Développement d'un outil d'optimisation des réseaux d'essais et de caractérisation des variétés

Arnaud GAUFFRETEAU

Arnaud Gauffreteau[1], Valérie Cadot[2], Aurélie Mailliard[2], Marie-Hélène Jeuffroy[1], Aurélia Luciani[2], Marianne Cerf[1], Marie Noël Mistou[1], Lorène Prost[1], Axel Olivier[3], Pierre Rigolle[3], Thierry Bebin[4], Denis Beghin[4], Thierry Demarquet[4] , Philippe Lonnet[4] , Odile Vion[4], Jean-bruno Beaufume[5] , Celine Duque[5], Laure Duchalais[6], Eric Margalé[6], Marcel Meausoone[6], Jean-Michel Moreau[6], Sebastien Caiveau[7], Patrice Senellart[7], Lamia Belkacem[3], Caroline Dizien[1], Sabrina Marques[2], Céline Robert[2,3]

(1) INRA, Bâtiment EGER - Campus de l'AgroParisTech, 78850 THIVERVAL-GRIGNON

(2) GEVES, 25, rue Georges Morel CS 90024, Beaucouzé Cedex

(3) GIE Club 5, 83, avenue de la Grande Armée, 75782 Paris Cedex 16

(4) Florimond Desprez, Hameau de Wattines, 59242 Cappelle en Pévèle

(5) Limagrain Verneuil Holding, Ferme de l'Etang - BP 3 - 77390 Verneuil l'Etang

(6) Serasem, 60, rue Léon Beauchamp BP 45, 59933 La Chapelle d'Armentières

(7) Syngenta Seeds, rue de Moyencourt, 78910 Orgerus

Contexte et objectifs

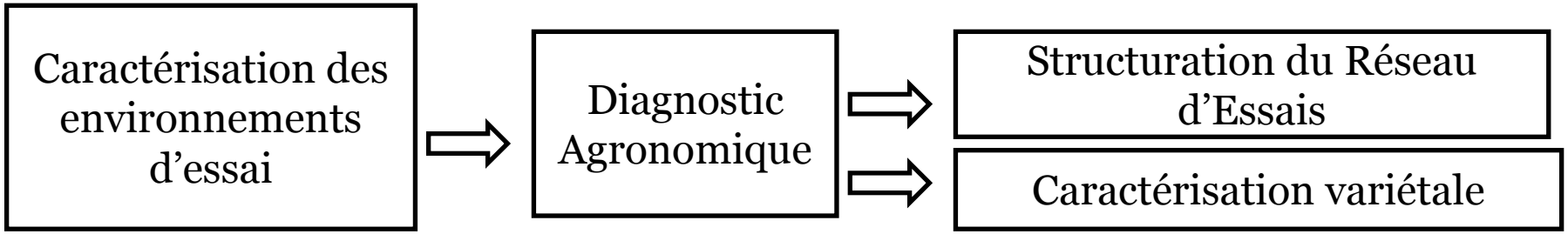
- Evolution climatique + Réduction des intrants

➔Diversification des milieux de culture avec plus grande expression des stress biotiques et abiotiques

➔Nécessité d'adapter le choix variétal à l'environnement de culture

➔Nécessité de caractériser les variétés pour leur résistance aux principaux stress environnementaux

- Diagvar (Lecomte, 2005 & Prost, 2008) : caractériser les variétés par une valorisation le l'IGE s'exprimant dans les réseaux d'essais



➔Faire progresser chaque étape de la démarche

➔Produire un outil informatique opérationnel

Faire évoluer le recueil et le traitement des données expérimentales pour permettre la caractérisation des réseaux d'essais variétaux

- Estimation du stade épi 1 cm
- Prédiction du statut azoté des plantes à partir de mesures SPAD
- Mise au point d'un indicateur fiable de la pression maladie
- Homogénéisation des pratiques de notation des maladies
- Analyse de l'effet de certains facteurs limitants par régression quantile

Estimation du stade épi 1 cm

Contexte : Identification du stade épi 1 cm peu fréquente dans les réseaux d'essais variétaux car mesure fastidieuse et destructive.

