

PRESENTATION D'UNE BASE DE DONNEES RELATIVE AUX ITINERAIRES
TECHNIQUES DU COLZA D'HIVER

P. Leterme (1), D. Debouzie (2), D. Wagner (3)

(1) CETIOM, 174 avenue Victor Hugo, 75116 Paris, France

(2) Université LYON I, 43 bld du 11 Novembre 1918, 69622 Villeurbanne, France

Depuis 1984, le CETIOM implante chaque année, dans plusieurs lieux, des dispositifs expérimentaux d'étude des itinéraires techniques. Ces dispositifs correspondent à l'organisation d'une enquête culturale avec contrôle d'une grande partie de l'hétérogénéité du milieu pédo-climatique et permettent d'étudier pour plusieurs variétés l'effet des variations des techniques employées.

La compilation des données caractérisant, à l'échelle de la parcelle, les conditions pédo-climatiques, l'itinéraire technique adopté, l'état sanitaire, l'élaboration du rendement et le rendement de la culture a été faite afin de réaliser une base de données.

Le tableau 1 présente les essais regroupés dans cette base ainsi que les nombres de parcelles. Aujourd'hui, les essais réalisés en 1990 n'ont pas encore été introduits: ils le seront prochainement.

Lieu	Année	Nombre de parcelles
Haute-garonne	1986	64
	1987	88
	1988	88
	1989	155
Indre	1985	104
	1986	68
	1987	86
	1988	102
	1989	159
Meurthe-et-Moselle	1985	98
	1986	110
	1987	91
	1988	97
	1989	155
Cher	1985	112
	1986	72

Tableau 1 : Liste des essais inclus dans la base (février 1991)

Présentation générale de la base de données :

Les 17 essais évoqués ci-dessus représentaient 1797 parcelles au total. Sur ces parcelles ont été mesurées ou évaluées 95 variables relatives à la caractérisation du milieu (sol, climat), du peuplement.

végétal, de l'itinéraire technique pratiqué et de l'état sanitaire. Ce fichier (1797 x 95) a été épuré des données sujettes à caution et, à cause des données manquantes quelquefois trop nombreuses pour certaines variables ou certains essais, ramené à une taille plus modeste.

Aujourd'hui, nous disposons de 1649 parcelles caractérisées complètement par 43 variables décrivant successivement :

- des caractéristiques pédologiques des parcelles dans un essai donné:
 - . la profondeur d'enracinement,
 - . la réserve utile,
 - . la richesse chimique du sol, (notée de 1 à 4),
 - . la teneur en matière organique,
- des caractéristiques climatiques de l'essai:
 - . le nombre de jours avec des températures moyennes inférieures respectivement à -15, -10, 0 et 5°C.
 - . la somme des pluies (en 1/10 mm) entre:
 - le semis et la levée,
 - la levée et le stade F1,
 - les stades F1 et F1+350°Cjours,
 - les stades F1+350 et F1+900°Cjours,
 - le stade F1+900°Cjours et la récolte.
 - . la somme des températures entre:
 - le semis et la levée,
 - la levée et le stade D2,
 - les stades D2 et E,
 - les stades E et F1,
 - le stade F1 et la récolte.
 - . la somme des rayonnements entre:
 - les stades D2 et E,
 - les stades E et F1,
 - les stades F1 et F1+350°Cjours,
 - les stades F1+350°Cjours et F1+900°Cjours,
 - le stade F1+900°Cjours et la récolte.
 - . le rapport ETR/ETM entre:
 - le semis et la reprise de végétation,
 - les stades F1 et F1+350°Cjours,
 - les stades F1+350°Cjours et F1+900°Cjours,
 - le stade F1+900°Cjours et la récolte.
- des caractéristiques de la plante:
 - . nombre de pieds au m² à l'automne,
 - . nombre de ramifications d'ordre 1 par plante,
 - . nombre de siliques sur la hampe principale par plante,
- des variables mesurées à la récolte:
 - . la teneur en huile de la graine (en %),
 - . le rendement (GPS) de la parcelle en q/ha,
 - . le poids de 1000 graines (en g),
 - . une note de verse, de 1 à 9,
 - . une note d'égrenage (78 parcelles).
- des caractéristiques de l'état sanitaire du colza (note de dégâts de 1 à 5):
 - . méligèthes,
 - . baris,
 - . phoma,

- . cylindrosporium,
- . pseudocercospora,
- . charançon de la tige,
- . charançon des siliques,
- . sclérotinia.

