



**HAL**  
open science

## Présentation d'une méthode pour améliorer la connaissance des aptitudes variétales, intérêt des modèles, besoins de développements nouveaux

Christophe Lecomte, Lorène Prost, Arnaud Gauffreteau

### ► To cite this version:

Christophe Lecomte, Lorène Prost, Arnaud Gauffreteau. Présentation d'une méthode pour améliorer la connaissance des aptitudes variétales, intérêt des modèles, besoins de développements nouveaux. Innovations Agronomiques, 2009, 7, pp.105-119. 10.17180/p0gw-5734 . hal-01173237

**HAL Id: hal-01173237**

**<https://hal.science/hal-01173237>**

Submitted on 6 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## Présentation d'une méthode pour améliorer la connaissance des aptitudes variétales, intérêt des modèles, besoins de développements nouveaux

C. Lecomte<sup>1</sup>, L. Prost<sup>2</sup>, A. Gauffreteau<sup>3</sup>

<sup>1</sup>: UMR LEG INRA / AgroSup Dijon, 17 rue Sully, 21065 Dijon

<sup>2</sup>: UMR SADAPT INRA / AgroParisTech, BP 01, 78850 Thiverval Grignon

<sup>3</sup>: UMR d'Agronomie INRA / AgroParisTech, BP 01, 78850 Thiverval-Grignon

Correspondance : lecomte@dijon.inra.fr

### Résumé

Les connaissances acquises en agronomie et statistiques permettent aujourd'hui d'aller beaucoup plus loin dans l'exploitation des résultats des réseaux d'évaluation variétale, et dans la connaissance des milieux et des variétés. Concernant les variétés, il est possible d'estimer les tolérances variétales à des facteurs limitants de la production dont les effets sont peu visibles ou difficilement mesurables. Cet article présente une méthode d'interprétation des résultats des essais variétaux, basée sur la description des facteurs limitants de la production. Celle-ci repose sur la collecte de données environnementales, sur l'observation de variétés – témoins et sur la mise en œuvre de modèles agronomiques. Des outils statistiques (régression linéaire, régression factorielle), apportant rapidité et richesse d'interprétation, sont également utilisés pour mettre en relation les variations d'une variable de production avec les variables décrivant les facteurs limitants. Les bases nécessaires à la mise en œuvre de cette méthode (données, observations) et ses limites actuelles sont présentées. Les travaux en cours sur l'amélioration des modèles agronomiques et statistiques, qui mobilisent tous les acteurs de l'évaluation des variétés, devraient permettre de réduire le travail supplémentaire à fournir dans le suivi des essais, et d'obtenir une meilleure convergence entre interprétation statistique et expertise agronomique sur les essais variétaux.

**Mots-clés:** évaluation variétale, interaction génotype x milieu, diagnostic agronomique, modèles agronomiques, modèles statistiques

### Abstract: Presentation of a method for improving knowledge on variety aptitudes, relevance of models and needs for new developments

Agronomy and statistics knowledge makes it possible to go further in the exploitation of trials for variety evaluation, and in describing the environment and the variety aptitudes. Thus, it is possible to assess variety tolerance to limiting factors of production whose effects are not visible or hardly measurable. This article presents a method for interpreting results of variety trials, based on the description of limiting factors of production. It depends on new information on the environment, on observations of probe genotypes, and on the use of agronomical models. Statistical tools (linear regression, factorial regression) provide quick and in-depth interpretations. They are used to link fluctuations of a given production variable and of variables describing limiting factors. The bases required for this method (data, observations) and its current limitations are presented. Present research aimed at improving agronomical and statistical models groups together different actors for variety evaluation. They should lead to a reduction in the work required in collecting data on the trials, and to a better concordance between statistical interpretation and agronomical expertise.

**Keywords:** variety evaluation, genotype x environment interaction, crop diagnosis, agronomical models, statistical models

## Introduction

Les agriculteurs et techniciens du monde agricole travaillent dans un milieu naturel dont ils ont généralement une bonne connaissance : ils connaissent ses atouts et ses faiblesses et savent quelles interventions techniques vont leur permettre de tirer parti des premiers et de compenser au mieux les secondes. Par exemple, les dates de travail du sol, les outils employés... sont raisonnés en fonction des caractéristiques des sols et du climat pendant la période de travaux (quantité d'argile, rapidité de ressuyage, pluviométrie hivernale...).

Le choix variétal est un des éléments importants qui permet d'adapter l'itinéraire technique aux contraintes les plus susceptibles d'apparaître dans le milieu. Mais pour étayer ce choix, les informations disponibles sont toujours limitées et assez variables selon l'espèce. Par exemple, pour le blé tendre d'hiver et le pois d'hiver, les catalogues variétaux fournissent des informations sur la précocité des variétés, sur la productivité et la qualité ; on a également des informations sur la résistance au gel et pour le blé (mais pas pour le pois), sur la résistance à la verse et la résistance aux principales maladies ; mais on n'a pas d'information sur la résistance des variétés à un grand nombre d'autres facteurs limitants comme une carence en azote, les excès de température, le stress hydrique (tableau 1). Pourtant, ces informations seraient bien utiles pour raisonner le choix variétal, d'autant plus qu'aujourd'hui, la nécessité de mettre en place des conduites de culture plus respectueuses de l'environnement va augmenter le risque d'apparition de certains facteurs limitants de la production.

Tableau 1 : Caractéristiques des variétés de pois d'hiver (*Semences et Progrès n°125, Septembre 2005*)

Variétés de pois d'hiver commercialisées en France																					
Variétés (et année d'inscription)	Obtenteur ou représentant	Type de plante	Cou- leur du grain	Tolérance au froid	Résultats Arvalis											Sur- face en se- mences (2005) (ha)					
					Hauteur à la récolte (cm)		Poids de 1 000 grains (grammes)		Teneur en pro- téines	Rendement (en % de Cheyenne)											
					2004	2005	2004	2005		2004	Bourgogne Barrois Lorraine		Centre Ile-de- France		Nord Picardie		Craie et cranettes		Sud		
									2000 à 2004 (1)		2005	2000 à 2004 (1)	2005	2000 à 2004 (1)	2005		2000 à 2004 (1)	2005	2000 à 2004 (1)	2005	
Apache (05)	GAE Semences	afila	jaune	bonne	43	36	180	180	23,0	-	113	-	109	-	108	-	105	-	-	-	
Beckii (ce)	Sem Partners	feuilles	jaune	bonne	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	11
Brévent (91)	Epis-Sem	feuilles	jaune	ass. faible	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	41
Cartouche (05)	Serasesem	afila	jaune	bonne	41	36	170	170	22,0	-	113	-	112	-	112	-	103	-	-	-	116
Cherokee (05)	GAE Semences	afila	jaune	très bonne	37	30	190	190	22,5	-	119	-	113	-	117	-	101	-	111	-	78
Cheyenne (00)	GAE Semences	afila	jaune	bonne	37	27	200	200	22,5	100	100	100	100	100	100	100	100	100	100	-	226
Dove (ce)	Agri-Obtentions	afila	vert	bonne	37	27	180	180	21,0	101	111	106	118	96	125	103	104	86	-	-	34
Iceberg (ce)	Lecureur	afila	jaune	bonne	43	-	170	-	22,5	-	-	98	-	96	-	-	-	99	-	-	6
Isard (05)	Agri-Obtentions	afila	jaune	très bonne	37	27	205	200	22,0	-	132	-	123	-	130	-	97	-	-	-	20
Lucy (ce)	GAE Semences	afila	vert	moyenne	44	36	175	180	23,5	95	-	105	111	87	110	-	94	106	100	73	73
Spirit (ce)	GAE Semences	afila	jaune	moyenne	37	-	210	-	23,0	-	-	101	-	97	-	-	-	100	-	-	46
Valeur du témoin Cherokee (ou Lucy, pour 2005 région Sud) (quintaux/ha)										44		49		54		45		45			
Nombre d'essais										5		9		2		6		6			

(1) Les résultats d'essais de rendement 2000-2004 sont des moyennes obtenues sur ces 5 années.

(2) Les résultats d'essais de rendements 2005 de la région Sud concernent, en fait, la région Pays de la Loire, et sont donnés en fonction de Lucy.

