STCLASS—Spatiotemporal Distance Class Analysis Software for the Personal Computer

Scot C. Nelson, Assistant Professor, University of Hawaii at Manoa, Department of Plant Pathology, Honolulu 96822

ABSTRACT

Nelson, S. C. 1995. STCLASS—Spatiotemporal distance class analysis software for the personal computer. Plant Dis. 79:643-648.

A software program (STCLASS) for DOS-based personal computers was developed to perform spatiotemporal distance class analysis of intensively mapped binary data. The program can be used to detect and quantify aspects of nonrandom patterns of disease increase in regularly spaced plant populations. Mapped disease incidence evaluations from two disease assessment dates (t_1 and t_2) are obtained for comparison with simulated (expected) data. Expected (random) maps for t_2 are generated by retaining diseased plants from t_1 in the exact spatial locations in which they were observed and by assigning the number of newly diseased plants observed at t_2 to random spatial positions among the nondiseased plant population observed at t_1. Distance class analysis is used to compare the expected pattern(s) to the observed pattern at t_2. Program output includes a map of observed data and $[X,Y]$ distance class matrices in numerical and graphical format. The distance class matrix consists of observed and expected standardized count frequencies for each $[X,Y]$ distance class, the level of significance, and 95% upper and lower confidence intervals on significance levels. The program also can perform two-dimensional distance class analysis and can superimpose distance class matrices from two-dimensional and spatiotemporal distance class analyses of the same data set. The program alerts the user to nonrandom patterns of disease increase and edge effects. The software and a detailed user's manual are available free from the author.

Computer-based technologies have become essential to botanical epidemiology. They facilitate data analysis and the simultaneous development of theory and application. From preplant prediction to postharvest repression of disease, computer programs and models have been used to explore, explain, and predict temporal and spatial aspects of epidemics. Coincident with the development of new theory and computer applications has been an enhanced understanding of epidemics.

An array of computer programs has been generated to study spatial patterns of disease in plant populations. Most general-purpose statistical software (e.g., 8,14) facilitate spatial data mapping and relatively routine calculations (e.g., indices of dispersion). Other programs were developed and distributed to perform more sophisticated spatial data analyses. They include programs for fitting discrete probability distributions (2,6) and for performing runs analysis (7,8), spatial lag autocorrelation analysis (3), spatiotemporal autocorrelation analysis (13), geostatistical analysis (1,15), and two-dimensional distance class analysis (12). The methods vary in sophistication, data requirements, use, interpretation, and application. Development of software for personal computers has allowed these and other forms of analysis to be used by a relatively large number of investigators in diverse pathosystems.

Recently, spatiotemporal distance class analysis was introduced as a method to detect and quantify attributes of nonrandom disease increase in plant populations (9,11). Intensively mapped binary data (presence/absence of disease) were used in computer simulations to compare observed with randomly generated (expected) data via distance class analysis. The method was an extension of Gray's two-dimensional distance class analysis (4,5) and represented a relatively robust method for spatiotemporal analysis of plant disease incidence data. The theories of spatiotemporal distance class analysis and of its relationship to other forms of spatiotemporal analysis were described (9,11).

A computer program (STCLASS) was written to allow more individuals to use spatiotemporal distance class analysis. The purpose of this article is to describe program data requirements and input/output, to outline the program control options available to the user, to present examples and guidelines on the use of the program, and to assist with interpretation of program output.

This work was supported by the University of Hawaii Office of Research Administration and the USDA Special Grants Program for Tropical and Subtropical Agriculture Research.

Corresponding author: Scot C. Nelson
E-mail: snelson@uhunix.uhcc.hawaii.edu

Accepted for publication 20 February 1995.

Program Language and Processing

Program language is Microsoft QuickBASIC. The STCLASS program requires DOS 2.0 or higher and will run with 8088/80286 and higher microprocessors. Extended memory and a math coprocessor are not necessary. Processing time is dependent on microprocessor type and computer clock speed.

