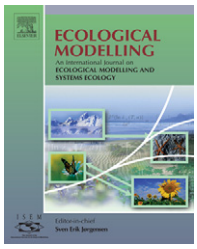
available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel



Using local spatial autocorrelation to compare outputs from a forest growth model

Michael A. Wulder[a,*], *Joanne C. White*[a], *Nicholas C. Coops*[b], *Trisalyn Nelson*[c], *Barry Boots*[d]

[a] *Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, Victoria, British Columbia, Canada V8Z 1M5*
[b] *University of British Columbia, Department of Forest Resource Management, Vancouver, British Columbia, Canada V6T 1Z4*
[c] *University of Victoria, Department of Geography, Victoria, British Columbia, Canada V8W 3P5*
[d] *Wilfrid Laurier University, Department of Geography and Environmental Studies, Waterloo, Ontario, Canada N2L 3C5*

ARTICLE INFO

Article history:
Received 26 July 2006
Received in revised form
19 April 2007
Accepted 29 June 2007
Published on line 17 August 2007

Keywords:
Physiological model
3PG
LAI
Stand volume
Local spatial autocorrelation
Getis statistic

ABSTRACT

Comparing model outputs is a critical precursor to successfully applying models to environmental issues. In this paper, we applied a calibrated physiological model (3PG) and predicted two fundamental forest growth attributes (leaf area index (LAI) and stand volume). As part of this simulation, we systematically changed two key model input parameters (soil water holding capacity and soil fertility rating) and compared the model outputs utilising a method that accounts for local spatial autocorrelation. The use of the Getis statistic (G_i^*) provides insights on the spatial ramifications of an aspatial change to model inputs. Specifically, the location of significant G_i^* values identified areas where the differences in LAI and stand volume occur and are spatially clustered. When soil water is doubled and soil fertility is unchanged, both LAI and stand volume increase; conversely, when soil water is doubled and soil fertility is halved, both LAI and stand volume decrease. The increase and decrease in these model outputs occurred differentially across the study area, although there is a similar pattern to the location of the significant G_i^* values ($p = 0.10$) in both LAI and stand volume outputs, for each model scenario. Analyzing the local spatial autocorrelation of the differences between model outputs identified those areas that have systematic sensitivity to specific model inputs. This information may then be used to aid in the interpretation of model outputs, or to direct the collection of additional data to refine model predictions.

1. Introduction

The past 10 years have seen significant new developments in the use of models investigating carbon dynamics in terrestrial ecosystems. In addition, recent advances in software and hardware technology have dramatically increased opportunities to undertake simulations and compare model assumptions and behaviours in a consistent and standardized way. The comparison of model outputs is therefore becoming an important, if not a critical step, in developing, testing, and ultimately applying models to environmental issues.

In basic situations where the same model framework is applied, yet input variables are varied (resulting in a range of output predictions), it is possible to simply subtract or ratio the two model predictions, and observe and analyse the differences in the predictions. In this paper, we describe

* *Corresponding author at*: 506 West Burnside Road, Victoria, British Columbia, Canada V8Z 1M5. Tel.: +1 250 363 6090; fax: +1 250 363 0775.
E-mail address: mwulder@nrcan.gc.ca (M.A. Wulder).
0304-3800/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.ecolmodel.2007.06.033

and apply a method of model comparison that allows an investigation of the differences in predictions, as well as the spatial patterns associated with the differences. The method employed is a measure of local spatial autocorrelation, which indicates whether the differences between the control and the model scenarios were randomly located over the study area, or followed some spatial pattern, thereby indicating some underlying physical or ecological process.

To develop and demonstrate the approach of using a measure of local spatial autocorrelation for comparison of model outputs, we examined predictions from a process-based model that simulates the growth of forest stands in terms of the underlying physiological processes. Process-based models are typically driven by climatic data and constrained by soil properties that affect the storage and availability of water and nutrients (see review by Makela et al., 2000). These models assume that primary production can be described in terms of radiation interception, photosynthesis, and carbon allocation (Landsberg and Gower, 1997). A key advantage of these types of light-use interception models is that by modeling the underlying physiological processes, the model can be applied at many locations over a landscape. In addition, by utilizing actual meteorological and environmental conditions, changes in forest structure resulting from climate change, management, or other effects, may also be modeled. Landsberg and Waring (1997) developed a deterministic forest growth model, 3PG (Physiological Principles for Predicting Growth) based on a number of established biophysical relationships and constants. 3PG differs from most process models in that it predicts stand properties measured by foresters (tree density, basal area, mean diameters, standing volume, and mean annual increment), as well as those of interest to ecologists (carbon allocation and water balance).

In this paper, we applied a calibrated version of the 3PG physiological model and predicted two fundamental forest growth attributes (leaf area index (LAI) and stand volume) for ponderosa pine (*Pinus ponderosa* Doug. ex Loud). As part of this simulation, we systematically changed two key model input parameters (soil water holding capacity and soil fertility rating) and compared the model outputs utilising a method of model comparison that accounts for local spatial autocorrelation. Our objective was to determine how aspatial changes in model inputs manifest spatially in the model outputs, or more specifically, were there spatial locations of systematic model sensitivity to aspatial changes in model inputs? The advantages of using this type of approach are then discussed.

2. Study area, data, and methods

The study area for this investigation is located on the western coast of the United States and spans the states of Washington, Oregon, and Northern California (Fig. 1). Within this region, ponderosa pine represents the major forest type, occurring in nearly pure stands in a 15–30 km wide band along the eastern flanks of the Cascade Mountains where annual precipitation is generally between 300 and 800 mm. Historically, ground fires at 8–20 year intervals maintained the ponderosa pine forest type free of other potential competing conifers. However, on more moist sites, ponderosa pine occurs in a mixture with Douglas–fir (*Pseudotsuga menziesii* (Mirb.) Franco), grand fir (*Abies grandis* (Dougl.) Lindl.), and other conifers. On more arid sites, juniper (*Juniperus occidentalis* Hook.) and/or sagebrush (*Purshia tridentata* (Pursh) D.C.) replace ponderosa pine as drought becomes more severe and fires more frequent (Franklin and Dyrness, 1973). Today, ponderosa pine occupies an extensive range, yet it maintains an ecologically precarious position, constrained to the east by more arid conditions that favour juniper woodlands, to the west by mountains with more moderate precipitation that favour a mixture of other conifers, and by elevation, where heavy snow loads can damage ponderosa pine branches (Waring et al., 1975). Ponderosa pine, along with Douglas–fir, has served as a basis for evaluating growth potential across a wide range of forests in the western United States (e.g., Waring et al., 2002). The ecological distributions and growth of ponderosa pine is therefore critically important (Franklin and Dyrness, 1973 and see special issue on "The ponderosa pine ecosystem and environmental stress: past, present, and future" published in *Tree Physiol*. vol. 21, 2001) and monitoring the current and future distribution of species in this region is a high priority for forest resource mangers concerned with future forest distribution and productivity (Coops et al., 2005).

2.1. *The 3PG model*

All ecosystem models are simplified versions of reality with the choice of which process based model to utilize being dependent upon their input and output parameters, minimum spatial and temporal units of operation, maximum spatial extent and time period of application (Nightingale et al., 2004). In addition, the scale at which the model operates (leaf–tree, plot–stand, regional and ecosystem levels) is also critical, with model complexity generally decreasing as the time-step and spatial extent of model operation increases (Nightingale et al., 2004; Coops et al., 2005). Given the large spatial extent of the study area, and the subsequent requirement for coarse spatial resolution input data we believe a monthly time step, stand-level, process based model is an appropriate choice for our analysis. Within this specification a number of process based models exist (Nightingale et al., 2004) including HYBRID (Friend et al., 1997), FOREST-BGC (Running and Coughlan, 1988), BIOME-BGC (Running and Hunt, 1993) amongst others). There are however two features that together distinguish 3PG from all other process-based models (some share one feature) include (Landsberg et al., 2003):

- The simplifying assumption that respiration is a fixed faction of gross photosynthesis (Waring et al., 1998; Gifford, 2003). This simplification removes the difficulty in predicting belowground growth, protein turnover rates, and separating carbon dioxide generated by microbial activity.
- Detailed forest inventory variables are readily predicted by the 3PG model (such as standing volume) and indirectly support its simplifying assumptions. Confidence in the 3PG structure and function is gained as it accurately predicts measured change in LAI, litterfall, stocking density, basal area, and mean tree diameters, in addition to annual growth in managed and unmanaged stands (Landsberg et al., 2003).

