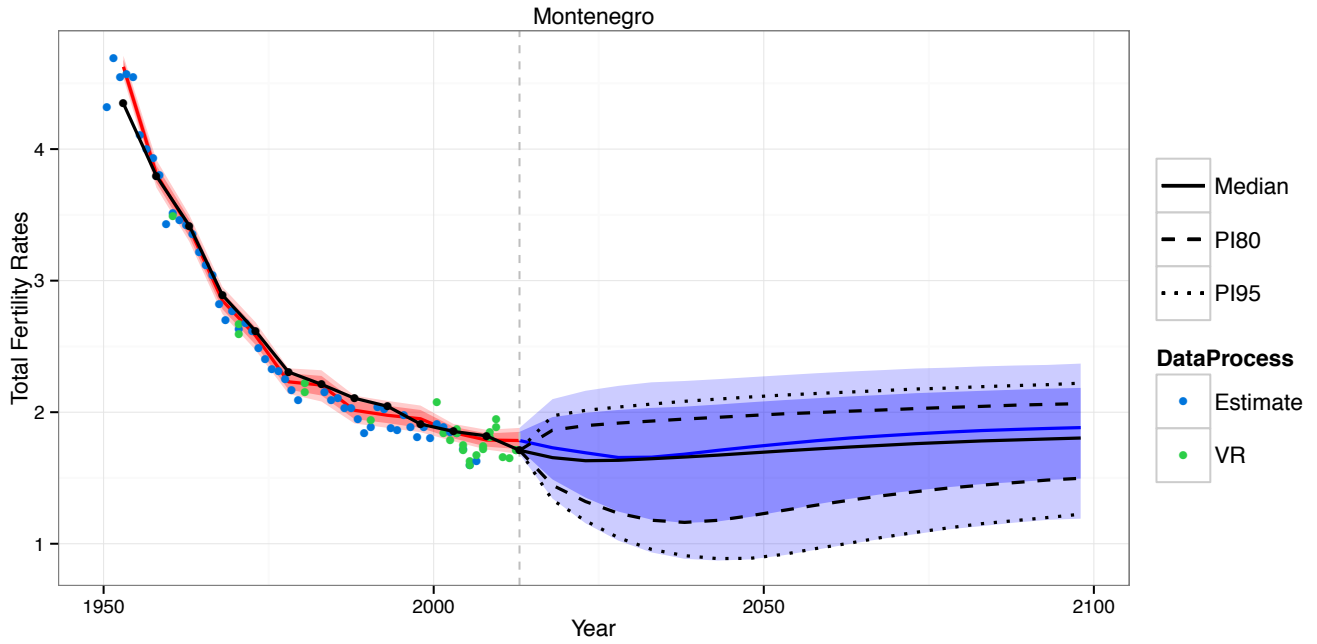


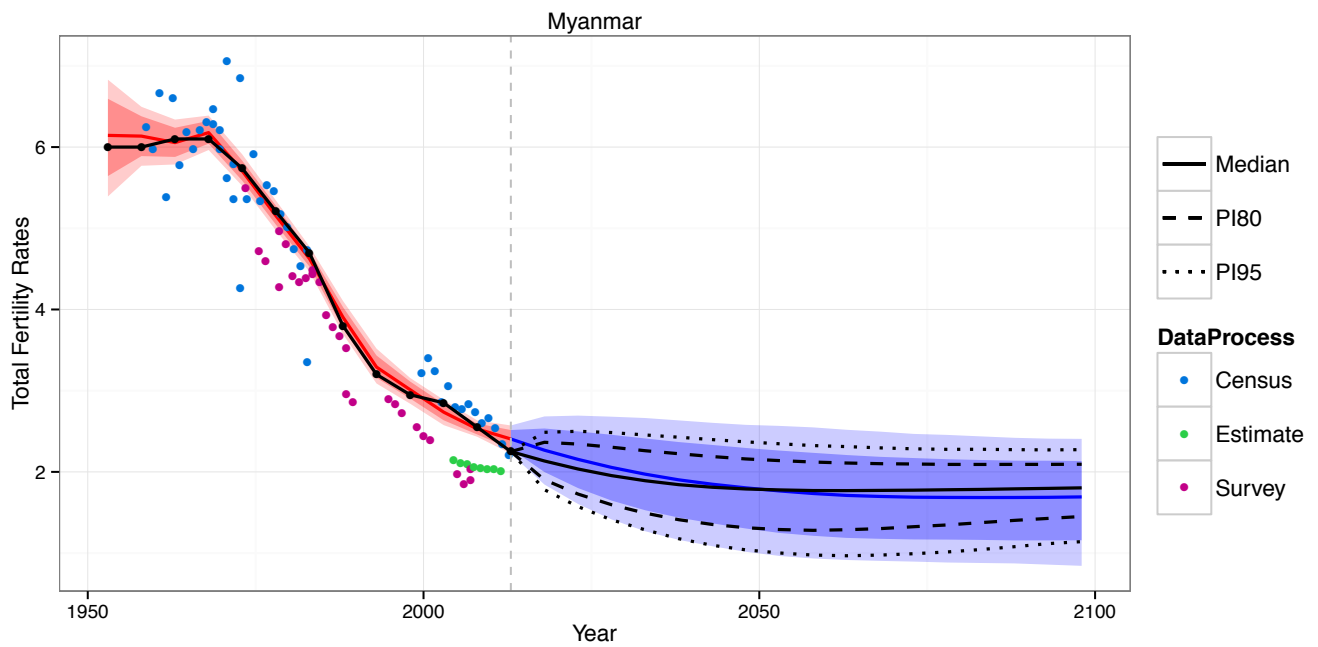
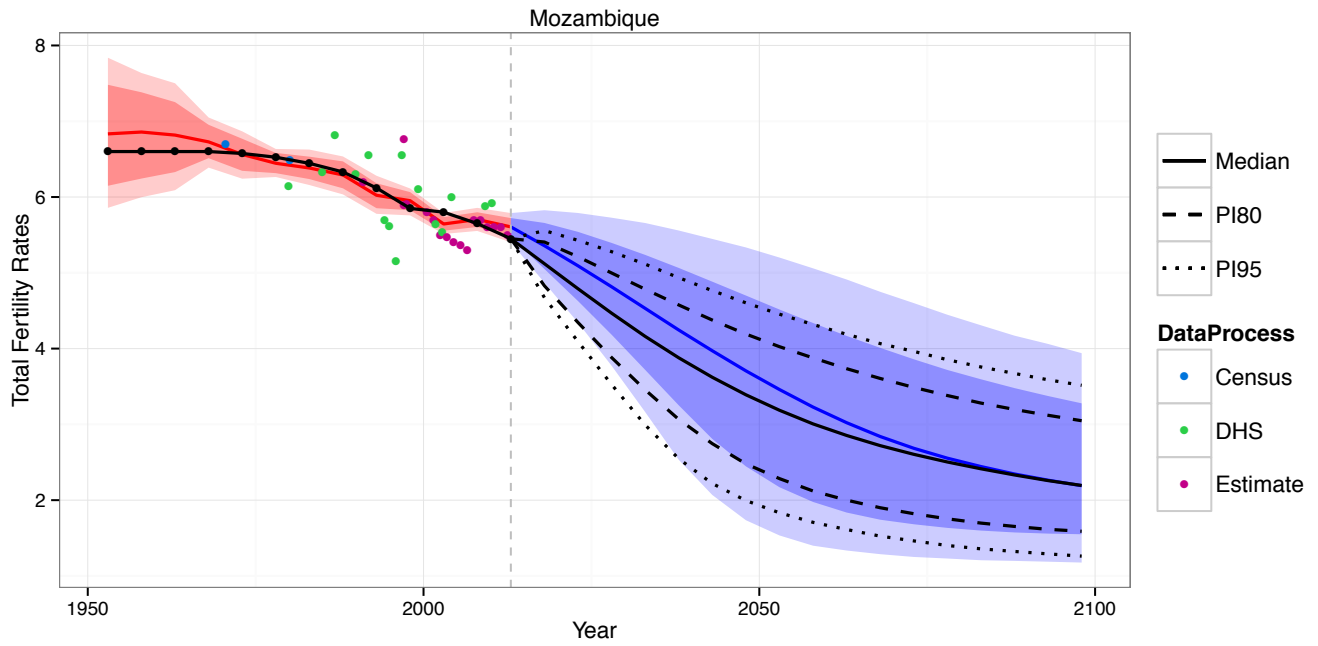


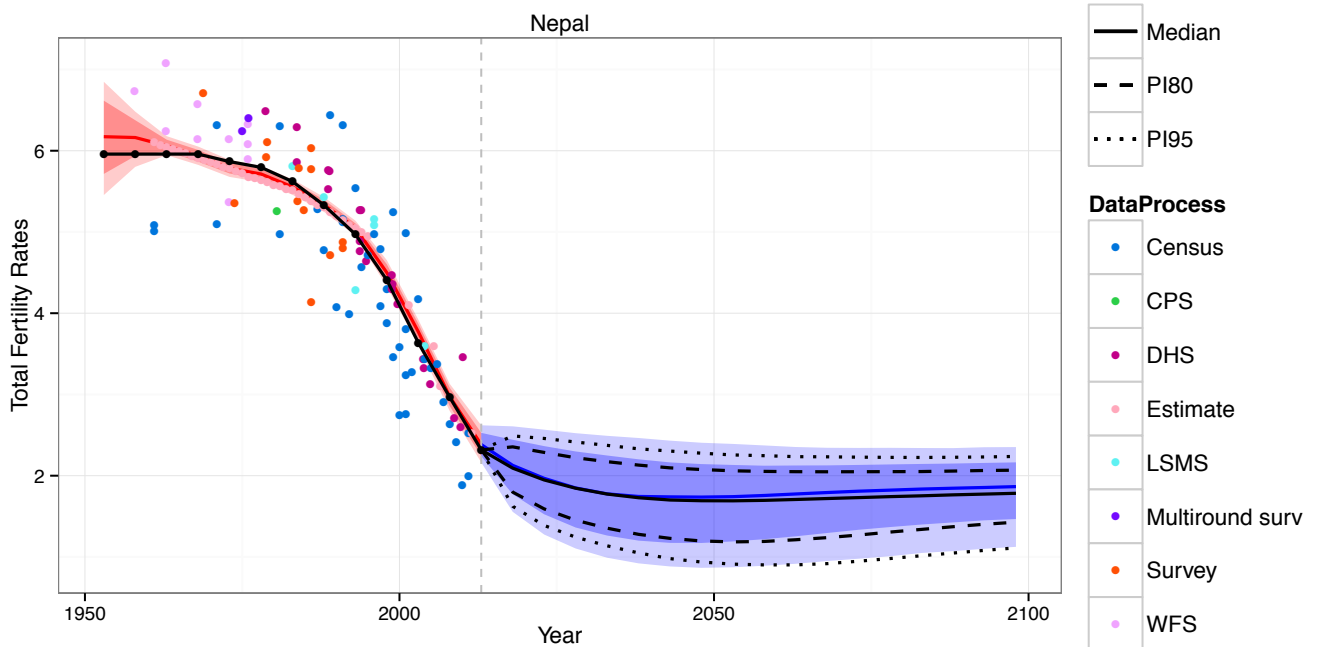
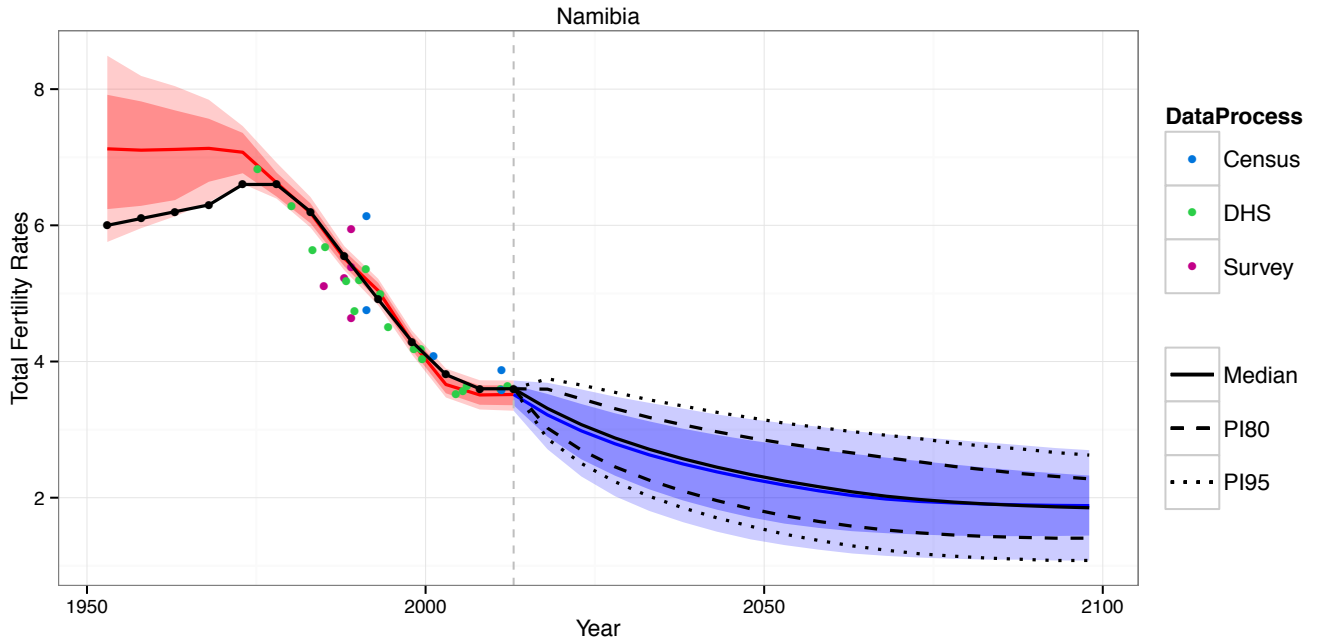


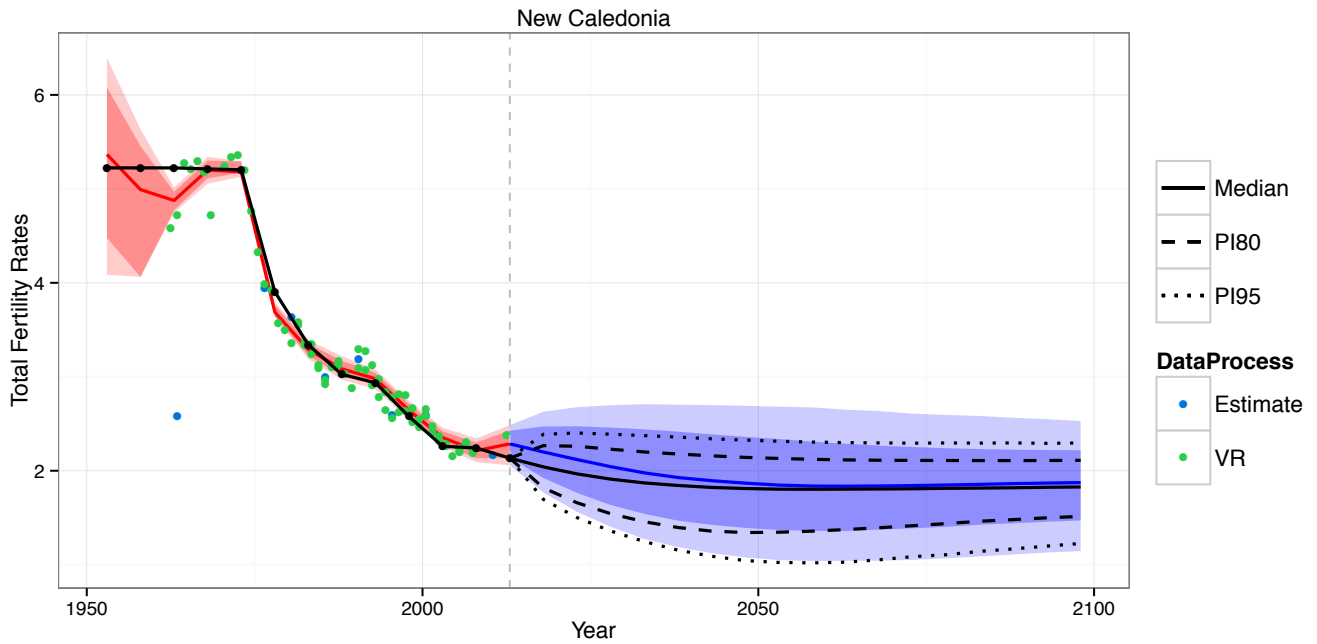
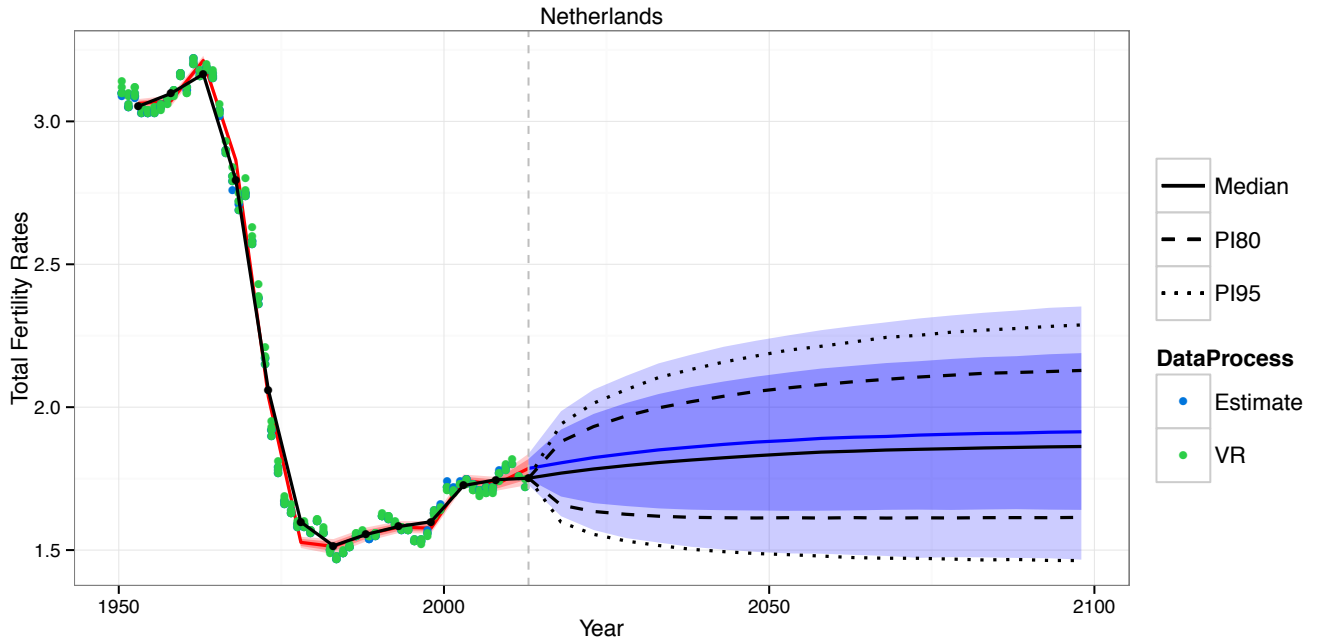


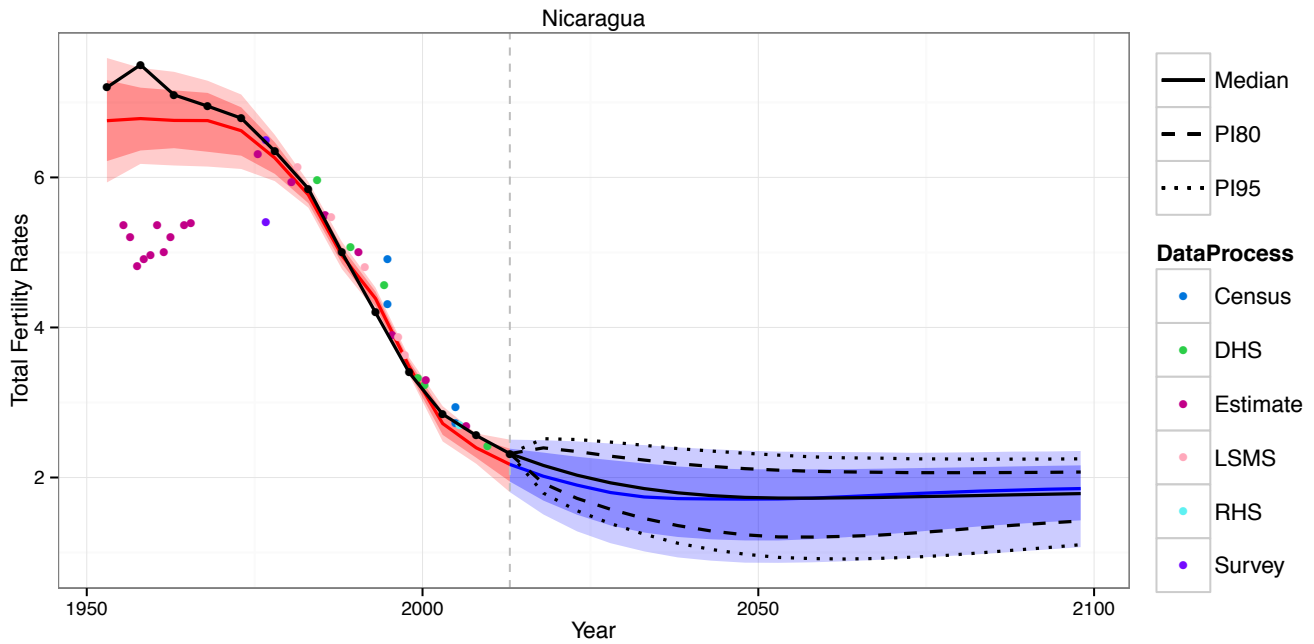
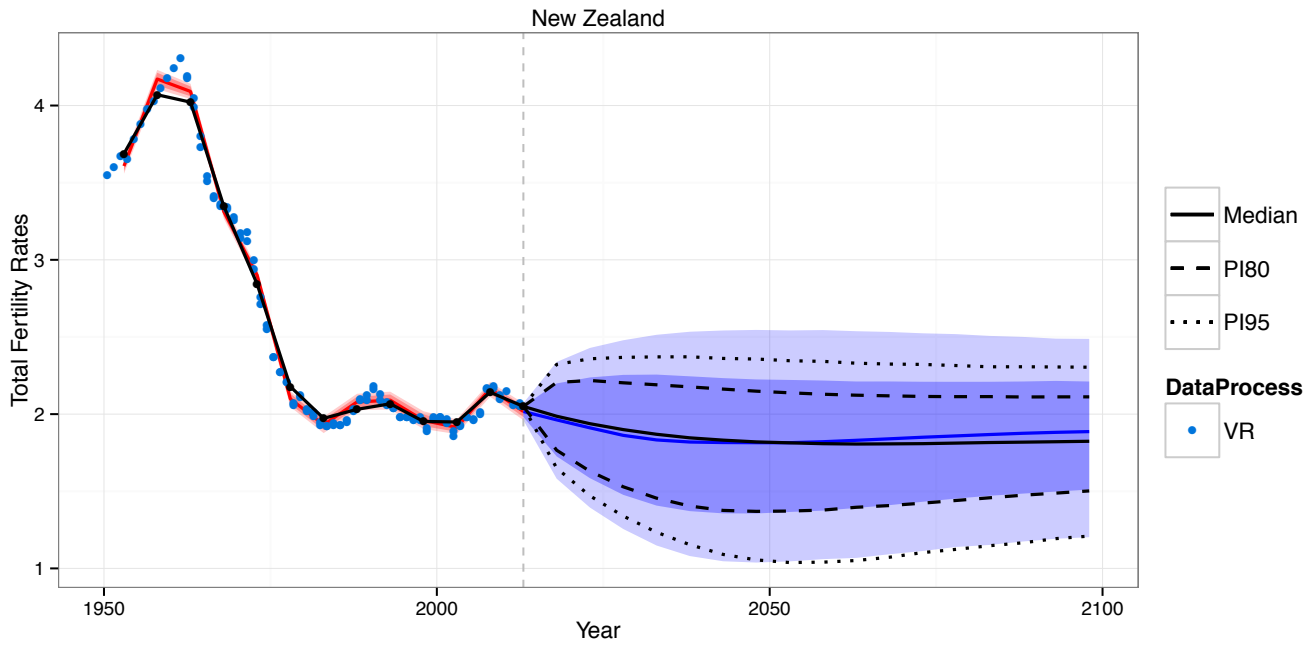


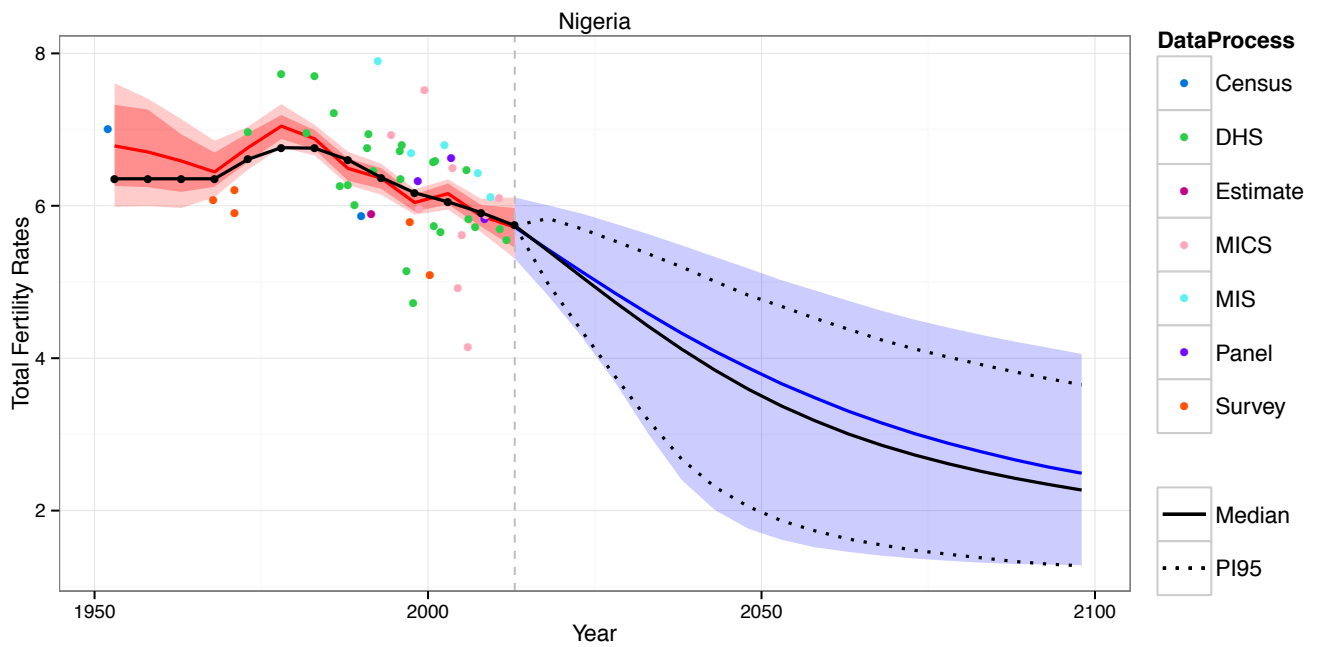
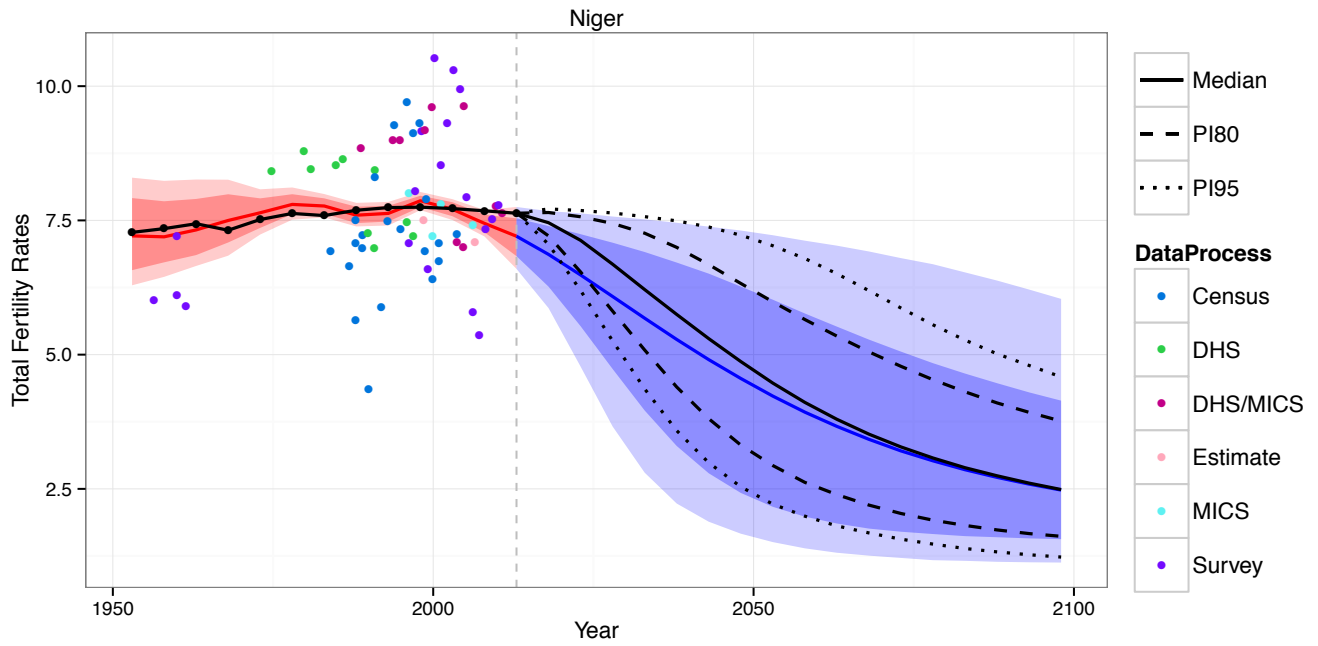


