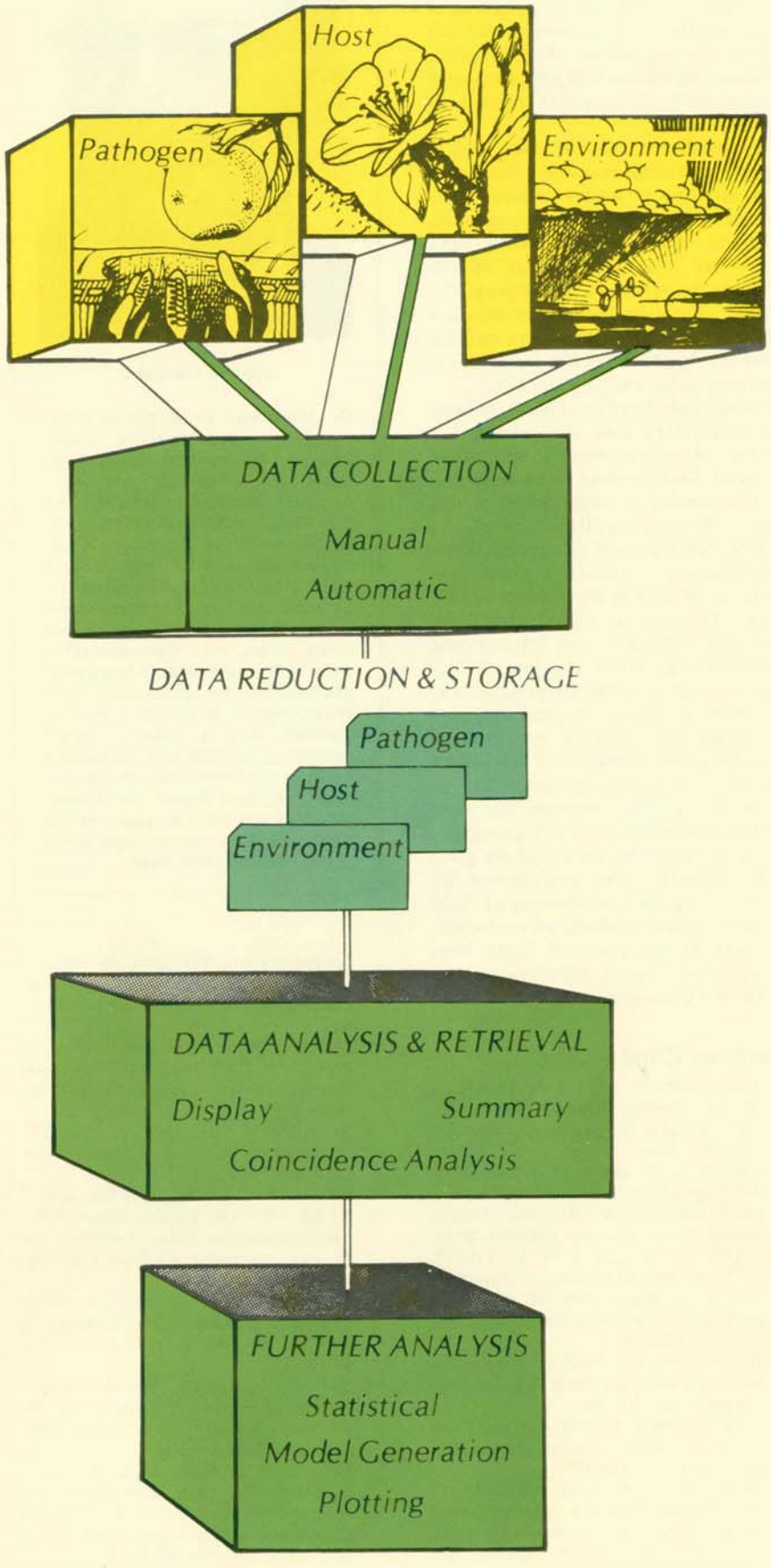
S. P. Eisensmith, R. Loria,
B. D. Olson, and A. L. Jones
Michigan State University,
East Lansing

EPISTAR: An Information



Computers are suited for storing and manipulating large, heterogeneous data bases, and facilities are accessible in many parts of the world. Powerful statistical packages (1,4,12,14) are available but do not meet all the needs of the plant pathologist. Although computer programs have been written for specific teaching (3,9,10), extension (7,8,13), and research purposes (11), little has been done to provide computer software to assist in handling general epidemiological data. At Michigan State University, we have prepared this type of computer program.