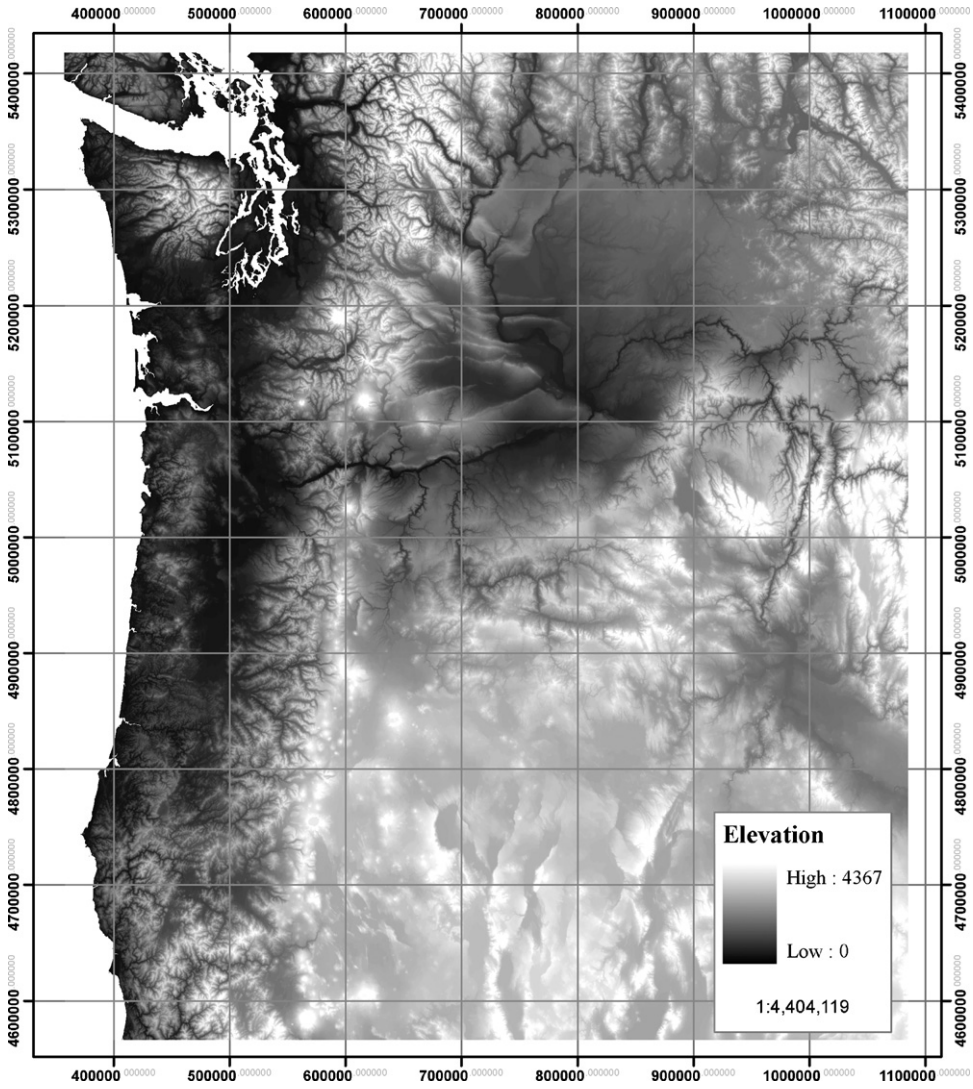


Fig. 1 – Extent of the study area in the Pacific Northwest with a Digital Elevation Model (DEM).

We therefore have confidence that the 3PG model occupies the middle ground between conventional mensuration-based growth and yield models, and process-based carbon balance models (Landsberg and Waring, 1997). Detailed information on the 3PG is available (Landsberg and Waring, 1997; Landsberg et al., 2003); however, for completeness a short overview is provided here. 3PG is a monthly time step model which requires average daily short-wave incoming radiation for each month, daily mean vapor pressure deficits (*D*), temperature extremes, total monthly precipitation, and estimates of soil water storage capacity (mm) and soil fertility. Absorbed photosynthetically active radiation (APAR) is estimated from global solar radiation and LAI; the utilized portion, APARu, is calculated by reducing APAR by an amount determined by a series of modifiers that take values between 0 (system 'shutdown') and 1 (no constraint) to limit gas exchange via canopy stomatal conductance (Landsberg and Waring, 1997). The modifiers include: (a) high averaged daytime *D*, (b) the frequency of subfreezing conditions, and (c) soil drought. Limitations on APARu are imposed each month by the modifier with the lowest value. Drought limitations are imposed as a function of soil texture when the total monthly precipitation and soil water supply is significantly less than transpiration estimated with the Penman–Monteith equation. Gross primary production (PG) is calculated by multiplying APARu by a canopy quantum efficiency coefficient (α), with a maximum value set by the soil fertility ranking and reduced monthly when mean temperatures are suboptimal for photosynthesis and growth. A major simplification in the 3-PG model is that it does not require detailed calculation of respiration from knowledge of root turnover, but rather assumes that autotrophic respiration (Ra) and total net primary production (PN) in temperate forests are approximately fixed fractions (0.53 and 0.47, S.E. $\pm$ 0.04) of PG (Landsberg and Waring, 1997, Waring et al., 1998; Law et al., 2000a). The model partitions PN into root and aboveground biomass. The fraction of total PN allocated to root growth increases from 0.25 to 0.6 as the ratio APARu/APAR decreases from 1.0 to 0.2. Under more favourable climatic conditions, the fraction of photosynthate allocated to roots increases with infertility of the soil (Landsberg and Waring, 1997).

The role of nutrition is an important variable within 3PG however our capacity to link soil nutrient status within quan-

titative models of plant growth is limited (Landsberg et al., 2003). This is partly due to a lack of good quality spatial information about soil physical and chemical properties, and second, the characterisation of simple relationships between standard measures of soil fertility and tree growth is difficult and depends upon geochemical cycling (see Waring and Schlesinger, 1985; Landsberg and Gower, 1997; and Waring and Running, 1998), particularly in relation to nitrogen. As a result in 3PG, although chemical analyses may provide a guide to fertility ranking, a degree of expert knowledge is required. The fertility ranking therefore (set between 0 and 1) can be used as a tuneable parameter in the model (Landsberg et al., 2003) and scaled according to the information on soil nutrient status available at a site or across a region.

3-PG (and variations of this model) has been used extensively to model the productivity of a wide range of forest types across regions of North America including: ponderosa pine (Law et al., 2000a, 2000b); lodgepole pine (*Pinus contorta* Dougl. ex Loud. var. *latifolia* Engelm.) (Hall et al., 2006); loblolly pine (*Pinus taeda*) (Landsberg et al., 2000); Douglas–fir (Coops et al., 2005); and jack pine (*Pinus banksiana*) (Peng et al., 2002). Model productivity estimates have exhibited a high degree of accuracy when compared to site index data and in addition, the model water balance has been shown to accurately demonstrate the general trends in regional soil water depletion using limited monthly climate datasets (Coops et al., 2001a; Coops and Waring, 2001a; Coops et al., 2001b; Coops and Waring, 2001b).

2.2. *Input data*

While many agencies are routinely producing average climate surfaces over large spatial areas using state of the art mathematical modeling approaches such as DAYMET in the United States (Thornton et al., 1997; Thornton and Running, 1999), and spline fitting in Canada (McKenney et al., 2001), the inclusion and integration of soil information is an ongoing issue.

