

Weighted average regression and environmental calibration as a tool for quantifying climate-driven changes in vegetation

Running title: Vegetation-based environmental reconstruction

Kuber Prasad Bhatta^{1,*}

Harry John Betteley Birks²

John-Arvid Grytnes²

Ole Reidar Vetaas¹

¹Department of Geography, University of Bergen, Postbox 7802, 5020 Bergen, Norway, ²Department of Biological Sciences, University of Bergen, Postbox 7803, 5020 Bergen, Norway

*Correspondence: Kuber Prasad Bhatta, Department of Geography, University of Bergen, Postbox 7802, 5020 Bergen, Norway.

Phone: +47 93425064

E-mail: kuber.bhatta@uib.no

Abstract

Aims

Studies of the climatic responses of plant assemblages via vegetation-based environmental reconstructions by weighted averaging (WA) regression and calibration are a recent development in modern vegetation ecology. However, the performance of this technique for plot-based vegetation datasets has not been rigorously tested. We assess the estimation accuracy of the WA approach by comparing results, mainly the root mean square error of prediction (RMSEP) of WA regressions for six different vegetation datasets (total species, high-frequency species, and low-frequency species as both abundance and incidence) each from two sites.

Methods

Vegetation-inferred environment (plot elevation) calibrated over time is used to quantify the elevational shift in species assemblages. Accuracy of the calibrations is assessed by comparing the linear regression models developed for estimating elevational shifts. The datasets were also used for the backward predictions to check the robustness of the forward predictions.

Important findings

WA regression has a fairly high estimation accuracy, especially with species incidence datasets. However, estimation bias at the extremes of the environmental gradient is evident with all datasets. Out of eight sets (each set with a model for total species, low-frequency species and high-frequency species) of WA regression models, the lowest RMSEPs are produced in the four models based on the total species datasets and in three models based on the high-frequency species only. The inferred environment mirrored the estimation precision of the WA regressions, i.e. precise WA regression models produced more accurate calibrated environmental estimates, which, in turn, resulted in

regression models with a higher adjusted r^2 for estimating the elevational shift in the species assemblages.

Reliable environmental estimates for plot-based datasets can be achieved by WA regression and calibration, although the edge-effect may be evident if species turnover is high along an extensive environmental gradient. Species incidence (0/1) data may improve the estimation accuracy by minimising any potential census and field estimation errors that are more likely to occur in species abundance datasets. Species data processing cannot guarantee the most reliable WA regression models. Instead, generally optimal estimations can be achieved by using all the species with a consistent taxonomy in the training and reconstruction datasets.

Keywords: Calibration, environmental reconstruction, RMSEP, temporal changes, weighted averaging regression

Introduction

The study of the climatic responses of plant assemblages is an important topic in modern vegetation ecology, palaeoecology, biogeography, and conservation science. Various approaches have been applied to study how mountain vegetation responds to recent climate warming. The most common approaches can be grouped into four broad types: changes in species richness or composition on mountain summits (e.g., [Odland *et al.* 2010](#); [Pauli *et al.* 2012](#)); changes in the highest elevational point of observation for a species (e.g., [Dolezal *et al.* 2016](#); [Vittoz *et al.* 2008](#)); changes in the elevational range or optimum of a species along an elevational gradient (e.g., [Felde *et al.* 2012](#); [Lenoir *et al.* 2008](#)); and shifts in the entire plant assemblage along the elevational gradient (e.g., [Bertrand *et al.* 2011](#); [Bhatta *et al.* 2018a](#); [Bhatta and Vetaas 2016](#); [Vetaas 1993](#)). Each of these approaches has its own strengths and limitations in assessing the climatic response of individual species or assemblages. Nonetheless, most have been used in modern vegetation ecology in conjunction with other analytical tools. Their strengths and limitations, and possible ways to resolve or minimise these limitations have been well documented (e.g., [Chytrý *et al.* 2014](#); [Dornelas *et al.* 2013](#); [Gotelli *et al.* 2010](#); [Jackson *et al.* 2012](#); [Kapfer *et al.* 2016](#); [Kopecký and Macek 2015](#); [van der Maarel 2007](#)).

Studies of changes in plant assemblages and their environment by vegetation-based environmental reconstructions have recently been developed in modern vegetation ecology (e.g., [Bertrand *et al.* 2011](#); [Bhatta *et al.* 2018a](#); [Brady *et al.* 2010](#)). Such vegetation-based environmental reconstructions, i.e. environmental estimations based on vegetation composition, mostly make use of the technique of weighted averaging (WA) regression and calibration. The basic assumption of the technique is that taxa in a training set are systematically related (with a unimodal relationship) to the physical environment in which they live ([ter Braak and Prentice 1988](#)), and therefore the environment of a species assemblage in a region (reconstruction set) can be predicted by using the species-environment relationship of the training set. The technique has been used extensively in

palaeoenvironmental reconstructions using species assemblage data from modern sediments and their major environmental factors as the training set, and fossil data of assemblages preserved in sediment cores as the reconstruction set (e.g., [Birks 1994](#); [Birks *et al.* 2010](#) and references therein; [Birks *et al.* 1990](#); [Cao *et al.* 2017](#); [ter Braak and van Dam 1989](#); [Tian *et al.* 2014](#)). In studies where WA regression and calibration are used in modern vegetation ecology, historical vegetation assemblages from known elevations or temperatures are usually used as the training set and the modern assemblages whose apparent elevational or temperature change is being inferred are the reconstruction set. This approach is here called ‘forward predictions’. The role of the modern and historical datasets can be reversed to give ‘backward predictions’ where the training set comprises the modern data, as in paleoecology.

In WA regression and calibration, the species assemblage of the training set is first fitted to the environmental gradient of interest by WA regression, and then the environmental parameter for the reconstruction set is predicted from the regression model by WA calibration ([Birks *et al.* 1990](#)). In modern vegetation studies, the plot-level temperature or elevation can be inferred for the present vegetation, which can then be used to assess the climatic responses of the species assemblages in relation to the historical data in the training set (e.g., [Bertrand *et al.* 2011](#); [Bhatta *et al.* 2018a](#); [Lenoir *et al.* 2013](#)). Difference in the measured and the predicted elevation for the present vegetation gives an estimate of elevational shifts in the species assemblages in response to temperature change. This approach uses the overall species composition of the modern and historical vegetation at a locality. It performs well irrespective of the multicollinearity among abundances of the taxa, of the number of zero values in the dataset which are ignored, of the proportion of the environmental gradient sampled, and of minor spatial inaccuracies in vegetation datasets ([ter Braak and Juggins 1993](#)). Moreover, it assumes a realistic non-linear taxon-environment response, works well with noisy and taxon-rich data, and is also relatively insensitive to outliers. Therefore, it is a theoretically sound, computationally straightforward, and robust analytical technique for environmental reconstructions

([Birks *et al.* 2010](#); [ter Braak and Looman 1986](#)). In modern vegetation ecology, the technique is particularly useful when the plot relocations and sampling intensity are not precise enough during vegetation re-surveys and hence are likely to produce biased estimations of climate-driven compositional changes with other analytical procedures. Moreover, the technique is useful also for testing the robustness of the climatic responses of vegetation quantified by other techniques such as ordinations and linear regressions ([e.g., Bhatta *et al.* 2018a](#))

Performance of WA has been tested in palaeoecology with different sediment core-based training sets by comparing the results (mainly the estimated prediction errors, RMSEP) for the datasets ([e.g., Birks 1994](#); [Birks *et al.* 1990](#); [Salonen *et al.* 2012](#); [Telford and Birks 2009](#); [Telford and Birks 2011b](#); [Tian *et al.* 2014](#)). However, despite being used in modern vegetation ecology for the same purpose, the performance of the technique for different types of plot-based vegetation datasets has not been explored rigorously. Although plot-based vegetation datasets mostly fulfil the basic assumptions of the technique ([as summarised by Birks *et al.* 2010](#)), plot-based data differ in several ways ([e.g., in spatial and temporal scale of study, methods of data collection, nature of collected data](#)) from the sediment core-based data used in palaeoecological studies. These differences in dataset types may lead to differences in the nature and degree of species-environment correspondence that ultimately might influence the estimation accuracy of WA regression and calibration. Therefore, it is important to assess whether the technique produces environmental estimates for plot-based vegetation data as reliably as in palaeoenvironmental reconstructions.

Both species incidence and cover-abundance (hereafter abundance) datasets have been commonly used in vegetation ecology for testing ecological hypotheses ([Jongman *et al.* 1995](#); [Magurran 2004](#)). However, incidence data are thought to be of limited scope in WA regression as the weights given to all taxa are either '0' or '1', whereas species abundance data do not have this limitation. Census or estimation bias, though, is likely to be more common in abundance data, especially if the data have been generated by different observers at different times (see [Burg *et al.* 2015](#); [Kapfer *et al.* 2016](#)).

Therefore, it is important to test whether the estimation accuracy of the technique is influenced by taxon incidences or abundances in the datasets.

