

MARIA GABRIELA ORZANCO

**EXPLORATION DE LA FUSION DES  
INFORMATIONS POUR AMÉLIORER LA  
FIABILITÉ LOCALE D'UNE CARTE FORESTIÈRE**

Thèse présentée  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de doctorat en Sciences Géomatiques  
pour l'obtention du grade de Philosophiae Doctor (Ph. D.)

FACULTÉ DE FORESTERIE ET DE GÉOMATIQUE  
UNIVERSITÉ LAVAL  
QUÉBEC

MARS 2006

© María Gabriela Orzanco, 2006

## **Résumé court**

L'objectif de cette recherche est d'explorer comment la complémentarité de différentes méthodes peuvent permettre de mieux caractériser du point de vue écologique et géomatique les éléments d'une carte forestière tels que les frontières entre les peuplements, les strates forestières et les estimations du volume ligneux et ce, en fusionnant les informations de la carte et des échantillons de terrain. L'emploi combiné de ces méthodes vise à : (1) détecter les secteurs du territoire où l'existence des frontières serait plus certaine, (2) regrouper les échantillons de terrain de façon à obtenir des estimations plus précises du volume ligneux que les estimations obtenues par la méthode actuellement employée au Québec, et (3) déterminer la concordance des strates forestières avec l'information de terrain. La fusion des frontières de la carte et du terrain, établie à partir des connaissances, confirme la certitude sur l'existence de certaines frontières cartographiques. La fusion des informations sur la variabilité des estimations et sur la concordance carte/échantillons de terrain des strates forestières permet d'obtenir des estimations du volume ligneux plus proches de la réalité.

## Résumé

Sur la carte forestière traditionnelle du Québec, les frontières entre les peuplements sont représentées par des lignes, les estimations du volume ligneux proviennent de regroupements des échantillons de terrain (p-e) et sont assignées aux strates forestières, qui sont considérées comme homogènes pour les attributs cartographiés. L'objectif général de cette recherche est d'explorer comment la complémentarité de différentes méthodes peuvent permettre de mieux caractériser les frontières entre les peuplements, les strates forestières et les estimations du volume ligneux et ce, en fusionnant les informations de la carte et des p-e. Le contraste et le contexte spatial des frontières sont quantifiés, tout comme le contexte des peuplements. Ces aspects sont complémentaires et ils permettent de détecter les secteurs du territoire où l'identification des frontières serait plus facile. Il existe une relation statistiquement significative entre le contraste des variables examinées et la probabilité d'existence d'une frontière. D'autres frontières sont identifiées à partir des p-e. Elles correspondent aux groupes de p-e semblables ainsi qu'aux discontinuités spatiales des variables. Les deux types de frontières coïncident lorsque les différences par rapport au voisinage sont intenses. Quant aux estimations du volume ligneux, les modèles de régression arborescente forment des groupes de p-e dont la variabilité est inférieure à celle des groupes définis par la méthode actuelle utilisée au Québec. Les strates forestières auxquelles ces estimations sont associées sont caractérisées en fonction de la concordance carte/p-e, aux échelles locale et du voisinage. Pour certaines strates, la concordance locale est élevée, tandis que celle du voisinage est faible et vice-versa. La concordance locale peut être attribuée au degré d'homogénéité interne des strates et à la représentativité des p-e, tandis que la concordance avec le voisinage rend compte de l'incertitude positionnelle des p-e, des frontières et/ou des conditions structurales de ces dernières. Finalement, la fusion des frontières cartographiques et de terrain peut se faire sur le plan des connaissances, étant données les différences de résolution spatiale et du degré de généralisation des deux cartes. De façon générale, les méthodes employées dans cette thèse permettent de mieux caractériser la carte forestière et ses éléments améliorant l'estimation locale du volume ligneux.

## **Abstract**

In the traditional forest maps in Quebec, sharp lines represent stand boundaries, forest strata are considered equally concordant in relation to field/cartographic attributes and the woody volume estimations issue from successive regrouping of field plots.

The general objective of this research is to explore how different methods can be combined to characterize stand boundaries, forest strata and woody volume estimations, from geomatic and ecological points of view, by map and field information fusion.

In order to qualify boundaries, their contrast and spatial contexts are quantified. The integration of these aspects allows the detection of forest sectors where boundary identification could be easiest. A significant relationship exists among the contrast of examined variables with the existence probability of boundaries. Other boundaries are identified from the field plots. These ones correspond to homogeneous plot clusters and to strong spatial discontinuities in the field variables. The two boundary types coincide spatially when neighbourhood differences are intense. The woody volume boundaries are, in this scale, spatially dependent on forest and topographic boundaries.

Regression trees are used for grouping field plots considering minimum volume variability. The most effective tree model is associated with field variables. The groups formed with this method present less spatial variability in comparison with the plot groups presently used in Quebec. The forest strata are qualified by their accord map/plot level. For many forest strata the local accord is high while the neighbourhood accord is low. The local accord is attributed to internal strata homogeneity and/or the fact that plots poorly represent the forest stratum. The neighbourhood accord is associated with positional uncertainty of field samples, of boundaries or both.

Finally, in order to permit cartographic and field boundary fusion at the element level, the resolution of heterogeneity between field and forest map data must be addressed first. The spatial resolution and semantic ecological significance of cartographic units are the most important of these heterogeneities.

## **Avant-propos**

Cette thèse n'aurait jamais pu être accomplie sans le soutien financier, scientifique et personnel d'un grand nombre de personnes. Je tiens tout d'abord à remercier mon directeur de recherche, Kim Lowell, pour son soutien financier, scientifique et surtout personnel, tout au long de ces années. Avec beaucoup de délicatesse, de respect et de passion, il a su me guider dans l'apprentissage de la recherche. En même temps, son attitude ouverte et respectueuse nous a permis de développer des rapports extrêmement enrichissants, autant sur le plan scientifique que personnel. Ma co-directrice, Marie-Josée Fortin a également enrichi ma formation et ce, sur tous les plans. Sa rigueur scientifique, sa disponibilité et son soutien ont été précieux pour moi. De plus, elle est pour moi un modèle et une référence en tant que femme et chercheur. Merci Josée! Je tiens aussi à remercier mon conseiller, Geoffrey Edwards, qui par des discussions passionnantes m'a stimulée à aller toujours un peu plus loin. Je remercie également le prélecteur, Marius Thériault, et l'évaluatrice externe, Benoît St-Onge, pour ses remarques et commentaires rigoureux et constructifs qui m'ont aidé à améliorer cette thèse, autant dans le fond que dans la forme.

Je remercie le Conseil national de la recherche en sciences et en génie (CNRSG), la Faculté de Foresterie et Géomatique, et le Réseau de Centres d'Excellence Geoïde pour leur soutien financier. Je remercie aussi bien sûr le Centre de Recherche en Géomatique ainsi que le Département de Sciences Géomatiques, qui fournissent un cadre de travail agréable, tout comme un soutien administratif très apprécié. Merci Jean-Jacques, Alfonso, Carmen, Annik, Sylvie, Michel et Edgar!

Pour mes amis et collègues, qu'est-ce que je peux ajouter? Vous faites que ma vie soit belle! Gilles, Marcelo, Mali, David, David, Hervé, Valérie, Michel, Pierre, Bousad, Mir, Orlando, Rodolphe, Benoît, Claudia et tant d'autres ... Merci particulièrement à Véronique pour ses corrections de français. Et un énorme merci à Anne-Marie... tu es un ange! Parmi les amis dont le soutien a été irremplaçable, je veux remercier Sylvie, Laura B., Christine, Daniela, Laura J., Rosana, Nancy, Maria Laura, Ana, Leo, Claude... Pesi, cette thèse n'aurait pas été possible sans toi.

Luc, ton amour m'as permis de me rendre au bout... merci infiniment, moïn rainmain wou!

Finalmente, quiero agradecer a mis padres, que me han comprendido y me han dado su apoyo y ayuda, aunque este proyecto implique el dolor de la distancia. Gracias Coca y Raul, porque si estoy aqui, es en parte, gracias a ustedes.

*A mis padres y a las mujeres que me han  
mostrado el camino : Jeanne, Ramona, Coca,  
Pesi, Cristina, José*

# Table des matières

Résumé court .....	ii
Résumé.....	iii
Abstract.....	iv
Avant-propos .....	v
Table des matières .....	viii
Liste des tableaux.....	xiv
Liste des figures .....	xviii
Liste des espèces d'arbres.....	xxii
Introduction.....	1
Objectifs et hypothèses de recherche.....	3
Structure de la thèse .....	5
Chapitre I - Cadre théorique .....	6
Le concept d'échelle .....	7
Approche écologique .....	9
1. Le paysage forestier.....	9
2. Les composantes du paysage forestier.....	10
2.1. Les parcelles .....	10
2.2. Les frontières .....	13
3. Les facteurs qui structurent les paysages forestiers.....	17
3.1. Les facteurs abiotiques.....	17
3.2. Les interactions biologiques .....	18
3.3. L'utilisation du territoire.....	18
3.4. Les perturbations et la succession écologique .....	18
Approche géomatique .....	22
1. Les modèles conceptuels de l'espace géographique.....	22
2. Qualité et incertitude des données géographiques.....	23
3. Les données forestières.....	25
3.1. La carte forestière traditionnelle.....	25
3.2. Les données de terrain .....	27
4. L'incertitude spatiale .....	28
4.1. Description de l'incertitude spatiale .....	29

4.2. Modélisation de l'incertitude spatiale.....	30
4.3. Modèles d'incertitude spatiale des cartes forestières.....	31
Approche analytique.....	33
1. Analyse spatiale.....	34
1.1. Mesures de dépendance spatiale.....	34
1.2. Détection de patrons spatiaux.....	35
1.3. Analyse de frontières.....	35
2. Analyse exploratoire non spatiale.....	37
3. La fusion des données.....	37
Résumé du chapitre.....	42
Chapitre II - Site d'étude et données disponibles.....	43
Site d'étude.....	43
1. Climat et physiographie.....	44
2. Végétation.....	44
3. Perturbations.....	46
3.1. Feu.....	46
3.2. Épidémies.....	46
3.3. Chablis.....	47
3.4. Coupes forestières.....	47
Données disponibles.....	48
1. Cartes forestières.....	48
2. Données de terrain.....	51
3. Autres cartes.....	53
Chapitre III - Les frontières de la carte forestière.....	54
Introduction.....	54
Méthodologie.....	56
1. Données.....	56
2. Méthodes.....	56
2.1. Quantification du contraste.....	57
2.2. Quantification du contexte.....	67
2.3. Le contexte des peuplements.....	68

2.4. Évaluation de la relation entre le contraste et le contexte, et la probabilité d'existence d'une frontière .....	69
Résultats et discussion .....	71
2. Quantification du contraste.....	71
1.1. Le contraste écologique quantifié au moyen de l'indice S .....	71
1.2. Le contraste quantifié par la force à différentes échelles.....	76
2. Quantification du contexte.....	84
2.1. Le contexte des frontières .....	84
2.2. Le contexte des peuplements .....	87
3. Évaluation du lien entre le contraste et le contexte, et la probabilité d'existence d'une frontière.....	91
3.1. Distributions de fréquence du contraste en fonction de la <i>Consistance</i> .....	91
3.2. Relation entre le contraste et le contexte, et la consistance d'une frontière .....	93
Conclusions.....	95
Résumé du chapitre.....	98
Chapitre IV - Les frontières issues des données de terrain.....	99
Introduction.....	99
Méthodologie .....	101
1. Données et variables .....	101
2. Méthodes.....	102
2.1. Identification des groupes spatiaux : frontières indirectes.....	102
2.2. Identification des différences : frontières directes .....	104
2.3. Analyse de frontières .....	108
2.4. Analyse de la structure spatiale locale des frontières .....	111
2.5. Relation non spatiale entre les frontières directes du Volume total et les variables forestières et environnementales .....	112
Résultats et discussion .....	113
1. Les groupes spatiaux et les frontières indirectes .....	113
2. Les frontières directes .....	117
3. Signification statistique des frontières et de leurs coïncidences spatiales .....	119
3.1. Analyses des sous-frontières.....	119
3.2. Analyses de recouvrement .....	120
4. Structure spatiale locale des variables .....	121

4.1. Volume total .....	122
4.2. <i>Type forestier</i> .....	124
5. Relation entre les frontières du Volume total et les variables forestières et environnementales .....	125
Conclusions.....	128
Résumé du chapitre.....	131
Chapitre V - Les estimations du volume de matière ligneuse .....	132
Introduction.....	132
Méthodologie.....	135
1. Le regroupement des p-e selon la méthode actuelle au Québec.....	135
2. Données et variables .....	136
3. Méthodes.....	138
3.1. La régression arborescente.....	139
3.2. Comparaison des groupes formés par les différentes méthodes en terme de la variabilité des estimations du <i>Volume total</i> .....	144
3.3. Comparaison des estimations du <i>Volume total</i> issues des différentes méthodes avec celles des p-e.....	145
3.4. Comparaison des groupes issus des différentes méthodes en terme de la structure spatiale des estimations du <i>Volume total</i> .....	146
Résultats et discussion .....	150
1. Caractéristiques générales des modèles.....	150
2. Les modèles de régression arborescente de taille optimale .....	152
2.1. Le modèle Cartographique – Carto28 et Carto21 .....	152
2.2. Le modèle Partiel de Terrain – PartielT17 et PartielT9.....	156
3. La variabilité des estimations du <i>Volume total</i> des groupes.....	160
3.1. Le Coefficient de variation (CV) du <i>Volume total</i> des groupes .....	160
3.2. Rapport entre le <i>Volume total</i> estimé pour les groupes et celui des p-e .....	162
3.3. Caractéristiques des groupes de faible et de forte hétérogénéité .....	164
4. La représentation spatiale des groupes issus des différents modèles.....	166
4.1. Comparaison des caractéristiques des cartes .....	166
4.2. Comparaison des groupes issus des différentes méthodes en terme de la structure spatiale des estimations du <i>Volume total</i> .....	168
Conclusions.....	169
Résumé du chapitre.....	172

Chapitre VI - Les strates forestières .....	173
Introduction.....	173
Méthodologie.....	176
1. Données et variables .....	176
2. Méthodes.....	177
2.1. Quantification de la concordance carte/p-e pour les strates de la carte forestière .....	179
2.2. Quantification de la concordance carte/p-e pour les combinaisons de variables .....	182
2.3. Quantification de la concordance carte/p-e pour les catégories des variables.....	184
2.4. Relation entre la concordance carte/p-e et l'incertitude spatiale des p-e.....	185
Résultats et discussion .....	186
1. La concordance des strates forestières.....	186
1.1. Analyse univariée .....	186
1.2. Analyse bivariée .....	189
1.3. Analyses multivariées .....	192
2. La concordance des combinaisons de variables .....	198
3. La concordance des catégories .....	199
4. La concordance carte/p-e et l'incertitude spatiale des p-e.....	203
Conclusions.....	206
Résumé du chapitre.....	208
Chapitre VII - Discussion .....	209
Fusion des informations spatiales de la carte forestière et des échantillons de terrain...209	
1. Différences et ressemblances liées au phénomène représenté.....	212
2. Différences et ressemblances liées à l'aspect sémantique .....	212
3. Différences et ressemblances liées à l'aspect syntactique .....	214
3.1. Différences qui ont trait à la résolution spatiale .....	215
3.2. Différences qui ont trait aux méthodes de saisie des variables.....	220
3.3. Différences liées à la structure des données .....	221
4. Différences liées à l'aspect schématique .....	222
5. Conclusion sur la fusion des deux représentations de la forêt.....	222
Les estimations du volume de matière ligneuse .....	224
Résumé du chapitre.....	227

Conclusions.....	229
Futures avenues de recherche .....	233
Utilité de cette recherche pour la planification des activités forestières.....	234
Références bibliographiques.....	236
Annexe I - Variables cartographiques et de terrain .....	250
Variables cartographiques.....	250
Variables de terrain.....	251
Annexe II Les frontières cartographiques.....	254
Annexe III - Modèles de régression arborescente .....	260
Modèle Cartographique général - CartoG .....	260
Modèle Cartographique taillé à 45 feuilles - Carto45.....	262
Modèle Cartographique de taille optimale – Carto28.....	264
Modèle Cartographique, version non géographique – Carto36.....	265
Modèle Cartographique de taille optimale, version non géographique – Carto21 .....	267
Modèle Complet de Terrain général - CompletTG.....	268
Modèle Partiel de Terrain général - PartielTG .....	271
Modèle Partiel de Terrain taillé à 45 feuilles - PartielT45 .....	273
Modèle Partiel de Terrain de taille optimale – PartielT17.....	275
Modèle Partiel de Terrain, version non géographique – PartielT42.....	276
Modèle Partiel de Terrain de taille optimale, version non géographique – PartielT9 ....	278
Annexe IV Les strates forestières .....	280
Exemple de matrice de confusion.....	280
Analyses univariées .....	281
Analyses bivariées .....	283
Analyses multivariées .....	285

## Liste des tableaux

Tableau 1. Catégories de <i>Types de couvert</i> et de <i>Groupements d'essences</i> présentes à la Forêt Montmorency. ST : surface terrière. (Ministère de Ressources naturelles du Québec 1995).....	49
Tableau 2. Catégories de <i>Classe de densité</i> (Ministère de Ressources naturelles du Québec 1995).....	50
Tableau 3. Catégories de <i>Classe de hauteur</i> (Ministère de Ressources naturelles du Québec 1995).....	50
Tableau 4. Catégories de <i>Classe d'âge</i> selon le type de peuplement (Ministère de Ressources naturelles du Québec 1995).....	50
Tableau 5. Catégories de la variable <i>Groupement d'essences (GE)</i> . Intervalle de couvert : pourcentage de la surface terrière du peuplement occupé par une essence ou type de végétation. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).....	60
Tableau 6. Matrice de similarité partielle entre les intervalles de couvert pour les essences S, E et Pb. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).....	61
Tableau 7. Matrice de similarité partielle entre les intervalles de couvert pour les essences Bb et Pe. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).....	61
Tableau 8. Matrice de similarité partielle entre les intervalles de couvert pour l'essence Bj. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).....	62
Tableau 9. Calcul de $S_{GE}$ entre les catégories SE et BbS. La valeur de $s_{12j}$ est extraite des matrices de similarité partielle de chaque descripteur. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence, dérivé de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).....	62
Tableau 10. Calcul de $S_{CD}$ pour les catégories A et C. La valeur de $k$ est égal à celle de $d$ entre les classes A et D (voir l'équation III.3).....	63
Tableau 11. Catégories de force d'une frontière.....	67
Tableau 12. Classification de l'indice $H$ d'hétérogénéité du contexte des peuplements en degrés d'hétérogénéité.....	69
Tableau 13. Niveaux de contraste et intervalles de valeurs de $S$ pour chacune des variables ( $n = 8018$ frontières entre 3729 peuplements).....	72
Tableau 14. Proportion des frontières indifférentes au changement de voisinage selon la catégorie de force, les valeurs sont exprimées en pourcentages ( $n = 8018$ frontières).....	80

Tableau 15. Pourcentage de frontières fortes et très fortes à l'échelle locale (c'est-à-dire dans les voisinages de niveau 1 ou 2) et régionale (c'est-à-dire dans les voisinages de niveau 4 ou 5) pour les variables et les combinaison de variables ( $n=8018$ ). .....	80
Tableau 16. Pourcentage de frontières fortes locales et régionales pour les variables $CD$ , $CH$ et $CA$ , qui coïncident avec des combinaisons de perturbations dans les peuplements adjacents. cht : chablis total; ct : coupe totale, es : épidémie sévère, sp : sans perturbation. ....	83
Tableau 17. Distributions des fréquence de $S_{GE}$ , $S_{CD}$ , $S_{CH}$ et $S_{CA}$ en fonction des degrés de <i>Consistance</i> : Test de Kolmogorov-Smirnov (KS) pour deux échantillons et sa probabilité (Prob. : Probabilité; $n=8018$ frontières).....	93
Tableau 18. Coefficients linéaires de Fisher des variables discriminantes dans la fonction de classification des frontières en degrés de <i>Consistance</i> et niveau de signification du modèle retenu ( $n=4515$ frontières). Prob. : Probabilité .....	93
Tableau 19. Valeur moyenne (M) et écart-type (ÉT) des variables en fonction des degrés de <i>Consistance</i> . Prob. : Probabilité.....	95
Tableau 20. Variables considérées pour l'analyse des frontières. ....	101
Tableau 21. Proportion de non coïncidences pour que les CEFs soient considérées comme des EFs.....	107
Tableau 22. Description des statistiques de sous-frontières. ....	109
Tableau 23. Statistiques de recouvrement. $G$ et $H$ sont deux ensembles de frontières. ....	110
Tableau 24. Signification des statistiques de sous-frontières pour les variables qualitatives ( $\alpha= *-0,01$ , $**--0,05$ et $***--0,10$ ). M : indique un paysage morcelé. C : indique une forte contiguïté des frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au Tableau 20 et les statistiques au Tableau 22.....	119
Tableau 25. Signification des probabilités des statistiques de recouvrement qui évaluent l'association des frontières directes et indirectes du <i>Volume Total</i> et de $MQ$ ( $\alpha= *-0,01$ , $**--0,05$ et $***--0,10$ ). R : indique répulsion entre les frontières. A : indique association entre les frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au le Tableau 20 et les statistiques au Tableau 23.....	120
Tableau 26. Signification des probabilités des statistiques de superposition qui évaluent la relation spatiale entre les frontières indirectes du <i>Volume total</i> et celles des autres variables ( $\alpha= *-0,01$ , $**--0,05$ et $***--0,10$ ). R : indique répulsion entre les frontières. D : indique dépendance des frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au Tableau 20 et les statistiques au Tableau 23.....	121
Tableau 27. Énumération des variables de terrain et cartographiques utilisées pour les modèles de régression arborescente.....	137
Tableau 28. Nombre de p-e avec des données manquantes des ensembles utilisés pour la définition (338 p-e) et la validation (183 p-e) des modèles de régression arborescente. ....	138

Tableau 29. Correspondance entre les variables de terrain et les variables cartographiques pour le transfert des règles des modèles de régression arborescente aux peuplements de la carte forestière de 1992.....	147
Tableau 30. Caractéristiques des modèles de régression arborescente. Dev. : Déviance... 151	
Tableau 31. Fréquences des catégories des variables estimées sur le terrain <i>Groupement d'essences</i> , <i>Classe de densité</i> , <i>Classe de hauteur</i> , <i>Classe d'âge</i> , <i>Perturbation d'origine</i> et <i>Perturbation moyenne</i> , dans les groupes RT, Carto45 et PartielT45 de faible et de forte hétérogénéité. La dernière colonne contient les fréquences des catégories pour l'ensemble des 338 p-e. La signification correspond au $\chi^2$ de chaque distribution. ...	165
Tableau 32. Caractéristiques non spatiales des cartes RT, Carto45, PartielT45, Carto28 et PartielT17.....	167
Tableau 33. Valeurs de l'effet pépète, du plateau et de l'étendue des variogrammes théoriques des estimations du <i>Volume total</i> issues des différents modèles ( $n=3792$ points). .....	168
Tableau 34. Catégories de la variable <i>Classe d'âge (CA)</i> sur la carte forestière et sur les p-e, ainsi que celles utilisées dans cette étude comme catégories des p-e.....	177
Tableau 35. Valeurs de l'indice de concordance (ICL et ICV) entre la catégorie de la carte et celle de la p-e des variables <i>Groupement d'essences (GE)</i> , <i>Classe de densité(CD)</i> , <i>Classe de hauteur(CH)</i> et <i>Classe d'âge (CA)</i> .....	186
Tableau 36. Indices de concordance carte/p-e des strates forestières pour les combinaisons de deux variables de <i>Groupement d'essences (GE)</i> , <i>Classe de densité (CD)</i> , <i>Classe de hauteur (CH)</i> et <i>Classe d'âge (CA)</i> , aux échelles locales (ICL) et de voisinage spatial (ICV).....	189
Tableau 37. Indices de concordance carte/p-e des strates forestières pour les combinaisons de trois variables parmi <i>Groupement d'essences (GE)</i> , <i>Classe de densité (CD)</i> , <i>Classe de hauteur (CH)</i> et <i>Classe d'âge (CA)</i> , aux échelles locales (ICL) et du voisinage spatial (ICV). .....	193
Tableau 38. Indices de concordance carte/p-e des strates forestières pour la combinaison de quatre variables, <i>Groupement d'essences (GE)</i> , <i>Classe de densité (CD)</i> , <i>Classe de hauteur (CH)</i> et <i>Classe d'âge (CA)</i> , aux échelles locales (ICL) et du voisinage spatial (ICV).....	195
Tableau 39. Indices de concordance carte/p-e des combinaisons de variables, toutes strates confondues, aux échelles locales (ICL) et de voisinage (ICV).....	198
Tableau 40. Indices de concordance carte/p-e des catégories de la variable <i>Groupement d'essences (GE)</i> , aux échelles locales (ICL) et de voisinage (ICV).....	199
Tableau 41. Indices de concordance carte/p-e des catégories de la variable <i>Classe de densité (CD)</i> , aux échelles locales (ICL) et de voisinage (ICV).....	200
Tableau 42. Indices de concordance carte/p-e des catégories des variables <i>Classe de hauteur (CH)</i> et <i>Classe d'âge (CA)</i> , aux échelles locales (ICL) et de voisinage (ICV). .....	202

Tableau 43. Valeurs de concordance globale, de concordance par catégorie et de coefficient kappa des variables <i>Groupement d'essences (GE)</i> , <i>Classe de densité (CD)</i> , <i>Classe de hauteur (CH)</i> et <i>Classe d'âge (CA)</i> correspondant à l'ensemble des p-e et aux p-e distants d'au moins 5 m, 10 m, 25 m ou 50 m de la frontière de la strate forestière. .204	204
Tableau 44. Différences et ressemblances au plan des aspects sémantiques, syntactiques et schématiques entre les deux représentations de la forêt. ....211	211
Tableau 45. Aire minimale de photo-interprétation (Ministère de Ressources naturelles du Québec 1995). ....250	250
Tableau 46. Catégories de <i>Type de couvert</i> (Ministère de Ressources naturelles du Québec 1995). ....250	250
Tableau 47. Catégories de <i>Classe de pente</i> (Ministère de Ressources naturelles du Québec 1995). ....251	251
Tableau 48. Catégories de <i>Stade de développement</i> (Conseil canadien des ministres des forêts 2004). ....251	251
Tableau 49. Catégories de <i>Versant</i> (Ministère des Ressources naturelles du Québec 1994). ....251	251
Tableau 50. Catégories de <i>Situation sur la pente</i> (Ministère des Ressources naturelles du Québec 1994). ....252	252
Tableau 51. Catégories de <i>Forme de la pente</i> (Ministère des Ressources naturelles du Québec 1994). ....252	252
Tableau 52. Catégories de <i>Texture de l'horizon B</i> et de <i>Texture de l'horizon C</i> (Ministère des Ressources naturelles du Québec 1994). ....252	252
Tableau 53. Catégories de <i>Type de humus</i> (Ministère des Ressources naturelles du Québec 1994). ....253	253
Tableau 54. Catégories d'épaisseur du <i>Dépôt de surface</i> (Ministère des Ressources naturelles du Québec 1994). ....253	253
Tableau 55. Catégories de <i>Classe de drainage</i> avec ses modificateurs (Ministère des Ressources naturelles du Québec 1994). ....253	253
Tableau 56. Valeurs de l'indice de similitude écologique $S_{GE}$ . ( $n = 8018$ frontières entre 3729 peuplements). Les valeurs en gras sont mentionnées dans le texte. ....257	257
Tableau 57. Valeur de l'indice de similitude écologique $S_{CD}$ ( $n = 8018$ frontières entre 3729 peuplements). ....258	258
Tableau 58. Valeurs de l'indice de similitude écologique $S_{CH}$ ( $n = 8018$ frontières entre 3729 peuplements). ....258	258
Tableau 59. Valeurs de l'indice de similitude écologique $S_{CA}$ ( $n = 8018$ frontières entre 3729 peuplements). ....258	258
Tableau 60. Matrice de confusion des catégories de la carte et de la p-e pour la variable <i>Classe de densité (CD)</i> , toutes strates forestières ensemble. ....280	280

## Liste des figures

Figure 1. Sources des estimations des superficies et du volume ligneux des strates forestières.....	2
Figure 2. Les régimes de perturbations, les réponses du biote et les unités de végétation à différentes échelles spatiales et temporelles (Modifiée de Delcourt et Delcourt 1988). 8	8
Figure 3. Localisation de la Forêt Montmorency (source : Bélanger et Bouliane 1992). ....	43
Figure 4. Étapes méthodologiques du Chapitre III. ....	57
Figure 5. Voisinages considérés pour quantifier la force de frontières. ....	66
Figure 6. Distribution des fréquences des frontières selon le niveau de contraste. a) <i>Groupement d'essences</i> ; b) <i>Classe de densité</i> ; c) <i>Classe de hauteur</i> ; d) <i>Classe d'âge</i> ; e) <i>Contraste global</i> . (n = 8018 frontières entre 3729 peuplements).....	74
Figure 7. Le contraste des frontières : valeurs de <i>S</i> reclassées en niveaux de contraste. a) <i>Groupement d'essences</i> , b) <i>Classe de densité</i> (n = 8018 frontières entre 3729 peuplements).....	77
Figure 8. Frontières fortes et très fortes locales et régionales. a) combinaison des trois variables à la fois; b) <i>CD</i> ; c) <i>CH</i> ; d) <i>CA</i> , e) <i>CD</i> ou <i>CH</i> ou <i>CA</i> (n=8018 frontières). ...	82
Figure 9. Valeurs d'autocorrélation spatiale locale du contraste de: a) <i>Groupement d'essences (GE)</i> , b) <i>Classe de densité (CD)</i> , c) <i>Classe de hauteur (CH)</i> , d) <i>Classe d'âge (CA)</i> et e) <i>Contraste Global (CG)</i> . ....	85
Figure 10. Caractère complémentaire du contraste et du contexte des frontières. ....	86
Figure 11. Proportion de surfaces et de peuplements correspondants à chaque niveau d'hétérogénéité ( <i>H</i> ), selon les variables <i>GE</i> , <i>CD</i> , <i>CH</i> et <i>CA</i> . ....	87
Figure 12. Hétérogénéité (quantifiée par <i>H</i> ) du contexte des peuplements. a) <i>GE</i> ; b) <i>CD</i> ; c) <i>CH</i> .....	89
Figure 13. Relation entre le contraste et le contexte des frontières, et le contexte des polygones.....	91
Figure 14. Distributions des fréquence des niveaux de contraste des frontières selon le degré de <i>Consistance</i> . a) <i>GE</i> ; b) <i>CD</i> ; c) <i>CH</i> ; d) <i>CA</i> (n=8018 frontières). ....	92
Figure 15. Étapes méthodologiques du Chapitre IV.....	102
Figure 16. Réseau de triangulation de Delaunay qui établit l'adjacence entre les p-e de la Forêt Montmorency. ....	103
Figure 17. Identification des frontières directes par la méthode de triangulation de Womble pour les variables quantitatives. a) Identification des CEFs et des EFs, b) angle du gradient entre les EFs, c) angle entre le gradient et la connexion entre les EFs. (Adaptée de Jacquez et Maruca 2001).....	106
Figure 18. Identification des frontières directes par la méthode de Womble pour des variables qualitatives (Adaptée de Jacquez et Maruca 2001). ....	108

- Figure 19. Frontières indirectes du *Volume total* : a) 100 groupes; b) 50 groupes, c) *Topo-q*, 100 groupes et d) *Topo-q*, 50 groupes. .... 114
- Figure 20. Frontières indirectes pour le *Volume total*. Valeurs minimales, maximales, moyennes et de variance du *Volume total* des groupes qui contiennent au moins 1 % des p-e pour : a) 100 groupes et b) 50 groupes. Répartition spatiale des : c) 100 groupes et d) 50 groupes. .... 115
- Figure 21. Frontières directes entre les p-e pour les variables : a) *Volume total* et b) *Type forestier*. .... 117
- Figure 22. Structure spatiale locale du *Volume total* ( $n=539$  p-e). .... 122
- Figure 23. Structure spatiale locale des frontières directes du *Volume total* : a) Répartition spatiale des valeurs d'*I* de Moran local. b) Détail du secteur central, c) détail du secteur sud. .... 123
- Figure 24. Structure spatiale locale des frontières directes du *Type forestier*. .... 124
- Figure 25. Distribution de fréquences des différences « maximum-minimum » de chaque variable dans les p-e associées aux EFs (graphiques à gauche) et aux CEFs (dont la valeur est inférieure au seuil du percentile 50) (graphiques à droite) du *Volume total*. a) et b) *Densité totale*. c) et d) *Surface terrière*. e) et f) *Hauteur des arbres dominants*. .... 126
- Figure 26. Fréquence des combinaisons de catégories observées dans les p-e associées aux EFs et aux CEFs pour *Volume total*. Combinaison de catégories de : a) *Groupement d'essences*, b) *Classe de hauteur*, c) *Classe d'âge*, d) *Perturbation d'origine*. ct : coupe totale, es : épidémie sévère, sp : sans perturbation. .... 127
- Figure 27. Étapes méthodologiques du Chapitre V. .... 138
- Figure 28. Variogramme général qui montre l'effet pépité, le plateau et la portée (Adaptée de Kaluzny *et al.* 1998). .... 149
- Figure 29. Dendrogramme du modèle Cartographique de taille optimale. a) Version géographique, Carto28 et b) version non-géographique, Carto21. La longueur des branches est proportionnelle à la déviance (Voir le Tableau 27 pour la définition des variables). .... 153
- Figure 30. Dendrogramme du modèle Cartographique de taille optimale, dans la version qui considère les variables géographiques, Carto28. La longueur des branches est uniforme. .... 155
- Figure 31. Dendrogramme du modèle Partiel de Terrain de taille optimale. a) version géographique, PartielT17 et b) version non-géographique, PartielT9. La longueur des branches est proportionnelle à la déviance. (Voir le Tableau 27 pour la définition des variables). .... 157
- Figure 32. Dendrogramme du modèle Partiel de Terrain de taille optimale, PartielT17. La longueur des branches est uniforme. .... 159
- Figure 33. Percentiles des coefficients de variation du *Volume total* des groupes issus de la méthode actuelle de regroupement (RT) et des modèles de régression arborescente

Cartographique et Partiel de Terrain : a) versions géographiques (Carto45, Carto28, PartielT45 et PartielT17) et b) versions non géographiques (Carto36, Carto21, PartielT42 et PartielT9). .....	161
Figure 34. Rapport entre le <i>Volume total</i> estimé pour les groupes formés par les différentes méthodes et celui des p-e. Les valeurs minimales, maximales et médianes sont présentées pour chaque méthode de groupement. Les p-e sont regroupées en quintiles selon le <i>Volume total</i> : a) 0,4-54 m <sup>3</sup> /ha, b) 54-97 m <sup>3</sup> /ha, c) 97-128 m <sup>3</sup> /ha, d) 128-173 m <sup>3</sup> /ha, e) 173-306 m <sup>3</sup> /ha et f) 205-306 m <sup>3</sup> /ha. ....	163
Figure 35. Les variables les plus importantes des deux modèles de régression arborescente. ....	171
Figure 36. Relation entre les échelles des informations (signaux) des polygones/peuplements de la carte forestière et les placettes-échantillons (p-e). Modifiée de Allen et Hoekstra (1991). ....	174
Figure 37. Étapes méthodologiques du Chapitre VI. ....	178
Figure 38. Les 20 strates les plus importantes en superficie, nombre de polygones et de p-e, sur la carte forestière de la Forêt Montmorency. ....	179
Figure 39. Échelle de voisinage spatial. Les points noirs sont les p-e (A, B et C, dans cet exemple). Les cercles indiquent le voisinage spatial de chaque p-e sur la carte forestière de 1992. Les couleurs des polygones représentent les strates forestières...	180
Figure 40. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons d'une variable, aux échelles locales (ICL) et du voisinage spatial (ICV). ....	188
Figure 41. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons de deux variables aux échelles locales (ICL) et du voisinage spatial (ICV). ....	191
Figure 42. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons de trois variables aux échelles locales (ICL) et du voisinage spatial (ICV). ....	194
Figure 43. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour la combinaison de quatre variables, aux échelles locales (ICL) et du voisinage spatial (ICV). ....	196
Figure 44. Gradients de concordance locale pour les catégories des variables individuelles et en combinaisons. ....	203
Figure 45. Exemples de superposition des frontières cartographiques aux frontières issues des p-e : a) et b) frontières directes de <i>Type forestier</i> , c) frontières directes et indirectes du <i>Volume total</i> . ....	216
Figure 46. Exemples de superposition de frontières cartographiques à fort contraste à l'échelle régionale (ou de voisinage) aux frontières issues de données de terrain : a) Frontières directes de <i>Type Forestier</i> , b) Frontières indirectes (100 groupes) du <i>Volume total</i> et c) Frontières directes du <i>Volume total</i> . ....	217

Figure 47. Exemples de la superposition des frontières cartographiques des groupes de p-e avec une même estimation du volume ligneux aux frontières de terrain pour le <i>Volume total</i> : a) frontières indirectes et b) frontières directes.....	219
Figure 48. Exemples de situations observées lorsque les frontières cartographiques sont superposées aux frontières directes de <i>Type forestier</i> . ....	221
Figure 49. Relation entre les frontières cartographiques et de terrain. ....	224
Figure 50. Procédure proposée pour la fusion des informations sur l'incertitude sur les estimations du volume ligneux avec celles liées à la concordance carte/p-e des strates forestières.....	227
Figure 51. Comparaison des intervalles de couvert des essences dans les différentes catégories de <i>Groupement d'essences</i> : a) Sapin baumier (S); b) Épinette (E); c) Pin blanc (Pb); d) Bouleau à papier (Bb); e) Bouleau jaune (Bj) et f) Peuplier (Pe).....	255
Figure 52. Exemple du calcul de la force d'une frontière (Csillag <i>et al.</i> 2001). ....	256
Figure 53. Pourcentage de chacune des catégories des variables <i>Groupement d'essences</i> , <i>Classe de densité</i> , <i>Classe de hauteur</i> et <i>Classe d'âge</i> des peuplements de la carte forestière de 1992 de la Forêt Montmorency.....	259
Figure 54. Dendrogramme du modèle Complet de Terrain général – CompletTG. La longueur des branches est proportionnelle à la déviance de chaque nœud.....	271
Figure 55. Indice de concordance carte/p-e des strates forestières pour les variables individuelles aux échelles locales (ICL) et du voisinage spatial (ICV).....	282
Figure 56. Indice de concordance carte/p-e des strates forestières pour les combinaisons de deux variables, aux échelles locales (ICL) et du voisinage spatial (ICV). ....	284
Figure 57. Indice de concordance carte/p-e des strates forestières pour les combinaisons de trois variables, aux échelles locales (ICL) et du voisinage spatial (ICV).....	287
Figure 58. Indice de concordance carte/p-e des strates cartographiques pour la combinaison des quatre variables, aux échelles locales (ICL) et du voisinage spatial (ICV).....	288

## Liste des espèces d'arbres

<b>Nom scientifique</b>	<b>Nom français</b>	<b>Abbréviation</b>
<i>Abies balsamea</i>	Sapin baumier	S
<i>Picea glauca</i>	Épinette blanche	E
<i>Picea mariana</i>	Épinette noire	E
<i>Pinus strobus</i>	Pin blanc	Pb
<i>Betula papyrifera</i>	Bouleau à papier ou Bouleau blanc	Bb
<i>Betula alleghaniensis</i>	Bouleau jaune	Bj
<i>Populus tremuloides</i>	Peuplier faux-tremble	Pe

## Introduction

La gestion des ressources naturelles, dont la forêt fait partie, s'appuie sur la connaissance de la répartition spatiale, des caractéristiques structurales, de la dynamique temporelle et de la productivité de la ressource. Cette connaissance dérive de l'information obtenue à partir de l'analyse des données portant sur différents aspects de la ressource. La fiabilité de cette information affecte donc la prise de décisions (Zhang et Goodchild 2002).

Au Québec, la carte forestière est le moyen utilisé pour synthétiser la position et les caractéristiques des écosystèmes forestiers. Il s'agit d'une première étape de l'inventaire forestier à partir duquel seront dérivées les estimations sur le volume ligneux, information de base pour l'établissement d'un plan d'aménagement forestier à long terme (Bérard et Côté 1996). Les peuplements forestiers, issus d'une stratification forestière préétablie, sont identifiés et délimités lors de la photo-interprétation. Le nombre et la position des placettes-échantillons (p-e) sur le terrain sont estimés de façon à assurer la fiabilité des estimations du volume ligneux de chaque strate forestière, en fonction de leur superficie et de leur distribution spatiale. Ainsi, dans le contexte de l'inventaire forestier, la carte forestière fournit les données sur la superficie et la distribution spatiale de chaque strate forestière, alors que l'ensemble des p-e appartenant à la strate forestière fournit les données qui permettent l'estimation de son volume ligneux (Figure 1) (Bérard et Côté 1996).

La fiabilité des estimations de superficie et du volume ligneux des strates forestières est acceptable pour la planification de certaines activités sur le territoire. Il s'agit d'une fiabilité globale, valable pour l'ensemble du territoire cartographié, étant donné que les erreurs locales d'estimations sont compensées à l'échelle globale.

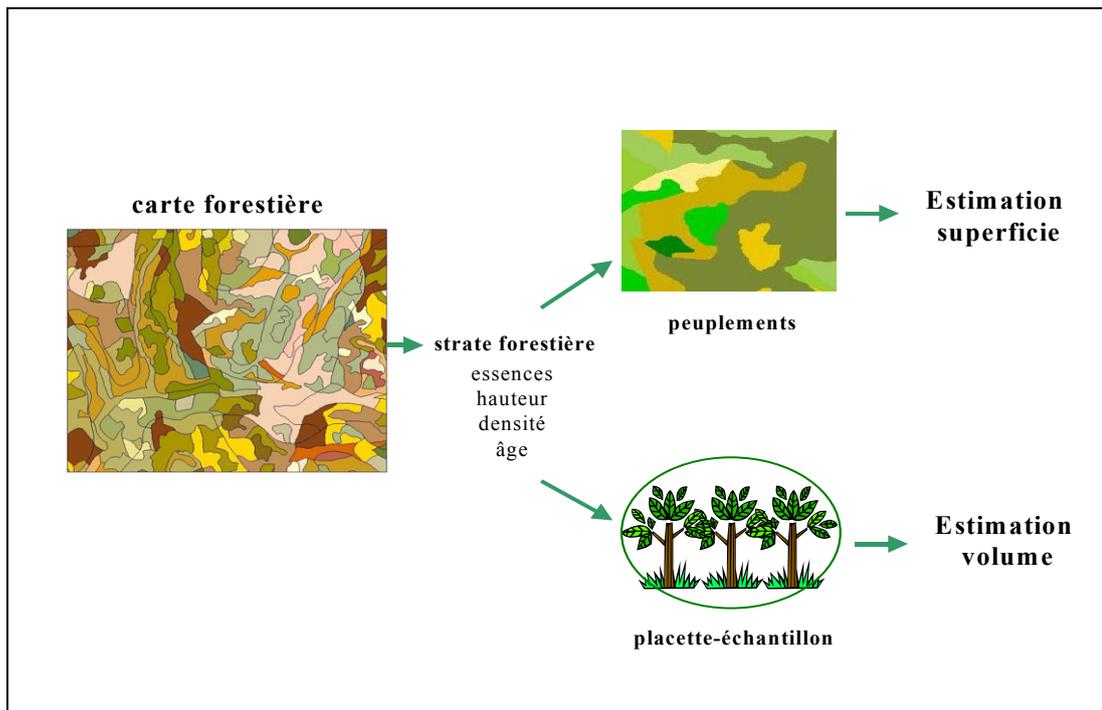


Figure 1. Sources des estimations des superficies et du volume ligneux des strates forestières.

La planification opérationnelle des activités forestières utilise le peuplement forestier comme unité de gestion (Côté 2003). À cette échelle, les estimations de superficie et du volume ligneux de la carte forestière deviennent peu fiables. Concrètement, cela signifie que la variabilité des estimations du volume ligneux d'une strate forestière devient trop grande lorsqu'elle s'applique à un peuplement particulier. De plus, l'estimation de la superficie d'un peuplement particulier est très incertaine puisque les attributs et les frontières des peuplements représentés sur la carte forestière sont eux-mêmes incertains (De Groeve 1999, De Groeve et Lowell 2000, Lowell *et al.* 2000, Lowell 1994). La carte forestière est donc un outil limité pour la gestion forestière à l'échelle locale.

Dans un contexte mondial où la valeur réelle nette de la production forestière a diminué et pour lequel la gestion durable des forêts est incontournable (Byron 2003), la précision de l'information sur la ressource forestière prend une importance économique majeure. Le défi est donc de trouver une méthode qui permettra d'obtenir une fiabilité acceptable pour la

gestion forestière locale, c'est-à-dire à l'échelle des peuplements, à partir d'informations cartographiques qui sont fiables à l'échelle des strates forestières.

## **Objectifs et hypothèses de recherche**

La recherche proposée vise à améliorer la fiabilité locale des estimations du volume ligneux et des surfaces d'une carte forestière.

Dans la carte forestière traditionnelle du Québec, toutes les frontières entre les peuplements sont représentées par des lignes identiques. Les estimations du volume ligneux, même si elles sont issues de regroupements successifs de placettes-échantillons, sont considérées comme inhérentes aux strates forestières. De plus, la concordance entre cette information de terrain et l'information cartographique est considérée comme similaire pour toutes les strates forestières.

**L'objectif général de cette recherche consiste à explorer différentes voies pour caractériser les éléments suivants d'une carte forestière : les frontières entre les peuplements, les strates forestières et les estimations du volume ligneux, en fonction des aspects liés à l'incertitude sur l'existence des frontières, à la précision des estimations et à la concordance entre l'information cartographique et de terrain; et ce, en fusionnant les informations provenant des deux sources.**

Cette recherche vise particulièrement à répondre à trois questions relatives à la carte forestière :

1. Est-il possible de déterminer le degré d'incertitude sur l'existence d'une frontière à partir de caractéristiques telles que le contraste écologique et les conditions du contexte spatial?

L'hypothèse de base est la suivante : l'identification d'une frontière par un photo-interprète serait facilitée dans le cas où le contraste de texture, d'intensité et de couleur des peuplements adjacents à la frontière est élevé (Edwards et Lowell 1996). La texture, intensité et couleur de la photographie sont utilisées pour interpréter les attributs des

peuplements (De Groeve 1999). En conséquence, le contraste des attributs des peuplements adjacents pourrait être relié à la probabilité d'existence de la frontière. Dans la mesure que le contraste des attributs des peuplements est plus fort, la frontière entre eux serait plus certaine.

2. Quelles sont les variables environnementales et forestières les plus importantes pour diviser l'ensemble des p-e selon des changements élevés de variance du volume ligneux?

Jusqu'à maintenant, les critères utilisés pour le regroupement des p-e sont fixés « *a priori* », c'est-à-dire qu'ils ne tiennent pas compte de la structure inhérente des données écologiques des p-e auxquelles ils seront appliqués. Selon une perspective différente, une technique exploratoire comme la régression arborescente peut aider à identifier les critères qui découlent de la structure interne des données (Clark et Pregibon 1992, De'Ath et Fabricius 2000). Ces critères peuvent être utilisés pour classer les p-e en groupes homogènes par rapport au volume ligneux.

L'hypothèse de base est que l'utilisation d'une technique exploratoire telle la régression arborescente pour regrouper les p-e en fonction du volume ligneux produira des estimations moins variables que celles de la méthode utilisée actuellement au Québec.

3. Toutes les strates forestières présentent-elles le même degré de concordance avec l'information de terrain? Ce degré dépend-il de la variable ou de l'ensemble des variables considérées?

Les strates de la carte forestière impliquent une certaine uniformité en termes des attributs qui les définissent. Toutefois, cette uniformité ou homogénéité est loin d'être réelle (Bérard et Côté 1996, Johnston 1999, Lowell et Liu 1999). Les méthodes de traitement de l'incertitude sur les attributs cartographiques ont été développées principalement pour évaluer les classifications des images satellites (Chrisman 1991, Verigin 1999). La majorité de ces méthodes sont basées sur la comparaison des classes de l'image à évaluer avec des données de référence provenant d'une source considérée comme plus fiable (Chrisman

1991, Foody 2002, Friedl *et al.* 2001, Stehman 1997, 1998, Veregin 1999). Les dissimilarités entre les deux jeux de données sont généralement attribuées aux erreurs de la classification évaluée, bien que la classification de référence puisse également contenir des erreurs (Foody 2002).

Dans le contexte de cette étude, la carte forestière et les p-e peuvent être considérées comme deux représentations de la même réalité, bien que de nature différente. Si l'on considère que les attributs notés dans les deux sont exacts, la comparaison entre les catégories des strates de la carte et celles des p-e peut donner de l'information sur la concordance qui, pour une même variable, existe entre les deux jeux de données.

L'hypothèse est la suivante : étant donné que l'uniformité des strates forestières cartographiques est relative et que rarement elle correspond à la réalité des conditions du terrain, la concordance entre les informations de la carte et du terrain varie en fonction des strates considérées.

## **Structure de la thèse**

Les notions fondamentales sont présentées dans le Cadre Théorique, au Chapitre I. Le chapitre suivant décrit le site à l'étude et les données utilisées. La première question de recherche, concernant l'incertitude sur l'existence de frontières cartographiques, est développée au Chapitre III. Les données de terrain sont utilisées pour établir des frontières pour les variables d'intérêt au Chapitre IV. L'utilisation des modèles de régression arborescente pour regrouper les p-e et obtenir des estimations sur le volume ligneux moins variables que celles produites par la méthode utilisée actuellement au Québec, est explorée au Chapitre V. La troisième question, sur la concordance entre les informations cartographiques et de terrain pour les strates forestières est développée au Chapitre VI. Une discussion sur la fusion des informations obtenues aux chapitres précédents est ensuite présentée au Chapitre VII et, finalement, les conclusions générales sont exposées.

## Chapitre I - Cadre théorique

La philosophie de l'aménagement du territoire et de la foresterie a beaucoup changé depuis une dizaine d'années, en ce sens qu'une perspective plus large des ressources forestières et un contexte à plus long terme sont devenus des critères indispensables aux enjeux actuels (Perera *et al.* 2000). L'approche traditionnelle pour l'aménagement forestier était basée sur le peuplement afin d'optimiser la production de bois. Cette approche a évolué et vise dorénavant à organiser les peuplements dans le paysage pour obtenir différents types de bénéfices dont la pérennité de la forêt, en tant qu'écosystème, et de la faune qui y vit (Boyce 1995). L'étude de l'écologie et de la dynamique spatiale des écosystèmes forestiers, à l'aide des concepts de l'écologie du paysage, peut fournir les connaissances nécessaires afin de formuler des politiques forestières et des stratégies qui permettront la gestion du territoire en favorisant la gestion durable de nos forêts (Perera et Euler 2000).

Évidemment, une meilleure gestion des ressources naturelles nécessite des données qui soient les plus précises possible. Or, comme pour tous les phénomènes géographiques, les données sur l'état de la forêt renferment de l'incertitude de divers types qui provient de sources variées. Cette incertitude est due, en premier lieu, au fait que le monde réel possède une hétérogénéité et une complexité inhérentes qui dépendent fortement de l'échelle considérée, qu'elle soit spatiale, temporelle ou thématique (Dungan *et al.* 2002). Des modèles conceptuels et opérationnels guident l'acquisition, le traitement, l'analyse et l'interprétation des données géographiques afin de saisir cette complexité. Cependant, des méthodes comme l'abstraction et la généralisation imposent à chaque étape un niveau de filtrage de l'information (Zhang et Goodchild 2002) qui peut entraîner des différences entre les modèles obtenus sur le monde réel et le monde réel en tant que tel. Ainsi, l'incertitude peut être définie, dans son sens le plus large, comme la connaissance de la relation entre ces deux éléments (Goodchild et Case 2001).

Dans ce contexte, le cadre théorique de cette recherche est structuré en trois approches reliées au degré et au type d'incertitude de l'information forestière : (1) l'approche

écologique, qui présente les concepts fondamentaux nécessaires à la compréhension de la structure et de la dynamique du paysage forestier; (2) l'approche géomatique, qui aborde les aspects reliés aux caractéristiques des données sur la ressource forestière; et (3) l'approche analytique, qui décrit les méthodes les plus importantes qui permettent le traitement et l'analyse des données afin d'extraire l'information sur la forêt. Chacune de ces approches est influencée par la notion d'échelle. Donc, les implications du concept d'échelle seront d'abord présentées et, par la suite, ces approches seront détaillées de manière plus spécifique.

## **Le concept d'échelle**

Le problème de la relation des phénomènes du point de vue des échelles spatiales est central en biologie et dans toutes les sciences en général (Levin 1992). Les facteurs environnementaux, les réponses biologiques et les patrons d'organisation des écosystèmes dans les paysages changent pour toutes les échelles spatiales et temporelles (Delcourt et Delcourt 1988). La terminologie reliée à l'échelle est souvent ambiguë et les concepts varient selon le domaine (Goodchild et Quattrochi 1997). En écologie, l'échelle est définie comme la dimension spatiale ou temporelle d'un objet ou d'un processus (Csillag *et al.* 2000, Dungan *et al.* 2002, Turner *et al.* 2001). Dans le domaine de la géographie et de la géomatique, le concept d'échelle est cartographique, indiquant la relation entre la taille d'un objet sur une carte ou une photographie aérienne et le monde réel (Burrough et McDonnell 1998).

Pour revenir au domaine de l'écologie, les concepts d'échelle et de hiérarchie sont étroitement reliés (Allen et Hoekstra 1992, Bian 1997). Une hiérarchie est un système d'interconnexions dans lequel le niveau supérieur impose des contraintes aux niveaux inférieurs à différents degrés (Turner *et al.* 2001). Les niveaux d'organisation sont des systèmes hiérarchiques. En conséquence, lorsque l'intérêt porte sur un niveau particulier, les niveaux supérieur et inférieur doivent être considérés. Le niveau supérieur impose le contexte pour le niveau visé, alors que le niveau inférieur fournit les détails sur les phénomènes observés (Merriam 1987). Le lien entre l'échelle et la hiérarchie est évident du

moment que les changements d'échelle spatiale ou temporelle ont une influence sur l'importance relative des variables qui expliquent les phénomènes. Dans les paysages, les régimes de perturbations et les réponses du biote déterminent les patrons de la végétation à des échelles temporelles et spatiales différentes (Figure 2; Delcourt et Delcourt 1988). Il n'existe donc pas d'échelle unique pour étudier un paysage (Turner *et al.* 2001).

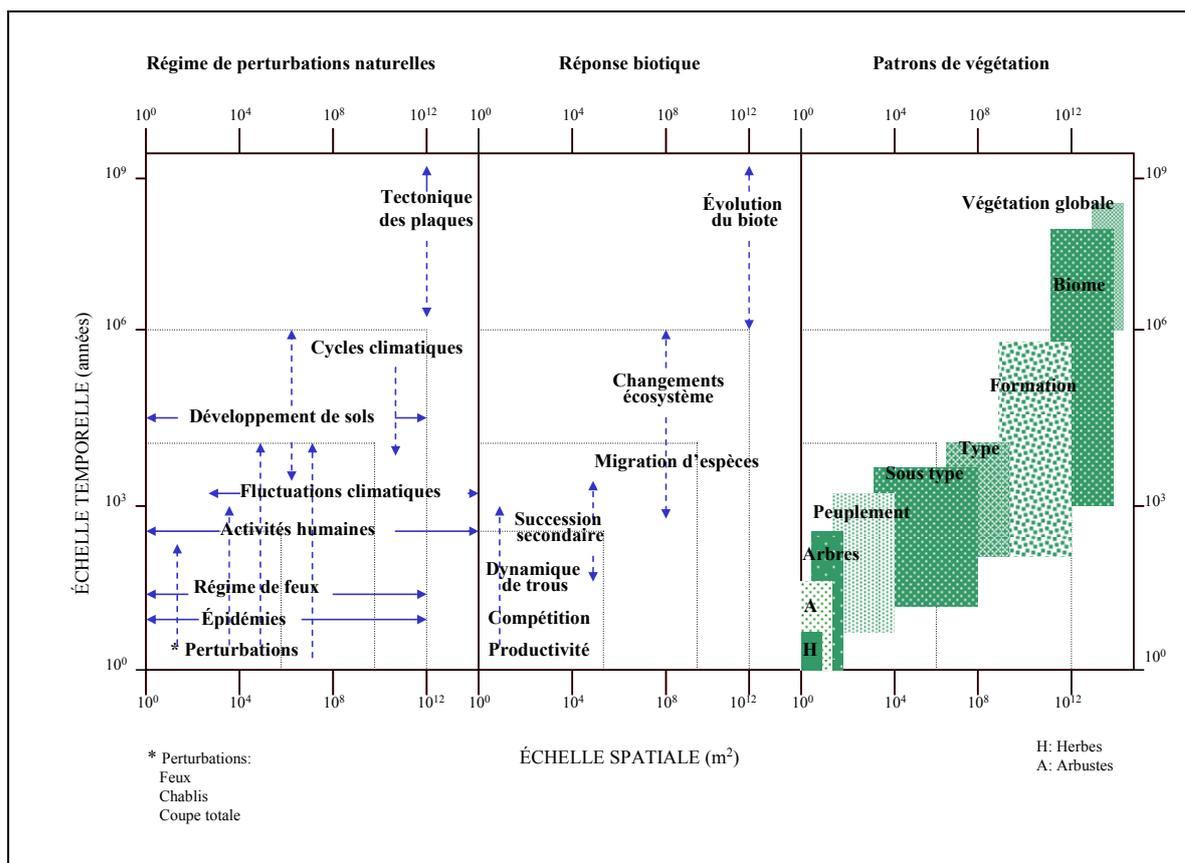


Figure 2. Les régimes de perturbations, les réponses du biote et les unités de végétation à différentes échelles spatiales et temporelles (Modifiée de Delcourt et Delcourt 1988).

Par ailleurs, la variation spatiale et temporelle d'un phénomène peut être détectée en fonction de l'échelle à laquelle le phénomène est échantillonné et de la taille de l'unité cartographique qui le représente (Gustafson 1998). Donc, du point de vue géomatique, l'échelle peut être décrite à travers le grain et l'étendue des observations. Le grain étant la taille de l'unité d'échantillonnage, il indique la résolution des données dans l'espace ou dans le temps. L'étendue, quant à elle, est la taille de la région d'étude ou de la période de

temps pour laquelle les données sont collectées (Csillag *et al.* 2000, Dungan *et al.* 2002, Gustafson 1998). Elle est indépendante du grain, bien qu'il existe une relation entre les deux qui détermine le volume des données (Goodchild et Hunter 1997, Turner *et al.* 2001).

Normalement, l'information échantillonnée à une certaine échelle peut être agrégée pour fournir de l'information à un niveau de détail inférieur (Bradshaw et Fortin 2000). Néanmoins, d'un point de vue analytique, plusieurs défis apparaissent face au changement d'échelle. Un des plus importants est celui de l'extrapolation, c'est-à-dire l'inférence des valeurs d'une variable au-delà de l'étendue des valeurs utilisées pour les estimations. Le transfert de l'information d'une échelle à une autre ou le transfert d'un ensemble de données à un autre ensemble pour une échelle similaire sont des exemples d'extrapolation. La normalisation par rapport à l'échelle et aux façons de collecter et d'organiser les données constitue un autre défi provoqué par le changement d'échelle (Milne 1998, Milne et Cohen 1999, Milne *et al.* 2002). Pour un même site, la différence de taille de la surface à laquelle correspondent les données ainsi que la différence de la fréquence temporelle et de la résolution des attributs posent des problèmes. Comme l'information écologique est rarement disponible aux échelles adéquates correspondant aux buts visés (Turner *et al.* 2001), des algorithmes et des méthodes d'extrapolation, de normalisation et de fusion de données ont été développés pour pouvoir changer les niveaux de détail.

## **Approche écologique**

Cette section présente d'abord une définition du paysage forestier, ensuite une caractérisation de ses composantes et pour finir, les principales causes des structures et configurations spatiales des paysages.

### **1. Le paysage forestier**

Un paysage est défini comme une aire qui est hétérogène du point de vue spatial, pour au moins une variable (Turner *et al.* 2001). Cette définition très générale présente les avantages de faire abstraction d'une échelle spatiale particulière ainsi que de pouvoir être

appliquée à tous les milieux et aux gradients d'écosystèmes qui ne sont pas récurrents dans l'espace (Turner *et al.* 2001). Elle permet également de considérer l'écologie du paysage comme un critère d'observation qui focalise sur l'hétérogénéité à n'importe quelle échelle spatiale (Cadenasso *et al.* 2003a). Dans ce cadre général, un paysage forestier est défini plus précisément comme un ensemble de peuplements forestiers qui sont caractérisés par une composition d'essences et qui possèdent des caractéristiques structurales différentes (comme l'âge et la biomasse des arbres; Perera et Euler 2000).

Plus spécifiquement, les paysages sont composés de deux types d'éléments : les parcelles et les frontières (Cadenasso *et al.* 2003b). Les parcelles peuvent être considérées comme des volumes qui se distinguent par leur composition, leur structure ou leur fonctionnement par rapport aux volumes adjacents à une échelle donnée. Les frontières sont, pour leur part, les zones entre deux parcelles adjacentes (Cadenasso *et al.* 2003b). Comme les parcelles, les frontières peuvent être détectées à plusieurs échelles spatiales et temporelles (Cadenasso *et al.* 2003a, Csillag *et al.* 2001, Fagan *et al.* 2003).

## **2. Les composantes du paysage forestier**

### **2.1. Les parcelles**

Les parcelles de la mosaïque sont associées à différents niveaux d'organisation. Pour l'activité forestière, les parcelles correspondent à une partie de la communauté forestière, soit le peuplement. Ce dernier est un concept opérationnel et administratif qui a longtemps constitué l'unité fondamentale de la gestion et de la planification forestière (Bérard et Côté 1996, Boyce 1995, Perera et Euler 2000). Un peuplement est donc défini comme un ensemble d'arbres ayant une uniformité jugée suffisante par rapport à sa composition floristique, sa structure, son âge, sa répartition dans l'espace, sa condition sanitaire, etc., pour se distinguer des peuplements voisins et pouvoir ainsi former une unité élémentaire sylvicole ou d'aménagement (Côté 2003). Les peuplements à différents stades de développement peuvent, d'une façon hiérarchique, être agrégés en structures à une plus

grande échelle en fonction des facteurs régionaux et ils peuvent être désagrégés à une plus petite échelle en fonction des structures intra-peuplements (Woodcock *et al.* 1994).

Ces structures à différentes échelles s'expliquent par le fait que les communautés forestières identifiées dans un paysage sont le résultat, d'une part, de l'adaptation des espèces végétales aux conditions locales et, d'autre part, des relations intra et interspécifiques qui se produisent à plusieurs échelles dans l'espace et dans le temps (Barnes *et al.* 1998, Milne 1991). À cette réalité écologique, s'ajoute le système de classification utilisé et la capacité des êtres humains à identifier les différents types forestiers. Toutefois, même si les communautés sont identifiées par les espèces d'arbres dominants, l'homogénéité associée à l'idée de communauté n'est jamais rencontrée sur le terrain. Malgré l'origine et l'évolution commune de la végétation, un peuplement est sujet à une variation spatiale qui peut rendre la distinction entre peuplements voisins difficile (Johnston 1999, Lowell 1999).

Cette variation intra-peuplement est la résultante de la réponse de la végétation à l'hétérogénéité environnementale locale. D'un point de vue général, « l'espace de croissance » est défini comme l'ensemble des facteurs nécessaires pour la croissance d'un arbre (Oliver et Larson 1996). Il présente une variation spatiale et temporelle et chaque espèce l'exploite différemment. Dans cet espace de croissance, la compétition intra et interspécifique est l'interaction primaire qui détermine la structure et la composition d'un peuplement (Barnes *et al.* 1998, Oliver et Larson 1996). Une fois que les plantes occupent l'espace de croissance disponible, la compétition pour étendre cet espace ou pour le maintenir permet à l'espèce ou à l'individu qui a un avantage particulier de se développer aux dépens des autres. Par conséquent, les arbres présentent différentes hauteurs de cimes et les peuplements multi-spécifiques, qui comportent des différences de taille maximale à atteindre et de tolérance à l'ombre, permettent le développement de strates (Oliver et Larson 1996). La compétition existe également dans le sous-bois. Certaines espèces d'arbres peuvent faire tout leur cycle de vie en sous-bois et d'autres, lorsque la chance se présente, occupent le couvert forestier. Les espèces « tolérantes » sont celles qui peuvent survivre et se développer dans le sous-bois, tandis que celles qui ne peuvent pas le faire

sont dites « intolérantes » aux conditions de lumière réduite en sous-bois (Barnes *et al.* 1998, Kimmins 1997). Cette caractéristique est importante pour comprendre l'évolution de la composition et de la structure d'un peuplement à la suite d'une perturbation.

Ces changements de structure et de composition d'un peuplement à travers le temps sont le résultat des interactions entre les arbres (Oliver et Larson 1996). Lorsqu'une perturbation majeure détruit la végétation existante, l'évolution du peuplement peut être divisée en ces quatre étapes (Barnes *et al.* 1998, Oliver et Larson 1996):

1. La mise en place du peuplement. La perturbation libère des espaces de croissance. Les plantes se régénèrent à partir des semences dispersées ou de manière végétative, tandis que la croissance des plantes opprimées s'accélère. L'invasion de l'espace se poursuit jusqu'à ce que les ressources deviennent limitées;
2. L'exclusion de tiges. Tout l'espace est utilisé, la futaie est fermée et la régénération des arbres est limitée ou arrêtée. Les arbres qui survivent sont de tailles et de diamètres différents et une ou deux espèces deviennent dominantes. Les arbres d'espèces tolérantes peuvent s'installer en sous-bois;
3. Régénération du sous-bois. La futaie commence à s'ouvrir dû aux trouées qui sont générées et il y a plus de lumière qui atteint le sous-bois. La régénération des arbres débute, tandis que les arbres tolérants commencent à prendre de l'expansion;
4. État stable. L'âge, la hauteur et la biomasse des arbres de la futaie augmentent. L'envahissement de l'espace et le processus de régénération se poursuivent. La taille des arbres tolérants augmente tant que les conditions du sous-bois le permettent. La sénescence, les perturbations et les maladies causent le dépérissement des cimes ou la mort d'arbres entiers. Des îlots d'arbres d'âge relativement uniforme se développent en lien avec les trouées.

La durée de chacune de ces étapes est extrêmement variable et dans certains cas, les perturbations peuvent interrompre la succession avant que l'étape finale ne soit atteinte. La composition spécifique de la première étape dépend fortement du type de perturbation, qui

détermine, en association avec d'autres facteurs comme les sols et le climat, les espèces qui auront des caractéristiques avantageuses pour se rendre à la prochaine étape (Oliver et Larson 1996).

## **2.2. Les frontières**

Les frontières sont des aires de transition, de contact ou de séparation entre les éléments contrastants d'une mosaïque ou d'un paysage (Cadenasso *et al.* 2003b, di Castri et Hansen 1992, Fortin *et al.* 2000, Gosz 1993, Risser 1995). Les frontières écologiques sont devenues l'objet d'un intérêt scientifique croissant, car elles régulent les flux de matière, d'organismes, d'énergie et d'information dans les paysages; mais leur importance dans cette régulation est souvent disproportionnée par rapport à l'espace qu'elles occupent (Cadenasso *et al.* 2003a). Cadenasso *et al.* (2003b) suggèrent six caractéristiques générales des frontières :

1. Les frontières peuvent présenter des caractéristiques communes avec les parcelles qu'elles séparent, mais elles peuvent aussi être complètement différentes;
2. Étant donné que les parcelles adjacentes à une frontière sont identifiées par une caractéristique particulière, le gradient de cette caractéristique est plus prononcé au sein de la frontière que dans chacune des parcelles voisines;
3. Les frontières peuvent être larges ou étroites, selon la magnitude du gradient entre les parcelles;
4. La magnitude et la localisation d'une frontière définie pour une caractéristique peut différer d'une autre frontière définie pour une caractéristique différente. Fortin *et al.* (2000) font la distinction entre les frontières environnementales et biotiques. Les premières correspondent aux changements physiques, alors que les deuxièmes correspondent aux réponses des espèces face aux changements environnementaux, aux interactions entre les espèces ou les deux;

5. La fonction d'une frontière est déterminée par l'organisme, le matériel, l'énergie, l'information ou le processus, qui peuvent tous être affectés par le gradient de la frontière. Les frontières peuvent servir d'habitat, de filtre ou de corridor selon la espèce (Forman 1995);

6. Les frontières sont tridimensionnelles, tout comme les parcelles. Forman (1995) reconnaît trois dimensions qui affectent les flux de matériaux, d'énergie et d'organismes : la largeur (entre la limite et l'intérieur de la zone de frontière), la hauteur (la hauteur et la stratification de la végétation) et la longueur (qui inclut les courbes et les lobes).

Les frontières peuvent également être classifiées en fonction de différents critères. Selon Strayer *et al.* (2003), quatre classes de critères peuvent être considérées : l'origine et le maintien de la frontière, sa structure spatiale, sa dynamique temporelle et sa fonction. Les trois premières classes sont particulièrement intéressantes dans le cadre de cette recherche.

Avant de considérer les frontières comme telles, mentionnons une première distinction qui existe en fonction de leur nature même. Elles peuvent être réelles ou résulter d'une conceptualisation. Les premières sont désignées comme tangibles ou « *bona fide* » alors que les deuxièmes sont désignées comme théoriques ou « *fiat* » (Jacquez *et al.* 2000, Smith 1995, Smith et Mark 1998, Strayer *et al.* 2003).

Donc, selon leur origine, les frontières peuvent être naturelles ou anthropiques et présenter ainsi des caractéristiques structurales très différentes (Forman 1995, Turner *et al.* 2001). L'origine et le maintien d'une frontière peuvent être endogènes ou exogènes, selon que les processus agissent à l'intérieur ou à l'extérieur du système. Des combinaisons de facteurs endogènes et exogènes y sont fréquentes. Le fait qu'un facteur soit considéré exogène ou endogène dépend toujours du système à l'étude (Strayer *et al.* 2003).

Les frontières possèdent également des propriétés structurales définies par les échelles spatio-temporelles et par le degré d'interactions entre les systèmes adjacents (Fortin *et al.* 2000). La taille, l'étendue, la forme ou la géométrie, le degré de netteté, la perméabilité et le contraste entre les parcelles adjacentes sont, entre autres, des caractéristiques structurales

importantes (Forman 1995, Strayer *et al.* 2003). Plusieurs des attributs comme la taille, la forme, l'étendue et la netteté sont dépendants de l'échelle considérée (Fagan *et al.* 2003, Fortin *et al.* 2000).

Par rapport à leur forme, les frontières naturelles sont généralement incurvées alors que les frontières d'origine anthropique tendent à être droites (Forman 1995). Par rapport à la netteté, le changement de conditions à travers la frontière peut être abrupt, ce qui crée une frontière nette, ou il peut être graduel. Dans le contexte écologique, les frontières nettes sont désignées comme des bordures alors que celles qui sont graduelles se nomment écotones (Fagan *et al.* 2003, Fortin *et al.* 2000). La magnitude du changement à travers la frontière, c'est-à-dire le contraste entre les parcelles adjacentes, dépend du processus étudié. Il s'agit d'un élément clé pour l'identification d'une frontière. Certaines méthodes ont été développées afin de quantifier ce contraste (Fagan *et al.* 2003, Fortin *et al.* 2000, Jacquez *et al.* 2000). Ces dernières seront présentées plus loin dans l'approche analytique. La combinaison d'attributs comme la forme et le contraste entre les parcelles adjacentes permet de reconnaître des frontières franches ou graduelles (Forman 1995). Les premières sont rectilignes et le contraste est élevé, alors que les deuxièmes sont plutôt sinueuses et le contraste est faible. Ces caractéristiques structurales affectent le fonctionnement des frontières et les flux qu'elles contrôlent.

De plus, les frontières sont dynamiques dû au fait que leur structure et leur localisation changent avec le temps (Fagan *et al.* 2003, Strayer *et al.* 2003). Une parcelle où la végétation est persistante tend vers l'expansion alors qu'une parcelle créée à la suite d'une perturbation tend vers la rétraction. Les frontières contrôlées par les variables biophysiques sont relativement stables (Forman 1995) mais elles peuvent aussi être éphémères ou persistantes à l'échelle temporelle considérée (Jordan 2002).

Quant à leur fonction, les frontières ont une perméabilité différentielle qui dépend de l'organisme, du matériel ou de l'énergie qui la traverse (Strayer *et al.* 2003). Elles peuvent être perméables pour certaines espèces ou matériaux et imperméables pour d'autres. En général, les espèces pénètrent mieux et plus loin à travers les frontières graduelles qu'à

travers des frontières rectilignes ou franches, qui peuvent diminuer les flux entre écosystèmes (Forman 1995).

Spécifiquement, les caractéristiques des frontières forestières varient selon les conditions de développement des peuplements, mais il demeure tout de même possible d'en identifier trois grands types (Oliver et Larson 1996) : (1) les frontières entre peuplements de même âge, (2) les frontières d'un peuplement adjacent à un terrain non forestier et (3) les frontières d'un peuplement jeune adjacent à un peuplement mature.

Les caractéristiques des frontières du premier type dépendent de celles des peuplements. Si ces derniers présentent la même composition spécifique, la frontière peut être le résultat d'une différence des taux de croissance ou d'espacement entre les arbres. Lorsque la composition spécifique est différente, la frontière peut présenter une stratification et les cimes de l'espèce dominante peuvent s'étendre sur la frontière, la surpasser et opprimer les espèces du peuplement voisin.

Dans le deuxième type de frontières, les cimes des arbres voisins aux terrains non forestiers se développent du côté lumineux. Cette différence de lumière par rapport à l'intérieur du peuplement permet aux arbres d'atteindre une taille élevée même s'ils sont plus rapprochés que ceux à l'intérieur du peuplement. La végétation non ligneuse peut perdre de la vigueur à cause de l'ombre, ce qui permet à la frontière de se déplacer vers l'espace ouvert. Les espèces occupant la frontière peuvent être différentes de celles à l'intérieur du peuplement si elles présentent des avantages compétitifs.

Le troisième type de frontières correspond à une situation dans laquelle une perturbation détruit la végétation d'une partie d'un peuplement ou lorsqu'un champ agricole est abandonné et qu'un nouveau peuplement se développe à côté d'un peuplement déjà établi. Dans le premier cas, les arbres nouvellement exposés peuvent subir des cassures ou tomber à cause du vent (Harper et Macdonald 2002). Les rejets de souches et les semences des arbres à proximité de la frontière permettent à une cohorte de s'installer. Les conditions d'ombre et d'occupation de l'espace de croissance changent progressivement selon la distance des jeunes plantes par rapport à la frontière. Le résultat est que la frontière présente

une bande d'arbres très rapprochés avec un gradient de présence d'espèces tolérantes qui augmente lorsqu'on approche du peuplement mature et un gradient de présence d'espèces non tolérantes et d'arbres plus grands qui augmente lorsqu'on approche du nouveau peuplement. Même lorsque le nouveau peuplement est mature, la frontière peut être reconnue par la présence d'arbres ou de branches mortes et par une différence dans l'espacement entre les arbres (Oliver et Larson 1996).

### **3. Les facteurs qui structurent les paysages forestiers**

Les paysages actuels sont le résultat de plusieurs processus et combinaisons de facteurs comme la variabilité des conditions abiotiques, les interactions biologiques, les patrons passés et actuels de l'utilisation du territoire, et la dynamique des perturbations et de successions écologiques (de Blois *et al.* 2001, Methven et Feunekes 1987, Turner *et al.* 2001). La variabilité à grande échelle de l'environnement abiotique impose des contraintes aux interactions biologiques et aux perturbations (Turner *et al.* 2001).

#### **3.1. Les facteurs abiotiques**

Les patrons des paysages sont le résultat, en partie, de la variabilité du climat et de la physiographie (Barnes *et al.* 1998, Perera *et al.* 2000, Turner *et al.* 2001). Le climat agit comme régulateur en distribuant l'énergie et l'eau. Les effets du climat sont modifiés par la physiographie (Kimmins 1997). Ensemble, le climat et la physiographie forment le modèle sur lequel les sols et le biote d'une région se développent (Rowe 1987, Turner *et al.* 2001).

La physiographie est la composante la plus stable des paysages et constitue donc une base solide pour la gestion des paysages forestiers (Barnes *et al.* 1998). Elle contrôle le régime climatique ainsi que le développement des sols et les processus pédologiques (Rowe 1987). La physiographie affecte la quantité et le flux de matériaux reçus par un écosystème, ainsi que les gradients gravitationnels qui guident les flux d'eau et d'air en plus d'opposer des barrières aux mouvements des matériaux. Les aspects physiographiques les plus importants à considérer pour les peuplements forestiers sont : l'élévation du terrain, sa forme, sa pente, ses dépôts de surface et sa position dans le paysage.

La physiographie, le climat, le matériel originel et la végétation déterminent les caractéristiques des sols à travers le temps. La texture et la structure des sols conditionnent, quant à eux, le drainage de l'eau et le contenu en nutriments. Parmi les caractéristiques qui affectent la croissance des arbres, la profondeur effective du sol et le drainage de l'eau et de l'air sont les plus importantes puisqu'elles affectent de façon directe et indirecte la disponibilité en nutriments et en eau pour les racines (Barnes *et al.* 1998, Rowe 1987).

### **3.2. Les interactions biologiques**

Les interactions entre les organismes peuvent produire une structure spatiale dans un espace complètement homogène. Par exemple, la compétition entre deux types de végétation pour l'espace peut provoquer la formation de frontières franches même si les conditions environnementales bénéfiques pour les deux types changent très peu à l'intérieur de la frontière (Turner *et al.* 2001). D'autres organismes définissent également le patron spatial, tels que les insectes dont les épidémies agissent comme une perturbation, ce qui crée de nouveaux peuplements (Turner *et al.* 2001).

### **3.3. L'utilisation du territoire**

Dès la préhistoire, l'influence humaine a transformé les paysages naturels en paysages culturels par les coupes, le feu, l'agriculture et la conversion des forêts en pâturages, dont les résultats demeurent encore visible dans les paysages actuels (Turner *et al.* 2001). Dans les régions gérées de façon intensive, l'histoire de l'utilisation du territoire s'avère plus importante que les perturbations naturelles pour déterminer les patrons de végétation (de Blois *et al.* 2001).

### **3.4. Les perturbations et la succession écologique**

Les perturbations et le développement subséquent de la végétation sont des facteurs clés dans la structure des paysages (Barnes *et al.* 1998, Turner *et al.* 2001). Les perturbations, des événements relativement discrets dans le temps (White et Pickett 1985), détruisent sélectivement les organismes et changent les conditions biophysiques locales. À long

terme, c'est l'ensemble des perturbations, souvent liées les unes aux autres, qui modèle les écosystèmes et les paysages (Barnes *et al.* 1998).

Ces perturbations peuvent être caractérisées à différentes échelles. Régionalement, le climat et la physiographie déterminent le type de perturbation propre à la région. À une échelle locale, les caractéristiques des écosystèmes décrits aux paragraphes précédents déterminent le type, la fréquence et la sévérité de la perturbation. Celles qui sont sporadiques et sévères peuvent causer des changements irréversibles dans l'écosystème, alors que celles qui sont régulières, avec des effets temporaires, permettent l'adaptation des composantes de l'écosystème. Ce dernier requiert finalement ces perturbations pour demeurer normal et résilient (Barnes *et al.* 1998). Souvent, l'interruption d'une perturbation est causée par une autre qui produit des changements irréversibles dans l'écosystème. Les perturbations les plus importantes qui affectent les paysages forestiers sont les feux, les chablis, les épidémies et les coupes forestières. Les paragraphes suivants décrivent de manière succincte les aspects les plus importants pour chacune de ces perturbations.

#### **3.4.1. Les feux**

Hormis les sites très humides, très chauds ou très froids, le feu agit dans presque tous les écosystèmes terrestres comme principal agent de leur structure (Kimmins 1997). Dans le contexte d'une communauté forestière établie, le feu interrompt le développement du peuplement existant. Il favorise plutôt l'établissement et la germination de semences en diminuant fortement la litière, ce qui réduit la quantité de sels minéraux dans la matière organique, et met généralement fin aux épidémies en tuant les arbres infectés. Le feu favorise les espèces qui dépendent de ce dernier pour se régénérer. Par conséquent, ce sont ces espèces qui seront présentes dans le peuplement même si elles ne sont pas des espèces particulièrement tolérantes. En effet, ces dernières ont plutôt tendance à être vulnérables au feu et elles occupent les sites qui y sont probablement les moins sujets. Durant les intervalles entre deux feux, ces espèces colonisent les aires adjacentes et peuvent éventuellement remplacer les espèces qui dépendent du feu en absence d'un brasier suffisamment intense pour les tuer (Barnes *et al.* 1998).

Le patron du feu dépend principalement des conditions du site et de la disponibilité de matériel combustible dans les écosystèmes qui composent le paysage. Un feu est décrit à partir de sa fréquence, de son intervalle de récurrence et de sa période de rotation, qui, ensemble, caractérisent le cycle du feu (Methven et Feunekes 1987). Les autres caractéristiques sont la taille, l'intensité et la sévérité du feu. La taille et le cycle sont des caractéristiques qui agissent plutôt à l'échelle du paysage, alors que l'intensité et la sévérité sont des attributs partiels du feu parce qu'elles dépendent de la végétation et qu'elles agissent à l'échelle locale (Methven et Feunekes 1987). L'intensité affecte les composantes aériennes de la végétation, alors que la sévérité, une mesure de l'impact du feu sur la communauté, affecte la partie souterraine. Ensemble, elles modifient la composition des espèces et des classes d'âge à l'intérieur du périmètre du feu. La taille influence l'échelle du patron de la végétation, alors que le cycle affecte la composition spécifique et la distribution des classes d'âge ou des stades de développement à travers le paysage. Un court cycle de feu favorise les espèces de régénération végétative et de courte vie alors qu'un cycle de feu long favorise les espèces longévives (Barnes *et al.* 1998).

### **3.4.2. Les chablis**

Le chablis correspond à la chute, au déracinement ou à la cassure d'arbres ou de groupes d'arbres à cause du vent, du poids de la neige ou de la glace et de l'âge des arbres (Côté 2003). Le vent, qui est un facteur particulièrement important sur le site à l'étude, affecte la structure des écosystèmes aux échelles régionale et locale (Barnes *et al.* 1998). Les vents locaux peuvent déraciner des arbres individuels ou des groupes d'arbres en générant des trouées dans le couvert forestier. Les rafales augmentent leur vitesse à mesure qu'elles entrent dans la forêt, tandis que les changements de végétation et d'altitude augmentent leur vitesse et leur turbulence (Kimmins 1997). La position du terrain dans le paysage, l'inclinaison et l'exposition d'une pente, le drainage du sol, le type de végétation, l'âge et la hauteur des arbres influencent la vulnérabilité d'un site aux dommages du vent (Tang *et al.* 1997). Ceux-ci tendent à être plus sévères à la lisière de la forêt, dans les chemins et dans les aires ouvertes.

### **3.4.3. Les épidémies**

Les épidémies peuvent jouer un rôle critique dans l'écologie des forêts boréales de l'Est du Canada (Déry *et al.* 2000). Elles constituent une perturbation très dispersée autant dans les événements catastrophiques que chroniques (Turner *et al.* 2001). En général, les épidémies tuent ou affaiblissent des arbres isolés ou des groupes d'arbres d'une espèce particulière. C'est un deuxième facteur qui produit des trouées dans la forêt en favorisant le développement des espèces tolérantes. Les épidémies interagissent souvent avec les feux et les vents, les arbres affaiblis étant plus vulnérables aux vents, alors que les arbres morts constituent le matériel combustible pour les feux (Barnes *et al.* 1998).

### **3.4.4. Les coupes forestières**

Les effets des coupes forestières dépendent de leur intensité et du type qui est appliqué. Par exemple, les coupes partielles ou légères favorisent les espèces tolérantes, en particulier celles qui sont déjà établies en sous-bois (Barnes *et al.* 1998). Les coupes sélectives changent également la composition du peuplement en récoltant seulement une espèce précise. Les récoltes intenses mais partielles ou en petits groupes favorisent, quant à elles, les espèces semi-tolérantes. Pour leur part, les coupes totales favorisent l'invasion d'espèces pionnières et changent les conditions du sol (Kimmins 1997). Les patrons de coupes totales peuvent affecter d'autres perturbations, comme le vent. Par exemple, la longueur du périmètre d'une coupe totale détermine la superficie de forêt qui sera exposée aux dommages liés aux vents (Tang *et al.* 1997).

### **3.4.5. La succession écologique**

Toutes les composantes des écosystèmes changent dans le temps. Ces changements peuvent être graduels ou rapides, selon les perturbations. La succession forestière, caractérisée par le changement séquentiel dans la structure, le type et l'abondance relative des espèces, est une manifestation du changement de l'écosystème (Kimmins 1997, Shugart 1984). À l'échelle du paysage, la dynamique est plus ou moins visible selon la résilience ou l'inertie des composantes qui sont perturbées (Barnes *et al.* 1998). Les séquences possibles

d'espèces en interaction, les régimes de perturbations et les processus sont contrôlés par le paysage et les écosystèmes, respectivement à l'échelle régionale et locale.

## **Approche géomatique**

La complexité du monde réel, comme on l'a constaté pour la forêt, requiert de l'abstraction, de la simplification et de la classification pour faciliter les analyses et la prise de décisions (Zhang et Goodchild 2002). Le processus d'abstraction et de généralisation des formes réelles, afin de les représenter en forme numérique, produit un modèle conceptuel du monde réel (Goodchild 1989). Il existe inévitablement une différence entre les phénomènes du monde réel et les données qui les représentent. Selon les auteurs, cette différence est l'incertitude ou l'erreur des données et peut être analysée à travers un ensemble de paramètres qui définissent la qualité de la représentation (Longley *et al.* 1999a).

Dans le cadre de ce travail, les modèles conceptuels sur l'espace, qui sont implicites dans les sources de données écologiques et forestières, sont d'abord présentés dans l'approche géomatique. Ensuite, les paramètres qui définissent l'incertitude et la qualité des données sont décrits. Par la suite, les sources de données forestières, la carte forestière utilisée au Québec et les données de terrain sont présentées. Finalement, les approches principales pour le traitement de l'incertitude spatiale sont examinées.

### **1. Les modèles conceptuels de l'espace géographique**

Les « objets » et les « champs » sont deux formes pour conceptualiser l'espace et les phénomènes géographiques (Burrough 1996, Burrough et McDonnell 1998). Ils constituent actuellement un des paradigmes de la science de l'information spatiale (Fisher 1997).

D'une part, les champs permettent de représenter des phénomènes continus, multivariés et dynamiques (Zhang et Goodchild 2002). Un champ représente l'espace en termes de coordonnées cartésiennes continues (avec trois ou quatre dimensions) et on suppose que les attributs y varient de façon continue (Burrough et McDonnell 1998). Selon le schéma d'échantillonnage, les champs peuvent être modélisés à partir de points réguliers ou

irréguliers, de lignes, de polygones, de grilles ou de TINs, pour des variables qualitatives ou quantitatives (Zhang et Goodchild 2002).

D'autre part, l'espace est conceptualisé sous forme d'objets, qui sont identifiés par des attributs non spatiaux, délimités par des frontières nettes et localisées de façon précise (Burrough et McDonnell 1998, Zhang et Goodchild 2002). Ces frontières sont importantes pour classer l'espace en catégories (Smith et Mark 1998). On suppose en effet que les valeurs des attributs sont constantes pour l'objet en entier (Burrough 1996).

Pour les phénomènes naturels, multivariés et dont la variation spatiale est multi-échelle, les modèles d'objets et de champs sont limités (Burrough 1996). On s'entend généralement pour représenter par des objets les phénomènes statiques qui sont qualitativement complexes et dont les éléments externes sont reconnaissables. Quant aux phénomènes mesurés par une seule variable quantitative, ils sont représentés par des champs (Burrough 1996, Zhang et Goodchild 2002). Cependant, les deux modèles peuvent coexister et même se compléter. Les phénomènes continus présentent des discontinuités qui peuvent être extraites et traitées comme des objets, tandis que des champs peuvent être définis à l'intérieur des objets (Jacquez *et al.* 2000, Zhang et Goodchild 2002).

## **2. Qualité et incertitude des données géographiques**

L'incertitude peut être définie comme l'état des connaissances sur la relation entre le monde réel et une formulation sur le monde (Goodchild et Case 2001). Les termes incertitude et qualité sont étroitement liés. L'incertitude est souvent restreinte au concept d'erreur (Heuvelink 1999), alors que la qualité possède un sens plus large qui est approprié dans un contexte de production/utilisation des données. Le producteur, qui peut être un chercheur, documente la qualité des données et l'utilisateur évalue cette qualité en fonction de son adéquation pour une application spécifique (Veregin 1999). Les composantes de la qualité sont aussi utilisées pour décrire l'incertitude.

Les observations géographiques, dont font partie les données forestières et écologiques, décrivent les phénomènes à partir des dimensions spatiale, temporelle et thématique. Les

composantes de la qualité s'appliquent à chacune de ces dimensions. Celles qui sont habituellement considérées sont l'exactitude, la précision, l'exhaustivité et la consistance logique (Veregin 1999).

L'exactitude mesure la ressemblance entre la valeur mesurée et la vraie valeur d'une variable. Le concept inverse est celui de l'erreur ou de l'incertitude qui, dans son sens le plus large, inclut la signification statistique de l'erreur (Heuvelink 1999). Comme la vraie valeur n'est pas toujours mesurable, celle qui est mesurée est habituellement comparée avec une référence établie selon les spécifications définies dans le modèle utilisé. L'exactitude est donc une mesure relative et non absolue (Vauglin 1998, Veregin 1999).

Pour sa part, la précision est utilisée selon deux sens, statistique et numérique. Au sens statistique, elle fait référence à la dispersion des observations par rapport à leur moyenne (Mowrer 1998). Au sens numérique, la précision ou résolution indique le nombre de détails qui peuvent être discernés. Du point de vue spatial, la précision implique la taille des éléments qui peuvent être discernés, alors que du point de vue thématique, elle implique le raffinement des catégories ou la résolution de l'instrument de mesure (Veregin 1999).

