Modelling interaction from top-cross design data and prediction of F1 hybrid value

A Charcosset [1], JB Denis [2], M Lefort-Buson [3], A Gallais [1]

[1] *INRA–INAPG–UPS, Station de Génétique Végétale, Ferme du Moulon, F91190 Gif-sur-Yvette;*
[2] *INRA, Laboratoire de Biométrie, Versailles, route de Saint-Cyr, F78026 Versailles;*
[3] *INRA, Département de Génétique et Amélioration des Plantes, 11, rue Jean-Nicot, F75007 Paris*

Summary — In breeding for hybrid varieties, top-cross mating designs are widely used to estimate the combining ability of inbred lines as a measure of their potential as parents of commercial hybrids. A multiplicative model is proposed to take into account interaction effects in addition to general combining ability effect (GCA) estimates to predict the values of potential line x line single cross hybrids. The efficiency of this approach was studied on a set of 58 maize inbred lines representing a wide range of germplasm. These lines were crossed to 4 tester lines and hybrids were evaluated for silage yield. The value of line x line hybrid values were predicted according to several models with and without multiplicative terms. The efficiency of the models was estimated by comparing the predicted and the observed values for a sample of 88 line x line reference hybrids. Predicted values were significantly correlated with observed values for every model. For a given set of testers, models which included multiplicative terms were most efficient. A simulation approach confirmed this conclusion for the best hybrids. Thus, this type of model provides a possibility to increase the efficiency of prediction without requiring additional measurements and should be tested for various germplasms.

hybrid value prediction / interaction model / top-cross design

Résumé — Modélisation des interactions et prédiction de la valeur d'hybrides F1 à partir de plans de croisements *top-cross*. *Dans les programmes de sélection de variétés hybrides, les plans de croisement de type* top-cross *sont largement utilisés pour estimer l'aptitude à la combinaison de lignées, et donc évaluer leur intérêt en tant que parents d'hybrides. Un modèle multiplicatif est proposé dans cette étude pour intégrer les effets d'interaction entre lignées et testeurs dans la prédiction de la valeur des hybrides pouvant être créés à partir des lignées testées. L'efficacité de cette approche a été étudiée pour un ensemble de 58 lignées de maïs représentant une large variabilité génétique. Ces lignées ont été croisées à 4 testeurs et la productivité des hybrides pour l'ensilage a été évaluée. La valeur des hybrides entre lignées a été prédite à partir de différents modèles, incluant ou non des termes multiplicatifs. L'efficacité des modèles a été évaluée en comparant les valeurs prédites et les valeurs observées pour un ensemble de 88 hybrides entre lignées. Les valeurs prédites et observées étaient corrélées de façon significative pour tous les modèles. Pour un ensemble de testeurs donné, les prédicteurs les plus efficaces faisaient intervenir des termes multiplicatifs. Ce résultat a été confirmé au niveau des meilleurs hybrides à l'aide d'une approche par simulation. Ce type de modèle offre ainsi la possibilité d'améliorer l'efficacité de la prédiction sans nécessiter d'observations supplémentaires, et devrait donc être testé pour différents types de matériel génétique.*

prédiction de la valeur hybride / modélisation des interactions / plans top-cross

INTRODUCTION

Since Shull (1908) introduced the basic concepts of hybrid breeding, inbred lines have been developed according to various schemes (see Hallauer, 1990 for a review) in order to provide parents for hybrid varieties of maize and other species. The increase in the number of parental lines available in a given species (or for given conditions) raises the problem of detecting the best F1 hybrids, since the test of all possible combinations ($N(N-1)/2$ if N is the number of lines) rapidly becomes unfeasible. Thus, the choice of tested combinations has a dramatic impact on the success of a hybrid breeding program.

Jenkins and Brunson (1932) suggested the use of top-cross tests (*ie* crosses between inbreds and a variety population) to achieve preliminary tests of new lines. The efficiency of this approach was illustrated by Sprague and Tatum (1942). These authors introduced the partitioning of hybrid values in terms of general and specific combining ability (GCA and SCA, respectively). They defined GCA of line *i* as the average performance of this line in hybrid combinations. SCA of the cross between line *i* and line *j* was defined as the deviation from the performance expected on the basis of the GCA of the lines (*ie* the interaction effect of the cross). Results illustrated that top-cross tests were efficient in selecting for GCA. Sprague and Tatum (1942) also suggested the simultaneous use of several testers to obtain GCA estimates not influenced by SCA effects with a given tester.

Once efficient screening of lines for general combining ability has reduced the number of hybrid potential parents, selection for SCA can be achieved by single-cross combination tests (Lonnquist and Rumbaugh, 1958). Most often, new lines are crossed to several classical commercial hybrid parents (elite lines), in order to find suitable hybrid combinations.

In a previous study, we illustrated that such devices could also be used to gauge with some accuracy the performance of untested line x line hybrids (Charcosset *et al*, 1990). From the basic hypothesis that line x tester SCA effects could provide information about line x line SCA, we established a significant relationship between F1 hybrid performance and an index of the distance between the parents computed from line x tester SCA estimates (this distance was derived from Hanson and Casas, 1968). However, theoretical limitations in the use of this method were pointed out: high distances are always associated with high heterotic values, but low distances may also be associated with high heterotic values. Discarding hybrids on the basis of low distances between parents may, therefore, be a hazardous practice. Thus, the aim of this study was to present a more attractive way to use line x tester SCA estimates to predict the SCA of line x line hybrids. In this preliminary study we explore a wide range of maize hybrids, including intragroup (flint, dent), and inter-group hybrids.

MATERIALS AND METHODS

Experimental design

Parental lines

We considered a set of 58 maize inbred lines representative of the lines which can be used in northern France. This set included typical European flint lines, such as lines F7 and Ep1, some of their derivatives and derivatives of the line F2, lines from Northern American origin (W401, CM7, A641 etc), mixed types resulting from crosses between lines of previous origin (*eg* Co255), and lines derived from exotic germplasm (*eg* F285, which results from a second backcross generation between the line F7 and the Argentine population F64). The *i*th line among the 58 will be further referred to as L_i.

