Spatial Distribution of Green Mold Foci in 30 Commercial Mushroom Crops

D. J. Royse, Department of Plant Pathology, College of Agricultural Sciences, The Pennsylvania State University, University Park 16802; K. Boomer, Y. Du, and M. Handcock, Department of Statistics, Eberly College of Science, The Pennsylvania State University; P. S. Coles, Giorgi Mushroom Company, Temple, PA 19560; and C. P. Romaine, Department of Plant Pathology, The Pennsylvania State University

ABSTRACT

Statistical analyses were performed on spatial distributions of mushroom green mold foci caused by Trichoderma spp. in 30 standard Pennsylvania doubles (743 m² production surface) selected at random from over 900 total crops mapped. Mapped production houses were divided into four tiers of six beds each with 16 sections per bed (total = 384 sections per double). Each section contained approximately 2 m². Green mold foci were mapped according to presence or absence in each section as they became visible during the course of the mushroom production. There was a trend toward higher disease incidence at the ends of the doubles, although this was not consistent from level to level. Spatial analysis revealed that green mold foci were more likely to occur in neighboring sections along the beds rather than above, below, or across from each other. Cultural practices that were associated with movement along the beds, i.e., nutrient supplementation, spawning, bed tamping, surface covering, etc., were considered the most likely factors influencing the incidence of green mold in spawned compost. Airborne contamination was considered a less likely source of inocula contributing to epidemic development. Sanitation practices that reduce spore loads along the beds are expected to provide the greatest degree of green mold control.

Additional keywords: Trichoderma harzianum, Agaricus bisporus

In the mid-1980s, new, virulent forms of the green mold fungus Trichoderma harzianum were discovered on mushroom (Agaricus bisporus) farms in Ireland. In the early 1990s, similar forms of T. harzianum were encountered on mushroom farms in North America (14,16). Economic losses worldwide from green mold in the last decade have exceeded (U.S.) $30 million (13,14,16).

Recent molecular work has shown that populations of the virulent forms of T. harzianum in Europe (Th2) and North America (Th4) are genetically distinct but closely related (3,14,15). The origins of both virulent populations (Th2 and Th4) are not currently known (3), although evidence to date indicates they are new forms of T. harzianum (13,14). Virulent forms have not been found in collections of T. harzianum associated with mushroom culture prior to 1990 (4).

On-farm sources of inocula that may have fueled the Irish epidemics of Th2 green mold in the mid-1980s were enumerated by Seaby (19). These sources included lagoon water, machinery, yards, floors, walls, wooden pallets, trailers, ladder runs, hand rails, spawn hoppers, shrink wrap used to bale plastic bags, restroom tables, chairs, boots, and clothing. Animal vectors of inocula included scarid flies (Lycoriella multa), pepper mites (Pygmephorus spp.), and mice (Mus musculus). Seaby (19) suggested that circumstantial evidence implicated airborne dust as the main source of contamination of compost or its packing machinery, along with transmission on workers’ clothing, pallets, load covers, and trailers, and vectors such as mites, mushroom flies, and mice.

Most mushroom production in Ireland is by the satellite method (20). Large, central compost producers supply several growers with bagged, spawned compost. These discrete units (20 kg of fresh compost) contrast with the bed method used by many North American mushroom growers. In North America, a typical structure (double; 19.5 m long × 12.2 m wide × 4.3 m high) used to produce mushrooms is composed of four tiered rows made up of six stacked beds (21). Thus, when compared with the Irish system, larger masses of compost are exposed to potential airborne contaminants in mushroom doubles. Disease epidermis may, therefore, develop differently in each system.

Analyses of spatial patterns of disease have been used for many years in attempts to reveal underlying causes of epidemic development in both human and plant populations (5,7,8,10). Spatial autocorrelations of green mold foci in mushroom doubles have not been conducted. The purpose of this work was to use spatial autocorrelation to elucidate factors that might influence patterns of green mold incidence in mushroom doubles.