La consistance logique est l'absence de contradictions apparentes dans les données et est considérée comme une mesure de la validité interne d'une base de données (Veregin 1999). L'exhaustivité qualifie la relation entre les objets représentés dans la base de données et « l'univers abstrait » de tous les objets. Cet univers est défini par le niveau d'abstraction et de généralisation désiré dans les spécifications de la base de données. Le concept d'exhaustivité peut être appliqué aux données, au modèle, aux attributs des objets ou aux valeurs des attributs (Veregin 1999).

Les sources d'incertitude sont variées. La première est la définition même des objets. La plupart des phénomènes naturels possèdent simultanément des caractéristiques de champs et d'objets. Même si un consensus existe sur le concept fondamental d'un objet, si les frontières spatiales et thématiques ne peuvent être localisées avec exactitude, l'objet est « mal défini » (Burrough 1996, Fisher 1999). Certains autres concepts sont, pour leur part, vagues ou ambigus (Fisher 1999). D'autres sources d'incertitude sont associées à

l'acquisition, au traitement, à la représentation des données (application de modèles spatiaux et non spatiaux, changements d'échelle, agrégations, changements de format, de structure ou de projection) et à la variabilité naturelle des phénomènes géographiques (Sklar et Hunsaker 2001, Stine et Hunsaker 2001). L'évaluation de l'incertitude comprend donc le développement de procédures pour estimer l'exactitude, la précision, le biais, la variance, la probabilité d'erreur de mesure d'un échantillonnage ou d'un calibrage et leur propagation aux prédictions portant sur les phénomènes naturels (Mowrer 1998).

### **3. Les données forestières**

#### **3.1. La carte forestière traditionnelle**

Les photographies aériennes et les images de télédétection constituent les données brutes à partir desquelles la carte forestière est élaborée. Pour les photographies aériennes, qui sont encore très utilisées pour les cartes forestières au Québec, il existe des standards établis pour la fenêtre spectrale, qui couvre le spectre visible et infrarouge, ainsi que pour des échelles comprises entre 1/5000 et 1/40000 (Bérard et Côté 1996).

La photo-interprétation est le processus d'extraction de l'information contenue dans les photographies aériennes. Le photo-interprète segmente la photographie en appliquant une stratification forestière. Cette dernière est l'ensemble des règles et des codes basés sur des critères reliés à l'estimation du volume ligneux, qui sont interprétables dans une photographie tels que la composition d'essences, la classe de densité, la classe de hauteur, la classe d'âge des arbres, la perturbation d'origine subie par le peuplement, la classe de pente, la défoliation et le dépôt de surface (Ministère des Ressources naturelles du Québec 1995). Pour le Québec, une taille d'unité cartographique minimale est fixée à huit hectares pour les peuplements productifs (Ministère des Ressources naturelles du Québec 1995).

La photo-interprétation au Québec implique deux étapes : d'abord le photo-interprète se forme une représentation mentale de la réalité forestière en observant la photographie aérienne et par la suite, il transfère cette représentation mentale sur une carte en utilisant des lignes et des polygones. Pour ce faire, il prend en compte la perception des variations de la

couleur, de l'intensité et de la texture de la photo (De Groeve 1999). Un objet est identifié à partir de la minimisation de la variation des indices visuels à l'intérieur de l'objet et de la maximisation de la variation entre les objets identifiés. Ceci signifie que même si les frontières sont floues, l'objet peut être identifié (Lowell 1999). Concrètement, l'essence et l'âge sont déterminés à partir de la couleur, de la texture et de l'expérience du photo-interprète, qui est complétée par des visites sur le terrain. La hauteur est déterminée par stéréoscopie, alors que la densité est estimée en utilisant des images de référence (De Groeve 1999, Johnston 1999). Un nombre limité de points de contrôle sont utilisés afin d'ajuster la photo-interprétation préliminaire. Les strates cartographiques ambiguës sont examinées et une vérification de la qualité de la carte est réalisée. L'interprétation est transférée sur une carte topographique de base. Si le terrain est accidenté, la monorestitution différentielle ou la stéréorestitution sont utilisées afin de réduire l'erreur sur la carte finale (De Groeve 1999, Johnston 1999).

Comme pour les données géographiques, les sources d'incertitude sont multiples. Une première source est la définition du peuplement qui, comme celle de la communauté écologique, est un concept complexe qui produit des objets « mal définis » dans une base de données (Fisher 1999). Une autre source est le modèle conceptuel utilisé. En effet, un ensemble de variables qualitatives détermine un champ discret avec des frontières nettes et abruptes entre les classes (Zhang et Goodchild 2002). Cette différence de netteté entre le modèle et la réalité amène à considérer les objets ciblés comme flous ou indéterminés (Smith 2001). Une autre source d'incertitude est la délimitation des peuplements et l'interprétation des paramètres qui sont des tâches cognitives très subjectives. Les frontières entre les peuplements sont déterminées par une combinaison de procédures au cours desquelles ces frontières et les attributs des peuplements sont importants (Lowell *et al.* 1996, Thierry et Lowell 2001). Par conséquent, les frontières et les attributs des polygones ont un degré d'incertitude élevé (Edwards et Lowell 1996), à tel point que si la même photographie est interprétée par différents interprètes, chacun en produira une version différente (Aubert 1995). La restitution du résultat de la photo-interprétation constitue également une autre source d'erreurs.

Malgré toutes ces sources d'incertitude, les cartes forestières ne présentent de l'information sur l'incertitude qu'à une échelle globale. Par exemple, les cartes forestières canadiennes ont une exactitude positionnelle des objets de la carte équivalente à 10-25 m (De Groeve 1999, Leckie et Gillis 1995).

Dans ce contexte, la cartographie forestière élaborée au Québec répond très bien aux exigences du ministère responsable de la gestion des forêts, dont le but est de connaître les stocks forestiers et leur dynamique et ce, afin d'en assurer la gestion et la conservation à long terme. Les programmes d'inventaires des ressources forestières sont prévus pour une durée variant entre dix et quinze ans. La mise à jour des cartes des peuplements forestiers consiste à refaire une photo-interprétation complète de tout le territoire. Pour cette opération, les photo-interprétations antérieures ne sont pas réutilisées dû, entre autres, aux changements des normes d'inventaires forestiers (Aubert 1995).

### **3.2. Les données de terrain**

Au Québec, les données de terrain les plus courantes sont obtenues à partir de l'inventaire forestier planifié sur la carte forestière. Différentes méthodes d'échantillonnage sont utilisées, qu'elles soient aléatoires simples, systématiques ou stratifiées (Bérard et Côté 1996). La dernière méthode est utilisée au Québec et les unités d'échantillonnage sont distribuées de façon conditionnelle dans les strates afin d'assurer une précision adéquate dans les estimations globales et dans certaines strates considérées comme particulièrement importantes.

Les unités d'échantillonnage peuvent être de surface définie ou indéfinie. Pour le premier type, les formes sont variées et celles qui sont circulaires sont les plus intéressantes, car elles sont facile à installer et qu'elles présentent un périmètre minimal (Bérard et Côté 1996). La dimension de la placette résulte d'un compromis entre le coût d'installation et la précision souhaitée dans les mesures. Certaines variables sont mesurées directement et d'autres sont estimées visuellement.

Les données de terrain saisies dans les placettes souvent contiennent à la fois des erreurs de type aléatoire et de type systématique. Les sources de ces erreurs correspondent à la saisie des données par la personne sur le terrain, à la transmission de ces données entre différentes personnes, à la nature des variables (par exemple, la hauteur des arbres est difficile à évaluer lorsque le peuplement est dense) et aux instruments de mesure. Le sondage est donc vérifié au moyen des re-échantillonnages aléatoires (Bérard et Côté 1996). Les unités d'échantillonnage présentent aussi de l'incertitude positionnelle, qui n'est habituellement pas estimée.

Les estimations du volume ligneux pour chaque placette-échantillon sont faites au moyen des données de terrain et des tarifs de cubage des essences qui sont élaborés pour le territoire échantillonné (Bérard et Côté 1996). Les strates insuffisamment échantillonnées sont regroupées en suivant des critères tels que le type forestier, les perturbations et les essences présentes (Johnston 1999).

#### **4. L'incertitude spatiale**

L'incertitude est endémique à l'information géographique. Dans un premier temps, des efforts ont été dirigés vers la description de l'incertitude des données, mais la perception de la variabilité spatiale des erreurs a permis de souligner l'importance de la modélisation de l'incertitude. Cette modélisation requiert des connaissances sur les distributions spatiales des erreurs qui sont variables et corrélées (Zhang et Goodchild 2002). La modélisation de l'incertitude vise à formuler des stratégies pour la gérer, pour prendre des décisions en présence des erreurs et pour réduire ou éliminer les erreurs dans les produits résultants. Les méthodes utilisées pour la description et la modélisation varient selon le modèle conceptuel de l'espace employé. Elles correspondent soit à des approches statistiques, probabilistes ou des simulations stochastiques, soit à d'autres approches comme les logiques de raisonnement spatial ou les techniques visuelles.

#### 4.1. Description de l'incertitude spatiale

Le modèle conceptuel d'objets permet souvent de séparer la description de l'incertitude en deux aspects : spatial et thématique. En ce qui concerne l'incertitude positionnelle, les méthodes de description de l'incertitude sont basées sur l'estimation de l'erreur quadratique moyenne (RMSE; Beard 2001). Les bandes d'erreur epsilon ont été introduites par Chrisman en 1982 pour représenter l'incertitude positionnelle des frontières des cartes qualitatives. Ce sont des zones qui contiennent la vraie position de la frontière avec une certaine probabilité et où les attributs sont considérés exacts dans les polygones et incertains dans la bande d'erreur. Différents auteurs ont travaillé sur ces bandes et ont proposé l'utilisation d'une distribution gaussienne de l'incertitude sur la bande (Mark et Csillag 1989). L'interprétation déterministe de la bande epsilon est critiquée parce que, premièrement, elle ne fournit pas un modèle de la distribution de l'erreur sur la bande et, deuxièmement, parce qu'elle postule que la vraie frontière est située à l'intérieur de la région epsilon (Alesheikh *et al.* 1998). Lowell (1994) a démontré que dans le cas des cartes forestières, les bandes peuvent être symétriques mais qu'elles peuvent aussi varier en largeur selon les frontières.

L'exactitude de la mesure des variables continues est, elle aussi, représentée par l'erreur quadratique moyenne (RMSE) à l'aide de la répétition des mesures ou de la comparaison de la moyenne et de l'écart-type d'un échantillon avec des valeurs de référence. Les méthodes géostatistiques sont aussi très utilisées, notamment le krigeage, qui fournit une surface de la variance des estimations au moment de l'interpolation (Zhang et Goodchild 2002).

L'erreur thématique d'une carte issue d'une classification est estimée, pour sa part, avec la matrice de confusion qui compare les résultats de la classification à une source de référence (McGwire et Fisher 2001). Des mesures comme le pourcentage correctement classifié, l'exactitude de l'utilisateur (l'erreur de commission), l'exactitude du producteur (l'erreur d'omission) et le coefficient kappa sont des dérivées de la matrice de confusion (Stehman 1998). Elles supposent un niveau d'erreur constant pour toutes les catégories et pour toute la carte.

## 4.2. Modélisation de l'incertitude spatiale

La modélisation a pour but de représenter la variation spatiale de l'incertitude ainsi que sa propagation à travers les opérations et les traitements jusqu'au produit final. Les approches statistiques conventionnelles qui supposent que les observations sont indépendantes et aléatoires, ne sont pas valides pour traiter les données spatiales, parce que la plupart présentent une dépendance spatiale (Zhang et Goodchild 2002).

La géostatistique est souvent utilisée pour modéliser l'erreur des variables continues et leur propagation. Pour ce faire, l'incertitude est considérée comme une réponse stochastique (Fisher 1998, Heuvelink 1993, 1996, Journel 1996, Kyriakidis 2001, Moeur et Hershey 1998, Shortridge 2001). Un ensemble de valeurs d'une variable continue est modélisé comme une fonction aléatoire, avec une distribution spécifique de probabilités qui reflète l'incertitude sur la valeur actuelle. La stationnarité de la variable est la base de toute la probabilité et de tous les algorithmes de prédiction. Cette stationnarité est une propriété du modèle et non des observations (Heuvelink 1998, Zhang et Goodchild 2002). Un ensemble de réalisations équiprobables de cartes qui respectent la structure spatiale des données originales est créé au moyen de simulations stochastiques, conditionnelles ou non. La distribution résultante des valeurs de réponse représente alors l'incertitude. Le modèle peut être mis à jour en y ajoutant de nouvelles données ou en incorporant d'autres types d'informations (Heuvelink 1999).

Quant à la modélisation de l'incertitude des variables qualitatives, ces dernières sont considérées comme des fonctions de la localisation avec une seule valeur (Goodchild *et al.* 1992). Cette approche permet de modéliser la variabilité spatiale de l'incertitude qualitative avec celle des catégories elles-mêmes. Les méthodes géostatistiques, particulièrement l'indicateur de krigeage, sont utilisées pour estimer, à partir d'un ensemble d'échantillons, les probabilités d'occurrence d'une classe pour des endroits non échantillonnés (Zhang et Goodchild 2002). D'autres approches utilisent, pour leur part, les simulations stochastiques en tenant compte des probabilités des catégories et du degré de dépendance spatiale. Parmi celles-ci, l'indicateur de simulation séquentiel, avec des formes conditionnelles et non conditionnelles, est un des modèles les plus utilisés pour la modélisation de l'incertitude

(Zhang et Goodchild 2002). Cependant, l'exactitude des réalisations dépend non seulement de la méthode d'estimation mais aussi de l'échantillonnage réalisé. Les méthodes probabilistes de classification fonctionnent pour l'incertitude de source aléatoire. Les incertitudes dues aux inexactitudes et aux imprécisions des définitions des catégories sont, pour leur part, mieux traitées au moyen d'autres approches. Les théories des ensembles flous (Altman 1994, Lowell 1994, McGwire et Fisher 2001, Zhu 2001), de l'entropie (Leung *et al.* 1992), des ensembles grossiers ou des ensembles multiples (Fisher 2001) permettent de classer les échantillons lorsque les catégories sont imprécises.

En ce qui concerne l'incertitude positionnelle des objets, Hunter et Goodchild (1996) ont développé un modèle de distorsion probabiliste des éléments à partir des champs d'erreur positionnelle indépendants en  $x$  et en  $y$ . Ceux-ci sont superposés avec les objets et des modifications aux nœuds et aux vertex sont appliquées afin d'obtenir de nouvelles versions équiprobables aux données originales. La variation des résultats permet d'évaluer l'incertitude des produits résultants (Hunter *et al.* 1996, 1998). Kiiveri (1997) a proposé, pour sa part, un modèle probabiliste d'incertitude positionnelle de frontières basé sur les déplacements de chaque point, en utilisant des fonctions dérivées des séries de Fourier. Le modèle définit une frontière floue de largeur variable autour de la frontière observée et les probabilités d'incertitude sont calculées pour les opérations d'union et d'intersection des cartes. Edwards et Lowell (1996) ont proposé, quant à eux, un modèle de variation floue de la largeur des bandes d'erreur epsilon pour les frontières forestières. De plus, Vauglin (1998) a proposé une approche géostatistique pour modéliser l'incertitude des éléments vectoriels en utilisant la distance de Hausdorff pour mesurer les incertitudes positionnelles relatives des lignes.

#### **4.3. Modèles d'incertitude spatiale des cartes forestières**

Le modèle d'incertitude à composantes positionnelles et thématiques est adéquat lorsque les frontières des objets sont nettes. Dans le cas des peuplements avec des frontières peu ou mal définies, une troisième composante doit être considérée, soit l'incertitude sur l'abstraction qui est rattachée à l'existence de la frontière (De Groeve 1999). La procédure

d'élaboration de la carte est une des sources de cette incertitude où certaines frontières sont clairement visibles sur la photo, mais où d'autres doivent être ajoutées dans les aires de changement graduel. Une autre source d'incertitude est l'agrégation, processus subjectif accompli par chaque interprète et qui consiste à regrouper de petites parcelles de forêt en polygones avec une classe prédominante (Edwards 1994).

Une procédure générale d'évaluation de l'incertitude consiste à superposer différentes réalisations cartographiques d'un même territoire, à faire coïncider les objets équivalents, à mesurer les différences entre les réalisations et à dériver une carte d'incertitude à partir de la cohérence spatiale des objets. Plusieurs auteurs ont suivi cette procédure, en situations contrôlées, pour obtenir des bandes d'erreur epsilon au moyen d'images ou de cartes vectorielles associées aux « vérités de terrain » (Aubert 1995, Hunter *et al.* 1996). L'évaluation de l'incertitude spatiale sans cette vérité de terrain réduit très fortement le détail de l'analyse. Legros *et al.* (1996) ont essayé de relier des causes cognitives à l'incertitude en utilisant des réalisations multiples mais ils n'ont pas réussi à obtenir de conclusions claires.

De Groeve (1999) a proposé une méthode qui tient compte de l'aspect cognitif impliqué dans la photo-interprétation pour développer une taxonomie de l'incertitude spatiale pour les trois composantes énumérées plus haut. L'incertitude spatiale est donc interprétée en termes de cohérence entre les éléments et les attributs des réalisations multitemporelles. La relation entre l'incertitude des frontières et leurs caractéristiques géométriques a également été explorée (De Groeve et Lowell 2001). Une autre approche s'est intéressée au processus de photo-interprétation comme tel. Pour ce faire, différents interprètes identifient seulement les polygones pour lesquels ils ont une certitude à 100%. Les résultats de l'ensemble des interprètes sont analysés et la carte est complétée en appliquant un algorithme basé sur le diagramme de Voronoï (Lowell 1994, Thierry et Lowell 2001). Cette approche demande un effort considérable afin d'obtenir plusieurs interprétations des photos qui permettent de mesurer la cohérence des polygones. Lowell *et al.* (2000) ont proposé une méthodologie qui combine les deux approches précédentes. Cette méthode emploie les interprétations multiples afin de déterminer la cohérence des frontières et des attributs en leur ajoutant une

valeur de certitude. L'algorithme basé sur les diagrammes de Voronoï est utilisé pour la production de surfaces de valeur de certitude pour les différents attributs. La structure de Voronoï a également été utilisée dans le cadre de la théorie des ensembles flous pour évaluer l'incertitude spatiale des frontières des cartes vectorielles et elle constitue une méthode alternative pour la représentation des phénomènes avec des frontières floues (Lowell 1994, Wang et Hall 1996).

Quant aux modèles de bandes epsilon, dont le postulat implicite suppose que la région d'incertitude thématique est près des frontières, ils ne sont pas adéquats pour tous les cas d'incertitude spatiale (Goodchild 2001). Dans plusieurs situations, les aires près des frontières sont plus caractéristiques de la classe de l'objet que les aires près du noyau. Pendant la photo-interprétation, l'interprète décide de la localisation des frontières et des petites parcelles qui seront séparées en polygones. Lorsque les petites parcelles sont situées près des frontières, elles sont probablement respectées, tandis que celles qui se trouvent au centre du polygone sont probablement effacées. Il en résulte une incertitude sur les attributs qui peut être supérieure au centre d'un polygone à celle près des frontières (Goodchild 2001). L'existence d'une bande de largeur homogène d'incertitude constitue un autre aspect critiqué. Cette homogénéité dans la largeur de la bande est peu réaliste et des largeurs floues ou différentielles, en fonction du type de frontières, ont été proposées (Edwards et Lowell 1996, Lowell 1994).

## **Approche analytique**

L'approche géomatique met en évidence le fait que les données géographiques fournissent des représentations imparfaites, incomplètes et incertaines du monde réel (Longley *et al.* 1999b). D'une part, l'extraction de l'information contenue dans les données géographiques exige des approches analytiques qui tiennent compte de la dépendance spatiale et de l'hétérogénéité inhérente aux données spatiales (Fischer 1999). D'autre part, l'analyse exploratoire de la structure non spatiale des données peut orienter le traitement de l'incertitude des estimations forestières. Dans un premier temps, différentes approches pour établir et caractériser les patrons spatiaux sont décrites dans les lignes qui suivent. Par la

suite, une approche de l'exploration de la structure non spatiale des données est abordée. Finalement, les méthodes de fusion des données de sources diverses sont présentées.

## **1. Analyse spatiale**

Le premier pas pour comprendre les processus écologiques est l'identification des patrons spatiaux (Fortin *et al.* 2002, 2000). Plusieurs méthodes d'analyse spatiale permettent d'établir, de mesurer et de modéliser leurs caractéristiques.

### **1.1. Mesures de dépendance spatiale**

Le degré d'association spatiale entre les données représentant des points du territoire est fondamental pour l'analyse statistique de la dépendance et de l'hétérogénéité dans les patrons spatiaux (Getis 1999). La dépendance spatiale a été conceptualisée à partir de deux perspectives principales, soient une approche géostatistique, qui prévaut dans les sciences physiques, et une approche matricielle, liée aux sciences sociales (Anselin 1999).

La perspective de la géostatistique considère les observations comme des échantillons de points issus d'une distribution spatiale continue sous-jacente (Fortin 1999). Cette surface est modélisée au moyen d'un variogramme qui représente la semivariance entre des paires de points comme une fonction continue de la distance (Anselin 1999). La semivariance mesure la différence moyenne des valeurs entre des paires de points localisés à une distance déterminée (Legendre et Legendre 1998).

La perspective matricielle, quant à elle, considère que les observations correspondent à des portions discrètes du territoire et que les données spatiales sont le résultat d'un processus stochastique (Anselin 1999). Les mesures d'autocorrélation dérivent de la semivariance ( $\gamma$ ). Elle est définie par le produit d'une matrice, qui représente les associations locales possibles entre tous les points, et d'une matrice d'association non spatiale. L'association spatiale peut être définie en termes de covariances ( $I$  de Moran), de différences ( $c$  de Geary) ou d'additions ( $G$  de Getis) (Getis 1999). Ces statistiques sont globales, car toutes les mesures entre les points sont prises en compte simultanément. Elles fournissent donc une

estimation moyenne et isotropique de l'intensité de l'autocorrélation à chaque intervalle de distance (Fortin *et al.* 2002).

Ces mesures globales reposent, comme la géostatistique, sur le postulat de stationnarité de la variable qui requiert que la moyenne et la variance soient constantes dans l'espace. Pour leur part, les mesures d'association locale permettent d'établir des patrons locaux et de détecter la présence de non-stationnarité localement. Elles utilisent la semivariance  $\gamma$  avec un vecteur colonne à la place de la matrice d'association. Les statistiques locales ( $I$ ,  $c$  ou  $G$  locaux) mesurent l'association de chaque point avec tous les autres points (Anselin 1995, Boots 2002, Boots et Tiefelsdorf 2000, Ord et Getis 1995).

## **1.2. Détection de patrons spatiaux**

Les approches statistiques usuelles pour la détection des patrons spatiaux se basent sur l'hypothèse d'un patron spatial complètement aléatoire. Elles utilisent le concept de distance au plus proche voisin, qui mesure la distance moyenne de chaque point au premier voisin. Cette distance est comparée à celle espérée selon le patron théorique (Fortin *et al.* 2002). Dans la même perspective, la fonction  $K$  de Ripley quantifie l'intensité du patron spatial dans un rayon circulaire, au moyen de la relation entre la somme de toutes les paires de points à l'intérieur du rayon et la somme de la totalité des paires de points (Fortin *et al.* 2002, Getis 1999).

## **1.3. Analyse de frontières**

Les approches d'analyse spatiale décrites précédemment sont appropriées pour le modèle conceptuel de champs. L'analyse de frontières cherche plutôt à définir des objets (qui sont en fait les frontières) dans les champs et ce, dans un contexte de reconnaissance de patrons fondé sur les modèles spatiaux nuls; c'est ce qui la distingue des techniques d'analyse d'images (Jacquez *et al.* 2000).

À partir de cette analyse, deux types de frontières peuvent être identifiés, les frontières indirectes, qui délimitent des zones homogènes, et les frontières directes, qui indiquent des

zones de changement intense de valeurs pour les variables considérées. Les premières sont identifiées au moyen d'algorithmes de regroupement, alors que les deuxièmes le sont avec des algorithmes de détection du taux de changement (Jacquez *et al.* 2000). Les frontières peuvent être identifiées pour des variables quantitatives ou qualitatives, dont les échantillons sont distribués régulièrement ou irrégulièrement.

Les algorithmes de regroupement à contraintes de contiguïté spatiale considèrent les relations spatiales entre les échantillons en terme de voisinage. Les groupes sont formés lorsque les sites adjacents montrent un certain degré de similarité (Legendre et Fortin 1989).

La méthode de détection de changement pour des points distribués régulièrement, mesure, quant à elle, le gradient d'une variable quantitative continue dans l'espace. Les valeurs de cette variable correspondent aux noeuds d'un réseau carré, alors que la magnitude du taux de changement est calculée pour le centroïde de chaque carré. La pente correspond, pour sa part, à l'intensité du changement. Un seuil arbitraire est établi (par exemple, 10 % des taux les plus élevés), afin que les taux soient considérés comme un élément de frontière (EF) (Fortin 1997, Fortin et Drapeau 1995). Les frontières sont déterminées en reliant les EFs adjacents (Barbujani *et al.* 1989). Lorsque les échantillons ont une distribution irrégulière, les relations de voisinage sont établies au moyen d'une triangulation (par exemple, la triangulation de Delaunay; Fortin 1994, Fortin et Drapeau 1995). L'algorithme de détection de changements calcule la première dérivée partielle en utilisant les trois échantillons qui forment un triangle et l'assigne au centroïde. La sélection des EFs est similaire à celle de la matrice régulière. Quant aux variables qualitatives, ce sont des mesures de dissimilitude qui sont estimées entre deux échantillons adjacents (Fortin et Drapeau 1995, Oden *et al.* 1993). La signification statistique des frontières est évaluée au moyen des statistiques de frontières, de superfluidité et de superposition (Fortin et Drapeau 1995, Jacquez *et al.* 2000).

## **2. Analyse exploratoire non spatiale**

L'analyse exploratoire vise à découvrir la structure à l'intérieur des données. Parmi les approches statistiques disponibles, les modèles basés sur les arbres de régression et de classification (CART) sont de plus en plus utilisés pour dériver des règles de prédiction facilement évaluables. Ils servent également à mettre en évidence les variables les plus importantes pour expliquer un phénomène, ainsi que pour évaluer l'adéquation des modèles linéaires et pour réduire des ensembles de données multivariées (Clark et Pregibon 1992). Les CART possèdent plusieurs avantages par rapport aux modèles linéaires et additifs. Ils permettent en effet de modéliser les interactions générales entre les variables prédictives au-delà des formes multiplicatives ou additives particulières. Par conséquent, ils sont plus adéquats pour capter les comportements non additifs. Ils sont également appropriés pour des ensembles de données multidimensionnelles et pour différents types de variables, de structures non gaussiennes et non homogènes (Breiman *et al.* 1984).

Un arbre de classification ou de régression est un ensemble de règles déterminées par une procédure de partition récursive (Clark et Pregibon 1992, Venables et Ripley 1999). La variable réponse est qualitative pour les arbres de classification et quantitative pour les arbres de régression. La procédure de partition récursive divise ainsi l'ensemble des données successivement jusqu'à l'obtention d'un modèle dont l'hétérogénéité est minimale.

## **3. La fusion des données**

La « fusion de données » couvre un ensemble très large de significations. Elle peut être définie comme le cadre formel dans lequel sont exprimés les moyens et les techniques permettant le regroupement des données provenant de différentes sources en vue de produire un certain type d'information. La fusion vise l'obtention d'information de plus grande qualité et la définition exacte de « plus grande qualité » dépend de l'application (Wald 1999). Dans cette définition, l'accent est mis sur le cadre de la fusion et la qualité du produit résultant. La qualité n'a pas une seule signification, c'est-à-dire qu'une meilleure qualité peut consister, par exemple, à augmenter l'exactitude d'un paramètre ou d'une classification. La qualité peut également être interprétée comme la production

d'informations plus utiles pour une application ou le développement d'une procédure opérationnelle plus robuste.

Dans un contexte spatial, la fusion des données peut être développée sur trois plans, celui du pixel, celui des éléments et celui des décisions (Gong 1994, Pohl et Genderen 1998). Les définitions de chacun des plans dépendent de la communauté scientifique concernée. Ainsi, dans un contexte de télédétection, le pixel est remplacé par les capteurs; dans un contexte de base de données, les éléments sont les données; tandis que, dans un contexte de logiques de raisonnements, on parle d'un niveau de connaissances (Edwards et Jeansoulin 2004). Wald (1999) considère que le terme pixel, très répandu dans la littérature, fait référence à un support d'information et qu'il n'a pas une signification sémantique. Cet auteur propose l'utilisation des termes « mesure », « observation » ou « signal ». Il considère également que la fusion ne doit pas être restreinte à un seul plan parce qu'elle est généralement développée sur plusieurs plans. Plus spécifiquement, la fusion des observations, pixels ou capteurs vise l'intégration des observations des deux sources. La fusion des éléments ou des données implique l'extraction de ces éléments à partir de chaque source de manière indépendante et leur fusion se fait dans une étape ultérieure. La fusion des connaissances consiste à la combinaison de deux bases de connaissances, plutôt que de deux bases de données (Edwards et Jeansoulin 2004). La fusion sur le plan des décisions consiste, quant à elle, à l'intégration de plusieurs versions de données qui résultent de procédures diverses et à leur interprétation dans un contexte préalablement défini.

Plusieurs termes, comme « intégration », sont souvent utilisés comme des synonymes de « fusion ». Cependant l'intégration est considérée un processus réversible dont le produit résultant maintient l'identité des sources, tandis que la fusion est un processus irréversible dont le produit final ne conserve pas l'identité des sources. La fusion, qui constitue donc un cas particulier de l'intégration, exige une correspondance et une complémentarité totale entre au moins un sous-ensemble des sources à fusionner (Edwards *et al.* 2000).

La fusion des données et des connaissances pose des difficultés liées aux hétérogénéités sémantiques, syntactiques et schématiques des sources d'information (Benslimane *et al.* 2000, Fonseca *et al.* 2000, Stock et Pullar 1999). Les premières font référence aux

différences dans la définition des concepts et des règles utilisées pour déterminer si une entité du monde réel est un exemple d'un concept donné. Les hétérogénéités syntactiques, pour leur part, concernent les différences liées aux structures des éléments utilisés pour représenter le monde réel. Dans le cas des données spatiales, ces structures correspondent aux types d'éléments géométriques et leurs caractéristiques, tandis que dans le cas des données non spatiales, elles correspondent aux détails des représentations, aux types, ainsi qu'aux contraintes et aux domaines des données. Finalement, l'hétérogénéité schématique se réfère aux éléments d'une base de données qui sont utilisés pour représenter une entité du monde réel comme : les attributs, les relations et les classes des objets (Stock et Pullar 1999).

L'hétérogénéité sémantique survient lorsque les différentes sources de données ne partagent pas un langage commun. Cependant, même lorsque les sources partagent un même langage, un terme peut avoir des significations diverses et, inversement, des termes différents peuvent avoir la même signification. De plus, des similarités syntactiques entre les éléments des sources à fusionner peuvent ne pas correspondre à des similarités sémantiques (Stock et Pullar 1999).

L'hétérogénéité syntactique prend son origine dans les méthodes de collecte, de manipulation et de présentation des données. Les structures que les données ont dans une base de données (matricielle ou vectorielle), les types de données (nominal, ordinal, intervalle, ratio), le degré de détail (résolution spatiale, temporelle et thématique), le degré d'incertitude, ainsi que de complétude et les formats incompatibles constituent les obstacles syntactiques les plus importants à la fusion.

L'hétérogénéité schématique, quant à elle, réfère au modèle spatial utilisé pour représenter les entités du monde réel. Goodchild (1992) propose deux modèles, le modèle-objet et celui des champs qui ont été décrits en détail aux sections précédentes.

Le processus de fusion implique la résolution des incompatibilités et des hétérogénéités entre les sources de données afin de les mettre en correspondance et de les combiner. La mise en correspondance et la combinaison des données se font par diverses procédures qui

utilisent des raisonnements dont les plus importants sont la conversion numérique ou symbolique, la mise en correspondance, la combinaison ou la décision. La représentation et le traitement des données sont modélisés au moyen de raisonnements dont les plus fréquents sont probabilistes, possibilistes, non monotoniques ou fondés sur des preuves.

Le raisonnement probabiliste permet de modéliser certains phénomènes aléatoires. Dans ce cadre, l'approche bayésienne sert à estimer la probabilité d'appartenance des observations à une classe. L'estimation se fait par la modélisation de la fonction de densité de probabilité de la classe et de la probabilité d'occurrence *a priori*. Ensuite, des règles de décision sont appliquées pour assigner chaque observation à une classe déterminée (Kittler *et al.* 1998).

Le raisonnement fondé sur des preuves estime les fonctions de « croyance » et de « plausibilité ». La première mesure l'importance des preuves d'un sous-ensemble parmi l'ensemble des possibilités, alors que la deuxième définit le rang dynamique de possibilité d'assignation à la classe (Wang et Civco 1994). Les croyances de chaque sous-ensemble sont combinées pour obtenir de nouvelles fonctions de croyance. Une fois qu'elles sont estimées pour chaque source et que les combinaisons sont développées, des règles de décision guident l'assignation de l'observation à une classe. Cette approche est très utilisée pour la classification des images satellites en intégrant les informations d'autres sources (Peddle *et al.* 1994, Wang et Civco 1994).

L'approche possibiliste permet, pour sa part, de raisonner sur des connaissances imprécises ou vagues en prenant en compte les incertitudes subjectives formulées sur les événements. L'information disponible sur la valeur d'un paramètre est représentée par une distribution de possibilités  $\pi$  qui est la fonction d'appartenance à un ensemble représentant une information imprécise.

Le raisonnement non monotone permet, lui, d'utiliser l'information par défaut. Cette information est inférée à partir de règles très générales et possède une exactitude plus faible. Avec ce type de raisonnement, le changement des inférences est possible à mesure que de nouvelles informations sont incorporées à la base de données (Hunter 1996).

La représentation ensembliste de l'information permet, pour sa part, d'effectuer la fusion à partir d'opérations entre différents ensembles. Les fusions conjonctive, disjonctive et adaptative sont les méthodes les plus utilisées. La première équivaut à l'intersection des ensembles flous et suppose que les sources d'information sont totalement fiables et concordantes. La fusion disjonctive correspond à l'union des ensembles flous et est utilisée lorsque les sources sont discordantes et qu'une seule des sources est fiable. La fusion adaptative est utilisée en fonction du degré de conflit, lorsque les sources ne sont pas fiables. Les formalismes développés dans le cadre de la théorie des possibilités sont donc de plus en plus utilisés, de par leur abord facile (Zhu et Band 1994).

Une approche de plus en plus utilisée dans les contextes de fusion et d'interopérabilité des bases de données et de connaissances est celle des ontologies. Il existe deux définitions très différentes de l'ontologie qui correspondent à deux communautés scientifiques distinctes. Dans un contexte philosophique, l'ontologie est définie comme la branche de la philosophie qui traite de la nature et de l'organisation de la réalité, c'est-à-dire le système particulier de catégories qui reflète une vision spécifique du monde (Fonseca *et al.* 2002, Mostafavi *et al.* 2004). Dans le domaine de l'intelligence artificielle, le terme « ontologie » est utilisé pour désigner un artefact d'ingénierie qui décrit une certaine réalité à l'aide de vocabulaire spécifique en utilisant un ensemble de postulats pour définir les termes (Fonseca *et al.* 2002). Les raisons de l'utilisation croissante de cette approche sont, d'une part, l'adéquation des ontologies pour gérer l'hétérogénéité sémantique des sources à fusionner et, d'autre part, leur potentiel pour manipuler des représentations multiples d'un même phénomène (Fonseca *et al.* 2002).

Plus proches du domaine philosophique que de celui de l'intelligence artificielle, des ontologies ont été développées depuis quelques années pour le domaine des sciences de l'information géographique; tandis que pour le domaine écologique, elles commencent à être examinées (Bittner et Edwards 2001, Fonseca *et al.* 2002, Fonseca et Martin 2004, Smith 1995, Smith et Mark 1998, 2001). Le développement des ontologies géographiques et écologiques est important pour comprendre la représentation des phénomènes dans les bases de données et la façon dont cette information est échangée entre les différentes

communautés d'utilisateurs. Les ontologies du domaine géographique mettent l'accent sur les relations topologiques et mérologiques (celles qui explicitent les relations entre les parties et le tout) des entités géographiques, ainsi que sur leurs relations hiérarchiques. Les ontologies écologiques doivent, pour leur part, ajouter à ces aspects spatiaux des phénomènes géographiques, des propriétés spécifiques permettant de saisir la nature auto-organisée et relativement ouverte des écosystèmes ainsi que le caractère temporel des phénomènes écologiques (Fonseca et Martin 2004, Fonseca *et al.* 2002). Des propriétés comme la récursivité, qui permet de saisir la dimension temporelle, et les nœuds, qui représentent les aspects fonctionnels des entités écologiques, caractérisent les éco-ontologies (Fonseca et Martin 2004, Fonseca *et al.* 2002).

## Résumé du chapitre

Tout au long de ce chapitre, les principaux concepts liés à cette recherche ont été présentés. Les différentes significations du terme échelle dans les domaines d'intérêt ont été discutées, car ce concept est à la base de plusieurs sujets abordés aux chapitres suivants. Les notions développées dans les trois approches permettent, d'une part, de comprendre la forêt réelle ainsi que les particularités et l'incertitude spatiale des données forestières qui la représentent et, d'autre part, elles donnent l'occasion de connaître les méthodes disponibles pour le traitement de cette incertitude, de même que pour l'extraction de l'information sur la ressource forestière.

Le chapitre suivant offre une description du site d'étude et présente les caractéristiques des données qui seront utilisées aux chapitres subséquents et qui permettront l'exploration des approches méthodologiques afin de réduire l'incertitude spatiale locale associée aux surfaces et aux estimations du volume ligneux d'une carte forestière.



## 1. Climat et physiographie

Le climat de la Forêt Montmorency est de type continental subpolaire humide avec des précipitations annuelles totalisant 1416 mm, ce qui explique la fréquence relativement basse des feux (Beaulieu et Lowell 1994). La température moyenne annuelle est de 0,3 °C et la saison moyenne de croissance est relativement courte, soit 143,8 jours par année (Bélanger et Bouliane 1992).

Le territoire est composé d'un plateau de 750 m d'altitude, surplombé de collines de plus de 900 m et découpé par deux rivières principales, la rivière Montmorency et son affluent, la rivière Noire (Bélanger et Bouliane 1992). Quatre lacs et plusieurs ruisseaux complètent le réseau hydrographique. Le terrain varie entre des pentes graduelles et des pentes abruptes, et celles exposées au sud sont, en général, plus accentuées que celles exposées au nord. Les arêtes divisent la forêt en trois bassins. Les vallées sont jeunes; les plus importantes sont typiquement glaciaires et les moins importantes ont un profil en V (Jurdant 1964).

Les dépôts dominants sont d'origine glaciaire et caractérisés par le matériel morainique d'hétérogénéité texturale considérable. Les dépôts glaciaires sont des tills d'ablation, de fond et délavés. Le long de la rivière Montmorency, les dépôts fluvioglaciaires et fluviaux sont prédominants (Bélanger et Bouliane 1992). Les sols varient en fonction du type de dépôt. Des podzols ferro-humiques orthiques ou gleyifiés se développent sur le till (Bérard et Côté 1996).

## 2. Végétation

La végétation appartient au domaine de la sapinière à bouleau à papier de la forêt boréale (Bérard et Côté 1996). Ce domaine est composé de forêts de conifères principalement dominées par le sapin baumier (*Abies balsamea* (L.) Mill) associé à l'épinette blanche (*Picea glauca* (Moench) Voss) et au bouleau à papier (*Betula papyrifera* Marsh.). Il est aussi possible d'y trouver de l'épinette noire (*Picea mariana* (Mill.) BSP) sur les sites où

les conditions sont extrêmes ou après un feu (Coulombe 1996). Les autres espèces d'arbres sont le *Sorbus americana*, l'*Acer spicatum* et le *Prunus pensylvanica* (Jurdant 1964).

Les types de végétation les plus importants dans la Forêt Montmorency sont les sapinières (Jurdant 1964). La sapinière à *Dryopteris* ainsi que la sapinière à *Dryopteris* et *Hylocomium* caractérisent les sites mésiques de milieu de pente (Bérard et Côté 1996). Elles possèdent un fort potentiel de croissance, puisqu'elles sont dominées par le sapin, qui est accompagné d'épinette blanche, de bouleau blanc et d'épinette noire, en proportions variables. Le fort recouvrement des strates herbacées est lié à une faible densité des strates arborescentes.

Le peuplier faux-tremble (*Populus tremuloides*) ou le bouleau blanc dominent les sites perturbés par le feu ou par la coupe sévère, pendant une période d'environ 100 ans (Bérard et Côté 1996). Dès l'âge de 30 ans, les sapins disséminés sous la couverture commencent à produire des semences tous les deux ans. La régénération est beaucoup moins abondante dans les peuplements où la litière feuillue est importante. Le dépérissement graduel des essences feuillues surannées fait en sorte que les essences résineuses prennent graduellement plus d'importance et forment des communautés mélangées ou résineuses. Les sapinières matures sont systématiquement attaquées par des caries de racine qui réduisent leur résistance aux vents et aux attaques des insectes.

Pour sa part, la sapinière à épinette noire occupe les sites mésiques de haut de pente, les sommets et les terrains plats peu fertiles, ainsi que les sites mésiques récemment perturbés par le feu (Jurdant 1964). La pessière noire à sphaignes caractérise, quant à elle, les sites hydriques sur sol organique, dans les dépressions humides. C'est une communauté dense dont le parterre est dominé par les mousses hypnacées et les sphaignes ombrotrophes. Vers le centre de la dépression, la formation devient arbustive ou herbacée avec des épinettes noires, des éricacées ou des cypéracées (Bérard et Côté 1996, Jurdant 1964).

### 3. Perturbations

#### 3.1. Feu

Le cycle du feu affecte la succession écologique et donc la composition de la mosaïque forestière. Dans la forêt boréale, un cycle court conduit à une mosaïque dominée par les essences feuillues de lumière ou par des essences résineuses pionnières comme le pin gris et l'épinette noire, il s'agit en fait d'une succession cyclique (Bergeron 2000). Un cycle de feu de plus de 200 ans favorise, pour sa part, le retour des sapinières (Bergeron et Dansereau 1993). Des études ont été faites sur la fréquence des feux dans le secteur sous-boréal Est de la province du Québec et elles ont mis en évidence une décroissance dans la fréquence des feux depuis le milieu du XIX<sup>e</sup> siècle, accentuée durant le XX<sup>e</sup> siècle (Bergeron *et al.* 2001). Il s'agirait d'une conséquence des changements climatiques liés au Petit Âge Glaciaire (*ca.* 1850). Dans ce contexte, la Forêt Montmorency n'est pas dans une région à haut risque d'incendies grâce aux fortes précipitations en été. Aucun feu n'y a été rapporté depuis 1963 (Bélanger et Bouliane 1992).

#### 3.2. Épidémies

Les épidémies de la tordeuse des bourgeons de l'épinette (*Choristoneura fumiferana* Clem) jouent un rôle important dans la composition des mosaïques dans le domaine de la sapinière à bouleau à papier. C'est l'insecte le plus destructeur des peuplements de la Forêt Montmorency, qui a subi quatre épidémies dans les derniers 200 ans (Beaulieu et Lowell 1994), les dernières datant de 1940 et de la période comprise entre 1975 et 1985 (Bélanger et Bouliane 1992).

Dans l'Est, où les essences résineuses abondent, les épidémies sont sévères et ont pour effet de régénérer les sapinières matures, dans une succession cyclique (Bérard et Côté 1996). Cependant, il y a des cas pour lesquels les sapinières immatures secondaires (après une coupe) sont remplacées par des bétulaies blanches à sapin (Déry *et al.* 2000).

Des études dendrochronologiques montrent que l'étendue et la sévérité des épidémies de la tordeuse des bourgeons de l'épinette semblent avoir augmenté au cours du XX<sup>e</sup> siècle. Parmi les raisons de cette augmentation, on peut citer la coupe des peuplements d'espèces résineuses pionnières, le contrôle des feux de forêt, l'utilisation de phytocides et l'augmentation des précipitations et des épidémies causées par d'autres insectes (Bérard et Côté 1996).

### **3.3. Chablis**

Les sapinières équiennes surannées accèdent à une phase de décrépitude lorsqu'elles ne sont pas ravagées par les épidémies. Elles sont vulnérables au chablis, étant donné l'enracinement superficiel et le développement de caries des racines. Les tiges de sapin tombent de façon isolée ou en petits groupes par l'action du vent et le phénomène s'accélère lorsque les sapinières dépassent l'âge de 80 ans et que leur densité diminue (Bérard et Côté 1996). Comme dans le cas des épidémies, le développement d'un peuplement dense et de structure équienne est favorisé par le chablis.

Aux chablis réguliers qui affectent généralement des superficies restreintes s'ajoutent des chablis catastrophiques, survenant lors de tempêtes automnales, caractérisés par de fortes précipitations en eau ou en neige mouillante et des sols détrempés qui diminuent la capacité de résistance des racines. Dans ces cas, tous les peuplements sont touchés, quels que soient l'exposition, le drainage, l'âge (plus de 30 ans) et la densité (Bérard et Côté 1996). Particulièrement, à la Forêt Montmorency, le chablis a été responsable de la réduction de la quantité de sapin baumier de plus de 60 ans (Beaulieu et Lowell 1994).

### **3.4. Coupes forestières**

La plus grande partie de la forêt a été coupée à blanc entre 1932 et 1944 (Vézina et Paillé 1969) et c'est pourquoi l'âge moyen des peuplements se situe entre 50 et 60 ans (Beaulieu et Lowell 1994). Des programmes de coupes ont été implantés par l'Université Laval entre 1965 et 1984 (Bélanger et Bouliane 1992). Pendant la première décennie, la récolte a été faite par des coupes à blanc et des coupes d'éclaircie. Par la suite, des éclaircies

commerciales, pré-commerciales et des coupes progressives par bandes ont été appliquées dans le cadre de projets subventionnés ou de recherche, affectant au total environ 400 ha (Bélanger et Bouliane 1992).

## **Données disponibles**

### **1. Cartes forestières**

Il existe trois cartes d'inventaire forestier de la Forêt Montmorency avec une couverture de 6625 ha à l'échelle 1/10000. Les cartes sont basées sur les photo-interprétations de photographies infrarouges à l'échelle 1/15000, obtenues dans le cadre des inventaires forestiers développés en 1973, 1984 et 1992, selon les normes de stratification forestière du moment. Leur projection cartographique est Mercator transverse modifiée (MTM), la surface de référence géodésique est l'ellipsoïde GRS80, tandis que le système de référence géodésique est le NAD83. Dans cette recherche, la carte de 1992 est utilisée comme une des sources principales d'information forestière.

La carte forestière de 1992 a été élaborée à partir de trois tailles d'aire minimale d'interprétation différentes (Annexe I, Tableau 45). Pour chaque polygone représentant un peuplement forestier, les variables qui sont considérées d'intérêt primaire pour cette recherche sont : le *Groupement d'essences (GE)* (Tableau 1), la *Classe de densité (CD)* (Tableau 2), la *Classe de hauteur (CH)* (Tableau 3) et la *Classe d'âge (CA)* (Tableau 4). Ce sont des variables qualitatives qui décrivent les aspects structuraux de la végétation forestière.

La *Perturbation d'origine* et la *Perturbation moyenne* sont des variables fréquemment utilisées dans cette thèse. La première indique le phénomène qui a entraîné la destruction des tiges sur plus de 75 % de la surface terrière du peuplement. Les catégories de *Perturbation d'origine* présentes à la Forêt Montmorency sont la coupe totale (ct), le chablis total (cht), l'épidémie sévère (es) et la plantation (pl). La *Perturbation moyenne* est le phénomène qui a détruit entre 25 et 75 % de la surface terrière d'un peuplement. Les

catégories présentes à la Forêt Montmorency sont la coupe partielle (cp), le chablis partiel (chp), l'épidémie légère (el), l'épidémie modérée (em), l'éclaircie pré-commerciale (epc) et l'éclaircie commerciale (ec) (Ministère des Ressources naturelles du Québec 1995).

Tableau 1. Catégories de *Types de couvert* et de *Groupements d'essences* présentes à la Forêt Montmorency. ST : surface terrière. (Ministère de Ressources naturelles du Québec 1995).

<i>Type de couvert</i>	<i>Groupement d'essences</i>	Description du Groupement d'essences
Feuille	Bb, Bj, FNC	Couvert de bouleau blanc ou de bouleau jaune ou de feuillus non commerciaux supérieur à 50 % de la ST de la partie feuillue
	Fi	Couvert de bouleau blanc et de peuplier supérieur à 50 % de la ST de la partie feuillue
Mélangé à prédominance feuillue	BbS, BjS, PeS	Couvert de bouleau blanc ou de bouleau jaune ou de peuplier à $\geq 50$ % de la ST de la partie feuillue, le sapin baumier couvre plus de 50 % de la ST de la partie résineuse
	BjR	Couvert de bouleau jaune entre 26 et 74 % de la ST de la partie feuillue, une ou plusieurs essences résineuses occupent entre 26 et 50 % de la ST du peuplement
	FiS	Couvert de bouleau blanc et de peuplier à $\geq 50$ % de la ST de la partie feuillue, le sapin baumier couvre plus de 50 % de la ST de la partie résineuse
Mélangé à prédominance résineuse	SBb, SBj, SPe	Couvert de sapin baumier à $\geq 50$ % de la ST de la partie résineuse, le bouleau blanc ou le bouleau jaune ou le peuplier couvre plus de 50 % de la ST de la partie feuillue
	SFi	Couvert de sapin baumier à $\geq 50$ % de la ST de la partie résineuse, le bouleau blanc et le peuplier en proportions égales couvrent plus de 50 % de la ST de la partie feuillue
Résineux	SS, EE	Couvert de sapin baumier ou d'épinette à $\geq 75$ % de la ST de la partie résineuse
	SE	Couvert de sapin baumier entre 50 et 74 % de la ST de la partie résineuse, l'épinette est la deuxième essence résineuse importante
	ES, EPb	Couvert d'épinette entre 50 et 74 % de la ST de la partie résineuse, le sapin baumier ou le pin blanc sont la deuxième essence résineuse importante

Tableau 2. Catégories de *Classe de densité* (Ministère de Ressources naturelles du Québec 1995).

Catégorie	Peuplement où le pourcentage de couvert est
A	supérieur à 80 %
B	61 à 80 %
C	41 à 60 %
D	25 à 40 %

Tableau 3. Catégories de *Classe de hauteur* (Ministère de Ressources naturelles du Québec 1995).

Catégorie	Hauteur moyenne des arbres dominants et codominants
1	supérieure à 22 m
2	entre 17 et 22 m
3	entre 12 et 17 m
4	entre 7 et 12 m
5	entre 4 et 7 m
6	Hauteur moyenne des tiges entre 1,5 et 4 m

Tableau 4. Catégories de *Classe d'âge* selon le type de peuplement (Ministère de Ressources naturelles du Québec 1995).

Type de peuplement	Catégories d'âge
Équienne	10 (0-20 ans) - 30 (21-40 ans) - 50 (41-60 ans) - 70 (61-80 ans) - 90 (81-100 ans) - 120 (plus de 101 ans)
Inéquienne	-Jeune : la plupart de tiges appartiennent à au moins 3 classes d'âge consécutives, de 10 à 70 ans. La <i>Perturbation d'origine</i> remonte à moins de 80 ans -Vieux : la plupart de tiges appartiennent à au moins 3 classes d'âge consécutives, de 10 à 120 ans. La <i>Perturbation d'origine</i> remonte à plus de 80 ans
Étagé	Les tiges forment deux étages différents, dont la hauteur diffère d'au moins 5 m, occupant chacun au moins 25 % de couvert. On indique les deux classes d'âge où la première est celle du peuplement dont la surface terrière est la plus importante.

D'autres variables d'intérêt qui font partie de la carte forestière et qui sont utilisées dans cette recherche sont : la *Classe de pente (CP)* (Annexe I, Tableau 47), l'*Année de coupe (AC)*, le *Stade de développement* du peuplement (Annexe I, Tableau 48) et l'*Appellation* de la strate forestière. Cette dernière variable est formée par la combinaison de catégories pertinentes de variables telles que le *Groupement d'essences*, la *Classe de densité*, la *Classe*

*de hauteur*, la *Classe d'âge*, la *Perturbation d'origine*, la *Perturbation moyenne* et l'*Année de coupe*.

Les variables dendrométriques de la carte forestière sont le *Volume moyen* (en m<sup>3</sup>/ha), la *Surface terrière* (en m<sup>2</sup>/ha) et la *Densité* (nombre de tiges/ha). Elles sont estimées séparément pour les essences feuillues et résineuses, ainsi que pour l'ensemble de la végétation forestière. Dans cette recherche, la variable dendrométrique d'intérêt principal est le *Volume total* de la végétation ligneuse (en m<sup>3</sup>/ha), toutes espèces feuillues et résineuses confondues. Les valeurs des variables dendrométriques ne correspondent pas à chaque peuplement, mais plutôt à la strate forestière à laquelle ils appartiennent. Chacune des strates forestières est donc associée à un groupe d'échantillons de terrain à partir desquels ces valeurs sont dérivées.

Les variables comme le *Groupement d'essences*, la *Classe de hauteur*, la *Classe de densité* et la *Classe d'âge* des cartes forestières de 1973 et de 1984 correspondent à d'autres systèmes de classification, avec une résolution thématique plus grossière que celle de la carte de 1992.

## **2. Données de terrain**

Les données des placettes d'échantillonnage (p-e) forestier datant de 1973 (231 placettes), 1986 (341 placettes), 1992 (342 placettes) et celles d'un réseau de p-e permanentes mesurées entre 1995 et 1999 (201 placettes) sont disponibles.

Pour cette recherche, les p-e de 1992 et celles qui sont permanentes (1995-1999) sont utilisées en raison de leur homogénéité par rapport aux systèmes de classification des variables, telles le *Groupement d'essences*, la *Classe de hauteur*, la *Classe de densité* et la *Classe d'âge*. Les p-e des années 1973 et 1984, tout comme les cartes forestières des mêmes années, utilisent des systèmes de classification qui possèdent une résolution thématique plus grossière que celles de 1992.

Les p-e sont circulaires et ont une superficie de 400 m<sup>2</sup>. Elles sont localisées sur des transects au nombre de trois ou quatre par transect. Les p-e qui se trouvent sur un même transect sont séparées par une distance d'environ 150 m. La distance moyenne entre toutes les p-e (de 1992 et du réseau permanent mesuré entre 1995-1999 confondues) est de 182,8 m où la valeur minimale est de 15 m et le maximum de 1014 m. La densité de cet ensemble est d'une p-e par 12 ha. Au Québec, cette relation est d'une p-e par 500 ha en moyenne, pour un inventaire global et d'une p-e par 100 ha environ, pour un inventaire d'exploitation (Beaulieu 1994).

Pour chaque p-e, le feuillet original contient la description de la strate de la carte forestière, qui correspond au peuplement dans lequel la p-e est localisée, ainsi que la strate reconnue sur le terrain. Ces deux strates sont définies à partir de variables telles que le *Type forestier*, le *Groupement d'essences*, la *Classe de densité*, la *Classe de hauteur*, la *Classe d'âge*, la *Classe de pente*, la *Perturbation d'origine* et la *Perturbation moyenne*. Même si les variables de terrain contiennent les mêmes catégories que les variables cartographiques, qui sont interprétées à partir de photographies aériennes et représentées sur la carte, elles ne sont pas identiques. La différence entre ces deux types de variables, est que les variables de terrain sont estimées sans prendre en considération une aire minimale comme c'est le cas des variables cartographiques.

La *Hauteur des arbres dominants* (en m) et la distribution de chaque essence forestière dans des classes de diamètre constituent les variables dendrométriques mesurées sur le terrain. Parmi les variables physiographiques estimées sur le terrain, les suivantes sont employées dans cette recherche : l'*Altitude* (en m), l'*Exposition* (en degrés), l'*Inclinaison* (en pourcentage), le *Versant* (Annexe I, Tableau 49), la *Situation sur la pente* (Annexe I, Tableau 50), la *Forme de la pente* (Annexe I, Tableau 51), la *Texture des horizons B et C* (Annexe I, Tableau 52), le *Type d'humus* (Annexe I, Tableau 53), l'*Épaisseur de la matière organique* (en cm), la *Pierrosité* (en pourcentage), l'épaisseur des *Dépôts de surface* (Annexe I, Tableau 54) et la *Classe de drainage* (Annexe I, Tableau 55). Les variables dont aucune unité n'est indiquée sont dites qualitatives (les détails de catégories de chacune des variables sont présentés aux tableaux de l'Annexe I).

### **3. Autres cartes**

Les autres sources d'information sont la carte topographique à l'échelle 1/10 000, qui présente une équidistance en élévation entre les courbes de niveau de 10 m, la carte du réseau hydrographique et la carte des chemins, toutes deux à l'échelle 1/10 000.

Ce chapitre a présenté de manière succincte une description du site à l'étude ainsi que les caractéristiques des données utilisées pour cette recherche. Le chapitre suivant, pour sa part, explore différentes méthodes pour quantifier le contraste des frontières de la carte forestière de 1992 à différentes échelles. Il aborde également la quantification du contexte spatial, autant des frontières que des peuplements, afin d'évaluer la possibilité de déterminer la probabilité d'existence d'une frontière cartographique à partir de ces aspects.

## Chapitre III - Les frontières de la carte forestière

### Introduction

Lors de la production d'une carte forestière par photo-interprétation, la norme forestière du Québec stipule que les frontières doivent être identifiées de façon hiérarchique : premièrement, celles qui séparent les zones de végétation des zones qui ont un autre type de couvert tels les routes, les lacs et les rivières; ensuite, les frontières entre les types de végétation forestière et non forestière; finalement, les frontières entre divers types forestiers (Ministère des Ressources naturelles du Québec 1995). Ce processus hiérarchique implique la reconnaissance d'un contraste décroissant entre la texture, la structure et la forme des surfaces adjacentes à la frontière, ainsi que certaines caractéristiques de la frontière elle-même comme le degré de netteté, la forme et la longueur.

Évidemment, à mesure que l'on descend dans la hiérarchie, le contraste entre les types de forêt devient moins évident et les frontières résultantes sont plus incertaines, non seulement par rapport à leur localisation, mais aussi en terme de leur existence. Cependant, toutes les frontières entre les peuplements sont représentées de façon similaire sur la carte forestière résultante. De plus, l'exactitude et la précision spatiale sont supposées être les mêmes sur toute la carte. Ce manque d'information sur l'incertitude, autant sur l'existence que sur la position des frontières, est de nature locale et limite fortement l'utilité de la carte forestière pour la planification opérationnelle des activités forestières.

Edwards et Lowell (1996) ont proposé un modèle d'incertitude spatiale locale pour les cartes forestières qui tient compte des différences de texture et du contexte des polygones adjacents à chaque frontière. Ils ont ainsi remarqué l'effet important des conditions locales de contexte telles que la simplicité et la linéarité de la frontière. Selon Csillag *et al.* (2001) le contraste entre les unités adjacentes représente le degré d'intensité (ou force) de la frontière et il peut être plus important à des échelles plus globales qu'à celle du peuplement forestier.

La quantification du contraste entre les types forestiers définis par des variables qualitatives ou semi-quantitatives exige une analyse plus poussée que le simple désaccord entre les catégories, surtout pour une variable complexe comme la composition d'essences. Par exemple, le degré de différence existant entre un peuplement mélangé et un peuplement résineux n'est pas le même que celui entre un peuplement résineux et un peuplement feuillu. Plusieurs indices ont été développés pour quantifier les similitudes entre les communautés écologiques à partir de variables qualitatives ou semi-quantitatives (Legendre et Legendre 1998). Ces indices peuvent être utilisés pour quantifier la similitude entre les peuplements en fonction d'une variable complexe comme la composition d'essences.

Ce chapitre vise à évaluer s'il est possible de déterminer le degré d'incertitude face à l'existence d'une frontière à partir de caractéristiques comme le contraste entre les attributs des peuplements séparés par la frontière, ainsi qu'à l'aide de certaines conditions du contexte ou du voisinage de la frontière.

L'hypothèse de base est la suivante : l'identification d'une frontière par un photo-interprète serait facilitée dans le cas où le contraste de texture, d'intensité et de couleur des peuplements adjacents à la frontière est élevé (Edwards et Lowell 1996). La texture, intensité et couleur de la photographie sont utilisées pour interpréter les attributs des peuplements (De Groeve 1999). En conséquence, le contraste des attributs des peuplements adjacents pourrait être relié à la probabilité d'existence de la frontière. Dans la mesure que le contraste des attributs des peuplements est plus fort, la frontière entre eux serait plus certaine.

Les objectifs du chapitre sont :

1. De quantifier le contraste entre les peuplements adjacents à chaque frontière, en fonction des différentes variables et pour des échelles locales et régionales;
2. De déterminer le contexte spatial de chaque frontière, plus précisément la configuration spatiale locale des valeurs de contraste;

3. De déterminer le contexte spatial de chaque peuplement en terme d'hétérogénéité des variables;
4. D'évaluer l'association entre le contraste et le contexte, et l'incertitude sur l'existence d'une frontière.

## **Méthodologie**

### **1. Données**

La carte forestière de la Forêt Montmorency produite en 1992 est utilisée pour l'estimation des indices de similitude écologique, afin d'être en mesure d'évaluer la force des frontières à différentes échelles et pour quantifier l'hétérogénéité du contexte spatial. Les variables sélectionnées sont le *Groupement d'essences (GE)*, la *Classe de densité (CD)*, la *Classe de hauteur (CH)* et la *Classe d'âge (CA)* dont les détails ont été présentés dans le chapitre précédent. Les cartes forestières de 1992, 1984 et 1973 sont utilisées pour quantifier la probabilité d'existence des frontières.

### **2. Méthodes**

Les méthodes utilisées dans cette étude sont présentées dans l'ordre des objectifs mentionnés auparavant : les méthodes de quantification du contraste écologique entre les frontières (1); celles de quantification du contexte des frontières (2) et des peuplements (3), et la méthode d'évaluation de la relation entre le contraste, le contexte et la probabilité d'existence d'une frontière (4). La Figure 4 illustre ces étapes méthodologiques.

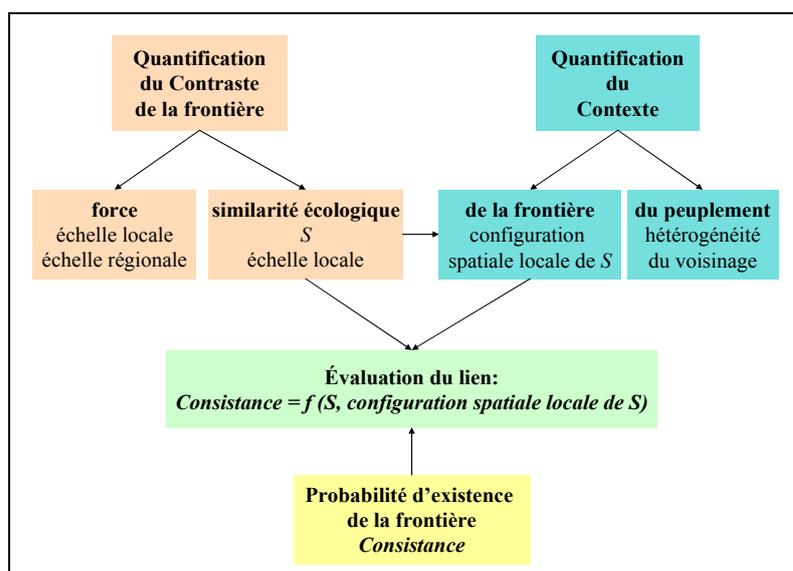


Figure 4. Étapes méthodologiques du Chapitre III.

## 2.1. Quantification du contraste

La quantification du contraste se fait à deux échelles, locale et régionale. À l'échelle locale, le contraste est quantifié au moyen de l'application d'un indice de similitude écologique entre les peuplements adjacents à la frontière. À l'échelle régionale, c'est-à-dire le voisinage de chaque frontière, la quantification du contraste se fait au moyen de la force des frontières à des tailles de voisinage distinctes.

### 2.1.1. La quantification du contraste à l'échelle locale : application d'un indice de similitude écologique aux frontières entre les peuplements

Le contraste entre les catégories de variables des peuplements adjacents à chaque frontière est quantifié par l'indice de similarité écologique «  $S$  » de Legendre et Chodorowsky qui varie entre 0 (similarité inexistante) et 1 (similarité maximale) (Legendre et Legendre 1998). Dans sa forme générale, cet indice calcule la similitude entre deux communautés écologiques comme la moyenne de la similarité partielle des espèces que les composent. Chaque espèce est considérée comme un descripteur de la communauté et son importance est estimée par une mesure d'abondance. L'indice  $S$  se calcule comme suit :

$$S(x_1, x_2) = \frac{\sum w_{12j} s_{12j}}{\sum w_{12j}}, \quad (\text{III.1})$$

où  $x_1$  et  $x_2$  sont les communautés à comparer;  $w_{12j}$  indique le poids de chaque descripteur  $j$  (0 indique que le descripteur est absent dans les communautés  $x_1$  et  $x_2$ , sinon sa valeur est de 1); et  $s_{12j}$  est la valeur de similarité partielle entre  $x_1$  et  $x_2$  pour le descripteur  $j$ . Cette similarité partielle a toujours une valeur entre 0 et 1 et elle peut être estimée de deux façons :

$$s_{12j} = f(\mathbf{x}_{1j}, \mathbf{x}_{2j}) \text{ selon une matrice de similarité partielle élaborée } ad \text{ hoc} \quad (\text{III.2})$$

$$s_{12j} = f(d, k) \rightarrow s_{12j} = 2(k + 1 - d) / 2k + 2 + dk, \text{ lorsque } d \leq k, \text{ sinon } \rightarrow s_{12j} = 0, \quad (\text{III.3})$$

où  $d$  est la distance entre les communautés  $x_1$  et  $x_2$  pour le descripteur  $j$  et  $k$  est un paramètre déterminé *a priori* pour chaque descripteur  $j$ , qui indique la valeur maximale de  $d$  pour laquelle la similarité partielle  $s_{12}$  est différente de zéro. La valeur de  $k$  peut différer selon le descripteur  $j$ , car elle dépend des valeurs d'abondance codées pour  $j$ .

Cet indice est utilisé dans cette étude pour quantifier la similitude entre les différentes catégories de la variable nominale *Groupement d'essences* ainsi qu'entre les catégories des variables ordinales *Classe de densité*, *Classe de hauteur* et *Classe d'âge*. Parmi les nombreux indices de similarité écologique, l'indice  $S$  est choisi parce qu'il ne prend pas en compte les doubles absences, c'est-à-dire que le fait qu'une essence soit absente de deux catégories du *Groupement d'essences* n'est pas considéré dans la mesure de similitude (Legendre et Legendre 1998).

### **Calcul de $S$ pour la variable *Groupement d'essences*, $S_{GE}$**

L'indice  $S$  pour la variable nominale *Groupement d'essences* ( $S_{GE}$ ) mesure la similitude entre les catégories ou classes que pour cette variable présentent les peuplements adjacents à chaque frontière. Pour ce faire, chaque catégorie ou classe de *Groupement d'essences* est traitée comme une communauté écologique, définie par l'abondance relative de l'essence principale et de l'essence secondaire (Ministère des Ressources naturelles du Québec

1995). Ainsi, chaque essence est un descripteur ( $j$ ) et le pourcentage de la surface du peuplement qui est couverte par l'essence est considéré une mesure de son abondance. Par conséquent, le calcul de  $S_{GE}$  consiste à mesurer, pour deux catégories de *Groupement d'essences*, la similarité entre les pourcentages de couvert des essences présentes dans les deux catégories. Cette similarité entre les pourcentages de couvert d'une même essence dans deux catégories différentes provient d'une matrice de similarité partielle (équation III.2).

Les pourcentages de couvert des essences sont, en fait, des intervalles de couvert de la surface terrière du peuplement, selon la définition des groupements d'essences des strates forestières (Tableau 5) (Ministère des Ressources naturelles du Québec 1995). À partir de ces définitions, il est possible d'estimer le pourcentage de couvert de chacune des essences principales et secondaires dans les catégories où elles interviennent (ce sont les intervalles de couvert comparés aux Tableaux 6, 7 et 8). Il faut souligner que les limites des intervalles de couvert du Tableau 5 découlent de la classification qui applique l'interprète lors de la photointerprétation et que le résultat est loin d'être précis car il repose sur l'évaluation subjective des pourcentages de couvert de chaque essence. Pour ce motif les valeurs inférieures et supérieures de chacun des intervalles de couvert ont été arrondis dans le Tableau 5, car ils servent surtout comme une guide lors de la photointerprétation.

Pour chaque catégorie de *Groupement d'essences*, outre l'essence principale et l'essence secondaire, l'intervalle de couvert des résineux est inclus dans cette étude afin de différencier les catégories résineuses, mixtes à prédominance résineuse, mixtes à prédominance feuillue et feuillues (Tableau 5). Cette distinction est importante du point de vue écologique, car elle concerne des types très différents de végétation et parce qu'elle ne peut être incorporée à la comparaison entre catégories à partir de l'essence principale et secondaire seulement.

Tableau 5. Catégories de la variable *Groupement d'essences (GE)*. Intervalle de couvert : pourcentage de la surface terrière du peuplement occupé par une essence ou type de végétation. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).

Catégorie de <i>GE</i>	Intervalle de couvert		
	Résineux	Essence principale	Essence secondaire
Bb, Bj, FNC	0 – 25	37 – 100	
Fi (Bb + Pe)	0 – 25	18 – 50	
F	0 – 25		
BbS, BjS, PeS	25 – 50	25 – 75	13 – 50
BjR (R=S+E+Pb)	25 – 50	25 – 75	4 – 16
FiS	25 – 50	12 – 37	13 – 50
M	25 – 75		
SBb, SBj, SPe	50 – 75	25 – 75	13 – 50
SFi (Fi= Bb + Pe)	50 – 75	25 – 75	6 – 24
SS, EE	75 – 100	56 – 100	
SE, ES, EPb	75 – 100	37 – 75	19 – 36
R	75 – 100		

La matrice de similarité partielle, élaborée *ad hoc*, contient les valeurs de similarité partielle ( $s_{12j}$ ) entre les pourcentages de couvert d'une même essence  $j$  en deux catégories différentes de *Groupement d'essences*. La similarité partielle concerne donc chaque essence dans les différentes catégories où elle intervient. Lorsque l'essence intervient seulement dans une des deux catégories à comparer, la valeur de  $s_{12j}$  est 0 pour cette essence ou descripteur.

Les valeurs de similarité partielle sont habituellement fixées par les experts, de façon à mieux décrire la similitude entre les états d'un descripteur lorsqu'ils sont nominaux ou partiellement ordinaux (Legendre et Legendre 1998). Dans le même souci de bien décrire la similitude entre les catégories de *Groupement d'essences*, dans cette étude la similarité partielle est estimée en considérant deux composantes, contribuant chacune à 50 % de la valeur : (1) la distance entre les moyennes des intervalles de couvert et (2) le degré de chevauchement des intervalles. De cette façon, lorsque la distance entre les moyennes de chaque intervalle de couvert est faible et le degré de chevauchement des intervalles est important, la similitude entre l'importance de l'essence dans les deux catégories est élevée. Le choix de deux composantes à considérer se base sur le fait que les pourcentages de couvert d'une même essence dans les différentes catégories montrent un chevauchement

important et variable (Tableau 5). Les détails du calcul de la similarité partielle sont exposés à l'Annexe II.

En raison de la ressemblance des intervalles de couvert de certaines essences dans les catégories de *Groupement d'essences* (par exemple S, E et Pb, Tableau 5), une même matrice de similarité partielle est élaborée pour ces essences. Ainsi, une seule et même matrice correspond aux trois essences résineuses S, E et Pb (Tableau 6), alors qu'une autre correspond aux essences feuillues Bb et Pe (Tableau 7) et une troisième correspond à l'essence feuillue Bj (Tableau 8).

Tableau 6. Matrice de similarité partielle entre les intervalles de couvert pour les essences S, E et Pb. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).

Intervalle de couvert	4,3-16,3	13-50	19,5-36,75	25-75	25-50	37,5-75	56,25-100
4,3-16,3	1,00						
13-50	0,09	1,00					
19,5-36,75	0,06	0,59	1,00				
25-75	0,03	0,26	0,16	1,00			
25-50	0,04	0,45	0,28	0,34	1,00		
37,5-75	0,02	0,14	0,04	0,54	0,19	1,00	
56,25-100	0,01	0,02	0,02	0,15	0,02	0,18	1,00

Tableau 7. Matrice de similarité partielle entre les intervalles de couvert pour les essences Bb et Pe. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).

Intervalle de couvert	6,5-24,5	12,5-37	13-50	18,75-50	25-75	37,5-100
6,5-24,5	1,00					
12,5-37	0,24	1,00				
13-50	0,20	0,49	1,00			
18,75-50	0,12	0,35	0,71	1,00		
25-75	0,03	0,14	0,25	0,30	1,00	
37,5-100	0,02	0,02	0,1	0,11	0,30	1,00

Tableau 8. Matrice de similarité partielle entre les intervalles de couvert pour l'essence Bj. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence. Information dérivée de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).

Intervalle de couvert	13-50	25-75	37,5-100
13-50	1,00		
25-75	0,24	1,00	
37,5-100	0,10	0,30	1,00

Le Tableau 9 illustre un exemple du calcul de  $S_{GE}$  entre les catégories SE et BbS. Les valeurs de  $s_{12j}$  sont établies pour chaque essence à partir de leurs matrices de similarité partielle (le Tableau 6 pour les essences S et E, le Tableau 7 pour l'essence Bb).

Tableau 9. Calcul de  $S_{GE}$  entre les catégories SE et BbS. La valeur de  $s_{12j}$  est extraite des matrices de similarité partielle de chaque descripteur. Intervalle de couvert : pourcentage de la surface terrière du peuplement couvert par l'essence, dérivé de la norme de stratification forestière du Québec (Ministère des Ressources naturelles du Québec 1995).

Descripteur $j$	Intervalle de couvert de $j$ dans les catégories		SSE-BbS	WSE-BbS
	SE	BbS		
R	75-100	25-50	0,28	1
S	37,5 – 75	13 – 50	0,14	1
E	19,5 – 36,5	0	0	1
Bb	0	25 – 74	0	1
$S_{SE-BbS} = [(0,28*1) + (0,14*1) + (0*1) + (0*1)] / 4 = 0,105$				

### Calcul de $S$ des autres variables

L'indice  $S$ , calculé de façon indépendante pour les variables *Classe de hauteur*, *Classe de densité* et *Classe d'âge*, revient à la similarité partielle  $s_{12j}$ , car pour chacune des variables il y a seulement un descripteur. La similarité partielle est estimée en fonction de  $d$  et  $k$  (équation III.3). La valeur de  $k$  est fixée à la valeur maximale de  $d$ , elle est donc différente pour chacune des trois variables. Les variables ordinales sont d'abord transformées en échelle quantitative en divisant la valeur moyenne de l'intervalle de classe de chaque catégorie par son étendue. Cette transformation est appliquée afin de préserver l'ordre des classes mais en considérant que les classes présentent des étendues différentes. Un exemple du calcul de  $S_{CD}$  illustre la procédure (Tableau 10).

Tableau 10. Calcul de  $S_{CD}$  pour les catégories A et C. La valeur de  $k$  est égal à celle de  $d$  entre les classes A et D (voir l'équation III.3).

Valeur de $CD$ pour la catégorie		$d$	$K$
A (81-100% couvert)	C (41-60% couvert)		
$90,5/20 = 4,52$	$50,5/20 = 2,52$	$4,52 - 2,52 = 2$	2,36
$S_{CD:A-C} = S_{A-C} = [2 (2,36 + 1 - 2)] / (2*2,36 + 2 + 2,36*2) = 0,237$			

L'indice  $S$  est calculé pour chacune des 8018 frontières (déterminées par photo-interprétation) de la carte forestière de 1992, dont les peuplements adjacents présentent des valeurs pour les variables considérées. Pour chaque frontière,  $S$  est calculé en comparant les catégories présentes dans les polygones adjacents à la frontière et ce, pour ces quatre variables : le *Groupement d'essences (GE)*, la *Classe de densité (CD)*, la *Classe de hauteur (CH)* et la *Classe d'âge (CA)*.  $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$  et  $S_{CA}$  sont les indices  $S$  respectifs de ces quatre variables. Un indice de similarité globale,  $S_G$ , dont les valeurs s'étendent entre 0 et 4, synthétise la similarité des quatre variables :

$$S_G = S_{GE} + S_{CH} + S_{CD} + S_{CA}. \quad (\text{III.4})$$

### Présentation et analyse des valeurs résultantes de $S_{GE}$ , $S_{CD}$ , $S_{CH}$ , $S_{CA}$ et $S_G$

Afin de simplifier l'interprétation des résultats, les valeurs résultantes de  $S$  pour les différentes variables sont classifiées en niveaux de contraste, lesquels sont présentés à la section des résultats.

#### 2.1.2. Quantification du contraste à différentes échelles : la force des frontières à des voisinages de différente taille

Une autre façon de mesurer le contraste entre deux peuplements adjacents pour une variable donnée est par la force de la frontière qui les sépare et à différentes échelles de voisinage (Csillag *et al.* 2001). La force des frontières est calculée au moyen d'une mesure locale, pour chacune des 8018 frontières de la carte forestière de 1992 dont les polygones adjacents présentent des valeurs pour les variables *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*. Ces variables sont transformées en échelle quantitative

en divisant la moyenne de l'intervalle de classe de chaque catégorie par son étendue, de façon similaire au calcul de l'indice  $S$ .

### L'algorithme

L'algorithme utilisé pour estimer la mesure de la force de chaque frontière a été proposé par Csillag *et al.* (2001). La force d'une frontière est déterminée par la différence des valeurs d'une variable quantitative des peuplements (polygones sur la carte) adjacents à la frontière. Cette différence de valeurs est exprimée en percentile de l'ensemble des frontières. La force d'une frontière particulière indique donc son rang parmi la totalité des frontières, en terme de la différence des valeurs d'une variable entre les peuplements adjacents à la frontière ou entre les groupes des peuplements, selon la taille du voisinage considéré. La taille du voisinage indique l'échelle à laquelle la force de la frontière est mesurée.

Ces concepts sont explicités dans un algorithme (Csillag *et al.* 2001) qui considère d'abord la définition d'une région d'étude,  $R$  et la définition de  $P = \{p_1, p_2, \dots, p_n\}$  comme un ensemble de polygones qui ne se superposent pas entre eux et qui couvrent la région, donc

$$R = p_1 \cup p_2 \cup \dots \cup p_n. \quad (\text{III.5})$$

Dans ce contexte, la valeur d'une variable quantitative du polygone  $p_i$  est  $a_i$  et  $b_{ij}$  constitue la frontière entre les polygones  $p_i$  et  $p_j$ . L'algorithme, qui prend en compte d'abord les polygones (étapes 1 à 3) et ensuite les frontières (étapes 4 et 5), comprend les étapes suivantes :

1. Un voisinage  $D$  est défini pour le polygone  $p_i$  en établissant une distance maximale topologique  $\delta$ . Si la distance topologique entre les polygones  $p_i$  et  $p_j$  est  $|p_i - p_j|$ , l'appartenance du polygone  $p_j$  au voisinage du polygone  $p_i$  est déterminée par la relation suivante :

$$D(p_i) = \{p_j : \|p_i - p_j\| \leq \delta; p_j \in P\}, \quad (\text{III.6})$$

selon laquelle le polygone  $p_j$  appartient au voisinage du polygone  $p_i$  lorsque la distance entre  $p_j$  et  $p_i$  est égale ou inférieure à la distance topologique  $\delta$ .

2. Un facteur de pondération  $w_{ij} \geq 0$  indique une association spatiale entre le polygone  $p_i$  et tous les polygones de  $D(p_i)$ , incluant  $p_i$ . Dans cette étude le facteur de pondération est 1, ce qui signifie que les valeurs des tous les polygones du voisinage ont la même importance.

3. Une statistique locale  $m_i$  est définie pour chaque polygone  $p_i$ .  $m_i$  est la somme pondérée par  $w_{ij}$  des valeurs de la variable quantitative  $a_i$  de tous les polygones qui appartiennent au voisinage  $D(p_i)$  :

$$m_i = \sum_{g=1}^{n_D} w_{ig} * a_{ig}, \quad p_g \in D(p_i), \quad (\text{III.7})$$

où  $n_D$  est le nombre de polygones dans  $D(p_i)$ .

4. Une statistique de différence  $d_{ij}$  est définie pour la frontière  $b_{ij}$  entre  $p_i$  et  $p_j$  :

$$d_{ij} = (m_i - m_j)^2. \quad (\text{III.8})$$

5. La statistique de différence  $d_{ij}$  est exprimée en percentile,  $q_{ij}$ , indiquant le rang de  $b_{ij}$  par rapport au reste des frontières.  $q_{ij}$  représente la particularité ou la force de la frontière à l'échelle d'analyse établie par la distance topologique  $\delta$ .

6. Les étapes 1 à 6 sont répétées en utilisant des valeurs croissantes de  $\delta$ .

Un exemple concret de l'application de l'algorithme est présenté à l'Annexe II.

### **Définition des voisinages**

Le voisinage peut-être défini en fonction de la distance métrique ou topologique, et ce choix influence fortement les résultats (Csillag *et al.* 2001). Dans cette étude le voisinage est défini sur la base de la distance topologique car les peuplements présentent des formes et des tailles très diverses. Cette façon de définir le voisinage permet de « contrôler » le degré d'adjacence des peuplements mais les distances métriques peuvent être très variables.

Cinq niveaux de voisinage sont considérés avec une distance topologique  $\delta$  qui varie entre 0 et 4 (Figure 5). Le niveau 1 est constitué par chaque polygone individuel ( $\delta = 0$ ). Pour sa part, le niveau 2 est constitué des polygones adjacents à chaque polygone, soit un voisinage de premier ordre ( $\delta = 1$ ), alors que le niveau 3 est constitué des polygones adjacents à ceux du niveau antérieur, soit un voisinage de deuxième ordre ( $\delta = 2$ ). Les niveaux 4 et 5 contiennent le même critère en déterminant un voisinage de troisième et quatrième ordre respectivement ( $\delta = 3$  et  $\delta = 4$ ).

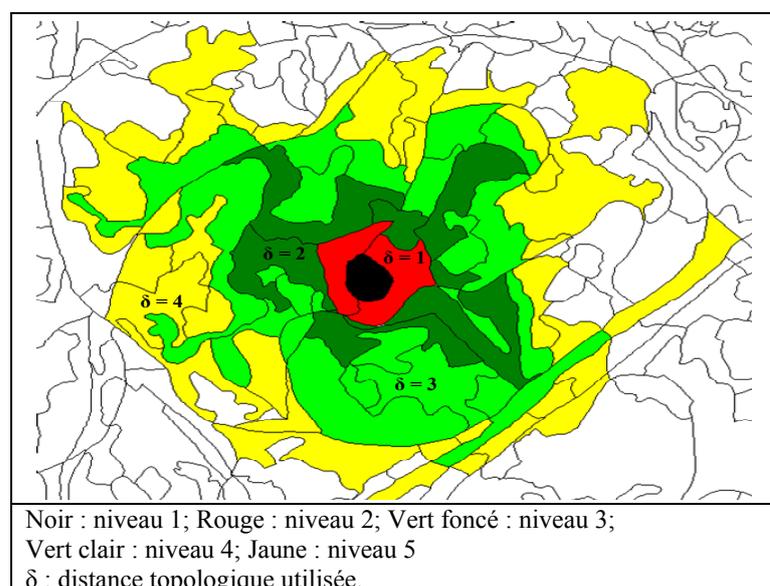


Figure 5. Voisinages considérés pour quantifier la force de frontières.

### Présentation et analyse des résultats

La mesure de la force étant exprimée en percentiles, les valeurs de  $q_{ij}$  s'étendent entre 0 et 1. Afin de faciliter l'interprétation des résultats, les valeurs de  $q_{ij}$  sont classifiées en catégories (Tableau 11).

L'interprétation de l'échelle à laquelle une frontière est forte tient compte des catégories de force dans les différents niveaux de voisinage. Pour chaque variable, les frontières indifférentes sont celles dont la force ne change pas lorsque la taille du voisinage varie. Les

Tableau 11. Catégories de force d'une frontière.

Catégories de force	Intervalle de valeurs de $q_{ji}$
Très faible	0,00 – 0,20
Faible	0,21 – 0,40
Moyenne	0,41 – 0,60
Forte	0,61 – 0,80
Très forte	0,81 – 1,00

frontières « localement fortes » sont celles très fortes ou fortes pour les voisinages de niveau 1 ou 2 et moyennes à très faibles pour les voisinages d'ordre supérieur (selon les catégories de force du Tableau 11). Les frontières « régionales » sont celles fortes ou très fortes pour les voisinages de niveau 4 et 5 et moyennes à très faibles pour ceux d'ordre inférieur.

Les résultats présentent les proportions de frontières indifférentes, localement fortes et régionales pour les variables isolées et en combinaisons de deux et de trois variables.

## 2.2. Quantification du contexte

### 2.2.1. Le contexte des frontières

Le but de cette analyse est d'établir la localisation et l'étendue des zones des frontières qui présentent des valeurs « anormales » de contraste (quantifié par  $S$ ) pour une variable donnée, c'est-à-dire des valeurs de  $S$  très différentes de la valeur moyenne pour l'ensemble des frontières. Il consiste également à déterminer si ces frontières sont de contraste faible ou élevé. L'expression « contexte d'une frontière » fait donc référence à la ressemblance entre la valeur de  $S$  de la frontière et les valeurs de  $S$  des frontières voisines.

Afin de caractériser cette ressemblance, une mesure d'autocorrélation spatiale locale de  $S$  est calculée pour chacune des 8018 frontières de la carte forestière de 1992. Puisque les statistiques locales sont développées pour des données ponctuelles, chaque frontière est représentée par le point central de la ligne, auquel est assignée la valeur de  $S$  de chaque variable, soit  $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$ ,  $S_{CA}$  et  $S_G$ .

L'autocorrélation spatiale locale de  $S$  est estimée au moyen de l' $I$  de Moran local, pondéré par la distance inverse entre les points et ajusté pour les distances faibles (Anselin 1995, Levine *et al.* 2002). Les valeurs d'autocorrélation locale significative et positive indiquent une ressemblance dans l'écart par rapport à la moyenne de  $S$  entre les frontières proches. Les valeurs significatives et négatives indiquent pour leur part que les valeurs de  $S$  des frontières proches sont différentes en ce qui concerne leur écart de la moyenne (Boots 2002). Suivant le critère proposé par Levine *et al.* (2002), les valeurs d'autocorrélation spatiale locale de  $S$  sont significatives lorsqu'elles appartiennent à l'intervalle (positif ou négatif) entre un et trois écarts-type de la moyenne. Les valeurs au-delà de trois écarts-type sont considérées extrêmes.

Lors de l'analyse, la carte des valeurs de l' $I$  de Moran local, où chaque point représente une frontière, est superposée à la carte des niveaux de contraste des frontières pour chaque variable, afin de disposer de deux types d'information.

### **2.3. Le contexte des peuplements**

Le but de cette analyse est d'identifier des frontières entre les zones de différent degré d'hétérogénéité pour une variable particulière. L'hétérogénéité concerne le contexte de chacun des peuplements. Ce contexte se définit comme l'ensemble des peuplements adjacents au peuplement, c'est-à-dire son voisinage de premier degré.

#### **2.3.1. L'indice $H$ d'hétérogénéité du contexte d'un peuplement**

L'hétérogénéité du contexte d'un peuplement est quantifiée au moyen de l'indice  $H$  propre au peuplement. Cet indice mesure, pour la variable d'intérêt, la proportion de voisins dont la valeur de la variable coïncide avec celle du peuplement en question. Les valeurs possibles de  $H$  s'étalent entre 0 (aucun des peuplements voisins coïncide avec le peuplement considéré), qui indique l'hétérogénéité maximale, et 1 (tous les peuplements voisins coïncident avec le peuplement), qui indique l'hétérogénéité minimale.

L'indice  $H$  est estimé pour le *Groupement d'essences* ( $H_{GE}$ ), la *Classe de densité* ( $H_{CD}$ ), la *Classe de hauteur* ( $H_{CH}$ ) et la *Classe d'âge* ( $H_{CA}$ ). Pour synthétiser les quatre variables, l'*Hétérogénéité globale* ( $H_G$ ) correspond à la somme des indices des variables :

$$H_G = H_{GE} + H_{CD} + H_{CH} + H_{CA}. \quad (\text{III.9})$$

Les valeurs de  $H_G$  s'étendent entre 0 et 4, l'hétérogénéité maximale et minimale respectivement. Afin de faciliter l'interprétation des résultats, les valeurs de  $H$  sont reclassées en cinq degrés d'hétérogénéité (Tableau 12).

Tableau 12. Classification de l'indice  $H$  d'hétérogénéité du contexte des peuplements en degrés d'hétérogénéité.

Degré d'hétérogénéité	$H_{GE} - H_{CD} - H_{CH} - H_{CA}$	$H_G$
Très haut	0,00 – 0,20	0,00 – 0,80
Haut	0,20 – 0,40	0,80 – 1,60
Moyen	0,40 – 0,60	1,60 – 2,40
Bas	0,60 – 0,80	2,40 – 3,20
Très bas	0,80 – 1,00	3,20 – 4,00

## 2.4. Évaluation de la relation entre le contraste et le contexte, et la probabilité d'existence d'une frontière

Le but de cette analyse est d'évaluer la relation entre le contraste des variables, ainsi que le contexte d'une frontière et sa probabilité d'existence.

