DOI: 10.1111/2041-210X.12950

RESEARCH ARTICLE

Methods in Ecology and Evolution

Functional ecology and imperfect detection of species

Tobias Roth^{1,2} I Eric Allan³ Peter B. Pearman^{4,5} Valentin Amrhein^{1,6}

¹University of Basel, Zoological Institute, Basel, Switzerland; ²Hintermann & Weber AG, Reinach, Switzerland; ³University of Bern Institute of Plant Sciences, Bern, Switzerland; ⁴Department of Plant Biology and Ecology, Faculty of Sciences and Technology, University of the Basque Country, UPV/EHU, Leioa, Spain; ⁵IKERBASQUE, Basque Foundation for Science, Bilbao, Spain and ⁶Swiss Ornithological Institute, Sempach, Switzerland

Correspondence

Tobias Roth Emails: roth@hintermannweber.ch, t.roth@unibas.ch

Funding information

Swiss National Science Foundation (grant no. 31003A_156294); Swiss Association Pro Petite Camargue Alsacienne, Fondation de bienfaisance Jeanne Lovioz, MAVA Foundation: Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 156294; Fondation de bienfaisance Jeanne Lovioz; Swiss Association Pro Petite Camargue Alsacienne

Handling Editor: Luísa Carvalheiro

Abstract

- 1. Functional traits are increasingly being used to understand the response of species to environmental change and their effects on ecosystem functioning. However, some ecologically important traits, such as plant height, influence the probability of species detection during field surveys. Imperfect detection of species could therefore bias measures of functional trait composition and diversity, leading to incorrect estimates of trait-environment relationships due to a process of "detection filtering." The importance of detection filtering for functional ecological studies remains unknown.
- 2. We used hierarchical models that account for detection filtering to analyse data on 1,296 vascular plant species sampled in 362 1-km² plots, distributed along a 2,460-m elevational gradient in Central Europe. We examined how detection filtering altered measures of functional diversity (multivariate functional richness and packing) and composition (community means of three traits). We also determined whether the strength of detection filtering varied over the gradient, to determine whether detection filtering biased trait-environment relationships.
- 3. Species detectability was correlated with all three functional traits tested in this study, meaning that short species with small seeds and high specific leaf area values were less likely to be detected. This suggests that imperfect detection has the potential to bias measures of functional composition. Generally, measures of functional composition were not strongly affected by detection filtering, but functional packing was underestimated when detection filtering was not accounted for. In addition to the traits, distributional characteristics were important; rare species and species occurring mainly at low elevations tended to have lower detection probabilities.
- 4. Overall, detection filtering did not strongly bias trait-environment relationships because the effects of the environment on functional composition and diversity were larger than the effects of detection.
- 5. Our results suggest that many measures of functional composition and diversity are robust to detection filtering, but some are likely biased. Functional ecologists should consider correcting for imperfect detection, and our approach provides a simple method to do so for a wide range of datasets.

KEYWORDS

altitudinal gradient, biodiversity monitoring, environmental filtering, functional diversity, multispecies occupancy models

1 | INTRODUCTION

Due to constraints in field surveys, it is usually not possible to obtain a complete list of species for a given community at a given time (Cardoso, Rigal, Borges, Carvalho, & Faith, 2014). Imperfect detection of species is therefore an important source of bias when comparing species richness among communities (Kéry & Schmid, 2004; Yoccoz, Nichols, & Boulinier, 2001). Recent studies have asked whether imperfect detection might also bias other important dimensions of biodiversity, such as functional diversity (Cardoso et al., 2014: Mihalievic, Joseph. & Johnson, 2015: van der Plas, van Klink. Manning, Olff, & Fischer, 2017). Jarzyna and Jetz (2016) suggested that detectability depends not only on site and survey conditions, but also on characteristics of individual species, that is on their traits. As measuring the diversity or composition of species traits within communities is at the core of functional ecology, imperfect detection could be of particular concern if detectabilities are correlated with certain species traits.

Some functional traits likely influence the probability of detecting species during field surveys. For example, plant height is commonly used in functional diversity research (Götzenberger et al., 2012), and larger plant species are likely more detectable than smaller species (Chen, Kéry, Plattner, Ma, & Gardner, 2013). Therefore, imperfect detection may "filter" the smaller species, or species with a specific growth habit, from a dataset, a process that we term detection filtering. Of course, detection will also depend on the type of community: in a heavily grazed grassland, all species are low growing, and therefore, short species are more likely to be detected than in an ungrazed system. We define detection filtering as a methodological process that selects which species are observed in a community, depending on their functional traits. Thus, we distinguish detection filtering from ecological filtering (or assembly rules sensu Keddy, 1992), caused by dispersal, environmental conditions or biotic interactions, which selects for or against species from the regional species pool depending on their functional traits. The detection filter could operate in a similar way to ecological filters by causing a signature of nonrandom functional composition.

This view of community assembly suggests a hierarchical process in which the species observed in a community depends on a series of ecological and detection filters. Hierarchical models (Kéry & Royle, 2016) could allow this process to be modelled. Hierarchical models contain two or more linear models that are conditionally related to each other; see Kéry and Royle (2016) for a detailed discussion of the framework of hierarchical models that can be fitted to presence/ absence or abundance data, accounting for imperfect detection. Guillera-Arroita (2017) provides a recent review, in which Figure 1 gives an overview of model structure and data needs of hierarchical models that account for imperfect detection.

To functionally characterize communities, occurrence data (occurrences of the species recorded in a set of sampling units) and trait data (a set of traits measured for the observed species) are typically combined to calculate metrics of functional diversity or composition. Often the trait data contain many missing values (Sandel et al., 2015).

FIGURE 1 Results of applying the method to meta-community data where 27% of 100 species were observed in less than 10% of 200 sites. Community means (CMs) of true communities (open orange dots, only known because data were simulated), CMs of observed communities (black dots) and CMs from detection-corrected communities (red +) along elevational gradient. The lines represent the regression lines from linear models with CMs as dependent variable and the gradient as predictor

Given that frequently measured species often have different trait values from rarely measured species (Sandel et al., 2015), gaps in trait data are likely nonrandom. Consequently, there is increasing concern that metrics of functional diversity are sensitive to gaps in the trait data (Májeková et al., 2016; Pakeman, 2014), and robust methodologies to impute missing values in trait data have been developed (Stekhoven & Buhlmann, 2012; Swenson, 2014). In contrast, only a few studies have investigated whether metrics of functional diversity are affected by missing species in the occurrence data (van der Plas et al., 2017). A reason for this might be that missing species lead to false absences (i.e. zeroes in the occurrence data), which are less obvious than gaps in trait data. Nonetheless, the consequences of missing species in occurrence data might be similar to those of missing trait values because in both cases, a nonrandom selection of species is excluded from the calculation of functional diversity measures.

