

Linking Fine-Scale Sub-Arctic Vegetation Distribution in Complex Topography with Surface-Air-Temperature Modelled at 50-m Resolution

Zhenlin Yang, Martin T. Sykes, Edward Hanna,
Terry V. Callaghan

Abstract Recent studies have shown that the complexities of the surface features in mountainous terrain require a re-assessment of climate impacts at the local level. We explored the importance of surface-air-temperature based on a recently published 50-m-gridded dataset, versus soil variables for explaining vegetation distribution in Swedish Lapland using generalised linear models (GLMs). The results demonstrated that the current distribution of the birch forest and snowbed community strongly relied on the surface-air-temperature. However, temperature alone is a poor predictor of many plant communities (wetland, meadow). Because of diminishing sample representation with increasing altitude, the snowbed community was under-sampled at higher altitudes. This results in underestimation of the current distribution of the snowbed community around the mountain summits. The analysis suggests that caution is warranted when applying GLMs at the local level.

Keywords Generalised linear model · Mountains · Vegetation distribution · Swedish sub-arctic · Scale

INTRODUCTION

A growing number of approaches have been proposed to investigate potential climate change impacts on vegetation distribution. One approach uses statistical methods such as the generalised linear model (GLM), the generalized additive model, classification tree analysis or an ensemble of several models (Araújo et al. 2005) to correlate current

distributions of plant species with a range of climatic variables and then project possible distributions under various climate scenarios. Applications of such models at the broad scale provide valuable scenarios for our understanding of the factors-driving global vegetation distributions, but their accuracy at the local level, especially for plant communities of areas with complex topography such as mountains and wet depressions, is often questioned (Trivedi et al. 2008; Randin et al. 2009).

It has been suggested that if statistical models are to be applied at the local level, the extent (the quantity of the data) and/or the scale (the quality of the data) of the study area need to be improved (Thuiller et al. 2004; Menke et al. 2009; Barbet-Massin et al. 2010; Braunisch and Suchant 2010). However, it is difficult to strike the balance between a proper spatial extent and scale in such models. On one hand, the macroscale model studies need to focus on the broad scale (Araújo et al. 2009), as a restricted study area will not capture the species potential distribution. This can lead for example to an over-prediction of the local extinction rates at the southern edge of a species' distribution (Barbet-Massin et al. 2010). On the other hand, the precision of climate/distribution data is reduced at the coarser scale. Macroscale models will underestimate the complexity of climate impacts at the local level (Trivedi et al. 2008; Randin et al. 2009).

Underestimating the complexities of the surface environment features in mountainous terrain may render many conclusions arising from macroscale (tens of kilometres' scale) studies inapplicable at the local level. For example, the extinction risks of alpine vegetation were exaggerated at the macroscale (Thuiller et al. 2004; Pearson et al. 2006; Botkin et al. 2007), because the macroscale climate data used was biased against local climate experienced by alpine plants. Other studies have shown that the accuracy

Electronic supplementary material The online version of this article (doi:10.1007/s13280-012-0307-0) contains supplementary material, which is available to authorized users.

of the modelling projection will be substantially improved by including topoclimatic variables (i.e. based on elevation ranges) (Luoto and Heikkinen 2008). In addition, the modelling of fine-grained spatial variation in warming may be more important at the finer landscape scale (Ashcroft et al. 2009). In many cases, a combination of biotic and topoclimatic variables may produce the best projection of vegetation distribution (Luoto and Heikkinen 2008).

Mountains are not the only feature of complex topography that controls the spatial distribution of surface-air-temperature and vegetation. Small-scale depressions and specific hydrological conditions that create wetlands provide added complexity (Reddy and DeLaune 2008). In the Arctic, proximity to the sea, and even sea ice conditions affect land surface temperatures (Hanna et al. 2004; Bhatt et al. 2010) and proximity to a large lake in the sub-arctic affects land surface temperatures with unknown effects on the vegetation distribution (Zhenlin et al. 2012).

The complexity of such factors at the landscape level in the context of current and projected climate changes require a re-assessment of progress in vegetation modelling approaches. The use of bioclimatic models is one approach. In spite of their robustness, as with all vegetation modelling approaches there are a number of uncertainties and assumptions in bioclimatic models, depending on the choice of modelling techniques (Pearson et al. 2006), the quality of modelling input data (Wiens et al. 2009) or important processes such as population dynamics or migration (Brooker et al. 2007). Theoretically, the input data need to be well distributed throughout the species' geographical range as well as its complete environmental envelope (Franklin 2009). In reality, the extent of a study is often truncated due to practical or geographical limits, and these limits therefore result in a biased representation of a species' geographical range in the input data. This bias limits the wider applicability of bioclimatic models (Menke et al. 2009; Barbet-Massin et al. 2010). However, the impact of the bias on model projection depends on the magnitude of the variability in the input data compared to that of the species' full geographical and environmental ranges (Kadmon et al. 2004).

As mountains can be considered to be cone shaped, with decreasing land surface with increasing elevation, the proportion of observations of alpine species at the summit will be smaller than their actual distributions within the same climatic space in low-elevation areas. Accordingly, random sampling will tend to under-sample high elevation areas, and thus by sampling randomly one will more often include low-elevation areas (Fig. 1). This, in turn, results in a 'climatic bias', denoted as 'the sampling bias with respect to climatic conditions in the observations available for a particular species' (Kadmon et al. 2003). While the impact of sampling strategies (and thus their associated bias) on

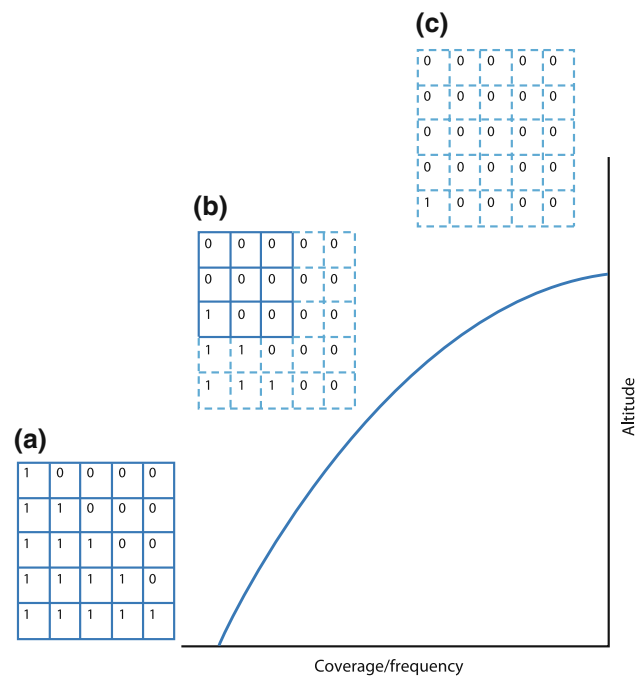


Fig. 1 Conceptual diagrams of the restricted sampling and incomplete sampling at different altitudes: **a** complete sampling. **b** Restricted sampling. **c** Incomplete sampling. Grids with the dotted lines are the hypothetical distribution while grids with the solid lines are the actual distribution. The hypothetical distribution represents the potential distribution without space limits

model projections was discussed by Hirzel and Guisan (2002), to the best our knowledge, the influence of such local-level climatic bias among the mountains has never been tested.

