Theor Appl Genet (1999) 99:1018–1030

M. Brancourt-Hulmel

Crop diagnosis and probe genotypes for interpreting genotype environment interaction in winter wheat trials

Received: 3 August 1998 / Accepted: 16 March 1999

Abstract Genotype*environment interaction has been analyzed with 12 genotypes and four probe genotypes in French wheat trials. An integrated approach was developed which combined crop diagnosis with the analysis of interaction by factorial regression. Crop diagnosis was helpful to characterize the environments and to select environmental variables. Such an approach succeeded in providing an agronomic explanation of genotype*environment interaction and in defining the responses or parameters for each genotype and each environment. Earliness at heading, susceptibility to powdery mildew and susceptibility to lodging were the three major genotypic covariates. Interaction could also be related to environment features, measured indirectly by the behavior of the four probe genotypes during the formation of yield, what we called the outputs of a simplified crop diagnosis, or described directly by indicators of yield-limiting factors. Two important crop diagnosis covariates were analyzed in order to characterize interaction during the formation of yield: the reduction in kernel number, which described the time-period until flowering, and the reduction in thousand kernel weight, which corresponded to the period after flowering. These variates were estimated for each probe genotype and allowed us to compare the behavior of the 12 genotypes to that of the probe genotypes. Both periods of the formation of yield contributed to the interaction, and 'Camp-Rémy' was the probe of particular interest for the comparisons. When true environmental variates were used, factorial regression revealed that water deficits during the formation of grain number and level of nitrogen were predominant. Such an integrated approach could be exploited when varieties are tested in a network where numerous and diverse yield-limiting factors may occur.

Communicated by H. Becker

M. Brancourt-Hulmel
Unité de génétique et d'amélioration des plantes, INRA,
F-80200 Estrées-Mons, France
e-mail: hulmel@mons.inra.fr
FAX: 33.03.22.85.69.96

Key words Genotype*environment interaction · Factorial regression · Crop diagnosis · Probe genotypes · Wheat

Introduction

Evaluation of genotypic performance in multi-environment trials has to deal with the analysis of genotype*environment interaction. During recent decades, new developments have been achieved in crop physiology, agronomy and modeling, and some approaches integrate these results in the evaluation of genotype*environment interaction. Bidinger et al. (1996) and Hammer et al. (1996) related genotype*environment interaction in crop adaptation to a physiological basis. Henderson et al. (1996) compared simulation and experimental approaches for the analysis of genotype*environment interactions in rice. According to these authors, the discrepancies between the two methods underline the effects of factors other than phenology. When simulation models are used in the analysis of genotype*environment interaction, they need to be built with a great accuracy, and this goal has not yet been completely achieved. The development of a strategy for integrating several components such as modeling, agronomy, meteorological data, genetics and plant breeding will contribute to an improvement in both crop evaluation and breeding methods.

Environments can be characterized using two approaches: (1) with a direct characterization by measuring environmental variables which can be physical, biological or nutritional, and (2) with an indirect characterization by measuring plant responses to capture the influence of environmental conditions on plant performance. Cooper and Fox (1996) distinguished two types of genotypes for indirect characterization: reference genotypes and probe genotypes. The concept of reference genotypes, a random sample of genotypes chosen on the basis of high environmental instability, is used for characterizing a target population of environments (Fox and Rosielle 1982). Probe genotypes are specific genotypes chosen for their known reaction to one or several environmental factors (Cooper and Fox 1996;

Desclaux 1996; Brancourt-Hulmel et al. 1999). The relative performance of the genotypes which comprise the probe set can then be used to judge the incidence of the environmental factor in multi-environment trials (Cooper and Fox 1996). Wade et al. (1996) used two rice cultivars of opposite drought resistance as probe lines for drought characterization. When the climatic conditions are very diverse in a network, as often occurs in France, it is impossible to choose the genotypes on the basis of their response to only one environmental factor. This would considerably increase the number of probes and limit the relevance of such an approach. Cooper and Fox (1996) exploited probe genotypes in indirect selection as a tool for identifying and quantifying the incidence of repeatable genotype*environment interactions. They distinguished these from interactions which are less repeatable. Probe genotypes can also be used for identifying the main limiting factors of yield occurring in wheat trials (Brancourt-Hulmel et al. 1999). The analysis is based on a simplified crop diagnosis which helps to identify *a posteriori* limiting factors of yield by examining losses of yield components. For an understanding of genotype*environment interaction, the probe genotypes can also be used as a comparison to a given set of genotypes. The objective of the present paper is to develop an integrated approach which combines the characterization of environments by means of a crop diagnosis with the analysis of genotype*environment interaction. The analysis of genotype*environment interaction will be performed by factorial regression (Denis 1980, 1988). In comparison to several models, this method is relevant when external information or covariates about both environments and genotypes are available and in many cases is helpful for a biological understanding of the genotype*environment interaction (Baril 1992, Baril et al. 1995; van Eeuwijk et al. 1995; Brancourt-Hulmel et al. 1997).

Materials and methods

Experimental data

A set of 12 winter wheat genotypes was evaluated for genotype*environment interaction, and their main features are given in Table 1. They differed in many traits such as earliness at heading and susceptibilities to diseases or to lodging. Most of these genotypes were commonly grown wheat varieties in France during the period under consideration.

The experiments were carried out over 2 years (1991, 1992) in France at five I.N.R.A. locations: Mons (MON), Rennes (REN), La Minière (MIN), Dijon (DIJ) and Ondes (OND). Two agronomic treatments were applied: medium/late sowing date at Dijon (IN, S2), treatment with/without fungicides at the other sites (IN, -F). At each of the five evaluation locations, the design was a randomized complete block. Overall, two blocks were used, except at Mons in 1992 where the experiment was conducted with three blocks. For the analysis of genotype*environment interaction, one environment was considered as a combination of year*location*treatment and thus a total of 20 environments were evaluated.

These environments have been fully characterized in a previous paper by a simplified crop diagnosis using four probe genotypes (Brancourt-Hulmel et al. 1999). 'Talent' (TAL), 'Soissons' (SOI), 'Camp-Rémy' (CAR) and 'Arminda' (ARM) constituted an independent specific set of probe genotypes and differed in earliness. Three of them were tested twice as they were common to the above-mentioned set of 12 genotypes.

Plant sampling and measurements

Grain yields were determined by harvesting mechanically all six rows of each plot at maturity, and susceptibilities to diseases and to lodging, height, earliness were also observed. About 1 day be-

Table 1 Main characteristics of the 12 genotypes. Susceptibilities were measured on a scale from 1 (resistant) to 9 (heavily damaged)

Genotype	HD[a]	HT[b]	PMK[c]	PMT[d]	PMTs[d]	LRT[e]	LBT[f]	LGBT[g]	FT_1[h]	FT_2[h]	LodgK[i]	LodgT[i]
APO[k]	149	102	5	7	1	4	4	1	2	4	3	4
ART	140	99	4	5	4	3	8	3	2	3	1	3
BAR	146	97	3	3	3	7	7	2	2	4	2	4
CAR	148	101	4	4	4	7	6	1	2	3	4	3
GEN	146	96	3	3	1	1	6	1	2	3	1	3
REC	138	87	5	6	4	2	9	1	4	3	1	2
REN	147	98	2	2	4	1	3	1	2	4	4	3
ROS	143	93	2	3	3	3	8	4	3	5	1	3
SOI	141	86	2	4	1	4	7	2	2	4	1	2
TAL	139	93	3	4	4	6	6	2	2	3	3	4
THE	143	93	4	5	3	6	8	4	3	4	2	2
VIK	148	99	4	4	4	3	7	1	3	4	1	3

[a] HD, Earliness at heading (expressed as days from the first of january)
[b] HT, Height (in cm)
[c] PMK, Susceptibility to powdery mildew during the formation of grain number
[d] Susceptibility to powdery mildew during grain-filling: PMT, on leaves; PMTs, on spikes
[e] LRT, Susceptibility to leaf rust during grain-filling
[f] LBT, Susceptibility to leaf blotch during grain-filling
[g] LGBT, Susceptibility to leaf and glume blotch during grain-filling
[h] Susceptibility to fusarium during grain-filling: FT_1, early observation; FT_2, late observation
[i] Lodging: LodgK, susceptibility to lodging during the formation of grain number; LodgT, susceptibility to lodging during grain-filling
[k] APO, 'Apollo'; ART, 'Artaban'; BAR, 'Baroudeur'; CAR, 'Camp-Rémy'; GEN, 'Génial'; REC, 'Récital'; REN, 'Renan'; ROS, 'Rossini'; SOI, 'Soissons'; TAL, 'Talent'; THE, 'Thésée'; VIK, 'Viking'