La principale source de connaissance des variétés est l'expérimentation variétale. Elle est réalisée dans des réseaux d'essais nationaux ou régionaux, par les obtenteurs avant l'inscription au catalogue officiel des variétés, par le Geves (Groupe d'étude des variétés et des semences) et le réseau de partenaires associés au cours des deux années d'épreuves d'inscription, par les organismes techniques et des entreprises de multiplication – distribution pendant les années d'inscription et au-delà. Ces réseaux sont globalement très peu reliés entre eux, la connaissance acquise lors d'une étape devant le plus souvent être reconstruite à l'étape suivante (Lecomte, 2005 ; Prost, 2008).

Lors de l'évaluation des variétés en réseaux d'essais, on constate que les performances obtenues par les variétés (rendement, qualité...) sont instables entre sites et entre années, les classements variétaux changent entre situations (c'est ce que l'on appelle l'interaction génotype x milieu : figure 1), ce qui impose de répéter les évaluations pour mieux appréhender la diversité des comportements variétaux en fonction des milieux d'évaluation. Or, l'évaluation traditionnelle des variétés repose sur leurs résultats moyens. De ce fait, on fait comme si elles avaient une réponse parallèle entre les différents milieux (figure 2). Autrement dit, on considère que l'interaction génotype x milieu n'existe pas, alors qu'elle révèle justement les différences d'adaptation des variétés aux milieux. Il apparaît donc particulièrement important de bien la décrire et l'exploiter.

Figure 1 : Rendement de 12 variétés de blé dans 14 milieux (réseau INRA 1995-1997).

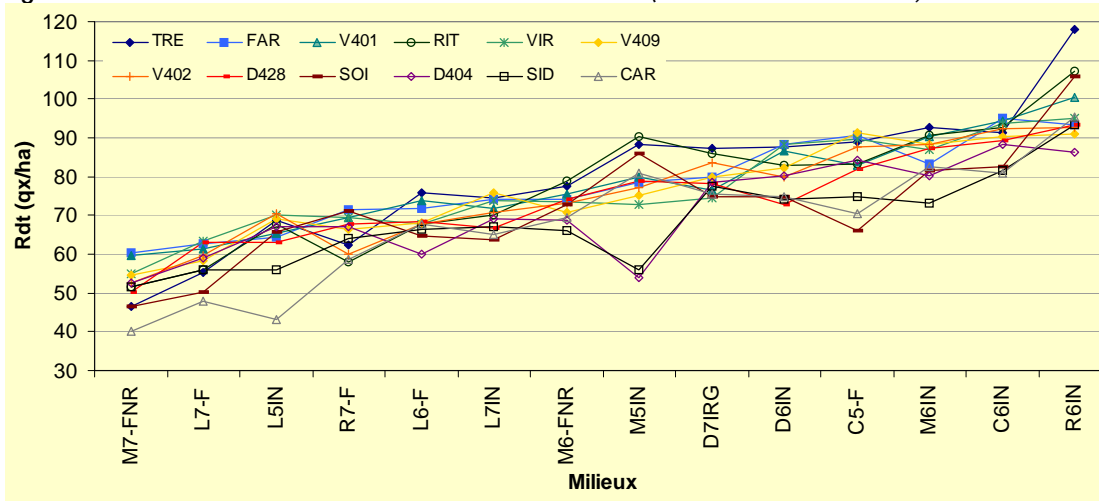
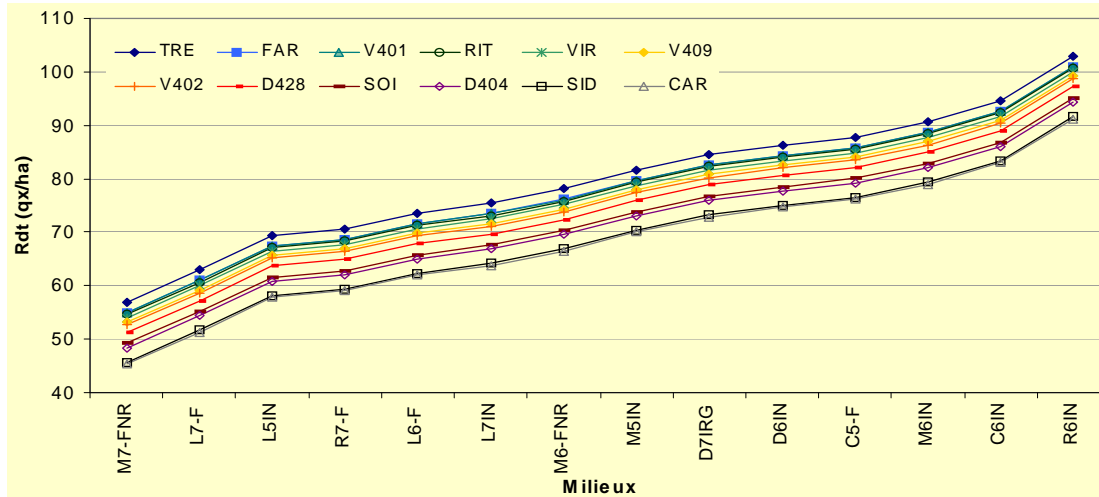


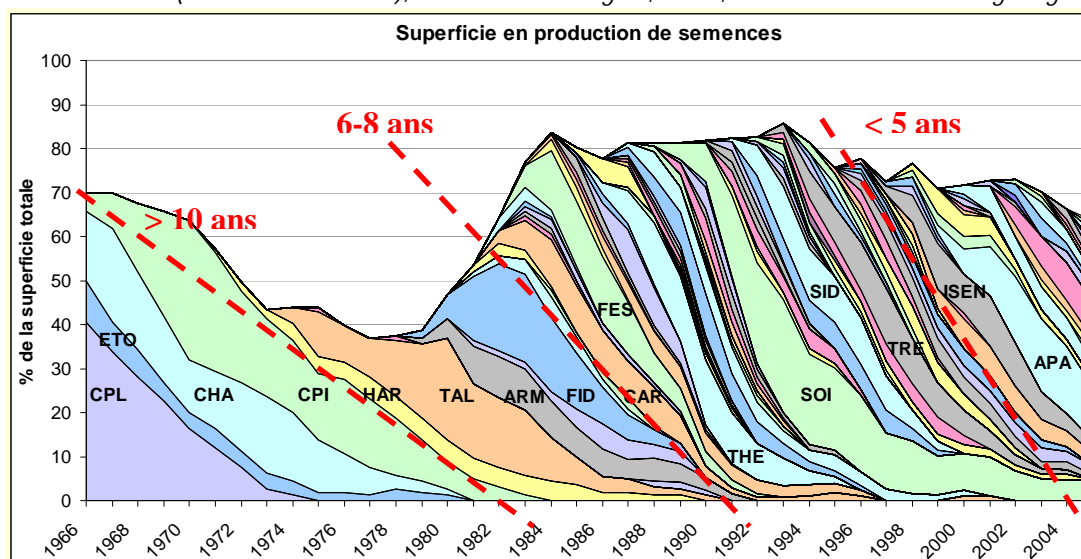
Figure 2 : Rendement des 12 mêmes variétés, décrit par les seules moyennes génotypiques et environnementales (réseau INRA 1995-1997).



Par ailleurs, les acteurs de l'évaluation variétale doivent faire face à un grand nombre d'essais, souvent délégués à des partenaires, et doivent réaliser un grand nombre de notations regroupées dans le temps. Ils ont aussi des délais de restitution des résultats très courts, ce qui leur laisse peu de latitude pour ajouter des observations et pour traiter les résultats. De plus, les variétés se renouvellent plus rapidement aujourd'hui qu'il y a 10 ou 20 ans, en particulier dans les espèces majeures de grande

culture (figure 3). Il faut donc aller vite pour apporter une connaissance de leur comportement qui soit suffisante pour que les acteurs de la mise en culture disposent des éléments utiles pour faire leur choix.

Figure 3 : Evolution de la surface en production de semences des variétés de blé tendre depuis les années 1960 (Source : GEVES (bulletin des variétés), Semences et Progrès, GNIS, Contrat de branche "Progrès génétique").