Persistence of noise is a common characteristic of community data matrices ([Legendre 1993](#)), which comes from a variety of sources, such as inadequate representation of response variables and their explanatory factors, census and measurement biases, taxonomic inaccuracies, and so on. In addition, sampling of exceptionally rare or particularly abundant taxa or particularly unusual sites may also add noise to the data. This in turn results in an inevitably high unexplained variation in the data ([ter Braak and Looman 1986](#); [ter Braak and Prentice 1988](#)). Therefore, processing of plot-based vegetation data prior to statistical analyses is a common practice in quantitative vegetation ecology, where species with very few occurrences and the sites with very few species are usually regarded as potential outliers. These species or sites are thought to convey little or no useful information or have information that differs from the main trend in the data, and therefore, may contribute to potentially spurious analytical results ([Jongman *et al.* 1995](#); [ter Braak and Prentice 1988](#)). However, it is uncertain whether or not these species are useful or how they may influence environmental predictions for whole plant assemblages using WA regression and calibration. It is important to answer these questions before one makes environmental reconstructions using such datasets because the original dataset together with potential outliers (= adding potential noise to the vegetation dataset) may influence the accuracy of the environmental estimations. Conversely, removal of infrequent species from the dataset may also result in a loss of ecological information represented by such species, leading to a compromise in the estimation accuracy.

In this study, we test the performance of WA regression and calibration for plot-based vegetation datasets with different species abundances and frequencies. Our research question is:

How does the estimation accuracy of WA regression and calibration vary between species abundance and species incidence datasets, and how do species of different frequencies influence the estimation accuracy?

To answer this, we compare WA regression and calibration estimates of species abundance and species incidence in three categories: total species, low-frequency (species with more than 50% frequency are removed), and high-frequency species (species with less than 5% frequency are removed) using vegetation data from two sites in Nepal. Estimation accuracy is assessed mainly by comparing the goodness-of-fit of the models for these different datasets.

Material and methods

Datasets

Species abundance and incidence data from two sites in central Nepal (Site 1, Phulchoki Mountain with temperate forest vegetation; Site 2, Langtang region with subalpine shrub to alpine meadow vegetation) are used in the analyses (Fig.1). Elevation of the sample plots (surrogate for temperature) forms the environmental data. For Site 1, the data are from 1993 and 2013, and for Site 2, they are from 1990 and 2014. The plots in the modern surveys were not the exact relocations of the historical plots, but they were in the same locality as the historical plots, and the historical and modern datasets for each site consist of equal numbers of sample plots over a similar elevational range. An overview of the data used in the analyses is given in Table 1.

For Site 1, we treated the dataset of 1993 as the training set and the species dataset of 2013 as the modern 'reconstruction' set during 'forward prediction', i.e. for the estimation of the environment for the 2013 species composition using the species and environmental variables of 1993. We also performed 'backward prediction', where we estimated the environment for the 1993 datasets using the species and environment of 2013 as the training set. A similar approach was applied to the datasets from Site 2. Backward predictions with the different dataset categories were performed to investigate the robustness of the environmental predictions made by forward predictions.

Species transformation

We transformed the relative abundance scores (0–4) of the species in Site 1 and Braun-Blanquet cover-abundance scores of the species in Site 2 to percentage cover values (Table 1) using the 'bb2num' function of 'simba 0.3-5' package ([Jurasinski and Retzer 2012](#)) in R version 3.4.1 ([R Core Team 2017](#)). Transformation of the factorial species scores to numeric values makes the dataset more suitable for WA regression.

Dataset categories

We divided the abundance datasets from each site into three categories as well as converting each into an incidence dataset. The three categories are 1) all data, 2) low-frequency where all species with more than 50% frequency are removed; and 3) high-frequency where species with less than 5% frequency are removed. Thus, we have three pairs (historical and modern) of datasets of species abundances as well as of species incidences for each of the two sites (= 24 datasets in total). The 24 datasets were further grouped into eight sets of data, each with three categories of data – all species, high-frequency and low-frequency species.

Environmental variable

We aim to estimate elevational shift in species assemblages in response to temporally increased temperature. Therefore, elevation of sample plots (= temperature for species assemblages) represent our main environmental variable. Elevation gradient as such may have minimal direct effect on ecology of species assemblages; however, it is a complex gradient of many important environmental variables such as temperature, rainfall, potential evapotranspiration (hence the number of growing days) and net primary productivity that exhibit very regular trends along elevation gradient of Nepal Himalaya ([Bhattarai and Vetaas 2003](#); [Bhattarai *et al.* 2004](#)). Climate records (from 1970-2000) of the weather stations along elevation gradient from 70 m to 4500 m a.s.l have been collected and published by the Department of Hydrology and Meteorology (DHM), Government of Nepal. Based on the regression analysis of these records, the most influential climatic variable, i.e. annual mean temperature reveals linear declining trend with increasing elevation ($r^2 = 0.98$; $P < 0.001$), with a lapse rate of 0.53 °C per 100m elevation ([Acharya *et al.* 2011](#); [Bhattarai and Vetaas 2003](#)). Temperature is a major gradient that directly affects potential evapotranspiration, net primary productivity and number of growing days (length of growing season) for plant assemblages. Therefore, we treated elevation of plots as a proxy of atmospheric temperature, and regarded the

temporally changed elevation of species assemblages as a response of species under climate warming.

Data analysis

Detrended Correspondence Analysis (DCA)

We performed DCA of the incidence data matrices of all species, where elevation and sampling data were fitted as supplementary variables. Temporal changes in species composition revealed by DCA were used to test the robustness of the temporal patterns inferred by WA regression and calibration. We also used DCA plots to detect any unusual plots before using the datasets in WA regression and calibration. DCA was performed using CANOCO version 5.04 ([ter Braak and Šmilauer 2012](#)).

Weighted Averaging (WA) regression and calibration

WA regression and calibration is a two-step analysis ([Birks *et al.* 1990](#); [ter Braak and van Dam 1989](#)). In the first step, a WA transfer function is derived using known environmental values associated with the species data in the training set using WA regression. In the second step, environmental values for a species assemblage are inferred by using the species-environment relationships of the training set, in WA calibration. The accuracy of such estimations depends on the predictive power of the datasets, i.e. goodness-of-fit of the WA regressions. We used 'rioja' package version 0.9-15 ([Juggins 2016](#)) in R version 3.4.1 for all WA regression and calibration analyses.

WA regression

To evaluate the goodness-of-fit of the WA regressions, cross-validated estimates in WA regression are more reliable because cross-validation provides the basis for producing more robust estimates of the prediction error associated with the environmental reconstruction ([Birks *et al.* 2010](#)). Therefore, we compared the cross-validated estimates (by bootstrapping, bootstrap cycles = 1000) of the WA

regression models for the different training sets. Root mean square error of prediction (RMSEP) in WA regression represents the goodness-of-fit of the model and is commonly used as a measure of the predictive abilities of training sets ([Birks *et al.* 1990](#)). Averages of the environmental parameter being used are taken twice in WA, once in WA regression and once in WA calibration, which results in a shrinkage of the range of inferred environmental values (elevation or temperature). To correct for this, deshinking of the inferred values is done by using different types of regression models that minimise the estimation error, such as edge effects. Environmental estimates were made using 'inverse', 'classical', and 'monotonic curvilinear' deshinking of the WA regressions, with taxa downweighted by their 'tolerances' ([Birks *et al.* 1990](#); [ter Braak and Juggins 1993](#)). Among these, we used the estimate with the lowest RMSEP for each dataset for making comparisons between the different datasets. We also used the WA regression estimate (with the lowest RMSEP) of each dataset for the environmental calibration of the reconstruction sets.

Environmental calibration

We used the species-environment relationships of the training sets in WA regression to predict the environment (elevation of plots) of the modern 'reconstruction' datasets. In most cases, cross-validated WA with a 'monotonic curvilinear regression deshinking' procedure (WA.mon) ([ter Braak and Juggins 1993](#)) reduced the edge effect and produced the lowest RMSEP among the regression models based on different deshinking procedures (supplementary Tables S1, S2). Therefore, we used WA.mon to predict the environment for the reconstruction sets, and used that predicted environment to estimate the temporal change in species composition. Differences in the observed and predicted environment of the modern datasets give an estimate of the temporal change in elevation of the plots, i.e. species assemblages. We also compared the pattern of temporal change revealed by the different dataset categories.

If the environmental calibrations are robust, the trends of elevational shift in species assemblages estimated from forward and backward predictions should complement each other. If forward prediction reveals an upslope shift in the current species assemblages, the backward prediction is expected to reveal a downslope shift for the historical species assemblages or vice versa.

We developed three linear regression models for each calibrated environmental dataset to assess the temporal pattern of elevational shift in the species assemblages:

M0: $\text{change} \sim \text{obs.elev}$ (Null model)

M1: $\text{change} \sim \text{obs.elev} + \text{time}$ (to test whether communities have shifted elevationally over the studied time period)

M2: $\text{change} \sim \text{obs.elev} * \text{time}$ (to test whether the communities have shifted more upwards at higher elevations)

where 'change' = change in elevation of the sample plots, i.e. difference in the observed and predicted elevation of the sample plots, obs.elev = observed elevation of the sample plots, and time = sampling time (year).

