

SELECTION VARIETALE ET MILIEU Sélection pour l'adaptation au milieu et prise en compte des interactions génotype/milieu

Selecting varieties for the environment Selecting the right varieties for the environment and taking account of genotype/environment interactions

Oléagineux, Corps Gras, Lipides. Volume 7, Numéro 6, 504-11, Novembre - Décembre 2000, Dossier : "Agriculture, recherche et territoire"

Auteur(s) : Maryse Brancourt-Hulmel, Unité de génétique et d'amélioration des plantes, Inra, 80200 Estrées-Mons, France.

Résumé : L'adaptation au milieu est un objectif de sélection recherché pour un grand nombre d'espèces végétales et elle fait le plus souvent appel à l'analyse du rendement. L'améliorateur peut rechercher des génotypes présentant une « adaptation spécifique », c'est-à-dire une adaptation à des milieux spécifiques, ou au contraire une « adaptation générale » à des conditions de milieux variés [1]. L'adaptation spécifique pourra être obtenue pour des stress particuliers, observés en l'occurrence dans des milieux particuliers : citons, par exemple, l'adaptation du maïs à des froids printaniers dans les régions françaises septentrionales, l'adaptation du blé tendre d'hiver à une alimentation azotée sub-optimale, la tolérance de l'orge à la mosaïque modérée, etc. L'adaptation générale, parfois appelée adaptabilité, est conférée par une adaptation simultanée à un ensemble de contraintes du milieu, telles que le froid, la sécheresse, le manque d'eau, le manque ou l'excès d'azote, les maladies, etc. C'est en quelque sorte une somme d'adaptations spécifiques. Mais le nombre de contraintes du milieu est tel qu'il est difficile de les étudier toutes. Il faudrait, en effet, des dispositifs factoriels très lourds à mettre en place car nécessitant l'étude d'un grand nombre de facteurs à la fois, avec toutes les combinaisons entre facteurs. Les conditions naturelles sont, de surcroît, difficiles à reproduire en enceintes contrôlées. Ainsi, l'adaptation générale s'observe le plus souvent en conditions naturelles dans des réseaux d'expérimentation regroupant un ensemble de milieux sur plusieurs années, les « réseaux multilocaux et pluriannuels ». La notion d'adaptation est à replacer dans le contexte des interactions génotype/milieu car des variations d'adaptation se traduisent par des interactions génotype/milieu. Lorsque plusieurs génotypes sont étudiés dans plusieurs milieux, le caractère considéré, par exemple le rendement, présente deux sources de variation : celle due aux génotypes, appelée « effet du génotype », et celle due aux milieux, appelée « effet du milieu ». Pour des caractères agronomiques complexes tels que le rendement, la teneur en protéines, etc., ces deux sources sont rarement additives.

En effet, le cas le plus général est celui où il existe un écart, une déviation entre la valeur observée et celle prévue par l'additivité des effets du milieu et des effets du génotype, ce qu'on appelle communément écart au modèle additif. Graphiquement, l'additivité se traduit par un parallélisme des réactions entre génotypes d'un milieu à un autre (figure 1A) tandis que l'interaction se manifeste par un non-parallélisme de ces réactions. On parle d'interaction quantitative (figure 1B) lorsque les classements des génotypes sont conservés entre les lieux et d'interaction qualitative (figure 1C) dans le cas d'une inversion de classement [2]. C'est une notion relative qui dépend à la fois des génotypes et des milieux considérés. Étudier l'adaptation d'un génotype revient donc à analyser l'interaction génotype/milieu.

Summary : Genotype-environment interaction is the expression of differential genotypic adaptation across environments. Statistical and agro-physiological techniques for quantifying genotype adaptation and for characterizing environments are presented. This will be illustrated by examples from wheat, focusing upon analyses of two-way tables of genotype-environment interaction and upon several regression models. Adaptation is caused by "genes of adaptation" or by the genetic structure. Several key aspects of a breeding programme are discussed in relation with genotype-environment interaction.

Keywords : genotype-environment interaction, adaptation, statistical techniques, genetics, breeding programme.

ARTICLE

Les principaux modèles statistiques pour analyser l'adaptation au milieu

Les méthodes statistiques les plus utilisées pour estimer l'adaptation sont fondées sur des analyses de la variance dérivées de la régression. L'adaptation générale pourra être mesurée par un modèle très simple, la régression conjointe, tandis que l'adaptation spécifique pourra faire appel à des modèles plus ou moins sophistiqués (régression conjointe, régression factorielle et régression factorielle biadditive) selon le nombre de facteurs ou contraintes du milieu à considérer. Notons aussi qu'il existe d'autres modèles pour décomposer l'interaction ([3-5] pour ne citer que les revues les plus récentes).

Régression conjointe

La régression conjointe, une méthode rendue célèbre par Finlay et Wilkinson [6], est l'outil le plus employé pour analyser l'interaction génotype/milieu du rendement chez de nombreuses espèces. Cette technique permet de décomposer l'interaction par une régression simple. Le plus souvent, le régresseur est « la potentialité » du milieu, estimée par l'effet moyen du milieu. Le coefficient de régression ou la pente renseigne sur l'adaptation ou l'adaptabilité des génotypes et permet de les classer en plusieurs catégories : les génotypes valorisant bien les milieux favorables (coefficients positifs), les génotypes bien adaptés à des milieux défavorables (coefficients négatifs) et les génotypes plus ou moins bien adaptés à un ensemble de milieux (coefficients proches de zéro).

De façon équivalente, les coefficients de régression précédents, ρ_{g_i} , peuvent être remplacés par $1 + \rho_{g_i}$ pour répartir les génotypes selon les mêmes catégories (respectivement avec des valeurs > 1 , ~ 1 et < 1). Il existe de nombreuses variantes pour cette modélisation tant au niveau des modèles statistiques qu'au niveau de son utilisation [5].

L'utilisation de ce modèle peut être illustrée à l'aide d'un exemple comparant l'adaptation de dix génotypes de blé tendre d'hiver qui ont été testés dans le réseau INRA. Deux variables ont été mesurées : le rendement en grain et la teneur en protéines des grains à la récolte. L'analyse des pentes uniquement permet de distinguer trois groupes de génotypes tandis que l'analyse simultanée des pentes et des performances moyennes en fait apparaître quatre (*tableau*) : les génotypes peu adaptés avec une pente voisine de 1 et une moyenne faible (cas de Récital pour le rendement), les génotypes bien adaptés avec une pente voisine de 1 et une bonne performance moyenne (cas de Soissons, Isengrain, Oratorio et Rumba), les génotypes adaptés aux milieux favorables (Hyno-Précia) et les génotypes adaptés aux milieux défavorables (Arche, Camp-Rémy, DI9714 et Renan). Le même type d'analyse peut s'effectuer pour d'autres caractères, par exemple la teneur en protéines (*tableau*).

