



HAL
open science

Identification des critères de qualité déterminant les aptitudes à l'emploi de grumes de hêtre (*Fagus sylvatica* L.) par un panel de professionnels

Fleur Longuetaud

► To cite this version:

Fleur Longuetaud. Identification des critères de qualité déterminant les aptitudes à l'emploi de grumes de hêtre (*Fagus sylvatica* L.) par un panel de professionnels. Sciences du Vivant [q-bio]. 2001. <hal-02833515>

HAL Id: hal-02833515

<https://hal.inrae.fr/hal-02833515v1>

Submitted on 7 Jun 2020

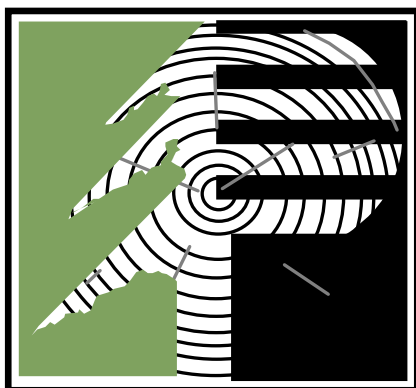
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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



HAL Authorization

D.E.A. Biostatistiques



**Identification des critères de qualité
déterminant les aptitudes à l'emploi
de grumes de Hêtre (*Fagus sylvatica* L.)
par un panel de professionnels.**

par Fleur LONGUETAUD

Soutenue publiquement le 27 Juin 2001 devant le jury composé par Messieurs :

Jean-Pierre DAURES
Gilles DUCHARME
Pierre JACOB
Jean-Dominique LEBRETON
Pascal MONESTIEZ
Robert SABATIER



LERFOB
UMR INRA - ENGREF 1092
EQUIPE DE RECHERCHES SUR LA
QUALITE DES BOIS

SOMMAIRE

AVERTISSEMENT	1
REMERCIEMENTS	2
RESUME	3
SUMMARY	4
0- INTRODUCTION	5
1- MATERIEL	7
1.1- Caractéristiques du bois de Hêtre visibles sur des sections de grumes prises en compte dans l'enquête	7
1.2- Caractéristiques de l'échantillon	9
2- MISE AU POINT DU QUESTIONNAIRE ET CHOIX DU PLAN D'EXPERIENCE	12
2.1- Mise au point du questionnaire d'enquête	12
2.2- Lien de notre étude avec l'analyse sensorielle	13
2.3- Choix du plan d'expérience	14
2.3.1- Plans en blocs incomplets équilibrés (BIE)	15
2.3.2- Plans en blocs incomplets partiellement équilibrés (BIPE)	16
2.3.3- Construction du plan d'expérience pour la réalisation de l'enquête	17
2.4- Saisie des données	18
2.5- Les données à analyser	19
3- METHODES STATISTIQUES UTILISEES POUR L'ANALYSE DE NOS DONNEES	21
3.1- Analyse préliminaire des variables de présence et d'intensité des défauts sur les rondelles	21
3.1.1- Généralités sur le type de variables utilisées	21
3.1.2- Analyse de la corrélation entre des variables ordinales	22

3.2- Méthodes pour l'analyse du lien entre les utilisations et la présence plus ou moins importante de défauts sur les rondelles	23
3.2.1- Nature des variables étudiées	23
3.2.2- Modèle linéaire généralisé pour une réponse binaire	24
3.2.2.1- Notations générales	24
3.2.2.2- Contexte du problème.....	24
3.2.2.3- Le modèle logistique (lien LOGIT)	26
3.2.3- La régression PLS (Partial Least Squares).....	31
3.2.3.1- Notations générales	31
3.2.3.2- Contexte du problème.....	31
3.2.3.3- La méthode de Tucker	32
3.2.3.4- La régression PLS.....	33
4- DEROULEMENT DE L'ENQUETE ET RESULTATS	36
4.1- Déroulement de l'enquête	36
4.2- Résultats.....	37
4.2.1- Analyses préliminaires	37
4.2.1.1- Evaluation de l'homogénéité des juges sur la perception de présence et d'intensité des défauts.....	37
4.2.1.2- Corrélations des variables « défauts » entre elles	38
4.2.1.3- ACP sur les variables « défauts »	39
4.2.1.4- Corrélations entre les variables « utilisations »	41
4.2.2- Analyse de l'incidence de la présence de défauts et de leur intensité sur le choix des utilisations par le modèle LOGIT	42
4.2.2.1- Utilisation en tranchage.....	43
4.2.2.2- Utilisation en déroulage	45
4.2.2.3- Utilisation en ébénisterie	46
4.2.2.4- Utilisation en menuiserie intérieure.....	47
4.2.2.5- Utilisation en carcasse de meuble.....	48
4.2.2.6- Conclusion	48
4.2.3- Utilisation de la régression logistique sur composantes PLS pour tester l'influence des défauts sur le choix des utilisations	49
4.2.3.1- Etape 1 : régression PLS	49
4.2.3.2- Etape 2 : régression logistique sur composantes PLS	51
4.2.4- Comparaison des modèles utilisés	51

5- CONCLUSION ET PERSPECTIVES	54
5.1- Critiques de l'enquête.....	54
5.2- Critiques des analyses effectuées sur les données.....	56
5.3- Synthèse des résultats obtenus	57
5.4- Perspectives	58
BIBLIOGRAPHIE	60
ANNEXES.....	62
TABLE DES ANNEXES	63

AVERTISSEMENT

Le travail de recherche du DEA Biostatistiques de Montpellier dont les résultats sont présentés ici s'inscrit dans un double contexte :

- notre volonté - en allant frapper aux bonnes portes - de préparer un profil de compétence pour notre équipe, et plus largement pour le Programme "Qualité des Bois" de l'INRA, dans le domaine qu'il est convenu d'appeler l'identification des besoins des transformateurs et utilisateurs du bois, que ce bois soit encore sur pied ou qu'il soit déjà entré dans le processus d'élaboration de produits. Ce domaine nous semble particulièrement important, sur le plan économique d'une part, mais aussi sur le plan scientifique puisque - ce travail de DEA nous semble le démontrer - il peut être le prétexte d'une recherche originale. Ne nous démentira pas sur ce point la montée en puissance au plan mondial de telles préoccupations et de tels travaux, des travaux auxquels notre collègue Gérard Janin, maintenant en retraite, a contribué avec son enthousiasme, son inventivité et son originalité ;
- le contexte "Tempêtes de 1999" qui a vu notre équipe s'engager, avec les moyens modestes dont elle dispose, dans la mobilisation générale autour de cet événement. Elle l'a fait en décidant de participer au projet dirigé par l'Office National des Forêts intitulé "Conservation en forêt de hêtres chablis sans exploitation", projet qui entraine dans le cadre des actions soutenues par la Direction de l'Espace Rural et de la Forêt au sujet des "Expérimentations d'exploitation forestière et de stockage des bois mises en place suite aux tempêtes de 1999". Afin de traiter cette question en sortant de l'empirisme, il nous est très vite apparu indispensable d'évaluer le plus objectivement possible l'impact sur l'aptitude à l'emploi des grumes de Hêtre à partir de ce dont le scieur, le dérouleur ou le trancheur (mais aussi le vendeur... et nous-mêmes dans le cadre de ce projet) disposent pour en juger, c'est-à-dire l'observation en forêt de la section des grumes .

Gérard NEPVEU
Responsable de l'Equipe « Qualité des Bois du LERFOB
19 novembre 2001

REMERCIEMENTS

Je tiens à remercier :

Gérard Nepveu pour m'avoir accueillie au sein de l'Équipe de Recherches sur la « Qualité des Bois » du Lerfob, ainsi que pour ses conseils et sa très grande disponibilité, même le week-end ;

Gilles Le Moguédec qui m'a apporté ses lumières sur l'aspect scientifique du travail tout au long du stage ;

Pierre Gelhaye, Christian Herbé et André Perrin qui ont activement participé à toutes les étapes techniques de mon stage ;

Ginette Jacquemot, Nathalie Morel pour le secrétariat ainsi que TOUS les autres membres de l'équipe Qualité des bois du Lerfob.

Robert Sabatier pour son aide et son soutien ;

ainsi que les personnes, extérieures au labo, qui m'ont consacré du temps pour la mise en place du questionnaire, Isabelle Planchais (STIR Nord-Est de l'ONF), René Keller (ENGREF-Nancy) et Jérôme Thaille (CTBA-Paris) ;

et enfin bien sûr tous les personnes, professionnels du bois qui ont accepté de répondre à cette enquête, malgré des emplois du temps très chargés, et sans qui ce travail n'aurait pu avoir lieu.

RESUME

Cette étude a été réalisée en vue d'identifier les critères de qualité conditionnant l'acceptation ou le refus du bois de Hêtre dans des utilisations de valorisations diverses. Pour ce faire, nous avons construit une enquête que nous avons menée auprès d'un panel de douze professionnels du bois.

Le travail de stage a consisté dans un premier temps à rédiger un formulaire d'enquête adapté au problème posé, et à mettre en place le plan d'expérience associé. Dans un deuxième temps, nous avons analysé les données ainsi collectées et nous avons cherché à mettre en place des modèles statistiques afin de prédire l'aptitude à l'emploi de grumes de Hêtre à partir de défauts observables sur des photos de sections de grumes. Pour cela nous avons essentiellement effectué la régression logistique des variables « défauts » sur les variables « acceptation/refus ».

L'enquête se situe dans le contexte de la tempête de Décembre 1999. Le marché actuel du bois est donc tout à fait particulier et a nécessairement influé sur les réponses des professionnels.

Mots-clés : Analyse sensorielle - Blocs Incomplets Partiellement Equilibrés - Enquête - Hêtre - Planification d'expérience - PLS - Qualité du bois - Régression logistique

SUMMARY

This study has been carried out with the aim of identifying the criteria of quality influencing the acceptance or rejection of Beech wood in uses of diverse values. For this purpose we have designed a questionnaire that we have addressed to a panel of wood professionals.

At first the work was to prepare a survey form adapted to problem posed, and to create the associated experimental design. Secondly, we have analysed the data thus collected and we have tried to establish some statistical models in order to predict the ability to use the Beech woods logs from some observable imperfections of photos of sections of the logs.

The survey occurs in the context of the storm of December 1999, the present market of wood is very particular, and has necessarily had an effect on the answers of the professionals.

Key-words : Sensory studies - Survey - Partially Balanced Incomplete Block Designs - Beech - Experimental Design - PLS - Quality of wood - Linear logistic model

0- INTRODUCTION

L'étude réalisée porte sur le bois de Hêtre (*Fagus sylvatica* L.). Le Hêtre représente 9.5% de la surface forestière¹ française (AFOCEL, 1999). En 1998, 1924 milliers de m³ de grume² de Hêtre ont été exploités soit une augmentation de 3,2% par rapport à l'année 1997, mais une diminution de 12,7% par rapport à 1974 (Cinotti, 1998). Le prix moyen du Hêtre varie suivant ses utilisations : 1504 F/m³ (prix au m³ sur écorce) en 1997 pour du Hêtre à déroulage³, 496 F/m³ pour du Hêtre à sciages (Cinotti, 1998). Les utilisations sont donc plus ou moins nobles et elles sont conditionnées par la présence de défauts sur le bois.

Des normes de classement en qualité existent déjà pour les bois ronds (CTBA⁴) qui tiennent compte de l'aspect de la section mais aussi de l'aspect de la grume. Les classes vont de A, la meilleure qualité, à D, qualité moindre (on retrouve en D les bois qui ne peuvent entrer dans aucune des classes A, B ou C). Cependant, les industriels du bois ne s'appuient pas toujours sur ces normes et en contestent même parfois la pertinence.

Le but de ce travail est d'étudier la perception des utilisateurs de bois de Hêtre à qui sont présentées des photographies de sections transversales de grumes prises au moment de leur tronçonnage. Les données seront donc essentiellement subjectives. Il s'agit de faire le lien entre les défauts perçus par les professionnels sur les photos et les utilisations auxquelles ils destinent le bois en question.

Pour cette étude, nous avons réalisé une enquête auprès d'industriels du bois de Hêtre, essentiellement lorrains (scieurs, dérouleurs, ...).

Notre travail de stage a consisté à réaliser tout le travail d'enquête ; depuis la conception du questionnaire, afin qu'il permette de répondre aux questions posées, et du plan d'expérience associé, jusqu'à l'analyse des données collectées par des méthodes

¹ Surface des forêts de production et de petits massifs selon l'essence dominante

² Grume : tronc ou section de tronc d'un arbre abattu, ébranché, recouvert ou non de son écorce (dictionnaire forestier ; Métro A., 1975)

³ Voir définition du déroulage au § 1.1

⁴ Centre Technique du Bois et de l'Ameublement - www.ctba.fr

statistiques préalablement identifiées, en passant par les phases d'interviews des industriels sur le terrain et de saisie des données.

Il est important de noter que l'enquête s'est déroulée au printemps 2001, soit un peu plus d'un an après la tempête de décembre 1999, tempête qui a particulièrement affecté la ressource française en bois de Hêtre, notamment ceux proches de l'exploitation en coupe de régénération. Dans ce contexte, le marché est donc encombré par un afflux très important de bois de Hêtre. Les réponses à l'enquête ont nécessairement été influencées par la conjoncture actuelle, mais également par l'enquête elle-même qui a porté largement sur des bois mis à terre par la tempête donc en partie affectés par des altérations consécutives à leur conservation en forêt sans exploitation.

Une autre question à laquelle l'enquête devait permettre de répondre, mais qui ne sera pas traitée ici, était de comparer la qualité du bois entre des arbres abattus par la tempête et conservés en forêt, et des arbres témoins restés debout. Notre travail peut toutefois être considéré comme ayant contribué au traitement de cette question dans la mesure où il s'est notamment enquis (annexes 15 et 16) de mettre en relation le pourcentage d'acceptation pour différentes utilisations du bois de Hêtre avec la note d'altération attribuée par les professionnels à la section des grumes examinées. Au cas d'espèce, les altérations en question étaient consécutives à la conservation en forêt et sans exploitation de hêtres chablis (et témoins) qui constituaient l'échantillonnage pris en compte dans notre recherche.

1- MATERIEL

1.1- Caractéristiques du bois de Hêtre visibles sur des sections de grumes prises en compte dans l'enquête

Les principaux défauts observables sur des sections de grumes de Hêtre et pris en considération dans la norme AFNOR (1997) de classement des bois ronds sont :

- des problèmes de dimension : **diamètre** trop petit ;
- des caractéristiques de structure : **largeur d'accroissement** trop faible. Si les cernes sont trop fins le bois sera réputé nerveux : il risquera de se fendre au tronçonnage ou au sciage. On recherche donc des largeurs d'accroissement assez grandes ;
- des singularités de structure :
 - **nœuds** : partie d'une branche englobée dans le tronc par suite de la croissance en diamètre de ce dernier. On distingue de nombreuses catégories de nœuds en fonction de leurs caractéristiques ;
 - **cœur excentré** : position excentrée du cœur par rapport au centre de la section ;
 - **méplat** : différence de longueur entre le diamètre maximum et le diamètre minimum de la section. La section n'est donc pas circulaire ;
 - **entre-écorce** : morceau d'écorce, parfois associé à un nœud, qui a été plus ou moins intégré par la croissance en diamètre du tronc ;
 - **fentes** : dans le cas de notre étude, il s'agissait essentiellement de fentes de cœur, c'est-à-dire de fentes radiales partant du cœur. Elles étaient provoquées par des problèmes de « nervosité » du bois ;
- des altérations dues à l'action des champignons (AFNOR, 1998) :
 - **cœur rouge** : discoloration rouge ou brune, affectant la partie centrale du bois de Hêtre, et nettement délimitée ;

- **cœur rouge flammé** : cœur rouge du Hêtre altéré qui a une forme étoilée ;
 - **discoloration** : modification de la coloration naturelle du bois non associée à une perte de résistance ;
 - **échauffure** : premier stade de la pourriture, caractérisé par des veines ou des discolorations dans le bois, la texture générale et les propriétés mécaniques restant encore plus ou moins inchangées ;
 - **pourriture** : décomposition du bois sous l'action de champignons ou d'autres micro-organismes créant un ramollissement, une perte progressive de masse et de résistance mécanique et, souvent, un changement de texture et de couleur ;
 - **moisissure** : développement fongique sous forme duveteuse ou poudreuse pouvant se former à la surface du bois en atmosphère humide ;
 - **maladie du T** : grave imperfection du bois de Hêtre dont la trace est en forme de T en coupe transversale. Il s'agit au départ d'une fissure de l'écorce entraînant des lésions. La plaie se cicatrise par recouvrement sur plusieurs années, ce qui est à l'origine de la trace en forme de T. Ce défaut se retrouve souvent à plusieurs endroits sur un même arbre, et nuit fortement à la commercialisation des grumes affectées (Perrin, 1981) ;
- Autres : blessures cicatrisées.

Des illustrations de certains de ces défauts figurent en Annexe 1.

Remarque : le « cœur rouge sain et homogène », le « cœur rouge flammé » et la « maladie du T » sont classés par l'AFNOR (1991, 1998) et le CTBA comme des altérations dues à l'action des champignons ; ce n'est pas nécessairement l'avis de tout le monde.

Les principales utilisations du bois de Hêtre sont par ordre décroissant de valorisation :

- le **tranchage** : découpage en feuilles discontinues d'épaisseur uniforme inférieure au millimètre. Il s'agit surtout d'une utilisation à but décoratif (façades de panneaux) ;
- le **déroulage** : consiste à dérouler la grume à l'aide d'une machine appelée « dérouleuse ». On obtient ainsi une feuille continue de placage (ordre de grandeur de son épaisseur : 2 à 3 mm) qui va servir ultérieurement, par collage à plis croisés des feuilles de placage, pour la fabrication de contre-plaqué, de mobilier scolaire, etc ;
- l'**ébénisterie** : fabrication de meubles massifs ;
- la **menuiserie intérieure** : lambris, plinthes, parquets, etc ;
- la **carcasse de meuble** : parties cachées de meubles ;
- la **palette, le bois de calage, le bois de traverses** de chemins de fer ;
- le **bois de trituration** : bois utilisé pour la fabrication de pâtes à papier et de panneaux de particules ;
- le bois de feu.

Dans le présent travail, nous nous proposons d'étudier le lien entre ces deux aspects, c'est-à-dire d'identifier, à partir de l'analyse des réponses des utilisateurs, l'importance des défauts dans la prise de décision quant à l'acceptation ou le refus des utilisations.

1.2- Caractéristiques de l'échantillon

L'étude, qui a fait l'objet d'un protocole très détaillé de prélèvements et de mesures par Gérard Nepveu de l'Equipe de Recherches sur la Qualité des Bois du Lerfob (Centre INRA de Nancy), portait sur 60 hêtres provenant de deux forêts situées en

Meurthe-et-Moselle (54) et ce, pour un total de quatre parcelles. Quarante de ces arbres avaient été abattus par la tempête de décembre 1999 : ce sont les arbres chablis² (Voir Tableau 1). Une partie (20 arbres) des arbres chablis étaient demeurés en forêt sous l'ombrage d'arbres restés debout dont sont issus les arbres témoins : ce sont les chablis sous couvert. Les autres étaient restés eux aussi en forêt mais sans ombrage : ce sont les chablis en plein découvert.

Les 60 grumes de bois d'œuvre issues de ces arbres ont été découpées en 3 billons³, le billon central d'une longueur d'environ 80 cm étant réservé à l'INRA. La situation de ce billon dans la grume était variable (Annexe 2), cela étant dû à des contraintes sur la longueur des deux autres billons pour des raisons commerciales. Ensuite, les billons destinés à l'INRA ont été emballés individuellement dans un film protecteur transparent afin de conserver leur taux d'humidité et ramenés à l'INRA (22 novembre au 4 décembre 2000), où ils ont été stockés jusqu'à leur débit (19 décembre 2000 au 11 janvier 2001).

Tableau 1 : Origine des 60 arbres ayant servi à l'étude

Situations géographiques	Conditions de conservation	Effectifs d'arbres
Forêt de Haye (54) Sol superficiel	Arbres témoins	10
	Chablis sous couvert	10
	Chablis plein découvert	10
Forêt de Natrou (54) Sol profond	Arbres témoins	10
	Chablis sous couvert	10
	Chablis plein découvert	10

Dans chacun des 60 billons, une rondelle a été découpée. Des photos numériques de ces rondelles, ainsi que des mesures d'humidité (sur des rondelles jumelles des précédentes) ont été réalisées immédiatement, c'est-à-dire dans l'état d'humidité au

² Chablis : arbre naturellement renversé et partiellement ou totalement déraciné, très souvent par le vent

³ Billon : tout tronçon de longueur relativement importante d'une grume résultant du tronçonnage de celle-ci

moment du débit. Plusieurs caractéristiques dimensionnelles ont été mesurées sur ces rondelles et le nombre de cernes représentés sur chacune d'elle a été compté. Nous avons pu assister à toute cette phase pendant deux jours au mois de décembre 2000. Les rondelles ont ensuite été mises à sécher dans un hangar bien ventilé naturellement. Par la suite, sur les photos, le pourcentage de surface altérée a été calculé avec le logiciel Visilog® (logiciel d'analyse d'images). Les résultats obtenus se trouvent en Annexe 3, mais nous n'avons pas eu le temps de les traiter ici.

Ce sont des tirages papiers couleur de ces photos qui ont été présentés aux industriels pour l'enquête. Quelques exemples de photos de rondelles se trouvent en Annexe 4.

2- MISE AU POINT DU QUESTIONNAIRE ET CHOIX DU PLAN D'EXPERIENCE

Pour la réalisation du questionnaire et du plan d'expérience, nous devons tenir compte de la contrainte de temps. En effet, le temps était une contrainte forte à plusieurs niveaux :

- pour nous, qui devons établir le questionnaire, réaliser l'enquête, réfléchir aux méthodes appropriées pour traiter notre problème, analyser les données et enfin rédiger le rapport, le tout en moins de 4 mois (durée du stage) ;
- pour les personnes interrogées, qui ont souvent d'importantes responsabilités au sein de leur entreprise et qui n'avaient donc pas une très grande disponibilité.

De plus, nous devons limiter la durée des interviews afin d'éviter la « saturation » des juges.

Dans ce chapitre, nous présenterons la mise au point du questionnaire puis celle du plan d'expérience.

2.1- Mise au point du questionnaire d'enquête

Le questionnaire (Annexe 5) était constitué d'une première partie portant à la fois sur la personne interrogée et sur son entreprise. Nous étions ainsi renseignée sur des points pouvant avoir leur importance par la suite lors de l'analyse des données : type de formation reçue, expérience de la personne, nature des produits sortis par l'entreprise, taille de l'entreprise, part du Hêtre dans le bois traité, etc.

Ensuite, pour chaque photo, le juge devait remplir un formulaire se composant de deux parties :

- une première partie permettait d'évaluer l'intensité des défauts que la personne considérait présents sur la rondelle ;

- une seconde partie permettait d'évaluer si le bois était d'une qualité suffisante, d'après le juge, pour pouvoir être valorisé dans les différentes utilisations considérées. En cas de refus de certaines utilisations, la personne devait préciser ce qui motivait son refus.

Pour la mise au point du questionnaire, plusieurs personnes, extérieures au laboratoire dans lequel s'est effectué notre stage, nous ont aidée : René Keller (Professeur à l' ENGREF⁴ Nancy), Isabelle Planchais (Ingénieur à la STIR² nord-est de l'ONF) et Jérôme Thaille (CTBA Paris).

2.2- Lien de notre étude avec l'analyse sensorielle

L'analyse sensorielle classique consiste à effectuer simultanément une étude d'appréciation des produits par des consommateurs (note d'appréciation du produit par exemple) et une étude de description des mêmes produits par des juges formés pour la description sensorielle. Des méthodes permettent ensuite d'étudier l'influence des descripteurs sur l'appréciation des produits (Robin-Huon De Kermadec, 1996).

Dans notre enquête, nous demandions à la fois aux « consommateurs » (ici des professionnels du Hêtre, en grande majorité des industriels de la première transformation de ce bois) d'apprécier le produit (acceptation ou refus pour les différentes utilisations), de décrire le produit du point de vue de ses défauts et de motiver leurs appréciations par les défauts qu'ils observaient.

Dans ce type d'enquête, plusieurs facteurs peuvent avoir de l'influence :

- influence du juge : la perception par un juge d'un défaut et de son intensité peut varier par rapport à la perception générale du groupe ;
- influence de la photo elle-même ;

⁴ ENGREF : Ecole Nationale du Génie Rural des Eaux et Forêts

² STIR : Section Technique Inter-Régionale de l'ONF

- influence de la position de présentation de la photo. Il peut par exemple y avoir un effet « fatigue » si le nombre de photos présentées est trop important.

D'autres facteurs non cités ici peuvent également intervenir comme par exemple les conditions dans lesquelles se déroulent l'interview (bruit, etc.).

Le choix du *plan d'expérience* doit permettre de tenir compte dans l'analyse de la présence des différents effets cités ci-dessus, ainsi que des contraintes liées à l'étude, notamment les contraintes de temps. En effet, les impératifs de temps et de disponibilité du panel, c'est-à-dire de l'ensemble des juges, sont les principales contraintes pour la réalisation d'études relevant de l'analyse sensorielle. Pour plus de détails sur les plans d'expérience les plus utilisés en analyse sensorielle, on pourra se référer au travail de thèse de Robin-Huon De Kermadec (1996).

2.3- Choix du plan d'expérience

Le plan d'expérience devait prendre en compte la contrainte du temps de réalisation d'une interview et la contrainte de durée du stage.

