Namibian Journal of Environment

Environmental Information Service, Namibia for the Ministry of Environment, Forestry and Tourism, the Namibian Chamber of Environment and the Namibia University of Science and Technology.

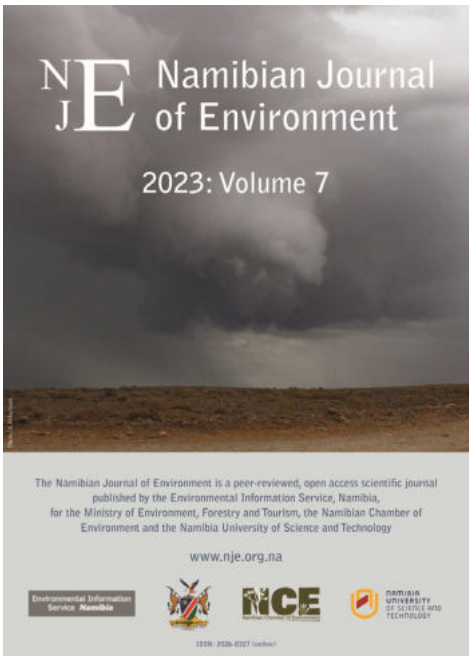
The *Namibian Journal of Environment* (NJE) covers broad environmental areas of ecology, agriculture, forestry, agro-forestry, social science, economics, water and energy, climate change, planning, land use, pollution, strategic and environmental assessments and related fields. The journal addresses the sustainable development agenda of the country in its broadest context. It publishes four categories of articles: **Section A: Research articles**. High quality peer-reviewed papers in basic and applied research, conforming to accepted scientific paper format and standards, and based on primary research findings, including testing of hypotheses and taxonomical revisions. **Section B: Research reports**. High quality peer-reviewed papers, generally shorter or less formal than Section A, including short notes, field observations, syntheses and reviews, scientific documentation and checklists. **Section C: Open articles**. Contributions not based on formal research results but nevertheless pertinent to Namibian environmental science, including opinion pieces, discussion papers, meta-data publications, non-ephemeral announcements, book reviews, correspondence, corrigenda and similar. **Section D: Monographs and Memoirs**. Peer-reviewed monographic contributions and comprehensive subject treatments (> 100 pages), including collections of related shorter papers like conference proceedings.

NJE aims to create a platform for scientists, planners, developers, managers and everyone involved in promoting Namibia's sustainable development. An Editorial Committee ensures that a high standard is maintained.

ISSN: 2026-8327 (online).

Chief Editor: R BECKER

Editor for this paper: JM MENDELSOHN



SECTION A: RESEARCH ARTICLES

Recommended citation format:

Weber MAB, Bryson U (2023) Ecological niche modelling of tree and wood pipits in southern Africa and adjacent countries may help to delimit distributions based on citizen science data. *Namibian Journal of Environment* 7 A: 14–20.

Cover photo: A Robertson

Ecological niche modelling of tree and wood pipits in southern Africa and adjacent countries may help to delimit distributions based on citizen science data

MAB Weber[1], U Bryson[2]

URL: https://www.nje.org.na/index.php/nje/article/view/volume7-weber
Published online: 17th October 2023

[1] ETH Zurich, Raemistrasse 101, 8092 Zurich, Switzerland; Onguma Nature Reserve, Namibia; manueweber@student.ethz.ch
[2] Becker-Gundahl-Str. 8, 81479 Munich, Germany

Date received: 25th June 2023; Date accepted: 14th October 2023.

ABSTRACT

Distribution maps are generally based on documented records rather than true occurrence patterns. This may be problematic for cryptic, under-reported species that occur in areas poorly covered by observers. Species distribution models may help overcome this challenge. Here, all available records of the migratory *Anthus trivialis* (tree pipit) and resident *Anthus nyassae* (wood pipit) for southern Africa and adjacent areas were assembled to train generalised linear models, random forest and gradient boosting machine species distribution models. Sampling pseudo-absences from a common species' similarly biased records helped to account for the spatial sampling bias present in the data. The model outputs suggest that *A. trivialis* and *A. nyassae* display a latitudinal habitat suitability gradient in the area of interest, opposing a latitudinal reporting gradient. The migratory behaviour of *A. trivialis* may blur its ecological niche. More and more reliable field observations are needed to confirm these findings. This study provides a clear framework to assist distribution delimitations from citizen science data by counteracting observer and sampling biases.