Régression factorielle

La régression factorielle, une méthode également dérivée de la régression, permet d'analyser l'adaptation spécifique. Pour décomposer l'interaction mais aussi les effets principaux (effet du milieu et effet du génotype), ce modèle peut intégrer plusieurs régresseurs qui sont des covariables décrivant les génotypes et/ou les milieux [7]. Les covariables du génotype, telles que la précocité, la résistance aux maladies, à la verse, etc., permettent d'identifier les origines génétiques des interactions génotype/milieu. Les covariables du milieu peuvent concerner le climat, le sol, les pressions de maladies, l'état nutritionnel de la culture, etc. Selon leur nature, il sera donc possible de préciser en quoi les génotypes sont plus ou moins bien adaptés.

Comme pour la régression conjointe, les génotypes bien adaptés présenteront des coefficients positifs, les génotypes mal adaptés des coefficients négatifs et les génotypes indifférents auront des coefficients nuls. Illustrons ces pentes par un exemple de douze variétés de blé testées dans un réseau multilocal et pluriannuel. Les milieux de ce réseau ont été caractérisés par différentes variables dont, par exemple, une variable décrivant la disponibilité en azote du milieu (le critère utilisé pour cette mesure étant la quantité d'azote absorbée par grain, quantité mesurée à la récolte). La *figure 2* représente, en ordonnée, les pentes de régression factorielle, avec comme régresseur du rendement la covariable « disponibilité en azote », et, en abscisse, les rendements moyens des génotypes sur l'ensemble des milieux ou plus exactement les effets des génotypes (rendements moyens centrés). Les ellipses (au seuil $\alpha = 5\%$) représentent la variabilité des paramètres portés en ordonnée (pentes) et en abscisse (effets). Plus les ellipses se superposent et plus les génotypes ont des comportements similaires. La variété Soissons (symbolisée par SOI) est un génotype répondant bien à des milieux riches en azote tandis que des génotypes comme Apollo (APO) ou Viking (VIK) répondent négativement : leur rendement chute dans les situations où l'azote, probablement en excès pour ces génotypes, a provoqué leur verse.

Régression factorielle biadditive

Pour des caractères complexes comme le rendement, cette situation de réaction à un seul facteur du milieu est très peu fréquente. Les modèles à analyser sont en fait un peu plus complexes qu'une régression à une covariable et font donc intervenir plusieurs covariables à la fois. Bien que la régression factorielle permette d'intégrer plusieurs covariables à la fois, il devient difficile cependant d'interpréter tous les paramètres produits (en l'occurrence les pentes) lorsque les covariables sont trop nombreuses. Une autre méthode, la régression factorielle biadditive [9], consiste à introduire les covariables du milieu non pas une par une comme en régression factorielle, mais sous forme de combinaisons linéaires afin d'en réduire le nombre.

La régression factorielle biadditive est une généralisation de la régression factorielle et du modèle multiplicatif, parfois appelé AMMI [10]. Comme pour le modèle multiplicatif, ce modèle décompose l'interaction en une somme de termes multiplicatifs, chaque terme étant la *multiplication* d'une fonction des génotypes (ou ensemble de paramètres génotypiques décrivant l'interaction) avec une fonction des milieux. Il diffère du modèle multiplicatif par les fonctions elles-mêmes qui correspondent à des combinaisons linéaires des covariables du génotype ou du milieu, combinaisons ou « variables synthétiques » qui maximisent la décomposition de l'interaction. À ces modèles est associée la représentation graphique dite du graphique double, plus communément appelée « biplot » [11]. Elle peut être adaptée dans le cas de la régression factorielle biadditive où les génotypes et les covariables des milieux sont représentés simultanément sur un graphique comprenant deux axes correspondant chacun à une variable synthétique du modèle (axe 1 en abscisses pour la première variable synthétique et axe 2 pour la seconde). De façon symétrique, les milieux peuvent être représentés avec les covariables des génotypes [12].

À partir du même exemple que précédemment, illustrons les paramètres de la régression factorielle biadditive à l'aide du « biplot ». Cette fois-ci, les covariables décrivant le milieu sont plus nombreuses : elles concernent la disponibilité en azote, la pression d'oïdium, la somme de rayonnement au moment de la méiose, le déficit hydrique, les fortes températures et la verse. Ces variables sont réparties selon les deux principales phases de l'élaboration du rendement du blé : les variables inscrites en vert influencent l'élaboration du nombre de grains tandis que les variables sur fond gris sont connues pour avoir un effet sur le remplissage des grains (*figure 3*).

L'information contenue dans le graphique est très riche. Commentons le comportement des génotypes. Comme pour le modèle multiplicatif, un génotype proche de l'origine présente une faible interaction - cas de Génial (GEN sur la *figure 3* - tandis qu'un génotype très éloigné est au contraire très interactif - cas de Soissons (SOI), Viking (VIK), Renan (REN) et Apollo (APO) sur fond jaune. Les points ou génotypes qui se superposent ont des comportements similaires tandis que les points éloignés représentent des génotypes au comportement très contrasté. Comme en analyse en composantes principales, les génotypes sont d'autant mieux représentés sur le graphique que la variabilité d'interaction est fortement expliquée par les deux axes. Ici, il faudrait considérer un troisième axe car les deux premiers ne totalisent qu'une partie de l'interaction (environ 60 % de la variabilité).

En quoi les quatre génotypes interactifs diffèrent-ils ? L'analyse graphique des covariables du milieu permet de retrouver les coefficients des régressions factorielles en projetant les lignes noires (symbolisant les génotypes) sur les lignes fléchées (symbolisant les variables du milieu). Avec la

covariable « disponibilité en azote », on peut expliquer le comportement de trois génotypes interactifs : Soissons (SOI) a une pente positive par rapport à Apollo (APO) ou Viking (VIK) qui montrent une pente négative. Quant à Renan (REN), il possède les plus grands coefficients positifs vis-à-vis de fortes températures pendant le remplissage, du rayonnement pendant la méiose ou de la pression d'oïdium avant la floraison alors que ses coefficients sont négatifs pour le déficit hydrique (pendant l'élaboration du nombre de grains ou pendant le remplissage).

Toutefois, les pentes ne sont pas rigoureusement identiques à celles observées dans l'exemple précédent avec la régression factorielle classique. Il existe deux raisons à cela : une raison liée à la méthode elle-même, car la régression factorielle biadditive revient à une approximation des coefficients de la régression factorielle, et une raison liée à l'exemple, car les milieux n'étaient pas tout à fait identiques entre les deux cas présentés. Comme pour bien d'autres outils statistiques, le problème posé est alors celui de la représentativité des paramètres (pentes dans les exemples cités, y compris ceux de la régression conjointe) au cours du processus d'évaluation multilocale et pluriannuelle. Afin de les utiliser efficacement en sélection, il faut en effet s'assurer que les estimations des paramètres seront comparables au fil des années et d'un ensemble de milieux à un autre.

