A framework for near-real time monitoring of diversity patterns based on indirect remote sensing, with an application in the Brazilian Atlantic rainforest

Andrea Paz[1,2,3], Thiago S. Silva[4,5] and Ana C. Carnaval[1,2]

[1] Department of Biology, City College of New York, New York, NY, United States of America
[2] Ph.D Program in Biology, City University of New York, Graduate School and University Center, New York, NY, United States of America
[3] Department of Environmental Systems Science, Institute of Integrative Biology, Swiss Federal Institute of Technology, Zurich, Zürich, Switzerland
[4] Instituto de Geociências e Ciências Exatas, Departamento de Geografia, Ecosystem Dynamics Observatory, Universidade Estadual Paulista, Rio Claro, São Paulo, Brazil
[5] Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom

ABSTRACT

Monitoring biodiversity change is key to effective conservation policy. While it is difficult to establish *in situ* biodiversity monitoring programs at broad geographical scales, remote sensing advances allow for near-real time Earth observations that may help with this goal. We combine periodical and freely available remote sensing information describing temperature and precipitation with curated biological information from several groups of animals and plants in the Brazilian Atlantic rainforest to design an indirect remote sensing framework that monitors potential loss and gain of biodiversity in near-real time. Using data from biological collections and information from repeated field inventories, we demonstrate that this framework has the potential to accurately predict trends of biodiversity change for both taxonomic and phylogenetic diversity. The framework identifies areas of potential diversity loss more accurately than areas of species gain, and performs best when applied to broadly distributed groups of animals and plants.

Subjects Biodiversity, Conservation Biology, Ecology, Climate Change Biology, Spatial and Geographic Information Science
Keywords Biodiversity, Prediction, Monitoring, Richness, Phylogenetic diversity

Submitted 17 February 2022
Accepted 12 May 2022
Published 29 June 2022

Corresponding author
Andrea Paz, paz.andreita@gmail.com

Academic editor
Rodolfo Jaffé

Additional Information and Declarations can be found on page 18

DOI 10.7717/peerj.13534

Copyright
2022 Paz et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

INTRODUCTION

Real-time monitoring of biodiversity is a fundamental goal of biological conservation. If implemented at large geographical scales, it can strongly impact policy and action. Yet the difficulty of establishing real- or near-real time diversity monitoring on the ground poses a significant challenge for large-scale conservation planning and management (*Scholes et al., 2008*; *Navarro et al., 2017*). The increased availability of free remote sensing data, encompassing broad geographic extents at fine spatial and temporal resolution,

How to cite this article Paz A, Silva TS, Carnaval AC. 2022. A framework for near-real time monitoring of diversity patterns based on indirect remote sensing, with an application in the Brazilian Atlantic rainforest. *PeerJ* **10**:e13534 http://doi.org/10.7717/peerj.13534

makes the promise of real- or near-real time monitoring of biodiversity from space more attainable (*Scholes et al., 2012*; *Turner, 2014*; *Reddy, 2021*). As such, research programs and services that automatically process periodically available satellite images to flag land-use change have begun to take advantage of this potential. The Global Forest Watch and the Amazon DETER initiative (*World Resources Institute, 2014*; *Diniz et al., 2015*) are examples of flagship programs that monitor deforestation and forest fires, issuing early alerts for conservation and providing examples of what can now be done in biodiversity monitoring.

Yet satellite imagery can only view so much—and often cannot provide species level identifications or information below the canopy level. Most proposals to measure essential biodiversity variables from remote sensing include only direct measurements such as canopy tree observations and, so far, do not integrate with the genetic dimension of diversity (*Reddy, 2021*; *Skidmore et al., 2021*). Even aerial photographs of canopy trees can fail to provide accurate identifications at the species level, particularly in the tropics (*Garzon-Lopez et al., 2013*). Although new techniques such as hyperspectral satellite imagery show better taxonomic discrimination, they can still only sense canopy plants or lower vegetation in open areas (*Papeş et al., 2010*). Furthermore, observations of animals are often unfeasible, or limited (but see *Fretwell & Trathan (2009)* for animal traces that may be detected from satellites in depauperate ecosystems, and *Schmaljohann (2020)* for an recent example of detection of animal movement with radar data). These limitations raise the question of whether and how species or general diversity patterns that are not directly observable from space may still be monitored remotely.

To overcome these challenges, it has been proposed that organisms such as understory plants and animals may be monitored indirectly, by tracking the environmental conditions that are statistically associated with their presence (*Turner et al., 2003*; *Turner, 2014*; *Paz et al., 2020*). Several studies have identified statistical relationships between metrics that quantify local biodiversity (*Rompré et al., 2007*; *Laurencio & Fitzgerald, 2010*; *Peters et al., 2016*; *Zellweger et al., 2016*) and environmental variables describing local climates (temperature and precipitation), or landscapes. Not only has the number of biological species (henceforth referred to as species richness) been related to climate and habitat configuration, but so have metrics that reflect the diversity of evolutionary lineages present in a given region, also known as phylogenetic diversity. Phylogenetic diversity combines information about which species are in a given region and how closely or distantly related they are (*Faith, 1992*). Measuring both dimensions of diversity allows for the identification of regions of mismatch between species richness and phylogenetic diversity, which indicate that the assembly of the local community was not random with respect to the species pool. For instance, when a given environment selects for species that are closely related to each other and therefore share crucial adaptations for that area (also referred to as environmental filtering, *Webb et al., 2002*), phylogenetic diversity will likely be lower than expected given the number of species. In contrast, whenever inter-specific competition is prevalent, species that are closely related to each other and fulfill similar ecological roles may not be able to co-exist; under these conditions, local species may be more distantly related than expected, and phylogenetic diversity would be high relative to species richness (*Webb et al., 2002*). Sometimes, however, these mismatches can be misleading: when distantly

related species share similar traits due to convergence (as opposed to common ancestry), then environmental filtering may be associated with phylogenetic overdispersion instead of clustering. Moreover, when competitive abilities are very different, then competition can lead to exclusion of species of lower competitive ability and thus only presence of species of the higher competitive ability with similar traits (*Cavender-Bares et al., 2009*; *Mayfield & Levine, 2010*).

Because many of the environmental variables relevant to models of diversity patterns under a static environment can be derived from remote sensing data, it has been proposed that near-real time indirect remote sensing may provide a lens to describe and predict the potential displacement of biodiversity in response to continuous environmental shifts (*Paz et al., 2021*; *Turner et al., 2003*). In other words, documentation of changes in environmental conditions *via* remote sensing may allow for the dynamic monitoring of changes of biodiversity components that cannot be directly observed. However, a platform that applies these concepts and allows governments and scientists to envision how ongoing environmental changes may potentially change the distribution of diversity, in near-real-time, is missing. Importantly, it remains to be uncovered how reliably this approach can generate predictions across different geographical areas, and for biological groups of varying levels of diversity, ecological constraints, and geographical range.

Here, we build and test such a platform for near-real time prediction of diversity change. By applying it to environmental and biodiversity data from the Brazilian Atlantic Forest hotspot, we critically evaluate its limitations and the conditions under which it may prove most useful. We focus on the Atlantic Forest to build on our previous studies (*Paz et al., 2020*; *Paz et al., 2021*), which demonstrated that climate variables can accurately predict local species richness and phylogenetic diversity at the community level. Here, we apply this new modeling framework to those two measures of biodiversity in the Atlantic Forest.

This new model for biodiversity monitoring through indirect remote sensing is built and implemented in a reproducible machine-learning framework that integrates Google Earth Engine (*Gorelick et al., 2017*), the R software (*R Core Team, 2022*), and data describing the diversity of nine taxonomic groups in the Atlantic forest known to differ in their geographic spread and sampling intensity. The input data include per-pixel descriptors of species richness and phylogenetic diversity of five groups of plants and four groups of animals. All groups have been well sampled throughout their range, adequately geo-referenced, and correctly identified by experts (*Paz et al., 2021*). By gathering and processing remote sensing data to describe local environmental conditions, and by integrating them with the input biological data (here, maps of species richness and phylogenetic diversity), the modeling framework allows the user to (i) describe potential changes in the distribution of species richness and phylogenetic diversity in near-real time, and (ii) predict potential trends in biodiversity change given changes in environments. These trends are predicted only within the region of model training, and within the 2002–2019 time interval. As such, it is here limited to the Atlantic Forest of Brazil, and for that period of time.

Because machine learning techniques are known to need a substantial amount of data (*Van Der Ploeg, Austin & Steyerberg, 2014*; *Kuhn & Johnson, 2016*), and given that species with smaller ranges might be more prone to be affected by history or chance

(*Paz et al., 2020*), we expect that those groups of organisms that have broad geographical ranges, and thus a high number of observations to train the model from, will be best suited for this modeling exercise. As such, we first apply the model to nine groups that differ in their geographical spread, and evaluate how sensitive it is to these distinct geographical ranges. Then, to provide a more stringent test of the modeling platform, we verify how accurately the model can retroactively predict changes in species richness that have been documented in the Atlantic Forest in decades past. Specifically, we use the model to reconstruct potential species loss and gain in two groups of organisms known to be closely impacted by climate, and which have been historically inventoried in this region. To that end, we use two databases that gather inventory data on species presence (but not phylogenetic diversity): ATLANTIC-epiphytes (*Ramos et al., 2019*) and the ATLANTIC-amphibians database (*Vancine et al., 2018*).

