ANALYSIS OF THRIPS DISTRIBUTION: APPLICATION OF SPATIAL STATISTICS AND KRIGING

John Aleong, Bruce L. Parker[1], Margaret Skinner[1] and Diantha Howard

Department of Mathematics and Statistics
Statistical and Computer Services
The University of Vermont
Burlington, Vermont USA

Abstract

Kriging is a statistical technique that provides predictions for spatially and temporally correlated data. Observations of thrips distribution and density in Vermont soils are made in both space and time. Traditional statistical analysis of such data assumes that the counts taken over space and time are independent, which is not necessarily true. Therefore, to analyze these data correctly we must account for the correlation structure in the data, which can be done with Kriging. The Kriging technique is reviewed and its use illustrated in determining the pattern of thrips distribution and density in Vermont by analysis of data from the Vermont Pear Thrips Soil Survey for the 1988-89 season.

Introduction

Pear thrips, *Taeniothrips inconsequens* (Uzel) (Thysanoptera: Thripidae), is a serious problem in Vermont and other eastern states, causing severe foliage damage to sugar maple trees in the early spring (Parker et al. 1988). A research/management project was initiated cooperatively by the University of Vermont Entomology Research Laboratory and the Vermont Department of Forests, Parks and Recreation in 1988 to develop effective methods to survey and

---

[1] Entomology Research Laboratory, The University of Vermont, South Burlington, Vt. 05403.

ultimately manage the pear thrips in sugar maple stands statewide. One of the highest priorities of this project was to develop a method to predict tree damage based on thrips population densities in the soil. A soil survey to determine the distribution and density of pear thrips was first conducted in Vermont in January 1989. Analysis of these data was necessary as a first step for management activities.

Kriging is a statistical technique developed for prediction of data values when the data are correlated in space and time. Thrips population data which are collected annually from soil surveys may be correlated spatially within and between time. Such a data set lends itself well to Kriging. The correlation in these data is a function of the distance between sample sites. Counts at neighboring or nearby sites should be more highly correlated then those at sites located further apart.

Cressie (1986, 1990) and Johnson (1990) give reviews of the theory and applications of Kriging, tracing its origin to geostatistics and the work of Matheron (1971) and Krige (1951). There are many applications for Kriging, including interpretation of rainfall data (Ord & Rees 1979), for soil mapping (Burgess & Webster 1980), and in groundwater pollution monitoring (Yates & Yates 1988). Kriging can also be a useful tool in the design and analysis of experiments, such as for uniformity trials to determine blocking mechanisms in agricultural experiments and to predict yields at unobservable points in fields using systematic samples (Johnson 1990). The purpose of this paper is to give a brief review of Kriging and to illustrate its application in the analysis of data on pear thrips population distribution.

Materials and Methods

Kriging Methodology

Following Johnson (1990), let x denote spatial location (i.e., the latitude and longitude of a point [sample site] at which the number of thrips are observed). Let $v(x)$ be the observed or unobserved number of thrips at location x, at a particular time. Thus, let

$$\mu(x) = E[v(x)] \tag{1}$$

where E denotes the expectation of $v(x)$.

$$\sigma^2(x) = \mathrm{Var}[v(x)] = E[v(x) - \mu(x)]^2 \tag{2}$$

$$c(x, x') = E[[v(x) - \mu(x)]\ [v(x')] - \mu(x')]] \tag{3}$$

and

$$\gamma(x,x') = 0.5E[v(x) - v(x')]^2 \tag{4}$$

Note that $c(x,x')$ is the covariance between the number of thrips at locations x and x' while $\gamma(x,x')$ is the *semivariogram*, with $2\gamma(x,x')$ as the *variogram*:

$$\gamma(x,x') = \gamma(x-x') = \gamma(h) \tag{5}$$

i.e., the variogram is a function of the distance between x and x'.

For a particular time, let us sample thrips at locations $x_1, \ldots, x_n$. Let v_i be the mean number of thrips per sample at location x_i, and $i = 1, \ldots, n$. Given the data (x_i, v_i), $i = 1, \ldots, n$, we want:

1. To estimate the number of thrips at an unobserved point x, with its standard error, and

2. To estimate the average number of thrips over some region in the state.

To estimate the number of thrips $\mu(x)$ at an unobserved point, we consider linear combinations of $\mu(x)$ at the observed locations. Let $\hat{v}(x) = \Sigma a_i \, v(x_i)$ and select a_i such that:

$$1. \quad E[\hat{v}(x) - v(x)] = 0 \tag{6}$$

and

$$2. \quad \mathrm{var}[\hat{v}(x) - v(x)] = E[\hat{v}(x) - v(x)]^2 \tag{7}$$

are minimized. If we find a vector a such that (6) and (7) are minimized then:

$$\hat{v}(x) = \Sigma a_i v(x_i) \tag{8}$$

is the Kriging estimate of $v(x)$ and the Kriging coefficients are the vector a.

