



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITE DE LORRAINE

Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires (ENSAIA)

Ecole doctorale Ressources Procédés Produits Environnement (RP2E)

Laboratoire Réactions et Génie des Procédés (LRGP)

THESE

Présentée à l'UL par

Budour AJIB

en vue d'obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Spécialité : Procédés Biotechnologiques et Alimentaires

**Contribution à la modélisation de la qualité de l'orge et
du malt pour la maîtrise du procédé de maltage**

Soutenue publiquement le 18 décembre 2013 devant la commission d'examen

Rapporteurs:

M. SINNAEVE Georges	Directeur de recherche	CRA-W Belgique
----------------------------	------------------------	----------------

M. SAULNIER Luc	Directeur de recherche	Angers-Nantes
------------------------	------------------------	---------------

Directeur de thèse:

M. FICK Michel	Professeur	ENSAIA-UL
-----------------------	------------	-----------

Co-directeur de thèse :

M. FOURNIER Frantz	Maitre de conférences	ENSAIA-UL
---------------------------	-----------------------	-----------

Invitées :

M. BOIVIN Patrick	Directeur de recherche	IFBM
--------------------------	------------------------	------

Mme PONS Marie-Noëlle	Directeur de recherche	ENSIC
------------------------------	------------------------	-------

Remerciement

Ces travaux ont été réalisés à l'Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires (ENSAIA) à Nancy, dans le Laboratoire Réactions et de Génie de Procédé (LRGP). Cette thèse a été réalisée en collaboration avec l'IFBM (Institut Français des Boisson de la Brasserie Malterie).

Je remercie messieurs Gabriel WILD et Laurent FALK, Directeurs successifs du Laboratoire Réactions et de Génie de Procédé de m'avoir accueilli au sein de ce laboratoire.

Je remercie vivement M. Michel FICK (Directeur de l'Ecole Nationale Supérieure d'Agronomie et des Industries Alimentaires ENSAIA) pour avoir accepté d'être mon directeur de thèse.

Je remercie sincèrement M. Frantz FOURNIER (Maitre de conférences à L'ENSAIA) d'avoir accepté d'être mon co-directeur de thèse. Je le remercie également pour sa présence, ses qualités humaines et ses précieux conseils.

Je remercie M. Ivan MARC M. Xavier FRAMBOISIER et tous les gens de plateforme SVS pour m'avoir accueilli parmi eux pendant une partie de ce travail.

Je remercie M. Patrick BOIVIN (Directeur scientifique de l'IFBM) et M. Marc SCHMITT pour m'avoir accueillie dans leurs locaux et pour m'avoir permis de réaliser la partie expérimentale de mon travail. J'associe à ces remerciements tous les personnels du laboratoire de la production à l'IFBM en particulier Nathalie OUARNIER- OLLOSIO, Erwan TROALEN, Catherine COLIN, Aude, Sandrine BAGNON pour leur présence et leur gentillesse.

Je remercie Mme Valérie GUILLARD (maitre de conférences université Montpellier 2) et Mme Valérie Lullien-Pellerin (Directeur de recherche à l'INRA Montpellier) de m'avoir accueilli dans leur laboratoire pour les études de porosité de l'orge.

Enfin je tiens à remercier toutes les personnes que j'ai côtoyé durant mes années de thèse au LRGP, en particulièrement Paola, Evelyne, Fabrice, Arnaud, Mme Terry WANGER, Mme Isabelle CHEVALOT, Mme Annie MARC.

Je remercie également mes amies, Laial, Kenza, Caroline, Claudia, Tatiana, Françoise, Mathilde, Amel, Florent, Abdelkrim, Bam Bou, Frédérique. pour tous les moments précieux qu'on a passé ensemble.

Je tiens à remercier le gouvernement Syrien, et l'université d'AL BAATH d'avoir m'offrir cette chance d'être là aujourd'hui et de réaliser mon rêve de suivre des études supérieures à l'étranger.

A mon père et ma mère, mes deux anges. Papa tu es une vrai école de la vie, je ne cesse pas d'apprendre tous les jours avec toi. Maman, une femme adorable qui n'a jamais cessé de me couvrir de tendresse. Pour votre dévouement incommensurable, soutien et encouragement pour que je donne le meilleur de moi-même, je ne vous remercierais jamais aussi

Remerciement

A mes frères et sœurs et leurs enfants et à ma belle-famille malgré la distance vous étiez toujours là pour moi. Chacun d'entre vous a son caractère différent mais avec vous tous j'avais tout ce dont j'ai besoin. Vous êtes toujours là pour me conforter, conseiller, faire rire, rêver, encourager.

A ma bien aimée sœur qui nous a quittée très tôt mais pas ses rêves. J'espère que tu es fière de moi là-haut où tu es.

A mes amies syrien ici et en Syrie, merci pour votre soutien, votre encouragement et pour tous les moments précieux que on a passé ensemble.

Enfin, Je remercie mon cher époux pour son soutien quotidien indéfectible et son enthousiasme contagieux à l'égard de mes travaux comme la vie général. Notre couple a grandi en même temps que mon projet scientifique, le premier servant de socle solide à l'épanouissement du second. Je remercie donc mes deux bouts d'chou Hadi et Adam qui ont ajouté un gout précieux dans ma vie.

Introduction générale	25
De l'orge au malt	31
1. Le marché mondial de l'orge, de malt et de la bière	31
2. Le procédé de maltage	32
2.1.1. La trempe.....	33
2.1.2. La germination	33
2.1.3. Le touraillage.....	34
3. L'orge brassicole	36
3.2.1. Embryon :	38
3.2.2. Endosperme :	38
3.2.3. La couche a aleurone :	38
3.2.4. Les glumelles.....	39
3.2.5. Le péricarpe testa.....	39
3.2.6. La paroi externe :	39
4. Méthode d'analyses.....	44
4.2.1. La spectroscopie infrarouge (IR)	45
5. Modélisation : caractérisation de l'orge et du malt, relation orge & malt	49
5.2.1. Choix d'un modèle-hypothèse.....	52
5.2.2. Fonction paramétrée pour la modélisation boîte noire	53
Matériel et méthodes.....	72
1. Transformation et analyses conventionnelles de l'orge et du malt	72
1.3.1. Mesure de l'humidité.....	74
1.3.2. Teneur en protéine de l'orge	74

Table de matière

1.3.3.	Taux de β -glucan	74
1.3.4.	Taux d'arabinoxylane.....	75
1.3.5.	Indice de germination (IG) :	75
1.3.6.	Pourcentage des grains piqués (GP) :	75
1.3.7.	Indice de trempe (test de Chapon) :	75
1.3.8.	La mesure de la porosité.....	76
1.3.9.	La mesure de la dureté des grains.....	77
1.4.1.	Taux de friabilité du malt:	77
1.4.2.	Extrait fine mouture:.....	78
1.4.3.	Taux de protéines solubles :	78
1.4.4.	La viscosité du mout.....	78
1.4.5.	Atténuation limite :	78
1.4.6.	Pouvoir diastasique :	79
2.	Méthodes d'analyses spectrales de l'orge	80
2.1.1.	Principes de spectroscopie infrarouge:	80
□	Mesures en transmission et en réflexion.....	82
□	Réflexion totale atténuée.....	83
2.2.1.	Déformation des spectres de réflexion	85
2.2.2.	Prétraitement des signaux MIR	86
Matériels et méthodes Analyses et modélisation des données (MLR, ACP, PLS)		89
1.	Méthodes d'analyse des données et de régression	89
1.1.1.	La modélisation par régression linéaire multiple MLR.....	90
1.1.1.1.1.	Le criblage (screening).....	93
1.1.2.	Méthode d'optimisation.....	95
2.	Les analyses multivariées: Analyse en composants principales (ACP)	100
3.	La régression en composantes principales (PCR)	105
4.	La régression en moindre carrée partielle (PLS)	106

Table de matière

1. Modélisation linéaire multivariable de la qualité du malt	113
2. Détermination et validation de conditions optimales de maltage à faible humidité	117
Les résultats de l'optimisation.....	118
3. Mise en évidence du découplage entre la composition de l'orge et l'imbibition des grains d'orge.....	122
4. Nouveaux facteurs orge pour améliorer les modèles de qualité du malt	124
4.1.1. La relation porosité-humidité des grains d'orge.....	125
5. Conclusion.....	126
Grain quality parameters	137
Malt quality parameters.....	138
Bibliography chapter 5	154
Abstract.....	160
Références chapitre 6.....	176
Conclusion	182
Perspectives	185

Liste des figures

Figure 1 schéma de processus de maltage	33
Figure 2 Structure d'un grain d'orge en germination et les processus biochimiques qui se produisent au cours de la partie "modification" du processus de maltage	34
Figure 3 à droite orge de deux rangs vs. A gauche orge de six rangs	37
Figure 4 Représentation schématique d'un grain d'orge en coupe longitudinale (adapté de Lewis et Young, 1995).	38
Figure 5 disposition des différents tissus d'un grain d'orge (Briggs, 1998)	40
Figure 6 des domaines d'application de MIR spectroscopie sur des produits alimentaires	Erreur ! Signet non défini.
Figure 7 des domaines d'application de MIR spectroscopie sur des produits alimentaires.....	48
Figure 8 Les étapes de modélisation statistique	52
Figure 9 les étapes de construire une plan d'expérience	55
Figure 10 protocole de touraillage	73
Figure 11 le protocole de mesure de la porosité d'orge a droite	76
Figure 12 le spectre électromagnétique.....	80
Figure 13 Représentation de vibrations des liaisons moléculaires (molécule triatomique). a. vibrations d'élongation. b. vibrations de déformation	81
Figure 14 principe de mesure a) en réflexion b) en transmission	82
Figure 15 Principe de mesure en réflexion totale atténuée	83
Figure 16 Schéma de principe d'un spectromètre FT-IR	84
Figure 17 Exemple de spectre FT-MIR de l'orge dans la gamme 1500 et 700 cm^{-1}	85
Figure 18 prétraitement des spectres a) centrage, b) normalisation SNV, c) dérivé première, d) dérivé second.....	88
Figure 19 Présentation de la matrice d'expérience	91
Figure 20 Schéma monte les étapes d'évaluation d'un algorithme génétique	96
Figure 21 Fonction objectif pour l'optimisation par domaine	99
Figure 22 Optimisation du maltage en utilisant l'algorithme génétique.....	100

Liste des figures

Figure 23 : description schématique d'une matrice de donnée pour une ACP	101
Figure 24 Présentation d'un nuage de point dans un espace de trois dimensions.....	102
Figure 25 Illustration de la rotation des axes pour la mise en place d'un nouveau repère	103
Figure 26 le carte factorielle d'un nuage de points	104
Figure 27 Présentation d'une régression PLS sur trois dimensions	106
Figure 28 schéma représente la corrélation facteurs-orge et qualité-malt par une modèle polynomiale	Erreur ! Signet non défini.
Figure 29 Localisation des points de mesure lors du procédé de maltage	114
Figure 30 Modèle de l'humidité en fin de trempe pour l'orge d'hiver 2006 Coefficient du modèle et intervalle de confiance (avec un niveau de confiance de 95%).....	115
Figure 31 Modèle de l'humidité en fin de trempe pour l'orge d'hiver 2006 : Coefficient de détermination entre données expérimentales et modélisées.....	116
Figure 32: Résultats du test de Bootstrap. Distribution des coefficients du modèle HFT/OH	117
Figure 33 Suivi du taux d'humidité pendant les phases de trempe à 16 °C pour quatre variétés d'orge : Azalea,	122
Figure 34 Location of the 16 cultivation site of barley in France	137
Figure 35 PCA plot on barley exploratory data from 2010 parameters without friability. a) Score plot □ winter barley ● spring barley. b) Loading plot. PCA was applied on original data. With 1 principal component 99% of variance total was explained.....	147
Figure 36 PCA score plot and loading plot on spring barley exploratory data from 2010. Parameters without friability. a) Score plot site A ● site B ● site C ● site D ○ b) loading plot. PCA was applied on normalized data [0-1]. With 3 principal components 80% of variance total was explained	148
Figure 37 PCA a) score plot present barley classification by site of cultivation: Site A ● site B ● site C ● site D ○. b) Loading plot on spring barley exploratory data from 2010 parameters without friability and diastatic power. PCA was applied on original data. With one principal component 90% of variance total was explained.	148
Figure 38 PCA a) score plot on barley spectra in 2D b) score plot in 3D 2009 harvest ■ 2010 harvest ■ 2011 harvest ■. PCA was performed on second derivative spectra in the interval [800-1000] cm ⁻¹ . PCA was applied on original data. Two principal components explain 87% of total variance	149

Liste des figures

Figure 39 a) PCA score plot on barley 2009 distinguishing □ winter barley ● spring barley. Two principal components explain 91% of the total variance.	150
Figure 40 a) PCA score plot on barley 2010 □ winter barley ● spring barley. Two principal components explain 89% of the total variance.....	151
Figure 41 a) PCA score plot on barley 2011 □ winter barley ● spring barley. Two principal components explain 87% of the total variance.....	151
Figure 42 PLS model quality index (coefficient of determination) vs. principal components. For spring barley of 2010 crop, with 30 principal components, PLS regression was applied on centered data. (o) coefficient of determination for PLS calibration models, (*) coefficient of determination for PLS cross validation models	169
Figure 43 FT-IR spectra in the interval 800-1700 with the regions of barley (protein, β-glucan, and arabinoxylan) signals.....	170
Figure 44 PLS criteria plot for β-glucan models of calibration from spring barley 2010: a) the percentage of variance explained in blue for spectra (X) and in red for the response β-glucan content y. b) the mean square error for spectra X model and β-glucan content y model.....	172
Figure 45 Comparison of fitted β-glucan content obtained by PLS regression vs. Observed β-glucan content for spring barley 2010. 10 principal components were used on centred spectra in the interval [800-1200] cm ⁻¹ , with $R^2_v=0.86$	173
Figure 46 Fitted protein by PLS regression vs. observed protein from spring barley 2010. With 10 principal components, centred spectra, in the interval [1300-1700] cm ⁻¹ , with $R^2_v=0.85$	174

Liste des tableaux

Tableau 1 classement des 10 premiers payes en production de la bière (Kirin Institute of Food and Life style Report Vol. 30, 2010).....	31
Tableau 2 la composition chimique du grain d'orge (Guiga, 2006)	40
Table 4 quelques exemples sélectionnés pour l'utilisation de MIR-spectroscopie.....	46
Tableau 5 Diagrammes de trempe pour les échantillons d'orge de quatre années de récolte..	73
Tableau 6 les intervalles des facteurs (procès et orges) et les intervalles des mesures en fin de trempe et du malt utilisés pour la normalisation des données dans les modèles.....	91
Tableau 7 le choix des dominances lors de la génération d'un nouvel individu.....	98
Tableau 8 indicateurs de qualité de malt à mesurer	114
Tableau 9 Intervalles de qualité pour un malt considéré comme satisfaisant	118
Tableau 10 Conditions optimales pour le maltage à faible humidité pour (OP).....	119
Tableau 11 Conditions optimales pour le maltage à faible humidité pour (OH)	120
Tableau 12 les nouvelles propriétés de l'orge 2006, mesurées en 2009 et les diagrammes de trempes proposés par l'optimisateur	120
Tableau 13 présente l'humidité fin de trempe et la friabilité du malt résultent d'application de diagramme de trempe optimale sur les 4 orges 2006 et celles prédites pour la modèle.	121
Tableau 14 les variétés d'orge récolte 2008 qui ont choisi pour réaliser la suivi d'humidité avec leurs teneur en protéine, leurs teneurs d'humidité en fin de trempe et la friabilité de malt résultant.	122
Tableau 15 les constants de cinétique de prise d'eau pour les quatre variétés d'orge dans la première phase de trempe.....	123
Tableau 16 Intégration de la porosité comme facteur du modèle MLR : coefficients significatifs du modèle pour chaque facteurs et coefficients de corrélation	125
Tableau 17 Répétition des mesures de porosité et du taux d'humidité après 30 jours de stockage en chambre froide (-12°C) pour 9 échantillons d'orge 2008	126
Table 18 mean and average of barley and malt data from 2009-2010-2011.....	140
Table 19 the best intervals and spectra pretreatments.....	145
Table 20 Classification capacity of PCA on normalized data with all parameters and all but one parameters. (+) parameter with significant influence (0) parameter without effect on barley classification, (ND) parameter with unclear effect	146

ANNEXE 1 les paramètres de qualité de 40 orges récolte 2006, pour cinq variétés : deux de printemps (Scarlett, Sebastian) et trois d'hiver (Carafe, Azurel, Esterel).....	189
ANNEXE 2 les paramètres de qualité de 40 malt des orges récolte 2006.....	190
ANNEXE 3 les paramètres de qualité du malt et d'orge d'hiver récolte 2009	191
ANNEXE 4 les paramètres de qualité du malt et d'orge de printemps récolte 2009	192
ANNEXE 5 les paramètres de qualité du malt et d'orge de printemps récolte 2010	193
ANNEXE 6 les paramètres de qualité du malt et d'orge d'hiver récolte 2010 (1/2)	194
ANNEXE 7 les paramètres de qualité du malt et d'orge d'hiver récolte 2010 (2/2)	195
ANNEXE 8 les paramètres de qualité du malt et d'orge de printemps récolte 2011 (1/2)	196
ANNEXE 9 les paramètres de qualité du malt et d'orge de printemps récolte 2011 (2/2)	197
ANNEXE 10 les paramètres de qualité du malt et d'orge d'hiver récolte 2011	198
ANNEXE 11 les mesure de humidités et de la porosité de 48 orges de printemps récolte 2007-2008.....	199
ANNEXE 12 Liste de 72 variétés d'orge utilisées dans cette étude et cultivés en France	200
ANNEXE 13 la carte et le nom des 16 lieux de cultiver pour les échantillons d'orge qu'on a utilisé dans cette étude	201
ANNEXE 14 Algorithme de NIPALS pour calculer la PCA	202

Liste des annexes

ANNEXE 15 L'algorithme pour le calcul de la régression PLS1 est présenté ci-dessous. Les composantes principales sont calculées successivement jusqu'au nombre souhaitée:.....	203
ANNEXE 16 Compilation des coefficients des modèles significatifs et des coefficients de déterminations pour l'orge de printemps.....	204
ANNEXE 17 Compilation des coefficients des modèles significatifs et des coefficients de déterminations pour l'orge d'hiver.....	204

Liste des abréviations

a_i	Coefficients du modèle linéaire ($i \in \{1, \dots, p\}$)
$\hat{A}$	Vecteur des coefficients estimés du modèle linéaire
A	Le vecteur des coefficients calculés du modèle linéaire
AR_1	Durée du premier sous-air ou <i>Air-Rest</i> (h)
Ara	Teneur en Arabinoxylanes (mg/l)
$AttLim$	Atténuation limite du malt (%)
BG	β _glucane (%)
E	vecteur des erreurs entre mesures expérimentales et prédites
$J(x)$	Fonction objectif à optimiser
F	vecteur de p fonctions qui modélise la relation entre la réponse et les facteurs
$Fria$	Friabilité du malt (%)
GA	Hormone acide gibbérellique
GP	Taux des grains piqués (%)
HFT	Humidité de l'orge fin de trempe (%)
HF	Humidité du malt vert (%)
$H_j + 1$	Humidité après un jour de germination (%)
H^*	La valeur cible de l'optimisation
H(x)	la valeur calculée la mesure
H1	le teneur en eau du moût
IG	Indice de germination de l'orge(%)
OH	Orge d'hiver
OP	Orge de printemps
P	Teneur en protéine initiale dans l'orge (%)

Abréviations

P_{dias}	Pouvoir diastasique du malt (WK sur sec)
P	Teneur en extrait dans le moût
R^2	Le coefficient de corrélation
T	Température (c°)
$U(t)$	Les entrées
$Visco$	viscosité (% malt sec)
x_{max}, x_{min}	La valeur maximum et minimum du facteur x
x_N	Facteur normalisé
x	valeur du facteur non normalisé
$X(t)$	Vecteur de l'état
X	Matrice des facteurs
$Y(t)$	Sorties du modèle
Y_{exp}	Réponse expérimentale
Y_m	Réponse estimée
Y	Le vecteur des réponses
WR_1	Durée du premier sous eau (h)
WR_2	Durée du deuxième sous eau (h)
α	Un nombre aléatoire
ε	Erreur entre la réponse calculée et la réponse estimée
σ_y^2	Variance de la réponse
σ_{ai}^2	Variance du coefficient ai

Chapitre 1
INTRODUCTION

Introduction générale

La consommation de la bière continue son ascension à travers le monde, emmenée par les marchés émergents pour atteindre un nouveau record. Le malt constitue la matière première principale de la fabrication de la bière et conditionne sa qualité finale. Répondre à cette augmentation de la consommation implique une augmentation des besoins en malt brassicole et donc en orge de brasserie de qualité.

Sur les marchés mondiaux de la bière, la véritable opportunité que peuvent saisir les brasseries actuellement est de différencier leurs produits en matière de goût. Il ne faut pas non plus oublier de mentionner le marché des bières sans alcool ni celui des boissons à base de bière. Ces produits sont essentiels pour les brasseries et sont la réponse de l'industrie aux nouveaux comportements des consommateurs. Les automobilistes et les sportifs sont prêts à renoncer à l'alcool, mais pas à abandonner la bière. Et la jeune génération aime toujours faire de nouvelles expériences en matière de boissons et de saveurs. La bière fascine, elle offre de multiples facettes, elle est dans l'air du temps. En cela, les artisans brasseurs sont des précurseurs et s'inspirant de cette tendance. La marge de manœuvre est encore grande en ce qui concerne les matériaux bruts ou le processus même du brassage.

Les tendances actuelles de la malterie sont donc de mieux maîtriser et contrôler le procédé afin de répondre aux exigences accrues du client et aux particularités des cahiers des charges établis. Les malteurs ont ainsi besoin de produire du malt de qualités physico-chimiques et sanitaires rigoureuses.

Une démarche de recherche et d'optimisation du procédé de maltage est un moyen de se différencier de la concurrence par une meilleure connaissance du procédé. Cette maîtrise passe par une limitation de la variabilité de la qualité du malt produit. Pour obtenir un malt de qualité homogène, on doit cependant être en mesure de prendre en compte la variabilité des variétés d'orge utilisées et celle du procédé de transformation appliqué.

Le maltage se déroule en trois étapes principales: la trempe, la germination et le touraillage. L'orge est tout d'abord trempée pour augmenter sa teneur en eau, puis elle subit une phase de germination pour développer les principales enzymes responsables de la désagrégation de l'endosperme et de l'amidon. Enfin, l'orge est touraillée pour arrêter la croissance du germe, développer les arômes et la couleur de la bière finale et aboutir à un produit stable. La trempe est une opération essentielle dont la conduite déterminera le déroulement de toutes les opérations ultérieures au maltage. Ainsi, l'optimisation de cette opération est-elle tout à fait justifiée dans la maîtrise de la qualité du malt.

L'impact de la trempe sur la qualité du malt peut être essentiellement résumé en deux types de facteurs: (i) les conditions opératoires de la trempe: la température de l'eau et la durée des phases sous-eau et sous-air, (ii) les facteurs de qualité liées à la matière première : les teneurs en eau, en protéines, en β -glucanes et arabinoxylanes et l'indice de germination. Pour contrôler la qualité du malt il est indispensable d'étudier et de comprendre l'impact de ces facteurs sur le malt. La prévision et la quantification des aptitudes et des cinétiques caractérisant. L'étude des effets de tous ces facteurs sur la qualité des produits du maltage est cependant très difficile en raison du nombre de facteur et de leurs interactions. Dans ce cadre,

L'utilisation de modèle est un bon outil pour représenter ces relations puis en déterminer les conditions optimales.

En termes de matière première, les conditions optimales signifient une qualité optimale de l'orge. En fait, il n'y a aucune orge parfaite pour un *bon* malt. En effet, la composition de la matière première et sa structure sont liées aux divers facteurs tels que l'année de récolte, les espèces de l'orge, le nombre de rangs dans l'orge, le site de la culture et la variété d'orge. La caractérisation de la matière première (l'orge) devient donc une étape clé pour mieux adapter les conditions opératoires de maltage de sorte qu'on maîtrise la qualité du malt et donc de la bière. Traditionnellement, une analyse de la qualité complète comprend mesurant de 4 à 7 paramètres physiques et chimiques de l'orge. Les méthodes conventionnelles impliquent des procédures longues, laborieuses et coûteuses, y compris l'analyse physique et chimique ou l'analyse de microscopie. Une méthode peu coûteuse, rapide et robuste basée sur la spectroscopie pour le processus de certification serait idéale pour la caractérisation de la qualité d'orge avant le procédé de maltage. Les méthodes de traitement des données classiques (régression) sont inadaptées pour traiter les données spectrales. En effet, elles contiennent habituellement un grand nombre d'information souvent très corrélées entre elles. Pour extraire les informations pertinentes des spectres, on doit faire appel à des techniques plus sophistiquées comme les méthodes d'analyses multivariées. Les méthodes multivariées les plus utilisés sont les analyses par composantes principales (ACP) et la régression des moindres carrés partiels (PLS). Ces méthodes permettent notamment de distinguer des échantillons apparemment semblables, en vue d'une classification ou de quantifier des propriétés difficiles à mesurer par prédiction.

Dans ce contexte, ce projet a pour double objectif de

1. Maîtriser la qualité de l'orge en proposant des conditions (durées et température) de trempage à faible humidité qui garantissent une qualité de malt acceptable. Il s'agit d'identifier l'impact des facteurs liés à la qualité de l'orge et des facteurs opératoires dits *facteurs process* sur la qualité finale des malts faiblement hydratés.
2. Caractériser l'orge en utilisant des techniques spectrales dans le but de mieux maîtriser la relation entre la matière première et le produit final, indépendamment de l'espèce, de la variété, de l'année de récolte, du lieu de récolte.

Ce travail a été réalisé au sein du Laboratoire Réaction et Génie de Procédé (LRGP), en collaboration avec les malteurs de France et l'IFBM (Institut Français des Boissons, de la Brasserie et de la Malterie).

Le document s'articule autour de 7 chapitres dont ce chapitre d'introduction et un chapitre conclusion et perspectives.

Le chapitre 2 est consacré à une étude bibliographique approfondie. Elle contribue à une meilleure connaissance des connaissances actuelles tant au niveau de la matière première, l'orge que du produit final, le malt. Elle décrit également les techniques d'analyse spectrales

de l'orge et les traitements du signal associés disponibles à ce jour. Enfin, elle présente les approches numériques d'analyse et de modélisation exploitables dans notre contexte.

Le chapitre 3 présente l'ensemble des techniques utilisées. Il est scindé en trois parties. Une première partie décrit les matériels et méthodes expérimentales et spectroscopiques appliqués à l'évaluation de la qualité de l'orge et du malt. Une seconde partie décrit les outils de modélisation et d'optimisation utilisés pour construire les plans d'expériences optimaux, définir les modèles et pour déterminer les conditions optimales de la trempe. Une troisième partie présente les approches numériques nécessaires aux traitements des grands ensembles de données (ACP, PLS)

Les chapitres 4, 5 et 6 présentent les principaux résultats obtenus dans cette étude.

Le chapitre 4 est consacré aux résultats obtenus par régression multilinéaire pour établir une relation entre les paramètres (orge et *process*) et la qualité du malt, dans le but de maîtriser la qualité du malt final en condition d'hydratation faible de l'orge. Il présente également les différentes hypothèses formulées pour essayer de consolider les modèles et améliorer l'approche précédente.

Le chapitre 5 est présenté sous forme d'un article. Cet article s'appuie sur un ensemble de données de caractérisation de l'orge et du malt obtenu pour un échantillonnage composé de 394 orges récoltés sur trois années. Il détaille les analyses des résultats expérimentaux par Analyse en Composantes Principales (ACP) réalisés à la fois sur les données expérimentales classiques puis sur les données spectroscopiques de l'orge. Il permet d'étudier l'impact de l'espèce, le lieu de culture, le cultivar ou la variété de l'orge et l'année de récolte sur la qualité d'orge brassicole.

Le chapitre 6 est également présenté sous la forme d'article et s'inscrit dans la continuité du chapitre précédent. Les données spectrales mentionnées ci-dessus ont été exploitées pour construire des modèles de régression de type *moindres carrés partiels* (ou Partial Least Square regression). Il discute en particulier des modèles pour la prédiction de la teneur en bêta-glucanes, en protéines et en arabinoxylanes dans l'orge.

La chapitre 7 résume les résultats obtenus et présente les principales conclusions et perspectives de ce travail.

Chapitre 2
SYNTHESE
BIBLIOGRAPHIQUE

De l'orge au malt

1. Le marché mondial de l'orge, de malt et de la bière

Les premières traces de La bière remonte à 2800 avant J.-C en Mésopotamie à Sumer. Des tables d'argile de l'époque témoignent de la présence de cette boisson appelée alors Sikaru.

La bière, boisson tant appréciée dans le monde entier, est le fruit de nombreuses heures de travail. C'est à partir de la sélection des meilleurs grains d'orge que le brasseur va produire un malt de haute qualité. Après un brassage rigoureux et une longue fermentation, la boisson que nous connaissons sous le nom de bière sera conditionnée en fûts et bouteilles puis sera vendue aux quatre coins du monde, pour le plus grand plaisir de nombreux consommateurs.

La consommation de bière continue son ascension à travers le monde, emmenée par les marchés émergents pour atteindre un nouveau record, d'après les dernières tendances dégagées par le département recherche du brasseur japonais Kirin en s'appuyant sur les chiffres de 2010. D'après les calculs des chercheurs publiés en décembre 2011, 182,69 millions de kilolitres de bière ont été consommés en 2010, ce qui représente une augmentation de 2,4% par rapport à l'année précédente et constitue un nouveau record pour la 25e année consécutive. Et cette ascension devrait se poursuivre d'après les chercheurs. Tableau 1 présente les dix premiers pays producteurs de bière dans le monde.

Tableau 1 classement des 10 premiers pays en production de la bière (Kirin Institute of Food and Life style Report Vol. 30, 2010).

Pays	Production Kilolitre	Part de la Production sur le Marché mondial %
Chine	44,830	24,4
Etats-Unis	22,816	12,3
Bresil	12,600	6,8
Russie	10,240	5,5
Allemagne	9,568	5,2
Mexique	7,988	4,3
Japon	5,850	3,2
Grande-Bretagne	4,499	2,4
Pologne	3,390	1,8
Espagne	3,337	1,7

Répondre à cette évaluation en consommation implique une augmentation des besoins en orge de brasserie de qualité et donc en malt. La production mondiale d'orge de brasserie, qui se

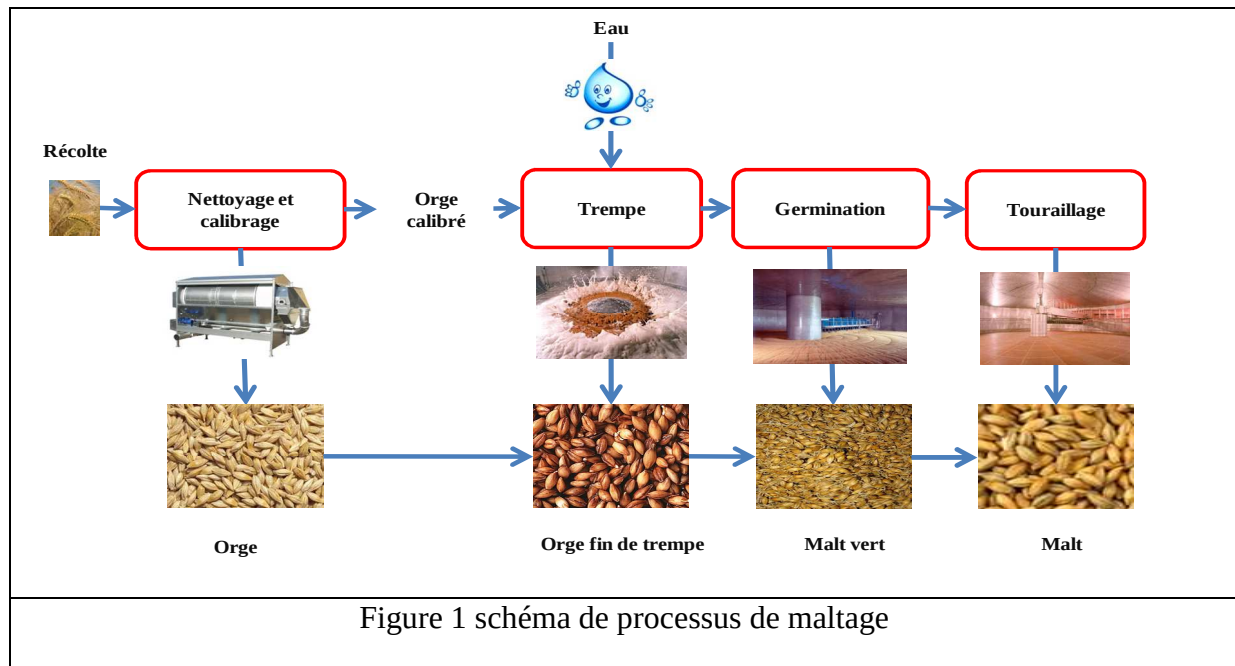
située aux alentours de 20 Mt, est essentiellement destinée à être transformée en malt puis en bière. Or la consommation mondiale de bière est en pleine expansion.

Pour obtenir une bière de qualité une matière première (malt) de bonne qualité est indispensable. Or la qualité de malt est conditionnée par deux éléments : 1) la matière première l'orge, 2) les conditions opératoires de maltage. Ce mémoire a fait l'objet de trois parties de recherche consécutives dont les objectifs peuvent être résumés comme suivant :

- ✓ Dans la première partie d'étude chapitre 4 : on utilise la régression multilinéaire pour :
 - identifier des influences significatives des conditions opératoires (facteurs processus) et des critères quantitatifs de l'orge (facteur matière) sur l'humidité finale de trempe.
 - Optimiser le procédé de maltage, pour arriver à un malt à faible humidité, et trouver des conditions opératoires compatibles pour tous les indicateurs (procédé- matière).
- ✓ Dans la deuxième partie d'étude chapitre 5 : l'analyse en composantes principales a été employée sur les données expérimentales et spectrales d'orge brassicole pour :
 1. Analyser et classer l'orge brassicole par année de culture et par espèce.
- ✓ La dernière partie est consacrée à la régression par les moindres carrés partiels :
 - Prédire et classer les paramètres de qualité d'orge brassicole (Protéines, β -glucanes et arabinoxylanes)

2. Le procédé de maltage

Le malt est l'ingrédient de base utilisé dans la production de la bière. Le malt fournit la majeure partie de l'amidon, les enzymes et les protéines qui sont nécessaires pour donner à la bière sa teneur en alcool et sa saveur distinctives. La qualité de malt est, donc, essentielle. Figure 1 présente un schéma du processus de maltage. L'orge est d'abord nettoyée et calibrée. Tous les grains ayant 2.5 mm et plus sont alors prêts à être maltés. Le processus de maltage comprend plusieurs étapes, les grains d'orge matures sont imprégnés ou trempés pour leur permettre de s'imprégner de l'eau. Ensuite, les grains sont étalés à germer, au cours de laquelle l'amidon à l'intérieur de l'endosperme sera hydrolysé par les α -amylases où l'embryon commence à se développer. Les grains germés, qui montrent un système racinaire bien développé et ont des acrospires, ce qui représente environ 75% de la longueur du grain, sont ensuite touraillés. Le malt modifié et touraillé fournit la matière première pour la fermentation. Le procédé de maltage permet de passer d'une matière première dont les réserves sont difficiles à exploiter pour la fabrication de la bière à une matière intermédiaire "le malt"; qui est farineux et très adapté à la fermentation pour obtenir une bonne boisson à la fin.



2.1. Les étapes de maltage

2.1.1. La trempe

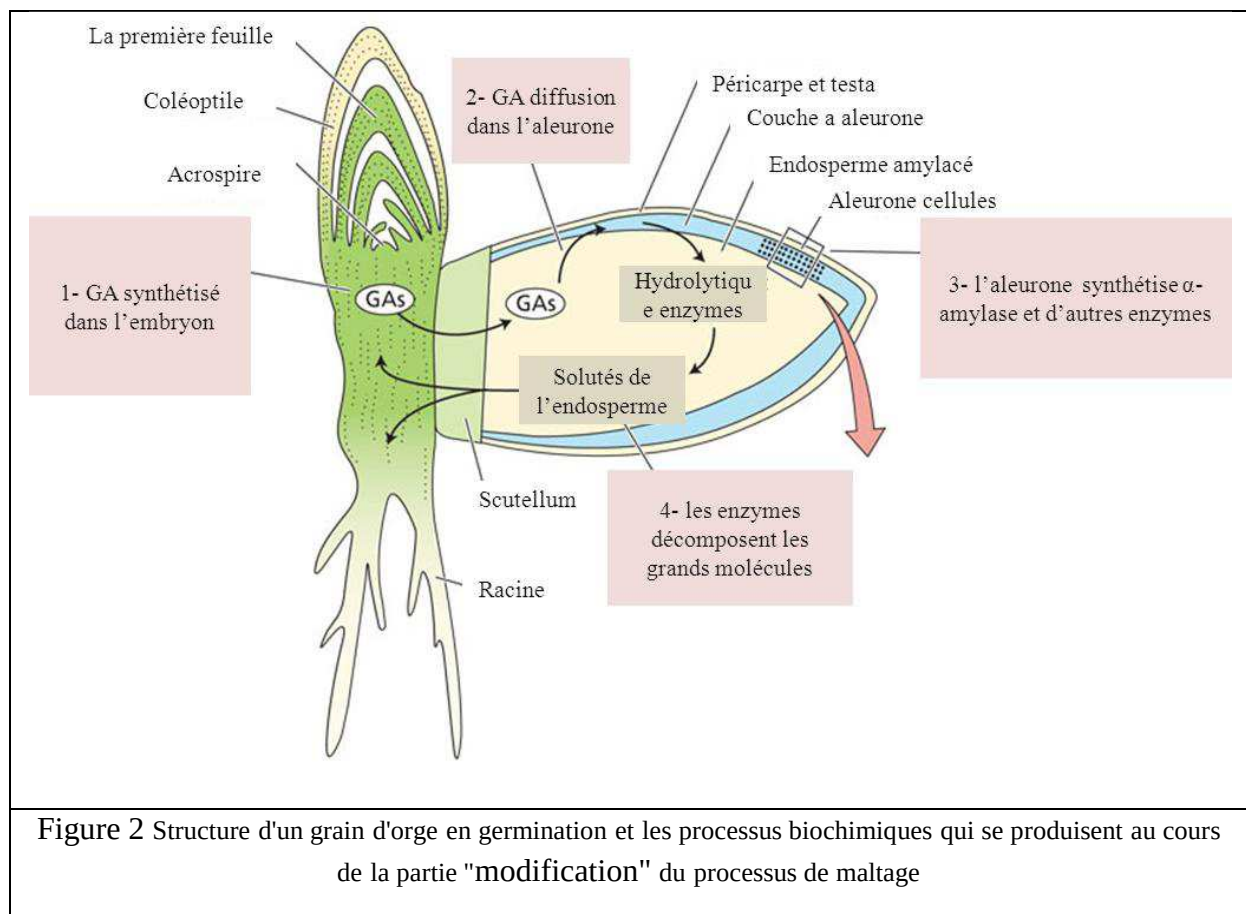
Le but de trempe est d'hydrater uniformément la masse de l'endosperme et de permettre une croissance uniforme pendant la germination. La trempe commence par mélanger les grains d'orge avec de l'eau pour augmenter le niveau d'humidité et activer les processus métaboliques du noyau en sommeil. La trempe est composée d'une séquence de phases dites sous-eau (water-reste WR) et sous-air (air-reste AR). Pendant les phases sous-eau, les grains sont immergés dans l'eau pour assurer la diffusion d'eau à l'intérieur des grains. Pendant les phases sous-air, les cuves de trempe sont vidées et aérées pour fournir l'oxygène nécessaire au grain et éliminer le CO₂ (Hariri, 2003).

L'eau passe d'abord dans l'embryon, ensuite il se répand dans les autres parties de grain. L'absorption d'eau dépend du temps de trempe, de la température, du volume du grain, de la variété, et des propriétés physiologiques. L'absorption d'eau est élevée au début de la trempe. Il augmente avec l'augmentation de la température de l'eau, et la diminution de taille des grains. La variété d'orge et l'année de culture ont également un effet important sur l'absorption d'eau (Manger, 2010). Pendant cette étape, les enzymes sont donc synthétisées et la migration commence, conduisant à la deuxième étape: la germination.

2.1.2. La germination

C'est une étape très importante étant donné qu'elle est à l'origine des principales modifications de grain pendant le maltage (Debourg, 1998). La principale intention de la

germination est la fabrication des enzymes. Figure 2 présente les différentes modifications qui se déroulent pendant la germination. Pendant la germination le noyau de germination développe radicules et un acrospire. Puis (1) l'embryon de grain excrète un facteur de croissance (Gibbérellique acide GA) qui (2) s'oriente vers l'aleurone et (3) induit la formation des enzymes telles que (α -amylases et dextrinases). Les β amylases existent déjà dans l'endosperme. (4) Ces enzymes sont nécessaires pour dégrader les grandes molécules aux simples consistantes dans l'endosperme et les faire migrer vers l'embryon en profitant de l'eau absorbée. Les enzymes les plus importantes du malt sont les β -glucanases et les amylases qui sont responsables de la dégradation de l'amidon. La quantité d'amylase dépend de la quantité d'eau absorbée. La formation des enzymes est plus rapide avec une température de trempe et de germination plus élevé, (Hjmaner ,2004).



2.1.3. Le touraillage

Le touraillage est appliqué sur l'orge germée pour éviter une transformation excessive et limiter les pertes. Obtenir un malt stable et conservable requière la réduction de la teneur d'eau de 40% à 5%. Cette étape arrête tous signe de vie et d'activité enzymatique dans le grain.

Au cours du touraillage un courant d'air chaud passe dans le malt pour éliminer l'eau. La qualité du malt est fortement liée au temps et à la température de touraillage (Manger, 2010). Enfin, le résultat est une matière intermédiaire "le malt", qui est en état stable et réservable?. Le malt doit être farineux et adapté à la fermentation. Le malt est la matière essentielle dans la fabrication de la bière. La qualité du malt qui dépend de plusieurs facteurs est un élément clé pour la qualité de la bière.

2.2. Qualité de malt

La qualité du malt est le résultat de beaucoup de facteurs, s'étendant de la qualité d'orge, à l'hygiène des installations, aux contrôles du processus?. La production mondiale d'orges brassicoles, qui est approximativement de 25 MT, est prévue pour être transformée en malt et puis en bière. Cependant, la consommation mondiale de bière est en pleine expansion (2 milliards de hl) (IFBM) et les malteurs cherchent constamment un meilleur contrôle du procédé de maltage, afin d'obtenir la meilleure qualité du malt possible. La qualité de malt est essentiellement liée à deux genres de facteurs: (i) les conditions opératoires du procédé de maltage: trempage (la température, diagramme de trempage), germination (la température, durée de germination), touraillage (les températures, durée de touraillage), (ii) qualités de matière première (orge) (contenu d'humidité, teneur en protéines, teneur en β -glucanes, et teneur en arabinoxylanes, index de germination, volume du grain, variété, année de culture, lieu de culture et des propriétés physiologiques). Pour commander la qualité de malt, il est nécessaire d'étudier et de comprendre l'impact de ces facteurs sur le malt.

En ce qui concerne les facteurs opératoires, le contrôle du maltage est un vrai défi en raison de divers paramètres opératoires pendant le maltage. La température de l'eau de trempage est très importante. Elle influence la durée de cette phase. Une trempage à 10-12 C° nécessite une prolongation de la durée de cette phase (Hudson, 1986). Plus la température de l'eau de trempage est élevée, plus la durée de trempage est longue et plus la vitesse d'absorption d'eau dans les grains est rapide (Chapon, 1960). Une température de trempage à plus de 20C° et la germination à 15C° provoque un choc, limite la vitesse et l'efficacité de la formation des enzymes pendant la germination (Narziss *et al.* 1970). La trempage chaud réduit le développement des enzymes responsables de la solubilisation des protéines de stockage et des parois cellulaires (Baxter *et al.* 1980). La diminution de l'oxygène consommé par l'embryon le rend incapable de respirer. Il favorise la fermentation, les concentrations de CO₂ et d'alcool augmentent et deviennent toxiques (Hough, 1985, Manger, 2010). Powell *et al.* (1993) ont montré l'effet positif de l'aération pendant la trempage sur presque tous les paramètres de malt. Elle (l'aération?) accélère l'activité physiologique de la germination, favorise le développement des enzymes. L'activité respiratoire de l'orge colonisée par la microflore augmente avec le niveau d'oxygène de l'air au cours de germination (Van Campenhout, 1999). Des études menées par Hariri (2003) ont montré qu'il n'y a pas un effet significatif de l'aération de la trempage sur la qualité de la trempage, une augmentation de la friabilité est constatée entre un système non aéré et un système aéré. Des corrélations entre le programme de trempage (la température, trempage ??? et périodes d'aération) et la qualité de germination ont été trouvées par Hariri, (2003) et Holmberg *et al.* (2005). Une augmentation de la durée de

germination favorise la désagrégation du grain et augmente l'activité enzymatique (Narziss, 1969). Une germination bien aérée encourage un développement plus précoce et plus importante des enzymes (Manger., 2010). Le temps et la température de germination affectent le développement des radicules et les pertes, plus l'air est chaud et la durée de germination est longue, plus les racines se développent et les plus les pertes deviennent importantes (Manger., 2010). La consommation d'oxygène est de 2 à 3 fois plus rapide en germination qu'en trempé (Dalshtrom *et al.* 1963). Les études ont prouvé qu'une teneur plus élevée en eau (trempé pendant plus longtemps) et une plus longue germination influencent l'extrait de malt et diminuent la filtrabilité Grujic *et al.* (2005). La température et la durée de touraillage ont un rôle important sur la qualité du malt et de la bière. Un touraillage à 50C° a un effet favorable sur la stabilité des saveurs dans la bière (Manger, 2010).

Bravi *et al.* (2012) ont étudié l'influence du procédé de maltage entier sur le contenu de lipides dans le malt. La trempé et la germination ont été montrés de fournir des conditions optimales pour la dégradation de lipide, tandis que le touraillage a été révélée n'avoir aucune influence. Mais, pour davantage de recherche sur la control de la qualité de malt, il n'est pas suffisant d'étudier seulement le procédé de maltage. La structure biologique de l'orge et les transformations complexes qu'elle subit pendant le procédé de maltage doit également être prise en considération.

Le contrôle du processus de maltage peut être réalisé en adaptant les conditions (température, durée) des trois étapes de maltage aux caractéristiques de qualité de l'orge. De ce fait, l'orge devient le facteur clé à étudier pour pouvoir contrôler et mieux mener la processus de maltage. L'orge est la matière première essentielle pour la production de bière. La qualité de l'orge a une influence décisive sur la qualité du produit final. La connaissance des propriétés de la matière première et de ses effets sur les processus et sur le produit final fournit la base pour sa manipulation et son traitement. Avec une telle connaissance il est possible de commander le processus technologique consciemment.

3. L'orge brassicole

Il est possible de faire de la bière à partir de toutes les céréales contenant de l'amidon (blé, avoine...). Cependant, les brasseurs du monde entier préfère l'orge pour son haut rendement en sucre, mais surtout parce que son grain est 'vêtu', c'est-à-dire que les enveloppes y adhèrent encore après le battage. Cette enveloppe pailleuse protège le germe pendant le maltage (Hariri, 2003 ; Manger, 2010).

3.1. Histologie du grain d'orge

La classification de l'orge peut être résumée comme suivent :

Règne	Végétal
Embranchement	Anthophytes
Classe	Monocotylédones
Sous-classe	Commelinidae
Ordre	Cyperales
Famille	Poaceae
Genre	<i>Hordeum</i>

Les brasseurs utilisent deux grandes espèces d'orge brassicole; les orges à deux rangs (*Hordeum distichum*), dont seul l'épillet central est fertile et les orges à six rangs (*Hordeum hexastichum*), dont tous les épillets sont fertiles. Généralement, les orges à six rangs sont des orges d'hiver (6RH) et les orges à deux rangs sont des orges de printemps (2RP). Mais il existe aussi des orges d'hiver à deux rangs (2RH) et des orges de printemps à six rangs (6RP). La Figure 3 présente les orges à deux et à six rangs.



Figure 3 A gauche orge de deux rangs vs. à droite orge de six rangs

3.2. Composition structurelle de grain d'orge

La structure physique du grain joue un rôle important sur les modifications de grain, et donc sur l'efficacité du maltage et la qualité du malt final. Les grains sont différents en taille et en structure?. Ce manque d'uniformité est une des sources principales de l'hétérogénéité du malt. Le grain d'orge a une structure complexe qui consiste en trois parties essentielles : embryon, endosperme, grain enveloppes Figure 4.

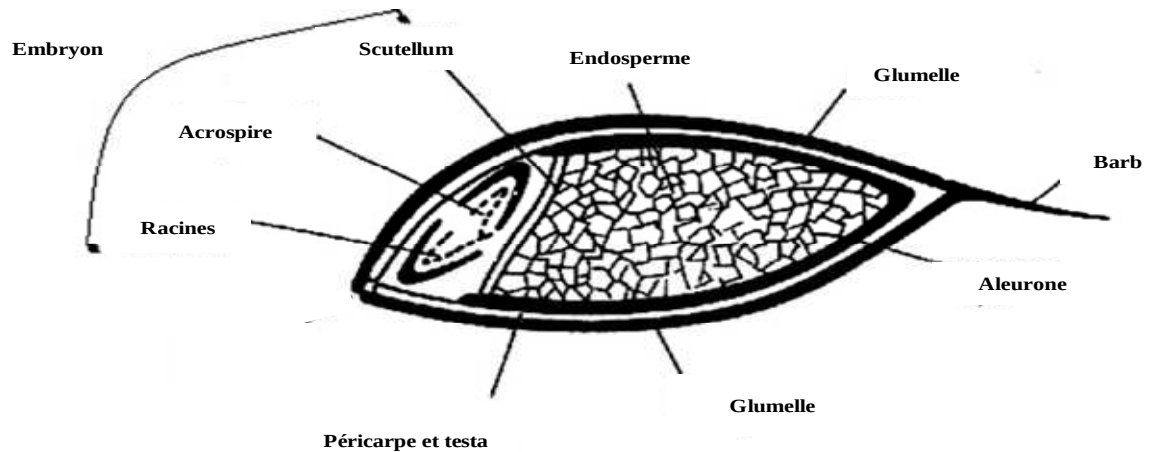


Figure 4 Représentation schématique d'un grain d'orge en coupe longitudinale (adapté de Lewis et Young, 1995).

3.2.1. Embryon :

Il se situe à la surface dorsale du grain. Il y a une couche amylacée limitée extérieurement par plusieurs assises successives. Le scutellum qui sépare l'embryon de l'endosperme permet la diffusion des enzymes hydrolytiques de l'embryon vers l'endosperme. Il n'y a pas de relation entre le développement de l'endosperme et l'embryon, de ce fait ils sont dissemblables en composition chimique. L'embryon contient un substrat respiratoire qui est le saccharose, des lipides et des protéines. Au cours du maltage, il produit des hormones (GA) qui traversent la couche extérieure du grain (l'aleurone). Les enzymes stimulées par les hormones vont être formées et sécrétées de l'aleurone vers l'endosperme (Hariri, 2003).

3.2.2. Endosperme :

Il est constitué de cellules mortes remplies de grands et petits granules d'amidon entourés de protéines. Cette matrice de protéines peut être dense ou absente complètement (Manger, 2010). L'endosperme forme la réserve d'amidon du grain. Une hydrolyse acide ménagée de cet organe met en évidence 3 niveaux de résistance (zone amorphe tendre, zone amorphe plus dure, zone cristalline acido-résistante). Les grains sont classés en fonction de la différence structurelle au niveau de l'endosperme.

3.2.3. La couche à aleurone :

La couche à aleurone renferme, dans sa partie la plus externe, des protéines, et dans sa partie interne, des lipides (30 %), des protéines (20 %), de l'acide phytique, des vitamines du groupe B, de la cellulose et des pentosanes, mais pas d'amidon. La couche à aleurone est constituée de 3 ou 4 couches de cellules vivantes riches en matières azotées et en globules lipidiques. Pendant le développement du grain, ces cellules vont se différencier. Les plus internes vont

permettre la synthèse de protéines . Les plus externes sont les cellules à aleurone. Les cellules à aleurone renferment de petits organes subcellulaires riches en protéines, acide phytique et sucres. Elles sont spécialisées dans la synthèse, lors de la germination, d'enzymes hydrolytiques qui vont solubiliser les polymères dans l'endosperme (Allosio-Ouarnier, 1999).

3.2.4. Les glumelles

Les glumelles protègent la plumule qui va se développer au contraire d'autres grains qui n'ont pas d'enveloppes. Les glumelles sont très riches en cellulose et hémicellulose. La glumelle dorsale possède une paroi épaisse alors que la glumelle ventrale a une paroi plus fine.

3.2.5. Le péricarpe testa

Le péricarpe testa est séparé de l'enveloppe par une couche cireuse et est soudé à la testa par une couche de cellules à parois lignifiées. Il est composé majoritairement de cellulose 20% et contient également des protéines 6% et des lipides 0,5%. Le testa est constitué de 2 bandes superposées composées de matériaux cireux et lipidiques. Le péricarpe testa est semi-perméable (perméable à l'eau mais formant une barrière pour de nombreux solutés). Il attribue la capacité sélective de l'absorption de grain. Cette imperméabilité diminue avec le temps de trempage.

3.2.6. La paroi externe :

La paroi externe entoure complètement le grain sauf aux deux extrémités. Elle doit être saine pour permettre une absorption d'eau homogène lors du trempage pour protéger l'embryon des infections pendant la germination. Elle possède une fonction protectrice et assure par capillarité la distribution d'eau sur la surface du grain. (Hariri, 2003). L'étape la plus importante dans le maltage c'est l'imbibition des grains dans l'eau, Il est intéressant de connaître que l'entrée de l'eau à l'intérieur des grains d'orge est gouvernée par le temps de maltage et en même temps par des autres conditions opératoires et par les caractéristiques de la matière de base.

3.3. Composition biochimique de grain d'orge

La composition du grain d'orge est très important pour la qualité de du malt. Ces compositions déterminent les modifications physico-chimiques du grain au cours du processus de maltage. La composition biochimique est présentée dans le Tableau 2.

Tableau 2 L a composition chimique du grain d'orge (Guiga, 2006)

Composition biochimique	Orge teneur en (% de poids sec)	Malt teneur en (% du poids sec)
Glucides	78-85	79-88
Amidon	63-65	85-60
Saccharose	1-2	3-5
Sucres réducteurs	0,1-0,2	3-4
Polysaccharides solubles dans l'eau (gommes)	1,1,5	2-4
Polysaccharides solubles dans les solvants organiques (hémicelluloses)	8-10	6-8
Cellulose	4-5	5
Autres	1	2
Lipides	2-3	2-3
Protéines	8-11	8-11
Albumine	0,5	2
Globuline	3	-
Hordéine	3-4	2
Glutéline	3-4	3-4
Acides aminés et peptides	0,5	1-2
Acides nucléiques	0,2-0,3	0,2-0,3
Sels minéraux	2	2,2
Autres	5-6	6-7

Le Figure 5 présente une coupe longitudinale dans un grain d'orge et montre les différents constituants de grain d'orge (Allosio-Ouarnier, 1999).

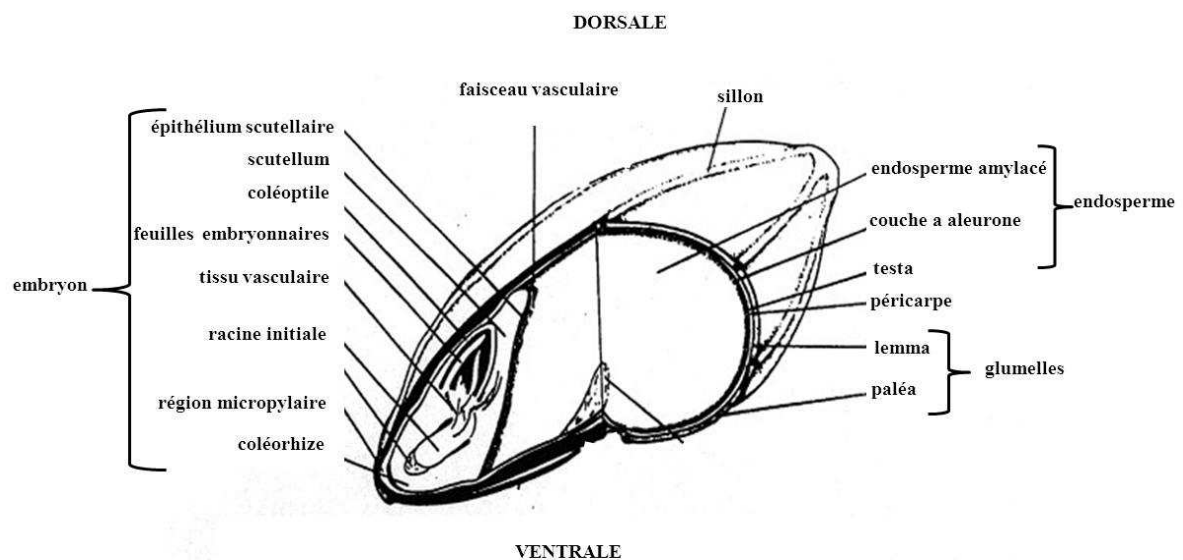


Figure 5 disposition des différents tissus d'un grain d'orge (Briggs, 1998)

3.4. Indicateurs de qualité de l'orge

L'orge brassicole doit donner un maximum de rendement en extraits solubles (sucres), pouvoir produire un maximum d'enzymes (pouvoir diastasique), avoir un type d'amidon se désagrégeant bien, ne pas contenir de composés indésirables (β -glucanes, tannins et polyphénols, anthocyanogènes, certaines protéines...). De ce fait, des critères définissant la qualité brassicole permettent de comprendre que toutes les orges ne sont pas utilisables pour faire une bière de qualité. Les différents critères de qualité de l'orge de brasserie peuvent être classés en trois familles de facteurs :

1. Facteurs liés à la composition : teneur en humidité, teneur en protéine, teneur en β -glucan et teneur en Arabinoxylane.
2. Facteurs liés à la structure: dureté vitrosité et porosité
3. Facteurs influençant la qualité : espèce, variété et lieu de culture.

L'influence de la composition chimique de l'orge sur la qualité du malt est détaillée dans le chapitre 4.

En ce qui concerne les facteurs liés à la structure, ce sont des facteurs physiques qui décrivent la structure de l'endosperme, cette dernière a une influence significative sur l'absorption l'eau pendant la trempe (Gamlath *et al.*2008; Psota *et al.*2007; Vejrazka *et al.*2008).