METHODS

To illustrate our modeling framework and evaluate its sensitivity to the geographical range of the group being modeled, we first applied it to nine organismal groups that differ in their spatial distribution. To this end, for each group, we combined its species richness map with temporally-explicit environmental data (temperature and precipitation), derived from remote sensing sources, to generate a model that describes how they are associated. We then performed the same procedure with the group's phylogenetic diversity map, to model this second biodiversity metric as well. We evaluated model fit for all groups and for both diversity metrics, exploring how they relate to the total range size of the group and the median individual species range size. Models with the best fit were then projected over time, yearly, based on that year's environmental conditions. Those yearly projections allowed us to map trends in diversity change across time (per-pixel loss or gain of species richness, and per-pixel loss or gain of phylogenetic diversity), for each group. They also allowed us to map areas of potential mismatch between species richness and phylogenetic diversity—which often indicate important ecological processes—and identify whether they may be changing over time. Finally, we applied our framework to an available dataset of observations of amphibians and epiphytes over time, to test the accuracy of the model predictions.

a. Biological datasets and maps of species richness and phylogenetic diversity

To illustrate our biodiversity monitoring model and evaluate its sensitivity to the geographical range of the group, we first applied it to nine groups that differ in their spatial distribution (five groups of plants, and four groups of animals; Table 1). The plants included a group within the Bromelioideae subfamily (*Aguirre-Santoro, Stevenson & Michelangeli, 2016*; *Aguirre-Santoro, 2017*; *Brown et al., 2020*), three groups in the family Melastomataceae [in the Miconieae tribe (*Goldenberg et al., 2008*; *Michelangeli et al., 2008*; *Caddah, 2013*; *Reginato & Michelangeli, 2016*; *Brown et al., 2020*), in the genus *Bertolonia* (*Bacci et al., 2020*; *Paz et al., 2021*), and in the tribe Cambessedesieae (*Bochorny et al., 2019*; *Paz et al., 2021*)], and a group including *Fridericia* and allies (*Kaehler, Michelangeli &*

Table 1 Biological data used for modeling. Biological data used to model species richness and phylogenetic diversity (five plant groups and four animal groups). Table includes the median area of distribution of species within each group, and the total area of distribution of each group. Data from *Paz et al. (2021)*.

Clade	Taxonomic level	Median individual area of distribution (km^2)	Total area of distribution for the group in the AF (km^2)
Thraupinae	Subfamily	680,022	1,330,613
Fridericia & allies	Group	85	1,282,733
Miconieae	Tribe	44,486	1,197,890
Ithomiini	Tribe	321,331	1,147,465
Boana	Genus	240	824,368
Proceratophrys	Genus	3,266	766,577
Cambessedesieae	Tribe	1,887	755,871
Bertolonia	Genus	1,970	240,376
Bromelioideae	Subfamily	41	195,639

Lohmann, 2019; *Paz et al., 2021*), in the plant family Bignoniaceae. The animals included the clearwing butterflies that compose the tribe Ithomiini (*Brown et al., 2020*), treefrogs of the genus *Boana* (*Vasconcelos et al., 2014*; *Brown et al., 2020*), horned frogs of the genus *Proceratophrys* (*Brown et al., 2020*), and a group of birds in the subfamily Thraupinae (*Burns et al., 2014*). These groups and their species differ in their distributions: some have small individual species distributions and small total geographical spread (like the Bromelioideae plants and the *Boana* frogs), others have broad individual species distributions and broad group-wide spatial coverage (like the Thraupinae), while others have discordant patterns (like the the Fridericia and allies, with small individual ranges but a broad group coverage; see Table 1, sorted from broader geographic distribution to smallest).

For each group, a map of species richness was downloaded from the Dryad repository of *Paz et al. (2021)*; here, we provide a brief description of the methods used in that publication. Those maps were originally based on individual species distribution maps developed at a 10 km resolution (*Paz et al., 2021*), and created through the superimposition of alpha hulls representing the ranges of each species. Alpha hulls are a generalization of convex hulls, allowing for more shape flexibility, with lines that are allowed to curve and match the point distributions more closely (*Meyer, Diniz-Filho & Lohmann, 2017*). Once alpha hulls representing the ranges of each species in group were obtained, they were stacked and summed to obtain a species richness map, which was used in the present analysis (*Paz et al., 2021*). As any outline-based maps, alpha hulls may overpredict individual species ranges; however, they have been shown to do so less often than other commonly used techniques, such as minimum convex. Although they may lead to overprediction of species richness, this issue is more prevalent at smaller resolutions (*Graham & Hijmans, 2006*; *Hurlbert & Jetz, 2007*).

For each group, a map of phylogenetic diversity, as measured with Faith's PD (*Faith, 1992*), was also downloaded from *Paz et al. (2021)*. To obtain that map of the spatial distribution of phylogenetic diversity, the species richness map and community composition information of each group were previously combined with phylogenetic information describing the evolutionary relationships among all species in the group,

using the software Biodiverse (*Laffan, Lubarsky & Rosauer, 2010*) as described in *Paz et al. (2021)*.

b. Environmental descriptors

Environmental data to train and project the near-real time models were obtained from remote sensing sources and used to create bioclimatic layers describing the conditions experienced at different time windows. For that, we collected Normalized Difference Vegetation Index (NDVI) data and temperature data from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS; *Vermote & Wolfe, 2015*) instrument, and precipitation data from the Climate Hazards Infrared Precipitation with Stations (CHIRPS; *Funk et al., 2015*) dataset. The MODIS data products are available bi-weekly since 2002, at a minimum resolution of 500m and 1km respectively; CHIRPS data are available monthly between 1981 and the present, at a minimum resolution of 0.05 degrees (∼5.5 km). Using a custom-created python script for Google Earth Engine (GEE, *Gorelick et al., 2017*), we obtained all NDVI and temperature MODIS data and CHIRPS precipitation data representing conditions experienced in the Atlantic Forest from 2002 to the present, using the forest shapefile from *Paz et al. (2021)*. In GEE, we summarized climate data as monthly means, maximum and minimum values and NDVI data as yearly medians at a 10 km resolution. All summary layers were downloaded into R for further processing and for model building. Based on those monthly data, sets of 19 bioclimatic layers representing summaries of temperature and precipitation (Table 2) were built using the *bioclim* function of the dismo package (*Hijmans et al., 2011*). To calibrate the machine learning model, we used built bioclimatic and NDVI layers for the 2002–2014 time period; for projections of the model, we used yearly layers from 2015 to 2019. The python script to summarize and download data from GEE, as well as the R scripts used to generate the bioclimatic layers, are available on the Github repository of this manuscript (https://github.com/andrepazv/NearRealTime). Before proceeding to modeling, we removed all highly redundant (correlated) variables (VIF > 10) using the *vifstep* function of the usdm R package (*Naimi et al., 2014*). This reduced the environmental dataset to ten variables that were used in all further analyses (Table 2).

c. The machine learning model and time projections

To predict the spatial distribution of species richness and phylogenetic diversity as a function of environmental change in near-real time, we independently modeled species richness and phylogenetic diversity with data describing temperature and precipitation. Those models with the best fit were then projected over time, yearly, based on each year's environmental conditions.

We first trained an ensemble machine learning model to describe species richness, and an ensemble machine learning model to describe phylogenetic diversity, for each of the nine groups (Table 1). For that, we used bioclimatic layers built from MODIS and CHIRPS and NDVI estimates (Table 2) summarized for the 2002–2014 time period. To improve model performance, we performed an ensemble of four machine learning algorithms: Random Forests (rf from *Liaw & Wiener, 2002*), Neural Networks (nnet from *Venables & Ripley,*

Table 2 **Variables used in the machine learning modeling and predictions.** A total of 19 bioclimatic variables were built for training and projecting periods using the dismo package in R. Yearly NDVI values were also included. Monthly temperature and precipitation data were obtained from remote sensing sources as described in the text. After removing all highly correlated variables, the following ten variables were used for model building.

Variable Name	Description
NDVI	Normalized difference vegetation index
bio2	Mean Diurnal Range (Mean of monthly (max temp–min temp))
bio3	Isothermality
bio4	Temperature Seasonality (standard deviation *100)
bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter

2002), support vector machines (svmRadial from *Karatzoglou et al., 2004*), and generalized linear models. These algorithms allow the prediction of continuous variables (as a form of regression), but the first three make no assumption of linearity in the relationship between predictor and response variables, thus allowing for more complex and biologically relevant relationships. Seeking good transferability across time projections, we used a spatial partition method for model tuning and validation (*Radosavljevic & Anderson, 2014*): for each model, we partitioned the biodiversity data (pixels) into geographic regions, instead of randomly. To do that, for each biodiversity metric (species richness or phylogenetic diversity) and taxon (each one of the nine groups of organisms), we created four spatial partitions using an adapted version of the function *get.block* in the package ENMeval, partitioning the data into bins across latitude and longitude lines (provided with the full script on GitHub; *Muscarella et al., 2014*). For each one of the four machine learning algorithms, we employed k-fold cross validation, building $k = 4$ models, each leaving out one different spatial partition that was used for model validation. Models were trained and tuned using the caretList function in the *caretEnsemble R* package, with partitions passed through the trainControl argument (*Deane-Mayer & Knowles, 2019*). Individual tuned models were selected using the root mean square error (RMSE) metric. Using the settings determined as optimal in the validation stage, we then built a full model with all data (full model from hereon). Finally, an ensemble model was built for each organismal group and biodiversity metric, based on a linear combination of the models for all four algorithms, weighted according to their RMSE values. For that, we used the caretEnsemble R package and repeated 10-fold cross-validation (*Deane-Mayer & Knowles, 2019*). The script used for modeling is available in the GitHub repository for this manuscript.

