

Supplementary Material: Estimating and Forecasting the Smoking-Attributable Mortality Fraction Jointly for Both Genders in Over 60 Countries

Yicheng Li and Adrian E. Raftery

October 20, 2019

1 MCMC Convergence Diagnostics

1.1 Hyperparameter Diagnostics

In this section, we present the MCMC convergence diagnostics of the hyperparameters in Level 4 of the model in terms of traceplots, Raftery diagnostic statistics (Raftery and Lewis, 1992), and Gelman diagnostic statistics (Gelman and Rubin, 1992). Table 1 provides the Gelman and Raftery diagnostic statistics of all hyperparameters. We use 3 chains with 2000 burnin and 8000 samples without thinning for the Gelman diagnostics, and randomly choose one of the chain to perform the Raftery diagnostics. Figure 1 shows the traceplots of all 8000 samples of hyperparameters.

Table 1: Diagnostic statistics for hyperparameters. PSRF and 95% UCI are the point estimator and upper bound of the 95% CI of the Gelman potential scale reduction factor. Burn1, Size1, and DF1 are the length of burn-in, required sample size, and dependent factor of Raftery diagnostics with parameters $q = 0.025, r = 0.0125, s = 0.95$. Burn2, Size2, and DF2 are the length of burn-in, required sample size, and dependent factor of Raftery diagnostics with parameters $q = 0.975, r = 0.0125, s = 0.95$.

Parameters	Gelman Diag		Raftery Diag					
	PSRF	95% UCI	Burn1	Size1	DF1	Burn2	Size2	DF2
a_1^m	1	1.00	6	1318	2.20	6	1164	1.94
a_2^m	1	1.00	3	710	1.18	6	1504	2.51
a_3^m	1	1.00	6	1584	2.64	6	1424	2.37
a_4	1	1.01	8	1750	2.92	9	2028	3.38
k^m	1	1.01	10	1952	3.25	6	1236	2.06
$\sigma_{a_2^m}^2$	1	1.00	2	640	1.07	6	1730	2.88
$\sigma_{a_4}^2$	1	1.00	21	4410	7.35	12	2132	3.55
$\sigma_{k^m}^2$	1	1.00	6	1334	2.22	8	1448	2.41
a_1^f	1	1.00	2	640	1.07	6	1318	2.20
Δ_{a_2}	1	1.00	4	756	1.26	6	688	1.15
a_3^f	1	1.00	8	1504	2.51	12	1852	3.09
k^f	1	1.00	5	895	1.49	2	640	1.07
$\sigma_{\Delta_{a_2}}^2$	1	1.00	3	696	1.16	4	839	1.40
$\sigma_{k^f}^2$	1	1.00	2	640	1.07	6	1376	2.29
ν	1	1.00	8	1518	2.53	8	1934	3.22
ρ^2	1	1.00	6	1270	2.21	6	1392	2.32
σ_h^2	1	1.00	10	1872	3.12	12	2337	3.90

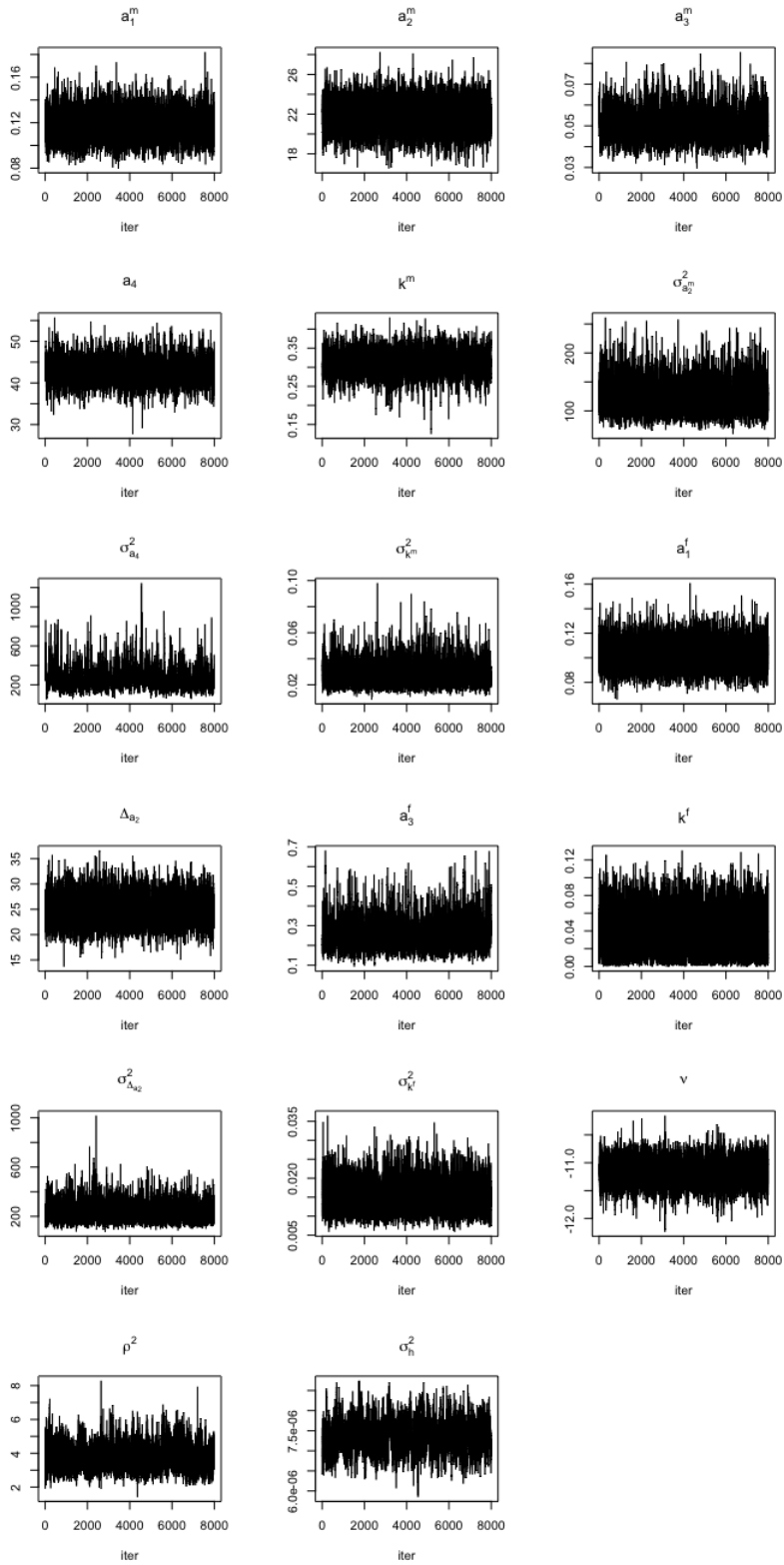


Figure 1: Traceplots for the hyperparameters.

1.2 Country-specific Parameter Diagnostics

In this section, we present the MCMC convergence diagnostics of country-specific parameters of the model in terms of traceplots, Raftery diagnostic statistics, and Gelman diagnostic statistics. Table 2 provides the Gelman and Raftery diagnostic statistics of country-specific parameters of the United States for male and female. The chains are the same as in the previous section. Figure 2 shows the traceplots of all 8000 samples of country-specific parameters for male and female of the United States.

Table 2: Diagnostic statistics for country-specific parameters for the United States. PSRF and 95% UCI are the point estimator and upper bound of the 95% CI of the Gelman potential scale reduction factor. Burn1, Size1, and DF1 are the length of burn-in, required sample size, and dependent factor of Raftery diagnostics with parameters $q = 0.025, r = 0.0125, s = 0.95$. Burn2, Size2, and DF2 are the length of burn-in, required sample size, and dependent factor of Raftery diagnostics with parameters $q = 0.975, r = 0.0125, s = 0.95$.

Parameters	Gelman Diag		Raftery Diag					
	PSRF	95% UCI	Burn1	Size1	DF1	Burn2	Size2	DF2
a_1^m	1	1.00	4	830	1.38	2	640	1.07
a_2^m	1	1.00	6	1326	2.21	4	756	1.26
a_3^m	1	1.00	4	822	1.37	2	633	1.06
a_4^m	1	1.00	2	614	1.02	4	772	1.29
k^m	1	1.00	6	1106	1.10	2	621	1.03
a_1^f	1	1.00	3	661	1.10	2	614	1.02
a_2^f	1	1.00	2	627	1.04	2	614	1.02
a_3^f	1	1.00	2	627	1.04	6	1314	2.19
a_4^f	1	1.00	3	661	1.10	4	848	1.41
k^f	1	1.00	3	668	1.11	6	1444	2.41
σ_c^2	1.01	1.03	96	15198	25.30	24	3834	6.39

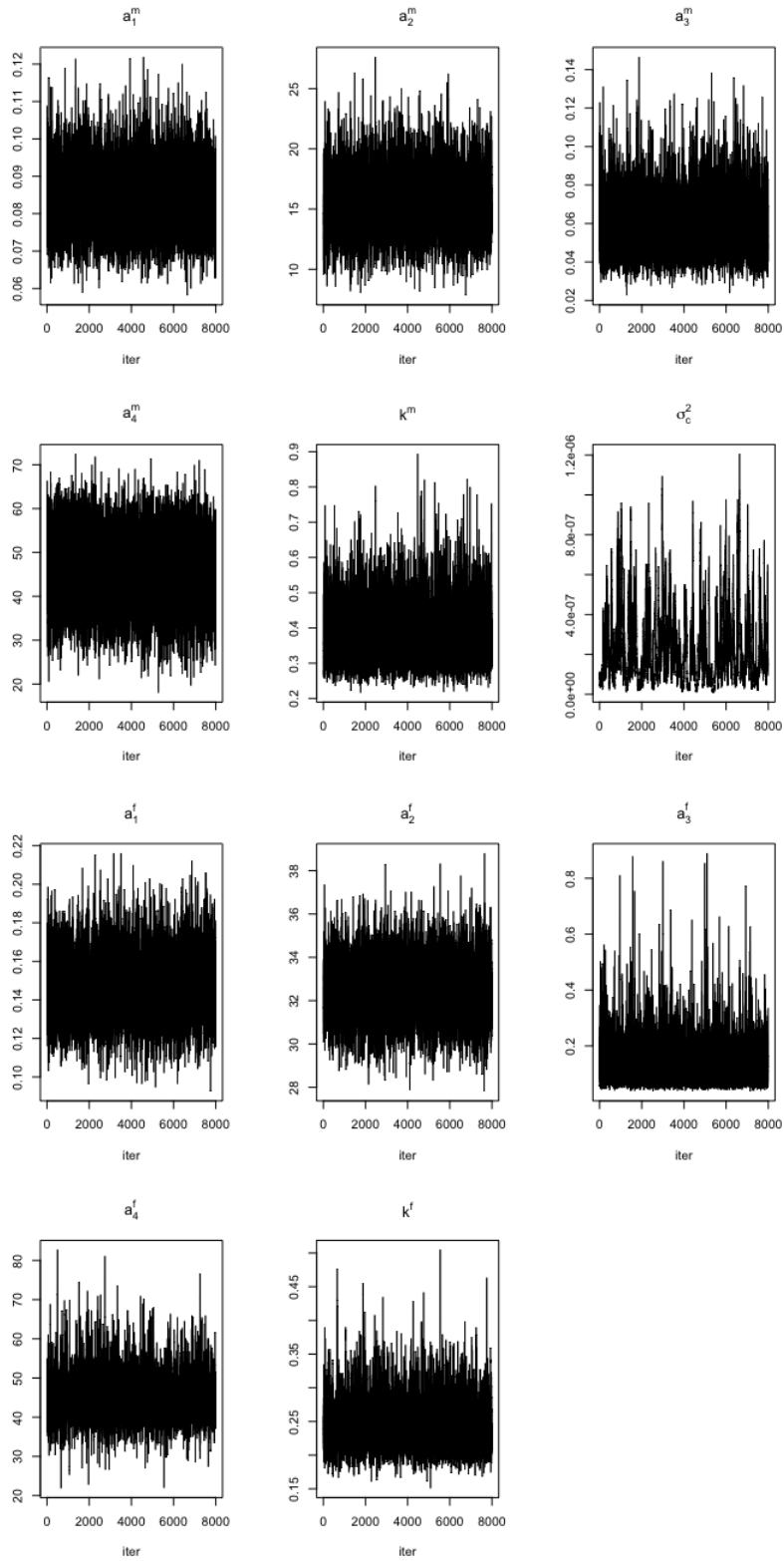


Figure 2: Traceplots for the country-specific parameters of the United States.

2 Hyperparameter Sensitivity Analysis

In this section, we present the sensitivity analysis for the hyperparameters set in $\pi(\cdot)$ on the posterior distributions of the global parameters ψ in Level 4 of our model. We use `rstansensitivity` package (Giordano, 2019) in R to perform the sensitivity analysis. The local sensitivity of the posterior mean of parameter θ under hyperparameters ζ (i.e., $\mathbb{E}(\theta|x, \zeta)$) to ζ at ζ_0 is defined as

$$S_{\zeta_0} := \left. \frac{d\mathbb{E}(\theta|x, \zeta)}{d\zeta} \right|_{\zeta_0},$$

where x is the observed data (cf. Basu et al. (1996), Gustafson (1996), and Giordano et al. (2018) for more discussions on local sensitivity in Bayesian analysis). By scaling the local sensitivity to be comparable with the possible range of the posterior distribution of θ , the normalized local sensitivity is defined as

$$\tilde{S}_{\zeta_0} := \left| \frac{S_{\zeta_0}}{\text{sd}(\theta|x, \zeta_0)} \right|.$$

As commented in Giordano et al. (2018), if the quantity \tilde{S}_{ζ_0} is less than 1 or if \tilde{S}_{ζ_0} is greater than 1 but the final results barely change when modifying the hyperparameters, then the model is robust. First of all, Table 3 investigates the normalized local sensitivity of the hyperparameters set in $\pi(\cdot)$ on posterior distributions of the global parameters ψ . For most hyperparameters, the normalized local sensitivity are much smaller than 1. For those whose normalized local sensitivity are greater than 1, we conduct out-of-sample validations for three five-year period prediction with the hyperparameters changed to evaluate the actual changes on the validation results. Table 4, 5, 6, and 7 show the out-of-sample validation results after modifying $\beta_{\sigma_{k_m}^2}$ (0.255^2 to 1), $\beta_{\sigma_{k_f}^2}$ (0.255^2 to 1), $\beta_{\sigma_h^2}$ (0.01^2 to 0.02^2), and α_{a_4} (38.362 to 20), respectively. All four cases show that the validation results barely change, and we conclude that model is robust under the current choices of hyperparameters.

Table 3: Normalized local sensitivity of hyperparameters on the global parameters. The bold numbers are local sensitivity \tilde{S}_{ζ_0} with absolute value greater than 1.

	α_1^m	α_2^m	α_3^m	α_4	k_m	$\sigma_{\alpha_4}^2$	$\sigma_{k_m}^2$	α_1^f	Δ_{α_2}	α_3^f	k_f	$\sigma_{\Delta_{\alpha_2}}^2$	$\sigma_{k_f}^2$	ν	ρ^2	σ_h^2
$\alpha_{\alpha_1^m}$	0.014	0.001	0.001	0.046	-0.006	-0.031	0.198	-0.000	-0.018	0.000	-0.001	-0.096	-0.000	0.003	0.002	-0.000
$\beta_{\alpha_1^m}$	-0.002	-0.000	-0.000	-0.006	0.001	0.003	-0.023	0.000	0.002	-0.000	0.000	0.011	0.000	-0.000	-0.000	0.000
$\alpha_{\alpha_2^m}$	0.000	0.012	0.000	-0.001	-0.000	0.005	0.003	0.000	0.003	-0.000	0.000	-0.002	0.000	0.000	-0.000	-0.000
$\beta_{\alpha_2^m}$	-0.000	-0.004	-0.000	0.000	-0.000	-0.000	-0.002	0.000	0.001	0.000	-0.000	0.002	-0.000	-0.000	0.000	0.000
$\alpha_{\alpha_3^m}$	0.002	0.019	0.011	0.059	-0.009	0.002	0.295	0.001	-0.034	0.002	-0.001	-0.094	-0.000	-0.001	-0.001	0.000
$\beta_{\alpha_3^m}$	-0.000	-0.001	-0.001	-0.003	0.000	-0.000	-0.015	-0.000	0.002	-0.000	0.000	0.005	0.000	0.000	0.000	-0.000
α_{α_4}	0.000	0.000	0.000	0.020	-0.001	-0.004	0.048	0.000	-0.003	0.000	-0.000	-0.011	-0.000	-0.000	-0.000	0.000
β_{α_4}	0.000	-0.001	0.000	0.025	-0.001	-0.005	0.064	0.000	-0.004	0.000	-0.000	-0.013	-0.000	-0.000	-0.000	0.000
α_{k_m}	-0.030	-0.004	-0.027	-0.617	0.168	0.290	-1.591	-0.032	0.186	-0.007	0.010	0.870	0.000	-0.003	0.005	-0.000
β_{k_m}	-0.012	0.015	-0.012	-0.316	0.062	0.162	-0.610	-0.000	0.085	-0.004	0.005	0.399	0.001	0.004	0.005	-0.000
$\alpha_{\sigma_{\alpha_4}^2}$	0.001	-0.025	-0.000	0.025	-0.002	-1.104	-0.051	0.000	0.002	-0.003	0.003	-0.136	-0.000	-0.002	0.007	0.000
$\beta_{\sigma_{\alpha_4}^2}$	-0.000	0.000	-0.000	-0.000	0.000	0.009	0.000	-0.000	0.000	-0.000	0.000	0.001	0.000	0.000	-0.000	-0.000
$\alpha_{\sigma_{k_m}^2}$	-0.008	-0.026	-0.007	-0.260	0.012	-0.066	-3.916	0.001	-0.014	0.010	-0.004	0.144	0.000	-0.006	0.006	0.000
$\beta_{\sigma_{k_m}^2}$	0.000	0.000	0.000	0.001	-0.000	0.000	0.015	-0.000	0.000	-0.000	0.000	-0.001	-0.000	0.000	-0.000	-0.000
$\alpha_{\sigma_{\Delta_{\alpha_2}}^2}$	-0.001	-0.001	-0.002	0.068	0.016	-0.010	0.178	-0.029	-0.001	0.051	-0.001	0.084	-0.000	-0.004	0.009	-0.000
$\beta_{\sigma_{\Delta_{\alpha_2}}^2}$	0.012	-0.063	0.055	-2.644	-0.468	-0.067	-10.187	1.004	0.058	-1.964	0.048	-0.065	0.004	0.123	-0.378	0.000
$\alpha_{\alpha_1^f}$	-0.000	0.001	-0.000	0.012	-0.001	-0.057	0.024	0.000	0.013	-0.051	0.004	-0.065	-0.000	0.002	0.001	-0.000
$\beta_{\alpha_1^f}$	0.000	-0.000	0.000	-0.001	0.000	0.006	-0.003	-0.000	-0.001	0.005	0.000	0.007	0.000	-0.000	-0.000	0.000
$\alpha_{\Delta_{\alpha_2}}$	-0.000	-0.005	-0.000	-0.006	0.001	0.001	0.008	-0.000	-0.001	0.040	-0.000	0.069	0.000	-0.001	-0.000	0.000
$\beta_{\Delta_{\alpha_2}}$	-0.001	-0.013	-0.001	-0.020	0.002	0.002	0.030	-0.001	-0.002	-0.001	0.002	0.222	0.000	-0.002	0.000	-0.000
$\alpha_{\alpha_3^f}$	0.000	0.002	0.001	0.003	-0.002	-0.074	-0.286	-0.000	0.004	-0.040	-0.006	-0.263	-0.002	-0.002	0.004	0.000
$\beta_{\alpha_3^f}$	-0.000	0.001	-0.000	-0.000	0.000	0.019	0.070	0.000	-0.001	0.008	0.001	0.059	0.000	0.001	-0.002	-0.000
α_{k_f}	-0.005	0.052	-0.002	-0.071	0.008	0.134	0.500	0.002	-0.007	0.249	0.110	0.909	-0.022	0.020	0.010	-0.000
β_{k_f}	0.012	-0.119	0.004	0.151	-0.019	-0.289	-1.155	-0.005	0.015	-0.551	-0.249	-1.948	0.049	-0.048	-0.027	0.000
$\alpha_{\sigma_{\Delta_{\alpha_2}}^2}$	0.003	0.013	0.003	0.063	-0.006	-0.129	0.217	0.001	0.002	-0.184	0.008	-2.611	-0.002	-0.004	0.001	0.000
$\beta_{\sigma_{\Delta_{\alpha_2}}^2}$	-0.000	-0.000	-0.000	-0.000	0.000	0.001	-0.001	-0.000	-0.000	0.001	0.000	0.012	0.000	0.000	-0.000	-0.000
$\alpha_{\sigma_{k_f}^2}$	0.000	-0.001	0.000	0.052	0.000	-0.021	0.044	-0.000	0.001	-0.018	0.007	-0.214	-0.016	0.002	0.008	0.000
$\beta_{\sigma_{k_f}^2}$	-0.025	-0.139	-0.021	-3.454	-0.022	1.940	-2.054	0.029	-0.042	1.225	-1.595	14.056	1.093	-0.124	-0.655	-0.000
α_{ν}	0.001	0.013	-0.000	-0.012	-0.000	0.018	0.034	0.001	0.001	-0.014	0.002	0.019	-0.000	0.079	-0.002	-0.000
β_{ν}	-0.001	-0.012	0.000	0.014	0.000	-0.029	-0.013	-0.001	-0.001	0.019	-0.001	-0.016	0.000	-0.098	0.005	0.000
α_{ρ^2}	-0.000	0.003	0.000	0.010	-0.000	0.051	0.039	0.000	-0.000	0.001	-0.002	0.017	0.000	0.002	-0.165	-0.000
β_{ρ^2}	0.000	-0.001	-0.000	-0.003	-0.000	-0.014	-0.013	-0.000	0.000	-0.000	0.000	-0.005	-0.000	-0.001	0.046	0.000
$\alpha_{\sigma_h^2}$	0.000	0.001	-0.000	-0.000	0.001	0.016	0.043	-0.000	0.001	-0.000	0.000	0.021	0.000	0.003	-0.007	-0.000
$\beta_{\sigma_h^2}$	-37.688	-116.830	16.011	17.100	-82.440	-2098.717	-5650.674	27.425	-71.201	10.789	37.786	-2863.039	-1.673	-434.118	958.762	4.054