Mieux connaître le milieu pour comprendre l'adaptation au milieu

À l'aide des exemples précédents, il est aisé de constater qu'il est important de connaître les milieux pour comprendre le comportement des génotypes évalués dans un réseau et valoriser leurs différences de comportement. Il paraît tout naturel au sélectionneur et au généticien de caractériser génétiquement le matériel végétal et de nombreux progrès sont actuellement réalisés grâce à la biologie moléculaire. En revanche, la caractérisation du milieu ne bénéficie pas d'autant d'efforts malgré la présence d'effet du milieu souvent beaucoup plus important que celui du génotype (par exemple sur blé [13], sur ray-grass [14], sur maïs [15] ou sur pois [16]) et malgré les progrès réalisés dans ce domaine (accès automatisé aux données météorologiques, développement continu de modèles agronomiques, etc.). Parmi les démarches proposées pour caractériser le milieu, une méthode s'appuie sur l'idée de restreindre à un petit nombre de génotypes les observations nécessaires à la caractérisation des milieux en termes de facteurs limitant le rendement [17, 18]. Ces génotypes, les « génotypes révélateurs », sont choisis *a priori* sur la base de leur précocité et de leurs comportements particuliers et complémentaires vis-à-vis d'un ou de plusieurs facteurs limitants du milieu pour en révéler l'apparition et l'intensité. La méthode, fondée sur le principe du diagnostic agronomique [19, 20], consiste à décomposer l'élaboration du rendement en caractères plus simples et à les comparer à des valeurs potentielles obtenues en conditions de milieu non limitantes.

Pour le blé tendre, le rendement en grain est constitué du produit de deux composantes : le nombre de grains par m² et le poids de mille grains. Les écarts ou déviations de ces composantes à des seuils (*NG seuil* pour le nombre de grains par m² et *PMG potentiel* pour le poids de mille grains) constituent des indicateurs du caractère plus ou moins favorable du milieu pendant leur phase de formation (*figure 4*). Ainsi, l'écart au nombre de grains seuil (*ENG* sur la *figure 4*) décrit les conditions du milieu pendant l'élaboration du nombre de grains tandis que la réduction du poids de mille grains par rapport au poids de mille grains potentiel (*RPMG* sur la *figure 4*) les décrit pendant le remplissage.

Les conditions du milieu peuvent être non limitantes pour tout le cycle de la plante (cas du milieu A sur la *figure 4*), être limitantes pour uniquement l'élaboration du nombre de grains (C) ou le remplissage (B) ou pour les deux phases (cas du milieu D).

Ces déviations de composantes peuvent être estimées sur une variété de blé, par exemple « Thésée » sur la *figure 5*, observée dans un réseau regroupant dix-huit milieux. Cette variété de blé a été expérimentée en tant que « génotype révélateur » de facteurs limitant le rendement. Les conditions ont été non pénalisantes pendant l'élaboration du nombre de grains pour des milieux tels que 91REN-F (voir décodage sur la *figure 5*) et 91RENIN, voire faiblement pénalisantes pour des milieux comme 91MININ. Le poids de mille grains a été plus ou moins fortement pénalisé dans l'ensemble des milieux. La plus faible réduction du poids de mille grains est observée dans le milieu 91DIJIN tandis que la plus forte est enregistrée dans le milieu 91MON-F. Aucun milieu n'a rencontré de situations potentielles pour l'ensemble du cycle tandis que les deux phases ont été pénalisées pour un grand nombre de milieux, comme par exemple 92RENIN et 92MON-F.

L'observation de plusieurs génotypes révélateurs a montré l'existence d'interactions génotype/milieu pour ces déviations de composantes. Il est possible de déterminer les facteurs limitants qui ont généré ces interactions à l'aide de la régression factorielle biadditive incluant plusieurs variables du milieu. L'analyse des corrélations de chaque variable environnementale aux variables synthétiques (qui sont les combinaisons linéaires de toutes les variables environnementales) permet de déterminer celles qui sont à l'origine des interactions. Comme pour l'analyse en composantes principales, les variables proches du cercle sont les plus importantes (exemple de la verse tardive sur la *figure 6*) alors que les covariables proches de l'origine contribuent très peu (exemple de la verse précoce). Ainsi, 75 % de l'interaction est expliquée par six principales variables : la verse tardive, la pression d'oïdium, la disponibilité en azote, la sécheresse et le rayonnement (en début de cycle et à la méiose).

Ainsi, en matière de caractérisation des milieux, la question importante est de déterminer non seulement si des contraintes du milieu sont apparues mais aussi si elles ont eu un impact sur l'élaboration du rendement des génotypes. Cela permet de repérer avec plus de pertinence les variables du milieu qui vont servir à mesurer l'adaptation des génotypes. En outre, le recours à des modèles de simulation tels les modèles dynamiques simulant le fonctionnement des cultures au cours du temps [22, 23] devrait permettre un progrès dans la caractérisation des milieux d'un réseau expérimental, voire dans la simulation ou prédiction des interactions génotype/lieu ou génotype/année.

Origine génétique de l'adaptation au milieu

L'adaptation d'un génotype au milieu peut avoir deux origines : la présence de « gènes d'adaptation », ou plus exactement de gènes qui gouvernent des caractères jouant un rôle dans l'adaptation, et le pouvoir tampon lié à la structure génétique de certains génotypes [24]. L'effet de la structure génétique peut se manifester soit par l'hétérozygotie conférée par l'état hybride, soit par l'hétérogénéité du peuplement cultivé (par exemple, une population, un mélange d'hybrides, une variété multilignée, une variété multiclone, etc.).

Les gènes d'adaptation spécifique sont nombreux : ils peuvent concerner l'évitement de la contrainte (gènes de précocité) ou bien l'adaptation à la contrainte elle-même (résistance au froid hivernal, etc.)

ou à l'agent pathogène. L'adaptation générale à différents milieux peut être contrôlée par des gènes non-allèles différents ou bien, comme dans le cas de la pléiotropie, par des gènes allèles produisant des effets variables (positifs ou négatifs) entre milieux [24]. Les génotypes les plus adaptés seront ceux cumulant un maximum de gènes à effets favorables. Les gènes peuvent être en outre plus ou moins spécifiques comme dans le cas des résistances aux maladies [25].

La supériorité de la structure hétérozygote par rapport à la structure homozygote a été étudiée depuis longtemps chez beaucoup d'espèces. L'effet de l'hétérozygotie sur la stabilité du rendement varie entre les espèces et dépend de leur système de reproduction [26]. L'effet est fort pour des espèces comme le maïs, le seigle ou le tournesol. Il est plus faible pour les espèces partiellement allogames comme le colza, la féverole ou le sorgho. Enfin, il n'a pas été mis en évidence de supériorité de la structure hétérozygote chez des espèces autogames comme le blé [26] ou le pois [27]. La supériorité de l'hybride est due au fait qu'il permet de réunir dans un même génotype un plus grand nombre de gènes, allèles ou non, codominants ou dominants favorables contrôlant l'adaptation à différents milieux [24].

