Global Change Biology

Global Change Biology (2013) **19**, 473–483, doi: 10.1111/gcb.12051

Modeling plant species distributions under future climates: how fine scale do climate projections need to be?

JANET FRANKLIN*, FRANK W. DAVIS†, MAKIHIKO IKEGAMI†, ALEXANDRA D. SYPHARD‡, LORRAINE E. FLINT§, ALAN L. FLINT§ and LEE HANNAH†¶

**School of Geographical Sciences and Urban Planning, Arizona State University, PO Box 875302, Tempe, AZ 85287-5302, USA, †Bren School of Environmental Science and Management, University of California – Santa Barbara, Santa Barbara, CA 93106, USA, ‡Conservation Biology Institute, 10423 Sierra Vista Ave., La Mesa, CA 91941, USA, §USGS California Water Science Center, 6000 J Street, Sacramento, CA 95819, USA, ¶Science & Knowledge Division, Conservation International, 2011 Crystal Drive, Arlington, VA 22202, USA*

Abstract

Recent studies suggest that species distribution models (SDMs) based on fine-scale climate data may provide markedly different estimates of climate-change impacts than coarse-scale models. However, these studies disagree in their conclusions of how scale influences projected species distributions. In rugged terrain, coarse-scale climate grids may not capture topographically controlled climate variation at the scale that constitutes microhabitat or refugia for some species. Although finer scale data are therefore considered to better reflect climatic conditions experienced by species, there have been few formal analyses of how modeled distributions differ with scale. We modeled distributions for 52 plant species endemic to the California Floristic Province of different life forms and range sizes under recent and future climate across a 2000-fold range of spatial scales (0.008–16 km^2). We produced unique current and future climate datasets by separately downscaling 4 km climate models to three finer resolutions based on 800, 270, and 90 m digital elevation models and deriving bioclimatic predictors from them. As climate-data resolution became coarser, SDMs predicted larger habitat area with diminishing spatial congruence between fine- and coarse-scale predictions. These trends were most pronounced at the coarsest resolutions and depended on climate scenario and species' range size. On average, SDMs projected onto 4 km climate data predicted 42% more stable habitat (the amount of spatial overlap between predicted current and future climatically suitable habitat) compared with 800 m data. We found only modest agreement between areas predicted to be stable by 90 m models generalized to 4 km grids compared with areas classified as stable based on 4 km models, suggesting that some climate refugia captured at finer scales may be missed using coarser scale data. These differences in projected locations of habitat change may have more serious implications than net habitat area when predictive maps form the basis of conservation decision making.

Keywords: biodiversity, California, climate change, downscaling, habitat, impacts, spatial resolution, terrain, topography

Received 19 April 2012 and accepted 18 September 2012

Introduction

Species distribution models (SDMs) are currently the most widely used of scientific approaches to examine potential climate-change impacts to biodiversity. These models, which are based on the statistical association of species distributions with mapped climatic and other environmental factors, have many well-known limitations (Wiens *et al.*, 2009; Dawson, 2011). One major concern is that the spatial scale of climate data used in most SDMs may be inadequate due to the coarse resolution of global climate models (GCM) or downscaled derivatives of those models (Seo *et al.*, 2009). In rugged terrain, climate grids produced at 1–100 km scale may be too coarse to capture topographically controlled climate variation at 10–100 m scale that can exert a strong influence on species distributions (Lookingbill & Urban, 2003; Ashcroft *et al.*, 2009; Dobrowski *et al.*, 2009). These local topoclimatic habitats may need to be accounted for when evaluating climate-change impacts on biodiversity as they may constitute refugia for local populations and increase habitat connectivity for dispersal and migration (Jackson & Overpeck, 2000; Ackerly *et al.*, 2010; Austin & Van Niel, 2011).

Two recent studies have shown that SDMs based on fine-scale climate projections provide markedly different estimates of climate-change impacts than coarse-scale projections; however, these studies only examined alpine plants and reached different conclusions. One study predicted greater persistence of suitable habitat

Correspondence: Janet Franklin, tel. + 480 965 9884, fax + 480 965 8313, e-mail: Janet.franklin@asu.edu

© 2012 Blackwell Publishing Ltd

under climate-change scenarios using finer resolution data (Randin *et al.*, 2009), whereas the other predicted greater persistence of habitat under warming scenarios using coarser data (Trivedi *et al.*, 2008). To help clarify the effect of model resolution on climate-change impact analysis, we modeled current and projected future ranges of 52 endemic plant species over the California Floristic Province, a large, topographically heterogeneous study area. Species were selected to represent a variety of life forms and range sizes. We analyzed the predicted distributions across a 2000-fold range of spatial resolutions (0.008–16 km^2) by independently downscaling 4 km climate models to 800, 270, and 90 m. We compared modeled distributions to ask "how fine-scale do climate projections need to be to capture the strong topoclimatic control on plant species distributions?"

Our specific questions include:

1 How does the extent and location of modeled suitable habitat under current climate depend on the spatial resolution of the climate data?
2 How does the predicted distribution of climate-change refugia (areas of future climatic suitability within the current species range) vary as a function of the scale (spatial resolution or grain) of climate grids?

We hypothesized that:

(a) On average, the predicted extent of modeled suitable habitat would increase with coarser climate data in a given time period owing to spatial generalization, and its location would have diminishing spatial congruence (overlap) with the area of suitable habitat predicted at the finest scale;
(b) climate-change refugia are more likely to be detected at finer spatial resolutions that better represent topographic effects on climate.

Materials and methods

Future climate scenarios

Global climate models were selected that realistically represent the spatial structure of precipitation and important orographic features, accurately simulate California's recent historical climate, and differ in their sensitivity to greenhouse gas forcing (Cayan *et al.*, 2008). These were the Parallel Climate Model (PCM) and the NOAA Geophysical Fluid Dynamics Laboratory (GFDL) CM2.1 model from the IPCC (2007) Fourth Assessment series. Greenhouse gas emissions scenarios examined were A2 (medium–high) and B1 (low) emissions (Nakićenović *et al.*, 2000). The B1 scenario assumes that global CO_2 emissions peak at approximately 10 Gt yr^{-1} in the mid-21st century followed by a leveling of the concentrations before dropping below current levels by 2100, resulting in a doubling of CO_2 relative to its preindustrial level by the end of the century (approximately 550 ppm). Under the A2 scenario, CO_2 emissions continue to climb throughout the 21st century, reaching almost 30 Gt yr^{-1}.