Nous avons évalué à 5 minutes le temps nécessaire pour examiner une photo et remplir le formulaire correspondant. Nous disposions de 60 photos. La durée d'une interview d'un juge ne devait, dans tous les cas, pas dépasser 2 heures, dans un souci d'économiser du temps et d'éviter la fatigue du juge. Il est évident qu'un juge ne pouvait pas voir toutes les photos lors d'une même séance (5 mn * 60 photos = 5 h !). Il n'était pas non plus envisageable de réaliser plusieurs séances avec la même personne, les personnes interrogées étant peu disponibles. En raison de ces contraintes, nous nous sommes orientés vers ce que l'on appelle les plans en *blocs incomplets*.

Un plan est dit en blocs incomplets lorsque certains blocs ne contiennent pas tous les traitements (Bergonzini & Duby, 1995). On parle de *blocs* quand les n unités expérimentales dont on dispose sont partitionnées en b sous-ensembles baptisés *blocs*.

Dans le cas de cette étude, nous considérerons que les juges sont les blocs et que les 60 photos constituent les traitements. Le nombre d'unités expérimentales sera le nombre total d'entités présentées à l'ensemble des juges.

Dans le cas de notre enquête, il s'agissait donc d'un plan en blocs incomplets (tous les juges ne voient pas toutes les photos).

Soient les notations suivantes :

- t le nombre de traitements (nombre de photos) ;
- b le nombre de blocs (nombre de juges) ;
- k_b le nombre d'unités expérimentales du bloc b (nombre de photos vues par le juge b) ;
- r_t le nombre d'occurrences du traitement t (nombre de fois qu'est vue la photo t) ;
- n le nombre total d'observations.

Un plan est dit *binnaire* quand un traitement ne peut être présent qu'au plus une fois dans chaque bloc. Un plan est dit *propre* si $\forall b, k_b = k$. Un dispositif est dit *équirépété* si tous les traitements t sont présents un même nombre de fois.

A partir de là, nous devons chercher à équilibrer le plan au maximum. Un plan *équilibré* permet de comparer les traitements deux à deux avec la même précision, et de ne pas trop perdre en précision par rapport à des dispositifs de « référence ». Il existe des listes de dispositifs équilibrés (Cochran & Cox, 1957), pour certaines valeurs des paramètres de construction (t, b, k, r et n).

Deux familles de plans en blocs incomplets peuvent être utilisées : les plans en *blocs incomplets équilibrés* (BIE) et les plans en *blocs incomplets partiellement équilibrés* (BIPE) (Das & Giri, 1979).

2.3.1- Plans en blocs incomplets équilibrés (BIE)

(1) Chaque traitement est répété un même nombre de fois $r_t = r$ et est présent, au plus, une fois par bloc. Il s'agit donc d'un plan équirépété et binaire ;

(2) Tous les blocs ont la même taille $k_b = k$;

(3) Chaque paire de traitements apparaît dans le même nombre λ de blocs : ceci présente l'avantage que toutes les comparaisons deux à deux s'effectuent avec la même précision.

Entre les paramètres, on a les relations suivantes :

$$t \leq b \quad tr = bk \quad \lambda = \frac{r(k-1)}{t-1}$$

Ces relations constituent une condition nécessaire mais non suffisante d'existence d'un tel plan.

Dans le cas de notre étude, nous avons $t = 60$ (le nombre de photos). Pour vérifier la première relation, il aurait fallu interroger plus de 60 personnes, ce qui n'était pas possible sur la durée du stage. Nous avons donc envisagé un plan en blocs incomplètement équilibrés.

2.3.2- Plans en blocs incomplètement équilibrés (BIPE)

Ici, les comparaisons deux à deux ne s'effectuent plus avec la même précision. Considérons un traitement donné : l'idée la plus immédiate est de partitionner l'ensemble des $t-1$ autres traitements en classes, chacune d'entre elles contenant les traitements que l'on souhaite comparer à ce traitement avec une précision donnée.

Les propriétés (1) et (2) des BIE sont également vérifiées par les BIPE. Les propriétés caractéristiques des BIPE sont :

(3) Chaque paire de traitements apparaît dans $\lambda_1, \dots, \lambda_m$ blocs.

Les schémas les plus simples et donc les plus usités sont les schémas à 2 classes d'association : chaque paire de traitements apparaît soit dans λ_1 blocs, soit dans λ_2 blocs. Les traitements des paires qui apparaissent dans λ_1 blocs sont dits « 1-associés », les autres sont dits « 2-associés » (on a donc 2 niveaux de précision pour la comparaison des paires). Il existe des répertoires de ce type de schémas (Cochran & Cox, 1957);

(4) Etant donné deux traitements qui sont « *i*-associés », le nombre de traitements qui sont « *j*-associés » au premier et « *k*-associés » au second est le même quelle que soit la paire de traitements « *i*-associés » considérée (Cochran & Cox, 1957). Cette propriété est la plus délicate à vérifier.

L'utilisateur des BIPE doit essayer de satisfaire ces propriétés, la construction d'un plan vérifiant les caractéristiques des BIPE n'est guère aisée. Dans le cas où une propriété n'est pas vérifiée, cela peut entraîner notamment des différences de précisions dans la comparaison entre les paires et une perte de robustesse.

2.3.3- Construction du plan d'expérience pour la réalisation de l'enquête

La difficulté de construction du plan dans le cadre de notre enquête était de respecter au maximum les contraintes ci-dessus tout en répartissant les photos au mieux entre les juges. Nous avons donc choisi de réaliser un plan en BIPE. Dans un premier temps, nous nous sommes intéressée aux schémas à 2 classes d'association, mais la bibliographie disponible sur le sujet ne nous a pas permis d'appliquer en peu de temps (il fallait démarrer les enquêtes rapidement) ce type de plans à notre problème. Plus particulièrement, nous avons réalisé des essais avec des schémas cycliques, mais sans aboutir.

Les paramètres de construction finalement choisis ont été les suivants :

$$t = 60 ; b = 12 ; k = 15 ; r = 3$$

Nous les avons choisis de façon à satisfaire :

- les contraintes de temps :

Le nombre de photos était fixé à $t = 60$. La période de temps réservée aux enquêtes sur le terrain étant limitée (2 semaines), et les rendez-vous avec les industriels restant encore à prendre, il ne fallait pas être trop ambitieux au niveau du nombre de juges, b , à interroger. De plus, la durée des interviews étant elle aussi limitée, le nombre de photos présentées par juge, k , devait rester raisonnable ;

- les contraintes de base des plans en blocs incomplets :

- (1) toutes les photos ont été vues un même nombre de fois $r = 3$, une photo étant vue au plus une fois par un juge donné;
- (2) tous les juges voyaient le même nombre de photos $k = 15$;
- (3) nous ne nous intéressions pas à certaines paires plus qu'à d'autres. Nous avons donc cherché à homogénéiser le nombre de fois où étaient vues les paires de photos avec nos paramètres. Chaque paire de photos pouvaient être vue au plus 3 fois car $r = 3$. Au total, nous avons $60 \cdot (60-1) / 2 = 1770$ paires possibles. Cependant, les paramètres de construction de notre plan ne nous permettaient au maximum de ne présenter que $[15 \cdot (15-1) / 2] \cdot 12 = 1260$ paires. Donc il y avait forcément des paires qui ne seraient pas vues. Nous avons essayé de rendre le nombre de paires présentées le plus grand possible. Pour cela, nous avons construit un algorithme (Annexe 6). Cet algorithme utilise de nouvelles contraintes facilitant la construction d'un plan, mais limitant également le nombre de plans possibles. Sous ces contraintes les paires de photos sont réparties de la façon la plus homogène possible ;
- (4) les paramètres vérifiaient bien la relation : $t \cdot r = b \cdot k$ ($60 \cdot 3 = 12 \cdot 15 = 180$).

Pour établir le plan, nous avons réalisé un programme en VBA® (Annexe 7). Après application du programme, nous avons obtenu une répartition des photos aux juges vérifiant les contraintes ci-dessus.

Pour l'ordre de présentation des photos, nous avons effectué une randomisation.

2.4- Saisie des données

Christian Herbé (équipe Qualité des Bois du Lerfob) a réalisé un masque de saisie en VBA pour une saisie plus rapide et plus sécurisée des données de l'enquête. Les données ont été stockées dans des tableaux Excel®.

Les données ont été saisies par Pierre Gelhaye (équipe Qualité des Bois du Lerfob) et nous-même.

Les quelques rares données manquantes observées lors de la saisie pour un défaut donné sur une photo donnée ont été remplacées par la moyenne de l'intensité du défaut sur les deux autres observations de la même photo.

2.5- Les données à analyser

Nous avons travaillé sur les variables suivantes :

◆ Variables d'utilisations :

Il s'agit de variables binaires 0/1 (1 si le bois est considéré comme étant de qualité suffisante pour être valorisé dans l'utilisation, 0 sinon).

Nom et code des variables d'utilisations :

Tranchage	Tranch
Déroulage	Deroul
Ebénisterie	Ebenist
Menuiserie intérieure	Menuis
Carcasse de meubles	Carca

Remarque : nous ne nous sommes pas intéressée aux utilisations les moins valorisantes (palette, trituration, bois de feu), car cela ne présentait pas un grand intérêt. En effet et sauf très rares exceptions, le bois est presque toujours de qualité suffisante pour ces utilisations.

◆ Variables de présence et d'intensité des défauts :

Elles sont au nombre de 13. Ces variables peuvent prendre 4 valeurs (1, 2, 3 ou 4) correspondant à 4 niveaux d'intensité (pas du tout, peu, moyennement, beaucoup).

Nom et code des variables de présence et d'intensité des défauts :

	Cœur rouge sain et homogène	CRSH
Cœur rouge flammé	CRF	
Discoloration	Disco	
Echauffure	Echauff	
Pourriture	Pourrit	
Nœuds	Nœuds	
Fentes		Fentes
Moisissure	Moisiss	
cœur excentré	CEX	
Entre-écorce		EntreEc
Maladie du T	T	
Blessures cicatrisées		Bless
Méplat		Meplat

3- METHODES STATISTIQUES UTILISEES POUR L'ANALYSE DE NOS DONNEES

3.1- Analyse préliminaire des variables de présence et d'intensité des défauts sur les rondelles

3.1.1- Généralités sur le type de variables utilisées

Il s'agit des variables de présence et d'intensité des défauts citées en 2.4.

Ce sont des variables *catégorielles*. On dit qu'une variable est *catégorielle* quand l'échelle de mesure est un ensemble de catégories (Agresti, 1990). Quand il n'y a pas d'ordre sur les catégories, la variable est dite *nominale* (exemple : variables de couleur) ; quand il y a un ordre, elle est dite *ordinaire*. Cependant, les distances entre les catégories demeurent inconnues. C'est le cas pour nos variables. L'intensité des défauts établit un ordre, mais rien ne permet d'évaluer la distance entre les différents niveaux d'intensités : pas du tout, peu, moyennement, beaucoup. Enfin, nous avons les variables pour lesquelles la distance entre catégories est connue (en général, des variables que l'on peut mesurer ou quantifier). Il s'agit des variables par intervalle et des variables quantitatives pures.

Ensuite, nous pouvons distinguer les variables *discrètes* et *continues*. En pratique, cela dépend du nombre de valeurs qu'elles peuvent prendre et de leur répartition.

Les variables nominales sont dites *qualitatives* ; les variables dont on connaît la distance entre les catégories sont dites *quantitatives*. Pour les variables ordinales, la distinction est plus floue. Elles sont un intermédiaire entre les variables qualitatives et quantitatives. Afin de simplifier les études, et bien qu'il existe des méthodes spécifiquement adaptées aux variables ordinales, il arrive souvent qu'on les considère

comme appartenant à l'une ou à l'autre de ces deux catégories. Le choix des méthodes est ainsi plus large mais les résultats sont alors à considérer avec précaution.

3.1.2- Analyse de la corrélation entre des variables ordinales

Pour des variables quantitatives continues, on utilise le *coefficient de corrélation de Pearson* qui permet d'établir le degré de la liaison linéaire entre deux variables.

Pour des variables ordinales, en l'absence de distance, parler de linéarité n'a pas beaucoup de sens. Cependant, la notion d'ordre permet tout de même de considérer le degré avec lequel la relation est monotone. Nous pouvons établir une relation d'ordre entre deux individus mais nous ne pouvons pas évaluer la distance qui les sépare.

Soient deux individus et deux variables. On parlera de paire d'individus *concordante* quand l'individu le mieux classé pour une variable l'est aussi pour l'autre ; de paire d'individus *discordantes* quand l'individu le mieux classé pour une variable est moins bien classé pour l'autre. On parlera d'*ex æquo* si les deux individus ont le même rang sur l'une au moins des variables de la paire. Les coefficients suivants sont adaptés aux données ordinales (Saporta, 1990) :

- ◆ le coefficient de corrélation des rangs de Spearman

L'expression de ce coefficient est celle du coefficient de corrélation habituel mais calculé sur les rangs.

- ◆ le coefficient de corrélation des rangs τ de Kendall

$$\text{Soit } S = \sum_{i < j} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j),$$

où x_i est la valeur de la variable x pour l'individu i ; y_i est la valeur de la variable y pour l'individu i .

on a :

$$\tau = \frac{S}{T}, \text{ où } T = \frac{n(n-1)}{2}$$

Dans le cas où il y a des classements *ex æquo*, la formule est légèrement modifiée :

$$\tau = \frac{S}{\sqrt{(T_0 - T_1)(T_0 - T_2)}} ;$$

où $T_0 = n(n-1)/2$, $T_1 = \sum t_i(t_i - 1)/2$ et $T_2 = \sum u_i(u_i - 1)/2$

Les t_i (les u_i) sont les nombres de valeurs *ex æquo* pour la variable x (respectivement pour la variable y) dans le $i^{\text{ème}}$ groupe d'*ex æquo* pour x (respectivement pour y), n est le nombre d'observations et $\text{sgn}(z) = 1$ si z est supérieur à 0, 0 si égal à 0 et -1 si plus petit que 0.

La valeur de τ est comprise entre -1 et 1.

3.2- Méthodes pour l'analyse du lien entre les utilisations et la présence plus ou moins importante de défauts sur les rondelles

3.2.1- Nature des variables étudiées

Les variables d'utilisation sont celles citées en 2.5. Il s'agit de variables binaires (acceptation ou refus pour telle utilisation). Quant aux variables de défauts, ce sont les variables ordinales précédemment citées (§ 2.5).

Nous désirons étudier l'impact des défauts sur l'acceptation ou le refus dans les différentes utilisations. Les *variables explicatives* sont les défauts et leur niveau d'intensité. Les *variables à expliquer* sont les variables binaires d'acceptation ou de refus des différentes utilisations. Nous aurons la possibilité d'étudier chaque utilisation indépendamment des autres (§ 3.2.2) ou bien conjointement (§ 3.2.3).

3.2.2- Modèle linéaire généralisé pour une réponse binaire

3.2.2.1- Notations générales

Soient :

- X la matrice ($n \times (k+1)$) des k variables explicatives (descripteurs + la constante). La première colonne n'est composée que de 1 ; $x_{i0} = 1, i = 1, \dots, n$. Les autres colonnes sont les variables de présence et d'intensité des défauts, l'intensité étant codée de 1 (pas du tout) à 4 (beaucoup) ; n est le nombre d'observations ;
- Y le vecteur des réponses ($n \times 1$) contenant uniquement des 0 et des 1, acceptation ou refus (« succès » / « échec ») pour une utilisation donnée.

3.2.2.2- Contexte du problème

Il s'agit de modéliser l'influence d'un ensemble de variables explicatives sur la réalisation d'une variable de Bernoulli (utilisations prises séparément les unes des autres) . Les variables explicatives seront traitées ici comme quantitatives ce qui permet de garder la notion d'ordre entre les intensités. Compte tenu du codage, cela revient à supposer que la distance entre deux modalités successives est constante.

Il y a plusieurs façons de considérer la variable de réponse :

- soit nous regroupons les observations par descripteurs identiques, puis nous nous intéressons à la proportion de « succès » sur l'ensemble des variables présentant des descripteurs identiques. Les variables ainsi considérées suivent une loi binomiale ;
- soit nous considérons les observations indépendamment les unes des autres. Les variables sont binaires et suivent une loi de Bernoulli, qui n'est autre qu'un cas particulier de la distribution binomiale.

Dans notre étude, nous avons un nombre assez important de descripteurs ayant chacun quatre modalités et il n'est pas intéressant de regrouper les observations par ensemble de descripteurs communs puisque chaque variable a en général un ensemble de

valeurs des descripteurs qui lui est propre. Nous nous sommes donc placée dans le cas où la variable de réponse Y est binaire et suit une loi de Bernoulli de paramètre $p = E(Y)$.

Nous pouvons alors faire plusieurs remarques :

- la variance de Y n'est pas constante puisqu'elle s'écrit comme une fonction de son espérance : $Var(Y) = p(1-p)$ alors que, dans le modèle linéaire classique, Y est distribuée suivant une loi normale et sa variance est donc constante ;
 - nous cherchons à prédire les valeurs 0 ou 1, ce que ne permet pas le modèle linéaire classique (Daudin, 1998). Le modèle linéaire classique nous fournirait des valeurs prédites assimilables à une probabilité de réalisation de l'événement ($Y = 1$), mais les valeurs pourraient alors être inférieures à 0 ou supérieures à 1 : que décider alors ?
- C'est en partie pour ces raisons que le modèle linéaire classique n'est pas adapté à nos données et doit être modifié.

Le *modèle linéaire généralisé* (McCullagh & Nelder, 1989) permet d'étendre le modèle linéaire à une classe plus large de distributions de la variable Y , appelée *famille exponentielle*. Si Y appartient à la famille exponentielle, alors il existe des fonctions a, α, b et δ telles que sa densité s'écrive sous la forme :

$$f(x, \theta) = \exp[a(x)\alpha(\theta) + b(x) + \delta(\theta)] ;$$

où θ est appelé paramètre naturel de la famille exponentielle, $E(Y) = \delta'(\theta)$ et $V(Y) = \delta''(\theta)$.

Trois éléments caractérisent un modèle linéaire généralisé : une *partie aléatoire* correspondant à la densité de probabilité de la variable de réponse, une *partie fixe* constituée d'une combinaison linéaire des variables explicatives et une *fonction de lien* g décrivant la relation entre la partie aléatoire et la partie fixe. La fonction de lien est une fonction monotone de classe C^2 .

Dans le modèle linéaire général (modèle classique), on cherche à exprimer l'espérance par une combinaison linéaire des colonnes de X . Dans le modèle linéaire généralisé, c'est une fonction de l'espérance que l'on cherche à écrire comme une combinaison linéaire des colonnes de X . Le modèle peut donc s'écrire sous la forme :

$$g(p) = X\beta,$$

où $p = E(Y)$ et β sont les paramètres du modèle à estimer.

La méthode la plus largement utilisée pour l'estimation des paramètres est la méthode du maximum de vraisemblance alors qu'en régression classique c'est généralement la méthode des moindres carrés qui est utilisée.

Notre but est de prédire la probabilité de réalisation de l'événement $\{Y = 1\}$; $p = E(Y) = P(Y = 1)$. Or, par définition, une probabilité prend ses valeurs entre 0 et 1. La réponse $X\beta$ peut prendre ses valeurs sur l'ensemble des réels. Pour répondre à cette contrainte, nous allons utiliser une fonction de lien bijective de $[0, 1]$ sur $(-\infty, \infty)$. Il existe une infinité de fonctions de lien a priori convenables mais les plus utilisées et les mieux documentées dans le cas d'une variable binaire sont : la fonction *logit* ou *logistique*, la fonction *probit* et la fonction *log-log*. Nous ne détaillerons ici que la fonction de lien LOGIT qui est celle que nous avons utilisée dans notre étude.

3.2.2.3- Le modèle logistique (lien LOGIT)

Le modèle logistique correspond à un modèle linéaire généralisé de fonction de lien LOGIT. Il s'applique à des variables de distribution binomiale donc également au cas particulier de variables de distribution de Bernoulli. Dans toute la suite, nous nous sommes placée dans le cas d'une loi de Bernoulli, la généralisation à une loi binomiale étant immédiate.

Y_i suit une loi de Bernoulli de paramètre p_i et, par conséquent, sa densité peut s'écrire sous la forme caractéristique d'une densité appartenant à la famille exponentielle :

$$P(Y_i = y_i) = \exp[y_i \log\left(\frac{p_i}{1-p_i}\right) + \log(1-p_i)],$$

où y_i prend les valeurs 0 ou 1, et avec $a(y_i) = y_i$ et $\alpha(p_i) = \log\left(\frac{p_i}{1-p_i}\right)$.

On parle de *lien naturel* quand $g(p) = \alpha(p)$, c'est-à-dire quand le lien apparaît dans l'expression de la vraisemblance. L'utilisation d'un lien naturel facilite les calculs lors de l'estimation du maximum de vraisemblance.

Le lien LOGIT, $g(p) = \text{logit}(p) = \log\left(\frac{p}{1-p}\right)$, est le lien naturel pour la loi de Bernoulli.

Le modèle de régression prend ainsi la forme suivante :

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = X\beta = \eta,$$

où β est le vecteur des paramètres du modèle à estimer.

La vraisemblance du modèle s'écrit :

$$L(\beta) = \prod_{i=1}^n p_i^{y_i} (1-p_i)^{1-y_i}$$

- Estimation des paramètres du modèle (Collett, 1991) :

L'estimation des paramètres β du modèle s'obtient par la méthode du maximum de vraisemblance.

La log-vraisemblance du modèle s'écrit :

$$\begin{aligned} \log L(\beta) &= \sum_{i=1}^n \left\{ y_i \log\left(\frac{p_i}{1-p_i}\right) + \log(1-p_i) \right\} \\ &= \sum_{i=1}^n \left\{ y_i \eta_i - \log(1 + e^{\eta_i}) \right\} \end{aligned}$$

On calcule les dérivées partielles de la log-vraisemblance par rapport à chacun des paramètres β_j :

$$\frac{\partial \log L(\beta)}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} e^{\eta_i} (1 + e^{\eta_i})^{-1}, \quad j = 0, \dots, k$$

En évaluant ces dérivées en $\hat{\beta}$ et en les égalant à 0, on est ramené à la résolution d'un système à $k+1$ équations non linéaires et $k+1$ inconnues.

Ce système ne peut être résolu que numériquement. Dans le logiciel SAS, deux algorithmes itératifs de calcul du maximum de vraisemblance sont disponibles :

- la méthode par défaut, appelée « Fisher-scoring » : elle est équivalente à la méthode des moindres carrés pondérés itérés ;
- l'algorithme de Newton-Raphson.

Les paramètres estimés, nous disposons donc d'un modèle permettant de prédire la valeur du « logit ». Il faut ensuite refaire une transformation pour obtenir les valeurs prédites des probabilités.

$$\hat{p} = \frac{1}{1 + e^{-x\hat{\beta}}}$$

- Qualité d'ajustement d'un modèle linéaire logistique dans le cas de variables binaires :

(1) La déviance

La déviance est un critère souvent utilisé pour tester la qualité d'ajustement d'un modèle et pour comparer des modèles entre eux (différence des déviances).

La déviance s'écrit $D = -2 \log\left(\frac{\hat{L}}{\hat{L}_0}\right)$, où \hat{L} est la vraisemblance du modèle considéré et \hat{L}_0 la vraisemblance du modèle saturé.

Cependant, elle n'est pas utilisable dans le cas d'une variable binaire ; en effet, elle s'écrit alors :

$$D = -2 \sum_{i=1}^n \{ \hat{p}_i \log \text{it}(\hat{p}_i) + \log(1 - \hat{p}_i) \}$$

Comme nous pouvons le voir dans l'expression ci-dessus, la déviance ne dépend des observations y_i qu'au travers des estimations des probabilités \hat{p}_i . Elle ne permet donc pas de mesurer la correspondance entre les observations et les prédictions. La déviance dans le cas de variables binaires ne peut pas être utilisée pour tester la qualité d'ajustement du modèle (Collett, 1991).

(2) Le χ^2 de Pearson

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \hat{p}_i)^2}{\hat{p}_i(1 - \hat{p}_i)}$$

Cette statistique peut être utilisée pour juger de la qualité du modèle. Sous l'hypothèse H_0 que les observations suivent une loi de Bernoulli, elle suit un χ^2 à $n - p$ degrés de libertés (n est le nombre d'observations et p le nombre de paramètres du modèle).

(3) Le reclassement

Si, par exemple, nous attribuons à une observation la valeur 1, si sa probabilité prédite est supérieure à un certain niveau (généralement 0,5), il suffira de regarder le tableau de répartition des observations entre les valeurs prédites pour avoir une idée de la qualité du modèle. Une inadéquation absolue est obtenue lorsque l'on est capable de ne reclasser que 50% des individus.

La qualité de la prédiction dépend bien sûr étroitement du seuil choisi.

Une autre méthode, un peu moins précise mais donnant une bonne idée de la qualité de prédiction du modèle, consiste à regarder le pourcentage de paires concordantes, c'est-à-dire que l'on prend des paires d'observations, et on regarde si l'ordre des valeurs observées est le même que celui des valeurs prédites. Exemple d'une paire concordante : si une observation A est acceptée en tranchage et qu'une observation B est refusée en tranchage, la probabilité prédite d'acceptation de A en tranchage doit être supérieure à celle prédite pour B. Pour une paire discordante, l'ordre entre les observations et les prédictions est inversé. Quand une paire n'est ni concordante, ni discordante, on parle d'ex æquo. Cela nous évite d'avoir à déterminer un seuil comme précédemment. C'est la méthode que nous avons choisie de retenir pour comparer les modèles entre eux du point de vue de leur qualité de prédiction.

La qualité du modèle peut être affectée pour plusieurs raisons :

- les variables explicatives : il se peut qu'il manque des variables explicatives qui auraient dues être incluses dans le modèle ou que certaines variables aient dû subir une transformation avant d'y être incluses ;

- la fonction de lien : celle utilisée n'est pas la plus appropriée pour traiter les données ;
- les observations : certaines observations peuvent être mal prédites par le modèle. Il se peut aussi que certaines observations aient un poids important dans l'ajustement du modèle alors qu'elles ne le devraient pas.