Test crosses

All lines were crossed to 4 elite tester lines: the widely-used European flint F2 (further referred to as T_2), a European flint derived from line F7: F283 (T_1), an early dent line: F252 (T_3) (derived from a cross between Co125 and F186) and a widely-used line from the io-dent group (T_4). Testers were used as female parents to obtain homogeneity in seed size. Some of the parental lines were related to testers T_1, T_2 or T_3; the average relationship coefficients (Malecot, 1969) between the 58 lines and these testers were 0.079, 0.052 and 0.054, respectively and never exceeded 0.50. Test-crosses (L_i x T_k), as well as tester x tester crosses and tester *per se* performances, were used to construct different predictors (described below) of the hybrid value associated with crosses between pairs of parental lines (L_i x L_j).

Estimation of the efficiency of predictors using reference hybrids

The 58 parental lines were used to generate a set of 88 hybrids (each line was used as a parent for ≈ 3 hybrids). This set was used to test the efficiency of predictors. The average relationship coefficient between hybrid parents was 0.04; here again, no relationship coefficient exceeded 0.5.

Yield trials

The experiment was conducted in 1988 at Gif-sur-Yvette (35 km south of Paris). Since a single location was used for this study, results should be considered as a support to present a method, and the actual efficiency of the approach should be re-examined in a multi-environmental study. Due to differences in maturity between testers, the test-crosses were evaluated in 4 different trials, each corresponding to a given tester. Reference hybrids were also evaluated in a separate trial. The 5 trials took place next to each other in a homogeneous experimental field and were sown on the same day. Each trial was planted in a randomized complete block design with 3 replications. Each plot consisted of one row 5-m long and with 0.8 m between rows, planted at a density of 90 000 pl/ha. Total green matter yield (silage fresh yield) was evaluated for each plot, when plant moisture averaged 65 to 70% for a given trial. Plant moisture was estimated for each plot, using a sample of $\approx$ 0.58 kg, dried at 35°C for 18 h and then 105°C for 12 h. Total dry matter (DM) yield was then calculated.

Variance analysis of the trials

Variance analyses were performed for each trial to test the significance of genotype effects. These effects were considered as fixed, and the variation among genotypes

$$\frac{\sum_{i=1}^{ng} g_i^2}{ng - 1}$$

where g_i is the centered genetic effect of genotype i, ng is the total number of genotypes) was evaluated, using the expectation of genotype mean square (see Scheffé, 1959). Moreover, an analysis of variance involving all 4 testers was performed, in order to investigate line x tester interaction effects using the classical model for factorial designs:

$$Y_{it} = \mu + \alpha_i + \beta_t + \varnothing_{it} \qquad [1]$$

where Y_{it} is the value of the cross between line i and tester t, α_i and β_t are the main effects of line i and tester t, respectively, and $\varnothing_{it}$ is the interaction effect. Although both trial and tester effects were confounded, preliminary investigations were carried out using check plots (hybrid cultivars Dea, Brûlouis and Mona). These analyses showed significant differences in trial effects. A negative relationship (over the trials) was observed between the mean of the checks and the mean of the test-crosses. That is to say that the average value of the checks in a given trial was high when the average value of the test-crosses in this trial was low. This could be explained by competition effects. However, when considering the checks, the interaction effects between cultivars and trials were not significant. Thus we may expect that, when considering parental lines in test-crosses, the interaction between lines on one hand and testers or trials on the other hand are mostly due to line x tester interactions. Henceforth we will call the interaction effect line x tester interaction, for reasons of simplicity, although strictly speaking this also involves line x trial interaction.

Computation of predicted values

The 4 tester trials were used to generate 15 data sets: 4 single tester data sets ($[T_1]$, $[T_2]$, $[T_3]$, $[T_4]$), 6 tester pairs ($[T_1, T_2]$, $[T_1, T_3]$, $[T_1, T_4]$, $[T_2, T_3]$, $[T_2, T_4]$, $[T_3, T_4]$), 4 tester triplets ($[T_1, T_2, T_3]$, $[T_1, T_2, T_4]$, $[T_1, T_3, T_4]$, $[T_2, T_3, T_4]$) and the 4 testers ($[T_1, T_2, T_3, T_4]$). For each data set, predicted values were computed using the classical linear model based on the estimation of the GCA of the lines. For data sets including at least 2 testers, a bilinear model of interaction effects was added to the model to take into account the effect of line x tester interaction (SCA).

Prediction using linear models

If the testers included in a given data set are 'representative' of the tested lines, the main effect of lines (model [1]) can be considered as an estimate of their general combining ability (see proposal of Jenkins and Brunson, 1932, in the introduction section). Thus, the predicted value of the hybrid between lines i and j ($\tilde{Y}_{ij}$) can be computed as:

$$\tilde{Y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\alpha}_j \qquad [2]$$

where $\hat{\mu}$ and $\hat{\alpha}$ are the parameters estimated for model [1]. If Y_{ij} is the observed value of the hybrid between lines i and j, $\tilde{E}_{ij} = Y_{ij} - \tilde{Y}_{ij}$ will subsequently be considered as an estimate of cross $i \cdot j$ SCA.

A similar approach can be used when a single tester *(T)* is considered. In this situation, $\hat{\alpha}_i$ will be the estimated genetic effect of the cross between line i and tester t. A total of 15 such predictors were computed (one for each data set).

Bilinear models of interaction effects

Bilinear models of interaction effects were described by Mandel (1971), and used to analyse genotype x environment interactions (see Freeman, 1973; Crossa *et al*, 1990). The basis of these models is to partition the interaction between 2 factors (for instance $\varnothing_{it}$ in model [1]) as a sum of multiplicative terms involving parameters specific to each of the interacting factors:

$$\varnothing_{it} = \sum_{r=1}^{R} \gamma_i^r \cdot \lambda_t^r \qquad [3]$$

where R is the number of multiplicative terms, γ_i^r and λ_t^r are the parameters specific to levels i and t of the 2

interacting factors, for the *r*th multiplicative term. This model shows analogies with principal component analysis (Perkins, 1972) and can be used to provide graphical representations of the levels of the interacting factors (*eg* the γ^1 versus γ^2 plots for lines). This model can be adapted to the particular case of diallel designs to take into account the dependency between the 2 interacting factors. If one assumes no reciprocal effects, interaction (SCA) between lines *i* and *j* will be written:

$$\varnothing_{ij} = \sum_{r=1}^{R} \gamma_i^r \cdot \gamma_j^r \qquad [4]$$

γ parameters can be estimated from a test-cross design if at least 2 testers are considered, and tester x tester crosses (including tester *per se* values) are included in the analysis. Otherwise γ_i^r and λ_t^r cannot be compared and the symmetry is not respected. If the preceding conditions are respected, the data set used for prediction can be represented as in table I (in a situation with 2 testers) and can be considered as a diallel design with missing data. As in the case of the linear model, the first step in prediction will be to use available data (Y_{it}) to estimate parameters for the complete model:

$$Y_{it} = \mu + \alpha_i + \alpha_t + \sum_{r=1}^{R} \gamma_i^r \cdot \gamma_t^r \qquad [5]$$

