**DATASET INFORMATION**

# TRUE ISLANDS

*General points*

The filenames given below match those that are stored in the data files within the repository. The numbers at the end of the filenames correspond to the dataset numbers in Table S1 in the Supplementary Information of the paper. The full references for the source papers can be found there also.

For all datasets, we have removed introduced species, but for some datasets we include alternative versions with them included. Given the large number of introduced species in certain datasets, and the provision of alternative versions with these included, we do not detail the number of introduced species removed from the true island datasets.

### Azeria(2004)\_noAliens – 26.

Twenty-six islands in the Dahlak Archipelago, Southern Red Sea, Eritrea.

We removed one island (Dahlak Kebir) from the dataset as the authors only sampled a very small part of it, whereas the others were fully surveyed. Focuses on resident, land bird breeding species, and migrants and vagrants excluded.

African Reed (Mangrove warbler)(*Acrocephalus\_* *baeticatu\_avicenniae*) taken to be African reed warbler (*Acrocephalus baeticatus*).

We removed Kori bustard also as we assumed it was a mistaken record.

### Baiser et al (2017) Cape Verde Birds\_current\_noAliens - 27

Birds on islands in Cape Verde.

This dataset was originally from Baiser et al. (2017). They “extracted presence/absence data from a database of (breeding) bird species on 152 oceanic islands compiled by Blackburn et al. (2004) and Cassey et al. (2007) from species lists, field guides, and literature.” Excluded are vagrants, birds that only overwinter, and migrants that don’t breed.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

We did not include I.Do.Rombo as it seems to be multiple islands.

The status of the kite species is very complex, with lots of conflicting literature. Essentially, some sources say there is just Red Kite (which is almost certainly extinct), but more recent sources say you have the endemic red kite subspecies (NW Islands only) and a breeding population of black kites (probably formerly widespread but only actually mentioned in the SE Islands). We consider both kites as extinct in the current dataset, but present on the various islands in the historic version.

Alternative versions including alien species (Baiser et al (2017) Cape Verde Birds\_current\_withAlien) and historic extinctions (Baiser et al (2017) Cape Verde Birds\_historic) were compiled.

Baiser, B., Valle, D., Zelazny, Z. & Burleigh, J.G. (2017) Non-random patterns of invasion and extinction reduce phylogenetic diversity in island bird assemblages. *Ecography*, 41, 361–374.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004) Avian extinction and mammalian introductions on oceanic islands. *Science*, 305, 1955–1958.

Cassey, P., Lockwood, J.L., Blackburn, T.M. & Olden, J.D. (2007) Spatial scale and evolutionary history determine the degree of taxonomic homogenization across island bird assemblages. *Diversity and Distributions*, 13, 458–466.

### Baiser et al (2017) Cook Islands Birds\_current\_noAliens - 28

Birds on islands in the Cook Islands.

This dataset was originally from Baiser et al. (2017). As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

Note that several of the “islands” here are atolls, and there is some discrepancy between sources in regard to the area of these islands.

Alternative versions including alien species (Baiser et al (2017) Cook Islands Birds\_current\_withAlien), historic extinctions (Baiser et al (2017) Cook Islands Birds\_historic) and prehistoric extinctions (Baiser et al (2017) Cook Islands Birds\_Prehistoric\_steadman) were compiled. A version of the current bird fauna excluding alien species and also excluding marine species (Baiser et al (2017) Cook Islands Birds\_Modern\_steadman) was also compiled. The latter two versions only included the eight main islands with fossil data in Steadman (2006).

### Baiser et al (2017) Galapagos Birds\_current – 29

Birds on islands in the Galapagos.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

### Baiser et al (2017) Hawaii Birds\_current\_noAliens - 30

Birds on islands in Hawaii.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

*Moho\_braccatus, Myadestes\_myadestinus, Moho\_bishopi*, and *Hemignathus\_lucidus* included in original current dataset but removed here as all three seem to be now definitely considered extinct.

Note that several species have been placed in new genera in recent years (e.g. *Hemignathus* species placed in *Akialoa*) but we have stuck with the BirdTree taxonomy. Several seabird species removed where they were only found to breed on offshore islets, not the islands themselves.

Alternative versions including alien species (Baiser et al (2017) Hawaii Birds\_current\_withAlien), historic extinctions (Baiser et al (2017) Hawaii Birds\_historic) and prehistoric extinctions (Baiser et al (2017) Hawaii Birds\_noMarine\_Prehistoric) were compiled. A version of the current bird fauna excluding alien species and also excluding marine species (Baiser et al (2017) Hawaii Birds\_noMarine\_Modern) was also compiled.

### Baiser et al (2017) Lesser Antilles Birds\_current\_noAliens - 31

Birds on islands in the Lesser Antilles.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

Guadeloupe considered to include both Grande-Terre and Basse-Terre.

San Andreas and Providencia were originally included as Lesser Antilles in Blackburn et al, but classified as "toward Colombia" in Baiser et al, so are not included here. Lots of species are either recent colonists (so in the current but not historic dataset) or are native to the archipelago but have been introduced to other islands by humans.