A 50-m resolution temperature model, based on historical measurements from 1913 to 2010 in Abisko, northern Sweden (Zhenlin et al. 2011), provided a unique opportunity to explore if vegetation distribution is related to fine-scale surface-air-temperature variation that is affected by local landscape features other than topography, for example proximity to a large lake. In this study therefore, our aims are: (a) to explore the importance of surface-air-temperature and soil variability as drivers of the vegetation distribution at 50-m resolution and (b) to assess sampling errors due to diminishing sample representation with increasing altitude. Our specific questions are: (a) to what extent are surface-air-temperature and soil variability at 50-m resolution associated with the current distribution of sub-arctic plant communities in a sampling area with complex topography? and (b) How does the topographically biased sampling affect the projection of the current vegetation distribution through the geographic range sampled?

We emphasise first that this study is not intended to generate a model to use in projections of future vegetation distribution but simply seeks to address the security of the foundations on which such projections are built. Second,

the spatial scale of 50 m was appropriate to reflect the quality for existing vegetation distribution data: conceptually, temperature at even smaller scales (e.g. mm) is appropriate for seed germination but our choice of 50-m scale is based on availability of a recently published, pioneering surface-air-temperature dataset as well as similar scale vegetation and soil data, and is sufficient to describe the distribution of major vegetation patches in the study area.

MATERIALS AND METHODS

Study Area

The study area is within the Torneträsk catchment (68°21'N, 18°49'E), which covers a region of 700 km² around the Abisko Scientific Research Station (ANS) in the sub-arctic zone of Fennoscandia (Fig. 2). There have been long-term human activities in this region such as reindeer herding, fishing and hunting. Although human activities have increased as the construction of the railway was completed in 1902, it is still considered to be one of the least disturbed natural areas in Europe. The east part of the catchment lies in the rain-shadow of the Norwegian mountains. As a result, there is a pronounced ocean-continental gradient from west to east in this region. The soil in

this region is mainly dominated by bare rocks and moraines, with sporadic peatlands. The soils are nutrient-poor in the west–east direction but nutrient rich in the central part. These climatic and edaphic conditions are associated with a distinctive vegetation composition and distribution: the most widespread plant communities are the birch forest (i.e. *Betula pubescens* ssp. *czerepanovii*), heath vegetation [e.g. *Arctostaphylos alpines* (L.) Sprengel, *Arctostaphylos uva-ursi* (L.) Sprengel, *Betula nana* L., *Empetrum hermaphroditum* ssp. *nigrum*], meadow [e.g. *Bistorta vivipara* (L.) Gray, *Calamagrostis lapponica* (Wahlend.) Hartm., *Filipendula ulmaria* (L.) Maxim, *Geranium sylvaticum* L.], and snowbed [e.g. *Carex bigelowii* Torr. ex Schwein, *Cassiope hypnoides* (L.) D. Don, *Gnaphalium supinum* L.], with other sparse plant communities such as bogs with and without mosses (e.g. *Dicranum fuscescens*, *Kiaeria Starkei*, *Oligotrichum hercynicum*, *Pohlia* spp., *Polytrichum hyperboreum*). The local forest limit is at an altitude range of about 550–650 m in the western part and 700–800 m in the eastern part, and the birch forest grades into the heath type at higher altitudes (Callaghan and Karlsson 1996).

Data

This study used 50-m resolution monthly average gridded surface-air-temperature data (1913–2010) for the Abisko region (Zhenlin et al. 2011). This dataset represents the

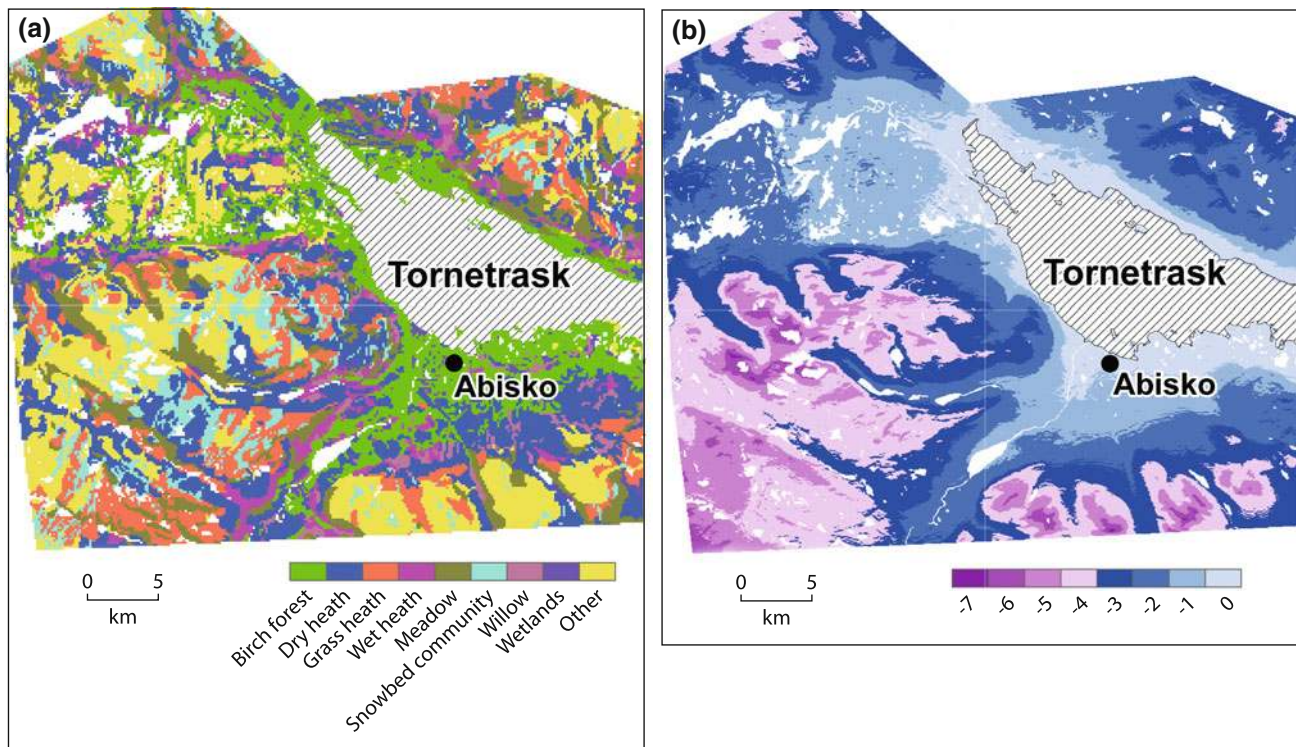


Fig. 2 **a** Vegetation map of the Abisko region. *Other* denotes glaciers, built-up areas and bare rocks. **b** The pattern of mean annual temperature during 1913–2010 of the Abisko region. *White areas* denote waters