Keywords: *Anthus nyassae*, *Anthus trivialis*, citizen science, distribution, ecological niche model, species distribution model

INTRODUCTION

Delimiting species distributions can be a challenging endeavour, as it attempts to discretize different levels of abundance from data that are often incomplete. The recent emergence of citizen-science platforms such as SABAP2 (Second Southern African Bird Atlas Project) (Brooks and Ryan 2023) has generated a wealth of data that can help improve the delimitation of distributions. However, sampling and observer biases in the data (Kosmala *et al.* 2016) may distort the results of these efforts. As such, distribution maps based on citizen science, while spatially detailed compared to broad, expert-drawn range maps, may condense observations rather than delimiting the true occurrence patterns of a species.

SDMs (Species Distribution Models, introduced by Guisan *et al.* 2017, Guisan and Zimmermann 2000) may help to overcome this challenge by generating habitat models through the correlation of a taxon's current presence or absence with prevailing environmental conditions. Thus, the models can be useful in detecting areas in which species may be under-recorded.

Anthus trivialis (tree pipit) is a non-breeding palearctic migrant in sub-Saharan Africa, mainly present from October to March. It shares its woodland habitat with the resident *Anthus nyassae* (wood pipit). Few studies have examined the status and distribution of *A. trivialis* and *A. nyassae* in sub-Saharan Africa (Clancey 1987, 1989, 1990, Adams *et al.* 2022). The available distribution maps of the two species vary significantly between sources due to poor observer coverage in certain areas (Clancey 1987, 1989, 1990, BirdLife International 2016, 2018). Furthermore, both species may be challenging to identify due to their cryptic appearance. They thus provide good examples and materials for testing whether available records spatially reflect their ecological niche, as modelled by SDMs.

METHODS

The study area covered Angola, Zambia, Malawi, Mozambique and countries to the south within which the area of interest was further delimited by both data availability and the centroids of the distribution of *A. nyassae* and the wintering distribution of *A. trivialis*, respectively. All available records of the two species in the area of interest were gathered by consulting eBird (Auer *et al.* 2022), GBIF (Global Biodiversity Information Facility) (multiple sources outlined below), accessed through rgbif (Chamberlain *et al.* 2023), iNaturalist (iNaturalist contributors 2023), ABAP (African Bird Atlas Project), accessed through SABAP2 (Brooks and Ryan 2023) and termed SABAP2 thereafter, SARBN (Southern African Rare Bird News) (SARBN 2023), SAFRING (South African bird ringing unit) (SAFRING 2023) and BirdPix (Navarro 2023).

Except for the SAFRING data and some GBIF entries, which are part of museum collections or other scientific occurrence datasets (multiple sources outlined below), all records are citizen science based. The datasets were cleaned of duplicates and merged (Table 1). Although partially overlapping, eBird and GBIF provided the most data, followed by SABAP2. The majority of data were recorded in recent years.

Based on the ecology of *A. nyassae* and *A. trivialis* (Chittenden *et al.* 2018), seven environmental predictors were selected (Table 2) that appeared meaningful in determining the ecological niche of the two species. Although *A. trivialis* is only present in the region during the local summer months, the selected predictors quantifying precipitation and temperature span the whole year for both species, as the prevailing environmental conditions during the local summer months are dictated by the climatic conditions across all seasons. For example, winter temperatures may affect food availability during the summer, and the vegetation in winter rainfall areas does not rely on precipitation during the presence of *A. trivialis*. Furthermore, rare overwintering birds have been recorded (Chittenden *et al*. 2018).

All data were acquired and converted into rasters with a spatial resolution of 1 km using Google Earth Engine (https://earthengine.google.com). The predictors and records were loaded into the R statistical package (R Core Team 2018). An autocorrelation analysis yielded Pearson's correlation coefficients below 0.6 for all combinations, ensuring limited covariance between predictors. Duplicate data points at 1 km resolution were removed. Sampled SABAP2 occurrences of the fork-tailed drongo (*Dicrurus adsimilis*) (Brooks and Ryan 2023), a common bird present throughout most of southern Africa, served as pseudo-absences to counter the spatial sampling bias in the available records of *A. nyassae* and *A. trivialis* (Kramer-Schadt *et al.* 2013). The apparent habitat requirement of *D. adsimilis* is the presence of wooded cover, the same prerequisite for the occurrence of *A. nyassae* and *A. trivialis* (Chittenden *et al.* 2018). Thus, this step assumes that the absence of an observation of *D. adsimilis* at a given location implies a high probability that the location has not been sufficiently covered by observers to detect the presence of *A. nyassae* and *A. trivialis*.

