Contingent effects of water balance variation on tree cover density in semiarid woodlands

Andrew J. Kerkhoff*, Scott N. Martens, Gregory A. Shore and Bruce T. Milne

ABSTRACT

Aim The local distribution of woody vegetation affects most functional aspects of semiarid landscapes, from soil erosion to nutrient cycling. With growing concern about anthropogenic climate change, it has become critical to understand the ecological determinants of woody plant distribution in semiarid landscapes. However, relatively little work has examined the determinants of local variation in woody cover. Here we examine water balance controls associated with patterns of tree cover density in a topographically complex semiarid woodland.

Location Los Pinos Mountains, Sevilleta National Wildlife Refuge LTER, New Mexico, USA.

Methods To explore the relationship between local water balance variation and tree cover density, we used a combination of high-resolution (1 × 1 m), remotely sensed imagery and quantitative estimates of water balance variation in space and time. Regression tree analysis (RTA) was used to identify the environmental parameters that best predict variation in tree cover density.

Results Using six predictor variables, the RTA explains 39% of the deviance in tree cover density over the landscape. The relationship between water balance conditions and tree cover density is highly contingent; that is, similar tree cover densities occur under very different combinations of water balance parameters. Thus, the effect of one environmental parameter on tree cover density depends on the values of other parameters. After tree cover density is adjusted for water balance conditions, residual variation is related to tree cover density in the neighbourhood of a particular location.

Conclusion In semiarid landscapes, vegetation structure is largely controlled by water supply and demand. Results presented here indicate that localized feedbacks and site-specific historical processes are critical for understanding the responses of semiarid vegetation to climate change.

Keywords Bioclimatology, dispersal limitation, disturbance, landscape pattern, piñon-juniper, regression tree analysis, savanna, vegetation feedbacks, USA.

INTRODUCTION

Semiarid landscapes can often be viewed as spatial mosaics of woody (tree, shrub) canopy patches and an intercanopy zone that is itself a mixture of grasses, herbs, and bare soil (Montana, 1992; Milne et al., 1996; Breshears & Barnes, 1999; Martens et al., 2000). In semiarid landscapes, the relative cover of woody plants affects many ecosystem properties, including rates of transpiration (Joffre & Rambal, 1993; Schulze et al., 1996), production (Belsky, 1994; Aguiar et al., 1996), infiltration (Rietkerk et al., 1997; Milne et al., 2002), erosion (Reid et al., 1999), and nutrient cycling (Padien & Lajtha, 1992; Schlesinger et al., 1996). The strong feedbacks between climate, soil, and vegetation in semiarid regions may make them particularly sensitive to anthropogenic changes in climate and land use. Thus, understanding the determinants of woody plant cover in semiarid landscapes is not only a basic ecological question, but also an important component of global change research.
Vegetation patterns are the manifestation of dynamic ecological processes (e.g. dispersal, disturbance, competition) operating in the context of abiotic environmental gradients (Urban et al., 2000). Consistently, and especially in semiarid regions, the most important environmental determinants of vegetation patterns are related to site water balance (Stephenson, 1990; Frank & Inouye, 1994; Knapp & Smith, 2001; Paruelo et al., 2001). However, few studies have examined the influence of local water balance variation on fine scale changes in vegetation structure (Stephenson, 1998; Urban et al., 2000). Here, we examine the spatial distribution of tree cover in a piñon-juniper woodland as a function of water balance variation over a topographically heterogeneous landscape, in order to understand its importance as a determinant of woodland vegetation structure.

Piñon-juniper woodlands are one of the most widespread vegetation types in western North America, covering approximately 24 million ha (Miller & Wigand, 1994). Major, persistent (c. 10 year) droughts have been shown to strongly affect the distribution and demography of woody species throughout the American South-west, both directly by increasing mortality and limiting recruitment (Allen & Breshears, 1998; Swetnam & Betancourt, 1998), and indirectly through interactions with regional fire regimes (Swetnam & Betancourt, 1990; Allen & Breshears, 1998; Swetnam & Betancourt, 1998). Since the mid 19th century, woodlands have reportedly been expanding into grasslands (Tausch et al., 1981), prompting concern about reductions in viable rangeland and accelerated soil loss due to increased rates of erosion (Belsky, 1996). The expansion of woody cover has traditionally been attributed to a combination of overgrazing, fire suppression, and climate change, but the fundamental physical conditions associated with the presence of woodland tree cover remain unquantified.

