**Functional traits in birds**

The AVONET dataset contains comprehensive functional trait data for all birds, including six ecological variables, 11 continuous morphological traits, and information on range size and location. Raw morphological measurements are presented from 90,020 individuals of 11,009 extant bird species sampled from 181 countries. There’s a ton of data here to explore! For more information, see: https://onlinelibrary.wiley.com/doi/full/10.1111/ele.13898

Diagram

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Description** | **Variable Types** | **Units** |
| Sequence | Unique sequence identifier used by BirdLife international. Ordering the "BirdLife taxonomy" dataset this way will convert from alphabetical to the sequence used by BirdLife International | ID |  |
| Species1 | Species taxonomy according to BirdLife international | categorical |  |
| Family1 | Family-level taxonomy according to BirdLife international | categorical |  |
| Order1 | Order-level taxonomy according to BirdLife international | categorical |  |
| Species2 | Species taxonomy according to eBird (following latest updates in the Clements checklist) | categorical |  |
| Family2 | Family-level taxonomy according to eBird (following latest updates in the Clements checklist) | categorical |  |
| Order2 | Order-level taxonomy according to eBird (following latest updates in the Clements checklist) | categorical |  |
| Species3 | Species taxonomy and nomenclature according to BirdTree, the most recent global bird phylogeny | categorical |  |
| Family3 | Family-level taxonomy and nomenclature according to BirdTree, the most recent global bird phylogeny | categorical |  |
| Order3 | Order-level taxonomy according to BirdTree, the most recent global bird phylogeny | categorical |  |
| Total.individuals | Number of individual specimens and live birds measured for each species | discrete |  |
| Female | Number of measured individuals identified as female | discrete |  |
| Male | Number of measured individuals identified as male | discrete |  |
| Unknown | Number of unsexed individuals measured | discrete |  |
| Complete.measures | Number of measured individuals with complete set of morphometric trait measurements | discrete |  |
| Beak.Length\_Culmen | Length from the tip of the beak to the base of the skull | continuous | mm |
| Beak.Length\_Nares | Length from the anterior edge of the nostrils to the tip of the beak | continuous | mm |
| Beak.Width | Width of the beak at the anterior edge of the nostrils | continuous | mm |
| Beak.Depth | Depth of the beak at the anterior edge of the nostrils | continuous | mm |
| Tarsus.Length | Length of the tarsus from the posterior notch between tibia and tarsus, to the end of the last scale of acrotarsium (at the bend of the foot) | continuous | mm |
| Wing.Length | Length from the carpal joint (bend of the wing) to the tip of the longest primary on the unflattened wing | continuous | mm |
| Kipps.Distance | Length from the tip of the first secondary feather to the tip of the longest primary | continuous | mm |
| Secondary1 | Length from the carpal joint (bend of the wing) to the tip of the first secondary, i.e. the outermost secondary adjacent to the innermost primary feather. Secondary1 is roughly equivalent to Wing length minus Kipp’s distance (measured in a fully folded and flat wing) | continuous | mm |
| Hand-Wing.Index | 100\*DK/Lw, where DK is Kipp’s distance and Lw is wing length (i.e., Kipp’s distance corrected for wing size). Species average HWI differ from estimates in Sheard et al. (2020) because of much higher sampling of individuals in some species, as well as taxonomic effects in the BirdLife list | continuous |  |
| Tail.Length | Distance between the tip of the longest rectrix and the point at which the two central rectrices protrude from the skin, typically measured using a ruler inserted between the two central rectrices | continuous | mm |
| Mass | Body mass given as species average (incorporating both male and female body mass) | continuous | gram |
| Mass.Source | Dunning (= Dunning, JB [2008] CRC Handbook of Avian Body Masses); EltonTraits\_Other (= mass published by Wilman et al. 2014 EltonTraits 1.0. Ecology, on basis of literature other than Dunning, usually Handbook of the Birds of the World [now Birds of the World, Cornell]) EltonTraits\_GenAvg (= average of the genus, published in Wilman et al. 2014 EltonTraits 1.0. Ecology);  EltonTraits\_Model (= inferred from mass and length data, with mass-length relationships parameterized at family level, published in Wilman et al. 2014 EltonTraits 1.0. Ecology); Updated\_literature (= any previous data that was calculated indirectly by Wilman et al. 2014 [Genavg or Modelled] or missing due to taxonomic gaps in BirdLife/eBird was replaced with new data where available from literature and online resources, including Birds of the World - specific sources are listed in the