Taylor & Francis
Taylor & Francis Group

Ringing & Migration

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tram20

The long and short of it: converting between maximum and minimum tarsus measurements in passerine birds

Anthony Caravaggi, Sam Bayley, Richard J. Facey, Iván de la Hera, Mike P. Shewring & Jez A. Smith

To cite this article: Anthony Caravaggi, Sam Bayley, Richard J. Facey, Iván de la Hera, Mike P. Shewring & Jez A. Smith (2022): The long and short of it: converting between maximum and minimum tarsus measurements in passerine birds, Ringing & Migration, DOI: 10.1080/03078698.2022.2050937

To link to this article: https://doi.org/10.1080/03078698.2022.2050937

Published online: 21 Apr 2022.

RINGING & MIGRATION
https://doi.org/10.1080/03078698.2022.2050937

The long and short of it: converting between maximum and minimum tarsus measurements in passerine birds

Anthony Caravaggi [a], Sam Bayley[b], Richard J. Facey [c], Iván de la Hera[d], Mike P. Shewring[c] and Jez A. Smith[a,c]

[a]School of Applied Sciences, University of South Wales, 9 Graig Fach, Pontypridd, CF37 4BB; [b]Siskin, Keelnacronagh West, Enniskeane, County Cork, Ireland, P47 NP90; [c]School of Biosciences, Cardiff University, The Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX; [d]Txepetxa Ringing Group, Vicente Aleixandre 10, 01003 Vitoria-Gasteiz, Spain

ABSTRACT
Tarsus length, minimum or maximum, is a commonly recorded metric used in a variety of ornithological studies. The nature of the relationships between minimum and maximum tarsus lengths and the accuracy with which one can be derived from the other is unknown. We explored the strength of the relationships between tarsus measurements across six species of passerines, deriving species-specific intercept equations and a minimum sample size (n = 49) required to support the accurate transformation of data for other species. The effective bidirectional transformation of measurements indicates that our method has broad applicability and utility and can be used to facilitate comparative studies, syntheses and collaborations.

ARTICLE HISTORY
Received 25 September 2021
Accepted 11 January 2022

Tarsus length is an important metric collected by researchers and bird ringers/banders and is used solely or in combination with other biometrics in a variety of studies. For example, tarsus measurements form the basis of indices of, or act as proxies for, body condition (e.g. Ormerod & Tyler 1990, Vafidis *et al* 2016, Wang *et al* 2019), and skeletal body size (e.g. Senar & Pascual 1997). Tarsus measurements are also commonly used to examine intraspecific growth rate (e.g. Cleasby *et al* 2011), and derive sexing criteria (e.g. Andersson & Wester 1971, Giammarino *et al* 2012, Hallgrimsson *et al* 2016).

The two most common methods used for measuring passerine tarsi in living birds (as opposed to museum specimens, see below) are: (i) from the inter-tarsal joint to the joint between tarsus and toes, with the toes bent at approximately 90° in relation to the tarsus ('minimum tarsus' hereafter; e.g. Richner 1989, Taborsky & Taborsky 1999, Eck *et al* 2011, Zuberogoitia *et al* 2011); and (ii) from the lower hind edge of the tibia to the tarsus/toes joint ('maximum tarsus' hereafter; e.g. Dougall 1997, Redfern & Clark 2001, Masello & Quillfeldt 2004, Li *et al* 2010; Figure 1). It is worth noting at this juncture that two of the most commonly used European field guides, Svensson (1992) and Demongin (2016), refer only to minimum tarsus but provide two measurement methods – the 'standard' and the 'alternative'. The 'standard' measure is defined as "*the notch on the back of the intertarsal joint to the lower edge of the last complete scale before the toes diverge*". This method is consistent with museum practices. The 'alternative' measure is more repeatable in the field, hence the definition given, above.

There is no global consensus, however, on a standard method for best practice; hence national ringing schemes and other authorities differ in their recommendations. For example, the European–African songbird migration network (Bairlein *et al* 1995), and the South African Ringing Scheme (SAFRING) (de Beer *et al* 2001) suggest that recorders should use the minimum tarsus method. In contrast, the British Trust for Ornithology (BTO) advocate the use of maximum tarsus (Redfern & Clark 2001). This picture is further complicated by the fact that historically the BTO recommended the minimum tarsus method (Redfern & Clark 2001).

Despite their widespread use, there is a lack of published information on intraspecific relationships between minimum and maximum tarsus measurements. Any assumptions of equivalency in development and bone ossification over time that suggest that the relative difference between minimum and maximum tarsus measurements would be conserved and consistent across species may be misplaced. For example, individuals within a species do not develop equally, despite being from the same

CONTACT Anthony Caravaggi ar.caravaggi@gmail.com

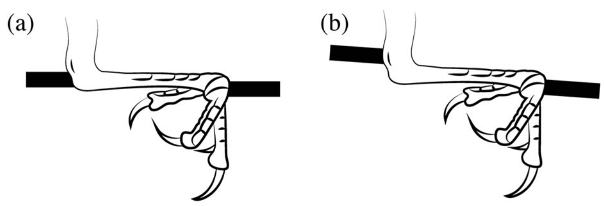

Figure 1. Methods for deriving (a) minimum and (b) maximum tarsus measurements. Artwork by Kim Caravaggi.

general genetic blueprint (e.g. Björklund 1997, Merilä & Fry 1998, Kunz & Ekman 2000). Thus, there is the potential for considerable non-linear, intraspecific variation in tarsal length and tibial width (Quisenberry *et al* 1941, Ren *et al* 2016, DeSesso & Scialli 2018). Therefore, projects and schemes that study the same species but use different methods of measuring tarsus length (e.g. Goodenough *et al* 2008 *vs* Norte *et al* 2009) may have limited scope for comparison, inferential extrapolation, or contributions to future data syntheses (e.g. McKechnie 2019, Weeks *et al* 2020).

This study describes the relationship between minimum and maximum tarsus measurements in living birds across six passerine species in Europe, derives correction factors that can be applied to existing data bilaterally to convert between minimum and maximum tarsus measures with a high degree of accuracy, and recommends a minimum sample size for determining correction factors for other species.