In this study, we therefore ask whether false absences of species in the occurrence data may cause a signature of nonrandom functional composition, similar to ecological filtering. If overlooking species during sampling results in consistently lower or higher functional diversity, or in a consistent change in composition values, then comparative studies may not be strongly affected (van der Plas et al., 2017). However, if these biases are stronger in certain environmental conditions, then detection filtering could affect studies that correlate functional diversity/composition with environmental gradients (e.g. Soliveres & Maestre, 2014), as well as observational studies that use functional composition to predict ecosystem function metrics (e.g. Allan et al., 2015). We propose hierarchical models as a ready-to-use method to estimate the missing species in occurrence data.

To test this, we analysed a large dataset containing occurrences of plant species in 362 1-km^2 plots from the Swiss Biodiversity



Methods in Ecology and Evolution 919

Monitoring programme (BDM). We estimated average detection probabilities for 1,296 plant species and tested whether species-specific detection probabilities were correlated with certain functional trait values. Detection filtering will occur if detection probability depends on the expression of one or more functional traits. We then compared measures of the functional composition and diversity of observed communities with the same measures calculated for communities with imputed missed species (i.e. detection-corrected communities). To test whether detection filtering obscured trait-environment relationships, we examined how the difference between detection-corrected and observed communities changed along an elevational gradient. We focused on three functional traits that are whole-plant properties (Laughlin, 2014; Westoby, 1998): specific leaf area (ratio of fresh leaf area to leaf dry mass), plant canopy height and seed mass. We also determined the impact of detection filtering on the functional composition of the communities, measured as community means of single traits. We used trait means rather than the more commonly employed community-weighted means because we do not have abundance data for the species. We also calculated the effect of detection filtering on two functional diversity measures that describe the functional space of plant communities in terms of filling (measured as the functional richness) and packing (measured as the mean nearest neighbour distance; Laliberte & Legendre, 2010; Swenson & Weiser, 2014).

2 | MATERIALS AND METHODS

2.1 | Study area and plant data

Data collection took place between 2010 and 2014 within the Swiss Biodiversity Monitoring (BDM) scheme (Weber, Hintermann, & Zangger, 2004). A total of 428 study plots of 1 km² were distributed in a regular grid across the whole of Switzerland. Each year, onefifth of the sample plots were surveyed. These plots were chosen so that they constituted a regularly spaced subsample of all plots. The botanists performing the surveys received special training to reduce among-observer variation. The presence or absence of 2,674 plant taxa was recorded in each plot, along a 2.5-km transect that followed existing trails wherever possible. If no trails existed, surveyors marked the transect route in the field and plotted it on a map. At a sample plot, the transect was inspected once in spring and again in summer, ensuring that data collection spanned variation in flowering phenologies (Pearman & Weber, 2007). Exceptions were made for sample plots at high elevations, where only one inspection per field season was conducted because of the short growing period (however, we excluded these plots from analyses, see below). During each inspection, a surveyor recorded all plant species within 2.5 m of the transect while walking it in both directions at a slow speed of approximately 3 km/h.

The robustness of the survey methods was assessed in a previous study (Plattner, Weber, & Birrer, 2004) in which two botanists independently recorded 23 transects. The mean species richness on these transects was 250 species. The number of species recorded by the two botanists differed by 19.7 ± 4.9 species ($M \pm SE$) per transect, which corresponds to a relative surveyor effect of 7.9%. An

unpublished study was designed to quantify the bias introduced by choosing transects that follow existing trails. Species were recorded in 15 1-km² plots, using randomly placed transects of 2.5 km rather than the regular transects. On average, 10% more species were recorded in the plots when the regular transects were followed, compared to the randomly placed transects. In general, ruderal and grassland species tended to be more common along the regular trails, while species of wet meadows tended to be more common on the random transects. However, estimated indicator values, derived from expert knowledge (values 1-5 in 0.5 steps) for Swiss plants (Landolt et al., 2010), did not strongly differ between regular and random transects. The average indicator values for light, moisture, nutrients, reaction (content of free H-ions in the soil, from acid soils to soils rich in bases) or temperature of recorded species did not differ significantly between the types of transects, indicating that the regular transects capture well the functional composition of the plots.

We only analysed plots that had been visited twice in the year they were surveyed, that is we removed the high Alpine plots. This resulted in a sample of 362 study plots. Median elevation within plots ranged from 250 m to 2,710 m (a.s.l.), while the mean plot elevation was 1,104 m ± 612 (*SD*). On average, there was a gap of 77.6 days between the first and the second visit to the sample plots, and this difference was relatively stable across the elevational gradient (the number of days between first and second visits decreased by 0.02 days per 100 m along the elevational gradient, linear model, *p* = .905).

2.2 | Trait data

We used the leaf-height-seed plant ecology strategy scheme (Westoby, 1998) to focus our study on traits from multiple organs (Laughlin, 2014; Westoby, 1998). We analysed three traits: specific leaf area (ratio of fresh leaf area to leaf dry mass, SLA), canopy height (CH) and seed mass (SM). Data for these traits came from the LEDA trait database (Kleyer et al., 2008). The traits were partly correlated (SLA and CH: r = -.18; SLA SM: r = -.07; CH and SM: r = .31, all n = 1,296).

We were able to find trait data for most of the species, but some trait values were missing (Table 1). If the species with missing trait values are not a random subset of all species, ignoring those species will lead to biased functional diversity measures (Pakeman, 2014; Penone et al., 2014). The trait coverage at the species level (i.e. the percentage of species with available trait values) was lower than the coverage at the observation level (the percentage of observations with available trait values; Table 1). This suggests that rare species were more likely to have missing trait values than common species. Therefore, removing species with missing trait values from the analyses would be similar to removing rare species from the analyses, which is likely to bias the results (Lyons & Schwartz, 2001). To avoid this source of bias, we imputed the missing trait values and included all species in the analyses.

Missing trait values were imputed with random forest estimation (R package MISSFOREST version 1.4; Stekhoven & Buhlmann, 2012). Random forests constitute a popular method for many machine **TABLE 1** Variables describing trait values of 1,733 plant species that were recorded between 2010 and 2014 on at least one of the 362 plots from the Swiss Biodiversity Monitoring scheme. Values are given as medians (ranges in brackets) of the recorded species. Trait coverage is given both as the percentage of species and as the percentage of observations for which trait values were available

Trait	Median	Coverage species (%)	Coverage observations (%)
Specific leaf area: the ratio of fresh leaf area to leaf dry mass (mm ² /mg)	22.1 (2.6-150.6)	65	88
Canopy height (m)	0.3 (<0.1-65.0)	80	92
Seed mass (mg)	0.9 (<0.1-10'612.0)	70	87

learning tasks. The basis of this analysis is decision trees that are constructed using the species with available trait values. Application of random forest constructs and combines a multitude of such decision trees to increase predictive performance. To predict a missing trait value, random forest uses extant values of the same and other traits. Traditional regression techniques could also be used for this analysis, but decision trees allow more complex models to be fitted, with many interactive and nonlinear effects (Breiman, 2001). This approach avoids bias and retains relationships among traits (Penone et al., 2014). The random forest approach outperforms other approaches for estimating missing values in trait databases (Penone et al., 2014; Stekhoven & Buhlmann, 2012).