These models allow for predictions of the response variable (here, species richness and phylogenetic diversity) to be transferred to a subsequent time period based on the same predictor variables (the bioclimatic variables + NDVI, Table 2, Fig. S4). To obtain an

estimate of trends in biodiversity change over time, we projected the models in time over five different years, generating species richness and phylogenetic diversity projections for 2015, 2016, 2017, 2018 and 2019. For this, we applied the corresponding model to each year's predictor variables. Using the model and the five projections over time (from 2015 to 2019), we stacked the maps together and computed the slope of a linear regression for each pixel to estimate trends in biodiversity change for each group and biodiversity metric. To obtain a spatial estimate of the expected mismatch between the biodiversity metrics, we computed the regression of phylogenetic diversity onto species richness for each year, and mapped the residuals of the regression for each year.

d. Ground truthing the modeling framework

We then used another dataset to ground-truth the model framework, contrasting the biodiversity change trends inferred by the model against trends documented in the field. For that, we used information available for two other groups of organisms—epiphytes and amphibians—for which survey data are available across multiple years and across a sufficient percentage of their ranges, using the ATLANTIC-epiphytes (*Ramos et al., 2019*) and ATLANTIC-amphibian datasets (*Vancine et al., 2018*). These datasets were made possible by a multi-institutional effort that gathered inventory data for different plant and animal groups, collected across several decades along the Atlantic Forest. Although geographical sampling along the years is sparse, these two datasets have survey information that spans most of the latitudinal extent of the Atlantic Forest and includes contemporary sampling that partly matches the available satellite information. The epiphyte dataset was composed of 89,270 observations from 1824 to 2018 (*Ramos et al., 2019*). That of the amphibians was composed of 17,619 records, with data from 1940 up to 2015 (*Vancine et al., 2018*). For both datasets, we created maps of species richness for both datasets at a resolution of 100 km. This resulted in maps of 188 and 137 pixels, respectively, with species richness per pixel varying from 1–365 epiphyte species and from 1–80 amphibian species. Because the ATLANTIC datasets do not include phylogenetic data, this analysis focused solely on species richness and not phylogenetic diversity.

For the purposes of ground-truthing the model, we limited our analyses to the years for which the datasets (i) documented the year of the observations, (ii) provided coordinates for the observation localities and (iii) provided a relatively broad sampling coverage of the region (*i.e.,* sampled sites were not clustered geographically). This resulted in the use of data collected up to the year of 2015 for the amphibians, and up to the year of 2018 for the epiphytes. We then used the available inventory data from all but the last four years of observations to generate maps of species richness (up to the year of 2011 for the amphibians, and 2014 for the epiphytes), measured as the number of species present in each pixel in each year. Following the same modeling procedure applied previously, we combined those species richness maps with bioclimatic variables derived from remote sensing sources as described above, and for the same period, to train a predictive model. The model was then projected onto the four following (most recent) years, for which observational data are available for ground-truthing. As such, we obtained yearly model projections of amphibian species richness for the years of 2012, 2013, 2014, and 2015, and

Table 3 R^2 of predictive models of species richness and phylogenetic diversity, based on environmental data, developed for nine target taxonomic groups.

Clade	R^2 of predictive model for species richness	R^2 of predictive model for phylogenetic diversity
Fridericia & allies	0.83	0.71
Thraupinae	0.72	0.68
Miconieae	0.71	0.74
Ithomiini	0.56	0.59
Bromelioideae	0.55	0.49
Boana	0.53	0.52
Cambessedesieae	0.47	0.34
Proceratophrys	0.29	0.08
Bertolonia	0.18	0.46

yearly model projections of epiphyte species richness for the years of 2015, 2016, 2017, and 2018. As done previously, we used these yearly model projections to create maps of inferred trends in species richness across time for each of the target groups, computing the slope of a linear regression for each group. We then compared these modelled trends with the observations recorded in the datasets, evaluating both directionality and magnitude of change. To evaluate if results were better than random predictions, we created 10,000 random predictions of gain and loss across the forest, and, for each one, computed the number of rightly predicted pixels. We used those results to create a null distribution of random predictions and compared the observed number of correct predictions with the null distribution. A script used to run the entire framework (in R and python) is available in the GitHub repository for this manuscript.

RESULTS

Model applicability and sensitivity to range size

The analysis of the first nine groups shows that prediction of biodiversity metrics from environmental variables describing temperature, precipitation and canopy cover during the years of 2002–2014 varied substantially in fit across datasets (species richness: R^2 values 0.18–0.83; phylogenetic diversity: 0.08–0.74; Table 3). In most cases, the fit of the ensemble model was better than that of any individual (tuned) model, with the fit of each individual model being similar across algorithms. In a few exceptions, however, the fit of the individual model was poor—in those cases, the fit of the ensemble model was even worse (Table S1).

Model fit increased with the total area of distribution of the group, when predicting both species richness and phylogenetic diversity (Fig. 1A). However, median individual species range does not appear to be related to model fit (Fig. 1B). Five out of the nine groups had well-fit models for both diversity measures ($R^2 > 0.5$): they included two plant groups (Miconieae and Fridericia & allies) and three animal groups (Thraupinae birds, Ithomiini butterflies and Boana frogs, Fig. 1, Table 3). These three groups are known for having the largest total geographical range in the Atlantic Forest (1,197,890

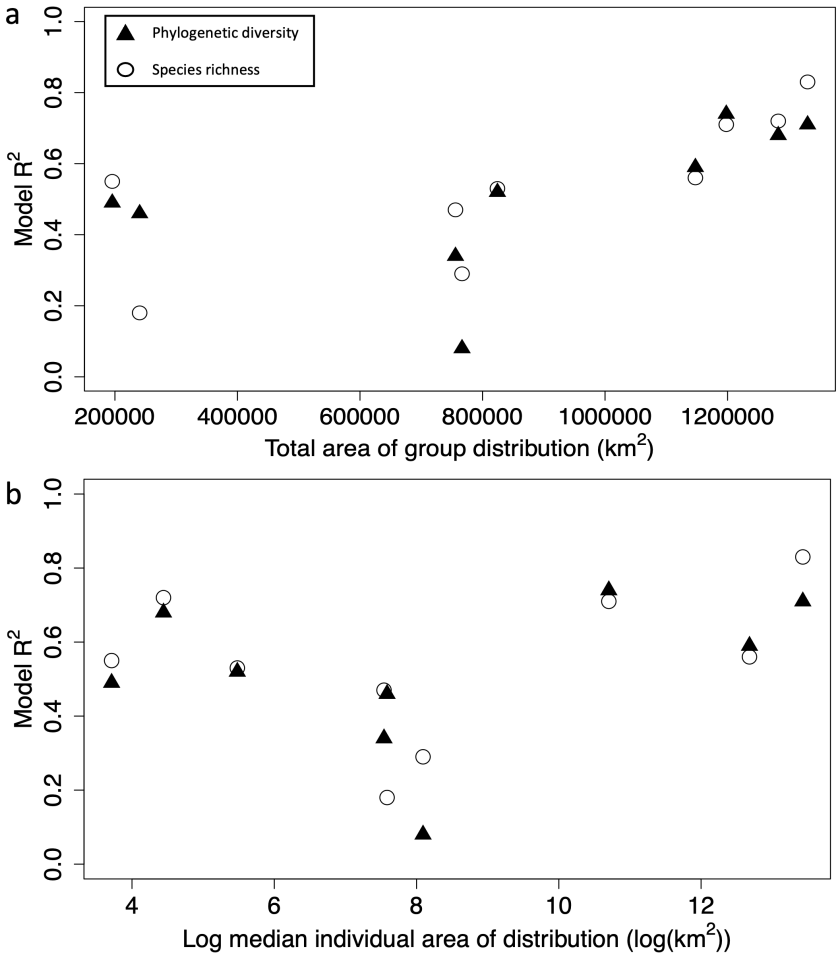

Figure 1 Relationship between model fit and total and individual distributions. (A) Relationship between the total distribution area (in km^2) and predictive model fit (R^2) for species richness (white circles) and phylogenetic diversity (black triangles), for nine target groups. (B) Relationship between the median area of distribution of individual species (log transformed km^2) and predictive model fit (R^2) for species richness (white circles) and phylogenetic diversity (black triangles). Distribution extent was obtained from the species richness maps, R^2 values were obtained from the machine learning predictive model.

Full-size DOI: 10.7717/peerj.13534/fig-1

km^2, 1,282,733 km^2, 1,330,613 km^2, 1,147,465 km^2 and, 824,368 km^2 respectively), but they have variable individual species ranges (median individual species ranges 44,486 km^2, 85 km^2, 680,022 km^2, 321,331 km^2 and 240 km^2 respectively, Table 2). Models of all four remaining groups had an R^2 smaller than 0.49 for at least one of the two diversity measures (Table 3). All four groups had comparatively small total ranges in the forest, but variable individual ranges (Table 2).

In the five groups for which the models were able to predict species richness with good accuracy (Miconieae, the *Fridericia* & allies, Thraupinae birds, Ithomiini butterflies and, *Boana* frogs; species richness R^2 values of 0.71, 0.83, 0.72, 0.56 and, 0.53 respectively; Table 3), the most important predictor variables were all related to precipitation. In Miconieae and the birds, precipitation of the warmest quarter (bio 18) stood up (explaining

25% and 18% of the variation respectively, Table S2). For *Fridericia* and the Ithomiini butterflies, precipitation of the wettest month (bio 13) was most important variable (explaining 18% of the variation). Finally, for the *Boana* frogs, precipitation of the driest month (bio 14) was most important (22%, Table S2). Patterns of phylogenetic diversity in these groups were similarly well correlated with local environments (R^2 values of 0.74, 0.71, 0.68, 0.59 and, 0.52 respectively) and, again, the most important predictor variables were all related to precipitation. Precipitation of the warmest quarter (bio 18) was the main predictor of phylogenetic diversity in the Miconieae and the *Boana* frogs (25% and 22% respectively), whereas precipitation of the wettest month was the main driver for the other three groups (15% for the *Fridericia*, 21% for the Thraupinae and 17% for the Ithomiini, Table S2).

