STUDIES IN SAMPLING TECHNIQUE

I. Estimation of Whitefly (*Aleurolobus Barodensis*) Incidence in Sugarcane

By K. L. Khanna, L. N. Nigam, B. R. Sehgal and K. S. Bandyopadhyay
(*Central Sugarcane Research Station, Pusa, Bihar*)

Received August 26, 1946

I. Introduction

The determination of the adequate sampling size in the efficient estimation of whitefly incidence is of much practical importance. The solution of the problem lies in being able to adopt the appropriate method of sampling and to determine its magnitude for estimation of the incidence with any desired accuracy in an infested plot of given area. The common practice of estimating the incidence of whitefly is in terms of the number of puparia per square inch of the leaf area. The process of calculating the incidence, therefore, involves the enumeration of the total puparia per leaf and the calculation of the leaf area. If this enumeration is required to be completed even in one clump in its entirety, the labour involved becomes obviously too huge. Thus the issue that arises is to investigate whether fewer canes in a clump, fewer leaves in a cane and further if partial examination of a leaf would serve the purpose of estimation.

The incidence of whitefly has been found to be highly variable even within a comparatively small area infested with it and it is thought that the sampling requirements might differ from plot to plot varying according to the intensity of incidence.

The aim of the present note is to answer with practicable accuracy the question of sampling as enunciated above. An attempt will be made to throw light on the nature and size of samples varying if at all with low, medium and severe infestation.

Three broad divisions have been made of the type of infestation to enable the survey staff easily recognise them under field conditions so that the right type of sampling procedure may forthwith be applied.

II. Material

The material for the present study consists of the complete enumeration of whitefly incidence in two plots of about 1/40th acre in size (60′ × 18′). The rows have been divided each into units of 3 ft. to facilitate the sampling

TABLE

Number of clumps, canes and leaves (affected

Units	Row 1								Row 2								Row 3							
	Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected		
			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.
1	3	9	89	36	125	76	7	83	2	12	95	96	191	79	27	106	1	3	28	10	38	21	6	27
2	2	6	70	47	17[illegible]	57	12	69	2	6	75	36	111	64	16	80	2	10	110	50	160	98	37	135
3	2	7	88	63	151	75	16	91	2	6	71	51	122	60	28	88	3	7	56	43	99	50	28	78
4	1	6	56	57	113	48	16	64	2	7	77	48	125	55	18	73	3	12	116	69	185	85	56	141
5	2	7	80	56	136	64	13	77	2	11	17	74	91	64	20	84	2	6	72	33	105	50	25	75
6	1	10	106	91	197	71	29	100	1	4	53	25	78	37	12	49	3	9	114	38	152	78	32	110
7	2	8	90	74	164	77	19	96	1	3	40	23	63	27	16	43	2	5	52	22	74	42	19	61
8	1	5	57	37	94	40	19	59	1	4	58	27	85	29	19	48	1	2	22	6	28	15	6	21
9	3	10	102	69	171	71	36	107	2	9	100	67	167	38	41	79	2	8	90	32	122	63	21	84
10	1	9	97	83	180	42	25	67	2	7	76	46	122	37	26	63	1	2	31	10	41	25	6	31
11	2	8	91	66	157	37	22	59	1	6	49	29	78	42	20	62	3	12	132	41	173	81	30	111
12	1	6	73	45	118	36	13	49	2	9	101	76	177	79	47	126	1	5	40	19	59	24	13	37
13	3	7	83	37	120	30	21	51	2	8	86	51	137	46	31	77	1	4	32	12	44	26	9	35
14	3	11	122	81	203	54	42	96	2	13	167	72	239	96	45	141	2	10	97	32	129	52	26	78
15	2	7	85	52	137	44	13	57	1	9	87	27	114	55	14	69	2	7	80	29	109	31	18	49
16	2	6	68	29	97	38	15	53	..	..	..	..	..	..	..	..	1	3	35	14	49	15	9	24
17	1	5	69	36	105	33	10	43	2	8	92	32	124	65	25	90	1	5	47	18	85	27	9	36
18	1	3	34	20	54	11	4	15	2	8	77	36	113	41	26	67	1	5	52	18	70	22	10	32
19	2	9	119	58	177	44	13	57	2	7	78	43	121	48	36	84	1	13	131	46	177	34	26	60
20	2	8	103	77	180	35	16	51	1	4	51	22	73	38	17	55	2	9	83	34	117	43	12	55
21	3	9	104	54	158	37	16	53	2	9	101	44	145	63	30	93	2	5	52	17	69	23	9	32
22	2	5	56	42	98	22	11	33	1	6	71	27	96	56	19	75	2	6	82	25	107	30	17	47
Total	42	161	1842	1210	3052	1042	388	1430	35	156	1622	952	2574	1119	533	1652	39	148	1554	618	2192	935	424	1359

% age affected Green leaves on the besis of Green leaves	56·57	68·99	60·17
% age affected Green leaves on the basis of total leaves	34·14	43·47	42·66
% age affected leaves on the basis of total leaves	46·85	64·18	62·00

Note. *G. L.* — Green leaves. *D. L.* — Dry leaves. *T. L.* — Total leaves.

For the whole plot % age affected green leaves = 58·51.

For the whole plot % age affected leaves = 49·43.

I

and otherwise) that existed in the plot—Plot (1)

Row 4								Row 5								Row 6							
Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected		
		G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.
3	13	104	33	137	72	1	73	1	7	84	55	139	68	5	73	2	12	138	63	201	120	30	150
2	6	48	16	64	32	0	32	2	7	75	44	119	52	10	72	2	5	51	27	78	44	12	56
2	7	56	14	70	28	0	28	1	7	69	46	115	49	13	62	1	6	64	88	152	54	17	71
2	4	16	9	25	23	0	23	2	4	41	26	67	28	5	33	1	10	136	55	191	107	15	122
2	5	35	17	52	18	2	20	2	9	106	66	172	62	8	70	2	5	49	25	74	42	5	47
1	4	18	13	31	4	3	7	1	7	77	56	133	26	11	37	2	6	56	38	94	46	1	47
2	12	110	64	174	66	43	109	2	8	84	54	138	39	7	47	2	5	33	34	67	27	7	34
2	10	114	42	156	51	25	76	3	10	100	71	171	36	17	53	3	7	60	45	105	54	9	63
2	7	72	19	91	35	12	47	1	3	32	30	62	16	6	22	3	7	67	49	116	54	5	59
1	7	48	20	68	34	2	36	2	10	110	84	194	37	9	46	3	8	67	44	111	53	6	59
1	6	42	24	66	12	10	22	1	4	46	29	75	20	3	23	3	6	35	28	63	33	2	35
2	6	45	15	60	16	9	25	1	5	50	37	87	21	7	28	3	12	53	45	98	76	16	92
2	8	72	31	103	13	15	28	1	4	39	27	66	8	4	12	2	3	31	16	47	29	4	33
2	8	63	38	101	18	11	29	2	7	52	40	92	12	1	13	3	11	83	53	136	61	7	68
2	8	64	25	89	8	0	8	2	6	53	48	101	13	2	15	3	9	77	34	111	57	6	63
2	6	30	14	44	8	0	8	2	8	74	58	132	14	4	18	1	1	11	6	17	6	0	6
2	8	48	22	70	3	0	3	1	3	19	21	40	7	0	7	1	5	37	24	61	30	2	32
1	5	24	9	33	4	0	4	2	11	88	80	162	22	1	23	2	8	49	89	88	33	3	36
1	4	25	6	31	5	0	5	2	5	44	37	81	14	0	14	2	7	55	39	94	36	0	36
2	5	32	10	42	5	0	5	2	9	87	86	173	22	0	22	2	3	20	18	38	13	0	13
1	5	32	15	47	6	0	6	2	7	80	67	147	13	0	13	1	3	19	6	25	15	2	17
1	4	28	5	33	5	0	5	1	5	45	31	76	8	0	8	2	5	36	14	50	15	1	16
38	148	1126	461	1587	466	133	599	36	146	1455	1093	2548	597	113	710	46	144	1227	790	2017	1005	150	1155
		41·39								41·03								81·91					
		29·36								23·43								49·83					
		37·74								27·86								57·26					

