EMU AUSTRAL ORNITHOLOGY

Emu - Austral Ornithology

ISSN: 0158-4197 (Print) 1448-5540 (Online) Journal homepage: https://www.tandfonline.com/loi/temu20

Revealing the potential of a huge citizen-science platform to study bird migration

Stephanie Caroline Schubert, Lilian Tonelli Manica & André De Camargo Guaraldo

EMU - AUSTRAL ORNITHOLOGY
2019, VOL. 119, NO. 4, 364–373
https://doi.org/10.1080/01584197.2019.1609340

Revealing the potential of a huge citizen-science platform to study bird migration

Stephanie Caroline Schubert, Lilian Tonelli Manica and André De Camargo Guaraldo

Behavioral Ecology and Ornithology Lab, Universidade Federal do Paraná (UFPR), Curitiba, Brazil

ABSTRACT
The currently best-known ornithological citizen-science platform is Cornell's eBird, which provides crucial information for bird migration studies. Considering the solid validity of eBird data, and after a validation process, we comparatively explored the data available in the Brazilian-wide platform WikiAves for bird migration studies. We selected five migratory and four resident species as models, controlling for likely sampling biases derived from efforts by the platform collaborators. If data in WikiAves were adequate for migration studies, we respectively expected, after a between-platform comparison, similar yearlong seasonal and non-seasonal occurrence records of all migratory and resident species. Data analysis supported our expectations: eBird and WikiAves data showed consistent temporal occurrence patterns for all evaluated species. Therefore, we selected another six model-species showing literature inconsistency on their migratory behaviour, demonstrating for the first time the potential of a Brazilian citizen-science database – WikiAves – in unveiling geographically seasonal occurrence patterns of the understudied migratory bird species in Brazil. Our study highlights the general public relevance on reducing knowledge gaps about bird migration in Brazil, revealing a feasible strategy to overcome some current logistic barriers that preclude advances in South American bird migration studies, a currently underexplored research area, especially by Brazilian researchers.

ARTICLE HISTORY
Received 24 August 2018
Accepted 16 April 2019

KEYWORDS
Birdwatching; circular statistical analyses; eBird; species' seasonality; WikiAves

Introduction

In the current 'big data' era, researchers must find efficient analytical tools for exploring huge databases that ensure trustworthy results (Hampton *et al.* 2013) and thus knowledge advances. Studies on bird migration, especially in the understudied South American intratropical migration system (Faaborg *et al.* 2010b), would largely benefit from identifying and describing seasonal occurrence patterns of species over broad geographical and temporal ranges amidst such large volumes of data. Acquiring these data in the field is time-consuming, costly, and often represents a barrier to scientific advances in understudied regions such as the Neotropics (Faaborg *et al.* 2010b). Thus, a suitable alternative for developing bird migration studies in knowledge-limited regions may be through using databases fed by volunteers and citizen-scientists (Silvertown 2009; Cooper *et al.* 2014). Citizen-science data can be used to feed structured, semi-structured, and unstructured databases, which may undergo different degrees of validation processes (Wiggins *et al.* 2011; Bonter and Cooper 2012; La Sorte *et al.* 2018). These records are freely available to the community, ultimately becoming an input to scientific studies which may feedback citizens' demand for knowledge on nature (McCaffrey 2005). This strategy is widely and successfully used by North American and European scientists (Audubon Society 2017; European Bird Census Council 2017), but yet underexplored in South America (but see Somenzari *et al.* 2018).

There are multiple online and worldwide citizen-science ornithological platforms, such as eBird (www.ebird.org) and xeno-canto (www.xeno-canto.org). In addition to submitting their records to these semi-structured platforms, Brazilian birdwatchers often do so to WikiAves (www.wikiaves.com.br). This unstructured 10 year-old citizen-science database hosts over 2 million photographs and >120 000 audio recordings of >1870 bird species, registered throughout the country and abroad by over 27 000 birdwatchers and academics (e.g. ornithologists) (WikiAves 2017). Surprisingly, this huge database has rarely been used by researchers and often merely as a depository of 'digital vouchers' (e.g. photographs) of new species' records in a given area (Biancalana *et al.* 2013; e.g. Godoi *et al.* 2012; but see Lees and Martin 2014 for an analytical approach to the WikiAves records). Such underuse of this platform

CONTACT Stephanie André de Camargo Guaraldo guaraldoac@gmail.com
Supplemental data for this article can be accessed here.

may be a consequence of the complex task of deriving conclusions based on the large amount of data available in it. In general, data in such datasets require proper treatment (Bonter and Cooper 2012) for removing likely biases in sampling efforts that may implicate noise and spurious correlations in data analysis results (Kelling *et al.* 2019; La Sorte *et al.* 2018).

Brazil has the third richest avifauna in the world (BirdLife International 2018) and a continental territorial area of >8.5 million km^2 (IBGE, I.B.d.G.E 2018). Nevertheless, it has only ~390 ornithologists with a Ph.D. (Plataforma Lattes 2017) and, as a likely consequence, several Brazilian bird species remain understudied even in relation to their breeding biology, one of their most conspicuous aspects (Xiao *et al.* 2016). These aspects highlight the challenge of deepening our knowledge of bird biology in this region, especially when dealing with movement and multiple sites such as occurs with migration (e.g. Somenzari *et al.* 2018). Thus to stimulate further research on this behaviour in Brazil and South America as a whole we tested a protocol for guiding the proper use of the WikiAves platform database in bird migration studies. We hypothesised that the WikiAves database is as reliable as eBird for studies on migratory species' seasonality based on occurrence records. Thus, we predicted that seasonal patterns of either resident or migrant species' occurrence would be highly similar between both platforms. Specifically, our goals were (i) to check the validity of the WikiAves data on defining species' annual seasonality, i.e. by correctly assigning a given species as resident or migrant; (ii) to compare the patterns found using WikiAves data with those revealed by the eBird database; and (iii) to further elucidate the migratory status of some bird species for which the literature fails to achieve an agreement (see Methods). In achieving our goals, we expect our study to help establish a landmark to bird migration studies in Brazil and South America, highlighting the unprecedented amount of data stored in the WikiAves platform and unveiling its potential, adequacy, and limitations to fill in several existing knowledge gaps on bird migration and thus aid advances in bird migration theory as a whole.