To use the Kriging method, we must estimate the variogram, i.e., $\gamma(x,x')$ or cov (x,x'). For the variogram we use the exponential model:

$$\gamma(h) = B + C[1 - e^{-h/a}], \text{ for all } h \geq 0. \tag{9}$$

The *sill* of this model, which is equal to $B + C$, is the maximum value that the variogram attains, and is the value of $\gamma(h)$ as h goes to infinity. The *range* is the distance beyond which two points are uncorrelated. The *nugget* is the value of the variogram at $h = 0$.

Because data are taken annually, we can expand the notation to include both space and time. Let v_{ij} be the number of thrips at location x_i, and time t_j for $i = 1, \ldots, n$ and $j = 1, \ldots k$. The data will be given in the form (x_i, t_j, v_{ij}), $i = 1, \ldots, n$ and $j = 1, \ldots, k$. For any particular time j, the process v_{ij} is purely spatial, whereas for any fixed location the process v_{ij} is temporal. Because we are illustrating the Kriging methodology here with one year's data (1988-89) we will consider only a purely spatial process (x_i, v_i), $i = 1, \ldots n$.

Thrips Data Collection

In January and February 1989, soil samples were taken with a hand-held bulb planter in sugar maple stands (those with sugar maple comprising more than 75% of the basal area) throughout the state to determine the distribution and relative density of pear thrips (Skinner & Parker 1989). Samples were taken in 91 stands in 13 of the 14 counties in Vermont (Fig. 1). Because pear thrips reside in soil from mid-June to mid-April (see Skinner et al., poster presentation, this publication), thrips in samples at this time reflect the potential population that entered the soil in 1988 and would emerge to cause foliar damage in the spring of 1989. In each stand soil samples were taken at 2 and 4 m from the south side of the bole of five sugar maple trees (Skinner & Parker 1989).

Stands were selected by personnel from the Vermont Department of Forests, Parks and Recreation using information from an aerial survey of pear thrips damage in 1988. Sites were selected in each county from each of three damage categories based on an estimate of leaf area reduction (light - 0 - 30% reduction, moderate - 31 - 60% reduction, and severe - 61 - 100% reduction).

Thrips were extracted from samples using a heptane flotation procedure (see Grehan & Parker, poster presentation, this publication). Residue from the extraction process was inspected under magnification to determine the number of pear thrips per sample. The total number of thrips per site was divided by the number of samples extracted from the site (generally 10 samples) and this mean number of thrips per sample was used for analysis.

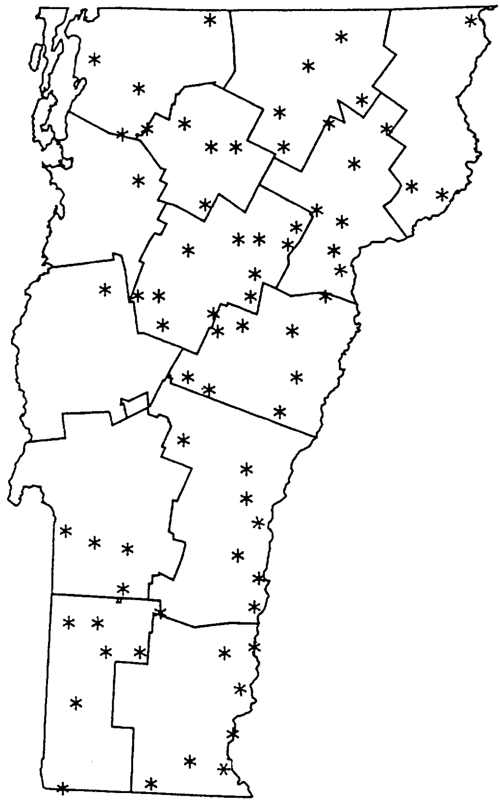


Figure 1. Townships (*) in which samples were taken for the Vermont Pear Thrips Soil Survey in 1988-89.

Data Analysis

The mean number of thrips per sample was determined for each township where soil samples were taken. Where more than one stand was sampled in the same town, the mean number of thrips was calculated by combining data from all stands within that town. Thrips data from 91 stands in 69 townships were used for Kriging analysis. Using the latitude and longitude coordinates for each town, mean thrips data were analyzed with software by Englund & Sparks (1988).

Results and Discussion

Figure 1 gives the spatial location, x_i of thrips sample collection sites in Vermont in 1988 while Figure 2 shows the scatter diagram of the mean number of thrips per sample, v_i, at each site.