Methods

Location and biometric methods

Field surveys were undertaken at various sites in the United Kingdom, Ireland and Spain as part of normal, licensed bird-ringing activities using a variety of mist nets and conventional traps, between February 2018 and July 2019, inclusive. Before beginning data collection, a power analysis was undertaken that suggested that a minimum of 16 samples were required to detect a correlative relationship of $r \geq 0.8$, with 95% power at $P = 0.01$. For each individual bird captured, age, sex, wing length, body mass and minimum and maximum tarsus length were recorded. Tarsus measurements were recorded, on the right leg, to the nearest 0.1 mm, using analogue callipers (Figure 1). The leg was temporarily released between measurements, thus 'resetting' both the leg and the observer and ensuring that the same process of leg and calliper manipulation was carried out for each measurement. Data were collected by all the authors, with the two tarsus measurements of any individual bird always being recorded by the same person.

Data analysis

Prior to analysis, outliers were categorised as those values that were found to lie above or below 1.5 times the interquartile range (Tukey 1977) and subsequently removed. For each species we calculated Pearson's correlation coefficient r for the relationship between the minimum and maximum tarsus measurements. The intercept equation facilitating the calculation of y (in the first instance) for a given value of x was subsequently calculated using the formula $y = mx + a$, where y indicates the y variable (e.g. minimum tarsus), and m represents the slope of the correlation. The unknown variable a was calculated using the formula $\underline{y} = m\underline{x}+a$. The slope of the correlation, m, was calculated as $m = r\frac{SDy}{SDx}$, where SD describes the standard deviation, and x indicates the x variable (e.g. maximum tarsus). Descriptive statistics for minimum and maximum tarsus measurements (the minimum, maximum and mean of each) derived from raw data were then compared to those derived from intercept equations.

Sequential bootstrapping and changepoint analyses were used to calculate the minimum number of observations required to support transformation between minimum and maximum tarsus measurements. Changepoint analyses aim to identify locations at which the statistical properties of a sequence of observations differs along its length. For example, given the sequence A–G, changepoint analysis may highlight D, i.e. the statistical properties of A–C differs substantially to the properties of E–G. The correlation coefficient for each species was calculated across 12 sampling instances, corresponding to the number of visits required during the BTO Constant Effort Survey (CES) methodology (BTO 2020), using groups of between one and 100 randomly sampled individuals, without replacement, thus simulating the uncertainty of actual site visits. The CES method provides a standardised baseline that placed the present study in an appropriately realistic context, as opposed to an arbitrary number of sampling instances, such as five or 10. Changepoint analysis was then applied using the power of the pruned exact linear time (PELT) method, with a manual penalty of $2 \times \log(n)$ (Wambui *et al* 2015), estimating the point or points at which the observed correlation coefficients before and after differed significantly (Killick & Eckley 2014, Killick *et al* 2016).

Due to the nature of changepoint analysis, a sequence of observations can contain several changepoints. Here, we apply the precautionary principle and interpret the maximum changepoint as indicating the minimum number of birds required to convert reliably between the different tarsus measurements. This process was repeated 1000 times, equating to a total of 12 000 randomly sampled iterations, from which the mean ± SD maximum changepoint was calculated.

Finally, we calculated the mean ± SD Pearson's correlation coefficient for each species, across 1000 iterations, based on random sampling of the original data, without replacement, where n was equal to (i) the mean changepoint, and (ii) the minimum number of observations required according to power analysis (n = 16).

All analyses were carried out and plots created in R v3.5.1 (R Core Team 2018), using the packages *changepoint* (Killick & Eckley 2014, Killick *et al* 2016), *tidyr* (Wickham & Henry 2019), *dplyr* (Wickham *et al* 2019) and *ggplot2* (Wickham 2016).

Results

A total 1238 individuals of 28 passerine species were caught and measured during the study. We focused on the six most commonly encountered species (n > 100), specifically Blue Tit *Cyanistes caeruleus* (n = 220; 217 minus outliers), Blackbird *Turdus merula* (n = 119; 113), Coal Tit *Periparus ater* (n = 106; 104), Chaffinch *Fringilla coelebs* (n = 191; 187), Great Tit *Parus major* (n = 307; 303) and Robin *Erithacus rubecula* (n = 108; 107). All results from these and the other species from which data were recorded can be found in the Appendix tables and, along with associated R code from this paper, in the online data repository at https://doi.org/10.5281/zenodo.6347026.

Table 1. Minimum and maximum tarsus measurements (mm) for six species of passerine birds. Actual (*a*) metrics derived from raw data and estimates (*e*) calculated from correlation slope intercept equations are given.

	Actual			Estimates		
Species	a_{min}	a_{max}	a_{mean} (± SD)	e_{min}	e_{max}	e_{mean} (± SD)
Minimum tarsus measurement						
Blackbird	31.9	36.7	34.3 ± 1.1	31.68	36.82	34.31 ± 1.22
Blue Tit	15.4	18.7	16.8 ± 0.6	15.01	18.29	16.74 ± 0.50
Chaffinch	16.7	19.9	18.3 ± 0.6	16.64	20.04	18.34 ± 0.62
Coal Tit	16.3	18.3	17.3 ± 0.5	15.67	18.38	17.25 ± 0.56
Great Tit	18.7	22.7	20.5 ± 0.7	16.22	22.69	20.49 ± 1.08
Robin	23.8	27.8	25.9 ± 0.9	23.55	27.99	25.87 ± 0.83
Maximum tarsus measurement						
Blackbird	36.6	41.5	39.1 ± 1.1	36.81	41.38	39.09 ± 1.22
Blue Tit	17.4	20.8	19.2 ± 0.7	17.81	21.23	19.26 ± 0.50
Chaffinch	19.2	22.8	21.0 ± 0.7	19.26	22.65	20.96 ± 0.62
Coal Tit	18.0	20.4	19.4 ± 0.5	18.56	20.33	19.45 ± 0.56
Great Tit	19.6	25.2	23.3 ± 0.8	21.75	25.21	23.31 ± 1.08
Robin	26.5	30.7	28.7 ± 0.9	26.74	30.52	28.73 ± 0.83

Tarsus measurements varied between species, from a minimum of 16.8 ± 0.6 mm and maximum of 19.2 ± 0.7 in Blue Tits to 34.2 ± 1.1 mm and 39.0 ± 1.1 mm in Blackbirds (Table 1). There were no differences in tarsus length between sexes or age classes in any of the focal species (Tables A1–A8). All focal species exhibited high correlations ($r > 0.7$; Taylor 1990) between minimum and maximum tarsus measurements. The strongest correlation was observed in Robins ($r = 0.93$), while Coal Tits had the lowest correlation ($r = 0.78$). There were no significant differences between raw tarsus measurements and those derived from intercept equations. Mean minimum and maximum tarsus measurements deviated by 0.01 mm (Blackbird in both measures) to 0.06 mm (Blue Tit and Chaffinch; Table 1). Correlation values between estimated minimum and actual maximum tarsus measurements and vice versa matched those of correlations using original data (see Figure 2, which also gives species-specific transformation equations).