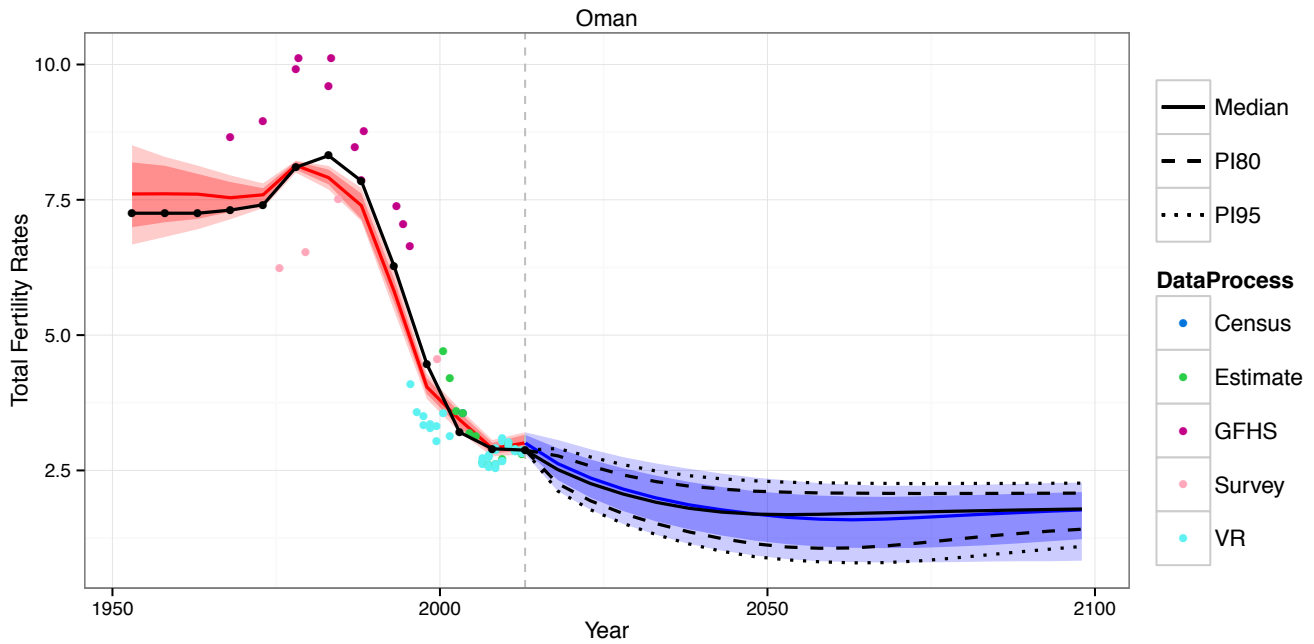
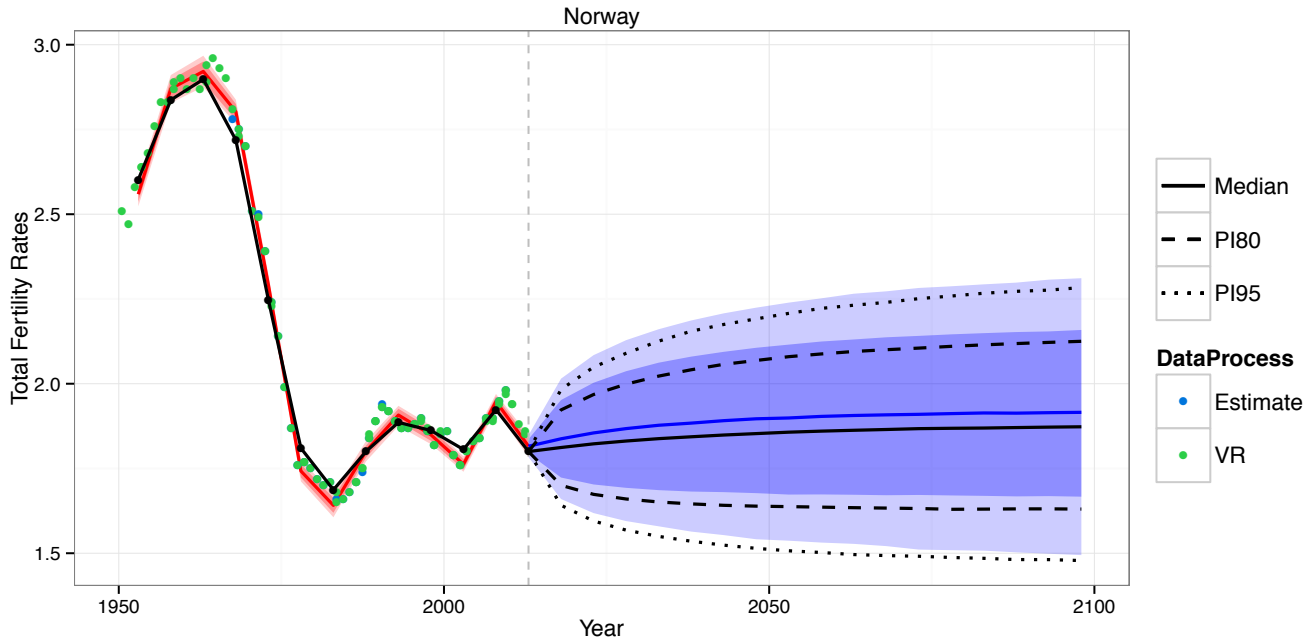


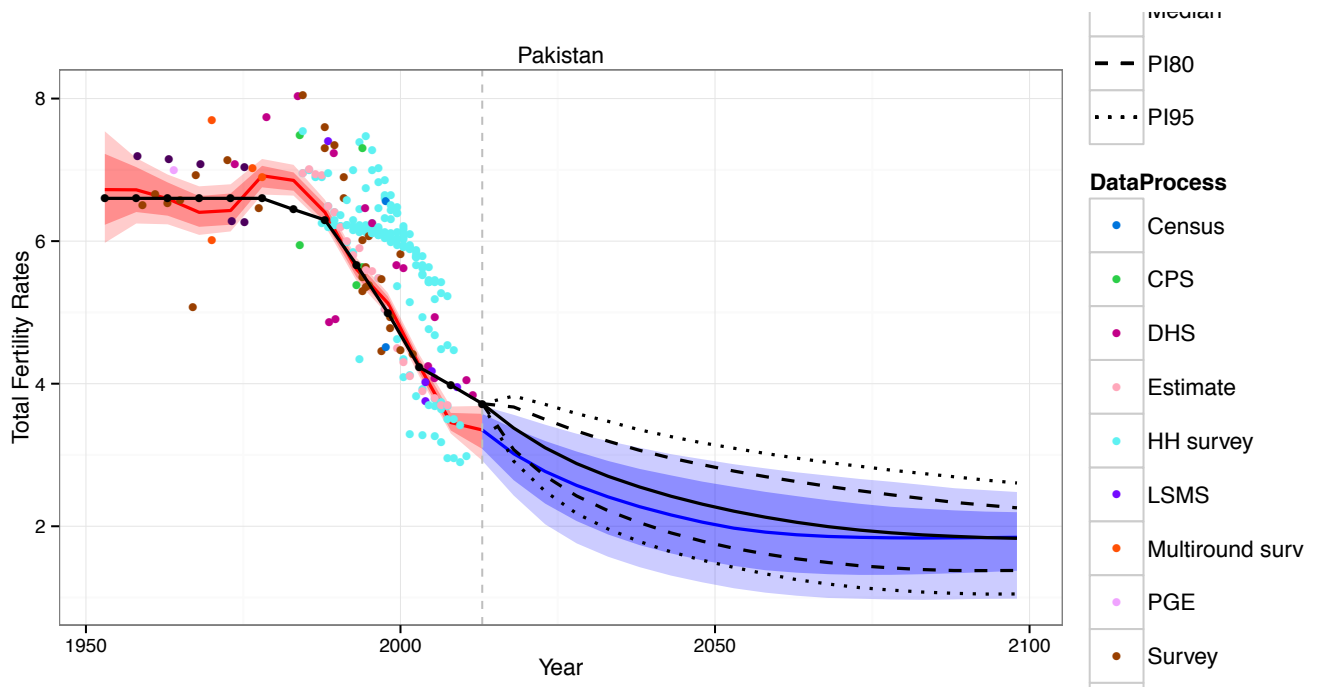
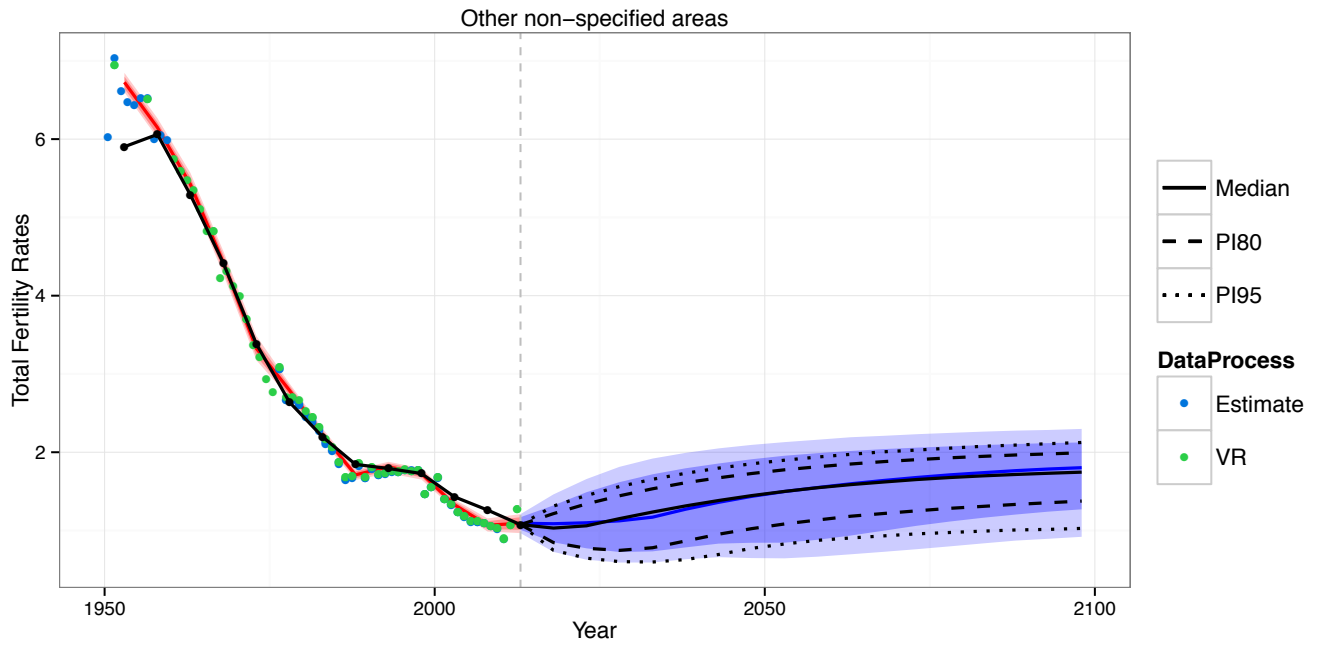


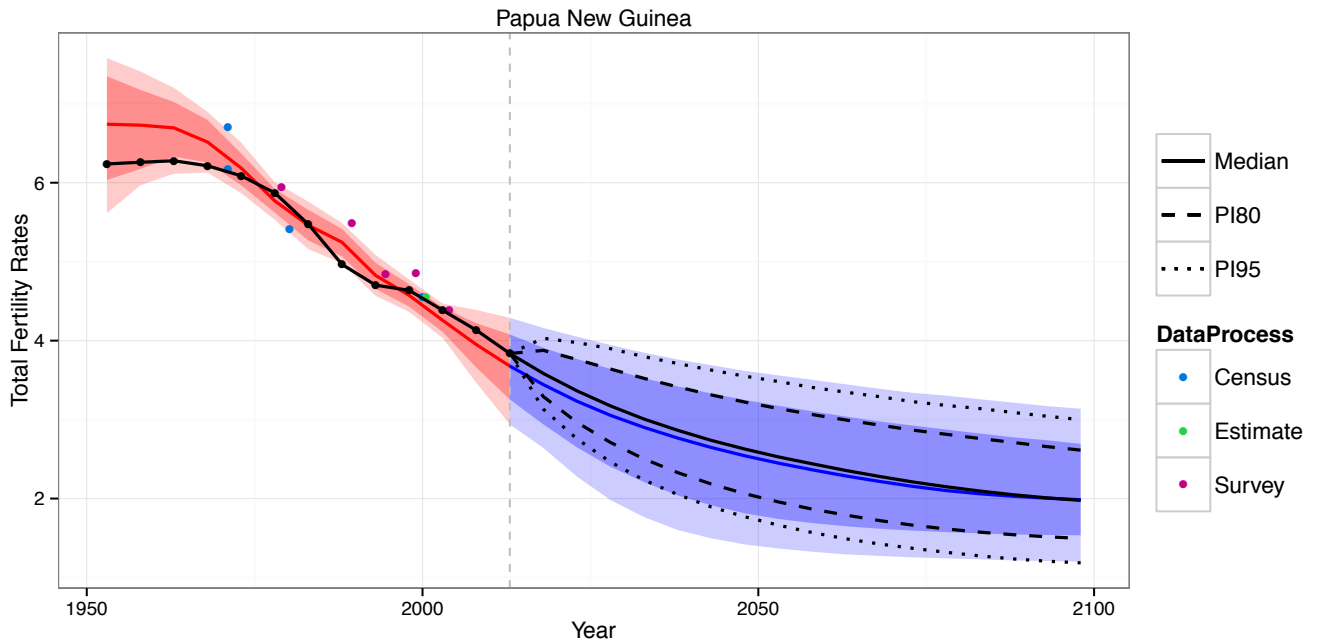
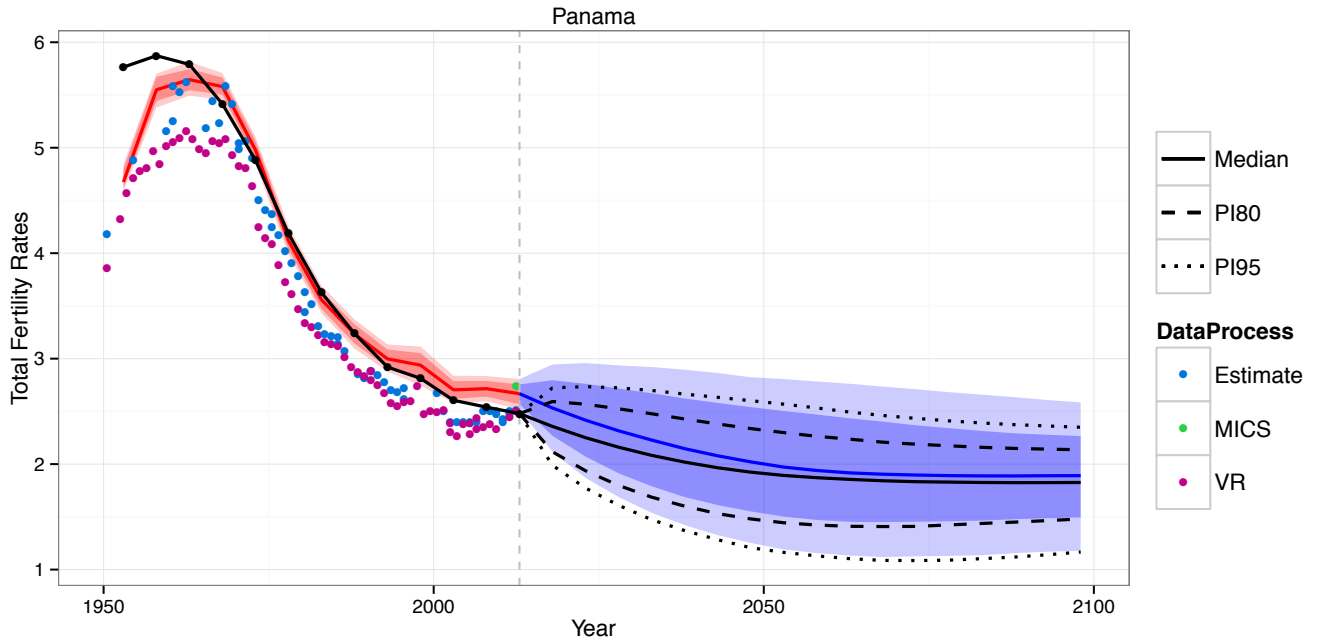


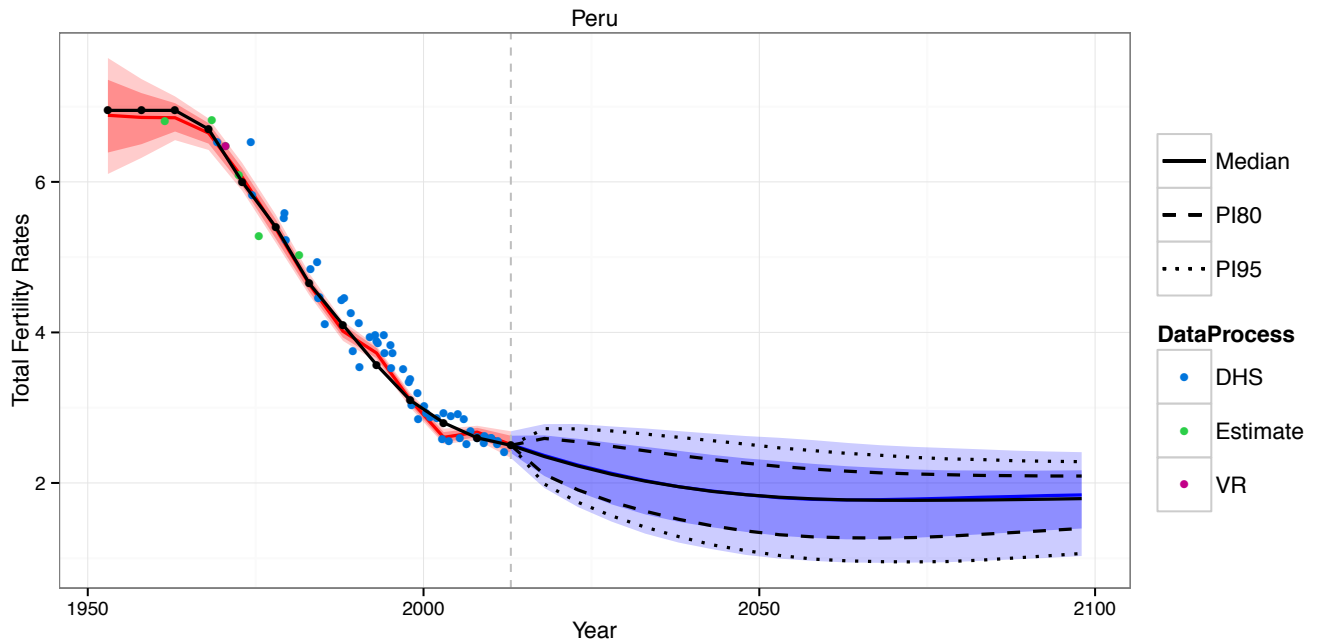
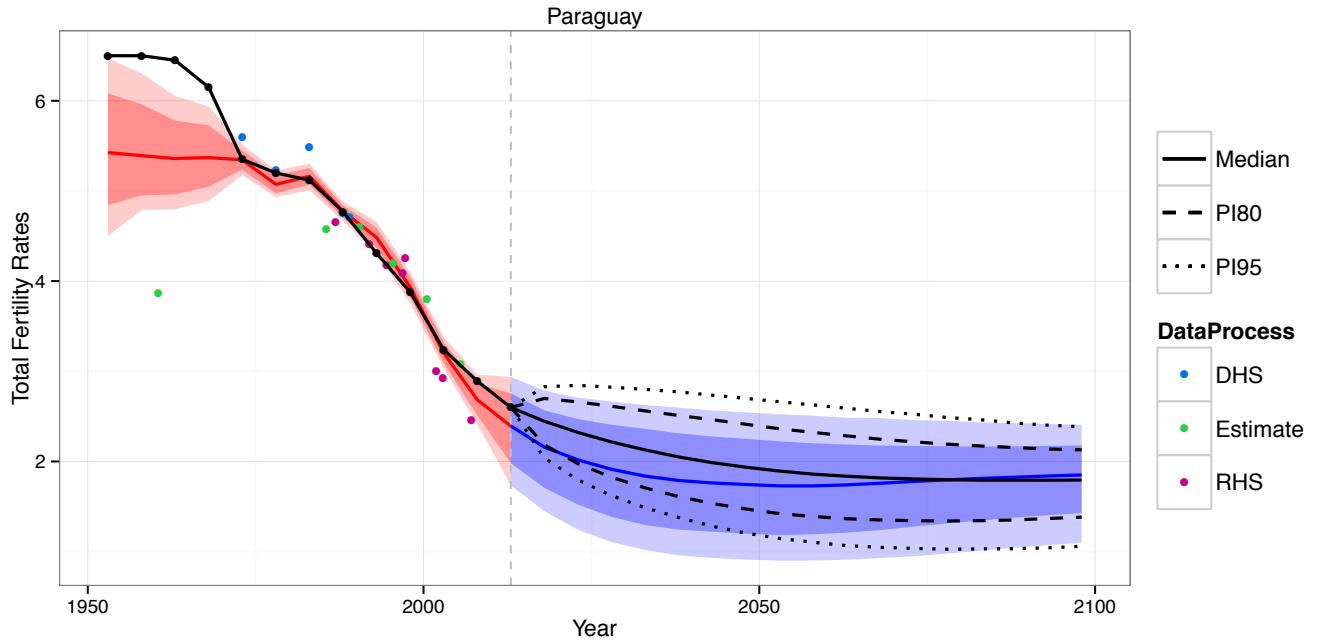


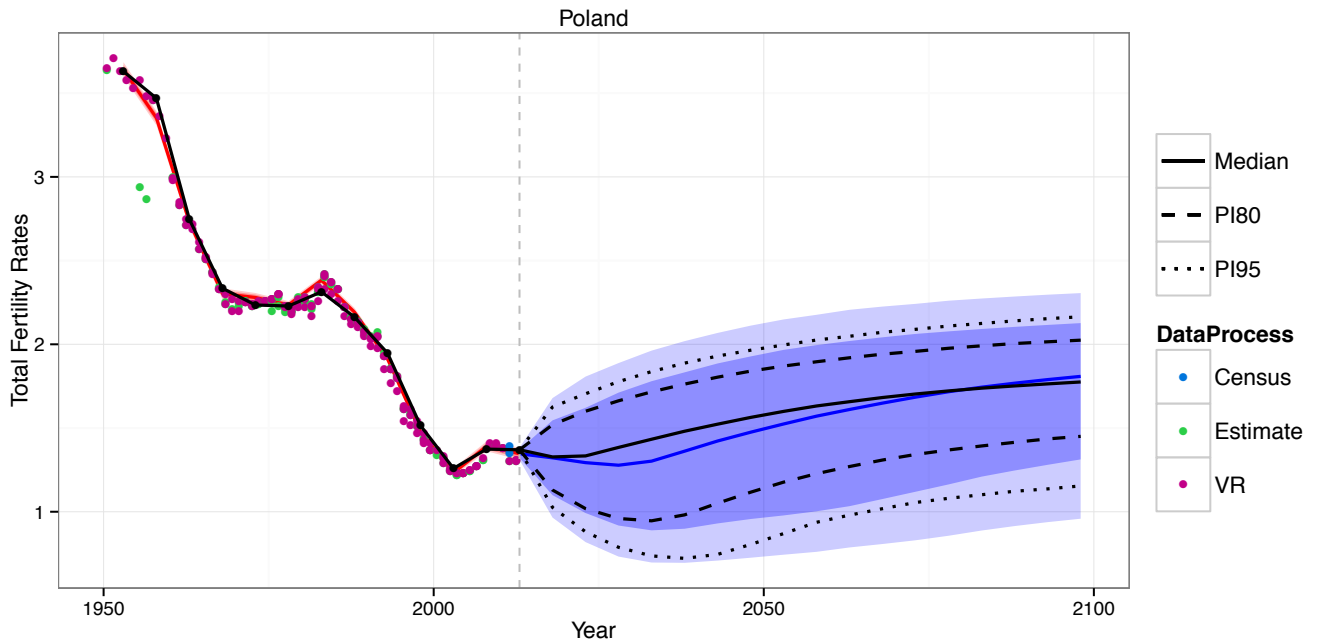
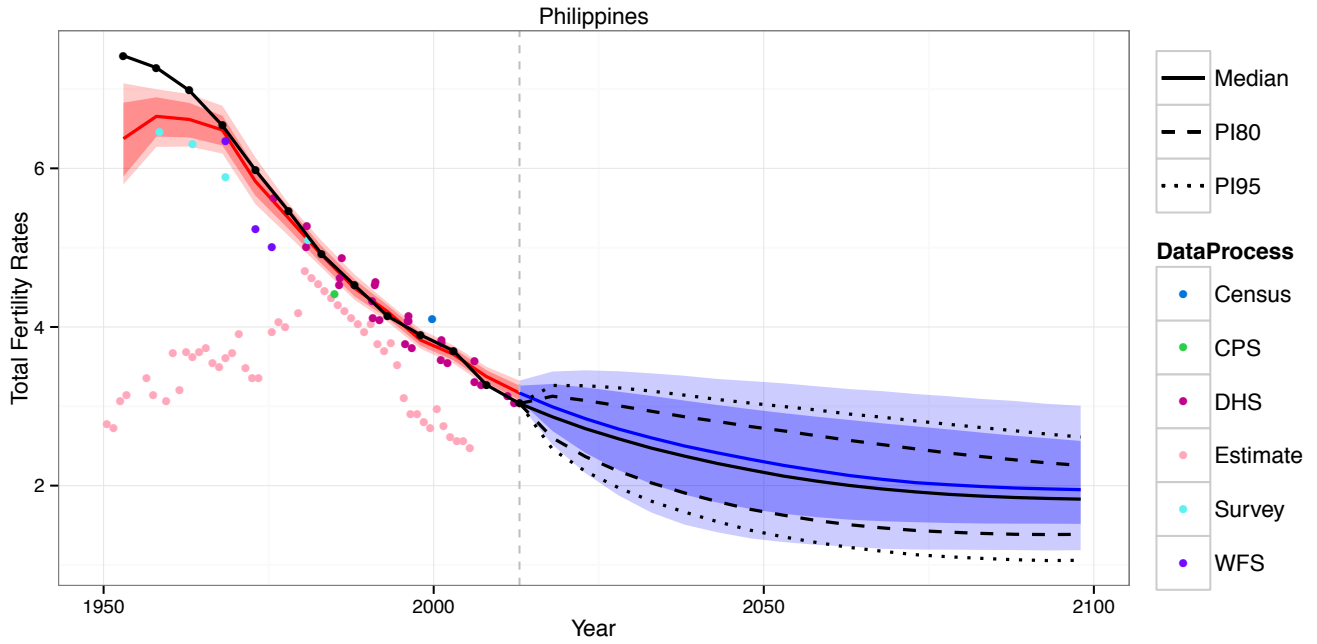


