Journal of Vegetation Science ■■ (2014)

Observation bias and its causes in botanical surveys on high-alpine summits

Sarah Burg, Christian Rixen, Veronika Stöckli & Sonja Wipf

Keywords
Between-observer variability; Long-term monitoring; Pseudo-turnover; Revisitation studies; Sampling error; Species detectability; Swiss Alps; Vegetation records

Nomenclature
Lauber and Wagner (2009)

Received 17 December 2013
Accepted 30 May 2014
Co-ordinating Editor: Alicia Acosta

Wipf, S. (corresponding author, wipf@slf.ch),
Burg, S. (sarahburg@gmx.ch),
Rixen, C. (rixen@slf.ch), **&**
Stöckli, V.(stoeckli@bergwelten21.ch): WSL Institute for Snow and Avalanche Research SLF, Fluelastrasse 11, Davos Dorf 7260, Switzerland
Burg, S. (sarahburg@gmx.ch): Department of Biology, ETH Zurich, Wolfgang-Pauli-Str. 27, Zürich 8093, Switzerland

Abstract

Question: To determine long-term vegetation changes in revisitation studies, it is crucial to know how much of measured species turnover over time can be attributed to pseudo-turnover (i.e. turnover caused by imperfect data acquisition), and which factors contribute to observation bias and pseudo-turnover. Independent simultaneous surveys provide a powerful tool to quantify pseudo-turnover and to indentify factors causing it, which may vary strongly between lowland and mountain areas.

Location: Alpine mountain summits (2616 m to 3418 m a.s.l.) in the south-eastern Swiss Alps.

Methods: Plant inventories of 48 summits were collected by two independent observers simultaneously. Pseudo-turnover between observers was compared to species turnover over one century based on historical species lists of the same summits. Variables linked to observer characteristics and external (observer-independent) factors were tested for their influence on pseudo-turnover and number of species missed by one of the observers, and plant characteristics were tested for their effect on species detection probability.

Results: Mean pseudo-turnover between observers (13.6%) was almost three times smaller than species turnover over one century (41.4%). Pseudo-turnover and the number of species missed increased with difference in botanizing time between observers and with a longer ascent to the summit, especially in combination with a high species richness on the summit. Species had a higher probability to be missed if occurring on many summits but with a low abundance, if small in stature and if belonging to certain taxonomic plant groups (e.g. Asteraceae).

Conclusions: Our critical evaluation of turnover over time vs pseudo-turnover confirms that floristic changes on alpine summits over time represent an ecological pattern. In mountainous terrain, factors related to observer characteristics play a major role, as we found the best correspondence between simultaneous records when the difference in botanizing time was small and the ascent was short. Our results help to improve data quality in mountainous terrain by pointing out possible causes for observation bias. Long-term vegetation studies in alpine ecosystems should make a strong effort to identify and minimize such causes in advance, for instance by reducing between-observer differences in botanical skills, fitness and time management through appropriate training.

Introduction

Ecosystems have been changing rapidly in recent decades due to climate and land-use changes. Revisitation studies are an effective tool to detect, monitor and analyse changes in vegetation composition (Grabherr et al. 1994; Klanderud & Birks 2003; Wipf et al. 2013). However, studies based on fieldwork are vulnerable to undesired influences affecting data quality. Variability is introduced by imperfect species detection, which may be influenced by traits of the species, environmental conditions or observer characteristics. In order to detect

Doi: 10.1111/jvs.12211

vegetation changes over time, it is crucial to determine and understand variability between records.

Already Smith (1944) stated that variability between observers had an influence on record quality, with high inter- and intra-daily variations for individual observers. MacArthur & Wilson (1967) defined the measure for changes in species composition of island communities as 'species turnover'. With the realization that part of this turnover can be attributed to observer errors, the term pseudo-turnover was then introduced (Lynch & Johnson 1974). In early studies (Lynch & Johnson 1974; Simberloff 1976; Nilsson & Nilsson 1982) pseudo-turnover seems to have been the major fraction of the calculated turnover, whereas in newer studies, pseudo-turnover usually ranges from 5% to 20% (review by Vittoz et al. 2010).

With the ongoing efforts to quantify and reduce pseudo-turnover, several studies tried to find out which factors are responsible for it. *Experience* was considered one of the most important factors potentially influencing record quality. However, some studies show that experience does not necessarily lead to a lower observation bias since even experienced botanists do overlook species, but the influence of experience on record quality remains unclear (Kirby et al. 1986; Vittoz & Guisan 2007; Milberg et al. 2008; Vittoz et al. 2010). There is more agreement on the importance of *botanizing time*, with higher detection rates when more time is invested into a record (Archaux et al. 2006; Kéry et al. 2006; Chen et al. 2009). *Seasonality* affected record quality in a range of studies, leading to decreased species detectability and consistency between records at the beginning and at the end of the field season (Kirby et al. 1986; Rich & Woodruff 1992; Kéry & Schmidt 2008). Also, *species detectability* is decreased in dense vegetation (Shefferson et al. 2001; Vittoz & Guisan 2007) and for species with lower abundance (Lepš & Hadincová 1992; Kercher et al. 2003; Vittoz & Guisan 2007). Other difficulties of species detectability are related to species taxonomy, morphology, life form or size (Scott & Hallam 2003; Kéry et al. 2006; Chen et al. 2009). *Plot area* also affected record quality in several studies, but results between studies are contradictory or lead to no conclusion (Nilsson & Nilsson 1985; Klimeš et al. 2001; Milberg et al. 2008; Chen et al. 2009). Only one study considered *difficulty of terrain*, which was found to have a negative influence on record quality (Rich & Woodruff 1992).