Alternative versions including alien species (Baiser et al (2017) Lesser Antilles Birds\_current\_withAlien) and historic extinctions (Baiser et al (2017) Lesser Antilles Birds\_historic) were compiled.

### Baiser et al (2017) Marianas Birds\_current\_noAliens - 32

Birds on islands in the Marianas.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

The status of Guam rail is complex. It is extinct on Guam, but has been introduced to Rota. However, it’s not considered entirely self-sustaining yet (i.e. it requires some management to survive). As such, we have not included it as introduced on that island.

Guam crow has very ambiguous recent data. The best data seem to suggest it went extinct on Guam, got reintroduced, but all the reintroduced birds died, so it’s extinct again (and treated as extinct in the current dataset). It’s still present on Rota.

Alternative versions including alien species (Baiser et al (2017) Marianas Birds\_current\_withAlien), historic extinctions (Baiser et al (2017) Marianas Birds\_historic) and prehistoric extinctions (Baiser et al (2017) Marianas Birds\_Prehistoric\_steadman) were compiled. A version of the current bird fauna excluding alien species and also excluding marine species (Baiser et al (2017) Marianas Birds\_Modern\_steadman) was also compiled. The latter two versions only included the five main islands with most fossil data in Steadman’s (2006) Table 8-4.

Steadman, D.W. (2006) *Extinction and biogeography of tropical Pacific birds*. University of Chicago Press, Chicago.

### Baiser et al (2017) Society Islands Birds\_current\_noAliens - 33

Birds on islands in the Society Islands.

As part of this paper, we have used a range of literature sources to update the datasets (see Appendix S1 in the supplementary information).

Species such as common myna are probably present on nearly all islands, as are things like reef egret, but given the lack of precise information we have erred on the side of caution and only included them on islands where we found definitive proof.

Alternative versions including alien species (Baiser et al (2017) Society Islands Birds\_current\_withAlien) and historic extinctions (Baiser et al (2017) Society Islands Birds\_historic) were compiled.

### Bengtson & Bloch (1983) Birds Breeding - 34

Birds on islands in the Faroe Islands.

The surveys observed 105 species (land birds, freshwater birds and larids) during the census, but they only present data (that we use) of the 41 regularly breeding land and freshwater species. They excluded 64 species include marine species, not regularly breeding species, non-breeding visitors, and rare visitors. We have taken definite breeding + and probably breeding (+) species. Hooded crow taken to be carrion crow.

We have sourced data from the appendix of the source paper.

### Bradstreet & McCracken (1978) noAliens - 35

Birds on the St. Lawrence Islands, Canada.

Data sourced from Hager (1997), originally from Bradstreet & McCracken (1978). Bird species were considered to breed on the islands if records indicated possible, probable or confirmed breeding evidence, or if the same species were seen repeatedly on an island during early summer months (i.e., nesting season) and if that island occurred within the species' breeding range.

Northern oriole taken to be Baltimore oriole as the former has been split; common gallinule taken to be common moorhen as the two were not fully split until 2011.

Bradstreet, M.S.W. & McCracken, J.D. (1978) Avifaunal survey St. Lawrence Islands National Park. Unpublished report for Parks Canada, c/o St. Lawrence Islands National Park Mailorytown Landing Ontario, Canada, p. 343.

Hager, H.A. (1997) Conservation of species richness: are all umbrella species of similar quality? Thesis. The University of Guelph, Ontario, Canada.

### Haila et al 1983\_birds - 36

Birds on islands in the Vargskär archipelago, Finland.

As one-visit censuses were used, it was difficult for the source paper authors to determine breeding status. Species list seemingly includes summer migrant breeders as well as residents. Not every island was sampled each year, so we took the records from 1979 as this was the only year all 44 islands were sampled. For the larger islands (> 16ha) that were sampled as transects, the authors list species recorded on the transects and additional species found during their sampling – we have included both.

Only possible to find 27 islands from the dataset in the Global Island Database, on which to calculate the isolation metrics.

### Nuddsetal(1996)FathomFiveIslandsBirds - 37

Birds on islands in the Fathom Five Islands.

Data taken from Hager (1997), originally from Nudds et al. (1996). Bird species were considered to breed on the islands if records indicated possible, probable or confirmed breeding evidence, or if the same species were seen repeatedly on an island during early summer months (i.e., nesting season) and if that island occurred within the species' breeding range.

From Hager’s Appendix B1: We have included possible (A), probable (B) and confirmed (c) breeding species. Common gallinule taken to be common moorhen as the two were not fully split until 2011.

Hager, H.A. (1997) Conservation of species richness: are all umbrella species of similar quality? Thesis. The University of Guelph, Ontario, Canada.

