



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Faculty of Science, Medicine and Health - Papers

Faculty of Science, Medicine and Health

2016

Bayesian methods for comparing species physiological and ecological response curves

Michael B. Ashcroft

University of Wollongong, ashcroft@uow.edu.au

Angelica Casanova-Katny

Universidad Catolica de Temuco

Kerrie Mengersen

Queensland University of Technology

Todd N. Rosenstiel

Portland State University

Johanna Turnbull

University of Wollongong, jdt17@uowmail.edu.au

See next page for additional authors

Publication Details

Ashcroft, M. B., Casanova-Katny, A., Mengersen, K., Rosenstiel, T. N., Turnbull, J. D., Wasley, J., Waterman, M. J., Zuniga, G. E. & Robinson, S. A. (2016). Bayesian methods for comparing species physiological and ecological response curves. *Ecological Informatics*, 34 35-43.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:
research-pubs@uow.edu.au

Bayesian methods for comparing species physiological and ecological response curves

Abstract

Many ecological questions require information on species' optimal conditions or critical limits along environmental gradients. These attributes can be compared to answer questions on niche partitioning, species coexistence and niche conservatism. However, these comparisons are unconvincing when existing methods do not quantify the uncertainty in the attributes or rely on assumptions about the shape of species' responses to the environmental gradient. The aim of this study was to develop a model to quantify the uncertainty in the attributes of species response curves and allow them to be tested for substantive differences without making assumptions about the shape of the responses. We developed a model that used Bayesian penalised splines to produce and compare response curves for any two given species. These splines allow the data to determine the shape of the response curves rather than making a priori assumptions. The models were implemented using the R2OpenBUGS package for R, which uses Markov Chain Monte Carlo simulation to repetitively fit alternative response curves to the data. As each iteration produces a different curve that varies in optima, niche breadth and limits, the model estimates the uncertainty in each of these attributes and the probability that the two curves are different. The models were tested using two datasets of mosses from Antarctica. Both datasets had a high degree of scatter, which is typical of ecological research. This noise resulted in considerable uncertainty in the optima and limits of species response curves, but substantive differences were found. *Schistidium antarctici* was found to inhabit wetter habitats than *Ceratodon purpureus*, and *Polytrichastrum alpinum* had a lower optimal temperature for photosynthesis than *Chorisodontium aciphyllum* under high light conditions. Our study highlights the importance of considering uncertainty in physiological optima and other attributes of species response curves. We found that apparent differences in optima of 7.5 °C were not necessarily substantive when dealing with noisy ecological data, and it is necessary to consider the uncertainty in attributes when comparing the curves for different species. The model introduced here could increase the robustness of research on niche partitioning, species coexistence and niche conservatism.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Ashcroft, M. B., Casanova-Katny, A., Mengersen, K., Rosenstiel, T. N., Turnbull, J. D., Wasley, J., Waterman, M. J., Zuniga, G. E. & Robinson, S. A. (2016). Bayesian methods for comparing species physiological and ecological response curves. *Ecological Informatics*, 34 35-43.

Authors

Michael B. Ashcroft, Angelica Casanova-Katny, Kerrie Mengersen, Todd N. Rosenstiel, Johanna Turnbull, Jane Wasley, Melinda J. Waterman, Gustavo E. Zuniga, and Sharon A. Robinson

Bayesian methods for comparing species physiological and ecological response curves

Running title: Comparing species response curves

Word count: 6367

Michael B. Ashcroft ^{a*},Angélica Casanova-Katny ^b,Kerrie Mengersen ^c,Todd N. Rosenstiel ^d,Johanna D. Turnbull ^a,Jane Wasley ^{a,e},Melinda J. Waterman ^{a,f},Gustavo E. Zúñiga ^f,Sharon A. Robinson ^a

^a Centre for Sustainable Ecosystem Solutions, School of Biological Sciences, University of Wollongong, Wollongong, New South Wales, 2522, Australia.

^b Núcleo de Estudios Ambientales, Universidad Católica de Temuco, Casilla 15-D, Temuco, Chile

^c Mathematical Sciences, Science and Engineering Faculty, Queensland University of Technology, Brisbane, Queensland, 4001, Australia.

^d Center for Life in Extreme Environments, Department of Biology, Portland State University, Portland, Oregon, 97207, USA.

^e Australian Antarctic Division, Department of Environment, Kingston, Tasmania, 7050, Australia.

^fDepartamento de Biología and CEDENNA, Universidad de Santiago de Chile, Alameda 3363, Estación Central, Santiago, Chile.

* Corresponding author. Email: ashcroft@uow.edu.au

Abstract

Many ecological questions require information on species' optimal conditions or critical limits along environmental gradients. These attributes can be compared to answer questions on niche partitioning, species coexistence and niche conservatism. However, these comparisons are unconvincing when existing methods do not quantify the uncertainty in the attributes or rely on assumptions about the shape of species' responses to the environmental gradient. The aim of this study was to develop a model to quantify the uncertainty in the attributes of species response curves and allow them to be tested for substantive differences without making assumptions about the shape of the responses. We developed a model that used Bayesian penalised splines to produce and compare response curves for any two given species. These splines allow the data to determine the shape of the response curves rather than making *a priori* assumptions. The models were implemented using the R2OpenBUGS package for R, which uses Markov Chain Monte Carlo simulation to repetitively fit alternative response curves to the data. As each iteration produces a different curve that varies in optima, niche breadth and limits, the model estimates the uncertainty in each of these attributes and the probability that the two curves are different. The models were tested using two datasets of mosses from Antarctica. Both datasets had a high degree of scatter, which is typical of ecological research. This noise resulted in considerable uncertainty in the optima and limits of species response curves, but substantive differences were found. *Schistidium antarctici* was found to inhabit wetter habitats than *Ceratodon purpureus*, and *Polytrichastrum alpinum* had a lower optimal temperature for photosynthesis than *Chorisodontium aciphyllum* under high light conditions. Our study highlights the importance of considering uncertainty in physiological optima and other attributes of species response curves. We found that apparent differences in optima of 7.5 °C were not necessarily substantive when dealing with noisy ecological data, and it is necessary to consider the

uncertainty in attributes when comparing the curves for different species. The model introduced here could increase the robustness of research on niche partitioning, species coexistence and niche conservatism.