The System

EPidemiological Information STorage And Retrieval (EPISTAR) is a computer-based system designed to assist researchers in collecting, reducing, storing, and analyzing data on pathogen, host, and environment. The system consists of: 1) forms for collecting and organizing data into a standard format, 2) an interactive FORTRAN V computer program that displays, summarizes, and performs coincidence analysis (11) on data, 3) a user's manual, and 4) a programmer's manual that describes program modification and transportability. The computer program is well commented and is easily modified to allow for new data processing commands or additional data types. For example, algorithms developed at Cornell University (5) that allow transcription of hourly values from hygrothermograph or

Michigan Agricultural Experiment Station Journal Article 9305.

Reprints of this article or information concerning acquisition of the EPISTAR package can be obtained from: EPISTAR, Department of Botany and Plant Pathology, Michigan State University, East Lansing 48824.

0191-2917/80/07064606/$03.00/0

Epidemiological Storage, Retrieval, and Analysis System

deWit leaf wetness charts (M. deWit, Hengelo, Netherlands) have been added to our version of EPISTAR at Michigan State University.

Data Types

EPISTAR accommodates 10 types of environmental measurements: air temperature, barometric pressure, leaf wetness, rainfall, relative humidity, soil moisture, soil temperature, solar radiation, wind direction, and wind speed. The system can handle up to 24 observations per day for each, and all output is in the same units of measurement as the original data. Biological data types processed by EPISTAR include catches of the Burkard spore trap (Burkard Scientific Sales Ltd., Rickmansworth, Hertfordshire, England) and Rotorod spore trap (Ted Brown Associates, Los Altos Hills, CA 94022), water trap catches, disease ratings, and stages of host phenology. Burkard spore trap catches may be recorded up to 24 times per day, whereas other biological data types may be recorded, at most, once daily. Data may be taken from multiple environmental sensors, spore traps, or plots during the same time period.

Data Collection

Fig. 1 shows the three data collection forms for recording biological and environmental data in the field or in the greenhouse (Fig. 2). Form I is for environmental and Burkard spore trap catch data, Form II is for Rotorod and water trap spore catch data, and Form III is for disease ratings and host phenology data. The forms also facilitate decoding strip charts from environmental monitoring equipment. To minimize transcription errors, data can be keypunched directly from the forms.

The Program

The three basic operations of the computer program are to: 1) display stored data for examination, 2) calculate summary statistics, and 3) search for relationships among variables.

Display. This allows convenient access to portions of the original data and facilitates creation of subfiles for plotting or statistical analyses. To obtain data, the user specifies the data type(s) and date range. The program then searches the data file and prints out the results.

Summary. This provides statistics useful in analyzing epidemiological information. Each biological and environmental factor is summarized on a daily basis. Summary calculations include minimum, maximum, and average values; number of time periods above or below user-specified thresholds; and number of missing values. The user may select subsets of data for

