MB 8.27
November 1985

THE VARIABILITY IN THE YIELD OF PEARL MILLET VARIETIES AND HYBRIDS IN INDIA AND PAKISTAN

J.R. Witcombe

Paper presented at the IFPRI/DSE Workshop on Sources of Increased Variability in Cereal Yields

Feldafing, Germany, November 26-29, 1985.

Pearl Millet Improvement Program
International Crops Research Institute for the Semi-Arid Tropics
ICRISAT Patancheru P.O.
Andhra Pradesh 502 324, India

THE VARIABILITY IN THE YIELD OF PEARL MILLET VARIETIES AND HYBRIDS IN INDIA AND PAKISTAN.

J.R. Witcombe[1]

Abstract

The International Pearl Millet Adaptation Trial (IPMAT) which has both hybrids and varieties as entries, has been grown multilocationaly in India and Pakistan. The grain yield data over five years have been analysed in a number of ways to examine the stability of the entries.

A regression analysis indicated that the breeder's procedure of selecting amongst the highest-yielding entries across environments is correct, as it also selects entries that perform well in poor environments. Selecting entries on predicted performance in the lowest-yielding environment appears to be a less reliable procedure.

The hybrids are generally higher-yielding than the varieties but are less stable. The most important source of genotype x environment interaction in the regression analysis was due to the deviation from the regressions (S^2_d values), rather than variation between the regressions. The varieties were superior to the hybrids in this respect, having lower than average S^2_d values. However, a mean-variance analysis showed that the highest-yielding genotype was always preferred among the risk efficient entries. Similarly, a first degree stochastic efficiency analysis indicated that the hybrids, despite their inferior stability in a regression analysis, were more risk efficient than the varieties.

---

1 Principal Pearl Millet Breeder, ICRISAT, Patancheru P.O., Andhra Pradesh 502 324, India.

One variety, ICMV 81111, from an advanced cycle composite combined both high yield and stability. Its advantage over hybrids in contributing to crop yield stability is discussed.

Introduction

The possibility that open-pollinated varieties may be more stable in yield than hybrids has been much discussed. Over a number of years and across many locations in India and Pakistan open-pollinated varieties and hybrids have been tested in the ICRISAT International Pearl Millet Adaptation Trial, IPMAT (Table 1). Consequently, it is possible to examine whether the customary procedure of selecting amongst the best entries on mean yield data is satisfactory. It is also possible to analyse the relative stability of hybrids and varieties and to examine the consistency of any differences over years. The hypothesis can then be tested that the narrowing of the genetic base by the adoption of uniform hybrids over large areas leads to increases in variability of yield.

Several statistical techniques have been developed to analyse the interaction of genotypes and environments. Apart from techniques which use analyses of variance, regression analyses have been extensively used. The first proposal of a regression analysis for the purpose of studying interactions was by Yates and Cochrane (1938). Modifications of this method have been used by Finley and Wilkinson (1963), Eberhart and Russell (1966) and Perkins and Jinks (1968). All these methods are essentialy the same, as the mean squares for the major components of variation are identical in all cases. Although regression analyses have been extensively used they have various limitations, some of which have been pointed out by Witcombe and Whittington (1971), by

Knight (1970), and by Binswanger and Barah (1980).

In all these analyses the average yield of all the genotyes in an environment is termed the environmental index and is used as an assessment of that environment. The individual yield data of the entries are then regressed on the environmental indicies, giving three main parameters that describe the performance of each entry:

1. The mean yield of the entry.
2. The regression coefficient of the entry. The average coefficient for all entries in the analysis is 1; entries with slopes above 1 are considered to be less stable, and those with slopes less than 1 more stable.
3. The remainder mean square (R.M.S.) of the individual regression analysis of an entry. This accounts for the deviations from the regression line and indicates how well an individual entry meets the linear model. The size of the R.M.S. indicates the degree of unpredictability of yield of an entry and lower values are desirable. Eberhart and Russell (1966) substract a constant, which is a measure of inexplicable environmental variation, from the R.M.S. The value thus obtained is termed $S^2{}_d$ but obviously bears a simple relationship to the R.M.S. The derivation of the R.M.S. is discussed and partially accounted for in Witcombe and Whittington (1971).

The regression analysis of variance also partitions the variance into its components, and the relative contributions of E, G and GxE can readily be seen.

Regression analyses have limitations when economic criteria are considered. Two further analyses have therefore been done using analyses more often used by economists rather than plant breeders.

Results and Discussion

Comparison of selection for yield in different environments

To examine the patterns in high, average and low-yielding environments the values for yield, regression coefficient and S^2_d of the entries have been plotted for the mean of the environments, and for the predicted yields in the highest and lowest-yielding environment in the trial (Figs 1-5). In this way the predictive value of the linear model has not been extrapolated beyond of the range of the experiments, and in all years, except 1983, the lowest yield of about 600 kg ha^{-1} which is typical of the average yield of farmer's fields in many regions of India.

The selection of the best entries which can be promoted to the next stage in the testing scheme is a most important consideration. How do the top-yielding entries vary between the environments? Clearly, there are large differences in the ranking of entries between the highest-yielding and lowest-yielding environments and usually the top two entries are not common in these two extreme environments. Of more interest is the comparison between the average-yielding environment (on which recommendations are usually based), and on the predicted performance in the lowest-yielding environment of the experiment (which is more typical of the farmers' fields). In three of the five years there is a common entry in the two highest-yielding entries in these two environments (hybrid 3 in 1979; variety 10 in 1983; and hybrid 23 in 1984) (Table 2).

Within these top two entries can selections be made for varieties in preference to hybrids, and can high-yielding entries be selected with low S^2_d values ? If selection is made in the lowest-yielding environment then in three out of five years a variety can be chosen, and in three out of the five years an entry with a low S^2_d value. In the mean-yielding environment a variety can only be selected in one of the five years, and there is little difference in the stability of the entries (Table 2).

We can also examine the data by selecting the only single best entry rather than the top two. We can then see how the highest-yielding entry over all environments performs in the lowest-yielding environment, or vice versa (Table 3). A marked contrast between these environments can be seen, and in only one year is the highest-yielding entry the same. However, it would appear that selection based on the mean of environments is more reliable for selecting entries that perform well in the poorest environment than vice versa. In the former case the worst yield rank is 5, but in the latter it is 13.

ri hybrids and varieties

In general, the hybrids are higher-yielding, differ little from the varieties in their regression coefficients, but are, on average, less stable for the S^2_d values (Figs 1-5). These conclusions are confirmed by the mean values of these parameters for hybrids and varieties (Table 4).