After imputation, trait values were log-transformed (Westoby, 1998) and then normalized to a mean of 0 and a *SD* of 1, allowing for direct comparison among traits (Schielzeth, 2010). Transformation also improves the normality of trait data (Májeková et al., 2016).

2.3 | Average species detection probabilities and detection-corrected meta-community

Hierarchical models may be expressed as f(y|z) because the outcome of the random variable y depends on the outcome of the random variable z. While z could be the true occurrence or the true abundance of a species, y would be the observed occurrence or observed abundance of that species. The observed occurrences depend on z and on some sort of measurement error (Kéry & Royle, 2016). In this study, we used the single-season occupancy model of MacKenzie et al. (2002), which is a particular type of hierarchical model in which the observations of a species at different sites (the data y) depend on the true occurrence for the species at the sites (occurrence z). We applied the single-season occupancy model to each recorded species separately. Each plot was visited twice during a single year. However, note that in each year between 2010 and 2014, one-fifth of the plots were surveyed and thus fieldwork lasted five years. We therefore considered the five-year span as the single closed season of the single-season occupancy model. Note that the single-season occupancy model assumes that it is

impossible to observe a species that does not occur in a site (i.e. no false positives). The single-season occupancy model is a hierarchical model with the form $f(y_{i,j}|z_i)$. The binary vector z_i , equal in length to the number of surveyed sites (indexed with *i*), describes the true occurrence of a species in the sites. The occupancy model that accounts for imperfect detection can concisely be formulated as

$$z_i \sim \text{Bernoulli}(\Psi_i)$$
 (1)

$$y_{i,j} \sim \text{Bernoulli}(z_i p_{i,j})$$
 (2)

where (1) describes occurrence, z_i , as a random variable that takes the value 1 with occurrence probability Ψ_i and where (2) is the methodological model that describes the data, $y_{i,i}$, as a random variable that takes the value 1 (i.e. a species is observed during visit j in site i) with probability $z_i p_{ii}$. Therefore, p_{ii} is the probability of detecting species *i*, assuming that it occurs in site *j*: a species that does not occur cannot be observed (because $z_i = 0$), and a species that does occur will be observed with detection probability p_{ij}. Predictors of occurrence probability Ψ_i and detection probability p_{ij} can then be added in a similar way as in standard generalized linear models. Because of the large elevational gradient, we incorporated linear and quadratic terms for elevation as predictors of occurrence probability Ψ_i . Elevation should summarize the effects of several environmental variables that affect the occurrence of the species (Chen et al., 2013). Because detection probability likely depends on phenology, we used the survey date (linear and quadratic terms) as predictors of detection probability p_{ii} (Chen et al., 2013).

Note that instead of applying the described single-season occupancy model to each of the observed species separately, an alternative approach would be to use a multispecies occupancy model to analyse the single-season occupancy of all species together in a single model. Such a multispecies occupancy model would provide a more synthetic framework to assess uncertainty in estimates and to share information across species about factors that might affect detection (Iknayan, Tingley, Furnas, & Beissinger, 2014). However, multispecies occupancy models are not implemented in accessible software packages and we therefore decided to apply single-season occupancy models to each species separately. To do so, we used the function *occu* of the R package UNMARKED (Fiske & Chandler, 2011).

We estimated the mean detectability of a species by averaging, across all plots, the probabilities of detecting it during at least one of the two surveys, using detection probabilities from the function *occu*. To test whether the average species detection probability was related to the trait values of the species, we applied a linear model with the logit-transformed $[log(\frac{p}{1-p})]$ average species detection probability as the dependent variable and specific leaf area, canopy height and seed mass as predictor variables. Further, widespread species are often locally common, making them easier to detect than sparsely distributed species. We added the estimated number of occupied plots per species as the fourth predictor in the linear model. Finally, we included the average elevation of plots occupied by a species as the fifth predictor to test whether detection probability varies with mean elevation of occurrence.

To estimate detection-corrected communities, we first calculated the occurrence of each species in each plot, that is the z_i values. The function *ranef* of the package UNMARKED was used to estimate posterior distributions of each z_i , using empirical Bayes methods, and the function *bup* of the package UNMARKED was used to extract the mode of the posterior probability for each z_i . After applying these functions to each observed species, we merged the z_i values from all species to obtain the estimated detection-corrected meta-community. Note that we consider the assemblage of species that occurs in a single plot as a community. Because $z_{i,k}$ describes 362 communities, we consider all plots together to be a meta-community, following the terminology of Kéry and Royle (2016).

2.4 | Functional composition and diversity

We calculated functional composition and diversity from the observed and detection-corrected meta-community and inferred the effect of detection filtering on functional diversity and composition along the elevational gradient. To estimate community functional composition, we calculated, for each community, the mean trait value across all species and did this separately for each of the three functional traits. We quantified functional diversity for each community as the multivariate convex hull volume, that is functional richness (FRic). The convex hull volume is a multidimensional volume that is the smallest convex hull enclosing a set of points (i.e. the trait data) and thus refers to the amount of functional space filled by the community (Villeger, Mason, & Mouillot, 2008). To calculate FRic, we used the function convhulln of the package FD (Laliberte & Legendre, 2010). Communities may fill a similar amount of the functional space but could differ in how tightly this volume is packed by species (Swenson & Weiser, 2014). Thus, we also calculated the mean nearest neighbour distance, using the Euclidean distance between species in multivariate trait space (Laliberte & Legendre, 2010; Swenson & Weiser, 2014).

2.5 | Simulation, worked example and R package DETECTIONFILTER

To make our approach more accessible, we applied the described method to simulated data. Note that extensive simulations to test the performance of hierarchical models have been carried out elsewhere (e.g. Bailey, Hines, Nichols, & MacKenzie, 2007); so our aim was merely to provide a worked example where modelled results can be compared with known values (i.e. the simulation setting) to show that our approach is able to reduce bias in the estimation of functional composition and diversity. We borrowed the ideas for the metacommunity simulation from chapter 11.2 in Kéry and Royle (2016) and added environmental and detection filtering to this simulation. To add environmental filtering, we assumed that the response of species to a gradient (i.e. the slope in the binomial GLM with occurrence as dependent variable and gradient as predictor variable) depends on their functional traits. To add detection filtering to the simulation, we assumed that the average detection probability of a species depends on its functional traits. See Data S1 where we describe the simulation in more detail, and Data S2 where we apply our method to the simulated data.

We bundled data and all R scripts into the R package DETECTIONFILTER, which will make it possible to reproduce our results and to employ the computational methods that are presented in this article. The package can be downloaded from www.github.com/TobiasRoth/detectionfilter. This package contains a function to estimate observed meta-community data from communities that are subject to ecological and detection filtering, the analysed plant data from the Swiss Biodiversity Monitoring, the values for the three functional traits (specific leaf area, canopy height and seed mass) for the recorded species, a vignette (i.e. a documentation to an R package) that develops the ideas behind the simulation of the meta-community, a vignette that describes the workflow to estimate the detection-corrected meta-community from observations, using the hierarchical models implemented in the R package UMARKED, and a vignette that describes all the analyses conducted in this project. The three vignettes are also available in Data S2.