Mass.Ref.Source column);  Updated\_live.sample (= any previous data that was calculated indirectly by Wilman et al. 2014 [Genavg or Modelled] or missing due to taxonomic gaps in BirdLife/eBird was replaced with new data where available from museum labels and/or field data of from wild birds weighed using a Pezola balance before release. Specific sources are listed in the Mass.Ref.Source column);  Inferred (= where EltonTraits GenAvg or Modelled data seemed misleading, and no published data were available for updating, we replaced it where possible with data inferred directly from a morphologically similar relative. Reference species and Traits inferred columns indicate where mass has been inferred and from which species); DataFromSplit (= daughter species body mass inferred from parent species after split; the parent or nominate form is typically morphologically similar and provides a reasonable estimate of body mass, particularly as mass data used to calculate the parent species body mass in earlier publications sometimes includes data from the recently split daughter species). | categorical |  |
| Mass.Refs.Other | Source of Updated mass data, including primary and secondary literature, online resources and museum specimen data. Primary and secondary literature citations or abbreviations are given in full in the sheet named "Mass\_Sources" | NA |  |
| Inference | NO = all biometric trait data were obtained from direct measurement of specimens, or in the case of body mass from published sources YES = at least some of the biometric traits were unavailable from specimens so these trait data were inferred from a closely related species; in some cases all morphological measurements were taken directly except for body mass, which can't be measured on museum specimens. | binary |  |
| Traits.inferred | List of traits inferred from the reference species | categorical |  |
| Reference.species | Species from which the biometric data were inferred | categorical |  |
| Habitat | Desert (= drylands and other open arid habitats, often sandy with very sparse vegetation);  Rock (= rocky substrate typically with no or very little vegetation, including rocky outcrops, rocky coastlines, arid stony steppes, rocky mountaintops and mountain slopes);  Grassland (= open dry to moist grass-dominated landscapes, at all elevations);  Shrubland (= low stature bushy habitats, included thornscrub, thorny or arid savanna, caatinga, xerophytic shrubland and coastal scrub);  Woodland (= medium stature tree-dominated habitats, including Acacia woodland, riparian woodlands, mangrove forests, forest edges, also more open parkland with scattered taller trees); Forest (= tall tree-dominated vegetation with more or less closed canopy, including palm forest);  Human modified (urban landscapes, intensive agriculture, gardens);  Wetland (= wide range of freshwater aquatic habitats including lakes, marshes, swamps and reedbeds);  Riverine (= associated with rivers and streams at all elevations);  Coastal (= intertidal zones within immediate vicinity of beaches, estuaries, brackish to salty marshes, including mudflats, lagoons, alkaline wetlands, coastal dunes and harbours); Marine (= pelagic, on sea near coasts, including species in the intertidal zone on beaches, and those pelagic species nesting near the sea on cliffs, islets and islands). | categorical |  |
| Habitat.Density | 1 = Dense habitats. Species primarily lives in the lower or middle storey of forest, or in dense thickets, dense shrubland etc. 2 = Semi-open habitats. Species primarily lives in open shrubland, scattered bushes, parkland, low dry or deciduous forest, thorn forest.  3 = Open habitats. Species primarily lives in desert, grassland, open water, low shrubs, rocky habitats, seashores, cities. Also applies to species living mainly on top of forest canopy (i.e. mostly in the open) | discrete |  |
| Migration | 1 = Sedentary.  2 = Partially migratory, i.e. minority of population migrates long distances, or most of population undergoes short-distance migration, nomadic movements, distinct altitudinal migration, etc. 3 = Migratory, i.e. majority of population undertakes long-distance migration | discrete |  |
| Trophic.Level | Herbivore = species obtaining at least 70% of food resources from plants; Carnivore = species obtaining at least 70% of food resources by consuming live invertebrate or vertebrate animals; Scavenger = species obtaining at least 70% of food resources from carrion or refuse; Omnivore = species obtaining resources from multiple trophic level in roughly equal proportion | categorical |  |
