TRAN. INST. INDIAN GEOGRAPHERS
Vol.23
No.1&2
2001

Cartographic Representation of the Results of Chi-Square Analysis for Spatial Explanation

SONALI SRIVANT AND
V. S. PHADKE

Abstract : *The quantitative techniques used for analysis in geography have been derived from mathematics or statistics. Some of them such as measures of central tendency or regression have been suitably amended to serve the needs of spatial analysis. The others such as chi-square or correlation continue to be discussed in an asptial manner. This makes the explanation incomplete. The present paper makes an attempt to suggest ways in which one of the techniques, viz. chi-square, could be amended to serve the needs of spatial explanation. It also indicated the need to supplement the result of quantitative analysis by mapping to provide spatial explanation.*

Introduction

Quantification is in vogue in geography over the last fifty years. Most of the techniques used for analysis in the discipline have been derived from statistics – descriptive as well as inferential, although some have come from mathematics. Some of these, such as, measures of central tendency, dispersion, regression, factor analysis have been suitably amended to serve the needs of spatial analysis (Sviatlovsky, 1937; Robinson et al, 1957; Taaffe et al, 1963; Berry, 1966; Murdie, 1969). Taaffe (1970) provides a short but valuable insight into some of the techniques used for spatial interpretation. Prior to this details were usually obtained form books of Yeates (1974) and Taylor (1977). Interpretation of the results of some other techniques such as Chi-square or correlation still continue to be discussed in a conceptual and functional manner. Some earlier attempts have been made to use spatial data for Chi-square or correlation exercises (Mackay, 1958; Daiton et al, 1972). McQuitty (1957) has also suggested a grouping technique, using correlation values to identify similarity among variables or areal units. No attempt has yet been made at spatial explanation of the results of these techniques. In fact, while learning these techniques, the students also face another difficulty, since the teacher generally starts with a readymade table; the students are hardly aware of the sources of the data used, nature of conditions prevailing in the field etc., with the result the explanation arrived at does not yield a very realistic picture of the situation. Since geography is a spatial discipline, there is need to discuss explanation in spatial terms. Such an interpretation, however, requires representation of the results on a map and collection of certain additional information. Though this calls for extra efforts, it also makes the interpretation more articulate and relevant. The primary objective of the present exercise is to suggest ways in which a more realistic interpretation of ground reality by spatial representation can be done. For this purpose Chi-square test has been selected as an example.

Data Source and Methodology

The discussion being mainly methodological, the study area selected is hypothetical and the data comprises rainfall and yield values for 100 randomly selected plots (Table 1). These data were plotted on a scattergram and the patterns of natural breaks were identified. Based on these natural breaks, the rainfall and yield values were grouped into three classes namely : high, moderate and low. The data were then represented on a map (refer to Figure 1). It was observed that the two variables did not produce identical spatial patterns. The data were then cross–tabulated to derive a frequency table (Table 2). This too did not yield a very clear pattern of distribution, though some trends towards a positive relationship could be identified. Based on these observations relevant hypotheses were formulated.

Research Hypothesis(H) : The amount of rainfall and the yield of rice are closely related.

Null Hypothesis(Ho) : There is no relationship between the amount of rainfall and the yield of rice.

The implication of the above null hypothesis was follows :

Implication(I) : The plots are proportionately distributed in different categories of rainfall/yield.

The Chi-square test was applied to verify the above hypotheses. The calculated value of Chi-square was 10.66.

Functional Interpretation

By referring to the table of critical values for Chi-square, it was observed that at four degrees of freedom, the critical value at 95% confidence level was 9.488, while that at 99% confidence level was 13.277. This means that the calculated value is significant at 95% but not at 99% confidence level. In other words, the null hypothesis could be rejected at 95% confidence level. Rejecting the null hypothesis at 95% confidence level meant that there existed a correlation between rainfall and yield of rice although this could be said with less confidence. The nature of this relationship could not, however, be determined by the Chi-square value itself and, therefore, a reference to the original table was made.

From Table 2, it was observed that the highest number of plots under the high yield category felt under high rainfall category, those under moderate yield received moderate rains and those under low yield were found to be associated with low rains. This suggested that there was a positive relationship between rainfall and the yield of rice.