Pour compléter l'analyse nous avons choisi de regarder les résidus du modèle. Cependant, le fait d'avoir à traiter des données binaires ($y_i = 0$ ou 1) présente des particularités dont il faut tenir compte.

1) Résidus de déviance

Dans le cas de variables binaires, la formule de la déviance pour la $i^{\text{ème}}$ observation s'écrit :

$$d_i = \text{sgn}(y_i - \hat{p}_i) \sqrt{\left\{ -2 \left[y_i \log \hat{p}_i + (1 - y_i) \log(1 - \hat{p}_i) \right] \right\}}$$

2) Résidus de Pearson

Le résidu de Pearson pour la $i^{\text{ème}}$ observation s'écrit :

$$\chi_i = \frac{y_i - \hat{p}_i}{\sqrt{\hat{p}_i(1 - \hat{p}_i)}}$$

Le fait d'avoir des variables binaires conduit à ce que la distribution des résidus ne suivra pas une loi normale. Le graphe des résidus en fonction des variables explicatives sera difficilement interprétable car il y aura généralement des tendances dans ces graphes, que le modèle soit correct ou non (Collett, 1991).

Cependant, la représentation des résidus en fonction des observations peut nous donner une information intéressante sur la qualité du modèle. Nous pouvons ainsi repérer des observations éventuellement aberrantes ou des observations qui jouent un rôle important dans la détermination des coefficients du modèle.

3.2.3- La régression PLS (Partial Least Square)

Historiquement la régression PLS est née de l'algorithme NIPALS développé par H. Wold (1966) pour l'analyse en composantes principales et de l'approche PLS proposée par H. Wold (1975) pour l'estimation de modèles d'équations structurelles sur variables latentes.

L'objectif est d'obtenir une régression entre un groupe de variables explicatives (corrélées entre elles) et un groupe de variables à expliquer (également corrélées entre elles) mais :

- en prenant des précautions au sujet du problème de stabilité des estimateurs ;
- en respectant autant que possible les liens entre les variables à expliquer.

3.2.3.1- Notations générales

Soient :

- X la matrice ($n \times (k+1)$) des variables explicatives : ce sont, comme précédemment, les variables de présence et d'intensité des défauts, l'intensité étant codée de 1 (pas du tout) à 4 (beaucoup) ;
- Y la matrice des réponses ($n \times q$), contenant uniquement des 0 et des 1, acceptation ou refus pour les différentes utilisations. Nous avons établi un classement a priori des utilisations suivant leur niveau de valorisation. Il serait intéressant d'utiliser une méthode qui puisse tenir compte de la hiérarchisation entre ces variables.

3.2.3.2- Contexte du problème

La *régression multiple* est une méthode adéquate pour prédire une variable Y suivant les valeurs de variables explicatives X . Cependant, les variables explicatives doivent être peu nombreuses et peu corrélées entre elles. Nous verrons ultérieurement (Chapitre 4.) que l'on peut identifier des groupes de variables corrélées parmi les

variables explicatives. Ceci posait, en régression classique, un problème de stabilité des estimateurs.

Des variables explicatives fortement corrélées entre elles conduisent à une instabilité des estimateurs. Soit $\hat{\beta}$ le vecteur des estimateurs du modèle de régression ($Y=X\beta+\varepsilon$), on a $Var(\hat{\beta})=(X'X)^{-1}\sigma^2$. Si les variables explicatives sont corrélées, $X'X$ sera mal conditionnée et son inverse aura des termes élevés. Les paramètres du modèle seront donc estimés avec une forte imprécision. D'autres méthodes semblent plus adaptées à notre problème, comme par exemple les méthodes PLS. Les méthodes PLS sont basées sur l'analyse inter-batterie de Tucker.

3.2.3.3- La méthode de Tucker

Soient deux triplets (X, Q_x, D) et (Y, Q_y, D) . Les variables X et Y sont deux ensembles de variables jouant un rôle symétrique, Q_x et Q_y respectivement, les métriques associées et D la métrique du poids des n individus. X et Y sont centrées pour la métrique D . En général, nous avons $Q_x=Id$, $Q_y=Id$ et $D=\frac{1}{n}Id$.

Le but de l'analyse est de rechercher l'information commune entre les deux ensembles de variables.

Pour cela, nous cherchons à maximiser le critère suivant :

$$Cov(XQ_x a, YQ_y b) = Cov(t, u) ;$$

(t et u sont appelées composantes ; ce sont des combinaisons linéaires des variables initiales)

sous les contraintes :

$$\|a\|_{Q_x} = 1 \text{ et } \|b\|_{Q_y} = 1.$$

Ce critère peut encore s'écrire :

$$\sqrt{Var(t)}\sqrt{Var(u)}Corr(t, u)$$

Nous retrouvons donc rassemblés ici les critères de l'ACP (maximisation de la variance), ainsi que de l'analyse canonique (maximisation de la corrélation). Les composantes

obtenues sont donc aussi corrélées que possible tout en étant aussi explicatives que possible de X et de Y .

Il s'agit d'un problème de maximisation sous contraintes. Nous pouvons le résoudre par la méthode des multiplicateurs de Lagrange (nous ne détaillerons pas les calculs ici).

Nous sommes ainsi ramenée à un problème aux valeurs propres :

- a est le vecteur propre de $X'DYQ_yYDXQ_x$ associé à la plus grande valeur propre ;
- b est le vecteur propre de $YDXQ_xX'DYQ_y$ associé à la plus grande valeur propre.

Les formules de transition entre a et b sont :

$$\begin{aligned} - \quad a &= \frac{1}{\lambda} X' D Y Q_y b ; \\ - \quad b &= \frac{1}{\lambda} Y' D X Q_x a . \end{aligned}$$

Un inconvénient de la méthode de Tucker réside dans le fait que les composantes obtenues ne sont pas orthogonales entre elles.

3.2.3.4- La régression PLS

La *régression PLS* est une méthode de régression sur variables latentes (Desbois, 1999). Elle permet de construire des modèles prédictifs lorsque les variables explicatives sont nombreuses et très corrélées, mais que l'on désire cependant toutes les garder dans le modèle. De plus, la régression PLS permet de traiter simultanément un ensemble de variables à expliquer Y et de tenir compte des corrélations entre elles ce que ne permet pas la régression multiple. Cette méthode est fréquemment utilisée en chimométrie ; son utilisation est étonnamment plus rare en analyse sensorielle (Robin-Huon de Kermadec, 1996).

Dans le cas de la régression PLS, les ensembles de variables X et Y ne jouent plus un rôle symétrique ; en effet, nous cherchons à expliquer le second par le premier. Nous nous placerons ici dans le cas classique où : $Q_x=Id$, $Q_y=Id$ et $D=\frac{1}{n}Id$.

Il s'agit de trouver H nouvelles variables (appelées aussi composantes, X-scores) orthogonales entre elles, combinaisons linéaires des variables de départ $t_h = Xa_h$, $h = 1, \dots, H$ (Wold, 1996). Ces composantes sont construites de manière itérative : on extrait un facteur à la fois.

L'algorithme de construction des composantes PLS basé sur la formule de Tucker est le suivant (Tenenhaus, 1995) :

Soient X_0 et Y_0 les matrices des variables de départ centrées-réduites.

- $t_1 = X_0 a_1$
- $t_2 = X_{(1)} a_2$, où $X_{(1)}$ est le résidu de la régression de X_0 sur t_1
- $t_H = X_{(H-1)} a_H$, où $X_{(H-1)}$ est le résidu de la régression de X_0 sur t_1, \dots, t_{H-1}

Les composantes sont construites avec $\|a_1\| = \dots = \|a_H\| = 1$ et de manière à maximiser le critère :

$$\sum_{h=1}^H \sum_{k=1}^P \text{cov}^2(Y_k, X_{(h-1)} a_h)$$

La contrainte d'orthogonalité sur les a_h est contenue dans ce critère (Tenenhaus, 1995). On obtient ainsi que le vecteur a_h est le vecteur propre normé de la matrice $(1/n^2) X_{(h-1)}' Y Y' X_{(h-1)}$ associé à la plus grande valeur propre.

Une fois que l'on a calculé les composantes t_1, \dots, t_H , on peut construire l'équation de régression de Y sur ces composantes. Et comme, de plus, chaque composante est elle-même combinaison linéaire des variables explicatives de départ, on obtient une équation reliant directement Y à X . C'est l'équation de la régression PLS.

Dans le cas général de la régression PLS de (Y, Q_y, D) sur (X, Q_x, D) , le critère à maximiser s'écrit :

$$\text{Cov}(XQ_x a, YQ_y b) = \text{Cov}(c_x, c_y)$$

¹ Il revient au même (et c'est souvent le cas en pratique) de prendre la matrice $(1/n^2) X_{(h-1)}' Y_{(h-1)} Y_{(h-1)}' X_{(h-1)}$, où $Y_{(h-1)}$ est le résidu de la régression de Y_0 sur t_1, \dots, t_{h-1}

sous les contraintes :

$$\|a\|_{Q_x} = 1 \text{ et } \|b\|_{Q_y} = 1,$$

et d'orthogonalité des composantes.

Avec les mêmes notations que précédemment (§ 3.2.3.3 et § 3.2.3.4), on obtient que a_h est le vecteur propre associé à la plus grande valeur propre de $X'_{(h-1)} D Y_{(h-1)} Q_y Y'_{(h-1)} D X_{(h-1)} Q_x$. Il existe bien sûr une formule symétrique pour calculer les b_h .

Nous avons également la formule de transition :

$$b_h = \frac{1}{\lambda} Y'_{(h-1)} D X_{(h-1)} Q_x a_h$$

Dans le cas de notre étude, les variables à expliquer sont binaires. Dans ce cas, nous pouvons utiliser la régression logistique PLS (Tenenhaus). Dans cet article, Tenenhaus présente la régression logistique PLS mais également la régression logistique sur composantes PLS qui semble donner des résultats semblables. Nous avons utilisé la régression logistique sur composantes PLS. Il s'agit d'effectuer la régression PLS des indicatrices de modalités de Y sur X puis d'effectuer une régression logistique de Y sur les composantes issues de la PLS.

4- DEROULEMENT DE L'ENQUETE ET RESULTATS

4.1- Déroulement de l'enquête

La réalisation des enquêtes s'est échelonnée sur les deux premières semaines d'avril. Dans un premier temps (mois de mars), il a fallu prendre des rendez-vous avec des industriels du bois de Hêtre (scieurs, dérouleurs) familiers du jugement des bois ronds. Cela n'a pas toujours été facile du fait de l'emploi du temps chargé des personnes ciblées.

Les photos des rondelles avaient été préalablement rangées dans des cahiers sous pochettes transparentes. Elles ont été réparties en groupes de 15 (Voir plan d'expérience § 2.3.3) dans l'ordre de présentation déterminé par la randomisation.

Les enquêtes se sont déroulées principalement en région Lorraine, mais rarement à proximité de Nancy, ce qui a impliqué des déplacements. Pendant ces déplacements, nous avons toujours été accompagnée d'un technicien de l'INRA (Pierre Gelhaye ou André Perrin).

Après les enquêtes et au vu de leur déroulement, nous avons déjà pu faire quelques remarques avant de commencer les analyses statistiques :

- les juges ne percevaient pas tous les photos de la même façon (Voir tableau en Annexe 9). Il a pu arriver, par exemple, que certains juges ne soient pas familiarisés avec certains défauts cités dans le questionnaire ;
- sur les photos, certains défauts étaient difficiles à distinguer les uns des autres. Par exemple, le type d'altération : discoloration, échauffure, pourriture ou moisissure. On pouvait donc aussi s'attendre à une confusion entre eux. Les analyses ultérieures nous permettront de confirmer ou d'infirmer ce point.

A noter aussi que la situation particulière du marché du Hêtre liée à la tempête de fin 1999 a été évoquée par tous les juges.

4.2- Résultats

4.2.1- Analyses préliminaires

4.2.1.1- Evaluation de l'homogénéité des juges sur la perception de présence et d'intensité des défauts

Nous avons cherché à déceler des différences possibles de notation entre les juges, c'est-à-dire à savoir si un juge note d'une façon générale plus sévèrement ou moins sévèrement que l'ensemble du groupe.

Pour cela, nous avons réalisé une analyse de variance à 2 facteurs (facteur « juge » et facteur « photo ») sur chacune des variables de défauts prises séparément. Les variables ont été considérées comme quantitatives. Nous n'avons malheureusement pas pu tester l'interaction entre les deux facteurs, ce qui nous aurait bien intéressée. Les différents niveaux d'altération : pas du tout, peu, moyennement, beaucoup, ont été codés en 1, 2, 3, 4. Nous sommes dans le cas d'un modèle *incomplet* (tous les juges ne voient pas toutes les photos), *sans répétition* (une photo est vue au plus une fois par un juge), *non équilibré* (un juge voit une photo donnée une fois ou pas du tout). Ce modèle s'écrit :

$$I_{ij} = \mu + P_i + J_j + \varepsilon_{ij}$$

où I_{ij} correspond à l'intensité du défaut considéré pour la photo i et le juge j , μ correspond à l'effet moyen, P_i correspond à l'effet principal de la photo i et J_j à l'effet principal du juge j .

Les tests de Bonferroni et de Scheffé ont été utilisés pour les comparaisons. Le fait de considérer les variables comme quantitatives peut constituer une simplification abusive. Il est important de garder cela à l'esprit et de considérer les résultats avec précaution. D'autres méthodes adaptées aux données ordinales existent mais, par manque de temps, elles n'ont pas été utilisées ici.

Pour ce qui est des variables **discoloration**, **échauffure**, **pourriture** et **moisissure** il apparaît un effet « juge ». Le juge n°6 note ces quatre défauts un peu plus sévèrement que les autres juges.

Pour les variables **méplat** et **cœur excentré**, le juge n°3 se distingue par sa sévérité.

Le juge n°9 ressort très fortement pour son appréciation de la « **maladie du T** ».

De même, le juge n°11 se distingue assez fortement pour son appréciation de la présence d'**entre-écorce** (Les résultats obtenus pour la « maladie du T » et l' « entre-écorce » se trouvent en Annexe 10, les autres sorties ne sont pas présentées).

Enfin les juges numéros 7 et 12 sont un peu sévères par rapport aux autres sur la présence de **fentes**.

Ces résultats nous ont permis de confirmer certaines observations (Annexe 9) faites pendant les enquêtes : le juge n°9 ne connaissait visiblement pas la maladie du T (pour lui il s'agissait d'une maladie liée au « temps » (temps qu'il fait dehors)); le juge n°11 avait une interprétation un peu particulière de la présence d'entre-écorce (ne correspondant pas à la définition).

Par la suite, nous remplacerons les observations des juges n°9 et 11 concernées par des données manquantes qui seront traitées en tant que telles par les logiciels utilisés.

4.2.1.2- Corrélations des variables « défauts » entre elles

Pour analyser la corrélation de variables ordinales, nous disposons du coefficient de corrélation de Spearman et du Tau de Kendall (§ 3.1.2 ; Voir les matrices de corrélations en Annexe 11).

Nous avons pu ainsi établir des groupes de variables assez corrélées entre elles : {**discoloration**, **échauffure**, **pourriture**, **moisissure**}, {**cœur excentré**, **méplat**} et dans une moindre mesure {**cœur rouge sain et homogène**, **cœur rouge flammé**}. A nouveau, ces résultats nous ont permis de confirmer certaines idées a priori sur les variables. En effet :

- nous avons pu remarquer, au travers des définitions ou pendant les enquêtes, certains liens entre les variables {discoloration, échauffure, pourriture, moisissure}. Nous pouvons en effet avoir une présence de discoloration sans échauffure, mais une rondelle échauffée est caractérisée par des discolorations dans le bois ; la pourriture survient à la suite de l'échauffure ; et enfin la moisissure est difficilement identifiable sur photo et peut être confondue avec l'échauffure ou la pourriture notamment. La discoloration et l'échauffure étaient en général bien identifiées par les juges. En revanche, pour la pourriture et la moisissure, les réponses étaient beaucoup plus hésitantes car, sur photo, ces défauts sont assez difficiles à évaluer ;
- ce qui est un peu plus étonnant est la corrélation (faible toutefois) observée entre le cœur rouge sain et homogène et le cœur rouge flammé. Ces deux défauts semblent pourtant, d'après leur définition, faciles à distinguer ;
- enfin, le cœur excentré et le méplat sont deux caractéristiques de forme de la rondelle (des illustrations de ces deux défauts se trouvent en Annexe 1). Il ne paraît pas très étonnant qu'il y ait une corrélation entre elles.

4.2.1.3- ACP sur les variables « défauts »

Dans cette partie, les variables ont à nouveau été considérées comme quantitatives ce qui nous a permis d'utiliser des méthodes descriptives telles que l'ACP. Nous avons utilisé le logiciel SPAD® qui est particulièrement adapté pour ce type d'analyses. (Remarque : SPAD traite les données manquantes en les remplaçant par la moyenne sur les données non manquantes ; les sorties, notamment les valeurs propres, se trouvent en Annexe 12).

Nous pouvons nous demander s'il vaut mieux effectuer une ACP normée ou une ACP non normée. Les variables étudiées variaient sur une même échelle d'intensité (de 1 à 4) ; à noter que ceci est en faveur d'une ACP non normée. Cependant, il n'est pas évident que tous les défauts aient été jugés avec la même sévérité. Par ailleurs, les juges n'utilisaient peut-être pas tous l'échelle en entier. Il est important de ne pas

privilégier les défauts pour lesquels l'intensité est très variable d'un individu (présentation d'une photo) à l'autre. Pour résoudre ce problème éventuel de variance, nous penchions pour l'utilisation d'une ACP normée (toutes les variances sont ramenées à 1).

De plus, dans le cas de l'ACP normée, le cercle des corrélations est la projection de l'ensemble des variables centrées-réduites sur le sous-espace engendré par les composantes principales. Le cercle des corrélations est donc facile à interpréter. Cela ne nous empêchera pas de regarder les sorties obtenues pour une ACP non normée.

Nous avons représenté le cercle des corrélations pour les variables « défauts ». Les variables « utilisations » sont représentées sur le cercle en tant que variables supplémentaires (les variables utilisations ont été recodées en 1, 2 car pour SPAD la valeur 0 correspond à une donnée manquante).

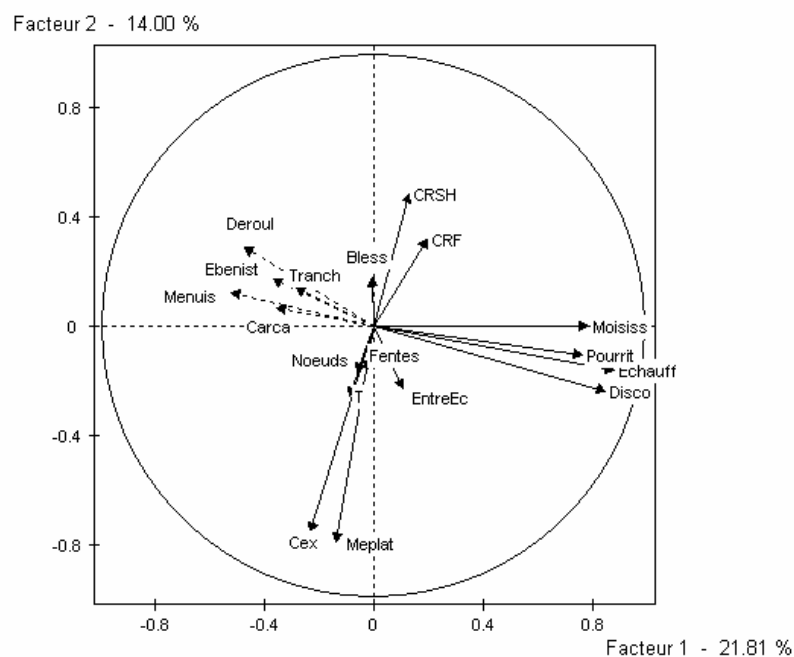


Figure 1 : Premier plan principal de l'ACP normée sur les variables « défauts »

L'examen de la Figure 1 nous permet d'interpréter les composantes principales et de repérer rapidement les caractères corrélés à condition toutefois qu'ils soient proches de la circonférence.

Dans le premier plan principal (expliquant 35.81% de l'inertie), nous retrouvons que les variables **discoloration**, **échauffure**, **pourriture** et **moisissure** sont assez fortement corrélées. Il en est de même pour les variables **cœur excentré** et **méplat**. Dans le deuxième plan principal (Annexe 12), on retrouve la corrélation un peu moins marquée des variables associées à la présence de **cœur rouge**.

Nous pouvons à présent donner une interprétation des axes principaux :

- le premier axe principal est associé au niveau d'altération des rondelles (discoloration, échauffure, pourriture, moisissure) ;
- le deuxième axe est plutôt associé à des caractéristiques de forme de la rondelle (cœur excentré et méplat) ;
- le troisième axe semble correspondre à la présence de cœur rouge (Annexe 12).

Les variables « utilisations » sont négativement corrélées à la présence des défauts **discoloration**, **échauffure**, **pourriture**, et **moisissure**. Dans une moindre mesure, elles semblent inversement corrélées à la présence de cœur rouge et à la présence de problèmes de forme.

Pour plus de détails se reporter aux sorties SPAD en Annexe 12.

Par l'ACP, nous avons retrouvé graphiquement les corrélations précédemment observées (§ 4.2.1.2).

Ces différentes analyses nous ont permis de représenter les corrélations existant entre les différentes variables « défauts » et d'établir des groupes de variables liées : {**discoloration**, **échauffure**, **pourriture**, **moisissure**}, {**cœur excentré**, **méplat**}, et {**cœur rouge sain et homogène**, **cœur rouge flammé**}.

4.2.1.4- Corrélations entre les variables « utilisations »

Précédemment (avant enquête), nous avons établi un classement a priori des utilisations par ordre décroissant de valorisation, à savoir : **tranchage**, **déroulage**, **ébénisterie**, **menuiserie**, **carcasse**. Après enquête, nous pouvons regarder les réponses

des juges quant à la valorisation dans les différentes utilisations. Le tableau 2 présente les fréquences d'acceptation et de refus dans les différentes utilisations.

Tableau 2 : Fréquences de l'acceptation et du refus dans les différentes utilisations pour les 180 observations

	Acceptation		Refus	
	effectifs	pourcentages	effectifs	Pourcentages
Tranchage	31	17.22	149	82.78
Déroulage	87	48.33	93	51.67
Ebénisterie	105	58.33	75	41.67
Menuiserie	143	79.44	37	20.56
Carcasse	165	91.67	15	8.33

Dans ce tableau issu de l'enquête, on retrouve le classement a priori des observations. Cela signifie que les personnes interrogées classent les utilisations dans le même ordre que celui déterminé avant enquête.

4.2.2- Analyse de l'incidence de la présence de défauts et de leur intensité sur le choix des utilisations par le modèle LOGIT

A présent, nous allons essayer d'expliquer l'incidence de la présence des différents défauts sur les utilisations à partir du lien LOGIT. Nous utiliserons pour cela la procédure LOGISTIC du logiciel SAS.

La matrice des variables explicatives X est constituée des variables de présence et d'intensité des défauts ; elle est de dimension : 180 x nombre de variables « défauts » du modèle. La matrice des variables à expliquer Y est constituée de l'acceptation ou du refus (1 ou 0) du bois dans une utilisation donnée ; elle est de dimension 180 x 1.

La phase descriptive réalisée précédemment nous a permis d'identifier des groupes de variables explicatives corrélées entre elles. Or, nous savons que la présence

de variables corrélées pose problème en régression pour la stabilité des estimateurs. Une solution consiste à supprimer des variables parmi celles qui sont très corrélées. Parmi chaque groupe de variables « défauts » corrélées, nous n'en n'avons gardé qu'une dans le modèle (celle qui donnait la qualité de prédiction la meilleure). Ceci ne veut pas dire que les autres défauts ne sont pas présents sur la rondelle. Par exemple, si nous gardons la variable « discoloration » dans le modèle cela ne veut pas dire qu'il n'y a pas d'échauffure, de pourriture ou de moisissure.

Pour sélectionner les variables, comparer les modèles entre eux, et évaluer le pouvoir de prédiction de ceux-ci, nous pouvions utiliser divers critères parmi ceux disponibles dans la procédure LOGISTIC du logiciel SAS® et mentionnés en 3.2.2.3.

Les estimations du maximum de vraisemblance sont effectués par la méthode du « Fisher-scoring », équivalente à l'algorithme des moindres carrés pondérés itérés.

Pour la validation des modèles, nous avons également regardé le graphe des résidus obtenus pour les différentes observations.

Pour chaque utilisation prise séparément, nous avons essayé de construire un modèle. Cela nous a déjà donné une première idée des défauts qui peuvent avoir leur importance sur les différentes utilisations. La probabilité modélisée est la probabilité qu'a la rondelle d'être refusée dans l'utilisation.

Cependant, les utilisations sont loin d'être indépendantes entre elles (§ 4.2.1.4) ; c'est pourquoi, dans un second temps, il nous est paru important de tenir compte de cette liaison.

4.2.2.1- Utilisation en tranchage

Lors de l'application de la procédure logistique pour l'utilisation en tranchage, nous avons rencontré une difficulté. En effet, les données étaient telles qu'elles posaient problème pour le calcul du maximum de vraisemblance. Nous avons l'existence de règles systématiques pour certaines des variables de défauts. Nous avons donc dû procéder à quelques modifications avant de pouvoir effectuer une régression logistique.

