Using Predictions Based on Geostatistics to Monitor Trends in Aspergillus flavus Strain Composition

Thomas V. Orum, Donna M. Bigelow, Peter J. Cotty, and Merritt R. Nelson

First, second, and fourth authors: Department of Plant Pathology, University of Arizona, Tucson 85721; and third author: Southern Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, New Orleans, LA 70179. Accepted for publication 23 May 1999.

ABSTRACT


Aspergillus flavus is a soil-inhabiting fungus that frequently produces aflatoxins, potent carcinogens, in cottonseed and other seed crops. A. flavus S strain isolates, characterized on the basis of sclerotial morphology, are highly toxicogenic. Spatial and temporal characteristics of the percentage of the A. flavus isolates that are S strain (S strain incidence) were used to predict patterns across areas of more than 30 km². Spatial autocorrelation in S strain incidence in Yuma County, AZ, was shown to extend beyond field boundaries to adjacent fields. Variograms revealed both short-range (2 to 6 km) and long-range (20 to 30 km) spatial structure in S strain incidence. S strain incidence at 36 locations sampled in July 1997 was predicted with a high correlation between expected and observed values (R² = 0.85, P = 0.0001) by kriging data from July 1995 and July 1996. S strain incidence at locations sampled in October 1997 and March 1998 was markedly less than predicted by kriging data from the same months in prior years. Temporal analysis of four locations repeatedly sampled from April 1995 through July 1998 also indicated a major reduction in S strain incidence in the Texas Hill area after July 1997. Surface maps generated by kriging point data indicated a similarity in the spatial pattern of S strain incidence among all sampling dates despite temporal changes in the overall S strain incidence. Geostatistics provided useful descriptions of variability in S strain incidence over space and time.

Additional keywords: cotton, geographic information systems, GIS.

Geostatistics provides a set of tools useful in characterizing variability over space. Fungi that produce aflatoxins are highly variable in diverse traits, and communities of these fungi frequently differ among regions. The need for tools to predict and monitor regional shifts in community structure of aflatoxin-producing fungi has increased because of severe periodic episodes of aflatoxin contamination and the development of control technologies directed toward changing the composition of fungal communities.

Aflatoxins are very toxic, carcinogenic compounds produced by Aspergillus flavus Link:Fr. (9). These toxins are highly regulated food contaminants, and methods to reduce contamination of cottonseed in the field are being actively investigated. A. flavus, the primary cause of aflatoxin contamination of cottonseed, can be divided into the S and L strains, based on sclerotial morphology (4). The L strain isolates are highly variable in aflatoxin production and include naturally occurring atoxigenic (do not produce aflatoxins) strains. The S strain isolates consistently (>98%) produce high aflatoxin levels. An aflatoxin control strategy directed at increasing the atoxigenic component of the A. flavus community is being developed (5,6,8,10).

The percentage of the A. flavus isolates that are S strain is termed the S strain incidence. Where S strain incidence is high, a decrease in S strain incidence in areas treated with an atoxigenic L strain may be a useful indicator of successful displacement. S strain incidences over 50% are common in certain regions. It is much less expensive to characterize S strain incidence than to genetically identify a specific atoxigenic L strain. Furthermore, the S strain produces very high aflatoxin concentrations, and a better understanding of regional fluctuations in S strain incidence may lead to additional aflatoxin management options.

The use of geostatistics and geographic information systems (GIS) in commercial agriculture and in agricultural research is increasing (17). Two important features of geostatistical analyses, smoothing and prediction, are used in the current work to describe trends in A. flavus strain composition. Smoothing is used in the creation of surface maps to reveal broad patterns that may be hidden by the scatter of local variability. Prediction is the result of estimating the value of a variable at unsampled locations by interpolating data from nearby sampled locations. Biological data are often spatially autocorrelated or patchy. This means that, on average, values of a variable from a cluster of nearby locations will be more similar to each other than to values from more distant locations. In geostatistics, spatial autocorrelation is characterized and used to improve smoothing and prediction.