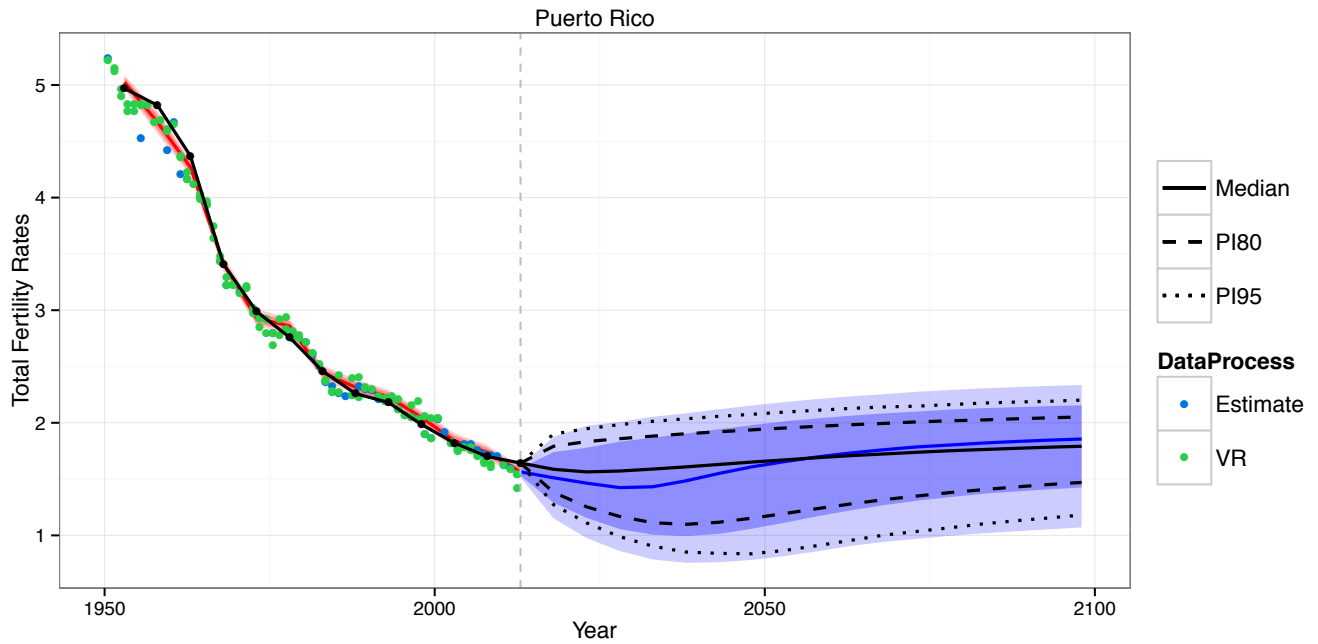
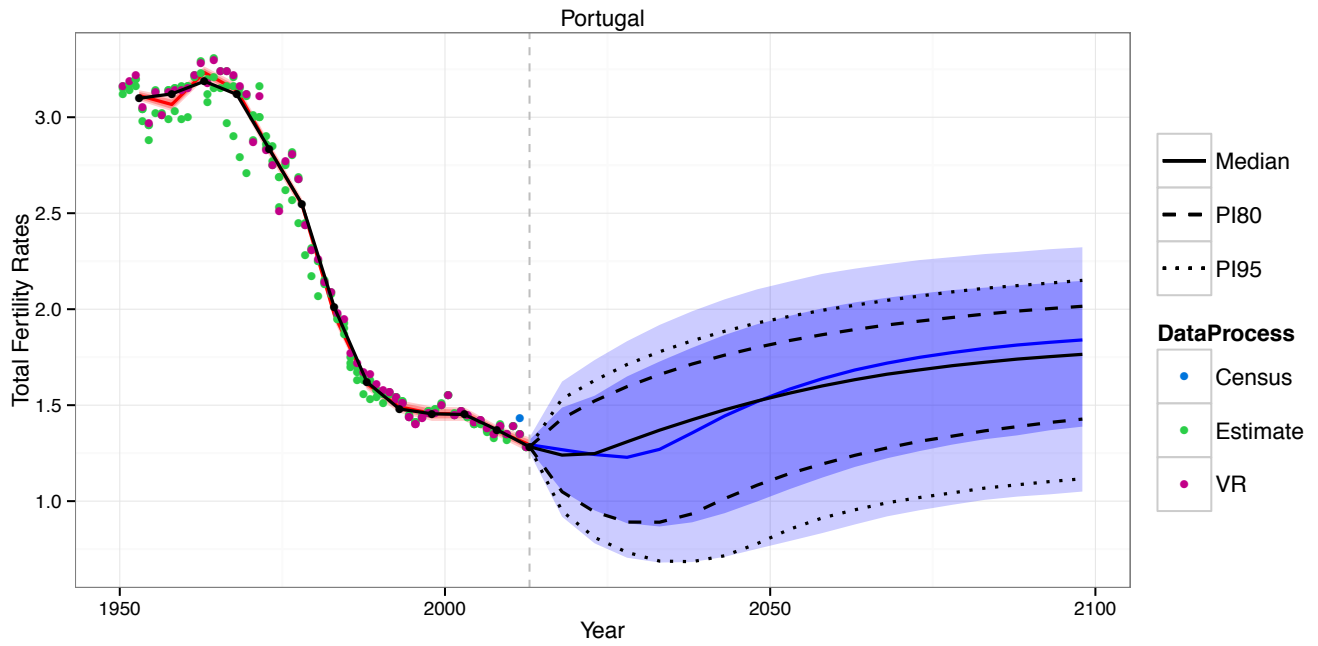


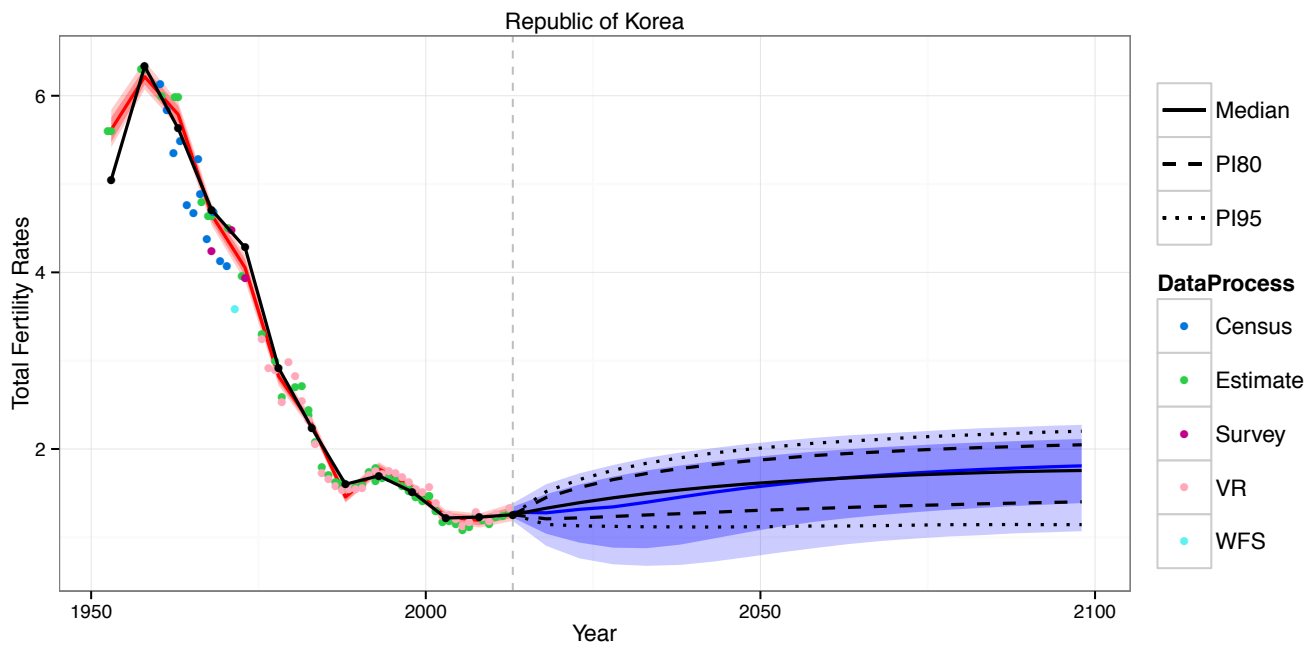
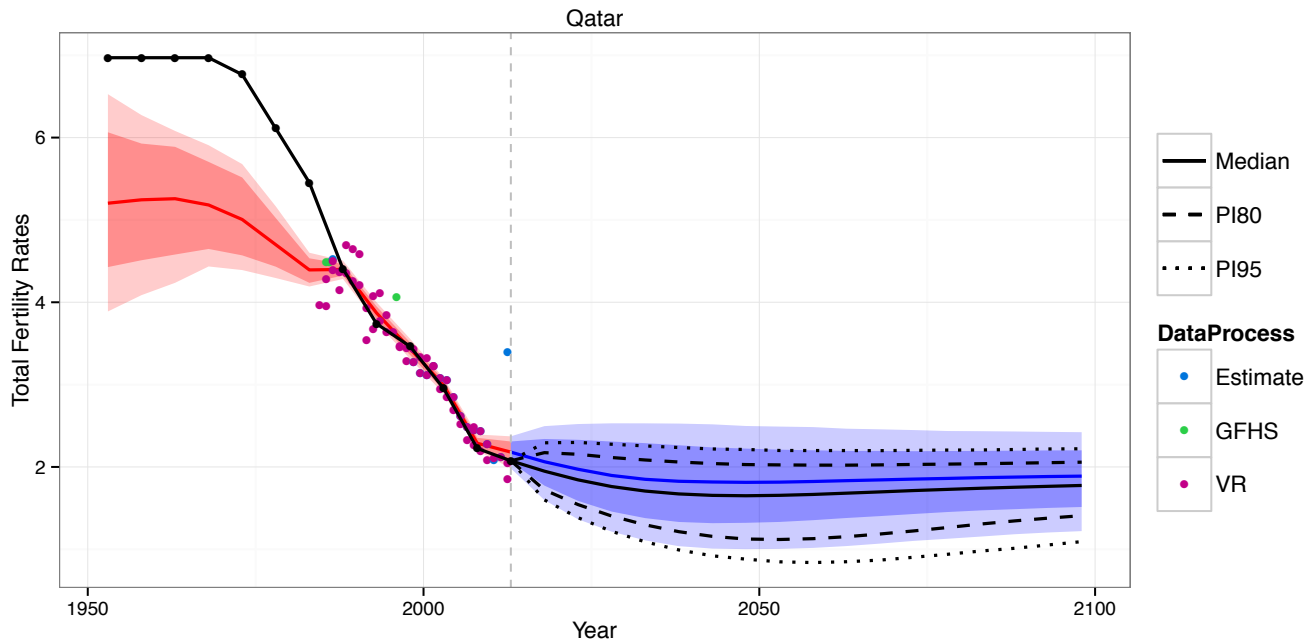


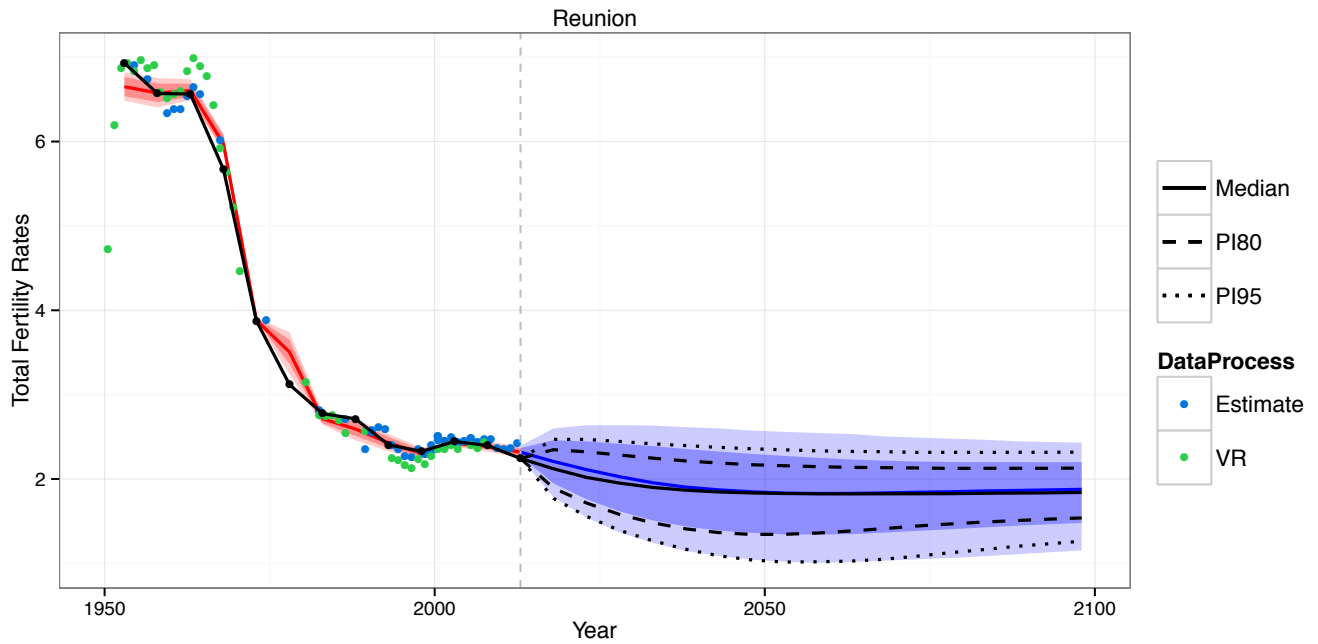
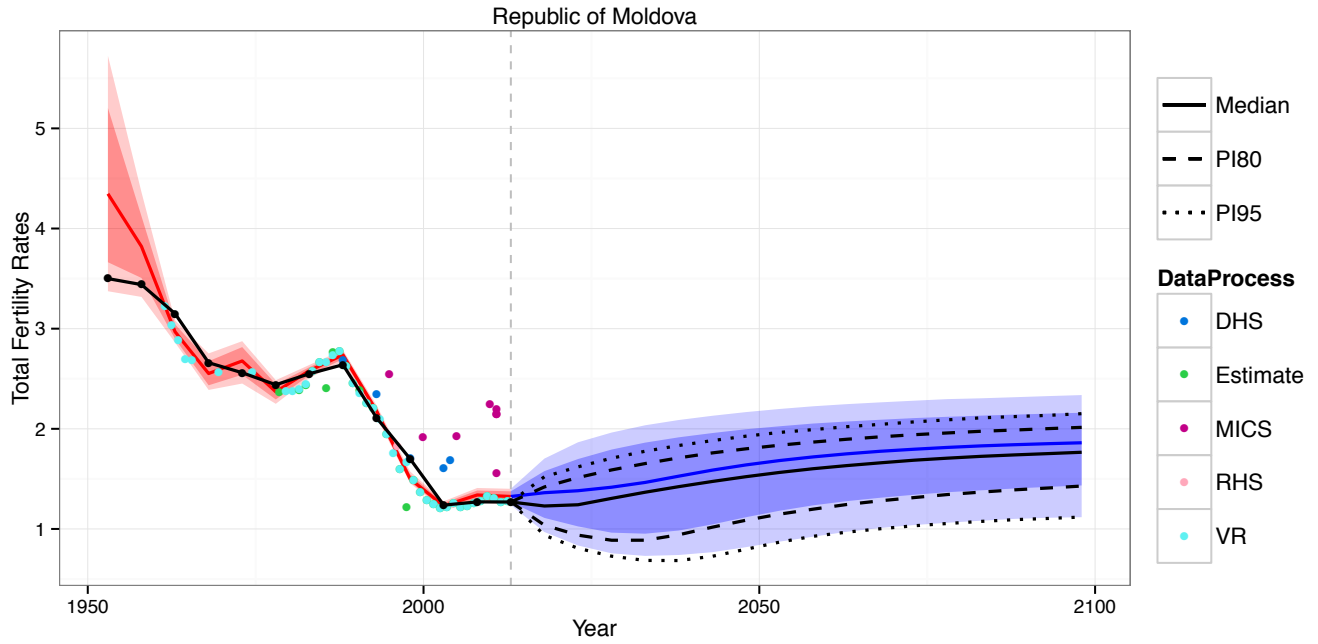


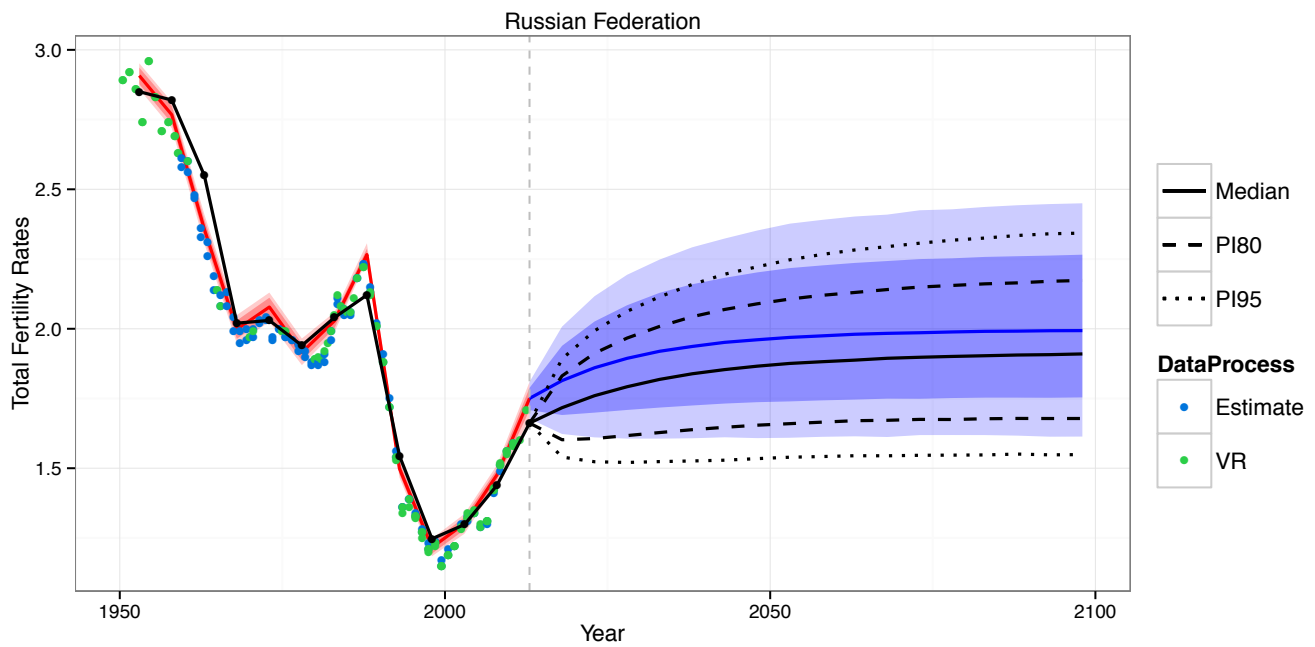
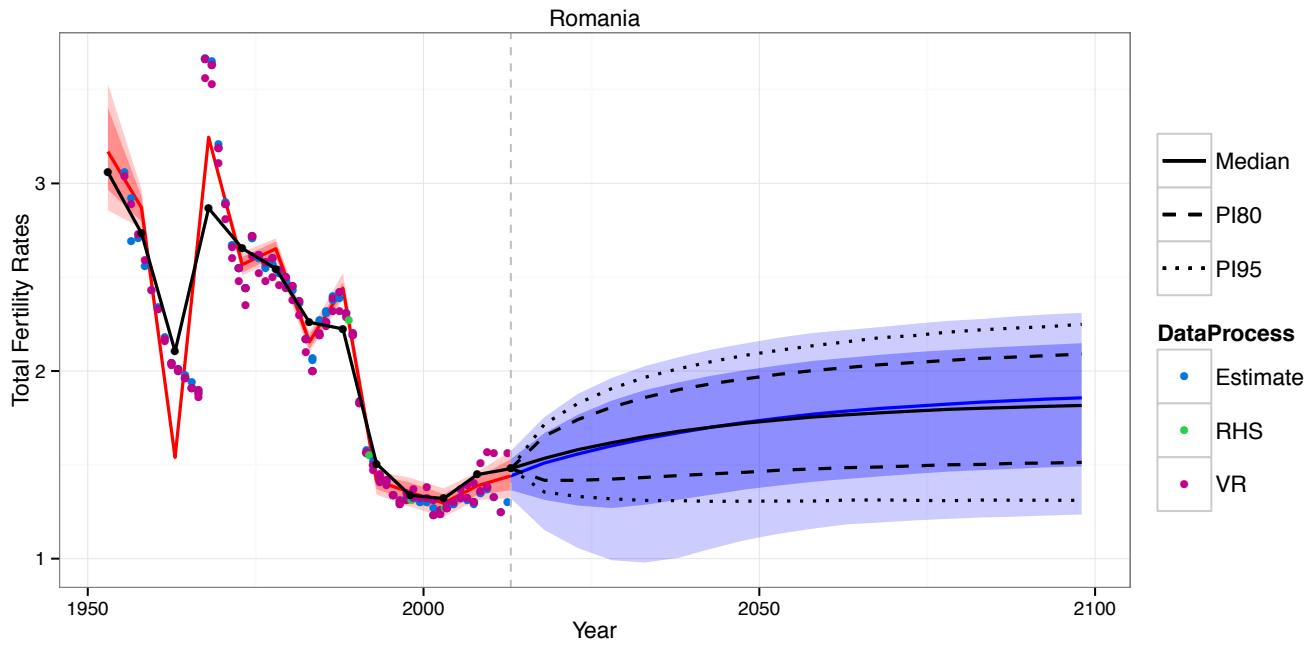


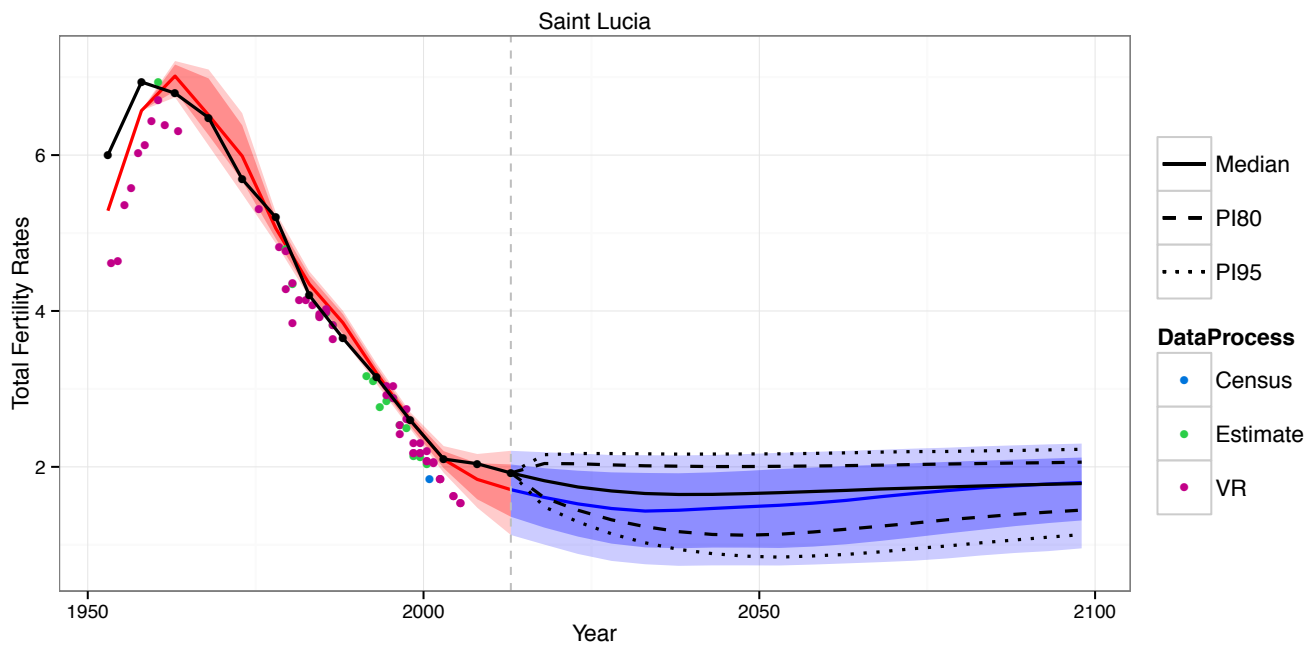
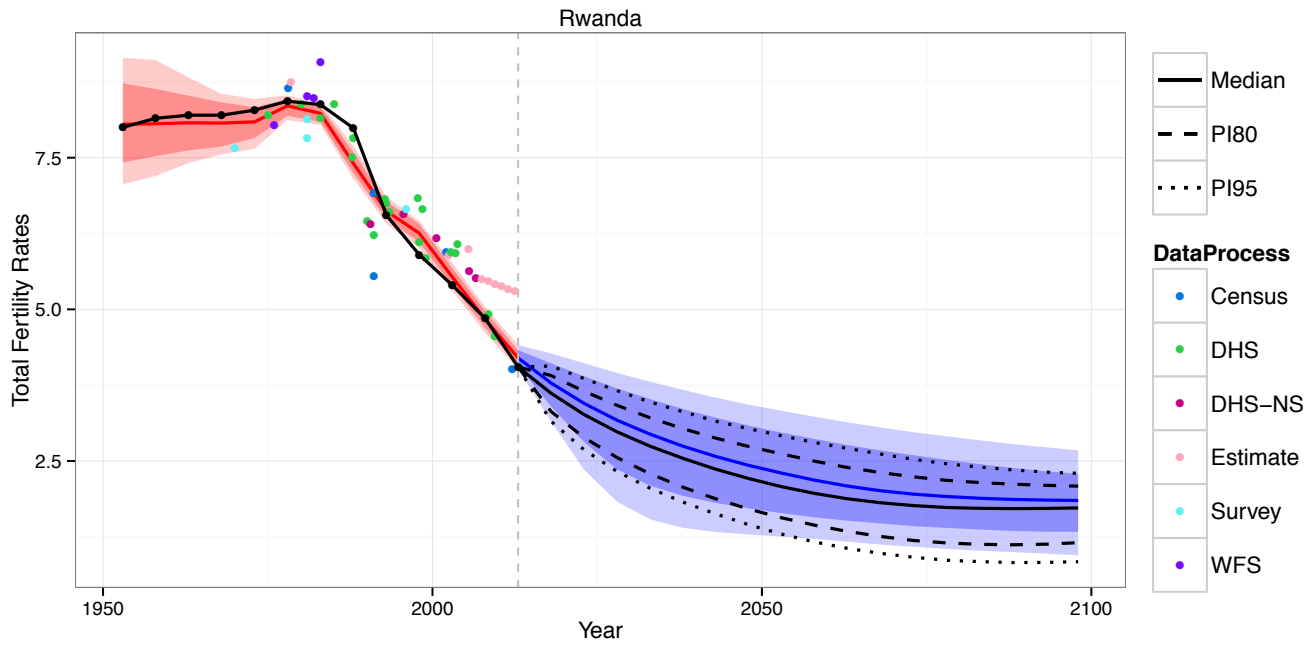