While some of the above-mentioned variables obviously influence data acquisition in the field, the direction and magnitude of their influence differs between studies, possibly due to the wide variety of study designs, vegetation types and terrain in which the studies were conducted. For instance, factors with minor influence in lowland situations may have a pronounced effect in alpine environments, where remoteness, exposed terrain and mountain climate may pose particular challenges to observers. To date, only a few studies have examined observation bias in alpine field studies (Vittoz & Guisan 2007; Vittoz et al. 2010), and no special attention has been paid to which factors related to this particular environment could have a special influence on data quality.

The Summit Flora Project (Stöckli et al. 2011) conducted a large revisitation study based on historical data in an alpine environment. By repeating historical vegetation records of summits that were visited by botanists one century ago, long-term changes in plant community composition, species richness and species distribution can be traced. For the Summit Flora Project as well as for other field studies, it is important to know how much of the calculated species turnover can be attributed to observation bias and which causing factors have to be considered carefully for future study designs to minimize variability in field records.

The objectives of this study are to scrutinize the results of the Summit Flora Project by (i) quantifying between-observer variability in simultaneous field records (so-called pseudo-turnover), (ii) comparing pseudo-turnover with species-turnover over time, and (iii) analysing the main factors causing variability between simultaneous field records. For this, vascular plant compositions of the uppermost 10 m of 48 summits were recorded simultaneously by two independent botanists. These records were then analysed for differences between observers and compared to the historical records of these summits. We then used general and generalized linear models to evaluate how variables related to attributes of the summit (e.g. summit area), observer characteristics (e.g. experience), botanizing circumstances (e.g. weather) and plant traits (e.g. size; Table 1) affected pseudo-turnover, the number of species missed per record and species detection probability. We hypothesized that in this complex alpine environment, observation bias would be high compared to other studies due to more difficult working conditions, but low compared to species turnover over time, which was found to be large on summits (summarized in Stöckli et al. 2011). Moreover, we hypothesized that variability in summit area, time constraints and differences in observer experience would have strong effects on observation bias and number of species missed, and that small species stature, low abundance and a short flowering period would reduce species detectability.

Methods

Study sites

In the field seasons of 2010 and 2011, the Summit Flora Project re-surveyed 120 summits in the southeastern Swiss Alps (Fig. 1), in which there were ca. 100 yr-old historical

Journal of Vegetation Science
Doi: 10.1111/jvs.12211

Table 1. Explanatory variables analysed in this study for their potential influence on observation bias (a,b) and species detectability (c). Factors of group (a) potentially vary strongly between individual observers, or in the effect they have on them. Factors of group (b) have no direct effect on observers, or influence individual observers in a similar way.

(a) Factors Directly or Differentially Affecting Observers		(b) Factors Not or Similarly Affecting Observers	(c) Species-Dependent Factors
Observer-Dependent	Mountain-Dependent		
Botanizing time per observer	Species number on summit	Botanizing time per summit	Species frequency
Experience	Ascent length	Area of summit	Average abundance
	Dangerousness of terrain	Habitat diversity	Minimum & average size
	Density of vegetation	Weather	Flowering period
		Date of record	Start & end of flowering
		Year of record	Taxonomic group
		Mountain altitude	Life form
			Functional group
			Habitat preferences

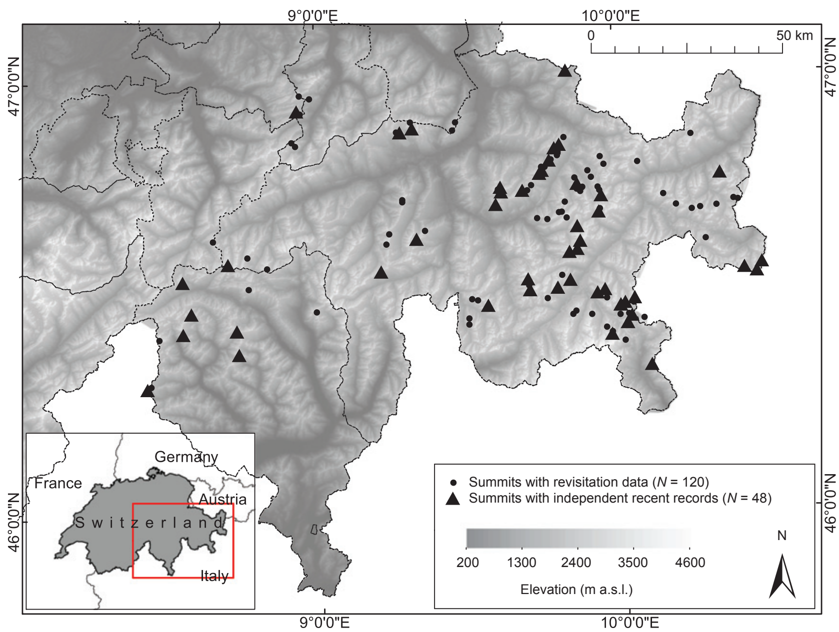

Fig. 1. Overview of study area and location of 120 summits sampled for their floristic composition in the Summit Flora re-survey study. On 48 of these summits, two independent, simultaneous species records were done and analysed for observation bias.

inventories of vascular plant species. The records comprised the lower alpine vegetation belt with its dwarf shrub-sized plants and closed vegetation cover, to the nival belt with its low-stature plants and patchy plant cover.