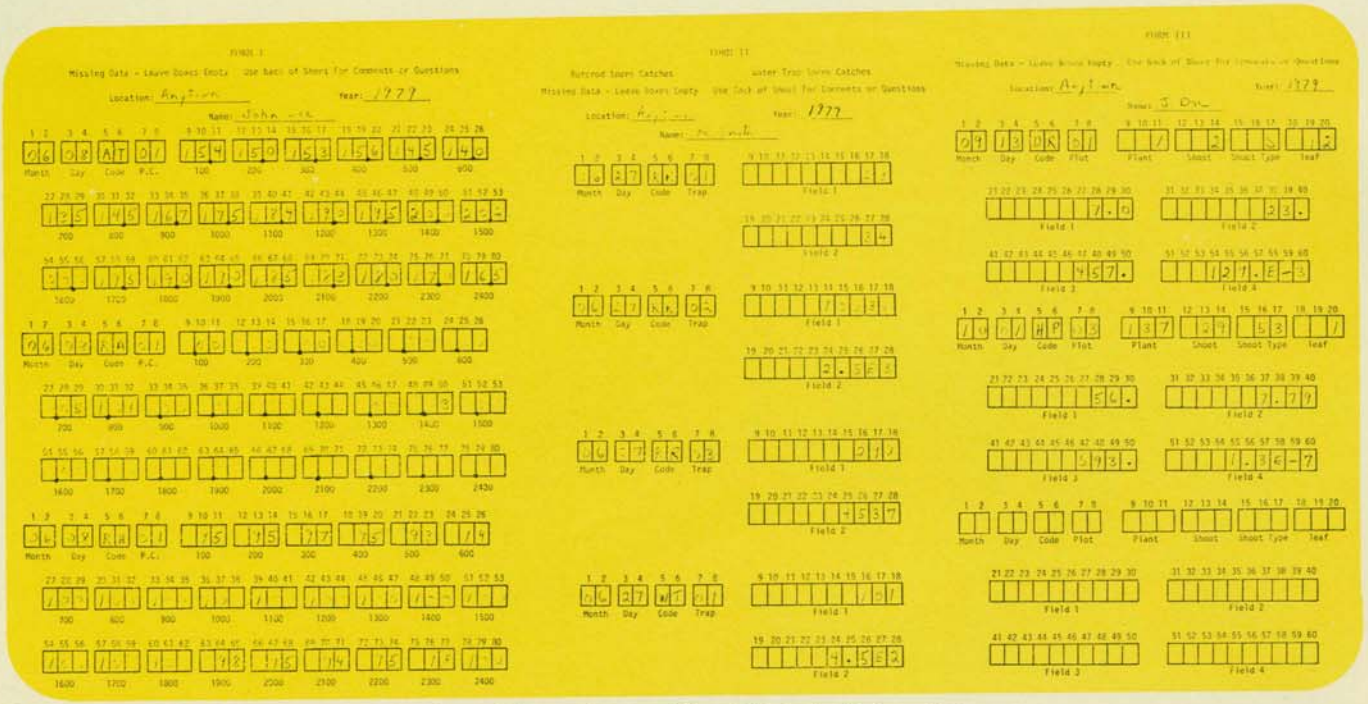


Fig. 1. Coded data forms for collecting and organizing environmental and biological information.

summarization, ie, data from a particular environmental sensor, plot, plant, shoot, shoot type, or leaf.

Coincidence analysis. This permits locating time periods within the data file with conditions that match user-specified values, for example, periods with greater than 0.01 in. of rainfall, temperatures less than 85 F, and a spore trap catch of at least 10 spores per hour. The conditions are specified using the FORTRAN relational operator codes LT, LE, EQ, NE, GT, and GE in reference to a user-specified value. Series of comparisons may be joined with the logical operators "and" or "or." Coincidence analysis may be used on all environmental data and Burkard spore trap catches and is particularly helpful in detecting associations between environmental conditions and biological events.

EPISTAR requires that data be placed in a permanent computer file. This is accomplished by punching the data from the coding forms onto computer cards, which are then read into a file via a card reader. Alternatively, the data can be typed directly into a file from an interactive terminal or dumped from a cassette tape.

Using the Program

Once the EPISTAR program is accessed, all the information necessary to run the program is requested from the user in the order outlined in Fig. 3. The user may choose to display, summarize, perform coincidence analysis, change the placement of the output, select a new data file, or terminate the program by entering the appropriate mnemonic command. EPISTAR accesses the data file and writes out the information, including collection date range, collector, collection location, and the data types in the file. If the user chooses to display, summarize, or perform coincidence analysis, the desired data types and probe, plot, or trap number(s) must be selected from those listed as available. If the data are to be summarized, appropriate thresholds should be provided for environmental data. Relational operators and values and possibly a logical operator must be chosen for coincidence analyses. When the data processing is completed, the EPISTAR program will again ask the user which procedure is desired—display, summary, coincidence analysis, new data file, or program termination. User responses are checked for errors, and invalid responses are rejected with a diagnostic message. No knowledge of FORTRAN programming is necessary to use EPISTAR.

Output

The form of output depends on the kind of processing and the type(s) of data

Fig. 2. Data collection in the greenhouse.