Keywords: Antarctic moss; community ecology; niche partitioning; photosynthesis; physiological response; uncertainty

1. Introduction

How species respond to environmental gradients is a fundamentally important topic in ecology, biogeography and evolution (Quintero & Wiens 2013). These responses, known as species response curves, are at the foundation of species distribution models (Guisan & Zimmermann 2000; Elith & Leathwick 2009), which have been increasingly used over the last 15 years to make ecological inferences based on species environmental niches. Recent machine-learning approaches are in widespread use because of their perceived superior performance in predicting species distributions (Elith *et al.* 2006), but the actual response curves in these complex models can often be difficult to visualise and interpret ecologically (Elith *et al.* 2005) and it can be challenging to estimate uncertainty when there is no underlying parametric model. These issues are problematic as evaluating species response curves is an integral part of interpreting species distribution models and assessing their ecological validity (Austin 2002, 2007).

However, despite their fundamental role in species distribution models, interest in species response curves extends far beyond this specific application. Quantifying the attributes of species response curves, such as the optima, niche breadth or limits where the response is above a certain threshold (Fig. 1a), is crucial for answering many physiological, ecological and evolutionary questions even if the objective is not to predict or explain species distributions *per se*. For example, scientists are often interested in species coexistence, niche separation or overlap (Silvertown 2004); the relative position or critical limits of species' niches on environmental gradients (Hernandez & Mulla 2008; Sinervo *et al.* 2010); niche conservatism over space and time (Peterson *et al.* 2009; Pellissier *et al.* 2013); or niche shifts of invasive species in new environments (Hill *et al.* 2013). For these topics, the shape and attributes of species response curves are the focus of the research, and it is often necessary to

compare and test for differences between different curves (e.g. niche conservatism, niche shifts of invasive species, niche separation and overlap).

The attributes of species response curves are difficult to compare using many common statistical models. Response curves can be presented with 95% confidence intervals (e.g. Fig. 1b), but this only quantifies the uncertainty in the response at a given position on the environmental gradient. There is generally no indication of uncertainty for the optima, limits or breadth of the curve as a whole, and there are limitless response curves that fit within the 95% confidence intervals but vary in optima or other attributes (Fig. 1b). This is an important limitation because it is problematic to test if the optima of two curves are different if we have not estimated their uncertainty (Hernandez & Mulla 2008). This would be analogous to directly comparing the absolute means of two data sets without considering the variances using Student's t-test or equivalent.

There are existing methods that can quantify the uncertainty in the optima of a response curve, but these generally need to make assumptions about the shape of the response curve and focus only on the attribute of interest (Vetaas 2000; Björnsson *et al.* 2001; Dreyer *et al.* 2001; Hernandez & Mulla 2008). For example, some methods rely on the assumption that there are Gaussian responses to environmental gradients, even though evidence suggests skewed or non-normal response curves are common, and even more complex shapes such as multi-modal curves are possible when dealing with realised rather than fundamental niches (Austin 2002, 2007; Oksanen & Minchin 2002). Other methods allow skewed responses by fitting parametric functions (e.g. Higgins *et al.* 2014), and this also allows scope for the uncertainty in some attributes of response curves to be quantified. However, there is still scope to estimate the uncertainty in the optima and other attributes of species response curves using a semi-parametric method that does not make *a priori* assumptions about the shape of a given response.

Bayesian models have the potential to address this deficiency. Bayesian models (McCarthy 2007) can be fitted using Markov Chain Monte Carlo (MCMC) random sampling which iteratively fits alternative species response curves that can explain the observed data. It does this by simulating sets of parameter values from their Bayesian posterior distributions, then generating values from the resultant response curves. After repeating this thousands of times, the collection of simulated values can be used to estimate not just the mean and 95% credible intervals (the Bayesian equivalent of confidence intervals; McCarthy 2007) for the species response curves (Fig. 1b), but also the mean and credible intervals for other attributes of the curve, such as the optima, niche breadth and limits.

In this paper we adapt the Bayesian penalised splines of Crainiceanu *et al.* (2005) to predict species response curves. Similarly to the Generalised Additive Models commonly used in ecological studies (GAMs; Hastie & Tibshirani 1990; Guisan & Zimmermann 2000; Austin 2002; Elith *et al.* 2006), the shape of penalised splines is determined by the data, so the shape of the response curves can be skewed or even multi-modal. However, the curves are smooth unlike methods such as classification and regression trees (Elith *et al.* 2008) and Maxent (Phillips *et al.* 2006), which can have discrete steps. The advantage of the Bayesian implementation of penalised splines is that the iteration involved allows us greater flexibility to examine the uncertainty in many attributes of the fitted curves. Our model was designed to estimate the means and 95% credible intervals for the peak and optima, as well as the niche breadth and limits based on a threshold of 80% of the peak (Fig. 1a). The model could also be modified to examine other attributes of species response curves or to use predefined response shapes such as Gaussian, Beta distribution, Huisman—Olf—Fresco (HOF) models (Oksanen & Minchin 2002) or the Arrhenius equation. If applied to parametric models it could be used to estimate the uncertainty in the parameter estimates.

The overall objective of this study was to develop a model that allowed users to quantify uncertainty in a wide range of attributes of species response curves and enable comparison with other curves. This represents an important advance from prior studies that primarily compare optima or limits of species response curves without considering the uncertainty in those estimates, or which estimate the uncertainty in a limited number of attributes by assuming, *a priori*, the shape of the response curve. Additionally, such a model also needs to be inherently flexible so that it can be applied to a wide variety of datasets or modified to address similar research questions.

The model is demonstrated here using two datasets of mosses in Antarctica. The first example uses an unrestricted continuous response variable, and examines whether the optimal temperature for photosynthesis varies between two species of moss from maritime Antarctica. The second uses a response variable that is constrained to a range of zero to one (proportion of presences in samples along a moisture gradient), and examines whether two species of moss from continental Antarctica have different moisture optima. The two examples illustrate how the model can be used to test for differences between the response curves of different species using contrasting types of response variables.