To quantify the difference between simultaneous records, independent records were performed on 48 summits, according to the protocol described below, by two botanists that visited the summit at the same time, but did not interact. The subset of summits was approximately representative of the 120 summits in terms of distribution in altitude (2616 to 3418 m a.s.l.), region (Fig. 1), bedrock type (70% siliceous and 30% calcareous) and measured species turnover.

Vegetation record protocol

For each of the 120 summits, a list of vascular plant species was assembled for the uppermost ten altitudinal meters without prior knowledge of the historical species list. The lower limit of the 10-m area was determined with barometric altimeters calibrated at the top of the summit (resolution ± 1 m; Garmin GPS, Olathe, US) and marked with cairns. Along a searching route starting at the top and leading in circles outward around the summit, each plant species was recorded with its highest occurrence. An inventory did not necessarily cover each spot within the 10-m area, but potential habitats for additional species

Doi: 10.1111/jvs.12211

were checked with particular care, as described in the historical literature we refer to (Braun 1913; Braun-Blanquet 1958). For safety reasons, areas that we judged dangerous, but were likely to have been accessed by historical botanists, were either botanized secured with a rope or searched with binoculars. Areas that were inaccessible or were likely not to have been accessed by earlier botanists were ignored. All plant species were recorded using a digital field book (custom-made program 'mIraculix' on a HTC® Sense Smartphone (Taoyuan, Taiwan); C. Suter, V. Stöckli & M. Gerber, unpublished resource), along with coordinates, time stamp (built-in GPS unit and watch) and altitude (estimated to the nearest meter). After the record, the abundance of each plant species was estimated on an ordinal scale with five categories (1 = 1 individual, 2 = 2–3 ind., 3 = 4–10 ind., 4 = 11–50 ind., 5 = >50 ind. per summit); 'individual' referring to spatially distinct ramets or clusters of ramets with a size of <1 m^2. In the case of larger patches (e.g. *Carex curvula* patches), each square meter covered was counted as one 'individual'. The protocol did not specify any recording time, thus each botanist decided when the summit was adequately surveyed. All summits were visited on foot from the nearest access point.

Fieldwork was performed between the end of June and mid-September in 2010 and 2011 by a team of three permanent researchers and eight student field assistants (two of them only taking part in 2010, three only in 2011 and three in both years). Field assistants went through an initial training phase of 2–4 wk until they proved capable of recording species lists comparable to those of more experienced team members. Records assembled during the training phase, incomplete records or records stated as unsatisfactory by the observer due to very bad weather (e.g. snow on the ground, thunderstorms) or health problems (i.e. headaches) were excluded from analyses, resulting in the 48 summits with simultaneous records to be analysed. Team composition changed almost every day. The species lists were corrected for obvious misidentifications (i.e. if out of two very similar species, the more unlikely species was found on a certain bedrock type by one observer, the more likely species by the other). Analyses were run both with corrected and uncorrected data, with negligible influences on the outcome and the same conclusions, thus we report results from the corrected data set here.

Potential factors influencing observation bias and species detectability

Of the multitude of factors that could contribute to observation bias in our study, the following variables were quantified (Table 1): time spent botanizing by an observer on a summit was defined as the time elapsed between the first and last species record, and difference in botanizing time between two observers calculated from this. Botanizing time per summit was calculated as the time when one or both of the observers started the record until the last one was finished. Observers were classified into three experience categories based on personal judgment of the authors. For each summit, total species number was calculated from the combined species lists of both observers. As a measure of strenuousness of ascent, ascent length was determined as the total altitudinal meters hiked before reaching the summit. Dangerousness of terrain for each visited summit was estimated in three classes (easily accessible terrain, terrain needing increased concentration, and parts of the summit with dangerous access) by each observer, as was the vegetation density (in two categories: 'open' for very scarcely vegetated summits, and 'patchy-to-dense' for summits at least partly covered with continuous vegetation. The area of the summit (uppermost 10 m) was determined in ArcGIS (v 10.0, ESRI, Redlands, CA, US) from a digital surface model with 2-m resolution. Habitat types and diversity (Simpson's index) of a summit were calculated from estimates of percentage cover of different habitat types (solid rock, coarse scree, fine scree, mineral soil vegetated and unvegetated, organic soil vegetated and unvegetated). Weather conditions were described during the record and subsequently categorized into three classes by the authors according to potential impact on the observers (fine weather; intermediate weather with low potential impact, e.g. moderate cold or threat of rain; conditions potentially experienced as stressful, i.e. rain, cold, strong winds, low visibility or threat of a thunderstorm).

Species frequency (as number of summits occupied by a species), species average abundance (mean abundance of a species over all summits where it occurred) and first and last date when a species was recorded were derived from the full data set of the Summit Flora Project (120 summits). Minimum and average plant size, beginning and end of flowering, length of flowering period, taxonomic group (family or order), functional group (graminoids, forbs, dwarf shrubs, cushion plants, ferns), Raunkiær life form and habitat preferences (wet, tall forb, grassy, rocky, woody vegetation) for each species were derived from Flora Indicativa (Landolt 2010).