### 2.4.1. La probabilité d'existence d'une frontière

La probabilité d'existence d'une frontière est estimée en superposant trois cartes forestières multitemporelles et en analysant les frontières individuelles de la carte résultante (De Groeve 1999). Les cartes forestières de 1973, 1984 et 1992 sont superposées et les cas de coïncidence de frontières pour une, deux ou trois cartes sont déterminés. Pour tenir compte de l'incertitude spatiale des frontières individuelles, une bande d'erreur est établie autour de chaque frontière. La largeur de cette bande d'erreur est l'erreur quadratique moyenne (RMSE) associée à l'erreur positionnelle de la cartographie forestière du Québec (Leckie et

Gillis 1995). La largeur de la bande est de 35 m, 30 m et 25 m pour les cartes de 1973, 1984 et 1992 respectivement.

Les frontières de la carte forestière de 1992 sont classifiées comme de « haute consistance » lorsque les bandes d'erreur des trois cartes se superposent, de « consistance moyenne » lorsque deux des trois bandes d'erreur se superposent et de « basse consistance » lorsque aucune des bandes d'erreur ne se superpose. Étant donné que les bandes d'erreur ne se superposent pas nécessairement sur toute la longueur de la frontière, le degré de consistance assigné à une frontière est le degré présent sur au moins 35 % de la longueur de la frontière. La probabilité d'existence d'une frontière résulte donc en une variable qualitative, *Consistance*, qui comprend trois degrés : haut, moyen et bas.

#### **2.4.2. Relation entre contraste, contexte et consistance d'une frontière**

Pour évaluer si les degrés de *Consistance* présentent une fréquence particulière de valeurs de contraste, les distributions de fréquences de  $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$  et  $S_{CA}$  sont analysées pour chaque degré de *Consistance* ( $n=8018$  frontières). La signification des différences entre les distributions de fréquences des valeurs de  $S$  est estimée au moyen du test de Kolmogorov-Smirnov pour deux échantillons.

De plus, la relation entre le contraste et le contexte, et la *Consistance* des frontières, est estimée au moyen de l'analyse discriminante (Klecka 1980). L'objectif de cette analyse est d'identifier les variables de contraste et de contexte les plus importantes pour discriminer les trois degrés de la variable réponse *Consistance*.

Les variables indépendantes de l'analyse discriminante représentent : (1) le contraste (les variables sont  $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$  et  $S_{CA}$ ), (2) le contexte de frontières, donné par les valeurs de l' $I$  de Moran locale de  $S$  (les variables sont  $IMLS_{GE}$ ,  $IMLS_{CD}$ ,  $IMLS_{CH}$  et  $IMLS_{CA}$ ) et (3) la longueur de la frontière en mètres (la variable est *Longueur*). Cette dernière variable est incluse dans l'analyse en raison de son importance dans les modèles d'incertitude positionnelle des frontières (De Groeve 1999). Les interactions entre les variables de

contraste ( $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$  et  $S_{CA}$ ) et de contexte ( $IMLS_{GE}$ ,  $IMLS_{CD}$ ,  $IMLS_{CH}$  et  $IMLS_{CA}$ ), ainsi qu'entre les variables de contraste et *Longueur* sont également considérées dans le modèle.

Étant donné que 51 % des frontières appartiennent au degré moyen de *Consistance*, l'analyse est faite sur un échantillon de 4545 frontières, dans lequel chaque degré de *Consistance* est représenté avec le même nombre de frontières (les frontières du degré moyen de *Consistance* sont sélectionnées aléatoirement). L'analyse utilise la méthode pas à pas où le critère d'incorporation d'une variable au modèle est une probabilité de  $F$  de 0,05 tandis que le critère d'exclusion d'une variable du modèle est une probabilité de  $F$  de 0,10. La classification des frontières en degrés de *Consistance* est faite en utilisant les variables discriminantes (Klecka 1980).

## Résultats et discussion

Les résultats sont présentés en suivant le même ordre que les méthodes exposées dans la section précédente. Premièrement, les résultats de la quantification du contraste de frontières sont détaillés, au moyen de l'indice  $S$  et de la force d'une frontière à différentes échelles. Les résultats concernant la quantification du contexte, tant des frontières que des peuplements, sont présentés par la suite. L'évaluation du lien entre le contraste, le contexte et la probabilité d'existence d'une frontière est présentée en dernier.

## 2. Quantification du contraste

### 1.1. Le contraste écologique quantifié au moyen de l'indice $S$

#### 1.1.1. Les valeurs de $S$ des frontières

L'indice  $S$ , calculé pour les variables *Groupement d'essences* ( $S_{GE}$ ), *Classe de densité* ( $S_{CD}$ ), *Classe de hauteur* ( $S_{CH}$ ) et *Classe d'âge* ( $S_{CA}$ ), s'étend théoriquement entre 0, similarité nulle, et 1 lorsque les peuplements adjacents à la frontière présentent la même catégorie pour la variable considérée.

Les valeurs obtenues pour les frontières de la carte forestière de 1992 n'atteignent pas les extrêmes théoriques de  $S$  sauf lorsque les catégories des deux peuplements adjacents sont identiques. Dans les cas de non égalité des catégories dans les peuplements adjacents, la valeur minimale (0,02) correspond à  $S_{GE}$  et la valeur maximale (0,96) correspond à  $S_{CH}$ . L'ensemble des valeurs de similarité obtenues est présenté à l'Annexe II (Tableau 56, Tableau 57, Tableau 58 et Tableau 59). Quant aux valeurs obtenues pour  $S_G$  elles s'étendent entre 0,88 et 4.

Pour simplifier l'interprétation subséquente des résultats, les valeurs de  $S$  sont reclassées en six niveaux de contraste selon des échelles adaptées aux valeurs obtenues pour chacune des variables (Tableau 13). Les valeurs obtenues pour  $S_{GE}$  et pour  $S_{CA}$  lorsque les peuplements adjacents sont différents, ne dépassent pas 0,60. Si l'échelle avec des intervalles de 0,20 est appliquée pour ces variables, toutes les frontières sont de contraste moyen à très haut. Afin de réduire cet effet, les intervalles pour ces variables sont fixés en 0,15. Il faut retenir aussi que  $S_G$  résulte de l'addition, pour chaque frontière, des valeurs de  $S$  calculées pour les quatre variables.

Tableau 13. Niveaux de contraste et intervalles de valeurs de  $S$  pour chacune des variables ( $n = 8018$  frontières entre 3729 peuplements).

Contraste	$S_{GE}$	$S_{CD}$	$S_{CH}$	$S_{CA}$	$S_G$
Très haut	0,00 – 0,15	0,00 – 0,19	0,00 – 0,19	0,00 – 0,15	0,00 – 0,80
Haut	0,16 – 0,30	0,20 – 0,39	0,20 – 0,39	0,16 – 0,30	0,81 – 1,60
Moyen	0,31 – 0,45	0,40 – 0,59	0,40 – 0,59	0,31 – 0,45	1,61 – 2,40
Bas	0,46 – 0,60	0,60 – 0,79	0,60 – 0,79	0,46 – 0,60	2,41 – 3,20
Très bas		0,80 – 0,99	0,80 – 0,99		3,21 – 3,99
Sans contraste	1,00	1,00	1,00	1,00	1,00

Au-delà des valeurs de  $S$  obtenues, l'intérêt porte sur leurs proportions dans la carte forestière de 1992. Les distributions de fréquence des  $S$  observées selon les niveaux de contraste varient en fonction de la variable considérée (Figure 6). La proportion de frontières correspondant au niveau « sans contraste » (Tableau 13) est assez similaire pour les quatre variables, entre 35 et 40 % des frontières. Cependant, pour le *Contraste global*, la proportion de frontières « sans contraste » est réduite à 12 %. Ce niveau « sans contraste » s'explique par le fait que les peuplements sont identifiés par des différences d'une ou de

plusieurs variables, parmi lesquelles seules quatre des plus importantes sont considérées dans cette étude (Bérard et Côté 1996, Ministère des Ressources naturelles du Québec 1995). Les frontières sans contraste pour les quatre variables considérées dans cette étude correspondent à des peuplements qui diffèrent par rapport à la pente, la perturbation, l'origine, la forme du terrain, la qualité de la station ou le stade de développement de la végétation.

En ce qui concerne les frontières qui présentent un contraste, pour *GE* les frontières à fort contraste sont les plus fréquentes (Figure 6a) tandis que pour *CD* et *CH* sont les frontières à contraste moyen qui sont plus fréquentes (Figure 6b et c). Pour *CA* les frontières à haut et moyen contraste présentent des proportions similaires (Figure 6d). Si la considération porte sur le *Contraste global*, la majorité des frontières (45 %) présente un contraste faible (bas et très bas) et seulement 11 % présentent un contraste élevé (Figure 6e).

Ces résultats suggèrent que l'identification de la majorité des frontières à la Forêt Montmorency, lors de la photo-interprétation, repose probablement sur le contraste d'une seule variable, c'est-à-dire qu'elle est de caractère local. De ce point de vue, 57 % des frontières sont de fort contraste (haut et très haut) pour l'une ou l'autre des variables.

Les fréquences des niveaux de contraste obtenues sont influencées par l'importance des catégories des variables présentes à la Forêt Montmorency. En effet, 75% des peuplements correspondent à trois catégories de *Groupement d'essences* tandis qu'environ 80 % des peuplements correspondent à deux catégories de *Classe de densité* et de *Classe de hauteur* et à trois catégories de *Classe d'âge* (Voir Figure 53 à l'Annexe II). Pour les trois dernières variables, ces catégories sont intermédiaires et consécutives. Ainsi l'examen des combinaisons de catégories montre que la majorité des frontières sépare des peuplements qui présentent ces catégories : B-C, 3-4 et 30-45 de *CD*, *CH* et *CA* respectivement. Cette homogénéité se reflète dans la proportion majoritaire de frontières qui présentent un faible *Contraste global*.

La proportion importante de frontières à fort contraste pour *GE* peut être attribuée à deux causes. D'une part, 15 catégories de cette variable sont présentes à la Forêt Montmorency

tandis que les autres variables ont entre quatre et sept catégories. Cette différence dans le nombre de catégories à comparer pourrait entraîner pour  $GE$  une plus forte probabilité de

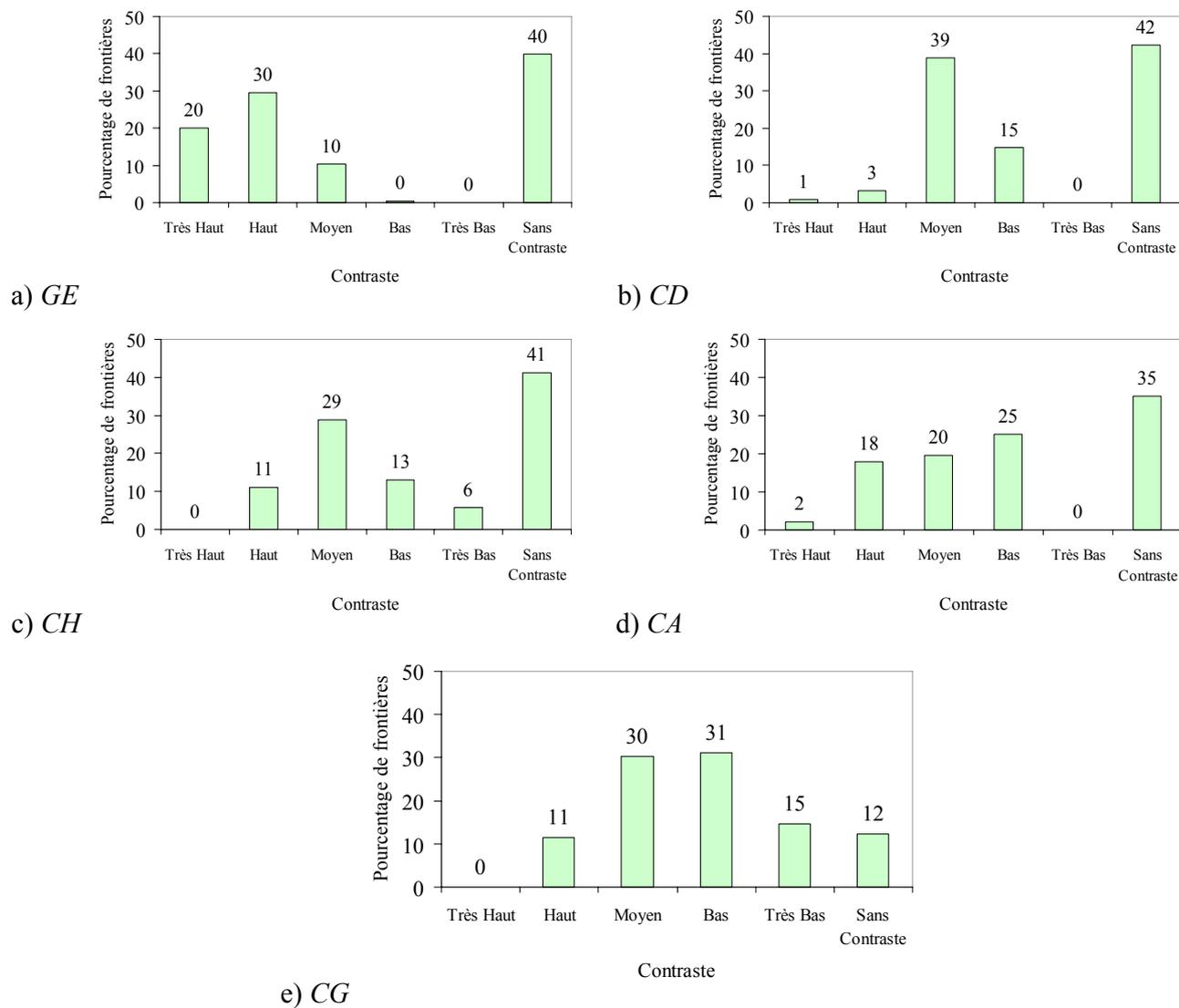


Figure 6. Distribution des fréquences des frontières selon le niveau de contraste. a) *Groupement d'essences*; b) *Classe de densité*; c) *Classe de hauteur*; d) *Classe d'âge*; e) *Contraste global*. ( $n = 8018$  frontières entre 3729 peuplements).

trouver des catégories plus dissemblables. D'autre part, les valeurs de  $S_{GE}$  de deux types de végétation qui semblent *a priori* assez proches sont relativement faibles. Par exemple, la valeur de  $S_{GE}$  de la combinaison SBb-SS est 0,23, une valeur relativement faible en terme

de l'étendue théorique de  $S$  (entre 0 et 1). Cette combinaison correspond à un peuplement où le sapin baumier (S) est l'essence principale et le bouleau à papier est l'essence secondaire (Bb), et qui est adjacent à un peuplement de sapin baumier pur (SS). Il s'agit d'une combinaison très fréquente à la Forêt Montmorency (20 % des frontières). Il est possible que la similarité partielle  $s_{12j}$  sur laquelle  $S_{GE}$  est calculé, surestime les différences entre les deux catégories. Toutefois, la définition des catégories indique qu'un peuplement classifié comme SBb peut présenter jusqu'à 50 % de la surface terrière couverte par le bouleau à papier (Bb), tandis qu'un peuplement classifié comme SS peut présenter jusqu'à 100 % de la surface terrière couverte par le sapin baumier (S) (Tableau 1 et Tableau 5). Ces situations extrêmes de chacune des catégories expliquent la valeur relativement faible de similarité entre les deux.

Des combinaisons des types de végétation « mixte à prédominance feuillue » – « résineux » correspondent à la majorité des frontières de très haut contraste. Parmi elles, la combinaison M-SS – qui correspond à un peuplement dans les premières étapes de son développement, dont la végétation est identifiée comme mixte ou mélangée (M), et qui est adjacent à un peuplement résineux de sapin baumier (SS) – est la plus fréquente (13 % des frontières).

Les résultats obtenus découlent de la méthode appliquée pour quantifier la similarité et des intervalles fixés dans la classification de  $S$  en niveaux de contraste. Il est évident qu'un changement dans la façon d'estimer la similarité partielle ou de fixer les bornes des intervalles de la classification entraîneront des résultats différents. Cependant ceci est un premier essai pour quantifier des différences entre variables qualitatives, qui peut être raffiné et amélioré par l'incorporation d'autres variables variables ou par la comparaison avec d'autres indices de similarité.

### **1.1.2. La répartition spatiale des niveaux de contraste**

La répartition spatiale des niveaux de contraste varie en fonction de la variable considérée, bien que l'analyse visuelle des cartes suggère que, pour toutes les variables, les frontières d'un même niveau tendent à être proches. Par exemple, les frontières de très haut contraste

sont proches de celles de haut contraste et les frontières de bas et très bas contraste exhibent la même tendance (Figure 7). La carte du *Contraste Global* synthétise ces tendances (Figure 7e) et les frontières « faibles » y sont prépondérantes et les frontières « fortes » sont rapprochées.

Comme il a déjà été mentionné, la reconnaissance des frontières par photo-interprétation tient compte d'une ou de plusieurs variables qui varient selon le cas. Pour certaines frontières il peut s'agir du contraste de *GE* et de *CH*, alors que pour d'autres il peut s'agir du contraste de *CA*, etc. La répartition spatiale de ces frontières de fort contraste pour l'une ou l'autre des variables est assez uniforme dans la zone d'étude (Figure 7f).

## **1.2. Le contraste quantifié par la force à différentes échelles**

Une autre manière de considérer le contraste entre les peuplements adjacents à une frontière est de prendre en considération la force de la frontière. Cette dernière mesure la différence entre les valeurs des variables quantitatives des peuplements qui appartiennent au voisinage de la frontière, avec les voisinages prenant différentes tailles. Le plus restreint consiste en deux peuplements adjacents à la frontière, tandis que le plus étendu comprend tous les peuplements jusqu'au voisinage de quatrième degré de chacun des peuplements adjacents à la frontière.

Les frontières considérées « indifférentes » sont celles dont la valeur de force ne change pas en fonction de la taille des voisinages. Les frontières dites « localement fortes » sont fortes ou très fortes dans les voisinages d'ordre 1 ou 2 et moyennes à faibles dans les voisinages d'ordre supérieur. Les frontières dites « régionales » sont fortes ou très fortes dans les voisinages d'ordre 4 et 5 et moyennes à faibles dans ceux d'ordre inférieur.

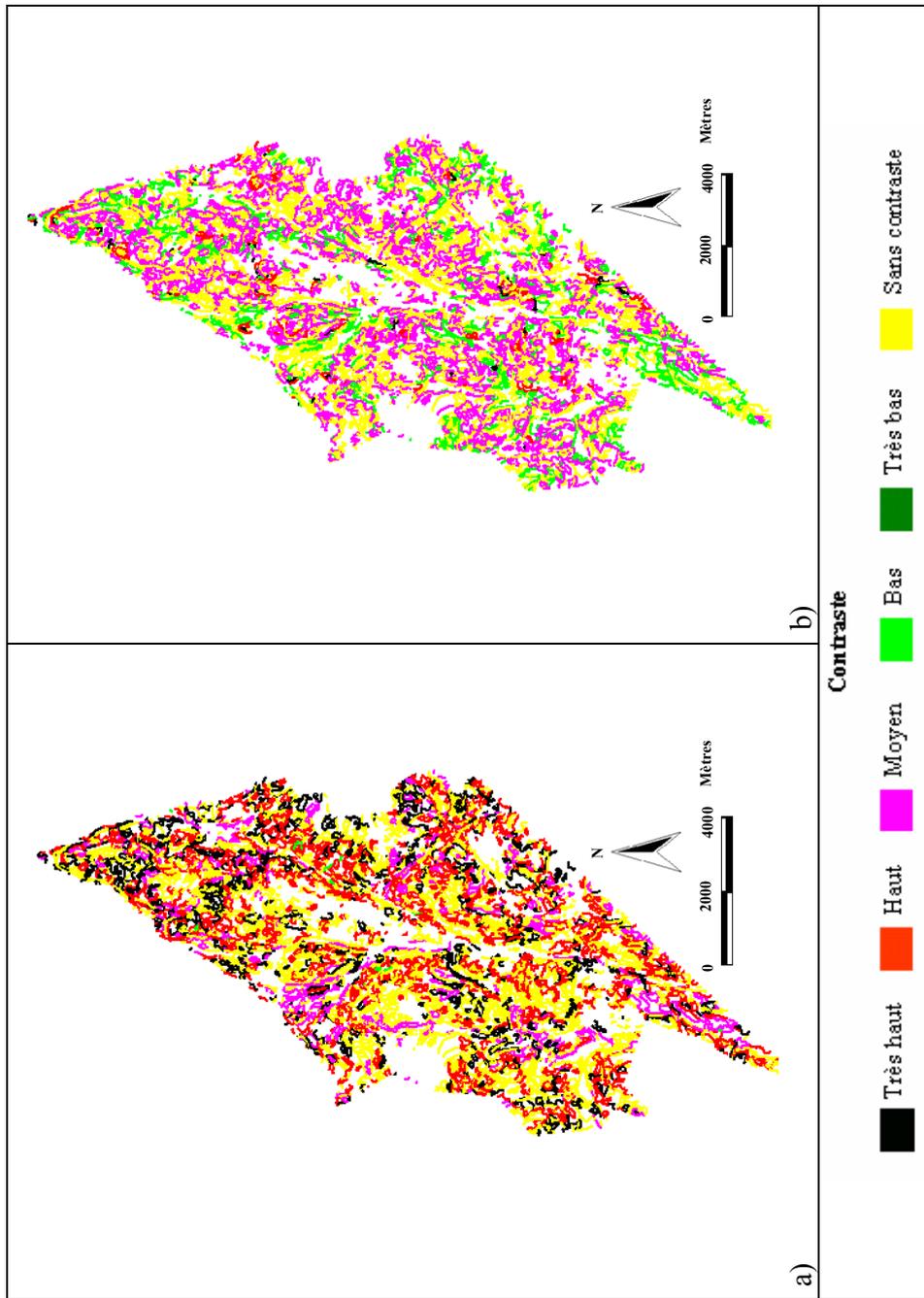


Figure 7. Le contraste des frontières : valeurs de  $S$  reclassées en niveaux de contraste. a) *Groupement d'essences*, b) *Classe de densité* ( $n = 8018$  frontières entre 3729 peuplements).

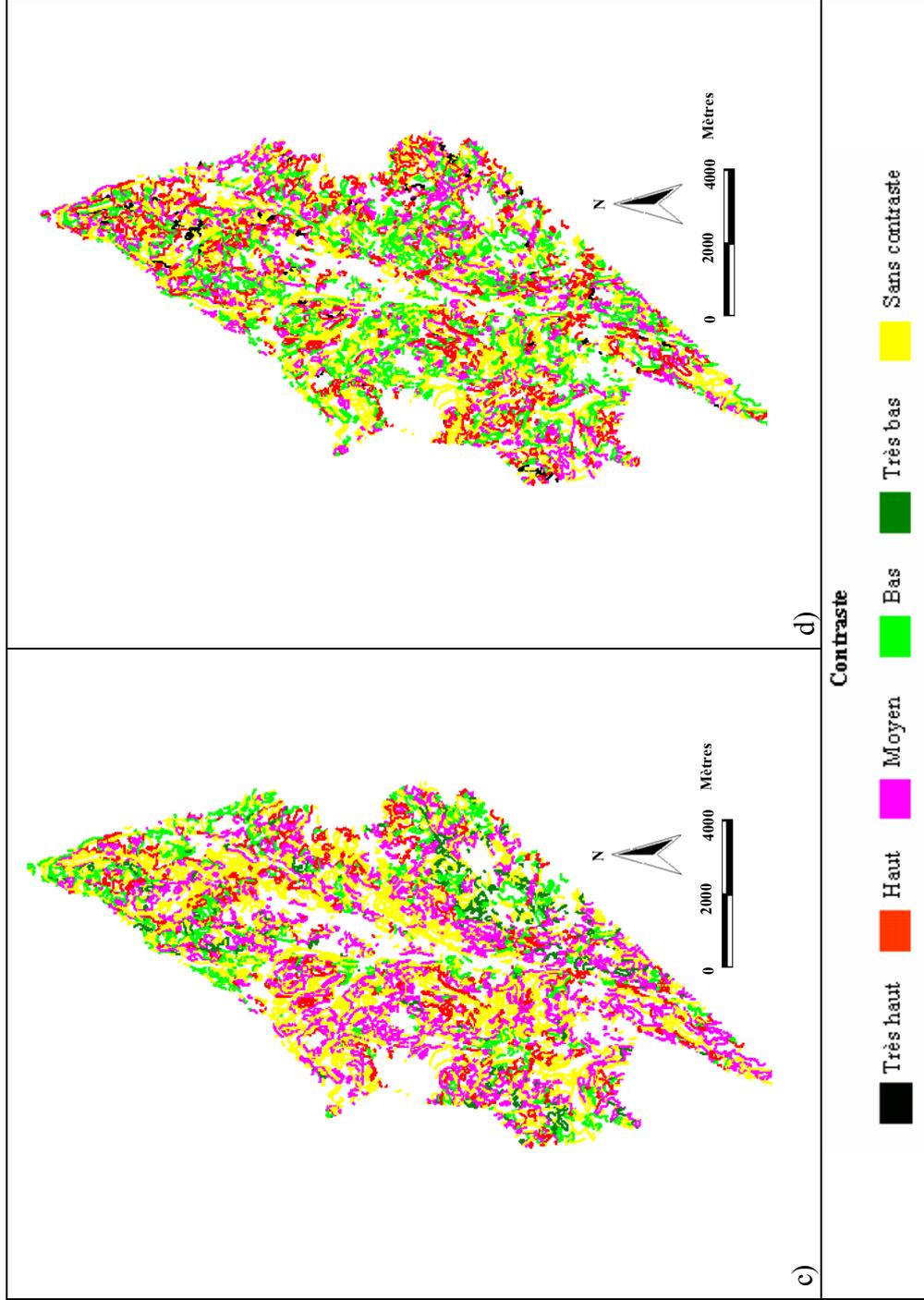


Figure 7. Suite. Le contraste des frontières : valeurs de *S* reclassées en niveaux de contraste. c) *Classe de hauteur*, d) *Classe d'âge* ( $n = 8018$  frontières entre 3729 peuplements).

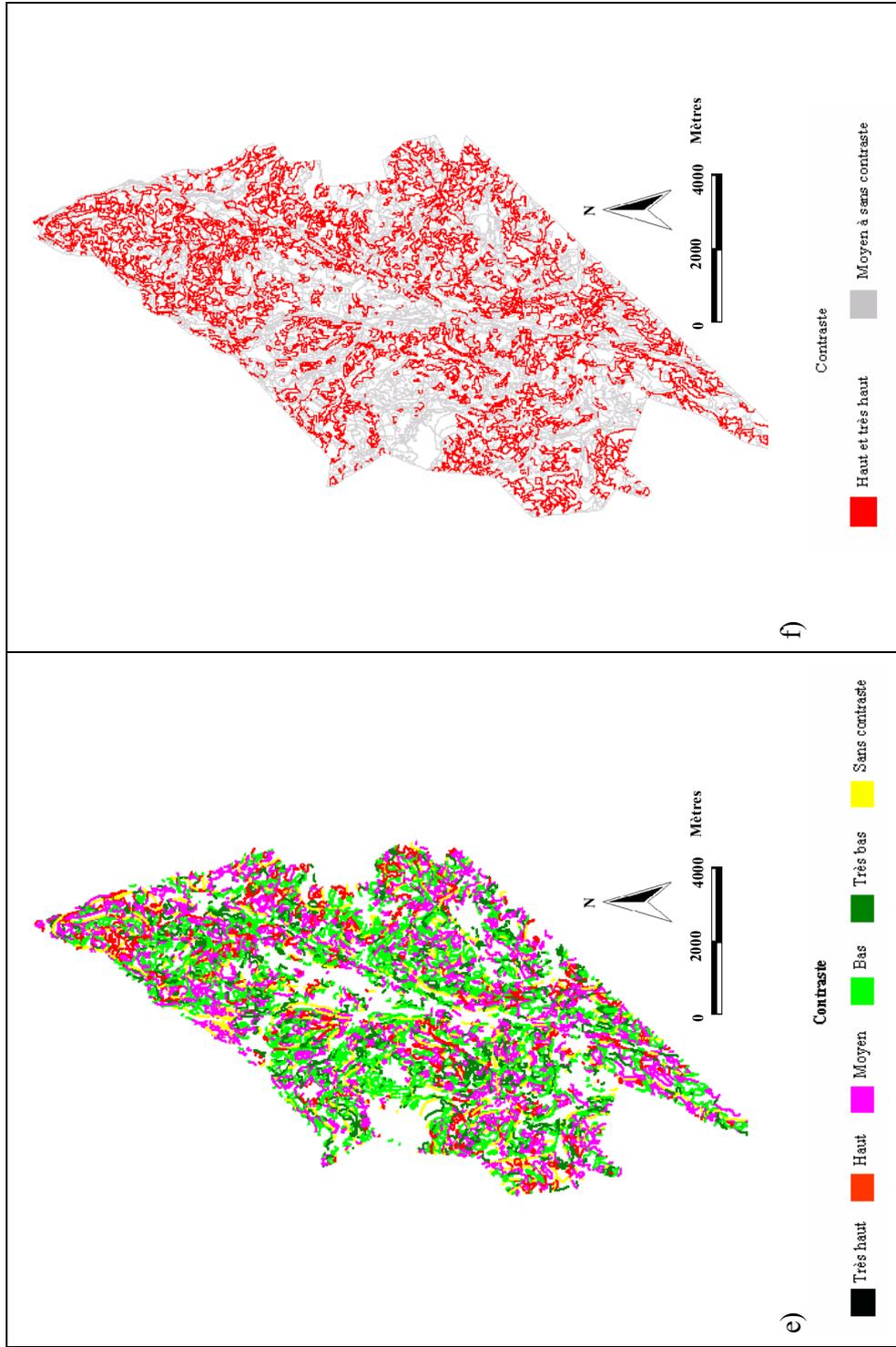


Figure 7. Suite. e) *Contraste global* des frontières : valeurs de  $S_G$  reclassées en niveaux de contraste. f) Les frontières fortes (de haut et très haut contraste) pour l'un ou l'autre des attributs : *Groupeement d'essences* ou *Classe de densité* ou *Classe de hauteur* ou *Classe d'âge* ( $n = 8018$  frontières entre 3729 peuplements).

### 1.2.1. Les valeurs de force résultantes dans la carte forestière de 1992

Les frontières indifférentes sont très minoritaires à la Forêt Montmorency (Tableau 14). La majorité des frontières est très faible pour toutes les combinaisons de variables, cependant la proportion de frontières très fortes augmente à mesure que la combinaison devient moins restrictive (soit d'une ou deux variables).

Tableau 14. Proportion des frontières indifférentes au changement de voisinage selon la catégorie de force, les valeurs sont exprimées en pourcentages ( $n=8018$  frontières).

Catégorie de force	<i>CD-CH-CA</i>	<i>CD-CH</i>	<i>CD-CA</i>	<i>CH-CA</i>	<i>CD</i>	<i>CH</i>	<i>CA</i>
Très forte	0,0	15,0	1,6	6,3	32,8	34,6	7,7
Forte	0,0	0,0	0,0	0,0	0,0	2,6	3,6
Moyenne	0,0	0,0	0,0	0,0	8,6	6,8	4,1
Faible	0,0	0,0	0,0	0,0	0,0	0,0	12,3
Très faible	100,0	85,0	98,3	93,6	58,6	55,9	72,3
% du total de frontières	0,6	0,9	0,7	1,4	3,2	3,3	2,4

La variable *CH* présente la plus grande proportion de frontières localement fortes et très fortes (Tableau 15). Pour les combinaisons de deux variables à la fois, *CH-CA* est la plus fréquente, alors que *CD-CA* est la moins représentée. Les frontières fortes pour la combinaison des trois variables sont très minoritaires.

Tableau 15. Pourcentage de frontières fortes et très fortes à l'échelle locale (c'est-à-dire dans les voisinages de niveau 1 ou 2) et régionale (c'est-à-dire dans les voisinages de niveau 4 ou 5) pour les variables et les combinaison de variables ( $n=8018$ ).

Échelle	<i>CD-CH-CA</i>	<i>CD-CH</i>	<i>CD-CA</i>	<i>CH-CA</i>	<i>CD</i>	<i>CH</i>	<i>CA</i>
Locale	0,6	1,5	0,9	5,3	4,6	10,2	7,0
Régionale	1,2	1,7	1,4	2,1	4,3	3,3	3,7

À l'échelle régionale, les proportions de frontières fortes et très fortes pour chaque variable sont inférieures à celles de l'échelle locale, la variable *CD* présentant le pourcentage le plus élevé (Tableau 15). Pour les combinaisons de deux variables, les proportions sont plus proches de celles de l'échelle locale et la combinaison *CH-CA* conserve le pourcentage le plus élevé. La proportion de frontières fortes pour la combinaison des trois variables, même si elle est très basse, est deux fois supérieure à celle de l'échelle locale.

### 1.2.2. Répartition spatiale des frontières fortes locales et régionales

Sur la base d'une analyse visuelle, les cartes des frontières fortes et très fortes aux échelles locales et régionales pour les variables et pour la combinaison des trois variables, montrent des frontières dispersées et les situations où les frontières sont enchaînées sont rares (Figure 8). Cette répartition spatiale plus ou moins homogène des frontières fortes indique également que si bien l'effet de bordure n'est pas considéré dans l'algorithme, il ne semble pas affecter les résultats. Lorsque les frontières fortes aux deux échelles, pour l'une ou l'autre des variables (*CD* ou *CH* ou *CA*), sont représentées ensemble, des chaînes de frontières locales ou régionales et très souvent des deux, commencent à se former (Figure 8e).

La recherche d'une relation entre les frontières fortes et très fortes à chaque échelle, et des combinaisons particulières de variables, comme la pente ou le type de perturbation des peuplements adjacents, suggère des singularités seulement pour les perturbations sévères. En effet, les frontières localement fortes pour *CH* et *CA* présentent une fréquence plus élevée que celle de l'ensemble, dans des situations où l'un des peuplements adjacents à la frontière a été atteint par une perturbation, soit une coupe totale, une épidémie sévère ou un chablis total (Tableau 16). Les frontières fortes à l'échelle régionale pour les mêmes variables présentent, au contraire, des fréquences inférieures à celles de l'ensemble pour ce type de situation. Cependant, ces frontières présentent une fréquence significativement supérieure ( $\alpha=0,05$ ) dans les situations où, soit aucun des peuplements soit les deux peuplements ont subi des perturbations (Tableau 16). Les frontières localement fortes pour *CD* diffèrent, pour leur part, de la tendance des autres variables. Les situations où un peuplement ayant subi un chablis total ou une épidémie sévère est adjacent à un peuplement non perturbé sont moins fréquentes parmi ces frontières que dans l'ensemble des frontières.

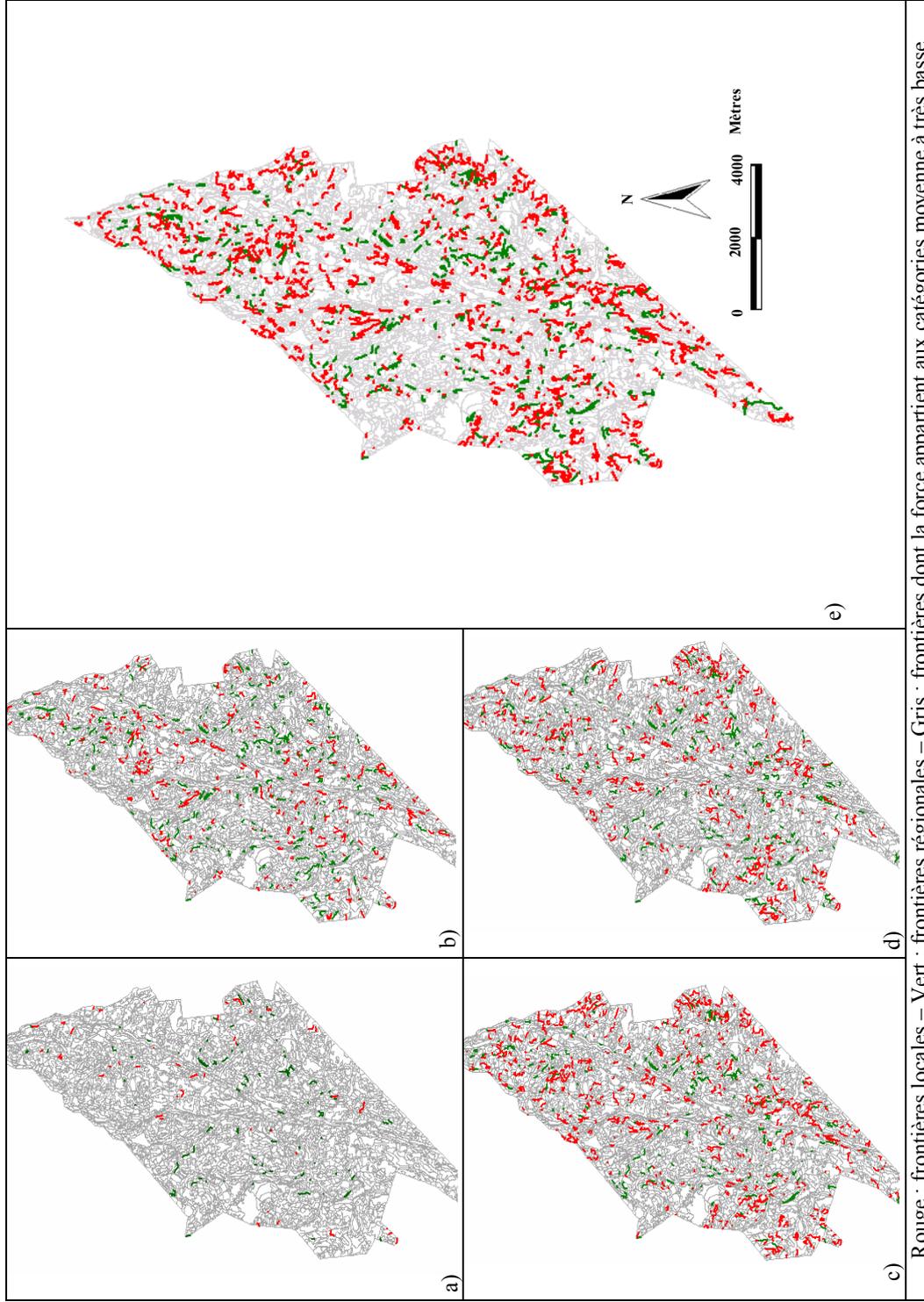


Figure 8. Frontières fortes et très fortes locales et régionales. a) combinaison des trois variables à la fois; b) *CD*; c) *CH*; d) *CA*, e) *CD* ou *CH* ou *CA* ( $n=8018$  frontières).

Tableau 16. Pourcentage de frontières fortes locales et régionales pour les variables *CD*, *CH* et *CA*, qui coïncident avec des combinaisons de perturbations dans les peuplements adjacents. cht : chablis total; ct : coupe totale, es : épidémie sévère, sp : sans perturbation.

Combinaison de types de perturbations	Pourcentage de frontières fortes à l'échelle locale			Pourcentage de frontières fortes à l'échelle régionale			Frontières totales (%)
	<i>CD</i>	<i>CH</i>	<i>CA</i>	<i>CD</i>	<i>CH</i>	<i>CA</i>	
cht – sp	4,9	16,0	15,2	4,7	1,1	2,0	5,9
ct – sp	32,2	70,6	64,3	26,0	6,7	18,9	24,8
es – sp	1,6	4,8	4,9	1,5	0,0	0,3	1,6
ct – ct	3,5	0,1	0,0	11,4	14,2	11,1	8,6
cht – cht	0,0	0,0	0,0	0,88	1,5	1,3	0,6
cht – ct	0,0	0,0	0,0	0,3	0,0	0,0	0,2
ct – es	0,0	0,0	0,0	0,6	0,4	0,0	0,2
sp – sp	61,2	8,57	15,5	54,6	76,0	66,3	58,1
$\chi^2$	20,7	1298,2	704,9	4,84	69,9	17,4	
Probabilité	0,000	0,000	0,000	0,424	0,000	0,001	

Quant au lien entre les deux méthodes de quantification du contraste, la majorité des frontières localement fortes coïncide spatialement avec les frontières de fort contraste quantifié par *S* et ce, pour les trois variables. Certaines coïncident avec les frontières de contraste moyen et une minorité avec celles de faible contraste. Ces différences peuvent être expliquées par les méthodes employées pour quantifier le contraste et la force.

Cette coïncidence spatiale des frontières, localement fortes et de haut contraste, est accompagnée d'une concordance dans les proportions de frontières appartenant à chaque niveau, surtout pour les variables *CD* et *CH*. Environ 4 % des frontières sont de fort contraste (soit mesuré par la force ou par *S*) pour *CD* et environ 10 % le sont pour *CH*. Les proportions sont différentes pour *CA*, dont 20 % des frontières sont de haut contraste et seulement 7 % des frontières sont localement fortes ou très fortes.

Malgré la faible proportion de frontières fortes à l'échelle régionale, la quantification du contraste à cette échelle est intéressante, car elle permet d'élargir la résolution spatiale de l'analyse en considérant non seulement les deux peuplements adjacents à la frontière, mais aussi des groupes de peuplements adjacents à la frontière.

## 2. Quantification du contexte

La quantification du contexte est faite sur deux plans différents. Le contexte des frontières, dont les résultats sont présentés en premier, fait référence à la configuration spatiale locale des valeurs de  $S$ , qui quantifie le contraste. Le contexte des peuplements, présenté par la suite, concerne l'hétérogénéité des variables du voisinage de chaque peuplement.

### 2.1. Le contexte des frontières

Dans le paragraphe antérieur, le contexte d'une frontière est défini en termes de configuration spatiale locale des valeurs de  $S$  pour chaque variable. La mesure d'autocorrélation spatiale locale de  $S$ , l' $I$  de Moran local, calculée pour chacune de frontières, permet de déterminer : (1) la localisation et l'étendue des groupes de frontières qui présentent des valeurs « anormales » et similaires de contraste par rapport à la moyenne de  $S$  pour l'aire d'étude (valeurs d' $I$  de Moran positives) et (2) les frontières dont les valeurs d'écart à la moyenne de  $S$  diffèrent des frontières voisines (valeurs négatives d' $I$  de Moran) (Fortin and Dale 2005, Boots 2002). Dans cette section, les résultats sont présentés en interprétant les valeurs de  $S$  en niveaux de contraste, selon les catégories décrites au Tableau 11 de la section méthodologique.

Plus précisément, les valeurs significatives d'autocorrélation locale du contraste, positives et négatives, montrent des répartitions spatiales différentes en fonction des variables (Figure 9). Par exemple, la *Classe de hauteur* affiche une concentration de frontières dont les valeurs de contraste sont très différents de la moyenne dans une bande qui traverse la Forêt Montmorency, tandis que le *Groupement d'essences* l'affiche dans le nord et le sud-ouest (Figure 9a et c).

Ces secteurs où les valeurs significatives d'autocorrélation spatiale locale se concentrent se superposent partiellement pour plusieurs variables. Le secteur nord présente des zones homogènes de frontières « fortes » (de haut et très haut contraste) pour le *Groupement d'essences* et la *Classe de densité*, ainsi que des frontières « faibles » (bas contraste) isolées

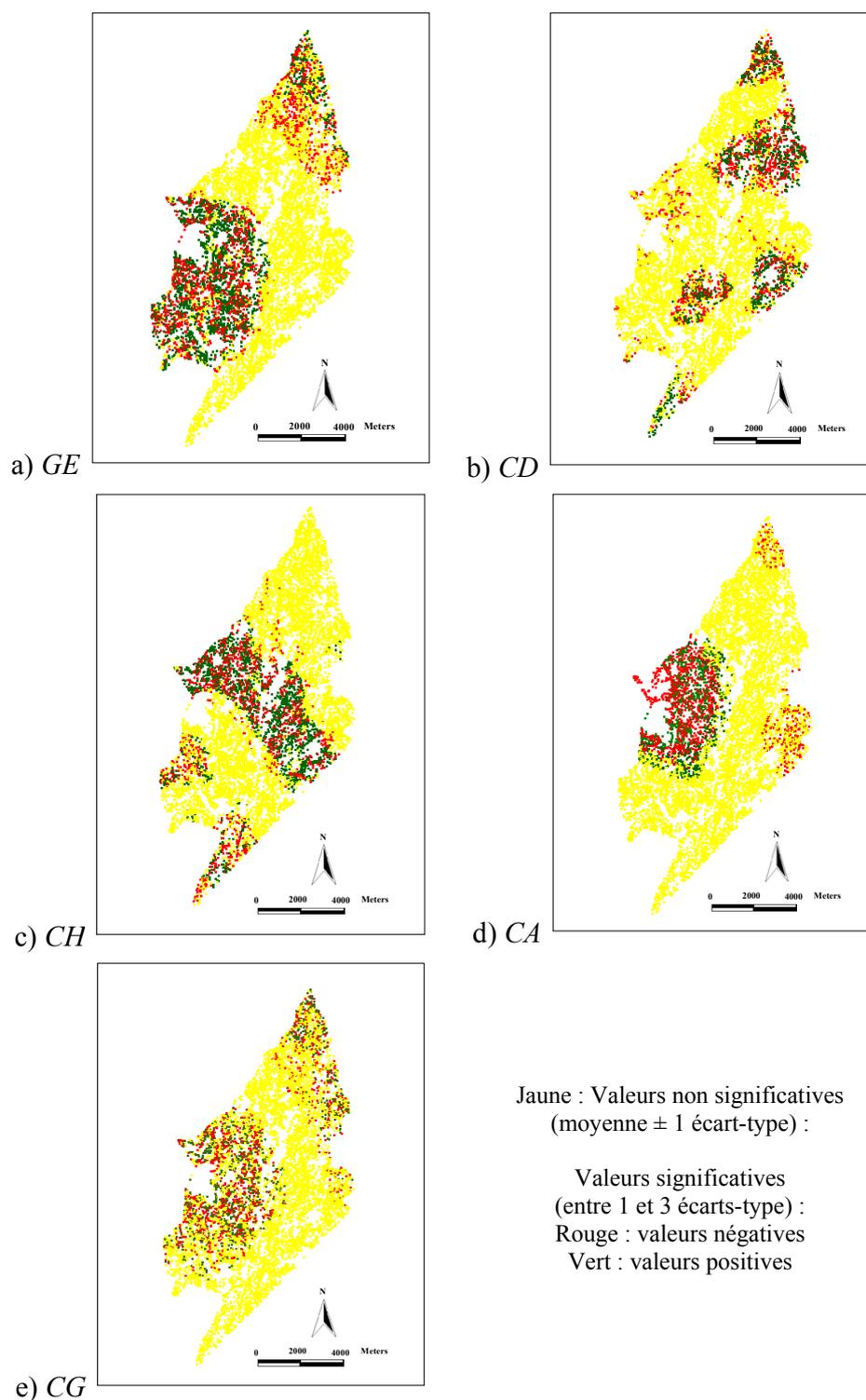


Figure 9. Valeurs d'autocorrélation spatiale locale du contraste de: a) *Groupement d'essences (GE)*, b) *Classe de densité (CD)*, c) *Classe de hauteur (CH)*, d) *Classe d'âge (CA)* et e) *Contraste Global (CG)*.

pour *Classe d'âge*. Le secteur ouest, qui comporte des zones homogènes de frontières faibles pour *Groupement d'essences* et *Classe d'âge*, coïncide partiellement avec la bande centrale qui comporte les mêmes caractéristiques de *Classe de hauteur*. Dans ce secteur, les frontières « fortes », qui se distinguent des frontières « faibles » de leur entourage, pourraient être plus facilement identifiables lors de la photo-interprétation. Le secteur sud, par contre, présente des zones homogènes de frontières « fortes » pour *Classe de hauteur* et « faibles » pour *Classe de Densité*. Dans ce secteur, comme celui du nord, les frontières « faibles » seraient difficilement identifiables, étant donné que leur contexte est constitué de frontières « fortes ».

Les informations sur le contraste et sur le contexte des frontières sont complémentaires. En fait, le contexte ne peut être interprété que de façon conjointe avec le contraste (Figure 10). Ces deux aspects permettent ainsi de déceler les zones particulières à la Forêt Montmorency où l'identification des frontières serait facilitée lors de la photo-interprétation.

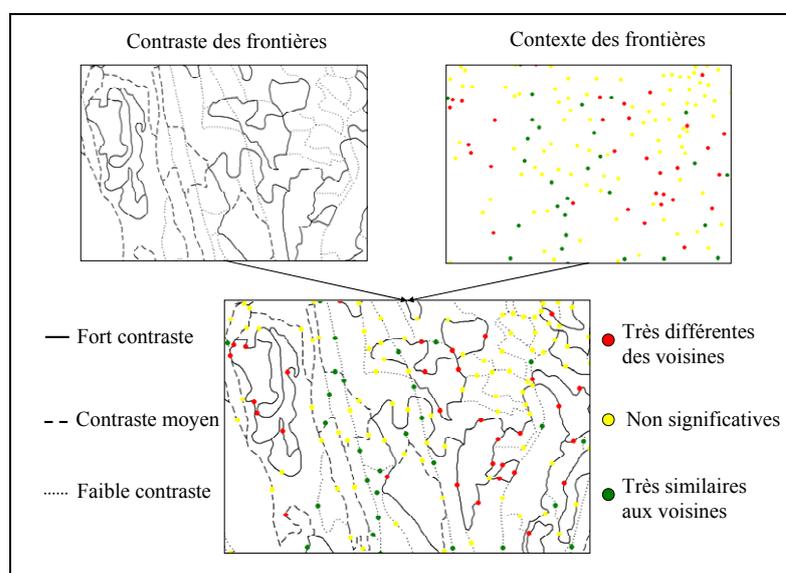


Figure 10. Caractère complémentaire du contraste et du contexte des frontières.

## 2.2. Le contexte des peuplements

Le contexte de chacun des peuplements est défini comme l'ensemble des peuplements adjacents. L'indice d'hétérogénéité du contexte de chaque peuplement ( $H$ ) indique, pour la variable d'intérêt, la proportion de voisins dont la valeur coïncide avec celle du peuplement en question. Pour faciliter l'interprétation des résultats, l'indice  $H$  est classifié en degrés d'hétérogénéité, selon les seuils présentés au Tableau 12.

La proportion des degrés d'hétérogénéité est assez similaire pour toutes les variables, qu'elle soit mesurée au moyen de la proportion de peuplements ou de la proportion de surface (Figure 11). Entre 60 et 70 % de la surface et des peuplements présentent une hétérogénéité haute et très haute du contexte, tandis qu'elle est basse et très basse pour 9 à 16 % de la surface et pour 11 à 18 % des peuplements (Figure 11). Cette prédominance de niveaux élevés d'hétérogénéité peut indiquer une forte fragmentation de la forêt.

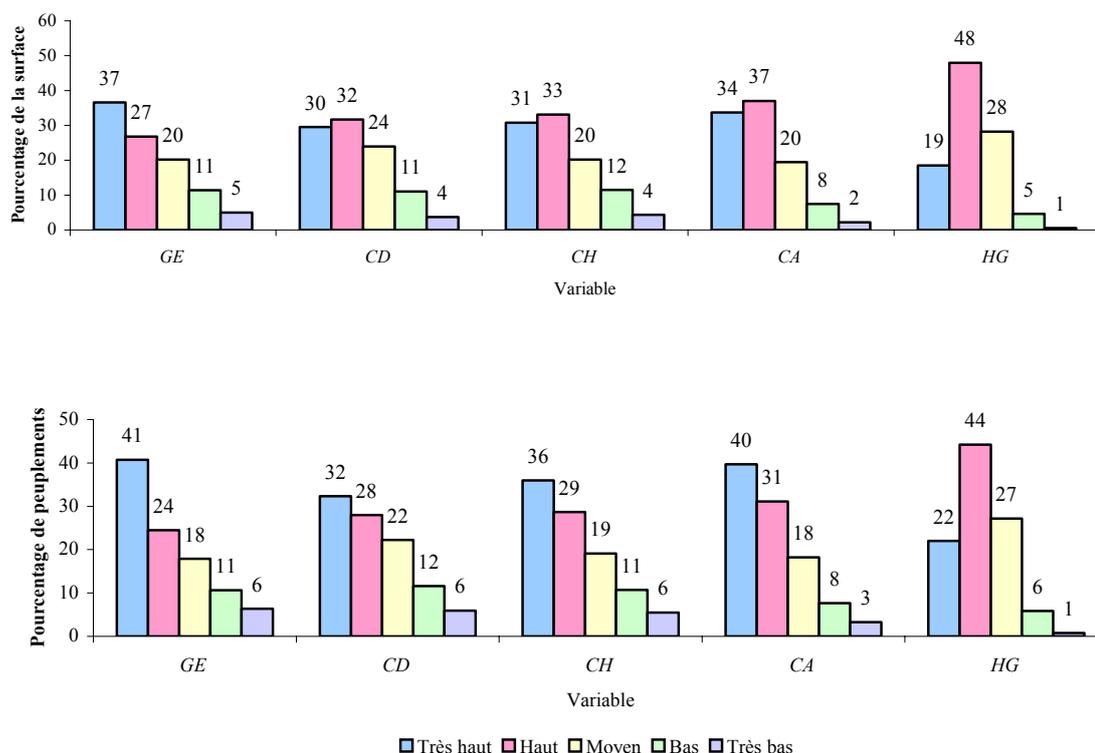


Figure 11. Proportion de surfaces et de peuplements correspondants à chaque niveau d'hétérogénéité ( $H$ ), selon les variables  $GE$ ,  $CD$ ,  $CH$  et  $CA$ .

La répartition spatiale de l'hétérogénéité est particulière à chaque variable (Figure 12). Des zones d'hétérogénéité très haute et très basse sont identifiables pour le *Groupement d'essences (GE)*, tandis que pour les autres variables ces zones sont plus fragmentées (Figure 12a). La plupart des zones d'hétérogénéité moyenne à très basse pour la *Classe de densité (CD)*, la *Classe de hauteur (CH)* et la *Classe d'âge (CA)* sont localisées à proximité du réseau hydrographique (Figure 12b, c et d). De plus, une coïncidence spatiale avec l'information sur le contexte des frontières est observée pour *GE*, *CD* et *CH*. En effet, pour *GE*, le secteur nord de haute hétérogénéité correspond à une zone de frontières de fort contraste, tandis que le secteur sud, de basse hétérogénéité, correspond aux frontières de faible contraste. Les deux types d'information sont donc complémentaires, car l'hétérogénéité n'implique pas un niveau de contraste, mais simplement que le voisinage présente des catégories différentes pour une variable donnée.

Du point de vue de l'*Hétérogénéité globale (HG)*, bien qu'une haute hétérogénéité soit prédominante (Figure 11), elle correspond à des zones fragmentées, tandis que les zones de basse hétérogénéité sont plus étendues (Figure 12e). Comme pour les variables isolées, les frontières de contraste plus élevé que la moyenne se localisent, en général, dans des zones de haute et de très haute hétérogénéité globale (Figure 9).

L'analyse des caractéristiques forestières des peuplements qui présentent un contexte avec un haut ou très haut degré d'hétérogénéité pour toutes les variables indique que ces peuplements présentent plus fréquemment des perturbations sévères que ceux de basse ou de très basse hétérogénéité. Par exemple, la proportion de peuplements qui ont subi un chablis total diminue, pour toutes les variables, à mesure que l'hétérogénéité du contexte diminue. La même tendance s'observe pour la proportion de peuplements ayant subi des coupes totales, par rapport à *GE* et *CH*, tandis que pour *HG* les peuplements avec une très basse hétérogénéité présentent la proportion maximale des coupes totales.

La considération conjointe du contraste et du contexte des frontières avec le contexte des peuplements permet d'identifier les zones particulières où la haute hétérogénéité des peuplements concorde avec des groupes de frontières de fort contraste (Figure 13). De plus, les coïncidences spatiales observées confirment la complémentarité de ces trois aspects.

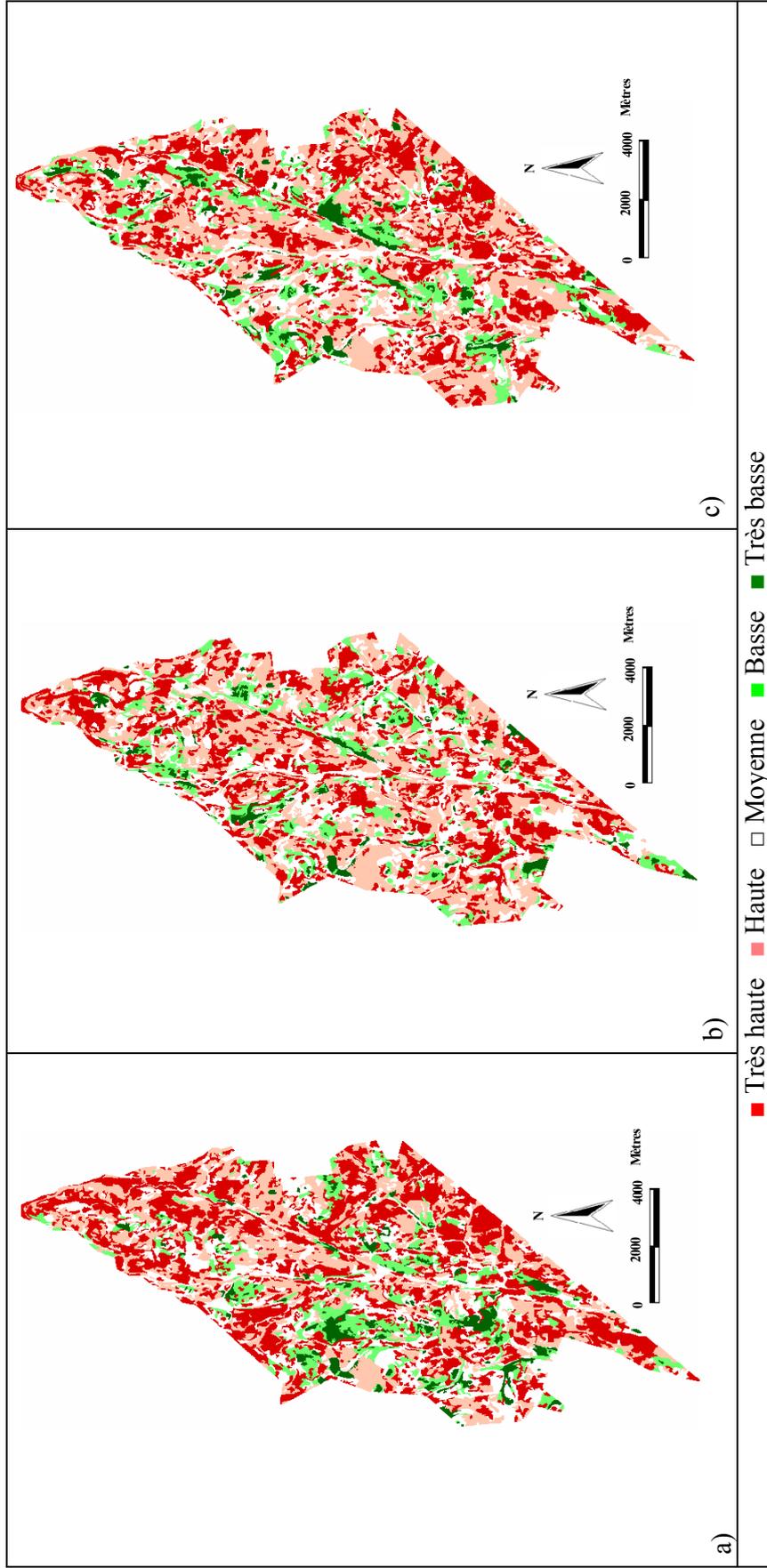


Figure 12. Hétérogénéité (quantifiée par  $H$ ) du contexte des peuplements. a)  $GE$ ; b)  $CD$ ; c)  $CH$ .

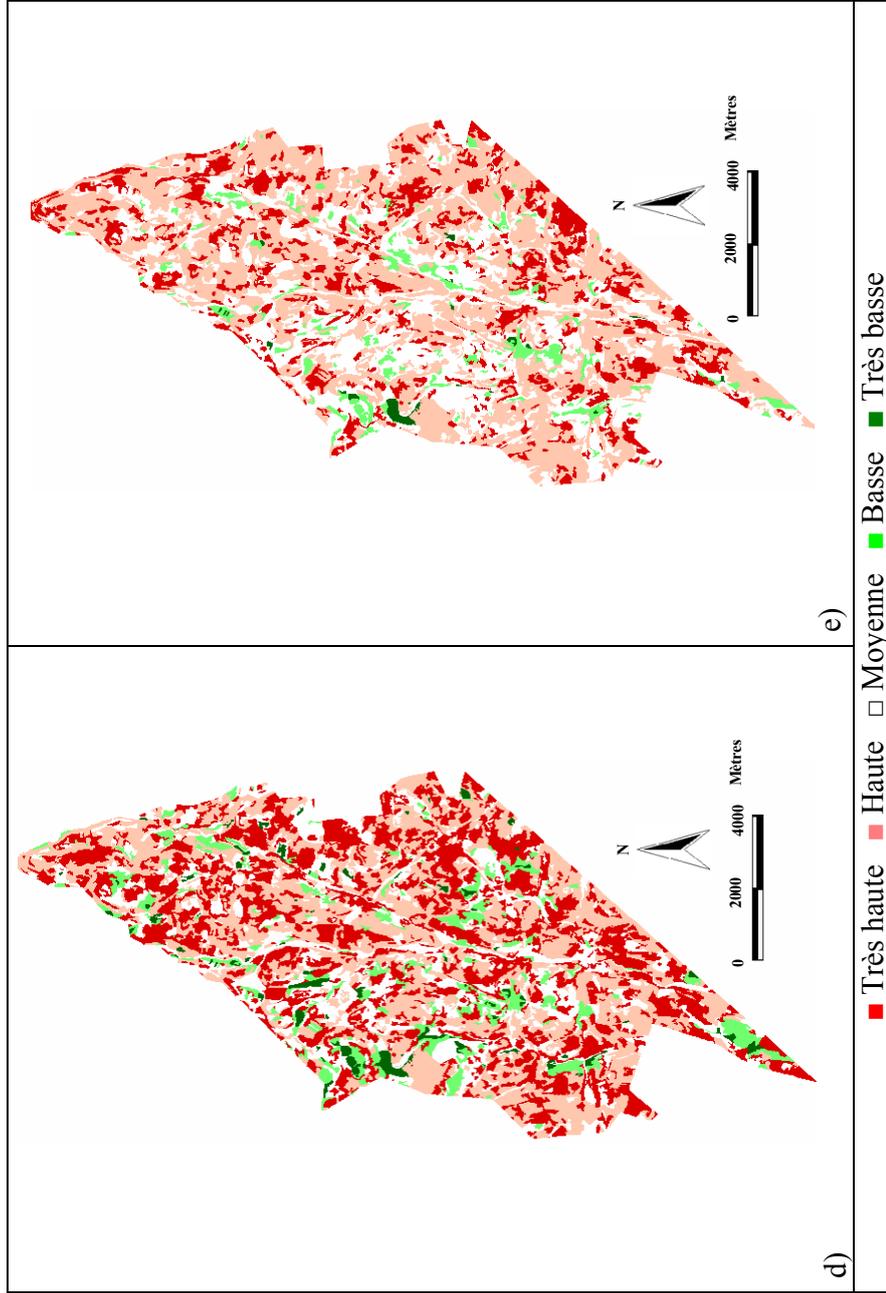


Figure 12. Suite. Hétérogénéité (quantifiée par  $H$ ) du contexte des peuplements. d)  $CA$ ; e) Hétérogénéité globale.

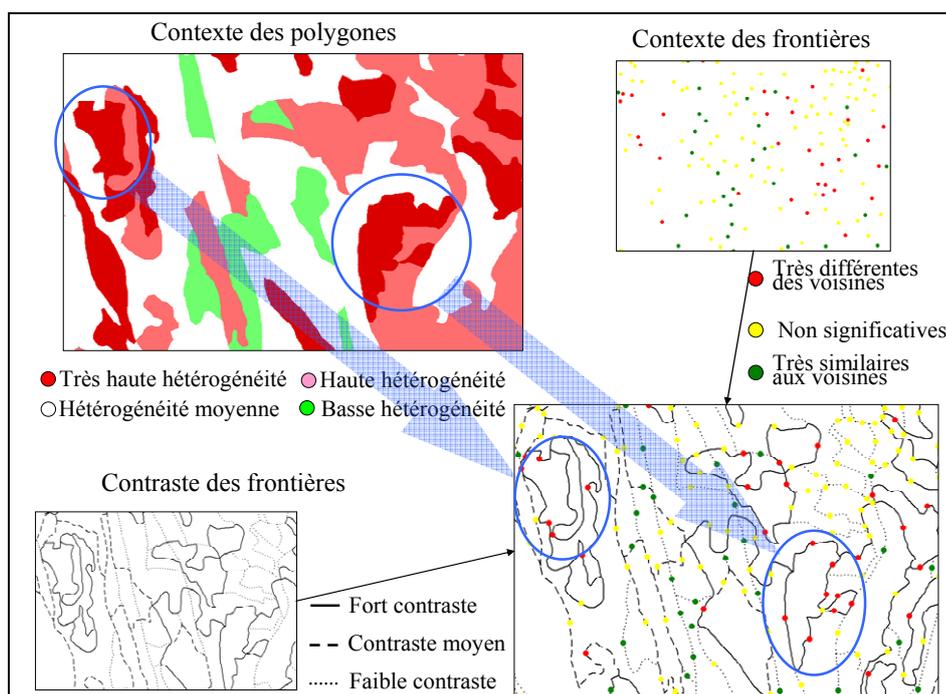


Figure 13. Relation entre le contraste et le contexte des frontières, et le contexte des polygones.

### 3. Évaluation du lien entre le contraste et le contexte, et la probabilité d'existence d'une frontière

Cette section vise à évaluer le lien entre le contraste, quantifié par  $S$ , le contexte et la probabilité d'existence d'une frontière, quantifiée par la *Consistance*. Premièrement, l'analyse des distributions de fréquences de  $S$  en fonction des trois degrés de *Consistance*, est présentée. Par la suite, le modèle qui évalue les variables les plus puissantes pour discriminer les degrés de *Consistance* est discuté.

#### 3.1. Distributions de fréquence du contraste en fonction de la *Consistance*

Selon l'hypothèse de départ, les frontières qui ont un contraste écologique fort (haut et très haut) devraient être celles qui présentent une haute *Consistance* et, à l'inverse, les frontières de faible contraste ou sans contraste devraient être celles qui présentent une basse *Consistance*.

La proportion de frontières dites « sans contraste » diminue à mesure que la *Consistance* augmente et ce, pour toutes les variables (Figure 14). La proportion de frontières dites « fortes » (contraste très haut et haut) augmente avec le degré de *Consistance* pour *GE* et *CA* et légèrement pour *CD*. Bien que ces tendances ne s'appliquent pas à la variable *CH* et que les différences dans les proportions ne soient pas très marquées, elles soutiennent quand même l'hypothèse de départ que l'existence d'une frontière sera plus certaine dans le cas où le contraste entre les peuplements adjacents à la frontière est élevé. De plus, les distributions de fréquence de *S* selon le degré de *Consistance* sont significativement différentes pour toutes les variables (test de Kolmogorov-Smirnov pour deux échantillons,  $\alpha = 0,05$ ,  $n = 8018$  frontières, Tableau 17).

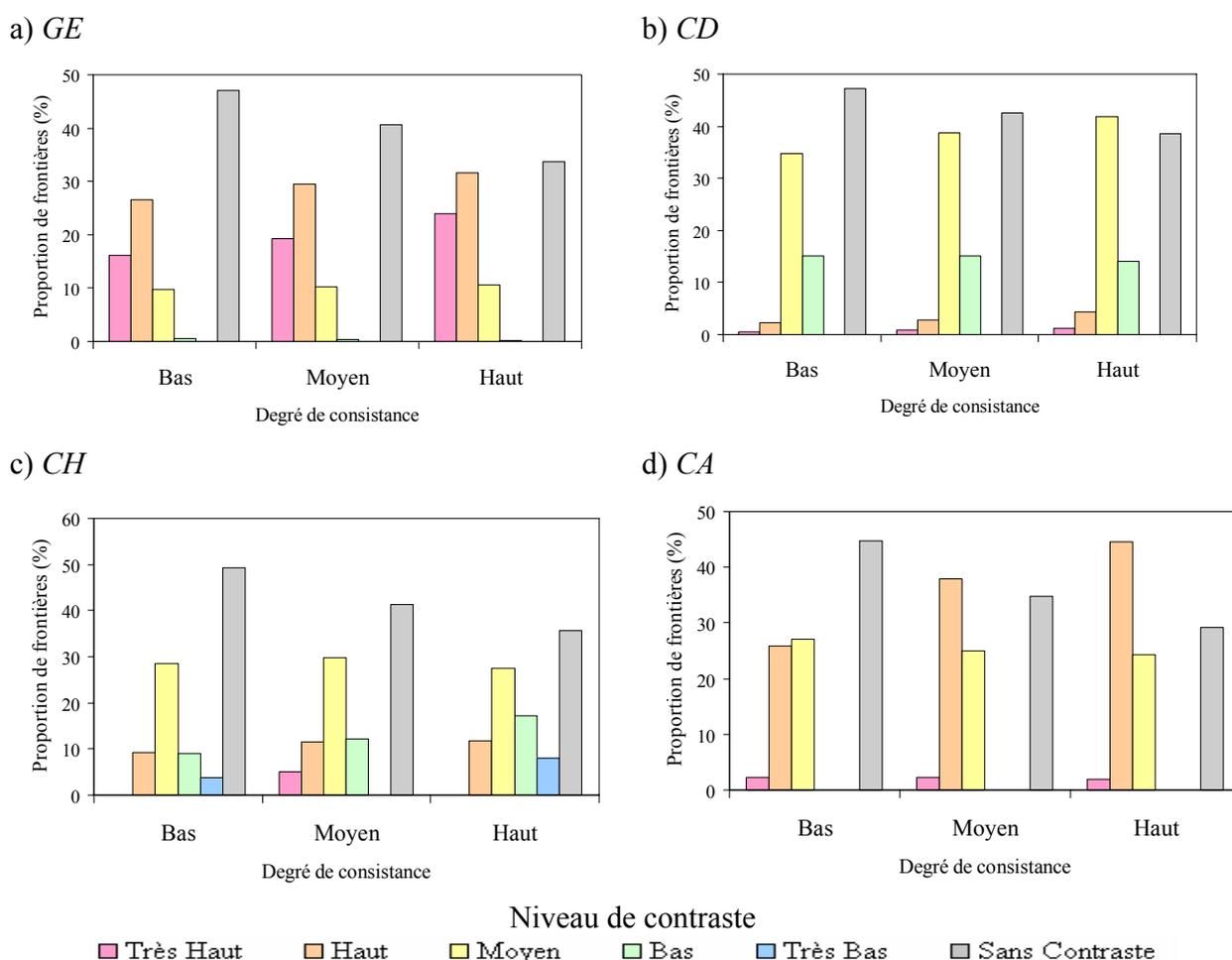


Figure 14. Distributions des fréquences des niveaux de contraste des frontières selon le degré de *Consistance*. a) *GE*; b) *CD*; c) *CH*; d) *CA* ( $n=8018$  frontières).

Tableau 17. Distributions des fréquences de  $S_{GE}$ ,  $S_{CD}$ ,  $S_{CH}$  et  $S_{CA}$  en fonction des degrés de *Consistance* : Test de Kolmogorov-Smirnov (KS) pour deux échantillons et sa probabilité (Prob. : Probabilité;  $n=8018$  frontières).

	Degrés de <i>Consistance</i> comparés											
	Bas – Moyen				Bas – Haut				Moyen – Haut			
	$S_{GE}$	$S_{CD}$	$S_{CH}$	$S_{CA}$	$S_{GE}$	$S_{CD}$	$S_{CH}$	$S_{CA}$	$S_{GE}$	$S_{CD}$	$S_{CH}$	$S_{CA}$
KS	0,06	0,04	0,08	0,12	0,13	0,09	0,13	0,18	0,07	0,05	0,05	0,06
Prob.	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

### 3.2. Relation entre le contraste et le contexte, et la consistance d'une frontière

Cette section évalue l'aptitude de l'ensemble des variables de contraste, de contexte ainsi que la *Longueur* des frontières pour discriminer les trois degrés de *Consistance* au moyen de l'analyse discriminante pas à pas. Le modèle final retient neuf variables, cependant son pouvoir discriminant est très bas (Wilks Lambda=0,948,  $F$  exact=13,531,  $p=0,000$ ) (Tableau 18). La fonction de classification des frontières en degrés de consistance utilise les variables discriminantes du modèle final.

Tableau 18. Coefficients linéaires de Fisher des variables discriminantes dans la fonction de classification des frontières en degrés de *Consistance* et niveau de signification du modèle retenu ( $n=4515$  frontières). Prob. : Probabilité

	Degré de Consistance				Prob. <sup>1</sup>	Niveau de signification			
	Bas	Moyen	Haut	Wilks' Lambda		$\chi^2$	dl	Prob.	
Constante	-13,63	-13,19	-12,77						
$IMLS_{CA}$	-0,19	-0,17	-0,21	*	0,948	240,55	18	0,000	
$S_{CA} * IMLS_{CA}$	0,05	0,06	0,10	*					
$S_{GE} * Longueur$	-0,02	-0,02	-0,02	0,062					
$S_{CA} * Longueur$	-0,05	-0,05	-0,05	**					
$S_{CA}$	6,23	5,86	5,38	**				Taux d'erreur	
$S_{CD}$	10,39	10,14	9,75	**				de la	
$S_{GE}$	4,23	3,94	3,81	*				validation	
$S_{CH}$	12,54	12,66	13,17	**				Taux d'erreur de	
$Longueur$	0,05	0,05	0,05	**				la classification	
								58,2 %	
								58,7 %	

<sup>1</sup>Indique un niveau de  $\alpha = *--0,05$  et  $**--0,01$ .

Avec une proportion similaire de frontières pour chaque degré de consistance, la classification présente un taux d'erreur de 58,2 %, tandis que celui de la validation croisée

est de 58,7 % (Tableau 18). Ces résultats indiquent que le pouvoir des fonctions discriminantes est relativement faible pour prédire le degré de *Consistance* d'une frontière.

Cependant, les résultats de l'analyse montrent que la *Consistance* est reliée au contraste des quatre variables, à la longueur de la frontière, au contexte de *Classe d'âge* ( $IMLS_{CA}$ ), à l'interaction contraste-contexte de *Classe d'âge* ainsi qu'aux interactions contraste-Longueur de *Groupement d'essences* et de *Classe d'âge*. Les quatre variables de contraste ( $S$ ) sont les plus importantes pour les fonctions de classification, la *Classe de hauteur* ( $S_{CH}$ ) et la *Classe de densité* ( $S_{CD}$ ) étant celles qui présentent les valeurs supérieures de signification (Tableau 18). Hormis la *Classe d'âge*, le contexte des autres variables semble être secondaire par rapport au contraste. Les résultats obtenus coïncident avec ceux reportés dans la littérature. En effet, il a été déterminé que l'incertitude sur l'existence d'une frontière est reliée principalement à sa longueur ainsi qu'aux différences de composition d'essences et de hauteur des peuplements adjacents à la frontière (De Groeve 1999, De Groeve et Lowell 2001).

Une des raisons qui explique la pauvre performance du modèle sont sans doute la proximité entre les valeurs moyennes des variables et leurs écarts-type dans les trois degrés de *Consistance* (Tableau 19). Cette proximité, même un certain chevauchement, des trois degrés dans l'espace des variables, rend la discrimination difficile et peut expliquer la faible valeur de Wilks Lambda et les taux élevés d'erreur des classifications.

De plus, deux autres aspects doivent être considérés. D'abord, la carte forestière est le résultat d'un processus cognitif complexe. Les peuplements forestiers sont délimités lors de la photo-interprétation selon une combinaison de règles et de procédures par lesquels les frontières et les attributs, isolés ou en combinaisons différentes, sont importants (Lowell *et al.* 1996, Thierry 1998, Lowell 2001). L'objet, dans le cas présent un peuplement forestier, est identifié et délimité à partir de la minimisation de la variation des indices visuels à l'intérieur de l'objet et de la maximisation de la variation entre objets et ainsi, même si ses frontières sont floues, l'objet peut être identifié (Lowell 1999). Le deuxième aspect concerne la quantification adéquate de la probabilité d'existence des frontières. La

Tableau 19. Valeur moyenne (M) et écart-type (ÉT) des variables en fonction des degrés de *Consistance*. Prob. : Probabilité.

Variable	Degré de Consistance						Test d'égalité des moyennes	
	Bas		Moyen		Haut		F	Prob.
	M	ÉT	M	ÉT	M	ÉT		
<i>IMLS<sub>CA</sub></i>	0,6	4,9	0,7	4,6	0,3	4,8	2,72	0,066
<i>S<sub>CA</sub>*IMLS<sub>CA</sub></i>	0,9	4,0	0,9	3,6	0,7	3,5	1,04	0,351
<i>S<sub>GE</sub>*Longueur</i>	85,3	98,6	81,3	107,8	56,9	75,0	39,71	0,000
<i>S<sub>CA</sub>*Longueur</i>	101,7	108,6	90,3	94,0	69,9	76,6	44,38	0,000
<i>S<sub>CA</sub></i>	0,7	0,3	0,6	0,3	0,6	0,3	59,89	0,000
<i>S<sub>CD</sub></i>	0,8	0,2	0,7	0,2	0,7	0,2	19,40	0,000
<i>S<sub>GE</sub></i>	0,6	0,4	0,5	0,4	0,5	0,4	36,90	0,000
<i>S<sub>CH</sub></i>	0,8	0,2	0,7	0,2	0,7	0,2	16,98	0,000
<i>Longueur</i>	141,8	121,6	147,3	126,1	126,5	115,3	11,97	0,000

consistance mesurée à partir de trois cartes multitemporelles est une mesure limitée de cette probabilité. Le fait que les cartes soient élaborées de manière indépendante avec une dizaine d'années d'écart ajoute de l'incertitude à un produit dont l'incertitude spatiale est déjà élevée (De Groeve 1999). Idéalement, la consistance devrait être mesurée à partir de nombreuses interprétations des mêmes photographies aériennes (Aubert 1995) pour qu'elle s'approche de l'expression d'une probabilité.

En lien avec le contraste et la probabilité d'existence de la frontière, mais à l'échelle régionale, la proportion de frontières fortes à cette échelle, pour les trois variables qui présentent une consistance moyenne à haute, est légèrement supérieure à la proportion de l'ensemble des frontières (57 et 51 % des frontières respectivement pour le degré moyen de consistance, 31 et 30 % respectivement pour le degré de haute consistance). Bien que les différences ne soient pas très marquées, elles laissent tout de même croire que ces frontières, régionalement fortes pour les trois variables, sont plus facilement identifiables lors de la photo-interprétation.

## Conclusions

Les méthodes utilisées dans ce chapitre pour quantifier le contraste écologique entre les peuplements ainsi que l'hétérogénéité de leur voisinage influencent fortement les résultats

obtenus, car elles reposent sur des nombreuses décisions arbitraires. Elles permettent, cependant, d'établir des différences entre les frontières que dans la carte forestières sont représentées de la même manière ainsi que de générer de nouvelle information quant aux aspects spatiaux des attributs des peuplements. Avec les limitations imposées par ces choix, les informations obtenues s'avèrent complémentaires et peuvent servir de guide pour la prise de décisions quant aux interventions comme les échantillonnages de terrain.

La quantification du contexte des peuplements indique les secteurs où la diversité des variables est élevée, alors que la quantification du contraste local des frontières indique la magnitude des différences entre deux peuplements voisins. La quantification du contraste à l'échelle du voisinage, quant à elle, met en évidence les frontières fortes entre deux groupes de peuplements. Finalement, la quantification du contexte des frontières informe sur les groupes de frontières qui présentent des valeurs de contraste « anormales », soit très haut ou très bas par rapport à la moyenne de l'ensemble des frontières.

L'indice de similarité écologique  $S$  permet de quantifier la ressemblance entre catégories complexes d'une variable nominale et d'estimer ainsi le contraste d'une façon plus raffinée, au-delà de la simple coïncidence de catégories. Toutefois, les résultats obtenus dépendent fortement des nombreuses décisions prises pour quantifier la similitude partielle ainsi que de la classification employée pour déterminer les niveaux de contraste.

Dans cette étude, 51 % des frontières présentent un fort contraste pour l'une ou l'autre des variables considérées et environ 11 % des frontières sont de fort contraste pour l'ensemble des variables. Le *Groupement d'essences* et la *Classe d'âge* présentent la plus grande proportion de frontières dites « fortes » (50 % et 40 % des frontières respectivement) et la *Classe de hauteur* a, pour sa part, la plus grande proportion de frontières dites « faibles » (19 %).

Néanmoins, l'importance des variables dépend des conditions locales. Lors de la photo-interprétation, chaque frontière est identifiée à partir d'une combinaison spécifique d'indices visuels, de la frontière même ou des peuplements adjacents, reliés à une ou à plusieurs variables. Donc, même les variables comme la *Classe de densité* ou la *Classe de*

*hauteur*, dont le contraste prédominant est moyen ou faible, sont nécessaires pour l'identification de certaines frontières. D'ailleurs, le même critère s'applique pour les frontières fortes à l'échelle régionale. Même si les proportions de ces frontières sont très faibles, elles complètent celles qui sont localement fortes, surtout lorsque la force concerne les trois variables.

La quantification du contexte de frontières est, pour sa part, complémentaire à celle du contraste. En effet, l'analyse intégrée des deux aspects permet de détecter les secteurs de la forêt où l'identification des frontières pourrait être facilitée, comme ceux des frontières fortes entourées de frontières faibles (c'est-à-dire un changement abrupte des attributs dans un contexte d'une certaine homogénéité pour ces attributs). L'identification des zones homogènes de frontières faibles pourrait également orienter le regroupement des peuplements lors de la planification des interventions forestières.

Quant à l'hétérogénéité du contexte des peuplements, environ 65 % des peuplements présentent un contexte d'hétérogénéité élevée pour toutes les variables. Une coïncidence est observée entre la répartition spatiale des degrés d'hétérogénéité des peuplements et le contraste des frontières. Ainsi, les frontières de fort contraste coïncident avec certaines zones de haute hétérogénéité. En fait, les deux types d'information sont complémentaires, l'hétérogénéité du contexte des peuplements informe sur la diversité des peuplements voisins par rapport aux variables d'intérêt, alors que le contraste des frontières indique l'intensité de ces différences.

En ce qui concerne la relation entre le contraste, le contexte et la probabilité d'existence des frontières quantifiée par la consistance, les résultats de l'analyse discriminante montrent une relation significative entre cette dernière et le contraste de toutes les variables, la longueur de la frontière, et le contexte de *Classe d'âge*. Toutefois, l'exploration de ce lien reste à approfondir, surtout par rapport aux moyens d'améliorer la quantification de la probabilité d'existence. Les résultats indiquent qu'il existe une relation significative entre le contraste et la consistance, c'est-à-dire que les frontières fortes montrent une tendance à avoir une consistance élevée, donc que la fiabilité sur l'existence d'une frontière varie en

fonction du contraste. Lorsque la probabilité d'existence d'une frontière sera mieux représentée, ce lien deviendra possiblement plus évident.

## **Résumé du chapitre**

Ce chapitre a traité des frontières entre les peuplements de la carte forestière. Différentes méthodes ont été utilisées pour quantifier des aspects reliés à l'incertitude sur l'existence des frontières, tels que le contraste écologique entre les peuplements adjacents à chaque frontière (à diverses échelles de voisinage), la configuration spatiale locale de ce contraste, ainsi que l'hétérogénéité du contexte des peuplements. Malgré les limitations entraînées par des nombreux choix plus ou moins arbitraires impliqués dans les méthodes, elles s'avèrent utiles et complémentaires pour identifier dans la zone d'étude des secteurs aux caractéristiques particulières.

Cependant, au-delà des frontières cartographiques existantes, il est possible de définir d'autres types de frontières à partir des données ponctuelles, telles les placettes-échantillons. Ces frontières permettent de mieux comprendre le patron spatial des variables d'intérêt comme le volume de matière ligneuse et peut aider à améliorer les interventions forestières. Tel est le sujet d'exploration du chapitre suivant.

## **Chapitre IV - Les frontières issues des données de terrain**

### **Introduction**

Les placettes-échantillons (p-e) relevées dans le cadre des inventaires forestiers au Québec (Bérard et Côté 1996) fournissent de l'information détaillée sur les caractéristiques structurales de la végétation ligneuse ainsi que sur un ensemble de variables topographiques et édaphiques. Cette source de données, d'un coût très élevé, est habituellement utilisée pour estimer le volume de matière ligneuse des strates de la carte forestière identifiées par photo-interprétation, dont les p-e constituent un échantillon représentatif (Bérard et Côté 1996). La répartition spatiale des estimations du volume ligneux sur le territoire est ainsi déterminée de manière indirecte, par la répartition spatiale des strates forestières. À la Forêt Montmorency il existe une densité exceptionnellement élevée de p-e. Serait-il possible, dans ce cas, d'établir la répartition spatiale des estimations du volume ligneux directement à partir des p-e ? Est-ce que cette représentation spatiale produite à partir des données de terrain pourrait-elle être fusionnée à la carte forestière pour obtenir une carte plus certaine par rapport aux surfaces des parcelles ? Ce chapitre vise à répondre à la première question afin de pouvoir explorer des réponses à la deuxième question au Chapitre VII.

Dans le contexte de reconnaissance des patrons de répartition spatiale des données, il est important de délimiter les zones du territoire forestier ayant des volumes ligneux similaires. L'analyse des frontières est donc employée pour analyser les données des p-e afin de délimiter de telles zones homogènes en termes du volume. Plus spécifiquement, l'analyse des frontières vise à identifier les discontinuités, ou frontières, dans les valeurs d'une variable mesurée sur des points répartis spatialement sur le terrain (c'est-à-dire sur les p-e). Cette type d'analyse est utilisée dans le domaine de l'écologie pour identifier des écotones dans la végétation et l'environnement, pour analyser la répartition spatiale de la faune en relation à la végétation ainsi que pour déterminer des aires dominées par un même

processus stationnaire (Fagan *et al.* 2003, Fortin 1994, Fortin et Drapeau 1996, Fortin *et al.* 1996, 2000, Hall et Maruca 2001, Wagner et Fortin 2005).

En principe, deux types de frontières peuvent être identifiés, d'une part les frontières indirectes qui résultent de la délimitation des zones homogènes et d'autre part, les frontières directes qui indiquent les zones de changement intense de valeurs des variables considérées et qui ne forment pas nécessairement des polygones fermés (Fortin 1994, Fortin et Drapeau 1995, Fortin *et al.* 2000, Jacquez et Maruca 2001, Jacquez *et al.* 2000).

L'objectif général de ce chapitre est donc de déterminer les caractéristiques spatiales des frontières des valeurs du volume ligneux mesurées sur divers points répartis sur le territoire, ainsi que de vérifier les frontières forestières ou environnementales auxquelles elles sont spatialement associées à l'échelle déterminée par l'espacement entre les points. Les possibilités de fusion de ces frontières issues des données de terrain avec les frontières cartographiques seront discutées au Chapitre VII.

Les objectifs spécifiques de ce chapitre sont :

1. D'identifier et de comparer les frontières indirectes (issues des groupes de p-e contigus semblables) et les frontières directes (issues des fortes différences entre les valeurs) pour les variables d'intérêt : volume ligneux, type forestier, topographie et sol;
2. De caractériser les coïncidences spatiales entre les deux types de frontières (directes et indirectes) pour le volume ligneux;
3. D'identifier, pour les variables d'intérêt, la localisation des groupes de frontières directes particulières en terme de leur intensité;
4. D'identifier les variables forestières et environnementales auxquelles sont associées spatialement les frontières directes et indirectes du volume ligneux.

## Méthodologie

### 1. Données et variables

Dans le cadre de cette étude, les données des placettes-échantillons (p-e) temporaires de l'année 1992 (342 p-e) et celles du réseau permanent mesuré entre 1995 et 1999 (201 p-e) sont utilisées. Les variables qualitatives forestières de ces deux ensembles de données correspondent au même système de classification. Sur l'ensemble des 543 p-e, 539 ont été retenues pour la détection des frontières. Quatre p-e ne contenant pas d'information pour les variables traitées ont été éliminées de l'analyse. Les variables ainsi que leurs catégories sont décrites au Chapitre II et à l'Annexe I.

Les frontières sont identifiées par les méthodes décrites à la section suivante pour des variables simples et pour des variables composites (qui représentent la valeur moyenne de plusieurs variables simples), de nature quantitative et qualitative (Tableau 20). Toutes les variables simples ainsi que celles qui intègrent les variables composites proviennent des p-e, c'est-à-dire qu'elles ont été estimées ou mesurées sur le terrain.

Tableau 20. Variables considérées pour l'analyse des frontières.

Nom	Nature	Type	Description/Définition
<i>Volume total</i>	Quantitative (m <sup>3</sup> /ha)	Simple	Volume de bois des essences résineuses et feuillues, estimé à partir des arbres mesurés dans la p-e
<i>Type forestier</i>	Qualitative	Composite	Ensemble des variables qualitatives <i>Groupement d'essences, Classe de densité, Classe de hauteur et Classe d'âge</i>
<i>Sol</i>	Qualitative	Composite	Ensemble des variables qualitatives : <i>Type d'humus, Texture de l'horizon B, Texture de l'horizon C, Dépôt de surface et Drainage</i>
<i>Topo-q</i>	Quantitative	Composite	Ensemble des variables quantitatives <i>Altitude, Exposition et Inclinaison</i>
<i>Topo-c</i>	Qualitative	Composite	Ensemble des variables qualitatives <i>Versant, Situation sur la pente et Forme de la pente</i>

La manière dont les frontières des variables composites sont estimées est détaillée dans la section suivante. En ce qui concerne la variable composite quantitative *Topo-q*, étant donné que les variables quantitatives qui la composent possèdent différentes étendues de valeurs, la variable composite est normalisée à un intervalle de valeurs variant entre 0 et 1, en utilisant l'étendue de valeurs de chaque variable (Jacquez et Maruca 2001).

