PREDICTION OF THE POPULATION DENSITY OF CITRUS
TREE PESTS BY A MULTIPLE REGRESSION AND A SYSTEM
MODEL

Masae SHIYOMI, Shota HIROSAKI and Ryoji KORENAGA
(National Grassland Research Institute; National
Institute of Environmental Studies; National
Fruit Tree Research Station)

Summary

In order to predict efficiently the population density of two most harmful pests infesting citrus trees, that is, arrowhead scale and red mite, a research group to introduce statistical or mathematical techniques was established in the Ministry of Agriculture, Forestry and Fishery, in 1971. This research group is organized by researchers belonging to Prefectural Horticulture or Fruit Tree Experiment Stations, including Kanagawa, Shizuoka, Hiroshima, Ehime, Saga and Kumamoto Prefectures, and several National Experiment Stations. The model for predicting arrowhead scale population density was constructed by the multiple regression analysis, and the model for red mite was constructed by the system dynamics. Now, these models are offered for testing the validity of prediction.

0. Introduction

An introduction of statistical or mathematical methods for efficient prediction of pest population density has been attempted in the Ministry of Agriculture, Forestry and Fishery since 1971. Several models for predicting the arrowhead scale(Unaspis yanonensis Kuwana) population density were constructed by the multiple regression analysis using biological and environmental data accu-

culated for many years at Kanagawa Horticulture Experiment Station, National Fruit Tree Research Station(Okitsu), Hiroshima Fruit Tree Experiment Station and Kumamoto Fruit Tree Experiment Station. On the other hand, five models for predicting red mite(Panonychus citri McGregor) population density is being constructed by the method of system dynamics at Shizuoka Citrus Research Station, Ehime Fruit Tree Experiment Station, Hiroshima Fruit Tree Experiment Station and Saga Fruit Tree Experiment Station.

1. Arrowhead scale

Linear multiple regression models to predict the population density of summer and autumn generations of the arrowhead scale were constructed. The population density at each specified developmental stage of the insect, average temperature and precipitation in each month of observations, the differences between districts, etc. were used as independent variables in these analyses (Table 1). These data had been accumulated from 1960 to 1972 at experiment stations mentioned previously, then the total number of samples provided for the analyses is over 150.

The following two types of model were built up by statistical(stepwise forward method) and/or biological significant variables. Type 1: Four separated models constructed for respective location by using data obtained at each of the experiment stations noted above. Type 2: A common model constructed by using all data from the 4 experiment stations.

In these 5 cases, the multiple correlation coefficients of the regression attained at least 0.85. The results in the 2nd generation are shown in Table 2.

These 5 models were used to predict the arrowhead scale population density in two years following the observations used in the analysis, i.e., 1973 and '74. Table 3 shows that type 2 model had high predictability but type 1 models did not. This is

Table 1. Variables used in analysis of the 2nd generation

Variable no.	Explanation
1*	Population of female adults of the 1st generation(late July)
2*	Population of female adults of the 1st generation on the initial occurrence date of larvae of the 2nd generation
3*	Population of female adults of the 2nd generation(October): dependent variable
4	Activity of natural enemy(none:0, middle:1, high:2)
5	Maximum number of fixed larvae of the 2nd generation/no.2
6	Mean temperature(July)
7	Mean temperature(August)
8	Mean temperature(September)
9	(no.6-25.63) ²
10	(no.7-27.44) ²
11	(no.8-23.81) ²
12*	A(>5mm/day)
13*	A(>15mm/day)
14	B(>5mm/day)
15	B(>15mm/day)
16*	A(>5mm/day)
17*	A(>15mm/day)
18	B(>5mm/day)
19	B(>15mm/day)
20*	A(>5mm/day)
21*	A(>15mm/day)
22	B(>5mm/day)
23	B(>15mm/day)
24	Kanagawa=1, others=0 (dummy variable)
25	Hiroshima=1, others=0
26	Kumamoto=1, others=0

* transformation of variables with $\log(x+1)$; A: amount of precipitation, B: number of rainy days; units in variables no. 1 to 3: the number of individuals/branch

Table 2. Multiple regression model for prediction (2nd generation; dependent variable: no.3)

Variable no.	Kanagawa	Okitsu	Hiroshima	Kumamoto	Pooled
2	0.522	0.513	0.710	0.652	0.722
5	0.029	-0.006	0.001	0.000	0.015
4	-	-0.162	-0.115	-0.146	-0.193
6	0.768	0.149	0.142	0.372	0.100
9	0.582	0.118	0.036	-0.204	-0.013
7	0.221	-0.482	-0.119	0.092	-0.012
10	-0.258	0.373	0.007	-0.181	-0.062
14	0.405	0.020	0.164	0.181	0.076
18	0.045	0.143	0.289	-0.298	0.071
22	-0.031	0.093	0.125	-0.063	0.073
16	-0.259	-0.520	-0.488	0.560	-0.051
24	-	-	-	-	-0.628
25	-	-	-	-	-0.024
26	-	-	-	-	0.074
constant	-25.149	10.714	0.222	-11.205	-1.646
R	0.929	0.875	0.944	0.959	0.922
SE	0.298	0.202	0.222	0.217	0.277

R: multiple correlation coefficient; SE: standard error.

Table 3. Prediction for 4 trees by the obtained multiple regression (2nd generation).