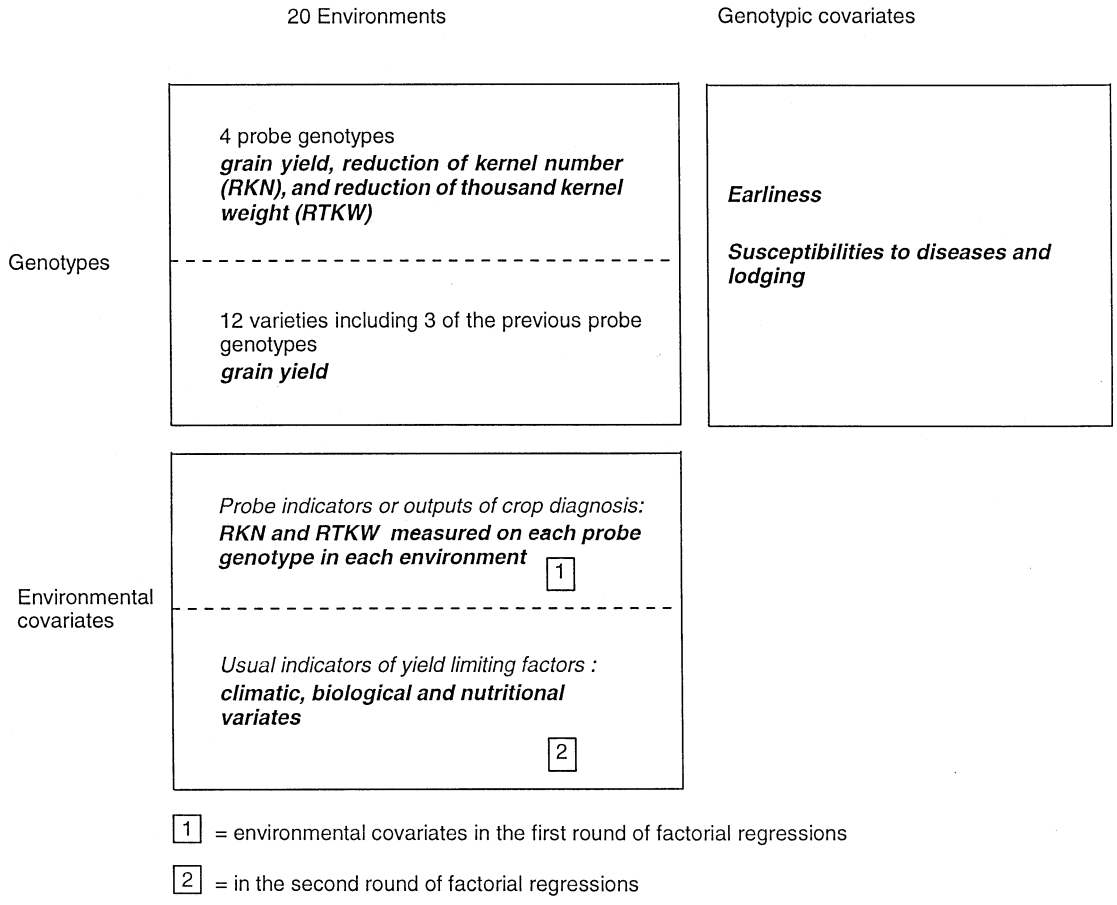

Fig. 1 General presentation of the data used in the different factorial regression models

fore mechanical harvest, 150 shoots were sampled in the same plot in the inside four rows in order to measure yield components: kernel number (KN) and thousand kernel weight (TKW).

Definition of covariates

The environmental covariates were of two kinds (Fig. 1): indicators of yield-limiting factors and outputs from crop diagnosis. Two series of factorial regression were carried out in order to determine the sensitivity of a given genotype to one or several indicators of yield-limiting factors and to compare it to the behavior of one or several probes during the formation of yield.

The characterization of the environments borrowed a method developed by Sebillotte (1980) and has been adapted to a breeding network (Brancourt-Hulmel et al. 1999). In winter wheat, grain yield comes from two main yield components: kernel number per square meter (KN) and thousand kernel weight (TKW). A procedure to estimate potential values for these yield components has been developed using a relationship between the two (Brancourt-Hulmel et al. 1999). Plants free from stress or well adapted to it would produce yield components close to the potential values. Two variates or outputs from crop diagnosis were calculated from the relationship TKW= f(KN) for each probe genotype in each environment:

reduction of kernel number :

RKN=max[0;100* ($KNthreshold - KN$) / $KNthreshold$]

reduction of thousand kernel weight :

$RTKW$=max[0;100* ($potentialTKW - TKW$) / $potentialTKW$]

RKN describes the kernel number formation period, whereas RTKW accounts for the grain filling period. Environments were considered as optimal for these periods when the KN and TKW of each probe were equal to threshold or potential values and were limited by factors when yield components were reduced. Therefore, these reduction values can be used as « probe indicators » for characterizing the environment for further investigation of the genotype*environment interaction. An equivalent criteria was also determined for grain yield (RGY). When introduced in a factorial regression model, the probe indicators, namely RKN and RTKW, enable a comparison of the genotypes to the behavior of the probe genotypes during the formation of the grain number (when RKN is introduced) or the grain filling period (when RTKW is introduced). Eight environmental covariates were provided as each variate was calculated for each probe, ARMrkn, CARrkn, SOIrkn and TALrkn for the reduction of kernel number and ARMrktw, CARrtkw, SOIrtkw and TALrtkw for the reduction of thousand kernel weight.

The other environmental covariates were physical (climatic or not), biological or nutritional indicators and were determined for the two main periods of the formation of yield: the formation of grain number, which corresponds to the time-period until flowering, and grain filling, which occurs after flowering. Several variates were used as indicators for yield-limiting factors for the period of grain number formation, namely the sum of daily water deficits (ET_m–ET_a) from ear at 1 cm to flowering (WDK), the ratio between total nitrogen absorbed by the plant and kernel number (BK), the sum of daily radiation from ear at 1 cm to flowering (RK), the sum of daily radiation +/- 3 days at meiosis (RKm) and infection of powdery mildew (PMK). BK was used to indicate nitrogen stress. During that period, no days with a minimum daily temperature below –4 °C occurred, and some varieties were lodged (LodgK). For the grain filling period, the sum of daily water deficits (WDT), the sum of daily radiation (RT) and high temperature estimated by the sum of degree-days based on 25°C

(HTT) were computed. Infections of powdery mildew (PMT on leaves and PMTs on spikes), leaf rust (LRT), leaf blotch (LBT), leaf and glume blotch (LGBT) and fusarium (early observation FT1 and late observation FT2) were observed. No stripe rust was recorded in the network during 1991 and 1992. Lodging occurred during grain filling (LodgT). Susceptibilities were the maximum scores noted on a given genotype (probe genotype or not), and other indicators were calculated according to the cycle of each probe genotype. Used as environmental covariates in a factorial regression model, these indicators enable us to determine the sensitivity of a given genotype to a yield-limiting factor. Earliness at heading and susceptibility to diseases and lodging were used as genotype covariates.

Statistical analysis

Genotype*environment interaction was first analyzed according to the following model 1 using the SAS Software System (1989):

$$E[Y_{gek}] = \mu + \alpha_g + \beta_e + b_{ek} + \alpha\beta_{ge} \quad \text{[model 1]}$$

where $E[Y_{gek}]$ is the expectation of a given yield Y_{gek} for genotype g grown in block k in environment e, μ is the grand mean, is the genotype main effect, β_e is the environment main effect, b_{ek} is the effect of block k in environment e and $\alpha\beta_{ge}$ is the interaction between genotype and environment. All terms in ANOVA were considered as fixed effects. Moreover, the environment effect was partitioned into year, location, treatment and the corresponding interaction terms. The partitioning was applied to environment main effect β_e as well as to the interaction term $\alpha\beta_{ge}$.

In a further investigation, covariates related to the genotypic factor as well as the environmental factor were introduced in a factorial regression for modeling the main effects (models 2 and 3) and the interaction term (models 4, 5 and 6).