MATERIALS AND METHODS

Disease assessment. Green mold foci from more than 900 mushroom doubles (743 m²; Giorgi Mushroom Farms, Temple, PA) were mapped throughout the mushroom production cycle. A double consists of four tiers of six beds each. That is, each tier is a stack of beds six units high. Although the compost is distributed throughout a bed, each bed is considered divided into 16 sections (divisions where uprights and joists serve as bed supports). Hence, a typical double consists of 384 sections in a 16 by 4 by 6 array. There is 0.99 m separating each of the four beds and 0.61 m on center between each level. That is, each section is 1.22 m long by 1.64 m wide. Figure 1 shows two tiers each with three beds, with three sections each.

Beginning at time of casing (12 to 14 days after spawning), foci were recorded as they became visible during the course of mushroom production. Sizes of foci were not quantified due to irregularity in shapes; rather, foci were mapped according to the section where they developed, and disease affected in each double. The data available indicated whether green mold was present at crop termination; a spatial-temporal analysis was not conducted.

A total of 30 maps were randomly selected from the set of 900 for spatial analysis. Each map provided data on 384 sections, with the following exceptions. One selected double had 18 sections in a bed, yielding data on 432 sections, and two doubles selected had 15 sections in a bed, yielding data on 360 sections. The disease incidences in the 30 doubles were considered independent events.

Statistical methodology. A rigorous modeling of the spatial process was used to describe the distribution of green mold within a double. Preliminary autocorrelation analysis, based on the work of Cliff and Ord (5), indicated that the distribution was not characterized by complete spatial randomness. Therefore, the current analysis was performed to provide insight into the distribution of green mold.
within a mushroom double. To specify the underlying spatial model, we first defined the spatial terms and then quantified the probability associated with any given realization of the spatial process (6,12).

**General description.** The data were modeled as a regular lattice of point sites (or sections). Each section was referenced with three indices: $i$ indicated distance along a bed, $i = 1, \ldots, 16$; $j$ indicated the bed, $j = 1, \ldots, 4$; and $k$ indicated the level, $k = 1, \ldots, 6$. (During the statistical analysis, $i$ indexed the total number of sections in each bed of each room. One room had 18 sections in a bed, two rooms had 15 sections in a bed, and the remaining 27 rooms had 16 sections.)

The response variable was binary, indicating presence or absence of green mold in each section. Specifically, the response variable was defined as

$$Z_{ijk} = \begin{cases} 1 & \text{if green mold is present at location } (i,j,k) \\ 0 & \text{if green mold is absent at location } (i,j,k) \end{cases}$$

**Neighbors.** In the terminology of spatial analysis, “neighbors” were considered adjacent sections. In the production house double, some sections adjoin (i.e., sections within a bed). Parallel sections in adjacent beds on the same level were neighbors separated by a distance of 0.61 m. Finally, sections at the same position on a bed on adjacent levels were neighbors separated by a vertical distance of 0.61 m. To account for these actual differences in spatial proximity, “first-order” and “second-order” neighbors were defined (Fig. 2).

![Fig. 1. A standard Pennsylvania double (approximately 743 m² mushroom production surface) consists of 24 beds, each with an average of 16 sections. Beds are stacked six levels high. This cross section cutout view shows two tiers, each with three beds, with three-plus sections each. Notice that there is no physical separation between sections within a bed, and there is space between beds and between levels.](image)

![Fig. 2. Definition of first- and second-order neighbors within a mushroom double, which allowed characterization of disease distribution.](image)

Within a bed, sections that were physically adjoined represented the closest possible neighbors and hence were termed first-order neighbors. A location $(i, j, k)$ had at most two first-order neighbors, namely $(i\pm1, j, k)$ and $(i, j\pm1, k)$.