Since the varieties have lower S^2_d values than the hybrids we need to examine the importance of the deviations from the regressions as a source of variability. Therefore, the amount of genotypic variability in the experiments has been related to the amount of variability between environments (Table 5). The environment is by far the largest component of variation, the interaction of the genotypes with the environment the next largest, and the difference between genotypes the smallest. When we consider the genotype x environment interaction, the variation due to differences in regression coefficients is invariably much smaller than that due to the R.M.S.'s (deviations from the regression). Since the varieties are superior to hybrids in respect of the deviations fsom the regressions (S^2_d values) then it is possible that the selection of varieties may reduce crop yield variability. When stability for both S^2_d and higher than average yield are the selection criteria varieties predominate, even though hybrids would most frequently be selected solely on the basis of yield. This predominance of varieties is demonstrated by the selection of entries on the basis of a selection index where entries have to be higher-yielding than average, both overall and in the lowest environment, and have a lower S^2_d value than average. In all cases the entry selected was a variety, although in 1984 it is impossible to make a selection using these criteria (Table 6).

What is not clear is how the superior stability of varieties translates into real terms for the farmer. How, in a specific location does a low S^2_d reduce the variation from year to year (the temporal variation) as compared to a low regression coefficient? How large a reduction in mean yield can be accepted to select an entry with a lower S^2_d? The answer to these questions are important but unatainable from a regression analysis. Binswanger and Barah (1980) have pointed out that regression analyses are not suitable for studies on risk aversion, instead they have used a mean-variance analysis where reduction in temporal variation (or improvement in adaptability) can be traded off against reduced yield. However, their analysis, unlike regression analysis, cannot predict the behaviour of the genotypes in different environments.

The variance analysis of Binswanger and Barah (1980) was carried out for all five years. The analysis essentially involves plotting the standard deviation against mean yield using the same scale for both axes (Figure 6), and it assumes that entries of equal utility lie on an iso-utility curve with a slope of 2.0 for representative risk averse farmers. Such an iso-utility curve is indicated in figure 7. In each year a line is drawn connecting the risk efficient entries i.e. those that for a given level of yield performance have the lowest standard deviation. That line maps a risk efficient frontier and it is also indicated in figure six. The best entry must lie on this frontier, and it is chosen with regard to the iso-utility curves. In no year was an entry chosen other than the one that was highest-yielding over all environments. Thus, in no case did the improved stability of the varieties over the hybrids compensate for their lower yield.

Stochastic dominance techniques (Anderson, Dillon, and Hardaker, 1977) can also be used to help predict on a risk averse farmer's choice of varietal type. A stochastic dominance analysis was carried out on pooled data for hybrids and varieties as separate groups on predicted yield in the lowest-yielding environment in 1979, 1980, 1981, and 1983. Data from 1984 were omitted because the hybrids in that year are qualitatively different from those in the earlier years. A focus on the lowest-yielding environment is justified because it is more likely to represent farmers' field conditions than the other locations. Over most of the range in predicted yields in the lowest yielding environments, hybrids were superior to varieties (Figure 7). For example, the cumulative probability that a variety would yield less than 400 kg ha^{-1} was about 0.2 whilst the same estimate for a hybrid was about 50% lower at 0.1. Because the two distributions crossed at the lowest yield interval, stochastic dominance analysis did not give clear predictions on whether risk averse farmers would prefer hybrids to varieties, but the estimated cumulative yield distributions do not indicate that the genetically more broadly based varieties as a group are more stable yielders than hybrids in the most adverse production environments in the trials.

Conclusions

Although there were big differences between the lowest and mean-yielding environments it is possible, when only the two highest-yielding entries are considered, to select in most years an entry which is common in both environments. Indeed, in three of the years the highest-yielding entry across all environments was also one of the two highest-yielding entries in the lowest-yielding environment. Moreover, good performance over all environments appears to indicate good performance of an entry in the low-yielding environment, whereas the reverse does not always apply. This is generally an encouraging result for the plant breeder in terms of selection procedure, as entries which perform well in both average and low-yielding environments possess a certain measure of stability.

Although regression analyses are helpful in testing selection procedure it is an inescapable conclusion that to obtain an overall picture of how stability and mean yield are to be traded off other analyses are required. The mean variance analysis convincingly showed that in all years no entry was selected other than on the basis of its mean yield. A stochastic dominance analysis showed that in general varieties would not be preferred to hybrids, even though varieties show superior stability in other analyses. It thus appears likely that these analyses demonstrate that the breeder's procedure of selecting among the highest-yielding entries across environments is satisfactory. Such an emphasis will usually select entries which perform well in poor environments and which would be chosen by risk averse farmers.

It is of interest that the best entry in 1983 is V 10 (Figures 4b, 4c and 6); this is an ICRISAT millet variety, ICMV 81111, derived from an advanced cycle composite, and it has both high yield and superior stability. It is currently in the All India Co-ordinated Millets Improvement Project (AICMIP) trials and, in the 1984 AICMIP Initial Population Trial, was the top-performing entry with a yield about 15% better than the released variety WC-C75 and with a very high level of downy mildew resistance. A variety which is both high-yielding and stable is a desirable alternative to a hybrid, particularly in view of the simpler seed multiplication procedures and the reduced susceptibility of varieties to ergot and smut. Moreover, individual hybrids in India have proven to be most unstable in yield from year to year due to their rapid increase in susceptibility to downy mildew. There is every reason to expect that the more genetically diverse variety would become susceptible in a less rapid and spectacular manner.

Acknowledgements

This paper would not have been possible without the help of our co-operators who have grown the trials, and their vital role is gratefully acknowledged. The IPMAT reports of 1979 and 1980 were prepared by Dr. Anand Kumar. Drs. Anand Kumar and K.N. Rai have co-ordinated the trials of IPMAT in different years.

I am grateful to Mr. M.M.V.R. Rao for his assistance in analysing the data using GENSTAT, and to Dr. Tom Walker for his valuable suggestions in the analysis of the data from an economist's viewpoint.

References

Anderson, J.R., Dillon, J.L. and Hardaker, J.B. 1977. Agricultural Decision Analysis. Iowa State University Press.

Binswanger, Hans P. and Barah, B.C. 1980. Yield risk, risk aversion, and genotype selection: conceptual issues and approaches. ICRISAT Research Bulletin 3, pp 23.

Eberhart, S.A. and Russell, W.A. 1966. Stability parameters for comparing varieties. Crop Science, 6, 36-40.