Hence, in the case of forward prediction, a positive trend in the elevational change of the plots of the modern dataset will indicate an upslope shift of the species assemblages, whereas a positive trend in the elevational change in backward prediction will indicate a downslope shift of the species assemblages over the time period. If there is a temporal trend in elevational shift in the species assemblages, we would expect the lowest adjusted coefficient of determination (adj. r^2) for the null model (model without the time factor or temporal trend included), and r^2 would increase with the addition of the time factor in the regression models.

Results

Detrended correspondence analysis (DCA)

DCA reveals a clear change in species composition along DCA axis 1 at both sites, irrespective of variation along the spatial/elevation gradient (Fig. 2). This axis corresponds to the temporal gradient. At both sites, the temporal gradient explains most of the variation in species composition (eigenvalue = 0.45 for site 1, 0.55 for site 2; explained variation (%) along axis 1 = 8.20 for site 1 and 4.24 for site 2; gradient length of compositional turnover = 2.60 standard deviations (sd) for site 1 and 5.81 sd for site 2), whereas the spatial gradient (elevation) at both sites is represented by the second DCA axis (eigenvalue = 0.21 for site 1, 0.49 for site 2; explained variation (%) along axis 2 = 3.83 for site 1 and 3.82 for site 2; gradient length = 2.20 sd for site 1, 5.45 sd for site 2).

Weighted averaging (WA) regression and calibration

a) WA regression

Goodness-of-fit of a WA regression model is assessed by the degree of concordance between the estimated and observed environment values for the training set. For all the training sets, there is high concordance ($r^2 > 0.90$ for Site 1; $r^2 > 0.84$ for Site 2) between the estimated and observed environment in the cross-validated models (Fig. 3, 4; supplementary Table S1, S2). For most datasets (all but two), WA regression models based on the monotonic curvilinear deshrinking procedure produced estimates with the lowest RMSEP for the environmental variable. However, model performance varies among the different categories of datasets (total species, low-frequency species, high-frequency species datasets for both abundance and incidence data; Fig. 3, 4; supplementary Table S1, S2).

1. Abundance versus incidence data

Species abundance datasets usually produce a higher RMSEP for the normal WA estimates, but a lower RMSEP for tolerance-downweighted WA estimates compared to the species incidence datasets (Fig. 3; supplementary Table S1, S2). The lowest RMSEP associated with the estimates based on the abundance datasets is always higher than that for the incidence datasets (supplementary Table S1, S2).

2. WA estimation for the datasets – total species vs. low-frequency species vs. high-frequency species

Among the datasets of Site 1, WA of the high-frequency species produces the lowest RMSEP in forward prediction (Fig. 3a; supplementary Table S1a). Furthermore, the lowest RMSEP for the incidence dataset of high-frequency species is significantly lower than for the abundance dataset.

The trend is similar for the incidence datasets in backward WA prediction, but with the abundance based datasets, WA of the dataset with total species produces the lowest RMSEP (Fig. 3b, supplementary Table S1b).

Based on the RMSEP of the WA regression estimates, datasets with the lowest to highest RMSEP (WA estimates with the lowest RMSEP being the most accurate) can be ranked as:

Abundance and incidence datasets in forward WA prediction: high-frequency species < total species < low-frequency species

Incidence datasets in backward WA prediction: high-frequency species < total species < low-frequency species

Abundance datasets in backward WA prediction: total species < low-frequency species < high-frequency species

Among the abundance datasets of Site 2, WA estimation for the dataset with all species produces the lowest RMSEP in forward prediction, whereas among the incidence datasets, estimation for the dataset of low-frequency species produces the lowest RMSEP (Fig. 3c; supplementary Table S2a).

However, the difference in RMSEP for the datasets with total species and low-frequency species is

very small. The lowest RMSEP for the incidence dataset is significantly lower than for the abundance dataset. Trends of WA estimates in the backward predictions are similar to those for the abundance datasets in forward predictions (Fig. 3d; supplementary Table S2b). Based on the RMSEP of the WA regression estimates, the datasets with the lowest to highest RMSEP can be ranked as:

Abundance datasets in forward prediction: total species < low-frequency species < high-frequency species

Incidence datasets in forward prediction: low-frequency species < total species < high-frequency species

Abundance as well as incidence datasets in backward prediction: total species < low-frequency species < high-frequency species

b) WA environmental calibration or reconstruction

1. Abundance versus incidence data

Accuracy of the environmental calibration for the different datasets was assessed by developing linear regression models for the elevational shifts based on the estimated or predicted environment factor.

Among the three contrasting regression models of environmental change for each dataset, the one with an interaction between observed elevation and time is the most significant (Table 2, Supplementary Table S3). However, differences are evident among the datasets regarding the diagnostics of the most significant regression models. The models for the incidence datasets mostly have higher adjusted r^2 than those for the abundance datasets (Table 2).

2. WA estimation for datasets – total species versus low-frequency species versus high-frequency species

At Site 1, significant elevational shifts of species assemblages are revealed by forward prediction for all abundance-based (Fig 5a-5c) as well as incidence-based (Fig. 5d-5f) datasets (Table 2a). For all the datasets, the elevational shift in plant communities is significantly explained by an interaction of the observed elevation (obs.elev) and sampling year (time) (Supplementary Table S3). Among the three frequency categories, the regression model for the dataset with high-frequency species is the most significant ($r^2 = 0.82, 0.85$ for abundance and incidence datasets, respectively; Fig. 5c, 5f; Table 2a3). The regression models from backward prediction are not as good as those from forward prediction (supplementary Fig. S1, Table 2b). The model for the total species dataset is most significant among those of the three frequency categories ($r^2 = 0.71, 0.76$ for abundance and incidence datasets, respectively; Fig. 5a, 5d; Table 2b1).

At Site 2, significant elevational shifts of species assemblages are revealed by calibrations for all abundance-based (Fig 6a-6c) as well as incidence-based (Fig. 6d-6f) datasets. Among the abundance datasets, the regression model for the dataset with high-frequency species is most significant ($r^2 = 0.79$; Table 2c3; Fig. 6c). In contrast, among the models for the incidence datasets, the model for the dataset of total species is most significant ($r^2 = 0.81$; Table 2c1; Fig. 6d). Trends of elevational shifts in assemblages estimated by forward prediction are similar to those estimated in backward prediction (Table 2d, Supplementary Fig. S2). However, the regression models for backward prediction have notably lower adjusted r^2 than those for forward prediction.

Discussion

Overall performance

Weighted Averaging (WA) regression and calibration reconstruct the environment of plot-based vegetation datasets with a fairly high accuracy. The inferred temporal trends in vegetation and environment revealed by the reconstructions are supported by ordination analyses that reveal significant changes in species composition at both sites along the temporal gradient. However, compositional turnover varies significantly between the two sites. This is largely due to the elevational effect on the compositional turnover, difference in the length of elevation gradient in both sites, and variation in census as well as estimation accuracy and sampling technique of the present and the past datasets in the sites.

Species turnover increases with increasing elevation, especially due to increased rate of plot-to-plot species loss with elevation that causes compositional dissimilarity among plots ([Bhatta *et al.* 2018b](#)). Due to this, the high-elevation site (Site 2) revealed higher compositional turnover along elevation gradient than that in the low-elevation site. Moreover, sample plots of site 1 are distributed along elevation gradient of c 500 m (approximately 2200 m a.s.l. to 2700 m a.s.l.) and cover only broad-leaved forest in temperate vegetation zone; whereas the site 2 spans over elevation gradient of c 1200 m (approximately 3800 m a.s.l. to 5000 m a.s.l.) and cover three vegetation types namely subalpine forest, alpine scrub and alpine meadows. This eventually resulted in higher compositional turnover in the site 2 datasets. Species' incidence-based datasets of site 1 were sampled using frequency of subplots (0-4) procedure during both surveys ([Vetaas 1997, cf. below](#)). However, datasets of site 2 were sampled on species' cover-abundance scale using different (subjective sampling for the past dataset and systematic sampling for the present dataset) sampling techniques and by different observers at different times. These factors lead to higher census and taxonomic

biases in the datasets resulting higher compositional dissimilarity between the past and present datasets of site 2 as compared to those of site 1.

Of the 12 pairs of datasets analysed, almost all suffer from 'edge-effects' that lead to an overestimation of optima at the low end of the gradient and an underestimation at the high end (ter Braak and Juggins 1993). This effect is most likely caused by the training sets that exhibited high compositional turnover along a dominant, long spatial gradient, which is a typical situation for marked edge effects (Birks *et al.* 2010; ter Braak and Juggins 1993). As recommended by previous studies (e.g., ter Braak and Juggins 1993), monotonic curvilinear deshrinking of WA regression models reduces, to some extent, the edge-effect and produces more accurate estimations with lower RMSEP than inverse or classical linear deshrinking.

Comparison of the accuracy of WA regression estimates across the different datasets of species abundance and frequency revealed several important trends, some of which are more pronounced and systematic than the other. These differences are also mirrored in the environmental calibrations and consequently, in the estimate of the temporal changes in species assemblages. Therefore, we focus our discussion on the WA regression estimates.

