



High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords

Nathalie Isabelle Chardon, Jacob Nabe-Nielsen, Jakob Johan Assmann, Ida Bomholt Dyrholm Jacobsen, Maya Guéguen, Signe Normand, Sonja Wipf

► To cite this version:

Nathalie Isabelle Chardon, Jacob Nabe-Nielsen, Jakob Johan Assmann, Ida Bomholt Dyrholm Jacobsen, Maya Guéguen, et al.. High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords. *Diversity and Distributions*, 2022, 28 (5), pp.956-975. <10.1111/DDI.13498>. <hal-04890915>

HAL Id: hal-04890915

<https://hal.science/hal-04890915v1>

Submitted on 13 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

RESEARCH ARTICLE

High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords

Nathalie Isabelle Chardon^{1,2,3}  | Jacob Nabe-Nielsen⁴  | Jakob Johan Assmann³  |
Ida Bomholt Dyrholm Jacobsen⁵  | Maya Guéguen⁶ | Signe Normand³  | Sonja Wipf^{7,8} 

¹Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

²WSL Institute for Snow and Avalanche Research, Davos Dorf, Switzerland

³Department of Biology, Aarhus University, Aarhus C, Denmark

⁴Department of Ecoscience, Aarhus University, Roskilde, Denmark

⁵Greenland Institute of Natural Resources, Nuuk, Greenland

⁶Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Laboratoire d'Ecologie Alpine, Grenoble, France

⁷Swiss National Park, Chastè Planta-Wildenberg, Zerne, Switzerland

⁸Climate Change, Extremes and Natural Hazards in Alpine Regions Research Centre CERC, Davos Dorf, Switzerland

Correspondence

Nathalie Isabelle Chardon, Biodiversity Research Centre, University of British Columbia, 2212 Main Mall Vancouver, BC, Canada.
Email: nathalie.chardon@gmail.com

Funding information

Greenland Climate Research Center; Swiss Polar Institute; Swiss National Science Foundation, Grant/Award Number: P400PB_194331; Basler Stiftung für biologische Forschung; Arctic Research Centre at Aarhus University; European Commission, Grant/Award Number: 730938; Det Frie Forskningsråd, Grant/Award Number: 7027-00133B; American Alpine Club; WSL Director's Fund

Editor: Susan Jarvis

Abstract

Aim: Improving species distribution models (SDMs) and species abundance models (SAMs) of woody shrubs is critical for predicting biodiversity changes in the Arctic, which is experiencing especially high warming rates. Yet, it remains relatively unexplored if SDMs and SAMs can explain local scale patterns. We aim to identify predictor differences for the distribution versus abundance of two widespread Arctic shrub species with high resolution models and to compare validation approaches to assess the models' predictive abilities.

Location: Nuup Kangerlua (NK) and Kangerluarsunnguaq (K), two adjacent fjords in Southwest Greenland.

Methods: We conducted two separate field surveys in either fjord to construct high resolution (~90 m) SDMs and SAMs for *Betula nana* and *Salix glauca*, analysing the predictive influences of local scale climate, topography and soil moisture indicators. We then alternatively trained and validated models in either NK or K fjord and compared these results with the common split-sample validation approach. Finally, we assessed if including local scale biotic predictors improves SAM performance.

Results: Temperature extremes and precipitation best predicted the distributions of both species, whereas insolation and soil moisture indicators best predicted abundances. Compared to split-sample validation, both SDM and SAM performance was substantially reduced with separate survey validation. Regardless of validation approach, models performed poor to moderately well, and including local scale biotic parameters improved SAM performance.

Main conclusions: Substantial differences in model performance between validation approaches highlight the usefulness of using a separate survey for validating model predictive performance. We discuss various factors that might have caused poor model performance, such as not capturing all relevant predictors or enough local scale heterogeneity in predictor or response variables. We emphasise the need to include predictors relevant at the spatial scale of study, such as local scale biotic interactions, for improved predictions at high spatial resolution.

Signe Normand and Sonja Wipf should be considered as joint senior authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

KEYWORDS

Arctic tundra, field survey methods, model validation, shrubs, spatial scale, species abundance model, species distribution model

1 | INTRODUCTION

With climate change-induced shifts in species geographic distributions observed in nearly every biome (Lovejoy & Hannah, 2019), spatial predictive models have become a dominant topic in ecological and conservation research (Araújo et al., 2019; García-Díaz et al., 2019; Pacifici et al., 2015). These models are important tools in understanding what factors influence species distributions (Smith & Santos, 2020) and assessing the effect of climate change on ecosystem composition and biodiversity loss (Bellard et al., 2012). This is particularly important in the Arctic, where exceptionally high warming rates (IPCC, 2014) are causing widespread, yet highly variable, changes in shrub distributions and abundances (Myers-Smith et al., 2020; Post et al., 2019). This 'shrubification' phenomenon has ecosystem-wide implications, as taller shrubs can dominate over smaller plants (Elmendorf, Henry, Hollister, Björk, Boulanger-Lapointe, et al., 2012), as well as influence soil moisture, temperature, and snow cover (Lawrence & Swenson, 2011; Liston et al., 2002; Sturm & Douglas, 2005). At a global scale, shifts in Arctic shrub distributions and abundances are expected to have a positive feedback on climatic warming (Pearson et al., 2013). However, our inability to accurately predict local scale distributional and abundance trends (e.g., Elith & Leathwick, 2009) hinders us from drawing robust conclusions on large-scale biodiversity changes in the Arctic as a result of shrubification (Myers-Smith et al., 2011).

Predicting where species ranges will expand or contract is most commonly done by constructing correlative species distribution models (SDMs; Guisan & Thuiller, 2005; Guisan & Zimmermann, 2000; Pacifici et al., 2015). These relate a species' occurrences with the environment in these regions to project the probability of occurrence across a landscape of interest, under current or future environmental scenarios (Booth & McMurtrie, 1988; Guisan et al., 2017; Wiens et al., 2009). Another approach is to construct species abundance models (SAMs) to understand how a species' abundance varies with environmental factors. Compared to SDMs, SAMs are expected to provide better predictions of trends in species performance under climate change (Morris & Ehrlén, 2015) and these predictions may be mismatched with those from SDMs (e.g., Jarnevich et al., 2021). However, specifically assessing differences in SDM versus SAM predictions has, to our knowledge, not been carried out in the Arctic. Regardless of model type, it remains contentious to what degree model outputs are related to empirical measures of population performance, such as population persistence (Araújo & Williams, 2000), abundance (Dallas & Hastings, 2018; Midolo et al., 2021; Weber et al., 2017), and demographic rates (Csörgő et al., 2017; Thuiller et al., 2014). Furthermore, examples of SDMs incorrectly predicting observed distributional trends (e.g., Sofaer et al., 2018) point to the more fundamental questions of how well these

models can accurately capture spatial patterns across the landscape and how to best validate this (e.g., Santini et al., 2021).

Given the highly variable response of Arctic shrubs to climate change (Elmendorf, Henry, Hollister, Björk, Bjorkman, et al., 2012; Myers-Smith et al., 2015, 2020), caused by local scale deterministic (e.g. moisture; Ackerman et al., 2017; Myers-Smith et al., 2015) and stochastic variation (e.g., disturbances; Prendin et al., 2020), it is difficult to predict shrub species' spatial patterns from broad-scale climate factors alone. What further complicates this is that predictors of occurrence may not be the best predictors of abundance for plant and animal species (e.g., Birch, 1953; Nielsen et al., 2005; Randin et al., 2006). Compared to occurrence, abundance might be best explained by factors difficult to measure, such as disturbance history (*sensu* Jarnevich et al., 2021). However, differences between the apparent drivers of species occurrence versus abundance may also arise as a function of the selected spatial scale of analysis, the available predictors, and other features of the study design. Furthermore, if there is a time lag in species establishment, then this might result in differences in the factors influencing occurrence versus abundance even when none exist.

Studies comparing different modelling resolution scales have observed that grid size can strongly impact conclusions on spatial patterning (e.g., Foxcroft et al., 2009) and projected amount of future suitable habitat (e.g., Gillingham et al., 2012), but we are not aware of studies that have compared how differences in predictors of occurrence versus abundance might change across scales. Observational studies that examine factors influencing both occurrence and abundance do not generally focus on these differences (e.g., Guarino et al., 2012; Lou et al., 2018) or do not find strong empirical trends (e.g., Nielsen et al., 2005; Økland, 1995). To our knowledge, such studies are lacking for the Arctic, except for recent work by Becher et al. (2018) in the Swedish Arctic showing that cryogenic (i.e. freeze-thaw) disturbances are associated with changes in the occurrences of plant species, but that their abundances are dependent on initial plant community composition.

A related, and long-standing, problem is the common mismatch between underlying mechanisms and modelling scale (Levin, 1992). SDMs and SAMs are often parameterised with large-scale regional climatic data (Fridley, 2009; Scherrer & Körner, 2010), but the substantial mismatch between regional versus microhabitat climate fails to capture the importance of microclimate in influencing range shifts (Franklin et al., 2013; Oldfather et al., 2020). However, there have been significant advances in modelling at higher resolutions within smaller geographic regions, illustrating this as a promising approach (D'Amen et al., 2015; Thuiller et al., 2015; Tomlinson et al., 2020). A mismatch can also exist between the factors relevant for species distributions and abundances, such as biotic interactions, and those included as predictors (Boulangeat et al., 2012; Mod & Luoto, 2016;

Wisn et al., 2013). While several frameworks for the inclusion of biotic interactions in spatial models have been proposed (e.g., Kissling et al., 2012; Staniczenko et al., 2017; Wisn et al., 2013), more work is needed to implement these in the Arctic (see e.g., Mod & Luoto, 2016).

Properly assessing model performance, and specifically the need to use spatially independent validation data, is a decades-long issue (Fielding & Bell, 1997). Randomly partitioning a dataset into training and validation sets has long been criticised (Chatfield, 1995) and often yields overly optimistic model performance (Roberts et al., 2017; Schratz et al., 2019). Yet, this 'split-sample' validation approach remains widely used in SDM studies (reviewed in Araújo et al., 2005; Chardon et al., 2020; Santini et al., 2021). Given the plethora of work examining how validation approaches affect model predictive ability (e.g., Araújo et al., 2005; Morrison et al., 1987; Raxworthy et al., 2003; Roberts et al., 2017), it is surprising that models are rarely validated with independent datasets collected within the training spatial region (Lee-Yaw et al., 2021; but see Angert et al., 2018; Elith et al., 2006), such as those from different survey designs. This is especially relevant in the Arctic, where many of the existing data have been collected with varying survey methods (Walker et al., 2016).

In this study, we use high resolution SDMs and SAMs to examine the distribution and abundance of two deciduous shrub species, *Betula nana* L. (dwarf birch, Betulaceae) and *Salix glauca* L. (grey willow, Salicaceae). Both species are widely distributed across the Arctic, and they are the most dominant species in the heaths of western Greenland (Fredskild, 1996; Rune, 2011), making them well-suited for this study. We specifically ask:

1. Do different environmental variables best predict the distributions versus abundances of *B. nana* and *S. glauca*?
2. Can SDMs and SAMs trained in one fjord predict distributions and abundances, respectively, in an adjacent fjord, as validated with separate survey data?
3. Does including biotic predictors improve SAM model performance?

To answer these questions, we first conducted separate field surveys in two adjacent fjords in Southwest Greenland, during which we compiled detailed datasets on shrub presence-absence, abundance, species composition, and canopy height. We started with a stratified systematic survey in one fjord and then used the resulting data to construct a preliminary SAM for a stratified random survey in the second fjord. Second, we used the data from each fjord to train high resolution (~90 m) SDMs and SAMs to assess the influence of local scale environmental predictors on *B. nana* and *S. glauca* distribution versus abundance within the fjord. We chose predictors based on their known influence on Arctic shrub distributions (Bjorkman et al., 2018; Elmendorf, Henry, Hollister, Björk, Bjorkman, et al., 2012; Gamm et al., 2018; Myers-Smith et al., 2015; Nabe-Nielsen et al., 2017; von Oppen et al., 2021), thus focusing on temperature, precipitation, and indicators of soil moisture. Third, we validated the models in the fjord not used in model training, making

ours one of few studies to compare separate survey and split-sample validation approaches. Finally, we included local scale biotic predictors representing plant-plant interactions in our models to assess their impact on model performance. Given the potential discrepancy between SDM and SAM outputs (e.g., Jarnevich et al., 2021), we explore the performance differences of both model types and validation approaches. These comparisons are relevant, for example, in selecting the best approach for upscaling model results for conservation purposes.