La porosité : est un facteur physique qui exprime la compacité de l'endosperme (Mayolle *et al.*, 2012). Ce facteur est rarement utilisé malgré l'intérêt qu'il peut fournir.

La dureté : est le degré de résistance à la déformation. La dureté a un rôle important dans la détermination de la qualité de grains. Elle affecte d'une manière significative les propriétés de mouture, la distribution de taille des particules, la dégradation de l'amidon et l'absorption de l'eau par la farine (Williams, 1967 ; Endey *et al.*2002). Flidel *et al.* (1989) ont montré que la dureté influence le rendement des grains, la quantité de brisures (parties éliminées avec les sons) est plus importante pendant le broyage des grains durs. Ce paramètre est sensible aux facteurs génétiques (Turnbull et Rahman, 2002)

La vitrosité : est un facteur physique qui influence les propriétés rhéologiques de l'endosperme. Les grains vitreux sont translucides et les granules d'amidons sont plus compactés dans la matrice protéique de l'endosperme, tandis que les grains farineux sont opaques avec des espaces vides dans la matrice de protéines-amidon. Elle est contrôlée par des facteurs environnementaux (Vejrazka *et al.*2008)

Les trois facteurs structurels des grains ont donc une influence significative sur la diffusion d'eau pendant la trempe. Mayolle *et al.*(2012), Chandra *et al.*(1999) ont montré que les variétés d'orges dont les grains sont plus durs et plus vitreux ont la diffusion de l'eau la plus lente. La diffusion d'eau augmente quand la porosité des grains est plus élevée (Maylloe *et al.*2012).

Dans le chapitre 3, nous avons cherché à exploiter certains de ces indicateurs pour caractériser le malt.

La qualité de l'orge a un effet déterminant sur la qualité du malt et de la bière produite. Dans un premier temps, il est ainsi important de comprendre les facteurs génétiques qui contribuent à la qualité de l'orge afin de sélectionner ou d'orienter vers les meilleures variétés. Un taux raisonnable de β -glucanes peut être obtenu en plantant des variétés précises dans des lieux spécifiquement adaptés (Zhang *et al.*2002; et Paynter *et al.*2010). Il est aussi nécessaire d'identifier les facteurs environnementaux qui affectent les différents indicateurs de qualité de l'orge. Si certains d'entre eux peuvent difficilement être maîtrisés (effet météorologique, variations saisonnières). Des études ont déjà montré l'effet de l'arrosage ou du chauffage en plein champs pour améliorer la qualité de l'orge. L'irrigation hebdomadaire augmente le poids du grain, le rendement en grains, et l'extrait de malt (Paynter and Young, 2004). Passarella *et al.*(2002) ont trouvé que des températures élevées en plein champs augmentent le taux d'azote, diminuent le taux de β -glucanes de l'orge et diminuent l'extrait de malt.

Dans un second temps, il est intéressant d'étudier l'effet des variations de qualité de l'orge sur la qualité du malt obtenu (dans des conditions opératoires équivalentes).

3.5. Relation entre la qualité de l'orge et la qualité du malt et de produit final

L'orge est un système complexe qui est caractérisé par un éventail d'index physiques et chimiques. L'orge est constamment exposée à des facteurs biotiques et abiotiques qui déterminent la fertilité, le rendement ainsi que des indicateurs qualitatifs. Le risque climatique peut inhiber la production des orges de printemps et d'hiver. La qualité est en train de changer d'année en année en fonction de la météo et du temps des semailles. Les malteurs recherchent l'orge qui germera de façon rapide et régulière. Diverses études et analyses ont été effectuées sur les facteurs affectant la qualité d'orge et, en conséquence, la qualité de malt :

1. Effet paramètres qualitatifs (variété, année, lieu...) sur les constituants de grain d'orge.

La structure et la composition d'un grain d'orge sont fortement liées aux paramètres qualitatifs discutés dans la section précédente. Oscarsson *et al.*(1999) , Tiwari et Cummins, (2008) et Dickin *et al.*(2011) ont montré que la variété d'orge a un effet important sur la teneur en β -glucanes de l'orge. Les semis d'automne et les fongicides diminuent la teneur en β -glucanes. En revanche ils diminuent la sensibilité aux variations environnementales. Des différences significatives de teneur en β -glucanes ont été soulignées par (Zhang *et al.*2002; et Paynter *et al.*2010). Parmi les facteurs essentiels, la variété avait le plus grand effet sur la teneur en β -glucanes, suivi par le lieu de culture, et la fertilisation azotée (N) ayant le plus petit effet. Un taux raisonnable de β -glucanes peut être obtenu en plantant des variétés précises dans des lieux spécifiquement adaptés. Xue *et al.*2008 ont montré que la teneur en protéines des trois fractions (albumine, hordéine et glutéline) dans les grains d'orge varie d'une variété à l'autre

et d'un lieu de culture à l'autre. Confirmé par Simic *et al.* 2008 ont confirmé l'effet de variété mais n'ont pas mis en évidence d'effet significatif de l'année ou du lieu de culture. L'effet environnemental et géographique sur la protéine d'orge et le contenu en arabinoxylane respectivement ont été étudiés (McTaggart *et al.* 1992 ; Fleury *et al.* 1997). De saisons courtes et sèches résultent des grains de plus haute teneur en protéines et des moindres teneurs en amidon (Aman et Newman, 1986). L'irrigation hebdomadaire augmente le poids moyen des grains et il diminue en même temps son calibre, conduisant ainsi à des meilleurs rendements (Paynter and Young, 2004). Passarella *et al.* 2002 ont trouvé que des températures élevées en plein champs augmentent le taux d'azote, diminue le taux de β -glucanes de l'orge et diminue l'extrait de malt

2. Effet des compositions de grain d'orge sur les modifications de l'endosperme, et la qualité de malt final.

Les constituants du grain d'orge sont ainsi influencés par des effets environnementaux et génétiques. Et affectent à leur tour la qualité du malt.

L'orge dure produit moins de masse de racine, plus d'azote aminé et de protéines solubles, et il en résulte un malt d'une structure moins friable. (Holpainen *et al.* 2005 ; Chandra *et al.* 1999).

La teneur en β -glucanes et en protéines est corrélée négativement avec la teneur en amidon (Oscarsson *et al.* 1998). Les niveaux élevés d'amidon, et faibles de protéines augmentent l'extrait de l'orge brassicole. En parallèle, la matrice de l'albumen de l'orge, essentiellement constituée de β -glucanes et de protéines, doit toujours être bien dégradée pour pouvoir exploiter ce potentiel d'extraction (Plamer *et al.* 1989).

Les taux élevés en protéines empêchent les modifications de l'endosperme, conduisant à un malt à faible extrait. Inversement, une teneur modérée en protéines augmente l'extrait de malt et diminue le taux de β -glucanes dans le moût. La filtration du moût et de la bière devient plus rapide (Sottnicova *et al.* 2013, Edney *et al.* 2012, Fox, 2002, Peltonen, 1994, Aman *et al.* 1985,).

Bravi *et al.* (2012) a étudié l'influence de la variété d'orge et de la réserve de lipides dans l'orge sur la teneur en lipides du malt final. La dégradation de plusieurs variétés d'orge a conduit à des compositions en acides gras dans le malt différentes. L'orge à teneur élevée en lipide produit un malt de faible friabilité et de faible fermentescibilité.

L'absorption de l'eau et sa distribution pendant le maltage a un important effet sur les modifications de l'endosperme. Le taux élevé des protéines et de β -glucanes freine l'absorption d'eau dans le grain, d'où les modifications de l'endosperme se ralentissent (Briggs, 1981, Duijnhower *et al.* 1993, Sheehy *et al.* 2009). Une forte teneur en B-Hordéine diminue l'absorption d'eau pendant le maltage, la présence des hordéines de petite taille permet à l'orge de se malter et d'absorber l'eau plus rapidement (Molina-Cano *et al.* 2002 ; Brookes *et al.* 1976).

En résumé, il n'y a aucune orge parfaite pour une bonne qualité de malt. En effet, la composition de la matière première est liée à divers facteurs tels que l'année de récolte, des espèces d'orge, le type d'orge (2 ou 6 rangs), de l'emplacement de la culture et de la variété d'orge.

La qualité de l'orge joue un rôle essentiel sur la qualité de malt final. L'évaluation des paramètres de qualité de l'orge est donc très importante pour pouvoir définir les meilleures conditions opératoires adaptées à ses besoins pour obtenir un produit final de qualité.

4. Méthode d'analyses

4.1. Méthodes classiques

Traditionnellement, une analyse complète de qualité d'orge inclut la mesure de 4 à 7 paramètres physiques et chimiques pour l'orge. Les 4 paramètres quantitatifs de qualité de l'orge les plus considérés sont (teneur en humidité (%), teneur en β -glucanes (mg/l), teneur en arabinoxylanes (%)) plus l'indice de germination (%) dans cette étude. Les paramètres de qualité sont obtenus par les méthodes normalisées par l'EBC (2010). Ces analyses de référence sont habituellement longues et chères.

4.2. Méthodes alternatives

Des méthodes alternatives ont été suggérées pour collecter facilement et rapidement l'information sur l'orge, et faciliter ainsi le choix de nouveaux cultivars par les malteurs. ???(2007) ont utilisé la résonance magnétique nucléaire de proton, pour étudier la mobilité de l'eau dans l'orge hydratée, il a constaté que la prise de l'eau et sa distribution dépend de la composition du grain d'orge. L'utilisation parallèle de deux techniques a été aussi envisagée dans la littérature. L'imagerie RMN a été couplée avec l'analyse morphologique (acquisition d'image, segmentation, extraction externe et interne de dispositif d'image, classification et interprétation) (Zapotocny *et al.* 2008). Ils ont constaté que des dispositifs morphologiques peuvent être utilisés avec succès pour l'identification variétale préliminaire des grains d'orge. Par contre le RMN technique requiert une pureté maximale de l'échantillon et une quantité suffisante de matériel. Cozzolino *et al.* (2012) ont utilisés d'autres types d'analyses, l'analyseur rapide de Viscosité couplé à l'analyse multi-variable qui peut être considéré comme outil analytique intéressant pour étudier la processus de transformation d'amidon et pouvait prévoir quelques importants facteurs de la qualité d'orge. Ces méthodes sont efficaces mais nécessitent du temps pour préparer les échantillons et analyser les résultats. Ce qui n'est pas apprécié dans le cas d'analyse de bases de données de grande taille.

Les méthodes conventionnelles impliquent des procédures longues, laborieuses et coûteuses, y compris l'analyse physique et chimique ou l'analyse par microscopie. Une méthode peu coûteuse, rapide et robuste basée sur la spectroscopie pour le processus de certification serait idéale pour la caractérisation de la qualité d'orge avant le procédé de maltage.

4.2.1. La spectroscopie infrarouge (IR)

Les méthodes spectroscopiques se décomposent globalement en deux grandes catégories, la spectroscopie des rayonnements – qui elle-même regroupe la spectrométrie d'absorption, la spectrométrie d'émission, la spectrométrie de diffusion Raman et la spectrométrie de résonance magnétique nucléaire – et la spectroscopie de masse. La spectroscopie et la spectrométrie des rayonnements regroupent un ensemble de méthodes d'analyse permettant d'accéder à la composition et à la structure de la matière fondées sur l'étude des spectres fournis par l'interaction des atomes et des molécules avec divers rayonnements électromagnétiques qu'ils émettent, absorbent ou diffusent. En spectroscopie infrarouge (IR), l'absorption résulte des phénomènes de vibration et rotation des molécules. Les spectres d'absorption infrarouge permettent donc de déterminer la nature des liaisons chimiques. Le proche infrarouge (NIR) et le moyen infrarouge (MIR) sont les deux méthodes les plus utilisées dans le domaine de la spectroscopie infrarouge. Le domaine du NIR est généralement désigné comme le spectre électromagnétique entre les longueurs d'onde de ~ 750 et ~ 2500 nm de longueur d'onde ($\sim 13\,500$ à ~ 4000 cm^{-1}) alors que celui du MIR est comprise entre $2,500$ et $\sim 23\,500$ nm (4000 à ~ 450 cm^{-1}) avec absorptions vibratoire, alors que NIR mesure les absorptions harmoniques, et les vibrations élevées de l'énergie.

4.2.1.1. La spectroscopie proche infrarouge (NIR)

Le NIR spectroscopie est un outil utilisé presque dans tous les domaines de la recherche. En pétrochimie : pétrole fractions, carburants, polymères (Yalvac *et al.*1997 ; Parisi *et al.*1990, Lachenal, 1998). Dans le secteur clinique: sérum humain, cancer (Malin *et al.*1999 ; Wallon *et al.*1994). En matière d'environnement, le NIR a été appliqué au recyclage (Alam *et al.*1995). En science de bois, de nombreux chercheurs ont rapporté que la technique NIR est utile pour la détection des propriétés chimiques et physiques des matériaux en bois (Tsuchikawa and Schwanninger, 2013). En pharmacie. La spectroscopie NIR et la chimiométrie ont été utilisées dans des plusieurs thèmes pharmaceutiques : analyses qualitatives, méthodes quantitatives et applications en ligne (Roggo *et al.*2007). Et bien sûr dans le domaine d'agriculture et des aliments : adultération, qualité et compositions (Ding *et al.*2000, Ridgway *et al.*1999, Kays *et al.*2000). Dans le domaine de l'analyse de l'orge, la technologie NIR est déjà bien établie dans l'évaluation de la composition de grain. Il a été conclu que le phénotype??? de grain de l'orge par NIR associée à la chimiométrie était un nouvel outil fiable pour caractériser les effets des mutant gènes sur la sélection de l'orge de qualité (Jacobsen *et al.*2005). Des modèles pour la prédiction rapide et sûre de la teneur en humidité et en lipides totaux des maïs par spectroscopie (NIR) proche-infrarouge ont été développés (Marte *et al.*2009). En outre, il s'est avéré que NIR a un potentiel pour prévoir la qualité du grain entier de malt (prédiction de la teneur en humidité et en azote, de l'extrait de malt et de la teneur en protéines, (Sileoni *et al.*2010 ; Tarr *et al.*2012).

4.2.1.2. La spectroscopie moyenne infrarouge (MIR)

Les applications de la spectroscopie MIR ont réalisées avec succès dans de très nombreux domaines de la science. Le Table 3 présente plusieurs exemples sélectionnés pour l'utilisation de MIR-spectroscopie dans 4 domaines scientifiques. Ces derniers ne représentent qu'une infime partie de la littérature associée à la spectroscopie MIR qui ne peut être développée de manière exhaustive dans ce document.

Table 3 quelques exemples sélectionnés pour l'utilisation de la spectroscopie MIR	
<u>Pharmacie</u> • Traçage - (Vigabatrin) as active component dans les produits pharmaceutiques (Torano et al., 2001) - Vitamin C dans les produits pharmaceutiques en poudre et solution (Yang and Irudayaraji, 2002)	<u>Médecine</u> • Biomédical pertinents compositions - DNA/RNA (Wang et al., 2008) - Protéine • Détecte des maladies (L'athérosclérose, Leucémie, ou la progression des maladies (Soyeurt et al., 2013)
<u>Aliments</u> • Discrimination par variété - Huile d'olive (Aouidi et al., 2012) - Jus de raisin (Cozzolino et al., 2012) • Qualité - Teneur en triglycérides et la composition d'acide gras de la Beurre de cacao (Maurer and rodriguez, 2013) • Adultération - Lactosérum, peroxyde d'hydrogène, et le lait synthétique dans le lait de vache (Santos et al., 2013). - La viande hachée avec la viande de dinde (Alamprese et al., 2013).	<u>Environnement</u> • Compositions - Bois carbone et nitrogène content (Olale et al., 2013) - Sol content en carbone (Wang et al., 2011)

Dans la littérature, la spectroscopie MIR a souvent été appliquée sur des produits alimentaires aussi variés comme le montre Figure 6, sur produits nombreuses pour chaque group d'aliments, et pour des applications différents :

Pour chaque domaine d'aliment plusieurs produits sont étudiés ils sont citées dans le paragraphe suivant et comme le montre le Figure 6.

1. Le lait et les produits laitiers
 - o Le lait

- o Le Fromage
- o Le Beurre
- 2. La viande et les produits carnés
 - o La viande de bœuf, poulet, porc...
 - o Le poisson
- 3. L'huile
- 4. Le miel
- 5. Le café
- 6. Les boissons alcoolisées
 - o Le vin
 - o La bière
- 7. Les céréales
 - o Le blé
 - o Le riz
 - o L'orge

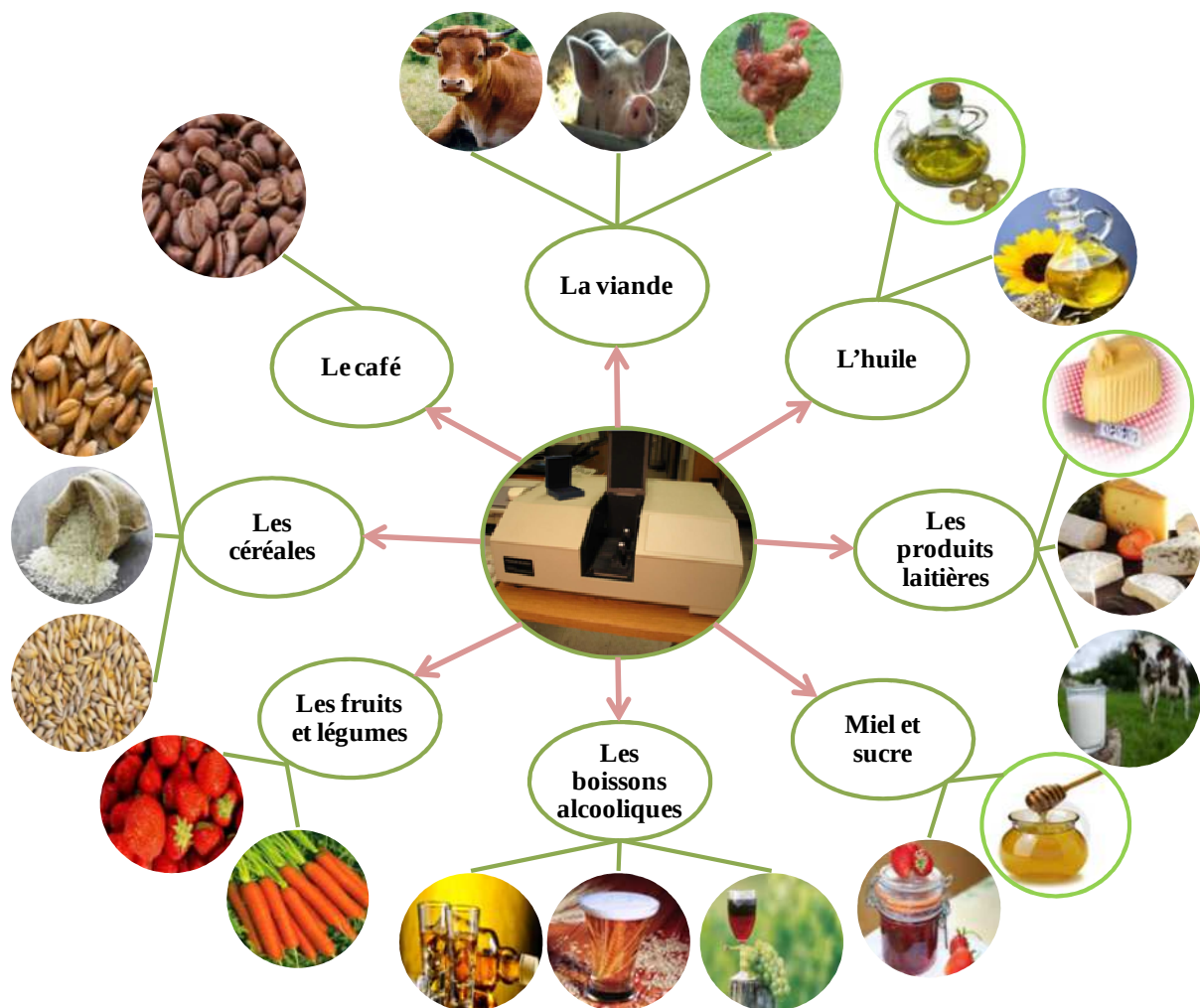


Figure 6 des domaines d'application de MIR spectroscopie sur des produits alimentaires

Viande et produits carnés, lait et produits laitiers, les boissons alcoolisées

Parmi les applications intéressantes sur ces différents produits alimentaires, on peut citer :

- ✓ Caractérisation qualitative d'un produit alimentaire :
 1. **Origine géographique** d'un produit : le fromage (Picque *et al.*2002 ; Pillonel *et al.*2003 ; Karoui *et al.*2004, 2005) ; le safran (Anastasaki *et al.*2010), le thé (Chen *et al.*2009), le café (Hashimoto *et al.*2004), le miel (Rough *et al.*2006).
 2. **Variété** : le café (Hashimoto *et al.*2004), le jus de pomme (Reid *et al.*2005), le miel (Rough *et al.*2006), le vin (Tarantilis *et al.*2008).
 3. **Année de récolte** : discrimination de purée de fruit d'une année par rapport à l'autre (Holland *et al.*1998). Par ailleurs, d'autres chercheurs ont montré que les spectres IR-MIR ne peuvent pas fournir une distinction par variété, géographique zone ou année de culture (Grurdeniz *et al.* 2009)
- ✓ Détermination de la composition d'un produit :
 - **Protéines** dans lait (Inon *et al.*2004, Etzion *et al.* 2004), dans le blé (Nunes *et al.*1997), dans le riz (Vansoest *et al.*1994).
 - **Les sucres** dans le vin (Egidio *et al.*2010), dans le miel (Lichtenberg *et al.*2002).
 - **Ethanol** dans le vin ou la bière (Lachenmeier, 2007 ; Egidio *et al.*2010).
 - **Les acides gras** dans les huiles de table (Inon *et al.*2003).
 - **La caféine et l'acide chlorogénique** dans le café (Hashimoto *et al.*2004)
- ✓ Analyser la qualité sensorielle d'un produit alimentaire :
 - Déterminer la qualité des viandes : viande hachée de poulet, de porc, de dinde (Al-Jowder *et al.*1997).
 - Discrimination des produits congelés de produits frais. Poisson (Al-Jowder *et al.* 1997 ; Karoui *et al.* 2007b)
 - Détection des fraudes alimentaires : détermination de l'adultération du beurre avec le margarine (Koca *et al.*2010), détection de la viande hachée de bœuf adultéré avec de viande de dinde (Alamprese *et al.* 2013). Investigation du miel irlandais et adultération avec un mélange glucose-fructose (Kelly *et al.*2006) . Détection de l'adultération des huiles d'olive avec de huile de tournesol (Tay *et al.*2002).

Plusieurs études sur l'utilisation de la spectroscopie MIR ont été réalisées dans le domaine des céréales. Elles concernent des analyses qualitatives : Reeves and Delwiche, (1997) ont montré que le MIR peut être un outil potentiel pour déterminer la teneur en protéine dans la farine de blé d'hiver avec des modèles très acceptables ($R^2 = 0.997$). Shao et al. (2011) ont montré que la teneur en protéines du riz peut être déterminée avec succès par le MIR et analyses chimiométriques avec un coefficient de détermination de $R^2 = 0.93$. Robert et al. (2005) ont proposé la spectroscopie MIR en tant qu'outil pour évaluer la proportion de β -glucanes et d'arabinoxylane dans les parois cellulaires des céréales. Plusieurs études ont été réalisées pour caractériser et étudier le malt ou le processus de maltage. Tarr et al. (2012) ont suggéré que le MIR pouvait être utilisé comme méthode complémentaire pour suivre le processus de maltage en ligne

A notre connaissance, peu d'études ont été réalisées sur l'orge. Toole *et al.* (2012) ont étudié la distribution de l'arabinoxylane dans l'endosperme des différentes céréales dont l'orge. Ils ont montré que la distribution de l'arabinoxylane de chaque céréale est différente. Ceci est probablement lié aux facteurs génétiques de chaque céréale. Lhotakova *et al.* (2010) ont réussi à discriminer l'orge par variété. Les résultats présentés confirment que les orges d'origines génétiques différentes peuvent être, rapidement et de manière fiable, distinguées par une analyse multivariée des spectres de vibration des grains d'orge moulus. Des échantillons d'orge contenant de petites quantités de β -glucanes et des teneurs relativement élevées de pentosanes ont été séparés par FT-IR et FT-NIR combiné avec des méthodes chimiométriques. Des déterminations quantitatives par FT-IR et FT-NIR de «masse de mille grains», amidon, N-substances, β -glucanes et pentosanes sont possibles.

Les données spectrales comme les spectres MIR contiennent habituellement un grand nombre d'information très corrélées entre elles. Les informations analytiques dans les spectres peuvent être extraites par des techniques analytiques multivariées. Les méthodes multivariées les plus utilisées permettent alors de distinguer des échantillons apparemment semblables (classification) et de quantifier des propriétés difficiles à mesurer par d'autres techniques). L'état des connaissances actuelles sur les applications de ces méthodes dans notre domaine sera décrit dans la section suivante.

5. Modélisation : caractérisation de l'orge et du malt, relation orge & malt

La modélisation, consiste à définir une (ou plusieurs) relations de nature mathématique (ou non), permettant de rendre compte, d'une manière suffisamment générale, des relations entre les grandeurs dans un système ou processus qu'il soit physique, biologique, économique ou autre.

Un modèle mathématique est une représentation ou une interprétation abstraite de la réalité qui est accessible à l'analyse et au calcul (Allaire, 2007). Les modèles peuvent être classifiés selon plusieurs paramètres.

5.1. Classification des modèles (Bouamama, 2004)

✚ Selon le mode de conception

Les modèles de connaissances : les modèles de connaissances sont construits à partir d'une analyse physique, chimique, biologique (ou autre suivant le type du processus), en appliquant soit les lois générales, fondées sur des principes (lois de la mécanique, de l'électromagnétisme, de la thermodynamique, de la physique quantique, etc.), soit les lois empiriques (finance, économie), qui régissent les phénomènes intervenant au sein des processus étudiés. Dans la pratique, il est toujours souhaitable d'établir un modèle de connaissance des processus que l'on étudie. Néanmoins, il arrive fréquemment que le processus soit trop complexe, ou que les phénomènes qui le régissent sont trop mal connus, pour qu'il soit possible d'établir un modèle de connaissances suffisamment précis pour l'application considérée. On est alors amené à concevoir des modèles purement empiriques, fondés exclusivement sur les résultats de mesures effectuées sur le processus. A notre connaissance, peu de modèles de connaissances ont été développés dans le domaine de la transformation de l'orge en malt. Des études sur la diffusion de l'eau dans les grains qui prennent en compte les différents compartiments du grain ont été réalisées. (Igathinathane, 1999) ont déterminé les coefficients de diffusion de l'eau dans l'endosperme pendant le séchage à différentes températures. Dans notre travail, en raison des différentes échelles étudiées (depuis les compartiments internes au grain jusqu'au lit fixe d'orge dans la cuve de trempage et de germination) et en raison d'un manque d'outils pour une analyse fine des grains d'orge et de l'hétérogénéité dans les procédés de transformation, nous n'avons pas utilisé de modèle de connaissances et nous nous sommes focalisés sur les approches de type "boîte noire".

Les modèles "boîte noire" : les modèles de type boîte noire sont construits sur la base de mesures exclusivement effectuées sur les entrées et les sorties du processus à modéliser. La modélisation consiste alors à utiliser des équations (algébriques, différentielles, récurrentes) pour décrire la relation entre entrées/sorties. Ces modèles ne postulent donc aucune connaissance concernant la structure interne (réelle) du modèle. (Kreis et al 2007 ; Toutain et al. 2007, Ouassar, 1998). Parmi les modèles de type "boîte noire", plusieurs choix sont possibles : polynômes, réseaux de neurones, modèles flous, systèmes experts.

On désigne également les modèles à mi-chemin entre les modèles de connaissance et les modèles "boîte noire" par le terme **modèles hybrides (ou boîte grise)**: Ils sont utilisés lorsque la connaissance du système à modéliser est incomplète.

✚ Selon l'utilisation de modèle (Ouassar, 1998 ; Bouamama , 2004)

Les objectifs d'une modélisation statistique peuvent être de différentes natures, que l'on peut tenter de répartir en deux classes, les objectifs prédictifs et les objectifs explicatifs :

- **Les modèles non adaptatifs** : il s'agit de modèles qui ne sont pas recalibrés en temps réel du processus qu'il représente. Ils doivent donc posséder un comportement aussi semblable que possible à celui du processus. Du point de

vue de la structure du modèle, les sorties passées, *mesurées sur le processus à modéliser*, ne peuvent constituer des entrées du modèle. L'estimation des paramètres et l'utilisation du modèle constituent deux phases successives et distinctes.

- **Les modèles adaptatifs:** il s'agit de modèles qui sont calibrés en parallèle du processus réel qu'ils modélisent. Ils prédisent la sortie du processus à une échelle de temps courte devant les constantes de temps du processus. Ces prédicteurs (observateurs d'état) sont utilisés pour la synthèse de lois de commande, ou dans le système de commande lui-même (commande avec modèle interne). Du point de vue de la structure du modèle, les sorties passées, mesurées sur le processus, peuvent constituer des entrées du modèle. L'estimation des paramètres et l'utilisation du modèle est effectuées simultanément si nécessaire.

✚ Selon l'évolution des paramètres

- Stochastique, déterministe : L'approche de modélisation déterministe suppose l'existence d'une relation fonctionnelle entre les variables observées. L'observation des variables se fait par une série de mesures entachées d'erreurs. L'hypothèse sous-jacente est que la relation fonctionnelle est simple et que les erreurs sont homogènes et indépendantes. Chaque observation est donc la somme d'une valeur déterministe et du tirage aléatoire de l'erreur de mesure. L'approche de modélisation stochastique généralise la démarche en considérant que les observations sont le résultat d'un tirage aléatoire mais que les variables considérées ne sont pas nécessairement indépendantes.

Ce mémoire présente la mise en œuvre de plusieurs modèles de type boîte noire (MLR, ANN, PCR, PLS) avec l'analyses en composantes principales (ACP) pour la modélisation statistique de la matière première (l'orge) et du processus de maltage. Les modèles de type boîte noire peuvent être utilisés indifféremment comme simulateurs ou comme prédicteurs.

5.2. Les différentes étapes de la conception d'un modèle (Ouassar, 1998 ; Jacques, 2009) :

La construction d'un modèle peut être résumée comme dans le schéma de la Figure 7 :



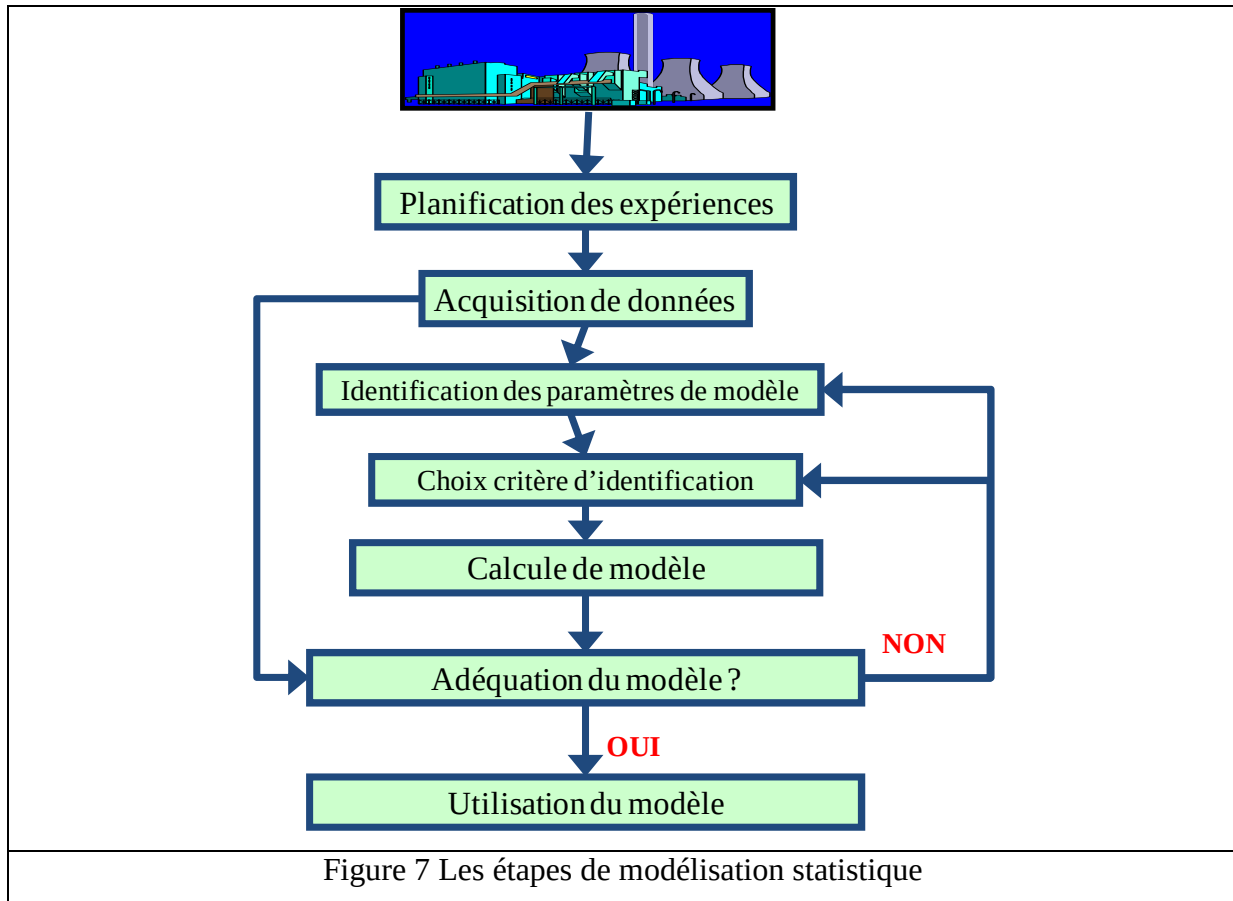


Figure 7 Les étapes de modélisation statistique

Dans la modélisation de type “boîte noire”, la construction du modèle nécessite les trois éléments suivants :

- Une hypothèse sur l’existence d’une relation déterministe liant les entrées à la (ou aux) sortie(s). Cette relation est caractérisée par une fonction appelée *fonction de régression* (ou plus simplement *régression*). L'expression formelle supposée adéquate pour représenter cette relation est appelée *modèle hypothèse*.
- Une séquence de mesures des entrées et de la sortie du processus.
- L'identification des paramètres du modèle par optimisation (régression linéaire, régression non-linéaire ou apprentissage).

Dans la suite de ce paragraphe, nous présentons les différents aspects qui doivent être pris en considération lors du choix d’un modèle-hypothèse.

5.2.1. Choix d’un modèle-hypothèse

Les connaissances dont on dispose a priori sur le processus doivent guider le concepteur dans le choix de la modélisation la plus appropriée (statique ou dynamique, linéaire ou non linéaire, ...). L'élaboration du modèle-hypothèse nécessite d'effectuer les choix suivants :

- **Modèle statique ou dynamique** : lorsque l'on cherche à modéliser un processus physico-chimique ou biologique, il est généralement facile de savoir si l'application envisagée nécessite de modéliser la dynamique du processus (c'est-à dire si l'on doit considérer une échelle de temps petite devant les constantes de temps du processus) ou si une modélisation statique suffit.
- **Modèle linéaire ou non linéaire** : la plupart des processus que l'on peut rencontrer nécessitent des modèles non linéaires s'il faut les décrire de manière précise dans la totalité de leur domaine de fonctionnement. Il est donc important de pouvoir élaborer un modèle non linéaire pour rendre compte du comportement d'un processus, non seulement autour de ses points de fonctionnement "habituels", mais également lors des passages d'un point de fonctionnement à un autre. La plupart des modèles linéaires constituent cependant des approximations valables lorsque le domaine est plus restreint.

5.2.2. Fonction paramétrée pour la modélisation boîte noire

En générale une fonction peut être linéaire ou non par rapport aux coefficients

✓ **Les fonctions paramétrées linéaires par rapport aux paramètres.**

Une fonction paramétrée est linéaire par rapport aux paramètres si elle est de la forme :

$$\psi(X) = \sum_{i=1}^N \theta_i \Phi_i(X) \quad (1)$$

où les $\Phi_i(X)$ sont des fonctions non paramétrées d'une ou plusieurs variables groupées dans le vecteur X , et où les θ_i sont des paramètres.

Les fonctions $\Phi_i(X)$ peuvent être quelconques ; traditionnellement on utilise des monômes ; des fonctions exponentielles, logarithmiques,... voire des fonctions plus complexes : Splines, gaussiennes dont les centres et les écarts types sont fixés...

✓ **Les fonctions paramétrées non linéaires par rapport aux paramètres.**

Les fonctions non linéaires par rapport aux paramètres sont de la forme

$$\psi(X) = \Phi(X, \theta) \quad (2)$$

Où θ_i est un vecteur de paramètres de la fonction Φ . Ainsi, la fonction réalisée est non linéaire par rapport aux coefficients θ_i .

Un modèle-hypothèse ayant été choisi, l'étape suivante consiste à déterminer les valeurs de ses paramètres à partir des observations expérimentales entrées-sorties.

A l'issue de la procédure d'estimation des paramètres (identification), il faut évaluer la capacité prédictive du modèle.

Dans le paragraphe précédent le processus de modélisation a été discuté en général (définition, classification des modèles, étapes de modélisation). L'état de l'art des modèles qui sont choisis pour réaliser ce mémoire (régression multilinéaire (MLR), analyses en composantes principales (PCA), régression sur les composantes principales (PCR), régression au sens des moindres carrés partiels (PLS)) est présenté dans la section suivante.

5.3. La régression multilinéaire (MLR)

Avant de commencer la modélisation, une étape très importante consiste à établir un plan d'expérience. Les plans d'expériences bénéficient d'une large audience et leur pratique se sont rapidement développées au cours des 20 dernières années dans les entreprises, de la conception à la fabrication. Un plan d'expérience bien choisi permet de quantifier et de classer les effets d'un grand nombre de facteurs d'un système étudié en minimisant le nombre d'essais, le temps et les coûts (Schimmerling, 1998). Selon Goupy (1989) les avantages des plans d'expériences peuvent être résumés par :

- La réduction de nombre d'essais, bonne détection des interactions entre facteurs.
- L'élargissement du nombre des facteurs qui peuvent être étudiés.
- La précision des résultats

La Figure 8 présente les étapes nécessaires pour établir une stratégie d'expérience

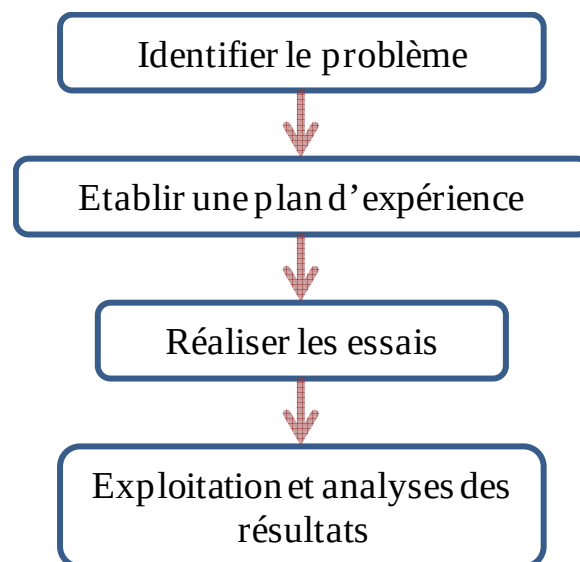


Figure 8 les étapes de construction d'un plan d'expérience

5.4. La régression multilinéaire (MLR)

Dans les démarches pour la maîtrise du maltage, des efforts de plus en plus forts sont imposés pour améliorer les connaissances autour du maltage, pour essayer de bien maîtriser ce processus. Dans la première partie de cette étude, il s'agit d'employer la régression linéaire multiple (MLR) comme un outil potentiel pour étudier et identifier les conditions opératoires et/ou les paramètres de qualité d'orge influençant la qualité de malt final.

La régression linéaire multiple est une méthode simple, pratique et très répandue de la modélisation. Elle consiste à rechercher une équation linéaire par rapport à ses paramètres,

reliant la variable à modéliser y au vecteur d'entrées X . Ces entrées peuvent être soit les variables primaires, soit des fonctions non paramétrées, ou à paramètres fixés, de ces variables (par exemple des monômes). Généralement ils sont des modèles polynomiaux linéaires par rapport aux coefficients mais qui peuvent être non linéaire par rapport aux paramètres.

L'identification des paramètres du modèle s'appuie sur la minimisation d'un critère de type moindre carré dont la solution optimale est déterminée par une expression analytique et explicite. Ces approches permettent ainsi de garantir que le meilleur jeu de paramètres est déterminé avant l'analyse et de vérifier l'adéquation du modèle avec le comportement expérimental observé.

Dans le domaine de la brasserie, de nombreuses études ont employé la MLR régression pour prédire des facteurs de qualités importants. Kettunen *et al.* (1996) ont trouvés des modèles MLR capables de prédire la concentration en β -glucanes dans la phase liquide de brassage (le moût). Ils présentent également des modèles prédisant l'effet de la température du brassage sur l'activité des β -glucanes et leur solubilisation. Holmberg *et al.* (2002) ont modélisé la germination des grains d'orge pendant et après la trempe en fonction du diagramme de trempe, de l'humidité des grains et de la sensibilité à l'eau. Ils ont réussi à prédire les performances de la trempe et de la germination en fonction de la durée des phases pour pouvoir ensuite planifier le programme de trempe pour atteindre une germination homogène pendant le maltage.

La régression linéaire multiple a été appliquée par Guido *et al.* (2007) pour explorer les relations entre les paramètres chimiques du malt et les données sensorielles de la bière, montrant le pouvoir anti-radicalaire comme la contribution majeure du malt sur la stabilité de la saveur de la bière. En outre, le pouvoir anti-radicalaire mesuré, qui est bien corrélée avec la teneur en polyphénols, a été jugé très similaire pour le malt et l'orge, en insistant sur le rôle clé des polyphénols endogènes de l'orge. Des modèles multilinéaires pour prédire la teneur en protéines de l'orge et du malt ont été établis par Fox *et al.* (2002) avec des erreurs standard de prédiction de 15% et 17% respectivement.

Des modèles multilinéaires ont aussi été développés pour prédire la thermo-stabilité de la β -amylase et de valeur de dextrinase limite en tant que facteurs de fermentescibilité du malt (Evans *et al.* 2010). Ils ont établi des modèles de type MLR et PLS. Les deux méthodes ont produit des identifications similaires des paramètres prédictifs de la fermentescibilité du moût à des niveaux semblables de capacité prédictive. Hariri (2003) a identifié et quantifié

l'influence du diagramme de trempe et de l'aération sur les qualités finales du malt, en particulier la friabilité du malt. Les conditions optimales permettant d'atteindre une friabilité entre 80 et 85%) (Point clé du cahier des charges de brasseurs) ont été identifiées à partir de ce modèle et validées expérimentalement en malterie semi-industrielle.

Les réseaux de neurones peuvent également être utilisés pour modéliser le maltage. Ce sont également des modèles de type boîte noire mais sans postulat de linéarité et dont la structure est plus complexe que les modèles polynomiaux. Aghajani *et al.* (2012) ont modélisé à l'aide de réseaux de neurones la diminution du taux d'humidité dans le malt vert en fonction du temps de séchage et de la température. Ils ont montré que les réseaux de neurones étaient plus précis que les modèles linéaires.

Lorsque la quantité d'information expérimentale augmente, les données sont souvent fortement corrélées entre elles et, les modèles de type régression multilinéaire deviennent inapplicable car l'inversion des matrices sous-jacentes n'est plus réalisable. L'analyse de telles bases de données avec des méthodes exploratoires ou spectrales est possible en appliquant les méthodes multivariées telles que l'Analyse en Composantes Principales (Principal Component Analysis PCA), la Régression au sens des moindres carrés partiels (Partial Least Square, PLS), ou la Régression en Composantes Principales (Principal Component Régression, PCR).

5.5. L'optimisation

L'optimisation est un processus, par le biais duquel, on cherche à déterminer des conditions opératoires viables qui tendent à améliorer une performance en rapprochant d'une solution optimale. (Foulods 1981, Fonseca et Fonteix, 1993).

Les méthodes d'optimisation sont nombreuses et sophistiquées et peuvent être classées selon deux groupes :

- ✓ Les méthodes déterministes : ces algorithmes permettent de trouver des solutions de manière méthodique et déterministe et ils sont utilisés pour des petits plans. Parmi ces algorithmes on peut citer le simplexe et le gradient conjugué.
- ✓ Les méthodes stochastiques : ces algorithmes permettent de se rapprocher de l'optimum recherché. Ils sont qualifiés de méthodes heuristiques. Parmi ces algorithmes, on peut citer le Recuit simulé et les algorithmes génétiques (AG).

Pour permettre une optimisation efficace des performances du procédé étudié plusieurs méthodes peuvent être utilisées. Le Recuit simulé est une heuristique très général d'optimisation, basée sur une analogie thermodynamique. L'algorithme procède par modifications aléatoires d'une solution de départ tirée au hasard. Les inconvénients de cet algorithme résident dans le fait qu'on ne peut trouver qu'un seul optimum et que sur une seule

réponse à la fois. De plus, la détermination de l'optimum recherché n'est pas garantie. Pour déterminer sûrement l'optimum il faut appliquer plusieurs fois le Recuit simulé, ce qui nécessite un temps considérable de calcul. Pour cette raison les algorithmes génétiques présentent une bonne alternative pour l'optimisation du procédé.

Les AG diffèrent des algorithmes classiques d'optimisation essentiellement en quatre points fondamentaux (Jones, 1995) :

1. Les AG utilisent un codage des éléments de l'espace de recherche et non pas les éléments eux-mêmes.
2. Les AG recherchent une solution à partir d'une population de points et non pas à partir d'un seul point.
3. Les AG n'imposent aucune régularité sur la fonction étudiée (continuité, dérivabilité, convexité...). C'est un des gros atouts des algorithmes génétiques.
4. Les AG ne sont pas déterministes, ils utilisent des règles probabilistes de transition.

5.6. L'analyse par composantes principales (PCA)

L'analyse en composantes principales (PCA) est utilisée pour la projection, l'extraction, et la compression des données multivariées (Brereton, 2000).

Les analyses PCA présentent deux avantages distincts : Tout d'abord, il s'agit d'une technique de réduction des données puissante qui peut condenser les données d'origine avec un grand nombre de variables initiales souvent inter-corrélées à un ensemble réduit de données indépendantes entre elles qui reflètent fidèlement l'information. Dans cette approche, les données expérimentales (par exemple, les spectres MIR) représentent un vaste système multidimensionnel. La méthode ACP projette ce vaste système en un espace de dimensions réduites caractérisé par des vecteurs orthogonaux appelés composantes principales. Ces composantes principales permettent d'expliquer le plus fidèlement possible la variance observée dans les données originales. Ce processus de réduction de dimensions des données est la base de nombreuses techniques d'analyses multivariées (ACP, PCR). Elle facilite en effet soit une classification des données et permet de réaliser des régressions dans des espaces de dimension réduites.

Dans un spectre, les signaux pour des longueurs d'onde voisines sont très fortement corrélés. On peut considérer que la PCA permet ainsi d'identifier les zones spectrales où l'information est la plus *explicative* des variations observées.

PCA peut accompagner les analyses conventionnelles aussi bien que les analyses spectrales. La PCA a été décrite en premier par Karl (1901) puis dans les applications pratiques par

Hotelling (1933). Nielsen and Munck (2003) ont analysé 50 orges -30 orges de printemps et 20 orges d'hiver- avec la PCA et ont montré que la PCA était un bon outil pour discriminer les orges de printemps des orges d'hiver (différentes teneurs en azote, différents quant au pouvoir diastasique), et pour prédire les effets génétiques, environnementaux des orges sur ces indicateurs de qualités des orges. Siebert (2012) a utilisé les analyses en composantes principales pour un très grand ensemble de données expérimentales récoltées sur l'orge et le malt. Ses modèles sont capables de classifier l'orge (avec/sans enveloppe, pays d'origine, variété) tandis que la classification par année de récolte et par rang est moins satisfaisante. Bravi *et al.* (2012) ont également différencié l'orge de printemps de l'orge d'hiver. Oscarsson *et al.* (1998) a réalisé une PCA sur des paramètres de qualité d'orge, ils ont trouvées à partir des *loadings* de cette PCA que les teneurs en protéines et en β -glucanes de l'orge sont corrélés négativement avec l'amidon et le rendement. La PCA est un outil utilisé pour la compression et la visualisation des données. Le PCA est également utilisée avec des données de spectroscopie Infrarouge. L'analyse PCA réalisée par Moller, (2004) a montré une discrimination des échantillons d'orge par années de culture. Cocchi *et al.* 2004 ont réussi à distinguer les farines de six différentes céréales en couplant la PCA et les analyses par spectroscopie MIR.

La recherche scientifique implique souvent d'utiliser des variables contrôlable et / ou facile à mesurer (facteurs) d'expliquer, de réglementer ou de prédire le comportement des autres variables (réponses). Lorsque les facteurs sont rares en nombre, ne sont pas significativement redondante (colinéaires), et ont une relation bien compris les réponses, puis la régression linéaire multiple (MLR) peut être un bon moyen de transformer les données en informations. Cependant, si l'une de ces trois conditions n'est pas rencontrée, les régressions MLR peuvent être inefficaces ou inappropriées. Le chercheur est souvent confronté à de nombreuses variables et à des relations mal comprises, et l'objet est simplement de construire modèle présentant une bonne faculté à prédire le constituant étudié.

5.7. La régression au sens des moindres carrés partiels (PLS)

La régression PLS est une extension du modèle de régression linéaire multiple. Il spécifie la relation linéaire entre la réponse Y et un ensemble de facteurs prédictifs X.

Dans le cas où les facteurs prédictifs sont : i) peu nombreux, ii) non corrélés significativement et iii) avec une relation claire avec la réponse, la régression linéaire multiple est la meilleure méthode. En l'absence de l'une de ces conditions, la régression linéaire multiple peut être inefficace et inappropriée.

En principe, la MLR peut être utilisé avec de très nombreux facteurs. Toutefois, si le nombre de facteurs devient trop important (par exemple, plus grand que le nombre d'observations), il est susceptible d'obtenir des modèles qui s'adaptent parfaitement au jeu de données utilisé lors du calibrage, mais qui ne parviennent pas à prédire un jeu de nouvelles données indépendantes ; on parlera alors d'un modèle sur-paramétré ou sur-ajusté. Dans ce cas, bien qu'il existe de nombreux facteurs manifestes, il peut y avoir seulement quelques facteurs latents qui représentent la majeure partie de la variation de la réponse. L'idée générale de la PLS est d'essayer d'extraire ces facteurs latents, représentant autant de la variation du facteur manifeste que possible tout en modélisant bien les réponses.

La régression au sens des moindres carrés partiels (MCP, ou PLS) (Wold, 1966 ; Hoskuldson, 1988) est également une méthode statistique utilisée pour construire des modèles prédictifs lorsque le nombre de variables est élevé et que celles-ci sont fortement corrélées. Cette méthode utilise à la fois des principes de l'analyse en composantes principales et de la régression multilinéaire. Elle consiste à remplacer l'espace initial des variables X par un espace de plus faible dimension, sous-tendu par un petit nombre de variables appelées « variable latentes » x , construites de façon itérative. Les variables X retenues sont orthogonales (non corrélées), et sont des combinaisons linéaires des variables initiales X . Contrairement aux variables résultant de l'analyse en composantes principales, les variables latentes sont obtenues à partir des variables initiales, mais en tenant compte de leur corrélation avec la variable modélisée y . Elles doivent ainsi expliquer le mieux possible la covariance entre les entrées X et la sortie y . Les variables latentes x sont alors les nouvelles variables explicatives d'un modèle de régression classique, telles que la régression linéaire multiple ($y = A x + b$).

Les modèles PLS qui exploitent des spectres MIR ont un grand succès dans le secteur agro-alimentaire et en particulier dans le domaine des boissons alcoolisées. Des modèles PLS basées sur les spectres MIR de la bière ont été réalisés par Castritius *et al.* (2012) pour prédire l'atténuation limite de la bière. Un test de Student n'a détecté aucune différence entre la précision obtenue par le couplage MIR-PLS et celle de la méthode de référence. Le couplage MIR-PLS a aussi été appliqué dans le contrôle de la qualité de bière. Iñón *et al.* (2006) ont déterminés avec succès la taux d'extrait réel, le taux d'extrait original et aussi la teneur en éthanol dans la bière par le couplage des spectroscopies MIR-NIR et la régression PLS. Des modèles au sens des moindres carrés partiels ont été établis par Lachenmeier, (2007) pour estimer la teneur en éthanol, la densité et le taux d'acide lactique dans la bière avec grande

précision ($R^2=0.97-0.98$). Les modèles PLS ont aussi été employées dans le domaine du vin : des modèles PLS pour la prédiction des tannins dans le vin ont été établis par Jensen *et al.* (2008) avec un bon coefficient de corrélation ($R^2=0.93-0.94$). Les composantes phénoliques de vin ont été quantifiés avec succès ($R^2>0.95$) en utilisant la technique FT-IR et la régression PLS (Fagoso *et al.* 2011).

Des modèles PLS ont également été établis dans le domaine de la malterie à partir de données expérimentales non spectrales. Anker-Nilssen *et al.* (2005) ont établis des modèles PLS sur quatre farines d'orge qui ont été cultivées dans des chambres à température contrôlée. Les modèles PLS ont révélés la corrélation négative entre la dégradation d'amidon et la température de croissance. Les modèles PLS ont montré que l'influence des teneurs en β -glucanes totaux et non solubles sur la dégradation d'amidon devient plus importante quand le temps d'incubation est prolongé. Nielson et Moller (2003) ont pu prédire l'extrait de malt en fonction de sept facteurs de qualité d'orge et six facteurs de qualité du malt. Leur modèle s'appuyait sur seulement deux composantes principales et conduisait à des coefficients de détermination élevés ($R^2= 0.91$). Siebert et Egi (2012) ont également établis un modèle pour prédire la teneur en protéines totales du malt en fonction de huit facteurs de qualité d'orge avec des coefficients de détermination *acceptables* ($R^2= 0.74$).

Cozzolino *et al.* (2013) ont montré que la régression PLS avec la spectroscopie NIR est un outil rapide, non destructif pour mesurer l'absorption de l'eau et la durée de trempé dans les grains d'orge complète pendant l'étape de trempé. En s'appuyant sur des mesures spectrales MIR de l'orge, Seefeldt (2009) a construit des modèles PLS pour la prédiction de la teneur en β -glucanes dans l'orge en cours de croissance en champs.

Malgré l'intérêt croissant de la chimiométrie dans le domaine agro-alimentaire et en particulier dans la filière Orge-Malt-Bière, il existe, à notre connaissance, encore peu d'applications de la régression PLS couplée à la spectroscopie MIR pour caractériser l'orge ou le malt.

Références bibliographiques du chapitre 2

Aghajani, N., Kashaninejad, M., Dehghani, A.A., Garmakhany, A.D., 2012. Comparison between artificial neural networks and mathematical models for moisture ratio estimation in two varieties of green malt. *Quality assurance and safety crops and foods* 4, 93-101.

Allaire, G., 2007. *Analyse numérique et optimization*. ©Editions de l'Ecole Polytechnique.

Al-Jowder, O., Kemsley, E.K., Wilson, R.H., 1997. Mid-infrared spectroscopy and authenticity problems in selected meats: a feasibility study. *Journal of food chemistry* 59, 195-201.

Anastasaki, E., Kanakis, C., Pappas, C., Maggi, L., del Campo, C.P., Carmona, M., Alonso, G.L., and Polissiou, M.G., 2010. Differentiation of saffron from four countries by mid-infrared spectroscopy and multivariate analysis. *European food research and technology*. 230- 571-577.

Anker-Nilssen, K., Faergestad, E.M., Sahlström, S., Uhlen, A.K., 2006. Interaction between barley cultivars and growth temperature on starch degradation properties measured in vitro. *Animal feed Science and technology* 130, 3-22.

Barr, A.R., Roumeliotis, R., Logue S.J., Jefferies, S.P., 1999. The heritability of malt extract predicted using NIR spectroscopy. *Australian Barley Technical Symposium online*.
<http://www.regional.org.au/au/abts/1999/barr2s.htm>.

Baxter, E.D., Reeves, S.G., Bamforth, C.W., 1980. The effect of increased steeping temperature on enzyme development in malt, *J. Inst. Brew*; 86, 182-185.

Bouamama, B. O. IMA 3 Automatique, 2004.
http://www.mocis_lagis.fr/wp-content/uploads/2011/10/IDENTIFICATION-DES-SYSTEMES_2012-Mode-de-compatibilit%C3%A9.pdf. www.mocis-lagis.fr/

Boubellouta, T. et Dufour, E., 2012. Cheese-Matrix Characteristics During Heating and Cheese Melting Temperature Prediction by Synchronous Fluorescence and Mid-infrared Spectroscopies. *Food and bioprocess technology* 5 273-284.

Bravi, E., Marconi, O., Perretti, G., Fantozzi, P., 2012. Influence of barley variety and malting process on lipid content of malt. *Food Chemistry* 135, 1112-1117.

Brereton, R.G., 2000. Introduction to multivariate calibration in analytical chemistry. *The analyst* 125, 2125-2154.

Briggs, D.E., Hough, J.S., Stevens, R., Young, T.W., 1981. *Malting and brewing science*. 1, malt and sweet wrot. Chapman and Hall, London, 42, 80, 121.

Brookes, P.A., Lovett, D.A. and MacWilliam, I.C., 1976. The steeping of barley. A review of the metabolic consequences of water uptake, and their practical implications. *Journal of the Institute of Brewing* 82, 14-26.

Castritius, S., Geier, M., Jochims, G., Stahl, U., Harms, D. Rapid Determination of the Attenuation Limit of Beer Using Middle-Infrared (MIR) Spectroscopy and a Multivariate Model. *Journal of agricultural and food chemistry* 60, 6341-6348.