Mean maximum changepoints calculated from iterative bootstrapping were also similar across all focal species, from 43 (Coal Tit) to 49 (Great Tit), with a mean of 45 (Figure 2). Following our prior rationale where the maximum changepoint was assumed to represent the minimum number of birds required for each species, the minimum number required to repeat this study reliably with other focal species was 49. Comparisons between r values derived from species-specific changepoints and those based on power analysis showed that using changepoints resulted in less dispersed data. For example, Chaffinch changepoint data had narrower errors (0.06) and a smaller interquartile range (0.07) than power analysis data for the same species (0.13 and 0.19, respectively; Figure 3).

Discussion

Measurements of tarsus length are a commonly recorded metric in passerine studies. These measurements, either minimum or maximum, are used to infer a variety of other individual characteristics and traits. Bootstrap analyses suggested that a minimum of 49 individual birds are required, per focal species, to support transformation while observing a strong correlation between metrics. This study is the first to explore relationships between minimum and maximum tarsus measurements in passerine birds, thereby facilitating intraspecific comparison between and synthesis across studies. This study was based on the observation that two

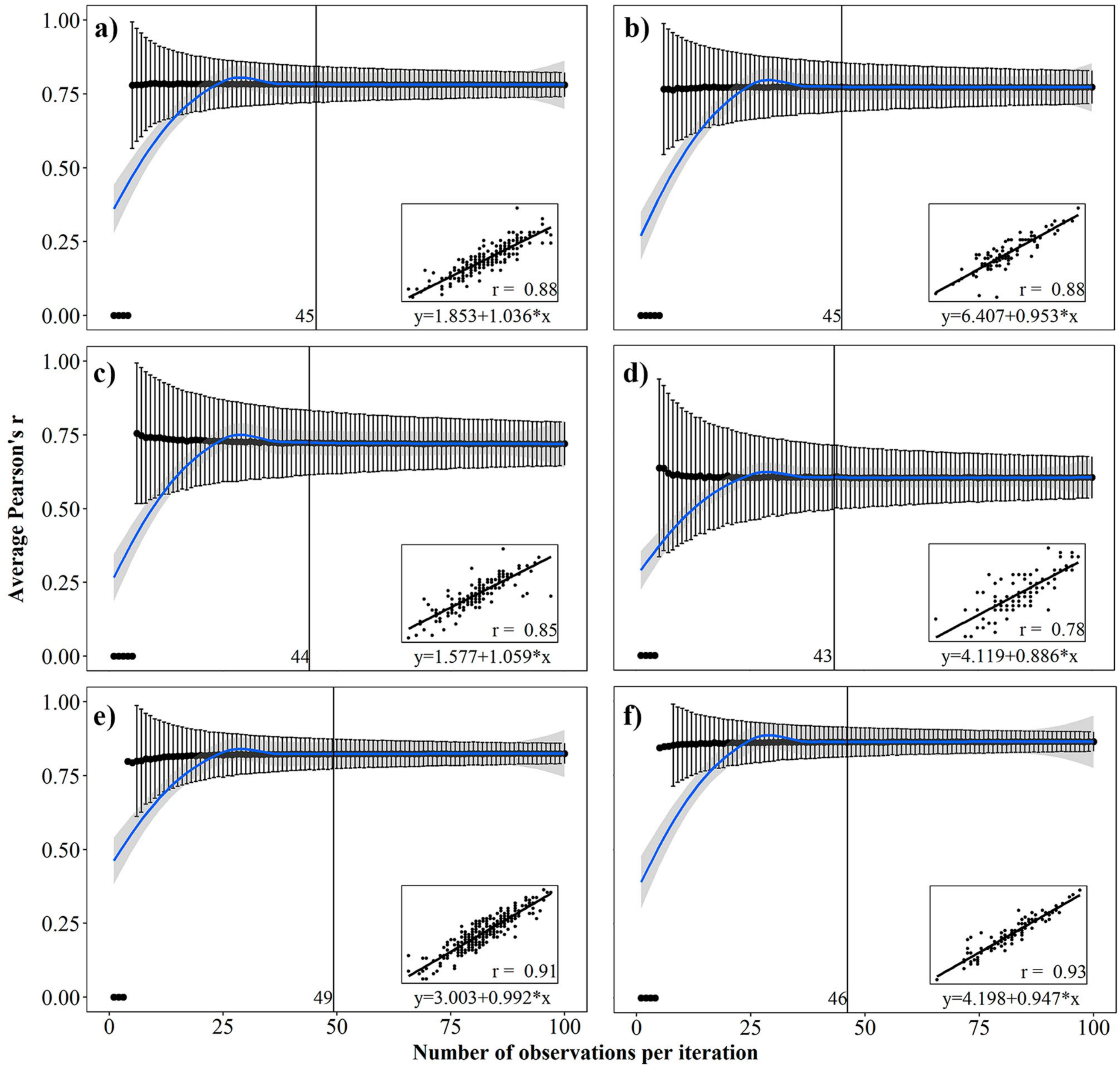

Figure 2. Bootstrapped Pearson correlations between minimum and maximum tarsus measurements in six passerine species: (a) Blue Tit; (b) Blackbird; (c) Chaffinch; (d) Coal Tit; (e) Great Tit; (f) Robin. Vertical lines and associated integers represent the mean maximum changepoint representing a significant difference between preceding and succeeding *r* values. Changepoints were calculated every 12 iterations. Inset plots show the correlations across all data for each species.

independent and geographically distant research groups, here anonymised, were collecting a large amount of data on the same focal species to answer very similar questions. However, each group was using a different method of measuring tarsus length, thereby limiting the potential for comparison and collaboration. Here, we have demonstrated that minimum and maximum tarsus measurements are directly related and that it is possible to convert reliably between the two metrics, given sufficient data. There was a high degree of correlation ($r > 0.7$) between minimum and maximum tarsus metrics across all six species included in the current study. Derived intercept equations supported transformation between tarsus metrics with no changes in correlation values and only small differences in actual values; e.g. for Robin the mean actual minimum tarsus was 25.9 ± 0.9 mm and the estimated value 25.87 ± 0.83 mm.