Parallèlement à ces variables, les techniques appliquées sont décrites: variété, date de semis, dose d'azote apporté à l'automne, au printemps, fractionnement de l'azote de printemps, régulateur de croissance, protection insecticide, fongicide, apport de soufre, irrigation. Soit en tout, 11 variables relatives à l'itinéraire technique.

Support informatique de la base de données :

Aujourd'hui, les données sont stockées sous deux formes:

- fichier standard aux normes ASCII,
- fichier au normes SAS.

Très prochainement, ces données seront gérées par ORACLE, système de gestion de base de données relationnel entièrement compatible avec SQL, ce qui facilitera la sélection et l'extraction de données particulières.

Types d'utilisation de la base de données :

La masse de données accumulées dans ce système d'information autorise une grande variété d'utilisation.

Actuellement, 3 types principaux d'utilisation peuvent être dégagés.

- utilisation pour la modélisation des fonctionnements du colza d'hiver.

Dans ce cadre, deux modalités peuvent être définies:

- . en temps que base de Test de modèles de fonctionnement élaborés par ailleurs. C'est le cas par exemple des modèles utilisés dans le logiciel COLIBRI (LETERME 1991) qui ont été conçus indépendamment des données de la base et ainsi peuvent être testés et paramétrés grâce à ces données.

- . en temps que corps de données permettant l'élaboration de modèles. Grâce aux données stockées dans cette base, complétées par quelques autres (photopériode), nous avons pu modéliser de façon satisfaisante les variations de dates du début floraison de la variété BIENVENU (LETERME, WAGNER, 1991).

- utilisation comme référentiel fournissant des éléments de décision technique.

Cette utilisation peut concerner la quantification statistique de l'effet d'une technique particulière, que ce soit en termes de rendement ou en termes économiques (marge brute par exemple). A titre d'exemple, le tableau 2 présente, pour 2 variétés, l'effet sur le rendement de l'application d'un régulateur de croissance au printemps. Cet effet est observé en situation de forte fertilisation azotée (>200 kg d'azote) mais en distinguant les situations où une partie de cet azote est apportée vers le stade E.

fertilisation tardive	non			oui		
	Variété	n	écart	test	n	écart
Bienvenu	52	+0.4	1.17 (NS)	48	-0.5	1.14 (NS)
Darmor	90	+2.2	5.04 (p<0.001)	78	+1.6	3.06 (p<0.001)

Tableau 2 : Effet de l'application d'un régulateur sur le rendement. L'écart entre parcelles homologues est calculé par $\text{Rend}(\text{Rég}+) - \text{Rend}(\text{Rég}-)$. n représente le nombre de couples de parcelles.

Les conclusions diffèrent selon les deux variétés: aucun effet pour Bienvenu, gain d'environ 2q/ha pour Darmor qu'il y ait fertilisation tardive ou pas.

Pour Bienvenu, aucune variable enregistrée dans la base (composantes du rendement, notes de verse ou d'attaques) n'est modifiée en moyenne par l'application d'un régulateur.

Pour Darmor, le gain de rendement, de 29.5 à 31.5 q/ha, s'accompagne, ou pourrait résulter d'une diminution de la note de verse (de 2.9 à 2.3; $p < 0.001$), des notes d'attaque par *Pseudocercospora* (de 1.26 à 1; $p < 0.001$) et par le charançon de la tige (de 1.43 à 1.16; $p < 0.001$).

Outre la mise en évidence des différences de comportement variétale, la disponibilité d'un grand nombre de données permet de poser de nouvelles questions: en quoi interviennent, ou bien sont influencées, les caractéristiques d'attaque par *Pseudocercospora* ou par le charançon de la tige. Nous voyons ici la valeur heuristique d'une telle base de données.