Genotypic main effect:

$$\alpha_g = \sum_k \rho_k \cdot G_{gk} \quad \text{[model 2]}$$

Environment main effect:

$$\beta_e = \sum_h \delta_h \cdot H_{eh} \quad \text{[model 3]}$$

Interaction:

$$\alpha\beta_{ge} = \sum_{hk} G_{gk} \cdot \theta_{kh} \cdot E_{eh} \quad \text{[model 4]}$$

$$\alpha\beta_{ge} = \sum_{hk} G_{gk} \cdot \theta_{kh} \cdot E_{eh} + \sum_h \alpha'_{gh} \cdot E_{eh} + \sum_k \beta'_{ek} \cdot G_{gk} \quad \text{[model 5]}$$

$$\alpha\beta_{ge} = \sum_h \alpha''_{gh} \cdot E_{eh} \quad \text{[model 6]}$$

θ_{kh}, α'_{gh}, α''_{gh} and β'_{ek} are regression parameters involving H environment covariates E_{eh} and K genotype covariates G_{gk}. θ_{kh} represent coefficients with respect of covariates cross-products which are not dependent on either genotype or environment. As such (in model 4), they provide the most simple way to deal with interaction (Baril et al. 1995). But very often they are not sufficient, and more elaborate models, as model 5, are needed where genotypic or environmental parameters are determined. To clarify the interpretation, we will first consider the simple model involving environmental covariates (model 6). For testing effects in models 2–6, we used the variance estimated in model 1, divided by 2, as the external error variance. The main reason is that INTERA (Decoux and Denis 1991) allows only means data. Since in 2 environments, there were three replications, and three values were missing, this variance is approximate. But this was considered as valid since it was noted that the individual error variances, resulting from two or three blocks, were similar between environments, and the design was slightly unbalanced as the three missing values were observed for 3 different genotype*environment combinations.

In the two series of factorial regression analyses, the model fitted to grain yield means was built by progressive addition of the most significant covariates according to a stepwise process proposed by Denis (1988). Several rounds of analyses using different orderings of covariates were carried out to determine the best subset. Factorial regressions and corresponding plots were performed using the INTERA package on a PC (Decoux and Denis 1991). Model 5 will be presented using recapitulative tables proposed by Denis (1991) which are more comprehensible than the usual presentation. Each cell of the table corresponds to one term of model 5 with its corresponding degrees of freedom, sum of squares and mean square.

Development of the models with 3 genotypic and 3 environmental covariates

To get more familiar with the previous models, we present some examples in this section. The optimal factorial regression model was obtained with 3 genotypic covariates, namely earliness at heading (HD), susceptibility to powdery mildew during the formation of grain number (PMK) and the susceptibility to lodging during the grain filling period (LodgT) and with 3 environmental covariates, reduction of kernel number for 'Camp-Rémy' (CARrkn), reduction of thousand kernel weight for 'Camp-Rémy' (CARrtkw) and reduction of thousand kernel weight for 'Soissons' (SOIrtkw).

Using these covariates, the models 2–6 become as follows where indices number the parameters:

$$\sum_k \rho_k \cdot G_{gk} = \rho_1 \cdot HD + \rho_2 \cdot PMK + \rho_3 \cdot LodgT \quad \text{[model 2]}$$

$$\sum_h \delta_h \cdot H_{eh} = \delta_1 \cdot CARrkn + \delta_2 \cdot CARrtkw + \delta_3 \cdot SOIrtkw \quad \text{[model 3]}$$

$$\begin{aligned}\sum_{hk} G_{gk} \cdot \theta_{kh} \cdot E_{eh} = {} & \theta_{11} \cdot HD \cdot CARrkn + \theta_{12} HD \cdot CARrtkw \\ & + \theta_{13} \cdot HD \cdot SOIrtkw + \theta_{21} \cdot PMK \cdot CARrkn \\ & + \theta_{22} \cdot PMK \cdot CARrtkw + \theta_{23} \cdot PMK \cdot SOIrtkw \\ & + \theta_{31} \cdot LodgT \cdot CARrkn + \theta_{32} \cdot LodgT \cdot CARrtkw \\ & + \theta_{33} \cdot LodgT \cdot SOIrtkw\end{aligned} \quad \text{[model 4]}$$

$$\sum_h \alpha''_{gh} \cdot E_{eh} = \alpha''_{g1} \cdot CARrkn + \alpha''_{g2} \cdot CARrtkw + \alpha''_{g3} \cdot SOIrtkw \quad \text{[model 6]}$$

Results

Analysis of variance

The results of ANOVA using model 1 are shown in Table 2. Genotype*environment interaction is statistically significant. The mean grain yield is 67.2 q/ha (at 0% moisture content), and the residual is small (3.6) indicating a good accuracy of the trials. There are three missing values, and 501 observations used in the analysis (*i.e.* 12 genotypes*20 environments*2 blocks plus a third block in 2 environments).

Results of factorial regression using model 5 are given in Table 3. A slight discrepancy between the estimations of the total interaction sum of squares between this table and the previous one is owing to the unbalanced design. In Table 3, environmental covariates are the outputs from a simplified crop diagnosis: reduction of kernel number (RKN) and reduction of thousand kernel weight (RTKW) determined for each probe. Earliness of heading is the first genotypic covariate to be introduced (HD). It gives the same result as a model using susceptibility to leaf blotch (LBT) since these

Table 2 ANOVA table of grain yield with model 1

Source	*df*	Type I SS	Mean square	*F* value	Pr > *F*
Genotype	11	4 765	433	33.4	0.0001
Environment	19	40 483	2131	164.4	0.0001
Block in environment	22	1 391	653	4.9	0.0001
Genotype*environment	209	14 200	68	5.2	0.0001
Residual	239	3 097	13		
	R-square	C.V.	Root MSE	GY mean	
	0.95	5.4%	3.6	67.2 q/ha	

Table 3 Decomposition of the interaction for first factorial regression. Case of environment covariates RKN and RTKW calculated fro Camp-Rémy (CARrkn and CARrtkw) and Soissons (SOIrtkw). Earliness at heading (HD), PMK (powdering mildew during shooting) ad LodgT (lodging during grain filling) were used as genotypic covariates. Each cell corresponds to one term of the model with degrees of freedom [in brackets], sum of squares in bold and mean square in italics

* Significant at the 0.05 probability level

*HD*CARrkn* [1] **7** *7*	*HD*CARrtkw* [1] **606** *606**	*HD*SOIrtkw* [1] **36** *36*	*HD*β'_{el}* [16] **1487** *93**
*PMK*CARrkn* [1] **1** *1*	*PMK*CARrtkw* [1] ***82*** *82**	*PMK*SOIrkw* [1] **218** *218**	*PMK*β'_{e2}* [16] ***439*** *27**
*LodgT*CARrkn* *[1]* **0** *0*	*LodgT*CARrkw* [1] **5** *5*	*LodgT*SOIrtkw* [1] **4** *4*	*LodgT*β'_{e3}* [16] **441** *28**
CARrknα'_{g1}* [8] **645** *81**	*CARrtkw*α'_{g2}* [8] **380** *47**	*SOIrtkw*α'_{g3}* [8] **210** *26**	Remaining interaction [128] ***2432*** *19**

Table 4 Correlations between genotypic covariates. Means of 12 genotypes. Abbreviations are as in the Table 1

	HD	HT	PMK	PMT	PMTs	LRT	LBT	LGBT	FT_1	FT_2	LodgK	LodgT
HD	1											
HT	**0.78***	1										
PMK	0.01	0.18	1									
PMT	–0.07	0.12	**0.89***	1								
PMTs	–0.19	0.16	0.06	–0.17	1							
LRT	0.08	0.15	0.00	0.09	0.15	1						
LBT	**–0.65***	–0.55	0.12	0.06	0.18	0.12	1					
LGBT	–0.41	–0.30	–0.40	–0.18	0.04	0.22	0.50	1				
FT_1	–0.31	–0.45	0.45	0.30	0.09	–0.25	**0.58***	0.22	1			
FT_2	0.13	–0.20	–0.36	–0.17	–0.40	–0.14	–0.03	0.47	0.27	1		
LodgK	0.47	0.51	–0.03	–0.05	0.27	0.32	**–0.77***	–0.39	–0.54	–0.24	1	
LodgT	**0.58***	**0.66***	–0.24	–0.27	0.07	0.26	–0.57	–0.19	**–0.60***	0.12	**0.60***	1

*Significant at the 0.05 probability level

variates are highly correlated (r = -0.65 in Table 4). This variate is followed by an environmental one, CARrkn. Then a genotypic one is incorporated: powdery mildew during the formation of grain number (PMK). Two other environmental covariates are added: CARrtkw and SOIrtkw. The other environmental covariates are not significant. Lodging during grain filling is the last genotypic covariate included (LodgT). Lodging during the formation of grain number could have been introduced instead since susceptibilities to early or late lodging are correlated (r = 0.60), and a similar situation is obtained for the susceptibility to powdery mildew (r = 0.89).