Statistical analyses

Species turnover (T) between historical and recent records of a summit served as a measure for vegetation change over time. The same metric was used to calculate turnover between two records made at the same site *and* the same time, so-called pseudo-turnover (Nilsson & Nilsson 1985), as $T = 100*(A + B)/(S_A + S_B)$, where A and B are the number of species exclusive to each of two records, and S_A and

Journal of Vegetation Science
Doi: 10.1111/jvs.12211

S_B are the total number of species in each record. We then compared species turnover over time (between historical and recent records) and pseudo-turnover between two observers with paired Student's *t*-test for 43 summits, since for five summits, historical records were not available or incomplete. While pseudo-turnover is a relative measure of observation bias including species richness in its calculation, we also calculated the number of species exclusively found by one of the observers ($A + B$ in the above formula) as an absolute measure of observation bias. Moreover, we calculated the probability that a species was missed as the proportion of missed occurrences at the number of total occurrences, for a total of 252 species. To determine which factors influence observation bias and probability to be missed, these dependent variables were tested for how they were affected by observer-dependent and observer-independent variables, or species-related variables, respectively (Table 1). Due to the different structures of these variables, we used different statistical methods for the analysis: pseudo-turnover was analysed with a linear model (lm), and number of exclusive species per summit with a generalized linear model (GLM) for Poisson distributed data with a log link, with observer and summit characteristics as explanatory variables (models M1 and M2, $N = 48$; Table 2). Probability to miss a species was analysed using a GLM for proportions with a logit link and a binomial error function with plant characteristics as explanatory variables (R-package 'stats'; model M3, $N = 252$; Table 2). Variables in all models were selected using step-wise forward selection. Multi-collinearities in the models were previously identified and eliminated using a Variance Inflation Factor-threshold of 3 (R-package AED; Zuur et al. 2009). Standardized residuals were visually checked for normality, independence and constant variance. All statistical analyses were performed in R (v 2.14.1; R Foundation for Statistical Computing, Vienna, Austria).

Results

Species turnover vs pseudo-turnover

For the 48 summits analysed, pseudo-turnover between two contemporary botanists ranged between 0% and 33.3%, with a mean of 13.5 ± 1.14 (±SE). Species turnover between historical and recent records was significantly higher than pseudo-turnover, with mean turnover over time exceeding pseudo-turnover between observers almost three-fold (41.4 ± 2.35; t paired = −11.0, $df = 42$, $P < 0.001$; Fig. 2).

The mean species richness on the 48 summits analysed was 38.4 ± 3.63 (±SE) species, with a range from three to 96 species per summit. An average of 4.6 ± 4.1 species per summit (range: 0–20 species) were missed by one observer but found by the other; i.e. 9.1 ± 7.3 species per summit were found by only one of the two observers (range: 0–31 species).

Influence of observer and mountain characteristics on observation bias

The major cause of a high pseudo-turnover was a large difference in botanizing time between observers (model M1; Fig. 3a, Table 2). This was confirmed in model M2, which showed that more species were missed by one of the observers when the difference in botanizing time was high (Fig. 3d). On summits with a higher species richness, a record generally took longer to complete (linear regression, $F_{1,46} = 112.4$, $P < 0.001$) and recording time differed more between observers (linear regression, $F_{1,46} = 7.42$, $P < 0.001$).

A long ascent to reach the summit significantly increased the number of species missed by one of the observers (model M2; Fig. 3e, Table 2) and marginally increased pseudo-turnover (model M1; Fig. 3b). Against expectation, a long ascent did not seem to have led to time constraints, since there was no significant relationship

Table 2. Summary of the model results, testing for the influence of different observer-, summit-, or species-related variables on pseudo-turnover (M1), number of species per summit found by only one observer (M2) and species detectability (M3) using forward selection. Significant and marginally significant (in brackets) explanatory variables after a step-wise model selection procedure are shown. In model M3, *z*-value instead of *F*-value was calculated.

	Model Type	Response Variable	Sample Size	Significant Explanatory Variables	*F* Value	*P* Value
M1	Linear model	Pseudo-turnover	48	Difference in botanizing time	8.71	0.005
				(Ascent length)	3.95	(0.053)
M2	Poisson generalized linear model	N species per summit only found by one observer	48	Species number on summit	10.80	<0.001
				Difference in botanizing time	3.79	<0.001
				Ascent length	3.05	<0.001
M3	Binomial generalized linear model	Probability for a species to be missed	252	Frequency	−6.17	<0.001
				Average abundance	−4.65	<0.001
				(Minimum size of plant)	−1.67	(0.082)
				Taxonomic group	1.7–2.3	<0.001

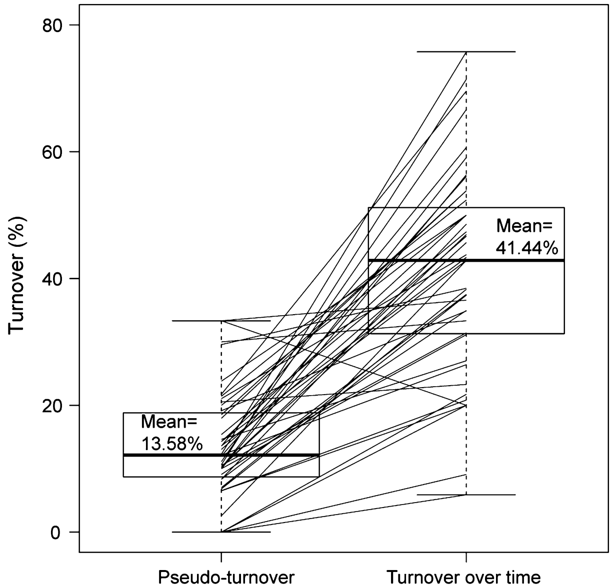

Fig. 2. Box plots comparing pseudo-turnover between floristic summit records of two independent simultaneous observers in 2010/2011 (left) with species turnover over time (between historical and recent summit records; right). Each line connects pseudo-turnover and species turnover over time on the same summit.

between ascent length and botanizing time (linear regression, $F_{1,46} = 0.19$, $P = 0.66$) or difference in botanizing time per record ($F_{1,46} = 0.56$, $P = 0.46$).