Methods

Model species and region of study

WikiAves users provide their personal data and metadata for every record uploaded to the platform, such as species' identification, record date, and coarse location (e.g. county or city). Then, each record undergoes a validation process by field-experienced moderators and senior users, attributing to them a dynamic weighted score (Wscore) of the species' taxonomic correctness, ensuring confidence in the identification of each bird record in the database (details of the scoring system are available in Portuguese only at http://www.wikiaves.com.br/wiki/wikiaves:peso_de_id; WikiAves 2017). Within this database, we harvested all photographic records of 15 species selected as models, namely the Swallow-tailed Kite (*Elanoides forficatus*), Plumbeous Kite (*Ictinia plumbea*), Vermilion Flycatcher (*Pyrocephalus rubinus*), Variegated Flycatcher (*Empidonomus varius*), Fork-tailed Flycatcher (*Tyrannus savana*), Roadside Hawk (*Rupornis magnirostris*), Great Kiskadee (*Pitangus sulphuratus*), House Wren (*Troglodytes musculus*), Rufous-collared Sparrow (*Zonotrichia capensis*), Southern Lapwing (*Vanellus chilensis*), Red-eyed Vireo (*Vireo olivaceus*), Grey-breasted Martin (*Progne chalybea*), Creamy-bellied Thrush (*Turdus amaurochalinus*), Blue-black Grassquit (*Volatinia jacarina*), and the Shiny Cowbird (*Molothrus bonariensis*).

After a thorough literature review, we defined the first five species in this list as migrants, the next four as residents and the latter six as showing inconsistent migratory behaviour (inconsistent, hereafter). This was achieved after considering six avifaunistic studies carried out for at least an entire year in the study region (Anjos and Graf 1993; Krügel and Anjos 2000; Galina and Gimenes 2006; Lopes and Anjos 2006; Scherer Neto and Bispo 2011; Vogel *et al.* 2011) plus one with a country-wide coverage and which is considered the best compilation of Brazilian ornithological data of its time (Sick 1997). We arbitrarily assigned a migratory or residency behaviour to those species with ≥80% consensus between the studies and inconsistent behaviour to those with <80% (details in Table S1). Not one of the six above-mentioned studies provides details on species' migratory patterns (i.e. obligatory or partial migration), and therefore we anticipate that inconsistencies may include partial migratory species (see, for example, the status of *Vireo chivi*, *Progne chalybea*, and *Turdus amaurochalinus* in Table 1 in Somenzari *et al.* 2018). Species' classification follows Remsen Jr *et al.* (Remsen *et al.* 2017).

Since WikiAves is a presence-only database, and thus with limited applications (Guillera-Arroita *et al.* 2015), we adopted a non-naive and more complex method to select model species (Isaac *et al.* 2014). For minimising likely biases in species' records due to variations in species-specific detectability and the wide range of observers' and photographers' expertise, we intentionally restricted our models to non-rare and conspicuous species that inhabit open vegetation areas and for which identification can be 100% verified based solely on a single photo. Since WikiAves lacks tools to ensure bulk data download, data collection demanded individual access to each record (detailed below) made possible through a semi-automated data-mining routine. Nevertheless, and for optimisation purposes, we restricted our sampling to Paraná state (~22.787426 to ~26.265134° S, ~53.382340 to ~48.690022° W), the region within Brazil which had a large number of records available in the platform. Also, the state has narrow latitudinal and longitudinal variation in comparison to other Brazilian states, an important aspect to consider for avoiding the likely regional variations in migratory patterns of species in the Neotropics (e.g. Alves 2007).

Data collection and treatment

We searched for every record of each model species available in WikiAves, from creation of the website in 2008 until 31 May 2016. Data harvesting and use were in accordance with the policies of the website (WikiAves 2017) and Brazilian copyright laws (Brasil 1998, 2013). We used data-mining tools implemented in package tm (Feinerer and Hornik 2015) for the R 3.3.2 environment (R Core Team 2018) for extracting the following data from each record (an R script commented sample is provided as supplemental material S1): record serial number, species, location, date, Wscore, and any additional observations provided by users. We excluded from analysis all records with unsure species ID (i.e. Wscore ≤0), comprising 0.01% of our dataset. All records compiled for this study are cited as supplemental material (supplemental material S2). WikiAves policy recommends users to upload a single photo of each individual bird. Nevertheless, we doubled-checked our database for excluding pseudo-replicates derived from records of a same individual bird made by both the same and multiple users (i.e. an individual bird record made in the same area and date by more than one user). Users' comments on their records (e.g. statements on other users that also recorded that individual bird) provided additional confidence in our procedure. Pseudo-replicates totalled ~6% of our initial dataset.

As in many other citizen-science databases, the monthly effort of WikiAves users photographing birds may vary (e.g. low activity in cold/rainy months; higher activity during birds' breeding season). Thus, we corrected the likely monthly bias of records of each model species (e.g. more records in hotter or drier months) by weighting the total number of records of a given species in each month by the total monthly records made of all bird species in the study area. Given the large number of monthly records available in the database, this process resulted in very low monthly ratios of each species, which were then amplified by a factor of 10 000 and rounded to the nearest integer for allowing us to run the circular statistics analysis (see Data analysis below).

We requested all records for the 15 model species from eBird administrators, which dated back from 2002 until 31 May 2016. Every record in this database is subjected to an intensive and thorough verification process for avoiding misreports (https://help.ebird.org/customer/en/portal/articles/1055676-understanding-the-ebird-review-and-data-quality-process). We filtered out the dataset for complete lists only, i.e. database entries in which collaborators uploaded only a full species' account of their recording activities, thus avoiding any partial records of species' occurrence. Despite the eBird database allowing the filtering out of pseudo-replicates by record ID, we double-checked and confirmed the lack of such an issue by looking for duplicate records with the same sampling event and group identifiers, values assigned by the eBird system to every uploaded record.