## 2. Méthodes

La Figure 15 illustre les étapes méthodologiques de ce chapitre, qui sont détaillées aux sections subséquentes.

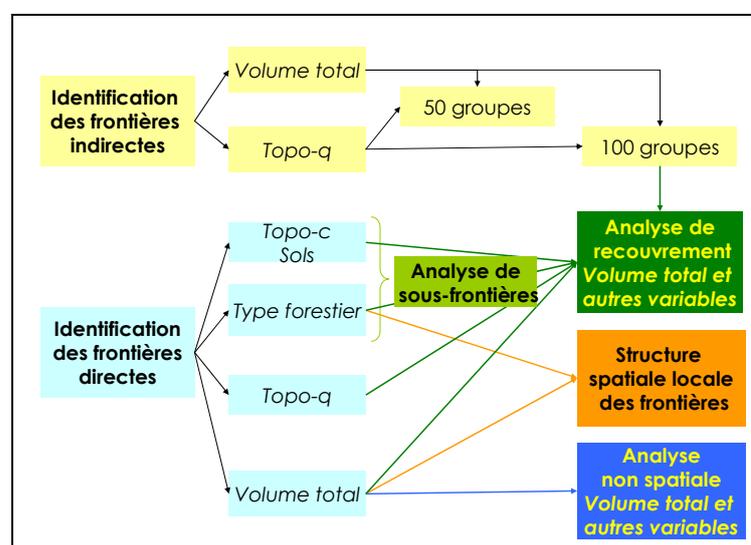
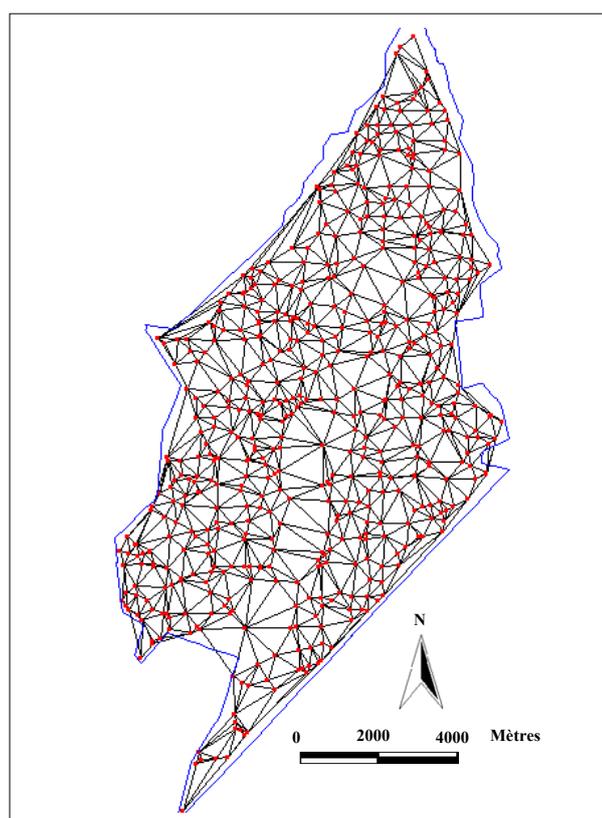


Figure 15. Étapes méthodologiques du Chapitre IV.

### 2.1. Identification des groupes spatiaux : frontières indirectes

Le but de cette méthode est de regrouper les p-e adjacentes dans l'espace selon le degré de ressemblance des valeurs des variables considérées et donc, d'identifier des groupes de p-e adjacentes homogènes par rapport aux valeurs des variables. La délimitation des groupes homogènes détermine de façon indirecte les frontières, qui représentent les bordures entre les groupes contigus (Fortin et Drapeau 1995).

Ces frontières indirectes sont détectées par la technique de regroupement à l'aide de contraintes spatiales (Jacquez et Maruca 2001). L'adjacence entre les p-e est déterminée au moyen d'une triangulation de Delaunay (Figure 16). La ressemblance entre deux p-e est déterminée par une mesure de dissimilarité, dans ce cas la distance euclidienne au carré, entre les valeurs de la variable dans les deux p-e. Cette mesure produit les mêmes résultats que la distance euclidienne mais elle est plus performante pour le traitement d'un grand nombre de données (Jacquez et Maruca 2001). Lorsqu'il s'agit d'une variable composite, la distance euclidienne au carré de cette variable correspond à la somme des distances de chacune des variables individuelles qui la composent (Jacquez et Maruca 2001).



Points rouges : p-e  
Lignes noires : liens entre les p-e.

Figure 16. Réseau de triangulation de Delaunay qui établit l'adjacence entre les p-e de la Forêt Montmorency.

La méthode de regroupement est agglomérative. Au début, chaque p-e constitue un groupe et, de façon itérative, l'algorithme agglomère ces groupes. À chaque itération, les groupes adjacents les plus semblables (c'est-à-dire ceux dont les valeurs de dissimilarité sont les plus faibles) sont fusionnés. Le processus d'itération s'arrête lorsque les groupes formés atteignent le nombre de groupes fixé au début par l'analyste, ce qui représente le critère d'arrêt. Afin de guider le choix du nombre de groupes à obtenir, un *indice d'ajustement* est calculé pour le groupement :

$$\text{Indice d'ajustement} = [B / (k-1)] / [W / (n-k)], \quad \text{IV.1}$$

où  $B$  est la somme des carrés des écarts entre tous les groupes,  $W$  est la somme des carrés des écarts intra-groupes,  $k$  est le nombre de groupes et  $n$  est le nombre de p-e. L'objectif est de choisir le nombre de groupes qui maximise la différence entre  $B$  et  $W$ , et qui correspond à la valeur la plus élevée possible de l'*indice d'ajustement* (Jacquez et Maruca 2001).

L'algorithme de regroupement utilisé est celui du lien flexible avec une connectivité de 0,5, dans lequel chaque p-e d'un groupe est comparée à chacune des p-e d'un groupe adjacent (Jacquez et Maruca 2001, Legendre et Legendre 1998). Finalement, les groupes sont raffinés en utilisant l'algorithme des centres de groupes (« k-means ») afin de minimiser l'hétérogénéité intra-groupe (Legendre et Fortin 1989, Fortin et Drapeau 1995, Jacquez et Maruca 2001).

Le nombre de groupes à former est fixé à 50 et 100, en tenant compte de l'*indice d'ajustement* et de la taille moyenne des groupes (soit entre 5 et 10 p-e). Les variables traitées sont *Volume total* et *Topo-q* (voir Tableau 20).

## **2.2. Identification des différences : frontières directes**

Cette méthode vise à identifier les secteurs pour lesquels les valeurs des variables mesurées dans les p-e adjacentes présentent des différences élevées. Les frontières sont ainsi identifiées de façon directe comme des emplacements spatialement contigus de changement intense dans les variables considérées (Fortin et Drapeau 1995).

Les frontières directes sont identifiées au moyen de la méthode de détection de changements de Womble pour des points distribués irrégulièrement (Barbujani *et al.* 1989, Fortin 1994, Fortin et Drapeau 1995, Jacquez et Maruca 2001). Selon que la variable est quantitative ou qualitative, les méthodes d'identification de frontières sont différentes, soient les méthodes de triangulation et qualitative de Womble respectivement.

### 2.2.1. Variables quantitatives

Pour les variables quantitatives, la méthode vise à mesurer la magnitude du changement entre les valeurs des p-e adjacentes dans l'espace. Il faut d'abord déterminer les relations d'adjacence spatiale entre les p-e et mesurer ensuite le degré de changement entre les p-e adjacentes.

Premièrement, les relations de proximité entre les p-e sont établies par la triangulation de Delaunay (Figure 16). Puis, le taux de changement, c'est-à-dire la première dérivée partielle des valeurs de la variable dans les trois p-e de chaque triangle, est calculé. Cette valeur indique la magnitude du changement de la variable dans l'espace défini par les trois p-e adjacentes formant un triangle. Elle se nomme « valeur de vraisemblance » et est assignée au centroïde du triangle, qui dévient ainsi un « Candidat à Élément de Frontière (CEF) » (Figure 17a) (Fortin 1994, Fortin et Drapeau 1995, Jacquez et Maruca 2001). Dans cette recherche, étant donnée l'irrégularité des surfaces et des formes des triangles, la valeur de vraisemblance originale est pondérée par la surface du triangle (en ha).

Les CEFs dont la valeur de vraisemblance est supérieure à un seuil déterminé par l'utilisateur sont considérées comme des « Éléments de Frontière (EFs) » (Figure 17a). Dans la littérature (Fortin 1994, 1997, Fortin et Drapeau 1995, Fortin *et al.* 1996), ce seuil est normalement fixé au 95<sup>e</sup> ou 90<sup>e</sup> percentile. Cela signifie que 5 % ou 10 % des CEFs dont les valeurs de vraisemblance sont les plus élevées sont considérés comme des EFs. Pour cette étude, vue la forte hétérogénéité environnementale (c'est-à-dire qu'il y a plus d'un processus agissant sur l'aire d'étude), le degré d'intervention forestière et le grand espacement entre les p-e, le seuil considéré correspond au 50<sup>e</sup> percentile, c'est-à-dire que

50 % des CEFs dont les valeurs de vraisemblance sont les plus élevées sont considérées comme des EFs.

Par la suite, les EFs adjacents sont connectés pour former des sous-frontières, à condition de satisfaire à deux critères (Figure 17a). Le premier établit l'angle maximal entre les deux gradients définis par les EFs adjacents (Figure 17b). Ce critère vise à éviter la connexion d'EFs dont la direction du gradient est opposée. Dans cette étude, l'angle maximal est fixé à  $180^\circ$  en raison de la forte hétérogénéité environnementale mentionnée plus haut. Le deuxième critère établit l'angle minimal entre le vecteur qui représente la direction du gradient et la connexion à établir (Figure 17c). Ce critère vise à s'assurer que les sous-frontières définies par la connexion entre les EFs adjacents ne se situent pas dans la même direction que le gradient de changement, étant donné qu'en principe, une frontière sépare deux aires dissimilaires. Dans le cas présent, cet angle minimal est fixé à  $30^\circ$ .

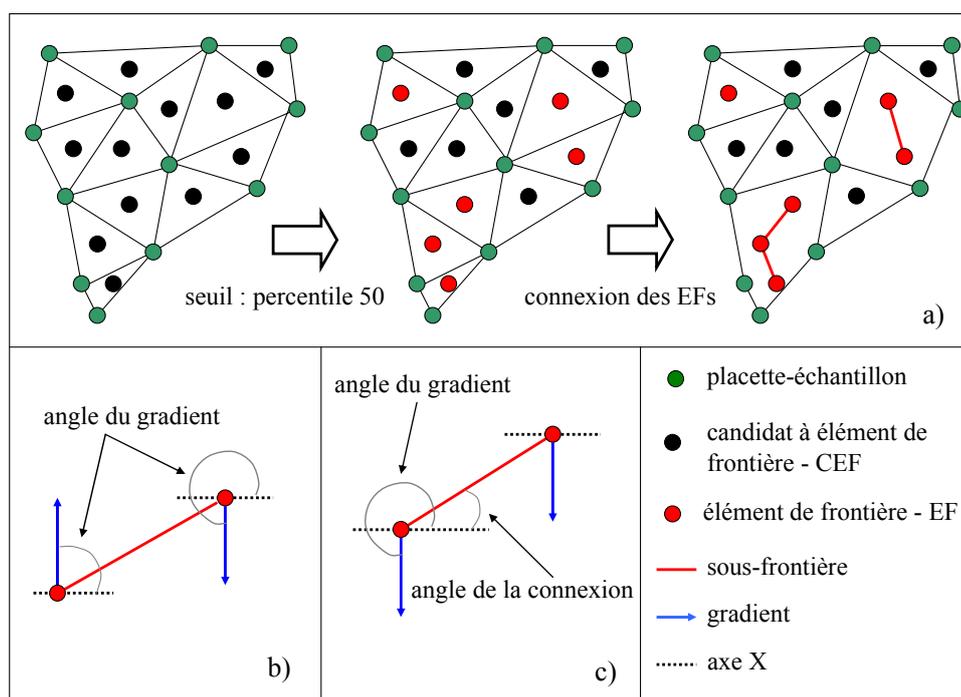


Figure 17. Identification des frontières directes par la méthode de triangulation de Womble pour les variables quantitatives. a) Identification des CEFs et des EFs, b) angle du gradient entre les EFs, c) angle entre le gradient et la connexion entre les EFs. (Adaptée de Jacquez et Maruca 2001).

En utilisant cette méthode, les variables pour lesquelles les frontières sont identifiées sont *Volume total* et *Topo-q* (Tableau 20). Dans le cas de la variable composite *Topo-q*, la valeur de vraisemblance est estimée pour chacune des variables individuelles et la moyenne des valeurs de vraisemblance individuelles est considérée comme la valeur de vraisemblance de la variable composite.

### 2.2.2. Variables qualitatives

Pour les variables qualitatives, la méthode vise à estimer une mesure de la coïncidence pour les catégories présentes dans les p-e adjacentes.

L'identification des frontières est faite au moyen de la méthode qualitative de Womble (Oden *et al.* 1993). Selon cette méthode, une mesure de dissimilarité est utilisée pour estimer la valeur de vraisemblance (VV) entre deux p-e adjacentes. La mesure de dissimilarité utilisée est la valeur de non-coïncidence obtenue comme suit :

$$VV = \text{nombre de variables pour lesquelles } z_i \neq z_j / \text{nombre total de variables}, \quad \text{IV.2}$$

où  $z_i$  et  $z_j$  sont les catégories d'une variable  $z$  dans les p-e  $i$  et  $j$ . Cette valeur de vraisemblance est assignée à la ligne équidistante des deux p-e adjacents, qui dévient une Candidate à Élément de Frontière (CEF). Les CEFs sont considérées comme des EFs en fonction de la proportion de catégories qui ne coïncident pas. Cette proportion dépend de la variable composite considérée (Tableau 21, Figure 18).

Tableau 21. Proportion de non coïncidences pour que les CEFs soient considérées comme des EFs.

Variable composite	Pourcentage minimal de non-coïncidence
<i>Type forestier</i>	75 (3 variables non coïncidentes / 4 variables)
<i>Topo-c</i>	66 (2 variables non coïncidentes / 3 variables)
<i>Sols</i>	66 (3 variables non coïncidentes / 5 variables)

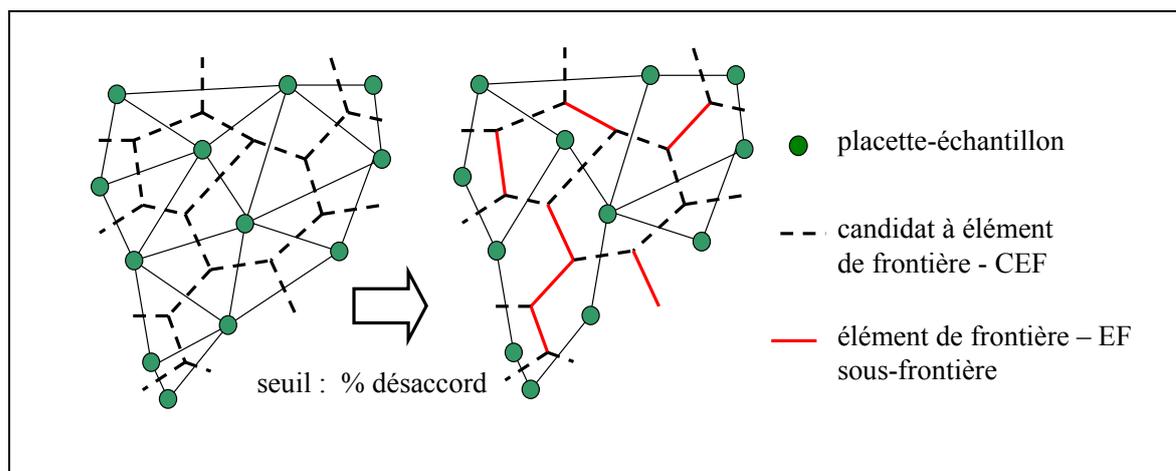


Figure 18. Identification des frontières directes par la méthode de Womble pour des variables qualitatives (Adaptée de Jacquez et Maruca 2001).

### 2.3. Analyse de frontières

Deux types d'analyses sont faites sur les frontières directes et indirectes identifiées précédemment : (1) l'analyse de sous-frontières, qui vise à déterminer si les connexions entre les EFs (les sous-frontières) sont significatives du point de vue statistique ou si elles sont dues au hasard, et (2) l'analyse de recouvrement, qui vise à déterminer si les coïncidences spatiales entre les frontières pour différentes variables ou entre les frontières directes et indirectes d'une même variable sont statistiquement significatives (Fortin et Drapeau 1995, Jacquez et Maruca 2001, Jacquez *et al.* 2000).

#### 2.3.1. Analyse de sous-frontières

L'analyse des sous-frontières évalue, au moyen des statistiques de sous-frontières, la contiguïté des frontières directes. Ces statistiques permettent de déterminer si les caractéristiques des sous-frontières d'une variable tels la longueur, le degré de ramifications et le diamètre sont statistiquement différentes d'un système aléatoire (Jacquez et Maruca 2001). Les statistiques des sous-frontières sont énumérées au Tableau 22.

Les distributions de référence des statistiques des sous-frontières sont obtenues au moyen de 1000 itérations Monte Carlo avec une répartition aléatoire spatiale complète (Jacquez et Maruca 2001). Dans cette procédure, les valeurs des variables sont assignées de façon

Tableau 22. Description des statistiques de sous-frontières.

<i>NS</i>	nombre de sous-frontières
<i>NI</i>	nombre de EFs isolés
<i>Lmax</i>	longueur maximale des sous-frontières (nombre de EFs connectés)
<i>Lmoy</i>	longueur moyenne des sous-frontières
<i>Dmax</i>	diamètre maximal des sous-frontières (le diamètre est la longueur du chemin le plus court entre chaque paire de EFs dans une sous-frontière)
<i>Dmoy</i>	diamètre moyen des sous-frontières
<i>D/L</i>	ratio moyen diamètre / longueur (indique le degré de ramification des sous-frontières)

aléatoire aux localisations des p-e à chaque itération, les valeurs de vraisemblance sont également recalculées, les EFs sont identifiés (avec les mêmes critères que les EFs des frontières testées), les sous-frontières sont définies et les statistiques de ces sous-frontières sont calculées. Ces statistiques obtenues sur les sous-frontières sont comparées à la distribution de référence afin d'évaluer la signification statistique des frontières directes ( $\alpha=0,05$ ).

Selon la procédure de répartition aléatoire spatiale complète, l'hypothèse nulle de cette analyse est la suivante : les sous-frontières formées sont le fruit du hasard et la répartition spatiale des EFs est complètement aléatoire. Deux hypothèses alternatives sont considérées :

1. Les frontières correspondent à un processus d'échelle globale, c'est-à-dire que les discontinuités spatiales de la variable reflètent un phénomène global. Cette hypothèse est retenue lorsque les valeurs des statistiques des sous-frontières sont significativement élevées et montrent une forte contiguïté des frontières;
2. Les frontières sont fragmentées, c'est-à-dire que la variable reflète un processus de morcellement dans le territoire. Cette hypothèse est retenue lorsque les valeurs obtenues des statistiques des sous-frontières sont significatives et faibles, en montrant une très faible contiguïté des frontières.

L'analyse des sous-frontières est appliquée aux frontières directes des variables qualitatives *Type forestier*, *Topo-c* et *Sol*.

### 2.3.2. Analyse de recouvrement

Cette méthode est employée pour évaluer la coïncidence spatiale entre les frontières indirectes (issues de groupes) et directes (issues de différences) d'une même variable et/ou la coïncidence spatiale entre les frontières de différentes variables (Fortin *et al.* 1996). L'objectif poursuivi est de déterminer si à cette échelle existe une association spatiale entre les frontières du *Volume total* et les frontières forestières et/ou environnementales.

À la suite de la superposition des deux ensembles de frontières à évaluer, la procédure consiste à calculer quatre statistiques de recouvrement basées sur la distance moyenne du plus proche voisin, afin d'évaluer leur association spatiale (Tableau 23) (Jacquez et Maruca 2001). Les analyses de recouvrement sont effectuées pour : (1) évaluer l'association spatiale entre les frontières indirectes et directes du *Volume total*, et (2) pour évaluer l'association spatiale des frontières indirectes du *Volume total* (issues de 100 groupes) avec les frontières directes de *Type forestier*, *Sol*, et *Topo-c*, ainsi qu'avec les frontières indirectes de *Topo-q*. Les frontières indirectes du *Volume total* issues de 50 groupes ne sont pas comparées en raison du nombre de frontières nettement inférieur par rapport à celui des autres variables.

Tableau 23. Statistiques de recouvrement. *G* et *H* sont deux ensembles de frontières.

$O_S$	nombre d'éléments de frontières (EFs) qui sont présents dans les deux ensembles de frontières G et H
$O_G$	distance moyenne entre les EFs de l'ensemble de frontières G et le EF le plus proche de l'ensemble de frontières H
$O_H$	distance moyenne entre les EFs de l'ensemble de frontières H et le EF le plus proche de l'ensemble de frontières G
$O_{GH}$	distance moyenne entre un EF de n'importe lequel des deux ensembles et le EF le plus proche de l'autre ensemble

Les distributions de référence des statistiques de recouvrement sont obtenues au moyen de 1000 itérations Monte Carlo selon deux hypothèses :

1. Les frontières indirectes du *Volume total* sont dépendantes des frontières des variables environnementales et forestières. Pour tester cette hypothèse, seule la variable dépendante est traitée avec une répartition aléatoire spatiale complète;

2. Les frontières indirectes du *Volume total* sont associées à celles des variables environnementales et forestières. Pour tester cette hypothèse, les deux variables sont soumises à une répartition aléatoire spatiale complète (Jacquez et Maruca 2001).

L'hypothèse nulle de cette analyse est la suivante : le recouvrement spatial des frontières est aléatoire. Trois hypothèses alternatives sont proposées :

1. Les frontières des deux variables sont spatialement associées. Cette hypothèse est retenue lorsque les deux variables sont soumises à une répartition aléatoire spatiale complète et que les statistiques indiquent un recouvrement entre les frontières supérieur à celui qui serait dû au hasard;
2. La superposition est directionnelle et les frontières du *Volume total* dépendent des frontières de l'autre variable testée. Cette hypothèse est retenue lorsque la variable dépendante est soumise à une répartition aléatoire spatiale complète et que les statistiques indiquent un recouvrement supérieur à celui qui serait dû au hasard;
3. Les frontières des deux variables se repoussent. Cette hypothèse est retenue lorsque les statistiques indiquent un recouvrement inférieur à celui qui serait dû au hasard (Jacquez et Maruca 2001).

#### **2.4. Analyse de la structure spatiale locale des frontières**

À l'aide des statistiques d'autocorrélation spatiale (Anselin 1995, Boots 2002, Fortin *et al.* 2002), la description de la structure spatiale locale permet d'établir la localisation des CEFs dont les valeurs de vraisemblance sont significativement différentes de la moyenne de l'ensemble des CEFs (ici CEF est utilisé dans un sens général qui inclut les EFs, voir Figure 17).

L'autocorrélation spatiale locale de chaque CEF est estimée au moyen de l'*I* de Moran local, pondéré par la distance inverse entre les points et ajusté pour les courtes distances (Anselin 1995, Levine *et al.* 2002). Une autocorrélation locale significative et positive indique une homogénéité des valeurs de vraisemblance des CEF voisins, alors qu'une

valeur négative indique que les valeurs de vraisemblance des CEFs voisins sont différentes (Boots 2002). Dans cette étude, d'après le critère de Levine *et al.* (2002), les valeurs d'*I* de Moran local sont considérées comme significatives lorsqu'elles appartiennent à l'intervalle, positif ou négatif, entre un et trois écarts-type de la moyenne. Les valeurs supérieures aux trois écarts-type sont considérées comme aberrantes.

La structure spatiale locale est décrite pour les frontières directes (CEFs) des variables *Volume total* et *Type Forestier*. Chaque CEF d'une variable quantitative ainsi que le centre de la ligne qui représente chaque CEF d'une variable qualitative constituent un point.

### **2.5. Relation non spatiale entre les frontières directes du Volume total et les variables forestières et environnementales**

Le but de l'analyse présentée dans cette section est de déterminer si les combinaisons d'attributs des p-e, qui forment les frontières directes du *Volume total*, présentent des caractéristiques forestières ou environnementales particulières par rapport aux combinaisons d'attributs des p-e qui ne déterminent pas de frontières directes.

En premier lieu, les Candidats à Élément de Frontière (CEFs) sont séparés en deux groupes, ceux qui dépassent le seuil fixé au 50<sup>e</sup> percentile et qui sont considérés comme des Éléments de Frontière (EFs), et ceux qui sont inférieurs au seuil et qui demeurent des CEFs. Pour chacun des deux groupes, les trois p-e associées à chaque CEF et à chaque EF sont identifiées. Les combinaisons de catégories des variables qualitatives (*Groupement d'essences*, *Classe de densité*, *Classe de hauteur*, *Classe d'âge* et *Perturbation d'origine*) des trois p-e sont identifiées et leur fréquence est établie pour les EFs et pour les CEFs. Les différences entre le maximum et le minimum des variables quantitatives (*Densité totale*, *Surface terrière* et *Hauteur dominante*) dans les trois p-e sont calculées pour chaque variable. Les distributions de fréquence de ces différences « maximum-minimum » sont estimées pour les EFs ainsi que pour les CEFs.

Les distributions de fréquence des combinaisons des variables qualitatives et des « maximum–minimum » des variables quantitatives sont comparées entre les EFs et les

CEFs. Les différences entre les distributions de fréquence sont évaluées avec le test de Kolmogorov-Smirnov sur deux échantillons (Insightful 2001).

## Résultats et discussion

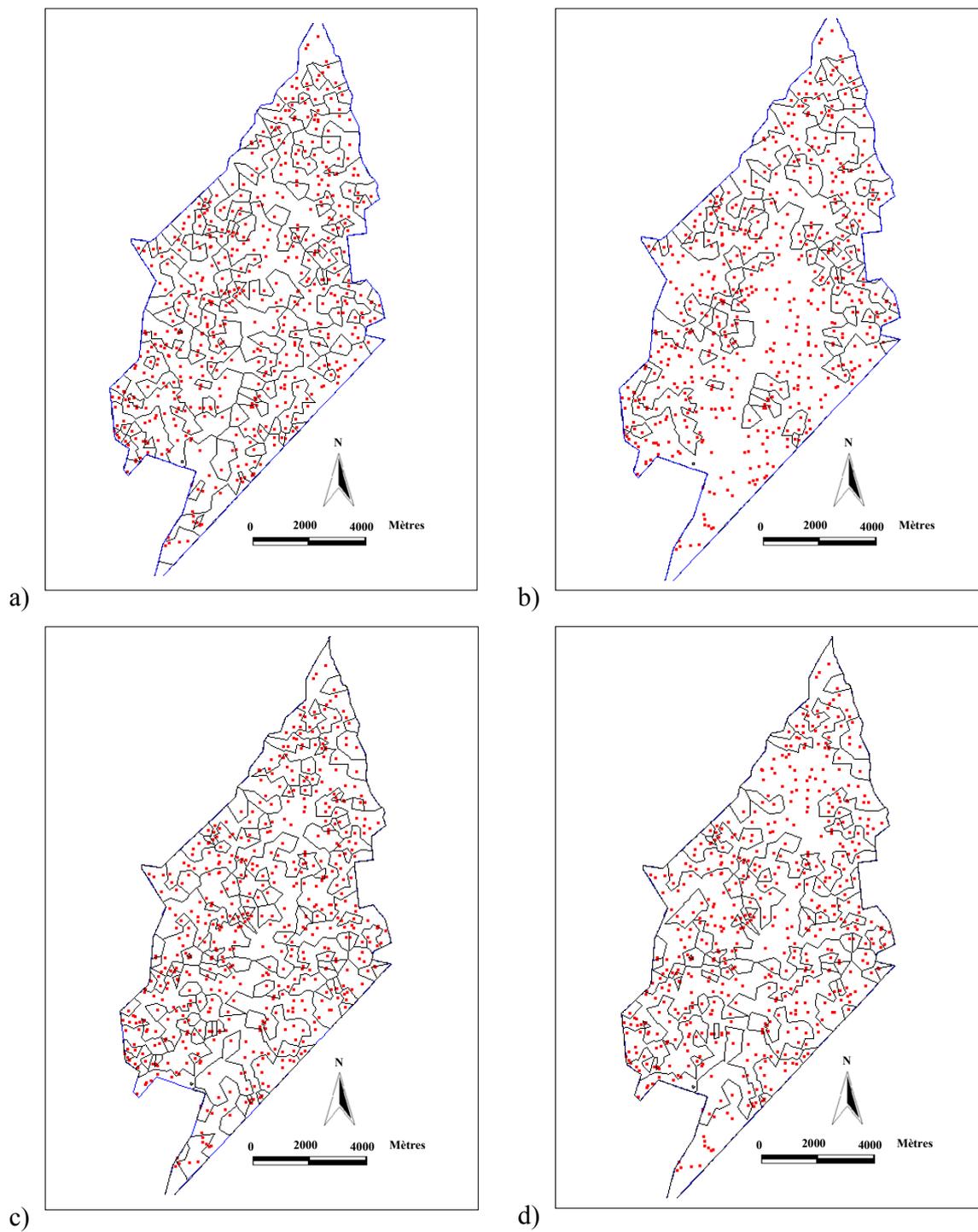
Les résultats sont présentés en quatre sections : (1) les frontières indirectes des groupes spatiaux; (2) les frontières directes résultant des différences entre les valeurs des variables; (3) les analyses de sous-frontières et de recouvrement de frontières et (4) la structure spatiale locale des frontières.

### 1. Les groupes spatiaux et les frontières indirectes

Ces résultats correspondent au regroupement des p-e adjacentes selon la ressemblance de leurs valeurs pour les variables *Volume total* et *Topo-q*. Les résultats du regroupement des p-e en 100 groupes, selon ces deux variables, sont semblables en ce qui a trait à la répartition spatiale des groupes et à la distribution de leurs tailles (Figure 19a, et c). Les groupes formés par une seule p-e prédominent (environ 50 % des groupes), tandis qu'entre trois et cinq groupes contiennent autour de 5 % des p-e chacun. La variable *Volume total* présente également un groupe qui contient la majorité des p-e (environ 40 % des p-e).

Lorsque les p-e sont regroupées en 50 groupes, la distribution des groupes par rapport à la taille est similaire à celle de 100 groupes et ce, pour les deux variables. Dans le cas du *Volume total*, des petits polygones sont concentrés dans le secteur nord de la zone d'étude (Figure 19b) tandis que pour *Topo-q* les groupes sont répartis de façon plutôt uniforme dans la zone d'étude (Figure 19d).

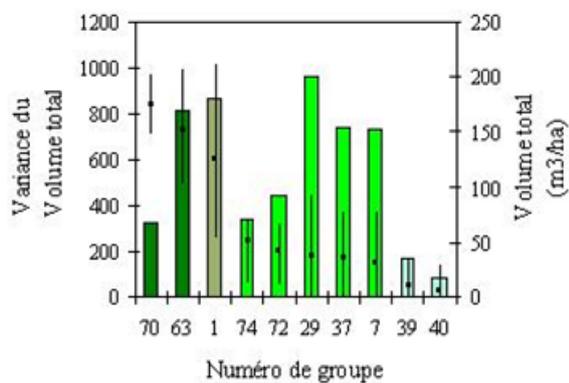
Afin de déterminer les caractéristiques forestières des groupes de p-e qui définissent des zones homogènes, les valeurs du *Volume total* des p-e qui composent les groupes contenant le plus grand nombre de p-e sont analysées. Cette analyse indique que les étendues des valeurs et les moyennes du *Volume total* sont similaires aux deux échelles (Figure 20). La variance des groupes de grande taille (avec de nombreux p-e) est supérieure dans le regroupement en 50 groupes que dans le regroupement en 100 groupes (Figure 20a et b).



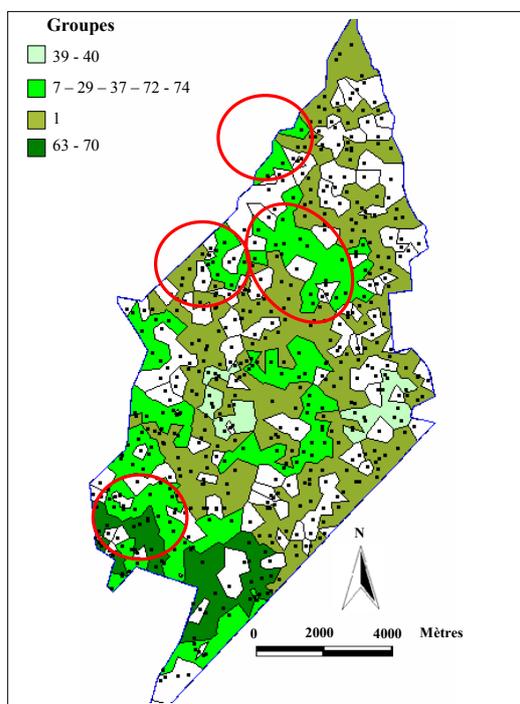
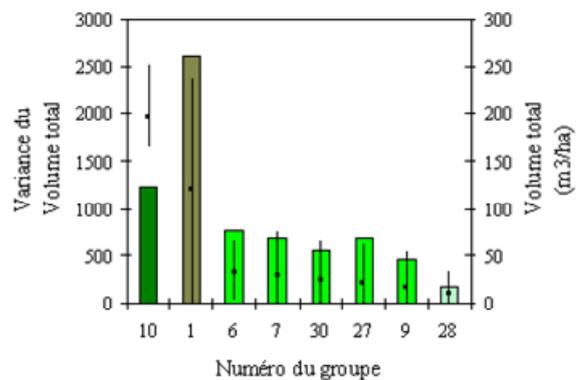
Lignes noires : frontières indirectes – Points rouges : p-e

Figure 19. Frontières indirectes du *Volume total* : a) 100 groupes; b) 50 groupes, c) *Topo-q*, 100 groupes et d) *Topo-q*, 50 groupes.

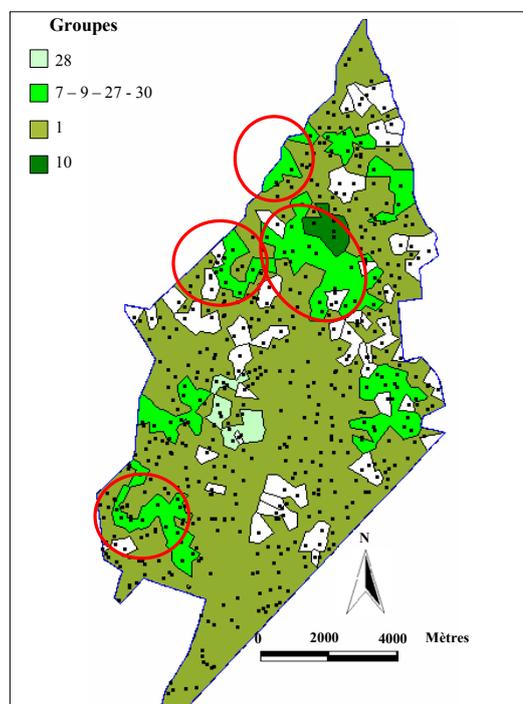
a)



b)



c)



d)

a et b : les lignes verticales indiquent l'étendue des valeurs du *Volume total* (minimum-maximum) de chaque groupe, le point central de chaque ligne indique sa moyenne et les barres jaunes indiquent sa variance.

c et d : les cercles rouges indiquent les groupes qui persistent dans les deux regroupements.

Figure 20. Frontières indirectes pour le *Volume total*. Valeurs minimales, maximales, moyennes et de variance du *Volume total* des groupes qui contiennent au moins 1 % des p-e pour : a) 100 groupes et b) 50 groupes. Répartition spatiale des : c) 100 groupes et d) 50 groupes.

Cette information permet d'identifier, à chaque échelle, des groupes qui se ressemblent en terme du *Volume total*. Par exemple, à l'échelle de 100 groupes, les groupes 37 et 7

(produits par l'algorithme, ces chiffres servent à identifier les groupes) sont semblables (Figure 20a et c). À l'échelle de 50 groupes, les groupes 6, 7, 30 et 27 se ressemblent (Figure 20b et d). Certains de ces groupes persistent lorsque l'échelle change (Figure 20c et d, groupes encerclés en rouge). Cette persistance indiquerait que ces groupes diffèrent suffisamment de leur entourage pour maintenir leur « identité » définie par la moyenne du *Volume total*. De plus, aux deux échelles, les groupes dont les moyennes sont plus élevées présentent fréquemment des p-e qui ont une végétation pré-mature, tandis que les groupes de faibles moyennes présentent, au contraire, des p-e qui ont une végétation en régénération à la suite d'une coupe totale.

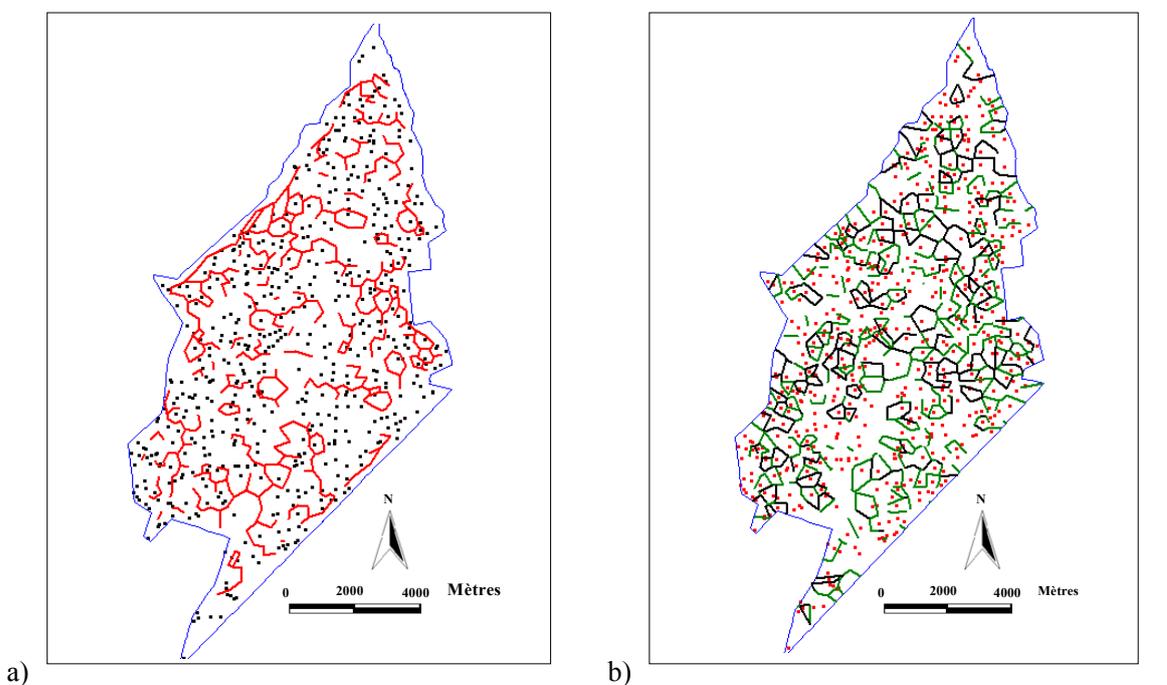
La distribution de fréquence de tailles des groupes déterminés pour le *Volume total* permet d'interpréter la zone d'étude comme une « matrice », dans le sens du terme donné par Forman (1995). Cette « matrice » correspond au groupe le plus étendu dans lequel les conditions forestières sont intermédiaires et hétérogènes. Une dizaine de parcelles de grande taille y sont incluses (correspondant aux groupes de taille intermédiaire détaillés à la Figure 20) avec des volumes ligneux plus élevés ou plus faibles que la matrice. De nombreuses parcelles de taille inférieure sont également incluses dans cette matrice et correspondent à des conditions très locales (les groupes formés d'une ou de deux p-e). Cette idée de matrice spatiale est accentuée lorsque l'échelle change à 50 groupes, car dans le secteur sud, la matrice dévient prédominante (Figure 20). L'hétérogénéité des conditions de cette matrice est alors mise en évidence par ses valeurs très supérieures de variance (Figure 20b).

La persistance de certains groupes aux deux échelles indiquerait des conditions suffisamment dissemblables des p-e voisines pour maintenir une stabilité en dépit du changement d'échelle. Les frontières de ces groupes pourraient être identifiables sur d'autres sources d'information comme la carte forestière, dont les unités sont délimitées à partir de variables comme la *Classe de hauteur* ou la *Classe d'âge*, pour lesquelles les groupes formés pour le *Volume total* exhibent une prédominance de catégories particulières.

## 2. Les frontières directes

Les frontières directes, issues des différences élevées dans les valeurs de variables entre p-e adjacentes, sont présentées, en premier, pour la variable *Volume total* et par la suite pour les variables composites « thématiques » *Type forestier*, *Topo-q*, *Topo-c* et *Sol* (voir le Tableau 20 pour la description des variables).

Les frontières directes de toutes les variables sont réparties sur toute l'étendue de la zone d'étude (Figure 21). En général, elles forment des chaînes de sous-frontières qui délimitent de nombreuses surfaces. La plupart de ces surfaces contiennent une p-e comme c'est le cas, par exemple, de celles du *Volume total* (Figure 21a).



a) Points noirs ou rouges : p-e

Lignes rouges : frontières des variables quantitatives

Lignes vertes : frontières de variables composites dont 66 à 99 % des variables sont en désaccord

Lignes noires : frontières des variables composites dont 100 % des variables sont en désaccord

Figure 21. Frontières directes entre les p-e pour les variables : a) *Volume total* et b) *Type forestier*.

La majorité des frontières de *Type forestier*, *Topo-c* et *Sol* présente un désaccord d'au moins 66 % entre les variables (soit de deux variables sur trois pour *Topo-c* et *Sol*, soit de trois variables sur quatre pour *Type forestier*) (Figure 21b, d et e). Entre 15 et 27 % des

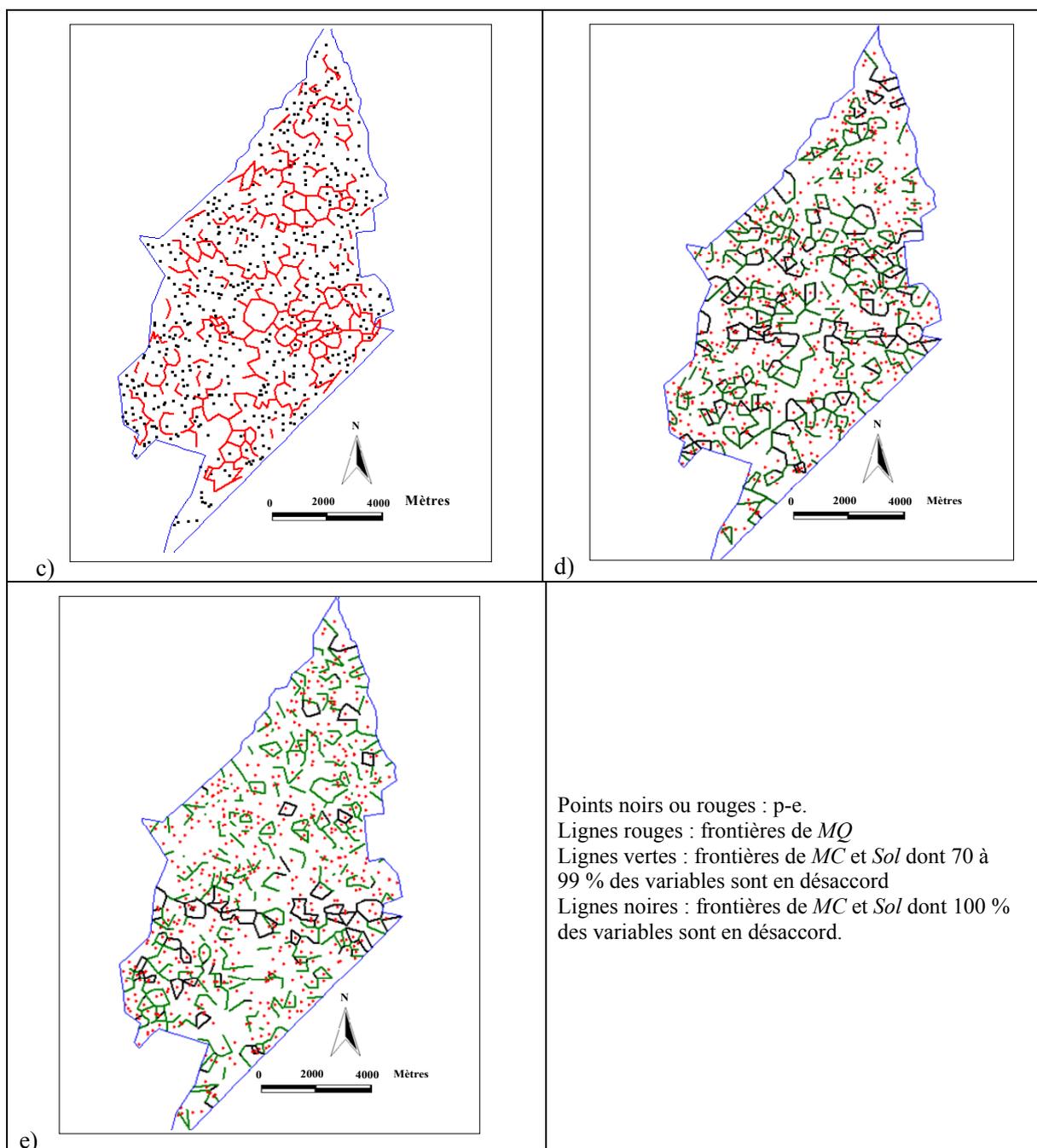


Figure 21. Suite. Frontières directes entre les p-e pour les variables : c) *Topo-q*, d) *Topo-c*, e) *Sol*, f) *MQ*, g) *MC*.

frontières directes présentent 100 % de désaccord entre les variables. Près de 60 % de ces frontières pour *Topo-c* et *Sol* sont localisées dans le secteur central de la Forêt Montmorency (Figure 21d et e). Comme pour les autres variables, les frontières montrent une tendance à s'enchaîner, tandis que les frontières « isolées » sont minoritaires.

### 3. Signification statistique des frontières et de leurs coïncidences spatiales

Les analyses des sous-frontières des variables qualitatives qui visent à établir leur signification statistique sont présentées en premier. Par la suite, les analyses de recouvrement des frontières sont décrites afin d'établir la signification statistique des coïncidences spatiales de différents ensembles de frontières.

#### 3.1. Analyses des sous-frontières

Les statistiques des sous-frontières permettent d'évaluer si les frontières directes des variables qualitatives sont significativement différentes de celles produites par hasard. Les distributions de référence des statistiques sont obtenues par répartition aléatoire.

Les statistiques de sous-frontières des variables *Type forestier* et *Topo-c* ne présentent pas de valeurs significativement élevées ou faibles par rapport aux distributions de référence ( $\alpha=0,10$ ) et l'hypothèse nulle selon laquelle les frontières sont un résultat du hasard ne peut pas être rejetée (Tableau 24). Pour la variable *Sol* ces statistiques indiquent, au contraire, que le nombre de EFs connectés en sous-frontières est faible, que le chemin entre chaque paire d'EFs est court et que les frontières sont très ramifiées (Tableau 24). Cette situation peut s'interpréter comme correspondant à un paysage morcelé (Jacquez et Maruca 2001).

Tableau 24. Signification des statistiques de sous-frontières pour les variables qualitatives ( $\alpha= *-0,01$ ,  $**--0,05$  et  $***--0,10$ ). M : indique un paysage morcelé. C : indique une forte contiguïté des frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au Tableau 20 et les statistiques au Tableau 22.

Variable	$N_s$	$N_l$	$L_{max}$	$L_{moy}$	$D_{max}$	$D_{moy}$	$D/L$
<i>Type forestier</i>							
<i>Topo-c</i>							
<i>Sol</i>	** C	*** M		** M		** M	** M

### 3.2. Analyses de recouvrement

Les analyses de recouvrement sont d'abord présentées pour évaluer l'association entre les frontières directes et indirectes du *Volume total* et ensuite, pour évaluer l'association et la dépendance des frontières indirectes du *Volume total* aux frontières directes des variables *Type forestier*, *Topo-c* et *Sol*, ainsi qu'aux frontières indirectes de *Topo-q*.

#### 3.2.1. Association entre frontières indirectes et directes du *Volume total*

Les analyses de recouvrement indiquent une association spatiale entre les frontières directes, issues des différences entre les variables, et les frontières indirectes, dérivées du regroupement des p-e en 100 et en 50 groupes pour la variable *Volume total* (Tableau 25).

Tableau 25. Signification des probabilités des statistiques de recouvrement qui évaluent l'association des frontières directes et indirectes du *Volume Total* et de *MQ* ( $\alpha = *$ --0,01,  $**$ --0,05 et  $***$ --0,10). R : indique répulsion entre les frontières. A : indique association entre les frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au le Tableau 20 et les statistiques au Tableau 23.

Variable	$O_G$	$O_H$	$O_{GH}$
<i>Volume total</i> - 100 groupes	* A		* A
<i>Volume total</i> - 50 groupes	* A	* A	* A

#### 3.2.2. Association entre les frontières du *Volume total* et des autres variables

Cette section présente les résultats de l'évaluation de l'association spatiale des frontières indirectes des 100 groupes de p-e selon leur *Volume total* avec les frontières directes des variables *Type forestier*, *Topo-c* et *Sol*, ainsi que les frontières indirectes des 100 groupes de *Topo-q*. L'hypothèse nulle de cette analyse est la suivante : les frontières coïncident spatialement de façon aléatoire. Deux hypothèses alternatives sont considérées : (1) les frontières indirectes du *Volume total* sont spatialement dépendantes des frontières des autres variables, et (2) les frontières indirectes du *Volume total* sont associées spatialement à celles des autres variables.

Les frontières indirectes du *Volume total* dépendent spatialement des frontières directes des variables *Type forestier* et *Topo-q* car la statistique  $O_S$  indique que le nombre d'EFs inclus

dans les deux ensembles de frontières est élevé (Tableau 26). La géométrie différente des frontières indirectes et directes pourrait expliquer les distances élevées entre elles (indiquées par  $O_G$  et d' $O_{GH}$ ). Les frontières de toutes les variables sont également associées aux frontières indirectes du *Volume total*, car le nombre d'EFs communs aux deux variables est plus élevé que ce qui serait dû au hasard (Tableau 26).

Tableau 26. Signification des probabilités des statistiques de superposition qui évaluent la relation spatiale entre les frontières indirectes du *Volume total* et celles des autres variables ( $\alpha = **0,01$ ,  $**0,05$  et  $***0,10$ ). R : indique répulsion entre les frontières. D : indique dépendance des frontières. Les répartitions aléatoires ont été répétées 1000 fois. Les variables sont décrites au Tableau 20 et les statistiques au Tableau 23.

Relation	Variable	$O_S$	$O_G$	$O_H$	$O_{GH}$
Dépendance	<i>Type forestier</i>	* D	* R		*** R
	<i>Topo-q</i>	* D		* D	* D
Association	<i>Type forestier</i>	* A	*** R		
	<i>Topo-q</i>	* A			
	<i>Topo-c</i>	* A			
	<i>Sol</i>	* A	*** R		

Ces résultats indiquent qu'à l'échelle spatiale déterminée par l'espacement des p-e, les frontières des groupes de p-e homogènes pour le volume ligneux sont associées aux frontières produites par un changement marqué du type de végétation forestière ainsi que de la topographie et du sol. Cette information peut servir à établir un lien avec les frontières de fort contraste à l'échelle régionale de la carte forestière, lesquelles signalent par une autre méthodologie des changements des conditions forestières des groupes de peuplements.

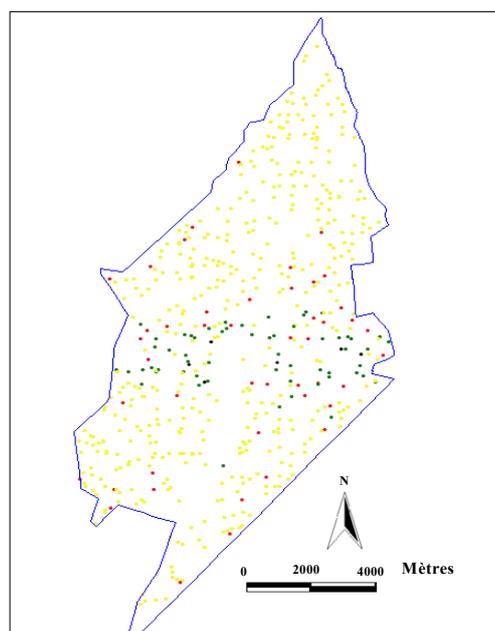
#### 4. Structure spatiale locale des variables

L'objectif de cette section est de déterminer la structure spatiale locale des frontières directes des variables *Volume total* et *Type forestier*, afin d'identifier les groupes de frontières particulières en terme de leur intensité. Le *Volume total* est la variable d'intérêt fondamental dans ce chapitre, mais la recherche d'une structure spatiale locale concerne également les valeurs du *Volume total* des p-e (et non seulement des frontières). Les résultats sont présentés par variable aux prochains paragraphes.

## 4.1. Volume total

### 4.1.1. Structure spatiale locale du *Volume total*

Les valeurs significatives d'autocorrélation spatiale locale du *Volume total* à partir des p-e sont réparties principalement sur une bande qui traverse le centre de la Forêt Montmorency d'est en ouest (Figure 22). Les valeurs significatives d'autocorrélation spatiale positive, localisées vers le centre de cette bande, correspondent aux p-e où le *Volume total* est faible à la suite de coupes forestières récentes. Les valeurs significatives d'autocorrélation spatiale négative correspondent, quant à elles, aux p-e où le *Volume total* est élevé, soit des situations de forêts matures. En fonction de ces résultats, on peut considérer que les aires où le volume de matière ligneuse est particulièrement faible à cause des coupes sont prédominantes dans la zone centrale et que parmi elles se trouvent des parcelles de forêt mature où le volume ligneux est particulièrement élevé.



Jaune : les p-e dont les valeurs d'*I* de Moran local sont non significatives.

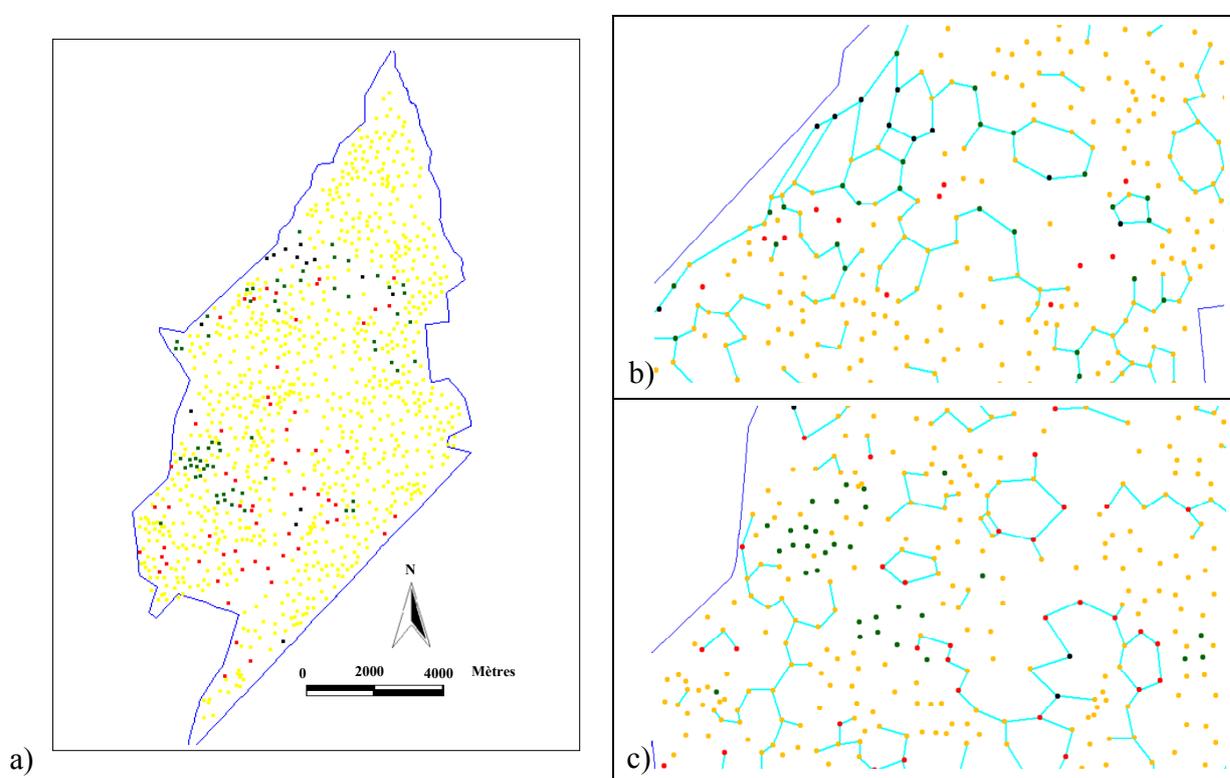
Vert : p-e dont les valeurs d'*I* de Moran local sont significativement positives.

Rouge : p-e dont les valeurs d'*I* de Moran local sont significativement négatives.

Figure 22. Structure spatiale locale du *Volume total* ( $n=539$  p-e).

#### 4.1.2. Structure spatiale locale des frontières du *Volume total*

Les valeurs significatives d'autocorrélation spatiale locale positive et négative sont réparties en deux secteurs (Figure 23a). Dans le secteur nord, les valeurs significatives et positives correspondent aux EFs, alors que les négatives correspondent aux CEFs inférieures au seuil des frontières (Figure 23b). Dans le secteur sud, les valeurs significatives et négatives correspondent, au contraire, aux EFs, alors que les positives correspondent aux CEFs (Figure 23c). Ces résultats permettent de dire que le secteur nord présente des discontinuités abruptes du volume ligneux, alors que le secteur sud est plus homogène et que les frontières se détachent de leur voisinage.



Jaune ou orange : frontières directes dont les valeurs d' $I$  de Moran local sont non significatives.  
 Vert : frontières directes dont les valeurs d' $I$  de Moran local sont significativement positives.  
 Rouge : frontières directes dont les valeurs d' $I$  de Moran local sont significativement négatives.  
 Lignes : frontières directes.

Figure 23. Structure spatiale locale des frontières directes du *Volume total* : a) Répartition spatiale des valeurs d' $I$  de Moran local. b) Détail du secteur central, c) détail du secteur sud.

#### 4.2. *Type forestier*

Cette variable présente des valeurs significatives d' $I$  de Moran local, autant négatives que positives, concentrées dans les secteurs central et sud-est (Figure 24). Le secteur central présente une prédominance de discontinuités marquées, tandis que dans le secteur sud les discontinuités peu marquées prédominent.

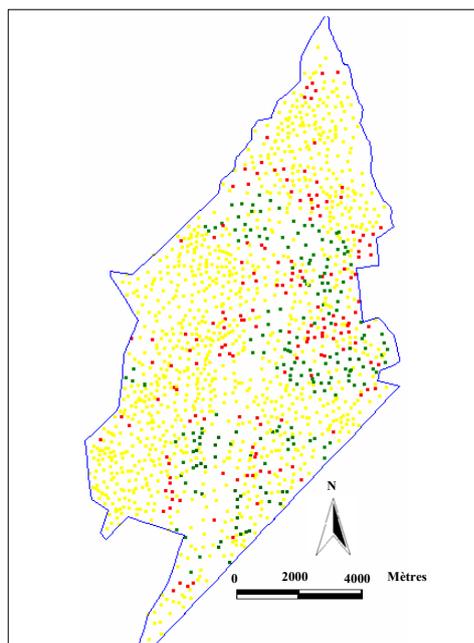


Figure 24. Structure spatiale locale des frontières directes du *Type forestier*.

Les résultats présentés permettent d'interpréter le secteur central comme une zone où prédominent les discontinuités marquées de la structure de la végétation forestière et de la production ligneuse. Il existe également une concentration de faibles valeurs du volume de matière ligneuse qui coïncident spatialement avec les frontières de la structure de la végétation (Figure 22, Figure 23 et Figure 24). Le secteur sud est, pour sa part, plus homogène en ce qui a trait à la structure de la végétation forestière et aux conditions productives.

## 5. Relation entre les frontières du Volume total et les variables forestières et environnementales

Au-delà des associations spatiales entre les frontières du *Volume total* et les frontières des variables forestières et environnementales, l'intérêt porte également sur la détection des tendances dans les combinaisons d'attributs forestiers et environnementaux des p-e associées aux frontières du *Volume total*. Ces tendances ont été détectées en comparant les CEFs (qui sont en dessous du seuil requis pour être considérées comme des EFs) et les EFs, par rapport aux fréquences des combinaisons de catégories de variables qualitatives et aux différences « maximum–minimum » des variables quantitatives.

En ce qui a trait aux variables quantitatives, les distributions de fréquence des différences « maximum–minimum » des EFs et des CEFs des variables topographiques (*Altitude* et *Inclinaison*) et édaphiques (*Épaisseur de la matière organique* et *Pierrosité*), ne sont pas significativement distinctes. Cependant, elles le sont pour les différences des variables dendrométriques : *Densité totale* (KS=0,2963, p=0,000), *Surface terrière* (KS=0,3099, p=0,000) et *Hauteur des arbres dominants* (KS=0,2729, p=0,000).

Les distributions de fréquences des différences « maximum–minimum » pour les variables dendrométrique indiquent que les EFs présentent plus fréquemment des différences élevées que les CEFs (Figure 25). Ainsi, pour la *Densité totale* le seuil entre les EFs et les CEFs peut se situer à 1000 tiges/ha (Figure 25a et b), pour la *Surface terrière* il peut se situer entre 20 et 40 m<sup>2</sup>/ha (Figure 25c et d), tandis que pour la *Hauteur des arbres dominants*, les différences à 8 m permettent de distinguer les EFs des CEFs (Figure 25e et f).

Les variables qualitatives dont les fréquences de combinaisons d'attributs diffèrent entre les EFs et les CEFs sont le *Groupement d'essences*, la *Classe de Hauteur*, la *Classe d'âge* et la *Perturbation d'origine* (Figure 26). Les différences les plus évidentes correspondent à la *Classe de hauteur* (Figure 26b) où les combinaisons de catégories non consécutives sont nettement plus fréquentes dans les EFs que dans les CEFs. Pour le *Groupement d'essences*, les EFs présentent plus fréquemment des combinaisons de végétation de type résineux et

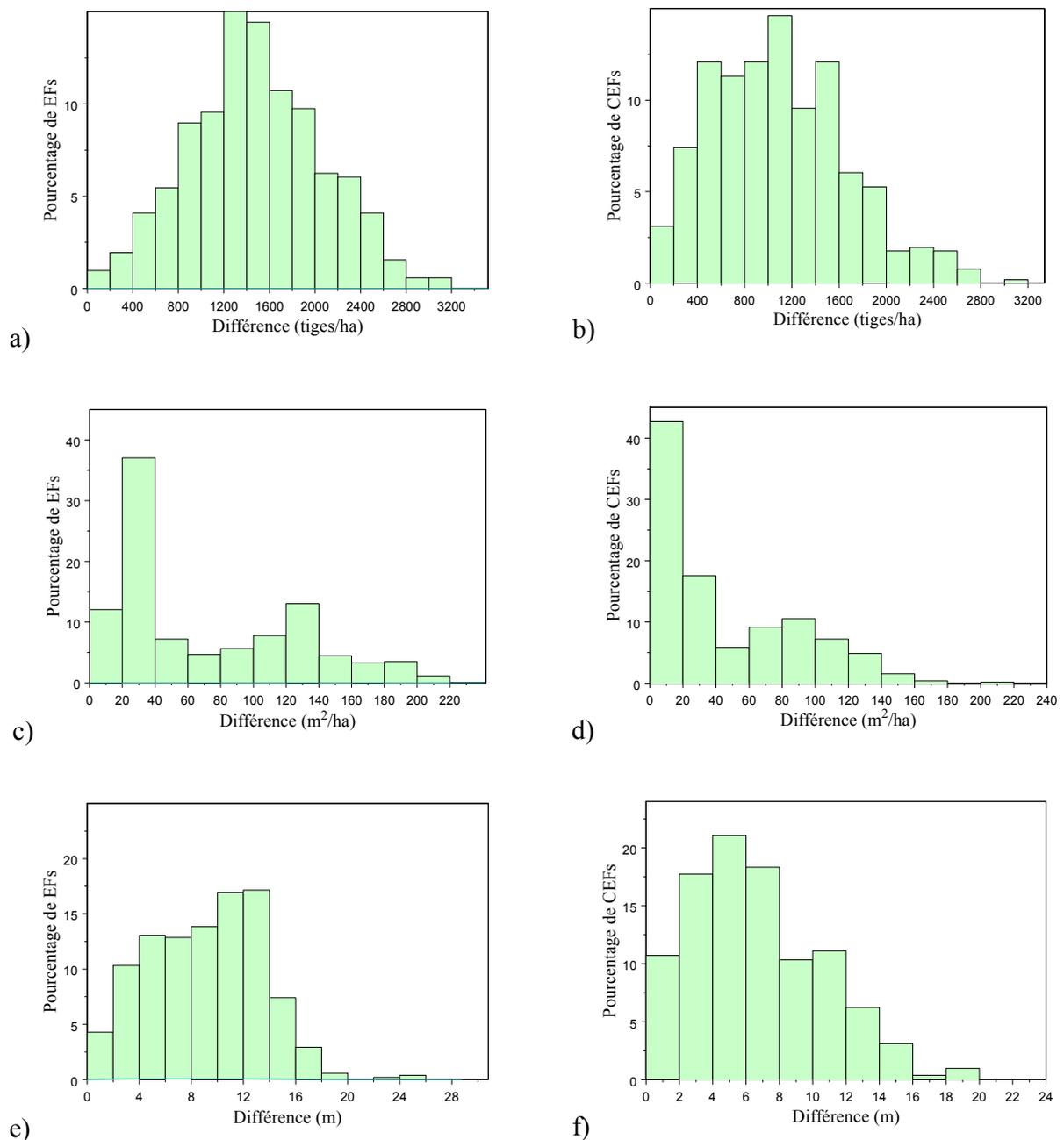


Figure 25. Distribution de fréquences des différences « maximum-minimum » de chaque variable dans les p-e associées aux EFs (graphiques à gauche) et aux CEFs (dont la valeur est inférieure au seuil du percentile 50) (graphiques à droite) du *Volume total*. a) et b) *Densité totale*. c) et d) *Surface terrière*. e) et f) *Hauteur des arbres dominants*.

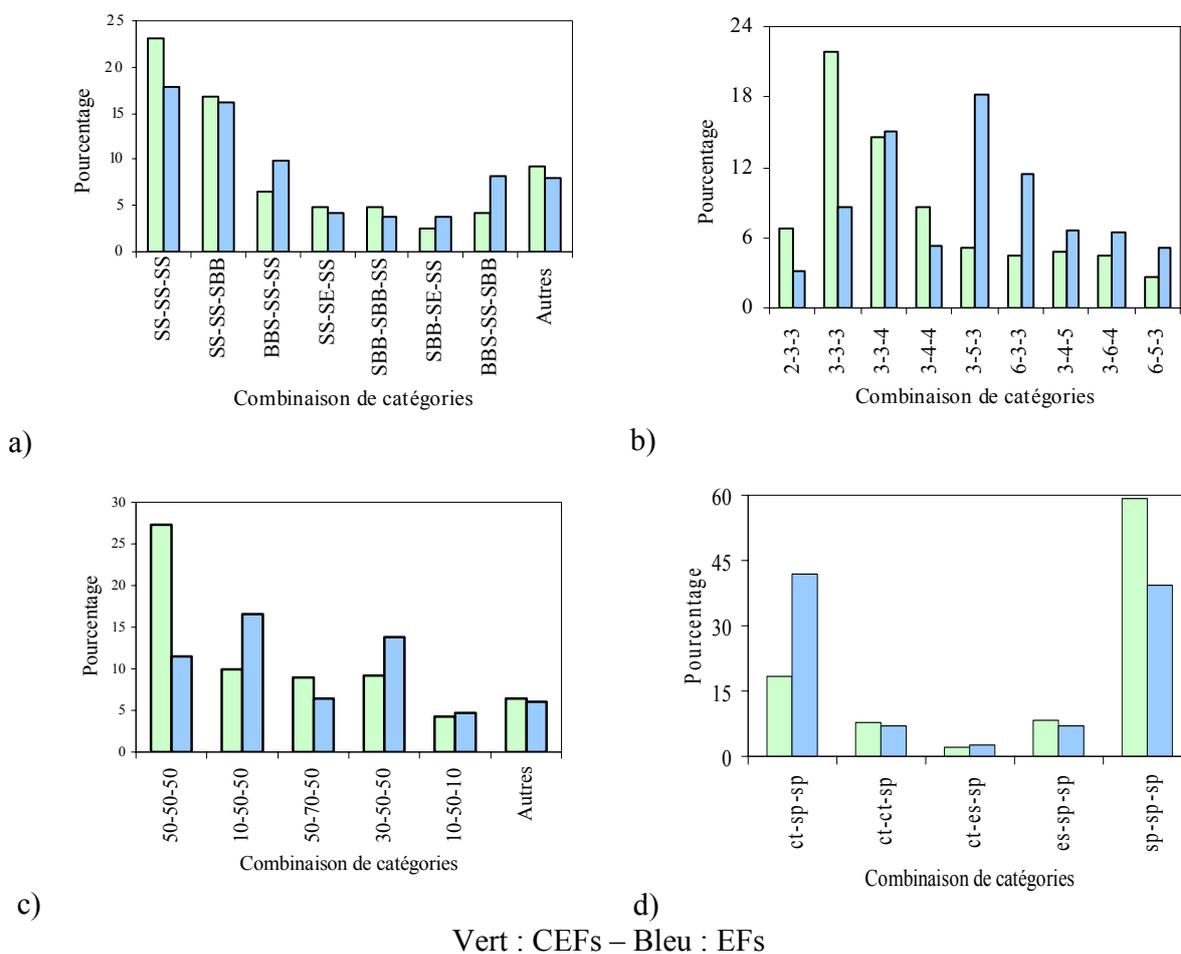


Figure 26. Fréquence des combinaisons de catégories observées dans les p-e associées aux EFs et aux CEFs pour *Volume total*. Combinaison de catégories de : a) *Groupement d'essences*, b) *Classe de hauteur*, c) *Classe d'âge*, d) *Perturbation d'origine*. ct : coupe totale, es : épidémie sévère, sp : sans perturbation.

mélangés à prédominance feuillue, alors que la combinaison de types de végétation résineux domine pour les CEFs. La tendance pour la *Classe d'âge* n'est pas aussi évidente (Figure 26c), bien que les combinaisons d'âges non consécutifs ou d'âges consécutifs, mais de végétation en développement, soient plus fréquentes dans les EFs que dans les CEFs. Finalement, pour la *Perturbation d'origine*, la combinaison d'une p-e affectée par une coupe totale (ct) et de deux autres p-e qui n'ont pas subi de perturbations sévères (sp), par exemple ct-sp-sp, est plus fréquente dans les EFs, tandis que pour les CEFs les p-e qui n'ont pas été affectées par des perturbations sévères dominent (Figure 26d).

Ces résultats fournissent des critères qui permettraient, avec l'information disponible, de mesurer l'incertitude sur l'existence de certaines frontières cartographiques, au moins de façon qualitative. En effet, les relations « non spatiales » entre les frontières du volume ligneux et les variables dendrométriques indiquent que les différences entre la valeur maximale et minimale des trois p-e qui déterminent un EF ou un CEF pourraient prendre une valeur indicatrice de l'existence possible d'une frontière. Cela est dû au fait que l'étendue des différences entre la valeur maximale et minimale est la même pour les deux et que seules les fréquences les différencient. Par exemple, si la différence entre les valeurs de *Densité totale* des trois p-e est de l'ordre de 2400 tiges/ha, il est probable qu'il existe entre elles une discontinuité pour le volume ligneux, car cette valeur est beaucoup plus fréquente dans les EFs que dans les CEFs (Figure 25).

Parmi les variables qualitatives, la *Classe de hauteur* est celle dont les combinaisons se démarquent fortement entre les CEFs et les EFs. Il est possible que, dans les p-e relevées, les conditions de site ne soient pas très différentes et que la hauteur des arbres, même si elle est estimée par classe, soit liée plus directement au volume ligneux qu'à une autre variable comme la *Classe d'âge*. En conclusion, dans la mesure où les caractéristiques structurales et la présence des perturbations déterminent des divergences écologiques entre les p-e, les discontinuités dans le volume ligneux sont évidentes.

## Conclusions

Les frontières directes et indirectes du volume ligneux ont été identifiées par la détermination de groupes de p-e avec des valeurs semblables pour cette variable, ainsi qu'à partir de la détection de discontinuités intenses entre p-e adjacentes. Les deux méthodologies utilisées permettent de mettre en évidence des aspects différents de la répartition spatiale des variables considérées.

La méthode de regroupement des p-e avec des contraintes spatiales, applicable seulement aux variables quantitatives, permet de découper le territoire à l'étude en fonction des données de terrain. Elle peut être appliquée à différentes échelles selon le nombre de

groupes formés. En conséquence, des cartes de la Forêt Montmorency ont été obtenues à partir des données de terrain, qui représentent la répartition spatiale du volume ligneux et des variables forestières et environnementales. Ces cartes, aux deux échelles, permettent d'identifier les zones qui sont tellement distinctes de leur voisinage qu'elles persistent lorsque le grain spatial augmente, c'est-à-dire lorsque le nombre de groupes est réduit.

Quant à la méthode de détection de changements, elle permet d'identifier les discontinuités des variables quantitatives et qualitatives. Ces discontinuités peuvent délimiter des zones sur le territoire, comme c'est le cas de la méthode de regroupements. Cette méthode indique l'emplacement des conditions forestières ou environnementales particulières qui créent ou maintiennent les discontinuités. Bien que la méthode s'avère très utile, elle demeure limitée par rapport aux variables qualitatives, puisque les discontinuités représentent le désaccord entre les catégories de deux p-e et non la dissemblance entre ces catégories.

Pour ce qui est de l'analyse de la structure spatiale locale des frontières, elle aide également à identifier des zones du territoire où les frontières sont concentrées. Cette information complète celle de la localisation des frontières directes et indirectes, d'une part en ajoutant la dimension du voisinage à celle de l'intensité de la frontière et, d'autre part, en ajoutant de l'information pour la qualification du territoire en fonction d'une ou de plusieurs variables.

Pour sa part, l'analyse de recouvrement permet d'identifier les associations spatiales entre les discontinuités des variables à l'échelle fixée par l'espacement entre les p-e. Cette analyse indique que les frontières des zones homogènes pour le volume ligneux coïncident spatialement avec les discontinuités intenses de la végétation forestière. Cette information peut être transférée à la carte forestière dont les frontières de fort contraste écologique sont identifiées à plusieurs échelles.

À la Forêt Montmorency, les deux types de frontières, directes et indirectes, coïncident fréquemment dans l'espace, surtout lorsqu'il s'agit de différences marquées par rapport au voisinage (représentées par des p-e isolées ou par des groupes de p-e). Cette association spatiale entre frontières indirectes et directes est significative du point de vue statistique.

Les frontières directes de toutes les variables sont généralement reliées en formant des chaînes et, lorsqu'elles forment des polygones fermés, ces derniers contiennent habituellement une seule p-e, qui représente des conditions locales particulières. Du point de vue des variables qualitatives composites, près de la moitié des frontières implique un désaccord pour la majorité des variables et les frontières où le désaccord concerne toutes les variables sont très fréquentes (entre 15 et 30 % des frontières). Cela signifie que, du point de vue de la structure de la végétation, des sols et de certains aspects topographiques, les p-e voisines montrent une certaine hétérogénéité.

Quant à la structure spatiale locale du *Volume total*, elle indique des secteurs où les valeurs faibles sont concentrées et associées aux coupes totales. Bien que les différences de volume ligneux ne soient pas suffisamment prononcées pour déterminer des frontières directes significativement différentes de la valeur moyenne, elles coïncident avec les frontières des variables associées à la structure de la végétation comme *Type forestier*. Ce fait indique que ces valeurs du *Volume total* seraient le produit d'autres facteurs comme les perturbations. Par ailleurs, l'analyse des distributions statistiques des différences des valeurs maximales et minimales des variables topographiques et édaphiques des trois p-e qui définissent les EFs et les CEFs ne permet pas de les différencier.

Cependant, les frontières du volume ligneux sont associées aux aspects structuraux de la végétation. En effet, la coïncidence spatiale des frontières indirectes du *Volume total* avec celles de *Type forestier*, confirmée par les résultats de l'analyse de recouvrement, est complétée par les résultats des analyses « non spatiales » des frontières directes. Les différences de valeurs des variables dendrométriques et des combinaisons d'attributs des variables forestières des trois p-e qui définissent les CEFs renforcent le lien qui existe entre le volume ligneux et les aspects structuraux de la végétation. Les frontières du *Volume total* seraient donc associées à des différences dans la *Classe de hauteur* et la *Classe d'âge*, ainsi qu'à la présence d'une perturbation sévère dans une des p-e. Les variables dendrométriques telles que la *Hauteur des arbres* et la *Densité* dépendent également des aspects structuraux. Par conséquent, les frontières du volume ligneux seraient associées à des p-e qui

correspondent à différentes étapes du développement d'un peuplement à la suite d'une perturbation sévère.

Les p-e fournissent des informations intéressantes par rapport au patron spatial du volume ligneux à une échelle déterminée par l'espacement entre les p-e. Les discontinuités de la structure de la végétation et d'un ensemble de variables environnementales et dendrométriques sont mises en évidence par les p-e disponibles. Elles rendent également possible l'identification de certains groupes de p-e dont le volume ligneux est suffisamment différent de celui des p-e voisines pour que ces groupes persistent lors des changements d'échelle.

## **Résumé du chapitre**

Ce chapitre a permis d'examiner les frontières définies à partir des données de terrain, afin d'être en mesure d'identifier les zones homogènes par rapport au volume ligneux, ainsi que les discontinuités spatiales qui y sont liées. Les différentes méthodes utilisées s'avèrent complémentaires pour extraire de l'information sur différents aspects de la structure spatiale des données de terrain. Ces informations permettent d'élaborer une représentation spatiale de la structure du volume ligneux à partir des données de terrain, laquelle pourra être comparée et éventuellement fusionnée à la carte forestière qui représente la même variable à partir des informations provenant de la photo-interprétation.

Les estimations du volume de matière ligneuse issues des données de terrain sont généralement regroupées afin d'obtenir des estimations du volume et de la variance pour différents groupes de p-e. L'exploration des critères qui permettent de guider ce regroupement afin d'obtenir des estimations du volume avec une faible variabilité constitue le sujet du chapitre suivant.

# **Chapitre V - Les estimations du volume de matière ligneuse**

## **Introduction**

Les inventaires forestiers réalisés par le Ministère de Ressources naturelles du Québec visent premièrement à produire une connaissance actualisée de la ressource forestière, notamment en vue de l'allouer aux divers intervenants (Coulombe 1996, Coulombe et Lowell 1995). La précision des estimations du volume de matière ligneuse obtenues par ces inventaires est adéquate à l'échelle du territoire à aménager. Toutefois, elles sont très souvent utilisées pour la planification opérationnelle des interventions forestières. Cette activité requiert un niveau de détail supérieur à celui de l'aménagement et nécessite, par conséquent, des estimations plus précises. La stratégie adoptée pour satisfaire à cette exigence consiste, dans la plupart des cas, à intensifier l'échantillonnage sur le terrain à l'aide d'inventaires d'exploitation. Cependant, même en quadruplant, par exemple, la densité des échantillons, l'amélioration de la précision des estimations demeure modeste (Beaulieu 1994, Beaulieu et Lowell 1994). Évidemment, d'autres approches doivent être explorées pour améliorer véritablement cette précision.

Une des approches explorées concerne la relation entre les variables dendrométriques (comme la surface terrière ou le volume de matière ligneuse) et les variables environnementales issues du terrain et de la cartographie, à l'aide de modèles de régression (Coulombe 1996, Coulombe et Lowell 1995). Les résultats obtenus indiquent que les modèles linéaires les plus performants sont basés sur des variables de terrain et ne peuvent donc pas être appliqués à la totalité de la zone d'étude, comme c'est le cas des modèles basés sur des variables cartographiques (Coulombe et Lowell 1995).

Une approche alternative est liée au processus d'élaboration des estimations dendrométriques. Au Québec, les critères de regroupement des échantillons de terrain (p-e) qui permettent d'estimer le volume de matière ligneuse des strates forestières sont variés.

Les p-e sont regroupées sur la base de leur ressemblance par rapport au type de végétation ligneuse, à la composition d'essences, aux caractéristiques structurales de la végétation et aux perturbations (Bérard et Côté 1996). Malgré ce fait, selon les objectifs de l'inventaire forestier effectué, d'autres critères peuvent être appliqués (Beaulieu 1994). Les strates forestières, dont le nombre de p-e est très bas ou nul, sont regroupées aux autres strates selon des critères de ressemblance.

Dans cette approche, la relation entre les variables dendrométriques et la configuration spatiale des strates forestières cartographiques a été examinée dans le but d'incorporer des critères spatiaux pour le regroupement des strates (Beaulieu 1994, Beaulieu et Lowell 1994). Les résultats obtenus indiquent qu'une autocorrélation spatiale dans les strates forestières, même faible, complique le regroupement des strates sur la base de la contiguïté spatiale.

Jusqu'à maintenant, les critères utilisés pour le regroupement des p-e sont fixés « *a priori* », c'est-à-dire qu'ils ne tiennent pas compte de la structure inhérente des données écologiques des p-e auxquelles ils seront appliqués. Selon une perspective différente, l'identification des critères qui découlent de la structure interne des données est plus appropriée pour définir des groupes de p-e homogènes en terme du volume de matière ligneuse.

Parmi les méthodes disponibles pour la formation des groupes homogènes d'observations, celles de classification ou de regroupement sont probablement les plus employées dans divers domaines. D'une part, de nombreux algorithmes regroupent les observations sur la base de leur ressemblance par rapport à une ou plusieurs variables (Legendre et Legendre 1998). Ces méthodes permettent ainsi de décrire et d'interpréter la structure des variables examinées. D'autre part, la régression et classification arborescente est une méthode exploratoire qui regroupe également les observations en groupes homogènes et son utilisation est de plus en plus répandue en écologie et en foresterie pour la classification, la prédiction et la cartographie (Cairns 2001, De'Ath et Fabricius 2000, Nigh et Love 2004, Rejwan *et al.* 1999, Rolling *et al.* 2004, Segal 1988). La régression arborescente permet de modéliser une variable quantitative à partir de variables explicatives de nature qualitative et/ou quantitative (Breiman *et al.* 1984, Clark et Pregibon 1992, Venables et Ripley 1999).

L'avantage de cette méthode, comparativement à celles de regroupement, est que les groupes formés sont homogènes en terme de la variable réponse, tandis que les variables explicatives sont utilisées pour la formation des groupes. De plus, comparativement aux modèles classiques de régression, la régression arborescente est plus adéquate pour saisir les contributions non additives des variables, qui n'exigent pas d'être explicitées sous une forme particulière (Clark et Pregibon 1992, Venables et Ripley 1999).

Le but de ce chapitre est d'explorer comment la combinaison de variables, issues de la cartographie et du terrain, permet de diviser l'ensemble des p-e d'un inventaire forestier en groupes selon des seuils de variance intergroupe élevés pour le volume de matière ligneuse. Les groupes de p-e formés de cette façon permettront d'obtenir des estimations du volume où la variabilité intragroupe est inférieure à celle des estimations des groupes issus de la méthode actuellement employée au Québec.

L'hypothèse de base est la suivante : en raison de la structure interne des données des p-e, certaines combinaisons de variables forestières et environnementales permettent leur regroupement et minimisent ainsi l'hétérogénéité interne des groupes en termes du volume de matière ligneuse.

Les objectifs spécifiques pour ce chapitre sont :

1. D'explorer, à l'aide de la régression arborescente, diverses combinaisons de critères en vue de regrouper les p-e selon leur ressemblance en terme du volume de matière ligneuse;
2. De comparer les groupes obtenus par la régression arborescente avec ceux issus de la méthode actuellement utilisée au Québec en termes de :
  - a. la variabilité intragroupe des estimations du volume;
  - b. la structure spatiale des estimations du volume.

## Méthodologie

### 1. Le regroupement des p-e selon la méthode actuelle au Québec

L'analyse de l'information des placettes-échantillons (p-e) temporaires de 1992 et de la carte forestière de 1992 de la Forêt Montmorency, fournies par ses gestionnaires, permet de déduire les étapes de la méthodologie qui a été employée pour (1) regrouper les p-e afin d'obtenir les estimations sur le *Volume total* moyen et (2) assigner ces estimations aux strates forestières de la carte de 1992.

L'inventaire forestier de 1992 a été planifié en utilisant la méthode d'échantillonnage aléatoire stratifiée, en assignant les échantillons en fonction de la surface des strates forestières (Bérard et Côté 1996, Coulombe 1996). Les strates de la carte sont prises en compte pour fixer le nombre d'échantillons requis sur le terrain.

Les 338 p-e de l'inventaire forestier ont été regroupées en 52 groupes par les gestionnaires de la Forêt Montmorency, en fonction des strates forestières observées sur le terrain, même si les strates cartographiques correspondantes sont différentes. La strate forestière de terrain, tout comme sa contrepartie cartographique, est dérivée de la combinaison, entre autres, des variables *Groupe d'essences*, *Classe de densité*, *Classe de hauteur*, *Classe d'âge*, *Perturbation d'origine* et *Perturbation moyenne*. Dans le cas des strates observées sur le terrain, toutes ces variables sont estimées à partir de l'observation des caractéristiques de la végétation, sans tenir compte d'une surface minimale d'observation, et normalement sur le point central de la p-e. Dans le cas des strates cartographiques, les variables sont estimées lors de la photo-interprétation, en prenant une superficie minimale de 8 ha (parfois 4 ha, selon le type de forêt) (Ministère des Ressources naturelles du Québec 1995).

Au Québec, les p-e sont généralement regroupées selon des critères comme la composition d'essences, la perturbation d'origine, la perturbation moyenne, la densité, la hauteur ou l'âge (Beaulieu 1994, Bérard et Côté 1996, Coulombe 1996). Des 52 groupes de p-e définis pour la Forêt Montmorency dans l'inventaire de 1992, 6 groupes contiennent une seule p-e. Les autres contiennent entre 2 et 29 p-e chacun. Pour chaque groupe les valeurs minimale,

maximale et moyenne, l'écart type et l'erreur type du *Volume total* des arbres de diamètre égal ou supérieur à 10 cm sont estimés à partir de l'information des p-e. Chaque groupe qui contient une seule p-e utilise la valeur du *Volume total* comme valeur moyenne.

Au Québec, une fois les p-e regroupées l'étape suivante consiste à attribuer les valeurs moyennes des groupes aux strates cartographiques pertinentes. Un des critères qui guident cette assignation, au-delà de l'appellation de la strate forestière, qui est composée de la combinaison des variables citées précédemment, semble être le stade de développement de la végétation ligneuse. Étant donné que certaines strates cartographiques ne sont pas échantillonnées, ces dernières sont associées aux strates les plus semblables (Bérard et Côté 1996).

## **2. Données et variables**

Les données utilisées dans ce chapitre sont les 338 placettes-échantillons (p-e) temporaires de 1992, les 183 p-e permanentes mesurées entre 1995 et 1999 pour lesquelles l'information des variables sélectionnées est disponible, ainsi que la carte forestière de 1992.

Les estimations du *Volume total* moyen des groupes formés par les régressions arborescentes subséquentes sont comparées avec les estimations du *Volume total* moyen effectuées selon la méthode décrite à la section précédente. Les estimations de la moyenne et de l'écart type du *Volume total* des arbres de plus de 10 cm de diamètre pour chacun des groupes ont été fournies par la Faculté de Foresterie et Géomatique (Jean Bégin, comm. pers.).

Les caractéristiques générales des p-e et de la carte forestière sont détaillées au Chapitre II et à l'Annexe III.

La variable dépendante des modèles de régression arborescente est le *Volume total* (m<sup>3</sup>/ha) estimé pour chacune des 338 p-e. Les variables indépendantes sont issues de deux sources : (1) le terrain, dont les données proviennent des p-e et (2) la cartographie, dont les données

correspondent au peuplement de la carte forestière de 1992 qui contient chaque p-e (Tableau 27). Les détails des variables qualitatives sont présentés au Chapitre II et à l'Annexe III. Les variables cartographiques de chaque p-e sont extraites de la carte forestière de 1992 avec l'extension StatMod de ArcView (Garrard 2002).

Tableau 27. Énumération des variables de terrain et cartographiques utilisées pour les modèles de régression arborescente.

Variable de terrain		Variable cartographique		Nature
Description	Symbole	Description	Symbole	
<i>Groupe ment d'essences</i>	<i>GET</i>	<i>Groupe ment d'essences</i>	<i>GE</i>	Nominale
<i>Classe de densité</i>	<i>CDT</i>	<i>Classe de densité</i>	<i>CD</i>	Ordinale
<i>Classe de hauteur</i>	<i>CHT</i>	<i>Classe de hauteur</i>	<i>CH</i>	Ordinale
<i>Classe d'âge</i>	<i>CAT</i>	<i>Classe d'âge</i>	<i>CA</i>	Ordinale
<i>Classe de pente</i>	<i>CPT</i>	<i>Classe de pente</i>	<i>CP</i>	Ordinale
<i>Perturbation d'origine</i>	<i>OrigT</i>	<i>Perturbation d'origine</i>	<i>Orig</i>	Nominale
<i>Perturbation moyenne</i>	<i>PertT</i>	<i>Perturbation moyenne</i>	<i>Pert</i>	Nominale
<i>Coordonnée X de la p-e</i>	<i>X</i>	<i>Coordonnée X de la p-e</i>	<i>X</i>	Quantitative
<i>Coordonnée Y de la p-e</i>	<i>Y</i>	<i>Coordonnée Y de la p-e</i>	<i>Y</i>	Quantitative
<i>Hauteur des arbres dominants</i>	<i>HD</i>			Quantitative
<i>Exposition</i>	<i>Exp</i>			Quantitative
<i>Inclinaison</i>	<i>Inc</i>			Quantitative
<i>Versant</i>	<i>Vrs</i>			Nominale
<i>Situation sur la pente</i>	<i>SP</i>			Nominale
<i>Forme de la pente</i>	<i>FP</i>			Nominale
<i>Type d'humus</i>	<i>TH</i>			Nominale
<i>Texture de l'horizon B</i>	<i>THB</i>			Nominale
<i>Texture de l'horizon C</i>	<i>THC</i>			Nominale
<i>Pierrosité</i>	<i>Pier</i>			Quantitative
<i>Dépôt de surface</i>	<i>DpS</i>			Nominale
<i>Drainage</i>	<i>Dr</i>			Nominale
<i>Épaisseur de la matière organique</i>	<i>EMO</i>			Quantitative

Parmi les p-e utilisées pour la construction et validation des modèles de régression arborescente, certaines comportent des données manquantes pour les variables qualitatives. Ces p-e sont utilisées en assignant aux données manquantes, une nouvelle catégorie SD (sans donnée). Les observations avec des données manquantes pour les variables quantitatives ne sont pas utilisées. Des 338 p-e utilisées pour la définition des modèles, 59

présentent certaines données manquantes pour les variables qualitatives alors que 38 les présentent parmi les 183 p-e utilisées pour la validation des modèles (Tableau 28).