Finlay, K.W. and Wilkinson, G.N. 1963. The analysis of adaptation in a plant breeding programme. Australian Journal of Agricultural Science, 14, 742-754.

Knight, R. 1970. The measurement and interpretation of genotype environment interactions. Euphytica, 19, 225-35.

Perkins, Jean M. and Jinks, J.L. 1968. Environment and genotype environmental components of variability. III. Multiple lines and crosses. Heredity, 23, 339-356.

Witcombe, J.R., and Whittington, W.J. 1971. A study of the genotype by environment interaction shown by germinating seeds of Brassica L. Heredity, 26, 397-411.

Yates, F., and Cochran, W.G. 1938. The analysis of groups of experiments. Journal of Agricultural Science, Camb., 28, 556-580.

Table 1. Summary of IPMAT trials analysed.

Year[1]	Number of entries analysed	Hybrids	Varieties	Number of locations analysed
1979	19	7	12	14
1980	19	8	11	19
1981	18	7	11	15
1983	21	8	13	17
1984	23	9	14	18

[1]Trial not held in 1982.

Table 2. Characteristics of the two highest-yielding entries in the mean and lowest-yielding environments.

	No. of entries	Mean environments						Lowest environment					
		First			Second			First			Second		
		Entry	Rank slope	Rank S^2_d	Entry	Rank slope	Rank S^2_d	Entry	Rank slope	Rank S^2_d	Entry	Rank slope	Rank S^2_d
1979	19	H 3	10	16	H 5	18	15	H 3	10	16	V 9	1	14
1980	19	H 1	16	3	H 8	12	13	H 2	4	6	H 4	3	12
1981	18	H 6	17	14	H 3	18	13	V 15	1	16	V 12	5	5
1983	21	V 10	11	2	H 17	14	9	H 2	3	18	V 10	11	2
1984	23	H 23	20	23	H 20	21	18	H 1	2	13	H 23	20	23

H = hybrid.
V = variety.

Table 3. Yield and stability characteristics of highest-yielding entries in the overall environments and the lowest-yielding environment.

	No. of entries	Mean of environments			Lowest-yielding environment	
		Yield rank	Regression rank	S^2_d rank	Predicted yield rank	Type
1979	19	1	10	16	1	H
1980	19	1	12	13	6	H
1981	18	1	17	14	5	H
1983	21	1	11	2	4	V
1984	23	1	20	23	2	H

	No. of entries	Lowest-yielding environment			Mean of environments	
		Predicted yield rank	Regression rank	S^2_d rank	Actual yield rank	Type
1979	19	1	10	16	1	H
1980	19	1	3	12	9	H
1981	18	1	1	16	13	V
1983	21	1	3	18	4	H
1984	23	1	2	13	6	H

Table 4. Mean yields, mean S^2_d values, and mean regression coefficients for the hybrids and varieties across years.

Year	Mean grain yield (kg ha^{-1})		S^2_d		Mean slope	
	Hybrids	Varieties	Hybrids	Varieties	Hybrids	Varieties
1979	2300	2237	163,533	61,222	1.02	0.98
1980	2096	1974	82,791	33,096	1.01	0.99
1981	2238	2262	117,970	33,189	1.04	0.97
1983	2068	2109	157,951	50,988	0.97	1.02
1984	2028	1773	82,228	40,188	1.07	0.98

Table 5. Sources of genetic variation relative to environmental variation standardised to 100.

	1979	1980	1981	1983	1984
Environment	100	100	100	100	100
Genotype (= Mean values)	2.2	1.5	4.1	2.8	6.6
G x E (linear = b values)	0.6	1.0	2.5	1.1	1.9
G x E (deviation = σ^2_d values)	14.9	12.8	16.3	19.7	19.0

Table 6. Entries selected on basis of complex criteria of yield and stability.

Entry	Mean E Rank for yield	Lowest E Rank for yield	Rank for σ^2_d	Rank for slope
1979 V 12	3	3	2	7
1980 V 10	5	10	1	13
1981 H 8	4	6	7	13
1983 V 10	1	4	2	12
1984 -	-	-	-	-

Legend for figures 1 to 5.

Relationship of regression coefficients, yield and S^2_d values in

(a) highest-yielding environment.
(b) average of environments.
(c) lowest-yielding environment.

Squares	- varieties
Triangles	- hybrids
Dotted symbols	- best ten ranked entries for S^2_d
Numbers	- refer to entry number in analysis of trial

