Better together? Assessing different remote sensing products for predicting habitat suitability of wetland birds

Koma, Z.; Seijmonsbergen, A.C.; Grootes, M.W.; Nattino, F.; Groot, J.; Sierdsema, H.; Foppen, R.P.B.; Kissling, W.D.

DOI
10.1111/ddi.13468

Publication date
2022

Document Version
Final published version

Published in
Diversity and distributions

Received: 25 June 2021 | Revised: 25 November 2021 | Accepted: 2 December 2021

DOI: 10.1111/ddi.13468

RESEARCH ARTICLE

Diversity and Distributions WILEY

Better together? Assessing different remote sensing products for predicting habitat suitability of wetland birds

Zsófia Koma[1,2] | Arie C. Seijmonsbergen[1] | Meiert W. Grootes[3] | Francesco Nattino[3] | Jim Groot[1] | Henk Sierdsema[4] | Ruud P. B. Foppen[4,5] | W. Daniel Kissling[1,6]

[1]Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Biology, Center for Sustainable Landscapes Under Global Change, Aarhus University, Aarhus, Denmark

[3]Netherlands eScience Center, Amsterdam, The Netherlands

[4]Sovon Dutch Centre for Field Ornithology, Nijmegen, The Netherlands

[5]Department of Animal Ecology and Ecophysiology, Institute for Water and Wetland Research, Radboud University, Nijmegen, The Netherlands

[6]LifeWatch Virtual Laboratory Innovation Center (VLIC), LifeWatch ERIC, Amsterdam, The Netherlands

Correspondence
Zsófia Koma, Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, P.O. Box 94240, 1090 GE Amsterdam, The Netherlands.
Email: komazsofi@gmail.com

Funding information
Netherlands eScience Center, Grant/Award Number: ASDI.2016.014

Editor: Yoan Fourcade

Abstract

Aim: The increasing availability of remote sensing (RS) products from airborne laser scanning (ALS) surveys, synthetic aperture radar acquisitions and multispectral satellite imagery provides unprecedented opportunities for describing the physical structure and seasonal changes of vegetation. However, the added value of these RS products for predicting species distributions and animal habitats beyond land cover maps remains little explored. Here, we aim to assess how metrics derived from different types of high-resolution (10 m) RS products predict the habitat suitability of wetland birds.

Location: North-eastern part of the Netherlands.

Methods: We built species distribution models (SDMs) with occurrence observations from territory mapping of two selected wetland bird species (great reed warbler and Savi's warbler) and metrics from a Dutch land cover map, country-wide ALS and Sentinel-1 and Sentinel-2 RS products. We then compared model performance, relative variable importance and response curves of the SDMs to assess the contribution and ecological relevance of each RS product and metric.

Results: Our results showed that ALS and Sentinel metrics improve SDMs with only land cover metrics by 11% and 10% of the Area Under Curve (AUC) for the great reed warbler and the Savi's warbler respectively. Assessments of feature importance revealed that all types of RS products contributed substantially to predicting the habitat suitability of these wetland birds, but that the most important variables vary among species.

Main conclusions: Our study demonstrates that metrics from different high-resolution RS products capture complementary ecological information on animal habitats, including aspects such as the proportional cover of habitat types, vegetation density and the horizontal variability of vegetation height. Land cover maps with detailed spatial and thematic information can already achieve high model accuracies, but adding metrics derived from ALS point clouds and Sentinel imagery further improve model accuracy and enhance the understanding of animal–habitat relationships.

[Correction added on 26 February 2022, after first online publication: Author name "Daniel Kissling" has been changed to "W. Daniel Kissling".]

KEYWORDS
Acrocephalus, habitat heterogeneity, LiDAR, *Locustella*, reedbed structure, SAR, Sentinel, temporal vegetation dynamics

1 | INTRODUCTION

The explanation and prediction of species distributions and biodiversity patterns of animals is a fundamental pursuit in biogeography, ecology and conservation biology. Species distribution models (SDMs) are commonly used to quantify the relationship between species occurrences and environmental conditions (Elith et al., 2006; Guisan & Zimmermann, 2000). In recent years, SDMs increasingly include remote sensing (RS) data to describe the environmental conditions of animal species at broad spatial extents (He et al., 2015). However, SDMs often focus solely on environmental predictor variables related to climate, topography, soil and land cover (Mod et al., 2016). Thus, information about the 3D configuration and physical structure of vegetation is rarely included (Kissling et al., 2017; Randin et al., 2020). In contrast, field observations show that the fine-scale structure of vegetation plays an important role in determining the distribution and diversity of animal species (Dunlavy, 1935; MacArthur & MacArthur, 1961; Tews et al., 2004). Moreover, structural variability and heterogeneity of vegetation has a positive effect on biodiversity by providing more niche space (Koma, Grootes, et al., 2021; Müller et al., 2014), by influencing resource use and habitat selection (Cody, 1985) and by creating diverse microclimates (Zellweger et al., 2019). Hence, the development of predictor variables from RS products that capture aspects of vegetation structure and habitat heterogeneity is required.

Categorical land cover maps are often used in SDMs to describe the coarse habitat use of animals (Pearson et al., 2004; Thuiller et al., 2004). The main reason is the wide availability and easy interpretability of land cover maps such as the CORINE Land Cover (https://land.copernicus.eu/pan-european/corine-land-cover). However, such products usually only distinguish relatively coarse land cover classes, which often do not fully align with the habitat types of animals (Kissling et al., 2017; St-Louis et al., 2014). National land cover maps can reveal finer details, but the predefined land cover classes still do not necessarily reflect the habitat requirements of species in relation to fine-scale vegetation structure. For instance, small and scattered habitats (e.g. linear vegetation elements such hedges) and the structural variation within land cover classes (e.g. density and height of reedbeds) are under-represented (Koma et al., 2021; Lucas et al., 2019) and hence do not properly describe the breeding and foraging requirements of animal species (Davies & Asner, 2014; Zellweger et al., 2016). Some studies suggest overcoming these limitations by using vegetation indices derived from satellite RS, which can describe spatial and temporal variation of vegetation growth and other aspects of habitat heterogeneity (Oeser et al., 2020; St-Louis et al., 2014). However, most studies so far have relied on RS products such as LANDSAT and MODIS spectral imagery, for which the spatial resolution is 30 m or coarser. The increasing availability of country-wide laser scanning data and the recent launch of the European Space Agency (ESA) Copernicus missions (Sentinel-1 and Sentinel-2) provide new opportunities for quantifying the fine-scale habitat variation at high (e.g. 10 m) spatial resolution (Schulte to Bühne & Pettorelli, 2018).

Light Detection and Ranging (LiDAR) technology provides the most direct and accurate way of obtaining detailed information about vegetation structure (Davies & Asner, 2014; Valbuena et al., 2020; Vierling et al., 2008). Airborne laser scanning (ALS) is an active RS technique using LiDAR where the scanner emits a laser pulse that is then reflected back from the vegetation (e.g. leaves, branches and stems) or from the terrain surface. As a result, the measurements provide a 3D point cloud with centimetre accuracy, representing vegetation structure or terrain. To derive ecologically relevant information, the obtained 3D point cloud needs to be further processed, for example into metrics which statistically aggregate the 3D point cloud information within raster cells (Bakx et al., 2019; Davies & Asner, 2014). These LiDAR metrics can then be used to model the fine-scale habitat suitability of animals such as birds, mammals and invertebrates (Davies et al., 2018; de Vries et al., 2021; Zellweger et al., 2013). Even though several studies have successfully used LiDAR metrics for characterizing habitat structure, the ALS data usually come with two major limitations. First, country-wide ALS campaigns are typically not repeated in different seasons, which limits the ability to monitor seasonal changes in vegetation structure caused by plant growth dynamics. Second, ALS datasets are often collected for terrain mapping in the leaf-off season and hence do not necessarily capture all aspects of vegetation structure during the time when animals are surveyed, for example in the breeding season (Koma, Grootes, et al., 2021; Koma, Seijmonsbergen, et al., 2021). It is thus important to explore the combination of LiDAR with other high spatial and temporal resolution satellite RS products to test which predictor variables can best describe habitat structure and seasonal dynamics of vegetation.

The Copernicus Sentinel programme provides an unprecedented opportunity to monitor the spatial and temporal changes of vegetation (Pettorelli et al., 2014; Schulte to Bühne & Pettorelli, 2018). For instance, the Synthetic Aperture Radar (SAR) sensor on board of Sentinel-1 can be used to derive ecologically meaningful information about vegetation structure (Bae et al., 2019). SAR data can be used to analyse the 3D structure of objects because the signals are sensitive to the orientation, volume and roughness of the surface (e.g. plants). SAR data have been used to map above-ground biomass (Sinha et al., 2015) and to provide information on vegetation structure in forested areas (Bae et al., 2019). Additionally, the Sentinel-2 satellite has a multispectral imaging sensor, which provides opportunities to derive common vegetation indices such as the Normalized Difference Vegetation Index (NDVI). The NDVI provides information

on the dynamics of primary production from the top of the vegetation canopy. This has been successfully used to predict animal species distributions (Pettorelli et al., 2005, 2011). Recent studies have also shown that textural metrics derived from multispectral imagery can provide information on habitat heterogeneity for predicting bird species richness (Farwell et al., 2021). Although the use of different types of satellite RS products in SDMs is increasing, a comprehensive assessment of how multiple types of high-resolution RS data (e.g. covering LiDAR, Sentinel-1 and Sentinel-2) can predict animal habitats is still lacking.

