SPATIAL ANALYSIS FOR MONITORING FOREST HEALTH

Francis A. Roesch, USDA Forest Service Southern Forest Experiment Station
Room T-10210, 701 Loyola Avenue, New Orleans, LA 70113 USA

KEY WORDS: Population dynamics, forest health

ABSTRACT

A plan for the spatial analysis for the sample design for the detection monitoring phase in the joint USDA Forest Service/EPA Forest Health Monitoring Program **(FHM)** in the United States is discussed. The spatial analysis procedure is intended to more quickly identify changes in forest health by providing increased sensitivity to **localized** changes. The procedure is initiated with a series of median sweeps along axes of interest. The resulting effect vectors and residuals are carried through time and analyzed for changes in trend. An example is given utilizing FHM data from New England.

INTRODUCTION

It is the goal of the detection monitoring phase of the Forest Health Monitoring **(FHM)** Program in the United States to be able to detect changes in the forest condition which might be indicative of serious forest health problems. It is hoped that this detection will be achieved by monitoring variables which would theoretically indicate a change in forest health. Efficient detection of changes in forest health requires an analytical approach somewhat different from those traditionally found in the forestry literature. We can recognize this if we think about the fairly well established ecological principle that stressors affect species at the limits of their ranges first. If we want to detect the early symptoms of stress on species of plants distributed over a wide geographical area, we have to pay attention to very localized changes, rather than say changes over the entire range of the species.

There are many alternative analytical treatments of the dimensions from a design which samples three dimensional space and time. One could ignore the fact that the data came from a spatially correlated system by treating the data as if they were independently derived at each point in time. This is a common approach in forestry, and this assumption of homogeneity across space is warranted by the usual intensive nature of forestry investigations. If, on the other hand, the survey is to be extensive as it is for the FHM Program, then some accounting for differences due to relative spatial position is necessary. Along these lines we could fully analyze the effect of all of the spatial dimensions or we could implicitly undermine the importance of one (or more) of the dimensions by collapsing it down into the remaining dimensions (as we will usually do for elevation). This paper explores potential spatial representations of the FHM data under the premise that most variables of interest will display spatial trend in some direction(s). It suggests that this exploratory spatial analysis is a necessary first step in the analytical procedure of FHM data to circumvent potential problems associated with non-stationarity.

THE FHM DESIGN

The FHM design actually varies slightly by region and state but in general it consists of a randomly placed triangular grid covering the United States, with some noise in the exact grid point locations. The grid was established by arbitrarily positioning a pattern consisting of approximately 28000 hexagons of 635 km^2 (each an individual HEX) over an area somewhat larger than the lower 48 states of the United States. A point was randomly selected to occupy the same position within each HEX to form the (noiseless) triangular grid. From each point of this grid, the nearest USDA Forest Inventory and Analysis (FIA) photo point within the HEX (or some surrogate point) is chosen for the location of a ground plot. If the photo point is classified as something other than forest, no variables are measured. If the photo point is classified as forest but landowner permission to access the plot is denied, the next closest photo point becomes the ground plot location. Note that this design chooses a single cluster sample of photo points which currently correspond to approximately **4500** forested plots over the entire United States (Scott et al., 1993). I will ignore the landowner permission problem in this paper. The entire set of photo points could be viewed as a finite set of points drawn from an **infinite** number of possible sets of points. (This infinite set is a super population.) A cluster sample of size 1 is then selected from this given set of photo points. (Note that I have simplified the sample design description from the one usually given by FHM program, which can be found in Roesch **(1994).)**

A proposed variation to the design is the incorporation of a **1/4** time interpenetrating feature in which one-fourth of the plots will be measured each year in a repeating systematic pattern as shown in Figure 1.