Nudds, T.D., Bogart, J.P., Britton, D., Hager, H.A., Middleton, A.L.A., Potter, D.N. et al. (1996) *Species-area relations of woody plants, vertebrates and selected invertebrates on islands in Georgian Bay: the role of national parks in evaluating null models of biodiversity. Final report of the Guelph Biological Inventory of the Georgian Bay Islands*. Report to Parks Canada, Fathom Five National Marine Park, Ontario, Canada.

### Nuddsetal(1996)GeorgianBayBirds\_noAliens - 38

Birds on islands in the Georgian Bay.

Data taken from Hager (1997), originally from Nudds et al. (1996). Bird species were considered to breed on the islands if records indicated possible, probable or confirmed breeding evidence, or if the same species were seen repeatedly on an island during early summer months (i.e., nesting season) and if that island occurred within the species' breeding range.

From Hager’s Appendix B2: We have included possible (A), probable (B) and confirmed (c) breeding species. Common gallinule taken to be common moorhen as the two were not fully split until 2011. There is an island in Appendix B2 (38) that is not listed in the island table list (p85) and this island has thus been excluded.

Only possible to find five islands from the dataset in the Global Island Database, on which to base isolation values.

Hager, H.A. (1997) Conservation of species richness: are all umbrella species of similar quality? Thesis. The University of Guelph, Ontario, Canada.

Nudds, T.D., Bogart, J.P., Britton, D., Hager, H.A., Middleton, A.L.A., Potter, D.N. et al. (1996) *Species-area relations of woody plants, vertebrates and selected invertebrates on islands in Georgian Bay: the role of national parks in evaluating null models of biodiversity. Final report of the Guelph Biological Inventory of the Georgian Bay Islands*. Report to Parks Canada, Fathom Five National Marine Park, Ontario, Canada.

### OConnell et al (2020) - Wakatobi - 39

This dataset summarizes all (breeding) bird species present on nine islands in the Wakatobi archipelago, South-east Sulawesi. The data are entirely derived from information provided in O’Connell *et al*. (2020). Details regarding data collection methods and determination of breeding and introduced species can be found in this paper. Taxonomy for all species in the dataset has been retrospectively matched to BirdTree.

Importantly, unlike other datasets which are derived from extensive fieldwork on the islands involved, the Wakatobi islands are poorly-explored. As such, the database is derived exclusively from fieldwork completed by the authors, plus a few historical records published by Hartert (1903). Fieldwork effort differed between islands; as such an appreciation of survey work completed on each island is important when interpreting the database. Table 1 summarizes the number of fieldwork days completed on each island.

Table 1 – *Number of days fieldwork completed in each of the nine islands in the Wakatobi archipelago dataset. Derived from supplementary information in O’Connell et al. (2020).*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Island name** | Wangi-wangi | Oroho | Kapota | Hoga | Kaledupa | Tomia | Lintea Selatan | Binongko | Runduma |
| **# days fieldwork** | 41 | 7 | 3 | 26 | 33 | 37 | 3 | 22 | 10 |

The number of days completed in some of the smaller islands is quite low, but species diversity on these small islets is also expected to be low. All introduced species (rock dove, Eastern spotted dove, and barred dove) have been removed. No historic or prehistoric extinct species are known from the islands. Three species in the dataset require special attention:

* Yellow-crested Cockatoo (*Cacatua sulphurea*) is a Critically Endangered species highly imperilled by the illegal pet trade. As such, this species has been removed from the published version of the dataset.
* Wangi-wangi White-eye (*Zosterops nov*.) is a unique, as-yet-undescribed species that cannot be placed on the Jetz et al (2012) phylogeny. As such, we replace it with another locally-occuring White-eye species (*Zosterops consobrinorum*) in this database as a proxy, to allow coarse comparative metrics to be applied in its stead. It should be noted, however, that although this rough proxy is appropriate for the analysis in this study, it is not assumed that *Z. consobrinorum* is a close evolutionary relative, and should not be treated as such.
* O'Connell et al. recognize both *Z. chloris* & *Z. flavissimus* (sister species) in the study area. *Z. flavissimus* is not recognized by Jetz (2012) so both species lumped together here.

Hartert, E. (1903). On the birds collected on the Tukang-Besi Islands and Buton, south-east of Celebes, by Mr. Heinrich Kuhn. *Novitates Zoologicae*, 10: 18–38.

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012). The global diversity of birds in space and time. *Nature*, 491, 444–448.

O’Connell, D.P.O., Kelly, D.J., Kelly, S.B.A., Analuddin, K., Karya, A., Marples, N.M. & Martin, T.E. (2020). An assessment of the avifauna of the Wakatobi Islands, South-east Sulawesi, Indonesia: species recorded and taxonomic considerations. *Raffles Bulletin of Zoology*, 68: 574–587.

### Power (1972) chanb - 40

Birds on islands off the coast of southern California and northwestern Baja California, Mexico.

Dataset used in the Nestedness Calculator set of presence-absence matrices (Atmar & Patterson, 1985). Breeding or summer resident species of land bird. Source paper author states these data relate to numerous years and likely overestimate the richness of any given breeding season. Source paper author has excluded large raptors and large carrion feeders.