La supériorité des mélanges par rapport aux cultures pures a été montrée pour un grand nombre d'espèces également. Pour reprendre l'exemple des hybrides, la stabilité varie selon leur structure : sur sorgho par exemple, elle est supérieure pour les hybrides doubles (qui sont assimilables à des mélanges d'hybrides simples), intermédiaire pour les hybrides trois-voies et plus faible pour les hybrides simples [28]. Une revue bibliographique comparant plusieurs espèces [26] montre, en outre, que l'effet de l'hétérogénéité varie selon le système de reproduction de l'espèce considérée. Chez les lignées, la supériorité des associations variétales (ou mélanges de lignées) par rapport aux lignées pures est conférée par l'assemblage de génotypes résistants à différentes races d'agents pathogènes [29]. La supériorité des mélanges et des associations est due au fait qu'ils permettent de cumuler des génotypes aux comportements complémentaires.

Les avancées de la biologie moléculaire, notamment la recherche de QTL (*Quantitative Trait Loci* ou régions du génome associées à l'expression « quantitative » de certains caractères agronomiques) et la possibilité nouvelle de séquencer le génome d'une espèce afin d'en faire l'inventaire des gènes, devraient permettre d'identifier des gènes d'adaptation, en particulier les gènes à effets pléiotropiques, jusque-là difficiles à détecter. Dans une revue bibliographique consacrée à l'identification de QTL influencés par le milieu [30], plusieurs travaux mettent en évidence une interaction QTL/milieu. La détection de ces interactions QTL/milieu devrait représenter les effets génétiques responsables de l'interaction génotype/milieu si ces différences sont bien biologiques et ne relèvent pas d'échantillonnage non représentatif de la variabilité génétique.

Prendre en compte l'adaptation en sélection

On peut préconiser la sélection pour l'adaptation spécifique lorsque les interactions sont répétables et la sélection pour l'adaptation générale dans le cas contraire [31]. Quelle que soit l'adaptation recherchée, le sélectionneur peut être amené à optimiser le dispositif expérimental surtout pour les expérimentations en conditions non contrôlées. D'une part, cette optimisation devra considérer des aspects liés au milieu. Le sélectionneur peut en effet devoir choisir un nombre de milieux représentatifs d'une région, d'un pays, etc. Pour établir ce choix, il faut tenir compte de la répétitivité des mesures de l'adaptation, de l'importance des effets dus à l'année et des effets dus aux lieux¹, de la possibilité de réduire le nombre de répétitions. D'autre part, l'optimisation devra tenir compte des

aspects liés à l'hérédité du caractère lui-même (héritabilité). Par exemple, il est nécessaire parfois d'augmenter le pouvoir discriminant des milieux lorsque l'héritabilité des paramètres est faible. La sélection pour l'adaptation pourra aussi s'appuyer sur l'observation de la performance moyenne car des résultats expérimentaux obtenus sur une quinzaine d'espèces montrent une liaison positive entre l'adaptation, mesurée par le coefficient de régression conjointe, et la performance moyenne [32].

Il est couramment établi que l'interaction génotype/année/lieu est plus importante que l'interaction génotype/année, elle-même plus importante que l'interaction génotype/lieu [33]. La présence de l'interaction génotype/année peut être à l'origine du manque de répétitivité des paramètres mesurant la stabilité constaté pour bien des espèces [32]. À volume expérimental constant, une variation du nombre d'années a plus d'influence sur la précision des performances d'un génotype entre milieux qu'un changement du nombre de lieux [33]. Par exemple, trois années d'expérimentation dans trois lieux donnent une meilleure précision qu'une seule année dans neuf lieux. Un équilibre est donc nécessaire entre le nombre d'années et le nombre de lieux. Toutefois, pour raccourcir la durée du processus de sélection, le sélectionneur préfère augmenter de façon importante le nombre de milieux pour compenser l'effet année. Des résultats obtenus sur orge montrent que l'interaction génotype/année est plus faible dans un réseau européen que dans le réseau du Royaume-Uni [33]. Toutefois, la triple interaction génotype/année/lieu reste de même importance dans les deux réseaux. En présence d'interaction génotype/milieu, il est économiquement plus intéressant de réduire les essais à une seule répétition, ce qui permet d'augmenter le nombre de milieux [34]. Ceci rejoint la constatation que la réponse à la sélection n'est pas trop affectée lorsque le nombre de répétitions n'est pas optimal [35]. En outre, la fiabilité des estimations peut dépendre des indicateurs utilisés pour caractériser les facteurs limitants. À l'heure actuelle, des questions importantes sur le choix de ces indicateurs, leur fiabilité, leur gamme de validité ont été très peu explorées dans le cadre de l'analyse des interactions génotype/milieu.

Lorsque le sélectionneur cherche à réduire le nombre de milieux, la question de la représentativité devient primordiale. Une stratégie consiste à échantillonner les milieux de façon à représenter un maximum de diversité pour les interactions observées sur le caractère considéré. Prenons pour exemple la mesure d'un caractère complexe comme la stabilité pour la qualité. Afin de réduire le coût des analyses technologiques, il est nécessaire de limiter le nombre de lieux et de choisir des lieux représentatifs d'une gamme de variation la plus large possible. Des analyses préliminaires sur une partie du matériel végétal (des témoins pour la qualité par exemple) peuvent être envisagées sur des caractères comme la teneur en protéines ou la dureté afin de sous-échantillonner les lieux destinés à fournir les échantillons pour des analyses plus complètes. Le principe consiste à choisir des lieux qui se ressemblent le moins possible du point de vue de l'interaction. La représentation graphique des paramètres multiplicatifs décrivant l'interaction des lieux permet d'identifier quatre groupes de lieux (symbolisés par des quadrants) au sein desquels les lieux se ressemblent, lorsqu'ils sont bien représentés [36]. Il est possible de choisir un lieu par groupe (en gris sur la *figure 7*) et d'appliquer cette méthode à l'échelle d'une année (pour le choix de lieux) ou de plusieurs années (pour le choix de milieux). Une alternative consiste à tenir compte de la représentativité d'une année à l'aide de génotypes de référence testés sur le long terme [37].

Dans le même esprit, les génotypes révélateurs peuvent être utilisés pour identifier des conditions particulières de milieu [31]. Notons aussi qu'il peut être intéressant, dans ce même contexte, de simuler les variations dues aux années, aux lieux ou de façon plus générale aux milieux (voir section précédente).

Des résultats obtenus sur le blé et l'orge dans un réseau allemand montrent que la pente de régression conjointe présente une héritabilité plutôt faible à modérée [28]. L'héritabilité du caractère peut être améliorée en augmentant le nombre de répétitions [38] ou en augmentant le nombre d'années [28], voire en augmentant le pouvoir discriminant des milieux par l'inoculation par les maladies. Dans certains cas, le caractère observé peut être complété par des caractères complémentaires de l'adaptation. Chez le maïs par exemple, la durée entre les floraisons mâle et femelle constitue une mesure indirecte de l'adaptation à la sécheresse [39].