Objectif : Tester des méthodes de prédiction du stade épi 1 cm par validation croisée sur des bases de données multi-locales et pluriannuelles.

Méthodes testées		**RMSEP (en degrés-jours)**
Prédiction à partir d'une mesure de taille d'épi (Gate, 1995)	Relation linéaire	<20 si 0.7cm<épi<1.7cm
	Relation exponentielle	<20 si 0.7cm<épi<2cm
Extrapolation des mesures faites sur un essai à tous les essais	∑T° épiaison - ∑T° épi 1cm	88 en moyenne [54,153]
	∑T° épi 1cm / ∑T° épiaison	89 en moyenne [53,163]
Pas de mesures : Relation linéaire entre ∑T° épiaison et ∑T° épi 1cm		72

➔ Mesures nécessaires pour prédire le stade épi 1cm avec une précision inférieure à la semaine

➔ Peu d'extrapolation possible du stade épi1cm d'un essai à l'autre

➔ Réduction possible du nombre de mesures dans chaque essai par prédiction du stade épi 1cm

Homogénéisation des pratiques de notation des maladies

- **3 actions** ont été menées pour améliorer l'homogénéité des notations maladies :

1. **Elaboration de 3 didacticiels sur le site Web d'Endure/ Quantipest** (collaboration GEVES - réseau INRA PIC) pour entraîner les expérimentateurs à se caler sur l'échelle de notation CTPS pour :
   a. ***Septoria tritici* :** 227 photos, 92 questions
   b. **Rouille brune :** 187 photos, 117 questions
   c. **Rouille jaune :** 136 photos, 49 questions

   http://194.94.61.13/quantipest/quantipest/quantipest_utilisateur/index_appli.php?portail=Agrescience&produit=quantipest&main=63&ssrub1=93&ssrub2=94&ssrub3=339&id_fiche=134
2. **Notations collectives au champ et formation sur l'identification des pathogènes** organisées par le GEVES tous les 2 à 3 ans pour les expérimentateurs du réseau VATE

2. **Une réflexion sur le rôle des régionaux GEVES**

Page d'entrainement à la notation rouille jaune

Mise au point d'une méthode opérationnelle d'analyse de l'interaction GxE

- Test de méthodes statistiques pour le diagnostic agronomique
- Adapter une analyse des milieux d'essais
- Test de méthodes statistiques pour l'analyse des IGE

Test de méthodes statistiques pour le diagnostic agronomique

Contexte : Instabilité des résultats du Diagnostic dans Diagvar largement due à une mauvaise gestion des données corrélées par la régression Stepwise

Objectif : Tester la robustesse d'un diagnostic basé sur l'utilisation de la régression PLS

Données : Caphorn cultivé selon 2 ITK sur 9 lieux en 2007 et 2008 (32 situations), les environnements sont décrits par 38 indicateurs

Protocole : 1- 1000 BdD générées par tirage aléatoire avec remise dans la BdD d'origine

2- Bour chaque BdD, réalisation de 2 diagnostics, l'un par Stepwise et l'autre par PLS

3- Comparaison des méthodes selon leur stabilité de sélection et la précision des estimations qu'elles fournissent

Méthodes	Diagnostic stable	Diagnostic instable
Stepwise	61%	11%
PLS	74%	5%

➔ Ecart-types associés aux estimations toujours plus faibles dans le cas d'un diagnostic sur PLS