2. Materials and methods

2.1 MODEL DEVELOPMENT

The Bayesian models were developed in R (R Core Team 2014) using the R2OpenBUGS package. This requires the OpenBUGS software application to be installed, which is a newer open-source version of WinBUGS. All software is freely available for download. The models require a data file (cwc2003data.csv, etrdatahigh.csv or etrdatalow.csv in Supplementary Material), a file which implements the OpenBUGS model (SplineModels.txt or SplineModelsLogit.txt in Supplementary Material) and a set of R commands to specify model

parameters, run the model and display results (Rcommands.txt in Supplementary Material). The analyses in this article can be repeated by downloading the files in the Supplementary Material, updating the 4th line of Rcommands.txt to point to the directory where the files are located, and adjusting lines 6–10 to select the desired case study. As the models are stochastic, results will not necessarily be identical every time the models are run.

Bayesian penalised splines are semiparametric models that had already been implemented in a model for WinBUGS in R by Crainiceanu *et al.* (2005), and this formed the basis for the model included here. Spline models are developed by fitting a number of polynomial segments that are joined together at knot points spread evenly along the environmental gradient. With four knots there is one polynomial from 0% to 20%, another from 20% to 40%, and so on. The polynomials join together to form a continuous smooth curve (spline) that has flexibility in its shape. More knots allow more complex shapes to be fitted, but more data is needed to reduce the potential for overfitting. The original model used 20 knots, although this was reduced to four in our model. Having too many knots allows complex responses that may be difficult to explain, and increases the uncertainty in the attributes of response curves as they can take more varied forms (see Results section).

The original model of Crainiceanu *et al.* (2005) fitted one spline based on a dataset containing a single response and covariate, but was modified to produce two splines for two different datasets. The model initially only predicted the response for the environmental conditions in the samples. To ensure complete response curves were produced for both species the model was modified to accept three sets of environmental samples: *X1* for the samples for the first response, *X2* for the samples for the second response, and *X* for the complete gradient. The complete gradient was chosen such that it extended 2% past the minimum and maximum environmental values from both data sets and was broken up into 300 samples to increase the precision with which the attributes could be estimated.

In each iteration of the MCMC algorithm, the two response curves were produced, and the peak, optima, niche limits and breadth were then calculated deterministically from the curves. A posterior mean and associated percentiles for these attributes were then calculated based on a large number of iterations. Similarly, the step function in OpenBUGS was used to estimate the probability that the attribute of one curve was larger than that of the other. This is referred to as a P-value in the remainder of this paper, although it is acknowledged that it does not have the same interpretation as a P-value obtained under a frequentist paradigm. We considered a result to be substantive if the P-value was less than 5%, which occurs if the attribute for one curve was higher than the other in 95% of the iterations. We acknowledge this is different to traditional frequentist P-values, where a P-value of 5% indicates that there is only a 5% chance that the results could have been obtained if there was actually no difference. As our model estimates a P-value, thresholds other than 5% could also be used.

2.2 CASE STUDY 1: CONTINUOUS DATA OF PHOTOSYNTHETIC RATE

The model was tested on an effectively unbounded continuous response variable (photosynthetic rate) using two species of moss from Ardley Island in the maritime Antarctic (62.2 °S, 58.9 °W), *Polytrichastrum alpinum* (Hedw.) G.L. Sm. and *Chorisodontium aciphyllum* (Hook.f. & Wilson) Broth. These two species of moss coexist in some communities and we were interested in comparing their photosynthetic responses to temperature. The two species have similar habitats, but *C. aciphyllum* has a very compact turf structure with its gametophytes densely packed together while the *P. alpinum* turf is more open, resulting in more space between adjacent gametophytes (see Fig 2). Antarctic moss turves can maintain temperatures well above air temperature, especially in sun-exposed but wind-sheltered sites (Smith 1988; Bramley-Alves *et al.* 2014). As a more open structure allows wind to penetrate into the moss beds and potentially keep temperatures lower, we

were interested in whether the optimal temperature for photosynthesis for *P. alpinum* was lower than that for *C. aciphyllum*.

Samples of moss tissue were collected in January 2015 and maintained in plastic containers under field conditions for up to 2 days prior to measurements being made in the laboratory of the Chilean Escudero Base on King George Island. Photosynthetic activity was measured for both species in a water bath with temperatures varying from 1 to 44 °C. Hydrated moss plugs (1 cm²; n = 6) were contained in aluminium cups to allow heat transfer within the water bath but prevent submergence and ensure the photosynthetic surface of the moss remained exposed to air. Moss was induced prior to measurements using a Dracast LED160a light source (Dracast, San Jose, CA, USA) which also provided the actinic light during the experiment.

Measurements were made under both low (non-saturating; 400 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) and high (saturating; 1000 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) light. Surface photosynthetically active radiation (PAR) was measured using an Apogee mq-200 quantum light sensor (Apogee Instruments, Logan UT, USA). Thermocouples were used to measure temperature of the photosynthetic tissue at the time of each photosynthetic rate measurement. Photosynthetic rates were quantified using chlorophyll fluorescence (measured as Electron Transport Rate; ETR; $\mu\text{mol e m}^{-2} \text{s}^{-1}$) with a Walz MINI-PAM Photosynthesis Yield Analyser (Walz, Effeltrich, Germany). ETR was calculated using the equation $\text{ETR} = \phi\text{PSII} \times \text{PAR} \times 0.5 \times 0.84$ (Maxwell & Johnson 2000), where ϕPSII is the quantum yield of photosystem II, estimated using chlorophyll fluorescence and PAR is the photosynthetically active radiation. This widely used equation makes the assumption that the leaf, or gametophyte, absorbs 84% of the incident light and that this light is shared equally between the two photosystems.

The Bayesian model was used to estimate the peak ETR, the optimal temperature, the temperature limits corresponding to a threshold of 80% of the maximum ETR, and the

breadth of the response curve between these limits. The model also tested if the two species were significantly different in any of these attributes. The model was repeated using the two response curves for each species to test whether the attributes for either species were substantively different under low and high light.

