Landscape Ecol (2007) 22:837–852
DOI 10.1007/s10980-006-9058-2

RESEARCH ARTICLE

Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models

John A. Kupfer · Calvin A. Farris

Received: 26 May 2006 / Accepted: 15 October 2006 / Published online: 12 December 2006

Abstract The results of predictive vegetation models are often presented spatially as GIS-derived surfaces of vegetation attributes across a landscape or region, but spatial information is rarely included in the model itself. Geographically weighted regression (GWR), which extends the traditional regression framework by allowing regression coefficients to vary for individual locations ('spatial non-stationarity'), is one method of utilizing spatial information to improve the predictive power of such models. In this paper, we compare the ability of GWR, a local model, with that of ordinary least-squares (OLS) regression, a global model, to predict patterns of montane ponderosa pine (*Pinus ponderosa*) basal area in Saguaro National Park, AZ, USA on the basis of variables related to topography (elevation, slope steepness, aspect) and fire history (fire frequency, time since fire). The localized regression coefficients exhibited significant non-stationarity for four of the five environmental variables, and the GWR model consequently described the vegetation-environment data significantly better, even after accounting for differences in model complexity. GWR also reduced observed spatial autocorrelation of the model residuals. When applied to independent data locations not used in model development, basal areas predicted by GWR had a closer fit to observed values with lower residuals than those from the optimal OLS regression model. GWR also provided insights into fine-scale controls of ponderosa pine pattern that were missed by the global model. For example, the relationship between ponderosa pine basal area and aspect, which was obscured in the OLS regression model due to non-stationarity, was clearly demonstrated by the GWR model. We thus see GWR as a valuable complement to the many other global methods currently in use for predictive vegetation modeling.

Keywords Geographically weighted regression · Ponderosa pine · Rincon Mountains · Arizona

J. A. Kupfer (✉)
Department of Geography, University of South Carolina, 709 Bull Street, Room 127, Columbia, SC 29208, USA
e-mail: kupfer@sc.edu

C. A. Farris
Laboratory of Tree-Ring Research, Department of Geography and Regional Development, University of Arizona, Tucson, AZ 85721, USA

Introduction

The rise of increasingly powerful statistical techniques and recent advances in geospatial technologies, particularly geographic information systems (GIS), have resulted in not only new approaches for understanding species–environment relationships but also novel methods for predicting and mapping the distribution of vegetation phenomena at broad spatial scales (Guisan and Zimmerman 2000; Miller 2005). Many of these methods fall under the heading of predictive vegetation mapping/modeling (PVM; Franklin 1995), a two step process in which: (1) relationships between vegetation properties and predictor variables are quantified using data collected from a number of sample sites, and (2) the geographic distribution of those properties is predicted and mapped for other locations on the basis of existing digital maps of the predictors (e.g., grids of topographic variables extracted from a digital elevation model (DEM); Davis and Goetz 1990). PVM has been used to predict patterns of species, communities and biodiversity across landscapes and regions (e.g., Bolstad et al. 1998; Wimberly and Spies 2001; Jensen et al. 2001; Meentemeyer et al. 2001; Ohmann and Gregory 2002; Humphries and Bourgeron 2003; Gavin and Hu 2005), project vegetation responses to climate change (Iverson and Prasad 1998), and enhance vegetation classifications (Franklin 2002).

While the results of PVM are often presented spatially (e.g., as a GIS-derived map of the predicted vegetation attribute), spatial information has not often been included in the model itself. Previous research, however, suggests that predictive models may benefit greatly from the inclusion of spatial information and the use of analytical methods that capture spatial attributes of the phenomena under study (e.g., Miller and Franklin 2002). For example, spatial dependence (i.e., autocorrelation) of vegetation characteristics is common and expected for a range of reasons (Koenig 1999), including autocorrelation of the underlying abiotic characteristics (King et al. 2004), spatially contagious processes and positive feedbacks (Malanson 1997; Dirnböck and Dullinger 2004), and the existence of locally contingent factors such as disturbance history (Phillips 2002; Hawkins and Porter 2003). Spatial dependence has a number of implications for traditional statistical techniques (Legendre 1993), which is especially relevant when the goal of a study is to relate species abundance to environmental variables. When the goal is prediction, disregarding autocorrelation also means that critical information which could be used to improve predictions of vegetation phenomena across space is ignored.

Another method for utilizing spatial information to improve the power of predictive models involves the incorporation of spatial non-stationarity, that is, allowing variability of relationships between variables across space. With ordinary least squares (OLS) regression, for example, the regression coefficients are 'global' in that they are assumed to apply equally to the entire study area. However, the coefficients may vary greatly in space for a range of reasons, limiting both the descriptive and predictive utility of global models (Foody 2004). More recently, 'localized' approaches such as geographically weighted regression (GWR; Fotheringham et al. 2002) have been developed that can detect and in some cases account for spatial non-stationarity in variable relationships.

A growing number of ecological studies have used various means for incorporating spatial dependence in their analyses (spatial autoregressive or spatial lag models: Lichstein et al. 2002; Miller 2005; spatial filters: Borcard and Legendre 2002; Diniz-Filho and Bini 2005), but fewer have addressed non-stationarity. In this paper, we compare the abilities of global (OLS) and local (GWR) regression to predict patterns of ponderosa pine (*Pinus ponderosa*) basal area in a montane environment on the basis of topography and fire history. Because GWR has not been used much in ecological studies, we focus especially on data analysis and interpretation. Our expectations were that: (1) GWR would describe vegetation–environment relationships significantly better than OLS regression; (2) the localized regression coefficients developed for the environmental variables would exhibit significant non-stationarity; (3) predictions for independent data locations not used in model development would be better when using GWR, and (4) an examination of spatial patterns of non-stationarity and GWR regression coefficients would help to clarify relationships between ponderosa pine basal area and environmental predictor variables that may not have been evident with OLS regression.

Methods

Study area

The study was conducted in high elevation conifer forests near the summit of Mica Mountain in the 23,443 ha Saguaro Wilderness Area of Saguaro National Park, AZ, USA (32° N, 111° W) (Fig. 1). First protected as a National Monument in 1933 and later designated a National Park in 1994, the park contains most of the Rincon Mountains as well as lower-lying desert areas to the south and west. Mica Mountain is the highest peak in the Rincon Mountains and is part of the Rincon-Catalina-Tortolita metamorphic core complex, a massive granitic dome of altered Precambrian rock with a carapace of metamorphic gneiss (Chronic 1983).

Summer temperatures in the desert valleys regularly exceed 40°C, but mean monthly temperatures at Manning Camp in the Rincons (2512 m; Rincon RAWS data station, 1994–2004) and at the Palisades Ranger Station in the Santa Catalina Mountains (2425 m; full period of record, 1965–1981) range from 2°C in January to 18°C in July (*Western Regional Climate Data Center: http://www.wrcc.dri.edu/index.html*). Mean annual precipitation at the Palisades Ranger Station is 78 cm, most of which occurs as a result of thunderstorms during a distinct summer monsoon season and steadier rains and snows from frontal storms during winter. Sub-freezing temperatures occur as late as June and as early as September.