In the face of global climate and land use changes, similar concerns have become common in semiarid regions around the world. Predicting the impacts of global change in semiarid landscapes requires a more complete accounting of water balance variation over complex terrain. Here, we show that woody cover is related to both the magnitude and variation in water balance over tens of metres in space and months to years in time. Further, our data imply that understanding local ecological feedbacks and historical processes may be very important for understanding the responses of semiarid landscapes to global change.

**METHODS**

**Study area**

Our study site is a piñon-juniper woodland covering c. 10,000 ha in the Los Pinos Mountains, New Mexico (34°21’N latitude, 106°32’W longitude, mean elevation 1919 m), within the Sevilleta National Wildlife Refuge (NWR). Originally a Spanish Land Grant, the Sevilleta was a private cattle ranch from the mid 1930s until 1973, when it was made a National Wildlife Refuge and grazing was ceased. Long-term tree ring chronologies from the central part of the Los Pinos woodland document episodes of mortality and recruitment mediated by decadal climate fluctuations (Swetnam & Betancourt, 1998; Milne et al., 2003), but only rare, localized fires (J. L. Betancourt, personal communication).

We delineated the boundaries of the Los Pinos woodland using the Sevilleta Long-term Ecological Research program (LTER) digital vegetation map, which was developed from a classification of multitemporal (1987–93) Landsat Thematic Mapper imagery (Muldavin et al., 1998). Using Arcview GIS software, we resampled the vegetation map from its original 28.5 m resolution to 30 m, in order to correspond to the digital elevation model (DEM) used in the estimation of many of the water balance parameters. We then selected all vegetation map cells in the Los Pinos region that were classified as either juniper savanna (dominated by Juniperus monosperma) or piñon woodland (dominated by Pinus edulis). Finally, we compiled environmental data and estimated tree cover density for each 900 m² cell, yielding a total of 113,842 sample points for analysis.

**Environmental parameters**

We compiled eight water balance variables that potentially influence the distribution of tree cover density (Table 1). While the list is not exhaustive, the variables represent many aspects of site water balance. We chose not to include elevation, slope, and aspect as direct predictors in the model, under the assumption that topography affects vegetation primarily through its effects on water supply and demand. To test this assumption, we performed a second analysis using elevation, slope, aspect and porosity as predictors, for comparison to the model that used estimates of water balance parameters.

All climatic data were drawn from the Sevilleta LTER data archive, which is collected at nine weather stations distributed throughout the Sevilleta NWR. Six of the stations have been in operation since 1989, with the remaining three added in 1992, 1998, and 1999. Two of the stations are within the Los Pinos woodland. Over the period of the available record (1989–99), we compiled daily values of total precipitation (PPT), mean temperature, mean saturation vapour pressure deficit, and mean wind speed. We modelled the spatial variation of climate parameters using a combination of empirical elevation lapse functions and inverse distance weighting. We developed lapse functions for

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPT</td>
<td>Annual Precipitation (cm)</td>
</tr>
<tr>
<td>APPTCV</td>
<td>PPT Coefficient of variation (%), annual time scale</td>
</tr>
<tr>
<td>MPPTCV</td>
<td>PPT Coefficient of variation (%), monthly time scale</td>
</tr>
<tr>
<td>APET</td>
<td>Annual potential evapotranspiration (cm)</td>
</tr>
<tr>
<td>APETCV</td>
<td>PET Coefficient of variation (%), annual time scale</td>
</tr>
<tr>
<td>MPETCV</td>
<td>PET Coefficient of variation (%), monthly time scale</td>
</tr>
<tr>
<td>POR</td>
<td>Soil porosity</td>
</tr>
<tr>
<td>TMI</td>
<td>Topographic moisture index (see text)</td>
</tr>
</tbody>
</table>
each Julian day using the station mean values for temperature and precipitation over the period of the record. For each day of the record, we lapsed all station measurements to an arbitrary reference elevation of 1625 m. Next, we estimated reference elevation values of the climate parameters for each cell as the inverse distance weighted mean of the station values. Finally, for temperature and precipitation, we lapsed each cell back from the reference elevation to its true elevation.