2.3 CASE STUDY 2: PROBABILITY OF PRESENCE ALONG MOISTURE GRADIENT

Sixty quadrats were surveyed in January and February 2003 across two long-term monitoring sites near the Australian Casey station (66.28 °S, 110.53 °E) in the Windmill Islands, East Antarctica (detailed site descriptions in Wasley *et al.* 2012 and Bramley-Alves *et al.* 2015). The monitoring of these sites forms Australia's State of the Environment Indicator 72 (<https://data.aad.gov.au/aadc/soe/>). Each site contained 10 transects along a moisture gradient, with three quadrats (25 × 25 cm) per transect. Quadrats were positioned in three distinct vegetation communities; bryophyte dominated, lichen dominated, and a transitional community in between (mix of moribund moss and encrusting lichens). In each quadrat, nine small samples (tweezer pinches of vegetation) were taken and scored for presence or absence of bryophyte and lichen species. These data were linearly scaled to a range of zero (not found in any of the nine samples) to one (found in all nine samples). In this analysis, we focus on how two of the dominant moss species (*Schistidium antarctici* (Card.) L. Savic. & Smirn. (syn. *Grimmia antarctici* (Card.)) and *Ceratodon purpureus* (Hedw.) Brid.) respond to a water availability gradient, and test the hypothesis that there is niche separation along this gradient.

Community water content (CWC) was estimated by inserting a sponge into the vegetation layer adjacent to each quadrat and measuring the amount of water it absorbed over 24 h (as described in Lucieer *et al.* 2014). This one-off measurement of water content may not represent the long-term water availability in any quadrat, but does serve to differentiate

the moistest and driest quadrats and is a relative estimate of water availability in the vegetation where the sponge was placed.

To model the proportion of samples where the species was present some minor changes were needed in the model. Namely, as the response variable was constrained to a range of zero to one a logit transform was added to four lines in `SplineModelsLogit.txt` (see Supplementary Material) to replace the corresponding lines in `SplineModels.txt`. These changes ensured the predicted response was not outside the range of zero to one after back transformation. The data file was replaced (`cwc2003data.csv` in Supplementary Material) and `Rcommands.txt` was modified to refer to the new data and model. The modified `SplineModelsLogit.txt` model would be suitable for analysing a range of presence-absence data where the response is constrained to zero to one, or percentage data, such as humidity, where the response is constrained to 0–100% (e.g. Ashcroft & Gollan, 2012).

3. Results

3.1 PHOTOSYNTHETIC TEMPERATURE RESPONSE CURVES

The ETR data for both species under both high and low light conditions contained a high degree of scatter, and there was considerable uncertainty in all the attributes for each of the photosynthetic temperature response curves (Fig. 3, S1). The curves were somewhat skewed and contained additional points of inflection that were not consistent with Gaussian response curves. With 20 knots the response curves contained additional complexity, and this increased the uncertainty in where the optima were located (Fig. S2). We therefore focus on results from models with four knots in this paper, although the method itself is equally valid with 20 knots and could be useful if there is sufficient data with less noise.

Under high light conditions, the peak ETR for *P. alpinum* was higher than that for *C. aciphyllum*, but not significantly so (mean peaks of 200.2 and 186.2 $\mu\text{mol e m}^{-2} \text{s}^{-1}$; $P =$

0.096; Fig. 3a). The optimal temperature for photosynthesis in *P. alpinum* was significantly lower than for *C. aciphyllum* (mean optima of 21.7 °C and 28.5 °C; $P = 0.012$), but there were no significant differences in the limits ($P = 0.122, 0.266$) or the breadth of the curves ($P = 0.221$; Fig. 3a).

Under low light conditions the peaks were reversed, with the peak for photosynthesis in *P. alpinum* significantly lower than that for *C. aciphyllum* (mean peaks of 86.6 and 95.4 $\mu\text{mol e m}^{-2} \text{ s}^{-1}$; $P = 0.022$; Fig. 3b). The optima were not significantly different (mean optima of 17.4 °C and 24.9 °C; $P = 0.055$) but there was a difference in the upper limits (mean upper limits of 31.8 °C and 36.9 °C; $P = 0.031$). The differences in lower limits and breadths were not significantly different ($P = 0.588, 0.247$; Fig. 3b). The hypothesis that the temperature optima of *P. alpinum* would be lower than that of *C. aciphyllum* was supported by both data sets, but it was only significant under high light conditions (low light 95% CI 12.3–25.0 °C, 20.9–28.6 °C, $P = 0.055$; high light 95% CI 19.2–25.0 °C, 24.2–34.3 °C, $P = 0.012$; Fig. 3).

When the same data were analysed to compare the low and high light photosynthetic temperature responses for each species (Fig. S1), the peaks and lower limits were significantly higher under high light conditions ($P < 0.05$). The temperature optima were also higher under high light conditions for both species, but these differences were not significant ($P = 0.121, 0.074$; Fig. S1). Both curves were noticeably flatter under low light conditions.

3.2 PRESENCE-ABSENCE ALONG A MOISTURE GRADIENT

The response curves for the presence of *C. purpureus* and *S. antarctici* along a moisture gradient were significantly different in terms of optima, peak, limits and breadth ($P \leq 0.002$; Fig. 4). *Ceratodon purpureus* presence had an optima at relatively low water availability and was absent from the moister quadrats, while *S. antarctici* was much more prevalent in the moister quadrats. Once again the response curves were noticeably skewed indicating that

Gaussian responses would not be a good fit. In terms of the niche separation hypothesis, *S. antarctici* inhabited significantly wetter microhabitats than *C. purpureus* (95% CI 0.27–0.57, >1.28, $P < 0.001$, Fig. 4). This finding is in keeping with previous physiological studies indicating that *C. purpureus* is more desiccation tolerant than *S. antarctici* (Robinson *et al.* 2000) but less tolerant of submergence (Wasley *et al.* 2006).

4. Discussion

A flexible and robust model has been successfully developed as a reliable strategy to compare species response curves whilst also incorporating uncertainties in the data sets. Ecological data from two case studies on Antarctic moss was visualised and interpreted using this model allowing certain ecological niches to be identified.