Vegetation within the park is structured along environmental gradients related to elevation and ranges from desert scrub and grassland/savanna at lower elevations to conifer forests at the highest elevations (Bowers and McLaughlin 1987). At the elevations examined in this study (2036–2606 m), communities include pine-oak-juniper (*Pinus-Quercus-Juniperus)* woodlands, ponderosa pine-oak (*P. ponderosa-Quercus*) woodlands, ponderosa pine forests, and mixed conifer forests with varying amounts of Douglas-fir, white fir and pine

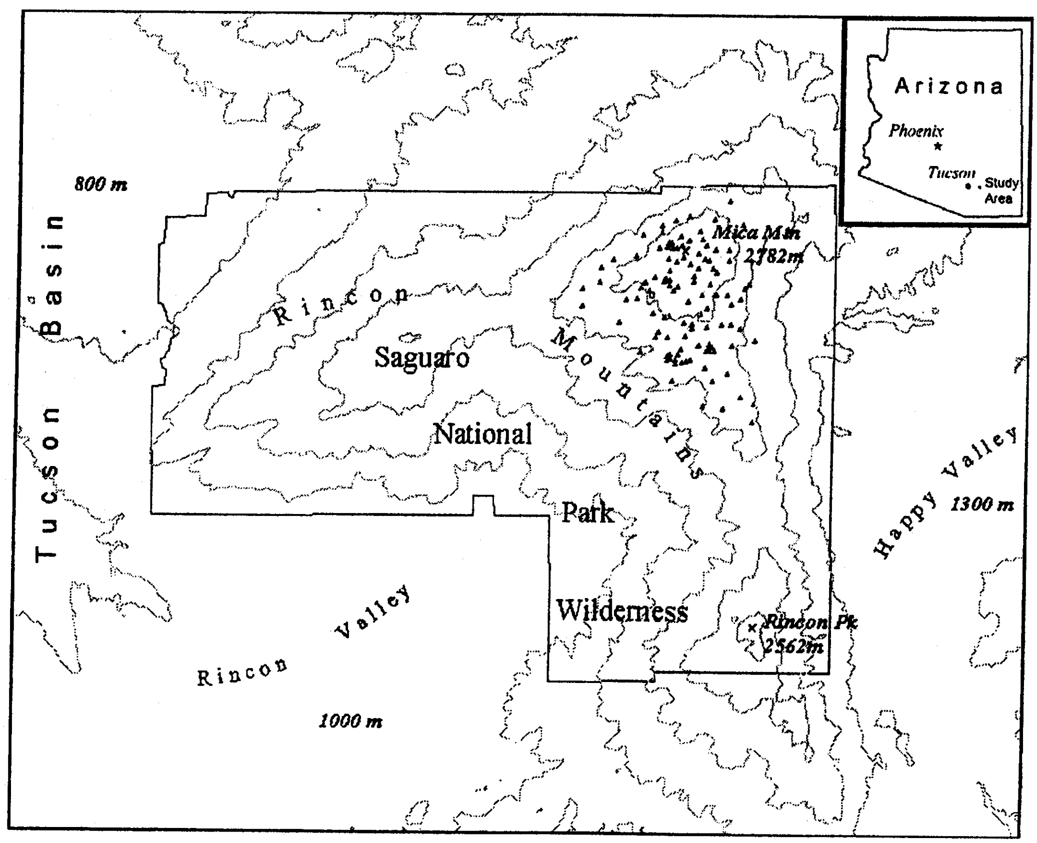

Fig. 1 Map of the Mica Mountain study area, Saguaro National Park, AZ. Locations of field plots are indicated

(*Pseudotsuga menziesii*, *Abies concolor*, *Pinus* species). Vegetation zones are generally higher on the south face of mountains due to differences in energy receipt, temperature and moisture, and differences in environmental conditions and vegetation types lead to (and are in turn influenced by) differences in fire history. Some ponderosa pine forests on Mica Mountain have experienced an unusually high frequency of fires in the 20th century compared to similar forests in the United States, with some stands burning at least nine times in the last 70 years.

Our specific 2,780 ha study area encompassed the relatively small portion of Mica Mountain containing montane conifer forests. Ponderosa pine is the dominant tree species above 2100 m, with southwestern white pine (*P. strobiformis*) as a co-dominant above 2300 m. White fir and Douglas-fir occur in small, isolated mixed stands on sheltered aspects of the north face of Mica Mountain while alligator juniper (*Juniperus deppeana*), border piñon (*P. discolor*) and oaks (e.g., *Q. arizonica*, *Q. hypoleucoides*) are increasingly dominant at the lower forest-woodland ecotone, with ponderosa pine being rare to locally absent.

Data collection

Data on forest overstory-(basal area by species for all individuals >10 cm in diameter at breast height: 1.37 m) and understory (density by species for all individuals 2–10 cm in diameter) composition were collected from 113 0.1 ha sample plots. To ensure complete spatial coverage, the first 23 plot locations were generated in the center of a 1.2 km grid extended across the entire study area. An additional 37 plots were established at randomly chosen locations between these initial grid points to increase the sampling density. Finally, we added compatible plot-level data collected at 53 National Park Service (NPS) fire effects monitoring plots (also 0.1 ha) to arrive at the grid of points shown in Fig. 1.

For each plot, elevation (meters above sea level; ELEV), slope orientation (azimuth degrees; ASPECT) and slope steepness (degrees; SLOPE) were extracted from a USGS 7.5′ DEM with 30 m resolution. Aspect was cosine transformed to scale the values between 0 (south-facing slopes) and 2 (north-facing slopes). Data on fire history were determined using two methods. For the 60 points that we sampled, Fire Frequency Since 1937 (number of fires; FF1937) and Time Since Fire (years; TSF) were determined by reconstructing fire history on the plots using fire scars and other proxy records. Within each plot, we collected from 3 to 14 fire scarred cross sections from living trees and remnant wood. Following sample preparation and cross-dating, all fire scars were assigned a calendar date, allowing us to calculate the number of fires since 1937 and the number of years since the last fire. For the 53 NPS plots, FF1937 and TSF were determined from the Mica Mountain fire atlas (Swantek et al. 1999). The NPS has ground mapped fires since 1937 to their nearest location and recorded information about the fire size, cause, origin date, and control efforts. Fires <10 ha were generally mapped as points at their site of ignition while larger fires were mapped as complete perimeters. FF1937 was determined for each plot based on the number of fire perimeters within which the point was located while TSF was assigned on the basis of the most recent fire perimeter.

Continuous surfaces of the predictor variables were made for each 1 ha pixel in the study area using the USGS DEM (ELEV, SLOPE, ASPECT) and the NPS fire atlas (FF1937, TSF) (Fig. 2). While we recognize the limitations of the fire atlas (including the absence of fire severity measures and uncertainty regarding whether all points inside a fire perimeter actually burned), these are some of the most comprehensive and spatially explicit fire data available in the western United States. Further, additional analyses have shown a strong correlation between the fire scar data and the corresponding fire atlas data for each plot (Farris, unpub. data).

Geographically weighted regression

GWR is an extension of the traditional regression framework in which variations in rates of change are allowed so that regression coefficients are specific to a location rather than being global estimates (Brunsdon et al. 1996). A brief overview of GWR is presented here, but more extensive mathematical treatments can be found