We calculated daily potential evapotranspiration (PET) for each cell using a version of the Penman equation (Penman, 1948).

\[
\lambda_{\text{PET}} = \frac{1}{\Delta + \gamma} \left[ \Delta (R_n - G) + K_v (a_v - b_v) D \right]
\]

Here \( \lambda \) is the latent heat of vapourisation of water (MJ/kg), \( \Delta \) is the slope of the vapour pressure curve (kPa/C), \( \gamma \) is the psychrometric constant (kPa/C), \( R_n \) is net solar radiation (shortwave + longwave, MJ/m²/d), \( G \) is soil heat flux (MJ/m²/d), \( a \) is wind speed (m/s) and \( D \) is the saturation vapour pressure deficit (kPa). The constant \( K = 6.43 \) balances units for PET in mm/day, while windspeed coefficients \( a = 1 \) and \( b = 0.357 \) were empirically derived by Penman (1948).

Daily shortwave radiation was derived from the inverse distance weighted mean measures of solar radiation, adjusted for the elevation (atmospheric transmissivity), slope, and aspect of each cell, drawn from the DEM. Daily net longwave radiation was calculated according to Allen et al. (1998). Wind speed and saturation vapour pressure deficit were drawn from the climate data described above. We then summed the daily PET and PPT values to produce monthly and annual totals. Average annual totals (APPT, APET, in cm) were used as predictor variables.

Because piñon-juniper woodlands span a large portion of the continuum from closed forest to treeless grassland, and because the transition from forested to nonforest ecosystems is related to interannual and seasonal variability in the inputs of water and energy (Frank & Inouye, 1994; Knapp & Smith, 2001), we also included measures of variability in the analysis. Thus, for both PET and PPT, we calculated coefficients of variation (standard deviation/mean) on two time scales for use as predictors in the model. The CVs of annual totals of PPT and PET (APPTCV, APETCV) reflect interannual climate variability, while seasonal (intra-annual) variability is captured using the CVs of monthly totals (MPPTCV, MPETCV).

While the low of precipitation (mean ± s.d. = 34.5 ± 3.1 cm/year) and large PET (mean ± s.d. = 154.1 ± 5.6 cm/year) of the site preclude the presence of flowing water, isolated, high-intensity storms may produce local overland flow. The capture and retention of these episodic moisture inputs may be important to tree establishment and growth. In order to index the propensity of a site to receive and capture runon we calculated a topographic moisture index (Moore et al., 1991; Quinn et al., 1995; Wolock & McCabe, 1995) as TMI = ln(\( \alpha / \tan(\beta) \)), where \( \alpha \) is the upslope contributing area per unit contour length, and \( \beta \) is the slope of the cell, both of which were drawn from the DEM using Arcview Spatial Analyst. Values of TMI are large in poorly drained basins and small on well drained ridges; negative values occur when \( \alpha > \tan(\beta) \).

Variation in soil properties may exert a large influence on site water balance as well as fertility. Unfortunately, spatially distributed soils data for the study site are limited to a coarsely drawn Soil Conservation Service map (Johnson, 1988), which was digitized by the Sevilleta LTER. We included only porosity (POR) values for each site, as other components of the soil description, e.g. percent sand, also varied discretely by soil class, and thus provided no additional information.

**Tree cover density estimation**

Tree cover density was estimated from high resolution Airborne Data Acquisition and Registration (ADAR) imagery, acquired by the Sevilleta LTER in late January 1998. The Positive Systems ADAR System 5500 sensor yields data in four spectral bands that span visual and near infrared. Image pixel resolution is 1 m, so each cell of the study site contained 900 ADAR cells, and tree canopies were clearly visible. We classified the image as tree/nontree using a supervised classification algorithm in the Imagine software package. We then calculated tree cover density as the areal fraction of each 30 m × 30 m cell covered by trees. Unfortunately, we could not discriminate reliably between juniper and piñon crowns. Classification errors undoubtedly occurred, due to unavoidable effects of slope shading and because the subimages that make up the ADAR image mosaic were not radiometrically corrected by the image provider before being registered. However, close visual comparison of several different regions of the classified image to the original ADAR image suggested no systematic patterns of deviation, owing primarily to the extremely strong signal of the evergreen trees in the midwinter scene.