3 | RESULTS

3.1 | Simulation example including rarely observed species

First, we applied our method to simulated meta-community data with 100 species and 200 study plots. In this example, two species were never observed and 27% of species were observed in less than 10% of plots (see "Effect of unobserved or rarely observed species" in Data S2 for more details). To simulate environmental filtering, we simulated a decrease in the community means (CMs) of a functional trait along an environmental gradient (open orange dots in Figure 1). Due to detection filtering, estimated CMs from observed communities (black dots in Figure 1) were biased towards high values and the estimated slope of CMs along the gradient was less steep than for true communities. If CMs are calculated from detection-corrected communities (red + in Figure 1), bias is reduced and the slope along the gradient more closely follows the slope for the true communities.

3.2 | Swiss plant communities

In total, 1,733 species were recorded on the 362 plots. Including the data of both visits, 256.0 \pm 52.0 (average \pm *SD*) species were observed per plot. The algorithm of the single-season occupancy model failed to converge consistently for species with fewer than four observations, so we analysed only the 1,296 species (75% of all recorded species) with at least four records. To determine the effect on functional diversity of removing the 437 species with fewer than four observations, we calculated all measures of functional composition and diversity for communities including all observed species and for communities with the rare (<4 plots) species excluded. We found removing the very rarest species made little difference, because measures of functional composition and diversity calculated from communities with all species were strongly correlated with measures from communities with the rare species excluded (all Pearson correlation coefficients >0.995).

The detection probability for a species with average trait values and an average number of occurrences (i.e. the number of plots a species was estimated to occur in), at average elevation, was 0.94. This value corresponds to the back-transformed intercept of the linear model that predicts species' detectabilities (Table 2). While detection probability decreased for plants with larger specific leaf area, it increased for plants with larger canopy height and higher seed mass. The effect of height on detection is clear, and most of the larger species (bushes and trees) had average detection probabilities >.95. A good example of how SLA affected detection probability is Sedum acre: while it is a relatively short and small-seeded species, it is still rather easy to detect (average detection probability = .92) due to its long-lived, fleshy leaves with low specific leaf area. The differences in detection between Convallaria majalis (average detection probability = .95) and Allium ursinum (average detection probability = .67) are also instructive: while canopy height and specific leaf area are similar for the two species, Convallaria majalis has conspicuous fruits with a few large seeds, in contrast to the inconspicuous fruits and small seeds of Allium ursinum. Compared to the functional traits, however, elevation and the number of occurrences (i.e. how widespread a species was) were stronger predictors of detection probability (Table 2). In general, species that were more widespread or occurred on average at higher elevations were more likely to be detected.

On average, 94.8% of the species that were estimated to occur in a community were actually observed. The proportion of observed species was highest at around 900 m and decreased with increasing elevation (Figure 2a). We also examined the characteristics of the species that were predicted to be missing from each community, and calculated the total occurrence and mean elevation of occurrence for each species. We then plotted the average characteristics (total occurrence and average elevation) of the observed and missing species, for each community, against elevation. Up to about 750 m, the missing species, that is those that remained undetected within a community, were

TABLE 2 Results from linear model predicting species' detection probabilities (logit-transformed) from the functional traits, the number of occurrences and the average elevation of occurrence of a species. Given are the results for the 1,296 plant species for which detection-corrected occurrences were estimated. Effect sizes are the estimated intercepts and slopes, *SE* is the standard error of the estimates, and the *p*-value is for the null hypothesis that the variables had no effect on species' detection probability. Note that the results are based on the normalized log-transformed variable data and thus allow for direct comparison of effect sizes

Intercept and predictor			
variables	Effect size	SE	p-value
Intercept	2.795	0.082	
Specific leaf area	-0.165	0.080	.040
Canopy height	0.219	0.092	.017
Seed mass	0.386	0.083	<.001
Number of occurrences	0.610	0.077	<.001
Average elevation of occurrence	0.564	0.097	<.001

generally species that occurred at lower elevations (Figure 2b) or that were rarer (i.e. occurred on fewer plots; Figure 2c) than the observed species. The species that most often remained undetected in communities below 750 m were *Buglossoides arvensis*, a weed of arable land, and *Helianthus tuberosus*, a currently spreading invasive species that flowers late. In contrast, above 1,000 m the missing species tended to be mostly those that occur at higher elevations than the observed species (Figure 2b), and in communities around 1,000 m, the missing species tended to be quite widespread ones that can occur at a range of elevations. The species that most often remained undetected in communities between 750 m and 1,250 m was *Descurainia sophia*, a ruderal plant of fields and dry rock faces.

The bias caused by detection filtering was relatively weak for measures of functional composition (Figure 3) but had larger effects on some functional diversity metrics (Figure 4). We believe that a study along an elevational gradient should be able to detect environmentally induced differences in functional composition or diversity that occur across a 100-m elevation difference. If detection filtering has a larger effect on functional composition than the effect of 100-m elevation difference, we consider that it could obscure traitenvironment relationships. We therefore (arbitrarily) defined detection filtering to have a substantial effect if it caused a larger change in functional composition/diversity than we observed per 100 m along the elevational gradient. Following this definition, imperfect detection biased the estimates of community mean specific leaf area in 17.1% of the studied communities (Figure 3a), biased the estimates of community mean canopy height in 2.8% of communities (Figure 3b) and biased the estimates of community mean seed mass in 31.5% of communities (Figure 3c). Following the same definition, the effect of removing rare (<4 observations) species was weaker: removing rare species was relevant in only 0.6% of communities for SLA, in 0.3% of communities for canopy height, in 1.4% of communities for seed mass, in 3.3% of communities for functional richness and in 13% of communities for functional packing.

Observed and detection-corrected estimates of community means of specific leaf area, canopy height and seed mass were all strongly correlated (all r > .99). For functional diversity, we found that detection filtering mainly affected functional packing and not functional richness: detection filtering affected functional richness in only 7.2% of communities but affected functional packing in 95.6% of communities. Still, correlations between estimates of observed and detectioncorrected communities were high (FRic: r = .997, mnnd: r = .987). Furthermore, random removal of species from communities led to a similar decline in functional packing as compared to the effect of detection filtering (see Figure 3.3 in Data S3).