It was not possible to quantify interobserver variability in this study due to the distribution of the study sites and to the locations and relative availability of observers. However, the current study does not aim to describe the 'true' lengths of individual tarsi under different measuring systems, but the relationships

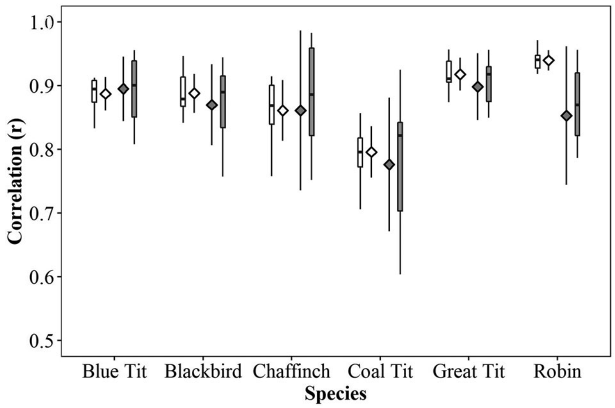

Figure 3. Pearson's correlations between minimum and maximum tarsus measurements in six passerine species, across 1000 iterations. The numbers of observations included in each iteration are 49 (see Results; open symbols) and 16 (the minimum number of observations required according to *a priori* power analysis; shaded). For each group, a box-and-whisker plot is given and separately the mean ± SD (diamonds).

between tarsus measurements within species. Further, intra-observer variability was not quantified due to welfare considerations associated with pre-processing retention and extended handling times, particularly during cold weather. However, studies have suggested that intra-observer error in experienced practitioners is typically low, including for linear measurements (Bulakhova *et al* 2011, Goodenough *et al* 2012). Hence, while the measurements described herein are likely to be subject to interobserver variability, individual observers may be consistent within themselves in their measurements of minimum and maximum tarsus for each individual and species (e.g. Broughton & Clark 2017; though note that substantial errors are still possible). Nevertheless, we strongly emphasise the importance of collecting reliable data via training, individual consistency checks, and verification by comparison with other ringers. Our results should, therefore, be taken as conservative estimates.

Synthesising data from contemporary and historical data sets, e.g. for comparative or meta-analyses, is made more difficult when the specific tarsus measurement method (minimum or maximum) is omitted from publications (e.g. Haywood & Perrins 1992, Riddington & Gosler 1995, Poissant *et al* 2016a). While it would appear reasonable to treat smaller values as minimum tarsus lengths and larger ones as maximum tarsus lengths, the current study shows that there is overlap between the two metrics across a population. As a result, data available for syntheses and comparisons may be subject to considerable truncation. For example, the minimum of the maximum tarsus length in Great Tits was 19.6 mm while the maximum of the minimum length was 22.7 mm, giving an overlap of 3.1 mm between metrics. Applying these thresholds to Great Tit data from Wytham Woods, where the method of tarsus measurement was not specified (Poissant *et al* 2016b; n = 2081), results in the necessary omission of 298 measurements, or 14.3% of the available data. This issue may be compounded by the rounding of tarsus measurements (e.g. to 0.5 mm, Poissant *et al* 2016a) that obfuscates fine-scale variation. This could result in the loss of important information relevant to species-specific conservation and management, particularly as some passerine species show morphological changes in response to climate change (Weeks *et al* 2020). We therefore encourage researchers and bird ringers to record and describe the specific method of tarsus measurement used in their work, thereby facilitating further study.

From a purely statistical point of view, neither minimum nor maximum tarsus measurements are inherently better than the other in studies of live birds (for discussions of specimen shrinkage inmuseum collections, see Bjordal (1983), Jenni & Winkler (1989) and Winker (1993)). Nevertheless, we have demonstrated that, with a sufficiently large data set, it is possible to convert between metrics with minimal error, thereby facilitating synthesis and comparative study. Hence, the choice of which method to use is arguably a matter of personal preference. Given the additional difficulties inherent in locating the intertarsal joint, and in smaller species in particular, however, we reaffirm the direction of the BTO (Redfern & Clark 2001) that ringers collecting passerine tarsus length measurements should use the maximum tarsus method.

Acknowledgements

The authors thank the Editor and reviewers, whose supportive feedback helped improve this manuscript.

Authors are sorted alphabetically, with the exception of AC who conceived, initiated and organised the project, and analysed the data. RF and JS contributed to the development of the analytical methodology. All authors contributed to the manuscript. All authors are active and experienced bird ringers licensed by the British Trust for Ornithology. AC, SB and IDLH operated under an additional National Parks and Wildlife licence in Ireland, pursuant to Section 32 of the Wildlife Act, 1976. No birds were injured during this study and appropriate welfare measures were followed to keep handling time and associated stress to a minimum.

ORCID

Anthony Caravaggi http://orcid.org/0000-0002-1763-8970
Richard Facey http://orcid.org/0000-0002-5496-2083

References

Andersson, J.S. & Wester, S.A. (1971) Length of wing, bill, and tarsus as a character of sex in the Dipper *Cinclus cinclus*. *Ornis Scandinavica* **1,** 75–79.

Bairlein, F., Jenni, L., Kaiser, A., Karlsson, L., van Noordwijk, A., Peach, W., Pilastro, A., Spina, F. & Walinder, G. (1995) *European–African Songbird Migration Network: manual of field methods*. European Science Foundation, Wilhelmshaven.

Bjordal, H. (1983) Effects of deep freezing, freeze-drying and skinning on body dimensions of House Sparrows (*Passer domesticus*). *Cinclus* **6**, 105–108.

Björklund, M. (1997) Variation in growth in the blue tit (*Parus caeruleus*). *Journal of Evolutionary Biology* **10**, 139–155.