Four ‘islands’ (Ancapa, Los Coronados, Todos Santos and San Benito) are actually groups of small very close together islets, and the area value for these is the summed value.

Atmar, W. & Patterson, B.D. (1995) The nestedness temperature calculator: a Visual Basic program. AICS Research Chicago.

### Si et al (2017) TIL - 41

Data for 36 islands in Thousand Island Lake, China.

Data presented in Appendix (Table S3) of Si et al (2017). Birds sampled in breeding season annually from 2010-2014. Only breeding species recorded. High flying species (passing over) excluded. Only terrestrial breeding birds included – diving birds, ducks, gulls, herons and shorebirds whose habitat extensive rely on water removed by source paper authors.

*Zosterops japonicus* kept but the systematics for this species complex are very messy and have been changed multiple times over the last decade – currently *Z. japonicus* is not in China.

*Garrulax cineraceus* and *Turdus merula* now split but splits not in Jetz. *Pomatorhinus\_erythrogenys* taken to be *Pomatorhinus\_swinhoei* and *Yuhina\_castaniceps* taken to be *Yuhina\_torqueola*.

### Simaiakis et al (2012) all islands- 42

Breeding land bird species on 77 Aegean islands. Data originally sourced from the literature, Greek ornithological databases, and field expedition reports over the last 20 years.

Eastern orphean warbler (*Sylvia crassirostris*) changed to Western orphean warbler (*S. hortensis*) as former until recently considered a subspecies of latter and is not in BirdTree.

*Oenanthe\_pleschanka* removed as we could find no evidence of breeding in the area.

### Simberloff & Martin (1991) Haida Gwaii - 43

Land (presumably breeding) bird species of the Haida Gwaii islands [Queen Charlotte Islands]. A version of this dataset is used in the Nestedness calculator (NC) set of presence-absence matrices (Atmar & Patterson, 1995), where they took the version presented in Simberloff & Martin (1991). We initially used the Simberloff & Martin version, but found an updated version with more information in Martin et al. (1995) and have thus mainly used this as the final analysed data, with *Sphyrapicus\_ruber* presences added in from Simberloff & Martin.

Data from Table 1 and the appendix of Martin et al. Following Simberloff & Martin, we focused on islands 1-40 (Laskeek Bay and Juan Perez Sound). Different islands were sampled in different years using point counts (but species recorded outside point counts during excursions were also included in the totals, following the single-visit sampling procedure).

Note that until 1967 the largest island was actually connected to another large island at low tide, until the channel was dredged.

*Certhia familiaris* taken as *Certhia americana* as these have been split and the split *is* in Jetz.

Martin, J.-L., Gaston, A.J. & Simon, H. (1995) The effect of island size and isolation on old growth forest habitat and bird diversity in Gwaii Haanas (Queen Charlotte Islands, Canada). *Oikos*, 72, 115–131.

Simberloff, D. & Martin, J.L. (1991) Nestedness of insular avifaunas: simple summary statistics masking complex species patterns. *Ornis Fennica*, 68, 178–192.

### Simberloff & Martin (1991) Maddabrd\_noAliens - 44

Breeding bird species of the Maddalena Archipelago. We only focused on the 8 islands in the Maddalena group proper; the rest are in the Tuscan archipelago / Lavezzi group, which is much further north and could have a different source pool. Budelli was not included as it had only possible breeders and no regular breeders (see below).

Edited version of dataset used in the Nestedness calculator (NC) set of presence-absence matrices, where they took the version presented in Simberloff & Martin (1991). Simberloff & Martin sourced the data from Thibault et al. (1990) [which is in French], which we were eventually able to obtain.

We then used the data present in Thibault et al (1990), focusing on regular breeders only (as done by the source paper authors in some of their analyses). This was to ensure the islands had relatively consistent data (some islands were sampled more frequently and consequently had higher numbers of irregular breeders). Species sampled were “the terrestrial avifauna that nest and feed on islands.”

Simberloff, D. & Martin, J.L. (1991) Nestedness of insular avifaunas: simple summary statistics masking complex species patterns. *Ornis Fennica*, 68, 178–192.

Thibault, J.C., Martin, J.L. & Guyot, I. (1990) Les oiseaux terrestres nicheurs des Iles Mineures des Bouches-de-Bonifacio: analyse du peuplement. *Alauda*, 58, 173-185.

### Sin et al (2022) – West Sumatra & Riau-Lingaa – 45 & 46

Data available for 172 islands in South East Aisa, but only 94 are used in Sin et al. (2022) (these are the islands with sufficient sampling effort). From these 94, we have created two datasets / archipelagos: West Sumatra (continental) and Riau Lingaa (mixed continental and deep sea islands). Note, we have only included islands in each of these groups that are present in the 94 islands with sufficient sampling effort.

Bird lists for each island were compiled by the source paper authors via a comprehensive assessment of published lists. Smaller islands were also surveyed manually between 2019 and 2020, with some islands in West Sumatra also supplemented through surveys taken between 2015 and 2020.