Aujourd'hui, il apparaît donc de plus en plus nécessaire de développer des méthodes d'analyse des réseaux expérimentaux qui permettent de tirer pleinement et rapidement parti de la variabilité des résultats. D'importants travaux ont été réalisés depuis longtemps, tant en agronomie qu'en statistiques, pour améliorer l'interprétation des résultats d'essais variétaux. Mais ces méthodes sont encore peu appliquées :

- Depuis les années 1970, des agronomes ont progressivement mis au point une méthode de diagnostic agronomique, qui permet d'identifier les contraintes apparues dans le milieu, à partir de l'analyse de l'élaboration du rendement de variétés bien connues (Meynard and David, 1992 ; Doré *et al.*, 1997). Mais cette méthode demande une expertise et est assez lourde à mettre en oeuvre.

- Des méthodes d'analyse statistique des résultats de réseaux d'essais variétaux ont été proposées dès les années 1960 pour interpréter la réponse différentielle des variétés aux milieux (régression conjointe, écovalences génotypiques ou environnementales..., Brancourt-Hulmel *et al.*, 1997). Proposée plus récemment, la régression factorielle permet d'expliquer l'interaction génotype x milieu en la décomposant par exemple au moyen de variables descriptives des milieux (Denis, 1980 & 1988). Si ces variables décrivent l'intensité des facteurs limitants apparus dans les milieux, les paramètres associés à ces variables dépendent des variétés et sont une estimation de la tolérance des variétés à ces facteurs limitants (van Eeuwijk *et al.*, 2004). Les méthodes d'analyse statistique sont rapides, elles permettent de traiter un grand nombre de données simultanément. Mais elles ne travaillent que par corrélation statistique, il n'y a pas de vérification agronomique des causes de variation des résultats.

L'objet de cet article est de présenter et de discuter une méthode d'analyse des résultats d'essais variétaux qui tire parti de ces deux grandes approches, l'approche agronomique et l'approche statistique, et qui s'appuie fortement sur des modèles proposés par ces deux disciplines, pour mieux connaître les milieux et les variétés.

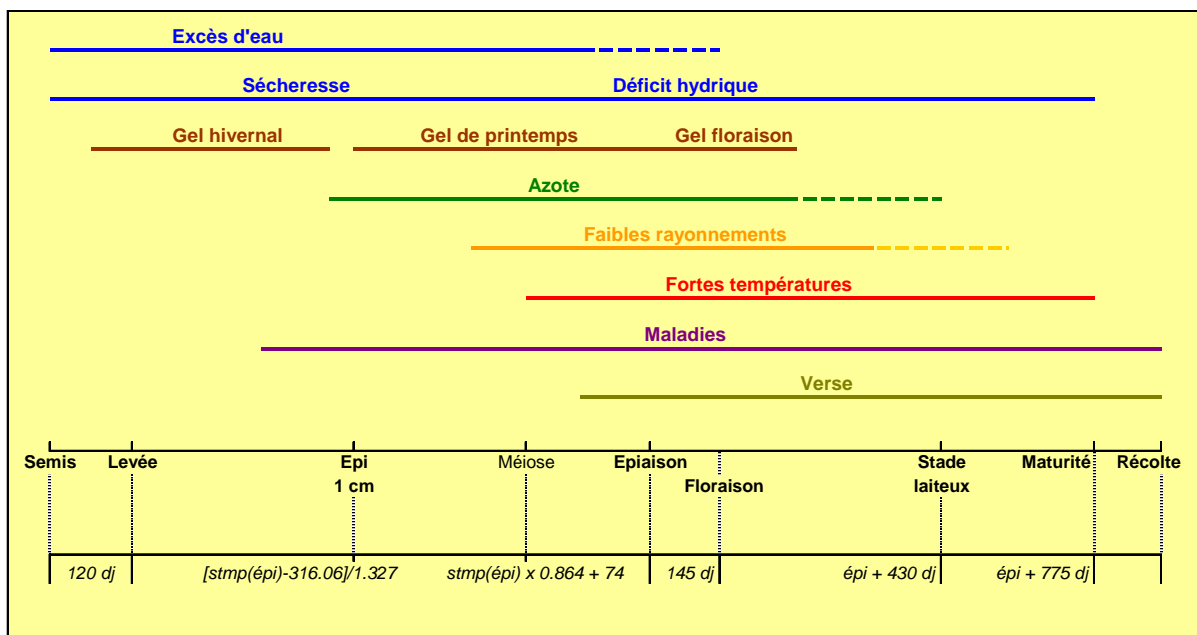
## 1- Informations que l'on peut retirer des essais variétaux

### 1.1- Il est nécessaire de décrire les milieux pour interpréter le comportement des variétés

La réponse des performances variétales aux milieux n'est pas parallèle (figure 1) : une variété peut montrer de meilleurs résultats que les autres dans un essai donné et de moins bonnes dans un autre. Ces variations sont dues aux contraintes (ou facteurs limitants) qui apparaissent dans les milieux, et aux différences de sensibilité des variétés à ces facteurs : une variété tolérante aux principaux facteurs limitants apparus dans un milieu montre de meilleures performances que les autres, alors qu'elle aura de moins bons résultats dans un milieu où ont agi des facteurs limitants auxquels elle est sensible. Il est donc nécessaire de décrire le plus complètement possible les facteurs limitants apparus dans les milieux d'essais pour expliquer les variations de comportement des variétés.

Les principaux facteurs limitants qui peuvent intervenir au cours du cycle cultural sont schématisés sur la figure 4 sur l'exemple du blé d'hiver. Ils se répartissent tout au long du cycle et peuvent être décrits par des observations ou des variables calculées par grandes phases délimitées par les stades – clés du développement de la culture (semis, levée, début de la phase reproductive, début floraison, début du remplissage des graines, maturité physiologique, récolte).

Figure 4 : Principaux facteurs limitants pouvant apparaître au cours du cycle du blé.



Il apparaît donc tout d'abord nécessaire de repérer les stades – clés du développement. Pour certains d'entre eux, ce repérage peut être coûteux en temps (plusieurs passages peuvent être nécessaires) et/ou imposer des prélèvements destructifs (par ex : repérage du stade épi 1cm chez le blé). Comme cela est souvent incompatible avec les contraintes de l'expérimentation variétale (il est difficile d'ajouter des parcelles supplémentaires pour des prélèvements et le calendrier de notations des expérimentateurs est souvent très chargé autour des stades – clés), des modèles agronomiques de développement sont progressivement mis au point, permettant d'estimer la date de certains stades à partir d'un autre (chez le blé, on prévoit ainsi la date du stade épi 1cm à partir de la levée et de l'épiaison : figure 4).

Les variables descriptives des facteurs limitants peuvent être issues d'observations visuelles (notes de verse, d'accidents, de maladies...), de mesures ou de prélèvements (par exemple : prélèvement de

plantes pour mesurer la teneur en azote à la floraison et déterminer l'indice de nutrition azotée), ou de variables calculées à partir des données météorologiques. Là aussi, les mesures peuvent être lourdes ou destructives. Des modèles agronomiques sont utilisés pour décrire des facteurs limitants dont le contrôle visuel est très difficile à réaliser (par exemple : établissement du bilan hydrique d'une culture et identification des périodes de stress hydrique), l'utilisation croissante des modèles devant permettre de gagner du temps et d'éviter des prélèvements (voir plus loin).

Le but de ces contrôles est de décrire les milieux expérimentaux. Il n'est donc pas nécessaire de répéter les mêmes observations sur toutes les variétés évaluées. Un choix de témoins, représentatifs de la diversité des variétés étudiées (notamment pour la précocité), permet de limiter la lourdeur de ces observations, les autres variétés restant décrites par les contrôles habituellement réalisés (date de certains stades et notations visuelles). La caractérisation des milieux repose donc sur quelques variétés, les témoins – révélateurs, représentatives de la diversité des autres variétés.