latest development of a surface-air-temperature distribution model and is built on empirical equations and a digital elevation model (DEM, Swedish National Land Survey). The dataset is output from this model based on: (1) gridded ERA-40 reanalysis data from European Centre for Medium-Range Weather Forecasts representing the large-scale/synoptic macroclimatology around the Abisko region; (2) surface-air-temperature observations from a long-term (40 years) network of 20 surrounding meteorological stations belonging to the Swedish Meteorological and Hydrological Institute (SMHI) or the Norwegian Meteorological Institute and 20 short-term (2 years) temperature loggers and (3) key microclimatic characteristics of the temperature distribution (i.e. exposure, aspect, lake effect). Output from the model was validated by independent temperature measurements, with an estimated error of 1 °C (Zhenlin et al. 2011). Due to the lack of high-altitude temperature measurements, the original model was limited to an altitude range of 335–800 m a.s.l. However, in this study, the same model was assumed to be applicable for the whole region, therefore the temperature data were developed for the whole altitude range of 335–1555 m a.s.l. The uncertainty of this will be discussed later.

Principal components analysis (PCA) was used to avoid multi-collinearity while keeping the monthly variance of the temperature variables. The sum of the first and second principal components (T-PCA1 and T-PCA2) represented 98.4 % of the total variance in the temperature variables (79.7, 18.6 %, respectively), with T-PCA1 and T-PCA2 capturing the temperature pattern during the summer and mainly from the winter, respectively (Table S1 in Electronic Supplementary Material).

The vegetation data were interpreted from the vegetation map by Metria May 1998 (25-m scale), produced by the Swedish National Land Survey from colour infrared photography and field visits (Tomas 1998). The lake Torneträsk, streams, water, and island within the lake were excluded from the map. For each 50-m grid, the presence/absence information was recorded from the vegetation map. To reduce the complexities of the research while keeping most of the information relevant to this study, the map were simplified into the eight categories of plant communities listed in Table S2 (Electronic Supplementary Material). The soil data were derived using cartography data from the Geological Survey of Sweden (Rodhe et al. 1999). Compared to the vegetation data, the quality of the soil data is relatively poor. Soil types with coverage of <0.1 % were removed. Each soil type's potential for supplying moisture and nutrients for plant growth was subjectively assessed into five categories (Table S3 in Electronic Supplementary Material).

Both the vegetation and the soil maps were corrected to the Swedish grid system (RT 90 25), first digitized as

vectors, and then converted to 50 × 50-m grid cells to match the scale of the temperature dataset. While we do not have detailed coverage data of each plant community/soil type at each gridcell, the presence/absence information was derived for each vegetation and soil type by the ArcGIS grid module, and stored in a binary layer.

Our surface-air-temperature model output is more appropriate to use here than also recently available satellite (thermal infrared) temperature scans because the latter: (1) only show relative temperature changes and are not based on actual surface temperature measurements; (2) variously measure surface skin and canopy temperatures depending on the vegetation cover of the ground area being observed, so are not consistent in their depiction of (near)-surface conditions; (3) are only available for a relatively short time period of the last two to three decades or so at most, giving a limited meteorological baseline and (4) have many missing gaps caused by adverse weather conditions (cloud cover) precluding surface viewing.

Statistical Analysis

Spatial Autocorrelation

Spatial autocorrelation can overestimate the degree of fitness in linear models. One strategy to reduce its effect is to resample the whole map at a distance where the spatial autocorrelation becomes negligible. Moran's *I* test (Rangel et al. 2006) was used to measure the degree of spatial autocorrelation within a neighbour matrix. According to Moran's *I* test, the spatial autocorrelation for the datasets in this study decreased monotonically after a lag of four grid cells (200 m). Consequently, the original datasets were re-sampled so that no grid cells were closer than 200 m both in the west–east and the north–south directions, and the spatial autocorrelation was assumed to become negligible thereafter.

Multi-collinearity

Hierarchical partition (HP) analysis was adopted to select those climate and soil variables, which independently explained the most variation in the vegetation distribution (Chevan and Sutherland 1991). The significance of the variables in the HP analysis was tested by 1000 randomizations of the variables using the 95 % confidence limit of the *z* scores ($z = [\text{observed value} - \text{mean}(\text{randomizations})]/\text{sd}(\text{randomizations})$, $z \geq 1.65$) (Mac Nally 2002; Kuhn et al. 2004). Consequently, only those variables that had significant impact on the vegetation distribution were included in the later analysis.

The Generalized Linear Model (GLM)

The GLM with a binomial distribution and logistics link is one of the most common methods to explore current climate/bioclimate variables that best explain species or vegetation distributions. To capture the potential nonlinear relationship while reducing data complexities to interpret the biological significance of higher order variables, polynomial terms up to the second order were included in the model.

In the GLM model, to reduce the model complexities, only those variables showing a significant relationship with the presence/absence of each plant community were entered into the model. The final model was selected by Akaike's information criterion (AIC) derived from the stepwise-based algorithm (Akaike 1973). Second, the Area Under the Receiver Operating Characteristic Curve (AUC) was used to assess the models' performance (Fielding and Bell 1997). Third, the original probability surface was kept for the model evaluation. The projected response curves were used to describe the probability of plant communities' occurrence in relation to environmental variables (Thuiller et al. 2004), and compare with the actual probability of plant communities' occurrence from the vegetation presence/absence data.

Variation Partitioning

Variation partitioning has been used often to identify the relative influences of independent variables in the model (Desdevises et al. 2003). In this study, the coefficients of the determination of the binary logistic regressions, Nagelkerke Pseudo R-Square (Nagelkerke 1991), were taken as measures of the percentage of the variation explained by different combinations of the variables, namely (1) both the temperature and the soil variables; (2) the temperature variables and (3) the soil variables.

All the variables, except the vegetation presence/absence data, were scaled according to their root-mean-squares, to simplify the comparison of the effects from different variables. The PCs, combined with non-continuous explanatory soil variables, helped to avoid potential correlation between the main effects in the further analysis. Consequently, the standardized regression slopes from the GLM were comparable independent of the scales at which the variables were measured. Statistical analyses were performed with R using the packages Hier.Part, Presenceabsence (Walsh and Mac Nally 2003; Freeman 2007) and SPSS for other standard statistical analyses. A random sample of 70 % of the dataset was used to calibrate the model while 30 % of the dataset was used to evaluate the modelling results.