And then to use the estimated parameters ($\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\gamma}_i^r$) to compute predicted values ($\tilde{Y}_{ij}$) for missing data:

$$\tilde{Y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\alpha}_j + \sum_{r=1}^{R} \hat{\gamma}_i^r \cdot \hat{\gamma}_j^r \qquad [6]$$

Model [5] requires the estimation of $(R + 1)(L + T - R/2)$ independent parameters, so up to $(T - 1)$ multiplicative terms can be introduced in the model for a given design (including *T* testers). If $(T - 1)$ multiplicative terms are introduced in the model, the number of independent parameters to be estimated is equal to the number of observed data and the parameter estima-

tions can be analytically written as functions of the data. In this situation, $\tilde{Y}_{it} = Y_{it}$.

If fewer than $(T - 1)$ multiplicative terms are introduced, estimation requires the use of an iterative alternating least squares procedure (see Denis, 1991). A special software package (INTERA) was written by Decoux and Denis (1991) and used in the present study. Problems in parameter estimations (aberrant estimates) were observed when using all the data of a given data set, because of the number of data involved and their structure. Thus, local estimates were computed: to predict Y_{ij} value (compute $\tilde{Y}_{ij}$), only tester x tester crosses (and tester *per se* values), *i* x testers and *j* x testers were used for parameter estimation. The estimations were made successively for each hybrid. The required number of iterations was very different according to the hybrid considered.

Seventeen new models of this kind were defined according to the subset of testers and the number *(R)* of multiplicative terms considered. The 32 possible models are further described using [*R;T*] notation (see table IV), *R* being the number of multiplicative terms introduced in the model (0 for the linear model) and *T* that of the tester considered.

Evaluation of the efficiency of the predictors

Predicted/observed plots and empirical correlation

The first approach for comparing the different models was to plot, for each of the 32 models, the predicted values of the 88 reference hybrids against the observed values. Note that predicted values do not depend on the observed values of the reference hybrids. We also computed the resulting (empirical) correlation coefficient.

Calculation of the prediction variances and "heuristic" correlation

In order to go further in the comparison of the different models, the variances of the predictions they give were calculated. When there was no multiplicative term $(R = 0)$, the model [5] is linear and the variance is known.

For the other models, the variances were calculated numerically using an asymptotic linear approximation (see *Appendix*). The experimental variation of testers x testers was not taken into account, since this information was used to compute the predicted value of every line x line potential hybrid. Thus, we shall say that the variances are calculated conditionally to the observation of the tester x tester crosses. Note however that the experimental variance in the estimation of the tester x tester crosses should affect the efficiency of the prediction when using multiplicative models.

Table I. Seven data values (●) are used to predict the value of the hybrid between line 2 (L_2) and line 5 (L_5) with models involving testers T_1 and T_3. The table was completed (O) according to the assumption of symmetry. There are 3 missing values (?).

	T_1	T_3	L_2	L_5
T_1	●	O	O	O
T_3	●	●	O	O
L_2	●	●	?	?
L_5	●	●	?	?

The estimates of the prediction variances were used to represent the 88 hybrids by confidence ellipses (at an α% risk level) in the plot of predicted vs observed values. The center of each ellipse corresponds to the predicted and observed values. Also, they were used to compute a 'heuristic' correlation coefficient; this coefficient is an estimation of the correlation which would be obtained, if there was no random variation (no error) disturbing the estimations of the observed values and initial data. This heuristic correlation shows analogies with the genotypic correlation (for the situation where genotypes are treated as random). If it is assumed that predicted and observed values are centered, then:

$$\rho_H = \frac{\Sigma_{i=1}^{N} P_i O_i}{\sqrt{\Sigma_{i=1}^{N} P_i^2 - N\hat{\sigma}_p^2} \sqrt{\Sigma_{i=1}^{N} O_i^2 - N\hat{\sigma}_o^2}} \quad [7]$$

where N is the total number of reference hybrids (88 in our experiment), P_i is the predicted value of the ith reference hybrid, O_i is its observed value, $\hat{\sigma}_p^2$ the estimation of the error variance of the predicted values (computed as the mean of the individual prediction variances), and $\hat{\sigma}_o^2$ the estimation of the error variance of the observed values. This formula is justified by the fact that the expectations of $\Sigma_{i=1}^{N} P_i O_i$, $\Sigma_{i=1}^{N} P_i^2 - N\hat{\sigma}_p^2$ and $\Sigma_{i=1}^{N} O_i^2 - N\hat{\sigma}_o^2$ are the sum of products and sums of squares that would be obtained if there was no random variation.

Simulation studies

Furthermore, to compare the efficiencies of the different models to detect the best F1 hybrid combinations, a simulation study of 500 experiments was conducted. For each experiment, random error terms were generated under the assumption of a normal distribution (with variance $\hat{\sigma}_o^2$) and added to the observed values to generate deviated observed values. Similar deviations were also applied to predicted values using the adequate variances. For each experiment the n hybrids ($n = 4, 8, 16, 24$), which had the highest deviated predicted values were determined. Their average deviated observed value was computed. For each model

and each selection intensity, these values were averaged over the 500 experiments.

RESULTS

Variance analysis of the trials

In all trials, the variation among hybrids was highly significant (table II). The variability among hybrids appeared greater for the T_1 (F283) and T_2 (F2) testers than for the other 2 testers (T_3 (F252) and T_4).

The values of the F-statistic for the effects of lines, testers and line x tester (more precisely line x (tester and trial); see previous section) interactions were highly significant (table III). The occurrence of interaction effects in this type of design was consistent with results reported for hybrid seed yield in maize (see Hallauer and Miranda, 1981, for a review).

Predicted/observed plots and empirical correlation

The observation of the predicted/observed plots has shown that the different models have very different efficiencies. Illustrations are given in figure 1 for models ([0; T_2], [0; T_4], [0; T_2, T_3], [2; $T_1 T_2 T_3$], [0; T_1, T_2, T_3, T_4] and [3; T_1, T_2, T_3, T_4]). For instance, the relationship appears much better for model [2; T_1, T_2, T_3] than for model [0; T_4]. Notice that for all the plots the predicted value of the hybrids is generally superior to the observed value. This is due to the fact that testers are elite lines, with a higher GCA than the tested germplasm. This is especially true for tester T_4.