Data analysis

We pooled the monthly records for each species from the entire data range of each platform (WikiAves: 2008–2016; eBird: 2002–2016), grouping data for each month into 30 categories (i.e. 0–29° for January records, 30–59° for February, and so on). There are multiple approaches for analysing cyclic data (i.e. general additive models with a cyclic cubic spline; Wood 2017, p. 202), but, given our study goals, we described each species' records' seasonality using circular distribution descriptors, a simpler and thus more parsimonious approach, specifically using the mean vector (μ), its length (r) and circular SD (Zar 2010a). The first refers to the mean occurrence date of a given species in the study area, and the second to data concentration around this mean date (i.e. resultant vector), ranging from 0 in the absence of a mean angle (i.e. data with uniform circular distribution) to 1 when all data converge to the same direction (Zar 2010a). We did so for data gathered from both WikiAves and eBird databases, expecting to find similar patterns between platforms, especially considering the already known validity of the eBird database for bird migration studies (e.g. Kelly *et al.* 2016; La Sorte *et al.* 2016). To check for data consistency between these platforms we performed a Wallraff test of angular distance which is based on the Kruskal–Wallis rank sum test (Wallraff 1979; Zar 2010b). Non-significant results for species' pairwise comparisons would confirm that each species has similar circular distribution of occurrence records in both platforms.

In order to evaluate the consistency of the WikiAves database and to compare species' temporal occurrence patterns between WikiAves and eBird we ran three additional sets of analysis . The first analysis validated the WikiAves data by individually assigning the migratory status of each bird species. This was done considering only those species with known migratory or resident behaviour (respectively, M and R in Table S1), and through a chi-squared test to compare the circular distribution of each species' data against a uniform distribution pattern. If WikiAves consists of a suitable citizen-science database for inferring species' migratory behaviour, we expected that resident species' records would show a uniform (i.e. non-seasonal) circular distribution pattern, whereas migrants would show a unimodal pattern (i.e. seasonal; Zar 2010a). In the second analysis, we ran a set of two-tailed Mann–Whitney's rank test for data from both platforms to compare vector length (r) values found for migratory and resident species, expecting that this parameter would be significantly higher for migrants. Finally, we ran another set of chi-squares analysis on data gathered from both platforms for clarifying the migratory status of the six species whose behaviour is inconsistent in the literature (Table S1). We ran all analysis in R (R Core Team 2018) using the packages circular (Agostinelli and Lund 2013) and ggplot2 (Wickham 2009). We assumed a 0.05 significance level for all tests except the pairwise comparisons, in which we adjusted Bonferroni corrected p-values.

Results

In total we collected 8778 valid records from the WikiAves database and 10 874 from eBird for the 15 model species (Table 1). Our analysis confirmed the migratory status for all nine model species unequivocally listed in the literature as residents or migrants, revealing the circular distribution congruency of the records between WikiAves and eBird platforms (Figure 1 and Table 1). Moreover, all species had a similar distribution pattern of records in both platforms (Wallraff tests:

Table 1. Descriptive circular statistics of occurrence data compiled from the WikiAves and eBird platforms for migrants (M), residents (R), and for species with inconsistent migratory status in the region (In; see Methods). Mean vectors (μ) represent species' mean occurrence dates in the studied region (Paraná state, Brazil), which were compared between platforms through the Wallraff test of angular distances. Concentration (r) shows the data distribution around the mean date, ranging from 0 (non-seasonal distribution) to 1 (highly seasonal). WA: WikiAves; eB: eBird

		n		r		μ ± SD (month)		Wallraff test
Species	Migratory status	WA	eB	WA	eB	WA	eB	χ^2 (*p*-value)
Swallow-tailed Kite (*Elanoides forficatus*)	M	163	116	0.54	0.61	284 ± 64° (Oct.)	305 ± 57° (Nov.)	1.82 (0.40)
Plumbeous Kite (*Ictinia plumbea*)	M	272	170	0.72	0.71	296 ± 45° (Oct.)	312 ± 48° (Nov.)	1.18 (0.55)
Vermilion Flycatcher (*Pyrocephalus rubinus*)	M	691	65	0.56	0.82	163 ± 62° (Jun.)	167 ± 36° (Jun.)	0.89 (0.64)
Variegated Flycatcher (*Empidonomus varius*)	M	465	323	0.64	0.51	318 ± 54° (Nov.)	311 ± 66° (Nov.)	2.48 (0.29)
Fork-tailed Flycatcher (*Tyrannus savana*)	M	685	516	0.70	0.69	298 ± 48° (Oct.)	321 ± 49° (Oct.)	0.73 (0.69)
Roadside Hawk (*Rupornis magnirostris*)	R	968	830	0.25	0.14	135 ± 95° (May)	123 ± 114° (May)	2.28 (0.32)
Great Kiskadee (*Pitangus sulphuratus*)	R	885	2048	0.08	0.09	235 ± 128° (Aug.)	113 ± 125° (Apr.)	2.40 (0.30)
House Wren (*Troglodytes musculus*)	R	616	1144	0.22	0.01	280 ± 100° (Oct.)	245 ± 163° (Sept.)	0.15 (0.93)
Rufous-collared Sparrow (*Zonotrichia capensis*)	R	744	1095	0.14	0.05	282 ± 114° (Oct.)	342 ± 139° (Dec.)	2.48 (0.29)
Southern Lapwing (*Vanellus chilensis*)	In	974	1626	0.11	0.07	256 ± 120° (Sept.)	58 ± 132° (Feb.)	0.33 (0.85)
Red-eyed Vireo (*Vireo olivaceus*)	In	239	552	0.57	0.52	321 ± 61° (Nov.)	320 ± 65° (Nov.)	0.94 (0.63)
Grey-breasted Martin (*Progne chalybea*)	In	312	635	0.30	0.33	306 ± 89° (Nov.)	320 ± 85° (Nov.)	0.76 (0.68)
Creamy-bellied Thrush (*Turdus amaurochalinus*)	In	583	786	0.29	0.09	269 ± 9° (Sept.)	251 ± 127° (Sept.)	1.90 (0.39)
Blue-black Grassquit (*Volatinia jacarina*)	In	521	310	0.48	0.51	339 ± 69° (Dec.)	12 ± 66° (Jan.)	0.62 (0.73)
Shiny Cowbird (*Molothrus bonariensis*)	In	660	658	0.25	0.08	289 ± 96° (Oct.)	257 ± 127° (Sep.)	1.46 (0.48)