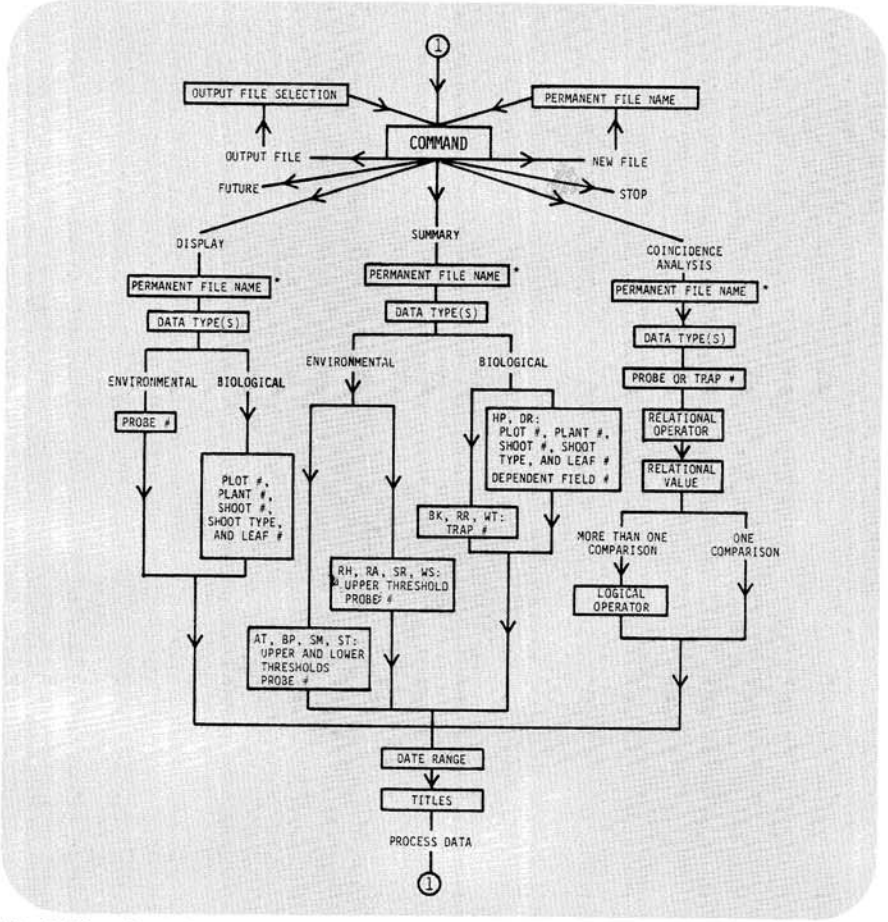


Fig. 3. Flowchart of user-supplied information and program execution.

```
   WELCOME TO EPISTAR -- AN EPIDEMIOLOGICAL INFORMATION STORAGE AND RETRIEVAL PACKAGE
                         SCOTT P. EISENSMITH
                         ROSEMARY LORIA
                         BRIAN D. OLSON
                         ALAN L. JONES
                    DEPARTMENT OF BOTANY AND PLANT PATHOLOGY
                    MICHIGAN STATE UNIVERSITY

ENTER COMMAND-help

LEGAL COMMANDS ARE-
    DISP
    SUMR
    COIN
    STOP
    NEWF
    OUTF


ENTER COMMAND-disp

ENTER PERMANENT FILE NAME
*cddaepistardatafile

        JOHN DOE, JR.        OF THE DEPARTMENT OF PLANT PATH
   COLLECTED THE FOLLOWING TYPES OF DATA AT   ANYTOWN, MICHIGAN
   BEGINNING  6/ 8/79 UNTIL  6/10/79

        AIR TEMPERATURE (CODE=AT)  1 PROBE(S)
        BAROMETRIC PRESSURE (CODE=BP)  1 PROBE(S)
        LEAF WETNESS (CODE=LW)  1 PROBE(S)
        RAINFALL (CODE=RA)  1 PROBE(S)
        RELATIVE HUMIDITY (CODE=RH)  1 PROBE(S)
        SOIL MOISTURE (CODE=SM)  1 PROBE(S)
        SOLAR RADIATION (CODE=SR)  1 PROBE(S)
        SOIL TEMPERATURE (CODE=ST)  1 PROBE(S)
        WIND DIRECTION (CODE=WD)  1 PROBE(S)
        WIND SPEED (CODE=WS)  1 PROBE(S)
        BURKARD CATCHES (CODE=BK)  1 TRAP(S)
        DISEASE RATING (CODE=DR)  1 PLOT(S)
        HOST PHENOLOGY (CODE=HP)  1 PLOT(S)
        ROTOROD CATCHES (CODE=RR)  3 TRAP(S)
        WATER TRAP CATCHES (CODE=WT)  2 TRAP(S)

ENTER DATA TYPE CODES, ONE PER LINE
OR TYPE E TO END
   1-at
   2-lw
   3-bk
   4-ra
   5-wt
      ENTER A 2 DIGIT NUMBER (I.E., 02 FOR 2)
      OR -1 FOR ALL VALUES
         TRAP NUMBER-01
   6-e

ENTER START DATE AS (MO/DA/YR)
*06/08/79

ENTER STOP DATE AS (MO/DA/YR)
*6/10/79

TITLES ON OUTPUT?
(Y OR N)-y

                1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19   20   21   22   23   24
 6/ 8 AT 1 15.4 15.0 15.3 15.6 14.5 14.0 13.5 14.5 16.7 17.5 18.4 19.0 19.5 20.0 20.2 20.2 19.5 19.0 19.0 18.5 18.3 18.0 17.0 16.5
 6/ 8 LW 1    1    1    1    1    1    1    2    2    2    2    2    2    2    2    2    2    2    1    0    0    0    0    0
 6/ 8 RA 1  .00  .00  .00  .00  .00  .00  .15  .21  .18  .15  .25  .21  .18  .15  .10  .9   .15  .20  .00  .00  .00  .00  .00  .00
 6/ 8 BK 1    0    0    0    0    0    0    0    4    5    8   11    9    8    5    6    7   12    4    2    1    2    4    3    4
 6/ 8 WT 1         10              4.5
 6/ 9 AT
 6/ 9 LW
 6/ 9 BK
 6/ 9 RA
 6/ 9 WT
 6/10 AT 1 16.0 15.7 15.8 16.0 16.2 16.5 16.0 17.0 18.0 20.0 21.2 22.3 24.3 25.4 27.5 25.6 28.6 27.5 26.5 25.4 24.5 25.4 24.5 24.3