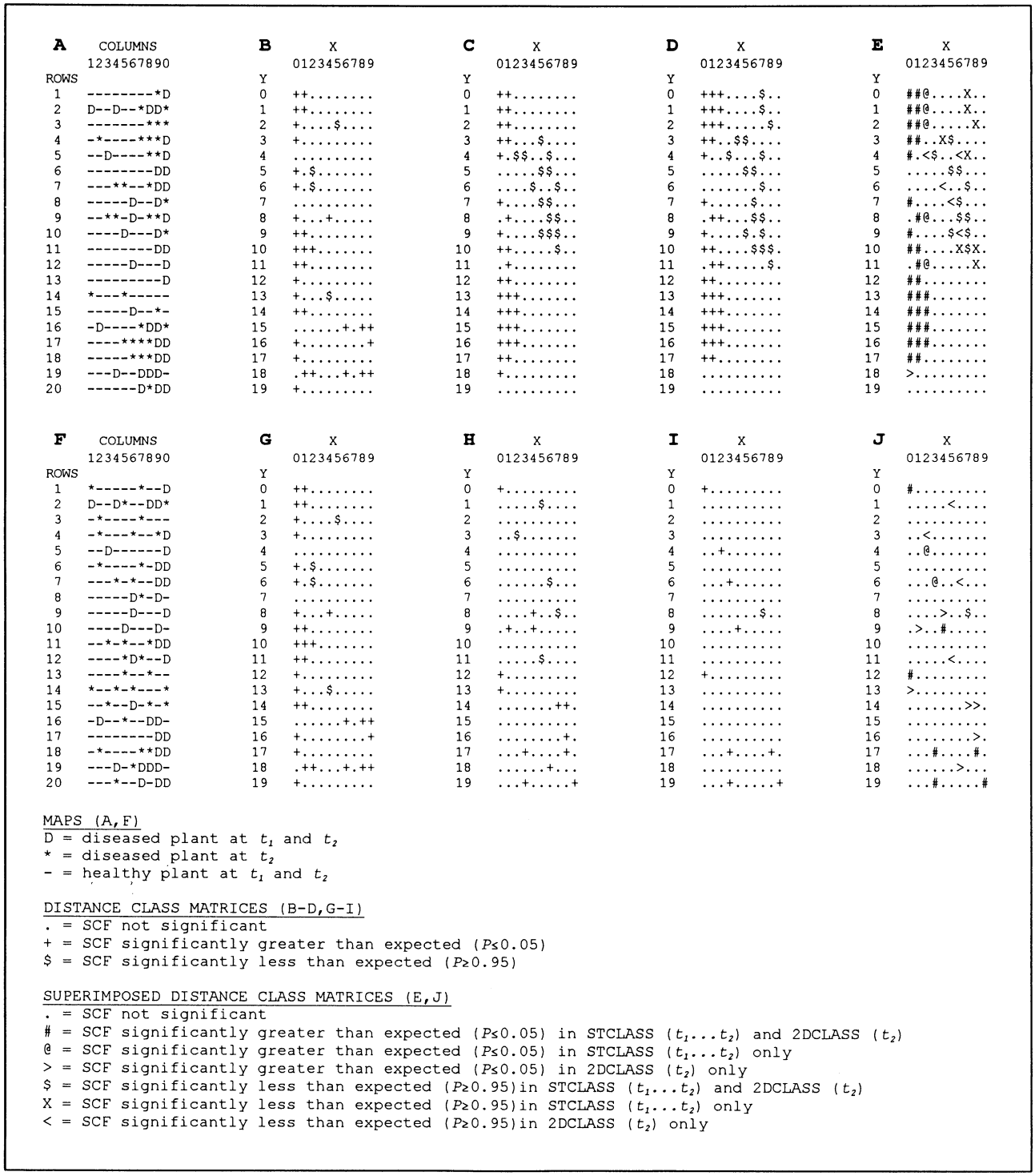
Data Sets

Input data sets are ASCII text. They are comprised of several lines of header information followed by several columns of disease-incidence data. The first two columns contain plant locations within the observed lattice, which are specified in the data set by their X (i.e., row number) and Y (i.e., column number) coordinate values. Successive data set columns contain disease values (1 = healthy, 2 = diseased, 3 = missing value or dead plant) for each plant or quadrat on each disease-assessment date. Data from an unlimited number of disease-assessment dates can be stored in a single data file. With the current version of the software, plot layouts with up to 1,600 lattice positions (e.g., quadrats, plants, positions) may be analyzed. Input data are stored by the program in a row × column matrix for mapping and for pattern comparison with simulated data sets.

Program Structure and Output

Control of program execution resides in a single main routine that calls up to nine subroutines. Among them are subroutines that produce maps of observed data, perform spatiotemporal distance class analysis, print output and information on screen, and send output to a printer. Error checking routines verify input data validity, data file format, printer paper and power supply, and program memory requirements based on data set size. The program alerts the user to data sets that do not meet recommended analytical criteria (e.g., for percent disease incidence, proportion of missing values).

STCLASS program output consists of graphical, textual, and numerical information. First, a map of observed data is produced that shows the spatial locations of healthy, diseased, and missing sampling units on two assessment dates. The map is followed by a set of simple statistics and a matrix of numerical output known as the distance class matrix. The distance class matrix contains information and statistics derived from a comparison of observed and simulated data sets. The distance class matrix comprises observed and expected standardized count frequency (SCF) values for each $[X,Y]$ distance class and a set of associated statistics (significance levels and confidence limits). The distance class matrix is followed by textual output from subroutines that identify strength of nonrandomness and significance of edge effects. A graphical summary of the distance class matrix is produced for user convenience. In addition, the user may select the option of performing two-dimensional distance class analysis (2DCLASS) for the second disease-assessment date. In that case, program output includes numerical



A COLUMNS 1234567890

ROWS	
1	--------*D
2	D--D--*DD*
3	-------***
4	-*----***D
5	--D----**D
6	--------DD
7	---**--*DD
8	-----D--D*
9	--**-D-**D
10	----D---D*
11	--------DD
12	-----D---D
13	---------D
14	*---*-----
15	-----D--*-
16	-D----*DD*
17	----****DD
18	-----***DD
19	---D--DDD-
20	------D*DD

B X 0123456789

Y	
0	++.....
1	++.....
2	+....$....
3	+.....
4	.....
5	+.$.....
6	+.$.....
7	.....
8	+...+.....
9	++.....
10	+++.....
11	++.....
12	+.....
13	+...$.....
14	++.....
15	......+.++
16	+.....+
17	+.....
18	.++...+.++
19	+.....