The data used in this study had a high degree of scatter, which is typical of many ecological studies. The models identified large differences in the attributes of the species response curves, but not all of these differences were significant given the uncertainty in the data and the semiparametric nature of the fitted model. For example, the mean low light optima for *P. alpinum* (17.4 °C; Fig. 3b) was 7.5 °C lower than that for *C. aciphyllum* (24.9 °C), but the models suggested this difference was marginally non-substantive using a traditional threshold of $\alpha = 0.05$ ($P = 0.055$) and there was overlap in the 95% credible intervals (12.3–25.0 and 20.9–28.6 respectively). This example highlights the importance of considering uncertainty in the optima and other attributes of species response curves. Our models produce a P-value for the probability the two optima are different, and with minor modifications could produce a 95% credible interval for what the difference is. This is much more informative than taking no account of the uncertainty and simply comparing the estimates of optima, and asserting a difference of 7.5 °C. Our P-values highlight it would be premature to make this conclusion with the given data. A larger sample size or broader

temperature range would be needed to reduce uncertainty and prove that differences between the curves are substantive.

The estimated uncertainty in optima is even higher if more complex response curves are fitted, for example by using 20 knots instead of four (Fig. S2). The uncertainty is especially high near the extremities of environmental gradients where it is uncertain if splines may continue to increase or decrease. To reduce this uncertainty, data should be gathered over as wide a range of environmental conditions as possible, however reducing the number of knots also limits the amount of extrapolation that is possible.

Numerous other approaches have been used to estimate the attributes of species response curves or compare them between different species. Some researchers have compared response curves by classifying them into a limited number of predefined categories based on the approximate position of the optima or the shape of the curve (Oksanen & Minchin 2002; Horsák 2006). Others have estimated the uncertainty in the optima or compared curves by assuming the shape of the response curve is known (Vetaas 2000; Björnsson *et al.* 2001; Dreyer *et al.* 2001; Hernandez & Mulla 2008). Another approach is to compare entire curves quantitatively but without considering uncertainty in the curves (Warren *et al.* 2008; Hill *et al.* 2013). Similarly, thermal tolerance limits are often determined from observational data without any estimate of uncertainty (Pellissier *et al.* 2013; Gouveia *et al.* 2014). The Bayesian semi-parametric model used here is advantageous because it quantifies the uncertainty in the attributes of species response curves without assuming a rigorous (parametric) shape of the response. This will increase the robustness of ecological studies on niche conservatism, niche partitioning and species coexistence. The method could also be extended to quantify the uncertainty in other attributes of species response curves, such as the probability the curve is above a certain threshold, determining the ranks of

multiple curves, or testing multiple curves for substantive differences if one species is used as a reference class (McCarthy 2007).

4.1 LIMITATIONS AND FUTURE ENHANCEMENTS

The model as it is currently implemented is restricted to a single covariate. As our photosynthesis results illustrate, the optima and limits on any gradient (e.g. temperature) may be affected by other conditions (e.g. light). For example, the optima we observed under low light were lower than the optima under high light (Fig. S1), even if these differences were marginally non-substantive ($P = 0.074, 0.121$). The shapes of the curves were also different: moss ETR did not change dramatically with temperature under low light conditions (flat red curves in Fig. S1) but did increase notably near the optimal temperature under high light conditions. It is possible to examine the optima along a single environmental gradient if other factors are held constant, but if multiple environmental factors are varied then the position of the optima may not be constant on any individual gradient.

Similarly, the current implementation only caters for a single pair of species. However, if one species is used as a reference then multiple species can be included in a Bayesian ANOVA analysis with each tested for substantive differences against the reference species (McCarthy 2007). This is not implemented in the current script but would be a relatively straight forward extension.

As noted in the methods section the ETR calculations were based on an assumption that turf absorbance was 0.84 for both species. This may not be the case. While this is not important for the methodology introduced in this paper, it may affect the comparisons of the peak in the first case study. If the actual absorbance has not been measured then it would be more correct to compare different species on the basis of relative ETR (scaled to % maximum ETR). This could easily be accommodated by applying the techniques used in the

presence/absence example where data are scaled to a range of zero to one, or by placing a prior distribution on the value to represent this uncertainty. This prior will then induce associated uncertainty in the posterior estimates and associated comparisons and inferences.

Another limitation is that the model does not predict optima that are outside the range of the environmental gradient examined. For example, it appears that the peak for *S. antarctici* occurred at a $\log(\text{CWC})$ higher than 1.28 (Fig. 4). It is relatively simple to extend the range of the environmental gradient in the model to make predictions at higher values but this would be highly uncertain, especially given the semiparametric setup of the model, and was deliberately avoided. In our study it is not clear if the curve would begin to fall at higher water contents, or whether it becomes limited by water depth, substrate availability or flow rate once the moss is completely saturated. To accurately predict optima or limits that are outside the range of observations it would be better to collect data over a broader range of conditions. Any 95% credible intervals that include either of the extremities on the environmental gradient may not be accurate and should be interpreted cautiously.

Despite these limitations the models developed in this study provide ecologists and physiologists with a powerful tool for assessing species response curves along single environmental gradients. The methods we present here allow quantification of attribute uncertainty, which represents an important advance of studies that currently compare species response curves qualitatively or subjectively without considering the uncertainties involved.

Acknowledgements

This research was primarily funded through Australian Research Council DP110101714 and Australian Antarctic Science grants AAS1313 and AAS4046 with logistical support from the Australian Antarctic Division for visits to Casey Station and Instituto Antártico Chileno (INACH) for Julio Escudero Station. ACK was supported by the grant FONDECYT

1120895, and INACH FR 0112 and the VRIDEI, USACH; GEZ was supported by the grant FONDECYT 1140189, CEDENNA and Proyectos Basales, VRIDEI, USACH. TNR was supported by the U.S. National Science Foundation (grants 1341742 and 1258225). KM acknowledges support from ARC DP and COE funding schemes and from the QUT Institute for Future Environments. MW held Australian Postgraduate and AINSE Awards during the time of the study. JW held an APA at the time of field data collection.

Data accessibility

The data and models used in this study are available in the Supplementary Material.