More species were missed on summits with higher species richness (model M2; Fig. 3c, Table 2), when the difference in botanizing time between two observers was high, and when the ascent was long (Table 2, Fig. 3d,e).

Influence of plant characteristics on probability that a species was missed

A species' detectability was mainly influenced by its frequency and abundance: species occurring on many mountains but with low abundance had a higher probability to be missed by an observer (model M3). The taxonomic group a species belongs to had a highly significant effect on species detectability (model M3); species were more often missed if they belong to Liliaceae/Orchidaceae ($z = 2.3$, $P = 0.021$), Asteraceae ($z = 1.89$, $P = 0.058$), Scrophulariaceae (genus *Veronica*; $z = 1.76$, $P = 0.078$) or Juncaceae ($z = 1.65$, $P = 0.098$). Finally, species with a small minimum size had a marginally higher chance of being missed (see Table 2).

Discussion

Species turnover vs pseudo-turnover

Since its introduction by Lynch & Johnson (1974), pseudo-turnover has been used as a relative measure for the quality of field observations. As such, it remains an approximation, as the true number of species, and thus the true bias introduced by each observer, remains unknown. Studies conducted under a wide variety of conditions reported pseudo-turnover values between 5% and 20% for vascular plants, with an average of 14.5% (calculated from: Nilsson & Nilsson 1985; Lepš & Hadincová 1992; Kercher et al. 2003; Scott & Hallam 2003; Archaux et al. 2006; Vittoz & Guisan 2007; Vittoz et al. 2010). Due to factors such as large and variable plot area, complex alpine terrain, potentially harsh environmental characteristics and large variation in team composition, we would have expected to find relatively high pseudo-turnover in our study; contrary to our hypothesis, however, pseudo-turnover in our study (13.6%) was well within the range of that found in the above-mentioned studies. As those were performed mainly in lowland and grassland sites, dense vegetation and high species numbers might have counterbalanced factors potentially raising observation bias at alpine sites.

Our results showed that species turnover between historical and recent records far exceeded pseudo-turnover between observers. Although we have no data to assess observer effort and bias in the historical surveys, the historical botanists we refer to were, at their time, highly esteemed for their expertise. They invested years into their research, with their stated goal of assembling exhaustive species lists during their relevés, some even resurveying their own records (Stöckli et al. 2011). Thus, their records were at least of equal quality to ours, and we confidently assume that the fraction of turnover between historical and recent records that can be attributed to variability between observers or recording protocol was within the same range as pseudo-turnover between recent surveys. Our measured species turnover over time can thus be accepted as corresponding mainly to an ecological rather than a methodological signal.

Influence of observer and mountain characteristics on observation bias

Of the multitude of variables that we tested for their potential influence on observation bias (Table 1), three stood out as having a significant effect: difference in time invested into the record, length of ascent to reach the record site and species richness on a summit. We considered all of these variables as depending on the observer or affecting individual observers differently – in contrast, none of the variables that we judged to be independent of observers or affecting both observers in the same way, had a detectable influence on record quality.

The time invested into a record has been considered a prevalent influence on observation bias in several studies.

