One, two and three-dimensional geometric constraints and climatic correlates of North American tree species richness.

Helen T Murphy¹, Jeremy VanDerWal² and Jon Lovett-Doust³

Author for correspondence:
1. CSIRO Sustainable Ecosystems
PO Box 780
Atherton QLD 4883
Australia
Helen.Murphy@csiro.au
Phone: 61-7-4091 8828, Fax: 61-7-4091 8888

2.  James Cook University
Centre for Tropical Biodiversity and Climate Change
School of Marine and Tropical Biology
James Cook University
Townsville QLD 4811
Australia
jvanderwal@gmail.com

3.  Nipissing University
100 College Dr, Box 5002
North Bay, Ontario P1B8L7
Canada
jld@uwindsor.ca

Keywords: mid-domain effect, species diversity, North America, null model, environmental gradients, latitude, climate, elevation
Abstract

The ‘mid-domain effect’ (MDE) has received much attention recently as a candidate explanation for patterns in species richness over large geographic areas. Mid-domain models generate a central peak in richness when species ranges are randomly placed within a bounded geographic area (i.e., the domain). The most common terrestrial mid-domain models published to date have been 1-D latitude or elevation models and 2-D latitude-longitude models. Here, we test 1-D, 2-D and, for the first time, 3-D mid-domain models incorporating latitude, longitude and elevation, and assess independent and concurrent effects of geometric constraints and climatic variables on species richness of North American trees. We use both the traditional ‘global’ regression models as well as geographically weighted regressions (‘local’ models) to examine local variation in the contribution of MDE and climatic variables to species richness across the domain. Our results show that in all dimensions the contribution of MDE to patterns of species richness can be quite substantial and we show that in most cases a combination of MDE and climate predicted empirical species richness best in both local and global models. For the North American domain, MDE in the elevation dimension is clearly important in describing patterns of empirical species richness. We also show that the assumption of stationarity in global models is not met in the North American domain and results of these models mask complex patterns in both the effect of MDE on richness and the response of species richness to climate. Our results support the hypothesis that geometric constraints contribute to species richness patterns and we suggest the mid-domain effect should be considered alongside more traditional environmental correlates in understanding patterns of species diversity.
Introduction

Biogeographic patterns in plant species richness have been widely documented and are usually explained by underlying geographical gradients in a variety of ecological and evolutionary factors (see e.g., Currie and Paquin 1987, Stevens 1989, Grytnes and Vetaas 2002, Hawkins et al. 2003, Willig et al. 2003, Zapata et al. 2003). Most commonly, a monotonic relationship with climatic and/or energy-related factors is considered the primary mechanism for explaining the decline in plant species richness with increasing latitude or elevation (Currie and Paquin 1987, Leathwick et al. 1998, Allen et al. 2002, Hawkins et al. 2003, Ohlemüller and Wilson 2003). However, Colwell and Hurtt (1994) showed that ‘nonbiological’ gradients in species richness can arise from the assumption of a random one-dimensional association between the size and placement of species’ ranges. A central peak in species richness is generated, in the absence of a direct effect of environmental gradients, when ranges (empirical or theoretical) are randomly placed within a geographic area bounded by features that limit dispersal (i.e., a domain). Colwell and Lees (2000) labeled the geometric theory of species richness gradients the ‘mid-domain effect’ (MDE) and defined it as ‘the increasing overlap of species ranges towards the centre of a shared geographic domain due to geometric boundary constraints in relation to the distribution of species’ range sizes and midpoints.’ In a review of 21 MDE studies, Colwell et al. (2004, 2005) concluded that the MDE may leave a substantial signature on patterns of species richness across a domain (but see Currie and Kerr 2008). The authors called for MDE to be considered on an equal statistical footing with other more traditional explanations for species richness patterns.