From the four probe genotypes, ‘Camp-Rémy’ (CAR) reveals the maximum of interaction. In a previous paper, we noticed indeed that ‘Camp-Rémy’ was the most interactive probe for the formation of yield. It was stable for grain yield but highly unstable for the formation of grain number and grain filling. This model is efficient because it partitions 65% of the interaction term with a moderate consumption of degrees of freedom (39%).

Regression on either genotype (model 2) or environment (model 3) covariates also explains a part of the corresponding main effects, 30% and 75%, respectively, but in both cases the rest of the main effect remains significant. Therefore, during the whole experiment, the variability of environments could be well accounted for by the behavior of the probe genotypes ‘Camp-Rémy’ and ‘Soissons’.

Table 5 Regression coefficient estimates for models 2, 3 and 4. Case of probe indicators as environmental covariates. Genotypic regression parameters of model 2 (in the first column), environmental regression parameters of model 3 (in the first line) and the nine cofactors of covariate products determined in model 4 (in the rest of the table). Codes of the covariates are the same as in Table 3

Genotype	Environmental component			
component	Main	CARrkn	CARrtkw	SOIrtkw
Main	μ=67.16	δ_1=–0.49*	δ_2=–2.28*	δ_3=–5.85*
HD	ρ_1=–0.37*	θ_{11}=*0.60*	θ_{12}=–2.62*	θ_{13}=*1.00*
PMK	ρ_2=–0.90*	θ_{21}=*0.97*	θ_{22}=0.65*	θ_{23}=–1.95*
LogdT	ρ_3=–1.53*	θ_{31}=*0.03*	θ_{32}=*0.43*	θ_{33}=*–0.33*

* Significant at the 0.05 probability level

Analysis of parameters from models 2 and 3 ($\alpha_g = \sum_k \rho_k \cdot G_{gk}$ and $\beta_e = \sum_h \delta_h \cdot H_{eh}$)

The regression coefficients of main effects (models 2 and 3) are presented in Table 5. Yield is limited by all of the environmental covariates as they show negative slopes. The varieties exhibit yield decrease when 'Camp-Rémy' is particularly affected during the formation of grain number (CARrkn) or during the grain filling period (CARrtkw) or when 'Soissons' is affected during the grain filling period (SOIrtkw). The same situation is observed for all genotypic covariates: the best yields are obtained with early varieties and varieties resistant to powdery mildew or to lodging. During the experiment, an advantage is observed for the earliest genotypes.

Analysis of genotypic parameters from model 6 ($\alpha\beta_{ge} = \sum_h \alpha''_{gh} \cdot E_{eh}$)

Figure 2 depicts genotypic parameters (on the ordinate) and genotypic main effects (on the abscissa) when deviations in the yield component measured on the probe genotypes are used as environmental covariates. Contrasts are noticeable for grain yield: 'Soissons' obtains the best average, while 'Camp-Rémy' yields poorly. Negative parameters or slopes (on the ordinate) correspond to genotypes whose yield decreases when conditions are poor: for 'Camp-Rémy' during the formation of grain number (top plot), for 'Camp-Rémy' during the grain filling period (middle plot) or for 'Soissons' during the grain filling (bottom plot). In the top figure, 'Apollo' and 'Renan' are opposite to 'Camp-Rémy', 'Baroudeur', 'Rossini' and 'Soissons'. When conditions are poor for Camp-Rémy during the formation of grain number, most varieties, including 'Camp-Rémy' itself, react badly, except for 'Apollo' and 'Renan'. 'Camp-Rémy' and 'Baroudeur' are the most affected.

When genotypic covariates are added to the environmental ones, such as in model 4, it is possible to relate the particular behavior of these genotypes to genotypic traits. This will be next considered.

Analysis of parameters from model 4 ($\alpha\beta_{eg} = \sum_{hk} G_{gk} \cdot \theta_{kh} \cdot E_{eh}$)

For better understanding the meaning of the interaction accounted for by the significant cross-products in Table 3, Fig. 3 gives a summary of the signs of the centered covariates (left part) with respect to the signs of the cross-products (right part). For earliness, date of heading is later as genotypes flower later. Then, after centering, early genotypes score negatively, while late ones score positively. About the environmental covariates, positive values correspond to high values in the reductions of yield components (CARrtkw, for instance), whereas negative ones are related to low values. Combining the signs of the cross-products with those of the parameters θ_{kh} gives a first interpretation of model 4 (bottom part of Fig. 3).

Yield decreases when varieties are late and subject to environments not favorable for 'Camp-Rémy' during grain filling. Early varieties behave better because they escape the bad conditions of 'Camp-Rémy'. Differences can be observed with the susceptibility to powdery mildew according to the conditions of the environments during the grain filling period. Environments not favorable for 'Camp-Rémy' during the grain filling are not favorable for the resistant genotypes. The opposite result is obtained in environments not favorable for 'Soissons'.

According to these additional results, the previous difference between 'Viking' and 'Récital' (middle plot in Fig. 2) can be explained by their differences in earliness, as the interaction HD*CARrtkw is significant (Tables 3 and 5). Indeed, 'Viking' flowered 10 days after 'Récital' on average (Table 1). 'Viking' yields poorly when conditions are unfavorable for 'Camp-Rémy' during grain filling. As 'Viking' and 'Camp-Rémy' have the same earliness, they have been subjected to the same climatic conditions. 'Viking' may be a better probe genotype than 'Camp-Rémy' itself because of its extreme behavior. 'Récital' escaped from these situations because of its earliness. Very similar conclusions could be addressed to 'Thésée' and 'Talent' which are also early varieties. The bottom plot of Fig. 3 illustrates coefficients obtained with the covariate SOIrtkw. Two main features can be drawn: a very poor behavior is observed for 'Récital' while 'Soissons' or even 'Camp-Rémy' are opposite to it. Under such conditions, sus-