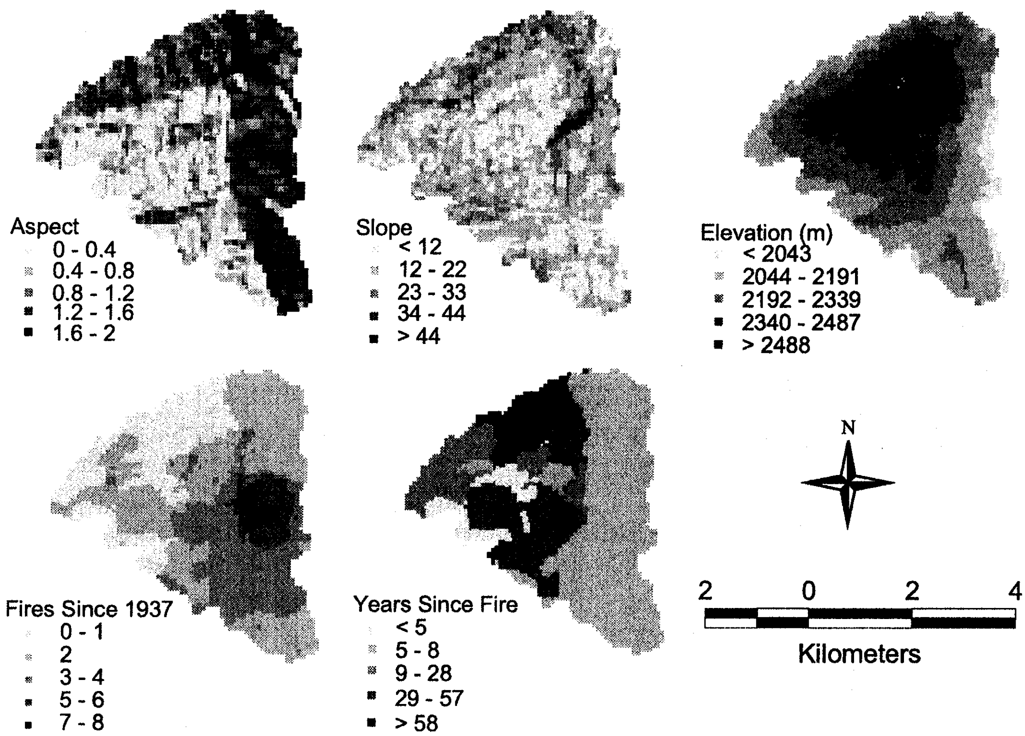

Fig. 2 Raw values of cosine-transformed aspect, slope (°), elevation (m asl), fire frequency (number of fires since 1937) and time since fire (years)

in Fotheringham et al. (2002) and other recent papers (e.g., Zhang and Shi 2004; Wang et al. 2005).

The global regression model can be stated as:

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \quad (1)$$

where y_i is the value of response variable y at location i, β_0 is the intercept, β_j is the slope coefficient for predictor variable j, x_{ij} is the value of predictor variable j at location i, and ε_i is the random error term. Model coefficients are estimated using criteria such as minimizing the squared error for each data point in OLS regression.

Unlike global models, GWR allows local coefficients for each of the predictor variables, meaning that the regression equation can be rewritten as:

$$y_i = \beta_0(u_i, v_i)_+ \sum_{j=1}^{p} \beta_j(u_i, v_i) x_{ij} + \varepsilon_I \quad (2)$$

where (u_i, v_i) are coordinate locations for each location i and $\{\beta_0(u_i, v_i), \beta_1(u_i, v_i), \ldots, \beta_p(u_i, v_i)\}$ are $p + 1$ continuous functions of the location (u_i, v_i) (Zhang et al. 2004). The localized variable coefficients are based on a weighting matrix in which observations around a sample point are weighted using a distance decay function, meaning that more proximal observations have greater influence on the resulting localized regression coefficients.

The estimated coefficients are a function of the bandwidth of the spatial kernel selected for the model, that is, the radius (e.g., in meters) or number of observations around each point included in the weighting matrix. Bandwidth controls the distance-decay in the weighting function and indicates the extent to which the resulting local calibration results are smoothed (Fotheringham et al. 2002). As bandwidth increases and distance decay decreases, the GWR model solution approaches that of the OLS regression solution; with a smaller bandwidth and steeper distance decay, the coefficient estimates increasingly depend on observations in close proximity to i (Brunsdon et al. 1996; Foody 2004).

As an alternative to using a spatial kernel with a fixed radius around each observation or a fixed number of points, the spatial kernels in GWR can be made to adapt themselves in size to variations in density of the data so that the kernels have larger bandwidths where the data are sparse and smaller bandwidths where the data are more dense. In this study, we used an adaptive kernel that utilizes the bi-square function:

$$w_{ij} = [1 - (d_{ij}^2/b^2)]^2 \text{ if } d_{ij} < b \quad w_{ij} = 0 \text{ otherwise} \tag{3}$$

where w_{ij} is the weight of location j in the space at which data are observed for location i, d_{ij} is the distance between locations i and j, and b is the bandwidth. This function has the property of not only decreasing the weight of more distant points from a sample location but also excluding all data points beyond the bandwidth. The selection of the weighting function and optimal bandwidth were accomplished by minimizing the corrected Akaike Information Criterion (AIC_c) as described in Fotheringham et al. (2002), although we also experimented with using generalized cross-validation, which in most cases resulted in comparable model results.

Data analysis

There were three specific questions of interest in this study: (1) Does a GWR model describe the data significantly better than a global model developed using OLS regression; (2) Do the localized regression coefficients exhibit significant non-stationarity and what are the consequences for spatial patterns of model error, and (3) Are predictions for independent data locations not used in model development better when using GWR than OLS regression? Additionally, we wanted to explore patterns of non-stationarity in the study area and assess the implications for predictions of ponderosa pine basal area.

We addressed the first question by calculating regression models of ponderosa pine basal area using all possible combinations of predictor variables, ranging from individual variable models to a model incorporating all five predictors, and testing for improvement in the GWR model over the OLS regression model. The model R^2 is not a meaningful metric for selecting the optimal models or for comparing GWR and OLS regression models because a model with many parameters will have a very good fit to the data but may have few degrees of freedom; the result is overfitting of the data and low predictive ability (Jetz et al. 2005). We therefore used a methodology that minimized the corrected Akaike Information Criterion (AIC_c) to find the smallest model that correctly explained the data. Lower AIC_c values indicate a closer approximation of the model to reality so the best model is the one with the smallest AIC_c. We also used an approximate likelihood ratio test, based on the F-test, to compare the ability of the models to replicate the observed data (Fotheringham et al. 2002). This test is based on a comparison of the residual sum of squares of the two regression models and tests the null hypothesis that the local model represents no improvement over the global model.

The presence and importance of non-stationarity were examined in two ways. First, the significance of the spatial variability in local coefficient estimates was tested formally using a Monte Carlo permutation test (*see* Fotheringham et al. 2002). Second, misspecification of a global model due to non-stationarity may be evident as clumping of the model residuals (e.g., pockets of consistent over- or underprediction) (Legendre 1993; Lennon 2000; Diniz-Filho et al. 2003; Jetz et al. 2005); therefore, another indicator of the improvement of a GWR model (and thus, the value of incorporating non-stationarity into predictive models) is a reduction in spatial autocorrelation of the residuals (Zhang et al. 2005).

To assess the clustering of model errors, we calculated Moran's I spatial autocorrelation coefficient for residual errors of both the OLS regression and GWR models (cf. Legendre and Fortin 1989). Values of Moran's I range from –1 (negative autocorrelation) to +1 (positive autocorrelation), with zero being the expected value for no autocorrelation (Cliff and Ord 1981). Our analyses were conducted using a spatial lag of 150 m, the finest distance that ensured each lag class had at least 20 neighbor pairs; all lag classes except 0–150 m actually had more than 75 neighbor pairs and nearly all had more than 175. The largest lag was 3,000 m, approximately half the maximum dimensions of the study area. The significance of the global correlogram and individual autocorrelation coefficients was tested using a Monte Carlo simulation method with a Bonferroni correction to adjust for repeated testing (Legendre and Fortin 1989).

While the ability of a model to describe relationships within a dataset is important, an additional criterion is whether it can improve prediction of sites not included in the model. One means of assessing this ability is to develop a model using one set of data and validate it using an independent dataset. To get a reasonably large validation dataset, however, required removing a large percentage of our data points, and subsequent GWR models based on the reduced training data displayed a significant increase in AIC_c. Therefore, we used a leave-one-out methodology (i.e., jack-knifing) to compare the predictive abilities of OLS regression and GWR. To do so, we removed one location, calculated the regression coefficients using the remaining 112 locations, predicted basal area for the left out point, and compared the predicted and actual basal areas to get a residual error for that point. We then put the first location back in, removed a second location, recalculated the model, predicted its basal area, and calculated the residual for that location. This process was repeated for all 113 points, and the mean residuals for the GWR and OLS regression models were analyzed.