Table 4: Out-of-sample validation results of ASAF for both male and female with $\beta_{\sigma_{k_m}^2}$ changed. “Bayes (mod)” is the BHM with changed hyperparameter.

Gender	Train	Test	num	Method	MAE	Coverage		
						80%	90%	95%
Male	1950-2000	2000-2015	63	Bayes(mod)	0.016	0.63	0.76	0.85
				Bayes	0.016	0.65	0.76	0.84
Female	1950-2000	2000-2015	63	Bayes(mod)	0.011	0.80	0.89	0.94
				Bayes	0.011	0.81	0.90	0.95

Table 5: Out-of-sample validation results of ASAF for both male and female with $\beta_{\sigma_{k_f}^2}$ changed. “Bayes (mod)” is the BHM with changed hyperparameter.

Gender	Train	Test	num	Method	MAE	Coverage		
						80%	90%	95%
Male	1950-2000	2000-2015	63	Bayes(mod)	0.016	0.64	0.77	0.85
				Bayes	0.016	0.65	0.76	0.84
Female	1950-2000	2000-2015	63	Bayes(mod)	0.011	0.82	0.90	0.95
				Bayes	0.011	0.81	0.90	0.95

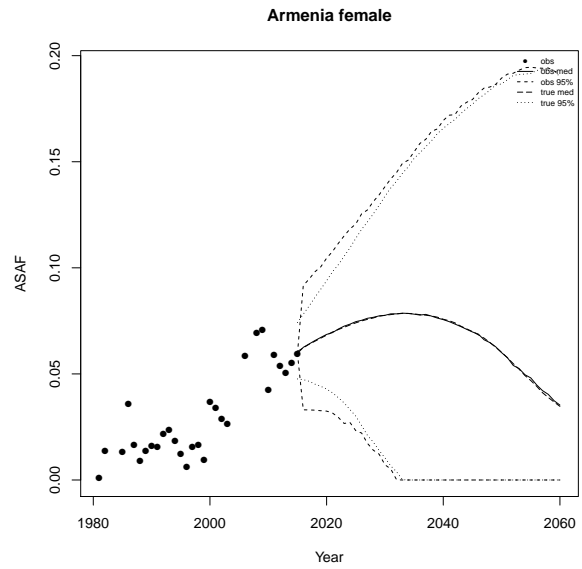
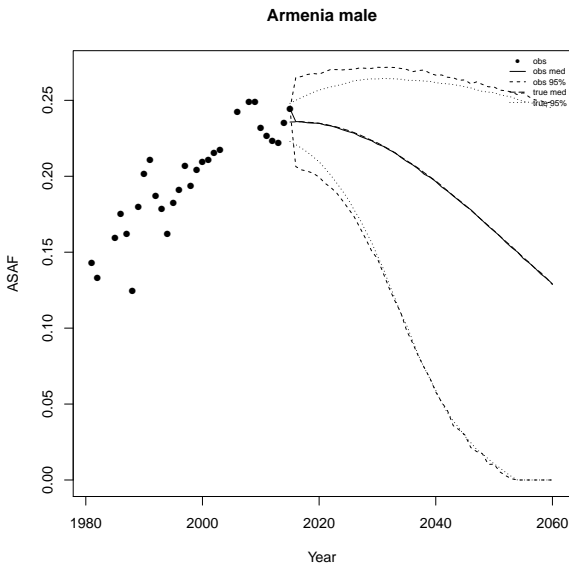
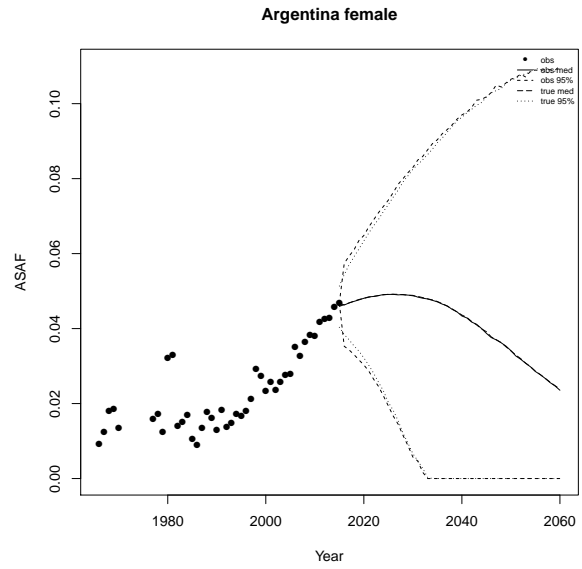
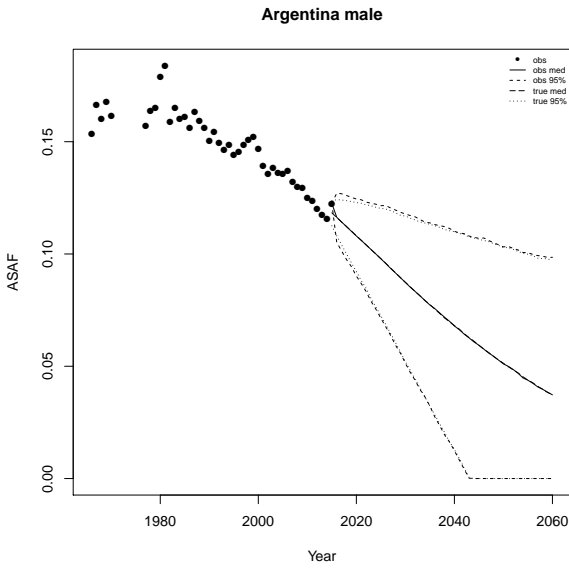
Table 6: Out-of-sample validation results of ASAF for both male and female with $\beta_{\sigma_h^2}$ changed. “Bayes (mod)” is the BHM with changed hyperparameter.

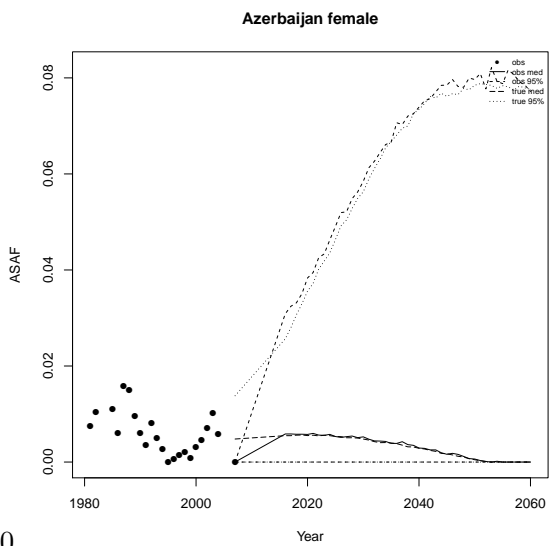
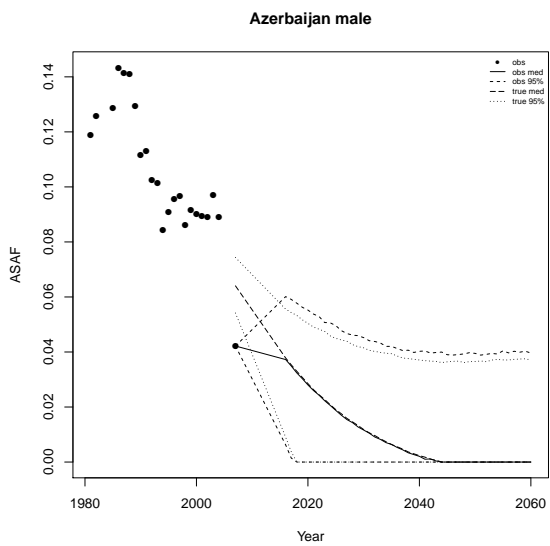
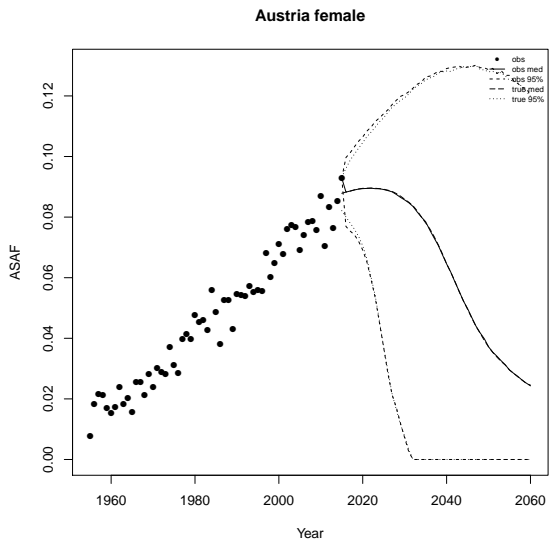
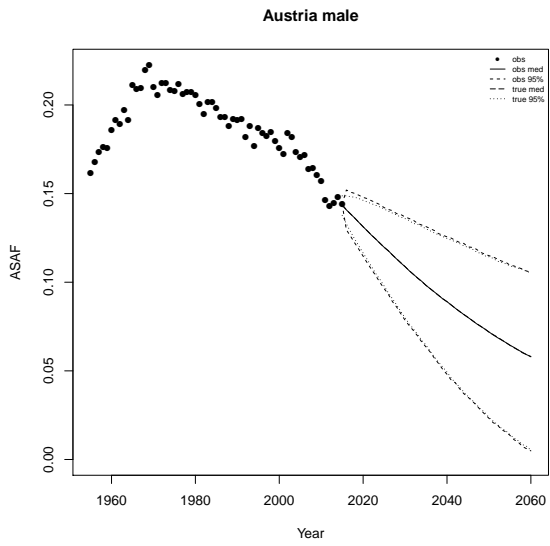
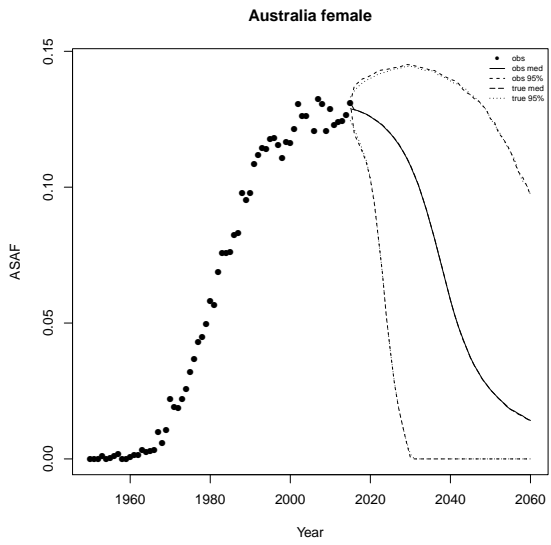
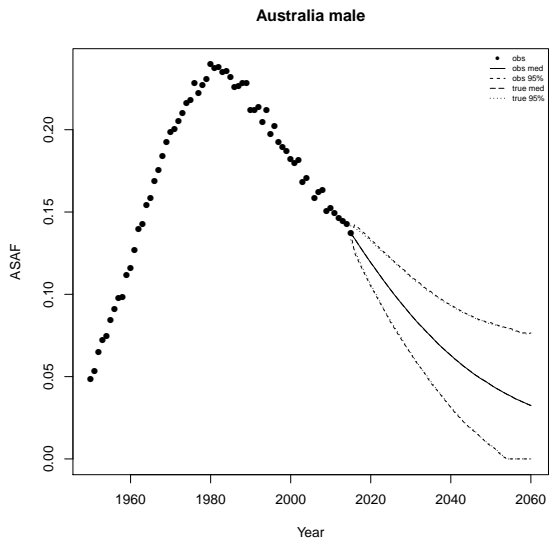
Gender	Train	Test	num	Method	MAE	Coverage		
						80%	90%	95%
Male	1950-2000	2000-2015	63	Bayes(mod)	0.016	0.68	0.80	0.88
				Bayes	0.016	0.65	0.76	0.84
Female	1950-2000	2000-2015	63	Bayes(mod)	0.011	0.82	0.90	0.95
				Bayes	0.011	0.81	0.90	0.95

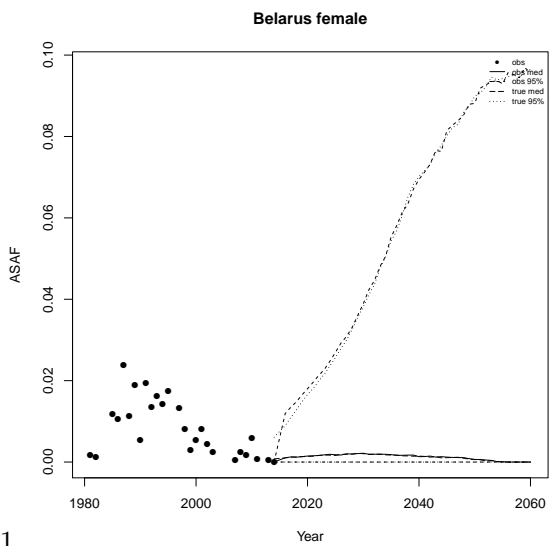
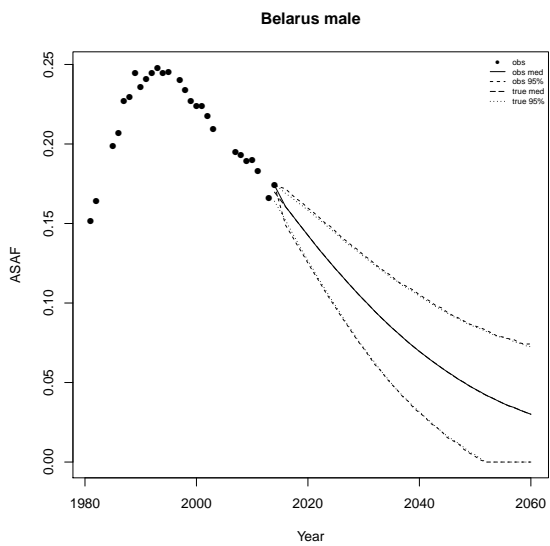
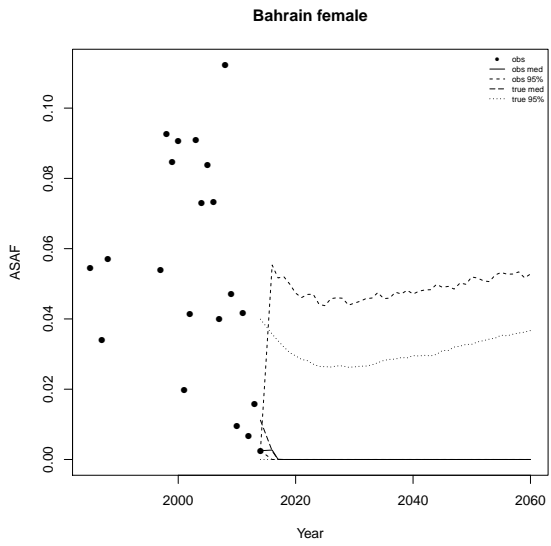
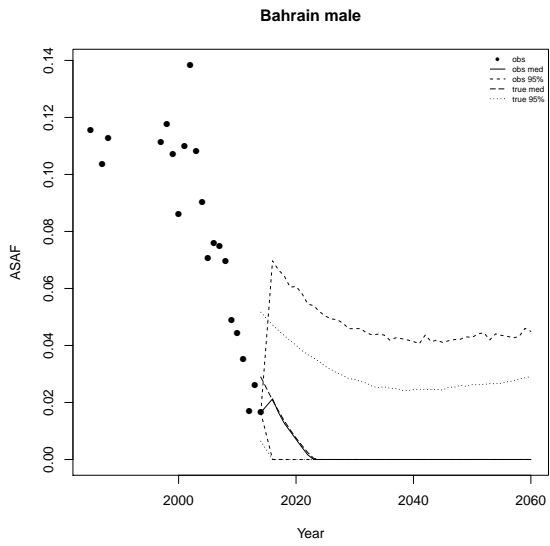
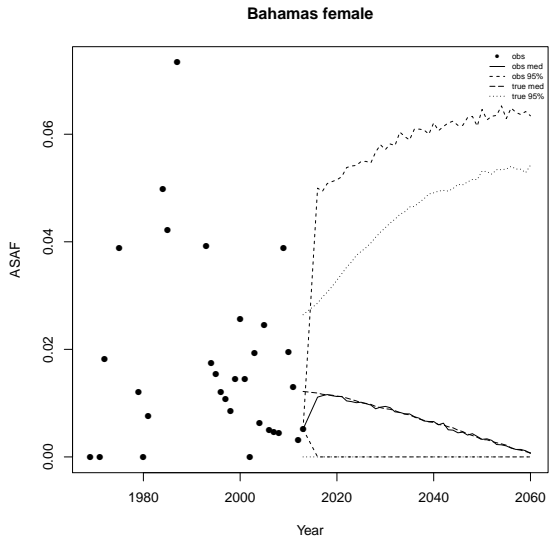
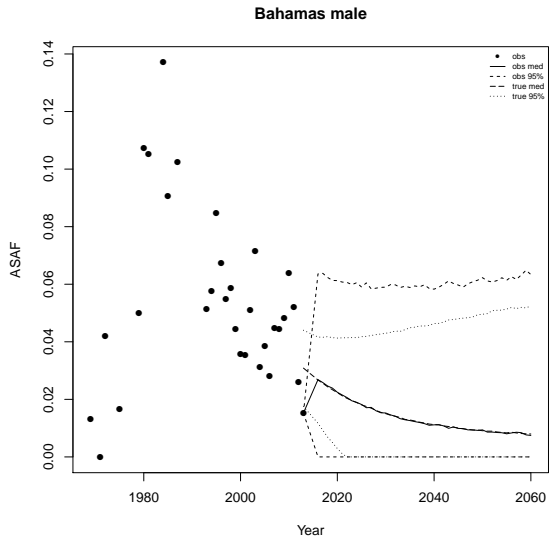
Table 7: Out-of-sample validation results of ASAF for both male and female with α_{a_4} changed. “Bayes (mod)” is the BHM with changed hyperparameter.