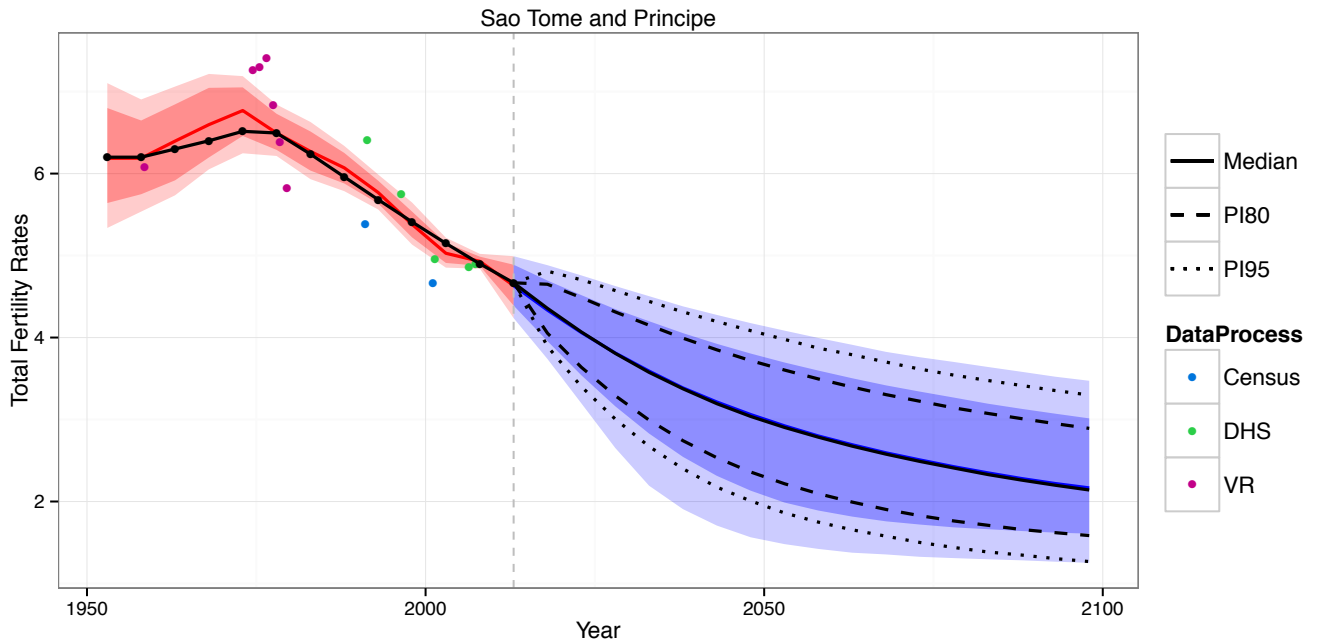
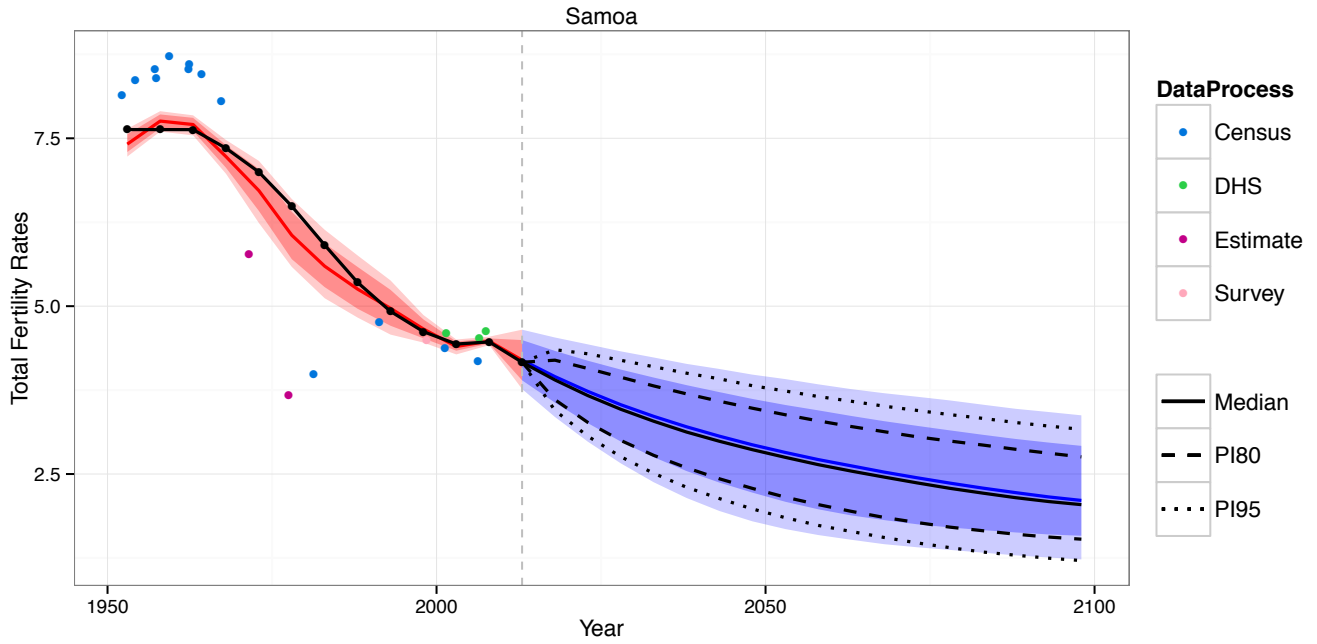


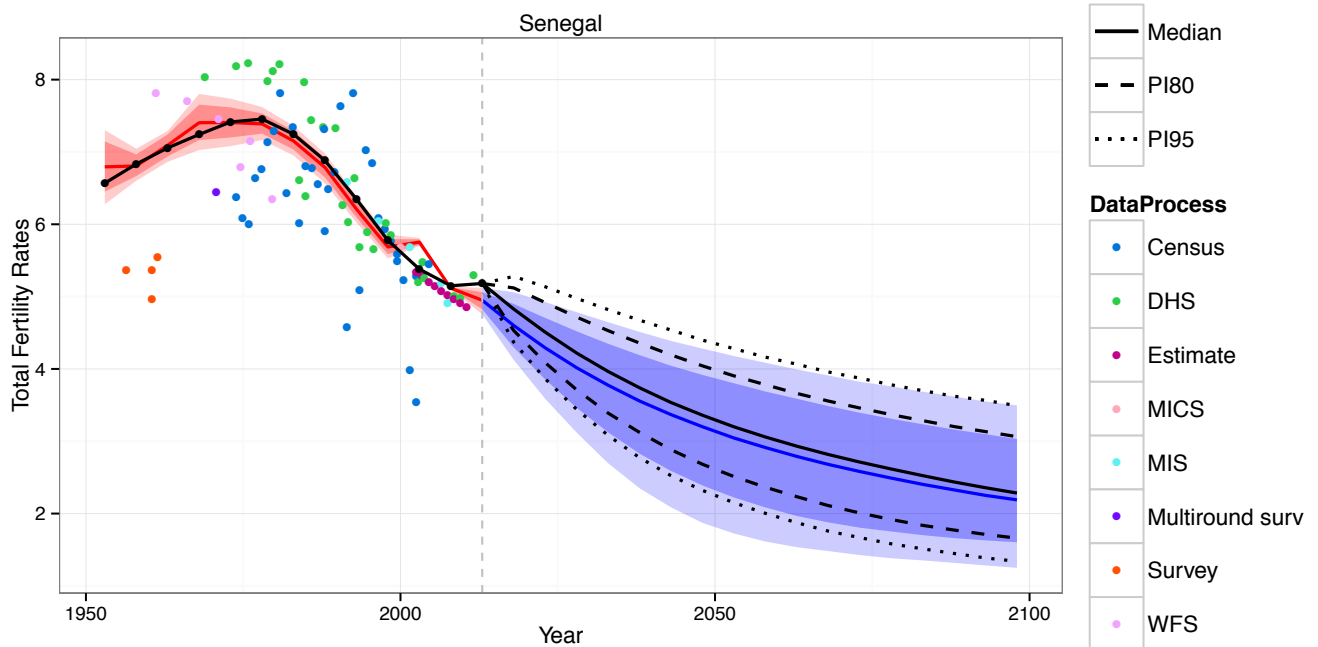
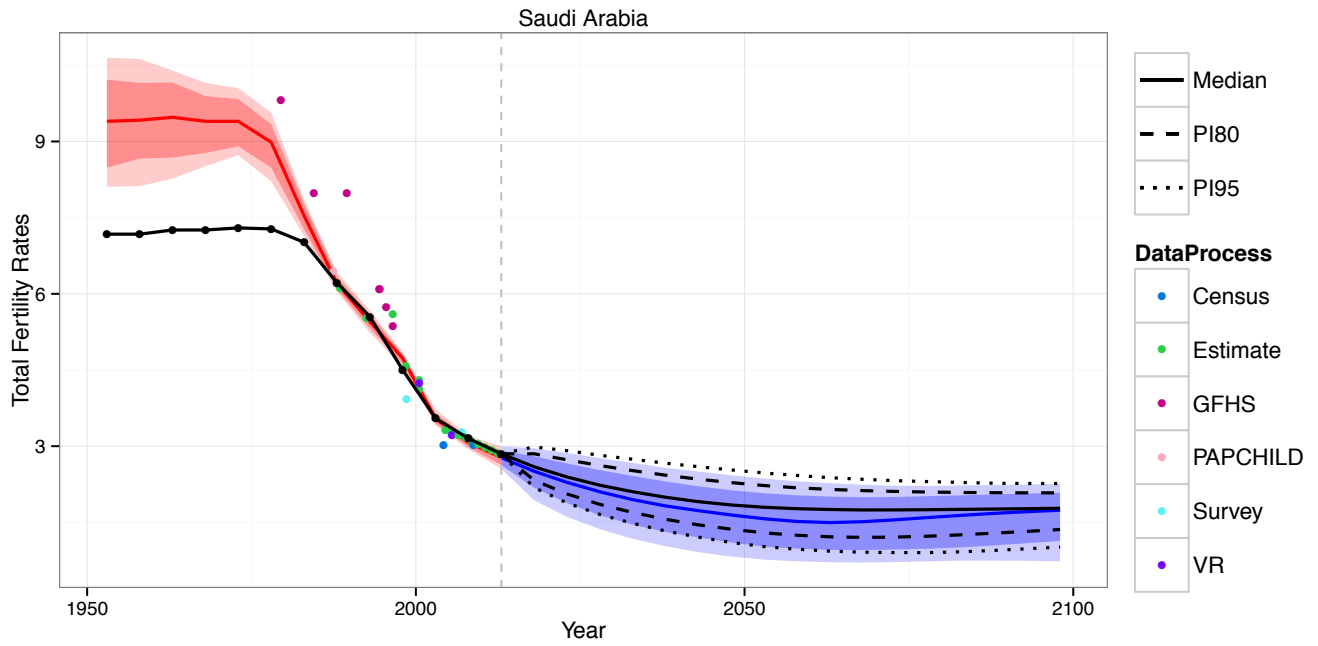


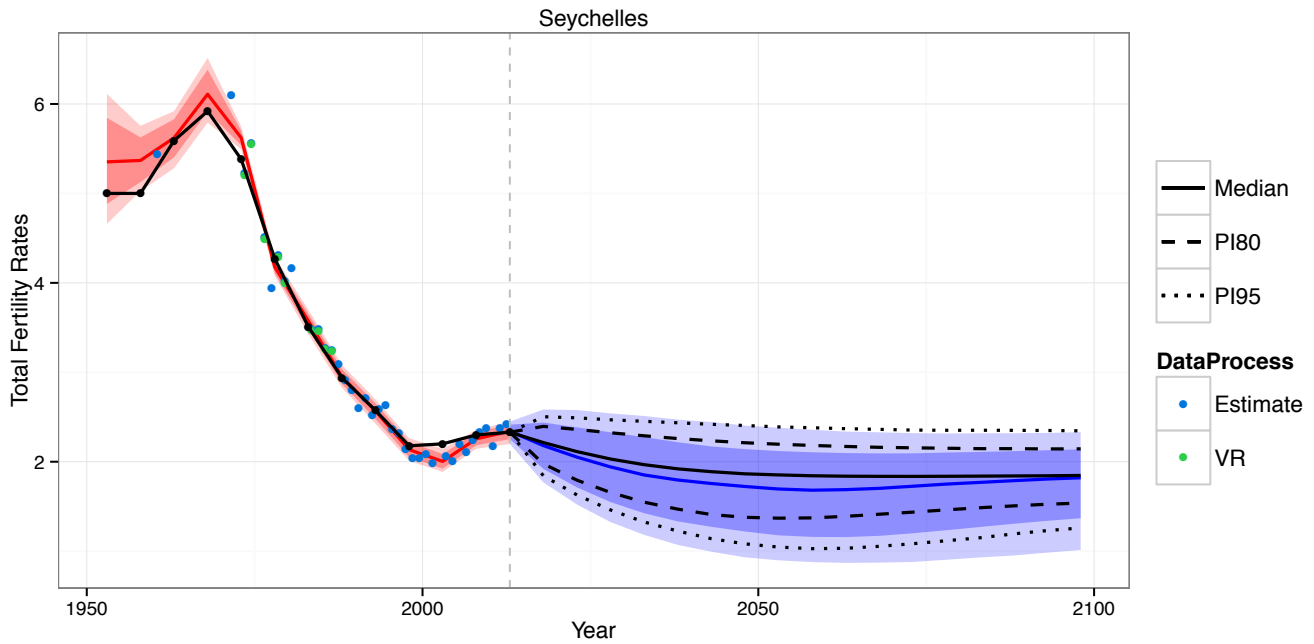
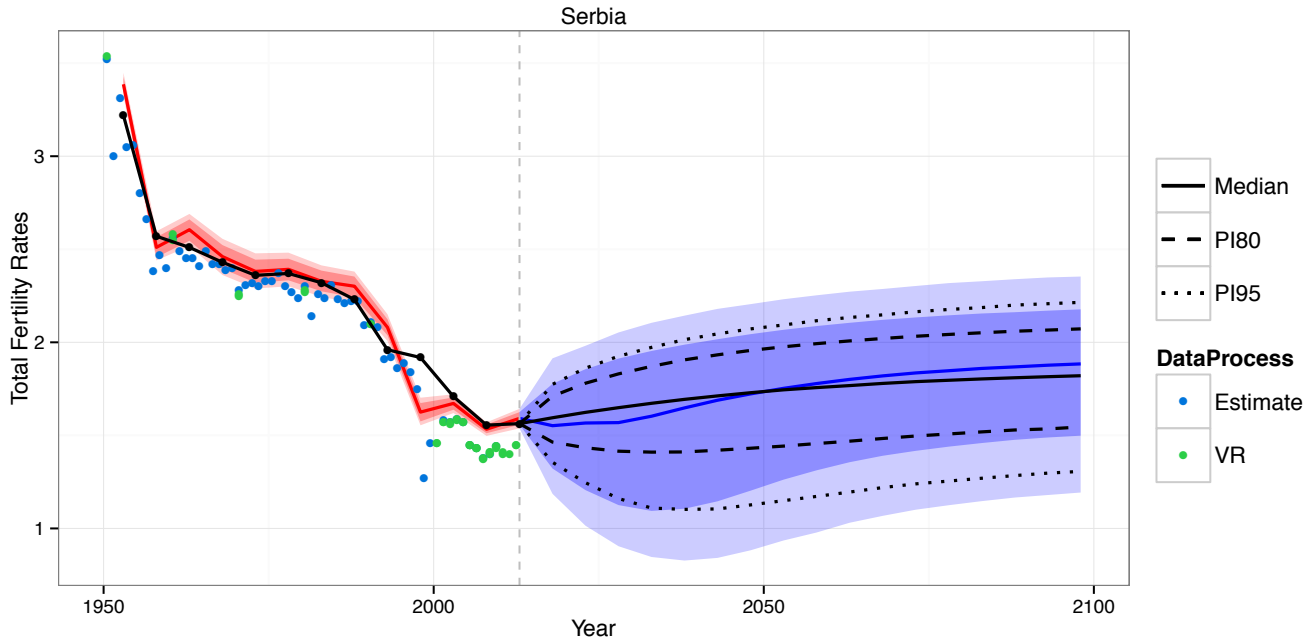


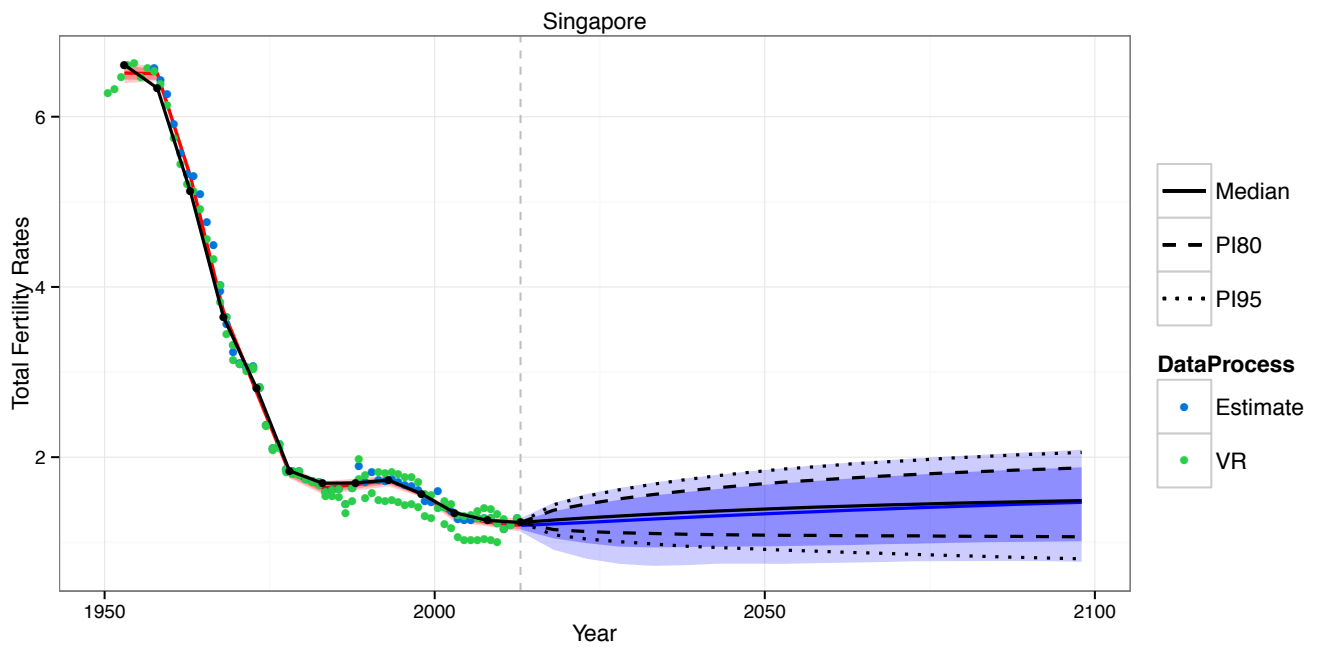
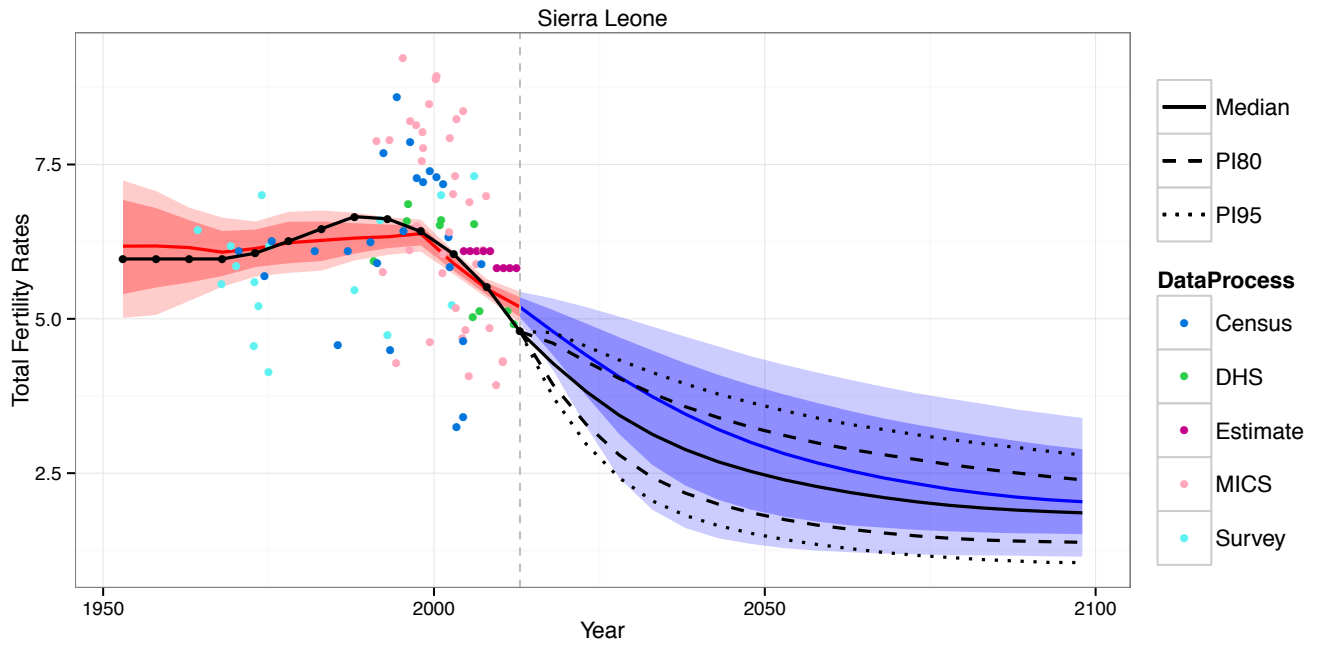


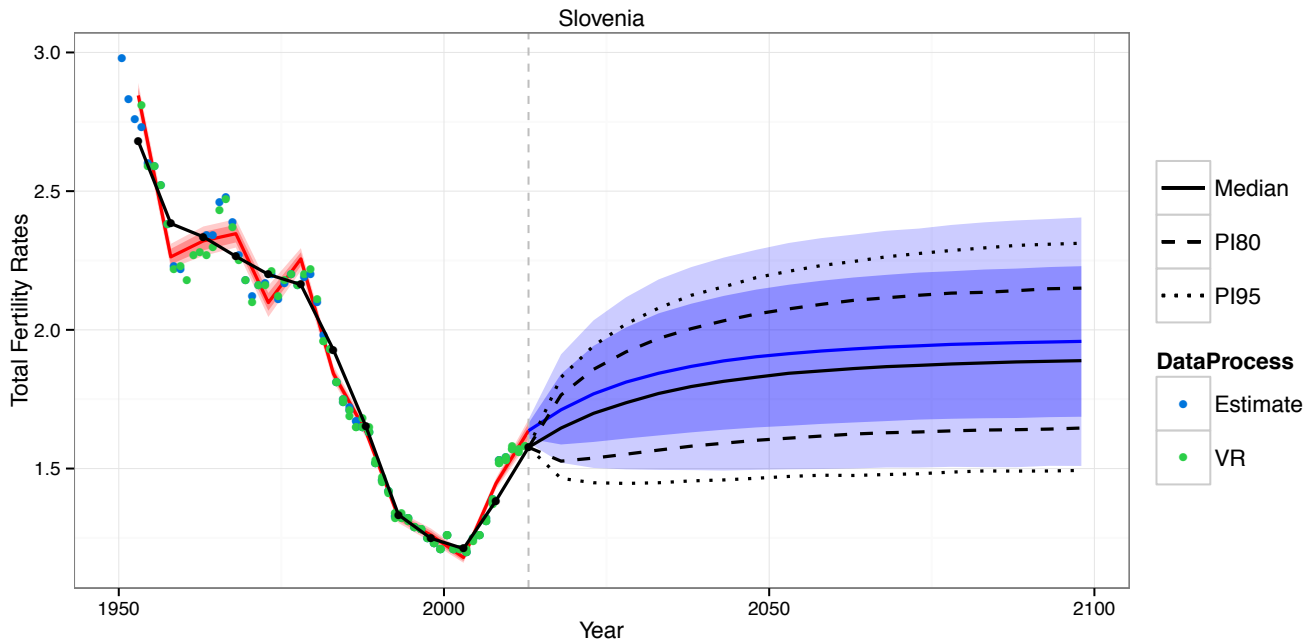
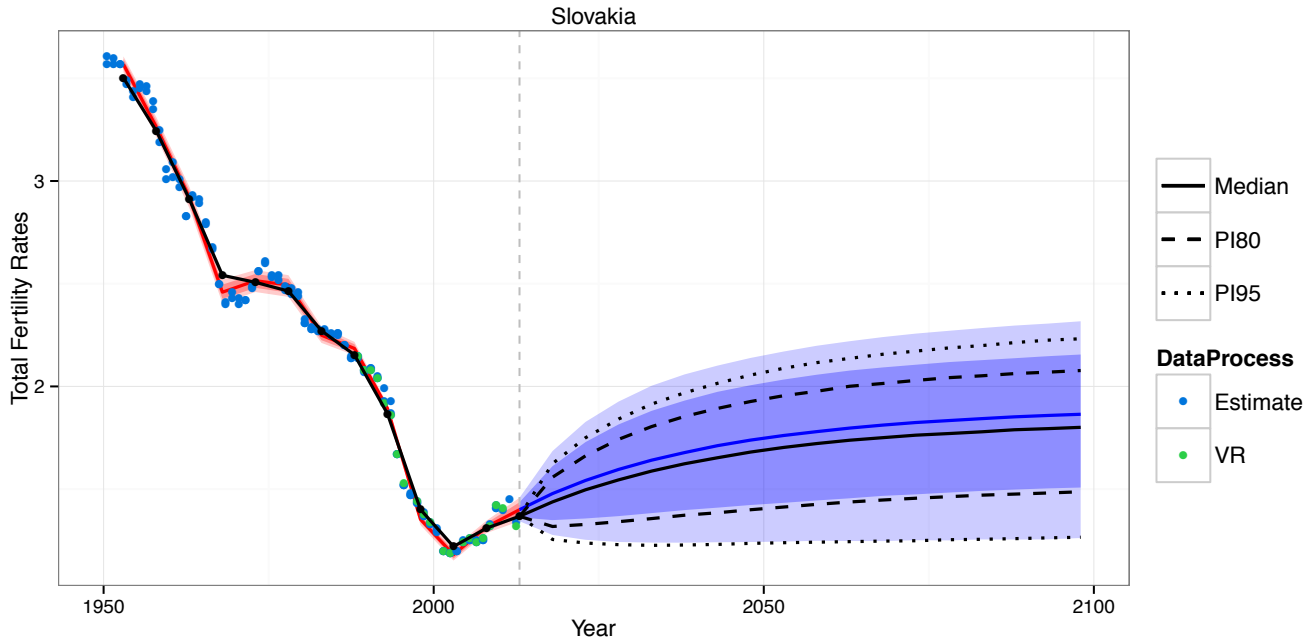