Climate-data downscaling and bias correction

Climate model outputs included daily and monthly temperature and precipitation maps through the end of the 21st century. These coarse-scale (275 km) grids of projected climate were first statistically downscaled to 12 km resolution using the method of constructed analogs (Hidalgo *et al.*, 2008). They were then spatially downscaled from 12 km with digital elevation grids to four resolutions: 4218 m (hereafter '4 km'), 800, 270, and 90 m. The data were first downscaled to 4 km and bias corrected using historical climate data for 1950–2000 available for the USA from PRISM (Daly *et al.*, 1994) as monthly maps (http://www.prism.oregonstate.edu/). The downscaling approach applied a spatial Gradient and Inverse Distance Squared weighting (GIDS) to monthly point data by developing multiple regressions for every fine-resolution cell for every month. Parameter weighting is based on the location and elevation of the coarse-resolution cells surrounding each fine-resolution cell to predict the climate variable of the fine-resolution cell (Flint & Flint, 2012; modified from Nalder & Wein, 1998). As discussed in Flint & Flint (2012), this procedure improves the spatial representation of air temperature over the landscape because at the finer scale it does not average the air temperature over the large area of the coarser scale and allows for the calculation of local adiabatic lapse rates to extrapolate to higher and lower elevations than the coarser averaged dataset. This is essentially a 'draping' of the climate variable over the landscape, which allows for comparisons of vegetation with air temperature estimates that are potentially more accurate at the finer scale than the coarser scale. The modified GIDS technique does not introduce additional uncertainty in the downscaling process, and may indeed improve the estimate of the climate variable by incorporating the deterministic influence (such as lapse rates or rain shadows) of location and elevation on climate. The details of the methodology and evaluation of uncertainty are discussed in Flint & Flint (2012). We derived mean values from downscaled grids of historical climate for the period 1971–2000 and for the future projections for the period 2071–2100 at each resolution.

Instead of including a large number of strongly correlated bioclimatic variables, we analyzed a small set of temperature and moisture variables (Table 1) expected to be physiologically relevant to (Austin & Van Niel, 2011) and found to be associated with western plant distributions (Stephenson, 1998; Rehfeldt *et al.*, 2006). These included minimum temperature of the coldest month (T_min), maximum temperature of the warmest month (T_max), mean temperatures of the wettest (T_wet) and driest (T_dry) quarters, growing degree days for days >5 °C, mean precipitation of the wettest (Ppt_wet) and warmest (Ppt_warm) quarters, and an aridity index [the ratio of annual precipitation to potential evapotranspiration (PET)]. Growing Degree Days (GDD5) is a cumulative annual sum of average temperature per day above a base temperature (5 °C). We calculated GDD5 using minimum and maximum monthly temperatures and the number of days per month (see Sork *et al.*, 2010).

Table 1 Bioclimatic variables used as predictors

Bioclimatic variable description	Abbreviation	Units
Max temperature of warmest period	T_max	°C × 10
Min temperature of coldest period	T_min	°C × 10
Mean temperature of wettest quarter	T_wet	°C × 10
Mean temperature of driest quarter	T_dry	°C × 10
Growing degree days above 5 °C*	GDD5	°C × 10 (cumulative)
Precipitation of wettest quarter	Ppt_wet	mm
Precipitation of warmest quarter	Ppt_warm	mm
Aridity index (annual ppt/PET)*	Aridity	×100

Bioclim = conventional names from WORLDCLIM (Hijmans *et al.*, 2005) (www.worlclim.org) based originally on ANUCLIM and BIOCLIM (Busby, 1986). *Not standard Bioclim variables. Units multiplied by 10 for faster processing using integer (not floating point) data while retaining precision.

Potential evapotranspiration was calculated from modeled insolation using topographic shading (Tian *et al.*, 2001; McCune & Dylan, 2002) and temperature using the modified Penman–Monteith methodology presented in Kay & Davies (2008).

Note that because local slope and aspect are used to calculate insolation the aridity index is especially sensitive to local terrain variation. Aridity is also the only bioclimatic variable analyzed that depends on modeled light regime, a factor that is important for modeling plant distributions (Davis & Goetz, 1990; Austin & Van Niel, 2011).

At each scale, temperature variables were moderately to highly correlated at the locations represented by our species locality data. For example, at 90 m for the period 1950–2000, squared Pearson correlations (r^2) among temperature variables ranged from 0.35 (T_max vs. T_min) to 0.96 (T_min vs. T_wet) (n = 70 989). Squared correlations between temperature and precipitation variables were much lower, ranging from 0.03 (T_max vs. Ppt_wet) to 0.15 (GDD5 vs. Ppt_wet). Aridity was most strongly correlated with Ppt_wet (r^2 = 0.64).

Species data

Fifty-two species (Table 2) were selected because their current ranges, or most of their ranges, are confined to the California Floristic Province. They were also chosen to represent a variety of range sizes (narrowly vs. widely distributed), life forms (trees, shrubs, herbs), and habitats representing different climate subregions (coastal, foothill, lower and upper montane). Range size was approximated by summing the area of ecological subregions within which species occurrences fell (Fig. 1). Species presence data came from a database of more than 82 000 locations that was compiled for ongoing studies of climate-change impacts on California plant species (Viers *et al.*, 2006; Hannah *et al.*, 2008). The number of presences ranged from 9 to 6359 at 90 m resolution for these 52 species, and one observation per grid cell was used for modeling at each resolution (Table 2). The majority of these occurrence records came from vegetation plot surveys conducted since the mid-1990s and their locational precision is on the order of tens of meters, well within the finest resolution considered in this study. Some of the records came from older natural history collections (≪1%) or surveys (18%) and therefore their locational uncertainty is on the order of hundreds of meters, contributing some noise to the analyses.

Species distribution modeling

We used the MaxEnt software (Phillips *et al.*, 2006; Elith *et al.*, 2011) for species distribution modeling because this method is particularly effective when species occurrence data comprise presence-only records, small samples, and are not a probability-based sample (Elith *et al.*, 2006; Phillips & Dudík, 2008). Seventy percent of occurrence data were used for model training and 30% for testing, and this process was repeated five times using bootstrap sampling from all occurrence data. We used the average of the five models for our final analyses. We conducted jackknife tests of predictor variable importance. To determine variable importance, in each iteration of the training algorithm the change in regularized gain is added to the contribution of the corresponding variable. That means the higher contributing parameters are used more frequently than others, but does not necessary mean they are more important biologically.

Ten thousand random background points were used for each model and were drawn from a 486 230 km^2 model domain based on river basin boundaries that extended up to 130 km outside of California (Fig. 1). At the 4 km scale the random sample covers 33% of the sample domain vs. 0.01% of the domain at the 90 m scale. The large difference in background sampling rate at different scales should not affect model parameterization. This is because the background samples in MaxEnt are considered random samples where the species could be present or absent, as opposed to 'pseudo-absences' (Phillips & Dudík, 2008). However, variation in background sampling will cause differences in model goodness-of-fit measures such as the area under the receiver-operating curve (AUC), which treats background samples as absences. The treatment of background samples as absences is certainly less appropriate for wide-ranging species at the 4 km scale. We report AUC values as an indication of model fits (Fielding & Bell, 1997), but do not compare AUC values across scales. Although there have been criticisms of the use of AUC to evaluate SDMs (Lobo *et al.*, 2008), and AUC based on a random background sample many be inflated (Hijmans, 2012), it is a useful metric for comparison with other studies (Phillips *et al.*, 2006; Elith & Graham, 2009).