TABLE

Number of clumps, canes and leaves

Units	Row 1								Row 2								Row 3							
	Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected			Clumps	Canes	Total			Affected		
			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.
1	1	3	27	8	35	13	4	17	2	6	22	10	32	17	0	17	2	9	40	20	60	13	3	16
2	1	2	13	3	16	9	0	9	2	6	20	9	29	11	0	11	1	9	11	5	16	7	2	9
3	2	8	38	15	53	19	7	26	2	8	23	15	38	17	1	18	1	4	14	7	21	7	0	7
4	2	7	45	15	60	22	6	28	2	11	39	15	54	15	1	16	2	5	32	14	46	16	5	21
5	1	2	12	4	16	6	1	7	2	8	36	18	54	14	3	17	1	5	34	9	43	23	2	25
6	2	12	86	30	116	42	1	43	3	8	40	14	54	20	1	21	..	..	..	..	..	..	..	..
7	1	4	24	10	34	12	2	14	2	7	51	16	67	15	0	15	2	12	101	30	131	35	2	37
8	1	1	4	3	7	2	3	5	1	3	13	6	19	4	0	4	1	8	67	22	89	14	0	14
9	2	12	44	29	73	26	1	27	1	11	16	28	88	16	0	16	1	5	53	51	104	17	6	23
10	2	9	52	52	104	19	0	19	1	3	29	8	37	13	0	13	2	5	83	78	161	16	5	21
11	2	5	46	38	84	13	1	14	1	5	25	7	32	8	4	12	1	9	87	84	151	26	6	32
12	2	7	76	57	133	29	2	31	1	3	21	10	31	6	1	7	2	6	64	52	116	20	8	28
13	2	15	113	98	211	30	7	37	2	12	86	34	120	32	6	38	1	7	42	31	73	18	7	25
14	1	7	71	58	129	15	5	20	1	2	19	6	25	7	1	8	2	6	61	52	113	18	1	20
15	1	5	27	22	49	6	0	6	1	4	25	16	41	12	3	15	1	9	75	61	136	29	4	27
16	3	7	37	38	75	13	0	13	1	4	44	17	61	14	5	19	2	7	73	54	127	20	0	20
17	2	13	28	72	170	20	15	35	1	7	54	28	82	22	5	27	1	3	38	27	65	9	0	9
18	2	10	93	68	161	23	4	27	1	3	20	11	31	7	1	8	1	7	82	69	151	20	1	21
19	1	4	26	23	49	8	4	12	1	6	56	21	77	14	1	15	1	3	30	29	59	8	2	10
20	1	3	79	64	143	19	9	28	1	9	95	44	139	22	6	28	1	4	47	42	89	15	2	17
Total	32	136	1011	707	1718	346	72	418	29	126	778	333	1111	286	39	325	26	123	1034	717	1751	326	56	382

% age affected Green leaves on the basis of Green leaves ..	34·22	36·76	31·53
% age affected Green leaves on the basis of total leaves ..	20·14	25·74	18·62
% age affected leaves on the basis of total leaves ..	24·33	29·25	21·82

II

(*affected and otherwise*) *that existed in the plot—Plot* (2)

Row 4								Row 5								Row 6							
Clumps	Canes	Total			Affected			Canes	Clumps	Total			Affected			Clumps	Canes	Total			Affected		
		G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L.			G.L.	D.L.	T.L.	G.L.	D.L.	T.L
..	..	..	..	..	..	..	..	2	5	61	45	106	12	5	17	1	2	15	9	24	6	0	6
1	5	47	28	75	14	2	16	1	11	11	90	201	33	15	48	1	7	49	22	71	21	0	21
2	8	86	39	125	18	19	37	1	3	31	21	52	10	6	16	1	7	41	18	59	21	5	26
2	7	73	31	104	24	2	26	..	..	..	..	..	..	..	..	1	3	26	10	36	11	3	14
2	5	61	13	74	20	0	20	1	10	91	81	172	42	7	49	1	6	42	19	61	25	2	27
1	4	52	22	74	17	0	17	1	4	36	35	71	9	4	13	1	5	20	12	32	18	0	18
1	3	32	12	44	7	3	10	2	4	45	32	77	21	10	31	2	8	55	22	77	35	1	36
1	7	96	32	128	18	3	21	1	2	22	19	41	8	1	9	2	5	50	16	66	33	0	33
1	6	69	36	105	15	3	18	2	6	51	46	97	12	6	18	1	4	40	15	55	14	0	14
1	3	26	12	38	9	0	9	2	9	81	57	138	20	8	28	1	6	42	11	53	13	0	13
1	6	48	21	69	11	0	11	1	8	69	66	135	18	7	25	1	4	51	15	66	19	0	19
2	5	52	27	79	16	1	17	1	4	49	48	87	11	5	16	1	10	57	21	78	18	0	18
..	..	..	..	..	..	..	..	2	6	62	47	109	20	11	31	1	2	26	7	33	5	0	5
1	9	57	22	79	18	0	18	1	3	26	23	49	8	4	12	2	6	57	15	72	27	1	28
1	6	48	24	72	12	0	12	2	3	35	29	64	11	2	13	..	..	..	..	..	..	..	..
1	7	47	22	69	14	0	14	1	3	32	30	62	9	3	12	1	5	47	15	62	13	3	16
2	5	40	17	57	10	0	10	1	6	54	48	102	10	6	16	2	8	89	27	116	28	0	28
1	6	64	25	89	15	0	15	1	4	35	36	71	11	1	12	2	8	77	35	112	33	1	34
..	..	..	..	..	..	..	..	1	5	46	45	91	12	6	18	..	..	..	..	..	..	..	..
1	8	81	31	112	23	1	24	1	5	47	39	86	13	5	18	..	..	..	..	..	..	..	..
22	100	979	414	1393	261	34	295	25	101	984	827	1811	290	112	402	22	96	784	289	1073	340	16	356

26·66	29·47	43·37
18·73	16·01	31·69
21·18	22·20	33·18

Note. *G. L.* — Green leaves. *D. L.* — Dry leaves. *T. L.* — Total leaves.

For the whole plot % age affected green leaves = 33·20

For the whole plot % age affected leaves = 24·59

work. In the first plot, the two border units also have been included—thus making up the total number of units equal to 22 in each row whereas the second plot consists of 20 units only in each of its rows. The number of clumps in each unit, the number of shoots per clump and the number of dry and green leaves (affected and otherwise) of each shoot have been counted and noted in the records. Tables I and II will show the details in these respects. The records of the enumeration of the total puparia on each affected leaf and the measurements of its length and maximum breadth form the actual observations that are subjected to detailed analysis.

The distribution of the puparia in each leaf has also been shown on its sketch drawn roughly on squared paper. The details of all the affected clumps, canes and leaves in the individual units which have been available with the data are entered in Tables III and IV whereas Tables I and II give details of what existed in the plots.

The leaf area has been calculated by multiplying the product of the length and the breadth (at its widest point) of the leaf by 0·7 according to the formula established by Khanna (1935). The population per unit area (square inch) of the affected leaf has been found by dividing the number of puparia on the affected leaf by its area.

III. Sampling Procedure and the Theoretical Background

(*a*) *Fourfold sampling and the estimation of the required zone-variances.*— The investigation has been divided into two parts, the first part combining the analysis on units, clumps, canes and leaves while the second showing the sufficiency or otherwise of partial examination of a leaf.

The first part is taken to be a case of fourfold Nested (or Hierarchical) sampling with the units, clumps, canes and leaves as the zones of successive orders. The variance of a mean obtained from the nested sampling is a function of the zone variances the estimation of which is, therefore, primarily essential to decide the question of sample size. The theory of Mathematical Expectation affords us the means of finding these effective variances.

