3 Simple estimation (Initialization)

3.1 Simulated Annealing

Consider a simple version of model $Y = A\sin(g(x)+b) - B\cos(2g(x)+2b)$ where we assume the growth-time process $g(x)$ is a piecewise linear function with a single breakpoint, to accommodate for a change in frequency:

$$g(x) = \begin{cases} a_1 x & x < x_0 \\ (a_1 - a_2)x_0 + a_2 x & x \geq x_0 \end{cases} \tag{7}$$

This is a non-stochastic model, where we have introduced 3 new parameters, namely breakpoint x_0 and frequencies a_1 and a_2 respectively. Let $\hat{y}_i$ be the fitted values in in this simple model, then we propose the following loss function:

$$\text{loss}(A, B, b, a_1, a_2, x_0) = \underbrace{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}_{\text{Prediction Error}} + \underbrace{\lambda\left((\frac{e_{min}}{e_N} - 1)\mathbf{1}(e_N < e_{min}) + (\frac{e_N}{e_{max}} - 1)\mathbf{1}(e_N > e_{max})\right)}_{\text{Penalization}}, \tag{8}$$

where the model estimate for the number of cycles is

$$e_N = \frac{1}{2\pi}g(x_N) = \frac{1}{2\pi}\int_0^{x_N} \xi_s \, ds \tag{9}$$

and the interval (e_{min}, e_{max}) puts a restriction on this estimate. Consequently, the penalization controls the number of cycles, thereby reducing the likelihood of fitting signals with a cycle count beyond this pres-specified interval and also reduces the susceptibility of the method to fit fluctuating noise. The hyperparameter λ is the standard penalization factor, which can for instance be calibrated using cross-validation or similar techniques.

Optimizing (8) requires robust optimization, and we propose Simulated Annealing, for example implemented in the R package GenSA [5]. From the optimization, we directly obtain initial estimates for the parameters A, B and b of the full model. Our guess at $g(x)$, ξ_x and σ can also be derived, whereas β, ω cannot and should be estimated in another way. This procedure is summarized in Algorithm 1.

Algorithm 1: Initialization using Simulated Annealing

Data: $\{(x_1, y_1), \ldots, (x_N, y_N)\}$

Result: $\hat{\xi}_x^0, \hat{a}^0, \hat{A}^0, \hat{B}^0, \hat{b}^0, \hat{\sigma}^0$

`Step(0)` Fit simple version of (**??**), (**??**) with $g(x)$ given by eq. (7). Use loss function (8) and Simulated Annealing. This provide estimates $\hat{A}^0, \hat{B}^0, \hat{b}^0$ and $\hat{g}(x), \hat{a}_1, \hat{a}_2$;

`Step(1)` Put $\hat{\xi}_x^0 = \hat{g}(x) - \hat{g}(x - \Delta x)$ with $\Delta x = x_i - x_{i-1}$;

`Step(2)` Put $\hat{a}^0 = \frac{1}{N}\sum_{i=1}^{N}\hat{\xi}_{x_i}^0$;

`Step(3)` Put $\hat{\sigma}^0 = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N} r_i^2}$ where r_i are the (simple) model residuals;

`Step(4)` Initialize β^0, ω^0, for example by drawing from distributions in equation (8), (9) and (11) and under the Feller condition (**??**)
