Theor Appl Genet (1991) 81:265–271

Procedures for selecting composites based on prediction methods

J. B. de Miranda-Filho [1,]* and L. J. Chaves [2]

[1] Department of Genetics, Escola Superior de Agricultura Luiz de Queiroz-ESALQ, C.P. 83, 13400-Piracicaba, SP, Brazil
[2] Department of Plant Breeding, Escola de Agronomia/Univ. Fed. de Goiás, C.P. 131, 74001-Goiânia, GO, Brazil

Received July 25, 1990; Accepted July 25, 1990
Communicated by A. R. Hallauer

Summary. Procedures for selecting among parental varieties to be used in the synthesis of composites are discussed. In addition to the criterion based on the mean and variance of composites of the same size (k) proposed by Cordoso (1976), we suggest the index $I_j = w_1 \hat{v}_j + w_2 \hat{h}_j$ or $I'_j = (2/k) I_j$ for a preliminary selection among parental varieties. We show that by increasing k (size of the composite (I'_j tends to g_j, the general combining ability effect. Such a criterion is particularly important when n, the number of parental varieties, is large, so that the number of possible composites ($N_c = 2^n - n - 1$) becomes too large to be handled when using the common prediction procedures. Yield data from a 9 × 9 variety diallel cross were used for illustration.

Key words: Composites – Prediction methods

Introduction

Predicting the performance of genotypes or a population of genotypes has been an important aim of genetics, particularly with respect to economic quantitative traits. Mather (1949) and Mather and Jinks (1971) provided the basis for predicting the mean of a quantitative trait in segregating generations (e.g., F_2, backcrosses) following the cross between two inbred lines. Wright (1922) studying the performance of random crosses obtained from inbred lines stated that "a random bred stock derived from n inbred families will have $(1/n)^{th}$ less superiority over its inbred ancestry than the first cross or a random bred stock from which the inbred families might have been derived without selection". Consequently, the mean of a quantitative trait should be predicted by $\bar{Y}_2 = \bar{Y}_1 - (\bar{Y}_1 - \bar{Y}_0)/n$, as shown by Kinman and Sprague (1945), where $\bar{Y}_1$ is the average performance of all single crosses among n inbred lines and $\bar{Y}_0$ is the average performance of all inbred lines.

Composite varieties have been defined as a population obtained by intercrossing and recombining two or more open pollinating varieties. The high genetic variability that is expected in the composite varieties make them particularly suitable as base populations to be used in breeding programs (Vencovsky et al. 1973; Miranda Filho 1974a). Eberhart et al. (1967) suggested the use of Wright's formula for the prediction of the mean of composite varieties, thereby allowing selection among the predicted populations. Vencovsky (1970) extended the theory of prediction procedures, making it possible to predict the performance of composite populations for unequal participation of the parent varieties. Miranda Filho (1974b) showed the contribution of heterosis components to the predicted composite mean. Hallauer and Miranda Filho (1988) showed many aspects of the utilization of prediction procedures in maize breeding.

Prediction procedures are useful when the number of derived populations is too large and their experimental evaluation is not feasible. The possible number of composite varieties derived from n parent varieties is $N_c = 2^n - (n+1)$ (Vencovsky and Miranda Filho 1972), when an equal contribution of the parental varieties is assumed. For an n relatively large (e.g., n = 10) N_c is also large ($N_c = 1013$), making it difficult to handle the predicted results. Auxiliary criteria for selection among the predicted composites should be helpful when a great number of varieties is to be used for the synthesis of new populations. Cardoso (1976) found algebraic formulas for calculating the mean and variance of composites of the same group of size k, thus allowing a preliminary identi-

* To whom correspondence should be addressed

fication of the most promising groups. In the present study we used a modification of the procedure given by Cardoso (1976) and introduced a complementary criterion for selecting among parental varieties to be used in the synthesis of composite populations. This new criterion was derived from the procedure used by Carvalho et al. (1979) and Carvalho (1980), with a different approach.

2 Materials and methods

In this study two criteria for selecting the most promising group of composites are used.