Tableau 28. Nombre de p-e avec des données manquantes des ensembles utilisés pour la définition (338 p-e) et la validation (183 p-e) des modèles de régression arborescente.

Variable	Ensemble de 338 p-e	Ensemble de 183 p-e
Terrain		
<i>Groupe ment d'essences (GET)</i>	2	2
<i>Classe de densité (CDT)</i>	27	23
<i>Classe de hauteur (CHT)</i>	27	4
<i>Classe d'âge (CAT)</i>	3	
Cartographique		
<i>Groupe ment d'essences (GE)</i>	22	25
<i>Classe de densité (CD)</i>	16	3
<i>Classe de hauteur (CH)</i>	16	3

### 3. Méthodes

La Figure 27 illustre les étapes méthodologiques détaillées aux sections subséquentes.

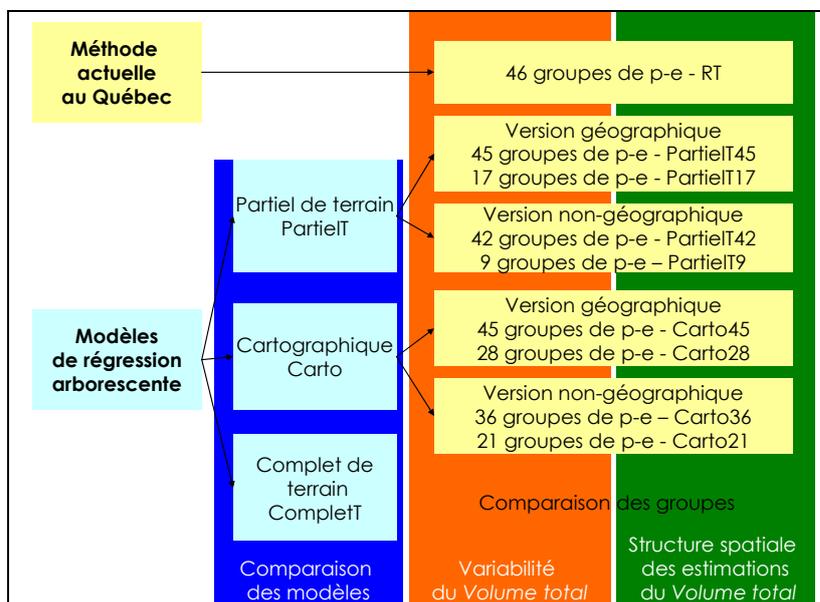


Figure 27. Étapes méthodologiques du Chapitre V.

### **3.1. La régression arborescente**

#### **3.1.1. Définition des modèles de régression arborescente ou arbres de régression**

La régression arborescente constitue une technique statistique visant à explorer et à modéliser des données complexes (De'Ath et Fabricius 2000). Elle modélise la variation d'une variable quantitative à partir d'une ou de plusieurs variables explicatives (ou indépendantes), qualitatives ou quantitatives, dont les interactions ne sont pas restreintes à une forme particulière (additive, multiplicative ou autre) (Breiman *et al.* 1984, Chambers et Hastie 1992, Clark et Pregibon 1992, De'Ath et Fabricius 2000, Insightful 2001).

Un modèle de régression arborescente ou arbre de régression est donc une collection de règles ou de critères structurés en forme d'arbre binaire. Ces règles sont déterminées par une procédure de partition récursive (Insightful 2001). L'arbre est construit à partir de la division répétée des observations (les p-e dans le cas de cette recherche). Chaque division est définie par une seule règle établie à partir d'une des variables explicatives. À chaque division, les observations sont partitionnées en deux groupes mutuellement exclusifs, chacun d'eux étant le plus homogène possible en terme de la variable réponse (*Volume total* dans cette étude). La procédure de division est appliquée ainsi à chaque groupe de façon séparée (Clark et Pregibon 1992). Des critères d'arrêt, décrits aux sections suivantes, sont établis pour que les divisions ne se poursuivent pas jusqu'à l'obtention d'un arbre où chaque groupe final serait constitué d'une seule observation.

#### **3.1.2. Structure des arbres de régression**

L'ensemble des observations de départ est considéré comme la « racine » de l'arbre et chacun des groupes formés par les divisions subséquentes est considéré comme un « nœud ». La taille d'un arbre est définie par le nombre de nœuds finaux ou « feuilles ». Chaque feuille est caractérisée par la valeur moyenne de la variable réponse, sa taille (nombre d'observations) ainsi que les variables indépendantes qui la définissent (Clark et Pregibon 1992).

### 3.1.3. Procédure de croissance d'un arbre

La façon dont les variables indépendantes sont traitées lors de chaque division dépend de leur nature. Pour les variables qualitatives avec plusieurs catégories, n'importe quelle combinaison de catégories peut être utilisée pour la division et pour  $k$  catégories, il existe  $2^{k-1}-1$  types de division possibles (De'Ath et Fabricius 2000). Pour les variables quantitatives, avec  $k$  valeurs distinctes, il existe  $k-1$  seuils de division possibles. Les valeurs ne sont pas utilisées en tant que telles, mais elles déterminent des rangs. Cet aspect rend les arbres de régression invariants face aux transformations monotones des variables indépendantes (Clark et Pregibon 1992). Parmi l'ensemble des divisions possibles, celle qui maximise l'homogénéité des deux nœuds résultants par rapport à la variable réponse est sélectionnée.

L'homogénéité d'un nœud est définie par l'impureté, une mesure qui prend la valeur de zéro pour les nœuds complètement homogènes (dans lesquels toutes les observations présentent la même valeur pour la variable réponse) et qui augmente à mesure que l'homogénéité diminue. Une mesure de l'impureté d'un nœud est la déviance par rapport au modèle statistique sur lequel la division est basée (De'Ath et Fabricius 2000). La fonction de déviance d'une observation  $i$  dans un nœud est définie comme :

$$D(\mu_i, y_i) = (y_i - \mu_i)^2, \quad \text{V.1}$$

où  $y_i$  est la valeur pour la variable réponse de l'observation  $i$ , et  $\mu$ , estimée par le maximum de vraisemblance et constante pour toutes les observations du nœud, est donnée par la moyenne du nœud (Clark et Pregibon 1992). La déviance d'un nœud est définie comme la somme de la déviance de toutes les observations du nœud :

$$D(\hat{\mu}, y) = \sum D(\hat{\mu}, y_i). \quad \text{V.2}$$

La déviance du nœud est de zéro si tous les  $y_i$  sont égaux (le nœud est pur) et elle augmente à mesure que les  $y_i$  s'éloignent de cet idéal.

La division d'un nœud se fait en comparant sa déviance avec celle des nœuds dits « enfants » candidats, issus de la séparation possible des observations à droite (R) et à gauche (G),

$$D(\mu_R, \mu_G; y) = \sum_R D(\mu_R; y_i) + \sum_G D(\mu_G; y_i). \quad \text{V.3}$$

La division sélectionnée est celle qui maximise le changement de la déviance :

$$\Delta D = D(\mu; y) - D(\mu_G, \mu_R; y). \quad \text{V.4}$$

Cette procédure est similaire à la sélection dite pas à pas des variables des modèles de régression. La différence réside dans le fait que dans les arbres de régression, la sélection se fait à chaque division d'un nœud. Il en résulte la possibilité que certaines des variables indépendantes proposées pour la construction d'un arbre de régression ne soient finalement pas utilisées.

### 3.1.4. Sélection de la taille optimale d'un arbre de régression

À mesure que les nœuds sont divisés selon le critère décrit dans la section précédente, l'homogénéité des nœuds résultants croît par rapport à la variable réponse. Finalement, un arbre peut avoir autant de feuilles qu'il existe d'observations, des critères d'arrêt sont donc nécessaires pour réduire sa taille. Deux critères ont été utilisés dans cette recherche : (1) la taille minimale d'un nœud à diviser est fixée à 10 observations (p-e) et (2) la déviance du nœud est égale ou inférieure à 0,01 % de la déviance de la racine. Ce pourcentage a été fixé en raison de la valeur élevée de déviance de la racine des arbres construits.

Les arbres construits selon ces contraintes sont postérieurement simplifiés par le taillage. Cette méthode consiste à obtenir, pour un arbre de régression donné, une séquence nichée de sous-arbres, lesquels sont établis par l'élimination des divisions les moins importantes. L'importance d'une division est déterminée par la mesure du coût-complexité :

$$Da(T') = D(T') + \alpha * \text{taille}(T'), \quad \text{V.5}$$

où  $D(T')$  est la déviance du sous-arbre  $T'$ ,  $taille(T')$  est le nombre de feuilles de  $T'$  et  $\alpha$  est un paramètre de coût-complexité (Clark et Pregibon 1992, Insightful 2001). Pour un  $\alpha$  particulier, le coût-complexité détermine, entre tous les sous-arbres  $T'$  possibles, celui qui minimise  $D\alpha(T')$ .

La sélection de la taille optimale requiert des estimations précises ou « honnêtes » de la déviance (Breiman *et al.* 1984). Ces estimations sont obtenues en testant un arbre de régression avec des données différentes de celles utilisées pour sa construction. Les nouvelles données peuvent provenir d'un ensemble indépendant de données ou d'une procédure de validation croisée. Cette dernière consiste à diviser l'ensemble des données disponibles en sous-ensembles, dont une partie est utilisée pour élaborer l'arbre de régression tandis que l'autre partie est utilisée pour le tester (Clark et Pregibon 1992). Au-delà de la méthode employée pour obtenir les données qui permettront de tester l'arbre de régression, ces données indépendantes sont utilisées pour élaborer une séquence d'arbres de régression de taille différente. La déviance résiduelle de chacun de ces arbres est estimée et l'arbre qui présente la valeur de déviance résiduelle la plus faible est considéré l'arbre de taille optimale (Clark et Pregibon 1992).

Dans cette étude, les p-e temporaires de l'inventaire forestier de 1992 (338 observations) sont utilisées pour construire les arbres de régression tandis que les p-e permanentes mesurées entre 1995 et 1999 (183 observations) sont utilisées pour les valider.

### **3.1.5. Les modèles de régression arborescente construits**

Trois modèles de régression arborescente sont construits : (1) Cartographique (Carto), (2) Complet de Terrain (CompletT) et (3) Partiel de Terrain (PartielT). Pour chaque modèle, deux versions sont élaborées, l'une considère les variables qui correspondent à la localisation géographique de la p-e (variables  $X$ ,  $Y$ ,  $XT$  et  $YT$ ) et l'autre ne les inclut pas. L'objectif d'inclure les variables géographiques est d'évaluer si elles peuvent jouer un rôle important dans la formation des groupes de p-e, étant donné qu'elles pourraient représenter des conditions locales des p-e qui ne sont pas saisies par les autres variables.

Le modèle Carto est élaboré en utilisant les variables cartographiques comme variables indépendantes. Les deux versions de Carto sont donc :

Version géographique :  $Volume\ total \sim f(AC, GE, CD, CH, CA, Orig, Pert, CP, SD, X, Y)$ ,

Version non-géographique :  $Volume\ total \sim f(AC, GE, CD, CH, CA, Orig, Pert, CP, SD)$ ,

où les codes correspondent aux variables énumérées au Tableau 27.

Quant à lui, le modèle CompletT est élaboré à partir des 23 variables de terrain disponibles pour les p-e. Les deux versions de CompletT sont :

Version géographique :  $Volume\ total \sim f(GET, CDT, CHT, CAT, OrigT, PertT, HD, CPT, Alt, Exp, Inc, FP, SP, Vrs, THB, THC, DpS, Dr, TH, EPO, Pier, X, Y)$ ,

Version non-géographique :  $Volume\ total \sim f(GET, CDT, CHT, CAT, OrigT, PertT, HD, CPT, Alt, Exp, Inc, FP, SP, Vrs, THB, THC, DpS, Dr, TH, EPO, Pier)$ ,

où les codes correspondent aux variables énumérées au Tableau 27.

Cet arbre n'a pas d'expression cartographique puisque les variables topographiques et pédologiques n'ont pas de variables qui leur correspondent dans la carte forestière. Pour cette raison, un deuxième modèle, PartielT, est élaboré avec les neuf variables de terrain qui possèdent une correspondance avec des variables cartographiques. Les deux versions de PartielT sont :

Version géographique :  $Volume\ total \sim f(GET, CDT, CHT, CAT, OrigT, PertT, CPT, XT, YT)$ ,

Version non-géographique :  $Volume\ total \sim f(GET, CDT, CHT, CAT, OrigT, PertT, CPT)$ ,

où les codes correspondent aux variables énumérées au Tableau 27.

Les arbres qui sont construits selon des conditions d'arrêt spécifiées par rapport au nombre d'observations (p-e) dans le nœud et selon la proportion de déviance du nœud par rapport à celle de la racine, sont considérés comme des arbres « généraux » : CartoG, CompletTG et PartielTG. Les arbres CartoG et PartielTG sont taillés à :

1. 45 feuilles (Carto45 et PartielT45), afin de rendre la variabilité du *Volume total* des feuilles comparables à celle des groupes obtenus par la méthode de regroupement actuelle. Ceci s'applique seulement aux versions géographiques des modèles car les versions non-géographiques ont une taille générale inférieure, soit de 36 feuilles pour Carto (Carto36) et de 42 feuilles pour PartielT (PartielT42);
2. la taille optimale, celle pour laquelle la déviance est minimale lors de la validation de chacun des arbres de régression avec des données différentes de celles utilisées pour sa construction. Comme il sera présenté plus tard, il s'avère que la taille optimale pour le modèle Carto dans sa version géographique est de 28 feuilles et dans sa version non-géographique est de 21 feuilles. Ils sont subséquemment référés comme Carto28 et Carto21 respectivement. Pour la version géographique du modèle PartielT la taille optimale est 17 feuilles et pour la version non-géographique cette taille est de 9 feuilles; subséquemment on les réfère comme PartielT17 et PartielT9 respectivement.

### **3.2. Comparaison des groupes formés par les différentes méthodes en terme de la variabilité des estimations du *Volume total***

L'objectif de cette analyse est de comparer les groupes obtenus par la méthode actuelle (groupes RT) avec les groupes obtenus à partir des régressions arborescentes (groupes Carto45, PartielT45, Carto28 et PartielT17) en terme de la variabilité des estimations du *Volume total*. La variabilité des estimations sur le *Volume total* est estimée au moyen du Coefficient de variation (CV) de chaque groupe :

$$CV = ( \text{Écart-type du } \textit{Volume total} / \text{Moyenne du } \textit{Volume total} ) * 100. \quad (\text{V.6})$$

Les distributions de valeurs de CVs des groupes issus de chaque méthode (RT, Carto et PartielT) sont comparées à partir des valeurs des percentiles 0 à 100.

Les groupes (RT, Carto45 et PartielT45) dont le CV est inférieur au 20<sup>e</sup> percentile sont considérés dans cette étude comme des groupes de faible hétérogénéité, alors que ceux dont le CV est supérieur au 80<sup>e</sup> percentile sont considérés comme de forte hétérogénéité. Afin de déterminer si ces groupes de variabilité « extrême » présentent des caractéristiques forestières particulières, les distributions de fréquence des variables *GE*, *CD*, *CH*, *CA*, *CP*, *Orig* et *Pert* (décrites au tableau 1) sont estimées et comparées pour ces deux groupes de variabilité. La comparaison statistique entre ces distributions de fréquence et la distribution de fréquence de l'ensemble des p-e est évaluée au moyen du  $\chi^2$  calculé sur le nombre de p-e de chaque catégorie. Pour atteindre le nombre minimal de p-e dans les cellules de l'ensemble des p-e (la fréquence attendue) et rendre l'analyse valide certaines catégories sont agrégées.

### 3.3. Comparaison des estimations du *Volume total* issues des différentes méthodes avec celles des p-e

Cette analyse est développée afin de vérifier si la variabilité des estimations du *Volume total* produites par les différentes méthodes diffère à l'échelle locale. Ainsi le rapport entre les estimations produites par les groupes RT, PartielT et Carto pour les peuplements de la carte forestière de 1992 et le *Volume total* des p-e est calculé selon la formule suivante :

$$Rapport_{Carte/p-e} = \left( \frac{VTmp}{VTp} \right) * 100, \quad \text{V.7}$$

où *VTmp* est le *Volume total* estimé par la méthode *m* pour le peuplement *p* et *VTp* est le *Volume total* consigné pour la p-e localisé dans le peuplement *p*. Un rapport égal à 100 indique l'égalité de valeurs entre la carte et la p-e tandis qu'un rapport supérieur ou inférieur à 100 indique une surestimation ou une sous-estimation du *Volume total* de la p-e respectivement.

### **3.4. Comparaison des groupes issus des différentes méthodes en terme de la structure spatiale des estimations du *Volume total***

Par définition, un arbre de régression est une collection de règles ou de critères structurés en forme d'arbre binaire où chaque critère concerne une variable explicative. Par conséquent, chacune des feuilles d'un arbre de régression peut être définie par les critères employés pour effectuer les divisions successives, en commençant par la racine (l'ensemble initial d'observations) jusqu'à la feuille.

Dans ce chapitre, les p-e sont utilisées pour explorer trois collections de critères afin de trouver la ou les collections qui permettent d'obtenir des groupes de p-e plus homogènes pour le *Volume total* que les groupes définis par la méthode actuelle (groupes RT). Cependant, les estimations du *Volume total* de ces groupes doivent être associées aux strates de la carte forestière comme le sont les estimations du *Volume total* des groupes RT. Donc, les collections de règles de Carto et PartielT sont appliquées aux peuplements de la carte forestière de 1992 pour les classer dans les feuilles correspondantes.

Dans le cas des modèles Carto, les critères dérivés de l'arbre de régression sont utilisés de façon directe pour classer les peuplements forestiers. Cette opération est possible parce que les critères de Carto concernent des variables cartographiques. Au contraire, pour les modèles PartielT, il est d'abord nécessaire d'établir une correspondance entre les variables de terrain utilisées (qui constituent les critères pour classer les peuplements) et les variables cartographiques qui caractérisent les peuplements. Les correspondances entre les deux types de variables sont présentées au Tableau 29.

Les cartes résultantes de Carto et PartielT sont comparées avec la carte des groupes issus de la méthode actuelle en terme de ses caractéristiques et selon la structure spatiale des estimations de chaque carte.

Tableau 29. Correspondance entre les variables de terrain et les variables cartographiques pour le transfert des règles des modèles de régression arborescente aux peuplements de la carte forestière de 1992.

Variable de terrain		Variable cartographique		
Description	Symbole	Description	Symbole	
<i>Groupe ment d'essences</i>	<i>GET</i>	→	<i>Groupe ment d'essences</i>	<i>GE</i>
<i>Classe de densité</i>	<i>CDT</i>	→	<i>Classe de densité</i>	<i>CD</i>
<i>Classe de hauteur</i>	<i>CHT</i>	→	<i>Classe de hauteur</i>	<i>CH</i>
<i>Classe d'âge</i>	<i>CAT</i>	→	<i>Classe d'âge</i>	<i>CA</i>
<i>Classe de pente</i>	<i>CPT</i>	→	<i>Classe de pente</i>	<i>CP</i>
<i>Perturbation d'origine</i>	<i>OrigT</i>	→	<i>Perturbation d'origine</i>	<i>Orig</i>
<i>Perturbation moyenne</i>	<i>PertT</i>	→	<i>Perturbation moyenne</i>	<i>Pert</i>
<i>Coordonnée X de la p-e</i>	<i>X</i>	→	<i>Coordonnée X du centroïde du polygone où se situe la p-e</i>	<i>XC</i>
<i>Coordonnée Y de la p-e</i>	<i>Y</i>	→	<i>Coordonnée Y du centroïde du polygone où se situe la p-e</i>	<i>YC</i>

### 3.4.1. Comparaison des cartes

Les cartes qui représentent spatialement les groupes RT, Carto et PartielT contiennent des parcelles qui correspondent à des ensembles de peuplements adjacents dont l'estimation du *Volume total* est la même. Les cartes sont comparées à partir des caractéristiques de ces parcelles, sans tenir compte de leur position :

1. Le nombre de parcelles (polygones de chaque carte);
2. Le pourcentage de parcelles avec une estimation du *Volume total*, parce que les peuplements dont les catégories qui ne correspondent pas aux règles dérivées des arbres de régression n'ont pas d'estimation du *Volume total*;
3. La densité des parcelles (le nombre de parcelles dans 100 ha);
4. La superficie moyenne des parcelles;
5. La superficie de la plus grande parcelle.

### 3.4.2. Comparaison de la structure spatiale des estimations du *Volume total*

Un autre aspect pertinent concerne la structure spatiale des estimations du *Volume total* des cartes RT, PartielT et Carto. Il permet de déterminer l'existence de différences par rapport à la variabilité spatiale du *Volume total*. Cette variabilité spatiale est mesurée à partir de la semivariance des valeurs du *Volume total* à des distances croissantes. La semivariance mesure la différence moyenne de ces valeurs entre des paires de points localisés à une distance déterminée (Legendre et Legendre 1998). Une semivariance faible signifie que les paires de points considérés présentent des valeurs du *Volume total* similaires, alors qu'une semivariance élevée indique que ces valeurs sont très différentes.

Trois éléments permettent de définir et de comparer les variogrammes : l'effet de pépite, le plateau et l'étendue (Figure 28) (Kaluzny *et al.* 1998). Le premier est la valeur de semivariance à 0 m de décalage entre les points (Figure 28). Cette semivariance, qui indique la variabilité locale, est considérée comme le bruit des données. Ce bruit est dû, d'une part, à la variabilité interne des données et, d'autre part, à des erreurs d'échantillonnage (Legendre et Fortin 1989, Legendre et Legendre 1998). Comme l'échantillonnage est le même pour toutes les méthodes, on peut considérer que les différences observées de l'effet pépite sont dues à la variabilité interne du *Volume total* de chacune des méthodes. Cette variabilité correspond à une échelle plus fine que les 100 m d'espacement entre les points de la grille. Le plateau, quant à lui, correspond à la valeur de semivariance lorsque celle-ci tend à la stabilité pour des distances croissantes (Figure 28). Il représente la semivariance lorsque la variable se comporte de façon aléatoire (Kaluzny *et al.* 1998). Pour sa part, la portée correspond à la distance à laquelle la semivariance commence à se stabiliser (Figure 28). Celle-ci est interprétée comme étant l'étendue de la structure spatiale du *Volume total* qui, au-delà de cette distance, se comporte comme une variable aléatoire.

Étant donné que la semivariance se calcule pour des données ponctuelles, une grille de points espacés de 100 m est superposée sur chacune des cartes, RT, Carto45, PartielT45, Carto36, PartielT42, Carto28, Carto21, PartielT17 et PartielT9. La valeur du *Volume total* de la parcelle sous-jacente de chacune des cartes est attribuée à chaque point. Cet espacement assure la présence d'au moins un point dans la majorité des parcelles de toutes

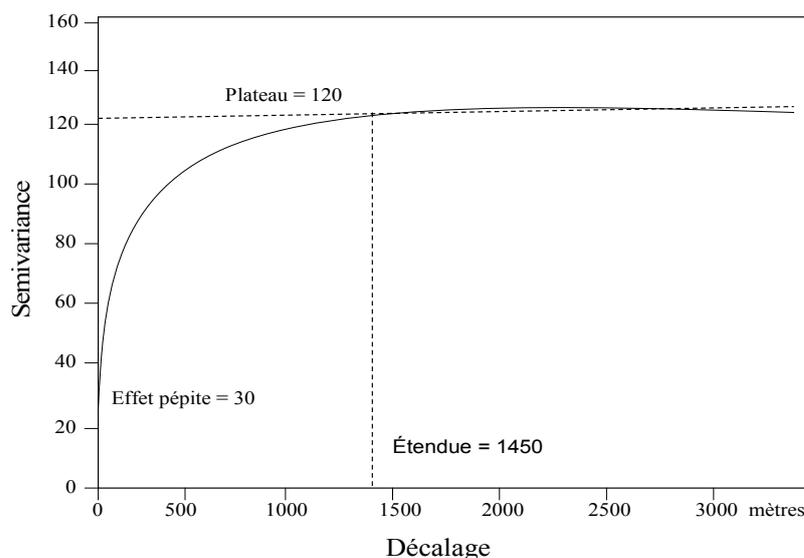


Figure 28. Variogramme général qui montre l'effet pépité, le plateau et la portée (Adaptée de Kaluzny *et al.* 1998).

les cartes. Pour le calcul de la semivariance, les 3792 points qui présentent des valeurs du *Volume total* issues de toutes les cartes sont considérés.

Des variogrammes isotropiques et anisotropiques sont construits pour chacune des modèles avec un décalage de 100 m entre les paires de points. Cette distance est fixée afin de considérer les voisins les plus proches en tenant compte de la taille moyenne des parcelles de la carte RT (2,36 ha). Les variogrammes s'étendent jusqu'à une distance de 5000 m, ce qui représente environ la moitié de l'étendue nord-sud de la Forêt Montmorency. Le variogramme isotropique implique que la variabilité du *Volume total* soit similaire dans toutes les directions tandis que le variogramme anisotropique implique que la variabilité spatiale est différente selon la direction. Les variogrammes sont construits avec des estimations robustes de la semivariance afin de réduire l'effet des valeurs extrêmes (Kaluzny *et al.* 1998). Des modèles théoriques sont ajustés à ces variogrammes expérimentaux.

## Résultats et discussion

En premier lieu, la description des trois arbres de régression construits en utilisant des ensembles différents de variables indépendantes est présentée. En deuxième lieu, les versions des arbres de taille optimale, dont les variables indépendantes permettent d'avoir une expression cartographique (Carto et PartielT), sont exposées. Par la suite, les groupes issus des régressions arborescentes sont comparés à ceux qui proviennent du regroupement actuel, en fonction de la variation du *Volume total*. Finalement, les représentations spatiales des groupes formés par les différentes méthodologies sont comparées.

### 1. Caractéristiques générales des modèles

Dans cette section, les informations générales des trois arbres de régression sont présentées. Les descriptions des arbres généraux, de Carto45 et PartielT45 ainsi que des versions non-géographiques de Carto et PartielT sont présentés à l'Annexe III.

La validation des arbres de régression Carto et PartielT avec des données indépendantes indique que les arbres de taille optimale sont ceux de 28 et de 17 groupes finaux ou feuilles pour les versions géographiques et ceux de 21 et de 9 feuilles respectivement pour les versions non-géographiques.

Les trois modèles de régression arborescente présentent des caractéristiques générales distinctes (Tableau 30). Les modèles généraux présentent une taille similaire, entre 50 et 53 feuilles pour les versions géographiques. Les coordonnées géographiques ne sont pas retenues dans le modèle CompletT. Les versions non-géographiques de Carto et de PartielT présentent une taille plus faible.

La proportion de déviance résiduelle (c'est-à-dire la déviance de l'ensemble des groupes formés) par rapport à la déviance de départ est similaire pour tous les modèles de terrain et inférieure à celle des modèles cartographiques (Tableau 30). Cette proportion indique la réduction de la variabilité de l'ensemble des p-e lors de leur regroupement en feuilles. Elle peut être interprétée comme la proportion de la variabilité du *Volume total* de l'ensemble

des p-e qui est « expliquée » par chacun des modèles. De ce point de vue, les modèles de terrain sont les plus performants parce qu'ils expliquent 56-57 % de la variabilité dans leur taille générale alors que cette proportion augmente à 62 % pour la taille optimale. Pour leur part, les arbres Carto expliquent entre 31 % et 34 % de cette variabilité. Il est important également de souligner que les modèles de 45 feuilles (comparables aux groupes RT) expliquent la même variabilité que les généraux.

Tableau 30. Caractéristiques des modèles de régression arborescente. Dev. : Déviance

Modèle		Nombre de variables indépendantes	Nombre de feuilles	Variabilité expliquée par le modèle [1- (Dév. résiduelle/Dév. départ)]
Carto version géographique	CartoG	9	52	0,311
	Carto45	9	45	0,310
	Carto28	8	28	0,331
Carto version non géographique	Carto36	8	36	0,319
	Carto21	6	21	0,342
	CompletTG	16	50	0,571
PartielT version géographique	PartielTG	9	53	0,562
	PartielT45	9	45	0,565
	PartielT17	6	17	0,621
PartielT version non géographique	PartielT42	7	42	0,560
	PartielT9	4	9	0,616

Ces résultats se rapprochent de ceux rapportés pour des modèles de régression linéaire construits avec un type similaire de variables pour la Forêt Montmorency (Coulombe et Lowell 1995). Dans leur étude, le modèle de régression linéaire le plus performant, construit également avec des variables de terrain, explique 50 % de la variabilité de la surface terrière. Une des raisons de l'efficacité supérieure des modèles de régression arborescente pourrait concerner leur capacité à saisir, d'une façon générale, les relations non linéaires entre les variables ainsi que leurs interactions (Clark et Pregibon 1992). Néanmoins, le but de ce chapitre n'est pas d'explorer les arbres de régression afin d'obtenir des estimations sur le volume ligneux mais plutôt d'explorer leur potentiel comme outil pour distribuer les p-e en groupes avec une variance du volume de matière ligneuse inférieure à celle des groupes traditionnels.

## 2. Les modèles de régression arborescente de taille optimale

Cette section présente les détails des modèles Carto et PartielT de taille optimale, c'est-à-dire, celle à laquelle la déviance est minimale lorsque le modèle est validé avec des données différentes de celles utilisées pour sa construction.

### 2.1. Le modèle Cartographique – Carto28 et Carto21

La méthode de taillage élimine certains des noeuds inférieurs du modèle général, par conséquent, les variables qui déterminent ces noeuds sont éliminées du modèle. Des neuf variables indépendantes du modèle CartoG (version géographique), Carto28 en retient huit : *Année de coupe (AC)*, *Groupe ment d'essences (GE)*, *Classe de pente (CP)*, *Classe de hauteur (CH)*, *Classe d'âge (CA)*, *Perturbation moyenne (Pert)*, *Coordonnée X (X)* et *Coordonnée Y (Y)*. Carto21, la version non-géographique retient pour sa part six des huit variables proposées : *Année de coupe (AC)*, *Groupe ment d'essences (GE)*, *Classe de pente (CP)*, *Classe de hauteur (CH)*, *Classe d'âge (CA)* et *Perturbation moyenne (Pert)*.

Les premières divisions dans les deux versions du modèle Carto de taille optimale sont similaires (Figure 29). Les variables les plus importantes sont *AC*, *CP*, *GE* et *CH* parce qu'elles interviennent dans les premières divisions qui réduisent fortement la déviance de la racine. La Figure 29 permet également d'apprécier que les feuilles qui correspondent aux noeuds générés par l'utilisation successive de *AC* (à gauche à la figure) sont beaucoup plus hétérogènes (les branches sont plus courtes) que celles dérivées de la combinaison *AC-CP* (la deuxième division).



Premièrement, la variable *AC* divise l'ensemble des p-e en deux sous-ensembles (Figure 30). Le premier correspond aux années de coupe qui ont eu lieu entre 1966 et 1979, entre 1987 et 1989 ainsi qu'aux p-e ayant subi des perturbations sévères telles un chablis total (cht) ou une épidémie sévère (es). Le deuxième sous-ensemble correspond aux années de coupe entre 1981 et 1983, entre 1992 et 1994 ainsi qu'aux p-e sans perturbation sévère (sp). Ce nœud de grande taille (278 p-e) contient curieusement des p-e où la végétation a subi des coupes très récentes et des p-e sans perturbation. Il semble discordant que les coupes des années proches ne produisent pas des valeurs semblables du *Volume total*. En effet, les p-e sans perturbation sont censées avoir des arbres de haute taille correspondant à une végétation ligneuse développée alors que la végétation d'une coupe récente est en régénération.

La même variable (*AC*) agit pour diviser le premier des sous-ensembles et diverses catégories se retrouvent dans chaque nouveau nœud (Figure 30). La division suivante utilise *GE* et sépare les p-e avec une végétation mélangée (SBb et BbS) des p-e avec une végétation résineuse où prédomine le sapin baumier (SS). Le mélange, dans un même nœud, des p-e correspondant à des situations discordantes détectées en premier pour *AC* se répète lors des autres divisions subséquentes.

Les différences entre les deux versions du modèle Carto de taille optimale se trouvent dans les branches inférieures, près des feuilles, où Carto28 utilise les variables géographiques et Carto21 utilise les variables *CA*, *CP*, *GE* et *CD*. Les détails de Carto21 sont présentés à l'Annexe III.

Il faut souligner que certaines des variables considérées *a priori* comme importantes selon la littérature pour guider le regroupement des strates forestières (Beaulieu 1994, Beaulieu et Lowell 1994, Bérard et Côté 1996, Coulombe 1996, De Groeve 1999), ne sont pas utilisées ou ne le sont que marginalement pour diviser des nœuds à un niveau inférieur, proche des feuilles. C'est le cas des variables *CD* (*Classe de densité*), *Orig* (*Perturbation d'origine*) et *SD* (*Stade de développement*). La première n'est pas retenue dans le modèle optimal

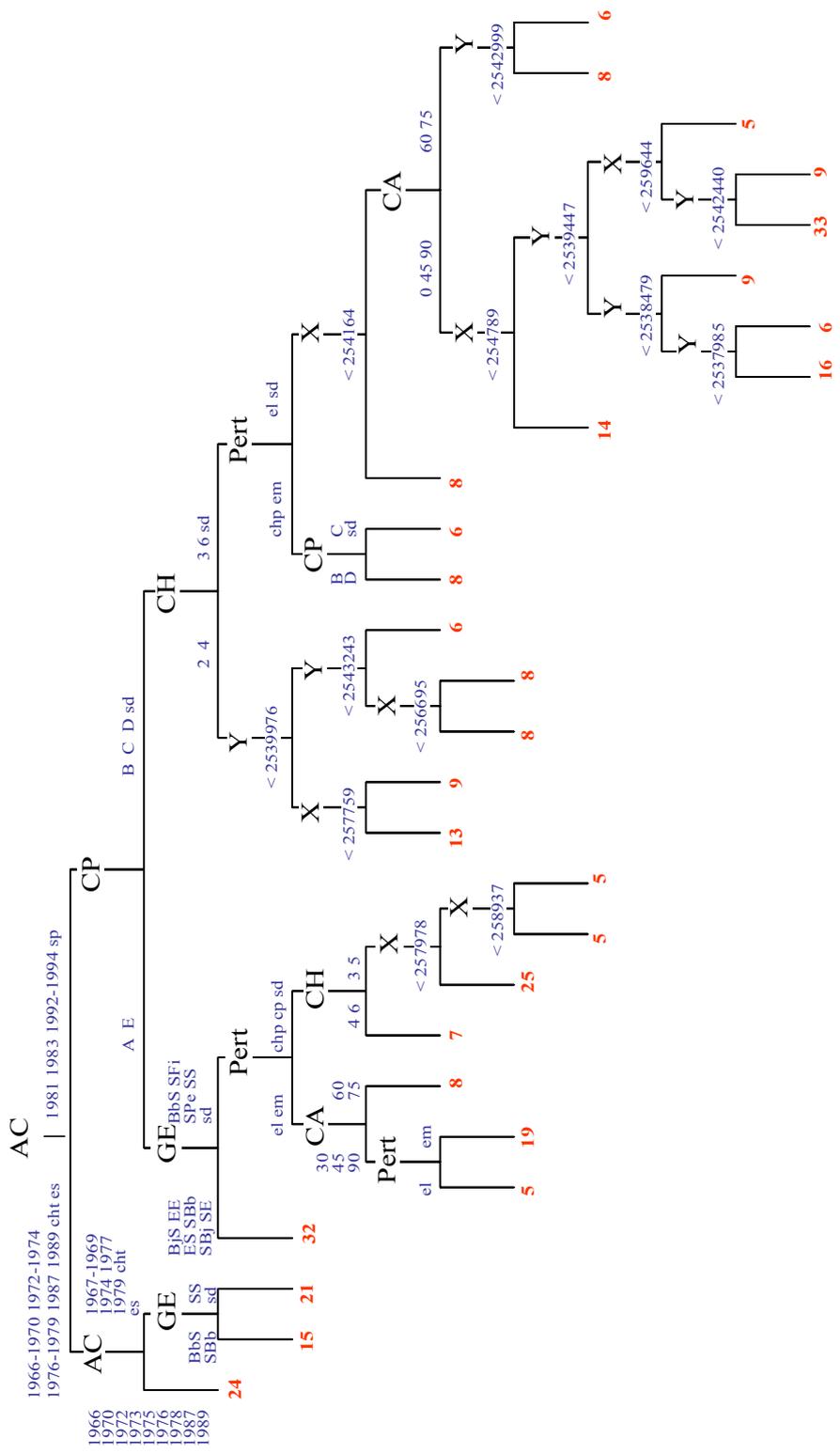


Figure 30. Dendrogramme du modèle Cartographique de taille optimale, dans la version qui considère les variables géographiques, Carto28. La longueur des branches est uniforme.

Carto28 et elle est utilisée une fois dans Carto21. *SD* est utilisée seulement dans la version non-géographique pour la définition des feuilles. Cette variable est présente seulement dans la cartographie et elle est employée par la méthode actuelle au Québec car tous les groupes formés sont homogènes pour cette variable. Le fait qu'elle est utilisée pour définir les feuilles indique que son effet sur la variabilité du volume ligneux dans la région à l'étude est plutôt local. *Orig*, pour sa part, n'est pas utilisée dans le modèle Carto, possiblement elle est remplacée par *AC*, l'année de coupe. En effet, les coupes sont le type de perturbation sévère le plus fréquente à la Forêt Montmorency. L'information fournie par cette variable serait donc redondante ou trop imprécise par rapport à celle de *AC*. Quoiqu'il en soit, les modèles de régression arborescente mettent nettement en évidence les variables clé pour expliquer les seuils de variabilité du volume de matière ligneuse des p-e dans la zone d'étude.

## 2.2. Le modèle Partiel de Terrain – PartielT17 et PartielT9

Le modèle Partiel de Terrain de taille optimale dans la version géographique, PartielT17, retient six des neuf variables du modèle général : *Classe de hauteur (CHT)*, *Groupe ment d'essences (GET)*, *Classe de densité (CDT)*, *Classe de pente (CPT)*, *Coordonnée X (X)* et *Coordonnée Y (Y)*. Le modèle dans la version non-géographique, PartielT9, retient quatre des sept variables du modèle général : *Classe de hauteur (CHT)*, *Groupe ment d'essences (GET)*, *Classe de densité (CDT)* et *Classe de pente (CPT)*.

Comme pour les modèles Carto, les premières divisions de PartielT17 et PartielT9 sont similaires. Les variables les plus importantes utilisées pour les divisions supérieures sont *CHT*, *CDT*, *GET* et *CPT*. Les trois premières divisions réduisent fortement la déviance de l'ensemble des p-e, en étant la première la plus importante (Figure 31). L'effet de cette première division a un impact sur la déviance plus accentué que la première division pour les modèles Carto, possiblement parce que *CHT* est plus directement reliée au volume ligneux qu'*AC*. Néanmoins, l'*Année de coupe (AC)* est d'une certaine façon reliée à la *Classe de hauteur (CHT)*. En effet, la première indique le temps écoulé depuis le début de la régénération de la végétation à la suite d'une perturbation sévère alors que la deuxième

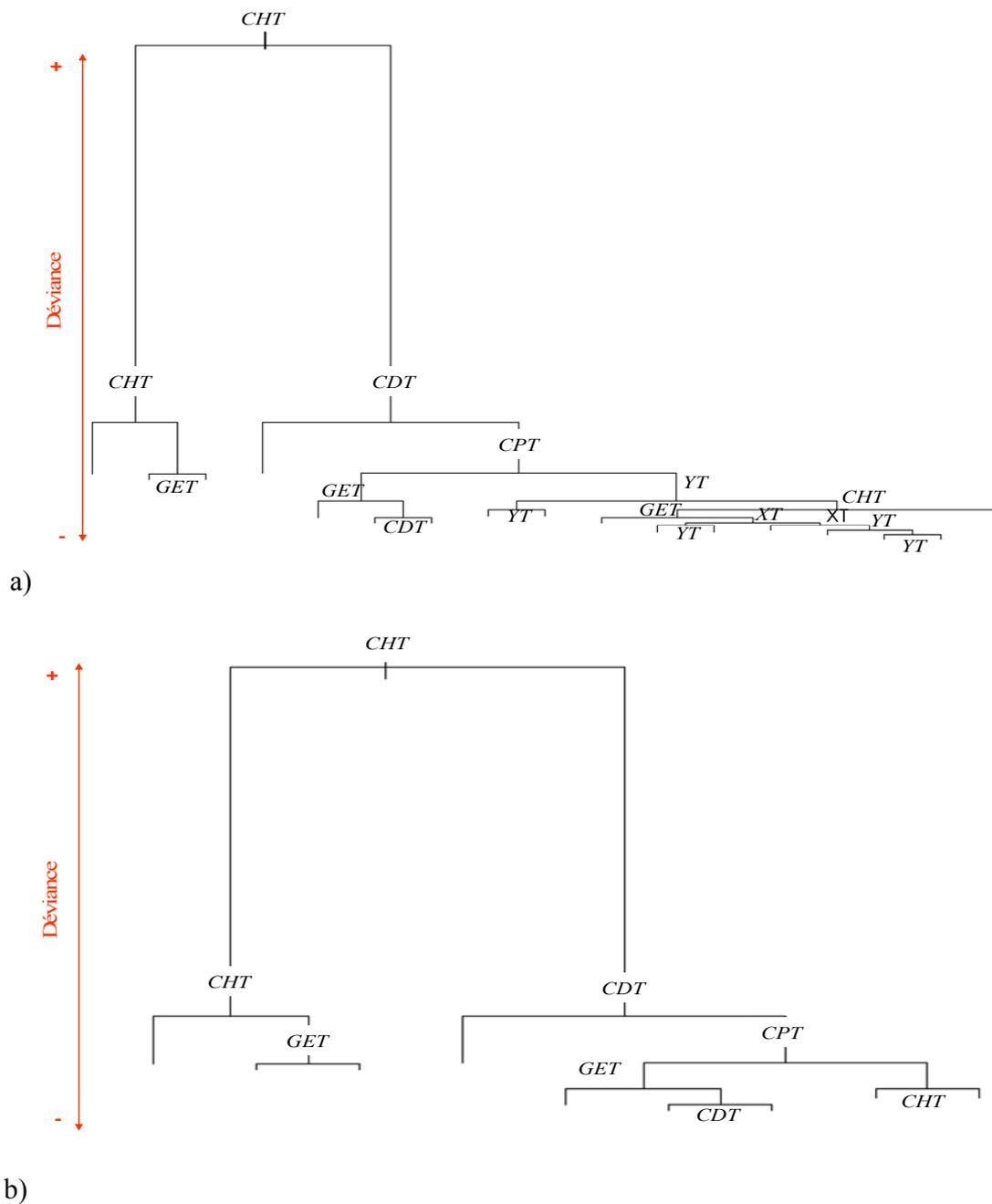


Figure 31. Dendrogramme du modèle Partiel de Terrain de taille optimale. a) version géographique, PartielT17 et b) version non-géographique, PartielT9. La longueur des branches est proportionnelle à la déviance. (Voir le Tableau 27 pour la définition des variables).

indique également, d'une façon indirecte, le temps d'évolution de la végétation ligneuse. En conséquence, on peut interpréter que l'aspect le plus important pour la réduction de la variabilité des p-e en ce qui a trait au volume de matière ligneuse est le degré de développement des arbres.

Dans la première partition, la variable *CHT* divise le nœud racine et le nœud résultant à gauche correspond aux p-e contenant des arbres de faible taille, d'une hauteur inférieure à 12 m (catégories 4, 5 et 6) alors que le nœud résultant à droite concerne les arbres plus grands (catégories 2 et 3). Dans le premier des nœuds, *CHT* sépare encore les p-e où les arbres dominants mesurent moins de 7 m des p-e contenant des arbres mesurant entre 7 et 12 m (Figure 32). En général on voit cette tendance pour les divisions « logiques » dans tout ce modèle ce qui contraste avec ce que nous avons vu pour Carto.

Le nœud dont les p-e contiennent des arbres qui mesurent plus de 12 m est divisé par la variable *CDT* où le nœud de gauche correspond aux p-e où le couvert ligneux varie entre 68 et 80 % (catégorie D) alors que le nœud de droite correspond aux p-e avec d'autres densités de couvert (Figure 32). La variable *CPT* est utilisée subséquemment pour les divisions des niveaux intermédiaires, ainsi que les variables géographiques.

Quant aux variables géographiques (*X* et *Y*) de PartielT17, elles sont utilisées lors de la définition des feuilles, c'est-à-dire à l'échelle locale, comme dans le modèle Carto28 (Figure 32). Le modèle PartielT9 utilise la variable *CHT* pour la division que dans PartielT17 est faite avec *YT* (les détails de PartielT9 sont présentés à l'Annexe III). La présence des variables géographiques dans les nœuds inférieurs suggère que ces variables représentent de façon indirecte des conditions locales déterminées par d'autres variables comme l'exposition ou le type de sol, qui correspondent à une échelle spatiale plus fine que celle des variables utilisées pour les modèles Carto et PartielT. Cette idée est soutenue par le fait que le modèle Complet de Terrain (CompletTG, dont les détails sont présentés à l'Annexe III) ne retient pas les variables géographiques et plusieurs des divisions inférieures sont déterminées à partir des variables telles que la texture des horizons B et C, le type de drainage et la pierrosité.



### 3. La variabilité des estimations du *Volume total* des groupes

Les résultats de cette section correspondent à la comparaison des groupes issus des modèles de régression arborescente avec les groupes issus de la méthode actuellement utilisée, en terme de la variabilité des estimations du *Volume total*. Les caractéristiques des groupes de faible et de forte variabilité, dont les CVs correspondent au 20<sup>e</sup> et au 80<sup>e</sup> percentile respectivement sont également présentées.

#### 3.1. Le Coefficient de variation (CV) du *Volume total* des groupes

La Figure 33 présente les valeurs des percentiles des Coefficients de variation (CVs) du *Volume total* des groupes issus de la méthode actuelle de regroupement (groupes RT) ainsi que ceux des groupes issus des deux versions des modèles Carto (Carto45, Carto28, Carto36 et Carto21) et PartielT (PartielT45, PartielT17, PartielT42 et PartielT9). La figure permet d'apprécier que, en général, les groupes formés par les régressions arborescentes sont plus homogènes que les groupes RT, surtout pour les groupes dont le CV est élevé. L'exception la constituent les groupes issus de la version non-géographiques du modèle Carto lesquels sont légèrement plus variables que les groupes RT si bien à partir du 80<sup>e</sup> percentile rejoignent les valeurs des autres groupes issus des régressions.

Le CV minimal des groupes formés par les versions géographiques des modèles arborescentes est 50 % du CV minimal des groupes formés par la méthode actuelle (RT) tandis que le CV maximal varie entre 60 et 74 % du CV maximal des groupes RT (Figure 33). Pour les groupes issus des versions non-géographiques ces différences diminuent. En effet, le CV minimal de ces groupes est entre 66 et 130 % du CV minimal des groupes RT tandis que le CV maximal varie entre 61 et 94 % du CV maximal des groupes RT.

Les groupes formés par les versions géographiques des deux modèles sont plus homogènes que ceux formés par les versions non-géographiques. Parmi les premiers, les groupes formés par le modèle PartielT sont plus homogènes que ceux formés par le modèle

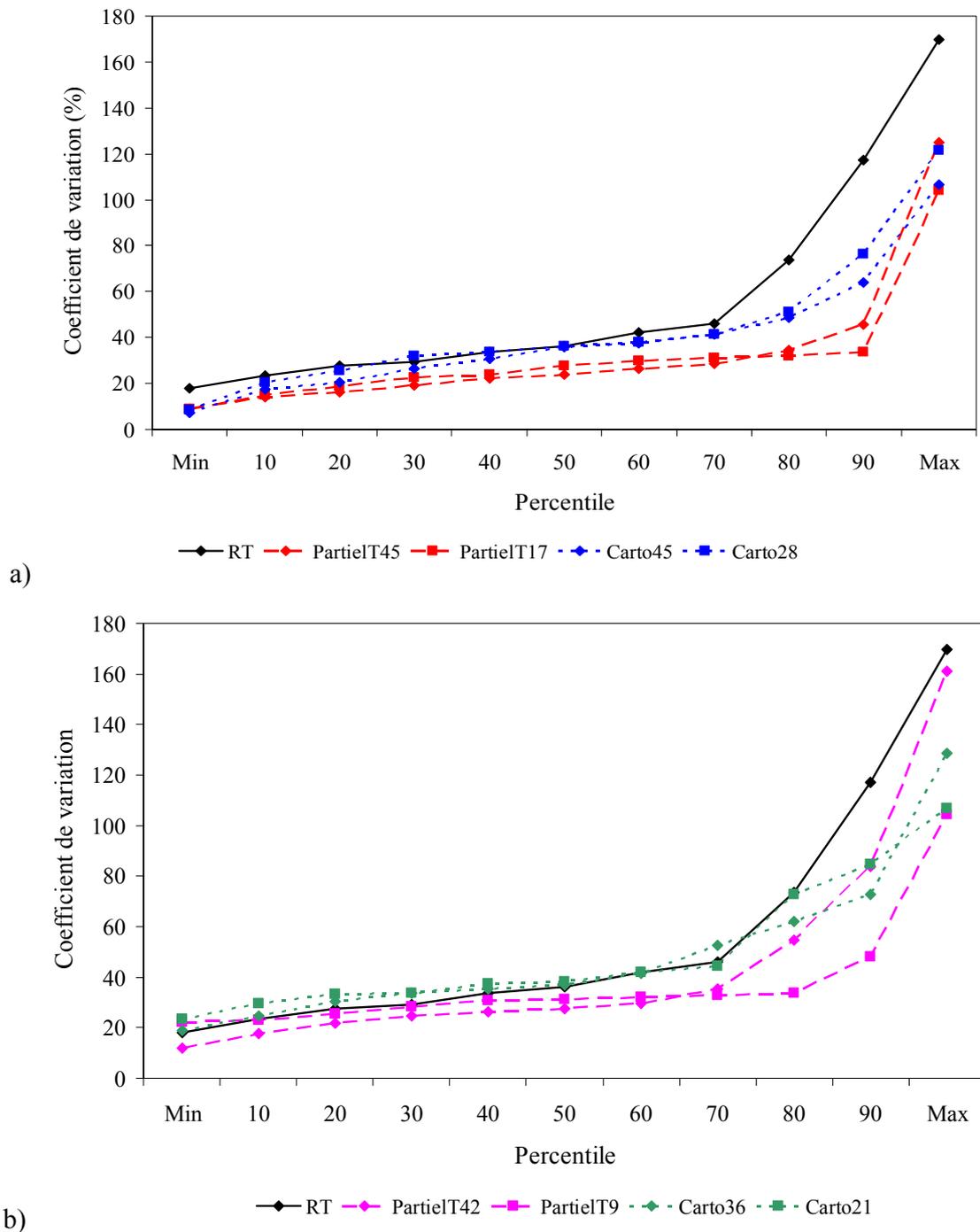


Figure 33. Percentiles des coefficients de variation du *Volume total* des groupes issus de la méthode actuelle de regroupement (RT) et des modèles de régression arborescente Cartographique et Partiel de Terrain : a) versions géographiques (Carto45, Carto28, PartielT45 et PartielT17) et b) versions non géographiques (Carto36, Carto21, PartielT42 et PartielT9).

Carto, spécialement à partir du 20<sup>e</sup> percentile. Cette différence s'amplifie jusqu'au 90<sup>e</sup> percentile, ce qui signifie que 90 % des groupes PartielT sont beaucoup moins hétérogènes que 90 % des groupes Carto.

La Figure 33 indique également que le nombre de groupes affecte faiblement les valeurs des CVs car pour la majorité des percentiles les CVs des groupes issus des modèles de 45 feuilles sont assez semblables à ceux des groupes issus des modèles de taille optimale. Par conséquent, les différences observées dans la variabilité des estimations découlent de la méthode de formation des groupes. Ces résultats soutiennent l'hypothèse que la régression arborescente permet de former des groupes plus homogènes de p-e en terme du volume de matière ligneuse que les groupes formés par la méthode actuellement utilisée. La section suivante va nous indiquer si cette variabilité inférieure se maintient à l'échelle des peuplements.

### **3.2. Rapport entre le *Volume total* estimé pour les groupes et celui des p-e**

Le rapport entre le *Volume total* estimé selon les différentes méthodes pour le peuplement où se localise chaque p-e et celui de la p-e indique la variabilité à l'échelle locale. L'examen des rapports obtenus indique qu'ils varient en fonction du *Volume total* des p-e, donc pour faciliter les comparaisons les p-e ont été regroupées en quintiles selon cette variable.

La Figure 34 permet d'identifier une relation inverse entre le volume ligneux des p-e et les estimations de toutes les méthodes. En effet, en général, lorsque le volume ligneux des p-e est très faible toutes les méthodes le surestiment tandis que lorsqu'il augmente toutes les méthodes tendent à le sous-estimer (Figure 34a et e) (un rapport supérieur à 100 indique une surestimation tandis qu'un rapport inférieur à 100 indique une sous-estimation). De plus, pour le volume ligneux faible les estimations des modèles PartielT sont les plus variables et celles de Carto les moins variables, les estimations RT étant intermédiaires (Figure 34a). Cette différence entre la variabilité et le degré de surestimation des modèles PartielT et Carto peut obéir au fait que pour le dernier le *Volume total* des p-e est expliqué par les caractéristiques des peuplements où les p-e sont situées. Ce modèle établit ainsi une

relation directe entre l'information des p-e et de la cartographie, informations que comme nous verrons dans le chapitre suivant ne sont pas toujours concordantes.

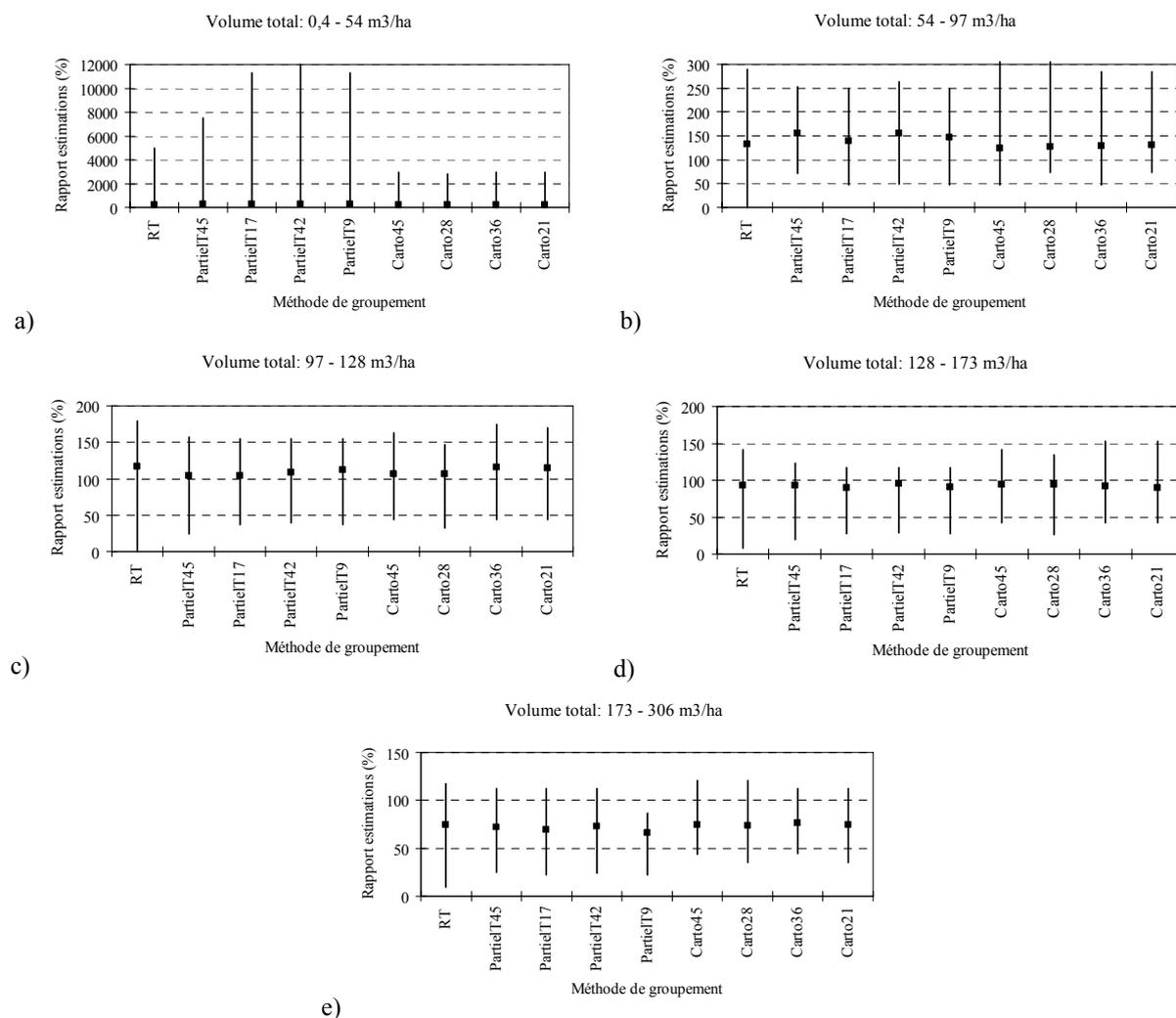


Figure 34. Rapport entre le *Volume total* estimé pour les groupes formés par les différentes méthodes et celui des p-e. Les valeurs minimales, maximales et médianes sont présentées pour chaque méthode de groupement. Les p-e sont regroupées en quintiles selon le *Volume total* : a) 0,4-54 m<sup>3</sup>/ha, b) 54-97 m<sup>3</sup>/ha, c) 97-128 m<sup>3</sup>/ha, d) 128-173 m<sup>3</sup>/ha, e) 173-306 m<sup>3</sup>/ha et f) 205-306 m<sup>3</sup>/ha.

Cependant, lorsque le volume ligneux augmente les estimations des modèles PartielT et Carto sont moins variables que celles de RT (Figure 34b à e). Les différences de variabilité des estimations de PartielT et de Carto sont faibles et ne montrent pas une tendance définie.

À l'exception des cas où le volume ligneux est très faible, les modèles de régression arborescente produisent, à l'échelle locale, des estimations moins variables que la méthode actuellement utilisée au Québec.

### **3.3. Caractéristiques des groupes de faible et de forte hétérogénéité**

Au-delà des différences de hétérogénéité parmi les groupes issus des différentes méthodes, il est pertinent de déterminer si les groupes de forte et de faible hétérogénéité se ressemblent en terme des caractéristiques forestières. Les premiers ont un CV inférieur au 20<sup>e</sup> percentile alors que les deuxièmes l'ont supérieur au 80<sup>e</sup> percentile.

À l'exception d'un cas, les comparaisons révèlent que les groupes de faible et de forte hétérogénéité sont significativement différents ( $\alpha=0,05$ ) de l'ensemble des p-e en ce qui a trait aux fréquences des catégories des variables *Groupement d'essences*, *Classe de densité*, *Classe de hauteur*, *Classe d'âge* et *Perturbation d'origine* (Tableau 31).

Ainsi, dans les groupes de faible hétérogénéité sont fréquentes les p-e avec une végétation de type résineuse où prédomine le sapin baumier (SS). Cette végétation ligneuse, d'un âge estimée en 50 ans, couvre plus du 61 % de la surface et elle est composée d'arbres d'une hauteur supérieure à 12 m. Les situations d'une perturbation sévère, parmi ces p-e, sont presque inexistantes (Tableau 31).

Dans les groupes de forte hétérogénéité, au contraire, les p-e présentent fréquemment une végétation ligneuse de type mélangé (comme par exemple SBb) dont la hauteur des arbres dominants est inférieure à 7 m. Ces p-e, dont l'âge la plus fréquente se situe en 10 ou 30 ans ont subi pour la plupart une perturbation sévère telle une coupe totale ou une épidémie.

Par rapport à la densité de couvert, les p-e de ces groupes, au delà de la méthode utilisée pour les former, présentent une fréquence élevée des données manquantes. Il faut souligner que la variable *Classe de densité* estimée sur le terrain (*CDT*) a été utilisée dans le modèle de régression PartielT45. Aucune information n'est disponible sur l'importance de cette

Tableau 31. Fréquences des catégories des variables estimées sur le terrain *Groupement d'essences*, *Classe de densité*, *Classe de hauteur*, *Classe d'âge*, *Perturbation d'origine* et *Perturbation moyenne*, dans les groupes RT, Carto45 et PartielT45 de faible et de forte hétérogénéité. La dernière colonne contient les fréquences des catégories pour l'ensemble des 338 p-e. La signification correspond au  $\chi^2$  de chaque distribution.

Catégorie	Groupes de faible hétérogénéité (20 <sup>e</sup> percentile)			Groupes de forte hétérogénéité (80 <sup>e</sup> percentile)			Ensemble des p-e
	RT	Carto45	PartielT45	RT	Carto45	PartielT45	
<i>Groupement d'essences (GET)</i>							
SBb	0,07	0,15	0,04	0,22	0,19	0,21	0,17
SS	0,75	0,62	0,87	0,33	0,36	0,35	0,59
Autres	0,03	0,23	0,09	0,44	0,45	0,44	0,24
Signification <sup>1</sup>	**		**	**	**	**	
<i>Classe de densité (CDT)</i>							
A + B	0,55	0,61	0,61	0,33	0,31	0,31	0,43
S/D	0,04	0,00	0,01	0,39	0,23	0,32	0,08
Signification <sup>1</sup>	*	**	**	**	**	**	
<i>Classe de hauteur (CHT)</i>							
2 + 3	0,74	0,84	0,93	0,19	0,16	0,18	0,63
5 + 6	0,05	0,00	0,00	0,78	0,67	0,70	0,18
Signification <sup>1</sup>	**	**	**	**	**	**	
<i>Classe d'âge (CAT)</i>							
10 + 30	0,08	0,03	0,00	0,81	0,77	0,80	0,23
50	0,79	0,75	0,86	0,14	0,20	0,15	0,60
Signification <sup>1</sup>	**	**	**	**	**	**	
<i>Perturbation d'origine (OrigT)</i>							
ct + es	0,05	0,00	0,01	0,72	0,58	0,64	0,16
sp	0,95	1,00	0,99	0,28	0,42	0,36	0,84
Signification <sup>1</sup>	**	**	**	**	**	**	

<sup>1</sup> Indique une distribution parmi les catégories non aléatoire au niveau de  $\alpha = *--0,05$  et  $**--0,01$ .

variable dans la définition des groupes par la méthode actuellement employée. Cependant, le fait que les proportions soient aussi élevées ou faibles dans les groupes issus des trois méthodes indique une tendance. Il est possible que les groupes de forte hétérogénéité contiennent des p-e qui correspondent à des situations très diverses et dont l'incorporation à d'autres groupes augmenterait leur hétérogénéité interne. Le fait d'inclure les p-e avec des données manquantes dans les modèles de régression arborescente obéit à un souci de

comparer sur des bases les plus semblables possibles les groupes ainsi formés avec ceux formés par la méthode traditionnelle, laquelle regroupe toutes les p-e.

Selon les résultats présentés, les estimations du volume de matière ligneuse seront moins variables lorsque la végétation est dans une étape de pré-maturité que lorsqu'elle est en régénération à la suite d'une perturbation sévère.

#### **4. La représentation spatiale des groupes issus des différents modèles**

Cette section explore la représentation visuelle des structures spatiales des regroupements de p-e effectués à l'aide des modèles de régression arborescente afin d'être en mesure de les comparer aux groupes issus de la méthode actuellement employée.

##### **4.1. Comparaison des caractéristiques des cartes**

Les cartes élaborées à partir des modèles de régression arborescente, soit les cartes Carto45, Carto28, PartielT45 et PartielT17, sont comparées avec la carte des groupes issus de la méthode actuellement utilisée (carte RT), au moyen d'indices sur les caractéristiques des cartes. Ces indices sont : le nombre, la densité et la surface moyenne des parcelles ainsi que la surface de la plus grande parcelle. Les parcelles sont représentées par les polygones de la carte et elles correspondent à des groupes de peuplements adjacents qui sont accordés la même estimation pour le *Volume total*.

Le nombre de parcelles est clairement plus élevé dans la carte RT que dans celles dérivées des arbres de régression (Tableau 32). Pour ces dernières, le nombre de parcelles diminue à mesure que le nombre de feuilles diminue. Cependant, les cartes Carto45 et Carto28 contiennent légèrement plus de parcelles que PartielT45.

Les cartes dérivées des modèles de régression ne contiennent pas d'estimation du *Volume total* pour tous les peuplements. La raison est que certains des peuplements de la carte forestière de 1992 ne s'ajustent pas aux règles de ces modèles. Par exemple, dans la carte forestière certains peuplements présentent un type de végétation mixte dominée par le

peuplier (catégorie PeS du *Groupement d'essences*) qui est absent dans les p-e. Ces peuplements ne seront pas classifiés parce que les règles ne contiennent pas cette catégorie.

Tableau 32. Caractéristiques non spatiales des cartes RT, Carto45, PartielT45, Carto28 et PartielT17.

Carte	Nombre de parcelles	% de parcelles avec estimation	Densité de parcelles (100 ha)	Surface moyenne des parcelles (ha)	Surface de la plus grande parcelle (ha)
RT	2709	100	42,35	2,36	161,69
Carto45	2093	90	32,10	3,11	103,20
PartielT45	2033	82	31,25	3,20	138,91
Carto28	2093	90	32,10	3,11	103,20
PartielT17	1851	91	28,45	3,52	175,67

De ce point de vue, les modèles Carto sont plus efficaces parce qu'ils fournissent des estimations pour 91 % des parcelles, de même que PartielT17, alors que PartielT45 le fait pour 82 % des parcelles (Tableau 32). Le modèle PartielT17 contient un nombre inférieur de critères (ou variables) que les modèles plus développés, ce qui permet l'inclusion d'un nombre supérieur de peuplements dans la classification. Du point de vue pratique, cette déficience des cartes dérivées des arbres de régression peut être palliée en assignant aux peuplements qui ne sont pas atteints par les critères de définition de groupes, les estimations du volume ligneux du nœud supérieur le plus proche.

La densité est en relation inverse à la surface moyenne des parcelles. La carte RT présente la surface moyenne la plus basse et la densité de parcelles la plus élevée. La carte Carto45, quant à elle, présente une densité de parcelles supérieure et une surface moyenne inférieure à celles de la carte PartielT45. La plus grande parcelle de la carte RT est, pour sa part, plus étendue que celles des autres cartes, hormis celle de la carte PartielT17.

En conclusion, les cartes des groupes issus des modèles de régression arborescente contiennent un nombre inférieur de parcelles que celle des groupes RT, cependant les différences sont marquées seulement lorsque le nombre de groupes est réduit.

#### 4.2. Comparaison des groupes issus des différentes méthodes en terme de la structure spatiale des estimations du *Volume total*

Les différences entre les variogrammes anisotropiques et isotropiques sont faibles, par conséquent seules ces derniers sont présentés. Ils indiquent des différences dans la structure spatiale du *Volume total* en fonction de la méthode utilisée pour former les groupes (Tableau 33, les variogrammes expérimentaux sont détaillées à l'Annexe III).

Les modèles théoriques qui correspondent aux variogrammes expérimentaux sont exponentiels. L'effet de pépité des estimations PartielT45 et PartielT42 ainsi que des versions non-géographiques de Carto indiquent une variabilité locale proche de celle des estimations RT (Tableau 33). Les versions géographiques de Carto (Carto45 et Carto28) présentent une variabilité locale nettement moindre, de même que PartielT17 et PartielT9. Comme mentionné dans la méthodologie, l'échantillonnage est similaire pour toutes les modèles, en conséquence les différences de variabilité locale peuvent être attribuées aux méthodes de formation des groupes.

Tableau 33. Valeurs de l'effet pépité, du plateau et de l'étendue des variogrammes théoriques des estimations du *Volume total* issues des différents modèles ( $n=3792$  points).

Modèle	Variogramme théorique	Effet de pépité	Plateau	Portée (m)
RT	Exponentiel	1222	1789	1117
PartielT45	Exponentiel	1025	2036	1659
PartielT17	Exponentiel	651	1624	1907
PartielT42	Exponentiel	1246	1611	1534
PartielT9	Exponentiel	668	1608	1904
Carto45	Exponentiel	685	2549	2054
Carto28	Exponentiel	608	2766	3119
Carto36	Exponentiel	1141	1460	2215
Carto21	Exponentiel	1027	1716	3246

La portée des estimations RT, c'est-à-dire la distance au-delà de laquelle la semivariance tend à la stabilisation, est la plus faible tandis que celle des estimations PartielT est intermédiaire et celle des estimations Carto est la plus étendue. Pour les estimations des modèles de régression arborescente la portée augmente lorsque le nombre de groupes diminue ce qui est consistante avec l'augmentation de la taille des parcelles présentée à la

section antérieure. Par rapport aux plateaux, ceux des versions géographiques de Carto sont les plus élevés, suivis par celui de PartielT45. Le plateau des estimations RT est intermédiaire et ceux de PartielT17, PartielT42 et partielT9 sont d'une valeur inférieure.

Les estimations issues des modèles présentent donc une structure spatiale plus étendue que les estimations issues de la méthode actuellement utilisée au Québec et celles des versions géographiques de Carto ainsi que des modèles de taille optimale de PartielT présentent également une variabilité locale beaucoup plus faible. Cette information peut être utile lors de la délimitation des aires à intervenir car elle indique que les estimations issues de PartielT17 sont, par exemple, moins variables dans l'espace que celles des groupes RT.

Nonobstant le fait que les modèles de terrain (PartielT) sont plus performants pour réduire en général la variabilité des estimations du volume de matière ligneuse que les modèles cartographiques, il est pertinent de considérer que ces derniers intègrent, d'une certaine manière, l'incertitude reliée à la concordance entre l'information de terrain et de la carte. Ceci se reflète dans le fait que la variabilité locale est similaire pour les deux modèles. Les variables cartographiques, bien qu'elles concernent les mêmes structures de végétation que les levés de terrain, décrivent une autre réalité. À l'étape de la photo-interprétation, les catégories sont assignées aux peuplements de la carte forestière en considérant une aire minimale. Cette généralisation entraîne une perte de détail qui croît de pair avec l'hétérogénéité de la végétation. Les variables cartographiques correspondent donc à un grain spatial plus large que les mesures de terrain. Par conséquent, la meilleure performance des modèles de terrain pour minimiser la variabilité du volume de matière ligneuse des groupes de p-e, peut être perdue lors du transfert des règles basées sur les variables de terrain aux variables cartographiques. Le chapitre suivant porte sur l'incertitude liée à la concordance entre ces deux sources d'information.

## **Conclusions**

Trois arbres de régression sont comparés. Ceux basés sur les levés de terrain (CompleT et PartielT) sont les plus performants du fait qu'ils expliquent, à leur taille optimale, 62 % de

la variabilité des observations. Cependant, l'utilisation du modèle de terrain construit avec un plus grand nombre de variables (CompleT) est restreinte en raison du caractère local et du coût de l'information.

Par rapport à la variabilité des estimations indiqué par le CV, les versions géographiques des modèles PartielT et Carto sont plus efficaces que les versions non-géographiques car elles produisent des estimations moins variables. Parmi les versions géographiques, les arbres PartielT présentent des CVs inférieurs du *Volume total* des groupes de p-e en comparaison aux CVs des groupes Carto. Cependant, les groupes de p-e issus des versions géographiques présentent des CVs du *Volume total* dont le maximum se situe entre 60 et 72 % du maximum des groupes formés par la méthode actuellement utilisée au Québec. De plus, les modèles Carto produisent à l'échelle locale des estimations moins variables que la méthode actuellement utilisée au Québec. Ainsi les modèles de régression arborescente produisent des groupes plus homogènes en terme du volume de matière ligneuse que cette dernière méthode.

Les variables les plus importantes pour diviser les observations selon des changements élevés de variabilité du *Volume total* sont reliées à la taille des arbres dans la forêt, dans le modèle du terrain (PartielT) et au temps écoulé depuis la dernière coupe dans le cas du modèle cartographique (Carto) (Figure 35). Les variables *Classe de densité*, *Classe de pente* et *Groupement d'essences* sont utilisées pour les divisions subséquentes. Les variables géographiques (*X* et *Y*) sont fréquemment utilisées dans les versions géographiques des modèles pour les divisions des nœuds intermédiaires et inférieurs, lors de la définition des feuilles. Ces variables représentent possiblement, d'une manière indirecte, des conditions environnementales locales non représentées par les variables liées à la structure de la végétation forestière.

Les groupes de faible et de forte hétérogénéité pour le *Volume total* présentent des fréquences de catégories des variables forestières (comme le *Groupement d'essences*, la *Classe de densité*, la *Classe de hauteur*, la *Classe d'âge* et la *Perturbation d'origine*) similaires, au-delà de la méthode utilisée pour les former. Les groupes de faible hétérogénéité correspondent à un type de végétation caractéristique des étapes de pré-

maturité de la forêt alors que ceux de forte hétérogénéité présentent des propriétés correspondant à la végétation forestière en régénération à la suite d'une perturbation sévère.

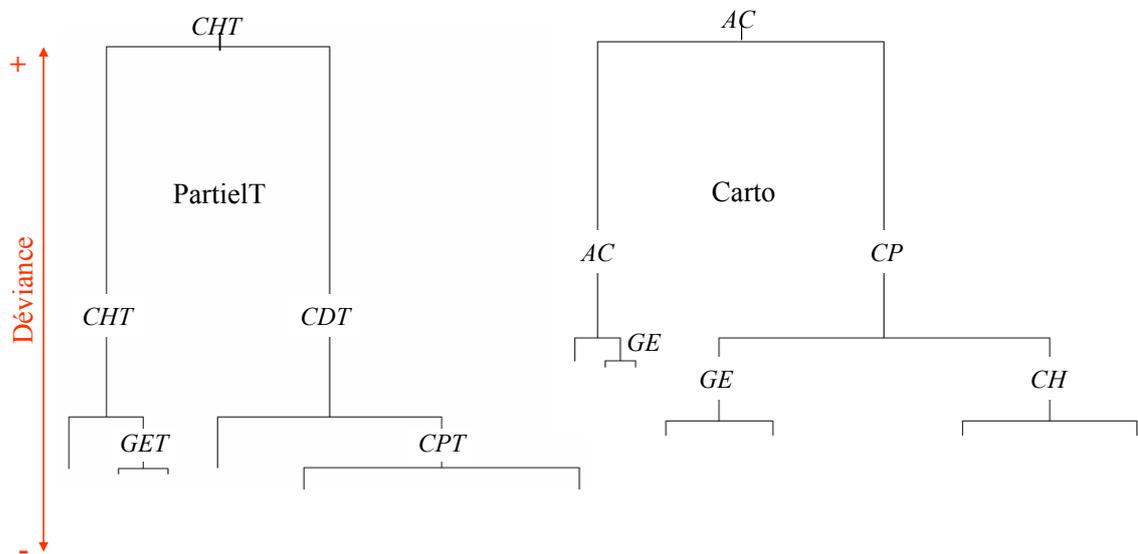


Figure 35. Les variables les plus importantes des deux modèles de régression arborescente.

Les différences concernant le nombre et la taille des parcelles entre les cartes dérivées des différentes méthodes sont peu accentuées. Elles s'accroissent lorsque les cartes correspondent à un nombre restreint de groupes.

Quant à la structure spatiale des estimations du volume ligneux des groupes issus des différentes méthodologies, celles basées sur le modèle de terrain de taille optimale (PartielT17 et PartielT9) sont nettement moins variables. Les estimations issues de la méthode de regroupement actuelle (RT) présentent la variabilité locale la plus élevée et la structure spatiale la moins étendue.

Tous les aspects énumérés conduisent à conclure que les groupes issus des arbres de régression, spécialement les versions géographiques, sont moins variables en terme du volume de matière ligneuse et cela concerne autant les groupes que leur représentation spatiale.

Les arbres de régression offrent une technique intéressante pour explorer la structure des données forestières en mettant en évidence les variables les plus importantes pour modéliser la variabilité du volume ligneux dans le territoire étudié. Les arbres de régression construits avec des variables de terrain sont aussi intéressants que ceux construits avec des variables cartographiques. Les premiers sont très performants pour expliquer la variabilité des données mais les seconds permettent de regrouper directement les p-e selon les variables de la carte forestière, ce qui diminue l'incertitude liée au transfert des règles basées sur des variables de nature différentes bien que semblables.

L'utilisation des modèles de régression arborescente permet également d'explicitier un processus qui repose actuellement sur la connaissance locale et l'expérience des experts, d'une valeur appréciable mais difficilement transférable. L'analyse des groupes et des variables de la carte forestière a permis d'inférer certains de ces critères utilisés pour leur formation. Cependant, la perception selon laquelle cette procédure repose principalement sur une connaissance approfondie de la forêt par les experts demeure persistante. Un autre fait intéressant est que les arbres de régression offrent la possibilité de comparer l'effet de différentes variables lors de la division des nœuds sur les estimations du volume de matière ligneuse et sa variabilité, de même que l'exploration des arbres de régression produits avec des ensembles de données d'autres types de forêts.

## **Résumé du chapitre**

Ce chapitre a examiné la régression arborescente comme outil pour regrouper les p-e afin d'obtenir des estimations du volume de matière ligneuse plus homogènes que celles obtenues par la méthode en usage actuellement au Québec. Ces estimations sont normalement associées aux strates de la carte forestière. Dans cette dernière, aucune information n'est disponible sur la concordance entre l'information de terrain et celle de la carte. Cet aspect est important lors du calcul du volume disponible sur une zone particulière parce qu'elle constitue une source supplémentaire de variabilité qui s'ajoute aux estimations. Le prochain chapitre propose une méthodologie pour quantifier la concordance carte/terrain des strates forestières.

# Chapitre VI - Les strates forestières

## Introduction

Un constat découle du chapitre précédent : au-delà de la méthode utilisée pour regrouper les placettes-échantillons (p-e), les estimations du volume de matière ligneuse correspondent à des groupes de p-e homogènes pour un ensemble de variables. Ces estimations sont associées aux strates de la carte forestière afin de les appliquer à toute l'étendue de la forêt et de rendre ainsi possible le calcul du volume de bois lors de la planification des activités extractives. Le transfert des estimations aux strates forestières ajoute une variabilité additionnelle aux estimations, qui est reliée à l'incertitude sur les attributs des strates. Cette incertitude est complexe et comprend des aspects qui vont au-delà de l'exactitude des attributs.

Les strates forestières correspondent aux différentes classes de forêt présentes sur un territoire. Ces entités, issues d'une taxonomie forestière, sont définies à partir d'un ensemble de variables qualitatives (Bérard et Côté 1996). Au moment de l'élaboration de la carte forestière, le photo-interprète associe des aspects comme la texture et la couleur des surfaces aux catégories des différentes variables afin de classer les polygones qu'il délimite sur la photo aérienne en strates forestières. La carte forestière présente donc les surfaces et la répartition spatiale des strates forestières photo-interprétées au moyen des polygones/peuplements qui appartiennent à chacune des strates.

Le traitement de l'incertitude thématique des cartes forestières vise habituellement la modélisation de l'erreur des attributs cartographiés (De Groeve 1999). Cependant, même si la définition des attributs est exacte, une incertitude ou variabilité liée aux caractéristiques des strates demeure. En effet, les strates de la carte forestière impliquent une certaine uniformité en termes des catégories qui les définissent. Toutefois, cette uniformité ou homogénéité est loin d'être réelle et ce, pour au moins trois raisons : (1) la variation spatiale des conditions environnementales à différentes échelles génère de l'hétérogénéité dans la végétation (Bérard et Côté 1996, Johnston 1999, Lowell et Liu 1999); (2) comme

discuté au Chapitre III, la définition des catégories utilisées pour déterminer les strates implique une diversité considérable de conditions de la végétation (Ministère des Ressources naturelles du Québec 1995) (Tableau 1, Tableau 2 et Tableau 3); et (3) les normes de cartographie forestière au Québec imposent une unité cartographique minimale de 8 ha (Ministère des Ressources naturelles du Québec 1995), ce qui entraîne la fusion des polygones de faible taille aux polygones plus étendus qui ont une catégorie prédominante (Goodchild 2001). Cette agrégation de petits polygones est reconnue comme une source d'incertitude sur les cartes forestières (Edwards 1994). Ces trois aspects influencent la classification (photo-interprétation) des strates dans la carte.

Les p-e, pour leur part, ont un grain beaucoup plus fin (la surface d'une p-e est de 400 m<sup>2</sup>) et contiennent de l'information à caractère beaucoup plus local. La relation entre les informations saisies par les deux sources est illustrée à la Figure 36. À l'échelle du polygone/peuplement, l'information (ou signal) locale des p-e apparaît comme une ligne qui représente les conditions prédominantes ou moyennes dans le polygone. La taille, le nombre et la localisation des p-e dans la strate détermineront leur représentativité des conditions du polygone/peuplement.

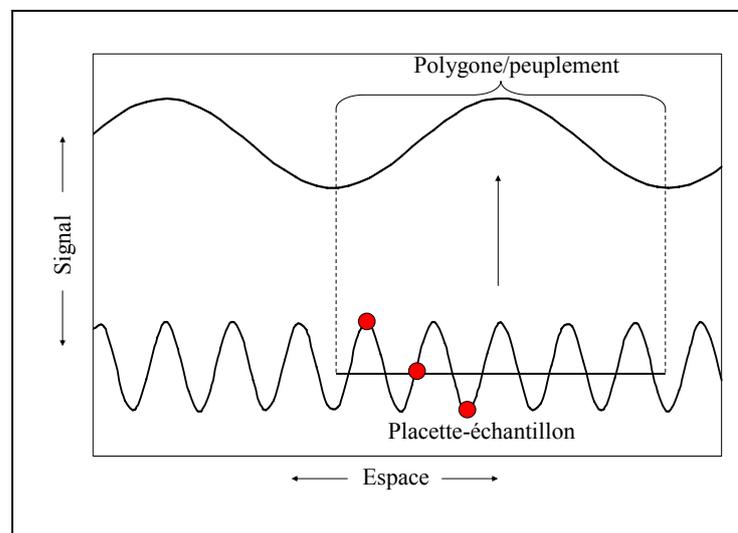


Figure 36. Relation entre les échelles des informations (signaux) des polygones/peuplements de la carte forestière et les placettes-échantillons (p-e). Modifiée de Allen et Hoekstra (1991).

Les méthodes de traitement de l'incertitude sur les attributs cartographiques ont été développées principalement pour évaluer les classifications des images satellites (Chrisman 1991, Veregin 1999). La majorité des méthodes d'évaluation et de communication de l'incertitude thématique sont basées sur la matrice de confusion. Cette dernière permet de comparer les classes de l'image à évaluer avec les classes des données de référence provenant d'une source considérée comme plus fiable (Chrisman 1991, Foody 2002, Friedl *et al.* 2001, Kyriakidis et Dungan 2001, Moghaddam *et al.* 2002, Stehman 1997, 1998, Veregin 1999). Parmi les méthodes d'évaluation de l'incertitude thématique, le pourcentage de concordance globale, le pourcentage de concordance de chaque catégorie et le coefficient kappa de Cohen (Foody 2002, Stehman 1997) sont les plus employés. Les dissimilarités entre les deux jeux de données sont généralement attribuées aux erreurs de la classification évaluée, bien que la classification de référence puisse également contenir des erreurs (Foody 2002).

Dans ce contexte et pour l'objet de cette étude, la carte forestière et les p-e peuvent être considérées comme deux représentations de la même réalité, bien que de nature différente. Si l'on considère que les attributs notés dans les deux sont exacts, la comparaison entre les catégories des strates de la carte et celles des p-e peut donner de l'information sur la concordance qui, pour une même variable, existe entre les deux jeux de données. De plus, l'extension de l'analyse aux strates voisines (de celle sur laquelle la p-e est établie) peut aider à mesurer l'influence que l'incertitude positionnelle des p-e et la proximité des frontières exercent sur la concordance entre les deux représentations. Ces informations seront utiles pour déterminer quelles strates auront les estimations du volume de matière ligneuse les plus incertaines en raison d'un lien plus faible entre les p-e à partir desquelles les estimations ont été faites et les conditions auxquelles elles seront appliquées.

L'objectif général de ce chapitre consiste à évaluer, pour les strates identifiées par photo-interprétation, le degré de concordance entre les catégories de certaines variables de la carte forestière et des p-e. L'hypothèse de base est la suivante : les strates de la carte forestière diffèrent par rapport à leur concordance avec les données de terrain pour des variables

comme la composition d'essences ou la hauteur des arbres, et ces différences sont associées à l'évolution de la végétation et aux perturbations.

Les objectifs spécifiques sont :

1. D'estimer pour les strates de la carte forestière la concordance avec les données de terrain par rapport aux variables les plus importantes;
2. D'identifier les strates photo-interprétées dont la concordance locale carte/p-e est élevée et celles dont la concordance est élevée avec les strates voisines;
3. D'élaborer un gradient de concordance locale carte/p-e pour les strates cartographiques les plus importantes sur le territoire étudié;
4. D'identifier les combinaisons de variables qui sont plus consistantes lors d'un changement d'échelle et de procédé de collecte de l'information.

## **Méthodologie**

### **1. Données et variables**

Les deux sources de données utilisées dans cette étude sont (1) la carte forestière de la Forêt Montmorency produite en 1992 et (2) les données des placettes d'échantillonnage (p-e) forestier temporaires de 1992 (342 p-e) et celles d'un réseau de p-e permanentes mesurées entre 1995 et 1999 (201 p-e). Sur l'ensemble des p-e, 539 ont été retenues pour l'analyse, car elles présentent toute l'information pour les variables examinées.

Au Québec, les strates forestières sont identifiées par des appellations issues de la combinaison de plusieurs variables, dont la plupart sont qualitatives (Ministère des Ressources naturelles du Québec 1995). Parmi ces variables, les plus importantes sont sélectionnées pour cette étude : *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*. Les détails des catégories de ces variables ont été décrits au Chapitre II.

Pour revenir à l'appellation des strates forestières, l'appellation SS B3 45 el, par exemple, fait référence à une strate résineuse dominée par le sapin baumier (catégorie SS de *GE*), où la végétation forestière couvre entre 61 et 80 % de la surface des polygones (catégorie B de *CD*), la hauteur des arbres dominants et codominants varie entre 12 et 17 m (catégorie 3 de *CH*), l'âge de la végétation est estimé à 45 ans (catégorie 45 de *CA*) et une épidémie légère est la perturbation qui a affecté la végétation (catégorie el de *Perturbation moyenne*).

La variable *Classe d'âge* est classifiée différemment sur la carte forestière et sur le terrain. Les catégories originales des p-e doivent donc être modifiées pour les ajuster à celles de la carte (Tableau 34).

Tableau 34. Catégories de la variable *Classe d'âge (CA)* sur la carte forestière et sur les p-e, ainsi que celles utilisées dans cette étude comme catégories des p-e.

Catégorie carte	Catégorie originale p-e	Catégorie utilisée
0	1	0
15	10	15
30	30	30
45	50	45
60	60	60
75	70	75
90	90	90

## 2. Méthodes

La base méthodologique de cette étude est la comparaison, à un même endroit, de la catégorie d'une variable estimée par photo-interprétation avec celle notée sur le terrain (Figure 37). Pour effectuer cette comparaison, les strates forestières de la carte de 1992 sont employées, car que l'échantillonnage duquel proviennent les p-e a été défini à partir de cette carte. Chaque p-e est associée à la strate forestière correspondant au polygone où elle est localisée sur la carte de 1992.

Étant donné que la carte forestière contient plus d'une centaine de strates, seules les plus importantes ont été retenues pour l'étude. Les trois critères qui déterminent l'importance

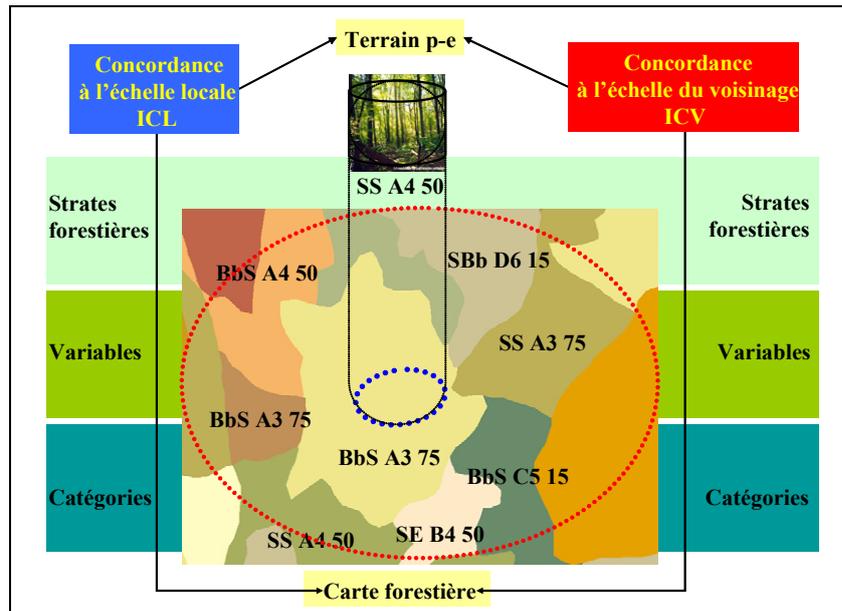


Figure 37. Étapes méthodologiques du Chapitre VI.

d'une strate sont : (1) la superficie qu'elle occupe à la Forêt Montmorency, (2) le nombre de polygones correspondants sur la carte et (3) le nombre de p-e qui la sondent. Parmi les 365 strates de la carte forestière de 1992, les 45 strates qui comptent au moins 10 polygones et trois p-e sont examinées dans cette étude. La Figure 38 montre les 20 strates les plus importantes selon les trois critères énumérés.

