S1 Appendix. Algorithms of calibration methods.

1 Algorithm for Rejection ABC

A description of the Rejection ABC algorithm is found below, as described in (1).

1. Randomly sample parameter values θ^* independently from the prior distribution $\pi(\theta)$.
2. Simulate data set y from the model, using sampled parameter θ^* and obtain a summary statistic(s) s.
3. If $d(t,s) \leq \epsilon$, retain θ^*, otherwise reject, where d expresses the Euclidean distance between target t and summary statistics s, with each statistic being normalized with a division by the standard deviation of the simulated set of the statistic, and ϵ represents a tolerance level.
4. Return to 1 and repeat until desired posterior size is attained. Where posterior size is the number of parameter combinations retained.

2 Algorithm for Bayesian Calibration followed by Sampling Importance Resampling

A description of the BcSIR algorithm as described in (2).

1. **Sampling Stage:** Draw initial random sample of N inputs $(\theta_i, i = 1, 2, ..., N)$ from the prior distribution $\pi(\theta)$.
2. **Importance Stage:**
   a) For each θ_i, run model, obtain output and calculate the likelihood L_i
   b) Compute importance weights:

$$w_i = \frac{L_i}{\sum_{j=1}^{N} L_j}. \tag{1}$$

3. **Resampling Stage:** Resample desired posterior size from the initial N parameter inputs with replacement, using the computed weights as sampling weights.

3 Algorithm for AbcSmc

A description of the AbcSmc algorithm is found below, as described in (3), where N is the number of parameter combinations sampled, ρ is the fraction retained in the posterior, and $n = \lfloor \rho * N \rfloor$ is the size of the posterior:

1. Set the SMC iterator $l = 1$.
    a) For sample iterator $i = 1, \ldots, N$:
        i. Sample parameter vector $\theta_i^{(1)}$ independently from prior distributions $\pi(\theta)$
        ii. Simulate data $x_i^{(1)} \sim p(x_i^{(1)} | \theta_i^{(1)})$.
2. Calculate Partial least squares regression (PLS) model for $\theta^{(l)}$ and $x^{(l)}$
3. Transform observed data and each $x_i^{(l)}$ to new independent, orthogonal metrics using PLS model
4. Calculate Euclidean distance between transformed observed and transformed simulated metrics
5. Set $\Theta^{(l)} =$ best $\theta^{(l)}$, the best n samples, ranked by smallest distance.
6. The user determines the convergence of the algorithm - when a desired posterior is attained. If $l > 1$ and Θ has converged, stop here. Otherwise:
    a) Set τ_{l+1}^2 equal to twice the variance of $\Theta^{(l)}$
    b) For $j = 1, ..., n$:
        i. Set weight $\omega_j^{(l)} \propto \begin{cases} 1/n & : l = 1 \\ \pi(\Theta_j^{(l)}) / \sum_{k=1}^{n} \omega_j^{(l-1)} K(\Theta_j^{(l)} | \Theta_k^{(l-1)}; \tau_l^2) & : l > 1 \end{cases}$
        where $K(\Theta_j^{(l)} | \Theta_k^{(l-1)}; \tau_l^2)$ is a Gaussian perturbation kernel with mean $\Theta_k^{(l-1)}$ and variance τ_l^2, evaluated at $\Theta_j^{(l)}$.
    c) Normalize $\omega_i^{(l)}$ to sum to 1.
    d) Set $l = l + 1$
        i. For $i = 1, \ldots, N$:
            A. Choose θ_i^* from $\Theta^{(l-1)}$ with probabilities $\omega^{(l-1)}$.
            B. Sample $\theta_i^{(l)}$ from $Gaussian(\theta_i^*, \tau_l^2)$
            C. Simulate data $x_i^{(l)} \sim p(x_i^{(l)} | \theta_i^{(l)})$.
    e) Go to step 2.

4 Algorithm for Seq ABC

A description of the Seq ABC algorithm, as described in (4), where N is the number of parameter values sampled, $N\alpha = \lfloor \alpha N \rfloor$ is the number of parameter values to retain at each step among the N parameter values ($\alpha \in [0,1]$) and $Pacc_{min}$, a positive number between 0 and 1, is the stopping criterion of the algorithm.

1. **for** l=1 **do**

   a) **for** $i = 1$ to N **do**

      i. Sample parameter value $\theta_i^{(0)}$ from the prior distribution $\pi(\theta)$ and simulate data set y from the model such that $y \sim f(y|\theta_i^{(0)})$.

      ii. Set $d_i^{(0)} = d(x,y)$, that is, distance between simulated data y and observed data x

      iii. Set the weight, $w_i^{(0)} = 1$

   b) **end for**

   c) Let $\epsilon_1 = Q_{p^{(0)}}(\alpha)$, the first α-quantile of the set of distance $d^{(0)}$, where $d^{(0)} = \left\{d_i^{(0)}\right\}_{1 \le i \le N}$

   d) Let $\left\{(\theta_i^{(1)}, w_i^{(1)}, d_i^{(1)})\right\} = \left\{(\theta_i^{(0)}, w_i^{(0)}, d_i^{(0)}) | d_i^{(0)} < \epsilon_1, 1 \le i \le N\right\}$

   e) Take σ_1^2 as twice the weighted empirical variance of $\{(\theta_i^{(1)}, w_i^{(1)})\}_{1 \le i \le N\alpha}$

   f) Set $Pacc = 1$, where $Pacc$ is the stopping criterion for the current step.