- Chandra, G.S., Proudlove, M.O., Baxter, E.D., 1999. The structure of barley endosperm-An important determinant of malt modification. *Journal of the Science of Food and Agriculture* 79, 37-46.
- Chapon, L., 1960. Structure de grain d'orge : aptitude à la trempe et friabilité de malt, Conference faite au colloque de la malterie a Maule.
- Chen, Q.S., Zhao, J.W., and Lin, H., 2009. Study on discrimination of roast green tea (*Camellia sinensis* L.) according to geographical origin by FT-NIR spectroscopy and supervised pattern recognition. *Spectrochim. Acta A*, 72, 845–850.
- Cocchi, M., Foca, G., Lucisano, M., Marchetti, A., Pagani, A., Tassi, L., Ulrici, A., 2004. , Classification of Cereal Flours by Chemometric Analysis of MIR Spectra. *Journal of agricultural and food chemistry* 52, 1062-16-067.
- Cozzolino, D., Alder, K., Roumeliotis, S., Eglinton, J., 2012. Feasibility study on the use of multivariate data methods and derivatives to enhance information from barley flour and malt samples analyzed using the Rapid Visco Analyzer. *Journal of Cereal Science* 56, 1-5.
- Cozzolino, D., McCarthy, J., Bartowsky, E., 2012a .Comparison of near infrared and mid infrared spectroscopy to discriminate between wines produced by different *Oenococcus Oeni* strains after malolactic fermentation: A feasibility study. *Food control* 26, 81-87.
- Cozzolino, D., Cynkar, W., Shah, N., Smith, P., 2012b. Varietal Differentiation of Grape Juice Based on the Analysis of Near- and Mid-infrared Spectral Data. *Food analytical methods*. 5, 381-387.
- Cozzolino, D., Roumeliotis, S., Eglinton, J., 2013. Monitoring water uptake in whole barley (*Hordeum vulgare* L.) grain during steeping using near infrared reflectance spectroscopy. *J. food Eng.* 114, 545-549
- Da-Wen, S. *Infrared Spectroscopy for Food Quality Analysis and Control*. London 2009. À compléter.
- Ding, H.B., Xu, R.J., 2000. Near-infrared spectroscopic technique for detection of beef hamburger adulteration. *Journal of Agricultural and Food Chemistry*. 48 , 2193-2198.
- Duijhouwer, I.C., Grashoff, C., Angelino, S.f., 1993. Kernel filling and malting barley quality. *Proc. European Brewing Convention*, Oslo, 121-128.
- Edney, M. J., Rossnagel, B. G., Endoll, Y., Ozawa, S., and Brophy, M. 2002. Pearling quality of Canadian barley varieties and their potential use as rice extenders. *Jornal of Cereal Science* 36, 295-305.
- Egidio, V., Sinelli, N., Giovanelli, G., Moles, A., Casiraghi, E., 2010. NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. *European food research and technology* 230, 947-955.
- Etzion, Y., Linker, R., Cogan, R., Shmulevich, U., 2004. Determination of protein concentration in raw milk by mid-infrared Fourier transforms infrared/attenuated total reflectance spectroscopy. *Journal of dairy science* 87, 2779-2788.
- Evans, D.E., Damberg, R., Ratkowsky, D., Li, C.D., Harasymow, S., Roumeliotis, S., Eglinton, J.K., 2010. Refining the Prediction of Potential Malt Fermentability by Including an Assessment of Limit Dextrinase Thermostability and Additional Measures of Malt Modification, Using Two Different Methods for Multivariate Model Development. *Journal of the institute of brewing* 116, 86-96.

Fleury, M. D., Edney, M. J., Campbell, L. D., Crow, G. H., 1997. Total, water-soluble and acid-soluble arabinoxylans in western Canadian barleys. *Journal of Plant Science* 77, 191-196.

FLiedel, G., Grenet, C., Gontrad, N., Pons, B., 1989. Dureté caractéristiques physicochimiques et aptitude au décortilage des grains de sorgho. *Céréales en régions chauds*. AUPELF-UREF, Eds John Libbey Eurotext, Paris © pp. 187-201.

Fonseca C.M. ET Fleming P.J., " Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization ", in *Genetic Algorithms: Proceedings of the Fifth International* 1993

Foulds L.R., *Optimisations technique, an introduction*. Springer Verlag, New York. 1981

Fonteix C., Viennet R., Marc I., New multicriteria optimisation method based on the use of diploid genetic algorithm: example of an industrial problem 1995.

Fox, G. P., Onley-Watson, K., Osman, A., 2002. Multiple linear regression calibrations for barley and malt protein based on the spectra of hordein. *Journal of the Institute of Brewing* 108, 155-159.

Gamlath, J., Aldred, G.P., Panozzo, I.F., 2008. Barley (1? 3; 1? 4)-b-glucan and arabinoxylan content are related to kernel hardness and water uptake. *Journal of Cereal Science* 47, 365–371.

Guido L.F., Curto, A.F., Boivin, P., Benismail, N. Goncalves, C.R., AquilesA. Barros, A.A., 2007. Correlation of Malt Quality Parameters and Beer Flavor Stability: Multivariate Analysis. *J. Agric. Food Chem.* 55, 728-733

Gunkel, J., Voetz, M., Rath, F., 2002. Effect of malting barley variety on the fermentability. *Journal of Institute of brewery* 108, 355-361.

Gurdeniz, G., Ozen, B., Tokatli, F., 2008. Classification of Turkish olive oils with respect to cultivar, geographic origin and harvest year, using fatty acid profile and mid-IR spectroscopy. *European Food Research and Technology* 227, 1275–128

Hariri, A., *Etude et modélisation de la trempe en malterie*, Thèse de l'Institut National Polytechnique de Lorraine, 2003.

Hashimoto, A., Mori, H., Kanou, M., Yamanaka, A., Kameoka, T., 2004. Mid-Infrared spectroscopic analysis on brewed coffee characteristics. In *Proc of 10th Asian Pacific Confederation of Chemical Engineering Conference (CD-ROM version)*, The Society of Chemical Engineers, Japan: Tokyo; 2004, p. 3P-01-068.

Hodson, O.P., 1986. Inter-relationship in malt analysis and malting parameters. *Proc. Conv. Inst. Brew. (Austr. And new Zealand Sect)*, 109-114.

Holland. J.K., Kemsley, E.K., Wilson, R.H., 1998. Use of Fourier Transform infrared spectroscopy and chemometrics for the detection of adulteration of strawberry purees. *Journal of science food and agriculture* 76, 263-269.

Holopainen, U.R., Willhelmsen, A., Marttila, M.S., Sainio, P.P., Arirajala, A., Reinikainen, P., Kotaviita, E., Simolin, H., Home, S., 2005. Endosperm Structure Affects the Malting Quality of Barley. *Journal of Agriculture and Food Chemistry* 53, 7279-7287.

Höskuldson, A., 1988. PLS regression methods. *Journal of Chemometrics* 2, 211-228.

Hough, J.S., 1985. *The biotechnology of malting and brewing*. Cambridge studies in biotechnology 1.

Igathinathane C., Chattopadhyay P.K., 1999. Moisture diffusion modelling of drying in parboiled paddy components. Part I: strachy endosperm, *Journal of Food Engineering*, 41, 89-101.

Inon, F., A., Garrigues, J.M., Garrigues, S., Molina-Diaz, A., de la Guardia, M., 2003. Selection of calibration set samples in determination of olive oil acidity by partial least square-attenuated total reflectance-Fourier transform infrared spectroscopy. *Journal of analytica chemica acta* 489, 59-75.

Inon, F., A., Garrigues, J.M., de la Guardia, M., 2004. Nutritional parameters of commercially milk samples by FTIR and chemometric techniques. *Journal of analytica chemica acta* 513, 401-412.

Inon A. F., Salvador Garrigues, S., Miguel de la Guardia, M., 2006. Combination of mid- and near-infrared spectroscopy for the determination of the quality properties of beers. *Journal of analytica chemica acta* 571, 167-174.

Jacques, J., Modélisation statistique. POLYTECH'LILLE <http://labomath.univ-lille.fr/~jacques/>

Jacobsen, S., Sondergaard, I., Moller, B., Desler, T., Munk, L., 2005. A chemometric evaluation of the underlying physical and chemical patterns that support near infrared spectroscopy of barley seeds as a tool for explorative classification of endosperm genes and gene combinations *Journal of Cereal Science* 42, 281-299.

Jensen, J.S., Edebo, M., Meyer, A.S., 2008. Identification of spectral regions for quantification of red wine tannins with fourier transform mid-infrared spectroscopy. *Journal of agricultural and food chemistry* 56, 3493-3502.

Karoui, R., Downey, G., Blecker, C., 2010. Mid-Infrared Spectroscopy Coupled with Chemometrics: A Tool for the Analysis of Intact Food Systems and the Exploration of Their Molecular Structure-Quality Relationships - A Review. ©American Chemical Society, 2010. In: <http://tin.dstc.uniud.it/~gambi/cr100090k.pdf>.

Karoui, R., Dufour, E., Pillonel, L., Picque, D., Cattenoz, T., Bosset, J.O., 2004. Determining the geographic origin of Emmental cheeses produced during winter and summer using a technique based on the concatenation of MIR and fluorescence spectroscopic data. *European food research and technology* 219, 184-189.

Karoui, R; Dufour, E; Pillonel, L., Schaller, E., Picque, D., Cattenoz, T., Bosset, J.O., 2005. The potential of combined infrared and fluorescence spectroscopies as a method of determination of the geographic origin of Emmental cheeses. *International dairy journal* 15, 287-298.

Karoui, R., Lefur, B., Grondin, C., Thomas, E., Demeulemester, C., De Baerdemaeker, J., Guillard, A.C., 2007. Mid-infrared spectroscopy as a new tool for the evaluation of fish freshness. *International journal of food science and technology*. 42, 57-64.

Karoui, R., Mazerolles, G., Bosset, J.O., de Baerdemaeker, J., Dufour, E., 2007. Utilization of mid-infrared spectroscopy for determination of the geographic origin of Gruyere PDO and L'Etivaz PDO Swiss cheeses. *Food chemistry* 105, 847-854.

Kays, S.E., Barton, F.I. Windham, W.R., 2000. Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. *J. Near Infra-red spectroscopy*. 8, 35-43.

Kelly, J.D., Petisco, C., Downey, G., 2006. Application of Fourier transform mid infrared spectroscopy to the discrimination between Irish artisanal honey and such honey adulterated with various sugar syrups. *Journal of agricultural and food chemistry* 54, 6166-6171.

Kettunen, A., Hamalainen, J.J., Stnholm, k., Pietila, K., 1996. An model for the prediction of β -glucanes activity and β -glucanes concentration during mashing, *Journal of Food Engineering*, 29,185-200.

Kirin Institute of Food and Life style, Global Beer Production by Country in 2010, Report 30, 2011. http://www.kirinholdings.co.jp/english/news/2011/0811_01.html.

Kreisz S., ,Hartmann K., , Zarknow M., Thiele F., Burberg F., Back W., 2007. A new statistical method to evaluate the malting performance of new barley varieties. Proceeding of the EBC congress-CD-ROM edition, congress, 31st, 20 pages.

Koca, N., Kocaoglu-Vurma, N.A., Harper, W.J., Rodriguez-Saona, L.E., 2010. Application of temperature-controlled attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy for rapid estimation of butter adulteration. *Journal of food chemistry* 121, 778-782.

Lachenal, G., 1998. NIR spectroscopy analysis and its applications to polymers analysis. *Analisis* 26, 20-28.

Lachenmeier, D.W., 2007. Rapid quality control of spirit and beer using multivariate data analysis of Fourier transform infrared spectra. *Journal of Food Chemistry* 101, 825-832.

Lewis, M.J., et Young. T.W., *Barley, Brewing*, Chapman & Hall, London (1995), 36–47.

Lichtenberg, B., Hedtke, K., Bienefeled, K., 2002. Infrared spectroscopy in routine quality analysis of honey *Apidologie* 33, 327-337.

Maamouri, O., Rouissi, H., Dridi, S., Kammoun, M., De Baerdemaeker, J., Karoui, R., 2008., Mid infrared attenuated total reflection spectroscopy as a rapid tool to assess the quality of Sicilo-Sarde ewe's milk during the lactation period after replacing soybean meal with scotch bean in the feed ration. *Journal of food chemistry* 106, 361-368.

Manger., H.J. *Technology Brewing and Malting*. 4th. Berlin(VLB) 2010.

March, T.J., Richter, D., Colby, T., Harzen, A., Schmidt, J.R., Pillen, K., 2012. Identification of proteins associated with malting quality in a subset of wild barley introgression lines. *Proteomics* 12, 2843–2851.

Marte, L., Belloni, P., Genorini, E., Sileoni, V., Perretti, G., Montanari, L., Marconi, O., 2009. Near-infrared reflectance models for the rapid prediction of quality of brewing raw materials. *Journal of Agriculture and Food Science*, 57, 326-333.

Maurer, N.E., Rodriguez-Saona, L., 2013. Rapid Assessment of Quality Parameters in Cocoa Butter Using ATR-MIR Spectroscopy and Multivariate Analysis. *Journal of the American oil chemist society* 90, 475-481.

Mayolle, J.E., Lullien-Pellerin, V., Corbineau, F., Boivin, P., Guillard, V., 2012. Water diffusion and enzyme activites during malting of balrey grains : a relationship assessment. *Journal of food engineering* 109, 358-365.

McParland, S., Banos, G., McCarthy, B., Lewis, E., O'Neill, B., O'Donovan, M., Wall, E., Berry, D.P., 2012. Validation of mid-infrared spectrometry in milk for predicting body energy status in Holstein-Friesian cows. *Journal of dairy science* 95, 7225-7235

McTaggart, I.P., Smith, K.A., 1992. The effect of fertiliser and soil nitrogen on the overall uptake of nitrogen in the plant, and the grain nitrogen content of spring-sown malting barley. HGCA Project Report 46.

Molina-Cano, J.L.; Sopena, A.; Polo, J.P.; Bergareche, C. ; Moralejo, M.A.; Swanston, J.S.; Glidewell, S.M., 2002. Relationships between barley hordeins and malting quality in a mutant of cv. triumph. II. Genetic and environmental effects on water uptake. *Journal of cereal science*, 36, 39-50.

Moller, B., 2004. Near infrared transmission spectra of barley of malting grade represent a physical-chemical fingerprint of the sample that is able to predict germinative vigour in a multivariate data evaluation model, *J. Inst. Brew.* 110, 18-33.

Narziss, L., et Friedrich. 1970. Influence of falling temperature malting an resteeeping on enzyme activity, *Brauwiss*; 23, 265-271.

Nielsen, J.P., and Munck, L., 2003. Evaluation of malting barley quality using exploratory data analysis. I. Extraction of information from micro-malting data of spring and winter barley. *Journal of cereal science* 38, 173-180.

Nunes, A., Taylor, J.N., Delgadillo, I., 1997. *Spectroscopy of biotechnological molecules: modern trends*; kluwer academic publisher: Dordrecht, 289.

Oscarsson, M., Andersson, R., Åman, P., Olofsson, S., Jonsson, A., 1998. Effects of cultivar, nitrogen fertilization rate and environment on yield and grain quality of barley. *Journal of the Science of Food and Agriculture* 78, 359–366.

Ouassar, Y., 1998. Réseaux d'ondelettes et réseaux de neurones pour la modélisation statique et dynamique de processus. These. Univ Pierre et Marie Curie, Paris1.

Palmer G.H., 1989. Cereals in malting and brewing, *Cereal Science and Technology* , 147–165.

Parisi, A.F., Nogueiras, L., Prieto, H., 1990. On-line determination of fuel quality parameters using near-infrared spectrometry with fibre optics and multivariate calibration *Analytical Chemistry Acta* 238, 95-100.

Passarella, V.S., Savin, R., Slafer, G.A., 2002. Grain weight and malting quality in barley as affected by brief periods of increased spike temperature under field conditions. *Australian journal of agricultural research* 53, 1219-1227.

Paynter, B.H., Young, K.J., 2004. Grain and malting quality in two-row spring barley are influenced by grain filling moisture. *Australian journal of agricultural research* 55, 539-550.

Paynter, B.H., Harasymow, S.E., 2010. Variation in grain β -glucan due to site, cultivar and nitrogen fertilizer in Western Australia. *CROP & PASTURE SCIENCE*, 61, 1017-1026

Peltonen, J., Rita, H., Aikasalo, R., Home, S., 1994. Hordein and malting quality in northern barleys. *Hereditas* 120, 231- 239.

Picque,D., Cattenoz, T., Corrieu, G., 2002. Discrimination of Emmental cheeses by infrared spectroscopy. *Milchwissenschaft*, 57, 202–204

Chapitre 2: Bibliographie

Powell, G.E., Chambley, J.D., Cole, N.M., 1993. The effect of oxygen and carbon dioxide levels on enzyme development in malting. *Proc. Conv. Inst. Brew. (Cent. And South African Sect)*, Somerset West, 4, 45-54.

Psota, V., Vejrazka, K., Famera, O., Hrcak, M., 2007. Relationship between grain hardness and malting quality of barley (*Hordeum vulgare* L.). *Journal of the Institute of Brewing* 113, 80–86

Reid, L.M., Woodcock, T., O'Donnell, C.P., Kelly, J.D., Downey, G., 2005. Differentiation of apple juice samples on the basis of heat treatment and variety using chemometric analysis of MIR and NIR data. *Food research international* 38, 1109-1115.

Reeves, J.B., and Delwiche, S.R. 1997. Determination of protein in ground wheat samples by mid-infrared diffuse reflectance spectroscopy. *Application Spectroscopy* 51, 1200–1204.

Ridgway, C., Chambers, J., Cowe, I.A., 1999. Detection of grain weevils inside single wheat kernels by a very near infrared two-wavelength model. *J. Near Infrared spectroscopy*. 7, 213-221.

Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., Jent, N., 2007. A review of near infrared spectroscopy and chemometrics. *Journal of pharmaceutical and biomedical analysis* 44, 683–700.

Rouf, K., Luginbuhl, W., Kunzli, R., Teresa, M., Bogdanov, S., Bosset, J.O., Von der Ohe, W., Amado, R., 2006. Authentication of the botanical and geographical origin of the honey by Mid-Infrared spectroscopy. *Journal of agricultural and food chemistry* 54, 6873-6880.

Seefeldt, H.F., Berg, F.V., Kockenberger, W., Engelsens, S.B., Wollenweber, B., 2007. Water mobility in the endosperm of high β -glucan barley mutants as studied by nuclear magnetic resonance imaging. *Magnetic Resonance Imaging* 25, 425-432.

Shao, Y., Zhao, C., He, Y., Bao, Y., 2008. Infrared spectroscopy and chemometrics for the starch and protein prediction in irradiated rice. *Journal of food chemistry* 126, 1856-1861.

Sheehy, M., Marafioti, A., Krawec, C., Löfqvist, B., Stewart, D., 2009. Active starch management-activity undoing malt quality-. *Proc. 14th Australian barley technical symposium*, Brisbane.

Sileoni, V., Perretti, G., Marte, Marte, L., Marconi, O., Fantozzi, P., 2010. Near-Infrared Spectroscopy for Proficient Quality Evaluation of the Malt and Maize Used for Beer Production. *Journal of the Institute of brewing* 116, 134-140.

Sotnikova, V., Hrivna, L., Hartman, I.v.o., 2013. Influence of Starch Content on the Germination Speed of Spring Barley Seeds. *Book Group Authors: CULS Prague. Source: Seed and Seedlings* 118-123.

Tarantilis, P.A., Troianou, V.E., Pappas, C.S., Kotseridis, Y.S., and Polissiou, M.G., 2008. Differentiation of Greek red wines on the basis of grape variety using attenuated total reflectance Fourier transform infrared spectroscopy. *Food Chemistry* 111: 192-196.

Tarr, A., Diepeveen, D., Appels, R., 2012. Spectroscopic and chemical fingerprints in malted barley. *Journal of Cereal Science* 56, 268-275.

- Tay, A., Singh, R.K. Krishnan, S.S., Gore, J.P., 2002. Authentication of olive oil adulterated with vegetable oils using Fourier transforms infrared spectroscopy. *Lebensmittel-Wissenschaft und-Technologie-food science and technology* 35, 99-103.
- Tinsson, W., *Plans d'expérience: constructions et analyses statistiques*, Mathématiques et Applications 67, DOI 10.1007/978-3-642-11472-4 1, Springer-Verlag Berlin Heidelberg 2010
- Tobias, R.D., SAS Institute Inc., Cary, N C. An Introduction to Partial Least Squares Regression. <http://www.ats.ucla.edu/stat/sas/library/pls.pdf>.
- Toole, G.A., Le Gall, G., Colquhoun, I.J., Drea, S., Opanowicz, M., Bedo, Z., Shewry, P.R., Mills, E.C., 2012. Spectroscopic analysis of diversity in the spatial distribution of arabinoxylan structures in endosperm cell walls of cereal species in the HEALTHGRAIN diversity collection. *Journal of cereal science* 56, 134-141.
- Torano, J.S., van Hattum, S.C. 2001. Quantitative analysis of active compounds in pharmaceutical preparations by use of attenuated total-reflection Fourier transform mid-infrared spectrophotometry and the internal standard method. *Journal of analytical chemistry* 371, 532-535.
- Tsuchikawa, S., Schwanninger, M., 2013. A Review of Recent Near-Infrared Research for Wood and Paper (Part 2). *Applied spectroscopy reviews* 48, 560-587
- Turnbull, K.M., Rahman, S., 2002. Endosperm texture in wheat. *Journal of Cereal Science* 36, 327-337.
- Van campenhout, L., Shen, H.Y., Iserentant et Verchtert, 1999. The gas environment of germination barely in various microbial states during malting. *Process biochemistry*, 34, 929-937.
- Vansoest, J.G., De wit, D., Tournois, H., Vliegthart, J.G., 1994. The influence of glycerol on structural changes in waxy maize starch as studied by Fourier transforms infra-red spectroscopy. *Polymer* 35, 4722-4727.
- Vejrazka, K., Psota, V., Ehrenbergerova, J., Hrstkova, P., 2008. Relationship between grain milling energy and malting quality of barley. *Cereal Research Communications* 36, 97-105.
- Viennet R., *Nouvelle technologie de planification expérimentale pour l'optimisation multicritères de procédés*. Thèse de l'Institut National Polytechnique de Lorraine, Nancy, France, 1997.
- Wallon, J., Yan, S.H., Tong, J., Meurens, M., Haot, J., 1994. Identification of breast carcinomatous tissue by near-infrared reflectance spectroscopy. *Applied Spectroscopy*. 48, 190-193.
- Wang, X.J., Yuan, X.F., Wang, H., Li, J., Wang, X.F., Zong jun Cui, Z.J., 2011. Characteristics and community diversity of a wheat straw-colonizing microbial community. *African Journal of Biotechnology* 10, 7853-7861.
- Williams, P.C., Sobering, D.C., 1986. Attempts at standardization of hardness testing of wheat. 1. The grinding sieving (particle size index) method. *Cereal Foods World* 31, 359-364.
- Wold, H. Estimation of principal components and related models by iterative least squares, in *Multivariate Analysis*, Krishnaiah, P.R., Editor. 1966, New York : Academic Press. p. 391-420.
- Woodcock, T., Fagan, C.C., O'Donnell, C.P., Downey, G., 2008. Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity. *Food and bioprocess technology* 1, 117-129.

Wu, D., Nie, P.C., He, Y., Bao, Y.D., 2012. Determination of Calcium Content in Powdered Milk Using Near and Mid-Infrared Spectroscopy with Variable Selection and Chemometrics. *Food and bioprocess rechnology* 5,140-1410.

Xue, D. W., Dong, J., Wu, F. B., Zhang, G. P., 2008. Genotypic and environmental variation in grain protein components and their relations to β -amylase and β -glucanase activity in malting barley. *Cereal research communications*, 36, 125-134.

Yalvac, E.D., Seasholtz, M.B., Beach, M.A., Crouch, S.R., 1997. Real-time analysis of light alkenes at elevated temperatures and pressures by fiber-optic near-infrared spectroscopy. *Applied. Spectroscopy*. 51, 1565-1572.

Yang H., Irudayaraj, J., 2002. Rapid determination of vitamin C by NIR, MIR and FT-Raman techniques. *Journal of Pharmacy Pharmacology*. 54, 1247-55.

Zapotocny, P., Zielinska, M., Nita, Z., 2008. Application of image analysis for the varietal classification of barley: Morphological features. *Journal of Cereal Science* 48, 104-110.

Zhang, G.P.; Wang, J.M.; Chen, J.X., 2002. Analysis of β -glucan content in barley cultivars from different locations of China. *Food chemistry*, 79, 251-254.

Chapitre 3

Matériel et méthodes

Matériel et méthodes

Ce chapitre décrit le matériel utilisé et les méthodes d'analyses mises en œuvre pour

- Le micromaltage
- La caractérisation de l'orge et du malt
- Les régressions et analyses statistiques des données expérimentales

1. Transformation et analyses conventionnelles de l'orge et du malt

1.1. Matière première : l'orge

Pour la régression MLR, 40 orges de la récolte 2006 ont été utilisées dans le micro maltage ANNEXE 1 et ANNEXE 2. Le lot est composé de deux variétés d'orge d'hiver OH (Azurel, Esterel), et trois variétés d'orge de printemps OP (Scarlett, Sébastien, Carafe).

Pour l'analyse par ACP/PLS: trois lots de trois années de récolte 2009, 2010 et 2011, représentant 394 orges ont été utilisés,

ANNEXE 3, ANNEXE 4, **Erreur ! Source du renvoi introuvable.**, ANNEXE 5, ANNEXE 6, ANNEXE 7 ANNEXE 8 ANNEXE 9, ANNEXE 10. Ils représentent 72 variétés différentes ANNEXE 12, se répartissent en 138 orges d'hiver et 256 orges de printemps, collectés en 16 lieux en France, ANNEXE 13.

1.2. Conditions de maltage

Le maltage a été réalisé en micro-malterie à l'IFBM (Institut Français des boissons de la brasserie et de la malterie). La micro malterie est constituée de trois enceintes où s'effectue pour chacune d'entre elles, une des étapes du maltage. L'ensemble est contrôlé par un automate programmable. On peut malter en simultané 8 échantillons d'orge de 1 kg. Dans cette micro-malterie on distingue :

La trempé : cette salle est climatisée avec une plage de température programmable de 10 C° à 25 C°. La température de la salle de trempé est régulée à 18 C°. Les échantillons d'orge sont placés dans des casiers de 2 kg. Tous les paramètres importants de la trempé (température, durée des phases sous eau, sous air pour chaque cabinet, température de trempé pour chaque casier) sont régulés. Les conditions de maltage (diagramme de trempé, température et durée de germination et de touraillage) dans cette étude sont présentées dans le Tableau 4:

Tableau 4 Diagrammes de trempe pour les échantillons d'orge de quatre années de récolte

Diagramme de trempe à 18 C°							
		Sous eau 1	Sous air 1	Sous eau 2	Sous air 2	Sous eau 3	Température
2006	Diagramme 1	11h	17h	7h			22C°
	Diagramme 2	8h	6h	5h			22C°
	Diagramme 3	11h	10h	14h			14C°
	Diagramme 4	6h	21h	5h			14C°
2009	Orge de printemps	8h	17h	6h			18C°
	Orge d'hiver	8h	16h	7h	10h	4h	18C°
2010	Orge de printemps	11h	17h	7h			18C°
	Orge d'hiver	10h	18h	7h	8h	3h	18C°
2011	Orge de printemps	10h	20h	7h			18C°
	Orge d'hiver	10h	18h	7h	8h	3h	18C°

À la fin de trempe, les paniers sont retirés de la cuve de trempe, pour être transférés en salle de germination

La germination : les orges sont ensuite placées dans des paniers de 1 kg. Ces paniers sont disposés autour d'un arbre central situé au milieu de la pièce et animé d'un mouvement de rotation continu. Cette rotation aide à éviter la prise en masse au cours de la gémiation, et favorise l'aération des grains. La température de la salle de germination est maintenue à 16 C°. Après cinq jours, les paniers cylindriques sont retirés du germoir. Le malt vert obtenu est alors transvasé dans les bacs de touraillage en inox et percés au fond-afin de laisser passer l'air chaud pour sécher les grains.

Le touraillage : Une *touraille* d'une capacité de huit fois 1 kg est utilisée. Tous les paramètres de séchage (nombre et durées et chaque palier de température) sont programmés et présentés dans le schéma suivant :

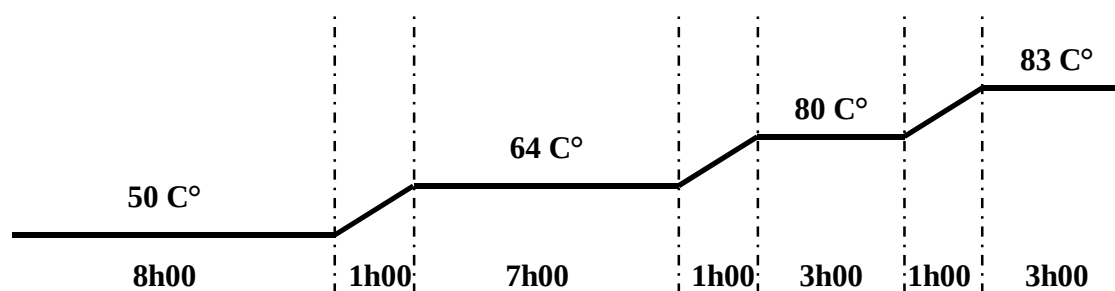


Figure 9 protocole de touraillage utilisé

1.3. Méthodes analytiques classiques de l'orge

Le processus de maltage a été suivi avec des analyses effectuées sur les grains d'orges initiaux, des analyses en cours de trempe, en cours de la germination, sur le malt vert et sur le malt final. Toutes les analyses suivent les recommandations de l'European Brewery Convention (EBC, 2010).

1.3.1. Mesure de l'humidité

La teneur en humidité pour les grains d'orge et le malt vert est mesurée par deux méthodes :

Une méthode *classique* [EBC 3.2]: deux échantillons de 5g broyés en fine mouture. La dessiccation se fait dans un four à 130 C° pendant deux heures. Les échantillons sont ensuite placés dans un dessiccateur une demi-heure avant d'être pesés. Le taux d'humidité est déterminé par différence de poids avant et après le séchage et exprimé en (%) par rapport à la masse sèche.

Une méthode *rapide* : la dessiccation des échantillons se fait à l'aide d'une balance à infrarouge (Sartorius®) à une température de 120 C°.

Les résultats des deux méthodes ont été comparés par le test d'égalité des moyennes et des variances et ont été jugé identiques. La deuxième méthode a été adoptée pour des raisons de gain de temps.

Une mesure de taux d'humidité toutes les heures avec la méthode rapide a été réalisée pour un suivi de taux d'humidité pendant la trempe décrite dans le chapitre 3.

1.3.2. Teneur en protéine de l'orge

Ce test a pour objectif contrôlé la qualité des matières premières. La mesure de la teneur en protéines s'effectue sur un échantillon de farine d'orge de granulométrie 0.2 mm selon la méthode de Kjeldahl [EBC 4.3.1]. L'échantillon est minéralisé en présence d'acide sulfurique et d'un catalyseur. L'ammoniac formé est déplacé de son sel par une base forte (NaOH), et entraîné par la vapeur dans une solution d'acide borique, puis titré par une solution diluée d'acide fort (H₂SO₄ 0.05 M) en présence d'un indicateur coloré. La quantité de protéines est obtenue en multipliant la valeur obtenue par 6.25 et est exprimée en % sur la matière sèche..

1.3.3. Taux de β -glucan

Selon [EBC 3.10.2]. Les β -glucanes présents dans l'orge doivent être rendus solubles avant l'analyse par le système. La solubilisation de des β -glucanes de l'orge est réalisée par une dégradation partielle en utilisant une hydrolyse acide douce. Le calcofluor fluorochrome forme des complexes avec les β -glucanes- ayant un poids moléculaire élevé (dessus de 104 dalton) en solution selon laquelle une augmentation de l'intensité de la fluorescent Colorant Calcofluor se produit. La fluorescence calcofluor est contrôlée dans un système automatique basé sur l'écoulement d'injection de l'analyse principale, la quantification reproductible de la fluorescence et donc de β -glucane peut être obtenue. L'appareil est calibré par rapport aux normes en β -glucane pure.

1.3.4. Taux d'arabinoxylane

La mesure du taux d'arabinoxylane consiste à hydrolyser l'arabinoxylane en présence d'acide sulfurique qui conduit essentiellement à libération de xylose. Les monosaccharides résultants sont transformés en milieu acide et à chaud en furfural. Le furfural formé est mélangé avec du phloroglucinol pour créer un complexe coloré. En utilisant une solution d'étalonnage de xylose et un spectromètre UV, les résidus de xylose montrent une forte absorption à 552 nm et une faible absorption à 510 nm. Comme les résidus de glucose montrent la même absorption à 552 nm et 510 nm, la teneur en arabinoxylane est déduite à partir de la différence d'absorption à 552 et 510 nm.

1.3.5. Indice de germination (IG) :

La mesure de l'indice de germination a pour but de déterminer la capacité germinative des grains d'orges en appliquant la méthode d'Aubry et Schönefeld qui consiste à mettre 500 grains d'orge sur des couches métalliques recouverts par du coton. Les couches humidifiées sont mises dans une armoire thermostatée à 20 °C. Les grains non germés sont comptés après 72 h et 120 h et à chaque comptage les grains germés sont retirés de l'échantillon. On exprime alors un pourcentage des grains germés après 72 h et 120 h.

1.3.6. Pourcentage des grains piqués (GP) :

Le mot piqué est utilisé pour désigner les grains où le germe a percé l'enveloppe du grain. On calcule le pourcentage de grains (parmi 100) ayant des piques en sortie de trempe. Les grains piqués sont comptés, un par un, à l'œil nu. Cet indicateur renseigne sur l'efficacité du diagramme appliqué pendant la trempe (activation de la germination).

1.3.7. Indice de trempe (test de Chapon) :

L'indice de trempe (Chapon, 1960), renseigne sur la distribution d'eau dans l'endosperme. Le test s'effectue en mettant 100 grains d'orge dans l'eau bouilli pendant 30 secondes. Les grains sont coupés en deux longitudinalement. L'examen des 100 demi-grains permet de classer les grains selon leur degré de translucidité en quatre catégories :

1^{ère} catégorie (a) : Complètement translucide

2^{ème} catégorie (b) : $\frac{3}{4}$ translucide

3^{ème} catégorie (c) : $\frac{1}{2}$ translucide

4^{ème} catégorie (d) : $\frac{1}{4}$ translucide

L'indice de trempe est calculé par la relation $I.T = 1 \times a + 2 \times b + 3 \times c + 4 \times d$

Selon la distribution de l'eau dans l'endosperme, cet indice peut varier entre 100 et 400. Pour atteindre une bonne distribution d'eau dans l'endosperme, l'indice de trempe doit avoir une valeur proche de 100.

1.3.8. La mesure de la porosité

Le protocole suivi pendant la mesure de porosité d'orge est présenté dans Figure 10

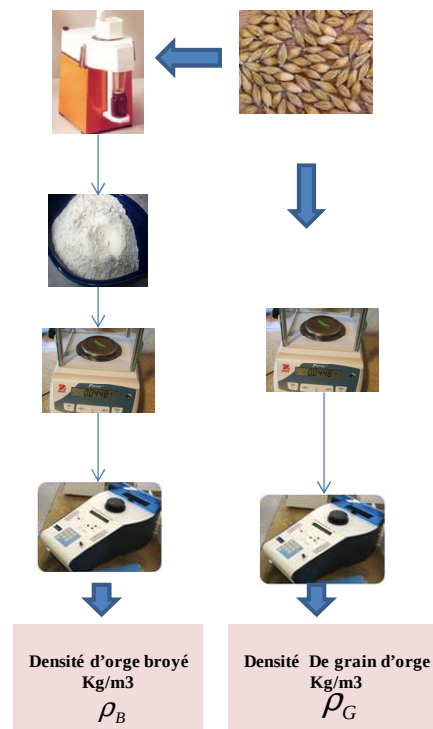


Figure 10 le protocole de mesure de la porosité d'orge a droite

D'abord, et à l'aide d'un pycnomètre pour gaz (QuantachromCrop., Syosset, NY) on calcule la densité des grains complets ρ_G puis la densité ρ_B des farines fines des grains broyés par un

moulin CyclotecTM 1093 muni d'une grille de 1.0 mm (FOSS). Dans une étape intermédiaire, on relève la teneur en eau dans les grains d'orge par séchage dans une étuve à 130 °C pendant 2h. La porosité des grains est alors donnée par la formule:

$$\varepsilon = \left[1 - \frac{\rho_G}{\rho_B} \right] \times 100 \quad (3)$$

ε est la porosité %, ρ_G est la densité des grains d'orge, ρ_B est la densité de farine de grain d'orge. Les résultats sont présentés dans l'ANNEXE 11.

Dans notre étude, les mesures de porosité sont réalisées sur des échantillons présentant un maximum de variabilité. Ils seront constitués de 12 orges de 2007(Meuse), 12 orges de 2008 (Nord) et 12 orges 2008 (Côte d'Or).

1.3.9. La mesure de la dureté des grains

Plusieurs techniques pour mesurer la dureté des grains sont disponibles dans quatre grands groupes selon la base de : broyage, concassage, abrasion ou pénétration au stylet (Anjum et Walker, 1991).

A notre connaissance aucun de ces méthodes n'a été standardisé. Dans cette étude, nous avons utilisé la méthode de pénétration d'un stylet.

1.4. Méthodes analytiques classiques du malt

1.4.1. Taux de friabilité du malt:

La mesure de la friabilité mesure le potentiel de désagrégation physique du grain selon [EBC 4.15]. On soumet 50 g de grains de malt à une érosion mécanique dans un tambour grillagé rotatif standardisé pendant 8 min. les petits fragments de grains physiquement modifiés passent à travers le grillage alors que les plus gros fragments sont retenus.

La masse de grains non modifiés A, retenue dans le tambour après 8 min est déterminée et la friabilité en est déduite par le calcul :

$$\text{Friabilité (\%)} = 100 - (2 * A) \quad (4)$$

Appréciation	Friabilité %
très bon	>81
bon	78-81
moyen	75-78
insuffisant	<75

1.4.2. Extrait fine mouture:

L'extrait fine mouture renseigne sur le rendement du brassage ou la quantité de matières solubilisées au cours de ce brassage conventionnel selon [EBC 4.5.1]. après brassin conventionnel, l'extrait est calculé à partir de la densité du moût à 20 C en utilisant un pycnomètre, et à l'aide des tables officielles de Plato. On détermine la teneur en extrait en g pour 100 g de mout en utilisant la formule :

$$E = \frac{[100 \times P \times (H_1 + 800)]}{[(100 - P) \times (100 - H_1)]} \quad (5)$$

Où P la teneur en extrait du moût

H_1 humidité du malt

800 la quantité d'eau distillée ajoutée au brassage pour 100 g de malt

1.4.3. Taux de protéines solubles :

Détermination de l'azote présent dans le mout. Le taux de protéines solubles permet d'examiner l'efficacité des enzymes protéolytiques du malt selon [EBC 4.9.1]. il commence par un brassin EBC puis une dosage de l'azote du mout par la méthode de Kjeldahl déjà décrite pour l'orge

1.4.4. La viscosité du mout

Cette méthode a pour but de contrôler la qualité des matières premières [EBC 4.8]: filtrabilité/désagréation. Le principe est de réaliser un brassin EBC. La mesure de la viscosité du mélange est calculée par la détermination instrumentale de la vitesse de chute d'une bille dans le liquide par l'équation:

$$M = K \times t(\mu_K - \mu_N) \quad (6)$$

Où M = viscosité du mout, K = constante de calibration, t =temps de chute de la bille, μ_K =masse volumique de la bille, μ_N =masse volumique du mout.

1.4.5. Atténuation limite :

L'atténuation réfère la qualité des sucres présents dans le mout (fermentiscibles ou non) [EBC 4.11].

Cette atténuation s'exprime par rapport à l'extrait de départ de la fermentation. Cet indice quantifie la fermentescibilité du malt. On utilise le mout conventionnel réalisé pour mesurer l'extrait. Les enzymes amylolytiques ensuite sont inactivés par maintien à une température supérieure à 90 °C pendant 15 min. après refroidissement le volume est corrigé. L'atténuation est calculée à partir des densités des mouts avant ensemencement et après fermentation.

1.4.6. Pouvoir diastasique :

L'objectif de ce test est de renseigner sur l'activité enzymatique du malt (α et β -amylases) sans spécificité d'action. Le but de cette analyse est de mesurer l'activité α - et β -amylasique dans des conditions de réaction normalisée. Méthode [EBC 4.12] Les enzymes du malt sont extraites dans l'eau distillée à 40 °C. Une solution d'amidon normalisée est alors hydrolysée par l'extrait obtenu. Les sucres réducteurs formés sont mesurés par Iodométrie. Les résultats sont exprimés en gramme de maltose par 100 g de malt ou en unité Windich Kolbach (W.K).

2. Méthodes d'analyses spectrales de l'orge

2.1. Techniques de spectroscopie Infrarouge (IR)

Les techniques spectroscopiques ont pour objectif principal la caractérisation (chimique ou structurale) de la matière. Elles mettent en évidence les propriétés de cette dernière en étudiant ses interactions avec un rayonnement électromagnétique. Les méthodes spectroscopiques se basent sur l'absorption ou l'émission de radiation du spectre électromagnétique par beaucoup de molécules quand les électrons changent de niveaux énergétiques. L'interaction de l'énergie de la lumière avec la matière est utile pour déterminer les composés et leur concentration dans des mélanges. Le spectre électromagnétique se situe dans la gamme des rayons- γ jusqu'aux ondes radio de très basse énergie (Figure 11)

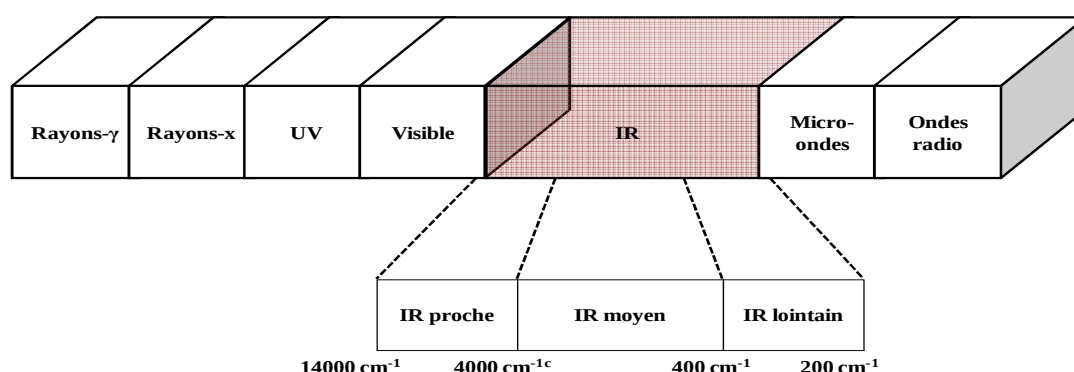


Figure 11 le spectre électromagnétique

Dans le cas de la spectroscopie infrarouge, les énergies mises en jeu sont les énergies vibrationnelles et rotationnelles des liaisons moléculaires. L'infrarouge est donc utilisé pour obtenir des informations sur la composition chimique des échantillons. La théorie de la spectroscopie infrarouge a été bien décrite par nombreux auteurs. Une description détaillée de la spectroscopie infrarouge est disponible dans les ouvrages de Banwell (1983) et Osborne (1993).

2.1.1. Principes de spectroscopie infrarouge:

La plage du spectre électromagnétique correspondant à l'infrarouge se situe entre le visible et les micro-ondes, 800 nm et 1 cm.

Trois zones sont distinguées :

- ✓ La proche infrarouge (PIR) entre 12500-4000 cm^{-1} où l'on observe les bandes d'absorption correspondant aux harmoniques et aux bandes de combinaisons des vibrations.

- ✓ Le moyen infrarouge (MIR) entre $4000-400\text{ cm}^{-1}$ où l'on observe les bandes d'absorption correspondant aux vibrations fondamentales.
- ✓ L'infrarouge lointain entre $400-1\text{ cm}^{-1}$

Les bandes d'absorption résultant de la vibration des liaisons moléculaires. On observe deux types principaux de vibration. D'une part, on distingue les vibrations où il y a un changement continu de la distance interatomique selon l'axe de la liaison (les élongations). D'autre part, les vibrations impliquant un changement d'angle : (les déformations) Figure 12.

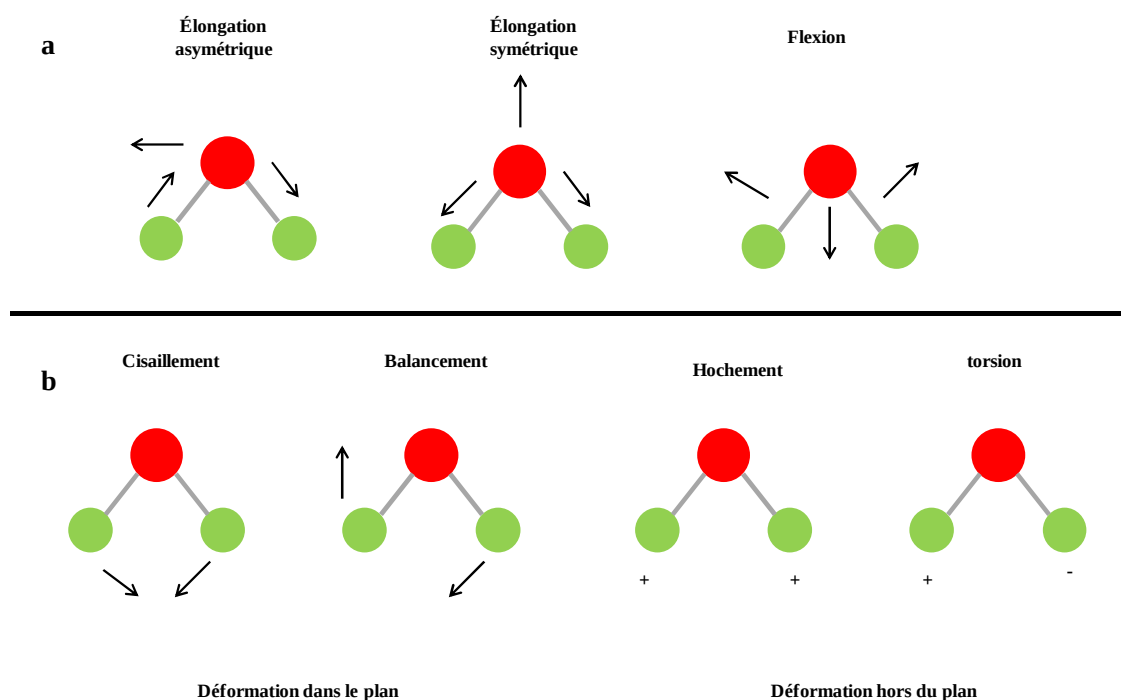


Figure 12 Représentation de vibrations des liaisons moléculaires (molécule triatomique). a. vibrations d'élongation. b. vibrations de déformation

2.1.1.1. Mesure de l'absorption lumineuse

Quand un rayonnement arrive sur un échantillon, il peut être soit absorbé, transmis ou réfléchi. Selon la loi de conservation d'énergie :

$$I_0 = I_a + I_r + I_t \quad (7)$$

Où I_0 l'intensité reçue par l'échantillon, I_a l'intensité absorbée, I_r l'intensité réfléchie, I_t l'intensité transmise.

La spectroscopie analytique consiste à mesurer la quantité d'énergie absorbée à chaque fréquence. Elle permet de déterminer les concentrations en différentes substances d'un mélange.

Dans le MIR les mesures sont effectuées en transmission, réflexion classique pour des échantillons peu absorbants et en réflexion totale atténuée(ATR).

✓ Mesures en transmission et en réflexion

Lors de mesures en transmission la source lumineuse et le détecteur de la lumière transmise se situent de part et d'autre de l'échantillon à analyser Figure 13b.

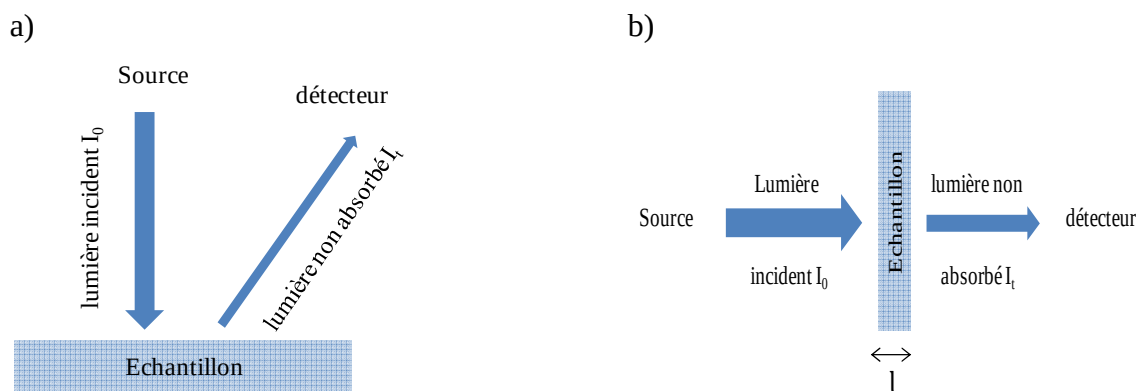


Figure 13 principe de mesure a) en réflexion b) en transmission

La loi de Beer-Lambert permet de relier l'intensité transmise et l'intensité incidente :

$$\log \frac{I_0}{I_t} = \varepsilon \quad (8)$$

Où l est le trajet optique dans l'échantillon, ε l'absorptivité molaire en $\text{mol}^{-1} \cdot \text{cm}^{-1}$, c'est la concentration en mol^{-1} du constituant considéré de l'échantillon.

L'absorbance : $\log \frac{I_0}{I_t}$

La transmittance : $\log \frac{I_t}{I_0}$

La source lumineuse et le détecteur de lumière transmise se situent du même côté de l'échantillon à analyser dans cette mesure Figure 13a. La loi de Beer-Lambert ne s'applique pas rigoureusement dans ce cas. En effet, le trajet optique l dans l'échantillon n'est pas connu avec précision et n'est pas constant.

✓ Réflexion totale atténuée

Cette technique n'est utilisée qu'en moyen infrarouge. En passant d'un milieu d'indice de réfraction élevé à un milieu d'indice inférieur, un rayon lumineux peut subir une réflexion totale à l'interface si l'angle de l'incidence à l'interface est supérieur ou égal à un angle α critique., tel que :

$$\alpha = \frac{n_2}{n_1} \quad (9)$$

n_1 , n_2 étant respectivement les indices des milieux, et $n_1 > n_2$.

Le rayonnement pénètre faiblement dans le milieu d'indice le plus faible et la profondeur de pénétration est donnée par la formule :

$$d_p = \frac{\lambda/n_1}{2\pi(\sin^2\theta - (n_2/n_1)^2)^{1/2}} \quad (10)$$

λ la longueur d'onde du rayon incident, n_1 l'indice du prisme, n_2 l'indice de l'échantillon, θ l'angle d'indice, n_2/n_1 pour avoir une réflexion totale.

Le principe de l'ATR repose sur la présence d'une onde évanescente dont l'intensité est maximale à l'interface d'un prisme d'indice de réfraction élevé et d'un échantillon d'indice plus faible.

Quand on place un échantillon sur un prisme, une partie du rayonnement incident est donc absorbée sélectivement Figure 14. La technique de réflexion totale atténuée ne permet que l'étude de la surface de l'échantillon en contact avec le cristal. Il sera donc nécessaire d'avoir une surface de contact maximale entre l'échantillon et le cristal.

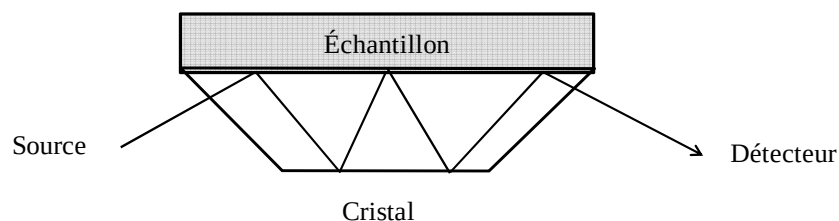


Figure 14 Principe de mesure en réflexion totale atténuée

Afin d'éviter les inconvénients des spectromètres dispersifs (temps d'analyse, manque de sensibilité, fragilité mécanique), nous avons utilisés dans cette étude un interféromètre à transformée de Fourier ou FT-IR.

Cette technique permet de mesurer toutes les fréquences *simultanément*. Le principe de fonctionnement s'appuie sur les cinq organes suivants (Figure 15) :

1. Une source de lumière.
2. Un procédé permettant de générer les interférences : l'interféromètre.
3. Une cellule échantillon (en mode réflexion ou transmission)
4. Un détecteur.
5. un convertisseur analogique-numérique

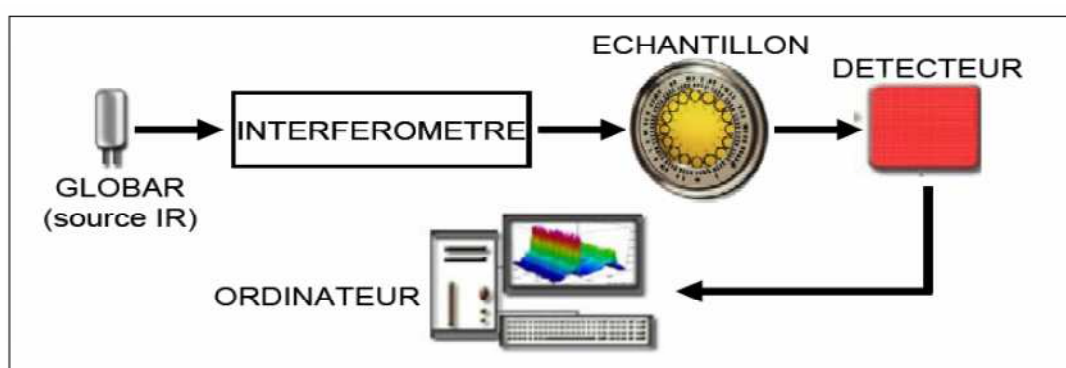


Figure 15 Schéma de principe d'un spectromètre FT-IR

L'intégration de l'interférométrie à la spectroscopie IR a été rendue possible par le développement d'outils de traitement du signal tels que la transformée de Fourier rapide. Elle a permis la résolution en temps réel de l'interférogramme, par le progrès des lasers améliorant la précision en fréquence de la mesure et par le couplage spectromètre-microordinateur.

2.2. Spécificités du spectre moyen infrarouge (MIR)

La spectroscopie MIR est une méthode physico-chimique basée sur la mesure la vibration d'une molécule soumise à un rayonnement IR dans la gamme de longueurs d'onde 4000 à 400 cm^{-1} . Toutes les molécules possèdent des fréquences caractéristiques d'absorbance et des vibrations moléculaires primaires dans cette gamme. Un exemple de spectre enregistré en réflexion totale atténuée entre 1500 et 700 cm^{-1} est présenté dans Figure 16 :

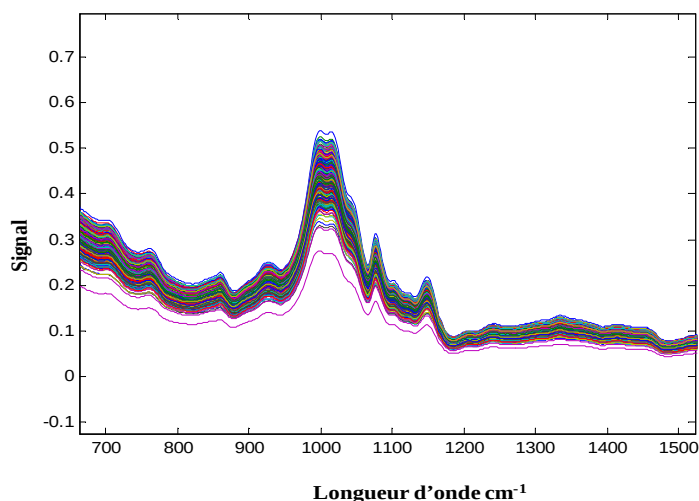


Figure 16 Exemple de spectre FT-MIR de l'orge dans la gamme 1500 et 700 cm^{-1}

Le spectre moyen infrarouge peut se décomposer en trois zones :

- Entre 4000 et 2000 cm^{-1} , les bandes d'absorption sont peu nombreuses et concernent les vibrations des groupements O-H, N-H, C-H.
- Entre 2000 et 1800 cm^{-1} , il n'y a pas pratiquement de bande d'absorption.
- Entre 1800 et 400, de nombreuses bandes d'absorption. Elles sont dues aux vibrations fondamentales de liaisons C-X. la zone 1200-800 cm^{-1} , est caractéristique des vibrations des liaisons C-C, C-O et est nommée "région d'empreinte digitale". En-deçà de 800 cm^{-1} , les oscillations observées sont dues essentiellement à des interférences engendrées par la composition du cristal.

2.2.1. Déformation des spectres de réflexion

Plusieurs facteurs influencent la qualité des spectres acquis :

- La variation de trajet optique.
- La variation d'absorbance liée à la variation de taille des particules.
- La densité de la poudre.
- L'épaisseur des cellules.
- Le taux d'humidité de l'appareil.
- L'appareil elle-même et sa température.
- Des effets dus au non linéarité du détecteur.

Pendant l'acquisition, il est nécessaire d'en limiter les effets qui perturbent le signal et masquent ou rendent l'information illisible. Les conséquences les plus fréquentes de ces perturbations sont : la variation de ligne de base, le changement de la pente (i.e. une dérive du signal) et la curvilinearité. Les spectres contiennent alors non seulement des informations chimiques sur l'échantillon à analyser mais aussi des informations physiques concernant les conditions d'acquisition du spectre. Dans la plupart des cas, il est donc important de procéder à des prétraitements sur ces spectres (correction de ligne de base, dérivées et autres traitement) pour atténuer les effets non désirés et extraire les informations chimiques importantes.

2.2.2. Prétraitement des signaux MIR

Un traitement adapté peut réduire, éliminer ou standardiser l'impact des sources de variations spectrales sans affecter l'information spectroscopique nécessaire à la réalisation du modèle de calibration. Les prétraitements de spectre sont nombreux, dans ce mémoire seuls les prétraitements utilisés sont présentés.

2.2.2.1. Centrage sur les Variables CV (Figure 17a)

Ce prétraitement est utilisé pour rendre les données indépendantes des unités de mesure. Il consiste à retrancher au signal brut, le signal moyen :

$$x_{ij} = m - \bar{m} \quad (11)$$

m est la valeur réelle du signal et $\bar{m}$ la valeur moyenne du spectre.

2.2.2.2. Normalisation Standard Normale Variante SNV (Figure 17b)

La normalisation SNV est utilisé pour compenser des variations en intensité de signal durant la collection des spectres, ces variations sont causés par la dispersion de la lumière et la taille des particules (Barnes *et al.* 1989, Candolfi *et al.* 1999). Le nouveau spectre X_{ij} est calculé de terme général :

$$x_{ij} = \frac{m_{ij} - \bar{m}_i}{\sqrt{\frac{\sum (m_{ij} - \bar{m}_i)^2}{p-1}}} \quad (12)$$

Où x_{ij} est le spectre normalisé à partir de spectre original m_{ij} , $\bar{m}_i$ est le moyen du $i^{\text{ème}}$ spectre et p est le nombre des variables dans le spectre.