Regardless of whether we used detection-corrected or observed meta-community data, we still found the same pattern of change in functional diversity and composition along the elevational gradient. Community mean values of SLA, height and seed size declined along the elevational gradient, with the steepest decline occurring for canopy height (Figure 3). The response of functional diversity metrics was less clear: while functional richness showed a rather continuous decline along the elevation gradient, the peak of functional packing



FIGURE 2 (a) Change in the observed proportion of occurring species plotted against community elevation. (b) Mean elevation of species occurrence averaged for observed species (red points) and for species that were estimated to occur in a community but that were not detected (i.e. overlooked species, blue points), plotted against community elevation. (c) Number of occurrences per species averaged for observed species (red points) and for overlooked species (blue points), plotted against community elevation. Smoothed curves are predictions from generalized additive models (GAMs)



FIGURE 3 Change in community means (CMs) of log-transformed and normalized (z-score) trait values along the elevational gradient for the three functional traits: (a) specific leaf area, (b) canopy height and (c) seed mass. Points give CMs of the 362 observed communities that are not corrected for detection filtering. Coloured points indicate communities where imperfect detection affected estimates of CMs more than the change in community composition we observed per 100 m along the elevational gradient (red points: observed CMs are lower than detection-corrected CMs; blue points: observed CMs are larger than detection-corrected CMs). The lines represent the predictions from the generalized additive model (GAM) applied to the observed communities (dotted line) and to the detection-corrected communities (solid line)

(i.e. smallest mean nearest neighbour distance) occurred at around 1,700 m (Figure 4a,b). The peak in functional packing coincided with the peak in the number of species in a community (Figure 4c).

4 | DISCUSSION

One of the greatest challenges when using ecological data is imperfect detection of species (Kéry & Schmid, 2004; Yoccoz et al., 2001). This is not only a problem for animals but is also an issue in studies on plants (Chen et al., 2013). Perfect detection should not be assumed in plant surveys, even for easily recognizable species under ideal survey conditions (Ng & Driscoll, 2015). Although there have been great efforts to account for imperfect detection in animal and plant surveys during the last decade (Bailey, MacKenzie, & Nichols, 2013; Chen et al., 2013; Kéry & Royle, 2016), imperfect detection has been largely neglected in community ecology research until recently (Cardoso et al., 2014; Jarzyna & Jetz, 2016; Mihaljevic et al., 2015; Tingley, Ruiz-Gutierrez, Wilkerson, Howell, & Siegel, 2016). Our study shows that imperfect detection may bias estimation of functional diversity in particular. We also show that hierarchical models can easily be applied to estimate detection filtering, even in cases where a high proportion of species are only rarely observed (Figure 1).



FIGURE 4 Changes in (a) functional richness (convex hull volume of the three functional dimensions specific leaf area, canopy height and seed mass), (b) functional packing (mean nearest neighbour distance) and (c) taxonomic diversity (number of species) along the elevational gradient. Points give the estimates of the 362 observed communities that are not corrected for detection filtering. Coloured points indicate communities where imperfect detection affected estimates more than the change in community diversity we observed per 100 m along the elevational gradient (red points: observed estimates are below the detection-corrected estimates; blue points: observed estimates are above the detection-corrected estimates). The lines represent the predictions from the generalized additive model (GAM) applied to the observed communities (dotted lines) and to the detection-corrected communities (solid line)

Species detectability depended on all three functional traits tested, meaning that short species with small seeds and high specific leaf area (SLA) values were less likely to be detected. Lower values of SLA are related to longer leaf life span (Reich, Walters, & Ellsworth, 1997), and leaves with a longer life span are likely to be present during both spring and autumn surveys, meaning that they may be more likely to be detected than leaves with short life spans. The reasons for the correlation between seed size and detectability are less clear. However, small-seeded species are usually ruderals with a high dispersal ability but short life span (Diaz et al., 2016). These plants may have morevariable population dynamics, consistent with occurring in relatively ephemeral habitats and making the species harder to detect. Finally, the correlation between canopy height and detectability of species is well known (Chen et al., 2013), as short species are less visible than large ones.

4.1 | Bias from nonrandom under-sampling

Our results suggest that imperfect detection may bias measures of some of the most commonly used functional traits of plants (Götzenberger et al., 2012). While many measures of functional diversity have been shown to be sensitive to random under-sampling, nonrandom under-sampling is expected to cause even more bias (van der Plas et al., 2017). Recent work on plant strategies shows two main axes of variation: a resource economics axis (related to specific leaf area) and a size axis related to height and seed size (Diaz et al., 2016; Westoby, 1998). The traits related to these axes are important predictors of plant responses to the environment and of effects on ecosystem function (Lavorel et al., 2011). Our results suggest that detection filtering may lead to underestimation of the occurrence of ruderal (short, small seeded) and fast-growing species, which could mean that functionally important species are overlooked. Detection filtering might bias functional trait proxies for ecosystem functions related to productivity and fast nutrient cycling, which are promoted by species with high SLA (Lavorel et al., 2011) or proxies for dispersal functions if these are promoted by small-seeded species.

Nevertheless, detection filtering only had marginal effects on functional composition and diversity in our study. Inclusion of the missed species mainly increased functional packing of communities (lower mean nearest neighbour distance of detection-corrected communities in Figure 4b) but hardly affected functional richness (Figure 4a). This suggests that the missing species were similar in terms of their combination of leaf-height-seed traits to the species that had been detected in that community. Indeed, detection effects on functional packing disappeared when we randomly removed species from detectioncorrected communities until they contained the same number of species as the observed communities. The functional similarity of observed and missed species is also evident when looking at the three traits separately, where the missed species hardly affected the community means of the three functional traits (Figure 3). This suggests that many functional trait metrics are relatively robust to detection filtering. However, care should be taken when estimating measures of functional packing as these metrics are sensitive to species detection. Even if the bias caused by missed species is not greater than that caused by randomly missing species, it is still important to know whether a lot of species are likely to be missing or not and whether the proportion of missed species is similar between communities.

4.2 | Applicability and adaptability of the approach

Overall, high detection probabilities in our study minimized the potential for detection filtering to be a major source of error. In cases where survey and sampling methods lead to lower detection rates or a species group has lower detectability than plants (e.g. birds), greater potential for inferential errors exists, and correcting functional diversity metrics for detection filtering is likely to be even more important. Our approach provides a method to correct functional diversity and composition metrics for detection filtering. As true occupancy can be estimated independently from detection probability using occupancy models (MacKenzie & Kendall, 2002), the derived estimates of functional composition and diversity, obtained from these corrected occupancy estimators, will be less biased by imperfect detection (Iknayan et al., 2014). This means that where repeated surveys have been carried out, it is possible to correct functional composition and diversity measures using the approach outlined here. However, methods to estimate detection probability are also available for single-visit surveys (Guillera-Arroita, 2017). Examples include time-to-detection methods (Bornand, Kéry, Bueche, & Fischer, 2014; Garrard, Bekessy, McCarthy, & Wintle, 2008; Guillera-Arroita, Morgan, Ridout, & Linkie, 2011). Like occupancy models, these methods provide an estimate of true occupancy. It should thus be possible to adopt the approach outlined here to these methods. Furthermore, methods such as distance sampling (Buckland, Anderson, Burnham, & Laake, 2005) or N-mixture models allow estimation of detection probability from abundance data (Rovle, 2004).