Second-order neighbors consisted of the next nearest neighbors. There were three types of second-order neighbors: within a bed, horizontally between beds (that is, beds on the same level), and vertically between beds (that is, between levels). A reference location had at most six second-order nearest neighbors. Type I (within bed) second-order neighbors were those sections two units from the reference location within the same bed. From a reference location of $(i,j,k)$, type I second-order neighbors were $(i\pm2,j,k)$ and $(i+2,j,k)$. Type II (horizontally between beds) second-order neighbors were located at the same horizontal level in adjacent beds. From a reference location of $(i,j,k)$, type II second-order neighbors were $(i,j\pm1,k)$ and $(i,j+1,k)$. Finally, type III (vertically between beds) second-order neighbors were those sections located vertically above or below each other. From a reference location of $(i,j,k)$, type III second-order neighbors were $(i,j,k\pm1)$ and $(i,j,k\pm1)$.

Sections on the boundary of the double were termed edge units, and the neighbors of such sections were calculated using only the available data. For example, from a reference location of $(1,1,1)$, the following neighbors were defined:

- First order $(2,1,1)$
- Type I second order $(3,1,1)$
- Type II second order $(1,2,1)$
- Type III second order $(1,1,2)$

No attempt was made to model the edge units separately.

**Specification of the Markov random field.** The Markov random field for some phenomena is specified by the joint distribution of the stochastic process $\{Z_{ijk}: i = 1, \ldots, 16; j = 1, \ldots, 4; \text{ and } k = 1, \ldots, 6\}$. Besag (1) explores the definition of a spatial model based on the conditional distribution of $Z_{ijk}$, given the incidence in the surrounding sections. Besag (1) notes that while the classical way to specify the model is through the joint distribution, processes are often naturally generated from conditional probability assumptions. Intuitively, the conditional distribution makes sense since it is natural to assume that the incidence of green mold at some location $Z_{ijk}$ is related to the incidence at the surrounding locations. The Markov random field model proposed by Besag (1) specifies that the conditional distribution of $Z_{ijk}$ depends only on those sections that are first- and second-order neighbors of section $(i,j,k)$. In general, higher orders could be specified if required, but they were not explored in this paper. From the conditional distributions, it was possible to arrive at the joint distribution that defines the Markov field.
In terms of the joint distribution for all sections, define

\[ Q(z) = \log \left( \frac{P(Z = z)}{P(Z = 0)} \right) \]  

(2)

For an individual section, equation 2 becomes

\[ Q(z_{ij}) = \log \left( \frac{P(\text{incidence at } i, j| \text{ given incidence at other sections})}{P(\text{absence at } i, j| \text{ given incidence at other sections})} \right) \]  

(3)

Following Besag's (1) general expansion of \( Q(z) \) and assuming only second-order pair-wise dependence, equation 2 can be expressed in so-called autologistic form:

\[ Q(z) = \sum_{\alpha} \sum_{\beta} \alpha_{ijk} z_{ijk} + \sum_{\alpha} \sum_{\beta} \beta_{ik} z_{ik} z_{j} \]  

(4)

In equation 4, the parameter \( \alpha_{ijk} \) represented the intensity of the process, and \( \beta_{ik} \)

---

Fig. 3. The sample proportion of sections with green mold infections was related to incidence in neighboring sections. For each section, the sample proportion was defined as the number of doubles in which that section was diseased divided by the total number of doubles sampled (30).
represented the spatial dependence. We further enumerated the spatial dependence using the definitions of first- and second-order neighbors. Let \( \beta^{(i)} \) denote the first-order effect and \( \delta^{(i)} \), \( \delta^{(2)} \), and \( \delta^{(3)} \) denote, respectively, the type I, II, and III second-order effects. Hence, we arrive at \( \beta = (\beta^{(1)}, \delta^{(1)}, \delta^{(2)}, \delta^{(3)}) \). Then, a simpler version of equation 4 is given by

\[
Q(z) = \sum_{i=1}^{16} \sum_{j=1}^{4} \left[ \alpha z_{ij} + \beta^{(1)}z_{ij} + \delta^{(1)}z_{ij} + \delta^{(2)}z_{ij} + \delta^{(3)}z_{ij} \right]
\]

where,

\[
x^{(1)j} = z_{i,j-1} + z_{i+1,j}
\]

\[
x^{(2)j} = z_{i-1,j} + z_{i+1,j}
\]

\[
x^{(3)j} = z_{i,j-1} + z_{i,j+1}
\]

\[
x^{(4)j} = z_{i-1,j} + z_{i+1,j-1}
\]