Fig 1a

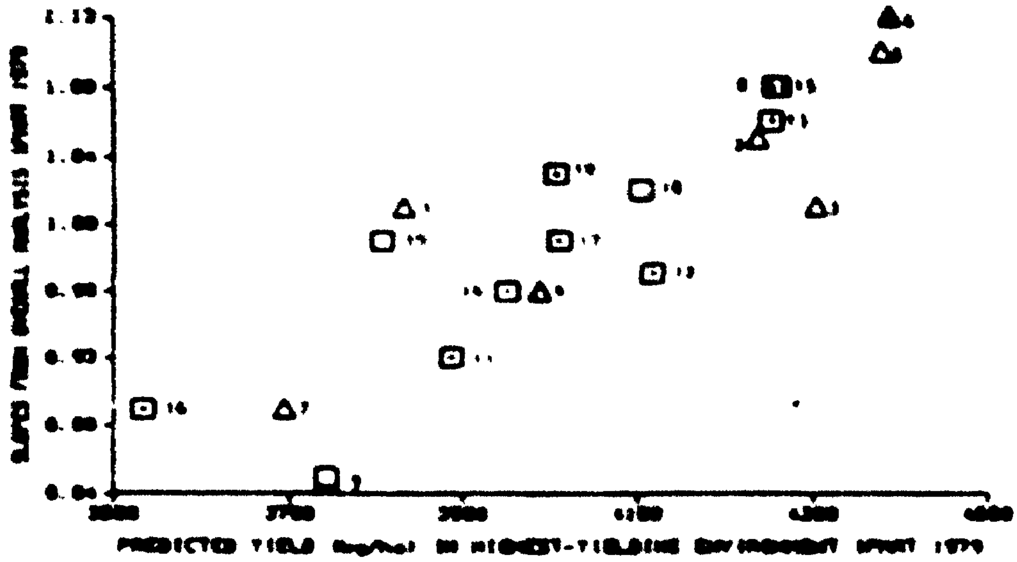


Fig 1b

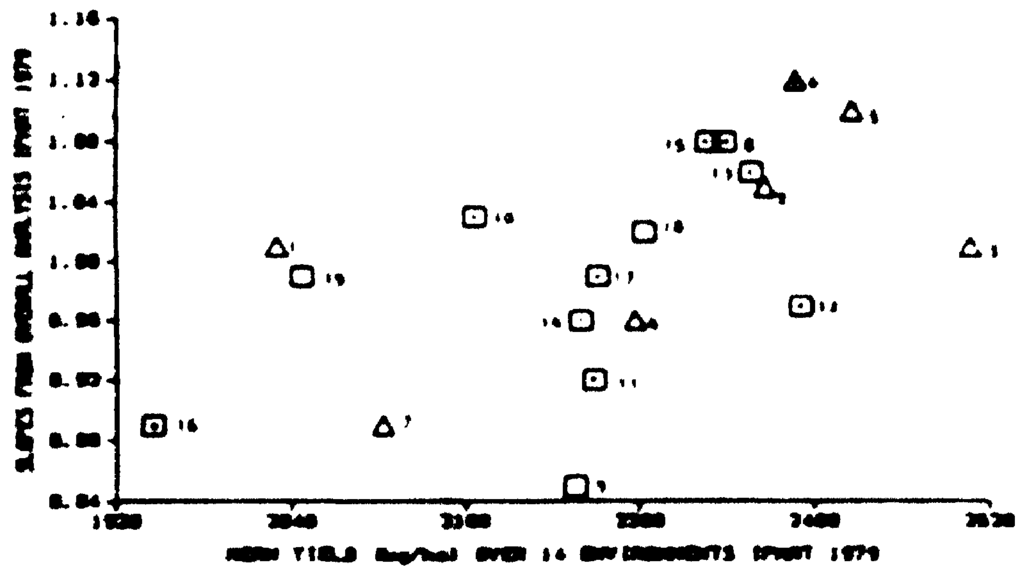


Fig 1c

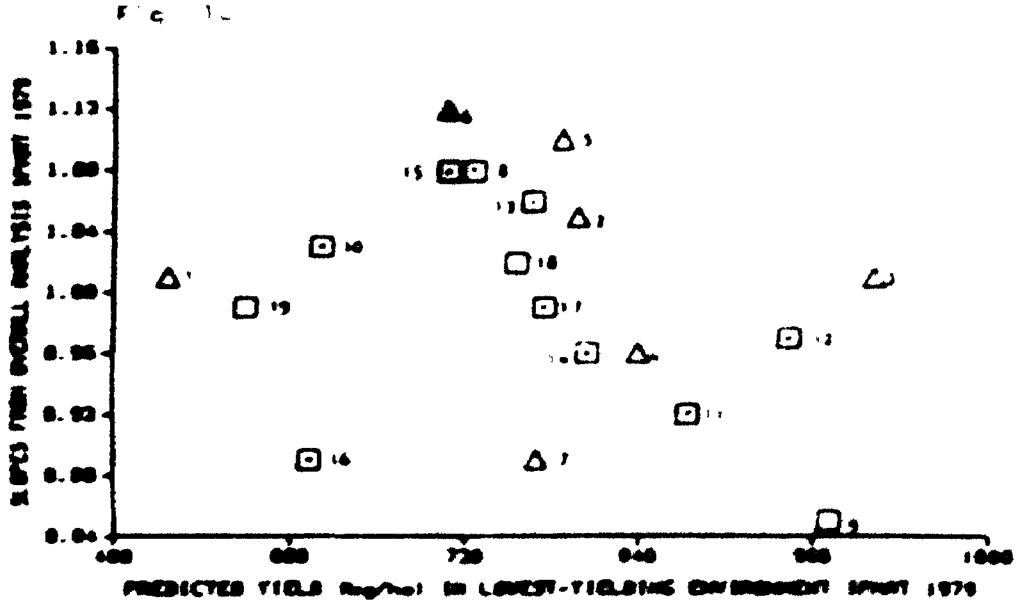


Fig. 2a