C X 0123456789

Y	
0	++.....
1	++.....
2	++.....
3	++...$....
4	+.$$..$...
5	.....$$...
6	....$..$..
7	+....$$...
8	.+....$$..
9	+....$$$..
10	++.....$..
11	.+.....
12	++.....
13	+++.....
14	+++.....
15	+++.....
16	+++.....
17	++.....
18	+.....
19	.....

D X 0123456789

Y	
0	+++....$..
1	+++....$..
2	+++.....$.
3	++..$$....
4	+..$...$..
5	.....$$...
6	.....$..$..
7	+.....$...
8	.++...$$..
9	+....$.$..
10	++....$$$.
11	.++.....$.
12	++.....
13	+++.....
14	+++.....
15	+++.....
16	+++.....
17	++.....
18	.....
19	.....

E X 0123456789

Y	
0	##@....X..
1	##@....X..
2	##@.....X.
3	##..X$....
4	#.<$..<X..
5	.....$$...
6	....<..$..
7	#....<$...
8	.#@...$$..
9	#....$<$..
10	##....X$X.
11	.#@.....X.
12	##.....
13	###.....
14	###.....
15	###.....
16	###.....
17	##.....
18	>.....
19	.....

F COLUMNS 1234567890

ROWS	
1	*-----*--D
2	D--D*--DD*
3	-*----*---
4	-*---*--*D
5	--D------D
6	-*----*-DD
7	---*-*--DD
8	-----D*-D-
9	-----D---D
10	----D---D-
11	--*-*--*DD
12	----*D*--D
13	----*--*--
14	*--*-*---*
15	--*--D-*-*
16	-D--*--DD-
17	--------DD
18	-*----**DD
19	---D-*DDD-
20	---*--D-DD

G X 0123456789

Y	
0	++.....
1	++.....
2	+....$....
3	+.....
4	.....
5	+.$.....
6	+.$.....
7	.....
8	+...+.....
9	++.....
10	+++.....
11	++.....
12	+.....
13	+...$.....
14	++.....
15	......+.++
16	+.....+
17	+.....
18	.++...+.++
19	+.....

H X 0123456789

Y	
0	+.....
1	.....$....
2	.....
3	..$.....
4	.....
5	.....
6	......$...
7	.....
8	....+..$..
9	.+..+.....
10	.....
11	.....$....
12	+.....
13	+.....
14	.....++.
15	.....
16	.....+.
17	...+....+.
18	......+...
19	...+.....+

I X 0123456789

Y	
0	+.....
1	.....
2	.....
3	.....
4	.....
5	.....
6	...+.....
7	.....
8	.....$..
9	....+.....
10	.....
11	.....
12	+.....
13	.....
14	.....
15	.....
16	.....
17	...+....+.
18	.....
19	...+.....+

J X 0123456789

Y	
0	#.....
1	.....<....
2	.....
3	..<.....
4	..@.....
5	.....
6	...@..<...
7	.....
8	....>..$..
9	.>..#.....
10	.....
11	.....<....
12	#.....
13	>.....
14	.....>>.
15	.....
16	.....>.
17	...#....#.
18	......>...
19	...#.....#

MAPS (A,F)
D = diseased plant at t_1 and t_2
* = diseased plant at t_2
- = healthy plant at t_1 and t_2

DISTANCE CLASS MATRICES (B-D,G-I)
. = SCF not significant
+ = SCF significantly greater than expected ($P \leq 0.05$)
$ = SCF significantly less than expected ($P \geq 0.95$)

SUPERIMPOSED DISTANCE CLASS MATRICES (E,J)
. = SCF not significant
# = SCF significantly greater than expected ($P \leq 0.05$) in STCLASS ($t_1 \ldots t_2$) and 2DCLASS (t_2)
@ = SCF significantly greater than expected ($P \leq 0.05$) in STCLASS ($t_1 \ldots t_2$) only
> = SCF significantly greater than expected ($P \leq 0.05$) in 2DCLASS (t_2) only
$ = SCF significantly less than expected ($P \geq 0.95$) in STCLASS ($t_1 \ldots t_2$) and 2DCLASS (t_2)
X = SCF significantly less than expected ($P \geq 0.95$) in STCLASS ($t_1 \ldots t_2$) only
< = SCF significantly less than expected ($P \geq 0.95$) in 2DCLASS (t_2) only

Fig. 1. Spatiotemporal distance class analysis (STCLASS) graphical program output from two hypothetical examples of disease increase (nonrandom increase in **A–E** and random increase in **F–J**) between two disease assessment dates (t_1 and t_2) in a 20 row × 10 column plant lattice. **(A, F)** Maps of nonrandom and random disease increase, respectively. **(B, C)** Distance class matrices from two-dimensional distance class analysis (2DCLASS) of data in **(A)** at t_1 and t_2, respectively. **(G, H)** Distance class matrices from 2DCLASS analysis of data in F at t_1 and t_2, respectively. **(D, I)** Distance class matrices from STCLASS analysis of data in **(A)** and **(F)**, respectively, for disease increase between t_1 and t_2. **(E, J)** Distance class matrices showing superimposed matrices from **(C)** and **(D)** and from **(H)** and **(I)**, respectively.

and graphical distance class matrices (from 2DCLASS analysis). Additionally, a graphical presentation of superimposed distance class matrices from 2DCLASS and STCLASS analyses is produced.