 6/10 LW 1    1    1    1    1    1    1    1    1    0    0    0    0    0    0    0    0    0    0    0    0    0    1    1
 6/10 RA 1  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00  .00
 6/10 BK 1    3    2    3    4    2    4    4    2    0    0    0    0    0    0    0
 6/10 WT 1         4.76E3              15.0
**********************************************************************************************************************************

ENTER COMMAND-sumr

ENTER DATA TYPE CODES, ONE PER LINE
OR TYPE E TO END
   1-at
ENTER UPPER THRESHOLD FOR TEMPERATURE
*20
ENTER LOWER THRESHOLD FOR TEMPERATURE
*15
   2-lw
   3-ra
ENTER UPPER THRESHOLD FOR RAIN
*0.03
   4-bk
   5-e

ENTER START DATE AS (MO/DA/YR)
*6,8,79

ENTER STOP DATE AS (MO/DA/YR)
*06/10/79

TITLES ON OUTPUT?
(Y OR N)-y

                AIR TEMPERATURE       LEAF WETNESS          RAINFALL                BURKARD CATCHES
MO DA YR DOY * NM  AVG   MIN  MAX AT RT * NM TO RA DE * NM  AVG   MAX  TOT DR AT * NM  MIN   MAX   AVG     SUM *
 6/ 8/79 159 *  0 17.30 13.5 20.2  3  5 *  0 19 12  7 *  0  .084  .25 2.02 12 12 *  0    0    12    4.0     95 *
 6/ 9/79 160 * 24-99.00-99.0-99.0-99-99 * 24-99-99-99 * 24 -.990 -.99 -.99-99-99 * 24  -99   -99  -99.0    -99 *
 6/10/79 161 *  0 21.67 15.7 28.6 15  0 *  0 10  0 10 *  0     0    0    0  0  0 * 10    0     4    1.7     24 *
**********************************************************************************************************************************

ENTER COMMAND-coin

ENTER DATA TYPE CODES, ONE PER LINE
OR TYPE E TO END
   1-lw

ENTER RELATIONAL OPERATOR FOR LW
*eq

ENTER RELATIONAL VALUE FOR LW
*1
   2-bk

ENTER RELATIONAL OPERATOR FOR BK
*gt

ENTER RELATIONAL VALUE FOR BK
*0
   3-e

ENTER LOGICAL OPERATOR FOR THIS SET
*a

ENTER START DATE AS (MO/DA/YR)
*6/8/79

ENTER STOP DATE AS (MO/DA/YR)
*6/10/79

TITLES ON OUTPUT?
(Y OR N)-y

  DATE    DOY     1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 SUM
 6/ 8/79 159                                                    ****                    1
 6/ 9/79 160                                                                            0
 6/10/79 161  ********************************                                       8
      TOTALS    1  1  1  1  1  1  1  1  1  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0
**********************************************************************************************************************************

ENTER COMMAND-stop
  STOP
     3.289 CP SECONDS EXECUTION TIME
```