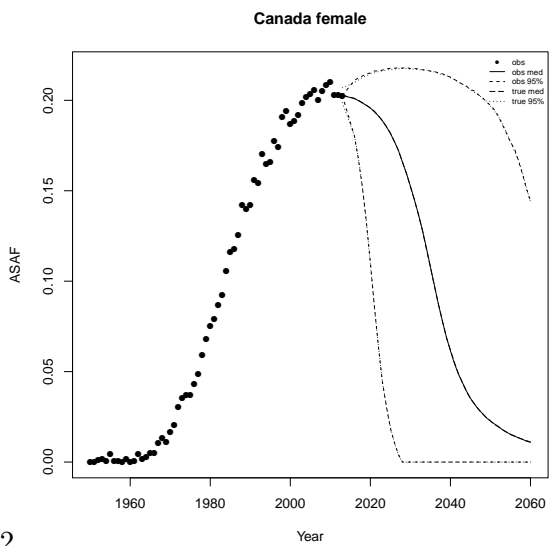
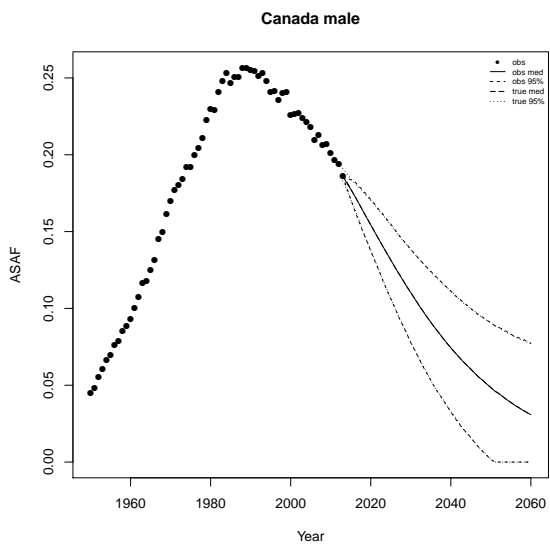
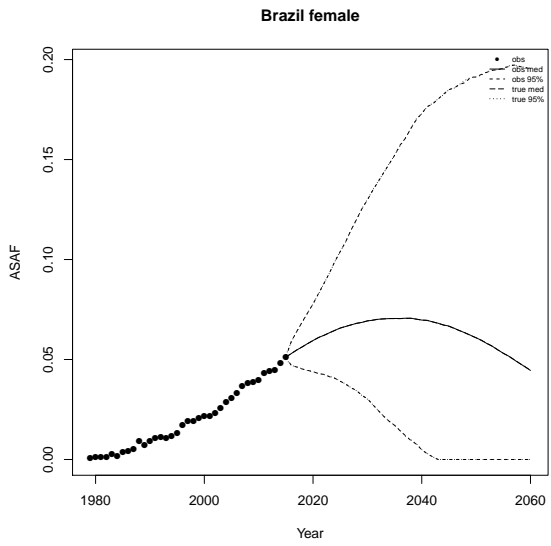
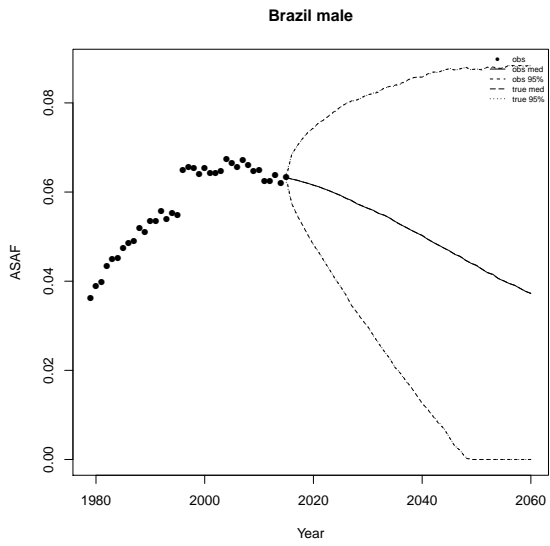
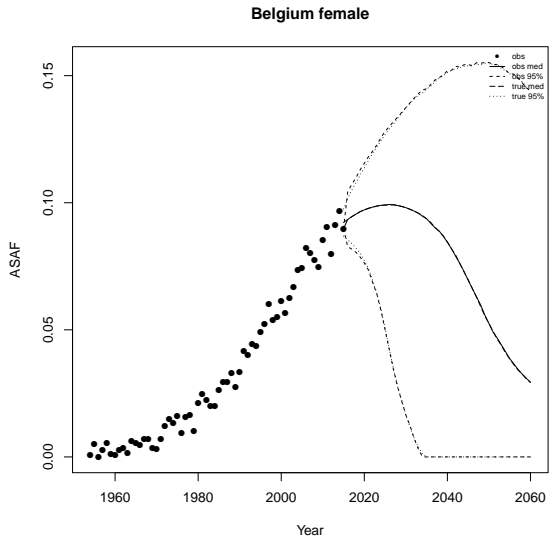
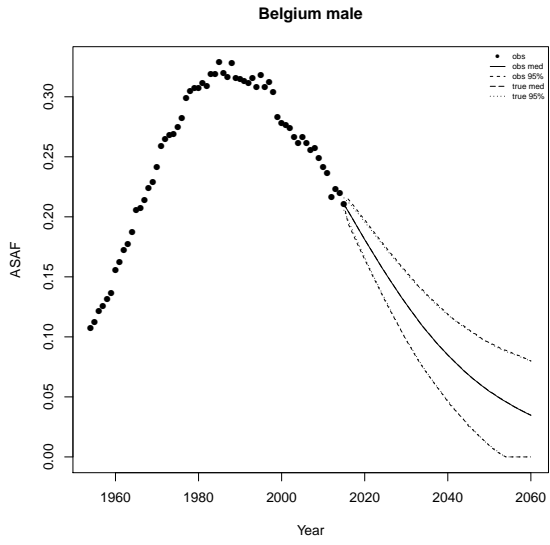
Gender	Train	Test	num	Method	MAE	Coverage		
						80%	90%	95%
Male	1950-2000	2000-2015	63	Bayes(mod)	0.016	0.65	0.76	0.85
				Bayes	0.016	0.65	0.76	0.84
Female	1950-2000	2000-2015	63	Bayes(mod)	0.011	0.81	0.89	0.95
				Bayes	0.011	0.81	0.90	0.95

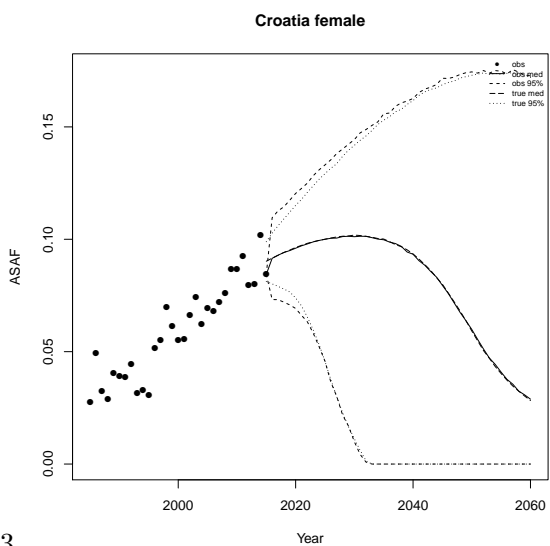
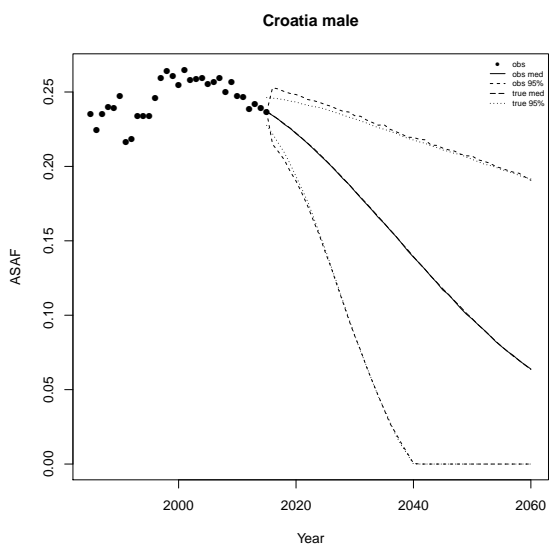
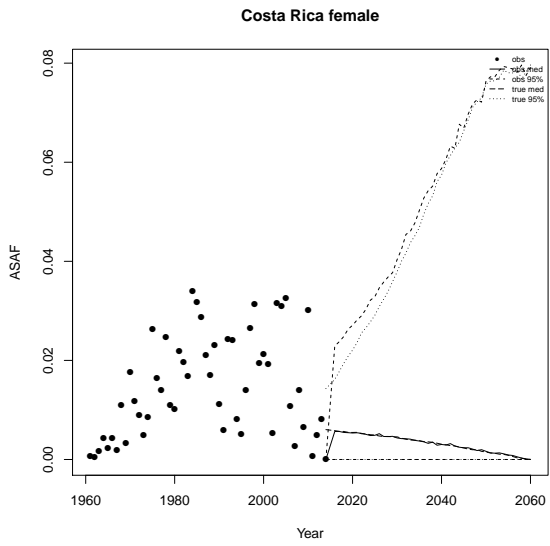
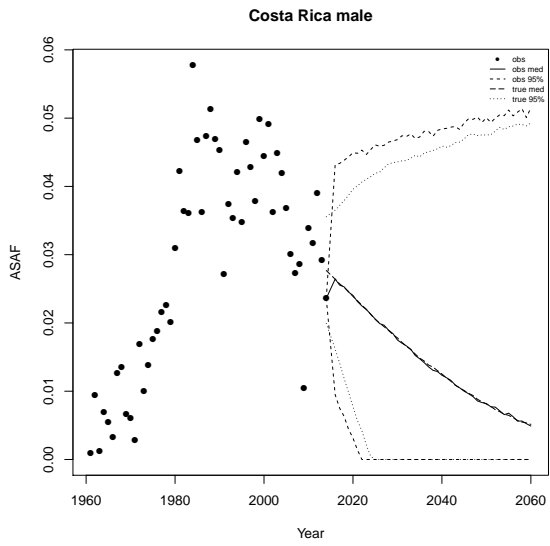
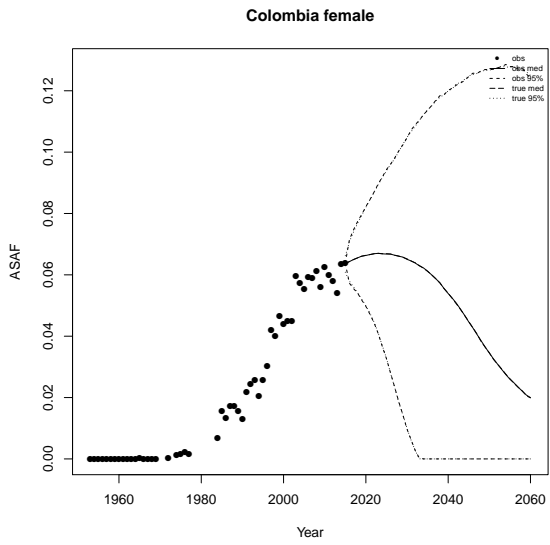
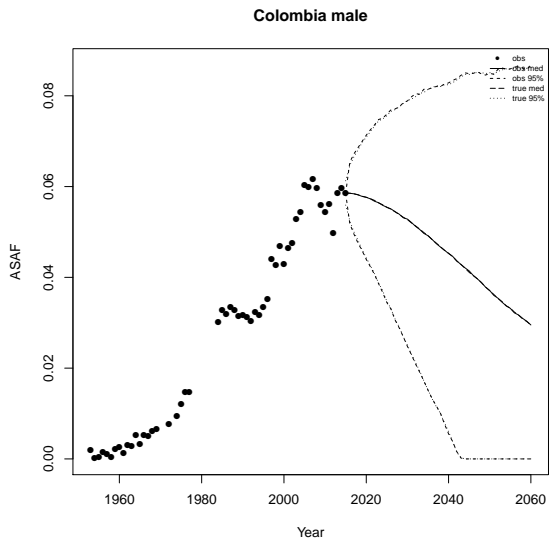
3 All-age Smoking Attributable Fraction Projection for Both Genders in Over 60 Countries

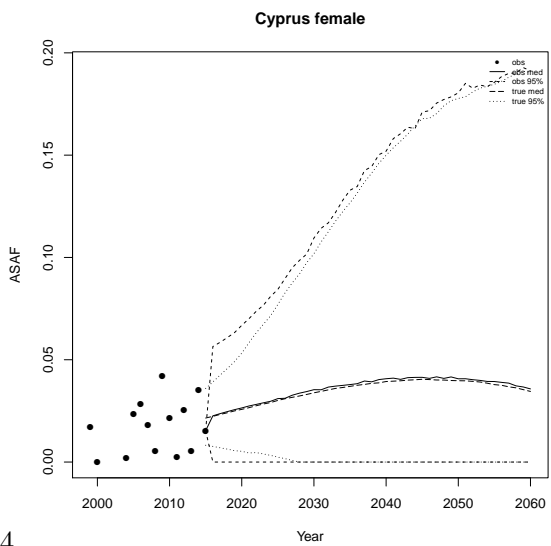
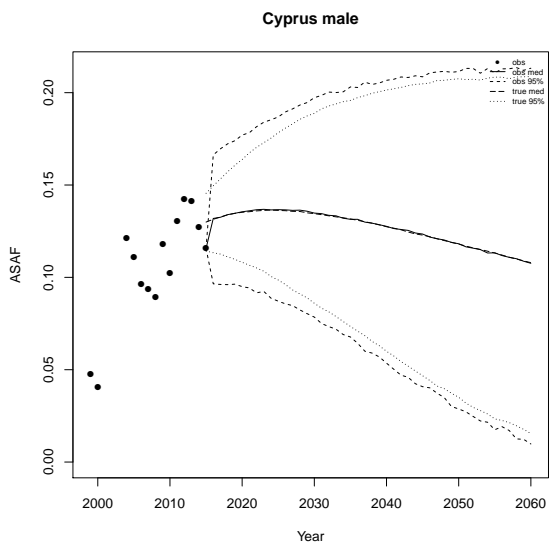
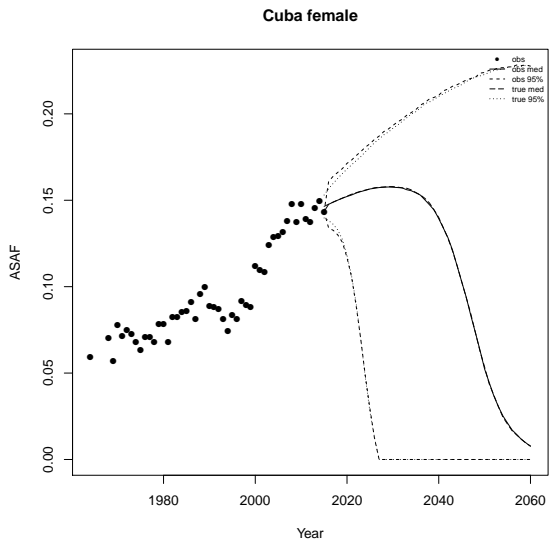
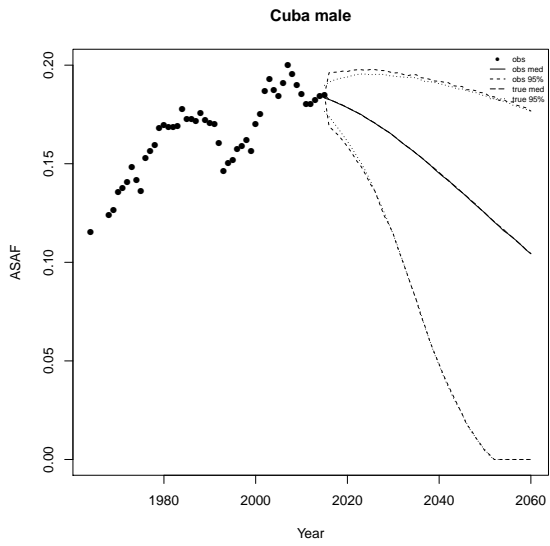
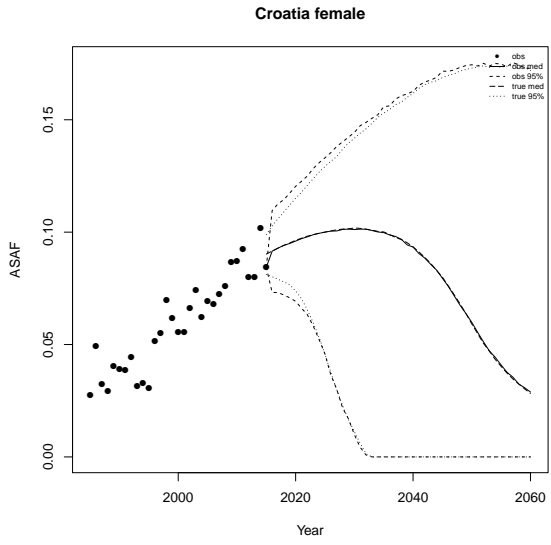
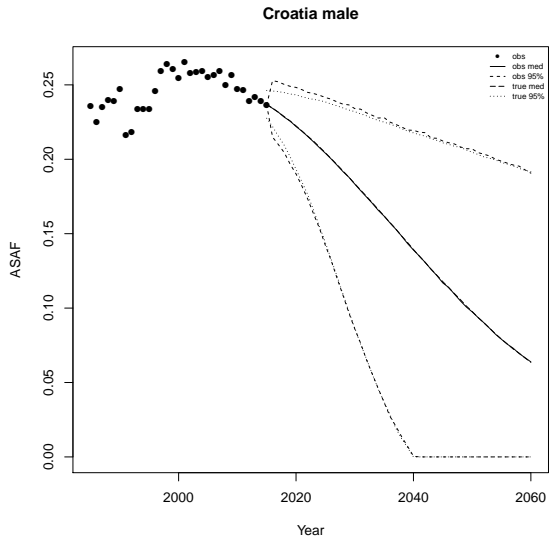