Projections in the five groups with good model fit show similarities and differences in the predicted trends of change in species richness and phylogenetic diversity from 2015 to 2019 (clusters of gain (red) and loss (blue) of predicted diversity across the forest; Fig. 2). In all groups, and for both diversity metrics, a trend toward diversity loss over time is identified in the central region of the forest, in the north of the São Paulo subtropical gap located between the coastal mountains (*Amaral et al., 2018*; *Thom et al., 2020*), and into the inland state of Minas Gerais. This is particularly strong in *Fridericia* & allies (Fig. 2, second row). On the other hand, a trend toward diversity gain (both in species richness and phylogenetic diversity) is detected along the coast, stretching northward from the state of Rio de Janeiro to Espírito Santo (Fig. 2). Models of all groups also detect trends of increasing diversity for both biodiversity metrics in the southern range of the forest (Fig. 2). Maps of the yearly predictions are available in the supplemental material (Figs. S1 and S2).

For all five groups, the predicted distribution of the residuals of the regression of phylogenetic diversity and species richness varies across space, and, to a lesser degree, over time (Fig. 3). In the Miconieae, the models suggest that more phylogenetic diversity is expected to accumulate given species richness (positive residuals expanding across time; Fig. 3, top row green colors). However, two areas of lower phylogenetic diversity are flagged over time: one to the northwest, where the deciduous forest is present (*Peres et al., 2020*), and a second one, along the coast, in the Serra do Mar region (Fig. 3, top row brown colors). In *Fridericia* & allies, the projections mostly suggest that phylogenetic diversity is expected to accumulate relative to species richness except for sites in the northwest (similar to Miconieae) and a large region in interior São Paulo (Fig. 3, second row, brown colors). The latter is the same region where both Miconiaeae and *Fridericia* & allies are inferred to undergo an increase in both diversity metrics over time. The animal groups (Thraupinae birds, Ithomiini butterflies and *Boana* frogs) differ from the two plant groups in the predicted pace of the changes; they appear less drastic over time (Fig. 3). The bird and frog models detect a large region accumulating more phylogenetic diversity than expected in the south of the forest, and suggest this trend to be mostly stable across time except for 2019 (Fig. 3, third and last row, green colors). These two groups also show a consistent prediction of lower phylogenetic diversity than expected in the north of the forest and a somewhat consistent trend in interior Sao Paulo (Fig. 3, third and last row, brown colors). For the butterflies, an opposite pattern emerges—with the south consistently predicted to have less

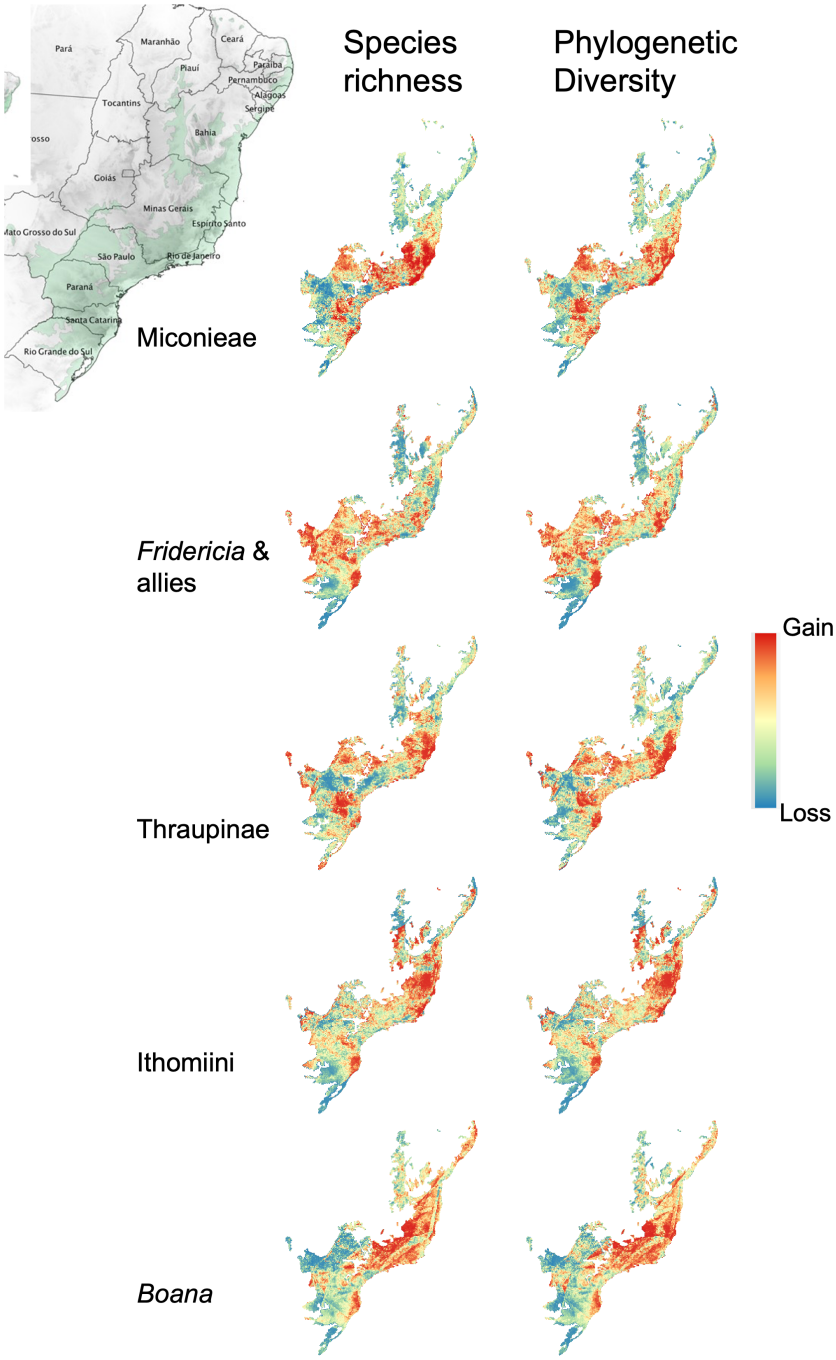

Figure 2 Models of temporal trends in species richness and phylogenetic diversity between the years of 2015 and 2019 for the five groups with the highest model predictive ability. Models of temporal trends in species richness (left) and phylogenetic diversity (right) between the years of 2015 and 2019 for the five groups with the highest model predictive ability (R2 > 0.5), from top to bottom: Miconieae, Fridericia & allies, Thraupinae birds, Ithomiinae butterflies and Boana frogs. Maps depict the slope of the regression between the model built with pre-2015 data and the yearly predictions of species richness and phylogenetic diversity between 2015 and 2019 using the framework presented in the paper. (continued on next page…)

Full-size DOI: 10.7717/peerj.13534/fig-2

Figure 2 (…continued)
Yellow areas depict no change, red areas represent gain of species or phylogenetic diversity, and blue areas represent loss of species or phylogenetic diversity. Darker shades represent larger regression slopes (more gain or loss). For species richness, the slope represents the number of species gained or lost per year; for phylogenetic diversity, the slope represents phylogenetic diversity gained or lost per year. Color scale for all maps is a quantile classification with 20 classes. For species richness, the slopes vary between −9.08 and 9.9 in the Miconieae, between −1.85 and 1.82 in the Fridericia & allies, between −1.09 to 1.55 in the Thraupinae, between −2.64 and 3.04 in the Ithomiinae and between −0.49 and 0.50 in the Boana frogs. For phylogenetic diversity, the slope varies between −33.41 and 43.71 in the Miconieae, between −0.016 and 0.010 in the Fridericia & allies, between -15.50 and 21.70 in the Thraupinae birds, between −0.12 and 0.13 in the Ithomiini butterflies, and between −0.085 and 0.11 in the Boana frogs.