Information on soil fertility and soil water holding capacity is critically important to models of plant production; however, consistent data on soil properties at a fine spatial resolution is typically unavailable for many regional studies. Soils maps delineated at scales of 1 km^2 or coarser generally mask significant spatial variation in physical and chemical properties. Even with more precise mapping the fertility of forest soils would be difficult to judge as the same soil type may be commercially fertilized, may support nitrogen-fixing vegetation, or receive significant atmospheric depositions of nutrients (and pollutants) (Nightingale et al., in press). For example, soil water holding capacity depends on the soil type and the depth to which the soil is exploited by tree root systems, with values typically ranging from 40 to 600 mm. Likewise, soil fertility indices represent changes in soil chemical composition that either stimulate or inhibit growth. Soil fertility is a particularly problematic variable, as there is only limited empirical data available regarding soil physical and chemical properties, and soil fertility varies extensively spatially (Ryan et al., 2000). Due to these difficulties, both soil water holding capacity and soil fertility are often set as constant mean values, over the entire study area (Running, 1994; Coops et al., 2005). The average values for these variables are usually derived from a limited set of field plots or based on past research in the study area. In previous studies considering ponderosa pine forests, soil water holding capacity has been set to a range of fixed values around 150 mm, which was the control value used for this study. Similarly, the soil fertility index was set to match values reported for previous studies (Coops et al., 2005) and hence, a constant value of 0.4 was used for the control.

DAYMET fine scale spatial climate coverages were used to drive the 3PG model. These layers provide 1 km × 1 km spatial estimates of averaged mean monthly data for precipitation, minimum and maximum temperature, frost occurrence, and short wave radiation for the conterminous United States (Thornton et al., 1997; Thornton and Running, 1999). Law et al. (2000a) undertook an evaluation of the 3PG model to examine its ability to estimate gross photosynthesis and net primary production of ponderosa pine. This work defined a number of key parameters required for the 3PG model and included a sensitivity analysis of selected variables. These parameter values were used for regional predictions undertaken by Coops et al. (2005) and in this study. Like most regions, input spatial coverages for soil water holding capacity and soil fertility, at an appropriate scale, were unavailable. As a result, these two critical inputs became the focus of the model comparison, and a set of simulations were undertaken to systematically vary these input coverages and quantify variation in the output predictions.

2.3. *Measuring spatial autocorrelation*

Spatial autocorrelation exists where there is a systematic spatial variation in values across a given area (Cliff and Ord, 1981). The emphasis is on the patterns in the values recorded at specific locations and not on the patterns of the locations themselves (Upton and Fingleton, 1985), with both the location and magnitude of the attribute considered simultaneously in the measurement of spatial autocorrelation (Goodchild, 1986). Measures of spatial autocorrelation are either global or local in nature (Boots, 2002). Global measures of spatial autocorrelation yield a single measure that summarizes the entire study area; however, such measures are rarely representative of the variation in spatial autocorrelation found over large areas (Wulder and Boots, 1998; Boots, 2002). This is because global measures assume spatial stationarity, which implies that the statistical properties of the spatial process in question are independent of absolute location (Bailey and Gatrell, 1995). In most cases, this assumption is violated, resulting in misleading measures of global spatial autocorrelation (Fotheringham and Brunsdon, 1999).

Local measures of spatial autocorrelation focus on identifying local variations within patterns of spatial dependence and are therefore useful for revealing spatial relationships which might otherwise be undetected (Anselin, 1995). Local measures may be influenced by the presence of global spatial autocorrelation, and since there is incomplete knowledge regarding the statistical distribution of measures of local spatial autocorrelation, caution must be exercised when testing for the statistical significance of these local measures in the presence of global spatial autocorrelation (Castro and Singer, 2006; Ord and Getis, 2001).

In the past decade, several techniques have been developed to measure local spatial autocorrelation. One such method is the Getis statistic. Although originally developed for identifying hot spots from point data (Getis and Ord, 1992; Ord and Getis, 1995), one form of the Getis statistic, G_i^*, has been modified for use with raster data (Derksen et al., 1998; Wulder and Boots, 1998; Wulder and Boots, 2001). Wulder and Boots (1998) provide a methodology for deriving G_i^* from raster data. Like all local measures of spatial autocorrelation, G_i^* may be influenced by the presence of global spatial autocorrelation (Boots, 2002). Global spatial autocorrelation should therefore be assessed prior to computing local measures; a measure such as Moran's I, which is a weighted correlation coefficient used to detect departures from spatial randomness (Cliff and Ord, 1981), may be used (Rogerson, 2002).

The Getis statistic yields a standardized value that indicates both the degree of positive spatial autocorrelation in the attribute values centered on a given pixel and the magnitude of these values in relation to those of the entire raster (Wulder and Boots, 1998). As per Wulder and Boots (1998), the Getis value was calculated using the following equation:

$$G_j^* = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}}{s[W_i^*(n - W_i^*)/(n-1)]^{1/2}} \quad (1)$$

where $w_{ij}(d)x_j$ is the sum of the variates within distance d of observation i (including i), W_i^* is the count of the pixels within distance d of pixel i and n is the total number of observations.

To calculate the G_i^*, five incrementally sized square cell windows (3×3, 5×5, 7×7, 9×9 and 11×11) are passed over the selected raster layer. The largest G_i^* value that occurs through the processing of the sequence of window sizes is utilized to obtain a maximized measure of local association. Then, for each cell, the maximum value is recorded. Large G_i^* values indicate spatial clusters of large values (relative to the mean); small G_i^* values indicate spatial clusters of small values (relative to the mean). In the analysis conducted for this paper, the Getis statistics were generated for normalized difference values between a control output and a model scenario output (Table 1); therefore, large G_i^* values indicate spatial clusters of large differences between the control output and the scenario output, and small G_i^* values indicate spatial clusters of minimal differences between the scenario output and the control output.

Since the G_i^* values are Z-score standardized, it is possible to assess the significance of spatial autocorrelation on a cell-by-cell basis, thereby identifying locations with more extreme levels of spatial dependence. The significance value is thus used as a threshold and may be altered according to the level of spatial dependence the analyst wishes to represent (Wulder and Boots, 2001). In situations such as that presented in this analysis, where G_i^* values are being generated for multiple model outputs over the same geographic area, cumulative hot spots may be identified. These hot spots represent those areas that are repeatedly identified as having significant G_i^* values, in multiple model outputs. In the analysis conducted for this paper, hot spots identified areas where statistically significant ($p = 0.10$) spatial clusters of large difference values consistently occurred.

2.4. *Approach*

In order to apply the model comparison the following procedure was used. First, a control set of simulations was created based on initial model conditions and the parameter values of Coops et al. (2005) and Law et al. (2001). Over the entire study area, available soil water holding capacity was set to 150 mm and soil fertility was set to 0.4. Two model outputs, LAI and stand volume were selected for analysis. LAI, defined as the amount of foliage per unit surface area (Fournier et al., 2003), is considered a fundamental parameter for ecological models and is coupled with photosynthesis, transpiration rates, and light and water interception (Breuer et al., 2003). Stand volume (m^3/ha) was selected as it provides a reliable indication of how the stand will develop over time, given various scenarios (Landsberg et al., 2003).

Two alternate model runs were then undertaken where soil water holding capacity and soil fertility were varied, with all other 3PG model inputs held constant. For the first model run, soil water was doubled to 300 mm and soil fertility remained unchanged (0.4); for the second model run soil water was doubled (300 mm) and soil fertility was halved (0.2). The underlying hypothesis is that when soil water is doubled and soil fertility unchanged, a commensurate increase in forest volume and LAI is expected due to an increase in water availability for plant growth. This is most likely in regions where soil water is currently a limiting factor to growth. For the second set of simulations, when soil water is doubled and soil fertility is halved, a similar increase in growth (to that of

Table 1 – Listing of comparisons made between the outputs for LAI and stand volume generated from the control with outputs generated from scenarios 1 and 2

Layer comparisons	
LAI (control) Soil water = 150 mm and soil fertility = 0.4	LAI (scenario 1) Soil water = 300 mm and soil fertility = 0.4
Stand volume (control) Soil water = 150 mm and soil fertility = 0.4	Stand volume (scenario 1) Soil water = 300 mm and soil fertility = 0.4
LAI (control) Soil water = 150 mm and soil fertility = 0.4	LAI (scenario 2) Soil water = 300 mm and soil fertility = 0.2
Stand volume (control) Soil water = 150 mm and soil fertility = 0.4	Stand volume (scenario 2) Soil water = 300 mm and soil fertility = 0.2