Most studies of the MDE have used one-dimensional models to study patterns of terrestrial species richness, in either a latitudinal (e.g., Willig and Lyons 1998, Koleff and Gaston 2001, McCain 2003, Romdal et al. 2005, Rahbek et al. 2007) or altitudinal dimension (e.g., Grytnes
and Vetaas 2002, Sanders 2002, Bachman et al. 2004, Carpenter 2005), or for the aquatic environment, depth (McClain and Etter 2005, Moreno et al. 2008). More recently a few studies have used two-dimensional latitude and longitude models (e.g., Bokma et al. 2001, Jetz and Rahbek 2001, Diniz-Filho et al. 2002, Hawkins and Diniz-Filho 2002, Kerr et al. 2006). One-dimensional latitudinal mid-domain models are criticized as simplistic and unrealistic because natural ranges are always constrained in at least two horizontal axes (Bokma and Mönkkönen 2000, Bokma et al. 2001, Zapata et al. 2003, Hawkins and Diniz-Filho 2004). Moreover, analysis solely of one-dimensional patterns restricts explanations to mechanisms that vary in that dimension and ignores mechanisms having more complex spatial patterns (Hawkins and Diniz-Filho 2004). Colwell et al. (2004) counter these arguments by noting that neither is the world two-dimensional, yet a reduction in dimensionality is a commonly used strategy for studying complex patterns in statistics and science. Colwell et al. (2004) suggest that, where possible, presenting both one- and two-dimensional analyses (in the absence at that time of any methodology for evaluating three-dimensional models) of the same data set seems the strongest approach.

VanDerWal et al. (2008) recently described an MDE model incorporating 2- and 3-D combinations of latitude, longitude and elevation reflecting a more realistic geometry of geographic range shapes. While a latitude-longitude 2-D MDE model produces a single peak in the middle of the domain, introducing elevation into either a 2-D or 3-D model results in multiple peaks in species richness across the domain because of the available combinations of latitude and longitude with elevation. VanDerWal et al. (2008) showed that for North American plant and animal taxa, geometric constraints may play a significant role in describing empirical species richness. However, the relative contribution of 2- and 3-D mid-
domain effects and other traditional predictors such as climate to observed patterns of species richness across the domain remains unexplored.

Almost all tests of the strength of predictors of species richness patterns, including MDE and environmental variables, to date have been via ‘global’ statistical tests such as linear regression, often now taking into account spatial autocorrelation through inclusion of a spatial term. However, spatially structured variation in the relative contribution of these predictors across a geographic domain may be masked in global models. Global models assume that the relationship under study is spatially stationary, and thus that the parameters of the derived regression model can be applied universally across the domain of interest (Fotheringham et al. 2002). Several recent studies point to non-stationarity in species-environment relationships, particularly across large domains. For example, Hawkins et al. (2003) and Whittaker et al. (2007) found that for a range of taxa in North America and Europe, water is the most important predictor of species richness at lower latitudes while energy becomes relatively more important at higher latitudes. Furthermore, spatially structured residuals are almost always evident in correlations between observed and MDE-expected species richness (Currie and Kerr 2008) suggesting MDE or environmental variables affect richness differently in different sub-regions of a domain. Geographically weighted regression (GWR) is a ‘local’ modeling approach designed to account for species-environment relationships that vary spatially (Fotheringham et al. 2002). GWR calculates a regression coefficient for each location in the domain by using surrounding observations within a particular distance (bandwidth) of the location, and allowing the model’s parameters to vary across the domain. As such it is useful for examining non-stationarity in relationships across a large geographic area (Osborne et al. 2007).
In this paper we use global models (ordinary least squares regression, OLS) to determine the relative contribution of MDE (1-D, 2-D and 3-D), and climate to observed patterns of diversity of 547 North American tree species. We also use GWR (‘local’ models) to examine local variation in the contribution of MDE and climatic variables to species richness across the domain.