En examinant les tableaux de contingences de l'acceptation ou du refus en tranchage en fonction des intensités des différents défauts, nous avons remarqué qu'il semblait y avoir quelques règles de sélection ; les rondelles (de notre étude) qui ont été acceptées en tranchage présentaient les caractéristiques suivantes : pas du tout de cœur rouge sain et homogène, pas du tout ou peu de cœur rouge flammé, pas du tout ou peu de « maladie du T ». Les rondelles avec beaucoup de méplat ou de cœur excentré sont exclues du tranchage, ainsi que celles qui présentaient beaucoup de discoloration, moyennement ou beaucoup d'échauffure et de pourriture. Nous avons procédé en deux étapes : ces règles nous ont permis de faire une première sélection ; puis, sur les rondelles qui ne présentaient pas les intensités rédhibitoires des défauts ci-dessus, nous avons effectué une régression logistique pour essayer d'affiner.

Tableau 3 : Résultats du modèle logistique pour l'incidence des défauts sur l'utilisation en tranchage après première sélection

Variable	Degré de liberté	Valeur estimée du paramètre	Estimation de l'écart type de l'estimateur	Test χ^2	Probabilité critique du test χ^2
Constante	1	-0,9645	0,5463	3,1165	0,0775
Echauffure	1	1,4820	0,3848	14,8337	0,0001

La variable « échauffure » a été conservée seule dans le modèle. Rappelons cependant que les variables qui lui sont corrélées (discoloration, ...) ont également un effet sur l'acceptation ou le refus de l'utilisation.

Conclusion : le cœur rouge, la maladie du T, l'échauffure, la pourriture apparaissent fortement rédhibitoires pour l'utilisation en tranchage. La discoloration, le cœur excentré, le méplat le sont également mais dans une moindre mesure.

4.2.2.2- Utilisation en déroulage

Le tableau 4 présente les coefficients du modèle prédisant la probabilité d'être refusé en déroulage.

Tableau 4 : Résultats du modèle logistique pour l'incidence des défauts sur l'utilisation en déroulage

Variable	Degré de Liberté	Valeur estimée du paramètre	Estimation de l'écart type de l'estimateur	Test χ^2	Probabilité critique du test χ^2
Constante	1	-5,3539	0,9434	32,2036	< 0,0001
Discoloration	1	1,3961	0,2332	35,8568	< 0,0001
Cœur excentré	1	0,7263	0,2324	9,7699	0,0018
T	1	1,2441	0,4605	7,2994	0,0069

Les sorties SAS pour cette utilisation se trouvent en Annexe 13.

Etude des résidus :

Nous avons analysé les graphes des résidus dans le cas du modèle contenant les 13 variables explicatives de départ et dans le cas du modèle du Tableau 4 (après sélection des variables). Une observation ressort dans les deux cas : il s'agit de l'observation numéro 95 (photo n°3496 vue par le juge n°7) ; cette photo ne comporte en effet aucun des défauts mentionnés pourtant elle a été refusée en déroulage par le juge n°7 en raison de sa taille trop petite. Nous pouvons donc voir ici que le fait de ne pas avoir tenu compte de la taille de la rondelle dans le modèle peut poser problème.

Conclusion :

nous avons retrouvé à travers le modèle les principales observations faites par les juges pour cette utilisation lors des enquêtes.

Les variables d'altérations (discoloration, échauffure,...) sont les plus rédhibitoires ($\chi^2 = 35,8568$) probablement pour des raisons esthétiques.

Ensuite, viennent les problèmes de forme (cœur excentré, méplat). En effet, en déroulage, il est important d'avoir des grumes bien circulaires pour avoir un bon rendement. Il s'agit ici de contraintes liées à la technique du déroulage.

Enfin la maladie du T ressort également comme pouvant influencer sur le refus de l'utilisation en déroulage.

4.2.2.3- Utilisation en ébénisterie

Il s'agit ici encore d'une utilisation considérée comme noble.

Tableau 5 : Résultats du modèle logistique pour l'incidence des défauts sur l'utilisation en ébénisterie

Variable	Degré de liberté	Valeur estimée du paramètre	Estimation de l'écart type de l'estimateur	Test χ^2	Probabilité critique du test χ^2
Constante	1	-3,4125	0,6427	28,1912	< 0,0001
Discoloration	1	0,9857	0,1833	28,9067	< 0,0001
T	1	0,8571	0,3604	5,6563	0,0174

Etude des résidus :

Les graphiques ne sont pas présentés dans ce rapport pour des raisons de place mais nous avons pu remarquer que :

- dans le cas du modèle contenant l'ensemble des 13 variables de départ, l'observation n°25 est celle qui est la plus mal ajustée par le modèle et qui est également la plus influente. Il s'agit de l'observation de la photo n°3474 par le juge n°2. Le bois a été accepté par le juge en ébénisterie, mais refusé en menuiserie qui est pourtant une utilisation moins noble. De plus, la rondelle a été jugée comme présentant beaucoup de discoloration et d'échauffure. Peut-être s'agit-il d'une erreur de la part du juge ?
- dans le cas du modèle obtenu après sélection des variables, c'est l'observation n°76 (observation de la photo n°3492 par le juge n°6) qui est mise en évidence. Cette rondelle a été considérée comme présentant une forte discoloration par le juge mais il l'a pourtant acceptée en ébénisterie.

Conclusion : les altérations (discoloration, échauffure, ...) sont encore la cause de rejet la plus forte.

La maladie du T ressort dans une moindre mesure avec une P-value de 0,0174.

4.2.2.4- Utilisation en menuiserie intérieure

L'utilisation en menuiserie intérieure est déjà moins valorisante que les utilisations précédentes.

Tableau 6 : Résultats du modèle logistique pour l'incidence des défauts sur l'utilisation en menuiserie intérieure

Variable	Degré de liberté	Valeur estimée du paramètre	Estimation de l'écart type de l'estimateur	Test χ^2	Probabilité critique du test χ^2
Constante	1	-5,3525	0,7474	51,2906	< 0,0001
Discoloration	1	1,5634	0,2458	40,4421	< 0,0001

Etude des résidus :

- dans le cas du modèle contenant l'ensemble des 13 variables de départ, l'observation n°106 est celle qui est la plus mal ajustée par le modèle, et qui a également le plus d'influence sur celui-ci. Cette observation correspond à l'observation de la photo n°3497 par le juge n°8. Le juge n°8 avait refusé ce bois en menuiserie intérieure à cause de la présence d'un peu d'échauffure (voir partie 2 du formulaire) ;
- dans le cas du modèle obtenu après sélection des variables (Tableau 6), c'est avant tout l'observation n°153 (observation de la photo n°3470 par le juge n°11) qui se distingue des autres par une valeur élevée des résidus et une forte influence. Le juge n°11 avait refusé le bois en menuiserie intérieure à cause de la présence de cœur rouge. Il n'est pas étonnant que cette observation se distingue des autres car le modèle ne prend en compte que les variables d'altération et non le refus pour cause de cœur rouge.

Conclusion : ici, il apparaît que ce soit seulement la présence d'altérations (discoloration, échauffure, ...) qui semble importante pour déterminer l'acceptation ou le refus en menuiserie intérieure.

4.2.2.5- Utilisation en carcasse de meuble

Tableau 7 : Résultats du modèle logistique pour l'incidence des défauts sur l'utilisation en carcasse

Variable	Degré de liberté	Valeur estimée du paramètre	Estimation de l'écart type de l'estimateur	Test χ^2	Probabilité critique du test χ^2
Constante	1	-7,1586	1,3780	26,9858	< 0,0001
Discoloration	1	1,6580	0,3872	18,3337	< 0,0001

Conclusion : comme en menuiserie intérieure, seules les variables d'altérations semblent avoir encore de l'importance ici.

4.2.2.6- Conclusion

Les modèles ci-dessus permettent de prédire de façon satisfaisante (se référer aux pourcentages de paires concordantes donnés au § 4.2.4) la probabilité de refuser un échantillon dans une utilisation donnée.

Un des reproches que l'on peut avoir à faire à cette façon de procéder est que, pour simplifier le modèle et rendre les estimations stables, nous avons supprimé des variables parmi les groupes de variables corrélées entre elles. Dans le modèle final, nous ne pourrions pas avoir l'ensemble des variables de départ. Par exemple, parmi le groupe de variables corrélées {discoloration, échauffure, pourriture, moisissure} , nous n'en garderons qu'une seule, celle qui nous donnera le meilleur modèle du point de vue de la vraisemblance, ou du pouvoir de prédiction. Mais cela ne signifie en aucun cas qu'une seule de ces variables est influente.

Par ailleurs, les utilisations présentent également des corrélations entre elles, et il serait intéressant d'utiliser une méthode permettant de tenir compte de ces liaisons.

La régression logistique PLS semble une méthode adaptée pour tenir compte de ces deux observations.

4.2.3- Utilisation de la régression logistique sur composantes PLS pour tester l'influence des défauts sur le choix des utilisations

Comme nous l'avons vu précédemment, nous avons des groupes de variables corrélées parmi les variables explicatives « défauts ». La régression PLS permet de s'accommoder de ces corrélations et d'éviter que l'estimation des coefficients ne soit perturbée. Nous avons ainsi pu obtenir un modèle contenant l'ensemble des variables explicatives de départ. De plus, dans cette partie, nous ne traiterons plus les utilisations indépendamment les unes des autres, mais simultanément.

Nous avons utilisé la méthode de la *régression logistique sur composantes PLS* (Tenenhaus, ?). Elle consiste à effectuer une régression logistique sur les composantes issues de la régression PLS des indicatrices des modalités de Y (matrice indicatrice des utilisations composée uniquement de 0 et de 1) sur X (matrice de présence et d'intensité des défauts). D'un point de vue pratique, cette approche était directement réalisable avec les méthodes déjà implémentées dans la procédure PLS du logiciel SAS, ce qui était primordial pour des raisons de temps.

4.2.3.1- Etape 1 : régression PLS

Dans un premier temps, nous avons effectué la régression PLS des variables indicatrices des utilisations sur les variables de présence des défauts.

La matrice des variables explicatives X est constituée des variables de présence et d'intensité des défauts (intensités codées de 1 (pas du tout) à 4 (beaucoup) ; elle est de dimension : 180 x nombre de variables « défauts » du modèle (ici 13). La matrice des variables à expliquer Y est constituée de l'acceptation ou du refus (1 ou 0) du bois pour les différentes utilisations ; elle est de dimension 180 x nombre d'utilisations

considérées (ici 5 : tranchage, déroulage, ébénisterie, menuiserie intérieure, carcasse de meubles).

Le choix du nombre de composantes à extraire a été déterminé par *validation croisée*. Les techniques de validation croisée consistent à couper l'échantillon de départ en deux : un *échantillon d'apprentissage* et un *échantillon-test*. L'échantillon d'apprentissage sert pour l'estimation du modèle tandis que l' *échantillon-test* permet d'évaluer l'efficacité prédictive du modèle. On répète ce processus plusieurs fois sur des groupes différents issus du même échantillon de départ.

Nous avons utilisé l'option CV de la procédure PLS de SAS : CV=one sélectionne une validation croisée où chaque observation constitue un échantillon test ; CV=block(n) sélectionne une validation croisée où chaque bloc de n observations est inclus dans l'échantillon test.

Par validation croisée, nous avons ainsi déterminé qu'il fallait extraire 3 composantes T1, T2, T3 (Annexe 14). Ces trois composantes expliquent 41,63% de la variation des variables « défauts » et 31,45% de celles des variables « utilisations » (Annexe 14).

Nous pouvons regarder la matrice des corrélations entre les utilisations et les composantes PLS d'une part et entre les défauts et les composantes PLS d'autre part. La première composante T1 est assez fortement corrélée (négativement) au groupe des variables {discoloration, échauffure, pourriture, moisissure} et légèrement corrélée (positivement) aux variables « utilisations ». La deuxième composante est légèrement corrélée (négativement) au groupe {cœur excentré, méplat}, ainsi qu'à la présence de T et de nœuds. Par contre, elle semble légèrement corrélée (positivement) à la présence de pourriture et de moisissure. Quant à la troisième composante, elle semble davantage corrélée (négativement) au groupe {cœur excentré, méplat}.

Pour chaque observation, nous avons donc obtenu les valeurs des composantes extraites T1, T2, T3 qui sont des combinaisons linéaires des variables « défauts » de départ centrées-réduites.

4.2.3.2- Etape 2 : régression logistique sur composantes PLS

Il s'agit à présent d'effectuer une régression logistique des utilisations sur les composantes PLS : T1, T2, T3. Dans un second temps, nous reviendrons aux variables d'origine, c'est-à-dire de présence et d'intensité des défauts, non centrées, non réduites. Cette méthode nous permet d'obtenir des estimations des coefficients stables pour chacune des variables explicatives de départ, ce que ne nous permettait pas la régression logistique effectuée précédemment.

Dans le tableau 8, nous présentons le résultat obtenu pour l'utilisation en tranchage.

Tableau 8 : Coefficients du modèle pour l'utilisation en tranchage

Variable	Constant <i>e</i>	CRSH	CRF	Discoloration	Echauffur <i>e</i>	Pourritur <i>e</i>	Moisissure
Valeur estimée du paramètre	-7,0459	0,3708	0,4805	1,0124	0,9191	0,8107	0,6539

Variable	Cœur excentré	Méplat	Maladie duT	Fentes	Nœuds	Entre- écorce	Blessures
Valeur estimée du paramètre	0,1739	0,3788	0,6488	0,1494	0,4389	0,7446	0,0581

Nous avons donc obtenu des coefficients pour chaque variable explicative de départ.

Les coefficients des modèles obtenus pour les autres utilisations se trouvent en Annexe 14.

4.2.4- Comparaison des modèles utilisés

Ci-dessous se trouve un tableau qui permet d'avoir une idée sur les différences du pouvoir de prédiction des différents modèles précédemment établis.

Tableau 9 : Tableau de comparaison des trois modèles établis du point de vue de la concordance des paires

Utilisation	Modèle	Pourcentage de paires		
		Concordantes	Discordantes	ex aequo
Tranchage	1 (*)			
	2 (*)			
	3	83,7	14,9	1,4
Déroulage	1	82,1	12,8	5,1
	2	89,8	10,0	0,2
	3	89,2	10,6	0,2
Ebénisterie	1	69,4	14,7	15,9
	2	82,4	17,2	0,4
	3	80,8	18,7	0,6
Menuiserie intérieure	1	79,6	6,3	14,1
	2	91,6	8,2	0,2
	3	91,4	8,4	0,2
Carcasse de meuble	1	80,4	4,9	14,6
	2	93,0	6,5	0,4
	3	90,0	9,4	0,6

(*) : le logiciel a rencontré un problème lors de l'estimation du maximum de vraisemblance en raison de l'existence d'un critère systématique. Nous n'avons pas pu utiliser une régression logistique sur l'ensemble des variables de départ comme pour les autres utilisations.

Modèle n°1 : modèle obtenu par régression logistique après sélection des variables influentes (§ 4.2.2) ;

Modèle n°2 : modèle obtenu par régression logistique sur l'ensemble des 13 prédicteurs de départ ;

Modèle n°3 : modèle obtenu par régression logistique sur composantes PLS.

Le modèle n°1 est plus faible au niveau du pourcentage de paires concordantes. C'est normal puisque dans ce modèle nous avons ôté des variables explicatives : nous avons donc perdu de l'information. Les modèles n°2 et n°3 contiennent l'ensemble des

variables explicatives de départ ; ils sont assez semblables au niveau de leur pouvoir de prédiction. Si l'on regarde les coefficients, ils sont par contre très différents d'un modèle à l'autre. Les coefficients du modèle n°2 étant instables à cause des corrélations entre les variables, le meilleur modèle semble donc être celui obtenu par régression logistique sur composantes PLS.

5- CONCLUSION ET PERSPECTIVES

5.1- Critiques de l'enquête

Tout d'abord, nous pouvons faire une première critique quant au support de l'enquête. Ce sont des tirages papiers couleur de photos qui ont été présentés aux industriels. Cela présente quelques inconvénients :

- absence de possibilité de toucher le bois par la personne interviewée ;
- problème du rendu des couleurs par l'imprimante ;
- les photos ne sont pas à l'échelle réelle (forte réduction par rapport à la taille réelle) ; la personne interviewée doit donc faire un effort supplémentaire pour s'imaginer la taille réelle de la rondelle ; cette difficulté peut biaiser la perception ;
- certains défauts peuvent être plus difficiles à distinguer ou à différencier les uns des autres sur des photos que sur des sections réelles.

Cependant, pour des raisons pratiques, il est bien évident qu'il n'aurait pas été possible de transporter des rondelles (beaucoup trop lourd et encombrant). De plus, les photos permettent d'avoir l'aspect du bois dans son état d'humidité lors du débit (la rondelle change d'aspect quand elle sèche) ce qui correspond à la situation rencontrée par le professionnel dans la pratique.

Par ailleurs, sur le terrain, les industriels du bois ne sont pas limités à voir juste une section de grume, mais ont accès à la grume dans sa totalité (forme, apparence extérieure).

Au fil des interviews, nous nous sommes notamment rendu compte que les personnes interrogées tenaient pour la plupart peu compte, voire pas du tout, de la taille de la rondelle (mentionnée pourtant en partie 2 du formulaire) pour le choix des différentes utilisations. Cette observation a été confirmée par l'analyse des résidus de la régression logistique (§ 4.2.2).

Or, la taille de la rondelle fait partie des critères de classement de qualité du CTBA, et a son importance pour l'acceptation dans les différentes utilisations, notamment pour savoir si le rendement va être suffisant. Cela vaut surtout pour le déroulage, où il y a déjà de la perte pour mettre le billon à dérouler « au rond ». Le fait que cette variable ne fasse pas partie des prédicteurs peut affecter la qualité du modèle. Il aurait fallu mentionner ce défaut déjà lors de la partie 1 pour obliger la personne interrogée à considérer le facteur « taille ». Il aurait sans doute fallu également indiquer pour chaque photo la valeur du diamètre de la rondelle afin que la personne n'ait pas à faire un effort supplémentaire.

Un autre problème de précision des questions de l'enquête est apparu auquel nous n'avions pas pensé à propos de l'acceptation ou du refus dans les différentes utilisations : Dans la pratique, les choses ne sont pas aussi tranchées que « acceptation » ou bien « refus » de la rondelle dans l'utilisation. En effet :

- très souvent les rondelles ont été acceptées « en partie » dans l'une ou l'autre des utilisations ; c'est-à-dire après purge. A la fois l'intensité et la position des défauts avaient ici leur importance ;
- la réponse pouvait également être variable en fonction du pays d'exportation ; chaque pays a ses propres exigences en matière de qualité du bois et certains sont moins difficiles que d'autres ;
- la marché du bois (notamment l'effet de la tempête), entraînait également des hésitations au niveau des réponses.

Nous n'avons pas tenu compte dans l'enquête du fait que les rondelles étaient prises dans le billon à des hauteurs différentes les unes des autres. Par défaut, les personnes interviewées ont considéré que la rondelle provenait de la base de l'arbre (c'est-à-dire de la section de la grume qu'ils voient effectivement dans la pratique). Il y a donc homogénéité de ce point de vue au niveau de juges. Il aurait été difficilement envisageable de demander aux personnes de tenir compte de la hauteur de prise de la rondelle sachant qu'en plus elle était différente d'une rondelle à l'autre. Certaines personnes ont cependant demandé des précisions à ce sujet pour deux raisons :

- certains défauts (notamment les nœuds) se trouvent rarement aussi bas dans l'arbre ;
- savoir à quelle hauteur se trouve un défaut dans la grume peut permettre de déterminer avec quelle intensité un défaut peut se retrouver ailleurs dans la grume.

Nous leur avons répondu que la section provenait de la base de l'arbre afin de garder une certaine homogénéité entre les juges.

5.2- Critiques des analyses effectuées sur les données

La plupart du temps, nous avons considéré les variables comme quantitatives. La notion d'ordre entre les niveaux d'intensité était ainsi conservée, mais nous avons fait l'hypothèse que les niveaux d'intensité étaient équidistants, ce qui n'est pas forcément le cas. Une idée, proposée par R. Sabatier, aurait été de refaire les analyses avec des distances différentes entre les niveaux, mais toujours le même ordre bien sûr, et de voir si les résultats obtenus restaient inchangés.

Nous avons manqué de temps au niveau des dernières analyses statistiques, certains points comme l'analyse des résidus auraient mérité d'être étudiés plus en détail.

De plus, nous n'avons pas eu le temps de mettre en pratique différentes méthodes de discrimination PLS sur nos données. Nous aurions voulu comparer les résultats obtenus par la méthode « PLS-DA » à ceux obtenus par des variantes des méthodes de discrimination PLS, notamment en choisissant de bonnes métriques.

5.3- Synthèse des résultats obtenus

Plusieurs méthodes statistiques ont été utilisées afin de mettre en place des modèles de prédiction d'acceptation ou non du bois dans les utilisations considérées.

Dans un premier temps, par des analyses préliminaires, nous avons pu repérer des groupes de variables « défauts » corrélées entre elles : {discoloration, échauffure, pourriture, moisissure}, {cœur excentré, méplat} et, dans une moindre mesure, {cœur rouge sain et homogène, cœur rouge flammé}.

Les variables « utilisations » étant binaires, nous avons effectué des régressions logistiques des variables « utilisations » prises indépendamment les unes des autres sur les variables « défauts ». Ces premières analyses nous ont permis de repérer les défauts intervenant effectivement dans la prise de décision. Cependant, les variables explicatives étant pour certaines corrélées entre elles, les estimations des paramètres des modèles obtenus étaient imprécises.

Nous avons alors utilisé une méthode plus adaptée à la présence de corrélations entre les variables explicatives, la régression PLS, et plus précisément dans notre cas la régression logistique PLS. Les modèles finalement obtenus permettaient de conserver l'ensemble des variables explicatives de départ tout en ayant des estimations des coefficients plus précises.

Nous avons obtenu les conclusions suivantes :

- pour l'utilisation en tranchage, les défauts apparaissant être les plus rédhibitoires sont : le cœur rouge, la maladie du T, l'échauffure et la pourriture. Dans une moindre mesure, interviennent également le méplat, le cœur excentré et la discoloration ; il s'agit de défauts nuisant à l'esthétique du bois ;
- pour l'utilisation en déroulage, les variables d'altérations (discoloration, échauffure, pourriture) sont déterminantes ainsi que, dans une moindre mesure, le méplat, le cœur excentré et la maladie du T ; en premier lieu interviennent donc des défauts liés à l'esthétique du bois, mais nous retrouvons également des défauts « dimensionnels » qui sont gênants pour la technique du déroulage ;

- pour l'utilisation en ébénisterie, qui est considérée également comme une utilisation noble, quasiment au même titre que le tranchage pour certains industriels, ce sont les défauts d'altérations qui interviennent de façon prépondérante ; la maladie du T a également son importance ;
- enfin pour les utilisations en menuiserie intérieure et en carcasse de meuble, seules les altérations semblent encore avoir de l'importance, et de façon moins marquée pour la carcasse de meuble que pour la menuiserie intérieure.

Nous avons pu remarquer que sur l'ensemble des défauts mentionnés dans le formulaire, ce sont toujours les mêmes qui influent sur le choix des utilisations. Les nœuds, fentes, blessures cicatrisées et l'entre-écorce n'apparaissent pas avoir un effet important. Il faut noter cependant que peu de rondelles présentaient des blessures cicatrisées ou de l'entre-écorce.

Une remarque est à faire sur la « maladie du T ». Ce défaut est considéré comme très rédhibitoire pour l'utilisation du bois, à la fois par les normes de classement des bois, que par les industriels interrogés. Cependant parmi les 14 personnes interviewées (nous n'en avons gardé que 12 dans l'étude), 4 (c'est-à-dire 29 %), ne connaissaient pas ou mal ce défaut.

5.4- Perspectives

Dans toute cette étude, nous avons considéré 180 observations (60 photos vues 3 fois chacune). Peu importait le fait que la photo était répétée trois fois sur des juges différents. Seul nous intéressait le lien entre la perception des défauts par un juge et sa décision vis-à-vis des utilisations.

Il serait intéressant d'étudier les réponses des personnes interrogées en tenant compte de certaines caractéristiques telles que la taille de l'entreprise qui l'emploie, la nature des produits sortis par l'entreprise et peut-être de pondérer les réponses en

fonction de ces indications. En effet, les personnes interrogées sont naturellement plus aptes à répondre pour les utilisations dont elles ont l'habitude. Pour les autres utilisations, elles doivent se baser sur ce qu'elles ont appris ou entendu dire.

Enfin, il aurait été intéressant de demander aux personnes de mettre une note à la rondelle, ou bien encore, à partir des utilisations où le bois est accepté, de lui attribuer un score afin de pouvoir discriminer une qualité de bois à partir des défauts visibles sur une section. Nous pourrions alors utiliser une version adaptée (métriques bien choisies) de l'analyse PLS discriminante. Nous n'avons malheureusement pas eu le temps de traiter ce point pendant la durée du stage.