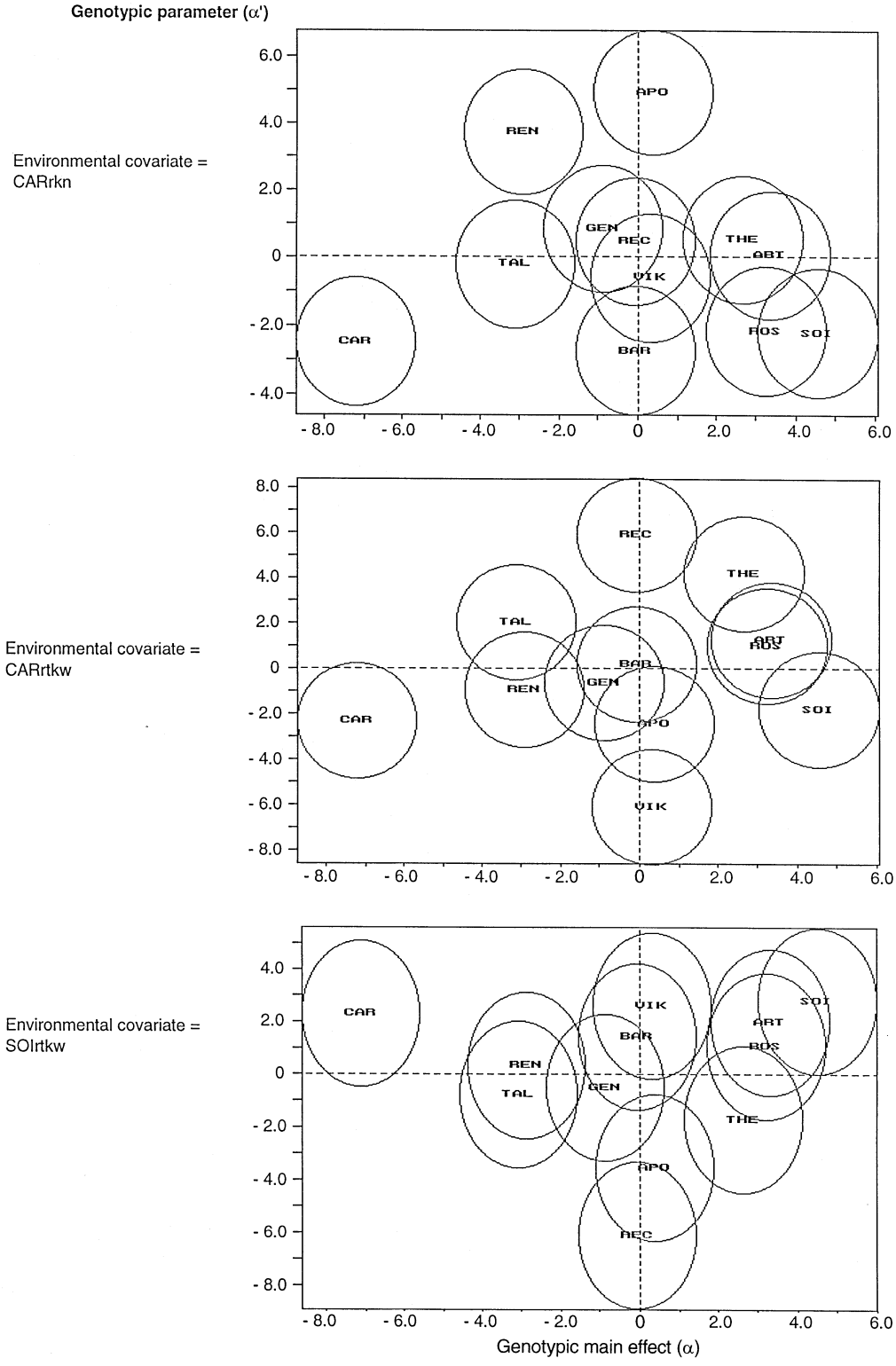

Fig. 2 Genotypic parameters (α') determined with probe indicators as environmental covariates in model 6. Variability of the estimates is represented by confidence ellipses determined at the 0.05 probability level

ceptibility to powdery mildew is regarded as a significant interaction that exists between PMK and SOIrtkw (Tables 3 and 5). In comparison to 'Soissons', 'Récital' was much more susceptible to an early infection of powdery mildew (Table 1). 'Soissons' behaves well even under poor conditions, and this can be attributed to its high resistance to an early infection of powdery mildew. This may explain its very good general performance.

The specific behavior of 'Apollo' or 'Renan' (top plot in Fig. 2) can not be related to any large degree to earliness at heading, to susceptibility to powdery mildew or to susceptibility to lodging as no interaction is observed between these covariates and CARrkn (Tables 3 and 5).

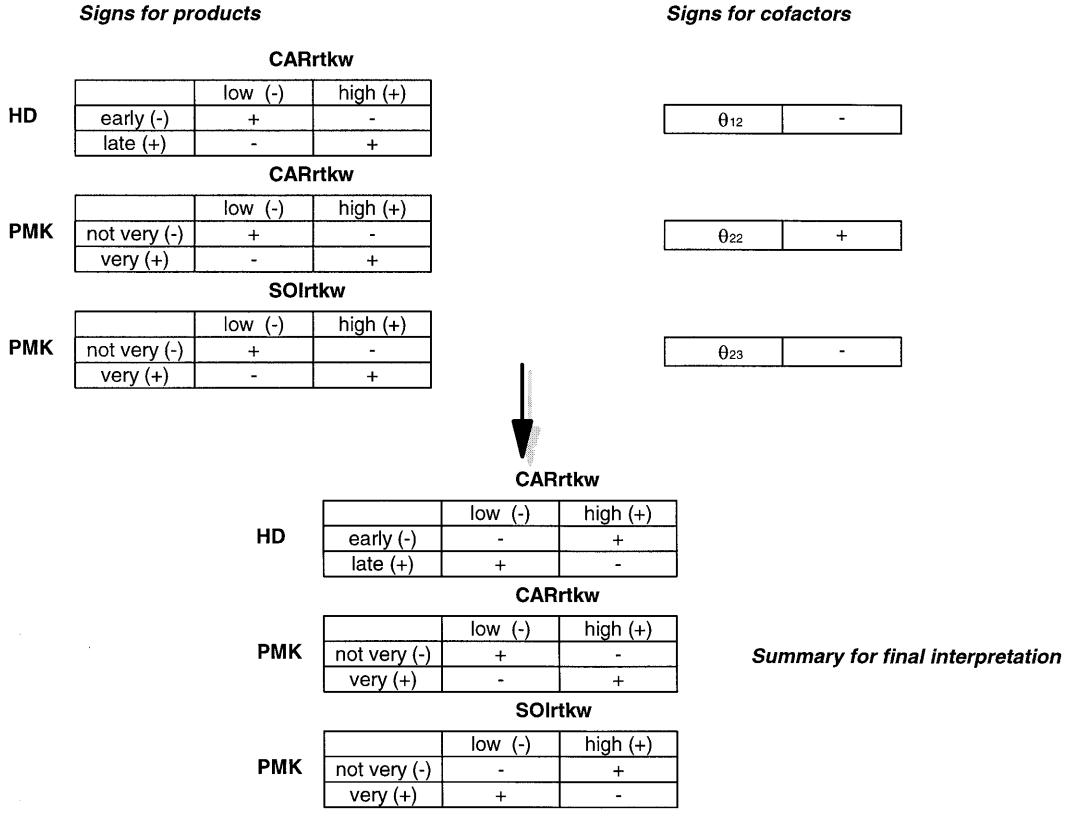

Fig. 3 Signs for products between centred covariates (*left*), and the corresponding significant coefficients θ (*right*) determined from model 4. Summary for final interpretation *below*. Case of probe indicators as environmental covariates

For these genotypes, other covariates might be involved as not all of the interaction is accounted for by the present model.

Analysis of environmental parameters from model 5 ($\alpha\beta_{ge} = \sum_{hk} G_{gk} \cdot \theta_{kh} \cdot E_{eh} + \sum_{h} \alpha'_{gh} \cdot E_{eh} + \sum_{k} \beta'_{ek} \cdot G_{gk}$)

Some environments are favorable to late varieties (Fig. 4): the four conditions at Mons and 91MIN-F. Ondes is favorable to early varieties in 1991. It is not the case in 1992 (92OND-F is hidden by 92ONDIN). Rennes is particularly suitable to early lines in 1991 (treatment IN) and, to a lesser extent, in 1992 (both treatments). Excepting 91MON-F, 92MON-F, 91MININ and 91RENIN, a slight tendency can be observed: yields increase when environments are favorable to late varieties. For powdery mildew, environments with negative coefficients correspond mostly to environments which did not receive fungicides in contrast to the others which display positive coefficients: yield decreases when susceptibility to powdery mildew increases. The reverse is obtained when fungicides are applied. Nevertheless, Rennes is an exception as powdery mildew was also observed on protected plots. These results indicate that powdery mildew infection may have occurred at an early stage in most of the trials during the study. In most of the unprotected environments, such an early observation is missing, and this result shows that it would be useful to record it more regularly. Similar conclusions could be drawn with an observation of powdery mildew during grain filling because of the high correlation between PMK and PMT. The bottom plot depicts coefficients obtained with susceptibility to lodging. The situation of Rennes and La Minière is of particular interest as severe lodging occurred in those locations during the experiment. The pressure was higher in 1992 than in 1991. In both locations, yields are higher in 1991 than in 1992 and in environments which received fungicides. However, a contrast can be detected: Rennes is favorable to resistant lines while La Minière is not. This could be attributed to the stage of the plants when lodging occurred: at Rennes, it was observed at an early stage, at La Minière, at a late one.