Methods

Species ranges and the MDE model

Range maps for 547 North American trees were downloaded (Digital representation of "Atlas of United States Trees" by Elbert L. Little, Jr., http://esp.cr.usgs.gov/info/veg-clim/). Between 1971 and 1977, Elbert Little, Chief Dendrologist with the U.S.D.A. Forest Service published a series of maps of tree species ranges based on inventory lists, detailed forest surveys, field notes and herbarium specimens (Prasad and Iverson 2003). These digitized maps have become the standard reference for most U.S. and Canadian tree species ranges. The domain utilized was continental North America, excluding the southern portion of Mexico. The southern portion of Mexico was omitted since this was the approximate limit to the arid / semi-arid zones in Mexico and our confidence that all Mexican endemic species were included decreased. Thus we used a combination of two methods (with continental edges and ecological regions) for defining the domain. Only species with >80% of their range within the defined domain, or of size >= a single domain cell (20 x 20 km), were included in the modeling and analysis (thus eliminating 132 species from the 679 range maps available on the on-line database). We truncated species ranges at the domain edge, using the remnant portion in the domain in the analysis (as did Diniz-Filho et al. 2002, McCain 2003). We have previously shown (in VanDerWal et al. 2008) that inclusion of non-endemics using this method makes little difference in the predicted mid-domain peak location or height.
A digital elevation model (DEM) of the domain was acquired from WorldClim
(http://www.worldclim.org/) at 5 arc-minute (~9 km) gridded resolution. The DEM was re-
projected from a geographic projection to Albers equal area projection to minimize distortions
in area. The total domain area was 16.3 million km$^2$ represented by 188535 20 x 20 km cells.
To ensure the model predictions were not confounded with area, the domain was divided into
equal-area longitude, latitude and elevational bands. Thus, longitude and latitude included 50
equal-area bands with a single band being 326,529 (±40 SE) and 326,531 (±36 SE) km$^2$
respectively, while elevation included 49 bands being 333,232 (±2087 SE) km$^2$. Although the
domain was in an equal-area projection, bands were created using the latitude and longitude
of each cell to produce banding patterns appropriately defined as latitudinal or longitudinal
bands. In order to account for slope, the area of each cell within the domain was estimated
based on the elevation of the cell with respect to its neighboring cells using equations defined
by Jenness (2004).

The longitudinal, latitudinal and elevational extents (in equal-area bands) of the 547 empirical
species ranges were measured based on their published range maps. Thus, the measured range
size (in km$^2$) of an empirical species (defined by its range extents in longitude, latitude and
elevation) was the surface area of the domain where the species range in longitude intersected
its range in latitude and its range in elevation. The empirical richness patterns for the domain
were created by mapping the number of overlapping ranges for each of the 1-, 2- and 3-
dimensional empirical species extents.

Mid-domain model
The mid-domain model is described in detail in VanDerWal et al. (2008) and the associated Python code is provided in the Supplemental Appendix to that manuscript. Briefly, one million theoretical species ranges were created by randomly selecting domain extents in each single dimension. Thus, as for the empirical species ranges, the range size of each theoretical species was the area of the domain where the species range in longitude intersected its range in latitude and its range in elevation. Then, for each empirical species one theoretical species from the pool (of 1 million) whose range size was within ±10% of the empirical species range size was placed randomly in the domain. The ±10% buffer was chosen, rather than a closest match approach, so that each iteration would not be identical to the previous. This approach allows the range-size frequency distribution (RSFD) of the MDE model to reflect the RSFD of the empirical species. This process was repeated 1000 times and the mean predicted richness for each 1-, 2- and 3-dimensional model was used for further analysis.

**Climatic variables**

Climatic data layers included in the analysis were sourced from WorldClim Version 1.3 (Hijmans et al. 2004). WorldClim includes 19 interpolated global climate layers on a square kilometer grid. Many of these layers are highly correlated, therefore we excluded from the analysis any variables correlated at >70%. Four layers remained in the analysis: (1) mean annual temperature, (2) mean temperature of the wettest annual quarter, (3) annual precipitation and (4) precipitation seasonality (coefficient of variation).