Cette faible héritabilité, qui traduit la forte influence du milieu sur le caractère, a pour conséquence que le sélectionneur ne peut agir que faiblement sur le caractère en début de sélection et qu'il doit procéder à la plus large évaluation multilocale par la suite. Considérer un nombre de milieux important peut en outre permettre de sélectionner pour une large adaptation générale. Même si ce type d'expérimentation est plus coûteux, cela permet de sélectionner dans des conditions extrêmes qui peuvent garantir une pression de sélection vis-à-vis de contraintes qui sont difficiles à reproduire en conditions contrôlées. Vis-à-vis de ces contraintes, il est possible de sélectionner de façon directe ou indirecte. Tout ceci peut conduire à différentes stratégies de sélection. L'évaluation dans les milieux favorables et défavorables peut être simultanée comme dans le cadre d'un programme de sélection pour l'adaptation du blé tendre à la sécheresse [40]. Elle peut être alternée entre les deux types de milieux dans les premières générations de sélection et suivie d'une évaluation multilocale la plus large possible dans les générations suivantes [41]. Une alternative consiste à sélectionner directement dans les milieux se rapprochant le plus des conditions visées [42]. Les conditions visées peuvent de plus être réparties en strates dans lesquelles des milieux représentatifs sont échantillonnés [43]. Chaque strate nécessite, pour sa constitution, de connaître au préalable les relations entre les milieux. Dans ce contexte, il est possible de procéder à une sélection par index en pondérant, par exemple, les milieux en fonction de leur représentativité vis-à-vis des conditions visées par la sélection [43].

Note

¹ Le milieu est considéré au sens large et peut désigner non seulement le lieu mais aussi la combinaison entre le lieu, l'année, le mode cultural.

CONCLUSION

Ainsi la recherche de génotypes adaptés au milieu requiert l'analyse des interactions génotype/milieu. Les méthodes les plus utilisées pour estimer l'adaptation sont fondées sur des analyses statistiques dérivées de la régression. Pour mieux cerner l'expression de la plante dans le milieu, les caractères complexes comme le rendement pourront être décomposés en caractères plus simples. L'origine des interactions est non seulement génétique, due à la présence entre autres de gènes d'adaptation, mais aussi liée aux milieux. Concernant la caractérisation génétique, les avancées de la biologie moléculaire, notamment la recherche de QTL, devraient permettre de progresser dans l'identification de gènes jusque là difficiles à détecter. La caractérisation des milieux, quant à elle, est

indispensable pour savoir si des contraintes sont apparues et si elles ont eu un impact sur le comportement des génotypes, en particulier sur celui des « génotypes révélateurs ». Enfin, la prise en compte de l'adaptation en sélection nécessite d'optimiser le dispositif d'évaluation de façon à ce qu'il soit discriminant et que les estimations de l'adaptation soient fiables et répétables. Notons que la modélisation des interactions génotype/milieu et que la caractérisation des milieux intéressent non seulement l'améliorateur mais aussi le généticien lorsqu'il est confronté à des interactions QTL/milieu. Une approche combinée de ces méthodes devrait l'aider à distinguer les interactions QTL/milieu représentant les effets génétiques responsables de l'interaction génotype/milieu de celles qui relèvent d'un échantillonnage non représentatif de la variabilité génétique. La recherche de l'adaptation demande donc un dialogue continu entre les généticiens, les agronomes et les biométriciens. La recherche de QTL, qui fait appel à de grandes populations de génotypes, va en outre amplifier cette demande vis-à-vis d'une meilleure évaluation phénotypique. Elle nécessite de continuer à adapter les méthodes, tant au niveau statistique qu'au niveau agronomique, à l'étude de grands effectifs pour, par exemple, structurer l'information, mettre en œuvre des dispositifs expérimentaux permettant la prise en compte d'hétérogénéité spatiale, adapter les modèles de fonctionnement des cultures à la caractérisation d'un réseau multilocal et pluriannuel, etc. Enfin, il faut souligner qu'une grande partie de ces conclusions intéresse non seulement le sélectionneur ou le généticien mais aussi l'expérimentateur et le prescripteur de variétés.

Remerciements

L'auteur remercie chaleureusement ses nombreux collègues pour leurs relectures attentives du manuscrit et leurs discussions intéressantes sur le sujet.

REFERENCES

1. GALLAIS A (1992). Adaptation et adaptabilité en amélioration des plantes. *Sel Fr*, 42 : 55-7.
2. ROMAGOSA I, FOX PN (1993). Genotype/environment interaction and adaptation. In : HAYWARD MD, BOSEMARK NO, ROMAGOSA I, eds. *Plant breeding: principles and prospects*. London : Chapman & Hall : 373-90.
3. VAN EEUWIJK FA, DENIS JB, KANG MS (1996). Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In : KANG MS, GAUCH HG, eds. *Genotype-by-environment interaction*. Boca Raton : CRC-Press : 15-49.
4. KANG MS, GAUCH HG (1996). *Genotype-by-environment interaction*. Boca Raton : CRC-Press ; 416 p.
5. 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 : 219-46.
6. FINLAY KW, WILKINSON GN (1963). The analysis of adaptation in a plant-breeding programme. *Aust J Agric Res*, 14 : 742-54.
7. DENIS JB (1980). Analyse de régression factorielle. *Biom Praxim*, 20 : 1-34.

8. BRANCOURT-HULMEL (1999). Crop diagnosis and probe genotypes for interpreting genotype/environment interaction in winter wheat trials. *Theor Appl Genet*, 99 : 1018-30.
9. DENIS JB (1991). Ajustement de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Rev Stat Appliquées*, 34 : 5-24.
10. GAUCH HG (1992). *Statistical analysis of regional yield trials : AMMI analysis of factorial designs*. Amsterdam : Elsevier ; 278 p.
11. KEMPTON RA (1984). The use of biplots in interpreting variety by environment interactions. *J Agric Sci*, 103 : 123-35.
12. BRANCOURT-HULMEL M, DENIS JB, LECOMTE C (2000). Determining environmental covariates which explain genotype/environment interaction in winter wheat through probe genotypes and biadditive factorial regression. *Theor Appl Genet*, 100 : 285-98.
13. BARIL CP (1992). Factor regression for interpreting genotype/environment interaction in bread wheat trials. *Theor Appl Genet*, 83 : 1022-6.
14. CHARMET G, BALFOURIER F, RAVEL C, DENIS JB (1993). Genotype/environment interactions in a core collection of French perennial ryegrass populations. *Theor Appl Genet*, 86 : 731-6.
15. ARGILLIER O, HÉBERT Y, BARRIÈRE Y (1994). Statistical analysis and interpretation of line/environment interaction for biomass yield in maize. *Agronomie*, 14 : 661-72.
16. BIARNÈS-DUMOULIN V, DENIS JB, LEJEUNE-HÉNAUT I, ÉTÉVÉ G (1996). Interpreting yield instability in pea using genotypic and environmental covariates. *Crop Sci*, 36 : 115-20.
17. 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 : 1798-808.
18. DESCLAUX D, HUYNH TT, ROUMET P (2000). Identification of soybean plant characteristics that indicate the timing of drought stress. *Crop Sci*, 40 : 716-22.
19. SÉBILLOTTE M (1980). An analysis of yield elaboration in wheat. In : *Wheat technical monograph*. Bâle : CIBA-GEIGY : 25-32.
20. MEYNARD JM, DAVID G (1992). Diagnostic sur l'élaboration du rendement des cultures. *Cahiers Agricultures*, 1 : 9-19.
21. LETERME P, MANICHON H, ROGER-ESTRADE J (1994). Analyse intégrée des rendements du blé tendre et de leurs causes de variation dans un réseau de parcelles d'agriculteurs du Thymerais. *Agronomie*, 14 : 341-61.
22. BRISSON N, MARY B, RIPOCHE D, JEUFFROY MH, *et al.* (1998). STICS : a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie*, 18 : 311-46.