2 | METHODS

2.1 | Study area

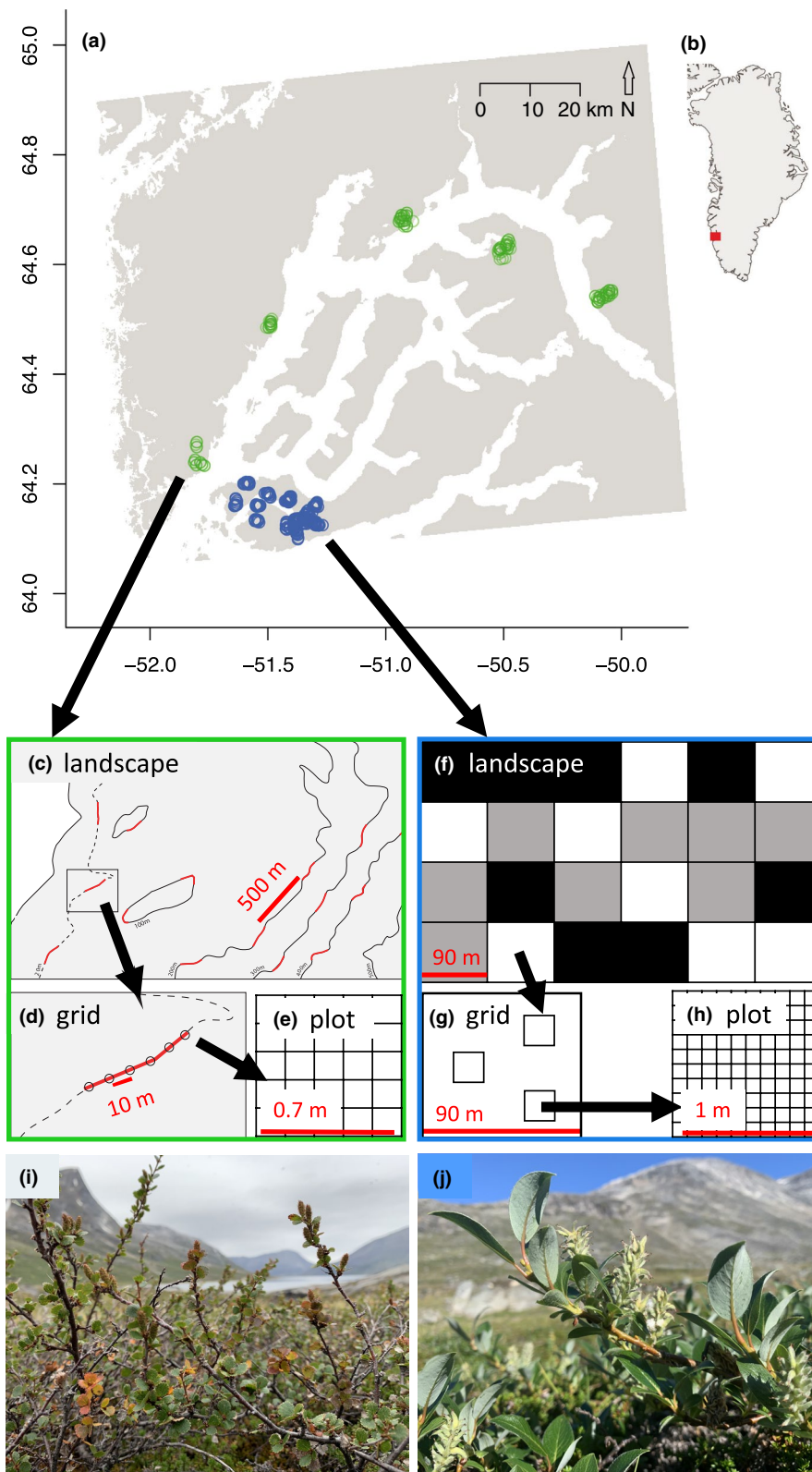
We chose a ~119 km × 97 km topographically complex study area in Southwest Greenland with feasible terrain access to encompass the Nuup Kangerlua (henceforth, NK; *Godthåbsfjord*) and Kangerluarsunnguag (henceforth, K; *Kobbefjord*) fjords with a general buffer of undefined width (Figure 1). The 160 km long NK and 16 km long K fjords partially overlap in their environmental spaces (see Figures S1.1, S1.2 in Appendix S1 in Supporting Information) despite the large differences in fjord length. One-third (18,800) of Greenland's population lives in the city of Nuuk (Grønlands statistik, 2021) at the entrance to these fjords, which are used recreationally for hiking and tourism (Elvarsdóttir, 2013).

2.2 | Study species

Our focal species, *Betula nana* and *Salix glauca* (see Figure 1i,j), are common and widely distributed deciduous shrub species in the Arctic. *B. nana* is a monoecious and clonal shrub distributed widely across the Northern hemisphere (Global Biodiversity Facility, 2020). In Greenland, it has a discontinuous distribution on both the west and east coasts, with no occurrences in the South (Böcher et al., 1978; Feilberg, 1984; Halliday, 2019). *S. glauca* is a low-arctic dioecious shrub with a circumboreal-arctic distribution (Elven et al., 2020; Feilberg, 1984), occurring abundantly along both west and east coasts.

Literature on Arctic shrub ecology and distribution patterns (e.g., Böcher, 2000; Born & Böcher, 2000) indicates that *B. nana* has a continental preference, possibly relating to precipitation and temperature amplitude, whereas *S. glauca* has very broad abiotic preferences. However, the role of soil moisture as a determinant of the distributions of these, and other Arctic shrubs, is increasingly acknowledged (Ackerman et al., 2017; Martin et al., 2017; Myers-Smith et al., 2015), as is that of microtopography (Scherrer & Körner, 2010; Stewart et al., 2016). Mycorrhizal fungi are also essential to certain species distributions in the Arctic, such as the obligate relationship between *Leccinum* Gray (Boletaceae) and *B. nana* (Born & Böcher, 2000). As an indicator species of low-arctic continental heaths, *B. nana* should prefer relatively warmer summers, colder winters, lower precipitation and acidic soils (Böcher, 2000), but this preference is not reflected

FIGURE 1 We conducted plant surveys in the Nuup Kangerlua (NK) (a; green points) and Kangerluarsunnguaq (K) (a; blue points) fjords in Southwest Greenland (b: red rectangle). At each of five sites in NK fjord, we selected three plot groups (c; red lines) for every 100 m increase in altitude, located 500 m apart on the same isocline (dashed lines). Within plot groups, we sampled in plots 10 m apart (d; black circles) using 0.7 m $\times$ 0.7 m pin-point frames with 25 grid intersects (e). In K fjord, we randomly selected spatial grids to survey in three predicted abundance classes (f; represented by grey scale) generated from a preliminary *Betula nana* SAM (see Figure S1.3). Within each grid, we chose (g) three randomly placed plots to (h) sample using a 1 m $\times$ 1 m quadrat with delineations of 1%. In both surveys, we collected data on (i) *Betula nana* and (j) *Salix glauca* as well as the overall shrub community. Maps in (a, b) are in Polar Stereographic (EPSG 4326) projection, and we revised (c–e) from Figure 2 in Nabe-Nielsen et al. (2017)



in the species' Greenlandic distribution (Bay, 1999; Fredskild, 1996). *S. glauca* is highly morphologically variable and is associated with a variety of shrub species compositional types. The species' only commonality is that it requires favourably warm summer conditions and moist soils, and suitable soil type ranges from acidic to neutral and general climate from coastal to inland (Böcher, 2000).

2.3 | Plant surveys

We conducted two separate plant surveys with different sampling designs in NK and K fjords to use in both model training and validation. In NK fjord, we conducted a stratified systematic plant survey along the length of the fjord over three northern hemisphere summer

seasons (July–August) in 2011–2013 (Figure 1; following Nabe-Nielsen et al., 2017). At five distinct sites, we sampled along elevational gradients to collect data on presences, absences, abundance and species composition of all shrub species using a 0.7 m × 0.7 m pin-point frame (Figure 1e). For model training, we converted these pin-point data into percent cover estimates based on the number of pins dropped ($n = 25$) for abundance data, and binarised abundance data for presence–absence data. To estimate how much information we lost with this binarisation, we calculated the ratio of Shannon entropy for abundance to presence data. We averaged the abundance data at the ~90 m grid size to match our analysis scale (see next section) for a total of 119 spatio-climatic grids encompassing the plot locations (for details see Appendix S2).

In K fjord, we conducted a random stratified plant survey in Summer (August) 2020. We used a preliminary SAM trained with summed pin counts of *B. nana* in NK fjord (see Figure S1.3) to stratify the ~27 km × 17 km fjord landscape into predicted low, medium and high abundances classes. We randomly selected ~90 m × 90 m spatio-climatic grids to survey in each class for a total of 200 grids, ensuring that they were accessible by foot or boat (for details see Appendix S2). Within each grid, we sampled three 1 m² quadrats arranged in a randomly rotated equilateral triangle centred on the mid-point of the cell. We used a gridded sampling quadrat with 1% delineations (Figure 1h) to record shrub species presences, absences and composition, estimated percent cover and measured maximum shrub species vegetative height. We calculated Shannon entropy ratios as above. At every plot, we also visually scanned the area in a 20 m radius from the plot and recorded the presence of any additional shrub species to estimate grid-level woody species richness. As in NK fjord, we averaged these data at the grid level (for details see Appendix S2).

2.4 | Predictor variables

We computed high resolution temperature, precipitation and solar radiation at local scales for the study area by statistically downscaling a climate time series (1982–2013) from the monthly CHLSA V1.2 data. CHLSA (Climatologies at High resolution for the Earth's Land Surface Areas) is a global temperature and precipitation climatology dataset downscaled to 30 arc sec (~400 m at the latitude of our study) (Karger et al., 2017, 2018). We downscaled these data to our target grid size of ~90 m with geographic weighted regression and using the MEaSUREs Greenland Ice Mapping Project (GIMP) Digital Elevation Model (DEM) v. 1 (Howat, Negrete, & Smith, 2014, 2015). We then calculated 30-year averages of the climate parameters: average summer (June–August) maximum temperature, yearly maximum temperature, yearly minimum temperature, temperature continentality (yearly max.–min. temperatures), cumulative spring (March–May) precipitation, cumulative summer precipitation and average summer incident solar radiation (henceforth, insolation). For calculation details see Appendices S2, S3 and Appendix S2 in von Oppen et al. (2021).

In addition to the climate data, we calculated several topographic and terrain wetness indices at a local scale. We derived slope, aspect and the SAGA wetness index (hereafter TWI; Boehner et al., 2002; Boehner & Selige, 2006) from the GIMP DEM. TWI is an inferred measure of how 'wet' an area is, based on water drainage from the surrounding landscape. We also calculated the tasseled cap wetness component (hereafter TCW; Crist & Ciccone, 1984) from satellite images (for details see Appendices S2 and S3) as an alternative measure of wetness (see Appendix S2 for data references, all of which are publicly available). We generated raster layers for each of these predictor variables for use in prediction mapping and extracted point values for each survey location for model training and validation.

To assess if biotic predictors improve model performance, we calculated local scale biotic variables from the plant survey data collected in NK and K fjords. We calculated shrub species richness, diversity and competition (i.e. sum of non-*B. nana* or non-*S. glauca* percent cover) with the grid-level averages from the survey data. In K fjord, we also calculated canopy height as the community weighted mean (by abundance) of maximum vegetative shrub height with the grid-level averages from the survey data, again excluding the target species to generate non-circular predictors.

2.5 | Model training and prediction

We constructed SDMs and SAMs with two contrasting modelling techniques, boosted regression trees (BRTs) and generalised linear models (GLMs). We fit the machine learning algorithm BRT (Freund, 1995; Freund & Schapire, 1997; Friedman et al., 2000; Schapire, 1990) and identified the most influential environmental predictors described above (see Table S1.1a) on *B. nana* and *S. glauca* distribution and abundance. We tuned the BRT hyperparameters learning rate and bagging fraction (see Table S1.2) following Elith et al. (2008). We fit the logistic regression GLM (McCullagh & Nelder, 1989) and identified the most parsimonious model containing uncorrelated (Pearson's correlation coefficient <0.5) variables with Akaike's information criterion (AIC; Akaike, 1973, 1974) for each species' distribution and abundance by fjord (see Appendix S3 for details). For each modelling technique, we generated four high resolution (~90 m) SDMs and SAMs per species for our study region, training models in either fjord. We tested spatial autocorrelation in the residuals with Moran's *I* in all models, ensured that the variance inflation factor for each GLM < 4, and checked GLM residual assumptions following Zuur et al. (2010).

The two modelling techniques we used have different advantages and disadvantages. The non-parametric BRT optimises predictive performance with a 'boosting' technique by fitting multiple simple regression trees in an ensemble modelling approach (Freund & Schapire, 1997). It makes no data assumptions and can fit complex nonlinear relationships (Elith et al., 2008). Because GLMs are parametric and are further constrained by the link function, they produce smoother response curves than BRTs (Elith & Graham, 2009). When extrapolating outside the training environmental space, BRTs predict