Based on the neighborhood structure, the conditional density for the underlying model is then

\[
\text{logit} \{ P[Z = z_{ij} \mid z_{ij}] \} = \alpha z_{ij} + \beta^{(1)}z_{ij} + \delta^{(1)}z_{ij} + \delta^{(2)}z_{ij} + \delta^{(3)}z_{ij}
\]

The maximum likelihood estimates of the parameters could be obtained using logistic regression, if the observations are conditionally independent. However, a primary objective in this analysis was to characterize the spatial dependencies. Besag (2) described a pseudolikelihood model that treated the likelihood function as the product of independent conditional probabilities. As a result, the parameters in equation 6 can be solved using standard logistic regression techniques. However, the standard errors associated with these parameters were not the estimated logistic regression standard errors, as those presume independence.

Pseudolikelihood standard errors were obtained using a parametric bootstrap procedure (9,11). A sample of 3,000 lattices was generated using a Gibbs sampling procedure, based on the estimated logistic parameters, and pseudolikelihood estimates were calculated for each lattice. This produced a sampling distribution for the estimated parameters from which we can make an assessment of the significance for each effect.

**Interpretation of parameters.** An advantage of the logistic regression model of equation 6 is that the estimated parameters have a direct interpretation for this spatial design. The parameters represent log-odds ratios; the exponential of the parameter represents the odds ratio. The odds ratio compares the odds of a section becoming infected when a neighbor is infected to the odds of a section becoming infected when a neighbor is not infected. An odds ratio greater than 1 indicates that the section with infected neighbors is more likely to be infected than a section with healthy neighbors. Hence, for each defined neighborhood structure, we were able to assess the association between neighboring an infected section and neighboring a healthy section.

For each estimated parameter, a standard error and 95% confidence interval were obtained from the parametric bootstrap procedure. The confidence intervals represent the range of estimated parameters for 95% of the 3,000 sampled lattices. A confidence interval that does not contain the value zero provides evidence that the parameter was significantly different from zero. As the estimated parameters represent log-odds ratios, an estimated parameter of zero corresponds to an odds ratio of 1, which suggests no association between disease incidence and disease incidence of neighbors.

**RESULTS**

Combined presence–absence data from the 30 mushroom doubles were used to calculate sample portions of sections with green mold infections (Fig. 3). Sample proportions ranged from 0.033 to 0.3, indicating that the sections with the most common infestations exhibited green mold in approximately one-third of the doubles. There was a trend toward higher sample proportions at the ends of the doubles, although this was not consistent from level to level. For example, high-numbered sections, such as sections 13 to 15, had a higher proportion of sections with green mold foci than mid- to low-numbered sections, such as sections 3 to 8. Further along a bed, a section tended to have a sample proportion of incidence similar to that of its immediate neighbors. On level 4, for example, several adjoining sections in a bed exhibited similar low proportions of disease.

To assess the potential for a section effect within the data, equation 6 was modified...
ified to include indicator variables for each section number. The results of this modified analysis indicated that the position of the section along the bed did have an effect on the odds ratio. Figure 4 plots the mean and the associated 95% confidence interval of the log-odds of infestation for each section number. Sections on the ends of the beds (i.e., sections 1, 2, and 15 and higher) have a significantly higher log-odds ratio of infestation than do sections in the middle of the bed.

The statistical analysis (Table 1) specifically assessed the association of disease in neighboring sections. Odds ratios indicated that disease incidences in neighboring sections were significantly associated with the incidence of disease in a given section (Table 1). Adjacent neighbors (first-order neighbors) had the strongest association for disease incidence. The odds of disease at a section were over four times higher if an adjacent neighbor contained disease. The effect attributed to neighbors in the same section in adjacent beds (type II second-order neighbors) was similar to the effect seen between neighbors in the same section but in immediately higher or lower beds (type III second-order neighbors). The p-value for the hypothesis test of the equality of these effects was 0.909. This suggests that a horizontal distance across open space of 0.99 m is similar in disease association to a vertical distance between beds of 0.61 m. A section was more than two times more likely to contain disease if a type II or III second-order neighbor was diseased. Thus, disease distribution along the beds had the strongest association for the neighbor types tested.