Analysis with true indicators of yield-limiting factors

Results of the second type of factorial regressions involving indicators of yield-limiting factors are displayed in Table 6. Indicators related to the formation of grain number are superior in the analysis of interaction: water deficits and the ratio between nitrogen absorbed and kernel number. However, this model is less efficient than the previous one because it explains only 57% of the interaction sum of squares with 35% of the degrees of freedom for the interaction term. In this model, only water

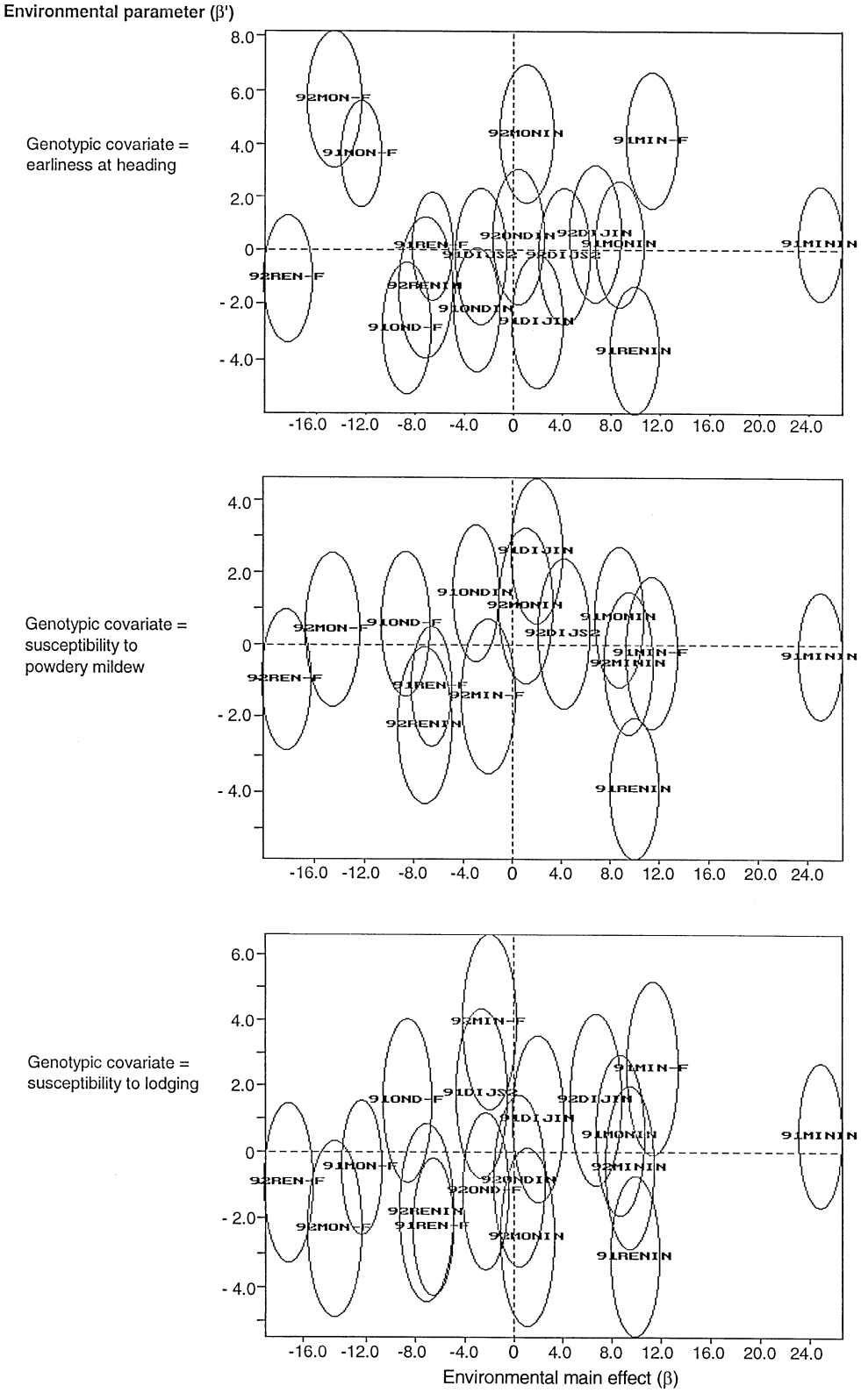

Fig. 4 Environmental parameters (β') determined with genotypic covariates in model 5. Variability of the estimates is represented by confidence ellipses determined at the 0.05 probability level

deficits explains significantly a part of the main environmental effect (results not shown). Some covariate products are significant: PMK*WDK, LodgT*WDK, HD*BK and PMK*BK. The regression coefficients of main environmental effects are presented in Table 7. Yield decreases in the presence of water deficits. No conclusion can be drawn for the second environmental indicator as the corresponding effect is non-significant. For covariate products, all cofactors are negative indicating negative interactions (Fig. 5). Yield decreases in situ-

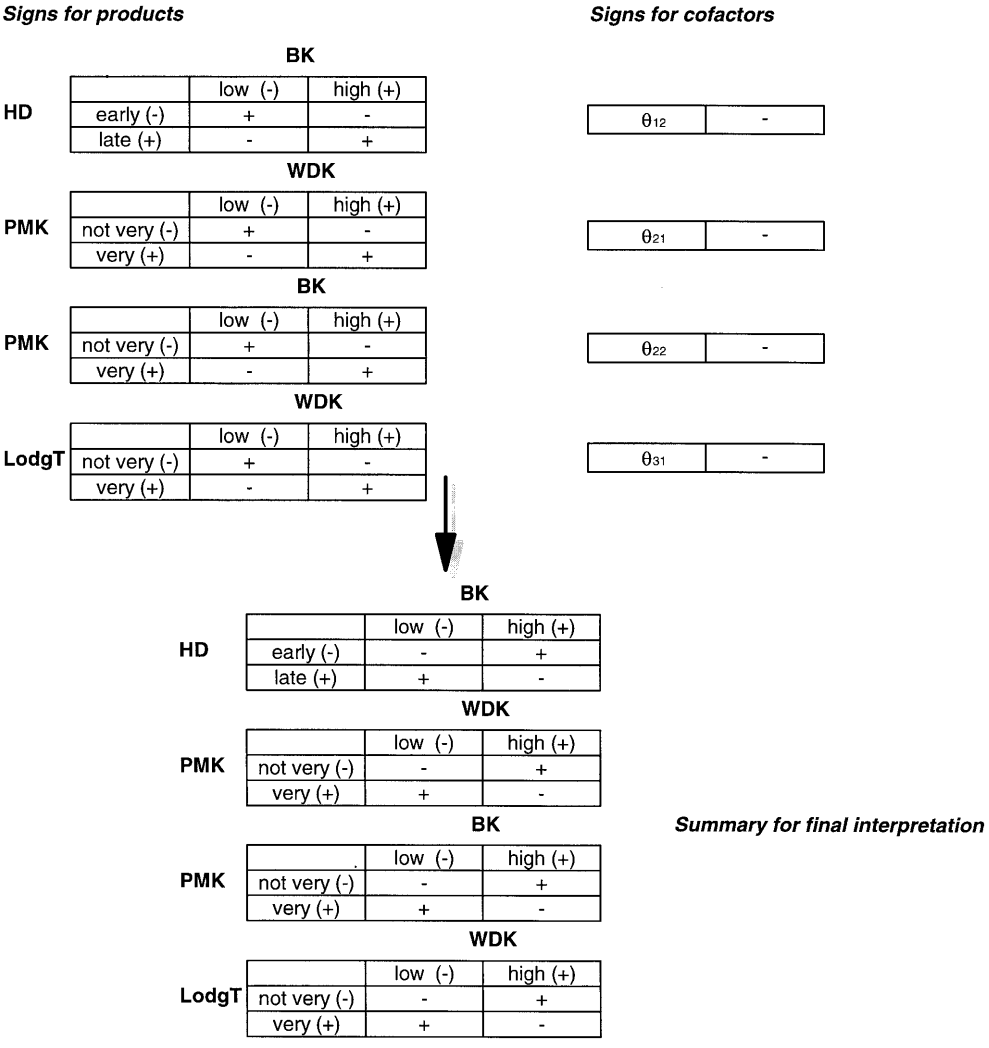

Fig 5 Signs for products between centered covariates (*left*), and the corresponding significant coefficients θ (*right*) determined from model 4. Summary for final interpretation *below*. Case of indicators of yield-limiting factors as environmental covariates

Table 6 Recapitulative table for second factorial regression model. Cae of environmental covariates water deficits (WDK), and nitrogen absorbed by kernel numbr (BK). Earliness at heading (HD), PMK (powdering mildew durig shooting) and LodgT (lodging during grain filling) were used as genotypic covariates. Each cell corresponds to one term of the model with degrees of feedom [in brackets], sum of squares of bold and mean square in italics

*HD*WDK* [1] **3** *3*	*HD*BK* [1] **271** *271**	*HD*β′$_{el}$* [17] **1861** *109**
*PMK*WDK* [1] **114** *114**	*PMK*BK* [1] **214** *214**	*PMK*β′$_{e2}$* [17] **412** *24**
*LodgT*WDK* [1] **42** *42**	*LodgT*BK* [1] **1** *1*	*LodgT*β′$_{e3}$* [17] **408** *24**
*WDK*α$_{g1}$* [8] **456** *57**	*BK*α′$_{g2}$* [8] **229** *29**	Remaining interaction [136] **2981** *22**

*Significant at the 0.05 probability level

Table 7 Regression coefficient estimates for models 2, 3 and 4. Case of indicators of yield-limiting factors as environmental covariates. Genotypic regression parameters of model 2 (in the first column), environmental regression parameters of model 3 (in the first line) and the six cofactors of covariate products determined in model 4 (in the rest of the table). Codes of covariates are the same as in Table 6