RESULTS

Climatic Misrepresentation

Within the study area, 50.8 % of the snowbed community was distributed around the mountain summits (colder climatic space) (T-PCA1 < -1.5), while the samples over the same climatic space around the mountain summits (T-PCA1 < -1.5) only represented 13.8 % of the total samples (Fig. 3). It becomes obvious that the climatic misrepresentation of the snowbed community distribution in this study violated the assumption that the input data need to be well distributed along the whole environmental gradient. This makes it necessary to evaluate further how the GLM model performance will be affected by such climatic misrepresentation of the input data.

Hierarchical Partition

The presence of the birch forest and snowbed community can be significantly explained by the temperature/soil variables, with up to 42 and 15 % of the total variance in the distribution of each plant community explained: the temperature variables were associated with 91 % of the total explained variance for the birch forest, while the soil variables explained much less (<10 %) of the total explained variance (Fig. 4); the explained variance of the snowbed community was in particular related to temperature (86 %), and far less to rock (6.7 %), till (6.7 %) and fluvial sediment (<1 %).

The GLM

Based on the grid-by-grid accuracy measurement using AUC values, the GLM agreed with the results from the hierarchical analysis in that some plant communities, such as the widespread birch forest and the more restrictedly distributed snowbed community which occupied 28 % of the whole area are well predicted and had AUC values up to 0.93 and 0.88, respectively. The model performances for other plant communities were, however, poorer, e.g. the meadow, the dry heath and wetland species (AUC ≤ 0.80, Table S4 in Electronic Supplementary Material).

The relative effect of temperature during summer on vegetation distribution was stronger than that during the winter and differed among alpine and subalpine vegetation: the birch forest tended to be more widespread the warmer the summer, while the presence of the snowbed community could be explained by the mild and shorter summers (Table S5 in Electronic Supplementary Material). Both the rock and the fluvial sediment were positively associated with the presence of the snowbed community.

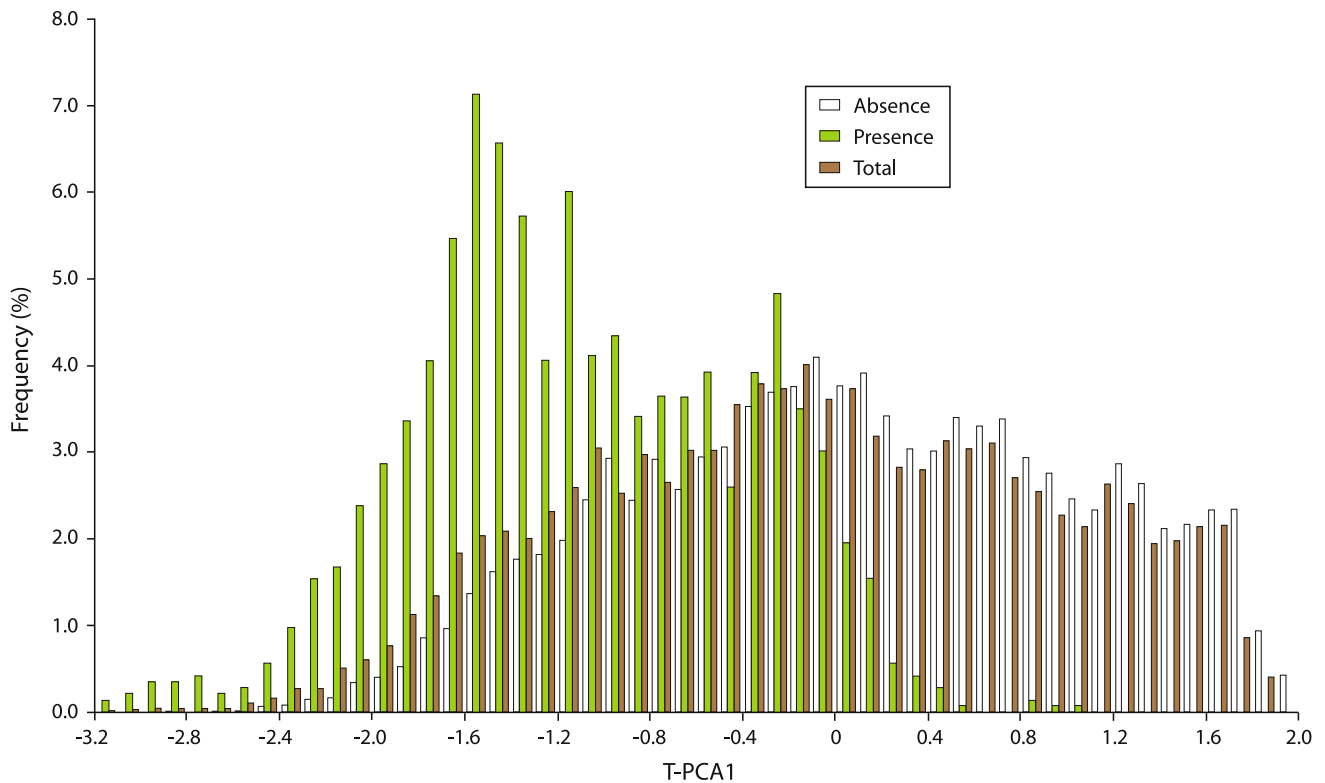
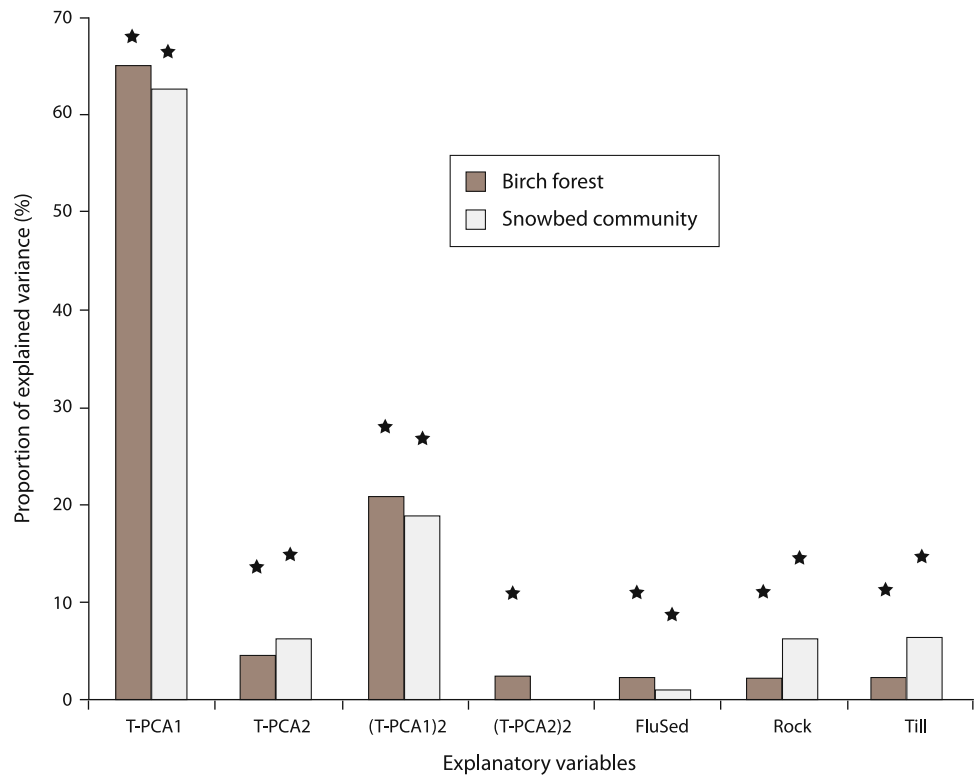