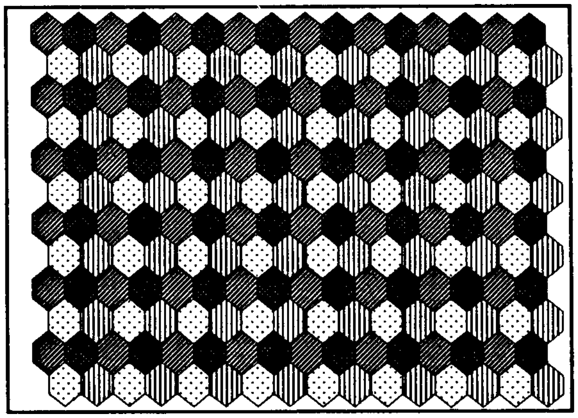

Figure 1 The interpenetrating feature. Plots in HEX's of the same pattern will be measured during the same year on a four-year rotation.

From a design-based sampling theoretic point of view, this design has both advantages and disadvantages for the purpose of the long-term monitoring of forest health. It will become clear below that most of the disadvantages can be moderated or even used to advantage through the use of spatial analytical techniques.

Some of the design weaknesses are:

1) There is no reliable variance estimator in systematic samples utilizing one random start;
2) There is no mechanism to preclude the confounding of space and time in the proposed 1/4 time interpenetrating feature;
3) It is less efficient than the design which would have resulted from an assumed integration with the FIA survey and the National Forest Surveys (NFS) because this integration could have gained a spatial advantage from the FIA and NFS plots and a temporal advantage from FHM plots (In this regard, a recent positive change has occurred for FHM's 1994 field season in that FHM plots will be overlaid onto FIA plots in Minnesota and possibly the other north central states.);
4) The triangular grid is closely aligned with major ecological systems- i.e. the NE/SW orientation of the Appalachian Mountains in the eastern United States.

The major strengths of the design include the frequency of sampling and the coverage efficiency of the regular grid. If the 1/4 time interpenetrating feature is incorporated it will save money and lessen the potential for visit effects such as trampling, at a cost of a reduced temporal measurement intensity coupled with the introduction of a spatial/temporal confounding.

ANALYSIS

An ideal analysis of this data would consider the spatial relationships of the individual plot values since autocorrelation is likely to be present for many of the variables of interest, and indeed a change in forest health is likely to be localized, at least initially. For this reason, the goal of early detection would benefit from at least an exploratory spatial analysis. The analysis I propose yields estimates of effects along axes of interest and an overall effect while preserving spatial identity in the form of a matrix of residuals, all of which can be carried through time. The overriding advantage of the median polish technique used below is that the effects are usually obtained without incorporating an undue influence from the outliers present. This leaves the outliers intact in the residuals, which is an advantage because the outliers may very well be the values of most interest. The exception to this occurs when too few observations exist in a particular row or column.

For the FHM data we could, at each point in time, arrange the plot values by their associated (noiseless) triangular grid points (HEX centers). This corresponds to what is known as a coarse mapping and serves to line up the values into rows and columns. A coarse mapping is usually accomplished by segmenting an area with a specific size grid and pooling the plots within each segment (e.g. see Cressie, 1991). The coarse mapping corresponding to the assignment of each plot value to its associated noiseless triangular grid point would result in a table of many missing values. In addition, for most variables, there will be missing values at non-forested grid points. The usual recommendation is to use a grid size for the coarse mapping which will result in a value in each cell. It could be argued, however, that the grid size selected for the coarse mapping should instead be a conscious compromise between two conflicting goals. These conflicting goals are to have confidence in each cell value and to preserve the spatial variation in the matrix. The coarser the grid the more support there is for each grid cell value and the finer the grid the finer the range of spatial variation that can be investigated. One might choose a coarser grid size to facilitate interpretation if one has reason to believe that only larger scale variation is important for a particular variable.

To start our spatial analysis, we decompose the value in cell i for time t ($Y_{i(jk)t}$) into its assumed components of an effect common to all plots, spatial effects in two directions and a residual:

$$Y_{i(jk)t} = A_t + B_{jt} + C_{kt} + R_{i(jk)t}$$

where:

A_t = *the* "all" effect at time t,
$t = 1, \ldots, T$

B_{jt} = *the jth* effect in direction B
$j = 1, \ldots, J; t = 1, \ldots, T$

C_{kt} = *the kth* effect in direction C
$k = 1, \ldots, K; t = 1, \ldots, T$

$R_{i(jk)t}$ = *the* residual in cell i
$i = 1, \ldots, N; t = 1, \ldots, T$

In T time periods, we end up with a **1xT** vector ***A*** of All effects, two matrices ***(B*** and ***C)*** of directional effects, and an i (within j and k) **xT** matrix ***R*** of residuals.