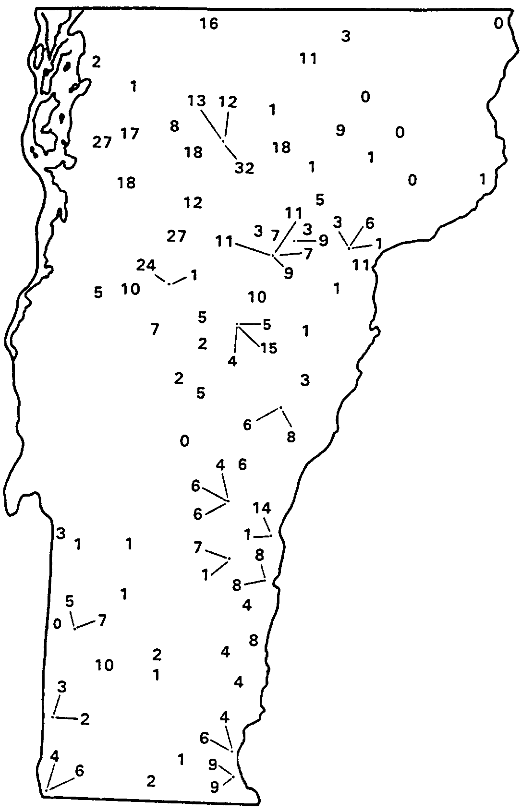


Figure 2. Scatter diagram of the mean number of thrips per sample in sites sampled for the Vermont Pear Thrips Soil Survey in 1988-89.

As expected, the histogram of thrips density data is highly skewed (Fig. 3a). Some of this skewness is corrected by the Anscombe transformation, $\sqrt{v_i + 3/8}$, where v_i is now the average thrips count at a location, making a more symmetrical histogram (Fig. 3b). (All subsequent reference to a square root transformation in this paper is the Anscombe transformation.)

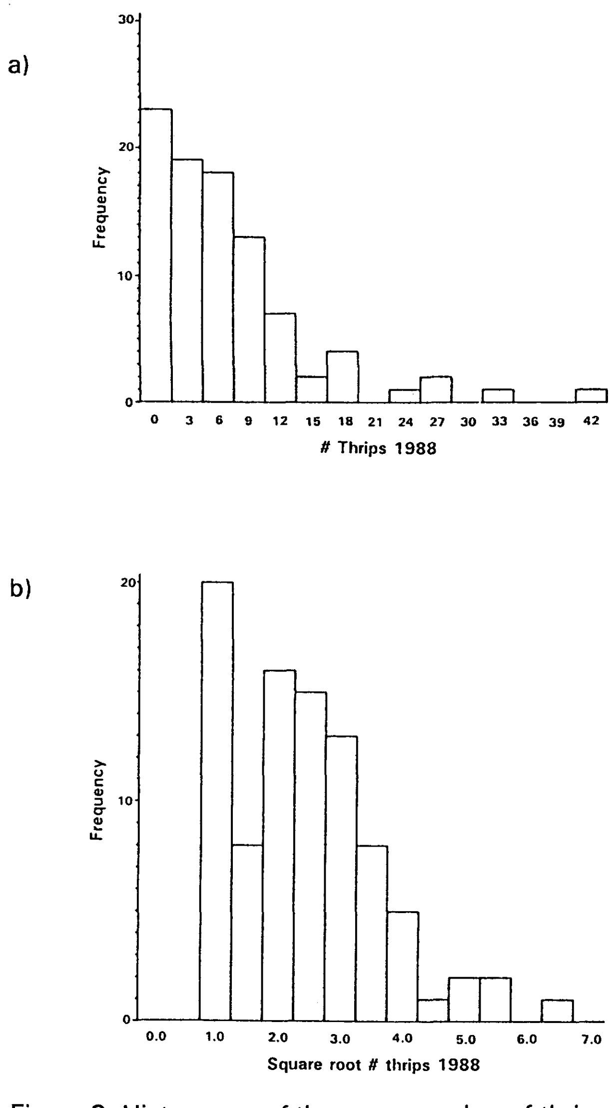
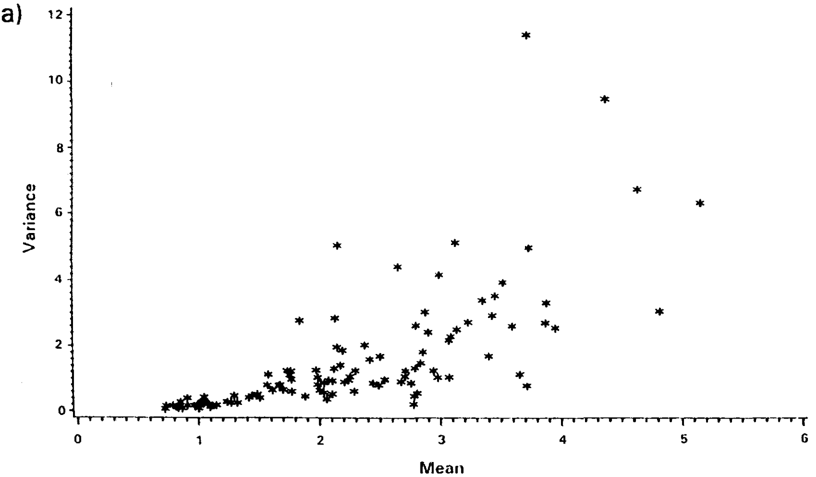


Figure 3. Histograms of the mean number of thrips per sample for each site sampled for the 1988-89 Vermont Pear Thrips Soil Survey, a) for the original data, b) for the transformed data, $\sqrt{v_i + 3/8}$.