BIBLIOGRAPHIE

- AFNOR (1998) - *Bois ronds et bois sciés ; Terminologie.*
- AFNOR (1997) - *Bois ronds feuillus ; Classement qualitatif.*
- AFOCEL (1999) - *Mémento annuel.*
- Agresti A. (1990) - *Categorical Data Analysis*, New York : John Wiley & Sons, Inc.
- Bergonzini J.-Cl. & Duby C. (1995) - *Analyse et planification des expériences ; Les dispositifs en blocs*, Masson.
- Callier P. & Schlich P. (1995) - *Extension de la cartographie des préférences aux plans incomplets*, 4^{ème} journée Agro-industrie et méthodes statistiques (A.S.U.), Dijon, 7-8 décembre 1995.
- Cinotti B. (1999) - *Récolte de bois en France en 1998 : poursuite de l'augmentation des volumes commercialisés*, Revue forestière française, 339-349.
- Cochran W.G. & Cox G.M. (1957) - *Experimental Designs*, New York : John Wiley & Sons, Inc.
- Collett D. (1991) - *Modelling Binary Data*, Chapman & Hall.
- CTBA - *Tableau synthétique du classement professionnel du chêne et du hêtre.*
- Das M.N. & Giri N.C. (1979) - *Design and analysis of experiments*, Wiley Eastern Limited.
- Daudin J.-J. (1998) - *Etude des données discrètes par le modèle linéaire généralisé*, Institut National Agronomique, Département OMIP.
- Desbois D. (1999) - *Les Cahiers Techniques de l'INRA n°41 (Bulletin de liaison interne) : Spécial Régression PLS*, INRA.
- Métro A. (1975) - *Dictionnaire forestier multilingue*, Association française des eaux et forêts.
- Droesbeke J.-J., Fine J., Saporta G (1997) - *Plans d'expériences ; Applications à l'entreprise*, Editions Technip.

Goupy J. (1998) - *La méthode des plans d'expériences ; Optimisation du choix des essais et de l'interprétation des résultats*, Dunod.

Marx B.D. (1996) - *Iteratively Reweighted Partial Least Squares Estimation for Generalized Linear Regression*, American Statistical Association and the American Society for Quality Control, *Technometrics*, 374-381.

McCullagh P. & Nelder J. A. (1989) - *Generalized Linear Models*, Chapman & Hall.

Myers R.H. & Montgomery D.C. (1997) - *A Tutorial on Generalized Linear Models*, *Journal of Quality Technology*, 274-291.

Nepveu G., Protocole 98 (note interne à l'équipe Qualité des Bois du Lerfob).

Perrin R. (1981) - Maladie du « T », Le Hêtre, monographie INRA, 508-510.

Robin-Huon de Kermadec F. (1996), *Méthodes statistiques permettant d'expliquer l'appréciation hédonique par les caractéristiques sensorielles*, thèse de l'université Montpellier II soutenue le 2 décembre 1996.

Sabatier R., *A propos de PLS discriminant : deux nouvelles propositions*.

Sabatier R. (2000-2001), cours au DEA biostatistique de Montpellier.

Saporta G. (1990) - *Analyse des données et statistiques*, Technip, Paris.

SAS/STAT Users' Guide, SAS OnlineDoc Version 8.

Tenenhaus M. (2000), *La régression logistique PLS*, 32^{ème} journées statistiques, 15-19 mars 2000.

Tenenhaus M. (1995) - *A partial least squares approach to multiple regression, redundancy analysis and canonical analysis*, Les cahiers de la recherche de HEC, CR 550/1995.

Tenenhaus M., *Une approche aux moindres carrés partiels de l'analyse canonique et de l'analyse des redondances*, Fondation HEC.

Wold S. (1996), *PLS Regression : The State of Art*.

ANNEXES

TABLE DES ANNEXES

- Annexe 1 :** Illustrations de certains défauts présents sur les rondelles
- Annexe 2 :** Caractéristiques des billons dont sont issues les rondelles de l'expérience
- Annexe 3 :** Calcul de la surface altérée observée sur les photos de sections
- Annexe 4 :** Exemples de photos de rondelles qui ont été soumises à l'appréciation du panel des douze professionnels
- Annexe 5 :** Formulaire d'enquête
- Annexe 6 :** Algorithme pour la construction du plan d'expérience
- Annexe 7 :** Programme VBA pour la répartition des photos entre les juges
- Annexe 8 :** Tableau de répartition des 60 photos sur les 12 juges
- Annexe 9 :** Remarques préliminaires sur les juges
- Annexe 10 :** Analyse de variance pour les effets « juge » et « photo »
- Annexe 11 :** Corrélations des variables défauts
- Annexe 12 :** Résultats complémentaires sur l'ACP
- Annexe 13 :** Le modèle LOGIT
- Annexe 14 :** Régression logistique sur composantes PLS
- Annexe 15 :** Pourcentage d'acceptation prédit pour différentes utilisations du bois de Hêtre (tranchage, déroulage, ébénisterie, menuiserie intérieure, carcasse de meuble) en fonction de la note d'altération attribuée par les professionnels à la section de grumes.
- Annexe 16 :** Photographies numériques de sections de grumes présentant des notes d'altération contrastées (notes 0 à 1, 2 à 3, 4 à 6, 7 à 9, 10 à 12)

ANNEXE 1.

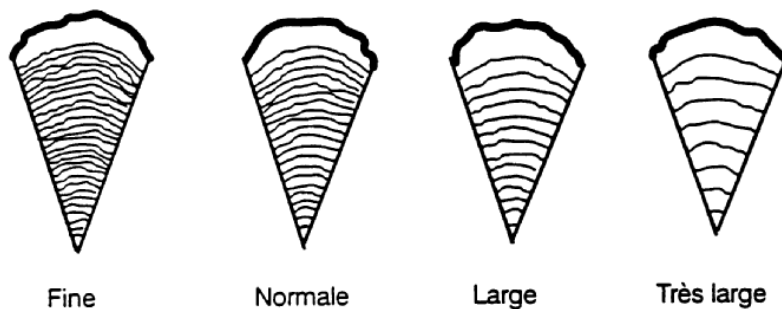
ILLUSTRATIONS DE CERTAINS DEFAUTS PRESENTS SUR LES RONDELLES

Les schémas ci-dessous sont empruntés à un document interne du CTBA
(Département Bois et Sciage).

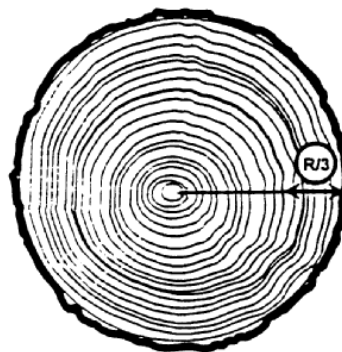
Caractéristiques de structure :

- Largeur d'accroissement

LA LARGEUR D' ACCROISSEMENT



METHODE DE MESURE



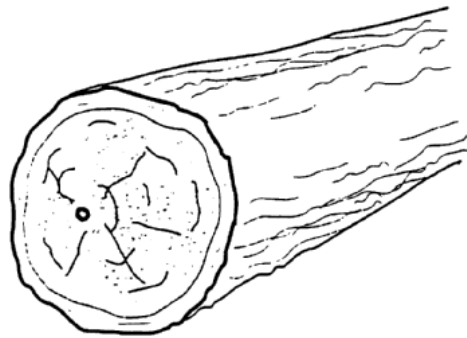
$$l = \frac{(R/3)}{n}$$

Dans la formule ci-dessus, c'est R/4 qui est à présent utilisé à la place de R/3.

Singularités de structure :

- **Cœur excentré**

COEUR EXCENTRE



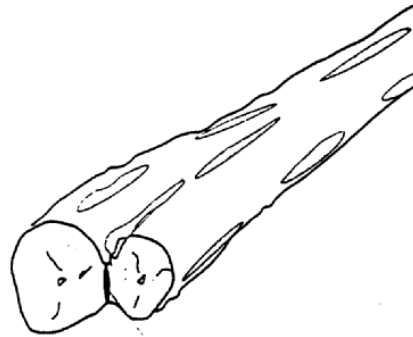
METHODE DE MESURE



Coeur excentré en % = $I/D \times 100$

- **Entre-écorce**

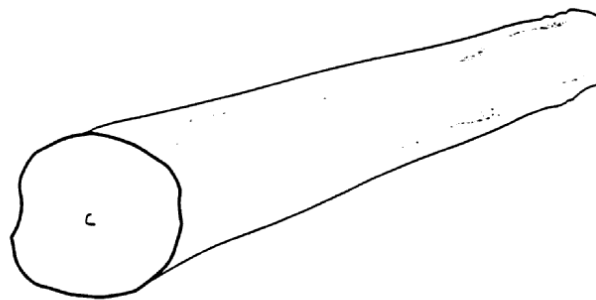
L' ENTRE ECORCE



Entre écorce

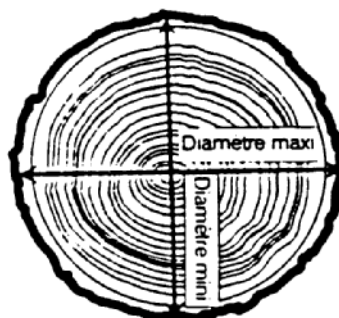
- **Méplat**

MEPLAT



Méplat

METHODE DE MESURE

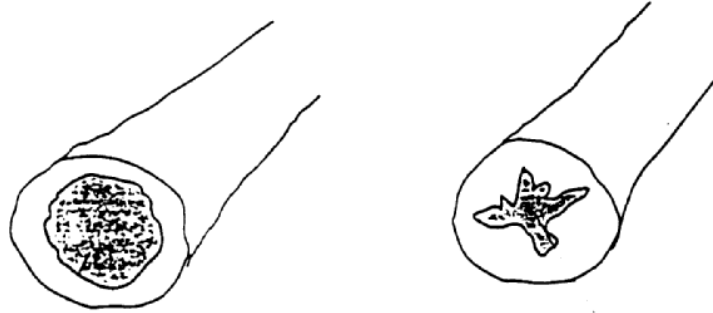


$$\text{Méplat en \%} = (d \text{ max} - d \text{ min}) / d \text{ max} \times 100$$

Altérations dues à l'action des champignons :

▪ **Cœur rouge**

COEUR ROUGE DU HÊTRE

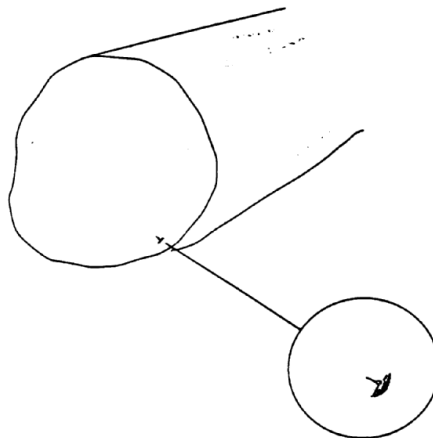


Coeur rouge
sain et homogène

Coeur rouge
flammé

▪ **maladie du T**

MALADIE DU T



ANNEXE 2.

CARACTERISTIQUES DES BILLONS DONT SONT ISSUES LES RONDELLES DE L'EXPERIENCE

Le tableau qui suit, réalisé par les techniciens de l'équipe « Qualité des Bois », contient les caractéristiques des billons utilisés dans l'expérience. Ce tableau ne concerne pas uniquement notre étude, toutes les colonnes ne nous intéressent pas, mais il n'en a pas été modifié pour autant.

Dans la troisième colonne « numéro QB », il s'agit du numéro de la photo.

La huitième colonne « D (cm) » correspond au diamètre de la rondelle en cm.

Les colonnes n°15 et n°16 correspondent respectivement à la hauteur du bas du billon, et à la hauteur du haut du billon (là où a été prise la rondelle photographiée), dans la grume.

En raison de difficultés techniques intervenues au moment du tirage du mémoire, le tableau annoncé n'a pu être reproduit.

ANNEXE 3.

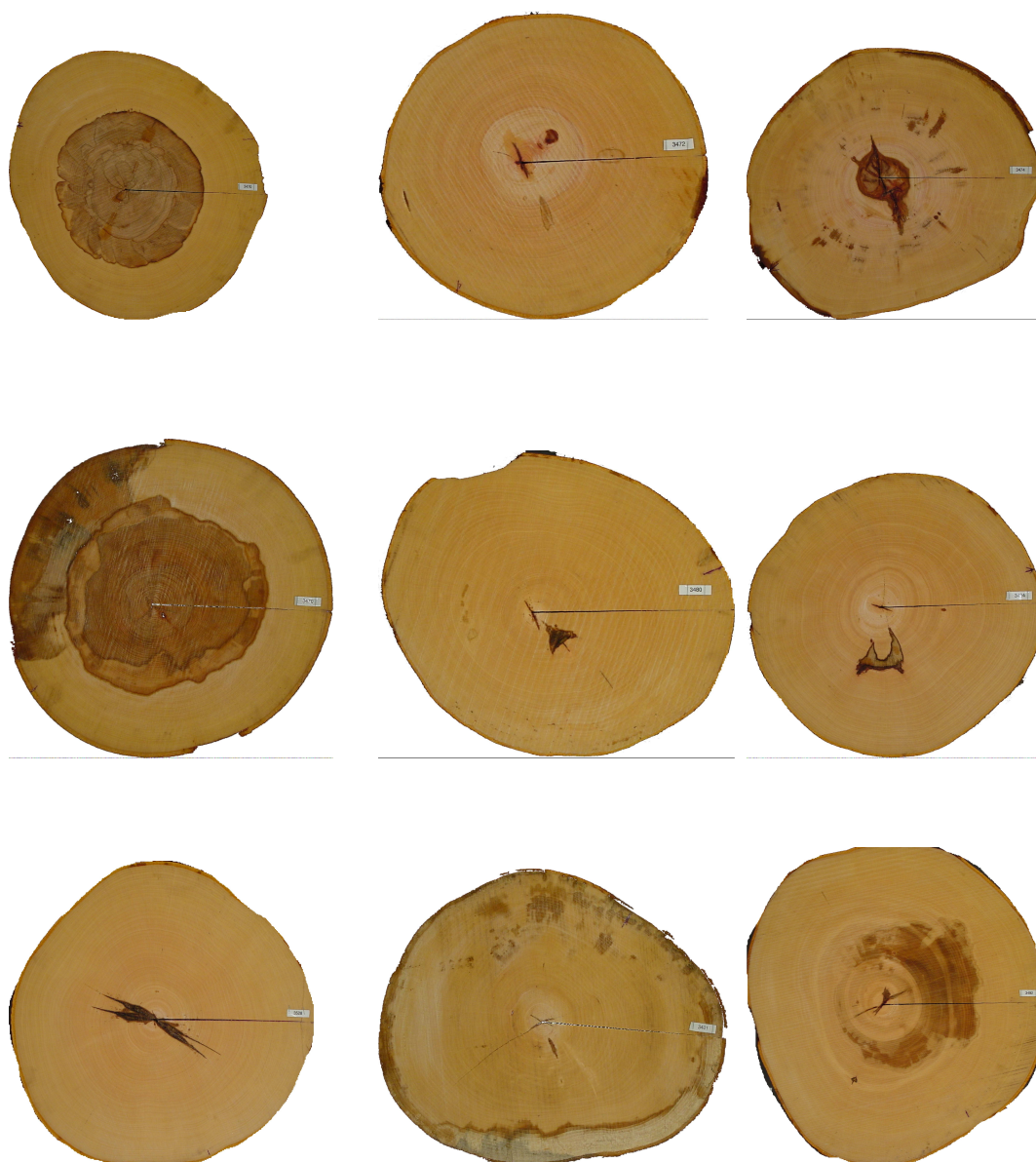
CALCUL DE LA SURFACE ALTEREE OBSERVEE SUR LES PHOTOS DE SECTIONS

Le tableau suivant contient pour chacune des photos, le pourcentage de la surface présentant des défauts. Il a été réalisé par Pierre Gelhaye (technicien INRA) à partir du logiciel Visilog.

En raison de difficultés techniques intervenues au moment du tirage du mémoire, le tableau annoncé n'a pu être reproduit.

ANNEXE 4.

EXEMPLES DE PHOTOS DE RONDELLES QUI ONT
ETE SOUMISES A L'APPRECIATION DU PANEL DES
DOUZE PROFESSIONNELS



ANNEXE 5.

FORMULAIRE D'ENQUÊTE

IDENTITE

Nom :

Prénom :

Age :

Quel est votre rôle dans l'entreprise ? :

Type de formation sur le bois :

Scolaire ou universitaire

Interne entreprise

Autodidacte

Autre, précisez : _____

Avez-vous suivi une formation sur le classement des bois ronds ?

Si oui, dans quel cadre était-ce ?

Nombre d'années d'expérience dans le milieu du bois :

Nombre d'années dans l'entreprise :

Nombre d'années dans le poste :

Nom de l'entreprise :

Nature des produits sortis par l'entreprise :

Le bois arrive-t-il sous forme de grumes dans l'entreprise ? :

Estimation du nombre de m³ de bois/an entrant dans l'entreprise :

Estimation de la part du Hêtre :

Estimation du nombre de salariés :

PARTIE 1

Indiquez le numéro à 4 chiffres présent sur la photo : _____

Indiquez le numéro d'ordre de la photo : _____

Cochez les caractéristiques que vous pensez présentes dans l'échantillon :

Cœur rouge sain et homogène	Pas du tout	<input type="checkbox"/>	Moisissure, Bleuissement	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Cœur rouge flammé	Pas du tout	<input type="checkbox"/>	Cœur excentré	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Discoloration (sauf cœur rouge)	Pas du tout	<input type="checkbox"/>	Entre-écorce	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Echauffure	Pas du tout	<input type="checkbox"/>	Maladie du T	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Pourriture	Pas du tout	<input type="checkbox"/>	Blessures cicatrisées	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Nœuds	Pas du tout	<input type="checkbox"/>	Méplat	Pas du tout	<input type="checkbox"/>
	Peu	<input type="checkbox"/>		Peu	<input type="checkbox"/>
	Moyennement	<input type="checkbox"/>		Moyennement	<input type="checkbox"/>
	Beaucoup	<input type="checkbox"/>		Beaucoup	<input type="checkbox"/>
Fentes	Pas du tout	<input type="checkbox"/>	Autres, précisez : _____		
	Peu	<input type="checkbox"/>			
	Moyennement	<input type="checkbox"/>			
	Beaucoup	<input type="checkbox"/>			

PARTIE 2

D'après vous, le bois dont est issue la photo est-il de qualité suffisante pour les usages suivants ?
(Attention en ce qui concerne la taille de la rondelle, la photo est une réduction de la taille réelle.
Les caractéristiques non visibles sur la photo seront considérées comme satisfaisantes.)
Si vous choisissez « non » pour un usage, cochez ce qui vous fait le refuser.

CATEGORIE A

Tranchage oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Déroutage oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Catégorie B

Ebénisterie (meubles massifs ...)

oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Menuiserie (lambris, plinthes, parquets, ...)

oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Carcasses de meubles oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Catégorie C

Palette, bois de calage, bois de traverses, ... oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture
- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

- Nœuds
- Fentes
- Moisissure
- Cœur excentré
- Entre-écorce
- Maladie du T
- Blessures
- Méplat
- Trop petit
- Autres, précisez : _____

Bois de trituration (panneaux de particules, pâte à papier, ...)

oui non

Si non, pourquoi ?

- Cœur rouge
- Discoloration
- Echauffure
- Pourriture

Catégorie D

Bois de chauffage, pièces difficilement valorisables oui non

Autres remarques éventuelles que vous auriez à faire sur l'échantillon :

ANNEXE 6.

ALGORITHME POUR LA CONSTRUCTION DU PLAN D'EXPERIENCE

Le but de cet algorithme est de répartir les 60 photos entre les 12 juges, sachant que chaque photo est vue 3 fois et qu'un juge en voit 15.

Pour les 8 premiers juges, les photos ont été attribuées manuellement de façon à ce que les 4 premiers juges voient une fois, à eux tous, les 60 photos et que les 4 suivants voient une nouvelle fois les 60 photos, et de façon à « casser » au maximum les associations tout en respectant les contraintes précisées au paragraphe 2.3. L'algorithme intervient pour l'attribution des photos aux juges 9 à 12. Le Tableau 6.1 ci-dessous correspond à la répartition obtenue pour les 24 premières photos.

Tableau 6.10 : Répartition des 24 premières photos sur les 12 juges. Un 1 signifie que le juge voit la photo.

	1 2 3 4	5 6 7 8	9 1 1 1 0 1 2	1 1 1 1 3 4 5 6	1 1 1 2 7 8 9 0	2 2 2 2 1 2 3 4
1	1 1 1 1	1 1 1 1	1 1 1 1	1 1 1 0	0 0 0 0	0 0 0 0
2	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 1	1 1 1 1	1 1 1 1
3	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
4	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
5	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0
6	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0
7	0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0
8	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1
9	1 0 0 0	0 1 0 0	0 0 1 0	0 0 0 1	0 0 0 1	0 0 1 0
10	0 1 0 0	1 0 0 0	0 0 0 1	0 0 1 0	1 0 0 0	0 1 0 0
11	0 0 1 0	0 0 0 1	1 0 0 0	0 1 0 0	0 1 0 0	1 0 0 0
12	0 0 0 1	0 0 1 0	0 1 0 0	1 0 0 0	0 0 1 0	0 0 0 1

Principe :

Pour les juges 9 à 12 nous raisonnons par blocs de 4 sur 4. Dans ces blocs, il faut qu'il y ait une sélection par ligne et par colonne (on introduit donc une contrainte supplémentaire). Pour chaque photo, nous allons déterminer quelles lignes sont encore

libres dans le remplissage du bloc en cours. Cette contrainte limite le nombre de plans possibles mais en facilite la construction. Le plan obtenu ne sera bien sûr pas optimal, mais il ne sera pas trop mauvais non plus.

Tableau 6.11 : Exemple de sélection dans un bloc de 4 photos. Un 1 à l'intersection d'un juge et d'une photo signifie que le juge voit la photo.

1	0	0	0
0	0	0	1
0	1	0	0
0	0	1	0

Exemple de fonctionnement de l'algorithme :

Nous construisons le tableau par ordre croissant du numéro de photo.

(1) Soit la photo n°7 (Voir Tableau 6.1). Pour cette photo, nous savons quels sont les juges parmi les juges 1 à 8 qui la verront : il s'agit des juges n°1 et n°7. Nous cherchons à définir le juge parmi les 4 derniers (numéros 9 à 12) qui la verra.

Dans le bloc de 4 auquel elle appartient, on regarde les lignes (juges) qui n'ont pas encore été sélectionnées. Il faut choisir parmi ces juges celui qui verra la photo n°7. Les juges numéros 9 et 10 du bloc sont déjà occupés (Voir Tableau 6.3 : ils voient déjà les photos n°5 et n°6). Il ne reste donc que les juges n°11 et n°12.

Tableau 6.12 : En lignes, nous avons les juges 9 à 12. En colonnes, nous avons les photos 5, 6 et 7. Une croix signifie que la case ne peut être occupée.

	5	6	7
9		1	×
10	1		×
11			
12			

Si nous choisissons le juge n°11 pour voir la photo n°7, la photo n°7 aura été vue 3 fois avec la photo n°3 sur l'ensemble des juges (Voir Tableau 6.13).

Tableau 6.13 : Nombre d'apparitions de la photo n°7 avec les photos n°1, 2, 3, 4, 5 et 6 si on choisit qu'elle soit vue par le juge n°11.

N° photo	1	2	3	4	5	6
Occurrences	1	1	3	1	1	1

Si nous choisissons que la photo n°7 soit vue par le juge n°12, alors les paires sont mieux réparties (Voir Tableau 6.14).

Tableau 6.14 : Nombre d'apparitions de la photo n°7 avec les photos n°1, 2, 3, 4, 5 et 6 si on choisit qu'elle soit vue par le juge n°12.

N° photo	1	2	3	4	5	6
Occurrences	1	1	2	2	1	1

Donc le juge n°12 verra la photo n°7.

(2) Ensuite on passe à la photo n°8. Comme c'est la dernière du bloc, on n'a pas le choix. D'après les contraintes, il n'y a que le juge n°11 qui puisse la voir.

(3) Photo n°9 : c'est la première d'un nouveau bloc de 4 photos. La comparaison aura lieu entre les juges n°9 à 12.

Tableau 6.15 : Nombre d'apparitions de la photo n°9 avec les photos n°1, 2, 3, 4, 5, 6, 7 et 8 si on choisit qu'elle soit vue par le juge n°9.

N° photo	1	2	3	4	5	6	7	8
Occurrences	3	1	1	1	2	2	1	1

Tableau 6.16 : Nombre d'apparitions de la photo n°9 avec les photos n°1, 2, 3, 4, 5, 6, 7 et 8 si on choisit qu'elle soit vue par le juge n°10.

N° photo	1	2	3	4	5	6	7	8
Occurrences	2	2	1	1	3	1	1	1

Tableau 6.17 : Nombre d'apparitions de la photo n°9 avec les photos n°1, 2, 3, 4, 5, 6, 7 et 8 si on choisit qu'elle soit vue par le juge n°11.

N° photo	1	2	3	4	5	6	7	8
Occurrences	2	1	2	1	2	1	1	2

Tableau 6.18 : Nombre d'apparitions de la photo n°9 avec les photos n°1, 2, 3, 4, 5, 6, 7 et 8 si on choisit qu'elle soit vue par le juge n°12.

N° photo	1	2	3	4	5	6	7	8
Occurrences	2	1	1	2	2	1	2	1

Pour les juges numéros 9 et 10, certaines paires sont présentées 3 fois.

Le choix se fera entre les juges 11 et 12 car les paires sont mieux réparties. Quand deux juges sont équivalents (les juges 11 et 12 ont le même nombre de paires contenant la photo n°9 présentées 2 fois), on choisit le premier. Donc ici, c'est le juge n°11 qui verra la photo n°9.

L'algorithme ci-dessus permet de déterminer quelles photos seront présentées à quels juges. Il a été programmé en VBA (voir Annexe 7). L'ordre de présentation est déterminé par randomisation, programmée elle aussi en VBA.

ANNEXE 7.