Table II. Characterization of individual test-cross and reference hybrid trials

Trial	Mean yield * (t·ha⁻¹)	No of hybrids	F-value	Residual variance	Estimated variation among hybrids
T_1 hybrids	17.5 ± 0.24	58	11.3**	2.7	9.57
T_2 hybrids	17.1 ± 0.33	58	6.2**	4.9	8.49
T_3 hybrids	16.9 ± 0.24	58	5.3**	2.7	3.9
T_4 hybrids	19.1 ± 0.35	58	3.0**	5.5	3.5
Ref hybrids	16.5 ± 0.24	88	5.8**	4.0	6.5

* Confidence interval at the 0.05 probability level. ** Effect significant at the 0.01 probability level.

Table III. Two-way analysis of variance of the test-cross values.

Effect	df	F-*value*
Line	57	42.7**
Tester (trial)	3	14.3**
Line x tester (trial)	171	3.4**
Error	$\sigma_e^2 = 3.5$	

df: Degrees of freedom ; ** effect significant at the 0.01 probability level. σ_e^2: residual variance estimation.

The empirical correlation coefficients between predicted and observed values depended on the testers considered (table IV). When considering only predictions obtained from single tester information, prediction appeared less efficient for T_4 than for other testers. This might be due to the fact that this tester was not related to any tested line (see theory by Hull, 1945; and Rawlings and Thompson, 1962).

When considering the models which took into account 2 testers and no multiplicative terms, the best result was obtained for the model [0; T_2, T_3] (F2 and F252). When 3 testers were considered, the best result was obtained for combination [0; T_1, T_2, T_3]. Results obtained for models [0; T_2, T_3] and [0; T_1, T_2, T_3] appeared rather similar, and superior to the model [0; T_1, T_2, T_3, T_4]. Although the results of these comparisons were not established statistically (because data implied in the estimation of ρ^2 for different models were not independent), the best predictor involving several testers was clearly superior to the best single tester predictor. This result is consistent with the suggestion of Sprague and Tatum (1942) (see *Introduction*). Even when several testers are taken into account, the choice of testers has an important incidence on prediction efficiency. Introducing tester T_4 information in prediction computation seemed to lead to a decrease of the relationship between predicted and observed values, when compared to predictions not including this tester's information. The best relationships were obtained when testers T_2 and T_3 were taken into account simultaneously, and when T_4 was not included in prediction computation.

For all tester combinations, the estimated value of the correlation increased with the increase of the number of multiplicative terms included in

Table IV. Empirical coefficient of determination (ρ_E^2) and 'heuristic' coefficient of determination (ρ_H^2) between predicted and observed values of the 88 reference hybrids; error variances (σ_e^2) of the predicted values. Each model is described by its number of multiplicative terms and the testers considered (described as T_1, T_2, T_3 and T_4) ; eg [1; T_1, T_2, T_4] designs the model with one multiplicative term and testers T_1, T_2 and T_4.

Model	ρ_E^2	ρ_H^2	σ_e^2
1 [0; T_1]	0.267	0.391	2.66
2 [0; T_2]	0.362	0.507	2.66
3 [0; T_3]	0,240	0.407	2.66
4 [0; T_4]	0.093	0.150	2.66
5 [0; T_1, T_2]	0.415	0.558	1.33
6 [1; T_1, T_2]	0.434	0.587	1.49
7 [0; T_1, T_3]	0.370	0.531	1.33
8 [1; T_1, T_3]	0.373	0.555	1.49
9 [0; T_1, T_4]	0.248	0.352	1.33
10 [1; T_1, T_4]	0.261	0.373	1.42
11 [0; T_2, T_3]	0.491	0.700	1.33
12 [1; T_2, T_3]	0.524	0.766	1.56
13 [0; T_2, T_4]	0.291	0.397	1.33
14 [1; T_2, T_4]	0.328	0.449	1.42
15 [0; T_3, T_4]	0.253	0.392	1.33
16 [1; T_3, T_4]	0.284	0.449	1.54
17 [0; T_1, T_2, T_3]	0.494	0.664	0.89
18 [1; T_1, T_2, T_3]	0.536	0.759	1.17
19 [2; T_1, T_2, T_3]	0.540	0.767	1.30
20 [0; T_1, T_2, T_4]	0.359	0.347	0.89
21 [1; T_1, T_2, T_4]	0.383	0.511	0.98
22 [2; T_1, T_2, T_4]	0.398	0.537	1.13
23 [0; T_1, T_3, T_4]	0.343	0.479	0.89
24 [1; T_1, T_3, T_4]	0.365	0.534	1.21
25 [2; T_1, T_3, T_4]	0.379	0.555	1.18
26 [0; T_2, T_3, T_4]	0.402	0.552	0.89
27 [1; T_2, T_3, T_4]	0.490	0.694	1.10
28 [2; T_2, T_3, T_4]	0.507	0.727	1.20
29 [0; T_1, T_2, T_3, T_4]	0.433	0.573	0.67
30 [1; T_1, T_2, T_3, T_4]	0.508	0.707	0.96
31 [2; T_1, T_2, T_3, T_4]	0.521	0.740	1.06
32 [3; T_1, T_2, T_3, T_4]	0.533	0.757	1.14

the model. This was not expected *a priori* since predicted and observed values were obtained independently (contrary to classical multiple regression analyses). The question of the test of the significance of the increase of the correlation raises methodological problems associated with the specificity of this approach which cannot be assimilated to a multiple regression study (in which observed values are used to determine regression coefficients). However, the fact that in all cases increases were observed (which was

not expected *a priori*) should be considered as an initial indication of the efficiency of this method. The largest increases were obtained for combination [T_2, T_3, T_4] (from ρ_E^2 = 40.2% for [0; T_2, T_3, T_4] to ρ_E^2 = 50.7% for [2; T_2, T_3, T_4]) and for combination [T_1, T_2, T_3, T_4] (from ρ_E^2 = 43.2% for [0; T_1, T_2, T_3, T_4] to ρ_E^2 = 53.3% for [3; T_1, T_2, T_3, T_4]). The best relationship was obtained for [2; T_1, T_2, T_3] (ρ_E^2 = 54.0%), which was superior to [0; T_1, T_2, T_3] (ρ_E^2 = 49.4%). Thus, the largest increases in prediction efficiency were pointed out when flint and dent testers were considered simultaneously. This is related to the fact these testers combinations generated the most important interaction effects (results not presented).