WA regression

Species abundance vs. incidence datasets

WA regression performed notably better with the species incidence datasets than with the species abundance datasets. The abundance scores of each species are standardised to 1 in the incidence datasets, otherwise the abundance and incidence datasets are identical regarding taxonomic diversity and species composition. Therefore, the difference in estimation accuracy is likely associated with field and recording methodological inaccuracies. In vegetation ecology, studies of temporal changes are mostly based on the resampling of previously sampled plots or vegetation or areas, mostly by different observer(s) (Chytrý *et al.* 2014; Kapfer *et al.* 2016). This may lead to

methodological differences between the surveys such as sampling design, sampling effort, census or estimation accuracy, individual observer skills and experience, and distribution of sample plots along an environmental gradient (e.g., [Archaux et al. 2006](#); [Kapfer et al. 2016](#); [Tingley and Beissinger 2009](#)). Consequently, noise may become prominent in the datasets. Census and estimation biases are pronounced in the cover-abundance datasets generated by different observers because it is unlikely that a relative cover-abundance score given to a species by two different observers at the same time is identical. Differences in the observer's experience and expertise may also influence the taxonomic accuracy of the species recorded ([Bhatta and Vetaas 2016](#)). Species abundance itself also fluctuates yearly due to seasonal changes in the weather conditions ([Diekmann 2003](#)). Such errors can contribute to high prediction errors in WA regression. Therefore, use of abundance data is sometimes avoided in the ecological studies that make WA estimations or compare the historical and current species composition ([Diekmann 1995](#); [Diekmann 2003](#)).

In a normal WA regression analysis, an estimate of a taxon's optimum for a particular environmental factor is an average of all the values of that factor for the sites in which the taxon occurs, weighted by the taxon's relative abundance ([ter Braak and van Dam 1989](#)). Therefore, the above-mentioned bias is incorporated into the WA estimates via weights (relative abundance) of the taxa. However, this bias is minimised in the incidence dataset because here the abundances of all taxa recorded from a site are standardised to 1, and therefore only an average of the environmental parameter for the sites in which the species occurs is used to estimate the optimum of the species for that parameter. Incidence datasets have been preferred in many studies for WA calculations because the quantitative response of a species does not only depend on the environmental conditions but also on its specific growth form ([Diekmann 1995 and references therein](#); [Ellenberg 1991](#)). Some species by nature, grow individually whereas the others from extensive populations irrespective of their response to environmental conditions, and hence the more frequent species are weighted comparatively more in WA calculations based on abundance datasets.

WA regressions for Site 1 datasets produced notably lower estimation errors than those for Site 2 because the relative abundance score (0 to 4) in the Site 1 datasets is only based on the absence or presence of a species within four subplots of a main plot ([Vetaas 1997](#)). The datasets are thus essentially incidence matrices and any error in cover-abundance estimation is absent. In contrast, the datasets from Site 2 were recorded using Braun-Blanquet's cover-abundance scale by different observers at different times. Therefore, census and taxonomic biases may be prevalent in the Site 2 datasets.

The 1990 survey of Site 2 was based on selective sampling, whereas that of 2014 used a non-preferential (systematic) sampling procedure. In contrast, the datasets of Site 1 were sampled by applying exactly the same (stratified-random) sampling technique in both surveys. Such differences in the sampling design can produce several important differences in the datasets regarding the representation of frequent and infrequent species in the datasets, sampled species heterogeneity, and species-environment concordance ([Bhatta *et al.* 2012](#); [Diekmann *et al.* 2007](#); [Grabherr *et al.* 2003](#); [Michalcová *et al.* 2011](#)). This may also ultimately lead to differences in the estimation accuracies of WA regression based on the datasets from both sites. An uneven distribution of the sample plots along an elevational gradient may also influence the performance of WA regression models ([Telford and Birks 2011a](#)). Accordingly, the higher number of sample plots in high elevation areas (4700–5000 m a.s.l.) compared to those between 3850 m and 4600 m a.s.l. at site 2 may have contributed to the higher RMSEP for the WA regression models compared to those for the elevationally uniformly distributed data of Site 1.

Datasets with total species vs. low-frequency species vs. high-frequency species

We found some inconsistency in the estimation accuracy of the WA regressions with the datasets processed in different ways. At Site 1, removal of infrequent species from the datasets mostly improved the accuracy of the regression models. In contrast, at Site 2, WA regression with the

unprocessed datasets (all species) performed better than with the reduced datasets. However, there are exceptions in both cases. Out of eight WA estimates (four for each Site: two forward and two backward, one each for abundance and incidence data), only four of the best estimates (lowest RMSEP) were for the training sets with all species. Nonetheless, WA regression with all species never had the worst performance (highest RMSEP) unlike the datasets with low-frequency or high-frequency species. Therefore, the findings of classical studies in palaeoenvironmental reconstructions that inclusion of the maximum number of taxa in the training sets minimises the predictions errors in WA regressions ([Birks 1994](#); [Birks *et al.* 1990](#)) is not fully supported by our analyses.

Removal of the 31 most frequent species from the full training set (175 species) of Site 1 during forward prediction reduced the estimation accuracy (increased RMSEP) compared to that for the total species dataset; whereas estimation accuracy increased (decreased RMSEP) when we removed the least frequent 49 species from the full training set. These findings contrast with those of previous studies that show that large prediction errors occur when only the commonest and numerically most abundant taxa are included in the WA regression ([Birks 1994](#)). In the full training sets, the most frequent taxa were identified with high taxonomic precision, whereas taxonomic inaccuracy (and hence noise) is more likely associated with the least frequent taxa. Removal of the most frequent species from the full training set is therefore likely to cause loss of valuable ecological information, whereas removal of the least frequent species intuitively minimises noise within the dataset.

At Site 2, WA regression with all species produced the most accurate environmental estimates, whereas removal of the infrequent species from the datasets reduced the estimation accuracy. There were 113 species in the 1990 dataset and 121 species in the 2014 dataset with less than 5% frequency. Despite a high taxonomic mismatch between both datasets, mostly due to these species, removal of almost half of the species from the dataset certainly contributed to a loss of environmental information associated with some of the correctly identified species. Moreover, low-

frequency species may sometimes have a narrow environmental tolerance and may therefore be optimal indicator species.

High accuracy in WA regression and environmental calibration requires high taxonomic accuracy, together with high spatial, temporal, and chronological precision ([Birks 1993](#); [Birks 1994](#); [Birks *et al.* 2010](#); [Munro *et al.* 1990](#); [ter Braak and Juggins 1993](#)). The development of modern and training sets with a high taxonomic precision ideally requires that the data be sampled, the taxa identified, and all the analyses be done by the same analyst, and that the analyst be skilled in taxonomy ([Birks 1994](#); [Munro *et al.* 1990](#)). Careful standardisation of taxonomy, censuses, and other methods are alternatives when the data have been collected by different observers or analysed by different analysts ([Birks *et al.* 2010](#)). However, differences probably linked to taxonomic and sampling biases persisted in our datasets despite standardisation prior to analyses. Thus, our analyses indicate that estimation accuracy of the WA regression is mainly influenced by taxonomic mismatches, census biases, and the number of taxa included or eliminated during data processing.

Calibration

As expected, environmental calibrations with different datasets revealed elevational shifts of sample plots (and thus of assemblages) over the studied time period. The ability of the regression models used to estimate the elevational shifts (indicated by adjusted r^2 of the models) mirrors the estimation accuracy of the WA regressions. This means that precise WA regression estimates produce more accurate calibrated environmental parameters, which, in turn, result in a robust regression model with a high adjusted r^2 for estimating elevational shift in plant assemblages.

In our analyses, backward predictions only broadly follow the trends of forward predictions, especially for the Site 2 datasets. One possible reason for this may be that not all the species were recorded in both the surveys (training and modern datasets) due to differences in the sampling effort and potential taxonomic inaccuracies. Another potential reason is that each species responds

individually to a particular environmental factor and therefore there will be changes in assemblages in response to the changed environmental conditions. Because WA regression is sensitive to the distribution of the environmental variable in the training set (ter Braak and Looman 1986), the backward prediction using the changed environment and assemblage as a training set may not properly complement the pattern revealed in forward predictions.

In conclusion, WA regression and environmental calibration with plot-based vegetation data perform with a fairly high accuracy, although edge-effects may be evident if species turnover is high along an extensive and monodominant environmental gradient. Species abundance datasets generated by different observers at different times may be prone to potential taxonomic and census errors that may result in high estimation errors in the models. However, the use of a species incidence matrix may improve the estimation accuracy. Species data processing cannot guarantee the most accurate WA regression estimates: instead it can lead to extreme estimates. However, most optimal estimates and calibrations can be achieved by using the full set of the species in the datasets.