DISCUSSION

Spatial autocorrelation of green mold foci from 30 Pennsylvania double mushroom crops revealed a nonrandom distribution pattern. Foci were more likely to occur in aggregated patterns in adjacent sections along the beds rather than in adjacent sections above or below the beds. This distribution pattern supports the contention that transmission of inocula on workers’ clothing and hands, spawning machinery, tamping equipment, and bed covers, which are factors associated with along-the-bed operations, are the most important variables in disease development.

Strong autocorrelations of disease foci along the bed do not support the circumstantial evidence of Seaby (19) that airborne contamination is a substantial factor contributing to disease development. However, spatial autocorrelations of disease development have not been conducted on Irish mushroom farms, so it is not possible to compare disease development patterns with Pennsylvania farms.

In a study of six mushroom farms in Canada, Rinker et al. (18) found that inside locations on farms with the highest incidence (35.7%) of green mold contamination were the common areas where personnel cross paths as they move from one activity to another (lunchrooms, washrooms, telephones, etc.). They also found that 15% of the samples collected from the spawn-fill areas and 17.7% of the samples from the tunnel pullers, head fillers, spawning machines, front end loaders, casing mixers, and transfer trucks contained Th4.

Fall-out capture (on nutrient agar plates) studies of airborne Th4 by Rinker (17) at the Horticultural Research Institute of Ontario have shown that few, if any, spores are carried in the air. Due to the sticky nature of Trichoderma spp. spores, movement (transport) is possible via dead mites, dry mycelium, flies, water droplets, or contaminated dust, and thus the potential for airborne inocula to contaminate freshly spawned compost is present. However, considering the spatial patterns of disease development reported in this study and the distribution of Th4 inside mushroom farms in Canada, it is much more likely that inocula are spread via workers and contaminated equipment rather than through airborne sources. Strict sanitation and hygiene practiced prior to and during spawning should provide the greatest control of green mold.

ACKNOWLEDGMENTS

We thank Janice Derr, Department of Statistics, The Pennsylvania State University, for assistance in statistical analysis. We also thank Cory Watson for data entry and technical assistance.

LITERATURE CITED


Table 1. Parameter estimates and odds ratios calculations from the logistic regression of green mold, Trichoderma harzianum (Th4), from 30 commercial mushroom doubles in Pennsylvania

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>Estimated parameter</th>
<th>Standard error</th>
<th>95% confidence intervals</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^{(1)}$</td>
<td>First-order neighbors; adjacent sections</td>
<td>1.55</td>
<td>0.101</td>
<td>(1.37, 1.76)</td>
<td>4.71</td>
</tr>
<tr>
<td>$\delta^{(1)}$</td>
<td>Type I second-order neighbors; sections within a bed separated by two sections</td>
<td>0.52</td>
<td>0.123</td>
<td>(0.28, 0.77)</td>
<td>1.68</td>
</tr>
<tr>
<td>$\delta^{(2)}$</td>
<td>Type II second-order neighbors; sections at the same position and level in adjacent beds</td>
<td>0.75</td>
<td>0.117</td>
<td>(0.50, 0.95)</td>
<td>2.12</td>
</tr>
<tr>
<td>$\delta^{(3)}$</td>
<td>Type III second-order neighbors; sections at the same position on the bed and at different levels</td>
<td>0.73</td>
<td>0.114</td>
<td>(0.47, 0.92)</td>
<td>2.08</td>
</tr>
</tbody>
</table>

* Standard errors and 95% confidence intervals were calculated from a parametric bootstrap (9). The p-values for these parameters are less than 0.001. The results indicate that disease incidence in neighbors affects the probability of disease in a given section.