Pelagic / marine species have not been included in either dataset as breeding data were not available for most islands.

*West Sumatra islands*

15 islands west of Sumatra. A mix of ‘deep sea’ (not connected to mainland during LGM) and continental islands. Here, night survey effort was relatively consistent across islands so we have included nocturnal species. Waterbirds also included.

*Riau-Lingaa islands*

26 islands in the Riau-Lingaa group, all classified as continental islands. Here, night survey effort not consistent, particularly on the smaller islands, and thus nocturnal species have been excluded. Waterbirds included.

### Borgesetal\_azores birds\_noAliens - 47

Birds on islands in the Azores.

Dataset originally produced by P.A.V.B and subsequently updated by us using a range of published sources (see Appendix S1 in the Supplementary information).

Residents and summer migrant breeders. Vagrants excluded.

*Puffinus\_baroli* changed to *Puffinus\_assimilis* as this is a new species and not in BirdTree.

An alternative version including alien species (Borgesetal\_azores birds\_withAlien) was compiled.

### Borgesetal\_canary birds\_noAliens – 48

Birds on islands in the Canaries.

Dataset originally produced by P.A.V.B and subsequently updated using a range of published sources (see Appendix S1 in the Supplementary information), particularly Illera et al. (2016). Note that Illera et al. class a lot of subspecies as species which we (as we are following BirdTree) do not.

Residents and summer migrant breeders. Vagrants excluded.

Several species are natural recent colonists and so are in the current but not historic datasets.

Alternative versions including alien species (Borgesetal\_canary birds\_withAlien), historic extinctions (Borgesetal\_canary birds\_historic) and prehistoric extinctions (Borgesetal\_canary birds\_noMarine\_Prehistoric) were compiled. A version of the current bird fauna excluding alien species and also excluding marine species (Borgesetal\_canary birds\_noMarine\_Modern) was also compiled.

Baiser, B., Valle, D., Zelazny, Z. & Burleigh, J.G. (2017) Non-random patterns of invasion and extinction reduce phylogenetic diversity in island bird assemblages. *Ecography*, 41, 361–374.

### Kubota et al. Birds Ryukus\_noAliens - 49

Birds on islands in the Ryukyu Islands.

The Ryukyu Islands dataset was compiled by Y.K. and colleagues, with data sourced from published literature, checklists and observation databases. These data are from the subset of Ryukyu Islands (Okinawa and Kagoshima Prefecture) where there are bird checklists or sufficient observation data (68 islands). All but two islands are continental shelf (two are oceanic).

Covers breeding bird species – residents and summer breeding migrants. All vagrants, species only recorded on passage (i.e. during migration) and species only wintering on the islands were excluded. All sub-species grouped to species level.

*Parus\_minor* changed to *Parus\_major*. *Hierococcyx\_hyperythrus* lumped with *Cuculus\_fugax* (latter was split into three species but after BirdTree published). *Anas zonorhyncha* taken as *Anas\_poecilorhyncha* as former not in BirdTree. *Ninox japonica* taken as *Ninox scutulata* as former not in BirdTree. *Otus lempiji* taken to be *Otus semitorques*. *Sittiparus (Poecille) varius olivaceus* taken as *Sittiparus (Poecille) varius* as split into Iriomote tit not in BirdTree. *Acridotheres javanicus* (introduced) taken as *Acridotheres\_grandis* as former not in BirdTree.

An alternative version including alien species (Kubota et al. Birds Ryukus\_withAlien) was compiled.

### Martin (2022) NZ\_current\_noAliens - 50

This dataset seeks to identify all bird species present (and historically/prehistorically present) on 22 New Zealand Islands, using a dataset previously compiled by Diamond (1984) as a foundation.

The presence of extant breeding species (including introduced species) on each island were determined by consulting maps and text in four sources: the IUCN Redlist (2021), Billerman et al. (2022), Robertson & Heather (2015) and New Zealand Birds Online (2022). Introduced species were classed as those with self-sustaining populations, as defined by these sources. Note that, in the case of New Zealand, native species within part of the archipelago can be introduced species in other areas due to the numerous conservation translocation projects which have been completed here. Taxonomy for all extant native and introduced species followed Jetz et al. (2012).

Lists of historic extinct breeding species (those considered extinct >1500AD) and prehistoric extinct breeding species (those considered extinct <1500AD) were obtained from Sayol et al. (2021) and additional literature searches. Once all species occurring outside of the study area range (e.g. outlying islands) were excluded, the spatial distribution of historical extinctions was determined by the above named sources, while spatial distributions of prehistoric extinctions was determined exclusively from New Zealand Birds Online (2022), which was the only source that provided detailed information on these species. Taxonomy for extinct species largely followed Sayol et al., with some changes to match with Jetz et al. where appropriate.

Note, as above, the status of birds in New Zealand can be complex, and in some cases a species can have a general status of “Native” but be introduced in certain Islands. Additionally, extant species can be extinct on some islands, and extinct species sometimes experienced staggered extinctions, with the last extinctions being historical but with extinctions on some islands being prehistorical.