References

- Acharya KP, Vetaas OR, Birks HJB (2011) Orchid species richness along himalayan elevational gradients. *J Biogeogr* **38**:1821–1833.
- Archaux F, Gosselin F, Bergès L, *et al.* (2006) Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *J Veg Sci* **17**:299–306.
- Bertrand R, Lenoir J, Piedallu C, *et al.* (2011) Changes in plant community composition lag behind climate warming in lowland forests. *Nature* **479**:517–520.
- Bhatta KP, Chaudhary RP, Vetaas OR (2012) A comparison of systematic versus stratified-random sampling design for gradient analyses: A case study in subalpine himalaya, nepal. *Phytocoenologia* **42**:191–202.
- Bhatta KP, Grytnes J-A, Vetaas OR (2018a) Downhill shift of alpine plant assemblages under contemporary climate and land-use changes. *Ecosphere* **9**:e02084.
- Bhatta KP, Grytnes J-A, Vetaas OR (2018b) Scale sensitivity of the relationship between alpha and gamma diversity along an alpine elevation gradient in central nepal. *J Biogeogr* **45**:804–814.
- Bhatta KP, Vetaas OR (2016) Does tree canopy closure moderate the effect of climate warming on plant species composition of temperate himalayan oak forest? *J Veg Sci* **27**:948–957.
- Bhattarai KR, Vetaas OR (2003) Variation in plant species richness of different life forms along a subtropical elevation gradient in the himalayas, east nepal. *Global Ecol Biogeogr* **12**:327–340.
- Bhattarai KR, Vetaas OR, Grytnes JA (2004) Fern species richness along a central himalayan elevational gradient, nepal. *J Biogeogr* **31**:389–400.
- Birks HJB (1993) Quaternary palaeoecology and vegetation science – current contributions and possible future developments. *Rev Palaeobot Palynol* **79**:153–177.
- Birks HJB (1994) The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions. *Rev Palaeobot Palynol* **83**:107–117.

- Birks HJB, Heiri O, Seppä H, *et al.* (2010) Strengths and weaknesses of quantitative climate reconstructions based on late-quaternary biological proxies. *The Open Ecology Journal* **3**:68–110.
- Birks HJB, Line JM, Juggins S, *et al.* (1990) Diatoms and pH reconstruction. *Phil Trans Royal Soc London B: Biol Sci* **327**:263–278.
- Brady TJ, Monleon VJ, Gray AN (2010) Calibrating vascular plant abundance for detecting future climate changes in oregon and washington, USA. *Ecol Indicators* **10**:657–667.
- Burg S, Rixen C, Stöckli V, *et al.* (2015) Observation bias and its causes in botanical surveys on high-alpine summits. *J Veg Sci* **26**:191–200.
- Cao X, Tian F, Telford RJ, *et al.* (2017) Impacts of the spatial extent of pollen-climate calibration-set on the absolute values, range and trends of reconstructed holocene precipitation. *Quat Sci Rev* **178**:37–53.
- Chytrý M, Tichý L, Hennekens SM, *et al.* (2014) Assessing vegetation change using vegetation-plot databases: A risky business. *App Veg Sci* **17**:32–41.
- Diekmann M (1995) Use and improvement of ellenberg's indicator values in deciduous forests of the boreo-nemoral zone in sweden. *Ecography* **18**:178-189.
- Diekmann M (2003) Species indicator values as an important tool in applied plant ecology – a review. *Basic Appl Ecol* **4**:493-506.
- Diekmann M, Kühne A, Isermann M (2007) Random vs non-random sampling: Effects on patterns of species abundance, species richness and vegetation-environment relationships. *Folia Geobotanica* **422**:179–190.
- Dolezal J, Dvorsky M, Kopecky M, *et al.* (2016) Vegetation dynamics at the upper elevational limit of vascular plants in himalaya. *Scientific Reports* **6**:24881.
- Dornelas M, Magurran AE, Buckland ST, *et al.* (2013) Quantifying temporal change in biodiversity: Challenges and opportunities. *Proc R Soc Lond B Biol Sci* **280**:20121931.
- Ellenberg H (1991) Zeigerwerte von pflanzen in mitteleuropa. *Scripta Geobotanica* **18**:1-248.

- Felde VA, Kapfer J, Grytnes J-A (2012) Upward shift in elevational plant species ranges in sikkilsdalen, central norway. *Ecography* **35**:922–932.
- Gotelli NJ, Dorazio RM, Ellison AM, *et al.* (2010) Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. *Phil Trans Royal Soc London B: Biol Sci* **365**:3621–3631.
- Grabherr G, Reiter K, Willner W (2003) Towards objectivity in vegetation classification: The example of the austrian forests. *Plant Ecol* **169**:21–34.
- Jackson MM, Turner MG, Pearson SM, *et al.* (2012) Seeing the forest and the trees: Multilevel models reveal both species and community patterns. *Ecosphere* **3**:1–16.
- Jongman RHG, ter Braak CJF, van Tongeren OFR (1995) Data analysis in community and landscape ecology: Cambridge University Press, Cambridge UK.
- Juggins S (2016) Rioja: Analysis of quaternary science data, r package version (0.9-9). (<http://cran.R-project.org/package=rioja>).
- Jurasinski G, Retzer V (2012) A collection of functions for similarity analysis of vegetation data. R package version 2.4-3.
- Kapfer J, Hédal R, Jurasinski G, *et al.* (2016) Resurveying historical vegetation data – opportunities and challenges. *App Veg Sci*:204–214.
- Kopecký M, Macek M (2015) Vegetation resurvey is robust to plot location uncertainty. *Divers Distrib* **21**:322–330.
- Legendre P (1993) Real data are messy. *Statistics and Computing* **3**:197–199.
- Lenoir J, Gégout JC, Marquet PA, *et al.* (2008) A significant upward shift in plant species optimum elevation during the 20th century. *Science* **320**:1768–1771.
- Lenoir J, Graae BJ, Aarrestad PA, *et al.* (2013) Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across northern europe. *Global Change Biol* **19**:1470–1481.

- Magurran AE (2004) Measuring biological diversity: Blackwell Publishing 350 Main Street, Maiden, MA 02148-5020, USA.
- Michalcová D, Lvončík S, Chytrý M, *et al.* (2011) Bias in vegetation databases? A comparison of stratified-random and preferential sampling. *J Veg Sci* **22**:281–291.
- Munro MAR, Kreiser AM, Battarbee RW, *et al.* (1990) Diatom quality control and data handling. *Phil Trans Royal Soc London B: Biol Sci* **327**:257–261.
- Odland A, Høitomt T, Olsen SL (2010) Increasing vascular plant richness on 13 high mountain summits in southern Norway since the early 1970s. *Arct Antarct Alp Res* **42**:458–470.
- Pauli H, Gottfried M, Dullinger S, *et al.* (2012) Recent plant diversity changes on Europe's mountain summits. *Science* **336**:353–355.
- R Core Team (2017) R: A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Salonen JS, Ilvonen L, Seppä H, *et al.* (2012) Comparing different calibration methods (wa/wa-pls regression and Bayesian modelling) and different-sized calibration sets in pollen-based quantitative climate reconstruction. *The Holocene* **22**:413–424.
- Telford RJ, Birks HJB (2009) Evaluation of transfer functions in spatially structured environments. *Quat Sci Rev* **28**:1309–1316.
- Telford RJ, Birks HJB (2011a) Effect of uneven sampling along an environmental gradient on transfer-function performance. *J Paleolimnol* **46**:99.
- Telford RJ, Birks HJB (2011b) A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages. *Quat Sci Rev* **30**:1272–1278.
- ter Braak CJF, Juggins S (1993) Weighted averaging partial least-squares regression (wa-pls) - an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* **269/270**:485–502.

- ter Braak CJF, Looman CWN (1986) Weighted averaging, logistic regression and the gaussian response model. *Vegetatio* **65**:3–11.
- ter Braak CJF, Prentice IC (1988) A theory of gradient analysis. In Begon M, et al. (eds) *Advances in ecological research*, Vol. 18: Academic Press, 271–317.
- ter Braak CJF, Šmilauer P (2012) Canoco reference manual and user's guide: Software for ordination, version 5.0: Microcomputer Power, Ithaca, USA.
- ter Braak CJF, van Dam H (1989) Inferring pH from diatoms: A comparison of old and new calibration methods. *Hydrobiologia* **178**:209–223.
- Tian F, Herzschuh U, Telford RJ, et al. (2014) A modern pollen–climate calibration set from central-western mongolia and its application to a late glacial–holocene record. *J Biogeogr* **41**:1909–1922.
- Tingley MW, Beissinger SR (2009) Detecting range shifts from historical species occurrences: New perspectives on old data. *Trends Ecol Evol* **24**:625–633.
- van der Maarel E (2007) Transformation of cover-abundance values for appropriate numerical treatment - alternatives to the proposals by podani. *J Veg Sci* **18**:767–770.
- Vetaas O (1997) The effect of canopy disturbance on species richness in a central himalayan oak forest. *Plant Ecol* **132**:29–38.
- Vetaas OR (1993) Spatial and temporal vegetation changes along a moisture gradient in northeastern sudan. *Biotropica* **25**:164–175.
- Vittoz P, Bodin J, Ungricht S, et al. (2008) One century of vegetation change on isla persa, a nunatak in the bernina massif in the swiss alps. *J Veg Sci* **19**:671–680.

List of figures, tables and supplementary material

Figures

Fig. 1. Location map of the study area. Site 1 = Phulchoki mountain, Site 2 = Langtang National Park.

Fig. 2. Detrended Correspondence Analysis (DCA) biplot of species composition. (a) Site 1. (b) Site 2.

Fig. 3. Root Mean Squared Error of Prediction (RMSEP) of the weighted averaging (WA) regression models for different training sets from **(a, b) Site 1 and (c, d) Site 2**, where the estimates produced by monotonic curvilinear regression deshrinking procedures were selected.