Estimates of a variable at unsampled locations may be achieved with inverse distance weighting. The inverse distance weighting procedure calculates a weighted average of nearby observed data values. The weighting is proportional to a power (exponent) of the inverse of the distance between the unsampled location and the nearby data locations. This is consistent with an assumption of spatial autocorrelation, because the closer a sample location is to an unsampled location, the greater the weighting. GIS software packages such as ArcView Spatial Analyst (Environmental Systems Research Institute, Inc. [ESRI]; Redlands, CA) permit the display of surface maps created from point data using inverse distance weighting. Advantages of inverse distance weighting are its accessibility in software packages and its easy-to-understand incorporation of spatial autocorrelation. In inverse distance weighting, a search radius needs to be specified to determine which points are included in the process. Inverse distance weighting is most effective when the search radius is approximately the range of spatial autocorrelation and the sample locations are uniformly spaced (14).

In geostatistics, spatial autocorrelation is usually characterized using variograms (13,20). Variograms can be explained in terms
of lag distances. A lag is the spacing between two points of data. Lag distances between all possible pairs of data points are calculated, and the pairs are grouped into classes according to lag distance. A variogram value characterizing the variability of a selected variable among pairs within each lag distance class is calculated (13,20). The variogram formula is such that a low variogram value corresponds to a high correlation value. The variogram values are plotted against the average lag distance for each class to produce an experimental variogram. Model variograms are functions of lag distance with parameters chosen so that the function “fits” the plot of the experimental variogram values. The model parameters are used as inputs to geostatistical programs for surface interpolation. Variogram plots can also be used to determine the spatial scale at which most of the variability in the selected variable occurs and to choose a search radius for surface interpolation by inverse distance weighting.

Kriging is the most common geostatistical procedure for surface interpolation. It is a linear regression technique that uses a model variogram (20) to determine weights based on separation distances. In contrast to inverse distance weighting, the kriging estimator not only takes into account the spatial autocorrelation between the data points and the location to be predicted, but also, to account for clustering, between each pair of data points. This is why kriging is preferred over inverse distance weighting when the sample points are not uniformly distributed. The weighting method for kriging is most easily expressed using a matrix formula: 

\[ w = C^{-1} \cdot D \]

in which \( w \), the solution, is an \( n+1 \) by 1 matrix giving the weights for the kriging estimate, \( C \) is an \( n+1 \) by \( n+1 \) matrix incorporating variogram values for distances between data locations, and \( D \) is an \( n+1 \) by 1 matrix incorporating variogram values for distances between the interpolation location and data locations (13). Kriging provides an unbiased estimate and minimizes the error variance. Other advantages of kriging are that directionality can be incorporated into the variogram models and, thus, into the kriging equations (13) and that an error variance is calculated along with each estimate (13,17).

If one has evidence for temporal as well as spatial autocorrelation, kriging can be used to predict, within limits, future values in previously unsampled locations. In the current study, the spatial and temporal characteristics of \( A. flavus \) S strain incidence in Yuma County, AZ, are described with geostatistics in order to assess the value of geostatistical methods in prediction and regional monitoring of \( A. flavus \) community structure.

**MATERIALS AND METHODS**

**Sampling procedures.** Soil samples were collected, stored, and processed according to methods described previously (19). Coordinates for sample locations were determined from Farm Services Agency (U.S. Department of Agriculture) field maps digitized in the Universal Transverse Mercator coordinate system (a metric equivalent of latitude and longitude), Zone 12, North American Datum 1927. A hand-held global positioning system unit (Garmin 45 XL; Garmin Corp., Olathe, KS) was used in conjunction with the field maps to determine the coordinates for the sample locations. To properly characterize S strain incidence at a location, at least 30 isolates (preferably more) are transferred from the modified rose bengal agar soil isolation plates to separate 5/2 agar plates (5% V8 juice and 2% agar) for strain characterization (19). Since 1994, this study has entailed the separate plating and strain characterization of 37,423 isolates.