Journal of Vegetation Science
Doi: 10.1111/jvs.12211

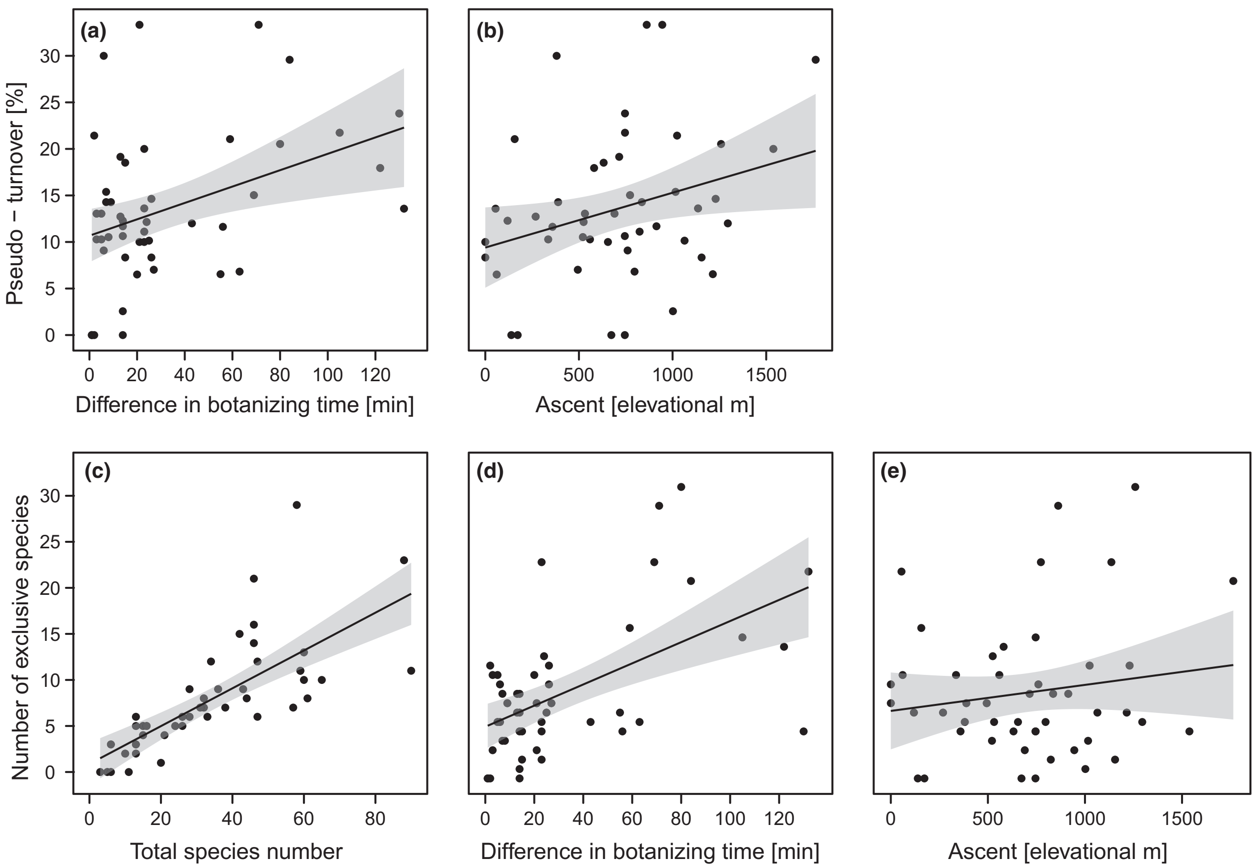

Fig. 3. Relationships between observation bias (*y*-axis) and explanatory variables (*x*-axis) with significant or marginally significant effects in general linear models (see Table 2). A high difference of botanizing time between the observers **(a)** and a long ascent (elevation gain in m to reach a summit) **(b)** increased pseudo-turnover. Number of exclusive species (number of species per summit missed by one of the observers) increased with species richness on the summit **(c)**, difference in botanizing time between the observers **(d)** and a long ascent **(e)**. Plots show linear regressions (solid line) and 95% confidence intervals (grey).

Detection probability generally increases with survey effort (Kéry et al. 2006; Chen et al. 2009) and, crucial for many biodiversity studies, rare species are generally found towards the end of a record (Archaux et al. 2006). It is thus remarkable that in our study, the absolute duration of the record had no effect on observation bias, but the difference in botanizing time between the observers did. The absolute duration to find a certain proportion of the species present on a summit is likely dependent on characteristics of the respective summit, for instance species number, summit area, topography or microhabitat distribution. Because of these differences between summits, botanizing time was not standardized nor prescribed in the record protocol. Therefore, the resulting difference in botanizing time between the two observers depended on the individual judgment of the observers working on the mountain, ergo their botanizing speed, searching method and perception of when the site was sufficiently covered.

The finding that the species richness of a summit had an effect on record quality can be explained in two ways. In our experience, species-rich sites pose a greater challenge to botanists than species-poor sites. Consequently, the results of two observers resolving a difficult task are more variable than when resolving an easy one. An alternative explanation could be that there is simply a higher likelihood to overlook species if more of them are present. In our study, species richness only influenced the number of species missed by one of the observers, but not pseudo-turnover. Thus, the number of species missed is more or less proportional to the total species number on the summit; i.e. the percentage of species missed was not increased in particularly species-rich areas.

Our study is, to our knowledge, the first to find a connection between physical effort of the fieldwork and the quality of resulting data, as we show that a longer ascent (in terms of elevation meters to reach a summit) led to a higher observation bias. Since ascent length was unrelated to the altitude of a mountain, species richness or time spent on a record (all correlations $P > 0.1$), any artifacts related to species richness or time constraints could be excluded.

Generally, physical exhaustion after a long ascent might impair a botanist's record. Whether an ascent is experienced as exhausting or not, however, is a matter of individual fitness. Additionally, it is likely that exhaustion interacts with other observer-related traits such as experience, as experienced observers may cope better with working when tired than less trained observers. Thus, it is likely that any two observers in a team will be affected differently by physical strain, resulting in a larger difference between records after more strenuous ascents. It is thus important that observers in alpine terrain not only train their botanical skills but also their physical fitness.

Contrary to our expectations, we did not find any effect of observer independent factors, such as plot area on observation bias, although several previous studies had found such influences (Nilsson & Nilsson 1985; Klimeš et al. 2001; Klimeš 2003; Milberg et al. 2008; Chen et al. 2009). Most of these studies worked with smaller, well-defined plots (usually 1–400 m^2), but others, such as Nilsson & Nilsson (1985) worked at a similar scale as we did (637–16720 m^2 for the 48 summits of this study). The importance of plot area might thus be higher in denser and more complex lowland vegetation than in the comparably open alpine vegetation. Neither did we find any effects of seasonality, contrary to other studies (Kirby et al. 1986; Rich & Woodruff 1992; McCune et al. 1997; Scott & Hallam 2003; Archaux et al. 2006; Kéry & Schmidt 2008). Its importance may be less pronounced at higher altitudes, most probably due to the absence of therophytes (all species are potentially visible during the whole season) and to the fact that our fieldwork generally started relatively late in the summer (July in our case) and was concentrated over a short vegetation period of maximum 10 wk.

In contrast to our hypothesis, a botanist's experience did not influence observation bias in this study, and neither did other observer-perceived factors such as dangerousness of terrain and density of vegetation. As these variables were classified on subjective assessments by the observers or authors themselves, it is possible that their explanatory power is too low or the assessments too variable to evaluate their influence on data quality. Such subjective classifications could be improved by a more standardized ranking of observer assessments and abilities, for instance by comparing subjective classifications made by all observers on the same summit. Nonetheless, because observer performance changes over time and shows inter- and intra-daily variation (Smith 1944), and because increased observer experience does not necessarily exclude bias (since even experienced botanists do overlook species), the effect of observer experience on record quality remains difficult to quantify (McCune et al. 1997; Scott & Hallam 2003; Milberg et al. 2008; Chen et al. 2009; Vittoz et al. 2010). In our project, all observers were trained in joint excursions at the start of the field season until their performance was judged satisfactory by more experienced peers. The lack of an observer effect related to experience in our study may thus also underline that time investment into rigorous training, even in the short alpine field season, very likely helps to reduce variability in the field data.

