Journal of Agrometeorology 5 (1) : 25-34 (June 2003)

Analysis of spatial and temporal wheat yield variability in Punjab and Haryana and identification of coherent yield zones

A. S. NAIN, V. K. DADHWAL* and T. P. SINGH[1]
Crop Inventory and Modeling Division, ARG, Space Applications Centre (ISRO), Ahmedabad-380 015.
[1]Amar Singh College, Lakhaoti, Bulandshahr-245 407

ABSTRACT

Coherent yield variability zones for wheat crop in Punjab and Haryana were identified using weather-induced yield variability index (WYVI). This new approach comprises three major steps: (a) generation of consistent district-level wheat yield series considering any reorganization of districts, (b) extraction of WYVI as ratio of observed yields to expected yield from technological trend, and (c) hierarchical agglomerative clustering of inter-district Squared Euclidean Distance (SED) matrix of WYVI. The procedure identified three major zones in Punjab and three major and one single district zones in Haryana, which had similar temporal yield variability. This zonation could act as a basis for grouping of districts for regional yield prediction.

Key words: Wheat yields, Weather induced yield variability index (WYVI), Cluster analysis, Punjab, Haryana, Coherent yield variability zones

Grouping of a geographical area on the basis of some criterian similarity, is an important step in a number of studies ranging from agro-ecology (Patel *et al*., 2000; Subramaniam, 1983; Sehgal *et al*., 1989; Jones and O'Toole, 1987), climatic classification, agricultural planning (Planning Commission, 1998), hydrology, soil suitability and crop distribution, crop growth potential assessment to crop yield forecasting (Dadhwal *et al*., 2001). A number of zonations have been made for India based on a diverse set of inputs and serving different applications. While all zonations have a similarity in a broad sense as they identify unique physical features such as deserts, islands and mountain ranges, they also have shown significant difference in how the agricultural plains of India are classified.

Studies on regional crop yield forecasting have either been done at state level or meteorological subdivision level, for which detailed models for district and smaller areas have been developed. The state level and meteorological subdivision level approach exhibit large within region variability with respect to crop distribution and its response to growing conditions. Thus regional yield forecasting requires identification of coherent regions, which respond similarly to regional weather

* Corresponding author

induced variability. NARP (127) and ARPU (72) zones which are based on similarity of physiography, climate and soil have been widely adopted for resource assessment and agricultural planning, but have not been evaluated against a spatial weather induced yield variability criterion.

The problem of identification of zones for yield forecasting is addressed through a new procedure, which is based on the similarity in the pattern of temporal variability in crop yield amongst the neighboring districts. The procedure was tested on historical wheat yield series for 12 districts of the Punjab and 12 districts of Haryana.

MATERIAL AND METHODS

The zonation of the districts on the basis of weather-induced variability in wheat yield has been done using following steps.

Generation of consistent wheat yield series

District-wise historical data of wheat area, production and yield for the period of 1974-1997 for Punjab and 1979-1998 for Haryana were obtained from State Agriculture Department and published estimates in various issues of "Agricultural Situations in India". The number of districts in both Punjab and Haryana was 12 in 1974, by 1998 this number increased to 17 and 19, respectively. A consistent historical yield series for the initial 24 districts of Punjab and Haryana was generated for further analysis. This involved acreage and production of new districts to be aggregated with statistics of their parent district or major parent district, when they were created from two or more districts. The resulting series is considered. While most of the new districts have been bifurcated from a single district, Nawasahar, Kaithal, Rewari and Yamunanagar in Haryana and Moga in Punjab were created by merging parts from adjoining old districts.