Three distinct SDMs were run for each species in R (R Core Team 2018), namely GLM (Generalised Linear Model), RF (Random Forest), and GBM (Gradient Boosting Machine) models. For the former, a set of 2,500 pseudo-absences was used, while for the two latter, this number was reduced to 500 to approximately match the number of presences, as recommended for tree-based algorithms (Guisan *et al.* 2017). The GLMs were fitted using linear and quadratic terms as well as a stepwise variable selection based on the AIC (Akaike Information Criterion). A minimum of 10 observations was kept for every node, and 500 trees were grown for the RF models using the ranger package (Wright *et al.* 2023). The GBMs were fitted with a minimum of 10 observations per node, 500 trees, a learning rate of 0.1, and 10 cross-validation folds using the gbm package (Greenwell *et al.* 2022). All models were trained on all available occurrences except for the

Table 1: *Number of reported occurrences of* Anthus nyassae *and* A. trivialis *and time spans covered by cleaned datasets used in species distribution models. The indicated time span for GBIF is based on a minority of dated records. N/A: not applicable.*

Data source	***Anthus nyassae***		***Anthus trivialis***	
	Records	**Years**	**Records**	**Years**
eBird	315	1971–2023	137	1971–2023
GBIF	268	(2015–2023)	206	(2015–2023)
iNaturalist	22	2011–2023	16	2014–2023
SABAP2	127	2008–2023	60	2010–2023
SARBN	N/A	N/A	6	2013–2019
SAFRING; BirdPix	33	2006–2021	14	1960–2022
Total	**717**	**1971–2023**	**418**	**1960–2023**

Table 2: *Environmental predictors included in the species distribution models for* Anthus nyassae *and* A. trivialis*. An autocorrelation test yielded Pearson's correlation coefficients below 0.6 for all combinations.*

Predictor	**Ecological Scale**	**Source**
Annual mean temperature	Climatic	Karger *et al*. (2017)
Annual precipitation	Climatic	Karger *et al*. (2017)
Elevation	Topographic	Amatulli *et al*. (2021)
Tree density	Ecological	Crowther *et al*. (2015)
Leaf area index	Ecological	Myneni *et al*. (2021)
Human development	Anthropogenic	Tuanmu & Jetz (2014)
Landscape intactness	Anthropogenic	Potapov *et al*. (2008)

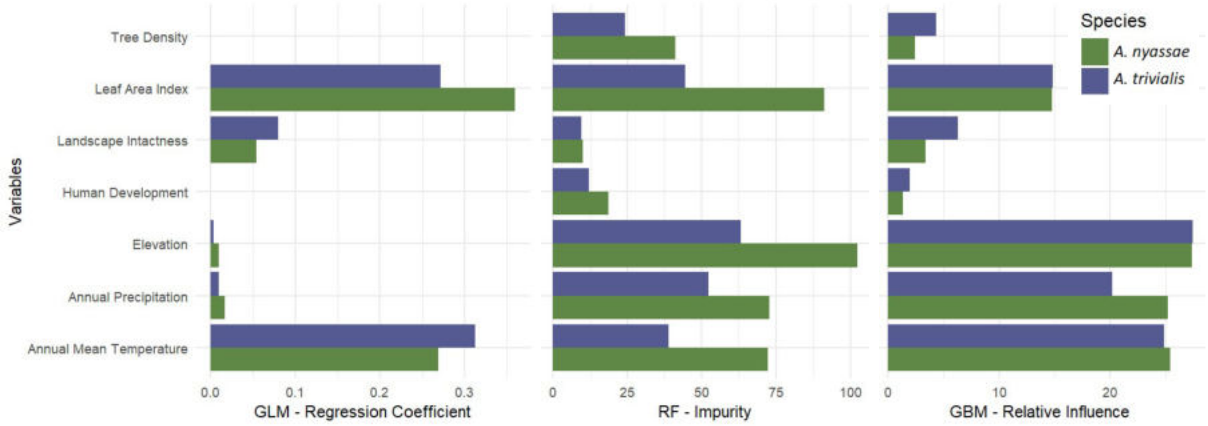


Figure 1: *Predictor importance across three species distribution models of two species of pipits (*Anthus nyassae *and* A. trivialis*). Generally, the climatic predictors (annual mean temperature and annual precipitation) performed best in explaining habitat suitability for both species, followed by topography and leaf area index. The tree-based algorithms (RF: Random Forest model; GBM: Gradient Boosting Machine model) yielded more balanced predictor importance values than the GLM (Generalised Linear Model). For the latter, only the coefficients of the linear terms are illustrated, as all regression coefficients of the quadratic terms were < 0.01. Furthermore, the stepwise variable selection based on Akaike Information Criterion excluded tree density and human development from the models.*