Fig. 4. Adjusted coefficient of determination (r^2) of the weighted averaging (WA) regression models for different training sets from **(a, b) Site 1 and (c, d) Site 2**, where the estimates produced by monotonic curvilinear regression deshrinking procedures were selected.

Fig 5a-5f. Change in the elevational distribution of the species assemblages from 1993 to 2013 at Site 1, where the difference in the observed and the predicted elevation of the sample plots of 2013 is the temporal change in elevation of the species assemblages (forward prediction

Fig 6a-6f. Change in the elevational distribution of the species assemblages from 1990 to 2014 at Site 2, where the difference in the observed and the predicted elevation of the sample plots of 2014 is the temporal change in elevation of the assemblages (forward prediction).

Tables

Table 1. Overview of the datasets compiled for the analyses.

Table 2. Elevational shift in plant communities over the time period (time) revealed by analyses of different datasets, where elev.change (elevational shift in metres) = predicted – observed elevation (obs.elev), SE = Residual standard error. The linear regression model for each dataset is fitted with

'elev.change' as a response variable and the interaction of 'obs.elev' and 'time' as predictor variables.

Supplementary material

Table S1. Summary statistics of the weighted averaging (WA) regression models for different training sets from **Site 1**, where the estimates are produced by inverse (WA.inv), classical (WA.cla), and monotonic curvilinear (WA.mon) regression deshrinking procedures.

Table S2. Summary statistics of the weighted averaging (WA) regression models for different training sets from Site 2, where the estimates are produced by inverse (WA.inv), classical (WA.cla), and monotonic curvilinear (WA.mon) regression.

Fig S1. Change in the elevational distribution of the species assemblages from 2013 to 1993 at Site 1, where the difference in the observed and the predicted elevation of the sample plots of 1993 is the temporal change in elevation of the species assemblages (backward prediction).

Fig S2. Change in the elevational distribution of the species assemblages from 2014 to 1990 at Site 2, where the difference in the observed and the predicted elevation of the sample plots of 1990 is the temporal change in elevation of the assemblages (backward prediction).

Table 1. Overview of the datasets compiled for the analyses.

Site	Category	1993/1990 data		2013/2014 data		Transformed cover scale	Incidence scale
		Plots	Species	Plots	Species		
1	Total species	63	175	64	182	0-87.5	0/1
1	Without frequent species	63	144	64	156	0-87.5	0/1
1	Without infrequent species	63	126	64	113	0-87.5	0/1
2	Total species	91	215	91	264	0-87.5	0/1
2	Without frequent species	91	211	91	252	0-87.5	0/1
2	Without infrequent species	91	102	91	143	0-87.5	0/1

Table 2. Elevational shift in plant communities over the time period (time) revealed by analyses of different datasets, where elev.change (elevational shift in metres) = predicted – observed elevation (obs.elev), SE = Residual standard error. The linear regression model for each dataset is fitted with ‘elev.change’ as a response variable and the interaction of ‘obs.elev’ and ‘time’ as predictor variables.

Site 1						
(a) 1993 as training set						
Model	Abundance dataset			Incidence dataset		
1. Total species						
	adj. r^2	SE	P	adj. r^2	SE	P
elev.change~obs.elev*time	0.79	37.22	<0.001	0.83	31.51	<0.001
2. Low-frequency species						
elev.change~obs.elev*time	0.69	45.46	< 0.001	0.67	39.20	< 0.001
3. High-frequency species						
elev.change~obs.elev*time	0.82	36.75	< 0.001	0.85	0.84	< 0.001
(b) 2013 as training set						
1. Total species						
elev.change~obs.elev*time	0.71	51.53	< 0.001	0.76	36.79	< 0.001
2. Low-frequency species						
elev.change~obs.elev*time	0.46	49.89	< 0.001	0.51	40.51	< 0.001
3. High-frequency species						
elev.change~obs.elev*time	0.63	56.11	< 0.001	0.67	37.98	< 0.001
Site 2						
(c) 1990 as training set						
1. Total species						

elev.change~obs.elev*time	0.77	88.02	< 0.001	0.81	78.58	< 0.001
2. Low-frequency species						
elev.change~obs.elev*time	0.75	100.50	< 0.001	0.78	88.77	< 0.001
3. High-frequency species						
elev.change~obs.elev*time	0.79	94.54	< 0.001	0.80	90.26	< 0.001
(d) 2014 as training set						
1. Total species						
elev.change~obs.elev*time	0.36	181.80	< 0.001	0.35	178.40	< 0.001
2. Low-frequency species						
elev.change~obs.elev*time	0.49	189.00	< 0.001	0.45	182.10	< 0.001
3. High-frequency species						
elev.change~obs.elev*time	0.24	181.00	< 0.001	0.28	161.00	< 0.001

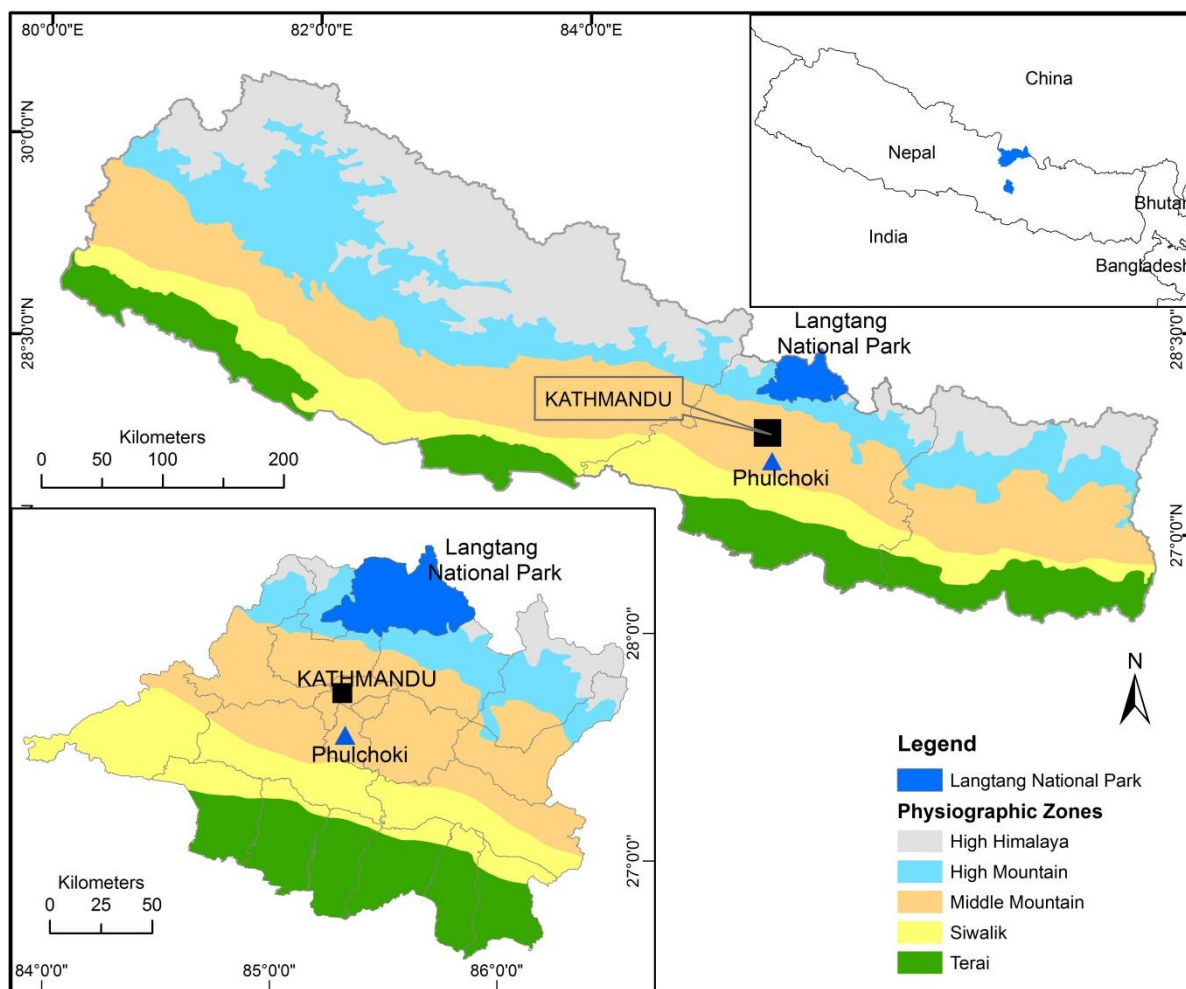


Fig. 1. Location map of the study area. Site 1 = Phulchoki mountain, Site 2 = Langtang National Park.