Table 2 – Summary statistics for modeled outputs for LAI

Model Outputs	Minimum	Maximum	Mean	Standard deviation	Moran's *I*
Control	0.000	5.781	2.424	0.951	
Model Scenario 1:					
Soil water doubled and soil fertility unchanged	0.000	6.360	2.721	1.102	
Normalized difference between control and doubled soil water and unchanged soil fertility	1.000	1.460	1.126	0.093	0.936 Z-score = 507.593
LAI (control) for locations with significant maximum G_i^* values ($p = 0.10$)	0.001	2.365	1.801	0.203	
LAI (scenario output) for locations with significant maximum G_i^* values ($p = 0.10$)	1.292	3.121	2.418	0.263	
Model Scenario 2:					
Soil water doubled and soil fertility reduced by half	0.000	4.079	1.671	0.747	
Normalized difference between control and doubled soil water and halved soil fertility	1.053	11.133	1.688	1.196	0.865 Z-score = 464.836
LAI (control) for locations with significant maximum G_i^* values ($p = 0.10$)	0.001	2.215	0.348	0.351	
LAI (scenario output) for locations with significant maximum G_i^* values ($p = 0.10$)	0.001	1.821	0.161	0.209	

the first set of simulations) is expected only where soil fertility is not considered a significant limiting factor to vegetative growth (Law et al., 2000a). All of the model simulations were run for a period of 30 years to generally coincide with canopy closure, but prior to stand self thinning.

LAI and stand volume are measured on different scales: control values for LAI in the study area range from 0 to 5.78, while stand volume ranges from 12 to 792 m^3/ha. To determine if the change in model inputs were having similar impacts on these two indicators, a normalized measure was required to facilitate interpretation. Furthermore, neither the absolute magnitude nor the direction of the difference between the control and the model output were of interest. A relative measure of difference was therefore calculated by dividing the control and model output, with the larger value of the two as the numerator. With this approach, relative change will always be greater than or equal to 1, and a value of 1 is indicative of perfect agreement between the control and the model output. Examining the difference values directly may identify single pixels of differences between the model outputs and the control that could be spurious or caused by noise. By examining significant G_i^* values, locations of systematic differences that are spatially clustered may be identified.

3. Results

Descriptive statistics for the base output and the two sets of simulations for LAI are shown in Table 2. Normalized difference values were generated by ratioing the two modified LAI and stand volume coverages with the LAI and stand volume estimates from the control. The table summarizes the minimum, maximum, mean, and standard deviation of the spatial coverages generated. In addition, Moran's *I* was calculated to assess if global spatial autocorrelation was present in the normalized difference outputs. Table 3 presents the

Table 3 – Summary statistics for modeled outputs for stand volume

Model outputs	Minimum	Maximum	Mean	Standard deviation	Moran's *I*
Control	12.924	791.709	300.833	126.428	
Model Scenario 1:					
Soil water doubled and soil fertility unchanged	12.924	890.758	340.154	149.976	
Normalized difference between control and doubled soil water and unchanged soil fertility	1.000	1.448	1.131	0.093	0.944 Z-score = 507.397
Stand volume (control) for locations with significant maximum G_i^* values ($p = 0.10$)	124.571	289.409	213.304	25.238	
Stand volume (scenario output) for locations with significant maximum G_i^* values ($p = 0.10$)	156.633	358.753	286.822	32.820	
Model Scenario 2:					
Soil water doubled and soil fertility reduced by half	12.854	519.303	209.715	89.803	
Normalized difference between control and doubled soil water and halved soil fertility	1.005	3.045	1.478	0.304	0.924 Z-score = 469.673
Stand volume (control) for locations with significant maximum G_i^* values ($p = 0.10$)	23.863	409.193	178.191	66.326	
Stand volume (scenario output) for locations with significant maximum G_i^* values ($p = 0.10$)	17.953	250.925	66.481	29.056	

same set of analysis for the standing volume estimates. Both Tables 2 and 3 indicate that significant global spatial autocorrelation was present in each of the four normalized difference layers (as described in Table 1).

3.1. *Increased soil water availability with no change to soil fertility*

Fig. 2 illustrates the top 5% of the normalized difference values (Fig. 2a) versus the significant ($p = 0.10$) G_i^* values (Fig. 2b) for a portion of the study area. The figure demonstrates the utility of analyzing the local spatial autocorrelation in model outputs. The top 5% of the normalized difference values show some general spatial coherence; however in general, the spatial distribution is heterogeneous with isolated pockets and single pixels with large values. The heterogeneity of these locations makes it difficult to assess whether these pockets and fragments are, in fact, regions of significant variation between the control output and the model scenario output, or are simply noise in the model prediction. By contrast, Fig. 2b shows the location of the significant G_i^* values, indicating where large difference values are spatially clustered. To summarize, the top 5% of difference values (Fig. 2a) shows us where the largest difference values are between the control and the model scenario, whereas the significant G_i^* values indicates where large difference values are spatially clustered on the landscape.

The impact on LAI of doubling the available soil water holding capacity and not changing the soil fertility ranking did not manifest in a uniform increase in LAI across the study area; rather, this scenario resulted in a highly variable spatial response in LAI values (Fig. 3, Table 2). For the area as a whole, doubling of available soil water increased the mean LAI from 2.424 in the control to 2.721 in the scenario output, and the maximum LAI from 5.781 to 6.360. The areas with significant G_i^* values are shown in red in Fig. 3. These areas had a mean LAI of 1.801 in the control, and a mean LAI of 2.418 in this scenario output (where soil water was doubled and soil fertility remained unchanged). The average difference between the control and the output across the entire study area was

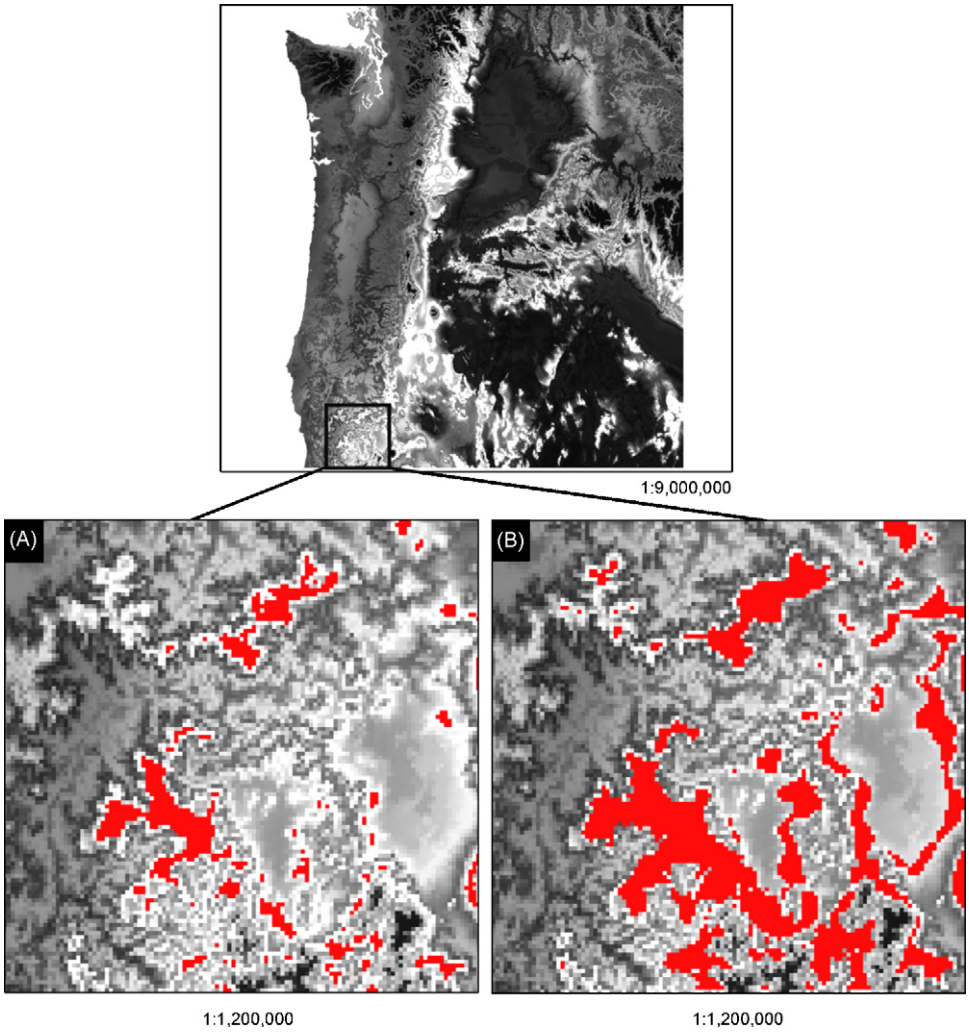


Fig. 2 – Comparison of top 5% of difference values (A) and significant ($p = 0.10$) Getis (G_i^*) values (B) for LAI. Output is generated from doubling soil water and not changing the control value for the soil fertility index. Background values (greyscale) are normalized difference values generated by ratioing the control output for LAI with the model scenario output for LAI.