a) One criterion consists in using the mean and variance of each group of composites of the same size k. We used the formulas given by Cardoso (1976) for estimating the mean ($\bar{Y}_k$) and variance (σ_k^2) involving composites of the same size k:

$$\bar{Y}_k = u + \left(\frac{k-1}{k}\right)\bar{h} \tag{1}$$

$$\sigma_k^2 = \frac{1}{(N_k-1)\,k^4}\left\{\binom{n-2}{k-1}\left[k^2\sum_{j=1}^{n}\hat{v}_j^2 + 4(k-1)^2\sum_{j=1}^{n}\hat{h}_j^2 + 4k(k-1)\sum_{j=1}^{n}\hat{v}_j\,\hat{h}_j\right] + 4\binom{n-4}{k-2}\sum_{j<j'}^{n}\hat{s}_{jj'}\right\} \tag{2}$$

for $k>2$ and $n-k>1$. Because the number of composites (N_k) varies among groups, we used the upper limit of the expected distribution of means to predict the value of the best composite of size k. This limit is given by $\bar{Y}_k + l_k\,\sigma_k$, where l_k corresponds to the middle range expressed in terms of standard deviation of a sample of size N_k taken from a normal population. The values of l_k were obtained making $l_k = \bar{w}_k/2$, where $\bar{w}_k$ corresponds to the mean range of samples of size N_k taken from a normal population (Table X from Tippet 1925).

b) Another criterion for selecting among composites is based on selection of the parent varieties on the basis of their contribution to the composite means. Here we used the prediction formula

$$Y_k = \bar{H} - (\bar{H} - \bar{V})/k \qquad \text{(Eberhart et al. 1967)} \tag{3}$$

and then replaced $\bar{V}$ (mean of k varieties) and $\bar{H}$ (mean of $k(k-1)/2$ crosses) by their corresponding components of means according to the following model, assuming no difference in reciprocal crosses:

$$\hat{Y}_{jj'} = \hat{u} + 1/2(\hat{v}_j + \hat{v}_{j'}) + \theta(\hat{\bar{h}} + \hat{h}_j + \hat{h}_{j'} + \hat{s}_{jj'}) \tag{4}$$

with $j = 1, 2, \ldots, n$; $\theta = 0$ for varieties ($j = j'$) and $\theta = 1$ for crosses ($j < j'$) (Gardner and Eberhart 1966). From Gardner and Eberhart (1966) and Gardner (1967), the following estimates are obtained:

$$\hat{u} = \bar{Y}; \quad \hat{\bar{h}} = \bar{Y}_h - \bar{Y}_v; \quad \hat{v}_j = Y_{jj} - \bar{Y}_v;$$

$$\hat{h}_j = \frac{n-1}{n-2}(\bar{Y}_j - \bar{Y}_h) + \frac{1}{2}(Y_{jj} - \bar{Y}_v).$$

Results (grain yield in t/ha) from a 9 × 9 diallel cross analysis reported by Miranda Filho and Vencovsky (1984) were used in this study. The analysis of variance according to the Gardner and Eberhart (1966) model showed significance for all effects except specific heterosis. Specific heterosis effects, however, were not neglected in our formulations.

3 Results and discussion

Mean and variance within groups of composites

Estimates of u, v_j, $\bar{h}$, h_j and $s_{jj'}$ from the study reported by Miranda Filho and Vencovsky (1984) are shown in Table 1. The following quantities are obtained:

$$\sum_{j=1}^{n} v_j^2 = 4.916179; \quad \sum_{j=1}^{n} \hat{h}_j^2 = 0.641187;$$

$$\sum_{j=1}^{n} \hat{v}_j\,\hat{h}_j = -1.022643; \quad \sum_{j<j'=2}^{n} \hat{s}_{jj'}^2 = 1.506675.$$

The mean and variance within groups of composites of size k, estimated from formulas (1) and (2), respectively, are shown in Table 2. The expected upper limits of the distribution are shown in the last column of Table 2. It is observed that the group mean increases and the variance decreases by increasing k.

Figure 1 gives the reference normal distributions (for infinite size, representing the finite distributions of composites) for composites of size $k=2$ to $k=8$. The expected upper limits (detached points in the curves) decrease by increasing k; this suggests that groups of smaller sizes will result in higher expected yielding composites, despite the lower mean. Therefore, composites of size $k=2$ would give a higher expected yield than composites of size $k>2$. The synthesis of composite populations has the primary objective of retaining high genetic variability so that they can be used as base populations in recurrent selection programs (Eberhart et al. 1967; Miranda Filho and Vencovsky 1984). It can be argued that the synthesis of composites of higher sizes would probably result in populations of higher genetic variability. This suggests that a fair balance between the mean and the genetic variability of the new composite must be investigated. Such a supposition, however, does not rule out the possibility of smaller size composites to show high variability, depending upon the degree of divergence between parents.

Several authors (Goodman 1965; Eberhart et al. 1967; Moll and Robinson 1967; Dudley and Moll 1969; Miranda Filho 1974a) have pointed out that the variability of a composite population is related to the genetic divergence between the parental varieties. It is also accepted that the genetic divergence between varieties is the basic condition for heterosis (Moll et al. 1962; Moll et al. 1965; Cress 1966). Miranda Filho (1974a) suggested that the genetic variability of a composite population can be roughly predicted by the average heterosis among the parental varieties as far as dominance genetic effects are concerned. He suggested that the best composites should have high values for predicted means and average heterosis. Following this procedure it is possible to select among composites by taking into account the predicted mean and the potential variability.