We shall, however, require in the present case, results up to the four-fold nested sampling and thus for ready reference the procedure for finding the variances is outlined for fourfold sampling only. Results on the two-fold and three-fold sampling will follow from symmetry:—

A variate in the fourth order zone may be defined as

$$x_{ijkl} = \mathrm{A} + b_i + c_{ij} + d_{ijk} + z_{ijkl},$$

TABLE III

Number of affected clumps, canes and leaves in the Indiviaual Units that are availcble in the Data Plot (1)

Row 1				Row 2				Row 3				Row 4				Row 5				Row 6			
U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf
1	3	9	74	1	2	11	97	1	1	3	7	1	3	13	70	1	1	7	72	1	2	12	48
2	2	6	67	2	2	6	73	2	2	10	131	2	1	4	15	2	2	7	71	2	2	5	54
3	2	7	87	3	2	5	66	3	3	7	78	3	2	7	28	3	1	7	62	3	1	6	70
4	1	6	43	4	2	7	64	4	3	12	126	4	2	4	23	4	2	4	33	4	2	10	115
5	2	7	77	5	2	11	82	5	2	6	74	5	2	3	19	5	2	9	70	5	1	4	47
6	1	10	95	6	1	3	33	6	3	9	105	6	1	4	7	6	1	7	37	6	2	6	46
7	2	8	94	7	1	3	41	7	2	5	58	7	2	12	88	7	2	8	46	7	2	5	32
8	1	5	59	8	1	4	47	8	1	2	20	8	2	10	68	8	2	9	51	8	3	7	61
9	3	10	107	9	2	9	76	9	2	8	84	9	1	4	31	9	1	3	22	9	3	7	56
10	1	9	57	10	2	7	63	10	1	2	31	10	1	6	28	10	2	9	46	10	3	8	58
11	2	8	58	11	1	6	60	11	3	12	105	11	1	6	20	11	..	..	..	11	2	5	35
12	1	6	40	12	2	9	106	12	1	5	37	12	2	6	25	12	1	5	27	12	3	12	89
13	3	7	49	13	2	7	76	13	1	4	33	13	2	8	21	13	1	4	12	13	1	2	32
14	3	11	96	14	2	13	139	14	2	10	74	14	2	8	26	14	2	4	13	14	3	11	68
15	2	7	51	15	1	9	69	15	2	7	49	15	1	2	2	15	1	5	15	15	2	8	63
16	2	6	52	16	..	..	..	16	1	3	23	16	2	4	8	16	2	7	17	16	1	1	6
17	1	5	41	17	2	8	88	17	1	5	34	17	1	3	3	17	..	..	..	17	1	5	32
18	1	3	14	18	2	7	66	18	1	5	32	18	1	3	4	18	2	8	22	18	2	8	36
19	2	9	57	19	2	7	82	19	1	13	57	19	1	3	5	19	2	4	14	19	2	7	36
20	2	8	51	20	1	3	38	20	2	9	54	20	1	3	5	20	2	8	22	20	2	2	13
21	3	9	51	21	2	9	88	21	2	5	32	21	1	4	6	21	2	6	12	21	1	3	17
22	2	5	33	22	1	6	75	22	2	6	45	22	..	..	..	22	1	3	8	22	2	4	16
Total 22	42	161	1353	21	35	150	1529	22	39	148	1289	21	32	117	502	20	32	124	672	22	43	138	1030
Average per unit:	1·91	7·32	61·50	1·67	7·14	72·81	1·77	6·75	53·59	1·52	5·57	23·90	1·60	6·20	33·60	1·95	6·27	46·81					

TABLE IV

Number of the affected clumps, canes and leaves in the Individual Units that are available in the Data Plot (2)

Row 1				Row 2				Row 3				Row 4				Row 5				Row 6			
U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf	U.	Cl.	Cane	Leaf
1	1	3	17	1	2	5	17	1	2	7	15	1	..	..	..	1	1	4	17	1	1	2	6
2	1	2	8	2	2	5	11	2	1	2	9	2	1	5	16	2	1	11	48	2	1	7	16
3	2	7	26	3	2	5	16	3	1	2	7	3	2	8	27	3	1	3	16	3	1	7	25
4	2	7	27	4	2	7	16	4	1	5	18	4	2	7	26	4	..	..	..	4	1	3	13
5	1	2	7	5	2	6	16	5	1	5	25	5	2	5	20	5	1	10	48	5	1	5	19
6	2	12	43	6	2	4	14	6	..	..	..	6	1	4	17	6	1	4	13	6	1	5	17
7	1	4	14	7	2	6	15	7	2	12	37	7	1	3	10	7	2	4	31	7	1	5	24
8	1	1	5	8	..	..	..	8	1	7	14	8	1	7	21	8	1	2	9	8	2	5	33
9	2	10	26	9	1	5	11	9	1	5	23	9	1	6	18	9	2	5	18	9	1	4	13
10	2	7	19	10	1	3	12	10	2	5	21	10	1	3	9	10	2	8	27	10	1	5	13
11	2	5	14	11	1	3	11	11	1	9	31	11	1	5	10	11	1	8	25	11	1	4	19
12	2	7	31	12	1	3	7	12	1	6	28	12	2	5	17	12	1	4	16	12	1	6	18
13	2	12	36	13	2	10	38	13	1	7	22	13	..	..	..	13	2	6	28	13	1	2	5
14	1	6	20	14	1	2	8	14	1	6	20	14	1	6	18	14	1	3	12	14	2	6	28
15	1	3	6	15	1	4	15	15	1	9	27	15	1	6	12	15	1	2	13	15	..	..	..
16	3	6	13	16	1	4	18	16	2	7	20	16	1	3	14	16	1	3	12	16	1	5	16
17	2	10	32	17	1	6	27	17	1	3	9	17	2	4	10	17	1	5	16	17	2	8	28
18	2	9	27	18	1	3	8	18	1	7	20	18	1	6	15	18	1	4	12	18	2	8	34
19	1	3	12	19	1	6	10	19	1	8	10	19	.	..	..	19	1	5	16	19	..	..	..
20	1	3	13	20	1	9	28	20	1	4	17	20	1	8	21	20	1	5	18	20	..	..	..
Total 20	32	119	396	19	27	96	298	19	23	111	373	17	22	94	281	19	23	96	395	17	21	87	327
Average per unit :		1·60	5·95	19·80	1·42	5·05	15·68	1·21	5·84	19·63	1·29	5·53	16·53	1·21	5·05	20·79	1·24	5·12	19·24				

Note.— *U* = Units. *Cl* = Clumps.

where satisfying the conditions of homoscedasticity we have

$$E(b_i) = 0;\ E(b_i)^2 = \sigma_1^2;\ E(c_{ij}) = 0;\ E(c_{ij})^2 = \sigma_2^2;$$
$$E(d_{ijk}) = 0;\ E(d_{ijk})^2 = \sigma_3^2;\ E(z_{ijkl}) = 0;\ E(z_{ijkl})^2 = \sigma_4^2.$$

From these relations can be found the expectation of the mean and its variance which are respectively:—

(1) $E(X....) = A$

and (2) $V(X....) = \frac{\Sigma p_i^2}{p^2}\sigma_1^2 + \frac{\Sigma_i \Sigma_j p_{ij}^2}{p^2}\sigma_2^2 + \frac{\Sigma_i \Sigma_j \Sigma_k p_{ijk}^2}{p^2}\sigma_3^2 + \frac{1}{p}\sigma_4^2.$

where p is the total number of variates and p_i, p_{ij} and p_{ijk} are the number of the fourth order variates in the selected units of the first, second and third order zones respectively. In equal case (that is where the number of variates is equal from unit to unit in a zone), the formula (2) reduces to

$$V(X....) = \frac{\sigma_1^2}{t} + \frac{\sigma_2^2}{n} + \frac{\sigma_3^2}{m} + \frac{\sigma_4^2}{p}$$

where t, n and m are the total number of variates in the successive zones.

Let v_1, v_2, v_3 and v_4 represent the successive variances of the zones in the structure of the analysis of variance.

Then v_4 will be the unbiassed estimate of σ_4^2.

$(m-n)\,v_3$ do do of $\left\{p - \Sigma_i\Sigma_j \frac{\Sigma_k p_{ijk}^2}{p_{ij}}\right\}\sigma_3^3 + (m-n)\,\sigma_4^2.$

$(n-t)\,v_2$ do do of $\left\{p - \Sigma_i \frac{\Sigma_j p_{ij}^2}{p_i}\right\}\sigma_2^2$

$$+ \left\{\Sigma_i\Sigma_j \frac{\Sigma_k p_{ijk}^2}{p_{ij}} - \Sigma_i \frac{\Sigma_j \Sigma_k p_{ijk}^2}{p_i}\right\}\sigma_3^2 + (n-t)\,\sigma_4^2.$$

$(t-1)\,v_1$ do do of $\left\{p - \frac{\Sigma_i p_i^2}{p}\right\}\sigma_1^2$

$$+ \left\{\Sigma_i \frac{\Sigma_j p_{ij}^2}{p_i} - \frac{\Sigma_i \Sigma_j p_{ij}^2}{p}\right\}\sigma_2^2 + \left\{\Sigma_i \frac{\Sigma_j \Sigma_k p_{ijk}^2}{p_i} - \frac{\Sigma_i \Sigma_j \Sigma_k p_{ijk}^2}{p}\right\}\sigma_3^2 + (t-1)\,\sigma_4^2.$$

It should, however, be remembered here that the question of estimating the σ's arises only when the null-hypothesis is found to be not true.