L'analyse suivante consiste à quantifier, pour les 45 strates retenues, la concordance carte/p-e pour chaque variable et pour les combinaisons de 2, 3 et 4 variables. Finalement, l'analyse de la variation de la concordance carte/p-e en fonction de l'incertitude spatiale des p-e utilisées, vise à établir l'influence de cette dernière sur les valeurs de concordance.

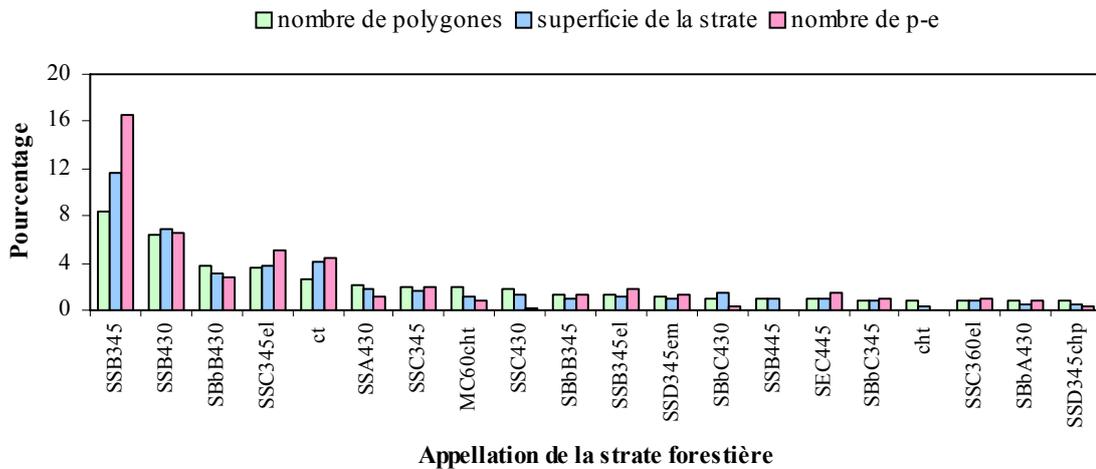


Figure 38. Les 20 strates les plus importantes en superficie, nombre de polygones et de p-e, sur la carte forestière de la Forêt Montmorency.

### 2.1. Quantification de la concordance carte/p-e pour les strates de la carte forestière

La concordance carte/p-e des 45 strates forestières sélectionnées est estimée pour les 15 combinaisons possibles des variables qualitatives *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*. Pour ce faire, la coïncidence entre la catégorie de la variable dans la carte forestière et dans les p-e associées à chaque strate est mesurée à l'aide d'un indice de concordance, tel que défini ci-dessous. Les quinze combinaisons sont : *GE*, *CD*, *CH*, *CA*, *GE-CD*, *GE-CH*, *GE-CA*, *CD-CH*, *CD-CA*, *CH-CA*, *GE-CD-CH*, *GE-CH-CA*, *GE-CD-CA*, *CD-CH-CA* et *GE-CD-CH-CA*. Cet indice est estimé aux échelles locale et de voisinage spatial.

L'échelle locale est définie comme la strate de la carte sur laquelle chaque p-e est localisée (le polygone où se situe chaque p-e). Le voisinage spatial est défini, dans cette étude, comme l'ensemble des polygones de la carte qui intersectent un rayon de 100 m autour de chaque p-e, pourvu qu'ils correspondent à des strates forestières différentes de celle du polygone qui contient la p-e. Étant donné que la taille et la forme des polygones sont variables, le rayon a été sélectionné de manière à inclure les strates adjacentes à celle qui contient la p-e (Figure 39).

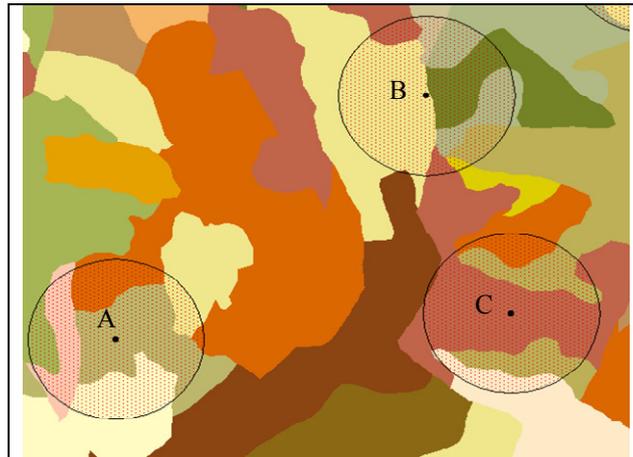


Figure 39. Échelle de voisinage spatial. Les points noirs sont les p-e (A, B et C, dans cet exemple). Les cercles indiquent le voisinage spatial de chaque p-e sur la carte forestière de 1992. Les couleurs des polygones représentent les strates forestières.

### 2.1.1. Indice de concordance à l'échelle locale

L'Indice de concordance locale (ICL) d'une strate pour une combinaison, correspondant à une variable, est dérivé du « Pourcentage de concordance des catégories », qui est un coefficient issu des matrices de confusion qui évalue l'incertitude des classifications (Foody 2002, Stehman 1997, 1998). L'ICL se calcul comme suit :

$$ICL_S = \frac{\sum_{i=1}^{N_S} c_i}{N_S} \text{ dans lequel } \begin{cases} c_i = 1 \rightarrow a_p = a_m \\ c_i = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.1}$$

où  $S$  est la strate forestière,  $N_S$  est le nombre de p-e associées à  $S$ ,  $c_i$  est une mesure de concordance entre la p-e  $i$  et la carte,  $a_p$  est la catégorie de la variable  $a$  sur la p-e  $i$ , alors que  $a_m$  est la catégorie de la variable  $a$  sur la carte.

Les valeurs d'ICL varient entre 0 et 1. La valeur 0 indique que les catégories de toutes les p-e de la strate forestière sont différentes de celle de la carte. La valeur 1 indique, pour sa part, que toutes les p-e de la strate présentent la même catégorie que la carte et sont interprétés comme ayant une concordance complète.

Pour une combinaison de deux variables,  $a$  et  $b$ , l'ICL est calculé selon la formule suivante :

$$ICL_S = \frac{\sum_{i=1}^{N_S} c_i}{N_S} \text{ dans lequel } \begin{cases} c_i = 1 \rightarrow a_p = a_m \cap b_p = b_m \\ c_i = 0 \rightarrow a_p \neq a_m \cup b_p \neq b_m \end{cases}, \quad \text{VI.2}$$

où  $S$  est la strate forestière,  $N_S$  est le nombre de p-e associées à  $S$ ,  $c_i$  est une mesure de concordance entre la p-e  $i$  et la carte,  $a_p$  et  $b_p$  sont les catégories des variables  $a$  et  $b$  sur la p-e  $i$ , et  $a_m$  et  $b_m$  sont les catégories des variables  $a$  et  $b$  sur la carte. Pour la combinaison de trois et de quatre variables, l'ICL est calculé de manière similaire.

### 2.1.2. Indice de concordance à l'échelle du voisinage spatial

Un Indice de concordance de voisinage (ICV) est élaboré pour cette étude. L'ICV est dérivé du coefficient « Pourcentage de concordance des catégories » issu des matrices de confusion (Foody 2002, Stehman 1997, 1998). Pour une strate forestière particulière, l'ICV d'une combinaison à une variable permet de considérer la concordance entre les catégories des p-e qui appartiennent à la strate et les catégories des strates du voisinage spatial. Cet indice se calcule de la façon suivante :

$$ICV_S = \frac{\sum_{i=1}^{N_S} c_i}{\sum_{i=1}^{N_S} V_i} \text{ où } c_i = \sum_{j=1}^{V_i} r_j \text{ dans lequel } \begin{cases} r_j = 1 \rightarrow a_p = a_m \\ r_j = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.3}$$

où  $S$  est la strate cartographique concernée,  $N_S$  est le nombre total de p-e qui appartiennent à  $S$ ,  $c_i$  est la mesure de concordance entre la p-e  $i$  et les strates de son voisinage spatial,  $V_i$  est le nombre de strates du voisinage spatial de la p-e  $i$ ,  $j$  est une strate du voisinage spatial de la p-e  $i$ ,  $r_j$  est la mesure de concordance entre la p-e  $i$  et la strate  $j$ ,  $a_p$  et  $a_m$  sont les catégories de la variable  $a$  de la p-e  $i$  et de la strate  $j$  respectivement.

Pour les combinaisons de deux, trois ou quatre variables, l'ICV est calculé comme suit (l'exemple est bivarié) :

$$ICV_S = \frac{\sum_{i=1}^{N_S} c_i}{\sum_{i=1}^{N_S} V_i} \text{ où } c_i = \sum_{j=1}^{V_i} r_j \text{ dans lequel } \begin{cases} r_j = 1 \rightarrow a_p = a_m \cap b_p = b_m \\ r_j = 0 \rightarrow a_p \neq a_m \cup b_p \neq b_m \end{cases}, \quad \text{VI.4}$$

où  $S$  est la strate cartographique concernée,  $N_S$  est le nombre total de p-e qui appartiennent à  $S$ ,  $c_i$  est la mesure de concordance entre la p-e  $i$  et les strates de son voisinage spatial,  $V_i$  est le nombre de strates du voisinage spatial de la p-e  $i$ ,  $j$  est une strate du voisinage spatial de la p-e  $i$ ,  $r_j$  est la mesure de concordance entre la p-e  $i$  et la strate  $j$ ,  $a_p$  et  $b_p$  sont les catégories des variables  $a$  et  $b$  de la p-e  $i$ , et  $a_m$  et  $b_m$  sont les catégories des variables  $a$  et  $b$  de la strate  $j$ .

Comme pour l'ICL, les valeurs d'ICV varient entre 0 et 1. La valeur 0 signifie qu'aucune des p-e de la strate forestière n'appartient à la même catégorie que les strates voisines. La valeur 1 indique, pour sa part, que toutes les p-e de la strate forestière appartiennent à la même catégorie que les strates voisines.

## 2.2. Quantification de la concordance carte/p-e pour les combinaisons de variables

Tout comme la concordance est quantifiée pour les strates forestières, il est pertinent de la quantifier pour les quinze combinaisons de variables, toutes strates confondues, et ce, aux deux échelles. Les indices de concordance sont similaires aux indices ICL et ICV présentés à la section précédente, les différences concernent les données qui sont prises en compte.

Pour une combinaison univariée, l'indice de concordance locale se calcule comme suit :

$$ICL_C = \frac{\sum_{i=1}^N c_i}{N} \text{ dans lequel } \begin{cases} c_i = 1 \rightarrow a_p = a_m \\ c_i = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.5}$$

où  $C$  est une combinaison de variables,  $N$  est le nombre de p-e (539 p-e),  $i$  est une p-e,  $c_i$  est une mesure de concordance entre la p-e  $i$  et la carte,  $a_p$  est la catégorie de la variable  $a$  sur la p-e  $i$ , alors que  $a_m$  est la catégorie de la variable  $a$  sur la carte.

Pour une combinaison bivariée, l'indice de concordance locale se calcule comme suit :

$$ICL_C = \frac{\sum_{i=1}^N c_i}{N} \text{ dans lequel } \begin{cases} c_i = 1 \rightarrow a_p = a_m \cap b_p = b_m \\ c_i = 0 \rightarrow a_p \neq a_m \cup b_p \neq b_m \end{cases}, \quad \text{VI.6}$$

où  $C$  est une combinaison de variables,  $N$  est le nombre de p-e (539 p-e),  $i$  est une p-e,  $c_i$  est une mesure de concordance entre la p-e  $i$  et la carte,  $a_p$  et  $b_p$  sont les catégories des variables  $a$  et  $b$  sur la p-e  $i$  et  $a_m$  et  $b_m$  sont les catégories des variables  $a$  et  $b$  sur la carte. Pour la combinaison de trois et de quatre variables, l' $ICL_C$  est calculé de manière similaire.

À l'échelle du voisinage, l'indice de concordance pour une combinaison univariée est le suivant :

$$ICV_C = \frac{\sum_{i=1}^N c_i}{N} \text{ où } c_i = \frac{\sum_{j=1}^{V_i} r_j}{V_i} \text{ dans lequel } \begin{cases} r_j = 1 \rightarrow a_p = a_m \\ r_j = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.7}$$

où  $C$  est la combinaison de variables,  $i$  est une p-e,  $N$  est le nombre total de p-e,  $c_i$  est la mesure de concordance entre la p-e  $i$  et les strates de son voisinage spatial,  $r_j$  est la mesure de concordance entre la p-e  $i$  et la strate  $j$ ,  $V_i$  est le nombre de strates du voisinage spatial de la p-e  $i$ ,  $a_p$  et  $a_m$  sont définies de la même façon que pour l' $ICL_C$ .

Pour les combinaisons de deux, trois ou quatre variables, l' $ICV$  est calculé comme suit (l'exemple est bivarié) :

$$ICV_C = \frac{\sum_{i=1}^N c_i}{N} \text{ où } c_i = \frac{\sum_{j=1}^{V_i} r_j}{V_i} \text{ dans lequel } \begin{cases} r_j = 1 \rightarrow a_p = a_m \cap b_p = b_m \\ r_j = 0 \rightarrow a_p \neq a_m \cup b_p \neq b_m \end{cases}, \quad \text{VI.8}$$

où  $C$  est la combinaison de variables concernée,  $N$  est le nombre total de p-e,  $c_i$  est la mesure de concordance entre la p-e  $i$  et les strates de son voisinage spatial,  $V_i$  est le nombre de strates du voisinage spatial de la p-e  $i$ ,  $r_j$  est la mesure de concordance entre la p-e  $i$  et la strate  $j$ ,  $a_p$  et  $b_p$  ainsi que  $a_m$  et  $b_m$  sont définies de la même façon que pour  $ICL_C$  bivariée.

Les indices de concordance aux deux échelles s'étendent entre 0 et 1 comme c'est le cas des indices des strates forestières.

### 2.3. Quantification de la concordance carte/p-e pour les catégories des variables

La concordance carte/p-e pour les catégories des variables *Groupement d'essences*, *Classe de densité*, *Classe de hauteur* et *Classe d'âge* est quantifiée, aux deux échelles, en utilisant des indices similaires à l'ICL et l'ICV. Les différences concernent les données qui sont prises en compte. À l'échelle locale la concordance se calcule comme suit :

$$ICL_A = \frac{\sum_{i=1}^{N_A} c_i}{N_A} \text{ dans lequel } \begin{cases} c_i = 1 \rightarrow a_p = a_m \\ c_i = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.9}$$

où  $A$  est une catégorie,  $N_A$  est le nombre de p-e qui présentent la catégorie  $A$ ,  $c_i$  est une mesure de concordance entre la p-e  $i$  (qui présente la catégorie  $A$ ) et la carte,  $a_p$  est la catégorie sur la p-e  $i$ , alors que  $a_m$  est la catégorie sur la carte.

Quant à la concordance à l'échelle du voisinage, elle se calcule comme suit :

$$ICV_C = \frac{\sum_{i=1}^{N_A} c_i}{N_A} \text{ où } c_i = \frac{\sum_{j=1}^{V_i} r_j}{V_i} \text{ dans lequel } \begin{cases} r_j = 1 \rightarrow a_p = a_m \\ r_j = 0 \rightarrow a_p \neq a_m \end{cases}, \quad \text{VI.10}$$

où  $A$  est la catégorie concernée,  $N_A$  est le nombre de p-e qui présentent la catégorie  $A$ ,  $c_i$  est la mesure de concordance entre la p-e  $i$  et les strates de son voisinage spatial,  $r_j$  est la mesure de concordance entre la p-e  $i$  et la strate  $j$ ,  $V_i$  est le nombre de strates du voisinage spatial de la p-e  $i$ ,  $a_p$  et  $a_m$  sont les catégories de la p-e  $i$  et de la strate  $j$  respectivement.

Comme les autres indices de concordance, l'ICL et l'ICV pour les catégories varient entre 0 et 1, c'est-à-dire une concordance nulle ou complète, respectivement.

#### 2.4. Relation entre la concordance carte/p-e et l'incertitude spatiale des p-e

L'objectif de l'analyse décrite dans cette section est d'explorer l'effet possible de l'incertitude positionnelle des p-e sur les valeurs de concordance globale carte/p-e et ce, pour chaque variable.

Une matrice de confusion est élaborée pour les variables *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*. Dans la matrice de confusion de chaque variable, on évalue le nombre de coïncidences entre la catégorie de la p-e et celle de la strate où elle est située. Cela signifie que seule l'échelle locale est considérée.

La concordance globale résulte de l'addition des éléments de la diagonale de la matrice de confusion, qui est divisée par le total des éléments de la matrice, tandis que la concordance de chaque catégorie est la proportion de coïncidences entre la catégorie de la carte et celle de la p-e. Un exemple est présenté pour la *Classe de densité* à l'Annexe IV. Le coefficient *kappa* est calculé à partir de chaque matrice de confusion (Chrisman 1991, Friedl *et al.* 2001).

À la différence de la quantification de la concordance pour les catégories, dans cette analyse, la concordance globale et par catégorie ainsi que le coefficient *kappa* sont calculés en considérant des ensembles différents de p-e : (1) la totalité des p-e (539 p-e), (2) les p-e qui sont localisées à une distance de la frontière de la strate qui est supérieure à 5 m (470 p-e), (3) les p-e à plus de 10 m de la frontière (406 p-e), (4) les p-e à plus de 25 m de la frontière (257 p-e) et finalement (5) les p-e à plus de 50 m de la frontière (49 p-e).

Les résultats de cette section indiqueront si le fait d'utiliser des p-e localisées à une distance croissante des frontières des polygones forestiers, c'est-à-dire de plus en plus certaines par rapport à leur appartenance à une strate forestière (en supposant que cette certitude est reliée à la distance croissante aux frontières), influence les valeurs de concordance globale, de concordance par catégorie et du coefficient *kappa*.

## Résultats et discussion

Les résultats concernant les strates forestières sont présentés en premier, suivis de ceux relatifs aux combinaisons de variables et aux catégories des variables. Les résultats présentés en dernier portent sur la relation entre la concordance carte/p-e et l'incertitude spatiale des p-e.

### 1. La concordance des strates forestières

Les résultats de l'analyse des concordances carte/p-e sont présentés en premier pour les combinaisons d'une seule variable (analyse univariée), ensuite pour les combinaisons de deux variables (analyse bivariée) et finalement pour les combinaisons de trois et quatre variables (analyse multivariée).

#### 1.1. Analyse univariée

Le Tableau 35 présente les résultats des indices de concordance (ICL et ICV) pour les variables individuelles, soient *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*.

Les ICL obtenus varient entre 0 et 1 pour toutes les variables, tandis que les ICV des variables *CD* et *CH* ont une valeur minimale légèrement au-dessus de 0 (Tableau 35). Les ICV des quatre variables ont des valeurs maximales inférieures à 1. Ce dernier résultat signifie qu'aucune des strates ne présente une concordance maximale entre la catégorie de la p-e et celles de son voisinage spatial. Par contre, les moyennes et les médianes d'ICL et d'ICV de toutes les variables sont semblables, hormis la médiane de *CA*. Par rapport aux valeurs modales, celles des ICL de toutes les variables sont égales à 0, alors que celles des ICV présentent des valeurs diverses, proches de la moyenne dans le cas de *GE* et de *CA* et proches du minimum dans le cas de *CH*.

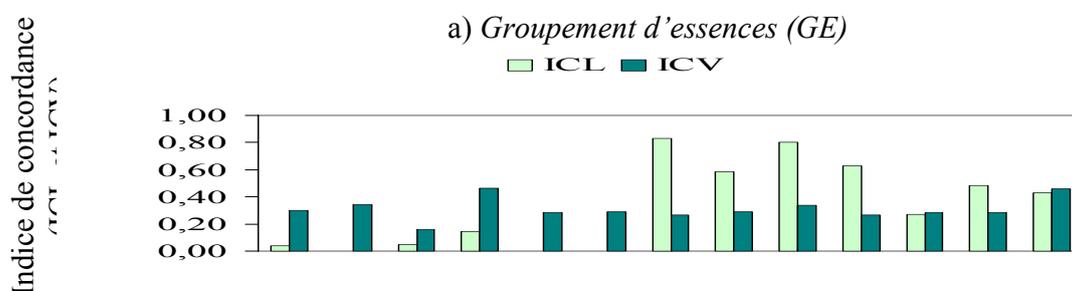
Tableau 35. Valeurs de l'indice de concordance (ICL et ICV) entre la catégorie de la carte et celle de la p-e des variables *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*.

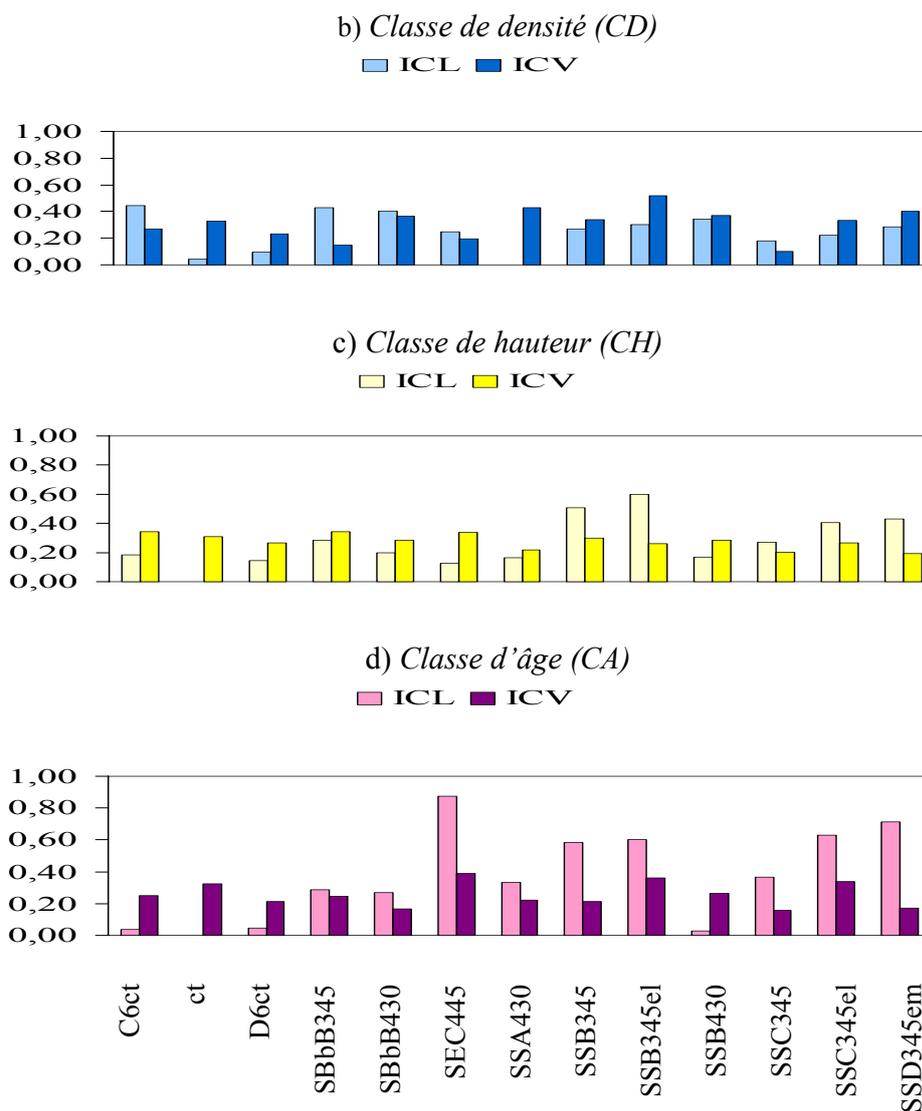
	ICL				ICV			
	GE	CD	CH	CA	GE	CD	CH	CA
Minimum	0,00	0,00	0,00	0,00	0,00	0,05	0,06	0,00
Moyenne	0,32	0,29	0,30	0,23	0,29	0,30	0,28	0,21
Médiane	0,27	0,27	0,25	0,03	0,29	0,31	0,28	0,20
Mode	0,00	0,00	0,00	0,00	0,28	0,15	0,06	0,17
Maximum	1,00	1,00	1,00	1,00	0,51	0,57	0,77	0,53

La Figure 40 présente les ICL et les ICV des dix strates forestières les plus importantes à la Forêt Montmorency, sondées par au moins cinq p-e (les résultats des autres strates sont présentés à l'Annexe IV). Seules trois strates présentent un patron similaire pour les quatre variables. D'une part, la strate résineuse dominée par le sapin baumier et perturbée par une épidémie légère, SS C3 45 el, présente une concordance locale supérieure à la concordance avec le voisinage. D'autre part, des strates correspondant à la végétation en régénération à la suite d'une coupe totale (D6 ct et ct) présentent une forte concordance avec le voisinage, tandis que la concordance locale est faible (Figure 40).

En général, les strates dont la concordance est forte localement affichent une faible concordance à l'échelle de voisinage et vice-versa. Par exemple, en terme de *Groupement d'essences*, les concordances locales des strates résineuses dominées par le sapin baumier (SS) sont supérieures à 0,48, tandis qu'avec le voisinage elles varient entre 0,27 et 0,34. Les strates qui présentent d'autres types de végétation, résineuse ou mélangée (SE et SBb), ou celles dont la végétation est en régénération (comme D6 ct, C6 ct et ct) concordent à l'échelle du voisinage et non à l'échelle locale (Figure 40a).

Il est remarquable qu'à l'exception de quelques strates et de certaines variables, comme le *Groupement d'essences*, les valeurs de concordance locale sont assez faibles, elles ne





#### Strate forestière

Figure 40. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons d'une variable, aux échelles locales (ICL) et du voisinage spatial (ICV).

dépassent pas 0,60 (c'est-à-dire que 60 % des cas présentent la même catégorie sur la carte et sur le terrain), et les valeurs nulles sont fréquentes. La concordance avec le voisinage spatiale n'améliore pas cette situation, car elle varie entre 0,10 et 0,46.

## 1.2. Analyse bivariée

La valeur minimale de deux indices est la même, tandis que les valeurs maximales sont très similaires pour toutes les combinaisons bivariées (Tableau 36). Cependant les maximums des ICL sont, en général, deux fois supérieurs à ceux des ICV. Quant aux moyennes des deux indices, elles sont assez semblables pour toutes les combinaisons. Les médianes d'ICL sont égales à 0, alors que celles d'ICV varient entre 0,07 et 0,13. Les valeurs modales des deux indices pour toutes les combinaisons sont égales à 0.

Les résultats des analyses bivariées des strates les plus importantes à la Forêt Montmorency sont semblables pour certaines caractéristiques à ceux de l'analyse univariée (Figure 41). En général, la concordance locale ne dépasse pas la valeur 0,50, c'est-à-dire que pour la majorité des strates les plus représentatives de la zone d'étude, moins du 50 % des p-e concordent avec la carte forestière.

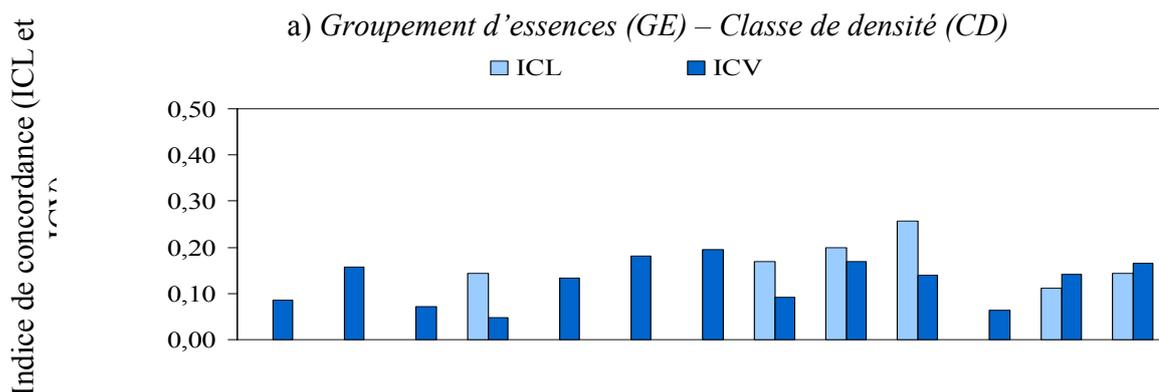
Tableau 36. Indices de concordance carte/p-e des strates forestières pour les combinaisons de deux variables de *Groupe ment d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*, aux échelles locales (ICL) et de voisinage spatial (ICV).

ICL	<i>GE-CD</i>	<i>GE-CH</i>	<i>GE-CA</i>	<i>CD-CH</i>	<i>CD-CA</i>	<i>CH-CA</i>
Minimum	0,00	0,00	0,00	0,00	0,00	0,00
Moyenne	0,10	0,15	0,09	0,08	0,06	0,13
Médiane	0,00	0,00	0,00	0,00	0,00	0,00
Mode	0,00	0,00	0,00	0,00	0,00	0,00
Maximum	0,67	0,67	0,67	0,67	0,40	0,50
<b>ICV</b>						
Minimum	0,00	0,00	0,00	0,00	0,00	0,00
Moyenne	0,11	0,10	0,08	0,13	0,12	0,12
Médiane	0,09	0,10	0,07	0,13	0,11	0,11
Mode	0,00	0,00	0,00	0,00	0,00	0,00
Maximum	0,38	0,34	0,23	0,33	0,33	0,33

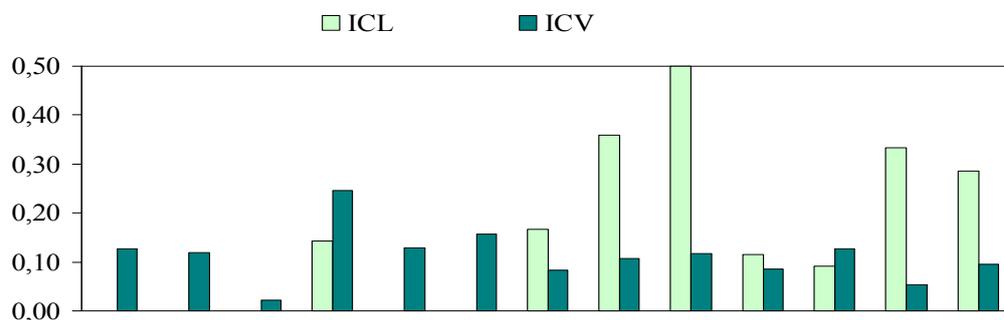
Seules deux strates présentent un patron constant de concordance pour toutes les combinaisons de variables. La première, soit la strate correspondant à la végétation en phase de régénération (ct), concorde seulement avec le voisinage. La deuxième, soit la strate résineuse dominée par le sapin baumier à couvert ligneux moyen, dont la hauteur des

arbres varie entre 12 et 17 m (SS B3 45), présente une plus forte concordance à l'échelle locale qu'à celle du voisinage. En effet, cette strate, qui est la plus importante de la Forêt Montmorency (Figure 38), obtient les valeurs les plus élevées de concordance locale pour plusieurs combinaisons (Figure 41).

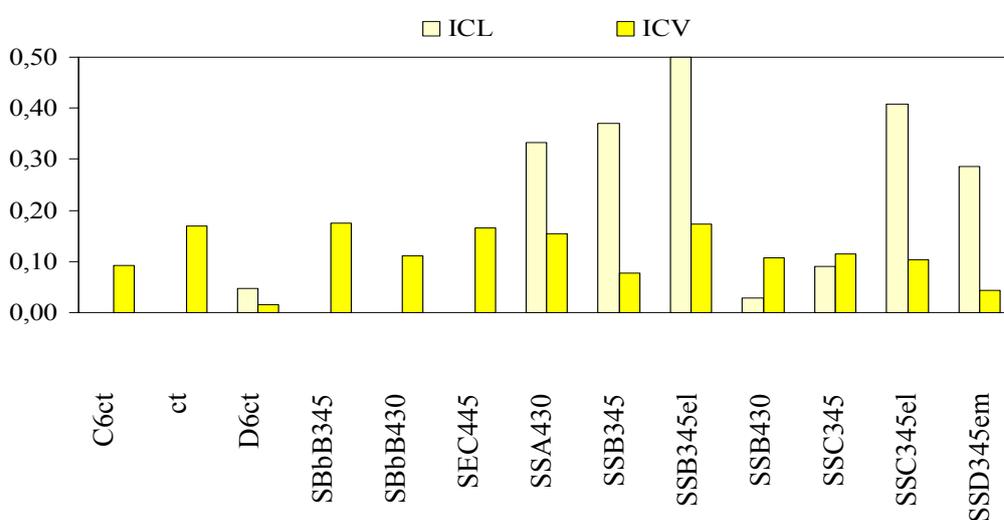
Les strates résineuses dominées par le sapin baumier (SS) présentent, en général, une concordance locale supérieure à celle du voisinage pour les combinaisons dans lesquelles interviennent la variable *Groupement d'essences* ainsi que la combinaison *Classe de hauteur-Classe d'âge* (Figure 41a, b, c et f). Pour les combinaisons où la *Classe de densité* intervient, la concordance à l'échelle du voisinage est supérieure, en général, à celle de l'échelle locale (Figure 41a, d et e). La concordance locale des strates résineuses autres que celles dominées par le sapin baumier ou mélangées (SE ou SBb) est nulle, tandis qu'à l'échelle du voisinage la concordance est présente pour les combinaisons où le *Groupement d'essences* intervient (Figure 41a, b et c). Pour les autres combinaisons, ces strates n'affichent pas de patron clair.



## b) Groupement d'essences (GE) – Classe de hauteur (CH)



## c) Groupement d'essences (GE) – Classe d'âge (CA)

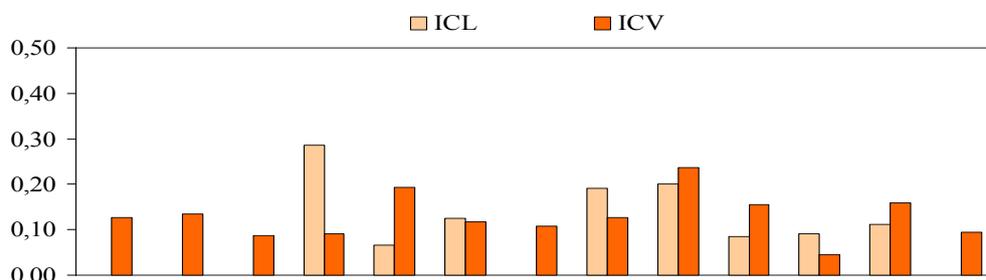


## Strate forestière

Figure 41. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons de deux variables aux échelles locales (ICL) et du voisinage spatial (ICV).

Indice de concordance (ICL et ICV)

## d) Classe de densité (CD) – Classe de hauteur (CH)



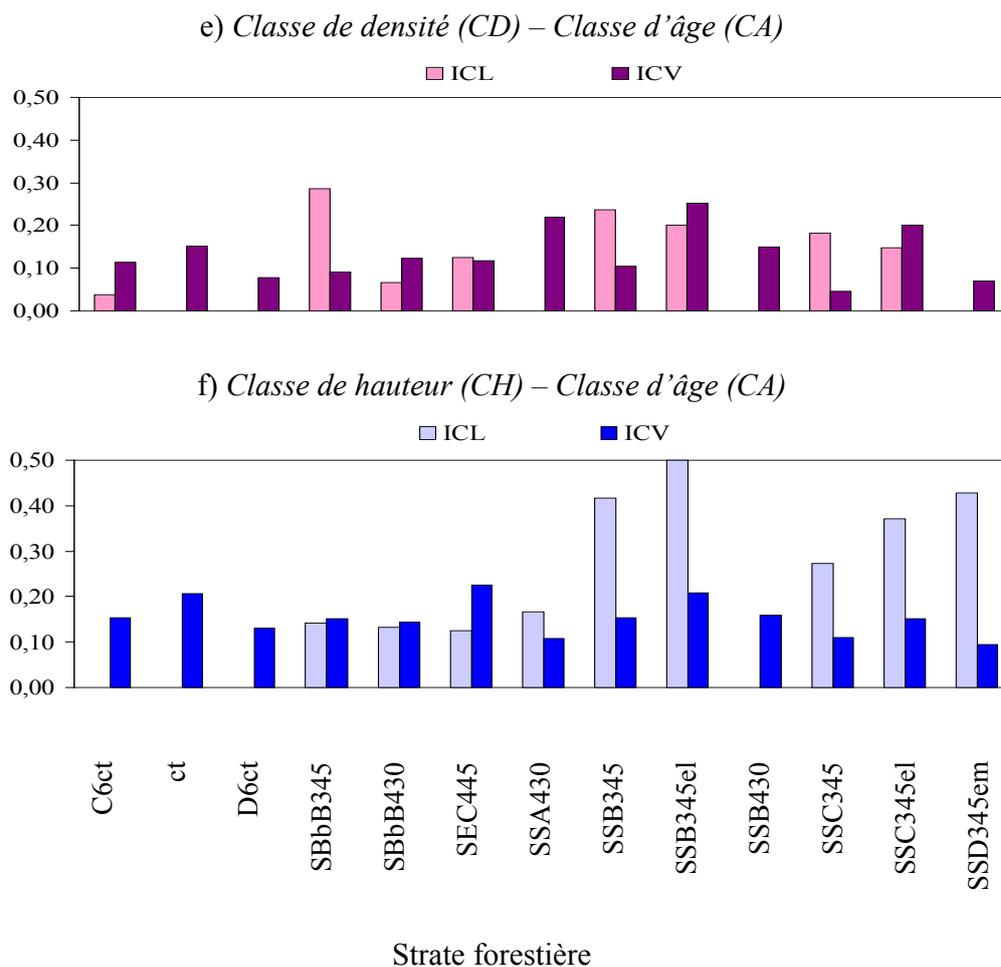


Figure 42. Suite. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons de deux variables aux échelles locales (ICL) et du voisinage spatial (ICV).

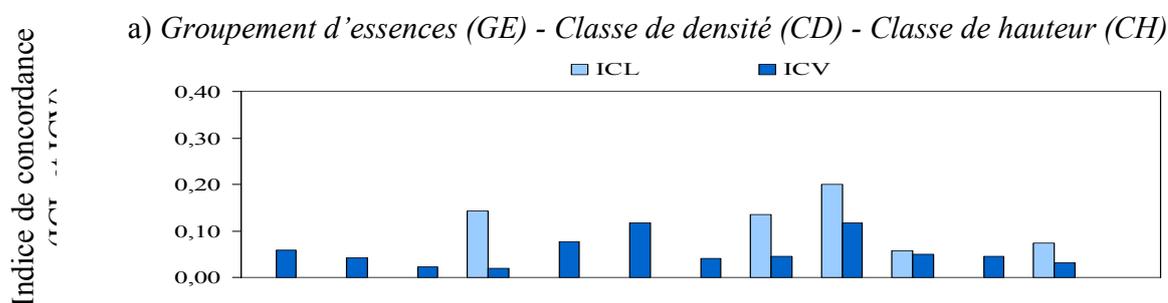
### 1.3. Analyses multivariées

La plupart des strates présentent des valeurs de concordance très faibles aux échelles locales (ICL) et du voisinage spatial (ICV) pour les combinaisons de trois variables (Tableau 37). Les valeurs minimales et modales d'ICL et d'ICV sont égales à 0, de même que la médiane d'ICL. Les moyennes sont similaires, variant entre 0,02 et 0,08. Ces valeurs sont similaires à celles des médianes d'ICV. Les maximums sont plus élevés à l'échelle locale que celles du voisinage, variant entre 0,33 et 0,50.

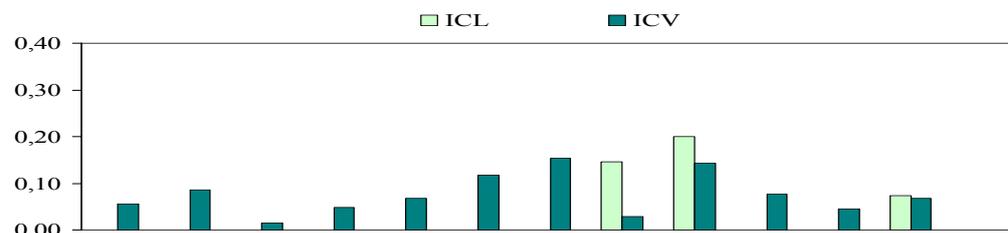
Tableau 37. Indices de concordance carte/p-e des strates forestières pour les combinaisons de trois variables parmi Groupement d'essences (GE), Classe de densité (CD), Classe de hauteur (CH) et Classe d'âge (CA), aux échelles locales (ICL) et du voisinage spatial (ICV).

ICL	<i>GE-CD-CH</i>	<i>GE-CD-CA</i>	<i>GE-CH-CA</i>	<i>CD-CH-CA</i>
Minimum	0,00	0,00	0,00	0,00
Moyenne	0,04	0,02	0,07	0,04
Médiane	0,00	0,00	0,00	0,00
Mode	0,00	0,00	0,00	0,00
Maximum	0,33	0,33	0,50	0,33
ICV				
Minimum	0,00	0,00	0,00	0,00
Moyenne	0,04	0,04	0,05	0,08
Médiane	0,04	0,02	0,04	0,07
Mode	0,00	0,00	0,00	0,00
Maximum	0,27	0,23	0,15	0,33

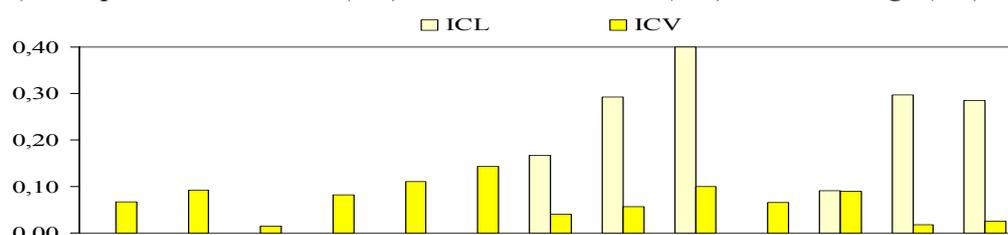
Parmi les strates les plus importantes de la Forêt Montmorency, cinq présentent un patron constant pour toutes les combinaisons de variables (Figure 42). Les strates dont la végétation est en régénération (C6 ct, ct et D6 ct) présentent une concordance seulement à l'échelle du voisinage. Par contre, deux des strates dominées par le sapin baumier, dont la végétation est dans une étape intermédiaire du développement (SS B3 45 et SS B3 45 el) concordent plus fortement à l'échelle locale qu'à celle du voisinage. Cependant, aucune des strates ne présente de concordance locale carte/p-e qui excède 40 % des cas. De plus, pour la combinaison *Groupement d'essences-Classe de densité-Classe d'âge*, la plupart des strates concordent seulement avec le voisinage, la concordance locale étant nulle (Figure 42b).



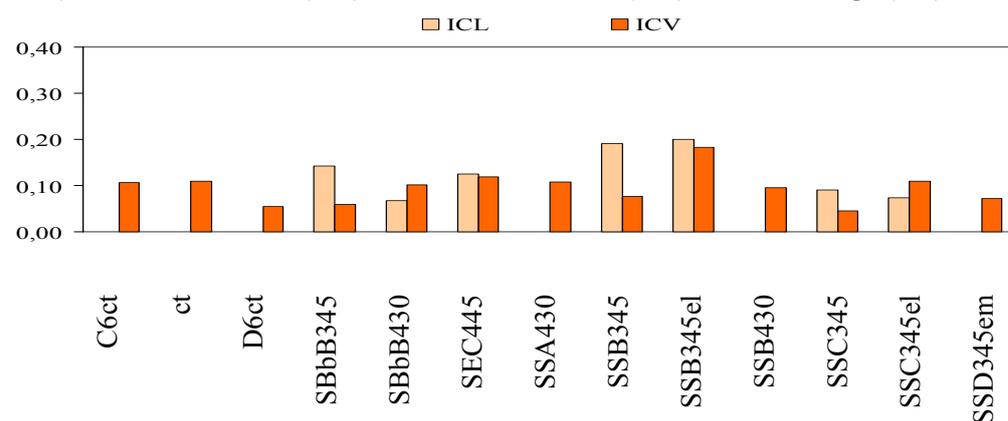
b) Groupement d'essences (GE) - Classe de densité (CD) - Classe d'âge (CA)



c) Groupement d'essences (GE) - Classe de hauteur (CH) - Classe d'âge (CA)



d) Classe de densité (CD) - Classe de hauteur (CH) - Classe d'âge (CA)



Strate forestière

Figure 42. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour les combinaisons de trois variables aux échelles locales (ICL) et du voisinage spatial (ICV).

Comme aux analyses précédentes, plusieurs strates dont la concordance locale est faible ou nulle présentent une concordance à l'échelle du voisinage plus élevée ou vice-versa. Certaines strates présentent également des valeurs similaires et faibles pour les deux échelles.

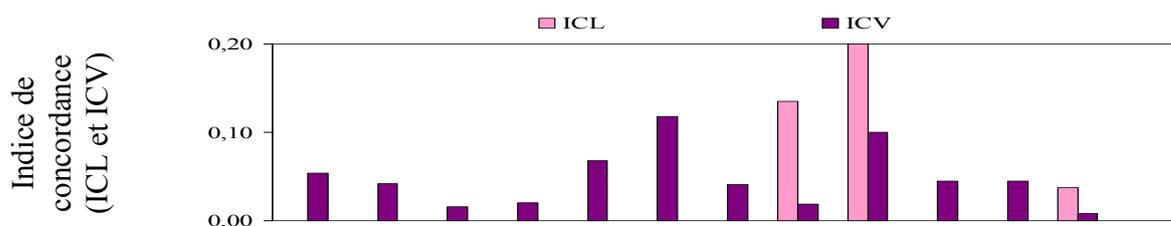
Finalement, pour la combinaison des quatre variables, la plupart des strates forestières présentent une concordance carte/p-e nulle, autant à l'échelle locale (ICL) qu'à l'échelle du voisinage spatial (ICV) (Tableau 38). Même les maximums sont faibles, celui de l'échelle locale (ICL) étant supérieur à celui du voisinage spatial (ICV).

Tableau 38. Indices de concordance carte/p-e des strates forestières pour la combinaison de quatre variables, *Groupe ment d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)*, aux échelles locales (ICL) et du voisinage spatial (ICV).

	ICL	ICV
Minimum	0,00	0,00
Moyenne	0,01	0,03
Médiane	0,00	0,00
Mode	0,00	0,00
Maximum	0,20	0,15

Seules trois strates concordent localement pour les quatre variables (Figure 43). Ces strates sont dominées par le sapin baumier d'une hauteur qui varie entre 12 et 17 m (SS B3 45, SS B3 45 el et SS C3 45 el). Cependant, la concordance locale maximale atteint une valeur de 0,20, ce qui indique que seulement 20 % des p-e concorde avec la carte forestière. Les autres strates concordent faiblement avec le voisinage spatial, la plupart des ICV étant inférieurs à 0,10.

La comparaison entre les valeurs de concordance locale et celles du voisinage permet de repérer certaines différences entre les deux échelles : (1) lorsque les variables individuelles sont considérées, les concordances minimales et modales du voisinage spatial sont supérieures aux concordances locales, (2) la concordance maximale à l'échelle du voisinage spatial est environ 50 % moins élevée que la concordance locale et (3) les concordances moyennes et médianes sont similaires pour les deux échelles. Ces différences s'affaiblissent



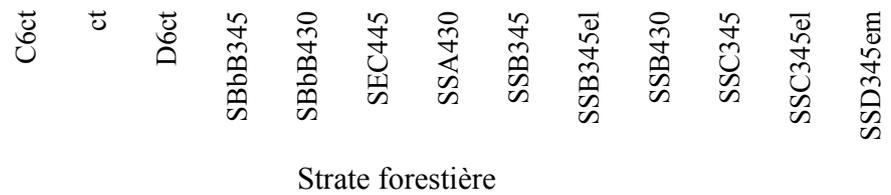


Figure 43. Indice de concordance carte/p-e des strates forestières les plus importantes à la Forêt Montmorency pour la combinaison de quatre variables, aux échelles locales (ICL) et du voisinage spatial (ICV).

à mesure que le nombre de variables considérées dans les combinaisons augmente, c'est-à-dire à mesure que l'indice de concordance est plus restrictif. Une explication possible est la très faible diversité des types de végétation à la Forêt Montmorency. En effet, l'examen de la Figure 38 indique que la majorité des strates les plus importantes en termes de superficie et de nombre de polygones/peuplements correspondent à deux groupements d'essences (SS et SBb) ainsi qu'à deux classes consécutives de densité (B et C), de hauteur (3 et 4) et d'âge (30 et 45). Cette homogénéité des conditions forestières a été observée au Chapitre III au sujet des frontières « sans contraste », correspondant aux peuplements adjacents qui présentent la même catégorie pour les variables. Il faut retenir que ces frontières sont très fréquentes, 12 % lorsque les quatre variables sont considérées. Il est donc très possible que si la catégorie de la p-e ne concorde pas avec celle de la strate où elle se situe, elle concorde avec les catégories des strates voisines. À mesure que l'indice devient plus restrictif les concordances aux deux échelles deviennent plus faibles.

La localisation des p-e par rapport aux frontières du polygone constitue une autre explication possible pour la concordance à l'échelle du voisinage. En effet, pour les strates ct et C6ct la distance minimale entre les p-e et les frontières des polygones varie entre 0,95 et 1,8 m, et entre 18 et 25 % des p-e sont situées à moins de 5 m de la frontière. Pour les strates qui présentent une forte concordance locale, les distances minimales observées sont supérieures à 9 m. Ces informations suggèrent que la concordance avec le voisinage peut être reliée à la proximité des p-e par rapport aux frontières. Cette proximité implique un risque accru de confusion spatiale en terme de l'appartenance des p-e à la strate où elles sont situées. Cette incertitude a deux sources possibles : (1) une erreur de localisation de la p-e et/ou de la frontière de la strate (plus précisément du polygone correspondant à la

strate) ou (2) l'évidence que les frontières nettes et minces de la carte forestière ne correspondent aux frontières forestières réelles qui s'étendent dans l'espace (Forman 1995). Il est probable que le photo-interprète soit influencé par ce dernier fait lors de la définition d'une frontière cartographique. Les deux causes sont possibles et elles expliqueraient les valeurs de concordance observées à cette échelle.

D'une façon générale et en lien avec l'objectif de ce chapitre, la considération de la concordance carte/p-e aux deux échelles permet de reconnaître quatre situations : (1) les strates dont la concordance locale est élevée, tandis que la concordance avec le voisinage est faible, (2) les strates dont la concordance locale est faible, tandis que la concordance avec le voisinage est élevée, (3) les strates dont la concordance est moyenne aux deux échelles et (4) les strates dont la concordance est faible aux deux échelles.

La situation où la concordance est forte à l'échelle locale, mais faible à l'échelle du voisinage, peut indiquer que le degré d'homogénéité interne de la strate est tel que les p-e qui l'échantillonnent sont représentatives de ses caractéristiques. Les estimations du volume de matière ligneuse seront dans ce cas plus certaines, car les conditions représentées dans les p-e sont assez étendues spatialement et, en conséquence, une variabilité spatiale moins élevée s'ajoutera à la variabilité intrinsèque des estimations (c'est le cas de SS B3 45 par exemple, Figure 40). Cette situation prédomine à la Forêt Montmorency, car les strates qui la présentent couvrent environ 15 % de la surface.

Quant à la deuxième situation, où seule la concordance avec le voisinage est forte, elle indique que les p-e représentent mal les caractéristiques de la strate, probablement à cause d'une incertitude élevée par rapport à l'appartenance des p-e à la strate considérée. Une autre cause probable est que la variabilité spatiale des conditions forestières de la strate soit élevée et que le nombre de p-e soit insuffisant pour bien représenter cette diversité. Dans ce cas, les estimations du volume seront très incertaines, car une variabilité spatiale qui n'est pas saisie par les p-e doit être ajoutée à la variabilité intrinsèque des estimations. Cette situation, qui est caractéristique de la végétation en régénération, est la deuxième en importance en terme de surface concernée, soit environ 10 % de la Forêt Montmorency.

La troisième situation, où la concordance est moyenne aux deux échelles, indique que les p-e représentent faiblement les conditions forestières, autant de la strate où elles se situent que des strates voisines. Il existe probablement un certain degré de confusion spatiale en terme de l'appartenance des p-e à la strate. Cependant, il est également possible que les strates voisines se ressemblent et qu'elles diffèrent pour une ou deux variables. Dans ce cas, la certitude des estimations du volume serait intermédiaire entre celles des deux situations précédentes et similaire à celle des strates voisines, surtout si les concordances aux deux échelles concernent les mêmes variables.

Enfin, la quatrième situation implique les estimations sur le volume ligneux les plus incertaines, étant donné que les p-e ne représentent pas les conditions forestières de la strate et qu'une variabilité spatiale élevée s'ajoute à la variance propre des estimations.

## 2. La concordance des combinaisons de variables

En se concentrant sur les combinaisons de variables, on observe que les valeurs d'ICL et d'ICV montrent des étendues similaires et assez réduites pour chaque échelle spatiale (Tableau 39). Les valeurs les plus élevées n'excèdent pas 34 % de concordance, alors que les plus faibles correspondent à 3 %.

Les différences des indices sont faibles pour les combinaisons de variables aux deux échelles. De plus, les indices diminuent à mesure que la combinaison implique un plus grand nombre de variables. Pour les combinaisons d'une variable, ils varient entre 0,22 et 0,34. *Groupe ment d'essences* et *Classe de hauteur* présentent les concordances locale et du voisinage spatial les plus élevées, tandis que *Classe de densité* présente la concordance la plus élevée pour le voisinage (Tableau 39).

Tableau 39. Indices de concordance carte/p-e des combinaisons de variables, toutes strates confondues, aux échelles locales (ICL) et de voisinage (ICV).

Variables et combinaisons de variables	ICL	ICV
<i>GE</i>	0,33	0,29
<i>CD</i>	0,28	0,30
<i>CH</i>	0,34	0,29

<i>CA</i>	0,27	0,22
<i>GE-CD</i>	0,11	0,11
<i>GE-CH</i>	0,16	0,10
<i>GE-CA</i>	0,13	0,08
<i>CD-CH</i>	0,10	0,12
<i>CD-CA</i>	0,10	0,11
<i>CH-CA</i>	0,16	0,13
<i>GE-CD-CH</i>	0,05	0,05
<i>GE-CD-CA</i>	0,04	0,04
<i>GE-CH-CA</i>	0,09	0,05
<i>CD-CH-CA</i>	0,06	0,08
<i>GE-CD-CH-CA</i>	0,03	0,03

Les indices de concordance diminuent à mesure que le nombre de variables dans la combinaison augmente. Également, les différences de concordance entre les deux échelles sont peu marquées. Il faut souligner que *Classe de hauteur* est la seule variable qui intervient dans les combinaisons dont l'ICL ou l'ICV est élevé.

Ces résultats laissent croire que les variables sont assez similaires en terme de consistance lorsque l'échelle d'analyse ou le procédé de collecte de l'information changent.

### 3. La concordance des catégories

À la différence des variables, les catégories affichent des différences marquées en terme de concordance. En effet, pour le *Groupement d'essences (GE)*, c'est la végétation résineuse dominée par le sapin baumier (SS) qui montre la concordance locale la plus élevée avec 54 % des p-e qui coïncident avec la carte forestière (Tableau 40). Les catégories résineuses où intervient l'épinette (EE, ES et SE) affichent une concordance locale nulle et une très faible concordance avec le voisinage.

Tableau 40. Indices de concordance carte/p-e des catégories de la variable *Groupement d'essences (GE)*, aux échelles locales (ICL) et de voisinage (ICV).

Catégorie	ICL	ICV
BbS	0,06	0,08
EE	0,00	0,02
ES	0,00	0,00
SBb	0,09	0,19

SE	0,00	0,02
SS	0,54	0,45

Pour les types mixtes de végétation (SBb et BbS), les faibles concordances locales peuvent être dues à plusieurs aspects inter-reliés :

- (1) Le patron spatial intra-peuplement des espèces qui composent la végétation, le sapin baumier et le bouleau à papier. La carte forestière ne fournit aucune information sur ce patron spatial;
- (2) Le fait que les catégories soient assignées au moment de la photo-interprétation en prenant comme base une unité minimale sur laquelle sont estimées les proportions de couvert de chaque espèce. Sur le terrain, ces catégories sont assignées par l'observation du couvert forestier sur la p-e et les zones environnantes;
- (3) La différence de taille entre l'unité cartographique minimale et la p-e (Bérard et Côté 1996). Ces différences peuvent influencer fortement la concordance entre les catégories estimées sur le terrain et celles sur les photos aériennes (Foody 2002).

La conjugaison des trois aspects implique que la probabilité que la p-e soit localisée sur le terrain à un endroit représentatif du patron de couvert observé sur la photographie aérienne est supérieure dans le cas d'un peuplement dominé par une espèce, surtout si elle présente un patron spatial uniforme et une densité élevée. Dans ce cas, l'estimation du volume ligneux serait plus certaine parce que la strate est plus uniforme et les échantillons à partir desquels ces estimations sont élaborées sont plus représentatifs de la strate.

Pour la *Classe de densité (CD)*, les catégories de couvert intermédiaire (C et D) présentent les concordances locales les plus élevées, alors que la catégorie du couvert plus dense (A) présente la plus faible (Tableau 41). À l'échelle du voisinage spatial, les quatre catégories montrent des valeurs similaires à celles de l'échelle locale (Tableau 41).

Tableau 41. Indices de concordance carte/p-e des catégories de la variable *Classe de densité (CD)*, aux échelles locales (ICL) et de voisinage (ICV).

Catégorie	ICL	ICV
-----------	-----	-----

A (81-100 % de couvert)	0,02	0,03
B (61-80 % de couvert)	0,35	0,32
C (41-60 % de couvert)	0,38	0,30
D (25-40 % de couvert)	0,18	0,16

La logique veut que, *a priori*, les catégories extrêmes A (81-100 % du couvert) et D (25-40 % du couvert) soient plus facilement identifiables sur le terrain que les catégories intermédiaires. Cependant, les résultats obtenus indiquent le contraire. Comme pour *Groupement d'essences*, la différence des méthodes de saisie de l'information entre la carte forestière et le terrain doit influencer fortement l'identification des catégories, qui repose sur l'estimation visuelle du couvert de la végétation. Même l'information saisie par les deux sources est de nature différente, l'une étant issue d'une évaluation du couvert forestier sur une unité minimale, l'autre étant issue d'une évaluation à partir du sol en absence de cadre spatial. Cependant, les deux sont organisées de façon similaire. Il est probable que la perception du terrain soit biaisée vers les valeurs centrales de couvert qui sont également les plus fréquents dans la zone d'étude. Les différences de fréquence des catégories sur la carte forestière peuvent également expliquer les valeurs de concordance avec le voisinage, les catégories plus fréquentes présentent des valeurs plus élevées comparativement aux autres catégories.

La classe d'arbres dont la hauteur varie de 7 à 12 m présente les meilleurs indices de concordance aux deux échelles, tandis que ceux des classes correspondant aux arbres de moindre taille sont nettement inférieurs (catégories 4, 5 et 6, Tableau 42). Comme pour variables précédentes, il est possible que la concordance soit affectée par la méthode de saisie de l'information et que les différences diminuent à mesure que les arbres grandissent.

La *Classe d'âge (CA)* présente un schéma similaire à la *Classe de hauteur*. La végétation dont l'âge est estimé à 45 ans affiche une concordance aux deux échelles supérieure à celle des autres catégories (Tableau 42). Ces résultats suggèrent que l'étape de développement du peuplement peut avoir une influence sur deux aspects. D'une part, il est possible qu'une étape intermédiaire du développement de la végétation ligneuse soit plus homogène du point de vue structural que les étapes initiales ou celles de maturité. Cette idée est reprise aux paragraphes subséquents qui traitent des strates dominées par le sapin baumier. D'autre

part, elle peut également influencer la facilité d'identifier et d'assigner une catégorie d'âge lors de la photo-interprétation.

Tableau 42. Indices de concordance carte/p-e des catégories des variables *Classe de hauteur (CH)* et *Classe d'âge (CA)*, aux échelles locales (ICL) et de voisinage (ICV).

<i>Classe de hauteur (CH)</i>			<i>Classe d'âge (CA)</i>		
Catégorie	ICL	ICV	Catégorie	ICL	ICV
1 (> 22 m)	0,00	0,00	15 (0-20 ans)	0,00	0,00
2 (17-22 m)	0,04	0,02	30 (21-40 ans)	0,14	0,18
3 (12-17 m)	0,49	0,39	45 (41-45 ans)	0,43	0,34
4 (7-12 m)	0,25	0,25	60 (46-60 ans)	0,00	0,00
5 (4-7 m)	0,11	0,12	75 (61-80 ans)	0,14	0,05
6 (1,5-4 m)	0,15	0,16	90 (81-100 ans)	0,00	0,00

Bien que les résultats obtenus ne permettent pas d'ordonner les strates selon un seul gradient de concordance locale, les tendances sont observées pour des catégories spécifiques. La Figure 44 illustre ces tendances et constitue une synthèse des résultats présentés précédemment.

En effet, pour la combinaison des quatre variables, seules quatre strates présentent une valeur de concordance différente de 0 et correspondent à la végétation résineuse dominée par le sapin baumier (catégorie SS de *GE*), dont les arbres varient entre 12 et 17 m de hauteur (catégorie 3 de *CH*), le couvert s'étend entre 41 et 80 % (catégories B et C de *CD*) et l'âge est estimé à 45 ans (catégorie 45 de *Classe d'âge*) (Figure 44). Les strates sont SS B3 45 el, SS C3 45 cp, SS B3 45 et SS C3 45 el. Sur la figure, seules les combinaisons des catégories sont représentées, et non les strates.

À l'opposé, pour la combinaison d'une variable, les catégories moins concordantes correspondent à la végétation en régénération (catégories 0 et 30 de *CA*, et de *GE*, D de *CD*) suivies de celles qui correspondent à la végétation plus développée (catégorie 75 de *CA*).

4 variables	<b>GE-CD-CH-CA</b> <b>SS B 3 45</b> <b>SS C 3 45</b>											
3 variables	<b>GE-CD-CH</b> <b>SS C 3</b> <b>SBb B 5</b> <b>SS B 3</b> <b>SBb B 3</b> <b>SS B 4</b>		<b>GE-CD-CA</b> <b>SS C 45</b> <b>SS B 45</b> <b>SS C 75</b> <b>SS C 60</b> <b>SBb B 15</b>		<b>GE-CH-CA</b> <b>SS 3 45</b> <b>SS 4 30</b>		<b>CD-CH-CA</b> <b>C 3 75</b> <b>B 3 45</b> <b>C 3 45</b>					
2 variables	<b>GE-CD</b> <b>SS C</b> <b>SS B</b> <b>SBb B</b> <b>SS D</b>		<b>GE-CH</b> <b>SS 3</b> <b>SBb5</b> <b>SS 4</b> <b>SBb 3</b>		<b>GE-CA</b> <b>SS 45</b> <b>SS 75</b> <b>SS 30</b> <b>SS 60</b>		<b>CD-CH</b> <b>SS C 3</b> <b>SS B 3</b> <b>SS B 4</b> <b>SS A 3</b>		<b>CD-CA</b> <b>C 45</b> <b>C 75</b> <b>B 45</b> <b>C 60</b> <b>B 60</b> <b>B 30</b>		<b>CH-CA</b> <b>3 45</b> <b>3 75</b> <b>4 30</b> <b>4 45</b>	
1 variable	<b>GE</b> <b>SS</b> <b>SBb</b> <b>ct</b>		<b>CD</b> <b>C</b> <b>B</b> <b>A</b> <b>D</b>		<b>CH</b> <b>3</b> <b>4</b> <b>5</b> <b>6</b>		<b>CA</b> <b>45</b> <b>75</b> <b>30</b> <b>0</b>					
Flèche rouge : indique la direction du gradient												

Figure 44. Gradients de concordance locale pour les catégories des variables individuelles et en combinaisons.

#### 4. La concordance carte/p-e et l'incertitude spatiale des p-e

Cette section vise à évaluer l'effet que l'incertitude sur l'appartenance des p-e aux strates, liée à la distance des p-e aux frontières des polygones, peut avoir sur les valeurs de concordance carte/p-e des variables examinées aux sections précédentes.

Le Tableau 43 présente les valeurs de concordance globale de chaque variable, de concordance des catégories les plus importantes en terme de fréquence de chacune des variables et les coefficients kappa. Elles sont obtenues en tenant compte de l'ensemble des p-e (comme cela a été fait aux sections précédentes) et en tenant compte des p-e qui sont localisées à différentes distances de la frontière de la strate. La distance croissante de la p-e à la frontière du polygone traduit une diminution du risque de confusion spatiale. En effet,

plus la p-e est située près de la frontière du polygone, plus grand est le risque que, par des erreurs positionnelles de la p-e et/ou de la frontière, elle appartienne au polygone voisin. De cette façon, la distance croissante par rapport à la frontière du polygone implique, dans le contexte de cette analyse, une certitude croissante que la p-e appartient à la strate représentée par le polygone.

Tableau 43. Valeurs de concordance globale, de concordance par catégorie et de coefficient kappa des variables *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)* correspondant à l'ensemble des p-e et aux p-e distants d'au moins 5 m, 10 m, 25 m ou 50 m de la frontière de la strate forestière.

		Toutes les p-e	p-e > 5 m	p-e > 10 m	p-e > 25 m	p-e > 50 m
<i>GE</i>	Concordance globale	0,40	0,40	0,41	0,42	0,44
	Kappa	-0,03	-0,02	0,04	-0,01	0,05
	SS	0,58	0,58	0,60	0,62	0,60
<i>CD</i>	Concordance globale	0,30	0,31	0,31	0,33	0,24
	Kappa	-0,018	0,005	-0,002	0,020	-0,084
	B (61-80 % de couvert)	0,33	0,35	0,35	0,36	0,22
	C (41-60 % de couvert)	0,34	0,37	0,36	0,37	0,36
<i>CH</i>	Concordance globale	0,35	0,35	0,37	0,36	0,40
	Kappa	0,002	0,010	0,015	-0,0008	0,041
	3 (12-17 m)	0,53	0,54	0,56	0,57	0,67
	5 (4-7 m)	0,10	0,08	0,09	0,10	0,00
	6 (1,5-4 m)	0,13	0,15	0,18	0,17	0,22
<i>CA</i>	Concordance globale	0,26	0,29	0,29	0,19	0,26
	kappa	-0,024	-0,009	0,0001	-0,015	0,048
	15 (0-20 ans)	0,02	0,02	0,03	0,00	0,00
	30 (21-40 ans)	0,10	0,11	0,13	0,13	0,08
	45 (41-45 ans)	0,56	0,60	0,62	0,62	0,69
	75 (61-80 ans)	0,13	0,14	0,13	0,19	0,33

La concordance globale de toutes les variables varie faiblement en fonction de la localisation ou « certitude » des p-e (Tableau 43). La concordance globale de *GE*, *CD* et *CH* tend à augmenter à mesure que les p-e utilisées sont localisées à une distance croissante de la frontière du polygone. Parmi ces trois variables, *GE* présente la meilleure concordance globale (Tableau 43).

Les coefficients kappa sont très près de 0 pour toutes les variables et pour toutes les situations, ce qui indique que les concordances globales sont probablement dues au hasard

(Tableau 43). Plusieurs variables (*GE*, *CD* et *CA*) présentent des valeurs négatives du coefficient kappa (cela signifie que la concordance globale observée est inférieure à celle produite par le hasard), particulièrement lorsque toutes les p-e sont considérées dans l'analyse. *CH* est la seule variable dont la valeur de kappa tend à augmenter à mesure que la distance des p-e par rapport aux frontières augmente. Lorsque la comparaison de valeurs implique seulement l'ensemble des p-e et les p-e localisées à 50 m ou plus de la frontière, la valeur de kappa de ces dernières est supérieure à celle de l'ensemble des p-e et ce, pour la plupart des variables. Il demeure néanmoins que toutes les valeurs de kappa sont très près de 0, c'est-à-dire que les concordances observées ne se démarquent pas de celles dues au hasard (non significatives).

Finalement, la concordance de certaines des catégories les plus importantes augmente légèrement lorsque seules les p-e localisées à 50 m de la frontière (dont l'appartenance à la strate est plus certaine) sont considérées (Tableau 43). Cette tendance concerne les catégories 3 et 6 de la *Classe de hauteur* et les catégories 45 et 75 de la *Classe d'âge*.

Deux interprétations de ces résultats sont possibles. Premièrement, en considérant seulement les valeurs de concordance globale pour chaque variable, ces résultats indiquent que l'incertitude sur la localisation des p-e n'affecte pas tellement les valeurs de concordance observées. Cependant, les valeurs du coefficient kappa suggèrent également que la carte forestière et les p-e sont deux sources d'information indépendantes sur une même réalité et qu'elles n'ont pas de relation particulière entre elles, car les concordances observées pourraient être le produit du hasard. Toutefois, cette analyse est faite à partir de toutes les strates forestières réunies. Le transfert de ces conclusions aux concordances des strates particulières ne peut donc pas être direct. Les valeurs élevées ou nulles de concordance de certaines des strates vont au-delà des valeurs aléatoires. En fait, il faudrait estimer le coefficient kappa pour chacune des strates pour vérifier si les concordances observées sont aléatoires. Le problème est que la plupart des strates comptent un nombre de p-e très réduit pour le faire.

## Conclusions

Les indices de concordance carte/p-e utilisés dans ce chapitre permettent d'établir un lien entre les deux sources d'information. L'indice de concordance locale indique la mesure dans laquelle les p-e représentent les conditions forestières de la strate cartographique. En effet, étant donnée la différence d'échelle et plus précisément de grain entre les unités cartographiques et les p-e, ainsi que l'homogénéité interne des conditions forestières des strates, les p-e représentent à divers degrés les caractéristiques relevées par la cartographie. Cette information est utile, car les estimations du volume sont dérivées de groupes de p-e homogènes du point de vue forestier. D'autant plus que lorsque ces estimations du volume sont assignées aux strates de la carte, il y a une incertitude ou variabilité qui s'ajoute à la variabilité intrinsèque des estimations. La concordance locale permet donc de quantifier cette incertitude pour les différentes strates.

La concordance à l'échelle du voisinage donne de l'information sur l'incertitude liée à l'appartenance des p-e à une strate particulière. Cette incertitude est elle-même issue de l'incertitude positionnelle, autant des p-e que des frontières des polygones, ainsi que des conditions structurales des frontières (espèces, hauteur des arbres, etc.). De plus, la concordance peut aussi indiquer également une homogénéité des conditions du voisinage pour les variables concernées.

L'application des deux indices permet de déterminer si l'incertitude ou variabilité liée aux strates d'une carte forestière est de caractère local, c'est-à-dire si elle est particulière à chaque strate. Par rapport aux objectifs spécifiques proposés, la concordance carte/p-e des strates forestières a été quantifiée à l'échelle locale et du voisinage spatial, pour les variables *Groupement d'essences (GE)*, *Classe de densité (CD)*, *Classe de hauteur (CH)* et *Classe d'âge (CA)* ainsi que pour leurs combinaisons.

Parmi les strates les plus importantes de la Forêt Montmorency, celles dominées par le sapin baumier avec un couvert ligneux moyen, dont la hauteur des arbres varie entre 12 et 17 m, et qui sont parfois affectées par des perturbations légères (SS B3 45, SS B3 45 el, SS C3 45), présentent une forte concordance locale. Les strates correspondant à la

végétation en phase de régénération concordent, au contraire, seulement avec le voisinage. Cette caractéristique serait reliée à la proportion de p-e de la strate cartographique dont la distance à la frontière du peuplement est de 5 m ou moins.

Quant aux gradients de concordance, ils sont applicables aux catégories de variables examinées, plutôt qu'aux strates forestières. Les catégories correspondant à la végétation résineuse dominée par le sapin baumier (SS), le couvert de la végétation ligneuse entre 41 et 80 % de la surface terrière (B et C), la hauteur des arbres variant entre 12 et 17 m (3) et à un âge estimé à 45 ans (45) sont celles qui présentent une plus forte concordance locale. Cette concordance serait reliée : (1) aux critères de définition des catégories de chaque variable, (2) aux procédures d'identification des catégories au moment de l'élaboration de la carte forestière et lors de l'échantillonnage de terrain, (3) aux patrons spatiaux de la végétation à l'échelle intra-peuplement, et (4) à la différence de taille entre l'unité cartographique et l'unité d'échantillonnage sur le terrain.

Pour les strates résineuses dont l'essence prédominante est le sapin baumier, qui sont les plus importantes en surface à la Forêt Montmorency, le gradient de concordance semble être lié aux étapes du développement de la végétation ligneuse. Les premières étapes présentent une faible concordance car la végétation présente une forte diversité d'essences, supérieure à celle des étapes postérieures. Les étapes intermédiaires, de pré-maturité, présentent la concordance maximale, possiblement en raison d'une plus grande homogénéité structurale de la végétation (qui est saisie de manière plus efficace par les p-e). Finalement, les étapes de maturité présentent à nouveau une faible concordance, possiblement en raison de l'accumulation de perturbations légères qui génèrent de l'hétérogénéité dans l'environnement et dans la structure de la végétation ligneuse, qui n'est pas saisie par l'échantillon. De ce point de vue, les estimations du volume ligneux des strates dans l'étape de pré-maturité seraient donc plus précises que celles des strates correspondant aux étapes antérieures ou postérieures, car la structure de la végétation relevée dans les p-e est plus semblable à celle représentée dans la cartographie.

Les variables montrent une consistance similaire lorsque l'information provient de la photo-interprétation ou du terrain. Cependant, les différences sont mises en évidence au

niveau des catégories. En effet, les catégories SS (*GE*), B et C (*CD*), 3 (*CH*) et 45 (*CA*) déjà mentionnées, sont plus stables, car elles obtiennent une forte concordance locale entre les deux sources d'information.

L'incertitude sur l'appartenance des p-e à une strate, représentée dans cette étude par la distance de la p-e à la frontière du polygone, ne semble pas affecter fortement les valeurs de concordance globale et par catégorie des variables. Néanmoins, les valeurs de coefficient kappa indiquent que la carte forestière et les p-e sont deux sources de données issues d'une même réalité mais qui n'ont aucune relation étroite entre elles. Il est possible que si la même analyse était réalisée pour chacune des strates forestières, les résultats de cet indice seraient différents. Cependant, le nombre de p-e disponibles pour chaque strate est trop faible pour permettre une telle analyse.

## **Résumé du chapitre**

Dans ce chapitre, une méthode pour qualifier les strates forestières en terme de la concordance entre l'information cartographique et de terrain a été appliquée aux échelles locales et du voisinage. Les informations obtenues aux deux échelles sont complémentaires et peuvent aider à mieux planifier les sondages de terrain ainsi qu'à déterminer le degré de variabilité des estimations du volume de matière ligneuse correspondant aux strates forestières. L'information obtenue dans ce chapitre complète, de ce point de vue, l'information du chapitre précédent. Les approches possibles pour fusionner ces informations de la carte forestière et des p-e sont discutées au chapitre suivant, de même que les obstacles à résoudre.

## **Chapitre VII - Discussion**

Le but de cette recherche est d'explorer différentes approches pour améliorer la caractérisation et la quantification de l'information reliée à l'incertitude locale d'une carte forestière et ce, en utilisant les données cartographiques et de terrain. Dans ce contexte, ce chapitre discute des possibilités et des contraintes de la fusion de ces informations.

Les concepts théoriques les plus importants sur la fusion des données spatiales ont été présentés au Chapitre I. Ce chapitre présente en premier une discussion sur les problèmes relatifs à la fusion de informations spatiales de la carte forestière et des échantillons de terrain. Par la suite, l'incertitude des estimations du volume de matière ligneuse des strates forestières est analysée en considérant les informations sur la concordance carte/p-e des strates forestières ainsi que la variabilité des estimations.

### **Fusion des informations spatiales de la carte forestière et des échantillons de terrain**

Cette section discute les possibilités de fusion des informations de la carte forestière avec celles issues des échantillons de terrain, à partir de l'analyse des hétérogénéités et des ressemblances les plus importantes entre les deux représentations du même territoire.

Le Chapitre III porte sur la qualification des frontières de la carte forestière en fonction du contraste écologique des peuplements voisins d'une frontière donnée. Cette qualification permet d'identifier les frontières dont l'existence est plus certaine et de façon indirecte, à travers d'une meilleure définition des surfaces des unités forestières, contribue à augmenter la précision des estimations du volume ligneux pour une zone particulière. Une autre façon d'arriver au même but consiste à identifier des zones homogènes et des discontinuités spatiales dans le volume ligneux à partir des échantillons de terrain, ce qui est l'objet du Chapitre IV. Ces deux chapitres sont reliés par le concept suivant : la fusion de

l'information obtenue par les p-e et par la carte forestière pourrait permettre de renforcer le degré de certitude sur l'existence de certaines frontières cartographiques ainsi que d'identifier d'autres frontières reliées de façon plus directe au volume ligneux. Cette procédure de fusion permettrait l'identification des zones dans lesquelles il est possible d'estimer le volume avec une meilleure précision en fonction des superficies.

Cependant, les deux sources d'information correspondent à deux représentations différentes du même territoire forestier. De nombreuses différences et ressemblances entre les deux au plan des aspects sémantiques, syntactiques et schématiques déterminent les possibilités de fusion sur le plan des éléments et des connaissances. Ces aspects correspondent partiellement aux caractéristiques suggérées par Hangouët (2004) pour décrire et comparer les différentes représentations des phénomènes géographiques. En effet, cet auteur considère qu'une représentation peut être décrite à partir du phénomène, de l'attention, du médium, de l'inscription et de la réception (PAMIR). Le phénomène réfère à la classe d'entités du monde réel sur laquelle porte la représentation, l'attention correspond au cadre théorique à partir duquel le phénomène est approché, le médium est l'ensemble des caractéristiques de saisie et d'emmagasinage des données sur le phénomène, l'inscription représente les données sur le phénomène, tandis que la réception est l'interprétation du phénomène à partir de la représentation (Hangouët 2004). Ainsi, l'attention est liée aux aspects sémantiques et schématiques, alors que le médium et l'inscription sont liés aux aspects syntactiques.

Le Tableau 44 présente de manière succincte les différences et les ressemblances entre les deux représentations, en ce qui a trait au phénomène représenté, ainsi qu'aux aspects sémantiques, syntactiques et schématiques. Ces caractéristiques sont expliquées aux sections suivantes et sont illustrées par des exemples.

Tableau 44. Différences et ressemblances au plan des aspects sémantiques, syntactiques et schématiques entre les deux représentations de la forêt.

	Cartographie	Terrain
<b>Phénomène</b>		
Différences	• Structure de la végétation	• Volume ligneux total
Similitudes	• Structure de la végétation, pour la variable <i>Type forestier</i>	
<b>Sémantique</b>		
Différences	• Unité délimitée : les peuplements	• Unité délimitée : des groupes de p-e adjacentes homogènes → frontières indirectes
Similitudes	<ul style="list-style-type: none"> <li>• Les frontières cartographiques peuvent être les bordures des zones homogènes (analogues aux frontières indirectes) ou des zones de changement intense de la structure de la végétation (analogues des frontières directes)</li> <li>• La quantification du contraste écologique permet d'identifier les frontières directes cartographiques</li> <li>• Il existe des coïncidences spatiales et non spatiales entre les frontières directes de <i>Type forestier</i> ainsi que du <i>Volume total</i> avec les frontières cartographiques</li> </ul>	
<b>Syntactique</b>		
Différences	<ul style="list-style-type: none"> <li>• Résolution spatiale : unité cartographique minimale de 8 ha. Taille moyenne des polygones de 1,72 ha.</li> <li>• Saisie de l'information : variables interprétées sur des photographies aériennes.</li> <li>• Structure des données vectorielle/ligne.</li> </ul>	<ul style="list-style-type: none"> <li>• Résolution spatiale : taille de la p-e de 0,04 ha. Espacement moyen entre les p-e de 182,8 m.</li> <li>• Saisie de l'information : variables estimées visuellement sur le terrain.</li> <li>• Structure des données matricielle (par exemple, un polygone de Voronoï)</li> </ul>
Similitudes	<ul style="list-style-type: none"> <li>• Étendue spatiale</li> <li>• Résolution temporelle et thématique</li> <li>• Variables nominales et ordinales</li> </ul>	
<b>Schématique</b>		
Différences	<ul style="list-style-type: none"> <li>• Modèle de l'espace sous forme d'objets</li> <li>• Les frontières sont le produit d'un processus cognitif complexe. Il n'y a pas de différences entre les frontières directes et indirectes.</li> </ul>	<ul style="list-style-type: none"> <li>• Modèle de l'espace en champs.</li> <li>• Les frontières sont le résultat des algorithmes mathématiques et statistiques qui cherchent des ressemblances ou des discontinuités entre les p-e, selon que les frontières soient indirectes ou directes.</li> </ul>

## 1. Différences et ressemblances liées au phénomène représenté

Au plan du phénomène écologique représenté, la carte forestière représente la structure de la végétation forestière décrite à partir d'un ensemble de variables comme le *Groupement d'essences*, la *Classe de densité*, la *Classe de hauteur* et la *Classe d'âge*, parmi d'autres. Quant au phénomène représenté par les données de terrain, il diffère en fonction des variables utilisées. Les frontières de la variable *Type forestier* représentent le même phénomène que la carte forestière, car le *Type forestier* combine les quatre variables mentionnées précédemment (Tableau 44). Les frontières de la variable *Volume total* représentent un autre phénomène : le volume ligneux de la végétation. Ce phénomène est relié à la structure de la végétation dans une relation complexe.

## 2. Différences et ressemblances liées à l'aspect sémantique

Dans le contexte de cette recherche, l'explicitation sémantique des entités à fusionner est un exercice qui doit être abordé en premier lieu pour obtenir des résultats interprétables du point de vue écologique et forestier. De ce point de vue, les différences entre les deux représentations de la forêt concernent principalement les entités délimitées par les frontières, car les frontières en tant que telles ne diffèrent pas au plan conceptuel entre la carte et les p-e lorsqu'il s'agit pour ces dernières des frontières de la structure de la végétation (correspondant à la variable *Type forestier*) (Tableau 44).

Les frontières indirectes ne délimitent pas une unité spatiale identifiable et définie *a priori* du point de vue écologique comme c'est le cas pour les peuplements. Or, la carte forestière représente, par définition, les écosystèmes forestiers (Bérard et Côté 1996). Au Québec et dans d'autres régions, la carte représente, plus précisément, les peuplements en tant qu'unités d'aménagement et de gestion forestière. Lors de l'utilisation de la carte, étape qui constitue la composante de la réception de la représentation (Hangouët 2004), l'attention est concentrée sur les unités forestières auxquelles des valeurs du volume ligneux peuvent être attribuées et non sur les discontinuités ou frontières en tant que telles. Par conséquent, la fusion des éléments, c'est-à-dire des peuplements et des unités homogènes par rapport à

la structure de la végétation selon les p-e, exige la résolution de cette différence entre les deux représentations.

Le développement des ontologies offre une approche intéressante pour résoudre ces différences sémantiques ainsi que pour effectuer la fusion. Une ontologie pour les entités et phénomènes représentés dans la carte forestière doit tenir compte de l'importante composante cognitive impliquée dans l'élaboration de la carte, tandis qu'une ontologie des entités représentées par les p-e doit tenir compte du caractère local de cette information de terrain. Ainsi, la fusion des deux types d'unités forestières doit être accomplie sur le plan des ontologies avant de poursuivre avec les aspects syntactiques et schématiques. Le développement d'une ontologie pour cette fusion dépasse le cadre de cette recherche. Cependant, elle s'avère une voie possible pour explorer la fusion des éléments.

En ce qui a trait aux ressemblances entre les représentations, les frontières cartographiques sont similaires du point de vue conceptuel aux frontières indirectes, mais également aux frontières directes issues des p-e. En effet, bien que ce ne soit pas explicité dans la carte forestière, les frontières cartographiques peuvent correspondre à des situations où il y a un changement abrupt de la structure de la végétation. Il a déjà été mentionné qu'à l'étape de la photo-interprétation, ce sont les frontières et les attributs des unités qui sont importants pour la délimitation des peuplements (Lowell *et al.* 1996, Thierry 1998, Lowell 2001). De plus, la quantification du contraste des frontières cartographiques, tel que développée au Chapitre III, permet d'identifier les frontières de fort contraste qui sont analogues aux frontières directes issues des p-e. La carte forestière contient donc, de façon indistincte, des frontières directes et indirectes dans le sens de Fortin et Drapeau (1995) et de Jacquez *et al.* (2000).

Cette ressemblance conceptuelle se reflète dans les coïncidences spatiales observées entre les frontières des deux sources de données. Par exemple, les frontières de terrain pour les variables *Type forestier* et *Volume total* ainsi que les frontières cartographiques de fort *Contraste global* (cette variable est formée par l'addition du contraste des variables qui composent le *Type forestier*) prédominent dans les secteurs centre-nord et est de la Forêt Montmorency. Le patron est inversé dans le secteur sud-ouest où prédominent les frontières

cartographiques de faible *Contraste global* ainsi que les discontinuités spatiales très faibles de *Type forestier* et du *Volume total*.

La ressemblance conceptuelle se reflète également dans les coïncidences observées lors de l'analyse non spatiale. De ce point de vue, les frontières directes du *Volume total* et les frontières cartographiques de fort contraste présentent une fréquence élevée de combinaisons de catégories résineuses et mixtes à prédominance feuillue, comme SS-BbS (sapin baumier pur-bouleau blanc et sapin baumier), pour la variable *Groupement d'essences*. Les discontinuités spatiales de très faible intensité pour le *Volume total* présentent, au contraire, une fréquence élevée des combinaisons de catégories similaires (par exemple SS-SS), qui correspondent au niveau « sans contraste » des frontières cartographiques.

Ces coïncidences non spatiales concernent également les autres variables et elles peuvent être utilisées comme des règles pour la fusion au niveau des connaissances. Par exemple, les frontières cartographiques entre un peuplement résineux où le sapin prédomine (SS) et un peuplement mixte où le bouleau prédomine (BbS) ont une existence plus certaine, de même que les frontières entre un peuplement où les arbres dominants mesurent entre 12 et 17 m et un peuplement où ils mesurent moins de 4 m. Au contraire, les frontières cartographiques entre deux peuplements où les arbres sont de taille semblable (par exemple, deux classes de hauteur consécutives) ou à composition d'essences similaire, sont plus incertaines. Cette fusion au niveau des connaissances est importante même si l'aspect spatial local de la fusion des éléments est perdu.

Les différences conceptuelles entre les unités spatiales de la carte et celles délimitées par les frontières de terrain sont étroitement liées à l'échelle spatiale de chacune des sources, ce qui correspond à l'aspect syntactique des données.

### **3. Différences et ressemblances liées à l'aspect syntactique**

Les différences syntactiques les plus importantes entre la carte forestière et les cartes dérivées à partir des échantillons de terrain portent sur le degré de détail ou de résolution

spatiale, ainsi que sur la méthode de mesure des variables et la structure des données correspondant aux frontières (Tableau 44). Ces différences sont traitées séparément dans cette section.

### 3.1. Différences qui ont trait à la résolution spatiale

La différence de la résolution spatiale ou du grain des données de la carte forestière et des p-e représente la plus importante source d'hétérogénéité et constitue l'obstacle principal à la fusion des deux représentations sur le plan des éléments. Elle affecte de façon directe l'aspect sémantique car, comme il a été mentionné dans le cadre théorique, la variation spatiale d'un phénomène dépend de l'échelle à laquelle il est échantillonné et de la taille de l'unité cartographique qui le représente (Gustafson 1998). Par conséquent, l'interprétation écologique des unités délimitées par les frontières de terrain et la définition ontologique de ces entités dépendent fortement de l'échelle spatiale à laquelle elles sont identifiées.

En ce sens, la relation entre l'espacement des p-e et la taille des polygones de la carte forestière détermine que les groupes de p-e adjacentes formés par l'algorithme de regroupement (auxquels correspondent les frontières indirectes) ont une étendue très supérieure à la surface des peuplements représentés par les polygones. Cette différence donne lieu à diverses situations lorsque les frontières issues des deux sources sont superposées.

Par exemple, les frontières directes issues des p-e pour le *Type Forestier* peuvent correspondre à une ou à plusieurs frontières cartographiques de fort contraste et même à des peuplements entiers, car la distance géographique entre les deux p-e est très grande par rapport à la taille des peuplements (Figure 45 a et b).

Une situation similaire est observée lorsque les frontières cartographiques de fort contraste sont superposées aux frontières du *Volume total*. La différence de grain est évidente pour les frontières indirectes, soit les frontières de 100 ou de 50 groupes de p-e (Figure 45c). Les frontières indirectes qui persistent lors du changement d'échelle, identifiées au Chapitre IV, coïncident fréquemment avec les frontières directes et, bien que toutes les frontières de

terrain soient placées de façon automatique entre deux p-e, l'étendue spatiale des groupes est très supérieure à la surface de la plupart des peuplements (Figure 45c).