Two hypothetical examples were generated to demonstrate STCLASS program output and interpretation. The first example comprised analysis of disease increase between disease assessments in a nonrandom and a random scenario. The second example comprised a disease-increase scenario that was characterized by radial expansion of a single disease focus.

Example one. Two data sets were generated via computer simulation to illustrate nonrandom (Fig. 1A) and random (Fig. 1F) disease increase between assessment dates. For each scenario, disease assessments were made on two dates (t_1 and t_2) for 200 sampling units in a 20 row × 10 column lattice. The sampling units could represent either contiguous quadrats or individual plants. Both patterns of disease increase evolved from an identical starting condition of nonrandomness of disease occurrence at t_1.

STCLASS program output for the nonrandom and random scenarios included maps of the input data (Fig. 1A and F, respectively), simple statistics (e.g., Table 1), and distance class matrices in numerical format (e.g., Table 1) and graphical format (Fig. 1B–E and G–J, respectively). Distance class matrices in graphical format produced by the program included 2DCLASS analysis of data at t_2 for the nonrandom and random scenarios (Fig. 1B–C and G–H, respectively), STCLASS analysis of disease increase between t_1 and t_2 (Fig. 1D and I, respectively), and superimposed distance class matrices from 2DCLASS (t_2) and STCLASS analyses (Fig. 1E and J, respectively). The distance class matrices from 2DCLASS analysis of data at t_1 for both scenarios are presented here for convenience and are not output normally during STCLASS analysis (Fig. 1B and G). The numerical program output consisted of a distance class matrix of SCF values from the observed pattern and expected patterns, and significance levels and confidence limits for each [X,Y] distance class (Table 1). In the numerical and graphical representations of the distance class matrices, significant SCF values were signified by the symbols + (significantly greater than expected, $P \leq 0.05$) and $ (significantly less than expected, $P \geq 0.95$).

Interpretation of program output. A general definition of STCLASS terminology and concepts, and guidelines for the use and detection of disease increase parameters is provided (Table 2). More complete illustration of selected concepts and parameters was provided elsewhere (9).

Scrutiny of the map of nonrandom (or random) disease increase revealed that at t_1

Table 1. Partial spatiotemporal distance class analysis (STCLASS) program output for analysis of disease increase between two disease assessment dates, including simple statistics and portion of the total distance class matrix in numerical format

Total number of positions in matrix	= 200
Number of vacancies (missing values)	= 0
Number of infected plants	= 71
Number of newly infected plants	= 34
Percentage of newly infected plants	= 21%
Number of healthy plants	= 129

LINE 1 STANDARDIZED NUMBER OF OBSERVED INFECTED PAIRS
LINE 2 STANDARDIZED NUMBER OF EXPECTED (SIMULATED) INFECTED PAIRS
LINE 3 SIGNIFICANCE LEVEL
LINE 4 LOWER CONFIDENCE LIMIT OF SIGNIFICANCE LEVEL
LINE 5 UPPER CONFIDENCE LIMIT OF SIGNIFICANCE LEVEL