## 1.2- Caractérisation des milieux

Des outils statistiques simples comme la régression linéaire multiple pas à pas permettent de relier les variations de performance de chaque variété – témoin (rendement, qualité...) avec les variables décrivant les facteurs limitants apparus dans les milieux expérimentaux. Appliquée aux pertes de rendement d'une variété par rapport à une valeur de référence, l'expression de la régression linéaire multiple est la suivante :

$$dR_{dtrj} = \lambda_0 + \lambda_1 \cdot Ve_{1j} + \lambda_2 \cdot Ve_{2j} + \dots + \lambda_n \cdot Ve_{nj} + E_j$$

avec:  $dR_{dtrj}$  : perte de rendement du témoin considéré dans le milieu  $j$

$\lambda_n \cdot Ve_{nj}$  : termes correspondant aux différentes variables décrivant les facteurs limitants

$E_j$  : représente le reste de la variation, non expliquée par le modèle

Les variables sont introduites dans le modèle de régression linéaire multiple par ordre de corrélation décroissante avec la variable expliquée. Chaque nouvelle variable est introduite étant donné toutes celles qui la précèdent. Comme la corrélation statistique ne garantit pas la validité agronomique d'une relation, un test sur le signe de l'effet peut être introduit dans cette procédure, n'étant retenues que les variables dont l'effet va bien dans le sens d'une diminution du rendement.

Dans différents réseaux ainsi décrits, nous avons pu obtenir des parts de variations du rendement expliquées de (Lecomte, 2005) :

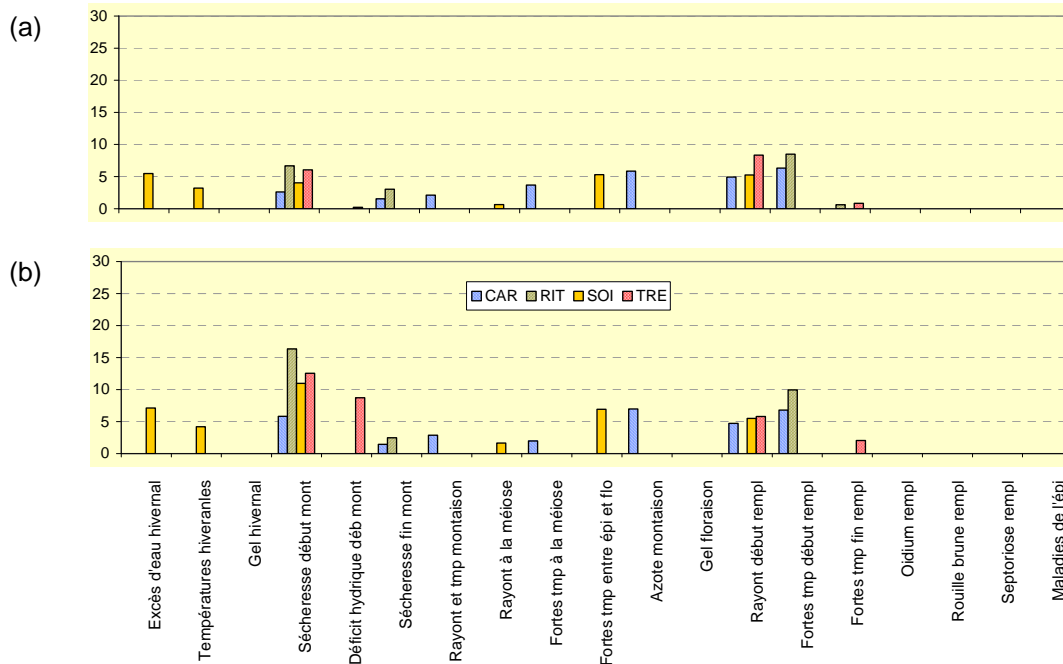
- 70 à 99 % selon le génotype révélateur, pour des réseaux de 10 à environ 35 milieux.
- 55 à 65 % pour des réseaux de 60 à 100 milieux.

L'outil statistique utilisé aboutit à un modèle mettant en relation par exemple un rendement et des variables décrivant les facteurs limitants apparus dans les milieux. En appliquant ce modèle dans chaque milieu, avec les valeurs observées des variables descriptives, on peut estimer la contribution de chaque facteur limitant aux pertes de rendement de la variété – témoin étudiée. La figure 5 donne l'exemple d'une caractérisation individuelle de deux milieux qui ne diffèrent que par la présence ou non d'une irrigation au début de la montaison du blé. On constate ici que l'absence d'irrigation a entraîné une aggravation de la sécheresse au début de la montaison. Ce facteur limitant était attendu, cette vérification constitue donc une validation de la démarche statistique effectuée.

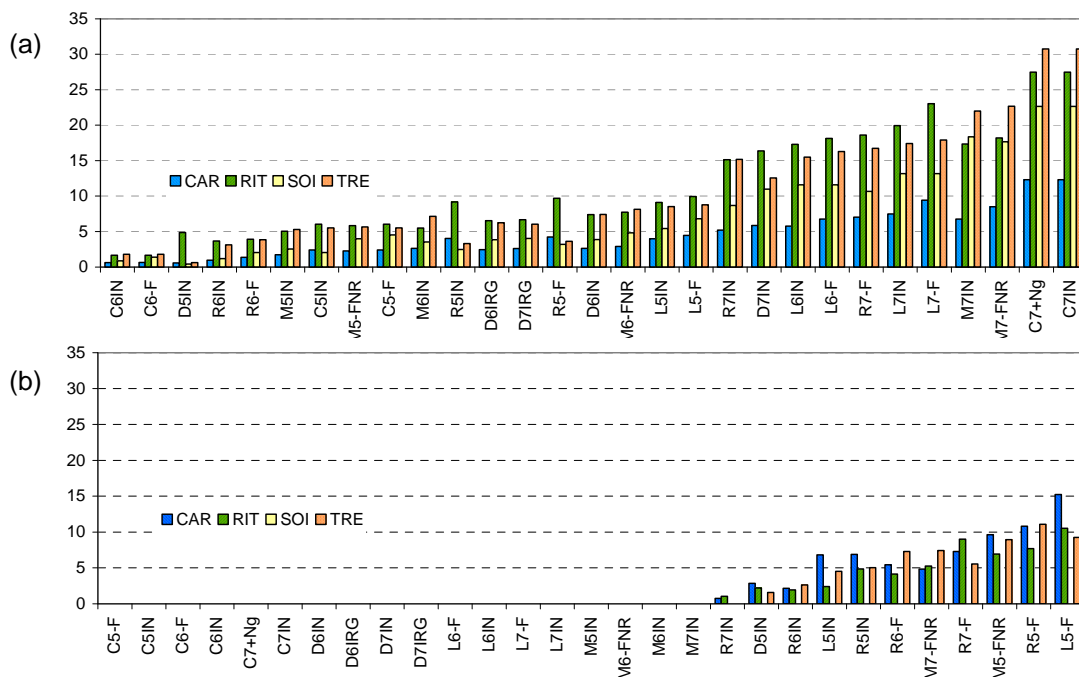
Dans le cas de réseaux pluriannuels, cette analyse peut ainsi permettre d'estimer la fréquence des facteurs limitants dans un site donné.

Une représentation commune de tous les essais d'un réseau peut être effectuée, pour l'intensité d'un facteur limitant. La figure 6 montre ainsi la classification des essais obtenue pour deux facteurs limitants du blé (réseau blé tendre de l'INRA de 1995 à 1997). Dans le premier cas (sécheresse au début de la montaison), tous les essais ont été plus ou moins affectés par le facteur limitant. Dans le second cas (septoriose pendant le remplissage des grains), un certain nombre d'essais n'ont pas été concernés. On constate aussi que dans les deux cas, les pertes de rendement peuvent être très importantes (30% de la valeur de rendement de référence dans le premier cas).

**Figure 5 :** Contribution des différents facteurs limitants aux pertes de rendement (en % par rapport à une valeur de référence) dans deux milieux d'évaluation : (a) Dijon 1997, conduite irriguée et (b) Dijon 1997 conduite non irriguée (réseau blé INRA 1995-1997). Les 4 variétés - témoins étaient Camprémy, Ritmo, Soissons et Trémie.