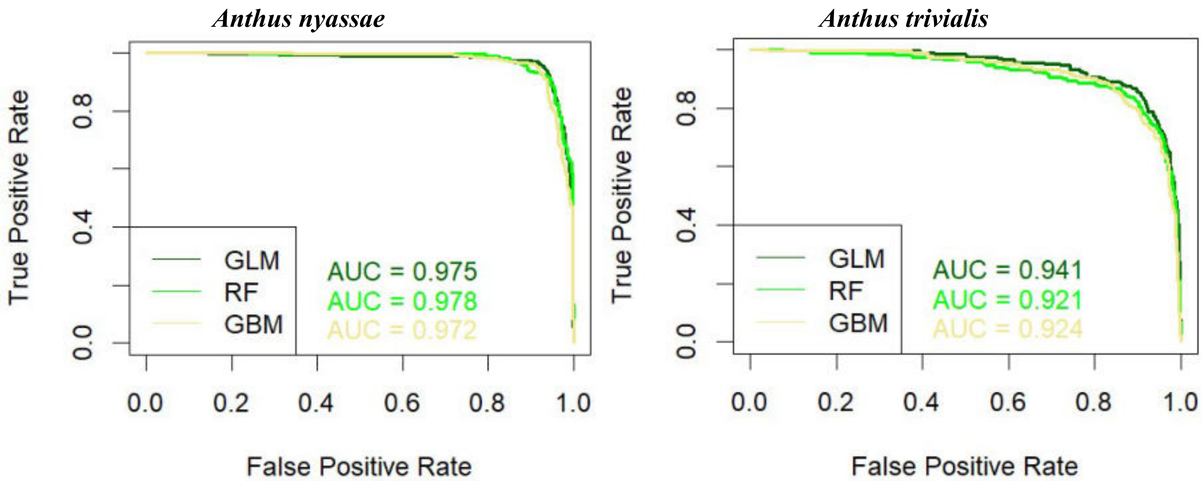


Figure 2: *Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC) values for three species distribution models of two species of pipits (*Anthus nyassae *and* A. trivialis*), generated from five-fold cross-validation. All models performed well for both species (Swets 1988). GLM: Generalised Linear Model; RF: Random Forest model; GBM: Gradient Boosting Machine model.*

SAFRING (SAFRING 2023) and BirdPix (Navarro 2023) datapoints. The latter were used for a visual comparison with the model predictions (Figure 1), as ringing data and reports covered by photographs are more reliable than ordinary citizen science records. The low number of data points and the sampling bias present in the SAFRING and BirdPix data prevented a computationally independent validation approach. Instead, a five-fold cross-validation was used to evaluate the models based on the ROC (Receiver Operating Characteristic) curve and AUC (Area Under Curve) values (Figure 2). The generated prediction layers were visualised using QGIS (QGIS Development Team 2023). For *A. nyassae*, the RF model was used to produce a projection, while for *A. trivialis*, the GBM was chosen based on model performance (high AUC values) and conservatism in predicting suitability (the model predicts high suitability less often).

RESULTS

All models performed well in terms of predictive performance. The Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) curves were all above 0.9, indicating good predictions (Swets 1988) (Figure 1). The climatic predictors performed best, and the anthropogenic variables (Table 2) performed worst in explaining habitat suitability for both species (Figure 2). Predictor