Broughton, R.K. & Clark, J.A. (2017) Understanding potential sources of bias and error in the biometric sexing of birds. *Ringing & Migration* **32**, 79–86.

BTO (2020) *Ringing Scheme: taking part*. British Trust for Ornithology, Thetford. https://www.bto.org/our-science/projects/ringing/surveys/ces/taking-part [accessed 2 Jan 2020]

Bulakhova, N.A., Böhme, W., Clasen, A., Orlova, V.F., Carretero, M.A., Kuranova, V.N., Ljubisavljevic, K., Roitberg, E.S., Shamgunova, R.R., Zinenko, O.I. & Fokt, M. (2011) Inter-observer and intra-observer differences in measuring body length: a test in the common lizard, *Zootoca vivipara*. *Amphibia–Reptilia* **32**, 477–484.

Cleasby, I.R., Burke, T., Schroeder, J. & Nakagawa, S. (2011) Food supplements increase adult tarsus length, but not growth rate, in an island population of house sparrows (*Passer domesticus*). *BMC Research Notes* **4**, 431.

de Beer, S.J., Lockwood, G.M., Raijmakers, J.H.F.A., Raijmakers, J.M.H., Scott, W.A., Oschadleus, H.D. & Underhill, L.G. (2001) *SAFRING Bird Ringing Manual*. South African Bird Ringing Unit, University of Cape Town, Cape Town.

Demongin, L. (2016) *Identification Guide to Birds in the Hand*. Privately published, Beauregard-Vendon.

DeSesso, J.M. & Scialli, A.R. (2018) Bone development in laboratory mammals used in developmental toxicity studies. *Birth Defects Research* **110**, 1157–1187.

Dougall, T.W. (1997) Biometrics and sex ratios of Skylarks *Alauda arvensis* in winter in south-east Scotland. *Ringing & Migration* **18**, 37–49.

Eck, S., van den Elzen, R., Fiebig, J., Fiedler, W., Heynen, I., Nicolai, B., Töpfer, T., Winkler, R. & Woog, F. (2011) *Vögel vermessen*. Deutsche Ornithologen-Gesellschaft, Wilhelmshaven.

Giammarino, M., Quatto, P. & Soglia, D. (2012) Analysis of biometric and DNA data to determine the sex of Hooded Crows *Corvus cornix* in northwest Italy. *Ringing & Migration* **27**, 38–42.

Goodenough, A.E., Hart, A.G. & Elliot, S.L. (2008) Variation in offspring quality with cavity orientation in the Great Tit. *Ethology, Ecology and Evolution* **20**, 375–389.

Goodenough, A.E., Smith, A.L., Stubbs, H., Williams, R. & Hart, A.G. (2012) Observer variability in measuring animal biometrics and fluctuating asymmetry when using digital analysis of photographs. *Annales Zoologici Fennici* **49**, 81–92.

Hallgrimsson, G.T., Helgason, H.H., Palsdottir, E.S. & Palsson, S. (2016) Sexing adult and fledgling Lesser Black-backed Gulls from morphometrics. *Ringing & Migration* **31**, 68–73.

Haywood, S. & Perrins, C.M. (1992) Is clutch size in birds affected by environmental conditions during growth? *Philosophical Transactions of the Royal Society London B Biological Sciences* **249**, 195–197.

Jenni, L. & Winkler, R. (1989) The feather length of small passerines: a measurement for wing length in live birds and museum skins. *Bird Study* **36**, 1–15.

Killick, R. & Eckley, I.A. (2014) *changepoint*: an R package for changepoint analysis. *Journal of Statistical Software* **58**, 1–19.

Killick, R., Haynes, K. & Eckley, I.A. (2016) changepoint*: an R package for changepoint analysis.* R package version 2.2.2. 2016. https://CRAN.R-project.org/package=changepoint [accessed 16 Nov 2019]

Kunz, C. & Ekman, J. (2000) Genetic and environmental components of growth in nestling blue tits (*Parus caeruleus*). *Journal of Evolutionary Biology* **13**, 199–212.

Li, J., Wang, N., Wang, Y., Lin, S., Li, Q., Liu, Y.Y., Ruan, X., Zhu, J., Xi, B. & Zhang, Z.W. (2010) Sexual size dimorphism and sex identification using morphological traits of two Aegithalidae species. *Zoological Science* **27**, 946–951.

Masello, J.F. & Quillfeldt, P. (2004) Consequences of La Niña phase of ENSO for the survival and growth of nestling Burrowing Parrots on the Atlantic coast of South America. *Emu* **104**, 337–346.

McKechnie, A.E. (2019) Physiological and morphological effects of climate change. In *Effects of Climate Change on Birds* (eds Dunn, P.O. & Møller, A.P.), p 120. Oxford Scholarship Online, London.

Merilä, J. & Fry, J.D. (1998) Genetic variation and causes of genotype–environment interaction in the body size of blue tit (*Parus caeruleus*). *Genetics* **148**, 1233–1244.

Norte, A.C., Ramos, J.A., Sousa, J.P. & Sheldon, B.C. (2009) Variation of adult Great Tit *Parus major* body condition and blood parameters in relation to sex, age, year and season. *Journal of Ornithology* **150**, 651.

Ormerod, S.J. & Tyler, S.J. (1990) Assessments of body condition in dippers *Cinclus cinclus*: potential pitfalls in the derivation and use of condition indices based on body proportions. *Ringing & Migration* **11,** 31–41.

Poissant, J., Morrissey, M.B., Gosler, A.G., Slate, J. & Sheldon, B.C. (2016a) Multivariate selection and intersexual genetic constraints in a wild bird population. *Journal of Evolutionary Biology* **29**, 2022–2035.

Poissant, J., Morrissey, M.B., Gosler, A.G., Slate, J. & Sheldon, B.C. (2016b) *Data from: multivariate selection and intersexual genetic constraints in a wild bird population*. Dryad. https://doi.org/10.5061/dryad.qt745

Quisenberry, J.H., Roberts, E. & Card, L.E. (1941) Genetic studies of skeletal dimensions and their relation to body weight and egg production in the domestic fowl (*Gallus domesticus*). *Poultry Science* **20**, 104–120.