PROGRAMME VBA POUR LA REPARTITION DES PHOTOS ENTRE LES JUGES

```
Sub plan()  
'  
' plan Macro  
' Macro enregistrée le 28/03/01 par longueta  
' Nom du fichier : Prog_plan.xls  
  
'Cette macro a pour but de déterminer les 15 échantillons  
' parmi les 60 qui seront présentés à chacun des 12 juges  
' chaque photo est vue 3 fois au total  
' un 1 signifie que la photo est vue par le juge,  
' un 0 signifie le contraire  
  
'déclaration des variables locales  
Dim occurrence(1 To 60) As Integer  
Dim occurrence_bis(1 To 4, 1 To 60) As Integer  
Dim occupe(1 To 4) As Integer  
Dim maxi(1 To 4) As Integer  
Dim nb_minimax(1 To 4) As Integer  
'le tableau occurrence contient le nombre de fois où la photo en  
' cours a déjà été vu avec les photos précédentes  
'le tableau occupe permet de savoir ou on peut remplir les blocs  
' de 4 sur 4  
'le tableau occurrence_bis va servir à la comparaison pour le  
' remplissage des 4 derniers juges  
  
'*****initialisation du plan à 0*****  
Range("A1").Select  
ActiveCell.Value = 0  
Selection.AutoFill Destination:=Range("A1:BH1"), Type:=xlFillDefault  
Range("A1:BH1").Select  
Selection.AutoFill Destination:=Range("A1:BH12"), Type:=xlFillDefault  
  
'*****4 premiers juges*****  
'photos vues par les 4 premiers juges  
'juge n°1  
Range("A1").Select  
ActiveCell.Value = 1  
Selection.AutoFill Destination:=Range("A1:O1"), Type:=xlFillDefault  
  
'juge n°2  
Range("P2").Select  
ActiveCell.Value = 1  
Selection.AutoFill Destination:=Range("P2:AD2"), Type:=xlFillDefault  
  
'juge n°3  
Range("AE3").Select  
ActiveCell.Value = 1  
Selection.AutoFill Destination:=Range("AE3:AS3"), Type:=xlFillDefault  
  
'juge n°4  
Range("AT4").Select  
ActiveCell.Value = 1  
Selection.AutoFill Destination:=Range("AT4:BH4"), Type:=xlFillDefault  
  
'*****4 juges suivants*****  
'photos vues par les 4 juges suivants  
'juge n°5  
Range("A5").Select  
j = 1  
ActiveCell.Value = 1
```

```

Do While j <= 60 - 4
ActiveCell.Offset(0, 4).Select
j = j + 4
ActiveCell.Value = 1
Loop

'juge n°6
Range("B6").Select
j = 2
ActiveCell.Value = 1

Do While j <= 60 - 4
ActiveCell.Offset(0, 4).Select
j = j + 4
ActiveCell.Value = 1
Loop

'juge n°7
Range("C7").Select
j = 3
ActiveCell.Value = 1

Do While j <= 60 - 4
ActiveCell.Offset(0, 4).Select
j = j + 4
ActiveCell.Value = 1
Loop

'juge n°8
Range("D8").Select
j = 4
ActiveCell.Value = 1

Do While j <= 60 - 4
ActiveCell.Offset(0, 4).Select
j = j + 4
ActiveCell.Value = 1
Loop

*****
*****4 derniers juges*****

'je place la première croix (un 1) pour le juge n°9
Range("A9").Select
ActiveCell.Value = 1

'numéro de la colonne courante
'initialisation
Range("B1").Select
ligne = 1
colonne = 2 'colonne de la photo n°2

'la boucle ci dessous est à faire tourner pour les différentes colonnes
Do While colonne <= 60 'boucle sur les colonnes
'on regarde avec quelle(s) photo(s) la photo n°colonne a déjà été vu
'par les 8 premiers juges

'initialisation des tableaux à 0
For i = 1 To 60
occurence(i) = 0
Next i

For i = 1 To 4
For j = 1 To 60
occurence_bis(i, j) = 0
Next j
Next i

For i = 1 To 4
occupe(i) = 0
maxi(i) = 100
nb_minimax(i) = 100
Next i
'fin initialisation

```

```

Do While ligne <= 8 'boucle sur les lignes <=8
'je calcule les occurrences de la photo n°colonne avec les photos
'qui la précèdent sur les 8 premières lignes (juges)

If ActiveCell.Value = 1 Then
For k = 1 To colonne - 1
If ActiveCell.Offset(0, -k) = 1 Then
occurrence(colonne - k) = occurrence(colonne - k) + 1
End If
Next k
End If
ligne = ligne + 1
ActiveCell.Offset(1, 0).Select
Loop 'fin boucle sur les lignes <=8
'à la sortie de la boucle ci-dessus on se situe dans la ligne n°9

'on veut remplir la dernière partie du tableau (4 derniers juges)
'on raisonne par blocs de 4 photos. Dans ces blocs de 4 sur 4
'on veut qu'il y ait une photo par juge
'dans un de ces blocs on ne peut donc remplir que les lignes
'qui ne l'ont pas encore été

'références lignes et colonnes
indice_colonne = colonne Mod 4 'on ne regarde que dans le bloc de 4/4
If indice - colonne = 0 Then
indice_colonne = 4
End If

Do While ligne <= 12 'boucle 1
indice_ligne = ligne Mod 4
If indice_ligne = 0 Then
indice_ligne = 4
End If
'si indice_colonne=1 on est dans la première colonne du bloc donc
'pas de contrainte pour les lignes
If indice_colonne > 1 Then 'on est dans la 2 ou 3 ème colonne du bloc
For k = 1 To indice_colonne - 1
If ActiveCell.Offset(0, -k) = 1 Then
occupe(indice_ligne) = 1 'on regarde à quelles lignes on a droit
End If
Next k
End If

'si on est dans la dernière colonne du bloc indice_colonne = 0
'il n y a plus qu'une seule possibilité pour placer la croix
If indice_colonne = 0 Then
For k = 1 To 3
If ActiveCell.Offset(0, -k) = 1 Then
occupe(indice_ligne) = 1
End If
Next k
End If

ligne = ligne + 1
ActiveCell.Offset(1, 0).Select
Loop 'fin boucle 1
'à la sortie de cette boucle on est dans la colonne courante et
'dans la 13 ème ligne

*****phase de remplissage*****
'maintenant il faut tenir compte des deux tableaux de contraintes
'"occurrence" et "occupe" afin de placer la croix
'on se place dans une ligne libre et on regarde les occurrences en
'supposant qu'on placerait une croix dans cette ligne
'après on choisira la ligne ou le nombre de répétitions des paires
'est le plus faible

ActiveCell.Offset(-4, 0).Select
ligne = 9 'on se replace dans la ligne n°9

n = 0
n = indice_colonne - 1

'cas où une seule ligne de libre --> pas le choix

```

```

If n = 3 Then 'If boucle 2
For i = 1 To 4
If occupe(i) = 0 Then libre = i
Next i
Cells(libre + 8, colonne).Select
ActiveCell.Value = 1

Else 'boucle 2 : plusieurs lignes possibles
Do While ligne <= 12 'boucle 3 sur lignes
indice_ligne = ligne Mod 4
If indice_ligne = 0 Then
indice_ligne = 4
End If

'on regarde si on peut placer une croix dans cette ligne
If occupe(indice_ligne) = 1 Then
'la ligne est déjà occupée
ligne = ligne + 1
ActiveCell.Offset(1, 0).Select

Else 'si la ligne pas encore occupée
For i = 1 To colonne - 1
occurrence_bis(indice_ligne, i) = occurrence(i)
Next i
For k = 1 To colonne - 1
If ActiveCell.Offset(0, -k) = 1 Then
occurrence_bis(indice_ligne, colonne - k) = occurrence_bis(indice_ligne, colonne - k) + 1
End If
Next k
ligne = ligne + 1
ActiveCell.Offset(1, 0).Select
End If

Loop 'fin boucle 3 sur lignes
'a la fin de cette boucle on est dans la ligne 13

'dans le cas ou on a plusieurs lignes de possibles
'il va falloir comparer les occurrences
For i = 1 To 4
If occupe(i) = 0 Then
maxi(i) = occurrence_bis(i, 1)
For j = 2 To colonne - 1
If occurrence_bis(i, j) > maxi(i) Then
maxi(i) = occurrence_bis(i, j)
End If
Next j
End If
Next i

'après on choisit la ligne dont le max est le plus petit parmi les
'lignes vides = 1er critere
mini = maxi(1)
ligne = 9
'si le min des max (minimax) a lieu dans la première ligne c'est
'elle qu 'on va remplir
For i = 2 To 4
If maxi(i) < mini Then
mini = maxi(i)
End If
Next i

'il peut y avoir des lignes qui ont le même minimax et dans ce cas
'il faut choisir la ligne que ce minimax soit présent le minimum de fois
For i = 1 To 4
If occupe(i) = 0 And maxi(i) = mini Then
'dans les lignes du minimax
'il faut compter le nombre de fois ou apparaît ce minimax
'ini
nb_minimax(i) = 0

For j = 1 To colonne - 1
If occurrence_bis(i, j) = mini Then
nb_minimax(i) = nb_minimax(i) + 1
'comptage du nombre de minimax
End If
Next j

```

```
End If
Next i

mini_nb = nb_minimax(1)
ligne = 9
For i = 2 To 4
If nb_minimax(i) < mini_nb Then
mini_nb = nb_minimax(i)
ligne = 8 + i
End If
Next i

'-->choix du juge pour la photo n°colonne
Cells(ligne, colonne).Select
ActiveCell.Value = 1

End If 'fin boucle 2

'on change de colonne et on recommence
colonne = colonne + 1
Cells(1, colonne).Select
ligne = 1
Loop 'fin de la boucle sur les colonnes

End Sub
```

ANNEXE 8.

TABLEAU DE REPARTITION DES 60 PHOTOS SUR LES 12 JUGES

En raison de difficultés techniques intervenues au moment du tirage du mémoire, le tableau annoncé n'a pu être reproduit.

ANNEXE 9.

REMARQUES PRELIMINAIRES SUR LES JUGES

Ces remarques ont été faites lors des enquêtes et ne sont pas le résultat d'analyses statistiques.

Tableau 9.1 : Remarques sur le comportement des personnes interrogées

Numéro du juge	Remarques préliminaires
1	Connaît tous les termes du questionnaire.
2	Ne traite que 1% de Hêtre → mauvaise connaissance des termes. On décide d'ôter ce juge de l'étude et d'en interroger un nouveau.
2 bis	Connaît tous les termes du questionnaire.
3	Connaît tous les termes du questionnaire. Juge sévère.
4	Connaît tous les termes du questionnaire. La largeur de cernes est un facteur de première importance pour ce juge.
5	Connaît tous les termes du questionnaire.
6	Connaît tous les termes du questionnaire.
7	Connaît tous les termes du questionnaire.
8	Bonne connaissance des termes. Il s'est arrêté sur la présence de points sur les photos. Etrange appréciation de l' excentricité du cœur .
9	Ne connaît pas la maladie du T → pour ce juge la maladie du T est une altération liée au temps (météo).
10	Connaît tous les termes du questionnaire.
11	Mauvaise connaissance de la maladie du T .
11 bis	Bonne connaissance des termes mais évaluation étrange de la présence d' entre-écorce .
12	Ne connaît pas la maladie du T . A complété quand même après explications.

Les analyses sont basées sur un panel de 12 juges.

Le juge n°2 qui connaissait assez mal le Hêtre a été remplacé par le juge n°2 bis ; de même, le juge n°11 a été remplacé par le juge n°11 bis.

ANNEXE 10.

ANALYSES DE VARIANCE POUR LES EFFETS

« JUGE » ET « PHOTO »

Tableau 19.1 : Analyse de variance pour la variable « entre-écorce »

General Linear Models Procedure					
Dependent Variable: ENTREEC					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	70	25.7869894	0.3683856	1.92	0.0011
Error	109	20.9505106	0.1922065		
Corrected Total	179	46.7375000			
	R-Square	C.V.	Root MSE	ENTREEC Mean	
	0.551741	37.84866	0.43841	1.15833	
Source	DF	Type I SS	Mean Square	F Value	Pr > F
PHOTO	59	16.2375000	0.2752119	1.43	0.0534
JUGE	11	9.5494894	0.8681354	4.52	0.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
PHOTO	59	11.5494894	0.1957541	1.02	0.4592
JUGE	11	9.5494894	0.8681354	4.52	0.0001
Bonferroni (Dunn) T tests for variable: ENTREEC					
NOTE: This test controls the type I experimentwise error rate, but generally has a higher type II error rate than REGWQ.					
Alpha= 0.05 df= 109 MSE= 0.192207					
Critical Value of T= 3.47					
Minimum Significant Difference= 0.5548					
Means with the same letter are not significantly different.					
Bon Grouping		Mean	N	JUGE	
A		2.0667	15	11bis	
B		1.2000	15	2bis	
B					
B		1.1333	15	7	
B					
B		1.1333	15	3	
B					
B		1.1333	15	4	
B					
B		1.1000	15	6	
B					
B		1.0667	15	5	
B					
B		1.0667	15	8	
B					
B		1.0000	15	12	
B					
B		1.0000	15	10	

B			
B	1.0000	15	1
B			
B	1.0000	15	9

Tableau 10.2 : Analyse de variance pour la variable « maladie du T »

General Linear Models Procedure					
Dependent Variable: T					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	70	63.2114906	0.9030213	3.55	0.0001
Error	109	27.7385094	0.2544817		
Corrected Total	179	90.9500000			
	R-Square	C.V.	Root MSE		T Mean
	0.695014	38.31356	0.50446		1.31667
Source	DF	Type I SS	Mean Square	F Value	Pr > F
PHOTO	59	36.2000000	0.6135593	2.41	0.0001
JUGE	11	27.0114906	2.4555901	9.65	0.0001
Source	DF	Type III SS	Mean Square	F Value	Pr > F
PHOTO	59	37.8614906	0.6417202	2.52	0.0001
JUGE	11	27.0114906	2.4555901	9.65	0.0001
Bonferroni (Dunn) T tests for variable: T					
NOTE: This test controls the type I experimentwise error rate, but generally has a higher type II error rate than REGWQ.					
Alpha= 0.05 df= 109 MSE= 0.254482					
Critical Value of T= 3.47					
Minimum Significant Difference= 0.6384					
Means with the same letter are not significantly different.					
Bon Grouping	Mean	N	JUGE		
A	2.5333	15	9		
B	1.3333	15	3		
B	1.3333	15	12		
B	1.2667	15	2bis		
B	1.2667	15	7		
B	1.2000	15	8		
B	1.2000	15	4		
B	1.2000	15	10		
B	1.1333	15	6		
B	1.1333	15	11bis		
B	1.1333	15	5		
B	1.0667	15	1		

ANNEXE 11.

CORRELATIONS DES VARIABLES DEFAUTS

Les variables de présence et d'intensité des défauts (§2.5 p18) étant de nature ordinaire, nous avons utilisé des méthodes adaptées : coefficient de corrélation de Spearman et Tau de Kendall.

Tableau 11.1 : Coefficients de corrélation de Spearman

Correlation Analysis						
13 'VAR' Variables:						
	CRSH	CRF	DISCO	ECHAUFF	POURRIT	
	NOEUDS	FENTES	MOISSISS	CEX	ENTREEC	
	T	BLESS	MEPLAT			
Simple Statistics						
Variable	N	Mean	Std Dev	Median	Minimum	Maximum
CRSH	180	1.2917	0.7729	1.0000	1.0000	4.0000
CRF	180	1.2222	0.6478	1.0000	1.0000	4.0000
DISCO	180	2.1111	1.0722	2.0000	1.0000	4.0000
ECHAUFF	180	1.9778	1.0881	2.0000	1.0000	4.0000
POURRIT	180	1.3222	0.7064	1.0000	1.0000	4.0000
NOEUDS	180	1.2889	0.5642	1.0000	1.0000	3.0000
FENTES	180	1.5000	0.8620	1.0000	1.0000	4.0000
MOISSISS	180	1.3111	0.7570	1.0000	1.0000	4.0000
CEX	180	1.6667	0.8969	1.0000	1.0000	4.0000
ENTREEC	165	1.0758	0.2739	1.0000	1.0000	2.5000
T	165	1.2061	0.5576	1.0000	1.0000	4.0000
BLESS	180	1.2861	0.5718	1.0000	1.0000	4.0000
MEPLAT	180	1.5667	0.8659	1.0000	1.0000	4.0000
Spearman Correlation Coefficients / Prob > R under Ho: Rho=0 / Number of Observations						
	CRSH	CRF	DISCO	ECHAUFF	POURRIT	
CRSH	1.00000 0.0 180	0.39287 0.0001 180	0.08081 0.2809 180	0.01685 0.8224 180	0.08542 0.2542 180	
CRF	0.39287 0.0001 180	1.00000 0.0 180	0.12314 0.0996 180	0.07325 0.3285 180	0.12193 0.1030 180	
DISCO	0.08081 0.2809 180	0.12314 0.0996 180	1.00000 0.0 180	0.72868 0.0001 180	0.47213 0.0001 180	
ECHAUFF	0.01685 0.8224 180	0.07325 0.3285 180	0.72868 0.0001 180	1.00000 0.0 180	0.53337 0.0001 180	
POURRIT	0.08542 0.2542 180	0.12193 0.1030 180	0.47213 0.0001 180	0.53337 0.0001 180	1.00000 0.0 180	

NOEUDS	-0.08137 0.2775 180	-0.04554 0.5438 180	0.04167 0.5786 180	0.02693 0.7197 180	-0.16922 0.0231 180
FENTES	-0.08915 0.2340 180	0.04732 0.5282 180	0.05069 0.4992 180	0.03321 0.6581 180	0.10507 0.1604 180
MOISSISS	0.01613 0.8298 180	0.09581 0.2007 180	0.51545 0.0001 180	0.46931 0.0001 180	0.43003 0.0001 180
CEX	-0.07898 0.2919 180	-0.05116 0.4952 180	0.03196 0.6702 180	-0.05137 0.4934 180	-0.02895 0.6996 180
ENTREEC	-0.05591 0.4757 165	0.04099 0.6011 165	0.07382 0.3460 165	0.07321 0.3500 165	0.01551 0.8432 165
T	0.01646 0.8338 165	-0.10641 0.1737 165	0.00621 0.9369 165	-0.02085 0.7904 165	-0.04704 0.5485 165
BLESS	0.09182 0.2203 180	0.05623 0.4534 180	0.04591 0.5406 180	-0.02281 0.7612 180	-0.06764 0.3669 180
MEPLAT	-0.17165 0.0212 180	-0.04348 0.5622 180	0.07837 0.2957 180	-0.04306 0.5660 180	-0.00779 0.9173 180
	NOEUDS	FENTES	MOISSISS	CEX	ENTREEC
CRSH	-0.08137 0.2775 180	-0.08915 0.2340 180	0.01613 0.8298 180	-0.07898 0.2919 180	-0.05591 0.4757 165
CRF	-0.04554 0.5438 180	0.04732 0.5282 180	0.09581 0.2007 180	-0.05116 0.4952 180	0.04099 0.6011 165
DISCO	0.04167 0.5786 180	0.05069 0.4992 180	0.51545 0.0001 180	0.03196 0.6702 180	0.07382 0.3460 165
ECHAUFF	0.02693 0.7197 180	0.03321 0.6581 180	0.46931 0.0001 180	-0.05137 0.4934 180	0.07321 0.3500 165
POURRIT	-0.16922 0.0231 180	0.10507 0.1604 180	0.43003 0.0001 180	-0.02895 0.6996 180	0.01551 0.8432 165
NOEUDS	1.00000 0.0 180	-0.00153 0.9838 180	-0.08706 0.2452 180	0.05499 0.4634 180	0.21146 0.0064 165
FENTES	-0.00153 0.9838 180	1.00000 0.0 180	-0.11295 0.1311 180	0.01064 0.8873 180	0.05368 0.4934 165
MOISSISS	-0.08706 0.2452 180	-0.11295 0.1311 180	1.00000 0.0 180	-0.19270 0.0096 180	0.07051 0.3682 165
CEX	0.05499 0.4634 180	0.01064 0.8873 180	-0.19270 0.0096 180	1.00000 0.0 180	0.13568 0.0823 165
ENTREEC	0.21146 0.0064 165	0.05368 0.4934 165	0.07051 0.3682 165	0.13568 0.0823 165	1.00000 0.0 165
T	0.13824	0.06768	-0.04699	0.07021	0.01627

CRSH	1.00000 0.0 180	0.38056 0.0001 180	0.07249 0.2813 180	0.01561 0.8174 180	0.08161 0.2508 180
CRF	0.38056 0.0001 180	1.00000 0.0 180	0.11148 0.0992 180	0.06538 0.3360 180	0.11693 0.1016 180
DISCO	0.07249 0.2813 180	0.11148 0.0992 180	1.00000 0.0 180	0.67825 0.0001 180	0.43071 0.0001 180
ECHAUFF	0.01561 0.8174 180	0.06538 0.3360 180	0.67825 0.0001 180	1.00000 0.0 180	0.49148 0.0001 180
POURRIT	0.08161 0.2508 180	0.11693 0.1016 180	0.43071 0.0001 180	0.49148 0.0001 180	1.00000 0.0 180
NOEUDS	-0.07765 0.2774 180	-0.04331 0.5467 180	0.03805 0.5730 180	0.02500 0.7126 180	-0.16143 0.0236 180
FENTES	-0.08400 0.2291 180	0.04479 0.5234 180	0.04481 0.4970 180	0.02937 0.6577 180	0.09743 0.1619 180
MOISSISS	0.01513 0.8312 180	0.09176 0.1985 180	0.46758 0.0001 180	0.42628 0.0001 180	0.40829 0.0001 180
CEX	-0.07274 0.2924 180	-0.04747 0.4942 180	0.02508 0.7008 180	-0.04770 0.4670 180	-0.02774 0.6874 180
ENTREEC	-0.05445 0.4737 165	0.04013 0.5991 165	0.06788 0.3456 165	0.06755 0.3496 165	0.01521 0.8410 165
T	0.01551 0.8362 165	-0.10252 0.1733 165	0.00547 0.9383 165	-0.01913 0.7882 165	-0.04462 0.5499 165
BLESS	0.08814 0.2184 180	0.05403 0.4528 180	0.04197 0.5349 180	-0.02063 0.7615 180	-0.06433 0.3678 180
MEPLAT	-0.16005 0.0215 180	-0.04042 0.5635 180	0.06971 0.2891 180	-0.03881 0.5570 180	-0.00710 0.9185 180
	NOEUDS	FENTES	MOISSISS	CEX	ENTREEC
CRSH	-0.07765 0.2774 180	-0.08400 0.2291 180	0.01513 0.8312 180	-0.07274 0.2924 180	-0.05445 0.4737 165
CRF	-0.04331 0.5467 180	0.04479 0.5234 180	0.09176 0.1985 180	-0.04747 0.4942 180	0.04013 0.5991 165
DISCO	0.03805 0.5730 180	0.04481 0.4970 180	0.46758 0.0001 180	0.02508 0.7008 180	0.06788 0.3456 165
ECHAUFF	0.02500 0.7126 180	0.02937 0.6577 180	0.42628 0.0001 180	-0.04770 0.4670 180	0.06755 0.3496 165
POURRIT	-0.16143 0.0236 180	0.09743 0.1619 180	0.40829 0.0001 180	-0.02774 0.6874 180	0.01521 0.8410 165
NOEUDS	1.00000	-0.00188	-0.08344	0.05100	0.20656

	0.0 180	0.9785 180	0.2416 180	0.4620 180	0.0068 165
FENTES	-0.00188 0.9785 180	1.00000 0.0 180	-0.10586 0.1283 180	0.00992 0.8836 180	0.05112 0.4916 165
MOISSISS	-0.08344 0.2416 180	-0.10586 0.1283 180	1.00000 0.0 180	-0.17731 0.0100 180	0.06871 0.3674 165
CEX	0.05100 0.4620 180	0.00992 0.8836 180	-0.17731 0.0100 180	1.00000 0.0 180	0.12821 0.0818 165
ENTREEC	0.20656 0.0068 165	0.05112 0.4916 165	0.06871 0.3674 165	0.12821 0.0818 165	1.00000 0.0 165
T	0.13344 0.0760 165	0.06318 0.3881 165	-0.04491 0.5485 165	0.06573 0.3667 165	0.01598 0.8421 150
BLESS	0.12403 0.0844 180	-0.01600 0.8197 180	0.03182 0.6557 180	-0.10180 0.1427 180	0.09652 0.2068 165
MEPLAT	0.00819 0.9067 180	0.06125 0.3695 180	-0.08109 0.2425 180	0.56650 0.0001 180	0.03258 0.6607 165
	T	BLESS	MEPLAT		
CRSH	0.01551 0.8362 165	0.08814 0.2184 180	-0.16005 0.0215 180		
CRF	-0.10252 0.1733 165	0.05403 0.4528 180	-0.04042 0.5635 180		
DISCO	0.00547 0.9383 165	0.04197 0.5349 180	0.06971 0.2891 180		
ECHAUFF	-0.01913 0.7882 165	-0.02063 0.7615 180	-0.03881 0.5570 180		
POURRIT	-0.04462 0.5499 165	-0.06433 0.3678 180	-0.00710 0.9185 180		
NOEUDS	0.13344 0.0760 165	0.12403 0.0844 180	0.00819 0.9067 180		
FENTES	0.06318 0.3881 165	-0.01600 0.8197 180	0.06125 0.3695 180		
MOISSISS	-0.04491 0.5485 165	0.03182 0.6557 180	-0.08109 0.2425 180		
CEX	0.06573 0.3667 165	-0.10180 0.1427 180	0.56650 0.0001 180		
ENTREEC	0.01598 0.8421 150	0.09652 0.2068 165	0.03258 0.6607 165		
T	1.00000 0.0 165	-0.02000 0.7906 165	0.18121 0.0134 165		

BLESS	-0.02000	1.00000	-0.02443
	0.7906	0.0	0.7270
	165	180	180
MEPLAT	0.18121	-0.02443	1.00000
	0.0134	0.7270	0.0
	165	180	180