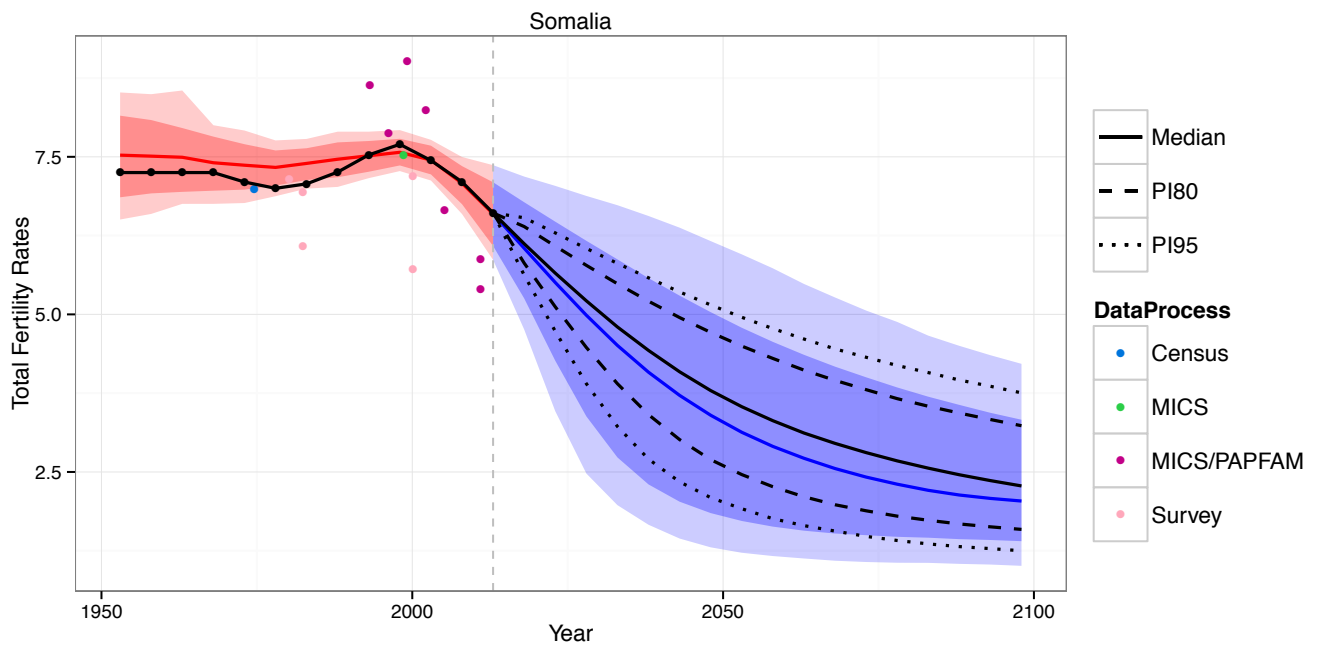
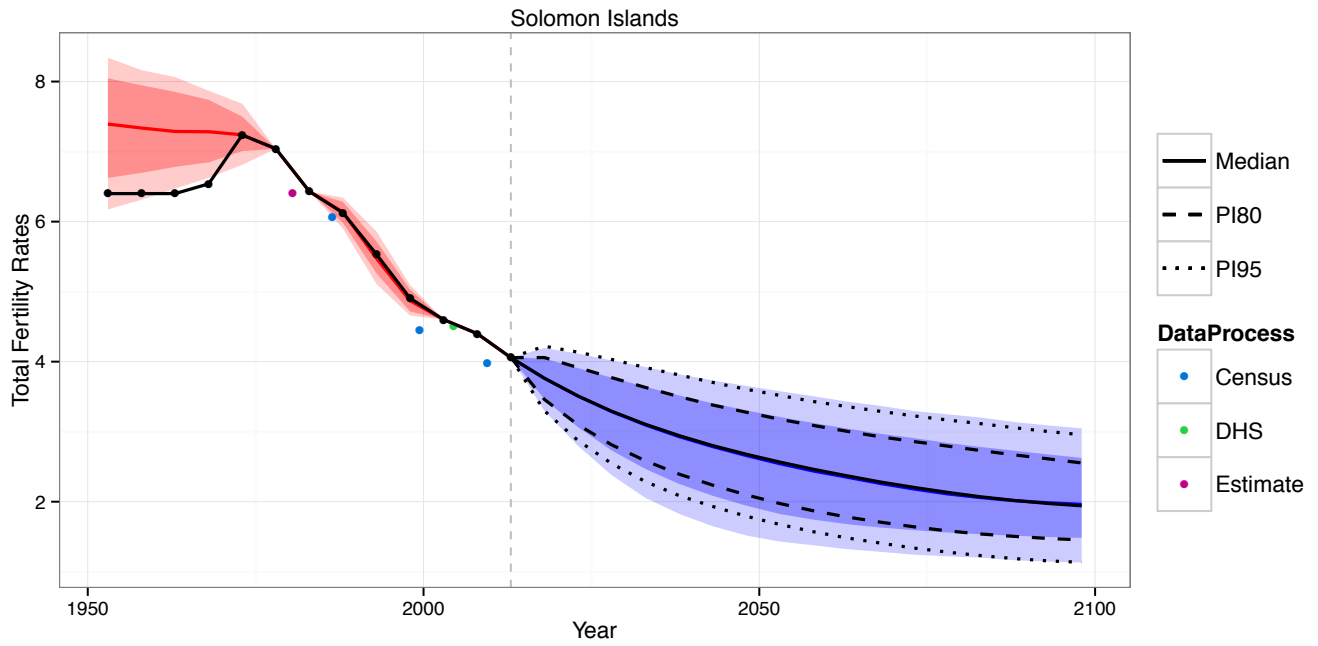


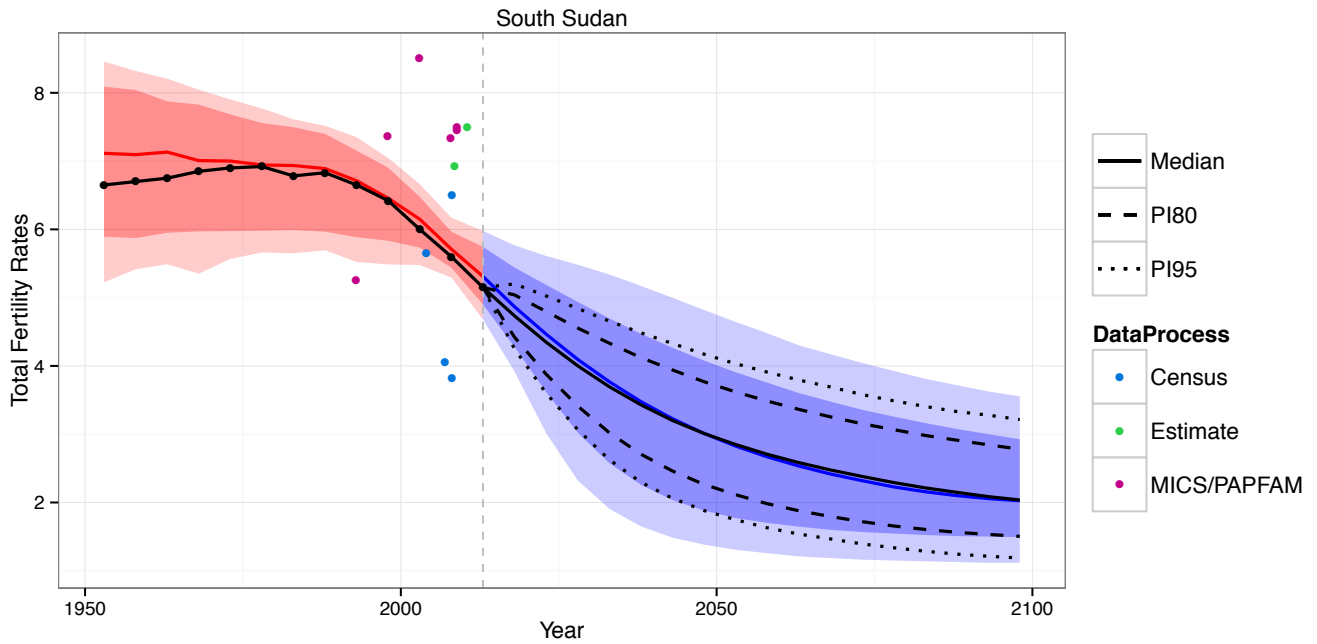
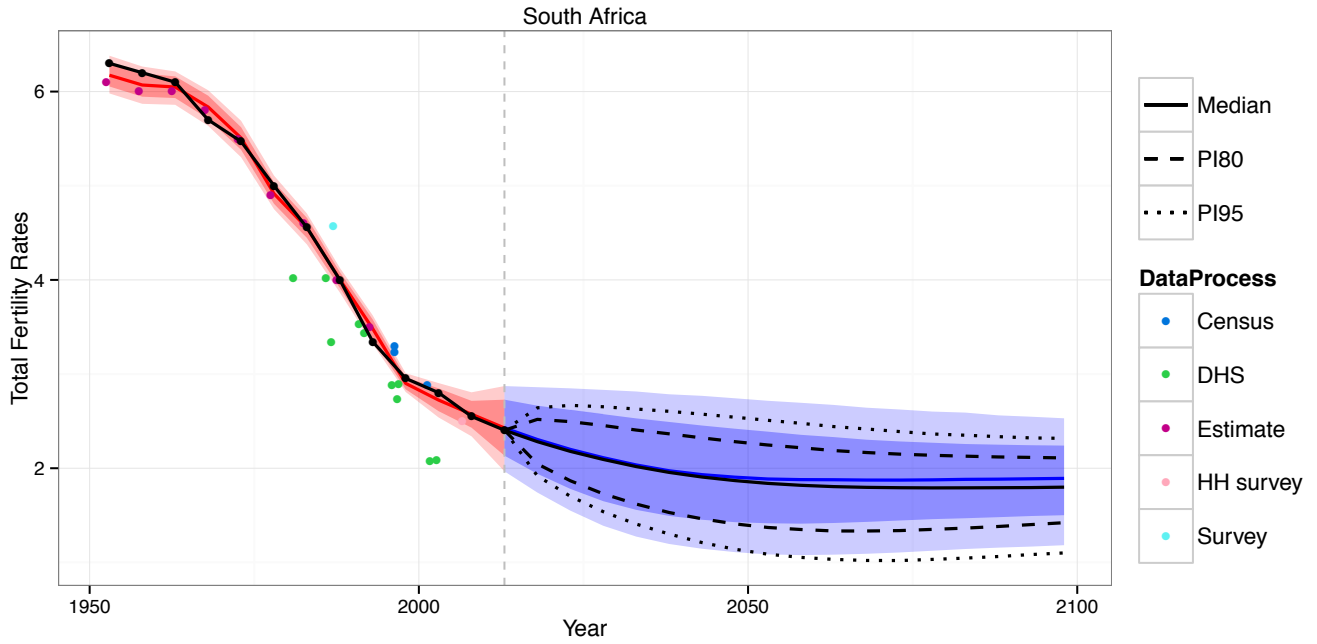


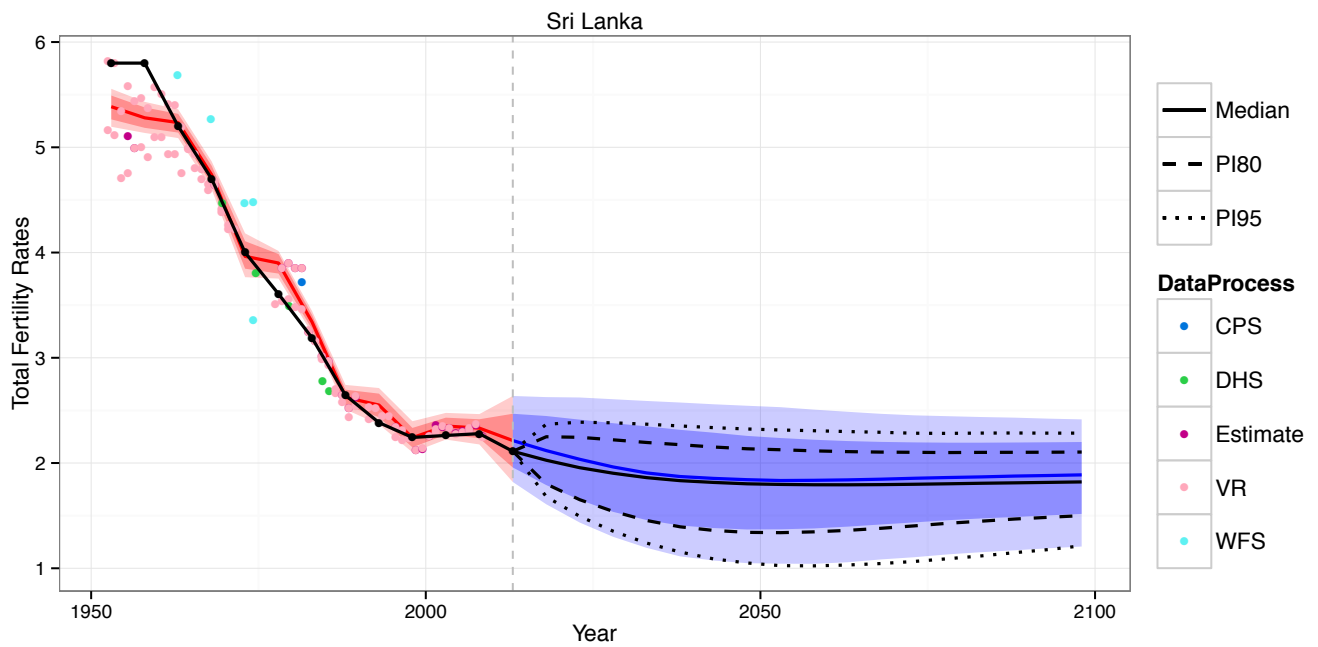
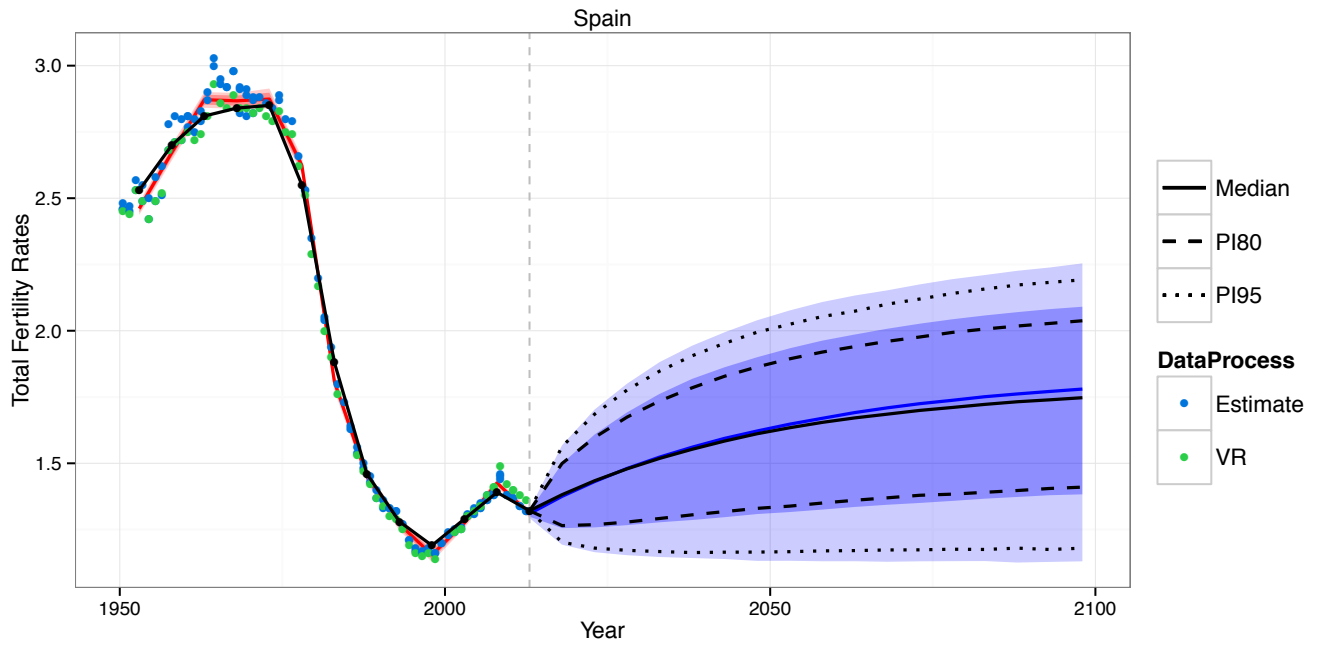


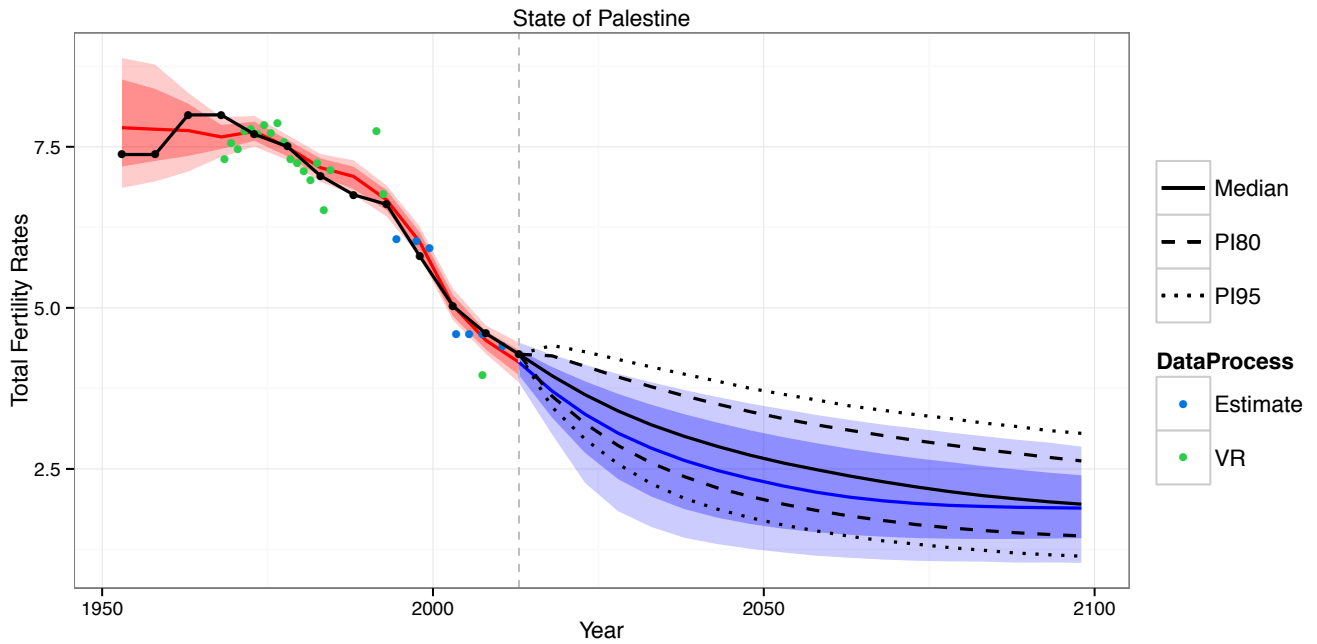
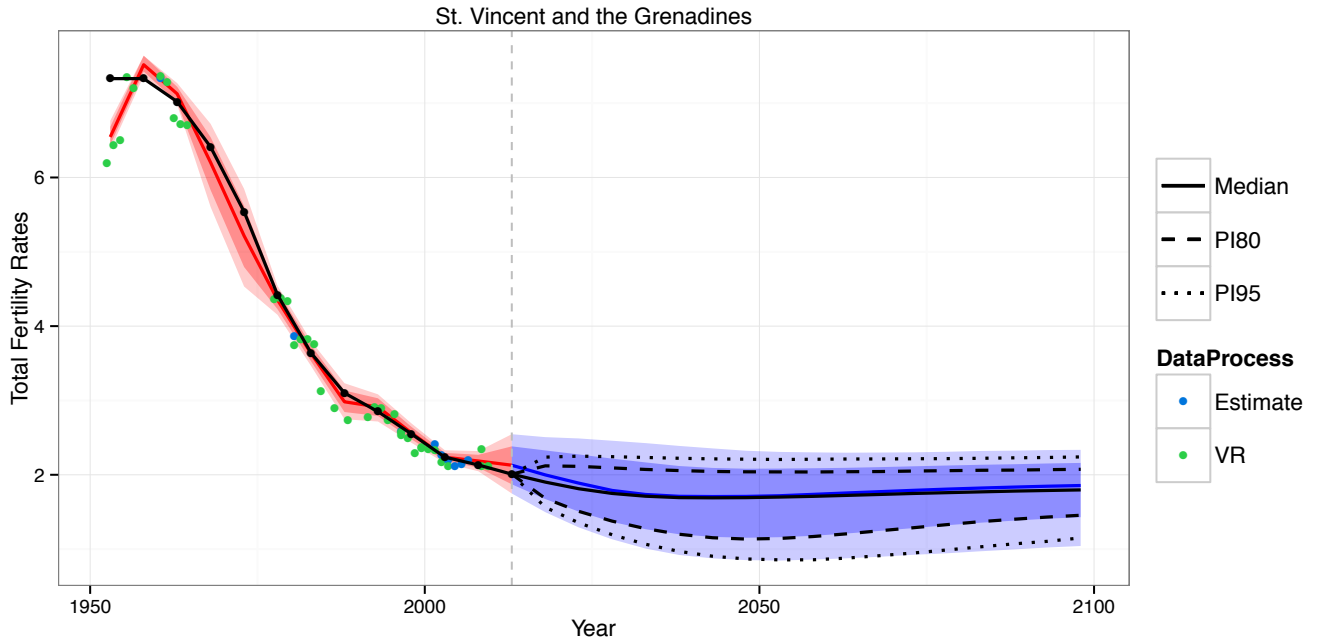


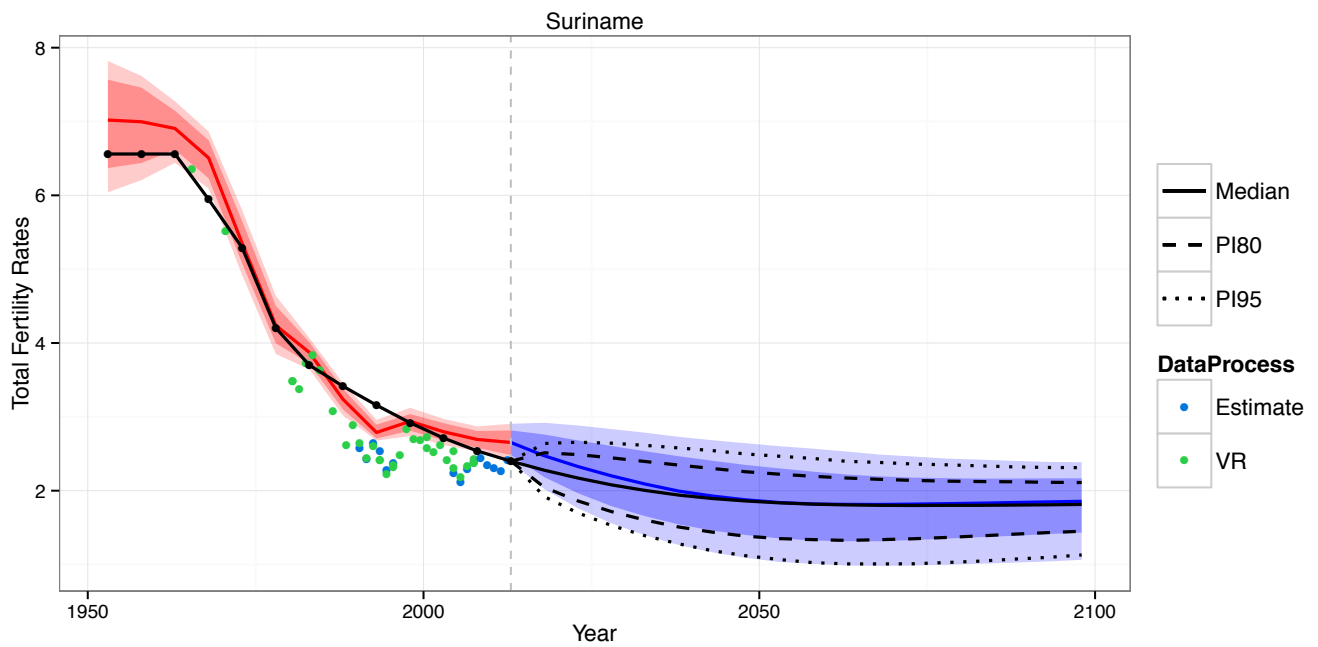
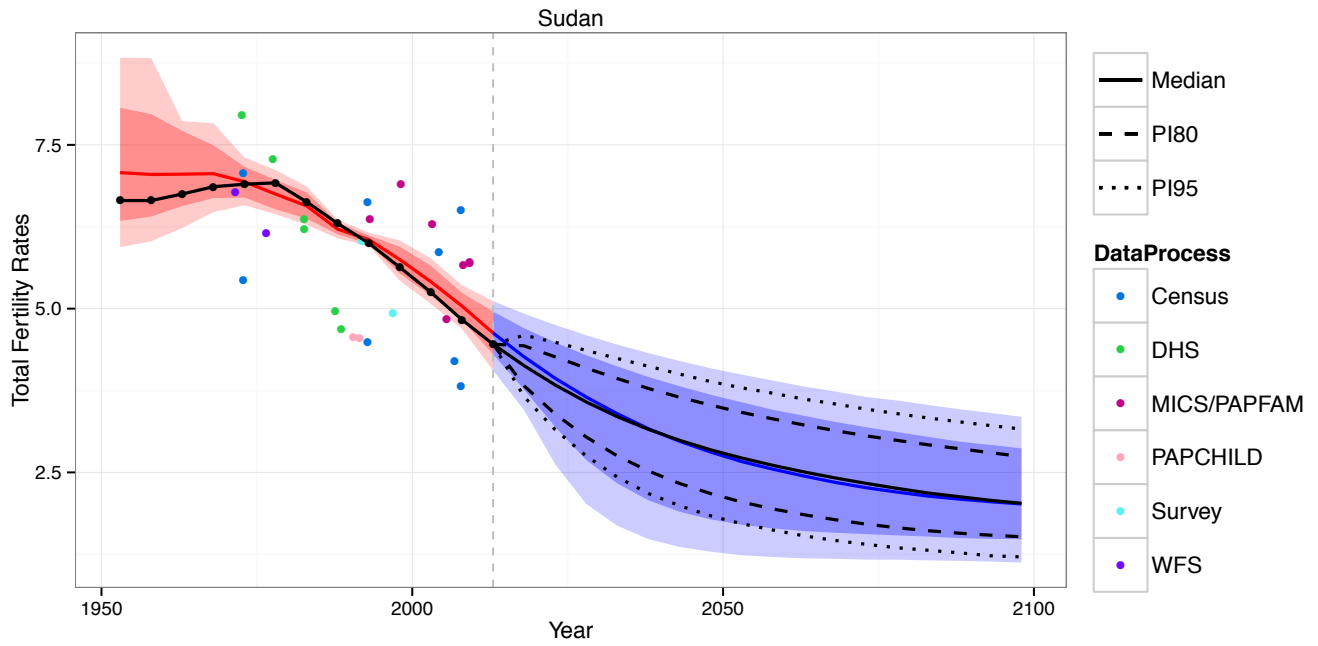