R Core Team (2018) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Redfern, C.P.F. & Clark, J.A. (2001) *Ringers' Manual*. British Trust for Ornithology, Thetford.

Ren, Y., Wang, H. & Zhang, Z. (2016) Postnatal variation of limb bones in the Japanese Quail, *Coturnix coturnix japonica*. *International Journal of Morphology* **34**, 708–712.

Richner, H. (1989) Habitat-specific growth and fitness in Carrion Crows (*Corvus corone corone*). *Journal of Animal Ecology* **58**, 427–440.

Riddington, R. & Gosler, A.G. (1995) Differences in reproductive success and parental qualities between habitats in the Great Tit *Parus major*. *Ibis* **137**, 371–378.

Senar, J.C. & Pascual, J. (1997) Keel and tarsus length may provide a good predictor of avian body size. *Ardea* **85**, 269–274.

Svensson, L. (1992) *Identification Guide to European Passerines*. Privately published, Stockholm.

Taborsky, B. & Taborsky, M. (1999) The mating system and stability of pairs in Kiwi *Apteryx* spp. *Journal of Avian Biology* **30**, 143–151.

Taylor, R. (1990) Interpretation of the correlation coefficient: a basic review. *Journal of Diagnostic Medical Sonography* **6**, 35–39.

Tukey, J.W. (1977) *Exploratory Data Analysis*. Addison-Wesley, Princeton.

Vafidis, J.O., Vaughan, I.P., Jones, T.H., Facey, R.J., Parry, R. & Thomas, R.J. (2016) The effects of supplementary food on the breeding performance of Eurasian Reed Warblers *Acrocephalus scirpaceus*; implications for climate change impacts. *PloS One* **11**, e0159933. https://doi.org/10.1371/journal.pone.0159933

Wambui, G.D., Waititu, G.A. & Wanjoya, A. (2015) The power of the pruned exact linear time (PELT) test in multiple changepoint detection. *American Journal of Theoretical and Applied Statistics* **4**, 581–586.

Wang, J., Liu, J., Zhang, Z., Ren, H., Gao, L. & Hou, J. (2019) Is male condition corrected with song features in Dusky Warblers (*Phylloscopus fuscatus*). *Avian Research* **10**, 18. https://doi.org/10.1186/s40657-019-0158-5

Weeks, B.C., Willard, D.E., Zimova, M., Ellis, A.A., Witynski, M.L., Hennen, M. & Winger, B.M. (2020) Shared morphological consequences of global warming in North American migratory birds. *Ecology Letters* **23**, 316–325.

Wickham, H. (2016) ggplot2*: elegant graphics for data analysis*. Springer-Verlag, New York.

Wickham, H. & Henry, H. (2019) tidyr*: easily tidy data with 'spread()' and 'gather()' functions.* R package version 0.8.3. https://CRAN.R-project.org/package=tidyr [accessed 15 Mar 2019]

Wickham, H., François, R., Henry, L. & Müller, K. (2019) dplyr*: a grammar of data manipulation.* R package version 0.8.3. https://CRAN.R-project.org/package=dplyr [accessed 18 Aug 2019]

Winker, K. (1993) Specimen shrinkage in Tennessee Warblers and "Traill's" Flycatchers. *Journal of Field Ornithology* **64**, 331–336.

Zuberogoitia, I., Alonso, R., Palomares, L.E. & Martínez, J.A. (2011) Sex determination in Eurasian Sparrowhawks (*Accipiter nisus*). *Journal of Raptor Research* **45**, 48–55.

Appendices

Appendix 1

Mean ± SD of minimum and maximum tarsus measurements, and their Pearson's correlation coefficients (*r*), for 28 species of passerine birds captured during this study. Data are presented separately by EURING age category (Redfern & Clark 2001)

Table A1. Fully grown birds whose year of hatching was completely unknown, and could have been the calendar year of capture: EURING age code 2.

Common name	Scientific name	Count	Minimum	Maximum	*r*
Goldcrest	*Regulus regulus*	1	17.8 0 (–)	19.0 0 (–)	–
Greenfinch	*Chloris chloris*	1	18.2 0 (–)	21.3 0 (–)	–
Long-tailed Tit	*Aegithalos caudatus*	17	17.22 (± 0.36)	18.91 (± 0.35)	0.61
Sand Martin	*Riparia riparia*	10	10.93 (± 0.49)	13.25 (± 1.10)	0.75

Table A2. Birds that had definitely hatched during the calendar year of capture: EURING age code 3.

Common name	Scientific name	Count	Minimum	Maximum	*r*
Blackbird	*Turdus merula*	38	34.34 (± 1.07)	39.05 (± 1.19)	0.92
Blackcap	*Sylvia atricapilla*	7	20.21 (± 0.62)	22.65 (± 0.77)	0.90
Blue Tit	*Cyanistes caeruleus*	97	16.85 (± 0.58)	19.38 (± 0.64)	0.90
Chaffinch	*Fringilla coelebs*	115	18.36 (± 0.58)	21.08 (± 0.66)	0.94
Chiffchaff	*Phylloscopus collybita*	5	19.55 (± 0.86)	21.17 (± 0.85)	1.00
Coal Tit	*Periparus ater*	8	17.15 (± 0.59)	19.44 (± 0.65)	0.98
Dunnock	*Prunella modularis*	3	21.50 (± 0.95)	23.63 (± 1.01)	0.83
Garden Warbler	*Sylvia borin*	3	20.79 (± 0.30)	23.42 (± 0.86)	1.00
Goldcrest	*Regulus regulus*	12	17.03 (± 0.40)	18.36 (± 0.61)	0.93
Goldfinch	*Carduelis carduelis*	1	14.4 0 (–)	16.7 0 (–)	–
Great Tit	*Parus major*	86	20.51 (± 0.73)	23.37 (± 0.82)	0.76
Melodious Warbler	*Hippolais polyglotta*	1	19.16 (–)	21.10 (–)	–
Pied Flycatcher	*Ficedula hypoleuca*	1	17.33 (–)	18.70 (–)	–
Robin	*Erithacus rubecula*	56	25.88 (± 0.67)	28.74 (± 0.73)	0.90
Starling	*Sturnus vulgaris*	1	29.60 (–)	34.20 (–)	–
Treecreeper	*Certhia familiaris*	1	16.50 (–)	18.00 (–)	–
Willow Tit	*Poecile montanus*	2	16.25 (± 0.07)	18.20 (± 0.28)	–
Wren	*Troglodytes troglodytes*	3	17.53 (± 1.19)	20.07 (± 0.87)	1

Table A3. Birds that had hatched earlier than the calendar year of capture, but whose exact year of hatching was unknown: EURING age code 4.