A positive relationship between these two variables was expected, as rice is a water – intensive crop and requires standing water. Higher amounts of rain ensures the availability of the necessary water requirements, under natural conditions. A reference to the table also indicated that all the plots did not hold the general relationship, because, if that were so, all the plots with high yield would have fallen under high rainfall category, all those with moderate yield would have received moderate rains, while those under low yield would have got clubbed under low rainfall category. But this was not the case and some significant departures were observed. For examples, six plots with high yield received low rains. This was probably because of the availability of irrigation facilities, as they could overcome deficiency of rainfall in a much more effective way. Seven plots under high rainfall category showed low yield. This indicated that other factors like slope, soil fertility, soil depth also influenced the yield. Departures from the normal require explanation in terms of extraneous factors, which had not been taken into consideration during Chi-square analysis, but did influence the yield of rice. Such factors could include nature of seeds used, fertilizer amount, type of farming practices employed, seriousness of the farmer in maintaining the schedule and others.

It was thought appropriate that plots found in areas of high/moderate/low yield respectively did not require any explanation, as their behaviour was normal; this is quite evident in the positive nature of correlation. The emphasis was, therefore, placed on explaining abnormal behaviour of the other

Table 1 : Malegaon Vicinity – Attributes of Experimental Plots

Sr. No.	Rainfall in mm	Yield in kg	Fertilizer Status	Farming Tech.	Sr. No.	Rainfall in mm	Yield in kgs	Fertilizer Status	Farming Tech.
1	425	1700	M	Mo	51	2400	3250	H	Mo
2	665	1900	M	Mo	52	2350	3600	H	Mo
3	630	3200	M	Mo	53	2210	3800	H	Mo
4	990	1900	M	Mo	54	610	3200	M	Mo
5	970	3600	H	Mo	55	1200	1900	M	Mo
6	960	2400	M	Mo	56	900	3200	H	Mo
7	1100	3900	H	Mo	57	1400	3400	H	Mo
8	500	1300	M	Mo	58	1700	3650	H	Mo
9	325	3400	M	Mo	59	600	1000	M	Mo
10	1600	3700	H	Mo	60	1800	1700	M	Mo
11	2290	900	M	Mo	61	2800	2100	M	Mo
12	910	1200	M	Mo	62	2250	2800	M	Mo
13	1200	3200	H	Mo	63	2340	3600	H	Mo
14	625	3700	M	Mo	64	2460	3900	H	Mo
15	890	2500	M	Mo	65	2700	3450	H	Mo
16	1600	2700	M	Mo	66	540	2200	M	Mo
17	1825	3200	H	Mo	67	420	2000	M	Mo
18	1420	3700	H	Mo	68	2500	3400	H	Mo
19	1275	3550	H	Mo	69	2900	3800	H	Mo
20	1390	3690	H	Mo	70	2700	3900	H	Mo
21	2400	1900	M	Mo	71	350	1700	M	Mo
22	255	1840	M	Mo	72	425	1900	M	Mo
23	465	2600	M	Mo	73	580	1850	M	Mo
24	590	3900	M	Mo	74	900	2200	M	Mo
25	1600	1000	L	T	75	625	1000	M	Mo
26	625	1800	M	Mo	76	2290	1300	M	Mo
27	925	2200	M	Mo	77	2400	1100	M	Mo
28	940	2600	M	Mo	78	2600	1000	M	Mo
29	1110	3400	H	Mo	79	2710	950	M	Mo
30	650	1200	M	Mo	80	2540	2800	M	Mo
31	1460	3600	H	Mo	81	2600	3600	H	Mo
32	1720	3500	H	Mo	82	2350	3700	H	Mo
33	2190	1200	M	Mo	83	2700	3900	H	Mo
34	2400	1750	M	Mo	84	2900	3500	H	Mo
35	2200	1910	M	Mo	85	2840	3750	H	Mo
36	2550	2100	M	Mo	86	460	1350	M	Mo
37	2900	3460	H	Mo	87	1500	2400	M	Mo
38	650	1750	M	Mo	88	580	950	M	Mo
39	700	3200	M	Mo	89	2250	1300	M	Mo
40	640	1900	M	Mo	90	2400	2600	M	Mo
41	900	2400	M	Mo	91	2460	2800	L	T
42	1150	2700	M	Mo	92	2520	3700	H	Mo
43	620	1200	M	Mo	93	2690	3800	H	Mo
44	1900	1750	M	Mo	94	650	1700	M	Mo
45	1600	1940	M	Mo	95	460	1900	M	Mo
46	2400	2350	M	Mo	96	510	1100	M	Mo
47	2500	2600	M	Mo	97	1700	2600	M	Mo
48	2700	3900	H	Mo	98	1420	2400	M	Mo
49	2650	3800	H	Mo	99	2700	2800	M	T
50	2950	3600	H	Mo	100	2800	3900	H	Mo