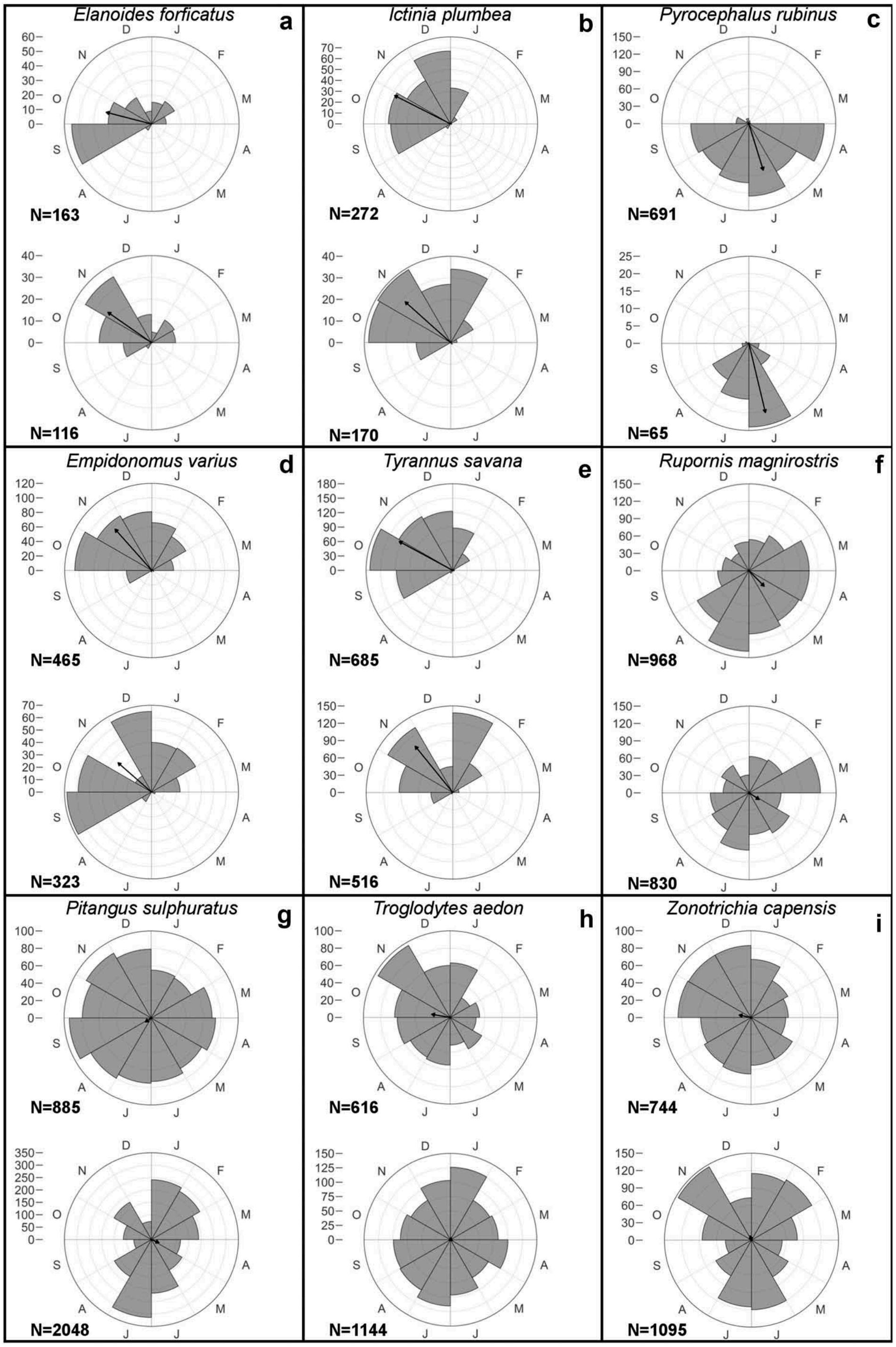

Figure 1. Circular histograms of records compiled from the WikiAves (upper graph) and eBird (lower graph) platforms for migratory (A–E) and resident species (F–I). The scales (pictured on the left) of the histograms vary between species and represent each species' weighted monthly proportion of records ([Monthly records].[10 000].[Total number of records of all species in the study region] − 1). Black arrows represent data concentration, which varies from 0 (non-seasonal distribution) to 1 (seasonal distribution).

$p > 0.05$ in all cases; Figure 1 and Table 1), with significant differences (WikiAves: U5,4 = 20; $p = 0.02$; eBird: U5,4 = 20; $p = 0.02$) between the mean vector length of migrants (median ± SD; WikiAves r = 0.64 ± 0.08; eBird: r = 0.69 ± 0.12) and residents (WikiAves: r = 0.18 ± 0.08; eBird: r = 0.07 ± 0.06).

Among the six species for which we found inconsistency in their migratory behaviour in the literature, the Red-eyed Vireo (*Vireo olivaceus*) showed an occurrence data distribution consistent with those of migrants, with a mean vector length >0.50 and low dispersion around the mean angle (Table 1 and Figure 2(B)). Four species showed evidences of residency in the study region, i.e. low concentration values and records spread throughout the year (Table 1 and Figure 2(A, C, D, F)), namely the Southern Lapwing, the Grey-breasted Martin, the Shiny Cowbird, and the Creamy-bellied Thrush. The Blue-black Grassquit had intermediate concentration values, with a decrease in partial occurrence from April to October (Table 1 and Figure 2(C)), suggesting a partial migratory behaviour in the study region.

Discussion

Our results showed that (i) data available in the unstructured citizen-science database WikiAves allow researchers to discriminate between the annual occurrence patterns of migratory and resident species; (ii) WikiAves and the semi-structured citizen-science platform eBird databases revealed consistent species' occurrence patterns throughout the year; and (iii) such databases can be explored through relatively simple analytical approaches to unveil long-standing uncertainties on the migratory status of many species. In that way our study validates for the first time the WikiAves platform as a trustworthy source of citizen-science data for bird migration studies in Brazil.