Computation of weather-induced yield variability index (WYVI)

The variability in wheat yield (Y_{ij}, i^{th} year, j^{th} district) originates from changes in technology (including varieties, fertilizer, irrigation and crop management) and current year's weather. Weather-induced wheat yield variability index (WYVI) was computed by normalizing observed yield with respect to expected yield from current technology using following steps:

(a) *Estimation of technology trend:* The change in technology was modelled as a single linear or segmented linear function of time. Segmented trends were necessary because in a large number of districts "green revolution fatigue" could be observed. The values for coefficients "a" and "b" are given in Table 1. The equation for technology (Yt_{ij}) trend is

For single trend

$$Yt_{ij} = a + b*T_i \qquad ......[1]$$

Where T_i = Years since 1974 for Punjab and 1979 for Haryana

For segmented trend

$$Yt_{ij} = a + b_1T_{1i} + b_2T_{2i} \qquad ......[2]$$

Where T_{1i} = Year since 1974 for Punjab and 1979 for Haryana or 0 if $T_i > Y_s$

T_{2i} = 0 if $T_i \leq Y_s$ Year since 1974 for Punjab and 1979 for Haryana

Y_s = Year of split for wheat yield series, a, b (b_1, b_2) are coefficients obtained by regressing observed yield (Y_{ij}) with time.

(b) *Generation of weather-induced yield variability Index (WYVI):* The relative deviations obtained in crop yield from modelled technology trend based yield (Yt_{ij}) for 12 districts each, for the period of 1974-1997 and 1979-1998 for Punjab and Haryana, respectively, was taken to represent WYVI for each district and year.

$$WYVIij = ((Y_{ij} - Yt_{ij})/Yt_{ij}) \quad ......[3]$$

Inter-district analysis of WYVI

(c) *Computation of inter-district Squared Euclidean Distance (SED) matrix:* Pair-wise squared Euclidean distance between year wise WYVI for each pair of districts (X, Y) was computed to measure the similarity in temporal characteristics of district's yield series.

$$D_{XY} = (WYVI_{ix} - WYVI_{iy})^2 \quad ......[4]$$

where,

$WYVI_{ix}$ = WYVI of x district for i^{th} year

$WYVI_{iy}$ = WYVI of y district for i^{th} year

n = Number of years

(d) *Agglomerative clustering of SED:* Cluster analysis of SED was carried out using hierarchical agglomerative cluster classification to group the districts into coherent zones of weather-induced yield variability. In this procedure, the two closest districts (i and j) were merged and the distances with remaining districts were updated using following relation.

$$d(i+j,k) = a_i d(i,k) + a_j d(j,k) + bd(i,j) + c \mid d(i,k) - d(j,k) \mid \quad ......[5]$$

where, i, j and k are clusters and d(i, j) is the distance between clusters i and j

Amongst the various methods of agglomerative clustering, Ward's method, which combines clusters with the smallest increase in the overall sum of the squared distance within cluster distance (Jain 1986) was adopted to identify the coherent yield variability clusters. The analysis was carried out using Statistical Package for Social Sciences (SPSS) version 6.1 (Marija 1994).

Ward's method for computation of a, b, and c is as follows:

$$a_i = \frac{n_k + n_i}{n_k + n_i + n_j}; a_j = \frac{n_k + n_j}{n_k + n_i + n_j}$$

$$b = \frac{n_k}{n_k + n_i + n_j}; c = 0 \quad ......[6]$$

where, n_i, n_j and n_k are the number of pattern in cluster i, j and k, respectively

(e) *Identification of coherent yield variability zones:* A dendrogram of inter-cluster distance indicating normalized distance at which various districts are merged, was generated, to identify major