(*b*) *Calculation of the sampling efficiency percentages so far as sampling in a unit is concerned.*—The random selection in the number of units having been made, the next question that arises is with regard to further sampling in the unit itself. The desirability of further sampling in clumps, canes and leaves from out of the selected units has to be tested on the merit of its efficiency. Lesser number of clumps in a unit, fewer canes in a clump and fewer

leaves in a cane will be preferred to complete enumeration only when the efficiency of the proposed sampling or the *information* afforded by it is high. So far as sampling in a unit is concerned, efficiency percentage may be defined as

$$\frac{100\left(\sigma_1^2+\frac{\Sigma_i\Sigma_j p_{ij}^2}{\Sigma_i p_i^2}\sigma_2^2+\frac{\Sigma_i\Sigma_j\Sigma_k p_{ijk}^2}{\Sigma_i p_i^2}\sigma_3^2+\frac{p}{\Sigma_i p_i^2}\sigma_4^2\right)}{\sigma_1^2+\frac{t}{n}\sigma_2^2+\frac{t}{m}\sigma_3^2+\frac{t}{p}\sigma_4^2}$$

The formula explains for itself what has been sought to be conveyed by sampling efficiency. The numerator in the formula is the mean variance in the unit where the σ's are substituted for by the respective estimated zone variances and the coefficients are based on the total number of variates that existed in the unit. The denominator is the mean variance in the unit based on the adopted sampling. This formula therefore takes account simultaneously of the number of variates and the respective variances and shows how near has been the size of the adopted sampling to the subpopulation existing in the unit.

IV. Formation of the Patches and Sampling in the Plots

Charts (1) and (2) show the distribution of the incidence over the 2 plots. It is quite interesting to see the way in which whitefly has distributed itself in different patches of intensity. Considering the following classifications for low, medium and severe incidence, the three grades of intensity are fairly distinct from one another except the formation of low patch in plot 2.

Grade	Number of puparia per square inch of affected leaves
Low	Less than 2·00
Medium	Between 2·00 and 5·51
Severe	Above 5·51

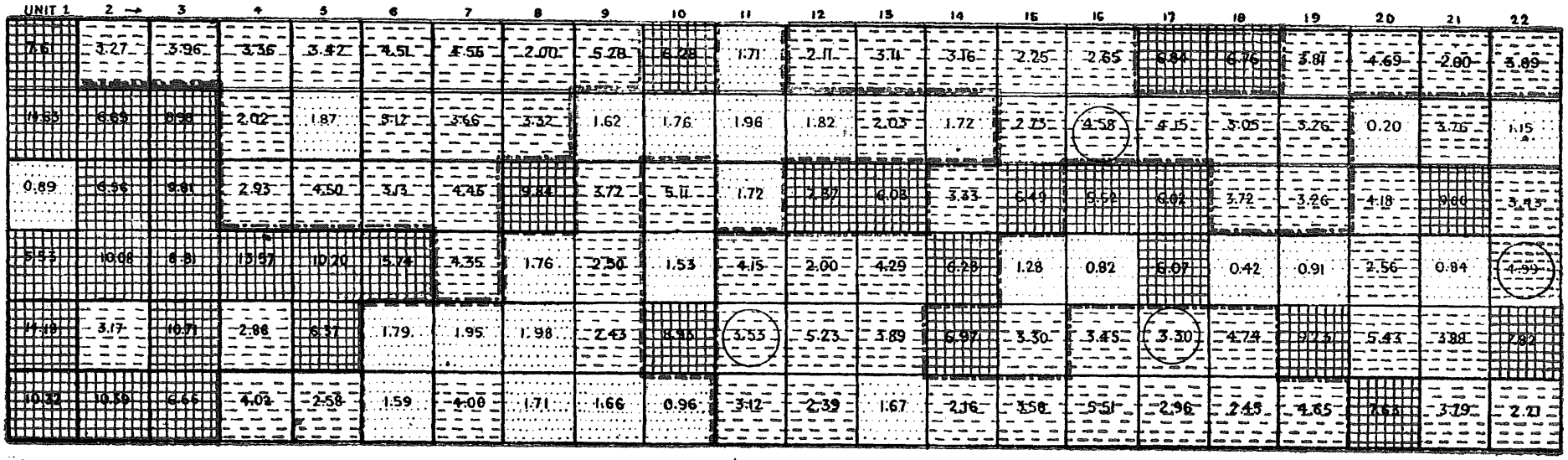
No puparia counts were available for 4 units in plot (1) and 9 units in plot (2) the incidence figures for which have been interpolated. The occurrence of two interpolated values in the low patch of plot (2) has partly impaired its well marked formation.

Sampling in plot (1).—Each of the patches so formed has been subjected to nested sampling and in each case 10 units have been chosen at random as the first order zones. Besides these, one more analysis in each plot has been done where 10 units have been selected at random out of the entire plot irrespective of the patches. This analysis will be referred to as the pooled analysis. In case of the low-patch in plot (1), the three units serving as the connecting links have not been considered for random selection.