4.3 | Effect of species distribution

Our results also suggest that attributes related to the distribution of a species are important in determining its detection probability (Table 2). The number of plots on which a species occurred correlated with detection probability, meaning that widespread species had generally higher detectabilities than species with restricted distributions. Species with a restricted distribution are often also locally rare (Gaston & Lawton, 1990), and it is well established that local abundance is a key variable predicting detection probability (Chen et al., 2013; Royle & Nichols, 2003). Note, however, that in communities at mid-elevations, at around 1,000 m, missed species were on average more widespread than observed species. Ruderals that can cope with human disturbance at low elevations but also with comparatively low temperatures at higher elevations might be relatively rare at midelevations, because disturbed habitat is rare at these elevations. As a consequence, relatively widespread species might be difficult to detect at mid-elevations. Taking abundance into account in indices of functional composition, that is by calculating community-weighted mean functional traits, offers promising insights into how species respond to environmental changes (de Bello, Leps, Lavorel, & Moretti, 2007), and community-weighted means are often good predictors of ecosystem function. As the probability of detecting a species is positively related to its abundance (Iknayan et al., 2014), using N-mixture models to infer how detection filtering alters abundance-weighted measures of functional diversity seems an important avenue for future research.

4.4 | Elevation and effects of imperfect detectability

More surprisingly, the species with high average detection probability tended to be those occurring at higher altitudes. Although the vegetation period is shorter at higher elevations, the period between first and second visit was about 78 days independent of elevation. We thus speculate that detection probabilities of species increased with elevation because the two visits spanned an increasing proportion of the vegetation period and that it was thus less likely to miss early or late flowering species. Although detection probabilities increased with elevation for individual species, at the community level, a lower proportion of the total species were detected at high elevation (Figure 2a). This apparent contradiction might be explained by the fact that the species level model is already corrected for the frequency of the species, as we included the number of occurrences of each species alongside the average elevation of occurrence. However, the species that occur at high elevations tended to be rarer (occur in fewer plots). This means that the decrease in community-level detectability with elevation (Figure 2a) could be driven by an increase in the proportion of rare species in high elevation communities.

There was only a weak indication that imperfect detection would change conclusions about how altitude affected functional composition and diversity. Similar to the effects at the species level, at low elevations, community means of height and seed size were slightly overestimated and SLA was slightly underestimated due to imperfect detection. Apparently, the botanists that walked the 2.5-km transects detected more of the large species and fewer of the short and smallseeded species. At high elevations, taller species with higher SLA and higher seed mass were slightly more likely to be missed (Figure 3). Species with this trait combination are not well adapted to high elevation environments (von Arx, Edwards, & Dietz, 2006) and thus may occur only in sheltered places, in dwarf forms, as seedlings, as young plants or simply in lower numbers. This could be one reason why the detection probabilities of these species were lower compared to other species in alpine environments. However, many studies are interested in the relative degree of change in functional composition/diversity along an environmental gradient, and in this case, it is mainly important that detection filtering does not change the correlation between observed and predicted values. The correlation between observed and detection-corrected measures of functional composition and diversity should be high if the effect of imperfect detection is constant across sites and if there are large differences among sites along the elevational gradient. For all tested measures of functional composition and diversity, we found correlations of >.99. These high positive correlations suggest that differences in the functional composition between sites along elevations are much more important than effects due to detection filtering. Thus, detection filtering did not obscure trait-environment relationships in our study, which is encouraging because many studies assessing trait-environment relationships have not corrected for imperfect detection. However, to compare functional composition values between studies with different sampling designs or sampling intensity, it might be important to correct for the bias caused by detection filtering.

Elevation was generally a major driver of functional composition and diversity in our study. Communities changed from containing larger plants with heavier seeds and higher specific leaf area in the lowlands to containing smaller plants with lighter seeds and lower specific leaf area (suggesting leafs with longer life span; Westoby, 1998) in alpine environments. These trends are consistent with earlier studies (Dubuis et al., 2013; Qi et al., 2015) and are likely to be caused by alpine plants adopting a more conservative growth strategy (von Arx et al., 2006). For example, alpine plants should allocate a higher proportion of their resources to vegetative growth (e.g. into leaves with smaller SLA), and because of trade-offs this may be reflected in reduced reproductive effort (e.g. the production of smaller seeds; von Arx et al., 2006; Nobis & Schweingruber, 2013). These gradual and consistent trends along the entire elevational gradient also suggest that climate warming will reduce populations of plants with slow growth strategies.

Patterns of functional traits can help to elucidate the underlying mechanisms of community assembly (Janecek et al., 2013; Mayfield et al., 2010), and functional diversity provides an indicator of ecosystem functioning (Diaz et al., 2007). Using traits as indicators of functioning in biodiversity experiments (e.g. Flynn, Mirotchnick, Jain, Palmer, & Naeem, 2011) is not likely to be biased by problems of imperfect detection. However, because ecosystems are managed and ecosystem services are delivered at landscape scales (Quijas et al., 2012), and because manipulative experiments are rarely feasible at such large scales (Barbaro, Giffard, Charbonnier, van Halder, & Brockerhoff, 2014), observational approaches using functional traits to map services offer much promise (de Bello et al., 2010; Violle, Reich, Pacala, Enquist, & Kattge, 2014). In these cases, imperfect detection may be the rule rather than the exception (Chen et al., 2013), and our results suggest that key functional traits are correlated with detection probabilities, meaning that measures of functional composition and diversity could be biased by detection filtering. Our analysis suggests this is not a major problem for functional composition, although this may or may not hold in other studies, but that diversity metrics based on the degree of packing of species in functional space are underestimated by detection filtering. Future studies should therefore consider correcting functional diversity measures for imperfect detection. Our approach provides a simple method to do this for a large number of species and communities.

ACKNOWLEDGEMENTS

Financial support for the study was provided by the Swiss National Science Foundation (grant no. 31003A_156294), the Swiss Association Pro Petite Camargue Alsacienne, the Fondation de bienfaisance Jeanne Lovioz, and the MAVA Foundation. We also thank the dedicated and qualified botanists who conducted fieldwork, and the Swiss Federal Office for the Environment (FOEN) for providing data from the Swiss Biodiversity Monitoring scheme. Fränzi Korner-Nievergelt gave valuable comments on the statistical analyses. Calculations were performed at the sciCORE (http://scicore.unibas.ch/) scientific computing core facility at the University of Basel. The authors have no conflict of interest to declare.

AUTHORS' CONTRIBUTIONS

T.R. and V.A. conceived the ideas and designed methodology, and additional analyses were suggested by E.A. and P.B.P. T.R. accessed

the trait data, performed the analyses and wrote the first draft of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

All data and code to run the analyses are available from the R package DETECTIONFILTER that can be downloaded from Github at https:// github.com/TobiasRoth/detectionfilter archived on Zenodo https:// doi.org/10.5281/zenodo.1065446.