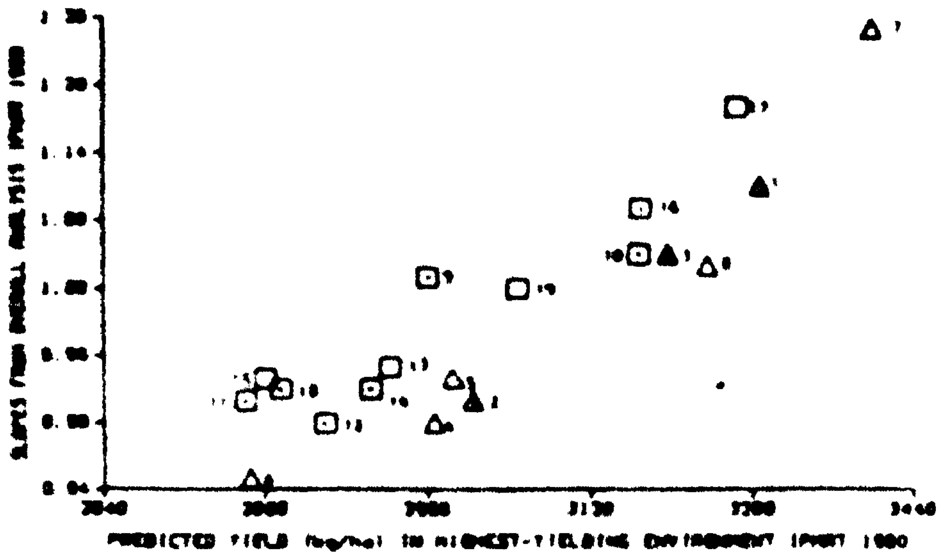


Fig. 2b

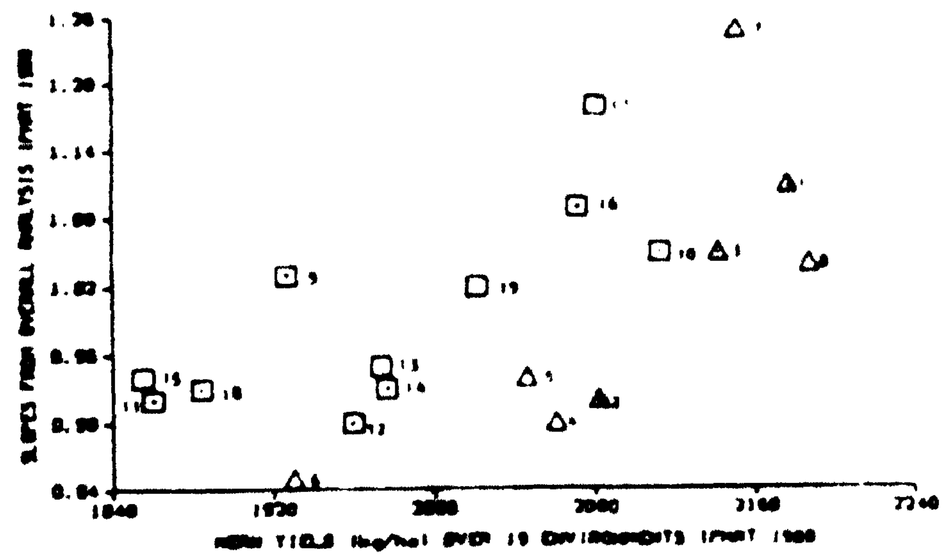


Fig. 2c

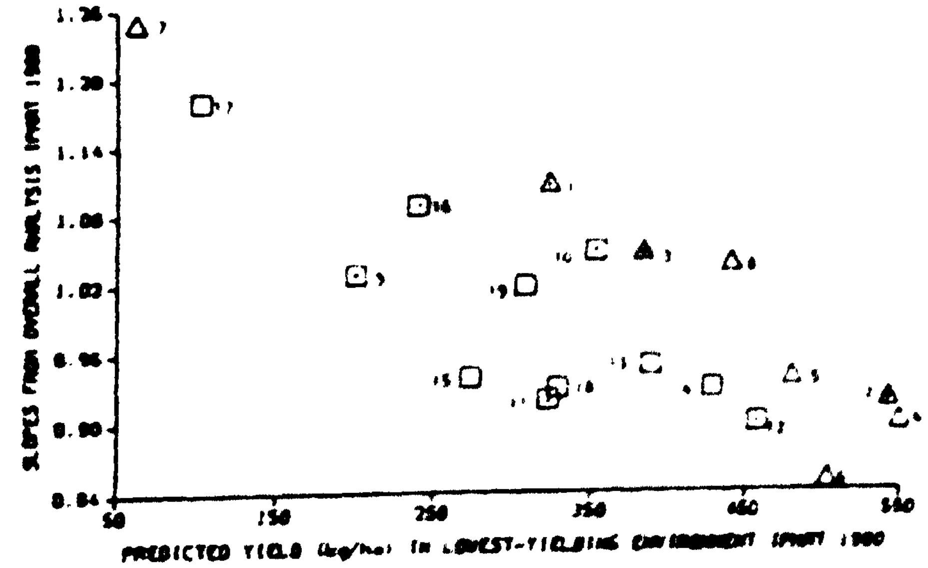


Fig 3a

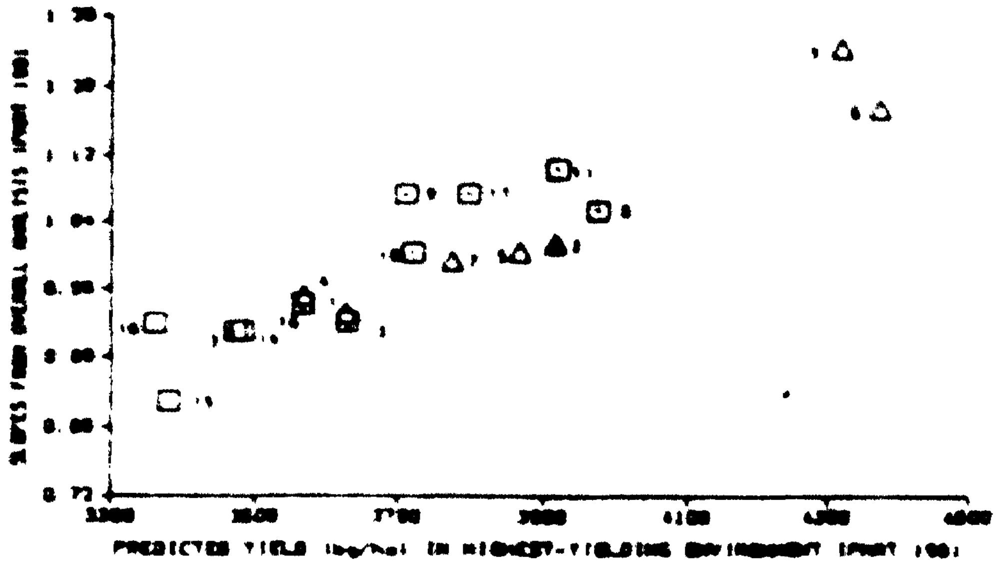