Finally, the use of global statistics implies that variable relationships in all parts of a study region can be accurately represented by a single coefficient; local statistics stress the local contingencies of data relationships and are useful in searching for exceptions or local 'hot spots' in the data (Fotheringham et al. 2002). We therefore produced and mapped GWR coefficient values to explore the spatial variability of relationships between ponderosa pine basal area and the environmental predictors. Because a standard error can be estimated for each location, it is possible to calculate localized *t*-values that are comparable in interpretation to standard *t*-values. *T*-values for the intercept and variables identified for the optimal GWR model were calculated and mapped for all 113 sample sites. These were then evaluated with respect to continuous surfaces of predicted ponderosa pine basal area by applying the optimal OLS regression and GWR models identified for our sample points to the entire set of grid cells for the study area.

GWR was conducted using GWR v. 3.0 (Fotheringham et al. 2002), OLS regression was conducted using SPSS v. 13.0, and spatial autocorrelation measures were calculated with Rook's Case v 0.95 (Sawada 1999). All maps were produced using ArcView GIS v. 3.3.

Results

Global versus local regression models

Compared to the corresponding OLS regression models, the GWR models exhibited a significant improvement in explained variance in all but two cases (Table 1). The AIC_c values for GWR models decreased by >5–15 in nearly all cases, suggesting that the variables showed a better fit in the GWR model even after accounting for issues of degrees of freedom. It is thus likely that the results of the global regression, with its assumption of spatial stationarity, failed to explain important local variation in ponderosa pine basal area, an interpretation that was borne out by subsequent analyses of coefficient stationarity.

The global linear regression model that included FF1937 and ELEV had the lowest AIC_c value ($AIC_c = 854.3$; Table 1). The coefficients for this model were estimated using OLS regression (Table 2), with the resulting model explaining 17.5% of the variance in ponderosa pine basal area. Both regression coefficients were positive and significant at $\alpha = 0.10$, indicating that ponderosa pine increased with greater fire frequency and at higher elevations.

We also calculated AIC_c values for all possible GWR models and opted to select two models with similar AIC_c values for further analysis, the first with just FF1937 ($AIC_c = 846.5$; GWR

Table 1 Corrected Akaike Information Criterion (AIC_c) and coefficient of determination (R^2) for ordinary least squares regression (OLS) and geographically-weighted regression (GWR) models of ponderosa pine basal area on Mica Mountain, AZ

Model	AIC_c	R^2	ANOVA	
	OLS ∣ GWR[a, b]	OLS ∣ GWR	DF	F-value[c]
x_1	$873.4 \mid 848.5^{41;\ 13.8}$	$0.005 \mid 0.378$	2, 99.2	5.04^{**}
x_2	$869.8 \mid 857.4^{57;\ 10}$	$0.035 \mid 0.266$	2, 103.0	4.04^{*}
x_3	$873.4 \mid 861.4^{22;\ 24.0}$	$0.004 \mid 0.467$	2, 89.0	3.52^{*}
x_4	$855.5 \mid 846.5^{54;\ 9.0}$	$0.150 \mid 0.319$	2, 104.0	5.88^{**}
x_5	$865.0 \mid 853.7^{42;\ 11.3}$	$0.075 \mid 0.309$	2, 101.7	3.71^{*}
x_1x_2	$871.9 \mid 853.6^{57;\ 14.4}$	$0.036 \mid 0.358$	3, 98.6	4.33^{**}
x_1x_3	$875.1 \mid 857.7^{60;\ 12.4}$	$0.008 \mid 0.302$	3, 100.6	4.51^{**}
x_1x_4	$856.5 \mid 847.3^{81;\ 8.6}$	$0.159 \mid 0.307$	3, 104.4	4.01^{**}
x_1x_5	$867.1 \mid 852.2^{42;\ 17.5}$	$0.076 \mid 0.412$	3, 95.5	3.78^{*}
x_2x_3	$871.6 \mid 862.5^{81;\ 9.0}$	$0.038 \mid 0.215$	3, 104.0	3.89^{*}
x_2x_4	$856.3 \mid 850.4^{81;\ 8.8}$	$0.160 \mid 0.291$	3, 104.2	3.33^{*}
x_2x_5	$866.5 \mid 860.0^{64;\ 11.4}$	$0.080 \mid 0.271$	3, 101.6	3.17^{*}
x_3x_4	$854.3 \mid 852.7^{81;\ 8.1}$	$0.175 \mid 0.266$	3, 105.9	$2.56^{n.s.}$
x_3x_5	$866.7 \mid 863.9^{31;\ 22.2}$	$0.080 \mid 0.426$	3, 90.8	2.85^{*}
x_4x_5	$855.8 \mid 849.1^{45;\ 14.3}$	$0.164 \mid 0.381$	3, 98.7	3.08^{*}
$x_1x_2x_3$	$873.7 \mid 858.0^{81;\ 11.8}$	$0.039 \mid 0.291$	4, 101.2	4.60^{**}
$x_1x_2x_4$	$857.8 \mid 850.8^{93;\ 10.0}$	$0.166 \mid 0.307$	4, 103.0	3.50^{**}
$x_1x_2x_5$	$868.7 \mid 855.7^{64;\ 15.2}$	$0.081 \mid 0.359$	4, 97.8	3.78^{**}
$x_1x_3x_4$	$855.3 \mid 851.2^{105;\ 7.6}$	$0.184 \mid 0.268$	4, 105.4	3.36^{*}
$x_1x_3x_5$	$868.7 \mid 858.3^{62;\ 14.7}$	$0.081 \mid 0.336$	4, 98.3	3.52^{**}
$x_1x_4x_5$	$857.3 \mid 848.7^{81;\ 10.8}$	$0.169 \mid 0.332$	4, 102.2	3.68^{**}
$x_2x_3x_4$	$855.4 \mid 853.2^{105;\ 7.9}$	$0.183 \mid 0.259$	4, 105.1	2.80^{*}
$x_2x_3x_5$	$868.1 \mid 864.2^{93;\ 9.9}$	$0.085 \mid 0.218$	4, 103.1	2.98^{*}
$x_2x_4x_5$	$857.5 \mid 852.5^{77;\ 11.6}$	$0.167 \mid 0.321$	4, 101.4	3.03^{*}
$x_3x_4x_5$	$855.0 \mid 856.9^{45;\ 18.9}$	$0.186 \mid 0.410$	4, 94.1	$2.40^{n.s.}$
$x_1x_2x_3x_4$	$856.9 \mid 852.4^{105;\ 9.9}$	$0.188 \mid 0.295$	5, 103.2	3.21^{**}
$x_1x_2x_3x_5$	$870.4 \mid 859.9^{81;\ 14.2}$	$0.086 \mid 0.319$	5, 98.8	3.66^{**}
$x_1x_3x_4x_5$	$856.4 \mid 863.0^{45;\ 24.6}$	$0.192 \mid 0.470$	5, 88.4	2.36^{*}
$x_2x_3x_4x_5$	$856.9 \mid 855.6^{105;\ 9.8}$	$0.188 \mid 0.273$	5, 103.2	2.53^{*}
$x_1x_2x_3x_4x_5$	$858.5 \mid 855.0^{105;\ 11.7}$	$0.193 \mid 0.309$	6, 101.3	2.95^{**}

Analysis of variance (ANOVA) results indicate whether the GWR model made significantly better predictions of basal area than the corresponding OLS regression model. Model variables: x_1 = ASPECT; x_2 = SLOPE; x_3 = ELEV; x_4 = FF1937; x_5 = TSF (*see text for further descriptions*)

[a] Optimal bandwidth, defined by minimization of Akaike Information Criteria

[b] Effective number of parameters

[c] Significance levels: *$0.01 < P < 0.05$; **$0.001 < P < 0.01$

Model 1) and the other with FF1937 and ASPECT ($AIC_c = 847.3$; GWR Model 2) (Table 3). Both models had significantly lower AIC_c values and a higher overall R^2 (0.319 and 0.307, respectively) than the optimal OLS regression model ($AIC_c = 854.3$; $R^2 = 0.175$) (Table 1).