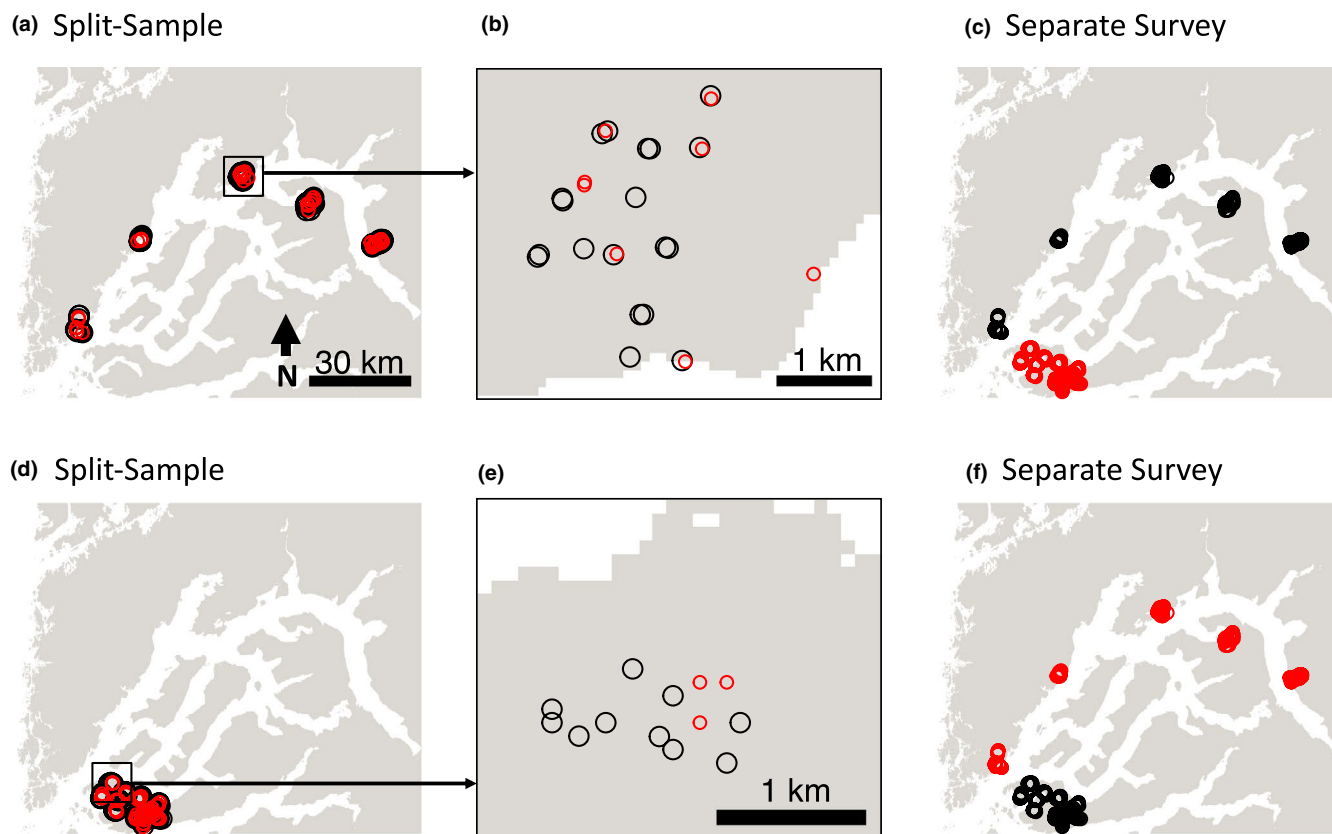


FIGURE 2 We validated our *Betula nana* and *Salix glauca* species distribution models (SDMs) and species abundance models (SAMs) with split-sample (a, d) and separate survey (c, f) approaches. A detailed example view of the split-sample validation is shown for Nuup Kangerlua (b) and Kangerluarsunnguaq (e) as indicated in (a) and (d), respectively. Shown are training (black) and validation data (red) on a background map in Polar Stereographic (EPSG 4326) projection. See Figures. S1.2, S1.4, S1.5 for comparisons of the environmental predictor space for each validation method

a constant value from the last observed environmental condition, whereas GLMs continue the fitted trend into the new environmental space (Elith & Graham, 2009). Both algorithms can perform well in new environmental spaces (reviewed in Yates et al., 2018).

We report all additional modelling details in Appendix S2, Table S1.2 and provide an ODMAP Protocol in Appendix S3 (following Zurell et al., 2020) for standardised and reproducible model documentation (Fitzpatrick et al., 2021). Unless otherwise noted, all these and subsequent data processing and analyses were conducted with R v. 4.0.2 and 4.1.1 (R Core Team, 2021). We specify R functions and packages in Appendix S2. The data corresponding to these analyses are available in Appendices S4 and S5 and stored in the public repository EnviDat (Chardon et al., 2022). Our computer code is available on GitLab (https://gitlab.com/nathaliechardon/gl_microclim).

2.6 | Model validation

We validated predictions from the final models with (i) five repetitions of random split-samples of the data used to train the model (retaining $\frac{1}{3}$ of the data for validation) and (ii) separate survey data

from the fjord not used in model training (Figure 2). To account for differences in the environmental predictor spaces between the two fjords (see Figure S1.1), we also validated models in the adjacent fjord using only the data within the same environmental space as the training data.

We compared validation approaches with standard statistics of the true skill statistic ($TSS = \text{sensitivity} + \text{specificity} - 1$; Allouche et al., 2006) and the area under the receiver operating characteristic curve (AUC; DeLeo, 1993; Elith et al., 2002; Zweig & Campbell, 1993) for SDMs, and correlation between observed and predicted values for SAMs (following Potts & Elith, 2006). To binarise SDM output, which predicts probability of occurrence, we chose the default presence-absence threshold (0.5) to facilitate comparisons of TSS, sensitivity (correctly predicted presences), and specificity (correctly predicted absences) between models (see Appendix S2 for details). This threshold also maximised sensitivity, thus reflecting our greater confidence in spotting species presences over absences. Eight other thresholds all gave the same qualitative results for BRTs (see Table S1.3). We calibrated SDM predictions following Dormann (2020), and we report these results in Table S1.4 due to only slight differences in model performance results compared to uncalibrated predictions.

2.7 | Including biotic predictors

To assess whether local scale biotic predictors improve model performance, we re-trained the final SAMs as described above in either NK or K fjord and added one of: richness, diversity, competition, or canopy height (see Table S1.1b). We calculated competition and canopy height for each focal species separately, excluding either values of *B. nana* or *S. glauca* in the predictor calculations used in that species' SAM. We added one variable at a time to assess if certain predictors play a greater role in improving model performance. We quantified model performance for each of the four resulting models with five repetitions of split-sample validation. We did not test if adding these biotic predictors improved the separate survey validation approach, because these biotic variables were collected with too different of methods in either fjord for such a prediction to be useful.

3 | RESULTS

In general, SDMs performed better than SAMs, GLMs outperformed BRTs, and models for *B. nana* were better than for *S. glauca* (Table 1). In almost all cases, the split-sample validation approach yielded slightly to substantially higher performance metrics than the separate survey validation approach. Split-sample validation of our models produced moderate to good *B. nana* SDMs (TSS 0.27–0.76; AUC 0.69–0.94) and weak SAMs (correlation 0.38–0.57). Calibrated SDM predictions yielded very similar results

(TSS 0.38–0.71; AUC 0.72–0.91; Table S1.4). Split-sample validation produced poor to moderate *S. glauca* SDMs (TSS 0.01–0.20; AUC 0.58–0.78) and weak SAMs (correlation 0.09–0.37) (Table 1). Calibrated SDM predictions again yielded very similar results (TSS 0.08–0.36; AUC 0.57–0.79; Table S1.4). In all cases except the GLM SDM and BRT SAM for *S. glauca* (both trained in NK fjord), separate survey validation resulted in slightly to substantially poorer SDMs (differences in TSS 0.01–0.76; AUC –0.08 to 0.40) and SAMs (differences in correlation –0.04 to –0.51) for both species (Table 1).

With the split-validation approach, models were trained and validated in almost identical environmental spaces (see Figures S1.4 and S1.5). With the separate survey validation approach, models were trained and validated in only partially overlapping environmental spaces (see Figures S1.1 and S1.2). Validating models within the subset of overlapping environmental space from the adjacent fjord produced similar results (Table S1.5), with the split-validation approach still far superior for all but the BRT and GLM *S. glauca* SAMs trained in NK fjord (differences in TSS 0.01–0.76; AUC 0.01–0.30; cor –0.09 to 0.44). A Moran's I test of model residuals showed no significant spatial autocorrelation in any of the models, except in both of the *B. nana* BRT SDMs trained in K fjord ($p < .001$) so we were cautious in comparing these model results. Our visual assessment of residual distribution plots showed that GLM residual assumptions were moderately well met. Binarising abundance data to presence–absence data yielded a substantial loss in information, with high Shannon entropy ratios of abundance to presence–absence data (*B. nana* in NK: 4.8; *S. glauca* in

TABLE 1 Validation results of *Betula nana* and *Salix glauca* species distribution models (a: SDMs) or species abundance models (b: SAMs) trained in Nuup Kangerlua (NK) or Kangerluarsunnguag (K) fjords and validated in the same fjord (split-sample) or the adjacent fjord (separate survey)

Split-sample							Separate survey				
(a) Train	Species	Validation	BRT		GLM		Validation	BRT		GLM	
			TSS (SD)	AUC (SD)	TSS (SD)	AUC (SD)		TSS	AUC	TSS	AUC
NK	<i>B. nana</i>	NK	0.27 (0.13)	0.74 (0.08)	0.33 (0.07)	0.69 (0.07)	K	−0.12	0.54	−0.02	0.66
NK	<i>S. glauca</i>	NK	0.09 (0.19)	0.58 (0.07)	0.20 (0.05)	0.59 (0.04)	K	−0.02	0.53	0	0.67
K	<i>B. nana</i>	K	0.76 (0.06)	0.94 (0.03)	0.64 (0.07)	0.90 (0.02)	NK	0	0.54	0	0.71
K	<i>S. glauca</i>	K	0.01 (0.02)	0.78 (0.06)	0.14 (0.18)	0.78 (0.03)	NK	0	0.63	0	0.58
Split-sample							Separate survey				
(b) Train	Species	Validation	BRT		GLM		Validation	BRT		GLM	
			Correlation (SD)		Correlation (SD)			Correlation		Correlation	
NK	<i>B. nana</i>	NK	0.57 (0.09)		0.50 (0.10)		K	0.06		0.2	
NK	<i>S. glauca</i>	NK	0.09 (0.13)		0.26 (0.9)		K	0.13		0.18	
K	<i>B. nana</i>	K	0.38 (0.15)		0.43 (0.4)		NK	0.34		0.41	
K	<i>S. glauca</i>	K	0.37 (0.07)		0.36 (0.03)		NK	0.19		0.12	

Note: We validated SDMs with the true skill statistic (TSS) and the area under the receiving operator characteristic Curve (AUC). We validated SAMs with Pearson's correlation coefficient calculated between the observed and predicted values. Validation is either with a random split-sample (average of five repetitions of ½ validation holdout with standard deviation (SD) reported) or separate survey approach. We based TSS calculations on the default threshold of 0.5. The higher performance metric(s) of the two validation approaches are in bold.

NK: 3.8; *B. nana* in K: 3.2; *S. glauca* in NK: 6.1), and we considered this when comparing SDM to SAM results.

3.1 | Different predictors of distributions vs. abundances

In all but one case for BRTs, different local scale environmental parameters best predicted *B. nana* and *S. glauca* distribution versus abundance (Table 2a). Temperature extremes and precipitation best predicted the distributions of both species, whereas insolation and soil moisture indicators best predicted abundance. However, a similar suite of environmental parameters in BRTs predicted distributions and abundances in both fjords, with many of the same predictors seen as influential in each. GLMs also used a similar predictor set for distributions versus abundances, and all but one predictor (maximum yearly temperature) used in GLMs was also used in final BRTs. The most parsimonious GLMs fit far fewer predictors ($n = 1\text{--}2$) than the BRTs ($n = 4\text{--}10$) (Table 2b). The relationships between environment and distribution or abundance depended on the training fjord and modelling technique, with BRTs showing less smooth response curves than GLMs (see Figures S1.6 and S1.7).

The BRTs showed yearly minimum temperature and spring precipitation as being the most important predictors of both *B. nana* and *S. glauca* distributions with the highest relative variable contributions in both models. Summer insolation or spring precipitation best predicted *B. nana* abundance in NK or K fjords, respectively. TCW best predicted *S. glauca* abundance, regardless of fjord. Out of all variables tested, only summer temperature and TWI were important for distribution, but not abundance. The most parsimonious GLMs also showed spring precipitation as being an important predictor for distribution and abundance, although not always in both fjords. In NK fjord, temperature continentality and slope best predicted *B. nana* distribution, and spring precipitation and summer insolation best predicted *B. nana* abundance. In K fjord, maximum yearly temperature and spring precipitation best predicted both *B. nana* distribution and abundance. In NK fjord, temperature continentality best predicted *S. glauca* distribution, and TCW and slope best predicted *S. glauca* abundance. In K fjord, spring precipitation best predicted both *S. glauca* distribution and abundance, with summer insolation being an additional best predictor of abundance.

3.2 | Separate survey validation yields lower model performance than split-sample validation

Compared to the split-sample validation approach, both BRTs and GLMs validated with separate survey data poorly predicted distribution and abundance in the fjord not used in model training (Table 1). The difference in performance metrics between the validation approaches was especially high for SDMs, which showed substantially

lower TSS (difference up to 0.76) and lower AUC (difference up to 0.40) values with the separate survey validation approach.

For both the BRT and GLM SAMs, the overall differences between validation approaches were lower than for SDMs, with only half of the models showing substantial differences in performance metrics. This difference was highest for the *B. nana* BRT SAM trained with the stratified systematic plant survey data in NK fjord (correlation difference 0.51), the *S. glauca* GLM SAM trained with the stratified random data from K fjord (0.24), and the *S. glauca* BRT SAM trained in NK fjord (0.18). The other two SAMs (*S. glauca* trained in NK and *B. nana* trained in K) showed only a slight difference in model performance (−0.04 to 0.08).

The *S. glauca* BRT SDM trained in NK fjord and validated in K fjord had good sensitivity and the other three SDMs had better specificity (Figure 3). SDMs trained in K fjord and validated in NK fjord yielded only presences when binarised with a threshold, and the *B. nana* SDMs had slightly higher sensitivity than the *S. glauca* SDMs (see Figure S1.8). SAMs generally overpredicted abundance and predicted a much smaller range of values than observed (Figure 4). SAM performance only improved in two cases when evaluating predictions at correctly predicted presences by the species' corresponding SDM.

3.3 | Including biotic predictors improves SAM performance

Including local scale biotic predictors at the shrub level (species richness, Shannon diversity, competition and canopy height) marginally to substantially improved both BRT and GLM SAM performance in 18 out of 32 cases (Table 3). Canopy height, competition, and diversity generally had the largest effects on model performance, whereas richness yielded no to slight increases. In 14 cases, adding a biotic predictor had no, or a negative, impact on model performance.

4 | DISCUSSION

We constructed two sets of high resolution SDMs and SAMs with BRTs and GLMs for *Betula nana* and *Salix glauca*, two widespread Arctic shrub species. We used data from NK fjord or the adjacent K fjord in turn to train the models and validated these with data from the other fjord. We found substantial model performance differences between the split-sample and separate survey validation approaches, illustrating that split-sample validation can overestimate performance metrics, and that separate survey approaches are valuable. Due to poor model performance with separate survey validation, especially for SAMs, we are cautious in our interpretations of which local scale factors best predict the distribution versus abundance of either species. Adding local scale biotic predictors improved model performance in most cases, illustrating the usefulness of including biotic predictors in predictions of species distributions and abundances.