There are many ways to effect this decomposition but because we don't want the effects to be overly influenced by any outliers present, the best choice is a **2-way** median polish, using any two directions of interest (say latitude and longitude). The matrix ***R* can *be*** evaluated for special cases much as any residual analysis, resulting in our desired site-specific sensitivity. Subsequent to the median polish we can obtain residuals which are not time detrended by adding the "all" effect for each time period back into the residuals:

$$W_{i(jk)t} = R_{i(jk)t} + A_t.$$

We could then treat **the** matrix ***W* as *an*** independent set of time-series observations and perform a Bayesian trend analysis such as the one found in Van **Deusen (1994),** or look for further spatial correlations.

As the effect vectors are collected over a period of time they could be analyzed for changes in distribution. Proposing a general solution to the analysis of these vectors is problematic for a number of reasons. It could be very a very informal analysis in that one could just look at the vectors and decide whether or not they are changing enough to warrant a closer look. One could also develop some maximum absolute change criterion, based on say known stand dynamics properties. Alternatively, we could use a Bayesian trend analysis, although a satisfactorily concise prior distribution for vectors of medians from varying-length vectors of values has **thus** far eluded **this** author. The amount of effort that is put into the analysis at this stage is going to be a tradeoff between the cost of taking a closer look versus the risk of not taking it.

Optionally, we could perform additional 2-way median polishes of the residuals along other pairs of axes (say for example in the NE/SW and NW/SE directions) if we think there may be spatial trend in other directions as well.

In addition applicable FIA and NFS data could be incorporated in one of two ways, depending on the variable. Variables measured by the same standards in all three systems should be averaged within the HEX's (or within the coarse mapping cells) during a given year and then treated exactly as described above for the FHM data. For variables collected by different standards one would have to choose an **appropriate** weighting system before combining the data.

In summary, pertinent questions that will have to be addressed by the analyst for a particular case include:

1) How coarse or fine should the mapping be?,
2) In which directions are there likely to be spatial effects?, and
3) If measurement standards are different, how should the data sets be combined?

Consideration of these questions is an extremely subjective process and should be undertaken with care. The optimal solution to question (1) would result in a map which was fine enough to still display all of the important spatial trends and coarse enough to combine homogenous areas, fill all of the internal cells, and eliminate as many "strays" from the 2-way table as possible. A stray is a datum which will be the sole contributor to an element of one of the effect vectors of interest.

The answer to question (2) should come from within the life sciences, as opposed to within statistics. Although one can fish around in a spatial data set for the axes of most significant trend, the result may just be an accident of that particular measurement (or set of measurements) and may lead to poor long-term detection capabilities. To have a strong theoretical argument in support of the choice of axes is the situation to be preferred.

Because, among other things, there will not be a good estimator of within grid-cell variance from the FHM sample alone, the ***best*** guidance to a solution for question (3) is found in the Bayesian literature (e.g. Berger, 1985).

EXAMPLE

In this example, I use data from the 1992 FHM plots in New England. I calculate the percent of total basal area attributable to spruce **(Picea** sp.) and fir **(Abies** sp.) trees for each plot. The map of these values by longitude and latitude is given in Figure 2(a).

