Modified Hot-Spot analysis for spatio-temporal analysis: a case study of the leaf-roll virus expansion in vineyards [1]

Yafit Cohen 1
Agricultural Research Organization, Israel, Yafitush@volcani.agri.gov.il

Rakefet SHARON, Tamar SOKOLSKY*, Tirza ZAHAVI 2
Northern R&D, Israel
*The Robert H. Smith Faculty of Agriculture, Food and environment

Abstract: Given a set of geo-referenced data points, the Getis–Ord Gi* statistic identifies hot-spots of points with values higher in magnitude than one might expect by a random chance. This tool works by looking at each data feature and its neighboring features in comparison to the overall spatial distribution of the phenomenon explored. If the difference between the local sum for a feature and its neighbors is highly larger than expected (the overall sum) a hot-spot is accepted. Leaf-roll virus (LRV) in vineyards appears in clusters which expand from year to year when no pest control is carried out. Exploring the spatio-temporal expansion of hot-spots of the LRV is limited with the Gi* statistics since relative hot-spots are accepted according to the infestation level of a specific year. A modified Gi* was developed which identifies year-to-year hot-spots which are relative to a year of reference. LRV symptoms were mapped yearly in a vineyard from 2005 to 2010. The Gi* indicated for a northern hot-spot only in 2007 and a southern one in 2009. Using the modified Gi* with 2005 as a year of reference, the northern and the southern clusters were identified in 2006 and 2007, respectively.

Keywords: Getis–Ord Gi*, Leaf-roll virus, Spatio-temporal dynamics

1. Introduction

Leafroll is one of the most widespread viral diseases of grapevine. Leafroll disease is an economically important graft-transmissible disease of grapevines and occurs in all grapevine-growing countries. Although grapevine leafroll virus (LRV) can affect the growth, development, longevity and yield of the vines, its most serious effect is on lowering the sugar content and raising the acidity of must. In the field, the spread of LRV associated with particular insect vectors has been reported in several countries, including Israel (Cabaleiro and Segura, 2006; Tanne et al., 1989). In vineyards with available virus inoculum and mealybugs present, LRV spreads quite quickly from vine to vine (Cabaleiro et al., 2008). Study of the spatio-temporal dynamics of the leafroll-infected vines in the vineyard scale may be helpful to determine whether or not spread

[1] Chief Scientist of the Ministry of Agriculture, Israel, research no 132-1502-09
.

of the viruses is occurring, and the best control measures to take. The spreading via the mealybugs creates clusters of infested vines which expanding from year to year. Local spatial statistics may assist with the identification of the infestation clusters. Local spatial statistics identifies those clusters with values higher in magnitude than is expected to be found by random chance. The Getis–Ord Gi* hotspot cluster statistic is one of the many possible approaches used for local spatial analysis (Getis and Ord, 1996). The Gi* statistic measures the degree of spatial clustering of a local sample and how different it is from the expected value which is the mean of the whole data set. Study of the annual expansion of the LRV can use the annual maps of hot spots and follow their expansion. Yet, since the Gi* statistic is a relative measure to the overall infestation level mean in a particular year and since the mean infestation increases annually, the discovery of hot spots is limited. The objective of this paper is to describe a modification of the Gi* statistic which enables a spatio-temporal analysis of the LRV hot-spots expansion.

2. Materials and Methods

The Gi* statistic

The Gi* statistic measures the degree of spatial clustering of a local sample and how different it is from the expected value (Equation 1). It is calculated as the sum of the differences between values in the local sample and the mean, and is standardized as a z-score with a mean of zero and a standard deviation of 1:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}^*}{s^*\sqrt{\frac{(nS_{1i}^*) - W_i^{*2}}{n-1}}}$$