Here, we aim to assess the potential of different high-resolution RS products to predict the fine-scale habitat suitability of wetland birds. We compare metrics related to the proportion, 3D structure, heterogeneity and seasonal variability of wetland vegetation using national land cover maps, country-wide ALS data, and SAR and multispectral imagery obtained from Sentinel-1 and Sentinel-2 satellites, respectively. We focus on wetlands because they are vital and structurally complex habitat types that offer breeding habitat for many species of conservation concern like marshland birds whose niches differ in vegetation composition and structure (Leisler & Schulze-Hagen, 2011). We used breeding bird information from a comprehensive national monitoring scheme together with metrics derived from different high-resolution (10 m) RS products and predicted species distributions of two wetland bird species, the great reed warbler (*Acrocephalus arundinaceus*) and the Savi's Warbler (*Locustella luscionides*). Both species have specialized habitat requirements within wetlands and breed almost exclusively within reedbeds (Cramp, 1992). In particular, we (1) compare model accuracies using different types of high-resolution RS products, (2) evaluate the relative importance of metrics derived from different RS products and (3) use response curves to assess the ecological relevance of the derived metrics for explaining the habitat suitability of the focal species. We expect that LiDAR and Sentinel-based metrics achieve higher overall accuracies in SDMs compared to SDMs using only categorical land cover maps. We further hypothesize that LiDAR metrics will contribute most to the models because they represent the fine-scale habitat structure more directly than metrics calculated from satellite RS. Furthermore, we expect that different RS products will provide complementary ecological information to explain the habitat suitability of both bird species. Overall, our comparative study offers new insights into how different high-resolution RS data can be used to model species distributions and thereby advances the characterization of fine-scale habitat structure and seasonal variability of vegetation in quantitative biodiversity modelling.

2 | METHODS

2.1 | Study area

The study area within the Netherlands comprised five Dutch provinces (Groningen, Drenthe, Overijssel, Gelderland and Flevoland) for which both ALS and Sentinel data were available for the same year (Figure 1). All derived metrics were georeferenced into the Dutch coordinate system (RD_New) and resampled according to the extent and resolution of the LiDAR metrics. Within the included provinces, we modelled the habitat suitability of the selected birds only within wetlands because the focal bird species only breed in wetlands. For this, we applied a wetland mask using the Dutch land cover map (Landelijk Grondgebruik Nederland, LGN2018), including only the six wetland-related LGN classes: reed, swamp, salt marsh, shrubs with low vegetation height, shrubs with high vegetation height and forests in wetlands.

2.2 | Breeding bird occurrence data

We selected two bird species that breed in wetlands: the great reed warbler (*Acrocephalus arundinaceus*) and the Savi's warbler (*Locustella luscinioides*). Both species almost exclusively breed within reedbeds, but have distinct habitat preferences in terms of vegetation structure and types of reedbeds. For instance, the great reed warbler breeds in areas close to open water (e.g. 5 m away from the water edge) (Báldi & Kisbenedek, 1999), whereas the Savi's warbler prefers large contiguous reedbeds mixed with other herbaceous perennial plants and with a few trees and bushes interspersed (Báldi, 2006). Within reedbeds, the vegetation structure and composition typically differs with distance to the water. Close to the water edge, reed vegetation (water reed) grows taller and thicker than in reedbeds that are located in the drier parts of a wetland (i.e. land reed) (Graveland, 1998; Koma, Seijmonsbergen, et al., 2021).

The breeding bird observations were collected by the Dutch Centre for Field Ornithology (Sovon, https://www.sovon.nl/en). We used bird observations obtained from the Breeding Bird Monitoring Program (BMP) between 2016 and 2019. This survey provides locations of territorial bird activity within survey plots across the Netherlands using a standardized field sampling protocol with a similar observation effort across species (Vergeer et al., 2016). Each year between March and July all survey plots (ranging from 10 to 500 hectares) are visited 5–10 times in the early morning by an experienced observer. During the field visits, the spatial locations of all birds with territorial and nest indicative behaviour (e.g. song, alarm, nest) are mapped. This sampling scheme thus provides high-resolution presence data (i.e. exact spatial locations of observed individuals during the breeding season). We generated 'background' points by randomly placing points within the survey plots and within the wetland-related land cover types. Furthermore, we randomly generated 'absences' within the survey plots and within the wetland-related land cover types, but at least 200 m away from the presence points. The 200 m radius was selected based on the habitat use, territory size and activity range of the selected bird species during the breeding season. We also applied spatial thinning on the presence data, which was set to 20 m to avoid sampling the same 10 m grid cell twice within the study area. In total, we had 146 presence and 730 absence or background points for the great reed warbler and 1483 presence points and 1231 absence or background points for

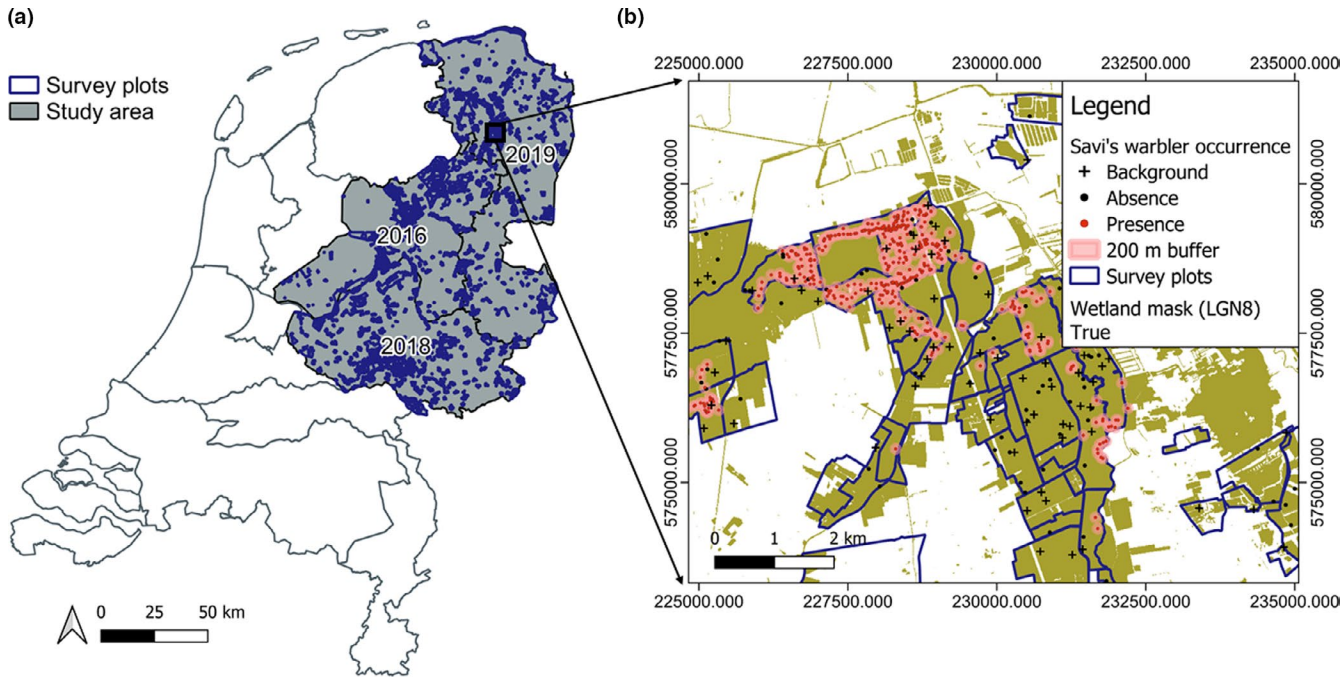

FIGURE 1 Study area, survey plots and species distribution information of birds within wetland land cover types. (a) Simultaneous coverage of airborne laser scanning (ALS) and Sentinel data in the north-eastern part of the Netherlands (grey = years indicating data acquisition), with survey plots of bird observations in blue. (b) Example area showing observed presences of birds (red dots, with 200 m buffer), generated background points (black cross) and generated absences (black dots, at least 200 m away from presences) of the Savi's warbler (SW) in survey plots within wetlands. The wetland mask (indicated with bronze colour) is generated using the six wetland-related classes from the Dutch land cover map (LGN8 = Landelijk Grondgebruik Nederland)

the Savi's warbler. To avoid an imbalanced training and testing dataset for the great reed warbler, we equalized the number of generated absence points with the number of presence points in each random fold of the modelling procedure (see below). Hence, we used different sets of absences in each modelling run (Barbet-Massin et al., 2012).

2.3 | Remote sensing data

We used different RS products to predict the habitat suitability of the selected wetland birds. This included a land cover map from the Netherlands, country-wide ALS data, SAR data (Sentinel-1) and optical imagery (Sentinel-2) from the Copernicus Sentinel missions of the European Space Agency (ESA). The sensors included active and passive RS techniques and thus contributed different information on the structure and dynamics of wetland vegetation. For instance, country-wide ALS provides 3D point clouds, which directly reflect the structure of vegetation, SAR sensors measure the scattering characteristic of the surface and thus capture changes in biomass and water content, and optical imagery captures spectral information from the top of the vegetation canopy. Both SAR and optical imagery from Sentinel further comes with a high temporal frequency (repeated measurements every 6 days). From each of these RS products, we extracted various metrics to capture different habitat aspects, including coverage of different land cover classes, vertical and horizontal vegetation structure, habitat heterogeneity, and seasonal and spatial variation in vegetation growth (Figure 2).