References

- Ashcroft, M.B. & Gollan, J.R. (2012) Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200 × 300 km) and diverse region. *International Journal of Climatology*, **32**, 2134–2148.
- Austin, M.P. (2002) Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101–118.
- Austin, M. (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.
- Bramley-Alves, J., King, D.H., Robinson S.A. & Miller, R.E. (2014) Dominating the Antarctic environment: Bryophytes in a time of change. Chapter 17. *Photosynthesis of Bryophytes and Early Land Plants, Advances in Photosynthesis and Respiration 37* (eds D.T. Hanson & S.K. Rice), pp. 309–324. Springer, Dordrecht, Netherlands.
- Bramley-Alves, J., Wanek, W., French, K. & Robinson, S.A. (2015) Moss $\delta^{13}\text{C}$: An accurate proxy for past water environments in polar regions. *Global Change Biology*, **21**, 2454–2464.
- Björnsson, B., Steinarsson, A. & Oddgeirsson, M. (2001) Optimal temperature for growth and feed conversion of immature cod (*Gadus morhua* L.). *ICES Journal of Marine Science*, **58**, 29–38.
- Crainiceanu, C.M., Ruppert, D. & Wand, M.P. (2005) Bayesian analysis for penalized spline regression using WinBUGS. *Journal of Statistical Software*, **14**, 1–24.
- Dreyer, E., Le Roux, X., Montpied, P., Daudet, F.A. & Masson, F. (2001) Temperature response of leaf photosynthetic capacity in seedlings from seven temperate tree species. *Tree Physiology*, **21**, 223–232.

- 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.
- Elith, J., Ferrier, S., Huettmann, F. & Leathwick, J. (2005) The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. *Ecological Modelling*, **186**, 280–289.
- Elith, J., Graham, C., Anderson, R., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S., Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.
- Gouveia, S.F., Hortal, J., Tejedo, M., Duarte, H., Cassemiro, F.A.S., Navas, C.A. & Diniz-Filho, J.A.F (2014) Climatic niche at physiological and macroecological scales: The thermal tolerance–geographical range interface and niche dimensionality. *Global Ecology and Biogeography*, **23**, 446–456.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hastie, T.J. & Tibshirani, R.J. (1990) *Generalized Additive Models*. Chapman and Hall, London.
- Hernandez, J.A. & Mulla, D.J. (2008) Estimating uncertainty of economically optimum fertilizer rates. *Agronomy Journal*, **100**, 1221–1229.

- Higgins, J.K., MacLean, H.J., Buckley, L.B. & Kingsolver, J.G. (2014) Geographic differences and microevolutionary changes in thermal sensitivity of butterfly larvae in response to climate. *Functional Ecology*, **28**, 982–989.
- Hill, M.P., Chown, S.L. & Hoffmann, A.A. (2013) A predicted niche shift corresponds with increased thermal resistance in an invasive mite, *Halotydeus destructor*. *Global Ecology and Biogeography*, **22**, 942–951.
- Horsák, M. (2006) Mollusc community patterns and species response curves along a mineral richness gradient: a case study in fens. *Journal of Biogeography*, **33**, 98–107.
- Lucieer, A., Turner, D., King, D.H. & Robinson, S.A. (2014) Using an Unmanned Aerial Vehicle (UAV) to capture micro-topography of Antarctic moss beds. *International Journal of Applied Earth Observation and Geoinformation*, **27**, 53–62.
- Maxwell K. & Johnson G.N. (2000) Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany*, **51**, 659–668.
- McCarthy, M.A. (2007) *Bayesian Methods for Ecology*. Cambridge University Press, Cambridge, UK.
- Oksanen, J. & Minchin, P.R. (2002) Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling*, **157**, 119–129
- Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, E.S., Zimmermann, N.E., Randin, C.F., Thuiller, W., Garraud, L., Van Es, J. & Guisan, A. (2013) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. *Global Ecology and Biogeography*, **22**, 933–941.
- Peterson, A.T., Soberón, J & Sánchez-Cordero, V. (1999) Conservatism of Ecological Niches in Evolutionary Time. *Science*, **285**, 1265–1267.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

- Quintero, I. & Wiens, J.J. (2013) What determines the climatic niche width of species? The role of spatial and temporal climatic variation in three vertebrate clades. *Global Ecology and Biogeography*, **22**, 422–432.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org> [accessed 7 April 2015].
- Robinson, S.A., Wasley, J., Popp, M. & Lovelock, C.E. (2000) Desiccation tolerance of three moss species from continental Antarctica. *Australian Journal of Plant Physiology*, **27**, 379–388.
- Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*, **19**, 605–611.
- Sinervo, B., Méndez-de-la-Cruz, F., Miles, D.B., Heulin, B., Bastiaans, E., Cruz, M.V., Lara-Resendiz, R., Martínez-Méndez, N., Calderón-Espinosa, M.L., Meza-Lázaro, R.N., Gadsden, H., Avila, L.J., Morando, M., De la Riva, I.J., Sepulveda, P.V., Rocha, C.F.D., Ibarngüengoytia, N., Puntriano, C.A., Massot, M., Lepetz, V., Oksanen, T.A., Chapple, D.G., Bauer, A.M., Branch, W.R., Clobert, J. & Sites Jr., J.W. (2010) Erosion of lizard diversity by climate change and altered thermal niches. *Science*, **328**, 894–899.
- Smith, R.I.L. (1988) Recording bryophyte microclimate in remote and severe environments. *Bryological methods workshop* (ed J. Glime), pp. 275–284. The Hattori Botanical Laboratory, Halle.
- Vetaas, O.R. (2000) Comparing species temperature response curves: population density versus second-hand data. *Journal of Vegetation Science*, **11**, 659–666.
- Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

Wasley, J., Robinson, S.A., Lovelock, C.E. & Popp, M. (2006) Some like it wet—biological characteristics underpinning tolerance of extreme water stress events in Antarctic bryophytes. *Functional Plant Biology*, **33**, 443–455.

Wasley, J., Robinson, S.A., Turnbull, J.D., King, D.H., Wanek, W. & Popp, M. (2012) Bryophyte species composition over moisture gradients in the Windmill Islands, East Antarctica: development of a baseline for monitoring climate change impacts *Biodiversity*, **13**, 257–264.