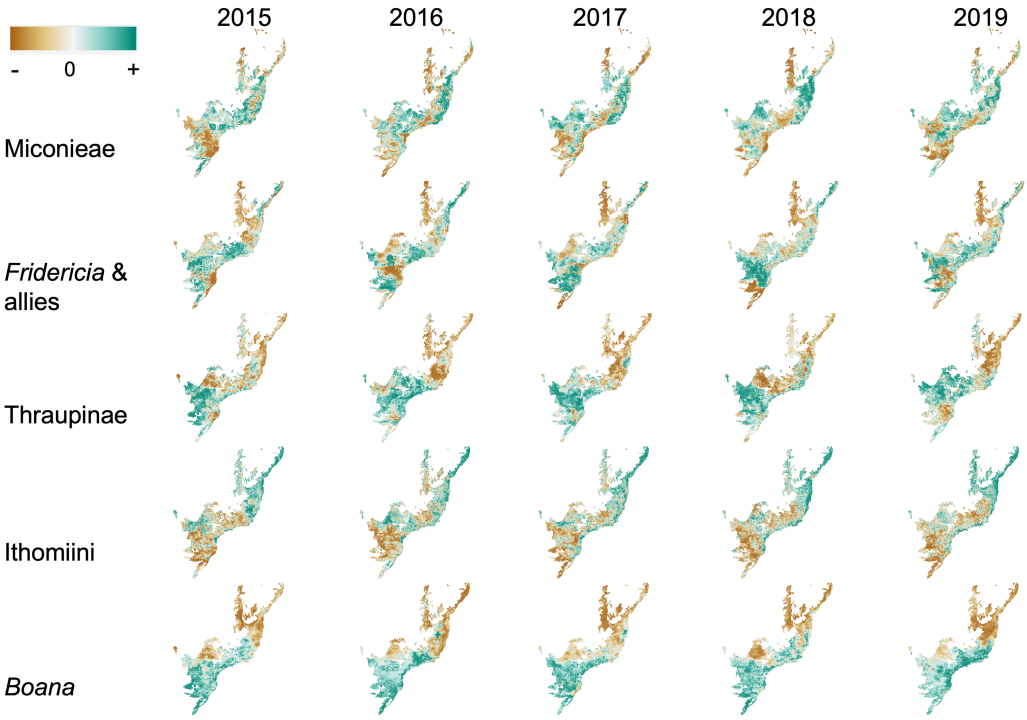

Figure 3 Modelled change in residuals of the regression of phylogenetic diversity on species richness across time between 2015 and 2019. From top to bottom: for the Miconieae, the *Fridericia* & allies, the Thraupinae birds, the Ithomiini butterflies, and the *Boana* frogs. Green colors represent positive residuals; brown colors negative residuals. Color intensity reflects the magnitude of the residuals (darker colors depicting larger residuals).

Full-size DOI: 10.7717/peerj.13534/fig-3

butterfly phylogenetic diversity than expected, and the north having more phylogenetic diversity than expected (Fig. 3, fourth row, brown and green colors respectively).

Lessons from model testing

When the model framework is applied to inventory data that include multiple records per year (enabled by the ATLANTIC-amphibian and ATLANTIC-epiphytes datasets), this framework predicts the directionality of biodiversity change well. Although the predictive power of the amphibian and epiphyte models is lower than those observed for the groups analyzed in the first part of the study (R^2 of 0.39 for epiphytes, 0.21 for amphibians), the

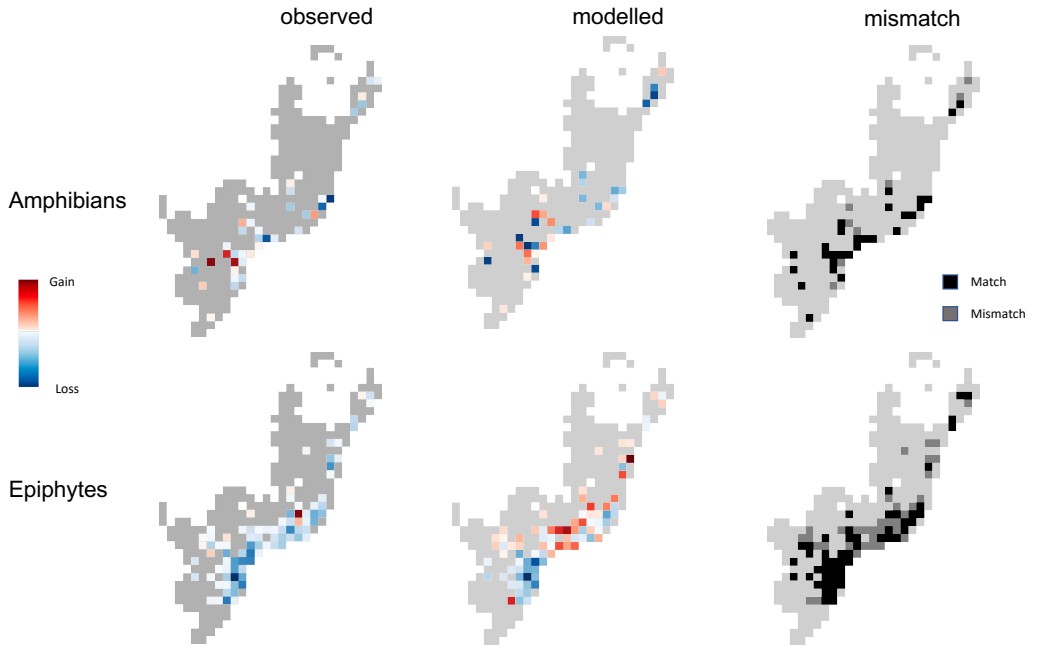

Figure 4 A comparison of the trends of change in species richness for two ATLANTIC datasets (top: amphibians, bottom: epiphytes). Maps show the slope of the regression between a training dataset (pooled records pre-2012 for amphibians, and pre-2015 for epiphytes) and the observed (left) and modeled (right) species richness in four subsequent years after that. Blue color depicts areas experiencing loss of species over time, whereas red denotes areas experiencing gain of species over time; gray denotes areas with no data. Darker shades denote larger slopes (stronger gain or loss). The last column shows the mismatch between the observed and predicted direction on of change, with black pixels being correctly predicted and dark grey pixels incorrectly predicted.

Full-size DOI: 10.7717/peerj.13534/fig-4

modelled trends of potential change in species richness are broadly congruent with those measured with empirical data, particularly regarding species loss (Fig. 4). A comparison of the modelled *vs.* observed trends in gain or loss of species richness indicates that the model framework correctly predicted the direction of change in 59% of the pixels (52/88 pixels, p-value = 0.08) in the epiphyte study, and 76% of the pixels (28/37 pixels p-value = 0.0008) in the amphibian group (Fig. 4). Even though richness was described using larger pixels in the ATLANTIC datasets relative to the nine other groups (100 km instead of 10 km), not all pixels contained enough information to compute observed trends; as such, fewer pixels are available for comparisons (88 for the epiphytes, 37 for the amphibians).

These modelled trends of change differ between the two groups and are spatially heterogeneous, with a tendency towards loss of species. A total of 46/88 pixels in the epiphytes model and 20/37 pixels in the amphibian model have a negative slope. The same tendency towards species loss is present in the observed data, although more strongly, especially in the epiphytes (79/88 pixels in the observed epiphyte data, and 26/37 pixels in the amphibian data; Fig. 4). While the predictive models of the epiphytes generally matched the observed trends of species loss, the magnitude of loss was generally underestimated; observed slopes ranged from -104 to 19, while modelled slopes varied from −75 to 28

(Fig. 4, bottom). In the amphibians, the patterns of biodiversity inferred by the model matched the direction of change reasonably well, with a similar estimated strength of change (observed slopes ranging from −5.62 to 5; modelled slopes spanning from −6.7 to 5.5, Fig. 4 top).

DISCUSSION

We propose a novel framework for indirect biodiversity monitoring based on correlations with remotely sensed environmental data, and find that it performs well in some, but not all groups of organisms for which data are available in the Brazilian Atlantic Forest. When applied to groups that are widely distributed—and for which abundant data are available for model training—the model performs very well, statistically. This is the case of the Miconieae, *Fridericia* & allies, the Thraupinae birds and, the Ithomiini butterflies. However, the ensemble model's predictive power decreases when applied to other groups, a result also observed when individual models are built (Table S1). The groups for which the models underperform are more spatially restricted, both in terms of overall group range, or, sometimes, including smaller individual species ranges (Table 3). This result is not surprising, as the performance of machine learning methods is expected to improve with the addition of more data (*Kuhn & Johnson, 2016*). Furthermore, the geographical partition method used for model validation also yields improved performance with better geographical coverage. It is possible that the use of environmental data at a resolution of 10 km may be inappropriate for those groups known to have more restricted distributions, and that those models may improve with environmental data at finer resolutions. Attempts to clarify this issue should nonetheless acknowledge that estimating species distributions at finer resolutions often results in an overestimation of species distributions (*Graham & Hijmans, 2006*; *Hurlbert & Jetz, 2007*).

Model outcomes of some groups, including the highly performing models for the *Boana* frogs, suggest that elements other than total group range are important for model performance. Models of the *Boana* tree frogs, which have small species ranges and a small total range in the forest, performed well (R^2 of 0.53 for species richness, 0.52 for phylogenetic diversity). Models for the two groups with the smallest total area, the *Bertolonia* and the Bromelioidae (the last one known for having species with the smallest ranges), had comparatively good performance for at least one of the two models (R^2 of 0.18 and 0.55 respectively for species richness, 0.46 and 0.49 for phylogenetic diversity). Although the small sample size precludes statistical evaluations ($N = 3$ groups), it is possible that these results reflect the impact of the ecology of these species, particularly their dependence on the environment, on model performance. Because the ecology and life history traits of treefrogs and bromeliads render them strongly dependent on the abiotic environment, we hypothesize that the models of these groups fit better than expected given their geographical ranges and those of their species. As such, we expect that this model framework will work best for groups of species that are tightly associated with specific environmental conditions (rather than landscape and historical factors), or broadly distributed in a biome of interest.