Common name	Scientific name	Count	Minimum	Maximum	r
Blackbird	*Turdus merula*	13	34.16 (± 1.17)	39.12 (± 1.02)	0.83
Blackcap	*Sylvia atricapilla*	1	20.60 (–)	23.40 (–)	–
Blue Tit	*Cyanistes caeruleus*	27	16.89 (± 0.57)	19.35 (± 0.67)	0.93
Bullfinch	*Pyrrhula pyrrhula*	1	16.40 (–)	19.00 (–)	–
Cetti's Warbler	*Cettia cetti*	1	21.11 (–)	24.04 (–)	–
Chaffinch	*Fringilla coelebs*	47	18.17 (± 0.63)	20.77 (± 0.76)	0.93
Chiffchaff	*Phylloscopus collybita*	1	17.40 (–)	19.50 (–)	–
Coal Tit	*Periparus ater*	11	17.10 (± 0.43)	19.37 (± 0.61)	0.70
Dunnock	*Prunella modularis*	3	20.67 (± 0.35)	23.43 (± 0.06)	0.82
Garden Warbler	*Sylvia borin*	1	20.44 (–)	23.50 (–)	–
Goldfinch	*Carduelis carduelis*	1	14.80 (–)	17.60 (–)	–
Great Tit	*Parus major*	62	20.40 (± 0.66)	23.32 (± 0.71)	0.93
Hawfinch	*Coccothraustes coccothraustes*	2	21.90 (± 0.71)	25.45 (± 0.78)	–
House Sparrow	*Passer domesticus*	11	18.95 (± 1.06)	21.15 (± 1.10)	0.72
Nuthatch	*Sitta europaea*	9	19.84 (± 0.95)	22.41 (± 0.89)	0.75
Robin	*Erithacus rubecula*	25	25.81 (± 0.83)	28.62 (± 0.91)	0.93
Willow Warbler	*Phylloscopus trochilus*	1	17.98 (–)	20.03 (–)	–

Table A4. Birds that had definitely hatched during the calendar year immediately previous to capture: EURING age code 5.

Common name	Scientific name	Count	Minimum	Maximum	*r*
Blackbird	*Turdus merula*	22	34.10 (± 0.93)	38.90 (± 1.03)	0.93
Blackcap	*Sylvia atricapilla*	2	20.80 (± 0.71)	24.05 (± 0.64)	–
Blue Tit	*Cyanistes caeruleus*	55	16.61 (± 0.66)	18.91 (± 0.82)	0.91
Brambling	*Fringilla montifringilla*	2	19.55 (± 0.21)	21.40 (± 0.57)	–
Chaffinch	*Fringilla coelebs*	13	18.25 (± 0.82)	20.85 (± 0.74)	0.63
Coal Tit	*Periparus ater*	42	17.29 (± 0.47)	19.43 (± 0.48)	0.83
Dunnock	*Prunella modularis*	1	21.20 (–)	24.40 (–)	–
Goldfinch	*Carduelis carduelis*	4	15.15 (± 0.37)	17.40 (± 0.55)	0.94
Great Tit	*Parus major*	75	20.41 (± 0.71)	23.16 (± 0.78)	0.91
Hawfinch	*Coccothraustes coccothraustes*	10	21.79 (± 1.30)	25.29 (± 1.09)	0.81
Pied Flycatcher	*Ficedula hypoleuca*	12	16.75 (± 0.81)	19.37 (± 0.65)	0.91
Robin	*Erithacus rubecula*	11	25.55 (± 1.57)	28.50 (± 1.35)	0.97
Wren	*Troglodytes troglodytes*	1	18.40 (–)	20.30 (–)	–

Table A5. Birds that had hatched earlier than the calendar year immediately previous to capture, but whose exact year of hatching was unknown: EURING age code 6.

Common name	Scientific name	Count	Minimum	Maximum	*r*
Blackbird	*Turdus merula*	46	34.20 (± 1.16)	39.07 (± 1.13)	0.96
Blackcap	*Sylvia atricapilla*	1	20.30 (–)	22.40 (–)	–
Blue Tit	*Cyanistes caeruleus*	41	16.88 (± 0.58)	19.33 (± 0.68)	0.81
Brambling	*Fringilla montifringilla*	1	18.10 (–)	21.30 (–)	–
Chaffinch	*Fringilla coelebs*	16	18.11 (± 0.75)	20.78 (± 1.20)	0.83
Coal Tit	*Periparus ater*	45	17.35 (± 0.45)	19.47 (± 0.53)	0.79
Dunnock	*Prunella modularis*	1	22.30 (–)	23.90 (–)	–
Goldfinch	*Carduelis carduelis*	13	14.56 (± 0.50)	17.08 (± 0.47)	0.07
Great Tit	*Parus major*	84	20.69 (± 1.10)	23.41 (± 1.02)	0.91
Hawfinch	*Coccothraustes coccothraustes*	24	21.10 (± 0.71)	24.41 (± 0.77)	0.94
Pied Flycatcher	*Ficedula hypoleuca*	15	16.86 (± 0.61)	19.70 (± 0.49)	0.61
Robin	*Erithacus rubecula*	16	25.82 (± 0.91)	28.61 (± 0.91)	0.96
Wren	*Troglodytes troglodytes*	1	17.90 (–)	19.50 (–)	–

Appendix 2

Mean (± standard deviation) minimum and maximum tarsus measurements and correlations thereof for 28 species of passerine birds according to sex. r = Pearson's correlation value.

Table A6. Known females, identified via established criteria (e.g. plumage, brood patch in non-cooperative breeding species).