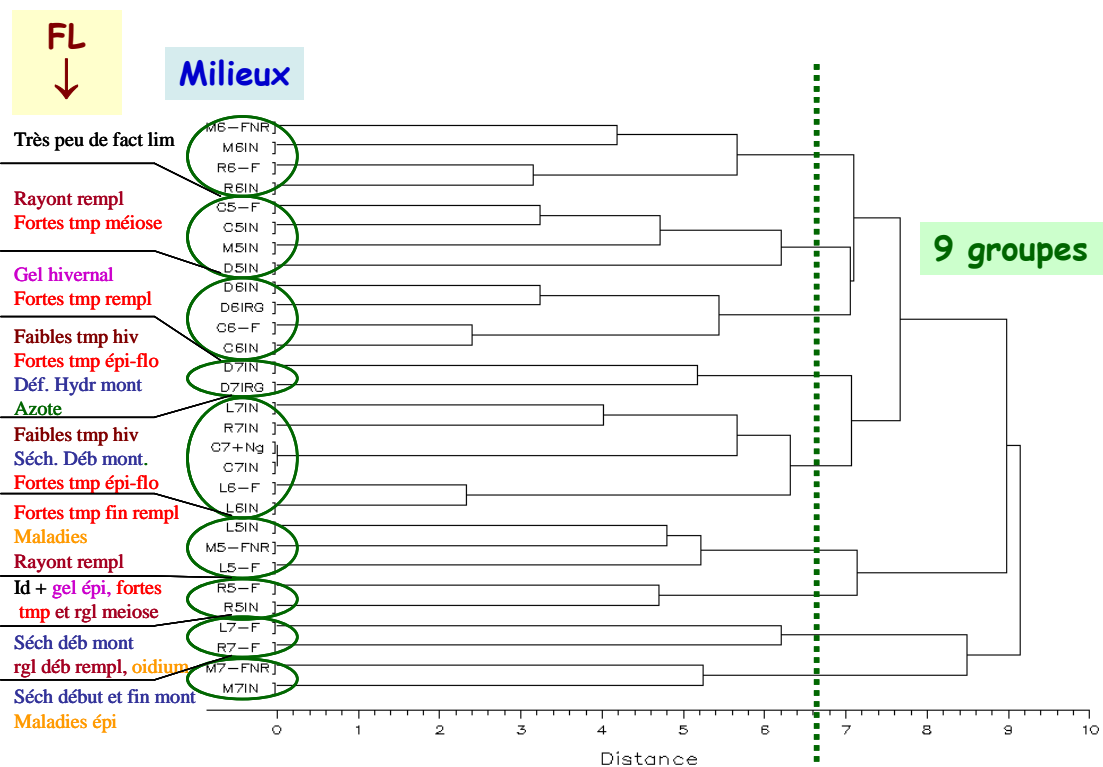
**Figure 6 :** Pertes de rendement (en % par rapport à une valeur de référence) occasionnées par (a) la sécheresse au début de la montaison et (b) la septoriose pendant le remplissage des grains, dans le réseau blé INRA 1995-1997. Les 4 variétés - témoins étaient Camprémy, Ritmo, Soissons et Trémie.





Le réseau expérimental peut aussi être représenté pour tous les facteurs limitants simultanément en utilisant une méthode de classification, les essais appartenant à un même groupe présentant une analogie de facteurs limitants (figure 7). Une analyse en composantes principales associée à cette classification permet d'identifier les facteurs limitants responsables de la constitution des groupes.

Figure 7 : Classification des milieux par analogie de facteurs limitants (FL) dans le réseau blé INRA 1995-1997.



Dans l'analyse des résultats d'un réseau expérimental, une telle approche permet de regrouper les essais sur une base agronomique et non pas géographique (ce deuxième mode de regroupement s'avérant souvent décevant dans la mesure où la variabilité à l'intérieur des groupes est du même ordre que la variabilité entre les groupes). Pour le réseau blé INRA 1995-1997, le mode de regroupement opéré de cette façon a été plus efficace pour structurer le réseau qu'un regroupement *a priori* sur la base de l'année, de la région, et même de la conduite de culture.

### 1.3- La description des milieux permet d'estimer les tolérances variétales aux facteurs limitants

Une fois que les milieux expérimentaux sont décrits, on peut utiliser cette description pour analyser le comportement de toutes les variétés étudiées. Comme indiqué en introduction, la réponse différentielle des variétés aux facteurs limitants est décrite par l'interaction génotype x milieu, c'est-à-dire les écarts aux valeurs moyennes des milieux ou des variétés.

Plusieurs méthodes de décomposition de l'interaction génotype x milieu ont été proposées par les statisticiens (Brancourt-hulmel *et al.*, 1997). Parmi elles, la régression factorielle présente l'avantage de la simplicité : il s'agit d'une décomposition linéaire de l'interaction, comme la régression linéaire multiple utilisée pour relier les pertes de rendement des témoins révélateurs aux variables décrivant les facteurs limitants. De plus, les paramètres associés aux variables environnementales utilisées (celles qui

décrivent les facteurs limitants) sont des paramètres variétaux qui représentent une estimation de la tolérance de chaque variété au facteur limitant correspondant.

Dans la régression factorielle, les variables sont introduites une par une, par ordre de corrélation décroissante avec le terme d'interaction, comme pour l'étape précédente (caractérisation des milieux par une régression linéaire multiple). Le modèle de régression factorielle s'écrit de la façon suivante :

$$Rdt_{ij} = Rdtm + G_i + M_j + G_i * M_j + E_{ij}$$

$$\text{et : } G_i * M_j = \rho_{1i}.Ce_{1j} + \dots + \rho_{ni}.Ce_{nj} + G_i.M'_j$$

avec :  $Rdt_{ij}$  : rendement de la variété  $i$  dans le milieu  $j$

$Rdtm$  : rendement moyen de toutes les variétés dans tous les milieux

$G_i$  : effet moyen du génotype (variété)  $i$

$M_j$  : effet moyen du milieu  $j$

$G_i * M_j$  : interaction entre le génotype  $i$  et le milieu  $j$

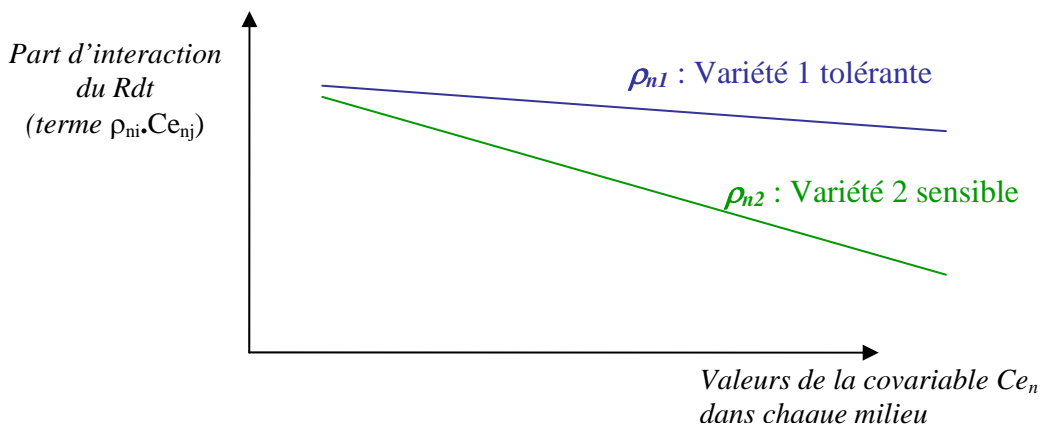
$E_{ij}$  : erreur résiduelle (ou aléatoire) de l'analyse

$\rho_{ni}.Ce_{nj}$  : interaction entre le génotype  $i$  et la covariable environnementale  $Ce_n$  dans le milieu  $j$ , le paramètre  $\rho_{ni}$  dépend du génotype  $i$

$G_i.M'_j$  : reste de l'interaction, non expliquée par les termes associés aux différentes covariables environnementales

La figure 8 donne la signification des paramètres génotypiques  $\rho_{ni}$  : ils représentent la pente de la relation entre la part de rendement expliquée par le terme d'interaction  $\rho_{ni}.Ce_{nj}$  pour chaque variété. Sur la figure, le rendement de la variété 1 décroît peu avec l'augmentation de l'intensité du facteur limitant  $Ce_n$  entre les différents milieux, alors que le rendement de la variété 2 décroît davantage. La variété 1 est donc tolérante au facteur limitant alors que la variété 2 est sensible.