2.2.2.3. Dérivées première et seconde (Figure 17c, 18d)

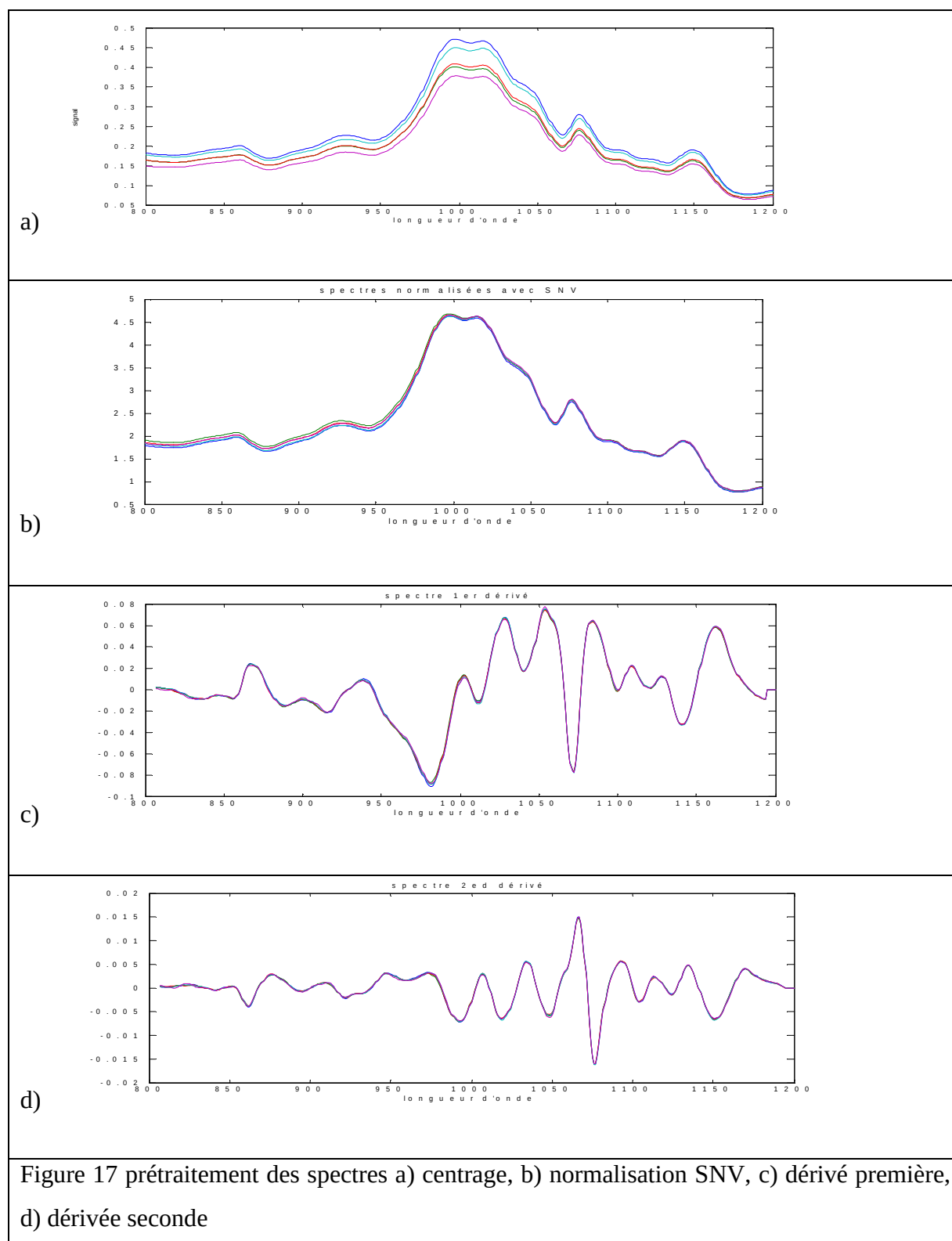
L'utilisation des dérivées permet d'éliminer la superposition des pics, de corriger la ligne de base et d'augmenter la résolution de spectre (Candofli *et al.* 1999).

Les dérivations ne peuvent être réalisées sur les spectres bruts. Il est indispensable de lisser les spectres afin de diminuer efficacement le rapport signal/bruit dans le spectre corrigé. La méthodologie chimiométrique la plus utilisée est celle proposée par Savitzky et Golay. Il s'agit d'une méthode de calcul de la moyenne de la fenêtre mobile. Le principe de la méthode est que, pour de petits intervalles de longueurs d'onde, les données peuvent être montées par un polynôme de degré suffisant, et que les valeurs ajustées sont une meilleure estimation que celles mesurées, parce qu'une partie de bruit a été supprimé. Pour la première fenêtre la méthode prend les premières $2m + 1$ points et s'adapte, par les moindres carrés, le polynôme correspondant de l'ordre m . La valeur ajustée pour le point dans la position m remplace la valeur mesurée. Après cette opération, la fenêtre est décalée d'un point et le processus est répété jusqu'à ce que la dernière fenêtre soit atteinte. Au lieu de calculer le polynôme correspondant à chaque fois, si les données ont été obtenues à des intervalles également espacés, le procédé utilise des coefficients tabulés de telle manière que la valeur ajustée pour le point central dans la fenêtre est calculée en tant que:

$$x_j^* = \frac{\sum_{k=-m}^m c_k x_{j+k}}{2m+1} \quad (13)$$

Où x_j^* la valeur calculé pour le point central de la fenêtre, x_{j+k} sont les $2m+1$ valeurs originaux dans la fenêtre, c_k les coefficients de Savitzky-Golay pour le $k^{\text{ème}}$ value de filtre, $2m+1$ la taille de fenêtre de lissage. La taille de la fenêtre et le degré du polynôme doit être définis ils sont [3,13] pour le lissage dans cette étude.

Dans la première partie de ce chapitre la discussion est faite sur le matériel et les méthodes pour la matière végétale. La seconde partie présente le matériel et les méthodes d'analyse et de régression. D'abord la régression multilinéaire principe et méthodologie sont présentées, ensuite la méthode d'analyse par la composantes principales est discutée, enfin la régression au sens des moindres carrés partiels est détaillée.



Matériels et méthodes Analyses et modélisation des données (MLR, ACP, PLS)

1. Méthodes d'analyse des données et de régression

Dans ce mémoire deux types d'approche de modélisation sont utilisées : la régression multilinéaire (MLR) et l'analyse multivariée (PCA/PLS)

La maîtrise des procédés agroalimentaires et biotechnologiques reste très difficile et imparfaitement réalisée. Cela peut être expliqué en grande partie par la nature et à la complexité de la matière biologique; où le contrôle des aptitudes et des cinétiques caractérisant ces procédés est difficile. Le besoin de maîtriser ces procédés implique cependant de connaître leurs réactions face aux changements de valeur des paramètres qui les influencent. Pour appréhender de tels phénomènes complexes, la mise au point d'un modèle est capable d'imiter fidèlement les comportements observés est généralement la meilleure approche. Pour mettre au point ces modèles, la réalisation d'expériences est indispensable pour disposer d'*observations* utilisées pour établir les capacités imitatrices ou prédictives du modèle. La planification expérimentale, est une approche qui permet de réaliser rapidement et efficacement cette phase d'acquisition d'information (Tinsson, 2010). Elle est associée à une phase de modélisation rapide et peut conduire rapidement à une phase d'optimisation simple. Elle implique cependant des modèles de tendances simples (polynômes d'ordre un ou deux).

Compte tenu à la grande quantité de données manipulées dans cette étude, l'utilisation d'un modèle est essentielle pour représenter les relations entre les facteurs et les mesures. Une méthode des plans d'expériences a été adoptée afin d'optimiser notre démarche.

Cette partie est consacrée à décrire dans un premier temps, les étapes et méthodes de planification expérimentale, puis le choix de modèle et son identification pour chaque type de régression.

1.1. Analyses par régression multilinéaire MLR

Dans la phase initiale de notre étude, l'objectif consistait à établir les conditions opératoires de trempage pour atteindre des propriétés acceptables de malt, en conditions de maltage à faible humidité, tout en intégrant les variations de qualité de la matière première : l'orge. Pour atteindre cet objectif, Des modèles polynomiaux, simples et pratiques ont été utilisés. : Peu de données étaient ainsi nécessaire à la mise au point des modèles. Cette étape de modélisation a été suivie par une phase d'optimisation.

1.1.1. La modélisation par régression linéaire multiple MLR

1.1.1.1. Modèles polynômiaux et leur présentation matricielle

Dans ce projet, ce type de modèle est utilisé étant donné qu'une connaissance assez limitée de ce que passe dans les grains d'orge pendant le maltage. Un modèle de connaissance reste un caractère particulièrement compliqué en raison de la complexité des phénomènes et des échelles spatiales (interne au grain, lit de grain) et temporelles (réaction enzymatiques rapide, maltage en 5 jours) mises en jeu.

Les modèles polynomiaux utilisés sont de la forme :

$$y = a_0 + \sum_{i=1}^{NF} a_i x_i + \sum_{i=1}^{NF} a_{ii} x_i^2 + \sum_{i < j} \sum_{j=1}^{NF} a_{ij} x_i x_j + \varepsilon \quad (14)$$

avec y , la réponse mesurée (sortie),

$a_0, a_1, \dots, a_k$, les paramètres du modèle à déterminer,

NF , représente le nombre de facteurs,

$x_0, x_1, \dots, x_{nf}$, les facteurs étudiés (entrées), et,

ε , l'erreur entre la valeur mesurée et celle estimée par le modèle.

Pour établir ce modèle, on réalisera N expériences conduisant ainsi à un système de N équations à P inconnues $NP = (NF+1)*(NF+2)/2$ pour un modèle d'ordre 2 complet. Cela conduit à un système matriciel simple qui est représenté schématiquement sur la Figure 18.

$$Y_M = AX + E \quad (15)$$

$Y_M (\in \mathbb{R}^N)$ le vecteur des réponses mesurées, $AX (\in \mathbb{R}^N)$ les valeurs modélisées, $A (\in \mathbb{R}^P)$ le vecteur des coefficients estimés par la régression, E le vecteur des erreurs de mesure, $X (\in \mathbb{R}^{N \times P})$ la matrice d'information

La valeur cherchée du vecteur A doit minimiser $E^T E$:

$$E = Y - AX \quad (16)$$

La structure du modèle et le plan d'expérience sont représentés sous la forme d'une matrice X composée de N lignes (égales au nombre d'expérience réalisées) et de NP colonnes (égales au nombre de paramètres du modèle) et un vecteur Y qui contient les valeurs des réponses mesurées pour chaque expérience du plan Figure 18.

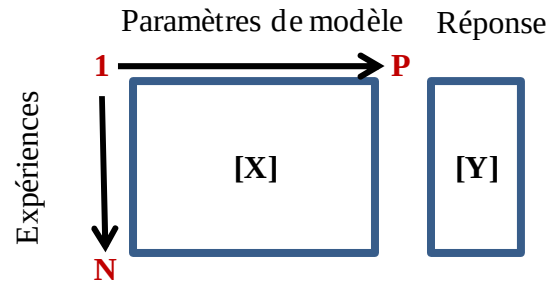


Figure 18 Présentation de la matrice d'expérience

Le meilleur vecteur A , au sens des moindres carrés, pour minimiser l'écart entre les valeurs calculées et les valeurs observées sont alors donné par l'expression:

$$\hat{A} = (X^T X)^{-1} X^T \quad (17)$$

$\hat{A}$ est la valeur estimée des coefficients du modèle. C'est la meilleure estimation des coefficients d'un modèle d'où on pourrait calculer les coefficients du modèle à partir des résultats expérimentaux. On montre ainsi un premier lien entre les choix des expériences, les observations et les coefficients du modèle.

Cependant, chaque facteur x ou chaque mesure y peut prendre des valeurs très différentes les unes des autres. Par exemple, la température varie entre 14 et 22 C tandis que les teneurs en β -glucanes varient entre 3 et 6 mg/L. Pour pouvoir comparer les différents effets, il est conseillé de les normaliser.

Dans ce travail, les facteurs et les mesures sont normés dans l'intervalle $[-1, +1]$ en utilisant la relation :

$$z_N = \frac{2z - (z_{max} + z_{min})}{(z_{max} - z_{min})} \quad (18)$$

Avec z_{min} La valeur minimum de z , z_{max} la valeur maximum de z , z la valeur du facteur, z_N la valeur normalisé de z

Les bornes pour chaque facteur orge et pour chaque mesure malt sont indiquées dans le Tableau 5.

Tableau 5 les intervalles des facteurs (procès et orges) et les intervalles des mesures en fin de trempé et du malt utilisés pour la normalisation des données dans les modèles

	Facteurs orges								Mesures fin de trempé et Malt							
	P	T	IG	BG	Ara	WR1	Ar1	WR2	HFT	H _{j+1}	HF	Extr	Visco	ProSol	Pdias	Attlim
Max	11.5	22	10	6	10	24	24	24	50	50	50	90	3	8	500	90
Min	9.5	14	1	3	6	4	4	4	30	30	30	75	1	2	100	60

Pour vérifier si les modèles sont significatifs et pour examiner la qualité des coefficients estimés, plusieurs tests de la régression multilinéaire peuvent être utilisés.

Une analyse complémentaire sur la variance, met également en évidence une *propagation* de la variance expérimentale $\sigma_{Y_M}^2 = Var(Y_M)$ et la variance des coefficients du modèle : $\sigma_A^2 = Var(A)$ à travers une matrice liée à X :

$$\sigma_A^2 = (X^T X)^{-1} \sigma_{Y_M}^2 \quad (19)$$

Non seulement la valeur des coefficients a_i mais également leur variance sont conditionnées par le choix des expériences à réaliser. Plusieurs démarches, regroupées sous l'intitulé planification expérimentale, visent à une minimisation de cette *propagation*.

1.1.1.2. Le plan d'expérience

Pour la régression MLR, la méthode des plans d'expériences cherche à déterminer et à établir les liens existant entre 2 types de variables : *La réponse (Y)*: les qualités de malt (friabilité, extrait, viscosité..); *les facteurs (X)*: les facteurs qualitatifs de l'orge pour cette étude (les teneurs en protéines, β -glucanes, arabinoxylanes...) qui sont censées induire des variations de la qualité de malt. Plus précisément, elle vise à identifier la relation liant la qualité du malt avec les facteurs *orge*. Une démarche expérimentale est établie pour répondre à nos questions. Notre plan d'expériences consisté à la réalisation d'une série de $N=40$ expériences, qui sont toutes: déterminées; réalisables ; indépendantes entre elles. La valeur de la réponse pour ces N expériences est ensuite déterminée expérimentalement avec précision. Un modèle polynomial en est alors déduit.

La planification des expériences a consisté à : i) choisir des échantillons d'orge pour lesquels les facteurs particuliers (des teneurs très variables en protéines, β -glucanes et arabinoxylanes, et à imposer des diagrammes de trempes diversifiés), ii) mesurer les variations induites de la qualité du malt, iii) déduire les relations entre facteurs orge et qualité du malt. L'objectif est de minimiser le nombre d'expériences sans sacrifier la qualité des résultats.

Il existe plusieurs types de plan d'expériences : les plans d'expériences (factoriels, orthogonaux, screening). Dans notre étude un plan de type screening a été adopté.

1.1.1.1. Le criblage (screening)

Cette méthode est utilisée quand les connaissances sur le procédé sont insuffisantes. La technique du screening permet de déterminer, parmi un ensemble initial de facteurs, les éléments qui influencent le procédé étudié. Il s'agit donc d'un procédé de sélection ou de criblage (screening) (Benoist, 1994, Schimmerling, 1998). Il est facile de déterminer expérimentalement le poids de chaque facteur dans le groupe choisi, de les classer et donc de pouvoir choisir pour une étude ultérieure plus fine, les facteurs qui sont les plus influents. Ces facteurs sont étudiés en deuxième étapes d'une façon plus précise en tenant compte les interactions possibles entre les différents facteurs. La technique du screening permet d'avancer dans la compréhension de tout système. Elle donne ainsi la possibilité de ne retenir que les grandeurs dignes d'intérêt.

Pour appliquer cette méthode, il est indispensable de cerner au mieux les mesures susceptibles d'agir sur le fonctionnement du procédé, et les causes possibles de variations observés sur la réponse. Les informations sur le phénomène étudié sont collectées à partir des essais qui sont les fruits d'une réalisation physique d'une configuration des paramètres d'entrée (Facteurs) de la boîte noire, pour lesquelles la valeur de la réponse est mesurée qui évoluent avec le temps (Walter et Pronzato 1994).

Dans cette étude, un plan d'expériences de type screening sera utilisé. Cependant, les conditions de micro maltage à l'IFBM permettent de réaliser simultanément 16 expériences complémentaires sans augmenter la durée expérimentale. Les expériences complémentaires sont extraites d'un plan de screening correspondant à un modèle avec interactions. La sélection des expériences complémentaires est basée sur l'équilibre du plan et sur la disponibilité de compartiments de micro maltage. Les 40 conditions opératoires du plan d'expériences résultant et les valeurs des mesures obtenues lors de sa mise en œuvre en micro malterie sont données dans les ANNEXE 1, ANNEXE 2 .

1.1.1.3. Les tests pour la significativité des modèles

La qualité des modèle MLR peut être évaluée à travers une série d'indicateur parmi lesquels on trouve notamment :

1. Le coefficient de détermination de la régression R^2 .
2. Le caractère significatif des coefficients du modèle par un Test de Student
3. La robustesse des coefficients du modèle par la technique du *Bootstrap*.

1. Coefficient de détermination R^2 :

Cet indicateur est utilisé pour étudier la corrélation entre les valeurs expérimentales et les valeurs modélisées. Il traduit la capacité d'imitation des tendances observées par le modèle. Dans un repère données estimées vs. données observées, on recherche la droite qui passe

au plus près de tous les points et dont le coefficient directeur est unitaire. Dans une régression, on souhaite minimiser un critère de type moindre carré

$$J = \sum_{j=1}^N (y_i^{obs} - y_i^{est})^2 \quad (20)$$

Le critère exprime le rapport entre la variance expliquée par la modèle et la variance totale qui doit être proche de 1.

$$R^2 = \frac{\text{la variance estimée par le modèle}}{\text{la variance totale}} \quad (21)$$

2. Test de Student :

Ce test paramétrique compare la moyenne observée d'un échantillon statistique (ici, celle des coefficients du modèle) à une valeur fixée. Dans le cas présent, on vérifie si l'intervalle des valeurs de coefficients du modèle contient la valeur 0.

La variance des valeurs des coefficients du modèle peut être estimée à partir de la relation :

$$s_{ai}^2 = (X^T X)^{-1} s_y^2 \quad (22)$$

Suivant la loi de Student, on fixe un seuil de signification α , ($\alpha = 5\%$), le degré de liberté v (nombre d'expérience-nombre de paramètres) permet de définir un $t_{v,\alpha}$ à partir des tables de Student. On peut alors tester la signification (nullité) d'un paramètre en construisant son intervalle de confiance :

$$a_i = a_i \pm t_{v,\alpha} s_{ai} \quad a_i = a_i \mp t_{v,\alpha} s_{ai} \quad (23)$$

La variance expérimentale s_y^2 est estimée soit à partir de répétition d'expériences, soit par approximation. Dans notre étude, la variance du résidu entre valeurs expérimentales et valeurs prédites est utilisée comme approximation

La variance entre les valeurs des sorties expérimentales et modélisés

$$s_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i^{obs} - \bar{y})^2 \quad (24)$$

3- Technique du Bootstrap :

Cette technique numérique poursuit deux buts : vérifier la robustesse d'un modèle, renseigner sur la stabilité des coefficients de ce modèle. Le principe de ce test est de reconstruire un plan d'expérience de la même taille en choisissant, parmi les N expériences réalisées, $N-1$ expériences au hasard, avec remise en jeu. On recalcule les coefficients du modèle à partir de cette nouvelle matrice d'information. Cette étape permet de justifier les stabilités des coefficients calculées et leur distribution et fournit, en plus une idée sur la robustesse du modèle

La régression multilinéaire ainsi que les tests ont été programmés sous Matlab, en utilisant un ordinateur (Pentium IV, 2,8 GHz, 1 Go Ram).

1.1.2. Méthode d'optimisation

Lorsque les modèles des différents indicateurs de qualité du malt sont établis, un outil d'optimisation est nécessaire pour déterminer les conditions opératoires qui conduisent simultanément à une humidité fin de trempe HFT faible (37-38%) et à une qualité acceptable du malt.

L'outil d'optimisation utilisé dans cette étude est un algorithme génético-évolutionnaire. Il est appelé ainsi car il met en œuvre les techniques issues à la fois des algorithmes génétiques (croisement des individus par des opérateurs "entier") et des algorithmes évolutionnaires (croisement des individus par des opérateurs "réels"). Les principales caractéristiques de cet algorithme sont présentées dans cette section.

1.1.2.1. L'algorithme génético-évolutionnaire

Il s'agit d'affecter l'évolution d'une population d'individus divers (généralement tirée aléatoirement au départ) à laquelle on applique différents opérateurs (recombinaisons, mutations...) et que l'on soumet à une sélection, à chaque génération. Si la sélection s'opère à partir de la fonction d'adaptation, alors la population tend à s'améliorer (Bäck, 1996) la Figure 19 montre un simple schéma avec les étapes essentielles pour l'évaluation d'un algorithme génétique en général

Un tel algorithme ne nécessite aucune connaissance du problème : on peut représenter celui-ci par une boîte noire comportant des entrées (les variables) et des sorties (les fonctions objectif). L'algorithme ne fait que manipuler les entrées, lire les sorties, manipuler à nouveau les entrées de façon à améliorer les sorties, etc. (Whitley, 1993).

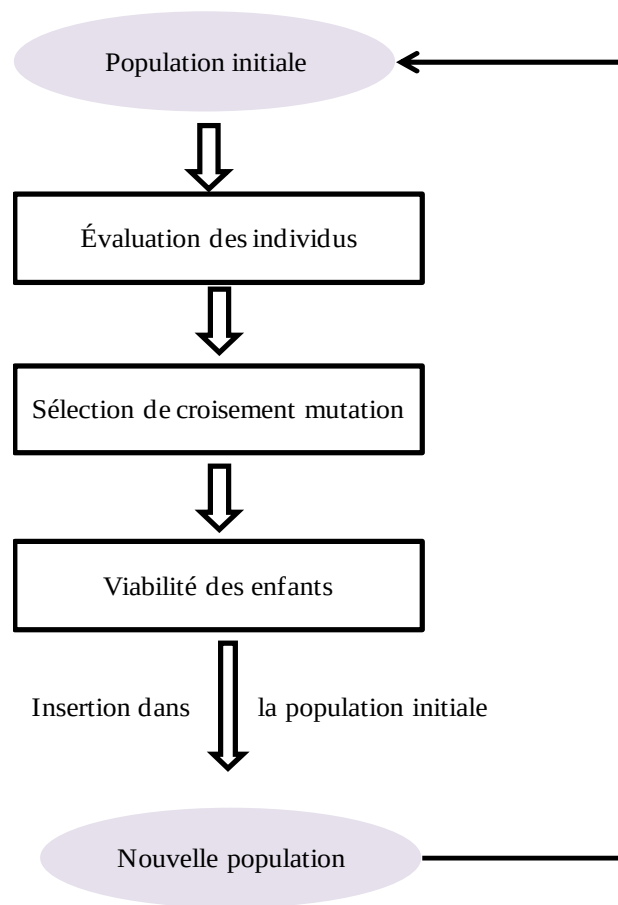


Figure 19 Schéma montrant les étapes d'évaluation d'un algorithme génétique

L'atout principal de ces algorithmes est leur capacité à travailler directement sur les fonctions sans nécessité de dérivation compliquée. C'est un processus itératif d'optimisation qui manipule une population d'individus plutôt qu'un point (Massebeuf, 2000). Leur évaluation se fait à partir de la connaissance du milieu.

1.1.2.2. Principe général

Pour une fonction multi variable, les variables sont représentées par vecteur désigné par son code génétique (chromosome(s)). Il existe des algorithmes haploïdes et d'algorithmes diploïdes. Ils se distinguent par leur codage et par leurs propriétés de convergence (temps de calcul et efficacité à trouver des optima globaux) (Fonteix *et al*, 1995)

- algorithme haploïde : chaque individu y représenté par un seul chromosome qui correspond exactement à son phénotype. La formule utilisée pour faire un croisement entre deux individus haploïdes est $X = \alpha X_1 + (1 - \alpha) X_2$ où α est un nombre aléatoire réel.
- algorithme diploïde : on a besoin de trois vecteurs pour codés les individus diploïdes deux chromosomes chacun entre eux a tous les paramètres à optimiser un vecteur de dominance qui est responsable de déterminer pour chaque gène quelle version va s'exprimer.

1.1.2.3. Initialiser la population

La population initiale correspond à l'ensemble des individus sélectionnés au lancement de l'algorithme. Et la taille de cette population affecte les performances et l'efficacité de l'A.G. Selon (Viennet ,1997) Dans l'algorithme utilisé, la population initiale est définie aléatoirement. Plusieurs études ont montrés l'influence de la composition de la population initiale sur la vitesse de convergence et la capacité à éviter les optimums locaux à l'optimum global (Bicking ,1994).

1.1.2.4. Evaluation de la population

C'est l'étape où les fonctions à optimiser sont évaluées pour chacun des individus d'une génération. Les valeurs obtenues permettent d'évaluer les individus les uns par rapport aux autres afin de les classer des meilleurs aux plus mauvais. Si un enfant est moins bon que le plus mauvais des parents de la génération précédente, il n'est pas conservé : c'est la mortalité infantile.

1.1.2.5. Génération des nouveaux individus

Il se trouve deux moyens principaux pour créer de nouveaux individus : d'un côté Le croisement qui est le moyen principal et privilégié pour générer un nouvel individu, permet d'élargir la population. Il s'effectue en choisissant aléatoirement deux parents de la population et donnent naissance à un nouvel individu (enfant).

Il existe 3 méthodes de génération de nouveaux individus : le croisement hétérozygotes, le croisement homozygotes et la mutation.

Dans le croisement hétérozygote, chaque parent donne aléatoirement la composition d'un de ses deux chromosomes à l'enfant. Selon les dominances des allèles, le vecteur de dominance de l'enfant suit les règles du Tableau 6.

Dans le croisement homozygote, les deux chromosomes du génotype des nouveaux enfants sont identiques. Le phénotype est calculé sur un principe identique au croisement de l'individu haploïde. Il est bon pour étendre les caractéristiques des meilleurs parents. Le croisement homozygote reste le meilleur pour une diversité plus grande.

Tableau 6 le choix des dominances lors de la génération d'un nouvel individu

		Domainance de l'allèle sélectionné pour le parent 1	
		Récessif	Dominant
Domainance de l'allèle sélectionné pour le parent 2	Récessif	Domminace aléatoire	Dominance de l'allèle du chromosome 1
	Dominant	Dominance de l'allèle du chromosome 2	Fonction du meilleur parent

La mutation est une autre méthode pour donner des nouveaux individus où les chromosomes et la dominance sont faits aléatoirement. Elle aide à élargir le domaine de recherche et à créer de diversité dans la population.

Dans tous les cas, les nouveaux individus sont soumis à un test de viabilité, les enfants qui ne sont pas meilleurs que le dernier parent de la population initiale est éliminé. Les nouveaux individus viables sont ajoutés à la population pour atteindre sa taille initiale.

1.1.2.6. Critère d'arrêt

L'algorithme lancé ne peut pas arrêter tout seul, il a besoin d'un critère d'arrêt capable de l'arrêter quand il est arrivé à construire une population répondant aux attentes. Pour cela, une valeur très petite (précision) est définie. Elle constitue un critère d'arrêt qui est utilisé quand la reconstruction de la population est terminée. Si la différence de performance entre le meilleur et le pire individu ??? est inférieure à cette précision, l'algorithme s'arrête.

L'algorithme génétique présenté constitue la base de la recherche de solutions optimales parmi une population d'individus. Cet algorithme, conçu pour trouver une solution optimale: le premier individu de la population.

L'algorithme génétique utilisé dans ce projet (AG, LSGC. Hariri ,2003) a une caractéristique spéciale qui a été ajoutée pour adapter la convergence de cet algorithme vers un domaine de solution plutôt que vers une seule solution.

Un exemple du modèle HFT / OH a été choisi pour expliquer comment on a fait pour adapter la fonction objectif pour une optimisation par domaine

On définit une fonction objective classique permettant d'atteindre une spécification donnée :

$$F(x) = (HFT(x) - 37.5)^2 \quad (25)$$

F(X) représente la différence entre le taux de humidité calculé et celui ciblé. Cette fonction conduit à un seul optimum. L'algorithme cherche les conditions opératoires qui donne une valeur optimale de (HFT = 37.5)

Dans la Figure 20 la courbe en pointillé illustre l'allure de la fonction d'optimisation $F(x)$. On choisit ensuite une précision ε qui décrit l'écart toléré sur la valeur à atteindre. La fonction objective utilisée par l'algorithme est alors définie comme suit :

Pour le modèle utilisé dans cet exemple, on utilise $\varepsilon = (0.5)^2$ c'est-à-dire les valeurs acceptables de l'humidité fin de trempe sont (37.5 ± 0.5) . La courbe en trait continu illustre l'allure de la fonction $J(x)$. Elle montre que toutes les valeurs de HFT entre (37-38) correspondent à une même valeur nulle de $J(x)$ et sont acceptables.

$$\text{Si } F(x) = (HFT(x) - 37.5)^2 \leq \varepsilon$$

$$\text{Alors } J(x) = 0$$

$$\text{Sinon } J(x) = F(x) - \varepsilon \quad (26)$$

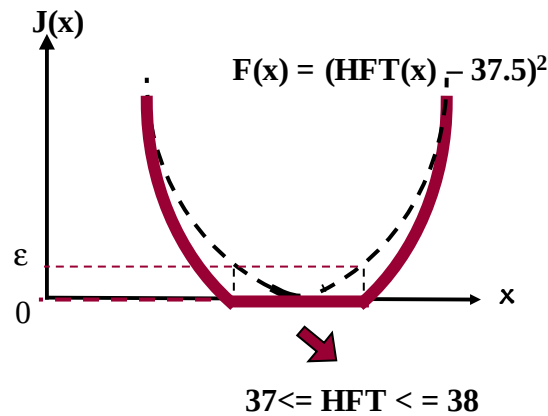


Figure 20 Fonction objectif pour l'optimisation par domaine

Ce travail est refait pour toutes les autres mesures, chacune avec sa valeur ciblée et sa précision ε propre. On peut résumer le procédé de modélisation et le procédé d'optimisation par un schéma dans la Figure 21

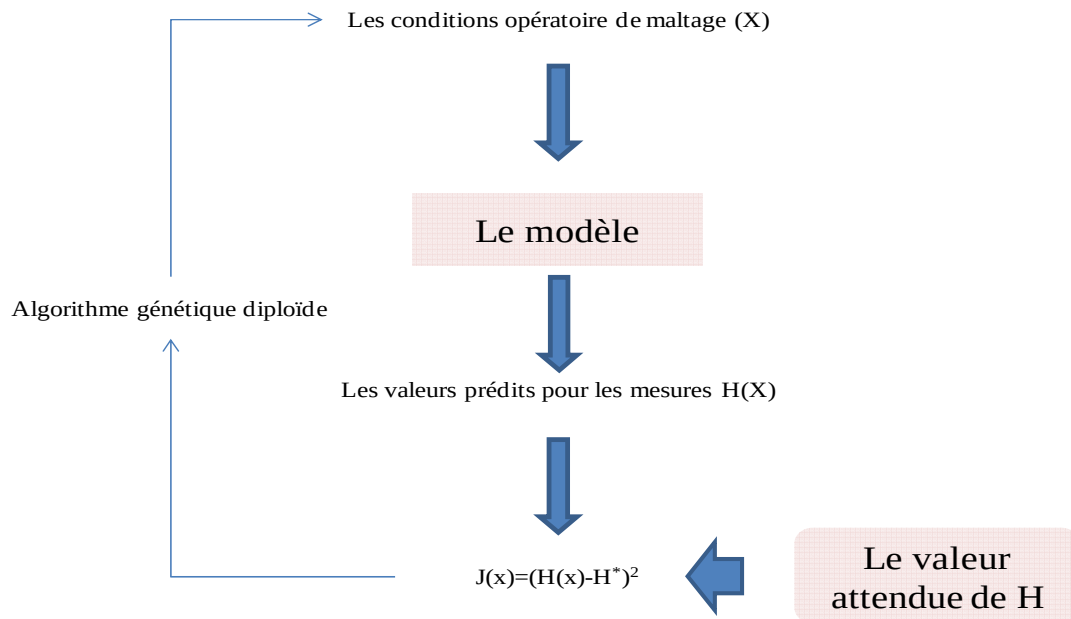


Figure 21 Optimisation du maltage en utilisant l'algorithme génétique

La valeur attendue

L'ensemble des outils d'optimisation ont été programmés sous Matlab (release 13) et les calculs sont réalisés sur un PC (Pentium IV, 2,8 GHz, 1 Go Ram)

2. Les analyses multivariées: Analyse en composants principales (ACP)

Les méthodes MLR se prêtent mal au traitement de données spectrales en raison des fortes corrélations intra-facteurs. Pour cette raison, les méthodes multivariées leur sont préférées pour le traitement numérique des spectres MIR.

Les analyses multivariées (factorielles) ont tous les mêmes objectifs : représenter des données multidimensionnelles à travers un nombre réduit de dimensions (idéalement 2 à 3) en concédant une perte *raisonnable* d'information (i.e. en conservant les informations les plus significatives).

Ces dimensions réduites, appelées composantes principales ou variables latentes, sont définies par combinaison linéaires des variables originales. L'information d'une centaine des données ou d'un spectre peut ainsi être condensée dans un petit nombre de composantes principales (Gottlieb *et al.* 2004 ; Munck and Moller, 2004). Ces composantes principales expliquent la majeure partie de la variance dans les données originales, et facilitent ainsi une interprétation efficace de propriétés des données par des analyses qualitatives telles que la discrimination ou la classification (Cordella *et al.* 2002 ; Downey, 1998). De plus, pour éviter d'avoir la même «information» dans plusieurs composantes principales, elles sont construites de façon à être orthogonales.

2.1. Présentation général de l'ACP

La théorie de l'ACP a été décrite par de nombreux auteurs (Lebart *et al.* 1977 ; Joliffe, 1986 ; Saporta *et al.* 1990).

L'analyse en composantes principales est une méthode descriptive qui s'applique à des tableaux de données quantitatives. Elle met en évidence des relations entre des variables et les observations pour lesquelles elles ont été mesurées. Les données enregistrées peuvent être présentées sous la forme d'un tableau ou d'une matrice dont les n lignes sont les différents échantillons/expériences et les p colonnes sont les variables ou conditions opératoires Figure 22.

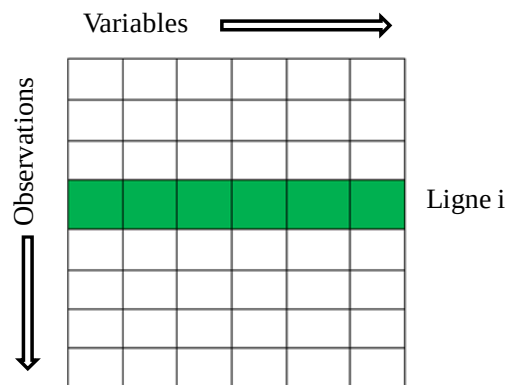


Figure 22 : description schématique d'une matrice de donnée pour une ACP

Une représentation géométrique de ces données expérimentales est possible dans un espace de dimension p où p désigne le nombre de variables p (*facteurs et/ou mesures*). *Projection*

Cette représentation géométrique dans le repère original ne met cependant pas en évidence les propriétés (corrélation) dans la population ainsi décrite. Un changement de repère permet parfois de mieux décrire l'ensemble des points. La difficulté réside dans la définition des nouveaux axes de ce repère réorienté (vecteurs propres) et dans le calcul des nouvelles coordonnées des points initiaux dans ce nouveau repère.

Dans un premier temps, chaque observation peut être projetée dans les plans de l'espace initial. L'ensemble de l'observation projetée forme un nuage de points de dimension réduite Figure 23 avec $p=3$.

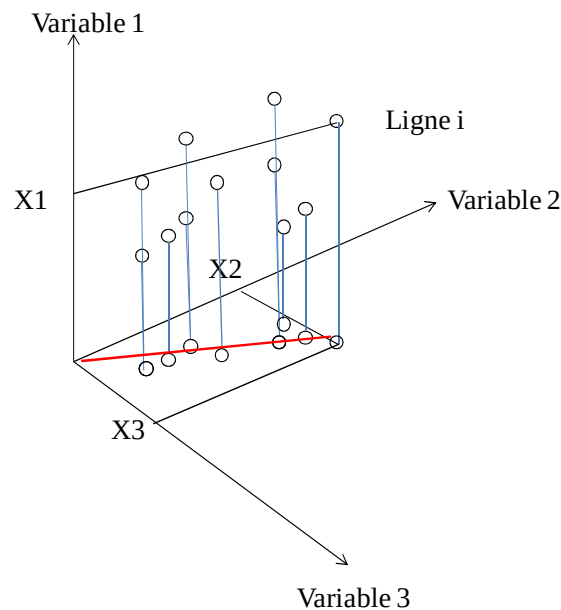


Figure 23 Présentation d'un nuage de points dans un espace de trois dimensions

L'observation graphique d'un tel nuage de points n'est possible que dans un espace à 2 ou 3 dimensions. Dans le cas des espaces de dimension supérieure, les méthodes multivariées permettent 1) de définir un nouvel espace permettant de mieux comprendre/observer les relations entre les individus et 2) de conserver uniquement les directions les plus informatives réduisant ainsi l'espace à quelques dimensions *pertinentes* en limitant les pertes d'information. D'un point de vue géométrique, l'ACP peut être considérée comme une méthode 1) de rotation des données suivie d'une méthode de 2) projection conduisant à une présentation plus efficace avec la dimension la plus petite possible (le minimum de composantes principales).

Pour définir la meilleure rotation des axes, on utilise comme direction principale celle qui présente le maximum de variance. On construit ainsi la première composante principale (CP1). On définit ensuite un deuxième axe, orthogonal au premier, qui décrit le maximum de variance dans la variance résiduelle. Cette direction est ainsi appelée 2^{ème} composante principale (CP2) (Figure 24). La procédure est répétée jusqu'à atteindre un niveau d'explication de la variance satisfaisant (80%, 90%; 95%, 99%). Si on utilise toutes les composantes principales, 100% de la variance est expliquée.

Cette rotation du nuage de points permet d'obtenir la représentation dans un nombre réduit de directions avec la plus grande fidélité possible.

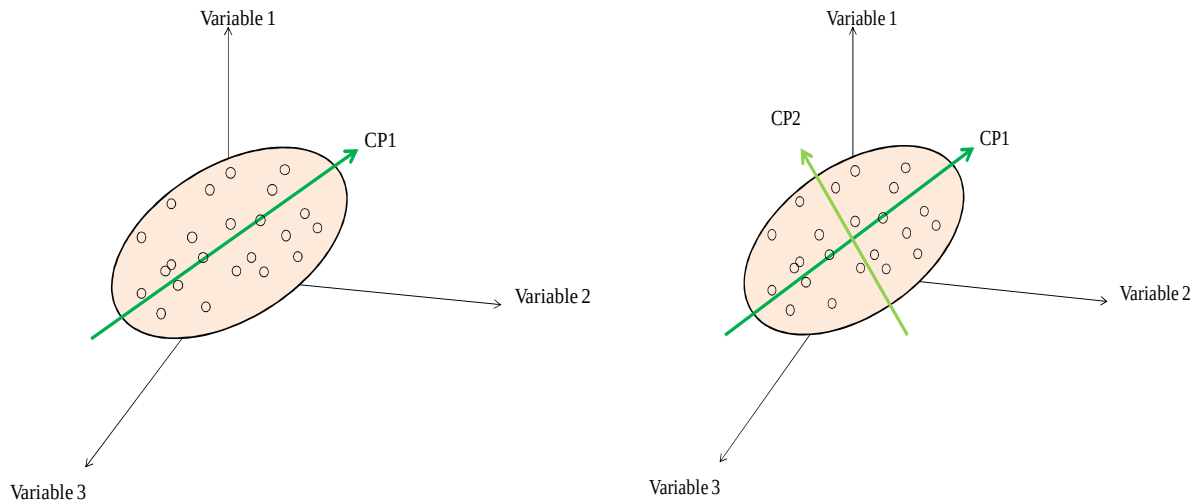


Figure 24 Illustration de la rotation des axes pour la mise en place d'un nouveau repère

Les coordonnées de tous les points du nuage dans le nouveau repère sont obtenues en projetant orthogonalement les points initiaux sur les nouveaux axes. Elles sont appelées coordonnées factorielles et elles constituent les valeurs des composants principales. Dans chaque couple d'axe, les points sont décrits par leurs coordonnées factorielles. On appelle ces représentations en 2D : cartes factorielles. Ces cartes ont un intérêt essentiel dans les analyses par composants principales. Elles permettent de projeter les individus sur des plans tel que on peut éventuellement détecter des répartitions structurées des objets, la formation de groupes ou la présence d'individus aberrants ce qui nous aide à la compréhension de la structure des données. Figure 25 illustre une ACP effectué dans cette étude sur des orges de printemps récolte 2010 dans quatre lieux différents en France : *Clermont-Ferrand*, *Issoudun*, *Bretenieres*, *Merignies*. L'ACP permet ainsi de distinguer les lieux et souligne, en même temps les points singuliers (deux points du lieu *Merignies* qui se trouvent avec les points du lieu *Bretenieres*).

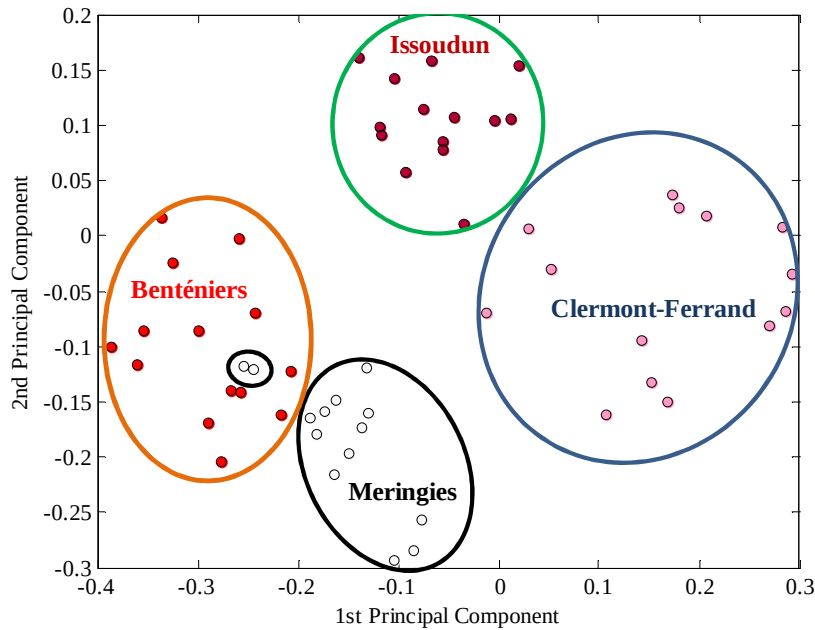


Figure 25 Carte factorielle d'un nuage de points

2.2. La mise en œuvre de l'ACP

La méthode ACP s'appuie sur le remplacement d'une matrice $\mathbf{X}$ de $\mathbf{n}$ lignes et $\mathbf{m}$ colonnes, par une nouvelle matrice $\mathbf{T}$ dérivée de $\mathbf{X}$ avec $\mathbf{n}$ lignes et un nombre de colonnes $\mathbf{k}$ très inférieur de $\mathbf{m}$. On peut écrire cette relation en terme matricielle :

$$\mathbf{T} = \mathbf{X} \mathbf{W} \quad (27)$$

où $\mathbf{W}$ est la matrice des coefficients qui définissent les combinaisons linéaires, avec $\mathbf{m}$ lignes et $\mathbf{k}$ colonnes, en d'autres termes, la matrice des vecteurs propres de $\mathbf{X}^T \mathbf{X}$ (*loadings*).

Les colonnes de la matrice $\mathbf{T}$ ont plusieurs particularités :

- 1 Ils sont des combinaisons linéaires des variables d'origine.
- 2 Ils sont orthogonaux entre eux.
- 3 Ils constituent les valeurs propres de $\mathbf{X}^T \mathbf{X}$.

L'ACP classe les valeurs propres (composantes principales) par ordre décroissant pour sortir les composantes les plus explicatives.

L'ACP permet donc de décrire l'essentiel de la variance sur les données $\mathbf{X}$ dans un nouvel espace de petite dimension (compression des données). De même temps l'ACP est un outil de classement de données (tri sélectif).

Dans l'ACP, on calcule la matrice carrée de covariance. La matrice initiale est toujours centrée et réduite, pour éviter les problèmes des unités et l'hétérogénéité des variances.

Les analyses ACP ont été appliqués d'abord sur tous les *données expérimentales* des organes récoltées de 2009, 2010, 2011. Ensuite les analyses en composantes principales sont

menés sur les *données spectrales* collectées sur les orges de 2009, 2010, 2011. Toutes les analyses ACP ont été programmées sous Matlab.

3. La régression en composantes principales (PCR)

La décomposition par ACP est la première étape de la plupart des méthodes linéaires multivariées basées sur des projections, telles que la PCR, PLS.

Pour la régression en PCR, on calcule $\hat{y}_i$ à partir de T calculé :

$$T = XW \quad (28)$$

$$y = Tb + e \quad (29)$$

et le vecteur des coefficients b de la régression est donné est

$$b = (T^t T)^{-1} T^t y \quad (30)$$

Le problème est de déterminer le nombre de composantes principales k à exploiter pour la régression. On utilise typiquement la méthode de validation croisée.

- On divise l'ensemble des données **X** et **y**, en **k** sous-ensembles
Dans notre étude, le sous-ensemble n'est constitué que d'un seul échantillon (technique du *Leave-One Out Cross Validation*)
- On construit un modèle avec (**k-1**) sous-ensembles.
- On prédit la valeur pour le sous-ensemble exclu de l'étape précédent et on calcule alors la racine de l'erreur quadratique moyenne de validation croisée (*Root Mean Square Error of Cross Validation*)

$$RMSECV_{k-1} = (y - Xb)^2 \quad (31)$$

- On répète ce calcul jusqu'à ce que tous les sous-ensembles soient exclus une fois.
- On calcule la somme de ces erreurs

$$RMSECV = \sum_k RMSECV_{k-1} \quad (32)$$

On choisit le nombre de composantes qui permet de minimiser la valeur de RMSECV.

En termes de prédiction, les composantes principales de PCR ne sont pas toutes nécessairement pertinentes. Elles le sont pour la description de la variance de la matrice X. En utilisant plus de composantes, cela peut également augmenter le RMSECV. Cela s'explique par le fait que ces composantes sont déterminées sans exploiter le vecteur des réponses y. Une meilleure approche consiste à construire les composantes en utilisant les informations de X et de y en même temps.

4. La régression en moindre carrée partielle (PLS)

La régression PLS est appliquée depuis quelques années dans plusieurs domaines en chimiométrie, dont la calibration multivariée, les relations structure-activité. Les variables prédictives (**X**) sont souvent des spectres mais peuvent aussi être des mesures bio-physico-chimiques, la plupart du temps avec une forte colinéarité entre variables. Les réponses (matrice **Y** ou vecteur **y**) peuvent être de nature variée, dont des concentrations des substances à prédire, mais aussi des propriétés physico-chimiques, des activités biologiques, des paramètres industrielles, des matrices spectrales d'un autre domaine. La régression PLS est une méthode utilisée couramment dans les analyses spectroscopiques.

Dans la régression PLS on calcule **T** en tenant en compte de **X** et de la variable à prédire **y**. La régression PLS effectue une double régression comme montre Figure 26.

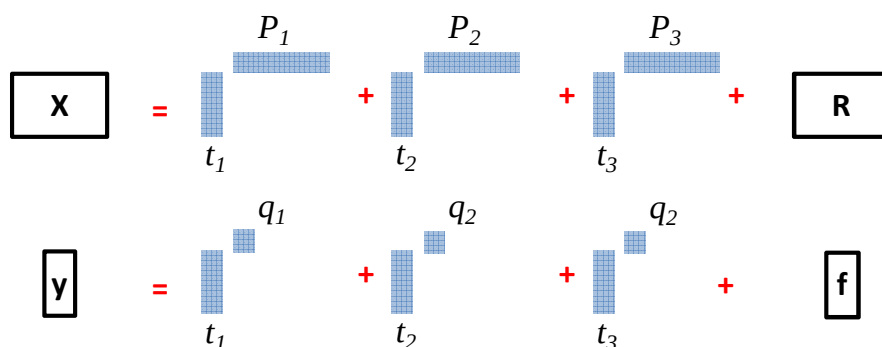


Figure 26 Présentation d'une régression PLS sur trois dimensions

A partir de la Figure 26 on comprend que les deux équations de régressions correspondantes à **X** et à **y** par les corrélations suivants :

$$X = TP + R \quad (33)$$

$$y = Tq + f \quad (34)$$

Dans le cas où il y a une seule réponse **y**, la méthode est appelée PLS1; C'est la méthode adaptée dans notre démarche scientifique.

Étape 1 : on calcule **t** :

$$t_{k1} = \sum_{j=1}^p X_{1j} W_{kj} \quad (35)$$

t_{k1} est la première colonne de la matrice **T** composée de lignes (scores) et des colonnes qui sont les **composantes principales**. **W** est la matrice des poids (**loadings**) qui s'appuie sur le calcul de la covariance entre **X** et **y** :

$$w = \frac{x^t y}{\|x^t y\|} \quad (36)$$

La matrice **T** maximise la variance entre **X** et **y**. la procédure consiste à remplir la matrice **T** colonne par colonne.

Étape deux : on réalise la régression :

$$y = t_1 q_1 + f \quad (37)$$

L'algorithme de SIMPLS pour la régression PLS1 est décrit dans ANNEXE 15. Pour les bases de données où plusieurs réponses y doivent être modélisées simultanément, la sortie de modèle est une matrice $\mathbf{Y}$. On utilise alors la méthode PLS2. Dans notre étude, nous avons essentiellement appliqué les méthodes MLR, ACP et PLS1 sur les mesures de qualité de l'orge et/ou du malt ainsi que sur les mesures de qualité de l'orge par spectroscopie MIR.

Références bibliographiques chapitre 3

- Anjum F.M., and Walker C.E., 1991. Review on the significance of starch and protein to wheat kernel hardness. *Journal of the Science of Food and Agriculture* 56, 1-13.
- Bäck T., *Evolutionary Algorithms in Theory and Practice*. Oxford: Oxford University Press, 1996
- Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and De-trending of near-infrared diffuse reflectance spectra. *Applied spectroscopy* 43, 722-777.
- Bensoit, D., Tourbier, Y., Germain, S., 1994. *Plans d'expériences : construction et analyse*.
- Candolfi, A., De Maesschalck, R., Jouan-Rimbaud, D., Haily, P.A., Massart, D.L., 1999. The influence of data preprocessing in the pattern recognition of excipients near-infrared spectra. *Journal of pharmaceutical and biomedical analysis* 21, 115-132.
- Cordella, C., Moussa, I., Martel, A.C., Sbirrazzuoli, N., Lizzani-Cuvelier, L., 2002. Recent developments in food characterization and adulteration detection: Technique-oriented perspectives, *J. Agri. Food Chemist.* 50, 1751-1764.
- Downey, D., 1998. Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics, *TrAC Trends in Analytical Chemistry* 17, 418-424
- Gottlieb, D.M., Schultz, J., Bruun, S.W., Jacobsen, S., Sondergaard, I., 2004. Multivariate approaches in plant science, *Photochemistry* 65, 1531-1548.
- Joliffe, I.T., *Principal component analysis*. Springer-Verlag. New-York. 1986.
- Munck, L., Moller, B., 2004. A new germinative classification model of barley for prediction of malt quality amplified by near infrared transmission spectroscopy calibration for vigour 'on line' both implemented by multivariate data analysis, *J. Inst. Brew.* 110, 3-17.
- Oomah, B.D., Reichert R.D., Youngs, C.G., 1981) A novel, multi-sample, tangential abrasive dehulling devise (TADD). *Cereal Chemistry* ; 58, 392-395.
- Saporta, G., *Probabilités, analyse des données et statistiques*. Technip. Paris. 1990.
- Schimmerling, P., Sisson, J.C., Zaidi, A., 1998. *Pratique des plans d'expériences*. Technique et documentation, Lavoisier, 61-438.
- Tinsson, W., 2010. *Plans d'expérience: constructions et analyses statistiques*, Mathématiques et Applications, Springer-Verlag Berlin Heidelberg.
- Lebart, L., Morineau, A., Tabard, N., *techniques de la description statistique, méthodes et logiciels pour l'analyse des grands tableaux*. Dunod. Paris. 1977.
- Walter, E., Pronzato, L. *Identification de modèles paramétriques : a partir de données expérimentales*. Masson, Paris, 1994.
- Whitley D., *A Genetic Algorithm Tutorial* ", Technical Report CS-93-103, Colorado State University, Department of Computer Science, 1993.

*Chapitre 4 étude et
modélisation du
maltage à faible
humidité*

Introduction

Ce chapitre présente les résultats complémentaires d'une démarche de modélisation et d'optimisation initiée dans une étude précédente dans le cadre d'un master en 2007-2008.

Ce chapitre contient trois parties :

- La première partie présente un rappel de la démarche de modélisation réalisée sur l'étape de trempe -pendant le maltage en vue d'obtenir un malt de qualité acceptable dans des conditions d'hydratation limitées des grains d'orge. Elle présente les modèles, les paramètres des modèles (facteurs opératoires (*process*) et les facteurs "matière première" (*orge*), mesures de la qualité du malt).
- La deuxième partie est consacrée à la démarche complémentaire d'optimisation des conditions du maltage en s'appuyant sur les modèles précédents.
- La dernière partie, à la lumière des difficultés rencontrées dans les deux parties précédentes, présente la possibilité d'intégrer de nouveaux facteurs dans le modèle pour caractériser l'orge et sa variabilité d'une année de récolte à l'autre est étudiée.

Les données utilisées dans cette partie sont présentées en ANNEXE 1 et ANNEXE 2. L'ANNEXE 1 contient tous les facteurs orge qui sont utilisés comme entrée de notre modèle de régression MLR. L'ANNEXE 2 contient les qualités malt, qui représentent les sorties de notre modèle.

1. Modélisation linéaire multivariable de la qualité du malt

L'objectif de cette partie est de trouver des modèles significatifs capables de décrire l'influence des facteurs considérés d'orge sur les qualités du malt faiblement hydraté humidité, et de prédire les autres indicateurs de qualité du malt **Erreur ! Source du renvoi introuvable.**

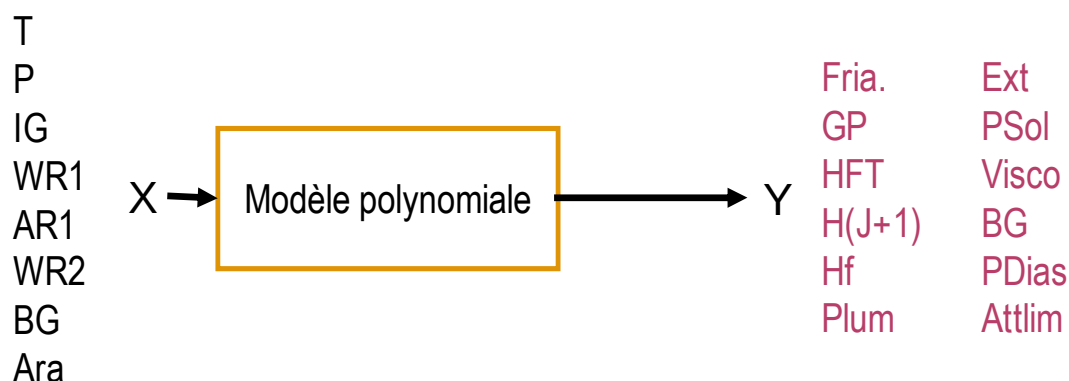


Figure 27 schéma représente la corrélation facteurs-orge et qualité-malt par un modèle polynomial

Dans cette étude, un plan d'expériences de type screening est utilisé. Cependant, les conditions de micro maltage à l'IFBM permettent de réaliser simultanément 16 expériences complémentaires sans augmenter la durée expérimentale. Les expériences complémentaires sont extraites d'un plan de screening correspondant à un modèle avec interactions. La sélection des expériences complémentaires est basée sur l'équilibre du plan et sur la disponibilité de compartiments de micro maltage.

1.1. La réalisation du plan d'expériences

Initialement, les facteurs considérés étaient : la variété d'orge, la température (T), le taux de protéines (P) et le diagramme de trempe (Diag). Par la suite, les informations ont été déclinées en facteurs plus discriminants en les étendant: au taux de protéines (P), à la température (T), à la durée des phases de trempe (WR1, AR1 et WR2), à l'indice de germination (IG) et aux taux initiaux de β -glucanes (BG) et d'Arabinoxylanes (Ara) dans les grains d'orge.

Les mesures relevées sont : l'humidité en fin de trempe (HFT), l'humidité à un jour de germination (Hj+1), humidité en fin de germination (HF), l'extrait sec (Extrait), la viscosité du moût (Visco), le taux de protéines solubles (Prot Sol), le pouvoir diastasique (Pdias), l'atténuation limite(AttLim), le taux de β -glucanes (BG) qui reste dans les grains à la fin de la

maltage, l'indice de plumule, le nombre de grains piqués (GP). Le Tableau 7 présente les indicateurs de qualité de malt mesurés. La Figure 28 précise les points de prélèvements de ces mesures au cours du maltage.

Tableau 7 indicateurs de qualité de malt à mesurer

GP	HFT	Hj+1	HF	Plumule	Fria	Extrait	Prot Sol	Visco	BG	Pdias	AttLim
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
%	%	%	%	%	%	%malt sec	%malt sec	mPA.sec	mg/l	WK sur sec	%

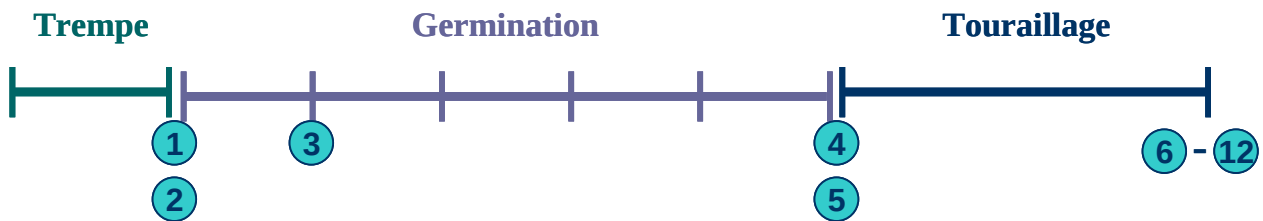


Figure 28 Localisation des points de mesure lors du procédé de maltage

Les 40 conditions opératoires du plan d'expériences résultant et les valeurs des mesures obtenues lors de sa mise en œuvre en micro malterie sont données dans les ANNEXE 1 et ANNEXE 2.