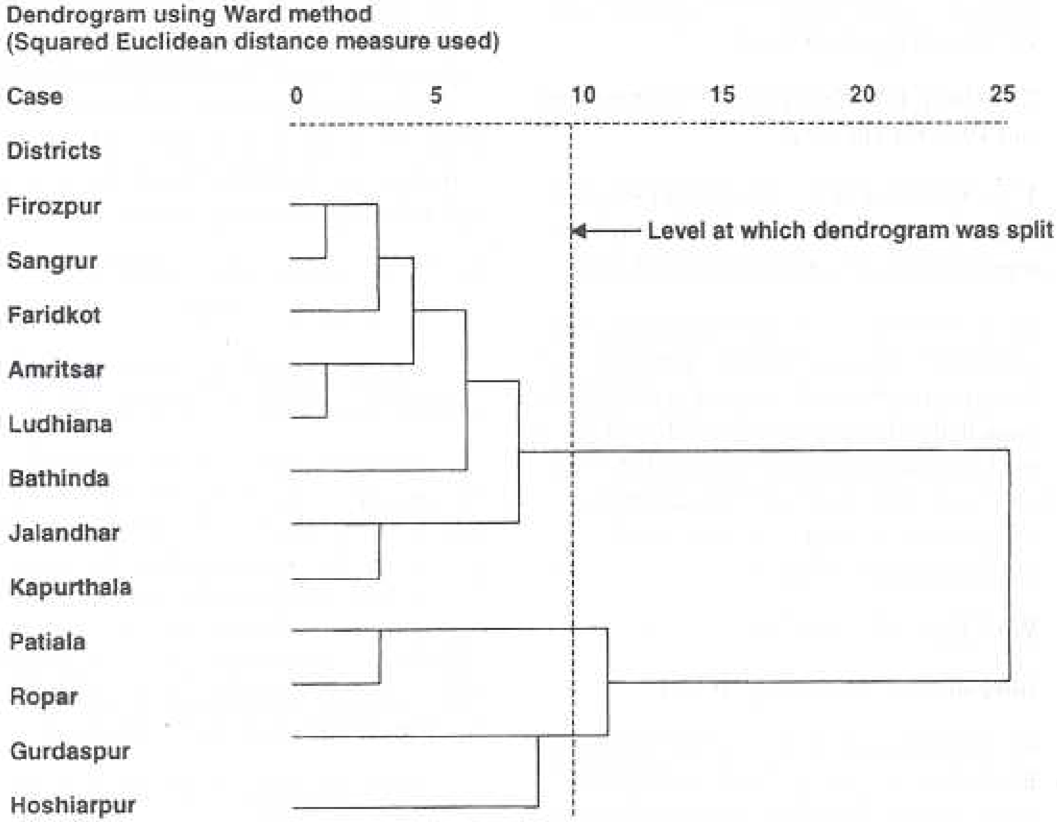

Fig. 1 : Dendrogram of 12 districts of Punjab with respect to weather-induced yield variability.

clusters which exist in Punjab and Haryana. The cluster labels were transferred to a district map of the respective state, and spatially contiguous districts having same cluster label were identified as forming a coherent yield variability zone. If a cluster did not show spatial contiguity, cluster was split to create contiguous zones.

RESULTS AND DISCUSSION

Analysis of yield series

District-wise average wheat yields for Punjab (1974-1997) and for Haryana (1979-1998) for 24 districts are summarized in Table 1. District wise wheat yield varied from 2.48 t ha^{-1} (Hoshiarpur) to 3.81 t ha^{-1} (Ludhiana) in Punjab, and 2.43 t ha^{-1} (Gurgaon) to 3.18 t ha^{-1} (Sirsa) in Haryana.

Variation in wheat yields in Punjab was the highest in district of Kapurthala with coefficient of variance 21.6% (SD 0.683 t ha^{-1}) while it was the lowest in Ludhiana. The yield range in Haryana was the highest in Mahendragarh the lowest in Faridabad. In 18 districts single technology trend could capture year-to-year variability in wheat yield. However, in 6 districts segmented regression explained a higher proportion of yield variability than the single trend. These districts are Bathinda, Ambala, Hisar, Karnal, Kurukshetra, and Mahendragarh. The slope in later period is lower in all these cases indicating slowing of year-wise yield increase due to technology advancement in current phase in comparison to earlier phase.

Identification of coherent zone

The district level inter-cluster distance depicted through agglomerative clustering using Ward's method for Punjab and Haryana are summarized in Figs. 1 and 3, respectively. The number of zones is determined by similarity level at which grouping is carried out. The approach could identify three major zones in Punjab and three major and one single district membership zones in Haryana as described below.