2.3.1 | Dutch land cover map

We used the Dutch land cover map from 2018 (LGN2018 or LGN8), which distinguishes 48 land cover types (e.g. agricultural crops, forests, water, wetland types and urban classes) at high (5 m) resolution. LGN8 is a combined land cover classification product based on high-resolution satellite SPOT imagery, aerial photographs and the Dutch cadastre data (TOP10NL, https://zakelijk.kadaster.nl/-/top10nl). Six LGN8 classes differentiate the wetland habitats of the selected bird species: reed, swamp, salt marsh, shrubs with low vegetation height, shrubs with high vegetation height and forests in wetlands. We used these six LGN8 classes to quantify the proportion of each land cover class using a moving window approach with a 100 m radius (see metric class 'Proportion of land cover' in Table 1).

2.3.2 | Airborne laser scanning data

The country-wide LiDAR point clouds were derived from the third Dutch national ALS flight campaign (AHN3, Actueel Hoogtebestand Nederland). The AHN3 dataset is openly accessible and captures multiple returns, with an average point density of 8 pt/m^2 (https://

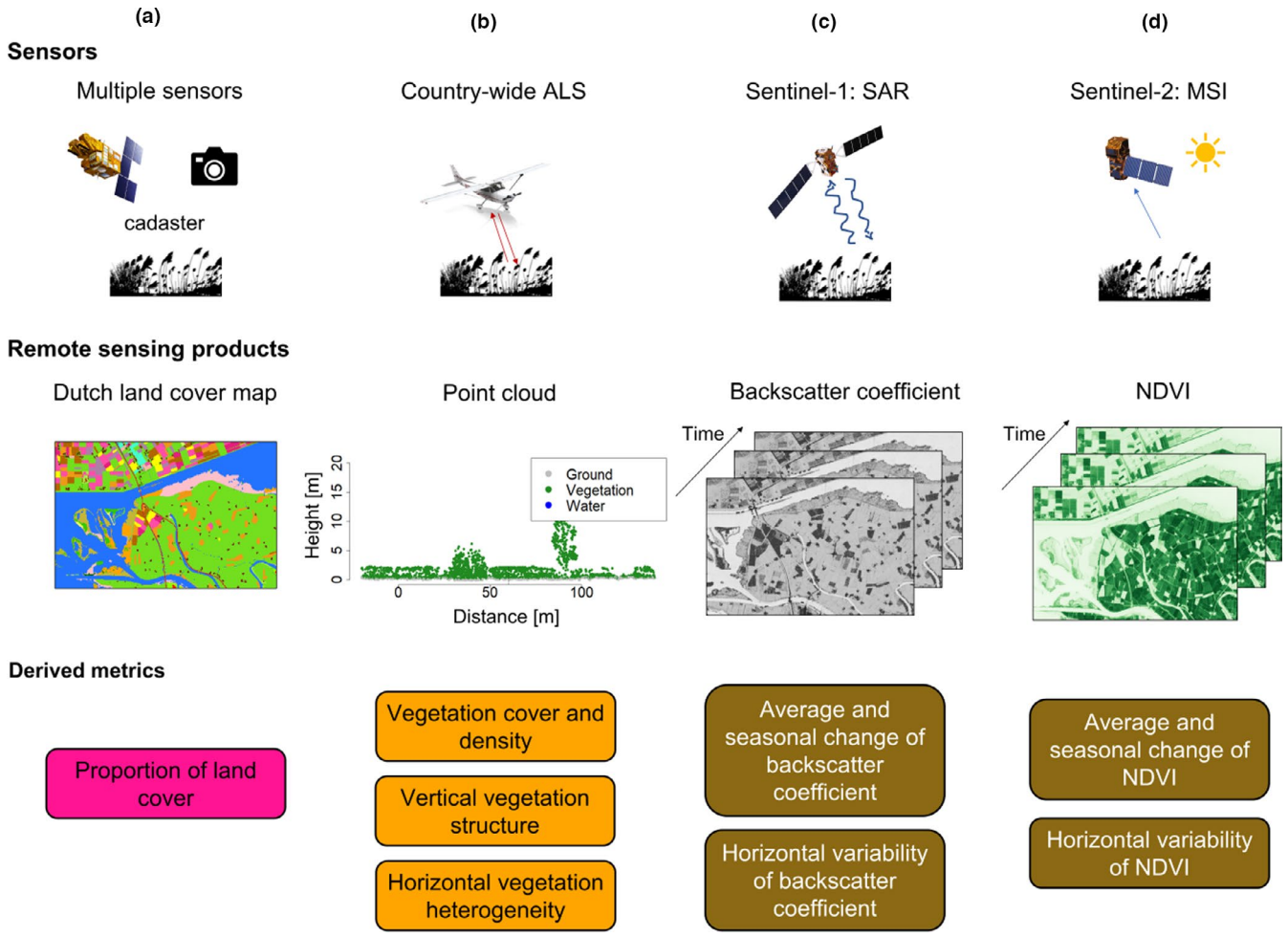

FIGURE 2 Overview of remote sensing sensors, products and derived metrics. (a) Multiple sensors (e.g. satellite images and aerial photographs in combination with cadastre data) are used to produce the Dutch land cover map from which several metrics capturing the proportion of wetland land cover types were derived. (b) Country-wide ALS (Airborne Laser Scanning) surveys provide 3D point clouds from which vegetation cover and density, vertical structure and horizontal heterogeneity were derived. (c) Satellite remote sensing data from Sentinel-1 time series of the backscatter coefficient from the synthetic aperture radar (SAR), which can be converted into metrics capturing annual average, seasonal change and horizontal variability of vegetation biomass and water content. (d) Sentinel-2 provides optical imagery from the Multispectral Instrument (MSI) sensor, which allows to calculate time series of the Normalized Difference Vegetation Index (NDVI) from which metrics of annual average, seasonal change and horizontal variability of vegetation greenness can be derived. See Table 1 for detailed information on metrics

ahn.arcgisonline.nl/ahnviewer/). The data were acquired during the leaf-off season between 2016 and 2019. The raw point cloud has been pre-processed by 'Rijkswaterstraat' (the executive agency of the Dutch Ministry of Infrastructure and Water Management) and comes with a classification of ground, building, water and non-ground points. The point clouds were processed using the python software 'laserchicken' (Meijer et al., 2020) and the High Performance Computing (HPC) workflow 'laserfarm' (https://laserfarm.readthedocs.io/en/latest/). This enabled the efficient, scalable and distributed processing of multi-terabyte LiDAR point clouds from the AHN3 surveys using a cluster of virtual machines from the Dutch national ICT facility (https://www.surf.nl/en/ict-facilities).

The workflow consisted of several modular pipelines. After downloading the AHN3 data to a local data storage, the raw data were re-tiled and then normalized by calculating the normalized z-value for each individual point as the height relative to the lowest point within a 1 m × 1 m cell. A number of LiDAR metrics were then calculated. We included a total of 11 LiDAR metrics to describe the fine-scale vegetation structure and habitat heterogeneity within wetlands (see metric class 'LiDAR metrics' in Table 1). These LiDAR metrics are commonly used in predictive habitat distribution models of birds (Bakx et al., 2019) and have also been shown to be important for separating the fine-scale habitat niches of our study species (Koma, Grootes, et al., 2021). Four of the LiDAR metrics represent measures of vegetation cover and density (e.g. pulse penetration ratio and density of vegetation points in certain strata) (see metric class 'Vegetation cover and density' in Table 1), and five additional metrics capture the vertical variability of vegetation (e.g. 95th percentile of vegetation height) (see metric class 'Vertical vegetation structure' in Table 1). All of these metrics

were calculated at 10 m spatial resolution. An additional three LiDAR metrics capture habitat heterogeneity (habitat proportion and horizontal variability) using a 100 m radius around focal grid cells (see metric class 'Horizontal vegetation heterogeneity' in Table 1). For this, we applied a moving window approach and used the 95th percentile of vegetation height as an input metric (Table 1). Horizontal variability was quantified using the standard deviation of the 95th percentile of height, either for all vegetation or low vegetation (<5 m). We further extracted the proportion of predominantly reed vegetation using the 95th percentile of height metric with a threshold of >1 m and <3 m. This differentiates reed (*Phragmites australis*) and other perennial wetland plants with >1 m height (e.g. *Typha angustifolia*) from high vegetation such as bushes and trees (>3 m).

2.3.3 | Sentinel data

The Sentinel-1 satellite data from the ESA Copernicus mission have been obtained with a C-band Synthetic Aperture Radar (SAR) sensor in multiple acquisition modes at different ground sampling distances (Torres et al., 2012). For our study, we used the Interferometric Wide (IW) swath mode at 10 m resolution. This offers polarization options such as vertical transmission with vertical receiving (VV) and vertical transmission with horizontal receiving (VH). Within Google Earth Engine (Gorelick et al., 2017), the available Sentinel-1 products consist of several pre-processing routines. These include removal of border noise, thermal noise, radiometric calibration, orthorectification of the images and the corrections of radiometric distortions (https://developers.google.com/earth-engine/guides/sentinel1). Overall, the retrieved images contained information of the backscatter coefficients measured in decibel (dB). We selected both VV and VH images for the relevant years (2016–2019) within our study area and calculated a total of 10 metrics (see metric class 'Sentinel metrics' in Table 1). These metrics used either the VV or the VH radar signals and captured average or seasonal changes or the horizontal variability (within a 100 m moving window) of the backscatter coefficients (Table 1).