Influence of plant characteristics on species detectability

Frequency and abundance of a species, its taxonomic plant group and size influenced detectability of plant species on our alpine summits. In agreement with other studies (Lepš & Hadincová 1992; Rich & Woodruff 1992; Kercher et al. 2003; Archaux et al. 2006; Vittoz & Guisan 2007; Milberg et al. 2008) is the finding that species with a low average abundance and a high frequency (occurring on many summits) are missed more often. It is likely that a frequent species is more often missed by one of the observers. As our summit records covered a relatively large area and our protocol did not prescribe covering each spot, it is also likely that species usually occurring at low abundance or as single individuals are more often missed than species that are usually dominant.

Smaller plants can be expected to be overlooked more often, and in fact our data confirm that the minimum size of a species was negatively related to its detectability. Trait information from databases usually spans values of a variety of populations of a species, but not necessarily the full range of plasticity of a species. We observed that at our sites, where several species are near their upper distribution limit, plant individuals are often particularly small in stature and non-flowering. The latter cannot be represented adequately in trait databases and was not systematically recorded in our data, although individuals are more conspicuous when flowering than in vegetative state.

That some taxonomic plant groups were particularly often overlooked may be related to the reasons stated above. Orchidaceae, Liliaceae and Scrophulariaceae (mainly the genus *Veronica*), for instance, are small in stature and often occur at low abundance. While Asteraceae are a notoriously difficult group for species determination, especially when not flowering, they were also generally frequent (Matteodo et al. 2013), increasing their chance to be missed. The same may be true for Juncaceae (genus *Luzula* and *Juncus*), which are also difficult to determine when not flowering. Although grass species of the genera *Festuca* and *Agrostis* were among those overlooked or misidentified most often, the Poaceae family as a whole was not significantly often missed, likely due to the fact that the very frequent *Poa laxa* and *P. alpina* were usually found by both observers.

Although our study identified a few general traits with an effect on species detectability, predictions of which species are particularly prone to being overlooked and should receive special attention remain difficult, because traits such as frequency and mean abundance are not known during the planning phase of a study. Expert knowledge of the studied vegetation type and species is essential in order to identify potentially difficult and rare taxa before the field season and to provide targeted training or determination keys for taxa that are difficult to find and identify.

Conclusions

The large data set of the Summit Flora Project and a representative subset of summits with two independent records enabled us to calculate the proportion of turnover that could be attributed to inter-observer variability. Moreover, we identified factors with major influence on observation bias out of our *a priori* set of variables potentially important in the alpine environment, several of which have never been considered in earlier studies. The influence of difference in botanizing time and, unexpectedly, strenuousness of ascent, stress the importance of observer-dependent factors in challenging working environments such as alpine summits.

The following recommendations should help to maximize the quality of field records and minimize the variability between observers. Issues related to botanizing time need to be carefully considered in the study design. First, field records should not be assembled under time pressure. Second, in many cases no standardized time frame for the record can be pre-defined, for instance due to varying survey area, structural complexity, vegetation types or working pace of observers. Instead, refined criteria on when to conclude a record could help to reduce inconsistencies between observers.

Rigorous training before sending observers out to the field should be a key component of each field project. Such training should lead to an in-depth understanding of particular recording techniques and protocols, but also improve species identification skills (particularly of difficult taxa) and, as was surprisingly demonstrated in our study, even include physical fitness. When working with several observers split up into small teams, frequent changes in team composition or regular common field days will counteract the development of observer-specific habits and help in equalizing record procedures and observer skills.

Although this study gives new insights into causes of observation bias in alpine summit relevés, it is likely that factors influencing record quality vary depending on the particular environment and circumstances of a study, and thus need to be considered anew for each study. However, simultaneous independent records should be a standard component of monitoring and revisitation studies to determine observation bias, analyse its causes and thus ensure high data quality.

Acknowledgements

We thank the Summit Flora Team A. Björkén, T. Gassner, S. Giovanettina, K. Herz, M. Mattéodo, C. Nilsson, P. Roux-Fouillet and a number of volunteers for help in the field; M. Dawes, E. Zenklusen-Mutter and J. Wheeler for statistical support; S. Güsewell and M. Baltisberger for helpful discussions; A. Mathis for language editing; and P. Kammer and two anonymous reviewers for thoughtful comments that helped to improve the manuscript. The Summit Flora Project is carried out on behalf of the Federal Office for the Environment (FOEN, Switzerland) and is supported by the VELUX Foundation. S.B. was additionally supported by the 'alpine flower fund' of the Swiss Botanical Society (SCNAT, Switzerland).

References

Archaux, F., Gosselin, F., Bergès, L. & Chevalier, R. 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* 17: 299–306.

Braun, J. 1913. Die Vegetationsverhältnisse der Schneestufe in den Rhätisch-Lepontischen Alpen. Ein Bild des Pflanzenleben an seinen äusseren Grenzen. *Neue Denkschriften der schweizerischen Naturforschenden Gesellschaft* 48: 1–333.