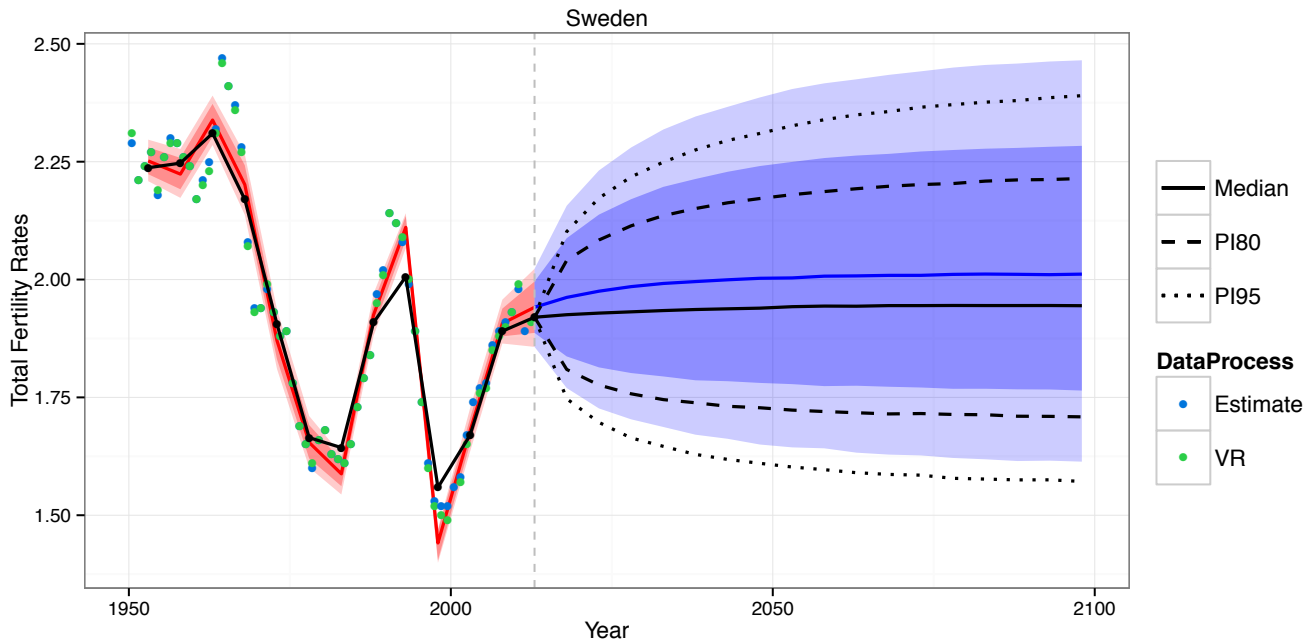
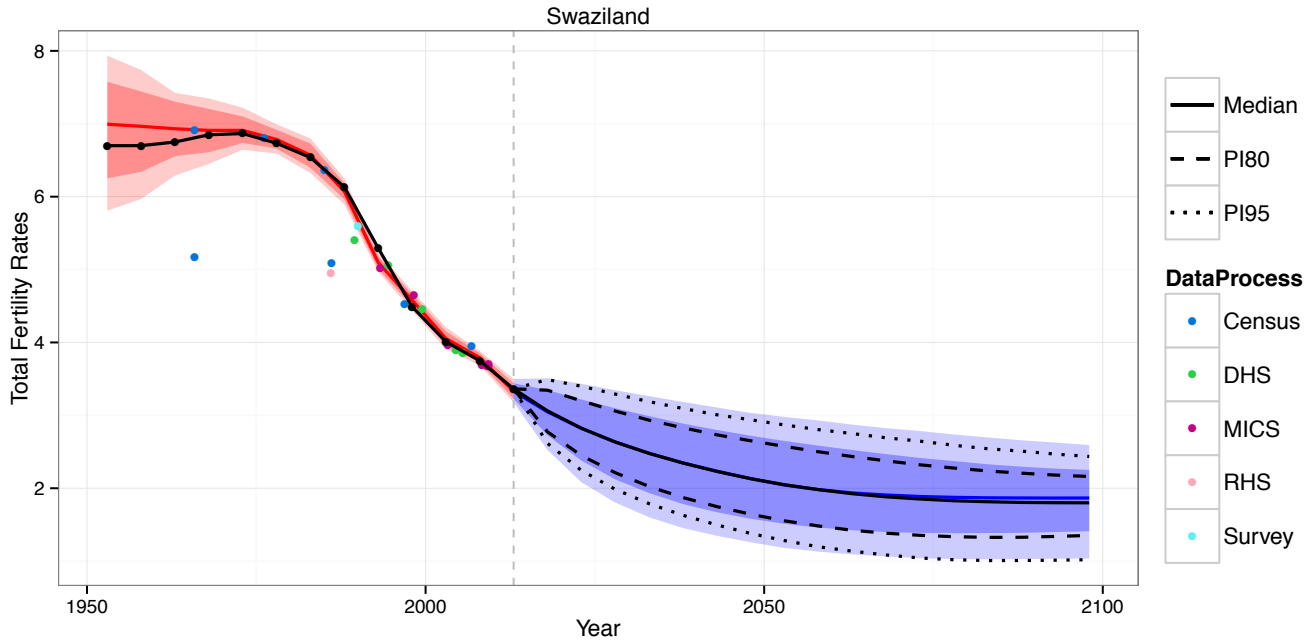


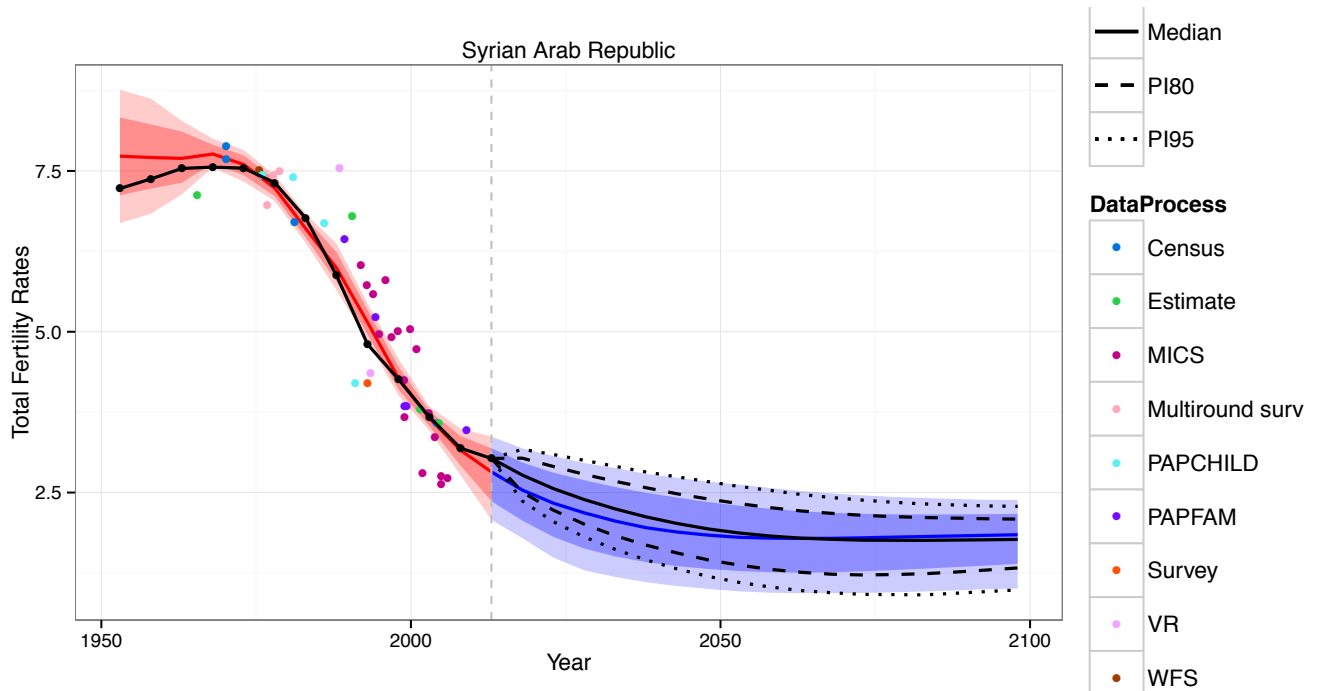
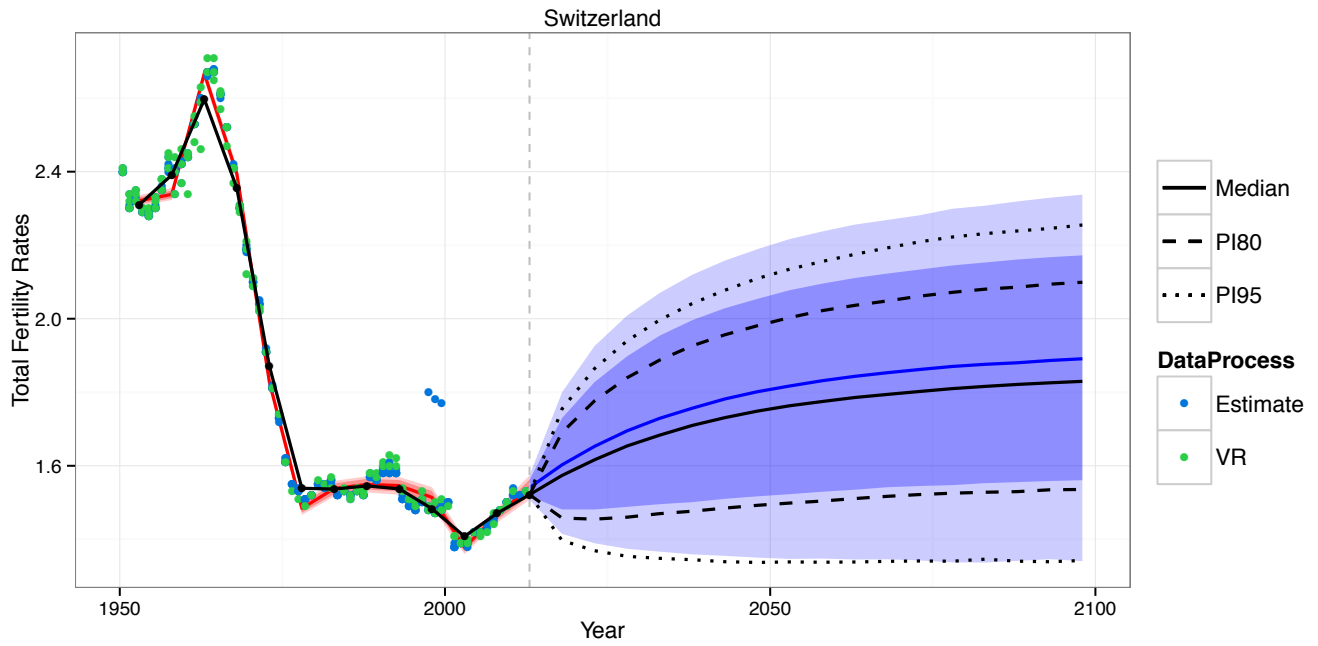


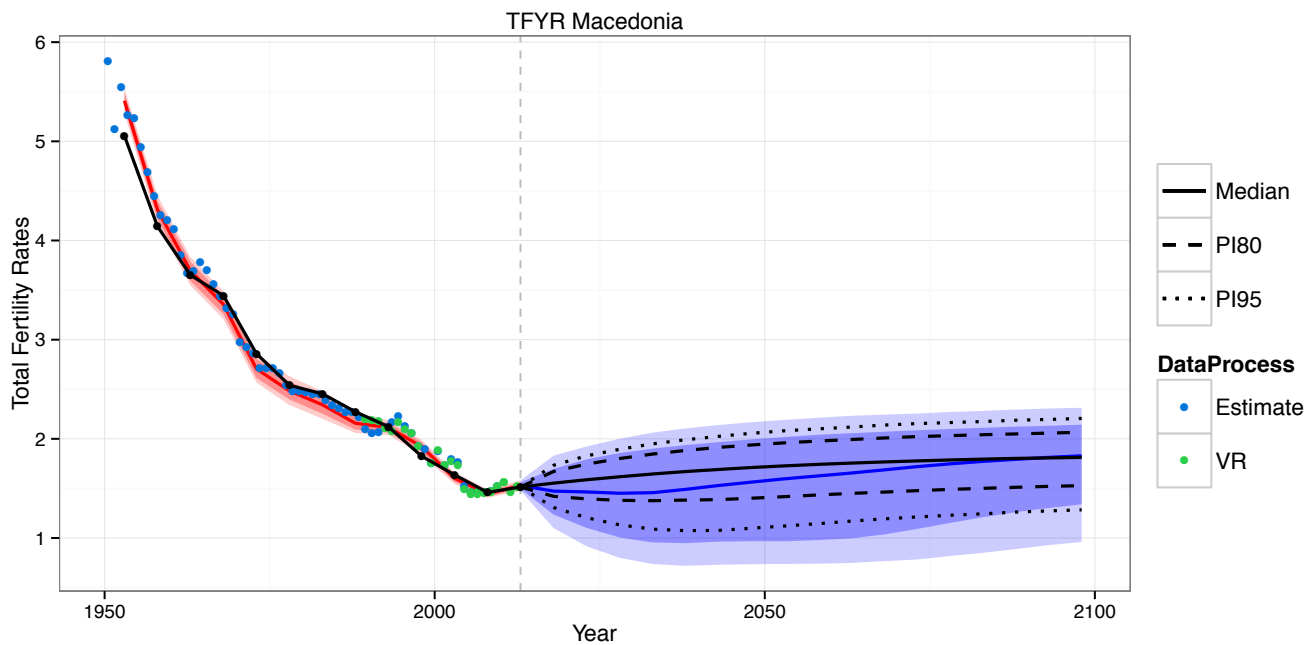
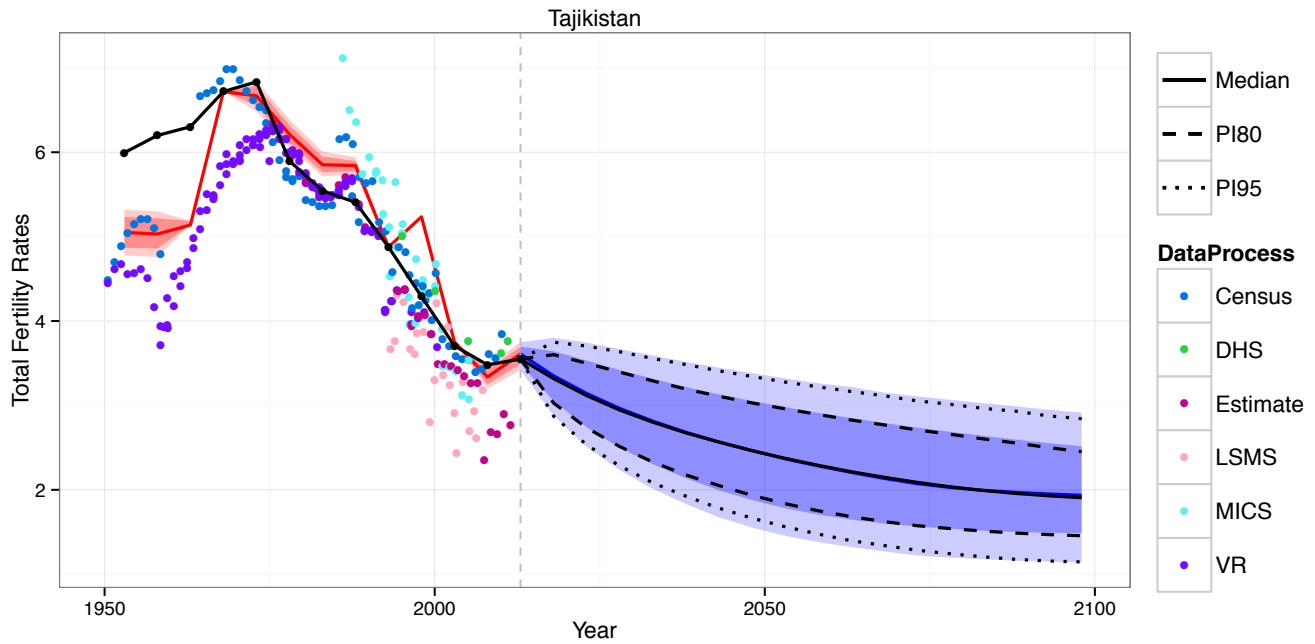


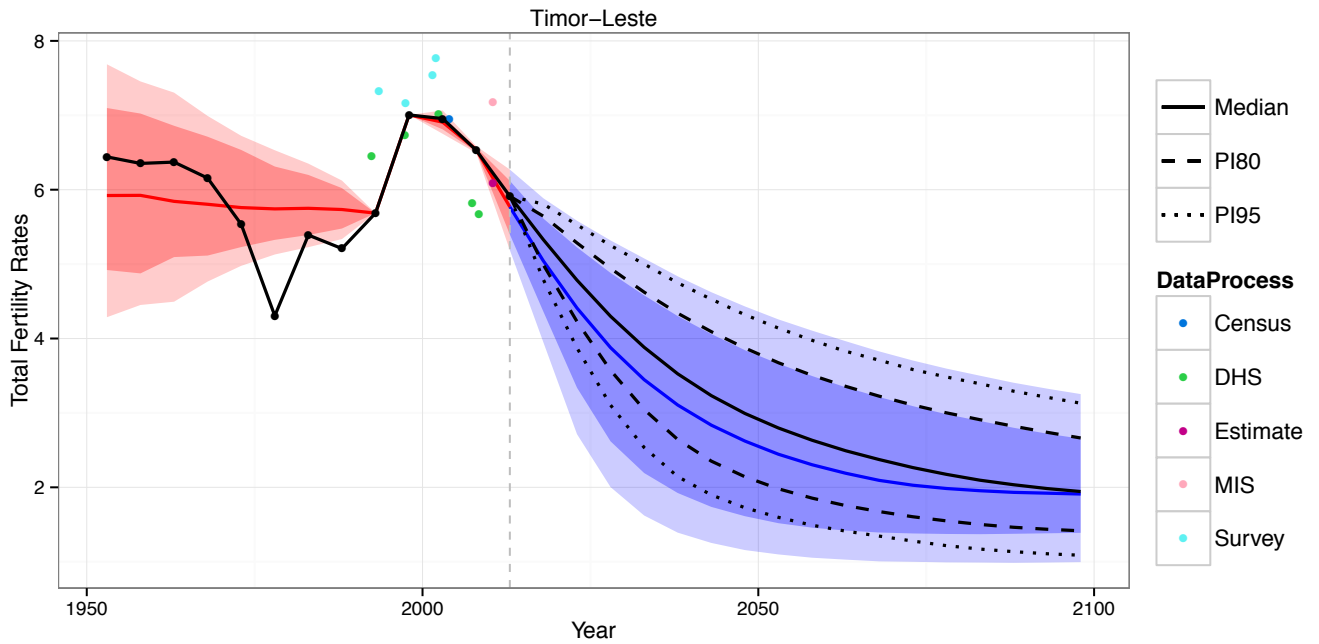
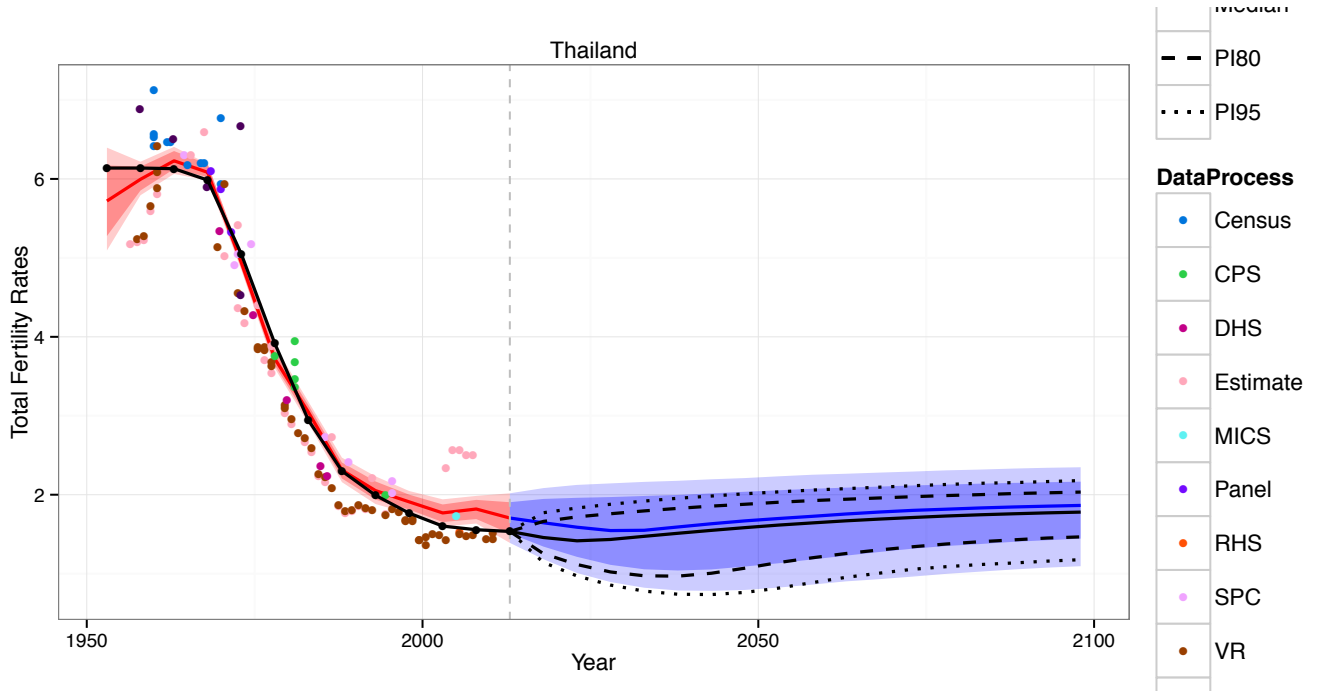


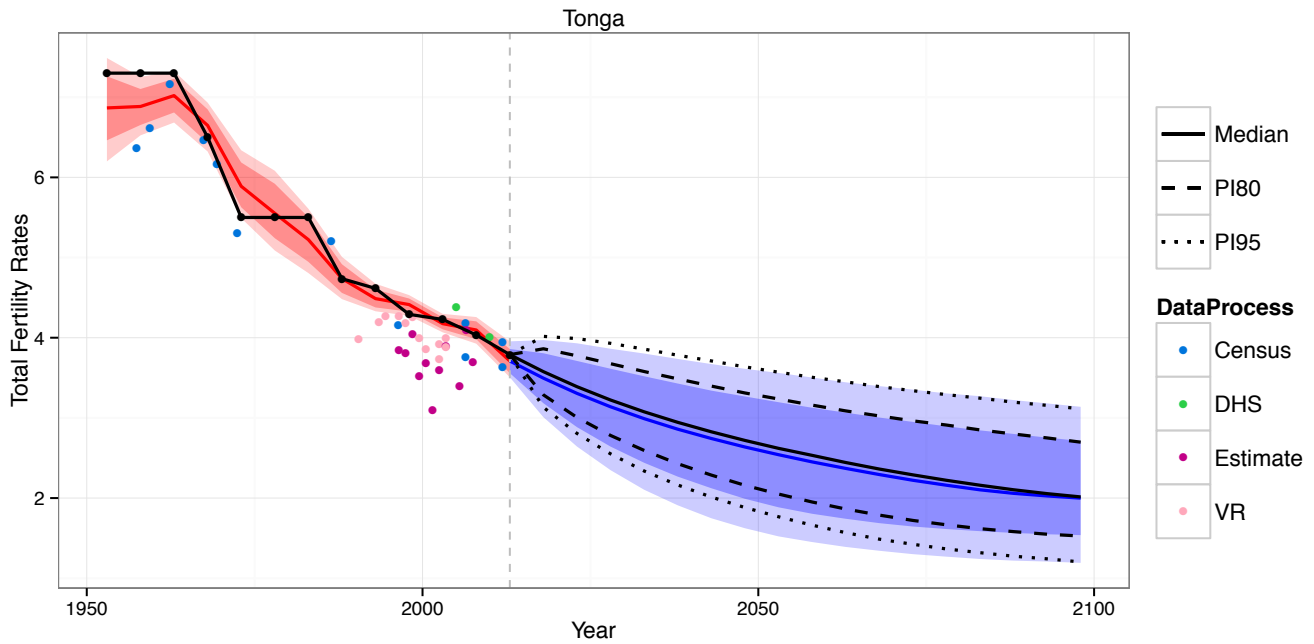
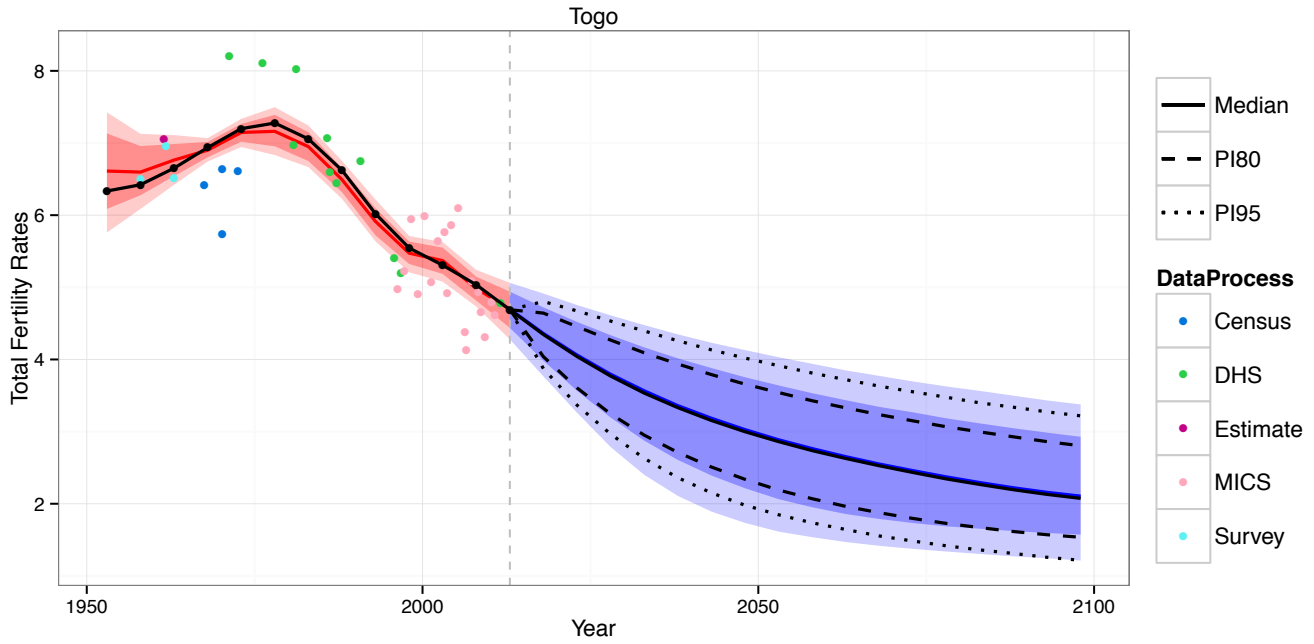