For each species, one model was developed at each scale using climate data from the period 1971–2000. This model was then projected to four sets of future climate maps (two GCMs,

Table 2 Study species, grouped according to range size class and life form. Numbers of occurrences at 90 m resolution are shown in parentheses. Although number of observations declined as resolution coarsened, it only fell below 9 (to 5 and 7) for two species at 4 km resolution

	Range		
Life Form	Narrow <10 000 km^2	Intermediate 10 000–50 000 km^2	Broad >50 000 km^2
Herb	*Acanthomintha ilicifolia* (71) *Chorizanthe orcuttiana* (9) *Deinandra conjugans* (40) *Delphinium hesperium* (24) *Eryngium aristulatum* (66)	*Galium angustifolium* (89) *Ptilagrostis kingii* (17) *Scutellaria californica* (120)	*Erigeron petrophilus* (115) *Poa stebbinsii* (17)
Shrub	*Arctostaphylos rainbowensis* (51) *Arctostaphylos rudis* (11) *Ceanothus megacarpus* (471) *Ceanothus verrucosus* (87) *Quercus dumosa* (173) *Viguiera laciniata* (25)	*Adenostoma sparsifolium* (243) *Arctostaphylos glandulosa* (267) *Arctostaphylos mewukka* (192) *Artemisia cana bolanderi* (19) *Artemisia rothrockii* (18) *Ceanothus greggii perlexans* (205) *Ceanothus tomentosus* (117) *Ericameria ericoides* (49) *Keckiella antirrhinoides* (56) *Quercus sadleriana* (703) *Ribes lasianthum* (22) *Salvia leucophylla* (25) *Xylococcus bicolor* (133)	*Ceanothus oliganthus* (76) *Chamaebatia foliolosa* (859) *Corylus cornuta californica* (2161) *Hazardia squarrosa* (420) *Heteromeles arbutifolia* (1870) *Lepechinia calycina* (72) *Pickeringia montana* (35) *Rhamnus rubra* (157) *Ribes malvaceum* (87) *Trichostema lanatum* (83)
Tree	*Cupressus forbesii* (14) *Juglans californica* (86)	*Picea breweriana* (218) *Pseudotsuga macrocarpa* (41) *Quercus engelmannii* (112)	*Abies magnifica* (3355) *Aesculus californica* (441) *Pinus attenuata* (427) *Pinus lambertiana* (6359) *Pinus sabiniana* (1875) *Quercus lobata* (527) *Torreya californica* (88) *Umbellularia californica* (1293)

Herbs include perennials and annuals and some small stature perennial subshrubs.

two emissions scenarios) at each of four scales. The threshold criterion used for analyses where binary maps of suitable vs. unsuitable habitat were required was the threshold where sensitivity equals specificity (Freeman & Moisen, 2008). R Development Core Team, 2012 packages rgdal (Keitt *et al.*, 2010) and raster (Hijmans & van Etten, 2011) were used for thresholding, comparing, and analyzing maps.

The resulting models, based on bioclimatic variables only, predict climatically suitable future habitat conditioned on present distributions and climate. Although terrain effects are taken into account in the aridity index, these models do not include other factors known to limit plant distributions, such as substrate and landform that affect nutrient and water availability (Franklin, 1995). The reason we restricted our analysis to bioclimatic variables was so that the effect of climate-data downscaling on the spatial characteristics of predictions could be examined independently of other factors (e.g., the fixed scale of other potential predictor maps).

Analysis

Generalized linear mixed models (GLMMs) were used to evaluate the effects of scale on attributes of SDM predictions describing habitat area, change in area, and spatial overlap between present and future scenarios (Fig. 2), and to determine if these relationships varied with species characteristics (range size, life form, and habitat); we used species as the random effect (multiple models comprising repeated measures on species), as in similar studies (Elith *et al.*, 2006). Generalized linear mixed models were Gaussian unless otherwise noted.

Predicted habitat area (current and future) and its projected net change, as well as range ratio and spatial congruence (based on current climate) were the response variables used to examine the effects of scale and species characteristics on the extent and location of modeled suitable habitat (based on the thresholded SDMs). Scale and species characteristics were the fixed effects. The range ratio was defined as the ratio of area predicted suitable at the coarser scales to the area predicted suitable at the finest scale for (a) 270 m vs. 90 m, (b) 800 m vs. 90 m, and (c) 4 km vs. 90 m. For this right-skewed response variable, a Poisson GLMM was estimated. For the same three pairs of spatial resolutions, the spatial overlap between suitable habitat at coarser scales vs. the finest was described by the spatial congruence index (based on a Dice Sorensen similarity measure, Legendre & Legendre, 1998),

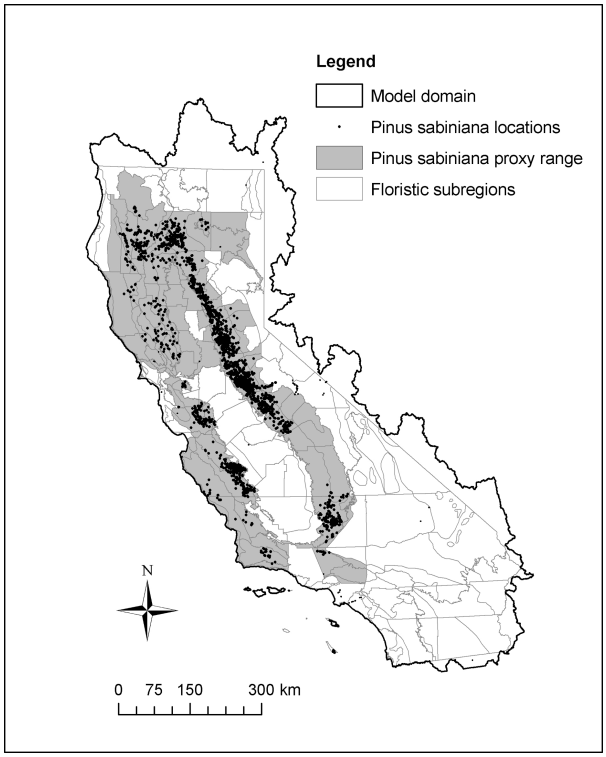

Fig. 1 Study area, showing the model domain (encompassing the state of California), occurrence data for an example species (*Pinus sabiniana*), and geographic units within the state of California (Floristic subregions) used to form proxy ranges (for assigning species to range size classes).