X		0	1	2	3	4	5	6	7
Y	0	0.0000	0.206+[a]	0.144+	0.1214	0.1083	0.1100	0.1000	0.083$
Y	0	0.0000	0.1350	0.1022	0.1148	0.1271	0.1155	0.1161	0.1503
Y	0	0.0000	0.0000	0.0125	0.3200	0.7575	0.5000	0.6325	0.9650
Y	0	0.0000	0.0000	0.0047	0.2873	0.7275	0.4650	0.5988	0.9521
Y	0	0.0000	0.0000	0.0203	0.3527	0.7875	0.5350	0.6662	0.9779
Y	1	0.216+	0.208+	0.145+	0.1015	0.1009	0.084$	0.0987	0.0877
Y	1	0.1503	0.1344	0.1117	0.1157	0.1153	0.1172	0.1234	0.1351
Y	1	0.0000	0.0000	0.0050	0.7925	0.7700	0.9675	0.8825	0.9425
Y	1	0.0000	0.0000	0.0001	0.7641	0.7405	0.9551	0.8600	0.9262
Y	1	0.0000	0.0000	0.0099	0.8209	0.7995	0.9799	0.9050	0.9588
Y	2	0.211+	0.201+	0.135+	0.0992	0.0926	0.1000	0.0972	0.0833
Y	2	0.1343	0.1231	0.1005	0.1120	0.1122	0.1063	0.1153	0.1271
Y	2	0.0000	0.0000	0.0200	0.7750	0.8400	0.5825	0.7775	0.9175
Y	2	0.0000	0.0000	0.0102	0.7458	0.8143	0.5480	0.7484	0.8982
Y	2	0.0000	0.0000	0.0298	0.8042	0.8657	0.6170	0.8066	0.9368
Y	3	0.188+	0.154+	0.1176	0.0966	0.093$[b]	0.0824	0.0956	0.0784
Y	3	0.1444	0.1223	0.1021	0.1070	0.1273	0.1156	0.1214	0.1236
Y	3	0.0025	0.0025	0.1375	0.7025	0.9675	0.9600	0.8625	0.9100
Y	3	−.0010	−.0010	0.1134	0.6705	0.9551	0.9463	0.8384	0.8900
Y	3	0.0060	0.0060	0.1616	0.7345	0.9799	0.9737	0.8866	0.9300
Y	4	0.169+	0.1250	0.0898	0.063$	0.083$	0.0938	0.063$	0.063$
Y	4	0.1331	0.1242	0.1093	0.1086	0.1204	0.1162	0.1094	0.1360
Y	4	0.0150	0.4600	0.8625	1.0000	0.9800	0.8600	0.9800	0.9925
Y	4	0.0065	0.4251	0.8384	1.0000	0.9702	0.8357	0.9702	0.9865
Y	4	0.0235	0.4949	0.8866	1.0000	0.9898	0.8843	0.9898	0.9985

Number of Distance Classes With SCFs Greater Than Expected: 35
Number of Distance Classes With SCFs Fewer Than Expected: 18
Strongly Non Random Data Set -- Proportion of Significant SCFs ≥ 0.08
Number Of Distance Classes With SCFs Greater Than Expected At The Edges Of The Distance Class Analysis Matrix = 0
Significant Edge Effect *** NOT *** Detected

[a] [X,Y] distance class with standardized count frequency significantly greater than expected ($P \leq 0.05$).
[b] [X,Y] distance class with standardized count frequency significantly less than expected ($P \geq 0.95$).

many diseased sampling units were located near the right-hand edge of the lattice (Fig. 1A and F). There were 37 diseased sampling units (i.e., plants) at t_1 and 34 additional diseased units at t_2. At t_2 many newly diseased sampling units were found in relatively close proximity to previously diseased units (Fig. 1A) or at random throughout the remaining nondiseased sampling locations in the lattice (Fig. 1F).

STCLASS detected a total of 53 distance classes with SCF values significantly greater ($P \leq 0.05$) and less than ($P \geq 0.95$) expected (35 and 18, respectively) for the case of nonrandom disease increase (Fig. 1D). This number exceeded the proportion criterion (>0.05) for concluding that disease increase was nonrandom (9). A significant edge effect was not detected (9). Conversely, for random disease increase between t_1 and t_2 (Fig. 1F), the percentage of significant SCF values was less than 5% of the total number of distance classes (excluding the [*X,Y*] distance class, [*0,0*]) (Fig. 1I). The [*0,0*] distance class was excluded from determination of overall lattice randomness due to the assumption that disease cannot spread from a plant to itself.

The graphical rendering of the STCLASS distance class matrix for the nonrandom scenario allowed for ready interpretation of pattern evolution between t_1 and t_2 (Fig. 1D). The size of the core cluster (or "core zone of disease increase") was 12 (i.e., there were 12 contiguous significant values, including the [*0,0*] class, in the approximate [*X,Y*] region, [*0–2,0–4*]). The [*0,0*] distance class was included in the calculation of core cluster size due to the assumption that the minimum cluster size in a lattice is equal to one sampling unit.

The relatively large and discrete core cluster for nonrandom disease increase (Fig. 1D) was taken as evidence that newly diseased plants tended to occur within four rows and three columns of other diseased and previously diseased plants. In addition, a reflected cluster (of size 24) was detected elsewhere in the distance class matrix in the approximate [*X,Y*] region, [*0–2,8–17*] (Fig. 1D). This was taken as evidence for localized disease increase near a second disease focus within the lattice. The $-symbols elsewhere in the distance class matrix indicated that disease increase was not likely within the [*X,Y*] distance regions delimited by these symbols and separated from the core zone (Fig. 1D).

When the distance class matrices from the STCLASS analysis of these nonrandom data and the 2DCLASS analysis of disease occurrence at t_2 were superimposed, additional information about pattern evolution became available (Fig. 1E). Superimposed STCLASS and 2DCLASS matrices can allow the user to derive additional spatiotemporal information. Distance class matrices from STCLASS and 2DCLASS vary in their similarity to one another for a given data set. By superimposing the distance class matrices from STCLASS and 2DCLASS analyses, inherent differences in pattern recognition between the two forms of analysis can emerge. This variance in pattern recognition can allow investigators to detect the general direction and magnitude of cluster expansion (discussed below) and/or the establishment and coalescence of secondary foci. For example, the @-symbols at the peripheries of the core and reflected cluster represented SCF values that were deemed significantly greater ($P \leq 0.05$) than expected in the STCLASS analysis but not significant in the 2DCLASS analysis. The position of these @-symbols in relation to each other and to #-symbols represented an approximate magnitude and general direction of cluster expansion between t_1 and t_2. For instance, in the region of the core cluster there were three @-symbols ([*2,0–2*]). These were interpreted as representing cluster expansion both within and across rows (up to three lattice positions in each direction, respectively).