AUTHOR'S COMMENTS

USER-SUPPLIED INFORMATION AND COMPUTER RESPONSES ARE IN LOWER AND UPPER CASE LETTERS, RESPECTIVELY.

A LIST OF THE LEGAL COMMANDS IS GIVEN IN RESPONSE TO 'HELP'.

USER SUPPLIES THE NAME OF THE DISK FILE UPON WHICH THE DATA ARE STORED.

THE PROGRAM DISPLAYS THE INFORMATION CONTAINED ON THE INFORMATION CARDS IN THE BEGINNING OF THE DATA DECK.

EXECUTING A DISPLAY RUN FOR AIR TEMPERATURE, LEAF WETNESS, BURKARD SPORE TRAP CATCHES, RAINFALL, AND WATER TRAP (# 1) CATCHES.

DATE RANGE OF DATA TO BE DISPLAYED.

TITLES FOR A DISPLAY RUN CONSIST OF THE HOURS OF THE DAY.

MISSING DATA ARE LEFT BLANK IN A DISPLAY RUN.

EXECUTING A SUMMARY RUN.

THE USER SUPPLIES THE UPPER AND LOWER THRESHOLDS FOR AIR TEMPERATURE AND RAINFALL THAT WILL BE USED IN THE SUMMARY CALCULATIONS.

DATE RANGE FOR SUMMARY RUN.

TITLES AID IN THE INTERPRETATION OF SUMMARIZED DATA.

EXECUTING A COINCIDENCE ANALYSIS RUN.

THE USER WISHES TO LOCATE THOSE TIME PERIODS WHEN LEAVES WERE WET FROM DEW (CODED AS 1) AND SPORES WERE CAUGHT IN THE BURKARD SPORE TRAP.

DATE RANGE FOR COINCIDENCE ANALYSIS RUN.

TITLES ARE THE HOURS OF THE DAY.

STARS INDICATE WHICH TIME PERIODS THE USER-SPECIFIED CONDITIONS WERE MET.

Fig. 4. Sample run of the EPISTAR program, with authors' comments.

Fig. 5. Collection of ascospores with a Burkard volumetric spore trap.

Fig. 6. Use of a digitizer to transcribe environmental data from strip charts.

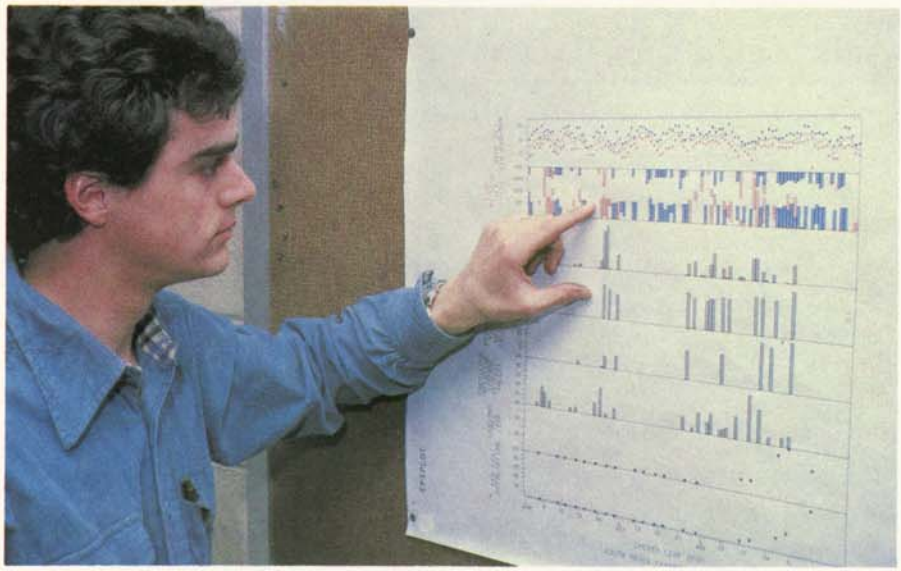


Fig. 7. Seasonal plot of environmental and biological data made from files organized by the EPISTAR computer program.