Un programme a été fait sur le logiciel Matlab © pour être capable de traiter le modèle choisi. Ce genre de programmation facilite le calcul et met en évidence toutes les liaisons possibles entre des facteurs du modèle et leurs effets en même temps. Cette méthode est plus efficace que d'utiliser un logiciel commercial, limité dans certains domaines d'utilisation, même si cette souplesse d'utilisation nécessite un plus grand investissement et une plus grande rigueur d'utilisation. Le modèle initialement considéré est le modèle
$$y = a_0 + \sum_{i=1}^{NF} a_i x_i + \sum_{i=1}^{NF} a_{ii} x_i^2 + \sum_{i < j} \sum_{j=1}^{NF} a_{ij} x_i x_j + \varepsilon \quad (14)$$
 du chapitre (3) page (90) sans termes quadratiques ni interactions pour

$$\hat{y} = a_0 + \sum_i a_i x_i \quad (38)$$

La modélisation avec les quarante expériences réalisées au sein de l'IFBM à partir du plan d'expériences proposé n'a pas permis d'identifier des modèles significatifs. De ce fait, les expériences ont été divisées en deux grands groupes : un groupe avec les 21 premières expériences impliquant les orges de printemps (OP), et l'autre avec les 19 expériences impliquant les orges d'hiver (OH). Cette distinction peut s'expliquer par les effets génétiques et environnementaux qui différencient les espèces. Dans une deuxième étape, nous avons étudié d'éventuels effets non linéaires entre facteurs et mesures. Des termes quadratiques et des interactions ont été ajoutés pour évaluer leur pertinence. Le nouveau modèle complet s'écrit ainsi :

$$\hat{y} = \sum_i a_i x_i + \sum_{i,j} a_{ij} x_i x_j + \sum_i a_{ii} x_i^2 \quad (39)$$

Pour faciliter l'analyse des résultats des régressions, les facteurs et les mesures ont été $zN = 2z - z_{max} + z_{min}$ (18) page 91 du chapitre 3). Les valeurs minimales et maximales de chaque facteur et de chaque mesure sont présentées dans le Tableau 5 (cf. page 91 chapitre 3).

Pour ne retenir que les modèles les plus pertinents, trois tests, décrits dans le chapitre 2, ont été utilisés.

La Figure 29 présente les coefficients du modèle de l'humidité en fin de trempe en fonction des 4 facteurs et de l'interaction considérés dans une relation du type :

$$HFT = a_1 T + a_2 WR1 + a_3 AR1 + a_4 BG + a_5 AR1 \times WR1 \quad (40)$$

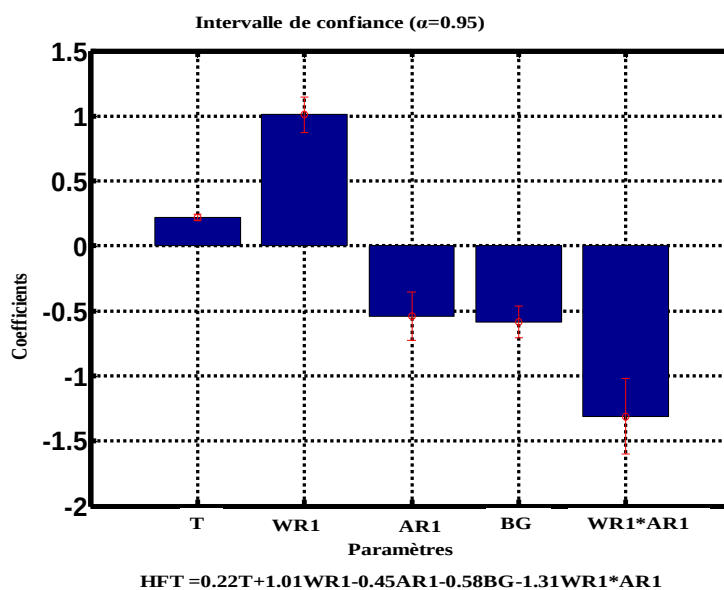


Figure 29 Modèle de l'humidité en fin de trempage pour l'orge d'hiver 2006
Coefficient du modèle et intervalle de confiance (avec un niveau de confiance de 95%)

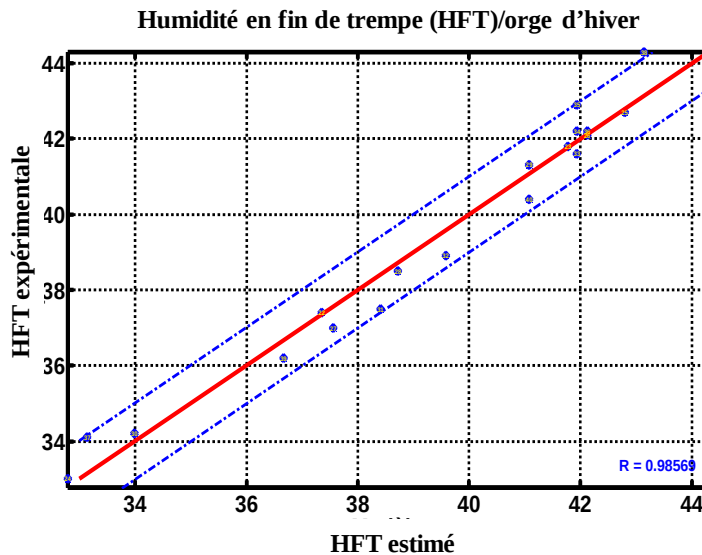


Figure 30 Modèle de l'humidité en fin de trempage pour l'orge d'hiver 2006 :
Coefficient de détermination entre données expérimentales et modélisées.

Les intervalles de confiance sur la Figure 29 confirment que les coefficients retenus dans le modèle de l'humidité fin de trempage sont significatifs. La Figure 30 indique une bonne corrélation ($R = 0.99$ contre R^2) entre les valeurs expérimentales de l'humidité HFT en fonction des valeurs prédites par le modèle. Une estimation de l'intervalle de confiance des valeurs prédites est présentée sur cette même figure à titre indicatif.

Le test de Bootstrap décrit dans le chapitre 2 pages 24 confirme la robustesse du modèle. Une distribution des coefficients du modèle de l'humidité HFT est illustrée sur la Figure 31. Elle montre que les coefficients présentent une distribution relativement homogène dans l'intervalle précisé et qu'ils ne sont donc pas trop sensibles aux choix des expériences.

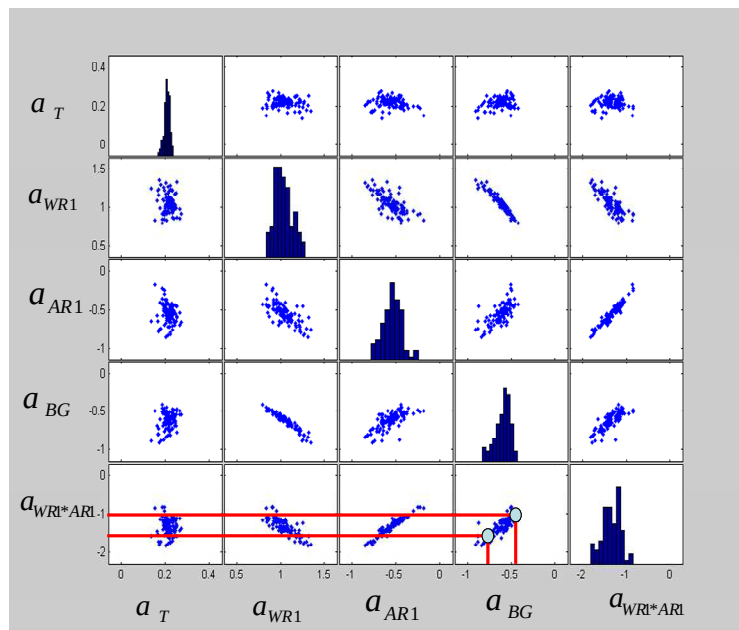


Figure 31: Résultats du test de Bootstrap. Distribution des coefficients du modèle HFT/OH

Les modèles *acceptables* sont résumés dans les ANNEXE 16 et ANNEXE 17. Ces tableaux présentent les coefficients significatifs des modèles ainsi que le coefficient de détermination la valeur de R^2 pour l'orge de printemps (OP) et l'orge d'hiver (OH) respectivement.

A l'issu de cette étude, des modèles souvent acceptables ont été établis pour prédire les qualités du malt des orges d'hiver (OH). Pour l'orge de printemps (OP), on constate que les modèles peuvent encore être améliorés.

2. Détermination et validation de conditions optimales de maltage à faible humidité

A l'aide des modèles significatifs précédents, une phase d'optimisation est lancée pour déterminer les conditions optimales qui permettent d'arriver, dans un maltage à faible humidité (c'est-à-dire avec une humidité en fin de trempé de 37-38%), à des indicateurs de qualité du malt *satisfaisants*. La définition de ces conditions *satisfaisantes* sont données sous forme d'intervalles de valeurs tolérées dans le Tableau 8

Tableau 8 Intervalles de qualité pour un malt considéré comme satisfaisant

Indicateur	Objectif
Humidité fin de trempé (%)	37-38
Friabilité (%)	OH(75-80) OP(80-85)
Extrait (% malt sec)	80%<
Protéine Soluble (% Mat sec)	3,2-5
Viscosité (mP.s)	1,45-1,65
β-glucane (mg/l)	<300
Pouvoir Diastatic (WK sur sec)	OH>300 OP>250
Atténuation Limite (%)	81-83

L'optimisation de ce procédé a été faite en employant l'algorithme génético-évolonnaire du LRGP.

Les résultats de l'optimisation

L'optimisation a été effectuée sur l'ensemble des critères modélisés considérés comme les plus importants pour le procédé de maltage : l'humidité en fin de trempé (HFT) et les qualités du malt suivantes : la friabilité, l'extrait, les teneurs en β-glucanes ou protéines solubles, le pouvoir diastasique et la viscosité du moût.

2.1. Conditions de maltage à faible humidité pour l'orge de printemps

Afin de réaliser les optimisations sur les modèles les plus fiables du Tableau 9 Conditions optimales pour le maltage à faible humidité pour (OP), seuls les modèles suivants ont été exploités pour l'orge de printemps : le modèle 1 de la friabilité, le modèle 3 de HFT, le modèle 1 de l'extrait, le modèle 1 des protéines solubles, et enfin le modèle 1 de la teneur en β -glucanes.

L'optimisation a été réalisée en trois étapes :

- Une première optimisation a permis de définir les conditions conduisant à une friabilité du malt acceptable (i.e. entre 80% et 85% pour une OP).
- Les conditions ainsi obtenue ont été analysées pour ne retenir que celles qui respectaient une humidité en fin de trempe dans l'intervalle acceptable (entre 37 et 38%)
- Ces conditions optimales ont ensuite été classées en grandes catégories en fonction de leur similitude.

Le Tableau 9 présent les cinq groupes des conditions opératoires ont mis en évidence.

Tableau 9 Conditions optimales pour le maltage à faible humidité pour (OP)

les mesures	P %	T (C°)	IG %	AR1 (h)	WR2(h)	BGi mg/l	HFT %	Fria %	Extrait %malt sec	BG mg/l	Prot Sol %malt sec
famille des cond. (1)	9,7	20	5	14	8,2	3,8	37,3	81,7	80,2	129,1	3,14
famille des cond. (2)	10,1	21	6	13	8,9	6,4	37,4	82,5	80,4	111,7	3,4
famille des cond. (3)	10,1	21	6	13	9,8	6,4	37,9	83	80,4	102,4	3,4
famille des cond. (4)	10,4	16	6,8	13,3	8	3,4	37,9	82,3	83,8	219,2	4,6
famille des cond. (5)	11,1	18	7,2	15,7	10,8	3,1	37,4	81,9	81,8	116,4	3,1
experience(12)	9,8	22	8,3	17	7	3,87	44,6	82,8	83,3	86	4,19

Une analyse détaillée de ces résultats a été développée dans le mémoire de master (Ajib, 2008).

2.2. Les résultats de l'optimisation pour l'orge d'hiver

De la même manière, seuls les modèles du Tableau 10 Conditions optimales pour le maltage à faible humidité pour (OH) suivants ont été retenus pour l'optimisation des conditions de maltage pour les orges d'hiver : le 1^{er} modèle pour la teneur en β -glucanes, la HFT, la friabilité, l'extrait, et le pouvoir diastasique, et le modèle (2) pour la viscosité.

L'optimisation a d'abord été réalisée à partir du meilleur modèle disponible, celui de la teneur β -glucanes. Cependant, ce modèle ne s'appuie pas sur l'ensemble des facteurs opératoires. Ainsi, pour chaque condition optimale identifiée, une deuxième phase d'optimisation est nécessaire pour déterminer les conditions opératoires *manquantes* qui conduisent à une humidité en fin de trempe HFT acceptable. Les conditions compatibles entre une teneur en β glucanes optimale et une l'humidité fin de trempe HFT acceptable sont ensuite utilisées pour évaluer tous les autres indicateurs de qualité du malt.

Les conditions satisfaisantes pour tous les indicateurs de qualité de malt sont résumées dans le Tableau 10.

Une analyse détaillée de ces résultats a été développée dans le mémoire de master (Ajib, 2008).

Tableau 10 Conditions optimales pour le maltage à faible humidité pour (OH)

les mesures	P%	T(C°)	IG %	WR1(h)	AR1(h)	WR2(h)	BG mg/l	HFT(%)	Fria%	Extrait %	Visco %	BGfinalmg/l	Pdias WK sur sec
famille des cond. (1)	10,5	16,1	9,5	(12-19)	21	9	3,9	(37,3_38)	77,2	80	1,62	175	398
famille des cond. (2)	10,4	16,3	9,3	13	20	8,5	3,9	37,9	75,1	80	1,64	194	386
famille des cond. (3)	10,4	17,4	8,5	(17-22)	20	9	4,2	(37,1_37,5)	75,8	80	1,63	178	375
famille des cond. (4)	10,1	15	8,4	24	20	15	4,5	37,2	78,2	80,1	1,61	187	393
famille des cond. (5)	11,3	18	7,6	22	19	17,3	4,5	38	77	80,2	1,64	229	388
experience(33)	11,3	22	9,5	6	21	5	3,41	41,6	84,2	80	1,6	88	400

2.3. Validations des conditions optimales

La modélisation et l'optimisation ont été réalisées sur des orges récoltées en 2006. L'objectif de cette validation est de vérifier expérimentalement les conditions optimales numériques présentées.

Les échantillons d'orge disponibles ont cependant évolués lors de stockage pendant la période de l'étude (de 2006 à 2009). Pour intégrer cette évolution, les indices de qualité (Indice de germination, teneurs en protéines et β -glucanes) des rares échantillons d'orge encore disponibles à l'IFBM ont été réévalués en 2009.

Le protocole de validation des conditions optimales est le suivant :

- 1 A partir des nouvelles caractéristiques des orges, les conditions opératoires ont été ré-optimisées. Les résultats sont rassemblés dans la Tableau 11, notamment les valeurs prédites par nos modèles, de l'humidité de fin de trempe et de la friabilité du malt.
- 2 Les diagrammes de trempe ont été appliqués expérimentalement.
- 3 Les échantillons d'orge 2006 ont été maltés selon les conditions opératoires optimales.
- 4 L'humidité en fin de trempe et la friabilité du malt expérimentales et prédites ont été comparées les résultats sont présentés dans Tableau 11 .

Tableau 11 les nouvelles propriétés de l'orge 2006, mesurées en 2009 et les diagrammes de trempes proposés par l'optimisateur

Variété	Facteurs orge			Facteurs processus			
	P	IG	BG	T	WR1	AR1	WR2
	%	%	%	C	h	h	h
Azurel	11.4	4.7	2.93	16	6	16	8
Esterel	9	5.4	2.59	15	6	14	5
Scarlett	10.2	5.5	4.28	18	11	14	10
Sebastian	11.7	4.8	2.95	21	8	15	7

Tableau 12 présente l'humidité fin de trempe et la friabilité du malt résultant d'application de diagramme de trempe optimale sur les 4 orges 2006 et celles prédites pour la modèle.

Variété	Humidité fin de trempe expérimentale	Humidité fin de trempe prédite	Friabilité expérimentale	Friabilité prédite
Azurel	37.2	37.5	76	83
Esterel	36.3	37.5	60	82
Scarlett	40.3	37.5	86	74
Sebastian	41	37.5	70	74

La validation expérimentale des conditions optimales sur les données restant de 2006 présenté dans le Tableau 15 confirme que les orges de printemps (Sebastian et Scarlett) sont difficiles à modéliser. A ce qui concerne les orges d'hiver, avec nos modèles on est arrivé à une humidité de fin de trempe très proche de celle prédite, mais avec une friabilité pas aussi bonne. Ça peut être expliqué par le fait que l'orge a subi des changements physico-chimiques pendant le stockage de longue durée. L'humidité en fin de trempe pour l'orge Esterel est basse par rapport à celle prédite : c'est très bien d'un point de vue économique et environnementale

mais un endosperme pas suffisamment hydraté conduit à du malt de mauvaise qualité comme le montre le taux de friabilité expérimental d'Esterel (60%).

L'ensemble de ces résultats montrent que :

- 1 Qu'il est possible de relier par des modèles robustes et significatifs, la plupart des indicateurs de qualité du malt aux conditions opératoires et aux indicateurs de qualité de l'orge
- 2 Qu'il faut cependant distinguer les orges d'hiver (OH) des orges de printemps (OP).
- 3 Qu'il est possible de déterminer des conditions de maltage à faible humidité conduisant à des qualités de malt acceptables,

Cependant, la variabilité des échantillons dans le temps et la variabilité des propriétés de l'orge d'une année de récolte à l'autre sont des facteurs de perturbation qui ne sont pas intégrés dans nos modèles. Cela exigerait une recalibration annuelle des modèles pour permettre l'étape d'optimisation des conditions opératoires. Cela s'avèrerait trop long et trop coûteux.

Par ailleurs, la teneur en protéines et en β -glucanes n'expliquent pas les variations observées. On peut d'ailleurs mettre en évidence ce résultat à travers la cinétique de prise d'eau des grains pendant la trempe.

3. Mise en évidence du découplage entre la composition de l'orge et l'imbibition des grains d'orge

Nous avons sélectionné un lot de quatre échantillons d'orge qui présentent des comportements antagonistes mais intéressants lors du maltage.

Tableau 13 les variétés d'orge récolte 2008 qui ont choisi pour réaliser la suivi d'humidité avec leurs teneur en protéine, leurs teneurs d'humidité en fin de trempe et la friabilité de malt résultant.

Variété	P %	HFT %	Friabilité %
Azalea	10.8	41.4	63.74
Thorgall	10.9	38.1	53.38
Sebastian	9.4	40.9	87.26
Henley	9.3	38.6	83.4

Notre sélection contient deux orges qui, pour un même processus de maltage et une humidité en fin de trempe optimale > 40 %, conduisent soit à un malt de très bonne qualité, soit à un malt *inacceptable*. Inversement, nos échantillons présentent deux orges qui, pour une humidité en fin de trempe faible < 40 %, conduisent également soit à un malt de très bonne

qualité, soit à un malt *inacceptable*. Tous ces échantillons ont des compositions du même ordre de grandeur mais différentes.

L'étape de trempe est l'étape clé pour la technologie de maltage. L'hydratation des grains a lieu dans cette partie avec ces phases (sous eau/sous air), elle est le facteur clé pour toutes les étapes suivantes et détermine donc la qualité de malt final. Afin d'étudier l'origine des différences de comportements observés, l'hydratation des grains en cours de trempe a été étudiée. Les cinétiques obtenues pendant le maltage sont présentées dans de la Figure 32

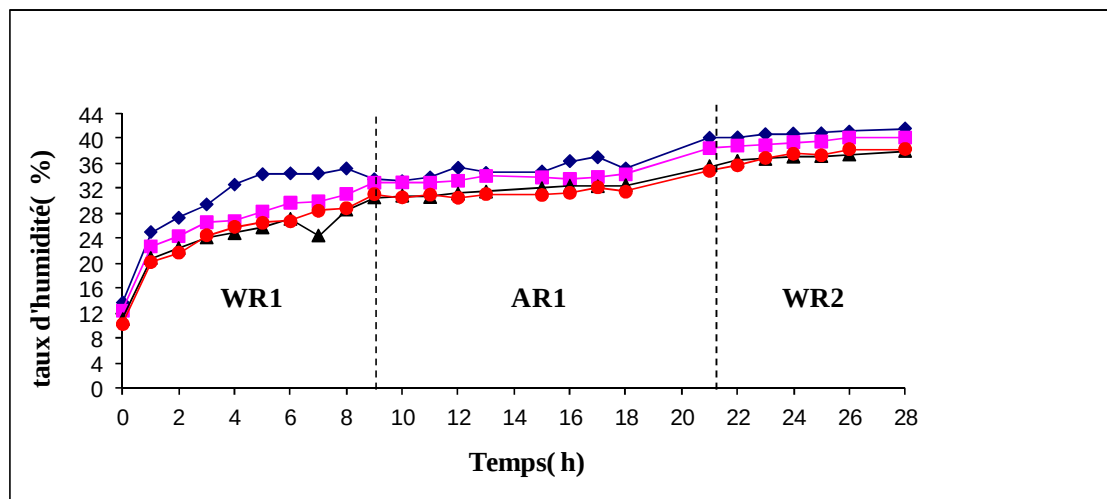


Figure 32 Suivi du taux d'humidité pendant les phases de trempe à 16 °C pour quatre variétés d'orge : ◆ Azalea,

■ Sebastian, ● Henley et ▲ Thorgall.

(WR1 premier sous eau, AR1 premier sous air et WR2 deuxième sous eau. avec quatre variétés d'orge)

A partir de la Figure 32, on constate que la première sous-eau est la seule phase qui révèle une différence dans le comportement des quatre variétés : la courbe d'absorption d'eau de la variété Azalea est bien distinguée des autres. Cette différence est moins perceptible pour Sebastian et les deux orges restantes qui ont un comportement très similaire. Cela nous a orienté vers le calcul des constantes cinétiques de prise d'eau k dans le premier phase sous eau WR1. Les résultats sont donnés dans Tableau 14. On constate que les deux orges ayant la plus haute constante cinétique de prise d'eau $k=0.0352$, $k = 0.0211$, Azalea et Throgall ont aussi les teneurs en protéines les plus hautes 10.8, 10.9 respectivement. Mais les deux orges se comportent différemment dans la prise d'eau. Cela peut être lié aux variétés (facteur génétique). Ce résultat a également été récemment présenté par Mayolle et al.(2012).

Ces résultats sont cependant en contradiction avec Molina-Cano et al.(2002) et Chandra et al.(1999) qui ont montré que la teneur en protéines ont un effet négatif sur la cinétique

d'absorption et avec Swanston et al.(2006) et Gamlath et al.(2008) qui ont montré que ce facteur n'influçait pas la prise d'eau.

Tableau 14 les constants de cinétique de prise d'eau pour les quatre variétés d'orge dans la première phase de trempe.

Variété	Constante de vitesse d'absorptin (k)	P %	HFT %	Friabilité %
Azalea	0.0352	10.8	41.4	63.74
Thorgal	0.0211	10.9	38.1	53.38
Sebastian	0.0191	9.4	40.9	87.26
Henley	0.0183	9.3	38.6	83.4

Aucune différence n'est ensuite détectable pendant le 1er sous air et le 2ème sous eau. La différence de composition en protéines a donc une influence sur la prise d'eau lors de la trempe mais elle ne peut pas seule expliquer les différences de comportement pendant le maltage. Il est donc nécessaire de rechercher de nouveaux facteurs pour établir nos modèles. Il est indispensable d'intégrer dans ces modèles des indicateurs qui permettraient de prendre en compte cette variabilité dans le temps et d'expliquer les cinétiques de prise d'eau. Nous avons donc cherché à ajouter des nouveaux facteurs caractérisant l'orge. Notre étude s'est orientée vers des facteurs structurels de l'orge : la porosité, la dureté des grains d'orge.

4. Nouveaux facteurs orge pour améliorer les modèles de qualité du malt

Il existe de nombreuses caractéristiques structurelles de l'orge. Dans le cadre du projet ANR MaltEco [ANR-PRNA 2006-000], les résultats internes ont montré que la porosité, la dureté, la vitrosité permettaient de distinguer les échantillons d'orge. Le comportement de différentes variétés d'orges en maltage semblait liés à ces propriétés structurelles, notamment en terme de capacité d'imbibition du grain d'orge. Notre étude s'est focalisée sur l'exploitation de ces facteurs dans nos modèles.

Cette même étude du projet MaltEco, a montré que la vitrosité et la dureté étaient corrélées à la porosité. Seul le facteur porosité a donc été considéré dans notre étude.

4.1. Mesure de porosité comme nouveau facteur orge

Afin d'étudier le rôle d'un facteur structurel sur la qualité du malt et sur les performances du maltage à faible humidité, nous avons sélectionné 36 échantillons d'orge de 2007 et 2008. Ces échantillons ont été notamment sélectionnés pour présenter des grandes variations de la qualité de l'orge, liées à l'année de récolte, à la variété d'orge, au lieu de cultiver. Ces échantillons conduisent aussi à une grande diversité de qualité de malt, pour des conditions de maltage standard. Les résultats expérimentaux associés à ces échantillons sont présentés dans les ANNEXE 11. Une porosité variée entre [3.73 et 9.15 %] pour une teneur en humidité dans l'intervalle [12.69 et 17.31%].

Des modèles polynomiaux de premier degré (sans interactions, ni terme quadratiques) ont été construits avec tous les facteurs orge initiaux (P, IG, BG, Ara) et la porosité pour améliorer la corrélation orge-malt. Les coefficients significatifs associés à chaque facteur et les coefficients de détermination R^2 des modèles des indicateurs de qualité du malt sont présentés dans le Tableau 15 pour un lot de 12 orges de 2008. Les meilleurs résultats obtenus montrent que la porosité n'apporte aucune amélioration dans la prédiction du modèle. ($R^2_{\max}=0.57$).

Tableau 15 Intégration de la porosité comme facteur du modèle MLR : coefficients significatifs du modèle pour chaque facteurs et coefficients de corrélation

Orge		Facteurs orge						R^2
	SUNT2008	Cst	P	IG	BG	Ara	Poro	
Qualités de malt	HFT		-0.14	0.1	0.03	-0.015	0.0003	0.13
	Fria	-1.74	-0.91	-1.78	-0.2		0.079	0.34
	Extr	-0.39	-0.04	-0.62	0.09		-0.1	0.18
	ProSol		-0.3	0.96	-0.13	-0.06	0.29	0.57
	Visco	-0.002	0.1	0.71	-0.15		0.09	0.48

L'influence de la porosité sur l'imbibition du grain ne semble donc pas avoir d'effet significatif sur la qualité du malt. Ce résultat est cependant en contradiction avec les observations réalisées dans le projet MaltEco. Des analyses complémentaires ont été réalisées pour expliquer cette incohérence apparente.

4.1.1. La relation porosité-humidité des grains d'orge

Les mesures de porosité ont été répétées sur plusieurs échantillons stockés à -12 °C dans la chambre froide. Ces répétitions ont mis en évidence une forte dépendance de la porosité au taux d'humidité des grains comme le montre les répétitions présentées dans le Tableau 16.

Des études sur la porosité ont confirmés cette résultat: La porosité des céréales augmente linéairement avec le taux d'humidité, pour le cumin, le tournesol, le *Vigna unguiculata* et l'orge par (Singh et Goswami, 1996 ; Gupta *et al.*1997 ; Yalçin, 2006 ; Oztirk et Esen, 2008) respectivement.

La densité des grains : est la masse de matériau contenue dans un volume donné. Il y a la densité apparente et la densité réelle.

Tableau 16 Répétition des mesures de porosité et du taux d'humidité après 30 jours de stockage en chambre froide (-12°C) pour 9 échantillons d'orge 2008

<i>Variété</i>	<i>Taux de humidité %</i>	<i>Porosité %</i>
Concerto-001	13.23	5.91
	15.96	7.01
Henley-020	11.05	3.36
	16.20	4.48
Victoriana-005	12.90	4.40
	15.75	6.89
Henley-007	16.22	5.19
	15.46	6.88
Signora-010	12.33	4.39
	15.62	7.31
Cellar-013	14.77	7.58
	15.65	7.13
Calcule-017	13.08	5.10
	15.10	5.98
Prestige-018	12.58	4.04
	16.56	4.60
Sebastian-019	11.06	4.35
	15.60	6.26

Le protocole de mesure de porosité ne permet donc pas de caractériser l'orge de façon unique. Les variations observées expliquent ainsi l'absence de corrélation apparente entre la porosité et les qualités du malt que nous cherchons à prédire dans les modèles précédents. Pour des raisons techniques et ne disposant pas de l'équipement de mesure de porosité, ce facteur a été abandonné.

5. Conclusion

L'objectif de ce chapitre de notre mémoire a été de faire un suivi pour finaliser et améliorer une démarche scientifique précédente consistant à maîtriser les coûts (énergétiques et environnementaux) du procédé de maltage en travaillant à faible humidité tout en conservant la qualité finale de malt.

Dans un premier temps, on a focalisé sur la phase de trempe, pour essayer d'étudier l'impact de diagramme de trempe sur la pénétration d'eau dans les grains, et donc sur la qualité du malt. Cela a été fait par une réalisation d'un suivi de taux d'humidité des grains d'orge (venant de quatre variétés d'orge différents avec des comportements très différents) pendant la trempe avec une mesure de qualité de malt final. Les résultats ont montrés que :

- Les quatre variétés montrent des comportements différents dans la première phase sous eau, les orges ayant le taux de protéines le plus élevé ont une prise d'eau plus importante que les autres orges. mais elle ne peut pas à elle seule expliquer les différences de comportement pendant le maltage

Il est indispensable d'intégrer dans ces modèles des indicateurs qui permettraient de prendre en compte cette variabilité dans le temps et d'expliquer les cinétiques de prise d'eau.

Ensuite, des facteurs structurels de qualité d'orge ont été ajoutés dans nos modèles pour tester les effets des autres qualités d'orge sur la qualité du malt final. La porosité, et la dureté deux qualités structurels de l'orge. En ajoutant les résultats de ces deux qualités obtenus dans nos modèles on peut conclure en deux points :

- On ne peut pas accepter les deux qualités d'orges comme facteurs indépendantes dans nos modèles car ils sont tous les deux fortement liée au taux d'humidité de grain.
- Les deux qualités n'apportent pas une amélioration significative sur nos modèles

La réponse à la première perspective de la modélisation a été faite.

Références bibliographiques du chapitre 4

- Ajib, B., 2008. Maltage à faible humidité. Master BAAN, ENSAIA.
- Bicking F., Algorithmes génétiques Commande de processus biotechnologiques 1994
- Brookes, P.A., Lovett D.A, Mac William I.C., 1976. The steeping of barley, a review of the metabolic consequences of water uptake and their practical implications, *Journal of institute of brewery* 82, 14-26.
- Chandra, G.S., Proudlove, M.O., Baxter, E.D., 1999. The structure of barley endosperm – an important determinant of malt modification. *Journal of the Science of Food and Agriculture* 79, 37–46.
- Debourg A., cours de techniques brassicoles pour la formation approfondie .1998
- De Francisco, A., Varriano-Marston, E., Hosney, R.C., 1982. Hardness of pearl millet and grain sorghum. *Cereal Chemistry*; 59, 5-8.
- Fliedel, G., Grenet, C., Gontard, N., Pons, B., Céréales en régions chaudes 3 : Dureté caractéristiques physicochimiques et aptitude au décorticage des grains de sorgho. AUPELEF, EDs John Libbey Eurotext, Paris©1989, pp.187-201.
- Gamlath, J., Aldred, G.P., Panozzo, I.F., 2008. Barley (1- 3; 1-4)-b-glucan and arabinoxylan content are related to kernel hardness and water uptake. *Journal of Cereal Science* 47, 365-371.
- Gibbons G.C., the effects of oxygen, nitrogen and carbon dioxide steeping regimes on endosperm modification and embryo growth of germinating barley seeds, 1983.
- Goupy J., La Méthode des plans d'expériences : optimisation du choix des essais et de l'interprétation des résultats, 1989.
- Hariri A., Etude et modélisation de la trempe en malterie, Thèse de l'Institut National Polytechnique de Lorraine, 2003.
- Hjmanger, Technology brewing and maltiry, 2004.
- Igathinathane C., Chattopadhyay P.K., Moisture diffusion modelling of drying in parboiled paddy components. Part I: strachy endosperm, *J. Food Engineering*, 41, 89-101. 1999
- Kim, Kish -Meng, influence de la teneur en eau sur les activités enzymatiques participant a l'élaboration des dérivés carbonylés au maltage ,1980.
- Kettunen A .J.J ;Hamalanen K., Stnholm et K.Pietila , an modèle for the prediction of β -glucanes activity and β -glucanes concentration during mashing ,*J.Food Engineering* ,29,185-200, 1996
- Kreisz, S., ,Hartmann, K., , Zarknow, M., Thiele, F., Burberg, F., Back, W., A new statistical method to evaluate the malting performance of new barley varieties,2007.
- Massebeuf, S., Optimisation multicritère de procédés discontinus d'homopolymérisation et de copolymérisation en émulsion, 2000.
- Miller B.S., Afework S., Hugues, J.W., Pomeranz, Y., 1981. Wheat hardness: time required to grind wheat with the Brabender automatic microhardness tester. *Journal of food science* 46, 1863-1865.

MODULAD, <http://www-rocq.inria.fr/axis/modulad/numero-35/Tutoriel-confais-35/confais-35.pdf>, 2006.

Molina-Cano, J.L., Sopena, A., Polo, J.P., Bergareche, C., Moralejo, M.A., Swanston, J.S., Glidewell, S.M., 2002. Relationships between barley hordeins and malting quality in a mutant of cv. triumph. II. Genetic and environmental effects on water uptake. *Journal of Cereal Science* 36, 39–50.

Morgan A G. Gilland A. A ., D.B Smith D.B., some barley grain and green malt properties and their influence on malt hot water extract ,1983.

Schimmerling P. , Sisson J.C., Zaidi A., *Pratique des plans d'expériences* 1998

Swanston, J.S., Newton, A.C., Hoad, S.P., Spoor, W., 2006. Variation across environments in patterns of water uptake and endosperm modification in barley varieties and variety mixtures. *Journal of the Science of Food and Agriculture* 86, 826–833.

Urion, E., Rouleau, H., *Bière et malt*. 1948.

Viennet R., *Nouvelle technologie de planification expérimentale pour l'optimisation multicritères de procédés*. Thèse de l'Institut National Polytechnique de Lorraine, Nancy, France, 1997.

Walter E. et Pronzato L., *Identification des modèles paramétriques à partir des données expérimentales*, Collection M. A. S. C, Ed Masson, Paris ,1994.

Chapitre 5

Classification of barley according to harvest year and species by using mid- infrared spectroscopy and multivariate analysis

Ce chapitre est présenté sous la forme d'un article. Les méthodes d'analyse de l'orge et du malt ont été décrites dans le chapitre 3. Les résultats de cet article ont été établis à partir des échantillons d'orge de 72 variétés d'orges, cultivées dans 16 lieux en France, récoltées en trois années.

Classification of barley according to harvest year and species by using mid-infrared spectroscopy and multivariate analysis

AJIB B.¹, FOURNIER F.¹, BOIVIN P.², SCHMITT M.², FICK M.¹,

⁽¹⁾ Université de Lorraine, LRGP, 2 Avenue de la Forêt de Haye, TSA 40602, 54518 Vandœuvre-lès-Nancy

⁽¹⁾ CNRS UMR 7274, LRGP, 2 Avenue de la Forêt de Haye, TSA 40602, 54518 Vandœuvre-lès-Nancy

⁽²⁾ IFBM, 7, rue du Bois de la Champelle B.P. 267, 54512 Vandœuvre-lès-Nancy Cedex

Abstract:

In order to monitor malt quality in the malting industry, despite yearly variations in the barley quality, a large number of barley samples were analysed using conventional (Moisture, Protein and β -glucan content) and FTIR Mid-InfraRed spectroscopy. The experimental dataset included barley from 3 harvest years, two barley species, 77 barley varieties, and two-row and six-row barley, from 16 cultivation sites. For each sample, the malt quality indices were also assessed according to EBC standards.

Principal Component Analysis (PCA) was carried out on mean-centred, normalized, and derivative spectra using 200 cm^{-1} width spectral bands. The most informative spectral bands were observed in the [800-1000] cm^{-1} and [1000-1200] cm^{-1} ranges. PCA revealed that barley harvested in 2010 and in 2011 had bands that were very close together, while 2009 harvest clearly displayed a difference in its quality. PCA analysis was performed on subsets defined according to harvest year or barley's specie and row. PCA made it possible to distinguish two species and confirmed that 2-row winter barley quality was closer to 2-row spring barley quality than to 6-row winter barley. Results indicate that MIR spectrometry could be a very useful and rapid analytical tool to assess barley qualitative quality.

Keywords: Malting barley, mean infrared spectroscopy, principal components analysis

1. Introduction

Malt is the basic ingredient used in the production of beer. Malt provides most of the starch, enzymes and proteins which are necessary to give beer its distinctive flavour and alcohol. Malt quality, therefore, is essential.

Malt quality is the result of many factors, ranging from barley quality, to plant sanitation, to process controls. The world production of barley for breweries, which is in the neighbourhood of 25 MT, is intended to be transformed into malt and then into beer. However, the world beer consumption is in full expansion (2 billion hl) (IFBM) and maltsters are constantly searching for a better understanding and controlling of the malting process in order to get the best quality of malt possible.

The purpose of malting is to produce endogenous enzymes, breaking down the matrix surrounding the starch granules to prepare it for conversion before stopping this action for storage [1].

The malting process includes three main steps: steeping, germination and kilning.

(i) Steeping: to increase moisture from 15% to 42-44% by alternating periods of soaking in water and aeration.

(ii) Germination: growing rootlets and acrospires with cool and humid airflow.

(iii) Kilning to reduce the moisture to 4-5% by dry airflow. The combination of low moisture grain and high temperatures denatures many enzymes while keeping those necessary for later starch conversion in brewing .

During steeping, alternating periods of soaking and aeration will allow the grain to take in water with the purpose of hydrating the endosperm to a sufficient level to allow for radicle protrusion, i.e. germination from a physiological viewpoint. At the end of the steeping phase, all kernels should be chitted, which means germinated, and the entire endosperm should be hydrated.

During the germination phase, by blowing cold and humid air to maintain moisture at around 42%, the embryo will produce gibberellins that will engage in producing enzymes in the aleurone layers. Then, the enzymes, particularly the β -glucanases, will need to move inside the endosperm to break down the cell wall. In order for the cell wall to break down properly, the endosperm should be sufficiently hydrated.

During kilning, which is a step that involves a huge amount of energy consumption, moisture is eliminated to get malt with 4 to 5 % moisture, a desired coloration and a specific aroma according to the type of beer.

To sum up, the malting process produces enzymes, some of which will be used by the brewer in the mashing step, such as amylases to convert starch into fermentable sugars, arabinoxylanase to hydrolyse malt arabinoxylans. Others, such as proteolytic enzymes and glucanases, will be involved in the malting process. So, malt will contain starch, proteins, modified proteins, and only small amounts of β -glucans because, during malting process, 90%

of β -glucan is hydrolysed to obtain friable malt. Modifying structures, particularly the endosperm cell wall, is crucial to get good malt.

Malt quality is essentially related to two kinds of factors: (i) the operatory conditions of the malting process: steeping (temperature, steeping diagram), germination (temperature, duration of germination), kilning (temperatures, duration of kilning), (ii) raw material qualities (moisture content, protein content, β -glucan content, and arabinoxylan content, germination index). To control malt quality one needs to investigate and understand the impact of these factors on the malt.

As far as operating factors are concerned, malting control is a real challenge because of the various operating parameters during malting. These factors have already been widely investigated. Correlations between the steeping program (temperature, soaking and aeration periods) to the germination quality were found [3]. Studies have shown that higher water content (longer soaking) and longer germination influences the malt extract and decreases filterability [4]. The influence of the whole malting process on the lipid content in the malt was studied [5]. The steeping and germination were shown to provide optimal conditions for lipid degradation, whereas kilning proved to have no influence. But, for further investigation into control malt quality, it is not sufficient to study only the malting process. The biological structure of barley and the complex transformations it undergoes during the malting process must also be taken into account.

As far as the raw material qualities are concerned, barley is a complex system which is characterized by a wide range of physical and chemical indices. Maltsters are looking for barley that will germinate evenly and easily. Various studies and analyses were carried out on the factors affecting barley quality and, therefore, malt quality. The environmental and geographical effect on barley protein and arabinoxylan content respectively were investigated ??? . Several protein genes influencing malt quality could be identified [8]. Other studies focused on the effect of barley quality factors on malt quality [10] and confirmed that high-protein contents decrease available carbohydrates, thus having a negative influence on the brewing process. The influence of barley genotype on the apparent attenuation limit (AAL) was investigated [11]. Barley variety was found to have more important influence on protein degradation and cytolytic degradation than the technological factors [4]. Other studies have found that biological composition of the barley endosperm affected malt quality [12][13]. Steelier barley produced less root mass, more free amino nitrogen and soluble protein, and malt of a less friable structure.

In fact, there is no perfect barley for good malt quality. Indeed, the composition of the raw material is related to various factors such as the year of harvest, the barley species, the number of rows in the barley, the site of cultivation and the barley variety. In this study, we aim to show a correlation between the variation in the barley's raw material and the differences in the malt it produces. A large dataset is required for such an analysis. Traditionally, a complete quality analysis includes measuring 4 to 7 physical and chemical parameters for raw barley. This analysis is usually long and expensive. Alternative methods

were suggested to easily and rapidly collect information on barley and thus to facilitate the selection of new cultivars by maltsters. Due to the sensitivity of proton nuclear magnetic resonance, water mobility in hydrated barley was studied by ??? and it was found that the uptake of water and its distribution depends on the composition of the grain. NMR-imagery was coupled with morphological analysis (image acquisition, segmentation, external and internal image feature extraction, classification and interpretation) [15]. They found that morphological features may be successfully employed for preliminary varietal identification of barley kernels. Moreover, it was proved that the Rapid Visco Analyser coupled with multivariate analysis was a very useful analytical tool to study the starch pasting process and were able to predict some important quality factors of barley [16].

Conventional methods involve time consuming, laborious and costly procedures, including physical and chemical analysis or microscopy analysis. A rapid and inexpensive method based on spectroscopy for the certification process would be ideal for characterising barley quality before the malting process.

Infrared spectroscopy is an easy and rapid method which requires little or no sample preparation. Both Near InfraRed (NIR), Mid-InfraRed (MIR) IR methods are based on vibrational spectroscopy which offers the possibility for fast and flexible analysis of a large number of samples. These methods have already provided analytical approaches for the agricultural and food industry as alternatives to traditional analytical methods. In addition, they can be used on samples of almost any morphology.

NIR spectroscopy is defined by a range of electromagnetic radiation in the region of about 780-2500 nm (12821-4000 cm^{-1}). The absorption of radiation by the molecules in this region is primarily caused by overtones and combinations of fundamental vibration bands that appear in the MIR region. NIR was already well established in evaluating grain composition. It was concluded that seed phenotyping of barley by NIR and chemometric was a new, reliable tool for characterising the effects of mutant genes in selecting barley for quality in plant breeding ???. Models for quickly and reliably predicting moisture and total lipid content in maize by means of near-infrared (NIR) spectroscopy were developed [18]. Furthermore, NIR was proved to have potential for predicting the quality of the whole malt grain (prediction of moisture and nitrogen content, of malt extract and protein content, of the progress of the malting process).

Mid-IR spectroscopy is a physico-chemical method based on measuring the vibration of a molecule excited by IR radiation at a specific wavelength range. The mid-IR (400-4000 cm^{-1}) is the most commonly used region for analysis as all molecules possesses characteristic absorbance frequencies and primary molecular vibrations in this range [22]. As IR radiation passes through a sample, specific wavelengths are absorbed, causing the chemical bonds in the material to undergo vibrations that make them stretch, contract, and bend. Functional groups present in a molecule tend to absorb IR radiation in the same wave number range regardless of other structures in the molecule, and spectral peaks are derived from the absorption of bond vibrational energy changes in the IR region [23]. The correlation between

IR band positions and chemical structures in the molecule can thus be exploited. Mid-IR spectrometry was widely used in the spirit drinks domain. The FTIR is proved to be a useful tool in the quality control of alcoholic beverages [24]. FTIR spectroscopy is routinely used in many oenological and wine cellars [25]. The mid-IR spectroscopy was also shown to be a valuable analytical tool for measuring beer quality [26]. It has also been applied in food biotechnology: the Mid-IR is used in the characterisation of edible oil and fat [27] [28] . The feasibility of the Mid-IR technique as a rapid analyser for dairy products was demonstrated [29]. FTIR was also used to study well-defined components such as plant cell wall polysaccharides [30] [31]. FTIR could usefully complement the NIR in monitoring grain transformation during the malting process.

Predicting compositions in spectra from large numbers of samples is very difficult because of the numerous interactions between the signals of the chemical components in the samples and by random differences in the spectral data. The goal of using IR spectra to predict chemical composition in complex samples has been made possible by applying multivariate approaches to data analysis. The simplest and possibly the most commonly used multivariate approach is PCA. This was first described in theoretical terms by Karl Pearson in 1901 [32] and later put into practical application by Hotelling in 1933 [33]. With PCA, the variance in the spectra data at hundreds of wave numbers is condensed into a much smaller number of new principal component-linear combinations of the original variables 00. These few principal components explain most of the variance in the original data, and facilitate an efficient interpretation of the sample spectral prosperities by qualitative analyses such as discrimination or classification 00. Due to there being fewer variables, the data can be more easily explored graphically or statistically for correlations with experimental, environmental and sample effects. PCA works best with highly correlated data and spectral data is highly correlated. Any single data point in a spectrum is highly influenced by the value of its immediate neighbours. PCA is of great value in helping to understand differences between samples in geographic location, growth year and species. Usually, only the first two or three PCs are considered. PCA thus presents the influence of the first principal components on a 2D or 3D graph [38] [39]. The PCA is widely combined to IR spectroscopy. It was concluded that NIR classification with the algorithm PCA is a strong tool for obtaining a preliminary overview of associations between variations in the malt quality and the barley samples 0. A comprehensive overview of an MIR application coupled with PCA analysed to determine the quality of several ingredients and food products was established by [41].

The aim of the present work is to analyse and classify a large set (394) of spring and winter barley, growing in 16 different locations, of 77 varieties, using Mid-IR spectroscopy in combination with multivariate analysis (PCA), according to year of harvest and species. Our goal is to demonstrate the utility of the Mid-IR as a rapid, inexpensive, and efficient classification tool in the malting and brewery domain.

2. Material and methods

2.1 Sample collection

A total of 394 Barley samples, provided by the French Institute of Brewing and Malting (IFBM), were studied. The samples were harvested in 2009, 2010 and 2011. They were composed of 256 spring barley and 138 winter barley, with 77 different varieties, grown in 16 different sites in France. Figure 33 depicts the locations of the harvested barley. The sample set thus offers a unique database for studying the barley quality index.

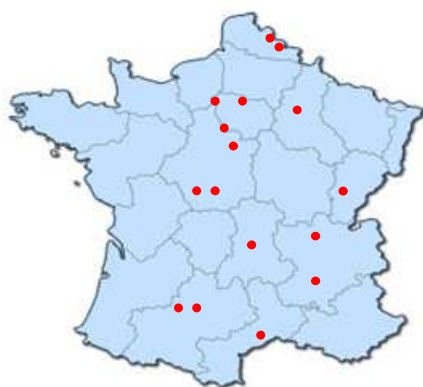


Figure 33 Location of the 16 cultivation sites of barley in France

Quality requirements for malting barley are reasonably strict and are directly related to processing efficiency and product quality for the malting and brewing industries. Many of the characteristics of barley are related to the malting process. Others are determined by weather conditions during the growing and harvesting seasons. There are two quality specifications for malting barley: (i) grain quality parameters, (ii) malt quality parameters.

Grain quality parameters

Under the same malting conditions, barley can give a wide range of malt qualities. Barley quality parameters change according to environmental and genetic factors. In this work, four barley parameters were chosen as indicators of barley quality: moisture content, protein content, β -glucan and arabinoxylan content. Moisture content is used to assure that the biological material is stored under suitable conditions to avoid raw material degradation. Barley that is too dry has excess oxygen which allows molds and yeasts to continue to grow, resulting in ruined feed from excessive mold and rotting. However, grain with high moisture

content will rot quickly [42]. Barley protein content is one of the essential barley components for beer quality. On the one hand, malting barley with high protein content results in lower extracts for the brewer. It also slows down water uptake during steeping, potentially affecting the final malt quality. On the other hand, a very low protein level results in a lack of the enzymes necessary to modify the barley kernel and to break down the starch during brewing [43][44]. Similarly, β -glucan content is a key factor for breeders and malt producers. In the barley grain, β -glucan is a component of the endosperm cell wall which needs to be degraded during malting so that starch will be accessible to amylase and other enzymes during brewing. It is thus closely associated with malt quality. β -glucan has long been recognized as having the potential to cause mash and biofiltration problems in the brewing process. This includes poor mash conversion, retarded and inefficient wort and beer filtration, and nonbacterial colloidal hazes in the final production. Therefore, better malting performance is expected to be associated with lower levels of β -glucan in grains [45][46]. Arabinoxylan content is also an important factor. In general, insufficient degradation of the cell wall hinders the diffusion of germination enzymes and the mobilization of kernel reserves, hence, reducing the quantity of malt extract. Arabinoxylan is one of the major components of the cell wall. High levels of arabinoxylan lead to highly viscous wort and beer, increasing filtration problems in the brewery. In malt barley breeding, either low arabinoxylan content or a high activity of xylosidase, which degrades arabinoxylan, are important targets. But for technical reasons, arabinoxylan content was only measured in a subset of all barley samples and thus was not considered in the current study. All the barley parameters that are presented here are presented as percentages (dry matter).

Malt quality parameters

Numerous malt quality parameters are used by maltsters and brewers. Malt extract percentage, malt friability, ratio of wort soluble protein, diastatic power, and wort viscosity are some of the most important parameters contributing to malting quality [47]. For example, when the extract is high, there is more soluble material and less husk and protein. Malt extract contains all soluble compounds and fermentable sugars; which are the most important compounds -around 80%- of malt extract. Fermentable sugars determine the amount of alcohol that can be made from one ton of grain. The higher the extract level, the more alcohol can be produced. Malt extract is measured by malting the grain and measuring the amount of soluble sugar (like glucose and maltose) in the wort (hot water mixed with ground malt). The measurement is given as a percentage (dry basis) [48].

Though Friability is only a mechanical property of malt grain; it is usually the first one which is measured to assess malt quality. It is measured through a quick and simple method, which usually gives the first information on the degradation of the barley structure and thus the transformation into malt. Friability is given as a percentage of the dry matter [48].

Diastatic power measures the amount of diastatic enzymes (α -amylase, β -amylase and limit-dextrinase). Diastatic power is measured by the enzymatic activity in the malt. The measurement is expressed in Windisch Kolbach units (WK). Diastatic enzymes convert the

starch of the grain into soluble sugars (or malt extract). The levels of the various diastase enzymes are important in achieving the quality standard. Low diastase levels are associated with a low potential for malt extract. As diastatic enzymes are proteins, their level is directly related to protein content in the grain [49]. The measurement of individual DP enzymes would most likely improve the consistency and predictability of the potential fermentability of malts supplied to brewers [50]. Soluble protein (“% of dry matter”) is the amount of protein or nitrogen in soluble form. It is an important indicator of malt modification. The higher the number, the more highly modified the malt. The balance of protein components in the wort is important because they contribute to beer foam and mouthfeel, beer color and flavor, and yeast metabolism. Problems can arise, however, when the protein level becomes excessive in wort or beer. This level depends on the process, product and variety. But problems might be expected when wort-soluble protein exceeds 5.5%. High levels of protein can lead to too much color development during wort boiling, filtration problems and the risk of haze formation [51].

Wort viscosity measures the gumminess of wort relative to water. It is essentially a measurement of the amount of stress a plant has undergone during grain filling. It depends on the cultivation site of barley. Genotype (hereditary constitution) also has a major influence on wort viscosity. Wort viscosity is determined from the time of flow in an Ostwald viscometer at 20 °C. Wort viscosity reflects the content and degradation states of various contributory factors such as β -glucan [52]. The measurement is expressed in (mPa.s). Highly viscous malt slows down the separation of the sugar-rich wort from husks during brewing. This slows the amount of beer processed in a brewery each day and increases production costs [49].

2.2 Malting operating conditions

For the malting process, only grains above 2.5 mm were to be converted. At this step, germination capacity was controlled and only grains which had a germination capacity above 95% were selected for malting. The barley grains were then analysed for barley moisture, protein, β -glucan content, according to EBC methods (Analysis by European Brewery Convention 2010). Once the barley analysis was performed, a micro-malting (650g) was run on all samples in the French Institute of Brewing and Malting (IFBM). Steeping was carried out at 18°C for 31-46 hours (two wet cycles for spring barley and three wet cycles for winter barley). Germination was carried out at 16 °C for 4 days for spring barley and 5 days for winter barley. The kilning diagram was 8 h at 50 °C, 1 h rise to 64 °C, 7h at 64 °C, 1h rise to 80 °C, 3 h at 80 °C, 1 h rise to 85 °C, 3 h at 85°C. Barley malt was then analysed for fine extract, viscosity, soluble protein, diastatic power and friability according to the methods of the 2010 EBC methods.

Table 17 presents the mean value and range of each barley and malt quality factor from the three years of cultivar and for both the winter and spring barley subsets.

The next step was a MIR spectra collection of the 394 barley samples

Table 17 mean and average of barley and malt data from 2009-2010-2011

ID	2009		2010		2011	
	Spring barley	Winter barley	Spring barley	Winter barley	Spring barley	Winter barley
1 Barley moisture (% dm)	13.7 [11.4-14.2]	12.7 [12.8-13]	14.01 [12.4-15.5]	13.3 [11.5-15]	14 [12.2-14.9]	11.9 [10.6-12.5]
2 Barley Protein ratio (% dm)	10.9 [9-12.6]	10 [9.4-11.5]	9.8 [8.7-11.9]	10.1 [8-12.4]	10 [8.9-12.9]	10.6 [8.9-13.2]
3 Barley β -glucan (% dm)	3.83 [2.18-5.95]	3.1 [2.09-4.18]	3.31 [2.18-7.73]	3.29 [2.25-4.99]	3.33 [1.95-5.62]	3.63 [2.7-4.4]
4 Malt Friabilty (% dm)	84.1 [51-98]	76.5 [62-85]	82 [77-98]	83.6 [44-99]	81.8 [64-97]	87.6 [72-99]
5 Malt Extract (% dm)	87.1 [80.3-84.3]	80 [79.4-82.5]	82.1 [76.9-83.7]	80.6 [75.6-82.8]	83.8 [80.4-85.8]	80.8 [79.4-82.8]
6 Malt Soluble Protein (% dm)	4.15 [3.30-5.30]	3.76 [3.25-4.53]	4.08 [3.24-5.17]	3.9 [1.69-5.46]	3.9 [3.15-5]	4.04 [3.39-4.96]
7 Malt Diastase power (WK)	334 [200-510]	339 [270-420]	289 [190-380]	292 [170-480]	239 [140-400]	312 [240-450]
8 Malt Viscosity (mPa.s)	1.61 [1.43-1.97]	1.45 [1.46-1.74]	1.49 [1.44-1.68]	1.59 [1.37-1.89]	1.50 [1.45-2.14]	1.61 [1.47-1.93]

2.3 Spectroscopic analysis

The MIR-spectroscopy requires powder samples. The first step in the spectra collection was milling the samples. To prepare the spectra samples, 10g of barley were milled in a Cyclotec™ 1093 for a 1 min. The grinding was carried out by a high-speed action where the sample was rolled against the inner circumference of a durable grinding surface and passed out through a screen. When the high grinding capacity reached 4 g/second, the recovery of the sample was complete.

Mid-infrared spectra were collected at room temperature with transmittance Fourier Transform InfraRed (FTIR) spectroscopy (PERKIN ELMER Spectrum One FT-IR spectrometer), in the interval (400-4000) cm^{-1} at resolution of 4 cm^{-1} . Every spectrum is the average of 5 spectra with 15 repetitions, and the spectra were corrected against the background spectrum of air.

2.4 Multivariate data analysis

The ACP is used to find a linear subspace which passes near the dataset: orthogonal projection of principal components T (latent variable) on a linear subspace of lower dimension M. The components Z represent the co-ordinates of the projection of X in this subspace of dimension M. This method is used quite often as preprocessing for extracting characteristics and visualizing them. The matrix X is decomposed into a structural part and an error part. The structural part contains the score-matrix (T) and the transposed loading-matrix (P^T) equation

$$X = T \times P^T + E \quad (1)$$

Scores are the relationship of PC to the samples (data rows); Loadings are the relationship of PC to the variables (data columns). The new variables are represented in decreasing order of importance, the first PC explains as much of the variation in the dataset as possible, thus the maximum variance is explained with only a few PC, and every PC is orthogonal to the others. The decomposition goes on until the maximum variance is explained. Data analysis was performed by MATLAB® version 7 (R2010a).

3. Results and discussions

The data used to study the barley and malt is shown in Table 1. It was collected from hundreds of samples over three years of harvest. The results are presented as average values and ranges for the spring and winter barley varieties as well as for the three years of cultivation (2009, 2010 and 2011). The large dataset was first analyzed for barley quality factors and then for malt quality factors. At this step, it is important to guarantee that the dataset describes simultaneously the standard raw materials as well as a wide variety of properties in order to obtain a pertinent multivariate analysis.

3.1 Barley quality analysis

As malting barley differs according to the year of harvest, the cultivar species, and the number of rows, these qualitative factors were first studied separately.

Firstly, a sample set from 2009 shows considerable variation in barley quality factors especially with the spring species from the other years. From the averages and ranges in Table 17, it is obvious that the 2009 spring varieties differ from the 2010 and 2011 spring varieties in having an overall higher malt quality index (higher protein content, β -glucan). Our test did not reveal any difference at 10% confidence level (data not shown). Secondly, for the same year of harvest, barley can be distinguished according to barley species. The 2009 spring barley presents higher barley quality in comparison to the 2009 winter barley. This difference is less pronounced in the 2010 barley; 2010 spring barley has higher barley moisture, β -glucan and lower Protein than the 2010 winter barley, while the 2011 spring barley has lower protein, β -glucan and higher moisture than the 2011 winter barley. Finally, the barley data was compared according to species. For 2009, 2-row winter barley was closer to 2-row spring barley than to 6-row winter barley, whereas in 2010, 2-row winter barley was closer to the 2-row spring barley in terms of malt quality factors, but not in terms of barley quality factors. It is more difficult to distinguish barley species for all harvest years.

Similarly, the barley quantitative index can be discussed as well. All barley quality indices show that the experimental dataset presents a large variation in barley quality parameters, including value, out of the usual acceptable quality range. It is thus a suitable base which provides the most pertinent information for studying the effect of barley composition on barley MIR spectra.