Fig. 3 Frequency diagram of the “Presence”/“Absence”/“Total” grid cells at different climatic spaces. “Presence”/“Absence” grid cells denote the proportion of grid cells with the presence/absence of

snowbed community. “Total” grid cells denote the proportion of grid cells (presence + absence). T-PCA1 denotes the first component of the monthly temperature

Fig. 4 HP of the proportions of explained variance in vegetation distribution using temperature and soil variables as explanatory variables. Stars indicate significant results (being outside the 0.95 confidence range) after 1000 randomizations. FluSed denotes the fluvial sediments as in Table S3 (Supplementary material). T-PCA1 denotes the first component of the monthly temperature. T-PCA2 denotes the second component of the monthly temperature. (T-PCA1)² denotes the square of T-PCA1. (T-PCA2)² denotes the square of T-PCA2



The GLM-predicted probability surfaces for the modelled present vegetation distributions were compared with the actual vegetation distribution (Fig. 5). Following Thuiller (2004), we present the response curves for the most significant variables among all plant communities, i.e. the first principle component for temperature (Fig. 6). For the birch forest, the shape of the predicted response curve agrees quite well with that of the actual response curve (Fig. 6b). However, there are marked differences in the tails of the response curve for the snowbed community: the probability of occurrence for the snowbed community was underestimated when the T-PCA1 was lower than -2.5 (Fig. 6a).

Variation Partitioning

In total, the variables explained 54.7 % of variance in the distribution of the birch forest, and 28.7 % of that in the snowbed community (Table 1). Compared to the soil variables, the temperature variables explained more of the variance in the vegetation distribution, especially for the birch forest (38.9 %) and snowbed community (20.6 %). For distributions of the wetland species, willow, wet heath, extreme dry heath and meadow, the explained portions of the variance were quite low. There were also some joint effects between temperature and soil variables for the distribution of the snowbed community and birch forest (up to 6.9 %) (Table 1).

DISCUSSION

Accurate representation of current vegetation distribution hinges on adequate sampling of the entire environmental range of conditions. However, data on all possible environmental conditions that affect vegetation are rarely if ever available and thus these factors affecting the current distributions of vegetation may not be fully understood. Our findings illustrate that caution is warranted when applying a GLM at the local level. For example if only restricted distribution data is available this may result in a truncated response curve and misrepresentation of the current, and therefore also future vegetation distributions.

In our study, the GLM model is driven by fine-scale topoclimate data. However, due to the conical geometry of the mountains (Fig. 3), the restricted sampling of the snowbed community among the mountains is not representative of the snowbed community's potential distribution, which extends into high Arctic regions (Walker et al. 2005). Such misrepresentation results in a truncated response curve and underestimation of the current distribution of the snowbed community at the mountain summits. This is the first study, however, that focuses on

uncertainties raised from the input data of the fine-scale GLMs rather than on uncertainties in the macroscale models.

Nevertheless, within even a restricted vegetation geographical range, it is still possible to build a useful vegetation model if the aim is to understand the local drivers of the vegetation distribution (Franklin 2009). Previous studies have shown that such local-scale models are different from macroscale models in at least two aspects (Trivedi et al. 2008; Randin et al. 2009). First, the heterogeneous topoclimate at the local level, for example resulting from different slopes, aspects, hollows etc. are likely to provide refugia for local species (Ashcroft et al. 2009). Studies that include topoclimate variability or topographic indices have therefore improved the projection of vegetation distribution (Illán et al. 2010). Second, biotic interactions (e.g. dispersal and interspecific competition) may play vital roles in determining the persistence of species at the local level (Heikkinen et al. 2006). The potential importance of biotic interactions has been advocated in a number of studies incorporating population dynamics processes (Travis et al. 2006; Brooker et al. 2007). While detailed information regarding biotic interactions will be difficult to quantify even at the local level, it is well known that in Abisko, biotic interactions such as herbivory and disturbance can over-ride climate controls on growth and range expansion (Van Bogaert et al. 2011) while facilitation of tree seedlings by shrubs has been demonstrated at treeline (Grau et al. 2012). However, such biotic interactions were not components of our current model.

In this study, topoclimate was ranked most highly in the variance explained in the GLM model (Table 1). The low explanatory power of the soil variables in Table 1 (<5 %) is not unexpected due to the poor quality of the soil map. With a more detailed investigation of the soil properties, both the edaphic and the climatic factors are likely to have substantial impacts on the vegetation distribution (Darmody et al. 2004). Our results are consistent with previous studies suggesting that climate, especially temperature is crucial for vegetation distribution over this region (Edenius et al. 2003), while influences from the soil, land use and topography are expected to shape the heterogeneity of the vegetation distribution at a local level (Darmody et al. 2004). For example, the peatland occurs mostly in open, relatively flat and slightly raised areas (Lang et al. 2009). For the distribution of the snowbed community, snow cover is obviously an important factor (Björk and Molau 2007) and technologies such as digital photography, process modelling and remote sensing be used to determine changes in snow cover and may provide additional improvement for this study (Christiansen 2001; Dankers and De Jong 2004; Liston and Hiemstra 2010).