ORCID

Tobias Roth () http://orcid.org/0000-0001-9742-8297

REFERENCES

- Allan, E., Manning, P., Alt, F., Binkenstein, J., Blaser, S., Blüthgen, N., ... Fischer, M. (2015). Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. *Ecology Letters*, 18, 834–843. https://doi.org/10.1111/ ele.12469
- von Arx, G., Edwards, P. J., & Dietz, H. (2006). Evidence for life history changes in high-altitude populations of three perennial forbs. *Ecology*, 87, 665–674. https://doi.org/10.1890/05-1041
- Bailey, L. L., Hines, J. E., Nichols, J. D., & MacKenzie, D. I. (2007). Sampling design trade-offs in occupancy studies with imperfect detection: Examples and software. *Ecological Applications*, 17, 281–290. https:// doi.org/10.1890/1051-0761(2007)017[0281:SDTIOS]2.0.CO;2
- Bailey, L. L., MacKenzie, D. I., & Nichols, J. D. (2013). Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5, 1269–1279.
- Barbaro, L., Giffard, B., Charbonnier, Y., van Halder, I., & Brockerhoff, E. G. (2014). Bird functional diversity enhances insectivory at forest edges: A transcontinental experiment. *Diversity and Distributions*, 20, 149–159. https://doi.org/10.1111/ddi.12132
- de Bello, F., Lavorel, S., Diaz, S., Harrington, R., Cornelissen, J. H. C., Bardgett, R. D., ... Harrison, P. A. (2010). Towards an assessment of multiple ecosystem processes and services via functional traits. *Biodiversity and Conservation*, 19, 2873–2893. https://doi.org/10.1007/ s10531-010-9850-9
- de Bello, F., Leps, J., Lavorel, S., & Moretti, M. (2007). Importance of species abundance for assessment of trait composition: An example based on pollinator communities. *Community Ecology*, 8, 163–170. https://doi. org/10.1556/ComEc.8.2007.2.3
- Bornand, C. N., Kéry, M., Bueche, L., & Fischer, M. (2014). Hide and seek in vegetation: Time-to-detection is an efficient design for estimating detectability and occurrence. *Methods in Ecology and Evolution*, 5, 433– 442. https://doi.org/10.1111/2041-210X.12171
- Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https:// doi.org/10.1023/A:1010933404324
- Buckland, S. T., Anderson, D. R., Burnham, K. P., & Laake, J. L. (2005). Distance sampling. New York: John Wiley & Sons, Ltd.
- Cardoso, P., Rigal, F., Borges, P. A. V., Carvalho, J. C., & Faith, D. (2014). A new frontier in biodiversity inventory: A proposal for estimators of phylogenetic and functional diversity. *Methods in Ecology and Evolution*, 5, 452–461. https://doi.org/10.1111/2041-210X.12173
- Chen, G., Kéry, M., Plattner, M., Ma, K., & Gardner, B. (2013). Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology*, 101, 183–191. https://doi. org/10.1111/1365-2745.12021

- Diaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., ... Gorne, L. D. (2016). The global spectrum of plant form and function. *Nature*, 529, 167–U173.
- Diaz, S., Lavorel, S., de Bello, F., Quetier, F., Grigulis, K., & Robson, M. (2007). Incorporating plant functional diversity effects in ecosystem service assessments. Proceedings of the National Academy of Sciences of the United States of America, 104, 20684–20689. https://doi.org/10.1073/ pnas.0704716104
- Dubuis, A., Rossier, L., Pottier, J., Pellissier, L., Vittoz, P., & Guisan, A. (2013). Predicting current and future spatial community patterns of plant functional traits. *Ecography*, *36*, 1158–1168. https://doi. org/10.1111/j.1600-0587.2013.00237.x
- Fiske, I. J., & Chandler, R. B. (2011). UNMARKED: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, 43, 1–23.
- Flynn, D. F. B., Mirotchnick, N., Jain, M., Palmer, M. I., & Naeem, S. (2011). Functional and phylogenetic diversity as predictors of biodiversityecosystem-function relationships. *Ecology*, 92, 1573–1581. https://doi. org/10.1890/10-1245.1
- Garrard, G. E., Bekessy, S. A., McCarthy, M. A., & Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology*, *33*, 986–998. https://doi.org/10.1111/j.1442-9993.2008.01869.x
- Gaston, K. J., & Lawton, J. H. (1990). Effects of scale and habitat on the relationship between regional distribution and local abundance. *Oikos*, 58, 329–335. https://doi.org/10.2307/3545224
- Götzenberger, L., de Bello, F., Brathen, K. A., Davison, J., Dubuis, A., Guisan, A., ... Zobel, M. (2012). Ecological assembly rules in plant communitiesapproaches, patterns and prospects. *Biological Reviews*, *87*, 111–127. https://doi.org/10.1111/j.1469-185X.2011.00187.x
- Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: Advances, challenges and opportunities. *Ecography*, 40, 281–295. https://doi. org/10.1111/ecog.02445
- Guillera-Arroita, G., Morgan, B. J. T., Ridout, M. S., & Linkie, M. (2011). Species occupancy modeling for detection data collected along a transect. *Journal of Agricultural Biological and Environmental Statistics*, 16, 301–317. https://doi.org/10.1007/s13253-010-0053-3
- Iknayan, K. J., Tingley, M. W., Furnas, B. J., & Beissinger, S. R. (2014). Detecting diversity: Emerging methods to estimate species diversity. *Trends in Ecology & Evolution*, 29, 97–106. https://doi.org/10.1016/ j.tree.2013.10.012
- Janecek, S., de Bello, F., Hornik, J., Bartos, M., Cerny, T., Dolezal, J., ... Klimesova, J. (2013). Effects of land-use changes on plant functional and taxonomic diversity along a productivity gradient in wet meadows. *Journal* of Vegetation Science, 24, 898–909. https://doi.org/10.1111/jvs.12012
- Jarzyna, M. A., & Jetz, W. (2016). Detecting the multiple facets of biodiversity. Trends in Ecology & Evolution, 31, 527–538. https://doi. org/10.1016/j.tree.2016.04.002
- Keddy, P. A. (1992). Assembly and response rules 2 goals for predictive community ecology. *Journal of Vegetation Science*, 3, 157–164. https:// doi.org/10.2307/3235676
- Kéry, M., & Royle, J. (2016). Applied hierarchical modeling in ecology. Analysis of distribution, abundance and species richness in R and BUGS. Amsterdam, the Netherlands: Academic Press.
- Kéry, M., & Schmid, H. (2004). Monitoring programs need to take into account imperfect species detectability. *Basic and Applied Ecology*, 5, 65–73. https://doi.org/10.1078/1439-1791-00194
- Kleyer, M., Bekker, R. M., Knevel, I. C., Bakker, J. P., Thompson, K., Sonnenschein, M., ... Peco, B. (2008). The LEDA traitbase: A database of life-history traits of the Northwest European flora. *Journal of Ecology*, 96, 1266–1274. https://doi.org/10.1111/j.1365-2745.2008.01430.x
- Laliberte, E., & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, *91*, 299–305. https://doi.org/10.1890/08-2244.1