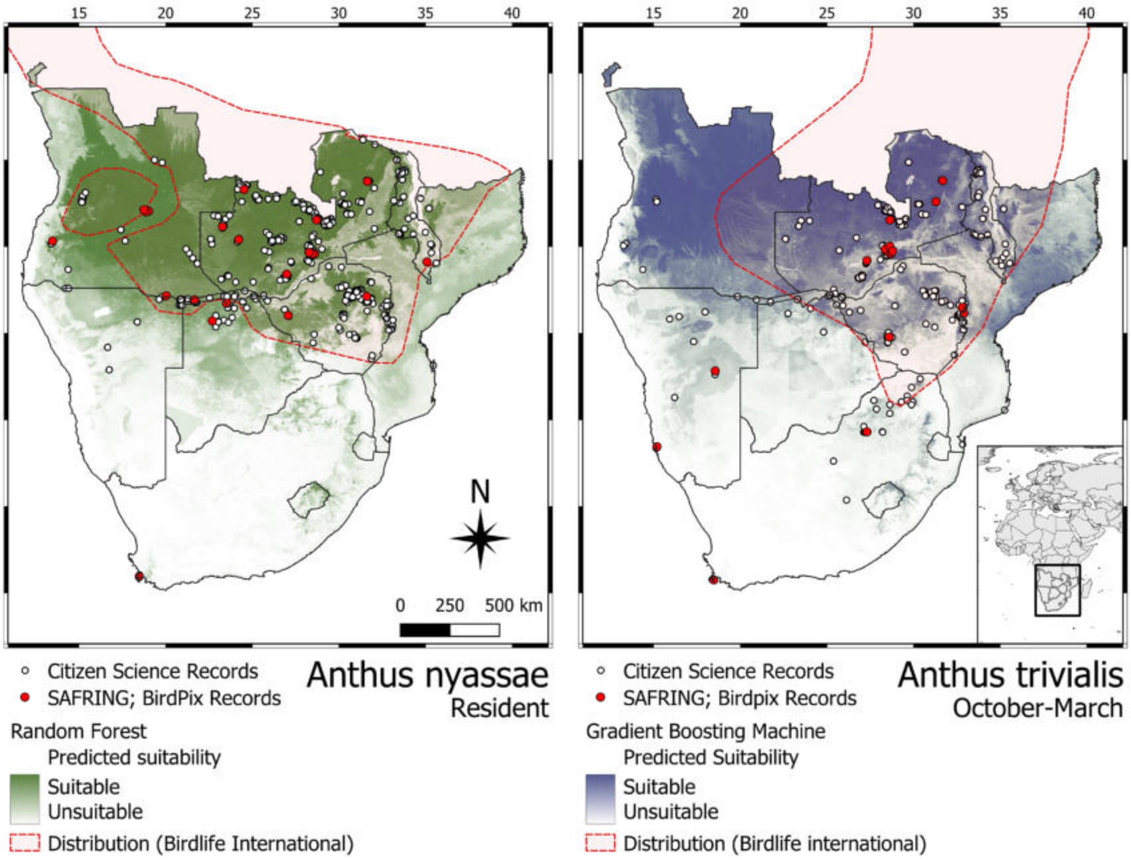


***Figure 3:** The white circles locate all available records for* A. nyassae *(left) and* A. trivialis *(right) in southern Africa and adjacent countries. The red polygons correspond to the (wintering for* A. trivialis*) distribution (BirdLife International 2016, 2018). The shades of green (*A. nyassae*) and blue (*A. trivialis*) show the potential distribution based on habitat suitability, as suggested by the random forest model (*A. nyassae*) and gradient boosting machine (*A. trivialis*) models.*

importance values were generally lower for *A. trivialis* than for *A. nyassae*, coinciding with slightly lower AUC values for *A. trivialis* (Figure 1, Figure 2).

The resulting habitat suitability maps (Figure 3) suggest that *A. nyassae* and *A. trivialis* display a latitudinal occurrence probability gradient: both species appear to be rare in the southern parts of the study area and more common further north. This contrasts with a latitudinal reporting gradient, as relatively few occurrences have been reported from potentially suitable areas such as Angola or northern Zambia.

DISCUSSION

The output suggests that current distribution maps exclude areas suitable for the potential occurrence of *A. trivialis* and *A. nyassae*, perhaps because they have been poorly covered by observers. Further exploration of the areas in question may yield new records of both species.

The strong model performances (Swets 1988) indicate a clear delimitation of the ecological niches of both species: the birds can generally be found in broadleaved woodlands at 800 or more meters above sea level, with at least 500 mm of annual rainfall. Unlike the resident *A. nyassae*, *A. trivialis* is only present in sub-Saharan Africa from October to March. The migratory behaviour may be reflected in the marginally poorer performance of the models in predicting its presence or absence (Figure 1) and the generally slightly lower predictor importances (Figure 2). Temporal fluctuations in the migratory patterns of *A. trivialis* due to varying food availability or weather conditions between years may further blur the picture.

Several sources of bias in the occurrence data need to be considered to contextualise the model outputs. Observer biases are inevitable in the context of both citizen science projects, such as eBird (Auer *et al.* 2022), iNaturalist (iNaturalist contributors 2023), or SABAP2 (Brooks & Ryan 2023), and platforms that are partially fed by citizen science, including GBIF (multiple sources outlined below). While the amount of data produced by many citizen scientists may make up for the quality trade-off (Kosmala *et al.* 2016), data quality control mechanisms may not always be able to flag faulty records. *Anthus* pipits can be notoriously difficult to identify even for specialists; hence, citizen science records should be approached with cautious scepticism. Although scarce, data provided by ringers or backed with photographs are far more reliable and may help to validate both other records and model outputs (Figure 2).