**Regression tree analysis**

We chose regression tree analysis (RTA, also known as classification and regression trees, or CART) to examine the effects of water balance gradients on the distribution of tree cover density. RTA is able to capture contingent relationships among predictor variables, and the resulting trees represent a natural hierarchy of influences on the response (Iverson & Prasad, 1998).

The RTA procedure recursively partitions observations, i.e. cells, into subsets that are maximally different from one another with respect to the response variable, i.e. tree cover density. The RTA algorithm finds the single best predictor variable and chooses a split that minimizes the residual sum of squares, summed over the resulting subsets. Variables may be used more than once, and the procedure continues recursively until no further reduction in residual deviance is attained. Each terminal subset of cells is called a leaf. The predicted value at each leaf is the mean of the response variable within the leaf. Because RTA models can be over-fitted, we used a process of tenfold cross-validation (Venables & Ripley, 1999) to reduce our tree to a statistically valid size. All calculations were performed using S-Plus statistical software.

We mapped residuals from the RTA procedure back onto the landscape in order to look for systematic patterns of variation unaccounted for by the environmental parameters. In particular,
we examined the relationship between residual tree cover density and elevation, slope, and aspect, to make sure that the effects of landscape topography were adequately reflected by our water balance estimates. Finally, to examine the feedback effects of trees themselves, we examined residual tree cover density as a function of the tree cover density in the neighbourhood around a particular cell. We tallied tree cover densities in windows of 90 × 90 m (3 × 3 cells) around each woodland cell. Cells outside the study site were assumed to contain no trees, and the tree cover within the focal cell was not used to calculate the neighbourhood tree cover density. Because the goal of the residual analysis was simply to look for patterns in the residual data, we used locally weighted linear regression (loess), with the linear fit calculated over 1% increments of the data.

RESULTS

Tree cover density in the Los Pinos woodland varied over the entire range from 0 to 1. Spatial variation in both tree cover density and water balance parameters was clearly visible (Fig. 1). Relationships among the various predictors and tree cover density were complex and frequently nonlinear (Fig. 2). Thus, RTA seemed appropriate to disentangle some of these highly convoluted relationships.

The original regression tree was trimmed to 104 leaves by the cross-validation procedure, and it explained approximately 50% of the variation in tree cover density, using all eight predictor variables. While the original regression tree was statistically valid, its size rendered it so complicated as to be uninterpretable. To simplify the analysis, we repeatedly ‘pruned’ the regression tree to find the optimal subtree for each size, from two to 103 leaves, and examined the deviance explained as a function of regression tree size (Fig. 3). Because larger regression trees clearly exhibited a diminishing return in explanatory power, we limited our final regression tree to 13 leaves (Fig. 4), which explained 39% of the variation in tree cover density, using six of the eight predictors (Table 2). In comparison, the RTA using elevation, slope, aspect, and porosity as predictors resulted in a regression tree of 121 leaves and explained only 34% of the variation in tree cover density.

Interpreting the regression tree is relatively straightforward. Each node (branching point) has an inequality attached to it. If the inequality is true, the data follow the left branch, if it is false, the right. Each leaf is labelled with the mean tree cover density of the cells in that leaf. The relative length of each branch corresponds to the proportion of the deviance explained by the preceding node. Thus, the ‘prediction’ that results from the regression tree amounts to assigning each cell to one of 13 tree
Figure 2 Scatterplots of the eight water balance parameters used in the RTA and tree cover density. A random 1% subsample of the data was used in all of the plots to make patterns in data density more visible.
cover density classes based on its particular combination of water balance parameters.