Figure 8 : Signification des paramètres génotypiques  $\rho_{ni}$ .



De cette façon, on peut donc estimer la tolérance des variétés étudiées à tous les facteurs limitants explicatifs de l'interaction. Cette démarche est intéressante pour appréhender la tolérance des variétés à des facteurs limitants dont on ne peut pas facilement observer les effets, comme le stress hydrique, les fortes températures... La figure 9 donne un exemple de représentation des notes de tolérance des variétés de blé aux fortes températures pendant le remplissage des grains (réseau INRA 1995-1997). L'utilisation d'une démarche statistique permet en outre de calculer et de représenter une incertitude autour des valeurs moyennes.

La validité de ces notes a pu être vérifiée pour des facteurs limitants dont les effets sont observables, comme certaines maladies. Dans le réseau étudié, nous avons obtenu une corrélation de 0.9 entre les

notes de tolérance à la rouille brune ainsi déterminées et les notes moyennes de résistance déduites des observations sur l'ensemble des essais du même réseau.

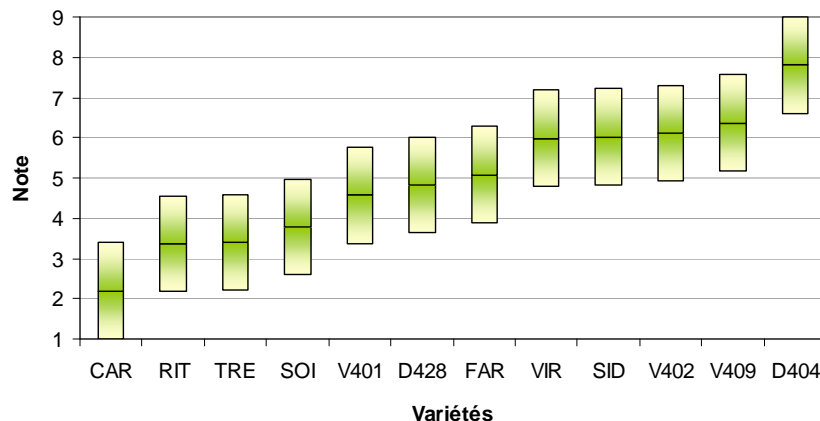


Figure 9 : Classification des variétés de blé étudiées dans le réseau INRA 1995-1997 pour leur tolérance aux fortes températures pendant le remplissage des grains.

## 2- Données complémentaires à acquérir dans les essais variétaux, possibilité de les alléger

La méthode d'analyse des résultats que nous venons de présenter requiert des informations non fournies dans le suivi habituel des essais variétaux. Mais nous avons vu que l'essentiel de ces informations peut être restreint à quelques variétés – témoins, représentatives de la diversité variétale étudiée.

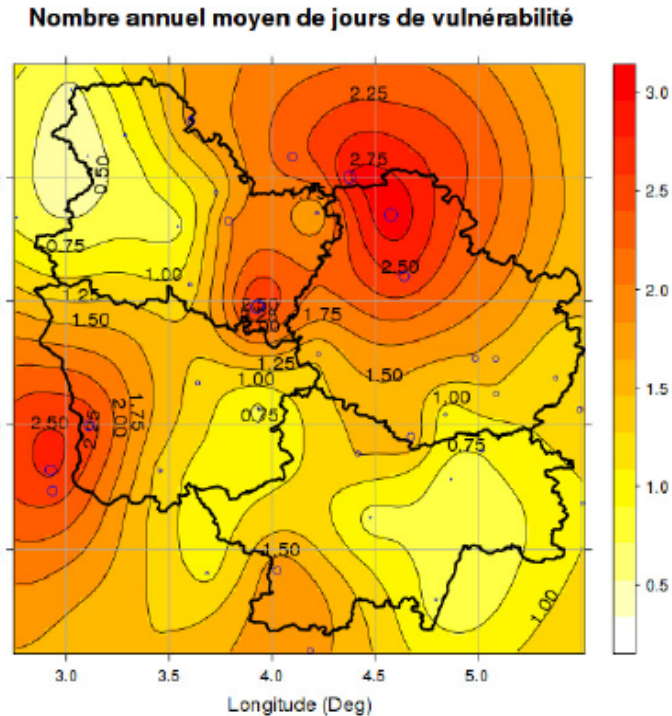
- *Il est tout d'abord nécessaire d'identifier les dates de plusieurs stades - clé*, qui permettent de délimiter des phases du cycle sur lesquelles sont calculées plusieurs variables descriptives des facteurs limitants : semis, levée, fin de la phase végétative, floraison (ou début floraison pour des espèces à floraison indéterminée), début du remplissage des grains, maturité physiologique. Dans le suivi habituel des essais variétaux, on connaît la date de semis, mais pas toujours la date de levée, et on connaît plus ou moins bien la date de floraison (pour les céréales, c'est plutôt la date d'épiaison qui est repérée). Il n'y a pas d'identification précise de la date du passage entre la phase végétative et la phase reproductrice (initiation florale, qui pour les céréales a lieu au voisinage du stade épi 1cm), ni de la date de début du remplissage (assimilée à la date de floraison chez les céréales, mais distincte pour les espèces à floraison indéterminée), ni de celle de la fin du remplissage.

Des modèles de développement sont requis pour estimer la date de ces stades, ou pour limiter le travail de repérage quand il est entrepris. Ces modèles fonctionnent le plus souvent sur la base de sommes de températures, par exemple en reliant la somme de températures entre le semis et la floraison à la somme de températures entre le semis et le stade correspondant au début de la phase reproductrice. Au bas de la figure 4 sont indiqués quelques modèles simples pour l'estimation des principaux stades de développement du blé, sachant que des modèles similaires sont utilisés ou en cours de développement chez d'autres espèces.

- *Cinq données météorologiques journalières sont également nécessaires* pour calculer les variables décrivant les principaux facteurs limitants d'origine climatique (excès ou manque d'eau, basses ou fortes températures, rayonnement) : les températures minimales et maximales, la pluviométrie et l'évapo-transpiration Penman, le rayonnement global.

Là aussi, les données météorologiques ne sont pas toujours disponibles sur le site expérimental. Des météorologistes conduisent des travaux pour tenter d'estimer certains paramètres par interpolation à partir de sites renseignés. L'interpolation est plus facile pour les données comme les températures ou le rayonnement global, qui varient peu géographiquement, ou pour des variables intégratives décrivant un risque particulier, par exemple les risques de gel (figure 10). Pour la pluviométrie, il semble nécessaire

aujourd'hui d'avoir une détermination aussi proche que possible du site à caractériser. L'évapotranspiration potentielle est moins variable à l'échelle d'une région, et l'absence d'une donnée précise comme l'évapotranspiration Penman peut être compensée de façon à peu près satisfaisante par la méthode de calcul de Turc, basée sur la température moyenne et le rayonnement global.



**Figure 10 :** Détermination du nombre de jours « à risque » pour le gel du pois d'hiver en Bourgogne sur la période 1988-2007, par interpolation entre stations météorologiques. Le critère retenu est le nombre de jours où la température minimale journalière est inférieure à  $-6^{\circ}$ , et présente un écart d'au moins  $5^{\circ}$  avec la température minimale du jour précédent (Cuccia, 2008).

Le calcul du bilan hydrique repose sur des hypothèses de développement de la culture et sur une estimation de la profondeur d'enracinement qui ne sont pratiquement pas vérifiées. De ce fait, ce calcul comporte de grosses imprécisions. Pour fiabiliser le calcul du bilan hydrique, nous proposons depuis plusieurs années d'implanter dans les essais variétaux des tensiomètres qui permettent de mieux repérer à quel moment la plante commence à restreindre sa consommation en eau. Il est alors possible d'ajuster les paramètres du bilan hydrique (notamment réserve en eau du sol) de façon à ce que le calcul du bilan indique le début d'un déficit en même temps que la valeur des tensiomètres commence à augmenter.