C12	0.05	0.06	0.50	0.60	1.00															
C13	-0.08	-0.02	0.08	0.03	-0.16	1.00														
C14	-0.13	-0.01	0.02	0.02	0.06	-0.06	1.00													
C15	0.04	0.10	0.57	0.55	0.50	-0.07	-0.15	1.00												
C16	-0.09	-0.03	-0.02	-0.12	-0.10	0.06	0.01	-0.15	1.00											
C17	-0.07	0.07	0.10	0.11	0.03	0.16	0.03	0.04	0.13	1.00										
C18	-0.04	-0.09	0.00	-0.04	-0.03	0.10	-0.02	-0.06	0.08	0.05	1.00									
C19	0.04	0.01	0.03	-0.04	-0.05	0.10	-0.05	-0.01	-0.14	0.05	0.08	1.00								
C20	-0.17	-0.07	0.06	-0.05	0.01	0.01	0.04	-0.10	0.65	-0.01	0.12	-0.07	1.00							

MATRICE DES VALEURS-TESTS													
	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
C8	99.99												
C9	6.50	99.99											
C10	0.56	1.36	99.99										
C11	-0.07	0.95	14.46	99.99									
C12	0.70	0.84	7.34	9.26	99.99								
C13	-1.06	-0.32	1.02	0.39	-2.23	99.99							
C14	-1.78	-0.13	0.32	0.24	0.86	-0.77	99.99						
C15	0.47	1.31	8.68	8.32	7.39	-0.91	-1.97	99.99					
C16	-1.19	-0.34	-0.26	-1.65	-1.39	0.79	0.19	-2.04	99.99				
C17	-0.92	0.90	1.32	1.38	0.32	2.10	0.42	0.50	1.69	99.99			
C18	-0.50	-1.17	0.06	-0.47	-0.44	1.27	-0.21	-0.83	1.08	0.69	99.99		
C19	0.51	0.11	0.46	-0.59	-0.67	1.31	-0.72	-0.18	-1.89	0.64	0.99	99.99	
C20	-2.29	-0.89	0.86	-0.70	0.14	0.08	0.60	-1.35	10.34	-0.11	1.60	-0.94	99.99

VALEURS PROPRES
 APERCU DE LA PRECISION DES CALCULS : TRACE AVANT DIAGONALISATION .. 13.0000
 SOMME DES VALEURS PROPRES 13.0000

HISTOGRAMME DES 13 PREMIERES VALEURS PROPRES			
NUMERO	VALEUR PROPRE	POURCENT.	POURCENT. CUMULE
1	2.8351	21.81	21.81
2	1.8206	14.00	35.81
3	1.3762	10.59	46.40
4	1.3155	10.12	56.52
5	1.0661	8.20	64.72
6	0.9737	7.49	72.21
7	0.8672	6.67	78.88
8	0.8248	6.34	85.22
9	0.5333	4.10	89.33
10	0.4700	3.62	92.94
11	0.4344	3.34	96.28
12	0.3096	2.38	98.67
13	0.1735	1.33	100.00

RECHERCHE DE PALIERS (DIFFERENCES TROISIEMES)			
PALIER ENTRE	VALEUR DU PALIER		
2-- 3	-572.19		
8-- 9	-313.08		
1-- 2	-186.57		
4-- 5	-170.89		

RECHERCHE DE PALIERS ENTRE (DIFFERENCES SECONDES)			
PALIER ENTRE	VALEUR DU PALIER		
1-- 2	570.18		
2-- 3	383.61		
8-- 9	228.17		
4-- 5	156.87		
6-- 7	64.05		
9-- 10	27.66		

INTERVALLES LAPLACIENS D'ANDERSON			
INTERVALLES AU SEUIL 0.95			
NUMERO	BORNE INFERIEURE	VALEUR PROPRE	BORNE SUPERIEURE
1	2.2477	2.8351	3.4225
2	1.4434	1.8206	2.1977
3	1.0911	1.3762	1.6613
4	1.0429	1.3155	1.5880
5	0.8453	1.0661	1.2870

ETENDUE ET POSITION RELATIVE DES INTERVALLES												
1
2
3
4
5

COORDONNEES DES VARIABLES SUR LES AXES 1 A 5																
VARIABLES ACTIVES																
IDEN - LIBELLE COURT	VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C8 - CRSH		0.13	0.48	0.67	0.12	0.06	0.13	0.48	0.67	0.12	0.06	0.08	0.36	0.58	0.10	0.06
C9 - CRF		0.20	0.32	0.73	0.17	-0.24	0.20	0.32	0.73	0.17	-0.24	0.12	0.24	0.62	0.15	-0.24
C10 - Disco		0.84	-0.24	0.01	0.12	0.00	0.84	-0.24	0.01	0.12	0.00	0.50	-0.18	0.01	0.11	0.00
C11 - Echauff		0.88	-0.16	-0.09	0.03	-0.04	0.88	-0.16	-0.09	0.03	-0.04	0.52	-0.12	-0.08	0.03	-0.04
C12 - Pourrit		0.76	-0.11	0.00	-0.22	0.01	0.76	-0.11	0.00	-0.22	0.01	0.45	-0.08	0.00	-0.19	0.01
C13 - Noeuds		-0.06	-0.17	-0.14	0.69	-0.10	-0.06	-0.17	-0.14	0.69	-0.10	-0.04	-0.12	-0.12	0.60	-0.10
C14 - Fentes		-0.03	-0.16	-0.14	-0.27	-0.76	-0.03	-0.16	-0.14	-0.27	-0.76	-0.02	-0.12	-0.12	-0.24	-0.73
C15 - Moisiss		0.78	0.00	-0.04	-0.02	0.20	0.78	0.00	-0.04	-0.02	0.20	0.47	0.00	-0.03	-0.02	0.19
C16 - Cex		-0.23	-0.75	0.44	-0.01	0.04	-0.23	-0.75	0.44	-0.01	0.04	-0.14	-0.56	0.38	-0.01	0.04
C17 - EntreEc		0.10	-0.23	0.02	0.51	-0.50	0.10	-0.23	0.02	0.51	-0.50	0.06	-0.17	0.02	0.44	-0.48
C18 - T		-0.09	-0.26	-0.08	0.36	0.28	-0.09	-0.26	-0.08	0.36	0.28	-0.05	-0.19	-0.07	0.32	0.27
C19 - Bless		-0.01	0.18	-0.16	0.51	0.14	-0.01	0.18	-0.16	0.51	0.14	0.00	0.13	-0.14	0.45	0.14
C20 - Meplat		-0.14	-0.79	0.34	-0.10	0.16	-0.14	-0.79	0.34	-0.10	0.16	-0.08	-0.59	0.29	-0.09	0.16

VARIABLES ILLUSTRATIVES																
IDEN - LIBELLE COURT	VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C2 - Tranch		-0.29	0.14	-0.18	-0.13	-0.05	-0.29	0.14	-0.18	-0.13	-0.05					
C3 - Deroul		-0.47	0.29	-0.20	-0.15	0.06	-0.47	0.29	-0.20	-0.15	0.06					
C4 - Ebenist		-0.37	0.17	-0.06	-0.21	0.06	-0.37	0.17	-0.06	-0.21	0.06					
C5 - Menuis		-0.53	0.12	-0.10	-0.22	0.03	-0.53	0.12	-0.10	-0.22	0.03					
C6 - Carca		-0.36	0.07	0.03	-0.04	0.05	-0.36	0.07	0.03	-0.04	0.05					

ANNEXE 13.

LE MODELE LOGIT

Exemples de sortie obtenue par application de la procédure LOGISTIC de SAS® :

Annexe 13.1- régression logistique de la variable « acceptation/refus » pour l'utilisation « déroulage » sur les variables de présence et d'intensité des défauts. Il s'agit du modèle obtenu après sélection des variables explicatives

Tableau 13.1- Régression logistique pour l'utilisation en déroulage après sélection des variables

```

The LOGISTIC Procedure
Model Information
Data Set                WORK.DONNEES
Response Variable       Deroul
Number of Response Levels 2
Number of Observations 165
Link Function           Logit
Optimization Technique  Fisher's scoring

Response Profile
Ordered Value   Deroul   Total
                Frequency
1               0       87
2               1       78

explanatory variables.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics
Criterion          Intercept Only   Intercept and Covariates
AIC                230.247         165.707
SC                 233.353         178.130
-2 Log L           228.247         157.707

Testing Global Null Hypothesis: BETA=0
Test              Chi-Square   DF   Pr > ChiSq
Likelihood Ratio  70.5408     3    <.0001
Score             58.7761     3    <.0001
Wald              41.4752     3    <.0001

Analysis of Maximum Likelihood Estimates
Standard
Parameter  DF   Estimate   Error   Chi-Square   Pr > ChiSq
Intercept  1   -5.3539    0.9434    32.2036     <.0001
Disco      1    1.3961    0.2332    35.8568     <.0001
Cex        1    0.7263    0.2324     9.7699     0.0018
T          1    1.2441    0.4605     7.2994     0.0069

Odds Ratio Estimates
Point          95% Wald

```

Effect	Estimate	Confidence Limits	
Disco	4.040	2.558	6.380
Cex	2.067	1.311	3.260

Odds Ratio Estimates			
	Point	95% Wald	
Effect	Estimate	Confidence Limits	
T	3.470	1.407	8.557

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	82.1	Somers' D	0.693
Percent Discordant	12.8	Gamma	0.730
Percent Tied	5.1	Tau-a	0.347
Pairs	6786	c	0.846

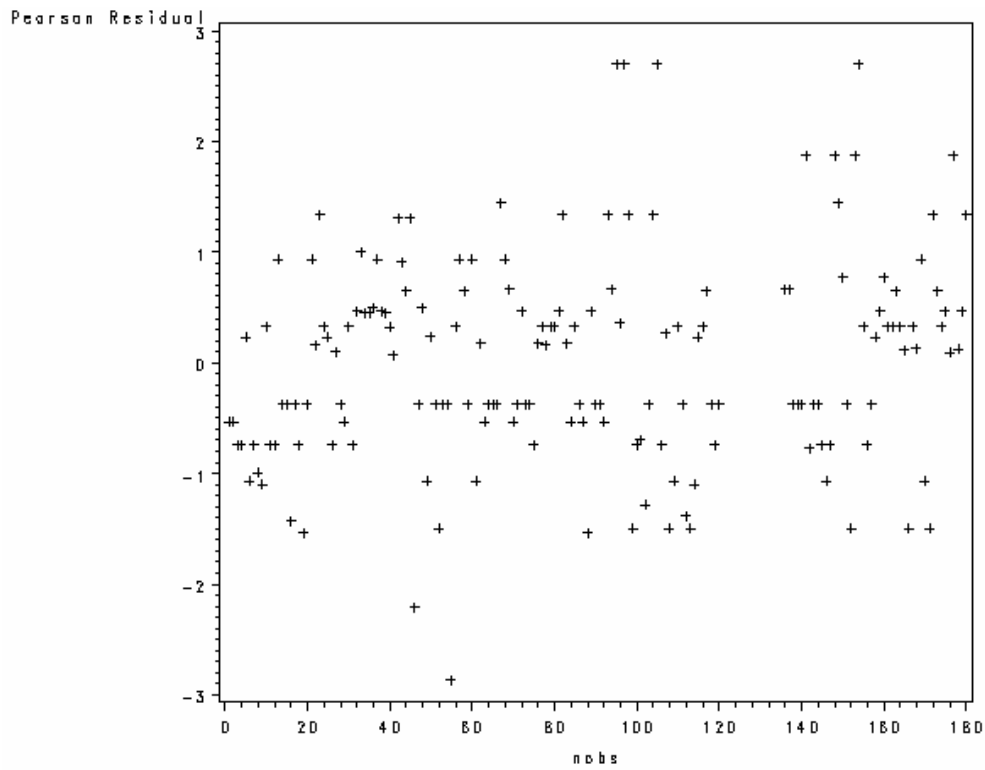


Figure 13.1 : Résidus de Pearson pour les différentes observations

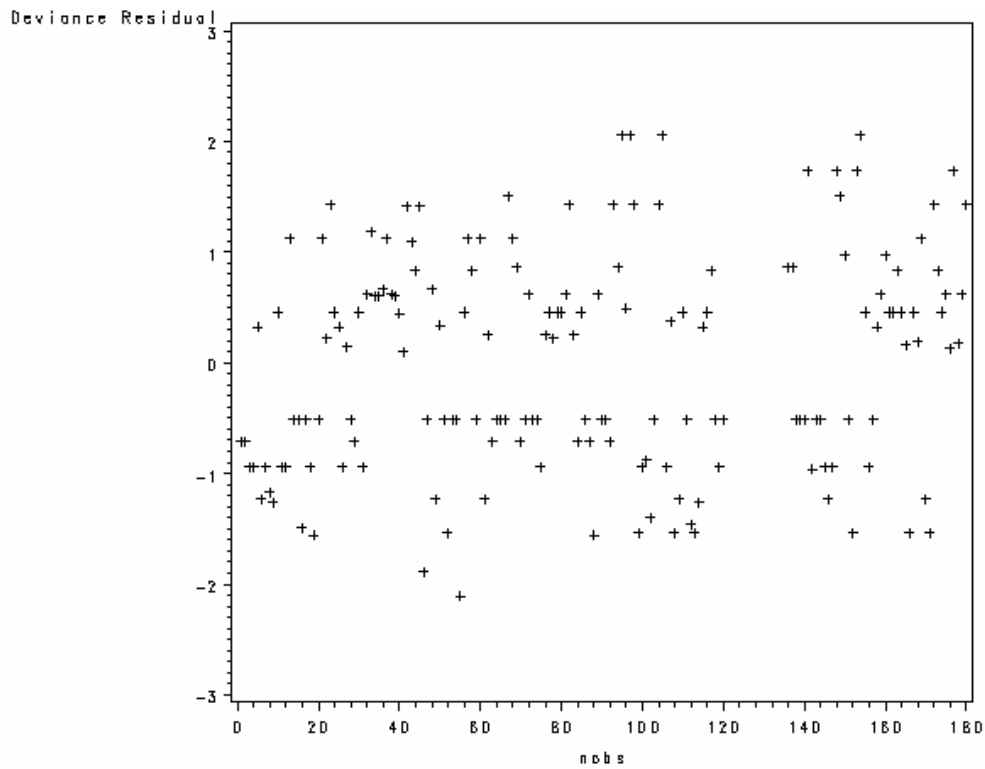


Figure 13.2 : Résidus de déviance pour les différentes observations

Les Figures 13.1 et 13.2 représentent les valeurs des résidus pour les différentes observations des photos par les juges.

Sur le graphe des résidus en Figures 13.1 et 13.2, nous pouvons remarquer que les observations n^{os} 55, 95, 97, 105 et 154 apparaissent assez mal ajustées par le modèle.

Annexe 13.2- résidus de la régression logistique de la variable « acceptation/refus » en déroulage sur l'ensemble des variables de présence et d'intensité des défauts

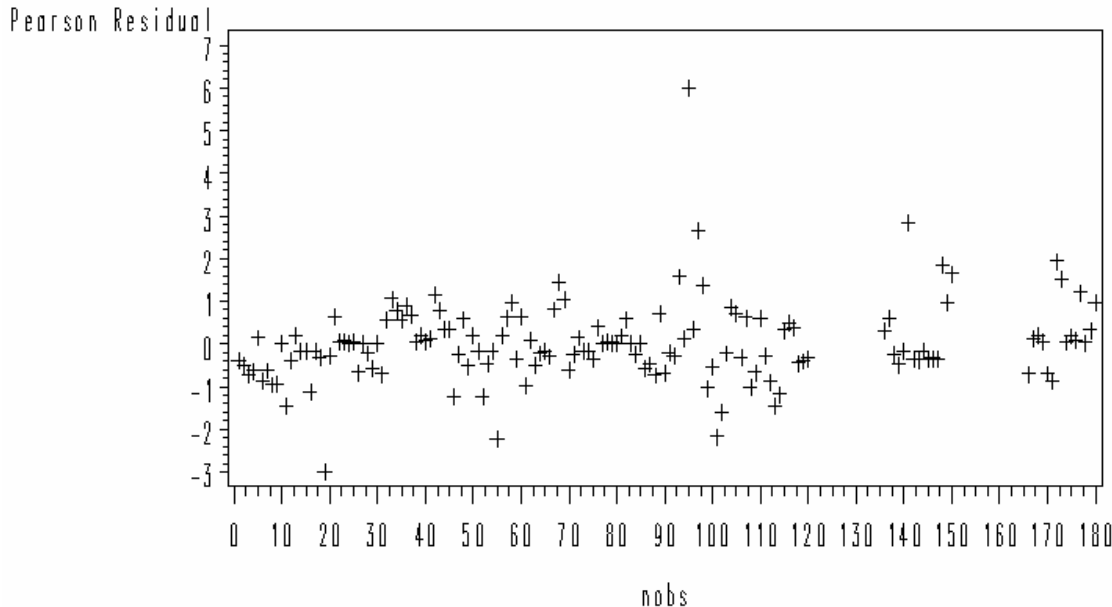


Figure 13.3 : Résidus de Pearson pour les différentes observations

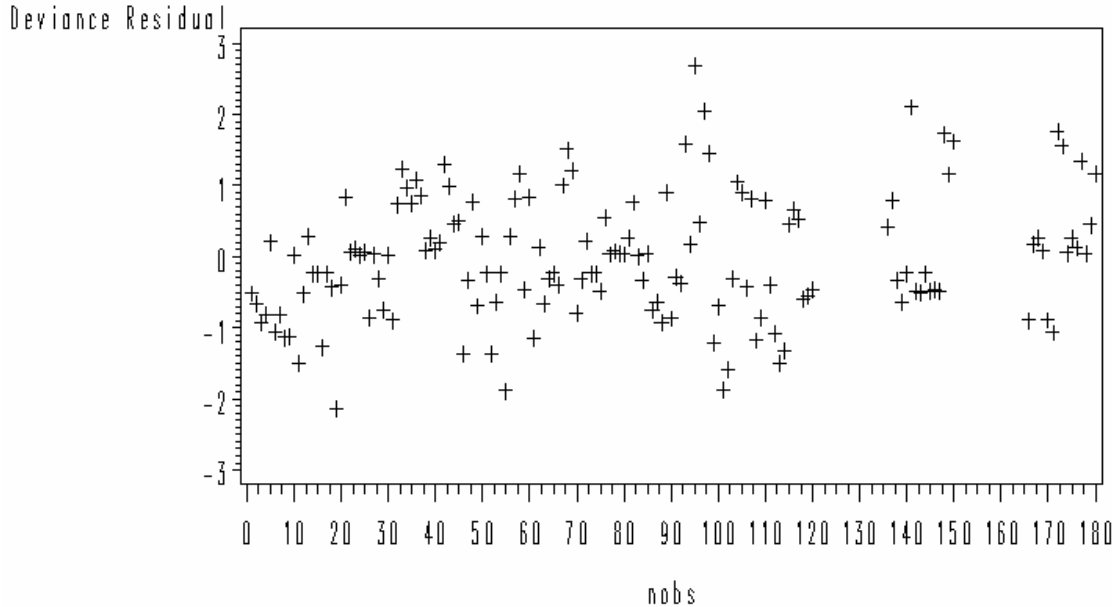


Figure 13.4 : Résidus de déviance pour les différentes observations

ANNEXE 14.

REGRESSION LOGISTIQUE SUR COMPOSANTES PLS

Annexe 14.1- Résultats de la validation croisée pour le choix du nombre de composantes dans la régression PLS des variables indicatrices des utilisations sur les variables de présence et d'intensité des défauts

La validation croisée nous permet de déterminer le nombre de composantes à extraire dans la régression PLS des utilisations sur les variables défauts. L'option CV de la procédure PLS du logiciel SAS est utilisée pour effectuer la validation croisée.

1/ Option CV = one

Cette option sélectionne une validation croisée effectuée observation par observation. Chaque observation constitue un échantillon test.

Programme SAS :

```
Proc PLS data = nom_table_données CV=one ;  
Model = CRSH CRF disco echauff pourrit moisiss cex meplat T  
entrec fentes nœuds bless ;  
Run ;
```

Tableau 14.1 : Validation croisée observation par observation

The PLS Procedure	
Cross Validation for the Number of Extracted Factors	
Number of Extracted Factors	Root Mean PRESS
0	1.006711
1	0.895649
2	0.884226
3	0.881108

4	0.885179
5	0.885437
6	0.887108
7	0.891907
8	0.902008
9	0.903324
10	0.901802
11	0.90271
12	0.903791
13	0.903857

La valeur absolue minimale du PRESS (« predicted residual sum of square ») (0,881108) est atteinte avec 3 composantes. Nous remarquerons cependant que la valeur du PRESS avec seulement 2 composantes n'est pas beaucoup plus élevée (0,884226).

2/ Option CV = block(n)

Les échantillons-test sont ici tous les groupes de n observations.

Tableau 14.2 : Validation croisée par blocs de n=7 observations

The PLS Procedure		
Blocked Cross Validation for the Number of Extracted Factors		
Number of Extracted Factors	Root Mean PRESS	
0	1.091053	
1	0.995121	
2	0.987711	
3	0.974503	
4	0.985222	
5	0.98635	
6	0.991933	
7	0.996404	
8	1.00033	
9	1.00301	
10	1.002694	
11	1.004155	
12	1.003812	
13	1.004584	
Minimum root mean PRESS		0.9745
Minimizing number of factors		3

Ici encore, on conclut à l'extraction de 3 composantes.

SAS nous donne également la table contenant le pourcentage de variance expliquée par chaque composante.

Tableau 14.3 : Pourcentage de variance expliquée par le modèle à 3 composantes

Number of Extracted Factors	Percent Variation Accounted for by Partial Least Squares Factors			
	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	20.9087	20.9087	23.7307	23.7307
2	11.2045	32.1132	4.7631	28.4938
3	9.5272	41.6403	2.9621	31.4559

La statistique proposée par Van Der Voet nous permet de comparer les différents modèles :

Tableau 14.4 : Test de comparaison des modèles PLS

Blocked Cross Validation for the Number of Extracted Factors				
Number of Extracted Factors	Root Mean PRESS	T**2	Prob > T**2	
0	1.091053	34.85387	<.0001	
1	0.995121	4.883279	0.4420	
2	0.987711	3.609083	0.6080	
3	0.974503	0	1.0000	
4	0.985222	14.00353	0.0100	
5	0.98635	12.62229	0.0170	
6	0.991933	17.94789	0.0010	
7	0.996404	15.40888	0.0060	
8	1.00033	15.55831	0.0050	
9	1.00301	17.26657	0.0030	
10	1.002694	17.29622	0.0030	
11	1.004155	17.2784	<.0001	
12	1.003812	17.65713	<.0001	
13	1.004584	17.86956	<.0001	
Minimum root mean PRESS			0.9745	
Minimizing number of factors			3	
Smallest number of factors with p > 0.1			1	
Percent Variation Accounted for by Partial Least Squares Factors				
Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	20.9087	20.9087	23.7307	23.7307

Avec ce test, c'est le modèle à une seule composante qui est retenu.

Cependant nous perdons beaucoup d'information en ôtant les composantes 2 et 3 ; nous passons de 41,6403% de variation expliquée à 20,9087% pour l'ensemble des variables explicatives (Tableau 14.3). C'est pourquoi nous avons décidé de les garder quand même dans le modèle.

Annexe 14.2- Coefficients des modèles obtenus par la régression logistique sur composantes PLS

En raison de difficultés techniques intervenues au moment du tirage du mémoire, le tableau annoncé n'a pu être reproduit.

ANNEXE 15

Pourcentage d'acceptation prédit pour différentes utilisations du bois de Hêtre (tranchage, déroulage, ébénisterie, menuiserie intérieure, carcasse de meuble) en fonction de la note d'altération attribuée par les professionnels à la section de grume

Nous avons travaillé sur 60 photos numériques de rondelles de Hêtre. Les photos ont été soumises au jugement de professionnels du bois de Hêtre (essentiellement des industriels de la première transformation). Chaque photo a été vue 3 fois (par 3 personnes différentes). Donc nous disposons de $60 \times 3 = 180$ observations de rondelles.

1- Pourcentages d'acceptation

Nous avons effectué la **régression PLS** de l'ensemble des 5 variables binaires « acceptation/refus des utilisations » (tranchage, déroulage, ébénisterie, menuiserie intérieure, carcasse de meuble) sur l'ensemble des 13 variables « présence et intensité des défauts ». On extrait 3 composantes PLS (validation croisée) : t1, t2 et t3.

Ensuite, nous avons réalisé une **régression logistique** de chacune des variables « utilisations » sur les composantes t1, t2, t3 issues de la régression PLS précédente. Parmi t1, t2 et t3 nous avons retenu dans le modèle uniquement les composantes influentes ($P_value < 0,05$) pour chacune des utilisations.

```
/* extraction des composantes PLS */
Proc PLS data=donnees nfac=3 ;
Model tranch deroul menuis ebenist carca = CRSH CRF disco echauff
pourrit moisiss T cex meplat bless entreec nœuds fentes ;
Output out=sortie Xscore=t ;
Run ;

/* regression logistique sur composantes PLS pour l'utilisation en
tranchage */
Proc logistic data=sortie ;
Model tranch = t1 t2 t3 ;
Output out=sortie2 predprobs=individual predicted=pred ;
Run;
```

Le tableau 15.1 donne pour chaque utilisation les composantes PLS qui ont été conservées dans le modèle car influentes.

Tableau 205.1: Composantes PLS conservées dans le modèle de régression logistique

Utilisations	Composantes
Tranchage	t1
Déroutage	t1, t3
Ebénisterie	t1, t2
Menuiserie intérieure	t1, t2
Carcasse de meuble	t1, t3

Puis, pour chaque observation (3 x 60 photos), nous nous sommes intéressée à la **probabilité** d'acceptation **prédite** par le modèle. Attention : le modèle obtenu par le logiciel SAS fournit une estimation du **logit** de p où p est la probabilité de refus. La probabilité d'acceptation sera donc de 1 - probabilité prédite. C'est cette probabilité que nous allons représenter en fonction des notes d'altération pour les 180 observations.

