

## Selecting varieties using a series of trials and a combinatorial ordering method

A Guénoche <sup>1</sup>, B Vandeputte-Riboud <sup>2\*</sup>, JB Denis <sup>3</sup>

<sup>1</sup> LIM-CNRS, 163, avenue de Luminy, F13288 Marseille cedex 9;

<sup>2</sup> DER-CETIOM, 80, avenue Victor-Hugo, F75116 Paris;

<sup>3</sup> INRA, Laboratoire de Biométrie, route de Saint-Cyr, F78026 Versailles cedex, France

(Received 3 December 1993 ; accepted 20 June 1994)

**Summary** — The interpretation of results from a series of experiments designed to evaluate varietal performance is frequently difficult due to the large interaction observed between varieties and location sites. In fact interactions are only relevant if a great variability in the varietal ranking is observed from one environment to another. However, classical statistical approaches do not take this fact into account, and so we have developed new strategies to establish median orders on the variety set. This raises the need for mathematical tools based on the search for total orders on the whole set of varieties. However, this search requires a large numerical (computer) effort and is impracticable because of the large number of possible orders. Thus, we propose a combinatorial method derived from social choice theory which gives orders minimizing a dissimilarity function between overall orderings and the observed orders in the individual trials. The proposed method is illustrated using an example of 14 rapeseed varieties grown in 60 trials. Its merits and limitations are discussed.

**tournament / median order / series of experiments / genotype–environment interaction / cluster / social choice theory / combinatorial algorithm**

**Résumé** — **Sélection de variétés expérimentées dans un réseau d'essais, à l'aide d'une méthode combinatoire.** L'interprétation des données issues de réseaux d'essais de variétés se heurte souvent aux interactions génotype–milieu. En fait, les interactions ne deviennent vraiment pertinentes que lorsqu'elles impliquent des changements de classement d'un essai à l'autre. Mais les approches statistiques classiques ne font pas cette distinction essentielle pour l'expérimentateur, d'où la nécessité de mettre en œuvre d'autres approches basées sur la recherche d'ordres préférentiels sur l'ensemble des variétés étudiées. En empruntant des procédures utilisées en théorie du choix social, nous présentons une méthode combinatoire qui énumère tous les ordres totaux à distance minimum de l'ensemble des ordres observés. L'utilisation de cette technique est illustrée sur un exemple réel de 14 génotypes de colza expérimentés dans 60 lieux. Son intérêt et ses limites sont discutés.

**tournoi / ordre médian / regroupement d'essais / classement / interactions génotype–milieu / théorie du choix social / algorithme**

\* Correspondence and reprints

## INTRODUCTION

Among the numerous challenges of applied agronomic research, varietal evaluation remains important. The objective is to give farmers indicators of the respective value of the varieties and, in each farming context, to answer the fundamental question: "Which variety should I choose?" Despite its apparent simplicity, the available answers to this question are not fully satisfactory, for 2 main reasons.

1) The definition of a good variety is far from simple. Many criteria may be used in order to determine the best variety: how and with what weighting should we take into account technical (yield, resistance to diseases, drought and lodging) and economic (net costs, inputs, and margins) criteria?

2) For most of these criteria, an interaction exists between the variety and the environment (type of soil, location, and year). The best varieties are not the same in all conditions (see Baker, 1988, for tests for crossover genetic-environmental interactions, and Byth *et al*, 1976, for their detection).

There is no fully satisfactory answer to this, but some solutions may be proposed, using the results given by a series of experiments (which might be incomplete) (see Denis and Vincourt, 1982, for a review of statistical methods for the analysis of interactions).

The first solution proposed is a statistical one, consisting of the development of a parametric yield model taking into account interaction (Freeman, 1973; Denis, 1983, 1992; Gauch, 1990). The advantage of this approach is to retain only one criterion for the decision, which seems quite global and objective if the yield and margin are maximized together. The modeling of interaction allows us to modulate the merits of the varieties to be adjusted, depending on the farming context and then on other technical criteria for the choice of varieties (resistance and precocity). Nevertheless the use of this method is limited by 2 facts:

1) The proposed models of interaction may be too rigid to correctly fit the observations.

2) In the new economic context, yield and margin seem to be more and more dissociated. The yield is then insufficient to evaluate varietal interest, and several decision criteria must be considered.

The approach we develop here is an alternative, based on the combinatorial techniques used in social choice theory (Fishburn, 1972; Barthélemy and Monjardet, 1981). The main idea

is to consider each trial as a 'voter' expressing the relative advantages of the tested varieties according to one or several criteria. The 'best' varieties can then be identified as respecting at the most the whole set of the votes. Only the rankings of the varieties are used in the expression of the results. This method is therefore less restrictive, and also allows several criteria to be used simultaneously. This paper presents the basic principles of this method and its application to the evaluation of rapeseed varieties (using yield as the only criterion for the demonstration).

## DESCRIPTION OF DATA

This study uses the observations from 60 trials in the CETIOM series of experiments for 1990. The experimental design is of the Fisher-block type with 4 replications. Not all the 'trial x variety' combinations have been observed, as there are only 600 of them for 60 trials and 14 varieties (840 possibilities). Table I shows the repartition and value of average yields as well as the least significant differences calculated for each trial.

## PROPOSED PROCESS AND TERMINOLOGY

This method is based on paired comparisons. For each couple  $(i, j)$  it is decided collectively whether variety  $i$  is better than the variety  $j$ , and the preference is quantified, a zero value showing indifference. The set of these preferences is called a tournament. The aim of the method is to determine the ordering of varieties that is most compatible with the entire set of preferences. Therefore, an order contradicting the fewest preferences is required, a perfect accordance usually being impossible. There may be several equivalent orders. The varieties that have the best ranks in these orders are then considered as being the best for the set of locations and criteria taken into account.

This method comprises 2 independent steps: (i) building a weighted tournament describing the comparisons between varieties; and (ii) computing total orders for the varieties of the tournament.

### **Step (i). Building the weighted tournament**

First a partial preorder of varieties can be drawn from the observations for each location. Only the

**Table 1.** Average yields of 14 oil rapeseed varieties (columns) checked in 60 trials (rows) in 1990. The last 2 columns show the least square difference at 5% and the number of varieties in each trial.