The generally lower predictive performance of the models, when applied to smaller-ranged groups (R^2 ranging from 0.08–0.55, Table 3), is consistent with the outcomes of

the models of the ATLANTIC-amphibians and ATLANTIC-epiphyte datasets. Because the latter had fewer pixels (observations), and hence a sparser geographical coverage, lower predictive performance was expected—and detected. Yearly coverage did not appear to dictate whether the direction of change predicted at a site was correctly or incorrectly inferred. Yet, despite these limitations, a comparison between model outputs and the empirical data demonstrates that the proposed framework can perform reasonably well when predicting the directionality of change in species richness—particularly when predicting species loss. This result is not surprising given that this study is evaluating near-real time changes in diversity (*i.e.*, over the span of a few years), when species gain might be more difficult to document *via* inventory work, and when species gain may lag behind environmental change (*Zanatta et al., 2020*). On the one hand, even if areas are gaining species, it might take time for populations to become established in enough numbers (or grow, in the case of plants) to be detected by *in situ* monitoring (*Alexander et al., 2018*). On the other hand, loss can happen faster in the face of extreme environmental conditions either from local population extinctions, or though movement of the species (*Davis & Shaw, 2001*; but extinction can also lag in long-lived species or under continued recruitment from seed banks or nearby populations; *Alexander et al., 2018*). In either case, we argue that the early detection of areas threatened with diversity loss is key to the survival of many species and a useful complement to the early detection of individual species at risk (*Stanton et al., 2015*).

Although some trends in diversity change are group specific, the models detect certain geographical regions of the forest—*e.g.*, coastal Espírito Santo—that are consistently predicted to gain diversity over time, irrespectively of biodiversity metric. This area is nonetheless under intense anthropogenic disturbance (*Tabarelli et al., 2010*), and habitat loss might prevent the species accumulation predicted by climate only. In the sense that this model framework does not incorporate land use and land use change, it is important to keep in mind that it so far only provides insight based on macroclimatic shifts. In contrast, the northwestern portion of the forest is consistently predicted to be losing both species and phylogenetic diversity in all groups but the butterflies (loss is only predicted in the northernmost area), and losing more phylogenetic diversity relative to species richness—thus potentially through the loss of species that are distantly related. While these conclusions are suggested by models of the community, they can be tested with species-specific models and targeted inventories.

Likewise, the models may be used to identify potential ecological processes at play—and test them. While patterns of predicted diversity change seem similar between species richness and phylogenetic diversity for all groups, the predicted residuals for each year show that their mismatch varies in time. Positive residuals are more widespread, showing a general pattern of more phylogenetic diversity than expected given the number of species. The latter can mean that the species to be gained locally are distantly related, or that species that are closely related are being lost. The gain of distant species or loss of redundant ones can be driven by species interactions driving similar species to compete (*Webb et al., 2002*). As such, studies that quantify and monitor Atlantic Forest species interactions and networks may be particularly relevant under current and future

environmental shifts (*Sales, Rodrigues & Masiero, 2021*). On the other hand, a few areas show consistent trends toward phylogenetic diversity loss relative to species richness (inland São Paulo for Miconieae plants, *Boana* frogs and Thraupinae birds, a cluster of sites in the northwestern forests for all but the butterflies, and the southern Atlantic Forest for the Miconiae plants and the butterflies). In those areas, the model is predicting the loss of distantly related species, or the gain of closely related ones. This might be pointing at areas where environmental conditions are becoming more extreme and so species persisting, or colonizing, will be more like each other than expected by chance. While the present study is unable to validate these hypotheses, they are testable through field observations and experimental work.

CONCLUSIONS AND FINAL REMARKS

The modeling framework proposed here allows for the flagging of areas that are at risk of losing diversity in the face of climatic changes. The incorporation of two dimensions of diversity—taxonomic and phylogenetic—enables the identification of areas that might be at risk of losing evolutionary diversity due to environmental change. However, the predictions of the model should be taken carefully when predicting diversity gain, as those seem to be overly optimistic. Because the strength of the predicted change is overestimated by the model, we propose that practitioners be careful when interpreting magnitude; it may be best to focus only on the expected direction of change. This would allow, for example, for a better design of new protected areas and for targeted monitoring of areas either predicted to lose diversity (in general) or to disproportionately lose evolutionary diversity.

Key to this indirect remote sensing effort is the periodicity of the remote sensing information to be correlated with—and used as proxy for—biological diversity. Remote sensing data are available at different periodicities, ranging from days to years, depending on the agencies that collect them, their instruments, and costs. Accessing free data has become easier with the advent of Google Earth Engine (*Gorelick et al., 2017*), a cloud-based service that not only stores multiple datasets but also allows for free data processing through open code. The framework is easily implemented and can be rerun with every pass of a satellite, incorporating giving us new environmental data and thus adjusting predictions to the real change in near real time. We propose that this use may be further explored as a tool for conservation planning and policy design, using the most up-to-date information.

This framework is also easily adaptable to other groups or areas of interest, by replacing the calibration biodiversity data. It may be applied to other diversity measures and dimensions of biodiversity, although users must carefully assess their degree of correlation with present-day environmental variables. Previous work has shown that metrics that largely reflect historical components or are associated with spatial restriction (*e.g.*, phylogenetic endemism) are not easily predicted by climate descriptors (*Paz et al., 2021*).

Importantly, our preliminary analyses demonstrate that this framework does not require extensive inventory data for the whole forest, but may be applied to species with curated occurrence data, allowing for indirect spatial monitoring of biodiversity in near real-time. Based on the findings of the present analysis, we propose that it be used only for groups of

species that are wide ranging and well sampled in a biome of interest. Because correlations between diversity patterns and environmental correlates are dependent on the area of model training, the model is not to be applied outside the region where it was trained (in this case, the Atlantic Forest biome). Diversity patterns can vary in different geographical areas, making model transfer problematic, especially in regions with area limitations, such as islands.

ACKNOWLEDGEMENTS

We would like to thank the members of the AF-Biota and NSF Dimensions of Biodiversity projects for discussions that inspired this work. We thank Rob Anderson, Brian Smith, and members of the Carnaval Lab at CUNY for comments on previous versions of this manuscript. We also thank the editor, Shan Kothari, Markus Gastauer and one anonymous reviewer for comments that greatly improved this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Andrea Paz was funded by a Fulbright-Colciencias fellowship and a CCNY Cluster for the study of biodiversity and environmental change seed award. This study used data collected by a collaborative project co-funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; 2013/50297-0), the National Science Foundation (NSF; DEB 1343578, DEB-1343612), and NASA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Fulbright-Colciencias fellowship.
CCNY Cluster for the study of biodiversity and environmental change seed award.
Fundação de Amparo à Pesquisa do Estado de São Paulo: 2013/50297-0.
National Science Foundation: DEB 1343578, DEB-1343612.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Andrea Paz conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Thiago S. Silva conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ana C. Carnaval conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The code is available at GitHub: https://github.com/andrepazv/NearRealTime.

The data used from *Paz et al. (2021)* is available at Dryad: Paz, Andrea (2021), Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest, Dryad, Dataset, https://doi.org/10.5061/dryad.6m905qfzp.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.13534#supplemental-information.

REFERENCES

Aguirre-Santoro J. 2017. Taxonomy of the Ronnbergia Alliance (Bromeliaceae: Bromelioideae): new combinations, synopsis, and new circumscriptions of Ronnbergia and the resurrected genus Wittmackia. *Plant Systematics and Evolution* **303**:615–640 DOI 10.1007/s00606-017-1394-y.

Aguirre-Santoro J, Stevenson D, Michelangeli F. 2016. Molecular phylogenetics of the Ronnbergia Alliance (Bromeliaceae, Bromelioideae) and insights about its morphological evolution. *Molecular Phylogenetics and Evolution* **100**:1–20 DOI 10.1016/j.ympev.2016.04.007.

Alexander JM, Chalmandrier L, Lenoir J, Burgess TI, Essl F, Haider S, Kueffer C, McDougall K, Milbau A, Nuñez MA, Pauchard A, Rabitsch W, Rew LJ, Sanders NJ, Pellissier L. 2018. Lags in the response of mountain plant communities to climate change. *Global Change Biology* **24**:563–579 DOI 10.1111/gcb.13976.

Amaral FRD, Alvarado-Serrano DF, Maldonado-Coelho M, Pellegrino KCM, Miyaki CY, Montesanti J, Lima-Ribeiro MS, Hickerson MJ, Thom G. 2018. Climate explains recent population divergence, introgression and persistence in tropical mountains: phylogenomic evidence from Atlantic Forest warbling finches. *bioRxiv* DOI 10.1101/439265.

Bacci LF, Amorim AM, Michelangeli FA, Goldenberg R. 2020. Flower morphology is correlated with distribution and phylogeny in Bertolonia (Melastomataceae), an herbaceous genus endemic to the Atlantic Forest. *Molecular Phylogenetics and Evolution* **149**:106844 DOI 10.1016/j.ympev.2020.106844.

Bochorny T, Michelangeli FA, Almeda F, Goldenberg R. 2019. Phylogenetics, morphology and circumscription of Cambessedesieae: a new Neotropical tribe of Melastomataceae. *Botanical Journal of the Linnean Society* **190**:281–302 DOI 10.1093/botlinnean/boz018.

Brown J, Paz A, Reginato M, Amaro C, Assis R, Lyra M, Caddah M, Aguirre-Santoro J, D'Horta F, Amaral FRD, Goldenberg R, Silva-Brandão KL, Freitas A, Rodrigues M, Michelangeli FA, Miyaki C, Carnaval A. 2020. Seeing the forest through many trees: multi-taxon patterns of phylogenetic diversity in the Atlantic Forest hotspot. *Diversity and Distributions* **26**:1160–1176 DOI 10.1111/ddi.13116.

Burns KJ, Shultz AJ, Title PO, Mason NA, Barker FK, Klicka J, Lanyon SM, Lovette IJ. 2014. Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of Neotropical songbirds. *Molecular Phylogenetics and Evolution* **75**:41–77 DOI 10.1016/j.ympev.2014.02.006.