A plot of the variance versus the mean number of thrips per sample for each site indicates that the variance is proportional to the mean (Fig. 4a), suggesting the suitability of a square root transformation for stabilizing the variance among sampling sites. The variance-mean plot of the transformed data show less proportionality (Fig. 4b).

a)



b)

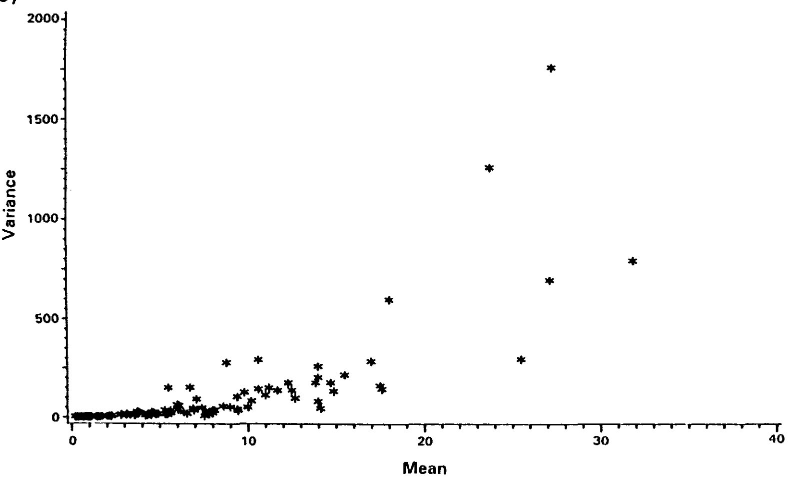


Figure 4. Plots of variance versus the mean number of thrips per sample from the Vermont Pear Thrips Soil Survey in 1988-89, a) for original data, b) for transformed data, $\sqrt{v_i + 3/8}$.

The variogram, as given in equations (5) and (9), is a function of the distance between sample sites and depends only on the relative position of the thrips count at locations x, and x'. The plot of the variogram versus distance for the original data shows that the variance of the difference in thrips counts between two locations is a function of the distance between the two locations (R^2 = 0.92) (Fig. 5a). Figure 5b gives the variogram using the transformed data.

The equation of the estimated variogram is $\gamma(h) = 10.13 + 35.66\ (1 - e^{-h/0.2218})$ on the original scale and $\gamma(h) = 0.3740 + 1.0576\ (1 - e^{-h/0.2694})$ on the square root scale. An examination of the plot of the transformed data shows that an exponential variogram seems to fit the data well with *nugget* = 0.37, *sill* = 1.43, *range* = 0.81 and R^2 = 0.92. The nonlinear regression models for the original and transformed data gave the same R^2 (= sum of squares for regression ÷ total sum of squares). However, because the Anscombe transformation appears to have stabilized the variance of the thrips counts between sampling sites (Fig. 4), we will use the variogram on the transformed scale. Figure 5b shows that the variance of the difference between two thrips counts is small for nearby or neighboring sites but increases exponentially as the distance between two sites increases, until the variance approaches its asymptote.

Using the Kriging technique, the number of thrips at unobserved points in Vermont, having habitat characteristics similar to that in the sample sites, can be predicted based on available data. The thrips count at the observed points are given in Figure 2 and the predicted number of thrips at unobserved points are given in Figures 6 (original data) and 8 (transformed square root scale). The contours in these figures show the areas of high (northern and central Vermont) and low (southern Vermont) infestation of thrips. For each predicted value of thrips, the standard deviations are given in the form of contours in Figures 7 and 9 for the original and square root counts, respectively.

As discussed, this data set should be analyzed on the square root scale. Therefore the practitioner should use the analyses done on this scale to predict thrips density. For illustration, using Figures 8 and 9, we can predict the mean number of thrips per sample in northeastern Vermont to fall between 3.2 ± 0.4 to 3.6 ± 0.4 (on the square root scale), which would represent a mean of approximately 9.9 to 12.6 thrips per sample after calculation back to the original scale.