	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8	VAR9	VAR10	VAR11	VAR12	VAR13	VAR14	LSD	Number of varieties
CV101	43.0	40.0	42.5	40.0	48.3	46.2	44.4	40.3							4.0	8
CV102	22.3	19.4	23.1	19.6	19.7	22.2	21.8	18.1	19.8	22.7					3.1	10
CV103	37.2	37.3	35.0		36.7	37.9	40.5	38.0	40.5	35.4	40.3				4.0	10
CV105	22.0	26.6	26.4		23.2	22.8	25.7	24.3							3.9	7
CV106	36.8	38.6	43.0		40.0	39.7	44.4	40.9	49.0	44.0	49.2				6.0	10
CV107	40.6	36.9	35.2		27.1	42.8	34.2	36.3	34.0	31.7					4.0	9
CV109	38.1	33.7	31.1		34.8	30.1	36.7	32.4			27.9				4.7	8
CV110	44.3	40.4	46.3		45.2	43.0	49.3	42.1			52.8				4.8	8
CV208		47.7	43.4		43.0	49.0	47.9	50.1	46.5	44.3	42.8	43.8	43.6	50.8	2.7	12
CV209		36.0	36.8		37.6	34.5	39.6	31.1		40.4		37.4	39.2	37.3	2.8	10
CV211		28.9	31.1		29.7	25.4	32.8	25.5				29.5	30.3	30.5	3.8	9
CV217		27.8	28.9		27.7	30.6	29.2	32.2				26.8	27.0	27.6	3.4	9
CV219		49.1	48.1			51.6	52.9	50.2	47.9	50.0		48.3	47.4	52.7	6.5	10
CV221		37.0	38.3	35.6	34.8	38.6	37.8	33.7		41.5		35.0	41.2	33.7	3.9	11
CV222		26.1	32.2		31.8	28.0	27.3	28.9		38.4		33.0		31.5	2.4	9
CV223		27.7	27.4	29.6	29.7	30.5	27.0	28.0	34.5	35.0		28.1	29.8	26.7	3.4	12
CV224		27.0	35.3	25.6	31.9	32.4	27.7	29.6				27.7	29.8	27.8	1.6	10
CV302	23.8	22.1	22.8		21.9	26.1	23.4	22.8			26.5				2.4	8
CV303	39.5	40.0	36.2	38.4	38.9	40.2	36.5	38.0	41.1	37.3					5.1	10
CV305	29.5	23.9	24.2	29.5	20.1	23.4	25.0	23.2	21.9	18.5	26.8			25.2	5.0	12
CV306	15.2	13.4	14.6		23.3	13.3	14.4	11.9							4.4	7
CV307	24.4	23.8	23.5		22.0	25.4	25.5	22.9							2.5	7
CV309	16.2	14.4	17.6		17.4	17.2	14.5	14.6							3.0	7
CV401		46.7	44.8	42.4	45.2	40.5	43.3	47.6	41.1	42.2		43.3	45.0	44.7	3.4	12
CV402		43.8	41.6		43.7	46.6	40.0	46.9				41.3	43.4	41.7	4.6	9
CV404		39.8	39.9		40.2	43.0	39.4	38.6				37.7	40.4	38.8	2.6	9
CV405		35.2	30.5		29.2	30.9	31.6	27.9				28.2	30.2	26.0	2.7	9
CV408		38.0	39.7	40.2	36.7	38.2	38.3	37.8				32.4	42.1	30.2	3.8	10
CV410		23.7	25.5		22.2	24.2	25.6	22.1				18.2	32.4	17.7	3.8	9
CV415		36.6	39.4		33.8	35.1	34.7	30.5				32.5	40.6	32.1	3.5	9
CV416		46.1	45.3		44.7	49.2	46.0	46.6				43.7	45.0	44.6	3.0	9
CV418		37.9	39.1		37.4	33.8	36.8	36.4				35.8	36.6	37.1	4.5	9
CV420		37.4	37.9		35.5	39.6	38.6	37.8	42.7	42.0	35.8	33.6	40.7	34.4	4.1	12
CV424		47.6	44.0		44.0	48.0	43.4	48.5				44.3	51.1	41.5	5.6	9
CV425		33.5	39.4		35.5	37.2	34.8	34.5	36.5			36.4	39.9	38.3	4.5	10
CV500		33.9	30.5		30.4	36.3	30.5	34.3	35.3	34.9		30.1	31.5	32.0	3.8	11
CV520		32.6	36.4	35.0	33.4	33.5	33.5	32.0	35.1	41.1	37.8	31.5	40.7	29.6	3.5	13
CV521		29.0	34.9	35.4	29.9	34.4	34.9	36.1	39.8	37.6	31.2	28.2	40.4	27.6	2.1	13
CV523		29.2	36.8	37.0	33.3	30.9	31.9		37.4	40.0	31.2	28.4	40.8	33.5	5.1	12
CV601	48.4	44.1	42.1	46.4	40.1	46.6	46.0	46.4			47.3	47.2	39.9		4.1	11
CV603	37.4	34.0	34.0	37.1	31.3	39.3	34.2	34.7	31.6	33.2	34.9	32.6			5.7	12
CV611	36.8	37.5	39.1	43.0	38.7	40.8	36.7	34.3			36.1	37.4			3.2	10
CV612	31.4	28.1	28.7	28.5	26.9	30.3	25.9	30.7	30.0		29.0	25.7			3.4	11
CV640	35.2	28.7	34.4	34.9	31.8	31.4	32.5	29.8	32.2		34.0	33.2			4.2	11
CV641	33.9	27.4	29.5	29.3	27.2	30.1	27.2	30.1			29.7	24.0			3.7	10
CV670	23.9	23.3	21.6	24.5	17.9	24.6	21.0	24.4	25.2	22.3	24.4	16.6		18.5	2.3	13
CV680	35.3	31.8	34.5	34.2	33.4	32.4	36.7	33.1			34.3	33.2		33.0	4.0	11
CV701	30.3	21.8	27.0		27.0	27.6	27.7	33.0		31.1	29.5	30.4			4.0	10
CV703	34.9	30.7	33.5	35.7	33.9	31.8	33.1	31.3	37.2	39.7	33.9	30.4		32.0	3.2	13
CV705	33.0	26.1	32.5	35.8	32.3	34.7	35.0	31.8			31.0	32.6			4.0	10
CV709	29.8	26.4	26.5	29.4	25.6	27.8	27.7	29.0	26.6	33.0	31.3	24.9		25.5	2.3	13
CV710	25.0	23.4	26.9	27.6	25.5	26.5	27.1	26.5			29.0	23.7			1.7	10
CV711	37.4	33.7	37.1	35.8	33.3	35.8	35.3	32.3			31.1	29.8			2.9	10
CV712	34.4	33.2	34.5	35.8	34.3	34.6	34.3	32.3	40.3	39.2	33.4	33.0			2.0	12
CV714	28.5	27.6	36.4	33.7	34.6	31.2	36.2	30.8	35.6	38.8	34.8	27.5	36.2	34.2	3.4	14
CV717	40.6	37.7	39.7		38.1	39.0	39.4	36.8			40.7	36.9			4.7	9
CV723		26.3	33.0		25.6	29.7	33.9	30.8				25.2		25.0	3.0	8
CV803	36.7	33.9	39.9		41.0	39.1	39.0	36.5							3.9	7
CV804	32.2	32.0	36.3	36.1	31.6	31.6	33.8	33.7		35.1	32.8				5.1	10
CV807	40.8	38.4	46.2		48.2	42.8	45.0	36.3							5.1	7

varieties grown at the location are ranked, in our example by decreasing yields; several varieties may have the same rank. Paired comparisons may be drawn from the whole of these preorders. Let  $T(i,j)$  be the number of locations where  $i$  is better than  $j$  associated with each  $(i,j)$ . These values can be stored in a non-symmetric matrix  $T$ , the number of rows and columns of which is  $n$ , the number of varieties compared. This information can be considered as a graph where each node is a variety and each arc between  $i$  and  $j$  is weighted

by the number of locations where  $i$  is better than  $j$ . It is admitted that there is a collective preference for  $i$  over  $j$  if  $T(i,j) > T(j,i)$ . If the varieties  $i$  and  $j$  have never been tried together in the same location then the 2 arcs have a zero value.