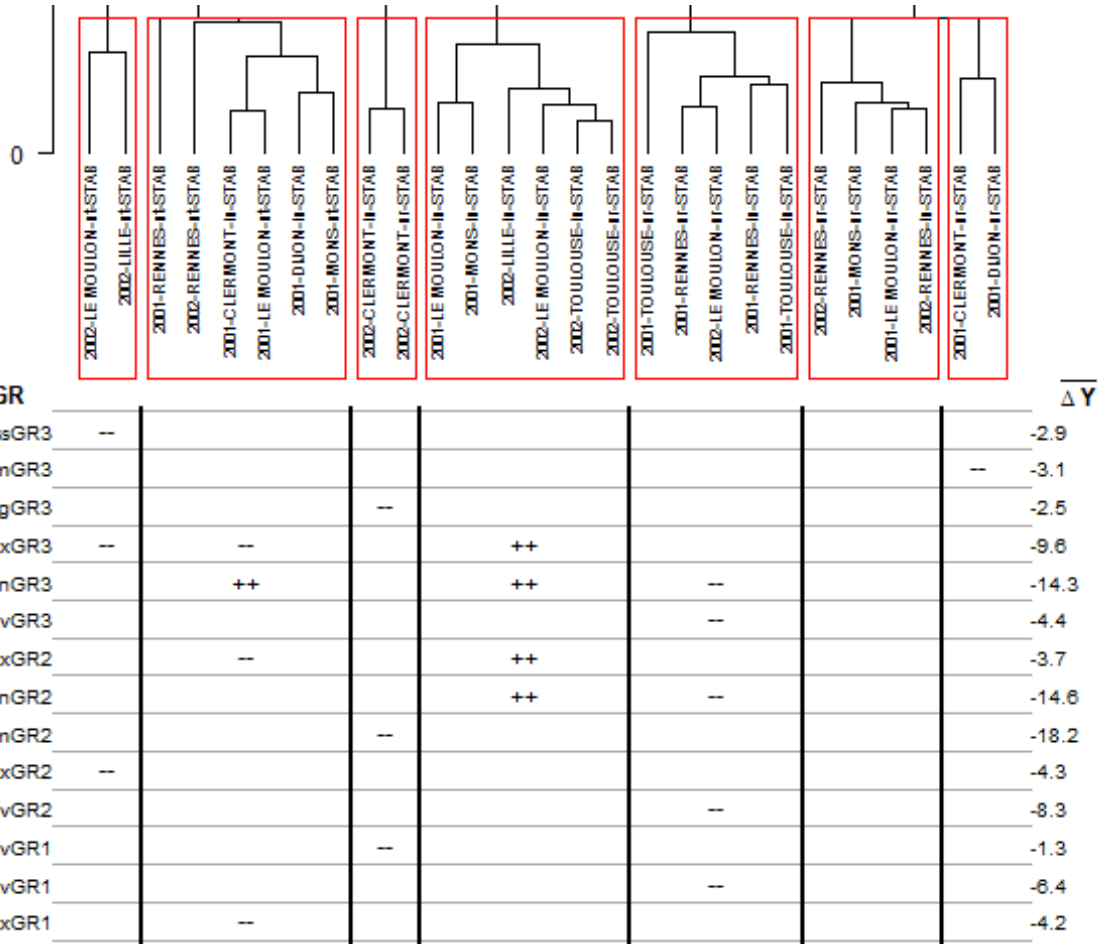
➔ Le rapport entre écart-types et estimations moyennes est plus faible avec la PLS dans 66% des cas

Adapter une analyse des milieux d'essais

Contexte : Suite au diagnostic agronomique, classification des milieux d'essais par CAH.

Objectif : Tester plusieurs méthodes de regroupement et faciliter l'interprétation des groupes

Données : Soissons cultivées dans 36 milieux avec 4 types de conduites, 9 lieux et 6 années



- Classification ascendante hiérarchique par méthode Ward
- Identification pour chaque groupe des facteurs limitants qui ont un effet moyen intra-groupe significativement différent de leur effet moyen à l'échelle du réseau d'essais

Test de méthodes statistiques pour l'analyse des IGE

Contexte : Régression factorielle avec contrainte de signe peu prédictive de l'IGE

Objectif : comparer la régression factorielle avec contrainte de signe avec une méthode en deux étapes couplant un modèle mixte additif et une régression PLS sur la matrice d'interaction