**Statistical Analysis**

We examined the separate (i.e. MDE alone and climatic variables alone) and combined (i.e. MDE and climatic variables together; henceforth the ‘combined’ model) contribution of the MDE and climatic variables to empirical species richness for each of the seven geometric dimensions.
Species diversity data usually have strong spatial structure or spatial autocorrelation (Legendre 1993) and this violates assumptions of standard multivariate statistical tests because observations are not independent from one another. Thus, the number of degrees of freedom is overestimated and Type 1 errors are inflated. We first performed standard ordinary least squares regression (OLS) and tested for spatial autocorrelation in the regression residuals; we found significant autocorrelation for all models, particularly at short-distances. We then ran simultaneous autoregressive models (SAR) which explicitly include the spatial autocorrelation structure of the dataset with incorporation of an additional term implemented through a spatial weights matrix. However, we have found that including a spatial term in the regression when both the dependent (i.e. empirical species richness) and independent variables (i.e. mid-domain modeled richness or climate) are strongly spatially structured, tends to depreciate the importance of the independents and account for nearly all the explanatory power of the models (VanDerWal et al. 2008). Furthermore, we also checked the regression residuals from the SAR models; most models still had significant residual spatial autocorrelation, particularly at the closest distance classes, which is not unusual in this type of analysis (see Bini et al. 2009). Hawkins et al. (2007) have shown that for gridded geographical data (as we have used here), coefficients estimated from OLS regression are not seriously affected by the presence of spatial autocorrelation. Therefore, here we sacrifice the assumption of spatial independence and show only results of OLS in order to assess the relative ‘global’ contribution of MDE and climatic variables in predicting empirical species richness. We compare the relative strength of OLS and GWR models via the Akaike’s Information Criterion (AIC) and although we show the R² values for completeness, we do not consider them in terms of assessing the ‘best’ models and they should be viewed with caution. We checked for multi-collinearity in the models (i.e. for the four climatic variables alone and
for climate with MDE) via the condition number (Lazaridis 2007); all models had a condition
number between 2.4 and 2.7 indicating some effect of multi-collinearity but not of a
magnitude likely to strongly impact the regression estimates.

We also use geographically weighted regression models (GWR) to examine local variation in
the contribution of MDE and climate variables to prediction of empirical species richness
across the domain of North America. The geographical weighting is implemented through a
spatial kernel function which defines the rate at which the weighting of every other
observation declines with distance from the focal observation. We used a 400 km
‘bandwidth’ in the spatial kernel function such that all observations are inversely weighted by
their distance to the focal observation point out to a distance of 400 km (see Fotheringham et
al. 2002 for details on GWR and weightings). An $R^2$ value can be calculated for each cell in
the domain and mapped along with the slope of the regression for visual interpretation of
local variation. While GWR is not designed to remove spatial autocorrelation, by allowing
parameter estimates to vary locally, local models do tend to capture much of the spatial
autocorrelation. We have not assessed the extent of remaining spatial autocorrelation in the
residuals for the GWR model.

All data extraction and modeling was done in Python (version 2.5; http://www.python.org/).
All statistics were done in R (version 2.9.2; http://www.r-project.org/), using the ‘Spatial
dependence: weighting schemes, statistics and models’ package (spdep, version 0.4-4) and the
‘Geographically weighted regression’ package (spgwr, version 0.6-2).

Results
Of the three 1-D global models, MDE alone performed best in the elevation dimension (Table 1; Figure 1, top row). In the 1-D latitude global model, MDE performed very poorly alone and when included in an OLS model with climate, failed to improve on a model with climate alone (Table 1; Figure 1, second row). In the 1-D longitude model, MDE alone performed better (for both the OLS global and local models) than climate alone (Table 1; Figure 1, row 3). In all three local (GWR) models, MDE alone performed better than climate alone, while the best model for all dimensions was obtained with the inclusion of both MDE and climate.