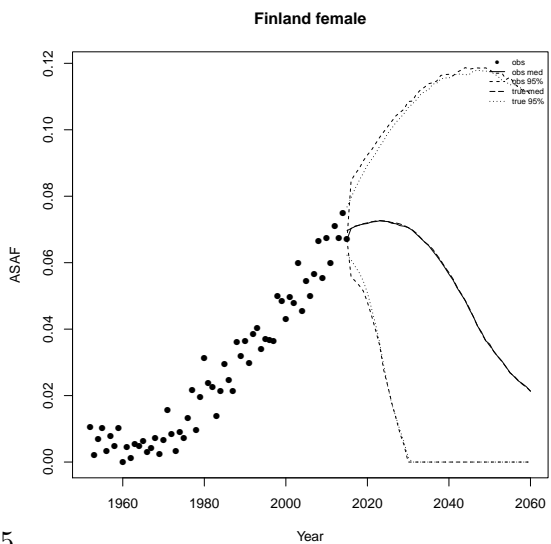
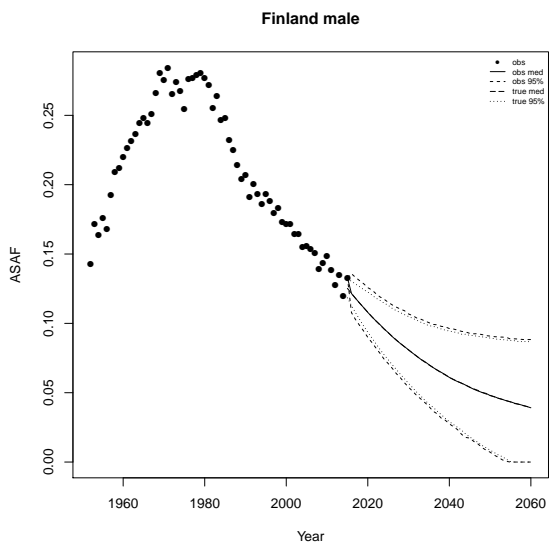
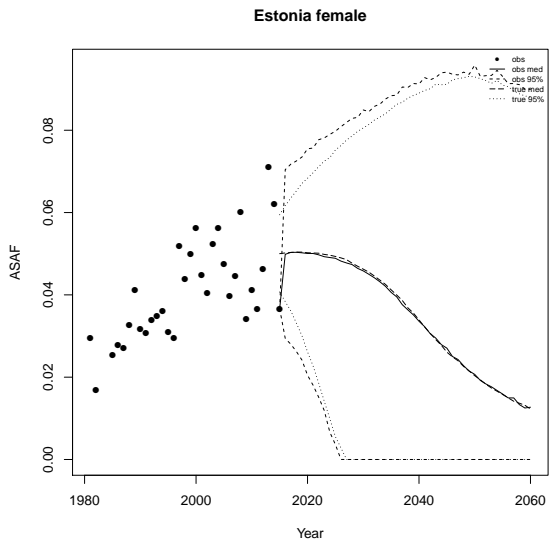
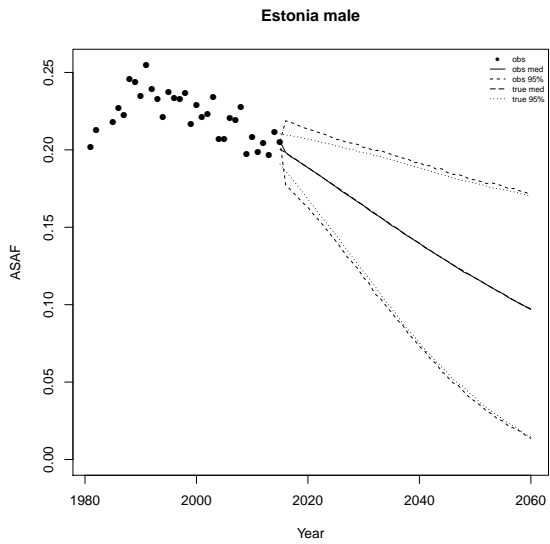
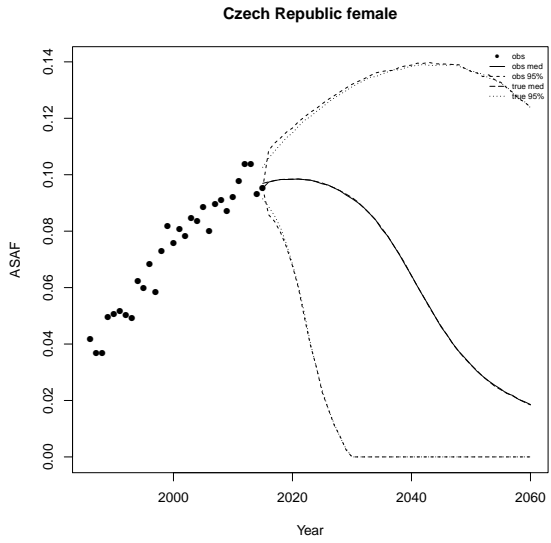
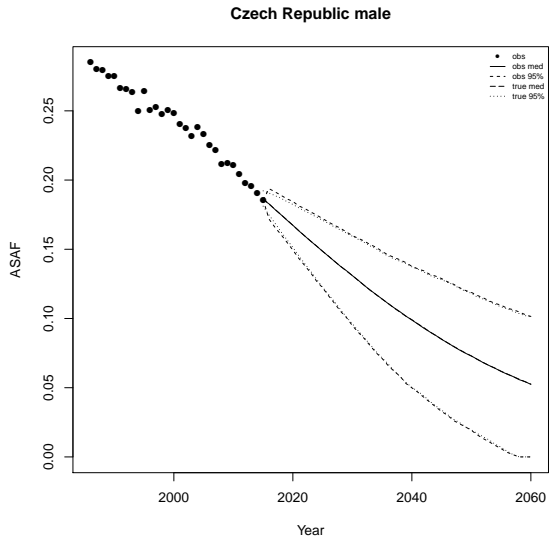


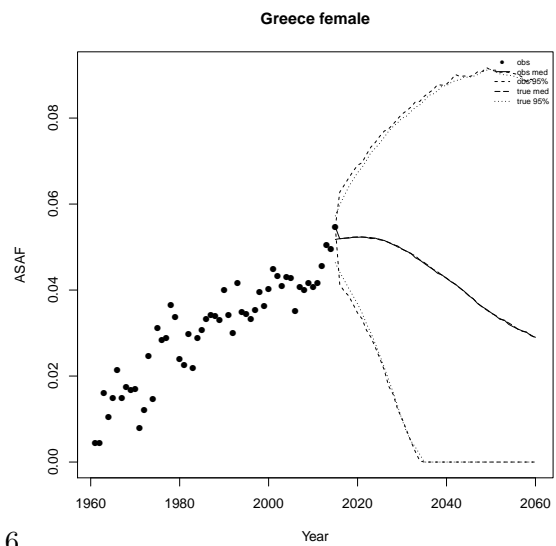
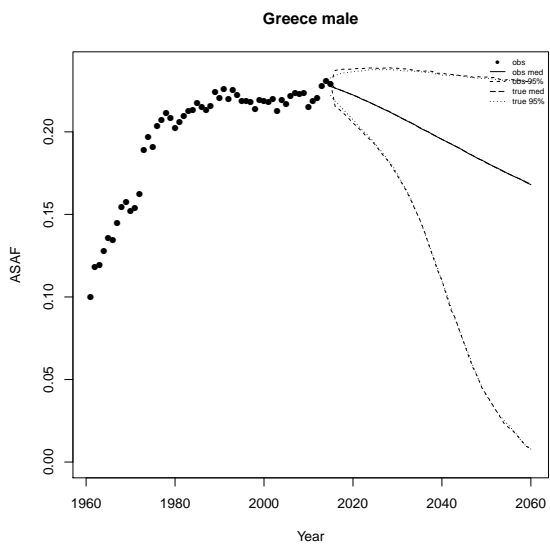
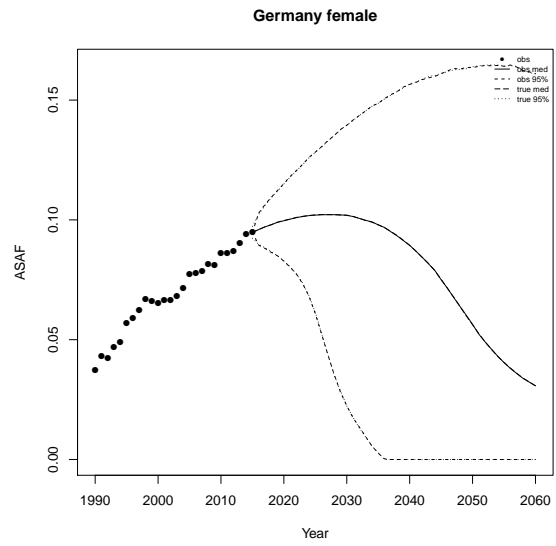
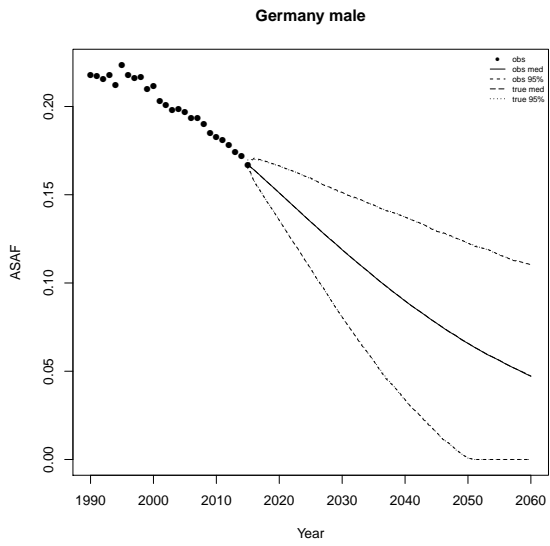
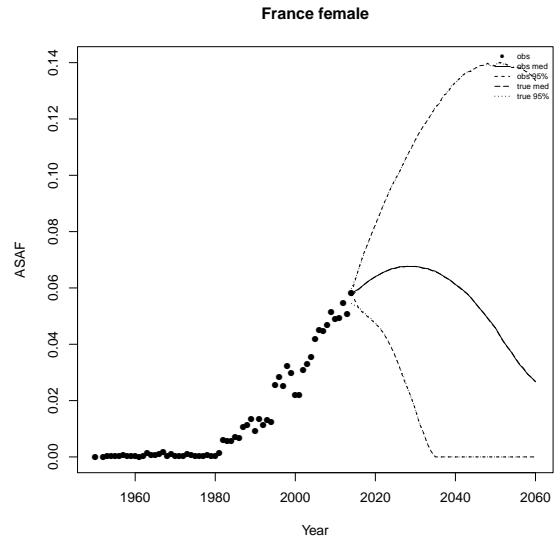
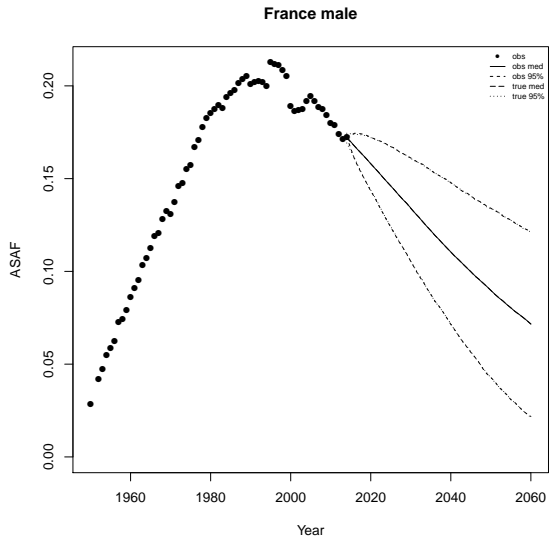


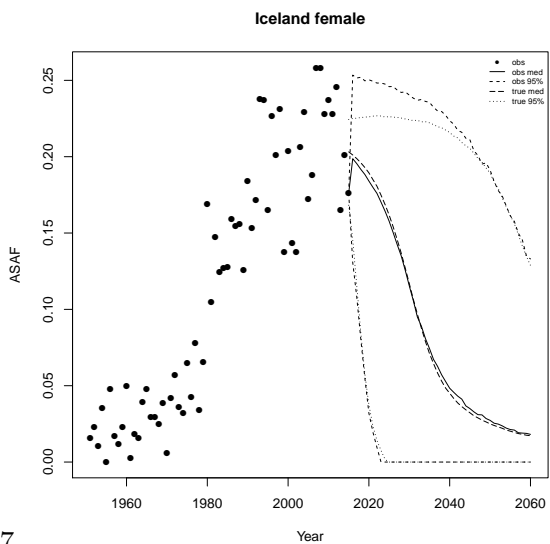
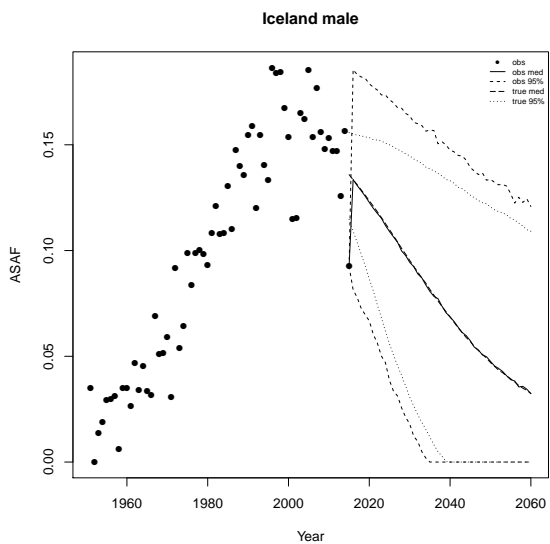
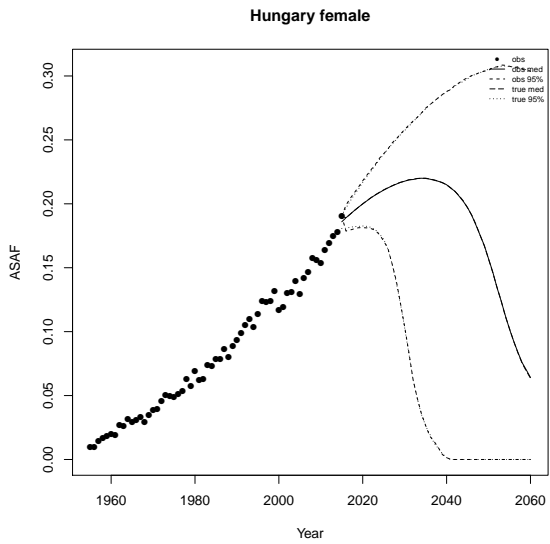
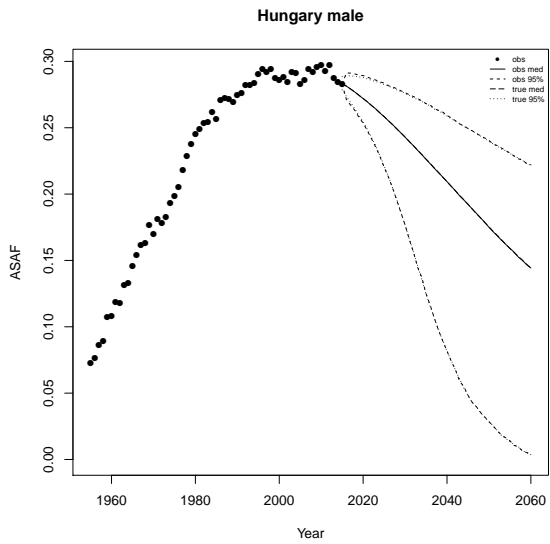
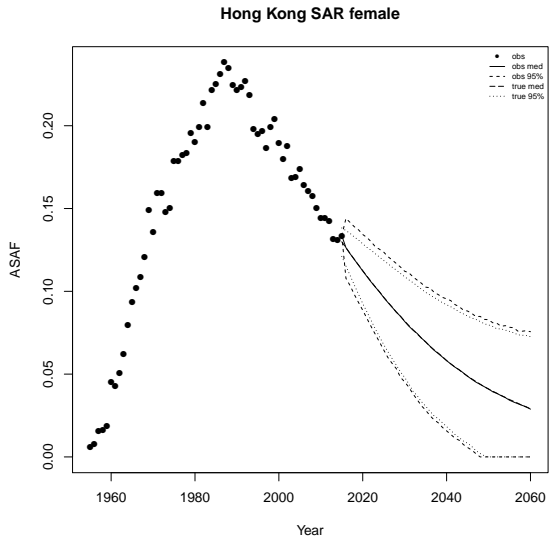
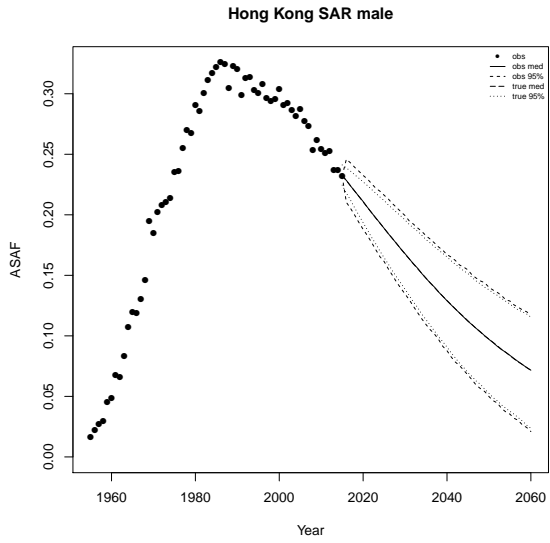