| Trophic.Niche | Frugivore = species obtaining at least 60% of food resources from fruit;  Granivore = species obtaining at least 60% of food resources from seeds or nuts;  Nectarivore = species obtaining at least 60% of food resources from nectar;  Herbivore = species obtaining at least 60% of food resources from other plant materials in non-aquatic systems, including leaves, buds, whole flowers etc.;  Herbivore aquatic = species obtaining at least 60% of food resources from plant materials in aquatic systems, including algae and aquatic plant leaves;  Invertivore = species obtaining at least 60% of food resources from invertebrates in terrestrial systems, including insects, worms, arachnids, etc.;  Vertivore = species obtaining at least 60% of food resources from vertebrate animals in terrestrial systems, including mammals, birds, reptiles etc.;  Aquatic Predator = species obtaining at least 60% of food resources from vertebrate and invertebrate animals in aquatic systems, including fish, crustacea, molluscs, etc;  Scavenger = species obtaining at least 60% of food resources from carrion, offal or refuse; Omnivore = Species using multiple niches, within or across trophic levels, in relatively equal proportions | categorical |  |
| Primary.Lifestyle | Aerial = species spends much of the time in flight, and hunts or forages predominantly on the wing;  Terrestrial = species spends majority of its time on the ground, where it obtains food while either walking or hopping (note this includes species that also wade in water with their body raised above the water);  Insessorial = species spends much of the time perching above the ground, either in branches of trees and other vegetation (i.e. arboreal), or on other raised substrates including rocks, buildings, posts, and wires;  Aquatic = species spends much of the time sitting on water, and obtains food while afloat or when diving under the water's surface; Generalist = species has no primary lifestyle because it spends time in different lifestyle classes | categorical |  |
| Min.Latitude | The minimum latitudinal extent of the species range (restricted to breeding and resident range) | continuous | decimal degrees |
| Max.Latitude | The maximum latitudinal extent of the species range (restricted to breeding and resident range) | continuous | decimal degrees |
| Centroid.Latitude | The geometric centre of the species range (restricted to breeding and resident range), calculated using the calcCentroid function from the R package PBSmapping (Schnute *et al.*, 2017). This function projects ranges into the Universal Transverse Mercator (UTM) projection, to accurately calculate centroids at high latitudes. To generate the breeding and resident range, we used the following settings: Presence: 1 (Extant). Origin 1 & 2 (Native & Reintroduced). Seasonality 1 & 2 (resident and breeding season) | continuous | decimal degrees |
| Centroid.Longitude | The geometric centre of the species range (restricted to breeding and resident range), calculated using the calcCentroid function from the R package PBSmapping (Schnute *et al.*, 2017). This function projects ranges into the Universal Transverse Mercator (UTM) projection, to accurately calculate centroids at high latitudes. To generate the breeding and resident range, we used the following settings: Presence: 1 (Extant). Origin 1 & 2 (Native & Reintroduced). Seasonality 1 & 2 (resident and breeding season) | continuous | decimal degrees |
| Range.Size | The total area of the mapped range of the species (not the Extent of Occurrence [EOO]). We used maps shared by BirdLife International and restricted our analysis to areas of the range coded as Presence = 1 (Extant only), Origin = 1 & 2 (Native & Reintroduced), Seasonal = 1 & 2 (Resident and breeding). We calculated the total combined mapped area of these parts of the range using the areaPolygon function from the R package geosphere (Hijmans *et al.*, 2011). This function accurately calculates the area in the World Geodetic System (WGS84) projection using spherical distances | continuous | km2 |
| Species.Status | Extant = the species is extant according to BirdLife international Extinct = the species is extinct according to BirdLife international  Invalid = the species is currently not recognised by BirdLife international on the basis of published information suggesting it does not exist (e.g. it may be a hybrid taxon, a genetic variant, or fails to meet criteria for species status) | categorical |  |
| Avibase.ID | Unique identifier for species (or species group) linking to the Avibase dataset which provide detailed information on taxonomy, synonymy and distribution for the world's birds. See https://avibase.bsc-eoc.org/ | ID |  |
| Avibase.ID1 | Unique Avibase identifier for the species concept according to BirdLife international | ID |  |
| Avibase.ID2 | Unique Avibase identifier for the species concept according to eBird | ID |  |
| eBird.species.group | Names of taxonomic units linking trait data in the AVONET raw data to eBird. Unlike the species-level data in AVONET 2, this list contains standard eBird names for individuals that cannot be assigned to species (e.g. *Acanthis* hornemanni/flammea/cabaret for redpolls that could not be assigned to a particular species) | categorical |  |