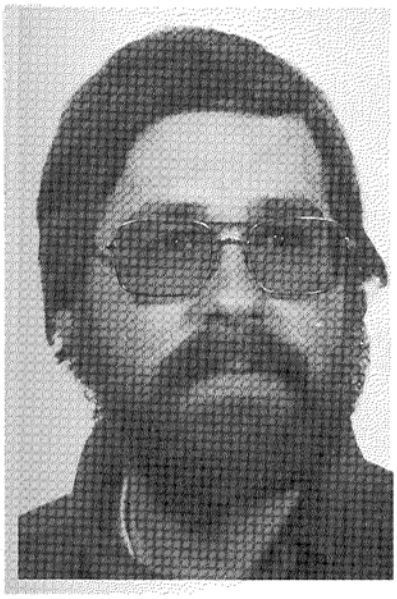
selected. Output from the program may be printed at a computer terminal or written onto a disk file for use with statistical analysis (1,12) and plotting (2,6) programs. A sample run (Fig. 4) illustrates output printed at the terminal. In display output, each line represents one day and begins with the date, data type code, and probe, plot, or trap number. The raw data follow, with decimal places inserted according to the standard formats. Each line of summary output begins with the date and the day of the year, followed by the summarized data values for that date. Summary values are calculated from data sets with missing data, but the number of missing values is indicated in the output. If all data values are missing, missing value indicators are given instead of the summarized value. Each data type is separated by a column of stars. Each line of coincidence analysis output begins with the date and the day of the year. Those time periods, represented in the title as the first through the 24th hour, for which the user-specified requirements have been satisfied are indicated with stars. The total number of time periods for which conditions have been satisfied for a day and for a particular time period are given as the sums of the rows and the columns, respectively.

An Application of EPISTAR

Cherry leaf spot, caused by *Coccomyces hiemalis* Higgins, is a major disease of sweet and sour cherries in Michigan and is characterized by necrotic spots on the foliage, leaf yellowing, and premature defoliation. In a study to define the environmental conditions necessary for infection and to develop a disease management model, environmental (rain, air temperature, leaf wetness, and relative humidity), host (number of leaf nodes with and without leaves), pathogen (number of spores caught in Rotorod, Burkard [Fig. 5], and water traps), and disease (percent defoliation and lesions per leaf) data were collected, organized, and stored using the EPISTAR format. Weather data were transcribed from strip charts onto EPISTAR data forms by hand or coded directly into a computer file via a digitizer (Fig. 6).

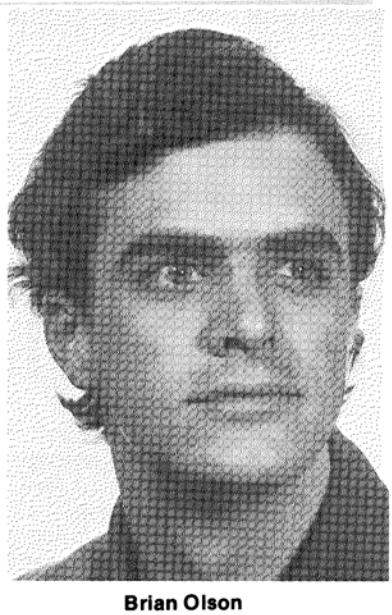
The display and summary options helped detect data transcription errors. The display option was also used to examine subsets of environmental data during periods when infection occurred on tagged shoots and exposed potted plants. The summary option was used to organize files for plotting (Fig. 7) and statistical analyses. Hypothesized environmental conditions necessary for infection, eg, rainfall and leaf wetness, were located within the data set using coincidence analysis.