**Step (ii). Search for median orders**

The total order that is the most compatible with these preferences is then required over all the

varieties. For this purpose a measure of distance is chosen, allowing the evaluation of the difference between any total order on varieties and the matrix  $\mathbf{T}$ . If  $i$  is better than  $j$  in an order compatible with tournament  $T$ , then  $T(i,j)$  is greater than or equal to  $T(j,i)$ . In the alternative case  $T(j,i) - T(i,j)$  is the number of majority opinions preferring  $j$  over  $i$ . The distance between an order  $O = (o_1, o_2, \dots, o_n)$  and tournament  $T$  is the sum of the majority opinions over all the pairs  $(i,j)$  that are not compatible with  $O$ . This point has been explained in more detail by Barthélémy *et al* (1989). Formally:

$$d(O, T) = \sum_{i < j} T(o_i, o_j) - T(o_j, o_i) \quad [1]$$

An order at distance zero only contains majority preferences, and then the tournament is transitive. Generally this is not the case. The order(s) minimizing this distance are then searched for. Such an order is termed a 'median order'. There is usually more than one median order; there may be more than the number of varieties in the tournament, and all these orders are equivalent as a collective opinion. This is why we need to enumerate all the median orders, and we propose keeping the varieties that are ranked first in at least one median order. In the literature these selected elements are termed winners (Kemeny, 1959).

The next sections give a detailed presentation of both steps. Table I presents data for their practical application. In this way, we can mention the problems encountered and develop solutions to overcome them.

## WEIGHTING THE TOURNAMENT

### Direct method

Let  $S$  be the set of locations and  $V$  the set of varieties.  $X_s(i)$  is the yield of the variety  $i$  in the  $s$  location. If it has not been grown, we set  $X_s(i) = 0$ . The roughest way to operate is to consider that the variety  $i$  is better than the variety  $j$  if  $X_s(i) > X_s(j) > 0$ . If  $X_s(i) = 0$  or  $X_s(j) = 0$  or  $X_s(i) = X_s(j)$ , and no preference is expressed between these varieties. The tournament can then be weighted:

$$T(i,j) = \text{Card}(\{\delta \in S \text{ such as } X_s(i) > X_s(j)\}) \quad [2]$$

In doing so, however, 2 facts must be considered: (i) small yield differences may not always

allow us to conclude that one variety is better than another; and (ii) large yield differences are treated like small ones.

As a consequence, no difference is made in the tournament between hasardous and sure yield differences. This practical problem led Frey and Yehia-Alcoutlabi (1986) to modify the previous distance between order and tournament so that the discrepancies observed within the data would be taken into account (see Young, 1986, and Charon and Hudry, 1992, for discussions on ranking methods for the weighting tournament).

Two possible approaches have been tested in order to obtain a better comparison of varieties in each location. The first uses the results of the analysis of variance undertaken in each trial and the associated multiple comparison of means. This point will be developed in the next section. The second uses a partition of varieties into ranked clusters obtained from a clustering method, and then defines the rank differences of the classes as a weight. This point will be developed later.

### Standardized deviations from the multiple comparison of averages in each trial

Before grouping, the results of each trial are individually interpreted. The variance analysis, which may be followed by a multiple comparison of means, makes it possible to establish a partial order on the expected yields of the varieties in each trial. Our proposal is to use these preorders to weight the tournament.

As the number of varieties tested may differ between trials, we used the least significant difference as a multiple comparison method, the results of which are unaffected by the number of varieties. In each  $s$  trial, we associate the  $t_s(i,j)$  value which is the score of  $i$  versus  $j$  in  $s$  to every  $(i,j)$  couple of varieties, so that

$$t_s(i,j) = (X_s(i) - X_s(j)) / \text{lsd}_s(\alpha),$$

$$\text{if } (X_s(i) - X_s(j)) / \text{lsd}_s(\alpha) > 1$$

$$t_s(i,j) = 0, \text{ elsewhere.}$$

In this formula,  $\text{lsd}(\alpha)$  is the least significant difference (Dagnélie, 1970) for a significance level  $\alpha$ , with  $\tau_k, 1 - \alpha/2$ , the value taken in Student's table for  $k$  degrees of freedom with a probability of  $1 - \alpha/2$ , and  $r$  the number of replications in each trial. Its value is given by

$$\text{lsd}(\alpha) = \sigma \cdot \tau_k, 1 - \alpha/2 \cdot (2/r)^{1/2}$$



**Table IV.** Tournament obtained with the standardized deviations from the least significant difference method.

	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8	VAR9	VAR10	VAR11	VAR12	VAR13	VAR14
VAR1	0.0	12.4	2.1	0.0	10.3	2.9	1.3	11.1	0.0	0.0	0.0	15.8	0.0	3.2
VAR2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.4	0.0	5.5
VAR3	0.0	30.3	0.0	2.2	11.6	0.0	6.2	17.1	0.0	0.0	0.0	31.1	0.0	21.6
VAR4	4.0	18.7	0.0	0.0	8.2	0.0	2.5	5.5	0.0	0.0	9.3	25.6	0.0	12.3
VAR5	0.0	12.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.3	0.0	5.8
VAR6	0.0	19.9	0.9	2.8	14.8	0.0	6.8	16.8	0.0	0.0	0.0	34.1	0.0	23.2
VAR7	0.0	20.8	0.0	0.0	6.9	0.0	0.0	10.7	0.0	0.0	1.0	22.7	0.0	15.0
VAR8														
VAR9	2.1	18.7	13.0	3.8	19.9	9.4	15.2	9.5	0.0	0.0	10.8	27.9	0.0	14.2
VAR10	5.5	31.1	18.3	10.9	31.5	16.5	17.0	22.6	1.5	0.0	10.2	37.0	0.5	25.3
VAR11	4.2	18.7	7.0	0.0	15.7	2.9	0.0	5.5	0.0	0.0	0.0	19.1	0.0	7.1
VAR12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
VAR13	2.3	19.5	3.9	8.1	17.8	10.4	12.7	17.1	0.7	0.0	7.5	28.9	0.0	26.1
VAR14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.6	0.0	0.0

other variety; VAR3, VAR1, VAR7 and VAR6 are an intermediate group; and VAR2, VAR4 and VAR8 have lower yields than the other varieties.

The partial order is then: (VAR5 > VAR3, VAR1, VAR7, VAR6 > VAR2, VAR4, VAR8).