Barley moisture values presented in Table 17 lie in the 10.6 – 15.5% dm range while most values are concentrated in the 13-14%dm interval.

Protein content for the brewing barley specification is assumed to lie in the interval 9-11.5%. In the studied dataset, experimental protein content varies between 8-13.2 %. β -glucan is one of the major components affected by the biochemical transformations during the malting process. It is expected to be between 3 and 5% dm. Experimental β -glucan content in barley is presented in table 1???. It shows a large variation range for β -glucan content between 2 and 8 %dm.

3.2 Malt quality analysis

In addition to the measurement of the barley quality index, the samples were malted according to standard steeping, germination and kilning conditions as presented before. The resulting malt quality factors such as friability, extract content, soluble protein content, diastatic power, and viscosity are presented in table 1. Malt with good quality is considered as follows: at least 80% extract, soluble protein content lower than 5%, diastatic power about 300 WK. Malt wort viscosity should be less than 1.65 mPa.s. Malt friability should be at least 80-85%.

However, the experimental results are in the following range: friability ranges from 44% up to 99% values. Extract values show extract contents in the range 75.6- 85.8%. Soluble protein ranges between 1.69% and 5.30%. Experimental diastatic power values vary from 140 to 510 WK. The viscosity of the wort varies between 1.37 and 2.14 mPa.s. The experimental results thus show a wide interval including mainly acceptable malt quality, but also some unacceptable ones. The malt quality dataset also shows the most suitable set for multivariate analysis and modeling.

The 2009 spring barley malt turned out to be better than the winter barley malt in almost all malt quality indices (except for diastatic power). In 2010, the malt extract and the soluble protein from the spring barley are higher than those from the winter barley but friability, viscosity, and diastatic power are lower. In 2011, the malt from the spring barley has higher friability, soluble protein, and viscosity, but has lower malt extract and diastatic power than the winter barley malt.

3.3 Spectra preprocessing

IR spectra contain information which is of analytical interest. However, in complex and multicomponent systems, the extraction of this information remains problematic due to the superposition of spectral information. It requires a rather complex mathematical treatment. Moreover, the information cannot be localized at one given place in the spectrum, but must be distributed on several spectral windows. For this reason, it is necessary to treat the spectra by statistical analysis multivariate methods, also called chemometric methods.

The rough spectral data, such as they are acquired by a spectrometer, must often be subjected to a process of pre-selection in order to eliminate the bad spectra (test of spectral quality).

Then, they must undergo a stage of pretreatment which includes a correction of the physical phenomena instrument changes (lamp, aging, environmental conditions) or sample variability (new genetics, new sample presentation) and experimental disturbances which contribute to spectral information. Preprocessing the raw Mid-infrared (MIR) spectral data is usually indispensable in multivariate calibration in order to detect and/or avoid the influence of such changes. There are various types of preprocessing.

- Derivatives

This makes it possible to reduce the shift of the base line, to separate the absorption bands more clearly and to boost the spectral information. The procedure, based on the convolution of Savitzky and Golay, requires two steps to calculate the derivative on a point: first, a polynomial of k degree is adjusted on at least $k+1$ points of the spectrum around the point, then the derivative of the polynomial at this point is calculated [53].

The original 3601 band dataset was used because smoothing is part of the processing. In the Savitzky-Golay derivative procedure, a first-order polynomial was fitted to spectral windows of 13 band widths. The derivative of this polynomial was assigned as the new value of the central band. The 2nd derivative was applied many times in remote sensing and spectroscopy, for example, to eliminate background signals and to differentiate overlapping signatures. The second derivative was calculated from the 1st derivative spectra. First derivatives result in less noisy spectra than higher-order derivatives directly applied to the original data, so the second derivatives were calculated as a derivative of the first derivative with identical settings [54].

- Smoothing

Smoothing removes short-term variations, or "noise", to reveal the important underlying unadulterated form of the data. Igor's smooth operation performs box 'moving average', binomial, and Savitzky-Golay smoothing. The different smoothing algorithms convolve the input data with different coefficients. This simply replaces each data value with the average of neighbouring values [55].

- Detrending

Toward the big wavelengths, the values of the absorption have a tendency to increase because of the effects of diffusion. According to the pressure of optical fiber on the sample, this tendency becomes nonlinear. The method of trending is a simple method of calculation which removes the curve of the second degree of the base line. A polynomial of the second degree is adjusted to the corrected spectrum. This treatment is often associated with a standardization method such as Standard Normal Variate (SNV)

- Spectra normalization

There are various kinds of normalization.

- I. Multiplication scatter correction (MSC)

MSC improves the linearity relation between the absorption and the concentration. To perform MSC correction, a reference spectrum is needed and then a linear model is established between the spectrum and the reference spectrum.

II. Standard Normal Variate (SNV)

SNV was developed by Barnes, *et al.* (1989) to remove multiplicative interferences of scatter and particle size, and to account for the variation in baseline shift and curvilinearity in diffuse reflectance spectra. Standard normal variate, also known as z-transformation or as centering and scaling, normalizes each spectrum to zero mean and unit variance by subtracting the mean of this spectrum and dividing the difference by its standard deviation.

In this study, spectra were pre-processed according to the following techniques: data centering, SNV normalization, Savitzky-Golay smoothing [3, 13], 1st and 2nd derivative. The raw or pre-processed MIR-spectrum theoretically contains information about all molecules vibrations. Due to the complexity of the barley composition, signals might interfere with each other. The whole spectrum was thus divided into smaller subintervals to focus the PCA analysis on elementary vibration zones. The width of the subintervals varied from 100 cm⁻¹ up to 400 cm⁻¹. For each subinterval, several pre-treatments were considered, as shown in Table 18. In this paper, only the best intervals are presented. The choice of the best intervals was made according to two criteria: explained percentage of total variation, and the minimum number of principal components. The approach is assumed to provide the most relevant spectral interval.

Table 18 the best intervals and spectra pretreatments

	Original spectra		Spectra mean centred & SNV normalisation		Spectra mean centred & first derivative		Spectra mean centred & second derivative	
	Compenets number	Explained variance	Compenets number	Explained variance	Compenets number	Explained variance	Compenets number	Explained variance
Interval 100 cm-1								
800-900	16	20%	12	40%	10	42%	4	66%
900-1000	15	25%	13	35%	11	40%	4	65%
1000-1100	14	27%	12	43%	12	44%	4	72%
1100-1200	16	32%	11	44%	10	45%	4	70%
Interval 200 cm-1								
800-1000	15	30%	12	40%	7	79%	3	80%
1000-1200	12	44%	11	45%	6	77%	3	80%
Interval 300 cm-1								
800-1100	11	54%	10	55%	8	77%	4	80%
1100-1400	10	60%	10	59%	6	70%	3	82%
Interval 400 cm-1								
800-1200	10	58%	9	57%	7	77%	2	85%

From the large combination of subinterval width and preprocessing methods, it turned out that the most relevant interval is 800-1200cm⁻¹. This agrees with the subinterval which is found in the literature. The FT-IR spectrum of a peptic polysaccharide sample shows characteristic absorbance in the region between 1200 and 850 cm⁻¹, which are within the so-called “fingerprint region”. The position and the intensity of the bands are specific for every polysaccharide [57][59].

3.4 PCA analysis

PCA multivariate tool was first used to evaluate the exploratory data, PCA was carried out on the whole dataset, including all-year samples with original and normalized data. Then, mid-infrared measurements of spectra were studied and finally a combination of both exploratory data and spectra was used to distinguish barley samples according to qualitative and quantitative parameters.

3.4.1 Global PCA analysis for malting barley and malt quality

In order to look for the pertinent information which might be *hidden* in such a large exploratory database while avoiding correlation between barley and malt factors, PCA was considered. PCA was firstly carried out on the whole dataset including all-year samples with original or normalized data. With one single principal component, PCA explained 99% of the total variation for the original data. Three principal components were needed when using normalized data. However, the score plot of the first and the second principal components did

not revealed any difference between spring and winter barley (results not presented). PCA was then performed on samples from each harvest year individually.

A PCA was then performed on the malting quality data, taking into account 8 quality parameters for the 394 available samples. Despite the fact that PCA is a powerful tool that can unravel large datasets with only a few orthogonal principal components, no satisfactory results were found. Therefore, the analysis was applied leaving one factor out of the list. The results are presented in Table 19, which shows the number of factors used the left-out factor, the number of principal components, and the explained variation (%). The classification results are presented according to two factors: (1) qualitative factors (species and site of cultivation), (2) quantitative factors (the 8 barley and malt quality indices). PCA was performed both on original data and normalized data in the range [0-1]. The PCA performed on original data showed that one single principal component explained more than 90% of the total variance. A clear difference between spring and winter barley was obtained. The classification by quantitative factors was successful. However, it could not identify the barley by site of cultivation

Table 19 Classification capacity of PCA on normalized data with all parameters and all but one parameters. (+) parameter with significant influence (0) parameter without effect on barley classification, (ND) parameter with unclear effect

Parameters Number	Excluded Parameter (s)	PC Number	Explained Percentage	Loading Influencing Parameters	Qualitative Factors		Barley Classification by Quantitative Factors							
					Specie	Site	Moist	Prot	BG	Fria	Extr	DP	Psol	Visco
8	-	4	84%	All parameters	ND	0	0	+	0	+	+	+	0	0
7	Friability	3	80%	Barley Moisture Barley β -glucan Barley Protein Malt diastatic power	0	+	+	ND	0		0	+	ND	ND
7	Diastatic power	4	88%	Barley moisture Barley β -glucan Barley protein Malt soluble protein	0	+	+	0	+	+	0		+	0
6	Friability Diastatic power	3	85%	Barley moisture Barley β -glucan Barley protein Malt soluble protein	0	+	+	+			ND		0	0
3	Malt parameters	3	99%	Barley Moisture Barley β -glucan Barley protein	0	+	+	+	0					

It was shown that normalized data required more principal components (2-4) to explain > 80% of the total variance. No separation between winter and spring barley was possible. On the contrary, normalized data allowed for distinguishing barley according to the site of cultivation. Moreover, classification of barley by more quantitative factors (i.e. moisture, extract, friability, etc.) was possible.

3.4.1.1 PCA on the barley quality factors excluding friability and/or diastatic power

PCA analysis was performed on a matrix of 168 observations and 7 factors of barley from the 2010 harvest. PCA was performed, first excluding the friability factor, then the diastatic power. Similar results were obtained in both cases. One example for PCA analysis on original and normalized 2010 data without friability is presented here.

Using original data, one principal component explained 90% of the total variance. The score plot of the first two principal components reveals a clear distinction between spring and winter barley. Figure 34 shows spring barley in the lower part of the plot. The corresponding loading plot for PC1 and PC2 is presented in Figure 34b. The dominant parameters are Diastatic Power which is located at the right side of the plot. Extract is in the lower part of the plot. When analyzing both scores and loadings, it is shown that barley samples with lower extract content are located in the upper part of the scores plot (Figure 34a). Thus, the 2010 spring barley representing good extract content is located in the lower part of the plot. In the same manner, barley samples with higher DP are located on the right part of the score plot, while samples with lower DP content are represented on the left.

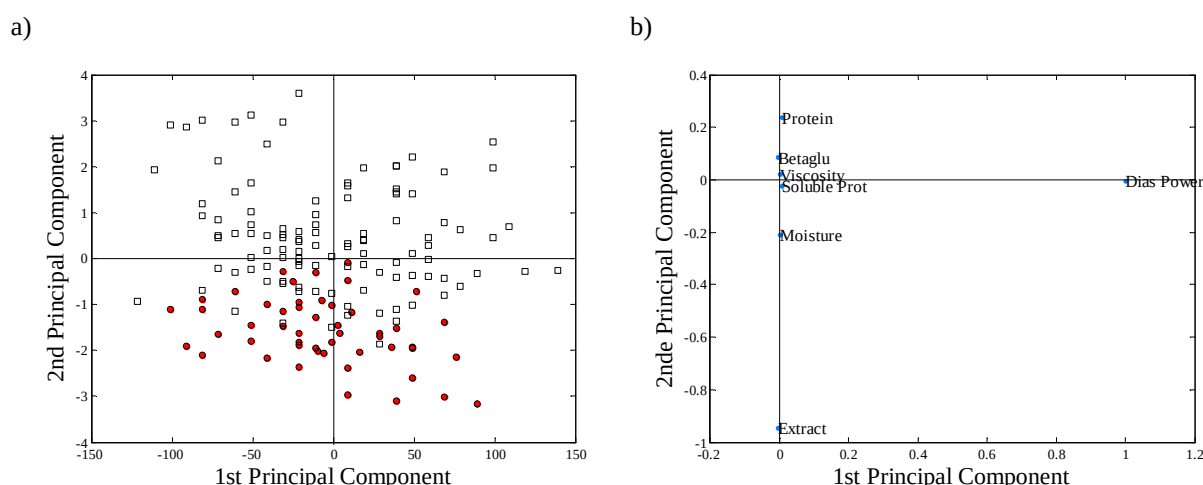


Figure 34 PCA plot on barley exploratory data from 2010 parameters without friability. a) Score plot □ winter barley ● spring barley. b) Loading plot. PCA was applied on original data. With 1 principal component 99% of variance total was explained.

When considering normalized data, three principal components were needed to explain 80% of the total variance. Figure 35a shows that the 2010 spring barley was separated according to the site of cultivation. PCA was also applied to the 2011 samples (136 observations with 7 factors), and on the 2009 barley (90 observations with 7 factors). The analysis is not presented here, but it showed the same trend for sample classification.

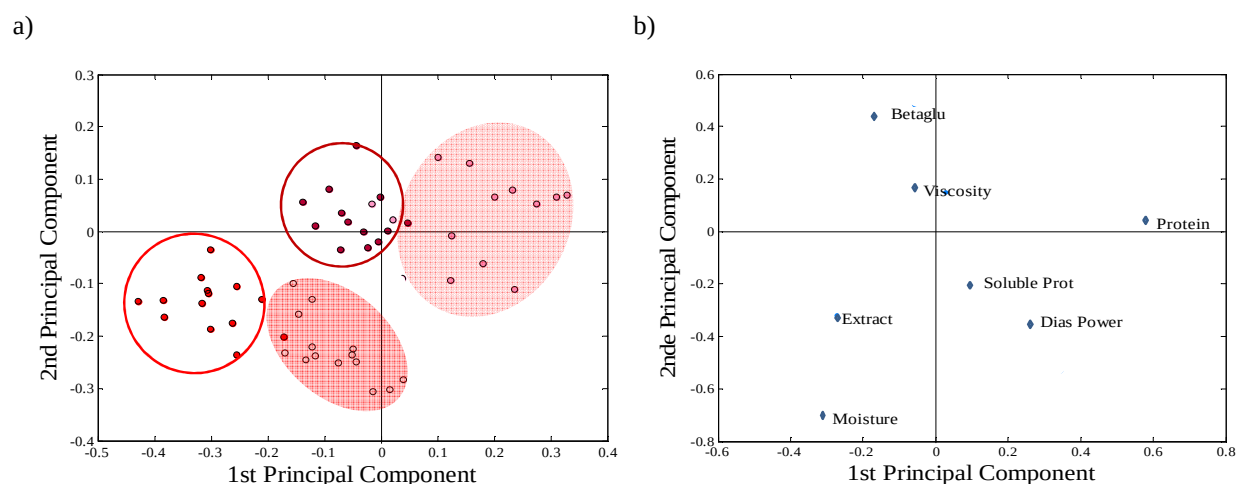


Figure 35 PCA score plot and loading plot on spring barley exploratory data from 2010. Parameters without friability. a) Score plot site A ● site B ● site C ● site D ○ b) loading plot. PCA was applied on normalized data [0-1]. With 3 principal components 80% of variance total was explained

The PCA analysis of the original data for each year, excluding friability and diastatic power, revealed that, in addition to the possibility of classification according to extract content, protein, moisture content and species, the 2010 spring barley could be distinguished according to the site of cultivation the classification is highlighted in Figure 36a. In this case, three principal components explained 90% of the total variance. The same results were found for the 2009 and the 2011 barley samples.

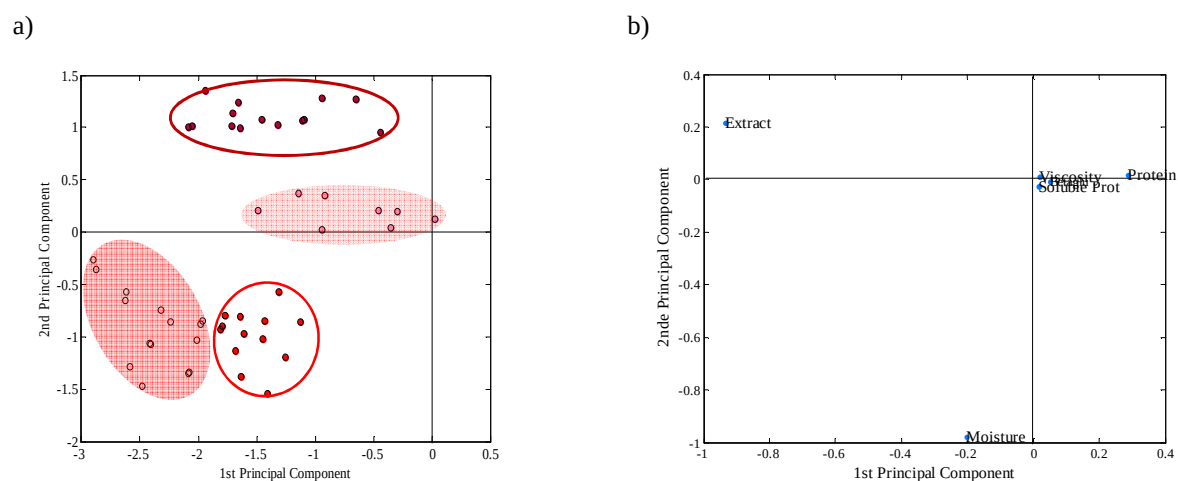


Figure 36 PCA a) score plot present barley classification by site of cultivation: Site A ● site B ● site C ● site D ○ b) Loading plot on spring barley exploratory data from 2010 parameters without friability and diastatic power. PCA was applied on original data. With one principal component 90% of variance total was explained.

PCA analysis applied to original exploratory data leads to the following conclusion:

1. PCA analysis for each year, excluding friability as a variable, allowed for distinguishing barley according to species, extract, and diastatic power.
2. PCA analysis for each year, excluding the diastatic power variable, made it possible to distinguish barley samples according to species, extract and friability content.
3. PCA performed without friability and diastatic power showed barley classification according to moisture, extract, protein and species.

Furthermore, with normalized data of the 2010 spring barley, the classification according to the site of cultivation was possible.

A classification of barley samples was possible according to qualitative factors as species and quantitative factors as extract by PCA analysis on exploratory data. However, the differences in the barley quality index year after year cannot be related to the changes in malt quality and do not allow for adapting the operating conditions to monitor the malting process. It is therefore interesting to evaluate if the multivariate analysis on the barley MIR-spectra allows for the same distinction/classification and could emphasize the yearly variation.

3.4.2 Global PCA analysis for barley-spectra

A PCA was initially performed on the whole 394 MIR spectra, i.e. in the interval $[400-4000]$ cm^{-1} . The second and third principal components explained, respectively, 87% and 90% of the total variance. The score of the first two components is shown in Figure 37. The 2009 barley samples appear in one isolated group on the right bottom part of the plot (blue zone). They appear to be very different from the 2010 (red zone) and 2011 (green zone) samples. The 2010 and 2011 samples could not be clearly separated on the 2D-PCA plot.

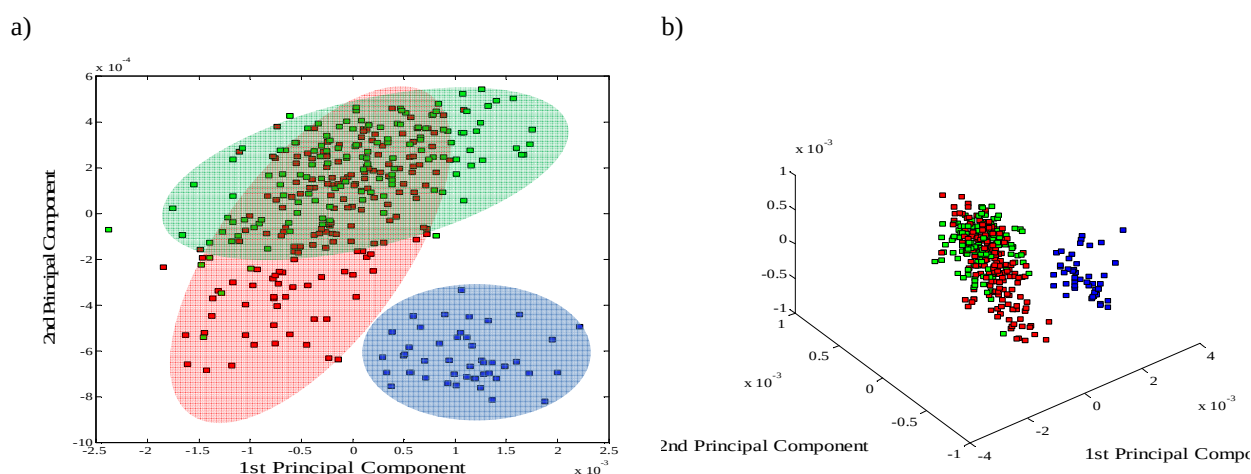


Figure 37 PCA a) score plot on barley spectra in 2D b) score plot in 3D 2009 harvest ■ 2010 harvest ■ 2011 harvest ■. PCA was performed on second derivative spectra in the interval $[800-1000]$ cm^{-1} . PCA was applied on original data. Two principal components explain 87% of total variance

The score distribution of the third principal component is presented in Figure 37b. The use of the first 3 principal components did not resolve the superposition of the 2010 and 2011 samples. Such similarity could be explained by the climatic conditions. Climatic data confirmed that the weather conditions in 2010 and 2011 were dominated by stormy rain at the end of the winter season and at the beginning of spring season, while the 2009 climate was normal with a regular rain period. The variation in the barley quality could therefore be associated to the climate conditions.

3.4.3 PCA analysis of MIR spectra to classify barley by species

In order to simplify the trends and emphasize the differences in the spectra exploration, PCA was performed on the samples for each year separately.

PCA was performed on the 2009 barley spectra (14 winter barley and 47 spring barley). 91% of the total variance was explained with two principal components. The score plot of the first and second principal components reveals a relatively clear difference between winter and spring barley as present in Figure 38a. The first principal component allows us to separate winter barley (on left plane) from spring barley (on the right plane).

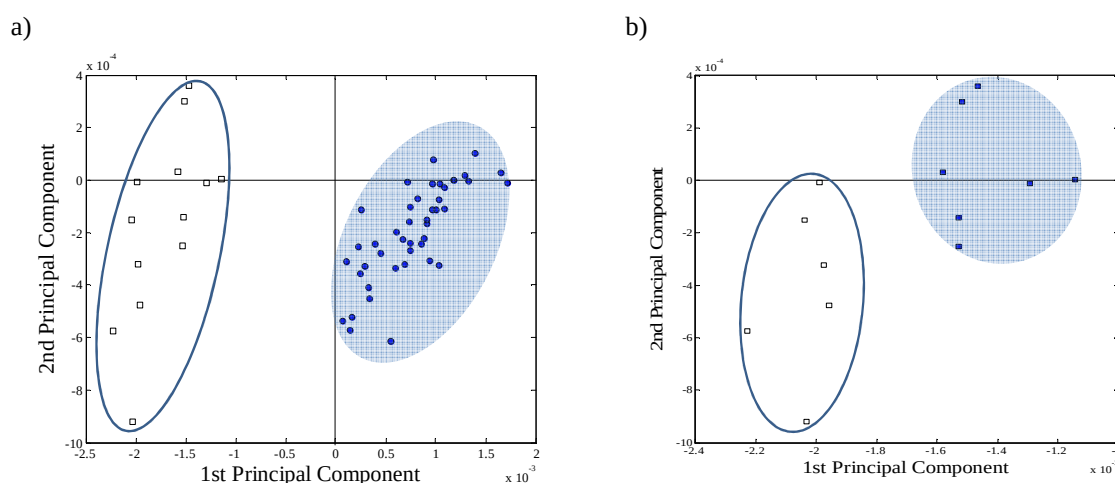


Figure 38 a) PCA score plot on barley 2009 distinguishing $\square$ winter barley $\bullet$ spring barley. Two principal components explain 91% of the total variance.

b) PCA score plot on winter barley 2009 distinguishing $\blacksquare$ 2-row and $\square$ 6-row barley. One principal component explains 83% of the total variance.

All PCA were performed on original data in the interval 800-1000 cm^{-1} with second derivative spectra

The score plot corresponding to the 2010 spectra PCA is presented in Figure 39a. PCA analysis was performed on 168 barley spectra (56 spring barley, 112 winter barley). The first two principal components explain 85% of the total variance. The spring barley is situated in the upper left part of the plot. From the exploratory data, it was pointed out that the 2010 winter barley had higher moisture, β -glucan, extract, and soluble protein content, and lower protein, friability, diastase power, and viscosity. In addition, quality indices for the 2010 winter barley were more spread out than for the 2010 spring barley. This indicates that the yearly growing conditions for winter barley induce larger variation in barley quality as

concluded by Nielsen et al (2003). Spring barley has relatively lower protein, friability, diastase power, viscosity and higher β -glucan, extract, and soluble protein.

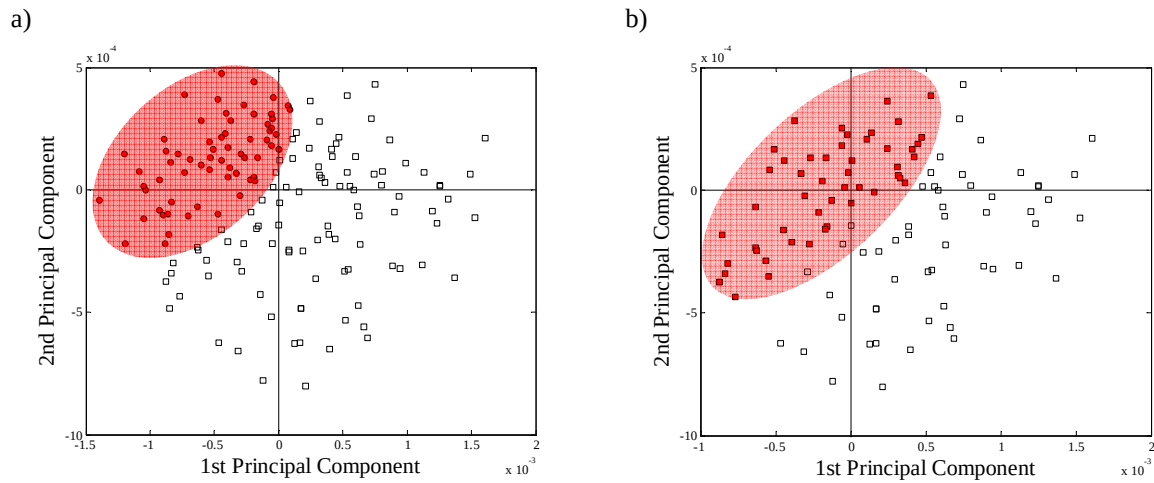


Figure 39 a) PCA score plot on barley 2010 $\square$ winter barley $\bullet$ spring barley. Two principal components explain 89% of the total variance.

b) PCA score plot on winter barley 2009 $\blacksquare$ 2-row $\square$ 6-row. Two principal components explain 89% of the total variance. The PCA was realized on original data in the interval 800-1000 cm^{-1} . With second derivative spectra

The score plot of the PCA on the 136 barley spectra of 2011 is presented in Figure 40a. The number of 2011 spring barley spectra (124) is much higher in comparison to the number of 2011 winter barley (12) spectra. The analysis for winter barley might be less significant. The score plot reveals that there is a difference between the two species of 2011; winter barley is located in the lower left part of the plot.

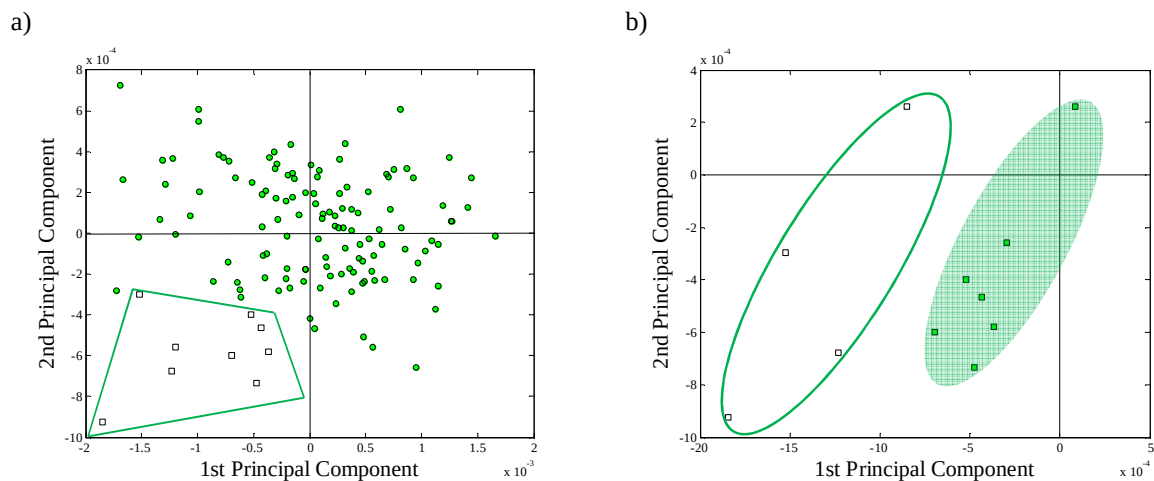


Figure 40 a) PCA score plot on barley 2011 $\square$ winter barley $\bullet$ spring barley. Two principal components explain 87% of the total variance

b) PCA score plot on winter barley 2009 $\blacksquare$ 2-row $\square$ 6-row. Two principal components explain 89% of the total variance
 . The PCA was realized on original data in the interval 800-1000 cm^{-1} . With second derivative spectra

Further investigation was carried out to study the influence of the number of rows on the MIR-spectra. As spring barley is 2-row only, the effect of rows was studied using winter barley only. It was difficult to distinguish any effect when analyzing the three years of samples together. Each year was studied individually. Thus, for 2009, the dataset contains 14 winter barley (6 samples of 6-row specie and 8 samples of 2-row specie). One principal component explained 83% of the total variance (Figure 38b). For the 2010 samples, PCA was applied on 112 spectra (64 samples of 6-row species and 48 samples of 2-row species). Results are presented in Figure 39b. With two components, PCA could explain 89% of the total variance. The 6-row barley was relatively different from the 2-row barley. PCA analysis was carried out in the same way on the winter barley 2011. There were few samples, only 12 were available (4 6-row samples and 8 2-row samples). The score plot reveals, however, a clear difference between 2-row and 6-row, which is presented in the left part in Figure 40b.

The influence of species and rows was thus successfully studied. The PCA results on MIR spectra confirmed results obtained from PCA on exploratory data. Barley could be distinguished according to species. MIR analysis thus contains essential information to characterise barley.

3.4.4 Conclusion on the PCA analysis

In this study, PCA was shown to be an efficient tool for classifying barley, using either exploratory or spectral data. It allows for distinguishing barley species (spring/winter barley or 2-row, 6-row barley), the harvesting year and, to some extent, the site of cultivation by using the second derivative of the spectra in the subintervals [800-1000] and [1000-1200] cm^{-1}

It could thus be expected that PCA could also be used to classify either barley or malt quality indices to offer maltsters an original tool for barley characterization. PCA analysis on the MIR-spectra was thus performed to classify samples according to three quantitative quality indices of barley: moisture content, protein content, and β -glucan content. Pre-processing and spectral interval were the same as in the previous study. However, no satisfactory classification according to these quantitative barley indices could be identified. Similarly, PCA was carried out in order to emphasize possible correlations between malt quality factors and the barley spectra. The malt quality factors used for this investigation were friability, extract, viscosity, soluble protein, diastatic power. No correlation could be drawn between the quantitative malt factors and the MIR spectra.

5. Conclusion

A large dataset of 394 barley samples consisting of 138 samples of winter barley and 256 samples of spring barley, harvested over three years at 16 sites of cultivation in France was studied. For each sample, 3 barley parameters, namely protein content, moisture content, β -glucan content were measured. The MIR-spectra of milled whole barley grains were collected in the 400-4000 cm^{-1} interval. After a standard malting process, the corresponding malt quality parameters such as friability, extract, soluble protein, diastatic power, viscosity were measured. Many pre-processing methods of the MIR spectra were considered and the spectral interval was divided into subintervals to identify the most relevant signal in the spectra that could be related to barley or malt quality. On the best subintervals coupled to the best preprocessing method, PCA was used to reduce this complex information to a few latent variables, using the second derivative of the spectra in the subintervals [800-1000] and [1000-1200] cm^{-1} .

PCA analysis was first performed on barley exploratory data, and then on barley MIR-spectra. Both PCA showed that the 2009 barley had different properties when compared to the 2010 and 2011 barley. The two-by-two parameter means (2009-2010/2009-2011/2010-2011) was compared using Student t-test. 2009 barley showed to be different from the 2010 and 2011 barley. In all cases, spring barley was clearly separated from winter barley for each year of harvest. Barley variations according to the number of rows were also distinguished. PCA is thus a powerful tool for barley classification according to qualitative indices such as year of harvest and barley species (spring/winter barley, 6/2 row).

Then, PCA analysis was carried out to correlate MIR-spectral data to barley or malt quantitative indices such as protein or friability. Though PCA analysis on exploratory data was successfully used to characterize the barley samples according to barley indices such as moisture, protein, content and malt indices such as friability, diastatic power, PCA on MIR spectra using the same pre-processing and the best spectral intervals did not provide satisfactory classification according to the same quantitative barley/malt indices.

In this work, spectra were collected from whole barley grain. As the structure of the grain involves many functional compartments (embryo, aleurone layer, albumen,...), it was assumed that PCA analysis on spectra collected from the different compartments of barley grain could better reveal the key barley characteristics. Focusing the analysis on the key compartments could indeed remove undesired interactions between barley compounds and improve PCA performance in barley classification according to quantitative indices.

Acknowledgement

Authors are grateful to Ms. T. Wagner for the proofreading

Bibliography chapter 5

- [1] D.E. Briggs, J.S. Hough, R. Stevens, *Malting and brewing science* 2nd edition. London, New York, 1981.
- [2] A. Hariri, *Etude et modélisation de la trempe en malterie*, Thesis de l'Institut National Polytechnique de Lorraine, 2003.
- [3] J.E. Holmberg, J.J. Hamalainen, P. Reinikainen, J. Olkku, A model for predicting the effects of a steeping program on the germination of barley with different water sensitivities, *J. Indus. Brew.* 108 (2002) 416-423.
- [4] O. Grujic, D.J. Pejin N.M. Pruzulj, The effect of winter barley variety and technological factors during malting on malt quality, *APTEFF* 36 (2005) 33-42.
- [5] E. Bravi, O. Marconi, G. Perretti, P. Fantozzi, Influence of barley variety and malting process on lipid content of malt, *J. Food Chem.* 135 (2012) 1112-1117.
- [6] I.P. McTaggart, K.A. Smith, The effect of fertiliser and soil nitrogen on the overall uptake of nitrogen in the plant, and the grain nitrogen content of spring-sown malting barley, *HGCA Project Report* 46 (1992).
- [7] M.D. Fleury, M.J. Edney, L.D. Campbell, G.H. Crow, Total, water-soluble and acid-soluble arabinoxylans in western Canadian barleys, *J. Plant Sci.* 77 (1997) 191-196.
- [8] T.J. March, D. Richter, T. Colby, A. Harzen, J.R. Schmidt, K. Pillen, Identification of proteins associated with malting quality in a subset of wild barley introgression lines, *Proteomics* 12 (2012) 2843-2851.
- [9] J. Peltonen, H. Rita, R. Aikasalo, S. Home, Hordein and malting quality in northern barleys, *Hereditas* 120 (1994) 231 -239.
- [10] G.P. Fox, K. Onley-Watson, A. Osman, Multiple linear regression calibrations for barley and malt protein based on the spectra of hordein, *J. Inst. Brew.* 108 (2002) 155-159.
- [11] J. Gunkel, M. Voetz, F. Rath, Effect of malting barley variety on the fermentability, *J. Inst. Brew.* 108 (2002) 355-361.
- [12] G.S. Chandra, M.O. Proudlove, E.D. Baxter, The structure of barley endosperm-An important determinant of malt modification, *J. Sci. Food Agri.* 79 (1999) 37-46.
- [13] U.R. Holopainen, A. Willhelmson, M.S. Marttila, P.P. Sainio, A. Arirajala, P. Reinikainen, E. Kotaviita, H. Simolin, S. Home, Endosperm Structure Affects the Malting Quality of Barley, *J. Agri. Food Chemist.* 53 (2005) 7279-7287.
- [14] H.F. Seefeldt, F.V. Berg, W. Kockenberger, S.B. Engelsen, B. Wollenweber, Water mobility in the endosperm of high β -glucan barley mutants as studied by nuclear magnetic resonance imaging, *Magnet. Res. Image.* 25 (2007) 425-432.
- [15] P. Zapotocny, M. Zielinska, Z. Nita, Application of image analysis for the varietal classification of barley: Morphological features, *J. Cereal Sci.* 48 (2008) 104-110.
- [16] D. Cozzolino, K. Alder, S. Roumeliotis, J. Eglinton, Feasibility study on the use of multivariate data methods and derivatives to enhance information from barley flour and malt samples analyzed using the Rapid Visco Analyzer, *J. Cereal Sci.* 56 (2012) 1-5.
- [17] S. Jacobsen, I. Sondergaard, B. Moller, T. Desler, L; Munk, A chemometric evaluation of the underlying physical and chemical patterns that support near infrared spectroscopy of barley seeds as a tool for explorative classification of endosperm genes and gene combinations, *J. Cereal Sci.* 42 (2005) 281-299.
- [18] L. Marte, P. Belloni, E. Genorini, V. Sileoni, G. Perretti, L; Montanari, O. Marconi, Near-infrared reflectance models for the rapid prediction of quality of brewing raw materials, *J. Agri. Food Sci.* 57 (2009) 326-333.
- [19] V. Sileoni, G. Perretti, L. Marte, O. Marconi, P. Fantozzi, Near-Infrared Spectroscopy for Proficient Quality Evaluation of the Malt and Maize Used for Beer Production, *J. Inst. Brew.* 116 (2010) 134-140.

- [20] A.R. Barr, R. Roumeliotis, S.J. Logue S.P. Jefferies, The heritability of malt extract predicted using NIR spectroscopy. Australian Barley Technical Symposium online, 1999.
- [21] A. Tarr, D. Diepeveen, R. Appels, Spectroscopic and chemical fingerprints in malted barley, *J. Cereal Sci.* 56 (2012) 268-275.
- [22] R. Davis, and L.J. Mauer, Fourier transform infrared (FT-IR) spectroscopy: A rapid tool for detection and analysis of foodborne pathogenic bacteria. Book 2 (2010) 1582-1594. Online published by: App. microbiology and microbial biotechnology © Formatex.
- [23] B. Smith, Fundamentals of Fourier Transform Infrared Spectroscopy. Second ed. CRC Press, Washington, 2011.
- [24] D.W. Lachenmeier, Rapid quality control of spirit and beer using multivariate data analysis of Fourier transform infrared spectra, *J. Food Chemist.* 101 (2007) 825-832.
- [25] R. Lleti, E. Meléndez, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, Outliers in partial least squares regression application to calibration of wine grade with mean infrared, *J. Analytica Chimica Acta* 544 (2005) 60-70.
- [26] E. Polshin, B. Aernouts, W. Saeys, F. Delvaux, F.R. Delvaux, D. Saison, H. Hertog, B.M. Nicolai, J. Lammertyn, Beer quality by FT-IR spectroscopy: impact of measurement strategies, data pre-processing and variable selection algorithms, *J. Food Eng.* 106 (2011) 188-198.
- [27] M.D. Guillén, N. Cabo, Infrared spectroscopy in the study of edible oils and fats, *J. Sci. Food Agri.* 75 (1997) 1-11.
- [28] M. Safar, D. Bertrand, P. Robert, M.F. ,Devaux, C. Genot, Characterization of edible oils, butters and margarines by Fourier transform infrared spectroscopy with attenuated total reflectance, *J. American Oil Chemists' Society* 71 (1994) 371-377.
- [29] N. Koca, N.A. Kocaoglu, W.J. Harper, L.E. Rodriguez, Application of temperature-controlled attenuated total reflectance-mid-infrared (ATR-MIR) Spectroscopy for rapid estimation of butter adulteration, *J. Food Chemist.* 121 (2010) 778-782.
- [30] P. Robert, M. Malanie, C. Barron, F. Guillon, L. Saulnier, FT-IR investigation of cell wall polysaccharides from cereal grains. Arabinoxylan infrared assignment, *J. Agri. Food Chemist.* 53 (2005) 7014-7018.
- [31] S. Philippe, P. Robert, C. Barron, L. Saulnier, F. Gullon, Disposition of cell wall polysaccharides in wheat endosperm during grain developing: Fourier transform-infrared microscopy study, *J. agri. Food Chemist.* 54 (2006) 2303-2308.
- [32] K. Pearson, K., On lines and planes of closest fit to systems of points in space, *Philosophical Magazine*, 2(1901) 559-572.
- [33] H. Hotelling, Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24 (1933) 417-441, 498-520.
- [34] D.M. Gottlieb, J. Schultz, S.W. Bruun, S. Jacobsen, I. Sondergaard, Multivariate approaches in plant science, *Photochemistry* 65 (2004) 1531-1548.
- [35] L. Munck, B. Moller, A new germinative classification model of barley for prediction of malt quality amplified by near infrared transmission spectroscopy calibration for vigour' 'on line'' both implemented by multivariate data analysis, *J. Inst. Brew.* 110 (2004) 3-17.
- [36] C. Cordella, I. Moussa, A.C. Martel, N. Sbirrazzuoli, L. Lizzani-Cuvelier, Recent developments in food characterization and adulteration detection: Technique-oriented perspectives, *J. Agri. Food Chemist.* 50 (2002) 1751-1764.
- [37] D. Downey, Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics, *TrAC Trends in Analytical Chemistry* 17 (1998) 418-424
- [38] J.L. Molina-Cano, T. Ramo, R.P. Ellis, J.S. Swanston, H. Bain, Uribe-Echevrria, A.M. Perzvendrell, Effect of grain composition on water uptake by malting barley: A genetic and environmental study, *J. Inst. Brew.* 101 (1995) 79-83.

- [39] J.P. Nielsen, L. Munck, Evaluation of malting barley quality using exploratory data analysis. I. extraction of information from micro-malting data of spring and winter barley, *J. Cereal Sci.* 38 (2003) 173-180.
- [40] B. Moller, Near infrared transmission spectra of barley of malting grade represent a physical-chemical fingerprint of the sample that is able to predict germinative vigour in a multivariate data evaluation model, *J. Inst. Brew.* 110 (2004) 18-33.
- [41] R. Karoui, G. Downey, C. Blecker, Mid-Infrared Spectroscopy Coupled with Chemometrics: A Tool for the Analysis of Intact Food Systems and the Exploration of Their Molecular Structure-Quality Relationships - A Review. ©American Chemical Society, 2010. In: <http://tin.dstc.uniud.it/~gambi/cr100090k.pdf>.
- [42] M. McClelland, Storage of High Moisture Barley. Alberta Agriculture and Rural Development, Agdex 114 (2008) 61-1.
- [43] G. Zhang, J. Chen, J. Wang, S. Ding, Cultivar and Environmental Effects on (1→3, 1→4)-β-D-Glucan and Protein Content in Malting Barley, *J. Cereal Sci.* 34 (2001) 295-301.
- [44] M. Brophy, Canadian Barley Malting and Brewing Technical Guide. Canadian Malting Barley Technical Centre, 2012.
http://www.cmbtc.com/CMBTC_Site/Technical_Guide_files/Barley%20Guide%202012.pdf
- [45] J. Wang, G. Zhang, J. Chen, F. Wu, The changes of β-glucan content and β-glucanase activity in barley before and after malting and their relationship to malt qualities, *J. Food Chemist.* 86 (2004) 223-228.
- [46] R.B. Vis, K. Lorenz, Malting and Brewing with a High β-Glucan Barley, *LWT, J. Food Sci. Tech.* 31 (1998) 20-26.
- [47] M. Munoz-Amatriaín, Y. Xiong, M.R. Schmitt H. Bilgic, A.D. Budde S. Chao, K.P. Smith, G.P. Muehlbauer, Transcriptome analysis of a barley breeding program examines gene expression diversity and reveals target genes for malting quality improvement, *J. BMC Genomics* 11 (2010) 1-15.
- [48] R. Pahl, Important raw materials quality parameters and their influence on beer production, Brewing Conference Bangkok 2011.
- [49] B. Paynter, Malt quality parameters for malting barley, *Research Agronomist (Barley)*, Farmnote 38/96, 2012.
- [50] D. Evans, L. Chengdao, Eglinton, K. Jason, Improved Prediction of Malt Fermentability by Measurement of the Diastatic Power Enzymes beta-Amylase, alpha-Amylase, and Limit Dextrinase: I. Survey of the Levels of Diastatic Power Enzymes in Commercial Malts, *J. American society brew. Chemists*, 66 (2008) 223-232.
- [51] P. Schwarz, H. Horsley, A Comparison of North American Two-Row and Six-Row Malting Barley. The brewer's market guide (BMG), 1997. In: <http://morebeer.com/brewingtechniques/bmg/schwarz.html>
- [52] J.S. Hough, D.E. Briggs, R. Stevens, T.W. Young, *Malting and brewing Science: Hopped Wort and Beer*, 2ed edition. Aspen Publishers, Inc. 2 (1999) 811.
- [53] A. Savitzky, M.J. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Journal of Analytical Chemistry* 36 (1964) 1627-1639.
- [54] H. Buddenbaum, M. Steffens, The Effects of Spectral Pretreatments on Chemometric Analyses of Soil Profiles Using Laboratory Imaging Spectroscopy, *App. Environ. Soil Sci.* 2012, 1-12
- [55] A. Candolfi, R. Maesschalck, D. Jouan-Rimbaud P.A. Hailey D.L. Massart, The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra, *J. Pharma. Biomedical Analy.* 21 (1999) 115-132
- [56] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra, *App. Spectr.* 43 (1989) 772-777.
- [57] M.A. Coimbra, A. Barros, M. Barros, N.R. Rutledge, I. Delgadillo, Multivariate analysis of uronic acid and neutral sugars in whole pectic samples by FT-IR spectroscopy, *J. Carbohydrate Polymers* 37 (1998) 241-248.

[58]X. Fang, B. Jiang, X. Wang, Purification and Partial Characterization of an Acidic Polysaccharide with Complement Fixing Ability from the Stems of *Avicennia Marina*, *J. Biochemist. Molecul. Biology*, 39 (2006) 546-555.

[59]L.M. Nollet, F. Toldrá, *Advances in Food Diagnostics*. First edition, lackwell Publishing Professional USA, 2007.

Chapitre 6

Middle InfraRed spectroscopy (MIR) and Partial Least Square Regression (PLSR) as promised tool for malting barley quality parameters classification and prediction

Ce chapitre est présenté sous la forme d'un article. Les méthodes d'analyse de l'orge et du malt ont été décrites dans le chapitre 3. Les résultats de cet article ont été établis à partir des échantillons d'orge de 72 variétés d'orges, cultivées dans 16 lieux en France, récoltées en trois années.

Middle InfraRed spectroscopy (MIR) and Partial Least Square Regression (PLSR) as promised tool for malting barley quality parameters classification and prediction

AJIB B.¹, FOURNIER F.¹, BOIVIN P.², SCHMITT M.², FICK M.¹,

⁽¹⁾ Université de Lorraine, LRGP, 2 Avenue de la Forêt de Haye, TSA 40602, 54518 Vandœuvre Cedex

⁽¹⁾ CNRS UMR 7274, LRGP, 2 Avenue de la Forêt de Haye, TSA 40602, 54518 Vandœuvre Cedex

⁽²⁾ IFBM, 7, rue du Bois de la Champelle B.P. 267, 54512 Vandœuvre-lès-Nancy Cedex

Abstract

This paper presents the classification and prediction of barley quality parameters by use of Mean InfraRed spectroscopy (MIR) and Partial Least Square regression (PLS). Several hundreds of barley samples from three years, spring and winter barley, different places of cultivations, and different varieties were studied for three barley quality parameters (protein, β -glucan, arabinoxylan) parameters.

In the first hand, PLS on all years samples together were very difficult to carry out with no significant or weak coefficient of determination R^2 . Therefore, PLS models on MIR spectra were established for each year separately. Excellent models were obtained for barley protein prediction; the best predictions were found in the interval 1300-1700 cm^{-1} ; with $R^2 = 0.75, 0.80, 0.76$ for spring barley 2009, 2010, 2011, respectively. Very good model were found for winter barley 2010 ($R^2 = 0.70$). As far as β -glucan prediction are concerned, very good models for spring barley were found with $R^2 = 0.72, 0.86$ for 2009, 2010, respectively. By cons, 2011 β -glucan prediction was more difficult to be modelled. No significant models were found for arabinoxylans prediction for any years.

This study shows that MIR spectroscopy with partial least square regression is a very powerful tool for prediction of barley quality.

1 Introduction

Breeding barley -from new cultivar, new growing sites, or new varieties- to be used for malting and brewing technologies requires the assessment of many grain quality parameters which contribute to the quality of malt and beer. These parameters depend on genetic and environment conditions [43]. A large number of parameters have been proposed to define malting barley quality. The most used parameters in malting and brewing industry are barley protein, β -glucan and arabinoxylan content. Barley protein content is one of the essential barley components for beer quality. Malting barley with high protein content result in a lower extracts for the brewer. It also slows down water uptake during steeping, potentially affecting final malt quality. While a very low protein level, results in a lack of enzymes necessary to modify the barley kernel and to break down the starch during brewing [43] [6]. β -glucan levels have long been regarded as one of the most influential characteristics in relation to malting potential and brewing yield [7][4]. In the barley grain, β -glucan compose the cell wall of the endosperm which should be degraded during malting to allow starch to be accessible to amylase and other enzymes during brewing. It is thus closely associated with malt quality. β -glucan has long been recognized as having the potential to cause a mash and filtration problems in the brewing process. These includes poor mash conversion, delayed and inefficient wort and beer filtration, and nonbacterial colloidal hazes in the final production. Therefore, better malting performance is expected to be associated with lower levels of β -glucan in grains[43][45][46]. Arabinoxylan content in barley grain also depends on genetic and environment conditions but appears to be less variable than β -glucan [10]. In general insufficient degradation of cell wall hinders the diffusion of germination enzymes, the mobilization of kernel reserves and hence reduces malt extract. Arabinoxylan is also one of the major components of cell wall. High level of Arabinoxylan leads to highly viscous wort and beer, rise filtration problems in the brewery. In malt barley breeding, low Arabinoxylan content or high activities of xylosidase, which degrades Arabinoxylan, is an important target.

Quality requirements for malting barley are reasonably strict and are directly related to processing efficiency and product quality for the malting and brewing industries. Malting process could be controlled by adapting the technology conditions to the samples qualities. Contrariwise, the control of the raw material still very difficult and veritable challenge for maltsters and brewers. This study is therefore focusing on studying and characterizing barley quality parameters.

As it is impossible to cultivate barley grains under similar conditions; characterisation of grain quality influenced by environment and biochemical complexity is very important for malt quality control. The ability to predict barley quality for malting purposes would be of great benefit to breeders and maltsters, allowing for selection of suitable lines to deliver malt highest quality. Many investigations have been realised for studying barley quality parameters. In few studies, simple correlations coefficients between various barley measurements were calculated. The inverse relationship between barley nitrogen and starch content was confirmed by [11] . The barley germination parameters “vigor” and “viability”

were modelled as a function of barley grains physical and chemical properties [12]. With $R^2 = 0.44$ for the viability as a function of hardness index, area, volume, 1000 kernel weight, and $R^2 = 0.71$ for vigor as function of protein, kernel roundness, length, width, and volume. On the one hand, many published studies used a limited number of cultivars or samples from small number of crop years or growing sites. A large site of data was utilised but to model malt quality [13]. On the other hand, in many cases the duration of the quality assessment is of major importance. Besides the traditional analytical techniques makes barley quality analysis expensive and time consuming and often require specialized equipment. Trends in food analysis are towards fast, simple and reliable analytical methods. These features are characteristic of a range of spectroscopic methods including mid-infrared (MIR) and near-infrared (NIR) spectroscopy. Near-infrared (NIR) spectroscopy is widely used in the investigation of cereals quality for breeding programmes since 1970 till now. Various studies concerning barley quality have been carried out with NIR spectroscopy Moisture [15], protein [16][17], β -glucan [18], [19]. Reported applications of NIR spectroscopy in this domain is still increasing. MIR technique was successfully employed to study and model alcoholic beverage quality such as wine and beer [24] [21]. MIR spectroscopy of barley was used by [22] [23] to identify the spectral interval where the specific bonds of protein, β -glucan, and arabinoxylan are located. It have been showed by [24] that MIR spectra predicate a good β -glucan with $R^2 = 0.92$ during barley filling.

Therefore, it could be of most interest to investigate MIR technology as a complementary tool for malting barley characterisation.

When a multiples measurements are made on a set of data chemometrics are required to extract maximum useful information. Chemometrics is the application of mathematical or statial methods to improve the understanding of chemical information and to correlate quality parameters or physical properties to analytical data. PLS regression is one of the most used methods of chemometrics. PLS regression was used to predicate malt quality as a function of barley and malt qualities [39]. Malt qualities were modelled as a function to barley qualities by PLS with high statistical significance [13]. Chemometric have been also developed to value the interpretation of the information contained within the spectral data. Partial least square regression (PLS) has successfully accompanied MIR analysis for spirit drinks analysis, quality control of beer [24][26], Wine [27][28]. PLS models to predict beta-glucan content of barley during grain filling were presented by [24].

This work aims at evaluating the potential of MIR infrared spectroscopy in combination with different signal pre-treatments and Partial Least Square Regression to classify and predict the most important barley quality parameters, protein, β -glucan, and arabinoxylan content.

The materiel and methods section describes the collection of barley samples, their MIR spectra analysis and multivariate analysis methods. In the Results and discussion section, PLS analysis of barley quality index is first performed before PLS regression on MIR spectra are presented and discussed. PLS regression is carried out over all crop year, then for each year

separately. Results are used to predict the barley quality of one crop year using the model based on the previous harvest.

2 Materiel and methods

2.1 Sample collection

A total of 394 Barley samples provided by the French Institute of Brewing and Malting (IFBM) were studied. The samples were harvested in 2009, 2010 and 2011. They are composed of 256 spring barley and 138 winter barley, with 77 different varieties, grown in 16 different sites in France. The sample set thus offers a unique database to study the barley quality index.

The MIR-spectroscopy requires powder samples. The first step in the spectra collection was the samples milling. Spectra samples preparation: 10g of barley milled in a Cyclotec™ 1093 for a 1 min, the grinding is carried out by a high-speed action where the sample is rolled against the inner circumference of a durable grinding surface and passed out through a screen. The high grinding capacity reaches 4 g/second and the recovery of the sample is complete.

Mid-infrared spectra were collected at room temperature with transmittance Fourier Transform InfraRed (FTIR) spectroscopy (PERKIN ELMER Spectrum One FT-IR spectrometer), in the interval $400\text{--}4000\text{ cm}^{-1}$ at resolution of 4 cm^{-1} . Every spectrum is the average of 5 spectra with 15 repetitions, and the spectra were corrected against the background spectrum of air.

2.2 Spectroscopic analysis

MIR spectroscopy has long been used for the elucidation of molecular structure. But thanks to interferometer of Michelson and Fourier transformation (FT) MIR was passed into quantitative analysis. Contrary to analytical methods by coloured or enzymatic reaction that react with a given molecule, the infrared analysis is a structural method which indicates the nature of the bonds between atoms in molecules. Organic compounds have the peculiarity of having inter-atomic bonds that come into vibration under the action of infrared radiation at characteristic wavelengths. This phenomenon is accompanied by a light energy in the considered wavelength. FTIR measures the absorbance of energy of chemical bonds of all components present in the analysed products, and correlated these components to their concentrations, since the intensity of the absorption is directly proportional to the concentration of the molecule. Chemical bonds in a sample (barley powder for this study) vibrate at specific frequencies, which are determined by atoms mass, molecule shape, bonds stiffness, and the period of the associated vibrational coupling. MIR interval of the electromagnetic spectrum lies between $4000\text{ and }400\text{ cm}^{-1}$. In this interval, some characteristic absorption bands are associated with some major components of barley. In the literature, protein is estimated using spectral interval related to amide I ($1650\text{--}1653\text{ cm}^{-1}$) and amide II ($1550\text{--}1567\text{ cm}^{-1}$) bonds [29]. β -glucan and arabinoxylan are showed to be in the region called fingerprint region $800\text{--}1200\text{ cm}^{-1}$, the region of carbohydrates. While β -glucan exhibited an

absorption at 895 and 991 cm^{-1} . Arabinoxylan presented a maximum absorption at 1045 and 958 cm^{-1} [22]. Figure 42 showed the fingerprint in the interval 800-1700 cm^{-1} . In this interval of spectrum, the specific spectral bands for each component of barley are depicted.

2.3 Multivariate data analysis

There are numerous data regression methods to process experimental data. Multiple Linear Regression (MLR) is a usual regression method to find the best linear relation between factors and one single measurement. When the number of predictors is large compared to the number of observations, $\mathbf{X}$ is likely to be singular and the regression approach is no longer feasible because of multicollinearity. This data configuration has been recently often called the '*small N large P problem*.' It is characteristic of recent data analysis domains such as, e.g., bio-informatics, brain imaging, chemometrics, data mining, and genomics.

Multivariate regression deals efficiently with tell data sets. All multivariate regression methods are preceded by a multivariate analysis, for taking high dimensional data, and using the dependencies between the variables to present it in more traceable, lower dimensional form, without losing too much information. Principal Component Analysis (PCA) is one of the simplest and more robust ways of doing such dimensionality reduction. The PCA can provide basis for the MIR-PLS analysis as it can be useful analytical strategy, in which complex data sets are explored with a minimum of suppositions, by mathematically reducing their dimensionality into fundamental underlying factors followed by validation against prior knowledge. Thus, with less a priori knowledge, data analysis can be exploited to achieve a better understanding of complex data sets [39]. PCA of the current data was established in a previous work (Ajib *et al.* 2013) and is not detailed here. The main results of the PCA are resumed in the results section.