An empirical model relating leaf wetness duration and mean air



Scott Eisensmith

Rosemary Loria



Brian Olson

Scott Eisensmith, Rosemary Loria, and Brian Olson are graduate students in the Department of Botany and Plant Pathology at Michigan State University, East Lansing. Scott Eisensmith is studying the epidemiology of cherry leaf spot and its application to disease management. Rosemary Loria is completing her dissertation on the epidemiology and seed indexing of loose smut of wheat. Brian Olson is investigating the survival and spread of bacterial canker pathogens in cherry orchards. These students are receiving their Ph.D.s under Dr. Alan Jones, who served as an advisor for the EPISTAR project.

temperature to conidial infection was developed to identify infection periods and forecast relative disease severity. Disease occurrence was monitored by counting lesions on tagged shoots every 4–7 days during three growing seasons. The model's predictions were correct in 95% of the 65 cases of infection and in 83% of the 18 cases of no infection. Fungicides were applied either on a calendar schedule or when environmental conditions were favorable for infection. When the disease prediction system was used in a research orchard during 1979, control comparable to that with a calendar spray schedule was achieved with four fewer fungicide sprays.

Advantages

The EPISTAR system may serve as a model for other standardized epidemiological data-processing systems. Data forms allow immediate organization of information into a computer-compatible format, thus facilitating the use of other computer software for data analysis or plotting. Subsets of the master data file are easily selected by date range and data type. Multiple observations for a single day can be summarized into daily values, and data meeting user-specified conditions can be quickly located within the data file.

The computer program is written in a modular format to facilitate modification, including addition of new data types or processing routines. We hope it may serve as a framework for additional epidemiologically oriented computer software.

Literature Cited

1. ANONYMOUS. 1974. Michigan State University STAT System: User's Guide VIII. Computer Laboratory, Michigan State University, East Lansing 48824. 287 pp.
2. ANONYMOUS. 1977. Statistical Plotting On-line Command System: User's Guide. Pest Management Technical Report 13. Department of Entomology, Michigan State University, East Lansing 48824. 54 pp.
3. ARNESON, P. A., T. R. OREN, R. LORIA, J. J. JENKINS, E. D. GOODMAN, and W. E. COOPER. 1979. APPLESCAB: A pest management game. NACTA Journal 23(3):61-62.
4. BARR, A. J., J. H. GOODNIGHT, J. P. SALL, and J. T. HELWIG. 1976. A User's Guide to SAS-76. SAS Institute, P.O. Box 10066, Raleigh, NC 27605. 329 pp.
5. BLUME, M. C., R. C. SEEM, and J. BARNARD. 1978. Two computer programs used in the analysis of rectangular and circular charts from continuously recording weather instruments. Search 9(1) Plant Pathology Department, No. 5, NY Agric. Exp. Stn., Geneva. 11 pp.
6. BOBZIN, J., S. P. EISENSMITH, and A. L. JONES. 1978. Plotting of Environmental and Biological Data: User's Guide. Pest Management Technical Report 16. Department of Botany and Plant Pathology, Michigan State University, East Lansing 48824. 82 pp.
7. COTTER, H. V. T., W. E. MacHARDY, and J. A. WARREN. 1979. Storing, retrieving, and reporting plant disease occurrence data using a computer data base management system. Plant Dis. Rep. 63:117-121.
8. CROFT, B. A., J. L. HOWES, and S. M. WELCH. 1976. A computer based extension pest management delivery system. Environ. Entomol. 5:20-34.
9. DIMOFF, K., and N. KILMER. 1978. SAMPLE User's Manual. Department of Entomology, Michigan State University, East Lansing 48824. 21 pp.
10. LOGAN, P. 1978. ORCHARD: A serious predator-prey interaction game. Department of Entomology, Michigan State University, East Lansing 48824. 26 pp.
11. MOGK, M. 1974. Automatic data processing in analysis of epidemics. Pages 55-77 in: J. Krantz, ed. Epidemics of Plant Disease. Springer-Verlag, New York. 170 pp.
12. NIE, H. C., H. HULL, J. G. JENKINS, K. STEINBRENNER, and D. H. BENT. 1975. Statistical Package for the Social Sciences. McGraw-Hill Publications, New York. 675 pp.
13. PENNYPACKER, S. P., H. D. KNOBLE, and J. L. LONGENECKER. 1976. Plant disease survey information (PDSI): A computer data storage and retrieval system. (Abstr.) Proc. Am. Phytopathol. Soc. 3:246.
14. RYAN, T. A., B. L. JOINER, and B. F. RYAN. 1976. MINITAB: A Student Handbook. Duxbury Press, Division of Wadsworth Publishing Co., North Scituate, MA. 341 pp.