TABLE 2 (a) Boosted regression tree relative percent variable contributions and (b) generalized linear model coefficient estimates (and standard errors) with *p*-values in final *Betula nana* and *Salix glauca* species distribution models (SDMs) and species abundance models (SAMs) trained in Nuup Kangerlua (NK) or Kangerluarsunnguak (K) fjords, chosen from the full parameter set examined (see Table S1.1a)

(a) Training data	<i>B. nana</i>				<i>S. glauca</i>			
	SDM		SAM		SDM		SAM	
	NK (%)	K (%)	NK (%)	K (%)	NK (%)	K (%)	NK (%)	K (%)
JJA temp ^a		2.8				11.9		
Max temp ^a								
Min temp ^a	41.4	3.9	17.8		48.8			
Temp cont ^a	16.6	6.1	9.3	25.5	8.8	2.8	21.0	11.4
JJA precip ^a		4.1				15.3		11.3
MAM precip ^a	16.7	64.6	21.1	55.5		44.8		19.3
JJA insol ^a	8.5	4.1	26.6			2.4	17.5	18.1
TWI ^a		3.3			4.7	4.1		
TCW ^a		4.0	15.5	19.0	19.2	7.7	41.0	21.6
Slope	16.8	4.4				7.0	20.5	13.3
Aspect		3.2	9.8		18.6	4.5		5.0
(b) Training data	<i>B. nana</i>				<i>S. glauca</i>			
	SDM		SAM		SDM		SAM	
	NK	K	NK	K	NK	K	NK	K
JJA temp ^a								
Max temp ^a		0.33 (0.14)*		0.94 (0.35)**				
Min temp ^a								
Temp cont ^a	0.57 (0.13)***				0.23* (0.09)			
JJA precip ^a								
MAM precip ^a		-0.53 (0.07)***	-0.50 (0.10)***	-0.59 (0.10)***		-0.2 (0.04)***		-0.44 (0.10)***
JJA insol ^a			0.63 (0.15)***					0.27 (0.07)***
TWI ^a								
TCW ^a							0.004 (0.002)*	
Slope	-0.09 (0.03)**						0.2 (0.13)*	
Aspect								

p-Values should only be interpreted as informational because we used a variable selection approach to identify best models. Note that climate variables are correlated with elevation and proximity to the fjord mouth, with spring (MAM) precipitation increasing with elevation and proximity to the fjord mouth. Variables with the highest relative contributions (a) are in bold.

p*: .001 (*), .001 (**), .01 (*).

^aJune–August average maximum temperature (JJA temp), yearly maximum temperature (Max temp), yearly minimum temperature (Min temp), temperature continentality (Temp cont = yearly maximum – minimum temperatures), June–August precipitation (JJA precip), March–May precipitation (MAM precip), June–August average insolation (JJA insol), SAGA Wetness Index (TWI), Tasseled Cap Wetness (TCW).

We conclude that using abundance data in our study system does not improve predictive power, even though SAMs are a useful tool in understanding how species respond to environmental changes (Morris & Ehrlén, 2015) and can generate good predictions within overlapping training and validation environments (Waldock et al., 2022). In our system, the information content was

substantially reduced when binarising the abundance data for the SDMs. The SDMs were therefore tasked with predicting simpler binary presence–absence data compared to the SAMs, which were tasked to predict the more complex abundance data. The increased complexity in the abundance data likely contributed to the reduced performance of the SAMs compared to the SDMs.

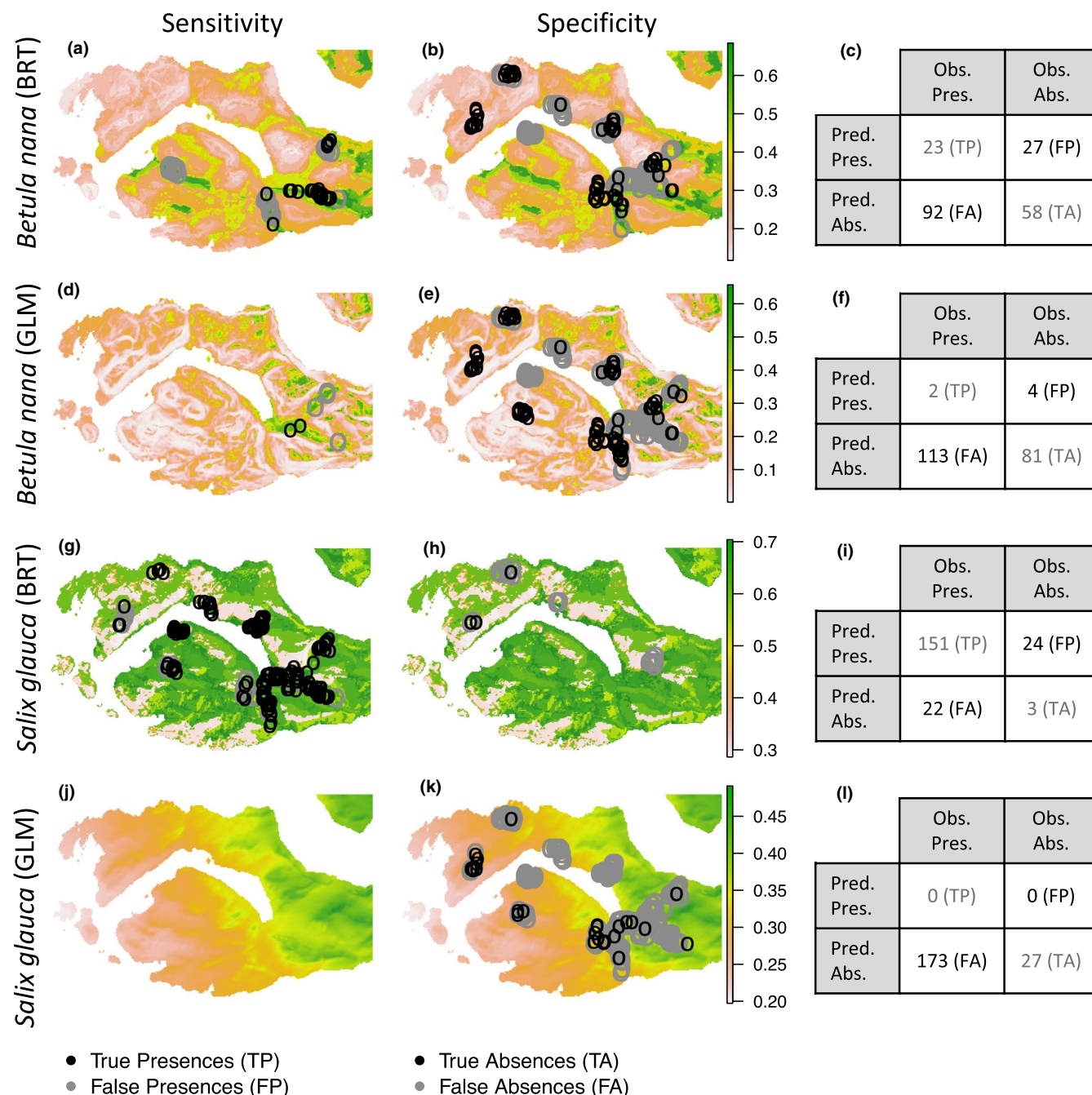


FIGURE 3 Sensitivity and specificity maps of *Betula nana* (a, b, d, e) and *Salix glauca* (g, h, j, k) boosted regression tree (BRT) and generalised linear model (GLM) species distribution models (SDMs) trained in Nuup Kangerlua (NK) fjord and validated in Kangerluarsunnguaq (K) fjord. As seen in the corresponding confusion matrices (c, f, i, l), the *S. glauca* BRT SDM was the only model that correctly predicted a substantial amount of true presences, whereas the other three models better predicted true absences, as based on a threshold of 0.5. SDMs trained in K fjord and validated in NK fjord yielded only presences after thresholding, and results are shown in Figure S1.8. Background maps are the predicted probabilities of occurrence (shown in colour) from the corresponding SDM in Polar Stereographic (EPSG = 4326) projection. Legends for (a, d, g) and (b, e, h) are as in (j) and (k), respectively, and suitability values correspond to both horizontal panels

4.1 | Seasonal precipitation and climatic extremes best predict distributions

Our SDMs, which showed reasonable AUC values with split-sample validation, indicate that *B. nana* and *S. glauca* distributions are best

predicted by the climate parameters yearly minimum and maximum temperatures, temperature range, temperature continentality, and spring precipitation. This suggests the importance of considering climate parameters beyond annual means, such as seasonal precipitation and climatic extremes (Stewart et al., 2021; Zimmermann et al.,

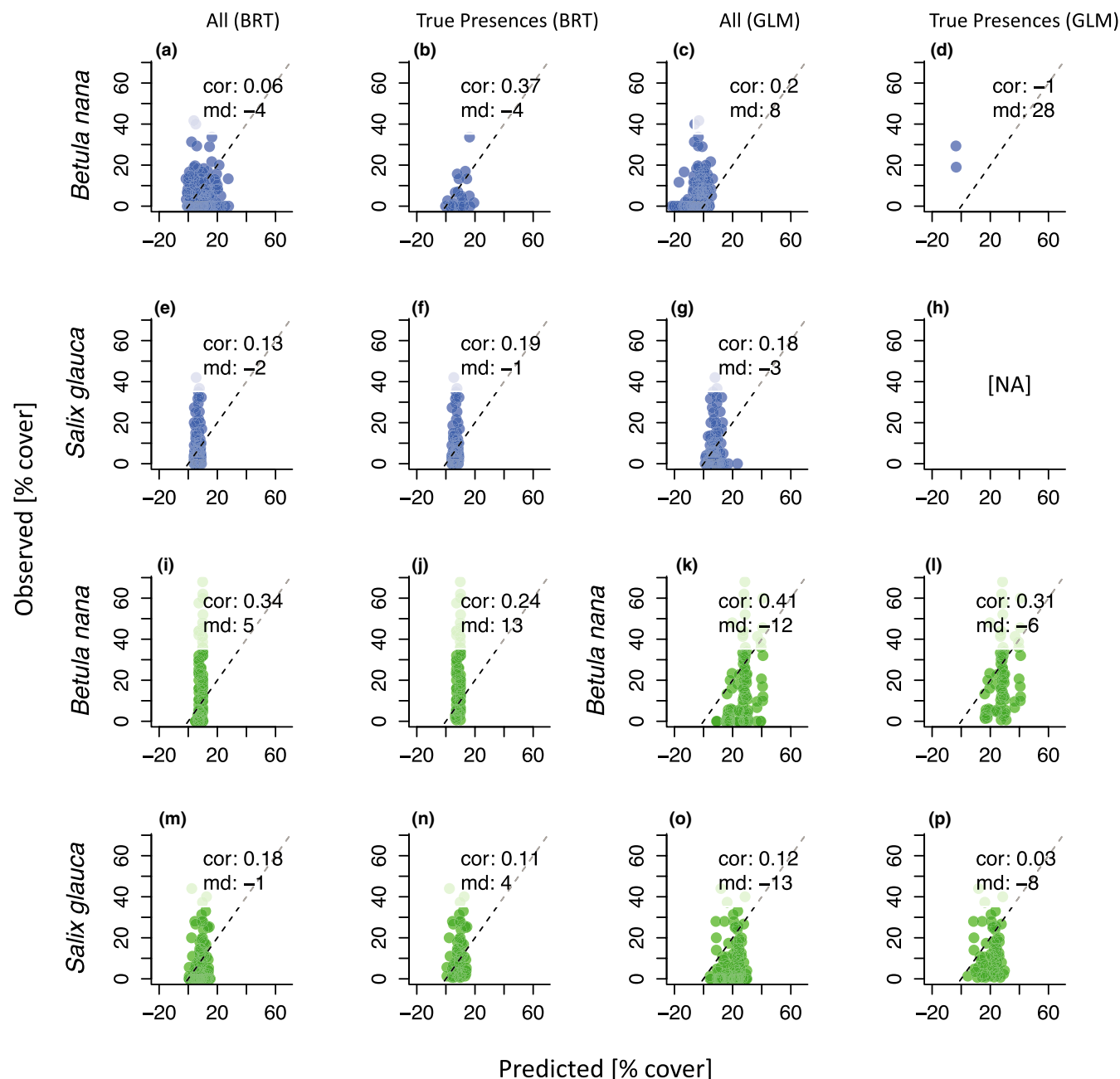


FIGURE 4 *Betula nana* and *Salix glauca* boosted regression tree (BRT) and generalised linear model (GLM) species abundance models (SAMs) generally underpredicted (as seen with negative md: mean deviation) in both Kangerluarsunnguak (K: blue points) fjord and Nuup Kangerlua (NK: green points) fjord. Model performance (cor: Pearson's correlation) did not generally increase when evaluated at correctly predicted presences ("True Presences") of the corresponding species distribution model, as evaluated with a default threshold of 0.5. Shown in blue (a–h) are SAMs trained in K fjord and in green (i–p) those trained in NK fjord. Each plot shows the line of perfect prediction (where Pearson's correlation = 1)

2009). This is reflected in other Arctic studies, which have found plant responses to both climatic extremes and seasonal variables (Berner et al., 2020; Niittynen et al., 2020; von Oppen et al., 2021). In particular, a recent exploratory analysis of the NK fjord data by some of the authors found that both *B. nana* and *S. glauca* abundances increase with temperature range, and that *S. glauca* abundance also increases with TWI, a proxy for soil moisture (von Oppen et al., 2021). We also found that temperature continentality and TCW, another

indicator of soil moisture, are influential predictors of both species' abundances. Our findings are in line with other work that emphasises the importance of soil moisture in limiting arctic shrub distributions and abundances (Ackerman et al., 2017; Elmendorf, Henry, Hollister, Björk, Björkman, et al., 2012; Elmendorf, Henry, Hollister, Björk, Boulanger-Lapointe, et al., 2012; Kemppinen et al., 2019; Myers-Smith et al., 2015; Nabe-Nielsen et al., 2017) as well as *B. nana* and *S. glauca* growth (Gamm et al., 2018).