$2a/(2a + b + c)$, where a is the area mapped as suitable at both resolutions (overlap), b is the area mapped suitable at the coarse resolution but not fine, and c is the area mapped suitable at the fine resolution but not coarse. High values of the index indicate higher degrees of spatial congruence.

A potential cause of the larger predicted ranges observed with coarse climate data (Seo *et al.*, 2009) is that response functions of species occurrence to bioclimatic predictors estimated from coarse-scale data may be broader and flatter, or less bounded (wider tolerances, less distinct limits, and optima and/or lower magnitude optima), than those estimated from fine-scale data. This is because coarse-scale data average the climate variables over broader areas, thus potentially missing some of the fine-scale variability that is relevant to the species' physiological limitations. We examined this hypothesis by comparing the 90 and 4 km marginal response curves from MaxEnt for each of the 52 species across the eight bioclimatic variables (examples shown in Figure S1). The marginal response curves show how the logistic prediction of habitat suitability changes as each bioclimatic variable is varied, keeping all other predictor variables at their average value. To provide a summary of response curve shape change with scale, we categorized the 416 comparisons of fine- and coarse-scaled response curves as either: response curves from models using fine-scale data changed to more generalized response curves from models using coarse-scale data (i.e., changing from unimodal to monotonic, lower peak, broader range) or not (no change in shape, changes not as above), and weighted the tally by variable importance.

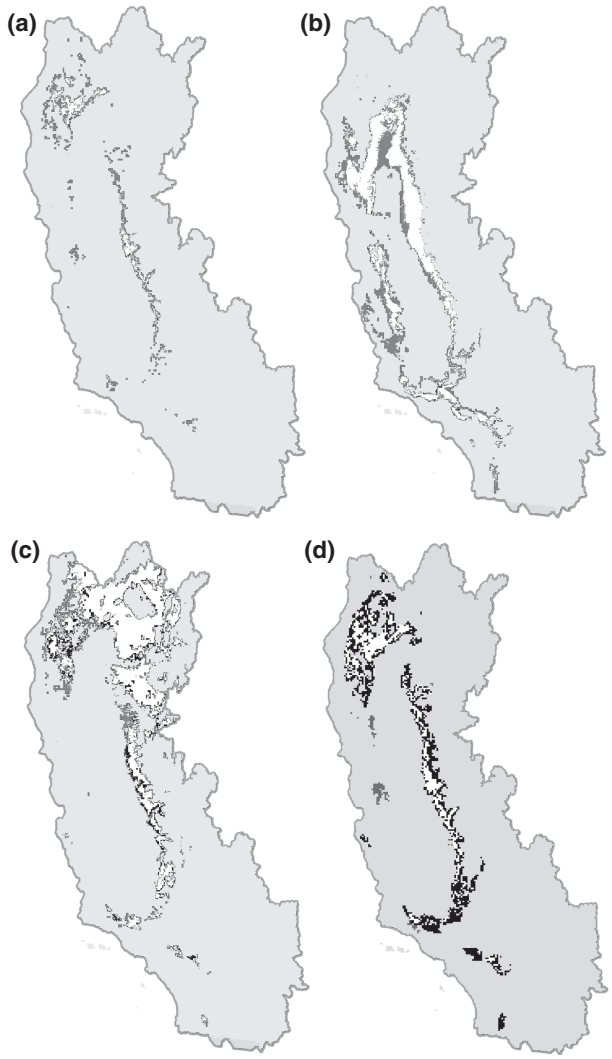

Fig. 2 Predicted habitat distribution for *Pinus sabiniana* in the modeling domain (Fig. 1), current vs. end of century, based on the Geophysical Fluid Dynamics Laboratory climate model and A2 emissions scenario, comparing predictions of (a) stable, (b) lost, and (c) gained (new) range based on 90 m vs. 4 km climate data. White, predicted by both 90 m and 4 km models; black, predicted by 90 m only; dark gray, predicted by 4 km only. Figure (d) compares end-of-century distribution predicted from 4 km climate data to a 4 km model where the cell is considered suitable range if it contains any suitable habitat based on 90 m climate data. Black cells in panel (d) are 4 km cells predicted unsuitable based on 4 km climate data, but containing suitable 90 m habitat. White cells are predicted suitable based on 4 km climate data and also containing suitable habitat at 90 m; dark gray = predicted by 4 km climate data only (do not contain 90 m suitable habitat).

At each scale, the net change in habitat area between present and future climate predictions, a measure of predicted

habitat gain or loss (e.g., Fig. 2), was calculated for each species. Net change is $(b - c)/(a + c)$ where a is the area predicted suitable in both present and future (stable), b is the area predicted suitable in future but not present (gain), and c is the area predicted suitable in present but not future (loss). The stable range, a measure of spatial habitat shift vs. stability, was calculated $a/(a + c)$. For these analyses, the response variables were percent net change and percent stable range, the fixed effects were scale and species characteristics, and species was again the random effect.

We examined the effect of scale on the predicted distribution of climate-change refugia using two different approaches. First, we analyzed the proportion of mapped habitat predicted to be stable at fine scale that was only predicted stable at fine scale but not the next coarsest scale, i.e., was missed at the coarser scale. We also analyzed the proportion of stable habitat predicted to be stable at a coarser scale but not using the finer scale predictors. As a second approach we quantified the frequency with which habitats identified in fine-scale models were nested within areas identified as suitable by coarse-scale models. To do so, we compared species suitable habitat maps that were based on 90 m models and generalized to 4 km resolution (a 4 km cell was classified as suitable if 90 m habitat was predicted to occur somewhere in the cell) vs. 4 km models (Fig. 2d). Here, we report results with 90 and 4 km range forecasts based on the GFDL model and A2 emissions scenario. Using these different approaches allows us to answer two related questions: (1) How do the amount and distribution of potential habitat refugia differ between fine-scale and coarser scale models? (2) How effective are coarse-scale models at predicting the same habitat as predicted by fine-scale models?

Results

SDM accuracy based on AUC ranged from 0.776 to 0.999 (median 0.982; Table S2). Model AUC was negatively correlated with number of species' observations ($r = -0.90$), and with range area ($r = -0.72$). This suggests that, although number of observations, based on a presence-only dataset from a composite of sources, was not perfectly correlated with range size ($r = 0.52$; see Table 2), it was correlated with the actual prevalence of the species on the landscape. There was no relationship between AUC and the proportion of species occurrences from older surveys (greater locational uncertainty) at any resolution.