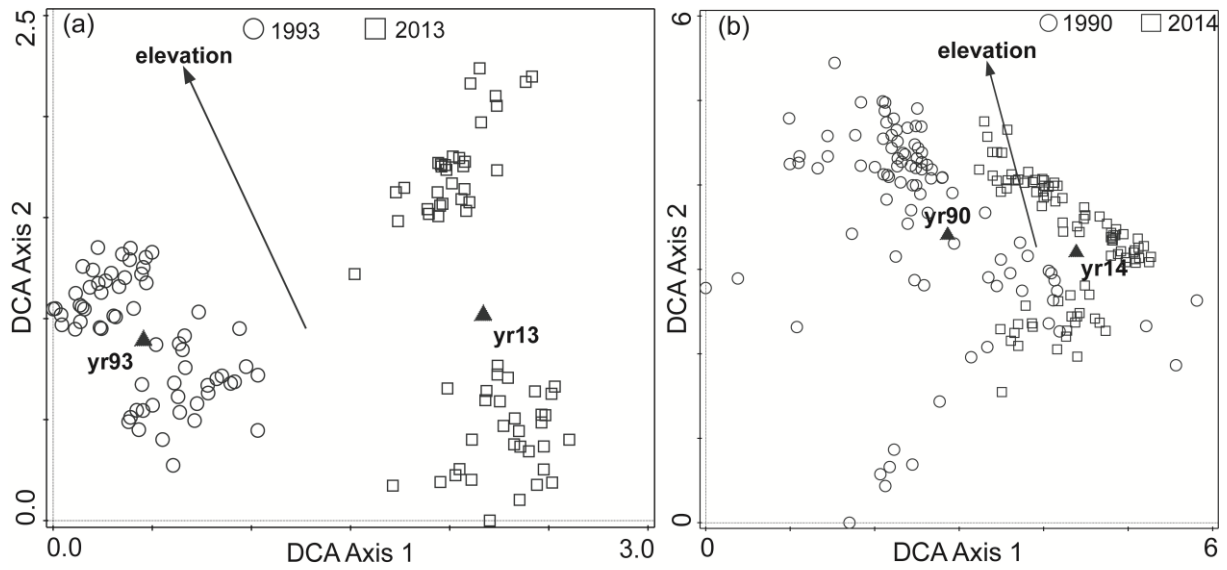


Fig. 2. Detrended Correspondence Analysis (DCA) biplot of species composition. (a) Site 1.

Eigenvalues: 0.45 and 0.21; gradient length: 2.60 and 2.20 standard deviations, for the first and

second axis, respectively. yr93 = year 1993, yr13 = year 2013. (b) Site 2. Eigenvalues: 0.55 and 0.49;
gradient length: 5.81 and 5.45 standard deviations, for the first and second axis, respectively. yr93 =
year 1993, yr13 = year 2013.

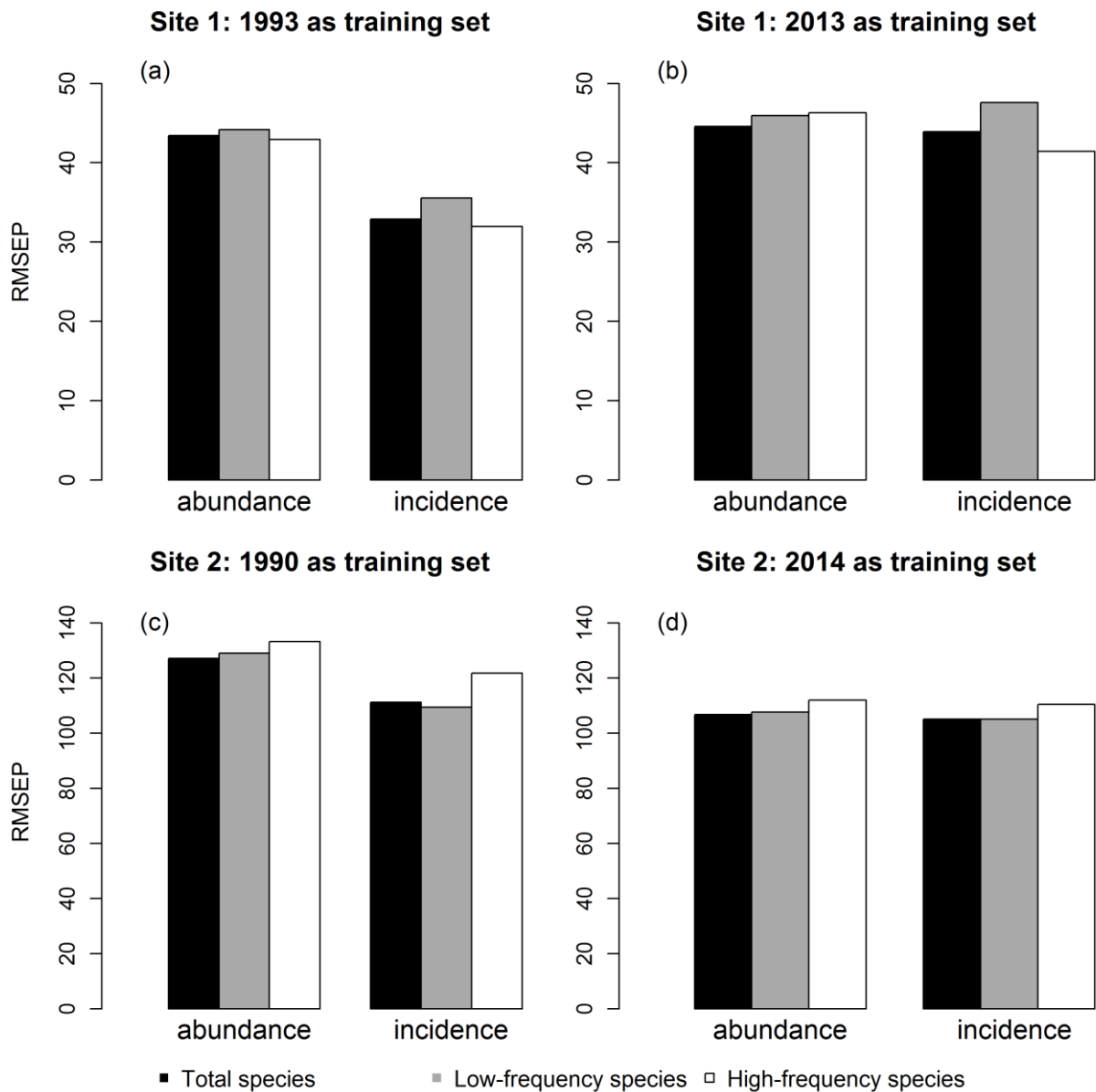


Fig. 3. Root Mean Squared Error of Prediction (RMSEP) of the weighted averaging (WA) regression models for different training sets from **(a, b) Site 1** and **(c, d) Site 2**, where the estimates produced by monotonic curvilinear regression deshinking procedures were selected. Cross-validated regression estimations are derived by bootstrapping with 1000 cycles.