In the 2D latitude-longitude model, MDE performed poorly alone (Table 1; Figure 1, row 4) and did not improve the fit of a model that also included climate. However, in the longitude-elevation model, including MDE in an OLS regression with climate provided a significant improvement over climate alone (Table 1; Figure 1, row 5). While the latitude-elevation MDE model performed poorly globally, locally it improved the fit of the model over climate alone and reached higher $R^2$ values over a larger proportion of the domain than any other model (Figure 1, row 6). Local MDE 2-D models performed much better than global models generally and also improved the fit of models that included climate over climate alone (Table 1). The 3-D global model that included MDE alone also performed poorly compared to the model that included climate alone, but again the best fit was obtained by the model that included both MDE and climate (Table 1; Figure 1, row 7). The local 3-D model also performed best with MDE and climate included.

Visual interpretation of the distribution of $R^2$ values for the GWR models indicates where MDE and climatic relationships with species richness are spatially inconsistent (Figure 1). There is considerable overlap in areas of inconsistency within models where climate and MDE were considered alone. In the latitudinal dimension the local MDE model has its
poorest fit in a band between 45-50°N (Figure 1, row 2). The local climate model performs poorly only at the western end of this band and at the very southern limit of the domain. The climate and MDE model combined improve the fit of the local model in these areas.

Longitudinally the region of poor fit for both the local MDE and climate models occurs between 100 and 110°W (Figure 1, row 3). This region remained an area of poor fit in the combined model although it was improved over either model alone. In the elevation model both climate and MDE achieve relatively poor fits in the mid-south (eastern Texas and Louisiana) and in the far north (Alaska) (Figure 1, row 1). However, the combined model virtually eliminates the region of poor fit in the south and reduces the area of poor fit in the far north.

The 2-D latitude-elevation combined local model appears to be the best predictive model with only a small region of relatively poorer fit (minimum $R^2$ 0.59) in the north-west US (Washington) (Figure 1, row 6). This area displayed the lowest $R^2$ values for all local models that contained latitude; the area was minimized in the 2- and 3-D models that also contained elevation.

For all models except 1-D elevation, the slope of the relationship between MDE and empirical species richness ranges from strongly negative to strongly positive across the domain. In the 1-D elevation model the slope ranges from weakly positive to strongly positive (see Appendix A).

**Discussion**

The most common terrestrial mid-domain models published to date have been 1-D latitude or elevation models and 2-D latitude-longitude models (see Table 1 in Currie and Kerr 2008).
Almost all have been assessed against empirical species richness using global statistical tests such as OLS or spatial linear regression. Most often, MDE has been treated as a simple null hypothesis to be accepted or rejected (Currie and Kerr 2008), rather than as a candidate explanatory variable (Colwell et al. 2009). When assessed as a null hypothesis, empirical support for MDE has been described as ‘questionable’ or ‘weak’ (Zapata et al. 2003, Zapata et al. 2005, Currie and Kerr 2008). We also show that for MDE alone (i.e. when treated as a null hypothesis), none of the global models, perhaps with the exception of the 1-D elevation model, predicts a pattern consistent with empirical species richness, and global $R^2$ is generally very low particularly for 2- and 3-D models. However, as Colwell et al. (2009) describe, MDE can be considered a candidate mechanism driving species richness in its own right since the simple stochastic phenomenon that drives the MDE pattern of richness may contribute to empirical species richness patterns. Therefore, the question is not whether geometry affects patterns of species richness, but what the magnitude of the contribution is (Colwell et al. 2004). Our results show that the magnitude of the contribution can be quite substantial and we show that in most cases MDE and climate combined predicted empirical species richness best in both local and global models. Other researchers have also found that when included in a multivariate framework, geometric constraints are an important contributor to observed species richness patterns (Storch et al. 2006, Carranza et al. 2008).