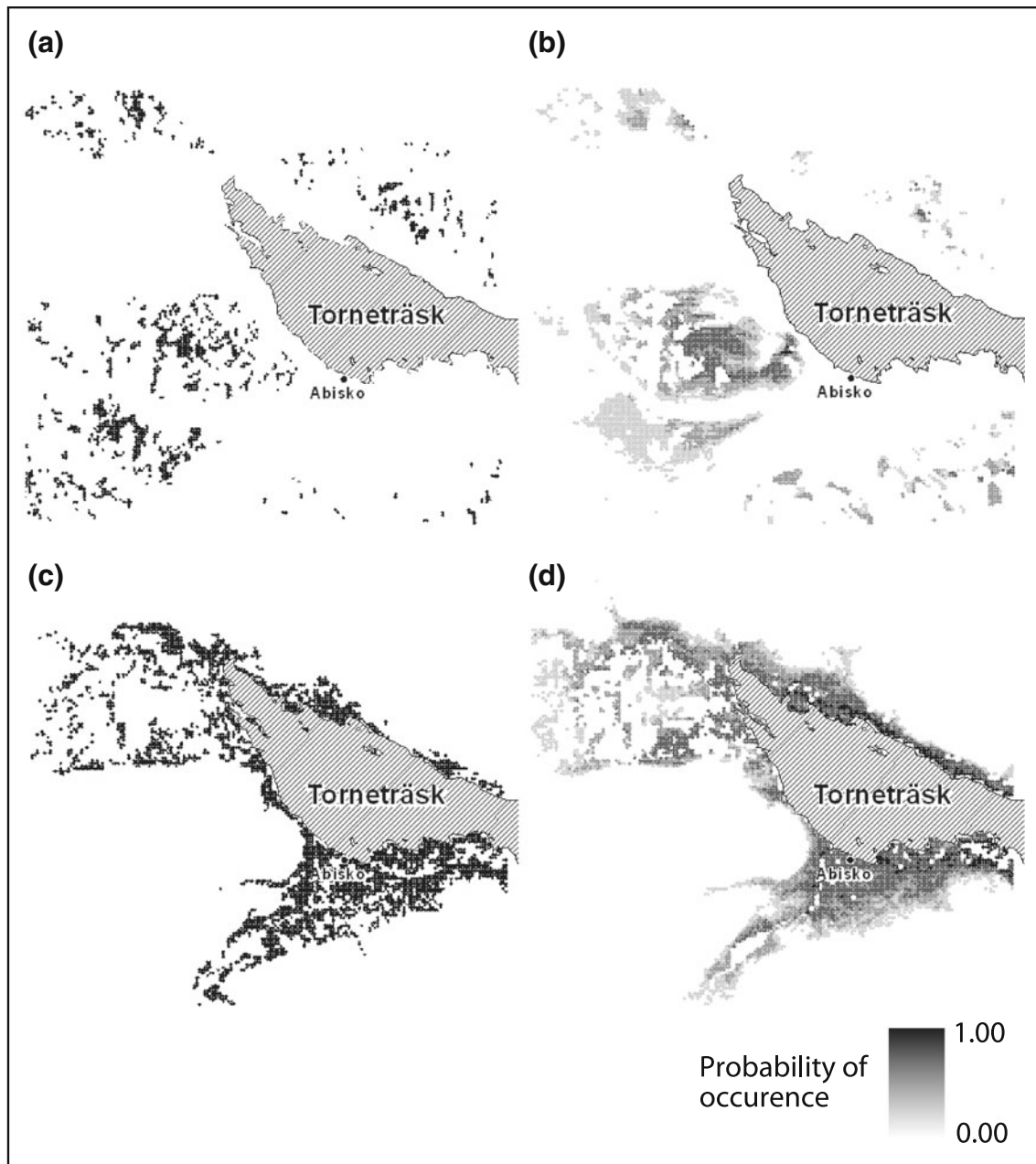


Fig. 5 Comparison of the actual and projected probability of occurrence for the vegetation distribution; **a** the actual probability of occurrence of the snowbed community, **b** the projected probability

of occurrence of the snowbed community, **c** the actual probability of occurrence of the birch forest and **d** the projected probability of occurrence of the birch forest

Many uncertainties with regard to the use of bioclimatic models arise from either the modelling techniques or the input data. While we are aware of the uncertainty due to the choice of the modelling techniques, such uncertainties are beyond the scope of this study and indeed have been discussed elsewhere (Heikkinen et al. 2006). Here, we mainly focus on the uncertainties raised by the climate/vegetation distribution data. The 50-m resolution surface-air-temperature data adopted here is a development of a well-based temperature dataset derived from the combination of

climatology datasets at different scales. Both the macro-scale climatology from the synoptic-scale climate dataset and the microclimatology from surrounding weather stations were integrated in this dataset, which therefore partly addressed the problem of the over-simplification of the microclimatology. Besides this, part of the validation temperature data is collected from re-samplings at the same locations after more than 20 years (Zhenlin et al. 2011). However, notwithstanding our effort to improve the quality of the temperature data, some uncertainties remain. For

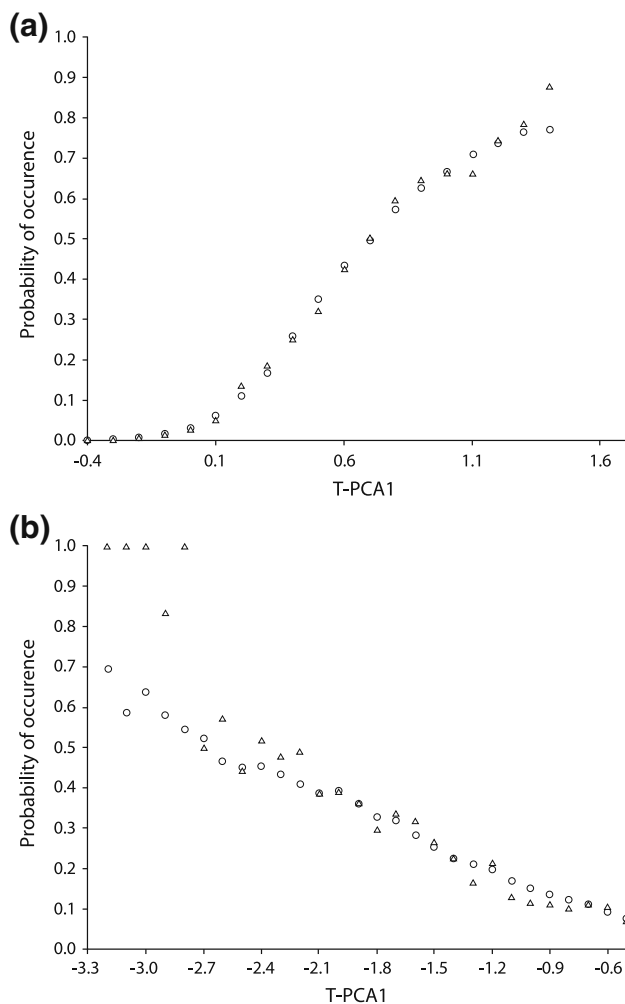


Fig. 6 Comparison of the actual (*circles*) and projected (*triangles*) response curves of **a** birch forest, **b** snowbed community. *T-PCA1* denotes the first principal component of the temperature

example, our understanding of the surface-air-temperature

Table 1 Partitioning of the variance in vegetation distribution explained by the binary logistic regressions model resulting from the stepwise selection of explanatory variables ($P < 0.05$)

R^2	Overall (%)	Temperature only (%)	Soil only (%)	Joint effect
Birch forest	54.7	38.9	1.5	4.3
Snowbed community	28.7	20.6	1.2	6.9
Grass heath	26.7	19.7	0.1	2.9
Wetland	15.8	12.7	0.8	2.3
Willow	15.1	8.1	2.3	4.7
Wet heath	13.3	12.5	0.8	0
Dry heath	<10	–	–	–
Meadow	<10	–	–	–

R^2 (Nagelkerke R-square) correspond to the amount of explained variance

at high altitudes is still poor in present temperature data due to the lack of high-altitude meteorological stations above 800 m a.s.l (Zhenlin et al. 2011). Moreover, surface-air-temperature is not an adequate substitute for climate as a whole, as other climate variables such as precipitation, wind etc. are known to be important as regulators of vegetation distribution but are not available as fine-scale datasets. Nor did we account for the influence of the snow cover, vegetation canopies, together with the local topographical and hydrological conditions that complicate the interaction in the energy exchange system among the mountains at scales down to centimetre resolution.