Several approaches have been developed to cope with the multicollinearity problem such as Principal component regression (PCR) which performs a principal component analysis (PCA) of the $\mathbf{X}$ matrix and then uses the principal components of $\mathbf{X}$ as the independent variables of a multiple regression model predicting $\mathbf{Y}$. This approach works well because the orthogonality of the singular vectors eliminates the multicollinearity problem. But, the problem of choosing an *optimum* subset of predictors remains. A possible strategy is to keep only a few of the first components. But these components were originally chosen to explain $\mathbf{X}$ rather than $\mathbf{Y}$, and so, nothing guarantees that the principal components, which 'explain' $\mathbf{X}$ optimally, will be relevant for the prediction of $\mathbf{Y}$. PLS finds components which make the two criteria of equal importance. It explains variance of the independent factors $\mathbf{X}$ while simultaneously correlating them to the dependant variables $\mathbf{y}$. It therefore often leads to models that are able to fit the response with fewer components. It could be considered as more efficient than MLR and PCR [30].

Partial least-squares (PLS) regression in combination with the spectroscopic technique was first introduced by [31] to construct predictive models. The method is based on the assumption of a linear relationship between the dependent variable of interest (e.g. protein content in barley grains) and the predictor variable (e.g. absorbance peaks in the spectra). The PLS regression analysis reduces the number of independent variables (i.e. spectra of samples) into a limited number of predictor variables, called eigenspectra (i.e. PLS loadings or factors) [31].

2.3.1 Spectra pre-treatment and intervals selection

Calibration models were first investigated on the whole spectra. Results were non satisfactory the raw or preprocessed MIR-spectrum was divided into subintervals to avoid the interaction between information and the superimposition of signals. The width of the subintervals was varied from 100 cm^{-1} up to 400 cm^{-1} . In addition, spectra were subjected to numerous pre-treatment. All the spectra pre-treatment tested in this work are explained by Ajib *et al.*(2013). In this study, the 394 spectra were pre-processed according to the following techniques: data centering, SNV normalization, Savitzky-Golay smoothing [3, 13], 1st and 2nd derivative to achieve the best performance possible, and extract the maximum of information. Best subintervals were selected according to 3 criteria (highest R^2 , minimum number of components and minimum root mean square error of cross validation).

In any empirical modeling, it is essential to determine the correct complexity of the model. With numerous and correlated X-variables there is a substantial risk for over-fitting i.e., getting a well-fitting model with little or no predictive power. Hence, a strict test of the predictive significance of each PLS component is necessary, and then stopping when components start to be non-significant. Cross-validation CV is a practical and reliable way to test this predictive significance [32]. The calibration model was developed using leave-one-out cross validation (LOOCV) method in PLS regression analysis. The LOOCV procedure estimates the prediction error by removing samples one by one from the calibration samples data and predicting them as unknown samples using the remaining samples in the data. The PLS model training process continues until the minimum root mean square error of cross validation (RMSECV) is attained, which is used with R_{CV}^2 to choose the optimal number of PLS components in order to diminish the possibility of over fitting the model [31].

It is also important to estimate the true errors in the prediction model and validate it using a set of unknown barley samples, i.e. validation set. The prediction coefficients obtained from the calibration model was applied to the validation set in order to predict the content of barley quality parameters in these samples. To evaluate the efficiency of the prediction-model, predicted qualities values were plotted against the measured barley qualities content. The following statistics was used to assess the prediction ability of calibration and validation models including, the coefficient of determination in calibration (R^2_C) and validation (R^2_V), the root mean square error in calibration set (RMSEC), the root mean square error of validation set (RMSEV) [33].

Partial least square regression (PLS) was carried out with MATLAB® version 7 (R2010a). Each barley quality index (protein, β -glucan and arabinoxylan content) was modelled separately as a function of MIR spectra.

3 Results and discussion

This work presents the results of PLS regression on MIR spectra of barley samples. The regression was performed in first time for each year alone. Then the prediction of one year from another was realized. Finally a general and total regression on samples from all years in the same time was investigated.

PCA analysis of all our spectra data set were realized in earlier work (Ajib *et al.* 2013). The PCA was realized on exploratory data and on pretreated spectra in sub-intervals. A very good result was found for spectra barley classification by year of crop and by specie.

In order to identify the most significant model through PLS regression, the samples were split into two subsets: one calibration set and one validation set. To avoid the disturbance which may be caused by the selection of the validation subset, a loop of 30 repetitions was programmed to randomly extract 4 samples as validation set from the whole set. The rest of the set was considered as the calibration set. For each loop (i.e. each random subset), calibration and validation R^2 were calculated, and at the end of the loop mean calibration and mean validation R^2 were evaluated.

3.1 Barley quality parameters

All barley quality indices show that the experimental dataset presents a large variation in barley quality parameters, including value out of the usual acceptable quality range. It is thus a suitable base which provides the most pertinent information to model the relationship between barley composition, barley MIR spectra and barley quality. The exploring barley dataset issued from 394 barley samples from three years of crop are presented and examined in details by Ajib *et al.* (2013). Investigation of MIR spectra by PCA analysis pointed out a significant difference between barley 2009 from those of 2010 and 2011 harvest.

Table 20 mean and interval of barley samples of 2009, 2010, and 2011 crops, from spring and winter species.

	Mean and range of spring and winter barley qualities of 2009-2010-2011 harvest					
	Moisture % dm		Protein % dm		Beta-glucan % dm	
Year	Spring barley	Winter barley	Spring barley	Winter barley	Spring barley	Winter barley
2009	13 (13-14.2)	12.7 (12.8-13)	10.4 (9-11.8)	10 (9.4-11.5)	3.6 (3.56-5.95)	3.1 (2.09-4.18)
2010	14.01 (12.4-15.5)	13.3 (11.5-15)	9.8 (8.7-11.9)	10.1 (8-12.4)	3.31 (2.18-7.73)	3.29 (2.25-4.99)
2011	14 (12.2-14.9)	11.9 (10.6-12.5)	10 (8.9-12.9)	10.6 (8.9-13.2)	3.33 (1.95-5.62)	3.63 (2.7-4.4)

The differences between mean values of barley parameters are investigated in Table 20. The two-by-two parameter means (2009-2010/2009-2011/2010-2011) was compared using Student t-test. Starting by the hypotheses $\mu_1 = \mu_2$ with the risk $\alpha = 0.05$ (5%) the test t of student was applied using the following formula:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (41)$$

with

$$s^2 = \frac{(n_1 - 1) \times \sigma_1^2 + (n_2 - 1) \times \sigma_2^2}{(n_1 + n_2 - 2)} \quad (42)$$

n Size of sample, σ standard deviation, $(n_1 + n_2 - 2)$ the degree of freedom (df)

Once t calculated it is compared to critical t of student found in the student table for the risk 0.05 and df (130, 198, and 178) for 2009-2010, 2009-2011, 2010-2011, respectively. From the literature for $df \geq 130$ - which is the case here- the critical $t \leq 1.66$. Results of student test β -glucan and protein means from 2009-2010-2011 crops were presented Table 21.

Table 21 Reveals the difference in β -glucan mean between 2009 and these of 2010 and 2011 as calculated t 2.36, 2.05 > critical t 1.65 for 2009-2010 and 2009-2011, respectively. In the same time, the table shows the difference of means between 2009 and 2010, 2011 for protein content with calculated t 3.46, 2.11 > critical t 1.65 for 2009-2010, 2009-2011, respectively. This test confirms the particularity of 2009 barley from these of 2010 and 2011.

Table 21 t-test of two-by-two means comparison between barley parameters from 2009, 2010 and 2011 crop

	Degree of	Calculated t		Critical t
	freedom (df)	Protein	β -glucan	
2009-2010	130	3.46	2.36	1.66
2009-2011	198	2.11	2.06	
2010-2011	178	1.23	0.15	

3.2 Best spectral intervals for PLS models

In many statistical tools, PLS regression can be performed with default settings. PLS regression quality can however be strongly influenced by these settings. In order to obtain the best results from PLS regression on our database, it was necessary to identify the best settings for: exploited spectra subinterval, number of principal components, and pre-processing of spectra. Pre-treatment settings were selected for each year, each species, and for each barley quality parameter. Results have shown that the same settings were required for all year spectra. In this section, only the settings for the 2010 spectra are presented.

In a first step, PLS regression on the whole spectra did not lead to satisfactory results. It is therefore necessary to analysis spectral subintervals for PLS regression. The whole spectra were scanned while performing PLS regression on small spectral interval. The width of the spectral subinterval were selected as either 100 cm^{-1} , 200 cm^{-1} 300 cm^{-1} or 400 cm^{-1} The

choice of the best intervals was based on three criteria: highest correlation coefficient (R^2) for the calibration set, best correlation coefficient for cross validation (R^2_{cv}), or minimum root mean square error of validation (RMSEV). In this paper, only the best intervals are presented for each spectral width. They are illustrated in Table 22 with the results for spring barley 2010 is summarized. The PLS regression was made on original spectra with 10 principal components. From Table 22, it is clear that the PLS regression on 100 cm^{-1} intervals gives a good coefficients of calibration $R^2_c=0.64$, 0.72 but weak coefficients for cross validation $R^2_{cv}=0.0075$, 0.0672 for 1000-1100 and 1600-1700 cm^{-1} intervals respectively. Results are therefore not acceptable. For 200 cm^{-1} and 300 cm^{-1} intervals very good regression and calibration models, but weak validation properties $R^2_v=0.32$ for the interval 1000-1300, 0.29 for the interval 800-1000. For 400 cm^{-1} width, only satisfactory one spectral interval was found. The most informative spectral band was located in 800-1200 cm^{-1} with a good values for cross validation, calibration and validation indicators (in bold character in Table 22)

Table 22 PLS model qualities indicators with 10 components for spectra of spring barley 2010. For β -glucan as quality.

		Indicators of PLS Models qualities vs. Interval width					
		R^2_c	R^2_v	R^2_{cv}	RMSEC	RMSEV	RMSECV
100 cm^{-1}	1000-1100	0.64	0.26	0.0075	0.0779	0.0859	0.0792
	1600-1700	0.72	0.63	0.0672	0.0602	0.0358	0.0552
200 cm^{-1}	800-1000	0.91	0.29	0.66	0.0352	0.0725	0.0362
	1000-1200	0.78	0.18	0.52	0.0621	0.0442	0.0631
	1600-1800	0.74	0.12	0.87	0.0668	0.0623	0.062
300 cm^{-1}	700-1000	0.93	0.29	0.53	0.0257	0.0547	0.0258
	1000-1300	0.82	0.32	0.79	0.0579	0.0215	0.0577
400 cm^{-1}	800-1200	0.88	0.82	0.55	0.0386	0.0233	0.0411

Once the spectral interval has been selected, the determination of the maximum number of principal components was carried out. The program was tuned with 30 principal components and the model quality indicators were calculated with β -glucan of barley as response. The retained number of components was defined as the trade-off between a minimum number of components, the highest R^2_c , and R^2_{cv} . Figure 41 presents the coefficient of determination R^2_v , and R^2_{cv} of PLS calibration and cross-validation models vs. principal components. The figure shows that with 10 principal components, PLS model reaches the maximum R^2_{cv} while more components do not enough increase the R^2_c .

In this study, 10 components are used for PLS regression on β -glucan data.

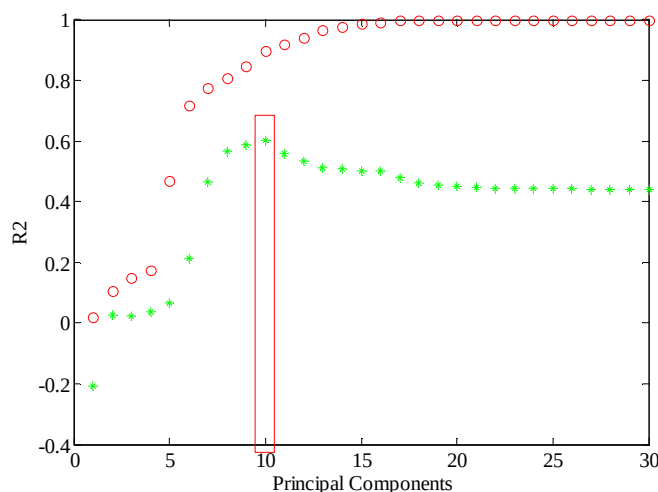


Figure 41 PLS model quality index (coefficient of determination) vs. principal components. For spring barley of 2010 crop, with 30 principal components, PLS regression was applied on centered data. (○) coefficient of determination for PLS calibration models, (*) coefficient of determination for PLS cross validation models

As the best interval width and best number of components were found to be 400 cm⁻¹ and ten components, respectively. The influence of spectra pre-treatment on the PLS results was studied and the most adapted spectra pre-treatment, leading to best results with PLS regression, was selected. Spectra of spring barley 2010 was subjected to PLS regression with various signal pre-treatments. Table 23 summarizes the best correlation obtained with respect to different signal pre-processing. It demonstrates that raw original spectra provide very good PLS model. However, using normalized/SNV pre-treatment decreases quality of the PLS regression (coefficient of determination of calibration down to 0.81). Similar trends are observed on root mean squared error (RMSE).

Table 23 PLS models qualities vs. Spectra pre-treatments for β -glucan of spring barley 2010, in the interval 800-1200 cm⁻¹. with ten PLS terms

Spectra	Indicators of PLS models qualities vs. Spectra pre-treatments					
	R^2_c	R^2_v	R^2_{cv}	RMSEC	RMSEV	RMSECV
Originals	0.89	0.86	0.55	0.0386	0.0886	0.0411
Centered	0.90	0.86	0.58	0.0366	0.0253	0.0386
Normalised with SNV	0.81	0.88	0.58	0.0525	0.0784	0.0537
1st derivative	0.86	0.76	0.46	0.0416	0.0451	0.0414
2ed derivative	0.84	0.51	0.16	0.0534	0.0234	0.0516

Using first and second derivative of the spectra, R^2 of cross validation and prediction models decreases. Derivation thus seems to cause a loss of information. According to table 5, it is

clear that centering spectra is the pretreatment leading to the highest determination coefficient for calibration models ($R^2_c=0.90$), and to a very high one for prediction models ($R^2_v=0.86$), and also for cross validation ones ($R^2_{cv}=0.58$), and minimum root mean squared error (RMSE) for calibration, validation and cross validation (0.036, 0.0253 and 0.038 respectively).

These approach were repeated for each barley quality (β -glucan, Protein, arabinoxylan), in each year, for spring and winter barley but results are not presented here.

In conclusion of the PLS settings section, the best results of PLS regression on barley spectra was obtained for all crop year with centered spectra, with 400 cm^{-1} spectral width, and with ten Principal Components.

Best prediction for β -glucan and arabinoxylan was found in the interval $800\text{-}1200\text{ cm}^{-1}$ Figure 42. The position and the intensity of such spectral bands are specific for every polysaccharide [34] [35]. This is within the so-called fingerprint region of carbohydrates.

As far as protein prediction is concerned, the interval $1300\text{-}1700\text{ cm}^{-1}$ was the best for spring barley of three years. It agrees with the subintervals in the literature, where amide **I** and amide **II** manifest. Figure 42 present the best spectral zones of each barley quality parameter in this work which are compared to the spectral zones presented in literature.

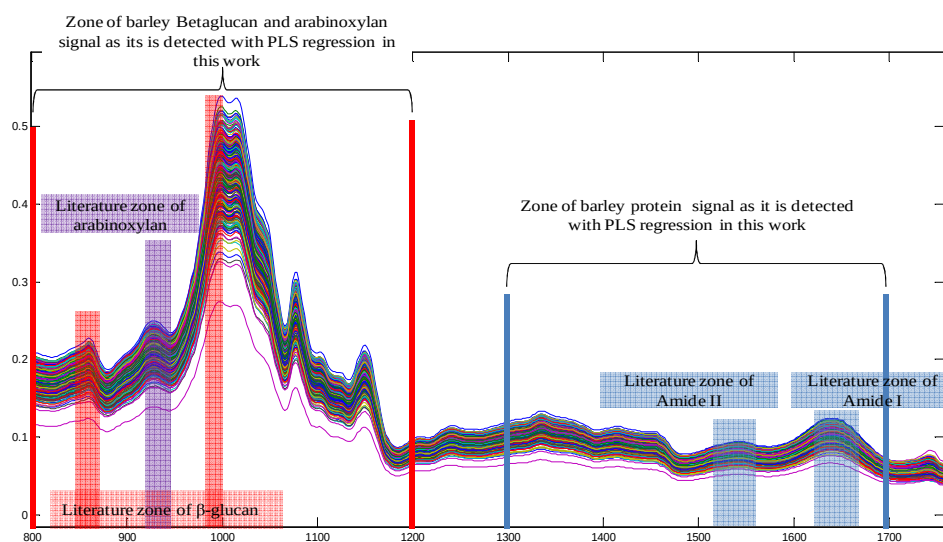


Figure 42 FT-IR spectra in the interval 800-1700 with the regions of barley (protein, β -glucan, and arabinoxylan) signals

3.3 PLS models

Once the PLS settings are set, PLS analysis was carried out to predict the conventional analysis (protein, β -glucan, arabinoxylan content) using spectral (MIR) data. In the first time regression models were obtained in applying PLS regression on the whole spectra.

In our first study, each barley parameter (protein, β -glucan, arabinoxylan) was modelled as a function of MIR spectra from spring barley of all years. For β -glucan, PLS regression was realised as mentioned before in the interval 800-1200 cm^{-1} with centred spectra. The best results - not presented here - were obtained with centred spectra in the interval 800-1200 cm^{-1} interval. Results were however not satisfactory ($R^2_c=0.57$).

Normally such a large data site forms a very suitable base for multivariate analysis. By cons, data in this work is coming from three years with different conditions climatic; therefore differences genetics from one year to another. The variance could be having as importance that it make very difficult to find a significant correlations.

As no satisfactory results were found and In order to better explore such a large database, PLS regression was then realised on each year spectra separately.

3.3.1 Regression by crop year

As it was shown in the preliminary study, barley species have different characters. PLS Regressions were carried out with samples from each single year. Spring barley was also modeled separately from winter barley. Best results of PLS regression on MIR spectra were obtained for 2010 spring barley. Only these resultants are detailed here. The results of PLS regression on the other years barley are summarized. It is worth to notice that no significant models for arabinoxylan prediction were found.

3.3.1.1 2010 barley modelling

56 centered spectra of spring barley 2010 were the base of PLS regression to predict barley qualities. PLS regression was performed on centered spectra in the zone 800-1200 cm^{-1} . PLS models were built on β -glucan and protein content, with ten principal components. 3 model quality indexes were studied: variance explained by principal components, the mean square error and determination coefficient R^2 . Figure 43a) presents mean residual square error of spectra (X) and β -glucan (y) with respect to the number of principal component and Figure 43b) shows the percentage of variance explained in the matrix of independents factors X and in the response y with respect to the number of Principal Component. In the light of figures 3,

10 PLS components explain 100% of spectra variation and 86% of the β -glucan variations, It confirm that the selected settings for PLS regression carried out in the zone 800-1200 cm^{-1} provides the best models.

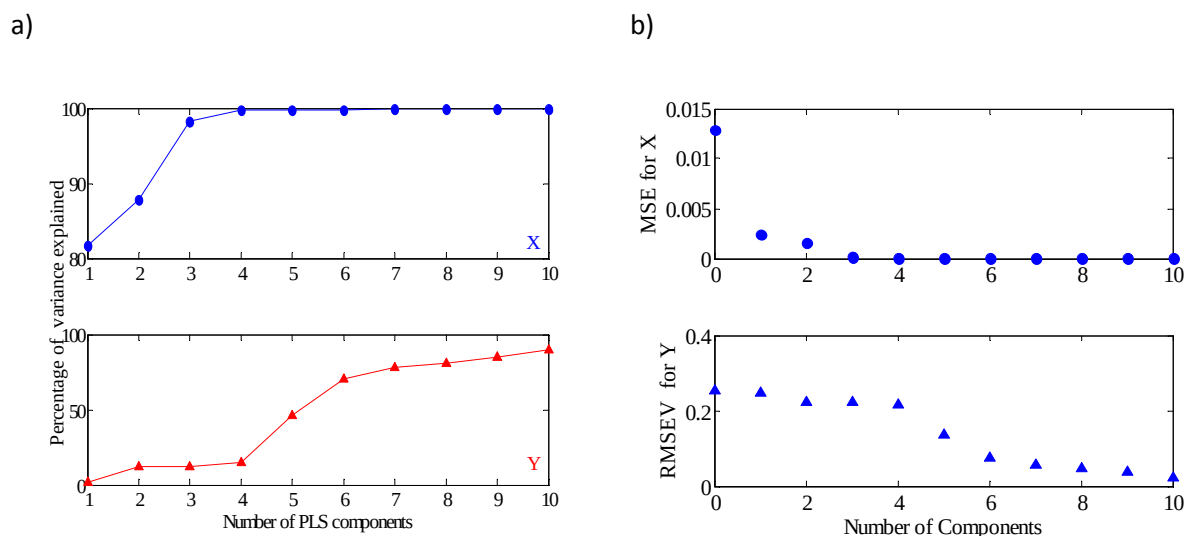


Figure 43 PLS criteria plot for β -glucan models of calibration from spring barley 2010: a) the percentage of variance explained in blue for spectra (X) and in red for the response β -glucan content y. b) the mean square error for spectra X model and β -glucan content y model

Another indicator of model quality is the coefficient of determination for the validation set R^2_v . It presents prediction capacity of the model. In the 56-sample dataset, 52 samples were employed to build the calibration model and 4 samples were used for the prediction of β -glucan. Validation samples were randomly chosen to assure that the validation is performed on the largest β -glucan range. Figure 44 presents fitted β -glucan vs. observed β -glucan for both calibration and validation sets. Figure 44 shows very satisfactory correlations: $R^2_c = 0.90$ and $R^2_v = 0.86$. The fitting and prediction capacity of PLS model for β -glucan are thus validated.

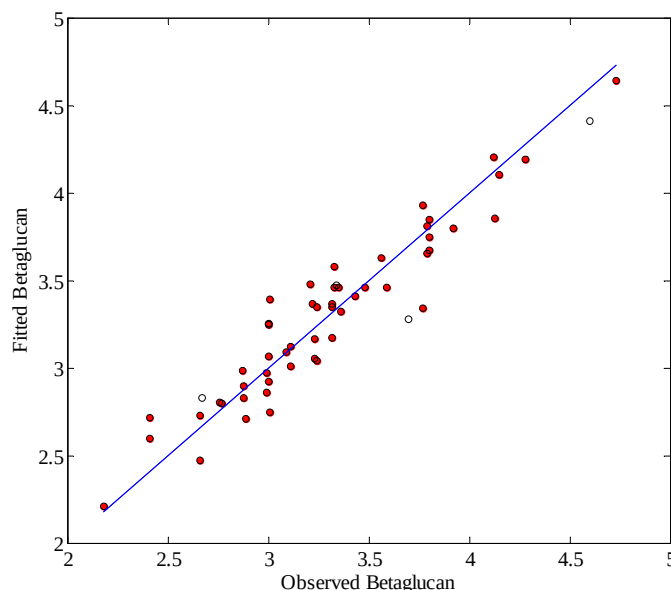


Figure 44 Comparison of fitted β -glucan content obtained by PLS regression vs. Observed β -glucan content for spring barley 2010. 10 principal components were used on centred spectra in the interval [800-1200] cm^{-1} , with $R^2_v = 0.86$.

The predictive capacity of PLS models was assessed. For the validation set, the prediction mean error was calculated. The resulting mean relative uncertainty was 7.47%. These statistical accuracy was compared to the analytical accuracy of β -glucan measurement (24% relative uncertainty according to EBC 3.10.2). It was found that the PLS model has a very good predictive capacity. All PLS models have shown the same accuracy of β -glucan prediction for every crop year and specie.

For protein prediction, the same protocol as for β -glucan prediction was followed. The best results were obtained in the interval [1300-1700] cm^{-1} with centred spectra and with 10 principal components. Comparison of predicted and observed protein content is presented on Figure 45 . It shows that fitted protein content is in very good accordance with observed ones ($R^2_v = 0.85$). Prediction accuracy (3.76%) was also satisfactory with respect to the experimental accuracy on protein measurement (30% incertitude relative issue EBC 3.10.2)

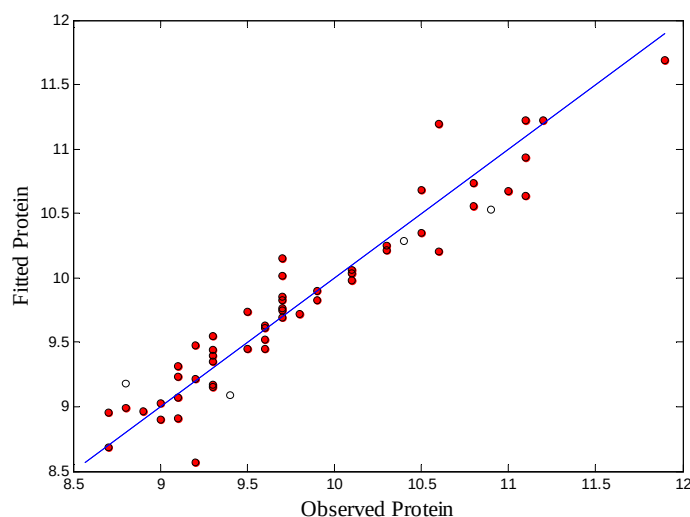


Figure 45 Fitted protein by PLS regression vs. observed protein from spring barley 2010. With 10 principal components, centred spectra, in the interval $[1300-1700] \text{ cm}^{-1}$, with $R^2_v=0.85$.

The same approach was performed to build model for all crop years and species. Results are summarized in the Table 24. It presents the model quality index (R^2_c , R^2_{cv} , R^2_v , RMSEC, RMSECV) for the best PLS models for each crop year for the prediction of both β -glucan and protein content. All PLS regressions were applied with ten components on centred spectra in each specific spectral range.

Table 24 PLS models quality index for each crop, specie. PLS models were performed on centered spectra with ten components, In the interval $800-1200 \text{ cm}^{-1}$, $1300-1700 \text{ cm}^{-1}$ for β -glucan, protein, respectively.

Model quality	2009 barley		2010 barley				2011 barley	
	Spring barley		Spring barley		Winter barley		Spring barley	
	Protein	Betaglucan	Protein	Betaglucan	Protein	Betaglucan	Protein	Betaglucan
R^2_c	0.83	0.70	0.93	0.90	0.80	0.60	0.81	0.68
R^2_{cv}	0.82	0.62	0.92	0.89	0.79	0.60	0.80	0.66
R^2_v	0.75	0.72	0.85	0.86	0.7	0.55	0.75	0.49
RMSEC	0.1210	0.1076	0.0312	0.1100	0.1012	0.1417	0.0843	0.1795
RMSECV	0.1240	0.1099	0.0320	0.1115	0.1045	0.1291	0.085	0.1813
RMSEV	0,2532	0,2785	0,2011	0,1203	0,2307	0,4526	0,2531	0,5187

For both β -glucan and protein, the same protocol was applied for each crop year and for each barley specie. Table 24 compiles the model performances. All coefficient of determination of calibration are very satisfactory or excellent, except for β -glucan fitting for 2010 winter and 2011 spring barley. Table 24 also demonstrates that the barley protein could be predicted

successfully with PLS models on MIR spectra with very good model quality index $R^2_v = 0.75$, 0.80, and 0.75 for spring barley 2009, 2010, and 2011, respectively. PLS prediction of winter barley protein was carried out only for crop 2010 with very good correlation ($R^2_v = 0.70$). Regression was not possible for 2009 and 2011 as there were no sufficient samples.

As far as β -glucan prediction is concerned, excellent model for 2010 spring barley and very good for 2009 spring barley with $R^2_v = 0.86$, 0.72, respectively. By cons, the prediction of β -glucan for 2011 crop was not satisfactory with a weak $R^2_v = 0.49$. As mentioned before, PLS regression on winter barley was done for 2010 crop only. Results are neither satisfactory ($R^2_v = 0.55$).

It can be concluded that PLS regression with MIR spectra is a successful tool for barley quality prediction by crop. It allows maltsters and brewers to use new relevant, rapid and nondestructive tool for barley quality parameters prediction in complement to conventional analysis.

Our results have also pointed out that barley qualities could be predicted for each year separately. It was however shown that the prediction of barley qualities from all crops year together was not satisfactory. It could therefore be interesting assess the possibility to predict barley qualities in applying one model to another year.

4 Conclusion and perspectives

In this study, barley samples issued from 3 years of harvest, 72 varieties, and 16 sites of cultivation, 2 species were analysed using conventional and spectral techniques. It thus provides a unique and very large dataset, mostly suitable for prediction of barley quality. The conventional data were first studied using Student test to highlight variations in samples according to years and species and number of row. In a previous work (Ajib *et al.* 2013) spectral data were analysed with PCA technique and barley samples could be classified according to crop year and species. In this work, MIR spectroscopy accompanied with PLS regression was used to classify and predict barley quality (protein, β -glucan, and arabinoxylan) content.

An exhaustive and systematic approach was performed in order to determine for the PLS regression: the minimum number of principal components, the best spectral intervals to be used and best spectra pre-treatment. Selection was based on the following model criteria: (i) highest coefficient of determination for calibration and cross validation models, (ii) minimum root mean error of cross validation and calibration. Best regressions were obtained for 10 principal components, on centred spectra in the 800-1200 cm^{-1} for β -glucan and arabinoxylan prediction and 1300-1700 cm^{-1} for protein prediction

Results have shown that regression with all the 3 year samples was not satisfactory. PLS models were built for each crop year and specie separately. Excellent models for protein prediction were found for each year, with $R^2 = 0.75, 0.80, 0.76$ for spring barley 2009, 2010, 2011, respectively. Very good model were found for winter barley 2010 with $R^2 = 0.70$. As far as β -glucan prediction are concerned, very good models for spring barley were found for with $R^2 = 0.72, 0.86$ for 2009, 2010, respectively. 2011 barley was difficult to be modelled. Predictions were not significant for arabinoxylan content, neither for 2009 nor for 2010 barley.

Such model allows not only classification of barley samples as good, acceptable or unacceptable for malt production but also prediction of quantitative index. It thus provides a cheap, rapid and simple alternative technique to assess barley quality. It could be very interesting to study the ability of prediction from one year to another.

Références chapitre 6

- [1] B. ajib, F. Fournier, P. Boivin, M. Schmitt, M. Fick. Classification of barley according to harvest year and species by using Mid-infrared spectroscopy and multivariate analysis. *Journal of Food Science and Engineering* 4 (2014) 36-54.
- [2] G. Zhang, J. Chen, J. Wang, S. Ding, Cultivar and Environmental Effects on (1 $\rightarrow$ 3, 1 $\rightarrow$ 4)- β -D-Glucan and Protein Content in Malting Barley. *Journal of Cereal Science* 34 (2001) 295-301.
- [3] O. Østergaard, S. Melchior, P. Roepstorff, B. Svensson, Initial proteome analysis of mature barley seeds and malt. *Proteomics* 2 (2002) 733-739.
- [4] C. Brennan, and L. Cleary, L., The potential of cereal (1-3, 1-4)- β -D-glucans as functional food ingredients. *Journal of cereal science* 42 (2005) 1-13.
- [5] O. Grujic, D.J. Pejin N.M. Przulj, The effect of winter barley variety and technological factors during malting on malt quality. *APTEFF* 36 (2005) 33-42.
- [6] M. Brophy, M. *et al.* Canadian Barley Malting and Brewing Technical Guide. Canadian Malting Barley Technical Centre, 2012.
- [7] S. Aastrup, and K. Erdal, Quantitative determination of endosperm modification and its relationship to the content of (1,3: 1,4)- β -glucans during malting of barley. *Carlsberg. Research Community*, 45 (1980) 369-379.
- [8] J. Wang, G. Zhang, J. Chen, F. Wu, The changes of β -glucan content and β -glucanase activity in barley before and after malting and their relationship to malt qualities. *Journal of Food Chemistry* 86 (2004) 223-228.
- [9] R.B. Vis, K. Lorenz, Malting and Brewing with a High β -Glucan Barley. *LWT - Food Science and Technology* 31 (1998) 20-26.
- [10] M. Gupta, Barley for brewing: characteristic changes during malting, brewing and applications of its by-products. *Comprehensive reviews in food science and food safety* 9 (2010) 318-328.

- [11] C. Agu., Some Links between and Among Quality Parameters Used in Assessment of Barley Adjuncts for Food Processing. Master Brewers Associations of the Americas Technical Quarterly (MBAA TQ) 45 (2008) 268-273.
- [12] L. Munck, and B. Moller, A new germinative classification model of barley for prediction of malt quality amplified by near infrared transmission spectroscopy calibration for vigour''on line'' both implemented by multivariate data analysis. J. Inst. Brew. 110 (2004) 3-17.
- [13] K.J. Siebert, A. Egi, R. McCaig, R., 2012. Chemometric Investigation of Barley and Malt Data. J. American societ. Brew. Chemists 70 (2012)163-175.
- [14] G. Downey, Estimation of moisture in undried wheat and barley by near infrared reflectance. J. Sci. Food Agric. 36 (1985) 951-958.
- [15] E. Roux, Near InfraRed spectroscopy for selection of malting barley in south African breeding programmes.Master Faculty of AgriSciences Department of Food Science. Stellenbosch University, 2011.
- [16] R.J. Henry, Near infrared reflectance analysis of carbohydrates and its application to the determination of (1→ 3), (1→ 4)-β-d-glucan in barley. Carbohydrate research 141 (1985) 13:19.
- [17] S. Trangoonrung, P.M. Haves, S.L. Brorch, Near-infrared reflectance estimates of grain protein and malt extract in hill and row plot evaluations of spring malting barley. Canadian J. Plant Sci. 70 (1990) 71-78.
- [18] M.J. Allison, I.A. Cowe, R. Mchale, The use of infrared reflectance for the rapid estimation of the soluble β-glucan content of barley. J. Inst. Brew. 84 (1978) 153-155.
- [19] Z. Czuchajowska, J. Szczodrak, Y. Pomeranz, Characterization and estimation of barley polysaccharides by near-infrared spectroscopy. I. Barleys, starches, and beta-D-glucans. J. Cereal. Chem. 69 (1992) 413-418.
- [20] D.W. Lachenmeier, Rapid quality control of spirit and beer using multivariate data analysis of Fourier transform infrared spectra. J. Food Chem. 101 (2007) 825-832.
- [21] F.A. Inon, S. Garrigues, M. de la Guardia, Combination of mid-and near-infrared spectroscopy for the determination of the quality properties of beers. Analytica chemistry acta 571 (2006) 167-174.
- [22] P. Robert, M. Malanie, C. Barron, C. Guillon, L. Saulnier, FT-IR investigation of cell wall polysaccharides from cereal grains. Arabinoxylan infrared assignment. J. Agri. Food Chem. 53 (2005) 7014-7018.
- [23] R. Karoui, G. Downey, C. Blecker, Mid-Infrared Spectroscopy Coupled with Chemometrics: A Tool for the Analysis of Intact Food Systems and the Exploration of Their Molecular Structure-Quality Relationships - A Review. ©American Chemical Society, 2010. In: <http://tin.dstc.uniud.it/~gambi/cr100090k.pdf>.
- [24] H.F Seefeldt, A. Blennow, B.M. Jespersen, B. Wollenweber, s.B. Engelsen, Accumulation of mixed linkage (1-3)(1-4)-beta-D-glucan during filling in barley: A vibrational spectroscopy study. J. Cereal Sci. 49 (2009) 24-31.

- [25] J.P. Nielsen, and L. Munck, Evaluation of malting barley quality using exploratory data analysis. I. extraction of information from micro-malting data of spring and winter barley. *J. Cereal Sci.* 38 (2003)173-180.
- [26] E. Polshin, B. Aernouts, W. Saeys, F. Delvaux, F.R. Delvaux, D. Saison, M. Hertog, B.M. Nicolai, J. Lammertyn, Beer quality by FT-IR spectroscopy: impact of measurement strategies, data pre-processing and variable selection algorithms. *J. Food Eng.* 106 (2011) 188-198.
- [27] R. Lleti, E. Meléndez, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, Outliers in partial least squares regression application to calibration of wine grade with mean infrared. *J. Anal. Chim. Acta* 544 (2005) 60-70.
- [28] D. Cozzolino, M. Holdstock, G. Damberg, W.U. Cynkar, P.A. Smith 2009. Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *J. food Chem.* 116 (2009) 761-765.
- [29] P. Yu, Molecular chemical structure of barley proteins revealed by ultra spatially resolved synchrotron light sourced FTIR microspectroscopy: comparison of barley varieties. *J. Biopolymers* 85 (2007) 308-317.
- [30] H. Abdi, Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Comp. Stat.* 2 (2010) 97-106.
- [31] D.M. Haaland, and E.V. Thomas, Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Journal of analytical chemistry* 60 (1988) 1193-1202.
- [32] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58 (2001) 109–130.
- [33] A. Pirie, B. Singh, K. Islam, Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. *Australian J. soil Res.* 43 (2005) 713-721.
- [34] M.A. Coimbra, A. Barros, M. Barros, D.N. Rutledge, I. Delgadillo, Multivariate analysis of uronic acid and neutral sugars in whole pectic samples by FT-IR spectroscopy. *Carbohydrate Polymers*, 37 (1998) 241–248.
- [35] X. Fang, B. Jiang, and X.L. Wang, Purification and partial characterization of an acidic polysaccharide with complement fixing ability from the stems of *Avicennia marina*. *J. Biochemist. Molecular Bio.* 39 (2006) 546-555.

Chapitre 7
Conclusions et
perspective

Conclusion

Dans cette étude, nous avons développé deux approches pour mieux maîtriser la qualité du malt. Pour cela nous avons mis en oeuvre une démarche de modélisation et/ou d'optimisation du procédé du maltage par deux approches:

Une première partie a consisté à compléter une démarche, initiée antérieurement, qui visait à maîtriser les coûts (énergétiques et environnementaux) du procédé de maltage en travaillant à faible humidité, tout en garantissant la qualité de malt final. Cette étude a été réalisée sur 40 orges de la récolte 2006 dont 21 orges de printemps de deux variétés (Scarlett, Sebastian) et 19 orges d'hiver de trois variétés (Azurel, Esterel et Carafe).

Des modèles polynomiaux ont été considérés. Ces modèles ont été construits par régression multilinéaire à partir des mesures expérimentales collectées lors de la mise en oeuvre d'un plan d'expériences. Les facteurs potentiels de ces modèles étaient : le taux de protéines initial des grains d'orge, la température de l'eau de trempe, le diagramme de trempe avec ses trois étapes (premier sous eau, premier sous air, deuxième sous eau), le taux de β -glucanes initiales, et le taux d'arabinoxylane. Nous avons considéré les indicateurs de qualité du malt et du maltage suivants : l'humidité de fin de trempe, l'humidité après un jour de germination, l'humidité finale, la friabilité de malt, l'extrait, la viscosité de moût, l'atténuation limite, les pouvoirs diastatiques, l'indice de germination, l'indice de la plumule, le taux de protéines soluble, le taux de β -glucanes finale, le taux des grains piquent.

A partir de ces modèles, des conditions optimales de maltage à faible humidité ont été proposées. Une étape de validation de ces conditions optimales a été mise en place. Les échantillons d'orge disponibles ont cependant évolué pendant la période de stockage (de 2006 à 2009). Pour intégrer cette évolution, les indices de qualité (Indice de germination, teneurs en protéines et β -glucanes) des rares échantillons d'orge encore disponibles à l'IFBM ont été réévalués en 2009. La mise en oeuvre expérimentale des conditions optimales numériques sur les orges 2006 a montré l'approche d'optimisation était faisable mais que les modèles manquaient de robustesse vis-à-vis de l'évolution de la matière première au cours du temps.

Par ailleurs, en sélectionnant un lot de quatre échantillons d'orge qui présentent des comportements antagonistes mais intéressants lors du maltage, nous avons montré que la teneur en protéines et en bêta-glucanes n'expliquent à eux seuls, les variations de comportement observées, notamment sur la cinétique de prise d'eau des grains pendant la trempe. Il est donc nécessaire de rechercher de nouveaux facteurs pour établir nos modèles. Notre étude s'est orientée vers des facteurs *structuraux* de l'orge : la porosité, la dureté des grains d'orge. Les mesures ont été faites pour la porosité seulement car la dureté est un facteur dépendant de la porosité. 48 échantillons d'orge ont été sélectionnés d'une manière à prendre en compte l'influence de l'année (2007-2008), l'influence de la variété (48 variétés), l'influence du lieu de culture (3 lieux de culture différents).

Nous avons ainsi montré i) qu'il existe bien une différence de qualité de la matière première qui se traduit par des porosités différentes ii) que le protocole de mesure de la porosité suivi dépend cependant trop fortement de la teneur en eau du grain et que ce facteur *structurel* n'a pas apporté d'amélioration significative sur nos modèles.

La conclusion de cette première partie est que la variabilité des échantillons dans le temps et la variabilité des propriétés de l'orge d'une année de récolte à l'autre, la variété, l'espèce ... sont des facteurs de perturbation qui sont mal intégrés dans nos modèles. La démarche d'optimisation est satisfaisante sur le principe mais elle nécessiterait de reconstruire les modèles chaque année pour être exploitée efficacement. Cela s'avèrerait trop long et trop coûteux.

Afin de mieux comprendre et maîtriser ces facteurs clé pour la production de bon malt, notre étude s'est focalisée sur la caractérisation de l'orge. En complément des techniques classiques de caractérisation de l'orge, nous avons introduit une technique spectroscopique MIR. Un large échantillon composé de 394 orges, incluant 138 orges d'hiver et 256 orges de printemps, récolté sur trois ans en 16 lieux de culture en France a été étudiée.

Dans un premier temps, nous analysé l'ensemble des données conventionnelles à l'aide d'outils de classification (ACP) puis nous avons cherché à corrélér les données conventionnelles et les données spectrales par des outils de régression adapté (PLS).

Étude et analyse de l'orge par des données expérimentales et spectrales et l'ACP analyses.

Un grand ensemble de données 394 échantillons d'orge composé de 138 échantillons d'orge d'hiver et 256 échantillons d'orge de printemps, récolté sur trois ans en 16 lieux de culture en France a été étudiée. Pour chaque échantillon, 3 paramètres de l'orge, la teneur en protéines, teneur en eau, teneur en β -glucanes ont été mesurés et cinq paramètres de qualité du malt, après un processus de maltage standard, ont été quantifiés : la friabilité, l'extrait, la protéine soluble, la viscosité et le pouvoir diastatic. Les spectres MIR de grains d'orge broyés ont été mesurés dans l'intervalle $[400 ; 4000] \text{ cm}^{-1}$.

Après avoir étudié l'influence des méthodes de prétraitement des spectres MIR, les meilleurs résultats ont été obtenus en considérant à l'aide de la dérivée seconde des spectres dans le sous-intervalle $[800-1000] \text{ cm}^{-1}$.

L'ACP a été utilisé en deux étapes :

1. analyses ACP pour caractériser l'orge à partir de ses données expérimentales (originales ou normalisés).
2. analyses ACP pour caractériser l'orge à partir de ses données spectrales.

Dans les deux cas, les résultats montrent que les orges de 2009 se distinguent très nettement des orges de 2010 et 2011. Il s'agit de la seule distinction qui a pu être exploitée à partir de l'ensemble des 394 données.

Une analyse plus approfondie a été réalisée en traitant les spectres par année de récolte.

Pour cela la recherche s'est orientée vers les analyses multivariées sur spectres MIR d'orge pour étudier si ces analyses permettent la même distinction/classification et pourrait mettre l'accent sur la variation annuelle.

En n'utilisant que les spectres d'une même année, l'ACP a montré qu'il était possible de distinguer les échantillons en fonction de l'espèce :

- orge de printemps ou orge d'hiver
- orge 2 rang ou orge 6 rang pour les orges d'hiver

L'ACP n'a pas permis de classer les orges selon leur indice quantitatifs de la qualité de l'orge ou de malt (teneur en protéine de l'orge, friabilité du malt...).

Analyse des données spectrales de l'orge par régression PLS

Les données spectrales mentionnées ci-dessus ont été exploitées pour construire des modèles de régression de type moindres carrés partiels (PLS). La régression a été réalisée sur chaque année seule.

Les pré-traitements des spectres qui conduisent aux meilleurs modèles consistent à utiliser des spectres centrés, dans l'intervalle $[800-1200] \text{ cm}^{-1}$ pour la prédiction des bêta-glucanes et $[1300-1700] \text{ cm}^{-1}$ pour la prédiction des protéines. Aucun modèle satisfaisant n'a été établi pour prédire la teneur en arabinoxylane de l'orge. Tous les modèles PLS s'appuient alors sur un maximum de 10 composantes PLS.

En comparaison des tendances publiées dans la littérature, nous pouvons affirmer que d'excellents modèles ont été établis pour la prédiction de la teneur en protéines de l'orge de printemps pour chaque année de récolte avec un coefficient de détermination ($0.75 \leq R^2_v < 0.80$). De très bons modèles pour la prédiction de la teneur en β -glucanes avec un coefficient de détermination R^2_v entre 0.86 et 0.72 ont également été mis au point. Par contre la prédiction de la teneur en β -glucanes pour l'orge de printemps 2011 n'est pas satisfaisante.

Le spectroscopie MIR associée à l'analyse multivariée constitue donc un outil à fort potentiel pour la classification des orges et la prédiction de la qualité de l'orge par année. En revanche, la prédiction de ces facteurs pour toutes les années de récolte réunies ne conduit pas à des résultats satisfaisants.

Perspectives

Les perspectives de ce travail peuvent être résumés en trois points;

Approche boîte noire (MLR, ANN)

1. L'étape de trempe - où la prise de l'eau dans les grains se produit -est une étape clé pour la qualité du malt final. Les facteurs *orge* qui interviennent sur la prise d'eau sont probablement nombreux :
 - a. **chimiques** : teneurs en protéines, en β -glucanes, en arabinoxylanes qui ont été étudiés dans ce mémoire.
 - b. **structurelles** : porosité, dureté, vitrosité pour lesquels une première étude succincte a été réalisée.
 - c. **physiques**: taille de grain, volume de grain, poids de 1000 grains...

Cette étude a largement étudié les premiers. Elle a aussi montré que la mesure de la porosité est fortement sensible à la teneur en eau dans les grains d'orge. Pour pouvoir l'étudier comme facteur un indépendant et pour mieux intégrer ce facteur structurel dans nos modèles, une **optimisation du protocole de mesure de la porosité** est nécessaire (avec une standardisation de l'humidité du grain pendant les mesures par séchage ou lyophilisation des grains).

Il pourrait être intéressant d'intégrer également des facteurs physiques pour enrichir nos modèles : taille de grain, volume de grain, poids de 1000 grain...

2. L'identification des facteurs *orge* comme les β -glucanes, Arabinoxylanes qui influencent la qualité de malt est très importante mais ce ne sont pas les seuls ayant un impact sur la performance du maltage. Des orges avec des teneurs en β -glucanes semblables, cultivées en deux pays différents ont montrés des performances de maltage différentes (Zhang *et al.*2002). Il serait donc intéressant d'identifier de nouveaux facteurs qui s'appuient sur une caractérisation plus fine de l'orge, notamment la nature **biologique** complexe des grains d'orge (activités enzymatiques des β -glucanases, α -amylases).
3. Notre étude nous a permis de construire une large base de données expérimentale et spectrale de 394 orges. Cette base de données reste, à notre connaissance unique et il conviendrait d'en exploiter toutes les facettes. Il serait par exemple intéressant d'essayer améliorer et de mieux modéliser (caractériser) la matière première (orge) ou la relation orge-malt avec d'autre méthode de régression, plus évoluée, et plus polyvalente comme les **réseaux de neurones artificiels** (ANN).

Perspectives liées à l'analyse spectrale.

Dans ce travail, des spectres ont été collectés pour des grains entiers d'orge. Comme la structure du grain implique beaucoup de compartiments fonctionnels (embryon, couche d'aleurone, albumen,...), on peut supposer que l'analyse multivariée sur des spectres de différents **compartiments du grain d'orge** pourrait mieux indiquer les caractéristiques principales d'orge. Concentrer l'analyse sur les compartiments principaux (endosperme avec la couche d'aleurone, l'embryon avec le scutellum) pourrait en effet enlever des interactions peu désirées entre les composés d'orge et améliorer l'analyse PCA ou PLS dans la classification d'orge ou la prédiction de la qualité de l'orge.

Il pourrait également être envisageable de compléter l'information sur les grains d'orge en exploitant des techniques complémentaires. Si des approches utilisant la spectroscopie NIR ont déjà été développées, d'autres pistes sont encore envisageables : **autres méthodes d'analyse**, spectrales ou non spectrales ou analyse d'images.

Dans cette étude, nous avons étudiés et identifiés des relations entre les spectres collectés et les facteurs qualitatifs de l'orge et certains des facteurs quantitatifs de l'orge. Les capacités de prédiction ne sont pas complètes mais compte tenu de ces premiers résultats très prometteurs, il semble pertinent de fusionner nos deux approches :

- Modèle MLR pour la relation entre les qualités de l'orge, les conditions opératoires et les qualités du malt
- Modèle PLS pour la relation entre spectres MIR et qualité de l'orge

Pour mettre en évidence d'éventuelles **relations entre les spectres orges et les qualités du malt** (friabilité, extrait...).

Enfin, si les données spectrales ont permis de caractériser certaines propriétés de l'orge, on peut supposer qu'elles permettraient également de caractériser les modifications de l'orge au cours du maltage dans le but de réaliser le **suivi du maltage par mesures spectroscopiques MIR**. Cette étude compléterait les études similaires antérieures mais avec la garantie d'une exploitation mieux ciblée des bandes spectrales pertinentes du spectre MIR.

Toutes les approches décrites ci-dessus ont en commun le fait qu'elles ne s'appuient que sur vision *boîte noire* et *globale* du système grain d'orge et de la transformation de l'orge en malt. Si ces approches permettent de capturer certaines tendances mais il n'en demeure pas moins que l'**approche phénoménologique** resterait l'approche idéale pour comprendre et résoudre les verrous identifiés lors du maltage. Les outils d'analyse apportent en effet de plus en plus d'informations fines et localisées (dans le grain) qui conduiront à la description fiable des phénomènes clés de la trempe (voire de la germination et du touraillage) et à une meilleure compréhension de l'interaction entre les qualités de l'orge, les conditions opératoires et les qualités du malt obtenu. On peut citer par exemple la simulation de la diffusion de l'eau dans un grain d'orge en 3 dimensions qui a été entreprise dans notre équipe mais qui n'a pas été

développée dans ce mémoire. Les premiers résultats de simulation de la **diffusion de l'eau dans un grain d'orge** doivent encore être affinés par une description plus réaliste de la géométrie du système et doivent être couplés à la simulation des phénomènes enzymatiques pour essayer d'expliquer les comportements observés mais cette voie est également très prometteuse.

References bibliographiques

Zhang, G., Wang, J., & Chen, J. 2002. Analysis of b-glucan content in barley varieties from different locations of China. *Food Chemistry*, 79, 179–182.