For the random scenario for disease increase between t_1 and t_2, there were only nine significant SCF values (excluding the [*X,Y*] distance class, [*0,0*]) in the STCLASS distance class matrix (Fig. 1I). Thus, the hypothesis of random disease increase between t_1 and t_2 could not be rejected. In addition, the nine significant distance classes were not grouped in a discrete pattern, and the core cluster size was the minimum value possible (one).

The superimposed distance class matrices for the example of random disease increase between t_1 and t_2 highlighted an important difference between two-dimensional distance class analysis and spatiotemporal distance class analysis. Two-dimensional distance class analysis of data at t_1 and t_2 (Fig. 1G and H, respectively) indicated nonrandom patterns of disease occurrence at both disease assessment dates. However, using only 2DCLASS analysis, it was not possible to test directly the hypothesis of randomness of disease increase between t_1 and t_2; one could only conclude that the strength of nonrandomness (i.e., proportion of significant SCF values) was greater at t_1 than at t_2 and that the core cluster size decreased between these dates. This reduction in core cluster size provided only indirect evidence of disease increase to distal plants rather than to proximal plants between t_1 and t_2. Conversely, application of STCLASS analysis allowed a direct test of the hypothesis of ran-

Table 2. Spatiotemporal distance class analysis parameters, their definition, use, and/or detection in quantifying spatiotemporal attributes of disease increase in plant populations

STCLASS parameter	Definition, use, and/or detection
Core cluster, or core zone of disease increase	Definition: the number of contiguous distance classes with SCF[a] values significantly greater than expected ($P \leq 0.05$) and adjacent to the [*0,0*] region of the distance class matrix
	Use: indication of nonrandom, proximal increase of disease and/or cluster expansion
Reflected clusters or zones of disease increase	Definition: groups of significant ($P \leq 0.05$) and contiguous SCF values for distance classes not adjacent to the [*0,0*] region of the distance class matrix
	Use: indirect evidence for increase of disease to proximal sampling units and the establishment of secondary disease foci
Cluster expansion	Detection: 1) superimpose the 2DCLASS and STCLASS matrices and look for @ symbols in distance class matrix. b) proximal increase (e.g., cluster expansion) indicated by the presence of discrete core and reflected clusters in the distance class matrix
Cluster filling in	Detection: significant ($P \leq 0.05$) SCF values in distance classes near the [*0,0*] region of the matrix of SCF values or within reflected zones of disease increase
Within-/across-row increase	Detection: significant ($P \leq 0.05$) and contiguous SCF values in *X*- and/or *Y*-directions in the distance class analysis matrix, especially near the [*0,0*] region
Cluster coalescence	Detection: significant ($P \leq 0.05$) SCF values in distance classes near/between two or more reflected zones of increase
Edge effect	Detection: significant ($P \leq 0.05$) SCF values at X_{max} and Y_{max} edges of distance class analysis matrix
Cluster maturity	Definition: time when core or reflected clusters decrease in size and density of SCF values

[a] SCF = Standardized Count Frequency.

domness of disease increase, a hypothesis that was not rejected on examination of the distance class matrix (Fig. 1D). Thus, for both the nonrandom and the random scenarios, disease occurrence at t_1 and t_2 was significantly nonrandom. The question of how disease increase occurred between t_1 and t_2 could only be addressed via STCLASS analysis.

Example two. Assume a rectangular focus of 42 diseased plants near the center of a 16 row × 16 column lattice at t_1 (Fig. 2A). A complete ring of adjacent plants that surrounded the initial focus was diseased by t_2. Thus, it appeared that radial expansion of the initial cluster occurred. A comparison of 2DCLASS distance class matrices from analysis of the patterns at t_1 and t_2 (Fig. 2B and C, respectively) indicated an increase in core cluster size between the assessment dates. These data also were submitted to spatiotemporal distance class analysis.

Strong evidence for disease increase to proximal plants (a phenomenon involved in cluster expansion) was found in the presence of significant ($P \leq 0.05$) SCF