For optical imagery, we used the ESA Copernicus Sentinel-2 satellite, which is equipped with a Multispectral Instrument (MSI) sensor (Drusch et al., 2012). This provides multispectral datasets across thirteen bands in visible, near-infrared and the short wave infrared part of the electromagnetic spectrum. We used the top of atmospheric reflectance (Level-1C product), which can be retrieved from Google Earth Engine (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2). To obtain cloud-free images, we only selected the images where the cloud cover was less than 2% based on the provided quality assurance bands for our study area. We then calculated for every image the NDVI at 10 m resolution. The NDVI is the ratio between the red and near-infrared values and quantifies vegetation greenness and vegetation density of plant canopies. To capture the seasonal and horizontal variability of vegetation greenness, we calculated a total of five Sentinel-2 metrics using the average, maximum and standard deviation of NDVI values (Table 1).

2.4 | Statistical analysis

We applied SDMs to assess and compare the predictive accuracy and feature importance of metrics from different RS products. We implemented this approach across different groups of metrics, that is using only the metrics from the Dutch land cover map, only the LiDAR metrics from the country-wide ALS data and only the combination of Sentinel-1 and Sentinel-2 metrics. Additionally, we also combined LiDAR and Sentinel metrics or all metrics together into one SDM to assess and compare the importance of fusing different RS products for predicting habitat suitability of the selected bird species. All statistical analyses were conducted with the R software v. 4.0.5 (https://www.r-project.org/). Furthermore, all developed R scripts are made available via GitHub (https://github.com/komazsofi/LiDAR_Sentinel_birdsdm_wetlands).

2.4.1 | Collinearity analysis

We performed a collinearity analysis with the derived metrics for each RS product separately (i.e. land cover, LiDAR and Sentinel) and then also across all remaining metrics. For assessing the collinearity between metrics, we carried out a stepwise Variance Inflation Factor (VIF) analysis using the R package 'usdm' v.1.1.-18 (Naimi et al., 2014). The VIF threshold was set to three to avoid collinearity between the derived metrics. Metrics with VIF >3 were excluded from further analysis, and all remaining 22 metrics had VIF <3 across all RS products (indicated with an asterisk in Table 1, see Table S1 for the VIF values of remaining metrics).

2.4.2 | Species distribution modelling

We built SDMs using three different algorithms: Generalized Linear Models (GLM; Guisan et al., 2002), Maximum Entropy (Maxent; Phillips et al., 2006) and Random Forest (RF; Breiman, 2001). We implemented the SDMs using the R package 'sdm' v.1.0-89 (Naimi & Araújo, 2016), which depends on the R package 'stats' for GLMs, the Java software for Maxent ('maxent.jar') and the R package 'randomForest' for RF. For the GLM and RF methods, we used the presence and absence data as described above, whereas in the case of Maxent we used the presence and background data. We further combined the 'sdm' package with the R package 'blockcv' v.2.1.1 (Valavi et al., 2019) for performing a spatial repetitive cross-validation procedure. This type of cross-validation allowed us to use geographically separated folds to calculate the model accuracies and hence avoid the underestimation of the prediction error by generating a spatially separated cross-validation (Valavi et al., 2019). We defined several 8 x 8 km blocks across our study area and set five-fold for the spatial cross-validation. This blocksize

TABLE 1 Overview of metrics calculated from different remote sensing products

Metric class	Name of the metric	Metric abbreviation	Description
Land cover metrics			
Proportion of land cover	Proportion of swamp*	landcover_propswamp	Proportion of land cover class 'swamp' (5 m resolution) within 100 m radius
	Proportion of salt marsh*	landcover_propsaltmarsh	Proportion of land cover class 'salt marsh' (5 m resolution) within 100 m radius
	Proportion of reed*	landcover_propreed	Proportion of land cover class 'reed' (5 m resolution) within 100 m radius
	Proportion of low shrub*	landcover_proplowshrub	Proportion of land cover class 'low shrub within wetlands' (5 m resolution) within 100 m radius
	Proportion of high shrub*	landcover_prophighshrub	Proportion of land cover class 'high shrub within wetlands' (5 m resolution) within 100 m radius
	Proportion of forest*	landcover_propforest	Proportion of land cover class 'forest within wetlands' (5 m resolution) within 100 m radius
LiDAR metrics			
Vegetation cover and density	Pulse penetration ratio*	lidar_C_ppr	Ratio of number of ground points to total number of points within 10 m grid cell
	Density of vegetation points below 1 m	lidar_VD_0_1	Ratio of number of points <1 m relative to number of total vegetation points within 10 m grid cell
	Density of vegetation points between 1 and 2 m*	lidar_VD_1_2	Ratio of number of points between 1–2 m relative to total number of vegetation points within 10 m grid cell
	Density of vegetation points between 2 and 3 m*	lidar_VD_2_3	Ratio of number of points between 2–3 m relative to total number of vegetation points within 10 m grid cell
Vertical vegetation structure	95^{th} percentile of height	lidar_VV_p95	95^{th} percentile of normalized z within 10 m grid cell
	25^{th} percentile of height*	lidar_VV_p25	25^{th} percentile of normalized z within 10 m grid cell
	Foliage height diversity	lidar_VV_FHD	Shannon entropy of normalized z within 10 m grid cell derived from height layers with 0.5 m thickness
	Kurtosis of height*	lidar_VV_kurt	Kurtosis of normalized z within 10 m grid cell
	Standard deviation of height*	lidar_VV_std	Standard deviation of normalized z within 10 m grid cell
Horizontal vegetation heterogeneity	Horizontal variability of total vegetation height*	lidar_HH_sd	Standard deviation (within 100 m radius) of 95^{th} percentile of normalized z (lidar_VV_p95 values of 10 m resolution grid cells)
	Horizontal variability of low vegetation*	lidar_HH_sd_low	Standard deviation (within 100 m radius) of 95^{th} percentile of normalized z below 5 m (lidar_VV_p95 values of 10 m resolution grid cells)
	Proportion of predominantly reed vegetation*	lidar_HH_reedveg_prop	Proportion of 10 m grid cells with normalized z > 1 m AND z < 3 m within 100 m radius, based on 95^{th} percentile of normalized z (lidar_VV_p95)

(Continues)

TABLE 1 (Continued)

Metric class	Name of the metric	Metric abbreviation	Description
Sentinel metrics			
Sentinel-1			
Average and seasonal change of backscatter coefficient	Annual median of VH polarization	radar_VHmed	The median of VH images within 10 m grid cells within a year
	Annual mean of VH polarization	radar_VHmean	The mean of VH images within 10 m grid cells within a year
	Annual standard deviation of VH polarization*	radar_VHstd	The standard deviation of VH images within 10 m grid cells within a year
	Annual maximum of VH polarization	radar_VHmax	The maximum of VH images within 10 m grid cells within a year
	Annual median of VV polarization	radar_VVmed	The median of VV images within 10 m grid cells within a year
	Annual mean of VV polarization	radar_VVmean	The mean of VV images within 10 m grid cells within a year
	Annual standard deviation of VV polarization*	radar_VVstd	The standard deviation of VV images within 10 m grid cells within a year
	Annual maximum of VV polarization*	radar_VVmax	The maximum of VV images within 10 m grid cells within a year
Horizontal variability of backscatter coefficient	Horizontal variability of the annual median of VH polarization*	radar_VHsd_hor_100m	
	Standard deviation of the annual median VH (10 m grid cells) within 100 m radius around the focal grid cell		
	Horizontal variability of the annual median of VV polarization	radar_VVsd_hor_100m	Standard deviation of the annual median VV (10 m grid cells) within 100 m radius around the focal grid cell
Sentinel-2			
Average and seasonal change of NDVI	Annual median of NDVI*	optical_NDVImed	The median of NDVI within 10 m grid cells within a year
	Annual mean of NDVI	optical NDVImean	The mean of NDVI within 10 m grid cells within a year
	Annual standard deviation of NDVI*	optical_NDVIstd	The standard deviation of NDVI within 10 m grid cells within a year
	Annual maximum of NDVI	optical_NDVImax	The maximum of NDVI within 10 m grid cells within a year
Horizontal variability of NDVI	Horizontal variability of the annual median NDVI*	optical_NDVIsd_hor_100m	Standard deviation of the yearly median of NDVI (10 m grid cells)within 100 m radius around the focal grid cell

Note: Land cover metrics were derived from the Dutch land cover map (LGN8), LiDAR metrics from a country-wide airborne laser scanning survey and Sentinel metrics from the SAR instrument of the Sentinel-1 mission and the multispectral imagery of the Sentinel-2 mission, respectively. The spatial resolution of all metrics is 10 m, but the metrics related to horizontal vegetation heterogeneity were calculated based on a 100 m moving window approach around focal 10 m grid cells. The asterisks (*) indicate which metrics were included after a collinearity analysis.

Abbreviations: ALS, Airborne Laser Scanning; LiDAR, Light Detection And Ranging; NDVI, Normalized Difference Vegetation Index; SAR, Synthetic Aperture Radar; VH, vertical transmitting with horizontal receiving dual polarization of the SAR signal; VV, vertical transmitting with vertical receiving dual polarization of the SAR signal.

was suggested by automatically fitting non-directional isotropic variogram models, which determine the effective range of autocorrelation using 5000 randomly selected points from each continuous predictor variables (Valavi et al., 2019). Additionally, for each spatial fold we randomly split the training and testing datasets 20 times. This resulted in 100 runs for each type of SDM method. Overall, model accuracy and

model performance were reported based on the combination of the three SDM algorithms by averaging the modelling results across the algorithms using the R package 'sdm' v.1.0-89. Additionally, we report the results separately for each algorithm (GLM, Maxent and RF) in the Appendices B and C. The parametrization of the SDM algorithms can be found in Appendix E.