Note that large deviations between observed values and their corresponding predicted values by model [0; T_1, T_2, T_3, T_4] were obtained for hybrids which were overestimated by the prediction (considering centered predicted and observed values). Predictor [3; T_1, T_2, T_3, T_4] led to smaller deviations for 3 hybrids whereas the values of 3 other hybrids were still overestimated: these hybrids had the same predicted values with [3, T_1, T_2, T_3, T_4] and with [0; T_1, T_2, T_3, T_4].

Error variances and 'heuristic' correlations

For models without multiplicative terms, the error variance is constant as illustrated in figure 2 by the constant width of the confidence ellipses. This is no longer true for models including multiplicative terms (see in figure 2 the differences in the width of the confidence ellipses). For a given combination of testers, the variance of the predicted values increased with the increase in the number of multiplicative terms; this is a general property associated with the introduction of additional terms in models.

The results obtained when comparing ρ_H^2 of the different models were similar to those obtained when comparing ρ_E^2. Even when no multiplicative terms were considered, the best predictors involving several testers were superior to the best single tester predictors. This illustrates the fact that the superiority observed for empirical correlations was not due solely to a greater number of replications but also had a genetic basis. The best results were obtained for predictors [2; T_1, T_2, T_3], [1; T_2, T_3], and [3; T_1, T_2, T_3, T_4].

When considering the 3 testers [T_1, T_2, T_3], the predictor [2; T_1, T_2, T_3] led to better results than the predictor [0; T_1, T_2, T_3] (ρ_H^2 = 76.7% and ρ_H^2 = 66.4%, respectively). Thus, the introduction of multiplicative terms in prediction computation could lead to an important improvement if testcross values were "perfectly known" (no experimental error). However, models including multiplicative terms lead to higher error variances. This is consistent with the differences between [2; T_1, T_2, T_3] and [0; T_1, T_2, T_3] which are greater for the heuristical correlation coefficient than for the empirical coefficient.

Simulation studies

Results of the simulation studies are given in table V for 4 selection intensities (4, 8, 16 and 24 hybrids selected out of the 88). In all cases, except model [0; T_4], the mean of selected hybrid was much larger than the general mean of the 88 hybrids (16.5, see table II). The models which lead to the higher values of selected hybrids were [1; T_2, T_3] and [2; T_1, T_2, T_3], regardless of the selection intensity; this was consistent with the correlation studies.

When comparing single tester predictors, T_3 and T_1 were more efficient than T_2, itself more efficient than T_4. So, the ranking of models on the basis of prediction efficiencies was not the same as the ranking obtained in the correlation studies: T_4 was still the least efficient predictor but T_2 was superior to T_3 and T_1. When no multiplicative terms were taken into account, the model [0; T_2, T_3] was the most efficient of models including 2 testers, whereas rather poor results were obtained for [T_2, T_4] (at least for high selection intensities); the best triplet was [T_1, T_2, T_3]. For a given number of testers, the best models were similar to the best ones based on correlation criteria.