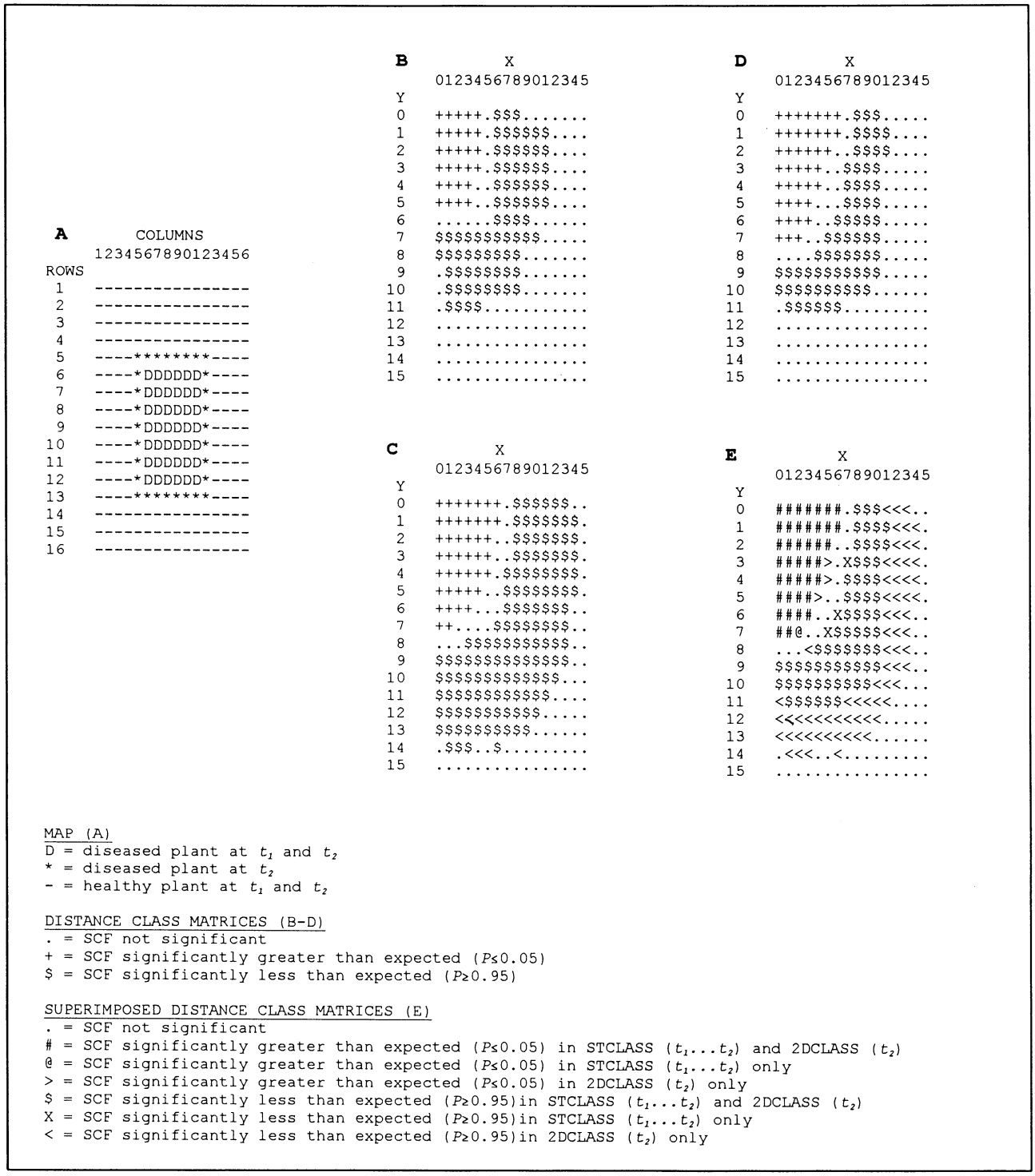


Fig. 2. Spatiotemporal distance class analysis (STCLASS) graphical program output for hypothetical data depicting nonrandom increase (e.g., radial expansion of a disease focus) in disease incidence between disease assessment dates (t_1 and t_2) in a 16 row × 16 column plant distribution lattice. **(A)** Map of data. **(B, C)** Distance class matrices from two-dimensional distance class analysis (2DCLASS) of data in **(A)** at t_1 and t_2, respectively. **(D)** Distance class matrix from STCLASS analysis of disease increase between t_1 and t_2. **(E)** Distance class matrix showing superimposed matrices from 2DCLASS **(C)** and STCLASS **(D)** analyses.

values in the core [0,0] region of the matrix (Fig. 2D). The [X,Y] dimensions of the core cluster and its isodiametric shape suggested radial disease increase from t_1 to t_2.

An additional indication of cluster expansion was found in the band of <-symbols that curved downward through the superimposed 2DCLASS/STCLASS matrices (Fig. 2E). The presence of 30 newly diseased plants at the focus perimeter (Fig. 2A) influenced the significance testing for those distance classes in the approximate [X,Y] region [*11–14,0–10*] and [*0–10,11–14*] of the distance class matrix. In the STCLASS analysis of these data, SCF values were not deemed significantly less than expected ($P \geq 0.95$) in this region of the matrix. Conversely, these distance classes did contain significant SCF values ($P \geq 0.95$) in the 2DCLASS analysis (Fig. 2C). Thus, 2DCLASS analysis indicated that diseased plants tended not to be separated by approximately 12 to 15 plants (within or across rows), whereas the STCLASS analysis detected the random (nonsignificant) tendency of newly diseased plants to fall within this range of distance separation. In other words, the distance separation between this band of SCF values and the [0,0] coordinate region was an indirect indication of the approximate range of distance(s) that separated newly diseased plants from other diseased and previously diseased plants. The shape and size of this band of <-symbols reflected the relatively symmetrical expansion of the initial disease focus.