**Sampling fungal communities in adjacent fields.** Twenty-four commercial fields were sampled in Yuma County, AZ, in July 1996, October 1996, and March 1997. The fields were selected to obtain a nested spatial pattern that included adjacent fields. In each of four areas (Fig. 1), two fields, 1 to 5 km apart, with a history of consistently different S strain incidence were selected. Two fields adjacent to each selected field were also chosen for sampling. Two locations were selected within each field at opposite corners (Fig. 2). Because these were commercial fields selected in part based on previous results, there was no control over the crops planted in the fields, and a variety of crops were observed at sample locations (Table 1).

**Area-wide sampling.** Beginning in July 1997, sampling locations were chosen to give a more uniform coverage of the North Gila and Texas Hill areas, to provide data for maps based on geostatistical techniques, and to provide data to compare predicted with observed S strain incidence across a region. In July and October 1997, 42 new locations for each date were sampled along with eight locations that had been sampled repeatedly since 1994. In March 1998 in the Texas Hill area, 36 new locations were sampled along with four locations that were previously sampled, and in the North Gila area, only four locations, all previously sampled, were sampled. In July 1998, 18 locations in the Texas Hill area, including four previously sampled locations, were sampled. Sampling patterns

![Fig. 1. Subregions of Yuma County, AZ, where commercial fields were sampled to analyze S and L strain compositions in communities of *Aspergillus flavus* in the soil.](image)

![Fig. 2. Nesting of sample locations for the analysis of variance in *Aspergillus flavus* S strain incidence among adjacent fields in the Dome and North Gila subregions. The “selected fields” (dark shading) were chosen on the basis of previous sampling. Samples were collected at locations indicated by a circled dot in July 1996, October 1996, and March 1997. The pattern was similar in Tacna West and Texas Hill (data not shown).](image)
in the Texas Hill area cover the region, but sampling had been clustered at different times to help define short-range spatial structure (Fig. 3). The irregular sampling patterns mean that kriging is preferred over inverse distance weighting for surface interpolation and prediction (14) at unsampled locations. Combining spatial and temporal analyses can be quite complex (2,11,15,21). In the current study, maps for each sampling date were generated to visually compare spatial patterns in a temporal context.

**Statistical analyses.** Spatial autocorrelation. A nested analysis of variance was used to get a basic understanding of the scale of the spatial structure (18,19,22) of S strain incidence using data from the adjacent fields’ samples collected in July 1996, October 1996, and March 1997. The terminology for the scale of nesting is the same as used in a previous study (19), and the SAS statistical package (SAS Institute, Cary, NC) was used for analysis of variance (procedure NESTED). The spatial scales evaluated were among subregions (>20 km), among widely separated fields (1 to 5 km), among adjacent fields (0.2 to 0.6 km), and among locations within fields (0.1 to 0.4 km). The SAS procedure CORR was used to determine the correlation among S strain incidences at different times at the same location. Only locations where at least 45 isolates were characterized for strain were included in this analysis. Microsoft Excel (Microsoft Corp., Redmond, WA) was used for the scatterplots and t tests.

Pair comparison files were generated, experimental variograms were plotted, and model variogram parameters were selected using GeoEAS (geostatistical environmental assessment software 1.2.1; U.S. Environmental Protection Agency [EPA], Environmental Monitoring Systems Laboratory, Las Vegas, NV) (12) and Variowin (version 2.2) (20). GeoEAS and a demonstration version of Variowin is available on-line. Other software is also available on-line and can be found by searching with the term “geostatistics.”

**Surface interpolation and prediction.** Surface maps of S strain incidence were generated from separate data sets from each sampling date after March 1997 and from combined data sets. Since previous analyses indicated seasonal variation in S strain incidence

### TABLE 1. *Aspergillus flavus* S strain incidence and crops in sets of adjacent fields with a history of low and high S strain incidence