2- Notes d'altération

Il s'agit d'une note reflétant l'état d'altération de la rondelle. Elle est basée sur la présence et l'intensité de discoloration, d'échauffure et de pourriture, les intensités variant de 0 (pas du tout) à 3 (beaucoup) :

$$\text{note d'altération} = 1 \times \text{discoloration} + 1,5 \times \text{échauffure} + 2 \times \text{pourriture}^{(1)}$$

Plus la note est élevée, plus la rondelle est altérée. La note d'altération est ainsi comprise entre 0 et 13,5.

Quelques exemples de correspondance entre photos et notes d'altération moyenne (calculée sur les 3 observations d'une même photo par des juges différents) :

Ces pondérations (1) permettent de tenir compte de la gravité du défaut considéré (par exemple l'échauffure est plus grave que la discoloration) mais les coefficients sont choisis arbitrairement. Nous aurions peut-être pu prendre, à la place, les coefficients de ces 3 variables sur le premier axe de l'ACP des variables « défauts » (le premier axe étant représentatif du niveau d'altération de la rondelle).

Au départ, nous avons attribué à chacune des 60 photos une note d'altération moyenne sur les 3 observations de la photo par des juges différents. Ici, nous avons cherché à relier le pourcentage d'acceptation (propre à chacune des 180 observations) à une note d'altération. Pour chacune des 180 observations, nous calculerons sa note d'altération propre à partir de la formule précédente.

Nous avons représenté l'histogramme (Figure 15.1) des notes d'altération ainsi obtenues pour les 180 observations.

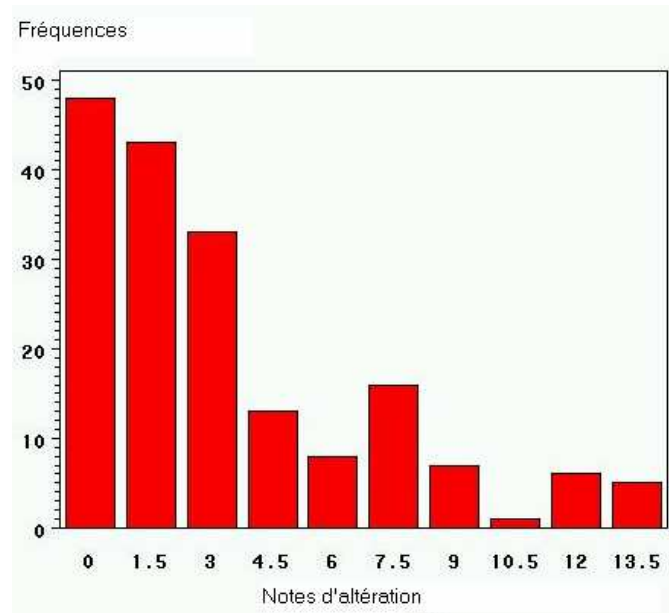


Figure 15.1: Histogramme des notes d'altération

3- Représentations graphiques

3.1- Représentation des probabilités d'acceptation prédites en fonction des notes d'altérations pour les 180 (60*3) observations de rondelles issues des interviews, tous les défauts (y compris ceux n'entrant pas dans le calcul de la note d'altération) intervenant dans les prédictions avec leurs intensités respectives.

Dans un premier temps, nous avons représenté les probabilités d'acceptation obtenues par les modèles, pour chacune des utilisations et pour les 180 observations, en fonction des notes d'altération pour chacune d'elles (Figures 15.2 à 15.6).

```
/* création de la variable « probabilité d'acceptation prédite » */  
Data sortie3 ;  
Set sortie2 ;  
Accep=1-pred ;  
Run ;  
  
/* proc gplot data=sortie3 */  
Plot accep*alter ;  
Run ;
```

Probabilités d'acceptation prédites par le modèle

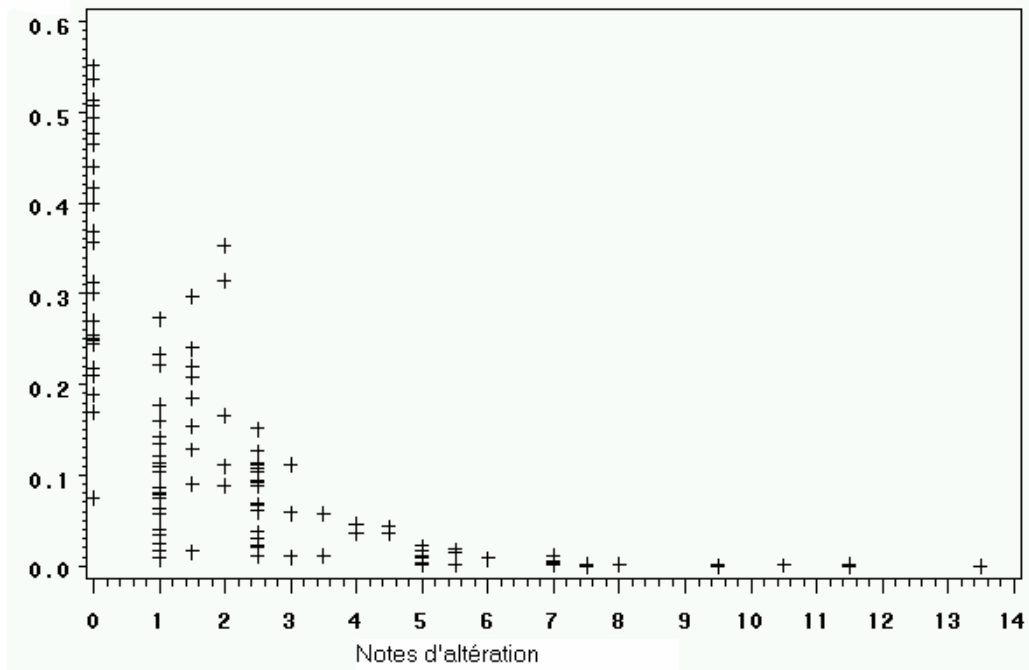


Figure 15.2 : Graphique obtenu pour l'utilisation en tranchage

Probabilités d'acceptation prédites par le modèle

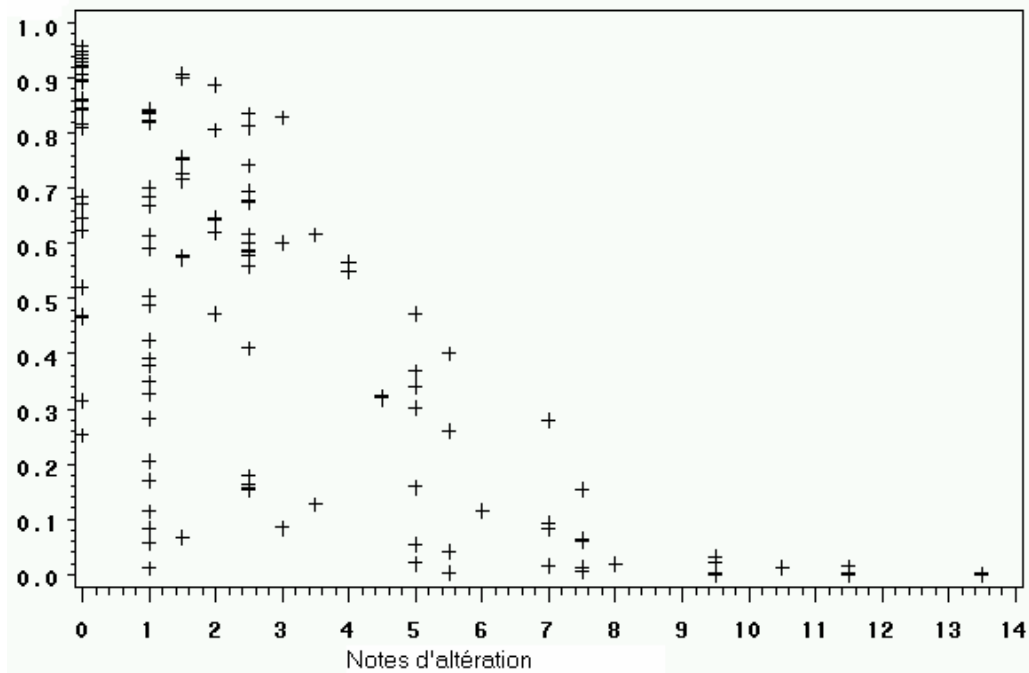


Figure 15.3 : Graphique obtenu pour l'utilisation en déroulage

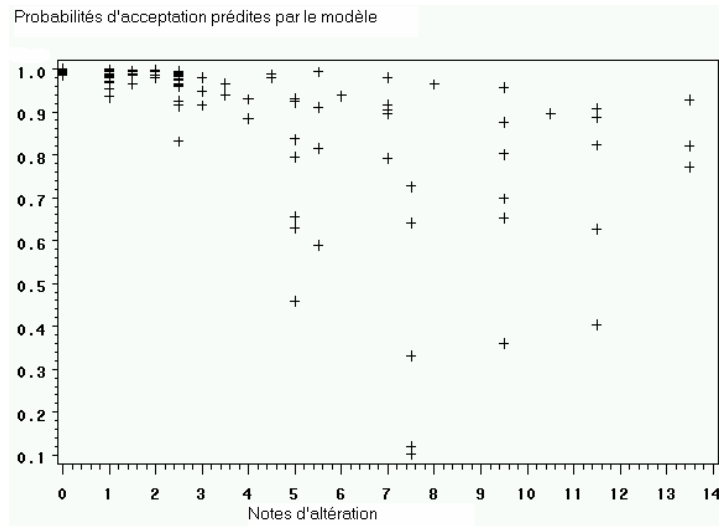


Figure 15.6 : Graphique obtenu pour l'utilisation en carcasse de meuble

Les graphes obtenus mettent en évidence une diminution de la probabilité d'acceptation quand la note d'altération augmente (tendance décroissante observée des nuages de points).

Cependant, même lorsque la note d'altération est faible, nous trouvons des probabilités d'acceptation basses pour un grand nombre de rondelles. Cela provient des autres défauts (non pris en compte dans la note d'altération) présents sur la rondelle et auxquels peuvent être imputés les refus.

Pour les utilisations en tranchage et en déroulage (Figures 15.2 et 15.3) nous remarquons que, pour des notes d'altération élevées, les probabilités d'acceptation sont très faibles, ce qui montre bien que la présence d'altérations est rédhibitoire pour l'utilisation en tranchage et en déroulage. En revanche, pour les trois autres utilisations, nous trouvons encore des probabilités d'acceptation assez fortes pour des notes d'altération élevées.

Nous remarquons aussi que, pour le tranchage, les probabilités les plus fortes d'acceptation sont inférieures à 0,6 (Figure 15.2) alors que, pour les autres utilisations, elles sont supérieures à 0,9. Cela provient du fait que le tranchage est l'utilisation la plus noble et que les industriels sont très "regardants" sur la qualité du bois pour cette utilisation.

Les rondelles contiennent des défauts autres que ceux intervenant dans la note d'altération (discoloration, échauffure et pourriture). Le choix de l'acceptation ou du

refus peut donc être imputé aux autres défauts. Ce qui nous intéresserait, c'est de représenter cette même probabilité d'acceptation en fonction des mêmes notes d'altération mais en considérant cette fois que les autres défauts ne sont pas présents sur les rondelles. Nous pouvons par exemple calculer les probabilités d'acceptation obtenues en considérant les autres défauts comme absents (c'est-à-dire mettre des valeurs d'intensité égale à 1 (pas du tout) pour ces défauts dans les modèles de prédiction précédents (les coefficients des modèles restent inchangés)). Cela fait l'objet du paragraphe 3.2.

3.2- Représentation des probabilités d'acceptation prédites en fonction des notes d'altérations pour les 180 (60*3) observations de rondelles issues des interviews, les défauts autres que ceux intervenant dans la note d'altération étant considérés comme absents.

Les modèles sont les mêmes que précédemment mais les intensités des défauts n'intervenant pas dans la note d'altération sont donc toutes remplacées par la valeur d'intensité 1 (absence du défaut dans le modèle). Nous avons ainsi obtenu les graphiques suivants (Figures 15.7 à 15.11) :

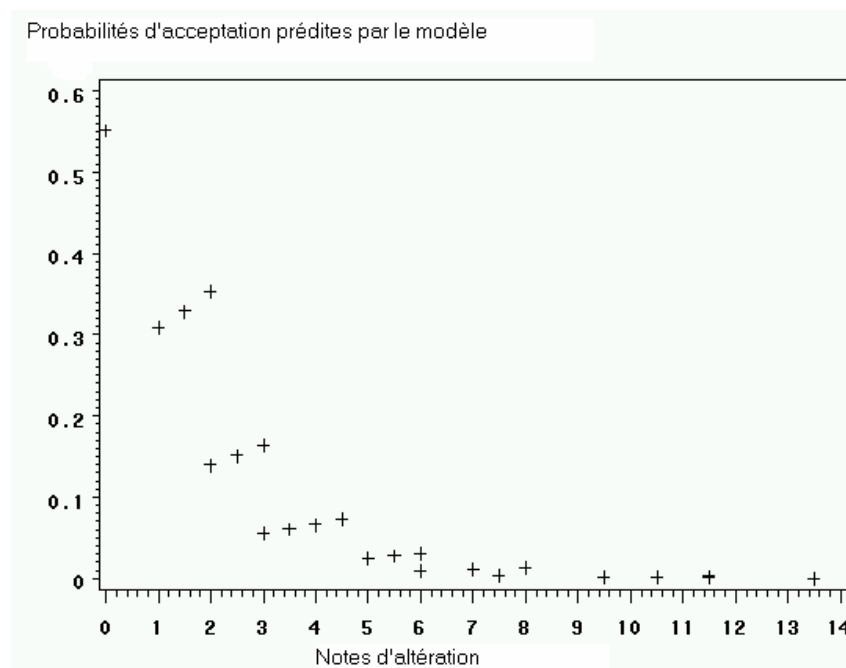


Figure 15.7 : Graphique obtenu pour l'utilisation en tranchage

Probabilités d'acceptation prédites par le modèle

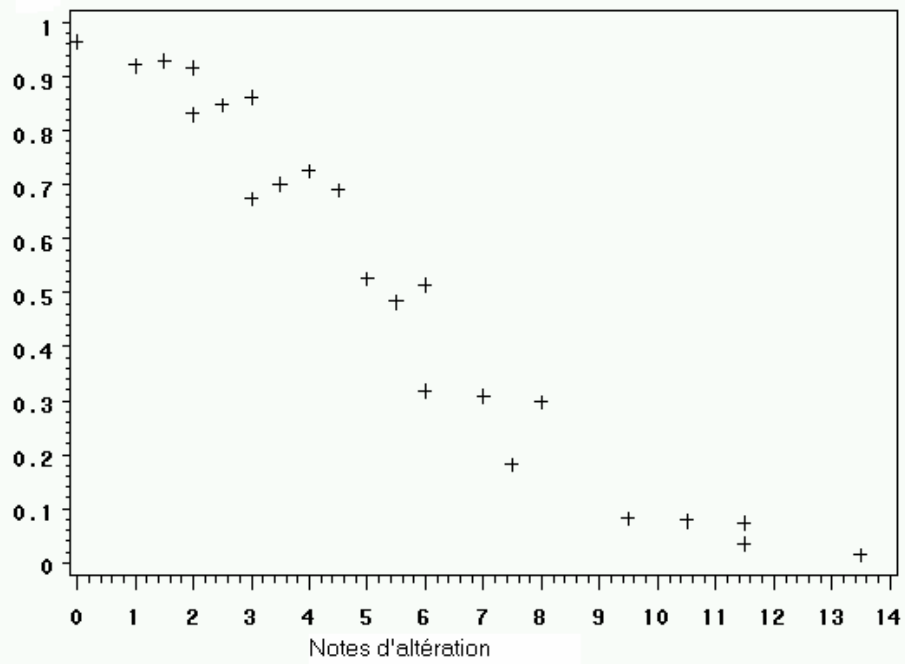


Figure 15.8 : Graphique obtenu pour l'utilisation en déroulage

Probabilités d'acceptation prédites par le modèle

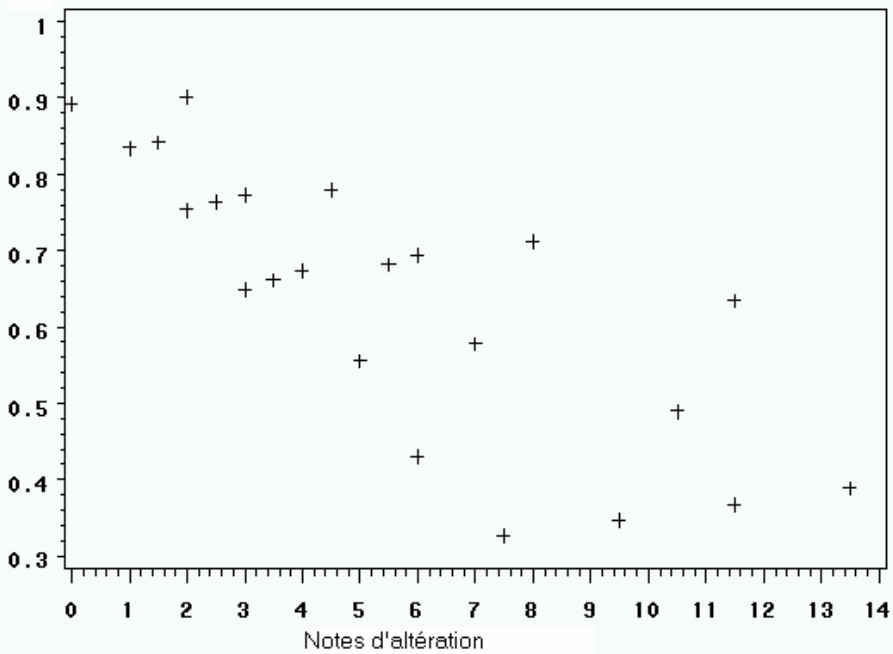


Figure 15.9 : Graphique obtenu pour l'utilisation en ébénisterie

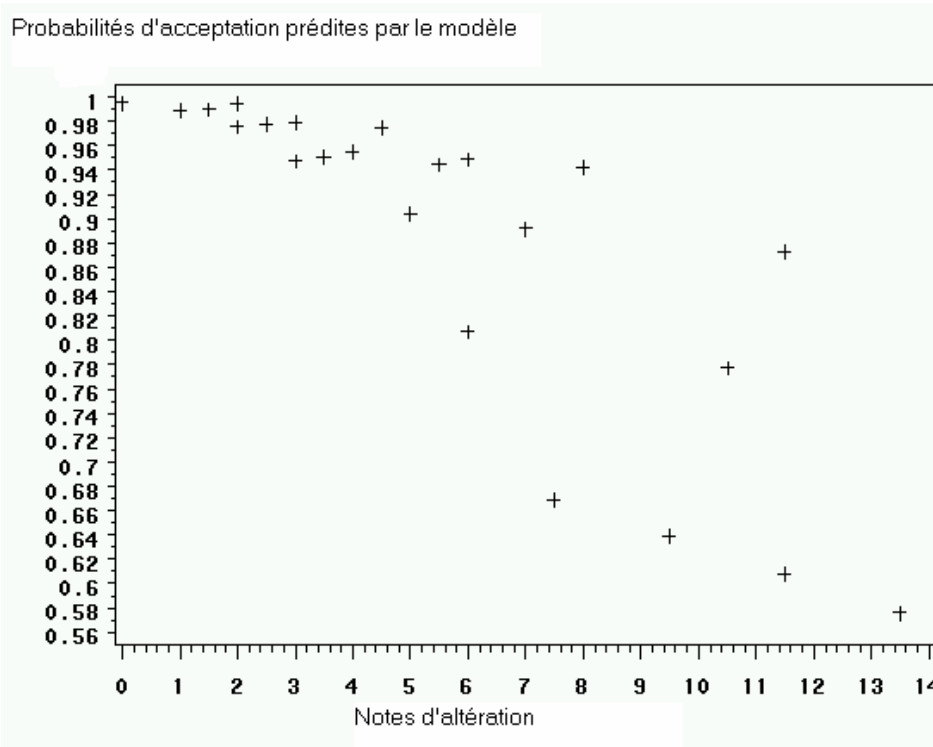


Figure 15.10 : Graphique obtenu pour l'utilisation en menuiserie intérieure

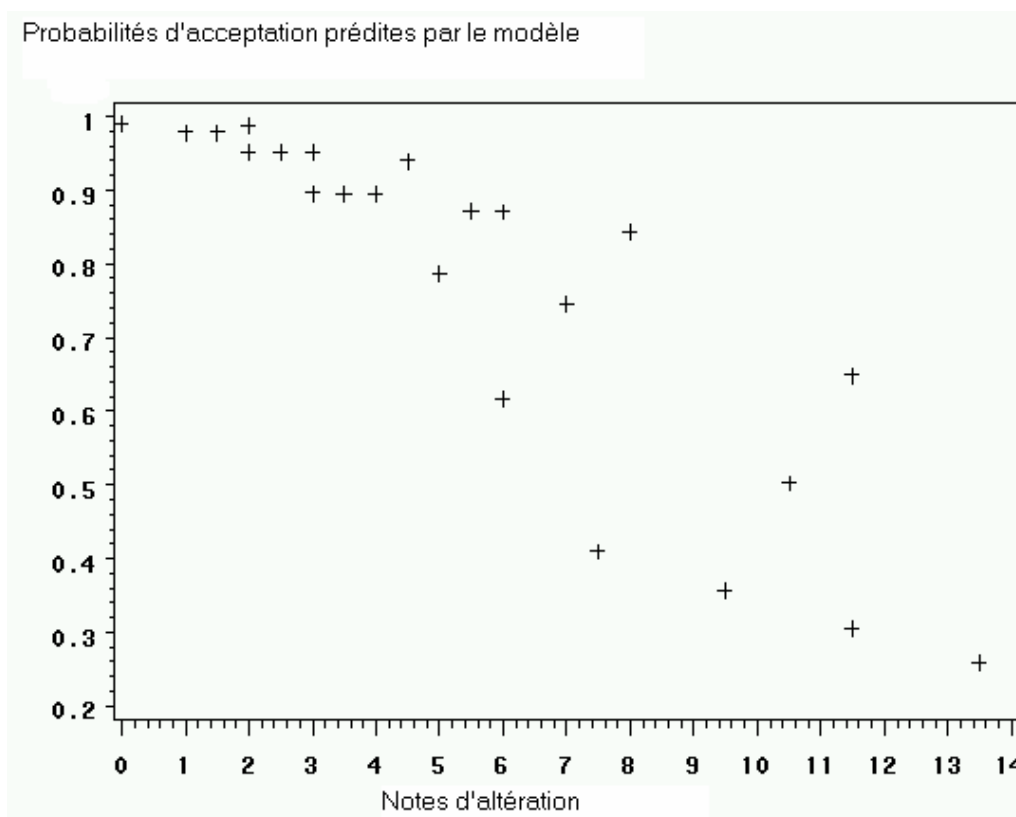


Figure 15.11 : Graphique obtenu pour l'utilisation en carcasse de meuble

Contrairement au § 3.1, pour des notes d'altération basses, nous ne trouvons plus de probabilités d'acceptation faibles, les autres défauts étant considérés comme absents ici.

Pour les utilisations en tranchage et en déroulage, nous constatons que la probabilité prédite d'acceptation est déterminée de façon assez précise par la note d'altération (Figures 15.7 et 15.8).

Comme dans le § 3.1, la probabilité prédite d'acceptation en tranchage ne passe pas au dessus de 0.6 dans le meilleur des cas et ce, contrairement aux autres utilisations où elle est proche de 1.

Nous remarquons aussi que les graphiques (Figures 15.7 à 15.11) comportent moins de points que ceux du § 3.1 : cela provient du fait que beaucoup de points sont ici confondus. En effet, dans le § 3.1 les valeurs des intensités des défauts (y compris ceux n'entrant pas dans la note d'altération) permettaient une meilleure différenciation des 180 observations.

Enfin, nous observons une plus grande variabilité des probabilités prédites pour les notes d'altération élevées que pour les notes d'altération faibles. Cela s'explique assez simplement : le modèle de prédiction ne prend pas en compte la note d'altération globale mais indépendamment chacune des trois variables (discoloration, échauffure, pourriture) la constituant. Or pour une note d'altération donnée, il peut, par exemple, y avoir plusieurs combinaisons possibles de ces trois variables. Par ailleurs, plus les notes d'altération sont élevées, plus les combinaisons des valeurs de ces trois variables peuvent être variées. Or pour chaque combinaison différente, nous obtenons une prédiction différente. Ceci explique que pour des notes d'altération élevées la variabilité des prédictions est plus importante.

4- Conclusion

Nous avons représenté l'influence des altérations sur les prédictions pour l'acceptation ou le refus du bois dans les différentes utilisations. La présence d'altérations est déterminante pour le choix de l'utilisation en tranchage et en déroulage. C'est encore vrai pour les utilisations moins nobles mais dans une moindre mesure.

Les coefficients utilisés pour le calcul de la note d'altération ont été choisis arbitrairement. Ils permettaient de tenir compte de l'ordre de gravité des défauts mais, ce que nous ne savons pas, c'est avec quel coefficient un défaut est plus grave qu'un autre. Ce que nous aurions pu faire c'est de recommencer l'étude en changeant la valeur de ces coefficients tout en gardant le même ordre de gravité et de vérifier que nous obtenions des résultats similaires.

Annexe 16 :

**Photographies numériques de sections de grumes
présentant des notes d'altération contrastées
(notes 0 à 1, 2 à 3, 4 à 6, 7 à 9, 10 à 12)**