| Data.type | 1 = museum specimen 2 = live individual | binary |  |
| Source | Abbreviation of the source collection (museum or research collection). See the Data Sources tab for the full list of institutional names.  Field1 = unpublished field data. Field2 = published field data. | categorical |  |
| Specimen.number | Specimen accession and/or collection number (including corrections to erroneous numbers) | ID |  |
| Sex | Sex of the individual sampled: M = Male; F = Female; U = Unknown | categorical |  |
| Age | 0 = all biometric trait data were obtained from direct measurement of an adult specimen.  1 = all biometric trait data were obtained from direct measurement of an immature specimen. Note that juvenile specimens or any specimen notably less than adult size were removed from the dataset. | binary |  |
| Locality | Locality at which the individual was collected or trapped-released. Infomation presented here is based on data submitted by the measurers, with some corrections where old locality assignments have been updated. For museum specimens this is taken from the label, hence some mis-transcriptions still need correcting. Data are embargoed for 1 year to allow completion of this process. | NA |  |
| Country\_WRI | Country in which the individual was collected or trapped/released following boundaries deignated by the World Resources Institute, for the purposes of mapping. | categorical |  |
| Country | Country (or other designation such as archipelago, state, dependency etc.) in which the individual was collected or trapped/released. Infomation presented here is based on data submitted by the measurers, with corrections where old country assignments have been updated. For museum specimens this information is typically transcribed from hand-written labels, hence some mis-transcriptions still need correcting. Data are embargoed for 1 year to allow completion of this process. | categorical |  |
| Measurer | Unique identifier (initials or other abbreviation) used to identify each individual measurer. See the Measurers tab for the full list of names. | categorical |  |
| Protocol | 0 = biometric measurements not taken using AVONET protocol. 1 = biometric measurements taken using AVONET protocol. | binary |  |
| Publication | In AVONET\_Raw\_Data: Citation for specimen measurements taken from published sources. | NA |  |
| Citation | In Mass\_Sources tab: Citation or abbreviation used in the Mass.Refs.Other column in AVONET\_1–3 |  |  |
| Full Reference | In Mass\_Sources tab: Full details of the source cited in the Mass.Refs.Other column in AVONET\_1–3 |  |  |
| Institution.name | In Data\_Sources tab: Full name of museum or collection from which traits were measured | NA |  |
| Specimens.sampled | In Data\_Sources tab: Number of specimens measured in each museum or collection | discrete |  |
| Institution.city | In Data\_Sources tab: City where museum or collection is located | categorical |  |
| Institution.country | In Data\_Sources tab: Country where museum or collection is located | categorical |  |
| Latitude | In Data\_Sources tab: Latitude of museum or collection | coordinate |  |
| Longitude | In Data\_Sources tab: Longitude of museum or collection | coordinate |  |
| Address | In Data\_Sources tab: Full address of museum or collection | NA |  |
| Measurer.Name | In Measurers tab: Full name of each individual measurer | categorical |  |
| Individuals.Measured | In Measurers tab: Number of museum specimens and/or individual live birds measured by each measurer | discrete |  |
| Match.type | In crosswalk tabs: Type of taxonomic match between source and target taxonomy. 1-to-1 indicates a direct match with like-for-like names; 1-to-1 marked with an asterisk (\*; e.g. 1BL to 1eBird\*) indicates that the match is imperfect, typically because one or more subspecies are assigned differently in the two treatments; 1-to-many indicates that a single taxon in the first taxonomy represents multiple taxa in the second taxonomy; many-to-1 indicates the reverse; many-to-many indicates a complex match with multiple taxa in both treatments; newly described species are cases which are not currently recognised by a taxonomy. | categorical |  |
| Match.notes | In crosswalk tabs: Notes on taxonomic matching in the BirdLife-eBird crosswalk. A conservative approach was taken to assigning eBird subspecies or species groups to BirdLife species. In many cases, some specimens of a particular BirdLife taxon can be assigned to an eBird species while other specimens of the same BirdLife taxon can't be assigned because of uncertain provenance (e.g. locality data are ambiguous or lacking). This means that different specimens of individual BirdLife taxa may map onto multiple eBird categories, often including both species and species groups. Notes also highlight cases involving imperfect matches between taxa with the same name but a different allocation of subspecies. | NA |  |