<table>
<thead>
<tr>
<th>Date</th>
<th>Area</th>
<th>S strain incidence (%)</th>
<th>Crops&lt;sup&gt;y&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low&lt;sup&gt;*&lt;/sup&gt;</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>July 1996</td>
<td>North Gila</td>
<td>22</td>
<td>Cotton, fallow</td>
</tr>
<tr>
<td></td>
<td>Dome</td>
<td>33</td>
<td>Cotton</td>
</tr>
<tr>
<td></td>
<td>Tacna West</td>
<td>36</td>
<td>Cotton</td>
</tr>
<tr>
<td></td>
<td>Texas Hill</td>
<td>47</td>
<td>Cotton, sudan</td>
</tr>
<tr>
<td>October 1996</td>
<td>North Gila</td>
<td>3</td>
<td>Fallow, lettuce</td>
</tr>
<tr>
<td></td>
<td>Dome</td>
<td>9</td>
<td>Fallow</td>
</tr>
<tr>
<td></td>
<td>Tacna West</td>
<td>38</td>
<td>Cotton, lettuce</td>
</tr>
<tr>
<td></td>
<td>Texas Hill</td>
<td>29</td>
<td>Fallow</td>
</tr>
<tr>
<td>March 1997</td>
<td>North Gila</td>
<td>5</td>
<td>Wheat, lettuce, alfalfa</td>
</tr>
<tr>
<td></td>
<td>Dome</td>
<td>7*</td>
<td>Lettuce, crucifer</td>
</tr>
<tr>
<td></td>
<td>Tacna West</td>
<td>22</td>
<td>Fallow</td>
</tr>
<tr>
<td></td>
<td>Texas Hill</td>
<td>35*</td>
<td>Cotton, crucifer, alfalfa</td>
</tr>
</tbody>
</table>

<sup>y</sup> Crops planted in the respective locations classified as low or high.<br>
<sup>*</sup> * = The mean S strain incidence for locations in the set of adjacent fields classified as low is significantly less (t test, P < 0.05) than the mean for locations classified as high. Each low and high value is an average of values from two locations in each of three fields.

Fig. 3. Location of sites in the Texas Hill subregion of Yuma County, AZ, where soil samples have been collected from August 1994 through July 1998 for analysis of the strain composition of *Aspergillus flavus* communities.
(19), data from appropriate seasons were combined: data from July 1995 and 1996 were used to predict July 1997 incidence; data from October 1995 and 1996 were used to predict October 1997 incidence; and data from March 1995, 1996, and 1997 were used to predict March 1998 incidence. Using GeoEAS (16), block kriging generated predicted S strain incidence values for 120 × 120-m grid cells, which were linked to polygon shapefiles in ArcView (version 3.0a; ESRI) for display as surface maps (16). The spatial join tool of ArcView was used to assign predicted values from the polygon shapefile to newly sampled locations, so that at each newly sampled location there was a predicted value interpolated from historical data. The number of locations with observed values less than predicted was counted and tested for statistical significance, using the SAS function PROBBNML, under the null hypothesis that an observed value less than predicted was as likely as an observed value greater than predicted. The kriging standard deviation (KSD) provided a measure of relative uncertainty for the grid cell values calculated by kriging (13,17). Exploratory analyses of correlations between observed and predicted values, as well as visual inspection of the spatial patterns of KSD (17), were used to choose KSD = 15% as a cutoff for accepting predicted values.

### RESULTS

#### Spatial autocorrelation
Variogram modeling is the primary geostatistical tool for the analysis of spatial autocorrelation. Prior to variogram modeling, two other lines of evidence suggested that S strain incidence is a spatially dependent variable: a correlation of S strain incidence at repeatedly sampled locations and a nested analysis of variance. The study of *A. flavus* communities in adjacent fields was based on fields selected because they previously (19) had consistently low S strain incidence compared with fields 1 to 5 km away. The pattern of consistently high and low S strain incidence continued in three of the four areas (Tacna West was the exception) in July 1996, October 1996, and March 1997 despite variability in crops growing in the fields (Table 1). Furthermore, S strain incidence at locations repeatedly sampled during these periods was highly correlated (July 1996 vs. October 1996, \( R = 0.696, P < 0.0001, N = 43 \) locations; July 1996 vs. March 1997, \( R = 0.788, P < 0.0001, N = 40 \) locations; and October 1996 vs. March 1997, \( R = 0.790, P < 0.0001, N = 42 \) locations). This is an indication of spatial autocorrelation at a short distance (within 20 m) across sampling times and independent of crop sequence.