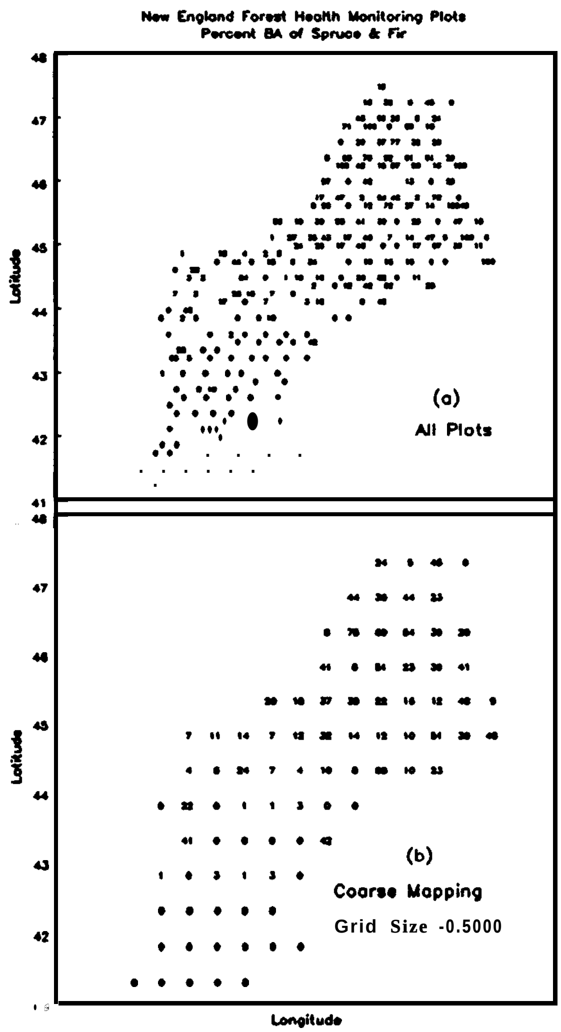

Figure 2 The percent of basal area in spruce and fir: (a) by plot, and **(b)** after plots are pooled following a coarse mapping of 0.5” in both longitude and latitude.

Note that if we made the mapping as **fine** as possible using the exact location of each plot (say to 1 minute of latitude and longitude), the median polish technique would usually be using only 1 or 2 plots for each marginal in any direction. These strays can create a problem in the analysis because they can give the appearance of an effect in the effect vectors even though one does not exist. We alleviated this by collecting the plot information to the noiseless grid point associated with the plot.

Since our goal is essentially exploratory in nature, and different stressors might be hypothesized to have effects at different scales, there are likely to be different optimal coarse grid sixes for the monitoring of different stressors on the same variable. A series of coarse mappings from **.125°** to **1°** in both longitude and latitude were conducted. Each of these resulting maps were subjected to the exploratory analysis described below. For this time period, the coarse mapping grid size of **.5°** seemed to be a good compromise in that the spatial information seemed to be about as complete as at the finer grid sixes, while the map was more comprehendible with more support for each cell. Figure 2(b) shows the percent basal area of spruce/fir after a coarse mapping by **0.5°in** both longitude and latitude. That is, the trees on plots for HEX centers which occupied common segments of a 30 minute longitude by 30 minute latitude grid were pooled:

$$P_i = \frac{\sum_{j=1}^{m_i} \sum_{k=1}^{n_j} I_k b_k}{\sum_{j=1}^{m_i} \sum_{k=1}^{n_j} b_k}$$

where

m_i = *the* number of plots in grid cell *i*
n_j = *the* number of trees in **plot *j***
I_k = **1 if tree *k* is a spruce or fir tree, and 0** otherwise, and
b_k = *the* basal area of tree ***k*.**

In consideration of what we **know** about the distribution of the spruce/fir type in New England, say as collected in Burns and **Honkala (1990),** it is of interest to examine effects along the diagonal axes (that is the NE/SW and NW/SE axes). Both the classical estimator of the variogram in the NE/SW direction

$$2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} (P_i - P_{(i+h)})$$

where

$N(h)$ = the number of distinct pairs of points separated by directed distance *h*,
$P_{(i+h)}$ = *the value* of the variable at the point separated from point *i* by directed distance *h*

and the Cressie-Hawkins robust estimator (Cressie and Hawkins (1980))

$$2\bar{\gamma}(h) = \frac{\left(\frac{1}{N(h)}\sum_{i=1}^{N(h)} |P_i - P_{(i+h)}|^{1/2}\right)^4}{\left(0.457+\frac{0.494}{N(h)}\right)}$$

were calculated and are shown in Figure 3(a). The trend in variogram estimates show the expected spatial correlation due to the trend in this variable. To estimate this trend and the trend in the NW-SE