Table 1. Estimates of u, v_j, $\bar{h}$, h_j and $s_{jj'}$ for yield (t/ha) in a 9×9 variety diallel cross (Adapted from Miranda Filho and Vencovsky 1984)

j	$\hat{v}_j$	$\hat{h}_j$	$\hat{s}_{jj'}$ (j' equals to)							
			2	3	4	5	6	7	8	9
1	0.346	0.121	−0.224	−0.248	−0.146	0.256	0.354	0.331	−0.114	−0.156
2	0.075	0.209		0.286	−0.304	0.036	−0.029	0.317	−0.118	0.089
3	0.180	−0.107			0.420	0.016	0.253	−0.358	−0.002	0.011
4	0.102	0.150				0.088	−0.121	−0.092	0.377	0.155
5	−0.383	0.345					−0.228	−0.266	−0.138	0.237
6	−0.931	0.141						0.136	−0.133	−0.232
7	−1.236	−0.100							0.081	−0.150
8	1.425	−0.616								0.046
9	0.421	−0.143								

$\hat{u}=4.122$; $\bar{h}=0.689$

Table 2. Estimates of the mean ($\bar{Y}_k$), variance (σ_k^2), standard deviation (σ_k) and expected mean of the best composite ($\bar{Y}_k+l_k\,\sigma_k$) within groups of composites of the same size (k)

k	N_k	$\bar{Y}_k$	$\hat{\sigma}_k^2\times 10^2$	$\hat{\sigma}_k$	l_k	$\bar{Y}_k+l_k\,\sigma_k$
2	36	4.467	18.637	0.432	2.118	5.381
3	84	4.581	9.807	0.313	2.445	5.347
4	126	4.639	5.947	0.244	2.589	5.270
5	126	4.673	3.756	0.194	2.589	5.175
6	84	4.696	2.339	0.153	2.445	5.070
7	36	4.713	1.352	0.116	2.118	4.959
8	9	4.725	0.645	0.080	1.485	4.844
9	1	4.734	–	–	–	4.734

Expected mean of the whole set (502) of composites = 4.642

The criterion based on the mean and variance of composites of the same size k may indicate the upper limit of k that would be adequate for the identification of the promising composites. For example, for n = 20 it may be concluded that the best composites are of a size not larger than k = 10, and then prediction procedures should be performed within this range. Table 3 shows the total number of possible composites and the number of composites up to size k, taken from n parent varieties, calculated by $N_c=\sum_{k=2}^{n}\binom{n}{k}$ which equals $N_c=2^n-(n+1)$ for k = n (Vencovsky and Miranda Filho 1972). It is seen that for n = 20 the total number of possible composites is N_c = 1,048,555, which reduces to N_c = 616,645 for k = 10

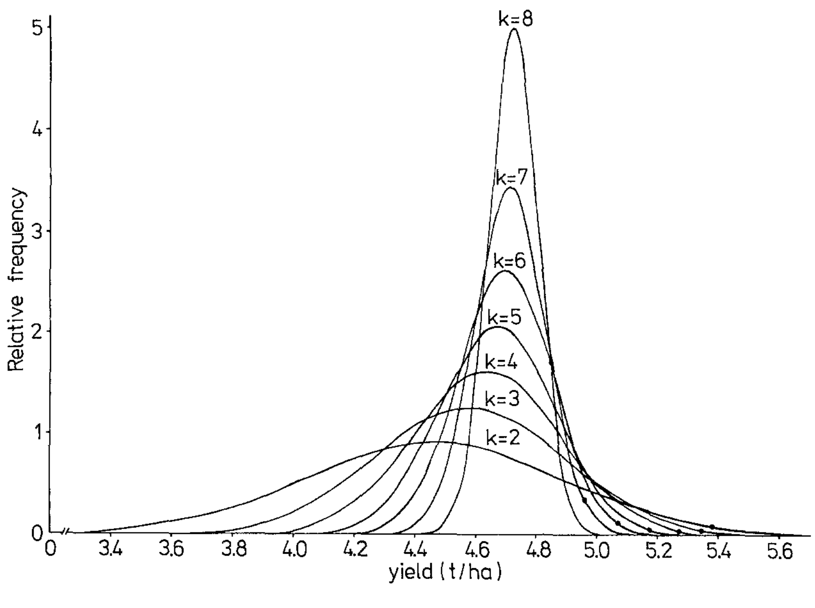

Fig. 1. Reference normal distributions for groups of composites of size k = 2 to k = 8