Area of suitable habitat, calculated from thresholded spatial predictions of probability, tended to be slightly larger for models that used coarser scale climate data under current climate (only the difference between 90 and 4 km was significant; $P = 0.001$) and under future climate for the A2 emissions scenarios (but not statistically significant) in both climate models (Table S1). Predicted net change in habitat area (the ratio of future to current habitat area at each scale) trended toward greater predicted loss at coarser scales under the GFDL scenarios, but while PCM scenarios tended to predict slight gains in habitat area, there was no apparent trend with scale for PCM (Table S1), and no significant effect of species characteristics.

The range ratio (ratio of area predicted suitable at coarser scales to the area predicted suitable at 90 m resolution in a single time period) increased as resolution coarsened for current climate (Fig. 3), especially at 4 km ($P = 0.05$). As resolution coarsened, the average spatial congruence (overlap) with the 90 m predictions also decreased from more than 80% at 270 m to about 65% at 4 km (significantly lower; $P \ll 0.01$; Fig. 3). There was no significant effect of species characteristics on range ratio, but spatial congruence was somewhat lower for narrow-range species than for those in the larger range classes ($P = 0.02$), was higher for the shrubs ($P = 0.0002$) and trees ($P < 0.0001$) than herbs, and did not differ among habitats.

Although the predicted area of suitable habitat under both present and future climates was generally larger at coarser scales, the percent net change was not signifi-

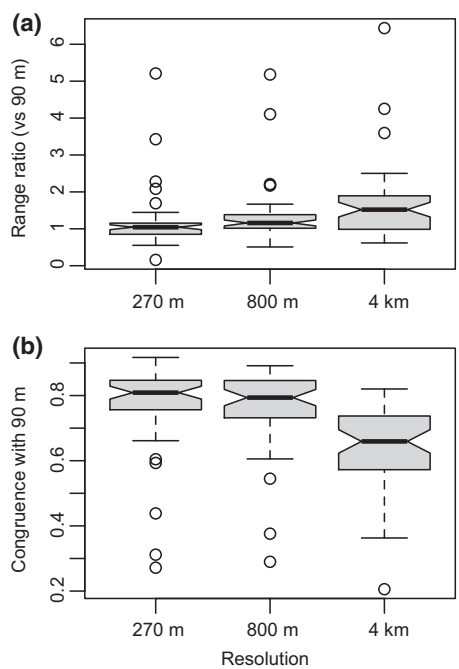

Fig. 3 Boxplots showing (a) the range ratio (ratio of suitable habitat area predicted at coarser resolutions vs. 90 m) and (b) spatial congruence (spatial overlap) of three coarser scales vs. finest 90 m scale predictions for current climate. Heavy line shows median, notches extend to ±1.58 interquartile range/$\sqrt{n}$ to indicate roughly a 95% confidence interval for the difference in two medians, and hinges (edges of box) encompass the quartiles. Whiskers extend to extremes of data (outliers shown as circles).

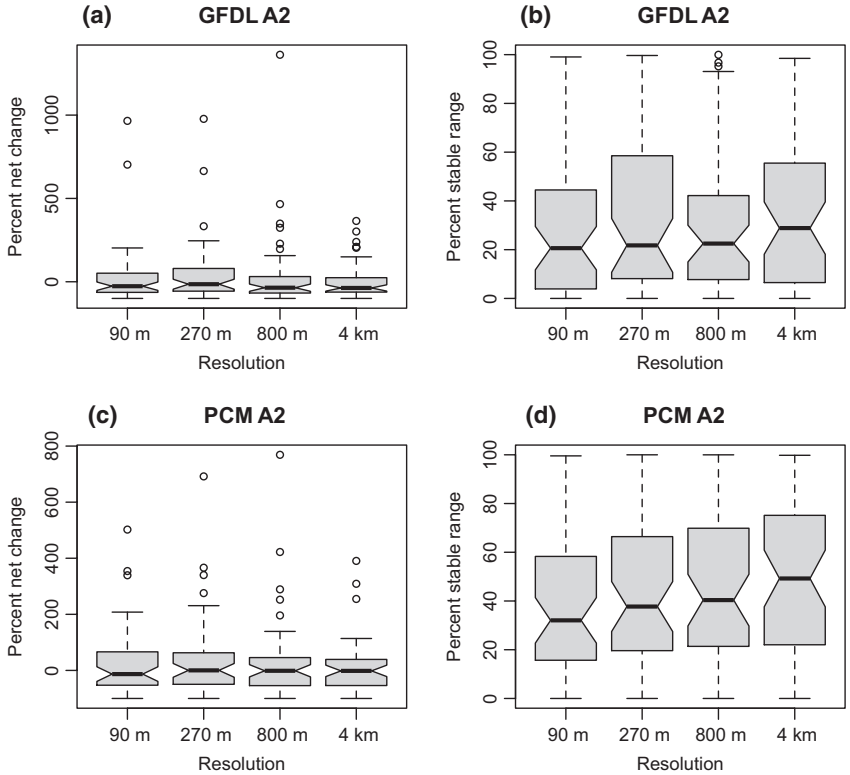

Fig. 4 Boxplots showing for the Geophysical Fluid Dynamics Laboratory (GFDL) climate model the (a) percent net change in predicted habitat area and (b) percent stable range under predicted end-of-century climate change as a function of scale, and the same for the Parallel Climate Model (PCM): (c) percent net change; (d) percent stable range. The A2 emissions scenario was used for both climate models.

cantly different among scales, shown for both GCMs and the A2 emissions scenario (Fig. 4). Under the GFDL A2 scenario, half of the models predicted net loss of suitable habitat of 30% or more, while for a few (6) species, large net gains were predicted at all scales (positive outliers in Fig. 4). The predicted net habitat changes for PCM were smaller (median −13% for 90 m and close to zero at the other resolutions). Predicted net change was greater for narrow-ranging species and lower for trees under both GFDL (P = 0.004 and 0.005, respectively) and PCM (P = 0.006 and 0.0004) A2 scenarios, and did not differ among habitats.

There was a trend for the percent stable range to increase with scale, from a median of 21% at 90 m to 28% at 4 km under the GFDL A2 scenario, and from 32% at 90 m to 49% at 4 km under the PCM A2 scenario; the trend is only significant for the PCM A2 scenario (P = 0.0341), and when scale is treated categorically, only for the 4 km data is percent stable range greater than for other scales (P = 0.009; Fig. 4). Predicted stable range was smaller for trees (P = 0.0009) and larger for foothills species (P = 0.001) under the PCM A2 scenario, but did not differ among range size classes.

On average, about 20% of species' climatically suitable habitat predicted to be stable using finer scale data was missed by coarser scale models (e.g., Fig. 2) using the most extreme GFDL A2 scenario (Fig. 5a). The average percent was slightly higher (24%; P = 0.10), so the loss of information slightly greater, for 270 m vs. 800 m, compared with 90 m vs. 270 m or 800 m vs. 4 km. In addition, an average of 28% of habitat predicted to be stable using coarser scale data was not predicted to be stable using finer scale predictors (Figs 2 and 5b). This larger area of stable habitat predicted by coarser scale data increased systematically with scale and was greatest when going from 800 to 4 km resolution (42%), significantly greater than the other scale comparisons ($P \ll 0.01$).