   g) $l \leftarrow l + 1$

2. **end for**

3. **While** $Pacc > Pacc_{min}$ **do**, where $Pacc_{min}$ is the stopping criterion for the algorithm

   a) **for** $i = N_\alpha + 1$ to N **do**

      i. Pick θ_i^* from $\theta_j^{(l-1)}$ with probability $\frac{w_j^{(l-1)}}{\sum_{k=1}^{N_\alpha} w_k^{(l-1)}}, 1 \le j \le N_\alpha$

ii. Generate $\theta_i^{(l-1)}|\theta_i^* \; N(\theta_i^*, \sigma^2_{(l-1)})$ and $x \; f(x|\theta_i^{(l-1)})$

iii. Set $d_i^{(l-1)} = d(S(x), S(y))$

iv. Set $w_i^{(l-1)} = \frac{\pi(\theta_i^{(l-1)})}{\sum_{j=1}^{N_\alpha}(w_j^{(l-1)} / \sum_{k=1}^{N_\alpha} w_k^{(l-1)})\sigma_{(l-1)}^{-1}\psi(\sigma_{(l-1)}^{-1}(\theta_i^{(l-1)} - \theta_j^{(l-1)}))}$

b) **end for**

c) Set $Pacc = \frac{1}{N - N_\alpha}\sum_{k=N\alpha+1}^{N} 1_{d_i^{(l-1)} < \epsilon_{l-1}}$

d) let $\epsilon_l = Q_{d^{(l-1)}}(\alpha)$ where $d^{(l-1)} = \left\{d_i^{(l-1)}\right\}_{1 \le i \le N}$

e) Let $\left\{(\theta_i^{(l)}, w_i^{(l)}, d_i^{(l)})\right\} = \left\{(\theta_i^{(l-1)}, w_i^{(l-1)}, d_i^{(l-1)}) | d_i^{(l-1)} \le \epsilon_l, 1 \le i \le N\right\}$

f) Take σ_l^2 as twice the weighted empirical variance of $\left\{(\theta_i^{(l)}, w_i^{(l)})\right\}_{1 \le i \le N_\alpha}$

g) $l \leftarrow l + 1$

4. **end while**

Where $\forall u \in [0,1]$ and $X = \left\{x_1, ..., x_n\right\}, Q_X(u) = inf\left\{x \in X | F_X(x) \ge u\right\}$ and $F_X(x) = \frac{1}{n}\sum_{k=1}^{n} 1_{x_k} \le x$.
Where $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

5 Algorithm for Incremental Mixture Importance Sampling

We used the BcIMIS algorithm as described in (5), where $N_0 = B * 10$ is the initial sample size, B is the incremental sample size at each iteration of the algorithm and k is the number of iterations.

1. **Initial stage**:

   a) Sample N_0 parameter values or inputs $\theta_1, \theta_2, ..., \theta_{N_0}$ from the prior distribution $p(\theta)$.

   b) For each θ_i, run model and calculate the likelihood L_i, and form the importance weights:

$$w_i^{(0)} = \frac{L_i}{\sum_{j=1}^{N_0} L_j}. \tag{2}$$

2. **Importance Sampling Stage:** For $k = 1, 2, ...,$, repeat the following steps:

   a) Choose the current maximum weight input as the center $\theta^{(k)}$. Estimate $\sum^{(k)}$ from the weighted covariance of the B inputs with the smallest Mahalanobis distances to $\theta^{(k)}$, where the distances are calculated with respect to the covariance of the importance weights and $\frac{1}{N_k}$.

   b) Sample B new inputs from a multivariate Gaussian distribution H_k with covariance matrix $\sum^{(k)}$.

   c) Calculate the likelihood of the new inputs and combine the new inputs with the previous ones. From the importance weights:

$$w_i^{(k)} = cL_i x \frac{p(\theta_i)}{q^{(k)}(\theta_i)}, \tag{3}$$

   Where c is chosen so that the weights add to 1, $q^{(k)}$ is the mixture sampling distribution $q^{(k)} = \frac{N_0}{N_k} p + \frac{B}{N_k} \sum_{s=1}^{k} H_s$, H_s is the $s-$th multivariate normal distribution, $N_0 = B * 10$ is the initial sample size and $N_k = N_0 + Bk$ is the total number of inputs up to iteration k.

3. **Resample Stage:**

   a) The algorithm stops when the expected fraction of unique parameter combinations in resample is at least $(1 - \frac{1}{e}) = 0.632$

   b) Once the stopping criterion is satisfied, resample J inputs with replacement from $\theta_1, ..., \theta_{N_k}$ with weights $w_1, ..., w_{N_K}$, where K is the number of iterations at the importance sampling stage.

List of references

[1] Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Molecular biology and evolution. 1999;16(12):1791–1798.

[2] Menzies NA, Soeteman DI, Pandya A, Kim JJ. Bayesian methods for calibrating health policy models: a tutorial. PharmacoEconomics. 2017;35(6):613–624.

[3] Hladish TJ, Pearson CA, Chao DL, Rojas DP, Recchia GL, Gómez-Dantés H, et al. Projected impact of dengue vaccination in Yucatán, Mexico. PLoS neglected tropical diseases. 2016;10(5).

[4] Lenormand M, Jabot F, Deffuant G. Adaptive approximate Bayesian computation for complex models. Computational Statistics. 2013;28(6):2777–2796.

[5] Raftery AE, Bao L. Estimating and projecting trends in HIV/AIDS generalized epidemics using incremental mixture importance sampling. Biometrics. 2010;66(4):1162–1173.