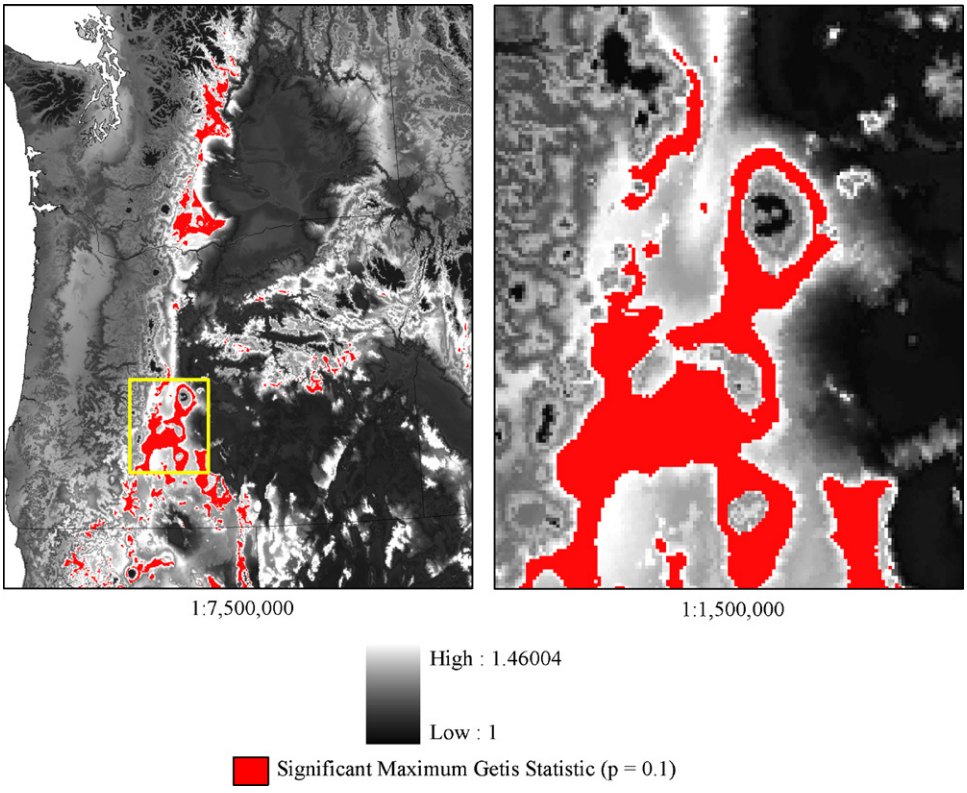


Fig. 3 – The impact on modeled LAI of doubling soil water and not changing soil fertility. Background values (greyscale) are normalized difference values generated by ratioing the control output for LAI with the model scenario output for LAI; the larger value of the two is always used as the numerator.

1.126 units, while the average difference in the areas where G_i^* values were significant was 1.343 units.

Similar results were found for standing volume. The mean stand volume for the region is 300 m^3/ha, with a maximum stand volume of 792 m^3/ha. Similar to the impact on LAI, the impact on stand volume of increasing the available soil water capacity and maintaining the soil fertility is also spatially diverse over the landscape (Fig. 4). For the region as whole, the mean stand volume increased to 340 m^3/ha, and the maximum stand volume to 891 m^3/ha. The areas with significant G_i^* values are shown in red in Fig. 4. In these areas, the mean stand volume was 213 m^3/ha in the control, and 286 m^3/ha in the scenario output. The mean difference in stand volume in these areas of significant G_i^* values was 1.321 units, while the mean difference across the study area as a whole was 1.131 units. Therefore, the relative impacts on model outputs of LAI and stand volume, from doubling soil water and not changing soil fertility, are similar.

3.2. *Increased soil water availability with reduced soil fertility*

The impact on LAI of increasing the available soil water and simultaneously reducing soil fertility produces varying spatial patterns over the landscape as shown in Fig. 5. The overall result was a reduction in mean LAI from 2.424 for the control, to 1.671 for the scenario output. In those areas with significant G_i^* values, the mean LAI of the control was 0.348, while the mean LAI in the scenario output was 0.161. The mean difference in LAI across the study area was 1.688 units, while the mean difference in areas with significant G_i^* values was 7.911 units.

The impact on stand volume is similar (Fig. 6), with a decrease in the average stand volume from 300 m^3/ha in the control to 210 m^3/ha in the scenario output. In those areas with significant G_i^* values, the mean stand volume in the control was 178 m^3/ha, while the mean stand volume in the scenario output was 66 m^3/ha. The mean difference between the control and the scenario output was 1.478 units, while in areas with significant G_i^* values, the mean difference was 2.719 units. Therefore, unlike the first scenario where the relative impacts of changing soil water alone had impacts of similar magnitude on both modeled LAI and stand volume, in this scenario where soil water was doubled and soil fertility was halved, the impact on LAI was of a greater magnitude than the impact on stand volume.

3.3. *Identification of cumulative hot spots and cold spots of extreme difference values*

Based on the results of the two simulations and two model outputs, it is possible to combine the areas of significant G_i^* values to identify cumulative hot spots from the multiple model outputs. These results are shown in Fig. 7. Hot spots indicate areas where model outputs consistently identify significant G_i^* values. The total area occupied by significant G_i^* values in the LAI

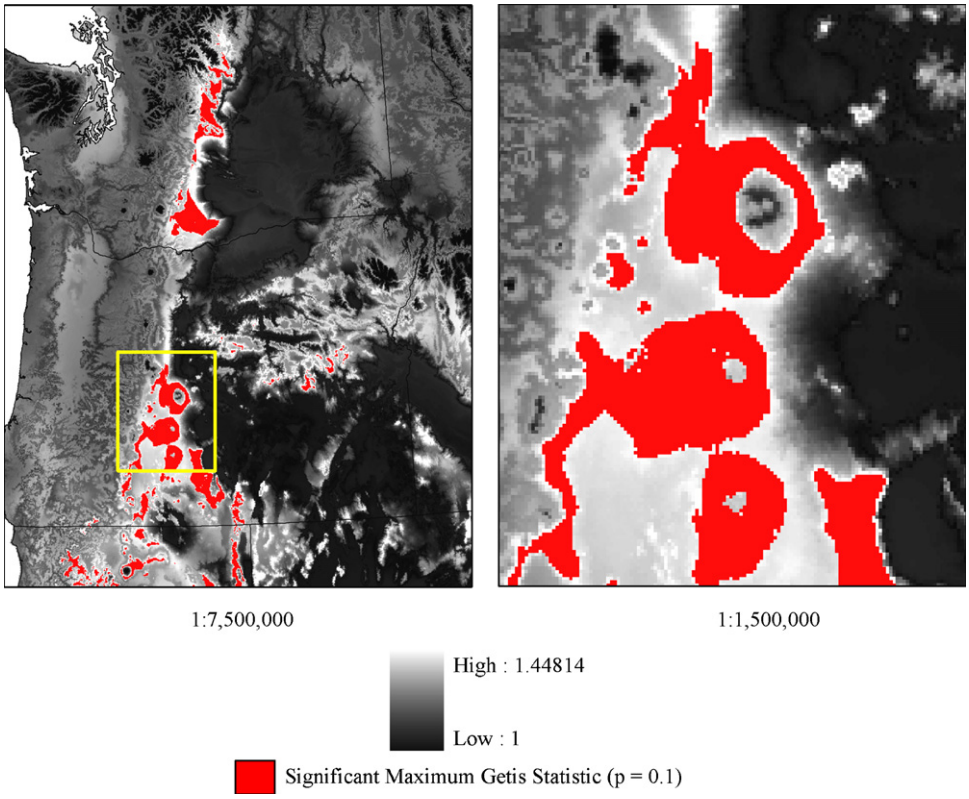


Fig. 4 – The impact on modeled stand volume of doubling soil water and not changing soil fertility. Background values (greyscale) are normalized difference values generated by ratioing the control output for stand volume with the model scenario output for stand volume; the larger value of the two is always used as the numerator.