To assess and compare the modelling results across different RS products, we calculated model accuracy using the Area Under Curve (AUC), the True Skill Statistic (TSS), Deviance (D) and Cohen's Kappa Statistics (Kappa). We reported all four accuracy measures to avoid biases when using only AUC (Lobo et al., 2008; Peterson et al., 2008). We extracted the mean and the standard deviation across the 300 runs for each accuracy measure separately. Accuracy values of AUC between 0.70 and 0.80 were interpreted as fair, between 0.80 and 0.90 as good and between 0.90 and 1.00 as excellent (Swets, 1988). For the evaluation of TSS values, we used the approach recommended by Landis and Koch (1977): between 0 and 0.40 were interpreted as poor, between 0.40 and 0.75 as good and between 0.75 and 1.00 as excellent. Accuracy values of Kappa between 0.21 and 0.40 were interpreted as fair, between 0.41 and 0.60 as moderate, between 0.61 and 0.80 as good and between 0.81 and 1.00 as excellent (Altman, 1990). We further used the relative variable importance and the response curves to interpret and compare the contribution of different metrics and their relevance for explaining the habitat preferences of the selected wetland bird species. The relative variable importance, which was measured by AUC improvements of model performance (Naimi & Araújo, 2016), was extracted for all metric group combinations. For every modelling technique, response curves for each metric were extracted to show the species-specific responses. These response curves represent the probability of occurrence of the species along the gradient of the selected metric while keeping all other predictor variables in the model at their mean (Elith et al., 2005). To address our research objectives, we compared (1) the different SDM accuracy measures (AUC, TSS, D and Kappa) across different RS products (only land cover, only LiDAR, only Sentinel, the combination of LiDAR and Sentinel, and the combination of LiDAR, Sentinel and land cover); (2) the relative feature importance of metrics within and across the different RS products; and (3) the response curves of the most important metrics to interpret the results ecologically based on the full SDM (consisting of land cover, LiDAR and Sentinel metrics). In the main text, we report the results from combining the three algorithms. All modelling results (accuracy measures, relative variable importance and response curves) were calculated using the R package 'sdm' v.1.0–89.

3 | RESULTS

3.1 | Comparison of modelling accuracies

The SDMs of the great reed warbler (Figure 3a) achieved a fair to good overall accuracy across the different RS data products (AUC = 0.73–0.84, TSS = 0.49–0.65, Kappa = 0.49–0.66). The lowest accuracy was achieved by using land cover metrics only (AUC = 0.73, TSS = 0.49, Kappa = 0.49), whereas the highest accuracy was achieved by using only Sentinel metrics (AUC = 0.84, TSS = 0.65, Kappa = 0.66). Compared to the SDM with land cover only, the SDM based on LiDAR metrics improved the overall accuracy by 3%, whereas the SDM with Sentinel metrics improved the overall accuracy by 11% according to the AUC (see Table S3 for details). The SDM with all metrics achieved slightly lower overall accuracies than the SDM using only Sentinel metrics (AUC = 0.80, TSS = 0.59, Kappa = 0.59). Deviance values were acceptable for all SDMs ($D < 7.17$).

The SDMs of the Savi's Warbler (Figure 3b) showed moderate to excellent overall modelling performance for all RS products (AUC = 0.80–0.90, TSS = 0.50–0.71, Kappa = 0.50–0.71). Among the different RS products, the SDM with only land cover metrics achieved good overall accuracy (AUC = 0.80, TSS = 0.50, Kappa = 0.51), whereas the SDM with only LiDAR metrics achieved excellent modelling performance (AUC = 0.90, TSS = 0.69, Kappa = 0.69). Hence, SDMs based on LiDAR metrics improved the overall accuracy by 10% (measured in AUC) compared to SDMs using only land cover. The combination of all three RS products showed the highest overall accuracy (AUC = 0.90, TSS = 0.71, Kappa = 0.71) and improved the overall accuracy by 10% compared to the SDM with only land cover metrics (see for more details Table S3). Deviance values were acceptable for all SDMs ($D < 1.12$).

3.2 | Evaluation of feature importance

The analysis of feature importance showed that metrics from all three types of RS products substantially contributed to the SDMs. However, the most important metrics varied between the focal species (Figure 4). The most important variable for the great reed warbler was the horizontal variability of the NDVI (optical_NDVIsd_hor_100m) extracted from the Sentinel-2 data (Figure 4a), followed by a vegetation cover-related LiDAR metric (lidar_C_ppr) and a land cover-related metric describing the proportion of swamp (landcover_propswamp). For the Savi's warbler (Figure 4b), the most important metric was the proportion of predominantly reed vegetation (lidar_HH_reedveg_prop) extracted from LiDAR, and, similarly to the great reed warbler, the proportion of swamp vegetation (landcover_propswamp) extracted from the land cover map. This was followed by a vegetation cover-related LiDAR metric (lidar_C_ppr). Metrics extracted from Sentinel-1 and Sentinel-2 data were generally less important (Figure 4b).

3.3 | Assessing the ecological relevance with response curves

The response curves of the five most important metrics (Figure 4c–l) revealed how the breeding occurrences of the two species varied along gradients of vegetation structure. For instance, the response

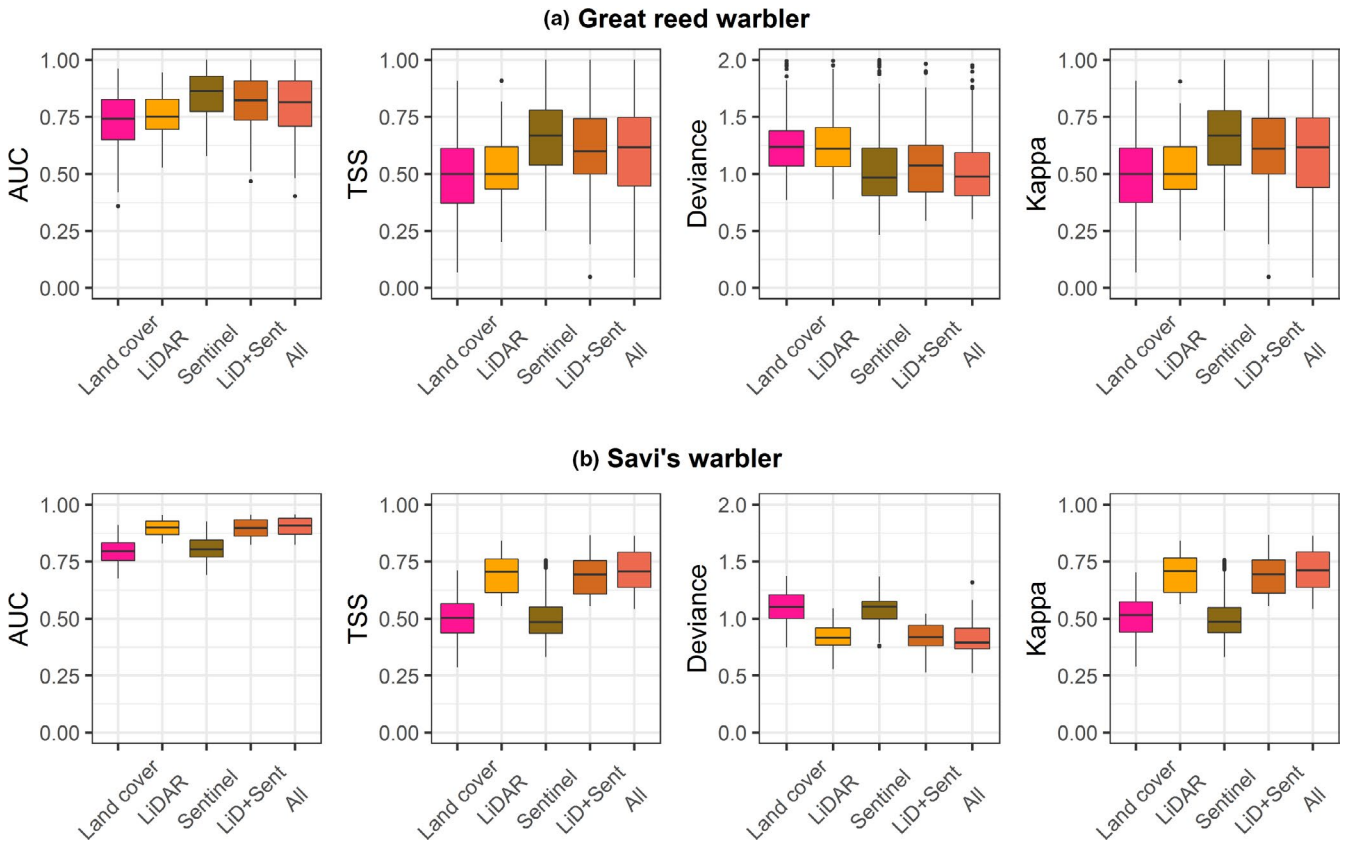

FIGURE 3 Performance of species distribution models (SDMs) in relation to metrics derived from different remote sensing products. Results are shown separately for (a) great reed warbler and (b) Savi's warbler. Model performance is assessed using the area under the curve (AUC), the true skill statistic (TSS), deviance and Cohen's Kappa (panels from left to right). Model accuracy was calculated across 300 model runs using three algorithms (Generalized Linear Model, Maxent and Random Forest). Colours indicate the different remote sensing products that were used in the SDMs. 'LiD+Sent' indicates the combination of LiDAR and Sentinel metrics, and 'All' indicates the combination of LiDAR, Sentinel and land cover metrics