Table 2 OLS regression model of ponderosa pine basal area, Mica Mountain, AZ[a]

Variables	Coefficient	Std. error	t-Value	Limits of 95% CI	
				Lower	Upper
(Intercept)	–18.3974	16.1421	$-1.140^{n.s.}$	–50.3870	13.5920
FF1937	2.7718	0.5805	$4.755^{P<0.001}$	1.6215	3.9222
Elevation	0.1210	0.0067	$1.818^{0.05<P<0.10}$	–0.0011	0.2529

[a] Model $R^2 = 0.175$; $P < 0.001$

Table 3 Descriptive statistics of the coefficient values for two GWR models of ponderosa pine basal area, Mica Mountain, AZ

	Coefficient values				
	Minimum	25% Quartile	Median	75% Quartile	Maximum
Model 1					
(Intercept)	3.9392	7.7201	9.7212	14.4029	19.5392
FF1937	–0.7658	1.9250	2.6501	3.5594	5.8821
Model 2					
(Intercept)	5.6361	9.6854	15.9336	17.8510	18.6918
FF1937	0.5052	1.8209	2.6926	3.1572	3.3962
Aspect	–6.9174	–6.2201	–2.8648	2.2123	4.5210

The mean absolute value, standard error and minimum error of model residuals were also all lower for the GWR models (*Initial model residuals*; Table 4). While several GWR models had a higher R^2 than the two models selected, those models generally suffered from overfitting, as was shown by a greater number of effective parameters and higher AIC_c values (e.g., the model with ELEV and TSF: $R^2 = 0.426$; $AIC_c = 863.9$; effective parameters = 22.2).

GWR model coefficient stationarity

Regression coefficients for four of the five predictor variables displayed significant non-stationarity (Monte Carlo test: ASPECT, ELEV and TSF, $P < 0.0001$; FF1937, $0.001 < P < 0.01$; SLOPE: n.s.), indicating spatial variation in the relationships between ponderosa pine basal area and the predictor variables. The intercept also exhibited significant non-stationarity in all GWR models (Monte Carlo test: ASPECT, ELEV, FF1937: $P < 0.05$; SLOPE, TSF: $0.05 < P < 0.10$).

Analysis of spatial autocorrelation in model residuals supported the better fit of the GWR models. All three models had residuals that were significantly correlated at a lag distance of 0–150 m (Fig. 3), but the strength of the correlations differed between the OLS regression model (Moran's I = 0.774; Monte Carlo test: $P < 0.001$) and the GWR models (Model 1: Moran's $I = 0.502$; Model 2: Moran's $I = 0.487$; both $0.01 < P < 0.05$). The low number of neighbor pairs ($n = 20$) may have influenced the results at this lag distance. Residuals were still highly correlated at the 150–300 m lag distance for the OLS regression model (Moran's $I = 0.372$; $0.001 < P < 0.01$; $n = 75$ neighbor pairs) but not the GWR models. Beyond this distance, there were only weak negative correlations at unpredictable distances in the OLS regression model.

Spatial modeling and data exploration

Variation in the regression coefficients and t-values for the GWR model with FF1937 and ASPECT illustrate the nature and strength of the non-stationary relationships between ponderosa pine and the predictors (Fig. 4). Rather than using t-values of 1.96, 2.58 and 3.29 (correspond-

Table 4 Descriptive statistics for the absolute values of residuals from the initial model calculations and the leave-one-out validations

Model	Mean/median	Std. error	Minimum	Maximum
Initial model residuals				
$OLS_{FF1937, ELEV}$	8.23/6.69	0.572	0.226	25.757
GWR_{FF1937}	7.47/6.29	0.521	0.118	26.175
$GWR_{FF1937, ASPECT}$	7.62/6.53	0.514	0.055	24.797
Validation data residuals				
$OLS_{FF1937, ELEV}$	8.83/9.19	0.609	0.310	28.928
$GWR_{FF1937, ASPECT}$	8.25/7.29	0.556	0.207	26.050

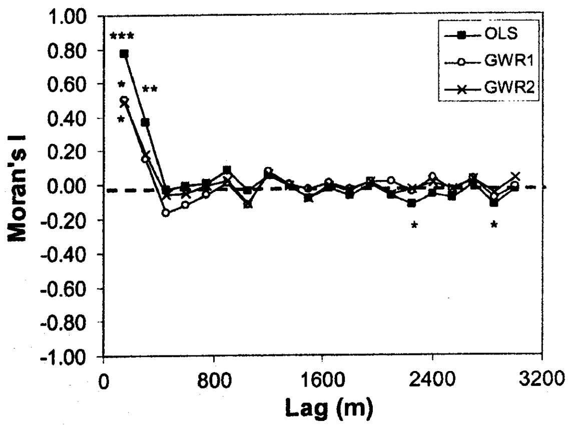

Fig. 3 Spatial autocorrelation (Moran's *I*) of model residuals for the OLS regression model with elevation and fire frequency (OLS), geographically weighted regression model with only fire frequency (GWR1) and GWR model with fire frequency and aspect (GWR2). Significant autocorrelation values: *0.01 < P < 0.05; **0.001 < P < 0.01; ***P < 0.001

ing to the 0.05, 0.01 and 0.001 significance levels in global models, respectively) as cutoffs in Fig. 4, we used Bonferroni-adjusted critical values of the *t*-distribution to avoid potential problems caused by multiple hypothesis tests. This conservative correction adjusts the critical value of the test upwards by setting a new significance level equal to the original significance level (e.g. 0.05) divided by the effective degrees of freedom in the GWR model (*see* Fotheringham et al. 2002). In this case, the critical values of t that we employed were 2.76 (0.05 level), 3.33 (0.01 level) and 3.99 (0.001 level), although 1.96 is also illustrated for reference.

With respect to individual variable patterns, regression coefficients for FF1937 were positive for the entire study area, supporting the global regression result that linked higher fire frequency to higher ponderosa pine basal area, but not significant everywhere. The greatest effect of fire frequency (as indicated by the highest regression coefficient) was at the highest elevations near the peak of Mica Mountain (an area of moderately high frequency; Fig. 2) and on the north-facing slope along the northern boundary of the study area (an area of low fire frequency; Fig. 2). Ecologically speaking, these are areas where fire-tolerant ponderosa pine potentially co-occurs with moderately fire-tolerant Douglas fir and fire-intolerant white fir, and the number of fires might thus be expected to exert a stronger influence on ponderosa pine abundance. Fire frequency was not a significant predictor on the large, lower elevation, southwest-facing slope below the summit.

Changes in the coefficient sign for ASPECT indicate that basal area was positively associated with slope orientation in some locations within the study area and negatively associated in others, although ASPECT was only a significant predictor near the summit of Mica Mountain and on the adjacent north-facing slope (Fig. 4). At these locations, decreased ASPECT values (which correspond to increased 'southness' and exposure) had large effects on predicted values of ponderosa pine. This result was intuitive in that ponderosa pine is relatively shade intolerant and prefers warmer and more xeric sites and would thus be favored on more exposed slopes in the highest elevation areas and on the largely sheltered north slope. The identification of areas where aspect is significantly related to ponderosa pine basal area is important because this relationship was missed by the global regression, which showed no significant relationship between the two.

Values of the intercept, which appeared to capture variability in ponderosa pine related to the elevation gradient (Fig. 2), varied significantly throughout the study area, ranging from a low of 5.6 in the southern third of the study area (where the intercept was only marginally significant) to more than 18.0 in the northern area (Fig. 4).