If we consider the highest selection intensity, and a given combination of testers, the predictor which includes the maximum of multiplicative terms appeared superior to the predictor with no multiplicative terms for 9 combinations out of 11. Thus, according to this criterion, the introduction of multiplicative terms in prediction computation seemed to improve the efficiency of selection.
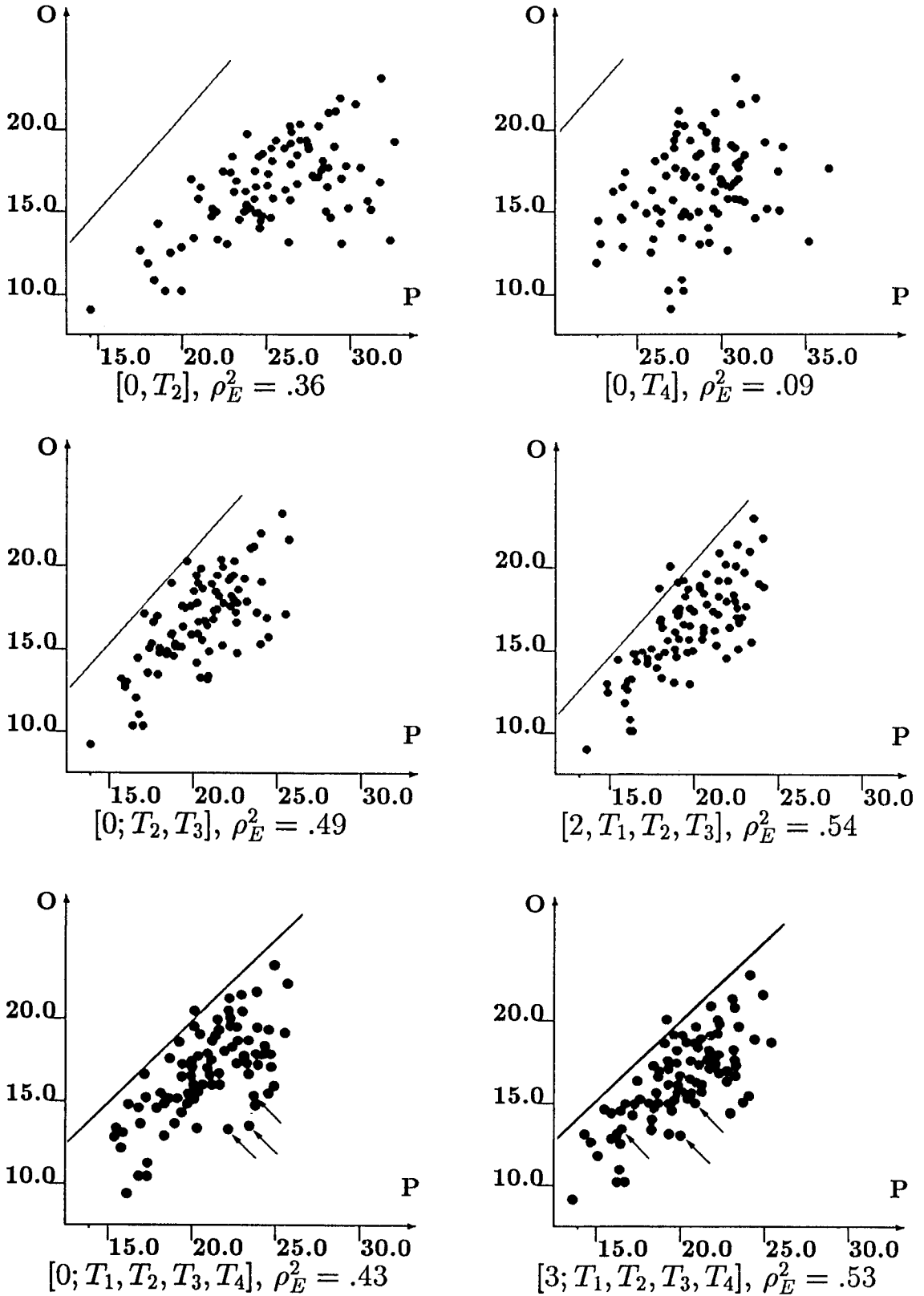
Fig 1. Plots of observed silage yield (t·ha^{-1}, **O**) *vs* predicted yield (**P**) for 6 models (see table IV for models notation). ρ_E^2 is the empirical coefficient of determination. The line is the **P** = **O** axis. For models [O; T_1, T_2, T_3, T_4] and (3; T_1, T_2, T_3, T_4] arrows indicate hybrids for which predicted values show the most important differences.
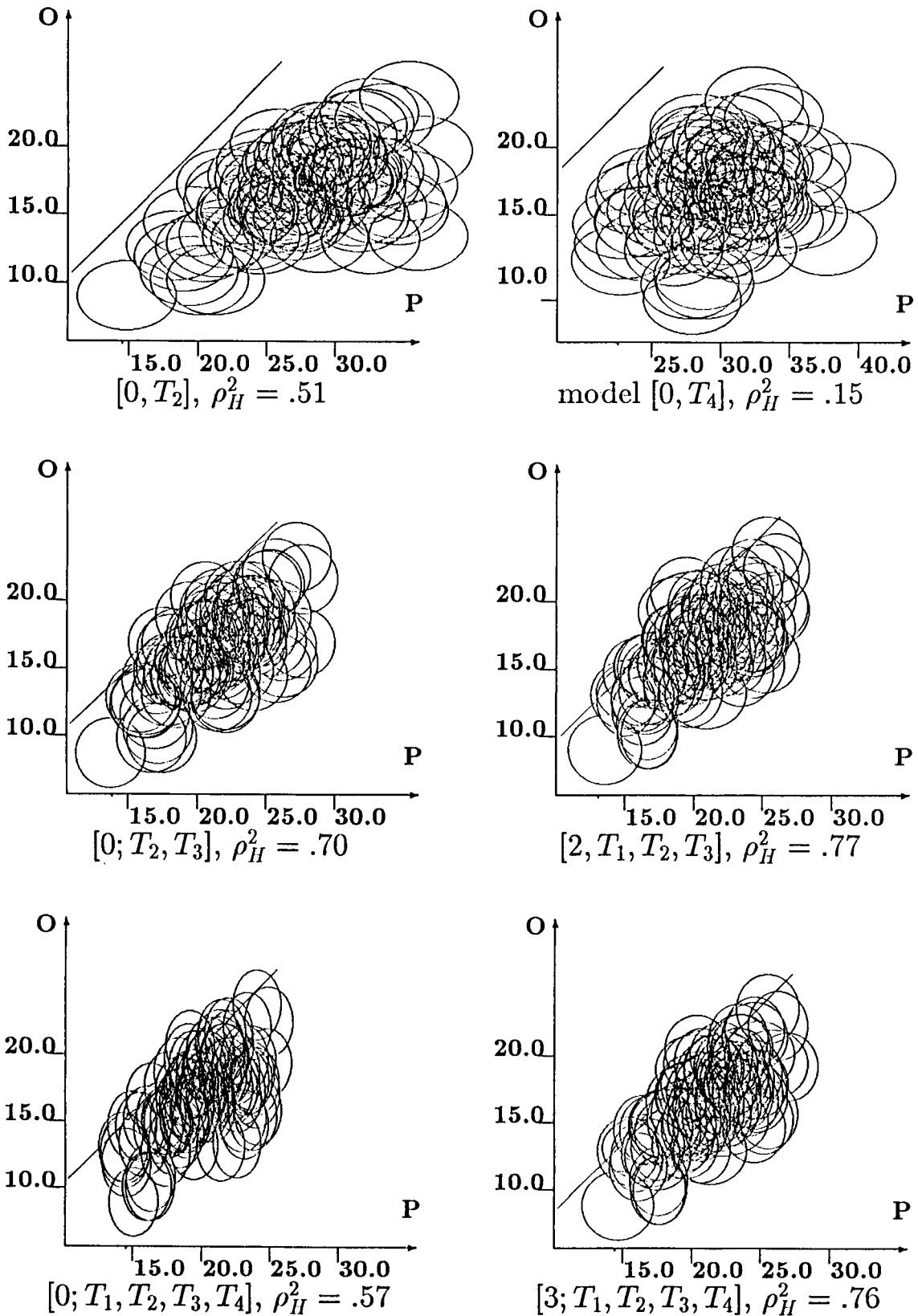
$[0, T_2], \rho_H^2 = .51$

model $[0, T_4], \rho_H^2 = .15$

$[0; T_2, T_3], \rho_H^2 = .70$

$[2, T_1, T_2, T_3], \rho_H^2 = .77$

$[0; T_1, T_2, T_3, T_4], \rho_H^2 = .57$

$[3; T_1, T_2, T_3, T_4], \rho_H^2 = .76$

Fig 2. Plots of observed silage yield (t·ha^{-1}, **O**) vs predicted yield (**P**) for 6 models (see table IV for models notation). The line is the **P** = **O** axis, ρ_H^2 is the 'heuristic' coefficient of determination. Each hybrid is represented by a 0.10 probability level confidence ellipse.

Tableau V. Simulation studies: mean production ($t \cdot ha^{-1}$) of the (Nselec) selected hybrids for each model. Each value is the mean of results obtained from 500 simulation experiments. See table IV for model description.