Scatterplot analysis of the repeatedly sampled locations indicates a significant (binomial test, \( P < 0.0001 \)) number (37 of 44) of locations with a lower incidence of the S strain in October 1996 than in July 1996 (Fig. 4A). A similar analysis indicated no significant (binomial test, \( P = 0.32 \)) shift in strain incidence between October 1996 and March 1997 (Fig. 4B). Mean S strain incidence across all locations was 50% in July, 32% in October, and 33% in March.

A nested analysis of variance of the July 1996, October 1996, and March 1997 data shows that most of the total variance can be attributed to the “among locations within fields” and the “among widely separated fields” categories (Table 2). Significantly less (\( P < 0.05 \) by \( t \) test with sampling date as replicates) of the total variance can be attributed to the variance component among adjacent fields (0 to 3%). In other words, the variability in S strain incidence associated with locations in adjacent fields is approximately the same as the

### Table 2. A nested analysis of variance of the spatial structure of *Aspergillus flavus* S strain incidence in Yuma County, AZ.

<table>
<thead>
<tr>
<th>Variance source</th>
<th>Spatial scale (km)</th>
<th>July 1996</th>
<th>October 1996</th>
<th>March 1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among subregions</td>
<td>&gt;20</td>
<td>11</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td>Among widely separated fields</td>
<td>1–5</td>
<td>46</td>
<td>58</td>
<td>49</td>
</tr>
<tr>
<td>Among adjacent fields</td>
<td>0.2–0.6</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Among locations within fields</td>
<td>0.1–0.4</td>
<td>40</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

#### Fig. 5. Variogram of *Aspergillus flavus* S strain incidence based on total samples collected in July 1995, 1996, and 1997. A, There are two components of the spatial structure. The solid line represents a variogram model consisting of a combination of two spherical functions with ranges of 2,500 and 25,000 m. B, The same points and model are shown with the x-axis changed to show detail at the shorter range. Variograms for October and March are similar.
variability within the same field over approximately the same spatial scale of 0.1 to 0.6 km (Table 2 and Fig. 2), despite differences in crops (Table 1). The variances “among widely separated fields” and “among subregions” are in addition to the “among locations within fields” variance. Additions to the total variance associated with an increase in the spatial scale in a nested analysis of variance, as seen here (Table 2), are characteristic of a spatially autocorrelated variable (18).

Variogram analysis of S strain incidence using all locations sampled in July 1995, 1996, and 1997 confirms the spatial structure observed in the nested analysis of variance, because the variogram values for S strain incidence generally increase with lag distance (Fig. 5). The points that are in Figure 5 represent variogram values calculated from the data, whereas the line represents a spherical function (13) whose parameters were selected so that the line “fits” the points (20). The same points shown at two different scales illustrate that there are two components of spatial structure (Fig. 5). The two components can be modeled by spherical functions with ranges of 2,500 and 25,000 m. Variograms for October (short range = 3,500 m and long range = 25,000 m) and March (short range = 2,500 m and long range = 25,000 m) are similar to those for July. The experimentally derived variogram points suggest variability at very short lag distances (<100 m), and the line representing the variogram model reflects this local variability by intersecting the y-axis at a variogram value of 250. This value is referred to as the nugget. The nugget describes spatial variability at a very short range (e.g., the range of repeat sampling at the “same” location).

**Surface interpolation and prediction.** The variogram models were used as input parameters to a kriging program (GeoEAS) for surface interpolation and prediction. Surface maps, generated from S strain incidence data at locations sampled in July 1995 and July 1996 in the North Gila and the Texas Hill areas (Fig. 6), give a prediction of incidence in July 1997 at all points across the interpolated surface. Predicted and observed S strain incidences are highly correlated (Fig. 7A, \( R = 0.85, P = 0.0001, n = 36 \) locations). The slope of the trend line is close to 1 (0.92), indicating that the observed values are, on average, near the expected values, but in only nine locations are the observed values greater than expected (Fig. 7A, to the left of the slope = 1 line). This is significantly less than what should be observed under a statistical hypothesis that observed values are equal to predicted values (binomial test, \( P = 0.002 \)).