Furthermore, occurrence data may often be subject to a strong sampling bias as more accessible areas attract more observers (Kosmala *et al.* 2016). As opposed to SABAP1 (1987–1991), SABAP2 (Brooks & Ryan 2023) does not entail spatially systematic observations throughout the region (Bonnevie 2011). Instead, the observer decides where to observe. As such, the discrepancy between the model outputs and the number of reported sightings of *A. trivialis* around Gauteng, South Africa, is somewhat expected. On the other hand, the lack of records from central Angola for both species may be due to a lower observer density and less to the species absence. Choosing to sample pseudo-absences from a common species that shows the same sampling bias as the target species appears to be an efficient strategy to counteract this constraint in the modelling process (Kramer-Schadt *et al.* 2013).

The conclusions suggested by the model outputs need to be considered with caution. SDMs attempt to model a taxon's fundamental niche from occurrences that reflect the realised niche (Guisan *et al.* 2017). This can be problematic when niche parameters that are not captured by the model dictate where a species can occur, such as biotic interactions. While the output of this study may guide efforts to locate species where they have not been previously recorded, field records are needed to generate the final evidence.

Applying this approach to other cryptic and potentially under-reported species occurring in areas yet poorly covered by observations may help to delimit distributions through citizen science. However, due to the correlative nature of SDMs, their output provides merely an indication of where a species may occur, while conclusive evidence remains dependent on field data.

ACKNOWLEDGEMENTS
Trevor Hardaker kindly provided access to the SARBN archives. Prof. Thomas Crowther granted access to an environmental composite used to extract predictors. Thomas Lauber assisted with accessing the composite through Google Earth Engine. Prof. Damaris Zurell made important suggestions during the review process. Dr. Ken Stratford, Dr. John Mendelsohn and the NJE editing team assisted with the publication.

REFERENCES

Adams TA, Ivande ST, Wilson M (2022) Aspects of the nonbreeding ecology of the Tree Pipit *Anthus trivialis* in and around Amurum Forest Reserve, Jos Plateau, Nigeria. *Ostrich* 93(3): 187–191. https://doi.org/10.2989/00306525.2022.2113167.

Amatulli G, Domisch S, Tuanmu M-N, Parmentier B, Ranipeta A, Malczyk J, Jetz W (2018) A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data* 5(1): 180040. https://doi.org/10.1038/sdata.2018.40.

BirdLife International (2016) IUCN Red List of Threatened Species: *Anthus nyassae*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en.

BirdLife International (2018) IUCN Red List of Threatened Species: *Anthus trivialis*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en.

Bonnevie BT (2011) Some considerations when comparing SABAP 1 with SABAP 2 data. *Ostrich* 82(2): 161–162. https://doi.org/10.2989/00306525.2011.603486.

Chamberlain S, Oldoni D, Barve V, Desmet P, Geffert L, McGlinn D *et al.* (2023) rgbif: Interface to the Global Biodiversity Information Facility API. https://github.com/ropensci/rgbif.

Chittenden H, Davies G, Weiersbye I (2018) *Roberts Bird Guide*. Jacana Media.

Clancey P (1987) The Tree Pipit *Anthus triviaus* (Linnaeus) in southern Africa. *Durban Museum Novitates*. https://journals.co.za/doi/pdf/10.10520/AJA0012723X_2071.

Clancey P (1989) Zonal distribution in Afrotropical montane pipits. *Durban Museum Novitates*. https://journals.co.za/doi/pdf/10.10520/AJA0012723X_2167.

Clancey P (1990) A review of the indigenous pipits of the Afrotropics. *Durban Museum Novitates*. https://journals.co.za/doi/pdf/10.10520/AJA0012723X_1940.

Crowther TW, Glick HB, Covey KR, Bettigole C, Maynard DS, Thomas SM *et al.* (2015) Mapping tree density at a global scale. *Nature* 525(7568): 201–205. https://doi.org/10.1038/nature14967.

Guisan A, Thuiller W, Zimmermann NE (2017) *Habitat Suitability and Distribution Models: With Applications in R*. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781139028271.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135(2): 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9.

Greenwell B *et al.* (2022) gbm: Generalized boosted regression models. https://cran.r-project.org/web/packages/gbm/index.html.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW *et al.* (2017) Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4(1): 170122. https://doi.org/10.1038/sdata.2017.122.

Kosmala M, Wiggins A, Swanson A, Simmons B (2016) Assessing data quality in citizen science. *Frontiers in Ecology and the Environment* 14(10): 551–560. https://doi.org/10.1002/fee.1436.

Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V *et al.* (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions* 19(11): 1366–1379. https://doi.org/10.1111/ddi.12096.

Myneni R, Knyazikhin Y, Park T (2015) MOD15A2H MODIS Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. http://doi.org/10.5067/MODIS/MOD15A2H.006.

Potapov P, Yaroshenko A, Turubanova S, Dubinin M, Laestadius L, Thies C *et al.* (2008) Mapping the World's Intact Forest Landscapes by Remote Sensing. *Ecology and Society* 13(2). https://doi.org/10.5751/ES-02670-130251.

QGIS Development Team (2023) QGIS Geographic Information System. Open Source Geospatial Foundation. http://qgis.org

R Core Team (2018) R: A language and environment for statistical computing. *MSOR connections*. https://www.R-project.org/.

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240(4857): 1285–1293. https://doi.org/10.1126/science.3287615.

Tuanmu M-N, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography* 23(9): 1031–1045. https://doi.org/10.1111/geb.12182.

Wright M, Wager S, Probst P (2023) ranger: A fast implementation of random forests. https://cran.r-project.org/package=ranger.

Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J *et al.* (2020) A standard protocol for reporting species distribution models. *Ecography* 43(9): 1261–1277. https://doi.org/10.1111/ecog.04960.

OCCURRENCE DATA SOURCES

Auer T, Barker S, Barry J, Charnoky M, Curtis J, Davies I *et al.* (2022) EOD – eBird Observation Dataset. Cornell Lab of Ornithology. Occurrence dataset. https://doi.org/10.15468/aomfnb. (Accessed on 24 June 2023).

Brooks M and Ryan P (2023) Southern African Bird Atlas Project 2. Version 1.56. FitzPatrick Institute of African Ornithology. Occurrence dataset. https://doi.org/10.2989/00306525.2022.2125097. (Accessed on 24 June 2023).

Cicero C (2023) MVZ Bird Collection (Arctos). Version 43.73. Museum of Vertebrate Zoology. Occurrence dataset. https://doi.org/10.15468/r1woj3. (Accessed via GBIF.org on 24 June 2023).

Calabuig I (2013) Western Palearctic migratory birds in continental Africa. Center for Macroecology, Evolution and Climate, University of Copenhagen. Occurrence dataset. https://doi.org/10.15468/qjmrtx. (Accessed via GBIF.org on 24 June 2023).

Conservation International (2023) Rapid Assessment Program (RAP) Biodiversity Survey Database. Occurrence dataset. https://doi.org/10.15468/tsrjm0. (Accessed via GBIF.org on 24 June 2023).

de Vries H and Lemmens M (2023) Observation.org, Nature data from around the world. Observation.org. Occurrence dataset. https://doi.org/10.15468/5nilie. (Accessed via GBIF.org on 24 June 2023).

Epperly K (2022) UWBM Ornithology Collection. University of Washington Burke Museum. Occurrence dataset. https://doi.org/10.15468/a2cmfl. (Accessed via GBIF.org on 24 June 2023).

Figueira R (2017) IICT Colecção Zoológica. Version 4.2. Instituto de Investigação Científica Tropical. Occurrence dataset. https://doi.org/10.15468/rbon8c. (Accessed via GBIF.org on 24 June 2023).

Finnish Biodiversity Information Facility (2023) Ringing and recovery database of birds (TIPU). Occurrence dataset. https://doi.org/10.15468/kht1r3. (Accessed via GBIF.org on 24 June 2023).

Fjeldså J (2015) Aves Tanzanian collection at the Natural History Museum of Denmark (SNM). Zoological Museum, Natural History Museum of Denmark. Occurrence dataset. https://doi.org/10.15468/xebb5f. (Accessed via GBIF.org on 24 June 2023).

Gall L (2023) Vertebrate Zoology Division - Ornithology, Yale Peabody Museum. Yale University Peabody Museum. Occurrence dataset. https://doi.org/10.15468/h25uz7. (Accessed via GBIF.org on 24 June 2023).

iNaturalist contributors (2023) iNaturalist research-grade observations. iNaturalist.org. Occurrence dataset. https://doi.org/10.15468/ab3s5x (Accessed on 24 June 2023).

Kamminga P and Creuwels J (2023) Naturalis Biodiversity Center (NL) - Aves. Naturalis Biodiversity Center. Occurrence dataset. https://doi.org/10.15468/dxmzbz. (Accessed via GBIF.org on 24 June 2023).