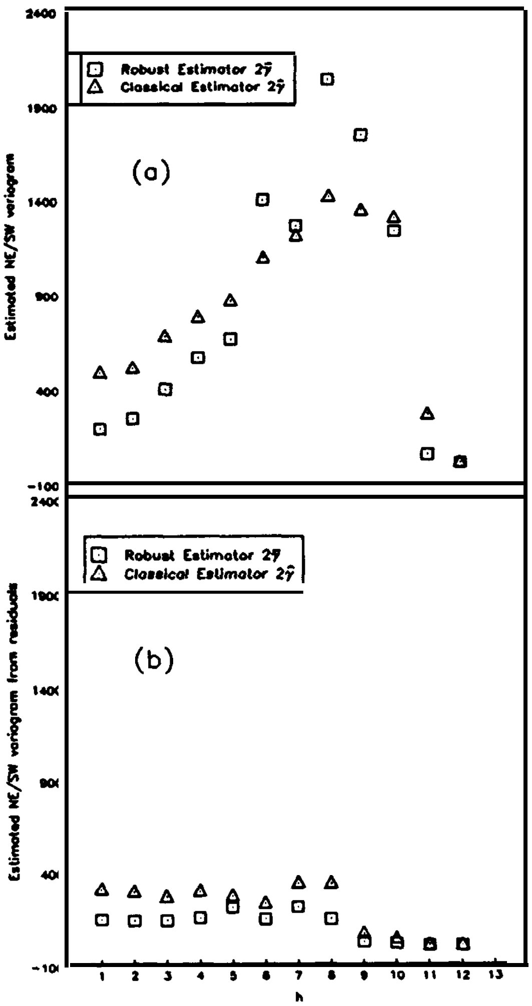

Figure 3 (a) Estimated NE/SW variogram for percent basal area spruce/fir, (b) Estimated NE/SW variograms of residuals from the median polish.

direction, the median polish technique was used. An algorithm for the median polish technique can be found in Cressie (1991, p. 186). The only data preparation necessary to use that algorithm for this application is to turn the map in Figure 2(b) 45” and fill in missing values to complete the rows and columns. Once that is done we have the NE-SW directed vectors in the rows of a kxl matrix and the NW-SE directed vectors in the columns. We then subtract the row median from each row and append the $kx1$ column vector of row medians to the right side of the matrix forming a $kx(l+1)$ matrix. Next, we subtract the column median from each of the $(l+1)$ columns and append the $1x(l+1)$ row vector of column medians to the bottom of the matrix forming a $(k+1)x(l+1)$ matrix. The values in the **first** k positions of the $(l+1)$ column are the initial row effects. The values in the first l positions of the $(k+1)$ row are the initial column effects and the value in the $(k+1), (l+1)$ position is the initial all effect. The values in the first k rows and first l columns are the initial residuals. In subsequent iterations we sweep across all of the rows and first l columns and add the resultant vector of medians to the $(l+1)$ column while subtracting it from the other columns. Likewise we sweep down all of the columns and first k rows, adding each column median to the $(k+1)$ row and subtracting it from the rest of the rows. This is done until some convergence criterion has been met. In this example we accepted convergence when the sum of the squared differences in effects (row, column, and all) was less than $1x10^{-7}$.

Figure 4 shows the results of performing the median sweeps along the diagonals which roughly correspond to the northeast to southwest direction of land types in New England. One could, for instance, postulate that a change in forest health might be detectable by analyzing change in the NW/SE effect vector of percent basal area in spruce and fir through time (the vector in the upper left part of Figure 4). This is likely to be an effective indicator of health because there is a noticeable trend in the vector initially (save for the “stray” effect at the top of the vector) and threats to the health of a population are usually first noticeable at the extremes of the areal range of the population.