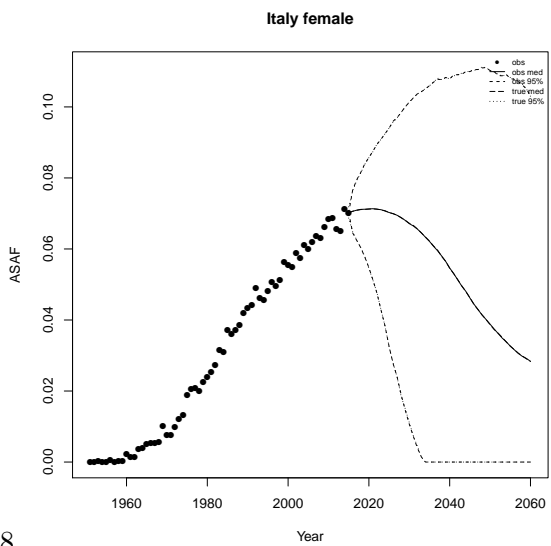
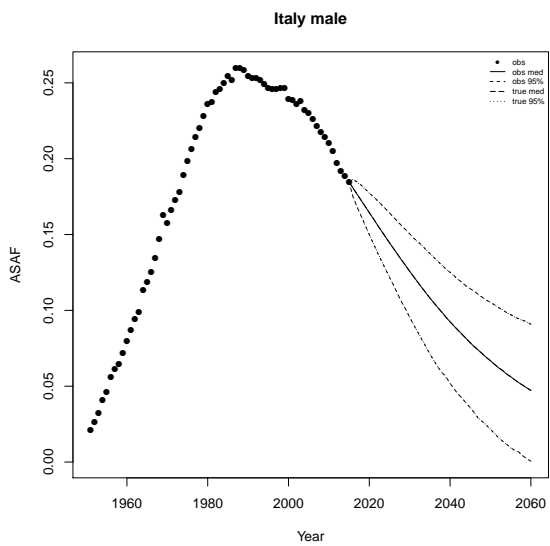
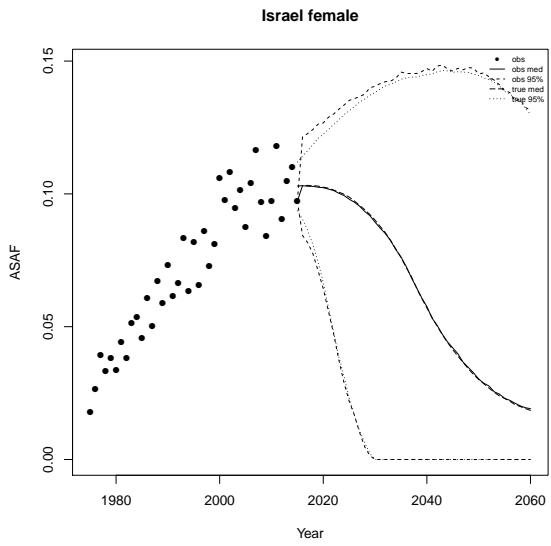
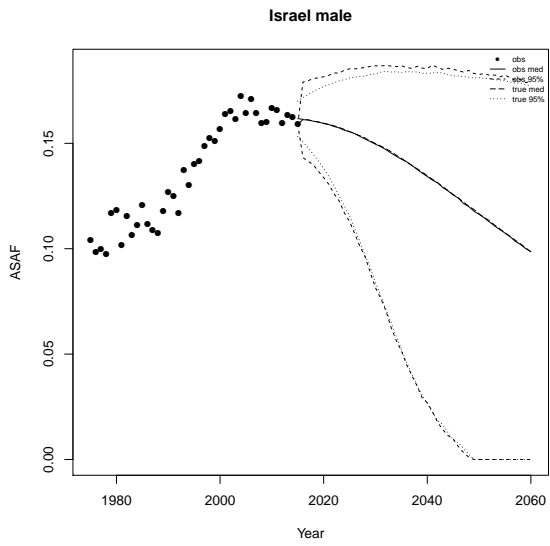
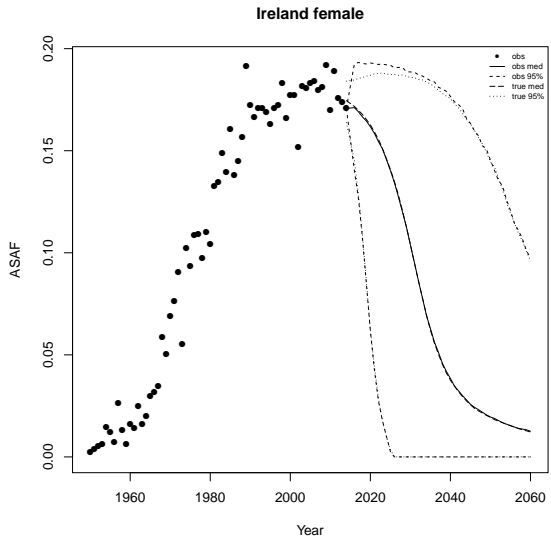
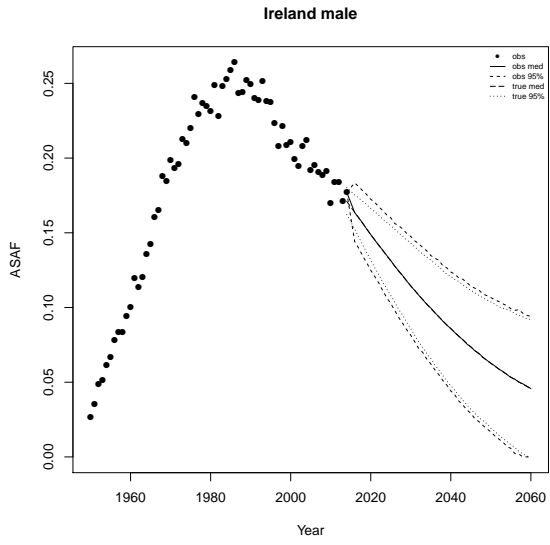


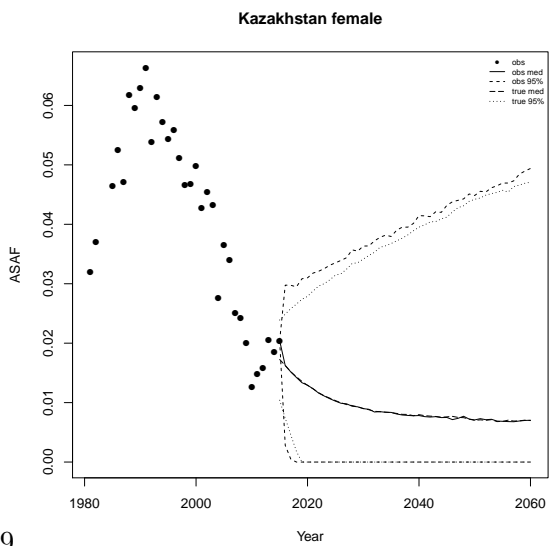
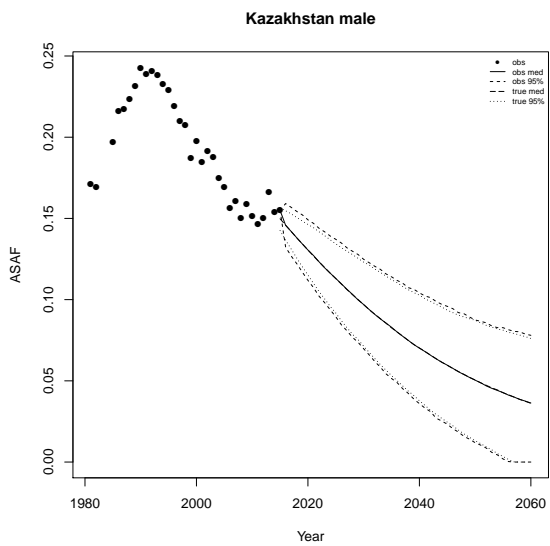
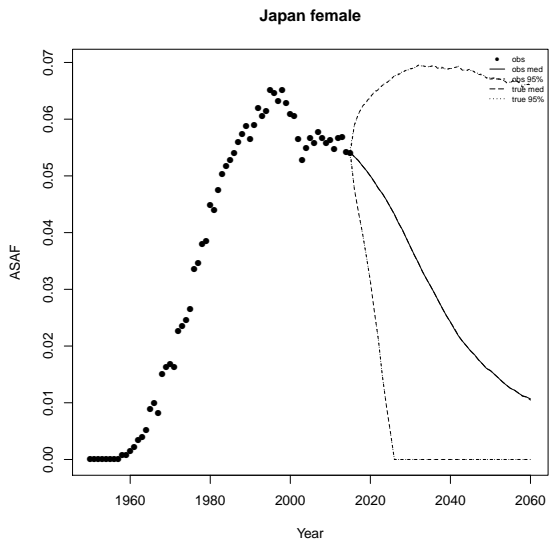
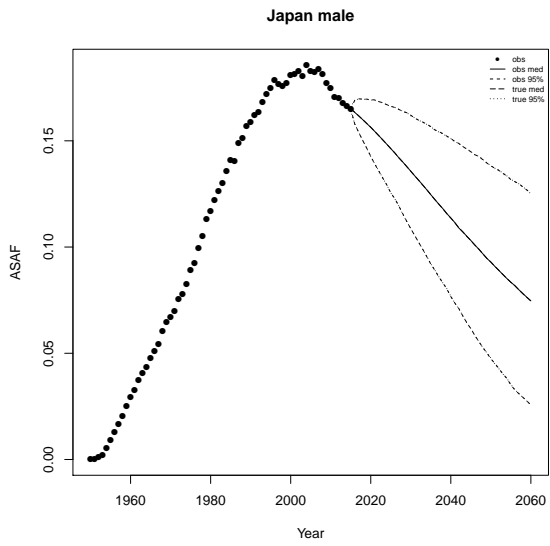
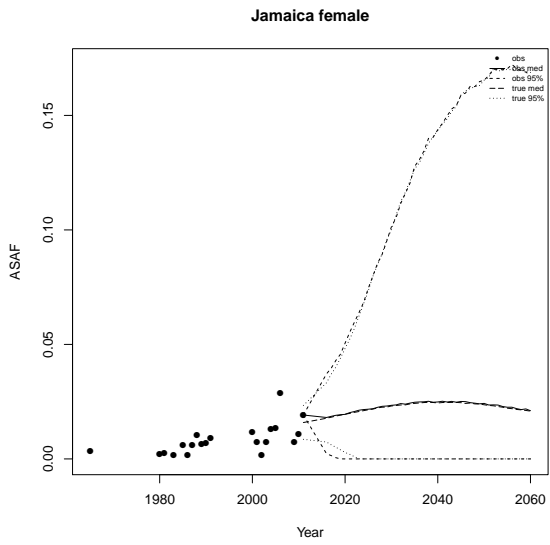
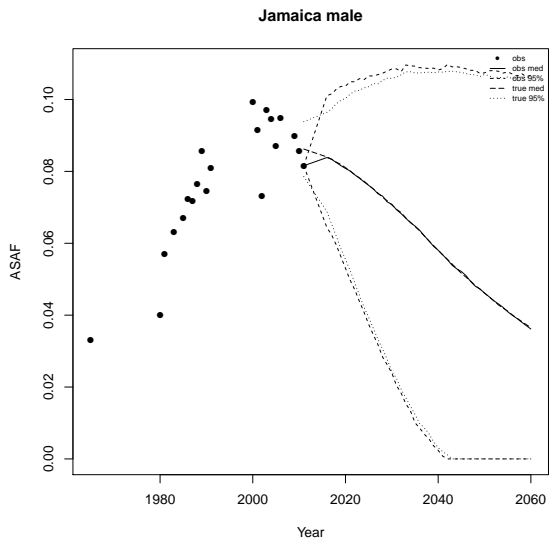


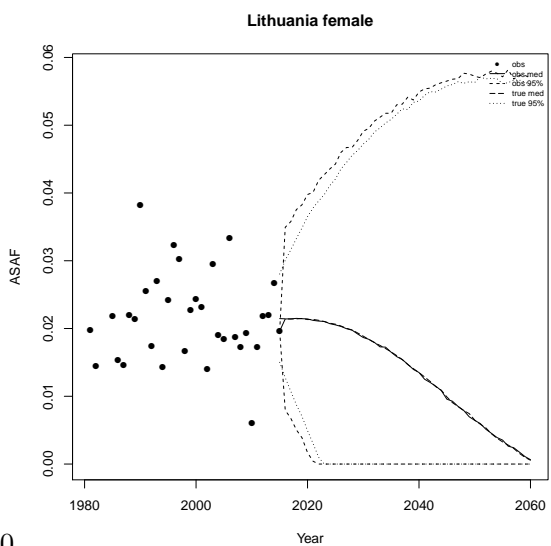
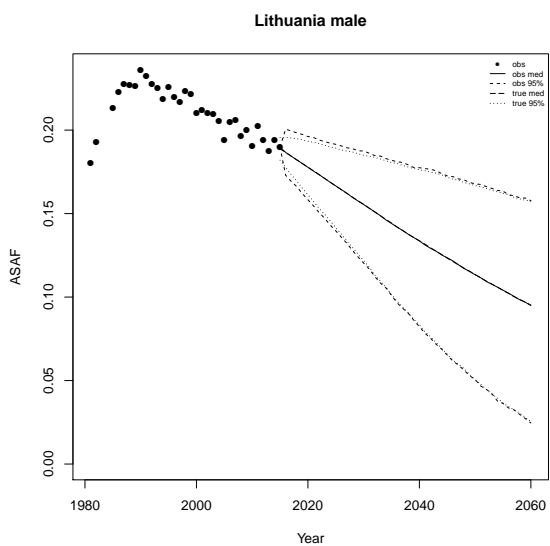
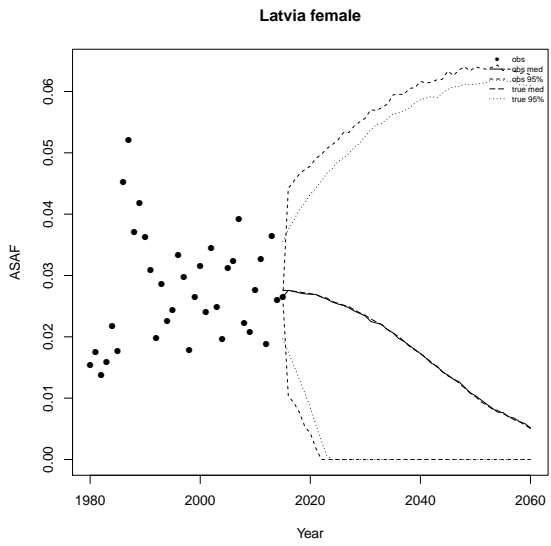
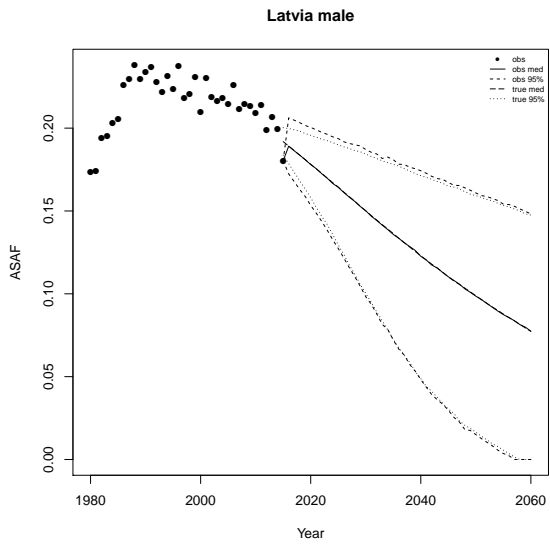
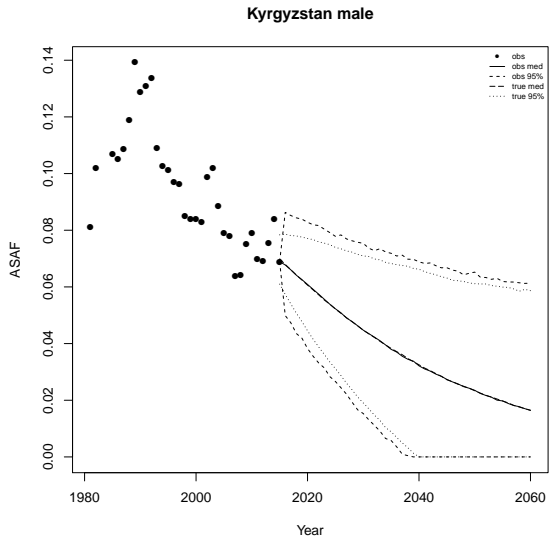
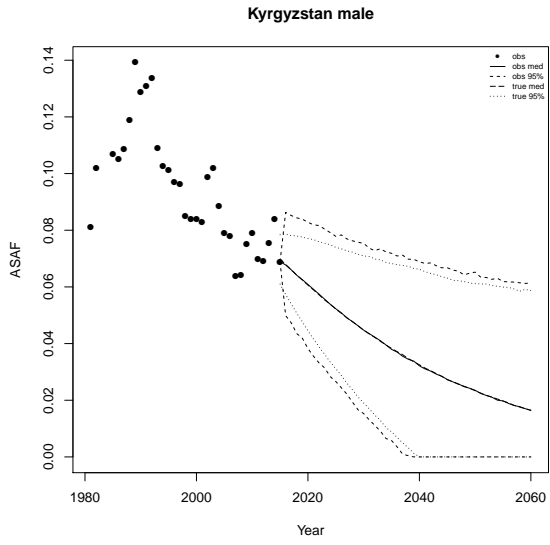


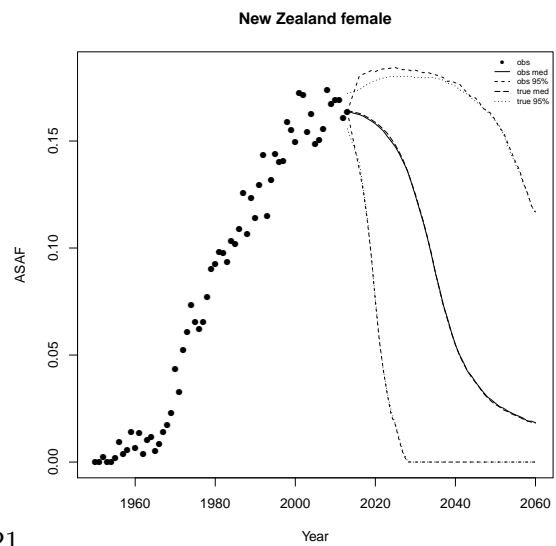
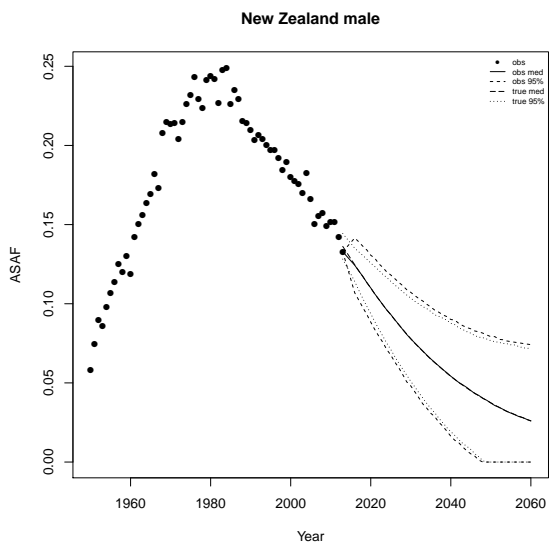
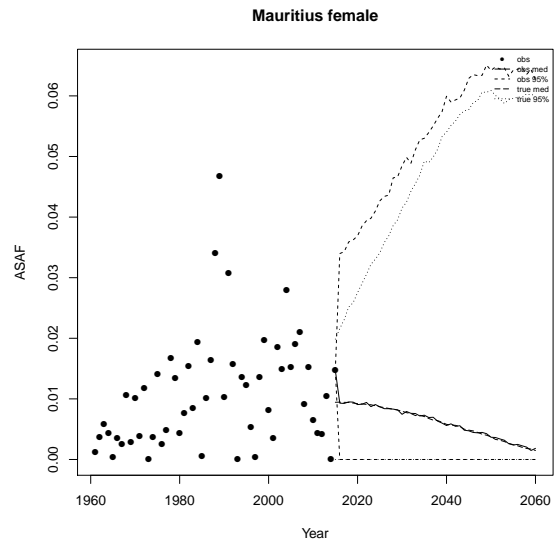
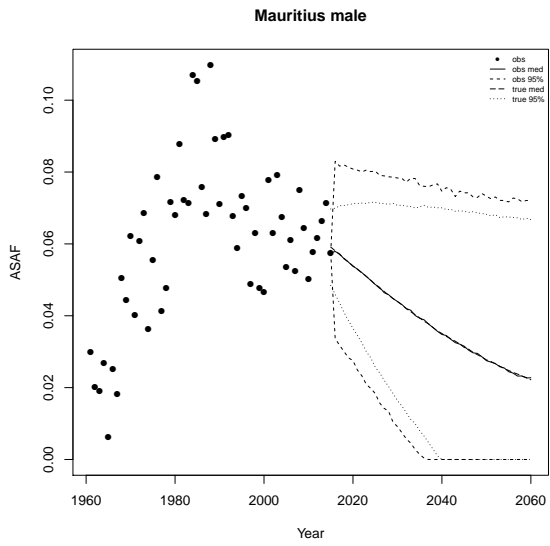
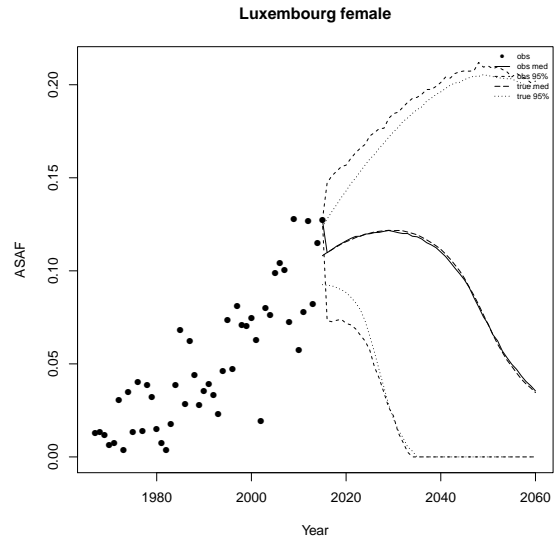
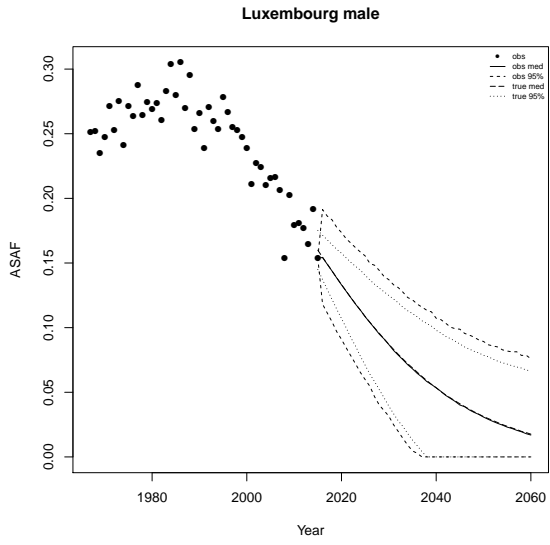