*- Il est également nécessaire de noter les accidents (gel, verse...) et maladies* pour ne pas manquer la cause possible d'une baisse de performances. Mais les observations ne sont jamais effectuées au même moment d'un essai à l'autre. Souvent une seule notation est effectuée alors que plusieurs seraient nécessaires pour juger l'évolution du phénomène. Aussi souvent que possible, les notateurs se réfèrent à une grille commune (par exemple échelle de notation du Geves), mais une part de subjectivité subsiste : par exemple, un technicien observe dans l'absolu ce qu'il voit à un moment donné, alors qu'un sélectionneur intègre dans sa notation la précocité des variétés, et donc le fait que certaines variétés « échappent » à l'accident grâce à une période de sensibilité décalée (Dizien, 2009).

De nombreux travaux visent à développer des modèles pour homogénéiser la description des accidents culturels entre sites et éventuellement remplacer des notations absentes. Ainsi, un modèle de prévision des dégâts de gel a été mis au point sur le blé tendre (Lecomte *et al.*, 2003). Il prend en compte la variété, le stade de développement de la plante, et le régime de température qui a précédé le coup de froid (degré d'endurcissement de la plante). Testé sur 11 campagnes expérimentales en conditions naturelles, ce modèle permet de prévoir l'apparition des dégâts de gel hivernal à  $1^{\circ}\text{C}$  près. Il est en cours d'adaptation sur le pois d'hiver, et une première évaluation sur deux variétés de sensibilité

contrastée a permis d'obtenir une corrélation entre dégâts de gel observés / dégâts prédits de 0.8 (Lecomte, travaux en cours).

Des modèles de développement des maladies ont été proposés par plusieurs organismes. Ils présentent l'avantage d'établir une cinétique d'évolution des maladies, qui n'est pas accessible par les observations habituelles. Ils ne prennent généralement pas en compte les sensibilités variétales et les conduites de cultures effectuées (traitement fongicides), mais l'évolution prévue, recalée par une ou deux observations, pourrait permettre d'estimer l'importance des maladies à un stade donné identique pour tous les essais.

- *L'alimentation azotée* figure parmi les facteurs limitants trophiques les plus importants, surtout quand on se situe dans une perspective de réduction des intrants, même pour une légumineuse dans la mesure où la fixation de l'azote de l'air peut être inhibée en cas de stress hydrique, d'excès d'eau, ou de faibles températures. Le statut azoté des plantes doit donc être décrit, la méthode la plus précise étant celle qui consiste à calculer l'indice de nutrition azotée (INN) au début de la floraison (Justes *et al.*, 1997). Cette détermination nécessite de prélever des plantes sur une surface de sol bien définie (pour calculer la biomasse produite par m<sup>2</sup>) et de doser la teneur en azote de ces plantes. C'est donc une mesure destructive, qui nécessite de prélever au plus près de la date de floraison, exigences auxquelles il est souvent difficile de répondre, même sur une seule variété.

Là aussi, des travaux sont en cours pour remplacer ce prélèvement par une mesure de réflectance infra-rouge (appareils de type Spad ou hydroN-tester). Aujourd'hui, on arrive assez bien à s'affranchir de l'effet variétal, mais il subsiste un effet année dans la relation INN / Indice Spad (rapport d'une mesure à la mesure sur une parcelle surfertilisée ; Prost et Jeuffroy, 2007). D'autres voies d'allègement sont envisagées, en utilisant des modèles de fonctionnement qui estiment la biomasse produite et la teneur en azote des plantes à un stade donné. Mais l'imprécision reste encore importante, notamment sur la teneur en azote.

### 3- Limites de cet exercice et perspectives pour les dépasser

Cette méthode d'analyse des résultats d'essais variétaux présente plusieurs limites ou contraintes dans sa mise en œuvre.

- Tout d'abord, des contraintes sont liées aux pratiques expérimentales elles-mêmes : les observations sur les essais requièrent du temps, et c'est souvent aux moments clés déjà très chargés (floraison...) qu'il faudrait investir dans de nouvelles observations. Les essais variétaux sont volumineux, il est parfois difficile d'ajouter des parcelles supplémentaires pour effectuer des prélèvements.

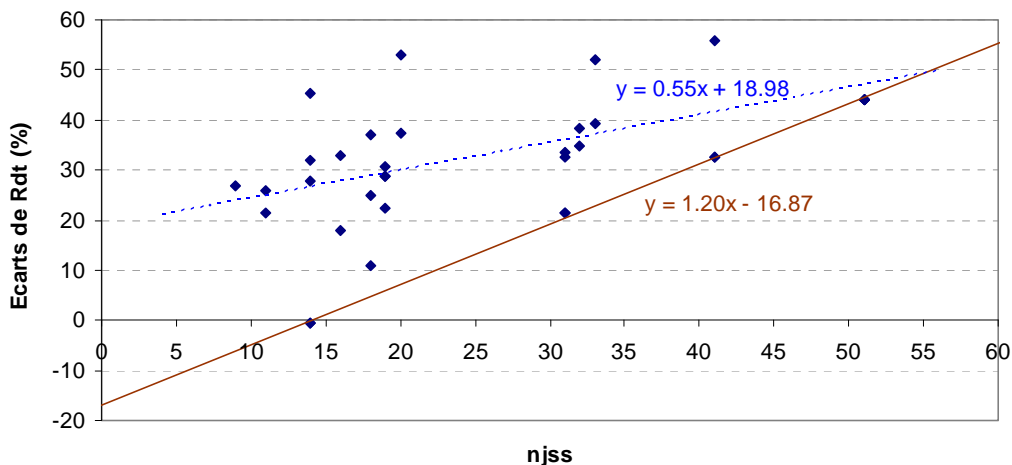
Ces obstacles peuvent toutefois être minimisés dans la mesure où les observations supplémentaires sont requises sur quelques variétés – témoins, et non pas sur l'ensemble des variétés étudiées. A terme, on espère réduire encore ces ajouts en travaillant le plus possible avec des modèles et des observations non destructives. De plus, l'investissement demandé, bien que non négligeable, peut être largement justifié par le gain important d'information que la mise en œuvre d'un diagnostic apporte. Il peut permettre ainsi d'alléger le réseau expérimental et d'en réduire le coût dans la mesure où l'on saura mieux en extraire l'information.

- Une limite actuelle importante à la démarche réside dans le choix des méthodes statistiques employées, qui s'appuient sur des modèles linéaires. Or, les relations entre les variables expliquées (rendement par exemple) et les variables décrivant les facteurs limitants ne sont généralement pas linéaires. De plus, les facteurs limitants agissent rarement seuls, et ils présentent le plus souvent un effet de seuil. Par exemple, la figure 11 illustre la relation observée entre les pertes de rendement et le nombre de jours secs autour du stade épi 1cm, dans le réseau blé INRA 1995-1997 (variété Soissons). La droite inférieure du graphe représente les situations qui n'ont été affectées que par la sécheresse au

début de la montaison. Les autres points, situés au-dessus de cette droite, représentent des situations où d'autres facteurs limitants ont également agi. Il n'y a donc que 3 situations (sur 29 au total dans ce réseau) qui n'ont été affectées que par la sécheresse au début de la montaison. La prise en compte correcte de l'effet de ce facteur limitant devrait se baser non pas sur la régression linéaire qui passe par le milieu du nuage de points (droite pointillée), mais sur la droite enveloppe, définie ici par 4 points. On constate également que, suivant cette droite, la sécheresse au début de la montaison commence à avoir un effet sur le rendement non pas à partir du premier jour, mais à partir du 15<sup>e</sup> jour. Par suite, tous les points situés à gauche de l'abscisse 15 n'ont pas été affectés par ce facteur.

Des travaux sont en cours pour déterminer le seuil d'effet des différentes variables, et pour établir leurs lois d'action sur la variable expliquée qui, contrairement au cas particulier de la figure 11, sont rarement linéaires. Pour cela, des outils statistiques comme la régression quantile sont mobilisés (Makowsky *et al.*, 2007), celle-ci permettant de modéliser la courbe enveloppe d'une variable d'intérêt par des quantiles (et non pas par l'espérance – ou moyenne) en fonction d'une ou plusieurs variables d'entrée (Arnaud Gauffreteau et Céline Robert, travaux en cours). Un travail d'amélioration des variables descriptives doit également être poursuivi, en recherchant si possible des variables dont l'effet sur les variables expliquées est linéaire.