The first branching was based on annual precipitation, and the next two nodes were both based on aspects of precipitation variability. Relatively dry cells segregated according to the seasonal variability of precipitation (MPPTCV), while the state of wetter cells depended on interannual variability (APPTCV). Subsequent nodes along the wetter main branch were decided by the magnitude and seasonality of evaporative demand (APET and MPETCV), while the terminal nodes depended either on precipitation or porosity. Mean tree cover densities in the leaves ranged from 0.10 to 0.53, and leaves in divergent parts of the regression tree frequently exhibited similar mean densities.

Residual variation in tree cover density was clearly nonrandom in space (Fig. 5), but it showed no trends associated with any topographic characteristics (Fig. 6). However, residual tree cover density did show a positive trend with neighbourhood tree cover density (Fig. 7). Tree cover density tended to be lower than predicted by the RTA when neighbourhood tree cover density was less than ~0.35 and higher than predicted as the neighbourhood density increased above this threshold.

**DISCUSSION**

Tree cover density was primarily related to annual precipitation and PET, as one would predict from many other gradient analyses of vegetation patterns (Stephenson, 1998; Urban et al., 2000) and the generally water-limited status of piñon and juniper trees (Lajtha & Getz, 1993; Linton et al., 1998). Direct water balance estimates and the inclusion of temporal variability allowed the

---

**Table 2** Water balance parameters selected in the regression tree bearing 13 leaves (Fig. 4) and the proportion of the deviance explained by each

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of nodes</th>
<th>Explained deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPT</td>
<td>2</td>
<td>0.255</td>
</tr>
<tr>
<td>MPPTCV</td>
<td>4</td>
<td>0.05</td>
</tr>
<tr>
<td>APPTCV</td>
<td>1</td>
<td>0.03</td>
</tr>
<tr>
<td>MPETCV</td>
<td>2</td>
<td>0.025</td>
</tr>
<tr>
<td>APET</td>
<td>1</td>
<td>0.015</td>
</tr>
<tr>
<td>POR</td>
<td>2</td>
<td>0.015</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>0.39</td>
</tr>
</tbody>
</table>
analysis to explain much more variation in tree cover density than the analysis using topographic and soil variables alone. Thus, even relatively simple estimates of water balance gradients are valuable for explaining fine scale patterns of tree cover.

Tree cover density did not vary in a simple way along any single environmental gradient (Fig. 2). The space filling relationship that tree cover density exhibited with many of the water balance parameters is similar to the distribution of basal area along an elevational gradient observed for several species of trees in the Sierra Nevada (Urban et al., 2000). These relationships suggest that water balance gradients constrain the upper limit of tree cover density while other ecological processes, such as disturbance, dispersal, and competition, determine the actual tree cover density realized beneath this constraint.

As documented by the regression tree, the effects of several water balance factors were contingent on the values of other factors. For example, the effect of the seasonality of precipitation (MPPTCV) differed systematically on the two main branches of the tree. Cutoffs selected by the RTA algorithm were always between 97 and 98, near the bottom of the range of the variable. Along the wetter main branch, cells with higher tree cover density consistently segregated above the cutoff, while along the drier branch, the cells above the cutoff fell into the leaf with the lowest mean tree cover density (0.10). This particular contingency is intuitive. Under conditions of low annual precipitation, high seasonal variability represents an increased chance of seasonal drought, which can be particularly detrimental to seedling establishment. Higher annual precipitation buffers the threat of drought, and increased seasonal variability represents the opportunity of an unusually wet season, not just the crisis of an unusually dry one.

Under these sorts of contingencies, different combinations, some highly divergent, lead to very similar mean tree cover densities. For example, leaves that averaged 0.25 and 0.26 tree cover were separated by seven nodes in the regression tree, and they lay on opposite sides of nearly identical cutoffs on both MPPTCV and POR. By comparison, the same number of nodes separates the leaf with the highest mean tree cover density (0.53) from that
with the lowest. Woodlands that are superficially similar structurally may thus emerge from very different combinations of water balance conditions. The strength of the RTA is that it renders some of these contingencies understandable.