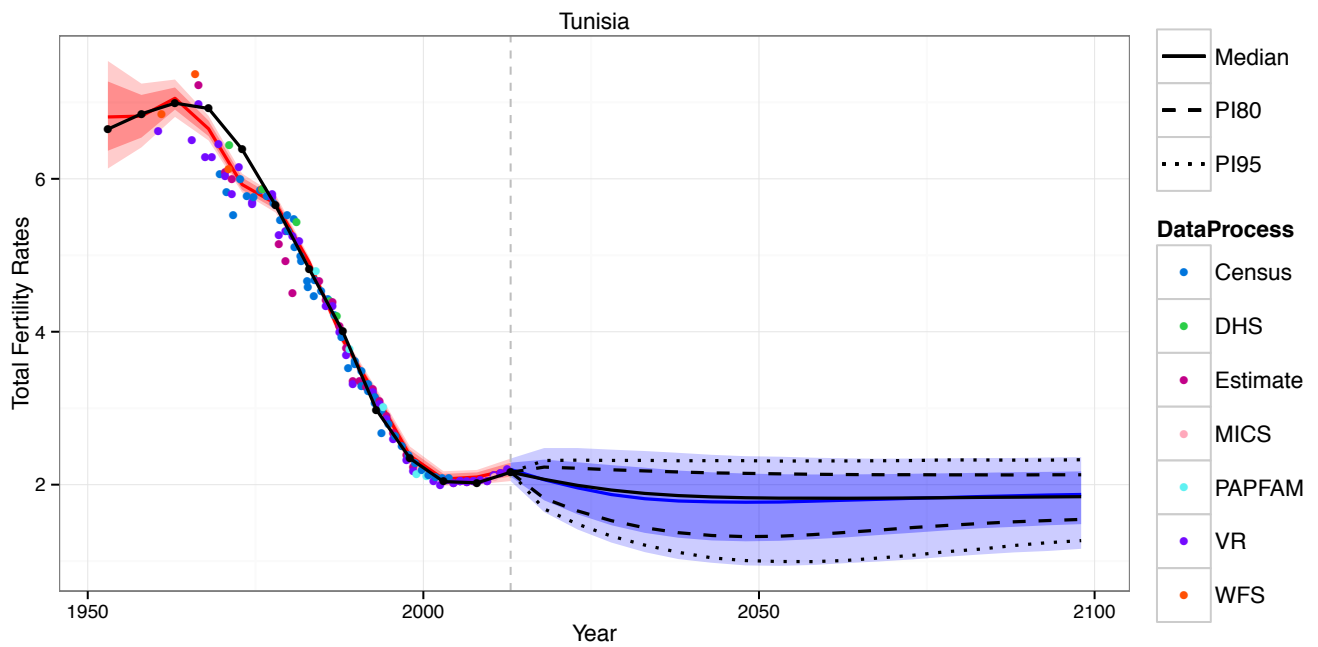
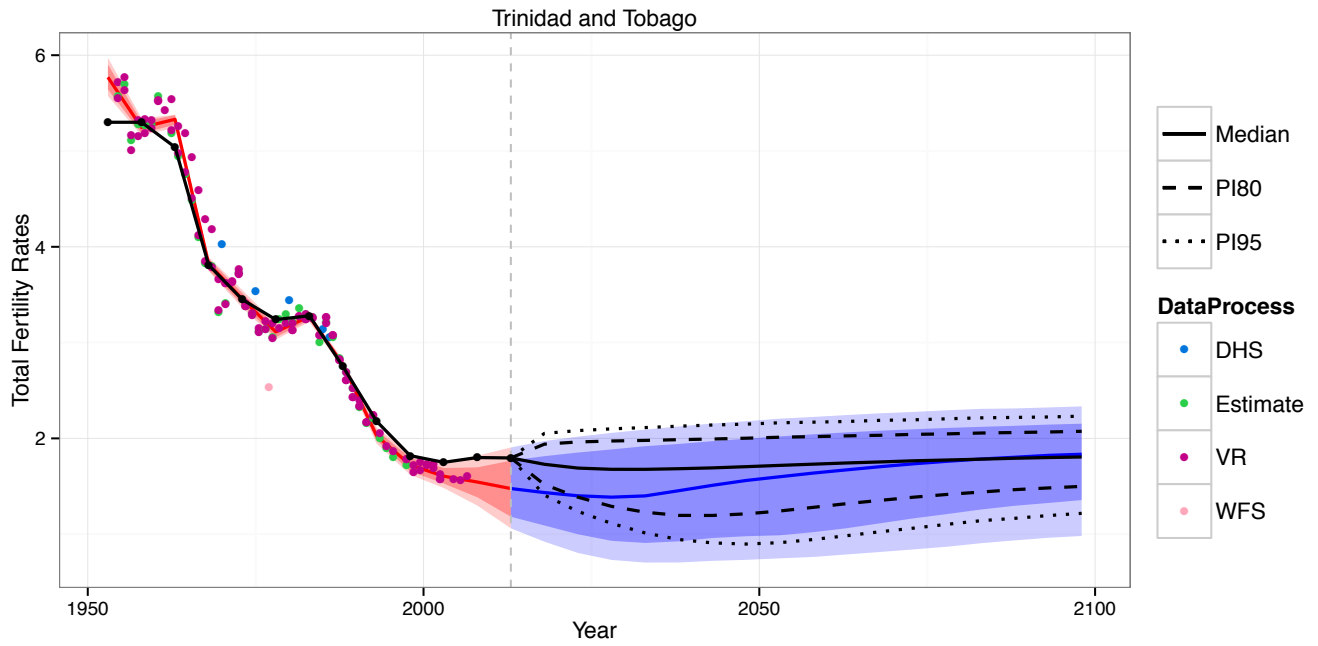


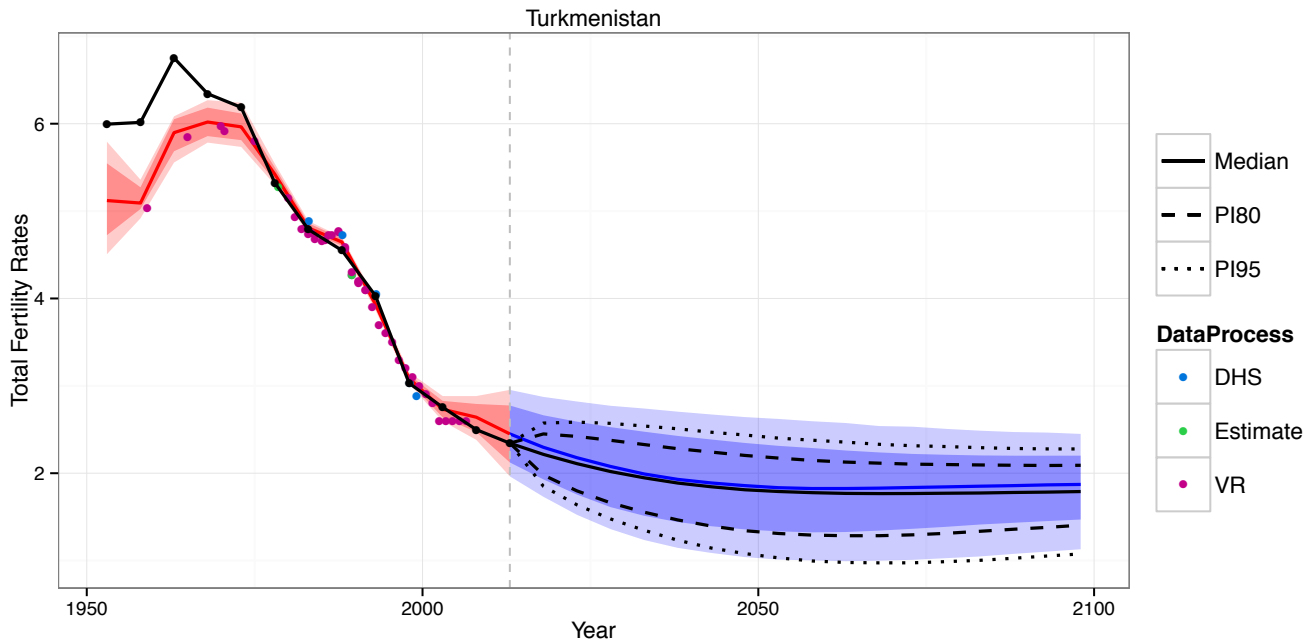
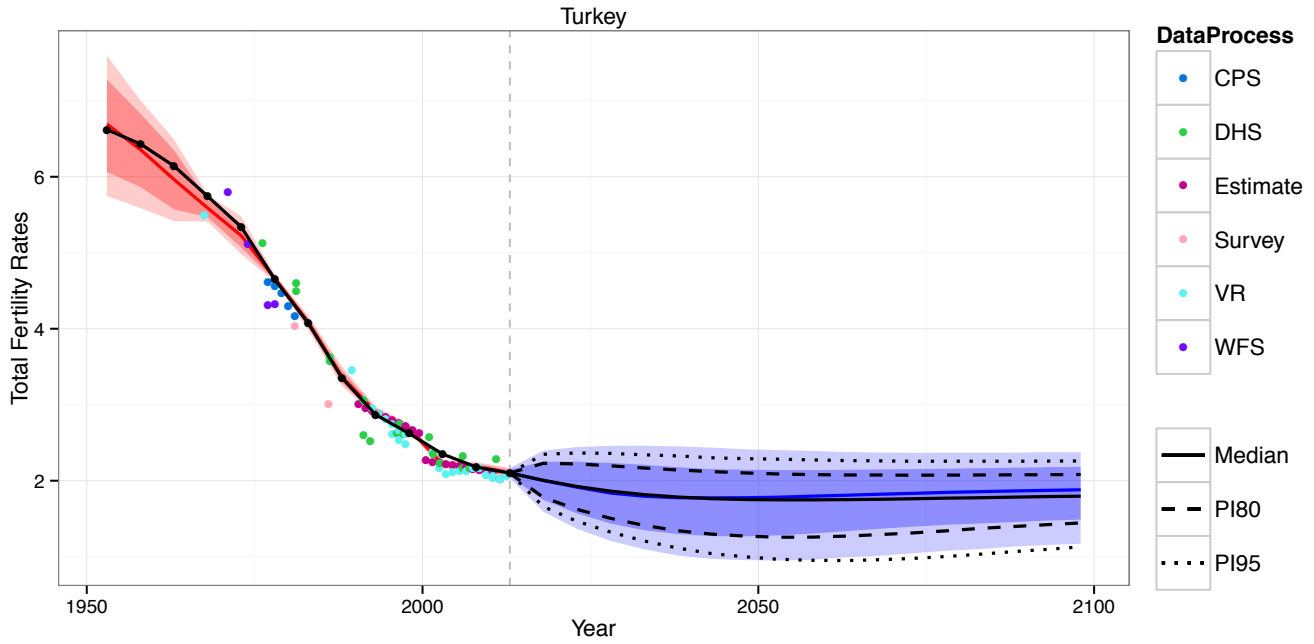


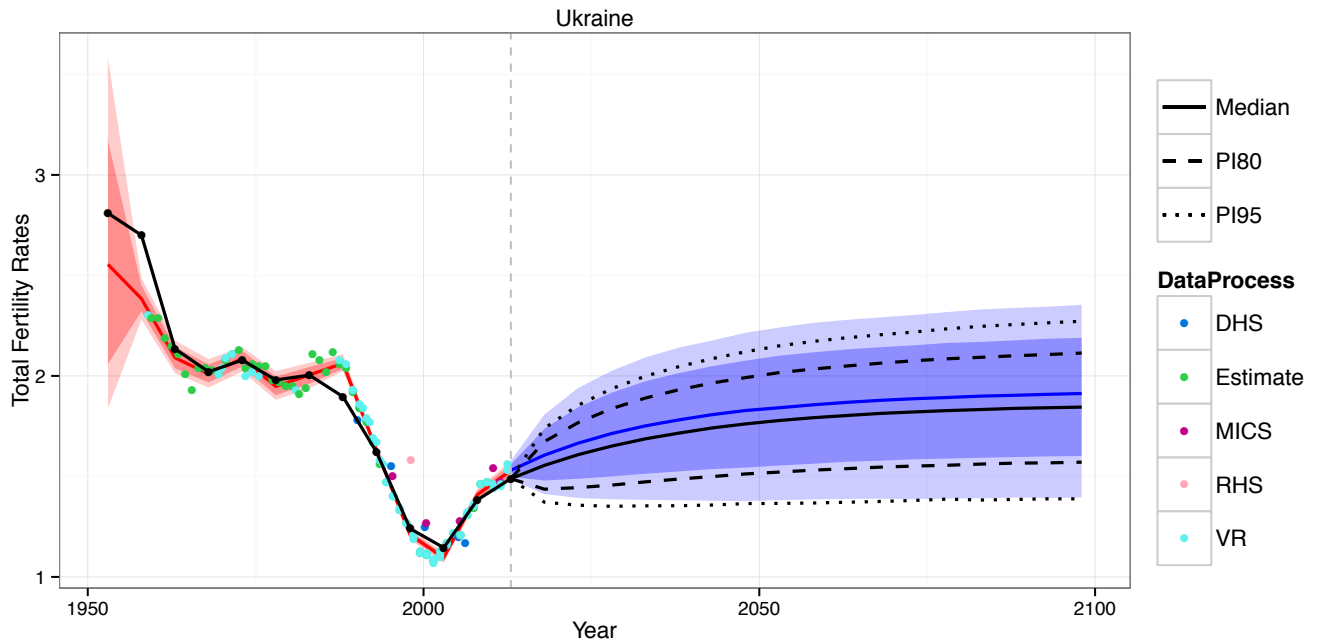
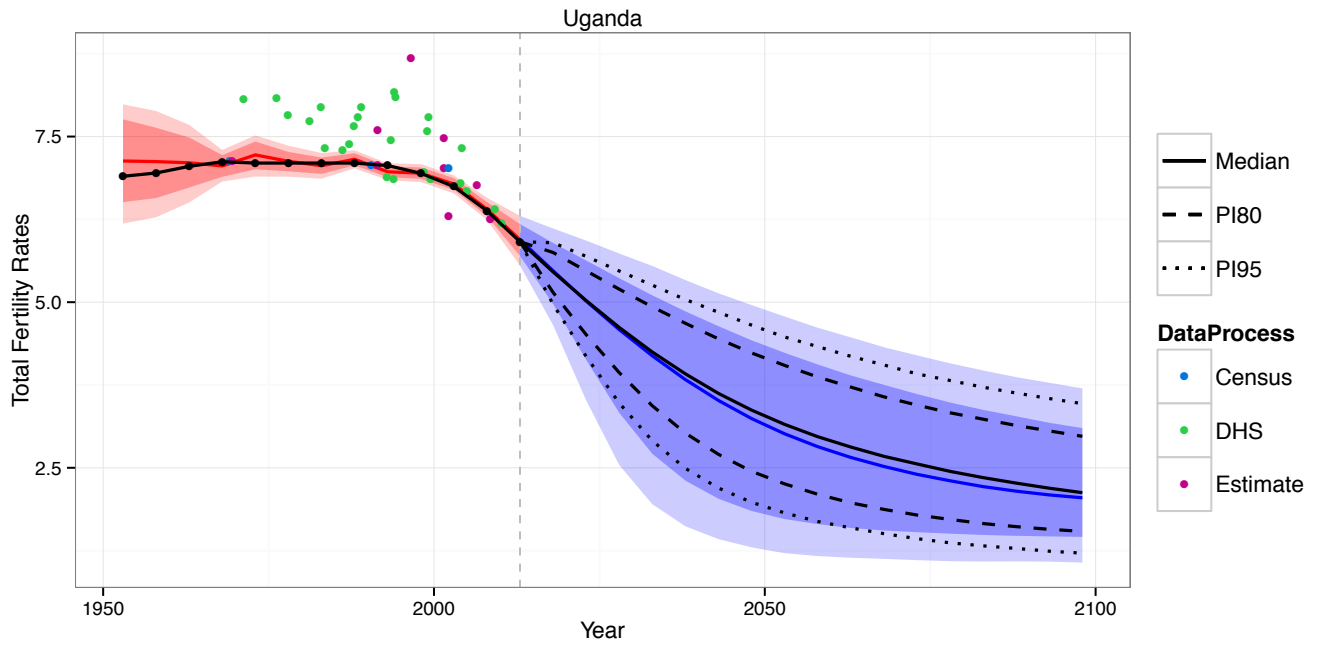


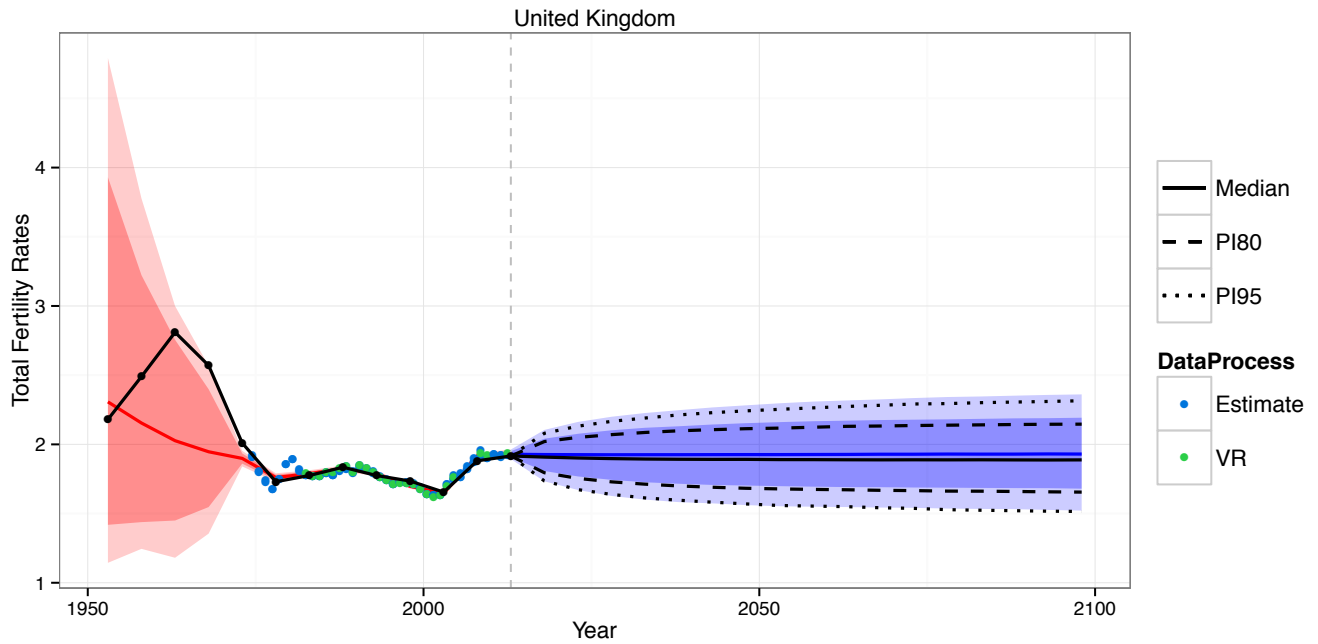
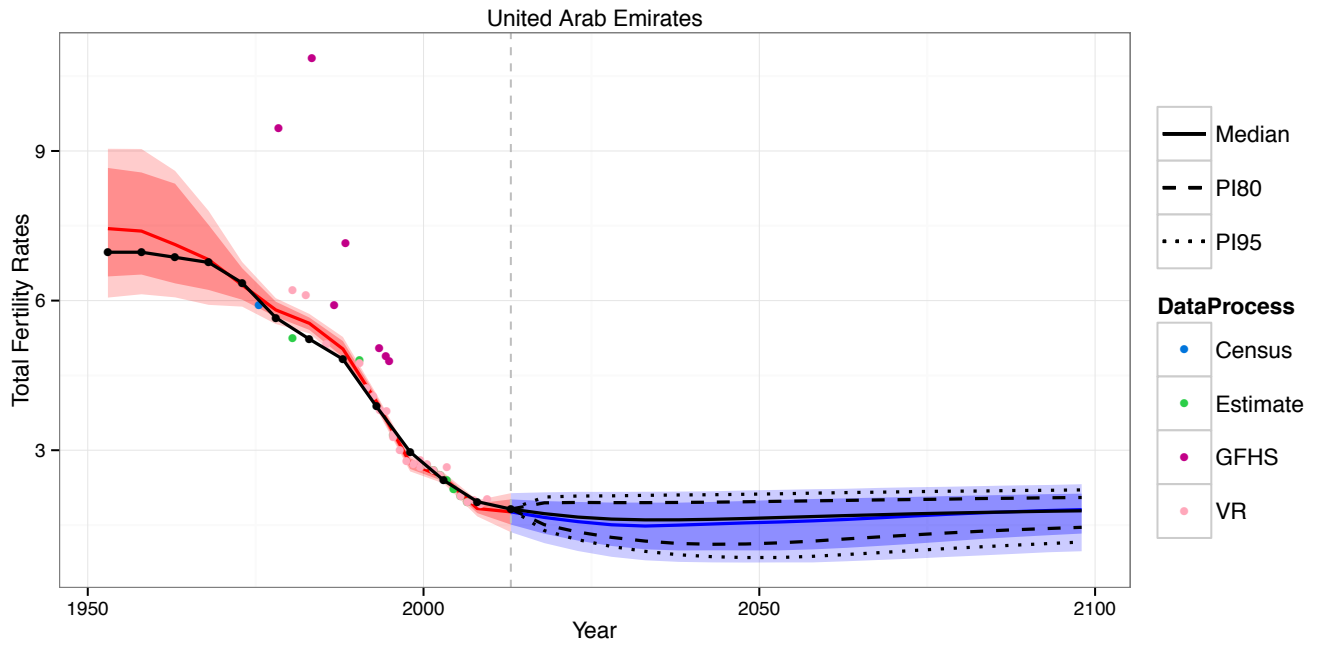


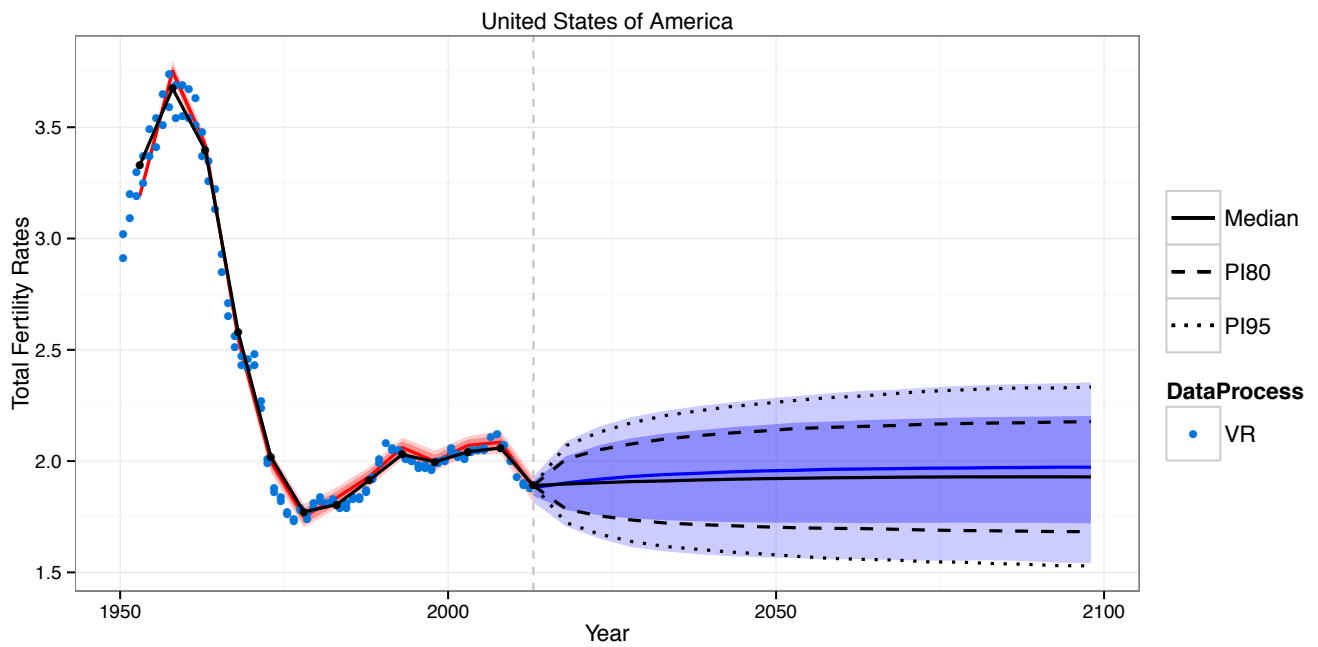
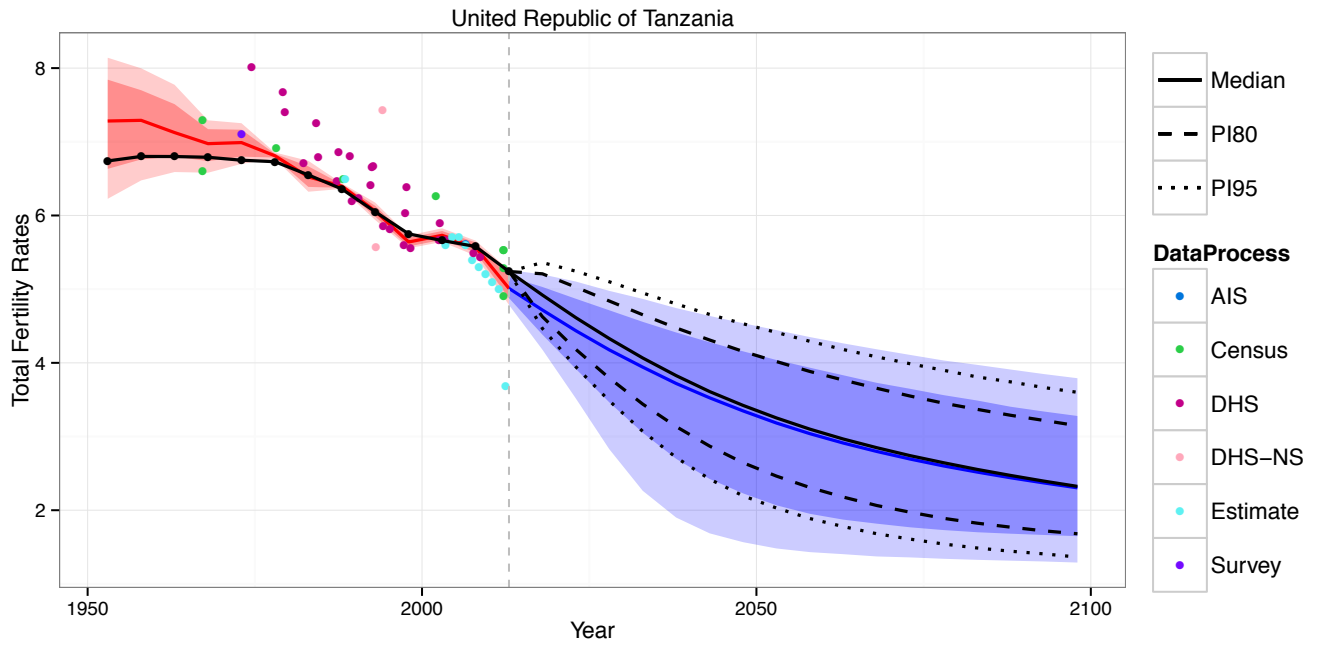