The quality of vegetation distribution data also affects the validity of the modelling results of this study. Depending on the data availability, the vegetation model can be built at either a community or species level. One of the main arguments in favour of the modelling at a community level is the difficulty to get detailed data such as spatial distributions at a species/plant functional type level due to the limits of both time and economic resources. Consequently, the grouping of plant species at a community level with similar long-term responses to environmental changes reflects a realistic trade-off between the information and the structural complexity.

It is important for stakeholders to have predictive models/assessments of climate change impacts so they can adapt to changes, for example in local ecosystem services. We have not produced such a model, but we have quantified the extent to which vegetation distribution is related to temperature and soil variability at the local level and identified the local complexity of achieving this. We have shown how caution is needed when modelling present vegetation distribution at the local level and have established some high-resolution databases on which future local-scale models can be developed.

Acknowledgments This study was conducted as part of the Marie Curie Early Stage Training network—Multiarc-supported by European Union FP7. This study was also partially supported by FORMAS projects “Climate change, impacts and adaptation in the sub-Arctic: a case study from the northern Swedish mountains” (214-2008-188) and “Advanced Simulation of Arctic climate change and impact on Northern regions” (214-2009-389). The authors wish to thank two anonymous reviewers for their comments. The authors are grateful to Eva Kuster, Jonas Åkerman, Christer Jonasson, and Jonathon Seaquist for valuable comments. We would like to thank Paul Coles for his help to redraw the graphs. We would like to thank Abisko Scientific Research Station staff for help and data collection.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. Paper presented at the 2nd international symposium on information theory, Akademiai Kiado. Budapest, Hungary.

- Araújo, M.B., R.J. Whittaker, R.J. Ladle, and M. Erhard. 2005. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* 14: 529–538. doi:10.1111/j.1466-822X.2005.00182.x.
- Araújo, M.B., W. Thuiller, and N.G. Yoccoz. 2009. Reopening the climate envelope reveals macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences of the United States of America* 106: E45–E46.
- Ashcroft, M.B., L.A. Chisholm, and K.O. French. 2009. Climate change at the landscape scale: Predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology* 15: 656–667.
- Barbet-Massin, M., W. Thuiller, and F. Jiguet. 2010. How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography* 33: 878–886. doi:10.1111/j.1600-0587.2010.06181.x.
- Bhatt, U.S., D.A. Walker, M.K. Reynolds, J.C. Comiso, H.E. Epstein, G. Jia, R. Gens, J.E. Pinzon, et al. 2010. Circumpolar arctic tundra vegetation change is linked to sea ice decline. *Earth Interactions* 14: 1–20. doi:10.1175/2010ei315.1.
- Björk, R.G., and U. Molau. 2007. Ecology of alpine snowbeds and the impact of global change. *Arctic, Antarctic, and Alpine Research* 39: 34–43.
- Botkin, D.B., H. Saxe, M.B. Araújo, R. Betts, R.H.W. Bradshaw, T. Cedhagen, P. Chesson, et al. 2007. Forecasting the effects of global warming on biodiversity. *BioScience* 57: 227–236. doi:10.1641/b570306.
- Braunisch, V., and R. Suchant. 2010. Predicting species distributions based on incomplete survey data: The trade-off between precision and scale. *Ecography* 33: 826–840. doi:10.1111/j.1600-0587.2009.05891.x.
- Brooker, R.W., J.M.J. Travis, E.J. Clark, and C. Dytham. 2007. Modelling species' range shifts in a changing climate: The impacts of biotic interactions, dispersal distance and the rate of climate change. *Journal of Theoretical Biology* 245: 59–65.
- Callaghan, T.V., and P.S. Karlsson. 1996. Plant ecology in subarctic Swedish Lapland: Summary and conclusions. *Ecological Bulletins* 45: 220–227.
- Chevan, A., and M. Sutherland. 1991. Hierarchical partitioning. *The American Statistician* 45: 90–96.
- Christiansen, H. 2001. Snow-cover depth, distribution and duration data from northeast Greenland obtained by continuous automatic digital photography. *Annals of Glaciology* 32: 102–108.
- Dankers, R., and S.M. De Jong. 2004. Monitoring snow-cover dynamics in Northern Fennoscandia with SPOT VEGETATION images. *International Journal of Remote Sensing* 25: 2933–2949. doi:10.1080/01431160310001618374.
- Darmody, R.G., C.E. Thorn, P. Schlyter, and J.C. Dixon. 2004. Relationship of vegetation distribution to soil properties in Karkevagge, Swedish Lapland. *Arctic, Antarctic, and Alpine Research* 36: 21–32.
- Desdèvises, Y., P. Legendre, L. Azouzi, and S. Morand. 2003. Quantifying phylogenetically structured environmental variation. *Evolution* 57: 2647–2652.
- Edenius, L., C.P. Vencatasawmy, P. Sandström, and U. Dahlberg. 2003. Combining satellite imagery and ancillary data to map snowbed vegetation important to reindeer *Rangifer tarandus*. *Arctic, Antarctic, and Alpine Research* 35: 150–157.
- Fielding, A.H., and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.
- Franklin, J. 2009. *Mapping species distributions: Spatial inference and prediction*. Cambridge: Cambridge University Press.
- Freeman, E. 2007. PresenceAbsence: An R Package for Presence-Absence Model Evaluation, USDA Forest Service.
- Grau, O., J.M. Ninot, J.M. Blanco-Moreno, R.S.P. van Logtestijn, J.H.C. Cornelissen, and T.V. Callaghan. 2012. Shrub-tree interactions and environmental changes drive treeline dynamics in the Subarctic. *Oikos*. doi:10.1111/j.1600-0706.2011.20032.x.
- Hanna, E., T. Jónsson, and J.E. Box. 2004. An analysis of Icelandic climate since the nineteenth century. *International Journal of Climatology* 24: 1193–1210. doi:10.1002/joc.1051.
- Heikkinen, R.K., M. Luoto, M.B. Araújo, R. Virkkala, W. Thuiller, and M.T. Sykes. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography* 30: 751–777.
- Hirzel, A., and A. Guisan. 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* 157: 331–341.
- Illán, J.G., D. Gutiérrez, and R.J. Wilson. 2010. The contributions of topoclimate and land cover to species distributions and abundance: Fine-resolution tests for a mountain butterfly fauna. *Global Ecology and Biogeography* 19: 159–173.
- Kadmon, R., F. Oren, and D. Avinoam. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications* 13: 853–867.
- Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14: 401–413.
- Kuhn, I., M. Brandenburg, and S. Klotz. 2004. Why do alien plant species that reproduce in natural habitats occur more frequently? *Diversity and Distributions* 10: 417–425.
- Lang, S.I., J.H.C. Cornelissen, A. Hölzer, C.J.F. Ter Braak, M. Ahrens, T.V. Callaghan, and R. Aerts. 2009. Determinants of cryptogam composition and diversity in Sphagnum-dominated peatlands: The importance of temporal, spatial and functional scales. *Journal of Ecology* 97: 299–310. doi:10.1111/j.1365-2745.2008.01472.x.
- Liston, G.E., and C.A. Hiemstra. 2010. Representing grass- and shrub-snow-atmosphere interactions in climate system models. *Journal of Climate* 24: 2061–2079. doi:10.1175/2010jcli4028.1.
- Luoto, M., and R.K. Heikkinen. 2008. Disregarding topographical heterogeneity biases species turnover assessments based on bioclimatic models. *Global Change Biology* 14: 483–494.
- Mac Nally, R. 2002. Multiple regression and inference in ecology and conservation biology: Further comments on identifying important predictor variables. *Biodiversity and Conservation* 11: 1397–1401.
- Menke, S.B., D.A. Holway, and R.N. Fisher. 2009. Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. *Global Ecology and Biogeography* 18: 50–64.
- Nagelkerke, N.J.D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78: 691–692.
- Pearson, R.G., W. Thuiller, M.B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, et al. 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography* 33: 1704–1711. doi:10.1111/j.1365-2699.2006.01460.x.
- Randin, C.F., R. Engler, S. Normand, M. Zappa, N.E. Zimmermann, P.B. Pearman, P. Vittoz, W. Thuiller, et al. 2009. Climate change and plant distribution: Local models predict high-elevation persistence. *Global Change Biology* 15: 1557–1569.
- Rangel, T.F.L.V.B., J.A.F. Diniz-Filho, and L.M. Bini. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography* 15: 321–327. doi:10.1111/j.1466-822X.2006.00237.x.
- Reddy, K.R., and R.D. DeLaune. 2008. *Biogeochemistry of wetlands: Science and applications*. Boca Raton: CRC Press.
- Rodhe, L., M. Pyykonen, and M. Krekula. 1999. Jordartskarta: Geological Survey of Sweden.