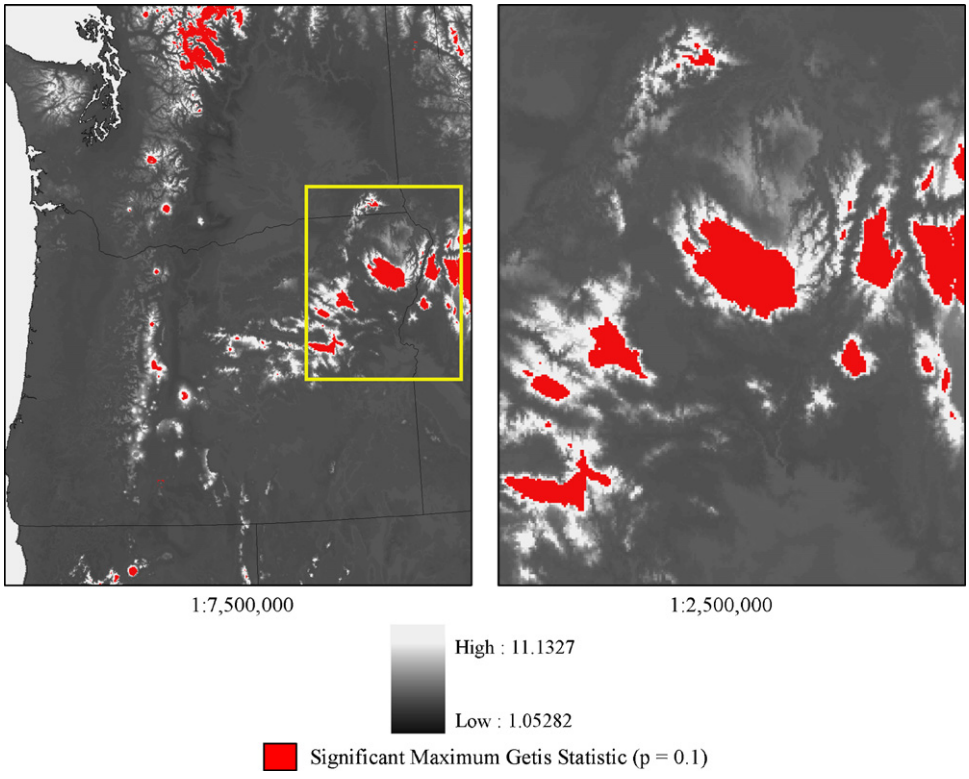


Fig. 5 – The impact on modeled LAI of doubling soil water and halving soil fertility. Background values (greyscale) are normalized difference values generated by ratioing the control output for LAI with the model scenario output for LAI; the larger value of the two is always used as the numerator.

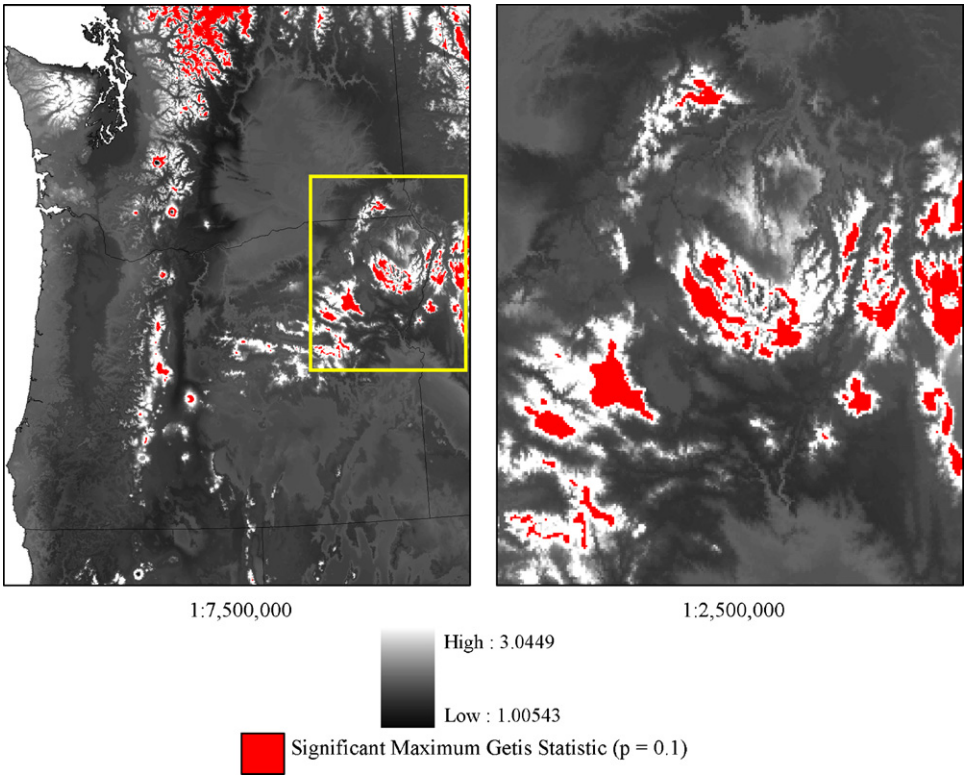


Fig. 6 – The impact on modeled stand volume of doubling soil water and halving soil fertility. Background values (greyscale) are normalized difference values generated by ratioing the control output for stand volume with the model scenario output for stand volume; the larger value of the two is always used as the numerator.

and stand volume predictions (all hotspots) was 3575 ha; 53% of this area was identified in at least two of the normalized difference layers (shown in green in Fig. 7).

4. Discussion

The results indicated that changes in model inputs of available soil water and soil fertility, applied uniformly across the study area, manifest as spatially variable estimates of LAI and stand volume. Examining the differences between the control and the two model scenarios facilitates the identification of areas of maximum change. However, by utilising techniques which also account for the local spatial autocorrelation of these differences (such as the Getis statistic) additional information is generated, such as where those areas of maximum change, relative to the mean value, are spatially clustered. This clustering, in turn, provides an indication of systematic sensitivity of the model, and thereby the modeled species, to underlying changes in soil water and soil fertility inputs.

The model outputs examined in this study represent stand conditions 30 years after stand establishment. The stand volume is cumulative whereas LAI is temporally variable and represents current conditions at 30 years, given a specific set of stand conditions. Generally, the development of a closed canopy may be expected for a 30-year-old ponderosa pine forest (Law et al., 2000a), and LAI is strongly constrained by a shortage of soil water (Landsberg and Waring, 1997). Both soil water holding capacity and soil fertility affect the allocation of growth above and below the ground. In addition, there are complex interactions between soil water holding capacity and soil fertility; for example, an improvement in soil fertility, which results in an increase in LAI, can induce drought (even a small increase in LAI results in a large increase in transpiration and interception).

The results for increased soil water availability with no change to soil fertility show consistent patterns that should be expected from the models (Figs. 3 and 4). Currently, the dry climate inland in Oregon and Idaho is too arid to readily support ponderosa pine. By doubling the available soil water, a consistent pattern of growth in both LAI and stand volume is predicted, as ponderosa pine is able to grow and compete with juniper, sagebrush, and grass. This results in an expansion of the geographical range predicted for ponderosa pine growth by approximately 100 km. Along the western side of its distribution, there is a slight expansion near the Californian and Oregon border where water appears to be a limiting factor in the distribution of ponderosa pine.