Punjab: Identification of clusters led to define 5 pairs of districts and 2 single districts, which successively merged into three zones having membership of 8, 2 and 2 districts (Fig. 1). Cluster I is the largest with 8 districts located in central and western plains zone of Punjab. While cluster II and III have two districts each falling in sub-mountainous undulating zone and central plain zone (Table 1). The average yield series for each identified zones were generated (Fig. 2) and percent change over previous year was computed. Zone III has lower yields than average state yields, while zone I has marginally higher yields. Zone II has lower yields in earlier years and higher than average state yield in following years. A representing set of 6 identified crop seasons, when the identified zones differed amongst themselves as well as with state level yield are presented in Fig 3.

Haryana: When cluster technique was applied on 12 Haryana districts, 4 pairs of districts and 4 single districts were identified. These 4 pairs and 4 single districts successively merged into three major clusters (Fig. 4). Cluster I with two districts was a part of central plain zone of Haryana. Cluster II includes four districts and Cluster III being the largest had membership of six districts grouped in four pockets in northern plains zone, southern plains zone, western plains zone and northern hilly zone of Haryana. As later cluster was not contiguous, it was split into two-part zone IIIa (covering Ambala) and zone IIIb (remaining 5 districts). Thus, in all 4 spatially contiguous zones could be identified from 3 original clusters. However, one of them (Zone IIIa) has single district (Ambala). Yield series of three zones and state of Haryana are shown in Fig. 5. and descriptions of six identified seasons with significant yield variation in identified zones are depicted in Fig 6. In general wheat yield series of zone II and IIIa have lower yields, while zone I and IIb have higher yields than state average yields of Haryana, except in a few cases when in earlier years yields of zone IIb has lower yields. Presence of geographically separated groups of districts in one cluster could also arise due to

Table 1 : Zonation of districts of Punjab and Haryana based on weather-induced variability in wheat yields.

Sr. No	District name	Avg. yield (t ha⁻¹)	Coefficient "a"*	Coefficient "b"**	SD (t ha⁻¹) & (CV %)	WYVI Zone
	Punjab					
1	Firozpur	3.376	0.079	2.387	0.565(16.7)	I
2	Sangrur	3.608	0.097	2.400	0.694(19.2)	I
3	Faridkot	3.220	0.067	2.237	0.499(15.5)	I
4	Amritsar	3.281	0.085	2.224	0.630(19.2)	I
5	Ludhiana	3.809	0.070	2.926	0.531(13.9)	I
6	Bathinda$	3.011	0.119	1.943	0.575(19.1)	I
			0.095	1.717		
7	Jalandhar	3.341	0.083	2.317	0.619(18.5)	I
8	Kapurthala	3.156	0.092	2.010	0.683(21.6)	I
9	Patiala	3.445	0.090	2.315	0.660(19.2)	II
10	Ropar	2.840	0.073	1.926	0.544(19.1)	II
11	Gurdaspur	2.914	0.083	1.876	0.621(21.3	III
12	Hoshiarpur	2.482	0.064	1.681	0.472(19.0)	III
	Haryana					
1	Karnal$	3.123	0.103	1.877	0.626(20.1)	I
			0.003	3.721		
2	Kurukshetra$	3.086	0.084	2.006	0.577(18.7)	I
			0.052	2.678		
3	Faridabad	3.022	0.087	1.675	0.542(17.9)	II
4	Rohtak	2.727	0.091	1.541	0.653(23.9)	II
5	Sonipat	2.862	0.089	1.703	0.655(22.9)	II
6	Gurgaon	2.433	0.077	1.433	0.588(24.2)	II
7	Ambala$	2.475	0.096	1.300	0.643(26.0)	IIIa
			0.126	0.383		
8	Jind	2.927	0.098	1.655	0.733(24.7)	IIIb
9	Hisar$	3.136	0.118	1.687	0.723(23.1)	IIIb
			0.005	3.883		
10	Mahendragarh$	3.024	0.089	1.774	0.749(24.8)	IIIb
			0.037	3.122		
11	Sirsa	3.176	0.100	1.821	0.725(22.8)	IIIb
12	Bhiwani	2.670	0.091	1.487	0.659(24.7)	IIIb

$ = Two values for coefficient "a" and "b" have been given for those districts having segmented technology trend.
* = Slope of the equation. ** = Intercept value of equation.