23. JEUFFROY MH, RECOUS S (1999). Azodyn : a simple model simulating the date of nitrogen deficiency for decision support in wheat fertilisation. *Eur J Agr*, 10 : 129-44.
24. GALLAIS A (1992). Bases génétiques et stratégie de sélection de l'adaptation générale. *Sel Fr*, 42 : 59-78.
25. LEFEBVRE V, PFLIEGER S (2000). L'approche gène candidat pour la caractérisation moléculaire et fonctionnelle de locus de résistance aux parasites chez les plantes. *Sel Fr*, 51 : 31-45.
26. LÉON J (1994). Mating system and the effect of heterogeneity and heterozygosity on phenotypic stability. In : VAN OOIJEN FW, JANSEN J, eds. *Biometrics in plant breeding*. Proc. Meeting of the eucarpia section biometrics in plant breeding 9th, 6-8 july 1994 : 19-31.
27. SARAWAT P, STODDARD FL, MARSHALL DR, ALI SM (1994). Hererosis for yield and related characters in pea. *Euphytica*, 80 : 39-48.
28. BECKER HC, LÉON J (1988). Stability analysis in plant breeding. *Plant Breed*, 101 : 1-23.
29. LANNOU C, POPE C (1997). Mechanisms of variety mixtures for reducing epidemics. In : *Variety mixtures in theory and practice*. Cost Action 817, Inra Grignon France : 206-7.
30. BEAVIS WD, KEIM P (1996). Identification of quantitative trait loci that are affected by environment. In : KANG MS, GAUCH HG, eds. *Genotype-by-environment interaction*. Boca Raton : CRC-Press : 123-49.
31. 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 international : 529-47.
32. WEBER WE, WRICKE G, WESTERMANN T (1996). Selection of genotypes and prediction of performance by analysing genotype-by-environment interactions. In : KANG MS, GAUCH HG, eds. *Genotype-by-environment interaction*. Boca Raton : CRC-Press : 353-71.
33. TALBOT (1997). Resource allocation for selection systems. In : KEMPTON RA, FOX PN, eds. *Statistical methods for plant variety evaluation*. London : Chapman & Hall : 162-74.
34. MOREAU L, LEMARIÉ S, CHARCOSSET A, GALLAIS A (2000). Economic efficiency of one cycle of marker-assisted selection. *Crop Sci*, 40 : 329-37.
35. WRICKE G, WEBER WE (1986). Selection between clones and homozygous lines. In : WRICKE G, WEBER WE, eds. *Quantitative genetics and selection in plant breeding*. Berlin, New York : De Gruyter : 171-94.
36. ROBERT N, DENIS JB (1996). Stability of baking quality in bread wheat using several statistical parameters. *Theor Appl Genet*, 93 : 172-8.
37. FOX PN, ROSIELLE AA (1982). Reference sets of genotypes and selection for yield in unpredictable environments. *Crop Sci*, 22 : 1171-5.

38. BOS I (1983). The optimum number of replications when testing lines or families on a fixed number of plots. *Euphytica*, 32 : 311-8.
39. BOLAÑOS J, EDMEADES GO (1996). The importance of the anthesis-silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research*, 48 : 65-80.
40. RAJARAM S, BRAUN HJ, VAN GINKEL M (1997). CIMMYT's approach to breed for drought resistance. In : TIGERSTEDT PMA, ed. *Adaptation in plant breeding*. The Netherlands : Kluwer Academic Publishers : 161-7.
41. BRAUN HJ, RAJARAM S, VAN GINKEL M (1997). CIMMYT's approach to breeding for wide adaptation. In : TIGERSTEDT PMA, ed. *Adaptation in plant breeding*. The Netherlands : Kluwer Academic Publishers : 197-205.
42. CECCARELLI S (1994). Specific adaptation and breeding for marginal conditions. *Euphytica*, 77 : 205-19.
43. COOPER M, BRENNAN PS, SHEPPARD (1996). A strategy for yield improvement of wheat which accommodates large genotype by environment interactions. In : COOPER M, HAMMER GL, eds. *Plant adaptation and crop improvement*. CAB international : 487-511.

Illustrations

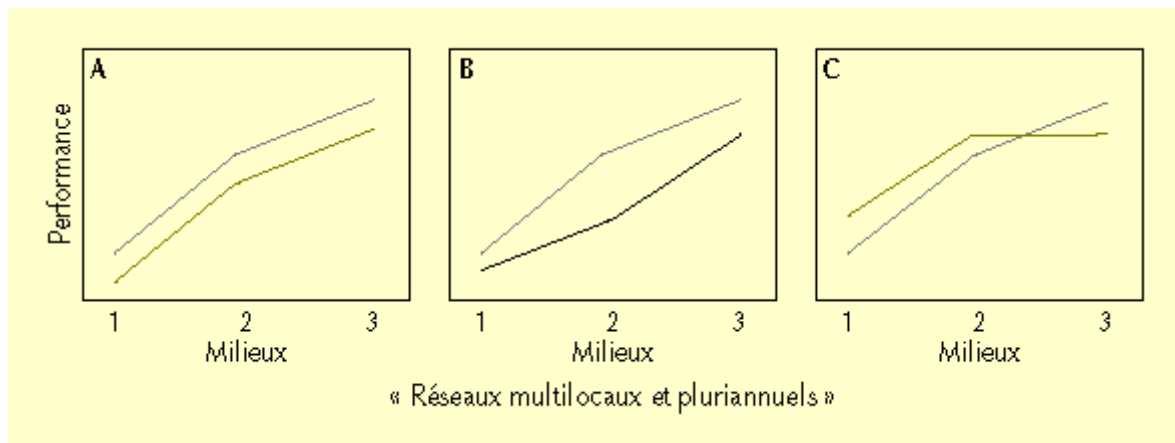


Figure 1. Performance de deux génotypes (l'un est représenté en gris et l'autre par une autre couleur) dans trois environnements (1, 2 et 3) : cas sans interaction (A), avec interaction quantitative (B) et avec interaction qualitative (C) (d'après Romagosa et Fox [2]).

