specieshindex: How scientifically popular is a species?

18 November 2020

Summary

Conservation efforts of species rely heavily on their existing research. Reintroduction projects, pest eradication, breeding programs, habitat restoration, etc. all require sound scientific knowledge for their proper implementation. Unfortunately, there is serious taxonomic bias (Bonnet, Shine, and Lourdais 2002) in ecological research, such that certain species receive more research interest (Donaldson et al. 2017; Santos et al. 2020; Seddon, Soorae, and Launay 2005; Titley, Snaddon, and Turner 2017; Troudet et al. 2017). Over time, it has widened the knowledge gaps of species and increased the patchiness of fundamental knowledge. This issue needs to be addressed by the scientific community as a whole to reduce the disparity of research and ensure its even distribution.

To remedy the problem of taxonomic bias, we must first quantify research interest in the species' publications. The challenge presented here is the lack of existing methods to quantify species-level research interest. Although this is a well-known problem, the few previous studies had only compared biases between higher taxonomic levels. Results showed that vertebrates in general attracted more research than invertebrates (Donaldson et al. 2017; Eisenhauer, Bonn, and Guerra 2019; Titley, Snaddon, and Turner 2017; Troudet et al. 2017). While vertebrate clades, i.e. Mammalia and Aves, had more related publications (Donaldson et al. 2017; Titley, Snaddon, and Turner 2017; Troudet et al. 2017). Species-specific investigations are vital in avoiding the over-generalisation of groups since variations in research interest may occur within clades.

Statement of need

The aim of specieshindex is to standardise the use of h-index in the context of measuring research popularity of species. The h-index was first introduced by Hirsch (Hirsch 2005) to compare the influence of academics (Hirsch and Buela-Casal 2014). It is obtained with the formula $h =$ total publications (n) that have at least been cited n times, after ranking the publications in a descending order by their number of citations. The h-index is now also being used to measure the research influence of the publications of different academic disciplines (Banks 2006; Harzing and Alakangas 2016), journals (Braun, Glänzel, and Schubert 2006), countries (Csajbók et al. 2007), species of animals (Fleming and Bateman 2016; McKenzie and Robertson 2015; Robertson and Mckenzie 2015) and pathogens (Cox et al. 2016). Using the h-index, specieshindex calculates the h-index of the publication of different species. There are currently 3 published studies (Fleming and Bateman 2016; McKenzie and Robertson 2015; Robertson and Mckenzie 2015) that adopted the species h-index as a measure of the species' research popularity. However, there are no standardised methods to achieve this at the moment.

specieshindex connects to the Scopus database and extracts citation records for analysis. It does it via the Scopus API, and returns information including the publication title, number of citations, publication type, etc. The binomial name of the species name must be used instead of their common names since they are less specific and can refer to larger groups of species. The 2 types of functions that connects to Scopus can be distinguished by their suffixes "T" and "TAK". "T" functions only extracts publications with the species' name in the title whereas "TAK" in the title, abstract and keywords. In the case that the user only wants to know the number of publications of a particular species, the `CountSpT()` and `CountSpTAK()` functions simply returns the total count without extracting any data. Apart from the h-index, specieshindex can also compute for other indices such as the *m*-index and the i10 index, which are also used to gauge research influence of authors. In this case, it will be used to gauge the research influence of species publication.

Implementation

The following packages are required for specieshindex to work.

```
# Installation from GitHub
install.packages("rscopus")
install.packages("taxize")
install.packages("XML")
install.packages("httr")
install.packages("dplyr")
install.packages("rlang")
devtools::install_github("jessicatytam/specieshindex", force = TRUE, build_vignettes = TRUE)

# Loading the libraries
library(rscopus)
library(taxize)
library(XML)
library(httr)
library(specieshindex)
```

Obtaining Scopus API key

An API key from Scopus is required to extract citation records from their database legally. Here are the steps to obtain the key.

1. Go to https://dev.elsevier.com/ and click on the button `I want an API key`.
2. Create an account and log in.
3. Go to the `My API Key` tab on top of the page and click `Create API Key`.
4. Read the legal documents and check the boxes.