The predicted distribution of stable range based on 4 km models shows only moderate agreement with the distribution of suitable habitat based on 90 m models that occurred within 4 km cells (GFDL A2, mean congruence = 0.41; Fig. 5c). Results vary widely among species, but congruence tended to be lower for species with small and medium range sizes (although the difference is not significant). Many of the predicted future local (90 m scale) habitat areas fall outside of areas

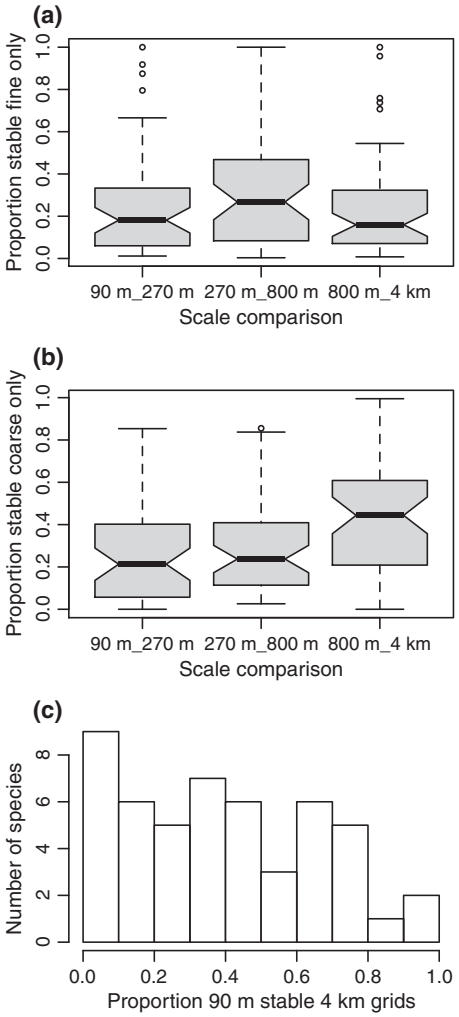

Fig. 5 Boxplots showing (a) the percent of stable habitat under climate change that is only predicted stable with the finer scale climate data and not predicted to be stable with the coarser scale data, (b) percent of stable habitat under climate change that is only predicted stable with the coarser scale climate data and not predicted to be stable with the finer scale data (calculated by spatial overlay of habitat predicted stable at each scale), and (c) frequency distribution for species of the proportion of those 4 km grid cells predicted to be stable by the 4 km model that contain some habitat that is predicted stable by the 90 m model (Geophysical Fluid Dynamics Laboratory A2 scenario).

modeled as stable range with 4 km data. Compared with 4 km models, 90 m models for 38 of 52 species predicted a greater extent of 4 km areas to include locally suitable habitats both now and at end of century (e.g., Fig. 2d).

On average, the most important environmental predictor in SDMs for all species was precipitation of the wettest period, and importance did not change much with scale (Table 3). For the next most important predictors, aridity, and mean temperature of the coldest period, their importance tended to increase at the coarser scales, although the variability in importance among species also increased, and the scale differences were not significant. For growing degree days above 5 °C and precipitation of the warmest period, there was not much of a change in importance with scale. The least important predictors on average were mean temperature of the wettest, warmest, and driest periods, and the importance of these predictors tended to diminish at coarser scales. Response curves were broader and flatter, or less bounded (indicating wider tolerances, less distinct optima and/or lower magnitude optima) when estimated from the 4 km vs. 90 m climate variables about half of the time for the eight predictors when weighted by variable importance (Table 3).

Discussion

There is growing concern over the potential effect of climate-data spatial resolution when using SDMs to predict biodiversity impacts of climate change. This is because climate varies dramatically at local scales, and this variation is likely to be reflected in spatial data. For example, Ackerly *et al.* (2010) reported a difference in range of January minimum temperatures of 3 °C using 800 m PRISM climate grids vs. 8 °C when topoclimatic effects were modeled at a 30 m scale.

We found that there was a trend for SDMs using coarser-resolution data to predict larger habitat area; however, this difference was significant only when finer scales were compared with 4 km. There was a disproportionately greater difference in a number of measures of habitat area and location between 800 and 4 km predictions than between 90 and 800 m predictions. These patterns, observed for a topographically diverse region and a broad range of species types, potentially suggest that there may be a threshold, between 800 and 4 km, beyond which further spatial generalization affects predicted habitat area. It may therefore be advisable for users interested in climate-change impacts on species distributions to use approximately 800 m or higher resolution climate data as it is becoming increasingly available.

Furthermore, we found diminishing spatial congruence (agreement or overlap) between fine- and increasingly coarser scale predictions. The congruence between fine- and coarse-scale predicted habitat was lowest for narrow-range species. Therefore, these rare or range-restricted species would be disproportionately affected by generalization errors introduced by using coarse-scale climate data to project habitat shifts under climate-change scenarios. Seo *et al.* (2009) reported a similar pattern using simple spatial averaging of climate

Table 3 Median variable importance (percent contribution of environmental variable to the MaxEnt model) shown for species distribution models (SDM) estimated at each scale. Range of values is shown in parentheses. Estimate of contribution of the each variable is based on increase in regularized gain in each iteration of the training algorithm. % species broader = the percent of the 52 species whose response curves were broader when estimated from 4000 m data than from 90 m data, and % weighted broader = that percent weighted by the variable importance in each species' 90 m model. Variables are defined in Table 1

Variable	90 m	270 m	800 m	4000 m	% species broader	% weighted broader
Ppt_wet	22 (0.1–66)	23 (0.1–65)	23 (0.1–69)	25 (0–67)	58	51
Aridity	11 (0.1–60)	12 (0.1–61)	14 (1–60)	17 (1–71)	44	44
T_min	8 (0.2–75)	9 (0.2–71)	6 (0.3–73)	12 (0.2–67)	38	57
GDD5	7 (0–69)	7 (0–71)	8 (0–71)	7 (0–63)	56	59
Ppt_warm	7 (0.4–61)	8 (0.6–58)	6 (0.2–59)	7 (0.3–40)	33	37
T_wet	5 (0–48)	4 (0–44)	5 (0–38)	3 (0–31)	37	44
T_max	6 (0–42)	5 (0–39)	4 (0.1–40)	4 (0.2–34)	38	33
T_dry	2 (0–19)	1 (0–17)	2 (0–18)	2 (0–19)	46	61

data. Our study confirms this trend using spatially downscaled very high-resolution climate data and comparing them with moderate-resolution climate data (4 km) that are widely available.