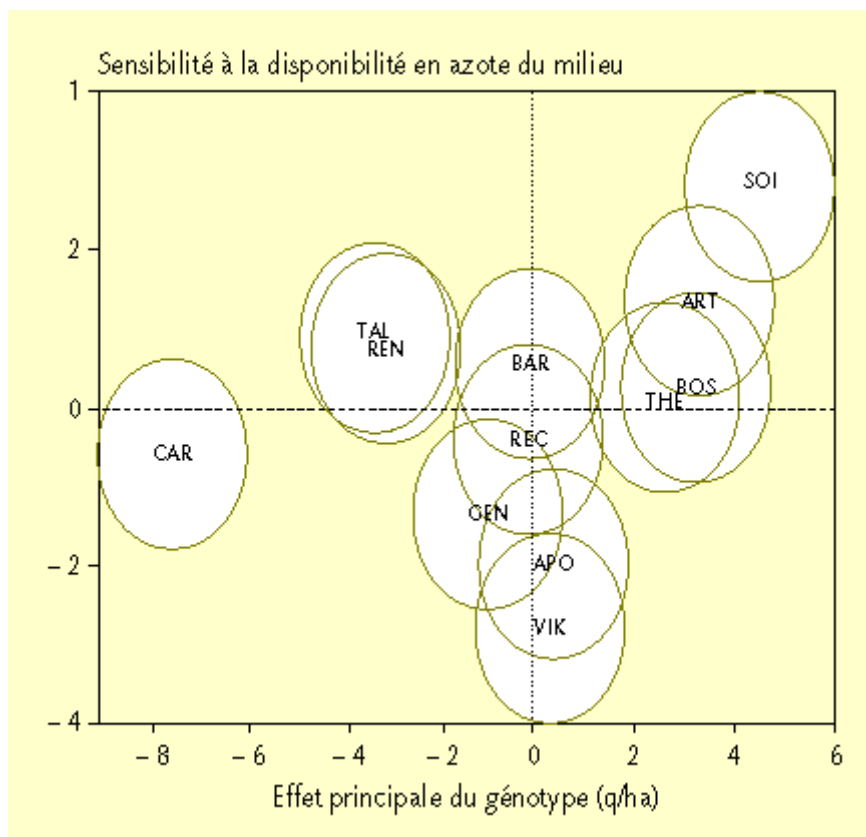


Figure 2. Sensibilité des génotypes à la disponibilité en azote de la culture dans le milieu (en ordonnée) en fonction de l'effet du génotype exprimé pour le rendement en grain (en abscisse). Les sensibilités ont été estimées par les pentes issues d'une régression factorielle incluant la covariable « disponibilité en azote du milieu », mesurée par la quantité d'azote absorbée par grain à la récolte. Les pentes positives correspondent aux génotypes dont le rendement augmente avec la disponibilité en azote alors que les pentes négatives correspondent aux génotypes dont le rendement diminue lorsque la disponibilité en azote s'accroît. Décodage des génotypes : APO = Apollo, ART = Artaban, BAR = Baroudeur, CAR = Camp-Rémy, GEN = Génial, REC = Récital, REN = Renan, ROS = Rossini, SOI = Soissons, TAL = Talent, THE = Thésée et VIK = Viking (d'après Brancourt-Hulmel [8]).

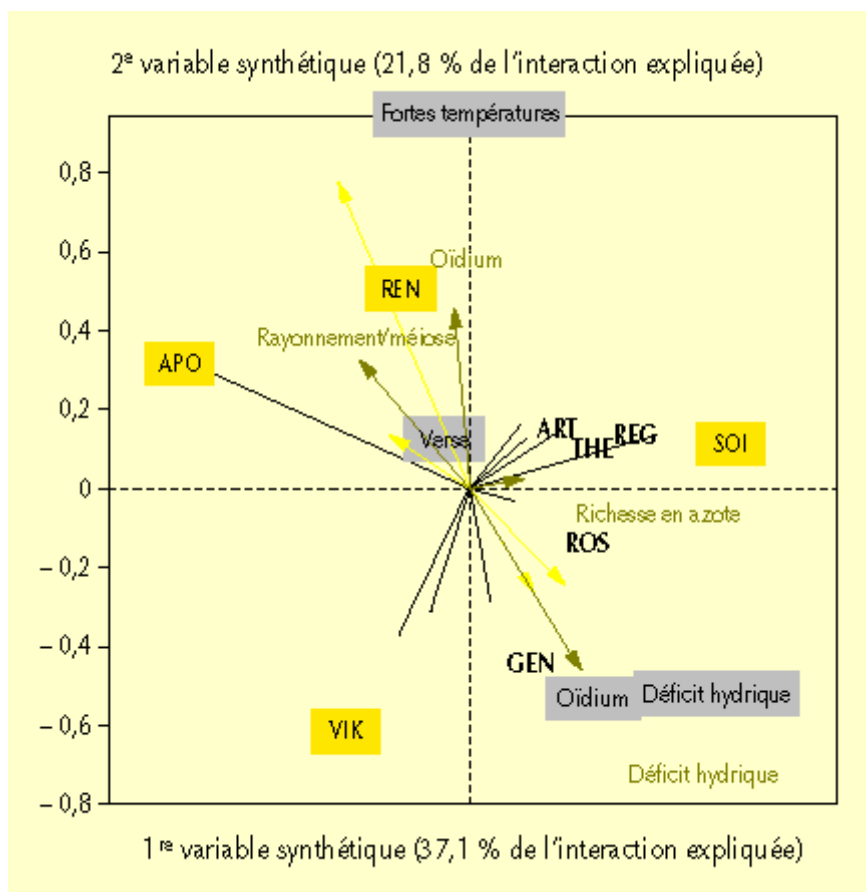


Figure 3. « Biplot » ou graphique représentant les pentes des génotypes (lignes noires) et les variables du milieu (lignes fléchées). Les génotypes les plus interactifs apparaissent sur fond jaune et les moins interactifs sont proches de l'origine. Les variables de l'élaboration du nombre de grains sont symbolisées en vert et celles du remplissage sur fond gris. Décodage des génotypes : APO = Apollo, ART = Artaban, BAR = Baroudeur, CAR = Camp-Rémy, GEN = Génial, REC = Récital, REN =

Renan, ROS = Rossini, SOI = Soissons, TAL = Talent, THE = Thésée et VIK = Viking
(d'après Brancourt-Hulmel et al. [12]).

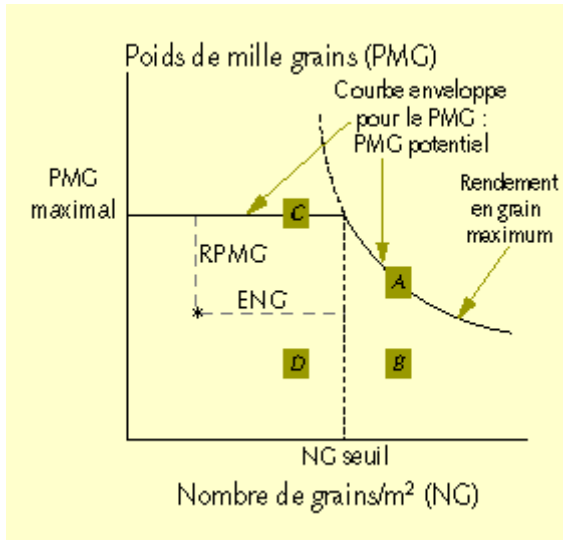


Figure 4. Principe du diagnostic agronomique et estimation des déviations de composantes : écart au nombre de grains seuil (ENG) et réduction du poids de mille grains (RPMG). Quatre cas pour l'interprétation : milieu A non limitant pour le nombre de grains (NG) et le poids de mille grains (PMG), milieu B limitant pour PMG, milieu C limitant pour NG et milieu D limitant pour NG et PMG (d'après Brancourt-Hulmel et al. [17] et Leterme et al. [21]).