Fertilizer Status: L – Low, M – Medium, H – High Farming Technique: T – Traditional, Mo – Modern

Table 2 : Distribution of Plots in Relation to Rainfall and Yield

Rf\Yield	High	Moderate	Low	Total
High	21	14	6	41
Moderate	12	16	14	42
Low	7	2	8	17
Total	40	32	28	100

plots. For this purpose, as a first step, the plots with abnormal behavior were studied with respect to relief and nature of deviation to be explained was sought ; some more information related to attributes such as fertilizer input and the nature of farming techniques used by the farmers which may be responsible for the abnormalities observed, were discussed.

Spatial Interpretation

A reference to the map shows that, six plots with low rainfall but high yields are located near the stream and therefore possess the benefit of irrigation that compensates for rainfall deficiency, moreover, the soil here is expected to be fertile and, being alluvial, is replenished during the annual floods. The seven plots characterized by high rainfall and low yields, are located on the steep slopes, which do not provide a deep soil cover. Furthermore, there is a possibility of heavy soil erosion and a lot of water is wasted through surface run-off, thereby accounting for the low yield, since these plots are located at or near the foothills; the soil here is coarse and therefore there is excessive percolation which makes it difficult for the shallow root system of rice to find adequate water. Two of the plots in this category, numbering 91 and 99, lie in the plain region but still have only moderate yield because of the traditional farming techniques adopted by the farmers (Table 1).

Fourteen plots characterized by moderate rainfall and high yield are located on the flat lands, on the tops of either a spur, a hill, a ridge or a plateau. The soil here is fine and therefore moisture retentive and thus compensates for the low rainfall to ensure high yields.

Of the two plots under moderate rainfall and low yields, one is located on a steep slope and the low yield therefore is explained. The other plot is located on a terrace and hence the low yield is rather surprising. A reference to table 1, suggests that this is due to low input of fertilizers and traditional farming techniques adopted by the farmer there.

Fourteen plots in the category of low rainfall and moderate yield are located either on terraces or in the plains area, away from the river. However, in all these plots moderate fertilizer use is coupled with moderate yield (Table1).

Conclusion

From the foregoing discussion it is seen that the explanations provided for the deviations, which appeared highly conceptual and aspatial in the beginning, became more realistic in geographical terms with the adoption of appropriate cartographic tools by the researchers to spatially explain the departures. Spatial interpretation of any phenomenon could be cross checked in this way. Cartographic techniques, if used conjunctively with quantitative techniques, would help in making explanations and interpretations more representative of ground reality. It is therefore, suggested that geographers should try to modify those statistical techniques which have not been amended so far to suit their requirements of spatial analysis. This would make their research findings more purposeful and relevant.

Acknowledgement

The authors wish to place on record their heartful thanks to Shri. V.S. Rao, Dr. D. Mukherji and Shri R.G. Gawade for their help in the preparation of the map using GIS.