**Figure 11** : Relation entre les pertes de rendement par rapport à une valeur de référence et le nombre de jours secs successifs autour du stade épi 1cm (jours où la pluie n'excède pas l'évapotranspiration potentielle entre épi 1cm-150 degrés.jours et épi 1cm+350 degrés.jours) ; réseau blé INRA 1995-1997, variété Soissons.



- La régression linéaire multiple introduit des biais importants quand il y a des corrélations entre variables. On risque de prendre un effet à la place d'un autre, simplement parce qu'ils sont corrélés entre eux. La méthode de choix des variables présente également une grande sensibilité à la base de données utilisée (nombre et nature des essais) : le retrait ou l'ajout d'un essai peut facilement faire disparaître ou apparaître un facteur limitant (Prost, 2008). Là aussi, des travaux sont en cours pour remédier à ce problème, notamment en travaillant par mélange de modèles (Prost *et al.*, 2008), ce qui améliore la stabilité de l'estimation des paramètres liés à chaque facteur limitant. Cette méthode est actuellement très lourde en temps de calcul, mais il existe des algorithmes de simplification qui sont en cours de test (Gauffreteau, travaux en cours).

- La méthode impose de mettre en place des variétés témoins – révélateurs sur lesquelles sont effectuées un certain nombre d'observations approfondies. Pour simplifier la démarche, on peut se demander si la caractérisation des milieux ne pourrait pas s'appuyer sur toutes les variétés étudiées (par exemple, la pression de maladie pourrait être estimée non sur des variétés - témoins, mais par la moyenne des notes obtenues sur l'ensemble des variétés). Dans le même ordre d'idées, on peut se

demander si la caractérisation des milieux ne pourrait pas utiliser directement les variables descriptives des milieux pour décrire l'intensité des facteurs limitants.

Le recours aux variétés – témoins est encore indispensable pour l'identification de certains stades, tant qu'on n'aura pas des modèles fiables pour estimer ces critères, et pour la caractérisation du statut azoté. Mais on peut espérer que le développement de modèles et de méthodes non destructives permettra de se passer des témoins révélateurs. Par ailleurs, la pondération des facteurs limitants par leur effet sur les plantes représente une sécurité, car pour la plupart des facteurs limitants on ne connaît pas bien la relation entre les pertes de rendement et la valeur de la variable qui les décrivent, et on ne maîtrise pas bien non plus l'effet des interactions entre facteurs limitants (l'effet de deux facteurs n'est pas en général la somme des effets individuels de ces facteurs). Mais, pour apporter une réponse à cette question, nous envisageons de comparer la structuration d'un réseau expérimental obtenue sur la base des variables descriptives des facteurs limitants avec celle qui est obtenue sur la base des contributions des facteurs limitants aux diminutions de rendement sur les géotypes révélateurs.

## Conclusion

Les réseaux d'évaluation variétale recèlent une très grande quantité d'information largement sous-utilisée, en terme de connaissance des milieux et des variétés. Pour les variétés en particulier, on peut aller beaucoup plus loin que ce que les catalogues actuels présentent, en estimant les sensibilités variétales pour tous les facteurs limitants qui peuvent affecter la production. Cette démarche est intéressante notamment pour les facteurs limitants dont on n'observe pas facilement les effets. De plus, dans un contexte de réduction des intrants et d'augmentation probable de certaines contraintes environnementales, il va devenir de plus en plus essentiel de connaître les aptitudes variétales par rapport à ces contraintes. Pour les gestionnaires de réseaux expérimentaux, cette approche rend possibles des économies et des optimisations grâce à une structuration plus optimale des réseaux. Ces bénéfices pourraient encore être amplifiés par une meilleure articulation des réseaux expérimentaux entre eux (notamment aux étapes de l'inscription et de la post-inscription), les références acquises sur un plus grand nombre d'années pouvant ainsi être regroupées et valorisées.

Cette méthode d'analyse a reçu le nom de DiagVar<sup>®</sup>. Elle a fait l'objet du développement d'un logiciel - pilote (Prost et Lecomte, 2005) et est en cours de développement sur un support informatique grand public (Gauffreteau, travaux en cours). Plusieurs travaux associant des organismes publics et privés visent à la développer également sur d'autres espèces (maïs, tournesol, pois, pommes de terre, betteraves).

Pour les différentes espèces concernées, des besoins importants subsistent pour améliorer les variables descriptives des facteurs limitants, alléger leur obtention, et pour cela mettre en œuvre des modèles agronomiques. Les méthodes statistiques employées semblent pouvoir répondre à l'attente des évaluateurs car elles permettent d'approfondir l'interprétation des résultats, et elles apportent puissance et rapidité de traitement. Leur application pose quand même un certain nombre de questions et les améliorations en cours doivent tendre à limiter la part de hasard dans les explications trouvées, et à aller toujours dans le sens d'une meilleure validation agronomique des explications. Ce travail met en évidence l'importance de poursuivre les interactions entre chercheurs agronomes et statisticiens, et avec toutes les personnes concernées par l'évaluation variétale.

## Références bibliographiques

- Brancourt-Hulmel M., Biarnès-Dumoulin V., Denis J.B., 1997. Points de repère dans l'analyse de la stabilité et de l'interaction génotype - milieu en amélioration des plantes. *Agronomie* 17, 219-246.
- Cuccia C., 2008. Changements climatiques observés en Bourgogne depuis 1961 : Etude des variations de températures et de précipitations. Mémoire de Master « Recherche Géobiosphère », CNRS, Université de Bourgogne, 52 p.
- Denis J.B., 1980. Analyse de régression factorielle. *Biom. Praxim.* 20, 1-34.
- Denis J.B., 1988. Two way analysis using covariates. *Statistics* 19, 123-132.
- Dizien C., 2009. Ajuster expérimentation et modélisation: le cas des notations maladies dans les réseaux d'essais variétaux. Mémoire d'ingénieur AgroParisTech, Paris (France), 91p.
- Doré T., Sebillotte M., Meynard J.M., 1997. A diagnostic method for assessing regional variations in crop yields. *Agric. Syst.* 54, 169-188.
- Justes E., Jeuffroy M.H., Mary B., 1997. Wheat, Barley and Durum Wheat. In G. Lemaire (Ed.) *Diagnosis of the Nitrogen Status in Crops.* Springer 4, p.73-92.
- Lecomte C., Giraud A., Aubert V., 2003. Testing a predicting model for frost resistance of winter wheat in natural conditions. *Agronomie* 23, 51-66.
- Lecomte C., 2005. L'évaluation expérimentale des innovations variétales. Proposition d'outils d'analyse de l'interaction génotype - milieu adaptés à la diversité des besoins et des contraintes des acteurs de la filière semences. Thèse de Docteur-Ingénieur de l'INAPG, Paris (France), 174p. + annexes.
- Makowski D., Doré T., Monod H., 2007, A new method to analyse relationships between yield components with boundary lines, *Agronomy for Sustainable Development*, 27, 119-128
- Meynard J.M., David G., 1992. Diagnostic sur l'élaboration du rendement des cultures. *Cahiers Agriculture* 1, 9-19.
- Prost L., Lecomte C., 2005. Cahier des charges pour un logiciel d'analyse des résultats des essais variétaux basé sur le diagnostic agronomique et la régression factorielle. Contrat FSOV INRA-GIE Club des 5-Arvalis 2003-2005. 39 p.
- Prost L., Jeuffroy M.H., 2007. Replacing the nitrogen nutrition index by the chlorophyll meter to assess wheat N status. *Agronomy for Sustainable Development* 27, 321-330.
- Prost L., Makovsky D., Jeuffroy M.H., 2008. Comparison of stepwise selection and Bayesian model averaging for yield gap analysis. *Ecological Modelling* 219. 66-76.
- Prost L., 2008. Modéliser en agronomie et concevoir des outils en interaction avec de futurs utilisateurs: le cas de la modélisation des interactions génotype x environnement et de l'outil DiagVar. Thèse de Doctorat, AgroParisTech Paris (France), 375p.
- van Eeuwijk F.A., Malosetti M., Yin X., Struik P.C., Stam P., 2004. Modelling differential phenotypic expression. In: "New discussions for a diverse planet", Proceedings of the 4th International Crop Science Congress, 26 Sep – 1 Oct 2004, Brisbane, Australia.