Model (type)	1973				1974			
	1	2	3	4	1	2	3	4
Kanagawa								
Observed	3	7	2	6	13	22	31	65
Type 1	20	49	21	31	1584	6027	2588	8707
Type 2	4	14	5	8	10	51	29	60
Okitsu								
Observed	725	274	143	458	351	182	160	185
Type 1	136	86	79	84	922	976	614	645
Type 2	433	249	215	499	115	128	65	67
Hiroshima								
Observed	451	291	245		183	71	124	161
Type 1	469	193	118		419	269	189	299
Type 2	220	128	75		318	467	182	198
Kumamoto								
Observed	8	4	3	31	44	566	80	
Type 1	56	26	30	94	839	20	83	
Type 2	33	13	15	50	204	127	101	

unit: the number of individuals/branch

because type 2 model is constructed by using data covering various environments. Although type 1 models fitted their own specific situations well, they were not sufficient to forecast future conditions because these models did not incorporate diverse experience which might occur in the future.

These models were further improved by incorporating data accumulated during 5 years following the previous 13 years. A new technique for determining multiple regression equations, that is, the method based on the smallest prediction sum of squares(PSS)(

Table 4. Multiple regression model for prediction (2nd generation; dependent variable: no. 3; by PSS)

Variable no.	Kanagawa	Okitsu	Hiroshima	Kumamoto	Pooled
2	0.868	0.610	0.821	0.853	0.874
5	0.025	0.011	0.030	0.025	0.023
4	-0.228	-0.152	-0.009	-0.236	-0.150
20	0.083	-0.104	0.139	0.306	0.108
24	-	-	-	-	-0.531
25	-	-	-	-	-0.072
26	-	-	-	-	-0.006
constant	-0.076	1.364	0.233	0.233	0.363
R	0.861	0.762	0.856	0.915	0.891
SE	0.322	0.238	0.294	0.321	0.323

R: multiple correlation coefficient; SE: standard error.

Hirosaki et al, 1979). Table 4 shows the estimated regression coefficients of variables selected by PSS and then adjusted by biological consideration, based on the respective set of data obtained at each of 4 experiment stations and the pooled samples.

The prediction for population densities in the 2nd generation in 1978 is shown in Table 5 compared with the observed data. This experience using PSS shows that this method has a possibility to improve the predictability by the linear regression model.

Table 5. Prediction for 4 trees by the obtained multiple regression (2nd generation; by PSS)

Model (type)	1978			
	1	2	3	4
Kanagawa				
Observed	57	20	8	15
Type 1	18	11	5	12
Type 2	20	11	5	13
Okitsu				
Observed	46	33	22	43
Type 1	107	64	67	162
Type 2	68	22	25	95
Kumamoto				
Observed	100	32	43	93
Type 1	36	19	28	141
Type 2	42	22	34	155

unit: the number of individuals/branch

The full outline of these studies are given in Korenaga et al(1978) and Korenaga et al(1981).

2. Red mite

A simplified ecosystem concerning a red mite population is assumed as in Fig. 1. Here, the red mite population has 3 stages, that is, egg, larva and adult; the population densities at these 3 stages at time t are expressed by y_1 , y_2 and y_3 , respectively; and the number of (abstract) natural enemies at time t is expressed by x . These state variables are affected by the following driving variables: precipitation, chemical control, leaf quality, air temperature and so forth.

Fig. 1. A simplified ecosystem concerning a red mite population.

Then, we have

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{x} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ x \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

where $\dot{y}=dy/dt$, a_{ij} 's are parameters representing effects of the 6 driving variables and functions of time t , and e_i 's are constants independent of time and the population densities.

The parameters a_{ij} 's and e_i 's are determined by experiments and experience in computation, but the detail is omitted in this report.

An example of red mite population changes is shown in Fig.2, which contains the real population growth in the field in Shizuoka in 1957 and its simulated result using parameters determined so that the system model fits well the real red mite population.

Table 6. Simulated population changes with time under the assumptions that the average air temperature from July 20 to August 19 is higher 2°C than in 1957 and lower 2°C than in 1957.

Month/day	6/15	6/30	7/15	7/31	8/15	8/31	9/15	9/30	10/15	10/31
1957										
(simulated)	0.0	3.5	16.0	11.8	7.7	11.8	14.8	23.5	19.0	16.5
+2°C	0.0	3.5	16.0	9.5	0.0	0.0	0.0	0.0	0.0	0.0
-2°C	0.0	3.5	16.0	19.0	24.0	4.6	13.0	18.0	17.0	13.0

unit: female adults/10 leaves.

Table 6 shows some population changes with time under the assumptions that the average air temperature from July 20 to August 19 is higher 2°C than in 1957 and lower 2°C than in 1957. If we can exactly incorporate the real world of ecosystem into our system model, we can succeed a dynamic prediction of pest occurrence.

3. Conclusion

We have had a lot of experience concerning mathematical or statistical techniques of biological prediction during these 10

Fig. 2. (a) Real population changes in Shizuoka(1957).

Fig. 2(continued). (b) Simulated results: an example.

years. In these experices, the followings becomes clear for us:
(1) Multiple regression models which utilize accumulated data for many years over many locations are less dynamic in pest prediction than system models. (2) It is impossible at the present time to estimate simultaneously all parameters contained in a system model.

References

- Hirosaki,S. & Kobayashi,F.(1979) Some selection methods of criterion variables by Prediction Sum of Squares (PSS) in regression analysis, Bull. Comp. Cent. Res. Agric. Forest. Fish. A15:45-104.
- Korenaga,R., Shiyomi,M., Hirosaki,S., Nakamura,K., Ito,S., Kimura, Y. & Uemura,M.(1978) Models for prediction of the population density of the arrowhead scale, Unaspis yannonensis Kuwana, on citrus tree by multiple regression, Jap. J. appl. Ent. Zool. 22:141-151.
- Korenaga,R., Hirosaki,S. & Shiyomi,M.(1981) Models for prediction of the population density of the arrowhead scale, Unaspis yannonensis Kuwana, on citrus tree by multiple regression analysis II, Jap. J. appl. Ent. Zool. 25:17-25.