While the general patterns of predicted ponderosa pine basal area were comparable for the optimal OLS regression (with FF1937 and ELEV; Fig. 5a) and GWR (with FF1937 and ASPECT; Fig 5b) models, the latter had a more complex pattern with finer spatial variability in ponderosa pine abundance. Predictions by OLS regression were higher on north-east facing slopes in the north-slope drainages and on the west-facing ridgeline arcing from the center of the study area to its southern-most extent (darker shaded areas in Fig. 5c) while predictions by GWR were greater in the western portion of the study area, on the steep east-facing slope along the eastern edge of the study area and on west-facing slopes in the dissected drainages running off the north face of Mica Mountain (darker shaded areas in Fig. 5d).

Validation

For the validation, we again compared results from the OLS regression model with FF1937 and ELEV and the GWR model with FF1937 and ASP. The relationship between predicted and observed values was significant for both models, but predictions using OLS regression had a lower R^2 (0.08) and varied little over the range of observed values, particularly underestimating values at the high end of the observed basal area range (Fig. 6a). In comparison, the R^2 for GWR predictions was 0.19, which increased to 0.21 if one outlier from an extreme topographic setting was removed. GWR did a much better job, in particular, of predicting locations with high ponderosa pine abundance (Fig. 6b).

Of the 113 plots, the GWR model provided a better prediction of observed ponderosa pine basal area for 49 plots (43.4%) while the OLS regression model was a better predictor for 38 (33.6%) (Fig. 6c). For the other 26 plots, predictions differed by less than 1 m^2 ha^{-1}, with 13 (50%) being better predicted by GWR. Overall, median residual error was 26% higher for the OLS model (*Validation Data Residuals*; Table 4). A comparison of model residuals for each plot shows that the amount of misprediction by OLS regression (i.e., the residual value) was typically much higher for plots predicted better by GWR while the reverse was not typically true (Fig. 6c).

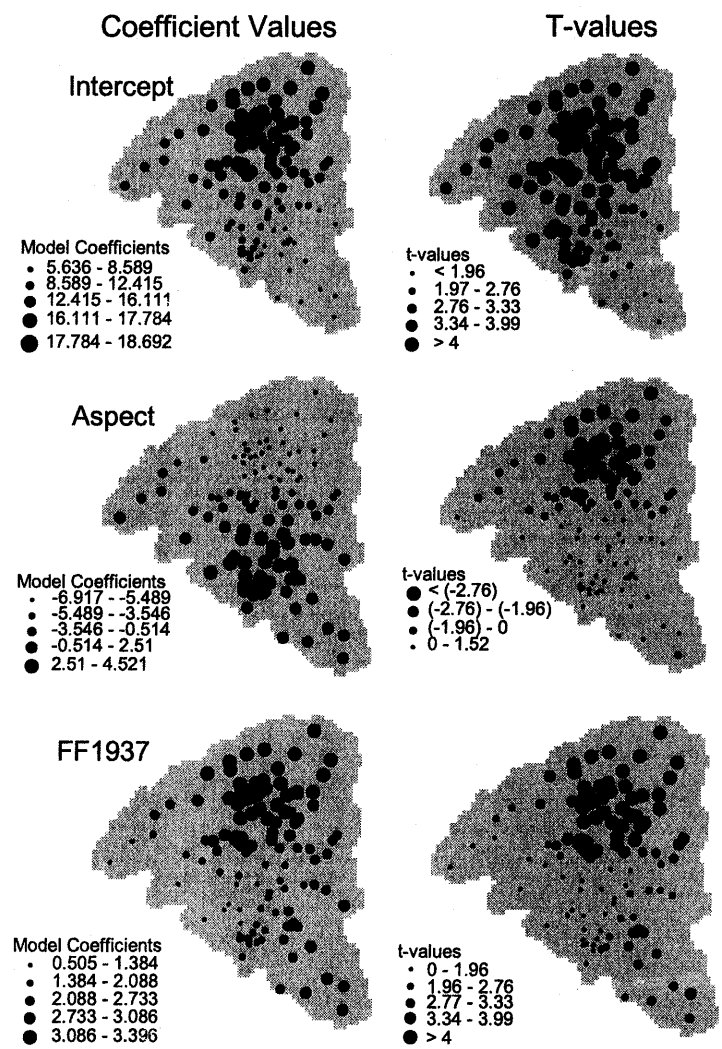

Fig. 4 Coefficients and *t*-values for the final GWR model incorporating fire frequency and aspect. Indicated locations are the 113 field sampled plots

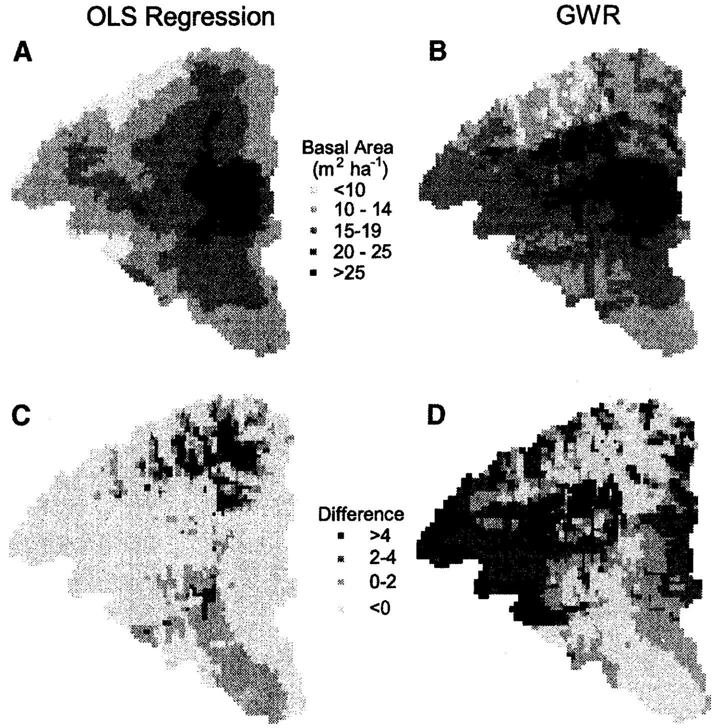

Fig. 5 Surfaces of ponderosa pine basal area within the Mica Mountain study area: (**a**) predicted by OLS regression with fire frequency and elevation, (**b**) predicted by GWR with fire frequency and aspect, (**c**) areas where predicted basal area was higher for the OLS regression model, including the difference between the models, and (**d**) areas where predicted basal area was higher for the GWR model, including the difference between the models. All values are in $m^2\ ha^{-1}$

In other words, OLS regression tended to badly predict values on sites that were predicted better by GWR while predictions by GWR on sites better predicted by OLS regression were often only slightly worse. It is similarly worth noting that most sites that were predicted better by OLS regression had low residuals for both models (e.g., residual values <10) whereas the harder to predict sites (those with a higher residual for both models) were predicted much better by GWR.

Discussion

Understanding and predicting ponderosa pine basal area

By comparing patterns of aspect and fire frequency (Fig. 2), GWR regression coefficients and *t*-values (Fig. 4), and predicted ponderosa pine abundance (Fig. 5), it is possible to better understand the complex relationships shaping ponderosa pine distribution in the study area than by using global data analysis alone. In particular, the importance of local variations in aspect, which were missed by the global model, becomes clear.

Along the northern edge of the study area, predicted basal area starts high on the basis of a high intercept value (>16 $m^2\ ha^{-1}$; Fig. 4). Even though the fire frequency coefficient is very high, predicted ponderosa pine basal area is not increased substantially because none of this area has experienced more than two fires. Conversely, predicted values are decreased substantially by the high ASPECT values associated with this north-facing slope, except on the more exposed west-facing slopes along drainages. These latter areas are obvious as locations where GWR predictions were much higher (Fig. 5d). The net result of these interactions is that predicted basal area is generally low to moderate in this area (Fig. 5b), as would be expected given the sheltered slope aspect and low fire frequency.