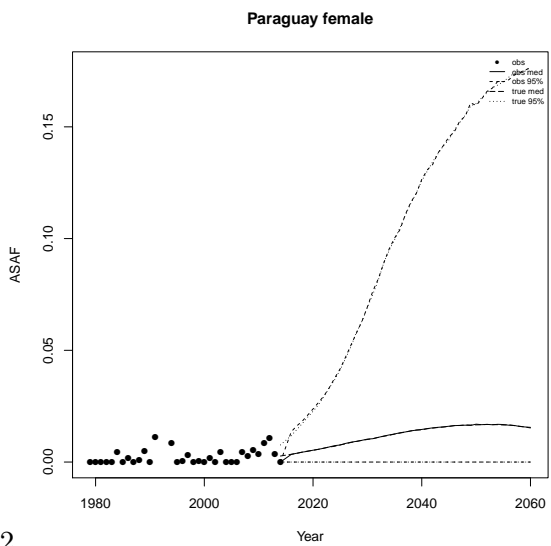
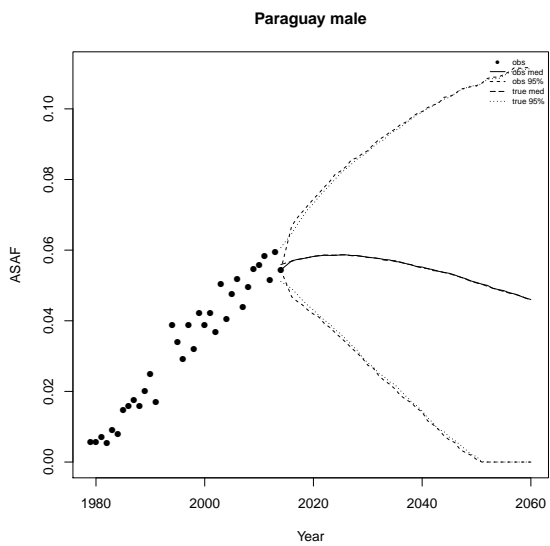
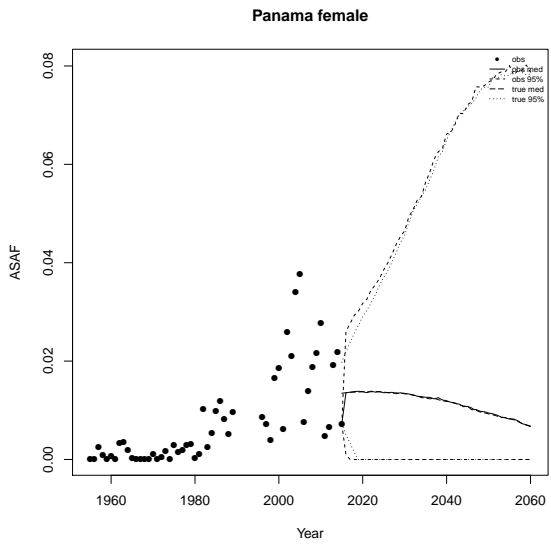
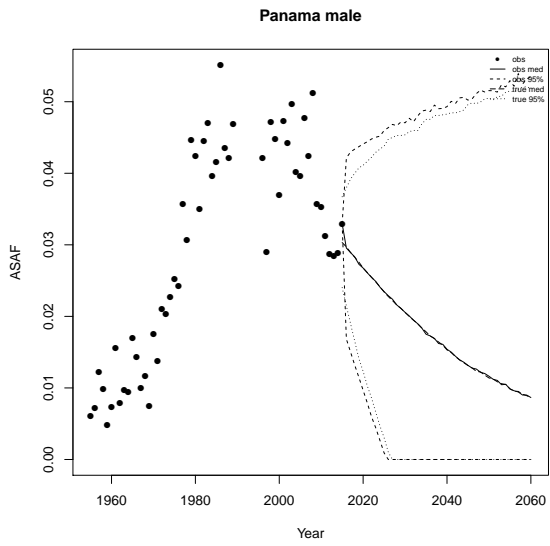
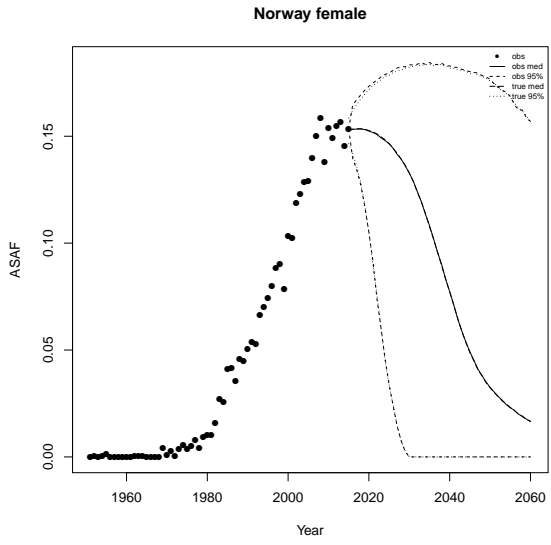
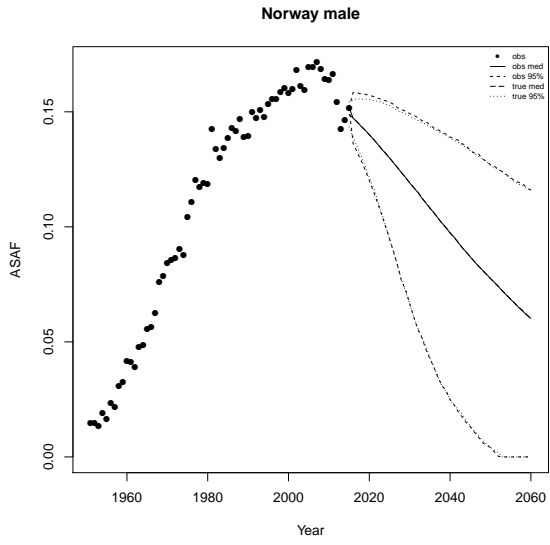


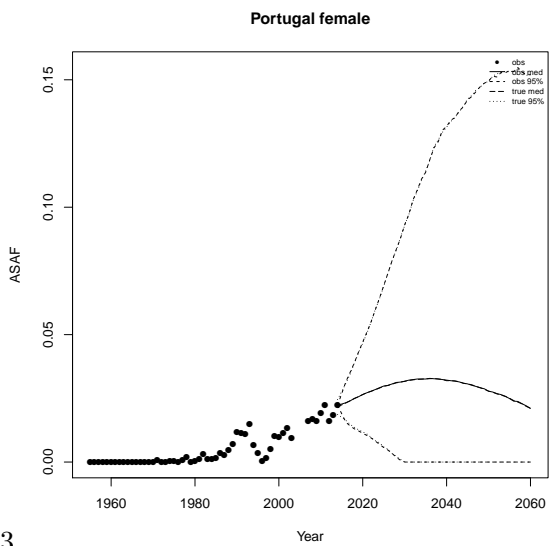
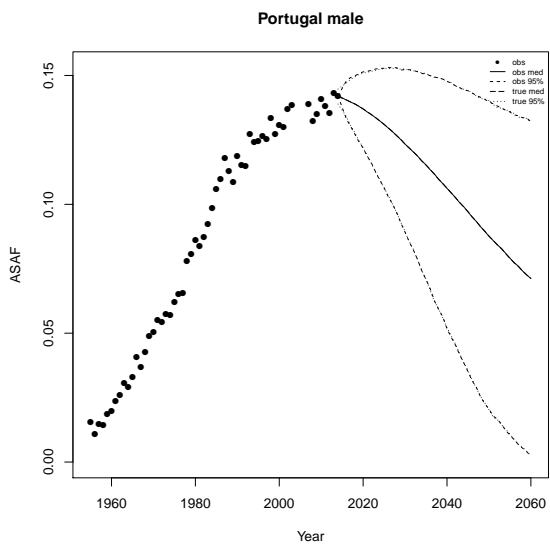
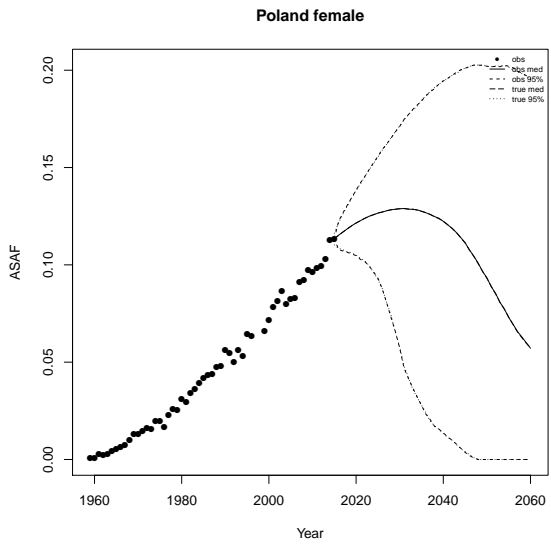
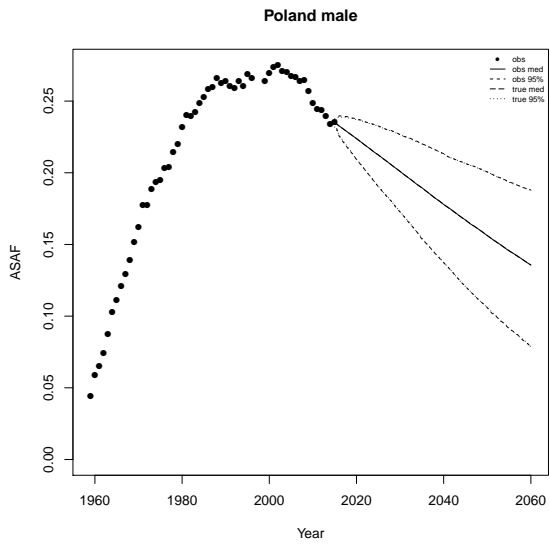
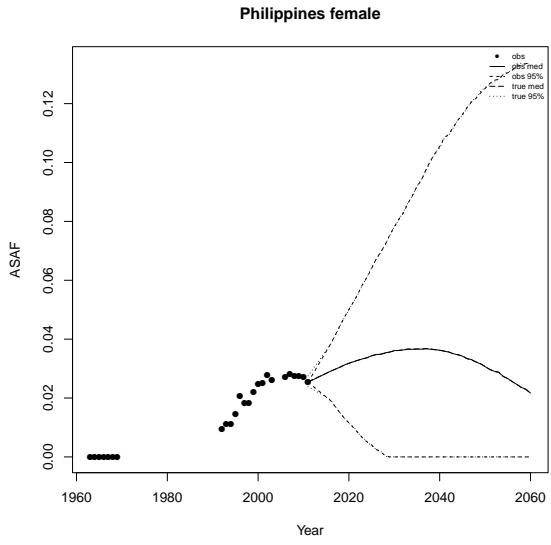
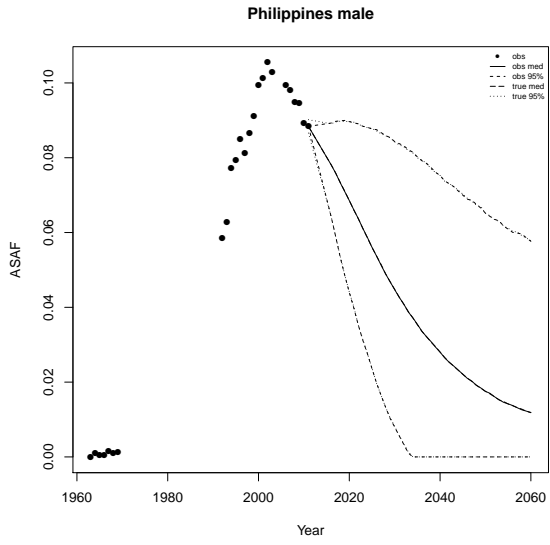


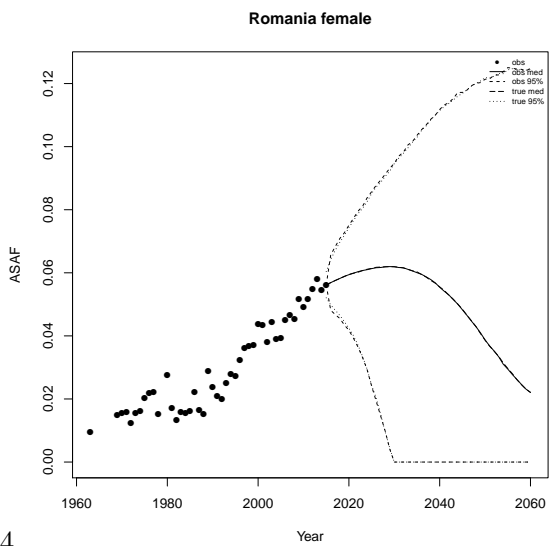
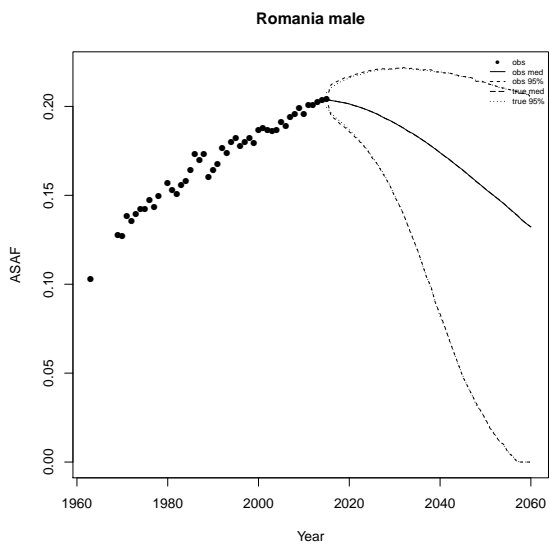
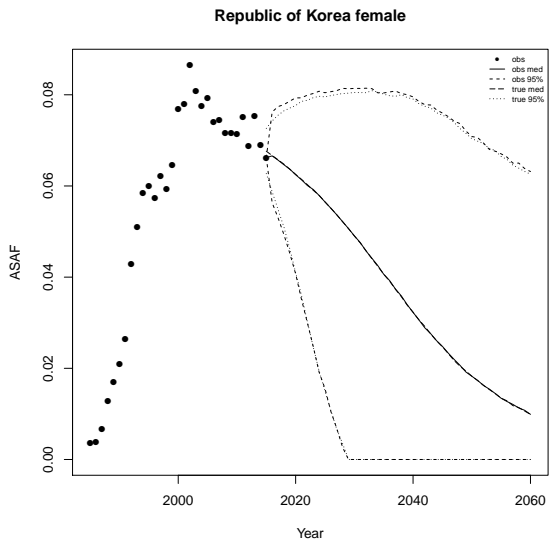
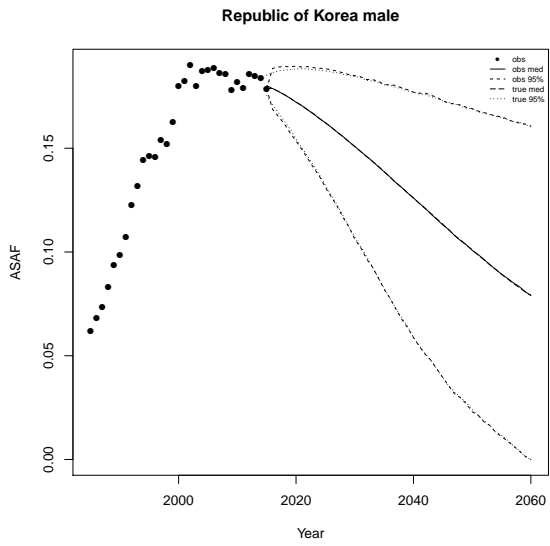
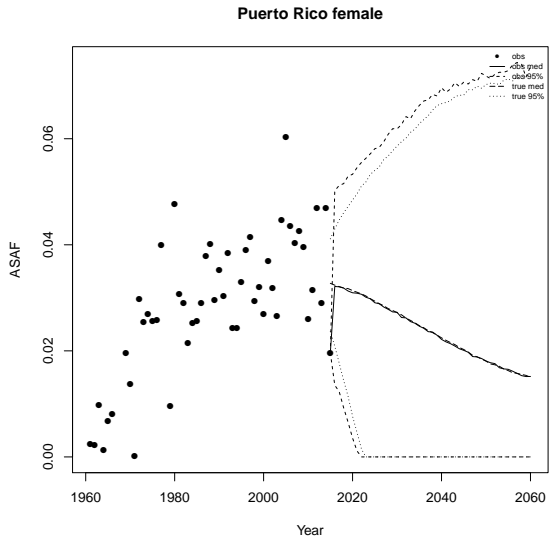
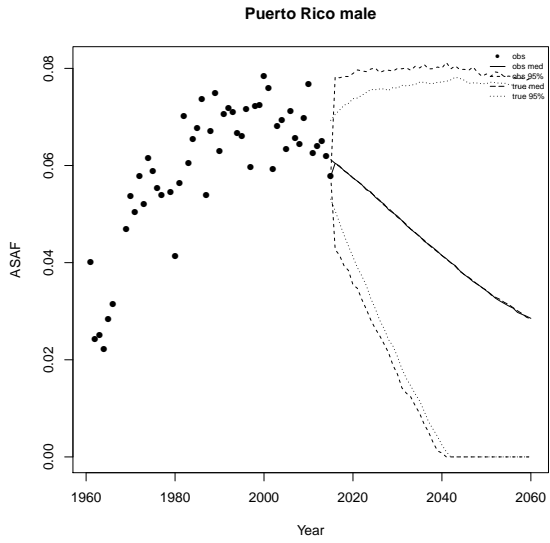