Common name	Latin name	Count	Minimum	Maximum	r
Blackbird	*Turdus merula*	37	34.02 (± 1.10)	38.78 (± 1.11)	0.95
Blackcap	*Sylvia atricapilla*	1	21.12 (± 0.00)	23.63 (± 0.00)	–
Blue Tit	*Cyanistes caeruleus*	46	16.45 (± 0.59)	18.69 (± 0.75)	0.90
Brambling	*Fringilla montifringilla*	2	19.55 (± 0.21)	21.40 (± 0.57)	–
Chaffinch	*Fringilla coelebs*	118	18.17 (± 0.59)	20.80 (± 0.76)	0.88
Chiffchaff	*Phylloscopus collybita*	2	19.12 (± 0.59)	20.97 (± 0.50)	–
Coal Tit	*Periparus ater*	20	17.38 (± 0.40)	19.29 (± 0.50)	0.73
Goldcrest	*Regulus regulus*	4	16.98 (± 0.25)	18.18 (± 0.13)	0.93
Goldfinch	*Carduelis carduelis*	7	14.41 (± 0.48)	17.07 (± 0.45)	0.37
Greenfinch	*Carduelis chloris*	1	18.20 (± 0.00)	21.30 (± 0.00)	–
Great Tit	*Parus major*	148	20.22 (± 0.63)	22.98 (± 0.72)	0.65
Hawfinch	*Coccothraustes coccothraustes*	15	21.93 (± 1.25)	25.60 (± 1.47)	0.89
House Sparrow	*Passer domesticus*	3	18.77 (± 0.84)	20.93 (± 1.12)	0.39
Nuthatch	*Sitta europaea*	3	19.93 (± 1.02)	23.03 (± 0.68)	0.93
Pied Flycatcher	*Ficedula hypoleuca*	27	16.81 (± 0.69)	19.55 (± 0.58)	0.77
Robin	*Erithcaus rubecula*	2	26.30 (± 0.71)	29.30 (± 0.85)	–

Table A7. Known males, identified established criteria (e.g. plumage, cloacal protuberance).

Common name	Latin name	Count	Minimum	Maximum	r
Blackbird	*Turdus merula*	71	34.28 (± 1.11)	39.17 (± 1.14)	0.91
Blackcap	*Sylvia atricapilla*	7	20.12 (± 0.49)	22.66 (± 0.76)	0.77
Blue Tit	*Cyanistes caeruleus*	37	17.08 (± 0.56)	19.45 (± 0.66)	0.90
Brambling	*Fringilla montifringilla*	1	18.10 (± 0.00)	21.30 (± 0.00)	–
Bullfinch	*Pyrrhula pyrrhula*	1	16.40 (± 0.00)	19.00 (± 0.00)	–
Cetti's Warbler	*Cettia cetti*	1	21.11 (± 0.00)	24.04 (± 0.00)	–
Chaffinch	*Fringilla coelebs*	68	18.45 (± 0.59)	21.15 (± 0.65)	0.85
Chiffchaff	*Phylloscopus collybita*	1	20.12 (± 0.00)	21.99 (± 0.00)	–
Coal Tit	*Periparus ater*	18	17.58 (± 0.52)	19.52 (± 0.57)	0.83
Dunnock	*Prunella modularis*	1	22.30 (± 0.00)	23.90 (± 0.00)	–
Goldcrest	*Regulus regulus*	9	17.14 (± 0.51)	18.51 (± 0.72)	0.94
Goldfinch	*Carduelis carduelis*	12	14.85 (± 0.46)	17.21 (± 0.52)	0.29
Great Tit	*Parus major*	144	20.84 (± 0.91)	23.70 (± 0.83)	0.91
Hawfinch	*Coccothraustes coccothraustes*	21	21.64 (± 1.22)	25.35 (± 1.33)	0.87
House Sparrow	*Passer domesticus*	8	19.03 (± 1.17)	21.24 (± 1.16)	0.78
Nuthatch	*Sitta europaea*	6	19.80 (± 1.09)	22.10 (± 0.75)	0.88
Robin	*Erithcaus rubecula*	6	25.65 (± 2.13)	28.57 (± 1.62)	0.97

Table A8. Individuals of unknown sex.

Common name	Latin name	Count	Minimum	Maximum	r
Blackbird	*Turdus merula*	11	34.52 (± 0.69)	38.93 (± 0.96)	0.95
Blackcap	*Sylvia atricapilla*	3	20.66 (± 0.62)	23.39 (± 1.11)	1.00
Blue Tit	*Cyanistes caeruleus*	137	16.95 (± 0.58)	19.38 (± 0.63)	1.00
Chaffinch	*Fringilla coelebs*	5	18.14 (± 0.22)	20.88 (± 0.26)	0.63
Chiffchaff	*Phylloscopus collybita*	3	19.65 (± 1.34)	20.95 (± 1.34)	1.00
Coal Tit	*Periparus ater*	68	17.21 (± 0.55)	19.51 (± 0.65)	0.93
Dunnock	*Prunella modularis*	7	21.30 (± 0.88)	23.77 (± 0.83)	0.80
Garden Warbler	*Sylvia borin*	4	20.71 (± 0.30)	23.44 (± 0.70)	0.77
Great Tit	*Parus major*	15	20.42 (± 0.74)	23.19 (± 0.93)	0.92
Long-tailed Tit	*Aegithalos caudatus*	17	17.19 (± 0.36)	18.94 (± 0.28)	0.21
Melodious Warbler	*Hippolais polyglotta*	1	19.16 (± 0.00)	21.10 (± 0.00)	–
Pied Flycatcher	*Ficedula hypoleuca*	1	17.33 (± 0.00)	18.70 (± 0.00)	0.00
Robin	*Erithcaus rubecula*	100	25.82 (± 0.75)	28.68 (± 0.89)	0.94
Sand Martin	*Riparia riparia*	10	10.93 (± 0.49)	13.25 (± 1.10)	0.75
Starling	*Sturnus vulgaris*	1	29.60 (± 0.00)	34.20 (± 0.00)	–
Treecreeper	*Certhia familiaris*	1	16.50 (± 0.00)	18.00 (± 0.00)	–
Willow Tit	*Poecile montanus*	2	16.25 (± 0.07)	18.20 (± 0.28)	–
Willow Warbler	*Phylloscopus trochilus*	1	17.98 (± 0.00)	20.03 (± 0.00)	–
Wren	*Troglodytes troglodytes*	5	18.45 (± 0.98)	20.55 (± 0.61)	1.00