Example

The species' binomial name is required to download the citation records from Scopus.

```
# Extract citation data
Woylie <- FetchSpTAK("Bettongia", "penicillata", myAPI)
Quokka <- FetchSpTAK("Setonix", "brachyurus", myAPI)
Platypus <- FetchSpTAK("Ornithorhynchus", "anatinus", myAPI)
Koala <- FetchSpTAK("Phascolarctos", "cinereus", myAPI)
```

These four datasets are readily available within the package. The API key is not required for the calculation of the indices. An efficient way to calculate all of the indices is to use the function `Allindices()`. The plot in Figure 1 illustrates one of the many ways to visualise the indices obtained below.

```
# Calculate indices
W <- Allindices(Woylie, genus = "Bettongia", species = "penicillata")
Q <- Allindices(Quokka, genus = "Setonix", species = "brachyurus")
P <- Allindices(Platypus, genus = "Ornithorhynchus", species = "anatinus")
K <- Allindices(Koala, genus = "Phascolarctos", species = "cinereus")

# Combine the citation records into a single dataframe
CombineSp <- rbind(W, Q, P, K)
CombineSp
```

```
##              genus_species     species          genus publications citations
## 1    Bettongia_penicillata penicillata       Bettongia          113      1903
## 2       Setonix_brachyurus  brachyurus         Setonix          242      3427
## 3 Ornithorhynchus_anatinus    anatinus Ornithorhynchus          321      6365
## 4   Phascolarctos_cinereus    cinereus   Phascolarctos          773     14291
##   journals articles reviews years_publishing  h     m i10 h5
## 1       55      110       3               43 26 0.605  54  7
## 2      107      237       5               66 29 0.439 121  4
## 3      153      308      13               67 41 0.612 177  7
## 4      227      744      29              139 53 0.381 427 14
```

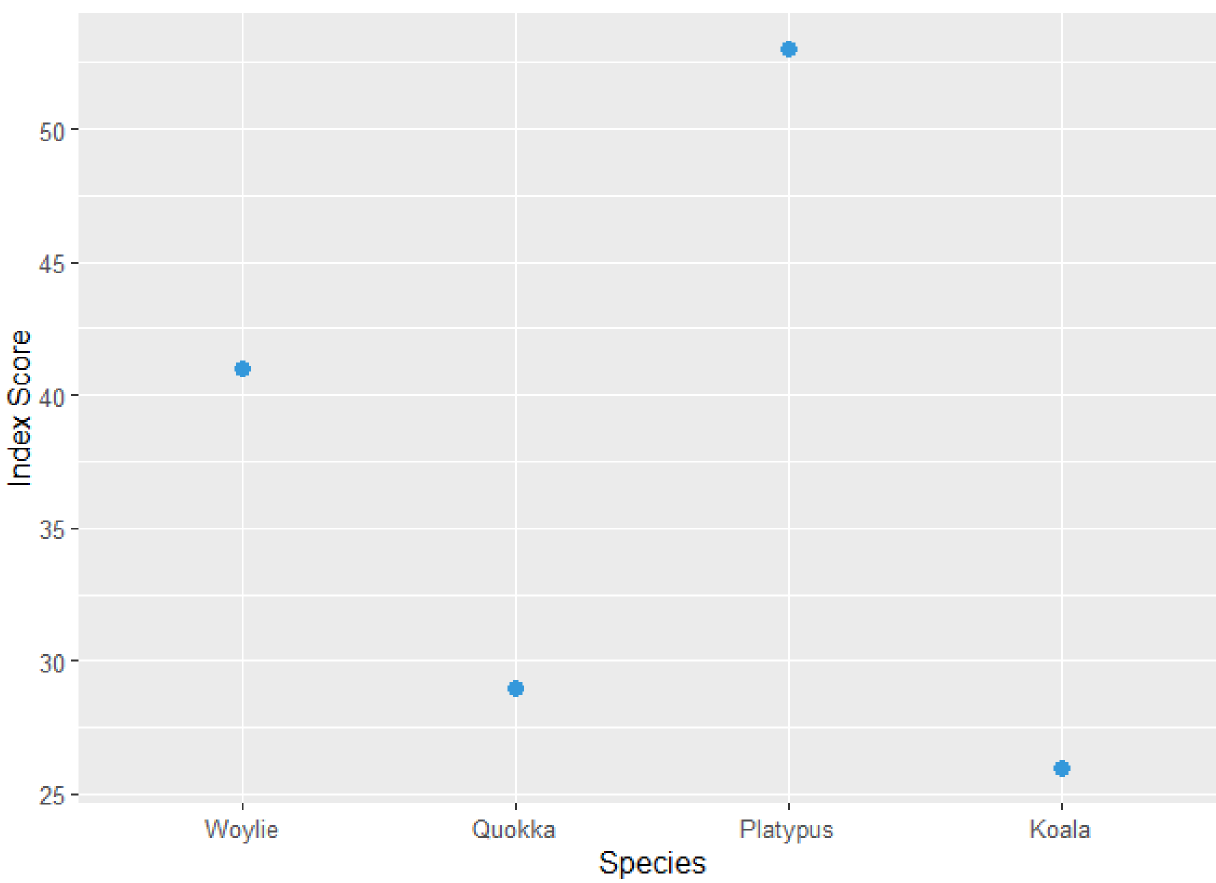

Figure 1. The *h*-index of the Woylie, Quokka, Platypus, and Koala.