PLOT No. 1

UNITS SHOWN ACCORDING TO THE INTENSITY OF INCIDENCE

Puparia per Sqr. Inch of Leaf Area



Less than 2.00

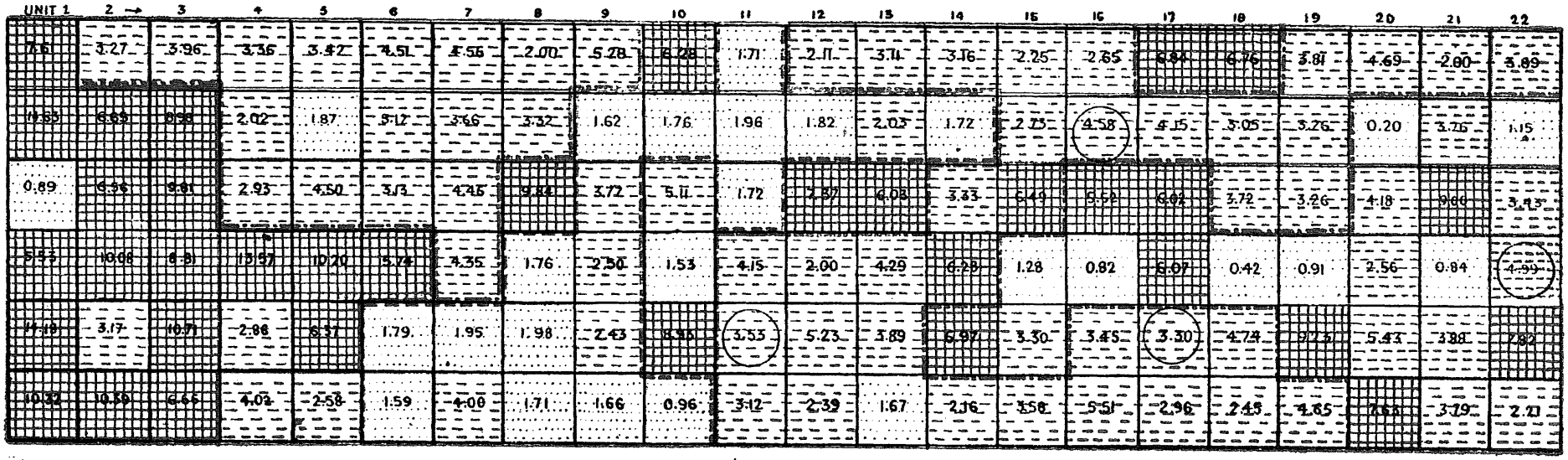
2.00 —— 5.51

Above 5.51

PLOT No. 1

UNITS SHOWN ACCORDING TO THE INTENSITY OF INCIDENCE

Puparia per Sqr. Inch of Leaf Area



Less than 2.00

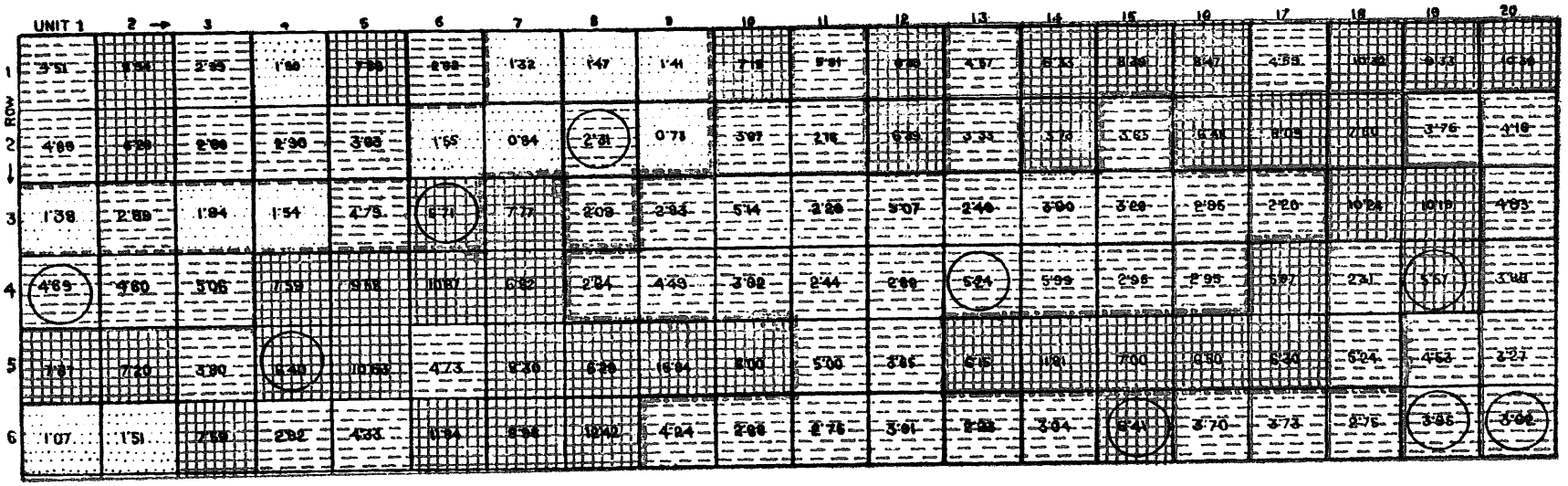
2.00 —— 5.51

Above 5.51

PLOT No. 2

UNITS SHOWN ACCORDING TO THE INTENSITY OF INCIDENCE

Puparia per Sqr. Inch of Leaf Area



Less than 2·00

2,00 ——— 5·51

Above 5.51

For the estimation of the zone variances and consequent analysis, two clumps per unit, two canes per clump and five affected leaves per cane have been successively selected at random in plot (1) in the 'low' and 'severe' patches. Owing to insufficient number of clumps per unit in the medium II-patch, only one clump at random from each unit has been included in the analysis. Again since the affected leaves per cane in the medium I-patch have not been sufficiently more than five in general, only 3 leaves per cane instead of five have been taken in this case to make allowance for sufficient choice in randomisation. In the pooled analysis, however, all clumps, all canes and all leaves have been included in the analysis for the estimation of the zone-variances.

Sampling in plot (2).—It will be apparent from Tables I and II that the crop growth in plot (2) is much poorer than that of plot (1) in respect of clumps per unit, canes per clump, leaves per cane and even in respect of the magnitude of leaf area. On account of this, sufficient choice for random selection of clumps, canes and leaves in the abovementioned proportions is not possible in plot (2). So in plot (2), all clumps, canes and leaves have been included in the analysis of variance to ensure sufficient degrees of freedom.

The question of transformation.—The ultimate variate used here is the mean number of puparia per square inch of the leaf surface, calculated on the basis of affected leaves. The question of some sort of transformation

Variances due to different zones

(With and without transformation)

Plot 1

	Mean with S.E.	σ_m exp. as % of mean	σ_1^2	σ_1^2 as % of total	σ_2^2	σ_2^2 as % of total	σ_3^2	σ_3^2 as % of total	σ_4^2	σ_4^2 as % of total
Severe	7·9131 ±0·643	8·1258	12·7883	9·8168	6·6294	5·0890	11·8286	9·0800	99·0236	76·0142
	2·60 ±0·061	2·3462	0·0150	4·9293	0·0049	1·6103	0·0282	9·2672	0·2562	84·1932
Medium	3·58 ±0·295	8·2402	0·0438	0·4164	0·00	0·00	0·00	0·00	10·4747	99·5836
	2·375 ±0·0026	0·1095	0·0085	4·6858	0·0053	2·9217	0·00	0·00	0·1676	92·3925
Low ..	1·6823 ±0·116	6·8953	0·00	0·00	0·00	0·00	0·3414	11·2968	2·6807	88·7032
	2·058 ±0·0019	0·0923	0·00	0·00	0·0021	1·35	0·0106	6·81	0·1430	91·84

N.B.—The upper figures are the results of the analysis of the actual incidence.
The lower figures are the results of the analysis of logarithm of (incidence × 100).

being needed for this variate has been examined and it is found that transformation is not essential in this case. Analysis of variance has been actually done in plot (1) with logarithmic transformation after multiplying each incidence figure by 100. The results are reproduced above.

It will be noticed in the above statement that leaves take away the maximum variation in both the cases and that there is also a distinct similarity in the distribution of the percentage variations in the other zones. There is therefore no point in making use of the transformation as the accruing advantage will not be of a considerably high order offsetting the extra inconvenience to be encountered in handling the transformed material and in interpreting the results in terms of the original units. Though not actual percentages the mean puparia per square inch very much resemble percentage figures where the necessity of a transformation is universally recognized. Even in such a domain transformation is not always needed. It is well known that no transformation would need to be made if the range in percentages is between about 25 and 75.

V. Discussion of the Results

(*a*) *High variability of the variate; largest variation due to leaves.*— Tables V and VI give the results of the analysis of variance in the different patches as well as in the pooled one. The estimates of the zone-variances in original and as percentages of the total variation are given at the bottom of Tables VII and VIII where $E(\sigma_1^2)$, $E(\sigma_2^2)$, $E(\sigma_3^2)$ and $E(\sigma_4^2)$ represent respectively the estimates of variances in units, clumps, canes and leaves.

A high degree of symmetry, however, exists in the two plots as will be noticed in the largest percentage variations due to leaves as also in the location of zero variances in some patches. σ_2^2 and σ_3^2 are zero respectively in the low and medium patches of both the plots. σ_3^2 in the medium II —patch of plot (1) expressed as percentage of the total variation comes out to oe 0·28 which also is negligibly small and may be taken to be of the zero order for all practical purposes. The percentage variation of σ_2^2 in the pooled analysis of plot (2) is 0 whereas the corresponding figure in plot (1) is 1·69 which also is obviously a small percentage.

In Tables VII and VIII will be found in respect of each type of patch, the errors expressed as percentages of the means corresponding to the different combinations of clumps, canes and leaves to be selected in a unit. The first three columns will show the selected combination, the next column gives the number of total leaves arising out of the combination in a unit and this expressed as percentage of the total affected leaves of the unit is entered in the adjacent column. The last column under each category will show the

TABLE V. *Results of the analysis of variance in the different patches—Plot* (1)

Due to	Low infested zone				Medium I				Medium II			
	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio
Between Units ..	9	17·8342	(V_1) 1·9594	0·73	9	99·0047	11·0050	1·05	..	..	..	..
Between clumps within a unit ..	10	40·7279	(V_2) 4·0738	1·52	10	79·0301	7·9031	0·76	..	..	..	..
Between clumps ..	19	58·5721	3·0827	1·15	19	178·0354	9·3703	0·90	9	74·9093	8·3233	1·637
Between canes within a clump ..	20	87·7590	(V_3) 4·3830	1·64	20	184·9874	9·2494	0·88	10	51·5952	5·1595	1·015
Between canes ..	39	146·3311	3·7521	1·40	39	363·0228	9·3082	0·89	19	126·5045	6·6581	1·390
Between leaves within a cane ..	160	428·9148	(V_4) 2·6807	..	80	837·9678	10·4746	..	80	406·7572	5·0845	..
Total ..	199	575·2459	..	..	119	1200·9906	..	..	99	533·2617	..	..

	Severe				Pooled			
	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio
Between units ..	9	4322·0375	(V_1) 480·2464	4·85	9	3020·8229	335·6476	9·054
Between clumps within a unit ..	10	2244·6077	(V_2) 224·4608	2·27	10	762·1847	76·2185	2·056
Between clumps ..	19	6566·6452	345·6129	3·49	19	3782·0076	199·1057	5·371
Between canes within a clump ..	20	3163·3303	(V_3) 158·1665	1·66	44	2499·6997	56·8114	1·532
Between canes ..	39	9729·9755	249·4865	2·52	63	6282·7073	99·7255	2·690
Between leaves within a cane ..	160	15843·7690	(V_4) 99·0236	..	488	18090·9912	37·0717	..
Total ..	199	25573·7445	..	..	551	24373·6985	..	..