TABLE 3 Adding biotic predictors to the final predictor set generally improved both boosted regression tree (BRT) and generalized linear model (GLM) Species Abundance Model (SAM) performance

Training data	Species	Model	Abiotic only	+ richness (grid) ^a	+ richness (plot) ^a	+ diversity	+ competition ^a	+ canopy height ^a
NK	<i>B. nana</i>	BRT	0.57	NA	0	0.06	0.08	NA
NK	<i>B. nana</i>	GLM	0.5	NA	-0.01	-0.02	-0.02	NA
NK	<i>S. glauca</i>	BRT	0.09	NA	0.01	0.07	-0.07	NA
NK	<i>S. glauca</i>	GLM	0.26	NA	-0.04	0.07	-0.2	NA
K	<i>B. nana</i>	BRT	0.38	0.02	0.08	0.12	0.09	0
K	<i>B. nana</i>	GLM	0.43	-0.01	0	0.01	0.04	-0.03
K	<i>S. glauca</i>	BRT	0.37	0.03	-0.06	0.01	0.04	0.19
K	<i>S. glauca</i>	GLM	0.36	0	0	0.02	0.12	0.27

Note: We report the differences in Pearson's correlation coefficient between the SAM with only abiotic predictors (first column, as reported in Table 1) and the corresponding SAM with the indicated added biotic predictor. We calculated the correlation coefficient using predicted and observed values in five repetitions of $\frac{1}{2}$ - $\frac{1}{2}$ split-sample validation of models trained in either Nuup Kangerlua (NK) fjord or Kangerluarsunnguak (K) fjord. Each biotic predictor represents indices at the shrub species community level (i.e. excluding non-shrub species), which we added to each SAM one at a time. Note that we did not collect all biotic data in NK fjord.

^aGrid = richness calculated from grid-level scans for species presences; plot = richness calculated from species presences within sampling plot; competition = sum of abundances of all shrub species excluding the target species; canopy height = community weighted mean (by abundance) of maximum vegetative shrub height excluding the target species.

SDM outputs are often used to infer species' population performance, such as abundance, or ecological function, such as plant functional traits (e.g., Araújo & Williams, 2000; Lee-Yaw et al., 2016; Thuiller et al., 2009; Van Couwenberghe et al., 2013; VanDerWal et al., 2009; Weber et al., 2017). As many others (Chardon et al., 2020; Csergő et al., 2017; Dallas & Hastings, 2018; Thuiller et al., 2014), we did not find strong support for such relationships (Table S1.6). Species height and shrub canopy height had the highest correlations with SAM outputs (Pearson's correlation coefficients 0.30–0.75), and species height might be a meaningful population performance parameter to infer from these model outputs since it can be correlated with growth rate (Paradis et al., 2016). We recommend further exploration of whether SAM predictions tend to be superior to those of SDMs when inferring population performance or ecological function, although we urge caution in relying on this inference given the many contrasting results that recent review studies have produced (e.g., Dallas & Hastings, 2018; Lee-Yaw et al., 2021; Weber et al., 2017).

4.2 | Separate survey data as a useful validation approach

In our study, split-sample validation yielded overly optimistic SDM and SAM performance compared to the separate survey validation approach, in line with other work showing that independent validation data yield more realistic, but weaker, results (e.g., Dobrowski et al., 2011; Eskildsen et al., 2013; Roberts et al., 2017; Schratz et al., 2019). This highlights the importance of using separate survey data for spatially independent validation (Hijmans, 2012). Yet, few studies use independent and distinct datasets collected across the range of predictions for model validation (Araújo et al., 2005; Elith et al., 2006, 2020; Lee-Yaw et al., 2021; Santini et al., 2021), even though

this would be ideal (Araújo & Guisan, 2006). Separate survey validation in our system illustrated poor model predictivity within fully overlapping to non-overlapping environmental predictor spaces (see Figures S1.1 and S1.2), even though the ranges of observed response variables are similar in both fjords (see Figure S1.9). While decreasing spatial dependence between training and validation data can be achieved by spatially or temporally cross-validating an available dataset with independent training-validation portions (Araújo et al., 2005; Gogol-Prokurat, 2011; Martínez-Meyer et al., 2004; Morán-Ordóñez et al., 2017; Rapacciuolo et al., 2012; Raxworthy et al., 2003; Sofaer et al., 2018), this would have yielded substantial model transference in NK fjord (see Figure S1.10).

Our separate survey validation was based on environmental spaces that only partially overlap with the respective training spaces (see Figure S1.1), and this could be a contributing reason to why our models predicted so poorly in adjacent fjords (e.g., Duque-Lazo et al., 2016; Pearson et al., 2006; Randin et al., 2006; Werkowska et al., 2017). In particular, the two contrasting modelling techniques we employed extrapolate differently outside the environmental training space. BRTs extrapolate at a constant value from the last observed point in environmental training space, and GLMs continue their fitted function (Elith & Graham, 2009). However, when we validated our models with the separate survey approach in the subset of the adjacent fjord's environmental space that overlaps with the training space, the separate survey validation still yielded substantially poorer results than split-sample validation (see Table S1.5). As model validation with this subset data decreased statistical power, we are cautious in our conclusions here. While others have used different datasets to train and validate models (e.g., Angert et al., 2018; Araújo et al., 2005; Elith et al., 2006; Eskildsen et al., 2013), to our knowledge we are the first to compare validation approaches within the overlapping

training environmental predictor space and we urge more research on this approach.

4.3 | Relevance of predictors and spatial scale

An additional explanation for poor model performance with the separate survey validation is that our models perhaps did not include all relevant predictors, and that we did not capture enough local scale heterogeneity in predictor and response variables with the relatively small extent of our two surveys. Including biotic predictors generally improved performance in our models, as it does in other studies (e.g., Mod et al., 2015), and is not surprising given the well-documented lack of relationship between the distribution of *S. glauca* and abiotic factors (Böcher, 2000; Born & Böcher, 2000). However, data on potentially important biotic predictors are often lacking at the landscape level needed to parameterise spatial models, although advances in remote sensing are closing this gap (Randin et al., 2020). This points to the need for data collection specifically aimed at understanding the relevant abiotic and biotic factors driving species responses, such as soil temperature (Lembrechts et al., 2020), herbivory (Prendin et al., 2020), or mycorrhizal associations (Gardes & Dahlberg, 1996).

Constructing high resolution models, especially in data-poor areas such as the Arctic, should be accompanied by careful consideration of response variables and meaningful predictors at relevant spatial scales (Levin, 1992). The potential difference in the drivers of species' distribution and abundances at broad (e.g., 1 km) versus local (e.g., 90 m) scales means that such models should be parameterised with scale-specific predictors (e.g., Gillingham et al., 2012). The broad climatic predictors used in low resolution models might not be the most relevant drivers of local scale distribution and abundance patterns (Marcer et al., 2012), which are often better described by microhabitat abiotic and biotic factors (Körner, 2003; Louthan et al., 2015; Suvanto et al., 2014). There might also be a scale mismatch between the variation in broad climate and that in local distribution and abundance patterns. Furthermore, down-scaled precipitation data are more error-prone in topographically complex areas, such as our study region, and can thus misrepresent local scale climate (Baker et al., 2017). Properly parameterising high resolution models thus requires an increased understanding of the factors driving species' variable, and sometimes even stochastic, responses at local scales.

Our interdependent survey design (see Appendix S2) did not allow us to test for the effects of sampling design and it remains insufficiently explored how sampling design affects model predictions, despite the potential for contrasting sampling designs to differentially capture landscape heterogeneity (e.g., Schweiger et al., 2016). One of the few studies that has explored the role of sampling design did so by computationally selecting data from global occurrences, but did not find an effect of different designs on model performance (Tessarolo et al., 2014). We encourage more work exploring the effects of sampling design on model predictions, to provide guidelines

for the optimal collection of species distribution data along ecological gradients in future studies.

There is momentum towards constructing spatial models at higher resolution (D'Amen et al., 2015; Thuiller et al., 2015; Tomlinson et al., 2020), which might better represent local scale drivers of species distributions and improve model transferability (Manzoor et al., 2018). However, such high resolution models can show little agreement with low resolution ones (Franklin et al., 2013) and even perform similarly to or worse than low resolution models (Pradervand et al., 2014; Thuiller et al., 2015). We suggest that poor model performance could be tied to predictor and response variables failing to capture relevant heterogeneity or the appropriate spatial scale, however this problem may also depend on algorithm choice. Although machine learning approaches, such as BRT, have several advantages (Elith et al., 2008), they can create uncertain estimates of variable contributions (Smith & Santos, 2020). Furthermore, while machine learning approaches generally resist overfitting, such tree-based models are not immune to overfitting the training data and they might poorly predict in new environmental spaces (Friedman et al., 2000; but see Yates et al., 2018). GLMs can outperform classification trees (Meynard & Quinn, 2007; Thuiller et al., 2003; Duque-Lazo et al., 2016; but see Valavi et al., 2021) but can provide unrealistic estimates outside the training space (reviewed in Yates et al., 2018). Ultimately, the choice of modelling approach largely depends on the aim (Merow et al., 2014) and generating robust predictions is best done with ensemble approaches (Araújo & New, 2007; Thuiller et al., 2019).

4.4 | Spatial models in biodiversity conservation

While SDMs have been widely used in conservation planning and management decisions (e.g., Guisan et al., 2013; Rodríguez et al., 2007; Tulloch et al., 2016), to our knowledge they are not applied in Arctic conservation. As vegetation is closely linked to the movement and distribution of caribou and muskox (e.g., CAFF, 2013; Raundrup et al., 2018), accurate modelling of changes in plant communities and their distribution patterns could play a central role in the long-term management for these large herbivores. Changes in dominant vegetation types, such as shrubs, could impact higher trophic levels (Klein et al., 2008), yet our poor understanding of vegetation shifts makes for uncertain estimates of how herbivore populations will change (e.g., Cuyler et al., 2020). Work on modelling the habitat suitability of muskox has indicated that their most suitable habitat is in areas that are vegetated, among other criteria (Beumer et al., 2019), and a decrease in caribou offspring success has been linked to a reduction in the spatial variability of plant phenology (Post et al., 2008). Our work suggests that, at the spatial resolution and with the predictors we examined, SDMs are a superior choice to SAMs in describing the spatial patterning of dominant shrub species.

We specifically focused on two contrasting algorithms and a split-sample validation approach to reflect common SDM modelling practices (Santini et al., 2021; Thuiller et al., 2019) but did not

find large differences between algorithm types. We further sought to compare SDM and SAM performance, and in our system SDMs outperformed SAMs. The pronounced differences in model performance between the standard split-sample validations with the rarely employed separate survey validation highlight the latter as a useful, more independent, validation approach. We encourage more work to further explore this, as well as the effect of sampling design on model performance to guide the collection of ecological data for predictive models. We furthermore urge ecologists and biogeographers to carefully choose relevant predictors at meaningful spatial scales. Finally, we recommend conducting more work to elucidate how species respond to their surrounding local and regional scale habitat, and taking great care when upscaling results to get a better understanding of the impacts of climate change on biodiversity.

ACKNOWLEDGMENTS

We pay our respects to the Kalaallit members past and present, who are the original and current inhabitants of the lands on which we conducted our field work. We are indebted to the Greenland Institute of Natural Resources and Kobbefjord Research Station, and especially Katrine Raundrup, for field work support and site access. Field work funding was awarded by Greenland Climate Research Centre (JNN), Arctic Research Centre at Aarhus University (JNN), European Union's Horizon 2020 project INTERACT grant no. 730938 (NIC), American Alpine Club (NIC) and the late Konrad Steffen via the WSL Director's Fund (NIC). NIC and SW were funded by the Swiss Polar Institute. NIC was also funded by the Basler Foundation for biological research, Switzerland and a Swiss National Science Foundation Postdoc. Mobility fellowship (project no. P400PB_194331). JJA was funded by the Independent Research Fund, Denmark (grant no. 7027-00133B to SN). We thank Isabelle Boulangeat, Amy Angert and Oliver Baines for helpful discussions, as well as Susan Jarvis for constructive editorial feedback.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Available in Appendix S4 in Supporting Information are the presence-absence and abundance data of *Salix glauca* and *Betula nana*, as well as species richness and competition on either species, at the 90 m grid level from the Nuup Kangerlua survey. Available in Appendix S5 are all field data collected in Kangerluarsunnguag at the 90 m grid level. These field data are stored in the public repository EnviDat (Chardon et al., 2022) and all other data used in this work are also publicly available (see Appendix S2 for all data references). Computer code necessary to reproduce this work can be found on GitLab (https://gitlab.com/nathaliechardon/gl_microclim).