The effect of increasing soil water holding capacity, and reducing the soil fertility ranking by half, provides an indication of regions where the LAI and standing volume predictions are highly sensitive to the current fertility setting, in addition to the soil water change. Figs. 5 and 6 present distinct spatial patterns of significant G_i values resulting from the increase in water holding capacity (Figs. 3 and 4). The more dynamic nature of the LAI estimates may explain why the

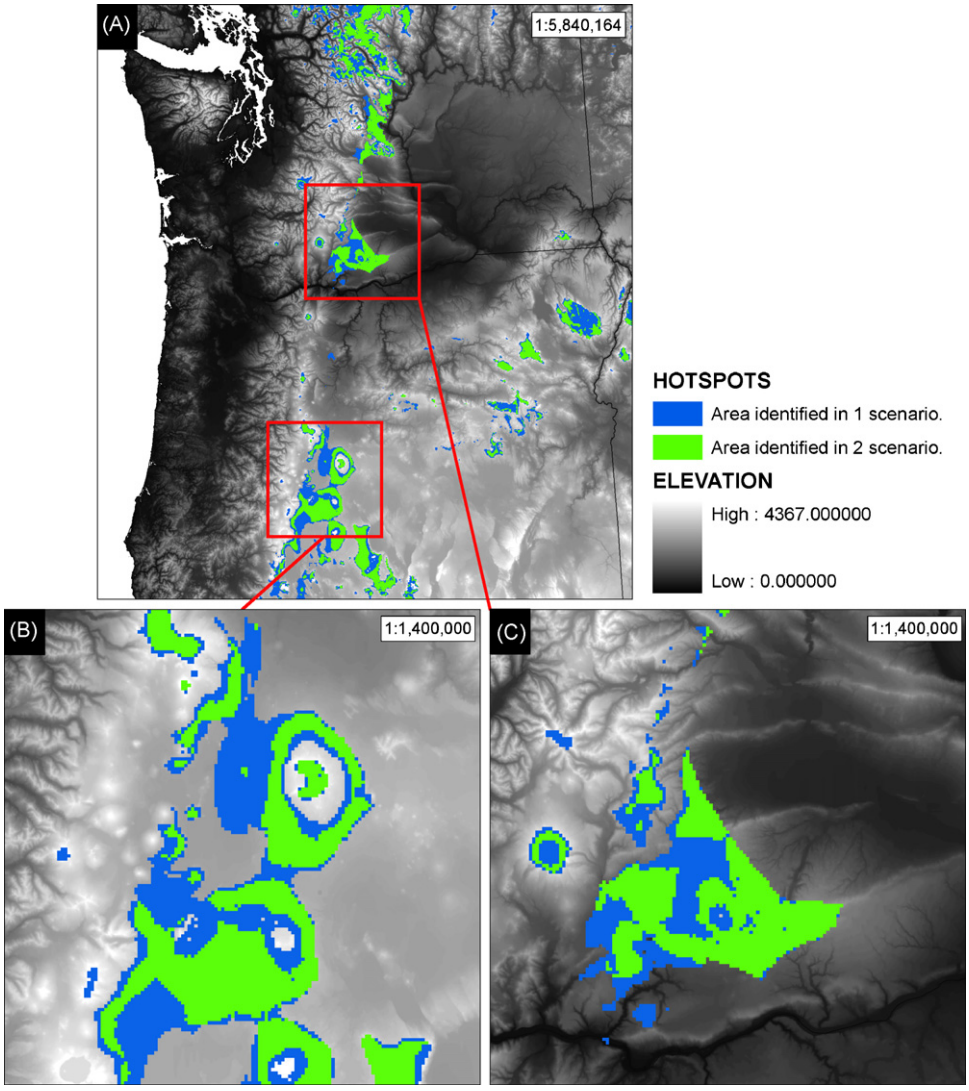


Fig. 7 – Identification of local spatial autocorrelation hot spots for significant ($p = 0.10$) Gi^* values. (A) These significant Gi^* values represent spatial clusters of maximum difference between the control and the model output generated from varying model inputs.

impact on LAI is greater when soil fertility is halved. The hot spots in Fig. 7 provide an indication of where the outputs for LAI and stand volume agree or differ. Fig. 7a and b show areas where model outputs consistently had spatial clustering of large normalized difference values. Again, the eastern side of the Cascades is the focus where both simulations show significant differences caused by increase in soil water alone, or with the simultaneous decrease in soil fertility. These results match those of other research, which has demonstrated that these hotspot locations coincide with areas of very poor soil nitrogen content (in some cases as low as 300 g/m^3 (Swensen et al., 2005) thereby presenting a major limitation to growth for Ponderosa pine. The hotspots verify that in each scenario, the clustered relative difference values tended to be located in the same areas for both LAI and stand volume. In other words, the same areas emerge with statistically significant clustering in the LAI and stand volume outputs.

The outputs from the method presented in this paper could serve several practical roles in a modelling context. For example, once areas sensitive to model inputs are identified, the corresponding input data layers for these areas could be queried to identify potential problems with input data (e.g. outliers, missing data, and scale of data collection, etcetera). Other pre-existing data sources may then be identified and substituted, or alternate data may be acquired through direct measurement. Knowledge of areas sensitive to model inputs may also provide a context within which the model outputs may be characterized. The output from the Getis approach can be used as a confidence rating, associated with the model outputs, indicating areas where model estimates may not be as robust (e.g., in a forest growth modelling context, unusual variations of LAI in a specific area may be attributed to sensitivity in model inputs). Increasingly, models of natural phenomena are incorporating spatial intelligence (e.g. Li et al., 2007; Miller et al., 2007; Swain et al., 2007). The Getis approach brings

spatial intelligence to an existing model, facilitating the comparison of different model scenarios, and allowing the spatial ramifications of aspatial changes to model inputs to be fully realized.

5. Conclusions

The use of the Getis statistic provides insights on the spatial ramifications of an aspatial change to model inputs. If the differences between a control and a model scenario are examined, only the locations of the dissimilarities between the control and the model scenario are learned; by examining the significant G_i^* values, the spatial patterns in the model differences are revealed. In turn, these spatial patterns may be indicative of some underlying physical or ecological process.

In the example presented in this paper, the location of significant G_i^* values indicate where the greatest differences in LAI and stand volume outputs are spatially clustered when soil water is doubled and soil fertility is unchanged, and when soil water is doubled and soil fertility is halved. For the model simulations run in this study, both LAI and stand volume increased (on average) when soil water was doubled and soil fertility was unchanged. Conversely, when soil water was doubled and soil fertility was halved, average values for both LAI and stand volume decreased. The increase and decrease in these model outputs occurred differentially across the study area. Areas that were consistently identified by significant G_i^* values were identified as hotspots of systematic sensitivity to model parameters.

A key advantage of this type of analysis is the capacity of the hot spot information to be used to detect areas of model sensitivity to changes in soil water and soil fertility inputs. This, in turn, can be used to aid in the interpretation of model predictions, or to guide field programs and data collection, with a higher priority being given to those areas identified as being most sensitive to changes in these attributes.

Acknowledgement

The authors are grateful to Dr. Richard Waring for his comments on an earlier draft of this manuscript.

REFERENCES

Anselin, L., 1995. Local indicators of spatial association—LISA. Geogr. Anal. 27, 93–115.

Bailey, T.C., Gatrell, A.C., 1995. Interactive Spatial Data Analysis. Longman, Essex, p. 413.

Boots, B., 2002. Local measures of spatial association. Ecoscience 9, 168–176.

Breuer, L., Eckhardt, K., Frede, H.G., 2003. Plant parameter values for models in temperate climates. Ecol. Model. 169, 237–293.

Castro, M.C., Singer, B.H., 2006. Controlling the false discovery rate: a new application to account for multiple and dependent tests in local statistics of spatial association. Geogr. Anal. 38, 180–208.

Cliff, A.D., Ord, J.K., 1981. Spatial Processes: Models and Applications. Pion, London, p. 266.

Coops, N., Waring, R.H., 2001a. Estimating forest productivity in the eastern Siskiyou Mountains of southwestern Oregon using a satellite driven process model, 3-PGS. Can. J. For. Res. 31, 143–154.

Coops, N.C., Waring, R.H., 2001b. The use of multiscale remote sensing imagery to derive regional estimates of forest growth capacity using 3-PGS. Remote Sens. Environ. 75, 324–334.

Coops, N.C., Waring, R.H., Brown, S.R., Running, S.W., 2001a. Comparisons of prediction and seasonal patterns in water use derived with two forest growth models in Southwestern Oregon. Ecol. Model. 142, 61–81.