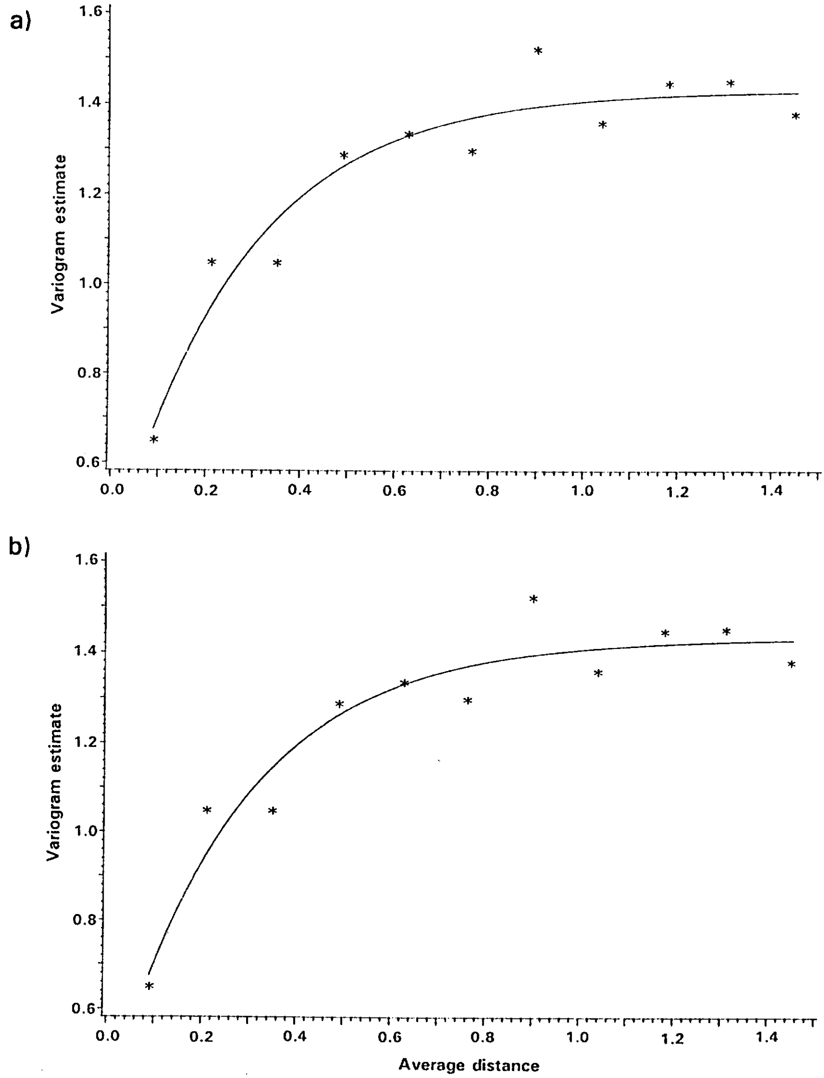


Figure 5. Variograms for data from the Vermont Pear Thrips Soil Survey in 1988-89, a) for original data, b) for transformed data, $\sqrt{v_i+3/8}$.

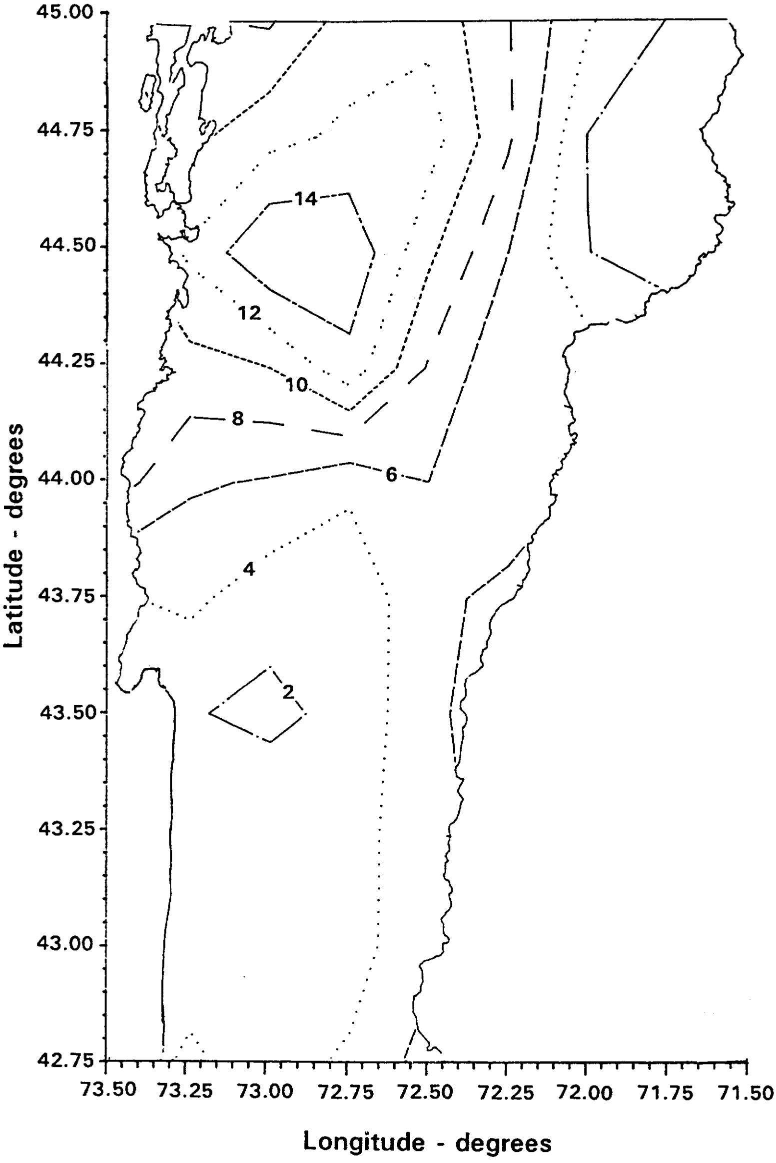


Figure 6. Map of Vermont showing contours for thrips density based on Kriging estimates from original data of the Vermont Pear Thrips Soil Survey in 1988-89.