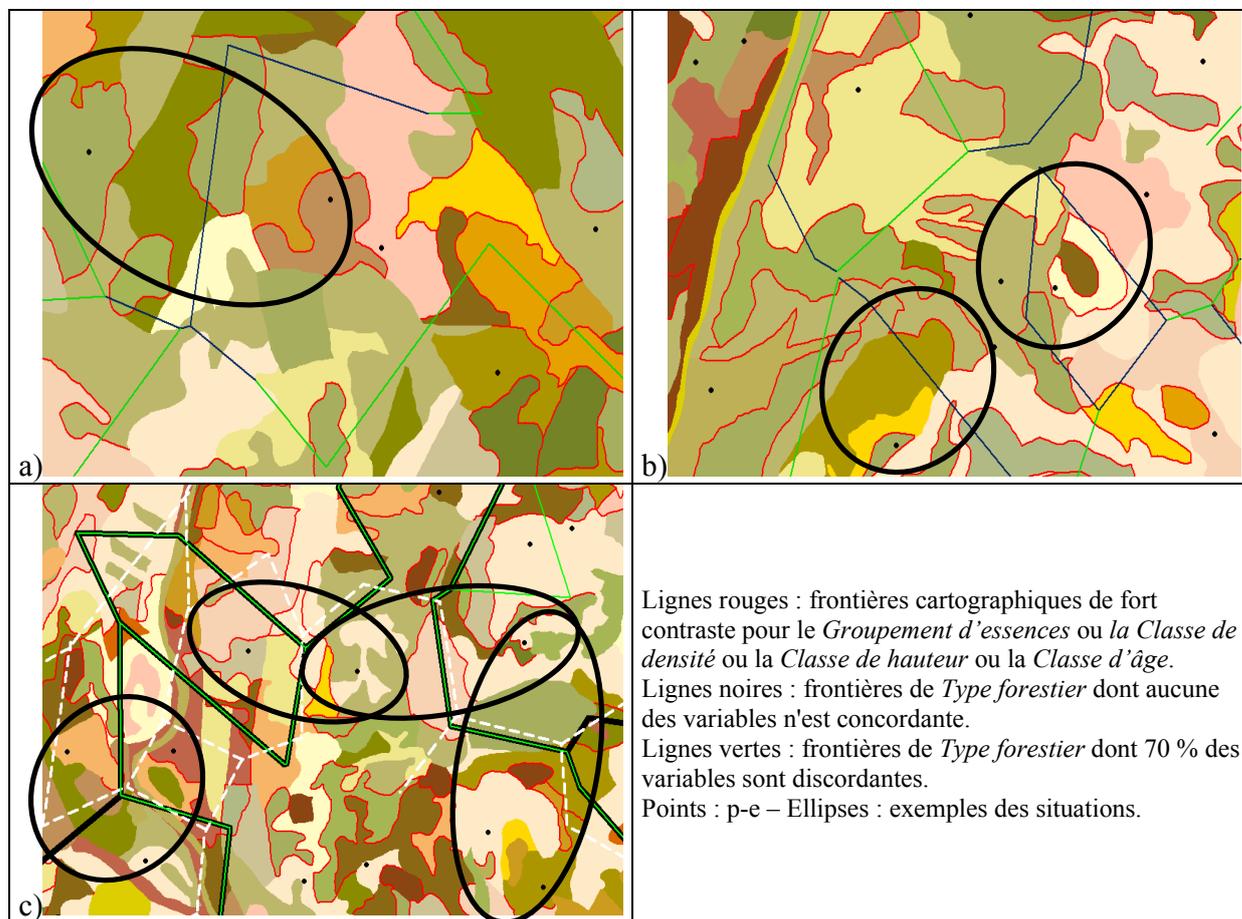


Figure 45. Exemples de superposition des frontières cartographiques aux frontières issues des p-e : a) et b) frontières directes de *Type forestier*, c) frontières directes et indirectes du *Volume total*.

Ces exemples illustrent comment la différence de grain entre la carte forestière et les p-e rend extrêmement difficile l'établissement d'une correspondance spatiale entre les frontières des deux sources de données, étape importante de la fusion au niveau des éléments.

Les exemples précédents correspondent à des frontières cartographiques de fort contraste local. Or, en considérant la différence de grain entre les deux représentations de la forêt, une correspondance entre les frontières de fort contraste à l'échelle régionale et les

frontières issues des données de terrain pourrait être envisagée car les premières sont déterminées à partir des groupes de peuplements adjacents, donc à un grain plus large que les peuplements isolés. La Figure 46 présente quelques exemples de la superposition des deux types de frontières où trois situations sont également fréquentes : (1) les frontières de la carte peuvent correspondre aux frontières issues des p-e, soit pour le *Type forestier* (Figure 46a) ou pour le *Volume total* (Figure 46b et c); (2) des frontières cartographique de fort contraste régional ne correspondent à aucune frontière de terrain ou (3) des frontières de terrain ne correspondent à aucune frontière cartographique.

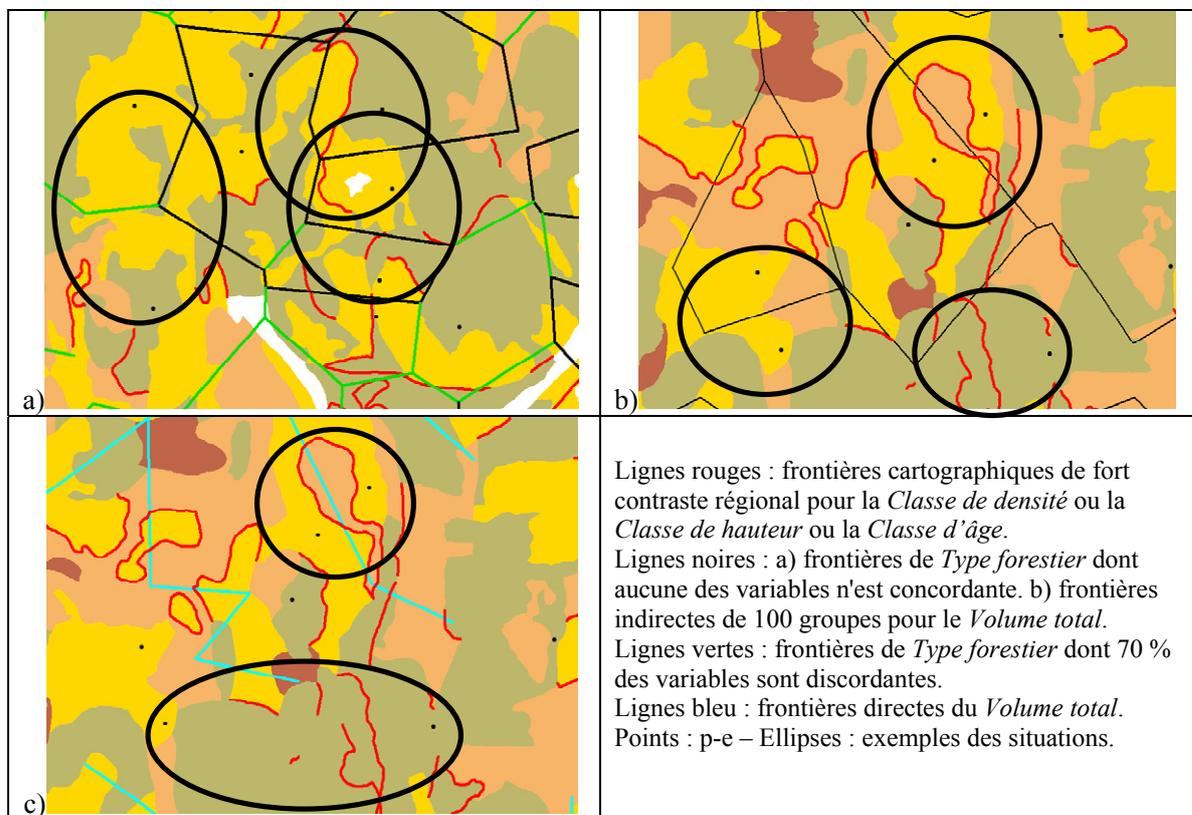


Figure 46. Exemples de superposition de frontières cartographiques à fort contraste à l'échelle régionale (ou de voisinage) aux frontières issues de données de terrain : a) Frontières directes de *Type Forestier*, b) Frontières indirectes (100 groupes) du *Volume total* et c) Frontières directes du *Volume total*.

Une autre manière de rapprocher le grain des deux représentations est de considérer les frontières cartographiques des groupes de peuplements dont la valeur de volume ligneux

estimée est similaire. Une correspondance avec les frontières de terrain pour le *Volume total* pourrait donc être envisagée. La Figure 47 présente quelques exemples de la superposition de ces frontières. Les peuplements ont été regroupés selon deux méthodes, celle actuellement utilisée au Québec (RT) et celle du modèle de régression arborescente employée au Chapitre V, dont la variabilité spatiale des estimations est la plus faible (PartielT17). Les peuplements ont été regroupés en dix-sept et en cinq classes, afin d'obtenir deux grains beaucoup plus larges que celui des peuplements isolés. Comme pour les exemples présentés précédemment, plusieurs situations sont identifiées : (1) une correspondance peut être établie entre les frontières de terrain et les frontières cartographiques, mais cette situation est la moins fréquente; (2) des frontières de terrain séparent des p-e localisées dans un même groupe de peuplements, cette situation est très fréquente, et (3) aucune frontière de terrain correspond à des frontières cartographiques de groupes de peuplements dont le volume ligneux estimé est très différent.

Les deux derniers exemples indiquent que si bien la différence de grain est très importante et constitue un obstacle pour établir une correspondance entre les frontières de deux représentations, lorsque ces différences de grain diminuent la correspondance spatiale demeure encore difficile à établir. La grande fragmentation de la végétation à la Forêt Montmorency pourrait expliquer en partie cette difficulté. En effet, même lorsque les peuplements sont regroupés en cinq classes de volume ligneux, demeurent fréquentes les parcelles peu étendues et la variabilité spatiale des estimations est élevée. Il est possible que les différences liées aux autres aspects entre les deux types de représentation jouent également un rôle important.

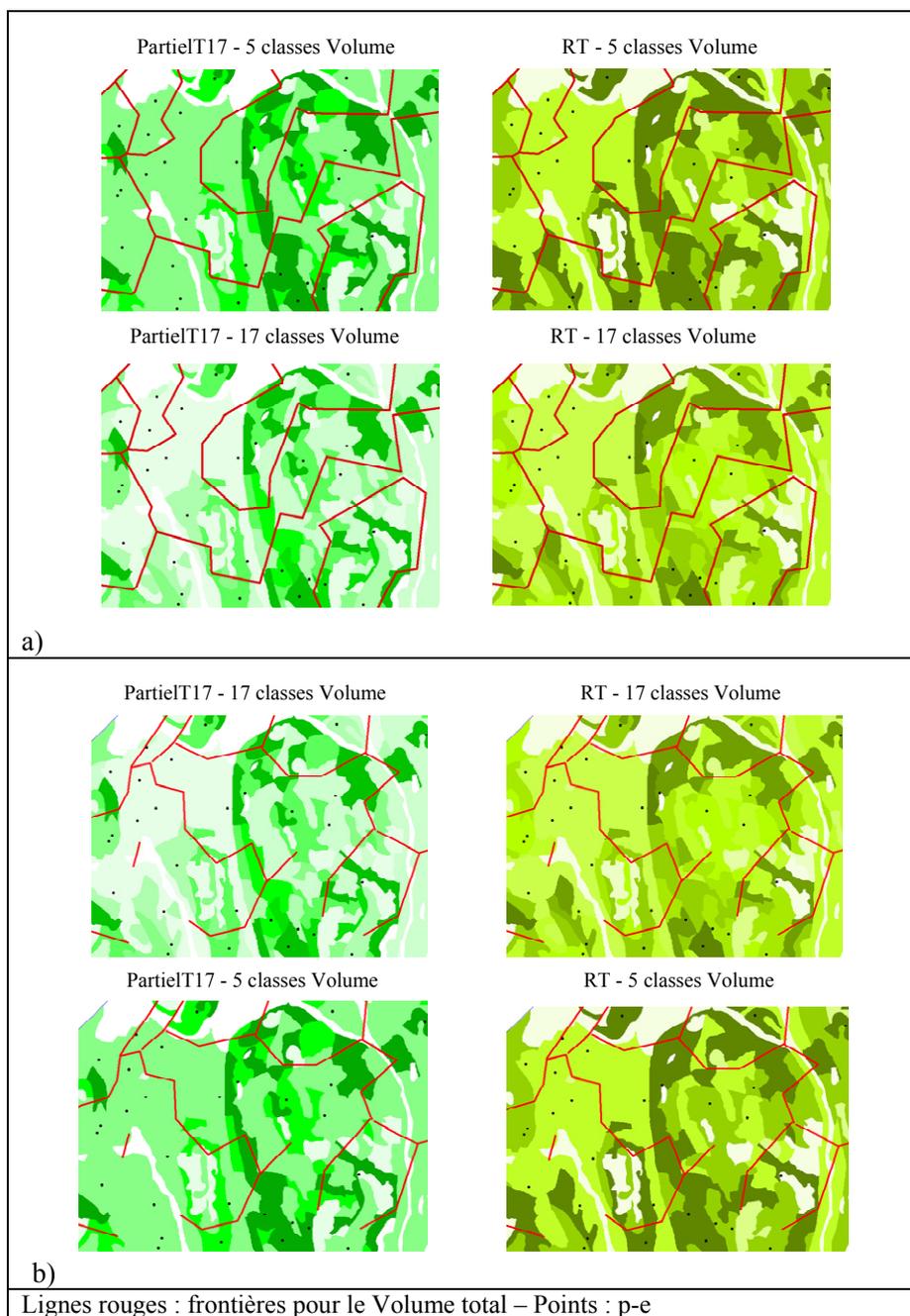


Figure 47. Exemples de la superposition des frontières cartographiques des groupes de p-e avec une même estimation du volume ligneux aux frontières de terrain pour le *Volume total* : a) frontières indirectes et b) frontières directes.

### 3.2. Différences qui ont trait aux méthodes de saisie des variables

Les principales variables utilisées pour définir les strates forestières et pour identifier, par la suite, les peuplements de la carte, sont similaires quant aux aspects suivants : la nature et la résolution thématique des variables notées sur les p-e et utilisées pour l'identification des frontières de terrain (Tableau 44). La résolution thématique réfère au degré de détail des classes ou des catégories.

Malgré ces faits, une différence subtile mais fondamentale concerne les méthodes par lesquelles les informations pour ces variables sont obtenues (Tableau 44). D'une part, les catégories de la carte forestière sont interprétées sur des photos aériennes, c'est-à-dire en analysant le couvert de la végétation ligneuse et en considérant une unité cartographique minimale (Ministère des Ressources naturelles du Québec 1995). D'autre part, les catégories pour les p-e sont estimées visuellement sur le terrain, à partir du sol à l'intérieur de la surface de la p-e (Saucier 1994). Cette différence explique que deux situations soient fréquemment observées lorsque les frontières directes de la variable *Type forestier* sont superposées aux frontières de la carte : (1) deux p-e situées dans deux peuplements séparés par une frontière de fort contraste concordent pour au moins 70 % des variables qui composent le *Type forestier*, c'est-à-dire que de ce point de vue elles sont suffisamment similaires pour ne pas avoir une frontière entre les deux (Figure 48 gauche); (2) deux p-e situées dans le même peuplement ou dans deux peuplements séparés par une frontière de faible contraste, présentent cependant un désaccord pour toutes les variables qui composent le *Type forestier*, c'est-à-dire qu'une frontière forte existe entre elles (Figure 48 droite).

Ces situations peuvent également être attribuées à la faible concordance entre l'information de la carte et du terrain mise en évidence au Chapitre VI, laquelle peut être due à l'hétérogénéité interne des conditions forestières et/ou à la représentativité des p-e qui échantillonnent la strate.

Dans la situation où deux p-e partagent une frontière pour le *Type forestier* et sont situées dans deux peuplements séparés par une frontière de faible contraste, un autre aspect doit être considéré. En effet, les frontières de *Type forestier* indiquent le simple désaccord entre

les catégories de variables, tandis que le contraste des frontières cartographiques quantifie ce désaccord. Par conséquent, même si les catégories de *Type forestier* sont différentes entre les deux p-e et déterminent une frontière entre elles, il peut s'agir des catégories adjacentes (par exemple, les catégories 3-4 de *Classe de hauteur*) et les p-e sont donc associées à une frontière cartographique de faible contraste.

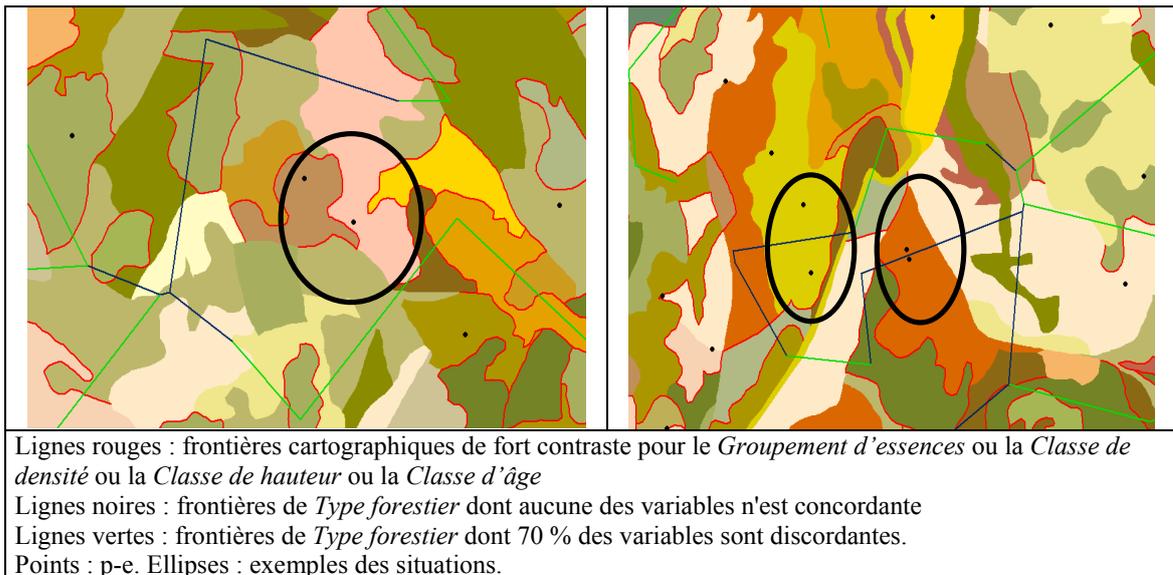


Figure 48. Exemples de situations observées lorsque les frontières cartographiques sont superposées aux frontières directes de *Type forestier*.

### 3.3. Différences liées à la structure des données

La structure des données, qui représentent les frontières dans les deux sources d'information, est différente. Les frontières cartographiques correspondent à une structure vectorielle, donc de lignes, tandis que celles issues des p-e correspondent à une structure de surface, comme la région délimitée par un polygone de Voronoï (Jacquez and Maruca 2001). Cette différence rend difficile l'établissement d'une correspondance entre les deux types de frontières, car les formes des lignes sont habituellement très différentes.

#### **4. Différences liées à l'aspect schématique**

Les différences schématiques concernent les modèles de l'espace auxquels correspondent la carte forestière et les frontières issues des p-e. La carte répond à une conception de l'espace occupé par des objets, tandis que les frontières correspondent à un modèle de champs, dans lequel une variable continue adopte différentes valeurs dans l'espace.

Les différences sont également reliées aux méthodes selon lesquelles les frontières sont identifiées. Celles de la carte forestière le sont par un processus cognitif très complexe et assez subjectif (Aubert 1995) qui repose fortement sur l'expertise de l'interprète ainsi que sur sa connaissance du type de forêt impliqué. La carte résultante ne contient pas d'information sur les caractéristiques des frontières. Les frontières issues des p-e sont, quant à elles, déterminées par des algorithmes basés sur la quantification mathématique et statistique des ressemblances ou des changements dans les valeurs d'une variable sur des p-e adjacentes. De plus, des changements de seuils et des conditions peuvent être ajoutés à ces algorithmes.

#### **5. Conclusion sur la fusion des deux représentations de la forêt**

Les différences énumérées entre les deux représentations de la forêt sont nombreuses et touchent toutes leurs composantes. Les différences sémantiques doivent être surmontées pour pouvoir accomplir la fusion lorsque les éléments concernent les unités forestières représentées dans les deux sources d'information. Cependant, ces différences ne concernent pas les frontières en tant que telles, donc une fusion au niveau des frontières pourrait être envisagée afin de valider la certitude sur l'existence de certaines des frontières cartographiques lorsqu'elles correspondent spatialement aux frontières de terrain.

Toutefois, la différence des résolutions spatiales et de méthodes de saisie des variables pose un obstacle majeur pour cette fusion. De plus, la différence de résolution spatiale augmente la différence de structure des données car une correspondance spatiale plus évidente pourrait être établie entre les frontières vectorielles et celles de polygones de Voronoï si les résolutions spatiales des deux sources étaient plus semblables.

Afin de surmonter cette différence de résolution spatiale, nous avons cherché une correspondance entre les frontières de fort contraste à l'échelle régionale et les frontières de terrain. Dans le même but, nous avons considéré également la correspondance entre les frontières des groupes de peuplements avec une même estimation de volume ligneux et les frontières de terrain. Cependant, dans aucun cas cette correspondance n'a pas pu être établie pour toutes les frontières. Une des raisons peut-être la fragmentation de la Forêt Montmorency qui entraîne des peuplements de faible taille et une variabilité spatiale élevée pour les estimations de volume ligneux.

Les différences de résolution spatiale et de méthode de saisie des variables génèrent une grande différence quant au degré de généralisation des informations des deux sources. En effet, les p-e contiennent de l'information de caractère très local sur la structure de la végétation et le volume ligneux mais leur espacement est beaucoup plus élevé que la taille moyenne des peuplements, ce qui résulte en des frontières très éparées mais correspondant à des situations très locales. Les peuplements de la carte forestière contiennent pour sa part, un certain degré de généralisation pour les mêmes variables, imposé par les normes de photo-interprétation. Ces différences, unies à la forte fragmentation de la forêt, déterminent, dans ce territoire et pour les variables étudiées, une faible correspondance spatiale entre les frontières cartographiques et de terrain et constituent l'obstacle majeur à leur fusion. La Figure 49 schématise ce concept et ses conséquences, c'est-à-dire les situations trouvées lorsque les deux représentations de la forêt sont superposées : (1) la correspondance spatiale entre les frontières (ou la non-existence de frontières) de deux représentations ; (2) des frontières de terrain à l'intérieur d'un peuplement ou d'un groupe de peuplements et (3) l'inexistante d'une frontière de terrain où une frontière cartographique de fort contraste existe. Cette figure reprend le concept de différence d'échelles et d'hétérogénéité entre les peuplements et les p-e de la Figure 36 au Chapitre VI.

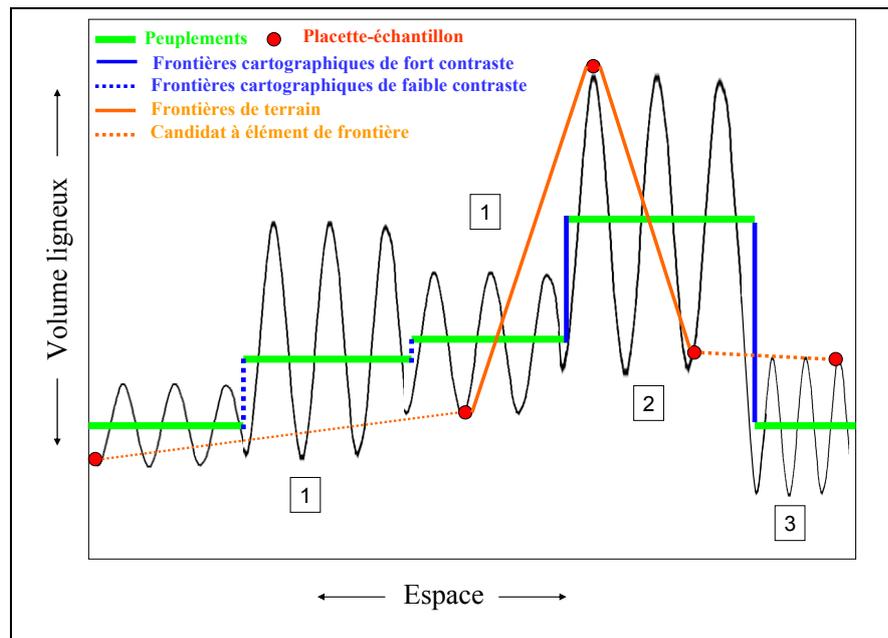


Figure 49. Relation entre les frontières cartographiques et de terrain.

Nonobstant ces faits, la fusion peut être envisagée au niveau des connaissances, basée sur les coïncidences spatiales et non spatiales qui découlent de la ressemblance sémantique entre les frontières des deux sources. Ainsi, les combinaisons d'attributs qui déterminent des frontières de terrain peuvent être transférées et appliquées aux frontières cartographiques. L'existence de ces dernières sera plus certaine dans les cas où le fort contraste déterminé par une combinaison d'attributs cartographiques est renforcé par le fait que la même combinaison d'attributs détermine des frontières lorsque l'information provient des échantillons de terrain. Par exemple, l'existence de frontières entre un peuplement résineux où le sapin prédomine (SS) et un peuplement mixte où le bouleau prédomine (BbS) ou celle de frontières entre un peuplement où les arbres dominants mesurent entre 12 et 17 m et un peuplement où ils mesurent moins de 4 m, est plus certaine du fait que les mêmes combinaisons de catégories déterminent des frontières de terrain.

## Les estimations du volume de matière ligneuse

La fusion des informations de la carte forestière et des p-e concerne également les estimations du volume ligneux. Par rapport à ces dernières, deux sources d'incertitude ont

été identifiées. Une des sources correspond aux estimations et elle est liée à la méthode de regroupement des p-e pour leur calcul (abordée au Chapitre V), alors que l'autre source d'incertitude est liée à la concordance carte/p-e des strates forestières auxquelles ces estimations sont associées (traitée au Chapitre VI).

La fusion des informations sur ces deux types d'incertitude doit prendre en compte le fait que leur origine est due à des processus indépendants. En effet, l'incertitude des estimations trouve son origine dans le processus analytique de la définition de critères pour regrouper les p-e et dans le nombre de groupes à former. Ce processus est incontournable pour obtenir des estimations de la moyenne et de la variance du volume ligneux qui soient applicables à des unités spatiales plus étendues que les p-e où les données dendrométriques ont été obtenues.

Quant à l'incertitude reliée à la concordance carte/p-e des strates forestières, elle comprend des aspects écologiques, géomatiques et analytiques. Du point de vue écologique, la forêt est hétérogène et la magnitude de l'hétérogénéité dépend de l'échelle spatiale considérée. Ainsi, les unités où la forêt est découpée pour permettre sa description et sa gestion contiennent toujours une certaine hétérogénéité issue des processus écologiques d'une échelle spatiale plus fine. L'aspect géomatique, quant à lui, est lié à la définition de l'unité cartographique minimale, ce qui implique que les zones hétérogènes de taille inférieure à cette unité minimale y soient intégrées. Il implique aussi, la différence de grain des polygones de la carte et des p-e, qui signifie que l'information des p-e est très locale comparativement à celle de la carte. Finalement, du point de vue analytique, les critères de la méthode de classification imposent une certaine uniformité, tout en permettant une diversité de situations, comme mentionné dans la discussion sur les groupements d'essences au Chapitre III et au Chapitre VI. De plus, l'impossibilité de recenser la forêt sur le terrain pour tenir compte de son hétérogénéité requiert des échantillonnages des strates forestières afin d'obtenir les données nécessaires pour estimer le volume ligneux. Ces sondages entraînent des problèmes liés à la taille de l'échantillon (ici le nombre de p-e de chaque strate) qui peut être plus ou moins représentatif des conditions forestières de la strate.

La considération de la représentation spatiale de chaque type d'incertitude permet de déterminer sur quel plan leur fusion est possible. En effet, bien que les phénomènes représentés soient différents, les deux représentations spatiales sont basées sur la même carte forestière. Dans le cas de l'incertitude liée aux strates, l'unité spatiale est le peuplement, tandis que dans le cas de la variabilité des estimations, l'unité spatiale est un groupe de peuplements adjacents avec une même estimation du volume ligneux. Par conséquent, les aspects syntactiques et schématiques des deux représentations sont semblables, ce qui rend possible leur mise en correspondance spatiale et leur combinaison.

Les deux représentations peuvent donc être considérées ensemble et l'estimation de chaque groupe de peuplements peut être pondérée par le degré de concordance locale carte/p-e des strates auxquelles appartiennent ces peuplements (Figure 50). La forme de la combinaison des deux informations doit être explorée afin de trouver celle qui reflète le mieux les magnitudes des deux types d'incertitude dans la valeur résultante. La nouvelle carte créée peut être considérée comme de meilleure qualité que les cartes individuelles, non parce que les estimations soient plus précises en général, sinon parce qu'elles intégreront les deux types d'incertitude et elles seront, par conséquent, plus proches de la réalité de la forêt.

Si la fusion des connaissances est envisagée, des liens non spatiaux peuvent être établis entre les deux types d'incertitude à partir de la considération des attributs autant des groupes de p-e auxquels correspondent les estimations que des strates forestières. En ce sens, selon les résultats présentés dans cette thèse, les groupes de p-e identifiés comme ayant peu de variabilité sont ceux où prédomine la végétation résineuse dominée par le sapin baumier avec un couvert forestier variant entre 41 et 80 % de la surface, où les arbres dominants ont une hauteur variant entre 7 et 12 m, où l'âge estimé se situe entre 50 et 70 ans et qui n'ont pas subi de perturbation. Les mêmes caractéristiques correspondent aux strates forestières les plus concordantes. Par conséquent, les estimations du volume de ces strates seraient les plus précises de tous les types présents à la Forêt Montmorency. À l'opposé, les estimations correspondant aux types de végétation en régénération seraient les moins précises.

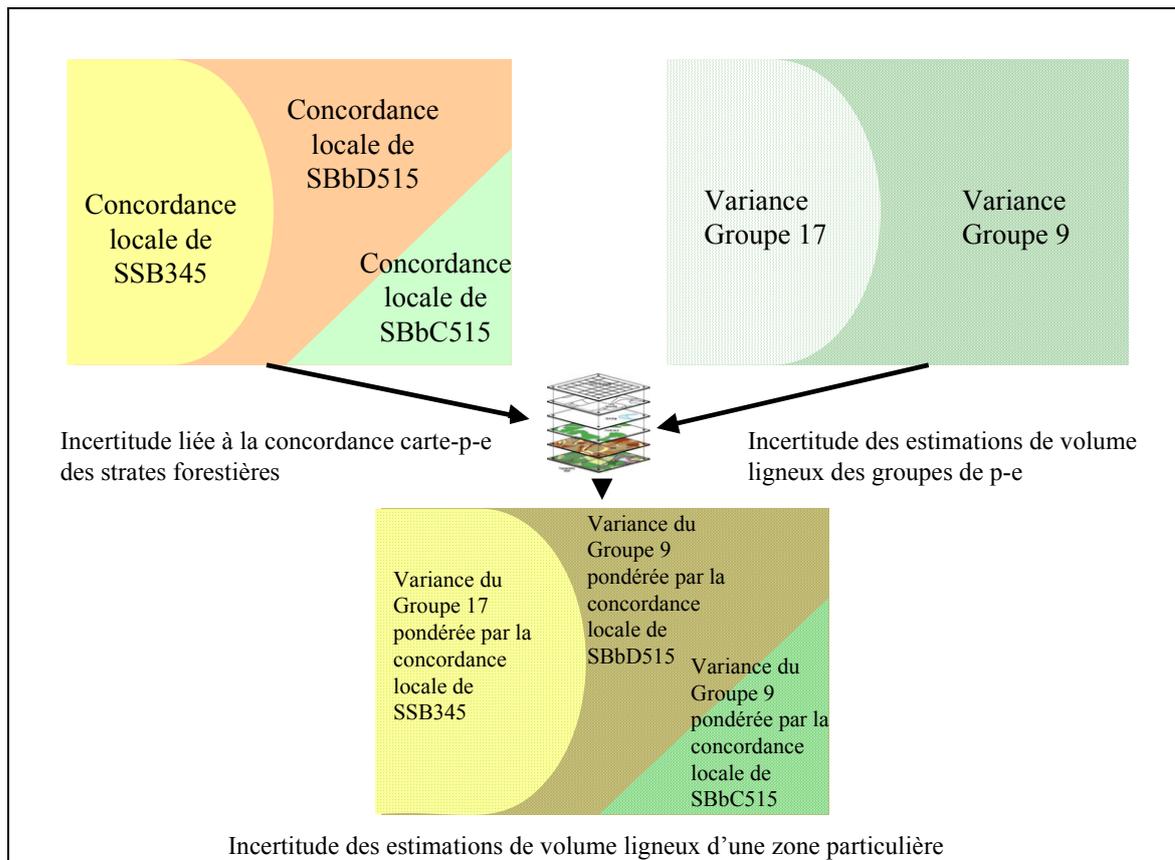


Figure 50. Procédure proposée pour la fusion des informations sur l'incertitude sur les estimations du volume ligneux avec celles liées à la concordance carte/p-e des strates forestières.

## Résumé du chapitre

La discussion développée dans ce chapitre permet de conclure que la fusion des informations spatiales de deux représentations de la Forêt Montmorency n'est pas possible sur le plan des éléments car elle exige une solution préalable de l'hétérogénéité des résolutions spatiales ainsi que de la conceptualisation des unités délimitées dans chaque représentation. Cependant, une fusion des connaissances est possible à partir de l'identification de coïncidences entre les combinaisons d'attributs qui déterminent des frontières dans les deux sources d'information. Quant aux estimations du volume de matière ligneuse des strates forestières, les deux types d'incertitude qui les affectent doivent être pris en compte afin d'obtenir des valeurs plus proches de la réalité lors du calcul du

volume ligneux disponible dans une zone donnée. Les représentations de ces deux types d'incertitude peuvent être fusionnées sur une base de correspondance spatiale, tandis qu'au niveau des connaissances, certaines règles peuvent être mises en évidence à partir de l'analyse des coïncidences entre les attributs des groupes de p-e à partir desquels sont issus les estimations et les attributs des strates forestières de faible et de forte concordance carte/p-e.

## Conclusions

La carte forestière est un outil de planification des interventions forestières sur un territoire, surtout pour l'extraction des ressources ligneuses. Dans la carte forestière traditionnelle du Québec, les frontières entre les peuplements sont toutes représentées par des lignes identiques, tandis que les strates forestières sont considérées comme bien représentées par les échantillons de terrain. Par conséquent, les incertitudes sur ces informations cartographiques affectent directement les décisions qui sont basées sur ces informations. Ces incertitudes concernent principalement la magnitude des surfaces des peuplements et les estimations du volume ligneux des strates auxquelles les peuplements appartiennent. La planification opérationnelle des interventions forestières nécessite de l'information sur ces incertitudes à une échelle locale, donc à l'échelle d'un peuplement en particulier ou d'une strate forestière précise.

Le but global de cette étude consistait à explorer comment améliorer la fiabilité locale des estimations du volume de matière ligneuse, de même que celles portant sur les surfaces de la carte forestière. Pour atteindre ce but, différentes approches analytiques ont été explorées afin de caractériser les éléments de la carte forestière, les frontières des peuplements et les strates forestières, en fonction des aspects liés à l'incertitude sur l'existence des frontières dans la réalité, ainsi que sur la concordance entre l'information de terrain et cartographique des strates. Les traitements analytiques employés sur les deux sources d'information ont cherché à mettre en évidence la réalité écologique qu'elles représentent, ainsi qu'à explorer les contraintes et possibilités qui découlent des aspects géomatiques de chacune d'elles, en visant leur fusion pour obtenir de l'information plus fiable à l'échelle locale.

En ce qui concerne les frontières de la carte forestière, les méthodes utilisées dans les analyses fournissent des informations complémentaires utiles qui permettent d'évaluer le degré d'incertitude sur l'existence des frontières. La quantification du contexte des peuplements permet de déterminer pour quels secteurs du territoire la végétation des peuplements voisins présente une grande diversité structurale, alors que la quantification du

contraste local des frontières indique la magnitude de ces différences, que les variables concernées soient qualitatives ou quantitatives. La quantification du contraste à l'échelle du voisinage met en évidence, quant à elle, les frontières à fort contraste entre deux groupes de peuplements. La quantification du contexte des frontières permet d'identifier les groupes de frontières qui présentent des valeurs de contraste « anormales », soit parce que ce contraste est très faible ou très élevé par rapport au contraste moyen. L'analyse intégrée du contraste et du contexte spatial des frontières permet de détecter les secteurs de la forêt où l'identification des frontières serait plus facile, notamment là où des frontières de fort contraste sont entourées de frontières de faible contraste.

Dans cette étude, 51 % des frontières présentent un contraste élevé pour l'une ou l'autre des variables considérées et autour de 11 % des frontières ont un contraste élevé pour l'ensemble des variables. Les variables *Groupement d'essences* et *Classe d'âge* présentent la plus grande proportion de frontières fortes (50 % et 40 % des frontières respectivement), tandis que la *Classe de hauteur* comporte, pour sa part, la plus grande proportion de frontières faibles (19 %). De plus, environ 65 % des peuplements présentent un contexte de forte hétérogénéité pour toutes les variables. Quant à la relation entre le contraste, le contexte et la probabilité d'existence des frontières, les résultats montrent une relation significative entre le contraste de toutes les variables, la longueur de la frontière et le contexte de *Classe d'âge*, ainsi que la probabilité d'existence d'une frontière, mesurée par la consistance. En effet, ces résultats indiquent que les frontières fortes ont tendance à présenter une consistance élevée.

Quant aux données de terrain, les deux méthodologies utilisées pour identifier les frontières mettent en évidence des aspects différents de la répartition spatiale des variables étudiées. La méthode de regroupement des p-e à contraintes spatiales, applicable seulement aux variables quantitatives, permet de découper le territoire à l'étude en zones homogènes pour les variables concernées et ce, à différentes échelles selon le nombre de groupes formés. De ce fait, les zones qui sont tellement distinctes de leur voisinage qu'elles persistent lorsque le grain spatial augmente peuvent être identifiées. La méthode de détection de changements, quant à elle, permet d'identifier les discontinuités des variables quantitatives et qualitatives.

Les emplacements où les conditions forestières ou environnementales particulières créent ou maintiennent les discontinuités sont ainsi détectés. Cependant, la méthode demeure limitée en ce qui concerne les variables qualitatives, car les discontinuités représentent le désaccord entre les catégories de deux p-e et non la dissimilitude entre ces catégories. L'analyse de la structure spatiale locale des frontières aide également à identifier les zones particulières du territoire où elles sont concentrées. Cette information complète celle de la localisation des frontières puisqu'elle ajoute la dimension du voisinage spatial à celle de l'intensité du changement de chaque frontière. Finalement, l'analyse de recouvrement permet d'identifier, à l'échelle déterminée par l'espacement entre les échantillons, les associations spatiales entre les frontières de différentes variables.

À la Forêt Montmorency, les frontières du volume ligneux issues du regroupement des p-e montrent une association spatiale avec les frontières des variables forestières et environnementales. Fréquemment, les deux types de frontières, celles des groupes de p-e et celles qui indiquent des discontinuités intenses, coïncident dans l'espace surtout lorsqu'il s'agit de différences marquées, soit des p-e isolées, soit de groupes de p-e, par rapport au voisinage. De plus, les frontières du volume ligneux sont associées à des situations où les p-e adjacentes correspondent à différentes étapes du développement de la végétation.

En ce qui concerne la variabilité des estimations du volume ligneux calculées à partir des données de terrain, les arbres de régression offrent une technique intéressante pour explorer la structure des observations forestières afin de mettre en évidence les variables les plus importantes dans les données pour la formation des groupes de p-e dont la variabilité par rapport au volume ligneux est minimale. Ils permettent également de rendre explicite un processus qui repose actuellement sur la connaissance locale et l'expérience des experts, d'une valeur inestimable et difficilement transférable.

Trois arbres de régression ont été construits à partir des données du volume ligneux des p-e. Ceux qui utilisent les variables estimées sur le terrain sont les plus performants puisqu'ils expliquent environ 69 % de la variabilité des observations. La variabilité des estimations des groupes de p-e formés par ces modèles est inférieure à celle des groupes formés par la méthode traditionnelle employée au Québec. Cette variabilité inférieure concerne

également la structure spatiale des estimations. Les variables les plus importantes qui permettent de diviser les observations selon de forts changements de variabilité du volume ligneux sont celles liées à la taille des arbres dans la forêt et au temps écoulé depuis la dernière coupe. D'autres variables importantes concernent le couvert de la végétation ligneuse, la pente et la composition de la végétation. Les groupes de p-e de faible hétérogénéité correspondent à un type de végétation caractéristique des étapes de pré-maturité de la forêt, alors que ceux de forte hétérogénéité présentent des propriétés correspondant à la végétation forestière en régénération à la suite d'une perturbation sévère.

Pour les strates forestières, les indices de concordance carte/p-e utilisés ont permis de quantifier l'accord entre l'information du terrain et l'information cartographique des strates et ce, à deux échelles. La concordance locale peut être attribuée au degré d'homogénéité interne des strates et à la représentativité des p-e qui les échantillonnent. La concordance à l'échelle du voisinage, quant à elle, peut être attribuée d'une part, à l'incertitude positionnelle autant des p-e que des frontières des strates forestières ainsi qu'aux conditions structurales des frontières. D'autre part, elle peut refléter une certaine homogénéité des attributs à cette échelle. L'application des deux indices permet de conclure que l'incertitude qui est liée à cette concordance entre les deux types d'information est particulière à chacune des strates.

Les résultats obtenus indiquent que les gradients de concordance sont applicables aux catégories de variables examinées plutôt qu'aux strates forestières. Ainsi, c'est la concordance carte/p-e des catégories correspondant à la végétation dominée par le sapin baumier, à un couvert moyen de végétation ligneuse, aux arbres variant entre 12 et 17 m et à un âge estimé à 45 ans, qui est la plus élevée. Pour les strates résineuses où le sapin baumier prédomine, qui sont les plus importantes en surface à la Forêt Montmorency, la concordance semble être liée aux étapes de développement de la végétation ligneuse. En effet, l'étape d'initiation du peuplement présente une faible concordance en raison d'une forte hétérogénéité due à la diversité d'essences, tandis que les étapes intermédiaires, de pré-maturité, présentent la concordance maximale. Finalement, l'étape de maturité présente à nouveau une faible concordance, possiblement en raison de l'accumulation de

perturbations légères qui génèrent de l'hétérogénéité dans l'environnement et dans la structure de la végétation ligneuse.

Finalement, la fusion des informations spatiales des deux représentations de la forêt n'est pas possible sur le plan des éléments, car les différences de résolution spatiale, de méthodes de saisie des variables et du degré de généralisation rendent très difficile l'établissement d'une correspondance spatiale entre les frontières des variables examinées dans une forêt très fragmentée. Malgré ces faits, une fusion des connaissances est possible à partir de l'identification des combinaisons d'attributs qui déterminent des frontières sur le terrain, lesquelles peuvent être transférées et appliquées aux frontières cartographiques. Quant à l'incertitude des estimations du volume ligneux des strates forestières, autant celle liée aux estimations que celle liée à la concordance carte/p-e, elles doivent être prises en compte afin d'obtenir des valeurs plus précises lors du calcul du volume ligneux pour une zone donnée. Une manière d'intégrer ces deux types d'incertitude peut reposer sur la fusion des informations basée sur la correspondance spatiale entre la carte forestière et la carte qui représente la répartition spatiale des groupes de p-e.

## **Futures avenues de recherche**

La recherche développée a permis d'explorer, à l'aide de différentes méthodes analytiques, certains aspects de la carte forestière et des données de terrain afin d'extraire de l'information utile pour mieux caractériser l'incertitude spatiale des éléments de cette carte et donc pour mieux décrire les écosystèmes forestiers. Cependant, de nombreux aspects demeurent encore ouverts à l'exploration.

En ce qui concerne les méthodes de quantification du contraste entre les peuplements, une possibilité intéressante à explorer consiste à appliquer le même indice écologique utilisé pour le contraste entre deux peuplements, c'est-à-dire à l'échelle locale, aux voisinages à des échelles diverses. Cette possibilité permettrait, d'une part, de mieux comparer les contrastes global et régional et, d'autre part, de rendre possible l'extension de l'analyse aux variables nominales, comme le *Groupement d'essences*. Un autre aspect à explorer est celui

de considérer le voisinage dans l'algorithme pour déterminer la force des frontières. En effet, dans sa forme actuelle les résultats sont très dépendants du nombre de voisins, lequel dans une situation réelle comme celle de la carte forestière est très variable. Un autre aspect à explorer concerne le transfert des différences d'hétérogénéité du contexte des peuplements aux frontières, afin d'intégrer cet aspect à l'incertitude sur l'existence de ces dernières. Une meilleure détermination de la consistance des frontières peut contribuer également à améliorer sa relation avec les variables indépendantes utilisées.

Quant aux méthodes de détection des frontières utilisées pour les données de terrain, l'utilisation d'un indice similaire à celui employé pour la quantification du contraste des frontières cartographiques permettrait de déterminer les frontières entre variables qualitatives de manière plus raffinée que la simple concordance entre les catégories. Les frontières obtenues seraient, par conséquent, plus faciles à comparer avec leurs équivalentes cartographiques.

L'utilisation d'autres sources de données forestières, telles que les images de très haute résolution spatiale (comme Ikonos et QuickBird), permettrait de disposer de l'information pour analyser, par exemple, la structure spatiale de la végétation ligneuse à l'intérieur des strates forestières ainsi que les caractéristiques des frontières, comme la largeur ou le degré de netteté. Cette information permettrait ainsi de quantifier l'homogénéité des strates pour mieux cerner la source d'incertitude des estimations du volume ligneux.

Le développement d'une ontologie pour les paysages forestiers qui prendrait en compte les informations des différentes sources de données constitue une voie possible pour atteindre la fusion de ces sources de données sur le plan des unités spatiales.

## **Utilité de cette recherche pour la planification des activités forestières**

Cette recherche est basée sur l'information disponible à partir de la carte forestière et les données de terrain obtenues dans le cadre d'un inventaire forestier. Les méthodes

analytiques employées sont facilement applicables à ces informations. De la même manière, les résultats obtenus permettent d'approfondir la connaissance sur les patrons spatiaux des variables forestières examinées dans le territoire étudié, connaissance qui est nécessaire pour la prise de décisions plus éclairées sur la forêt.

La caractérisation des frontières de la carte forestière en ce qui concerne leur contraste, ainsi que la détection des frontières à partir des données de terrain permettent de déterminer les secteurs du territoire qui présentent des caractéristiques particulières au plan de l'hétérogénéité des attributs ou à la présence de groupes de frontières de fort contraste, par exemple.

Les arbres de régression constituent, pour leur part, une méthode puissante pour explorer la structure des variables d'intérêt (comme le volume ligneux) estimées sur les p-e, afin de mettre en évidence les variables que dans les données sont importantes pour regrouper les p-e en groupes homogènes. Les résultats de l'application de cette méthode sont relativement simples à interpréter et à exprimer de façon spatiale. Ces résultats permettent également d'explicitier les critères utilisés pour le regroupement de p-e. De cette façon, l'utilisation d'autres critères peut être explorée et les résultats comparés afin d'obtenir un ensemble de règles qui soit optimal pour un territoire donné.

Finalement, l'application d'une méthode pour quantifier la concordance carte/p-e des strates forestières permet de caractériser ces strates et de considérer de façon explicite cette source d'incertitude qui affecte fortement les estimations du volume ligneux.

Cette recherche a ciblé les cartes forestières étant donné l'impact économique associé aux incertitudes de cet outil sur les activités forestières. Cependant, l'information à caractère local obtenue sur la forêt peut être utilisée à d'autres fins, comme pour l'aménagement écologique de la forêt, la conservation de l'habitat faunique et la conservation des écosystèmes particuliers.

## Références bibliographiques

- Alesheikh, A., J. Blais, M. Chapman and H. Karimi. 1998. Rigorous geospatial data uncertainty models for GIS. Lowell, K. and A. Jaton (Ed.). *Third International Symposium on Spatial accuracy assessment in natural resources and environmental sciences*. Québec. Ann Arbor Press. pp.: 195-202.
- Allen, T. F. H. and T. W. Hoekstra. 1992. *Toward an Unified Ecology*. Columbia University Press, New York. 384 pp.
- Altman, D. 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *International Journal Geographical of Information Systems*, 8 (3): 271-289.
- Anselin, L. 1995. Local indicators of spatial association - LISA. *Geographical Analysis*, 27 (2): 93-115.
- Anselin, L. 1999. Interactive techniques and exploratory spatial data analysis. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 253-266.
- Aubert, É. 1995. Quantification de l'incertitude spatiale en photo-interprétation forestière à l'aide d'un sig pour le suivi spatio-temporel des peuplements. Thèse de maîtrise. Faculté de Foresterie et Géomatique, Laval. Québec. 92 pp.
- Barbujani, G., N. L. Oden and R. R. Sokal. 1989. Detecting regions of abrupt change in maps of biological variables. *Systematic Zoology*, 38: 376-389.
- Barnes, B. V., D. R. Zak, S. R. Denton and S. H. Spurr. 1998. *Forest Ecology*. John Wiley & Sons, Inc., 774 pp.
- Beard, K. 2001. Roles of meta-information in uncertainty management. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 363-378.
- Beaulieu, P. 1994. Étude de l'autocorrélation spatiale sur un territoire forestier en vue d'établir une méthodologie de regroupement des strates de base. Thèse de Maîtrise. École de Gradués, Université Laval. Québec. 125 pp.
- Beaulieu, P. and K. Lowell. 1994. Spatial autocorrelation among forest stands identified from the interpretation of aerial photographs. *Landscape & Urban Planning*, 29 (2-3): 161-169.
- Bélangier, L., P. Bouliane et J.-P. Jetté. 1990. Plan général d'aménagement de la Forêt Montmorency. Le plan de production de bois (Révision 1990)Annexe IV. Université Laval. Québec. pp.: 22 pp.

- Bélangier, L. et P. Bouliane. 1992. *Plan général d'aménagement de la forêt Montmorency (Révision 1990)*. Département des sciences forestières, Faculté de foresterie et de géomatique, Université Laval, Québec. 215 pp.
- Benslimane, D., E. Leclerc, M. Savonnet, M. Terrasse and K. Yétongnon. 2000. On the definition of generic multi-layered ontologies for urban applications. *Computers, Environment and Urban Systems*, 24: 191-214.
- Bérard, J. A. et M. Côté. 1996. *Manuel de Foresterie*. L'Ordre des ingénieurs forestiers du Québec - Les Presses de l'Université Laval, Québec. 1428 pp.
- Bergeron, Y. and P. R. Dansereau. 1993. Predicting forest composition under different fire cycles in the southeastern boreal forest of Canada. *Journal of Vegetation Science*, 4: 827-832.
- Bergeron, Y. 2000. Species and stand dynamics in the mixed woods of Quebec's southern boreal forest. *Ecology*, 81 (6): 1500-1516.
- Bergeron, Y., S. Gauthier, V. Kafka, P. Lefort and D. Lesieur. 2001. Natural fire frequency for the eastern Canadian boreal forest: consequences for sustainable forestry. *Canadian Journal of Forest Research*, 31: 384-391.
- Bian, L. 1997. Multiscale nature of spatial data in scaling up environmental models. In Quattrochi, D. A. and M. F. Goodchild (Ed.), *Scale in Remote Sensing and GIS*. CRC Press, Lewis Publishers, Inc. pp.: 13-26.
- Bittner, T. and G. Edwards. 2001. Toward an ontology for geomatics. *Geomatica*, 55 (4): 475-490.
- Boots, B. and M. Tiefelsdorf. 2000. Global and local spatial autocorrelation in bounded regular tessellations. *Journal of Geographical Systems*, 2 (4): 319-348.
- Boots, B. 2002. Local measures of spatial association. *Ecoscience*, 9 (2): 168-176.
- Boyce, S. G. 1995. *Landscape Forestry*. John Wiley & Sons, New York. 239 pp.
- Bradshaw, G. A. and M.-J. Fortin. 2000. Landscape heterogeneity effects on scaling and monitoring large areas using remote sensing data. *Geographic Information Sciences*, 6 (1): 61-68.
- Breiman, L., J. H. Friedman, R. A. Olshen and C. J. Stone (Ed.). 1984. *Classification and Regression Trees*. Wadsworth International Group, Belmont, California. 358 pp.
- Burrough, P. A. 1996. Natural objects with indeterminate boundaries. In Burrough, P. A. and A. U. Frank (Ed.), *Geographic Objects with Indeterminate Boundaries*. Taylor & Francis Ltd. pp.: 3-28.

- Burrough, P. A. and R. A. McDonnell. 1998. *Principles of Geographical Information Systems*. Oxford University Press Inc., New York. 333 pp.
- Byron, N. 2003. Relancer la foresterie mondiale. *XII Congrès forestier mondial*. Québec. pp.: 10-18.
- Cadenasso, M. L., S. T. A. Pickett, K. C. Weathers, S. S. Bell, T. L. Benning, M. M. Carreiro and T. E. Dawson. 2003a. A interdisciplinary and synthetic approach to ecological boundaries. *Bioscience*, 53 (8): 717-722.
- Cadenasso, M. L., S. T. A. Pickett, K. C. Weathers and C. G. Jones. 2003b. A framework for a theory of ecological boundaries. *Bioscience*, 53 (8): 750-751.
- Cairns, D. M. 2001. A comparaison of methods for predicting vegetation type. *Plant Ecology*, 156: 3-18.
- Chambers, J. M. and T. J. Hastie. 1992. *Statistical Models in S*. Wadsworth & Brooks, Pacific Grove. 608 pp.
- Chrisman, N. R. 1991. The error component in spatial data. In Maguire, D. J., M. F. Goodchild and D. W. Rhind (Ed.), *Geographical Information Systems. Principles and Applications*. Longman Scientific and Technical. New York. pp.: 165-174.
- Clark, L. A. and D. Pregibon. 1992. Classification and regression trees. In Chambers, J. M. and T. J. Hastie (Ed.), *Statistical Models in S*. Wadsworth & Brooks. Pacific Grove. pp.: 377-419.
- Conseil canadien des ministres des forêts. 2004. Terminologie de la sylviculture au Canada. Glossaire. Programme national de données sur la forêt. 2004. [http://nfdp.ccfm.org/silviterm/silvi\\_f/silvitermetohf.htm#F](http://nfdp.ccfm.org/silviterm/silvi_f/silvitermetohf.htm#F).
- Côté, M. (Ed.). 2003. *Dictionnaire de la Foresterie*. Ordre des ingénieurs forestiers du Québec, Québec. 744 pp.
- Coulombe, S. and K. E. Lowell. 1995. Ground-truth verification of relations between forest basal area and certain ecophysiographic factors using a geographic information system. *Landscape & Urban Planning*, 32 (2): 127-136.
- Coulombe, S. 1996. Étude des relations entre les caractéristiques dendrométriques de la forêt et certains facteurs forestiers et éco-physiographiques estimés à l'aide d'un S.I.G. Thèse de maîtrise. Faculté de Foresterie et Géomatique, Université Laval. Québec. 156 pp.
- Csillag, F., M.-J. Fortin and J. Dungan. 2000. On the limits and extensions of the definition of scale. *Bulletin of the Ecological Society of America*, 81: 230-232.

- Csillag, F., B. Boots, M.-J. Fortin, K. Lowell and F. Potvin. 2001. Multiscale characterization of boundaries and landscape ecological patterns. *Geomatica*, 55 (4): 509-522.
- de Blois, S., G. Domon and A. Bouchard. 2001. Environmental, historical and contextual determinants of vegetation cover: a landscape perspective. *Landscape Ecology*, 16: 421-436.
- De Groeve, T. 1999. L'incertitude spatiale dans la cartographie forestière. Thèse de doctorat. Département de Sciences Géomatiques, Faculté de Foresterie et Géomatique, Laval. Québec. 217 pp.
- De Groeve, T. and K. E. Lowell. 2000. Improving local forest volume estimates by fusion of multi-temporal forest type maps. *Environmental Modelling and Software*, 15 (4): 373-385.
- De Groeve, T. and K. E. Lowell. 2001. Boundary uncertainty assessment from a single forest-type map. *Photogrammetric Engineering and Remote Sensing*, 67 (6): 717-726.
- De'Ath, G. and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81 (11): 3178-3192.
- Delcourt, H. R. and P. A. Delcourt. 1988. Quaternary landscape ecology: Relevant scales in space and time. *Landscape Ecology*, 2 (1): 23-44.
- Déry, S., L. Bélanger, S. Marchand and S. Côté. 2000. Succession après épidémie de la tordeuse des bourgeons de l'épinette (*Choristoneura fumiferana*) dans des sapinières boréales pluviales de seconde venue. *Canadian Journal of Forest Research*, 30: 801-816.
- di Castri, F. and A. J. Hansen. 1992. The environment and development crises as determinants of landscape dynamics. In Hansen, A. J. and F. Di Castri (Ed.), *Landscape Boundaries. Consequences for Biotic Diversity and Ecological Flows*. Springer-Verlag. New York. pp.: 3-18.
- Duckham, M., K. Mason, J. Stell and M. Worboys. 2001. A formal approach to imperfection in geographic information. *Computers, Environment and Urban Systems*, 25: 89-103.
- Dungan, J. L., J. N. Perry, M. R. T. Dale, P. Legendre, S. Citron-Pousty, M.-J. Fortin, A. Jakomulska, M. Miriti and M. S. Rosenberg. 2002. A balanced view of scale in spatial statistical analysis. *Ecography*, 25 (5): 626-640.
- Edwards, G. 1994. Modeling of fuzzy data: aggregation and desaggregation of fuzzy polygons for spatio-temporal modeling. *Advanced Geographic Data Modeling Workshop (AGDM'94): Spatial Data Modeling and Query Languages for 2D and 3D applications*. Delft, Netherlands. pp.: 141-154.

- Edwards, G. and K. Lowell. 1996. Modelling uncertainty in photointerpreted boundaries. *Photogrammetric Engineering & Remote Sensing*, 62 (4): 337-391.
- Edwards, G., R. Jeansoulien, Y. Bédard, S. Rouabah, F. Savopol, J. Pouliot, F. Vincent, M. Mostafavi, J. Brodeur, T. D. Groeve, M. Plante, R. Devillers, M. Gervais et G. Cotteret. 2000. Groupe de discussion sur la fusion de données. Université Laval. Québec. pp.: pp.
- Edwards, G. and R. Jeansoulin. 2004. Data fusion - from a logic perspective with a view to implementation. *International Journal Geographical Information Science*, 18 (4): 303-307.
- Fagan, W. F., M.-J. Fortin and C. Soykan. 2003. Integrating edge detection and dynamic modeling in quantitative analyses of ecological boundaries. *Bioscience*, 53 (8): 730738.
- Fischer, M. M. 1999. Spatial analysis: retrospect and prospect. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 283-292.
- Fisher, P. 1997. Concepts and paradigms of spatial data. In Craglia, M. and H. Couclelis (Ed.), *Geographic Information Research : Bridging the Atlantic*. Taylor & Francis Ltd. pp.: 297-307.
- Fisher, P. 1998. Improving error models for elevation digital data. In Goodchild, M. F. and R. Jeansoulin (Ed.), *Data Quality in Geographic Information, From Error to Uncertainty*. Ed. Hermes. pp.: 55-62.
- Fisher, P. 2001. Alternative set theories for uncertainty in spatial information. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 351-362.
- Fisher, P. F. 1999. Models of uncertainty in spatial data. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 191-205.
- Fonseca, F., M. Egenhofer, C. Davis and K. Borges. 2000. Ontologies and knowledges sharing in urban GIS. *Computers, Environment and Urban Systems*, 24: 251-271.
- Fonseca, F., J. Martin and A. Rodriguez. 2002. From geo to eco-ontologies. Egenhofer, M. and D. Mark (Ed.). *Geographic Information Science-Second International Conference Giscience 2002*. Lecture Notes in Computer Science, Springer Verlag. pp.: 93-107.
- Fonseca, F. and J. Martin. 2004. Space and time in eco-ontologies. *AI Communications - The European Journal on Artificial Intelligence* 17(4): pp. 259-269.
- Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80: 185-201.

- Forman, R. T. T. 1995. *Land Mosaics : the Ecology of Landscapes and Regions*. Cambridge University Press, Cambridge. 632 pp.
- Fortin, M.-J. 1994. Edge detection algorithms for two-dimensional ecological data. *Ecology*, 75 (4): 956-965.
- Fortin, M.-J. and P. Drapeau. 1995. Delineation of ecological boundaries: comparison of approaches and significance tests. *Oikos*, 72 (3): 323-332.
- Fortin, M.-J., P. Drapeau and G. M. Jacquez. 1996. Quantification of the spatial co-occurrences of ecological boundaries. *Oikos*, 77 (1): 51-60.
- Fortin, M.-J. 1997. Effects of data types on vegetation boundary delineation. *Canadian Journal of Forest Research*, 27 (11): 1851-1858.
- Fortin, M.-J. 1999. Effects of sampling unit resolution on the estimation of the spatial autocorrelation. *Ecoscience*, 6: 636-641.
- Fortin, M.-J., R. J. Olson, S. Ferson, L. Iverson, C. Hunsaker, G. Edwards, D. Levine, K. Butera and V. Klemas. 2000. Issues related to the detection of boundaries. *Landscape Ecology*, 15 (5): 453-466.
- Fortin, M.-J., M. R. T. Dale and J. v. Hoef. 2002. Spatial analysis in ecology. In El-Shaarawi, A. and W. W. Piegorsch (Ed.), *Encyclopedia of Environmetrics*. Wiley. Chichester. pp.: 2051-2058.
- Friedl, M. A., K. C. McGwire and D. M. McIver. 2001. An overview of uncertainty in optical remotely sensed data for ecological applications. In Hunsaker, C. T., M. F. Goodchild, M. A. Friedl and T. J. Case (Ed.), *Spatial Uncertainty in Ecology. Implications for Remote Sensing and GIS Applications*. Springer-Verlag. New York. pp.: 258-283.
- Garrard, C. 2002. StatMod. <http://bioweb.usu.edu/gistools/statmod/>. 10/05/2005.
- Getis, A. 1999. Spatial statistics. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 239-251.
- Gong, P. 1994. Integrated analysis of spatial data from multiple sources: an overview. *Canadian Journal of Remote Sensing*, 20 (4): 349-359.
- Goodchild, M. and G. J. Hunter. 1997. A simple positional accuracy measure for linear features. *International Journal Geographical Information Science*, 11 (3): 299-306.
- Goodchild, M. F. 1989. Modeling error in objects and fields. In Goodchild, M. F. and S. Gopal (Ed.), *The Accuracy of Spatial Databases*. Taylor & Francis. pp.: 107-113.
- Goodchild, M. F., G. Sun and S. Yang. 1992. Development and test of an error model for categorical data. *International Journal of Geographical information Systems*, 6 (2): 87-104.

- Goodchild, M. F. and D. A. Quattrochi. 1997. Introduction: Scale, multiscaling, remote sensing and GIS. In Quattrochi, D. A. and M. F. Goodchild (Ed.), *Scale in Remote Sensing and GIS*. CRC Press, Lewis Publishers, Inc. pp.: 1-11.
- Goodchild, M. F. 2001. A geographer looks at spatial information theory. Montello, D. (Ed.). *Spatial Information Theory, Foundations of Geographic Information Science, International Conference COSIT 2001*. Morro Bay, California. Lecture Notes in Computer Science, Springer. pp.: 1-13.
- Goodchild, M. F. and T. Case. 2001. Introduction. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 3-13.
- Gosz, J. R. 1993. Ecotone hierarchies. *Ecological Applications*, 3 (3): 369-376.
- Gustafson, E. J. 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems*, 1: 143-156.
- Hangouët, J.-F. 2004. Geographical multi-representation: striving for the hyphenation. *International Journal Geographical Information Science*, 18 (4): 309-326.
- Harper, K. A. and S. E. Macdonald. 2002. Structure and composition of edges next to regenerating clear-cuts in mixed-wood boreal forest. *Journal of Vegetation Science*, 13 (4): 535-546.
- Heuvelink, G. 1993. Error propagation in quantitative spatial modelling. *Applications in Geographical Information Systems*. 151 pp.
- Heuvelink, G. 1996. Identification of field attribute error under different models of spatial variation. *International Journal of Geographical information Systems*, 10 (8): 921-935.
- Heuvelink, G. 1998. *Error Propagation in Environmental Modelling with GIS*. Taylor & Francis, London. 127 pp.
- Heuvelink, G. 1999. Propagation of error in spatial modelling with GIS. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 207-217.
- Hunter, A. 1996. *Uncertainty in Information Systems: An Introduction to Techniques and Applications*. McGraw-Hill, 152 pp.
- Hunter, G., B. Hock, M. Robey and M. F. Goodchild. 1996. Experimental development of a vector data uncertainty. *Second Symposium of Spatial Data Accuracy Assessment in Natural Resources and Environmental Sciences*. Fort Collins, Colorado. pp.: 217-224.
- Hunter, G. J., J. Qiu and M. F. Goodchild. 1998. Application of a new model of vector data uncertainty. Lowell, K. and A. Jaton (Ed.). *Third International Symposium on Spatial*

*accuracy assessment in natural resources and environmental sciences*. Québec. Ann Arbor Press. pp.: 203-208.

Insightful, C. 2001. *S-PLUS 6 for Windows Guide to Statistics*. Seattle. 712 pp.

Jacquez, G. and S. Maruca. 2001. *Boundary Seer. Software for geographic boundary analysis. User guide*. TerraSer Inc., 189 pp.

Jacquez, G. M., S. Maruca and M.-J. Fortin. 2000. From fields to objects: a review of geographic boundary analysis. *Journal of Geographical Systems*, 2 (3): 221-241.

Johnston, D. 1999. Étude des caractéristiques géostatistiques du volume forestier. Essai. Département de Sciences Géomatiques, Université Laval. Québec. pp.

Jordan, G. J. 2002. Space, time, and uncertainty : detecting and characterizing boundaries in forest fire disturbances. Ph. D. School of Resource and Environmental Management, Simon Fraser University. 246 pp.

Journel, A. 1996. Modelling uncertainty and spatial dependence : stochastic imaging. *International Journal of Geographical Information Science*, 10 (5): 517-522.

Jurdant, M. 1964. *Carte phytosociologique et forestière de la Forêt Expérimentale de Montmorency*. Direction des Recherches Forestières, Ministère des Forêts du Canada, pp.

Kaluzny, S. P., S. C. Vega, T. P. Cardoso and A. A. Shelly. 1998. *S+SpatialStats. User's Manual for Windows and UNIX*. Springer-Verlag, New York. 327 pp.

Kimmins, J. P. 1997. *Forest Ecology : a Foundation for Sustainable Management*. Prentice Hall, Upper Saddle River, N.J. 596 pp.

Kittler, J., M. Hatef, R. P. W. Duin and J. Matas. 1998. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 (3): 226-239.

Klecka, W. R. 1980. *Discriminant Analysis*. SAGE Publications Inc., Newbury Park. 70 pp.

Kyriakidis, P. C. 2001. Geostatistical models of uncertainty for spatial data. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 175-213.

Kyriakidis, P. C. and J. L. Dungan. 2001. A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data in ecological model predictions. *Environmental and Ecological Statistics*, 8: 311-330.

Leckie, D. G. and M. D. Gillis. 1995. Forest inventory in Canada with emphasis on map production. *Forestry Chronicle*, 71 (1): 74-88.

- Legendre, P. and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. *Vegetatio*, 80 (2): 107-138.
- Legendre, P. and L. Legendre. 1998. *Numerical Ecology*. Elsevier Science B.V., 853 pp.
- Leung, Y., M. F. Goodchild and C. C. Lin. 1992. Visualization of fuzzy scenes and probability fields. *Fifth International Symposium on Spatial Data Handling*. Charleston. pp.: 480-490.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology*, 73 (6): 1943-1967.
- Levine, N. and Associates. 2002. *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 2.0)*. National Institute of Justice, Washington, DC. pp.
- Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind. 1999a. Data quality: Introduction. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 175-176.
- Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind. 1999b. Spatial analysis: Introduction. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 235-237.
- Lowell, K., G. Edwards and G. L. Kucera. 1996. Modelling heterogeneity and change in natural forests. *Geomatica*, 50 (4): 425-440.
- Lowell, K. and G.-J. Liu. 1999. Modelling forest productivity as a function of digital ecophysiological data and distance from a mapped forest boundary. *Geographical and Environmental Modelling*, 3 (2): 163-177.
- Lowell, K., T. De Groeve and B. Thierry. 2000. Combining map geometry and polygon attributes to estimate spatial uncertainty from multiple realisations. (Ed.). *Spatial Data Handling*. pp.:
- Lowell, K. E. 1994. A fuzzy surface cartographic representation for forestry based on Voronoi diagram area stealing. *Canadian Journal of Forest Research*, 24 (9): 1970-1980.
- Lowell, K. E. 1999. Effects of adjacent stand characteristics and boundary distance on density and volume of mapped land units in the boreal forest. *Plant Ecology*, 143 (1): 99-106.
- Mark, D. and F. Csillag. 1989. The nature of boundaries on "area-class" maps. *Cartographica*, 26: 65-78.
- McGwire, K. and P. Fisher. 2001. Spatially variable thematic accuracy: Beyond the confusion matrix. In Hunsaker, C., M. Goodchild, M. Friedl and T. Case (Ed.), *Spatial*

- Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 308-329.
- Merriam, G. 1987. Landscape ecology: the ecology of heterogeneous systems. Moss, M. (Ed.). *Landscape Ecology and Management. Proceedings of the First Symposium of the Canadian Society for Landscape Ecology and Management*. University of Guelph. Polyscience Publications Inc. pp.: 43-50.
- Methven, I. and U. Feunekes. 1987. Fire games for park managers: exploring the effect of fire on landscape vegetation patterns. Moss, M. (Ed.). *Landscape Ecology and Management. Proceedings of the First Symposium of the Canadian Society for Landscape Ecology and Management*. University of Guelph. Polyscience Publications Inc. pp.: 101-110.
- Milne, B. T. 1991. Heterogeneity as a multiscale characteristic of landscapes. In Kolasa, J. and S. T. A. Pickett (Ed.), *Ecological Heterogeneity*. Springer-Verlag. pp.: 69-84.
- Milne, B. T. 1998. Motivation and benefits of complex systems approaches in ecology. *Ecosystems*, 1: 449-456.
- Milne, B. T. and W. B. Cohen. 1999. Multiscale assessment of binary and continuous landcover variables for MODIS validation, mapping, and modeling applications. *Remote Sensing of Environment*, 70 (1): 82-98.
- Milne, B. T., V. K. Gupta and C. Restrepo. 2002. A scale invariant coupling of plants, water, energy, and terrain. *Ecoscience*, 9 (2): 191-199.
- Ministère des Ressources naturelles du Québec. 1994. *Le Point d'Observation Écologique*. Gouvernement du Québec, Québec. 116 pp.
- Ministère des Ressources naturelles du Québec. 1995. *Norme de Stratification Écoforestière Troisième Programme de Connaissance de la Ressource Forestière*. Ministère de Ressources naturelles du Québec, Gouvernement du Québec, Québec. 79 pp.
- Moeur, M. and R. R. Hershey. 1998. Preserving spatial and attribute correlation in the interpolation of forest inventory data. Lowell, K. and A. Jatton (Ed.). *Third International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Québec. Ann Arbor Press. pp.: 419-430.
- Moghaddam, M., J. L. Dungan and S. Acker. 2002. Forest variable estimation from fusion of SAR and multispectral optical data. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (10): 2176-2187.
- Mostafavi, M. A., G. Edwards and R. Jeansoulin. 2004. Ontology-based method for quality assessment of spatial data bases. *Proceedings of ISSDQ'04*. Bruck/Leitha, Austria. GeoInfo Series, pp.: 49-66.

- Mowrer, H. 1998. Accuracy (Re)assurance: selling uncertainty assessment to the uncertain. Lowell, K. and A. Jaton (Ed.). *Third International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Québec. Ann Arbor Press. pp.: 3-10.
- Nigh, G. D. and A. B. Love. 2004. Predicting crown class in three western conifer species. *Canadian Journal of Forest Research*, 34 (3): 592-599.
- Oden, N. L., R. R. Sokal, M.-J. Fortin and H. Goebel. 1993. Categorical wombling: determining regions of abrupt change in categorical variables. *Geographical Analysis*, 25: 315-336.
- Oliver, C. D. and B. C. Larson. 1996. *Forest Stand Dynamics*. John Wiley & Sons, Inc., 520 pp.
- Ord, J. K. and A. Getis. 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27 (4): 286-306.
- Peddle, D., G. Foody, A. Zhang, S. Franklin and E. Ledrew. 1994. Multi-source image classification II: An empirical comparison of evidential reasoning and neural network approaches. *Canadian Journal of Remote Sensing*, 20 (4): 396-407.
- Perera, A. H. and D. L. Euler. 2000. Landscape ecology in forest management: An introduction. In Perera, A. H., D. L. Euler and I. D. Thompson (Ed.), *Ecology of a Managed Terrestrial Landscape: Patterns and Processes of Forest Landscapes in Ontario*. UBC Press. Vancouver. pp.: 3-11.
- Perera, A. H., D. L. Euler and I. D. Thompson (Ed.). 2000. *Ecology of a Managed Terrestrial Landscape: Patterns and Processes of Forest Landscapes in Ontario*. UBC Press, Vancouver. 336 pp.
- Pohl, C. and J. L. V. Genderen. 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19 (5): 823-854.
- Rejwan, C., N. C. Collins, L. J. Brunner, B. J. Shuter and M. S. Ridway. 1999. Tree regression analysis on the nesting habitat for smallmouth bass. *Ecology*, 80: 341-348.
- Risser, P. G. 1995. The status of the science examining ecotones. *Bioscience*, 45 (5): 318-325.
- Rolling, M. G., R. E. Keane and R. A. Parsons. 2004. Mapping fuels and fire regimes using remote sensing, ecosystem simulation and gradient modelling. *Ecological Applications*, 14 (1): 74-95.
- Rowe, J. S. 1987. Landscape ecology: the study of terrain ecosystems. Moss, M. (Ed.). *Landscape Ecology and Management. Proceedings of the First Symposium of the Canadian*

- Society for Landscape Ecology and Management*. University of Guelph. Polyscience Publications Inc. pp.: 35-42.
- Segal, M. R. 1988. Regression trees for censored data. *Biometrics*, 44: 35-47.
- Shortridge, A. 2001. Characterizing uncertainty in digital elevation models. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 238-257.
- Shugart, H. H. 1984. *A Theory of Forest Dynamics. The Ecological Implications of Forest Succession Models*. Springer-Verlag, New York. 278 pp.
- Sklar, F. and C. Hunsaker. 2001. The use and uncertainties of spatial data for landscape models: an overview with examples from the Florida Everglades. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 15-46.
- Smith, B. 1995. On drawing lines on a map. Frank, A. U. and W. Kuhn (Ed.). *Spatial Information Theory: A Theoretical Basis for GIS. International Conference COSIT '95*. Austria. Lecture notes in computer science, Springer. pp.: 475-484.
- Smith, B. and D. Mark. 1998. Ontology and geographic kinds. *Proceedings International Symposium on Spatial Data Handling*. Vancouver. pp.: 308-320.
- Smith, B. 2001. Peter A. Burrough and Andrew U. Frank (eds.), Geographic objects with indeterminate boundaries (London and Bristol, PA. 1996). *Dialectica*, 55 (1): 72-74.
- Smith, B. and D. Mark. 2001. Geographical categories: an ontological investigation. *International Journal Geographical Information Science*, 15: 591-612.
- Stehman, S. V. 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62: 77-89.
- Stehman, S. V. 1998. Alternative measures for comparing thematic map accuracy. Lowell, K. and A. Jaton (Ed.). *Third International Symposium on Spatial accuracy assessment in natural resources and environmental sciences*. Québec. Ann Arbor Press. pp.: 45-51.
- Stine, P. and C. Hunsaker. 2001. An introduction to uncertainty issues for spatial data used in ecological applications. In Hunsaker, C., M. F. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 91-107.
- Stock, K. and D. Pullar. 1999. Identifying semantically similar elements in heterogeneous spatial databases using predicate logic expressions. In Vikouski, A., K. Brassel and H. Schek (Ed.). *Interoperating Geographic Information Systems. Second International Conference INTEROP'99*. Zurich. Lecture Notes in Computer Science, Springer. pp.: 231-252.

- Strayer, D. L., M. E. Power, W. F. Fagan, S. T. A. Pickett and J. Belnap. 2003. A classification of ecological boundaries. *Bioscience*, 53 (8): 723-729.
- Tang, S. M., J. F. Franklin and D. R. Montgomery. 1997. Forest harvest patterns and landscape disturbance processes. *Landscape Ecology*, 12: 349-363.
- Thierry, B. and K. E. Lowell. 2001. An uncertainty-based method of photointerpretation. *Photogrammetric Engineering and Remote Sensing*, 67 (1): 65-72.
- Turner, M. G., R. H. Gardner and R. V. O'Neill. 2001. *Landscape Ecology in Theory and Practice. Pattern and Process*. Springer-Verlag, New York. 401 pp.
- Vauglin, F. 1998. Statistical representation of relative positional uncertainty for geographical linear features. In Goodchild, M. F. and R. Jeansoulin (Ed.), *Data Quality in Geographic Information, From Error to Uncertainty*. Ed. Hermes. pp.: 87-96.
- Venables, W. N. and B. D. Ripley. 1999. *Modern Applied Statistics with S-Plus*. Springer Verlag, New York. pp.
- Veregin, H. 1999. Data quality parameters. In Longley, P. A., M. F. Goodchild, D. J. Maguire and D. W. Rhind (Ed.), *Geographical Information Systems*. John Wiley & Sons, Inc. pp.: 177-189.
- Vézina, P.-E. et G. Paillé. 1969. *Aménagement Intensif des Sapinières Mûres et Surannées à la Forêt Montmorency*. Fonds de Recherches forestières de l'Université Laval, Québec.
- Wagner, H. et M.-J. Fortin. 2005. Spatial analysis of landscapes : concepts and statistics. *Ecology*, 86(8): 1975-1987.
- Wald, L. 1999. Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (3): 1190-1193.
- Wang, F. and G. B. Hall. 1996. Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical information Systems*, 10 (5): 573-590.
- Wang, M. and P. J. Howarth. 1994. Multi-source spatial data integration: problems and some solutions. *Canadian Journal of Remote Sensing*, 20 (4): 360-367.
- Wang, Y. and D. Civco. 1994. Evidential reasoning-based classification of multi-source data for improved land cover mapping. *Canadian Journal of Remote Sensing*, 20 (4): 381-395.
- White, P. S. and S. T. A. Pickett. 1985. Natural disturbance and patch dynamics: an introduction. In White, P. S. and S. T. A. Pickett (Ed.), *The Ecology of Natural Disturbance and Patch Dynamics*. Academic Press. San Diego. pp.: 3-13.

Woodcock, C., J. Collins, S. Gopal, V. Jakabhazy, X. Li, S. Macomber, S. Ryherd, Y. Wu, V. Harward, J. Levitan and R. Warbington. 1994. Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. *Remote Sensing of Environment*, 50 (3): 240-254.

Zhang, J. and M. Goodchild. 2002. *Uncertainty in Geographical Information*. Taylor & Francis, London. 266 pp.

Zhu, A. 2001. Modeling spatial variation of classification accuracy under fuzzy logic. In Hunsaker, C., M. Goodchild, M. Friedl and T. Case (Ed.), *Spatial Uncertainty in Ecology, Implications for Remote Sensing and GIS Applications*. Springer. pp.: 330-350.

Zhu, A.-X. and L. E. Band. 1994. A knowledge-based approach to data integration for soil mapping. *Canadian Journal of Remote Sensing*, 20 (4): 408-418.

## Annexe I - Variables cartographiques et de terrain

### Variables cartographiques

Tableau 45. Aire minimale de photo-interprétation (Ministère de Ressources naturelles du Québec 1995).

Aire minimale	Situation
8 ha	-Terrain forestier productif -Lorsqu'on tient compte des limites de dépôts de surface -Lorsqu'on tient compte de la classe de pente (sauf pour F et S) -Lorsqu'on tient compte de l'historique du peuplement
4 a	-Îles -Peuplement qui se distingue nettement de son entourage : a) par le type de couvert b) par un écart d'au moins deux classes pour la hauteur, densité et âge
2 ha	-Terrain non forestier -Terrain forestier improductif -Îlot boisé de hauteur moyenne de 7 m et plus situé sur a) une aire de coupe, b) une zone ravagée par le feu, c) par le vent ou d) une épidémie grave -Îlot déboisé par une coupe totale, un feu, un chablis total ou une épidémie grave et entouré de peuplements de hauteur de 7 m ou plus Superficiés avec une pente de classe F ou S sur les photographies.

Tableau 46. Catégories de *Type de couvert* (Ministère de Ressources naturelles du Québec 1995)

Type de couvert	Définition
Résineux	Les essences résineuses occupent 75% ou plus de la surface terrière du peuplement
Mélangé	Les essences résineuses couvrent entre 26 et 74% de la surface terrière du peuplement. À prédominance résineuse : les résineuses couvrent entre 50 et 74% de la surface terrière. À prédominance feuillue : les feuillues couvrent entre 50 et 74% de la surface terrière.
Feuillu	Les essences résineuses couvrent 25% ou moins de la surface terrière du peuplement

Tableau 47. Catégories de *Classe de pente* (Ministère de Ressources naturelles du Québec 1995).

Classe	Pente (%)	Aire minimale à cartographier (ha)
A	0 à 3 – nulle	20
B	4 à 8 – faible	20
C	9 à 15 – douce	8
D	16 à 30 – modérée	8
E	31 à 40 – forte	8
F	41 et plus – abrupte	4
S	superficiés entourées de pentes de 41% et plus	4

Tableau 48. Catégories de *Stade de développement* (Conseil canadien des ministres des forêts 2004).

Catégorie	Définition
Semis	Peuplement composé de jeunes plantes provenant de la germination d'une graine, dont le diamètre à la hauteur de la poitrine (dhp) est égal ou inférieur à 1 cm ou la hauteur est égale ou inférieure à 1,5 m
Fourré	Peuplement touffu de jeunes brins dont les branches s'entrecroisent, provenant le plus souvent d'une régénération naturelle
Gaulis	Peuplement composé de jeunes arbres de 1 à 3 m de hauteur et entre 1 et 10 cm de dhp
Perchis	Stade qui suit au gaulis, dans lequel les arbres ont de 12 à 20 cm de dhp
Futaie	Peuplement normalement constitué d'arbres issus de semences qu'on laisse se développer jusqu'à la fermeture du couvert.

## Variables de terrain

Tableau 49. Catégories de *Versant* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e
Haut versant	H
Moyen versant	M
Bas versant	B

Tableau 50. Catégories de *Situation sur la pente* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e
Terrain plat (0 à 3% de pente)	0
Escarpement	2
Sommet arrondi	3
Haut de pente	4
Mi-pente	5
Replat	6
Bas de pente	7
Dépression ouverte	8
Dépression fermée	9

Tableau 51. Catégories de *Forme de la pente* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e
Concave	C
Convexe	X
Régulière	R
Pente irrégulière	I

Tableau 52. Catégories de *Texture de l'horizon B* et de *Texture de l'horizon C* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e	Catégorie	Code dans la p-e
Sable très grossier	STG	Loam sableux moyen	LSM
Sable grossier	SG	Loam sableux fin	LSF
Sable moyen	SM	Loam sableux très fin	LSTF
Sable fin	SF	Loam sablo argileux	LSA
Sable très fin	STF	Loam	L
Sable très grossier loameux	STGL	Loam limoneux	LLI
Sable grossier loameux	SGL	Limon	LI
Sable moyen loameux	SML	Loam argileux	LA
Sable fin loameux	SFL	Loam limono argileux	LLIA
Sable très fin loameux	STFL	Argile sableuse	AS
Loam sableux très grossier	LSTG	Argile limoneuse	ALI
Loam sableux grossier	LSG	Argile	A

Tableau 53. Catégories de *Type de humus* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e
Mull	MU
Moder	MD
Mor	MR
Tourbe	TO
Anmoor	AN
Sol organique	SO
Ne s'applique pas	NA

Tableau 54. Catégories d'épaisseur du *Dépôt de surface* (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e
Dépôt épais (épaisseurs modales supérieures à 1 m. Affleurement rocheux absentes ou très rares)	1A
Dépôt d'épaisseur moyenne (épaisseur modale entre 50 cm et 1 m. Affleurements rocheux rares ou très rares)	1AY
Dépôt mince (épaisseur modale entre 25 cm et 50 cm. Affleurements rocheux rares ou peu fréquentes)	1AM
Dépôt très mince (épaisseur modale inférieure à 25 cm. Affleurements rocheux peu fréquentes)	M1A
Dépôt allant de mince à très mince (épaisseur modale inférieure à 50 cm. Affleurements rocheux fréquentes)	R1A
Dépôt très mince ou absent (affleurements rocheux très fréquents, représentent plus de 50 % de la surface)	R

Tableau 55. Catégories de *Classe de drainage* avec ses modificateurs (Ministère des Ressources naturelles du Québec 1994).

Catégorie	Code dans la p-e	Modificateurs du drainage	Code dans la p-e
Excessif	0	Aucun modificateur	0
Rapide	1	Drainage latéral	1
Bon	2	Horizon gelé	2
Modéré	3	Amélioration d'origine	
Imparfait	4	anthropique	3
Mauvais	5	Ralentissement d'origine	
Très mauvais	6	anthropique	4

## Annexe II Les frontières cartographiques

### Calcul de la similarité partielle $s_{12j}$

La Figure 51 présente une comparaison des intervalles de couvert des différentes essences dans les catégories de *Groupement d'essences*. Cette figure indique les valeurs minimales, moyennes et maximales de chacun des intervalles de couvert. Elle permet aussi d'apprécier l'important chevauchement des intervalles de couvert d'une essence dans les catégories. Afin de bien décrire la similarité des intervalles de couvert des catégories, la similarité partielle est donc calculée en prenant compte de la distance entre les moyennes de chacun des intervalles ainsi que de leur degré de chevauchement.

Le principe qui guide ce calcul est qu'à mesure que la distance entre les valeurs moyennes des deux intervalles de couvert augmente, leur similarité s'affaiblit tandis qu'à mesure que le chevauchement entre les deux intervalles augmente, leur similarité augmente aussi. Ainsi la similarité partielle entre les catégories 1 et 2 de *Groupement d'essences* pour l'essence ou descripteur  $j$  se calcul comme :

$$s_{12j} = \left( -\frac{|Moy_1 - Moy_2|}{200} + 0,5 \right) + \left[ \left( \frac{Max_1 - Min_2}{Max_2 - Min_1} \right) * 0,5 \right], \quad \text{AI.VII.1}$$

où  $Moy_1$  et  $Moy_2$  sont les valeurs moyennes,  $Max_1$  et  $Max_2$  sont les valeurs maximales et  $Min_1$  et  $Min_2$  sont les valeurs minimales des intervalles de couvert des catégories 1 et 2 respectivement. Dans le premier terme de l'équation AI.1 la distance entre les moyennes qui peut prendre des valeurs entre 0 et 100 correspond à des valeurs de similarité partielle entre 0,5 et 0 respectivement. Dans le deuxième terme de l'équation AI.1, le chevauchement entre les intervalles qui peut prendre des valeurs entre 0 et 1 correspond à de valeurs de similarité partielle entre 0 et 0,5 respectivement. Ainsi, les deux termes de l'équation contribuent chacun en 50 % à la valeur de similarité partielle.

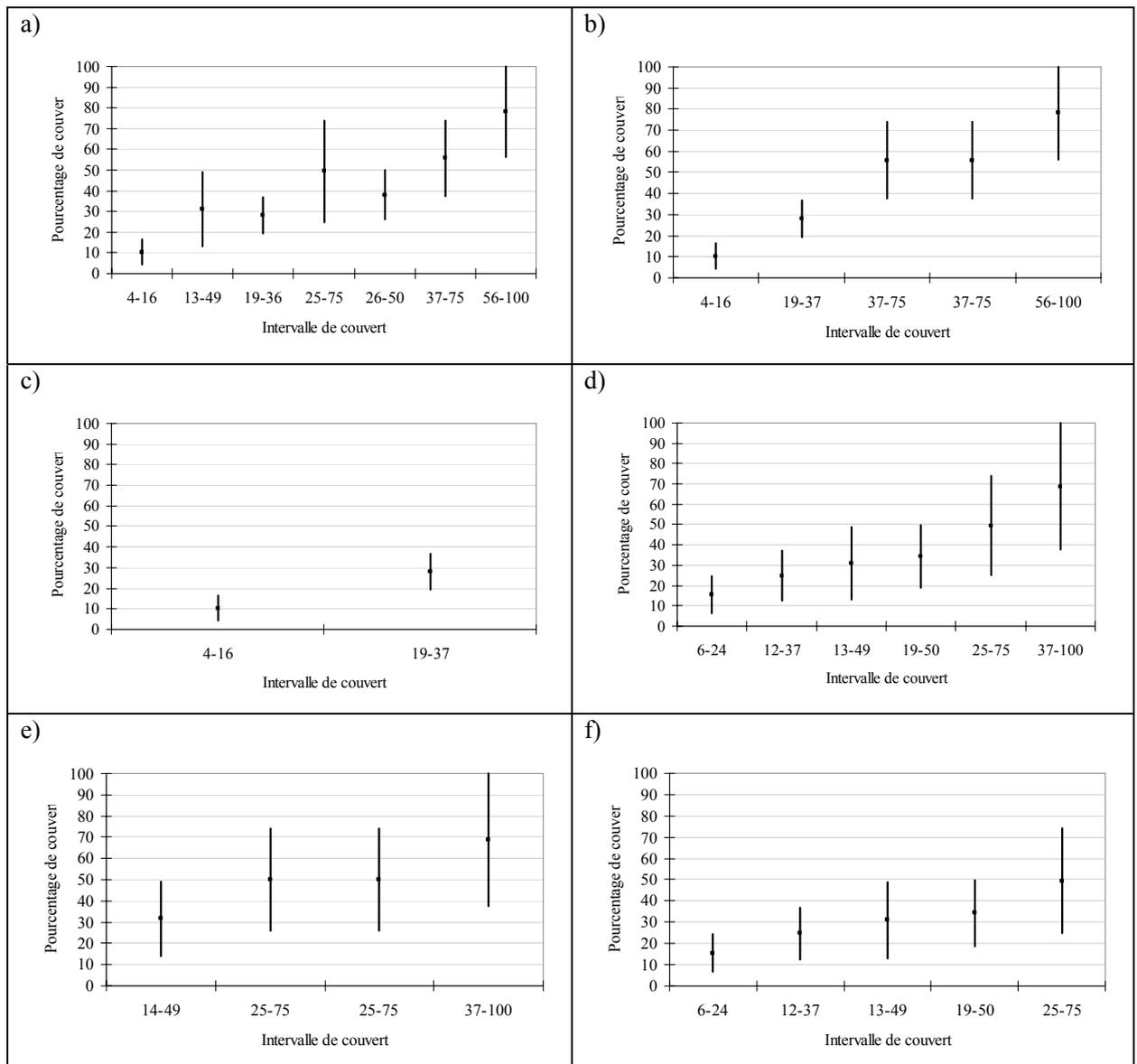


Figure 51. Comparaison des intervalles de couvert des essences dans les différentes catégories de *Groupe ment d'essences* : a) Sapin baumier (S); b) Épinette (E); c) Pin blanc (Pb); d) Bouleau à papier (Bb); e) Bouleau jaune (Bj) et f) Peuplier (Pe).