Table 3. Total number of possible composites (N_c) and number of composites up to size k from n parent varieties

n	N_c	Arbitrary limit of size (k)[a] of composites								
		2	3	4	5	6	7	8	9	10
2	1	1	–	–	–	–	–	–	–	–
3	4	3	4	–	–	–	–	–	–	–
4	11	6	10	11	–	–	–	–	–	–
5	26	10	20	25	26	–	–	–	–	–
6	57	15	35	50	56	57	–	–	–	–
7	120	21	56	91	112	119	120	–	–	–
8	247	28	84	154	210	238	246	247	–	–
9	502	36	120	246	372	456	492	501	502	–
10	1,013	45	165	375	627	837	957	1,002	1,012	1,013
11	2,036	55	220	550	1,012	1,474	1,804	1,969	2,024	2,035
12	4,083	66	286	781	1,573	2,497	3,289	3,784	4,004	4,070
13	8,178	78	364	1,079	2,366	4,082	5,798	7,085	7,800	8,086
14	16,369	91	455	1,456	3,458	6,461	9,893	12,896	14,898	15,899
15	32,752	105	560	1,925	4,928	9,933	16,368	22,803	27,808	30,811
16	65,519	120	680	2,500	6,868	14,876	26,316	39,186	50,626	58,634
17	131,054	136	816	3,196	9,384	21,760	41,208	65,518	89,828	109,276
18	262,125	153	969	4,029	12,597	31,161	62,985	106,743	155,363	199,121
19	524,268	171	1,140	5,016	16,644	43,776	94,164	169,746	262,124	354,502
20	1,048,555	190	1,330	6,175	21,679	60,439	137,959	263,929	431,889	616,645

[a] $k=2, 3, \ldots, K$

or to $N_c=21{,}679$ for $k=5$. It is clear that for large n, the number of predicted means still remains too high, and then an auxiliary criterion for selecting among composites may be helpful.

Contribution of varieties to the composite mean

From (2) and (4) it was found that the predicted mean of a composite of size k is given by:

$$Y_k=\hat{u}+\frac{1}{k}\sum_{j=1}^{k}\hat{v}_j+\left(\frac{k-1}{k}\right)\hat{\bar{h}}+\frac{2(k-1)}{k^2}\sum_{j=1}^{k}\hat{h}_j + \frac{2}{k^2}\sum_{j<j'=2}^{k} s_{jj'}\,. \quad (5)$$

For selecting among composites of the same size k, the terms $\hat{u}$ and $\left(\frac{k-1}{k}\right)\bar{h}$ are constants, and $\frac{2}{k^2}\sum\hat{s}_{jj'}$ is negligible for k sufficiently high. For large values of k, $\frac{2}{k^2}$ decreases rapidly and $\sum\hat{s}_{jj'}$ tends to zero (for $k=n$ and $k=n-1$, $\sum\hat{s}_{jj'}=0$). In addition, specific heterosis seems to be an unimportant source of variation in many variety diallel crosses (Hallauer and Miranda Filho 1981). Therefore, important terms for discriminating among composite means are those depending on $\hat{v}_j$ and $\hat{h}_j$ and their respective weighting coefficients. In other words, the index $I_j=w_1\hat{v}_j+w_2\hat{h}_j$, where $w_1=1/k$ and $w_2=2(k-1)/k^2$, seems to be adequate for selecting varieties as parents of composite populations of size k.

The general combining ability effect (g_j) as defined by Griffing (1956) is widely used as a measure of the contribution of a genotype in crosses. Gardner and Eberhart (1966) showed that $g_j=\frac{1}{2}v_j+h_j$. For large values of k, I_j tends to $\frac{2}{k}\left(\frac{1}{2}v_j+h_j\right)$ or $I_j\simeq\frac{2}{k}g_j$. For comparison we take the index $I'_j=\frac{k}{2}I_j$ or $I'_j=\frac{1}{2}v_j+\left(\frac{k-1}{k}\right)\hat{h}_j$, which approximates to g_j by increasing k. Table 4 shows the estimates of g_j and I'_j for the nine parental varieties by considering composites of several sizes. Such estimates are also shown graphically in Fig. 2, where it is readly seen that varieties 6 and 7 are consistently the poorest as parents for composites. The seven remaining varieties showed changes in their relative values for varying sizes of composites. Nevertheless, their values (I'_j and $\hat{g}_j$) are relatively close one to another, thus suggesting that a composite of size $k=7$ (after discarding varieties 6 and 7) would be the best choice.