Genotype component	Environmental component		
	Main	WDK	BK
Main	μ=67.16	δ_1=–0.63*	*δ_2=–0.30*
HD	ρ_1=–0.36*	*θ_{11}=0.38*	θ_{12}=–1.17*
PMK	ρ_2=–0.91*	θ_{21}=–0.75*	θ_{22}=–1.00*
LodgT	ρ_3=–1.54*	θ_{31}=–0.59*	θ_{32}=0.10

* Significant at the 0.05 probability level

ations where varieties susceptible to powdery mildew or to lodging are tested in dry environments (PMK*WDK = -0.75 and LodgT*WDK = -0.59). In the presence of a high level of nitrogen (high values for BK), similar results are obtained with late varieties (HD*BK = –1.17) or with genotypes susceptible to an early infection of pow-

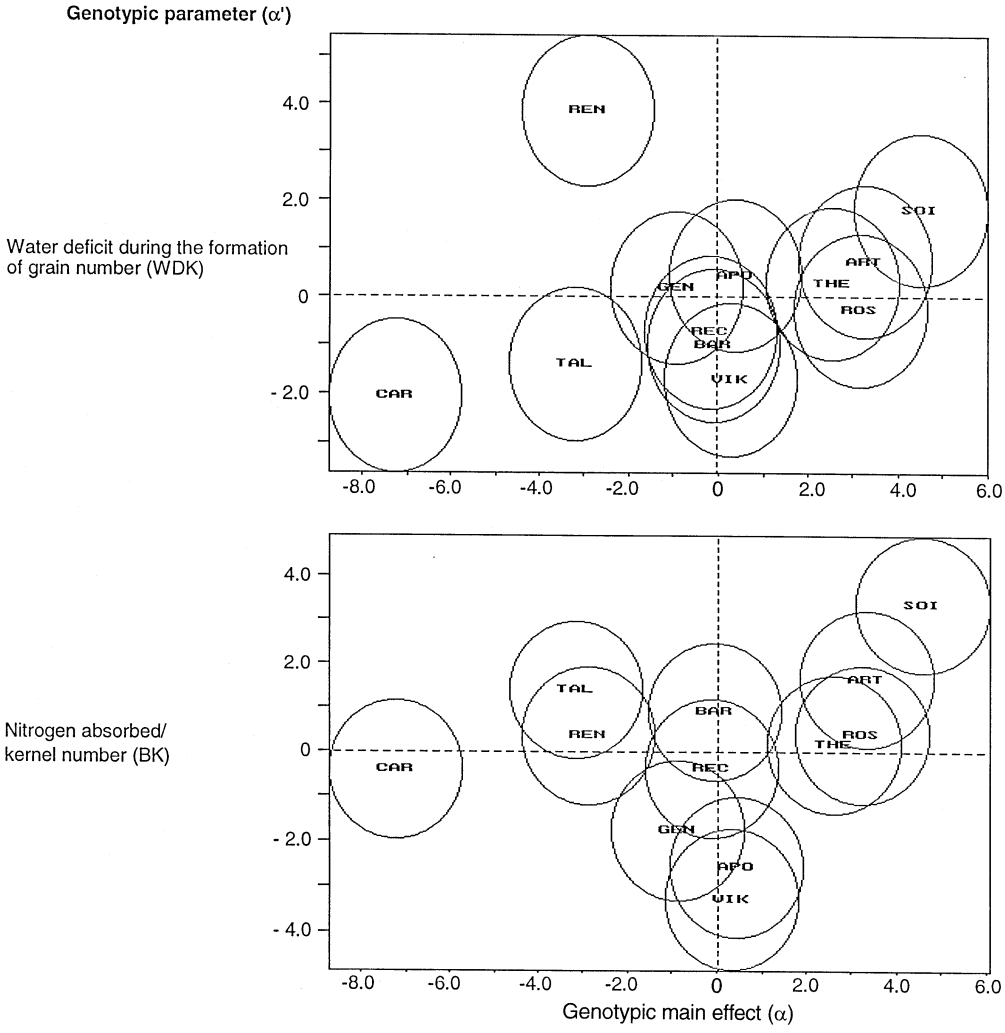

Fig. 6 Genotypic parameters (α') determined with physical indicators as environmental covariates in model 6. Variability of the estimates is represented by confidence ellipses determined at the 0.05 probability level

dery mildew (PMK*BK = –1.00). Figure 6 depicts the sensitivity of the 12 genotypes to water deficits (top plot) and nitrogen absorbed per kernel (bottom plot). In the upper plot, yield can be well related to the sign of the slope: positive values are related to the best yields, while negative ones correspond to poor yields. 'Renan' does not follow this rule as its high positive coefficient is not associated with a good yield. This is owing to the model which accounts for a part of interaction (57%). Under conditions of water deficits during the formation of grain number, this genotype behaves differently from the others. As PMK*WDK is significant, this particular result could be related to the good resistance of 'Renan' to powdery mildew during the formation of grain number during the study. In comparison to the others, this genotype was subject to fewer yield-limiting factors. Another feature is the contrast between 'Soissons' and 'Camp-Rémy'. 'Soissons' is more suited to dry conditions than 'Camp-Rémy'. Very little reaction can be observed for the other genotypes because their values are not significantly different from zero. The nitrogen per kernel, as an indicator of the nitrogen status of the crop, also identifies several contrasting genotypes: 'Soissons' is different from 'Viking' and 'Apollo'. 'Soissons' is well suited to a crop management system involving high levels of nitrogen inputs as its coefficient is positive. During the study, a high level of nitrogen absorbed per kernel were observed at Rennes (both years) and La Minière (in '1992'). The converse is observed with 'Viking' or 'Apollo': an excess of nitrogen does not contribute to increase their performances. These features can be attributed to the earliness of the genotypes (as HD*BK is significant) and to their susceptibility to powdery mildew (as PMK*BK is significant also). 'Apollo' and 'Viking' are late varieties susceptible to an early infection of powdery mildew, while 'Soissons' is early and less susceptible (Table 1).

Discussion and conclusion

The present study succeeded in providing an agronomic explanation of genotype*environment interaction: earliness at heading and susceptibilities to powdery mildew and to lodging were the genotypic covariates of major importance during the experiment. Genotype*environ-

ment interaction could also be related to environment features, measured indirectly by the behavior of the probe genotypes during the formation of yield or described directly by the usual indicators of yield-limiting factors. Both periods of the formation of yield contributed to genotype*environment interaction: the formation of grain number as well as the grain filling period. In a previous paper, we noticed that these two periods influenced grain yield since grain yield was significantly correlated to RKN ($r = –0.63$) and to RTKW ($r = –0.71$) and that no correlation was found between RKN and RTKW (Brancourt-Hulmel et al. 1999). These goals were successfully achieved by combining the results of a simplified crop diagnosis with factorial regressions.

Factorial regression analysis succeeded in explaining 57–65% of the genotype*environment interaction with 35% and 39% degrees of freedom, respectively. The first factor regression model with outputs of a simplified crop diagnosis as environmental covariates allows us to partition the interaction term as well as the environmental main effect. Such a modeling is useful for comparing the results of a given set of genotypes to the performance of probe genotypes. However, these probe genotypes have to be well characterized by a preliminary crop diagnosis. In our analysis, 'Camp-Rémy' was a probe of particular interest. A preliminary study showed that this variety was stable for grain yield but very unstable during the formation of grain number and during grain filling. Our results show that other genotypes like 'Viking' or 'Récital' could also be used as probe genotypes. This method could be helpful to identify probe genotypes for further analyses.

The second round of factorial regressions using the indicators of yield-limiting factors is less convincing. The introduction of more environmental covariates would be necessary in order to explain more interaction than 57% and to more effectively partition the environmental main effect. In our study, we identified only two major environmental covariates: water deficits and level of nitrogen. It was not possible to identify other covariates, such as high temperatures or radiation during several periods, owing to several reasons. The first one is that the environments are subject to numerous yield-limiting factors and the effect of each factor may be small, sometimes too small to be detected. Secondly, the results are also related to how the indicators are defined. For instance, short heat stresses can not be detected during grain filling with indicators computed during the whole period. Improvement is still needed to determine more suitable indicators of yield-limiting factors, and the approach presented here could be helpful in this respect. For example, weather variables defined in the same way as by Feyerherm et al. (1992), who estimated climatic elements (precipitation or temperature) calculated over particular periods of the crop, could improve the analysis itself. Such variables defined over short periods were discarded in our study because of the great number of putative yield-limiting factors. A third reason is that biological variates, i.e. susceptibility to powdery mildew and to lodging, affected yield more than the climatic variates. Finally, it is difficult to introduce and interpret linear factor regression using numerous covariates. In such a network, it is obvious that numerous factors affect yield. Further modeling, such as biadditive regression which combines the AMMI model and factorial regression, has to be investigated in order to take more covariates into account (Denis 1991; van Eeuwijk 1995).