Citizen-science platforms are currently widespread in ornithology, thus enabling important advances in studies with bird migration, from regional-scale initiatives, such as in the Bird Track project in the UK

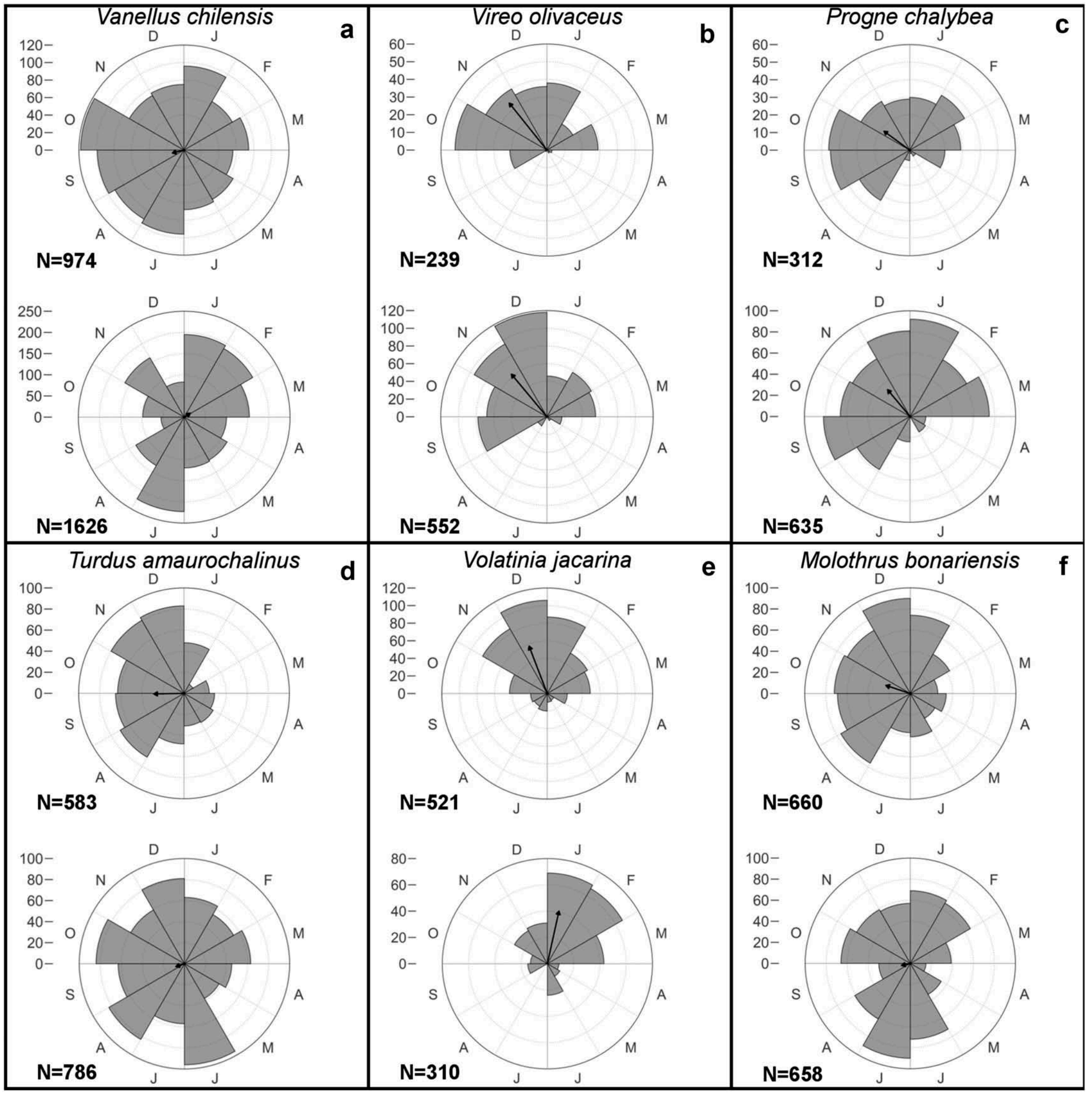

Figure 2. Circular histograms of records compiled from the WikiAves (upper graph) and eBird (lower graph) platforms for species whose migratory status is inconsistent in the literature. The scales (pictured on the left) of the histograms vary between species and represent each species' weighted monthly proportion of records ([Monthly records].[10 000].[Total number of records of all species in the study region] − 1). Black arrows represent data concentration, which varies from 0 (non-seasonal distribution) to 1 (seasonal distribution).

(http://www.bto.org/birdtrack) and Migration Watch (Baillie *et al.* 2006), to global ones such as the eBird online checklist program (http://www.ebird.org). Since researchers are currently facing an increasing demand for large geographic and temporal-scale datasets for addressing high-priority global ecological questions such as the impacts of climate change on bird migration (e.g. Pearce-Higgins *et al.* 2015; Tellería *et al.* 2015; Nourani *et al.* 2017), citizen-science data have been playing an important role in ornithology (e.g. Cooper *et al.* 2014). Thus, it is striking to realise that, to the best of our knowledge, only a single researcher explored the enormous database from WikiAves to study bird migration in Brazil (Lees 2016). Through explicitly highlighting the potential of WikiAves for assessing species' migratory patterns, we expect a shift in this scenario in the coming years and further advances to fill in the gaps on the migratory status of a number of species that occur in Brazil (as highlighted by Somenzari *et al.* 2018). We acknowledge that the limited use of WikiAves likely arises from the fact that, in comparison to other citizen-science platforms such as eBird, downloading data from it is a difficult task. Thus, by providing a protocol that enables an easier download process of bulk data from WikiAves we expect other projects to thrive in exploring this almost pristine and fruitful database.

Partial migration is the commonest migratory behaviour of birds worldwide (Berthold 2001, p. 52) and in the Neotropics (Stotz *et al.* 1996, p. 76; Dingle 2008). This migratory behaviour occurs when part of a population migrates to winter away from the breeding grounds while some individuals remain residing in that area throughout the year (Berthold 2001). Although studying partial migration is very promising for advances in bird migration theory (Jahn *et al.* 2006), this is a largely understudied behaviour especially because its confirmation usually depends on a year-long evaluation of population density (Stotz *et al.* 1996, p. 76; Berthold 2001, p. 52; Dingle 2008). A closer look at our results suggests that three model species – the Vermilion Flycatcher, the Grey-breasted Martin, and the Blue-black Grassquit – have occurrence record distributional patterns consistent with a partial migratory behaviour. Partial migration behaviour and breeding records of the Vermilion Flycatcher are novel findings for the species in the studied region where, to date, it has been described only as a wintering austral migrant (Stotz *et al.* 1996; Sick 1997). A closer inspection of WikiAves data for this species revealed a nesting activity never reported in previous studies for the study region (e.g. http://www.wikiaves.com.br/516635). Nevertheless, such a pattern may also result from resident populations in the study region receiving wintering austral migratory individuals. This population substitution behaviour has already been observed in austral populations of the nominal subspecies and in *P. r. saturatus* which inhabits northern Brazil (Sick 1997). Thus, it would be worth testing the alternative hypothesis that this pattern arises from chain or leapfrog migration (Berthold 2001, p. 67), i.e. a local breeding population being replaced by austral migrants during the winter, a task that is also unlikely to be accomplished without fieldwork collection of individual-level data in different populations.

