

Fixing parameter shift times in a birth-death skyline analysis

Alexandra Gavryushkina

In a birth-death skyline model [2] there are n intervals defined by $n + 1$ times: the time of the start of the process, $n - 1$ shift times, and the present time (or the time of the most recent sample). To fix these intervals in an analysis one can specify an n -dimensional vector of change times in the `-RateChangeTimes` argument for each parameter. There are two ways to specify the shift times using the `-RateChangeTimes` arguments:

- In forward time, meaning that the change times are given as a distance from the start of the process which can be the root (if the origin time is not specified) or the time of origin. In this case, the first change time should be 0.0. For example, if we know that there were two shifts in a parameter value: one shift happened 3 years after the start of the process and another 5 years after the start of the process then the correct vector of times is: “0.0 3.0 5.0” and the respective value of the `reverseTimeArrays` should be set to “false.” Note, in this case, the remaining time will be defined by the start of the process: either the root height or the origin time, which are always given as a distance from the present (the most recent sample). That means that the calendar times of the shifts will change whenever the origin or the root time is changed in MCMC. Thus, if the most recent sample is in 2009 and the time of origin is 8 then the three intervals in calendar times will be [2001, 2004], [2004, 2006], and [2006, 2009]. If during MCMC the origin time changes to, say, 6 then the intervals will be [2003, 2006], [2006, 2008], and [2008, 2009].
- In backward time, meaning that the change times are given as a distance from the present (the most recent sample). In this case, the last change time should be 0.0. For example, if the first shift happened 8 years ago and the second shift happened 7 years ago, the correct vector of times is: “8.0 7.0 0.0” and the respective value of the `reverseTimeArrays` should be set to “true.” Again the remaining time is defined by the time of origin or the root. That is, if the most recent sample is in 2009 and the time of origin is 9.5 then the three

intervals will be [1999.5,2001], [2001,2002], and [2002,2009]. In this case, the calendar times of the shifts will not change with the time of origin (or the root) but only the start of the first interval. That is, if the time of origin changes to 11 then the three intervals will be [1998, 2001],[2001,2002], and [2002,2009].

The order of the time shifts in the `-RateChangeTimes` argument is not really important. But it is important that one of the shifts is always 0.0. One could avoid specifying 0.0 every time as it is constant but the developers chose to keep the number of shift times the same as the number of intervals.

Also note that if one uses the alternative parameterisation with R , δ , and s then `birthRateChangeTimes` will define the change times for `reproductiveNumber`, `deathRateChangeTimes` for `becomeUninfectiousRate` and `samplingRateChangeTimes` for `samplingProportion`.

If all of the parameters (λ , μ , ψ , ρ , and r for SA model [1]) change at the same times then instead of specifying change times for each of the parameters one can specify `IntervalTimes` argument. The rules for forward/backward time are the same as for `-RateChangeTimes` arguments. Make sure you change the `reverseTimeArrays` to “true” for all parameters if you specify `IntervalTimes` in backward times. If in addition to `IntervalTimes`, `-RateChangeTimes` is specified for some parameters (for example, `deathRateChangeTimes` is also specified) then the change times specified by `-RateChangeTimes` will be used for these parameters (that is, all parameters but `deathRate` will change at the times specified by `IntervalTimes` and `deathRate` will change at the times specified by `deathRateChangeTimes`). You may specify `IntervalTimes` in forward times and `-RateChangeTimes` in backwards times and vice versa. Make sure that the `reverseTimeArrays` values for each parameter are set to “true” or “false” depending on the type of shift times specified for each parameter.

References

- [1] A. Gavryushkina, D. Welch, T. Stadler, and A. J. Drummond. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Computational Biology*, 10(12):e1003919, 2014.

- [2] T. Stadler, D. Künert, S. Bonhoeffer, and A. J. Drummond. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc Natl Acad Sci USA*, 110:228–33, 2013.