

Supplementary Material for: `iucn_sim` -  
Improved predictions of future extinctions using  
IUCN status assessments

Tobias Andermann<sup>1,2</sup>, Søren Faurby<sup>1,2</sup>, Robert Cooke<sup>1,2</sup>,  
Daniele Silvestro<sup>1,2,\*</sup> and Alexandre Antonelli<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Biological and Environmental Sciences, University of  
Gothenburg, Göteborg, Sweden

<sup>2</sup>Gothenburg Global Biodiversity Centre, Göteborg, Sweden

<sup>3</sup>Royal Botanic Gardens, Kew, Richmond, UK

\* Joint last authorship

December 16, 2019

```
# Download and install iucn_sim (requires conda package manager) 1
conda config --add channels conda-forge 2
conda config --add channels bioconda 3
conda install iucn_sim 4
# Estimate status transition rates 5
iucn_sim get_rates \ 6
  --input_data data/example_data/gl_data_all_birds.txt \ 7
  --reference_group Aves \ 8
  --reference_rank class \ 9
  --outdir data/example_data/birds_output/transition_rates/ 10
# Run future simulations and estimate extinction rates 11
iucn_sim run_sim \ 12
  --indir data/example_data/birds_output/transition_rates/ \ 13
  --outdir data/example_data/birds_output/future_simulations/ \ 14
  --n_years 100 \ 15
  --n_sim 10000 16
```

Code sample 1: Installation and usage of `iucn_sim`, using a bash command line.

```

# Print the get_rates help function
iucn_sim get_rates -h
--input_data INPUT_DATA
    Path to generation length (GL) data: first column taxon
    list, followed by n columns of GL values.
--reference_group REFERENCE_GROUP
    Name of taxonomic group (or list of groups) to be used for
    calculating status transition rates (e.g. 'Mammalia' or
    'Rodentia,Chiroptera'). Alternatively provide path to text
    file containing a list of species names, compatible with
    IUCN taxonomy (>1000 species recommended). If none
    provided, the input species list with GL data will be used
    for calculating transition rates. Tip: Use precompiled
    group for significantly faster processing (see available
    groups at github.com/tobiashofmann88/
    iucn_extinction_simulator/data/precompiled/iucn_history/)
--reference_rank REFERENCE_RANK
    Provide the taxonomic rank of the provided reference
    group(s). E.g. in case of 'Mammalia', provide 'class' for
    this flag, in case of 'Rodentia,Chiroptera' provide
    'order,order'. Has to be at least 'Family' or above. This
    flag is not needed if species list is provided as
    reference_group or if reference group is already
    precompiled.
--n_rep N_REP
    How many different transition-rate and extinction risk
    estimates to produce for simulations (default == number of
    provided GL value columns per species under 'input_data'
    flag).
--iucn_key IUCN_KEY
    Provide your IUCN API key
    (see https://apiv3.iucnredlist.org/api/v3/token) for
    downloading IUCN history of your provided reference group.
    Not required if using precompiled reference group.
--outdir OUTDIR
    Provide path to outdir where results will be saved.
--status_list STATUS_LIST
    Provide a text file containing a valid IUCN status
    (LC,NT,VU,EN,CR,DD) for each species, separated by newline
    (same order as species names provided under --input_data).
--allow_precompiled_iucn_data ALLOW_PRECOMPILED_IUCN_DATA
    Set this flag to 0 if you want to avoid using precompiled
    IUCN history data. By default (1) this data is used if
    available for your specified
    reference organism group.
--n_gen N_GEN
    Number of generations for MCMC for transition rate
    estimation (default=100000).
--burnin BURNIN
    Burn-in for MCMC for transition rate estimation
    (default=1000).

```

Code sample 2: The get\_rates help-function.

```

1 # Print the run_sim help function
2 iucn_sim run_sim -h
3 --indir INDIR Path to directory created by get_rates function.
4 --outdir OUTDIR Provide path to outdir where results will be saved.
5 --n_years N_YEARS How many years to simulate into the future.
6 --n_sim N_SIM How many simulation replicates to run. By default
7 (value 0) as many simulation replicates are being produced
8 as there are available rate estimates, resulting from the
9 get_rates function (set by --n_rep flag in get_rates). If
10 the number of simulation replicates exceeds the number
11 of available transition rate estimates, these rates will
12 be randomly resampled for the remaining simulations.
13 --status_change STATUS_CHANGE
14 Model IUCN status changes in future simulations. 0=off,
15 1=on (default=1).
16 --conservation_increase_factor CONSERVATION_INCREASE_FACTOR
17 The transition rates leading to improvements in IUCN
18 conservation status are multiplied by this factor.
19 --threat_increase_factor THREAT_INCREASE_FACTOR
20 Opposite of conservation_increase_factor, multiplies the
21 transition rates leading to worsening in IUCN conservation
22 status.
23 --model_unknown_as_lc MODEL_UNKNOWN_AS_LC
24 Model new status for all DD and NE species as LC (best
25 case scenario). 0=off, 1=on (default=0).
26 --extinction_rates EXTINCTION_RATES
27 Estimation of extinction rates from simulation results:
28 0=off, 1=on (default=1).
29 --n_gen N_GEN Number of generations for MCMC for extinction rate
30 estimation (default=100000).
31 --burnin BURNIN Burn-in for MCMC for extinction rate estimation
32 (default=1000).
33 --plot_diversity_trajectory PLOT_DIVERSITY_TRAJECTORY
34 Plots the simulated diversity trajectory: 0=off, 1=on
35 (default=1).
36 --plot_histograms PLOT_HISTOGRAMS
37 Plots histograms of simulated extinction times for each
38 species: 0=off, 1=on (default=0).
39 --plot_posterior PLOT_POSTERIOR
40 Plots histograms of posterior rate estimates for each
41 species: 0=off, 1=on (default=0).