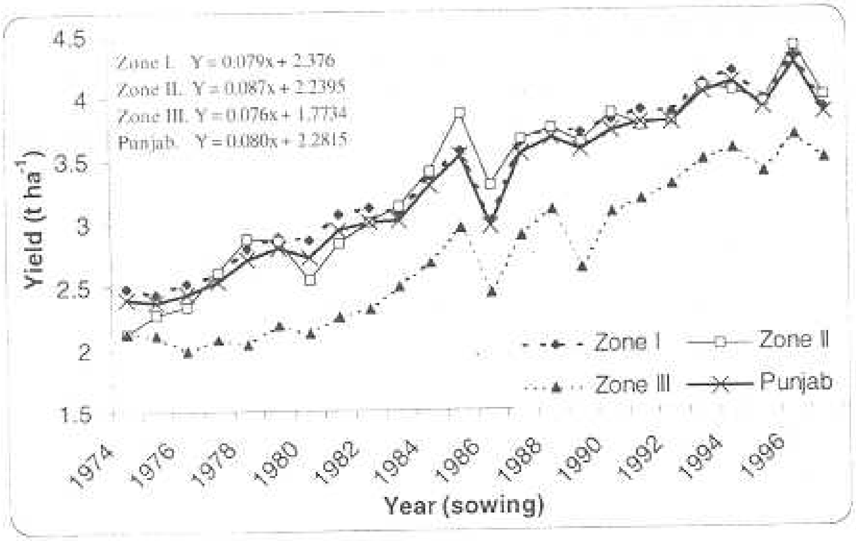

Fig. 2 : Wheat yields series of Punjab and 3 zones identified in this study.

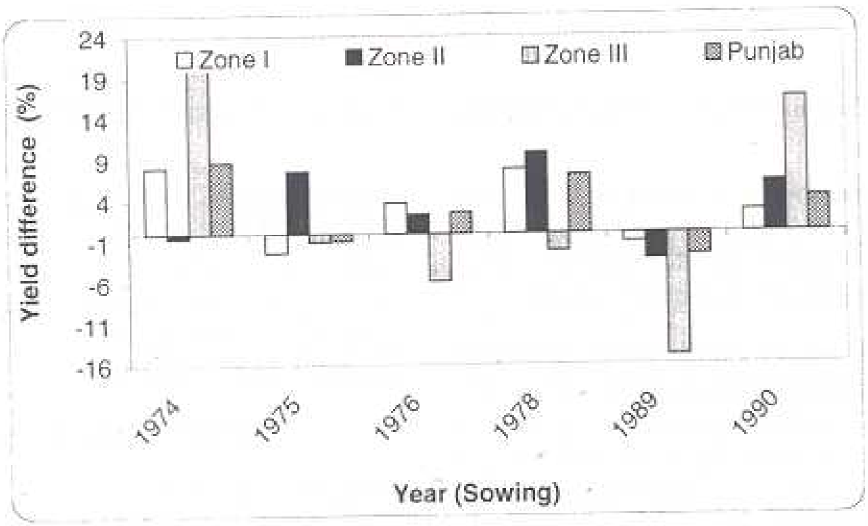

Fig. 3 : Year-to-year differences in wheat yield for selected seasons demonstrating differential temporal behaviour of the identified zones of Punjab.