Fig 3b

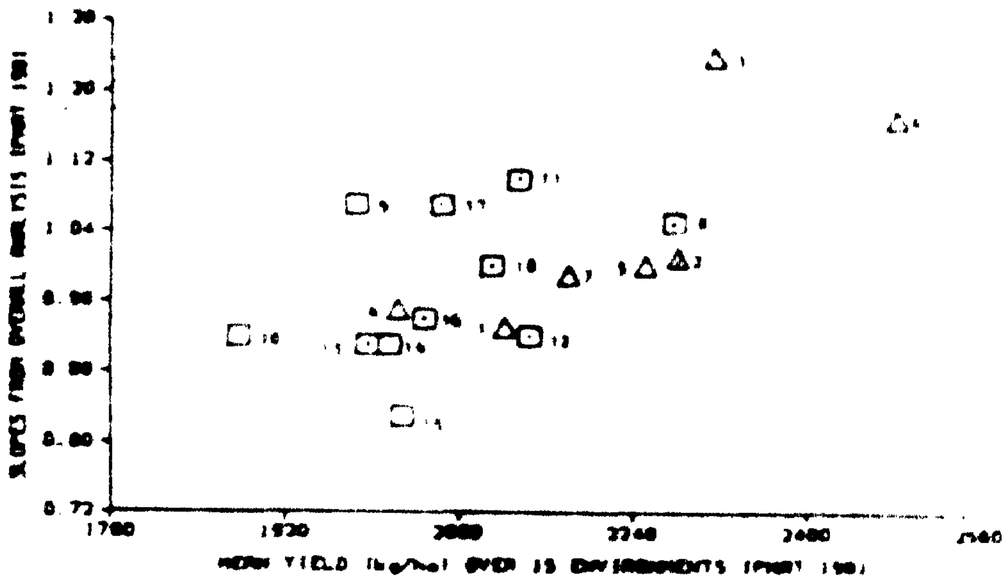


Fig 3c

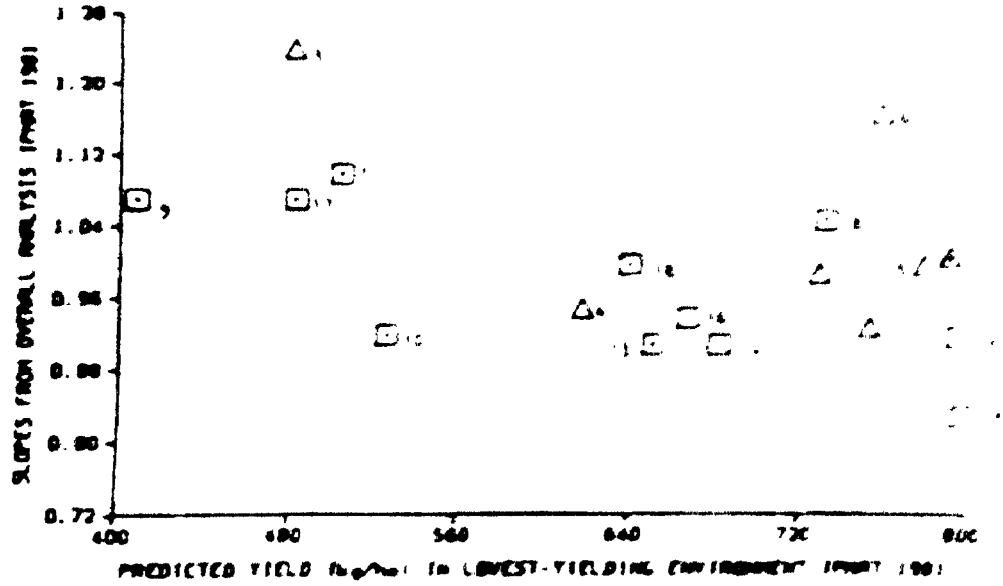


Fig. 4a

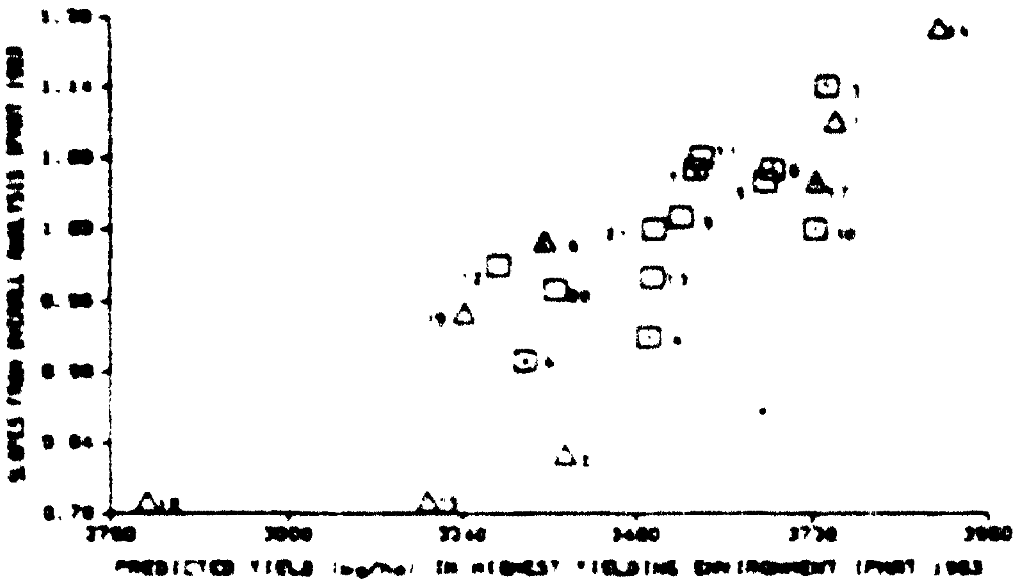


Fig. 4b

Fig. 4c