*Area values*

The areas given in Diamond seemed quite coarse, so we collated area values from a range of sources (published papers, NZ Government statistics, Department of Conservation statistics, Google earth area measurer; all sources in separate file in DARs folder).

It is worth nothing that for a few of the smaller islands, Diamond may have actually been referring to island groups (Aldermen), i.e. groups of small islands. Here, we have only focused on specific islands, taking the largest island in each group (for species lists and area values).

Alternative versions including alien species (Martin (2022) NZ\_current\_withAlien), historic extinctions (Martin (2022) NZ\_historic) and prehistoric extinctions (Martin (2022) NZ\_prehistoric\_NOMarine) were compiled. A version of the current bird fauna excluding alien species and also excluding marine species (Martin (2022) NZ\_noMarine\_Modern) was also compiled.

Literature cited:

Billerman, S.M., Keeney, B.K., Rodewald, P.G. & Schulenberg, T.S.(2022). *Birds of the World* Cornell Lab of Ornithology, Ithaca, NY, USA. Accessed 07/05/22 from: https://birdsoftheworld.org/

Diamond, J.M. (1984) Distributions of New Zealand birds on real and virtual islands. *New Zealand Journal of Ecology*, 7, 37–55.

International Union for Conservation of Nature (2021). *The IUCN Red List of Threatened Species. Version 2021-3*. Accessed 07/05/22 from: https://www.iucnredlist.org/

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012). The global diversity of birds in space and time. *Nature*, 491, 444–448.

New Zealand Birds Online (2022). New Zealand Birds Online. Accessed 07/05/22 from: <https://www.nzbirdsonline.org.nz/>

Robertson, H. & Heather, B. (2015). *The Hand Guide to the Birds of New Zealand. Second edition*. Penguin: Auckland.

### Zhao et al (2022) Zhoushan - 51

34 islands in Zhoushan archipelago, China. Data have been embargoed for a period of six months as are being used in a separate publication.

Seventy transects were surveyed for birds on 34 study islands, including 39 in forest habitats, 19 on farmland habitats, and 12 in mixed habitats. This study only considered breeding birds (resident and summer species) that mainly use terrestrial habitats on islands, excluding diving birds, ducks, and gulls whose habitats rely on or include water. In addition, nocturnal species (one species in this study: *Caprimulgus indicus*) were excluded.

# HABITAT ISLANDS

*General Points*

The filenames match those that are stored in the data files within the repository. The numbers at the end of the filenames correspond to the dataset numbers in Table S1 in the Supplementary Information of the paper. The full references for the source papers can be found there also.

Several datasets taken from the FragSAD database (Chase et al., 2019; further meta-data are available from there).

Chase, J.M., Liebergesell, M., Sagouis, A., May, F., Blowes, S.A., Berg, Å. et al. (2019) FragSAD: a database of diversity and species abundance distributions from habitat fragments. *Ecology*, 100, e02861.

### battisti et al (2009) Anzio - 1

Birds in 13 oak forest fragments in central Italy (Anzio – Nettuno: southern sector of province of Rome). Only breeding birds surveyed. Data taken from Table 2 and the Appendix.

*Corvus corone cornix* taken as *Corvus corone* as hooded crow not in BirdTree.

### battisti et al (2009) cornicolan hills\_noAliens - 2

Birds in 20 oak forest fragments in central Italy (Cornicolan hills: between the Central Apennines and suburban area of Rome). Only breeding birds surveyed. Data taken from Table 2 and the Appendix.

One introduced species (*Phasianus colchicus*) removed.

*Corvus corone cornix* taken as *Corvus corone*; *Passer\_italiae* not in BirdTree and taken as *Passer domesticus*.

### Berg\_1997\_birds - 3

Birds in 12 forest fragments (only forest fragment data used) in central Sweden. Data taken from FragSAD database (code = Berg\_1997).

### Blake (1991) birds Illinois\_noAliens - 4

Birds in woodland fragments in Illinois, USA. Species are all breeding residents (and the list does not include flyovers, wide ranging species [e.g. birds of prey, crows] or nocturnal species). We use the species list for the year 1980. Two introduced species removed.

### CSV-brotons and herrando 2001 ECJ - 5

Birds in secondary pine forest fragments, in Spain, surrounded by agricultural land. Dataset provided by L.B. Only includes breeding species.

*Picus viridis* now split but the split is not in BirdTree.

### CSV-Cieslak & Dombrowski 1993 ECJ - 6

Fragments come from two forest “archipelagos” (one sampled in 1980 and one in 1983) in Eastern Poland, separated by about 45 km. We have followed the authors in combining them into one dataset. Only breeding species.

### Daily et al (2001) birds - 7

Birds in forest fragments within agricultural land in Costa Rica. We only included the eight surveyed fragments. Birds that fly high overhead (e.g. raptors, swifts, swallows, parrots), owls, and birds closely associated with water were not included.