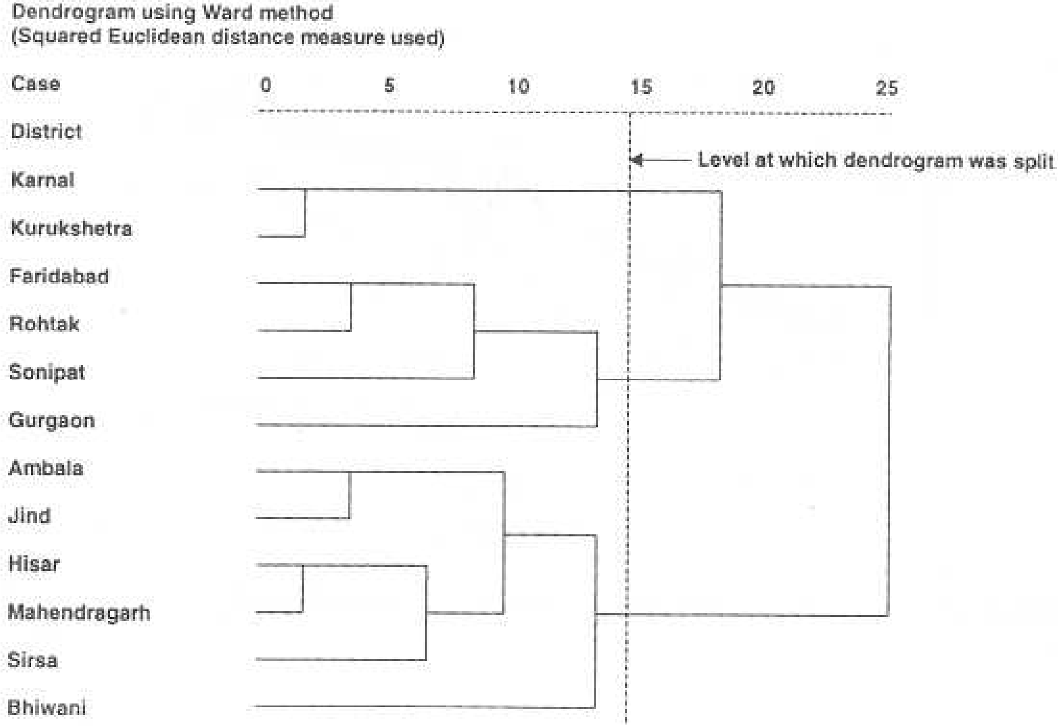

Fig. 4 : Dendrogram of 12 districts of Haryana with respect to weather-induced yield variability.

teleconnections in weather behavior, especially of parameter influencing wheat yield. But in crop assessment context, these zones need to be analyzed separately.

The identified zones could form a basis for developing regional yield model for use in wheat yield forecasting in Punjab and Haryana. Similar approach needs to be tested on other crops as this zonation may vary in different crops.

ACKNOWLEDGEMENT

Authors acknowledge encouragement received from Shri J. S. Parihar, Group Director, ARG, SAC (ISRO), Ahmedabad. ASN acknowledges Dr. S. P. Vyas for providing the wheat statistics, and grant of Junior Research Fellowship by Space Applications Centre (ISRO).

REFERENCES

Dadhwal, V. K., Singh T. P. and Nain, A. S. 2001. Identification of coherent yield variability zones for wheat in Uttar Pradesh, India. Submitted to Agricultural Situations in India. In press.