Données : 9 Variétés conduites de façon intensive et intégrée sur 9 sites en 2007 et 2008

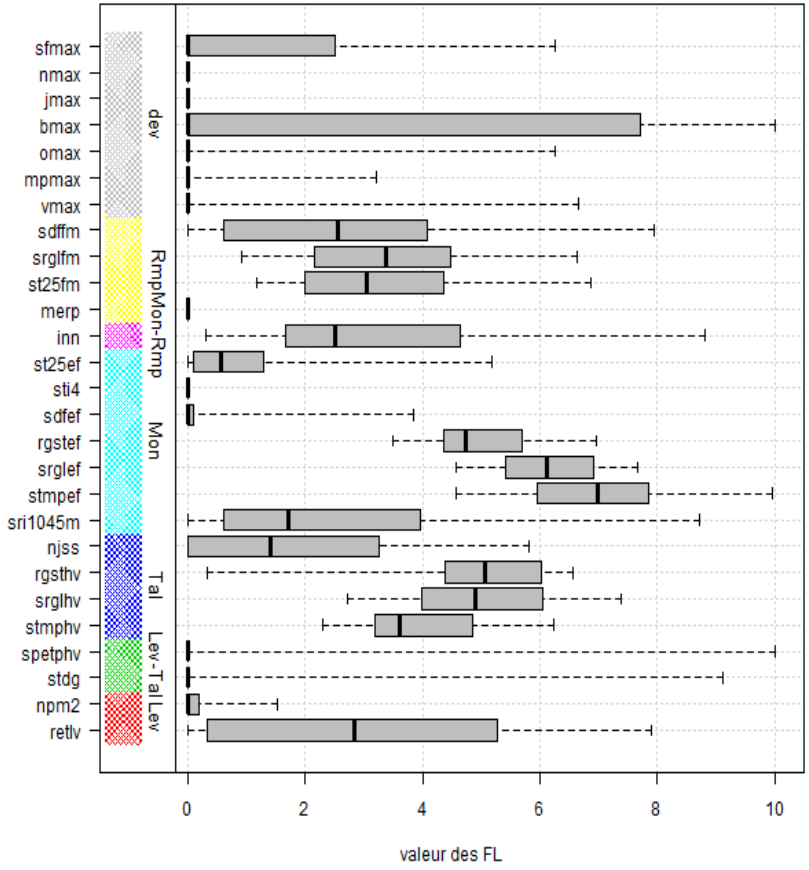
Protocole : 1- Validation croisée sur les environnements . 2- Evaluation de la capacité des modèles à identifier la meilleure variété parmi 2 sur tous les couples possibles. 3- Valeur de référence = valeur fournie par des prédictions à partir d'un simple modèle mixte additif

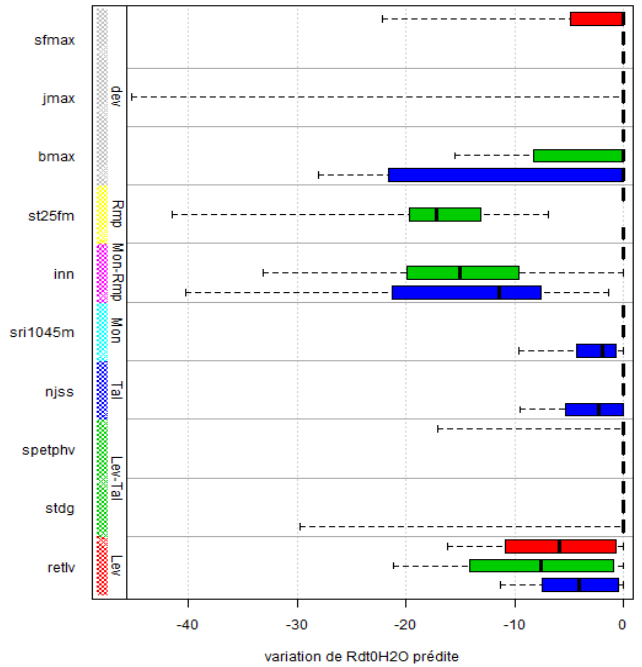
Méthodes d'analyse	Erreur d'identification
Référence	1.22 q/ha
Régression factorielle + contrainte de signe	1.26 q/ha
Modèle mixte + PLS	0.98 q/ha