In fact, Table 5 also indicates $k=7$ as the best choice because the predicted mean does not differ greatly from the best composites of smaller sizes and also because it comprises a number of parental varieties that will assure a good amount of genetic variability in the composite population. In addition, the average heterosis among the seven selected varieties is relatively high ($\bar{h}=0.684$ or 15.4%) and six-sevenths of this component is expected to be retained in the composite. The other values shown in Table 5 refer to the predicted means of the best composites within groups of size k. Table 5 also shows the identification of the parent varieties and their average heterosis, which roughly indicates the potential genetic variability of the resulting composite population.

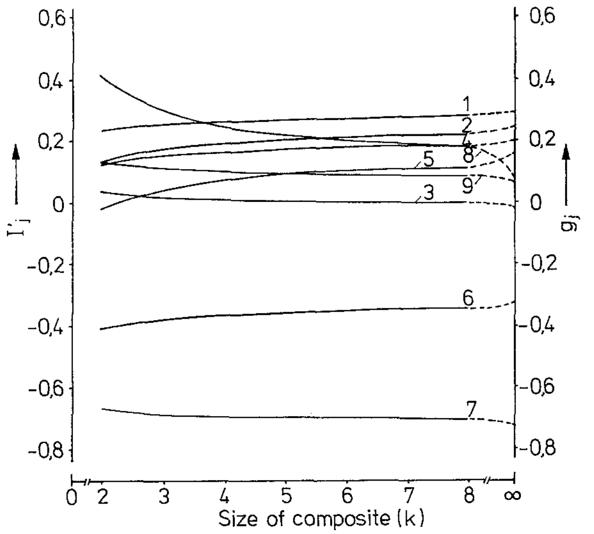

Fig. 2. Relative contribution of nine parental varieties to composite means

Table 4. Estimates ($\times 10^3$) of the general combining ability effect (g_j) and the index I'_j of varieties in composites of varying sizes (k)

j	$\hat{g}_j$	I'_j						
		k=2	k=3	k=4	k=5	k=6	k=7	k=8
1	294	233	254	264	270	274	277	279
2	274	143	177	195	205	212	217	221
3	−17	37	19	10	5	1	−2	−3
4	201	126	151	164	171	176	180	182
5	154	−19	39	67	85	96	104	111
6	−325	−395	−372	−360	−353	−348	−345	−342
7	−719	−668	−685	−694	−698	−702	−704	−706
8	97	405	302	251	220	199	185	174
9	68	139	115	103	96	91	88	86

Table 5. Predicted yield mean (Y_k) of the best composite of size k and its components $\sum \hat{v}_j$, $\sum \bar{h}_j$, $\sum s_{jj'}$, and average heterosis ($\bar{h}_k$) in crosses of the parental varieties

k	Y_k	$\sum_{j=1}^{k} \hat{v}_j$	$\sum_{j=1}^{k} \bar{h}_j$	$\sum_{j<j'=2}^{k} \hat{s}_{jj'}$	Parental[a] varieties	$\bar{h}_k$	$\bar{h}_k$%
2	5.048	1.7705	−0.4947	−0.1137	81	0.081	1.6
3	4.957	1.8459	−0.2852	−0.5088	812	0.329	6.9
4	5.002	1.9475	−0.1350	−0.5815	1824	0.525	11.4
5	5.022	2.3685	−0.2779	−0.4472	18249	0.533	11.6
6	5.048	1.9860	0.0670	0.0313	128459	0.714	16.0
7	5.018	2.1664	−0.0400	0.1361	1284593	0.684	15.4
8	4.902	1.2356	0.1006	0.0003	12485936	0.714	16.7
9	4.734	0	0	0	All	0.689	16.7

[a] Sequence is the order of selection

Y_k: Predicted according to formula (V), where $\hat{u}=4.122$ and $\bar{h}=0.689$

Predicting the mean of quantitative traits in composite populations has been used for the synthesis of new base populations to be used in recurrent selection programs in maize (Vencovsky et al. 1983; Miranda Filho 1974; Carvalho 1980). The prediction procedure is based on information obtained from diallel crosses among varieties, according to a formulation given by Eberhart et al. (1967) and extended by Vencovsky (1970). However, for large n (number of parent varieties) the number of predicted means is too great, which results in difficulty in handling data and making decisions about the performance of the predicted populations. Modern maize breeding programs tend to explore a larger portion of the great genetic variability of the maize species, as represented by more than 250 races and thousands of varieties of populations (Hallauer and Miranda Filho 1988), so that a large n can be expected to be used in the synthesis of new populations.