### Exemple du calcul de la force d'une frontière

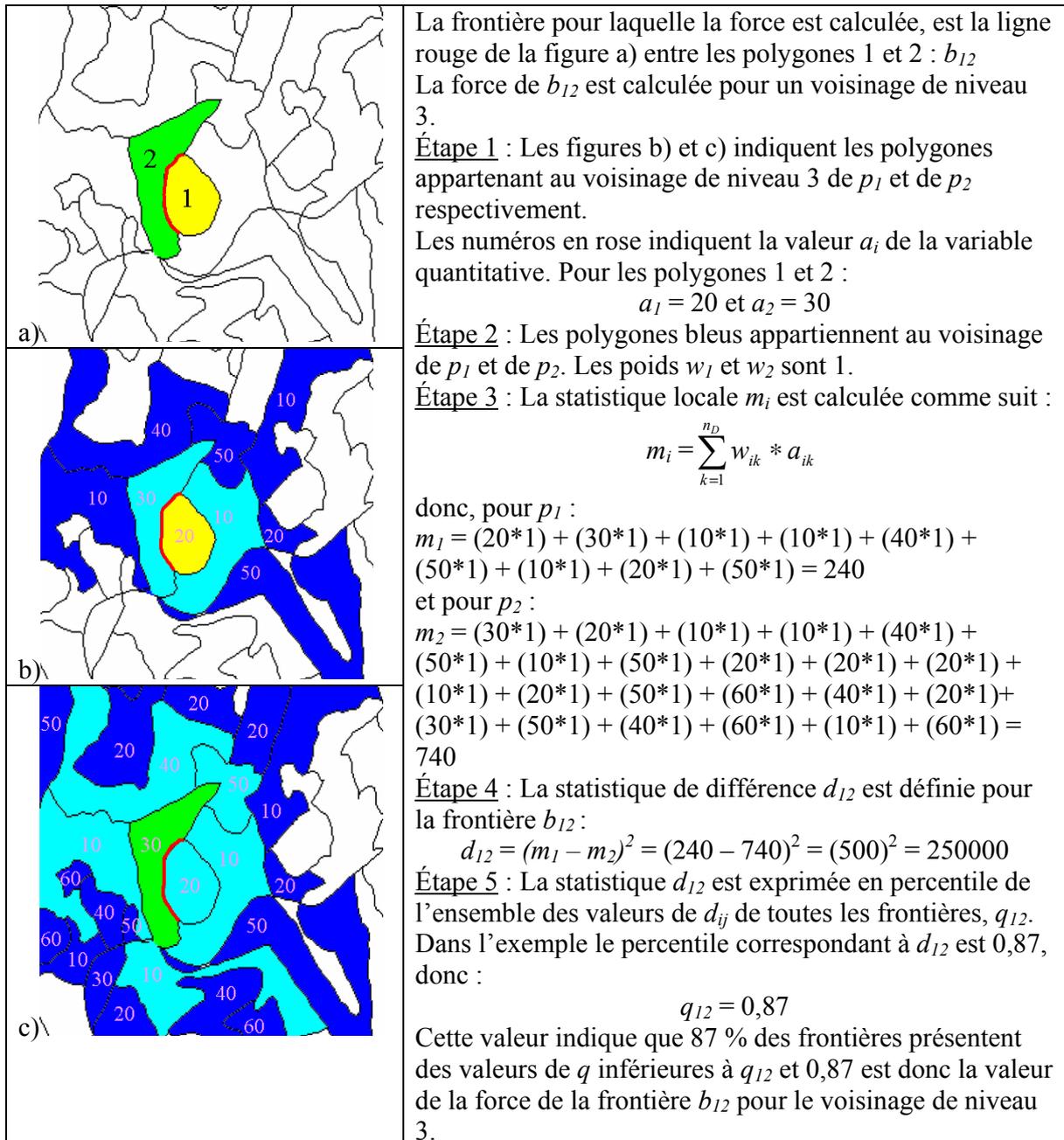


Figure 52. Exemple du calcul de la force d'une frontière (Csillag *et al.* 2001).

Tableau 56. Valeurs de l'indice de similitude écologique  $S_{GE}$ . ( $n = 8018$  frontières entre 3729 peuplements). Les valeurs en gras sont mentionnées dans le texte.

	SS	EE	SE	ES	EPb	SBb	SBj	SFi	SPe	BbS	BjS	FiS	PeS	BjR	BbR	FiR	FNC	R	M	F
SS	1	0,33	0,39	0,34	0,25	0,23	0,23	0,17	0,23	0,10	0,10	0,08	0,10	0,06	0,04	0,03	0,04	0,36	0,14	0,03
EE		1	0,34	0,39	0,39	0,14	0,14	0,11	0,14	0,07	0,07	0,06	0,07	0,06	0,04	0,03	0,04	0,36	0,14	0,03
SE			1	0,36	0,26	0,27	0,27	0,22	0,27	0,11	0,11	0,08	0,11	0,07	0,03	<b>0,02</b>	0,03	0,36	0,14	0,03
ES				1	0,50	0,18	0,18	0,14	0,18	0,22	0,22	0,17	0,22	0,07	0,03	<b>0,02</b>	0,03	0,36	0,14	0,03
EPb					1	0,11	0,11	0,09	0,11	0,06	0,06	0,05	0,06	0,07	0,03	<b>0,02</b>	0,03	0,36	0,14	0,03
SBb						1	0,50	<b>0,55</b>	0,50	0,35	0,22	0,32	0,20	0,10	0,13	0,07	0,25	0,07	0,17	0,30
SBj							1	0,40	0,50	0,20	0,37	0,16	0,20	0,16	0,07	0,13	0,06	0,07	0,17	0,30
SFi								1	0,55	0,21	0,18	0,32	0,21	0,08	0,08	0,06	0,13	0,06	0,17	0,30
SPe									1	0,20	0,22	0,32	0,35	0,10	0,07	0,07	0,25	0,07	0,17	0,30
BbS										1	0,36	0,54	0,50	0,18	0,22	0,14	0,21	0,14	0,10	0,30
BjS											1	0,29	0,36	0,41	0,14	0,28	0,11	0,14	0,10	0,30
FiS												1	0,54	0,16	0,21	0,11	0,31	0,11	0,10	0,30
PeS													1	0,18	0,14	0,14	0,21	0,14	0,10	0,30
BjR														1	0,09	0,17	0,08	0,09	0,10	0,30
Bb																0,33	0,37	0,33	0,03	0,13
Bj																	0,25	0,33	0,03	0,13
Fi																		0,25	0,03	0,13
FNC																			0,03	0,13
R																			0,03	0,13
M																			0,14	0,03
F																			1	0,13

Tableau 57. Valeur de l'indice de similitude écologique  $S_{CD}$  ( $n = 8018$  frontières entre 3729 peuplements).

Catégorie	A	B	C	D
A	1	0,52	0,24	0,16
B		1	0,52	0,40
C			1	0,79
D				1

Tableau 58. Valeurs de l'indice de similitude écologique  $S_{CH}$  ( $n = 8018$  frontières entre 3729 peuplements).

Catégorie	1	2	3	4	5	6
1	1	0,57	0,32	0,17	0,16	0,08
2		1	0,57	0,32	0,31	0,20
3			1	0,57	0,54	0,36
4				1	0,96	0,63
5					1	0,66
6						1

Tableau 59. Valeurs de l'indice de similitude écologique  $S_{CA}$  ( $n = 8018$  frontières entre 3729 peuplements).

Catégorie	0	15	30	45	60	75	90
0	1	0,60	0,38	0,25	0,16	0,09	0,04
15		1	0,60	0,38	0,25	0,16	0,09
30			1	0,60	0,38	0,25	0,16
45				1	0,60	0,38	0,25
60					1	0,60	0,38
75						1	0,60
90							1

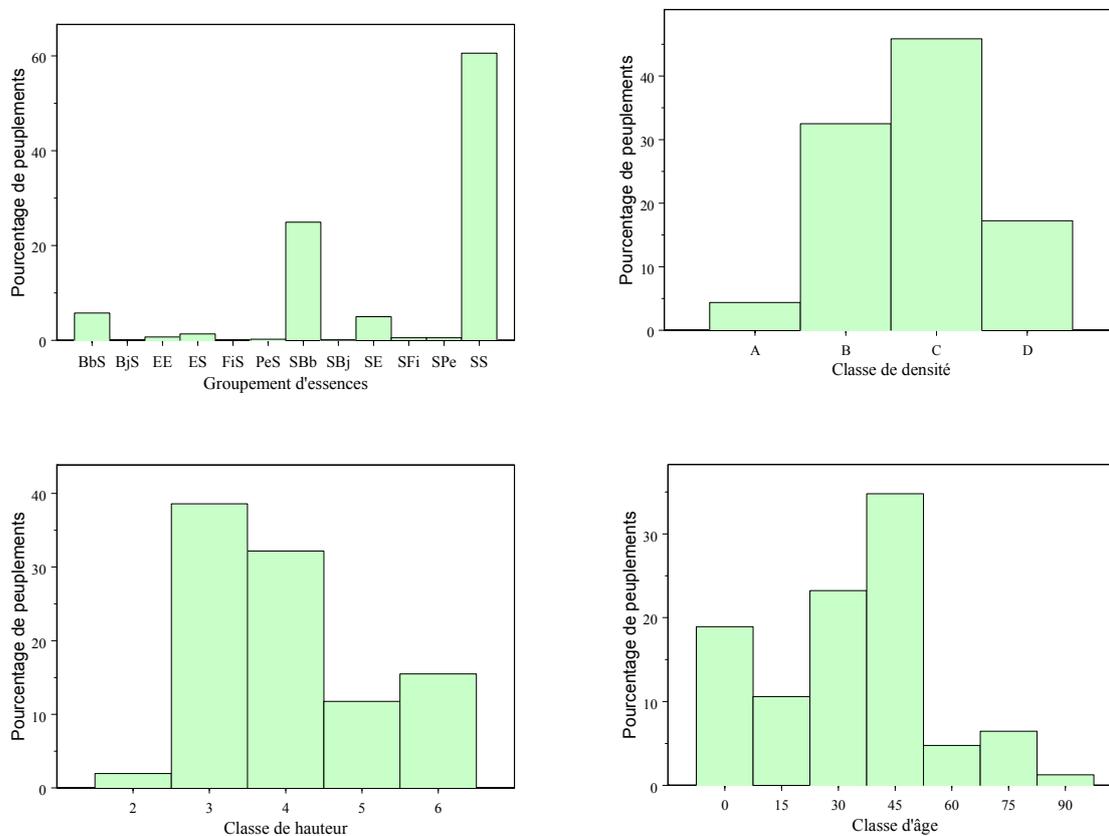


Figure 53. Pourcentage de chacune des catégories des variables *Groupement d'essences*, *Classe de densité*, *Classe de hauteur* et *Classe d'âge* des peuplements de la carte forestière de 1992 de la Forêt Montmorency.

## Annexe III - Modèles de régression arborescente

### Modèle Cartographique général - CartoG

$Volume\ total \sim AC + CA + CD + GE + CH + Orig + CP + Pert + SD + X + Y$

Variabiles utilisées pour la construction de l'arbre :

*AC GE CD Y X CP CA CH Pert*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) root 338 1.472e+006 114.3000
  2) AC:1966,1967,1968,1969,1970,1972,1973,1974,1975,1976,1977,1978,
1979,1987,1989,CHT,ES 60 1.303e+005 39.4600
  4) AC:1966,1970,1972,1973,1975,1976,1978,1987,1989 24 3.137e+003
11.3000
    8) GE:BbS,CT,SD 14 7.629e+002 5.9570
    16) CD:D 6 1.568e+000 0.6167 *
    17) CD:B,C,sd 8 4.619e+002 9.9630 *
    9) GE:SBb,SS 10 1.415e+003 18.7800
    18) Y<37051.5 5 2.454e+002 13.4000 *
    19) Y>37051.5 5 8.801e+002 24.1600 *
  5) AC:1967,1968,1969,1974,1977,1979,CHT,ES 36 9.543e+004 58.2300
  10) GE:BbS,SBb 15 3.274e+004 39.8100
    20) AC:1967,1969,1979,CHT,ES 9 2.478e+003 23.0900 *
    21) AC:1968,1974 6 2.397e+004 64.9000 *
  11) GE:SD,SS 21 5.397e+004 71.3900
    22) X<56997.9 8 2.625e+004 97.0100 *
    23) X>56997.9 13 1.923e+004 55.6200
    46) AC:1967,CHT 6 2.055e+003 31.4800 *
    47) AC:1977,1979,ES 7 1.068e+004 76.3000 *
  3) AC:1981,1983,1992,1993,1994,SP 278 9.332e+005 130.4000
  6) CP:A,E 106 2.427e+005 104.5000
    12) GE:BjS,EE,ES,SBb,SBj,SE 32 2.761e+004 78.1100
    24) X<55602.6 13 9.643e+003 63.7300
    48) GE:BjS,EE,SE 8 2.629e+003 47.8600 *
    49) GE:ES,SBb,SBj 5 1.776e+003 89.1200 *
    25) X>55602.6 19 1.343e+004 87.9500
    50) CA:30,60,90 6 3.202e+003 73.4000 *
    51) CA:45,75 13 8.375e+003 94.6700
    102) CH:3 6 9.372e+002 80.3200 *
    103) CH:2,4 7 5.143e+003 107.0000 *
  13) GE:BbS,SD,SFi,SPe,SS 74 1.831e+005 116.0000
    26) Pert:el,em 32 6.853e+004 95.8100

```

52) CA:30,45,90 24 4.434e+004 84.5500  
 104) Pert:em 5 5.847e+003 45.1600 \*  
 105) Pert:el 19 2.869e+004 94.9200  
 210) Y<37284.3 13 1.129e+004 85.5000  
 420) Y<35596.3 6 4.440e+003 91.0000 \*  
 421) Y>35596.3 7 6.515e+003 80.7900 \*  
 211) Y>37284.3 6 1.375e+004 115.3000 \*  
 53) CA:60,75 8 1.202e+004 129.6000 \*  
 27) Pert:SD,chp,cp 42 9.162e+004 131.3000  
 54) CH:4,6 7 6.274e+003 95.5300 \*  
 55) CH:3,5 35 7.458e+004 138.5000  
 110) X<57978.3 25 4.244e+004 127.8000  
 220) X<54617.3 5 1.534e+004 164.3000 \*  
 221) X>54617.3 20 1.878e+004 118.7000  
 442) Y<38817.8 11 1.005e+004 111.3000  
 884) X<55690.3 6 2.163e+003 103.8000 \*  
 885) X>55690.3 5 7.147e+003 120.3000 \*  
 443) Y>38817.8 9 7.402e+003 127.7000 \*  
 111) X>57978.3 10 2.212e+004 165.2000  
 222) X<58937.3 5 7.021e+003 201.7000 \*  
 223) X>58937.3 5 1.819e+003 128.8000 \*  
 7) CP:B,C,D,SD 172 5.756e+005 146.4000  
 14) CH:2,4 44 1.177e+005 123.6000  
 28) Y<39976.8 22 5.472e+004 107.2000  
 56) X<57759.8 13 2.614e+004 127.3000  
 112) X<56422.2 6 3.162e+003 105.7000 \*  
 113) X>56422.2 7 1.777e+004 145.8000 \*  
 57) X>57759.8 9 1.576e+004 78.1400 \*  
 29) Y>39976.8 22 5.112e+004 140.0000  
 58) Y<43243.5 16 1.828e+004 153.9000  
 116) X<56695.2 8 7.989e+003 130.6000 \*  
 117) X>56695.2 8 1.620e+003 177.1000 \*  
 59) Y>43243.5 6 2.151e+004 102.9000 \*  
 15) CH:3,6,sd 128 4.270e+005 154.3000  
 30) Pert:chp,em 14 3.957e+004 115.2000  
 60) CP:B,D 8 7.950e+003 90.5000 \*  
 61) CP:C 6 2.027e+004 148.0000 \*  
 31) Pert:SD,el 114 3.634e+005 159.1000  
 62) X<54164.3 8 6.972e+003 120.2000 \*  
 63) X>54164.3 106 3.434e+005 162.0000  
 126) CA:0,45,90 92 2.667e+005 157.7000  
 252) X<54788.3 14 2.334e+004 181.0000  
 504) Y<39083.5 8 6.751e+003 197.5000 \*  
 505) Y>39083.5 6 1.155e+004 159.1000 \*  
 253) X>54788.3 78 2.343e+005 153.5000  
 506) Y<39447.3 31 8.681e+004 136.8000  
 1012) Y<38479.5 22 5.871e+004 152.2000  
 2024) Y<37985.5 16 4.195e+004 139.1000  
 4048) X<56030.2 5 1.077e+004 111.0000 \*  
 4049) X>56030.2 11 2.545e+004 151.8000  
 8098) CD:B,D 6 2.053e+004 137.5000 \*  
 8099) CD:C 5 2.207e+003 169.0000 \*  
 2025) Y>37985.5 6 6.602e+003 187.3000 \*  
 1013) Y>38479.5 9 1.011e+004 99.1600 \*

```

507) Y>39447.3 47 1.332e+005 164.5000
1014) X<59644.9 42 1.096e+005 170.8000
2028) Y<42440 33 6.902e+004 181.1000
4056) GE:BbS,CT,SD,SFi 12 1.395e+004 160.2000
8112) X<59088.1 7 9.772e+003 145.8000 *
8113) X>59088.1 5 7.187e+002 180.3000 *
4057) GE:SBb,SS 21 4.683e+004 193.0000
8114) X<56241.3 5 1.601e+004 168.9000 *
8115) X>56241.3 16 2.699e+004 200.6000
16230) X<57307.3 5 8.045e+003 226.7000 *
16231) X>57307.3 11 1.397e+004 188.7000
32462) Y<41211.5 5 7.057e+003 174.5000 *
32463) Y>41211.5 6 5.062e+003 200.5000 *
2029) Y>42440 9 2.444e+004 133.3000 *
1015) X>59644.9 5 7.531e+003 110.8000 *
127) CA:60,75 14 6.373e+004 190.4000
254) Y<42999.3 8 3.435e+004 217.8000 *
255) Y>42999.3 6 1.541e+004 154.0000 *

```

## Modèle Cartographique taillé à 45 feuilles - Carto45

$Volume\ total \sim AC + CA + CD + GE + CH + Orig + CP + Pert + SD + X + Y$

Variables utilisées pour la construction de l'arbre :

*AC GE X CP CA CH Pert Y CD*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) root 338 1472000.0 114.30
2) AC:1966,1967,1968,1969,1970,1972,1973,1974,1975,1976,1977,1978,
1979,1987,1989,CHT,ES 60 130300.0 39.46
4) AC:1966,1970,1972,1973,1975,1976,1978,1987,1989 24 3137.0
11.30 *
5) AC:1967,1968,1969,1974,1977,1979,CHT,ES 36 95430.0 58.23
10) GE:BbS,SBb 15 32740.0 39.81
20) AC:1967,1969,1979,CHT,ES 9 2478.0 23.09 *
21) AC:1968,1974 6 23970.0 64.90 *
11) GE:SD,SS 21 53970.0 71.39
22) XT<56997.9 8 26250.0 97.01 *
23) XT>56997.9 13 19230.0 55.62
46) AC:1967,CHT 6 2055.0 31.48 *
47) AC:1977,1979,ES 7 10680.0 76.30 *
3) AC:1981,1983,1992,1993,1994,SP 278 933200.0 130.40
6) CP:A,E 106 242700.0 104.50

```

12) GE:BjS,EE,ES,SBb,SBj,SE 32 27610.0 78.11  
 24) XT<55602.6 13 9643.0 63.73  
 48) GE:BjS,EE,SE 8 2629.0 47.86 \*  
 49) GE:ES,SBb,SBj 5 1776.0 89.12 \*  
 25) XT>55602.6 19 13430.0 87.95  
 50) CA:30,60,90 6 3202.0 73.40 \*  
 51) CA:45,75 13 8375.0 94.67  
 102) CH:3 6 937.2 80.32 \*  
 103) CH:2,4 7 5143.0 107.00 \*  
 13) GE:BbS,SD,SFi,SPe,SS 74 183100.0 116.00  
 26) Pert:el,em 32 68530.0 95.81  
 52) CA:30,45,90 24 44340.0 84.55  
 104) Pert:em 5 5847.0 45.16 \*  
 105) Pert:el 19 28690.0 94.92  
 210) Y<37284.3 13 11290.0 85.50 \*  
 211) Y>37284.3 6 13750.0 115.30 \*  
 53) CA:60,75 8 12020.0 129.60 \*  
 27) Pert:SD,chp,cp 42 91620.0 131.30  
 54) CH:4,6 7 6274.0 95.53 \*  
 55) CH:3,5 35 74580.0 138.50  
 110) X<57978.3 25 42440.0 127.80  
 220) X<54617.3 5 15340.0 164.30 \*  
 221) X>54617.3 20 18780.0 118.70 \*  
 111) X>57978.3 10 22120.0 165.20  
 222) X<58937.3 5 7021.0 201.70 \*  
 223) X>58937.3 5 1819.0 128.80 \*  
 7) CP:B,C,D,SD 172 575600.0 146.40  
 14) CH:2,4 44 117700.0 123.60  
 28) Y<39976.8 22 54720.0 107.20  
 56) X<57759.8 13 26140.0 127.30  
 112) X<56422.2 6 3162.0 105.70 \*  
 113) X>56422.2 7 17770.0 145.80 \*  
 57) X>57759.8 9 15760.0 78.14 \*  
 29) Y>39976.8 22 51120.0 140.00  
 58) Y<43243.5 16 18280.0 153.90  
 116) X<56695.2 8 7989.0 130.60 \*  
 117) X>56695.2 8 1620.0 177.10 \*  
 59) Y>43243.5 6 21510.0 102.90 \*  
 15) CH:3,6,sd 128 427000.0 154.30  
 30) Pert:chp,em 14 39570.0 115.20  
 60) CP:B,D 8 7950.0 90.50 \*  
 61) CP:C 6 20270.0 148.00 \*  
 31) Pert:SD,el 114 363400.0 159.10  
 62) X<54164.3 8 6972.0 120.20 \*  
 63) X>54164.3 106 343400.0 162.00  
 126) CA:0,45,90 92 266700.0 157.70  
 252) X<54788.3 14 23340.0 181.00  
 504) Y<39083.5 8 6751.0 197.50 \*  
 505) Y>39083.5 6 11550.0 159.10 \*  
 253) X>54788.3 78 234300.0 153.50  
 506) Y<39447.3 31 86810.0 136.80  
 1012) Y<38479.5 22 58710.0 152.20  
 2024) Y<37985.5 16 41950.0 139.10  
 4048) X<56030.2 5 10770.0 111.00 \*

```

4049) X>56030.2 11 25450.0 151.80
8098) CD:B,D 6 20530.0 137.50 *
8099) CD:C 5 2207.0 169.00 *
2025) Y>37985.5 6 6602.0 187.30 *
1013) Y>38479.5 9 10110.0 99.16 *
507) Y>39447.3 47 133200.0 164.50
1014) X<59644.9 42 109600.0 170.80
2028) Y<42440 33 69020.0 181.10
4056) GE:BbS,CT,SD,SFi 12 13950.0 160.20
8112) X<59088.1 7 9772.0 145.80 *
8113) X>59088.1 5 718.7 180.30 *
4057) GE:SBb,SS 21 46830.0 193.00
8114) X<56241.3 5 16010.0 168.90 *
8115) X>56241.3 16 26990.0 200.60
16230) X<57307.3 5 8045.0 226.70 *
16231) X>57307.3 11 13970.0 188.70 *
2029) Y>42440 9 24440.0 133.30 *
1015) X>59644.9 5 7531.0 110.80 *
127) CA:60,75 14 63730.0 190.40
254) Y<42999.3 8 34350.0 217.80 *
255) Y>42999.3 6 15410.0 154.00 *

```

## Modèle Cartographique de taille optimale – Carto28

$Volume.total \sim AC + CA + CD + GE + CH + Orig + CP + Pert + SD + X + Y$

Variables utilisées pour la construction de l'arbre :

*AC GE CP Pert CA CH X Y*

noeud), Variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) root 338 1472000 114.30
2) AC:1966,1967,1968,1969,1970,1972,1973,1974,1975,1976,1977,1978,
1979,1987,1989,CHT,ES 60 130300 39.46
4) AC:1966,1970,1972,1973,1975,1976,1978,1987,1989 24 3137
11.30 *
5) AC:1967,1968,1969,1974,1977,1979,CHT,ES 36 95430 58.23
10) GE:BbS,SBb 15 32740 39.81 *
11) GE:SD,SS 21 53970 71.39 *
3) AC:1981,1983,1992,1993,1994,SP 278 933200 130.40
6) CP:A,E 106 242700 104.50
12) GE:BjS,EE,ES,SBb,SBj,SE 32 27610 78.11 *
13) GE:BbS,SD,SFi,SPe,SS 74 183100 116.00
26) Pert:el,em 32 68530 95.81

```

52) CA:30,45,90 24 44340 84.55  
 104) Pert:em 5 5847 45.16 \*  
 105) Pert:el 19 28690 94.92 \*  
 53) CA:60,75 8 12020 129.60 \*  
 27) Pert:SD,chp,cp 42 91620 131.30  
 54) CH:4,6 7 6274 95.53 \*  
 55) CH:3,5 35 74580 138.50  
 110) X<57978.3 25 42440 127.80 \*  
 111) X>57978.3 10 22120 165.20  
 222) X<58937.3 5 7021 201.70 \*  
 223) X>58937.3 5 1819 128.80 \*  
 7) CP:B,C,D,SD 172 575600 146.40  
 14) CH:2,4 44 117700 123.60  
 28) Y<39976.8 22 54720 107.20  
 56) X<57759.8 13 26140 127.30 \*  
 57) X>57759.8 9 15760 78.14 \*  
 29) Y>39976.8 22 51120 140.00  
 58) Y<43243.5 16 18280 153.90  
 116) X<56695.2 8 7989 130.60 \*  
 117) X>56695.2 8 1620 177.10 \*  
 59) Y>43243.5 6 21510 102.90 \*  
 15) CH:3,6,sd 128 427000 154.30  
 30) Pert:chp,em 14 39570 115.20  
 60) CP:B,D 8 7950 90.50 \*  
 61) CP:C 6 20270 148.00 \*  
 31) Pert:SD,el 114 363400 159.10  
 62) X<54164.3 8 6972 120.20 \*  
 63) X>54164.3 106 343400 162.00  
 126) CA:0,45,90 92 266700 157.70  
 252) X<54788.3 14 23340 181.00 \*  
 253) X>54788.3 78 234300 153.50  
 506) Y<39447.3 31 86810 136.80  
 1012) Y<38479.5 22 58710 152.20  
 2024) Y<37985.5 16 41950 139.10 \*  
 2025) Y>37985.5 6 6602 187.30 \*  
 1013) Y>38479.5 9 10110 99.16 \*  
 507) Y>39447.3 47 133200 164.50  
 1014) X<59644.9 42 109600 170.80  
 2028) Y<42440 33 69020 181.10 \*  
 2029) Y>42440 9 24440 133.30 \*  
 1015) X>59644.9 5 7531 110.80 \*  
 127) CA:60,75 14 63730 190.40  
 254) Y<42999.3 8 34350 217.80 \*  
 255) Y>42999.3 6 15410 154.00 \*

## Modèle Cartographique, version non géographique – Carto36

*Volume.total ~ AC + CA + CD + GE + CH + Orig + CP + Pert + SD*

Variables utilisées pour la construction de l'arbre :

*AC GE SD CP Pert CA CH CD*

noeud), Variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) root 338 1472000.0 114.300
  2) AC:1966,1967,1968,1969,1970,1972,1973,1974,1975,1976,1977,1978,
1979,1987,1989,CHT,ES 60 130300.0 39.460
  4) AC:1966,1970,1972,1973,1975,1976,1978,1987,1989 24 3137.0
11.300
  8) GE:BbS,CT,SD 14 762.9 5.957 *
  9) GE:SBb,SS 10 1415.0 18.780 *
  5) AC:1967,1968,1969,1974,1977,1979,CHT,ES 36 95430.0 58.230
 10) GE:BbS,SBb 15 32740.0 39.810
  20) AC:1967,1969,1979,CHT,ES 9 2478.0 23.090 *
  21) AC:1968,1974 6 23970.0 64.900 *
 11) GE:SD,SS 21 53970.0 71.390
  22) AC:1967,1977,CHT 14 27490.0 62.260
  44) SD:GAULIS 5 1978.0 33.080 *
  45) SD:FOURRE,PERCHIS 9 18890.0 78.470 *
  23) AC:1969,1979,ES 7 22980.0 89.640 *
3) AC:1981,1983,1992,1993,1994,SP 278 933200.0 130.400
  6) CP:A,E 106 242700.0 104.500
 12) GE:BjS,EE,ES,SBb,SBj,SE 32 27610.0 78.110
  24) GE:BjS,EE,ES,SBj,SE 24 16170.0 72.440
  48) Pert:chp,el 7 5813.0 59.290 *
  49) Pert:SD 17 8644.0 77.860
  98) CA:60,75,90 7 1071.0 70.940 *
  99) CA:45 10 7004.0 82.700 *
  25) GE:SBb 8 8351.0 95.120 *
 13) GE:BbS,SD,SFi,SPe,SS 74 183100.0 116.000
  26) Pert:el,em 32 68530.0 95.810
  52) CA:30,45,90 24 44340.0 84.550
 104) Pert:em 5 5847.0 45.160 *
 105) Pert:el 19 28690.0 94.920
  210) SD:FUTAIE 6 5287.0 79.720 *
  211) SD:PERCHIS 13 21380.0 101.900 *
  53) CA:60,75 8 12020.0 129.600 *
  27) Pert:SD,chp,cp 42 91620.0 131.300
  54) CH:4,6 7 6274.0 95.530 *
  55) CH:3,5 35 74580.0 138.500
 110) CA:30,60 7 4757.0 112.900 *
 111) CA:15,45,75,90 28 64080.0 144.900
  222) Pert:chp,cp 5 4390.0 133.900 *
  223) Pert:SD 23 58960.0 147.300 *
7) CP:B,C,D,SD 172 575600.0 146.400
 14) CH:2,4 44 117700.0 123.600
  28) CA:15,45,75 13 28620.0 98.980

```

56) CP:B 7 8867.0 70.600 \*  
 57) CP:D 6 7530.0 132.100 \*  
 29) CA:30 31 77910.0 133.900  
 58) CD:A 6 18800.0 117.100 \*  
 59) CD:B,C 25 57010.0 137.900  
 118) GE:SS 19 42240.0 135.200 \*  
 119) GE:BbS,SBb,SPe 6 14180.0 146.500 \*  
 15) CH:3,6,T 128 427000.0 154.300  
 30) Pert:chp,em 14 39570.0 115.100  
 60) CP:B,D 8 7950.0 90.500 \*  
 61) CP:C 6 20270.0 148.000 \*  
 31) Pert:SD,el 114 363400.0 159.100  
 62) CA:0,45,90 98 281100.0 155.400  
 124) CD:C,D 29 64310.0 143.600  
 248) Pert:SD 14 36960.0 139.800  
 496) GE:SS 7 30260.0 132.000 \*  
 497) GE:SBb,SD,SFi 7 5861.0 147.500 \*  
 249) Pert:el 15 26940.0 147.300  
 498) SD:PERCHIS 8 13640.0 144.900 \*  
 499) SD:FUTAIE 7 13200.0 149.900 \*  
 125) CD:A,B,T 69 211000.0 160.400  
 250) GE:CT,SS 63 190500.0 156.700  
 500) CP:C,D 23 80460.0 136.700 \*  
 501) CP:B,SD 40 95540.0 168.200  
 1002) AC:1993,1994 14 38340.0 160.600 \*  
 1003) AC:SP 26 55940.0 172.400 \*  
 251) GE:BbS,SBb 6 10780.0 198.800 \*  
 63) CA:60,75 16 73190.0 181.300  
 126) CP:C 8 29500.0 159.500 \*  
 127) CP:B,D 8 36130.0 203.000 \*

## Modèle Cartographique de taille optimale, version non géographique – Carto21

*Volume total* ~ *AC* + *CA* + *CD* + *GE* + *CH* + *Orig* + *CP* + *Pert* + *SD*

Variables utilisées pour la construction de l'arbre :

*AC GE CP Pert CA CH CD*

noeud), Variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

1) root 338 1472000 114.30

- 2) AC:1966,1967,1968,1969,1970,1972,1973,1974,1975,1976,1977,1978,  
1979,1987,1989,CHT,ES 60 130300 39.46
- 4) AC:1966,1970,1972,1973,1975,1976,1978,1987,1989 24 3137
- 11.30 \*
- 5) AC:1967,1968,1969,1974,1977,1979,CHT,ES 36 95430 58.23
- 10) GE:BbS,SBb 15 32740 39.81
- 20) AC:1967,1969,1979,CHT,ES 9 2478 23.09 \*
- 21) AC:1968,1974 6 23970 64.90 \*
- 11) GE:SD,SS 21 53970 71.39 \*
- 3) AC:1981,1983,1992,1993,1994,SP 278 933200 130.40
- 6) CP:A,E 106 242700 104.50
- 12) GE:BjS,EE,ES,SBb,SBj,SE 32 27610 78.11 \*
- 13) GE:BbS,SD,SFi,SPe,SS 74 183100 116.00
- 26) Pert:el,em 32 68530 95.81
- 52) CA:30,45,90 24 44340 84.55
- 104) Pert:em 5 5847 45.16 \*
- 105) Pert:el 19 28690 94.92 \*
- 53) CA:60,75 8 12020 129.60 \*
- 27) Pert:SD,chp,cp 42 91620 131.30
- 54) CH:4,6 7 6274 95.53 \*
- 55) CH:3,5 35 74580 138.50 \*
- 7) CP:B,C,D,SD 172 575600 146.40
- 14) CH:2,4 44 117700 123.60
- 28) CA:15,45,75 13 28620 98.98
- 56) CP:B 7 8867 70.60 \*
- 57) CP:D 6 7530 132.10 \*
- 29) CA:30 31 77910 133.90 \*
- 15) CH:3,6,T 128 427000 154.30
- 30) Pert:chp,em 14 39570 115.10
- 60) CP:B,D 8 7950 90.50 \*
- 61) CP:C 6 20270 148.00 \*
- 31) Pert:SD,el 114 363400 159.10
- 62) CA:0,45,90 98 281100 155.40
- 124) CD:C,D 29 64310 143.60 \*
- 125) CD:A,B,T 69 211000 160.40
- 250) GE:CT,SS 63 190500 156.70
- 500) CP:C,D 23 80460 136.70 \*
- 501) CP:B,SD 40 95540 168.20 \*
- 251) GE:BbS,SBb 6 10780 198.80 \*
- 63) CA:60,75 16 73190 181.30
- 126) CP:C 8 29500 159.50 \*
- 127) CP:B,D 8 36130 203.00 \*

## Modèle Complet de Terrain général - CompletTG

*Volume total ~ Alt + CAT + CDT + CHT + CPT + DpS + Dr + EMO + Exp + FP + GET  
+ HD + Inc + OrigT + PertT + Pier + SP + THB + THC + TH + Vrs + X + Y*

Variables utilisées dans le modèle :

*HD CHT OrigT...THB GET THC X Y DpS Dr CDT Pier CPT Alt EMO  
Exp*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

- 1) racine 329 1,441e+006 115,1000
- 2) HD<12,5 117 1,578e+005 52,5100
- 4) CHT: 5,6 56 3,884e+004 25,2800
- 8) OrigT: CT,sd 43 2,385e+004 19,5700
- 16) THB: L,LI,LSM,SF,SFL,STFL,STGL,sd 19 1,966e+003 8,7420
- 32) GET: Bb,BbS,SE 11 4,660e+002 3,3360
- 64) THC: LSF,LSG,LSM,SML,STGL 6 2,128e+000 0,4833 \*
- 65) THC: SGL,sd 5 3,564e+002 6,7600 \*
- 33) GET: SBb,SS 8 7,361e+002 16,1800 \*
- 17) THB: LSF,LSG,LSTF,SGL,SML 24 1,789e+004 28,1400
- 34) X<57155,6 5 3,016e+003 55,1800 \*
- 35) X>57155,6 19 1,026e+004 21,0200
- 70) Y<38885,5 14 1,761e+003 15,0200
- 140) THC: LSF,SML 5 1,246e+002 5,5800 \*
- 141) THC: LSM,SFL,SGL,sd 9 9,432e+002 20,2700 \*
- 71) Y>38885,5 5 6,584e+003 37,8200 \*
- 9) OrigT: ES 13 8,957e+003 44,1500
- 18) THC: LSM,SG,STG 5 9,741e+002 21,4000 \*
- 19) THC: LSF,SFL,SML,STGL,sd 8 3,776e+003 58,3800 \*
- 5) CHT: 3,4 61 3,931e+004 77,5100
- 10) GET: BbS,EE,SBb,SE,SFi 24 8,347e+003 61,9800
- 20) THC: L,LSF,LSM,LSTF,SM 6 2,614e+003 43,7000 \*
- 21) THC: SF,SG,SGL,SML,STG,sd 18 3,058e+003 68,0800
- 42) DpS: 1A,1AY,7AM,8C 11 1,388e+003 61,7000
- 84) X<56560,2 6 6,710e+002 68,5300 \*
- 85) X>56560,2 5 1,011e+002 53,5000 \*
- 43) DpS: 1AD,1AM,7TM 7 5,194e+002 78,1000 \*
- 11) GET: ES,SS 37 2,143e+004 87,5800
- 22) THB: LLI,LSM,LSTF 13 2,238e+003 66,9700
- 44) Dr: 10,30,40 8 7,085e+002 59,4100 \*
- 45) Dr: 20,41,50 5 3,418e+002 79,0600 \*
- 23) THB: L,LSA,LSF,SFL,SML,STGL 24 1,067e+004 98,7500
- 46) CDT: A,C,D 14 4,060e+003 88,8600
- 92) HD<11,5 6 7,585e+002 73,9300 \*
- 93) HD>11,5 8 9,605e+002 100,1000 \*
- 47) CDT: B 10 3,334e+003 112,6000
- 94) X<54804,1 5 8,439e+002 123,6000 \*
- 95) X>54804,1 5 1,275e+003 101,6000 \*
- 3) HD>12,5 212 5,717e+005 149,7000
- 6) CDT: D 29 3,016e+004 99,5400
- 12) THB: LSF,LSM,LSTF 15 6,987e+003 78,1200
- 24) Pier<20 7 1,472e+003 95,5100 \*
- 25) Pier>20 8 1,544e+003 62,9000 \*
- 13) THB: L,LLI,SFL,SGL,SML,STGL 14 8,920e+003 122,5000

26) THC: LSF,SFL,SML,sd 9 4,544e+003 112,1000 \*  
 27) THC: SGL,STGL 5 1,631e+003 141,3000 \*  
 7) CDT: A,B,C 183 4,572e+005 157,6000  
 14) HD<15,5 105 2,162e+005 139,7000  
 28) CHT: 4,5 17 2,840e+004 96,2400  
 56) THB: LLI,LSF,LSM,SFL 8 8,937e+003 70,8500 \*  
 57) THB: L,LSA,LSTF,SGL,SML 9 9,720e+003 118,8000 \*  
 29) CHT: 3 88 1,494e+005 148,1000  
 58) GET: EE,ES,SBb,SE 18 1,761e+004 116,6000  
 116) THB: LLA,LSF,STGL,sd 12 7,372e+003 100,6000  
 232) THC: LSM,SGL,SML 7 3,636e+003 87,9600 \*  
 233) THC: LSF,SG,sd 5 1,043e+003 118,3000 \*  
 117) THB: L,LSTF,STFL 6 1,082e+003 148,4000 \*  
 59) GET: BbS,SS 70 1,093e+005 156,3000  
 118) Pier<37,5 48 6,553e+004 167,1000  
 236) CPT: B,C,D,F 43 4,568e+004 160,8000  
 472) Y<37617,8 11 6,555e+003 135,7000  
 944) Pier<10 6 1,011e+003 120,5000 \*  
 945) Pier>10 5 2,487e+003 154,0000 \*  
 473) Y>37617,8 32 2,985e+004 169,4000  
 946) Alt<815,5 18 1,311e+004 157,1000  
 1892) THC: LSG,SF,SGL,SML,sd 9 4,263e+003 138,0000 \*  
 1893) THC: LLI,LSF,LSM,SFL,SM 9 2,253e+003 176,2000 \*  
 947) Alt>815,5 14 1,051e+004 185,2000  
 1894) EMO<6,5 5 3,326e+003 210,9000 \*  
 1895) EMO>6,5 9 2,057e+003 171,0000 \*  
 237) CPT: A,E 5 3,284e+003 221,6000 \*  
 119) Pier>37,5 22 2,567e+004 132,5000  
 238) THC: LA,LSF,SF,SML 6 4,501e+003 96,4000 \*  
 239) THC: LSG,LSM,LSTF,NON OBSEVAB,SFL,SG,SGL,sd 16  
 1,040e+004 146,1000  
 478) Y<39488,3 9 2,173e+003 133,1000 \*  
 479) Y>39488,3 7 4,755e+003 162,8000 \*  
 15) HD>15,5 78 1,624e+005 181,6000  
 30) THC: LLI,LSA,LSF,LSTF,STFL,STGL,sd 27 3,777e+004 154,3000  
 60) Dr: 21,30,31,40 20 1,616e+004 142,3000  
 120) THB: L,LSF,LSTF,SFL,SML 13 8,155e+003 131,6000  
 240) EMO<8 7 4,257e+003 143,7000 \*  
 241) EMO>8 6 1,667e+003 117,5000 \*  
 121) THB: LLI,LSA,LSTG,STFL,STGL,sd 7 3,781e+003 162,1000 \*  
 61) Dr: 20 7 1,041e+004 188,7000 \*  
 31) THC: L,LA,LSG,LSM,SFL,SGL,SM,SML,STG 51 9,381e+004 196,1000  
 62) THB: L,LSG,LSM,SFL,STFL,STGL 25 2,905e+004 175,7000  
 124) CPT: B,C,E 11 3,939e+003 159,0000  
 248) Exp<269,5 6 2,029e+003 170,1000 \*  
 249) Exp>269,5 5 2,915e+002 145,7000 \*  
 125) CPT: D,F 14 1,965e+004 188,8000  
 250) THB: L,STFL,STGL 7 9,724e+003 167,9000 \*  
 251) THB: LSM,SFL 7 3,818e+003 209,7000 \*  
 63) THB: LLI,LSA,LSF,LSTF,SM,SML 26 4,434e+004 215,7000  
 126) Alt<849 21 2,906e+004 226,5000  
 252) Alt<682,5 5 6,568e+003 184,2000 \*  
 253) Alt>682,5 16 1,079e+004 239,7000  
 506) HD<17,5 7 1,231e+003 220,2000 \*



17) Y>35164,8 38 14240,00 16,340  
 34) Y<38885,5 32 4821,00 13,490  
 68) CPT: B,E,F 8 118,00 4,063 \*  
 69) CPT: A,C,D 24 3755,00 16,630  
 138) GET: BbS,Fi,SE 11 1509,00 10,960  
 276) X<57263,8 6 1095,00 14,730 \*  
 277) X>57263,8 5 225,70 6,440 \*  
 139) GET: Bb,sd,SBb,SS 13 1593,00 21,430  
 278) Y<37177,5 5 249,90 13,460 \*  
 279) Y>37177,5 8 826,90 26,410 \*  
 35) Y>38885,5 6 7776,00 31,520 \*  
 9) OrigT: ES 13 8957,00 44,150  
 18) X<58057,5 7 4889,00 32,500 \*  
 19) X>58057,5 6 2009,00 57,750 \*  
 5) CHT: 4 67 59940,00 80,980  
 10) GET: BbS,EE,ES,SE 22 7921,00 63,020  
 20) CAT: 30,70 5 1677,00 42,860 \*  
 21) CAT: 50 17 3614,00 68,950  
 42) CPT: A,C 12 1818,00 63,630  
 84) Y<42061,5 5 98,83 56,620 \*  
 85) Y>42061,5 7 1298,00 68,640 \*  
 43) CPT: B,D 5 641,20 81,720 \*  
 11) GET: SBb,SS 45 41460,00 89,750  
 22) CDT: A,C,D 27 14530,00 80,770  
 44) CPT: A,B,D,F 21 7043,00 74,350  
 88) CPT: A,F 5 990,80 60,100 \*  
 89) CPT: B,D 16 4719,00 78,810  
 178) Y<43437,3 11 3144,00 73,590  
 356) X<56620 6 399,30 72,320 \*  
 357) X>56620 5 2723,00 75,120 \*  
 179) Y>43437,3 5 617,50 90,280 \*  
 45) CPT: C,E 6 3598,00 103,200 \*  
 23) CDT: B 18 21480,00 103,200  
 46) CPT: B,C,D 13 8346,00 93,920  
 92) Y<40517,5 8 5755,00 86,340 \*  
 93) Y>40517,5 5 1396,00 106,000 \*  
 47) CPT: A,F 5 9078,00 127,400 \*  
 3) CHT: 2,3,sd 212 561400,00 149,700  
 6) CDT: D,sd 32 33500,00 98,900  
 12) Y<33940,5 7 1475,00 68,160 \*  
 13) Y>33940,5 25 23560,00 107,500  
 26) CPT: A,B,C,D 18 13450,00 97,760  
 52) X<56386,4 12 10370,00 90,590  
 104) CAT: 50,70120,sd 5 2141,00 66,640 \*  
 105) CAT: 70,90 7 3308,00 107,700 \*  
 53) X>56386,4 6 1232,00 112,100 \*  
 27) CPT: E,F 7 4006,00 132,600 \*  
 7) CDT: A,B,C 180 430700,00 158,700  
 14) CPT: A,B,C 61 124300,00 134,800  
 28) GET: EE,ES,SBb,SE,SFi 13 9787,00 91,180  
 56) Y<37252,3 5 3677,00 70,680 \*  
 57) Y>37252,3 8 2696,00 104,000 \*  
 29) GET: BbS,SS 48 83080,00 146,700  
 58) CDT: A,C 24 41270,00 132,300

116) PertT: CA,sd 15 23500,00 121,400  
 232) X<55213,2 5 6230,00 140,700 \*  
 233) X>55213,2 10 14490,00 111,800  
 466) X<55850,8 5 2715,00 94,480 \*  
 467) X>55850,8 5 8783,00 129,100 \*  
 117) PertT: EL 9 13050,00 150,400 \*  
 59) CDT: B 24 31890,00 161,000  
 118) X<56187,3 15 22250,00 152,400  
 236) Y<37480,8 5 744,60 136,500 \*  
 237) Y>37480,8 10 19630,00 160,300  
 474) X<54649,1 5 16080,00 165,100 \*  
 475) X>54649,1 5 3314,00 155,500 \*  
 119) X>56187,3 9 6646,00 175,500 \*  
 15) CPT: D,E,F 119 253600,00 171,000  
 30) Y<37708,5 26 43980,00 148,200  
 60) Y<35020,5 5 11220,00 189,000 \*  
 61) Y>35020,5 21 22480,00 138,500  
 122) X<57160,9 16 8102,00 129,700  
 244) X<55534,4 8 3379,00 138,200 \*  
 245) X>55534,4 8 3569,00 121,200 \*  
 123) X>57160,9 5 9184,00 166,600 \*  
 31) Y>37708,5 93 192400,00 177,400  
 62) CHT: 3 86 162200,00 173,800  
 124) GET: Bbs,SBb,SE 23 40970,00 156,500  
 248) CDT: A,B 15 22160,00 143,800  
 496) CPT: D 7 5127,00 121,700 \*  
 497) CPT: E,F 8 10630,00 163,100 \*  
 249) CDT: C 8 11860,00 180,300 \*  
 125) GET: SS 63 111800,00 180,200  
 250) X<56500,5 19 20780,00 195,300  
 500) Y<40271 14 7792,00 210,500  
 1000) Y<38368,8 5 2285,00 198,000 \*  
 1001) Y>38368,8 9 4287,00 217,500 \*  
 501) Y>40271 5 758,90 152,900 \*  
 251) X>56500,5 44 84760,00 173,600  
 502) X<57110,4 5 2373,00 133,700 \*  
 503) X>57110,4 39 73400,00 178,800  
 1006) Y<39947 12 27030,00 157,700  
 2012) X<58527,9 6 15230,00 179,000 \*  
 2013) X>58527,9 6 6380,00 136,500 \*  
 1007) Y>39947 27 38700,00 188,100  
 2014) Y<43199,8 18 20980,00 201,200  
 4028) Y<41225 9 5630,00 182,200 \*  
 4029) Y>41225 9 8797,00 220,300 \*  
 2015) Y>43199,8 9 8400,00 161,800 \*  
 63) CHT: 2 7 16110,00 220,400 \*

## Modèle Partiel de Terrain taillé à 45 feuilles - PartielT45

*Volume total ~ CAT + CDT + CHT + CPT + GET + OrigT + PertT + X + Y*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) racine 338 1472000,0 114,30
2) CHT:4,5,6 126 197600,0 54,72
4) CHT:5,6 59 39050,0 24,90
8) OrigT:CT,sd 46 23910,0 19,46
16) Y<35164,8 8 7539,0 34,31 *
17) Y>35164,8 38 14240,0 16,34
34) Y<38885,5 32 4821,0 13,49 *
35) Y>38885,5 6 7776,0 31,52 *
9) OrigT:ES 13 8957,0 44,15
18) X<58057,5 7 4889,0 32,50 *
19) X>58057,5 6 2009,0 57,75 *
5) CHT:4 67 59940,0 80,98
10) GET:BbS,EE,ES,SE 22 7921,0 63,02
20) CAT:30,70 5 1677,0 42,86 *
21) CAT:50 17 3614,0 68,95
42) CPT:A,C 12 1818,0 63,63 *
43) CPT:B,D 5 641,2 81,72 *
11) GET:SBb,SS 45 41460,0 89,75
22) CDT:A,C,D 27 14530,0 80,77
44) CPT:A,B,D,F 21 7043,0 74,35
88) CPT:A,F 5 990,8 60,10 *
89) CPT:B,D 16 4719,0 78,81 *
45) CPT:C,E 6 3598,0 103,20 *
23) CDT:B 18 21480,0 103,20
46) CPT:B,C,D 13 8346,0 93,92
92) Y<40517,5 8 5755,0 86,34 *
93) Y>40517,5 5 1396,0 106,00 *
47) CPT:A,F 5 9078,0 127,40 *
3) CHT:2,3,sd 212 561400,0 149,70
6) CDT:D,sd 32 33500,0 98,90
12) Y<33940,5 7 1475,0 68,16 *
13) Y>33940,5 25 23560,0 107,50
26) CPT:A,B,C,D 18 13450,0 97,76
52) X<56386,4 12 10370,0 90,59
104) CAT:50,70,120,sd 5 2141,0 66,64 *
105) CAT:70,90 7 3308,0 107,70 *
53) X>56386,4 6 1232,0 112,10 *
27) CPT:E,F 7 4006,0 132,60 *
7) CDT:A,B,C 180 430700,0 158,70
14) CPT:A,B,C 61 124300,0 134,80
28) GET:EE,ES,SBb,SE,SFi 13 9787,0 91,18
56) Y<37252,3 5 3677,0 70,68 *
57) Y>37252,3 8 2696,0 104,00 *
29) GET:BbS,SS 48 83080,0 146,70
58) CDT:A,C 24 41270,0 132,30
116) PertT:CA,sd 15 23500,0 121,40
232) X<55213,2 5 6230,0 140,70 *
233) X>55213,2 10 14490,0 111,80

```

466) X<55850,8 5 2715,0 94,48 \*  
 467) X>55850,8 5 8783,0 129,10 \*  
 117) PertT:EL 9 13050,0 150,40 \*  
 59) CDT:B 24 31890,0 161,00  
 118) X<56187,3 15 22250,0 152,40  
 236) Y<37480,8 5 744,6 136,50 \*  
 237) Y>37480,8 10 19630,0 160,30 \*  
 119) X>56187,3 9 6646,0 175,50 \*  
 15) CPT:D,E,F 119 253600,0 171,00  
 30) Y<37708,5 26 43980,0 148,20  
 60) Y<35020,5 5 11220,0 189,00 \*  
 61) Y>35020,5 21 22480,0 138,50  
 122) X<57160,9 16 8102,0 129,70  
 244) X<55534,4 8 3379,0 138,20 \*  
 245) X>55534,4 8 3569,0 121,20 \*  
 123) X>57160,9 5 9184,0 166,60 \*  
 31) Y>37708,5 93 192400,0 177,40  
 62) CHT:3 86 162200,0 173,80  
 124) GET:BbS,SBb,SE 23 40970,0 156,50  
 248) CDT:A,B 15 22160,0 143,80  
 496) CPT:D 7 5127,0 121,70 \*  
 497) CPT:E,F 8 10630,0 163,10 \*  
 249) CDT:C 8 11860,0 180,30 \*  
 125) GET:SS 63 111800,0 180,20  
 250) X<56500,5 19 20780,0 195,30  
 500) Y<40271 14 7792,0 210,50  
 1000) Y<38368,8 5 2285,0 198,00 \*  
 1001) Y>38368,8 9 4287,0 217,50 \*  
 501) Y>40271 5 758,9 152,90 \*  
 251) X>56500,5 44 84760,0 173,60  
 502) X<57110,4 5 2373,0 133,70 \*  
 503) X>57110,4 39 73400,0 178,80  
 1006) Y<39947 12 27030,0 157,70  
 2012) X<58527,9 6 15230,0 179,00 \*  
 2013) X>58527,9 6 6380,0 136,50 \*  
 1007) Y>39947 27 38700,0 188,10  
 2014) Y<43199,8 18 20980,0 201,20  
 4028) Y<41225 9 5630,0 182,20 \*  
 4029) Y>41225 9 8797,0 220,30 \*  
 2015) Y>43199,8 9 8400,0 161,80 \*  
 63) CHT:2 7 16110,0 220,40 \*

## Modèle Partiel de Terrain de taille optimale – PartielT17

*Volume total ~ CAT + CDT + CHT + CPT + GET + OrigT + PertT + X + Y,*

Variables utilisées pour la construction de l'arbre :

*CHT...GET...CDT...CPT...Y...X*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

- 1) racine 338 1472000,0 114,30
- 2) CHT:4,5,6 126 197600,0 54,72
- 4) CHT:5,6 59 39050,0 24,90 \*
- 5) CHT:4 67 59940,0 80,98
- 10) GET:BbS,EE,ES,SE 22 7921,0 63,02 \*
- 11) GET:SBb,SS 45 41460,0 89,75 \*
- 3) CHT:2,3,sd 212 561400,0 149,70
- 6) CDT:D,sd 32 33500,0 98,90 \*
- 7) CDT:A,B,C 180 430700,0 158,70
- 14) CPT:A,B,C 61 124300,0 134,80
- 28) GET:EE,ES,SBb,SE,SFi 13 9787,0 91,18 \*
- 29) GET:BbS,SS 48 83080,0 146,70
- 58) CDT:A,C 24 41270,0 132,30 \*
- 59) CDT:B 24 31890,0 161,00 \*
- 15) CPT:D,E,F 119 253600,0 171,00
- 30) Y<37708,5 26 43980,0 148,20
- 60) Y<35020,5 5 11220,0 189,00 \*
- 61) Y>35020,5 21 22480,0 138,50 \*
- 31) Y>37708,5 93 192400,0 177,40
- 62) CHT:3 86 162200,0 173,80
- 124) GET:BbS,SBb,SE 23 40970,0 156,50 \*
- 125) GET:SS 63 111800,0 180,20
- 250) X<56500,5 19 20780,0 195,30
- 500) Y<40271 14 7792,0 210,50 \*
- 501) Y>40271 5 758,9 152,90 \*
- 251) X>56500,5 44 84760,0 173,60
- 502) X<57110,4 5 2373,0 133,70 \*
- 503) X>57110,4 39 73400,0 178,80
- 1006) Y<39947 12 27030,0 157,70 \*
- 1007) Y>39947 27 38700,0 188,10
- 2014) Y<43199,8 18 20980,0 201,20 \*
- 2015) Y>43199,8 9 8400,0 161,80 \*
- 63) CHT:2 7 16110,0 220,40 \*

## Modèle Partiel de Terrain, version non géographique – PartielT42

*Volume total* ~ *CAT* + *CDT* + *CHT* + *CPT* + *GET* + *OrigT* + *PertT*

Variables utilisées pour la construction de l'arbre :

*CAT CDT CHT CPT GET OrigT PertT*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```

1) root 338 1472000.0 114.300
  2) CHT:4,5,6 126 197600.0 54.720
    4) CHT:5,6 59 39050.0 24.900
      8) OrigT:CT,SD 46 23910.0 19.460
        16) CPT:B,C,F 15 1566.0 9.767
          32) GET:BbS 8 284.7 3.962 *
          33) GET:SD,SBb,SS 7 704.2 16.400 *
        17) CPT:A,D,E 31 20250.0 24.150
          34) GET:Bb,Fi,SE 5 338.3 12.040 *
          35) GET:BbS,SBb,SS 26 19040.0 26.480
            70) PertT:EPC 6 1521.0 19.980 *
            71) PertT:SD 20 17190.0 28.430
              142) GET:BbS,SBb 15 10820.0 24.870
                284) CDT:A,C 6 1147.0 17.880 *
                285) CDT:B,T 9 9181.0 29.520 *
              143) GET:SS 5 5609.0 39.140 *
        9) OrigT:ES 13 8957.0 44.150
          18) CPT:D,F 6 1907.0 34.770 *
          19) CPT:B,C,E 7 6068.0 52.200 *
    5) CHT:4 67 59940.0 80.980
      10) GET:BbS,EE,ES,SE 22 7921.0 63.020
        20) CAT:30,70 5 1677.0 42.860 *
        21) CAT:50 17 3614.0 68.950
          42) CPT:A,C 12 1818.0 63.630
            84) CDT:C,D 6 510.5 57.850 *
            85) CDT:B 6 906.4 69.420 *
          43) CPT:B,D 5 641.2 81.720 *
    11) GET:SBb,SS 45 41460.0 89.750
      22) CDT:A,C,D 27 14530.0 80.770
        44) CPT:A,B,D,F 21 7043.0 74.350
          88) CPT:A,F 5 990.8 60.100 *
          89) CPT:B,D 16 4719.0 78.810
            178) CPT:B 5 461.9 72.260 *
            179) CPT:D 11 3945.0 81.780
              358) CDT:A 5 325.5 75.980 *
              359) CDT:C 6 3311.0 86.620 *
          45) CPT:C,E 6 3598.0 103.200 *
    23) CDT:B 18 21480.0 103.200
      46) CPT:B,C,D 13 8346.0 93.920
        92) CPT:B,D 8 5347.0 91.260 *
        93) CPT:C 5 2852.0 98.160 *
        47) CPT:A,F 5 9078.0 127.400 *
    3) CHT:2,3,T 212 561400.0 149.700
      6) CDT:D,T 32 33500.0 98.900

```

12) CPT:A,B,D,F 24 21240.0 92.880  
 24) PertT:CHP,EL 19 13520.0 86.590  
 48) GET:BbS,Bj,BjR,SBb 8 2291.0 71.970 \*  
 49) GET:SE,SS 11 8282.0 97.220 \*  
 25) PertT:SD 5 4105.0 116.800 \*  
 13) CPT:C,E 8 8789.0 116.900 \*  
 7) CDT:A,B,C 180 430700.0 158.700  
 14) CPT:A,B,C 61 124300.0 134.800  
 28) GET:EE,ES,SBb,SE,SFi 13 9787.0 91.180  
 56) GET:EE,SBb,SFi 5 4940.0 73.840 \*  
 57) GET:ES,SE 8 2405.0 102.000 \*  
 29) GET:BbS,SS 48 83080.0 146.700  
 58) CDT:A,C 24 41270.0 132.300  
 116) PertT:CA,SD 15 23500.0 121.400  
 232) CPT:C 10 18860.0 115.500 \*  
 233) CPT:A,B 5 3588.0 133.300 \*  
 117) PertT:EL 9 13050.0 150.400 \*  
 59) CDT:B 24 31890.0 161.000 \*  
 15) CPT:D,E,F 119 253600.0 171.000  
 30) CHT:3 109 217600.0 167.400  
 60) CPT:D,F 88 175300.0 162.900  
 120) GET:BbS,SBb,SE 25 43280.0 147.200  
 240) CDT:A,B 12 14920.0 132.800 \*  
 241) CDT:C 13 23520.0 160.600  
 482) PertT:CHP,EL 7 7523.0 145.900 \*  
 483) PertT:SD 6 12730.0 177.700 \*  
 121) GET:SS 63 123400.0 169.100  
 242) CAT:120,30,90 5 7419.0 144.700 \*  
 243) CAT:50,70 58 112800.0 171.200  
 486) CPT:F 6 9311.0 156.700 \*  
 487) CPT:D 52 102000.0 172.900  
 974) PertT:EC,EL 13 35880.0 169.200 \*  
 975) PertT:SD 39 65910.0 174.200  
 1950) CDT:B 23 42520.0 170.900 \*  
 1951) CDT:A,C 16 22790.0 178.900 \*  
 61) CPT:E 21 33160.0 186.200  
 122) CDT:C 6 8523.0 160.000 \*  
 123) CDT:A,B 15 18840.0 196.700 \*  
 31) CHT:2 10 19270.0 210.100 \*

## Modèle Partiel de Terrain de taille optimale, version non géographique – PartielT9

*Volume total ~ CAT + CDT + CHT + CPT + GET + OrigT + PertT*

Variables utilisées pour la construction de l'arbre :

*CHT GET CDT CPT*

noeud), variable pour la division, nombre de p-e, déviance, valeur prédite du Volume total

\* dénote une feuille

```
1) root 338 1472000 114.30
2) CHT:4,5,6 126 197600 54.72
4) CHT:5,6 59 39050 24.90 *
5) CHT:4 67 59940 80.98
10) GET:BbS,EE,ES,SE 22 7921 63.02 *
11) GET:SBb,SS 45 41460 89.75 *
3) CHT:2,3,T 212 561400 149.70
6) CDT:D,T 32 33500 98.90 *
7) CDT:A,B,C 180 430700 158.70
14) CPT:A,B,C 61 124300 134.80
28) GET:EE,ES,SBb,SE,SFi 13 9787 91.18 *
29) GET:BbS,SS 48 83080 146.70
58) CDT:A,C 24 41270 132.30 *
59) CDT:B 24 31890 161.00 *
15) CPT:D,E,F 119 253600 171.00
30) CHT:3 109 217600 167.40 *
31) CHT:2 10 19270 210.10 *
```

## Annexe IV Les strates forestières

### Exemple de matrice de confusion

La concordance globale résulte de l'addition des éléments de la diagonale de la matrice de confusion, qui est divisée par le total d'éléments de la matrice. Dans l'exemple du Tableau 60, la valeur est égale à  $(1+68+73+10+0) / 515$ ). La concordance de chaque catégorie est la proportion de coïncidences entre la catégorie de la carte et celle de la p-e. Dans l'exemple du Tableau 60, pour la catégorie A la valeur est  $1/21$ .

Tableau 60. Matrice de confusion des catégories de la carte et de la p-e pour la variable *Classe de densité (CD)*, toutes strates forestières ensemble.

Catégorie de la p-e	Catégorie de CD de la carte					Total
	A	B	C	D	SC	
A	1	15	16	5	0	37
B	10	68	82	25	0	185
C	8	75	73	26	0	182
D	1	23	21	10	0	55
SC	1	22	23	10	0	56
Total	21	203	215	76	0	515

Le coefficient kappa (Chrisman 1991, Friedl *et al.* 2001) vise à évaluer la matrice de confusion par rapport à une matrice de confusion dont les valeurs sont aléatoires. Il se calcule selon la formule suivante (Friedl *et al.* 2001) :

$$kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})}$$

où  $N$  est le nombre d'observations de la matrice de confusion (dans l'exemple du Tableau 60,  $N=515$ ),  $r$  est le nombre de rangées de la matrice,  $x_{i+}$  et  $x_{+i}$  sont les totaux marginaux de la rangée  $i$  et de la colonne  $i$  respectivement, alors que  $x_{ii}$  est le nombre d'observations dans la colonne  $i$  et la rangée  $i$ . Pour une matrice de confusion donnée, une valeur de kappa égale à 1 indique une exactitude parfaite alors qu'une valeur de 0 indique une concordance similaire à celle obtenue aléatoirement. Une valeur négative indique, pour sa part, que la concordance est inférieure à celle obtenue par hasard.

## Analyses univariées

Cette section contient les résultats de l'analyse de concordance univariée pour les vingt strates forestières qui ont obtenu les valeurs les plus élevées d'ICL et ICV.

Les strates qui présentent la concordance maximale à l'échelle locale pour *GE* sont SS C3 75 et SS C3 60 chp (1,0), SSA430 (0,83), SS B3 45 el et SS C3 60 el (0,80) (Figure 55a) tandis que celles qui la présentent à l'échelle du voisinage spatial sont SBb C3 45 (0,51), SBb C4 30 (0,50), C6 es (0,48), SBb B3 45 et SS D3 45 em (0,46). En général, les strates SS exhibent une concordance élevée à l'échelle locale et faible à l'échelle du voisinage. En effet, SS C3 75 qui présente la valeur maximale d'ICL présente également la valeur minimale d'ICV (Figure 55a).

Par rapport à *CD*, les strates dont la concordance locale est maximale sont SS C5 15 ct (1,0) et SBb C3 45 (0,80) alors que celles dont le maximum est à l'échelle du voisinage spatial sont SBb C4 30 (0,57), SBb C3 45, SS C3 60 el, SS B3 45 el (0,52) et SS C4 30 (0,50). De plus, pour cette variable, les strates qui ont des ICL élevés présentent des ICV faibles et vice-versa (Figure 55b).

Relativement à *CH*, la concordance maximale à l'échelle locale correspond à SBb D3 75 el (1,0), SS C3 75 el (1,0) et SS C3 45 cp (0,80) alors que celle à l'échelle du voisinage correspond à SBb D3 75 el. Pour cette variable, certaines strates comme BbS B5 15 ct et SS C3 60 présentent un ICL et un ICV faibles. D'autres strates, telle Bb D3 75 el, ont un ICL et ICV élevés (Figure 55c).

Les strates cartographiques qui montrent la concordance locale la plus élevée pour *CA* sont SS C3 45 chp (1,0) et SE C4 45 (0,88) alors que celles qui la révèlent pour le voisinage spatial sont SS C3 45 chp (0,53) et BbS A4 15 ct (0,48). La strate SS C3 45 chp présente les valeurs maximales d'ICL et d'ICV, tandis que BbS A4 15 ct obtient la valeur minimale d'ICL et la maximale d'ICV (Figure 55d).

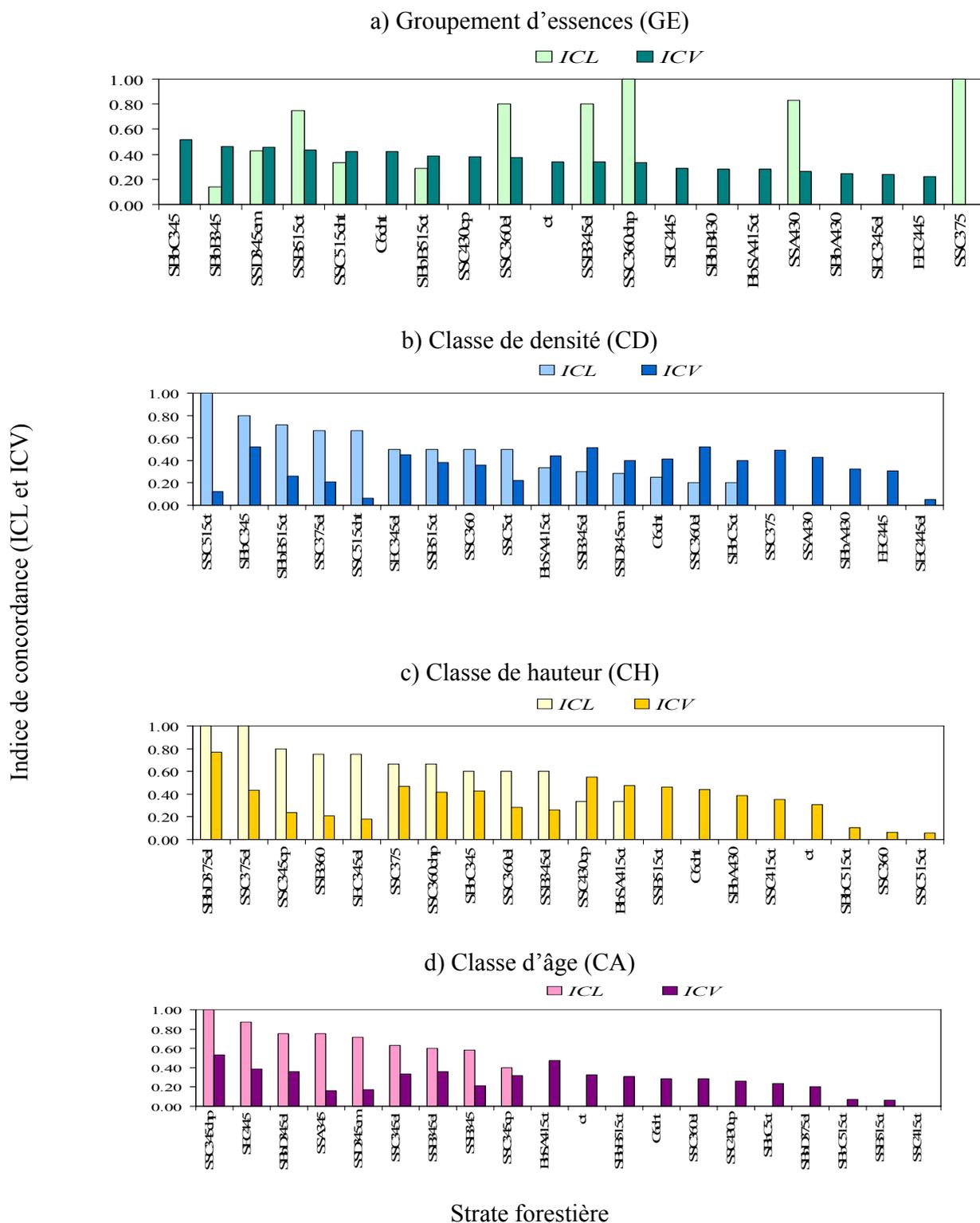


Figure 55. Indice de concordance carte/p-e des strates forestières pour les variables individuelles aux échelles locales (ICL) et du voisinage spatial (ICV).

## Analyses bivariées

La concordance locale maximale pour la combinaison *GE-CD* correspond à SS C5 15 ct (0,67) alors que les autres strates présentent des valeurs beaucoup plus faibles. À l'échelle du voisinage spatial, la concordance maximale est présentée par les strates SBb C3 45 (0,38) et C6 cht (0,35) (Figure 56a).

Les strates qui présentent une concordance locale élevée pour *GE-CH* sont SS C3 60 chp et SS C3 75 (0,67) (Figure 56b) tandis que celles qui présentent une concordance élevée à l'échelle du voisinage sont SS C4 30 cp (0,34) et SBb C3 45 (0,31). Comme pour d'autres cas mentionnés précédemment, la strate SS C3 75 présente un ICL maximal et un ICV minimal. Par contre, les strates SS C4 30 cp et SBb C3 45 ont une valeur d'ICL égale à 0 (Figure 56b).

Par rapport à la combinaison de variables *GE-CA*, les strates SS C3 45 chp (0,67), SS A3 45, SS B3 45 el et SS D3 45 chp (0,50) présentent la concordance locale maximale (Figure 56c). D'autres strates comme C6 cht (0,23), SS C3 45 chp et SS C3 45 cp (0,22) ont les ICV les plus élevés. Comme pour les cas précédents, certaines strates telles que C6 cht et SS A3 45 présentent la concordance maximale pour une des deux échelles et minimale pour l'autre. SS C3 45 chp présente, quant à elle, la concordance maximale pour les deux échelles (Figure 56c).

Pour la combinaison de variables *CD-CH*, la concordance locale maximale (ICL) correspond à SS C3 75 el (0,67) et SE C3 45 el (0,50) (Figure 49d) alors que la concordance de voisinage maximale correspond à BbS A4 15 ct, SBb C3 45, SS C4 30 (0,33) ainsi qu'à SBb A4 30 (0,28) et SBb D3 75 el (0,27). Les strates BbS A4 15 ct, SS C4 30 et SBb D3 75 el présentent une valeur d'ICL égale à 0. Les strates qui présentent les ICL les plus élevés pour la combinaison *CD-CA* sont SBb C3 45 (0,40), SS C3 45 chp et SS C3 75 el (0,33) (Figure 49e). D'autre part, les strates qui présentent les ICV les plus élevés sont SS B3 45 el (0,25) et BbS A4 15 ct (0,24). Cette dernière présente également la valeur minimale d'ICL.

Finalement, par rapport à la combinaison *CH-CA*, les strates qui présentent la concordance maximale à l'échelle locale sont SBb D3 45 el, SE C3 45 el, SS A3 45 et SS B3 45 el (0,50) (Figure 49f) alors que celles qui la présentent à l'échelle du voisinage sont BbS A4 15 ct (0,31) et SS C3 60 el (0,28), la concordance locale de ces dernières étant nulle.

En résumé, pour toutes les combinaisons de deux variables, plusieurs des strates qui présentent une concordance locale élevée présentent également une faible concordance à l'échelle du voisinage et vice-versa.

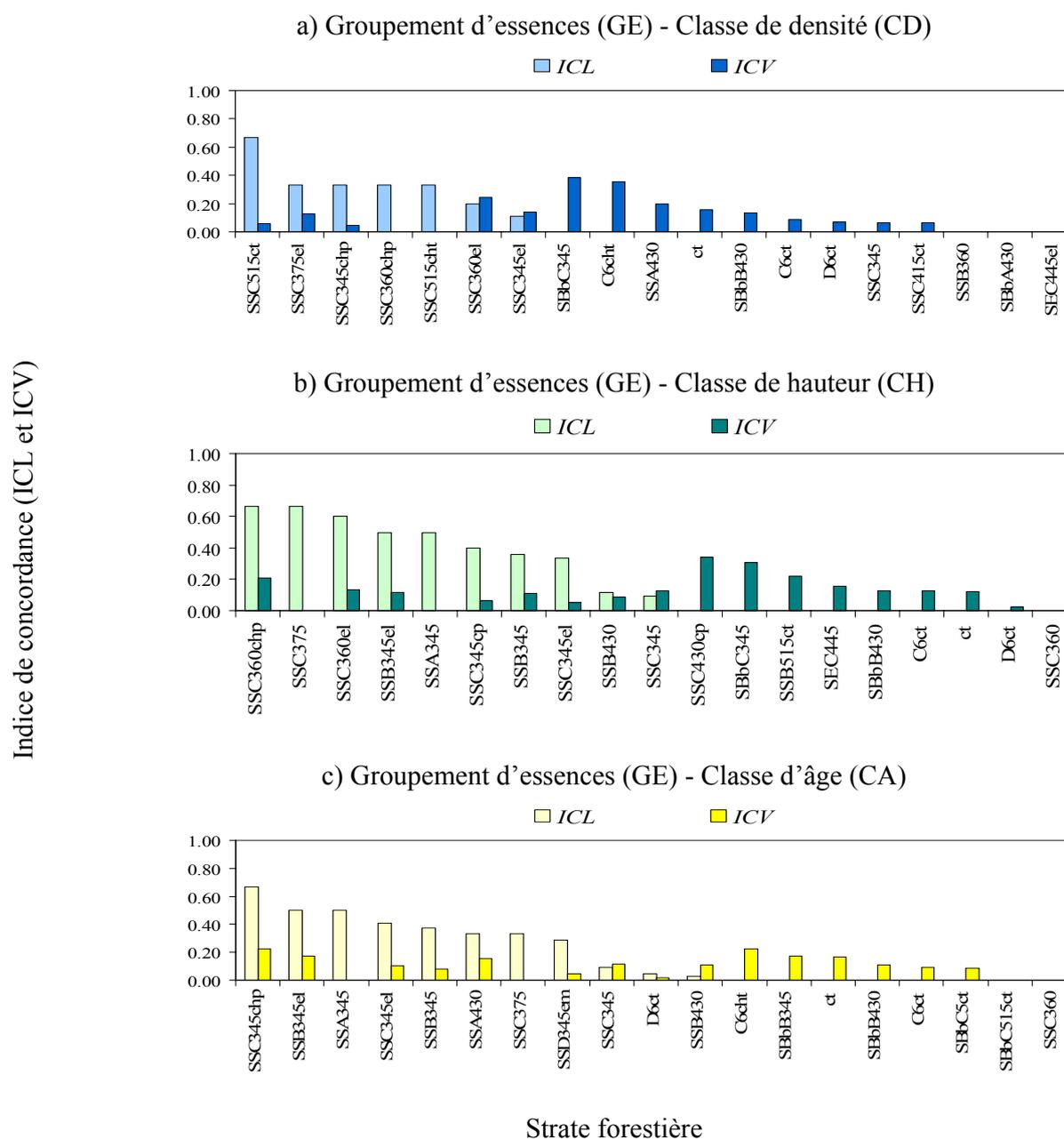


Figure 56. Indice de concordance carte/p-e des strates forestières pour les combinaisons de deux variables, aux échelles locales (ICL) et du voisinage spatial (ICV).

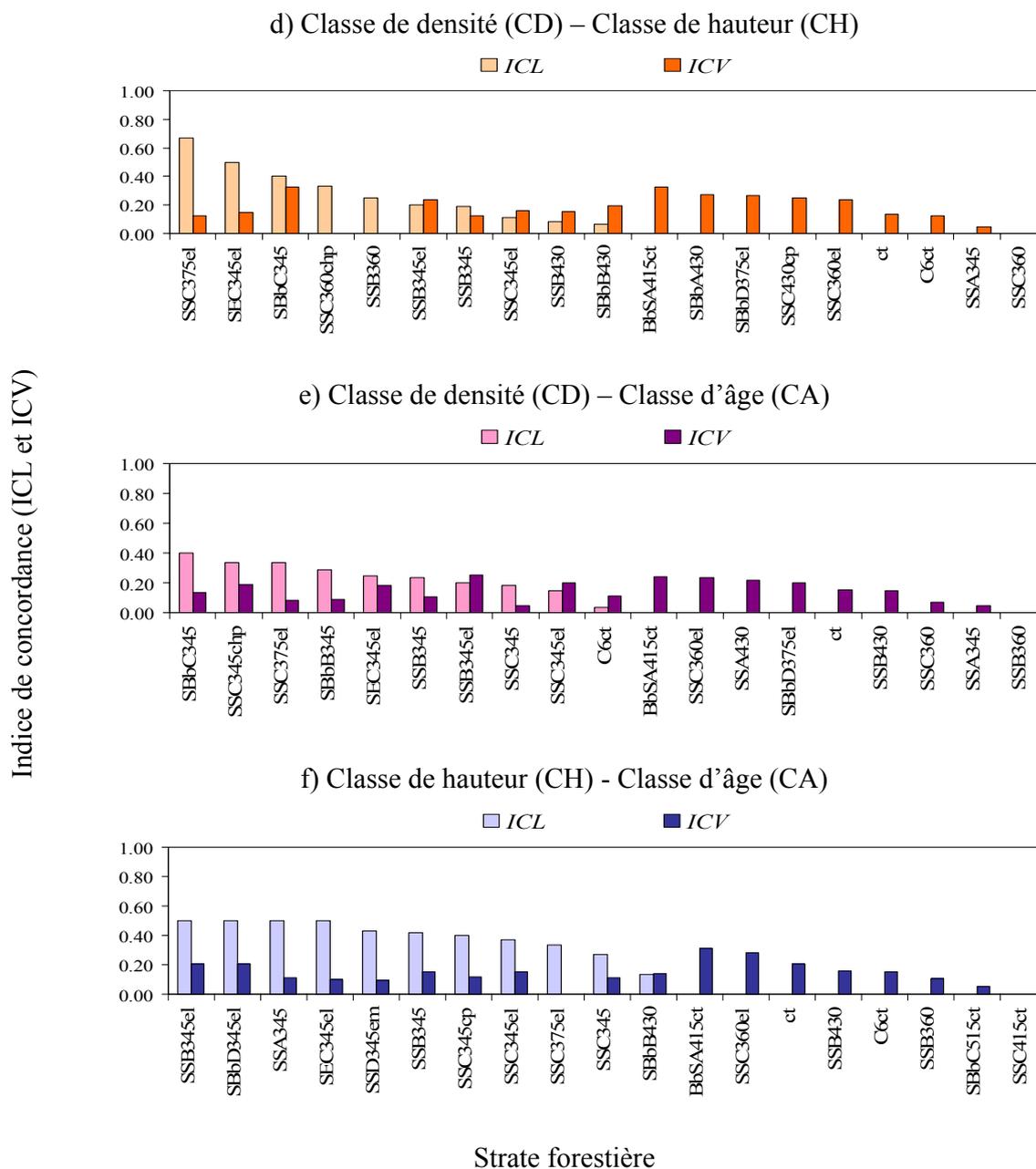


Figure 63. Suite. Indice de concordance carte/p-e des strates forestières pour les combinaisons de deux variables, aux échelles locales (ICL) et du voisinage spatial (ICV).

## Analyses multivariées

Les strates qui présentent la concordance locale la plus élevée pour la combinaison *GE-CD-CH* sont SS C3 60 chp et SS C3 75 el (0,33) et SBb B5 15 ct (0,29) (Figure 57a) alors que SBb C3 45 l'affiche à

l'échelle du voisinage spatial (0,27). Cette dernière présente aussi un ICL égal à 0. SS C3 60 chp, pour sa part, obtient une valeur maximale d'ICL et minimale d'ICV (Figure 57a).

Pour la combinaison *GE-CD-CA*, la concordance locale maximale correspond à SS C3 45 chp (0,33) et celle du voisinage spatial à C6 cht (0,23) (Figure 57b). Cette dernière présente un ICL minimal tandis que la première strate présente un faible ICV.

Les strates SS A3 45 et SS D3 45 chp présentent, pour la combinaison *GE-CH-CA*, la concordance locale maximale (0,50) tandis que C6 cht et SE C4 45 l'obtiennent pour le voisinage spatial (0,15 et 0,14 respectivement) (Figure 57c). Suivant la même tendance que d'autres combinaisons de variables précédentes, les dernières strates affichent une concordance locale nulle tandis que SS D3 45 chp concorde peu à l'échelle du voisinage spatial (Figure 57c).

Les strates SS C3 75 el et SE C3 45 el possèdent la concordance locale maximale pour la combinaison *CD-CH-CA* (0,33 et 0,25 respectivement) tandis que BbS A4 15 ct et SS C3 60 el l'obtiennent pour le voisinage spatial (0,24 et 0,23 respectivement) (Figure 57d). Ces dernières montrent également une concordance locale nulle, de même que SS C3 75 el à l'échelle du voisinage spatial.

Pour la combinaison des quatre variables, les strates SS B3 45 el et SS C3 45 cp (0,20) présentent la concordance locale maximale tandis qu'à l'échelle du voisinage, elle correspond à C6 cht (0,15), SS C3 60 el (0,13) et SE C4 45 (0,12) (Figure 58). Comme pour les autres combinaisons, certaines strates dont la concordance est élevée pour une des échelles présentent une faible concordance pour l'autre.

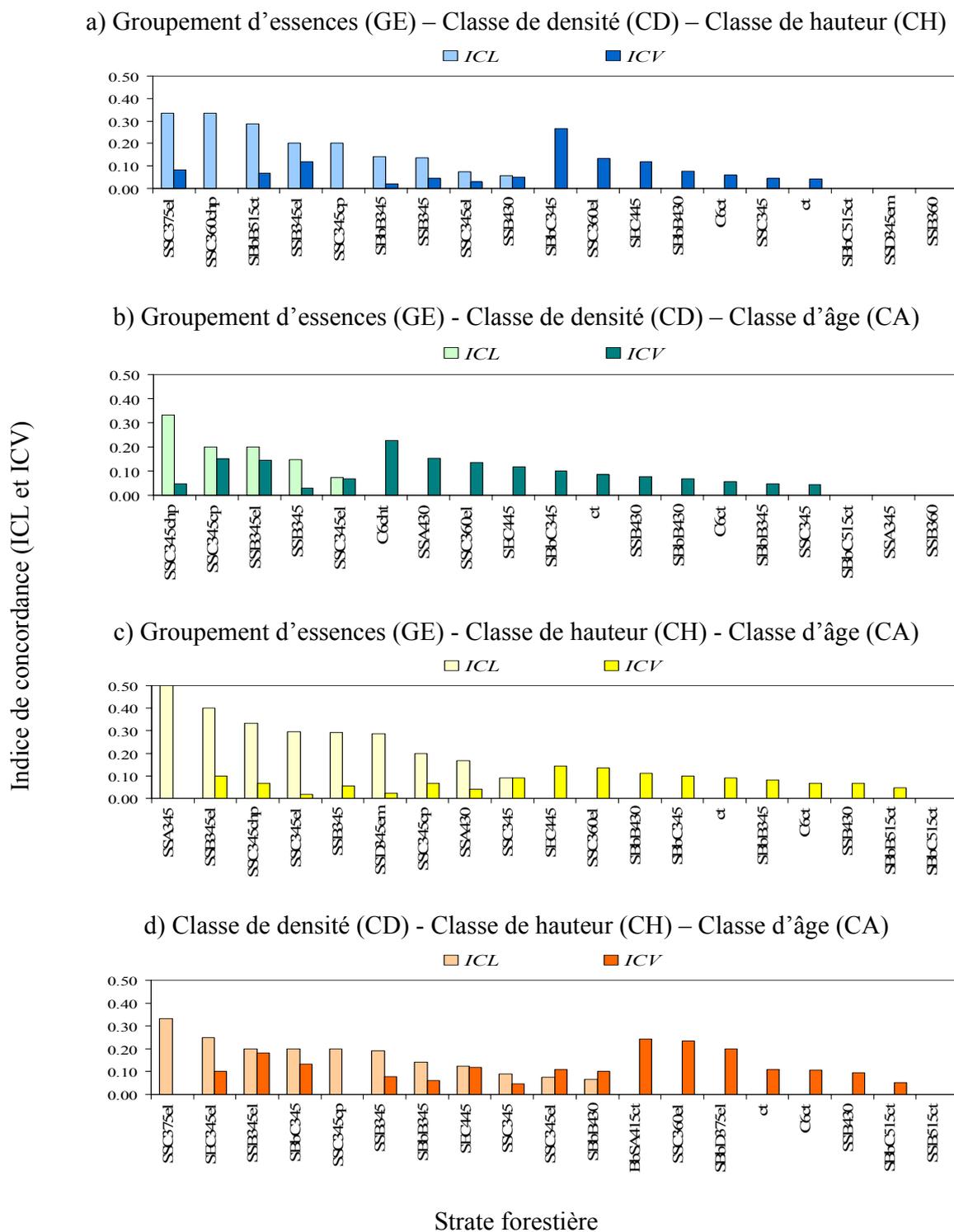


Figure 57. Indice de concordance carte/p-e des strates forestières pour les combinaisons de trois variables, aux échelles locales (ICL) et du voisinage spatial (ICV).

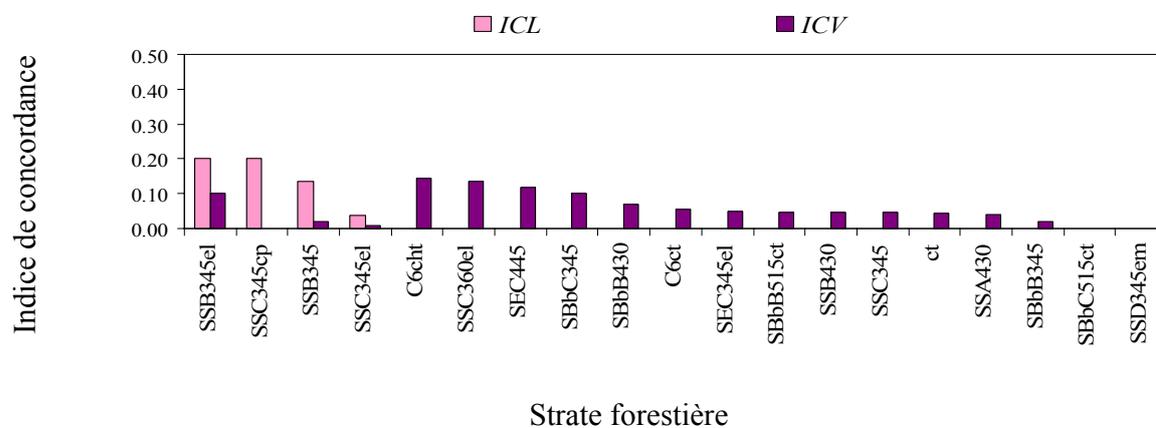


Figure 58. Indice de concordance carte/p-e des strates cartographiques pour la combinaison des quatre variables, aux échelles locales (ICL) et du voisinage spatial (ICV).