The western part of the study area is primarily a south-facing, low-to mid-elevation slope with a low-to moderate-fire frequency (mostly 1–2 fires since 1937) and intermediate time since fire. Predicted basal area (ca. 15–20 $m^2\ ha^{-1}$; Fig. 5b) is shaped largely by the intercept value (ca. 15 $m^2\ ha^{-1}$; Fig. 4) as neither ASPECT nor FF1937 has an especially strong relationship to basal area in this area (Fig. 4).

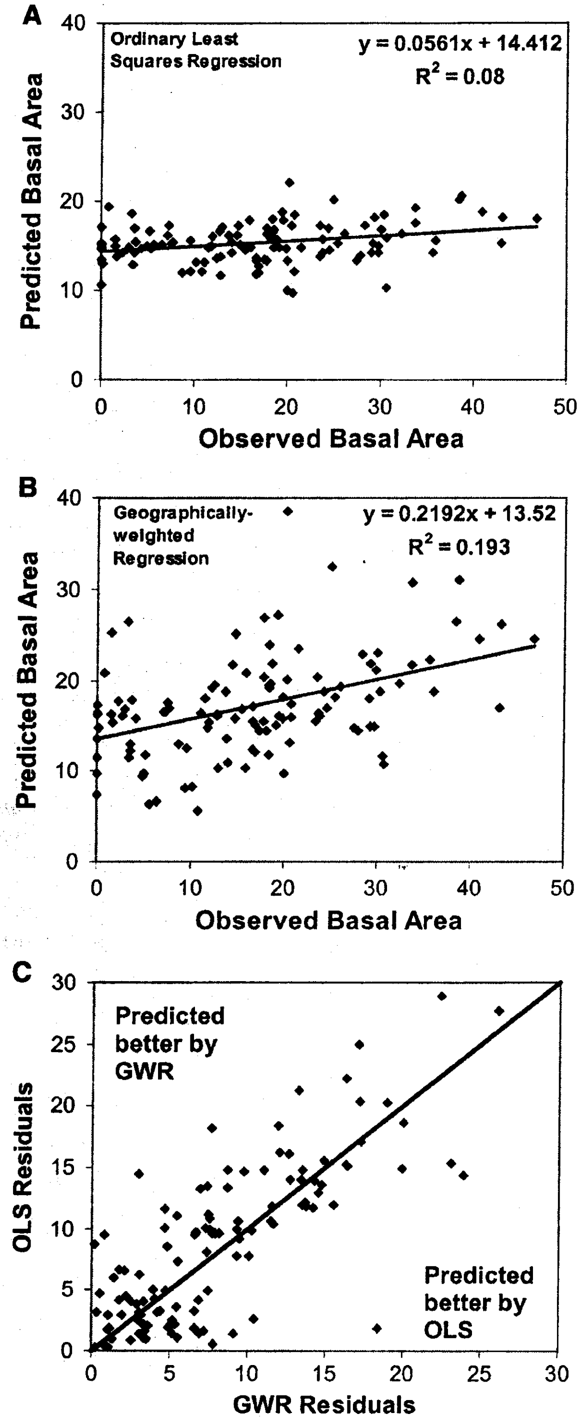

Fig. 6 Comparisons of: (**a**, **b**) actual versus predicted values of ponderosa pine from 113 leave-one-out regressions, and (**c**) residuals from the OLS regression models versus the GWR models. For the ordinary least squares regression analyses, a model with elevation and fire frequency was used; for the GWR analyses, a model with fire frequency and aspect was used

The southern portion of the study area straddles a ridge running from Mica Mountain to Rincon Peak (Fig. 1) and includes lower elevation, southwest- and east-facing slopes with a moderate fire frequency (2–4 fires since 1937) (Fig. 2). Most of this area has burned within the last 10 years. Predicted basal area is initially low as this area has the lowest intercept value (ca. 5–10 m^2 ha^{-1}). While ASPECT is not significantly related to basal area in this location, fire frequency exerts a strong and significant positive influence. The result is a low- to moderate-predicted basal area.

Finally, the areas of highest ponderosa pine basal area predicted by both GWR and OLS regression are located in the central portion of the study area (Fig. 5a, b), which includes the southern flank of the Mica Mountain summit (Fig. 1). Intercept values are nearly as high as those for the first area (ca. 12–16 m^2 ha^{-1}; Fig. 4), but these values are increased substantially because of the high fire frequency (generally 5–7 fires since 1937; Fig. 2). Although the strength of the relationship was weak, aspect generally increased basal area on more southerly slopes.

The value of examining non-stationarity

GWR analyses illustrated the non-stationary relationships between ponderosa pine and topography and fire history, and in so doing, predicted significant local variations of vegetation pattern that were missed by the global model. For example, GWR predicted variations in ponderosa pine abundance associated with fine-scale drainage patterns on the north-facing slope of Mica Mountain. Differences in abundance between opposing exposed and sheltered slopes are ecologically intuitive but were missed by OLS regression, which showed no significant relationship between ponderosa pine and aspect.

While we chose to compare results from GWR with those from OLS regression, we recognize that more sophisticated methodologies than OLS regression exist, including classification and regression tree analysis, multivariate adaptive regression splines, artificial neural networks, and spatial autoregressive models, among many others (e.g., Prasad and Iverson 2000; Moisen and

Freschino 2002). Such techniques are generally superior to OLS regression in terms of both explanation and prediction for a range of reasons, although which technique is 'best' is highly case specific. While comparisons of GWR with other methodologies are certainly warranted, we did not do so in this study because our emphasis was explicitly on demonstrating the significance of non-stationarity, and GWR represents a direct extension of the OLS regression model. Further, one of the reasons that the above-mentioned techniques are often superior to OLS regression is that they are not constrained by the assumption of linearity between the predictors and the response variable. We examined scatterplots of the raw data, analyzed model residuals, and included log-transformed and quadratic versions of the predictors in the OLS regression models and found no evidence of non-linearity in any of these analyses, which would suggest that in this case, GWR outperformed OLS regression because of non-stationarity rather than a better ability to accommodate non-linear relationships. However, in the circumstances of what is likely an under-specified model, it is not surprising that GWR provides a better solution because it 'flexes' to fit the data locally. The case for GWR would be made much stronger if it was shown to improve even very good OLS regression models.

An additional limitation to consider when assessing the performance of GWR is the location-specific nature of the variable relationships. Specifically, we would expect GWR to outperform OLS regression only on independent data that lies within the polygon formed by the locations in the model and only then if the spread and density of locations is adequate to encompass local variations in the response and predictor variables.

Finally, in comparing three methods of predictive vegetation modeling (classification trees, artificial neural networks and generalized linear models), Cairns (2001: 3) makes the observation that: "The variability in predictive ability of the three methods tested here indicates that there may not be a single best predictive method. Rather it may be important to use a suite of predictive models to help understand the environment—vegetation relationships." We believe that the usage of GWR offers an analogous lesson in that the analysis and mapping of localized regression coefficients and associated *t*-values provided unique insights on species controls and predictions beyond those gleaned by OLS regression alone. GWR analyses thus fall in the realm of exploratory spatial data analysis where the emphasis is on developing hypotheses from the data (Fotheringham et al. 2002). In this way, we see GWR as a valuable complement to the many other global methods currently in use for predictive vegetation modeling.

Conclusions

We examined the value of incorporating non-stationarity of regression coefficients into models of high elevation ponderosa pine basal area through the use of GWR. Our results indicated that GWR explained species–environment relationships better than OLS regression and that predictions for withheld data points were more accurate. The latter is an important finding because few studies applying GWR to ecological phenomena have used it to predict values for independent data points.