There have been only a few other mid-domain analyses focused on North America; certainly none in 3-D, and none that include elevation in a 2-D modeling approach. Hawkins and Diniz-Filho (2002) reported for Nearctic birds that species richness was only weakly correlated (in a global modeling approach) with a latitude-longitude 2-D MDE model which compared poorly with a climate model (Currie 1991, Currie and Kerr 2008). Our 2-D latitude-longitude global MDE model also does not predict the species richness of North
American trees well compared with climate. However, for the North American domain, MDE in the elevation dimension is clearly important in describing patterns of empirical species richness. For both local and global models, MDE in the 1-D elevation model performed better than climate alone. The complexity in the patterns of species richness predicted with the 3-D model also demonstrates the major effect of including elevation in the model. Rather than the single peak in species richness in the centre of the domain, predicted by the 2-D latitude-longitude model, inclusion of elevation results in multiple richness peaks and non-uniform gradients in species richness. It is not surprising that elevation should have a strong effect on species richness; Rahbek (1995, 2005) reviewed species richness patterns in relation to elevation and found that the most common pattern, representing approximately half of the studies, was a mid-elevation or ‘hump-shaped’ peak in richness. In the elevation dimension, our models actually create a peak in species richness that is shifted towards lower elevations because we use equal area bands, and disproportionately more of the North American continent occurs at low to mid elevations than at high. Thus when the MDE model draws randomly from allowable range endpoints in the elevation dimension, low elevations are much more likely to be drawn. In fact, the hump-shaped elevational peaks Rahbek (1995) noted are also shifted towards lower elevations when the data are standardized for area.

It is clear that the assumption of stationarity in global models describing the relationship between MDE and empirical species richness is not met in the North American domain and results of these models mask complex patterns in both the effect of MDE on richness and the response of species richness to climate. Using GWR we are able to show where, geographically, the relationship between MDE, climate and empirical species richness breaks down. It is interesting that there is considerable overlap in the geographic areas where the relationship between empirical species richness and both MDE and climate is weakest.
(tending towards a negative slope). In the latitude and longitude dimensions, the ‘weak’ areas occur immediately to the north (between 45-50°N in the latitudinal dimension) and to the west (between 100-110°W) of the south-eastern peak in empirical species richness. Empirical tree species richness in both these areas drops sharply over relatively short distances compared with changes in MDE or climatic gradients and richness is over-predicted. In the elevation dimension, none of the models predicts species richness in Alaska well, with the area of poor fit for the climate model extending down through western Canada. This is an area that is relatively heterogeneous in terms of elevation, with MDE and climatic predicted minimum and maximum species richness values varying over relatively short distances. This results in poor fits over the bandwidth (400 km) of the local models. However, this region is generally very well fitted by the latitude and longitude models, and adding either (in 2-D) or both (in 3-D) to a model with elevation greatly improves the fit. Generally it can be concluded that poor fits of the local models are achieved wherever the empirical species richness varies over relatively shorter gradients than the predictor variables, or vice versa. Again, stationarity is clearly not a good assumption for fitting global models to describe species richness over the North American domain.

Currie and Kerr (2008) have argued that collinearity of MDE with climatic variables confounds many tests of the MDE effect. These authors suggest that species richness tracks MDE in the latitudinal direction in North America, but not the longitudinal direction because MDE and climate are collinear in the north-south direction but not in the east-west direction. While we found low levels of multi-collinearity in our regression models which included MDE and climate, we suggest that geographic gradients of climatic variables across the North American domain in latitude and longitude, as well as elevation, are actually far more complex than Currie and Kerr (2008) imply. In fact, collinearity between MDE and climate
is higher in our models in the longitudinal dimension (condition number 2.7) than in the
latitudinal dimension (2.5), and is relatively low (2.4) in the elevation dimension where MDE
alone had the greatest explanatory power. Moreover, MDE alone performed better than
climate alone in the 1-D global longitude model.