Développement d'un outil informatique sous R

- Des sorties graphiques enrichies
- Un outil accessible

Des sorties graphiques enrichies





	intercept	bmax	bmax_HP	inn	inn_HP	retlv	sfmax	sfmax_HP	sri1045m	stdg	NS
ISENGRAIN	6.76	-1.24	-0.28	-1.11	-0.35	0.18	1.8	0.88	-0.75	-0.37	
YNOPRECIA	6.68	1.31	0.69	0.38	-0.11	-0.22	-0.66	-0.06	0.43	-0.8	
ORATORIO	4.1	0.26	1.05	-1.23	-0.6	0.62	-0.56	-1.32	0.27	-0.59	
ARCHE	3.1	-0.24	-1.02	0.18	-0.44	-0.55	-0.72	0.04	-0.1	0.16	
RUMBA	2.46	0.63	0.04	0.42	-0.05	0.04	-2.08	-1.51	0.68	0.11	st25fm
SOISSONS	-3.08	-1.39	-0.5	-1.11	-0.39	0.69	1.12	0.26	-0.23	-0.05	
RENAN	-4.02	2.55	1.67	1.2	0.49	-0.18	-0.58	0.27	0.26	0.48	
CAMPREMY	-5.68	0.51	-1.16	1.8	0.46	0.02	-0.65	0.96	0.11	0.99	
RECITAL	-10.31	-2.39	-0.48	-0.53	1	-0.6	2.32	0.48	-0.67	0.07	

Un outil accessible

- Choix du logiciel statistique R (http://cran.r-project.org/)
  - Logiciel gratuit
  - Puissant et riche en méthodes d'analyses
  - Code facile à maintenir par le développeur non informaticien
- Format de l'outil
  - 34 fonctions R
  - Structures compartimentée permettant l'adaptation à de nouvelles espèces et à des fichiers d'entrées non issus du calcul par Diagvar
- Utilisation facilitée
  - Peu de lignes de codes sous R
  - Utilisateur guidé par des boites de dialogues tout au long de la démarche
  - Export de l'ensemble des tables de résultats et des graphiques produits
  - Guide d'utilisation décrivant précisément chaque étape de la démarche

Perspectives

- Retour vers les partenaires pour évaluer et comprendre le niveau d'utilisation de l'outil
- Adaptation de la démarche à d'autres espèces (Pois, Pomme de Terre, Betterave, Tournesol)
- Travail méthodologique pour améliorer l'analyse des IGE (cadre du GIS GC HP2E)
- Travail sur la caractérisation des stress abiotiques (eau et azote) dans les essais variétaux (projet CTPS en partenariat avec le GEVES, l'UFS, Arvalis institut du végétal, le CETIOM, l'ITB et l'INRA).

Prédiction du statut azoté des plantes à partir de mesures SPAD

Contexte : INN : mesure du statut azoté d'une culture précise mais lourde et destructive

Objectif : Tester la possibilité de prédire l'INN à partir de mesures SPAD sur une modalité Bas Niveau d'Intrants dans les essais du réseau CTPS (4 lieux en 2009, 11 lieux en 2010)