TABLE VI. *Results of the Analysis of variance in the different patches—Plot* (2)

Due to	Low				Medium				Severe I			
	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio
Between units ..	9	300·5585	(V_1) 33·3933	12·146	9	111·5204	12·3911	1·426	9	720·1908	80·0213	3·152
Between clumps within a unit ..	5	7·8973	(V_2) 1·5795	0·574	2	38·9395	19·4697	2·242	3	232·0041	77·3347	3·047
Betwecn clumps ..	14	30844558	22·0326	8·013	11	150·4599	13·6782	1·574	12	952·1949	79·3496	3·126
Between canes within a clump ..	34	129·6351	(V_3) 3·8128	1·387	40	301·5856	7·5397	0·868	29	424·1010	14·6242	0·576
Between canes ..	48	438·0909	9·1268	3·319	51	452·0485	8·8637	1·020	41	1376·2959	33·5682	1·322
Between leaves within a cane ..	100	274·9634	(V_4) 2·7496	..	120	102·2829	8·6857	..	108	2741·2032	25·3815	..
Total ..	148	713·0543	..	..	171	1494·3314	..	..	149	4117·4991	..	..

Due to	Severe II				Pooled			
	D.F.	S.S.	M.S.	Ratio	D.F.	S.S.	M.S.	Ratio
Between units ..	9	1363·6460	(V_1) 151·5162	4·988	9	887·9436	98·6604	6·115
Between clumps within unit ..	2	7·4006	(V_2) 3·7003	0·122	3	66·2994	22·0998	1·369
Between clumps ..	11	1371·0466	124·6406	4·102	12	954·2430	79·5202	4·928
Between canes within a clump ..	38	1739·8522	(V_3) 45·7855	1·507	25	635·6103	25·4244	1·575
Between canes ..	49	3110·8988	63·4877	2·090	37	1589·8533	42·9690	2·663
Between leaves within a cane ..	154	4678·2193	(V_4) 30·3780	..	91	1468·2388	76·1345	..
Total ..	203	7789·1181	..	..	128	3058·0921	..	..

TABLE VII. *Error as percentage of the Mean in the different types of sampling with the corresponding sampling efficiencies (in percentage) so far as sampling in a unit is concerned and the estimates of the zone-variances*

Plot (1)

Low						Medium I			Medium II				Severe				Pooled		
Sampling Combination			No. of leaves	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	Sampling combination per clump		%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	No. of leaves	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age
Clump	Canes	Leaves								Cane	Leaf										
1	1	1	1	1·28	32·69	2·29	2·04	28·77	2·08	1	1	1·49	30·38	7·69	1	1·24	45·62	14·27	1·81	41·49	13·50
1	1	3	3	3·85	20·21	5·60	6·11	16·48	6·26	1	3	4·48	18·47	20·55	3	3·71	31·85	28·92	5·43	27·79	30·04
1	1	5	5	6·42	17·83	7·89	..	..	..	1	5	7·46	15·19	30·88	5	6·18	28·43	36·41	9·06	24·07	39·78
1	2	1	2	2·57	23·18	4·58	4·07	20·39	4·15	2	1	2·99	22·17	14·53	2	2·47	34·50	24·83	3·62	31·12	24·02
1	2	3	6	7·70	14·86	11·21	12·22	11·73	12·24	2	3	8·96	13·96	35·55	6	7·42	25·91	44·43	10·87	22·11	47·07
1	2	5	10	12·84	12·48	15·77	..	..	..	2	5	14·92	11·90	50·02	10	12·36	23·76	52·75	18·12	19·96	58·25
2	1	1	2	2·57	23·18	4·58	4·07	20·39	4·15							2·47	33·74	25·98	3·62	30·92	24·39
2	1	3	6	7·70	14·86	11·21	12·22	11·73	12·24							7·42	24·77	48·25	10·87	19·77	48·50
2	1	5	10	12·84	12·48	15·77	..	..	..							12·36	22·62	58·23	18·12	19·57	60·44
2	2	1	4	5·13	16·05	9·16	8·15	14·53	8·23							4·94	26·41	42·42	7·25	24·07	40·35
2	2	3	12	15·40	10·70	22·42	24·44	8·38	23·90							14·83	20·85	68·05	21·74	18·40	68·53
2	2	5	20	25·67	8·92	31·54	..	..	..							24·72	19·59	77·40	36·23	17·03	79·65
					5·35			4·47					7·39				19·08			16·83	

	E (61^2)	E (62^2)	E (63^2)	E (64^2)	Variances	Improved	As %age	Variations
Low ..	0	0	0·3414	2·6807	0	0	11·30	88·70
Medium I ..	0·0438	0	0	10·4747	4·16	0	0	95·84
Medium II ..	..	8·3164	0·0150	5·0845	..	5·84	0·28	93·88
Severe ..	12·7883	6·6294	11·8286	99·0236	9·82	5·09	9·08	76·01
Pooled ..	4·8159	0·7584	2·2551	37·0717	10·73	1·69	5·02	82·56

TABLE VIII. *Error as percentage of the mean in the different types of sampling with the corresponding sampling efficiencies (in percentage) so far as sampling in a unit is concerned and the estimates of the zone-variances—Plot (2)*

Low							Medium I			Severe I			Severe II			Pooled		
Sampling combinations			No. of leaves	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age	%age leaves sampled	Error as %age of the mean	Sampling efficiency as %age
Clump	Cane	Leaf																
1	1	1	1	6·71	36·86	44·01	5·81	27·22	13·28	6·67	27·95	24·47	11·90	26·31	18·07	7·75	33·94	30·50
1	1	3	3	20·13	29·48	68·76	17·44	17·42	32·50	20·00	19·69	49·23	14·71	18·36	37·08	23·26	25·64	53·42
1	1	5	5	33·56	27·77	77·48	29·07	14·67	45·74	35·33	17·57	61·77	24·51	16·37	46·96	38·76	23·43	64·01
1	2	1	2	13·42	30·66	63·57	11·63	20·31	23·88	13·33	22·03	39·26	9·80	19·86	31·90	15·50	26·63	49·56
1	2	3	6	40·27	26·38	85·91	34·88	13·90	50·93	40·00	17·00	65·98	29·41	14·66	58·28	46·51	21·40	76·78
1	2	5	10	67·11	25·45	92·40	58·14	12·22	65·88	66·67	15·80	76·37	49·02	13·41	69·58	77·52	20·19	86·26
2	1	1	2	13·42	30·66	63·58	11·63	19·24	26·55	13·33	19·76	48·84	9·80	19·86	31·90	15·50	26·63	49·56
2	1	3	6	40·27	26·38	85·91	34·88	12·33	64·99	40·00	13·92	98·46	29·41	14·66	58·28	46·51	21·40	76·78
2	1	5	10	67·11	25·45	92·40	58·14	10·38	91·48	66·67	12·43	..	49·02	13·41	69·58	77·52	20·19	86·26
2	2	1	4	26·85	27·05	81·75	23·26	14·37	47·72	26·67	15·58	78·42	19·61	15·56	51·68	31·01	22·07	72·09
2	2	3	12	80·54	24·68	98·15	69·77	9·83	..	80·00	12·02	..	58·82	12·41	81·60	93·02	19·02	97·14
2	2	5	20	..	..	..	..	..	..	..	..	..	98·04	11·67	92·28	..	..	..
					25·29			11·07			12·25			12·54			20·21	

	E (61^2)	E (62^2)	E (63^2)	E (64^2)	Variances	Expressed	As %age	Variations
Low ..	1·9584	0	0·3848	2·7496	38·45	0	7·56	53·99
Medium ..	0	1·1038	0	8·6857	0	11·28	0	88·72
Severe I ..	0	8·2031	0	25·3815	0	24·43	0	75·57
Severe II ..	5·2465	0	3·8748	30·3780	13·28	0	9·81	76·91
Pooled ..	5·6889	0	2·8302	16·1345	23·07	0	11·48	65·45

percentages of sampling efficiency corresponding to the individual types of combination, efficiency being calculated on the same procedure as outlined in Section 3 (*b*). While considering the results in plot (2), it should be remembered that though the average number of canes per unit is mostly greater than 4, the average number of clumps per unit is hardly equal to 1·5. Combination with two clumps per unit, therefore, will necessarily have its limitation. The size of the adopted sampling in many of the combinations will cover nearly all the available clumps and as such will be nearer the total elements available in the units from which samples were drawn. This is a reason why efficiency percentage in plot (2) has sometimes been very high. The line corresponding to the combination of 2 clumps $\times$ 2 canes $\times$ 5 leaves has been kept blank since the average number of leaves per unit is in no case equal to 20.