Several researchers have reviewed and critiqued the application and relevancy of mid-domain
models in recent years (Hawkins and Diniz-Filho 2002, Zapata et al. 2003, Zapata et al. 2005,
Currie and Kerr 2008). Some of the criticisms are based on methodological issues with
formulation and testing of MDE models; some on grounds that empirical species richness
patterns are simply not consistent with the mid-domain hypothesis. Our MDE models
address most of the methodological criticisms; we incorporate realistic geometry of both
geographic ranges in three dimensions, as well of the domain itself, and we utilize realistic
range size frequency distributions. Our results demonstrate that MDE does leave a signature
on tree species richness in some dimensions of the domain and across a large portion of the
domain, and often improves the predictions of models which also include climatic variables.
We suggest elevation in particular is an important factor in describing species diversity across
the North American continent. Finally, it has not escaped our notice that for a complete
evolutionary perspective on these biogeographic matters there is a need for the temporal
perspective, taking our three-dimensional geospatial analysis into four dimensions.

**Acknowledgements**

JVDW was supported by the Research Advancement Program at James Cook University and
an Australian Research Council Discovery Grant. We thank David Westcott and Cameron
Fletcher for providing constructive reviews of the manuscript.
References


Rahbek, C. 2005. The role of spatial-scale and the perception of large-scale species-richness patterns. – Ecol. Lett. 8:224-239.


Willig, M. R. et al. 2003. Latitudinal gradients of biodiversity: pattern, process, scale and


Table 1 – AIC and $R^2$ results for OLS regression and GWR for the seven models (three times 1-D models, three times 2-D models and one 3-D model) with the dependant being empirical species richness. The lowest AIC value for each of the global and local models is shown in bold. For local models, the minimum and maximum $R^2$ values (i.e., the range of values across the domain) is shown.

<table>
<thead>
<tr>
<th></th>
<th>1-D</th>
<th></th>
<th></th>
<th>2-D</th>
<th></th>
<th></th>
<th>3-D</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIC</td>
<td>$R^2$</td>
<td>AIC</td>
<td>$R^2$</td>
<td>AIC</td>
<td>$R^2$</td>
<td>AIC</td>
<td>$R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE</td>
<td>-575</td>
<td>0</td>
<td>-3284</td>
<td>0.32</td>
<td>-4984</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate</td>
<td>-3161</td>
<td>0.75</td>
<td>-3157</td>
<td>0.28</td>
<td>-4111</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE &amp; Climate</td>
<td>-3160</td>
<td>0.75</td>
<td>-4261</td>
<td>0.6</td>
<td>-5188</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE</td>
<td>-8342</td>
<td>0.15-1</td>
<td>-8796</td>
<td>0-1</td>
<td>-7483</td>
<td>0.28 - 0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate</td>
<td>-7796</td>
<td>0.26-1</td>
<td>-7974</td>
<td>0.11-1</td>
<td>-6858</td>
<td>0.16 - 0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE &amp; Climate</td>
<td>-9074</td>
<td>0.50-1</td>
<td>-9426</td>
<td>0.15-1</td>
<td>-8457</td>
<td>0.52 - 0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat-long</td>
<td>-4010</td>
<td>0.05</td>
<td>-1534</td>
<td>0</td>
<td>-3332</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat-elev</td>
<td>-6473</td>
<td>0.75</td>
<td>-4078</td>
<td>0.74</td>
<td>-3587</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>long-elev</td>
<td>-6548</td>
<td>0.76</td>
<td>-4122</td>
<td>0.75</td>
<td>-4015</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat-long-elev</td>
<td>-10589</td>
<td>0 -1</td>
<td>-8007</td>
<td>0 -1</td>
<td>-8144</td>
<td>0 -1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat-elev</td>
<td>-11328</td>
<td>0.42 -1</td>
<td>-8235</td>
<td>0.31 -1</td>
<td>-8262</td>
<td>0.09 - 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>long-elev</td>
<td>-11894</td>
<td>0.43 -1</td>
<td>-9104</td>
<td>0.59 -1</td>
<td>-9067</td>
<td>0.11 - 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 – Empirical richness (column 1) and predicted MDE richness (column 2) with $R^2$ results of GWR for MDE alone (column 3), climate alone (column 4) and MDE and climate combined (column 5) for each of the seven geometric constraints models (rows). The first two figures in each row share a common colour scale representing the lowest richness value in brown and the highest in green (but note that the actual lowest and highest values vary between rows). Columns 3, 4, and 5 are all scaled from $R^2 = 0$ (brown) to 1 (green). Further figures, including fitted values for global models and GWR slope for MDE alone, are shown in Appendix A.
Appendix A