Coops, N.C., Waring, R.H., Landsberg, J.J., 2001b. Estimation of potential forest productivity across the Oregon transect using satellite data and monthly weather records. Int. J. Remote Sens. 22, 3797–3812.

Coops, N.C., Waring, R.H., Law, B.E., 2005. Assessing the past and future distribution and productivity of ponderosa pine in the Pacific Northwest using a process model, 3PG. Ecol. Model. 183, 107–124.

Derksen, C., Wulder, M., LeDrew, E., Goodison, B., 1998. Associations between spatially derived autocorrelated patterns of SSM/I-derived prairie snow cover and atmospheric circulation. Hydrol. Process 12, 2307–2316.

Fotheringham, A.S., Brunsdon, C., 1999. Local forms of spatial analysis. Geogr. Anal. 31, 340–358.

Fournier, R., Mailly, D., Walter, J.M.N., Soudani, K., 2003. Indirect measurements of forest canopy structure from *in situ* optical sensors. Pages 77-114. In: Wulder, M.A., Franklin, S.E. (Eds.), Remote Sensing of Forest Environments: Concepts and Case Studies. Kluwer, Boston, MA, p. 519.

Franklin, J.F., Dyrness, C.T., 1973. Natural vegetation of Oregon and Washington. Oregon State University Press, Oregon, USA, p. 417.

Friend, A.D., Stevens, A.K., Knox, R.G., Cannell, M.G.R., 1997. A process-based, biogeochemical, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). Ecol. Model. 95, 249–287.

Getis, A., Ord, J., 1992. The analysis of spatial association by distance statistics. Geogr. Anal. 24, 189–206.

Gifford, R.M., 2003. Plant respiration in productivity models: conceptualisation, representation and issues for global terrestrial carbon-cycle research. Funct. Plant Biol. 30, 171–186.

Goodchild, M., 1986. Spatial autocorrelation. Concepts and Techniques in Modern Geography. Geobooks, Norwich 47, 3–6.

Hall, R.J., Raulier, F., Price, D.T., Arsenault, E., Bernier, P.Y., Case, B.S., Guo, X., 2006. Integrating remote sensing and climate data with process-based models to map forest productivity within west-central Alberta's boreal forest: Ecoleap-West. For. Chron. 82, 159–178.

Landsberg, J.J., Gower, S.T., 1997. Applications of Physiological Ecology to Forest Management. Academic Press, San Diego, CA, p. 354.

Landsberg, J.J., Johnsen, K.H., Albaugh, T.K., Allen, A.L., McKeand, S.E., 2000. Applying 3-PG, a simple process-based model designed to produce practical results, to data from Loblollly pine experiments. For. Sci. 47, 43–51.

Landsberg, J.J., Waring, R.H., 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecol. Manage. 95, 209–228.

Landsberg, J.J., Waring, R.H., Coops, N.C., 2003. Performance of the forest productivity model 3PG applied to a wide range of forest types. Forest Ecol. Manage. 172, 199–214.

Law, B.E., Goldstein, A.H., Anthoni, P.M., Unsworth, M.H., Panek, J.A., Bauer, M.R., Fracheboud, J.M., Hultman, N., 2001. Carbon dioxide and water vapor exchange by young and old *P. ponderosa* ecosystems during a dry summer. Tree Physiol. 21, 299–308.

Law, B.E., Waring, R.H., Anthoni, P.M., Abers, J.D., 2000a. Measurements of gross and net ecosystem productivity and water vapour exchange of a *Pinus ponderosa* ecosystem, and an evaluation of two generalized models. Global Change Biol. 6, 155–168.

Law, B.E., Williams, M., Anthoni, P.M., Baldocchi, D.D., Unsworth, M.H., 2000b. Measuring and modeling seasonal variation of carbon dioxide and water vapor exchange of a *Pinus ponderosa* forest subject to soil water deficit. Global Change Biol. 6, 613–630.

Li, Y., White, R., Chen, D., Zhang, J., Li, B., Zhang, Y., Huang, Y., Edis, R., 2007. A spatially referenced water and nitrogen management model (WNMM) for (irrigated) intensive cropping systems in the North China Plain. Ecol. Model. 203, 395–423.

Makela, A., Landberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., Puttonen, P., 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. 20, 289–298.

McKenney, D.W., Kesteven, J.L., Hutchinson, M.F., Venier, L., 2001. Canada's plant hardiness zones revisited using modern climate interpolation techniques. Can. J. Plant Sci. 81, 139–143.

Miller, J., Franklin, J., Aspinall, R., 2007. Incorporating spatial dependence in predictive vegetation models. Ecol. Model. 202, 225–242.

Nightingale, J.M., Coops, N.C., Waring, R.H., Hargrove, H.H. Comparison of MODIS gross primary production estimates for U.S. forests with those generated by 3-PGS, a model that accounts for spatial variation in soil properties. Remote Sens. Environ.

Nightingale, J.M., Phinn, S.R., Held, A.A., 2004. Ecosystem process models at multiple scales for mapping tropical forest productivity. Prog. Phys. Geogr. 28, 241–281.

Ord, J., Getis, A., 1995. Local spatial autocorrelation statistics: distributional issues and an application. Geogr. Anal. 27, 286–306.

Ord, J., Getis, A., 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. J. Reg. Sci. 41, 411–432.

Peng, C.H., Liu, J.X., Dang, Q.L., Apps, M.J., Jiang, H., 2002. TRIPLEX: a generic hybrid model for predicting forest growth and carbon and nitrogen dynamics. Ecol. Model. 153, 109–130.

Rogerson, P.A., 2002. Change detection thresholds for remotely sensed images. J. Geogr. Syst. 4, 85–97.

Running, S.W., 1994. Testing FOREST-BGC ecosystem process simulations across a climatic gradient in Oregon. Ecol. Appl. 4, 238–247.

Running, S.W., Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications. I. Hydrologic balance, canopy gas exchange and primary production processes. Ecol. Model. 42, 125–154.

Running, S.W., Hunt Jr., E.R., 1993. Generalization of a forest ecosystem process model for other biomes. BIOME-BGC and an application for global scale models. In: Ehleringer, J.R., Field, C. (Eds.), Scaling Physiological Processes: Leaf to Globe. Academic Press, San Diego, pp. 141–158.

Ryan, P.J., McKenzie, N.J., O'Connell, D., Loughead, A.N., Leppert, P.M., Jacquier, D., Ashton, L., 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. For. Ecol. Manage. 138, 139–157.

Swain, D.L., Hutchings, M.R., Marion, G., 2007. Using a spatially explicit model to understand the impact of search rate and search distance on spatial heterogeneity with an herbivore grazing system. Ecol. Model. 203, 319–326.

Swensen, J.J., Waring, R.H., Fan, W., Coops, N.C., 2005. Predicting site index with a physiologically based growth model across Oregon, USA. Can. J. For. Res 35, 1697–1707.

Thornton, P.E., Running, S.W., 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agric. For. Meteorol. 93, 211–228.

Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. J. Hydrol. 190, 214–251.

Upton, G., Fingleton, B., 1985. Spatial Data Analysis by Example, Volume 1: Point Pattern and Quantitative Data. John Wiley and Sons, New York, NY, p. 410.

Waring, R.H., Coops, N.C., Ohmann, J.L., Sarr, D.A., 2002. Interpreting woody plant richness from seasonal ratios of photosynthesis. Ecology 83, 2964–2970.

Waring, R.H., Emmingham, W.H., Running, S.W., 1975. Environmental limits of an endemic spruce *Picea breweriana*. Can. J. Bot. 53, 1550–1613.

Waring, R.H., Landsberg, J.J., Williams, M., 1998. Net primary production of forests: a constant fraction of gross primary production? Tree Physiol. 18, 129–134.

Waring, R.H., Running, S.W., 1998. Forest Ecosystems: Analysis at Multiple Scales. Academic Press, p. 370.

Waring, R.H., Schlesinger, W.H., 1985. Forest ecosystems: Concepts and management. Academic Press, Orlando, San Diego, NY, p. 340.

Wulder, M., Boots, B., 1998. Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. Int. J. Remote Sens. 19, 2223–2231.

Wulder, M., Boots, B., 2001. Local spatial autocorrelation characteristics of Landsat TM imagery of managed forest area. Can. J. Remote Sens. 27, 67–75.