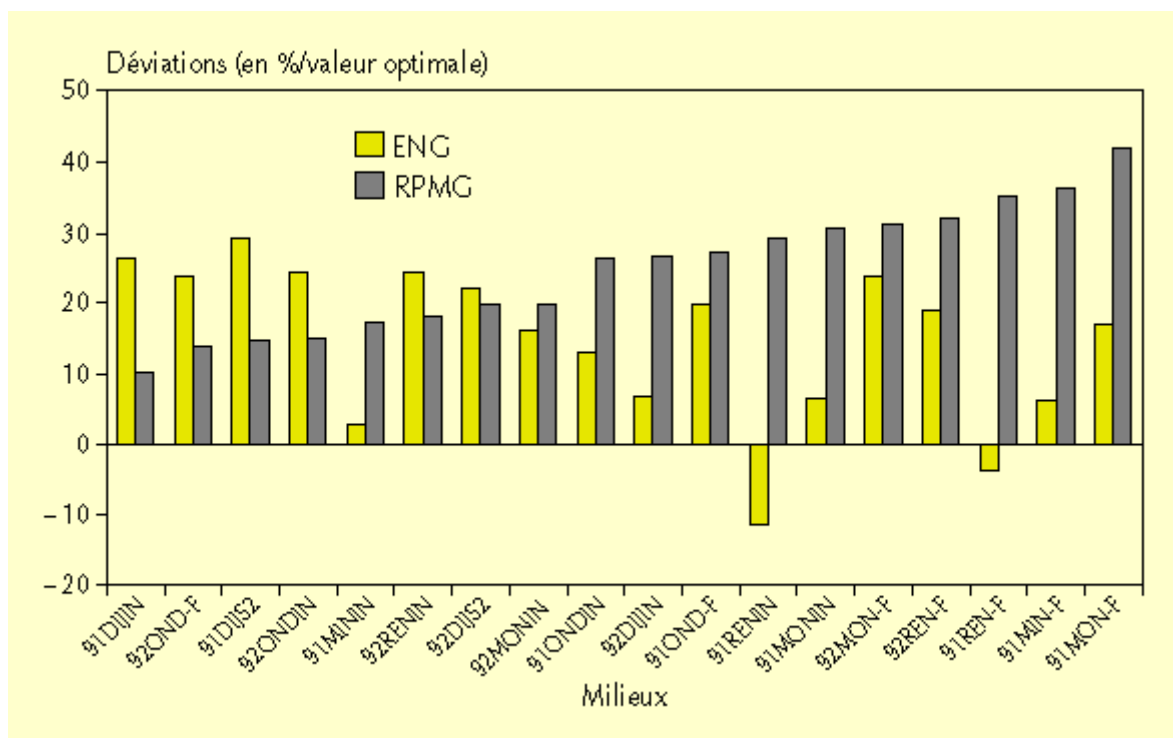


Figure 5. Illustration des déviations de composantes mesurées pour Thésée : écart au nombre de grains seuil (ENG) et réduction de poids de mille grains (RPMG). Certains milieux ont subi peu ou pas de contraintes pendant l'élaboration du nombre de grains ($ENG < 0$) ou des contraintes plus ou moins fortes ($0 < ENG < 29\%$) tandis que les milieux ont tous subi l'action de contraintes plus ou moins importantes pendant le remplissage ($0 < RPMG < 42\%$). Décodage des milieux : les deux premiers chiffres repèrent l'année (91 ou 92), les trois caractères suivants le lieu (DIJ = Dijon, MIN = La Minière, MON = Mons, OND = Ondes, REN = Rennes) et les deux derniers le mode cultural (IN = mode standard, -F = mode sans fongicides, S2 = semis tardif).

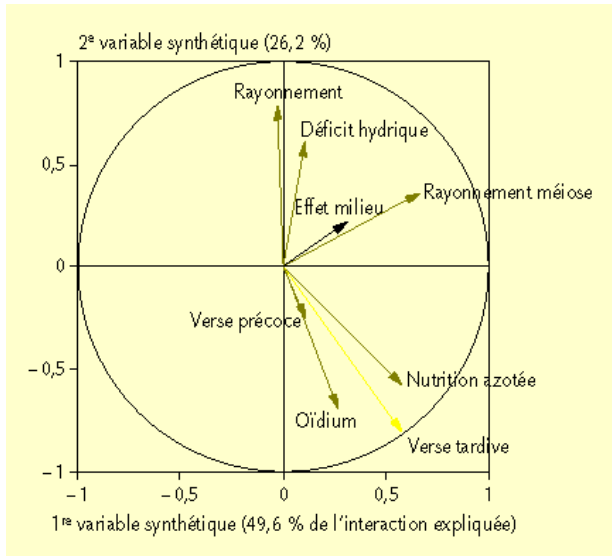


Figure 6. Contributions des variables environnementales à l'interaction observée pour l'écart au nombre de grains seuil (ENG). Les variables synthétiques sont les combinaisons linéaires des variables environnementales issues de la régression factorielle biadditive. Une variable proche du cercle contribue beaucoup à l'interaction tandis qu'une variable proche de l'origine y contribue peu (d'après Brancourt-Hulmel et al. [12]).

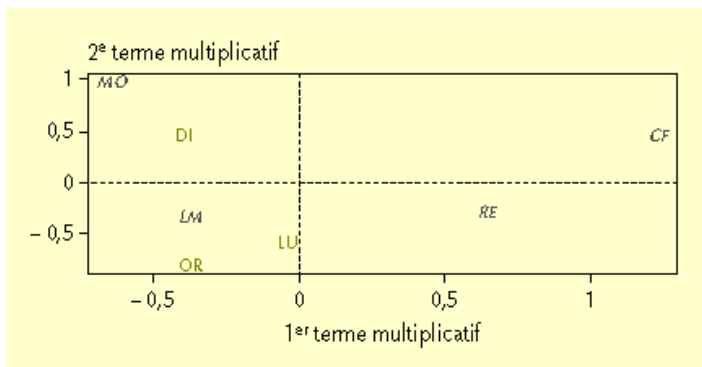


Figure 7. Sous-échantillonnage de milieux à partir de la modélisation de l'interaction par un modèle multiplicatif. Pour représenter un maximum de cas de figure pour l'interaction, les milieux sont sélectionnés (en gris) dans chaque quadrant. Codes des milieux : MO = Mons, DI = Dijon, CF = Clermont-Ferrand, LM = Le Moulon, OR = Orgeval, LU = Lusignan et RE = Rennes.