A more theoretical justification for this is that a maximum split partition is built, the smallest inter-class distance is maximized, and the partition is then optimal according to this criterion.

The number of classes may of course be adapted according to the difference between the extreme values, or the absolute or relative differences between yield values, for example, in order to get a larger split than a defined percentage of the maximum difference. However, we must emphasize that this maximal difference would tend to give more importance to trials where large variations are observed.

Let us denote by  $C_s(i)$ , the rank of the class containing the variety  $i$  in site  $s$ , with the convention that the best varieties are in the class ranked 1, and the worst in the class ranked 3. The pre-orders obtained through this method are shown in table V.

### From orders to tournament

Now, since we have ordered classes of varieties for each trial, the natural process is to weight the tournament taking into account the difference between class ranks. If in trial  $s$ , 2 varieties  $i$  and  $j$  are in the same class, the contribution of this trial to the tournament for the comparison of  $i$  and  $j$  is null. If the varieties are in 2 different classes,

the trial contributes to the tournament through the difference between class ranks. The tournament thus is weighted:

$$T(i,j) = \text{Max}\{0, \sum_{s \in S} (C_s(j) - C_s(i))\} \quad [4]$$

A positive value for  $T(i,j)$  shows a collective preference for  $i$  over  $j$ , and in this case  $T(j,i) = 0$ . At least one of any 2 symmetrical values is positive. If both are zero, it means collective indifference.

The tournament over the 14 varieties is shown in table VI. There again variety 10 is better than all the others and variety 12 is the worse.

### The multi-criteria case

The evaluation may be undertaken using several criteria weighted according to their relative importance. For each criterion, the process is the same as above, and then a weighted sum of the criterion-specific tournaments gives the final tournament.

### ENUMERATION OF MEDIAN ORDERS

For the computation of all the orders at minimum distance orders of the weighted tournament, we cannot try all possible orders, even for a small number of varieties. In our example, there are  $14! \gg 87 \times 10^9$  possibilities.

**Table V.** Partial preorders through 3 classes maximum separation partition. 0 indicates that the variety was not experimented in the trial.

	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1
CV101	2	3	2	3	1	2	2	3	0	0	0	0	0	0	0
CV102	1	2	1	2	2	1	1	3	2	1	0	0	0	0	0
CV103	2	2	3	0	2	2	1	2	1	3	1	0	0	0	0
CV105	3	1	1	0	3	3	1	2	0	0	0	0	0	0	0
CV106	3	3	2	0	3	3	2	3	1	2	1	0	0	0	0
CV107	1	2	2	0	3	1	2	2	2	2	0	0	0	0	0
CV109	1	2	2	0	2	2	1	2	0	0	3	0	0	0	0
CV110	3	3	3	0	3	3	2	3	0	0	1	0	0	0	0
CV208	0	1	3	0	3	1	1	1	2	3	3	3	3	1	1
CV209	0	2	2	0	2	2	1	3	0	1	0	2	1	2	2
CV211	0	2	2	0	2	3	1	3	0	0	0	2	2	2	2
CV217	0	3	3	0	3	2	3	1	0	0	0	3	3	3	3
CV219	0	3	3	0	0	2	1	3	3	3	0	3	3	1	1
CV221	0	2	2	3	3	2	2	3	0	1	0	3	1	3	3
CV222	0	3	2	0	2	3	3	3	0	1	0	2	0	2	2
CV223	0	3	3	2	2	2	3	3	1	1	0	3	2	3	3
CV224	0	3	1	3	2	2	3	3	0	0	0	3	3	3	3
CV302	2	3	2	0	3	1	2	2	0	0	1	0	0	0	0
CV303	2	2	3	2	2	2	3	2	1	2	0	0	0	0	0
CV305	1	2	2	1	3	2	2	2	2	3	2	0	0	0	2
CV306	2	2	2	0	1	2	2	3	0	0	0	0	0	0	0
CV307	2	2	2	0	3	1	1	2	0	0	0	0	0	0	0
CV309	2	3	1	0	1	1	3	3	0	0	0	0	0	0	0
CV401	0	1	2	3	2	3	3	1	3	3	0	3	2	2	2
CV402	0	2	3	0	2	1	3	1	0	0	0	3	2	3	3
CV404	0	2	2	0	2	1	2	2	0	0	0	3	2	2	2
CV405	0	1	2	0	2	2	2	2	0	0	0	2	2	3	3
CV408	0	1	1	1	1	1	1	1	0	0	0	2	1	3	3
CV410	0	2	2	0	2	2	2	2	0	0	0	3	1	3	3
CV415	0	2	1	0	2	2	2	3	0	0	0	2	1	2	2
CV416	0	2	2	0	2	1	2	2	0	0	0	3	2	2	2
CV418	0	2	1	0	2	3	2	2	0	0	0	2	2	2	2
CV420	0	2	2	0	3	2	2	2	1	1	3	3	2	3	3
CV424	0	2	3	0	3	2	3	2	0	0	0	3	1	3	3
CV425	0	3	1	0	3	3	3	3	0	0	3	1	2	2	2
CV500	0	2	3	0	3	1	3	2	2	2	0	3	3	3	3
CV520	0	2	2	2	2	2	2	2	2	1	2	2	1	3	3
CV521	0	2	1	1	2	1	1	1	1	1	2	2	1	2	2
CV523	0	3	2	2	3	3	3	3	0	2	1	3	3	1	3
CV601	1	1	2	1	3	1	1	1	0	1	1	3	0	0	0
CV603	2	3	3	2	3	1	3	3	3	3	3	3	0	0	0
CV611	2	2	2	1	2	2	2	3	0	0	2	2	0	0	0
CV612	1	1	1	1	2	1	3	1	1	0	1	3	0	0	0
CV640	1	3	1	1	1	1	1	2	1	0	1	1	0	0	0
CV641	1	2	2	2	2	2	2	2	0	0	2	3	0	0	0
CV670	1	1	1	1	2	1	1	1	1	1	1	3	0	2	2
CV680	2	3	3	3	3	3	1	3	0	0	3	3	0	3	3
CV701	1	3	2	0	2	2	1	0	1	1	1	0	0	0	0
CV703	3	3	3	3	3	3	3	2	1	3	3	0	3	3	3
CV705	2	3	2	1	2	1	1	2	0	0	2	2	0	0	0
CV709	3	3	3	3	3	3	3	3	1	2	3	0	3	3	3
CV710	2	3	2	2	2	2	2	2	0	0	1	3	0	0	0
CV711	1	2	1	1	2	1	1	2	0	0	2	3	0	0	0
CV712	3	3	3	2	3	3	3	3	1	1	3	3	0	0	0
CV714	3	3	2	2	2	3	2	3	2	1	2	3	2	2	2
CV717	1	2	2	0	2	2	3	3	0	0	1	2	0	0	0
CV723	0	3	1	0	3	2	1	2	0	0	0	3	0	3	3
CV803	2	3	1	0	1	1	1	2	0	0	0	0	0	0	0
CV804	3	3	1	1	3	3	3	3	0	2	3	0	0	0	0
CV807	1	2	1	0	1	1	1	3	0	0	0	0	0	0	0

limitation in the exploration of the tree. The initial parts are progressively eliminated in the development of the tree as soon as their distance to the tournament becomes too large to make it possible for them to give a solution. This method of building orders was detailed by Barthélemy *et al* (1989), and we present it only briefly here with some improvement due to the fact we are only interested in the best varieties that correspond to the first-ranked ones.