It will be noticed in Tables VII and VIII that the largest variation is due to leaves. In plot (1) the variation ranges from 76·01 per cent. to 95·84 per cent. while in plot (2) it ranges from 53·99 per cent. to 88·72 per cent. This obviously indicates that the totality of leaves is the most predominant factor in the effective reduction of the error of the estimate. The error percentage column will give an idea as to how the error falls off with the increase in the number of leaves. It demonstrates, therefore, the relative merit of the different types of sampling combination so far as reducing the error is concerned.

The contributions that can be made by zones other than leaves have fully reflected themselves in the errors of the different sampling combinations. The trend of errors under these combinations which have been sufficient in number would throw light on the potentialities of the different zones in the error reduction.

If the estimated variances are taken as truly representative of the zone variances, reduction of the error may be carried beyond the size of the sample and extended up to the total number of leaves, canes and clumps that were available in the 10 units. The last (isolated) figure shown below the error percentage column represents this minimum reducible error in the different patches. The lowest of these figures in plot (1) is 4·47% and occurs in the medium 1-patch whereas the highest value 19·08% occurs in the severe patch. In the pooled analysis, the minimum error has come out to be 16·83% of the mean. These figures amply demonstrate the highly variable nature of whitefly incidence from leaf to leaf. That the variation is really very wide will be clear from the fact that in case of 19·08% and

16·83% that is in the severe and the pooled patches the number of leaves available in the units are respectively 809 and 552.

Owing to less number of leaves occurring in plot (2), the minimum reducible error in this plot has been of still higher order. The sample size in leaves varies from 129 in the pooled analysis to 204 in the severe II-patch, the corresponding minimum errors being 20·21% and 12·54%, tne lowest error being 11·07% in the medium patch for a sample size of 172 leaves. These minimum errors are more important in plot (2) in so far as they will fix up the upper limits of the sampling combinations, to be adopted. In the error column no figure should be less than this minimum error. In case where the minimum error is greater than any figure in the error column, the corresponding sampling combinations is not valid in the particular instance but it has none the less been entered to show the extent ana possibility of error reduction.

(*b*) *Varied sampling requirements under different grades of intensity.*— Examinatien of Table VII—particularly the percentage variations of the zones will reveal that when sampling has to be done in a region of low or medium infestation, the sampling selection may be confined to leaves alone. Sampling need not be of the nested type. That is to say, the affected leaves of all the clumps (or canes) in a unit may alone be subjected to random selection without any consideration to see whether or not the leaves drawn represent all the clumps in a unit or all the canes in a clump. In such a case, the variation due to leaves is the only variation that matters. But in case of severe infestation or in cases where it is fairly high (represented by the type of pooled analysis), sampling procedure should necessarily be of the Nested-type so that the clumps and canes in a unit may be equitably represented. Percentage variations of the different zones as given in Table VIII also exhibit the same results in respect of plot (2) except in the low-patch where the results are in favour of Nested sampling. It has been already taken into notice that the formation of the low patch in plot (2) has not been well enough nor has it been sufficiently large.

Another important feature that has come out prominently is the occurrence of the lowest error under the medium patch in both the plots. Therefore, for a required error percentage, the sample size to be needed in such a region will be comparatively small. Fixing the limit of the error to be near about 15% of the mean, the following table is prepared to furnish the information on the relative requirements of the sample size under different grades of infestation. Percentages of leaves required in the medium patch to those required on the average in the severe and pooled patches form roughly about 23 per cent. in plot (2) and 11 per cent. in plot (1).

Classification	Plot 1					Plot 2				
	Unit	Clump	Cane	Leaf	Error %	Unit	Clump	Cane	Leaf	Error %
Low ..	10	2	1	3	14·86	15	2	2	5	14·98
Medium (I) ..	10	2	2	1	14·53	10	1	1	5	14·67
Medium (II)· ..	10	1	1	5	15·19	..	..	..	..	..
Severe (I) ..	15	2	3	6	14·89	10	2	2	1	15·58
Severe (II) ..	..	..	..	..	..	10	2	1	3	14·66
Pooled ..	13	2	2	5	14·98	15	2	2	5	14·98

(*c*) *Percentage area to be sampled.*—The 3 ft. units of the plots served as the first order zones. Random selection in the units gives representation to the spread of the plot area. For the sake of uniformity, 10 units have been selected from each patch as well as in the pooled analysis which covering as it does the whole plot, will be of practical interest. Table IX will show that except for Medium II-patch in plot (1) the difference between the whole-patch mean and the one obtained from the sample is fairly within the limits of the standard error. In the pooled analysis ten 3 ft. units form respectively 7·81% and 9·01% of the total number of units available in the plots. The corresponding error percentages are, however, 16·83 and 20·21 in the two cases. If 15 units (about 13·51% of the total units) be selected in plot (2) with the same combination, (2 clumps × 2 canes × 5 leaves per unit), the error will reduce from 20·21 per cent. to about 15 per cent. whereas in plot (1) 13 units (forming about 10·16%) will bring down the error from 16·83 per cent. to roughly about 15%. Further reduction in the error per cent. will necessitate more units to be taken in the sample. Variability of incidence of white-fly is of such a magnitude that if the estimate is required within an error of 5 per cent., almost all the units of the plot (1/40th acre) will have to be sampled. 5 per cent. in error therefore will appear to be too much to expect in practice in a plot of 1/40th acre. 10 per cent. of units from an $\frac{1}{4}$th acre plot is likely to reduce the error in estimation to about 5 per cent. provided, however, the variation in incidence in the larger plot (such as $\frac{1}{4}$th acre) does not correspondingly increase in dimension.

In plot (1), 20 leaves (2 clumps × 2 canes × 5 leaves) in the pooled analysis have formed about 17 per cent. of the affected leaves. In case where the combination, 2 clumps × 2 canes × 5 leaves is not available in a unit in the plot, about 20 leaves equitably representing the clumps and canes should be taken. Alternatively, about 17–20 per cent. of the affected leaves (*i.e.*, one in every five) per unit should be taken if this gives larger number of leaves in the sample.

TABLE IX. *Sample-mean and the mean from the whole patch—Plot* (1)

Plot I						Plot II					
Grades	Mean for the whole patch	Sample mean	Diff. as %age of the mean	S. E. as %age of the mean	%age units sampled	Grades	Mean for the whole patch	Sample mean	Diff. as %age of the mean	S. E. as %age of the mean	%age units sampled
Low ..	1·87	1·68	11·31	8·92	58·82	Low ..	1·82	1·94	6·19	23·39	83·33
Medium I ..	3·53	3·58	1·40	8·38	29·41	Medium ..	3·50	3·63	3·58	11·67	29·41
Medium II ..	3·65	2·44	50·83	11·90	55·56	Severe I ..	7·50	6·56	14·33	15·25	47·61
Severe ..	8·27	7·91	4·55	15·69	47·62	Severe II ..	7·41	7·55	1·85	12·54	45·45
Pooled ..	4·38	5·11	14·29	16·83	7·81	Pooled ..	5·08	4·63	9·72	20·21	9·01

It may be pointed out that the sampling procedure evolved in this note would strictly apply to a plot of 1/40th acre in size, but the contribution to error made by plots or by any other upper hierarchies will perhaps be small for practical purposes as the maximum of the total variation has been consumed by leaves in all the patches except in the low patch of plot (I). Study of the material divided into patches of low, medium and severe infestation has thrown sufficient light on the extent of variations that might possibly exist in a bigger plot or field where the infestation has to be either low or medium or severe. Results of investigations made in this direction will be furnished in a subsequent communication.

VI. Results on the Number of Inch-Units to be Selected in a Leaf

It now remains to see whether partial enumeration of a leaf will serve the purpose of estimation with fair accuracy. It is difficult indeed to select random square inch units from the area of the leaf since the breadth of the leaf is not uniform throughout. If a leaf is divided into inch units along its length each unit covering the entire breadth of that portion of the leaf, the units so formed will more easily lend themselves to random selection. If these units be chosen at random in sufficient number, the total area of these units may be calculated for practical purposes exactly in the same way as the net area of the leaf itself was found for working out the puparia per square inch.

For the purpose of estimating the variances in the present analysis, two leaves per unit and 10 random inch-units per leaf have been selected at random from each of the regions of low, medium and severe infestation.

Table X will show the estimates of the variances between leaves as well as within leaves. The last line in the table gives the standard error expressed as percentage of the mean; the sample in each case consists of 20 leaves and 200 inch units in all.