In comparison to previous studies, crop diagnosis makes the choice of the environmental covariates easier and less arbitrary. Biarnès-Dumoulin et al. (1996) found that water balance was also a major component of the genotype*environment interaction in pea, but no comparisons with other environmental covariates were mentioned. In a study of genotype*environment interaction in wheat, Baril (1992) assumed that interaction was essentially due to yield-limiting factors occurring at the grain-filling stage, but the action of these factors could not be confirmed by a crop diagnosis. The present study shows that crop diagnosis has to be combined with the analysis of interaction in order to relate pre-existent limiting factors of yield to the major factors responsible for the interaction. Results indicate that genotypic covariates are more important than environmental ones in the case of susceptibilities to diseases or to lodging owing to a greater variability among genotypes than among environments for these traits. This could be an additional reason for the small number of environmental covariates identified. Such a result was true in the present study, and one can expect other covariates to be significant with another set of genotypes or with another climatic sequence. Concerning susceptibilities to diseases, our results indicated that powdery mildew and, indirectly, septoria (which was correlated to earliness) were superior in explaining interaction. This is partly in agreement with the study of Chevalier-Gérard et al. (1994) who identified that stripe rust, septorias (*Septoria tritici* and *Stagonospora nodorum*) and powdery mildew were of major importance in the estimation of yield losses in a winter wheat network located in the north of France. However, no stripe rust variate was selected as no infection occurred in our study. In another experiment, susceptibility to stripe rust might have been identified. In addition to the results of Chevalier-Gérard et al. (1994), we also showed that these susceptibilities generate genotype*environment interactions. With respect to susceptibility to lodging, Baril (1992) also identified this genotypic covariate in the same wheat breeding network as in our study, but with another set of genotypes.

Other authors have associated environmental variables with genotype*environment interaction using the biadditive model, such as Nachit et al. (1992) in durum wheat. Their study associated morpho-physiological traits and genotypic scores by rank order correlation coefficients and by multiple regression analyses. Similarities were found between both models by van Eeuwijk et al. (1993) who selected the same variables using these two methods in perennial ryegrass. Furthermore, factorial regression provides genotypic and environmental pa-

rameters. The former corresponds to the sensitivity of a genotype to environmental characteristics, while the latter describes the response of an environment to genotype features. Such parameters are very important for the breeder and complement the information given by the ecovalence, defined as the contribution of a genotype or an environment to the interaction (von Wricke 1962).

A stepwise and integrated approach has also been explored by Desclaux (1996) in soybean with water-limited environments. In the present study, we investigated other environmental stresses, since the climatic conditions are more diverse in the wheat breeding network. This is due to the fact that wheat is grown almost everywhere in France and that the growing season is longer than that of soybean. This diversity allows a direct application of the approach to the interpretation of the lines tested in the current wheat breeding network. Moreover, the introduction of outputs of the simplified crop diagnosis gives a better alternative to the use of the more usual indicators for several reasons. Interaction as well as environmental main effects are better partitioned with these covariates. They also offer a more global explanation than the usual indicators. In addition, they are interesting for establishing a variety type based on the behavior of the probe genotypes. It is possible to compare any given set of genotypes to the behavior of probe genotypes that are well characterized during the formation of yield. When probe genotypes are commonly grown cultivars, comparisons become straightforward.

Acknowledgments The author thanks Paul Bataillon, Denis Beghin, Michel Leleu and Claude Sausseau for their technical assistance and is indebted to Christophe Lecomte for providing climatic data. Helpful comments and suggestions from Gilles Charmet, Jean-Baptiste Denis, Gérard Doussinault, Christophe Lecomte, Marie Purcell and from the reviewers were appreciated.

References

Baril CP (1992) Factor regression for interpreting genotype-environment interaction in bread wheat trials. Theor Appl Genet 83:1022–1026

Baril CP, Denis JB, Wustman R, van Eeuwijk FA (1995) Analysing genotype by environment interaction in Dutch potato variety trials using factorial regression. Euphytica 82:149–155

Biarnès-Dumoulin V, Denis JB, Lejeune-Hénaut I, Etévé G (1996) Interpreting yield instability in pea using genotypic and environmental covariates. Crop Sci 36:115–120

Bidinger FR, Hammer GL, Muchow RC (1996) The physiological basis of genotype by environment interaction in crop adaptation. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement, CAB Int, Wallingford, UK, pp 329–347

Brancourt-Hulmel M, Biarnès-Dumoulin V, Denis JB (1997) Points de repère dans l'analyse de la stabilité et de l'interaction génotype-milieu en amélioration des plantes. Agronomie 17:219–246

Brancourt-Hulmel M, Lecomte C, Meynard JM (1999) A diagnosis of yield-limiting factors on probe genotypes for characterizing environments in winter wheat trials. Crop Sci 39 (in press)

Charmet G, Balfourier F, Ravel C, Denis JB (1993) Genotype x environment interactions in a core collection of French perennial ryegrass populations. Theor Appl Genet 86:731–736

Chevalier-Gérard C, Denis JB, Meynard JM (1994) Perte de rendement due aux maladies cryptogamiques sur blé tendre d'hiver. Construction et validation d'un modèle de l'effet du système de culture. Agronomie 14:305–318

Cooper M, Fox PN (1996) Environmental characterization based on probe and reference genotypes. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement, CAB Int, Wallingford, UK, pp 529–547

Decoux G, Denis JB (1991) INTERA. Logiciels pour l'interprétation statistique de l'interaction entre deux facteurs. Laboratoire de Biométrie, INRA, Route de Saint-Cyr, 78026 Versailles Cédex, France

Denis JB (1980) Analyse de régression factorielle. Biom Praxim 20:1–34

Denis JB (1988) Two way analysis using covariates. Statistics 19:123–132

Denis JB (1991) Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. Rev Stat Appl 34:5–24

Desclaux D (1996) De l'intérêt de génotypes révélateurs de facteurs limitants dans l'analyse des interactions génotype*milieu chez le soja (*Glycine max.* L. Merill). PhD thesis, Institut national polytechnique de Toulouse, Toulouse, France

Feyerherm AM, Sears RG, Paulsen GM (1992) Utilizing genotype-by-weather interactions for selecting wheat cultivars. Agron J 84:391–394

Fox PN, Rosielle AA (1982) Reference sets of genotypes and selection for yield in unpredictable environments. Crop Sci 22:1171–1175

Hammer GL, Butler DG, Muchow RC, Meinke H (1996) Integrating physiological understanding and plant breeding via crop modelling and optimization. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement, CAB Int, Wallingford, UK, pp 419–441

Henderson SA, Fukai S, Jongdee B, Cooper M (1996) Comparing simulation and experimental approaches to analysing genotype by environment interactions for yield in rainfed lowland rice. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement, CAB Int, Wallingford, UK, pp 443–464

Nachit MM, Sorrells ME, Zobel RW, Gauch HG, Fischer RA, Coffman WR (1992) Association of environmental variables with sites' mean grain yield and components of genotype-environment interaction in durum wheat. II J Genet Breed 46:369–372

SAS Institute Inc. (1989) SAS/STAT user's guide, Version 6. SAS Institute, Cary, N.C.

Sebillotte M (1980) An analysis of yield elaboration in wheat. In: Wheat technical monograph. CIBA-GEIGY Bâle 25–32

van Eeuwijk FA (1995) Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. Euphytica 84:1–7

van Eeuwijk FA, Elgersma A (1993) Incorporating environmental information in an analysis of genotype by environment inteaction for seed yield in perennial ryegrass. Heredity 70:447–457

van Eeuwijk FA, Keizer LCP, Bakker SS (1995) Linear and bilinear models for the analysis of multi-environments trials: II. An application to data from the Dutch Maize Variety Trials. Euphytica 84:9–22

von Wricke G (1962) Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. Z Planzenzuecht 47:92–96

Wade LJ, McLaren CG, Samson BK, Regmi KR, Sarkarung S (1996) The importance of environmental characterization for understanding genotype by environment interactions In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement, CAB Int, Wallingford, UK, pp 549–562