curves of the great reed warbler (Figure 4c–g) indicated that the probability of occurrence increased when habitat heterogeneity measured as the horizontal variability of the NDVI (optical_NDVIsd_hor_100m) was high (>0.1 difference in NDVI, Figure 4c). The probability of occurrence further decreased when vegetation cover derived from LiDAR (lidar_C_ppr) was less dense (>0.65 ratio between ground and vegetation points, Figure 4d) and when the proportion of swamp derived from land cover increased (Figure 4e). Analysis of Sentinel-1 data also indicated that habitat suitability of the great reed warbler increased when the annual maximum value of VV (radar_VVmax) was >−8 dB (Figure 4f), indicating a high biomass accumulation and wetness during the growing season. Moreover, habitat suitability of the great reed warbler also increased when the proportion of predominantly reed vegetation derived from LiDAR (lidar_HH_reedveg_prop) was >0.50 (Figure 4g).

For the Savi's warbler, the response curves (Figure 4h–l) showed that the probability of occurrence increased when the proportion of predominantly reed vegetation derived from LiDAR (lidar_HH_reedveg_prop) was greater than 25% (Figure 4h). When the proportion of swamp derived from the land cover map (landcover_propswamp) was greater than 65%, the probability of occurrence of this species strongly declined (Figure 4i). Furthermore, the Savi's warbler was likely to be absent when vegetation cover derived from LiDAR (lidar_C_ppr) was sparse (> 0.75 ratio between ground and vegetation points, Figure 4j). The species further mainly occurred when the horizontal variability of vegetation from LiDAR (lidar_HH_sd) was low (Figure 4f). The probability of occurrence also decreased when the 25th percentile of height from LiDAR (lidar_VV_25p) was >3 m (Figure 4l).

4 | DISCUSSION

We assessed the potential of using different high-resolution RS products (including a Dutch land cover map, country-wide ALS data, Sentinel-1 SAR and Sentinel-2 optical imagery) for predicting the fine-scale habitat suitability of two wetland birds (great reed warbler and Savi's warbler). Our results show that LiDAR and Sentinel metrics can improve the accuracy of SDMs compared to models using only land cover metrics, even if the land cover map has a 5 m spatial resolution and a high thematic detail. Moreover, our assessment of feature importance and the interpretation of response curves revealed that LiDAR and Sentinel metrics contribute substantially to a better understanding of animal–habitat relationships within wetlands.

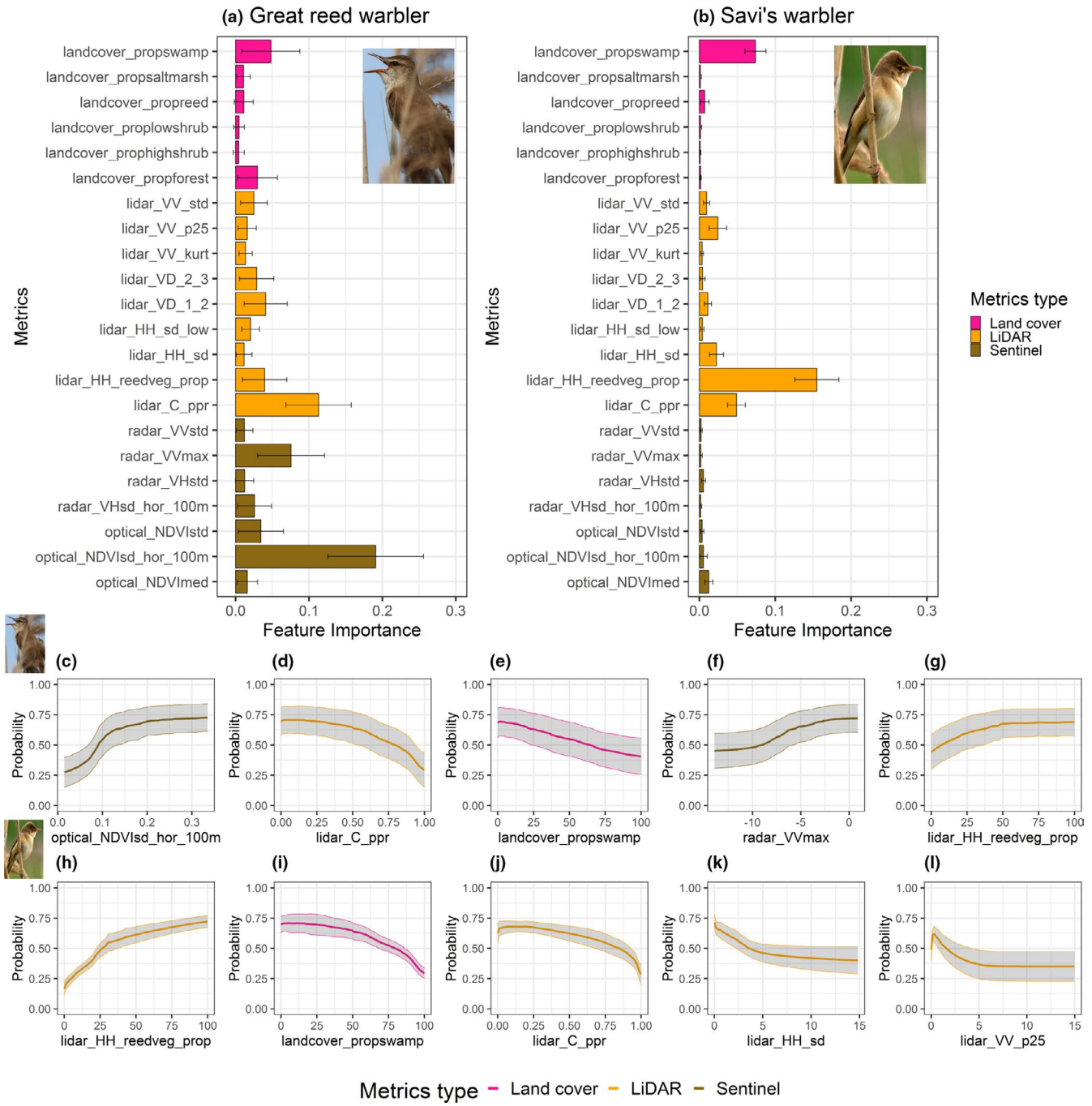

FIGURE 4 Importance of remote sensing metrics for predicting habitat suitability of two warbler species. (a and b) Feature importance of land cover (pink), LiDAR (orange) and Sentinel (brown) metrics for modelling the distribution of the great reed warbler and Savi's warbler, respectively. (c–l) Response curves of the five most important metrics for each species. Both the feature importance and response curves were calculated across 300 model runs using three algorithms (Generalized Linear Model, Maxent and Random Forest). Metric abbreviations are explained in Table 1. Photo credits: great reed warbler photo by Michele Lamberti (CC-BY), source: Flickr; Savi's warbler photograph by Ron Knight (CC-BY), source: Wikipedia

After adding predictor variables from LiDAR and Sentinel-1 and 2 products, the accuracy of SDMs improved by 11% and 10% in AUC for the great reed warbler and the Savi's warbler, respectively. This shows that predictive habitat mapping can be improved when variables other than land cover are added. Previous research from forest ecosystems supports this by showing that the addition of LiDAR and satellite RS metrics improve SDMs, either by adding statistically significant predictors or with increases in AUC values (Burns et al., 2020; Farrell et al., 2013; Ficetola et al., 2014; Müller et al., 2014; St-Louis et al., 2014; Tattoni et al., 2012). Nevertheless, our results also suggest that using only land cover maps with a high spatial resolution (e.g. 5 m) and a high thematic detail (e.g. six wetland-related land cover classes) can already achieve a fair to good overall predictive accuracy for breeding occupancy of

wetland birds. This finding is also confirmed by another study in the Netherlands which showed that a land cover map with 25 m resolution already provides an adequate representation of the habitat of a forest bird (crested tit) (Ficetola et al., 2014). Beyond land cover, our study further suggests that using Sentinel-1 and Sentinel-2 metrics can achieve similar improvements in model accuracy as additional LiDAR metrics. To our knowledge, this has not been demonstrated previously.

Our analysis of feature importance revealed that LiDAR metrics—contrary to our initial expectation—are not necessarily the most important variables in predicting habitat suitability of wetland birds. Previous studies suggested that LiDAR metrics are influential predictor variables in describing the habitat suitability of bird species, especially in forests because they provide direct information about the vertical and horizontal variation in vegetation structure (Bakx et al., 2019; Davies & Asner, 2014; Valbuena et al., 2020). In contrast, our study suggests that each type of RS product can substantially contribute to predicting habitat suitability, not only LiDAR. Moreover, our analysis shows that the most important variables differ among species. For instance, the most important predictor variable for the great reed warbler was a Sentinel metric (optical_NDVIsd_hor_100m), whereas for the Savi's warbler it was a LiDAR metric (lidar_HH_reedveg_prop). Various other studies confirm that the feature importance varies depending on the focal species because each species might prefer (or be adapted) to different habitat structures (Cord et al., 2014; Tattoni et al., 2012). In cases where species are highly dependent on specific fine-scale habitat elements, such as isolated trees, LiDAR metrics might turn out as the best predictor variables in SDMs (Tattoni et al., 2012).