The pattern of the NE/SW effect vector (vector at upper right of Figure 4) corresponds roughly to the elevation gradient in New England, reflecting the noted affinity for cool, moist climates by the spruce/fir complex in general. The “all” effect at the intersection of the two diagonal effect vectors in Figure 4 is only meaningful once an adequate time series has accumulated to test for overall trends in the proportion

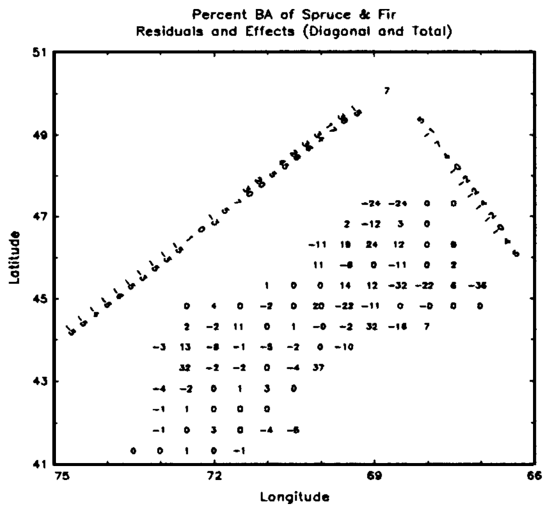

Figure 4 Residuals and Effects after median polish along NW-SE and NE-SW axes.

of basal area in spruce and **fir.** The residuals are plotted at their grid cell centers. There seem to be some interesting small-scale patterns in the residuals, but nothing as obvious as the trend in the original coarse mapping in Figure 2(b). The estimates of the NE/SW variogram of the residuals appears in Figure 3(b) and it is clear that the median polish **served** to remove the trend in the variogram.

The coarse mapping with a square of **.5°** sides resulted in an average of less than 3 plots per cell. We see that the estimates of the variogram at $h=2$ is very close to the same as at h =I in Figure 3(a). This suggests that our coarse grid mapping size could have been a little larger without a significant loss of spatial trend information. This indicates that a move to the **1/4** time interpenetrating scheme might be feasible given certain caveats. These include (1) recognition that the triangular grid is already sparse, and therefore the decision to measure **only 1/4** of the plots every year should be accompanied by a data augmentation plan such as the incorporation of the related measurements on FIA and NFS plots, and (2) the non-stationarity problem created by systematically displacing plots from one year to the next will have to be addressed.

This example has shown the use of the median polish technique subsequent to a coarse mapping for data summarization of a forest monitoring variable. It has resulted in a fairly good picture of the current state of the variable which can be carried through time to monitor changes in the variable. Since we are also keeping the plot data, we can analyze change in the forest condition at other coarse mapping scales whenever it seems appropriate.

This work has been somewhat successful to date in that we have a fairly good exploratory data analysis system for the detection monitoring of forest health. Up to the point of obtaining the effects and the residuals, the system can be readily understood by **non-**statisticians. The system also has the desirable characteristic of **being** flexible enough to incorporate any external information.

LITERATURE CITED

Berger, J.O., 1985: ***Statistical Decision Theory and Bayesian Analysis.*** Springer-Verlag, New York. 617 pp.

Bums, R.M. and **Honkala,** B.H., 1990: ***Silvics of North America, V&me 1. Conifers.*** USDA Forest Service, Agriculture Handbook 654, Washington, D.C. 675 pp.

Cressie, N., 1991: ***Statistics For Spatial Data.*** Wiley, New York, 900 pp.

Cressie, N, and Hawkins, D.M., 1980: Robust estimation of the variogram. ***Journal of the International Association for Mathematical Geology 12:* 115-125.**

Roesch, F.A. 1994: Design and analysis for detection monitoring of forest health. Paper presented at the Monte **Verità** Conference on Forest Survey Designs: “Simplicity versus Efficiency” and Assessment of Non-Timber Resources. May 2-7, 1994. Ascona, Switzerland.

Scott, C.T., Cassell, D.L. and Hazard, J.W., 1993: Sampling design of the U.S. National Forest Health Monitoring Program. **In:** ***Proceedings Ilvessalo Symposium on National Forest Inventories.*** Helsinki, Finland, August 17-21, pp. 150-157.

Van Deusen, P.C., 1994: Bayesian trend estimation for repeated surveys. In review.

ACKNOWLEDGEMENT

The data were provided by Charles Bamett and Andrew Gillespie of the USDA Forest Service Northeastern Forest Experiment Station. The author thanks Charles E. Thomas for a review of the manuscript.