*Phaethornis superciliosus* taken as *Phaethornis longirostris*; *Phaethornis longuemareus* taken as *Phaethornis\_striigularis*; *Pipra\_cornuta* taken as *Pipra\_mentalis*; *Schiffornis turdine, Momotus momota, Zimmerius vilissimus* and *Platyrinchus mystaceus* recently split but splits not in BirdTree; *Piranga flava* taken as *Piranga\_lutea*; *Microcerculus philomela to Microcerculus marginatus; Saltator albicollis* taken as *Saltator\_striatipectus*; *Microcerculus luscinia* is a subspecies (not in BirdTree) and so we use the full species in BirdTree.

### Dami\_2012 - 8

Birds in forest fragments in Nigeria. Data taken from FragSAD database (code = Dami\_2012).

*Schoenicola platyurus* taken as *Schoenicola\_brevirostris*; *Emberiza tahapisi* is a new split not in BirdTree; *Bubo\_africanus* taken as *Bubo cinerascens*; *Anthus novaeseelandiae*, *Sheppardia bocagei* and *Emberiza tahapisi* now split but splits not in BirdTree.

CSV-dos Anjos and Boçon, 1999 ECJ - 9

Birds in naturally fragmented rainforest patches in southern Brazil. We have included the largest patch (A), which in some places is described as continuous forest but in others as a patch and fragment. But the authors do warn that sampling intensity is lower in (A) due to its larger size, thus the richness will be underestimated. Dataset provided by L.dA.

*Basileuterus\_culicivorus*, *Hirundinea ferruginea*, *Myiodynastes maculatus, Elaenia obscura,* and *Tityra cayana* recently split but splits not in BirdTree. *Turdus nigriceps* changed to *Turdus subalaris; Euphonia musica* to *Euphonia cyanocephala*

### dos Anjos (2004) human-fragments - 10

Birds sampled in 14 fragments from the same region as dos Anjos & Boçon, but these are anthropogenically fragmented islands rather than naturally fragmented islands. The sampled area (i.e., number of point counts) was kept the same in each fragment, meaning proportionally more area of the smaller fragments was sampled and the species richness of the larger fragments will likely be underestimated. Species only recorded once in a single point count in a fragment were marked (by the source paper author) as absent. Dataset provided by L.dA.

*Basileuterus\_culicivorus*, *Tityra cayana* and *Myiodynastes maculatus* now split but splits not in BirdTree. *Turdus nigriceps* taken as *Turdus subalaris.*

### Edwards\_2010 – 11

Birds in logged forest fragments in Malaysia, surrounded by oil palm. Data taken from FragSAD database (code = Edwards\_2010). Apodidae and Hirundinidae not included.

*Arachnothera modesta* taken as *Arachnothera affinis* as split not in BirdTree; *Lonchura\_malacca* taken as *Lonchura atricapilla*. *Luscinia calliope* is a very rare record for north Borneo but is possible so kept in.

### CSV-Fernandez Juiricic 2000 ECJ\_noAliens - 12

Breeding birds in wooded urban parks in Madrid, Spain. Waterfowl and swifts excluded by source paper authors. Two introduced species removed by us.

*Picus viridis* now split but the split is not in BirdTree.

### CSV-Ford 1987 ECJ\_noAliens - 13

Birds in small woodland fragments in Oxfordshire, UK. We have included species classed either as resident or transient, as the definition of transient here could include species seen in the fragment in two sampling visits (i.e., properly recorded, not just flying over). Two water bird species and a winter visitor excluded by the source paper author. Two introduced species removed by us.

### CSV-Gillespie & Walter, 2001 ECJ - 14

Patches are tropical dry forest protected areas in Costa Rica and Nicaragua, surrounded by agricultural land. However, the forest cover is below 100% in each reserve. Authors give both reserve area and forest cover area – we have used the former. Resident breeding birds included, with migratory and nocturnal species excluded.

*Momotus momota, Aratinga strenua, Sittasomus griseicapillus,* and *Campylorhynchus rufinucha* recently split but splits not in BirdTree; *Caracara plancus* taken as *Caracara\_cheriway*; *Dendrocolaptes certhia* taken as *Dendrocolaptes\_sanctithomae.*

### CSV-Langrand (1995) ISAR\_justFRAGMENTS - 15

Forest birds in forest fragments in Madagascar. We included the seven smaller fragments but not the larger forest block (control site). Mist netting effort per unit area decreased with increasing fragment area. The author was focusing primarily on forest species and has excluded some non-forest species.

*Apus\_barbatus* taken as *Apus\_balstoni*. *Riparia paludicola* now split but split not in BirdTree.

### Martensen et al\_2012\_Cau - 16

Birds in forest fragments in the Atlantic Forest, Brazil: Caucaia do Alto landscape (intermediate landscape forest cover). Data taken from FragSAD database (code = Martensen\_2012). Understorey bird species sampled using mist nets.

*Basileuterus\_culicivorus* recently been split and nominate not in study area, but split not in BirdTree.