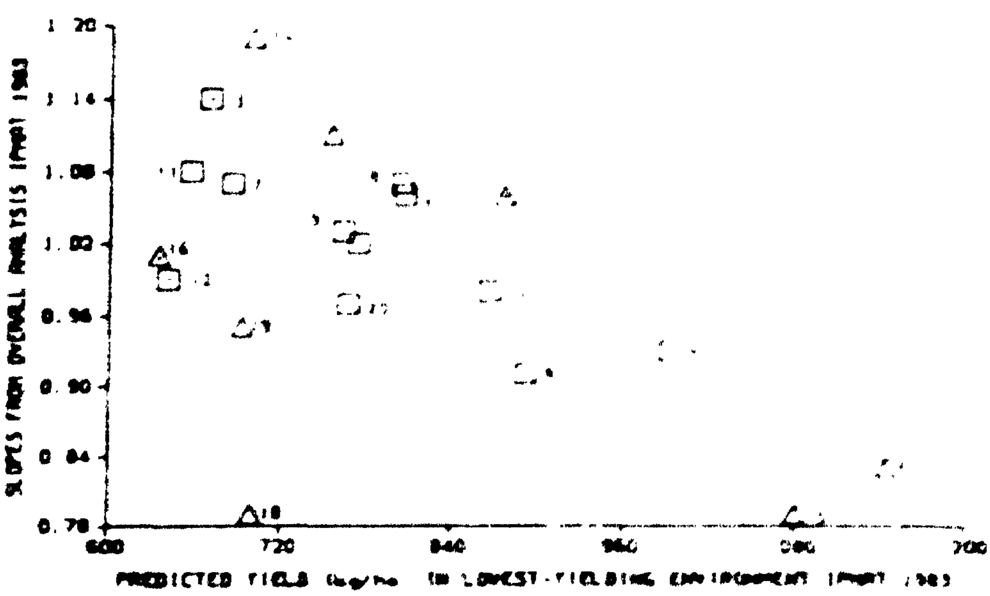


Fig. 5a

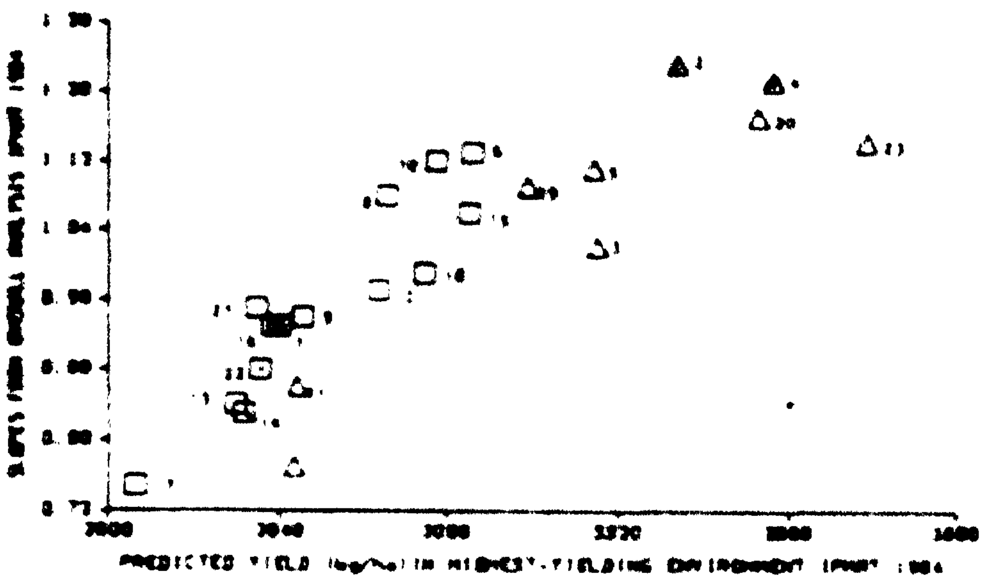


Fig. 5b

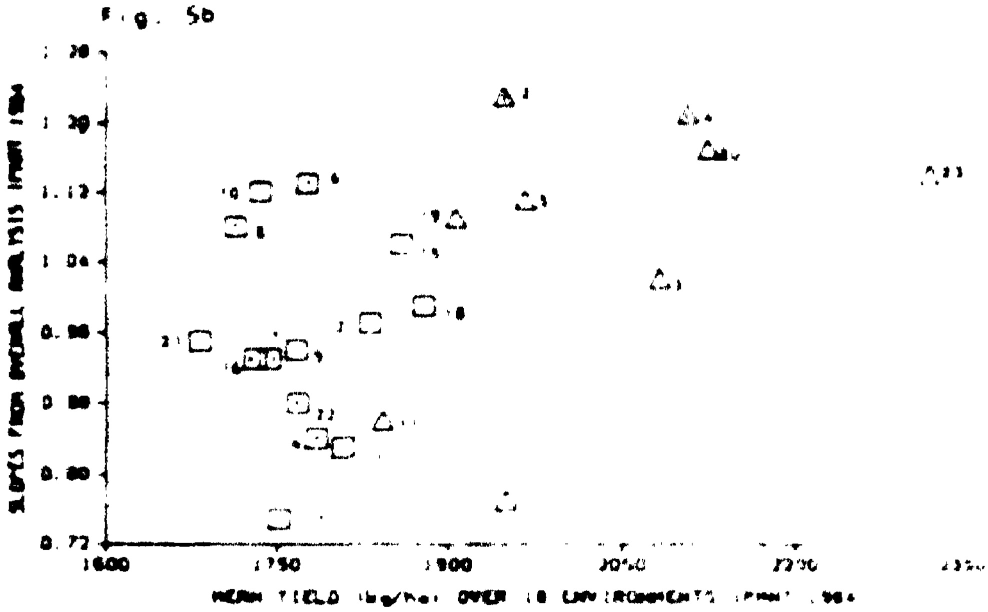
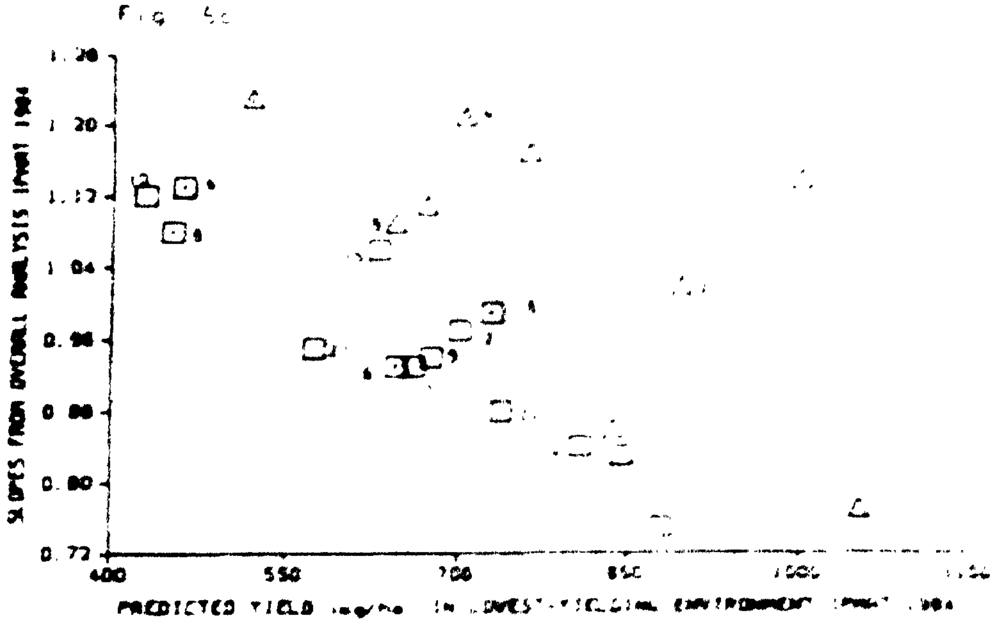