The same procedure used to predict October 1997 S strain incidence based on October 1995 and October 1996 incidences gave a different result. Observed and predicted values are less well correlated (Fig. 7B, \( R = 0.51, P < 0.001, n = 50 \) locations), and the slope of the trend line is much less than 1 (0.51). The observed values are less than predicted in October at 40 of the 50 sample locations (binomial test, \( P < 0.001 \)). Similarly, in a high proportion of the locations, observed S strain incidence in March 1998 is less than predicted based on kriging March 1995, 1996, and 1997 S strain incidences (binomial test for 26 of 35 locations, \( P = 0.003 \)). The correlation between observed and predicted is statistically significant (Fig. 7C, \( R = 0.50, P = 0.0021, n = 35 \) locations). The Texas Hill and North Gila areas have a high proportion of locations in which the observed incidence is less than the expected incidence (Fig. 7). In the 2 years prior to October 1997, only 2 of 97 locations (2%) sampled in the Texas Hill area had an S strain incidence of 20% or less. By contrast, 9 of 29 (31%) locations in October 1997, 11 of 38 locations (29%) in March 1998, and 5 of 18 locations in July 1998 (28%) had an S strain incidence of 20% or less. In the four locations in the Texas Hill subregion repeatedly sampled from April 1995 through July 1998, a major reduction in S strain incidence occurred after July 1997 (Fig. 8).

The results of over 3.5 years of sampling in the Texas Hill area are summarized graphically using surface maps (Fig. 9). The maps

![Fig. 6. Surface map of S strain incidence in the Texas Hill subregion of Yuma County, AZ, based on July 1995 and 1996 data. Only the portion of the surface where the kriging standard deviation is less than 15% is shaded. The map shading represents a spatial average, which has a smoothing function. Because local variability is high, S strain incidence at a specific location would vary from the average values shown on the surface.](image-url)
show a generally higher S strain incidence in July (Fig. 9A and D) than in October and March (Fig. 9B, C, E, and F) across most of the area. S strain incidence was particularly low in October 1997 (Fig. 9E) and March 1998 (Fig. 9F). In all six maps, S strain incidence was relatively low around field 11 and high around field 1. Similarly, there was consistency in pattern in the North Gila area (data not shown) in which the surface around field 5 was lowest in S strain incidence. The patterns of S strain incidence from 1995 to 1998 (Fig. 9) can be combined into a single map of four zones characterized as consistently high, intermediate to high, variable but including low, and low (Fig. 10). Retrospectively, sampling sites were assigned to the zone in which they were located and S strain incidence and propagule density were tabulated by zone (Table 3). An analysis of variance of the mean S strain incidence in each zone from April 1995 through July 1998 shows a significant difference among the zones but no significant difference in propagule density (Table 3). Temporal patterns of S strain incidence and propagule density averaged across zones for each sampling date in Texas Hill are correlated ($R = 0.77, P = 0.006$) but not identical (Table 3).

**DISCUSSION**

In our previous study (19), the spatial and temporal patterns of *A. flavus* strain incidence were compared with patterns of propagule density in Yuma County, AZ. Nested analysis of variance was used to analyze the spatial structure of both S strain incidence and *A. flavus* propagule density. For both variables, the largest component of variance occurred among fields within areas at a spatial scale of 1 to 5 km. The prior analysis suggested that crop sequence does not influence S strain incidence, so we suspected that the S strain incidence was dependent on one or more factors that extend beyond field boundaries (19). The sampling of adjacent fields reported here confirmed that suspicion. Less than 3% of the S strain incidence variance occurred among adjacent fields, even though some of the sampled fields were in different crops. This result suggests the observed variance component at 1 to 5 km is due to factors related to the distance between sample locations in a manner that is independent of crop or agronomic practices that may change at field boundaries. This contrasts with *A. flavus* propagule density, which is correlated with crop sequence (19). In other words, spatial autocorrelation of S strain incidence extends beyond field boundaries. This lays the foundation for using the geostatistical tools of variogram analysis and kriging to create regional surface maps of S strain incidence.