```

Code sample 3: The run\_sim help-function.

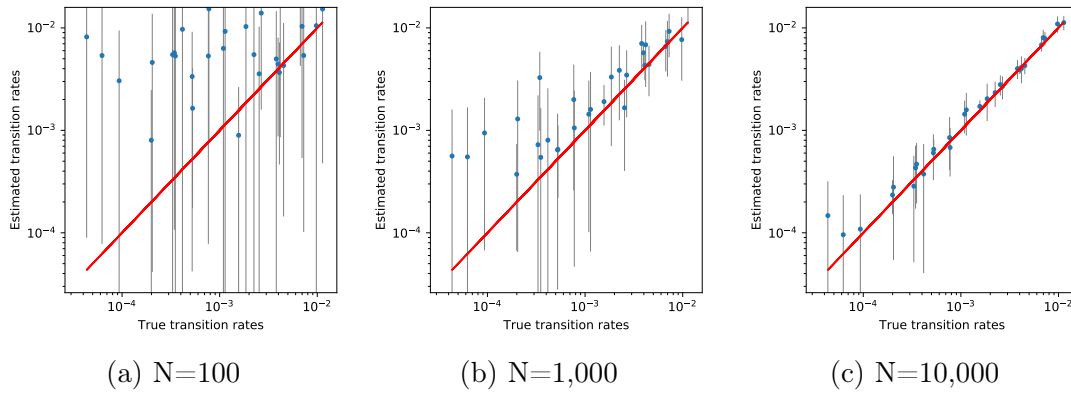


Figure 1: The estimates of status transition rates are improving with increasing size of the chosen reference group. We drew a random rate for each of the 30 possible changes between the six valid IUCN statuses LC, NT, VU, EN, CR, and DD, and used these rates to simulate transition counts between IUCN conservation statuses, mimicking the empirical IUCN history data. From these simulated transition count data we estimated transition rates using the `get_rates` function. The plots show the 30 transition rates used for simulation (x-axis) plotted against the estimated rates (y-axis). We simulated transition counts for reference groups of different sizes: (a) 100 species, (b) 1,000 species, and (c) 10,000 species. Mean values are represented as blue dots and the 95% credible interval of rate estimates are shown as grey vertical lines. The diagonal red line shows a theoretical perfect correlation for reference.