**Initialization of the algorithm using heuristic methods**

The initialization consists of the estimation of an upper bound for the distance between a median order and the tournament, thus defining a rejection threshold. The smaller this bound, the greater the efficiency of the algorithm, and the quicker the solution. First, we search for orders close to the tournament, using heuristic methods to evaluate the bound.

The strategy of classical heuristics is the association of a score to each variety, the varieties being ranked by increasing scores. The scores for varieties may be the number of varieties collectively preferred to *i*, or the sum of the arc values coming from *i*, etc. These heuristics are generally inefficient compared with those based on the elimination of cycles, which are the adaptation to a weighted tournament of Smith and Payne's heuristics (1974). Some stochastic optimization techniques, such as simulated annealing (Hudry, 1989), can also be used.

For example, on the tournament in table IV, the order achieved by Smith and Payne's heuristics is: VAR10 > VAR13 > VAR9 > VAR4 > VAR11 > VAR1 > VAR6 > VAR3 > VAR7 > VAR8 > VAR5 > VAR2 > VAR14 > VAR12. The distance of this to the tournament is 6, which is kept as the upper bound L of the minimum distance of an order to the tournament in the branch and bound tree.

**'Branch and bound' algorithm**

This is not the place to detail how branch-and-bound methods work, nor to explain all the details of our algorithm. We refer to Barthélemy *et al* (1989) or Reinelt (1985). Like any branch-and-bound method, this algorithm builds up a tree, the vertices of which are the beginning parts

In order to reduce the number of examined permutations without omitting any minimum distance order, a 'branch and bound' algorithm is used, based on the connection of 2 processes. The first one is an arborescent search. The solution order can be seen as a tree, the nodes of which correspond to the initial orders on parts of subsets of varieties. The second is a progressive

**Table VI.** Tournament obtained with the 3 classes of the maximum separation partition.

	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	VAR8	VAR9	VAR10	VAR11	VAR12	VAR13	VAR14
VAR1	0	18	3	0	13	0	0	17	0	0	0	14	0	2
VAR2	0	0	0	0	0	0	0	0	0	0	0	16	0	8
VAR3	0	19	0	0	19	0	0	17	0	0	0	26	0	13
VAR4	2	13	2	0	11	2	6	11	0	0	8	19	0	7
VAR5	0	0	0	0	0	0	0	0	0	0	0	13	0	4
VAR6	0	23	4	0	22	0	4	22	0	0	0	30	0	13
VAR7	0	19	0	0	17	0	0	18	0	0	1	25	0	12
VAR8	0	1	0	0	1	0	0	0	0	0	0	19	0	8
VAR9	4	12	9	2	17	4	9	10	0	0	7	16	0	5
VAR10	5	18	16	8	25	11	13	16	1	0	9	25	3	14
VAR11	2	13	4	0	15	1	0	8	0	0	0	13	0	1
VAR12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VAR13	1	8	6	5	13	4	8	7	1	0	5	22	0	16
VAR14	0	0	0	0	0	0	0	0	0	0	0	0	0	0

of total orders, which are chains of varieties, and which may lead to a median order.

Every time a branch  $X$  is extended, its distance  $D(X)$  to the tournament is calculated and compared with  $L$ , the current upper bound of the minimal distance of an order to the tournament. If this bound is overpassed, the branch is erased because it can only generate orders that will exceed the minimum distance.

This is an iterative algorithm, the elementary step of which is the following: Let  $X = (o_1, \dots, o_p)$  be a node of the tree with minimum value. This is a total order on  $p$  varieties ( $p < n$ ), where  $o_k$  is the variety ranked  $k$  in the  $X$  order. Its value  $D(X)$  is defined by:

$$D(X) = \sum_{k>l} T(o_l, o_k) \tag{5}$$

If  $D(X) > L$ , any order beginning with the  $X$  varieties cannot be a median order; this branch is stopped. If  $D(X) \leq L$ ,  $X$  may be extended with a variety  $i$  which does not belong to  $X$ . If  $p = n - 1$ ,  $i$  is the last unused variety and the order is at the minimum distance of  $T$ , which is the current value  $D(X)$ . If this is not the case the tree is developed from  $X$ . The variety  $i$  can be added at the end of  $X$  if it can be verified that:  $i$  is not included in  $X$ ;  $i$  is not preferred to  $o_p$ , since a median order is a Hamiltonian path in the tournament.

$$i \in V - \{o_1, o_2, \dots, o_p\}; \tag{6}$$

This must thus verify that  $T(i, o_p) = 0$ ; if this is not the case, the initial section  $(o_1, \dots, o_{p-1}, i, o_p)$

would be a best ordering on  $X \cup \{i\}$ ; and the distance value of  $(X, i) = (o_1, o_2, \dots, o_p, i)$  is lower than  $L$ . This value is the distance of  $X$  increased by the new cost coming from the integration of  $i$ :

$$V[(X, i)] = D(X) + \sum_{k \in V - \{o_1, \dots, o_p\}} T(k, i) \tag{7}$$

In order to enumerate all the median orders, the tree is developed until any terminal node has a value exceeding  $L$ .

For example, for the tournament calculated from the yield differences exceeding the least significant difference (table IV), there is only one median order, which is that found by the heuristic method. For the tournament calculated from the differences between yield classes (table VI), there are 12 solutions at distance 1. As varieties VAR2 and VAR5 are consecutive and as there is no collective preference between them, we show only 6 orders in table VII. To get the other 6, VAR2 and VAR5 have to be permuted.

This kind of method encounters a major difficulty, due to the complexity of the consensus function in order relations (Fishburn, 1972; Barthélémy and Monjardet, 1981) and more precisely to the computing of a median order (Hudry, 1989). The amount of computer memory needed is proportional to the size of the tree and the above-mentioned algorithm allows a limitation of this size. However, it is not always sufficient practically to solve the problem. The next sections set out 2 techniques allowing further restrictions on the tree size, thus increasing the possibilities of finding satisfactory solutions.



**Table VII.** Six solution orders at distance 1.

(1)	VAR10 > VAR13 > VAR9 > VAR4 > VAR6 > VAR7 > VAR11 > VAR1 > VAR3 > VAR8 > VAR2 > VAR5 > VAR14 > VAR12
(2)	" >VAR11 > VAR6 > VAR7 > VAR1 > VAR3 "
(3)	" >VAR11 > VAR1 > VAR6 > VAR7 > VAR3 "
(4)	" >VAR11 > VAR6 > VAR1 > VAR7 > VAR3 "
(5)	" >VAR11 > VAR1 > VAR6 > VAR3 > VAR7 "
(6)	" >VAR11 > VAR6 > VAR1 > VAR3 > VAR7 "

**Decreasing the number of terminal vertices**

A set of varieties ordered in different ways may create several nodes in the tree. Our proposal is to keep only the order(s) of least value in every section.