- Thuiller, W., L. Brotons, M.B. Araújo, and S. Lavorel. 2004. Effects of restricting environmental range of data to project current and future species distributions. *Ecography* 27: 165–172.
- Travis, J.M.J., R.W. Brooker, and E.J. Clark. 2006. The distribution of positive and negative species interactions across environmental gradients on a dual-lattice model. *Journal of Theoretical Biology* 241: 896–903.
- Trivedi, M.R., P.M. Berry, M.D. Morecroft, and T.P. Dawson. 2008. Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Global Change Biology* 14: 1089–1103.
- Tomas, P. 1998. Fjällvegetation, vektorformat för 30I (Abisko). Metria, Lantmäteriet GSD.
- Van Bogaert, R., K. Hanece, J. Hoogesteger, C. Jonasson, M.D. Dapper, and T.V. Callaghan. 2011. A century of tree line changes in sub-Arctic Sweden shows local and regional variability and only a minor influence of 20th century climate warming. *Journal of Biogeography* 38: 907–921.
- Walker, D.A., M.K. Raynolds, F.J.A. Daniëls, E. Einarsson, A. Elvebakk, W.A. Gould, A.E. Katenin, S. Kholod, et al. 2005. The circumpolar Arctic vegetation map. *Journal of Vegetation Science* 16: 267–282.
- Walsh, C., and R Mac Nally. 2003. The hier.part Package: Hierarchical Partitioning. (Part of: Documentation for R: A language and environment for statistical computing.)
- Wiens, J.A., D. Stralberg, D. Jongsomjit, C.A. Howell, and M.A. Snyder. 2009. Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences of the United States of America* 106: 19729–19736.
- Zhenlin, Y., E. Hanna, and T.V. Callaghan. 2011. Modelling surface-air-temperature variation over complex terrain around Abisko, Swedish Lapland: Uncertainties of measurements and models at different scales. *Geografiska Annaler: Series A, Physical Geography* 93: 89–112. doi:10.1111/j.1468-0459.2011.00005.x.
- Zhenlin, Y., E. Hanna, T.V. Callaghan, and C. Jonasson. 2012. How can meteorological observations and microclimate simulations improve understanding of 1913–2010 climate change around Abisko, Swedish Lapland? *Meteorological Applications*. doi:10.1002/met.276.

AUTHOR BIOGRAPHIES

Zhenlin Yang (✉) is a PhD student at Lund University and the University of Sheffield. The main research interests of Z. Y are the ecological modelling and climate impact assessment.
 Address: Department of Physical Geography and Ecosystem Science (ENES), Lund University, Sölvegatan 12, Lund, Sweden.
 e-mail: Zhenlin.Yang@nateko.lu.se

Martin T. Sykes is an experimental and theoretical ecologist with broad interests in plant ecology focused on modelling the responses of vegetation to climate, and on the maintenance of plant species diversity.
 Address: Department of Physical Geography and Ecosystem Science (ENES), Lund University, Sölvegatan 12, Lund, Sweden.
 e-mail: Martin.Sykes@nateko.lu.se

Edward Hanna is a Reader in Climate change at the University of Sheffield, UK. He specializes in climatology, applied/historical meteorology including instruments and observations, astronomy and planetary science.
 Address: Department of Geography, University of Sheffield, Winter Street, Sheffield S10 2TN, UK.
 e-mail: e.hanna@shef.ac.uk

Terry V. Callaghan is a Distinguished Research Professor at the Royal Swedish Academy of Sciences and Professor of Arctic Ecology at Universities of Sheffield, UK and Lund, Sweden. He specializes in arctic ecology, and climate and UV-B radiation impacts on arctic ecosystems.
 Address: Royal Swedish Academy of Sciences, Lilla Frescativägen 4 A, 114 18 Stockholm, Sweden.
 e-mail: terry_callaghan@btinternet.com