Mertz W (2023) LACM Vertebrate Collection. Version 18.13. Natural History Museum of Los Angeles County. Occurrence dataset. https://doi.org/10.15468/77rmwd. (Accessed via GBIF.org on 24 June 2023).

Millen B (2018) Ornithology Collection Passeriformes - Royal Ontario Museum. Version 15.10. Royal Ontario Museum. Occurrence dataset. https://doi.org/10.15468/ahiyvz. (Accessed via GBIF.org on 24 June 2023).

Morris P J (2023) Museum of Comparative Zoology, Harvard University. Version 162.377. Museum of Comparative Zoology, Harvard University. Occurrence dataset. https://doi.org/10.15468/p5rupv. (Accessed via GBIF.org on 24 June 2023).

Museum of Comparative Zoology, Harvard University. Occurrence dataset. https://doi.org/10.15468/p5rupv. (Accessed via GBIF.org on 24 June 2023).

Natural History Museum (2023) Natural History Museum (London) Collection Specimens. Occurrence dataset. https://doi.org/10.5519/0002965. (Accessed via GBIF.org on 24 June 2023).

Naturalis Biodiversity Center (NL) - Aves. Naturalis Biodiversity Center. Occurrence dataset. https://doi.org/10.15468/dxmzbz. (Accessed via GBIF.org on 24 June 2023).

naturgucker.de (2023) naturgucker. Occurrence dataset. https://doi.org/10.15468/uc1apo. (Accessed via GBIF.org on 24 June 2023).

Navarro R (2023) BirdPix. FitzPatrick Institute of African Ornithology. Occurrence dataset. http://vmus.adu.org.za/?vm=BirdPix. (Accessed via GBIF.org on 24 June 2023).

Navarro R and Jackson C (2023) Kenya Virtual Museum Records. Version 1.19. FitzPatrick Institute of African Ornithology. Occurrence dataset. https://doi.org/10.15468/jolylt. (Accessed via GBIF.org on 24 June 2023).

NMNH Extant Specimen Records (USNM, US). Version 1.69. National Museum of Natural History, Smithsonian Institution. Occurrence dataset. https://doi.org/10.15468/hnhrg3. (Accessed via GBIF.org on 24 June 2023).

Orrell T (2023) NMNH Extant Specimen Records (USNM, US). Version 1.69. National Museum of Natural History, Smithsonian Institution. Occurrence dataset. https://doi.org/10.15468/hnhrg3. (Accessed via GBIF.org on 24 June 2023).

Oschadleus D and Ranwashe F (2017) SAFRING: Historical Bird Ringing Records (2005-2009). Version

1.1. South African National Biodiversity Institute. Occurrence dataset. https://doi.org/10.15468/jclwaz (Accessed on 24 June 2023).
SAFRING (2023) https://safring.birdmap.africa. (Accessed on 24 June 2023).
SARBN (2023) Southern African Rare Bird News. sa-rarebirdnews@googlegroups.com. (Accessed on 24 June 2023).
Scholes E (2015) Macaulay Library Audio and Video Collection. Cornell Lab of Ornithology. Occurrence dataset. https://doi.org/10.15468/ckcdpy. (Accessed via GBIF.org on 24 June 2023).
South African National Biodiversity Institute. Southern African Bird Atlas Project. Version 1.0. Occurrence dataset. https://doi.org/10.15468/ry4ux6. (Accessed via GBIF.org on 24 June 2023).
Tushabe H (2021) National Biodiversity Data Bank. Observation records, 1900-2014. Version 1.6. National Biodiversity Data Bank. Occurrence dataset. https://doi.org/10.15468/djzgie. (Accessed via GBIF.org on 24 June 2023).
UWBM Ornithology Collection. University of Washington Burke Museum. Occurrence dataset. https://doi.org/10.15468/a2cmfl. (Accessed via GBIF.org on 24 June 2023).
Vellinga W (2023) Xeno-canto - Bird sounds from around the world. Xeno-canto Foundation for Nature Sounds. Occurrence dataset. https://doi.org/10.15468/qv0ksn. (Accessed via GBIF.org on 24 June 2023).
White J, White N, Barker D, Evans D, Steer T, David M *et al.* (2023) Birda - Global Observation Dataset. Version 1.14. Birda. Occurrence dataset. https://doi.org/10.15468/6kud7x. (Accessed via GBIF.org on 24 June 2023).

ADDITIONAL RESOURCES

The following GitHub repository provides this study's Google Earth Engine script that was used to acquire and preprocess the predictors, as well as the R script that was used to preprocess the data and run the models: https://github.com/Manuel-Weber-ETH/Anthus_nyassae_trivialis.git