ANNEXE

ANNEXE 1 les paramètres de qualité de 40 orges récolte 2006, pour cinq variétés : deux de printemps (Scarlett, Sebastian) et trois d'hiver (Carafe, Azurel, Esterel)

ID	Orge	Protéine	Température	Diagramme	Sous eau 1	Sous air 1	sous eau 2	Indice de germination	β-glucan0	Arabinoxylan0
--	--	%	°C	--	h	h	h	%	mg/l	%
1	Scarlett	11,6	22	1	11	17	7	8,8	5,57	8,16
2	Scarlett	11,6	22	2	8	6	5	8,8	5,57	8,16
3	Scarlett	11,6	14	3	11	10	14	8,8	5,57	8,16
4	Scarlett	11,6	14	4	6	21	5	8,8	5,57	8,16
5	Scarlett	9,8	14	1	11	17	7	9	5,12	7,89
6	Scarlett	9,8	22	4	6	21	5	9	5,12	7,89
7	Sebastian	11,5	14	1	11	17	7	8,3	3,97	6,57
8	Sebastian	11,5	22	2	8	6	5	8,3	3,97	6,57
9	Sebastian	11,5	22	3	11	10	14	8,3	3,97	6,57
10	Sebastian	11,5	14	4	6	21	5	8,3	3,97	6,57
11	Sebastian	11,5	22	4	6	21	5	8,3	3,97	6,57
12	Sebastian	9,8	22	1	11	17	7	8,3	3,87	6,91
13	Sebastian	9,8	22	2	8	6	5	8,3	3,87	6,91
14	Sebastian	9,8	14	3	11	10	14	8,3	3,87	6,91
15	Sebastian	9,8	14	4	6	21	5	8,3	3,87	6,91
16	Carafe	11,7	22	1	11	17	7	9,3	3,69	6,69
17	Carafe	11,7	14	2	8	6	5	9,3	3,69	6,69
18	Carafe	11,7	22	4	6	21	5	9,3	3,69	6,69
19	Carafe	10,7	14	1	11	17	7	9,1	3,53	8,37
20	Carafe	10,7	22	2	8	6	5	9,1	3,53	8,37
21	Carafe	10,7	22	3	11	10	14	9,1	3,53	8,37
22	Azurel	12,2	14	1	11	17	7	9,1	3,72	8,82
23	Azurel	12,2	22	1	11	17	7	9,1	3,72	8,82
24	Azurel	12,2	14	2	8	6	5	9,1	3,72	8,82
25	Azurel	12,2	22	3	11	10	14	9,1	3,72	8,82
26	Azurel	10,2	22	1	11	17	7	6,2	3,63	7,59
27	Azurel	10,2	22	2	8	6	5	6,2	3,63	7,59
28	Azurel	10,2	14	3	11	10	14	6,2	3,63	7,59
29	Azurel	10,2	22	4	6	21	5	6,2	3,63	7,59
30	Esterel	11,3	14	2	8	6	5	9,5	3,41	7,52
31	Esterel	11,3	22	2	8	6	5	9,5	3,41	7,52
32	Esterel	11,3	14	3	11	10	14	9,5	3,41	7,52
33	Esterel	11,3	22	4	6	21	5	9,5	3,41	7,52
34	Esterel	11,3	22	4	6	21	5	9,5	3,41	7,52
35	Esterel	11,3	22	4	6	21	5	9,5	3,41	7,52
36	Esterel	10,3	22	1	11	17	7	9,7	3,63	8,13
37	Esterel	10,3	14	2	8	6	5	9,7	3,63	8,13
38	Esterel	10,3	22	3	11	10	14	9,7	3,63	8,13
39	Esterel	10,3	14	4	6	21	5	9,7	3,63	8,13
40	Esterel	10,3	22	4	6	21	5	9,7	3,63	8,13

ANNEXE 2 les paramètres de qualité de 40 malts issus des orges récoltés en 2006

ID	Indice Grains Piqués	Himidité initial	Himidité jour+1	Himidité final	Indice de plumule	Friabilité	Extrait	Protéine Soluble	Viscosité	β-glucan	PouvoirDia static	Attenuation Limite
--	%	%	%	%	%	%	%	%	mPs.a	mg/l	W.K	%
1	99	42,8	46,7	44,9	81	79,9	81,6	5,12	1,51	47	330	84,7
2	85	37,3	41,7	40,6	79	66,7	81,2	4,42	1,67	535	280	82,2
3	84	39,7	42,7	42	87	72,7	81,9	4,54	1,6	311	300	82,2
4	88	36,5	40,6	39,3	20	54	81,2	3,98	1,74	870	270	79,8
5	94	39,7	42,4	41,5	79	75,1	82,7	3,86	1,62	326	250	81,8
6	94	44,5	47	45,3	86	90,6	82,2	4,36	1,52	45	300	83,1
7	88	40	42,6	41,7	64	76,1	82,2	4,46	1,56	195	360	82,9
8	92	39,1	44	41,1	75	69,4	82,1	4,34	1,57	244	340	83,1
9	94	45,4	48,5	46,1	88	77,2	81,3	4,74	1,54	88	420	83,6
10	90	38,5	41,6	40,1	76	67,5	82,1	4,29	1,55	221	350	82,9
11	96	42,2	45,4	44,7	84	85,2	81,5	4,53	1,58	53	390	82,7
12	95	44,6	47,8	45,3	87	82,8	83,3	4,19	1,54	86	340	84,3
13	87	38,9	42,7	42,1	88	80,7	83	4,01	1,63	225	310	81,6
14	85	40,9	43,9	42,8	95	81,5	83,8	4,16	1,55	147	330	83,9
15	89	39	42,5	40,7	76	75,7	83,2	3,86	1,59	284	300	82
16	96	44	44,8	46	78	89,7	84	5,1	1,53	33	370	83,8
17	70	37,1	39,8	38,7	30	65,8	83,6	4,37	1,6	336	310	80,7
18	95	46,4	48,3	45,6	75	87,1	83,9	5,3	1,51	23	370	84,1
19	97	39,8	43	42,3	68	85,1	84,2	4,85	1,49	137	320	83,7
20	89	40,1	44,5	42	75	82,5	83,9	4,61	1,5	133	320	83,4
21	95	45,8	49,4	47,4	78	84,9	83,4	5,15	1,49	39	340	84,6
22	77	37,4	41,1	40,2	2	62,1	80,5	4,64	1,72	605	390	82,2
23	89	41,8	45	44,4	78	85,9	80,6	5,88	1,59	72	410	83,3
24	35	33	37,3	36,8	0	42,8	79,1	3,93	2,1	1173	360	78,1
25	89	42,7	46,3	45	77	86,5	80,2	6	1,58	47	420	83,8
26	75	42,1	44,4	42,5	15	85,3	80,8	4,63	1,64	210	300	83,8
27	20	37	41,4	39,6	3	65,3	80,5	4,36	2,04	1154	290	81,9
28	38	38,5	41,9	41	3	73,1	81	4,51	1,86	546	320	80,4
29	77	41,3	43,5	41,4	11	83,8	80,3	4,6	1,64	240	260	80,9
30	38	34,2	38,8	36,6	0	55,7	79,5	3,87	1,79	538	360	79,8
31	66	37,5	41,6	39,9	43	69,1	80,2	4,43	1,67	224	400	81
32	83	38,9	42,6	40,3	55	80,3	80,2	4,5	1,57	131	440	82,9
33	99	41,6	45	42,8	85	84,2	80	4,34	1,6	88	400	81,6
34	97	42,2	42,9	42,9	68	84	79,7	4,3	1,58	60	420	82,2
35	96	42,9	43,4	42,1	71	82,3	79,6	4,61	1,59	60	420	82,3
36	98	42,2	45,2	43,7	82	92,1	79,7	4,34	1,54	24	420	83,8
37	12	34,1	37,6	37,3	0	47,7	79,8	3,73	2,23	1996	260	78,7
38	99	44,3	46,8	45,3	98	93,2	79,3	4,27	1,51	<20	430	84,8
39	92	36,2	38,7	38,5	0	69,9	79,9	3,76	1,72	336	380	79,8
40	99	40,4	42,5	40,4	78	90,6	79,6	3,92	1,59	82	400	81,7

ANNEXE 3 les paramètres de qualité du malt et d'orge d'hiver récolte 2009

ID	Orge	Humidité d'orge	Protéine	β -gucan	Arabinoxylan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité	Atténuation Limite
--	--	%	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s	%
1	Vanessa	12.8	11.5	2.23	10.9	80.0	80.9	4.35	370	1.49	82.5
2	15633 FH	13.0	9.90	3.47	8.40	68.0	82.5	3.68	320	1.58	81.2
3	AC 03-271-4	12.9	10.0	3.78	6.74	84.0	80.2	3.72	380	1.49	83.8
4	MH 02 BS 10	13.0	10.7	4.18	6.74	62.0	80.3	3.82	420	1.60	82.0
5	Esterel	12.7	10.6	3.22	7.95	67.0	80.3	3.88	370	1.71	81.7
6	CCB 07 W 901	12.3	9.90	2.57	6.87	65.0	80.7	3.78	360	1.74	80.9
7	MH 02 GG 78	12.1	9.70	3.64	6.59	80.0	79.5	3.80	340	1.53	83.1
8	Esterel	12.5	9.40	2.33	10.9	83.0	80.0	3.42	350	1.57	83.3
9	CCB 07 W 901	12.7	9.70	2.47	10.7	75.0	79.9	3.63	330	1.53	82.6
10	MH 02 GG 78	12.8	9.40	3.28	9.48	85.0	79.4	3.30	290	1.46	84.3
11	Vanessa	12.9	10.1	2.09	10.3	85.0	81.1	4.18	270	1.46	86.6
12	15633 FH	12.7	9.70	2.89	8.33	84.0	81.8	3.86	300	1.48	84.8
13	AC 03-271-4	12.8	9.70	3.48	8.91	85.0	80.2	3.62	320	1.49	84.1
14	MH 02 BS 10	12.9	10.3	3.77	7.28	69.0	80.4	3.67	330	1.51	83.0

ANNEXE 4 les paramètres de qualité du malt et d'orge de printemps récolte 2009

ID	Orge	Humidité d'orge	Protéine	β-gulcan	Arabinoxylan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité	Atténuation Limite
--	--	%	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s	%
1	Scarlette	13.6	11.0	4.31	6.95	75.0	82.1	3.97	340	1.58	83.5
2	Sebastian	13.6	9.20	3.23	6.89	91.0	84.2	3.77	340	1.53	84.4
3	SYN407143	13.3	9.60	3.57	6.87	83.0	83.1	3.61	260	1.51	83.6
4	LDSTRG67804	13.4	9.50	4.28	7.01	83.0	83.3	3.88	290	1.52	83.8
5	LP1057604	13.5	9.50	2.18	8.19	96.0	83.9	4.15	310	1.48	83.6
6	UNS071085	13.5	9.50	3.59	8.12	97.0	84.2	4.31	350	1.48	84.7
7	UNB8701D4	13.6	9.80	3.74	6.83	98.0	83.8	4.12	310	1.47	85.0
8	26141H1	13.6	9.30	3.69	7.21	73.0	83.8	3.45	270	1.54	82.8
9	9021G	13.7	9.80	4.09	6.90	74.0	83.8	3.72	250	1.54	82.4
10	30202G	13.7	9.40	4.12	6.86	89.0	83.4	4.20	280	1.48	84.6
11	Scarlette	12.7	11.4	3.83	8.51	53.0	81.4	3.54	240	1.69	80.5
12	Sebastian	12.6	10.4	2.72	8.37	76.0	82.7	3.25	260	1.52	81.8
13	SYN407143	13.0	11.1	2.83	8.46	83.0	81.9	4.13	220	1.50	81.5
14	LDSTRG67804	13.2	10.3	3.37	7.51	86.0	82.3	3.85	270	1.49	83.2
15	LP1057604	13.2	10.4	2.48	9.16	91.0	82.0	4.19	240	1.46	81.9
16	UNS071085	13.3	10.5	3.10	9.22	89.0	83.8	4.08	290	1.43	80.5
17	UNB8701D4	13.3	10.8	3.32	7.58	87.0	82.6	4.65	320	1.47	83.3
18	26141H1	13.1	10.0	3.11	8.77	69.0	82.4	3.59	210	1.56	80.6
19	9021G	13.2	10.4	3.01	9.60	74.0	81.8	3.94	200	1.55	81.5
20	30202G	13.3	11.2	3.27	9.06	82.0	82.4	3.82	220	1.51	81.9
21	Scarlette	13.4	10.2	4.51	5.90	80.0	83.5	3.59	270	1.52	82.7
22	Sebastian	13.3	9.60	3.33	4.90	88.0	84.1	3.63	330	1.49	83.0
23	BELLINI	13.2	9.60	3.09	6.00	92.0	84.2	3.88	350	1.46	84.3
24	NORD07-2403	13.2	10.0	3.22	5.70	91.0	82.2	3.80	350	1.43	83.7
25	KWS07-33	13.4	9.90	3.10	5.00	98.0	83.5	4.23	360	1.44	84.8
26	AC04-506-12-8	13.3	9.80	3.46	4.60	85.0	84.0	3.69	270	1.47	82.7
27	SYN408-177	13.1	9.10	3.10	2.90	90.0	83.3	3.31	280	1.47	82.1
28	180-02-5B	13.2	9.60	3.22	3.90	96.0	80.3	4.57	320	1.45	83.7
29	65-03NZ22G	13.1	9.40	3.22	3.70	94.0	83.3	4.17	320	1.44	83.7
30	62931 I 2	13.3	9.30	3.23	3.40	87.0	83.3	3.25	260	1.49	82.0
31	SB055559	13.3	9.30	3.11	4.20	91.0	83.7	3.67	400	1.47	82.8
32	UNS071152	13.2	9.20	3.45	3.50	92.0	83.3	3.79	300	1.48	83.8
33	S15501	13.0	9.40	3.33	1.50	97.0	84.3	3.86	330	1.44	85.0
34	PF12079-51	13.0	9.60	3.21	3.20	91.0	83.7	3.97	340	1.45	82.2
35	Scarlette	11.8	12.6	5.11	4.30	57.0	81.8	4.22	350	1.64	80.5
36	Sebastian	11.5	11.4	3.68	3.70	76.0	82.6	4.11	500	1.51	82.5
37	BELLINI	11.6	11.6	4.26	3.40	71.0	82.8	4.30	510	1.50	82.1
38	NORD07-2403	11.6	11.8	4.27	4.80	85.0	82.9	4.55	500	1.47	82.5
39	KWS07-33	11.6	11.2	3.68	2.70	83.0	83.5	4.34	430	1.49	82.4
40	SYN408-177	11.4	10.5	3.90	4.40	73.0	82.8	3.62	330	1.44	81.2
41	180-02-5B	11.4	12.0	3.76	4.40	79.0	83.3	5.30	340	1.50	82.5
42	65-03NZ22G	11.6	11.0	3.91	2.80	88.0	83.1	4.64	390	1.50	83.1
43	62931 I 2	11.4	11.7	4.63	5.40	76.0	81.9	4.14	450	1.51	81.5
44	SB055559	11.4	10.4	4.87	5.30	75.0	82.9	4.11	490	1.51	81.9
45	UNS071152	11.7	10.7	4.27	4.50	88.0	83.6	3.78	300	1.47	81.4
46	S15501	11.5	10.7	3.91	4.10	82.0	83.2	4.14	300	1.48	83.2
47	PF12079-51	11.7	11.8	4.03	4.50	75.0	81.8	4.49	420	1.51	81.4
48	Scarlette	13.6	11.7	5.95	2.00	55.0	81.2	3.75	420	1.74	80.9
49	Sebastian	13.9	10.7	3.80	1.80	69.0	81.3	3.74	260	1.55	81.1
50	BELLINI	13.9	10.5	4.27	2.00	73.0	82.5	3.93	370	1.51	82.5
51	NORD07-2403	13.9	11.0	4.51	2.20	70.0	81.7	4.07	360	1.51	82.5
52	KWS07-33	13.7	11.0	3.21	2.20	78.0	82.6	3.92	360	1.52	82.1
53	AC04-506-12-8	13.6	11.0	3.57	4.30	73.0	82.8	3.97	310	1.53	81.9
54	SYN408-177	13.7	11.0	3.56	2.30	66.0	80.7	3.34	310	1.56	80.1
55	180-02-5B	14.2	10.6	3.32	3.80	63.0	82.9	4.45	260	1.65	81.6
56	65-03NZ22G	14.0	10.5	3.46	1.50	91.0	81.5	4.34	380	1.46	83.4
57	62931 I 2	14.2	11.1	3.68	2.20	70.0	81.3	3.65	380	1.63	81.4
58	SB055559	13.8	10.0	3.58	1.60	80.0	81.9	3.84	400	1.49	83.1
59	UNS071152	13.6	10.2	3.57	1.60	83.0	81.6	3.79	270	1.52	81.4
60	S15501	13.9	10.8	3.32	1.60	70.0	82.0	3.91	290	1.56	82.5
61	PF12079-51	14.1	10.9	3.20	1.80	67.0	81.2	3.86	250	1.65	80.6
62	Scarlette	13.1	11.0	4.88	3.60	51.0	81.9	3.66	250	1.97	80.3
63	Sebastian	13.0	9.60	3.43	4.00	68.0	83.0	3.47	310	1.69	81.6
64	BELLINI	12.9	10.5	3.44	2.90	71.0	82.7	3.98	370	1.60	82.2
65	UNS071152	12.9	9.60	3.08	2.00	81.0	82.2	3.68	250	1.55	81.4
66	S15501	12.7	10.0	3.20	2.80	79.0	83.2	3.71	280	1.52	82.1
67	PF12079-51	12.5	9.70	3.20	2.00	69.0	82.2	3.61	220	1.64	79.5
68	AC04-506-12-8	11.8	12.0	4.27	4.40	76.0	82.5	4.76	430	1.51	82.1
69	NORD07-2403	13.0	10.2	3.68	3.50	73.0	82.2	3.99	300	1.60	82.1
70	KWS07-33	13.1	9.60	3.33	3.20	92.0	83.0	3.76	270	1.48	82.5
71	AC04-506-12-8	13.0	10.1	3.44	1.50	69.0	83.3	3.47	230	1.73	81.4
72	SYN408-177	12.6	9.80	3.92	3.50	74.0	82.5	3.34	260	1.58	80.6
73	180-02-5B	13.0	10.7	3.57	2.80	78.0	82.9	4.30	210	1.56	81.7
74	65-03NZ22G	12.8	9.90	3.44	3.10	94.0	81.6	4.19	250	1.45	83.0
75	62931 I 2	13.0	9.70	3.68	1.80	65.0	82.5	3.38	240	1.68	79.8
76	SB055559	13.0	9.00	3.56	2.80	83.0	83.5	3.65	320	1.49	82.2

ANNEXE 5 les paramètres de qualité du malt et d'orge de printemps récolte 2010

ID	Orge	Humidité d'orge	Protéine	β -gulcan	Arabinoxylan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité	Attenuation Limite
--	--	%	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s	%
1	Scarlette	15.2	10.4	4.73	3.76	89.0	38.5	4.07	250	1.53	80.9
2	Sebastian	14.8	9.70	3.23	2.61	98.0	83.2	4.07	340	1.48	82.5
3	BELLINI	14.7	9.90	3.11	2.49	96.0	83.6	4.45	380	1.48	82.2
4	NORD07-2403	14.7	9.80	3.11	2.16	82.6	82.3	4.30	320	1.46	97.0
5	KWS07-33	14.9	10.3	2.99	3.49	82.9	82.6	4.49	340	1.48	99.0
6	AC04-506-12-8	15.2	10.6	3.35	2.57	81.9	83.5	4.21	270	1.49	92.0
7	SYN408-177	14.8	9.90	3.48	1.08	81.2	83.6	3.96	360	1.49	98.0
8	180-02-5B	14.9	10.1	2.67	3.30	83.3	83.2	4.98	300	1.47	99.0
9	65-03NZ22G	14.8	9.50	3.33	2.53	83.4	82.6	4.91	290	1.46	99.0
10	62931 I 2	14.7	9.70	3.23	3.59	81.1	82.5	4.15	260	1.51	95.0
11	SB055559	14.5	9.20	3.00	3.61	82.1	83.7	4.35	330	1.48	96.0
12	UNS071152	14.4	9.70	3.00	3.71	81.2	82.2	4.10	260	1.55	81.0
13	S15501	14.4	9.10	2.66	3.19	83.4	83.7	4.22	300	1.48	99.0
14	PF12079-51	14.8	9.70	2.77	2.77	82.8	82.8	4.50	270	1.47	98.0
15	Scarlette	15.4	9.70	4.28	1.83	85.0	83.4	3.66	210	1.56	79.5
16	Sebastian	14.8	9.00	3.00	3.61	94.0	82.6	3.34	270	1.50	79.3
17	BELLINI	15.0	9.60	3.34	2.14	92.0	82.8	3.83	280	1.45	81.1
18	NORD07-2403	15.0	9.60	3.59	2.22	80.5	82.0	3.56	250	1.49	90.0
19	KWS07-33	15.3	9.30	3.24	3.16	79.5	82.7	3.56	220	1.46	99.0
20	AC04-506-12-8	15.5	8.90	3.92	3.02	79.7	83.0	3.27	200	1.50	93.0
21	SYN408-177	14.6	9.10	3.79	2.39	77.1	81.8	3.26	230	1.49	93.0
22	180-02-5B	15.3	9.30	3.01	3.73	80.9	82.7	3.95	240	1.46	98.0
23	65-03NZ22G	15.3	8.70	3.24	3.08	82.3	81.8	3.52	210	1.44	99.0
24	62931 I 2	14.7	9.30	3.22	1.71	79.8	82.5	3.38	240	1.68	65.0
25	SB055559	14.8	8.70	4.15	3.55	81.1	82.0	3.47	280	1.47	95.0
26	UNS071152	15.2	8.80	3.36	2.05	79.9	82.1	3.24	210	1.52	92.0
27	S15501	14.8	8.80	3.80	3.45	82.6	82.4	3.58	270	1.48	98.0
28	PF12079-51	15.0	9.20	3.01	2.44	79.5	82.3	3.69	190	1.50	94.0
29	Scarlette	13.6	11.9	4.60	3.48	79.0	81.7	4.33	300	1.51	81.0
30	Sebastian	13.2	11.1	3.80	3.61	86.0	77.5	4.22	380	1.48	79.9
31	BELLINI	13.2	10.5	3.79	4.30	91.0	78.4	4.38	350	1.46	80.7
32	NORD07-2403	13.4	11.1	3.80	2.67	80.5	76.9	4.59	350	1.45	90.0
33	KWS07-33	13.3	11.1	3.32	2.97	80.1	77.8	4.73	310	1.46	96.0
34	AC04-506-12-8	13.4	10.8	4.12	3.24	79.6	78.1	4.33	290	1.47	87.0
35	SYN408-177	13.4	10.6	3.77	3.80	78.9	77.8	4.03	260	1.48	91.0
36	180-02-5B	13.8	10.8	3.33	2.42	80.9	82.3	5.17	290	1.46	94.0
37	65-03NZ22G	13.5	10.9	3.43	3.09	81.9	81.7	4.63	280	1.45	96.0
38	62931 I 2	13.5	10.5	3.77	3.41	80.0	82.5	4.36	270	1.47	91.0
39	SB055559	13.7	10.1	3.56	2.31	81.2	82.7	4.89	320	1.44	90.0
40	UNS071152	13.5	10.3	3.70	3.32	80.6	81.7	3.97	260	1.49	86.0
41	S15501	13.5	11.0	3.32	2.20	83.2	82.4	4.99	360	1.45	90.0
42	PF12079-51	13.7	11.2	3.21	2.66	80.1	81.8	4.46	300	1.47	91.0
43	Scarlette	12.8	10.1	4.13	3.93	86.0	82.5	3.64	270	1.53	82.9
44	Sebastian	12.7	9.60	2.88	3.00	87.0	82.9	3.77	340	1.53	81.3
45	BELLINI	12.8	9.20	3.09	4.76	96.0	82.4	3.75	330	1.49	82.9
46	NORD07-2403	12.7	9.60	2.99	2.21	82.6	81.6	4.10	342	1.48	86.5
47	KWS07-33	12.8	9.50	2.87	3.77	83.8	82.9	4.17	327	1.48	93.0
48	AC04-506-12-8	13.0	9.30	3.00	3.30	81.7	83.2	3.91	281	1.52	83.4
49	SYN408-177	12.9	9.10	3.32	2.95	82.6	82.7	3.57	295	1.50	85.4
50	180-02-5B	13.0	9.70	2.88	2.58	84.6	83.3	4.69	285	1.50	88.1
51	65-03NZ22G	12.7	9.30	2.66	3.12	82.9	82.2	4.36	302	1.46	95.0
52	62931 I 2	12.8	9.40	2.89	2.71	82.0	82.6	3.74	293	1.54	83.0
53	SB055559	12.9	9.30	2.41	3.00	84.8	82.8	4.00	367	1.48	88.3
54	UNS071152	12.4	9.10	2.41	3.42	83.7	81.7	3.74	266	1.50	89.1
55	S15501	12.6	9.00	2.18	2.46	83.8	83.0	4.04	307	1.49	89.1
56	PF12079-51	12.5	9.70	2.76	3.17	80.9	82.2	4.06	283	1.52	85.3

ANNEXE 6 les paramètres de qualité du malt et d'orge d'hiver récolte 2010 (1/2)

ID	Orge	Humidité d'orge	Protéine	β -gulcan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité
--	--	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s
1	Vanessa	14.4	10.2	2.90	93.0	82.0	4.46	320	1.51
2	HEGW525-1	12.6	9.50	3.57	87.0	81.1	3.66	220	1.68
3	SY209-66	12.6	9.10	2.65	98.0	82.8	4.28	320	1.51
4	S-HD4480	12.3	9.90	3.58	84.0	82.2	3.71	280	1.71
5	UNS063654	13.1	10.1	2.89	97.0	81.9	4.82	260	1.51
6	SC16477IH122	13.0	10.1	3.01	92.0	82.8	4.14	290	1.57
7	NSL07-7425B	12.7	10.0	2.77	89.0	82.1	4.15	270	1.57
8	AC04-062-6	12.7	9.80	2.78	93.0	82.0	3.74	290	1.56
9	Vanessa	14.1	10.7	2.68	90.0	81.2	4.34	300	1.51
10	HEGW525-1	12.8	9.40	3.01	87.0	80.6	3.94	220	1.64
11	SY209-66	13.4	10.4	2.33	94.0	80.7	4.79	410	1.49
12	S-HD4480	14.1	11.0	3.15	84.0	80.1	4.19	330	1.55
13	UNS063654	13.7	10.6	2.79	94.0	80.8	5.05	350	1.50
14	SC16477IH122	13.7	10.9	2.79	88.0	81.4	4.32	370	1.54
15	NSL07-7425B	13.2	10.0	2.32	87.0	81.1	4.20	310	1.57
16	AC04-062-6	13.2	10.4	3.02	91.0	80.7	4.23	430	1.52
17	Vanessa	14.2	10.9	2.80	87.0	80.4	4.52	280	1.52
18	HEGW525-1	13.6	10.4	3.47	84.0	80.0	4.06	230	1.58
19	SY209-66	13.8	10.3	2.78	93.0	81.5	4.51	360	1.51
20	S-HD4480	14.7	10.2	3.73	78.0	81.1	3.43	260	1.64
21	UNS063654	13.6	10.2	2.91	96.0	81.2	4.67	270	1.49
22	SC16477IH122	14.6	9.90	3.27	90.0	81.9	3.31	300	1.54
23	NSL07-7425B	14.2	11.2	3.03	78.0	80.7	3.89	280	1.58
24	AC04-062-6	14.6	10.5	3.27	82.0	81.2	3.37	330	1.52
25	Vanessa	14.2	11.2	2.67	87.0	80.5	3.78	310	1.54
26	HEGW525-1	13.7	9.70	3.13	85.0	79.6	3.12	180	1.64
27	SY209-66	13.4	10.3	2.55	91.0	81.2	4.11	350	1.53
28	S-HD4480	13.6	10.1	3.37	86.0	80.9	3.26	280	1.57
29	UNS063654	13.8	10.1	3.02	92.0	81.0	3.85	290	1.39
30	SC16477IH122	13.3	11.0	3.25	90.0	81.0	3.62	340	1.37
31	NSL07-7425B	13.3	10.5	2.78	77.0	81.0	3.72	310	1.41
32	AC04-062-6	13.1	10.6	3.25	84.0	81.4	3.39	360	1.39
33	Vanessa	13.8	10.7	2.60	87.0	81.7	3.89	240	1.50
34	SYN208-51	12.9	10.2	2.86	71.0	81.9	4.39	230	1.59
35	Vanessa	13.2	11.4	2.61	75.0	80.2	4.60	360	1.51
36	SYN208-51	12.8	12.4	2.41	63.0	80.0	4.44	330	1.62
37	Vanessa	12.9	10.6	2.25	60.0	81.5	4.17	280	1.56
38	SYN208-51	12.8	10.9	2.52	65.0	81.0	4.22	300	1.67
39	Vanessa	13.8	11.8	2.90	59.0	80.4	3.62	390	1.55
40	SYN208-51	13.4	12.3	2.90	44.0	79.9	3.88	340	1.63
41	Esterel	14.1	11.2	2.67	82.0	78.2	4.42	240	1.57
42	CCB 09 W 907	14.0	11.7	2.44	78.0	77.7	5.46	480	1.53
43	S 34257	13.8	10.4	2.79	78.0	79.7	4.55	400	1.67
44	S 34196	13.5	10.7	2.44	87.0	79.7	4.63	390	1.55
45	UN 7118	14.2	11.6	2.78	62.0	75.6	4.06	320	1.65
46	MH 04 CL 16	14.1	10.8	3.01	82.0	79.4	4.42	330	1.55
47	MH 04 CG 28	13.3	10.6	2.77	82.0	78.6	4.27	340	1.60
48	MH 04 IC 28	13.7	11.0	3.36	67.0	78.9	4.43	360	1.66
49	SC 1205 IH 23	13.4	10.8	2.66	82.0	78.0	4.11	390	1.56
50	SC 1905-2	13.6	10.6	2.78	84.0	78.5	4.13	390	1.55
51	Esterel	13.5	9.20	3.02	86.0	81.3	3.44	250	1.62
52	CCB 09 W 907	13.5	8.80	3.14	88.0	80.5	4.09	270	1.59
53	S 34257	13.4	10.1	3.49	86.0	81.1	4.69	270	1.56
54	S 34196	13.7	9.10	3.15	93.0	81.6	3.85	270	1.60
55	UN 7118	13.2	9.40	3.83	66.0	78.6	3.33	190	1.73
56	MH 04 CL 16	13.4	9.20	3.71	86.0	80.9	3.68	220	1.55
57	MH 04 CG 28	13.3	9.50	3.48	85.0	79.6	3.44	240	1.60
58	MH 04 IC 28	13.3	9.00	4.29	83.0	80.2	3.22	210	1.61
59	SC 1205 IH 23	13.4	8.90	3.94	86.0	80.7	3.27	230	1.60
60	SC 1905-2	13.4	8.50	3.72	86.0	82.0	3.37	210	1.57

ANNEXE 7 les paramètres de qualité du malt et d'orge d'hiver récolte 2010 (2/2)

61	Esterel	13.8	9.20	3.14	84.0	80.5	3.52	300	1.67
62	CCB 09 W 907	13.6	9.30	2.89	89.0	80.7	4.45	330	1.56
63	S 34257	14.1	9.30	3.48	89.0	81.6	4.08	340	1.62
64	S 34196	13.7	9.00	2.78	95.0	81.7	3.92	330	1.56
65	UN 7118	13.1	9.30	3.92	68.0	78.3	3.15	230	1.87
66	MH 04 CL 16	13.8	9.10	4.17	92.0	81.6	3.77	260	1.55
67	MH 04 CG 28	13.8	9.10	3.70	87.0	80.5	3.37	260	1.70
68	MH 04 IC 28	14.2	9.40	4.40	83.0	81.2	3.53	270	1.64
69	SC 1205 IH 23	13.9	9.40	3.20	78.0	79.6	3.45	280	1.65
70	SC 1905-2	13.9	9.10	3.80	79.0	79.9	3.47	280	1.65
71	Esterel	14.1	10.0	3.80	77.0	80.5	3.64	310	1.73
72	CCB 09 W 907	13.4	9.80	3.20	78.0	80.4	4.17	350	1.61
73	S 34257	13.9	9.00	3.80	88.0	81.0	3.64	270	1.65
74	S 34196	13.9	9.80	3.50	85.0	81.6	3.85	310	1.67
75	UN 7118	13.8	9.50	3.90	64.0	78.5	3.10	200	1.89
76	MH 04 CL 16	13.9	9.80	4.41	83.0	80.6	3.45	240	1.59
77	MH 04 CG 28	13.1	9.00	4.16	86.0	81.1	3.11	250	1.68
78	MH 04 IC 28	13.6	10.3	4.99	70.0	79.5	3.25	220	1.69
79	SC 1205 IH 23	13.6	9.40	3.95	86.0	80.5	2.78	260	1.68
80	SC 1905-2	13.6	9.20	3.37	85.0	80.6	3.05	270	1.67
81	Esterel	13.1	10.9	3.74	71.0	79.1	3.77	330	1.57
82	83113EH	12.7	10.9	4.20	62.0	79.4	3.61	310	1.71
83	MH03BZ29	12.9	10.9	4.67	70.0	77.9	3.88	270	1.57
84	MH03DJ54	13.3	10.8	3.39	81.0	79.5	4.52	330	1.52
85	Esterel	11.8	10.5	3.86	66.0	80.2	3.64	300	1.68
86	83113EH	11.9	11.4	4.20	83.0	81.6	3.72	250	1.65
87	MH03BZ29	11.7	10.6	3.86	79.0	79.3	3.74	250	1.63
88	MH03DJ54	11.9	10.4	3.83	82.0	80.8	4.70	240	1.57
89	Esterel	12.5	10.5	4.09	71.0	79.8	3.65	300	1.61
90	83113EH	11.5	10.4	3.96	63.0	79.9	3.81	300	1.70
91	MH03BZ29	12.0	10.7	3.84	76.0	78.7	3.36	260	1.58
92	MH03DJ54	12.0	10.8	4.06	75.0	81.2	3.99	270	1.58
93	Esterel	12.4	11.1	3.86	69.0	79.3	4.03	330	1.56
94	83113EH	12.2	9.60	4.20	77.0	81.1	3.50	240	1.62
95	MH03BZ29	12.3	9.60	3.96	84.0	78.6	3.80	210	1.58
96	MH03DJ54	12.7	9.70	2.99	91.0	81.2	1.69	260	1.54
97	Esterel	14.0	8.60	3.35	98.0	81.9	3.59	300	1.53
98	S34196	13.8	8.00	3.03	99.0	82.2	3.67	170	1.49
99	Esterel	13.5	10.2	2.56	84.0	80.9	4.08	380	1.60
100	S34196	12.9	10.2	2.71	77.0	80.1	3.63	370	1.61
101	Esterel	14.6	10.4	3.09	81.0	80.4	3.92	310	1.61
102	S34196	13.9	9.80	2.70	89.0	81.0	3.99	340	1.57
103	Esterel	13.8	9.70	3.02	86.0	81.1	3.60	320	1.65
104	S34196	15.0	9.70	3.21	89.0	81.9	4.42	330	1.58
105	Vanessa	12.7	9.20	2.91	99.0	82.7	4.04	260	1.47
106	SZD-A7619	12.9	8.90	3.84	99.0	82.6	3.36	230	1.51
107	Vanessa	12.6	9.80	3.26	90.0	81.5	4.44	240	1.49
108	SZD-A7619	12.2	9.30	3.61	93.0	81.9	3.32	220	1.64
109	Vanessa	11.8	11.1	3.26	85.0	80.9	4.29	350	1.55
110	SZD-A7619	12.0	10.4	3.85	85.0	81.3	3.43	260	1.65
111	Vanessa	13.5	10.5	3.26	92.0	80.7	4.49	210	1.52
112	SZD-A7619	13.3	10.0	4.08	87.0	82.0	3.74	250	1.65

ANNEXE 8 les paramètres de qualité du malt et d'orge de printemps récolte 2011 (1/2)

ID	Orge	Humidité d'orge	Protéine	β-gulcan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité
--	--	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s
1	Sebastian	14.4	11.5	3.03	79.0	82.5	3.98	350	1.51
2	Bellini	14.3	12.3	3.27	72.0	81.9	4.83	400	1.47
3	SC65-03NZ7C	14.6	10.4	3.40	91.0	82.8	4.51	280	1.47
4	KWS09-328	14.5	9.70	3.63	95.0	83.8	3.46	260	1.48
5	SC15269K3	14.3	11.4	4.21	89.0	81.5	4.42	320	1.47
6	SC3883K2	14.1	11.6	3.73	88.0	81.2	4.79	310	1.47
7	KWS09-410	14.6	12.3	3.04	80.0	81.9	4.18	270	1.47
8	MH06NI34	14.6	12.7	3.75	84.0	80.5	5.00	310	1.50
9	NSL08-2649B	14.5	11.6	3.28	73.0	81.6	3.83	260	1.49
10	SER35131-16	14.5	11.6	3.99	82.0	82.1	4.17	330	1.52
11	LSB0315-5	14.4	11.8	3.16	75.0	81.8	4.49	350	1.48
12	AC06-509-41-5	14.4	11.0	3.98	80.0	80.6	3.49	230	1.53
13	AC07-624-34	14.2	10.3	3.51	93.0	82.6	3.67	250	1.47
14	UNS108036	14.4	10.4	3.63	92.0	83.4	4.58	370	1.45
15	UNS108077	14.7	11.1	3.63	94.0	82.3	4.66	280	1.48
16	NORD09-2421	14.5	11.0	3.27	92.0	82.9	4.04	260	1.50
17	SY409-226	14.4	10.8	3.63	90.0	83.1	4.46	280	1.47
18	SY410-244	14.3	11.4	3.63	82.0	81.7	4.11	330	1.46
19	Sebastian	12.7	11.0	3.01	77.0	84.4	4.10	290	1.69
20	Bellini	12.4	10.8	3.12	76.0	84.2	4.02	270	1.66
21	SC65-03NZ7C	12.5	10.9	3.47	81.0	85.1	4.58	220	1.67
22	KWS09-328	12.4	10.3	3.58	85.0	84.9	4.52	260	1.60
23	SC15269K3	12.6	10.7	3.93	82.0	83.4	4.09	230	1.69
24	SC3883K2	12.2	10.9	3.46	79.0	83.2	4.47	210	1.70
25	KWS09-410	12.6	10.7	2.66	85.0	85.2	4.04	190	1.66
26	MH06NI34	12.4	10.3	3.00	85.0	84.0	4.36	190	1.73
27	NSL08-2649B	12.7	10.5	2.77	78.0	85.0	3.95	210	1.71
28	SER35131-16	12.5	10.3	3.81	83.0	85.8	3.83	210	1.79
29	LSB0315-5	12.4	10.7	2.65	78.0	83.7	3.94	260	1.68
30	AC06-509-41-5	12.4	11.2	3.11	64.0	83.0	3.92	210	1.99
31	AC07-624-34	12.4	10.2	3.00	79.0	84.2	3.81	240	1.65
32	UNS108036	12.5	10.6	3.00	90.0	85.1	4.72	330	1.55
33	UNS108077	12.5	10.7	3.00	88.0	84.8	4.90	230	1.65
34	NORD09-2421	12.4	10.7	2.53	82.0	84.6	4.15	190	1.68
35	SY409-226	12.4	10.2	2.54	85.0	85.5	4.30	200	1.60
36	SY410-244	12.5	10.8	2.66	85.0	84.9	4.10	220	1.61
37	Sebastian	14.2	10.9	2.92	72.0	81.6	3.41	250	1.53
38	Bellini	14.1	11.0	2.80	72.0	82.0	3.71	260	1.51
39	SC65-03NZ7C	14.5	9.90	3.04	86.0	82.0	3.73	200	1.49
40	KWS09-328	14.3	9.80	3.14	86.0	83.1	3.33	180	1.52
41	SC15269K3	14.1	10.3	3.62	87.0	81.7	3.50	180	1.51
42	SC3883K2	14.1	10.0	3.38	80.0	81.2	3.15	150	1.56
43	KWS09-410	14.5	9.90	3.04	87.0	82.8	3.37	140	1.54
44	MH06NI34	14.4	10.4	2.80	83.0	80.8	4.06	160	1.51
45	NSL08-2649B	13.9	9.90	2.91	79.0	82.3	3.38	140	1.55
46	SER35131-16	14.1	10.2	3.15	86.0	82.2	3.76	220	1.51
47	LSB0315-5	14.1	10.4	2.57	78.0	81.9	3.56	250	1.51
48	AC06-509-41-5	14.3	11.0	3.38	76.0	80.4	3.39	170	1.54
49	AC07-624-34	14.3	10.0	3.80	83.0	81.6	3.32	150	1.50
50	UNS108036	14.5	9.60	3.60	94.0	83.1	4.04	240	1.46
51	UNS108077	14.4	10.4	4.10	93.0	82.6	4.02	220	1.48
52	NORD09-2421	14.0	10.5	3.30	82.0	82.3	3.53	160	1.52
53	SY409-226	14.3	10.0	2.50	91.0	82.9	3.50	210	1.47
54	SY410-244	14.0	10.3	3.80	81.0	81.7	3.31	200	1.49
55	Sebastian	14.6	10.3	2.80	81.0	83.2	4.44	330	1.58
56	Bellini	14.7	10.5	3.30	83.0	83.4	4.36	280	1.56
57	SC65-03NZ7C	14.8	10.2	2.94	94.0	83.0	4.32	200	1.53
58	KWS09-328	14.5	9.10	2.96	91.0	83.9	3.61	250	1.54
59	SC15269K3	14.7	9.60	3.64	92.0	83.1	3.78	230	1.57
60	SC3883K2	14.6	10.0	2.95	89.0	81.9	4.31	250	1.57
61	KWS09-410	14.3	9.60	4.23	83.0	84.1	3.21	190	1.67
62	MH06NI34	14.7	10.3	4.10	87.0	82.7	4.50	230	1.59
63	NSL08-2649B	14.3	11.1	3.86	65.0	81.7	3.44	200	1.58
64	SER35131-16	14.5	10.0	4.32	85.0	83.3	3.85	230	1.63
65	LSB0315-5	14.5	10.3	4.09	83.0	82.3	3.80	240	1.59
66	AC06-509-41-5	14.6	9.70	4.45	80.0	82.2	3.79	230	1.66
67	AC07-624-34	14.6	9.40	4.23	85.0	83.1	3.44	200	1.60
68	UNS108036	14.6	9.50	4.23	96.0	84.6	4.19	290	1.50
69	UNS108077	14.9	10.0	4.23	94.0	82.9	4.35	250	1.56

ANNEXE 9 les paramètres de qualité du malt et d'orge de printemps récolte 2011 (2/2)

70	NORD09-2421	14.8	9.70	4.11	89.0	84.2	4.42	240	1.58
71	SY409-226	14.9	9.70	4.23	86.0	83.6	3.98	250	1.57
72	SY410-244	14.9	9.90	4.22	89.0	83.4	4.03	280	1.49
73	Sebastian	14.0	11.3	2.75	80.0	82.1	3.74	350	1.51
74	Bellini	13.9	11.9	2.86	74.0	82.3	4.51	380	1.47
75	KWS09330	14.0	10.3	2.12	97.0	83.1	4.18	370	1.47
76	SC28317K2	13.9	11.7	2.47	71.0	81.4	4.67	270	1.48
77	KWS10-206	14.0	12.2	2.91	71.0	80.7	4.20	290	1.51
78	LN1075	14.1	12.9	3.60	70.0	80.8	4.21	310	1.51
79	MH065572-2	14.3	10.9	3.10	87.0	82.0	3.98	300	1.47
80	AC07-551-181	14.1	10.6	2.25	93.0	82.2	4.17	230	1.48
81	NORD09-1113	14.0	11.1	2.39	85.0	82.5	3.78	240	1.47
82	UN-C514	13.9	10.5	2.60	79.0	82.0	3.58	300	1.52
83	NORD09-2531	14.1	10.1	2.68	94.0	82.2	4.26	180	1.46
84	BR10168G2	14.0	10.7	3.34	85.0	82.2	3.79	270	1.57
85	SY409-228	14.3	10.3	2.26	86.0	82.9	3.99	250	1.52
86	Sebastian	12.8	10.7	2.00	80.0	84.4	4.29	280	1.67
87	Bellini	13.0	10.4	2.04	82.0	84.4	4.38	250	1.61
88	KWS09330	13.0	10.0	1.95	96.0	85.1	4.69	230	1.53
89	SC28317K2	12.8	10.6	2.31	79.0	84.3	4.52	240	1.65
90	KWS10-206	12.8	10.0	2.69	81.0	84.4	3.87	200	1.74
91	LN1075	13.0	10.4	3.47	77.0	84.3	3.84	220	1.82
92	MH065572-2	13.0	10.7	2.66	79.0	83.6	3.92	260	1.70
93	AC07-551-181	13.1	10.4	2.19	92.0	83.6	4.36	200	1.60
94	NORD09-1113	13.1	10.8	2.46	85.0	84.7	4.30	200	1.66
95	UN-C514	13.2	10.2	2.27	80.0	83.6	3.73	250	1.69
96	NORD09-2531	13.0	10.5	2.78	88.0	83.4	4.23	160	1.61
97	BR10168G2	13.1	10.4	2.99	72.0	84.4	3.60	200	2.14
98	SY409-228	13.1	10.1	2.62	80.0	85.3	3.98	190	1.73
99	Sebastian	14.2	10.7	2.86	80.0	81.9	3.40	230	1.54
100	Bellini	14.4	10.8	2.59	78.0	82.1	3.66	260	1.52
101	KWS09330	14.4	10.0	2.64	93.0	82.6	4.10	210	1.47
102	SC28317K2	14.4	10.3	3.16	79.0	82.7	3.63	240	1.50
103	KWS10-206	14.4	10.5	3.55	78.0	81.8	3.20	210	1.51
104	LN1075	14.2	10.0	3.03	80.0	82.8	3.20	170	1.55
105	MH065572-2	14.3	10.5	3.57	82.0	81.6	3.44	200	1.53
106	AC07-551-181	14.4	9.90	2.90	92.0	82.2	3.72	190	1.48
107	NORD09-1113	14.2	10.4	4.09	85.0	82.4	3.37	190	1.52
108	UN-C514	14.3	10.5	3.85	76.0	82.3	3.80	290	1.50
109	NORD09-2531	14.1	10.3	4.43	90.0	81.8	3.93	140	1.47
110	BR10168G2	14.3	10.5	4.67	82.0	82.4	3.55	220	1.61
111	SY409-228	14.0	9.70	4.43	85.0	82.8	3.21	180	1.55
112	Sebastian	14.5	9.60	4.10	84.0	84.0	3.93	320	1.56
113	Bellini	14.7	9.10	3.98	88.0	84.4	4.07	300	1.56
114	KWS09330	14.8	8.90	3.99	91.0	84.2	4.03	250	1.54
115	SC28317K2	14.6	9.30	4.22	90.0	84.1	4.05	270	1.53
116	KWS10-206	14.6	9.50	4.32	86.0	83.7	3.81	220	1.63
117	LN1075	14.7	9.50	3.97	81.0	83.7	3.45	210	1.71
118	MH065572-2	14.6	9.70	3.75	82.0	83.6	3.40	230	1.66
119	AC07-551-181	14.7	9.10	3.28	95.0	83.7	3.66	180	1.55
120	NORD09-1113	14.8	9.60	4.11	86.0	84.0	3.39	170	1.61
121	UN-C514	14.5	9.40	4.11	76.0	83.0	3.23	230	1.72
122	NORD09-2531	14.5	9.40	4.56	97.0	83.1	3.87	140	1.52
123	BR10168G2	14.5	9.50	5.62	83.0	83.6	3.61	200	1.76
124	SY409-228	14.7	9.50	4.22	86.0	84.4	3.55	220	1.60

ANNEXE 10 les paramètres de qualité du malt et d'orge d'hiver récolte 2011

ID	Orge	Humidité d'orge	Protéine	β -gulcan	Friabilité	Extrait	Protéine soluble	Pouvoir Diastatic	Viscosité
--	--	%	%	%	%	% m.s	%m.s	WK m.s	mPa.s
1	Vanessa	12.2	13.2	2.72	83.0	79.9	4.96	450	1.53
2	SZD-A7619	12.2	12.3	2.74	86.0	79.9	3.86	380	1.57
3	Vanessa	12.5	10.6	3.34	93.0	81.8	4.42	330	1.52
4	SZD-A7619	12.5	10.6	3.92	94.0	81.2	3.87	340	1.56
5	Vanessa	12.0	10.5	3.45	93.0	80.7	4.74	280	1.49
6	SZD-A7619	11.8	8.90	3.67	99.0	81.3	3.39	240	1.52
7	Vanessa	11.2	10.0	3.67	91.0	82.8	4.14	270	1.64
8	SZD-A7619	11.0	10.7	3.90	78.0	82.1	3.68	270	1.78
9	Esterel	12.5	9.30	4.27	89.0	80.6	3.46	310	1.67
10	Esterel	12.1	11.2	3.18	81.0	79.5	4.04	380	1.65
11	Esterel	10.6	10.5	4.36	72.0	81.0	3.68	300	1.93
12	Esterel	11.8	9.60	4.36	93.0	79.4	4.23	320	1.47

ANNEXE 11 les mesures de humidités et de la porosité de 48 orges de printemps récolte 2007-2008

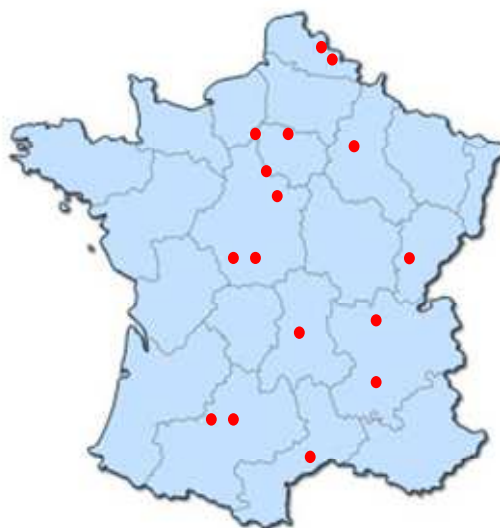
Variété	Année	Humidité %	Porosité %
Concerto-001	2008	15.96	7.01
Henley-020	2008	16.20	4.48
Victoriana-005	2008	15.75	6.89
Sebastian-006	2008	16.77	6.20
Henley007	2008	15.46	6.88
Signora-010	2008	15.62	7.31
Cellar -013	2008	15.65	7.13
Throgall-016	2008	16.04	3.18
Calcule-017	2008	15.10	5.98
Prestige-018	2008	16.56	4.60
Sebastian-019	2008	15.60	6.26
Tipple-021	2008	15.73	3.73
Beatrix-011	2008	12.51	5.54
Azalea-014	2008	12.69	5.78
1009796-1	2008	14.55	8.30
1009796-3	2008	16.55	9.70
1009796-8	2008	16.62	8.71
1009796-9	2008	13.11	8.42
1009796-12	2008	16.63	8.68
1009796-15	2008	16.12	7.52
1009796-16	2008	12.91	8.45
1009796-20	2008	15.29	9.15
1009796-24	2008	15.29	8.23
1009796-26	2008	15.92	8.49
1009782-1	2007	13.25	5.15
1009782-3	2007	14.46	5.96
1009782-6	2007	14.07	3.85
1009782-7	2007	14.44	4.38
1009782-8	2007	13.94	5.60
1009782-9	2007	15.02	5.37
1009782-12	2007	14.66	5.65
1007782-15	2007	14.70	5.01
1009782-16	2007	12.99	4.09
1009782-20	2007	15.09	5.56
1009782-24	2007	14.86	4.55
1009782-26	2007	15.19	5.41
1008664-1	2008	14.73	8.47
1008664-4	2008	17.31	7.02
1008664-6	2008	14.73	7.47
1008664-9	2008	15.01	8.06
1008664-10	2008	14.63	8.26
1008664-11	2008	14.63	8.40
1008664-15	2008	15.64	8.04
1008664-19	2008	15.42	6.78
1008664-22	2008	12.17	8.22
1008664-28	2008	14.78	8.93
1008664-37	2008	14.76	7.09
1008664-42	2008	14.36	8.61

ANNEXE 12 Liste de 72 variétés d'orge utilisées dans cette étude et cultivées en France
en 2009, 2010, 2011

Variétés d'orge			
1	Vanessa	37	AC07-624-34
2	Esterel	38	UNS108036
3	Scarlette	39	UNS108077
4	Sebastian	40	NORD09-2421
5	BELLINI	41	SY409-226
6	15633 FH	42	SY410-244
7	AC 03-271-4	43	HEGW525-1
8	MH 02 BS 10	44	SY209-66
9	CCB 07 W 901	45	S-HD4480
10	MH 02 GG 78	46	UNS063654
11	SYN407143	47	SC16477IH122
12	LDSTRG67804	48	NSL07-7425B
13	LP1057604	49	AC04-062-6
14	UNS071085	50	SYN208-51
15	UNB8701D4	51	CCB 09 W 907
16	26141H1	52	S 34257
17	9021G	53	S 34196
18	30202G	54	UN 7118
19	NORD07-2403	55	MH 04 CL 16
20	KWS07-33	56	MH 04 CG 28
21	AC04-506-12-8	57	MH 04 IC 28
22	SYN408-177	58	SC 1205 IH 23
23	180-02-5B	59	SC 1905-2
24	65-03NZ22G	60	83113EH
25	62931 I 2	61	MH03BZ29
26	SB055559	62	MH03DJ54
27	UNS071152	63	S34196
28	S15501	64	SZD-A7619
29	PF12079-51	65	SC65-03NZ7C
30	SC3883K2	66	KWS09-328
31	KWS09-410	67	SC15269K3
32	MH06NI34	68	KWS09330
33	NSL08-2649B	69	SC28317K2
34	SER35131-16	70	KWS10-206
35	LSB0315-5	71	LN1075
36	AC06-509-41-5	72	MH065572-2

ANNEXE 13 la carte et le nom des 16 lieux de culture pour les échantillons d'orge qu'on a utilisé dans cette étude

Lieu de cultivar
RIVERES
MONTPELLIER
ABLIS
ATTRAY
AUZEVILLE TOLOSANE
BRETENIERES
BUZANCAIS
CAPPELLE EN PEVELE
CLERMONT FERRAND
ESTREES ST DENIS
ISSOUDUN
LENS LESTANG
LES GRANDES LOGES
MAULE
MERIGNIES
MONBEQUI



ANNEXE 14 Algorithme de NIPALS pour calculer la PCA

L'algorithme itératif NIPALS pour calculer la PCA 17 est présenté ci-dessous. Le vecteur des coordonnées factorielles $\mathbf{t}$ et les vecteurs propres $\mathbf{p}$ sont déterminés de façon à minimiser la somme des carrés des résidus $\mathbf{E}$. Une contrainte doit être observée, c'est-à-dire que la norme de $\mathbf{p}$ soit égale à 1. Les PCs sont calculées successivement l'une après l'autre. En considérant une matrice de données $\mathbf{X}$ prétraitée de la façon souhaitée, et centrée par variable:

- (1) $\mathbf{t} = \mathbf{x}_i$ Définir initialement $\mathbf{t}$ égal à une colonne de la matrice $\mathbf{X}$
- (2) $\mathbf{p} = (\mathbf{X}^T \cdot \mathbf{t}) / (\mathbf{t}^T \cdot \mathbf{t})$ Calculer le poids de chaque variable de $\mathbf{X}$ dans cette PC par projection de chaque colonne de $\mathbf{X}$ sur $\mathbf{t}$
- (3) $\mathbf{p} = \mathbf{p} / \|\mathbf{p}\|$ Normaliser le vecteur des poids pour chaque variable
- (4) $\mathbf{t}_{\text{ancien}} = \mathbf{t}$ Garder cette valeur de $\mathbf{t}$ pour comparaison ultérieure
- (5) $\mathbf{t} = (\mathbf{X} \cdot \mathbf{p}) / (\mathbf{p}^T \cdot \mathbf{p})$ Calculer la projection de chaque échantillon sur cette PC
- (6) $d = \|\mathbf{t} - \mathbf{t}_{\text{ancien}}\|$ Vérifier la convergence. Si $\mathbf{t}$ ne change plus ($d < \text{seuil}$), on continue vers l'étape (7), sinon on revient à (2)
- (7) $\mathbf{E} = \mathbf{X} - \mathbf{t} \cdot \mathbf{p}^T$ On enlève la partie de $\mathbf{X}$ expliquée par les vecteurs $\mathbf{t}$ et $\mathbf{p}$ calculées
Pendant cette itération
- (8) $\mathbf{X} = \mathbf{E}$ Pour calculer un nouveau PC, on met $\mathbf{X}$ égal aux résidus, et on recommence à partir de (1)

Les variances des PCs peuvent finalement être obtenues sur la diagonale de $\mathbf{S} = \mathbf{T}^T \cdot \mathbf{T}$.

ANNEXE 15 L'algorithme pour le calcul de la régression PLS1 est présenté ci-dessous. Les composantes principales sont calculées successivement jusqu'au nombre souhaitée:

- | | |
|--|---|
| (1) $\mathbf{y} = \mathbf{y}_i$ | Définir initialement $\mathbf{y}$ égal est la vecteur y de réponse |
| (2) $\mathbf{w} = \mathbf{y}^T \cdot \mathbf{X} / \ \mathbf{y}^T \cdot \mathbf{X}\ $ | Les poids de chaque colonne de $\mathbf{X}$ dans cette composante principale sont calculés |
| (3) $\mathbf{w}_k = \mathbf{w}$ | (alternativement, on peut garder $\mathbf{w}$ pour vérifier la convergence) |
| (4) $\mathbf{w}_k = \mathbf{w}_k / \ \mathbf{w}_k\ $ | Les poids sont normalisés |
| (5) $\mathbf{t}_k = \mathbf{X} \cdot \mathbf{w}_k$ | Projections des objets de $\mathbf{X}$ sur l'espace du vecteur $\mathbf{w}$ |
| (6) $\mathbf{t}_k = \mathbf{t}_k / \ \mathbf{t}_k\ $ | les scores de PLS sur $\mathbf{X}$ sont normalisée |
| (7) $\mathbf{q}_k = \mathbf{t}_k^T \cdot \mathbf{Y}$ | Les poids de chaque colonne de $\mathbf{Y}$ dans ce LV sont calculés |
| (8) $\mathbf{y} = \mathbf{Y} \cdot \mathbf{q}_k / \ \mathbf{q}_k^T \cdot \mathbf{q}_k\ $ | Projections des objets de $\mathbf{Y}$ sur l'espace du vecteur $\mathbf{q}$ |
| (9) $d = \ \mathbf{y} - \mathbf{y}_k\ $ | Vérifier la convergence. Si $\mathbf{y}$ (ou $\mathbf{w}$) ne change plus ($d < \text{seuil}$), on continue vers l'étape (9) sinon on revient à (2) |
| (9) $\mathbf{p}^T = \mathbf{t}_k^T \cdot \mathbf{X}$ | On calcule le vecteur $\mathbf{p}$ de contributions factorielles de $\mathbf{X}$ |
| (10) $\mathbf{E} = \mathbf{X} - \mathbf{t} \cdot \mathbf{p}^T$ | On soustrait la variance de $\mathbf{X}$ calculée pendant l'itération |
| (11) $\mathbf{F} = \mathbf{Y} - \mathbf{t} \cdot \mathbf{q}^T$ | On soustrait la variance de $\mathbf{Y}$ calculée pendant l'itération |
| (12) $\mathbf{X} = \mathbf{E}$ et $\mathbf{Y} = \mathbf{F}$ | On garde tous les vecteurs $\mathbf{t}$, $\mathbf{y}$, $\mathbf{p}$, $\mathbf{w}$, $\mathbf{q}$. On met les matrices $\mathbf{X}$ et $\mathbf{Y}$ égales aux résidus respectivement, et on recommence à partir de (1) pour une nouvelle variable latente |

À partir des vecteurs obtenus, on peut calculer la matrice des coefficients de régression B , selon l'équation :

$$B = W \cdot Q^t$$

La prédiction des nouvelles objets est obtenue par :

$$Y_{\text{estimé}} = X \cdot B + E$$

ANNEXE 16 Compilation des coefficients des modèles significatifs et des coefficients de détermination pour l'orge de printemps

Indicateur	Modèle	N°exp	R ²	a ₀	P	T	IG	WR1	AR1	WR2	BG	Ara	Interac.	Quad.
HFT	1	21	0.85			0.23	0.6		0.29	0.54				
	2	21	0.85	0.21		0.23			0.29	0.55				0.43*IG ²
	3	21	0.92			0.22	0.9		0.27	0.59				0.63*IG ²
Fria	1	21	0.8		-0.17	0.23	0.56		0.26				0.42(WR2*BG)	0.53*(WR2) ²
ProtSol	1	21	0.83		-0.31	0.07		0.22	0.09				0.70(P*IG)	(-0.20)*P ²
	2	21	0.72	-0.6	-0.07	0.06	0.68			0.14				
Extr	1	21	0.84	-0.57	-0.07		0.84				-0.18			
BG	1	19	0.77	-0.93		-0.06			-0.06	-0.07				
Pdias	1	21	0.78		0.13	0.07					-0.12	-0.94	(IG*AR1)	
HF	1	21	0.95			0.19	0.87		0.21	0.44				(-0.22)*BG ²
	2	21	0.96	0.26		0.19	0.43		0.21	0.43				

ANNEXE 17 Compilation des coefficients des modèles significatifs et des coefficients de détermination pour l'orge d'hiver

Indicateur	Modèle	N°exp	R ²	a ₀	P	T	IG	WR1	AR1	WR2	BG	Ara	Interac.	Quad.
HFT	1	19	0.87			0.22		0.78	0.28		-0.5			
	2	19	0.97			0.22		1.01	-0.54		-0.58		1.31(WR1*AR1)	
Fria	1	19	0.94			0.35			0.5	0.93	-0.9			
	2	19	0.95		-0.08	0.34			0.5	0.93	-0.97			
ProtSol	1	19	0.72	-0.13	0.07	0.05	-0.2							
	2	19	0.93	-0.47		0.06	3.58	0.24		0.08			0.09P*T	(-3.47)*IG ²
Extr	1	19	0.72	-0.28		0.07	-0.14			-0.07				
Visco	1	16	0.82	-0.29		-0.07	-0.15		-0.06	-0.12				
	2	16	0.77	-0.35		-0.09			0.1	-0.29				(-0.17)*IG ²
BG	1	16	0.9	-0.76		-0.11	-0.2		-0.06	-0.13			(-1.05)P*WR2	
Pdias	1	19	0.93		0.09	0.08	0.68			0.31				
AttLim	1	19	0.73			0.09		0.37	0.66		-1.02	0.21		
	2	19	0.82			0.08		0.46	0.66		-1.1	0.21	(-0.11)WR1*AR1	
GP	1	19	0.92	0.28		0.2	0.77		0.67	0.66				
	2	19	0.92			0.19	0.76		0.15	0.67	-0.49			
H _{j+1}	1	19	0.94			0.2			0.09	0.55	-0.97		0.84WR1*AR1	
	2	19	0.94			0.2		0.46		0.29	-1.02			
	3	19	0.96			0.2				-0.68	-0.77			(-1.29)WR2 ²
HF	1	19	0.84			0.22				0.39	-0.51	0.29		
	2	19	0.9			0.19				0.42	-0.54	0.32	(-0.18)WR1*AR1	
	3	19	0.95			0.18				0.28	-0.63	0.23	(-0.14)AR1*BG	(-0.34)AR1 ²

Résumé: Dans un marché en permanente progression et pour répondre aux besoins des brasseurs en malt de qualité, la maîtrise du procédé de maltage est indispensable. La qualité du malt est fortement dépendante des conditions opératoires, en particulier des conditions de trempage, mais également de la qualité de la matière première : l'orge. Dans cette étude, nous avons établi des modèles polynomiaux qui mettent en relation les conditions opératoires et la qualité du malt. Ces modèles ont été couplés à nos algorithmes génétiques et nous ont permis de déterminer les conditions optimales de maltage, soit pour atteindre une qualité ciblée de malt (friabilité), soit pour permettre un maltage à faible teneur en eau (pour réduire la consommation en eau et maîtriser les coûts environnementaux de production) tout en conservant une qualité *acceptable* de malt. Cependant, la variabilité de la matière première est un facteur limitant de notre approche. Les modèles établis sont en effet très sensibles à l'espèce d'orge (printemps, hiver) ou encore à la variété d'orge utilisée. Les modèles sont surtout très dépendants de l'année de récolte. Les variations observées sur les propriétés d'une année de récolte à une autre sont mal caractérisées et ne sont donc pas intégrées dans nos modèles. Elles empêchent ainsi de capitaliser l'information expérimentale au cours du temps. Certaines propriétés structurelles de l'orge (porosité, dureté) ont été envisagées comme nouveaux facteurs pour mieux caractériser la matière première mais ils n'ont pas permis d'expliquer les variations observées en malterie. Afin de caractériser la matière première, 394 échantillons d'orge issus de 3 années de récolte différentes 2009-2010-2011 ont été analysés par spectroscopie MIR. Les analyses ACP ont confirmé l'effet notable des années de récolte, des espèces, des variétés voire des lieux de culture sur les propriétés de l'orge. Une régression PLS a permis, pour certaines années et pour certaines espèces, de prédire les teneurs en protéines et en bêta-glucanes de l'orge à partir des spectres MIR. Cependant, ces résultats, pourtant prometteurs, se heurtent toujours à la variabilité. Ces nouveaux modèles PLS peuvent toutefois être exploités pour mettre en place des stratégies de pilotage du procédé de maltage à partir de mesures spectroscopiques MIR.

Abstract: In a continuously growing market and in order to meet the needs of Brewers in high quality malt, control of the malting process is a great challenge. Malt quality is highly dependent on the malting process operating conditions, especially on the steeping conditions, but also the quality of the raw material: barley. In this study, we established polynomial models that relate the operating conditions and the malt quality. These models have been coupled with our genetic algorithms to determine the optimal steeping conditions, either to obtain a targeted quality of malt (friability), or to allow a malting at low water content while maintaining *acceptable* quality of malt (to reduce water consumption and control the environmental costs of malt production). However, the variability of the raw material is a limiting factor for our approach. Established models are very sensitive to the species (spring and winter barley) or to the barley variety. The models are especially highly dependent on the crop year. Variations on the properties of a crop from one to another year are poorly characterized and are not incorporated in our models. They thus prevent us to capitalize experimental information over time. Some structural properties of barley (porosity, hardness) were considered as new factors to better characterize barley but they did not explain the observed variations. To characterize barley, 394 samples from 3 years of different crops 2009-2010-2011 were analysed by MIR spectroscopy. ACP analyses have confirmed the significant effect of the crop-years, species, varieties and sometimes of places of harvest on the properties of barley. A PLS regression allowed, for some years and for some species, to predict content of protein and beta-glucans of barley using MIR spectra. These results thus still face product variability, however, these new PLS models are very promising and could be exploited to implement control strategies in malting process using MIR spectroscopic measurements.