Equation 1:

where *i* is the centre of the local neighborhood, *d* is the lag distance (radius), w_{ij} is the weight for neighbor *j* from location *i*, *n* is the number of samples in the data set, W_i* is the sum of the weights, S_{1i}* is the number of samples within *d* of the central location, x* is the mean of the whole data set, and *s** is the standard deviation of the whole data set. The Gi* statistic is two-tailed, so a score of ±2 represent strong clustering, as 95% of the data under a normal distribution should be within 2 standard deviations of the mean. Values between ±2 may be interpreted as weakly clustered, with values being less than 2 standard deviations away from what one would expect if there were no spatial clustering (Laffan, 2006). While positive values of Gi* represent clusters that are, on average, greater than the mean (Hot-spots) the negative values represent clusters that are less than the mean (Cold-spots). The Gi* statistic is a relative measure and the existence of hot spots of LRV infested vines is highly depended on the overall mean. Since LRV is expanding annually the overall mean increases and local expansions of infestation might not be observed. This attribute limits the use of the Gi* statistic for spatio-temporal analysis of the LRV expansion. A modified Gi* is suggested which calculates Gi* using a pre-defined overall mean (*x**) and standard deviation of (*s**) infestation which function as a common baseline. In this way local expansion of hot spots may be observed.

To explore the potential in using the modified Gi* (*m*Gi*) real data of LRV infested vines were used. Ten rows in a vineyard in the Golan Heights in Israel were monitored for LRV symptoms from 2005 to 2010 (overall of 1142 vines). The study area was divided into 4X4 meters cells (3-4 vines). Each cell was set with the number of infested vines inside it in each year. Gi* statistic was calculated for a radius of 24 meters (d in equation 1) for each year allowing at least 30 data features which are required for a valid analysis (Getis and Ord, 1996). No weights were set to the data. For the calculation of mGi* the overall mean (x*) which is the proportion of the infested vines the standard deviation (s*) of the year 2005 were used. The calculations were made in Matlab, transferred into shapfiles and mapped in ArcGIS 9.3.1.

3. Results

Table 1 shows the average and standard deviation of LRV infestation levels in the years 2005-2010. The infestation level increased over the years. A notable increase was between the years 2008 and 2009 and between 2009 and 2010.

Year	Infestation mean	STD
2005	0.162	0.421
2006	0.219	0.492
2007	0.266	0.592
2008	0.320	0.694
2009	0.563	0.964
2010	0.823	1.159

Table 1: The mean and standard deviation of infestation levels in the years 2005-2010.

Figure 1 presents maps of number of LRV infested vines in 2005, 2007 and 2009. Visual interpretation of the maps indicates for clustered infestation distribution. There was a group of infested vines in the northern part of the vineyard which expanded along the years. In 2009 another distinctive group of infested vines was located in the southern part.

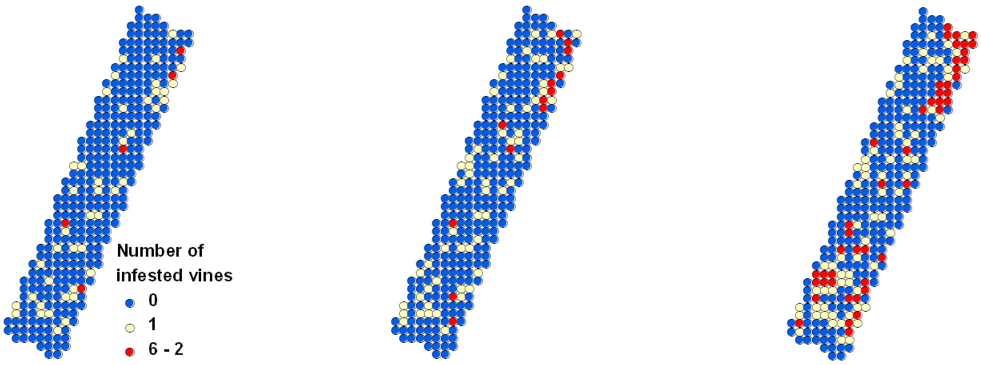

Figure 1: Maps of infested vines in 2005 (left), 2007 (center), and 2009 (right)

Figure 2 presents maps of hot and cold spots in 2005, 2007 and 2009 using the Gi*. The hot spot analysis using the Gi* generally agrees with the visual analysis. Nonetheless, an increase in cold spots was also observed and this is despite the fact that the average infestation level was doubled from 2007 and 2009 (Table 1). Additionally, the southern hot spot in 2009 is relatively small in comparison with the infestation map of 2009. The respective maps created based on the *m*Gi* calculations (Figure 3) shows the advantages and weakness of the *m*Gi*. In 2007 the *m*Gi* discovered a much larger cluster in the northern part on comparison with the Gi*. The *m*Gi* also discovered the southern cluster already in 2007 while by the Gi* this cluster was discovered only in 2009. On the other hand, in 2009 almost all the vineyard was mapped as hot spot.