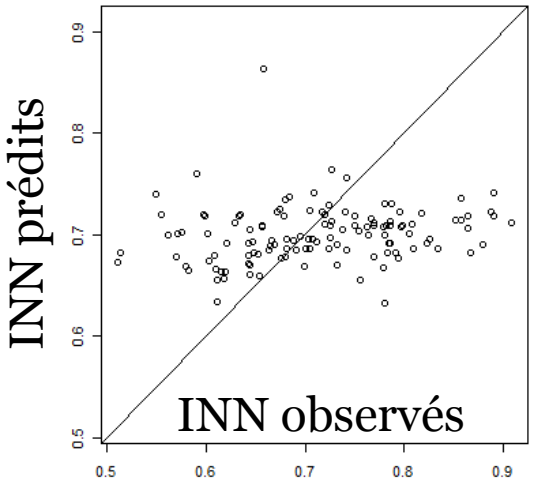
Relation testée : **$INN = \alpha.exp(\beta.X)$**

avec X=SPAD, X=IndexS=SPAD/$SPAD_S$, X=IndexT=SPAD/$SPAD_T$

Impact de la variété (Var) et de l'environnement (Env) sur α et β

		SPAD	IndexS	IndexT
α	Env	0.0000	0.0000	0.0000
	Var	0.2379	0.2033	0.3132
β	Env	0.7661	0.7795	0.7793
	Var	0.0721	0.9296	0.9554

Qualité prédictive du modèle Exemple de l'IndexT



RMSEP = 0.09

Mise au point d'un indicateur fiable de la pression maladie

- 5 indicateurs du niveau de pression biotique pour 4 pathogènes (**septoriose, rouille brune, rouille jaune et oïdium**) ont été comparés selon leur qualité prédictive

Indicateurs testés
I1=Moyenne des notes d'attaque sur l'ensemble des variétés de l'essai
I2= Moyenne des notes d'attaque des témoins de l'essai
I3= Note d'attaque maximale sur l'essai
I4= 5 + (note de sensibilité du témoin sensible dans l'essai – note de sensibilité théorique du témoin sensible)
I5= Moyenne des notes d'attaques sur le quart des variétés les plus attaquées de l'essai

Bioagresseurs	Méthode de validation croisée	I1	I2	I3	I5
Septoriose	Coefficient corr. Pearson	0.76	0.74	0.67	0.73
	% corrél. significatives	77,1	72,8	65,8	71,3
Rouille brune	Coefficient corr. Pearson	0.77	0.73	0.61	0.72
	% corrél. significatives	81,3	80,1	63,1	78,8
Oïdium	Coefficient corr. Pearson	0.72	0.69	0.54	0.68
	% corrél. significatives	49,4	44,5	25,8	43,3
Rouille jaune	Coefficient corr. Pearson	0.57	0.49	0.28	0.52
	% corrél. significatives	35,3	35,6	10,8	32,8

➔Le premier indicateur s'avère être le plus prédictif

Analyse de l'effet de certains facteurs limitants par régression quantile

Contexte : Les seuils d'action de certains stress abiotiques (T°, RG…) sont mal connus et difficiles à établir à partir d'essais non factoriels où plusieurs stress agissent simultanément

Objectif : Tester la possibilité d'établir ces seuils par régression quantile – Exemple du stress haute température à montaison et remplissage

Données : 552 situations : 10 variétés et 200 environnements intégrant 10 années

Protocole : 1- décompte du nombre de jours à montaison et remplissage où la température maximale est comprise entre : 19 et 20, 21 et 22, 23 et 24, 25 et 26 °C

2- Pour chaque intervalle : réalisation de régressions quantiles (τ=0.80, τ=0.85 et τ=0.90)

3- Identification des pentes négatives dont l'IC à 95% ne contient pas 0

	τ = 0.8			τ = 0.85			τ = 0.9		
	Coef.	Borne inf.	Borne sup.	Coef.	Borne inf.	Borne sup.	Coef.	Borne inf.	Borne sup.
19-20°C	-0.45	-0.84	0.03	-0.65	-1.36	0.14	-1.34	-1.77	0.81
21-22°C	1.19	0.35	1.48	0.59	0.22	1.79	0.88	-0.66	1.31
23-24°C	-3.23	-3.70	-1.04	-2.46	-4.49	-0.63	-2.34	-6.79	-0.50
25-26°C	-3.83	-4.34	-3.33	-4.61	-5.11	-1.58	-3.91	-7.43	-1.08