Our findings for the Blue-black Grassquit add the study region within those where the species is already referred to as a partial migrant, the central Brazil area and northern region of Rio Grande do Sul state in southern Brazil (Negret and Negret 1981; Sick 1997; Bencke and Kindel 1999). These data also highlight that future review studies must update Somenzari *et al.* (2018), including the Blue-black Grassquit in the Brazilian list of migratory species. Our data show that the Creamy-bellied Thrush occurs in the study region throughout the year, suggesting a residency behaviour that contrasts to the obligatory (Stotz *et al.* 1996) or partially migratory (Somenzari *et al.* 2018 and references therein) status described for the species in the literature. A similar finding applies to the Grey-breasted Martin, often stated as an austral migrant (e.g. Sick 1997; Bencke and Kindel 1999), but also likely a partial rather than an obligatory migrant in the study region. Nevertheless, and similar to the Vermilion Flycatcher, both species may show alternative obligatory migratory patterns in the region such as chain or leapfrog migration, thus deserving further individual- and population-level field evaluations. Altogether, our findings stress the need for pioneer coordinated fieldwork studies on the annual variation in population densities of the above-mentioned species in addition to a deeper exploration of citizen-science databases. Specifically, researchers should look at likely between-population variations in the migratory behaviour of these species, either longitudinally or latitudinally.

Our analytical approach to species' data compiled from the WikiAves and eBird platforms adequately discriminated between residents and migrants, with converging analytical results from data gathered in both platforms. With that, and by providing a protocol that eases data download, our approach certifies the use of WikiAves as an unprecedented and largely underexplored citizen-science database for bird migration studies. Data sampling has the essential premise that samples truly represent the statistical population from which they were extracted (Krebs 2014). When sampling is made by citizen-scientists, protocols are often less standardised than a scientist would likely employ. Therefore, the validation of database quality in relation to the needs of given research is essential prior to any analysis (Dickinson *et al.* 2010 and references therein). With the proper above-mentioned data sampling and treatment techniques we expect future studies to increase the precision of country-wide data on the behaviour of migratory species in Brazil, thus ensuring solid country-wide advances in bird migration studies and, considering the geographical range of Brazil, in South America as a whole.

We also expect our results to reinforce the importance of citizen-science initiatives, increasing the participation of citizens in such still-incipient platforms, projects and initiatives for enabling studies in regions where the contribution of citizen-science is still modest. For instance, as of July 2018, the eight southernmost Brazilian states (Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Mato Grosso, Rio de Janeiro, Minas Gerais and Espírito Santo) held 1 700 788 records, whereas there were only 244 313 records within a similar area comprising the 10 northernmost states (~1 800 000 km^2; Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia and Tocantins). Such a 7:1 ratio is similar when the southernmost states were compared to those seven comprising the Amazon Forest domain (242 530 records; ~4 480 000 km^2: Acre, Amazonas, Roraima, Rondônia, Amapá, Pará and Mato Grosso). Through revealing the value of WikiAves data, we expect to draw researchers' and platform administrators' attention to the need to stimulate further citizens' engagement in recording species in those low-sampled areas as well.

Current theories on bird migration are based largely on the data of a few species (Faaborg *et al.* 2010b), among which intratropical migrants are historically the least studied. Therefore, we urge researchers to explore the valuable WikiAves platform data, following the lead of other ornithological citizen-science databases (e.g. eBird). In this way we expect advances in at least some basic questions such as the identification of which bird species indeed migrate, and where and when they do so, thus allowing continuous updating and refinement of the reviewed list of migratory bird species that occur in Brazil (Somenzari *et al.* 2018). In parallel, we hope for the dawn of a new era of studies on bird migration ecology, which should incorporate a more representative universe of South American species, ultimately leading to papers revisiting at least some of the general bird migration theories. Moreover, such studies should be able to consider data for a geographic range largely neglected in previous studies (e.g. lack of data for Brazil, the Guianas, and Suriname in Figure 1 in La Sorte *et al.* 2016), but which should be revealed soon as hosts of a number of regional and intratropical migrants (Somenzari *et al.* 2018). Our study's approach allows use of the WikiAves database to identify which species show migratory behaviour in Brazil, thus optimising and complementing efforts in acquiring data in the field and museums, plus the use of scarce funding resources, to deepen our knowledge of the ecology of those migrants. For instance, WikiAves data analysis may reveal the wintering grounds of migrants in Brazil. Despite being the most critical period for the conservation of migratory birds, this is a time in which ecological knowledge of most species is far below minimum (Faaborg *et al.* 2010a). In the ongoing 'big data' era, ornithologists worldwide should not resist extracting the best from every existing database. This seems to us a plausible strategy to avoid the problem that the development of bird migration theories continues to be guided by the ecological and evolutionary aspects of only a handful of northern hemisphere long-distance migratory species.

Acknowledgements

The authors wish to thank all citizen-scientists who have contributed to WikiAves and its founder Reinaldo Guedes and his crew. We are indebted to Drs Isabela G. Varassin, Rosana M. Rocha and Kate Buchanan, and two anonymous reviewers for their invaluable comments and advice. This study results from S.C.S.'s undergraduate thesis.

Funding

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under fellowship 1459754 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [455908/2014-1].

ORCID

Stephanie Caroline Schubert http://orcid.org/0000-0003-0068-8573
Lilian Tonelli Manica http://orcid.org/0000-0001-6005-7103
André De Camargo Guaraldo http://orcid.org/0000-0003-1705-2926

References

Agostinelli, C., and Lund, U. (2013) R Package Circular: Circular Statistics (version 0.4-3). https://r-forge.r-project.org/projects/circular

Alves, M. A. (2007). Sistemas de migrações de aves em ambientes terrestres no Brasil: Exemplos, lacunas e propostas para o avanço do conhecimento. *Revista Brasileira de Ornitologia* **15**(2), 231–238.