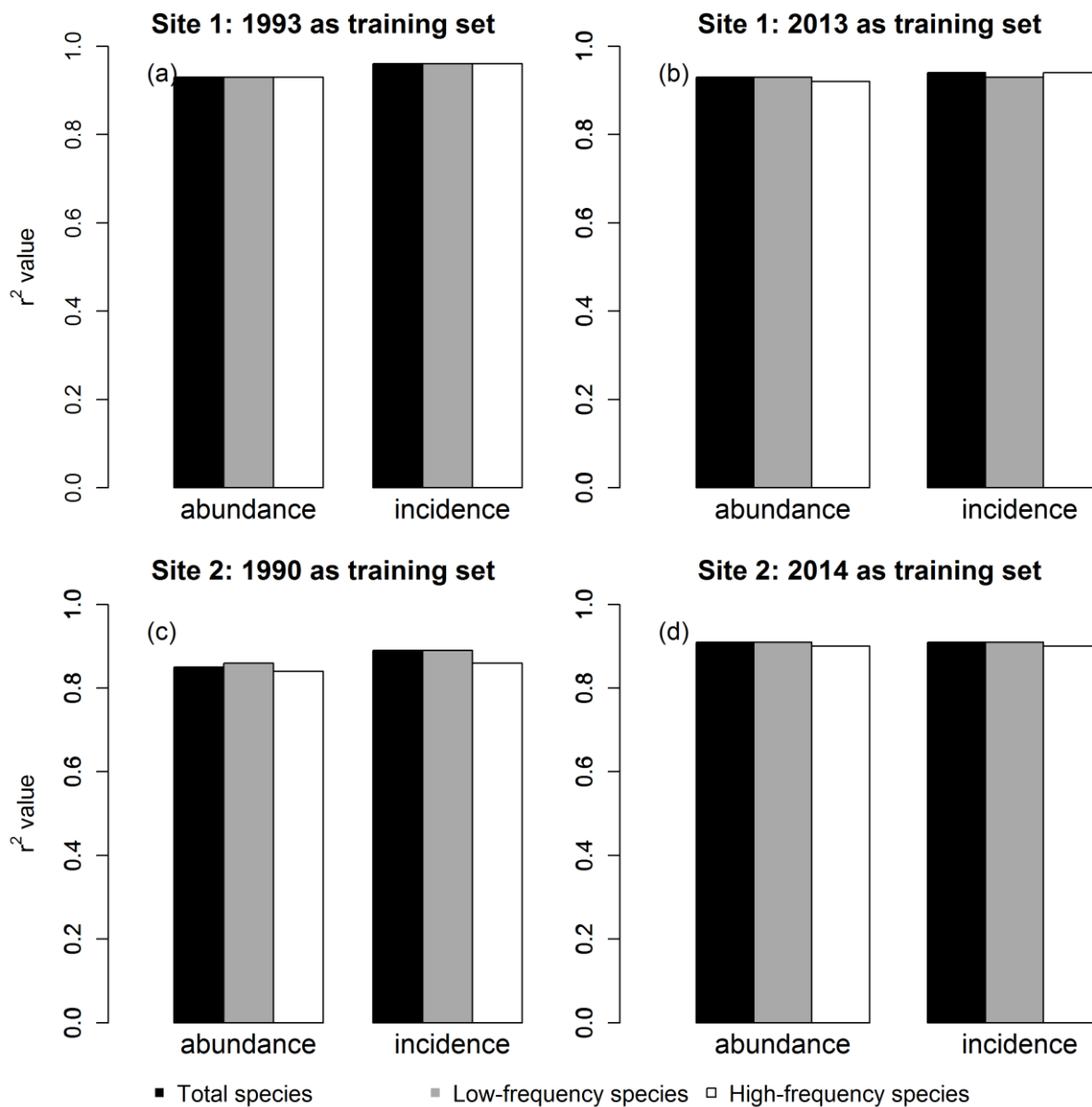


Fig. 4. Adjusted coefficient of determination (r^2) of the weighted averaging (WA) regression models for different training sets from **(a, b) Site 1** and **(c, d) Site 2**, where the estimates produced by monotonic curvilinear regression deshinking procedures were selected. Cross-validated regression estimations are derived by bootstrapping with 1000 cycles.

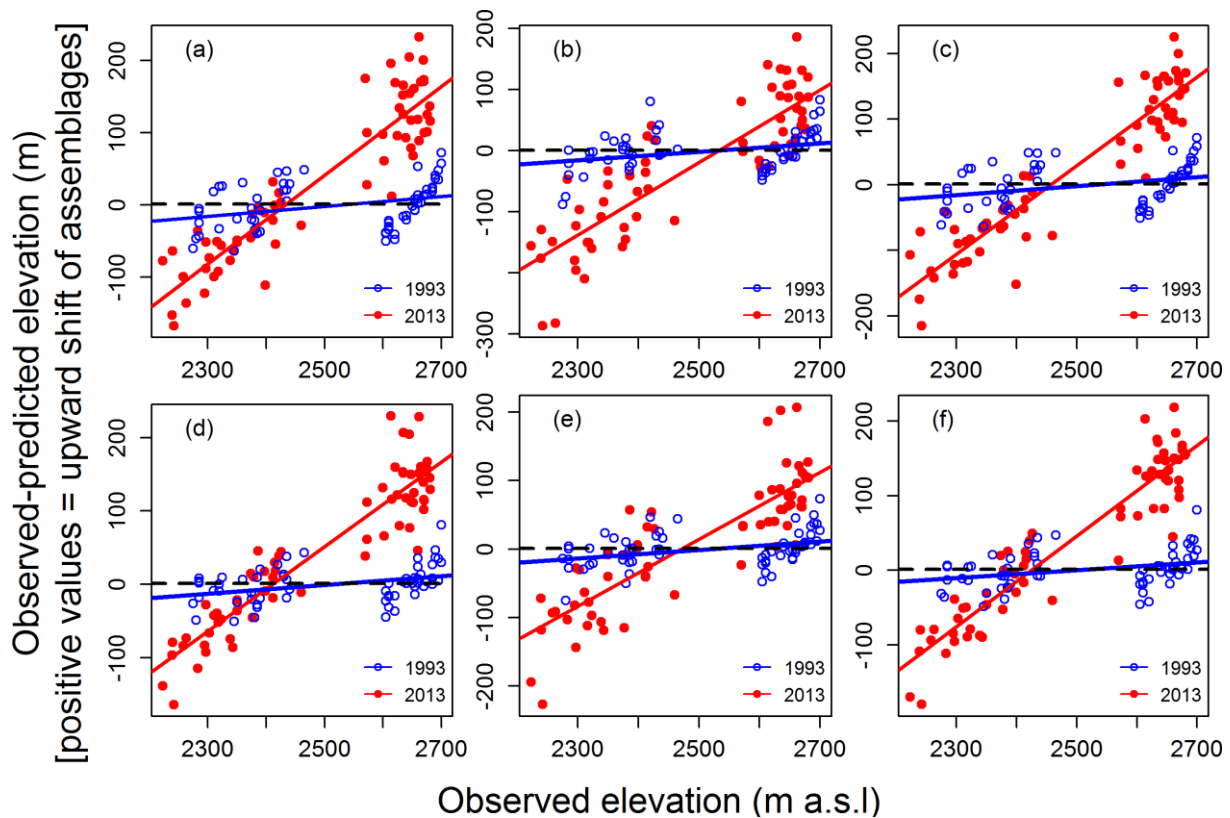


Fig 5. Change in the elevational distribution of the species assemblages from 1993 to 2013 at Site 1, where the difference in the observed and the predicted elevation of the sample plots of 2013 is the temporal change in elevation of the species assemblages (forward prediction). Elevation for the 2013 data was calibrated using the species-environment relationship of the training set (year 1993). The patterns of the elevational change are given for **(5a-5c) abundance datasets** (a) with total species, (b) with low-frequency species, and (c) with high-frequency species; **(5d-5f) incidence datasets** (d) with total species, (e) with low-frequency species, and (f) with high-frequency species.

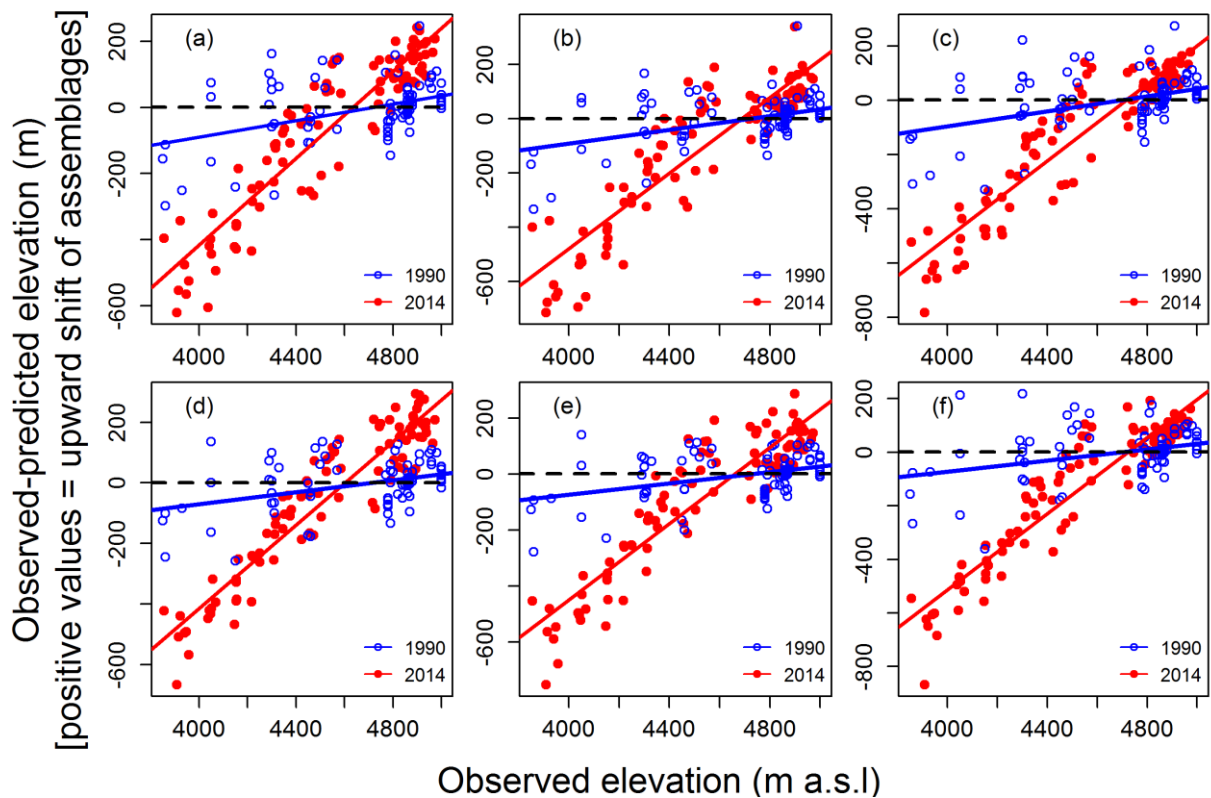


Fig 6. Change in the elevational distribution of the species assemblages from 1990 to 2014 at Site 2, where the difference in the observed and the predicted elevation of the sample plots of 2014 is the temporal change in elevation of the assemblages (forward prediction). Elevation for the 2014 data was calibrated using the species-environment relationship of the training set (year 1990). The patterns of the elevational change are revealed for **(6a-6c) abundance datasets** (a) with total species, (b) with low-frequency species, and (c) with high-frequency species; **(6d-6f) incidence datasets** (d) with total species, (e) with low-frequency species, and (f) with high-frequency species.