Braun-Blanquet, J. 1958. Über die obersten Grenzen Pflanzlichen Lebens im Gipfelbereich des Schweizerischen National parks. *Kommission der Schweizerischen Naturforschende Gesellschaf zur wissenschaftlichen Erforschung des Nationalparks* 6: 119–142.

Chen, G.K., Kery, M., Zhang, J.L. & Ma, K.P. 2009. Factors affecting detection probability in plant distribution studies. *Journal of Ecology* 97: 1383–1389.

Grabherr, G., Gottfried, M. & Pauli, H. 1994. Climate effects on mountain plants. *Nature* 369: 448.

Kercher, S.M., Frieswyk, C.B. & Zedler, J.B. 2003. Effects of sampling teams and estimation methods on the assessment of plant cover. *Journal of Vegetation Science* 14: 899–906.

Kéry, M. & Schmidt, B. 2008. Imperfect detection and its consequences for monitoring for conservation. *Community Ecology* 9: 207–216.

Kéry, M., Spillmann, J.H., Truong, C. & Holderegger, R. 2006. How biased are estimates of extinction probability in revisitation studies? *Journal of Ecology* 94: 980–986.

Kirby, K.J., Bines, T., Burn, A., Mackintosh, J., Pitkin, P. & Smith, I. 1986. Seasonal and observer differences in vascular plant records from British woodlands. *Journal of Ecology* 74: 123–131.

Klanderud, K. & Birks, H.J.B. 2003. Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. *Holocene* 13: 1–6.

Klimeš, L. 2003. Scale-dependent variation in visual estimates of grassland plant cover. *Journal of Vegetation Science* 14: 815–821.

Klimeš, L., Dancak, M., Hajek, M., Jongepierova, I. & Kucera, T. 2001. Scale-dependent biases in species counts in a grassland. *Journal of Vegetation Science* 12: 699–704.

Landolt, E. 2010. *Flora Indicativa*. Haupt, Bern, CH.

Lauber, K. & Wagner, G. 2009. *Flora Helvetica*, 4th edn. Haupt, Bern, CH.

Lepš, J. & Hadincová, V. 1992. How reliable are our vegetation analyses? *Journal of Vegetation Science* 3: 119–124.

Lynch, J.F. & Johnson, N.K. 1974. Turnover and equilibria in insular avifaunas, with special reference to the California Channel Islands. *Condor* 76: 370–384.

MacArthur, R.H. & Wilson, E.O. 1967. *The theory of island biogeography*. Princeton University Press, Princeton, NJ, US.

Matteodo, M., Wipf, S., Stöckli, V., Rixen, C. & Vittoz, P. 2013. Elevation gradient of successful plant traits for colonizing alpine summits under climate change. *Environmental Research Letters* 8: 024043, 10p.

McCune, B., Dey, J.P., Peck, J.E., Cassell, D., Heiman, K., Will-Wolf, S. & Neitlich, P.N. 1997. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. *The Bryologist* 100: 40–46.

Milberg, P., Bergstedt, J., Fridman, J., Odell, G. & Westerberg, L. 2008. Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science* 19: 633–644.

Nilsson, I. & Nilsson, S. 1982. Turnover of vascular plant species on small islands in lake Möckeln, South Sweden 1976–1980. *Oecologia* 53: 128–133.

Nilsson, I.N. & Nilsson, S.G. 1985. Experimental estimates of census efficiency and pseudoturnover on islands: error trend and between-observer variation when recording vascular plants. *Journal of Ecology* 73: 65–70.

Rich, T.C.G. & Woodruff, E.R. 1992. Recording bias in botanical surveys. *Watsonia* 19: 73–95.

Scott, W.A. & Hallam, C.J. 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology* 165: 101–115.

Shefferson, R.P., Sandercock, B.K., Proper, J. & Beissinger, S.R. 2001. Estimating dormancy and survival of a rare herbaceous perennial using mark-recapture models. *Ecology* 82: 145–156.

Simberloff, D. 1976. Species turnover and equilibrium island biogeography. *Science* 194: 572–578.

Smith, A.D. 1944. A study of the reliability of range vegatation estimates. *Ecology* 25: 441–448.

Stöckli, V., Wipf, S., Nilsson, C. & Rixen, C. 2011. Using historical plant surveys to track biodiversity on mountain summits. *Plant Ecology & Diversity* 4: 415–425.

Vittoz, P. & Guisan, A. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. *Journal of Vegetation Science* 18: 413–422.

Vittoz, P., Bayfield, N., Brooker, R., Elston, D.A., Duff, E.I., Theurillat, J.-P. & Guisan, A. 2010. Reproducibility of species lists, visual cover estimates and frequency methods for recording high-mountain vegetation. *Journal of Vegetation Science* 21: 1035–1047.

Wipf, S., Stöckli, V., Herz, K. & Rixen, C. 2013. The oldest monitoring site of the Alps revisited: accelerated increase in plant species richness on Piz Linard summit since 1835. *Plant Ecology & Diversity* 6: 447–455.

Zuur, A.F., Ieno, E., Walker, N.J., Saveliev, A.A. & Smith, G.M. 2009. *Mixed effects models and extensions in ecology with R*. Springer, New York, NY, US.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Significant relationships between species number on the summit and different aspects of botanizing time.

Appendix S2. Relationships between ascent length to reach a summit and variables that were also tested for their influence on observation bias.

Appendix S3. (a,b) Summary of observation bias, mountain characteristics, observer-related variables and external variables for each target summit.

Appendix S4. List of the species most often missed during our observation bias study, with trait variables that were included in the models.

Journal of Vegetation Science
Doi: 10.1111/jvs.12211