Although coarser scale models generally predicted more habitat area than finer scale models, they also potentially underestimate the extent of localized suitable habitats because many of those habitats were not nested within areas classified as suitable at coarser scale. In other words, it is possible that, especially in mountainous terrain, 90 m data capture habitat variability that represents opportunities for local persistence that coarser scale data cannot resolve.

Our study used climate data that were statistically downscaled to four resolutions spanning three orders of magnitude, all of which would be considered fine-scaled compared with climate data that have previously been used for climate-change impact projections (e.g., 12–50 km). The 4 km scale averages ridge-to-ridge variability in complex terrain, whereas the 800 m scale begins to capture the larger topographic features, highlighting the differences between large valley bottoms and ridgetops. The 270 m scale captures climate at the hillslope scale in California, whereby several grid cells generally describe an individual hillslope (Flint & Flint, 2012). If these spatial scales are accurately reflected by the downscaling procedure, the 90 m scale may have improved the correlation of climate with plant distributions, especially in locations with more topographic complexity. Because the range of temperatures is smallest at fine resolutions, this effect, whereas important at a local scale, may not have been detectable at a regional scale. Also, the ability to increase the resolution of climate models by downscaling is limited in flat locations where adiabatic lapse rates are small, vs. locations with topographic complexity, where rates are large over short distances.

The strongest predictors of current distributions for the study species were wet season precipitation and the aridity index. Both wet season precipitation and aridity are probably serving as surrogates for seasonal or annual climatic water deficits that are of direct physiological importance to plants (see also Stephenson, 1998; Flint & Flint, 2012). Because we did not incorporate soils information, we could not model seasonal soil water balance across the landscape. Precipitation in the study region is mainly frontal cyclonic, and presumably less sensitive to local topography than temperature-related variables (Dobrowski, 2011). Although the aridity index exhibited greater sensitivity to spatial scale than other variables in our study, its influence in the distribution models increased at coarser scales. Given its dependence on local hillslope orientation, this result is contrary to our expectations and therefore difficult to explain. Possibly, the 90 m scale is still too coarse to resolve important local topographic facets that control radiation regimes, whereas at slightly coarser scales, the aridity index more effectively captures combinations of larger scale terrain orientation and precipitation gradients. Moreover, the downscaling methods used here did not account for important drivers of local temperature variation and soil moisture balance such as shading from surrounding hillslopes, fine-scale cold air pooling, and surface water convergence. We think it is likely that congruence between fine and coarse scales will decrease even further as local topoclimates are resolved with finer topographic data and additional refinements to the downscaling procedure.

Other studies have suggested that spatial averaging in coarse-scale climate data may result in an imprecise estimate of species response functions (Trivedi *et al.*, 2008; Randin *et al.*, 2009). In this study, models estimated from coarser climate data had broader, less bounded environmental response curves about half of

the time, so the trend was not consistent across species and predictors. Nevertheless, models using coarser-resolution climate data did yield predictions of larger, more generalized habitat areas both for current climate and future climate scenarios, consistent with expectations. Due to this generalization across scales, it is not surprising that the estimated percent change in habitat area from current to future did not differ significantly as a function of scale. This suggests that summary projections of, for example, the number of species expected to gain or lose climatically suitable habitat under climate-change scenarios might not differ greatly if they were based on 4 km vs. much finer scale climate data. However, these summary projections do not account for scale effects on the predicted location of stable habitat. The proportion of the species range predicted to be stable, i.e., the amount of spatial overlap between predicted current and future climatically suitable habitat, increased with coarser scale data. On average, 4 km data overestimated stable habitat by 42% compared with 800 m data.

The more spatially generalized estimates of stable habitat at coarser scales relative to finer scales, as well as the lack of spatial congruence between habitat forecasted to be stable at coarse vs. fine scale, both have more serious implications than percent change in area when SDMs form the basis of conservation decision making with regard to climate-change adaptation. In this study, SDMs based on coarser scale data would yield more optimistic projections than fine-scale predictions about the total amount of habitat likely to be stable, and therefore likely occupied, especially if these spatial forecasts of suitable habitat were used to support models accounting for dispersal (reviewed in Franklin, 2010), particularly for poorly dispersing species (Keith *et al.*, 2008; Regan *et al.*, 2010). At the same time, coarser scale models underestimate the presence of locally suitable habitats that could allow species persistence, i.e., without requiring dispersal to areas that may become climatically suitable in the future.

Another study examining climate-data scale effects also determined that coarse-scale data predicted persistence of habitat for plant species under warming scenarios where fine-scale models predicted loss (Trivedi *et al.*, 2008). While they painted an optimistic picture of stable habitat with their broad brush, in our study coarser-resolution models failed to capture much of the habitat predicted to be stable under climate-change scenarios at finer resolutions, assuming that the fine-scale predictions are accurate (see also Flint & Flint, 2012). These spatial forecasts of the locations of future suitable habitat are crucial for conservation planning and adaptation policies.

In conclusion, the potential importance of local climate refugia for species persistence and rapid migration makes understanding the effect of climate-data spatial resolution especially critical for predicting the impacts of climate change on biodiversity in complex terrain. Determining how analysis scale and downscaling approach affect projected changes in habitat distributions under climate-change scenarios is necessary to establish risk-based conservation priorities and adaptation strategies. We found only small differences in estimated range extent based on downscaled climate grids at 800, 270, and 90 m resolution. However, the actual location of forecasted suitable or stable habitat varied considerably as a function of model resolution, especially comparing 4 km or 800 m model with 270 or 90 m model. The difference in location of modeled suitable habitat at 90 m vs. 270 m is minor, suggesting that, at least based on environmental data and climate downscaling methods used here, the answer to the question posed in the title of this study is that on average '270 m is fine enough,' but results vary widely among species. In rugged terrain even finer models (e.g., 10–30 m) may be needed to represent ecologically significant microclimates associated with cold air pooling, topographic convergence, and insolation patterns (Dobrowski *et al.*, 2009). These scale dependencies may also vary according to climate-change scenario and species range characteristics.

Acknowledgments

This research was supported by grants from the U.S. National Science Foundation (BCS-0824708 and DEB-0824708 to J. F and EF-1065826 to F. D.), the U.S. Department of Energy (DE-FC02-06ER64159) to J. F., the California Energy Commission (500-08-020) to F. D., M. I., and L. H., and by The Nature Conservancy, for which we are grateful to R. Shaw. We also thank H. M. Regan., D. D. Ackerly, and J. Ripplinger for commenting on an earlier draft of this article, which was also greatly improved by the comments of three anonymous reviewers.

References

Ackerly DD, Loarie SR, Cornwell WK, Weiss SB, Hamilton H, Branciforte R, Kraft NJB (2010) The geography of climate change: implications for conservation biogeography. *Diversity and Distributions*, **16**, 476–487.