ORCID

Nathalie Isabelle Chardon  <https://orcid.org/0000-0001-9120-4778>

Jacob Nabe-Nielsen  <https://orcid.org/0000-0002-0716-9525>

Jakob Johan Assmann  <https://orcid.org/0000-0002-3492-8419>

Ida Bomholt Dyrholm Jacobsen  <https://orcid.org/0000-0003-2481-8061>

Signe Normand  <https://orcid.org/0000-0002-8782-4154>

Sonja Wipf  <https://orcid.org/0000-0002-3492-1399>

REFERENCES

- Ackerman, D., Griffin, D., Hobbie, S. E., & Finlay, J. C. (2017). Arctic shrub growth trajectories differ across soil moisture levels. *Global Change Biology*, 23(10), 4294–4302. <https://doi.org/10.1111/gcb.13677>
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proc. 2nd Int. Symp. Information Theory, Supp. to Problems of Control and Information Theory*. 267–281.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Angert, A. L., Bayly, M., Sheth, S. N., & Paul, J. R. (2018). Testing range-limit hypotheses using range-wide habitat suitability and occupancy for the Scarlet Monkey flower (*Erythranthe cardinalis*). *The American Naturalist*, 191(3), E76–E89. <https://doi.org/10.1086/695984>
- Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5(1), 1–12. <https://doi.org/10.1126/sciadv.aat4858>
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10), 1677–1688. <https://doi.org/10.1111/j.1365-2699.2006.01584.x>
- Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, 22(1), 42–47. <https://doi.org/10.1016/j.tree.2006.09.010>
- Araújo, M., Pearson, R., Thuiller, W., & Erhard, M. (2005). Validation of species-climate impact models under climate change. *Global Change Biology*, 11(9), 1504–1513. <https://doi.org/10.1111/j.1365-2486.2005.001000.x>
- Araújo, M. B., & Williams, P. H. (2000). Selecting areas for species persistence using occurrence data. *Biological Conservation*, 96(3), 331–345. [https://doi.org/10.1016/S0006-3207\(00\)00074-4](https://doi.org/10.1016/S0006-3207(00)00074-4)
- Baker, D. J., Hartley, A. J., Pearce-Higgins, J. W., Jones, R. G., & Willis, S. G. (2017). Neglected issues in using weather and climate information in ecology and biogeography. *Diversity and Distributions*, 23(3), 329–340. <https://doi.org/10.1111/ddi.12527>
- Bay, C. (1999). Floristiske Provinser. In D. B. Jensen (Ed.), (*Grønlands biodiversitet - et landestudie* (p. 25). Greenland Institute of Natural Resources.
- Becher, M., Olofsson, J., Berglund, L., & Klaminder, J. (2018). Decreased cryogenic disturbance: One of the potential mechanisms behind the vegetation change in the Arctic. *Polar Biology*, 41(1), 101–110. <https://doi.org/10.1007/s00300-017-2173-5>
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, 15(4), 365–377. <https://doi.org/10.1111/j.1461-0248.2011.01736.x>
- Berner, L. T., Massey, R., Jantz, P., Forbes, B. C., Macias-Fauria, M., Myers-Smith, I., Kumpula, T., Gauthier, G., Andreu-Hayles, L., Gaglioti, B. V., Burns, P., Zetterberg, P., D'Arrigo, R., & Goetz, S. J. (2020). Summer warming explains widespread but not uniform greening in the Arctic tundra biome. *Nature Communications*, 11, 4621. <https://doi.org/10.1038/s41467-020-18479-5>

- Beumer, L. T., van Beest, F. M., Stelvig, M., & Schmidt, N. M. (2019). Spatiotemporal dynamics in habitat suitability of a large Arctic herbivore: Environmental heterogeneity is key to a sedentary lifestyle. *Global Ecology and Conservation*, 18, e00647. <https://doi.org/10.1016/J.GECCO.2019.E00647>
- Birch, L. C. (1953). Experimental background to the study of the distribution and abundance of insects. I. The influence of temperature, moisture and food on the innate capacity for increase of 3 grain beetles. *Ecology*, 34, 698–711.
- Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Normand, S., Rüger, N., Beck, P. S. A., Blach-Overgaard, A., Blok, D., Cornelissen, J. H. C., Forbes, B. C., Georges, D., Goetz, S. J., Guay, K. C., Henry, G. H. R., HilleRisLambers, J., Hollister, R. D., Karger, D. N., Kattge, J., Manning, P., ... Weiher, E. (2018). Plant functional trait change across a warming tundra biome. *Nature*, 562(7725), 57–62. <https://doi.org/10.1038/s41586-018-0563-7>
- Böcher, T. W. (2000). *Det Grønne Grønland* [Green Greenland] (pp. 146, 148, 174, 186, 189). B. Fredskild & J. J. Böcher (Eds.). Rhodos.
- Böcher, T. W., Holmen, K., & Jakobson, K. (1978). *Grønlands Flora* [The Flora of Greenland] (Vol. 10, No. 1). P. Haase & Søn's Forlag. <https://doi.org/10.14430/arctic3887>
- Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A., & Selige, T. (2002). Soil regionalisation by means of terrain analysis and process parameterisation. In E. Micheli, F. Nachtergaele, & L. Montanarella (Eds.), *Soil classification 2001* (pp. 213–222). European Soil Bureau, Research Report No. 7, EUR 20398 EN.
- Boehner, J., & Selige, T. (2006). Spatial prediction of soil attributes using terrain analysis and climate regionalisation. In J. Boehner, K. R. McCloy, & J. Strobl (Eds.), *SAGA - analysis and modelling applications* (pp. 13–28). Goettinger Geographische Abhandlungen.
- Booth, T. H., & McMurtrie, R. E. (1988). Climatic change and *Pinus radiata* plantations in Australia. In G. I. Pearman (Ed.), *Greenhouse: Planning for climate change* (pp. 534–545). CSIRO Australia.
- Born, E. W., & Böcher, J. (2000). *The ecology of Greenland* (pp. 240, 241, 268). Ministry of Environment and Natural Resources.
- Boulangeat, I., Gravel, D., & Thuiller, W. (2012). Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, 15(6), 584–593. <https://doi.org/10.1111/j.1461-0248.2012.01772.x>
- CAFF (2013). *Arctic biodiversity assessment. Status and trends in Arctic biodiversity*. Conservation of Arctic Flora and Fauna.
- Chardon, N. I., Nabe-Nielsen, J., Assmann, J., Dyrholm Jacobsen, I. B., Guéguen, M., Normand, S., & Wipf, S. (2022). Greenland shrubs and microclimate. *EnviDat*. <https://doi.org/10.16904/envidat.286>
- Chardon, N. I., Pironon, S., Peterson, M. L., & Doak, D. F. (2020). Incorporating intraspecific variation into species distribution models improves distribution predictions, but cannot predict species traits for a wide-spread plant species. *Ecography*, 43(1), 60–74. <https://doi.org/10.1111/ecog.04630>
- Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158(3), 419. <https://doi.org/10.2307/2983440>
- Crist, E. P., & Cicone, R. C. (1984). A physically-based transformation of Thematic Mapper Data—The TM Tasseled Cap'. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(3), 256–263. <https://doi.org/10.1109/TGRS.1984.350619>
- Csörgő, A. M., Salguero-Gómez, R., Broennimann, O., Coutts, S. R., Guisan, A., Angert, A. L., Welk, E., Stott, I., Enquist, B. J., McGill, B., Svenning, J.-C., Violle, C., & Buckley, Y. M. (2017). Less favourable climates constrain demographic strategies in plants. *Ecology Letters*, 20(8), 969–980. <https://doi.org/10.1111/ele.12794>
- Cuyler, C., Rowell, J., Adamczewski, J., Anderson, M., Blake, J., Bretten, T., Brodeur, V., Campbell, M., Checkley, S. L., Cluff, H. D., Côté, S. D., Davison, T., Dumond, M., Ford, B., Gruzdev, A., Gunn, A., Jones, P., Kutz, S., Leclerc, L. M., ... Ytrehus, B. (2020). Muskox status, recent variation, and uncertain future. *Ambio*, 49(3), 805–819. <https://doi.org/10.1007/S13280-019-01205-X/FIGURES/2>
- Dallas, T. A., & Hastings, A. (2018). Habitat suitability estimated by niche models is largely unrelated to species abundance. *Global Ecology and Biogeography*, 27(12), 1448–1456. <https://doi.org/10.1111/geb.12820>
- D'Amen, M. D., Pradervand, J., & Guisan, A. (2015). Predicting richness and composition in mountain insect communities at high resolution: A new test of the SESAM framework. *Global Ecology and Biogeography*, 24, 1443–1453. <https://doi.org/10.1111/geb.12357>
- DeLeo, J. M. (1993). Receiver operating characteristic laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty. In 1993 (2nd) *International Symposium on Uncertainty Modeling and Analysis*. 318–325. <https://doi.org/10.1109/ISUMA.1993.366750>
- Dobrowski, S. Z., Thorne, J. H., Greenberg, J. A., Safford, H. D., Mynsberge, A. R., Crimmins, S. M., & Swanson, A. K. (2011). Modeling plant ranges over 75 years of climate change in California, USA: Temporal transferability and species traits. *Ecological Monographs*, 81(2), 241–257. <https://doi.org/10.1890/10-1325.1>
- Dormann, C. F. (2020). Calibration of machine-learning probability predictions. *Global Ecology and Biogeography*, 29(4), 760–765. <https://doi.org/10.1111/geb.13070>
- Duque-Lazo, J., van Gils, H., Groen, T. A., & Navarro-Cerrillo, R. M. (2016). Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. *Ecological Modelling*, 320, 62–70. <https://doi.org/10.1016/j.ecolmodel.2015.09.019>
- Elith, J., Burgman, M. A., & Regan, H. M. (2002). Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling*, 157(2–3), 313–329. [https://doi.org/10.1016/S0304-3800\(02\)00202-8](https://doi.org/10.1016/S0304-3800(02)00202-8)
- Elith, J., & Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32(1), 66–77. <https://doi.org/10.1111/j.1600-0587.2008.05505.x>
- Elith, J., Graham, C., Valavi, R., Abegg, M., Bruce, C., Ford, A., Guisan, A., Hijmans, R. J., Huettmann, F., Lohmann, L., Loiselle, B., Moritz, C., Overton, J., Peterson, A. T., Phillips, S., Richardson, K., Williams, S., Wiser, S. K., Wohlgemuth, T., & Zimmermann, N. E. (2020). Presence-only and presence-absence data for comparing species distribution modeling methods. *Biodiversity Informatics*, 15(2), 69–80. <https://doi.org/10.17161/bi.v15i2.13384>
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., ... E. Zimmermann, N. (2006). Novel methods improve prediction of species distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Elmendorf, S. C., Henry, G. H. R., Hollister, R. D., Björk, R. G., Bjorkman, A. D., Callaghan, T. V., Collier, L. S., Cooper, E. J., Cornelissen, J. H. C., Day, T. A., Fosaa, A. M., Gould, W. A., Grétarsdóttir, J., Harte, J., Hermanutz, L., Hik, D. S., Hofgaard, A., Jarrad, F., Jónsdóttir, I. S., ... Wookey, P. A. (2012). Global assessment of experimental climate warming on tundra vegetation: Heterogeneity over space and time. *Ecology Letters*, 15(2), 164–175. <https://doi.org/10.1111/j.1461-0248.2011.01716.x>