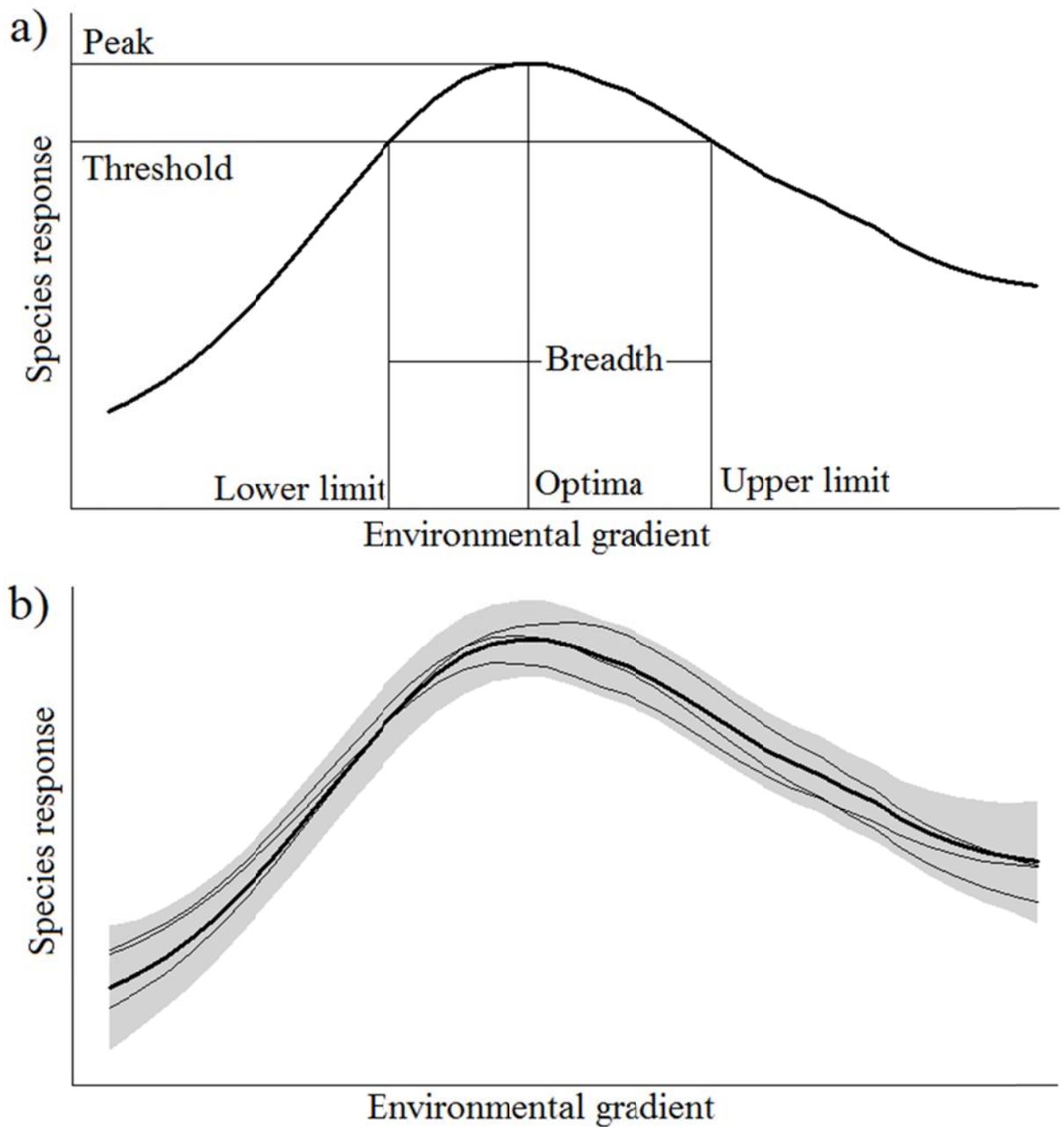


Fig. 1 A species response to an environmental gradient (bold line) has attributes such as its peak (maximum Y value), optima (corresponding X value), and for a given Y threshold, a niche breadth and upper and lower limit (a). Traditional statistical methods can capture the uncertainty in response for given environmental conditions (95% confidence interval – grey shading in (b)) but there are limitless alternative curves (e.g. thin lines in (b)) that fit within this zone of uncertainty and there is no estimate of uncertainty for the optima, peak or other attributes of the curves unless you make assumptions about the shape of the response.



Fig. 2 The open turf structure of *Polytrichastrum alpinum* (top), with gametophytes more widely spaced from one another, contrasts with the densely packed *Chorisodontium aciphyllum* (bottom) illustrated in cross section as well as from above. We hypothesised that the different turf structures could influence the photosynthetic temperature response curves of the two species (Fig. 3). Photographs taken on King George Island by SR and MW.