In this paper we presented two procedures for helping selection among the $2^n-(n+1)$ possible compósities from n parent varieties. The use of the mean and variance of composities of the same size (k) provides the basis for comparisons among and within groups (varying sizes) of composites and helps in the choice of the desirable size and choice of the best composites within that group without the need for handling hundreds or thousands of predicted means. Conversely the use of an index based on $\hat{v}_j$ (variety effect) and $\hat{h}_j$ (variety heterosis effect) makes it feasable to use the choice of the best parents for the synthesis of new composite populations. The actual power and advantages of the suggested procedures will depend upon several factors, such as the number of parent varieties, the desirable size of composites, the relative importance of non-additive genetic effects, the precision of the experimental data used for prediction, and others. It is worth noting that both the preliminary analysis of variance and the analysis of the variety diallel crosses (Gardner and Eberat 1966) provide much valuable information with respect to the prediction procedures and their basic assumptions.

Appendix

Symbols used that are not explained in the text

u	expected mean of n parental varieties used in a diallel test
$\bar{h}$	average heterosis of all crosses among n parental varieties
v_j	effect of j^{th} variety
h_j	heterosis effect of j^{th} variety
$s_{jj'}$	specific heterosis of the cross between j^{th} and j'^{th} varieties
g_j	general combining ability effect of j^{th} variety
$\binom{n}{k}$	combinations of n elements, k at a time, where $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$
Y_{jj}	observed mean of j^{th} variety
$Y_{jj'}$	observed mean of the cross between j^{th} and j'^{th} varieties

$\bar{Y}_v$ mean of n parental varieties
$\bar{Y}_h$ mean of all single crosses among n parental varieties.

Mean and variance of composites of the same size k

The model for the mean of a quantitative trait for varieties and variety crosses in a diallel mating scheme is given by:

$$Y_{jj'} = u + 1/2\,(v_j + v_{j'}) + \theta(\bar{h} + h_j + h_{j'} + s_{jj'}) + \bar{e}_{jj'}$$

(Gardner and Eberhart 1966). The predicted mean of a composite of size k from a set of n varieties is given by:

$$Y_k = \bar{Y}_{h*} - \frac{1}{k}(\bar{Y}_{h*} - \bar{Y}_{v*})$$

(Eberhart et al. 1967), where $\bar{Y}_{v*}$ and $\bar{Y}_{h*}$ stand for the means of varieties and variety crosses within the set of k varieties. The same formula can be rewritten as

$$Y_k = \frac{1}{k^2}\left(\sum_j^{k*} Y_{jj} + 2\sum_{j<j'}^{k*} Y_{jj'}\right);$$

k* means that the summation is restricted to the set of k varieties. In terms of the diallel model, it follows

$$Y_k = \frac{1}{k^2}\left[k^2 u + k(k-1)\bar{h} + k\sum_j^{k*} v_j + 2(k-1)\sum_j^{k*} h_j + 2\sum_{j<j'}^{k*} s_{jj'}\right].$$

The general mean of the N_k composites of the same size is:

$$\bar{Y}_k = \frac{1}{N_k}\sum_{i=1}^{N_k}\left[u + \frac{k-1}{k}\bar{h} + \frac{1}{k}\sum_j^n v_j + \frac{2(k-1)}{k^2}\sum_j^n h_j + \frac{2}{k^2}\sum_{j<j'}^n s_{jj'}\right],$$

or

$$\bar{Y}_k = u + \frac{k-1}{k}\bar{h}, \text{ because } \sum_j^n v_j = \sum_j^n h_j = \sum_{j<j'}^n s_{jj'} = 0$$

(for simplicity, we omitted the hat in the estimates of v_j, h_j and $s_{jj'}$).

The variance among composites of the same size k is estimated by:

$$\sigma_k^2 = \frac{1}{N_k - 1}\left\{\sum_{i=1}^{N_k}\left[\frac{1}{k^2}\left(\sum_j^{k*} Y_{jj} + 2\sum_{j<j'}^{k*} Y_{jj'}\right) - \bar{Y}_k\right]^2\right\},$$

where $\sum_{i=1}^{N_k}$ is a summation over all composites of the same size k; $N_k = \binom{n}{k}$. By ignoring $u + \frac{k-1}{k}\bar{h}$ and Y_k, that are constants and cancel out each other in the sum of squares, we have:

$$\sigma_k^2 = \frac{1}{(N_k-1)k^4}\Bigg[k^2\binom{n-1}{k-1}\sum_j^n v_j^2 + 2k^2\binom{n-2}{k-2}\sum_j^n v_j v_{j'}$$
$$+4(k-1)^2\binom{n-1}{k-1}\sum_j^n h_j^2 + 8(k-1)\binom{n-2}{k-2}\sum_j^n h_j h_{j'}$$
$$+4k(k-1)\binom{n-1}{k-1}\sum_j^n v_j h_j + 4k(k-1)\binom{n-2}{k-2}\sum_{j \neq j'}^n v_j h_{j'}$$
$$+4k\binom{n-2}{k-2}\sum_j^n v_j s_{jj'} + 4k\binom{n-3}{k-3}\sum_j^n v_j s_{j'j''}$$
$$+8(k-1)\binom{n-2}{k-2}\sum_j^n h_j s_{jj'} + 8(k-1)\binom{n-3}{k-3}\sum_{j',j''}^n h_j s_{j'j''}$$
$$+4\binom{n-2}{k-2}\sum_{j<j'}^n s_{jj'}^2 + 8\binom{n-3}{k-3}\sum_j^n s_{jj'} s_{jj''}$$
$$+8\binom{n-4}{k-4}\sum_j^n s_{jj'} s_{j''j'''}\Bigg].$$