Ashcroft MB, Chisholm LA, French KO (2009) Climate change at the landscape scale: predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology*, **15**, 656–667.

Austin MP, van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1–8.

Busby JR (1986) A biogeoclimatic analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. *Australian Journal of Ecology*, **11**, 1–7.

Cayan DR, Maurer EP, Dettinger MD, Tyree M, Hayhoe K (2008) Climate change scenarios for the California region. *Climatic Change*, **87**, S21–S42.

Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Applied Meteorology*, **33**, 140–158.

Davis FW, Goetz S (1990) Modeling vegetation pattern using digital terrain data. *Landscape Ecology*, **4**, 69–80.

Dawson TP (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science*, **332**, 664–664.

Dobrowski SZ (2011) A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology*, **17**, 1022–1035.

Dobrowski SZ, Abatzoglou JT, Greenberg JA, Schladow SG (2009) How much influence does landscape-scale physiography have on air temperature in a mountain environment? *Agricultural and Forest Meteorology*, **149**, 1751–1758.

Elith J, Graham C (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, **32**, 1–12.

Elith J, Graham CH, Anderson RP *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Fielding A, Bell J (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Flint AL, Flint LE (2012) Downscaling future climate scenarios to fine scales for hydrologic and ecologic modeling and analysis. *Ecological Processes*, **1**, 2, doi: 10.1186/2192-1709-1-2.

Franklin J (1995) Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, **19**, 474–499.

Franklin J (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321–330.

Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence. *Ecological Modelling*, **217**, 48–58.

Hannah L, Midgley G, Davies I *et al.* (2008) *BioMove-Improvement and Parameterization of a Hybrid Model for the Assessment of Climate Change Impacts on the Vegetation of California*. California Energy Commission, Public Interest Energy Research Program. CEC-500-02-004.

Hidalgo H, Dettinger M, Cayan D (2008) *Downscaling with Constructed Analogues: Daily Precipitation and Temperature Fields over the United States*. California Energy Commission PIER Energy-Related Environmental Research. CEC-500-2007-123.

Hijmans RJ (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, **93**, 679–688.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans RJ, van Etten J (2011) *Raster: Geographic Analysis and Modeling with Raster Data*. R package version 1.7-46. Available at: http://CRAN.R-project.org/package=raster (accessed 30 June 2011).

IPCC (2007) *IPCC Fourth Assessment Report: Climate Change 2007 (AR4)*. Cambridge University Press, Cambridge, U.K.

Jackson ST, Overpeck JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194–220.

Kay AL, Davies HN (2008) Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts. *Journal of Hydrology*, **358**, 221–239.

Keith DA, Akçakaya HR, Thuiller W *et al.* (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560–563.

Keitt TH, Bivand R, Pebesma E, Rowlingson B (2010) *rgdal: Bindings for the Geospatial Data Abstraction Library*. R package version 0.6-29. Available at: http://CRAN.R-project.org/package=rgdal (accessed 30 June 2011).

Legendre P, Legendre L (1998) *Numerical Ecology*. Elsevier, Amsterdam.

Lobo JM, Jimenez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

Lookingbill TR, Urban DI (2003) Spatial estimation of air temperature differences for landscape-scale studies in montane environments. *Agricultural and Forest Meteorology*, **114**, 141–151.

McCune B, Dylan K (2002) Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science*, **13**, 603–606.

Nakićenović N, Alcamo J, Davis G *et al.* (2000) *Intergovernmental Panel on Climate Change Special Report on Emissions Scenarios*. Cambridge University Press, Cambridge, UK.

Nalder IA, Wein RW (1998) Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agricultural and Forest Meteorology*, **92**, 211–225.

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Randin CF, Engel RE, Normand S *et al.* (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557–1569.

Regan HM, Crookston JB, Swab R, Franklin J, Lawson DM (2010) Habitat fragmentation and altered fire regime create trade-offs for an obligate seeding shrub. *Ecology*, **91**, 1114–1123.

Rehfeldt GE, Crookston NL, Warwell MV, Evans JS (2006) Empirical analyses of plant-climate relationships for the western United States. *International Journal of Plant Sciences*, **167**, 1123–1150.

Seo C, Thorne JH, Hannah L, Thuiller W (2009) Scale effects in species distribution models: implications for conservation planning under climate change. *Biology Letters*, **5**, 39–43.

Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang HF, Grivet D (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Nee) in the face of climate change. *Molecular Ecology*, **19**, 3806–3823.

Stephenson NL (1998) Actual evapotranspiration and deficit: biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography*, **25**, 855–870.

Tian YQ, Davies-Colley RJ, Gonga P, Thorrold BW (2001) Estimating solar radiation on slopes of arbitrary aspect. *Agricultural and Forest Meteorology*, **109**, 67–74.

Trivedi MR, Berry PM, Morecroft MD, Dawson TP (2008) Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Global Change Biology*, **14**, 1089–1103.

Viers JH, Thorne JH, Quinn JF (2006) A spatial distribution database of CalFlora and Jepson plant species. *San Francisco Estuary and Watershed Science*, **4**, art1.

Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA (2009) Niches, models, and climate change: assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, **106**, 19729–19736.

Supporting Information

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

Table S1. Predicted habitat area for 52 species as a function of climate-data resolution for present climate and four future climate scenarios for the period 2071–2100. Median and standard deviation (km^2) for areas (Net change is the ratio of future to current and is unitless). (a) indicates significant differences (within rows) at alpha = 0.01 ($P < 0.01$).

Table S2. AUC [Area under the curve (AUC) of the receiver-operating characteristic (ROC) plot] shown for each Species Distribution Model and each climate date resolution. AUC is a measure of model performance. It is interpreted as the probability that a random positive instance (species occurrence) and a random background observation are correctly discriminated by the model.

Figure S1. Examples of increasing generalization in marginal response curve shape estimated using MaxEnt from 90 m to 4 km scale are shown here. In the first row, the response of *Ceanothus tomentosus* to precipitation in the wettest quarter (Bio_16) is strongly unimodal and skewed at 90 m resolution, with a peak of 0.8 at around 200 mm (A). At 4 km resolution (B) the response could be described as piecewise linear and unbounded, with a positive response up to 200 mm, but then one that remains flat at higher values. In the second example, the marginal response of *Deinandra conjugans* to minimum temperature of the coldest period (Bio_6) shows an optimum of about 0.8 at about 6 °C and is steeply negative above that at 90 m scale (C), but a lower peak (0.7) and is more or less flat (constant high probability) above that at 4 km scale (D). Under a scenario of increasing winter temperatures, habitat suitability could be predicted to increase at 4 km scale but not 90 m scale for this species.