Acknowledgements

I acknowledge the contributions from my supervisors Dr. Malgorzata (Losia) Lagisz, Professor Shinichi Nakagawa, and Associate Professor Will Cornwell.

References

Banks, Michael G. 2006. "An Extension of the Hirsch Index: Indexing Scientific Topics and Compounds." *Scientometrics* 69 (1): 161–68. http://isiknowledge.com.

Bonnet, Xavier, Richard Shine, and Olivier Lourdais. 2002. "Taxonomic Chauvinism." *Trends in Ecology & Evolution* 17 (1): 1–3. https://doi.org/10.1016/S0169-5347(01)02381-3.

Braun, Tibor, Wolfgang Glänzel, and András Schubert. 2006. "A Hirsch-Type Index for Journals." *Scientometrics* 69 (1): 169–73.

Cox, R., K. M. Mcintyre, J. Sanchez, C. Setzkorn, M. Baylis, and C. W. Revie. 2016. "Comparison of the H-Index Scores Among Pathogens Identified as Emerging Hazards in North America." *Transboundary and Emerging Diseases* 63 (1): 79–91. https://doi.org/10.1111/tbed.12221.

Csajbók, Edit, Anna Berhidi, Lívia Vasas, and András Schubert. 2007. "Hirsch-Index for Countries Based on Essential Science Indicators Data." *Scientometrics* 73 (1): 91–117. https://doi.org/10.1007/s11192-007-1859-9.

Donaldson, Michael R., Nicholas J. Burnett, Douglas C. Braun, Cory D. Suski, Scott G. Hinch, Steven J. Cooke, and Jeremy T. Kerr. 2017. "Taxonomic Bias and International Biodiversity Conservation Research." *FACETS* 1 (1): 105–13. https://doi.org/10.1139/facets-2016-0011.

Eisenhauer, Nico, Aletta Bonn, and Carlos A Guerra. 2019. "Recognizing the Quiet Extinction of Invertebrates." *Nature Communications* 10 (1): 1–3.

Fleming, Patricia A., and Philip W. Bateman. 2016. "The Good, the Bad, and the Ugly: Which Australian Terrestrial Mammal Species Attract Most Research?" *Mammal Review* 46 (4): 241–54. https://doi.org/10.1111/mam.12066.

Harzing, Anne Wil, and Satu Alakangas. 2016. "Google Scholar, Scopus and the Web of Science: A Longitudinal and Cross-Disciplinary Comparison." *Scientometrics* 106 (2): 787–804. https://doi.org/10.1007/s11192-015-1798-9.

Hirsch, J. E. 2005. "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of the National Academy of Sciences of the United States of America* 102 (46): 16569–72. https://doi.org/10.1073/pnas.0507655102.

Hirsch, Jorge E, and Gualberto Buela-Casal. 2014. "The Meaning of the H-Index." *International Journal of Clinical and Health Psychology* 14 (2): 161–64.

McKenzie, Ailsa J., and Peter A. Robertson. 2015. "Which Species Are We Researching and Why? A Case Study of the Ecology of British Breeding Birds." Edited by Antoni Margalida. *PLOS ONE* 10 (7): e0131004. https://doi.org/10.1371/journal.pone.0131004.

Robertson, Peter A., and Ailsa J. Mckenzie. 2015. "The Scientific Profiles of Terrestrial Mammals in Great Britain as Measured by Publication Metrics." *Mammal Review* 45 (2): 128–32. https://doi.org/10.1111/mam.12038.

Santos, Janisson W dos, Ricardo A Correia, Ana CM Malhado, JV Campos-Silva, D Teles, P Jepson, and RJ Ladle. 2020. "Drivers of Taxonomic Bias in Conservation Research: A Global Analysis of Terrestrial Mammals." *Animal Conservation*.

Seddon, Philip J., Pritpal S. Soorae, and Frédéric Launay. 2005. "Taxonomic Bias in Reintroduction Projects." *Animal Conservation* 8 (1): 51–58. https://doi.org/10.1017/S1367943004001799.

Titley, Mark A., Jake L. Snaddon, and Edgar C. Turner. 2017. "Scientific Research on Animal Biodiversity Is Systematically Biased Towards Vertebrates and Temperate Regions." Edited by Bernd Schierwater. *PLOS ONE* 12 (12): e0189577. https://doi.org/10.1371/journal.pone.0189577.

Troudet, Julien, Philippe Grandcolas, Amandine Blin, Régine Vignes-Lebbe, and Frédéric Legendre. 2017. "Taxonomic Bias in Biodiversity Data and Societal Preferences." *Scientific Reports* 7 (1): 1–14. https:

//doi.org/10.1038/s41598-017-09084-6.