Caddah M. 2013. Estudos taxonomicos e filogenéticos em Miconia sect. Discolor (Meslastomataceae, Miconieae). PhD Thesis, UNICAMP, Campinas.

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW. 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters* **12**:693–715 DOI 10.1111/j.1461-0248.2009.01314.x.

Davis MB, Shaw RG. 2001. Range shifts and adaptive responses to quaternary climate change. *Science* **292(5517)**:673–679 DOI 10.1126/science.292.5517.673.

Deane-Mayer ZA, Knowles JE. 2019. caretEnsemble: ensembles of caret models. R package version 6.1.1. *Available at https://cran.r-project.org/package=caretEnsemble*.

Diniz CG, Souza AADA, Santos DC, Dias MC, Da Luz NC, De Moraes DRV, Maia JSA, Gomes AR, Narvaes IDS, Valeriano DM, Maurano LEP, Adami M. 2015. DETER-B: the new amazon near real-time deforestation detection system. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **8**:3619–3628 DOI 10.1109/JSTARS.2015.2437075.

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**:1–10 DOI 10.1016/0006-3207(92)91201-3.

Fretwell PT, Trathan PN. 2009. Penguins from space: faecal stains reveal the location of emperor penguin colonies. *Global Ecology and Biogeography* **18**:543–552 DOI 10.1111/j.1466-8238.2009.00467.x.

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J. 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* **2**:150066 DOI 10.1038/sdata.2015.66.

Garzon-Lopez CX, Bohlman SA, Olff H, Jansen PA. 2013. Mapping tropical forest trees using high-resolution aerial digital photographs. *Biotropica* **45**:308–316 DOI 10.1111/btp.12009.

Goldenberg R, Penneys DS, Almeda F, Judd WS, Michelangeli F. 2008. Phylogeny of Miconia (Melastomataceae): Initial insights into broad patterns of diversification in a megadiverse neotropical genus. *International Journal of Plant Sciences* **169**:963–979 DOI 10.1086/589697.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. 2017. Remote sensing of environment google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**:18–27 DOI 10.1016/j.rse.2017.06.031.

Graham CH, Hijmans RJ. 2006. A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography* **15**:578–587 DOI 10.1111/j.1466-822x.2006.00257.x.

Hijmans R, Phillips S, Leathwick J, Elith J. 2011. Package 'dismo'. *Available at http://cran.r-project.org/web/packages/dismo/index.html*.

Hurlbert AH, Jetz W. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences of the United States of America* **104**:13384–13389 DOI 10.1073/pnas.0704469104.

Kaehler M, Michelangeli F, Lohmann LG. 2019. Fine tuning the circumscription of Fridericia (Bignonieae, Bignoniaceae). *Taxon* **68**:751–770 DOI 10.1002/tax.12121.

Karatzoglou A, Smola A, Hornik K, Zeileis A. 2004. kernlab—An S4 package for kernel methods in R. *Journal of Statistical Software* **11**:1–20.

Kuhn M, Johnson K. 2016. *Applied predictive modeling*. New York: Springer DOI 10.1007/978-1-4614-6849-3.

Laffan SW, Lubarsky E, Rosauer DF. 2010. Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography* **33**:643–647 DOI 10.1111/j.1600-0587.2010.06237.x.

Laurencio D, Fitzgerald LA. 2010. Environmental correlates of herpetofaunal diversity in Costa Rica. *Journal of Tropical Ecology* **26**:521–531 DOI 10.1017/S0266467410000222.

Liaw A, Wiener M. 2002. Classification and regression by random forest. *R News* **2**:18–22.

Mayfield MM, Levine JM. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* **13**:1085–1093 DOI 10.1111/j.1461-0248.2010.01509.x.

Meyer L, Diniz-Filho JAF, Lohmann LG. 2017. A comparison of hull methods for estimating species ranges and richness maps. *Plant Ecology and Diversity* **10**(**5–6**):389–401 DOI 10.1080/17550874.2018.1425505.

Michelangeli F, Judd WS, Penneys DS, Skean Jr J, Becquer E, Goldenberg R, Martin C. 2008. Multiple events of dispersal and radiation of the tribe Miconieae (Melastomataceae) in the Caribbean. *Botanical Review* **74**:53–77 DOI 10.1007/s12229-008-9004-x.

Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP. 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution* **5**:1198–1205 DOI 10.1111/2041-210X.12261.

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG. 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**:191–203 DOI 10.1111/j.1600-0587.2013.00205.x.

Navarro LM, Fernández N, Guerra C, Guralnick R, Kissling WD, Londoño MC, Muller-Karger F, Turak E, Balvanera P, Costello MJ, Delavaud A, El Serafy GY, Ferrier S, Geijzendorffer I, Geller GN, Jetz W, Kim ES, Kim HJ, Martin CS, McGeoch MA, Mwampamba TH, Nel JL, Nicholson E, Pettorelli N, Schaepman ME, Skidmore A, Sousa Pinto I, Vergara S, Vihervaara P, Xu H, Yahara T, Gill M, Pereira HM. 2017. Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability* **29**:158–169 DOI 10.1016/j.cosust.2018.02.005.

Papeş M, Tupayachi R, Martínez P, Peterson AT, Powell GVN. 2010. Using hyperspectral satellite imagery for regional inventories: a test with tropical emergent trees in the Amazon Basin. *Journal of Vegetation Science* **21**:342–354 DOI 10.1111/j.1654-1103.2009.01147.x.

Paz A, Brown JL, Cordeiro CLO, Aguirre-Santoro J, Assis C, Amaro RC, Raposodo Amaral F, Bochorny T, Bacci LF, Caddah MK, D'Horta F, Kaehler M, Lyra M, Grohmann CH, Reginato M, Silva-Brandão KL, Freitas AVL, Goldenberg R, Lohmann LG, Michelangeli FA, Miyaki C, Rodrigues MT, Silva TS, Carnaval AC. 2021. Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest. *Journal of Biogeography* **48**:1377–1391 DOI 10.1111/jbi.14083.

Paz A, Reginato M, Michelangeli FA, Goldenberg R, Caddah M, Aguirre-Santoro J, Kaehler M, Lohmann L, Carnaval AC. 2020. Predicting patterns of plant diversity and endemism in the tropics using remote sensing data: a study case from the Brazilian Atlantic rainforest. In: Cavender-Bares J, Gamon J, Townsend P, eds. *Remote sensing of plant biodiversity*. Cham: Springer, 255–266 DOI 10.1007/978-3-030-33157-3_11.

Peres EA, Pinto-da Rocha R, Lohmann LG, Michelangeli FA, Miyaki CY, Carnaval AC. 2020. Patterns of species and lineage diversity in the Atlantic rainforest of Brazil. In: Rull V, Carnaval AC, eds. *Neotropical diversification: patterns and processes*. Cham: Springer, 415–447 DOI 10.1007/978-3-030-31167-4_16.

Peters MK, Hemp A, Appelhans T, Behler C, Classen A, Detsch F, Ensslin A, Ferger SW, Frederiksen SB, Gebert F, Haas M, Helbig-Bonitz M, Hemp C, Kindeketa WJ, Mwangomo E, Ngereza C, Otte I, Röder J, Rutten G, Schellenberger Costa D, Tardanico J, Zancolli G, Deckert J, Eardley CD, Peters RS, Rödel M-O, Schleuning M, Ssymank A, Kakengi V, Zhang J, Böhning-Gaese K, Brandl R, Kalko EKV, Kleyer M, Nauss T, Tschapka M, Fischer M, Steffan-Dewenter I, Gaston KJ, Brown JH, Fine PVA, Rhode K, Mittelbach GG, Allen AP, Brown JH, Gillooly JF, Hawkins BA, Kreft H, Jetz W, Hutchinson GE, Brown JH, Hurlbert AH, Stegen JC, Belmaker J, Jetz W, Romdal TS, Grytnes J-A, Colwell RK, Rahbek C, Gotelli NJ, Brehm G, Colwell RK, Kluge J, Novotny V, Basset Y, Wiens JJ, Crisp MD, Kühsel S, Blüthgen N, Classen A, Chao A, Colwell RK, Lin C-W, Gotelli NJ, Chao A, Chao A, Burnham KP, Anderson DR, Burnham KP, Anderson DR, Huyvaert KP, Grace JB, Colwell RK, Brehm G, Cardelus CL, Gilman AC, Longino JT, Duan M, Kessler M, Rahbek C, Rahbek C, Wiens JJ, Graham CH, Jablonski D, Merckx VSFT, Wu Y, Kozak KH, Wiens JJ, Tittensor DP, Baur B, Meier T, Baur A, Schmera D, McCain CM, Colwell RK, Forister ML, Zhou J, Allen AP, Gillooly JF, Savage VM, Brown JH, Rolland J, Condamine FL, Jiguet F, Morlon H, Bagchi R, Johnson DJ, Beaulieu WT, Bever JD, Clay K, Hemp A, Hemp A, Trautwein MD, Wiegmann BM, Beutel R, Kjer KM, Yeates DK, Misof B, Grafen A, Zancolli G, Steffan-Dewenter I, Rödel M-O, Hemp C, Hemp C, Wang D, Atzberger C, Eilers PHC, Colwell RK, Hurtt GC, Johnson JB, Omland KS, Hsieh TC, Ma KH, Chao A, Rosseel Y. 2016. Predictors of elevational biodiversity gradients change from single taxa to the multi-taxa community level. *Nature Communications* **7**:13736 DOI 10.1038/ncomms13736.

R Core Team. 2022. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. *Available at https://www.R-project.org/*.

Radosavljevic A, Anderson RP. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography* **41**:629–643 DOI 10.1111/jbi.12227.