- Elmendorf, S. C., Henry, G. H. R., Hollister, R. D., Björk, R. G., Boulanger-Lapointe, N., Cooper, E. J., Cornelissen, J. H. C., Day, T. A., Dorrepaal, E., Elumeeva, T. G., Gill, M., Gould, W. A., Harte, J., Hik, D. S., Hofgaard, A., Johnson, D. R., Johnstone, J. F., Jónsdóttir, I. S., Jorgenson, J. C., ... Wipf, S. (2012). Plot-scale evidence of tundra vegetation change and links to recent summer warming. *Nature Climate Change*, 2(6), 453–457. <https://doi.org/10.1038/nclimate1465>
- Elvarsdóttir, H. S. (2013). *An Approach Towards Sustainable Coastal Tourism Management: Nature-based Tourism in Nuup Kangerlua, Greenland* [Master Thesis]. University of Akureyri, Iceland.
- Elven, R., Murray, D. F., Razzhivin, V. Y., & Yurtsev, B. A. (2020). *Pan-arctic flora: Salix glauca. Annotated Checklist of the Panarctic Flora (PAF) Vascular Plants*. Retrieved from <http://panarcticflora.org/results?biogeographic=&bioclimatic=®ion=&name=Salix+glauc+a#paf-580213>
- Eskildsen, A., le Roux, P. C., Heikkinen, R. K., Høye, T. T., Kissling, W. D., Pöyry, J., Wisz, M. S., & Luoto, M. (2013). Testing species distribution models across space and time: High latitude butterflies and recent warming. *Global Ecology and Biogeography*, 22(12), 1293–1303. <https://doi.org/10.1111/geb.12078>
- Feilberg, J. (1984). A phytogeographical study of south Greenland vascular plants. *Meddelelser Om Grønland*, 15, 1–70.
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1), 38–49. <https://doi.org/10.1017/S0376892997000088>
- Fitzpatrick, M. C., Lachmuth, S., & Haydt, N. T. (2021). The ODMAP protocol: A new tool for standardized reporting that could revolutionize species distribution modeling. *Ecography*, 44, 1–4. <https://doi.org/10.1111/dom.2013.15.issue-1>
- Foxcroft, L. C., Richardson, D. M., Rouget, M., & MacFadyen, S. (2009). Patterns of alien plant distribution at multiple spatial scales in a large national park: Implications for ecology, management and monitoring. *Diversity and Distributions*, 15(3), 367–378. <https://doi.org/10.1111/j.1472-4642.2008.00544.x>
- Franklin, J., Davis, F. W., Ikegami, M., Syphard, A. D., Flint, L. E., Flint, A. L., & Hannah, L. (2013). Modeling plant species distributions under future climates: How fine scale do climate projections need to be? *Global Change Biology*, 19(2), 473–483. <https://doi.org/10.1111/gcb.12051>
- Fredskild, B. (1996). *A phytogeographical study of the vascular plants of West Greenland (62°20'–74°00'N)*. Meddelelser om Grønland. Bioscience; 45. Copenhagen, Denmark: Commission for Scientific Research in Greenland.
- Freund, Y. (1995). Boosting a weak learning algorithm by majority. *Information and Computation*, 121(2), 256–285. <https://doi.org/10.1006/inco.1995.1136>
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119–139. <https://doi.org/10.1006/jcss.1997.1504>
- Fridley, J. D. (2009). Downscaling climate over complex terrain: High finescale (<1000 m) spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains). *Journal of Applied Meteorology and Climatology*, 48(5), 1033–1049. <https://doi.org/10.1175/2008JAMC2084.1>
- Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting (With discussion and a rejoinder by the authors). *The Annals of Statistics*, 28(2), 337–407. <https://doi.org/10.1214/aos/1016218223>
- Gamm, C. M., Sullivan, P. F., Buchwal, A., Dial, R. J., Young, A. B., Watts, D. A., Cahoon, S. M. P., Welker, J. M., & Post, E. (2018). Declining growth of deciduous shrubs in the warming climate of continental western Greenland. *Journal of Ecology*, 106(2), 640–654. <https://doi.org/10.1111/1365-2745.12882>
- García-Díaz, P., Prowse, T. A. A., Anderson, D. P., Lurgi, M., Binny, R. N., & Cassey, P. (2019). A concise guide to developing and using quantitative models in conservation management. *Conservation Science and Practice*, 1(2), e11. <https://doi.org/10.1111/csp2.11>
- Gardes, M., & Dahlberg, A. (1996). Mycorrhizal diversity in arctic and alpine tundra: An open question. *New Phytologist*, 133(1), 147–157. <https://doi.org/10.1111/j.1469-8137.1996.tb04350.x>
- Gillingham, P. K., Palmer, S. C. F., Huntley, B., Kunin, W. E., Chipperfield, J. D., & Thomas, C. D. (2012). The relative importance of climate and habitat in determining the distributions of species at different spatial scales: A case study with ground beetles in Great Britain. *Ecography*, 35(9), 831–838. <https://doi.org/10.1111/j.1600-0587.2011.07434.x>
- Global Biodiversity Facility (2020). *Occurrence search Betula nana L.* Accessed on 10–02-2021. Retrieved from https://www.gbif.org/occurrence/map?taxon_key=5332004
- Gogol-Prokurat, M. (2011). Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. *Ecological Applications*, 21(1), 33–47. <https://doi.org/10.1890/09-1190.1>
- Grønlands statistik (2021). *Grønlands befolkning* [Greenland's population]. Retrieved from <https://stat.gl/dialog/main.asp?lang=da&version=202101&sc=BE&subthemecode=GRØNLANDSBEFOLKNIN&G2021&colcode=G>
- Guarino, E. D. S. G., Barbosa, A. M., & Waechter, J. L. (2012). Occurrence and abundance models of threatened plant species: Applications to mitigate the impact of hydroelectric power dams. *Ecological Modelling*, 230, 22–33. <https://doi.org/10.1016/j.ecolmodel.2012.01.007>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models*. Cambridge University Press. <https://doi.org/10.1017/9781139028271>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3), 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Halliday, G. (2019). *A flora of the east Greenland central fjord region, 70°N–77°N*. Trollius Publications.
- Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, 93(3), 679–688. <https://doi.org/10.1890/11-0826.1>
- Howat, I. M., Negrete, A., & Smith, B. E. (2014). The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets. *The Cryosphere*, 8(4), 1509–1518. <https://doi.org/10.5194/tc-8-1509-2014>
- Howat, I. M., Negrete, A., & Smith, B. E. (2015). *MEASUREs Greenland Ice Mapping Project (GIMP) Digital Elevation Model, Version 1*. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/NV34YU1XLP9W>
- IPCC (2014). *Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. Retrieved from https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-FrontMatterA_FINAL.pdf
- Jarnevich, C. S., Sofaer, H. R., & Engelstad, P. (2021). Modelling presence versus abundance for invasive species risk assessment. *Diversity and Distributions*, 27, 2454–2464. <https://doi.org/10.1111/ddi.13414>
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, P., & Kessler, M. (2017). Climatologies

- at high resolution for the Earth land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122G>
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2018). Data from: Climatologies at high resolution for the earth's land surface areas. Dryad Digital Repository. <https://doi.org/10.5061/dryad.kd1d4>
- Kemppinen, J., Niittynen, P., Aalto, J., le Roux, P. C., & Luoto, M. (2019). Water as a resource, stress and disturbance shaping tundra vegetation. *Oikos*, 128(6), 811–822. <https://doi.org/10.1111/oik.05764>
- Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., McInerney, G. J., Montoya, J. M., Römermann, C., Schiffers, K., Schurr, F. M., Singer, A., Svenning, J.-C., Zimmermann, N. E., & O'Hara, R. B. (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, 39(12), 2163–2178. <https://doi.org/10.1111/j.1365-2699.2011.02663.x>
- Klein, D. R., Bruun, H. H., Lundgren, R., & Philipp, M. (2008). Climate change influences on species interrelationships and distributions in high-arctic Greenland. *Advances in Ecological Research*, 40, 81–100. [https://doi.org/10.1016/S0065-2504\(07\)00004-9](https://doi.org/10.1016/S0065-2504(07)00004-9)
- Körner, C. (2003). *Alpine plant life: Functional plant ecology of high mountain ecosystems*, 2nd ed. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-18970-8>
- Lawrence, D. M., & Swenson, S. C. (2011). Permafrost response to increasing Arctic influence of shrubs on local soil cooling versus large-scale climate warming. *Environmental Research Letters*, 6(4), 045504. <https://doi.org/10.1088/1748-9326/6/4/045504>
- Lee-Yaw, J. A., Kharouba, H. M., Bontrager, M., Mahony, C., Csörgő, A. M., Noreen, A. M. E., Li, Q., Schuster, R., & Angert, A. L. (2016). A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecology Letters*, 19(6), 710–722. <https://doi.org/10.1111/ele.12604>
- Lee-Yaw, J. A., McCune, J. L., Pironon, S., & Sheth, S. N. (2021). Species distribution models rarely predict the biology of real populations. *Ecography*, 44, 1–16.
- Lembrechts, J. J., Aalto, J., Ashcroft, M. B., De Frenne, P., Kopecký, M., Lenoir, J., Luoto, M., Maclean, I. M. D., Rouspard, O., Fuentes-Lillo, E., García, R. A., Pellissier, L., Pitteloud, C., Alatalo, J. M., Smith, S. W., Björk, R. G., Muffler, L., Ratier Backes, A., Cesarz, S., ... Nijs, I. (2020). SoilTemp: A global database of near-surface temperature. *Global Change Biology*, 26(11), 6616–6629. <https://doi.org/10.1111/gcb.15123>
- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73(6), 1943–1967. <https://doi.org/10.2307/1941447>
- Liston, G. E., Mcfadden, J. P., Sturm, M., & Pielke, R. A. (2002). Modelled changes in arctic tundra snow, energy and moisture fluxes due to increased shrubs. *Global Change Biology*, 8(1), 17–32. <https://doi.org/10.1046/j.1354-1013.2001.00416.x>
- Lou, Y., Gao, C., Pan, Y., Xue, Z., Liu, Y., Tang, Z., Jiang, M., Lu, X., & Rydin, H. (2018). Niche modelling of marsh plants based on occurrence and abundance data. *Science of the Total Environment*, 616–617, 198–207. <https://doi.org/10.1016/j.scitotenv.2017.10.300>
- Louthan, A. M., Doak, D. F., & Angert, A. L. (2015). Where and when do species interactions set range limits? *Trends in Ecology & Evolution*, 30(12), 780–792. <https://doi.org/10.1016/j.tree.2015.09.011>
- Lovejoy, T. E., & Hannah, L. (2019). *Biodiversity and climate change*. Yale University Press.
- Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size-finer may not always be better. *Scientific Reports*, 8(1), 1–9. <https://doi.org/10.1038/s41598-018-25437-1>
- Marcen, A., Pino, J., Pons, X., & Brotons, L. (2012). Modelling invasive alien species distributions from digital biodiversity atlases. Model upscaling as a means of reconciling data at different scales. *Diversity and Distributions*, 18(12), 1177–1189. <https://doi.org/10.1111/j.1472-4642.2012.00911.x>
- Martin, A. C., Jeffers, E. S., Petrokofsky, G., Myers-Smith, I., & Macias-Fauria, M. (2017). Shrub growth and expansion in the Arctic tundra: An assessment of controlling factors using an evidence-based approach. *Environmental Research Letters*, 12, 085007. <https://doi.org/10.1088/1748-9326/aa7989>
- Martínez-Meyer, E., Peterson, A. T., & Hargrove, W. W. (2004). Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*, 13(4), 305–314. <https://doi.org/10.1111/j.1466-822X.2004.00107.x>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall.
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37, 1267–1281. <https://doi.org/10.1111/ecog.00845>
- Meynard, C. N., & Quinn, J. F. (2007). Predicting species distributions: A critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, 34(8), 1455–1469. <https://doi.org/10.1111/j.1365-2699.2007.01720.x>
- Midolo, G., Wellstein, C., & Faurby, S. (2021). Individual fitness is decoupled from coarse-scale probability of occurrence in North American trees. *Ecography*, 44(5), 789–801. <https://doi.org/10.1111/ecog.05446>
- Mod, H. K., le Roux, P. C., Guisan, A., & Luoto, M. (2015). Biotic interactions boost spatial models of species richness. *Ecography*, 38(9), 913–921. <https://doi.org/10.1111/ecog.01129>
- Mod, H. K., & Luoto, M. (2016). Arctic shrubification mediates the impacts of warming climate on changes to tundra vegetation. *Environmental Research Letters*, 11, 124028. <https://doi.org/10.1088/1748-9326/11/12/124028>
- Morán-Ordóñez, A., Lahoz-Monfort, J. J., Elith, J., & Wintle, B. A. (2017). Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: What factors influence the reliability of predictions? *Global Ecology and Biogeography*, 26(3), 371–384. <https://doi.org/10.1111/geb.12545>
- Morris, W. F., & Ehrlén, J. (2015). Predicting changes in the distribution and abundance of species under environmental change. *Ecology Letters*, 18(3), 303–314. <https://doi.org/10.1111/ele.12410>
- Morrison, M. L., Timossi, I. C., & With, K. A. (1987). Development and testing of linear regression models predicting bird-habitat relationships. *The Journal of Wildlife Management*, 51(1), 247–253. <https://doi.org/10.2307/3801663>
- Myers-Smith, I. H., Elmendorf, S. C., Beck, P. S. A., Wilmking, M., Hallinger, M., Blok, D., Tape, K. D., Rayback, S. A., Macias-Fauria, M., Forbes, B. C., Speed, J. D. M., Boulanger-Lapointe, N., Rixen, C., Lévesque, E., Schmidt, N. M., Baittinger, C., Trant, A. J., Hermanutz, L., Collier, L. S., ... Vellend, M. (2015). Climate sensitivity of shrub growth across the tundra biome. *Nature Climate Change*, 5, 887–891. <https://doi.org/10.1038/NCLIMATE2697>
- Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., Macias-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. A., Schmidt, N. M., ... Hik, D. S. (2011). Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities. *Environmental Research Letters*, 6(4), <https://doi.org/10.1088/1748-9326/6/4/045509>
- Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S. A., Berner, L. T., Bhatt, U. S., Björkman, A. D., Blok, D., Bryn, A., Christiansen, C. T., Cornelissen, J. H. C., Cunliffe, A. M., Elmendorf, S. C., ... Wipf, S. (2020). Complexity revealed in the greening of the Arctic. *Nature Climate Change*, 10, 106–117. <https://doi.org/10.1038/s41558-019-0688-1>
- Nabe-Nielsen, J., Normand, S., Hui, F. K. C., Stewart, L., Bay, C., Nabe-Nielsen, L. I., & Schmidt, N. M. (2017). Plant community composition