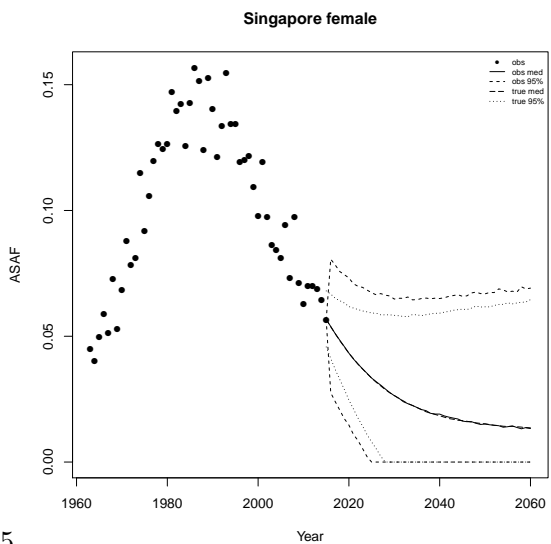
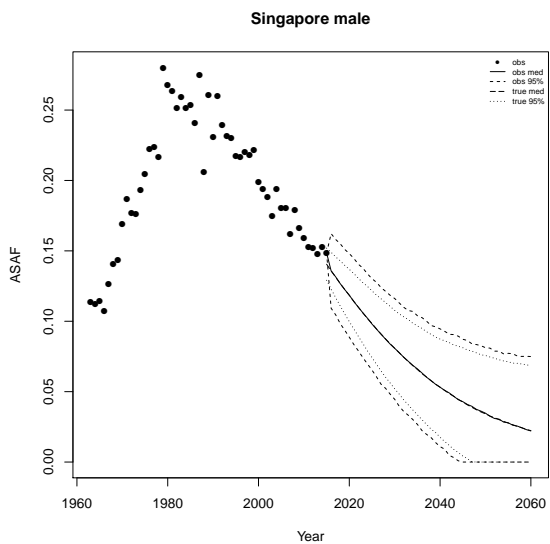
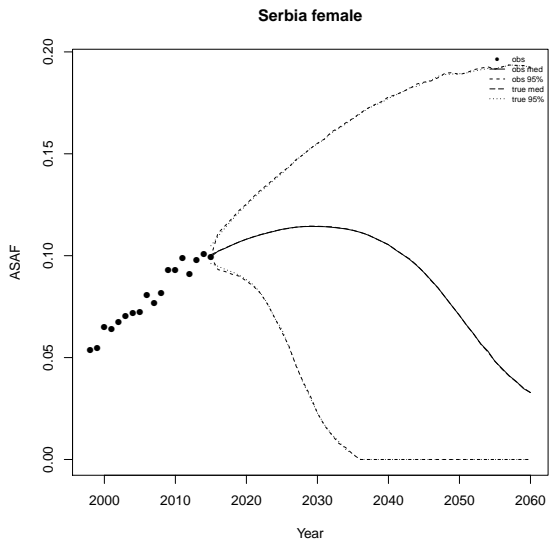
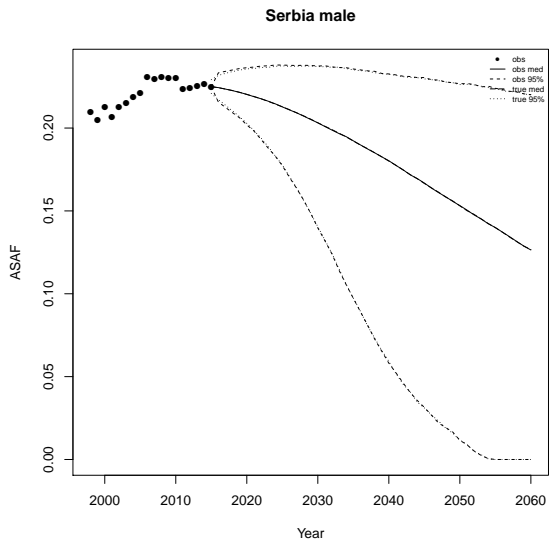
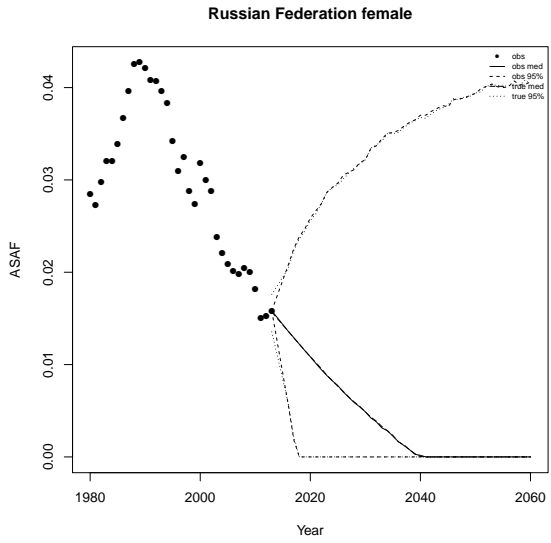
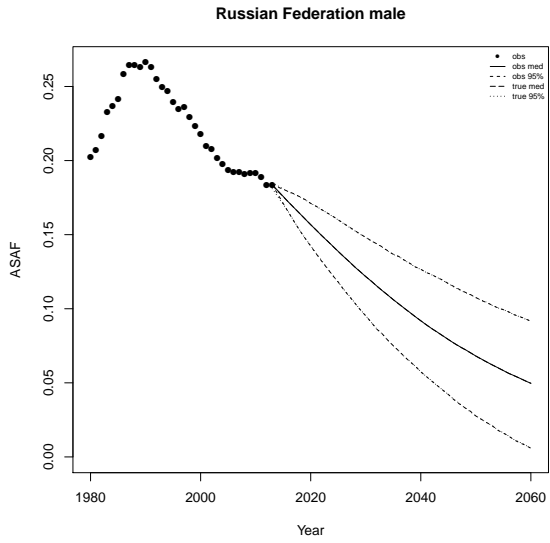


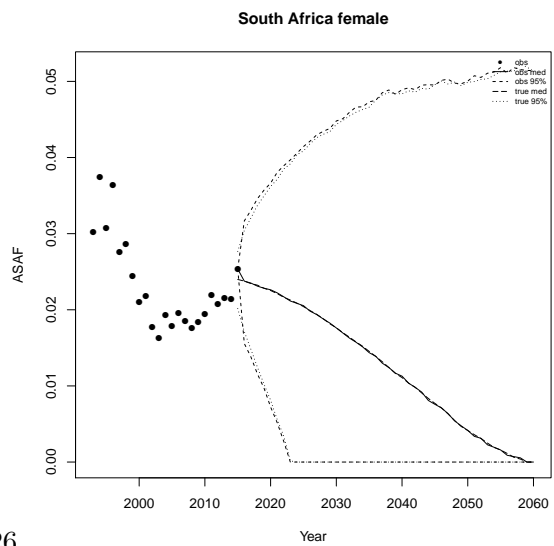
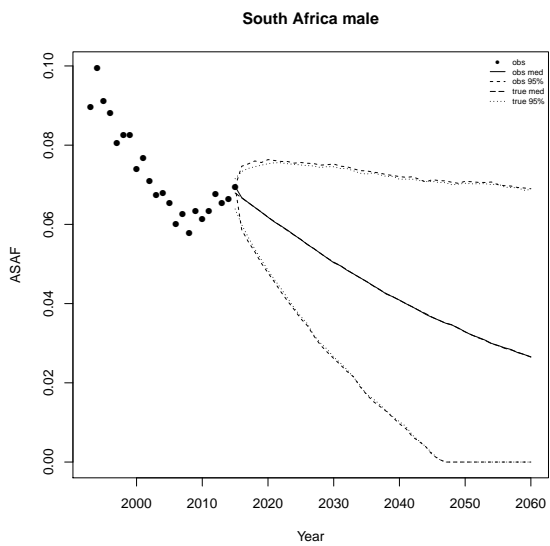
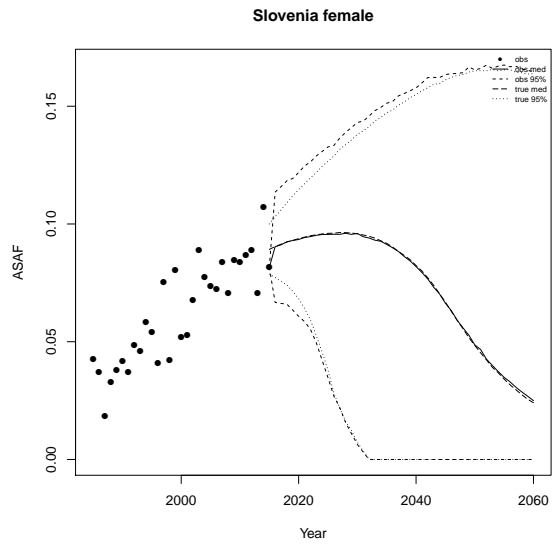
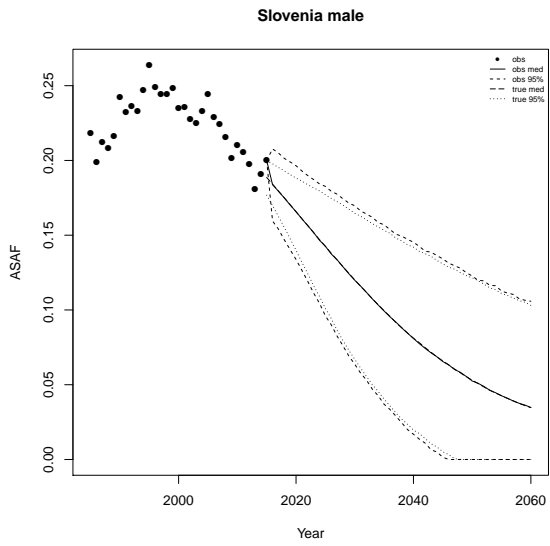
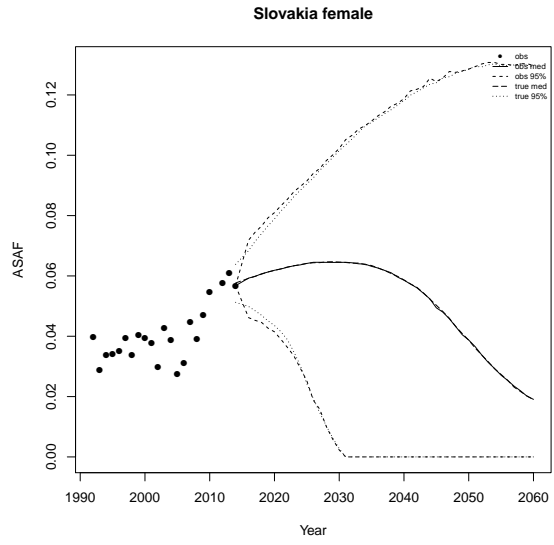
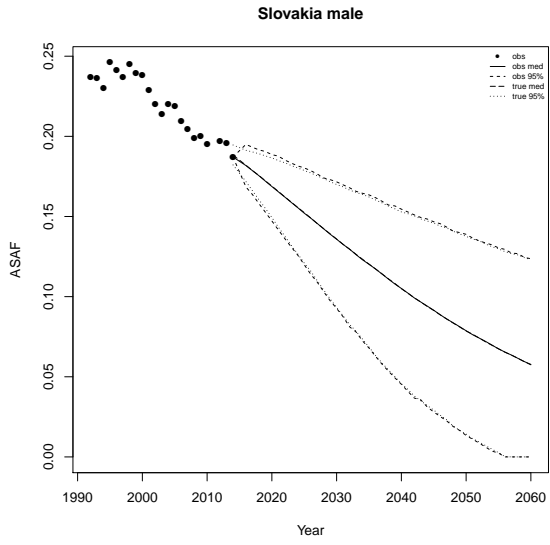


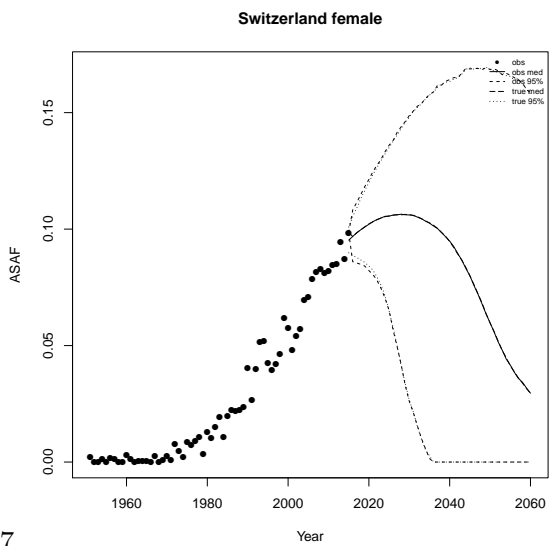
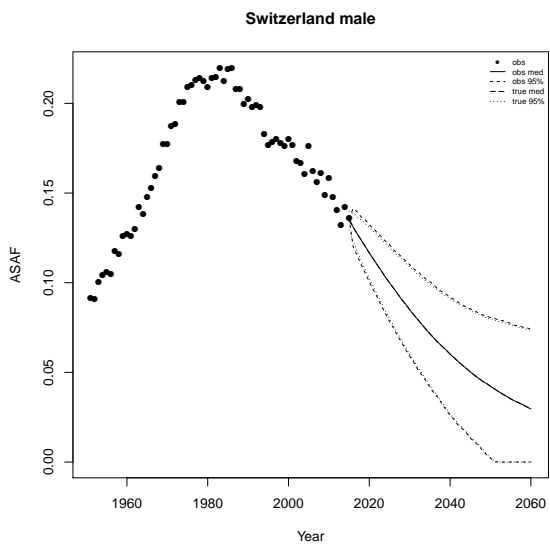
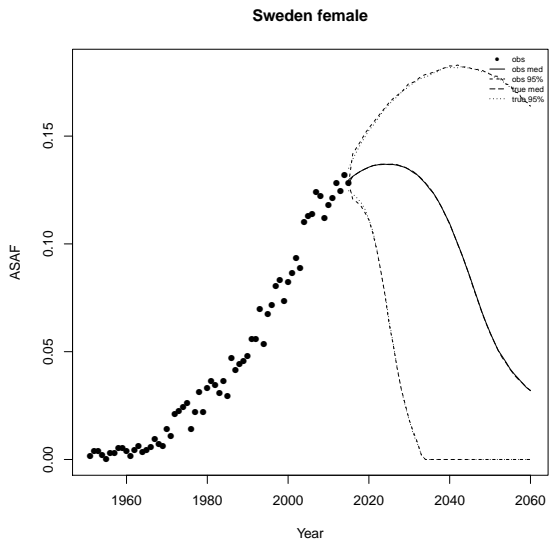
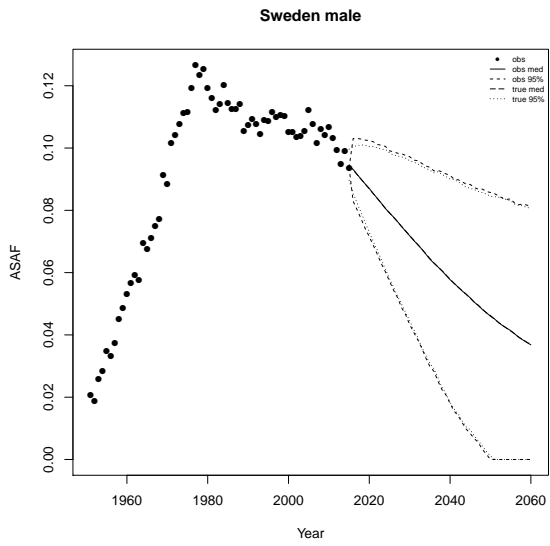
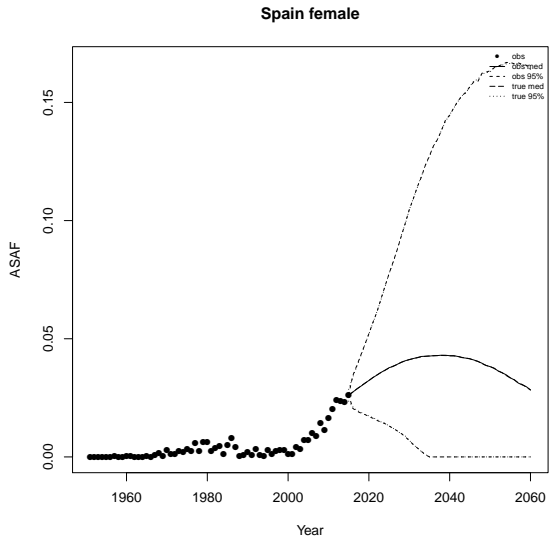
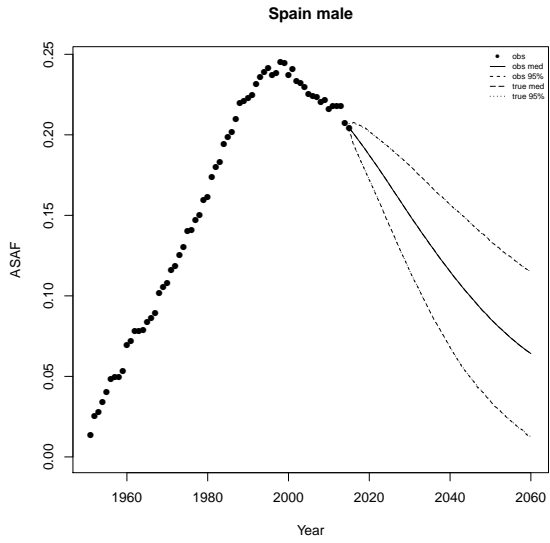