Ramos FN, Mortara SR, Monalisa-Francisco N, Elias JPC, Neto LM, Freitas L, Kersten R, Amorim AM, Matos FB, Nunes-Freitas AF, Alcantara S, Alexandre MHN, De Almeida-Scabbia RJ, De Almeida OJG, Alves FE, De Oliveira Alves RM, Alvim FS, De Andrade ACS, De Andrade S, Aona LYS, Araujo AC, De Araújo KCT, Ariati V, Assis JC, De Azevedo CO, Barbosa BF, Barbosa DEF, Barbosa FDR, De Barros F, Basilio GA, Bataghin FA, Bered F, Bianchi JS, Blum CT, Boelter CR, Bonnet A, Brancalion PHS, Breier TB, Brion CDT, Buzatto CR, Cabral A, Cadorin TJ, Caglioni E, Canêz L, Cardoso PH, De Carvalho FS, Carvalho RG, Catharino ELM, Ceballos SJ, Cerezini MT, César RG, Cestari C, Chaves CJN, Citadini-Zanette V, Coelho LFM, Coffani-Nunes JV, Colares R, Colletta GD, Corrêa NDM, Da Costa AF, Da Costa GM, Costa LMS, Costa NGS, Couto DR, Cristofolini C, Da Cruz ACR, Del Neri LA, Di Pasquo M, Dos Santos Dias A, Dias LDCD, Dislich R, Duarte MC, Fabricante JR, Farache FHA, De Faria APG, Faxina C, Ferreira MTM, Fischer E, Fonseca CR, Fontoura T, Francisco TM, Furtado SG, Galetti M, Garbin ML, De Gasper AL, Goetze M, Gomes-da Silva J, Gonçalves MFA, Gonzaga DR, Silva ACGE, Guaraldo ADC, Guarino EDSG, Guislon AV, Hudson LB, Jardim JG, Jungbluth P, Kaeser SDS, Kessous IM, Koch NM, Kuniyoshi YS, Labiak PH, Lapate ME, Santos ACL, Leal RLB, Leite FS, Leitman P, Liboni AP, Liebsch D, Lingner DV, Lombardi JA, Lucas E, Luzzi JDR, Mai P, Mania LF, Mantovani W, Maragni AG, Marques MCM, Marquez G, Martins C, Martins LDN, Martins PLSS, Mazziero FFF, Melo CDA, De Melo MMF, Mendes AF, Mesacasa L, Morellato LPC, Moreno VDS, Muller A, Murakami MMDS, Cecconello E, Nardy C, Nervo MH, Neves B, Nogueira MGC, Nonato FR, De Oliveira-Filho AT, De Oliveira CPL, Overbeck GE, Marcusso GM, Paciencia MLB, Padilha P, Padilha PT, Pereira ACA, Pereira LC, Pereira RAS, Pincheira-Ulbrich J, Pires JSR, Pizo MA, Pôrto KC, Rattis L, Reis JRDM, Reis SGD, Da Rocha-Pessôa TC, Rocha CFD, Rocha FS, Rodrigues ARP, Rodrigues RR, Rogalski JM, Rosanelli RL, Rossado A, Rossatto DR, Rother DC, Ruiz-Miranda CR, Saiter FZ, Sampaio MB, Santana LD, Santos JSD, Sartorello R, Sazima M, Schmitt JL, Schneider G, Schroeder BG, Sevegnani L, Júnior VOS, Da Silva FR, Da Silva MJ, Silva MPP, Silva RG, Silva SM, Singer RB, Siqueira G, Soares LE, De Sousa HC, Spielmann A, Tonetti VR, Toniato MTZ, Ulguim PSB, Van den Berg C, Van den Berg E, Varassin IG, Da Silva IBV, Vibrans AC, Waechter JL, Weissenberg EW, Windisch PG, Wolowski M, Yañez A, Yoshikawa VN. 2019. ATLANTIC EPIPHYTES: a data set of vascular and non-vascular epiphyte plants and lichens from the Atlantic Forest. *Ecology* **100**:1–2 DOI 10.1002/ecy.2541.

Reddy CS. 2021. Remote sensing of biodiversity: what to measure and monitor from space to species? *Biodiversity and Conservation* **30**:2617–2631 DOI 10.1007/s10531-021-02216-5.

Reginato M, Michelangeli F. 2016. Untangling the phylogeny of Leandra sensu str. (Melastomataceae, Miconieae). *Molecular Phylogenetics and Evolution* **96**:17–32 DOI 10.1016/j.ympev.2015.11.015.

Rompré G, Robinson WD, Desrochers A, Angehr G. 2007. Environmental correlates of avian diversity in lowland Panama rain forests. *Journal of Biogeography* **34**:802–815 DOI 10.1111/j.1365-2699.2006.01657.x.

Sales LP, Rodrigues L, Masiero R. 2021. Climate change drives spatial mismatch and threatens the biotic interactions of the Brazil nut. *Global Ecology and Biogeography* **30**:117–127 DOI 10.1111/geb.13200.

Schmaljohann H. 2020. Radar aeroecology—a missing piece of the puzzle for studying the migration ecology of animals. *Ecography* **43**:236–238 DOI 10.1111/ecog.04807.

Scholes RJ, Mace GM, Turner W, Geller GN, Jürgens N, Larigauderie A, Muchoney D, Walther BA, Mooney HA. 2008. Toward a global biodiversity observing system. *Science* **321**:1044–1045 DOI 10.1126/science.1162055.

Scholes RJ, Walters M, Turak E, Saarenmaa H, Faith DP, Mooney HA, Heip CHR, Ferrier S, Jongman RHG, Harrison IJ, Yahara T, Pereira HM, Larigauderie A, Geller G. 2012. Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability* **4**:1–8 DOI 10.1016/j.cosust.2011.12.005.

Skidmore AK, Coops NC, Neinavaz E, Ali A, Schaepman ME, Paganini M, Kissling WD, Vihervaara P, Darvishzadeh R, Feilhauer H, Fernandez M, Fernández N, Gorelick N, Geijzendorffer I, Heiden U, Heurich M, Hobern D, Holzwarth S, Muller-Karger FE, Van De Kerchove R, Lausch A, Leitão PJ, Lock MC, Mücher CA, O'Connor B, Rocchini D, Turner W, Vis JK, Wang T, Wegmann M, Wingate V. 2021. Priority list of biodiversity metrics to observe from space. *Nature Ecology and Evolution* **5**:896–906 DOI 10.1038/s41559-021-01451-x.

Stanton JC, Shoemaker KT, Pearson RG, Akçakaya HR. 2015. Warning times for species extinctions due to climate change. *Global Change Biology* **21**:1066–1077 DOI 10.1111/gcb.12721.

Tabarelli M, Aguiar AV, Ribeiro MC, Metzger JP, Peres CA. 2010. Prospects for biodiversity conservation in the Atlantic Forest: lessons from aging human-modified landscapes. *Biological Conservation* **143**:2328–2340 DOI 10.1016/j.biocon.2010.02.005.

Thom G, Smith BT, Gehara M, Montesanti J, Lima-Ribeiro MS, Piacentini VQ, Miyaki CY, Do Amaral FR. 2020. Climatic dynamics and topography control genetic variation in Atlantic Forest montane birds. *Molecular Phylogenetics and Evolution* **148**:106812 DOI 10.1016/j.ympev.2020.106812.

Turner W. 2014. Sensing biodiversity. *Science* **346**:301–302 DOI 10.1126/science.1256014.

Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M. 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution* **18**:306–314 DOI 10.1016/S0169-5347(03)00070-3.

Van Der Ploeg T, Austin PC, Steyerberg EW. 2014. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Medical Research Methodology* **14**:1–13 DOI 10.1186/1471-2288-14-137.

Vancine MH, Duarte KDS, De Souza YS, Giovanelli JGR, Martins-Sobrinho PM, López A, Bovo RP, Maffei F, Lion MB, Ribeiro Júnior JW, Brassaloti R, Da Costa COR, Sawakuchi HO, Forti LR, Cacciali P, Bertoluci J, Haddad CFB, Ribeiro MC. 2018. ATLANTIC AMPHIBIANS: a data set of amphibian communities from the Atlantic Forests of South America. *Ecology* **99**:1692 DOI 10.1002/ecy.2392.

Vasconcelos TS, Prado VHM, Da Silva FR, Haddad CFB. 2014. Biogeographic distribution patterns and their correlates in the diverse frog fauna of the atlantic forest hotspot. *PLOS ONE* **9(8)**:e104130 DOI 10.1371/journal.pone.0104130.

Venables WN, Ripley BD. 2002. *Modern applied statistics with S*. New York: Springer.

Vermote E, Wolfe R. 2015. MYD09GA MODIS/Aqua surface reflectance daily L2G Global 1 km and 500 m SIN Grid V006. NASA EOSDIS Land Processes DAAC. NASA LP DAAC. DOI 10.5067/MODIS/MYD09GA.006.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* **33**:475–505 DOI 10.1146/annurev.ecolsys.33.010802.150448.

World Resources Institute. 2014. Global forest watch. *Available at https://www.globalforestwatch.org/* (accessed on 23 June 2020).

Zanatta F, Engler R, Collart F, Broennimann O, Mateo RG, Papp B, Muñoz J, Baurain D, Guisan A, Vanderpoorten A. 2020. Bryophytes are predicted to lag behind future climate change despite their high dispersal capacities. *Nature Communications* **11**:5601 DOI 10.1038/s41467-020-19410-8.

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**:1080–1090 DOI 10.1111/jbi.12696.