Model		Nselec			
		4	8	16	24
1	$[0; T_1]$	19.05	18.55	18.12	17.87
2	$[0; T_2]$	17.53	17.59	17.71	17.79
3	$[0; T_3]$	18.97	18.91	18.64	18.34
4	$[0; T_4]$	16.54	16.96	17.29	17.33
5	$[0; T_1, T_2]$	18.28	17.98	17.79	17.83
6	$[1; T_1, T_2]$	18.31	18.05	17.89	17.87
7	$[0; T_1, T_3]$	19.17	18.83	18.53	18.35
8	$[1; T_1, T_3]$	19.09	18.81	18.53	18.39
9	$[0; T_1, T_4]$	18.32	17.83	17.54	17.58
10	$[1; T_1, T_4]$	18.51	17.97	17.58	17.61
11	$[0; T_2, T_3]$	19.28	18.91	18.66	18.43
12	$[1; T_2, T_3]$	19.58	19.17	18.78	18.51
13	$[0; T_2, T_4]$	16.65	17.37	17.71	17.74
14	$[1; T_2, T_4]$	16.66	17.38	17.73	17.77
15	$[0; T_3, T_4]$	18.82	18.57	18.26	17.99
16	$[1; T_3, T_4]$	18.95	18.73	18.44	18.18
17	$[0; T_1, T_2, T_3]$	19.16	18.85	18.53	18.34
18	$[1; T_1, T_2, T_3]$	19.19	18.97	18.69	18.53
19	$[2; T_1, T_2, T_3]$	19.43	19.11	18.71	18.49
20	$[0; T_1, T_2, T_4]$	17.52	17.57	17.58	17.77
21	$[1; T_1, T_2, T_4]$	17.48	17.56	17.60	17.77
22	$[2; T_1, T_2, T_4]$	17.60	17.56	17.61	17.81
23	$[0; T_1, T_3, T_4]$	19.03	18.53	18.23	18.13
24	$[1; T_1, T_3, T_4]$	18.98	18.56	18.33	18.26
25	$[2; T_1, T_3, T_4]$	18.99	18.63	18.41	18.33
26	$[0; T_2, T_3, T_4]$	18.85	18.51	18.15	18.06
27	$[1; T_2, T_3, T_4]$	18.97	18.70	18.43	18.32
28	$[2; T_2, T_3, T_4]$	19.21	18.88	18.57	18.39
29	$[0; T_1, T_2, T_3, T_4]$	18.85	18.49	18.15	18.06
30	$[1; T_1, T_2, T_3, T_4]$	18.91	18.60	18.37	18.33
31	$[2; T_1, T_2, T_3, T_4]$	18.81	18.66	18.46	18.40
32	$[3; T_1, T_2, T_3, T_4]$	19.16	18.83	18.56	18.43

DISCUSSION

Choice of testers

Although a single location was considered, and results on the choice of testers should be reconsidered in a multi-environmental study, our results were consistent with published studies (see Hallauer and Miranda, 1981) which reported that the choice of the testers has a strong influence on selection efficiency. Theoretical considera-

tions about the choice of testers have been provided by Hull (1945) and Rawlings and Thompson (1962) who showed that the best tester should carry the greatest amount of recessive alleles segregating in the tested germplasm. Empirical studies did not really confirm the usefulness of the proposed criteria (Hallauer and Miranda, 1981). However, the fact that in our study T_4 was the least efficient tester might be related to the fact that it was not related to any tested lines (contrary to other testers).

The simultaneous use of several testers appeared to increase prediction efficiency when compared to best single tester predictors, even when using models based only on GCA estimation. This was consistent with the proposal of Sprague and Tatum (1942) that the use of several testers should provide greater security when estimating tested germplasm GCA. In our study, the best tester combinations included lines which were related to the tested germplasm and which are 'complementary' (from different origins).

The use of line x tester specific effects through multiplicative models

The use of several testers also makes it possible to estimate line x tester specific effects (SCA). A previous experiment had shown that these effects could be taken into account to predict the values of line x line hybrids (Charcosset *et al*, 1990), in addition to GCA estimates. The method which was proposed was to use SCA effects to compute a distance index similar to Hanson and Casas' distance (Hanson and Casas, 1968). This distance was correlated significatively to line x line SCA. However, it was discussed theoretically that when 2 testers are considered, this relationship is 'triangular': the highest distances correspond to the highest SCA but the lowest distances may correspond to any SCA value (high, intermediate or low). Thus the use of this distance in linear combination with GCA estimates may lead to an underestimation of the value of some hybrids.

A theoretical study (Charcosset, 1990) of this problem showed that multiplicative models were a more convenient way to take line x tester specific effects into account. It was demonstrated that the highest predicted SCA corresponded to the highest actual SCA, the lowest predicted SCA corresponded to the lowest actual SCA and the intermediate predicted SCA (zero) corre-

sponded to any actual SCA value (high, intermediate or low). Thus, all major changes in predicted values associated with the use of a multiplicative model in addition to GCA estimates should lead to an increase in prediction accuracy.

Even though results should be reconsidered in a multi-environmental study, the present study is consistent with preceding theoretical conclusions: according to all criteria, the use of multiplicative models increased prediction efficiency. Contrary to multiple regression studies, no systematic increase in the correlation was expected with the introduction of multiplicative terms in the prediction formula, since observed values are not taken into account in the computation of the coefficients associated with multiplicative terms (these coefficients are computed from test-cross data). Moreover, the study of models [0; T_1, T_2, T_3, T_4] and [3; T_1, T_2, T_3, T_4] illustrated that most important changes in predicted values due to the introduction of the multiplicative model lead to a better concordance with observed values. Thus, even if the question of a statistical test remains open, this study should be considered as an indication of the efficiency of multiplicative models of interaction to enhance prediction efficiency without requiring additional measurements.

The usefulness of multiplicative models should depend on several factors, including the germplasm of interest (see further) and the precision of the experimentation. The introduction of multiplicative terms in prediction increases the error variances of predictors, depending on the estimated error variance in the initial data. Thus, the difference between the efficiencies of multiplicative and non-multiplicative predictors depends on the precision of the initial data. In this study, using 3 replications for a trait of relatively low heritability, this precision appeared good enough to underline the efficiency (estimated by empirical correlations) of the approach.