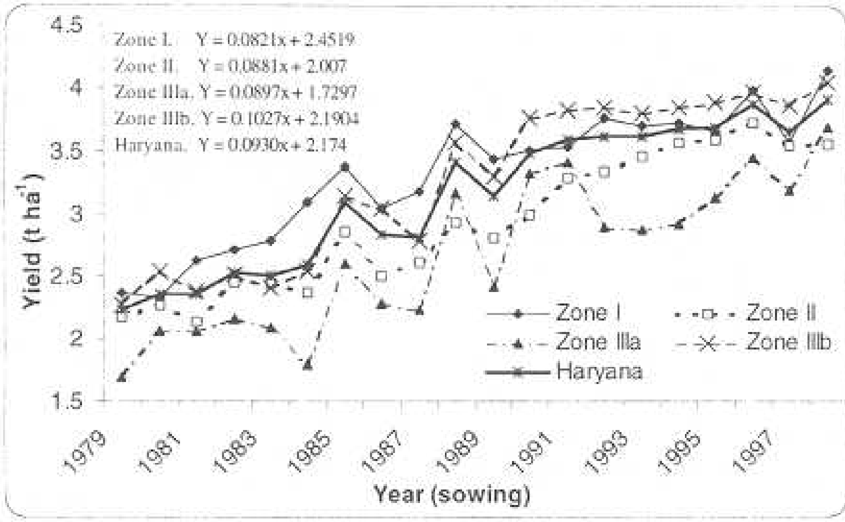

Fig. 5 : Wheat yield series of Haryana and 4 zones identified in this study.

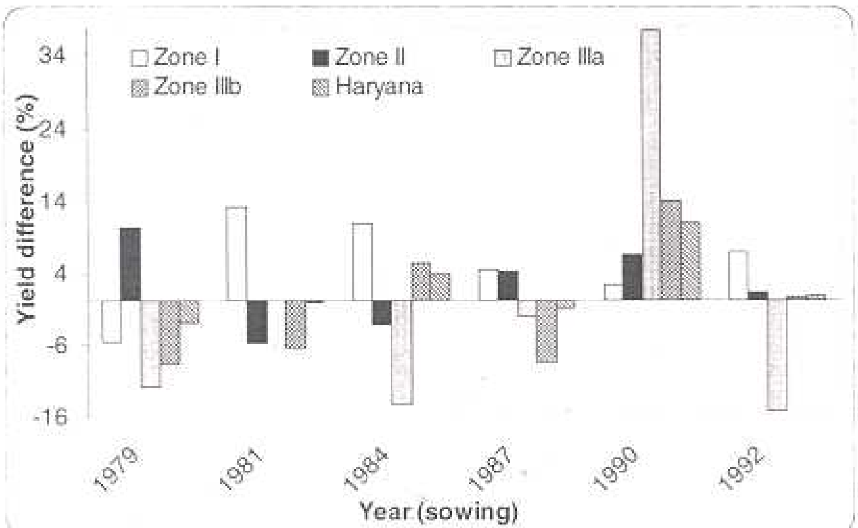

Fig. 6 : Year-to-year differences in wheat yield for selected seasons demonstrating differential temporal behaviour of the identified zones of Haryana.

Ghosh, S. P. (ed.) 1991. Agro-climatic zone Specific Research: Indian perspective under NARP. ICAR publication, New Delhi.

Jain, A. K. 1986. Cluster analysis. Handbook of pattern recognition and image processing. Academic Press Inc. 1986. 33-57.

Jones, C. A. and O:Toole, J. C. 1987. Application of crop production models in agro-ecological characterization: Simulation models for specific crops. *Agricultural environments-characterization, classification and mapping*. Bunting, A. H. (ed.). CAB International UK. 199-209.

Marija J. Norusis 1994. *Manual of SPSS Professional Statistics 6.1*. SPSS Inc. Michigan Chicago USA 1994. pp-385.

Patel, N. R., Mandal, U. K. and Pande, L. M. 2000. Agro-ecological zoning system - A remote sensing and GIS perspective. *J. Agrometeorol*, 2(1): 1-13.

Planning Commission 1998. Agro-climate Regional Planning - Recent developments. *Working Paper No. 10*, Ahmedabad, ARPU.

Sehgal, J.L., Mandal, D.K., Mandal, C. and Vadivelu, S. 1989. Agro-ecological Region of India, NBSS&LUP (ICAR) Publication 24, Nagpur 76pp.

Subramaniam, A.R. 1983. Agro-ecological Zones of India. *Arch. Met. Geoph. Biocl. Ser.* B32: 329-333.