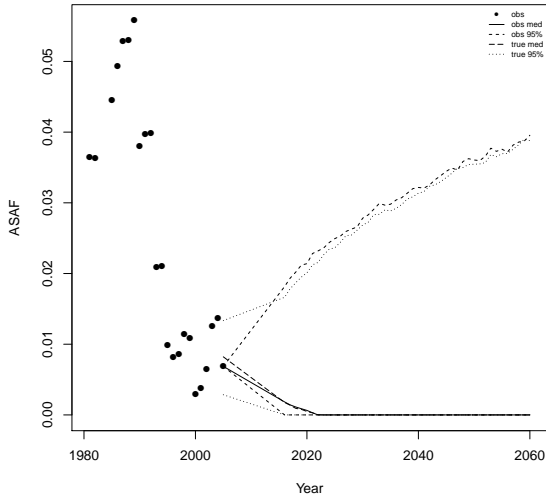




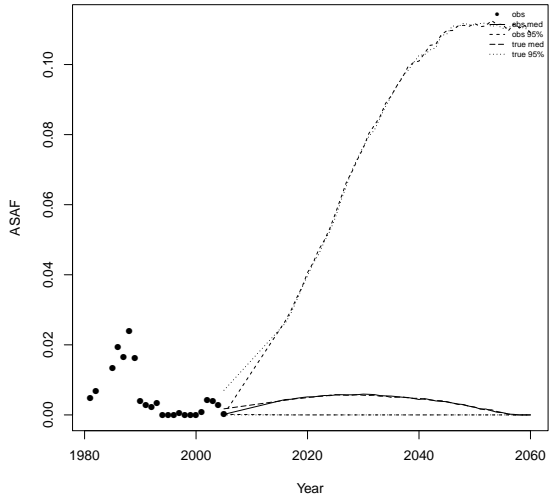




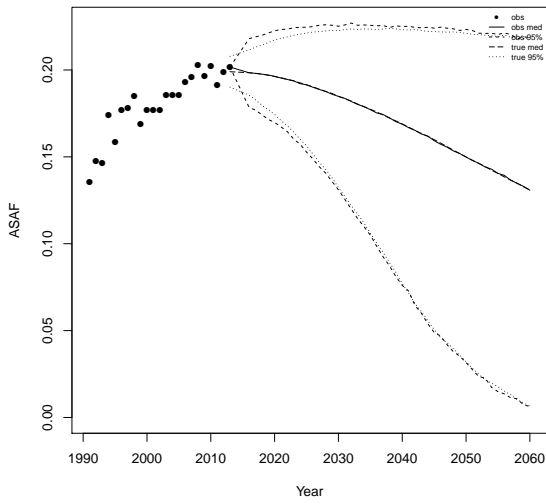
Tajikistan male



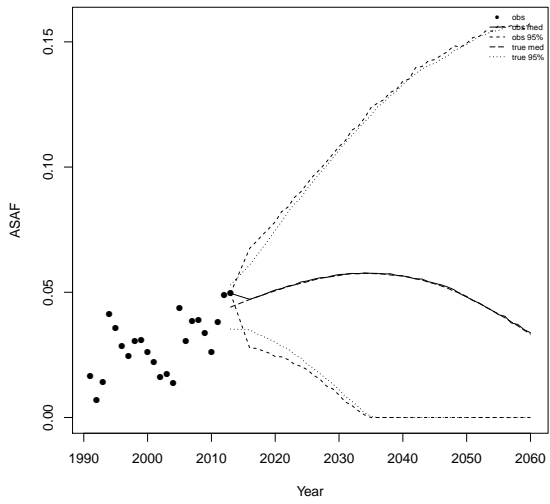
Tajikistan female



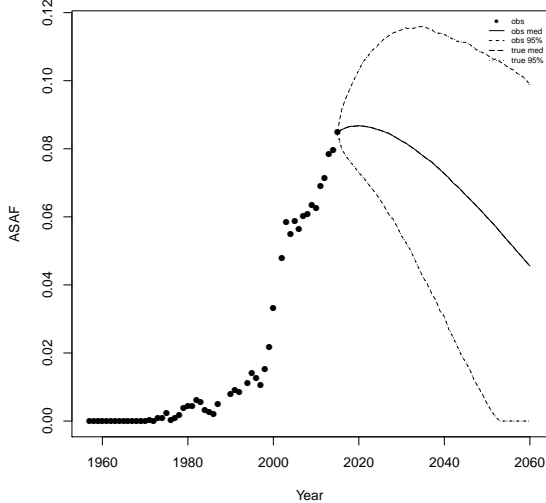
TFYR Macedonia male



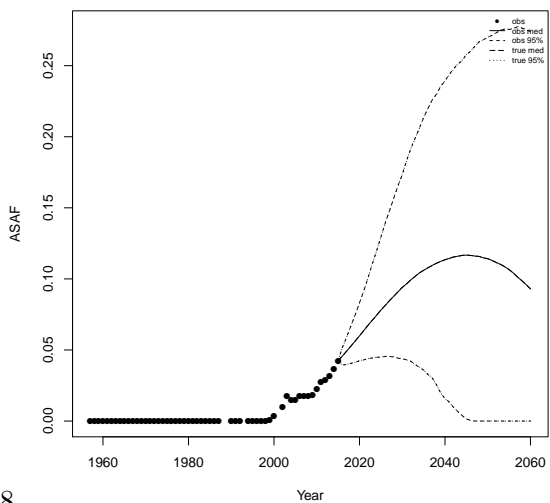
TFYR Macedonia female

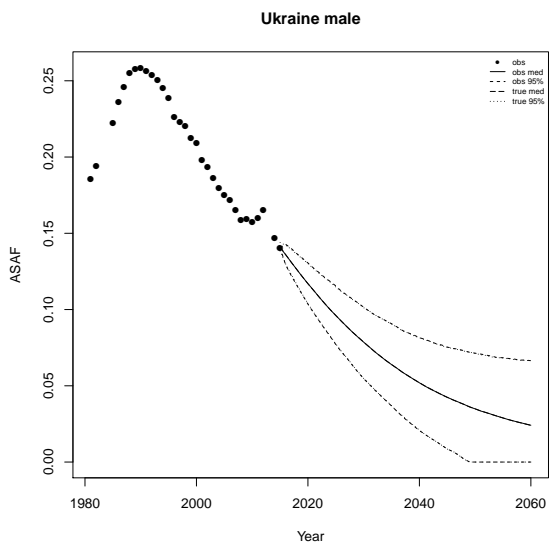
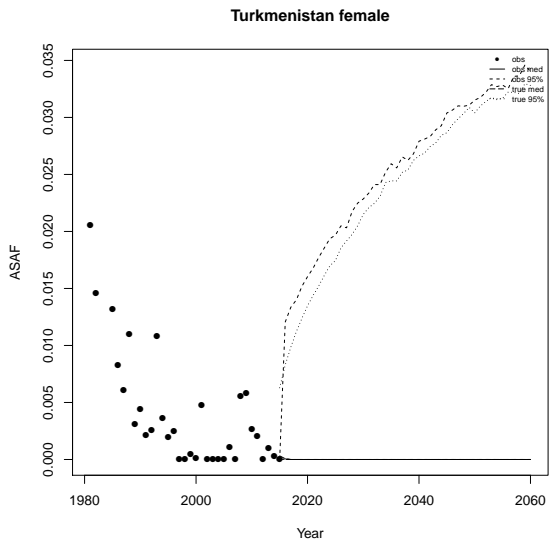
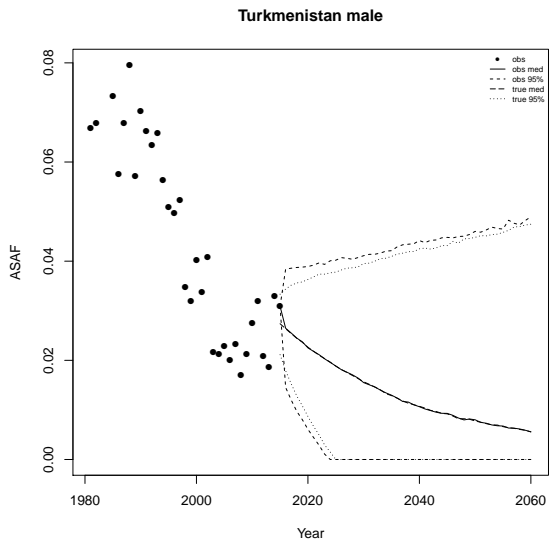
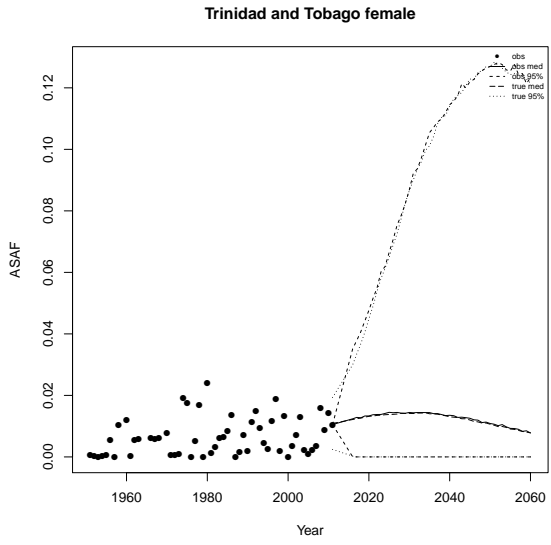
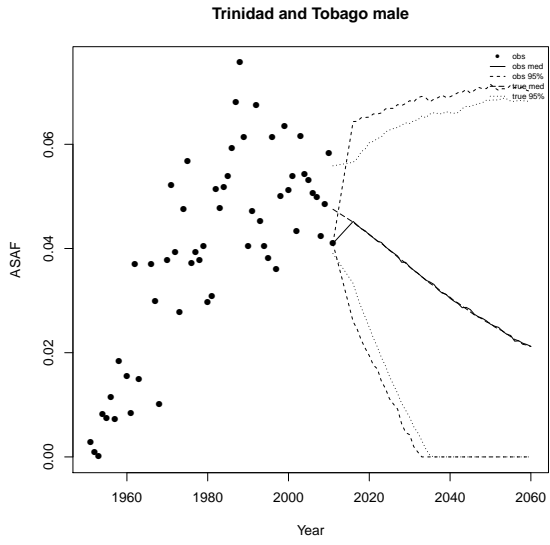


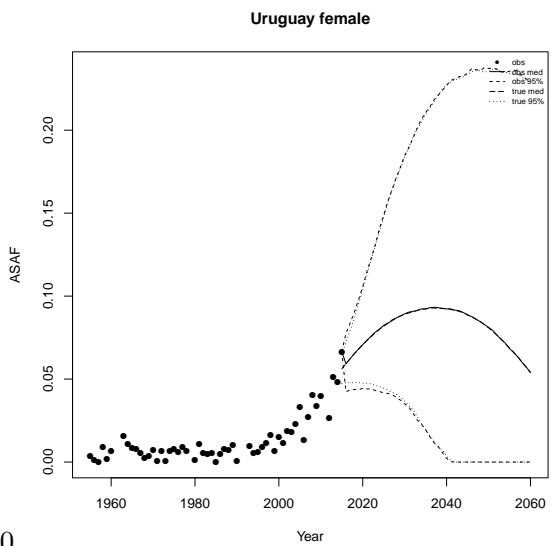
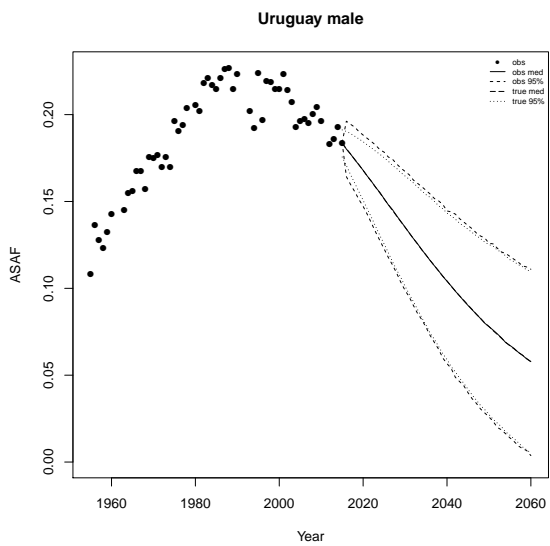
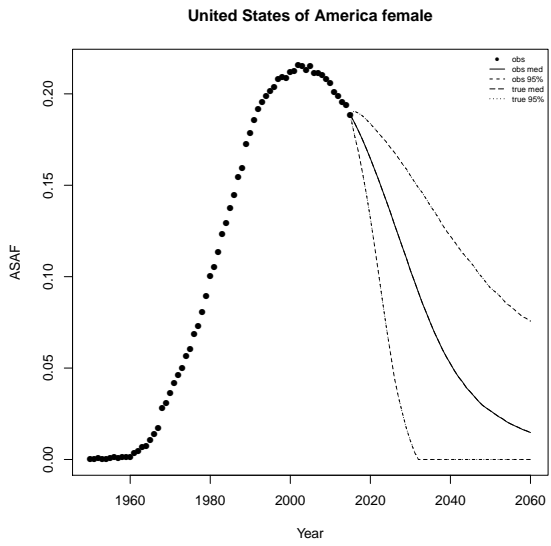
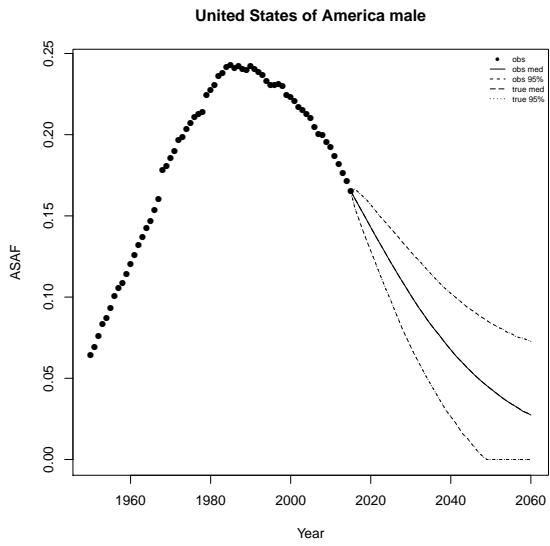
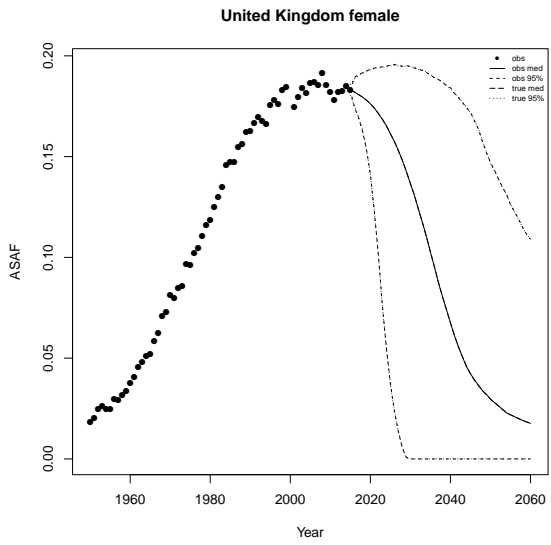
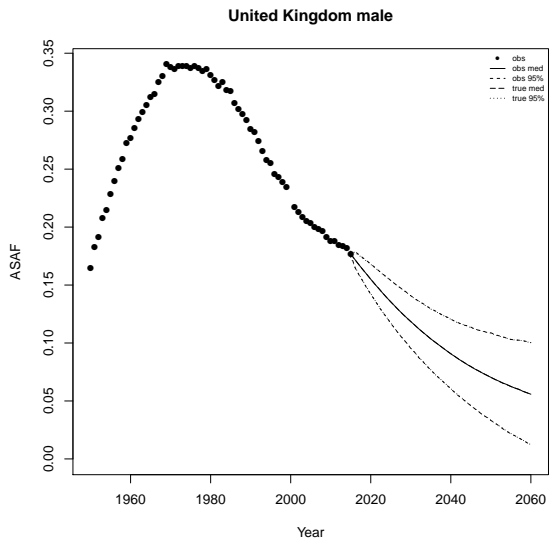
Thailand male

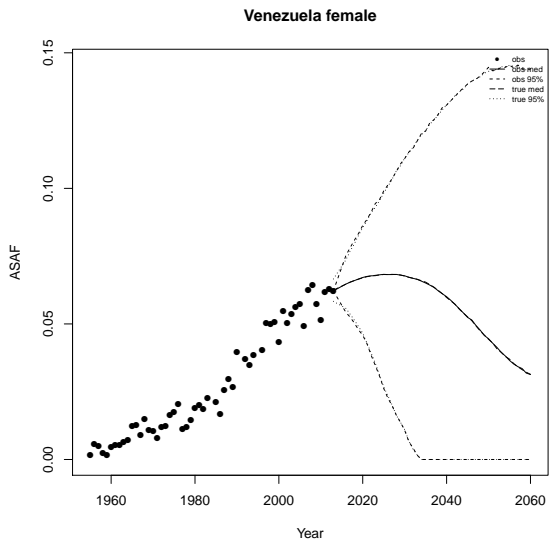
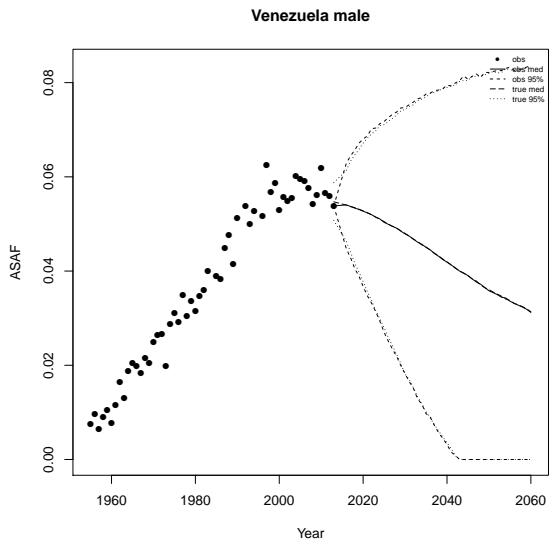
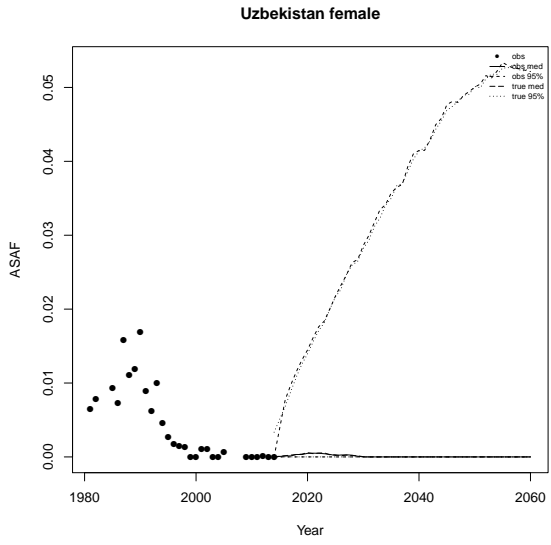
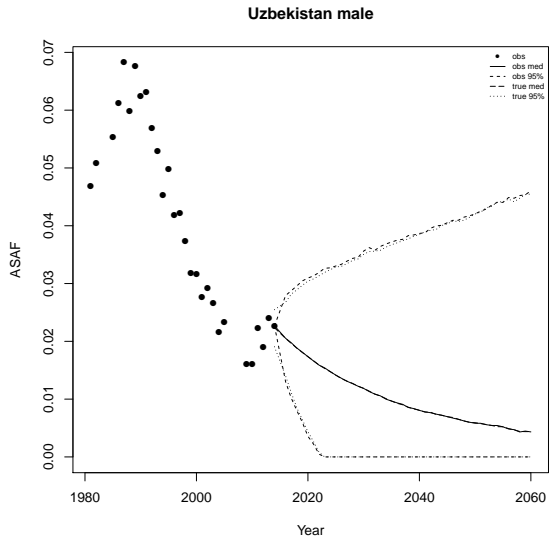


Thailand female









References

- Basu, S., Jammalamadaka, S. R., and Liu, W. (1996). Local posterior robustness with parametric priors: maximum and average sensitivity. In *Maximum Entropy and Bayesian Methods*, pages 97–106. Springer.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.
- Giordano, R. (2019). *rstansensitivity: Tools for calculating hyperparameter sensitivity in Stan*. R package version 0.0.0.9000.
- Giordano, R., Broderick, T., and Jordan, M. I. (2018). Covariances, robustness and variational bayes. *The Journal of Machine Learning Research*, 19(1):1981–2029.
- Gustafson, P. (1996). Local sensitivity of posterior expectations. *The Annals of Statistics*, 24(1):174–195.
- Raftery, A. E. and Lewis, S. M. (1992). One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, 7(4):493–497.