- and species richness in the High Arctic tundra: From the present to the future. *Ecology and Evolution*, 7(23), 10233–10242. <https://doi.org/10.1002/ece3.3496>
- Nielsen, S. E., Johnson, C. J., Heard, D. C., & Boyce, M. S. (2005). Can models of presence-absence be used to scale abundance? Two case studies considering extremes in life history. *Ecography*, 28(2), 197–208. <https://doi.org/10.1111/j.0906-7590.2005.04002.x>
- Niittynen, P., Heikkinen, R. K., Aalto, J., Guisan, A., Kemppinen, J., & Luoto, M. (2020). Fine-scale tundra vegetation patterns are strongly related to winter thermal conditions. *Nature Climate Change*, 10, 1143–1148. <https://doi.org/10.1038/s41558-020-00916-4>
- Økland, R. H. (1995). Changes in the occurrence and abundance of plant species in a Norwegian boreal coniferous forest, 1988–1993. *Nordic Journal of Botany*, 15(4), 415–438. <https://doi.org/10.1111/j.1756-1051.1995.tb00174.x>
- Oldfather, M. F., Kling, M. M., Sheth, S. N., Emery, N. C., & Ackerly, D. D. (2020). Range edges in heterogeneous landscapes: Integrating geographic scale and climate complexity into range dynamics. *Global Change Biology*, 26(3), 1055–1067. <https://doi.org/10.1111/gcb.14897>
- Oppen, J., Normand, S., Bjorkman, A. D., Blach-Overgaard, A., Assmann, J. J., Forchhammer, M., Guéguen, M., & Nabe-Nielsen, J. (2021). Annual air temperature variability and biotic interactions explain tundra shrub species abundance. *Journal of Vegetation Science*, 32(2), e13009. <https://doi.org/10.1111/jvs.13009>
- Pacifici, M., Foden, W. B., Visconti, P., Watson, J. E. M., Butchart, S. H. M., Kovacs, K. M., Scheffers, B. R., Hole, D. G., Martin, T. G., Akçakaya, H. R., Corlett, R. T., Huntley, B., Bickford, D., Carr, J. A., Hoffmann, A. A., Midgley, G. F., Pearce-Kelly, P., Pearson, R. G., Williams, S. E., ... Rondinini, C. (2015). Assessing species vulnerability to climate change. *Nature Climate Change*, 5, 215–225. <https://doi.org/10.1038/nclimate2448>
- Paradis, M., Lévesque, E., & Boudreau, S. (2016). Greater effect of increasing shrub height on winter versus summer soil temperature. *Environmental Research Letters*, 11, 085005. <https://doi.org/10.1088/1748-9326/11/8/085005>
- Pearson, R. G., Phillips, S. J., Loranty, M. M., Beck, P. S. A., Damoulas, T., Knight, S. J., & Goetz, S. J. (2013). Shifts in Arctic vegetation and associated feedbacks under climate change. *Nature Climate Change*, 3(7), 673–677. <https://doi.org/10.1038/nclimate1858>
- Pearson, R. G., Thuiller, W., Araújo, M. B., Martínez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T. P., & Lees, D. C. (2006). Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33(10), 1704–1711. <https://doi.org/10.1111/j.1365-2699.2006.01460.x>
- Post, E., Alley, R. B., Christensen, T. R., Macias-Fauria, M., Forbes, B. C., Gooseff, M. N., Iler, A., Kerby, J. T., Laidre, K. L., Mann, M. E., Olofsson, J., Stroeve, J. C., Ulmer, F., Virginia, R. A., & Wang, M. (2019). The polar regions in a 2°C warmer world. *Science Advances*, 5(12), eaaw9883. <https://doi.org/10.1126/sciadv.aaw9883>
- Post, E., Pedersen, C., Wilmers, C. C., & Forchhammer, M. C. (2008). Warming, plant phenology and the spatial dimension of trophic mismatch for large herbivores. *Proceedings of the Royal Society B: Biological Sciences*, 275(1646), 2005–2013. <https://doi.org/10.1098/RSPB.2008.0463>
- Potts, J. M., & Elith, J. (2006). Comparing species abundance models. *Ecological Modelling*, 199(2), 153–163. <https://doi.org/10.1016/j.ecolmodel.2006.05.025>
- Pradervand, J.-N., Dubuis, A., Pellissier, L., Guisan, A., & Randin, C. (2014). Very high resolution environmental predictors in species distribution models: Moving beyond topography? *Progress in Physical Geography*, 38(1), 79–96. <https://doi.org/10.1177/0309133313512667>
- Prendin, A. L., Carrer, M., Karami, M., Hollesen, J., Bjerregaard Pedersen, N., Pividori, M., Treier, U. A., Westergaard-Nielsen, A., Elberling, B. O., & Normand, S. (2020). Immediate and carry-over effects of insect outbreaks on vegetation growth in West Greenland assessed from cells to satellite. *Journal of Biogeography*, 47(1), 87–100. <https://doi.org/10.1111/jbi.13644>
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Randin, C. F., Ashcroft, M. B., Bolliger, J., Cavender-Bares, J., Coops, N. C., Dullinger, S., Dirnböck, T., Eckert, S., Ellis, E., Fernández, N., Giuliani, G., Guisan, A., Jetz, W., Joost, S., Karger, D., Lembrechts, J., Lenoir, J., Luoto, M., Morin, X., ... Payne, D. (2020). Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment*, 239, 111626. <https://doi.org/10.1016/j.rse.2019.111626>
- Randin, C. F., Dirnböck, T., Dullinger, S., Zimmermann, N. E., Zappa, M., & Guisan, A. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography*, 33(10), 1689–1703. <https://doi.org/10.1111/j.1365-2699.2006.01466.x>
- Rapacciuolo, G., Roy, D. B., Gillings, S., Fox, R., Walker, K., & Purvis, A. (2012). Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. *PLoS One*, 7(7), e40212. <https://doi.org/10.1371/journal.pone.0040212>
- Raundrup, K., Josephine, N., Tøttrup, C., & Jacob, N.-N. (2018). Resource selection and drivers of habitat choice in West Greenland caribou. In *Industrial PhD: Movement patterns and habitat selection – insights from West Greenland caribou* (pp. 35–57). Department of Bioscience, Aarhus University.
- Raxworthy, C. J., Martínez-Meyer, E., Horning, N., Nussbaum, R. A., Schneider, G. E., Ortega-Huerta, M. A., & Peterson, A. T. (2003). Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, 426(6968), 837–841. <https://doi.org/10.1038/nature02205>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Aroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929. <https://doi.org/10.1111/ecog.02881>
- Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, 13(3), 243–251. <https://doi.org/10.1111/j.1472-4642.2007.00356.x>
- Rune, F. (2011). *Wild Flowers of Greenland - Grønlands Vilde Planter* (p. 187, 193). Gyldenlund Naturforlag.
- Santini, L., Benítez-López, A., Čengić, M., Maiorano, L., & Huijbregts, M. A. J. (2021). Assessing the reliability of species distribution projections in climate change research. *Diversity and Distributions*, 27(6), 1035–1050. <https://doi.org/10.1101/2020.06.10.143917>
- Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197–227. <https://doi.org/10.1023/A:1022648800760>
- Scherrer, D., & Körner, C. (2010). Infra-red thermometry of alpine landscapes challenges climatic warming projections. *Global Change Biology*, 16(9), 2602–2613. <https://doi.org/10.1111/j.1365-2486.2009.02122.x>
- Schratz, P., Muenchow, J., Iturrutxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>
- Schweiger, A. H., Irl, S. D. H., Steinbauer, M. J., Dengler, J., & Beierkuhnlein, C. (2016). Optimizing sampling approaches along ecological gradients. *Methods in Ecology and Evolution*, 7(4), 463–471. <https://doi.org/10.1111/2041-210X.12495>
- Smith, A. B., & Santos, M. J. (2020). Testing the ability of species distribution models to infer variable importance. *Ecography*, 43(12), 1801–1813. <https://doi.org/10.1111/ecog.05317>
- Sofaer, H. R., Jarnevich, C. S., Flather, C. H., & Kark, S. (2018). Misleading prioritizations from modelling range shifts under climate change. *Global Ecology and Biogeography*, 27(6), 658–666. <https://doi.org/10.1111/geb.12726>

- Staniczenko, P. P. A., Sivasubramaniam, P., Suttle, K. B., & Pearson, R. G. (2017). Linking macroecology and community ecology: Refining predictions of species distributions using biotic interaction networks. *Ecology Letters*, 20(6), 693–707. <https://doi.org/10.1111/ele.12770>
- Stewart, L., Alsos, I. G., Bay, C., Breen, A. L., Brochmann, C., Boulanger-Lapointe, N., Broennimann, O., Bültmann, H., Bøcher, P. K., Damgaard, C., Daniëls, F. J. A., Ehrich, D., Eidesen, P. B., Guisan, A., Jónsdóttir, I. S., Lenoir, J., le Roux, P. C., Lévesque, E., Luoto, M., ... Pellissier, L. (2016). The regional species richness and genetic diversity of Arctic vegetation reflect both past glaciations and current climate. *Global Ecology and Biogeography*, 25(4), 430–442. <https://doi.org/10.1111/geb.12424>
- Stewart, S. B., Elith, J., Fedrigo, M., Kasel, S., Roxburgh, S. H., Bennett, L. T., Chick, M., Fairman, T., Leonard, S., Kohout, M., Cripps, J. K., Durkin, L., & Nitschke, C. R. (2021). Climate extreme variables generated using monthly time-series data improve predicted distributions of plant species. *Ecography*, 44, 626–639. <https://doi.org/10.1111/ecog.05253>
- Sturm, M., & Douglas, T. (2005). Changing snow and shrub conditions affect albedo with global implications. *Journal of Geophysical Research*, 110(G1), G01004. <https://doi.org/10.1029/2005jg000013>
- Suvanto, S., Le, P. C., & Luoto, M. (2014). Arctic - alpine vegetation biomass is driven by fine - scale abiotic heterogeneity. *Geografiska Annaler: Series A, Physical Geography*, 96(4), 549–560. <https://doi.org/10.1111/geoa.12050>
- Tessarolo, G., Rangel, T. F., Araújo, M. B., & Hortal, J. (2014). Uncertainty associated with survey design in species distribution models. *Diversity and Distributions*, 20(11), 1258–1269. <https://doi.org/10.1111/ddi.12236>
- Thuiller, W., Albert, C. H., Dubuis, A., Randin, C., & Guisan, A. (2009). Variation in habitat suitability does not always relate to variation in species' plant functional traits. *Biology Letters*, 6(1), 120–123. <https://doi.org/10.1098/rsbl.2009.0669>
- Thuiller, W., Araújo, M. B., & Lavorel, S. (2003). Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, 14(5), 669–680. <https://doi.org/10.1111/j.1654-1103.2003.tb02199.x>
- Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N., & Zimmermann, N. E. (2019). Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications*, 10, 1446. <https://doi.org/10.1038/s41467-019-09519-w>
- Thuiller, W., Münkemüller, T., Schiffrers, K. H., Georges, D., Dullinger, S., Eckhart, V. M., Edwards, T. C., Gravel, D., Kunstler, G., Merow, C., Moore, K., Piedallu, C., Vissault, S., Zimmermann, N. E., Zurell, D., & Schurr, F. M. (2014). Does probability of occurrence relate to population dynamics? *Ecography*, 37(12), 1155–1166. <https://doi.org/10.1111/ecog.00836>
- Thuiller, W., Pollock, L. J., Gueguen, M., & Münkemüller, T. (2015). From species distributions to meta-communities. *Ecology Letters*, 18(12), 1321–1328. <https://doi.org/10.1111/ele.12526>
- Tomlinson, S., Lewandowski, W., Miller, B. P., Turner, S. R., & Elliott, C. P. (2020). High-resolution distribution modeling of a threatened short-range endemic plant informed by edaphic factors. *Ecology and Evolution*, 10(2), 763–777. <https://doi.org/10.1002/ece3.5933>
- Tulloch, A. I. T., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P. M. B., Possingham, H., Guisan, A., & Rhodes, J. R. (2016). Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation*, 199, 157–171. <https://doi.org/10.1016/j.biocon.2016.04.023>
- Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2021). Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecological Monographs*, 92(1), 0–3. <https://doi.org/10.1002/ecm.1486>
- Van Couwenberghe, R., Collet, C., Pierrat, J. C., Verheyen, K., & Gégout, J. C. (2013). Can species distribution models be used to describe plant abundance patterns? *Ecography*, 36(6), 665–674. <https://doi.org/10.1111/j.1600-0587.2012.07362.x>
- VanDerWal, J., Shoo, L. P., Johnson, C. N., & Williams, S. E. (2009). Abundance and the environmental niche: Environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, 174(2), 282–291. <https://doi.org/10.1086/600087>
- Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., Tjiputra, J., & Pellissier, L. (2022). A quantitative review of abundance-based species distribution models. *Ecography*, 2022(1), e05694. <https://doi.org/10.1111/ecog.05694>
- Walker, D. A., Daniëls, F. J. A., Alsos, I., Bhatt, U. S., Breen, A. L., Buchhorn, M., Bültmann, H., Druckenmiller, L. A., Edwards, M. E., Ehrich, D., Epstein, H. E., Gould, W. A., Ims, R. A., Meltotte, H., Reynolds, M. K., Sibik, J., Talbot, S. S., & Webber, P. J. (2016). Circumpolar Arctic vegetation: A hierarchic review and roadmap toward an internationally consistent approach to survey, archive and classify tundra plot data. *Environmental Research Letters*, 11, 055005. <https://doi.org/10.1088/1748-9326/11/5/055005>
- Weber, M. M., Stevens, R. D., Diniz-Filho, J. A. F., & Grelle, C. E. V. (2017). Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. *Ecography*, 40(7), 817–828. <https://doi.org/10.1111/ecog.02125>
- Werkowska, W., Márquez, A. L., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental Reviews*, 25(1), 127–133. <https://doi.org/10.1139/er-2016-0045>
- Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (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(Supplement 2), 19729–19736. <https://doi.org/10.1073/pnas.0901639106>
- Wis, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J.-A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E., Schmidt, N. M., ... Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88(1), 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., ... Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology and Evolution*, 33(10), 790–802. <https://doi.org/10.1016/j.tree.2018.08.001>
- Zimmermann, N. E., Yoccoz, N. G., Edwards, T. C., Meier, E. S., Thuiller, W., Guisan, A., Schmatz, D. R., & Pearman, P. B. (2009). Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences*, 106(Supplement 2), 19723–19728. <https://doi.org/10.1073/pnas.0901643106>
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., ... Merow, C. (2020). A standard protocol for reporting species distribution models. *Ecography*, 43(9), 1261–1277. <https://doi.org/10.1111/ecog.04960>
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14. <https://doi.org/10.1111/j.2041-210x.2009.00001.x>
- Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(4), 561–577. <https://doi.org/10.1093/clinchem/39.4.561>

BIOSKETCH

NI Chardon (<https://nathaliechardon.weebly.com>) combines extensive field work with spatial modelling approaches to understand what drives species geographic distributions and how these will shift with global changes. She is also interested in improving models of current and projected species spatial patterns by seeking solutions to common modelling problems. She works primarily in arctic, alpine and montane regions and is currently a Swiss National Science Foundation postdoctoral fellow at the University of British Columbia in Canada.

Author contributions: N.I.C., S.W. and S.N. conceived the presented idea. N.I.C. and J.N.-N. designed and directed the field surveys with support from all authors. N.I.C., J.N.-N and I.B.D.J. conducted the field work. N.I.C., M.G. and J.A. performed the computations with support from all authors. J.A. and all authors aided in interpreting the results, and N.I.C. wrote the paper with input from all authors.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Chardon, N. I., Nabe-Nielsen, J., Assmann, J. J., Dyrholm Jacobsen, I. B., Guéguen, M., Normand, S., & Wipf, S. (2022). High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords. *Diversity and Distributions*, 28, 956–975. <https://doi.org/10.1111/ddi.13498>