Variogram analysis suggests two ranges of spatial structure and these may reflect two distinct sets of factors influencing S strain incidence. Within subregions, variability in climate and soil type may have perennial influence. Between subregions, irrigation water source (in Yuma County, agriculture is totally dependent on irrigation), drainage, and variation in agronomic practices may play roles. Agriculture occupies less than 4% of the land area in Yuma County and there are significant *A. flavus* communities throughout the native habitats (3). Therefore, factors from outside the agricultural fields may have significant influences on S strain incidence for both scales of variability.

Surface interpolation (by kriging) may be particularly valuable in recognizing shifts in regional patterns. Correlations between values of S strain incidence predicted from surface maps and observed values are highly significant ($P < 0.0005$, in all cases), because many locations ($N > 35$) were included. However, the scatterplots
(Figs. 4 and 7) also show that variability is high. Prediction at any one location is not particularly good, but with a large number of points (30 to 40), patterns emerge. Shifts in regional patterns can be detected visually by comparison of surface maps and statistically by correlation analysis and the binomial test of the number of observations that are less than (or greater than) predicted.

In Yuma County, S strain incidence ranges from under 5% to over 90%. The spatial structure observed here might be restricted

Fig. 9. Surface maps of Aspergillus flavus S strain incidence generated by ordinary block kriging of point data from A, July 1995 and 1996; B, October 1995 and 1996; C, March 1995, 1996, and 1997; D, July 1997; E, October 1997; and F, March 1998. The two locations sampled in fields 1 and 11 are indicated on each map as reference points for patterns common to the six maps. Cells with a kriging standard deviation (KSD) less than 15% are displayed with no outline and cells with a KSD ≥15% are displayed with a strongly contrasting outline (indicating greater uncertainty in the estimate of incidence). The map shading represents a spatial average, which has a smoothing function. Because local variability is high, S strain incidence at a specific location would vary from the average values shown on the surface.
to regions where high S strain incidence occurs at least in some fields. There are areas in the United States Cotton Belt where S strain incidence is consistently low (less than 5%) (7; P. J. Cotty, unpublished data). Variogram models appropriate for Yuma County may not be appropriate for low S strain incidence areas, because spatial and temporal variability of S strain incidence would be less. In those areas, variogram models would have a longer range, and fewer locations would be required to characterize the S strain incidence.

Bayman and Cotty (1) observed large shifts in the pattern of A. flavus vegetative compatibility groups within a single field from one season to the next. The data reported here suggest a shift in the fungal community structure extending over the entire Texas Hill study area, which is about 30 km². This area-wide shift may indicate an event that differentially influences the A. flavus strains and, as such, may reflect an aspect of the population biology of A. flavus not previously observed. Weather-related factors may partly explain the unexpected drop in S strain incidence in the Texas Hill area from October 1997 through July 1998. The El Niño weather pattern greatly influenced the 1997 to 1998 period with extended cool weather. Typically, S strain incidence peaks during the warmest periods and rapidly declines during cool periods (19). Changes in temperature patterns over large areas may be expected to cause regional changes in S strain incidence such as that observed.

Our results may be most meaningful in the context of ongoing efforts to use atoxigenic L strain isolates of A. flavus to manage aflatoxin contamination of cotton. Commercial-scale field evaluations (120 to 500 acres per annum) of the use of atoxigenic strains are being undertaken in Arizona under an EPA-approved experimental use registration. The Arizona cotton industry is seeking to expand registration and use of atoxigenic strains in the near future. The economics of treatment improves as the required frequency of treatment declines. Even taking into account the unexpected shift in S strain incidence in the fall of 1997, the relative stability of S

---

**TABLE 3. Retrospective analysis of S strain incidence and Aspergillus flavus propagule density in zones generated to predict S strain incidence in the Texas Hill subregion of Yuma County, AZ**