Fig. 5c



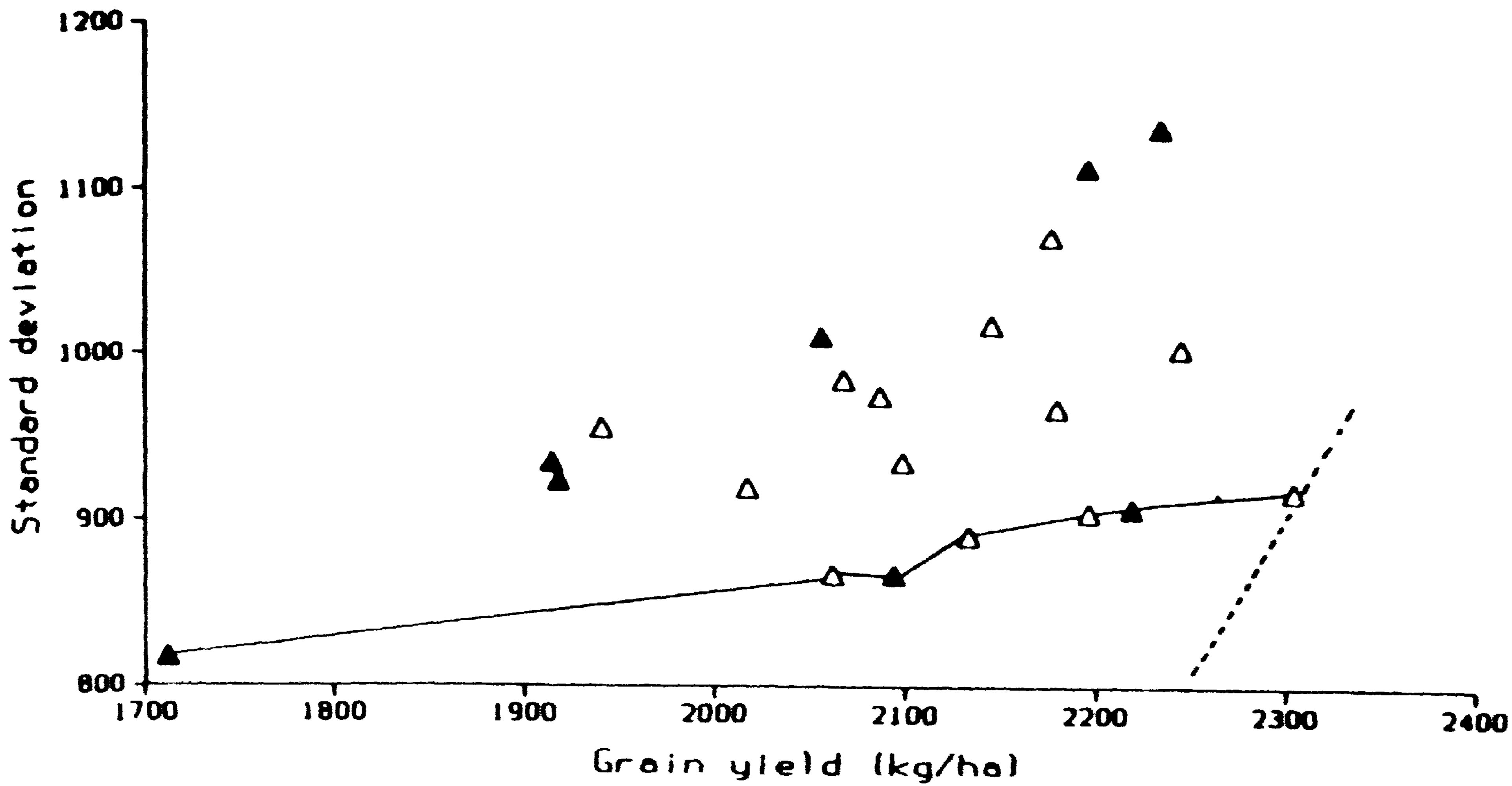


Fig. 6. Mean yield and standard deviation of the entries in IPMAT, 1983. (Iso-utility curve is indicated by a dotted line and the risk efficient frontier by a solid line).

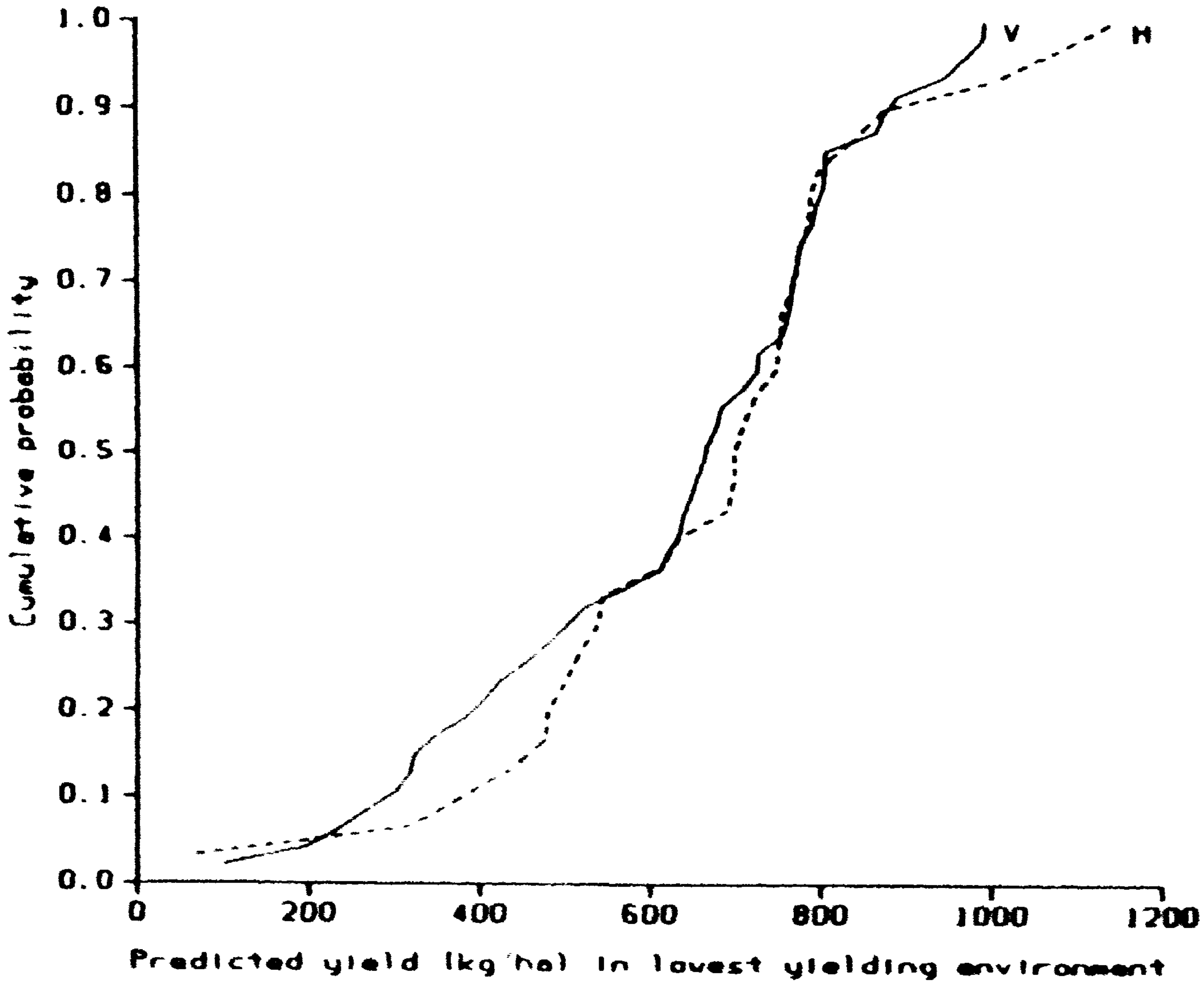


Fig. 7. Cumulative probability and predicted yield in the lowest-yielding environment in IPMAT 1979, 1980, 1981 and 1983.