Let the beginning section ( $i > j > k > l$ ) be weighted  $v$ . If in the further development of the tree, the same subset (differently ordered) is obtained, for example ( $j > k > l > i$ ) weighted  $v'$ , there are 3 possibilities: i) if  $v' = v$ , this section is kept in the tree; ii) if  $v' > v$ , this vertex is not relevant as it cannot lead to a median order; and (iii) if  $v' < v$ , the value of subset  $\{i, j, k, l\}$  is updated to  $V(\{i, j, k, l\}) = v'$ , and only section  $\{j, k, l, i\}$  is kept in the tree.

Testing this idea for the enumeration of all median orders shows that the resulting tree has roughly 3 times less vertices. However, the data structure must be adapted (characteristic of function of any subset of varieties as a binary chain) in order to keep the best order for each subset, but the global result is by far the better.

With a set of varieties for which a consensus on the order of relations is hard to find, the first step may be the search of a single median order. Possibilities (i) and (ii) may then be merged, since we already have an optimal order on this subset and it is useless to get another equivalent order. In another computation, the tree size is then further reduced, increasing the possibilities of finding a solution. With this median order, L can be initialized at its exact value, which may allow us to get the enumeration of the other solutions through reprocessing with the former algorithm.

**Use of minimum ranks**

Let L be the exact distance of a median order, or its approximation. We term a L-order, as a total order having a distance to the tournament not

greater than L. For such an order, a minimum (or maximum) rank for each variety may be determined. This rank is the lowest (or the highest) position it can have in an order, the distance of which is lower than or equal to L. Maybe there is no L-order where it has that position, but in any L-order it cannot be placed before (or after) that rank.

If variety  $k$  may be ranked first in a median order, which means that:

$$\sum_{i \in V} T(i, k) \leq \Delta \tag{8}$$

If it may be ranked after  $j$ , this means that:

$$\sum_{i \neq j} T(i, k) \leq L \tag{9}$$

However  $j$  may be such that  $T(j, k) = \text{Max}_i T(i, k)$

$$\sum_{i \in V} T(i, k) - \max_{i \in V} T(i, k) \leq L \tag{10}$$

Generally speaking, let  $l_p$  be the set of indexes of the  $p$  greatest values in  $\{T(i, k)\}_{i=1, \dots, n}$ . The variety  $k$  may be ranked  $(p + 1)$  in the median order if the sum of the  $n - p + 1$  smallest values does not exceed [11].

$$\sum_{i \in V - l_p} T(i, k) \leq L \tag{11}$$

Using the same reasoning on the rows of matrix  $T$ , if  $k$  is placed at rank  $p + 1$  in a L-order, the sum of the  $p$  smallest values of  $\{T(i, k), i = 1, \dots, v\}$  is smaller than L. For every variety, we can then calculate minimum and maximum ranks in a median L-order.

For the minimum rank, which is the only interesting one here, as we are looking for the best varieties, the values of  $\{T(i, k)\}_{i=1, \dots, n}$  have only to be set in the increasing order and then the smallest values are added until L is reached. If

$p$  values can be added, there are  $n - p$  varieties that must be placed before  $k$  so that it can disappear;  $k$  is then at the best ranked  $n - p + 1$ . If  $k$  arise before this place such an order will be at a distance to the tournament greater than  $L$ .

For example, following the heuristic approach on the second tournament, we looked for orders with the distances inferior or equal to 1. The variety 1 cannot be first in such an order, as it is dominated by 4, 9, 10, 11 and 13 and an order starting by 1 would be at least at distance 14 from the tournament. However if 4, 9, 10 and 11 are ranked ahead, 1 can be put before 13 and the distance does not exceed 1. So, without trying to estimate the cost of an order ranking 4, 9, 10 and 11 before 1, it can be said that 1 cannot be ranked better than fifth.

In the worst case the same variety will come eighth. As a matter of fact, 4, 9, 10 and 11 have to be ranked before 1. Varieties 6, 7 and 13, which are not dominated by 1, may also be placed, but with any other variety the distance of this starting section to the tournament would exceed 1. The minimum and maximum ranks of the varieties in median orders are the following:

Varieties:

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Minimum rank:

5 10 8 4 10 5 6 10 2 1 5 14 2 12

Maximum rank:

8 12 9 4 12 7 9 11 3 2 7 14 3 13

We shall remark that only VAR10 may be ranked 1; VAR9 and VAR13 may be ranked 2. Only VAR13 is actually ranked second in a median order, as it dominates VAR9. VAR1 never reaches its minimum rank, being sixth at the best.

**Sub-tournament among the first  $k$  ones**

The distance value between an order and the tournament is not increased by the part of distance coming from the last ranked varieties. However to select the best varieties, we are not concerned by this part of distance. We would like to consider the sub-tournament of the varieties which are always, or most of the time (or at least once), in the first half or even first in a median order. We thus eliminate those which will never have a good rank. If we search for a winner according to Kemeny (the winner in a median order), we may consider only the sub-tournament of varieties with minimum rank 1, and all the oth-

ers may be eliminated. Here only variety 10, which is certainly the best (according to this type of evaluation), would subsist.

If we want to know which varieties are ranked in the first half, we consider the sub-tournament of those with minimum rank lower than or equal to 7, *ie* VAR1, VAR4, VAR6, VAR7, VAR9, VAR10, VAR11 and VAR13, which are the first 8 ones. We then obtain 4 median orders, which will be initial sections of the former orders.

Thus the risk of having a non-computable tournament, which is considered as a penalty when using combinatorial methods, is no longer a problem when only the best candidate(s) is (are) searched for.

In this case, only an upper boundary of the distance of a median order to the tournament is needed. The better this boundary, the sharper is the interval between minimum and maximum ranks for each variety, and the quicker the lower ranks will be discarded.

**LEAST-SQUARES ESTIMATION**

Having determined a median order for the whole set of varieties, it may be useful to fit the data under the constraint of respecting this order. Such an adjustment may be obtained using the classical least-squares criterion, as demonstrated by Denis (1979). Table VIII shows the residual sums of squares for the 12 median orders from table VII. They are between 1 852 and 1 950 and are all smaller than the sum of squares of the adjustment obtained by interaction modelling with one multiplicative term (2 210).

**Table VIII.** Sums of squared error from the 12 models under order restrictions (table VII), and from additive and multiplicative interaction models.