### Model test – Phylogenetic imputation for Aves

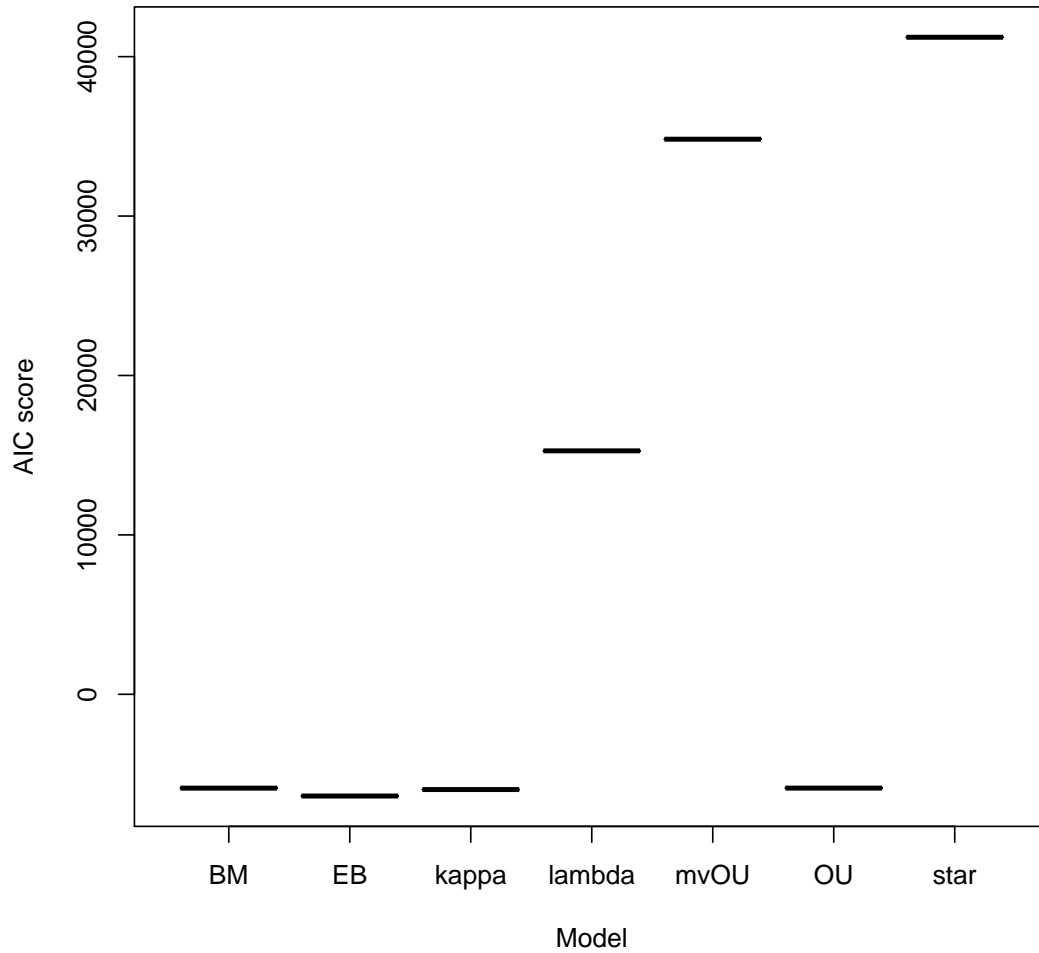


Figure 2: Results of model test in rphylopars. The plot shows the AIC scores of all tested models for the multivariate phylogenetic imputation of GL in birds. Based on these results we selected ‘EB’ as the best model.

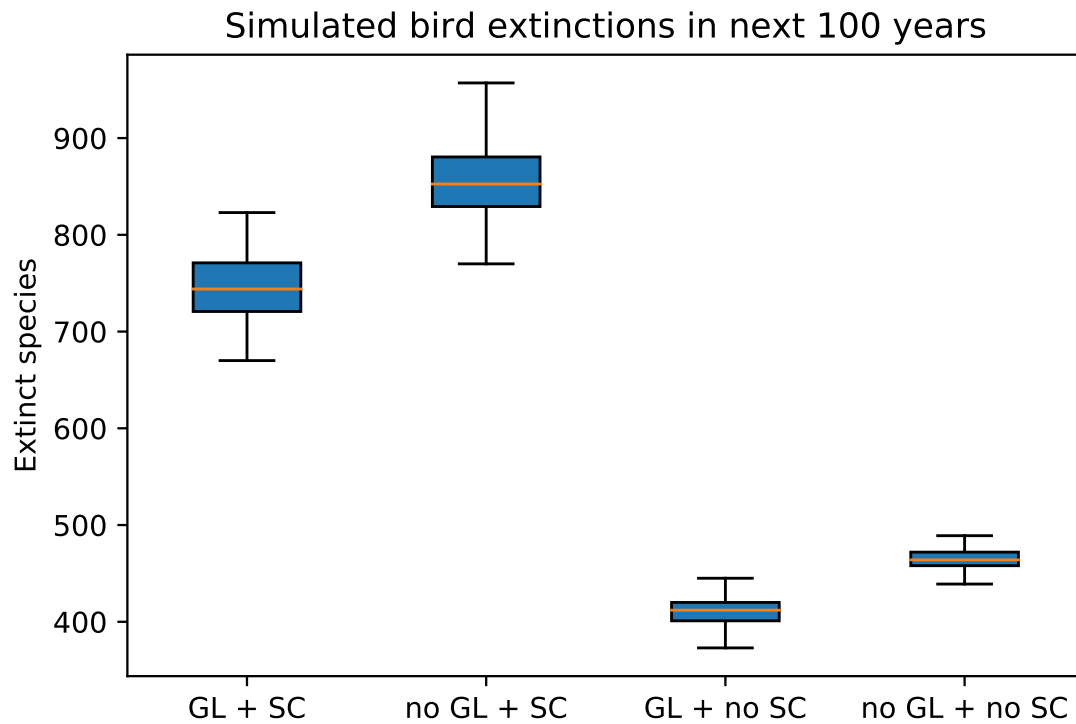


Figure 3: Number of predicted extinctions for birds in the next 100 years under different simulation scenarios across 100 simulation replicates. The blue boxes show the lower to upper quartile values of the predicted extinctions, with a orange line at the median. The whiskers show the full range of the predictions. Including generation length (GL) and conservation status changes (SC) into future simulations, leads to a significant increase in the number of predicted extinctions, compared to ignoring this information (compare first and last box-plot column). The individual effect of adding GL information to the simulations is a decrease of the predicted extinctions (3rd box-plot column), while only modeling SC leads to very high numbers of predicted extinctions (2nd box-plot column).