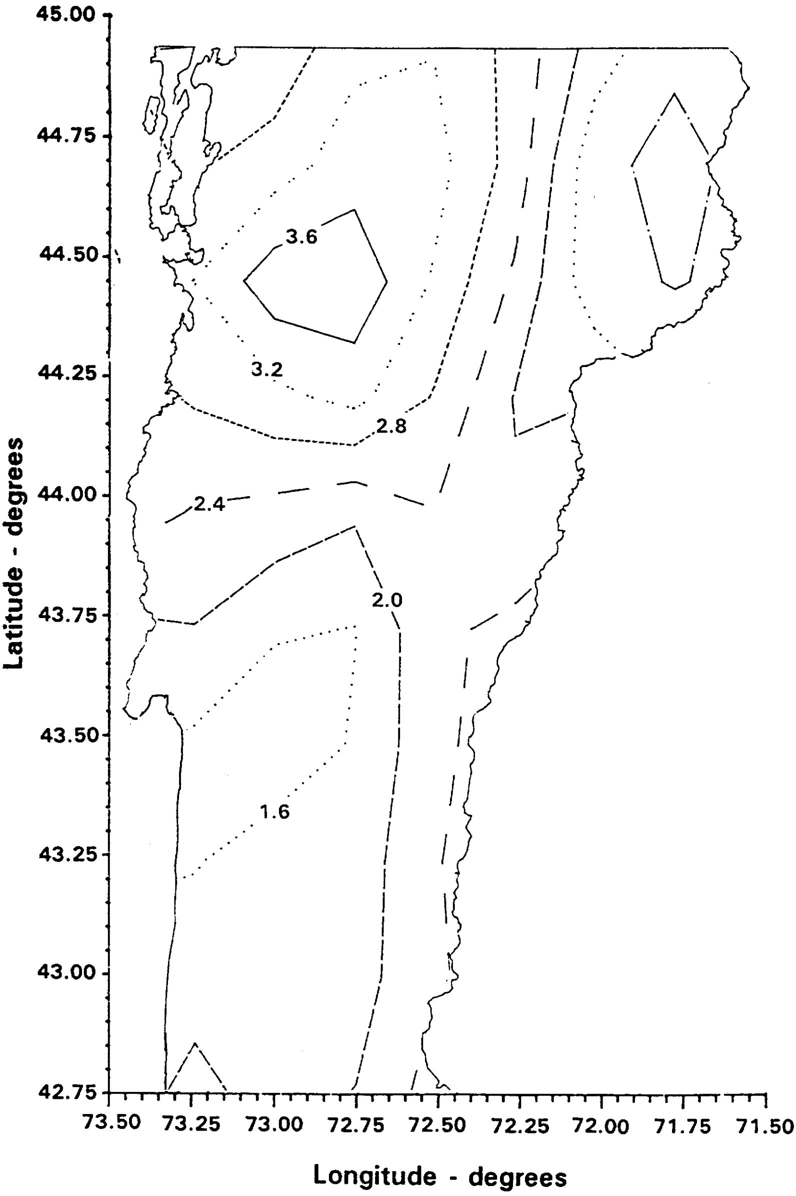


Figure 7. Map of Vermont showing contours for standard deviations of thrips density based on Kriging estimates from original data of the Vermont Pear Thrips Soil Survey in 1988-89.