Anjos, L., and Graf, V. (1993). Riqueza de aves da Fazenda Santa Rita, região dos campos gerais, Palmeira, Paraná, Brasil. *Revista Brasileira de Zoologia* **10**(4), 673–693. doi:10.1590/S0101-81751993000400013

Audubon Society (2017) 'Christmas Bird Count.' Available at http://www.audubon.org/conservation/science/christmas-bird-count [Verified 13 Jul 2017].

Baillie, S. R., Balmer, D. E., Downie, I. S., and Wright, K. H. M. (2006). Migration watch: an Internet survey to monitor spring migration in Britain and Ireland. *Journal of Ornithology* **147**(2), 254–259. doi:10.1007/s10336-006-0062-8

Bencke, G. A., and Kindel, A. (1999). Bird counts along an altitudinal gradient of Atlantic forest in northeastern Rio Grande do Sul, Brazil. *Revista Brasileira De Ornitologia-Brazilian Journal of Ornithology* **7**(11), 91–107.

Berthold, P. (2001). '*Bird Migration: A General Survey*,' 2nd ed. (Oxford University Press: New York.)

Biancalana, R. N., Nogueira, W., Bessa, R., Pioli, D., and Lees, A. (2013). Range extensions and breeding biology observations of the Sooty Swift (Cypseloides fumigatus) in the states of Bahia, Goiás, Minas Gerais and Tocantins. *Revista Brasileira De Ornitologia-Brazilian Journal of Ornithology* **20**(48), 87–92.

BirdLife International. (2018) 'Country Profile: Brazil.' Available at http://www.birdlife.org/datazone/countrybrazil [Verified 2 March 2018].

Bonter, D. N., and Cooper, C. B. (2012). Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* **10**(6), 305–307. doi:10.1890/110273

Brasil. (1998) Lei 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Diário Oficial [da República Federativa do Brasil], Brasília, DF: 3.

Brasil. (2013) Lei 12.853, de 14 de agosto de 2013. Altera os arts. 5º, 68, 97, 98, 99 e 100, acrescenta arts. 98-A, 98-B, 98-C, 99-A, 99-B, 100-A, 100-B e 109-A e revoga o art. 94 da Lei no. 9.610, de 19 de fevereiro de 1998, para dispor sobre a gestão coletiva de direitos autorais, e dá outras providências. Diário Oficial da União, Brasília, DF: 1.

Cooper, C. B., Shirk, J., and Zuckerberg, B. (2014). The invisible prevalence of citizen science in global research: migratory birds and climate change. *PloS One* **9**(9), e106508. doi:10.1371/journal.pone.0106508

Dickinson, J. L., Zuckerberg, B., and Bonter, D. N. (2010). Citizen science as an ecological research tool: challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics* **41**(1), 149–172. doi:10.1146/annurev-ecolsys-102209-144636

Dingle, H. (2008). Bird migration in the Southern Hemisphere: a review comparing continents. *EMU* **108**(4), 341–359. doi:10.1071/MU08010

European Bird Census Council (2017) 'Euro Bird Portal.' Available at http://www.eurobirdportal.org/ebp/en/ [Verified 13 July 2017].

Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux, S. A., Heglund, P., *et al.* (2010a). Conserving migratory land birds in the New World: do we know enough? *Ecological Applications* **20**(2), 398–418. doi:10.1890/09-0397.1

Faaborg, J., Holmes, R. T., Anders, A. D., Bildstein, K. L., Dugger, K. M., Gauthreaux, S. A., Heglund, P., *et al.* (2010b). Recent advances in understanding migration systems of New World land birds. *Ecological Monographs* **80**(1), 3–48. doi:10.1890/09-0395.1

Feinerer, I., and Hornik, K. (2015). tm: Text mining package. R package version 0.7–6. https://CRAN.R-project.org/package=tm

Galina, A. B., and Gimenes, M. R. (2006). Riqueza, composição e distribuição espacial da comunidade de aves em um fragmento florestal urbano em Maringá, Norte do Estado do Paraná, Brasil. *Acta Scientiarum. Biological Sciences* **28**(4), 379–388.

Godoi, M. N., Costacurta, M. D. B., Nunes, A. P., Patrial, E. W., and Morante Filho, J. C. (2012). First records of the Crested Black-Tyrant (Knipolegus lophotes, Tyrannidae) in the State of Mato Grosso do Sul, Brazil. *Biota Neotropica* **12**, 311–314. doi:10.1590/S1676-06032012000300030

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., *et al.* (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography* **24**(3), 276–292. doi:10.1111/geb.2015.24.issue-3

Hampton, S., Strasser, C., Tewksbury, J., Gram, W., Budden, A., Batcheller, A., Duke, C., *et al.* (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment* **11**(3), 156–162. doi:10.1890/120103

IBGE, I.B.d.G.E. (2018) 'Área Territorial Brasileira.' Available at http://www.ibge.gov.br/home/geociencias/cartografia/default_territ_area.shtm [Verified 2 Mar 2018].

Isaac, N. J. B., van Strien, A. J., August, T. A., de Zeeuw, M. P., Roy, D. B., and Anderson, B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* **5**(10), 1052–1060. doi:10.1111/2041-210X.12254

Jahn, A. E., Levey, D. J., Johnson, J. E., Mamani, A. M., and Davis, S. E. (2006). Towards a mechanistic interpretation of bird migration in South America. *El Hornero* **21**(2), 99–108.

Kelling, S., Johnston, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., Bonn, A., Fernandez, M., Hochachka, W., Julliard, R., Kraemer, R., *et al.* (2019). Using semistructured surveys to improve citizen science data for monitoring biodiversity. *BioScience* **69**(3), 170–179.

Kelly, J. F., Horton, K. G., Stepanian, P. M., Beurs, K. M., Fagin, T., Bridge, E. S., and Chilson, P. B. (2016). Novel measures of continental-scale avian migration phenology related to proximate environmental cues. *Ecosphere* **7**(9). doi:10.1002/ecs2.1434

Krebs, C. J. (2014) 'Ecological methodology.' Available at http://www.zoology.ubc.ca/~krebs/books.html [Verified 20 Mar 2018].