Guidelines for Program Use

Relatively conservative, albeit arbitrary, guidelines for use and application of spatiotemporal distance class analysis are recommended. The guidelines are modified from those developed for two-dimensional distance class analysis (10,12). Diseased plants must remain diseased throughout the entire epidemic (or be rated as missing values). Thus, analyses are limited to virus epidemics and those other pathosystems not normally characterized by a remission of symptoms (e.g., via defoliation) in the host plant. However, the analysis may be used to characterize spatiotemporal attributes of any intensively mapped physical or biological system in which sampling units are in a rectangular grid and for which binary data are relevant simplifications of the system. Disease incidence on the first disease assessment date can be any value greater than or equal to 1%. However, the incidence of newly diseased plants at t_2 should be approximately 10% of the healthy population at t_1. If the incidence of newly diseased plants is too low, then unusually high numbers of significant SCF values are yielded by the analysis (nearly every occupied distance class is deemed significant). Similarly, if the percentage of newly diseased plants exceeds approximately 90% of the previously healthy plant population, a uniform pattern of disease increase/occurrence is indicated. Missing values are tolerated by the analysis. The suggested maximum percentage of missing values for a data set is 20%. A minimum of 400 simulations is recommended to stabilize the statistics and to enhance repeatability of results. However, slight variability in significance of SCF values will occur when the same data are analyzed repeatedly due to the slight differences in the generation of the randomly simulated maps. With the current DOS version of the program, data set size is limited to lattices with approximately 1,600 sampling units (due to memory limitations of the programming language). The minimum percentage of distance classes with significant SCF values needed to indicate nonrandomness equals approximately 5% of the total number of distance classes (excluding the [0,0] distance class). Strongly nonrandom data sets contain greater than 8% significant SCF values. Edge effects are deemed significant if 12.5% of the SCF values at the right hand (X_{max}) and bottom edges (Y_{max}) of the distance class matrix are significantly greater than expected.

The STCLASS program source code, executable file, and user's guide are available free of charge from the author on request.

LITERATURE CITED

1. Englund, E., and Sparks, A. 1991. GEO-EAS, Geostatistical Environmental Assessment Software, Version 1.2.1. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Las Vegas, NV.
2. Gates, C. E. 1988. Discrete, a computer program for fitting discrete frequency distributions. Pages 458-466 in: Lecture Notes in Statistics. Vol. 55, Estimation and Analysis of Insect Populations. L. McDonald, B. Manly, J. Lockwood, and J. Logan, eds. Springer-Verlag, Berlin.
3. Gottwald, T. R., Richie, S. M., and Campbell, C. L. 1992. LCOR2—Spatial correlation analysis software for the personal computer. Plant Dis. 76:213-215.
4. Gray, S. M., Moyer, J. W., and Bloomfield, P. 1986. Two-dimensional distance class model for quantitative description of virus-infected plant distribution lattices. Phytopathology 76:243-248.
5. Gray, S. M., Moyer, J. W., Kennedy, G. G., and Campbell, C. L. 1986. Virus-suppression and aphid resistance effects on spatial and temporal spread of watermelon mosaic virus 2. Phytopathology 76:1254-1259.
6. Madden, L. V., and Hughes, G. 1994. BBD—Computer software for fitting the beta-binomial distribution to disease incidence data. Plant Dis. 78:536-540.
7. Madden, L. V., Louie, R., Abt, J. J., and Knoke, J. K. 1982. Evaluation of tests for randomness of infected plants. Phytopathology 72:195-198.
8. Minitab. 1991. MINITAB Reference Manual, Release 8. Minitab Inc., State College, PA.
9. Nelson, S. C. 1995. Spatiotemporal distance class analysis of plant disease epidemics. Phytopathology 85:37-43.
10. Nelson, S. C., and Campbell, C. L. 1993. Comparative spatial analysis of foliar epidemics on white clover caused by viruses, fungi, and a bacterium. Phytopathology 83:288-301.
11. Nelson, S. C., Ferreira, S. F., Pitz, K. Y., and Sanaka, V. 1993. Spatio-temporal distance class analysis of epidemics. Phytopathology 83:1362.
12. Nelson, S. C., Marsh, P. L., and Campbell, C. L. 1992. 2DCLASS, a two-dimensional distance class analysis software for the personal computer. Plant Dis. 76:427-432.
13. Reynolds, K. M., and Madden, L. V. 1988. Analysis of epidemics using spatiotemporal autocorrelation. Phytopathology 78:240-246.
14. SAS Institute. 1988. SAS/STAT User's Guide, Release 6.03 ed. SAS Institute, Cary, NC.
15. Yost, R. S., Trangmur, B. B., Ndiaye, J. P., and Yoshida, N. S. 1989. Geostatistical Software for PC-DOS and MS-DOS: User's Guide. Part I: Semivariograms. Hawaii Institute of Tropical Agriculture and Human Resources Ext. Ser. 108, Honolulu, HI.