### Martensen et al\_2012\_RG - 17

Birds in forest fragments in the Atlantic Forest, Brazil: Ribeirao Grande landscape (low landscape forest cover). Data taken from FragSAD database (code = Martensen\_2012). Understorey bird species sampled using mist nets.

*Basileuterus\_culicivorus* recently been split and nominate not in study area, but split not in BirdTree.

### Martensen et al\_2012)\_TAP - 18

Birds in forest fragments in the Atlantic Forest, Brazil: Tapirai landscape (high landscape forest cover). Data taken from FragSAD database (code = Martensen\_2012). Understorey bird species sampled using mist nets.

*Basileuterus\_culicivorus* recently been split and nominate not in study area, but split not in BirdTree. *Onychorhynchus coronatus* changed to *Onychorhynchus swainsoni.*

CSV-McCollin 1993 ECJ\_noAliens - 19

Birds in woodland fragments, UK, surrounded by farmland. This dataset is in FragSAD but we used the version direct from the source paper. The data are for confirmed breeding species – the author removed four species that were considered to not be breeding. Woodpigeon removed by source paper author due to difficulties in estimating territories. One introduced species removed by us.

### CSV-Simberloff & Martin 1991 ECJ - 20

Land bird species in old coniferous taiga patches in southern Finland, embedded in a matrix of clear-cuts and young managed forest. Original data in Haila et al. (1987) – we have used the data as presented in Simberloff & Martin (1991). Reading Haila et al. (1987), it appears that the data are for breeding species, and they removed ‘transients’ where they could (but this was only possible in fragments that they sampled more than twice).

Haila, Y., Hanski, I.K. & Raivio, S. (1987) Breeding bird distribution in fragmented coniferous taiga in southern Finland. *Ornis Fennica*, 64, 90-103.

Simberloff, D. & Martin, J.L. (1991) Nestedness of insular avifaunas: simple summary statistics masking complex species patterns. *Ornis Fennica*, 68, 178–192.

### Ulrich\_2016\_birds - 21

Birds in cloud forest fragments in Kenya, surrounded by human settlements and pasture land. Data taken from FragSAD database (code = Ulrich\_2016). Sampled understorey birds using mist nets.

*Andropadus milanjensis, Zosterops abyssinicus, Zosterops poliogastrus,* and *Trochocercus cyanomelas* recently split but this is not in BirdTree.

### wang et al (2013) birds (resident and breeding) - 22

Birds in urban woods within an urban matrix in Hangzhou, China. Birds flying over not included, nor were raptors, common swift, Hirundinidae, some nocturnal and crepuscular species, certain exclusively migratory species (e.g. that use the woodlots only during migration passage), and waterfowl. We only included breeding bird species (residents and breeding migratory species), and not wintering species (data in Appendix A).

*Zosterops japonicus* kept as this but the systematics for this species complex are very messy and have been changed multiple times over the last decade – currently *Z. japonicus* is not considered present in China. *Turdus merula* and *Terpsiphone paradisi* now split but splits not in BirdTree. *Garrulax poecilorhynchus* taken as *Garrulax\_berthemyi*; *Phylloscopus reguloides* taken as *Phylloscopus goodsoni*.

### CSV-Watson, 2003 ECJ - 23

Birds in naturally fragmented mountain pine oak fragments in Oaxaca, Mexico. Patches span a large area gradient (2 ha to 159,246 ha). The original source paper only focuses on 60 forest species, but D.M.W. provided the full dataset, which includes all species recorded in the patches. This comprises mostly species from the D.M.W.’s own field sampling, augmented with records found in the literature and museum specimens.

*Aphelocoma\_coerulescens* taken as *Aphelocoma californica* (should actually now be *Aphelocoma woodhouseii* but this split is not in BirdTree). *Phaethornis superciliosus* taken as *Phaethornis longirostris*. *Pipilo erythrophthalmus* taken as *Pipilo maculatus*. *Piranga flava* taken as *Piranga\_hepatica*. *Colaptes auratus* and *Trogon elegans* now split but splits not in BirdTree.

### Wethered & Lawes 2005 Balgowan ECJ – 24

Birds in Afromontane forest fragments in South Africa. One of two datasets from this source paper; here (Balgowan landscape) the fragments are surrounded by a natural grassland matrix. Data are for the ‘total compliment’ of species in Table 3.

*Guttera edouardi* taken as *Guttera\_pucherani*; *Zosterops capensis* is a subspecies of *Z. virens*, which is a recent split and not in BirdTree, so *Z. pallidus* used instead.

### Wethered & Lawes 2005 Gilgoa ECJ – 25

Birds in Afromontane forest fragments in South Africa. One of two datasets from this source paper, here (Gilboa landscape) the fragments are surrounded by a plantation pine forest matrix. Data are for the ‘total compliment’ of species in Table 3.

*Guttera edouardi* taken as *Guttera\_pucherani*; *Zosterops capensis* is a subspecies of *Z. virens*, which is a recent split and not in BirdTree, so *Z. pallidus* used instead.