Further, GWR provided insights into localized controls of ponderosa pine patterns that were not evident when using a global model and thereby helped to demonstrate the value of examining the non-stationarity of regression coefficients. Non-stationarity is not the same as spatial dependence, although the two may be related. Instead, spatial regression models are sometimes called 'semi-local' models because they incorporate the spatial covariance structure of error terms and are therefore able to capture some local variation in the response variable, but the parameter estimates are nonetheless global in that they apply equally to all data points in the study regardless of location (Jetz et al. 2005).

Landscape ecologists have been quick to embrace the use of not only 'traditional' landscape metrics but also spatial statistics and methods adopted from the physical sciences (e.g., semi-variance, lacunarity, kriging) (e.g., Haines-Young and Chopping 1996; Gustafson 1998). The acceptance of relevant methods from the social sciences, however, has been more limited (e.g., local indices of spatial association such as Moran's I or Geary's C; Anselin 1995). Anselin (2002) recently noted a remarkable growth in the application of spatial models in applied economics, spatial econometrics and geography (see also papers in Getis et al. 2004), some of which would likely be of great interest to landscape ecologists. We believe that this paper and others like it (e.g., Foody 2004, 2005; Zhang and Shi 2004; Zhang et al. 2004, 2005) illustrate the potential contributions that GWR, which has its roots and most of its applications in the social sciences, can make to the study and prediction of broad-scale vegetation patterns.

Acknowledgments We gratefully acknowledge the cooperation of Kathy Schon and Mark Holden from Saguaro National Park as well as the field assistance of Chris Baisan, Jose Iniguez, Ellis Margolis, James Riser, Devin Petry, and Jeff Balmat. JAK would particularly like to thank Gordon Mulligan for comments on the final draft of this paper and for our stimulating discussions of GWR and spatial econometrics. This research was supported in part by funds provided by the Aldo Leopold Wilderness Research Institute of the Rocky Mountain Research Station, Forest Service, U.S. Department of Agriculture.

References

Anselin L (1995) Local indicators of spatial autocorrelation – LISA. Geogr Anal 27:93–115

Anselin L (2002) Under the hood – Issues in the specification and interpretation of spatial regression models. Agr Econ 27:247–267

Bolstad PV, Swank W, Vose J (1998) Predicting Southern Appalachian overstory vegetation with digital terrain data. Landscape Ecol 13:271–283

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol Model 153:51–68

Bowers JE, McLaughlin SP (1987) Flora and vegetation of the Rincon Mountains. Desert Plants 8: 51–94

Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr Anal 28:281–298

Cairns DM (2001) A comparison of methods for predicting vegetation type. Plant Ecol 156:3–18

Chronic H (1983) Roadside geology of Arizona. Mountain Press Pub Co., Missoula, MT, USA

Cliff AD, Ord JK (1981) Spatial processes: models and applications. Pion, London

Davis F, Goetz S (1990) Modeling vegetation pattern using digital terrain data. Landscape Ecol 4:69–80

Diniz-Filho JAF, Bini LM (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. Global Ecol Biogeogr 14:177–185

Diniz-Filho JAF, Bini LM, Hawkins BA (2003) Spatial autocorrelation and red herrings in geographical ecology. Global Ecol Biogeogr 12:53–64

Dirnböck T, Dullinger S (2004) Habitat distribution models, spatial autocorrelation, functional traits and dispersal capacity of alpine plant species. J Veg Sci 15:77–84

Foody GM (2004) Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. Global Ecol Biogeogr 13:315–320

Foody GM (2005) Mapping the richness and composition of British breeding birds from coarse spatial resolution satellite sensor imagery. Int J Remote Sens 26:3943–3956

Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester

Franklin J (1995) Predictive vegetation mapping: Geographic modeling of biospatial patterns in relation to environmental gradients. Prog Phys Geog 19:474–499

Franklin J (2002) Enhancing a regional vegetation map with predictive models of dominant plant species in chaparral. Appl Veg Sci 5:135–146

Gavin DG, Hu FS (2005) Bioclimatic modelling using Gaussian mixture distributions and multiscale segmentation. Global Ecol Biogeogr 14:491–501

Getis A, Mur J, Zoller HG (eds.) (2004) Spatial econometrics and spatial statistics. Palgrave Macmillan, Basingstoke

Guisan A, Zimmermann N (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Gustafson EJ (1998) Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1:143–156

Haines-Young R, Chopping M (1996) Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. Prog Phys Geogr 20:418–445

Hawkins BA, Porter EE (2003) Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America. Global Ecol Biogeogr 12:475–481

Humphries HC, Bourgeron PS (2003) Environmental responses of *Pinus ponderosa* and associated species in the south-western USA. J Biogeogr 30:257–276

Iverson LR, Prasad AM (1998) Predicting abundance of 80 tree species following climate change in the eastern United States. Ecol Monogr 68:465–485

Jensen ME, Dibenedetto JP, Barber JA, Montagne C, Bourgeron PS (2001) Spatial modeling of rangeland potential vegetation environments. J Range Manage 54:528–536

Jetz W, Rahbek C, Lichstein JW (2005) Local and global approaches to spatial data analysis in ecology. Global Ecol Biogeogr 14: 97–98

King RS, Richardson CJ, Urban DL, Romanowicz EA (2004) Spatial dependency of vegetation-environment linkages in an anthropogenically influenced wetland ecosystem. Ecosystems 7:75–97

Koenig WD (1999) Spatial autocorrelation of ecological phenomena. Trends Ecol Evol 14:22–26
Legendre P (1993) Spatial autocorrelation: problem or new paradigm?. Ecology 74:1659–1673
Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. Vegetatio 80:107–138
Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. Ecography 23:101–113
Lichstein JW, Simons TR, Shriner SA, Franzreb KE (2002) Spatial autocorrelation and autoregressive models in ecology. Ecol Monogr 72:445–463
Malanson GP (1997) Effects of feedbacks and seed rain on ecotone patterns. Landscape Ecol 12:27–38
Meentemeyer RK, Moody A, Franklin J (2001) Landscape-scale patterns of shrub-species abundance in California chaparral – the role of topographically mediated resource gradients. Plant Ecol 156:19–41
Miller J (2005) Incorporating spatial dependence in predictive vegetation models: residual interpolation methods. Prof Geogr 57:169–184
Miller J, Franklin J. (2002) Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. Ecol Model 157:227–247
Moisen GG, Frescino TS (2002) Comparing five modelling techniques for predicting forest characteristics. Ecol Model 157:209–225
Ohmann JL, Gregory MJ (2002) Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. Can J Forest Res 32:725–741
Phillips JD (2002) Global and local factors in earth surface systems. Ecol Model 149:257–272
Prasad AM, Iverson LR (2000) Predictive vegetation mapping using a custom built model-chooser: comparison of regression tree analysis and multivariate adaptive regression splines. In: 4th international conference on integrating GIS and environmental modeling (GIS/EM4): problems, prospects and research needs. Banff, Alberta, Canada
Sawada M. (1999) ROOKCASE: An Excel 97/2000 Visual Basic (VB) add-in for exploring global and local spatial autocorrelation. B Ecol Soc Amer 80:231–235
Swantek PJ, Halvorson WL, Schalbe CR (1999) GIS database development to analyze fire history in southern Arizona and beyond – an example from Saguaro National Park. Technical Report 61, U.S. Geological Survey, Cooperative Park Studies Unit, Tucson, AZ
Wang Q, Ni J, Tenhunen J (2005) Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. Global Ecol Biogeogr 14:379–393
Wimberly MC, Spies TA (2001) Modeling landscape patterns of understory tree regeneration in the Pacific Northwest, USA. Appl Veg Sci 4:277–286
Zhang L, Shi H (2004) Local modeling of tree growth by geographically weighted regression. Forest Sci 50:225–244
Zhang L, Bi H, Cheng P, Davis CJ (2004) Modeling spatial variations in tree diameter–height relationships. Forest Ecol Manage 189:317–329
Zhang LJ, Gove JH, Heath LS (2005) Spatial residual analysis of six modeling techniques. Ecol Model 186:154–177