The combinations of parameters can be arranged in two-way tables (rows and columns) so that for each row it follows that:

1. $$k^2\binom{n-1}{k-1}v_j^2 + k^2\binom{n-2}{k-2}\sum_{j=j'}^n v_j v_{j'}$$
$$= k^2\left[\binom{n-1}{k-1} - \binom{n-2}{k-2}\right]v_j^2 + k^2\binom{n-2}{k-2}v_j\sum_j^n v_j = k^2\binom{n-2}{k-1}v_j^2,$$

and for the whole table we have

$$k^2\binom{n-2}{k-1}\sum_j^n v_j^2.$$

2. $$4(k-1)^2\left[\binom{n-1}{k-1}h_j^2 + \binom{n-2}{k-2}\sum_{j'}^n h_j h_{j'}\right]$$
$$= 4(k-1)^2\left[\left(\binom{n-1}{k-1} - \binom{n-2}{k-2}\right)h_j^2 + \binom{n-2}{k-2}h_j\sum_j^n h_j\right]$$
$$= 4(k-1)^2\binom{n-2}{k-2}h_j^2,$$

and for the whole table we have

$$4(k-1)^2\binom{n-2}{k-1}\sum_j^n h_j^2.$$

3. $$4(k-1)\left[\binom{n-1}{k-1}v_j h_j + \binom{n-2}{k-2}\sum_{j'}^n v_j h_{j'}\right]$$
$$= 4(k-1)\left[\left(\binom{n-1}{k-1} - \binom{n-2}{k-2}\right)v_j h_j + \binom{n-2}{k-2}v_j\sum_j^n h_j\right]$$
$$= 4(k-1)\binom{n-2}{k-1}v_j h_j,$$

and for the whole table we have

$$4k(k-1)\binom{n-2}{k-1}\sum_j^n v_j h_j.$$

4. $$4k\left[\binom{n-2}{k-2}\sum_{j'}^n v_j s_{jj'} + \binom{n-3}{k-3}\sum_{j',j''}^n v_j s_{j'j''}\right]$$
$$= 4k\Bigg[\left(\binom{n-2}{k-2} - \binom{n-3}{k-3}\right)v_j\sum_{j<j'}^n s_{jj'}$$
$$+\binom{n-3}{k-3}v_j\left(\sum_{j'}^n s_{jj'} + \sum_{j''}^n s_{j'j''}\right)\Bigg] = 0;$$

in the whole table the coefficient also is zero.

5. $$8(k-1)\left[\binom{n-2}{k-2}h_j\sum_j^n s_{jj'} + \binom{n-3}{k-3}h_j\sum_{j',j''}^n s_{j'j''}\right] = 0, \text{ as in 4.}$$

6. $$4\Bigg[\binom{n-2}{k-2}s_{jj'}^2 + \binom{n-3}{k-3}\sum_{j''}^n s_{jj'} s_{jj''} + \binom{n-3}{k-3}\sum_{j''}^n s_{jj'} s_{j'j''}$$
$$+\binom{n-4}{k-4}\sum_{j'',j'''}^n s_{jj'} s_{j''j'''}\Bigg] = 4\Bigg[\left(\binom{n-2}{k-2} - \binom{n-4}{k-4}\right)s_{jj'}^2$$
$$+\left(\binom{n-3}{k-3} - \binom{n-4}{k-4}\right)s_{jj'}\sum_{j''}^n s_{jj''}$$
$$+\left(\binom{n-3}{k-3} - \binom{n-4}{k-4}\right)s_{jj'}\sum_{j''}^n s_{j'j''}$$
$$+\binom{n-4}{k-4}s_{jj'}\left(s_{jj'} + \sum_{j''}^n s_{jj''} + \sum_{j''}^n s_{j'j''} + \sum_{j',j'''}^n s_{j''j'''}\right)\Bigg]$$