Even with the largest statistically valid regression tree (104 leaves), half of the variation in tree cover density remains unexplained by water balance variation. The residual analysis suggests that the explanatory value of topographically mediated water balance gradients has been exhausted. The positive trend between residual tree cover density and neighbourhood tree cover density (Fig. 7) indicates some local autocorrelation in tree cover density that is not accounted for by spatial variation in our water balance parameters. This pattern may reflect the influence of some unmeasured environmental parameter, e.g. soil nutrients. Alternatively, it suggests that local ecological feedbacks and site history play a very important role in determining vegetation patterns. Sparse neighbourhoods might represent previously inhospitable or recently disturbed areas. Under such a scenario, tree cover density is lower than expected from site water balance due to time lags associated with dispersal and establishment (Chambers et al., 1999). Dense neighbourhoods, on the other hand, likely provide more abundant seed sources and may facilitate seedling establishment and persistence through the effect of tree cover on understory light (Breshears et al., 1997; Martens et al., 2000), nutrient availability (Klopatek, 1987; Padien & Lajtha, 1992), and soil moisture conditions (Rietkerk et al., 1997; Breshears et al., 1998; Breshears & Barnes, 1999; Klausmeier, 1999). Such feedbacks between vegetation and local water balance were beneath the resolution of this analysis, but our results suggest that they may be very important for explaining residual variation in tree cover density.

The importance of historical processes in determining woodland distribution is apparent in the dynamic responses of pífon-juniper woodlands to past changes in both climate and human land use (Swetnam & Betancourt, 1990; Miller & Wigand, 1994; Allen & Breshears, 1998; Swetnam & Betancourt, 1998). Tree ring records document periodic wet and dry episodes that persist on the order of decades at our site (Milne et al., 2003). During wet periods, e.g. the 1980s, woodland trees may establish in previously inhospitable areas (Swetnam & Betancourt, 1998). During extended droughts, e.g. the 1950s, the same areas are more stressful for trees, and mortality tends to increase, resulting in local contractions in the woodland distribution. A large-scale drought mortality episode at the Los Pinos site beginning in 2002 (Kerkhoff and Milne, personal observation) provides an opportunity to improve understanding of the relationship between water balance variation, disturbance, demography, and woodland tree distributions.

Primary human concerns about global change in semiarid regions are the degradation of rangeland, increased soil erosion and loss, and desertification. The magnitudes and rates of these responses to climate change depend, at least in part, on the distribution and dynamics of woody vegetation. Our study demonstrates that contingencies among water balance conditions generate complex patterns of tree canopy distribution that are only partially predictable; and ecological processes and site history are important determinants of woodland structure within the constraints of water balance. Thus, while the forces precipitating climate and land use change are global, the responses of semiarid landscapes emerge from locally interacting ecological and hydrological processes. Water balance gradients provide clues to the broad-scale, regional ramifications of climate and land use change, but the detailed ecological knowledge necessary for predicting the dynamics of semiarid landscapes presents a serious challenge for global change research.

ACKNOWLEDGEMENTS

This research was supported by the NSF (grants DEB-0073139 to AJK and DEB-991023 to BTM) and NASA (grant NGR-30227 to AJK), and the Sevilleta LTER program. Valuable feedback on earlier drafts of this work was provided by M. Rietkerk, A.P. Allen, C. Restrepo, J.H. Brown, D.D. Breshears, and an anonymous reviewer. Sevilleta LTER publication number 293.

REFERENCES


**BIOGRAPHICAL SKETCHES**

**Drew Kerkhoff** is currently a postdoctoral research associate in Ecology and Evolutionary Biology at the University of Arizona. His research interests include plant allometry, the macroecology of plant functional traits, the evolution of reproductive variability and synchrony in plant populations, and ecohydrology.

**Scott Martens** was Visiting Research Associate Professor of Biology at University of New Mexico with interests in ecosystem and landscape ecology. He is currently at Sierra Science in Three Rivers, CA, USA.

**Greg Shore** is System Administrator at the US National Science Foundation Long-term Ecological Research (LTER) Network Office and former GIS/RS Manager of the Sevilleta LTER Program. He specializes in the management and analysis of geospatial and ecological databases.

**Bruce Milne** is Professor of Biology at the University of New Mexico, where he focuses on landscape ecology, biocomplexity, the dynamical origins of scaling behaviour, and sustainability.