Krügel, M. M., and Anjos, L. (2000). Bird communities in forest remnants in the city of Maringá, Paraná State, southern Brazil. *Ornitologia Neotropical* **11**(4), 315–330.

La Sorte, F. A., Fink, D., Hochachka, W. M., and Kelling, S. (2016). Convergence of broad-scale migration strategies in terrestrial birds. *Proceedings of the Royal Society B: Biological Sciences* **283**(1823), 1–9. doi:10.1098/rspb.2015.2588

La Sorte, F. A., Lepczyk, C. A., Burnett, J. L., Hurlbert, A. H., Tingley, M. W., and Zuckerberg, B. (2018). Opportunities and challenges for big data ornithology. *The Condor* **120**(2), 414–426. doi:10.1650/CONDOR-17-206.1

Lees, A. C. (2016). Evidence for longitudinal migration by a 'sedentary' Brazilian flycatcher, the ash-throated Casiornis. *Journal of Field Ornithology* **87**(3), 251–259. doi:10.1111/jofo.2016.87.issue-3

Lees, A. C., and Martin, R. W. (2014). Exposing hidden endemism in a neotropical forest raptor using citizen science. *IBIS* **157**(1), 103–114. doi:10.1111/ibi.12207

Lopes, E. V., and Anjos, L. D. (2006). The avifaunal composition of Universidade Estadual de Londrina, northern Paraná, Brazil. *Revista Brasileira de Zoologia* **23**(1), 145–156. doi:10.1590/S0101-81752006000100006

McCaffrey, R. E. (2005). Using citizen science in urban bird studies. *Urban Habitats* **3**(1), 70–86.

Negret, A. J., and Negret, R. A. (1981). 'As Aves Migratórias Do Distrito Federal,' p. 64. (Ministério da Agricultura, Instituto Brasileiro de Desenvolvimento Florestal: Brasília.)

Nourani, E., Yamaguchi, N. M., and Higuchi, H. (2017). Climate change alters the optimal wind-dependent flight routes of an avian migrant. *Proceedings of the Royal Society B: Biological Sciences* **284**(1854), 1–6. doi:10.1098/rspb.2017.0149

Pearce-Higgins, J. W., Eglington, S. M., Martay, B., and Chamberlain, D. E. (2015). Drivers of climate change impacts on bird communities. *The Journal of Animal Ecology* **84**(4), 943–954. doi:10.1111/1365-2656.12364

Plataforma Lattes (2017) 'Plataforma Lattes.' Available at http://lattes.cnpq.br/ [Verified 15 March 2018].

R Core Team. (2018). 'R: A Language and Environment for Statistical Computing.' (R Foundation for Statistical Computing: Vienna, Austria.)

Remsen, J. V., Jr, Cadena, C. D., Jaramillo, A., Nores, M., Pacheco, J. F., Robbins, M. B., Schulenberg, T. S., *et al.* (2017) 'A classification of the bird species of South America.' Available at http://www.museum.lsu.edu/~Remsen/SACCBaseline.htm [Verified 22 April 2017].

Scherer Neto, P., and Bispo, A. A. (2011). Avifauna do Parque Estadual de Vila Rica do Espírito Santo, Fênix, Paraná. *Biota Neotropica* **11**(3), 317–329. doi:10.1590/S1676-06032011000300026

Sick, H. (1997). '*Ornitologia Brasileira*,' p. 912. (Nova Fronteira: Rio de Janeiro, Brasil.)

Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution* **24**(9), 467–471. doi:10.1016/j.tree.2009.03.017

Somenzari, M., Amaral, P. P. D., Cueto, V., Guaraldo, A. C., Jahn, A. E., Lima, D. M., Lima, P. C., *et al.* (2018). A review of Brazilian migratory birds. *Papéis Avulsos de Zoologia (online)* **58**, 1–66.

Stotz, D. F., Fitzpatrick, J. W., Parker, T. A., III, and Moskovits, D. K. (1996). '*Neotropical Birds: Ecology and Conservation*,' p. 478. (The University of Chicago Press: Chicago, MI, USA.)

Tellería, J. L., Ramírez, Á., and Aguirre, J. I. (2015). Are European birds leaving traditional wintering grounds in the Mediterranean? *Journal of Avian Biology* **46**(3), 1–7. doi:10.1111/jav.00588

Vogel, H. F., Metri, R., Zawadzki, C. H., and Moura, M. O. (2011). Avifauna from a campus of Universidade Estadual do Centro-Oeste, Guarapuava, Paraná State, Brazil. *Acta Scientiarum, Biological Sciences* **33**(2), 197–207. doi:10.4025/actascibiolsci.v33i2.7710

Wallraff, H. G. (1979). Goal-oriented and compass-oriented movements of displaced homing pigeons after confinement in differentially shielded aviaries. *Behavioral Ecology and Sociobiology* **5**(2), 201–225. doi:10.1007/BF00293306

Wickham, H. (2009). 'Ggplot2: Elegant Graphics for Data Analysis.' (Springer-Verlag: New York.)

Wiggins, A., Newman, G., Stevenson, R. D., and Crowston, K. (2011) 'Mechanisms for data quality and validation in citizen science.' (IEEE Seventh International Conference on e-Science Workshops, Stockholm, Sweden) 14–19. doi:10.1177/1753193411414352

WikiAves (2017) 'WikiAves – A enciclopédia das aves do Brasil.' Available at www.wikiaves.com.br [Verified 9 July 2017].

Wood, S. N. (2017). 'Generalized Additive Models: An Introduction with R,' 2nd ed. (Boca Raton, Florida: Chapman and Hall/CRC.)

Xiao, H., Hu, Y., Lang, Z., Fang, B., Guo, W., Zhang, Q., Pan, X., and Lu, X. (2016). How much do we know about the breeding biology of bird species in the world? *Journal of Avian Biology* **48**(4), 513–518. doi: 10.1111/jav.00934

Zar, J. H. (2010a). Circular distributions: descriptive statistics. In '*Biostatistical Analysis*.' (Ed. J. H. Zar.) pp. 605–623. (Prentice-Hall International: London, England.)

Zar, J. H. (2010b). Circular distributions: hypothesis testing. In '*Biostatistical Analysis*.' (Ed. J. H. Zar.) pp. 624–668. (Prentice-Hall International: London, England.)