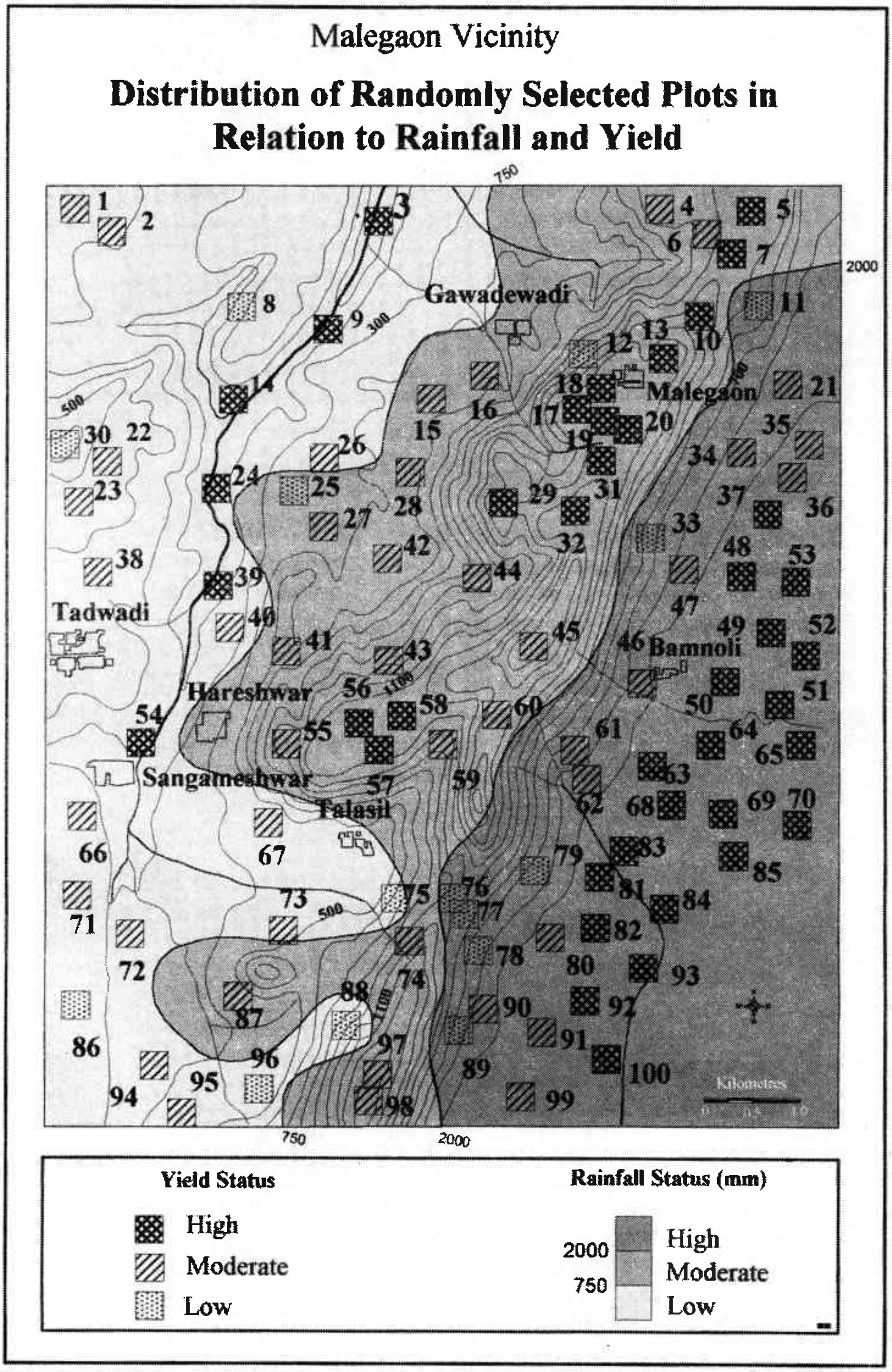

Fig. 1 : Malegaon Vicinity - Distribution of Randomly Selected Plots in Relation to Rainfall and Yield

References

Berry, B.J.L. (1966), Essays on Commodities and the Spatial Structure of the Indian Economy, *Research Paper III*, Department of Geography, University of Chicago.

Dalton, R., Garlick, J., Minshull, R. and Robinson, A. (1972), *Correlation Techniques in Geography*, George Philip and Son Ltd., London.

Mackay, J.R. (1958), Chi-Square as a tool for Regional Studies, Annals AAG, 48, 164-166.

McQuitty, L.L. (1957), Elementary Linkage Analysis for Isolating Orthogonal and Oblique Types and Typal Relevancies, *Educational and Psychological Measurement*, 17, 207-229.

Murdie, R.A. (1969), *Factorial Ecology of Metropolitan Toranto-1951-61*, Research Paper 116, Department of Geography, University of Chicago.

Robinson, A.H. and Bryson, R.A. (1957), A Method for Describing Quantitatively the Correspondence of Geographical distribution, *Annal AAG*, 47, 379-391.

Saviatlovsky, E.E. and Eells, W.C. (1937), The Centrographical Method and Regional Analysis, *Geographical Review*, 27, 240-254.

Taaffe, E.J. (1970), *Geography*, Prentice Hall Inc. Englewood Cliffs, N.J., 44-49.

Taaffe, E.J., Morrill, R.L. and Gould, P.R. (1963), Transport Expansion in Underdeveloped Countries: A Comprehensive Analysis, *Geographical Review*, 53, 503-529.

Taylor, P.J. (1977), *Quantitative Methods in Geography – An Introduction to Spatial Analysis*, Houghton Mifflin Com., Boston.

Yeates, M. (1974), *Introduction to Quantitative Techniques in Human Geography*, McGraw Hill, New York.

Sonali Srivant
and
Dr. V. S. Phadke
both at the Department of Geography,
University of Mumbai, Mumbai.