$$=4\left[\left(\binom{n-2}{k-2}-2\binom{n-3}{k-3}+\binom{n-4}{k-4}\right)s_{jj'}^2\right.$$
$$+\left(\binom{n-3}{k-3}-\binom{n-3}{k-4}\right)s_{jj'}\left(s_{jj'}+\sum_{j''}^{n}s_{jj''}\right)$$
$$\left.+\left(\binom{n-3}{k-3}-\binom{n-4}{k-4}\right)s_{jj'}\left(s_{jj'}+\sum_{j''}^{n}s_{j'j''}\right)+0\right]$$
$$=4\left[\binom{n-4}{k-2}s_{jj'}^2+0+0+0\right];$$

in the whole table we have

$$4\binom{n-4}{k-2}\sum_{j<j'}^{n}s_{jj'}^2.$$

7. Putting all terms together:

$$\sigma_k^2=\frac{1}{(N_k-1)k^4}\left\{\binom{n-2}{k-1}\left[k^2\sum_j\hat{v}_j+4(k-1)^2\sum_j\hat{h}_j^2\right.\right.$$
$$\left.\left.+4k(k-1)\sum_j\hat{v}_j\hat{h}_j\right]+4\binom{n-4}{k-2}\sum_{j<j'}\hat{s}_{jj'}^2\right\}.$$

Reference

Cardoso AA (1976) Variação entre compostos e sua utilização no melhoramento do milho. PhD thesis, ESALQ-USP, Piracicaba

Carvalho HWL (1980) Predição de médias de compostos de milho (*Zea mays* L.) para a microregião homogêna 131 do Estado da Bahia. MSc thesis, ESALQ-USP, Piracicaba

Carvalho HWL, Miranda Filho JB, Chaves LJ (1979) Critérios para a seleção entre compostos com base nos métodos de predição. Rel Cient Instit Genét (ESALQ-USP) 13:83–94

Cress CE (1966) Heterosis of the hybrid related to gene frequency differences between two populations. Genetics 53:269–274

Dudley JW, Moll RH (1969) Interpretation and use of estimates of heritability and genetic variances in plant breeding. Crop Sci 9:257–262

Eberhart SA, Harrison MN, Ogada F (1967) A comprehensive breeding system. Der Zucht 37:169–174

Gardner CO (1967) Simplified methods for estimating constants and computing sums of squares for a diallel cross analysis. Fitotec Latinoam 4:1–12

Gardner CO, Eberhart SA (1966) Analysis and interpretation of the variety cross diallel and related populations. Biometrics 22:439–452

Goodman M (1965) Estimates of genetic variance in adapted and exotic populations of maize. Crop Sci 5:87–90

Hallauer AR, Miranda Filho JB (1988) Quantitative genetics in maize breeding. Iowa State Univ Press, Ames, Iowa

Kinman ML, Sprague GF (1945) Relation between number of parental lines and theoretical performance of synthetic varieties of corn. Agron J 37:341–351

Mather K (1949) Biometrical genetics. Methuen, London

Mather K, Jinks JL (1971) Biometrical genetics. Chapman & Hall, London

Miranda Filho JB (1974a) Cruzamentos dialélicos e síntese de compostos de milho (*Zea mays* L.) com ênfase na produtividade e no porte da planta. PhD thesis, ESALQ-USP, Piracicaba

Miranda Filho JB (1974b) Predição de médias de compostos em função das médias das variedades parentais e das heteroses dos cruzamentos. Rel Cient Instit Genét (ESALQ-USP) 8:134–138

Miranda Filho JB, Vencovsky R (1984) Analysis of diallel crosses among open-pollinated varieties of maize (*Zea mays* L.). Maydica 29:217–234

Moll RH, Robinson HF (1967) Quantitative genetic investigations of yield of maize. Der Zucht 37:192–199

Moll RH, Salhuana WS, Robinson HF (1962) Heterosis and genetic diversity in variety crosses of maize. Crop Sci 2:197–198

Moll RH, Lonnquist JH, Fortuno JV, Johnson EC (1965) The relationship of heterosis and genetic divergence in maize. Genetics 52:139–144

Tippett LHC (1925) On the extreme individuals and the range of samples taken from a normal population. Biometrika 17:364–387

Vencovsky R (1970) Alguns aspectos teóricos e aplicados relativos a cruzamentos dialélicos de variedades. "Livre Docente" thesis, ESALQ-USP, Piracicaba

Vencovsky R, Miranda Filho JB (1972) Determinação do número de possíveis compostos e pares de compostos. Rel Cient Instit Genét (ESALQ-USP) 6:120–123

Vencovsky R, Zinsly JR, Vello NA, Godoy CRM (1973) Predição da média de um caráter quantitativo em compostos de variedades e cruzamento de compostos. Fitotec Latinoam 8:25–28

Wright S (1922) The effects of inbreeding and crossbreeding on guinea pigs. III. Crosses between highly inbred families. USDA Bull 1121