4. Concluding remarks

The *m*Gi* may be used for spatio-temporal study of LRV expansion but better reference values for its calculations need to be defined.

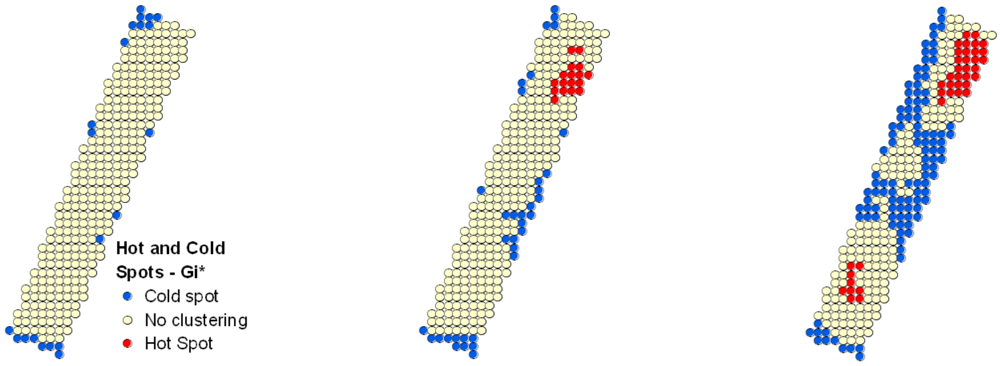

Figure 2: Maps of hot and cold spots using Gi* in 2005 (left), 2007 (center), and 2009 (right)

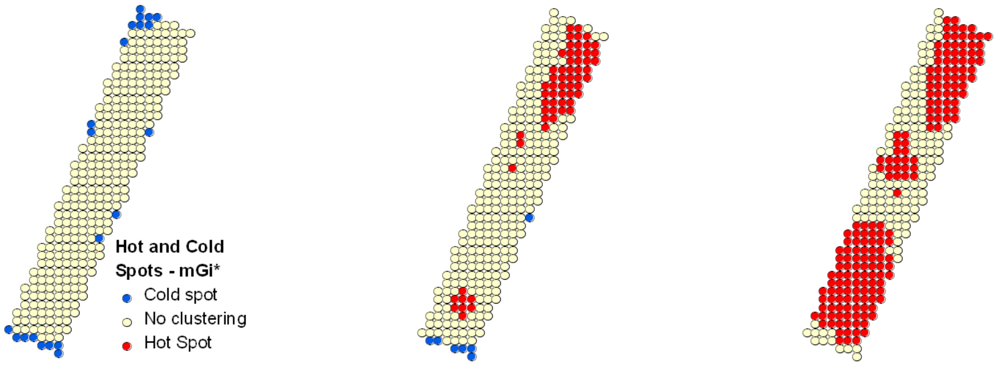

Figure 3: Maps of hot and cold spots using *m*Gi* in 2005 (left), 2007 (center), and 2009 (right)

References

Cabaleiro, C., Couceiro, C., Pereira, S., Cid, M., Barrasa, M., and Segura, A. (2008). Spatial analysis of epidemics of Grapevine leafroll associated virus-3. *European Journal of Plant Pathology* **121**, 121-130.

Cabaleiro, C., and Segura, A. (2006). Temporal Analysis of Grapevine leafroll associated virus 3 Epidemics. *European Journal of Plant Pathology* **114**, 441-446.

Getis, A., and Ord, J. (1996). Local spatial statistics: an overview. *In* "Spatial Analysis: Modeling in a GIS environment " (P. Longley and M. Batty, eds.), pp. 269-294. GeoInformation International, Cambridge.

Laffan, S. W. (2006). Assessing regional scale weed distributions, with an Australian example using Nassella trichotoma. *Weed Research* **46**, 194-206.

Tanne, E., Ben-Dov, Y., and Raccah, B. (1989). Transmission of closterolike particles by mealybugs (Pseudococcidae) in Israel. *In* "Proceedings 9th Meeting ICVG", pp. 71-73, Kiryat Anavim.