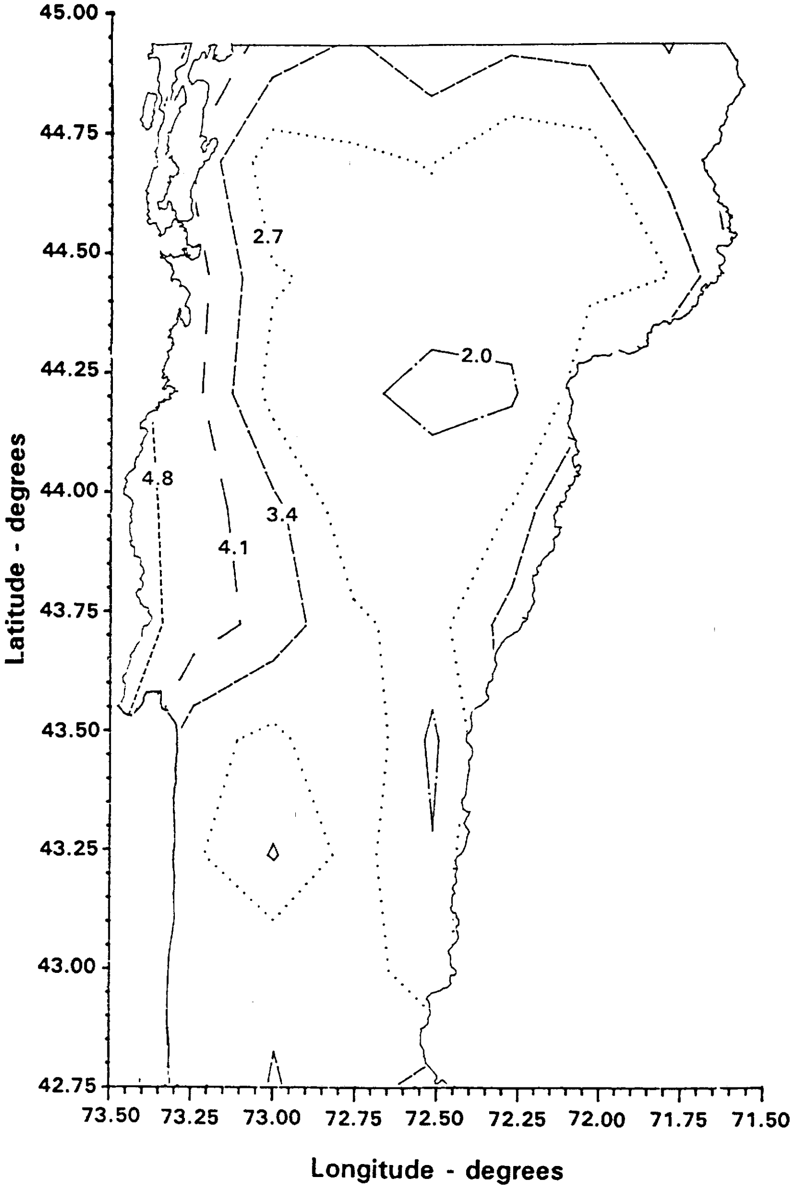


Figure 8. Map of Vermont showing contours for thrips density based on Kriging estimates from transformed data, $\sqrt{v_i + 3/8}$, of the Vermont Pear Thrips Soil Survey in 1988-89.