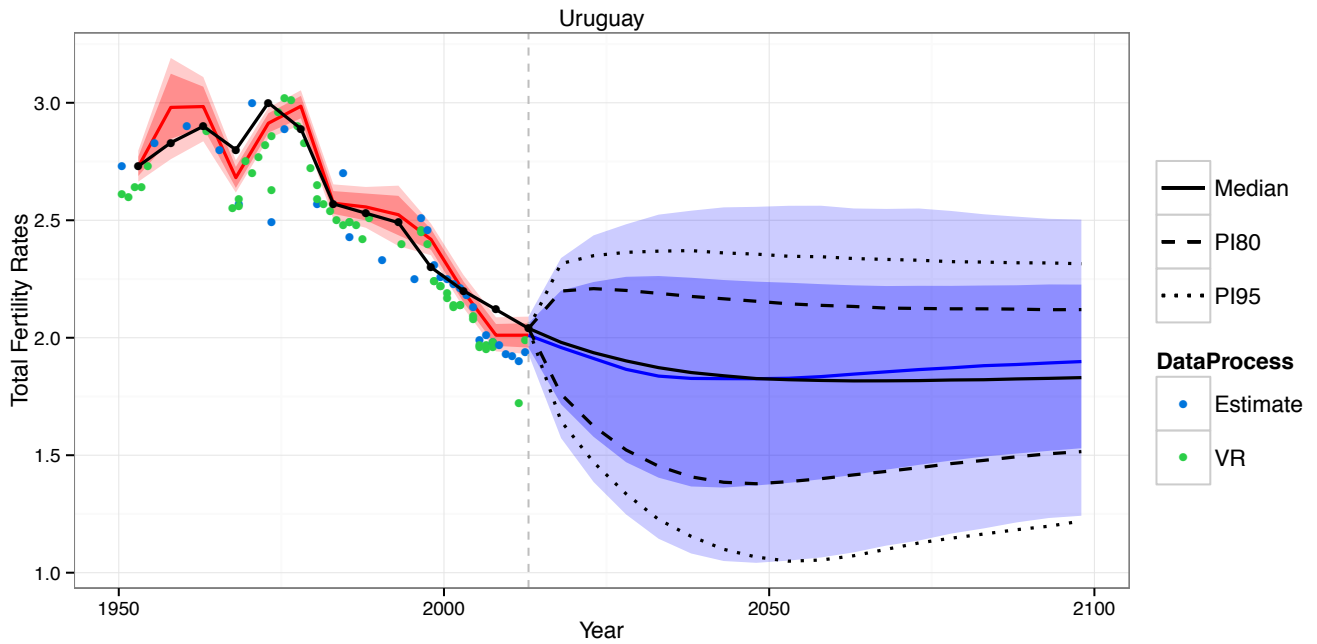
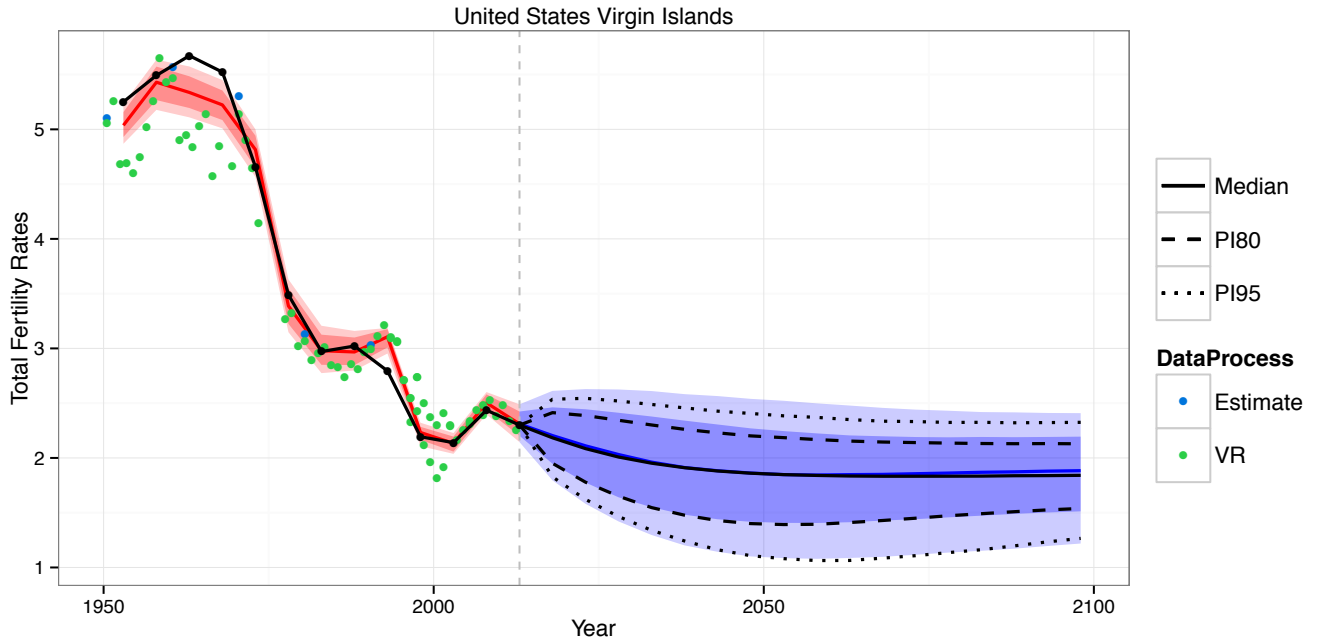


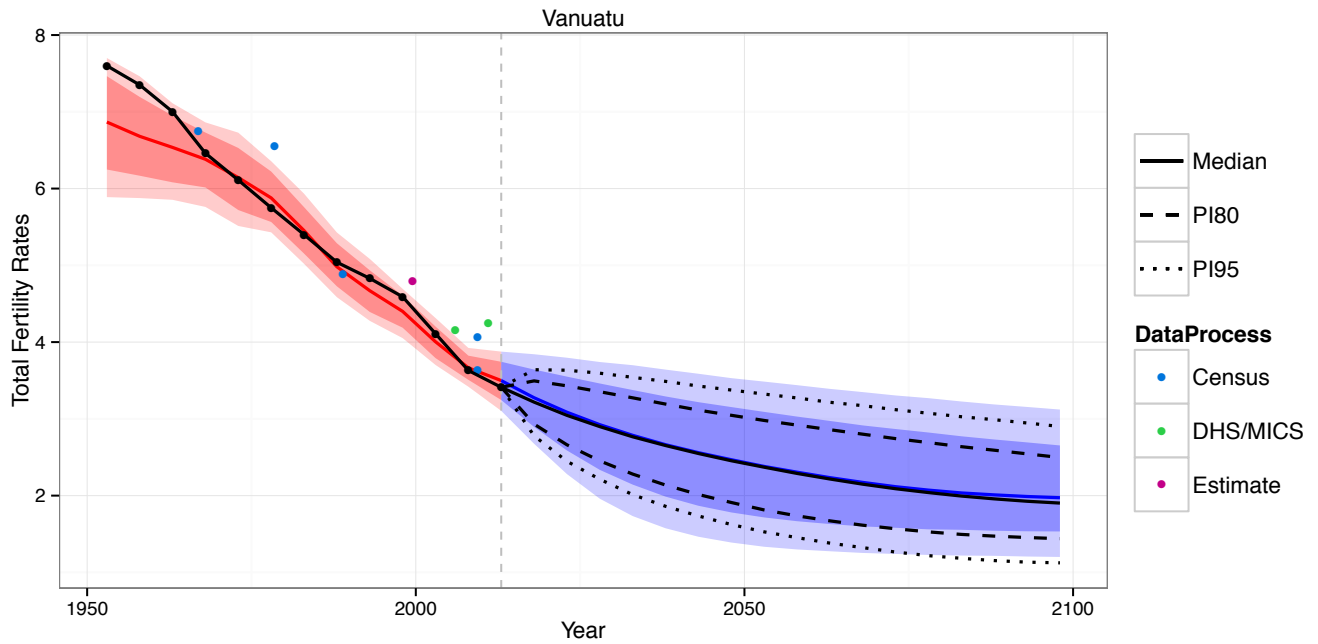
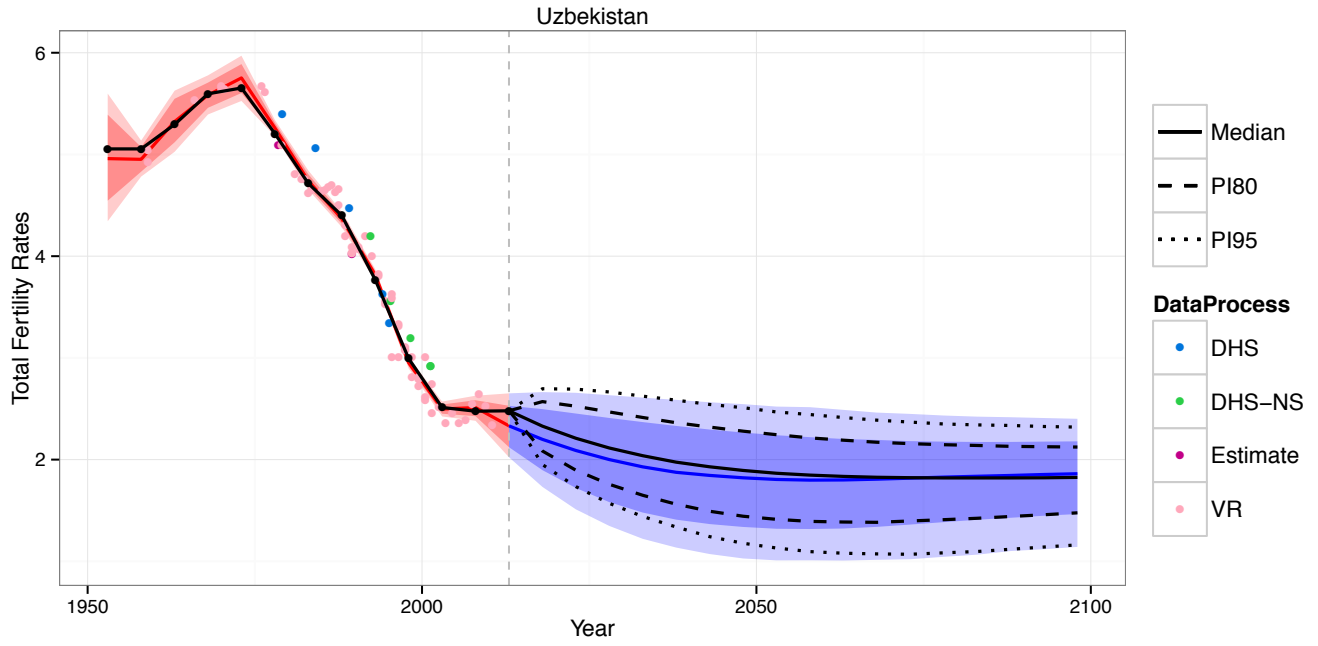


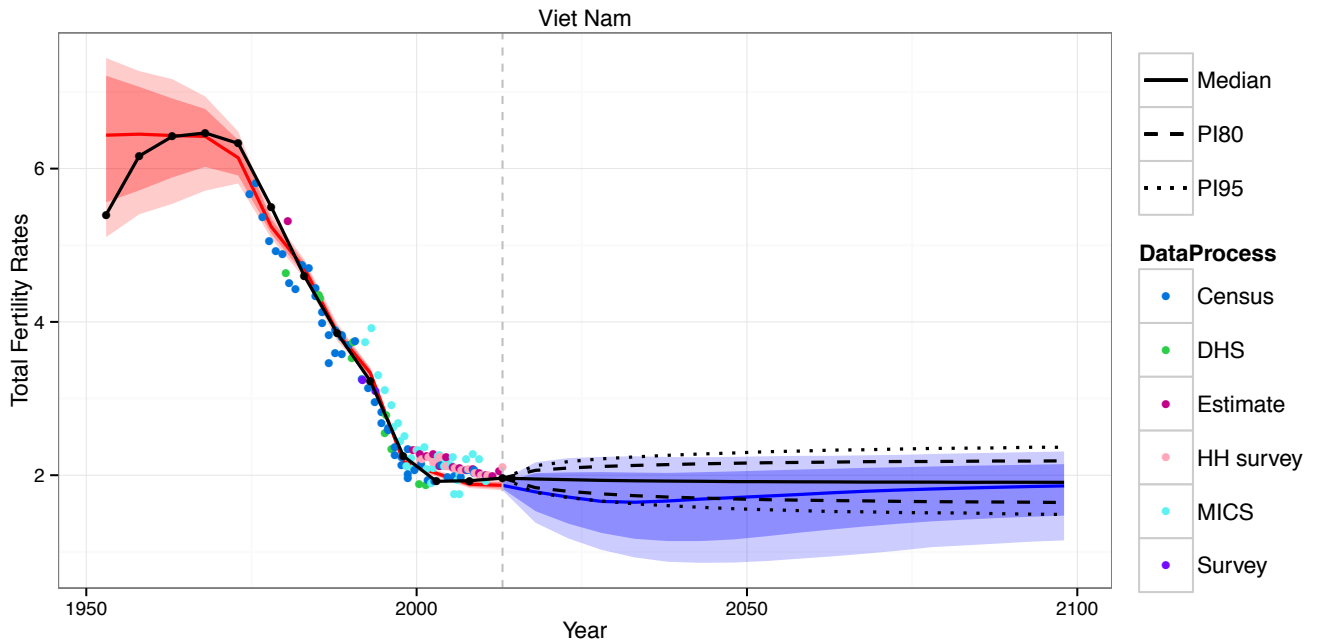
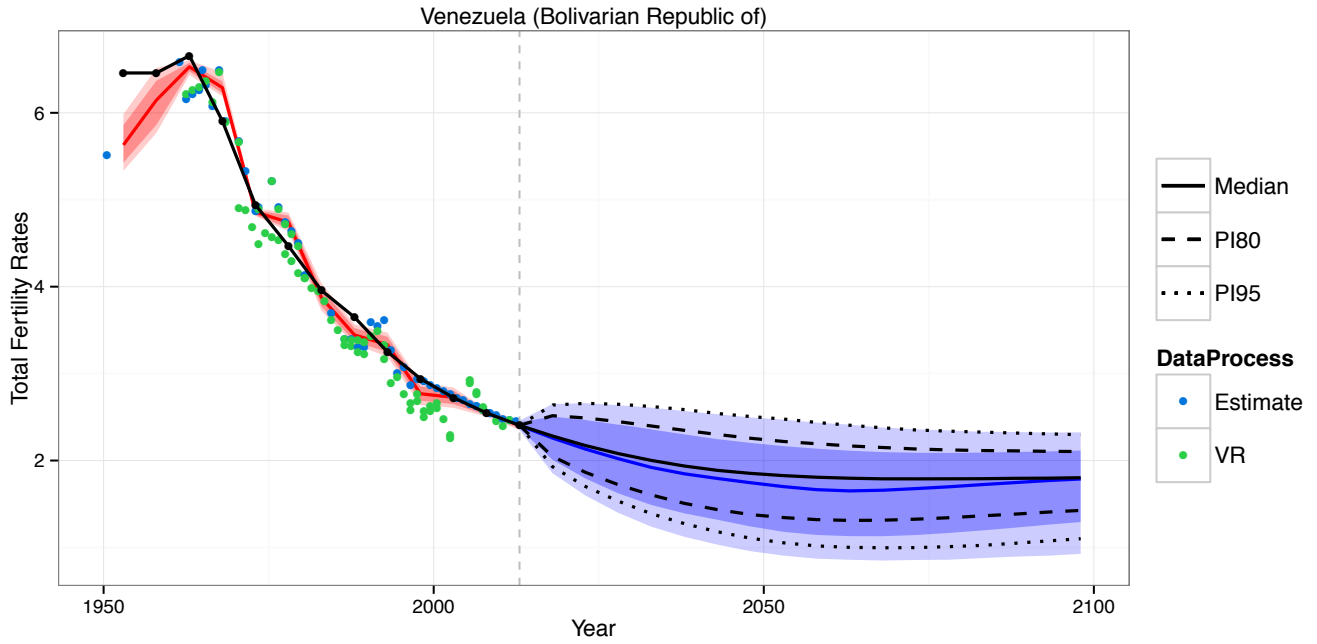


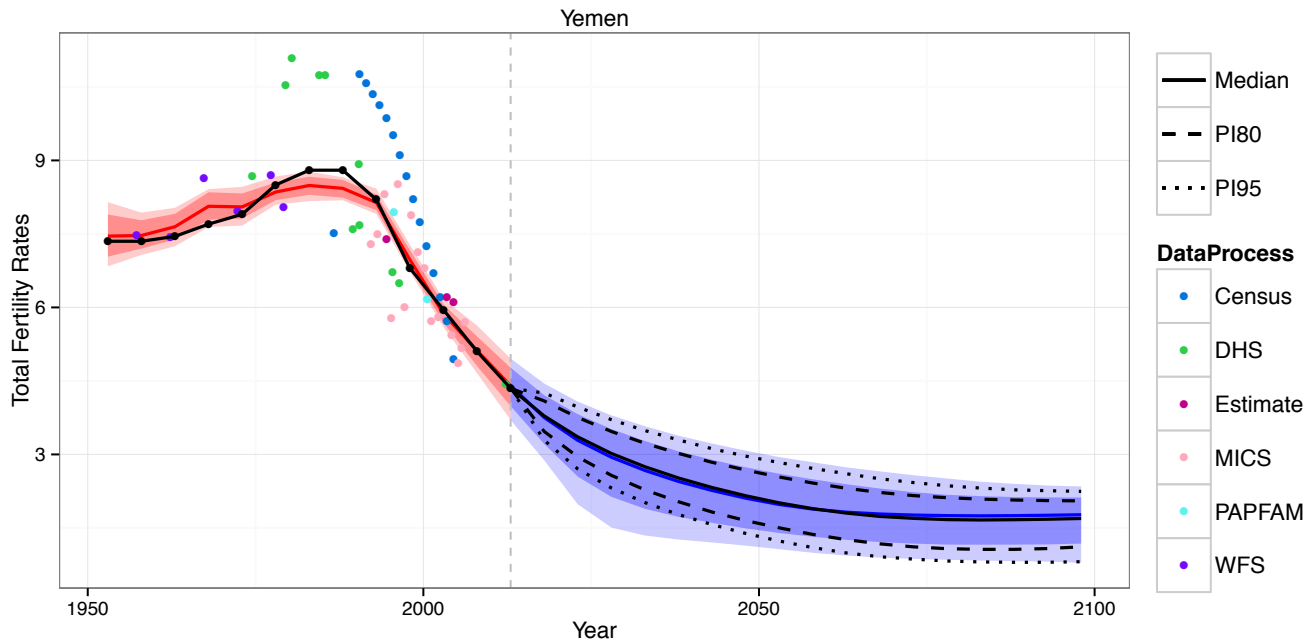
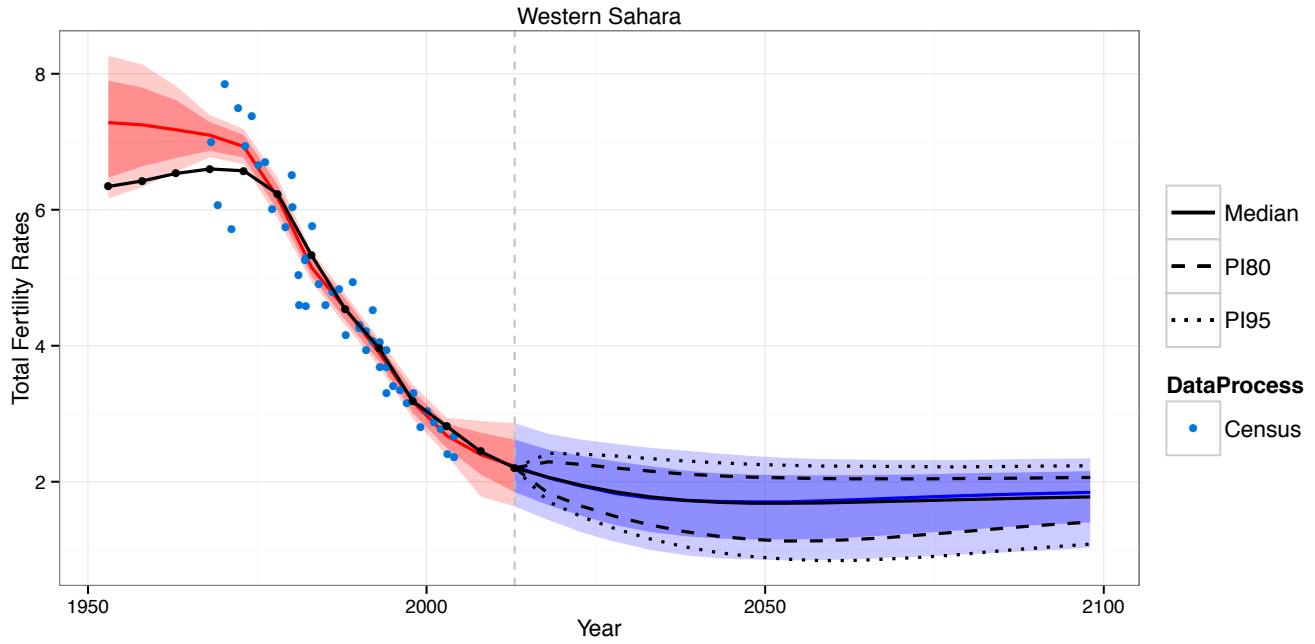


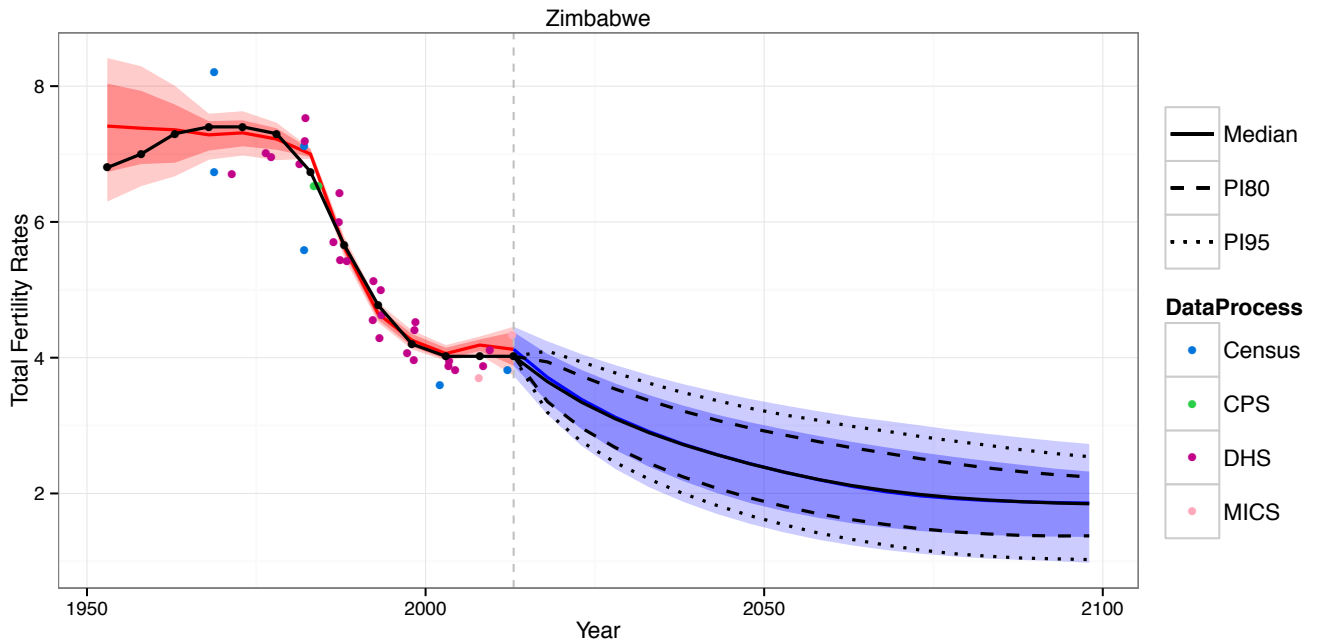
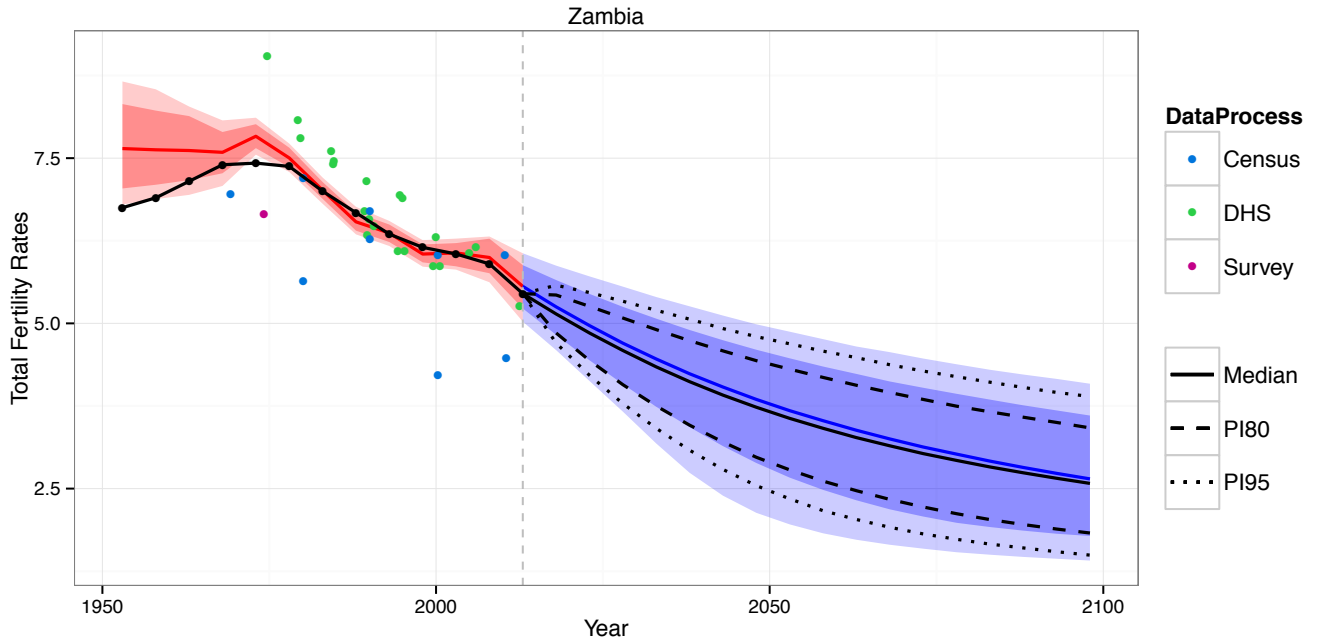












References.

- ALKEMA, L., RAFTERY, A. E., GERLAND, P., CLARK, S. J., PELLETIER, F., BUETTNER, T. and HEILIG, G. K. (2011). Probabilistic Projections of the Total Fertility Rate for All Countries. *Demography* **48** 815–839.
- BROOKS, S. P. and GELMAN, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* **7** 434–455.
- GELMAN, A. and RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 457–472.
- RAFTERY, A. E. and LEWIS, S. M. (1996). Implementing MCMC. In *Practical Markov Chain Monte Carlo* (W. Gilks, D. J. Spiegelhalter and S. Richardson, eds.) 115–130. Chapman and Hall, London.