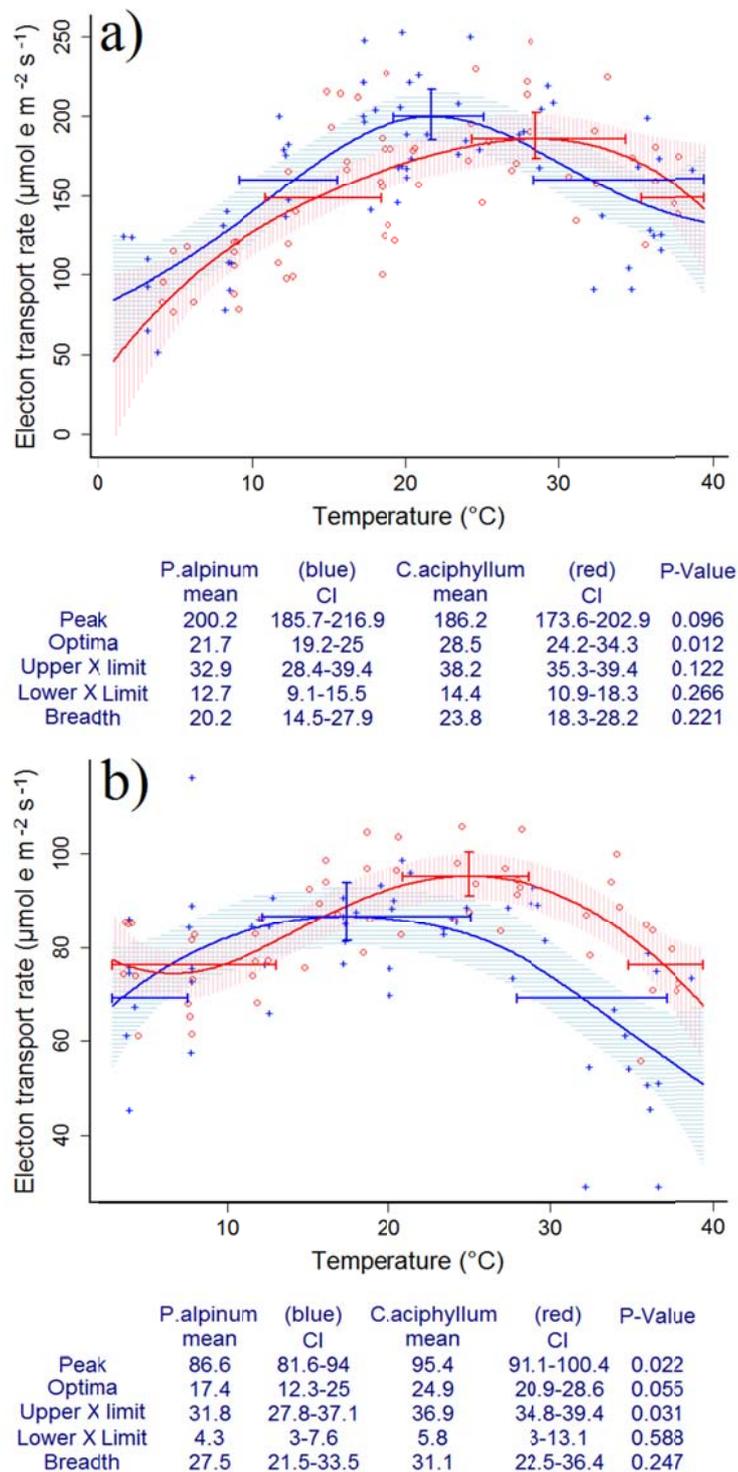
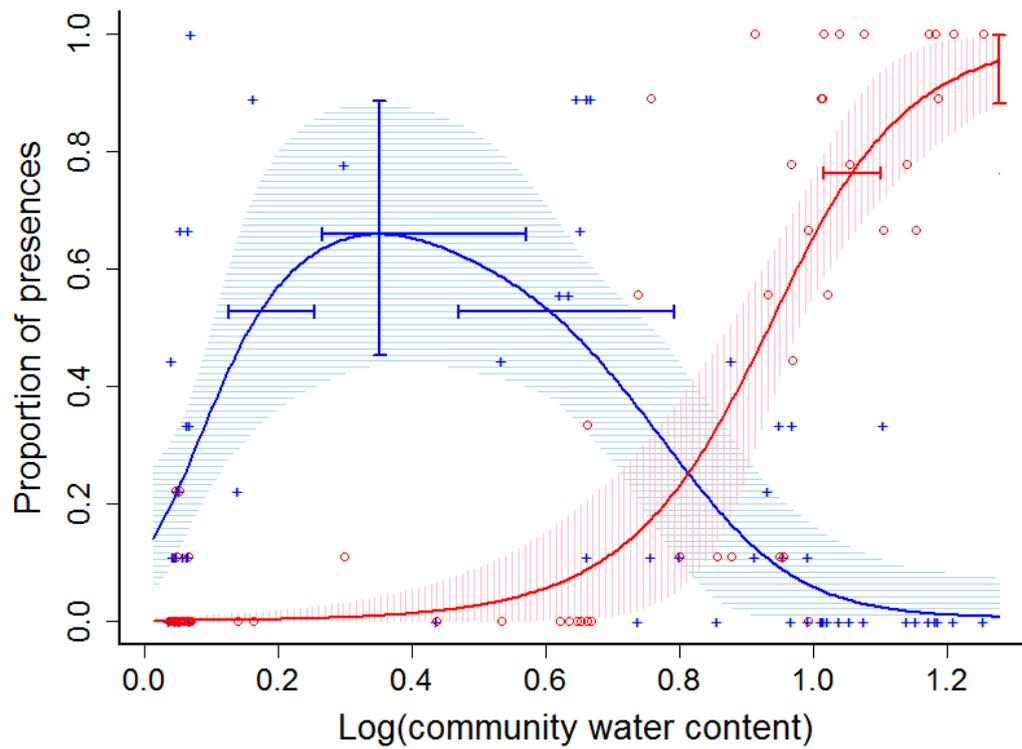


Fig. 3 Photosynthetic temperature response curves under high (a) and low (b) light. Data fitted with Bayesian models using splines with four knots. The responses are shown for *Polytrichastrum alpinum* (blue crosses, lines and horizontal shading) and *Chorisodontium aciphyllum* (red circles, lines and vertical shading). Shaded bands highlight the 95% credible intervals for the responses. Error bars highlight the 95% credible intervals for the peak and optima as well as the upper and lower limits based on a threshold of 80% of the peak. The posterior means and 95% credible intervals are provided in the tables underneath, with a P-value given for the probability that the two curves are different with respect to each attribute.



	C.purpureus mean	(blue) CI	S.antarctici mean	(red) CI	P-Value
Peak	0.66	0.45-0.89	0.95	0.88-1	0.002
Optima	0.35	0.27-0.57	1.28	1.28-1.28	<0.001
Upper X limit	0.61	0.47-0.79	1.28	1.28-1.28	<0.001
Lower X Limit	0.18	0.13-0.25	1.06	1.01-1.1	<0.001
Breadth	0.43	0.32-0.61	0.22	0.18-0.26	<0.001

Fig. 4 The proportion of presences in nine pinch samples in a quadrat that contained *Ceratodon purpureus* (blue crosses, lines and horizontal shading) or *Schistidium antarctici* (red circles, lines and vertical shading) from samples collected over a moisture gradient in 2003. Species response curves have been fitted with Bayesian models using splines with four knots. See Fig. 3 for full explanation of error bars.