- Landolt, E., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F., Lämmler, W., ... Theurillat, J. (2010). Flora indicativa, ecological indicator values and biological attributes of the flora of Switzerland and the Alps. Bern, Switzerland: Haupt-Verlag.
- Laughlin, D. C. (2014). The intrinsic dimensionality of plant traits and its relevance to community assembly. *Journal of Ecology*, 102, 186–193. https://doi.org/10.1111/1365-2745.12187
- Lavorel, S., Grigulis, K., Lamarque, P., Colace, M.-P., Garden, D., Girel, J., ... Douzet, R. (2011). Using plant functional traits to understand the landscape distribution of multiple ecosystem services. *Journal of Ecology*, 99, 135–147. https://doi.org/10.1111/j.1365-2745.2010.01753.x
- Lyons, K. G., & Schwartz, M. W. (2001). Rare species loss alters ecosystem function – invasion resistance. *Ecology Letters*, 4, 358–365. https://doi. org/10.1046/j.1461-0248.2001.00235.x
- MacKenzie, D. I., & Kendall, W. L. (2002). How should detection probability be incorporated into estimates of relative abundance? *Ecology*, 83, 2387–2393. https://doi.org/10.1890/0012-9658(2002)083[238 7:HSDPBI]2.0.CO;2
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*, 2248–2255.
- Májeková, M., Paal, T., Plowman, N. S., Bryndová, M., Kasari, L., Norberg, A., ... de Bello, F. (2016). Evaluating functional diversity: Missing trait data and the importance of species abundance structure and data transformation. *PLoS ONE*, 11, e0149270. https://doi.org/10.1371/journal. pone.0149270
- Mayfield, M. M., Bonser, S. P., Morgan, J. W., Aubin, I., McNamara, S., & Vesk, P. A. (2010). What does species richness tell us about functional trait diversity? Predictions and evidence for responses of species and functional trait diversity to land-use change. *Global Ecology and Biogeography*, 19, 423–431.
- Mihaljevic, J. R., Joseph, M. B., & Johnson, P. T. J. (2015). Using multispecies occupancy models to improve the characterization and understanding of metacommunity structure. *Ecology*, 96, 1783–1792. https://doi. org/10.1890/14-1580.1
- Ng, K., & Driscoll, D. A. (2015). Detectability of the global weed Hypochaeris radicata is influenced by species, environment and observer characteristics. *Journal of Plant Ecology*, *8*, 449–455. https://doi.org/10.1093/jpe/ rtu032
- Nobis, M. P., & Schweingruber, F. H. (2013). Adult age of vascular plant species along an elevational land-use and climate gradient. *Ecography*, 36, 1076–1085. https://doi.org/10.1111/j.1600-0587.20 13.00158.x
- Pakeman, R. J. (2014). Functional trait metrics are sensitive to the completeness of the species' trait data? *Methods in Ecology and Evolution*, 5, 9–15. https://doi.org/10.1111/2041-210X.12136
- Pearman, P., & Weber, D. (2007). Common species determine richness patterns in biodiversity indicator taxa. *Biological Conservation*, 138, 109– 119. https://doi.org/10.1016/j.biocon.2007.04.005
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., ... Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods in Ecology and Evolution*, *5*, 961–970. https://doi. org/10.1111/2041-210X.12232
- Plattner, M., Weber, D., & Birrer, S. (2004). Data quality in monitoring plant species richness in Switzerland. *Community Ecology*, 5, 135–143. https://doi.org/10.1556/ComEc.5.2004.1.13
- Qi, W., Zhou, X., Ma, M., Knops, J. M. H., Li, W., & Du, G. (2015). Elevation, moisture and shade drive the functional and phylogenetic meadow communities' assembly in the northeastern Tibetan Plateau. *Community Ecology*, 16, 66–75. https://doi.org/10.1556/168.2015.16.1.8
- Quijas, S., Jackson, L. E., Maass, M., Schmid, B., Raffaelli, D., & Balvanera, P. (2012). Plant diversity and generation of ecosystem services at the landscape scale: Expert knowledge assessment. *Journal of Applied Ecology*, 49, 929–940. https://doi.org/10.1111/j.1365-2664.2012.02153.x

- Reich, P. B., Walters, M. B., & Ellsworth, D. S. (1997). From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 13730–13734. https://doi.org/10.1073/pnas.94.25.13730
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60, 108–115. https://doi. org/10.1111/j.0006-341X.2004.00142.x
- Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, 84, 777–790. https:// doi.org/10.1890/0012-9658(2003)084[0777:EAFRPA]2.0.CO;2
- Sandel, B., Gutierrez, A. G., Reich, P. B., Schrodt, F., Dickie, J., & Kattge, J. (2015). Estimating the missing species bias in plant trait measurements. *Journal of Vegetation Science*, 26, 828–838. https://doi.org/10.1111/ jvs.12292
- Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1, 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x
- Soliveres, S., & Maestre, F. T. (2014). Plant-plant interactions, environmental gradients and plant diversity: A global synthesis of community-level studies. *Perspectives in Plant Ecology Evolution and Systematics*, 16, 154– 163. https://doi.org/10.1016/j.ppees.2014.04.001
- Stekhoven, D. J., & Buhlmann, P. (2012). MISSFOREST-non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597
- Swenson, N. G. (2014). Phylogenetic imputation of plant functional trait databases. *Ecography*, 37, 105–110. https://doi. org/10.1111/j.1600-0587.2013.00528.x
- Swenson, N. G., & Weiser, M. D. (2014). On the packing and filling of functional space in eastern North American tree assemblages. *Ecography*, 37, 1056–1062.
- Tingley, M. W., Ruiz-Gutierrez, V., Wilkerson, R. L., Howell, C. A., & Siegel, R. B. (2016). Pyrodiversity promotes avian diversity over the decade following forest fire. *Proceedings of the Royal Society B-Biological Sciences*, 283, 20161703. https://doi.org/10.1098/ rspb.2016.1703

- van der Plas, F., van Klink, R., Manning, P., Olff, H., & Fischer, M. (2017). Sensitivity of functional diversity metrics to sampling intensity. *Methods in Ecology and Evolution*, *8*, 1072–1080. https://doi. org/10.1111/2041-210X.12728
- Villeger, S., Mason, N. W. H., & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290–2301. https://doi.org/10.1890/07-1206.1
- Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J., & Kattge, J. (2014). The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 13690–13696. https://doi.org/10.1073/pnas.1415442111
- Weber, D., Hintermann, U., & Zangger, A. (2004). Scale and trends in species richness: Considerations for monitoring biological diversity for political purposes. *Global Ecology and Biogeography*, 13, 97–104. https:// doi.org/10.1111/j.1466-882X.2004.00078.x
- Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. PlantandSoil, 199, 213–227. https://doi.org/10.1023/A:1004327224729
- Yoccoz, N. G., Nichols, J. D., & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution*, 16, 446–453. https://doi.org/10.1016/S0169-5347(01)02205-4

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Roth T, Allan E, Pearman PB, Amrhein V. Functional ecology and imperfect detection of species. *Methods Ecol Evol*. 2018;9:917–928. <u>https://doi.</u> org/10.1111/2041-210X.12950