**SUPPORTING INFORMATION 1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Workflow | Category | What to report | Your method | Your rationale |
| (A) Obtaining and processing occurrence data | metadata | (A1) source of occurrence data | GBIF, speciesLink, MMA, literature, NeoXen database,  group experts |  |
| (A2) download date; version of data source | 2018, 2021 |  |
| (A3) basis of records | presence data |  |
| (A4) spatial extent | Point data |  |
| (A5) temporal range |  |  |
| processing | (A6-1) duplicate coordinates | Removed |  |
| (A6-2) spatial/environmental outlier; error | Removed |  |
| (A6-3) spatial/coordinate uncertainty | N/A |  |
| (A7-1) sampling bias | N/A |  |
| (A7-2) spatial autocorrelation | Points occurring in the same 50km2 cell were considered duplicates and removed |  |
| (B) Obtaining and processing environmental data | metadata & processing | (B1) source | ecoClimate |  |
| (B2) download date; version of data source | 2017 |  |
| (B3) spatial resolution | 50 km2 |  |
| (B4) temporal range | Modern (1950-1999) |  |
| (C) Model calibration | data input | (C1) modeling domain | Mainland Neotropics | Our species only occur mainland |
| (C2) number of background data | 50% the number of occurrences |  |
| (C3) sampling method for background data | Randomly | Chosen randomly outside the occurrence data |
| (C4) variable selection | Factor analysis | Choose the 5 variables that better fit the area |
| algorithm | (C5) name | BioClim, Maxent, SVM, GLM, Gower Distance, Random Forest |  |
| (C6) version of algorithm and software | RStudio 1.3.1094 , packages: “dismo”, “kernlab”, “randomForest” |  |
| (C7) parameterization |  |  |
| (D) Model transfer and evaluation | evaluation | (D1) evaluation index | Threshold-independent: models with AUC < 0.7 values were considered poor performing. |  |
| (D2) threshold for evaluation index | > 0.7 |  |
| (D3) dataset used to evaluate models | Random partition of occurrence and background data into testing and training |  |
| output | (D4) format/transformation | Binary presence and absence data |  |
| (D5) threshold | Prevalence: when modeled prevalence is closest to observed prevalence |  |
| extrapolation | (D6) novelty of projected environments compared with training environments |  |  |
| (D7) collinearity shift between training and projected environments |  |  |
| (D8) extrapolation strategy |  |  |
| metadata | (D9) source |  |  |
| (D10) download date; version of data source |  |  |
| (D11) spatial resolution |  |  |
| (D12) temporal range |  |  |