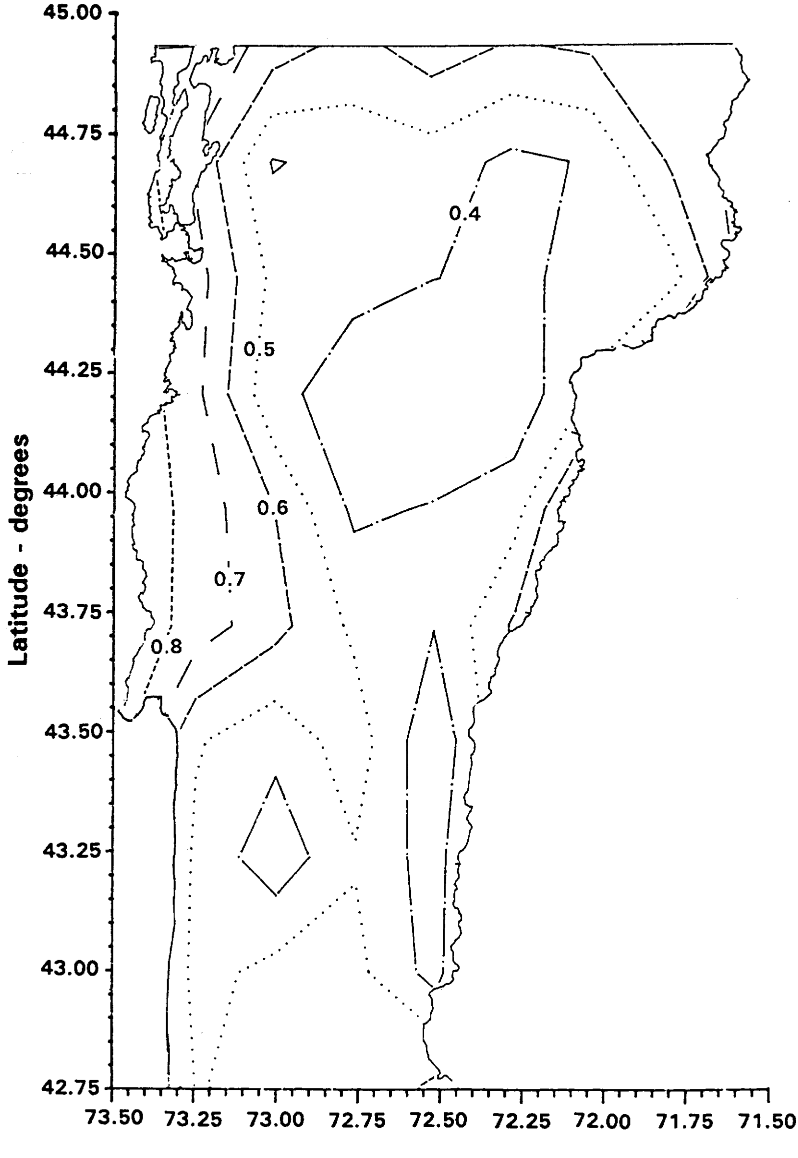


Figure 9. Map of Vermont showing contours for standard deviations of thrips density based on Kriging estimates from transformed data, $\sqrt{v_i + 3/8}$, of the Vermont Pear Thrips Soil Survey 1988-89.

This statistical method shows promise for estimation of pear thrips densities in the soil in areas where sampling has not been done, and ultimately for prediction of the extent of thrips damage statewide in the spring. More research is needed to assess the value of this methodology for pear thrips management. Firstly, verification of thrips density based on Kriging values is needed to determine if in fact it can be used to accurately estimate existing thrips densities.

Secondly, the relationship between thrips density in the soil and the resultant damage in the spring must be investigated. Preliminary results suggest that this relationship may be fairly weak, i.e., a low number of thrips in the soil does not guarantee that damage will not occur in the spring, and visa versa. Considering this, the value of thrips density as a predictor of damage may be questionable.

Thirdly, this Kriging method can be extended to a disjunctive Kriging method whereby we can calculate the conditional probability that the mean number of thrips per sample is greater than a critical level, i.e., the mean number of thrips at which severe damage would be certain to occur and when pest suppression is deemed an economic or environmental necessity. If we can model thrips population levels in the soil in relation to the damage in the spring, then these probabilities can be used in management of pear thrips by helping pest managers determine whether suppression action is warranted. In the future we intend to analyze thrips soil survey data from several years. This will provide information on the pattern of thrips population trends over time as well as space.

The Kriging methodology could be applied to other forest pests as well as the pear thrips. Information on thrips and other pest populations on a large scale is essential for effective pest management implementation. However, resources for monitoring pests on a statewide scale is often limited. Therefore, development of statistical methods that reduce the need for extensive sampling but also provide reliable predictions on pest population levels over a large area would be invaluable.

Conclusion

Kriging is an easy graphical method that can assist entomologists design and analyze their experiments. It can be used to predict thrips infestations statewide from well-designed sample data. This will provide valuable information for making decisions for the management of pear thrips and other important insect pests. The validity of prediction should now be validated under field conditions.

Acknowledgment

The research described in this article was supported in part by the Vermont Agricultural Experiment Station and the Vermont Department of Forests, Parks and Recreation. Thanks go to the many technicians who collected and extracted the soil samples and Peggy Verville for word processing.

References Cited

Burgess, T. M. & R. Webster. 1980. Optimal interpolation and isarithmetic mapping of soil properties, I: the semi-variogram and punctual Kriging. J. Soil Sci. 31: 315-331.

Cressie, N. A. C. 1986. Kriging nonstationary data. J. Am. Statistical Assoc. 81: 625-634.

Cressie, N. A. C. 1990. Spatial prediction from networks. Research Report, Dept. of Statistics, Iowa State Univ., Ames, Iowa.

Englund, E. & A. Sparks. 1988. Geostatistical environmental assessment software users guide. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S.E.P.A., Los Vegas, Nev.

Johnson, D. E. 1990. (unpublished). Introduction to spatial statistics and Kriging with some possible application to experimental designs. Paper presented at the Joint Statistical Meeting of the Am. Statistical Assoc. and the Biometric Soc., August 1990, Anaheim, Calif.

Krige, D. G. 1951. A statistical approach to some mine valuations and allied problems in Witwatersrand. Master's thesis. University of the Witwatersrand, Johannesburg, South Africa.

Matheron, G. 1971. The theory of regionalized variables and its applications. Cahiers du Centre de Morphologie Mathematique, No. 5, Fontainebleau, France.

Ord, J. K. & M. Rees. 1979. Spatial processes: recent developments with applications to hydrology, pp. 95-118. *In* E. H. Lloyd, T. O'Donnell & J. C. Wilkinson [eds.], The mathematics of hydrology and water resources. Academic Press, London.

Parker, B. L., M. Skinner & H. B. Teillon. 1988 Proceedings, regional meeting: "The 1988 thrips infestation of sugar maple". Bennington, VT June 23. Vt. Agr. Exp. Sta. Bull. 696. Univ. Vt., Burlington. 113 pp.

Skinner, M. & B. L. Parker. 1989. Vermont soil survey for pear thrips, protocol for statewide sampling program. Research report, Entomology Research Laboratory, South Burlington, Vt.

Yates, S. R & M. V. Yates. 1988. Disjunctive Kriging as an approach to management decision making. Soil Sci. Soc. Am. J. 52: 1554-1558.