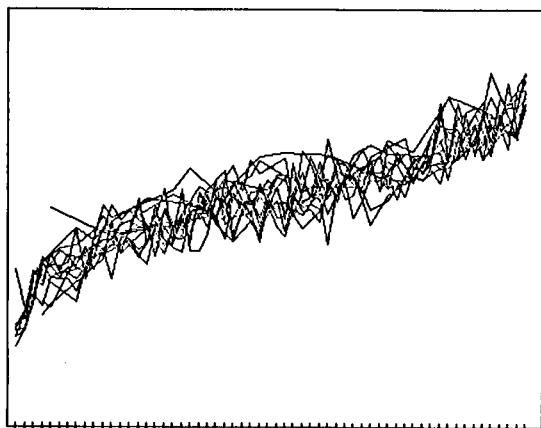
<i>Fitting</i>	<i>SSE</i>	<i>SSE with VAR2 and VAR5 permuted</i>
Order 1	1 935	1 950
Order 2	1 884	1 899
Order 3	1 867	1 880
Order 4	1 874	1 887
Order 5	1 852	1 870
Order 6	1 860	1 878
Additive model	3 445	—
Multiplicative	2 210	—

Table IX. Estimations of the adjustment for the fifth order.

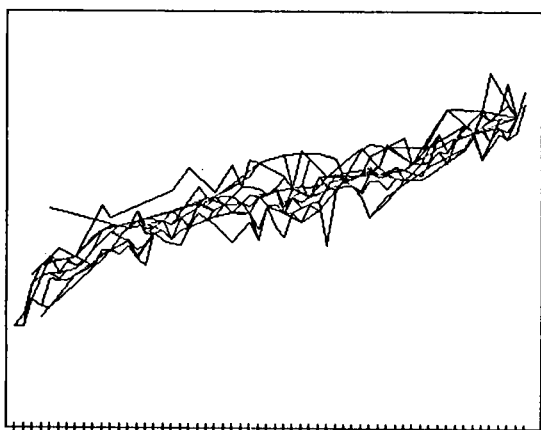
	V10	V13	V9	V4	V11	V1	V6	V3	V7	V8	V2	V5	V14	V12
101				43.2		43.2	43.2	43.2	43.2	42.9	42.9	42.9		
102	22.7	21.5	21.5	21.5	21.5	21.5	21.5	19.1	19.1	19.1				
103	38.7		38.7		38.7	37.7	37.7	37.7	37.7	37.7	37.3	36.7		
105							24.6	24.6	24.6	24.6	24.6	24.6	23.2	
106	47.4		47.4		47.4	41.0	41.0	41.0	41.0	40.9	39.3	39.3		
107	37.3		37.3			37.3	37.3	35.6	35.6	35.6	35.6	27.1		
109					33.1	33.1	33.1	33.1	33.1	33.1	33.1	33.1		
110					52.8	45.7	45.7	45.7	45.7	42.6	42.6	42.6		
208	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	46.3	43.8		
209	40.4	39.2	36.9	36.9	36.9	35.9	35.9	35.9	35.9	35.9				
211		30.3				29.8	29.8	28.8	28.8	28.8	28.8	28.8		
217		29.6					29.6	29.6	29.6	29.6	27.8	27.7	27.6	26.8
219	50.0	50.0	50.0				50.0	50.0	50.0	50.0	50.0		50.0	48.3
221	41.5	41.2		37.6			37.6	37.6	37.6	35.3	35.3	34.8	34.4	34.4
222	38.4						30.1	30.1	29.8	29.8	29.8	29.8	29.8	29.8
223	35.0	32.2	32.2	30.0			30.0	28.0	28.0	28.0	28.0	28.0	27.4	27.4
224		30.8		30.8			30.8	30.8	29.0	29.0	29.0	29.0	27.8	27.7
302					26.5	25.0	25.0	23.1	23.1	22.8	22.1	21.9		
303	39.3		39.3	39.3		39.3	39.3	37.9	37.9	37.9	37.9	37.9		
305	25.2		25.2	25.2	25.2	25.2	24.2	24.2	24.2	23.6	23.6	22.6	22.6	
306							15.2	15.1	15.1	15.1	15.1	15.1	15.1	
307					24.9	24.9	24.5	24.5	23.3	23.3	22.0			
309						17.0	17.0	17.0	15.2	15.2	15.2	15.2		
401	44.0	44.0	44.0	44.0			44.0	44.0	44.0	44.0	44.0	44.0	44.0	43.3
402		45.0					45.0	43.2	43.2	43.2	43.2	43.2	41.7	41.3
404		41.7					41.7	39.9	39.5	39.5	39.5	39.5	38.8	37.7
405		31.1					31.1	31.1	31.1	31.1	31.1	29.2	27.1	27.1
408		42.1		40.3			39.0	39.0	38.3	37.9	37.9	36.6	31.3	31.3
410		32.4					25.1	25.1	25.1	22.9	22.9	22.2	18.0	18.0
415		40.6					37.3	37.3	34.7	33.6	33.6	33.6	32.3	32.3
416		47.1					47.1	46.0	46.0	46.0	46.0	44.7	44.6	43.7
418		36.9					36.9	36.9	36.9	36.9	36.9	36.9	36.9	35.8
420	42.0	41.7	41.7		37.9		37.9	37.9	37.9	37.8	37.4	35.5	34.4	33.6
424		51.1					48.0	45.9	45.9	45.9	45.9	44.0	42.9	42.9
425		39.9	37.7				37.7	37.7	35.5	35.5	35.5	35.5	35.5	35.5
500	34.9	34.3	34.3				34.3	32.3	32.3	32.3	32.3	31.2	31.2	30.1
520	41.1	40.6	36.0	36.0	36.0		35.0	35.0	33.5	32.7	32.7	32.7	30.5	30.5
521	39.3	39.3	39.3	35.4	34.3		34.3	34.3	34.3	34.3	29.4	29.4	27.9	27.9
523	40.4	40.4	37.4	37.0	33.0		33.0	33.0	32.0		32.0	32.0	32.0	28.4
601	47.3			47.3	47.3	47.3	46.6	44.8	44.8	44.8	44.1	40.1	39.9	
603	35.6		35.6	35.6	35.6	35.6	35.6	34.3	34.3	34.3	34.0	32.0		32.0
611				43.0	38.2	38.2	38.2	38.2	36.9	36.9	36.9	36.9		36.9
612			29.9	29.8	29.8	29.8	29.8	28.7	28.3	28.3	28.1	26.9		25.7
640			34.1	34.1	34.1	34.1	32.9	32.9	32.5	30.9	30.9	30.9		30.9
641				31.0	31.0	31.0	30.1	29.5	28.6	28.6	27.4	27.2		24.0
670	24.1		24.1	24.1	24.1	24.1	22.6	22.6	22.6	22.6	22.6	18.2	18.2	16.6
680				34.6	34.6	34.6	34.5	34.5	34.5	33.1	32.9	32.9	32.9	32.9
701	31.1				29.9	29.9	28.8	28.8	28.8	28.8	26.4	26.4		26.4
703	39.7		37.2	35.7	34.4	34.4	32.8	32.8	32.8	32.0	32.0	32.0	32.0	30.4
705				35.8	33.2	33.2	33.2	33.2	33.2	31.8	30.3	30.3		30.3
709	33.0		29.3	29.3	29.3	29.3	27.8	27.8	27.8	27.8	26.4	25.6	25.5	24.9
710				28.3	28.3	26.4	26.4	26.4	26.4	26.4	24.1	24.1		23.7
711				35.8	35.4	35.4	35.4	35.4	35.3	33.1	33.1	33.1		29.8
712	39.8		39.8	35.8	34.2	34.2	34.2	34.2	34.2	33.2	33.2	33.2		33.0
714	38.8	36.1	35.6	34.3	34.3	33.1	33.1	33.1	33.1	31.8	31.8	31.8	31.8	27.4
717					40.7	40.6	39.4	39.4	39.4	37.5	37.5	37.5		36.9
723							32.2	32.2	32.2	30.8	26.3	25.6	25.1	25.1
803						38.7	38.7	38.7	38.7	37.1	37.1	37.1		
804	35.6			35.6	33.4	33.4	33.4	33.4	33.4	33.4	32.0	31.6		
807						43.7	43.7	43.7	43.7	41.0	41.0	41.0		