<table>
<thead>
<tr>
<th>Sampling date</th>
<th>Z1 = Consistently High</th>
<th>Z2 = Intermediate to High</th>
<th>Z3 = Variable to Low</th>
<th>Z4 = Consistently Low</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 1995</td>
<td>70 (4)</td>
<td>34 (2)</td>
<td>46 (2)</td>
<td>50 bc</td>
<td>148</td>
</tr>
<tr>
<td>July 1995</td>
<td>90 (4)</td>
<td>87 (2)</td>
<td>61 (2)</td>
<td>79 a</td>
<td>9,144</td>
</tr>
<tr>
<td>October 1995</td>
<td>73 (4)</td>
<td>62 (2)</td>
<td>51 (2)</td>
<td>62 ab</td>
<td>1,099</td>
</tr>
<tr>
<td>March 1996</td>
<td>76 (4)</td>
<td>78 (2)</td>
<td>28 (2)</td>
<td>61 ab</td>
<td>153</td>
</tr>
<tr>
<td>July 1996</td>
<td>89 (12)</td>
<td>82 (11)</td>
<td>40 (3)</td>
<td>64 ab</td>
<td>3,389</td>
</tr>
<tr>
<td>October 1996</td>
<td>69 (12)</td>
<td>59 (11)</td>
<td>38 (3)</td>
<td>49 bc</td>
<td>404</td>
</tr>
<tr>
<td>March 1997</td>
<td>67 (10)</td>
<td>65 (2)</td>
<td>39 (1)</td>
<td>52 bc</td>
<td>20</td>
</tr>
<tr>
<td>July 1997</td>
<td>78 (12)</td>
<td>80 (13)</td>
<td>42 (2)</td>
<td>64 ab</td>
<td>1,952</td>
</tr>
<tr>
<td>October 1997</td>
<td>46 (6)</td>
<td>32 (18)</td>
<td>45 (6)</td>
<td>32 c</td>
<td>139</td>
</tr>
<tr>
<td>March 1998</td>
<td>50 (12)</td>
<td>31 (26)</td>
<td>45 (4)</td>
<td>35 c</td>
<td>84</td>
</tr>
<tr>
<td>July 1998</td>
<td>52 (5)</td>
<td>36 (7)</td>
<td>70 (2)</td>
<td>44 bc</td>
<td>216</td>
</tr>
<tr>
<td>Mean</td>
<td>69 a</td>
<td>59 a</td>
<td>46 b</td>
<td>35 b</td>
<td>1,653</td>
</tr>
</tbody>
</table>

---

* Locations were assigned to zones retrospectively based on an analysis of surface maps of S strain incidence from April 1995 through March 1998. Zones are illustrated in Figure 10. Zone 1 = consistently high, zone 2 = intermediate to high, zone 3 = variable to low, and zone 4 = consistently low.

* The number of locations for each date and zone is in parentheses.

* Means followed by the same letter are not significantly different (*P* < 0.05) based on Duncan’s multiple range test using the interaction of zones by sampling date as the error term.
strain incidence from year to year suggests it may not be necessary to re-treat an area with atoxigenic strains each year.

We would expect the spatial structure to change under treatment with atoxigenic strains. In particular, one might anticipate the within-field component of variance to decrease relative to the other components in the nested analysis of variance because of the uniform application of a single strain across a field. Also, treatments should decrease S strain incidence overall, and we have observed lower within-field variance associated with lower S strain incidence (Table 2, March and October). Monitoring both maps of S strain incidence and changes in the short-range elements of spatial structure (nugget and within-field variance) should be helpful in evaluating the impact and persistence of atoxigenic L strains. The spatial structure and regional patterns described here provide an important baseline for future assessments.

The low component of variance among adjacent fields (Table 1) and variogram ranges of 2 to 3 km and 20 to 30 km (Fig. 5) suggest a regional form of site-specific management for S strain-induced aflatoxin contamination. Such a program would contrast with typical precision agriculture in that the site management would be based on patches (square kilometers) within a region rather than patches (square meters) within a field. The GIS and geostatistical tools reported here help lay a foundation for an integrated approach to the management of aflatoxin contamination so that multiple tools (e.g., genetically modified insect-resistant cotton, atoxigenic strains, and early harvest) may be used to maximize efficacy.

ACKNOWLEDGMENTS

This work was supported by a grant from the Multicrop Working Group of the Aflatoxin Elimination Program through U.S. Department of Agriculture cooperative agreement #58-6435-4-115. We thank D. E. Myers, Department of Mathematics, University of Arizona, for reviewing drafts of this manuscript.

LITERATURE CITED