The calculation of the puparia per square inch of the leaf-area has already been found out. Now, the puparia per square inch of the leaf has been found from consideration of the area of these random inch-units, which are 10 in number in each case. The net area of these 10 units has, however, been found by the same process as was used in finding out the area of the leaf itself. The puparia per square inch of the leaf worked as such as well as the puparia found from consideration of the entire leaf area, have been shown side by side in Table X. The difference in between these two estimates has been nowhere significant as revealed by the 't' tests. It, therefore, justifies the method used in finding the net area of these random units besides showing that 10 random units are sufficient in giving an estimate

TABLE X. *Leaf-wise puparia per square-inch on the basis of the affected area of the leaf as well as on the basis of 10 random inch-units with the 't'-values of the differences in each patch (or zone); Estimates of the variances and the error per cent. of the estimate of incidence—Plot (1), Plot (2)*

Leaf No.	Low		Medium		Severe		Pooled		Low		Medium		Severe		Pooled	
	Pup. per sq. in. of leaf	Pup. per sq. in. on the basis of 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units	Pup. per sq. in. of leaf	Pup. per sq. in. from 10 units
1	2·97	3·14	1·31	0·67	4·93	6·73	12·45	17·01	2·41	2·57	14·82	2·03	8·31	4·93	6·51	1·29
2	0·99	0·91	7·85	6·67	16·38	10·00	3·84	3·81	0·63	1·71	2·61	6·86	2·98	3·29	0·76	0·89
3	1·29	1·43	2·35	2·98	4·22	3·00	0·27	0·12	0·48	..	3·77	3·71	19·35	22·00	9·18	11·60
4	2·30	2·27	5·70	5·40	5·79	0·55	0·97	1·00	0·20	0·18	2·18	1·34	2·39	3·06	6·14	6·14
5	0·31	0·22	0·64	0·71	4·12	3·69	0·54	1·29	1·03	0·63	1·82	1·30	4·95	6·19	2·98	2·86
6	1·08	1·57	1·45	1·79	6·01	4·29	13·25	8·29	0·57	0·57	3·23	1·67	2·64	2·86	3·21	3·86
7	1·91	1·75	2·17	3·00	2·60	6·12	4·00	5·97	0·50	0·91	5·39	3·78	2·66	0·95	1·60	2·14
8	1·26	2·29	1·35	1·69	4·91	6·71	21·18	18·30	1·29	1·86	2·89	3·65	4·40	5·36	6·11	0·79
9	1·39	1·57	1·10	0·71	34·08	50·71	35·87	32·72	7·48	8·81	1·08	1·63	2·95	1·61	5·82	4·60
10	0·86	0·39	0·62	0·90	22·64	17·57	6·41	7·43	1·08	2·29	4·57	5·00	3·39	2·74	2·64	2·29
11	0·63	..	4·32	4·29	19·36	13·39	2·07	1·82	0·22	0·10	6·57	7·14	7·00	5·29	15·56	30·35
12	1·42	1·81	0·64	1·07	2·88	2·57	0·22	0·29	1·04	0·89	0·99	0·11	27·53	8·80	0·48	0·24
13	0·46	1·54	0·90	0·82	7·61	12·22	0·22	0·36	4·14	5·45	0·49	0·95	4·52	3·21	15·66	15·71
14	1·15	1·05	1·28	0·45	6·21	9·17	3·23	3·21	1·39	1·43	4·33	5·71	1·98	0·82	4·62	6·00
15	2·84	1·10	1·75	1·43	7·06	7·29	6·86	6·00	0·69	0·29	1·51	2·86	7·32	6·35	2·93	2·62
16	0·87	7·79	1·17	1·00	7·87	8·14	2·08	1·05	3·38	2·21	3·10	2·26	1·75	2·00	1·28	2·38
17	8·99	9·21	8·05	6·75	5·02	5·45	1·10	0·20	1·08	0·41	2·15	1·87	10·31	5·31	2·07	0·92
18	1·89	1·56	6·69	5·83	5·22	7·43	1·47	1·21	1·71	1·43	0·67	1·02	3·20	33·90	1·08	2·32
19	1·63	1·82	1·69	0·82	2·54	2·60	3·69	1·96	0·67	0·86	2·31	3·57	6·32	7·43	1·90	3·02
20	0·56	0·54	9·13	7·98	5·01	1·43	4·12	3·33	0·44	0·79	2·16	2·86	12·79	20·32	3·32	2·43
Total	34·80	50·14	60·10	54·96	174·46	179·06	122·84	115·37	30·43	33·39	66·64	59·32	136·74	146·51	93·85	102·45
't' ..		1·437 N.S		1·82 N.S.		0·205 N.S.		0·883 N.S.		0·970 N.S.		0·512 N.S.		0·254 N.S.		0·876 N.S.
E (σ_1^2) ..		1·2712		2·7217		50·6402		38·3455		1·296		9·9252		35·6738		5·0888
E (σ_2^2) ..		2·0772		7·9630		48·6444		32·7656		1·513		8·5010		30·0040		12·3900
Error % of mean		21·59		21·27		26·28		32·77		11·42		27·08		26·97		20·40

reasonably near the true value leaf by leaf. But the error per cent. with only 20 leaves in the whole patch and 10 units per leaf is not sufficiently low as will be noticed in the last line of the table. If instead of 2 leaves per unit, as many as 20 leaves per unit be selected (as recommended before) and then 10 inch units (about 25 per cent. of the inch-units) from each leaf be chosen at random, the error per cent. will come down to near about 5 per cent.

But from practical point of view, the selection of 10 random inch-units in a leaf followed by the counting of the puparia unit by unit will probably be more difficult than counting the puparia of the entire leaf itself. Even if these two methods involve almost equal labour, enumeration of the whole leaf is preferred since complete enumeration of a leaf will avoid the possibility of the further error introduced by sampling.

VII. Summary of Conclusions

(i) The incidence of white-fly is highly variable and the maximum variation occurs in leaves. The variations between the clumps and the canes if and when they exist are of a much lesser magnitude. Therefore, increase in the number of leaves alone in a sample is pre-eminently effective in the reduction of error of an estimate.

(ii) In a field where the infestation is either high or fairly high (puparia per square inch of the affected area more than 5·00), Nested sampling has to be adopted so as to ensure the equitable representation of clumps, canes and leaves in a sample. In such a field, 10–13 per cent. of the 3 ft. units should first be selected and then Nested sampling done in each unit in that particular form of the alternatives, which gives the larger number of leaves in the sample, namely 2 clumps × 2 canes × 5 leaves per unit or 20 affected leaves equitably representing the clumps and canes in a unit as far as possible or 18-20 per cent. of affected leaves (*i.e.*, about one in every five) equitably representing the clumps and canes in a unit.

(iii) In a plot of low or mild infestation, Nested sampling is not essential. Only the leaves as a whole may be subjected to random selection, other requirements being the same as mentioned under number (ii) above.

(iv) For a required error percentage, a medium infested field (puparia, 2·00–5·51 per square inch) requires a smaller sample size than that needed in a plot of high infestation or of an infestation which is fairly high. A difference in the sample size in the number of leaves by even about 75% gave the same efficiency. This suggested inherent homogeneity in variation occurring in mild form of infestation.

(v) If in each leaf, 10 random inch units (about 25% of the total inch units) be taken instead of enumerating the whole leaf, the estimate of the incidence will be fairly reliable with a very slight additional error; but for practical consideration and working convenience, the complete enumeration of a leaf may be preferred to counting puparia in 10 random inch-units.

VIII. Acknowledgments

The work was carried out as part of the Sugarcane Research Scheme in Bihar being financed jointly by the Bihar Government and the Imperial Council of Agricultural Research, to whom grateful thanks are due. Sincere appreciation of the facilities afforded to field staff by the management of the Rohtas Industries Ltd., Dalmianagar in whose area the field study was made, is also recorded. Similarly assistance rendered in field collection and population counts by Mr. A. C. Sen, Senior Entomological Assistant, is acknowledged.

References

1. Cochran, W. G. .. "Some difficulties in the statistical analysis of replicated experiments," *Emp. Jour. Exp. Agri.*, 1938, 6, 157–75.
2. ———— .. "The use of the analysis of variance in enumeration by sampling," *J. Amer. Statist. Assn.*, 1939, 34, 492–510.
3. Khanna, K. L. .. *Bull* 6, *Department of Agric.*, Bihar, 1935, 17.
4. Roy, S. N., and Banerjee, K. S. .. "On Hierarchical sampling, Hierarchical variances and their connection with other aspects of statistical theory,"—*Science and Culture*, 1940, 6, 189.
5. Yates, F., and Zacopanay; I. .. "The estimation of the efficiency of sampling with special reference to sampling for yield in cereal experiments," *J. Agric. Sci.*, 1935, 25, 545–77.