Table IX shows that the estimates of the adjustment for the fifth order are only averages of consecutive varieties. The incompatibilities between observed yields and the chosen median order are solved by giving equal values to the estimates. With the adjustment of this model, Denis (1982) proposed a test of additivity. However, the approximations he applied are invalid here as his simulations were not for more than 30 varieties and 20 trials. Moreover, the proposed formulas have been established only for complete data sets.

Figures 1 and 2 present our observations and estimates. The appearance of the adjusted data is much clearer. The 14 graphs associated with the varieties never cross each other and many graphs are identical for a certain number of points.



**Fig 1.** Plot of the observed yields by variety and trial versus the ranks of the average yield of the sites. Each variety is represented by a line which joins the points  $(\text{rank}(s), Xs(i))$  concerned. The missing values are interpolated except at the extremities of the line.



**Fig 2.** Plot of the yield estimations of the adjustment for the fifth order versus the ranks of the average yield of the sites.

## DISCUSSION

This approach does not fit the usual statistical framework of parametric models, which leads to some gaps in the interpretation of the results. The result is actually limited to the proposal of an order (or more often a set of orders) supposed to show the collective preference in all the trials. However the algorithm will always give such an order, which can only be qualified through the number of incompatibilities with the majority tournament. An important remark is the fact that the building of the majority tournament represents some reduction in information given by the trials. In order to see this better, let us suppose the extreme situation in which 31 of our trials give an order  $O$  and the others give a similar but different order  $O'$ . For every pair of varieties, the comparison will follow the majority opinion coming from order  $O$ , which is the only median order at distance zero. The 29 minority trials will be ignored, which obviously is not a good representation of the data. Still the advantage of orders (for all varieties or for the best of them) is to provide a direct answer to a question of great importance to research workers. This is why we think that it may be extremely useful if it is used reasonably. Moreover, as said before, it may be used in complex situations, involving numerous missing data and especially polycriterion evaluation, where classical methods are inoperative. The only condition is the building of the majority tournament allowing the easy consideration of many points of view.

What does being reasonable mean? It is not simply trusting blindly the proposed median orders, and corroborating them with complementary information. When the information allowing the elaboration of the majority tournament is of the continuous type, the first thing to do is to supplement it with a least-squares analysis under an order constraint and compare the sum of squared errors (SSE) with that obtained from the classical model (eg, interaction model with one multiplicative term). If it is not lower, we may suppose that the proposed model is an unsatisfactory arrangement. In order to have a more precise idea, the trials can be separated in order to check that none of them bring an abnormal contribution to the SSE. This could be completed by the examination of the residuals. The differences between the median order and the data will then be clear, leading in some cases to the building up of several median orders for distinct groups of trials. It is clear that if the interaction is such that orders

of varieties are inverted in different trials, no median order can provide a full solution, and several orders will have to be considered.

In this view, the search for median orders is more exploratory than confirmatory. This is not a failure, as data analysis has often shown. Our point of view is that it offers one more tool available to research workers.

## REFERENCES

- Baker RJ (1988) Tests for crossover genotype-environmental interactions. *Can J Plant Sci* 68, 405-410
- Barthélemy JP, Guénoche A, Hudry O (1989) Median linear orders: heuristics and a branch-and-bound algorithm. *Eur J Oper Res* 42, 313-325
- Barthélemy JP, Monjardet B (1981) The median procedure in cluster analysis and social choice theory. *Math Soc Sci* 1, 235-267
- Byth DE, Eiseman RL, de Lacy IL (1976) Two-way pattern analysis of a large data set to evaluate genotypic adaptation. *Heredity* 37, 215-30
- Charon I, Hudry O (1992) Utilisation des scores dans les méthodes exactes déterminant des ordres médians de tournois. *Math Inf Sci Hum* 119, 1992
- Dagnelie P (1970) Les comparaisons multiples de moyennes. In: *Théorie et Méthodes Statistiques vol 2*, Presses Agronomiques de Gembloux, Gembloux, 241-253
- Denis JB (1979) Sous-modèles interactifs respectant des contraintes d'ordres. *Biométrie-Praximétrie* 19, 49-58
- Denis JB (1982) Test de l'interaction sous contraintes d'ordre. *Biométrie-Praximétrie* 22, 29-45
- Denis JB (1983) Interaction entre deux facteurs. Thèse de Docteur Ingénieur en Sciences Agronomiques, INA-PG, 222 p
- Denis JB (1992) Analyse statistique de l'interaction entre deux facteurs. Internal document of the Laboratoire de Biometrie, INRA Versailles, 240 p
- Denis JB, Vincourt P (1982) Panorama des méthodes statistiques pour l'étude de l'interaction génotype-milieu. *agronomie* 2, 219-230
- Fishburn PC (1972) *Mathematics of Decision Theory*. Mouton, Paris
- Freeman GH (1973) Statistical methods for the analysis of genotype-environment interactions, *Heredity* 31, 339-354
- Frey JJ, Yehia Alcoutlbi A (1986) Comparaisons par paires, une interprétation et une généralisation de la méthode des scores. *RAIRO* 20, 213-227
- Gauch HG (1990) Full and reduced models for yield trials. *Theor Appl Genet* 80, 153-160
- Hudry O (1989) Recherche d'ordres médians: complexité algorithmique et problèmes combinatoires. Thèse ENST, Paris
- Kemeny JG (1959) Mathematics without numbers. *Deadalus* 88, 577-591
- Reinelt G (1985) *The Linear Ordering Problem: Algorithms and Applications*, Helderman Verlag, Berlin
- Smith AFM, Payne CD (1974) An algorithm for determining Slater's  $i$  and all nearest adjoining orders. *Br J Math Stat Psychol* 27, 49-52
- Young HP (1986) Optimal ranking and choice in paired comparisons. In: *Information Pooling and Group Decision Making* (B Grofman, G Owen, eds) JAI Press, Greenwich, UK, 113-122