Ecological field studies suggest that the great reed warbler prefers wetland habitats, which are structurally diverse, specifically with reed vegetation that is tall and close to open water (Dyrcz, 1981; Graveland, 1998). The response curves from our SDMs confirm that this species prefers a high habitat heterogeneity, represented by the horizontal variability of the annual median NDVI metric from Sentinel-2 (optical_NDVIsd_hor_100m). This metric measures the standard deviation of the annual median NDVI within a 100 m radius and has high values in structurally diverse and tall reedbeds, that is those that include some trees or shrubs besides reed and other herbaceous plants. Additionally, a LiDAR-based vegetation cover metric (lidar_C_ppr) was the second most important predictor variable for the great reed warbler, capturing dense reed vegetation (low lidar_C_ppr values) close to water (Koma, Seijmonsbergen, et al., 2021). Moreover, a high proportion of swamp within a 100 m radius as derived from land cover (landcover_propswamp) allowed to capture locations where the species tends to be absent because it does not prefer to breed within swamps (Cramp, 1992; Graveland, 1998). The high values of the maximum VV backscatter coefficient from Sentinel-1 (radar_VVmax) also revealed suitable habitats for the great reed warbler, maybe because the SAR signal reflects biomass variability of structurally diverse reedbeds (Slagter et al., 2020), or a high water content of water reed. Overall, the results confirm that both optical and SAR RS products from Sentinel can provide important information for quantifying habitat structure and vegetation heterogeneity of breeding birds in wetlands.

In contrast to the great reed warbler, the Savi's warbler prefers large, homogeneous reedbeds, which can be mixed with other helophyte vegetation, but only with a few bushes and trees (Báldi, 2006; Neto, 2006). This habitat preference is confirmed by the species' response to the proportion of reed vegetation (lidar_HH_reedveg_prop) and the 25th percentile of height (lidar_VV_25p), both derived from LiDAR. The results show that this species prefers areas that have (within a 100 m radius) a large proportion of vegetation with 1–3 m height. In addition, horizontal variability of vegetation height in a 100 m radius (lidar_HH_sd) has to be low, reflecting homogenous reedbeds or other similar types of wetland vegetation (Koma, Grootes, et al., 2021). Moreover, a high proportion of swamp derived from the land cover map (landcover_propswamp) indicated a low probability of occurrence of that species because it tends to avoid swamp vegetation for breeding (Cramp, 1992). The vegetation cover metric derived from LiDAR (lidar_C_ppr) indicates the density of reed, where the Savi's warbler prefers more open reed vegetation such as land reed (Koma, Seijmonsbergen, et al., 2021). Overall, the results of the Savi's warbler encourage the synergetic use of different RS products for modelling the habitat suitability of wetland birds.

5 | CONCLUSION

Our study shows that different metrics derived from high-resolution RS products (land cover maps, country-wide ALS, Sentinel-1 and Sentinel-2) can capture unique and complementary ecological information for modelling habitat suitability of wetland birds. While SDMs based on land cover maps with high thematic detail (e.g. many land cover classes) and high spatial resolution (e.g. 5 m) can already produce accurate predictions of habitat suitability, we demonstrate that the synergy of land cover maps with other metrics derived from LiDAR and Sentinel products can improve SDMs and thus enhance the ecological interpretation of animal–habitat relationships. We therefore expect that the increasing availability of high-resolution RS products will greatly improve predictive habitat and species distribution modelling across a range of taxa and ecosystems. We further suggest that future studies should combine metrics from land cover maps, LiDAR, Sentinel-1 and Sentinel-2 to yield a comprehensive understanding of animal–habitat relationships.

ACKNOWLEDGEMENTS

This work is a part of the project 'eScience infrastructure for Ecological applications of LiDAR point clouds' (eEcoLiDAR) (Kissling et al., 2017), funded by the Netherlands eScience Center (https://www.esciencecenter.nl), grant number ASDI.2016.014.

CONFLICT OF INTEREST

None declared.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13468.

DATA AVAILABILITY STATEMENT

All scripts for processing the different remote sensing products, and the R scripts to reproduce the statistical analyses, are available on GitHub (https://github.com/komazsofi/LiDAR_Sentinel_birdsdm_wetlands). The calculated remote sensing metrics (raster layers) are available on Zenodo (https://doi.org/10.5281/zenodo.5772673).

ORCID

Zsófia Koma https://orcid.org/0000-0002-0003-8258

REFERENCES

Altman, D. G. (1990). *Practical statistics for medical research*. Chapman and Hall/CRC.

Bae, S., Levick, S. R., Heidrich, L., Magdon, P., Leutner, B. F., Wöllauer, S., Serebryanyk, A., Nauss, T., Krzystek, P., Gossner, M. M., Schall, P., Heibl, C., Bässler, C., Doerfler, I., Schulze, E.-D., Krah, F.-S., Culmsee, H., Jung, K., Heurich, M., … Müller, J. (2019). Radar vision in the mapping of forest biodiversity from space. *Nature Communications*, *10*(1), 4757. https://doi.org/10.1038/s41467-019-12737-x

Bakx, T. R. M., Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2019). Use and categorization of light detection and ranging vegetation metrics in avian diversity and species distribution research. *Diversity and Distributions*, *25*(7), 1045–1059. https://doi.org/10.1111/ddi.12915

Báldi, A. (2006). Factors influencing occurrence of passerines in the reed archipelago of Lake Velence (Hungary). *Acta Ornithologica*, *41*(1), 1–6. https://doi.org/10.3161/068.041.0105

Báldi, A., & Kisbenedek, T. (1999). Species-specific distribution of reed-nesting passerine birds across reed-bed edges: Effects of spatial scale and edge type. *Acta Zoologica Academiae Scientiarum Hungaricae*, *45*(2), 97–114.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Burns, P., Clark, M., Salas, L., Hancock, S., Leland, D., Jantz, P., Dubayah, R., & Goetz, S. J. (2020). Incorporating canopy structure from simulated GEDI lidar into bird species distribution models. *Environmental Research Letters*, *15*(9), 95002. https://doi.org/10.1088/1748-9326/ab80ee

Cody, M. L. (1985). *Habitat Selection in Birds*. Academic Press.

Cord, A. F., Klein, D., Mora, F., & Dech, S. (2014). Comparing the suitability of classified land cover data and remote sensing variables for modeling distribution patterns of plants. *Ecological Modelling*, *272*, 129–140. https://doi.org/10.1016/j.ecolmodel.2013.09.011

Cramp, S. (the late) (Ed.). (1992). *Handbook of the Birds of Europe, the Middle East, and North Africa: The Birds of the Western Palearctic Volume VI: Warblers*. Oxford University Press.

Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in Ecology & Evolution*, *29*(12), 681–691. https://doi.org/10.1016/j.tree.2014.10.005

Davies, A. B., Gaylard, A., & Asner, G. P. (2018). Megafaunal effects on vegetation structure throughout a densely wooded African landscape. *Ecological Applications*, *28*(2), 398–408. https://doi.org/10.1002/eap.1655

de Vries, J. P. R., Koma, Z., WallisDeVries, M. F., & Kissling, W. D. (2021). Identifying fine-scale habitat preferences of threatened butterflies using airborne laser scanning. *Diversity and Distributions*, *27*(7), 1251-1264. https://doi.org/10.1111/ddi.13272

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Dunlavy, J. C. (1935). Studies on the phyto-vertical distribution of birds. *The Auk*, *52*(4), 425–431. https://doi.org/10.2307/4077518

Dyrcz, A. (1981). Breeding ecology of great reed warbler *Acrocephalus arundinaceus* and reed warbler *Acrocephalus scirpaceus* at fish-ponds in SW Poland and lakes in NW Switzerland. *Acta Ornithologica*, *18*, 307–334.

Elith, J., Ferrier, S., Huettmann, F., & Leathwick, J. (2005). The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. *Ecological Modelling*, *186*(3), 280–289. https://doi.org/10.1016/j.ecolmodel.2004.12.007

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., … E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Farrell, S. L., Collier, B. A., Skow, K. L., Long, A. M., Campomizzi, A. J., Morrison, M. L., Hays, K. B., & Wilkins, R. N. (2013). Using LiDAR-derived vegetation metrics for high-resolution, species distribution models for conservation planning. *Ecosphere*, *4*(3), art42. https://doi.org/10.1890/ES12-000352.1

Farwell, L. S., Gudex-Cross, D., Anise, I. E., Bosch, M. J., Olah, A. M., Radeloff, V. C., Razenkova, E., Rogova, N., Silveira, E. M. O., Smith, M. M., & Pidgeon, A. M. (2021). Satellite image texture captures vegetation heterogeneity and explains patterns of bird richness. *Remote Sensing of Environment*, *253*, 112175. https://doi.org/10.1016/j.rse.2020.112175

Ficetola, G. F., Bonardi, A., Mücher, C. A., Gilissen, N. L. M., & Padoa-Schioppa, E. (2014). How many predictors in species distribution models at the landscape scale? Land use versus LiDAR-derived canopy height. *International Journal of Geographical Information Science*, *28*(8), 1723–1739. https://doi.org/10.1080/13658816.2014.891222

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone–*Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Graveland, J. (1998). Reed die-back, water level management and the decline of the Great Reed Warbler *Acrocephalus arundinaceus* in the Netherlands. *Ardea*, *86*(2), 187–201.

Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, *157*(2–3), 89–100. https://doi.org/10.1016/S0304-3800(02)00204-1

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

He, K. S., Bradley, B. A., Cord, A. F., Rocchini, D., Tuanmu, M., Schmidtlein, S., Turner, W., Wegmann, M., & Pettorelli, N. (2015). Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation*, *1*(1), 4–18. https://doi.org/10.1002/rse2.7

Kissling, W. D., Seijmonsbergen, A., Foppen, R., & Bouten, W. (2017). eEcoLiDAR, eScience infrastructure for ecological applications of LiDAR point clouds: Reconstructing the 3D ecosystem structure for animals at regional to continental scales. *Research Ideas and Outcomes*, *3*, e14939. https://doi.org/10.3897/rio.3.e14939

Koma, Z., Grootes, M. W., Meijer, C. W., Nattino, F., Seijmonsbergen, A. C., Sierdsema, H., Foppen, R., & Kissling, W. D. (2021). Niche separation of wetland birds revealed from airborne laser scanning. *Ecography*, *44*(6), 907–918. https://doi.org/10.1111/ecog.05371

Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2021). Classifying wetland-related land cover types and habitats using fine-scale lidar metrics derived from country-wide Airborne Laser Scanning. *Remote Sensing in Ecology and Conservation*, *7*(1), 80–96. https://doi.org/10.1002/rse2.170

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Leisler, B., & Schulze-Hagen, K. (2011). *The reed warblers: diversity in a uniform bird family*. KNNV Publishing.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*(2), 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Lucas, C., Bouten, W., Koma, Z., Kissling, W. D., & Seijmonsbergen, A. C. (2019). Identification of linear vegetation elements in a rural landscape using LiDAR point clouds. *Remote Sensing*, *11*(3), 292. https://doi.org/10.3390/rs11030292

MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, *42*(3), 594–598. https://doi.org/10.2307/1932254

Meijer, C., Grootes, M. W., Koma, Z., Dzigan, Y., Gonçalves, R., Andela, B., van den Oord, G., Ranguelova, E., Renaud, N., & Kissling, W. D. (2020). Laserchicken—A tool for distributed feature calculation from massive LiDAR point cloud datasets. *SoftwareX*, *12*, 100626. https://doi.org/10.1016/j.softx.2020.100626

Mod, H. K., Scherrer, D., Luoto, M., & Guisan, A. (2016). What we use is not what we know: Environmental predictors in plant distribution models. *Journal of Vegetation Science*, *27*(6), 1308–1322. https://doi.org/10.1111/jvs.12444

Müller, J., Bae, S., Röder, J., Chao, A., & Didham, R. K. (2014). Airborne LiDAR reveals context dependence in the effects of canopy architecture on arthropod diversity. *Forest Ecology and Management*, *312*, 129–137. https://doi.org/10.1016/j.foreco.2013.10.014

Naimi, B., & Araújo, M. B. (2016). sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*(4), 368–375. https://doi.org/10.1111/ecog.01881

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, *37*(2), 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Neto, J. M. (2006). Nest-site selection and predation in Savi's Warblers *Locustella luscinioides*. *Bird Study*, *53*(2), 171–176. https://doi.org/10.1080/00063650609461430

Oeser, J., Heurich, M., Senf, C., Pflugmacher, D., Belotti, E., & Kuemmerle, T. (2020). Habitat metrics based on multi-temporal Landsat imagery for mapping large mammal habitat. *Remote Sensing in Ecology and Conservation*, *6*(1), 52–69. https://doi.org/10.1002/rse2.122

Pearson, R. G., Dawson, T. P., & Liu, C. (2004). Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. *Ecography*, *27*(3), 285–298. https://doi.org/10.1111/j.0906-7590.2004.03740.x

Peterson, A. T., Papeş, M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, *213*(1), 63–72. https://doi.org/10.1016/j.ecolmodel.2007.11.008

Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., & Turner, W. (2014). Satellite remote sensing for applied ecologists: Opportunities and challenges. *Journal of Applied Ecology*, *51*(4), 839–848. https://doi.org/10.1111/1365-2664.12261

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jędrzejewska, B., Lima, M., & Kausrud, K. (2011). The Normalized Difference Vegetation Index (NDVI): Unforeseen successes in animal ecology. *Climate Research*, *46*(1), 15–27. https://doi.org/10.3354/cr00936

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, *20*(9), 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Randin, C. F., Ashcroft, M. B., Bolliger, J., Cavender-Bares, J., Coops, N. C., Dullinger, S., Dirnböck, T., Eckert, S., Ellis, E., Fernández, N., Giuliani, G., Guisan, A., Jetz, W., Joost, S., Karger, D., Lembrechts, J., Lenoir, J., Luoto, M., Morin, X., … Payne, D. (2020). Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment*, *239*, 111626. https://doi.org/10.1016/j.rse.2019.111626

Schulte to Bühne, H., & Pettorelli, N. (2018). Better together: Integrating and fusing multispectral and radar satellite imagery to inform biodiversity monitoring, ecological research and conservation science. *Methods in Ecology and Evolution*, *9*(4), 849–865. https://doi.org/10.1111/2041-210X.12942

Sinha, S., Jeganathan, C., Sharma, L. K., & Nathawat, M. S. (2015). A review of radar remote sensing for biomass estimation. *International Journal of Environmental Science and Technology*, *12*(5), 1779–1792. https://doi.org/10.1007/s13762-015-0750-0

Slagter, B., Tsendbazar, N.-E., Vollrath, A., & Reiche, J. (2020). Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, *86*, 102009. https://doi.org/10.1016/j.jag.2019.102009

St-Louis, V., Pidgeon, A. M., Kuemmerle, T., Sonnenschein, R., Radeloff, V. C., Clayton, M. K., Locke, B. A., Bash, D., & Hostert, P. (2014). Modelling avian biodiversity using raw, unclassified satellite imagery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1643), 20130197. https://doi.org/10.1098/rstb.2013.0197

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

Tattoni, C., Rizzolli, F., & Pedrini, P. (2012). Can LiDAR data improve bird habitat suitability models? *Ecological Modelling*, *245*, 103–110. https://doi.org/10.1016/j.ecolmodel.2012.03.020

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures. *Journal of Biogeography*, *31*(1), 79–92. https://doi.org/10.1046/j.0305-0270.2003.00994.x

Thuiller, W., Araújo, M. B., & Lavorel, S. (2004). Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography*, *31*(3), 353–361. https://doi.org/10.1046/j.0305-0270.2003.00991.x

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B. Ö., Floury, N., Brown, M., Traver, I. N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., … Rostan, F. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An r package for generating spatially or environmentally

separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, *10*(2), 225–232. https://doi.org/10.1111/2041-210X.13107

Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., Silva, C. A., Almeida, D. R. A., Danks, F., Morsdorf, F., Chirici, G., Lucas, R., Coomes, D. A., & Coops, N. C. (2020). Standardizing ecosystem morphological traits from 3D information sources. *Trends in Ecology & Evolution*, *35*(8), 656–667. https://doi.org/10.1016/j.tree.2020.03.006

Vergeer, J.-W., van Dijk, A., Boele, A., van Bruggen, J., & Hustings, F. (2016). *Handleiding Sovon Broedvogelonderzoek: Broedvogel Monitoring Project en Kolonievogels*. Sovon Vogelonderzoek Nederland. www.sovon.nl

Vierling, K. T., Vierling, L. A., Gould, W. A., Martinuzzi, S., & Clawges, R. M. (2008). Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, *6*(2), 90–98. https://doi.org/10.1890/070001

Zellweger, F., Baltensweiler, A., Ginzler, C., Roth, T., Braunisch, V., Bugmann, H., & Bollmann, K. (2016). Environmental predictors of species richness in forest landscapes: Abiotic factors versus vegetation structure. *Journal of Biogeography*, *43*(6), 1080–1090. https://doi.org/10.1111/jbi.12696

Zellweger, F., Braunisch, V., Baltensweiler, A., & Bollmann, K. (2013). Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. *Forest Ecology and Management*, *307*, 303–312. https://doi.org/10.1016/j.foreco.2013.07.023

Zellweger, F., Frenne, P. D., Lenoir, J., Rocchini, D., & Coomes, D. (2019). Advances in microclimate ecology arising from remote sensing. *Trends in Ecology & Evolution*, *34*(4), 327–341. https://doi.org/10.1016/j.tree.2018.12.012

BIOSKETCH

The research team is interested in understanding the spatial distribution of biodiversity and the habitat preferences and ecological niches of species, using a combination of in situ (species) observations and remote sensing.

Author Contributions: Conceptualization: ZK, ACS and WDK; methodology: ZK and WDK; data processing: ZK, MWG, FN and JG; data curation: ZK, HS and RPBF; analysis: ZK; writing initial draft: ZK; review and editing: WDK, ACS, HS and RPBF; supervision: WDK and ACS.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Koma, Z., Seijmonsbergen, A. C., Grootes, M. W., Nattino, F., Groot, J., Sierdsema, H., Foppen, R. P. B., & Kissling, W. D. (2022). Better together? Assessing different remote sensing products for predicting habitat suitability of wetland birds. *Diversity and Distributions*, 28, 685–699. https://doi.org/10.1111/ddi.13468