Results of OLS and GWR for the seven geometric constraints models. Figures show empirical richness and MDE predicted richness, with fitted values across the domain for global and local models as well as the minimum and maximum richness values. $R^2$ values across the domain are shown for the GWR model. The slope of the GWR regression across the domain is also shown for MDE alone.

Figure A.1 – Longitude

Figure A.2 – Latitude

Figure A.3 – Elevation

Figure A.4 – Latitude-longitude

Figure A.5 – Longitude-elevation

Figure A.6 – Latitude-elevation

Figure A.7 – Longitude-latitude-elevation
Figure A.1 - Empirical and predicted 1D longitude only richness patterns for North American Trees. Fitted values, slope (mDE only regression), AIC and R-squares are based on geographically weighted regressions between empirical richness and mDE predictions, climate or mDE & climate. Common color scale (brown to green) represents -0.62 to 0.63 for richness and 0 to 1 for R-sqr.
Figure A.2 - Empirical and predicted 1D latitude-only richness patterns for North American Trees. Fitted values, slope (MDE only regression), AIC and R-squares are based on geographically weighted regressions between empirical richness and MDE predictions, climate or MDE & climate. Common color scale (brown to green) represents -0.15 to 0.65 for richness and 0 to 1 for Freq.
Figure A.3 - Empirical and predicted 1D-elevation only richness patterns for North American Trees. Fitted values, slope (MDE only regression), AIC and R-squares are based on geographically weighted regressions between empirical richness and MDE predictions, climate or MDE & climate. Common color scale (brown to green) represents 0 to 0.87 for richness and 0 to 1 for R-sqr.
Figure A.4 - Empirical and predicted 3D longitude x latitude richness patterns for North American Trees. Fitted values, slope (MDE only regression), AIC and R-squares are based on geographically weighted regressions between empirical richness and MDE predictions, climate or MDE & climate. Common color scale (brown to green) represents -0.05 to 0.45 for richness and 0 to 1 for R-sq.
Figure A.5 - Empirical and predicted 3D longitude x elevation richness patterns for North American Trees. Fitted values, slope (only regression), AIC and R-square are based on geographically weighted regressions between empirical richness and MIDE predictions, climate or MIDE & climate. Common color scale (brown to green) represents 0 to 0.40 for richness and 0 to 1 for R-sqr.
Figure A.9 - Empirical and predicted 3D bivariate elevation richness patterns for North American trees. Fitted values, slope (full regression), AIC and R-squares are based on geographically weighted regressions between empirical richness and MDE predictions, climate or MDE & climate. Common color scale (brown to green) represents -3.13 to 0.95 for richness and 0 to 1 for Rsq.
Figure A.7 - Empirical and predicted 3D longitude x latitude x elevation richness patterns for North American Trees. Fitted values are from only regression. AIC and R-squares are based on geographically weighted regressions between empirical richness and IMDE predictions, climate or IMDE & climate. Common color scale (brown to green) represents -0.08 to 0.36 for richness and 0 to 1 for R-sq.