A côté des effets sur les résultats techniques ou économiques, les données contenues dans la base permettent de quantifier statistiquement l'impact direct de caractéristiques de l'environnement sur la production (milieu pédo-climatique ou état sanitaire) ainsi que celui des techniques sur cet environnement. Le tableau 3 présente la technique employée pour élaborer la note d'attaque par le charançon de la tige. Elle résulte de la combinaison d'une note de fréquence et d'une note de gravité toutes 2 allant de 1 (nul) à 5 (très élevé).

		Fréquence				
		1	2	3	4	5
Gravité	1	①				
	2		②			
	3			③		
	4				④	
	5					⑤

Tableau 3 : Note globale d'attaque par le charançon de la tige (encerclée sur le tableau). Les cases hachurées sont impossibles.

La régression entre le rendement et la note d'attaque ainsi déterminée a été calculée sur 143 situations de la variété BIENVENU. On obtient:

$$\text{Rend} = -2.4 \times \text{note} + 35.3 \quad r=0.45$$

soit une perte moyenne de 2,4 q/ha par point d'attaque.

Le tableau 4 illustre l'effet de la lutte insecticide sur la note de dégâts du charançon de la tige en comparant des parcelles protégées (Ins+) à des parcelles non traitées (Ins-).

Note d'attaque		1	2	3	4	5	n	Moyenne
Bienvenu	Ins -	57	52	23	9	5	146	1.93
	Ins +	222	49	0	2	0	273	1.19
Darmor	Ins -	26	39	79	12	13	169	2.69
	Ins +	235	20	0	1	0	256	1.09

Tableau 4 : Distribution des notes d'attaque par le charançon de la tige, selon l'application ou non d'un insecticide. n est le nombre de parcelles.

On constate que, globalement, les insecticides sont bien appliqués; néanmoins 18.7% des parcelles Bienvenu (51/273) et 8.2% des parcelles Darmor (21/256) ont été attaquées, en général avec une note égale à 2, donc assez faible malgré tout.

- à côté des utilisations agronomiques dont quelques exemples viennent d'être donnés, la base fournit des données utiles aux statisticiens pour mettre au point et tester diverses techniques de traitement des données. Trois grands types de techniques biométriques sont travaillées au laboratoire de biométrie de l'Université de LYON (DEBOUZIE):

. les techniques de traitements multitableaux comme STATIS (ESCOUFIER, 1980) permettent d'étudier simultanément plusieurs tableaux de données à 2 entrées. Un tableau correspond ici aux résultats de p variables sur n parcelles, la répétition des mesures mêmes différentes au cours du temps fournissant la suite des tableaux.

. les techniques de recherche d'hétérogénéité spatiales internes aux essais: mise en évidence, test de significativité, étude de stabilité.

. les techniques de couplage mathématiques de données de nature différente (tableau de variables qualitatives, de variables quantitatives).

Par ailleurs, les données permettent de procéder à des études d'échantillonnage, comme par exemple GEFFRAY (1985) à propos du nombre de graines par silique.

CONCLUSION :

Dans un futur proche, les données accumulées au cours des expérimentations d'itinéraires techniques du colza d'hiver seront accessibles au public d'une manière aisée (SGBD relationnel ORACLE). Le CETIOM a accompli la même démarche à propos du tournesol et du

soja: la mise en forme et la validation des données est moins avancée pour ces 2 espèces mais à terme le même travail que pour le colza sera effectué. Les praticiens de l'agriculture disposeront alors d'un système d'information complet relatif à la culture des oléagineux.

Références bibliographiques:

ESCOUFIER Y., 1980 - L'analyse conjointe de plusieurs matrices de données

in : Biometrie et Temps, Jolivet et al. Eds, 59-77

GEFFRAY B., 1985 - Etude de la variabilité du nombre de graines par silique chez le colza d'hiver.

Rapport interne Lab. de Biométrie LYON I - CETIOM 13p + figures.

LETERME P., 1991 - Modélisation du fonctionnement du colza d'hiver: utilisation pour la conduite de la culture.

8ème congrès intern. sur le colza - Saskatoon, Canada -9-11/07/91

LETERME P., WAGNER D., 1991 - Contribution à la modélisation de la date de floraison du colza d'hiver.

à paraître dans Info. Techn. CETIOM.