Incidence of the tested germplasm on prediction efficiency

The germplasm considered for this study can be considered as rather wide, and reference hybrids included both intra-group and inter-group hybrids. In this situation, it was estimated that if there was no error variation disturbing the estimation of yield values, the most efficient predictor would explain up to 77% of the variation among F1 hybrids when using only 3 tester lines. For this data set, multiplicative models appeared to increase prediction efficiency from 66 to 77%. Increases in prediction efficiency associated with multiplicative terms should be dependent on the germplasm of interest and the testers which are considered. First, possible increases in prediction efficiency will depend on the relative magnitude of GCA and SCA effects. They cannot be important if SCA effects are small when compared to GCA effects. Secondly, it can be demonstrated theoretically (Charcosset, 1990) that multiplicative models are very efficient for predicting the value of hybrids for which one parent line at least is highly related to a tester line, and are not efficient for original germplasm (the 2 parent lines being poorly related to any tester line). Thus, multiplicative models of interaction should be extremely efficient for sets of lines which are structured in homogeneous groups, each group being represented by a tester line.

Further experimental studies should be carried out with various sets of lines to define in which conditions multiplicative models will be efficient. Two kinds of studies can be considered. If heterotic groups are not clearly established for the germplasm of interest, testers should be chosen to give the best representation of the tested germplasm, and prediction can be achieved as presented in this study. If heterotic groups are clearly established for the germplasm of interest, the lines and the testers belonging to one group could be crossed to the testers belonging to the complementary group, and *vice versa*. This is classically done in many breeding programs, the testers being elite lines which represent the diversity within the heterotic groups. Following an approach similar to that described in this study, this design can be considered as a factorial design between the testers and lines of one group and the testers and lines of the other group, with missing data (line x line hybrids). For this factorial, interactions can be analyzed following model [3] and used to predict the SCA of the line x line hybrids. Comparison with the actual values of a set of line x line hybrids would contribute to analyzing the efficiency of interaction multiplicative models to increase prediction efficiency.

for useful discussions on prediction strategies. We are grateful to C Lariagon, H Mitteul, D Pacault F, Pfeiffer and S Roger for their technical assistance, and JP Sampoux for his help in experiment management.

APPENDIX

Computation of variances

For a given model, the predicted value of the hybrid *(i,j)* can be written as:

$$Y_{ij} = \Phi\,(Y_{it},\,Y_{jt},\,Y_{tt'}) \qquad [1]$$

where Y_{it}, (Y_{jt}) are the observed values of crosses between line *i* and testers (line *j* and testers) and $Y_{tt'}$ are the observed values of the crosses between testers. The function Φ is not known explicitly and is only numerically computed, so it is impossible to derive an expression of the variance. The variance computation is based on the linear approximation of the Φ function about the observed values:

$$\Delta\Phi(Y_{it},\,Y_{jt},\,Y_{tt'}) \approx \underset{it}{\Sigma}\,a_{it}\Delta Y_{it} + \underset{jt}{\Sigma}b_{jt}\Delta Y_{jt} + \underset{tt'}{\Sigma}\,c_{tt'}\Delta Y_{tt'}\;[2]$$

As we only want to compare the predictions of the different hybrids, the variance of the contrasts has been approximated as:

$$\text{Var}\,(\Phi(Y_{it},\,Y_{jt},\,Y_{tt'}) - \Phi(Y_{i't},\,Y_{j't},\,Y_{tt'}))$$
$$\approx (\underset{it}{\Sigma}\,a_{it}^2 + \underset{jt}{\Sigma}\,b_{jt}^2 + \underset{i't}{\Sigma}\,a_{i't}^2 + \underset{jt}{\Sigma}\,b_{j't}^2)\sigma_e^2 \qquad [3]$$

a_{it} and b_{jt} values were computed numerically using [2] with 2 values for ΔY_{it} ($\Delta_1 = 0.05\sigma_e$ and $\Delta_2 = 0.10\sigma_e$). The validity of the linear approximation in [2] is supported by the fact that $a_{it}(\Delta_1)$ ($b_{jt}\,(\Delta_1)$) and $a_{it}(\Delta_2)$ ($b_{jt}(\Delta_2)$) are highly correlated.

REFERENCES

Charcosset A (1990) Étude de l'hétérosis chez le maïs : prédiction de la valeur d'hybrides F1 sur la base d'informations agronomiques, biochimiques et moléculaires. INAPG, Thèse de doctorat

Charcosset A, Lefort-Buson M, Gallais A (1990) Use of top-cross designs for predicting performance of maize single cross hybrids. *Maydica* 35, 23-27

Crossa J, Gauch HG Jr, Zobel RW (1990) Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. *Crop Sci* 30, 493-500

Decoux G, Denis JB (1991) *INTERA: Logiciels Pour l'Interprétation Statistique de l'Interaction Entre Deux Facteurs.* INRA, Versailles, pp 175

Denis JB (1991) Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Rev Stat Appliquée* 39(2), 5-24

Freeman GH (1973) Statistical methods for the analysis of genotype–environment interactions. *Heredity* 31, 339-354

Hallauer AR (1990) Methods used in developing maize inbreds. *Maydica* 35, 1-17

Hallauer AR, Miranda JB (1981) *Quantitative Genetics in Maize Breeding.* Iowa State Univ Press, Ames, IA, pp 468

Hanson HC, Casas E (1968) Spatial relationship among eight populations of *Zea mays* L utilizing information from a diallel mating design. *Biometrics* 24, 867-880

Hull HF (1945) Recurrent selection for specific combining ability in corn. *J Am Soc Agron* 37, 134-145

Jenkins MT, Brunson AM (1932) Methods of testing inbred lines of maize in crossbred combinations. *J Am Soc Agron* 24, 523-530

Lonnquist JH, Rumbaugh MD (1958) Relative importance of test sequence for general and specific combining ability in corn breeding. *Agron J* 50, 541-544

Malecot G (1969) *The Mathematics of Heredity.* WH Freeman and Co, San Francisco, CA

Mandel J (1971) A new analysis of variance model for non-additive data. *Technometrics* 13, 1-8

Perkins JM (1972) The principal component analysis of genotype–environmental interactions and physical measures of the environment. *Heredity* 29, 51-70

Rawlings JO, Thompson DL (1962) Performance level as criterion for the choice of maize testers. *Crop Sci* 2, 228-231

Scheffé H (1959) *The Analysis of Variance.* John Wiley and Sons, Inc, NY

Shull GH (1908) The composition of a field of maize. *Ann Breeders'Assoc Rep* 4, 296-301

Sprague GF, Tatum LA (1942) General *vs* specific combining ability in single crosses of corn. *J Am Soc Agron* 34, 923-932